A dual-activity Topoisomerase Complex Regulates both Transcription and Translation

Shuaikun Su
National Institute on Aging, NIH
Yutong Xue
Alexei Sharov
National Institute on Aging
Yongqing Zhang
NIA/NIH
Seung kyu Lee
National Institute on Aging
Jennifer Martindale
Wen Li
Peking University
Wai Lim Ku
NHLBI, NIH https://orcid.org/0000-0002-8859-6771
Keji Zhao
National Institutes of Health https://orcid.org/0000-0001-5559-6233
Supriyo De
National Institute on Aging, NIH
Payel Sen
National Institute on Aging, NIH https://orcid.org/0000-0003-2809-0901
Myriam Gorospe
National Institutes of Health
Dongyi Xu
Peking University https://orcid.org/0000-0001-5711-2618
Weidong Wang (wangw@grc.nia.nih.gov)
National Institute on Aging https://orcid.org/0000-0002-0658-7928

Article

Keywords: Topoisomerase, transcription, translation, TOP3B, TDRD3, FMRP, autism, CHD8

Posted Date: November 23rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1054387/v1

Abstract

Topoisomerases solve topological problems for DNA, but whether they function similarly for RNA remains unclear. Here we show that the Topoisomerase 3b (TOP3B)-TDRD3 complex acts on both DNA and RNA. On DNA, TOP3B-TDRD3 binds transcription start or termination regions; and inactivation of TOP3B topoisomerase activity decreases expression of some though not all genes downregulated in TOP3B-KO and TDRD3-KO cells. On mRNA, TOP3B-TDRD3 preferentially binds coding regions of long mRNAs and stabilizes them independently of TOP3B topoisomerase activity. Moreover, ablation of the complex or inactivation of TOP3B catalytic activity alters translation for some mRNAs, suggesting that TOP3B-TDRD3 may solve RNA topological problems. Finally, several schizophrenia and autism-risk genes are bound and regulated by TOP3B-TDRD3 at either transcription or translation. Our data suggest that TOP3B-TDRD3 can regulate transcription and translation by topoisomerase activity dependent or independent mechanisms; and disruption of either function could contribute to pathogenesis of psychiatric disorders in TOP3B mutation carriers.

Introduction

Topoisomerases are “magicians” of the DNA world, with critical roles in relieving topological stress produced during DNA replication, transcription, and chromosome segregation (*1*). Mutations in or deregulation of topoisomerases can cause defective development, lethality, and human diseases, such as neurological disorders and cancer (*1*). A fundamental question that has remained unanswered is whether RNA metabolism also produces topological stress that requires resolution by a topoisomerase. Recent discoveries that a Type IA topoisomerase in animals, TOP3B, is a dual-activity topoisomerase that can change topology of both DNA and RNA (*2, 3*), have led to the hypothesis that RNA metabolism may indeed produce topological problems resolved by TOP3B (*4*). Moreover, the findings that the RNA topoisomerase activity is prevalent in Type IA topoisomerases in all domains of life (*5*) further suggest that RNA topological problems occur in a wide range of species that select for retention of RNA topoisomerase activity through millions of years of evolution (*4*).

The notion that TOP3B can act in both DNA and RNA metabolism is supported by accumulating evidence. For DNA, TOP3B has been shown to promote neuronal activity-dependent transcription in the mouse brain (*6*); and to be recruited to specific promoters by its partner, TDRD3, to reduce R-loops and maintain genome stability (*7, 8*). For RNA, TOP3B is the only topoisomerase in animals containing a conserved RNA binding domain, RGG-box, which is critical for TOP3B binding to mRNAs (*9*), and catalyzing RNA topoisomerase reactions (*3*). TOP3B works with the siRNA machinery to promote heterochromatin formation and silencing of transposons (*10*); and promotes neural development (*9*). In addition, TOP3B possesses a nuclear export sequence that allows it to shuttle between nucleus and cytoplasm (*2*). Moreover, overexpression of a TOP3B mutant can produce cleavage complexes on both DNA and RNA in cells (*11*), suggesting that TOP3B can catalyze reactions on both types of nucleic acid *in vivo*. Notably, TOP3B has been recently shown to be required for replication of positive-strand RNA viruses including SARS-CoV2, raising the possibility that TOP3B could be a feasible antiviral target (*12*).

Indirect evidence supports a role of the TOP3B-TDRD3 complex in mRNA translation (*2, 3*). First, the complex can biochemically and/or genetically interact with translation regulators, including Fragile-X mental retardation protein (FMRP) (*2, 3*), and the exon-junction complex (EJC) (*2, 13*). Second, TDRD3 itself has been suggested to regulate translation of some mRNAs (*14*). Third, a large fraction of TOP3B is localized in the cytoplasm, where it co-localizes with TDRD3 and FMRP in RNA stress granules (SGs) in response to cellular stress (*2, 3*). SGs are compartments consisting of transiently inactivated mRNAs and translation machinery, which are formed as a cytoprotective mechanism to shut down non-essential translation during stress. Fourth, TOP3B co-fractionates with TDRD3 and FMRP in polyribosomes (*2, 3, 5*), whose mRNAs under active translation. Finally, *TOP3B* mutations have been linked to multiple psychiatric and cognitive disorders, including schizophrenia and autism (*2, 15-17*). This feature also resembles that of FMRP, with the Fragile X syndrome identified as the leading cause of autism (*18*). This has led to the proposal that TOP3B works with FMRP to regulate translation of mRNAs important for these disorders (*2, 3*). However, critical evidence has been lacking to show that ablation of TOP3B protein or inactivation of its catalytic activity can alter translation of mRNA bound by the topoisomerase.

This study aims to address three basic questions regarding the roles of TOP3B-TDRD3 complex in DNA transcription and mRNA translation. First, which genes are directly bound by the complex at DNA or mRNA levels? Second, is transcription or translation of these TOP3B-TDRD3 bound genes altered when the complex is inactivated? Third, is the topoisomerase activity of TOP3B required for expression of these genes? The answer to the last question is particularly important because it determines whether topological changes on DNA and RNA catalyzed by TOP3B are important for optimal transcription and translation. To address these questions, we identified TOP3B-TDRD3 target genes at either DNA or mRNA levels using genome-wide approaches; created *TOP3B* or *TDRD3* knock-out and knock-in mutant HCT116 cell lines; and determined gene expression changes at the level of transcription, mRNA abundance, and translation. Our data show that TOP3B-TDRD3 can stimulate transcription of a subset of its bound genes by preventing early dissociation of RNA polymerase II (pol II). It can also alter translation efficiency of bound mRNAs either positively or negatively. Moreover, TOP3B catalytic activity is needed for normal transcription and translation of some - though not all - of its target genes, suggesting that TOP3B-TDRD3 can solve topological problems on both DNA and RNA for some genes; and it can also regulate gene expression in a manner independent of topoisomerase activity. Finally, several schizophrenia and autism risk genes are bound and regulated by the TOP3B-TDRD3 complex, suggesting that their dysregulation could be a mechanism by which *TOP3B* mutation contributes to pathogenesis of psychiatric disorders.

Results

Creating *TOP3B-TDRD3 KO* and *KI* cell lines to study their function

To study the function of the TOP3B-TDRD3 complex, we generated isogenic HCT116 cell lines that are individually inactivated for *TOP3B* (*TOP3B-KO*); *TDRD3* (*TDRD3-KO*); the catalytic activity of TOP3B (*Y336F-KI*); or their interacting protein FMRP (*FMR1-KO*), using the CRISPR-Cas9 technology (Fig. 1a-d & Fig. S1a-f) (*19*). We then utilized these cell lines to identify TOP3B-TDRD3 DNA and RNA targets by ChIP-

seq and eCLIP-seq analyses, respectively. We further analyzed the effect of *TOP3B-TDRD3* mutation on transcription by pol II ChIP-seq and PRO-seq; mature mRNA levels by RNA-seq; and translation by Ribo-seq (Fig. 1a) (*20-22*). The HCT116 cell line has earlier been extensively used for comparable transcription and translation studies, enabling us to compare our findings for TOP3B with those for other topoisomerases. Genomic DNA sequencing revealed the occurrence of cleavage-directed frame-shift mutations in various *KO* clones (Fig. S1b & S1e), as well as the targeted homozygous Y336F-substitutions in the *KI* clone (Fig. 1c). Western blotting confirmed the absence of each target protein in their respective *KO* cells (Fig. 1b) and the presence of the Y336F mutant protein in *TOP3B-KI* cells (Fig. 1d). The level of Y336F was about 50% of that of *WT* cells, suggesting that the catalytic reaction may help to stabilize TOP3B protein. We noticed that the levels of TOP3B and TDRD3 were lower in *KO* cells of each other (~70%), but not in *FMR1-KO* cells (Fig. 1b), which is consistent with earlier observations that TOP3B-TDRD3 depends on each other for stability (*7*).

We generated an additional *TOP3B-KO* clone (*KO2*) using a different guide RNA and found that it expresses a mutant protein carrying a partial deletion of the Toprim domain (Fig. S1c-d). Because this highly conserved domain is critical for topoisomerase activity (*23*), the mutant protein is expected to show loss of function and was used as an internal control in our translation studies (see below).

TOP3B-TDRD3 complex regulates transcription of a subset of genes.

We first studied whether genome-wide transcription is altered in cells with a *KO* or *KI* mutant TOP3B-TDRD3 complex by pol II ChIP-seq and PRO-seq analyses. The former method detects both transcribing and stalled pol II binding genome-wide, whereas the latter measures the levels of nascent transcripts produced by an active polymerase, and thus can more specifically reflect transcription (*21*). Scatter plot analysis revealed that the PRO-seq and pol II signals at transcription start sites (TSS) of all genes in *TOP3B-KO*, *TDRD3-KO*, and *TOP3B-Y336F-KI* cells strongly correlate with those of *WT* cells, with a correlation coefficient > 0.9 (Fig. 1e; Fig. S2a; Fig. S2c), indicating that inactivation of TOP3B-TDRD3 complex leaves global transcription largely unperturbed in basal conditions.

In fact, calculation of PRO-seq signals at TSS and exons revealed that the majority of genes remained unchanged (>66%, 1.2-fold threshold), whereas a smaller percentage of genes showed consistently decreased or increased in two independent biological replicates (Fig. 1f; Fig. S2b; Table S1). A Venn graph analyses identified about 244 genes consistently decreased in both *TOP3B-KO* and *TDRD3-KO* cells; 224 genes that are decreased in both *TOP3B-KO* and *Y336F-KI* cells; 132 genes that are decreased in all three cell lines (Fig. 1g & Table S1). We infer that these 132 genes are likely regulated by the TOP3B-TDRD3 complex in a topoisomerase activity dependent manner, either directly or indirectly.

We noticed that the number of upregulated genes in *TOP3B-KO* and *TDRD3-KO* cell lines is about 3 times greater than that of the down-regulated genes (Fig. 1f). This was somewhat unexpected, because

topoisomerase is known to enhance but not repress transcription. One explanation could be that the upregulated (and some downregulated) genes are indirectly regulated by TOP3B-TDRD3.

A gene ontology (GO) analysis of the downregulated genes showed enrichment of functional terms related to regulation of transcription (Fig. 1h), consistent with the notion that the upregulated genes might result from the alleviation of secondary effects of TOP3B-TDRD3 inactivation. The altered transcription of several genes in *TOP3B-TDRD3 KO* and *KI* cells by PRO-seq was verified by bedGraph analysis and subsequently by RNA-seq analysis of polyA-selected RNA (Fig. 1i; Fig. S2d-e). The altered PRO-seq signals at TSS suggest that the initiation of transcription is regulated by the TOP3B-TDRD3 complex in a topoisomerase dependent manner. One of them is *FGF2*, which encodes a fibroblast growth factor that is reduced in the postmortem brains of depressed patients (*24*). Another example is *APOBEC3H*, which encodes a protein functioning importantly in innate anti-viral immunity (*25*). Dysregulation of these genes may contribute to the autism observed in individuals carrying a *TOP3B* mutation, and to the autoimmune phenotype observed in *TOP3B-KO* mice (*26*).

TOP3B-TDRD3 preferentially binds TES regions of highly active genes

We identified TOP3B-TDRD3-bound genes in HCT116 cells by ChIP-seq, using two different anti-TDRD3 antibodies. Mock ChIP-seq in *TDRD3-KO* cells was included as a negative control, which is critical because TOP3B-TDRD3 preferentially binds highly transcribed genes (*6*) that are “hotspots” in ChIP-seq (*27*). We calculated the difference in TDRD3 ChIP signals at TSS, exons, introns and transcription end sites (TES) between *WT* and *TDRD3-KO* cells from three replicates and identified a total of 55 genes that have >1.2-fold difference at any of the four regions (Table S2). We considered these as TOP3B-TDRD3-bound genes. Interestingly, the number of genes with TOP3B-TDRD3 binding at TES (23) is greater than the number with binding at TSS (20) or other regions (15). BedGraph visualization confirmed the presence of strong TDRD3 peaks in several representative genes from *WT* but weaker signals from *TDRD3-KO* cells. These include *EGR1* and *AREG*, in which TDRD3 binds at TES; and *ACSL5*, in which TDRD3 binds at TSS (Fig. 2a). The TDRD3 peaks at either TES or TSS always colocalized with the position of pol II, consistent with our previous findings that TOP3B-TDRD3 associates with pol II during the entire transcription process in mouse brains (*6*). As a negative control, no obvious difference between *WT* and *TDRD3-KO* cells was detected for a TOP3B-TDRD3 unbound gene, *GAPDH* (Fig. 2a).

We found that the average PRO-seq and Pol II signals of TOP3B-TDRD3 bound genes were significantly higher compared to the average signals over all genes (Fig. 2b), indicating that the complex preferentially binds highly transcribed genes, in accord with our previous findings in mouse brains under neuronal activation (*6*).

TOP3B topoisomerase activity facilitates transcription of some of its target genes

Analysis of the TOP3B-TDRD3 bound genes revealed that their PRO-seq and pol II signals at TSS, exons and TES in *TOP3B-KO* cells strongly correlate with those of the corresponding regions in *WT* cells (Fig. 2c). The signals for only a subset of genes showed consistent reduction. The results resemble our previous mouse data that many TOP3B-TDRD3-bound genes exhibit no obvious decrease in basal transcription.

Among the altered genes, *EGR1*, *AREG* and *ACSL5* showed reduced nascent transcript levels at TSS, exons and TES, ranging from -20% to -80%, in both *TOP3B-KO* and *TDRD3-KO* cells, by bedGraph visualization (Fig. 2a) and quantification from two independent replicates of PRO-seq (Fig. 2d), suggesting that the TOP3B-TDRD3 complex binds and stimulates transcription of these genes. The reduced transcription of these genes in *TOP3B-KO* cells was confirmed by pol II ChIP-seq (Fig. 2a, e), and RT-qPCR (Fig. 2f) analyses. As a control, no obvious difference was observed for *GAPDH* (Fig. 2a, d, e). The strongest reduction was observed in the TES of *AREG* (80%), greater than that at TSS or exons of the same gene, suggesting that the complex enhances transcription elongation of this gene.

The reduction of PRO-seq signal in *TOP3B-Y336F* mutant cells was comparable to those in *TOP3B-KO* cells at TSS and exons, but not at TES for the above genes (Fig. 2d). Analysis of *EGR1* in this cell line showed that its pol II levels displayed significant reduction at exons ($p<0.05$), and a strong trend of reduction at TES ($p=0.07$) (Fig. 2e); and *EGR1* mRNA levels also exhibited a strong trend for reduction ($p=0.08$) (Fig. 2f), suggesting that TOP3B catalytic activity contributes to transcription of this gene. In contrast, for *AREG*, its pol II and RNA-seq signals exhibit no obvious decrease (or perhaps a small increase), indicating that TOP3B catalytic activity is dispensable for transcription of this gene (Fig. 2e-f). For *ACSL5*, its pol II and RT-qPCR levels showed reduction, but not statistically significant. Taken together, these data suggest that the TOP3B-TDRD3 complex can coordinately bind and regulate transcription of specific genes using topoisomerase activity dependent and independent mechanisms.

TOP3B inactivation disrupts translation and mRNA levels for a small number of genes

We examined mRNA translation in HCT116 cells harboring mutations of either *TOP3B* or *TDRD3* by Ribosome profiling (Ribo-seq), which monitors translation of mRNAs genome-wide by quantifying ribosome-protected RNA fragments (RPFs) (*20*). We also analyzed the levels of mRNAs genome-wide in the same experiment by RNA-seq. As a proof of validity of our assays, the levels of RPFs and RNA-seq for *TOP3B* and *TDRD3* genes were reduced in *TOP3B-KO1* and *TDRD3-KO* cells, respectively (Fig. S3a-b & Fig. S3e-f), consistent with absence of these proteins in their respective *KO* cells (Fig. 1b). In addition, no obvious difference was observed for *TOP3B* mRNA in *TOP3B-KO2* cells (Fig. S3d), which is also consistent with the finding that these cells express a truncated TOP3B protein (Fig. S1c-d).

We found that the levels of RPFs and RNA-seq from both *TOP3B-KO1* (3 replicates) and *TOP3B-KO2* (2 replicates) cells strongly correlated with those of *WT* cells (R>0.9) (Fig.3a, b & Fig. S4a), indicating that inactivation of TOP3B again did not alter global translation or mRNA levels. The levels of RPFs also

correlated well with those of RNA-seq (R=0.84), but poorly with those of PRO-seq (R=0.41, Fig. S4e), which are consistent with previous findings (*29*). The heatmaps (Fig. 5e-f) also showed strong co-clustering between RPFs and RNA-seq signals, but not with those of PRO-seq, both in the *TOP3B-KO* and also in *TOP3B-Y336F* and *TDRD3-KO* cells. These data suggest that translation is largely controlled by mRNA levels, but not with *de novo* transcription.

We identified about 1400 differentially expressed genes (DEGs, threshold >1.2-fold, $p<0.05$) in *TOP3B-KO1* and *KO2* cells by either RNA-seq or Ribo-seq (Fig. 3c, d; Table S3). A Venn diagram identified DEGs that were commonly reduced or increased in both *KO1* and *KO2* cells (Fig. S4f & Table S3). We used these common genes to represent the DEGs in *TOP3B-KO* cells in subsequent calculations to reduce the effect of clonal variation. In total, there were 884 DEGs by RNA-seq and 550 by Ribo-seq (Table S3).

Examination of the 550 DEGs identified by Ribo-seq showed that the percentage of decreased vs. increased genes were comparable (47% vs. 53%), implying that TOP3B may affect translation either positively or negatively. Because translation of an mRNA is often affected by its transcription and degradation (*30*), some of these DEGs could be regulated by TOP3B at those steps rather than at translation *per se*. To identify mRNAs that were dependent on TOP3B only for translation, we divided the decreased DEGs identified by Ribo-seq into 4 groups, based on whether these genes show concomitant decrease in RNA-seq (which reflects steady-state levels) and PRO-seq (which reflects nascent transcripts) (Fig. 3e & Table S3). Group 1 DEGs showed no concomitant decrease in RNA-seq and PRO-seq levels, which should represent genes regulated by TOP3B at translation level only. Group 2 DEGs exhibited decreased signals in RNA-seq, but no change in PRO-seq signals. This group consists of transcripts like *TOP3B* mRNA, whose decreased translation also reduces mRNA levels through no-go decay or nonsense-mediated decay mechanisms (*31-33*). The group may also contain genes of which translation reduction is due to decreased mRNA levels caused by accelerated mRNA degradation. Group 3 DEGs showed concomitant decrease in PRO-seq but no change in RNA-seq, whereas Group 4 DEGs showed decrease in both RNA-seq and PRO-seq (Fig. 3e). The last two groups likely consist of genes regulated at the transcription step.

This analysis revealed that majority of the decreased DEGs belong to group I (60%) and II (22%), suggesting that TOP3B mainly acts post-transcriptionally to regulate translation and mRNA stability. Analysis of the increased DEGs obtained similar results, as the most of the DEGs belong to post-transcriptional groups (I and II) (Fig. 3e). The number of decreased DEGs in group 1 and 2 was comparable to that of increased DEGs (212 vs. 189), supporting the notion that TOP3B can both promote and repress translation.

Several TOP3B-regulated mRNAs are important for mental disorders

Notably, the DEGs of Group 1 and 2 consist of multiple autism or schizophrenia risk genes, including *CHD8*, *FGFR1*, *SGSH* and *FAT1* (Figure 3e-f & Table S4). BedGraph visualization confirmed the decrease or

increase of the RPF signals across the gene body (Fig. 3f). Their patterns resemble that of the positive control gene (*TOP3B*) (Fig. S3a), suggesting that translation of these genes was disrupted in *TOP3B-KO* cells. The difference in RPF signals between *TOP3B-KO* and *WT* cells was significant (p<0.05) (Fig. 3g). As a control, the mRNA levels of *CHD8*, *SGSH* and *FAT1* in *TOP3B-KO* cells exhibited no significant difference comparing to those of *WT* cells (Fig. 3f-g), suggesting that translation efficiency but not mRNA levels were disrupted by TOP3B inactivation. The nascent transcript levels (PRO-seq) for these genes showed no obvious difference relative to those of *WT* cells (Fig. 3f-g), indicating that these genes are regulated by TOP3B post-transcriptionally.

We further performed polysome profiling coupled with RT-qPCR and immunoblotting to verify changes in mRNA translation in *TOP3B-KO* cells. The former method measures the translation by separating the mRNAs based on the numbers of ribosomes on them through sucrose gradient (*34*). For *CHD8* mRNA, the peak in the heavy polysome fraction in *TOP3B-KO* cells was decreased, whereas that in the light polysome fraction was increased (Fig. 3h), consistent with the Ribo-seq data that its translation was reduced. Immunoblotting confirmed a reduced CHD8 protein level and increased FAT1 protein level in *TOP3B-KO* cells (Fig. 3i-j & Fig. S5c). As controls, the level of *TOP3B* mRNA was decreased in the polysome fractions but increased in the monosomes in *TOP3B-KO* cells; and the level of *GAPDH* mRNA was largely unchanged (Fig. 3h & Fig. S3c). Moreover, the overall polysome profile was indistinguishable between *TOP3B-KO* and *WT* cells (Fig. S4g), supporting the findings from Ribo-seq that TOP3B inactivation does not affect global mRNA translation. Together, these data suggest that TOP3B inactivation disrupts translation of multiple mRNAs important for mental disorders.

TOP3B topoisomerase activity regulates RNA levels and translation

We analyzed *TOP3B-Y336F-KI* cell line to study the unresolved question of whether TOP3B depends on its topoisomerase activity for its function. The RNA-seq and Ribo-seq signals of *Y336F-KI* cells strongly correlated with those of *WT* cells (Fig. S4b; R>0.95), which resemble those of *TOP3B-KO* cells (Fig.3a-b), suggesting that inactivating either TOP3B topoisomerase activity or the entire protein does not affect global mRNA levels or translation. Our RNA-seq and Ribo-seq analyses identified about 1000 decreased and 2000 increased DEGs in *Y336F-KI* cells (Fig. 4a; Table S5). Comparison of these DEGs with those of *TOP3B-KO* cells revealed the percentages of commonly decreased and increased DEGs: 24% and 50%, respectively by RNA-seq; and 27% and 51%, respectively by Ribo-seq (Fig. S5a-b; Fig. S6a; Table S6-7). We then assessed whether the DEGs of *TOP3B-Y336F* cells showed the same direction of alteration in *TOP3B-KO* cells by chance using identical numbers of randomly selected genes as comparisons. We found that the observed percentages of the commonly decreased or increased DEGs are about 2.5 to 3.5 -fold greater than those of the randomly selected genes (Fig. 4b; Table S7), implying that the commonly altered DEGs do not likely occur by chance but are most probably co-regulated by TOP3B and its catalytic activity. We refer to these genes as the TOP3B-Catalytic-Activity-dependent Genes (abbreviated as TCAGs) (marked as “+” in Fig. 4c-d). Conversely, the percentages of DEGs that are altered in opposite

directions between *TOP3B-KO* and *Y336F* cells are lower than those of randomly selected genes (Fig. 4b), indicating that TOP3B protein and its catalytic activity usually function in the same direction. This conclusion is further supported by evidence that the percentages of DEGs altered in the same direction are 2 to 10-fold higher than those in the oppositive direction (Fig. 4b).

Our heatmap analysis showed that a large fraction (about 50-75%) of the decreased or increased DEGs by RNA-seq or Ribo-seq in *TOP3B-KO* cells did not overlap with those in *Y336F-KI* cells (Fig. 4c-d), again suggesting that TOP3B may also regulate RNA levels and translation using mechanisms independent of its topoisomerase activity. Similarly, a large fraction (about 90%) of DEGs in *Y336F-KI* cells do not overlap with those of *TOP3B-KO* cells, hinting that the *TOP3B-Y336F* mutant protein may alter mRNA levels and translation in ways that are different from that caused by loss of the protein.

We also examined the representative mental disease genes that are altered in *TOP3B-KO* cells and found that they exhibit a similar pattern in *TOP3B-Y336F* cells: decreased or increased by Ribo-seq and RNA-seq analyses, and unchanged in PRO-seq analyses (Fig. 4e-f), indicating that normal translation or RNA level control of these genes depend on TOP3B topoisomerase activity. This conclusion is further supported by immunoblotting, showing similar alteration of CHD8 and FAT1 proteins between *TOP3B -KO* and *KI* cells (Fig. 3i-j & Fig. S5c).

We noticed that the levels of *CHD8* and *FGFR1* mRNAs by RNA-seq and Ribo-seq analyses were concomitantly reduced in *Y336F-KI* cells (Fig. 4e-f). The reduction of *CHD8* mRNA level was further confirmed by RT-qPCR analysis ($p<0.05$) (Fig. S7a-b), suggesting that this mRNA may be bound and stabilized by TOP3B. We studied whether this reduction of mRNA levels is due to altered translation by pre-treating cells with a translational elongation inhibitor, cycloheximide (CHX), which is expected to increase the levels of those mRNAs subject to translation-associated mRNA decay (*35*). The *CHD8* mRNA in *TOP3B-KO* and *Y336F-KI*, but not *WT* cells, was significantly increased by CHX treatment (Fig. S7c-e), indicating that the reduced level of mRNA is due to reduced translation caused by inactivation of TOP3B catalytic activity. As a positive control, *TOP3B* mRNA level was also increased by CHX treatment in *TOP3B-KO1* but not *WT* cells (Fig. S7c). This is expected because the *TOP3B* frameshift mutations in *KO1* cells create premature stop codons in the earlier exons, making this mRNA a substrate of NMD that can be stabilized by inhibiting translation elongation (*35*). In summary, these data suggest a hierarchical action: TOP3B topoisomerase activity regulates translation of an mRNA, and this process stabilizes the mRNA.

TOP3B co-regulates mRNA levels and translation with TDRD3 and FMRP

We studied whether TOP3B and its two partners, TDRD3 and FMRP, co-regulate a common set of genes. We analyzed *TDRD3-KO* and *FMR1-KO* cells using the same methods described above. The levels of RNA-seq and Ribo-seq of the two *KO* cells strongly correlated with those of *WT* cells (R>0.95; Fig. S4c-d). These features resemble those of *TOP3B-KO* cells, suggesting that inactivation of TOP3B-TDRD3 and

FMRP affects neither global mRNA levels nor translation. Volcano plots identified approximately 2500 DEGs in *TDRD3-KO* cells by each assay; and about 800~1300 in *FMR1-KO* cells (Figure 5a-d). The number of DEGs in *TDRD3-KO* cells is about twice those of *TOP3B-KO* and *FMR1-KO* cells (Fig. 3c-d), consistent with the earlier data that TDRD3 can function independently of the TOP3B-TDRD3 complex (*36*).

A Venn diagram showed that the number of commonly decreased or increased genes between *TOP3B-KO* and *TDRD3-KO* is larger than that between *TOP3B-KO* and *FMRP-KO* cells (Fig. S6b-c). Heatmaps revealed that there are more commonly downregulated (blue) or upregulated (*37*) DEGs between *TOP3B-KO* and *TDRD3-KO* than those between *TOP3B-KO* and *FMRP-KO* in both RNA-seq and Ribo-seq assays (Fig. 5e-f; compare column 4 and 7). For example, the percentage of commonly decreased DEGs between *TOP3B-KO* and TDRD3-KO cells is about 2-fold higher than those between *TOP3B-KO* and *FMR1-KO* in RNA-seq (37% vs. 18%) and Ribo-seq (56% vs. 27%) (Fig. 5e-f). The data suggest that TOP3B coregulates more genes with TDRD3 than with FMRP. The findings correlate with the interaction data that TOP3B forms a stoichiometric complex with TDRD3, and only a minor fraction of this complex interacts with FMRP (*2, 3, 38*).

We then assessed whether the observed percentages of DEGs of *TOP3B-KO* cells showing the same alteration in *TDRD3-KO* or *FMR1-KO* cells could happen by chance, using randomly selected genes as control. The results showed that the observed percentages are 3 to 5-fold or 2 to 3-fold higher respectively, than those of control genes in *TDRD3-KO* or *FMR1-KO* cells, indicating that these concomitantly altered DEGs are less likely to occur by chance but are more probably genes co-regulated by TOP3B and its partners (Fig. 5g-h). In contrast, the percentages of DEGs showing the opposite alterations between *TOP3B-KO* and *TDRD3-KO* were lower than those of randomly selected genes, indicating that TOP3B and TDRD3 largely acts in the same direction in RNA level control and translation (Fig. 5g-h). This conclusion is further supported by that the percentages of DEGs showing the same direction of alteration are about 6-10-fold greater than those of DEGs showing opposite direction (Fig. 5g-h). Similarly, the percentages of DEGs showing the oppositive changes between *TOP3B-KO and FMR1-KO* are roughly equal to those of randomly selection genes, hinting that these changes could occur by chance, and these genes may not be co-regulated by TOP3B and FMRP (Fig. 5g-h). The percentages of DEGs altered in the same direction between *TOP3B-KO* and *TDRD3-KO* cells were about 3-fold higher than those between *TOP3B-KO* and *FMR1-KO*, suggesting that TOP3B and TDRD3 coregulate more genes than TOP3B and FMRP.

We also analyzed DEGs of *FMR1-KO* cells and found a stronger overlap with those of *TDRD3-KO* than *TOP3B-KO* or *Y336F* cells, in both RNA-seq (downregulated: 39% vs. 32% or 35%, respectively) and Ribo-seq (downregulated: 38% vs. 20% or 21%, respectively) (Fig. S8a-b). The percentages of concomitantly decreased or increased DEGs between *FMR1-KO* and *TDRD3-KO* or *TOP3B-KO* were about 2 to 5-fold higher than those of randomly selected genes, whereas those of DEGs that are altered in opposite directions were either similar or lower than those of randomly selected genes (Fig. S8c). The data support the notion that TOP3B-TDRD3 and FMRP co-regulate mRNA levels and translation.

We examined the four mental disease genes with altered expression in *TOP3B-KO* and *Y336F* mutant cells and found that they all display the same direction of alteration in *TDRD3-KO* cells, whereas only two of the four exhibit the same trend in *FMR1-KO* (*FGFR1* and *FAT1*) (Fig. 5i-j). Together, these data suggest that TOP3B co-regulates more mRNAs in conjunction with TDRD3 than with FMRP at post-transcriptional steps.

TOP3B-TDRD3 regulates mRNAs in both topoisomerase-dependent and independent mechanisms

We hypothesized that the co-regulated mRNAs in both *TOP3B-KO* and *TDRD3-KO* cells (Fig. 6a-b & Table S8) are controlled by the entire TOP3B-TDRD3 complex and investigated how these mRNAs are regulated by the topoisomerase activity and FMRP. Comparing to *WT* cells, the percentage of decreased mRNAs is roughly equal to the increased mRNAs in RNA-seq (49% vs. 51%) and about 50% fewer (40% vs. 60%) in Ribo-seq (Fig. 6a-b), suggesting that the TOP3B-TDRD3 complex can affect gene expression either positively or negatively. There is strong overlap between RNA-seq and Ribo-seq data in both *TOP3B-KO* and *TDRD3-KO* cells: 64%-88% for the decreased DEGs, and 77%-93% for increased DEGs, suggesting a coordinated regulation by TOP3B-TDRD3 (Fig. 6a-b).

To determine whether the complex acts transcriptionally or post-transcriptionally, we analyzed PRO-seq levels for these DEGs. For the decreased DEGs from both RNA-seq and Ribo-seq assays, less than 38% exhibit concomitant decrease in PRO-seq levels (Fig. 6a-b). This contrasts with the increased DEGs, in which more than 51% exhibited concomitant increase in PRO-seq levels. The data resembles findings above (Fig. 5e-f) and suggests that the TOP3B-TDRD3 complex suppresses gene expression at the transcriptional level but enhances gene expression mainly post-transcriptionally.

We analyzed the decreased DEGs of the 2 *KO* cell lines and found that more than 40% of them were concomitantly decreased in *Y336F-KI* cells, whereas about 30% were decreased in *FMR1-KO* cells (Fig. 6a-b). The data thus suggest that some but not all mRNAs regulated by the TOP3B-TDRD3 complex are under control of the topoisomerase activity of TOP3B and FMRP.

The mRNAs altered in both *TOP3B-KO* and *TDRD3-KO* cells include 44 mental disorder-related genes (Fig. 6c). Half of them (22) were also reduced in *Y336F* cells. In addition to the four representative transcripts (*CHD8*, *FGFR1*, *SGSH* and *FAT1*) that were altered in *Y336F* cells (Fig. 4e), three more transcripts (*PRR12*, *SMC3* and *BCORL1*) were unchanged in *Y336F* cells (Fig. 6d-e). These findings further reinforce the notion that the TOP3B-TDRD3 complex can regulate the expression of genes important for mental disorders in topoisomerase activity dependent and independent manners.

TOP3B preferentially binds long mRNAs enriched in mental disorder genes

We performed eCLIP-seq to identify TOP3B-bound mRNAs and determine whether they overlap with those regulated by TOP3B-TDRD3 described above (*22*). We identified 1106 TOP3B-bound mRNAs based on two independent experiments in HCT116 cells (Table S9). About 1/3 of these mRNAs matched those identified from HeLa cells by TOP3B HITS-CLIP(*3*) (Fig. S9a). Those with higher eCLIP-seq signals tend to have higher HITS-CLIP signals (Fig. 7a), suggesting that TOP3B may recognize specific features of these mRNAs. We found several features that are common in TOP3B-bound mRNAs identified in the current and previous studies, which are also shared by FMRP-bound mRNAs. First, the largest fraction of TOP3B CLIP-reads (~50%) are localized in coding regions of mRNAs (Fig. 7b) (*3*), consistent with findings that TOP3B-TDRD3 associates with polyribosomes and regulates translation. Second, the average lengths of these mRNAs are significantly longer than that of randomly selected unbound mRNAs (Fig. 7c) (*3, 39*). This also resembles findings that TOP1 and TOP2-regulated genes tend to have longer average lengths and are enriched in genes important for autism and neurological disorders (*40*). Third, a fraction of TOP3B-bound mRNAs (17%) from both current and previous studies overlapped with those bound by FMRP (Fig. S9b) (*41*), including many autism (https://gene.sfari.org/), and schizophrenia-related (http://bioinfo.mc.vanderbilt.edu/SZGR/) mRNAs (109 and 59, respectively) (Fig. 7d-e; Table S9), supporting the proposal that TOP3B may work with FMRP to regulate translation of mRNAs important for mental disorders (*3*). To verify that TOP3B can indeed bind these mRNAs, we conducted RNA immunoprecipitation (RIP) analysis coupled with RT-qPCR and found that TOP3B antibody immunoprecipitated *CHD8* and *FAT1*mRNAs, but not *GAPDH* or *ACTB* mRNAs (Fig. S9g). The data indicate that TOP3B can directly bind and regulate mRNAs important for autism or schizophrenia.

TOP3B binding but not catalytic activity stabilizes target mRNAs

We next examined the effects of TOP3B inactivation on its bound mRNAs by RNA-seq or Ribo-seq analyses. In both cases, the levels of RNA and RPFs in *TOP3B-KO* cells strongly correlate with those of *WT* cells (Fig. S9c); and a minor fraction of TOP3B-bound mRNAs (12% or less; 135 out of 1106) displayed altered RNA-seq levels in *TOP3B-KO* cells treated with CHX to inhibit translation (Fig. 7f). Interestingly, among those with altered mRNA levels, 96% (130) were decreased, whereas only 4% (5) were increased (Fig. 7f & 7h). Only 8% of the decreased DEGs showed this trend by PRO-seq analysis (Fig. 7h), indicating that this decrease in RNA-seq signals is not due to reduced transcription. The cells without CHX treatment showed only a mild trend (about 60% decreased) (Fig. S9e). The data suggest that TOP3B binding preferentially stabilizes its target mRNAs when translation elongation is inhibited. In support of this notion, when cells were treated with CHX, TOP3B-bound mRNAs exhibited overall reduced RNA-seq signals in *TOP3B-KO* cells than *WT* cells (Fig. S9d).

We then analyzed these TOP3B-bound mRNAs in *Y336F* cells but did not observe a similar overall decrease (Fig. S9d & 9f), suggesting that stabilization of these mRNAs is independent of the topoisomerase activity. As examples, two representative mRNAs, *AGRN* and *KIAA0100*, were bound by TOP3B; and their RNA-seq, but not PRO-seq, signals were significantly reduced in both *TOP3B-KO* and

TDRD3-KO, but not in *Y336F-KI* cells (Fig. 7j-k), indicating that they are stabilized by TOP3B-TDRD3 binding but not topoisomerase activity.

Our analysis of TOP3B-bound mRNAs in Ribo-seq found that the numbers of mRNAs showing decreased or increased signals are about equal (37 vs. 34) (Fig. 7g and 7i), suggesting that TOP3B binding may either enhance or suppress translation of its target mRNAs. In addition, we found that 30 of the TOP3B-bound mRNAs display altered Ribo-seq signals in *Y336F-KI* cells. Among them, the number of mRNAs showing decreased abundance (11) is about 1.7-fold lower than those with increased levels (19) (Table S10), suggesting that TOP3B topoisomerase activity may either enhance or suppress translation of its bound mRNAs.

Discussion

TOP3B functions in topoisomerase activity-dependent and independent mechanisms

Our findings are relevant to the challenging question for the topoisomerase field: whether mRNA metabolism produces topological stress that depends on a topoisomerase for resolution. TOP3B has already been speculated to regulate translation, but direct evidence has been lacking. Our findings show that about 1/3 of TOP3B-bound mRNAs exhibiting reduced translation in *TOP3B-KO* also have reduced translation in *Y336F-KI* cells, indicating that they are regulated by TOP3B in topoisomerase activity-dependent manner. The data support our earlier model that mRNA translation may produce topological problems that depends on a topoisomerase to solve (Fig. 8e-f). Conversely, about 2/3 of TOP3B-bound mRNAs show reduced translation in *TOP3B-KO* but not *Y336F-KI* cells, suggesting that they are regulated by TOP3B topoisomerase activity independent mechanism.

In addition to translation, we observed that TOP3B-TDRD3 complex but not its topoisomerase activity is needed for transcription of the *AREG* gene. The finding that a topoisomerase can function independently of its enzymatic activity has precedence (*42*). We propose several such mechanisms for TOP3B-TDRD3. In one, observed in transcription, TOP3B-TDRD3 may stabilize the elongating pol II complex through protein-protein interactions between TDRD3 and pol II C-terminal domain (CTD) (*43*), and through protein-RNA interactions between TOP3B and nascent transcript (Fig. 8d). In another manner, in translation, TOP3B-TDRD3 may bind mRNA through the RNA-binding domains of TOP3B, and recruit translation factors, including FMRP and EJC, through the Tudor and CTD domains of TDRD3. FMRP may repress, whereas EJC may enhance, the first round of translation (Fig. 8g-h). In still a third manner, in mRNA stabilization, TOP3B-TDRD3 as an RNA binding protein complex may simply protect mRNAs from nuclease attack or may otherwise assemble the mRNAs into a higher order complex that is more stable (Fig. 8h). Detailed analysis uncovered several new features TOP3B-TDRD3 on DNA and mRNA (below).

TOP3B-TDRD3 mainly binds to TES and promotes both initiation and elongation of transcription

On DNA, TOP3B-TDRD3 only binds about 55 genes in HCT116 cells, which are far fewer than its bound genes in mouse brains under neuronal activation (about 5000) and TOP1-regulated genes in HCT116 (>15000) (*44*). The data suggest that although TOP3B-TDRD3 and TOP1 can both interact with pol II (*43, 44*), the former may play a more important role in brains and a minor role in other tissues; whereas the latter may play a more important role in colon cancer-derived HCT116 cells. In addition, TOP3B-TDRD3 binds many genes at TES (23), which differs from a previous study that it mainly binds TSS (*45*). We have re-analyzed our previous mouse brain data (*6*), and observed Top3b-TDRD3 binding at TES for many neuronal early response genes, with or without neuronal activation (Fig. S10). Whereas all forms of pol II bind primarily TSS, the elongating form binds mainly the TES (*46*). Our findings thus suggest that in HCT116 cells, TOP3B-TDRD3 may preferentially associate with the elongating pol II to promote elongation. This suggestion is supported by findings that both *TOP3B-KO* and *TDRD3-KO* cells show stronger reduction of pol II and PRO-seq signals at TES and exons than TSS for *AREG* and other genes (Fig. 2a & 2c-e). Moreover, our findings that pol II levels are reduced in gene bodies in *TOP3B* mutant HCT116 cells and mouse brains (*6*) differ from previous data that pol II levels are increased when TOP1 is inhibited (*44*) (Fig. S11a-b). This difference could be due to different mechanisms by the two topoisomerases: TOP3B can release only negative supercoils generated by the polymerase during transcription (*47*), whereas TOP1 can release both negative and positive supercoils (*1*) (Fig. 8a). Possibly, positive supercoil accumulation caused by TOP1 inhibition may stall the polymerase, resulting in increased level of the enzyme in gene bodies (Fig. 8b). In contrast, negative supercoil accumulation caused by TOP3B inactivation may lead to dissociation of the polymerase from DNA template (Fig. 8c). Our findings that some TOP1-regulated genes (*44*) are also controlled by TOP3B (such as *EGR1*) (Fig. S11a-b) suggest that the two topoisomerases can work together to stimulate transcription of the same genes (Fig. 8a).

TOP3B-TDRD3 directly binds mRNAs and regulates their translation and stability

On mRNAs, the TOP3B-TDRD3 binding sites are enriched in coding regions and these mRNAs have longer average lengths and overlap with many mental disease genes. The data suggest that TOP3B-TDRD3 mainly binds and regulates translation of longer mRNAs enriched in genes important for autism and neurological disorders (*2, 3*). In support of this notion, about 70 of these mRNAs display altered levels of RPFs in *TOP3B-KO* cells. Surprisingly, the percentage of mRNAs displaying increased and decreased levels of RPFs are roughly equal, suggesting that TOP3B-TDRD3 may either enhance or suppress translation. It is somewhat rare that a topoisomerase may suppress a process on nucleic acid, but TOP1 can suppress transcription initiation (*42*). Possible mechanisms for TOP3B to suppress translation may include catenation of different mRNAs and introduction of knot-like structures (Fig. 8f). It should be mentioned that the main effect of TOP3B binding is to stabilize mRNAs during translation (Fig. 7f). This feature resembles that of FMRP (*30, 48*), suggesting that TOP3B can act as an RNA binding protein in mRNA stabilization.

Our findings that multiple genes regulated by the TOP3B-TDRD3 are related to autism or schizophrenia support the notion that *TOP3B* mutations contribute to the pathogenesis of psychiatric disorders (*2, 15, 16*). In addition, the data that TOP3B topoisomerase activity is needed for translation imply that TOP3B may form covalent cleavage complexes on RNA in cells, which is supported by data that a self-cleavage mutant of TOP3B can form such complexes in cells (*11*). We propose that TOP3B-RNA cleavage complexes could be a target of anti-viral drugs. Because the resulting RNA damages consist of protein-RNA crosslink and a strand break (Fig. S12a-b), we expect that these damages should block viral RNA replication in a mechanism analogous to that employed by TOP1-DNA cleavage complex to block DNA replication. TOP3B has been proposed to be an anti-viral target, based on findings that its loss of function reduces viral RNA replication (*12*). Our proposal does not rely on the requirement of TOP3B for RNA viral replication, and postulates that as long as TOP3B can form cleavage complexes on viral RNAs, it should be a viable anti-viral target.

Materials And Methods

Cell lines

HCT116 cells were cultured in Dulbeccos' Modified Eagle Medium (DMEM, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (HyClone) and 1% antibiotics (Penicillin-Streptomycin, Sigma). For generation of the *TOP3B-KO* HCT116 cell lines, pX330 plasmid containing the guide RNA sequence was transfected in HCT116 cells. Eight to ten days later, single clones were selected for screening of absence of TOP3B protein and mutations at *TOP3B* genomic locus by Western blotting and genomic sequencing analysis, respectively. CRISPR-Cas9 mediated ablation of the *TDRD3* and *FMR1* was achieved with PX459 plasmid as the targeting vector (Addgene) (*19*). Guide RNA sequences were designed using the http://crispr.mit.edu website, and cloned into the targeting vector. Plasmids containing the guide RNA sequence were transfected into cells using FuGENE HD (E2311; Promega). Twenty-four hours post-transfection, the cells were added with fresh medium containing 1 μg/mL puromycin (P8833-10MG, Sigma). After 48 ~72 hours, cells were diluted with new medium (without puromycin) into 96-well plate. After one week, cells from single clones were transferred to 24-well plate. Then, Western blotting and Sanger sequencing were used to screen for clones that lack the target protein and harbor mutations in the target genes. The HCT116 cell line expressing TOP3B catalytic point mutant protein (*TOP3B-Y336F*) was generated by a modified CRISPR-Cas9 gene editing strategy (*19*). A single-strand DNA template (100 nt) designed to produce Y336F mutant protein was synthesized by Integrated DNA Technologies company. Plasmid containing the guide RNA sequence and the DNA template were transfected into HCT116 cells using FuGENE HD. Clone selection and validation were performed as described above. The guide RNA and DNA template sequences are described in Table S11 and Figure S1.

Ribosome profiling and RNA sequencing

Ribosome profiling protocol was based on a published method (*20*) with small modifications. Briefly, about 5~10 × 10^6 HCT116 cells were cultured in 10 cm dish for 12~16 hours. Cycloheximide (2112S, New England Biolabs) was added to a final concentration of 100 µg/mL for 10 min. Cells were washed with 2 mL cold phosphate buffered saline (PBS) with CHX, and then were collected on ice. Cells were lysed in 400 µL polysome lysis buffer by trituration through a 25-gauge needle. After centrifugation, 100 µL cell of lysates were used to extract RNA with TRIzol (15596026, Invitrogen). Then, 5~10 µg RNA was used to purify mRNAs using oligo dT beads (61012, Invitrogen). The purified mRNAs were used for RNA-seq library preparation by following a published protocol (*6*). Three hundred microliters of cell lysates were used for ribosome profiling. One thousand units of RNase T1 (EN0541, Thermo Fisher Scientific) were added to the lysates, and incubated at 25 °C for 1 hour at 500 rpm rotation. Two thousand units of RNase inhibitor (AM2696, Thermo Fisher Scientific) were added to the lysates to stop the digestion. Ribosomes were collected by sedimentation through 0.9 mL 1 M sucrose cushion. RNAs were extracted from the purified ribosomes with TRIzol. Then, 24-34 nt RNA fragments were purified using a 15% TBE-Urea gel (Thermo Fisher Scientific). Ribosomal RNAs were removed from the purified RNA fragments with NEBNext rRNA Depletion Kit (E6310S, New England Biolabs). The ribosome-protected fragments (RPFs) were treated with T4 Polynucleotide Kinase (M0201S, New England Biolabs) at 37 °C for 30 minutes with 20 units RNase inhibitor. After precipitation with isopropanol, the RPFs were used to prepare library with NEBNext® Multiplex Small RNA Library Prep Set for Illumina (E7300S, New England Biolabs). The RNA-seq and Ribo-seq libraries were sequenced by using the Hi-seq 2000 system (Illumina).

Because RNase I digestion has been commonly used in Ribo-seq, we also attempted to use RNase I to digest polysomes. However, ribosomes from HCT116 cells were very sensitive to RNase I digestion, and we were unable to obtain high quality data of RPFs (data not shown). As an alternative, RNase T1 was used to digest polysomes to monosomes without loss of ribosomes (data not shown) (*49*).

eCLIP-seq

eCLIP-seq protocol was based on a previous publication (*50*) with some modifications. *TOP3B-KO* HCT116 cells were used as negative control instead of IgG. Five µg of TOP3B antibody (WH0008940M1-100UG, Sigma Aldrich) was used for immunoprecipitation. RNA 5' Pyrophosphohydrolase (RppH) (New England Biolabs) was used to replace Tobacco Acid Pyrophosphatase (TAP) to treat RNA. All other steps and reagents were identical as described in the published method (*50*).

ChIP-seq

ChIP-seq was performed as previously described (*51*). One 15cm plate of HCT116 cells cultured in 20 ml DMEM containing 10% FBS was crosslinked with 0.54 ml of 37% formaldehyde (final concentration is 1%) at 37°C for 15 minutes. The crosslinking reaction was terminated by removing the medium. The cells

were washed once with 10 ml of 1x PBS, and then the buffer was removed leaving 0.5 ml in the plate. Cells were harvested by scrapping and concentrated by centrifugation at 3000 rpm at 4 °C for 2 minutes. The cell pellet was resuspended in 5 volumes of RIPA buffer (150 mM NaCl, 20 mM Tris pH 7.5, 2 mM EDTA, 1% NP40, 0.5% Na-Deoxycholate acid, 0.1% SDS) containing 1mM PMSF and 1x protease inhibitor cocktail. The resuspended cells were sonicated in ice-cold water bath with the Diagenode Bioruptor Sonication System to break chromatin to about 0.3-1kb. The setting of the Bioruptor is for 2 x 4 minutes with 15 seconds ON and 15 seconds OFF cycles at High power setting. After sonication, the cells were centrifuged in a 4 °C microfuge at 14,000 rpm for 10 minutes. The supernatant containing soluble chromatin fraction is saved for ChIP.

The above chromatin fraction was precleared by incubation with 100 µL of protein A beads for 1 hour at room temperature. The ChIP mixture contains 500 µL of the precleared chromatin fraction, 2.5 µL of 20 mg/ml BSA, 30 µL of the protein A beads, and the indicated amounts of antibodies. The antibodies include: 2 µg of ANTI-RNA POLYMERASE II CTD CL8WG16 (Sigma Cat#: 05-952-I), 5 µg TDRD3 rabbit polyclonal antibody described in (*3*), 5 µg of TDRD3 (D3O2G) antibody (Cell Signaling Cat#: 5942). The mixture was incubated at 4 °C overnight with rotation. The beads were collected by centrifugation using a microfuge for 1 minute at 3000 rpm at 4 °C. They were washed twice with RIPA buffer supplemented with 0.3M NaCl, twice with RIPA buffer supplemented with 0.15M NaCl, twice with LiCl buffer (1x TE, 0.25M LiCl, 0.5% NP-40, 0.5% Na-Deoxycholate), and twice with 1 x TE buffer (Tris-HCl, pH 8.0, 10mM; EDTA, 1mM). The beads were resuspended in 100 µL of 1 x TE buffer. 2.5 µL of 10% SDS, 5 µL of 10 mg/ml Proteinase K were then added, and incubated at 65 °C overnight to release DNA from the bound proteins and beads. About 100 µL of chromatin fraction was saved as the “input” which was also treated with SDS and proteinase K. The supernatant was collected by centrifugation. The remaining beads were eluted one more time with 100 µL of 1x TE. The eluted supernatants were combined and extracted twice with 1 volume of Phenol/Chloroform to remove proteins. The DNA from aqueous fraction containing was precipitated with 1 µL of 20 mg/ml glycoblue, 20 uL of 3M NaOAc (pH 5.2) and 500 µL of 100% ethanol followed by incubation for 10 minutes on dry ice. The precipitated DNA was collected by centrifugation for 20 minutes at 4 °C in a microfuge. The DNA pellet was washed once with 75% ethanol, air dried, and resuspended in 40 µL of 1x TE. The DNA was then used for library preparation and sequencing. Library preparation was described in a previous publication (*6*).

PRO-seq

PRO-seq was performed as previously described (*21*). Briefly, HCT116 cells were seeded at a concentration that will enable them to reach ~80% confluency on a 15 cm plate in 24 hours. Cells (about 10-20 million) were then collected, permeabilized, and processed for PRO-seq.

RNA Immunoprecipitation

Cytoplasmic lysates of HCT116 cells were prepared in polysome extraction buffer (20 mM Tris-HCl at pH 7.5, 100 mM KCl, 5 mM $MgCl_2$ and 0.5% NP-40) containing protease and RNase inhibitors. The supernatants were incubated with protein A-Sepharose beads coated with TOP3B antibody (WH0008940M1-100UG, Sigma Aldrich) for 2 h at 4 °C. After three washes with ice-cold NT2 buffer (50 mM Tris-HCl at pH 7.5, 150 mM NaCl, 1 mM $MgCl_2$, 0.05% NP-40), bound RNAs were extracted from the beads using TRIzol and subjected to RT-qPCR (*52*).

Polysome profiling

Polysome profiling was performed as previously described (*3, 5*). Briefly, 5×10^6 HCT116 cells were cultured in a 10 cm dish for 12~16 hours. CHX (2112S, New England Biolabs) was added to the medium to a final concentration of 100 μg/mL for 10 min. Cells were collected on ice and lysed in polysome extraction buffer (20 mM Tris-HCl at pH 7.5, 100 mM KCl, 5 mM $MgCl_2$ and 0.5% NP-40). After centrifugation, the lysate was separated through 10% to 50% sucrose gradients, 12 fractions were collected and RNA extracted to perform RT-qPCR. The distribution of mRNAs was quantified by RT-qPCR analysis and plotted as a percentage of the specific mRNA in each fraction relative to the total amount of that mRNA in the gradient.

Nascent RNA extraction

Nascent RNA was extracted based on a previously published method (*53*). Briefly, $1\sim2 \times 10^7$ HCT116 cells were collected and washed two times with ice-cold PBS. The cell pellet was resuspended with 10 mL ice-cold Hypotonic Buffer (10 mM HEPES, pH 8.0; 10 mM KCl; 2 mM $MgCl_2$; 1 mM DTT, add before use.) and incubated on ice for 15 min. The cell suspension was centrifuged at 200 ×g for 10 min at 4 °C and the supernatant was discarded. The pellet was resuspended in 2 mL cold Hypotonic Buffer and then transferred into a precooled 2 mL Dounce tissue grinder. The cell suspension was homogenized with a tight pestle using 20 strokes. The nuclei were pelleted by centrifugation at 600 × g for 10 min at 4 °C, and the supernatant was discarded. The nuclei pellet was then washed twice with 1 mL ice-cold Nuclei Wash Buffer (10 mM HEPES, pH 8.0; 250 mM Sucrose; 1 mM DTT and 20 U/mL RNase inhibitor). 1 × NUN Buffer (with RNase-free water) was prepared by mixing 50% volume of 2 M urea solution (fresh and filtered) and 50% volume of ice-cold 2 × NUN Buffer (40 mM HEPES, pH 8.0; 15 mM MgCl2; 0.4 mM EDTA; 600 mM NaCl; 2% v/v NP-40) with DTT and RNase inhibitor. The nuclei was suspended and disrupted using 1 × NUN buffer by pipetting up and down vigorously for 10 times followed by incubation on a rotating wheel for 5 min at 4 °C. The chromatin was pelleted by centrifugation at 1000 x g for 3 min at 4 °C. The chromatin pellet was washed three times with 1 × NUN buffer at 4 °C. And 1 mL of TRIzol Reagent was added into chromatin pellet followed by vortex for 30 s. The homogenized sample was incubated for 5~10 min at 50 °C on a nutator to dissolve the chromatin pellet. Then the rest steps of RNA

extraction are the same as the standard protocol of RNA purification using TRIzol. DNA contamination was removed with DNase I (AMPD1-1KT, Sigma Aldrich). Then the RNA was purified with TRIzol. RT-qPCR was performed to detect the nascent RNAs with the primers amplifying the exon-intron junction.

RNA extraction and RT-qPCR

HCT116 cells were cultured in 6-well plates for 12~16 h. For CHX treatment groups, 100μg/mL CHX was added into the medium for 3 hours before extracting RNA. After removing the medium, 1 mL TRIzol (Invitrogen, 15596026) was added to the plate directly to lyse the cells. The cell lysate was transferred into 1.5 mL tubes. Two hundred μL chloroform was added into the lysate, vortexed, and followed by centrifugation at 18,000 × g for 10 min at 4 °C. The aqueous phase was transferred to a new tube. RNA was precipitated in 1.5-fold volume of 2-propanol together with 10% volume of 3 M sodium acetate. cDNA was synthesized from 1 μg RNA using Taqman Reverse Transcription Reagents (Applied Biosystems, N8080234). After 10-fold dilution, the cDNA was used as a template to perform qPCR with SYBR Green PCR Master Mix (Applied Biosystems, 4309155). The PCR primer sequences can be found in Table S11.

Western blot and antibodies

Whole cell lysates were prepared using RIPA buffer (10 mM Tris-HCl at pH 7.4, 150 mM NaCl, 1% NP-40, 1 mM EDTA, 0.1% SDS, 1 mM dithiothreitol) containing protease inhibitor (11697498001, Roche). The protein concentrations were measured using Bradford (#500-0205, Bio-Rad). 2 × Laemmli Sample Buffer with 5% β-Mercaptoethanol were added into the protein lysates. After boiling at 95 °C for 10 min, the protein lysates were separated on a 4–20% Mini-PROTEAN® TGX™ Gel (Bio-Rad) and transferred to nitrocellulose membrane using Trans-Blot® Turbo™ Transfer System (Bio-Rad). Incubations with primary antibodies to detect TOP3B (WH0008940M1-100UG, Sigma Aldrich), TDRD3 (5942S, Cell Signaling Technology), FMRP (MAB2160, Sigma Aldrich), CHD8 (ab114126, Abcam), FAT1 (A304-403A, Thermo Fisher), GAPDH (2118s, Cell Signaling Technology) and ACTB (ab8226, Abcam) were followed by incubations with appropriate secondary antibodies conjugated with HRP (GE Healthcare). Signals were developed using Enhanced Chemiluminescence (ECL).

Bioinformatics analysis

Ribo-seq data analysis was performed as previously published (*20*). Briefly, low-quality reads and the adapter sequence (AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC) were removed by FASTX-Toolkit from FastQ files. Reads less than 25 nt were also removed by FASTX-Toolkit. The remaining reads were aligned to an rRNA reference (hg38) using the Bowtie short-read alignment program. The rRNA

alignments were discarded and the unaligned reads were collected. The non-rRNA sequencing reads were aligned to the human genomic reference (hg38) using the TopHat splicing-aware short-read alignment program (*54*). Perfect-match alignments were extracted from TopHat output.

RNA-seq reads were mapped to hg38 human genome with TopHat. Non-unique reads with N≤10 hits in the genome were assigned weights of 1/N, respectively. Reads with >10 hits in the genome were not counted. The frequency of reads was calculated in 100-bp intervals across the genome in fpkm units and visualized using the Integrative Genomics Viewer (IGV). Normalized numbers of sequenced reads (fpkm) matching to exons of a gene were used as a measure of gene expression with a custom Perl program. We filtered out low expressed genes by removing genes with zero fpkm value in wild type or knock-out cells. Because we focused on protein coding genes, non-coding RNAs were also removed. After filtering, 11,159 protein coding genes were used to do further analysis. The *p*-values were calculated by t-test using log-transformed fold changes (normalized to wild type group). For Ribo-seq and RNA-seq, only the genes with *p*-value less than 0.05 and the average of the fold changes > 1.2-fold were selected as differentially expressed genes.

ChIP-seq and PRO-seq reads were mapped to human genome with Bowtie2 allowing maximum of 2 mismatches (*55*). We calculated the intensities of protein binding to specific sub-regions of each gene: TSS region (-700 to +700 bp from TSS), TES region (+100 bp to +3000 bp from TES), and exons excluding TSS region. The TES regions did not include TSS regions or exons of other nearby genes. Statistical analysis for sub-regions of a gene was performed in the same way as for the full genes (all exons).

eCLIP-seq reads were mapped to the genome using TopHat. Other steps were similar to the RNA-seq analysis pipeline. Genes with criterion (WT-IP_fpkm/WT-Input_fpkm > 1.5 and WT-IP_fpkm/TOP3B-KO-IP_fpkm > 1.5) in both biological replicates were selected as TOP3B eCLIP targets. RSeQC was used to analyze the read distribution of Ribo-seq, RNA-seq and eCLIP-seq on CDS, 5'UTR and 3'UTR (*56*). Statistical significance was assessed using Student's *t*-test.

Declarations

Data availability

All next-generation sequencing data will be deposited at GEO, and the accession number will be released.

Contributions

SS, DX, MG and WW designed experiments. SS, YX, JLM, and DX performed experiments. SS, SKL, YZ, AS, WLK, YD, SD, PS and WW analyzed the data. AS, KZ, SD, MG, DX, AS, PS and WW supervised the project. SS and WW wrote the manuscript with the input from all authors.

Acknowledgements

We'd like to thank Dr. David Schlessinger for reading the manuscript and giving helpful advice. This work utilized the computational resources of the NIH HPC Biowulf Cluster and NIA computer servers. This work is supported in part by the Intramural Research Program of the National Institute on Aging (Z01 AG000657-08) and the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Disease (ZO1 DK015602-09), National Institutes of Health.

Competing Interest Statement

The authors declare that we have no competing interest in publishing this work.

References

1. Y. Pommier, Y. Sun, S. N. Huang, J. L. Nitiss, Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat Rev Mol Cell Biol* **17**, 703-721 (2016).

2. G. Stoll *et al.*, Deletion of TOP3beta, a component of FMRP-containing mRNPs, contributes to neurodevelopmental disorders. *Nat Neurosci* **16**, 1228-1237 (2013).

3. D. Xu *et al.*, Top3beta is an RNA topoisomerase that works with fragile X syndrome protein to promote synapse formation. *Nat Neurosci* **16**, 1238-1247 (2013).

4. M. Ahmad, D. Xu, W. Wang, Type IA topoisomerases can be "magicians" for both DNA and RNA in all domains of life. *RNA Biol* **14**, 854-864 (2017).

5. M. Ahmad *et al.*, RNA topoisomerase is prevalent in all domains of life and associates with polyribosomes in animals. *Nucleic Acids Res* **44**, 6335-6349 (2016).

6. Y. Joo *et al.*, Topoisomerase 3beta knockout mice show transcriptional and behavioural impairments associated with neurogenesis and synaptic plasticity. *Nat Commun* **11**, 3143 (2020).

7. Y. Yang *et al.*, Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. *Mol Cell* **53**, 484-497 (2014).

8. T. Zhang *et al.*, Loss of TOP3B leads to increased R-loop formation and genome instability. *Open Biol* **9**, 190222 (2019).

9. M. Ahmad *et al.*, Topoisomerase 3beta is the major topoisomerase for mRNAs and linked to neurodevelopment and mental dysfunction. *Nucleic Acids Res* **45**, 2704-2713 (2017).

10. S. K. Lee *et al.*, Topoisomerase 3beta interacts with RNAi machinery to promote heterochromatin formation and transcriptional silencing in Drosophila. *Nat Commun* **9**, 4946 (2018).

11. S. Saha *et al.*, DNA and RNA Cleavage Complexes and Repair Pathway for TOP3B RNA- and DNA-Protein Crosslinks. *Cell Rep* **33**, 108569 (2020).

12. K. R. Prasanth *et al.*, Topoisomerase III-beta is required for efficient replication of positive-sense RNA viruses. *Antiviral Res* **182**, 104874 (2020).

13. I. Kashima *et al.*, SMG6 interacts with the exon junction complex via two conserved EJC-binding motifs (EBMs) required for nonsense-mediated mRNA decay. *Genes Dev* **24**, 2440-2450 (2010).

14. A. Morettin *et al.*, Tudor Domain Containing Protein 3 Promotes Tumorigenesis and Invasive Capacity of Breast Cancer Cells. *Sci Rep* **7**, 5153 (2017).

15. I. Iossifov *et al.*, De novo gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285-299 (2012).

16. B. Xu *et al.*, De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet* **44**, 1365-1369 (2012).

17. J. D. Riley, C. Delahunty, A. Alsadah, S. Mazzola, C. Astbury, Further evidence of GABRA4 and TOP3B as autism susceptibility genes. *Eur J Med Genet* **63**, 103876 (2020).

18. T. W. Yu, E. Berry-Kravis, Autism and fragile X syndrome. *Semin Neurol* **34**, 258-265 (2014).

19. F. A. Ran *et al.*, Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308 (2013).

20. N. T. Ingolia, G. A. Brar, S. Rouskin, A. M. McGeachy, J. S. Weissman, The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* **7**, 1534-1550 (2012).

21. D. B. Mahat *et al.*, Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). *Nature Protocols* **11**, 1455-1476 (2016).

22. E. L. Van Nostrand *et al.*, Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* **13**, 508-514 (2016).

23. L. Aravind, D. D. Leipe, E. V. Koonin, Toprim--a conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Res* **26**, 4205-4213 (1998).

24. A. F. T. van Scheltinga, S. C. Bakker, R. S. Kahn, Fibroblast Growth Factors in Schizophrenia. *Schizophrenia Bull* **36**, 1157-1166 (2010).

25. A. Takaori, [Antiviral defense by APOBEC3 family proteins]. *Uirusu* **55**, 267-272 (2005).

26. K. Y. Kwan *et al.*, Development of autoimmunity in mice lacking DNA topoisomerase 3beta. *Proc Natl Acad Sci U S A* **104**, 9242-9247 (2007).

27. T. B. Leonid, D. M. Thurtl, J. Rine, A. van Oudenaarden, Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *P Natl Acad Sci USA* **110**, 18602-18607 (2013).

28. A. G. Barbosa *et al.*, Assessment of BDNF serum levels as a diagnostic marker in children with autism spectrum disorder. *Sci Rep-Uk* **10**, (2020).

29. W. R. Blevins *et al.*, Extensive post-transcriptional buffering of gene expression in the response to severe oxidative stress in baker's yeast. *Sci Rep-Uk* **9**, (2019).

30. B. Liu *et al.*, Regulatory discrimination of mRNAs by FMRP controls mouse adult neural stem cell differentiation. *Proc Natl Acad Sci U S A* **115**, E11397-E11405 (2018).

31. S. Brogna, J. Wen, Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol* **16**, 107-113 (2009).

32. Y. Harigaya, R. Parker, No-go decay: a quality control mechanism for RNA in translation. *Wiley Interdiscip Rev RNA* **1**, 132-141 (2010).

33. T. Kurosaki, M. W. Popp, L. E. Maquat, Quality and quantity control of gene expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol* **20**, 406-420 (2019).

34. C. A. Piccirillo, E. Bjur, I. Topisirovic, N. Sonenberg, O. Larsson, Translational control of immune responses: from transcripts to translatomes. *Nat Immunol* **15**, 503-511 (2014).

35. Y. Ishigaki, X. J. Li, G. Serin, L. E. Maquat, Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* **106**, 607-617 (2001).

36. N. Narayanan, Z. H. Wang, L. Li, Y. Z. Yang, Arginine methylation of USP9X promotes its interaction with TDRD3 and its anti-apoptotic activities in breast cancer cells. *Cell Discov* **3**, (2017).

37. P. H. Frederikse, A. Nandanoor, C. Kasinathan, Fragile X Syndrome FMRP Co-localizes with Regulatory Targets PSD-95, GABA Receptors, CaMKIIalpha, and mGluR5 at Fiber Cell Membranes in the Eye Lens. *Neurochem Res* **40**, 2167-2176 (2015).

38. S. Goto-Ito, A. Yamagata, T. S. Takahashi, Y. Sato, S. Fukai, Structural basis of the interaction between Topoisomerase III beta and the TDRD3 auxiliary factor. *Sci Rep-Uk* **7**, (2017).

39. K. Sawicka *et al.*, FMRP has a cell-type-specific role in CA1 pyramidal neurons to regulate autism-related transcripts and circadian memory. *Elife* **8**, (2019).

40. I. F. King *et al.*, Topoisomerases facilitate transcription of long genes linked to autism. *Nature* **501**, 58-62 (2013).

41. J. C. Darnell *et al.*, FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247-261 (2011).

42. A. Merino, K. R. Madden, W. S. Lane, J. J. Champoux, D. Reinberg, DNA Topoisomerase-I Is Involved in Both Repression and Activation of Transcription. *Nature* **365**, 227-232 (1993).

43. R. J. Sims *et al.*, The C-Terminal Domain of RNA Polymerase II Is Modified by Site-Specific Methylation. *Science* **332**, 99-103 (2011).

44. L. Baranello *et al.*, RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. *Cell* **165**, 357-371 (2016).

45. Y. Z. Yang *et al.*, TDRD3 Is an Effector Molecule for Arginine-Methylated Histone Marks. *Molecular Cell* **40**, 1016-1023 (2010).

46. K. Anamika, A. Gyenis, L. Poidevin, O. Poch, L. Tora, RNA Polymerase II Pausing Downstream of Core Histone Genes Is Different from Genes Producing Polyadenylated Transcripts. *Plos One* **7**, (2012).

47. T. M. Wilson, A. D. Chen, T. S. Hsieh, Cloning and characterization of Drosophila topoisomerase III beta - Relaxation of hypernegatively supercoiled DNA. *Journal of Biological Chemistry* **275**, 1533-1540 (2000).

48. S. Das Sharma *et al.*, Widespread Alterations in Translation Elongation in the Brain of Juvenile Fmr1 Knockout Mice. *Cell Rep* **26**, 3313-3322 e3315 (2019).

49. M. V. Gerashchenko, V. N. Gladyshev, Ribonuclease selection for ribosome profiling. *Nucleic Acids Res* **45**, e6 (2017).

50. E. L. Van Nostrand *et al.*, Robust, Cost-Effective Profiling of RNA Binding Protein Targets with Single-end Enhanced Crosslinking and Immunoprecipitation (seCLIP). *Methods Mol Biol* **1648**, 177-200 (2017).

51. A. Barski *et al.*, High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).

52. L. O. Penalva, S. A. Tenenbaum, J. D. Keene, Gene expression analysis of messenger RNP complexes. *Methods Mol Biol* **257**, 125-134 (2004).

53. F. X. Chen, S. A. Marshall, Y. Deng, S. Tianjiao, Measuring Nascent Transcripts by Nascent-seq. *Methods Mol Biol* **1712**, 19-26 (2018).

54. C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111 (2009).

55. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-U354 (2012).

56. L. Wang, S. Wang, W. Li, RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184-2185 (2012).

57. G. E. Siaw, I. F. Liu, P. Y. Lin, M. D. Been, T. S. Hsieh, DNA and RNA topoisomerase activities of Top3beta are promoted by mediator protein Tudor domain-containing protein 3. *Proc Natl Acad Sci U S A* **113**, E5544-5551 (2016).

Figures

Figure 1. Generating *TOP3B-TDRD3 KO* and *TOP3B-Y336F-KI* HCT116 cell lines to study their function.

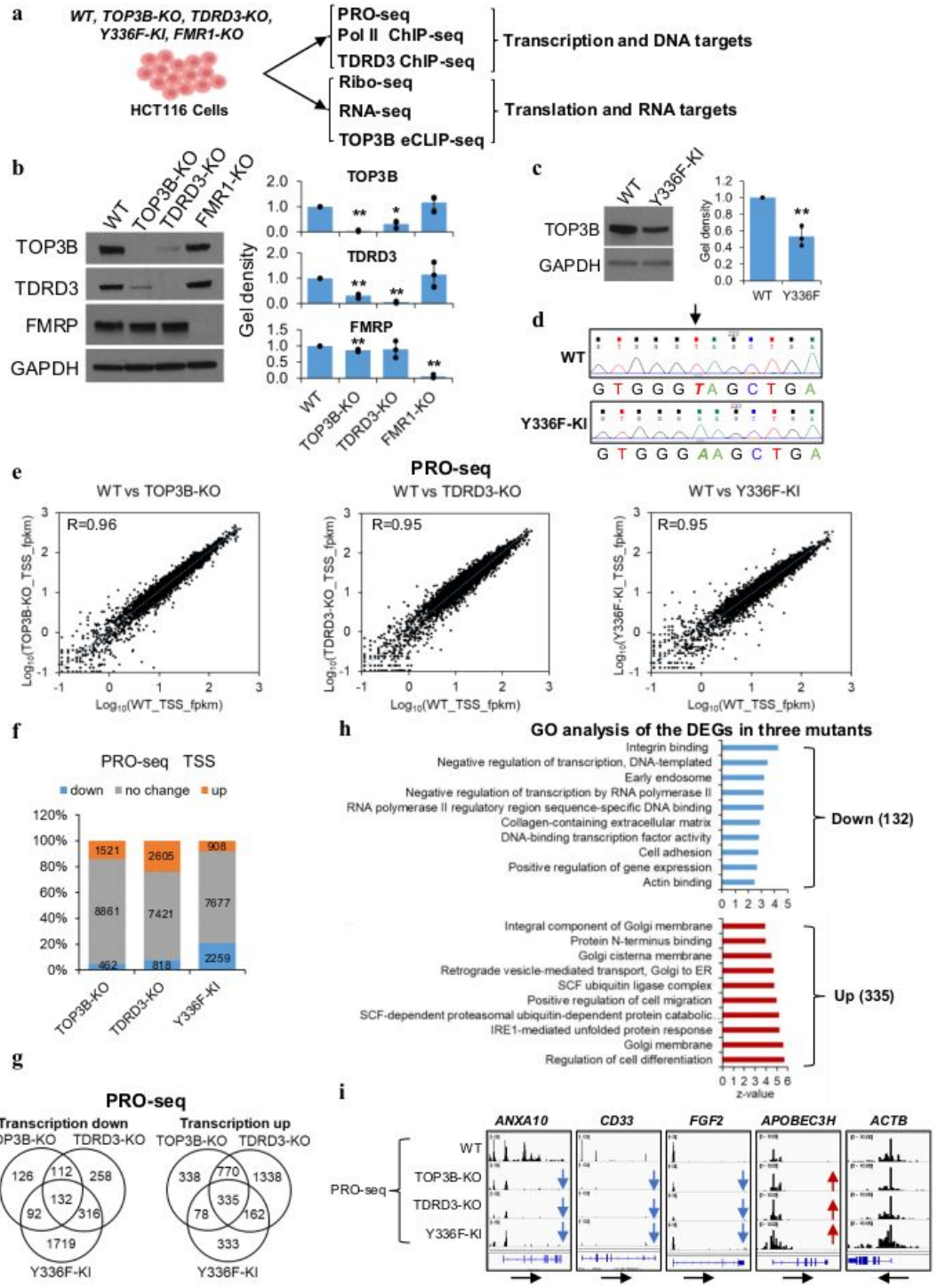


Figure 1

The TOP3B-TDRD3 complex inactivation alters transcription of specific genes a A schematic diagram displaying the experimental design procedure. b, c Immunoblotting gel images (left) and quantification (right) confirm the absence of TOP3B, TDRD3 and FMRP proteins in their respective HCT116 knockout (KO) cells generated by CRISPR-Cas9 gene targeting method. The level of TOP3B-Y336F (c) was about 50% of that of WT cells, suggesting that the catalytic reaction may help to stabilize the topoisomerase

protein. d Genomic sequencing results show the T to A substitution in TOP3B-Y336F-KI cell line generated by CRISPR-Cas9. e Scatterplots showing strong correlations of log-transformed PRO-seq signals at TSS regions between WT and three different mutants. The correlation efficient (R) is shown in each graph. f A bar graph shows the numbers of the genes with altered PRO-seq TSS signals in three mutants (fold change > 1.2-fold by two biological replicates). g A Venn diagram shows the numbers of overlapped and non-overlapped genes with altered PRO-seq signals at TSS regions between TOP3B-KO, TDRD3-KO, and TOP3B-Y336F cells. h A graph shows enriched cellular functions of the 132 genes with downregulated PRO-seq signals (up) or the 335 genes with upregulated PRO-seq signals (down) at TSS regions in all three mutants cell lines by Gene Ontology (GO) analysis. i BedGraphs showing four genes with altered PRO-seq signals in three mutants. ACTB gene is a negative control with similar PRO-seq signals in WT and mutants. For the bar charts, all the values are normalized to WT and log-transformed before P-value calculation. * = P-value < 0.05, ** = P-value < 0.01 (Student's t-test).

Figure 2. The TOP3B-TDRD3 complex binds to both transcription start and termination regions to promote transcription.

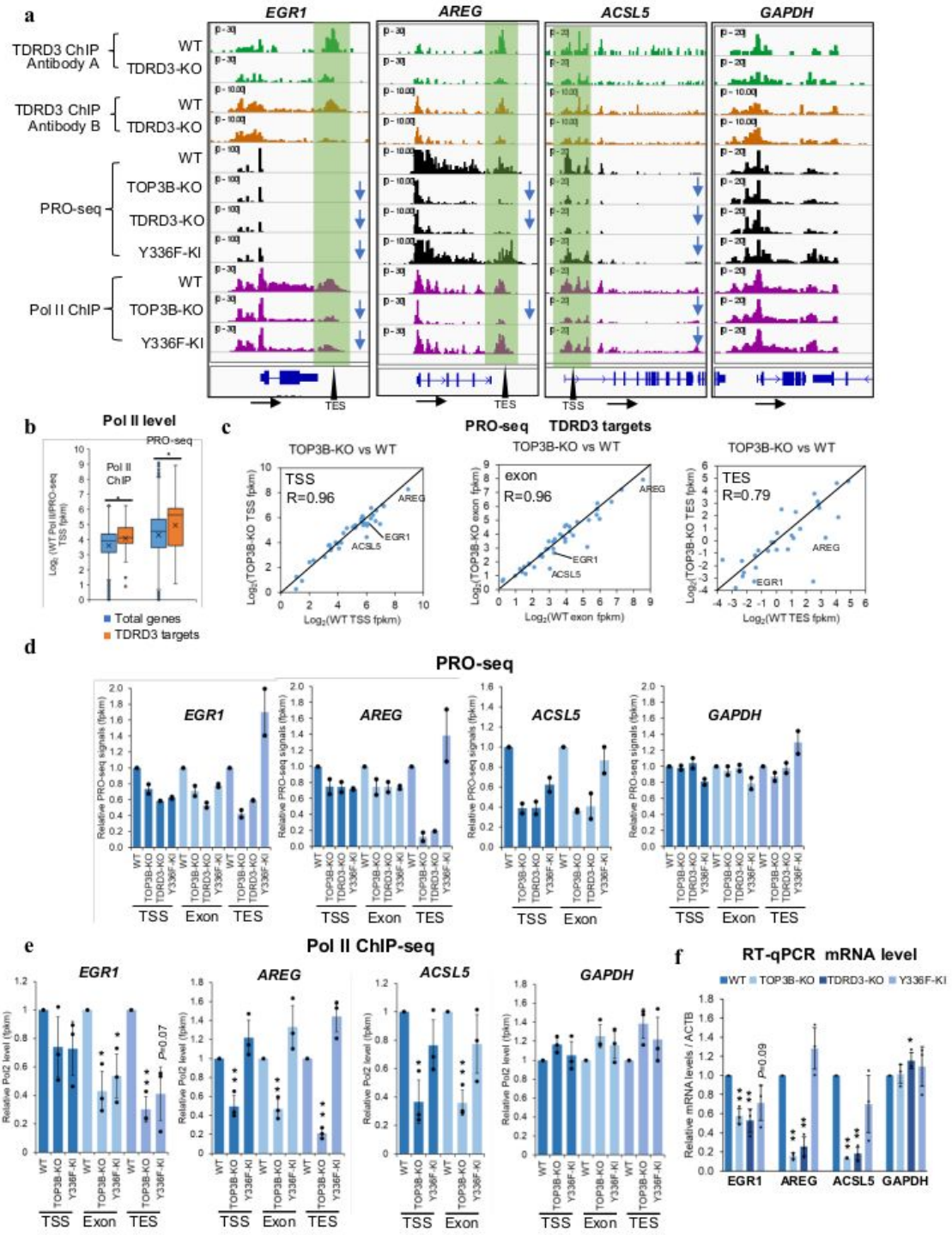


Figure 2

The TOP3B-TDRD3 complex binds to both transcription start and termination regions to promote transcription. a. BedGraphs show that three representative genes that were bound at TES (EGR1 and AREG) or TSS (ACSL5) by the TOP3B-TDRD3 complex, and their transcription was reduced in TOP3B-KO and TDRD3-KO cells based on decreased PRO-seq and/or Pol II ChIP-seq signals. Their transcription was also reduced in TOP3B-Y336F cells for EGR1, partially reduced for ACSL5, but unchanged for AREG gene.

GAPDH gene was included as a negative control. The green color highlights TDRD3 binding peaks that were strong in WT but weak in TDRD3-KO cells (which serve as a negative control). Two different anti-TDRD3 antibodies (A and B) were used in ChIP-seq to ensure the specificity of the antibodies. b A Box and Whisker plot shows that genes bound by the TOP3B-TDRD3 complex based on TDRD3 ChIP-seq have higher transcription levels than those of total genes in WT HCT116 cells. c Scatterplots showing that for the TDRD3 bound genes, their log-transformed PRO-seq signals at TSS, exons or TES correlate with those of WT cells. Several genes with consistent decreased PRO-seq signals were marked (their data points were below the equal line). d, e Bar graphs showing the altered PRO-seq (d) or Pol II ChIP-seq (e) signals of four representative genes at TSS, exons and TES regions in WT and mutant cells. f A bar graph shows the mRNA levels of three TDRD3 target genes by RT-qPCR analysis. For the bar graphs, all the values are normalized to WT and log-transformed before P-value calculation. * = P-value < 0.05, ** = P-value < 0.01 (Student's t-test).

Figure 3. TOP3B regulates translation of specific genes including those important for mental disorders.

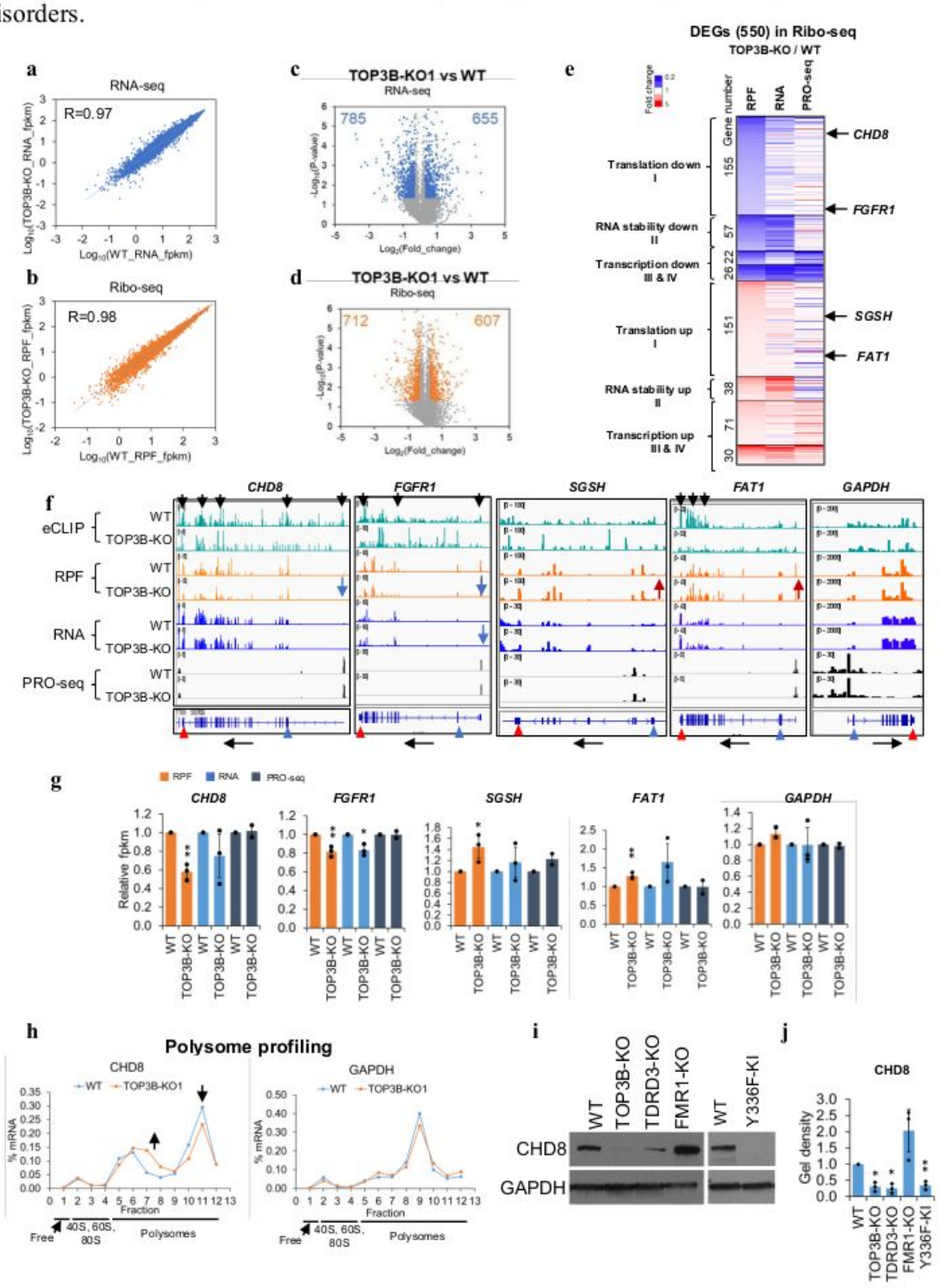


Figure 3

TOP3B regulates translation of specific genes including those important for mental disorders. a, b Scatter plots showing strong correlations of log-transformed RNA-seq (a) or Ribo-seq (b) signals (fpkm) between WT and TOP3B-KO1 cells. The correlation coefficients were marked. c, d Volcano plots showing the numbers of down-regulated or upregulated DEGs in TOP3B-KO1 by RNA-seq (c) or Ribo-seq (d) from three independent experiments. Fold change > 1.2 and P-value < 0.05. e A heatmap shows expression changes

(in fold) between TOP3B-KO vs. WT for the differentially expressed genes (DEGs) identified by Ribo-seq. The DEGs with decreased or increased signals were marked by blue and red, respectively. The colors were transformed from the averages of fold changes. The DEGs were divided into eight groups based on their RPF, mRNA and PRO-seq levels as described in the main text. Several mental diseases related genes altered at the translation level are marked on the right. f, g BedGraphs (f), and their quantifications (g), show that four representative genes with altered RPF levels in TOP3B-KO cells. eCLIP-seq results show the binding of TOP3B to three of them. The RNA-seq and PRO-seq were included for comparison. GAPDH was included as a control. The black arrows above the bedgraphs mark the TOP3B eCLIP peaks that were higher in WT and lower in TOP3B-KO cells (a negative control), whereas those below mark the transcription direction. Blue and red arrows (below) mark the positions of translation start and stop codons separately. The data in (g) were shown in relative fpkm. h Graphs from polysome profiling analysis show the reduced translation of CHD8 mRNA, as evidence by its decreased level in the heavy polysome fractions, and increased level in light polysome fractions. The relative distributions (%) of CHD8 and GAPDH (a control) mRNAs on the sucrose gradients were quantified by RT-qPCR analysis. The results were reproducible and one representative result was shown here. i, j Immunoblotting images (i) and quantification (j) show that CHD8 protein level is reduced in TOP3B and TDRD3 mutant cells, as indicated on top. For the bar graphs, all the values are normalized to WT and log-transformed before P-value calculation. * = P-value < 0.05, ** = P-value < 0.01 (Student’s t-test).

Figure 4. TOP3B requires its topoisomerase activity to regulate translation and mRNA level.

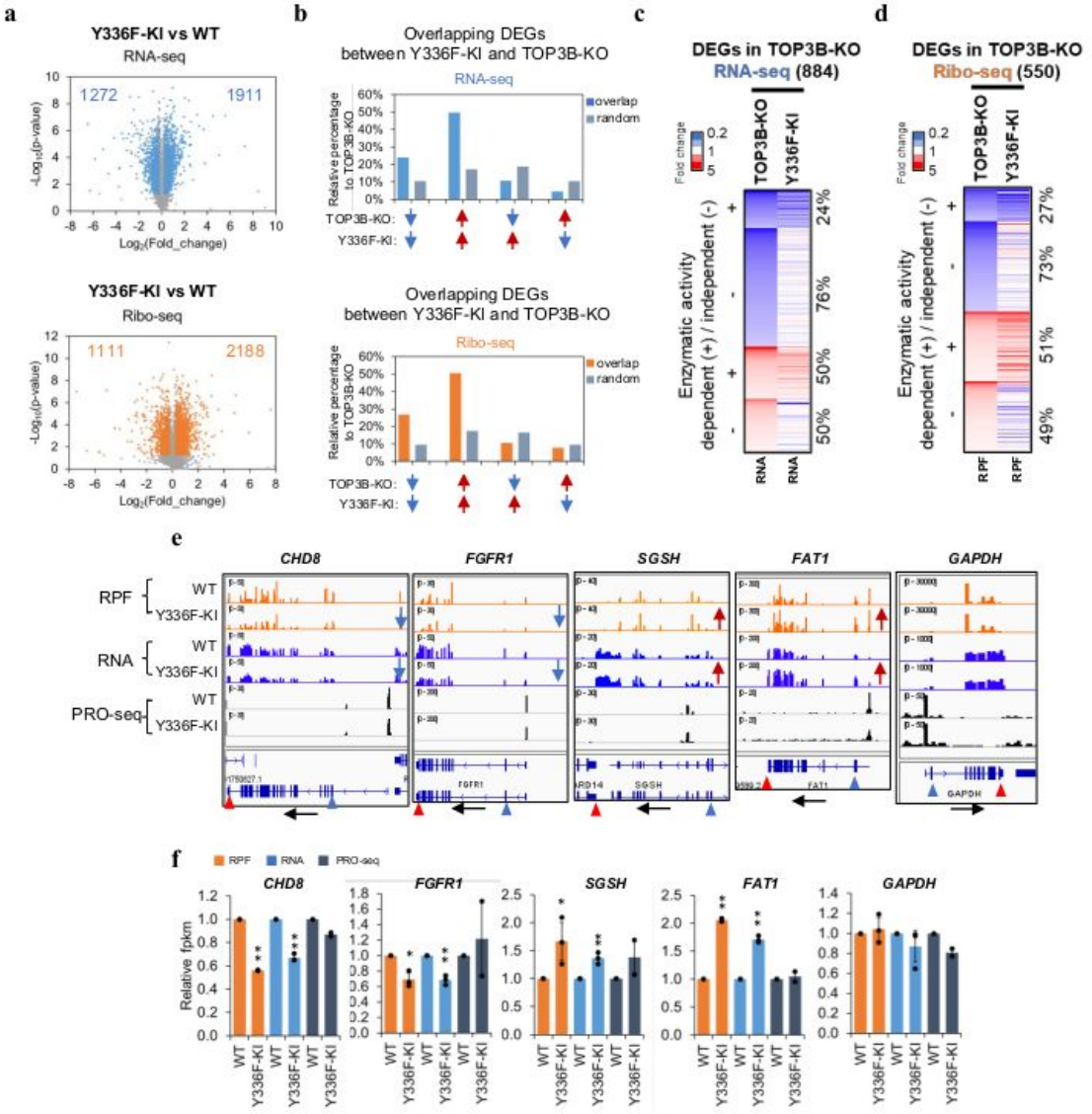


Figure 4

TOP3B requires its topoisomerase activity to regulate translation of specific genes. a Volcano plots showing the number of DEGs in TOP3B-Y336F by RNA-seq (up) or Ribo-seq (down). The threshold of cutoff is: Fold change > 1.2 and P-value < 0.05. b Graphs to compare the percentages of overlapping DEGs in the same or opposite direction of alteration between TOP3B-Y336F and TOP3B-KO groups vs. those randomly selected genes. The percentages were relative to the DEGs numbers of TOP3B-KO group.

The numbers of randomly selected genes were identical to that of the decreased or increased DEGs of TOP3B-Y336F cells. Blue arrows represent reduced, whereas red arrows represent increased DEGs. Arrows in the same direction depicted DEGs that were altered in the same direction in TOP3B-KO and Y336F cells. c, d Heatmaps display the overlapping expression changes (in fold) of the DEGs between TOP3B-KO and Y336F cells by RNA-seq (c) or Ribo-seq (d). The DEGs with decreased or increased signals were marked by blue and red, respectively. The overlapped DEGs that are altered in the same directions are marked as “+”, whereas the others are marked as “-”. The percentages of the overlapping or non-overlapping DEGs were shown on the right. e, f BedGraphs of sequencing read distributions (e), and bar graphs of quantification of these reads (f), show altered Ribo-seq signals for five representative genes. The RNA-seq and PRO-seq data were included for comparison. The alteration for some of them in TOP3B-KO cells were described in Fig. 3g. The bar graphs show the relative fpkm calculations from three biological replicates. * = P-value < 0.05, ** = P-value < 0.01 (Student’s t-test).

Figure 5. TOP3B co-regulates mRNA levels and translation with TDRD3 and FMRP.

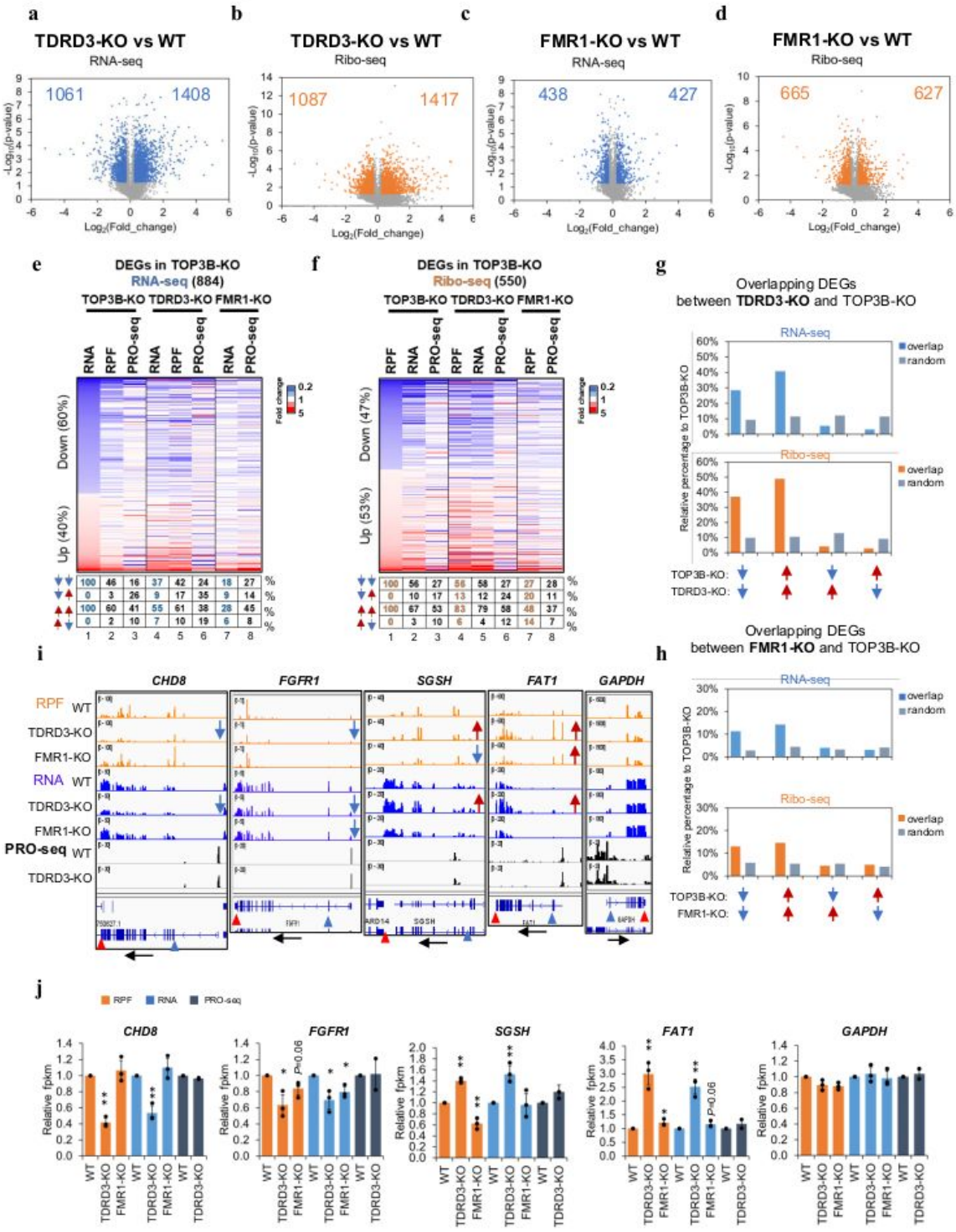
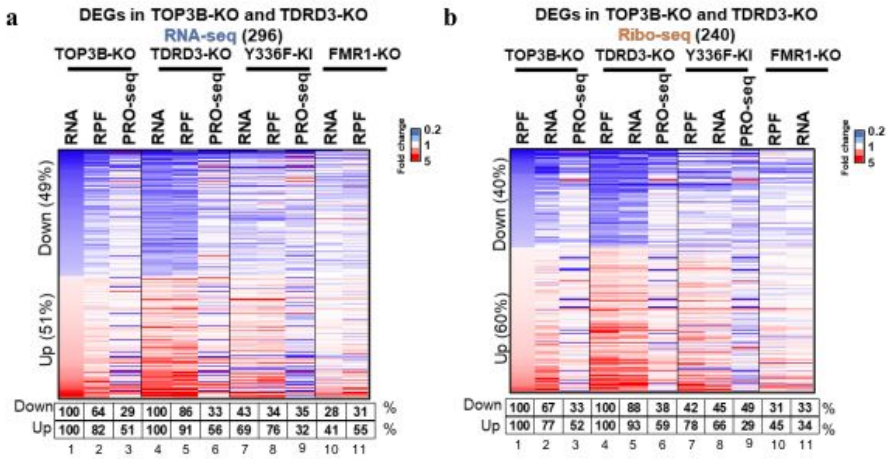


Figure 5

TOP3B co-regulates mRNA levels and translation with TDRD3 and FMRP. a-d Volcano plots showing the number DEGs of TDRD3-KO and FMRP1-KO cells by RNA-seq or Ribo-seq, as indicated on top of the plots. e, f Heatmaps showing the concomitantly decreased (blue color) or increased expression changes of the DEGs in TOP3B-KO vs. those of TDRD3-KO and FMR1-KO cells by RNA-seq (e) or Ribo-seq (f). The percentages were calculated by artificially setting the decreased or increased DEGs of TOP3B-KO cells as

100% (column 1 of each graph). The percentages of the DEGs of TOP3B-KO cells that were altered in the same or opposite directions in other KO cells were shown in a table below the figure. The DEGs identified by different Seq methods from each cell type were included in the analysis. The cutoff threshold for the increased or decreased DEGs is 1.2-fold. Notably, a stronger co-clustering was observed between TOP3B-KO and TDRD3-KO than between TOP3B-KO and FMR1-KO (column 4 vs. 7 in each figure). In addition, a stronger co-clustering was also detected between the levels of RPF and RNA than with PRO-seq for each cell type. g Graphs to compare the percentages of overlapping DEGs in the same or opposite direction of alteration vs. those randomly selected genes between TOP3B-KO and TDRD3-KO cells. The DEGs were identified by either RNA-seq (up) or Ribo-seq (bottom). The percentages were relative to the DEGs numbers of TOP3B-KO cells. The numbers of randomly selected genes were identical to that of the decreased or increased DEGs of TDRD3-KO cells. The percentages of these genes that were decreased or increased in RNA-seq or Ribo-seq of TDRD3-KO cells were then calculated. Blue arrows represented reduced, whereas red arrows represented increased DEGs. Arrows in the same direction depicted DEGs that were altered in the same direction in TOP3B-KO and TDRD3-KO cells. h The DEGs between TOP3B-KO and FMR1-KO were analyzed using the same method as (g). i, j BedGraphs of sequencing read distributions (i), and bar graphs of quantification of these reads (j), show altered Ribo-seq signals for five representative genes in TDRD3-KO and FMR1-KO cells. The RNA-seq and PRO-seq data were included for comparison. The alteration for some of them in TOP3B-KO and Y336F cells are described in Fig. 3g and 4f. The Bar graphs show the relative fpkm calculations from three biological replicates. * = P-value < 0.05, ** = P-value < 0.01 (Student’s t-test).

Figure 6. The TOP3B-TDRD3 complex regulates mRNAs in topoisomerase activity dependent and independent manners.



c DEGs associated with autism or schizophrenia in both *TOP3B-KO* and *TDRD3-KO* HCT116 cells

RPF/RNA	Gene expression	autism	schizophrenia
RPF	down	***TTN***, ***CHD8***, *PRR12*, *SBF1*, *SMC3*, *SYN2*, *PER1*, *TET2*, *KDM5B*	*PRR12*, *SYN2*, *ANKRD13A*, ***FGFR1***
	up	*STK39*, ***ALDH1A3***, *GNAS*, ***SGSH***, ***FAT1***, ***SYT1***, ***IRF2BPL***, *GRIN1*, *SLC1A1*, *BCORL1*, ***CEP290***	***PDE4B***, ***MDK***, *B9D1*, *NT5DC2*, *GPM6A*, ***INPP4B***, ***BDNF***
RNA	down	***ADORA2A***, *PER1*, ***SATB2***, *SBF1*, ***GALNT14***, *MBD6*, ***BCL11A***, *CCDC88C*, *TAF6*	***SATB2***, *OVOL1*, *DUS2*, ***CTRL***, *EDC4*, *DDX28*
	up	*STK39*, ***ALDH5A1***, ***ALDH1A3***, *CAMK4*, ***SYT1***	***PDE4B***, ***AIG1***, ***BDNF***

The genes in bold are regulated by TOP3B-TDRD3 complex in topoisomerase catalytic activity dependent manner.

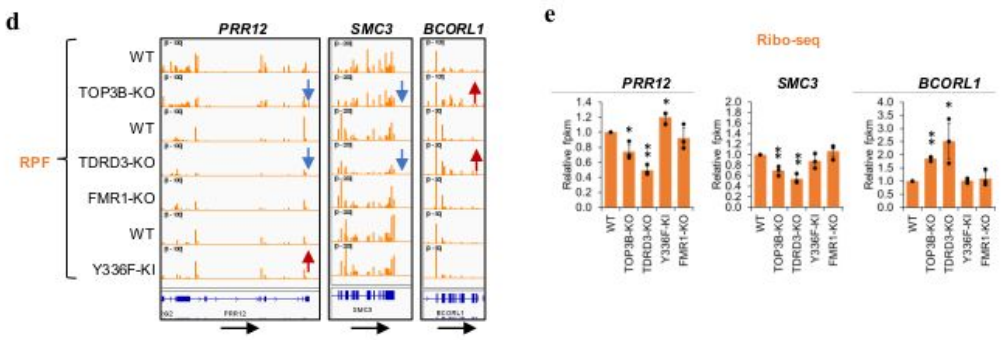


Figure 6

The TOP3B-TDRD3 complex regulates mRNAs in topoisomerase activity dependent and independent manners. a, b Heatmaps showing how the concomitantly decreased (blue) or increased (red) DEGs in both TOP3B-KO and TDRD3-KO cells were overlapped with the DEGs in TOP3B-Y336F and FMR1-KO cells by RNA-seq (a) or Ribo-seq (b). The percentages below the maps were calculated by artificially setting the decreased or increased DEGs of both TOP3B-KO and TDRD3-KO cells as 100% (column 1 and 4 of each

graph). The percentages of the DEGs of TOP3B-Y336F or FMR1-KO cells that were altered in the same directions were shown in a table below the figure. The DEGs identified by different Seq methods from each cell type were included in the analysis. The cutoff threshold for the increased or decreased DEGs is 1.2-fold. Notably, a stronger co-clustering was observed between TOP3B-KO/TDRD3-KO and TOP3B-Y336F than that between TOP3B-KO/TDRD3-KO and FMR1-KO (column 7 vs. 10). In addition, a stronger co-clustering was also detected between the levels of RPF and RNA than with PRO-seq for each cell type. c A table showing the DEGs associated with autism or schizophrenia in both TOP3B-KO and TDRD3-KO HCT116 cells. d, e BedGraphs and bar graphs showing the expression levels of three representative genes with altered RPFs levels in both TOP3B-KO and TDRD3-KO cells. For the bar graphs, all the values were normalized to WT and log-transformed before P-value calculation. The bar graphs show the relative fpkm calculations from three biological replicates. * = P-value < 0.05, ** = P-value < 0.01 (Student's t-test).

Figure 7. eCLIP-seq reveals that TOP3B prefers to bind coding region of long mRNAs and its binding stabilizes these mRNAs.

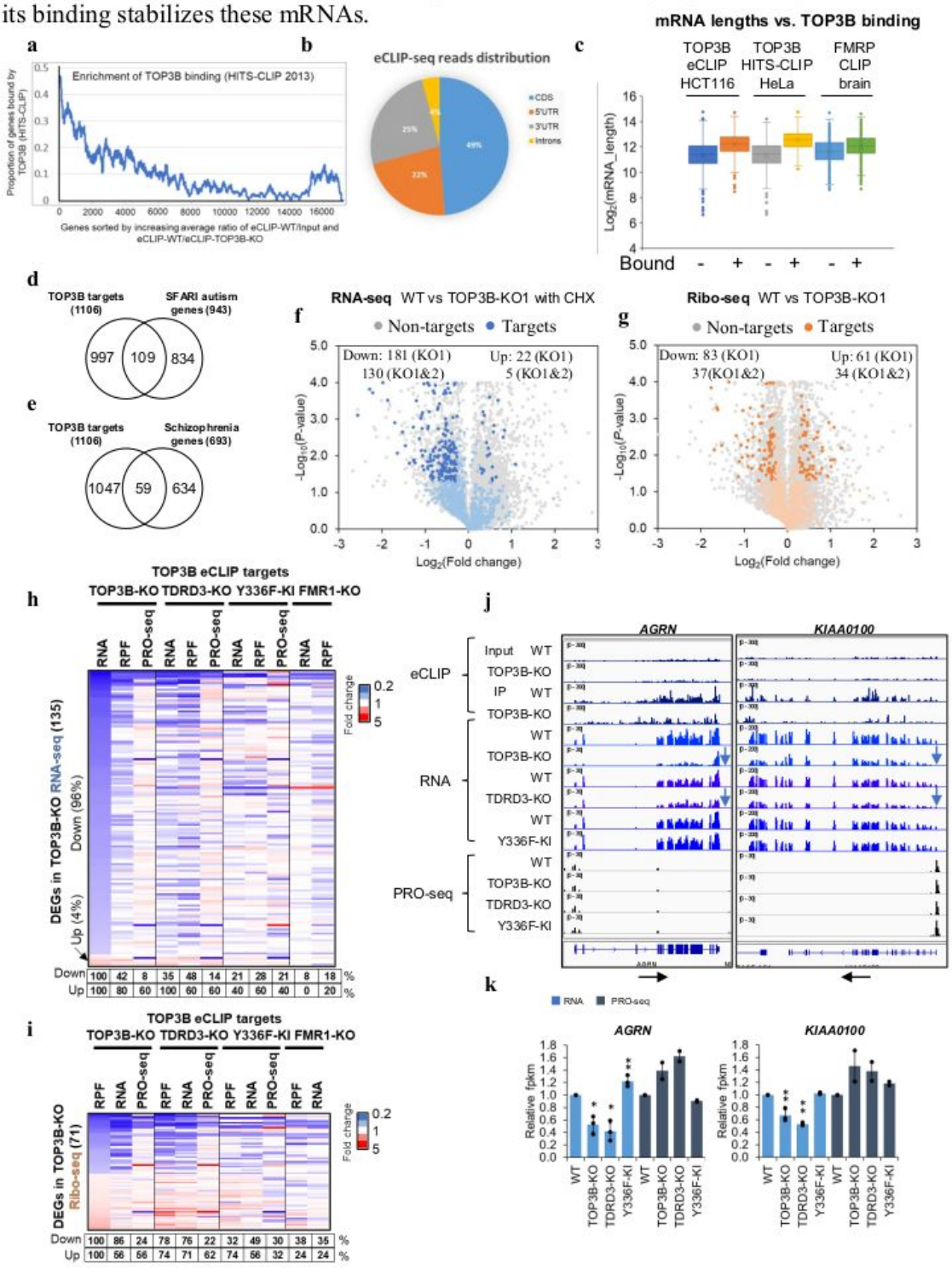


Figure 7

eCLIP-seq reveals that TOP3B prefers to bind coding region of long mRNAs and its binding stabilizes these mRNAs. a A ranked plot shows that genes containing high frequency of TOP3B eCLIP-seq tags from HCT116 cells were enriched with TOP3B HITS-CLIP targets from our previous study in HeLa cells (3). b A pie chart displaying eCLIP-seq read density (tags/kb) distribution in CDS, 5'UTR, 3'UTR and introns. Reads from WT immunoprecipitation group were analyzed using RSeQC. c Box and Whisker plot

showing longer average of lengths of mRNAs bound by TOP3B identified by eCLIP in HCT116, HITS-CLIP in HeLa, or those bound by FMRP identified by HITS-CLIP in mouse brains, when compared to the non-targets (3, 41). The mRNA lengths were log-transformed. d, e Venn diagrams showing the overlapped gene numbers between TOP3B targets with those of SFARI autism genes (d) or those of schizophrenia risk genes (e). f, g Volcano plots showing the expression changes of TOP3B targets by RNA-seq (f) or Ribo-seq (g). h, i Heatmaps showing the expression changes of the TOP3B-bound mRNAs determined by RNA-seq (h) or Ribo-seq (i), in different KO and KI mutant cells indicated on the top. The percentages below the maps were calculated by artificially setting the decreased or increased DEGs of TOP3B-KO cells as 100% (column 1 of each graph). The percentages of these DEGs altered in the same directions in other KO or KI mutant cells were shown in a table below the figure. The DEGs identified by different Seq methods from each cell type were included in the analysis. The cutoff threshold for the increased or decreased DEGs is 1.2-fold. Notably, the percentage of decreased TOP3B-bound mRNAs was much higher than that of increased by RNA-seq (96% vs. 4%; h), whereas about the two percentages by Ribo-seq were roughly equal (i). In addition, a stronger co-clustering was also detected between the levels of RPF and RNA than with PRO-seq for each cell type. j, k BedGraphs of sequencing read distributions (j), and bar graphs of quantification of these reads (k), show altered RNA-seq signals for two TOP3B-bound mRNAs. The bar graphs show the relative fpkm calculations from three biological replicates. All the values were normalized to WT and log-transformed before P-value calculation. * = P-value < 0.05, ** = P-value < 0.01 (Student's t-test).

Figure 8. Models showing the roles of the TOP3B-TDRD3 complex in transcription and translation.

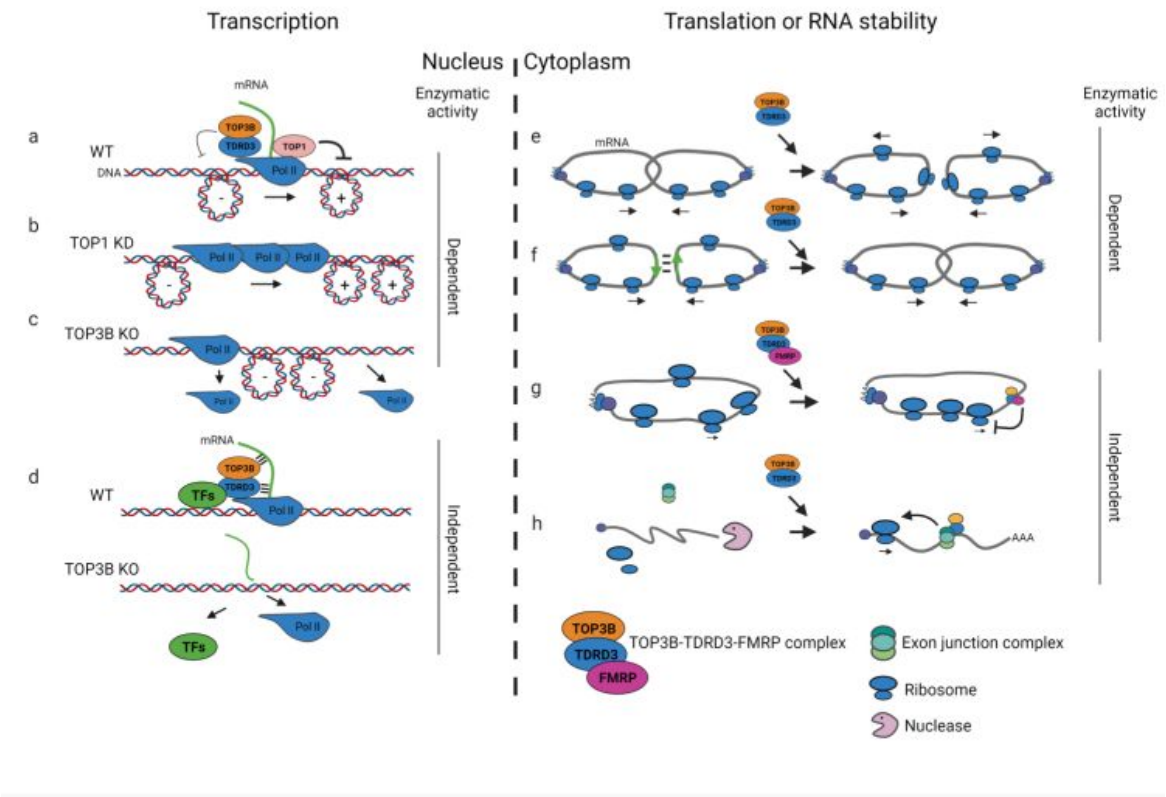


Figure 8

Models showing how the TOP3B-TDRD3 complex regulates transcription and translation. a. A model showing how the TOP3B-TDRD3 complex and TOP1 may stimulate transcription by interacting with pol II (43) and relax positive (TOP1) or negative (TOP3B) supercoils (1). b TOP1 knockdown or inhibition results in accumulation of positive supercoils, which block progression and results increased pol II accumulation in gene bodies and TSS regions (44). c TOP3B knockout increases negative supercoils

(47), which may promote pol II dissociation with DNA template. d The TOP3B-TDRD3 complex may stabilize pol II, mRNA, and DNA complex by interacting with mRNA, pol II and other transcription factors (TFs) in its enzymatic activity independent manner. e The TOP3B-TDRD3 complex may resolve (decatenate) two tangled mRNAs to promote their translation. f For two translating mRNAs with reverse complementary regions, the TOP3B-TDRD3 complex may induce catenation of the two mRNAs to suppress their translation. These models are supported by in vitro assays showing TOP3B can catalyze catenation of RNA circles (3, 57). g, h The TOP3B-TDRD3 complex may suppress translation by recruiting FMRP to its target mRNAs (g), or promote mRNA translation or stability by recruiting exon junction complex (EJC) (h) in enzymatic activity independent manner. Previous evidence has shown that TOP3B-TDRD3 can biochemically and/or genetically interact with FMRP and EJC (2, 3, 13).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementarytableS111.xlsx
- SupplementaryFiguresS1S12.pdf
- accessionnumbersandreviewertokens.pdf
- nrreportingsummary.Su.v2flatten.pdf