Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2024 May 02.

Published in final edited form as:

Cell Rep. 2024 February 27; 43(2): 113809. doi:10.1016/j.celrep.2024.113809.

Mapping catalytically engaged TOP2B in neurons reveals the principles of topoisomerase action within the genome

Amir Segev[1,2,3], **Lance Heady**[1,2,3], **Morgan Crewe**[1,2], **Ram Madabhushi**[1,2,4,*]

[1]Departments of Psychiatry, Neuroscience, and Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX, USA

[2]Peter O' Donnell Brain Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA

[3]These authors contributed equally

[4]Lead contact

SUMMARY

We trapped catalytically engaged topoisomerase IIβ (TOP2B) in covalent DNA cleavage complexes (TOP2Bccs) and mapped their positions genome-wide in cultured mouse cortical neurons. We report that TOP2Bcc distribution varies with both nucleosome and compartmental chromosome organization. While TOP2Bccs in gene bodies correlate with their level of transcription, highly expressed genes that lack the usually associated chromatin marks, such as H3K36me3, show reduced TOP2Bccs, suggesting that histone posttranslational modifications regulate TOP2B activity. Promoters with high RNA polymerase II occupancy show elevated TOP2B chromatin immunoprecipitation sequencing signals but low TOP2Bccs, indicating that TOP2B catalytic engagement is curtailed at active promoters. Surprisingly, either poisoning or inhibiting TOP2B increases nascent transcription at most genes and enhancers but reduces transcription within long genes. These effects are independent of transcript length and instead correlate with the presence of intragenic enhancers. Together, these results clarify how cells modulate the catalytic engagement of topoisomerases to affect transcription.

In brief

Segev et al. map sites of catalytically engaged topoisomerase IIβ (TOP2B) in neurons and observe that TOP2B activity varies with nucleosome configuration and chromosome organization into compartments. Surprisingly, inhibiting TOP2B stimulates nascent transcription at many genes, enhancers, and cryptic sites, revealing how TOP2B catalytic engagement regulates transcription.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

*Correspondence: ram.madabhushi@utsouthwestern.edu.

AUTHOR CONTRIBUTIONS

Conceptualization, R.M.; methodology, R.M., A.S., and L.H.; investigation, all authors; writing, all authors; funding acquisition, R.M., L.H., and M.C.; resources, R.M.; supervision, R.M.

DECLARATION OF INTERESTS

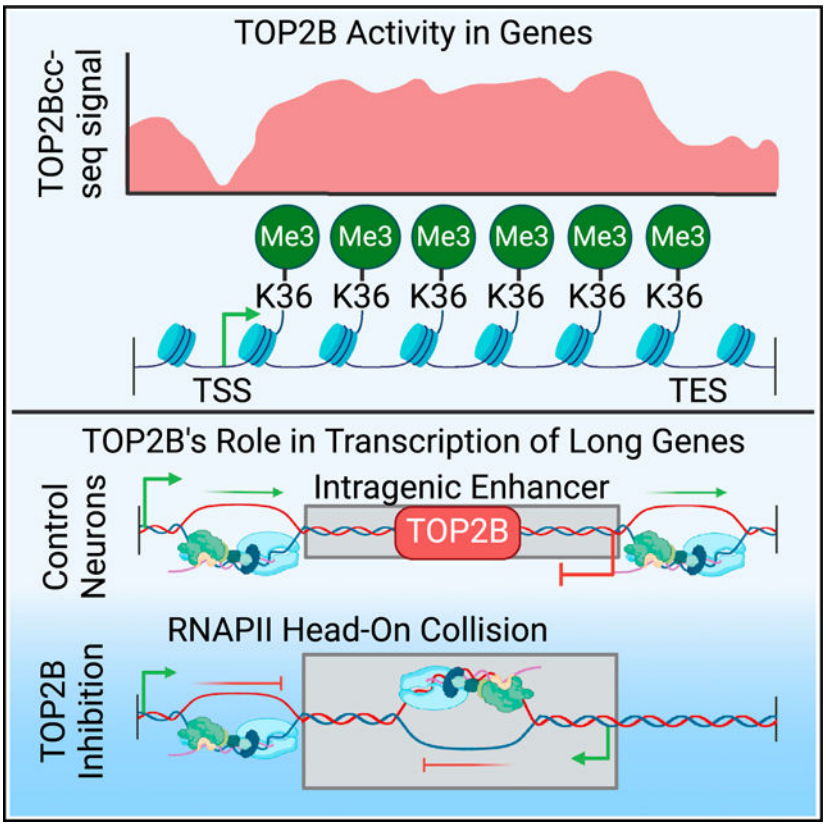
The authors declare no competing interests.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.113809.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Graphical Abstract



INTRODUCTION

Topoisomerases resolve chromosome entanglements and torsional stress that arise during transcription, DNA replication, and chromosome organization by transiently breaking and passing DNA strands through each other.[1,2] Mammalian cells express six topoisomerases, four of which break and rejoin one strand of DNA and are referred to as type I topoisomerases. The remaining two type II enzymes, TOP2A and TOP2B, create and reseal enzyme-linked double-strand breaks (DSBs), through which they pass duplex DNA.[3,4] TOP2A and TOP2B show similar structural and catalytic properties; however, TOP2A is expressed only in proliferating cells, whereas TOP2B is expressed ubiquitously.[5–7] TOP2A activity is crucial for chromosome condensation and segregation,[8,9] while TOP2B is dispensable for cell-cycle progression and cannot compensate for the loss of TOP2A.[10] These results suggest that TOP2A and TOP2B perform distinct functions; yet, the precise cellular roles of TOP2B are poorly understood.

Analysis of *Top2b* knockout mouse models showed that TOP2B is essential for the development and function of postmitotic neurons.[11,12] Because postmitotic cells are free of cell-cycle-related topological problems, these studies alluded to a role for TOP2B in resolving transcription-generated torsional stress. However, only selective developmental and long genes showed reduced expression in the absence of TOP2B activity.[13–16]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Additionally, TOP2B was shown to facilitate the rapid transcription of selective stimulus-responsive genes by forming DSBs within their promoters.[17–24] Yet, how TOP2B-mediated DSBs stimulate transcription and the extent to which this phenomenon defines the functions of TOP2B are unclear. Recent studies have attempted to decipher the roles of TOP2A and TOP2B by mapping sites of either chromatin-bound (using chromatin immunoprecipitation [ChIP]-based methods) or catalytically engaged TOP2.[24–30] The latter typically involves employing the TOP2 poison etoposide (ETP) to trap catalytically active TOP2 in covalent cleavage complexes with the DNA (TOP2ccs). These TOP2ccs are then isolated and processed using various strategies to yield free DNA ends that are then labeled and sequenced. These studies indicate that TOP2B binding and activity are enriched within open chromatin regions, including promoters, the bodies of actively transcribed genes, and at sites occupied by cohesin and CTCF, which mediate chromatin looping.[31] Additionally, loop anchors located within and around transcriptionally active genes were found to be vulnerable to TOP2-induced DNA breakage and chromosome translocations that underlie ETP chemotherapy-induced acute myeloid leukemia.[25,30,32]

Despite these insights, systems and methods used to map TOP2ccs and TOP2-mediated DNA breaks thus far either have not differentiated between the activities of TOP2A and TOP2B or have used array-based methods, which lack sufficient resolution. Here, to determine how TOP2B acts within the genome, we immunoprecipitated TOP2Bccs from ETP-treated primary mouse cortical neurons (which do not express TOP2A) and sequenced the associated DNA. We report that TOP2B activity within the genome correlates with nucleosome configuration and compartmental organization of chromosomes but not with topologically associated domains (TADs) and loop anchors. Like TOP1,[33] TOP2Bccs are enriched within actively transcribed gene bodies but depleted near transcription start sites (TSSs), indicating the existence of a common strategy to restrain TOPs at the TSS. Surprisingly, nascent transcription at highly expressed genes is enhanced by either the poisoning or inhibition of TOP2B, revealing a role for TOP2B in regulating transcription initiation.

RESULTS

Mapping genome-wide sites of catalytically engaged TOP2B in neurons

To decipher the roles of TOP2B, we incubated cultured mouse cortical neurons (13 days *in vitro*) with the TOP2 poison ETP (50 μM for 30 min) to trap catalytically engaged TOP2B in covalent DNA cleavage complexes (TOP2Bccs). Neurons were then rapidly lysed with sarkosyl, and genomic DNA was sheared and centrifuged through cesium chloride gradients to separate TOP2Bccs from non-covalently bound proteins (Figure 1A).[34] The pelleted genomic DNA was resuspended and sonicated to ensure uniform fragmentation, and TOP2Bcc-containing fragments were purified by performing native TOP2B ChIP. Because DNA fragments containing TOP2Bccs are capped by TOP2B at one end (Figure 1A), Illumina-barcoded adapters were first ligated to the free DNA end. Attempts to amplify libraries for sequencing after this step alone proved unsuccessful, consistent with the successful isolation of TOP2Bccs. Following the first round of adapter ligation, the 5′ strand was digested with λ exonuclease, and primers against the ligated adapter were used

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

to perform ten rounds of primer extension, Ribo-G tailing of the single-stranded DNA products, and ligation of a double-stranded adapter with a CCC overhang, as previously described[35] (Figure 1A). Libraries were then amplified and sequenced to map genome-wide sites of catalytically engaged TOP2B (hereafter called TOP2Bcc-seq). As a control, input DNA was also sequenced, and input-normalized signals from replicate experiments were concordant (Figure 1B). Based on this, the MACS2 pipeline[36] was utilized to call peaks (n = 35,505) of TOP2B activity genome-wide and to generate a merged input-normalized TOP2Bcc-seq signal file from two replicate experiments that was used for subsequent analysis (STAR Methods) (Figure 1C).

TOP2Bcc-seq signals generally correlated with TOP2B -ChIP-seq signals generated previously from cultured cortical neurons (Spearman = 0.64; Figure 1D),[24] although TOP2Bcc-seq peaks were relatively depleted at promoters and elevated within gene bodies compared to TOP2B ChIP-seq peaks (Figure 1E). Analysis of motifs enriched at TOP2Bcc-seq peaks using MEME-ChIP[37] revealed an AT-rich sequence that is consistent with findings that TOP2B regulates the transcription of neuronal genes within AT-rich environments[38] (Figure 1F). Processing of TOP2Bccs stabilized by ETP form DNA single-stranded breaks and DSBs that trigger the phosphorylation of the histone variant, H2AX, in the vicinity of DSB sites.[24,39] Phosphorylated H2AX (γH2AX) ChIP-seq signals were elevated near TOP2Bccs in ETP-treated neurons compared to γH2AX signals in control neurons, indicating DSB formation at TOP2Bcc peaks (Figure 1G). These results validate that TOP2Bccs mapped in our study correspond to sites of catalytically engaged TOP2B in neurons.

TOP2B activity is enriched within transcribed regions but depleted in active promoters

To further investigate how catalytically engaged TOP2B distributes within the neuronal genome, we first classified neuronal chromatin into 15 distinct states (various classes of promoters, enhancers, transcribed regions, and heterochromatin) based on the combinatorial occupancy patterns of eight key histone marks using ChromHMM (Figures S1A, S1B, and 2A).[40,41] The distribution of TOP2Bccs within each of these states was then assessed relative to their surrounding regions and relative to TOP2B occupancy based on available TOP2B ChIP-seq data.[24] Both TOP2Bcc-seq and TOP2B ChIP-seq signals were broadly distributed within the genome and were detectable to some extent at each of the 15 chromatin states (Figures 2A–2G); however, several trends became apparent. Chromatin states annotated as no signal heterochromatin (NS) and transcription permissive (Tr-P), which are depleted in all eight of the assessed histone marks and cover most of the genome (Figure S1A), also showed reduced TOP2Bcc-seq signals compared to their surrounding regions (Figures 2A and 2B). Meanwhile, a unique pattern of TOP2Bcc-seq signals was detected in H3K9me3-rich heterochromatin (Hc-H) (Figure 2C). When compared to other states, such as H3K27me3-rich polycomb heterochromatin (Hc-P), TOP2Bcc-seq signals were markedly lower at Hc-H and their surrounding regions (Figure 2C). However, within this context, TOP2Bcc-seq signals were enriched at Hc-H both when compared to adjoining regions (Figure 2A) and relative to TOP2B ChIP-seq signals within the same regions (Figure S1C). These results suggest that H3K9me3 could facilitate the catalytic engagement of TOP2B in chromatin environments that are otherwise less permissive to TOP2B. Notably,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the three states with the lowest TOP2Bcc-seq signals (Hc-H, NS, and Tr-P) are also depleted in active transcription, suggesting that TOP2Bcc levels within chromatin could be influenced by transcription. Accordingly, chromatin states associated with actively transcribed regions, such as strong transcription (Tr-S), transcription initiation (Tr-I), and promoter flanking (Pr-F), which are typically found in gene bodies, as well as several enhancer states, were enriched for TOP2Bccs (Figures 2A, 2D–2F, and S1D). In fact, relative to TOP2B occupancy, TOP2Bccs were highly enriched at Tr-S (Figure S1C).

Previous reports suggest that TOP2 activity within gene promoters could support transcription by resolving negative supercoils behind RNA polymerase II (RNAPII).[42] TOP2A[43] and TOP2B[24,28] binding was shown to be enriched within gene promoters in several cell types, although others found that 2%–3% of TOP2B ChIP-seq peaks reside within 5 kb upstream of the TSS.[27] These results prompted us to compare TOP2Bcc-seq and TOP2B ChIP-seq distributions at neuronal gene promoters. TOP2Bccs were depleted at both active and weak promoters in neurons despite enriched TOP2B occupancy (Figures S1C, 2G, and 2H). By contrast, TOP2Bccs were enriched in bivalent promoters (Pr-B) (Figure 2G). These results suggest that TOP2B bound at active promoters is largely catalytically disengaged. Intriguingly, previous reports indicate that TOP1 is held inactive at promoter-proximal regions containing paused RNAPII.[33] To further explore TOP2B catalytic engagement in the vicinity of RNAPII, the distribution of TOP2Bcc-seq signals was assessed relative to RNAPII ChIP-seq peaks from publicly available ENCODE data.[44] As expected, a significant fraction of RNAPII peaks were located within promoters, and RNAPII ChIP-seq signals were enriched at the TSS of expressed genes (Figures S1D and S1E). TOP2Bcc-seq signal intensity was reduced at RNAPII peaks as a function of RNAPII peak strength (Figure 2I). These results indicate that whereas topoisomerases act within gene bodies to resolve supercoiling, their catalytic engagement is restrained within active promoters.

TOP2Bccs are not enriched at most TAD boundaries and chromatin loop anchors

Beyond nucleosomes, chromosomes fold into loops and into units of self-interacting chromatin segments called TADs.[45–47] Chromatin loops and TADs are formed by cohesin, which reels in DNA bidirectionally and extrudes loops. Extrusion of loops halts when cohesin encounters the architectural protein CTCF, bound to inwardly oriented CTCF motifs.[31,46] TOP2B interacts with CTCF and cohesin, and TOP2 activity is enriched at CTCF and cohesin binding sites, indicating that TOP2 could resolve topological constraints at loop anchors.[24–26,28,29] TOP2-mediated DNA breaks were enriched at loop anchors in B cells[26]; however, its activity did not correlate with loop strength in human retinal pigment epithelial (RPE1) cells.[29] These observations prompted us to assess whether catalytically engaged TOP2B is enriched at TAD boundaries and chromatin loop anchors in neurons.

To this end, we overlaid TOP2Bcc-seq signals onto Hi-C maps generated from cultured mouse cortical neurons.[48] Heatmap analysis showed that TOP2Bcc-seq signals do not vary as a function of TAD strength, while aggregate plots revealed a reduction in TOP2Bccs immediately outside TADs (Figures 3A and S2A). To identify factors that affect TOP2Bcc distribution at TAD boundaries, we used k-means clustering to inspect TOP2Bcc-seq signals

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

at TADs and their adjoining regions and isolate TADs whose boundaries are particularly depleted in TOP2Bccs. This analysis indicated that TOP2Bcc patterns at TADs partition into two broad groups and that most TADs show no change in TOP2Bccs at their boundaries (group 1, 71.4%; Figure 3B). The remaining TADs (group 2; 28.6%) and their adjacent regions were relatively depleted of TOP2Bccs, and here, TOP2Bcc-seq signal intensity was again elevated within TADs and reduced immediately outside the TAD boundaries (Figure 3B). Comparison of chromatin states that populate group 1 and group 2 TADs revealed that group 2 TADs are enriched in Hc-H and depleted of chromatin states containing higher TOP2Bccs, such as Pr-B, Pr-F, Tr-I, and Tr-S, while group 1 TADs contain most chromatin states but are depleted of Hc-H (Figure 3C). Like at TADs, TOP2Bcc-seq signals were also unchanged at most loop anchors (group 1; 66.5%) that were identified from available Hi-C data and were elevated at a subset of loop anchors (group 2; 33.5%) that were defined by low TOP2Bcc-seq signals overall (Figures 3D and S2B). Group 2 loop anchors were also enriched in Hc-H, although unlike group 2 TADs, they were also enriched for promoter and promoter-proximal enhancer states (Figure S2C). These results suggest that levels of catalytically engaged TOP2B are largely independent of loop anchors and TAD boundaries, except for those found in H3K9me3 heterochromatin.

As mentioned above, TOP2B associates with CTCF, and while CTCF binding outnumbers TADs and chromatin loop anchors within the genome, it borders a subset of TADs and loops and is important for chromatin looping.[47,49] To explore how TOP2Bccs distribute relative to CTCF and chromatin loops, we performed CTCF ChIP-seq in cultured cortical neurons.[50] By intersecting intervals of consensus CTCF binding motifs within the mouse genome with CTCF peaks, we obtained a list of CTCF motifs that were occupied by CTCF in neurons. As previously reported for TOP2,[26,29] analysis of TOP2Bcc-seq signals relative to the center of the oriented CTCF motif revealed two peaks where TOP2Bcc-seq signals were concentrated (Figure S2D). Out of 32,390 CTCF motifs containing CTCF peaks, 15,046 peaks (~46%) were found at loop anchors in neurons, and ~28% loop anchors (11,567/42,000) contained CTCF motifs with CTCF peaks. The pattern and intensity of TOP2Bcc-seq signals at CTCF peaks were similar regardless of their presence at loop anchors (Figure 3E). However, the relationship between CTCF and TOP2Bccs was context dependent. At group 1 loop anchors, where TOP2Bccs were invariant compared to surrounding regions, the presence of CTCF peaks within loop anchors (13,838/34,040) did not lead to further increases in TOP2Bccs (Figure 3F). By contrast, TOP2Bccs were relatively elevated at group 2 loop anchors containing CTCF peaks compared to those lacking CTCF peaks, indicating that CTCF occupancy correlates with elevated TOP2Bccs at group 2 loop anchors (Figure 3F). Finally, reanalysis of END-seq data generated from ETP-treated neurons[26] suggested that END-seq signals were highly enriched at CTCF peaks but relatively unaffected at loop anchors (Figure S2E). Taken together, our data indicate that while TOP2Bccs are elevated at CTCF binding sites, their levels are unchanged at most TAD boundaries and loop anchors. Furthermore, the patterns of TOP2Bcc-seq signals in group 2 TADs and loop anchors are consistent with their chromatin state composition and suggest that chromatin marks within TADs and loop anchors, and not the TAD/loop organization itself, could affect TOP2B activity within these regions.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Chromosome compartments define the threshold for TOP2B catalytic engagement

In addition to TADs, transcriptionally active and inactive regions segregate into chromosome compartments, called compartments A and B, respectively, that visualize as plaid patterns in Hi-C maps and exhibit distinct biophysical properties.[46,51] To understand whether TOP2Bccs vary according to chromosome compartments, principal-component analysis of Hi-C data from cultured cortical neurons was performed and aligned with available ATAC-seq data to define chromosome compartments in cultured cortical neurons (STAR Methods). TOP2Bcc-seq signals were elevated within chromosome compartment A compared to compartment B, indicating that levels of catalytically engaged TOP2B partition according to chromosome compartments (Figures 4Afig4, 4B, and S3A). To further assess how compartmental organization affects TOP2Bccs, we grouped intervals comprising the 15 distinct chromatin states in neurons according to their presence in either compartment A or B and assessed the distribution of TOP2Bccs. For any given state, TOP2Bcc-seq signals were always lower for intervals in compartment B compared to those in compartment A, but the distribution of TOP2Bcc-seq signals at and around each chromatin state was consistent between compartments. For instance, TOP2Bccs were elevated within Tr-S intervals relative to neighboring regions in both compartments A and B, while TOP2Bcc-seq signal intensity was reduced in Tr-S intervals found within chromosome compartment B (Figure 4C). Similar patterns were observed for other chromatin states (Figures S3B–S3O), and the effects of compartmental organization on overall TOP2Bcc-seq signals was particularly evident in low transcriptional activity states, such as Hc-H and NS (Figures S3N–S3P). Finally, a majority of group 1 TADs and group 1 loop anchors mapped to compartment A, whereas group 2 TADs and group 2 loop anchors were more prevalent within compartment B (Figures 4D and 4E). These observations suggest that whereas TOP2Bccs vary with nucleosome configuration in chromatin states, the threshold of catalytically engaged TOP2B within these regions is set at the level of chromosome compartments.

Nucleosome configuration, and not transcription itself, predicts TOP2Bcc levels within genes

The enrichment of TOP2Bccs within gene bodies and transcriptionally active compartment A suggests that TOP2B could affect transcription by resolving transcription-generated supercoiling. To assess how TOP2Bcc-seq signals relate to transcription, we performed global run-on sequencing (fastGRO) under basal conditions and following ETP treatment of cultured cortical neurons.[52] Unlike array-based methods and RNA-seq, which measure steady-state RNA, fastGRO and related methods precisely capture the positions, levels, and orientation of transcriptionally engaged RNAPII.[53] We reasoned that comparing how TOP2Bccs distribute relative to transcriptionally engaged RNAPII and how ETP treatment affects RNAPII could elucidate the roles of TOP2B in transcription. To this end, neuronal nuclei were rapidly isolated from ETP-treated neurons (50 μM, 30 min) and untreated controls and permeabilized at 4°C, and run-on reactions were performed in the presence of 4-thio ribonucleotides (4-S-UTP). Reverse biotinylation was then employed to label and purify nascent RNAs that were subsequently identified through strand-specific library preparation and sequencing.[52] fastGRO signals between replicate experiments were highly correlated (Figures S4A and S4B), and aggregate profiles of fastGRO signals at highly transcribed genes that were defined by performing RNA-seq showed coverage throughout

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the gene body (Figure S4C). fastGRO signals were also reduced at RNA-seq-defined silent genes compared to expressed genes (Figure S4D). These results indicate that fastGRO signals mark sites of transcriptionally engaged RNAPII in neurons.

Following this initial validation, fastGRO FPKMs (fragments per kilobase million) from untreated neurons were used to rank and group genes into quartiles of expression (Q1 – low expression, Q4 – high expression). fastGRO signals from untreated and ETP-treated neurons were then examined at genes within each expression quartile. Surprisingly, fastGRO signals in ETP-treated neurons trended higher than those in untreated controls across expression quartiles (Figures 5A, 5B, and S4E). The difference in fastGRO signals between the two groups increased with increasing transcriptional activity within each quartile (Figures 5A and 5B). fastGRO signals in ETP-treated neurons were also elevated throughout the length of transcripts, including at the TSS, indicating that TOP2B poisoning stimulates new rounds of transcription, especially at highly expressed genes. In addition to gene bodies, transcription occurs at numerous proximal and distal gene regulatory elements, such as enhancers,[54] and TOP2Bcc-seq signals are also elevated at several enhancer classes. fastGRO signals at enhancer states defined by chromatin marks (Figure S1A) were elevated in ETP-treated neurons compared to controls, suggesting that TOP2B poisoning also could also stimulate nascent transcription at enhancers (Figure S5A). To further address this issue, we utilized dREG,[55] a machine learning program that employs support vector regression, and identified active transcriptional regulatory elements (TREs) from fastGRO signals in untreated and ETP-treated neurons (Figure S5B). Implementation of dREG identified 16,785 and 16,336 TREs in untreated controls and ETP-treated neurons, respectively. TREs in ETP-treated neurons distributed similarly to TREs in untreated controls across the 15 chromatin states (data not shown). fastGRO signals at TREs identified by dREG were elevated in ETP-treated neurons compared to controls (Figures S5C and S5D), indicating that TOP2B inhibition also stimulates transcription at TREs, including enhancers.

Next, to investigate whether there were genes whose transcription requires TOP2B activity, raw fastGRO counts from ETP-treated neurons and controls were directly compared using DESeq2.[56] This analysis identified 531 genes that showed at least a 1.5-fold reduction in transcriptionally engaged RNAPII in ETP-treated neurons (hereafter referred to as ETP-downregulated genes; Figure S5E). Comparison of fastGRO signals at ETP-downregulated genes in untreated neurons with genes segregated according to quartiles of expression (Q1, low expression; Q4, high expression) revealed that ETP-downregulated genes are expressed at levels higher than those of Q3 genes but lower than Q4 genes (Figure 5C). Gene body TOP2Bcc-seq signals tracked with the level of transcription in each quartile (Q4 > Q3 > Q2 > Q1) (Figure 5D). However, TOP2Bcc-seq signal intensity in ETP-downregulated genes was lower than that expected for their expression level and was similar to Q2 genes (Figure 5D). Furthermore, despite being highly expressed in untreated neurons, ETP-downregulated genes were depleted of chromatin states that mark strong transcription, including Pr-F, Tr-I, and Tr-S, compared to Q3 genes (Figure 5E). These results indicate that ETP-downregulated genes display a distinct chromatin signature compared to other similarly expressed genes.

The depletion of TOP2Bccs at the highly expressed ETP-downregulated genes suggested that TOP2Bcc levels within gene bodies may not be related to their level of transcription per

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

se. Although aggregate TOP2Bcc-seq signals were elevated within gene bodies according to their quartile of expression (Figure 5D), TOP2Bcc-seq signals within individual genes correlated neither with fastGRO FPKMs nor with RNA-seq TPMs from untreated neurons (Figures 5F and 5G). Because the bodies of actively transcribed genes are also enriched in the chromatin states Pr-F, Tr-I, and Tr-S, we next examined how the enrichment of these chromatin states affects TOP2Bcc-seq signals within gene bodies. Mean TOP2Bcc-seq signal intensity within gene bodies tracked with the proportion of genes containing Pr-F, Tr-I, and Tr-S states within each quartile (Figure 5H). Coordinates for genes within each quartile were then intersected with Pr-F, Tr-I, and Tr-S intervals, and TOP2Bcc-seq signals were assessed at genes that either contained or lacked these states. Remarkably, whereas TOP2Bcc-seq signals tracked with the level of transcription in genes that contained Pr-F, Tr-I, and Tr-S, this association was lost in genes that lacked these states (Figures 5I and 5J). These results suggest that TOP2B activity within gene bodies is influenced by their nucleosome configuration. Genes depleted in specific chromatin states (Pr-F, Tr-I, Tr-S) display reduced TOP2Bccs despite being highly expressed, and this imbalance could cause them to be particularly reliant on TOP2B activity within their bodies. On the other hand, the elevated nascent transcription in ETP-treated neurons suggests that TOP2B tempers transcriptional output, especially at highly expressed genes.

The presence of intragenic enhancers affects long gene transcription in ETP-treated neurons

Whereas nascent transcription in Q3 and Q4 genes was elevated in ETP-treated neurons, how other genes, which are also highly expressed in untreated neurons, instead become downregulated following ETP treatment remained unclear. Previous reports indicate that TOP1 and TOP2B are dispensable for the expression of most genes but are essential for the transcription of long genes, particularly in neurons.[15,57] Consistent with these observations, the median length of ETP-downregulated genes was substantially greater than those of all expressed genes in neurons, including the similarly expressed Q3 genes and ETP-upregulated genes (Figure S6A). To further understand how gene length relates to their regulation by ETP, expressed genes in neurons were grouped into various classes according to their length. fastGRO signals in these genes were then compared as a function of ETP treatment. Mean fastGRO signals in ETP-treated neurons were elevated across gene length categories except for genes longer than 80 kb, which showed the opposite trend (Figure S6B). fastGRO signal intensity was elevated at the TSS and beyond the zone of proximally paused RNAPII in ETP-treated neurons regardless of gene length category, indicating that TOP2B inhibition stimulates new cycles of transcription (Figure 6A). At genes shorter than 20 kb, fastGRO signals were elevated throughout the gene body in ETP-treated neurons compared to controls (Figure 6A). Thereafter, however, fastGRO signals were gradually reduced in ETP-treated neurons as a function of gene length, with decreases relative to controls first observed at transcription end site (TES)-proximal regions, which then extended toward TSS-proximal regions in genes longer than 80 kb (Figure 6A).

A potential explanation for these results is that the transcription of long genes generates enough dynamic supercoiling to stall RNAPII in ETP-treated neurons. If this were the case, then it can be expected that TOP2Bccs would be elevated within the bodies of long

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

genes compared to short genes. Furthermore, this scenario implies that in sufficiently long genes, newly elongating RNAPII should stall at relatively similar distances from the TSS in ETP-treated neurons. To test whether this was the case, we first assessed how TOP2Bcc-seq signals distribute within and around genes of various length categories. TOP2Bcc-seq signals within gene bodies were elevated in short genes compared to long genes (Figure 6B), suggesting that the levels of TOP2Bccs within long genes are inconsistent with greater torsional stress within their gene bodies compared to short genes. Next, we compared fastGRO signals from ETP-treated neurons and controls within the first 35 kb in genes longer than 40 kb. The estimated elongation rate of RNAPII is ~2 kb/min, indicating that a newly elongating RNAPII could cover a distance of ~60 kb during the duration of ETP treatment (30 min). Based on this and buffering for edge effects in genes with the minimum length of 40 kb, assessments within the first 35 kb allowed us assess how ETP affects new transcription. In genes ranging 40–80 kb, ETP fastGRO signals converged with those from untreated controls at ~ 28 kb from the TSS but did not dip further (Figure 6C). The distance of such intersection decreased to ~22 kb in genes ranging between 80 and 160 kb, and further to ~14 kb in genes longer than 160 kb, and in both cases, ETP fastGRO signals declined below control signals shortly thereafter (Figure 6C). Reductions in ETP fastGRO signals beyond this point were greater for genes longer than 160 kb than genes ranging between 80 and 160 kb (Figure S6C). Meanwhile, aggregate fastGRO signals within the first 10 kb were largely similar between the analyzed gene length categories in both ETP-treated neurons and untreated controls (Figure S6D), suggesting that reductions in RNAPII activity downstream are less likely to emerge from increased transcription-generated supercoiling in longer genes. Thus, while transcription in long genes is more vulnerable to ETP, transcript length itself does not seem to account for why TOP2B activity is needed at these genes.

Based on these results, we explored properties of genes that could cause their transcription to be affected by ETP. Chromatin state analysis revealed an enrichment of distal and weak enhancers within the bodies of genes longer than 80 kb (Figure 6D). Similar enrichment of enhancers was also observed in TOP2B-dependent genes, including in the small subset of ETP-downregulated genes that were shorter than 80 kb (Figure 5E). To understand how the presence of intragenic enhancers affects the transcription of long genes, the locations of intragenic enhancers were identified by intersecting intervals for strong distal (En-Sd), poised distal (En-Pd), and weak (En-W) enhancers with those for genes longer than 80 kb. fastGRO signals were then examined relative to these genic enhancers while accounting for strand orientation. Whereas fastGRO signals in untreated neurons were unaffected by the presence of intragenic enhancers, fast-GRO signals in regions distal to enhancers were reduced compared to proximal regions in ETP-treated neurons (Figure 6E), indicating that intragenic enhancers disrupt transcription specifically in ETP-treated neurons. To further characterize the effects of intragenic enhancers on long gene transcription, genes longer than 80 kb were subdivided into two groups based on whether they either contained or lacked En-Sd, En-Pd, or En-W states within them, and fastGRO signals from untreated and ETP-treated neurons were compared within these gene sets. Long genes containing intragenic enhancers displayed a significant reduction in fastGRO signals following ETP treatment compared to controls than long genes that lacked intragenic enhancers (Figure 6F).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ETP is a TOP2 poison that causes the formation of DNA DSBs in neurons, and ETP-mediated DSBs could interfere with transcription and potentially underlie the transcriptional changes described above. To address a potential role for ETP-mediated DSBs in the observed transcriptional changes, cultured cortical neurons were incubated with a catalytic inhibitor of TOP2B, ICRF-193 (20 μM for 30 min), and the effects of ICRF-193 on nascent transcription were assessed using fastGRO. Data from replicate experiments were validated as before (Figure S4) and fastGRO signals between replicate experiments were highly concordant (Figures S7A and S7B; data not shown). Importantly, analysis of fastGRO signals from ICRF-193-treated neurons largely recapitulated findings from ETP-treated neurons. Specifically, fastGRO signals were elevated in ICRF-193-treated neurons in genes across most expression quartiles, including the most highly expressed Q4 genes (Figures S7C, S8A, and S8B). fastGRO signals in ICRF-193-treated neurons were also elevated at enhancers compared to controls (Figure S8C) but were reduced in ETP-downregulated genes (Figure S8D). Like in ETP-treated neurons, mean fastGRO signals in ICRF-193-treated neurons were elevated across gene length categories, except for genes longer than 80 kb, which showed the opposite trend (Figures S8E–S8G). Furthermore, long genes containing intragenic enhancers displayed a significant reduction in fast-GRO signals following ICRF-193 treatment compared to controls than long genes that lacked intragenic enhancers (Figure S8H). These results suggest that reductions in TOP2B activity, and not DSBs, downregulate transcription of long genes in ETP-treated and ICRF-193-treated neurons. Because both ETP and ICRF-193 stimulate transcription at enhancers, we propose that cryptic transcription from internal enhancers under these conditions could interfere with genic transcription and that the enrichment of intragenic enhancers within long genes could explain their reliance on TOP2B.

DISCUSSION

In this study, we attempted to decipher the cellular roles of TOP2B by mapping genome-wide sites of catalytically engaged TOP2B in postmitotic neurons. We show that catalytic TOP2B engagement varies according to the presence of specific nucleosome conformations within the bodies of actively transcribed genes. The apparent attunement of TOP2Bccs to the level of transcription is lost in genes that lack Pr-F-, Tr-I-, and Tr-S-related chromatin marks, suggesting that these chromatin states, and not transcription-generated torsional stress, could affect the catalytic engagement of TOP2B within gene bodies. Acting atop these chromatin states, the higher-order organization of chromosomes into compartments sets the threshold of TOP2B activity within chromatin like a rheostat. In contrast to compartments, we were unable to detect differences in TOP2Bccs at most loop anchors and TAD boundaries, indicating that these areas either do not accrue significant torsional stress or that torsional stress at these sites is not resolved by TOP2B in neurons. The minority of loop anchors and TAD boundaries that did show elevated TOP2Bccs compared to surrounding regions were enriched for Hc-H, and the level of TOP2Bccs at these sites matched those observed around Hc-H sites genome-wide. How the catalytic engagement of TOP2B is stimulated at specific chromatin states is unclear. It was reported that TOP2A is recruited to H3K9me3-rich chromatin by the multidomain reader UHRF1, which recognizes H3K9me3 through its Tudor and PHD domains.[60] Similarly, UHRF1 and other readers of histone posttranslational

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

modifications could engage TOP2B at Hc-H and other chromatin states in neurons. Previous studies also indicate that inhibition of histone deacetylases causes the redistribution of TOP2B from heterochromatin to euchromatin.[61] It would be intriguing to test how dynamic changes in histone acetylation and other chromatin modifications also affect the catalytic engagement of TOP2B genome-wide.

Whereas TOP2Bccs are enriched within the bodies of strongly transcribed genes, they are depleted at active promoters. Similar activity profiles were reported for TOP1 in human colon cancer cells,[33] suggesting that TOP1 and TOP2B activity are similarly organized at transcribed genes. Reduced topoisomerase activity could enhance negative supercoiling at active promoters and facilitate DNA melting and the loading of RNAPII. TOP2Bcc-seq levels negatively correlate with the levels of RNAPII at the TSS, indicating that components of the RNAPII complex could curtail its activity.

Classical models predict that topoisomerase activity within gene bodies is important for resolving supercoiling that would otherwise stall RNAPII.[62] Surprisingly however, nascent transcription was elevated, and not diminished, at many genes after ETP treatment. Either inhibition or poisoning of TOP2B resulted in increased transcriptional initiation and elongation, and we see no evidence of increased RNAPII pausing under these conditions. Because TOP2B was only inhibited/poisoned for short periods of time (30 min) in our experiments, it is unclear whether the resultant increase in transcriptionally engaged RNAPII is sustainable or whether increased supercoiling in ETP-/ICRF-193-treated neurons would eventually stall RNAPII. However, our results are largely consistent with those from RNA-seq experiments, which show that transcription at relatively few genes is reduced even after prolonged inhibition/knockdown of topoisomerase activity in mammalian cells.[14,15,57] These results call for a reevaluation of the relationship between transcription-generated supercoiling, RNAPII stalling, and topoisomerase activity. While excess supercoiling was shown to arrest RNAPII *in vitro* and in bacteria,[63,64] it is unclear whether this scenario applies to eukaryotic cells with long linear chromosomes, where supercoiling could transmit and equilibrate over long distances and reduce the need for topoisomerases to support transcription. Moreover, positive supercoiling was shown to displace nucleosomes ahead of RNAPII and aid transcriptional elongation.[35] We propose that reduced catalytic engagement of TOP2B could cause unresolved negative supercoiling generated by RNAPII to propagate to the TSS and stimulate the initiation of new transcription. In this regard, our observations that nascent transcription at highly transcribed genes is stimulated by ETP and ICRF-193 are also consistent with previous observations that TOP2 activity is particularly essential for resolving supercoiling at the most highly expressed genes.[35,42] Direct assessments of DNA supercoiling and RNAPII activity following prolonged topoisomerase inhibition should elaborate whether increased nascent transcription in ETP-treated neurons can be sustained over long periods.

Analysis of nascent transcription in ETP-treated neurons also clarifies that RNAPII stalling from torsional stress is unlikely to explain why topoisomerases are selectively required for the transcription of long neuronal genes.[15,57] Nascent transcription in ETP-treated neurons declines below controls before the first 40 kb in genes longer than 80 kb, while transcription in genes many ranging between 40 and 80 kb in length remains unaffected. Moreover, both

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RNAPII and TOP2B activity is generally higher in genes ranging between 40 and 80 kb than in genes longer than 80 kb. These observations suggest that the characteristics of genes themselves, and not their length alone, mediate their reliance on topoisomerases. Similarly, our observations that ICRF-193 largely recapitulates the transcriptional effects of ETP on long genes suggest that the downregulation of long genes is unlikely to be caused by the accumulation of DSBs within long genes in ETP-treated neurons.

We show here that in addition to stimulating new rounds of transcriptional initiation at genes, ETP and ICRF-193 also stimulate transcription at TREs, including enhancers, and that genic enhancers are enriched within genes longer than 80 kb. Promiscuous transcription at these intragenic enhancers following reductions in TOP2B activity could lead to collisions with RNAPII involved in genic transcription, leading to their downregulation.[65] Interestingly, both H3K4me2 and H3K36me3, which are enriched in chromatin states with high TOP2B activity (Pr-F, Tr-I, and Tr-S), also suppress cryptic intragenic transcription in yeast and mammalian cells.[66–71] These chromatin states and TOP2B activity are both depleted in ETP-downregulated genes, indicating that these genes could be particularly susceptible to cryptic transcription initiation. Whereas the role of cryptic transcription in suppressing long gene transcription was not directly tested in this study, nascent transcription declined in ETP-treated neurons (but not controls) at positions that follow genic enhancers, and the length-related effects of ETP on transcription correlated with the proportion of genes that contained intragenic enhancers in each length category. Overall, these results reveal how increased nascent transcription in ETP-/ICR-193-treated neurons could also explain why a subset of long neuronal genes rely on TOP2B for their transcription.

Limitations of the study

One limitation of this study is that although TOP2B catalytic engagement correlated with the occupancy of several combinatorial histone marks, a direct role for these marks in regulating TOP2Bccs was not directly demonstrated. For instance, readers of specific histone marks could either recruit or stimulate TOP2B catalytic activity within chromatin. Future studies that assess TOP2Bccs under conditions that perturb the distribution of histone marks associated with Pr-F, Tr-I, Tr-S, and Hc-H and assessments of interactions between readers of such marks and TOP2B should clarify this issue. While our data suggest that RNAPII is unlikely to be stalled by torsional stress from excess supercoiling in most long genes, supercoiling levels were not directly assessed in this work. Such efforts are the subject of a separate ongoing study and should further clarify the relationship between topoisomerases and transcription in long genes.

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ram Madabhushi. (ram.madabhushi@utsouthwestern.edu).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Materials availability—This study did not generate new unique reagents.

Data and code availability—All sequencing data generated in this study has been deposited at GEO (GSE227463) and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. The existing, publicly available data that was analyzed for this study are listed in the key resources table along with accession numbers. This paper does not report any original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Mouse primary cortical neurons—Dissociated primary cortical neurons from E16 Swiss-Webster mice (Charles River) were plated at a density of 5 million cells/plate in 10 cm plates. The plates were pre-coated with poly-D-lysine (0.5 mg/mL) overnight at 37°C, followed by washing twice with dH_2O. Neurons were maintained in neurobasal media (GIBCO) supplemented with L-glutamine, penicillin/streptomycin, and B27. Etoposide (ETP, 50 μM, Sigma E1383) and ICRF-193 (20 μM, Sigma I4659) were dissolved in DMSO. Cells were treated by adding drug directly to the culture media 30 min prior to collection.

S2 Drosophila cell line—Drosophila Schneider's Line 2 Cells (S2 Cells) were cultured at 25°C and 5% CO_2 in Schneider media supplemented with 10% heat inactivated fetal bovine serum (GIBCO). S2 cells were incubated for 5 min with 50 mM 4-thiouridine prior to collection for spike in sample.

METHOD DETAILS

TOP2Bcc-seq—Following ETP treatment, covalently bound TOP2B was isolated via the *in vivo* complex of enzyme (ICE) assay as described previously.[34] Two biological replicates of DIV13 primary cortical neurons (~60 million cells per replicate) were lysed in 1% sarkosyl (750 μL per 10 cm plate) and sheared via 10 passes through a 1 mL syringe with a 25-gauge needle. Sheared lysates (1.5 mL) were layered atop 150% CsCl (2 mL) in 13 × 51 mm polycarbonate ultracentrifuge tubes (Beckman) and centrifuged at 71,000 RPM in an SLA-100.3 rotor (Beckman) for 20 h at room temperature. Pellets containing DNA and covalently bound proteins were washed with 70% ethanol and resuspended in 300 μL SDS lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS, protease and phosphatase inhibitor cocktail). Following resuspension, samples were sonicated to shear DNA to uniform lengths between 200 and 500 bp using a bioruptor (Diagenode; high power, 30 s on/30 s off for 40 cycles). Samples were then centrifuged at 13,000 RPM for 10 min at 4°C to remove cellular debris, and the supernatant was diluted up to 2 mL in ChIP dilution buffer (16.7 mM Tris-HCl pH 8, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS).

TOP2Bccs were isolated from sonicated lysates containing covalently bound proteins via native TOP2B ChIP. Prior to immunoprecipitation, 1% of each sample was transferred to a clean tube to be used for input normalization. Samples were then incubated with anti-TOP2B antibody (1:50, Santa Cruz (H-8), RRID: AB_628384) overnight at 4°C followed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

by incubation with ChIP-grade protein A/G magnetic beads (Thermo) for 3 h at 4°C. Antibody-conjugated beads were sequentially washed with low salt buffer (20 mM Tris-HCl pH 8, 1% Triton X-100, 2 mM EDTA, 0.1% SDS, 150 mM NaCl), high salt buffer (20 mM Tris-HCl pH 8, 1% Triton X-100, 2 mM EDTA, 0.1% SDS, 500 mM NaCl), LiCl buffer (10 mM Tris-HCl pH 8, 1% IGEPAL-CA630, 1% deoxycholic acid (sodium salt), 1 mM EDTA, 0.25 M LiCl), and twice with TE buffer (50 mM Tris-HCl pH 8, 1 mM EDTA) for 5 min each with rotation at room temperature. Immunoprecipitated TOP2B-DNA complexes were eluted from the magnetic beads in 500 μL ChIP elution buffer (1% SDS, 0.1 M $NaHCO_3$) and purified along with input DNA via phenol-chloroform extraction and ethanol precipitation. Precipitated DNA was washed with 70% ethanol and resuspended in TE buffer.

To prepare libraries for sequencing, immunoprecipitated DNA was first end-repaired, A-tailed, and ligated with paired-end, Illumina barcoded adapters (KAPA biosciences) according to the manufacturer's instructions. Attempts to amplify libraries for sequencing after this step alone proved unsuccessful, implying the successful isolation of TOP2Bccs. Following the first round of adapter ligation, the 5′ strand was digested with λ exonuclease (NEB). The resulting 3′ strand was used as a temple for 10 rounds of primer extension (cycling conditions: 95°C for 3 min, ten rounds of linear amplification (95°C for 1 min, 57°C for 1 min and 72°C for 3 min), 95°C for 1 min, hold at 4°C) using primers against the ligated adapter in a solution containing 1X HiFi Phusion buffer (NEB), 0.8 mM dNTPs, 2 U Phusion DNA polymerase (NEB) and 40 nM of P7 extension primer (CAAGCAGAAGACGG CATACGA*G; *denotes phosphorothioate linkage). The resulting single-stranded products were purified using 0.8X Ampure XP beads (Beckman), eluted in 35 μL TE dH_2O, and concentrated to a volume of 18 μL using a vacuum centrifuge (Eppendorf). The purified products were appended with ribo-G in 1X terminal deoxynucleotidyl transferase buffer (NEB), 20 U terminal deoxynucleotidyl transferase enzyme (NEB), and 4 mM of rGTP at 37°C for 15 min followed by 0.8X Ampure XP bead clean-up and elution as before. The single-stranded, ribo-tailed products were ligated to double-stranded adapters with CCC overhang (oligo 1, AATGATACGGCG ACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC.

TCCC; oligo 2, (Phosphate)GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG) as described previously.[35] Ligated products were used as templates for standard paired-end library amplification (KAPA biosciences) with an additional extension step prior to linear amplification to create double-stranded DNA templates (cycling conditions: 60°C for 3 min, 98°C for 30 s, eleven cycles of amplification (98°C for 10 s, 65°C for 30 s, and 60°C for 30 s), 60°C for 5 min and hold at 4°C). The final PCR products were purified using 0.8X Ampure XP beads and eluted in 40 μL EB (Qiagen). DNA fragment size was assessed via Tapestation (Agilent) and additional Ampure XP bead clean-up steps were performed to obtain a final library size of 350–450 bp. DNA concentration was quantified by fluorometry (dsDNA high sensitivity Qubit assay, Thermo) to ensure equal DNA concentration for each replicate before replicates were pooled and sequenced on an Illumina NextSeq 500 instrument using V2.5 reagents to a depth of 33.3 million reads (replicate 1) and 35.6 million reads (replicate 2).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

TOP2Bcc-seq analysis—Raw FASTQ sequencing reads were analyzed for quality control using FastQC,[72] trimmed using Trimmomatic,[73] and aligned to the mm10 reference genome (GRCm38.p6 assembly) using BWA-MEM.[74] BAM files were marked and filtered for duplicate reads using Sambamba[75] and Samtools,[76] respectively. Peaks were called using MACS2[36] from biological duplicates relative to input DNA. MACS2-generated bedGraph files were converted to bigWig files using UCSC Kent tools[77] for individual replicates and pooled replicates. The TOP2B ChIP-seq data was from cultured mouse cortical neurons and as previously described.[24] The data were reanalyzed and aligned to mm10 using BWA.[74] Peaks were called using MACS2 with $q < 0.05$, pseudoreplicates were generated and replicated peaks were called from those and the two biological replicates (raw data is available at GEO accession GSE61887).

CTCF ChIP-seq—CTCF ChIP was performed for two biological replicates using a ChIP assay kit (MilliporeSigma) as described by the manufacturer. Briefly, DIV13 primary cortical neurons (~15M cells per replicate) were crosslinked with 1% formaldehyde for 10 min at 37°C, washed twice with ice-cold PBS, scraped into 1.5 mL tubes, and resuspended in 300 mL SDS lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS, protease and phosphatase inhibitor cocktail). Cells were sonicated to shear DNA to 200–500 bp using a bioruptor (Diagenode; high power, 30 s on/30 s off for 40 cycles). Samples were then centrifuged at 13,000 RPM for 10 min at 4°C to remove cellular debris, and the supernatant was diluted up to 2 mL in ChIP dilution buffer (16.7 mM Tris-HCl pH 8, 167 mM NaCl, 1.2 mM EDTA, 1.1% Triton X-100, 0.01% SDS). Prior to immunoprecipitation, 1% of each sample was transferred to a clean tube to be used for input normalization. Samples were then incubated with anti-CTCF antibody (1:200, CST, #3418) overnight at 4°C followed by incubation with ChIP-grade protein A/G magnetic beads (Thermo) for 3 h at 4°C. Antibody-conjugated beads were washed with low salt buffer (20 mM Tris-HCl pH 8, 1% Triton X-100, 2 mM EDTA, 0.1% SDS, 150 mM NaCl), high salt buffer (20 mM Tris-HCl pH 8, 1% Triton X-100, 2 mM EDTA, 0.1% SDS, 500 mM NaCl), LiCl buffer (10 mM Tris-HCl pH 8, 1% IGEPAL-CA630, 1% deoxycholic acid (sodium salt), 1 mM EDTA, 0.25 M LiCl), and twice with TE buffer (50 mM Tris-HCl pH 8, 1 mM EDTA) for 5 min each with agitation at room temperature. Immunoprecipitated CTCF-DNA complexes were eluted from the magnetic beads in 500 μL ChIP elution buffer (1% SDS, 0.1 M $NaHCO_3$). DNA-protein crosslinks for immunoprecipitated samples and input DNA were reversed via incubation with an additional 20 mM NaCl for 4 h at 37°C and residual protein was digested with Proteinase K (NEB) for 1 h at 45°C. DNA was precipitated using phenol/chloroform isolation and ethanol precipitation, washed with 70% ethanol, and resuspended in 50 μL TE buffer. Samples were then prepared for sequencing using a paired-end library preparation kit compatible with Illumina sequencing instruments (KAPA biosciences) according to the manufacturer's instructions. Following library amplification, DNA fragment size was assessed via Tapestation (Agilent) to ensure a final library size of 350–450 bp. DNA concentration was quantified by fluorometry (dsDNA high sensitivity Qubit assay, Thermo) to ensure equal DNA concentration for each replicate before replicates were pooled and sequenced on an Illumina NextSeq 500 instrument using V2.5 reagents to an average depth of 25–35 million reads per replicate.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ChIP-seq analysis—ChIP sequencing reads were processed in accordance with ENCODE analysis guidelines.[89] Read alignment and all subsequent analysis requiring chromosome length files were performed using GRCm38.p6 assembly (mm10) reference genome. Peaks were called using MACS2 from biological duplicates and pseudo-replicates where each ChIP sample was normalized to input. IDR (irreproducible discovery rate[90]) thresholding was performed on MACS2 peak calls using an IDR threshold of 0.05.

RNA-seq—Total RNA was extracted from two biological replicates of DIV13 primary cortical neurons (~5M cells per replicate) using the Qiagen RNeasy Plus Universal Kit (Qiagen) according to the manufacturer's instructions. RNA quality was assessed using an Agilent Tapestation 4200, and only RNA with RIN score of 8 or higher was used for library preparation. Prior to library preparation, samples were treated with DNase to remove any residual DNA, and RNA concentration was measured by fluorometry (RNA high sensitivity Qubit assay, Thermo). RNA libraries were prepared for sequencing using a TruSeq stranded mRNA library prep kit with single-end adapters (Illumina) according to the manufacturer's instructions. Prior to sequencing, library quality was assessed using an Agilent Tapestation 4200, and sample concentration was measured by fluorometry (dsDNA high sensitivity Qubit assay, Thermo). Samples were pooled and sequenced on an Illumina NextSeq 500 instrument with V2.5 reagents to an average depth of 35–45 million reads per replicate.

RNA-seq analysis—Raw FASTQ sequencing reads were trimmed using Trimmomatic[73] and aligned to the mm10 reference genome (GRCm38.p6 assembly) using TopHat.[78] The aligned reads were processed by the HTSeq-Count pipeline and relative transcript abundance was measured by transcripts per million mapped fragments (TPM).

fastGRO—The recently developed fastGRO protocol[52] was adapted to perform GRO-seq in cultured primary cortical neurons. DIV13 primary cortical neurons for two biological replicates (~20 million cells per replicate) were washed twice with ice-cold PBS and collected in a 15 mL conical tube. Cells were incubated in 10 mL swelling buffer (10 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$, 2 U/ml Superase-In (Invitrogen)) for 5 min on ice, washed with swelling buffer +10% glycerol, and resuspended in 10 mL lysis buffer (10 mM Tris-HCl pH 7.5, 2 mM $MgCl_2$, 3 mM $CaCl_2$, 10% glycerol, 1% IGEPAL, 2 U/ml Superase-In) to isolate nuclei. Isolated nuclei were washed twice with lysis buffer and resuspended in 100 μL freezing buffer (50 mM Tris-HCl pH 8.3, 5 mM $MgCl_2$, 0.1 mM EDTA, 40% glycerol). An equal volume of pre-warmed, 2X nuclear run-on reaction buffer (10 mM Tris-HCl pH 8, 5 mM $MgCl_2$, 300 mM KCl, 1 mM DTT, 0.5 mM ATP, 0.5 mM GTP, 0.5 mM 4-thio-UTP (Trilink), 2 μM CTP, 200 U/ml Superase-in, 1% sarkosyl) was added, and the reaction was incubated for 7 min at 30°C. Nuclear run-on (NRO) RNA was extracted with TRIzol LS reagent (Invitrogen) following the manufacturer's instructions and precipitated with ethanol. Precipitated NRO-RNA was resuspended in 100 μL dH_2O, and concentration was determined using a Qubit 4.0 fluorimeter (Thermo). RNA (120–150 μg) was transferred to a new tube and 5%–10% spike-in RNA derived from S2 cells was added (ETP fastGRO seq) or omitted (ICRF-193 fastGRO seq). RNA was then fragmented using a bioruptor (Diagenode; high power, 30 s on/30 s off for 3 cycles) Fragmentation efficiency was analyzed by running fragmented and unfragmented RNA on an Agilent 4200

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

TapeStation instrument using High Sensitivity RNA ScreenTape following manufacturer's instructions. Fragmented RNA was incubated in biotinylation solution (20 mM Tris pH 7.5, 2 mM EDTA, 40% dimethylformamide, 200 μg/mL EZ-link PDP Biotin (Thermo)) for 2 h at 25°C with rotation (800 rpm) in a dark room then purified via ethanol precipitation and resuspended in dH_2O. M280 streptavidin Dynabeads (Invitrogen) were washed twice with 2 volumes of freshly prepared wash buffer (100 mM Tris pH 7.5, 10 mM EDTA, 1M NaCl, 0.1% Tween 20) and resuspended in 1 volume of wash buffer. Thereafter, biotinylated RNA was affinity purified via incubation with 100 μL washed Dynabeads for 15 min at 4°C. After binding, Dynabeads were washed three times with wash buffer pre-warmed at 65°C, and three times with wash buffer at room temperature. RNA was eluted with 100 mM DTT buffer and purified with RNA Clean and Purification kit (Zymo Research) with in-column DNase reaction to eliminate traces of genomic DNA. The eluted RNA was quantified with Qubit High Sensitivity Assay kit (Invitrogen) and used to produce barcoded RNA sequencing libraries using the Kapa RNA HyperPrep kit (KAPA biosystems) according to the manufacturer's instructions. Libraries were sequenced on an Illumina NextSeq 500 instrument using V2.5 reagents.

fastGRO analysis—Raw fastGRO sequencing reads were processed as described previously.[52] Briefly, reads were trimmed and poly-A sequences were removed using Cutadapt[79] then aligned to the mm10 reference genome using BWA-ALN[74] and filtered/sorted using Samtools[76] (q = 10). Sorted read files were converted to BAM files and normalized by counts per million (CPM) to create bigWig coverage files using deepTools *BamCoverage*. BAM files were also used to generate tag directories using HOMER.[81] Tag directories were used to call transcripts with HOMER *findpeaks*[82] (options: -style groseq -tssFold 6 -bodyFold 5 -pseudoCount 0.5 -minBodySize 500 -maxBodySize 100000) and corresponding FPKM values for known genes using HOMER *analyzeRepeats.pl*[82] in RNA mode (analyzeRepeats.pl rna mm10 -count genes -strand both -rpkm -condenseGenes). HOMER *findpeaks*[18] derived raw counts were analyzed for differential expression using DEseq2[56] with $q < 0.05$ and cut-offs of Log2FC < −0.6 for downregulated transcripts and >0.6 for upregulated transcripts. Discriminative regulatory elements (dREGs) were identified using the dREG gateway.[55,84,83] Stranded bigWig signal files were generated from fastGRO reads using a dedicated fastGRO/Pro-seq pipeline (*RunOnBamToBigWig*[84]) Unnormalized, strand separated bigWigs were uploaded to the dREG[83] gateway (https://dreg.dnasequence.org) and analyzed using default parameters. Resulting dREG probability bigWigs were used for track visualization, and dREG called peak files were used for plotting TOP2Bcc-seq signal surrounding identified regulatory elements. Where indicated, fastGRO FPKM values were used to group transcripts into quartiles of expression (Q1 = lowest expression, Q4 = highest expression).

Hi-C analysis—Previously published Hi-C data from mouse cortical neurons were used to identify chromosome compartments, topologically associating domains (TADs), and chromatin loops using the Juicer suite[85] Chromosome compartments were called using the *eigenvector* command with a 500 kb resolution and KR normalization. Eigenvalue bins were converted to bedGraphs using the *cut* command on the Linux command line and then to bigWig UCSC Kent tools *bedGraphToBigWig*[77] to be used for visualization. Chromatin

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

loops were called using HiCCUPS[47] with default parameters, where resulting 5 kb loop anchors were used for downstream analysis. TADs were called using Arrowhead from the Juicer suite (5 kb bins, KR normalized), and corner scores were used for approximating TAD strength.

Motif analysis—Peak files containing peak summit information were generated by MACS2. Peak summits defined by MACS2 were extended +/−250bp and converted to mm10 (GRCm38.p6 assembly) FASTA sequence using bedtools *getfasta*.[87] FASTA sequences were analyzed via MEME-ChIP[86] using default parameters.

Chromatin state annotation—ChromHMM[40] was used to segment the genome into 15 chromatin states based on ENCODE ChIP-seq data from P0 mouse forebrain for eight histone modifications (H3K27ac, H3K9ac, H3K4me3, H3K4me1, H4K4me1, H3K27me3, H3K9me3, H3K36me3). Annotation of chromatin states was as described previously.[40] ChromHMM *OverlapEnrichment* was used to determine enrichment of each chromatin state within intervals of interest using the 15 chromatin state model. ChromHMM *NeighborhoodEnrichment* was used to visualize chromatin state enrichment as a function of distance from the TSS. BED files corresponding to each chromatin state were also used to examine TOP2Bcc-seq or TOP2B ChIP-seq signal surrounding each chromatin state. Where indicated, intervals corresponding to each chromatin state were intersected with chromosome compartments or fastGRO transcripts using bedtools *intersect* (-wa) to separate each compartment into 15 chromatin states. To identify transcripts with or without distal intergenic enhancers, fastGRO transcripts were intersected with strong TSS distal enhancer (En-Sd), poised TSS distal enhancer (En-Pd), and weak enhancer (En-W) chromatin states annotated by ChromHMM using bedtools *intersect* (-wa for overlapping, -v for non-overlapping). In addition to ChromHMM analysis, chromatin state enrichment of TOP2B ChIP-seq and TOP2Bcc peak files generated by MACS2 were analyzed using ChIPseeker *annotatePeak*[88] using the UCSC mm10 genome annotation database and default parameters.

Data visualization—To assess the agreement between TOP2Bcc-seq replicates or TOP2Bcc-seq and TOP2B ChIP-seq signals, correlations between bigWig files were computed using deepTools *multiBigwigSummary* and plotted using *plotCorrelation*[80] with default parameters. All aggregate plots and heatmaps throughout the manuscript were generated from score matrices using deepTools *computeMatrix* in either reference-point or scale-regions mode.[80] For aggregate plots of TOP2Bcc-seq or TOP2B ChIP-seq signal surrounding annotated chromatin states and loop anchors, chromatin state or loop anchor intervals were scaled to 200 bp and TOP2Bcc-seq signal was averaged over 200 bp bins with 3 kb flanking regions. For aggregate plots of TOP2Bcc-seq signal surrounding TADs, TAD intervals were scaled to 10 kb and TOP2Bcc-seq signal was averaged over 200 bp bins with 50 kb flanking regions. For aggregate plots of TOP2Bcc-seq signal in each chromosome compartment, compartment intervals were scaled to 10 kb and TOP2Bcc-seq signal was averaged over 200 bp bins. For TOP2Bcc-seq signal at TADs and loop anchors, the k-means clustering option in *plotProfile* was used to separate TOP2Bcc-seq signal into two distinct clusters. Further increase in cluster number did not result in identification

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of additional clusters that significantly differed from clusters in the two-state model. For aggregate plots of TOP2Bcc-seq signal surrounding annotated chromatin states within each chromosome compartment, chromatin state intervals were scaled to 1 kb and TOP2Bcc-seq signal was averaged over 200 bp bins with 5 kb flanking regions. For aggregate plots of TOP2Bcc-seq signal at gene expression quartiles or gene length categories, gene intervals were scaled to 1 kb and TOP2Bcc-seq signal was averaged over 200 bp bins with 600 bp flanking regions. For aggregate plots of fastGRO signal at gene expression quartiles or gene length categories, gene intervals were scaled to 1 kb and fastGRO signal was averaged over 50 bp bins with 300 bp flanking regions as indicated on the x axes with the exception of genes longer than 80 kb, which were scaled to 30 kb. Aggregate plots of fastGRO signal surrounding intergenic enhancers were averaged over 50bp bins in reference-point mode using the center of the enhancer interval and 5 or 10 kb flanking regions. All other aggregate plots were averaged over 200 bp bins in reference-point mode using the center of the region of interest and flanking regions as indicated on the x axes with the exception of TOP2Bcc-seq signal surrounding CTCF peaks, which were averaged over 50 bp bins. Matrices from *computeMatrix* outputs were plotted using the *plotProfile* or *plotHeatmap* commands for visualization. Median signals for each bin were calculated using kentUtils *bigWigAverageOverBed*[77] for significance testing. TOP2Bcc-seq, chromosome compartment, and dREG signal tracks from bigWig files were visualized using the UCSC genome browser.[91]

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were generated from independent biological replicates of cultured cortical neurons isolated from independent mouse litters, as described in the method details. TOP2Bcc-seq and all fastGRO seq experiments consisted of two biological replicates while the RNA-seq consisted of three biological replicates. Differences in signal intensities between two groups was done using Mann-Whitney-U or between multiple groups using One-way ANOVA with Tukey's post-hoc tests for pairwise comparisons. All TOP2Bcc-seq signal intensity plots show the mean signal between the two reps and averaged over the indicated states/locations. All fastGRO-seq signal plots show signal of a single replicate averaged over the indicated state/locations. Bar graphs depicting the quantification of signal intensity were generated and tested for significance using GraphPad Prism 9 and show median signal with 95% confidence intervals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was supported by CPRIT recruitment award RR170010, Welch Foundation grant I-1960-201803324, and NIMH R01-MH120132 to R.M.; NIH T32GM109776 and T32HL139438 to L.H.; and CPRIT training grant RP210041 to M.C.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REFERENCES

1. Pommier Y, Nussenzweig A, Takeda S, and Austin C. (2022). Human topoisomerases and their roles in genome stability and organization. Nat. Rev. Mol. Cell Biol. 23, 407–427. 10.1038/s41580-022-00452-3. [PubMed: 35228717]
2. Pommier Y, Sun Y, Huang SYN, and Nitiss JL (2016). Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. Nat. Rev. Mol. Cell Biol. 17, 703–721. 10.1038/nrm.2016.111. [PubMed: 27649880]
3. Wang JC (2002). Cellular roles of DNA topoisomerases: a molecular perspective. Nat. Rev. Mol. Cell Biol. 3, 430–440. 10.1038/nrm831. [PubMed: 12042765]
4. Vos SM, Tretter EM, Schmidt BH, and Berger JM (2011). All tangled up: how cells direct, manage and exploit topoisomerase function. Nat. Rev. Mol. Cell Biol. 12, 827–841. 10.1038/nrm3228. [PubMed: 22108601]
5. Austin CA, Lee KC, Swan RL, Khazeem MM, Manville CM, Cridland P, Treumann A, Porter A, Morris NJ, and Cowell IG (2018). TOP2B: The First Thirty Years. Int. J. Mol. Sci. 19, 2765. 10.3390/ijms19092765. [PubMed: 30223465]
6. Capranico G, Tinelli S, Austin CA, Fisher ML, and Zunino F. (1992). Different patterns of gene expression of topoisomerase II isoforms in differentiated tissues during murine development. Biochim. Biophys. Acta 1132, 43–48. 10.1016/0167-4781(92)90050-a. [PubMed: 1380833]
7. Juenke JM, and Holden JA (1993). The distribution of DNA topoisomerase II isoforms in differentiated adult mouse tissues. Biochim. Biophys. Acta 1216, 191–196. 10.1016/0167-4781(93)90144-3. [PubMed: 8241259]
8. Woessner RD, Mattern MR, Mirabelli CK, Johnson RK, and Drake FH (1991). Proliferation- and cell cycle-dependent differences in expression of the 170 kilodalton and 180 kilodalton forms of topoisomerase II in NIH-3T3 cells. Cell Growth Differ. 2, 209–214. [PubMed: 1651102]
9. Nielsen CF, Zhang T, Barisic M, Kalitsis P, and Hudson DF (2020). Topoisomerase IIalpha is essential for maintenance of mitotic chromosome structure. Proc. Natl. Acad. Sci. USA 117, 12131–12142. 10.1073/pnas.2001760117. [PubMed: 32414923]
10. Grue P, Grässer A, Sehested M, Jensen PB, Uhse A, Straub T, Ness W, and Boege F. (1998). Essential mitotic functions of DNA topoisomerase IIalpha are not adopted by topoisomerase IIbeta in human H69 cells. J. Biol. Chem. 273, 33660–33666. 10.1074/jbc.273.50.33660. [PubMed: 9837951]
11. Lyu YL, and Wang JC (2003). Aberrant lamination in the cerebral cortex of mouse embryos lacking DNA topoisomerase IIbeta. Proc. Natl. Acad. Sci. USA 100, 7123–7128. 10.1073/pnas.1232376100. [PubMed: 12773624]
12. Yang X, Li W, Prescott ED, Burden SJ, and Wang JC (2000). DNA topoisomerase IIbeta and neural development. Science 287, 131–134. 10.1126/science.287.5450.131. [PubMed: 10615047]
13. Lyu YL, Lin CP, Azarova AM, Cai L, Wang JC, and Liu LF (2006). Role of topoisomerase IIbeta in the expression of developmentally regulated genes. Mol. Cell Biol. 26, 7929–7941. 10.1128/MCB.00617-06. [PubMed: 16923961]
14. Tiwari VK, Burger L, Nikoletopoulou V, Deogracias R, Thakurela S, Wirbelauer C, Kaut J, Terranova R, Hoerner L, Mielke C, et al. (2012). Target genes of Topoisomerase IIbeta regulate neuronal survival and are defined by their chromatin state. Proc. Natl. Acad. Sci. USA 109, E934–E943. 10.1073/pnas.1119798109. [PubMed: 22474351]
15. King IF, Yandava CN, Mabb AM, Hsiao JS, Huang HS, Pearson BL, Calabrese JM, Starmer J, Parker JS, Magnuson T, et al. (2013). Topoisomerases facilitate transcription of long genes linked to autism. Nature 501, 58–62. 10.1038/nature12504. [PubMed: 23995680]
16. Austin CA, Cowell IG, Khazeem MM, Lok D, and Ng HT (2021). TOP2B's contributions to transcription. Biochem. Soc. Trans. 49, 2483–2493. 10.1042/BST20200454. [PubMed: 34747992]
17. Ju BG, Lunyak VV, Perissi V, Garcia-Bassets I, Rose DW, Glass CK, and Rosenfeld MG (2006). A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. Science 312, 1798–1802. 10.1126/science.1127196. [PubMed: 16794079]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

18. Bunch H, Lawney BP, Lin YF, Asaithamby A, Murshid A, Wang YE, Chen BPC, and Calderwood SK (2015). Transcriptional elongation requires DNA break-induced signalling. Nat. Commun. 6, 10191. 10.1038/ncomms10191. [PubMed: 26671524]
19. Trotter KW, King HA, and Archer TK (2015). Glucocorticoid Receptor Transcriptional Activation via the BRG1-Dependent Recruitment of TOP2-beta and Ku70/86. Mol. Cell Biol. 35, 2799–2817. 10.1128/MCB.00230-15. [PubMed: 26055322]
20. Wong RHF, Chang I, Hudak CSS, Hyun S, Kwan HY, and Sul HS (2009). A role of DNA-PK for the metabolic gene regulation in response to insulin. Cell 136, 1056–1072. 10.1016/j.cell.2008.12.040. [PubMed: 19303849]
21. Lutz H, Nguyen TA, Joswig J, Rau K, and Laube B. (2019). NMDA Receptor Signaling Mediates cFos Expression via Top2beta-Induced DSBs in Glioblastoma Cells. Cancers 11, 306. 10.3390/cancers11030306. [PubMed: 30841565]
22. Stott RT, Kritsky O, and Tsai LH (2021). Profiling DNA break sites and transcriptional changes in response to contextual fear learning. PLoS One 16, e0249691. 10.1371/journal.pone.0249691.
23. Delint-Ramirez I, Konada L, Heady L, Rueda R, Jacome ASV, Marlin E, Marchioni C, Segev A, Kritskiy O, Yamakawa S, et al. (2022). Calcineurin dephosphorylates topoisomerase IIbeta and regulates the formation of neuronal-activity-induced DNA breaks. Mol. Cell 82, 3794–3809.e3798. 10.1016/j.molcel.2022.09.012. [PubMed: 36206766]
24. Madabhushi R, Gao F, Pfenning AR, Pan L, Yamakawa S, Seo J, Rueda R, Phan TX, Yamakawa H, Pao PC, et al. (2015). Activity-Induced DNA Breaks Govern the Expression of Neuronal Early-Response Genes. Cell 161, 1592–1605. 10.1016/j.cell.2015.05.032. [PubMed: 26052046]
25. Canela A, Maman Y, Huang SYN, Wutz G, Tang W, Zagnoli-Vieira G, Callen E, Wong N, Day A, Peters JM, et al. (2019). Topoisomerase II-Induced Chromosome Breakage and Translocation Is Determined by Chromosome Architecture and Transcriptional Activity. Mol. Cell 75, 252–266.e8. 10.1016/j.molcel.2019.04.030. [PubMed: 31202577]
26. Canela A, Maman Y, Jung S, Wong N, Callen E, Day A, Kieffer-Kwon KR, Pekowska A, Zhang H, Rao SSP, et al. (2017). Genome Organization Drives Chromosome Fragility. Cell 170, 507–521.e18. 10.1016/j.cell.2017.06.034. [PubMed: 28735753]
27. Manville CM, Smith K, Sondka Z, Rance H, Cockell S, Cowell IG, Lee KC, Morris NJ, Padget K, Jackson GH, and Austin CA (2015). Genome-wide ChIP-seq analysis of human TOP2B occupancy in MCF7 breast cancer epithelial cells. Biol. Open 4, 1436–1447. 10.1242/bio.014308. [PubMed: 26459242]
28. Uusküla-Reimand L, Hou H, Samavarchi-Tehrani P, Rudan MV, Liang M, Medina-Rivera A, Mohammed H, Schmidt D, Schwalie P, Young EJ, et al. (2016). Topoisomerase II beta interacts with cohesion and CTCF at topological domain borders. Genome Biol. 17, 182. 10.1186/s13059-016-1043-8. [PubMed: 27582050]
29. Gittens WH, Johnson DJ, Allison RM, Cooper TJ, Thomas H, and Neale MJ (2019). A nucleotide resolution map of Top2-linked DNA breaks in the yeast and human genome. Nat. Commun. 10, 4846. 10.1038/s41467-019-12802-5. [PubMed: 31649282]
30. Gothe HJ, Bouwman BAM, Gusmao EG, Piccinno R, Petrosino G, Sayols S, Drechsel O, Minneker V, Josipovic N, Mizi A, et al. (2019). Spatial Chromosome Folding and Active Transcription Drive DNA Fragility and Formation of Oncogenic MLL Translocations. Mol. Cell 75, 267– 283.e12. 10.1016/j.molcel.2019.05.015. [PubMed: 31202576]
31. Davidson IF, and Peters JM (2021). Genome folding through loop extrusion by SMC complexes. Nat. Rev. Mol. Cell Biol. 22, 445–464. 10.1038/s41580-021-00349-7. [PubMed: 33767413]
32. Cowell IG, Sondka Z, Smith K, Lee KC, Manville CM, Sidorczuk-Lesthuruge M, Rance HA, Padget K, Jackson GH, Adachi N, and Austin CA (2012). Model for MLL translocations in therapy-related leukemia involving topoisomerase IIbeta-mediated DNA strand breaks and gene proximity. Proc. Natl. Acad. Sci. USA 109, 8989–8994. 10.1073/pnas.1204406109. [PubMed: 22615413]
33. Baranello L, Wojtowicz D, Cui K, Devaiah BN, Chung HJ, Chan-Salis KY, Guha R, Wilson K, Zhang X, Zhang H, et al. (2016). RNA Polymerase II Regulates Topoisomerase 1 Activity to Favor Efficient Transcription. Cell 165, 357–371. 10.1016/j.cell.2016.02.036. [PubMed: 27058666]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

34. Nitiss JL, Kiianitsa K, Sun Y, Nitiss KC, and Maizels N. (2021). Topoisomerase Assays. Curr. Protoc. 1, e250. 10.1002/cpz1.250. [PubMed: 34606690]
35. Teves SS, and Henikoff S. (2014). Transcription-generated torsional stress destabilizes nucleosomes. Nat. Struct. Mol. Biol. 21, 88–94. 10.1038/nsmb.2723. [PubMed: 24317489]
36. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, and Liu XS (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137. 10.1186/gb-2008-9-9-r137. [PubMed: 18798982]
37. Bailey TL, Johnson J, Grant CE, and Noble WS (2015). The MEME Suite. Nucleic Acids Res. 43, W39–W49. 10.1093/nar/gkv416. [PubMed: 25953851]
38. Sano K, Miyaji-Yamaguchi M, Tsutsui KM, and Tsutsui K. (2008). Topoisomerase IIbeta activates a subset of neuronal genes that are repressed in AT-rich genomic environment. PLoS One 3, e4103. 10.1371/journal.pone.0004103. [PubMed: 19116664]
39. Dobbin MM, Madabhushi R, Pan L, Chen Y, Kim D, Gao J, Ahanonu B, Pao PC, Qiu Y, Zhao Y, and Tsai LH (2013). SIRT1 collaborates with ATM and HDAC1 to maintain genomic stability in neurons. Nat. Neurosci. 16, 1008–1015. 10.1038/nn.3460. [PubMed: 23852118]
40. Ernst J, and Kellis M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. Nat. Protoc. 12, 2478–2492. 10.1038/nprot.2017.124. [PubMed: 29120462]
41. Gorkin DU, Barozzi I, Zhao Y, Zhang Y, Huang H, Lee AY, Li B, Chiou J, Wildberg A, Ding B, et al. (2020). An atlas of dynamic chromatin landscapes in mouse fetal development. Nature 583, 744–751. 10.1038/s41586-020-2093-3. [PubMed: 32728240]
42. Kouzine F, Gupta A, Baranello L, Wojtowicz D, Ben-Aissa K, Liu J, Przytycka TM, and Levens D. (2013). Transcription-dependent dynamic supercoiling is a short-range genomic force. Nat. Struct. Mol. Biol. 20, 396–403. 10.1038/nsmb.2517. [PubMed: 23416947]
43. Thakurela S, Garding A, Jung J, Schübeler D, Burger L, and Tiwari VK (2013). Gene regulation and priming by topoisomerase IIalpha in embryonic stem cells. Nat. Commun. 4, 2478. 10.1038/ncomms3478. [PubMed: 24072229]
44. ENCODE Project Consortium; Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, Kawli T, Davis CA, Dobin A, et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature 583, 699–710. 10.1038/s41586-020-2493-4. [PubMed: 32728249]
45. Dekker J, and Misteli T. (2015). Long-Range Chromatin Interactions. Cold Spring Harbor Perspect. Biol. 7, a019356. 10.1101/cshperspect.a019356.
46. Mirny LA, Imakaev M, and Abdennur N. (2019). Two major mechanisms of chromosome organization. Curr. Opin. Cell Biol. 58, 142–152. 10.1016/j.ceb.2019.05.001. [PubMed: 31228682]
47. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, and Aiden EL (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell 159, 1665–1680. 10.1016/j.cell.2014.11.021. [PubMed: 25497547]
48. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A, and Cavalli G. (2017). Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell 171, 557–572.e24. 10.1016/j.cell.2017.09.043. [PubMed: 29053968]
49. de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, Teunissen H, Splinter E, Wijchers PJ, Krijger PHL, and de Laat W. (2015). CTCF Binding Polarity Determines Chromatin Looping. Mol. Cell 60, 676–684. 10.1016/j.molcel.2015.09.023. [PubMed: 26527277]
50. Crewe M, Segev A, Rueda R, and Madabhushi R. (2023). Atypical Modes of CTCF Binding Facilitate Tissue-Specific and Neuronal Activity-Dependent Gene Expression States. Preprint at Mol. Neurobiol. 10.1007/s12035-023-03762-5.
51. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293. 10.1126/science.1181369. [PubMed: 19815776]
52. Barbieri E, Hill C, Quesnel-Vallières M, Zucco AJ, Barash Y, and Gardini A. (2020). Rapid and Scalable Profiling of Nascent RNA with fastGRO. Cell Rep. 33, 108373. 10.1016/j.celrep.2020.108373. [PubMed: 33176136]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

53. Wissink EM, Vihervaara A, Tippens ND, and Lis JT (2019). Nascent RNA analyses: tracking transcription and its regulation. Nat. Rev. Genet. 20, 705–723. 10.1038/s41576-019-0159-6. [PubMed: 31399713]
54. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptewicz M, Barbara-Haley K, Kuersten S, et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. Nature 465, 182–187. 10.1038/nature09033. [PubMed: 20393465]
55. Danko CG, Hyland SL, Core LJ, Martins AL, Waters CT, Lee HW, Cheung VG, Kraus WL, Lis JT, and Siepel A. (2015). Identification of active transcriptional regulatory elements from GRO-seq data. Nat. Methods 12, 433–438. 10.1038/nmeth.3329. [PubMed: 25799441]
56. Love MI, Huber W, and Anders S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. 10.1186/s13059-014-0550-8. [PubMed: 25516281]
57. Solier S, Ryan MC, Martin SE, Varma S, Kohn KW, Liu H, Zeeberg BR, and Pommier Y. (2013). Transcription poisoning by Topoisomerase I is controlled by gene length, splice sites, and miR-142–3p. Cancer Res. 73, 4830–4839. 10.1158/0008-5472.CAN-12-3504. [PubMed: 23786772]
58. Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, and Kraus WL (2013). Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. Mol. Cell 50, 212–222. 10.1016/j.molcel.2013.02.015. [PubMed: 23523369]
59. Jonkers I, Kwak H, and Lis JT (2014). Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. Elife 3, e02407. 10.7554/eLife.02407. [PubMed: 24843027]
60. Lu LY, Kuang H, Korakavi G, and Yu X. (2015). Topoisomerase II regulates the maintenance of DNA methylation. J. Biol. Chem. 290, 851–860. 10.1074/jbc.M114.611509. [PubMed: 25451918]
61. Cowell IG, Papageorgiou N, Padget K, Watters GP, and Austin CA (2011). Histone deacetylase inhibition redistributes topoisomerase IIbeta from heterochromatin to euchromatin. Nucleus 2, 61–71. 10.4161/nucl.2.1.14194. [PubMed: 21647300]
62. Liu LF, and Wang JC (1987). Supercoiling of the DNA template during transcription. Proc. Natl. Acad. Sci. USA 84, 7024–7027. 10.1073/pnas.84.20.7024. [PubMed: 2823250]
63. Chong S, Chen C, Ge H, and Xie XS (2014). Mechanism of transcriptional bursting in bacteria. Cell 158, 314–326. 10.1016/j.cell.2014.05.038. [PubMed: 25036631]
64. Ma J, Bai L, and Wang MD (2013). Transcription under torsion. Science 340, 1580–1583. 10.1126/science.1235441. [PubMed: 23812716]
65. Cinghu S, Yang P, Kosak JP, Conway AE, Kumar D, Oldfield AJ, Adelman K, and Jothi R. (2017). Intragenic Enhancers Attenuate Host Gene Expression. Mol. Cell 68, 104–117.e6. 10.1016/j.molcel.2017.09.010. [PubMed: 28985501]
66. Li B, Gogol M, Carey M, Pattenden SG, Seidel C, and Workman JL (2007). Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. Genes Dev. 21, 1422–1430. 10.1101/gad.1539307. [PubMed: 17545470]
67. Neri F, Rapelli S, Krepelova A, Incarnato D, Parlato C, Basile G, Maldotti M, Anselmi F, and Oliviero S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. Nature 543, 72–77. 10.1038/nature21373. [PubMed: 28225755]
68. Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, and Workman JL (2012). Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. Nature 489, 452–455. 10.1038/nature11326. [PubMed: 22914091]
69. Venkatesh S, Li H, Gogol MM, and Workman JL (2016). Selective suppression of antisense transcription by Set2-mediated H3K36 methylation. Nat. Commun. 7, 13610. 10.1038/ncomms13610. [PubMed: 27892455]
70. Kim T, and Buratowski S. (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5’ transcribed regions. Cell 137, 259–272. 10.1016/j.cell.2009.02.045. [PubMed: 19379692]
71. Kim T, Xu Z, Clauder-Münster S, Steinmetz LM, and Buratowski S. (2012). Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. Cell 150, 1158–1169. 10.1016/j.cell.2012.08.016. [PubMed: 22959268]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

72. Andrews S. A Quality Control Tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
73. Bolger AM, Lohse M, and Usadel B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. 10.1093/bioinformatics/btu170. [PubMed: 24695404]
74. Li H, and Durbin R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760. 10.1093/bioinformatics/btp324. [PubMed: 19451168]
75. Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, and Prins P. (2015). Sambamba: fast processing of NGS alignment formats. Bioinformatics 31, 2032–2034. 10.1093/bioinformatics/btv098. [PubMed: 25697820]
76. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, and Li H. (2021). Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. 10.1093/gigascience/giab008.
77. Kuhn RM, Haussler D, and Kent WJ (2013). The UCSC genome browser and associated tools. Briefings Bioinf. 14, 144–161. 10.1093/bib/bbs038.
78. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, and Salzberg SL (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36. 10.1186/gb-2013-14-4-r36. [PubMed: 23618408]
79. Martin M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. j. 17, 10–12.
80. Ramírez F, Du€ndar F, Diehl S, Gru€ning BA, and Manke T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 42, W187–W191. 10.1093/nar/gku365. [PubMed: 24799436]
81. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589. 10.1016/j.molcel.2010.05.004. [PubMed: 20513432]
82. Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y, Qiu J, Liu W, Kaikkonen MU, Ohgi KA, et al. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. Nature 474, 390–394. 10.1038/nature10006. [PubMed: 21572438]
83. Wang Z, Chu T, Choate LA, and Danko CG (2019). Identification of regulatory elements from nascent transcription using dREG. Genome Res. 29, 293–303. 10.1101/gr.238279.118. [PubMed: 30573452]
84. Chu T, Wang Z, Chou SP, and Danko CG (2019). Discovering Transcriptional Regulatory Elements From Run-On and Sequencing Data Using the Web-Based dREG Gateway. Curr. Protoc. Bioinformatics 66, e70. 10.1002/cpbi.70. [PubMed: 30589513]
85. Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, and Aiden EL (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Syst. 3, 95–98. 10.1016/j.cels.2016.07.002. [PubMed: 27467249]
86. Machanick P, and Bailey TL (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics 27, 1696–1697. 10.1093/bioinformatics/btr189. [PubMed: 21486936]
87. Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. 10.1093/bioinformatics/btq033. [PubMed: 20110278]
88. Yu G, Wang LG, and He QY (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 31, 2382–2383. 10.1093/bioinformatics/btv145. [PubMed: 25765347]
89. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. 22, 1813–1831. 10.1101/gr.136184.111. [PubMed: 22955991]
90. Li Q, Brown JB, Huang H, and Bickel PJ (2011). Measuring Reproducibility of High-Throughput Experiments. Ann. Appl. Stat. 5, 1752–1779. 10.1214/11-Aoas466.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

91. Nassar LR, Barber GP, Benet-Pagès A, Casper J, Clawson H, Diekhans M, Fischer C, Gonzalez JN, Hinrichs AS, Lee BT, et al. (2023). The UCSC Genome Browser database: 2023 update. Nucleic Acids Res. 51, D1188–D1195. 10.1093/nar/gkac1072. [PubMed: 36420891]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- The distribution of TOP2B activity in chromatin relates to nucleosome configuration
- TOP2B activity varies relative to chromosome compartments but not loops and TADs
- Inhibiting TOP2B stimulates nascent transcription at active genes and enhancers
- TOP2B promotes long gene expression by suppressing cryptic intragenic transcription

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

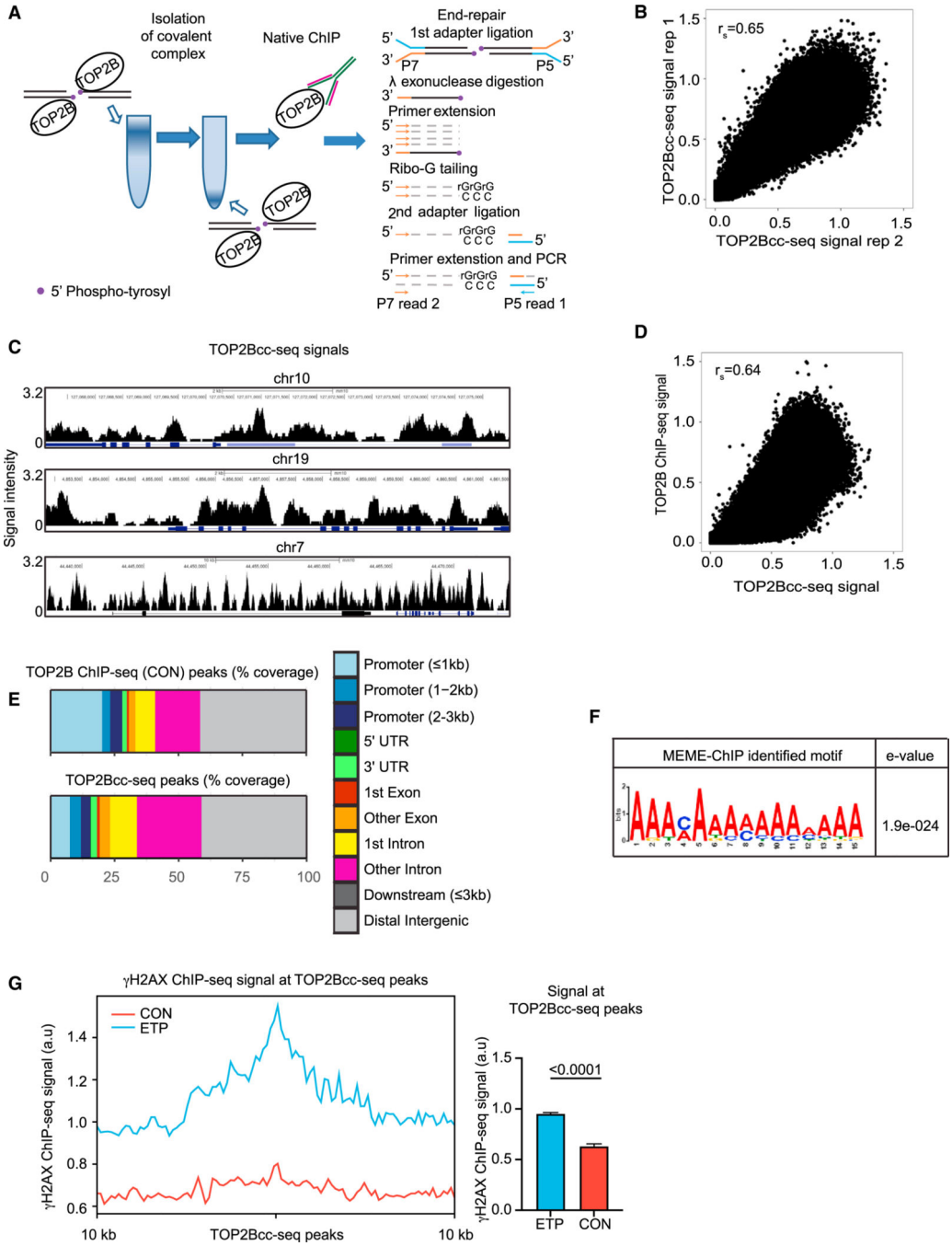


Figure 1. Mapping sites of catalytically active TOP2B in mouse primary cortical neurons with TOP2Bcc-seq

(A) Diagram of TOP2Bcc-seq protocol.

(B) Concordance between the two biological replicates of TOP2Bcc-seq experiments across genomic intervals (Spearman rank, $R^2 = 0.65$).

(C) Representative UCSC browser view of TOP2Bcc-seq signals at three genomic intervals.

(D) Concordance between TOP2Bcc-seq and TOP2B-ChIP seq data[24] (Spearman rank, $R^2 = 0.64$).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(E) Distribution of TOP2Bcc-seq (n = 2 biological replicates) and TOP2B ChIP-seq (n = 3 biological replicates) peaks across chromatin states annotated by ChIPseeker.
(F) Top DNA binding motif enriched at TOP2Bcc-seq peaks (n = 35,505) discovered via MEME-ChIP.
(G) Average γH2AX signal intensity in control (CON) and ETP-treated neurons surrounding the top 500 TOP2Bcc-seq peaks (left) with quantification of median γH2AX signal intensity (right, Mann-Whitney U, $p < 0.0001$).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

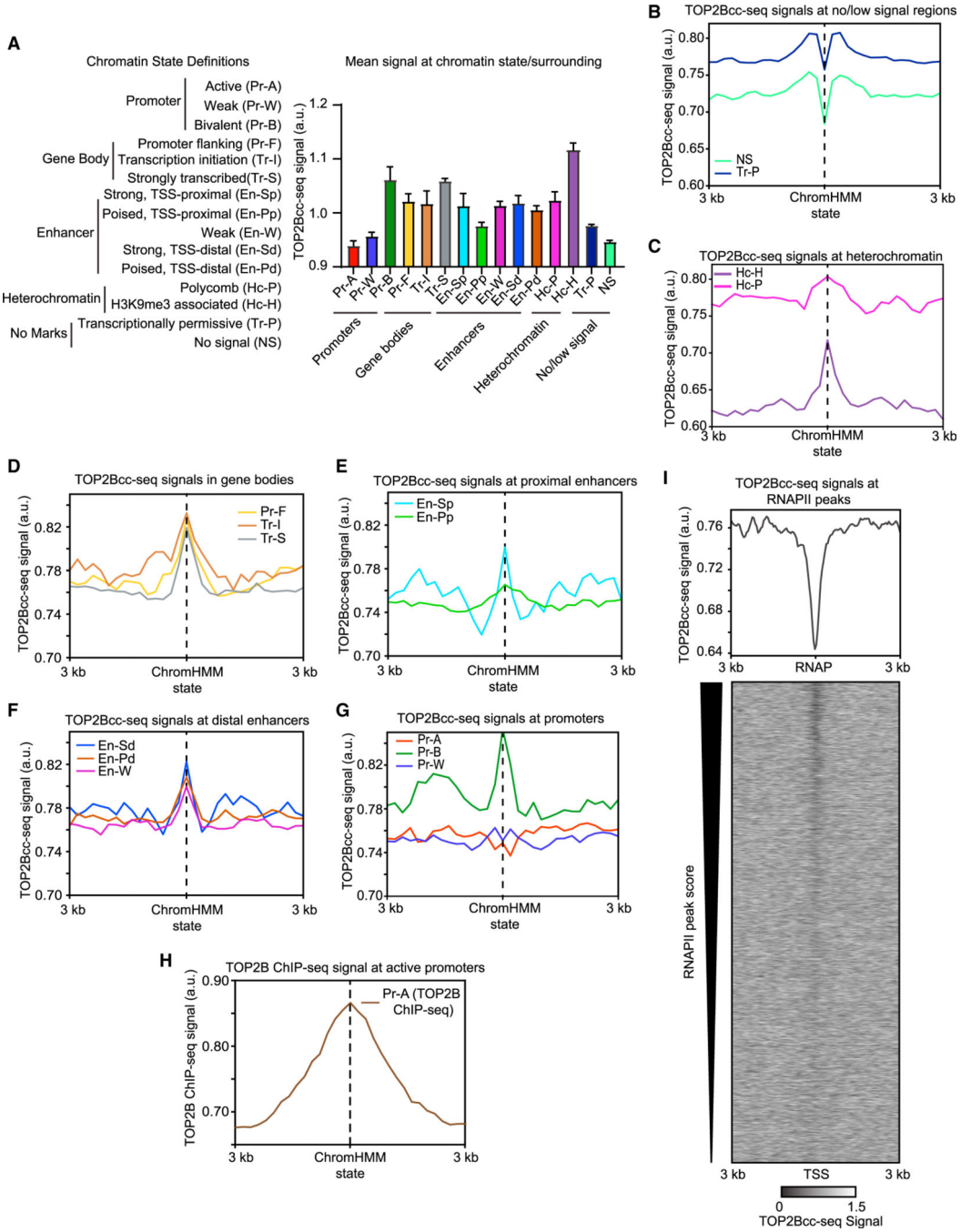


Figure 2. Assessing the distribution of TOP2B activity across annotated chromatin states

(A) Enrichment of TOP2Bcc-seq signal intensity within each chromatin state annotated by ChromHMM compared to the signal intensity in surrounding regions.

(B–G) Aggregate plots of mean TOP2Bcc-seq signals at and surrounding 5 kb bins for the indicated chromatin states.

(H) Aggregate plot of mean TOP2B ChIP-seq signals surrounding active promoters (Pr-A).

(I) (Top) Aggregate plot of mean TOP2Bcc-seq signals surrounding RNAPII peaks. (Bottom) Heatmap depicting the intensity of TOP2Bcc-seq signals at RNAPII peaks ranked in descending order of their peak score.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

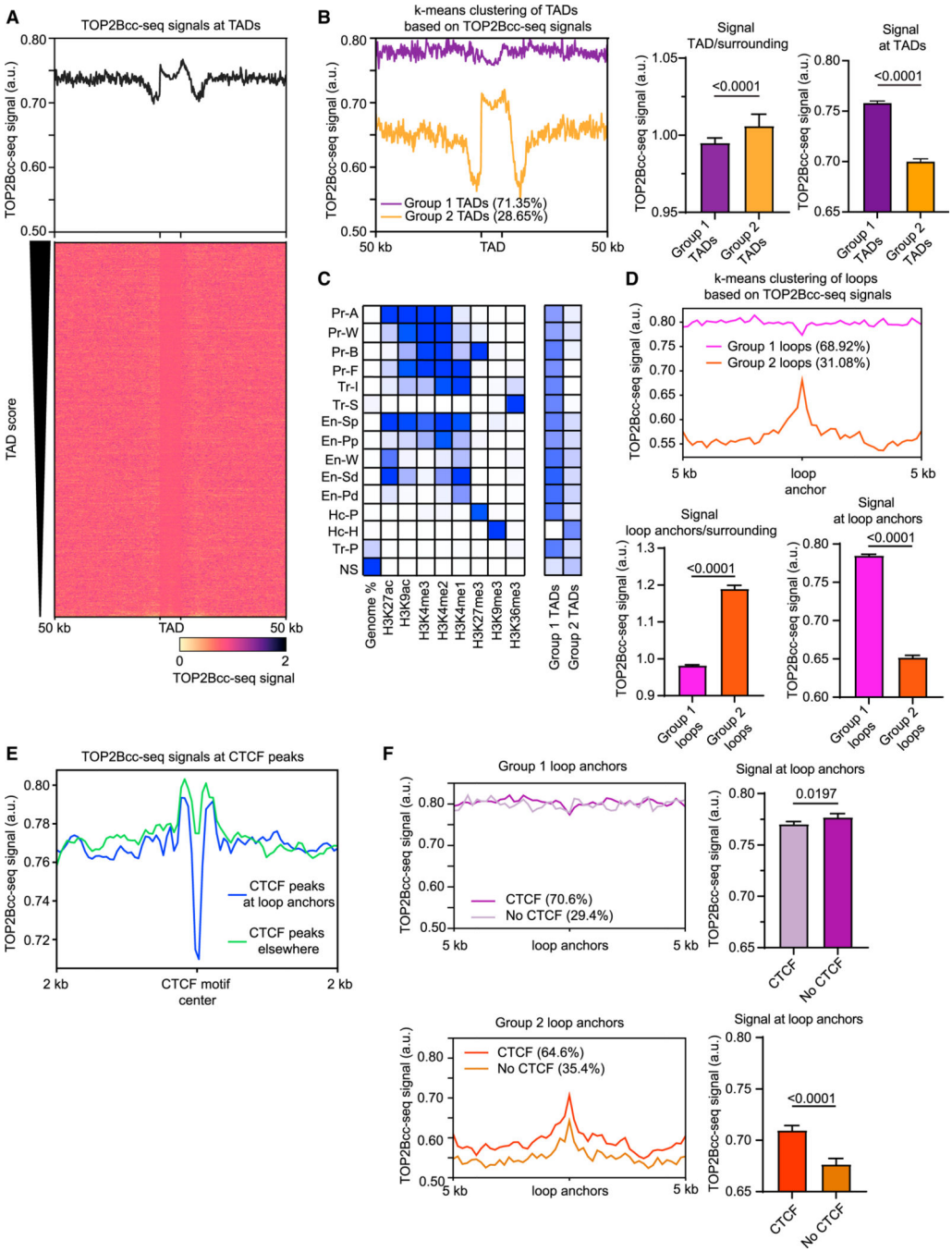


Figure 3. Investigating TOP2B activity at TAD boundaries and loop anchors

(A) (Top) Aggregate plot of mean TOP2Bcc-seq signals within/surrounding TAD boundaries identified from previously published Hi-C data in primary cortical neurons.[48] (Bottom) Heatmap view of TOP2Bcc-seq signals within/surrounding each TAD boundary ranked according to TAD strength.

(B) (Left) Mean TOP2Bcc-seq signal within/surrounding two groups of TAD boundaries identified by k-means clustering of TOP2Bcc-seq signals at TAD boundaries. (Middle) Quantification of median TOP2Bcc-seq signals within TAD boundaries compared to signals

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in surrounding regions (Mann-Whitney U, $p < 0.0001$). (Right) Quantification of median TOP2Bcc-seq signals within TADs (Mann-Whitney U, $p < 0.0001$).

(C) Heatmap representation of chromatin states annotated by ChromHMM that are enriched within group 1 and group 2 TADs that were identified by k-means clustering, as shown in (B).

(D) (Top) Aggregate plot of TOP2Bcc-seq signals within/surrounding two groups of chromatin loop anchors identified by k-means clustering of TOP2Bcc-seq signals at chromatin loop anchors. (Bottom left) Quantification of median TOP2Bcc-seq signal within loop anchors compared to signals in surrounding regions (Mann-Whitney U, $p < 0.0001$). (Bottom right) Quantification of median TOP2Bcc-seq signal within loop anchors (Mann-Whitney U, $p < 0.0001$).

(E) Distribution of TOP2Bcc-seq signals in 2 kb windows surrounding oriented CTCF motif sites that occur either at chromatin loop anchors (blue line) or elsewhere in the genome (green line).

(F) (Left) Aggregate plots comparing mean TOP2Bcc-seq signals at group 1 (top) and group 2 (bottom) loop anchors that either do or do not overlap with CTCF ChIP-seq peaks. (Right) quantification of median TOP2Bcc-seq signals of the corresponding loop anchors (Mann-Whitney U, group 1 $p = 0.0197$, group 2 $p < 0.0001$).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

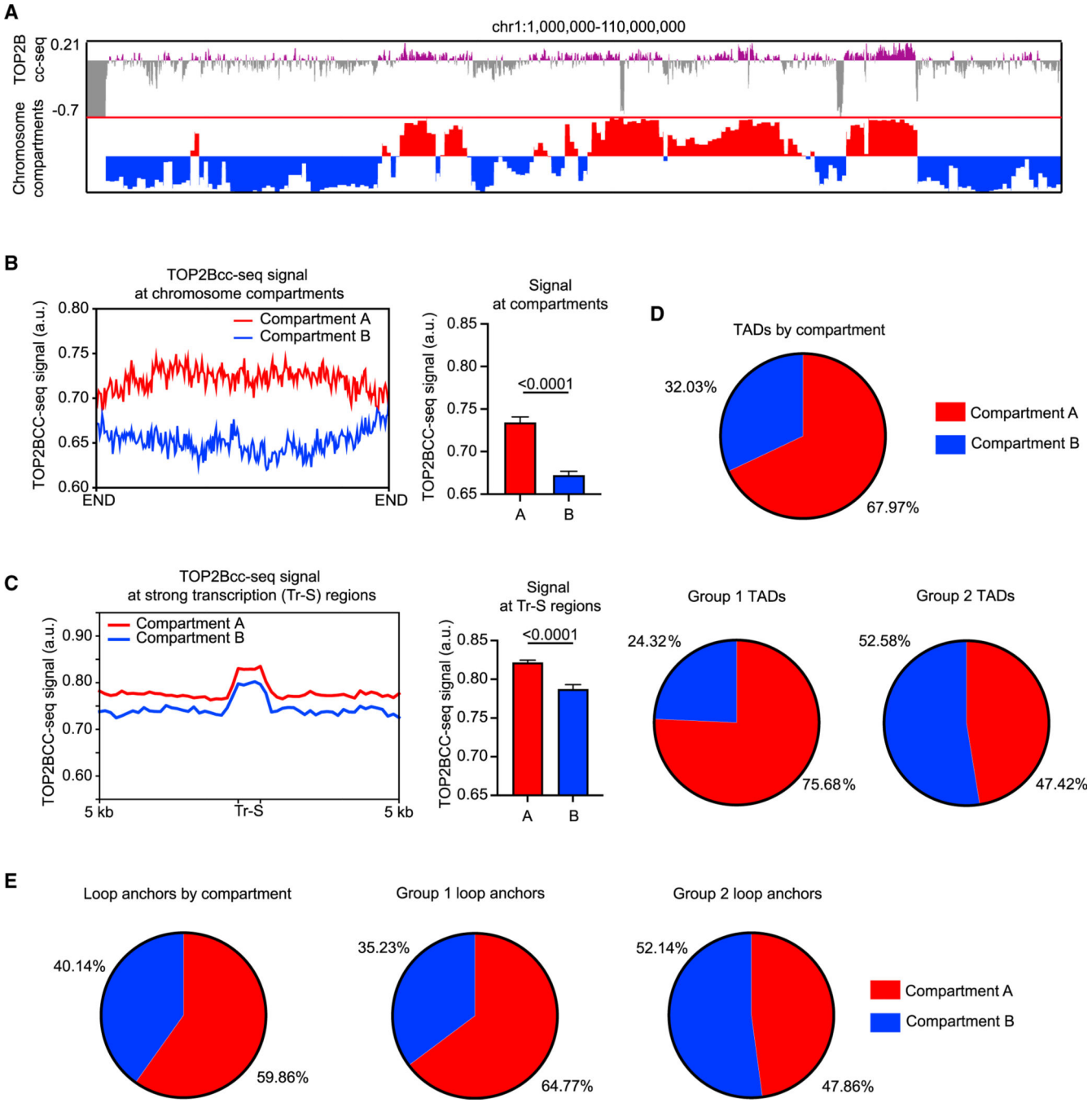


Figure 4. The role of chromosome compartmental organization in the regulation of TOP2B activity

(A) UCSC browser view of a portion of chromosome 1 showing the alignment between TOP2Bcc-seq signals and chromosome compartment designation (red: compartment A, blue: compartment B).

(B) Aggregate plots of mean TOP2Bcc-seq signals within chromosome compartments A and B (left) with quantification of median TOP2Bcc-seq signals for each compartment (right, Mann-Whitney U, $p < 0.0001$).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(C) Comparison of mean TOP2Bcc-seq signals within/surrounding regions of strong transcription (Tr-S) annotated by ChromHMM in chromosome compartments A and B (left) with quantification of median TOP2Bcc-seq signal for each compartment (right, Mann-Whitney U, $p < 0.0001$).

(D) Percentage of all TADs that occur within chromosome compartment A or B (top) compared to the percentage of group 1 (bottom left) or group 2 (bottom right) TADs that were identified by k-means clustering that occur within each compartment.

(E) Percentage of all chromatin loop anchors that occur within either chromosome compartment A or B (left) compared to the percentage of group 1 (middle) or group 2 (right) loop anchors identified by k-means clustering that occur within each compartment.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

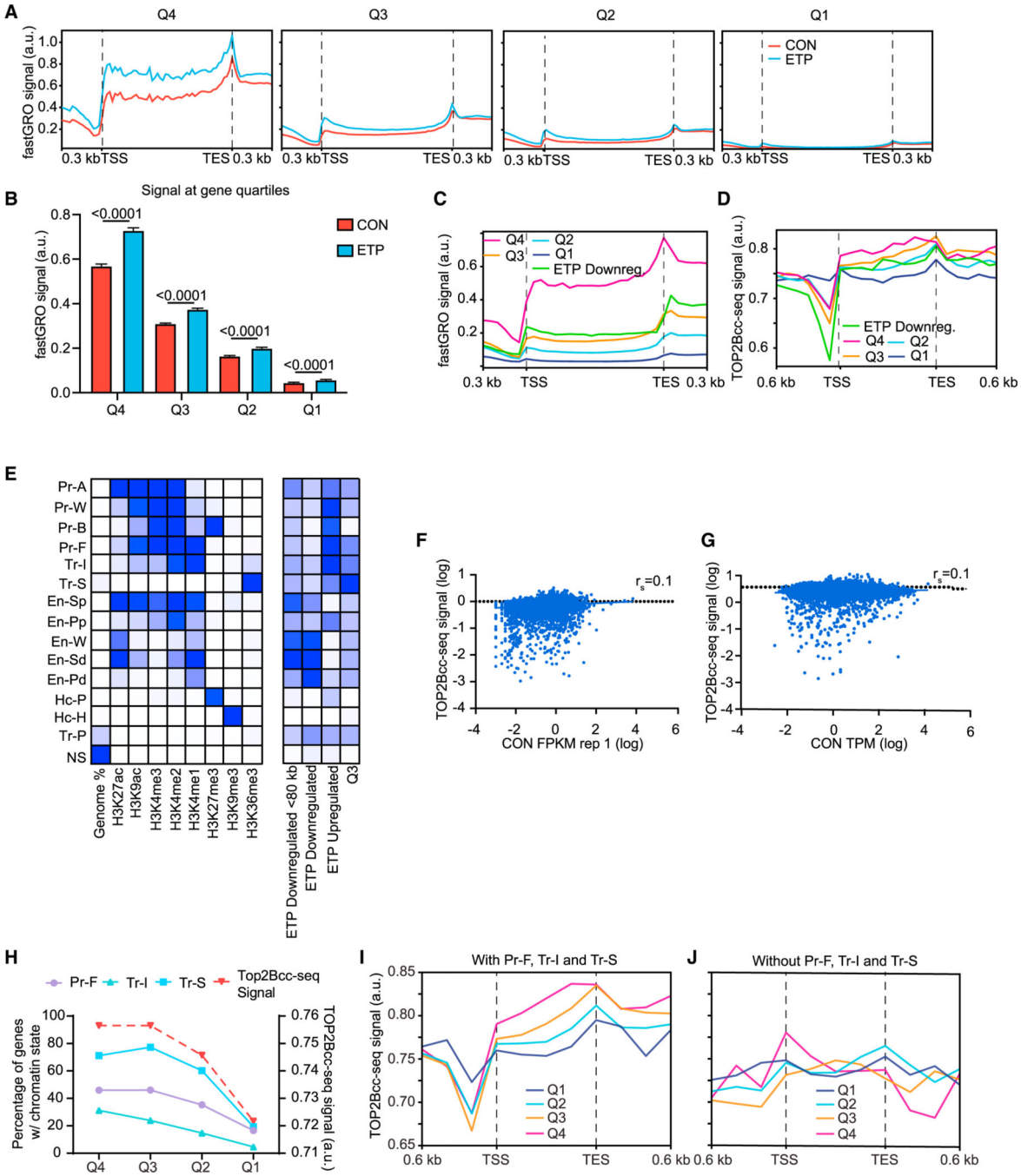


Figure 5. Assessing the influence of transcriptional activity and nucleosome configuration on TOP2B activity

(A) Mean fastGRO signal within/surrounding gene bodies for CON and ETP-treated neurons within each expression quartile (Q4, highest expression; Q1, lowest expression).

(B) Quantification of median fastGRO profiles in (A) (Mann-Whitney U, all $p < 0.0001$).

(C and D) Mean fastGRO signals in untreated neurons (C) and mean TOP2Bcc-seq signals (D) within/surrounding genes in each expression quartile were compared to corresponding signals in ETP-downregulated genes (ETP downreg.).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(E) Heatmap of chromatin states enriched in various classes of ETP-up- and -downregulated genes and Q3 genes, which resemble expression levels of ETP-downregulated genes.
(F and G) Correlation between TOP2Bcc-seq signal intensity and either fastGRO FPKMs (F) or RNA-seq TPMs (G) for all transcripts.
(H) Percentage of genes within each expression quartile that contain promoter flanking (Pr-F), transcription initiation (Tr-I), and strong transcription (Tr-S) chromatin states annotated by ChromHMM were compared to the mean TOP2Bcc-seq signal intensity for each expression quartile.
(I and J) Aggregate plots of TOP2Bcc-seq signals in each expression quartile in genes that either (I) contain (I) or (J) lack Pr-F, Tr-I, and Tr-S chromatin states.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

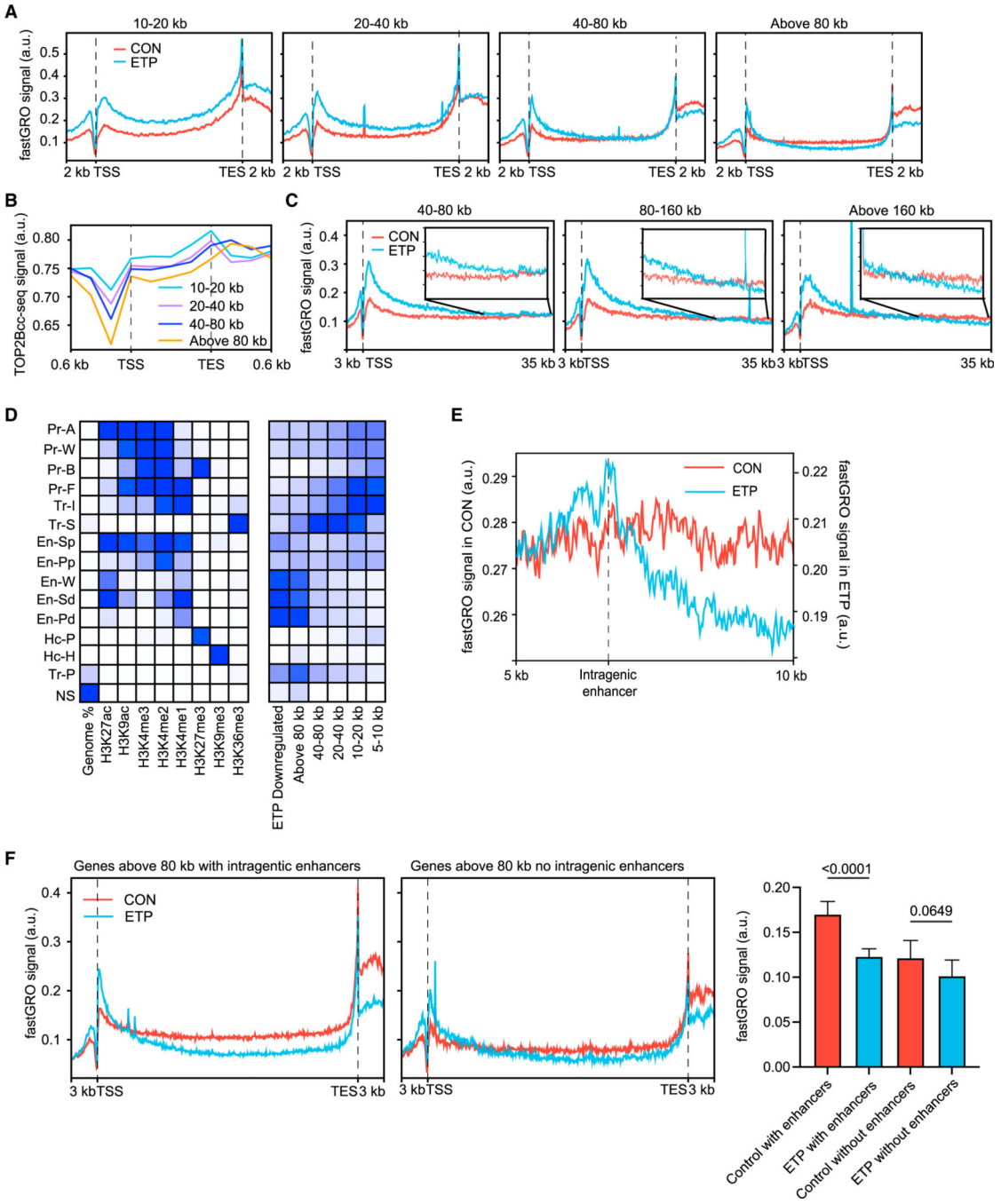


Figure 6. Effect of intergenic enhancer enrichment on the expression of long neuronal transcripts in ETP-treated neurons

(A) Aggregate plots of mean fastGRO signals in CON and ETP-treated neurons within/surrounding gene bodies that were grouped according to their length as indicated.

(B) Mean TOP2Bcc-seq signals within/surrounding gene bodies that were grouped according to their length.

(C) Comparison of mean fastGRO signals in the first 35 kb of gene bodies across the indicated length categories in CON and ETP-treated neurons.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(D) Heatmap representation of chromatin states enriched within ETP-downregulated genes and genes in the indicated length categories.
(E) Comparison of mean fastGRO signals in intervals immediately preceding and succeeding intragenic enhancers in ETP-treated neurons and untreated CONs. The signals are overlaid to depict relative trends, and the signal intensity for each condition is depicted on separate y axes (CON, left; ETP, right).
(F) Aggregate plot of mean fastGRO signals at genes longer than 80 kb in length that either contain (left) or lack intragenic enhancers (middle) in ETP-treated neurons and untreated CONs. (Right) Quantification of median fastGRO signal intensity for each category (Mann-Whitney U, CON vs. ETP with enhancers $p < 0.0001$, CON vs. ETP without enhancers $p = 0.06495$).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Cell Rep. Author manuscript; available in PMC 2024 May 02.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-TOP2B	Santa Cruz	RRID:AB_628384
Rabbit monoclonal anti-CTCF	Cell Signaling	RRID:AB_2086791
Chemicals, peptides, and recombinant proteins		
Etoposide	Sigma	Cat#:E1383
ICRF-193	Sigma	Cat#I4659
Pierce ChIP-grade Protein A/G magnetic beads	ThermoFisher	Cat#:26162
Lambda Exonuclease	NEB	Cat#:M0262S
Phusion DNA Polymerase	NEB	Cat#:M0530S
AMPure XP Beads	Beckman	Cat#:A63881
rGTP	VWR	Cat#:PAE6031
TdT	NEB	Cat#:M0315S
Proteinase K	NEB	Cat#:P8107S
Superase-IN	ThermoFisher	Cat#:AM2694
Trizol LS	ThermoFisher	Cat#:10296010
4-thio-UTP	FisherScientific	Cat#:NC0091659
4-thiouridine	Sigma	Cat#:T4509
EZ-link HPDP-biotin	ThermoFisher	Cat#:A35390
M280 Streptavidin Dynabeads	ThermoFisher	Cat#:11205D
Critical commercial assays		
KAPA HyperPrep Kit	Roche	Cat#: 07962363001
Qiagen RNeasy Plus Universal Kit	Qiagen	Cat#:74034
TruSeq stranded mRNA library prep	Illumina	Cat#:20020594
Zymo Clean and Concentrator Kit	Zymo	Cat#R1013
KAPA RNA HyperPrep Kit	KAPA	Cat#:08098093702
Deposited data		

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
H3K9me3 – P0 Forebrain	ENCODE	ENCSR093DWU
H3K27me3 – P0 Forebrain	ENCODE	ENCSR070MOK
H3K36me3– P0 Forebrain	ENCODE	ENCSR069TDC
H3K4me2– P0 Forebrain	ENCODE	ENCSR906UJW
H3K4me3– P0 Forebrain	ENCODE	ENCSR258YWW
H3K4me1– P0 Forebrain	ENCODE	ENCSR465PLB
H3K9ac– P0 Forebrain	ENCODE	ENCSR369RBO
H3K27ac– P0 Forebrain	ENCODE	ENCSR094TTT
POL2A – E14.5 Forebrain	ENCODE	ENCSR000CEI
HiC – Cortical Neurons	Bonev et al.[48]	GEO: GSE96107
END-seq	Canela et al.[26]	GEO: GSE99194
TOP2B-ChIP seq	Madabhushi et al.[24]	GEO: GSE61887
γH2AX ChIP-seq	Madabhushi et al.[24]	GEO: GSE61887
TOP2Bcc-seq	This Paper	GEO: GSE227463
fastGro seq – Control Neurons	This Paper	GEO: GSE227463
fastGro seq – ETP treated Neurons	This Paper	GEO: GSE227463
fastGro seq – ICRF-193 treated Neurons	This Paper	GEO: GSE227463
RNA-seq Control Neurons	This Paper	GEO: GSE227463
Experimental models: Cell lines		
Schneider's Drosophila Line 2 (S2)	ATCC	CRL-1963
Experimental models: Organisms/strains		
Swiss Webster Mice	Charles River	Strain 024
Oligonucleotides		
P7 extension primer: CAAGCAGAAGACGGCATACGA*G (* denotes phophorothioate linkage)	This paper, supplied by IDT	N/A
CCC overhang adapter oligo1: AATGATACGGCGACCACCGAGATCTACA CTCTTTCCCTACACGACGCTCTTCCGATC	Teves et al.,[35] supplied by IDT	N/A

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CCC overhang adapter oligo2: (Phosphate)GA TCGGAAGAGCGGTTCAGCAGGAATGCCGAG	Teves et al.,[35] supplied by IDT	N/A
Software and algorithms		
FastQC	S. Andrew[72]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Trimmomatic	Bolger et al.[73]	http://www.usadellab.org/cms/index.php?page=trimmomatic
BWA	Li et al.[74]	https://github.com/lh3/bwa
Sambamba	Tarasov et al.[75]	https://github.com/biod/sambamba/releases
Samtools	Danecek et al.[76]	http://www.htslib.org/
MACS2	Zhang et al.[36]	https://github.com/macs3-project/MACS
UCSC Kent tools	Kuhn et al.[77]	http://hgdownload.soe.ucsc.edu/downloads.html#source_downloads
TopHat	Kim et al.[78]	https://ccb.jhu.edu/software/tophat/index.shtml
Cutadapt	M. Martin[79]	https://github.com/marcelm/cutadapt/
Deeptools	Ramirez et al.[80]	https://github.com/deeptools/deepTools
HOMER	Heinz et al.[81] Wang et al.[82]	http://homer.ucsd.edu/homer/download.html
DESeq2	Love et al.[56]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
dREG	Wang et al.[83]	https://dreg.dnasequence.org/
RunOnBamToBigWig	Chu et al.[84]	https://github.com/Danko-Lab/RunOnBamToBigWig
Juicer	Durand et al.[85]	https://github.com/aidenlab/juicer
HiCCUPS	Rao et al.[47]	https://github.com/aidenlab/juicer/wiki/HiCCUPS
Arrowhead	Rao et al.[47]	https://github.com/aidenlab/juicer/wiki/Arrowhead
MEME-ChIP	Machanick et al.[86]	https://meme-suite.org/meme/doc/meme-chip.html
ChromHMM	Ernst et al.[40]	http://compbio.mit.edu/ChromHMM/#:~:text=ChromHMM%20is%20based%20on%20a,one%20or%20more%20cell%20types.
Bedtools	Quinlan et al.[87]	https://github.com/arq5x/bedtools2
ChIPseeker	Yu et al.[88]	https://bioconductor.org/packages/release/bioc/html/ChIPseeker.html