NAR Cancer, 2024, **6**, zcae016
https://doi.org/10.1093/narcan/zcae016
Advance access publication date: 8 April 2024
Cancer Genomics

TDP2 is a regulator of estrogen-responsive oncogene expression

Nicholas Manguso[1,†], Minhyung Kim[2,†], Neeraj Joshi[1], Md Rasel Al Mahmud[3], Juan Aldaco[1], Ryusuke Suzuki[1], Felipe Cortes-Ledesma[4], Xiaojiang Cui[1,5], Shintaro Yamada[3], Shunichi Takeda[3], Armando Giuliano[1,5], Sungyong You[2,5] and Hisashi Tanaka[1,5,6,*]

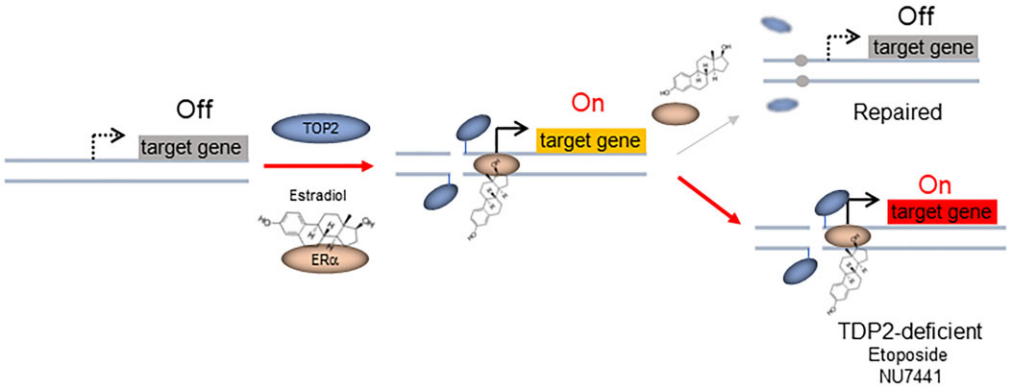
[1]Department of Surgery, Cedars-Sinai Medical Center, West Hollywood, CA 90048 USA
[2]Department of Urology and Computational Biomedicine, Cedars-Sinai Medical Center, West Hollywood, CA 90048 USA
[3]Department of Radiation Genetics, Graduate School of Medicine, Kyoto University, Kyoto 606-8501, Japan
[4]Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER), CSIC-Universidad de Sevilla-Universidad Pablo de Olavide, Sevilla, 41092, Spain
[5]Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, West Hollywood, CA 90048, USA
[6]Department of Biomedical Sciences, Cedars-Sinai Medical Center, West Hollywood, CA 90048, USA

[*]To whom correspondence should be addressed. Tel: +1 310 423 0551; Email: Hisashi.Tanaka@cshs.org
[†]The first two authors should be regarded as Joint First Authors.
Present addresses:
Rasel Al Mahmud, Developmental Therapeutics Branch and Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA.
Shunichi Takeda, Shenzhen University School of Medicine, Shenzhen, Guangdong, China.
Felipe Cortes-Ledesma, Topology and DNA breaks Group, Spanish National Cancer Centre (CNIO), Madrid, Spain.

Abstract

With its ligand estrogen, the estrogen receptor (ER) initiates a global transcriptional program, promoting cell growth. This process involves topoisomerase 2 (TOP2), a key protein in resolving topological issues during transcription by cleaving a DNA duplex, passing another duplex through the break, and repairing the break. Recent studies revealed the involvement of various DNA repair proteins in the repair of TOP2-induced breaks, suggesting potential alternative repair pathways in cases where TOP2 is halted after cleavage. However, the contribution of these proteins in ER-induced transcriptional regulation remains unclear. We investigated the role of tyrosyl-DNA phosphodiesterase 2 (TDP2), an enzyme for the removal of halted TOP2 from the DNA ends, in the estrogen-induced transcriptome using both targeted and global transcription analyses. *MYC* activation by estrogen, a TOP2-dependent and transient event, became prolonged in the absence of TDP2 in both TDP2-deficient cells and mice. Bulk and single-cell RNA-seq analyses defined *MYC* and *CCND1* as oncogenes whose estrogen response is tightly regulated by TDP2. These results suggest that TDP2 may inherently participate in the repair of estrogen-induced breaks at specific genomic loci, exerting precise control over oncogenic gene expression.

Graphical abstract



Introduction

With a lifetime risk of one in eight women, breast cancer is the most common female cancer and is the most frequent cause of cancer-related death in women. About two-thirds of breast tumors express estrogen receptor alpha (ERα), and tumor growth depends on the hormone estrogens (1). Estrogens,

Received: April 28, 2023. **Revised:** February 19, 2024. **Editorial Decision:** March 6, 2024. **Accepted:** March 14, 2024

particularly 17beta-estradiol (E2), play a crucial role in female reproductive system development. Secreted E2 diffuses into the cell nucleus, where it binds to a receptor ERα (2). The E2–ERα complex then binds to DNA, acting as a potent regulator of gene expression.

In vitro studies have shown that transcriptional activation occurs within an hour after E2 stimulation and involves Type II topoisomerase (TOP2), in particular, TOP2 beta (TOP2β) (3,4). TOP2 resolves torsional stress during DNA metabolism processes, including replication, chromosome segregation, and transcription, by transiently cleaving one duplex DNA to introduce a DNA double-stranded break (DSB) and passing another DNA duplex through the DSB (5–7). These reactions involve a transient intermediate with covalent bonds between the catalytic tyrosine in TOP2 and the 5′ phosphate groups of the DNA ends (TOP2 cleavage complex, TOP2cc), followed by the re-ligation of the DSB ends. In general transcription, TOP2 relieves over-wound DNA ahead of RNA polymerase II (PolII), facilitating transcriptional elongation by releasing paused PolII (8–10). Because of the critical and broad roles of TOP2 in nucleotide metabolisms, small molecule inhibitors that intervene in the TOP2 activity have been developed. TOP2 catalytic inhibitors block a cascade of conformational changes and enzymatic activities, including duplex DNA cleavage, and TOP2-poisons stabilize TOP2cc and prevent the re-ligation of TOP2-induced DSBs (11–13).

In addition to relieving torsional stress in general transcription, TOP2, in particular TOP2β, plays a crucial role in the activation of a specific set of genes in response to exogenous stimuli, including E2, heat shock, serum starvation, and neuronal activity (8,10,14,15). DSBs induced by TOP2β at gene promoters and transcription start sites are followed by the recruitment of DSB repair proteins, DNA-PK and Ku70 (10,14). The recruitment is somewhat surprising, given that TOP2 is a self-sufficient enzyme and catalyzes and relegates DSBs (5,16). TOP2β occupies more than 2000 sites in the genome upon E2 treatment in MCF7 breast cancer cells (17), with a fraction of these sites possibly recruiting repair proteins. At least two processes are involved in the repair of DSBs with trapped Top2; (i) proteins that process (remove) TOP2 that remains covalently attached to the 5′ DSB ends (TOP2 cleavage complex, TOP2cc) and (ii) those that repair the resulting clean DSB ends. The first process was executed either by 5′-tyrosyl-DNA phosphodiesterase (TDP2) or the Mre11–RAD50–Nbs1 (MRN) complex (18–20). The resulting clean ends are repaired by non-homologous end-joining (NHEJ) (21). BRCA1 and RAD54L2 are also implicated in the resolution of TOP2cc (20,22,23). While these processes are generally seen as backups for instances where TOP2cc is trapped at DNA ends, this perspective may not fully explain the recruitment of repair proteins to specific gene promoters in response to exogenous stimuli.

Intriguingly, when left unrepaired, TOP2-induced DSBs can result in transcriptional activation of a specific set of genes. For example, etoposide treatment results in the upregulation of early response genes to neuronal activity in cultured primary neurons (14). The mechanism underlying the upregulation is of great interest, and it is also crucial to define which genes are upregulated in response to other exogenous stimuli. The upregulation of target genes could cause a serious problem when external stimuli, such as E2, promote cell proliferation, potentially leading to abnormal cell proliferation and, ultimately, tumorigenesis. Such a possibility was shown at the protein level in mice lacking NHEJ repair; E2 stimulated the expression of MYC protein in mammary epithelial cells of mice lacking DNA-PKcs (24).

We wished to identify a set of E2-induced genes that becomes upregulated in cells deficient in the resolution of trapped TOP2cc. The list of such genes would inform gene loci where trapped TOP2cc is inherently repaired by DSB repair proteins. To do so, we created ER-positive breast cancer cell lines that lack TDP2 and investigated a global transcriptional program. We chose TDP2 because TDP2 knockout cells are viable and amenable to E2 stimulation experiments. Using digital PCR, bulk- and single cell-RNAseq, we identified *MYC* as one of the several genes that exhibited enhanced transcription in response to E2 when TDP2 was deficient. The gene list included another ER-positive breast cancer oncogene, *CCND1*. These results suggest that DSB repair genes are inherently configured to resolve TOP2cc and tightly regulate oncogenic gene expression.

Materials and methods

Cell culture and digital PCR

MCF7 and MCF7 TDP2KO cells were maintained with MEM with 10% FBS. T47D, T47D TDP2KO and MDA-MB-231 cells were cultured with DMEM with 10% FBS. TDP2KO cells were created by targeting exon 2 of the *TDP2* gene (5′-GGCTCAGAGATGGTTTCAGGT-3′). The targeting vector was constructed using pSpCas9(BB)-2A-GFP(PX458), a gift from Feng Zhang (Addgene plasmid #48138).

Etoposide sensitivity was determined using the colony formation assay. Briefly, 3000 cells were plated in each well of a 6-well plate, and cells were cultured for 8 days with or without etoposide before staining with crystal violet. Experiments were performed in triplicate. Colonies were counted using the Hybrid Cell Function of the Keyence BZ-X700.

RNA was extracted using the RNeasy mini kit, and first-strand cDNA synthesis was performed using Superscript III (Invitrogen). Cells were cultured in phenol red-free medium supplemented with 5% FBS-charcoal stripped for 24 h then treated with 10 nM 17β-estradiol for either 1, 2 or 6 h before extracting RNA. MCF7 and T47D cells were treated with increasing concentrations of ICRF-193 for 1 h then treated with 10 nM E2 for 2 h. NU7441 was added to the media 3 h before adding 10 nM E2. Etoposide was added to the media 1 h before adding E2. Digital PCR was done using the QuantStudio 3D Digital PCR System. A TaqMan probe labeled with FAM (MYC-Hs00153408, IL20-Hs00218888 or OLFML3-Hs01113293 and VIC (TBP-Hs00427620 or TFRC-Hs00951083) were mixed in a reaction, and the amount of transcript (copies/microliter) was determined using the Analysis Suite (Thermo Fisher Scientific).

Mouse experiments

TDP2 Wild-type and *TDP2*$^{-/-}$ strains (21) came from the C57BL6/J strain. E2 (300 μg/kg body weight) was administered via intraperitoneal injection (IP) into three independent two-month-old female mice at 100 μl volume. As a control, the same volume of an E2 solvent (PBS) was injected.

Mice were sacrificed to harvest mammary gland tissues, and frozen blocks were prepared. After a brief wash with cold PBS, tissues were fixed with paraformaldehyde (4%, Cat# 163-20145, Wako, Japan) for 15 min at 4°C and then washed 3 times with 1× PBS. To preserve the tissue integrity, incu-

bation was done in 30% sucrose (in PBS) for 30 min to 3 h. Tissues were then embedded into optimal cutting temperature (OCT) compound (Cat# 4583, Sakura, Japan) into Cryomolds (Cat# 4565, Sakura, Japan). For cryo-sectioning of frozen blocks, a cryostat (CM1850, Leica, Germany) was used at −25°C and sectioned into 10 μm slices. Then, the slides were heated at 55°C for 10–30 min to dry. Dried slides were washed three times with Tween-20 in PBS (0.05%, PBS-T). For blocking 1% BSA and 5%, goat serum in PBS-T was used for 3 h at room temperature, then washed once with PBS-T. Slides were then incubated with both α-Cytokeratin-8/18 (1/10, rat monoclonal, University of Iowa, USA) and α-c-Myc antibody (1/200 [Y69], ab32072, rabbit monoclonal, Abcam) overnight at 4°C. After washing with PBS-T, slides were incubated with both α-rabbit (Alexa Fluor 546) and α-rat (Alexa Fluor 488) secondary antibodies (Molecular Probe, USA). c-Myc and Cytokeratin-8/18 signals were detected using LEICA SP8.

Bulk RNA-seq

MCF7, MCF7 with 10 nM E2 treatment for 2 h, MCF7 TDP2KO, MCF7 TDP2KO with 10 nM E2 treatment for 2 h were subject to RNA-seq at the Cedars-Sinai Applied Genomics, Computation and Translational Core. Biological duplicates were taken for all of the processes. Our quality control pipeline utilizes the Illumina standards to assess base call quality and truncates low-quality reads. The quality of sequencing data was assessed using MultiQC software (25). Sequence alignment and quantification were performed using the STAR-RSEM pipeline (26,27). Reads overlapping exons in the annotation of Genome Reference Consortium Human Build 38 (GRCh38) will be identified. Genes were filtered out and excluded from downstream analysis if they failed to achieve raw read counts of at least two across all the libraries. Batch effects were corrected by ComBat (28). The trimmed mean of M-values normalization method (TMM) (29) (version 1.6.1) was used for calculating normalized count data.

Differential expression analysis of bulk RNA-seq data

For the two comparisons, (i) E2 treated versus non-treated in MCF7 cells and (ii) E2 treated versus non-treated in MCF7 TDP2KO cells, we used the integrated hypothesis testing method as previously reported (30). Briefly, for each gene, two *P*-values were computed by performing a two-tailed Student's *t*-test and a two-tailed median different test using the empirical distributions that were estimated by random permutations of the samples. We then combined two *P*-values into a combined *P*-value using Stouffer's method. Multiple testing corrections were done using Storey's correction method (31). Finally, the DEGs were selected, with the ones having a combined *P*-value <0.05 and fold-change ≥ 1.5. The GSEA was used to check the enrichment of gene sets (32).

Single-cell RNA-seq data preprocessing and differential expression analysis

Single-cell RNA-seq was performed using the 10× Genomics scRNA-seq platform at the Cedars-Sinai Applied Genomics, Computation and Translational Core. The Cell Ranger Single-Cell Software Suite was used for demultiplexing, barcode assignment, and unique molecular identifiers (UMIs) quantification. The reads were aligned to the hg38 reference genome using a pre-built annotation package obtained from the 10× Genomics website. All lanes per sample were processed using the 'cellranger count' function. The output from different lanes was then aggregated using 'cellranger aggr' with –normalise set to 'none.' The quality control and analysis were conducted using the Seurat packages in R (version 3.2.2) (33). Poor-quality cells were identified by filtering out the cells with the number of detected genes <2000. In addition, all cells with 11.89% (three standard deviations) or more of UMIs mapping to mitochondrial genes were defined as non-viable or apoptotic and removed from the analysis. Genes detected in at least three cells were used for further analysis, resulting in a total of 19 043 genes remaining for the analysis. We also ensured that none of the reads in our data set were derived from index swapping. For this, we labeled cells with barcodes that appeared in more than one sample (non-unique barcodes) as 'NULL' and excluded these cells for further analysis. Gene expression values were normalized by log normalization using the 'NormalizeData' function in Seurat. We performed two comparisons of (i) E2 treated versus non-treated in MCF7 cells and (ii) E2 treated versus non-treated in MCF7 TDP2KO cells. The 'FindMarkers' function in Seurat was used with filtered data to identify DEGs between the cell groups. DEGs were selected with adjusted *P*-value <0.05 and log fold change ≥ 0.1 (1.11-fold). The fold change threshold was determined at 95 percentile value of the empirical null distribution of fold changes, which was generated by 1000 times random permutation of the cells.

Statistical analysis

We performed principal component analysis (PCA) for visualization of the samples to assess the distribution of the samples by gene expression profile. We used MATLAB (v.9.0; Mathworks, Natick, MA, USA), R (v.3.5), and Python (v.3.7) for bioinformatic analysis.

Ethics statement

The experimental protocol with mice was approved by the Ethical Committee for Animal Experimentation of the University of Seville (34).

Results

MYC induction by 17beta-estradiol (E2) depends on TOP2

MYC is a well-studied oncogene and is known to transcriptionally respond to E2 (35). To assess this response, we collected RNA from two ER-positive breast cancer cell lines, MCF7 and T47D, before and after treatment with E2 (Figure 1A). We employed digital PCR assays to measure the expression of *MYC* mRNA. In this assay, a *MYC* Taqman probe labeled with a dye FAM and a control probe, either *TBP* or *TFRC,* labeled with a dye VIC were mixed with cDNA from cell lines and PCR-amplified in approximately 20 000 × 1 nl reaction (Figure 1B). *MYC* expression levels (the number of *MYC*-positive 1 nl reactions) were normalized by the internal control (the number of internal control-positive reactions). We added 10 nM E2 and incubated cells for two h before extracting RNA. The fold induction of *MYC* was calculated as the ratio of normalized *MYC* expression with E2 treatment to normalized *MYC* expression without E2 treatment.

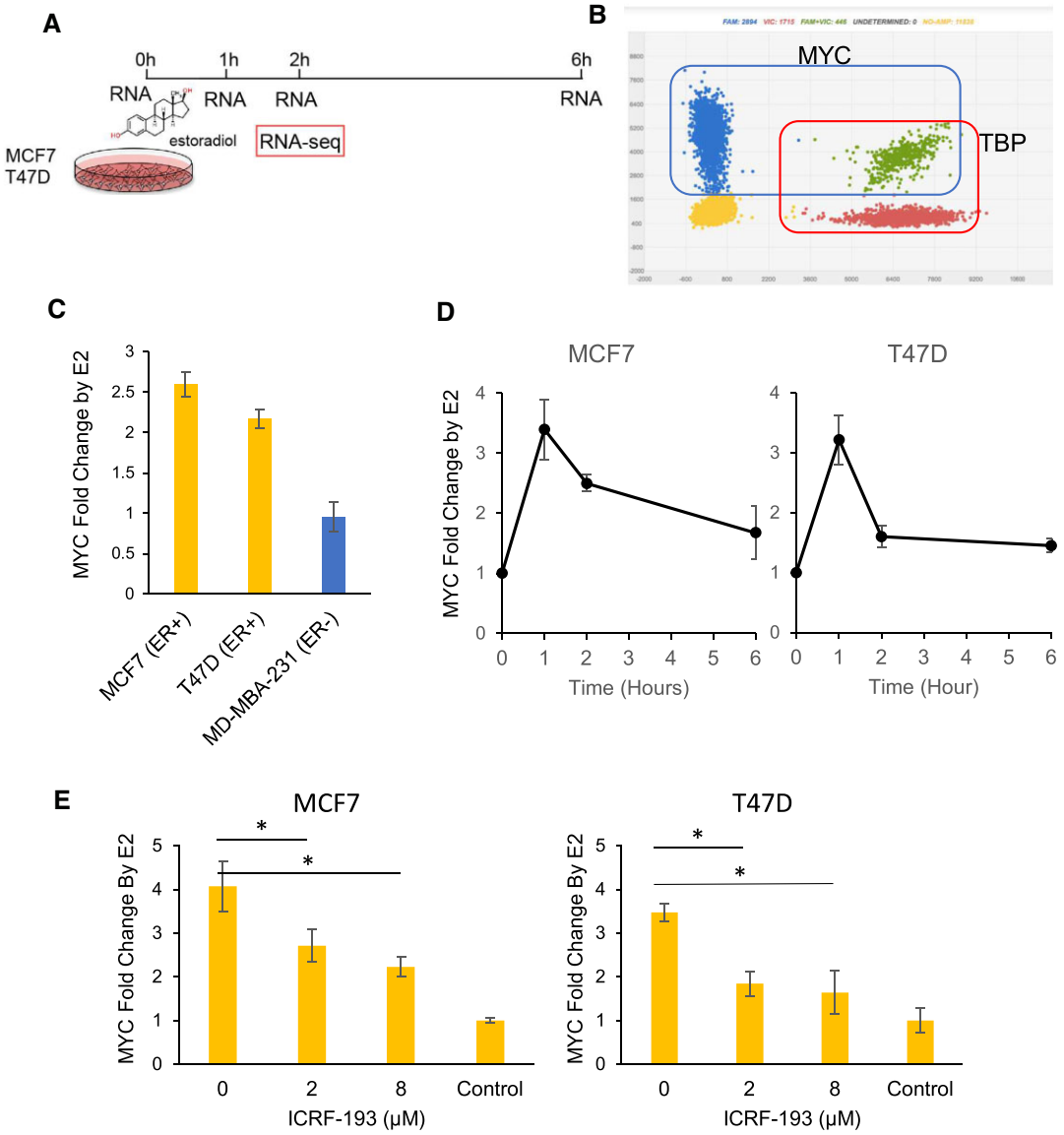


Figure 1. Estrogen and TOP2-dependent induction of *MYC*. (**A**) Experiment design. MCF7 and T47D cells were treated with 10 nM 17β-estradiol for 1, 2 and 6 h. RNA was extracted at each time point for 3D PCR analysis and RNA-seq at 2 h. (**B**) 3D PCR output from Analysis Suites. Blue dots represent wells only with the amplification of genes detected by FAM (*MYC*), red dots represent wells only with the amplification detected by VIC (*TBP* or *TFRC*) probe (internal control), yellow dots represent wells without amplification, green dots represent wells with the amplification of both probes. Normalized *MYC* expression was calculated by the ratio of FAM (*MYC*) to VIC (*TBP* or *TFRC*). (**C**) *MYC* induction in cells after 2 h of treatment with 10 nM E2 relative to cells with no E2 in two ER-positive breast cancer cell lines (MCF7 and T47D) and an ER-negative cell line (MD-MBA-231). Error bars represent the standard deviation from three independent measurements. (**D**) Dynamic responses of *MYC* induction. Cells were treated with 10 nM E2 for 0, 1, 2 and 6 h. *MYC* fold-induction peaked at 1 h and remained elevated until 6 h. Error bars represent the standard deviation from three independent measurements. (**E**) A TOP2 catalytic inhibitor ICRF-193 abrogated the induction of *MYC* mRNA. MCF7 and T47D cells were treated with increasing concentrations of ICRF-193 for 1 h then treated with 10 nM E2 for 2 h. *MYC* induction was progressively decreased in a concentration-dependent manner in both cell lines. Error bars represent the standard deviation from three independent measurements. * $P < 0.05$, ** $P < 0.01$.

ER-positive (ER+) cell lines, MCF7 and T47D, increased *MYC* expression after E2 treatment (2.59-fold and 2.17-fold, respectively). As expected, the ER-negative (ER–) cell line MDA-MB-231 did not show the induction (0.93-fold change) (Figure 1C). Further, we sought to understand the temporal regulation of *MYC* expression and measured *MYC* mRNA at 1, 2 and 6 h after the addition of 10 nM E2. In both the MCF7 and T47D cell lines, *MYC* expression was highest at 1 h (MCF7 2.89-fold, T47D 2.79-fold), followed by gradual decreases until 6 h (Figure 1D). A previous study determined that TOP2 mediates the induction of E2-responsive genes (3). Therefore, we next tested whether *MYC* induction is also TOP2-dependent. We used a TOP2 catalytic inhibitor, ICRF-193, which prevents TOP2 from cleaving a duplex DNA (12,13). We added ICRF-193 to cell culture media 1 h before adding 10 nM E2 and incubated cells for 2 h. We observed a dose-dependent inhibition of *MYC* induction in both MCF7 and T47D (Figure 1E), strongly suggesting that E2 stimulation of *MYC* expression in ER-positive cells is TOP2-dependent.

MYC transcriptional activation in TDP2 knockout cells

Since ICRF-193 inhibits enzyme turnover and DSB induction, DSBs could play a crucial role in the transcriptional activation of *MYC*. When left unrepaired, TOP2-induced DSBs activate transcription for a specific set of genes (14). Therefore, we next sought to determine whether unrepaired DSBs could further enhance the *MYC* induction by E2. To test the idea, we knocked out *TDP2* in ER-positive cell lines. When TOP2 fails to re-ligate DSBs and is aborted, TOP2cc is processed by a ubiquitin-proteasome pathway or, alternatively, a proteosome-independent pathway (36,37). TDP2 removes the remnant aborted TOP2 and creates a clean DSB for NHEJ machinery to repair (21). We used CRISPR-Cas9 to knockout TDP2 in both MCF7 (MCF7 TDP2KO-1) and T47D cell lines. In addition, we employed an independently created TDP2 knockout MCF7 cell line (MCF7 TDP2KO-2)(20). We confirmed the lack of TDP2 protein in knockout cells by western blotting (Figure 2A). We also confirmed the TDP2 deficiency functionally using the TOP2 poison, etoposide. Etoposide traps TOP2cc at DSBs, which would require TDP2 to repair. Both the MCF7 and T47D cells deficient in TDP2 were significantly more sensitive to etoposide than parental cell lines (Figure 2B).

We followed the dynamics of *MYC* induction using digital PCR, as shown in Figure 1C. We noticed that the highest level of induction at 1 h after the addition of E2 (10 nM) was similar between MCF7 and MCF7 TDP2KO-1 cells (Figure 2C). However, the *MYC* expression continued to be high in MCF7 TDP2KO cells, whereas it declined quickly in MCF7 cells. We further measured the induction of *MYC* at 2 h after the addition of 10 nM E2 and found the increased fold induction of *MYC* in all three TDP2 knockout cell lines relative to their parental TDP2 wild-type counterparts (Figure 2D). Because the expression was measured by total RNA of entire cell populations, these results suggest that *MYC* induction by E2 was enhanced in a significant proportion of the TDP2KO cell population. The delayed repair of TOP2-induced DSBs could promote the *MYC* response to E2 at 2 h, as previously described for early response genes to neuronal activity (14). Consistently, we observed that etoposide promotes the induction of *MYC* in MCF7 cells (Figure 2E, left). Inhibiting NHEJ could also promote the *MYC* response, as previously shown *in vivo* in DNA-PKcs deficient mice (24). To test this possibility, we inhibited NHEJ by DNA-PKcs inhibitor NU7441 (38). Indeed, *MYC* induction at 2 h after the addition of E2 was increased by NU7441 (5 μM) (Figure 2E, right).

Induction of MYC by estrogen *in vivo* in mice mammary glands

To address the relevance of the above observation *in vivo*, we examined *MYC* transcriptional response to E2 in mammary glands in mice. We injected E2 (300 μg/kg body weight) or PBS (control) intraperitoneally into three *TDP2* wild-type and $TDP2^{-/-}$ mice (21). Mice before injection (0 h) and 6 and 20 h after injection were examined for MYC protein expression in CK8/18-positive mammary epithelial cells by immunostaining (Figure 3A). We counted MYC-positive epithelial cells in 3–12 mammary ducts at each timepoint and scored the fraction of MYC-positive cells in each mammary duct (Figure 3B). MYC-positive cells increased in both wild-type and $TDP2^{-/-}$ mice at 6 h. MYC-positive cells become less frequent at 20 h, showing transient induction of MYC in mammary epithelial cells. MYC-positive cells in $TDP2^{-/-}$ mice at 20 h were significantly more frequent than in *TDP* wild-type mice ($P < 0.01$) at 20 h or in $TDP2^{-/-}$ mice without E2 at 0 hour ($P < 0.01$). The prolonged induction of MYC in $TDP2^{-/-}$ mice is reminiscent of TDP2KO MCF7 cells (Figure 2C) and is consistent with our observation *in vitro*.

MYC and MYC downstream targets induced by E2

MYC is a beta Helix-Loop-Helix transcription factor that, with a partner protein Max, binds to DNA at CACGTG E-Box sequence and modulates global gene expression (39). To examine the expression of MYC-regulated genes in MCF7 and MCF7 TDP2KO cells, we conducted RNA-seq analysis using total RNA before and 2 h after the addition of 10 nM E2 (bulk RNA-seq). We found that a large number of genes become differentially expressed in each MCF7 (2021) and MCF7 TDP2 KO cells (1034) (Supplementary Tables S1 and S2; fold change ≥1.5 and combined *P*-value < 0.05), and *MYC* was one of them in both cell types. Consistent with the results from digital PCR, Myc expression showed a 3.16-fold upregulation by E2 in MCF7 TDP2KO cells compared to a 1.94-fold in parental MCF7 cells (Figure 4A), confirming the prolonged response of *MYC* to estrogen when TDP2 is deficient. *MYC* transcriptional response was slightly more pronounced in MCF7 cells treated with etoposide (2.17-fold, Supplementary Table S3), suggesting the unrepaired TOP2-induced DSBs contributed to the enhanced response.

To assess whether *MYC* is activated after E2 treatment in MCF7 TDP2KO and parental MCF7 cells, we performed Gene Set Enrichment Analysis (GSEA) (32) with predefined MYC target genes. GSEA determines whether a defined set of genes shows statistically significant, concordant differences between two biological states. We looked for the significant changes in the expression of genes whose promoters have MYC binding sites within the 4 kb of the transcription start sites (MYC_Q2) and, therefore, potentially the direct MYC targets (Supplementary Figure S1, A, top). When treated with E2 for 2 h, the vast majority of genes in MYC_Q2 were non-randomly distributed and primarily found to be upregulated in both MCF7 and MCF7 TDP2KO cells. We further calculated the normalized enrichment score (NES) (Supplementary Figure S1, B). NES reports the degree of overrepresentation of genes at the extremes of all the ranked genes. We found that MYC_Q2 genes were slightly more clustered in highly upregulated genes in MCF7 TDP2KO cells than in MCF7 cells. We also examined the gene sets (Hallmark MYC target v2 and MYC_UP.V1_UP), the expression of which was increased by overexpressing MYC in human primary mammary epithelial cells (40). Genes in both sets were enriched in upregulated genes, with equivalent NES values between MCF7 and MCF7 TDP2KO. These results suggest that the upregulation of MYC by E2 modulated global transcriptional programs in both cell types.

MYC induction at the single-cell level

Bulk RNA-seq offers insights into the average RNA expression across an entire cell population. The observed upregulation of *MYC* in MCF7 TDP2KO cells could stem from either (i) a scenario in which only a subset of cells robustly upregulates *MYC*, creating a bimodal distribution that elevates the population average or (ii) a situation where *MYC* is in-

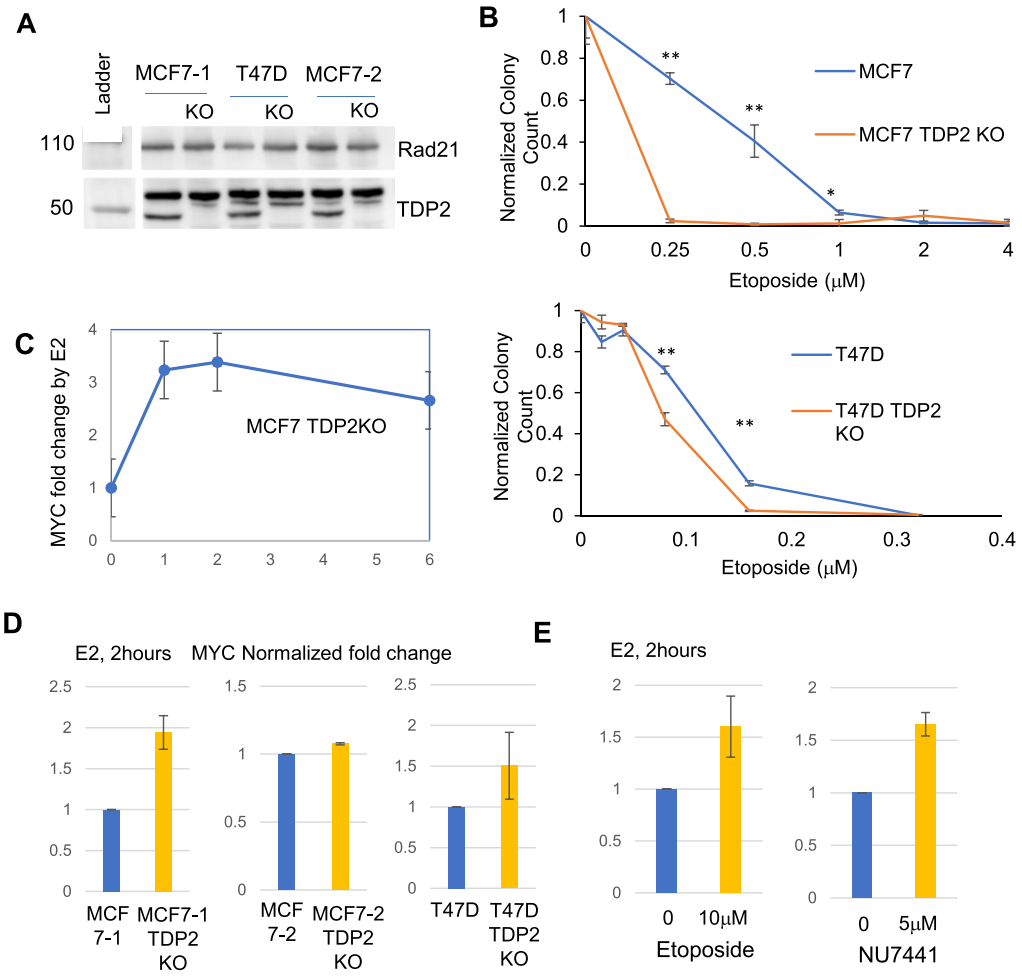


Figure 2. TDP2 suppresses the volatile and prolonged induction of *MYC* by estrogen. (**A**) Western blot showing MCF7-1, T47D and MCF7-2 cell lines and TDP2 KO cells for each cell line. (**B**) Etoposide sensitivity measured by colony formation assay for MCF7 (top) and T47D (bottom). Results from three independent experiments. Error bars represent the standard deviation from three independent measurements. * $P < 0.05$, ** $P < 0.01$. (**C**) *MYC* transcriptional responses persist in TDP2-deficient cells. MCF7 and MCF7 TDP2 KO cells were treated with 10 nM E2 for 1, 2 and 6 h. After a similar level of induction at 1 h, high level of response persisted in *MCF7* TDP2 KO cells persisted through 6-h time point. Error bars represent the standard deviation from two independent experiments. * $P < 0.05$, ** $P < 0.01$. (**D**) *MYC* transcriptional responses at 2 h. MCF7-1, MCF7-2 and T47D, and their TDP2 KO counterparts were treated with 10 nm E2 for 2 h. Note that two independent TDP2 KO clones of MCF7 cells were tested. Error bars represent the standard deviation from two independent experiments. (**E**) NU7441 and etoposide increased the *MYC* transcriptional response. MCF7 cells, cultured with or without NU7441 or etoposide were then treated with 10 nM E2 for 2 h. Error bars represent the standard deviation from two independent experiments.

duced more robustly in the vast majority of MCF7 TDP2KO cell population than in MCF7 cell population, resulting in the normal distribution. The first scenario suggests that, for those DSBs that are involved in *MYC* regulation after E2 treatments, TOP2-induced DSBs are primarily re-ligated by TOP2 itself, and TDP2 plays a role in repair only in a subset of cells with abortive TOP2ccs. The second scenario implies that TDP2 is programmed in the repair of TOP2-induced DSBs that regulate *MYC* expression. These possibilities can be distinguished by single-cell RNA-seq (sc RNA-seq) experiments.

Four groups of cells were subject to scRNA-seq: MCF7 parental cells, MCF7 parental cells treated with E2 for 2 h, MCF7 TDP2KO cells, and MCF7 TDP2KO cells treated with E2 for 2 h. Quality control and normalization of scRNA-seq data were executed using the Seurat pipeline (33). Following the exclusion of poor-quality cells and genes, the transcriptome data comprised 354 genes in E2-treated MCF7, 585 genes in untreated MCF7, 708 genes in E2-treated TDP2KO, and 399 genes in untreated TDP2KO. For the filtered and normalized data, we conducted two independent comparisons: E2-treated MCF7 versus untreated MCF7 and E2-treated TDP2KO versus untreated TDP2KO. Applying an adjusted *P*-value <0.05 and log fold change ≥0.1 (1.11-fold), we identified a total of 111 genes that were differentially expressed between untreated and E2-treated MCF7; 100 upregulated and 11 down-regulated genes, respectively (Supplementary Table S4). For TDP2-KO cells, E2 treatment resulted in gene expression changes in 327 genes: 138 up- and 189 down-regulated genes (Supplementary Table S5).

Consistent with the findings from digital PCR and bulk RNA-seq, *MYC* expression was induced by E2 treatment in both MCF7 TDP2KO and MCF7 cells, with $\log_2$-fold changes of 0.43 (1.54-fold) and 0.17 (1.19-fold), respectively (Figure 4A). Notably, the *MYC* induction by E2 was more pronounced in MCF7 TDP2KO cells than in MCF7. Normalized expression data revealed that MCF7 TDP2KO cell populations dis-

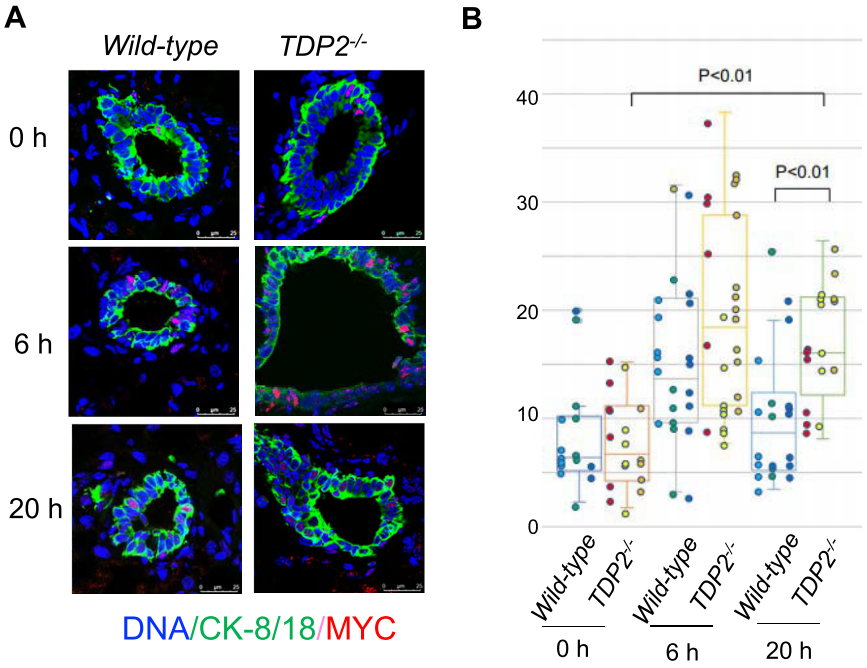


Figure 3. The induction of MYC protein by E2 *in vivo*. (**A**) Representative images of mammary ducts from wild-type (left) and TDP2KO (right) mice immunostained with MYC (red) and CK-8/18 (green) proteins before injection (0 h) and 6 and 20 h after injection of E2. (**B**) BOX plots showing the fraction of MYC-positive cells in (CK8/18-positive) mammary epithelial cells. Each plot represents the distribution of the fraction of MYC-positive epithelial cells in a mammary duct. MYC-positive epithelial cells were scored in 3–12 mammary ducts in each mouse at each time point.

played continuous distributions without distinct subpopulations and uniquely included cells with very high *MYC* expression (normalized expression > 1.5) after E2 treatment (Figure 4B). Histograms of raw read count were consistent with the skewed unimodal distribution across all four populations (Figure 4C). The normal distribution of expression values at a single cell level suggests that most cells in the population contributed equally to the mean increase, providing supports for the scenario (2).

We next examined whether the enhanced transcriptional response of the *MYC* in MCF7 TDP2KO cells is associated with cellular processes, such as cell cycle phases. Cell cycle phases are considered confounding factors in single-cell studies (41). To address this concern, we allocated individual cells to G1, S and G2/M phases using the cell cycle prediction tool based on single transcriptome data (42). We then examined whether cell cycle distribution might confound the observed difference in *MYC* transcriptional responses by collecting *MYC* expression profiles with and without E2 in each cell cycle phase (Figure 4D). Based on the cell cycle phase prediction, we found that MCF7 TDP2KO cells exhibited modestly higher enrichment in G1 cell populations than MCF7 cells; >60% of TDP2 KO cells were in G1, while 50% of MCF7 cells were in G1 (Supplementary Figure S2). However, the difference of *MYC* transcriptional induction by E2 was less pronounced between MCF7 and MCF7 TDP2KO cells in G1 cells (1.26-fold with $P < 0.0001$ in MCF7 versus 1.47-fold with $P < 0.0001$ in MCF7 TDP2KO, Wilcoxon rank-sum test) than in S (1.08-fold with $P = 0.0332$ versus 1.52-fold with $P < 0.0001$) and G2M cells (1.24-fold with $P = 0.0014$ versus 1.88-fold with $P < 0.0001$) (Figure 4C). Therefore, the difference in the distribution of cell cycles was not a major contributor to increased *MYC* induction in MCF7 TDP2KO cells. Instead, cell cycle profiling revealed that the increased response to E2 would likely stem from cells in S/G2M phases.

The dominance of upregulated genes by the short period of E2 treatment

The global transcriptional responses to E2 have been extensively studied in the MCF7 cell line, and both upregulated and downregulated genes have been documented. However, variations in the amounts of E2 and treatment durations across experiments, as well as the differences in genomic platforms (e.g. microarray, RNA-seq, and scRNA-seq), could influence the repertoire of E2-responsive genes, as previously discussed (43,44). To establish a reliable set of E2 (10 nM)—responsive genes at an early time point (2 h), we sought the intersections between different genomic platforms, bulk RNA-seq and scRNA-seq.

Bulk RNA-seq analysis identified a large number of differentially expressed genes, including 2021 for MCF7 and 1034 for MCF7 TDP2KO, between untreated and E2-treated cells. In contrast, the numbers of differentially expressed genes in our scRNA-seq data were modest: 111 genes for MCF7 and 327 genes for MCF7 TDP2KO (Figure 5A). Thirty-two genes were common between bulk and single-cell RNA-seq in MCF7, while 62 genes were shared in MCF7 TDP2KO cells. The majority of these common genes were upregulated genes in both cell lines (Figure 5B; Supplementary Tables S6 and S7). In MCF7, 28 genes were upregulated, and four genes were repressed, with one gene, GPRC5A, in opposite directions between bulk- and sc-RNA-seq. In MCF7 TDP2KO, 58 genes were upregulated, and four genes were repressed. There were 8 common differentially expressed genes between MCF7 and MCF7 TDP2KO, including *MYC*, *CCND1*, *NRIP*, *SH3BP5*, *IGFBP4*, *RET*, *NBPF1* and *CA12* (Figure 5C). Notably, *MYC*

A

Comparison	Bulk RNA-seq		Single-cell RNA-seq	
	log2-FC	P-value	log-FC	P-value
MCF7 with E2 vs without E2	0.9760	0.0194	0.1722	<0.0001
MCF7 TDP2KO with E2 vs without E2	1.6595	0.0001	0.4267	<0.0001

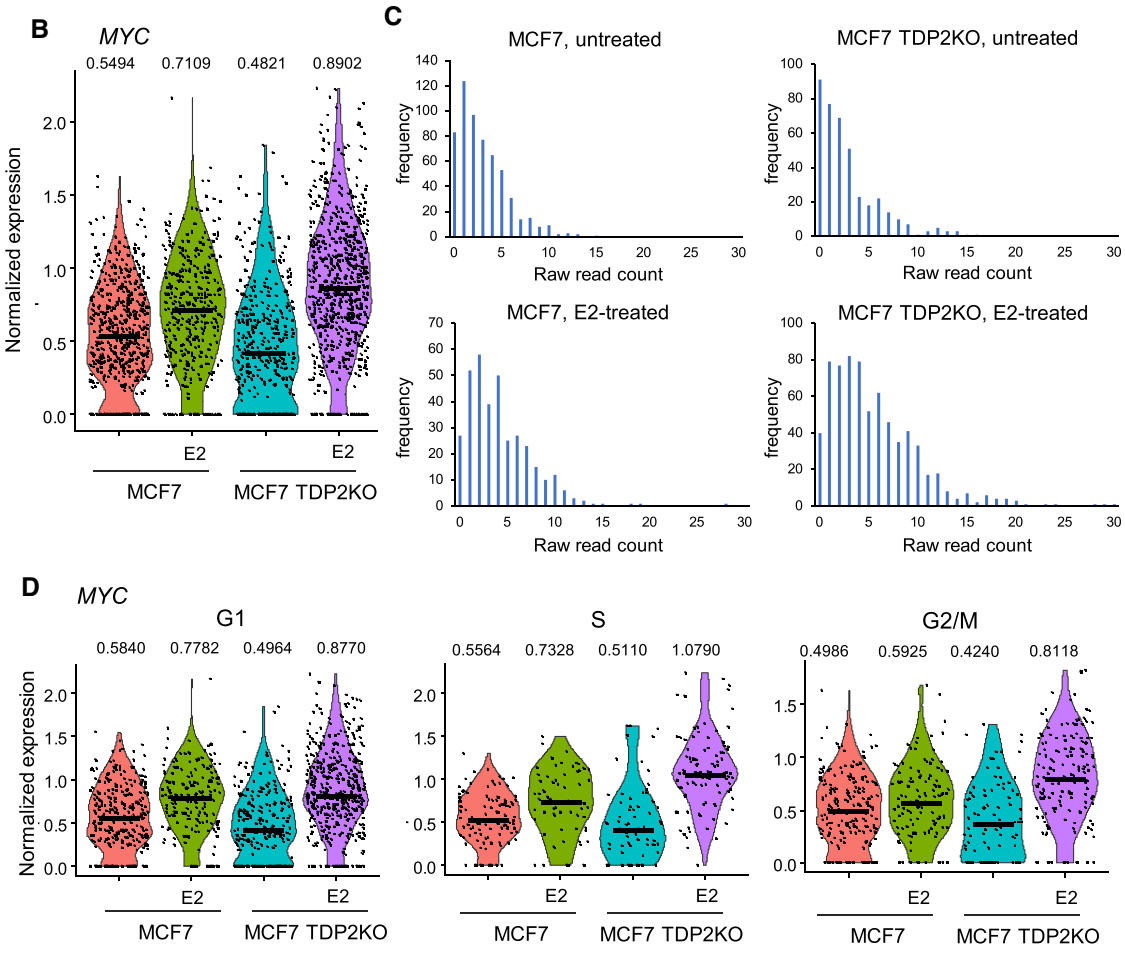
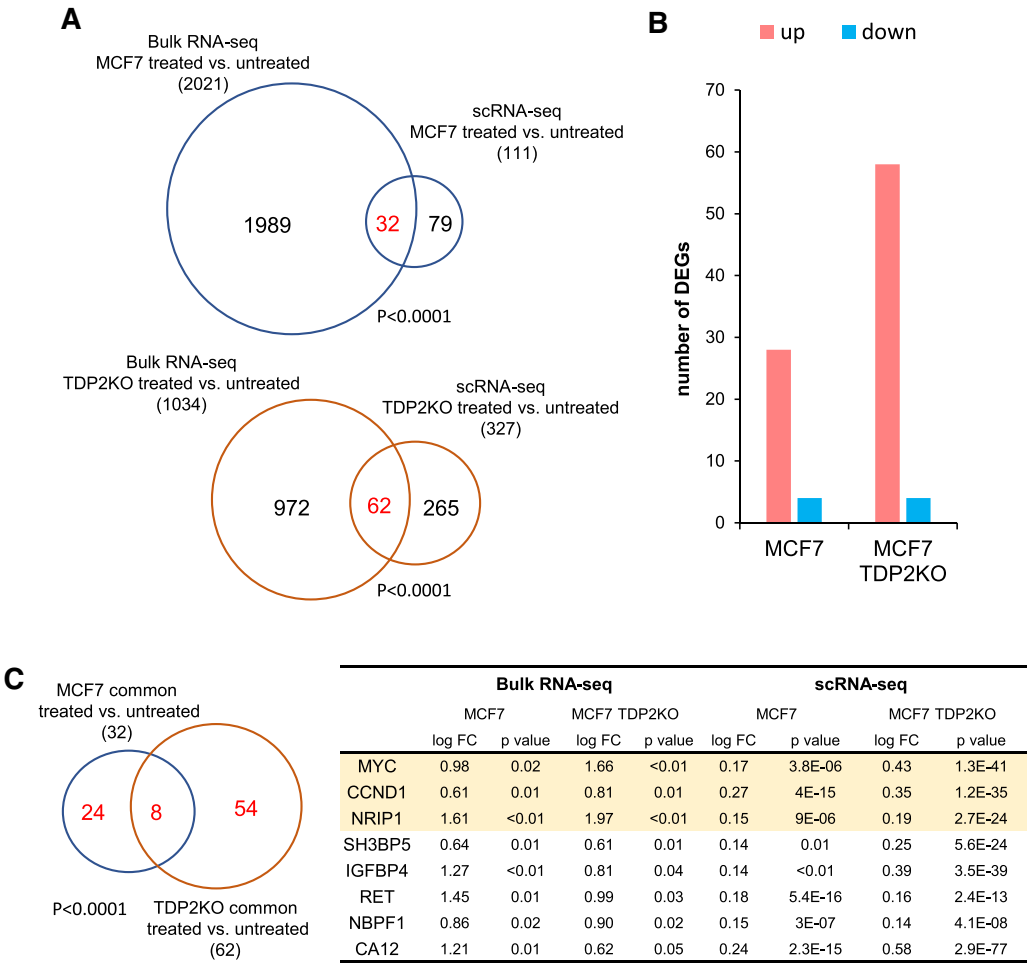


Figure 4. Bulk and Single-cell RNA sequencing analysis of *MYC* transcriptional response by E2. (**A**) *MYC* transcriptional response to E2 (2 h) in MCF7 and MCF7 TDP2KO cells, measured by bulk RNA-seq and single-cell RNA-seq. (**B**) Violin plots show normalized MYC expression distributions in MCF7 and MCF7 TDP2KO cells with or without E2 (2 h), obtained from single-cell RNA-seq data. (**C**) Histograms showing the skewed unimodal distribution of MYC expression in MCF7 and MCF7 TDP2KO cells with or without E2 (2 h), obtained from single-cell RNA-seq data. (**D**) Violin plots showing the normalized expression of *MYC* at the single-cell level in G1, S and G2/M cell populations.

and *CCND1* are well-established oncogenes in breast cancer (45–48). In summary, a reliable set of estrogen-responsive genes at an early time point (2h) were predominantly upregulated genes, with TDP2 potentially regulating responses to E2 for both *MYC* and *CCND1*.

Discussion

In this study, we assessed the role of TDP2, an enzyme involved in processing TOP2cc at the 5′ DSB ends, in the transcriptional responses to 17beta-estradiol (E2) in ER-positive breast cancer cells. Defects in the repair of TOP2-induced DSB have been shown to enhance the transcriptional response to external stimuli in a defined set of genes (14). Therefore, by observing E2 response in TDP2 wild-type and TDP2KO cell lines, we aimed to understand whether TDP2 inherently participates in the repair of TOP2-induced DSBs. We found that the *MYC* transcriptional response to E2 was upregulated and prolonged in cells deficient in TDP2. The upregulation was detected using techniques that measured the average expression level in the entire cell population, such as digital PCR and bulk mRNA-seq. Therefore, TDP2 seems involved in the transcriptional regulation of *MYC* in a significant fraction of the cell population. This observation was also seen at the protein level *in vivo* in breast epithelial cells in mice.



	Bulk RNA-seq				scRNA-seq			
	MCF7		MCF7 TDP2KO		MCF7		MCF7 TDP2KO	
	log FC	p value	log FC	p value	log FC	p value	log FC	p value
MYC	0.98	0.02	1.66	<0.01	0.17	3.8E-06	0.43	1.3E-41
CCND1	0.61	0.01	0.81	0.01	0.27	4E-15	0.35	1.2E-35
NRIP1	1.61	<0.01	1.97	<0.01	0.15	9E-06	0.19	2.7E-24
SH3BP5	0.64	0.01	0.61	0.01	0.14	0.01	0.25	5.6E-24
IGFBP4	1.27	<0.01	0.81	0.04	0.14	<0.01	0.39	3.5E-39
RET	1.45	0.01	0.99	0.03	0.18	5.4E-16	0.16	2.4E-13
NBPF1	0.86	0.02	0.90	0.02	0.15	3E-07	0.14	4.1E-08
CA12	1.21	0.01	0.62	0.05	0.24	2.3E-15	0.58	2.9E-77

Figure 5. Commonly upregulated genes between bulk- and single-cell RNA-seq. (**A**) Venn diagrams showing the differentially expressed genes by E2 treatment in MCF7 and MCF7 TDP2KO cells. 32 genes in MCF7 and 62 genes in MCF7 TDP2KO are common between the two platforms. The significances of overlap were calculated using Fisher's exact test. (**B**) Bar plots showing the number of common DEGs by E2 (2 h). 28 and 4 genes were commonly up- and down-regulated genes in E2 treated MCF7 versus untreated MCF7, respectively. 58 and 4 genes were commonly up- and downregulated in E2 treated MCF7 TDP2KO versus untreated MCF7 TDP2KO, respectively. (**C**) Venn diagram of 32 and 62 common DEGs from bulk and single-cell RNA-seq data and table of 8 overlapping genes. The significance of overlap was calculated using Fisher's exact test.

Using single-cell RNA-seq (scRNA-seq), we validated that the observed upregulation of *MYC* was a result of widespread upregulation across the cell population, as indicated by the unimodal distribution of *MYC* mRNA in MCF7 TDP2KO cells (Figure 4C). This unimodality contrasts with the bimodal expression pattern observed for many genes in a previous scRNA-seq study of E2-stimulated transcriptional responses in MCF7 cells (49). However, the numbers of cells captured for transcriptomic analysis are smaller in the study (87 MCF7 cells) relative to our study (354 cells from MCF7 with E2, and 585 cells from MCF7 without E2) and may confound the difference. With the scRNA-seq results along with the regulation of *MYC* expression response to E2 by TOP2 (Figure 1E), we propose that TDP2 is programmed to repair TOP2-induced DSBs and regulates *MYC* expression when stimulated by E2. Unrepaired DSBs could suppress the transcription of nearby genes (50–52). However, unrepaired DSBs far from the target genes, such as enhancer elements, would not directly disrupt the elongating mRNA of target genes. This scenario potentially applies to *MYC*, given the dispersed enhancer elements within gene deserts surrounding *MYC* (53,54). In particular, an enhancer located at 135kb distal to *MYC*, known as Myc+ 135 kb enhancer, contains an ER binding site and responds to estrogen treatments (55). The estrogen-induced expression of enhancer RNA and the physical interaction with *MYC* promoter were greatly enhanced when the repair of TOP2-induced DSBs was inhibited by an ATM inhibitor. However, ATM and TDP2 function independently to repair etoposide-induced DSBs (56), and the spatial and temporal regulations of these two repair pathways remain unclear.

scRNA-seq also allowed us to investigate the responses to E2 in each cell cycle phase. We found that TDP2 KO cells responded to E2 more robustly in the S and G2/M phases than in the G1 phase. The removal of abortive TOP2ccs involved at least two sets of proteins: TDP2 and Mre11/BRCA1. How these pathways coordinate their tasks remains to be determined. BRCA1 facilitates the nuclease activity of Mre11 to remove abortive TOP2ccs (20). This activity was shown to be active in G1 cells. Given the robust upregulation of *MYC* in S and G2/M cells in our study, we speculate that these pathways repair DSBs with abortive TOP2cc in different cell cycle phases. A recent study showed that TDP2 binds to K63

and K27 poly-ubiquitinated cellular proteins, and this binding stimulates TDP2-catalyzed removal of TOP2 from DNA (57). While the identity of the K27 and/or K63 poly-ubiquitinated proteins bound to TDP2 remains unknown, it is conceivable that such proteins could regulate TDP2 activity in a cell-cycle-dependent manner.

We cataloged E2-responsive genes regulated by E2 using bulk and scRNA-seq. The list of common genes between the two platforms was strongly restricted by scDNA-seq data, as only small fractions of differentially expressed genes detected by bulk RNA-seq were also called by scRNA-seq. This is likely because scRNA-seq data are, in general, very heterogeneous with many zero values (58). Nevertheless, the genes that were consistently differentially expressed across platforms displayed a notable bias towards upregulation, notably including oncogenes such as *MYC* and *CCND1*, which are frequently overexpressed in ER-positive breast tumors. Estrogen induces mammary ductal hyperplasia through MYC in mice deficient in NHEJ (24), which is downstream of TDP2-mediated repair of TOP2cc. CCND1 is more frequently upregulated in ER-positive tumors than other subtypes and promotes cell cycle progression (59,60). Given its involvement in regulating these oncogenes, TDP2 emerges as a potential suppressor of the most common subtype of breast tumors.

Data availability

Bulk and single-cell RNA-seq data are available in Gene Expression Omnibus (accession # GSE225423).

Supplementary data

Supplementary Data are available at NAR Cancer Online.

Acknowledgements

We thank Cedars-Sinai Cancer, Applied Genomics, Computation and Translational Core and Flow Core for technical support. This work is supported by NIHR01CA149385 and R01GM141698, Cedars-Sinai Cancer Biology Development Fund and Institutional Support (to H.T.), Cedars-Sinai CTSI Clinical Scholar Award (to N.M.), Leukemia & Lymphoma Society (to N.J.) and Margie and Robert E. Petersen Foundation, The Fashion Footwear Charitable Foundation of New York, Inc., Avon Foundation and Associates for Breast and Prostate Cancer Studies (to A.G.), National Key R&D Program of China 2022YFA1302800, National Natural Science Foundation- General project 82072947, National Natural Science Foundation- International Senior Scientists 32250710138, and Pearl River Talent Plan to introduce high-level talents 802-014967 (to S.T.).

Author contributions: Conceptualization, N.M., M.K., S.T., S.Y.Y. and H.T.; Data collection, N.M., N.J., R.A.M., J.A., R.S. and H.T.; Data analysis and interpretation, N.M., M.K., R.A.M., J. A., S.Y.Y. and H.T.; Materials, F.C-L., S.Y., X.C.; Writing the manuscript, N.M., M.K., S.T., S.Y.Y. and H.T.; Funding Acquisition, N.M., A.E.G. and H.T.

Funding

National Cancer Institute [CA149385].

Conflict of interest statement

None declared.

References

1. Katzenellenbogen,B.S. and Katzenellenbogen,J.A. (2000) Estrogen receptor transcription and transactivation: estrogen receptor alpha and estrogen receptor beta: regulation by selective estrogen receptor modulators and importance in breast cancer. *Breast Cancer Res.*, **2**, 335–344.
2. Fuentes,N. and Silveyra,P. (2019) Estrogen receptor signaling mechanisms. *Adv. Protein Chem. Struct. Biol.*, **116**, 135–170.
3. Ju,B.G., Lunyak,V.V., Perissi,V., Garcia-Bassets,I., Rose,D.W., Glass,C.K. and Rosenfeld,M.G. (2006) A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science*, **312**, 1798–1802.
4. Hah,N., Danko,C.G., Core,L., Waterfall,J.J., Siepel,A., Lis,J.T. and Kraus,W.L. (2011) A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell*, **145**, 622–634.
5. Pommier,Y., Sun,Y., Huang,S.N. and Nitiss,J.L. (2016) Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat. Rev. Mol. Cell Biol.*, **17**, 703–721.
6. Morimoto,S., Tsuda,M., Bunch,H., Sasanuma,H., Austin,C. and Takeda,S. (2019) Type II DNA topoisomerases cause spontaneous double-strand breaks in genomic DNA. *Genes (Basel)*, **10**, 868.
7. Roca,J. and Wang,J.C. (1992) The capture of a DNA double helix by an ATP-dependent protein clamp: a key step in DNA transport by type II DNA topoisomerases. *Cell*, **71**, 833–840.
8. Tubbs,A. and Nussenzweig,A. (2017) Endogenous DNA damage as a source of genomic instability in cancer. *Cell*, **168**, 644–656.
9. King,I.F., Yandava,C.N., Mabb,A.M., Hsiao,J.S., Huang,H.-S., Pearson,B.L., Calabrese,J.M., Starmer,J., Parker,J.S., Magnuson,T., *et al.* (2013) Topoisomerases facilitate transcription of long genes linked to autism. *Nature*, **501**, 58–62.
10. Bunch,H., Lawney,B.P., Lin,Y.-F., Asaithamby,A., Murshid,A., Wang,Y.E., Chen,B.P.C. and Calderwood,S.K. (2015) Transcriptional elongation requires DNA break-induced signalling. *Nat. Commun.*, **6**, 10191.
11. Nitiss,J.L. (2009) Targeting DNA topoisomerase II in cancer chemotherapy. *Nat. Rev. Cancer*, **9**, 338–350.
12. Roca,J., Ishida,R., Berger,J.M., Andoh,T. and Wang,J.C. (1994) Antitumor bisdioxopiperazines inhibit yeast DNA topoisomerase II by trapping the enzyme in the form of a closed protein clamp. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 1781–1785.
13. Huang,K.-C., Gao,H., Yamasaki,E.F., Grabowski,D.R., Liu,S., Shen,L.L., Chan,K.K., Ganapathi,R. and Snapka,R.M. (2001) Topoisomerase II poisoning by ICRF-193*. *J. Biol. Chem.*, **276**, 44488–44494.
14. Madabhushi,R., Gao,F., Pfenning,A.R., Pan,L., Yamakawa,S., Seo,J., Rueda,R., Phan,T.X., Yamakawa,H., Pao,P.C., *et al.* (2015) Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell*, **161**, 1592–1605.
15. Calderwood,S.K. (2016) A critical role for topoisomerase IIb and DNA double strand breaks in transcription. *Transcription*, **7**, 75–83.
16. Nitiss,J.L. (2009) DNA topoisomerase II and its growing repertoire of biological functions. *Nat. Rev. Cancer*, **9**, 327–337.
17. Manville,C.M., Smith,K., Sondka,Z., Rance,H., Cockell,S., Cowell,I.G., Lee,K.C., Morris,N.J., Padget,K., Jackson,G.H., *et al.* (2015) Genome-wide ChIP-seq analysis of human TOP2B occupancy in MCF7 breast cancer epithelial cells. *Biol Open*, **4**, 1436–1447.
18. Ledesma,F.C., El Khamisy,S.F., Zuma,M.C., Osborn,K. and Caldecott,K.W. (2009) A human 5′-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature*, **461**, 674–678.

19. Hoa,N.N., Shimizu,T., Zhou,Z.W., Wang,Z.-Q., Deshpande,R.A., Paull,T.T., Akter,S., Tsuda,M., Furuta,R., Tsutsui,K, *et al.* (2016) Mre11 Is essential for the removal of lethal topoisomerase 2 covalent cleavage complexes. *Mol. Cell*, **64**, 580–592.
20. Sasanuma,H., Tsuda,M., Morimoto,S., Saha,L.K., Rahman,M.M., Kiyooka,Y., Fujiike,H., Cherniack,A.D., Itou,J., Callen Moreu,E., *et al.* (2018) BRCA1 ensures genome integrity by eliminating estrogen-induced pathological topoisomerase II–DNA complexes. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, E10642–E10651.
21. Gómez-Herreros,F., Romero-Granados,R., Zeng,Z., Alvarez-Quilón,A., Quintero,C., Ju,L., Umans,L., Vermeire,L., Huylebroeck,D., Caldecott,K.W., *et al.* (2013) TDP2-dependent non-homologous end-joining protects against topoisomerase II-induced DNA breaks and genome instability in cells and in vivo. *PLoS Genet.*, **9**, e1003226.
22. Aparicio,T., Baer,R., Gottesman,M. and Gautier,J. (2016) MRN, CtIP, and BRCA1 mediate repair of topoisomerase II–DNA adducts. *J. Cell Biol.*, **212**, 399–408.
23. D'Alessandro,G., Morales-Juarez,D.A., Richards,S.L., Nitiss,K.C., Serrano-Benitez,A., Wang,J., Thomas,J.C., Gupta,V., Voigt,A., Belotserkovskaya,R., *et al.* (2023) RAD54L2 counters TOP2-DNA adducts to promote genome stability. *Sci. Adv.*, **9**, eadl2108.
24. Itou,J., Takahashi,R., Sasanuma,H., Tsuda,M., Morimoto,S., Matsumoto,Y., Ishii,T., Sato,F., Takeda,S. and Toi,M. (2020) Estrogen induces mammary ductal dysplasia via the upregulation of myc expression in a DNA-repair-deficient condition. *iScience*, **23**, 100821–100821.
25. Gan,W., Guan,Z., Liu,J., Gui,T., Shen,K., Manley,J.L. and Li,X. (2011) R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.*, **25**, 2041–2056.
26. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2012) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
27. Ewels,P., Magnusson,M., Lundin,S. and Käller,M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, **32**, 3047–3048.
28. Johnson,W.E., Li,C. and Rabinovic,A. (2006) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, **8**, 118–127.
29. Robinson,M.D. and Oshlack,A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.*, **11**, R25.
30. You,S., Yoo,S.-A., Choi,S., Kim,J.-Y., Park,S.-J., Ji,J.D., Kim,T.-H., Kim,K.-J., Cho,C.-S., Hwang,D., *et al.* (2014) Identification of key regulators for the migration and invasion of rheumatoid synoviocytes through a systems approach. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 550–555.
31. Storey,J.D. and Tibshirani,R. (2003) Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 9440–9445.
32. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S., *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 15545–15550.
33. Stuart,T., Butler,A., Hoffman,P., Hafemeister,C., Papalexi,E., Mauck,W.M. 3rd, Hao,Y., Stoeckius,M., Smibert,P. and Satija,R. (2019) Comprehensive integration of single-cell data. *Cell*, **177**, 1888–1902.
34. Al Mahmud,M.R., Ishii,K., Bernal-Lozano,C., Delgado-Sainz,I., Toi,M., Akamatsu,S., Fukumoto,M., Watanabe,M., Takeda,S., Cortés-Ledesma,F., *et al.* (2020) TDP2 suppresses genomic instability induced by androgens in the epithelial cells of prostate glands. *Genes Cells*, **25**, 450–465.
35. Shang,Y., Hu,X., DiRenzo,J., Lazar,M.A. and Brown,M. (2000) Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell*, **103**, 843–852.
36. Mao,Y., Desai,S.D., Ting,C.Y., Hwang,J. and Liu,L.F. (2001) 26 S proteasome-mediated degradation of topoisomerase II cleavable complexes. *J. Biol. Chem.*, **276**, 40652–40658.
37. Schellenberg,M.J., Lieberman,J.A., Herrero-Ruiz,A., Butler,L.R., Williams,J.G., Munoz-Cabello,A.M., Mueller,G.A., London,R.E., Cortes-Ledesma,F. and Williams,R.S. (2017) ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein cross-links. *Science*, **357**, 1412–1416.
38. Zhao,Y., Thomas,H.D., Batey,M.A., Cowell,I.G., Richardson,C.J., Griffin,R.J., Calvert,A.H., Newell,D.R., Smith,G.C. and Curtin,N.J. (2006) Preclinical evaluation of a potent novel DNA-dependent protein kinase inhibitor NU7441. *Cancer Res.*, **66**, 5354–5362.
39. Grandori,C., Cowley,S.M., James,L.P. and Eisenman,R.N. (2000) The Myc/Max/Mad network and the transcriptional control of cell behavior. *Annu. Rev. Cell Dev. Biol.*, **16**, 653–699.
40. Bild,A.H., Yao,G., Chang,J.T., Wang,Q., Potti,A., Chasse,D., Joshi,M.-B., Harpole,D., Lancaster,J.M., Berchuck,A., *et al.* (2006) Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature*, **439**, 353–357.
41. Hsiao,C.J., Tung,P., Blischak,J.D., Burnett,J.E., Barr,K.A., Dey,K.K., Stephens,M. and Gilad,Y. (2020) Characterizing and inferring quantitative cell cycle phase in single-cell RNA-seq data analysis. *Genome Res.*, **30**, 611–621.
42. Nestorowa,S., Hamey,F.K., Pijuan Sala,B., Diamanti,E., Shepherd,M., Laurenti,E., Wilson,N.K., Kent,D.G. and Gottgens,B. (2016) A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood*, **128**, e20–e31.
43. Cheung,E. and Kraus,W.L. (2010) Genomic analyses of hormone signaling and gene regulation. *Annu. Rev. Physiol.*, **72**, 191–218.
44. Jagannathan,V. and Robinson-Rechavi,M. (2011) Meta-analysis of estrogen response in MCF-7 distinguishes early target genes involved in signaling and cell proliferation from later target genes involved in cell cycle and DNA repair. *BMC Syst. Biol.*, **5**, 138.
45. Escot,C., Theillet,C., Lidereau,R., Spyratos,F., Champeme,M.H., Gest,J. and Callahan,R. (1986) Genetic alteration of the c-myc protooncogene (MYC) in human primary breast carcinomas. *Proc. Natl. Acad. Sci. U.S.A.*, **83**, 4834–4838.
46. Sicinski,P., Donaher,J.L., Parker,S.B., Li,T., Fazeli,A., Gardner,H., Haslam,S.Z., Bronson,R.T., Elledge,S.J. and Weinberg,R.A. (1995) Cyclin D1 provides a link between development and oncogenesis in the retina and breast. *Cell*, **82**, 621–630.
47. Weinstat-Saslow,D., Merino,M.J., Manrow,R.E., Lawrence,J.A., Bluth,R.F., Wittenbel,K.D., Simpson,J.F., Page,D.L. and Steeg,P.S. (1995) Overexpression of cyclin D mRNA distinguishes invasive and in situ breast carcinomas from non-malignant lesions. *Nat. Med.*, **1**, 1257–1260.
48. D'Cruz,C.M., Gunther,E.J., Boxer,R.B., Hartman,J.L., Sintasath,L., Moody,S.E., Cox,J.D., Ha,S.I., Belka,G.K., Golant,A., *et al.* (2001) c-MYC induces mammary tumorigenesis by means of a preferred pathway involving spontaneous Kras2 mutations. *Nat. Med.*, **7**, 235–239.
49. Zhu,D., Zhao,Z., Cui,G., Chang,S., Hu,L., See,Y.X., Lim,M.G.L., Guo,D., Chen,X., Poudel,B., *et al.* (2018) Single-cell transcriptome analysis reveals estrogen signaling coordinately augments one-carbon, polyamine, and purine synthesis in breast cancer. *Cell Rep.*, **25**, 2285–2298.
50. Harding,S.M., Boiarsky,J.A. and Greenberg,R.A. (2015) ATM dependent silencing links nucleolar chromatin reorganization to DNA damage recognition. *Cell Rep.*, **13**, 251–259.
51. Larsen,D.H., Hari,F., Clapperton,J.A., Gwerder,M., Gutsche,K., Altmeyer,M., Jungmichel,S., Toledo,L.I., Fink,D., Rask,M.B., *et al.* (2014) The NBS1-treacle complex controls ribosomal RNA transcription in response to DNA damage. *Nat. Cell Biol.*, **16**, 792–803.
52. Shanbhag,N.M., Rafalska-Metcalf,I.U., Balane-Bolivar,C., Janicki,S.M. and Greenberg,R.A. (2010) ATM-dependent chromatin changes silence transcription in cis to DNA double-strand breaks. *Cell*, **141**, 970–981.

53. Bahr,C., von Paleske,L., Uslu,V.V., Remeseiro,S., Takayama,N., Ng,S.W., Murison,A., Langenfeld,K., Petretich,M., Scognamiglio,R., *et al.* (2018) A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. *Nature*, **553**, 515–520.
54. Ahmadiyeh,N., Pomerantz,M.M., Grisanzio,C., Herman,P., Jia,L., Almendro,V., He,H.H., Brown,M., Liu,X.S., Davis,M., *et al.* (2010) 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 9742–9746.
55. Najnin,R.A., Al Mahmud,M.R., Rahman,M.M., Takeda,S., Sasanuma,H., Tanaka,H., Murakawa,Y., Shimizu,N., Akter,S., Takagi,M., *et al.* (2023) ATM suppresses c-myc overexpression in the mammary epithelium in response to estrogen. *Cell Rep.*, **42**, 111909.
56. Álvarez-Quilón,A., Serrano-Benítez,A., Ariel Lieberman,J., Quintero,C., Sánchez-Gutiérrez,D., Escudero,L.M. and Cortés-Ledesma,F. (2014) ATM specifically mediates repair of double-strand breaks with blocked DNA ends. *Nat. Commun.*, **5**, 3347.
57. Schellenberg,M.J., Appel,C.D., Riccio,A.A., Butler,L.R., Krahn,J.M., Liebermann,J.A., Cortes-Ledesma,F. and Williams,R.S. (2020) Ubiquitin stimulated reversal of topoisomerase 2 DNA-protein crosslinks by TDP2. *Nucleic Acids Res.*, **48**, 6310–6325.
58. Risso,D., Perraudeau,F., Gribkova,S., Dudoit,S. and Vert,J.-P. (2018) A general and flexible method for signal extraction from single-cell RNA-seq data. *Nat. Commun.*, **9**, 284.
59. Roy,P.G., Pratt,N., Purdie,C.A., Baker,L., Ashfield,A., Quinlan,P. and Thompson,A.M. (2010) High CCND1 amplification identifies a group of poor prognosis women with estrogen receptor positive breast cancer. *Int. J. Cancer*, **127**, 355–360.
60. Ortiz,A.B., Garcia,D., Vicente,Y., Palka,M., Bellas,C. and Martin,P. (2017) Prognostic significance of cyclin D1 protein expression and gene amplification in invasive breast carcinoma. *PLoS One*, **12**, e0188068.

Received: April 28, 2023. **Revised:** February 19, 2024. **Editorial Decision:** March 6, 2024. **Accepted:** March 14, 2024