

# Topoisomerases facilitate transcription of long genes linked to autism

Ian F. King<sup>1</sup>, Chandri N. Yandava<sup>2</sup>, Angela M. Mabb<sup>1</sup>, Jack S. Hsiao<sup>3</sup>, Hsien-Sung Huang<sup>1†</sup>, Brandon L. Pearson<sup>1</sup>, J. Mauro Calabrese<sup>4</sup>, Joshua Starmer<sup>4</sup>, Joel S. Parker<sup>4,5</sup>, Terry Magnuson<sup>4,5</sup>, Stormy J. Chamberlain<sup>3</sup>, Benjamin D. Philpot<sup>1,2,6</sup> & Mark J. Zylka<sup>1,2,6</sup>

Topoisomerases are expressed throughout the developing and adult brain and are mutated in some individuals with autism spectrum disorder (ASD). However, how topoisomerases are mechanistically connected to ASD is unknown. Here we find that topotecan, a topoisomerase I (TOP1) inhibitor, dose-dependently reduces the expression of extremely long genes in mouse and human neurons, including nearly all genes that are longer than 200 kilobases. Expression of long genes is also reduced after knockdown of *Top1* or *Top2b* in neurons, highlighting that both enzymes are required for full expression of long genes. By mapping RNA polymerase II density genome-wide in neurons, we found that this length-dependent effect on gene expression was due to impaired transcription elongation. Interestingly, many high-confidence ASD candidate genes are exceptionally long and were reduced in expression after TOP1 inhibition. Our findings suggest that chemicals and genetic mutations that impair topoisomerases could commonly contribute to ASD and other neurodevelopmental disorders.

Autism is a neurodevelopmental disorder with symptoms that include repetitive behaviours and deficits in social interactions. Hundreds of genes are now associated with ASD<sup>1,2</sup>, suggesting that there are diverse genetic risk factors for autism. Environmental factors, including chemicals that are ingested during critical periods of brain development<sup>3</sup>, can also increase autism risk. Many ASD candidate genes regulate synapse function<sup>4–6</sup>; however, whether there are additional mechanisms that unite ASD patients or expression of ASD genes is unclear.

Recently, we found that topoisomerase inhibitors can transcriptionally un-silence the paternal allele of *Ube3a* in mouse cortical neurons<sup>7</sup>. *Ube3a* is located adjacent to a cluster of imprinted genes, is normally expressed only from the maternal allele in neurons, and regulates synaptic function<sup>8</sup>. In humans, *UBE3A* is associated with two distinct neurodevelopmental disorders. Specifically, deletion or mutation of maternal *UBE3A* causes Angelman syndrome, whereas duplication of the chromosomal region containing maternal *UBE3A* is frequently detected in individuals with autism<sup>9,10</sup>.

Intriguingly, mutations in topoisomerases were recently identified in some individuals with ASD<sup>11,12</sup>. However, precisely how topoisomerases regulate the expression of *Ube3a* and possibly other genes associated with autism is unknown. Topoisomerases, including *Top1* and *Top2b*, are expressed throughout the developing and adult mouse brain<sup>13,14</sup>. Topoisomerases are integral to gene expression, as they resolve DNA supercoiling that is generated during transcription<sup>15–18</sup>. Here we sought to determine whether topoisomerases preferentially regulate the expression of additional imprinted genes in neurons, or whether they have broader effects on gene expression. Using genome-wide approaches, we unexpectedly found that topoisomerases facilitate the expression of long genes, including numerous long genes associated with synaptic function and ASD. In addition, our study uncovers a transcriptional mechanism that is particularly important for maintaining the expression of numerous ASD genes at normal levels.

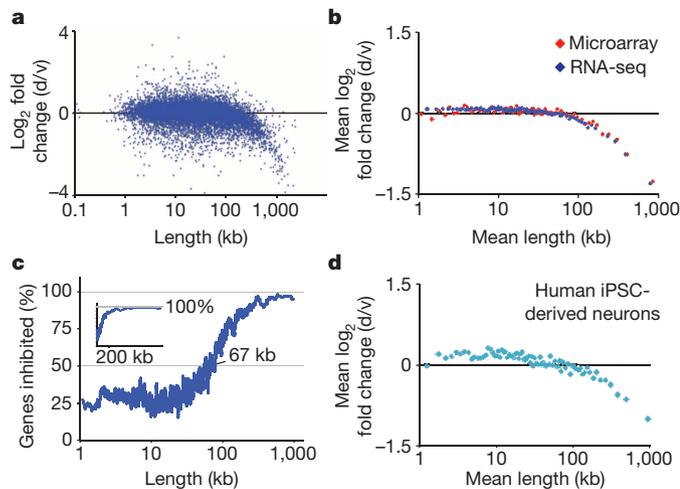
## Gene length effects

To determine whether topotecan, a TOP1 inhibitor, altered the expression of imprinted genes, we treated cultured cortical neurons from C57BL/6J (B6) × CASTEi/J (CAST) F<sub>1</sub> hybrid mice with vehicle or 300 nM topotecan, then used high-throughput transcriptome sequencing (RNA-seq) to survey changes in gene expression genome-wide. Single nucleotide polymorphisms were used to determine the parent-of-origin of sequence reads for autosomal genes<sup>19</sup>. We defined imprinted genes as those displaying statistically significant parent-of-origin expression bias in reciprocal B6 × CAST crosses (Fisher's exact test,  $P < 0.05$  after adjustment for multiple comparisons). We found that cortical neurons expressed 49 known autosomal imprinted genes (Extended Data Table 1), but *Ube3a* was the only imprinted gene that showed a significant change in parental allele bias in reciprocal crosses upon topotecan treatment (Fisher's exact test,  $P < 0.05$  after correction; Extended Data Table 1). Indeed, topotecan increased expression of the paternal allele of *Ube3a*, driving *Ube3a* levels significantly above wild-type levels (Extended Data Fig. 1a, b).

As we found previously<sup>7</sup>, topotecan reduced the expression of an extremely long (>1 megabase), paternally expressed antisense transcript that overlaps *Ube3a* (*Ube3a-ATS*; Extended Data Fig. 1a, b). The *Ube3a-ATS* transcript is required for paternal *Ube3a* silencing<sup>20,21</sup>. Other imprinted genes in the same genomic region as *Ube3a* did not show changes in allelic expression following topotecan treatment (Extended Data Fig. 1b and Extended Data Table 1). Notably, topotecan also reduced the expression of *UBE3A-ATS* and increased the expression of *UBE3A* in induced pluripotent stem cell (iPSC)-derived neurons from an Angelman syndrome patient (Extended Data Fig. 1c). Topotecan thus had similar transcriptional effects at the *Ube3a* locus in mouse and human neurons.

Because *Ube3a-ATS* is extremely long and was strongly downregulated, we hypothesized that topotecan might reduce the expression of

<sup>1</sup>Department of Cell Biology and Physiology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>2</sup>Carolina Institute for Developmental Disabilities, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>3</sup>Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06032, USA. <sup>4</sup>Department of Genetics, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>5</sup>Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. <sup>6</sup>UNC Neuroscience Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. †Present address: Graduate Institute of Brain and Mind Sciences, College of Medicine, National Taiwan University, Taipei 10051, Taiwan.



**Figure 1 | TOP1 inhibition reduces expression of long genes in neurons.** **a**, Mouse cortical neurons treated with vehicle (v) or 300 nM topotecan (drug; d) for 3 days ( $n = 5$  biological replicates). RNA-seq gene expression versus gene length. **b**, Mean expression change in bins of 200 genes by length. **c**, Percentage of genes that were reduced in expression by topotecan; plotted as a sliding window of 100 genes by length, RNA-seq data (log scale). Inset, same data on linear scale. **d**, iPSC-derived human neurons treated with 1  $\mu$ M topotecan for 6 days relative to vehicle, RNA-seq data in bins of 200 genes by length ( $n = 2$  biological replicates).

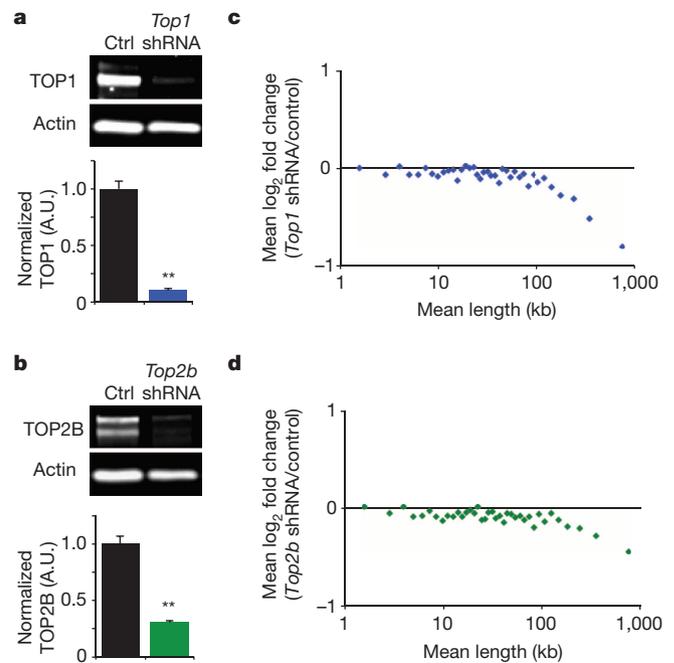
other long genes. Notably, using RNA-seq and Affymetrix microarrays to quantify gene expression, we found that topotecan reduced the expression of nearly all extremely long genes in mouse cortical neurons (Fig. 1a–c), with a strong correlation between gene length and reduced expression (for genes longer than 67 kilobases (kb); Pearson's correlation coefficient ( $R$ ) =  $-0.69$ ). Topotecan also reduced the expression of long genes in iPSC-derived human neurons (Fig. 1d). Topotecan did not exclusively reduce the expression of extremely long genes, but instead acted over a continuum of gene lengths (Fig. 1c). Specifically, the percentage of genes that were inhibited (to any extent) by 300 nM topotecan increased from 50% for genes 67 kb in length to nearly 100% for genes  $\sim$ 200 kb and longer. In addition, inhibition of long genes by topotecan was highly dose dependent (Extended Data Fig. 2).

By contrast, topotecan increased the expression of most of the genes that were  $<67$  kb in length (Fig. 1c), although the magnitude of this increase was very small for most genes (Fig. 1a, b). For some genes, this increase may reflect regulation by longer overlapping transcripts, like for *Ube3a*, or it might reflect other stimulatory effects of topoisomerase inhibitors<sup>22,23</sup>.

The length-dependent effects on gene expression were not due to cell death or persistent DNA damage, as topotecan (300 nM for 3 days) did not kill neurons or damage DNA (Extended Data Fig. 3a, b). Moreover, agents that damage DNA in neurons (paraquat and  $H_2O_2$ ) did not reduce the expression of long genes (Extended Data Fig. 3b–d). Notably, all length-dependent effects were fully reversible upon drug washout (Extended Data Fig. 3e), ruling out the possibility that gene expression changes were due to permanent effects (such as irreversible DNA damage and/or killing neurons).

A different TOP1 inhibitor, irinotecan, had a highly correlated length-dependent effect on gene expression in cortical neurons (Extended Data Fig. 4). In addition, we re-analysed published data from other laboratories and found that irinotecan and camptothecin (another TOP1 inhibitor) strongly reduced the expression of long genes and moderately increased the expression of shorter genes in several human cell lines (Extended Data Fig. 5a–e). Thus, the length-dependent effects we observed were not unique to postmitotic neurons and could be reproducibly detected in expression data acquired by other laboratories.

We found that lentiviral delivery of short hairpin RNA (shRNA) to knockdown *Top1* also reduced the expression of long genes in neurons

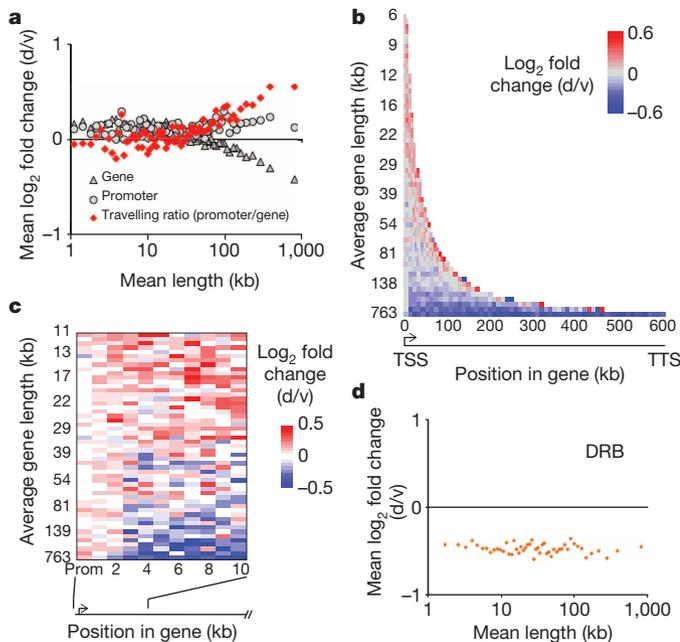


**Figure 2 | Lentiviral shRNA knockdown of *Top1* or *Top2b* in mouse cortical neurons reduces the expression of long genes.** **a**, **b**, Representative western blots and quantification of TOP1 (**a**) and TOP2B (**b**) 7 days post infection. Normalized to  $\beta$ -actin in arbitrary units (A.U.).  $**P < 0.01$  relative to scrambled control (Ctrl), Student's  $t$ -test. Error bars, s.e.m.  $n = 3$  biological replicates. **c**, **d**, Gene expression from Affymetrix microarrays, relative to scrambled control shRNA. Plotted as mean expression change in bins of 200 genes by length.

(Fig. 2a, c), providing independent genetic support that TOP1 facilitates the expression of long genes. These gene knockdown results also rule out the possibility that TOP1–DNA covalent complexes, which form only in the presence of TOP1 inhibitors<sup>18</sup>, block the expression of long genes. Unlike TOP1 inhibitors (Fig. 1a–c and Extended Data Figs 4a and 5a–e), *Top1* knockdown did not globally increase the expression of shorter genes (Fig. 2c). Thus TOP1 inhibitors probably increase the expression of shorter genes through a drug-specific effect that is unrelated to TOP1 depletion.

TOP2 enzymes (particularly TOP2B) also participate in gene transcription<sup>15,16,24</sup>. We next tested whether genetic or pharmacological inhibition of TOP2 enzymes could reduce the expression of long genes. Indeed, with new experiments and by re-analysing data from others<sup>14,25</sup>, we found that the TOP2A/TOP2B inhibitor ICRF-193 reduced gene expression in a length-dependent manner in cultured mouse cortical neurons, embryonic stem (ES) cells and ES-cell-derived neurons (Extended Data Figs 6a and 7a, b). There was extensive overlap between genes affected by ICRF-193 and topotecan in cortical neurons, particularly for long genes, and the magnitudes of these effects were highly correlated (Extended Data Fig. 6b–e). Thus, TOP1 and TOP2 enzymes regulate the expression of many of the same genes.

*Top2b* is the predominant TOP2 expressed in neurons<sup>25</sup>. We next knocked down *Top2b* with shRNA (Fig. 2b, d) and found that this manipulation reduced the expression of long genes (Fig. 2d). Moreover, re-analysis of published data sets showed that expression of long genes was reduced in embryonic brain and ES-cell-derived neurons from *Top2b*<sup>−/−</sup> mice<sup>14</sup> (Extended Data Fig. 7c, f, g). By contrast, long genes were expressed normally in *Top2b*<sup>−/−</sup> ES cells and neuronal progenitors<sup>25</sup> (Extended Data Fig. 7d, e), presumably because these cell types express *Top2a* in addition to *Top2b*<sup>25</sup>. Last, two additional TOP2 inhibitors (doxorubicin and etoposide) reduced the expression of long genes in a human cancer cell line<sup>26</sup> (Extended Data Fig. 5f, g). Together, our data show that TOP1 and TOP2 enzymes are required for proper expression of long genes in mammalian cells. This stands in contrast to yeast, in



**Figure 3 | Topotecan impairs transcription elongation of long genes.**

**a**, Mouse cortical neurons treated with vehicle or topotecan (300 nM for 3 days). Pol II density in gene bodies and promoter regions, averaged for bins of 200 genes by length. **b**, **c**, Change in Pol II density across genes (**b**) and across the first 10 kb (**c**), averaged for groups of 200 genes by length. Gene bins aligned relative to the transcription start site (TSS). TTS, transcription termination site. **d**, Mouse cortical neurons treated with 100  $\mu$ M DRB for 3 days. Affymetrix microarray expression compared to controls in bins of 200 genes by length.

which length-dependent transcriptional effects specific to Top2 were observed<sup>27</sup>.

### Length-dependent effect on elongation

Topoisomerases interact directly with RNA polymerase II (Pol II) and are required for transcription elongation<sup>17,28,29</sup>. To study genome-wide effects of topotecan on transcription in neurons, we measured Pol II distribution by chromatin immunoprecipitation followed by high-throughput DNA sequencing (ChIP-seq). We calculated the travelling ratio for all genes bound by Pol II, defined as the ratio of Pol II density (read counts per million mapped reads) near transcription start sites to Pol II density in the body of the gene, a parameter that is increased when transcription elongation is inhibited<sup>30–32</sup>. We found that topotecan produced a distinct length-dependent increase in travelling ratio for very long genes, consistent with an impairment of transcription elongation (Fig. 3a and Extended Data Fig. 8).

The change in travelling ratio observed with topotecan treatment could reflect either a progressive inhibition of Pol II as it transcribes long genes, or it could reflect a block in the transition to productive elongation. To distinguish between these possibilities, we examined the change in Pol II density across the entire length of genes at high resolution (Fig. 3b). We found that Pol II density was slightly increased throughout the gene bodies of smaller genes, consistent with the modest increase in expression seen for smaller genes and with studies showing that TOP1 inhibitors can stimulate the transition to elongation of two genes (both shorter than 67 kb)<sup>5,6</sup>. However, for longer genes, topotecan strongly reduced Pol II density across the entire length of the gene body (Fig. 3b). These results are consistent with topotecan affecting the transition to productive elongation at long genes. On average, all genes showed a slight increase in Pol II density in the promoter-proximal region (Fig. 3c) and a stronger increase in Pol II density near the transcription termination site (Fig. 3b). However, these changes were independent of length, making it unlikely that they contributed to differential expression of short versus long genes.

TOP1 inhibitors can stimulate the release of the positive elongation factor P-TEFb<sup>17</sup>. Thus, we next tested whether P-TEFb inhibition would affect expression of longer genes differently from shorter genes by treating neurons with 5,6-dichloro-1- $\beta$ -D-ribofuranosylbenzimidazole (DRB; a P-TEFb inhibitor that can block the transition to elongation<sup>17,30</sup>). We found that DRB (100  $\mu$ M) reduced gene expression equally across all gene lengths (Fig. 3d), suggesting that the requirement for P-TEFb is not influenced by length. (Note that 1  $\mu$ M and 10  $\mu$ M DRB had no length-dependent effects on expression (data not shown), ruling out the possibility that lower DRB concentrations preferentially impair expression of longer genes.)

### Numerous long ASD genes affected

To investigate further the biological consequences of TOP1 inhibition in neurons, we defined a list of genes that were differentially expressed with high confidence. From our RNA-seq expression data, we found that topotecan significantly downregulated 155 genes and significantly upregulated 28 genes (Benjamini–Hochberg method, 5% false discovery rate) (Supplementary Data 1). The topotecan-downregulated genes were significantly longer (mean 591 kb, median 548 kb) than all expressed genes in cortical neurons (mean 59.3 kb, median 23.5 kb) and were significantly longer than topotecan-upregulated genes (mean 29.3 kb, median 16.4 kb) (one-way analysis of variance versus all expressed genes  $P = 2.2 \times 10^{-16}$ , versus upregulated genes  $P = 3.7 \times 10^{-14}$ ), further indicating that topotecan has pronounced effects on long genes.

On the basis of Gene Ontology and functional annotation terms, we found that many topotecan-downregulated genes were involved in neuronal development and synaptic function (Supplementary Data 2). Because ASD is thought to be a neurodevelopmental disorder that affects synapses, we cross-referenced our list of downregulated genes with known ASD candidate genes, combining genes in the SFARI Gene database with candidates identified in recent exome sequencing studies<sup>11,12,33–35</sup> (Supplementary Data 3). Notably, 27% ( $n = 49$ ) of the 183 differentially expressed genes are known ASD candidate genes (Table 1 and Supplementary Data 3), a proportion that is highly significant compared to chance ( $P = 4.4 \times 10^{-8}$ , Fisher's exact test). Independent microarray experiments showed that these ASD genes were dose-dependently downregulated by topotecan (Extended Data Fig. 9). Notably, ASD candidate genes are exceptionally long as a group; genes in the SFARI Gene database (as of June 20, 2013) are 3.7-fold longer on average than all genes expressed in cortical neurons (means of 217.3 kb versus 59.3 kb). Thus, mutations that alter topoisomerase activity might reduce expression of numerous long ASD genes and might contribute significantly to ASD. Consistent with this hypothesis, recent sequencing studies of autism patient cohorts uncovered rare *de novo* missense mutations in *TOP1*, *TOP3B*, *TOPORS* (a TOP1-SUMO ligase<sup>36</sup>) and several other genes that directly connect to TOP1<sup>11,12</sup>.

### Discussion

Our study shows that topoisomerases facilitate the expression of a large number of ASD candidate genes, including many that are long and that are thought to have large effects on ASD pathology in isolation<sup>1,2,37</sup>. Pharmacological inhibition of topoisomerases also reduced expression of long genes in other cell types, suggesting that this length-dependent transcriptional effect is fundamental to all mammalian cells. Our data rule out numerous possibilities as to why topoisomerase inhibitors reduce the expression of long genes (for example, cell death, DNA damage, formation of covalent complexes), and instead implicate a gene-length-dependent block in transcription elongation. Pol II and topoisomerases dynamically form and remodel large supercoiling domains<sup>38</sup>, and the effects of topoisomerases on gene expression are strongly influenced by genomic structure and context<sup>23,39</sup>. Thus, we speculate that higher order structure differentially constrains shorter and longer genes, and that this creates distinct length-dependent requirements for topoisomerases in transcription elongation.

**Table 1 | Topotecan reduces expression of numerous ASD candidate genes in neurons**

Gene	Length in mouse (kb)	Length in human (kb)	Log <sub>2</sub> fold change (mouse)	P (adjusted)	SFARI Gene	Mutation
<b>Cntnap2*</b>	2241.3	2304.6	-4.5718	0.0055	Yes	H275R <sup>34</sup>
<b>Csmd1</b>	1642.8	2059.4	-3.4896	<0.0001	No	Q2254R <sup>33</sup>
<b>Nrxn3</b>	1612.1	1460.7	-2.9494	<0.0001	Yes	
<b>Fhit</b>	1611.9	1502.1	-2.1229	0.0490	Yes	
<b>Rbfox1</b>	1527.7	1694.2	-1.3178	0.0441	Yes	
<b>Grid2</b>	1409.4	1468.1	-2.4100	0.0204	Yes	
<b>Il1rapl1</b>	1368.4	1368.3	-3.0444	0.0213	Yes	
<b>Il1rapl2</b>	1275.4	1200.8	-2.6076	0.0034	Yes	
<b>Park2</b>	1223.0	1380.2	-2.8055	0.0003	Yes	
<b>Cntn5</b>	1243.9	1337.9	-2.3360	0.0009	Yes	
<b>Ptprt</b>	1133.5	1117.2	-2.0706	<0.0001	Yes	
<b>Frmpld4</b>	1105.9	586.1	-1.5332	0.0028	Yes	
<b>ErbB4</b>	1068.1	1162.9	-3.3890	<0.0001	Yes	
<b>Nrxn1</b>	1059.2	1114.0	-2.1914	<0.0001	Yes	Y587STOP <sup>12</sup>
<b>Gpc6</b>	1054.2	1181.2	-2.3530	<0.0001	Yes	
<b>Astn2</b>	1023.7	989.8	-2.7763	0.0001	Yes	
<b>Cntn4</b>	1021.6	959.1	-2.0717	0.0092	Yes	
<b>Pard3b</b>	1003.5	1074.4	-2.9457	<0.0001	Yes	
<b>Epha6</b>	952.0	934.3	-1.8366	0.0301	Yes	
<b>Grm7</b>	921.6	880.4	-2.2647	<0.0001	No	R622Q <sup>33</sup>
<b>Nckap5</b>	917.0	896.7	-1.3453	0.0089	Yes	
<b>Nlgn1</b>	900.1	884.9	-1.8339	<0.0001	Yes	H795Y <sup>34</sup>
<b>Pcdh9</b>	875.2	927.5	-1.8314	0.0002	Yes	
<b>Pcdh15</b>	827.9	998.5	-3.2033	<0.0001	Yes	
<b>Grid1</b>	761.0	766.9	-1.4922	0.0448	Yes	
<b>Mdga2</b>	756.5	835.3	-1.6942	0.0012	Yes	
<b>Dpp10</b>	713.4	1402.4	-1.7914	0.0007	Yes	
<b>Kcnma1</b>	705.4	768.2	-1.9021	0.0004	Yes	F372V <sup>11</sup>
<b>Grik2</b>	689.3	669.2	-1.5500	0.0028	Yes	
<b>Plcb1</b>	689.1	752.3	-1.4067	0.0075	Yes	
<b>Ptprm</b>	687.6	839.5	-1.7610	0.0057	No	A535T <sup>12</sup>
<b>Rbms3</b>	679.7	729.1	-1.6432	0.0191	Yes	
<b>Vps13b</b>	559.7	864.3	-1.0465	0.0317	Yes	
<b>Nbea</b>	558.5	730.5	-1.3664	0.0147	Yes	
<b>Grm5</b>	550.9	559.1	-1.3240	0.0343	Yes	Δaa679 <sup>12</sup>
<b>Kalrn</b>	545.0	626.5	-1.3646	0.0186	Yes	
<b>Ptprk</b>	522.6	551.9	-1.6458	0.0005	No	R784H <sup>34</sup>
<b>Atrnl1</b>	522.3	855.3	-1.2194	0.0211	Yes	
<b>Robo2</b>	518.6	1712.5	-1.4921	0.0017	Yes	
<b>Cacna1c</b>	515.9	644.7	-1.1665	0.0236	Yes	
<b>Lrrc7</b>	477.8	363.3	-1.6330	0.0340	Yes	
<b>Reln</b>	460.3	517.7	-2.0687	<0.0001	Yes	Q417STOP <sup>11</sup> R2290C <sup>12</sup>
<b>Exoc6b</b>	451.0	646.7	-0.9691	0.0285	Yes	
<b>Gabrb1</b>	437.2	395.2	-1.4909	0.0041	Yes	
<b>Nfia</b>	341.3	385.5	-1.8218	0.0186	Yes	R30STOP <sup>12</sup>
<b>Nxph1</b>	298.6	319.0	-1.8182	0.0181	Yes	
<b>Myb</b>	36.1	37.9	-3.8779	0.0011	No	R45Q <sup>33</sup>
<b>C3</b>	24.1	42.8	-2.2892	<0.0001	No	I1569M <sup>11</sup>
<b>C4b</b>	15.5	1.7	-1.5240	0.0121	Yes	

\* Isoform a (*uc009bst.2*).  
aa, amino acid.

Some long genes were not strongly reduced in expression following topotecan treatment (Fig. 1a). In many cases this reflected ambiguity in gene annotation (data not shown). For example, a number of long genes also express shorter transcripts, making it difficult to distinguish expression of short isoforms from long isoforms. Alternatively, some long genes might be located within genomic regions that are more permissive for expression when TOP1 is inhibited.

Intriguingly, numerous genes associated with transcription are mutated in autism patients<sup>11,40,41</sup>, although how these diverse transcriptional regulators contribute to autism is unclear. Our study highlights a mechanistic link between a critical step in transcription elongation and expression of numerous long ASD candidate genes. Our data suggest that chemicals or genetic mutations that impair topoisomerases, and possibly other components of the transcription elongation machinery that interface with topoisomerases, have the potential to profoundly affect the expression of long ASD candidate genes. Length-dependent impairment of gene transcription, particularly in neurons and during critical periods of brain development, may thus represent a unifying cause of pathology in many individuals with ASD and other neurodevelopmental disorders.

## METHODS SUMMARY

All procedures using vertebrate animals were approved by the University of North Carolina Institutional Animal Care and Use Committee. Embryonic (E13.5–15.5) cortical neuron cultures from C57BL/6J × CASTEi/J mice (The Jackson Laboratory) were used for RNA-seq and ChIP-seq studies. C57BL/6J mice were used for all other mouse experiments. Cultures were prepared as previously described<sup>7</sup>. Affymetrix Mouse Genome 430 2.0 arrays and Mouse Gene 1.1 ST 24-array plates were used for microarray studies. ChIP-seq experiments were performed essentially as described<sup>30,42</sup> using RNA Pol II N20 antibody (sc-899, Santa Cruz Biotechnology).

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 January; accepted 24 July 2013.

Published online 28 August 2013.

1. Abrahams, B. S. & Geschwind, D. H. Advances in autism genetics: on the threshold of a new neurobiology. *Nature Rev. Genet.* **9**, 341–355 (2008).
2. State, M. W. & Levitt, P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nature Neurosci.* **14**, 1499–1506 (2011).
3. Christensen, J. G. T. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *J. Am. Med. Assoc.* **309**, 1696–1703 (2013).

4. Delorme, R. *et al.* Progress toward treatments for synaptic defects in autism. *Nature Med.* **19**, 685–694 (2013).
5. Betancur, C., Sakurai, T. & Buxbaum, J. D. The emerging role of synaptic cell-adhesion pathways in the pathogenesis of autism spectrum disorders. *Trends Neurosci.* **32**, 402–412 (2009).
6. Peça, J. & Feng, G. Cellular and synaptic network defects in autism. *Curr. Opin. Neurobiol.* **22**, 866–872 (2012).
7. Huang, H.-S. *et al.* Topoisomerase inhibitors silence the dormant allele of *Ube3a* in neurons. *Nature* **481**, 185–189 (2012).
8. Mabb, A. M., Judson, M. C., Zylka, M. J. & Philpot, B. D. Angelman syndrome: insights into genomic imprinting and neurodevelopmental phenotypes. *Trends Neurosci.* **34**, 293–303 (2011).
9. Cook, E. H. *et al.* Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *Am. J. Hum. Genet.* **60**, 928–934 (1997).
10. Moreno-De-Luca, D. *et al.* Using large clinical data sets to infer pathogenicity for rare copy number variants in autism cohorts. *Mol. Psychiatry* <http://dx.doi.org/10.1038/mp.2012.138> (2012).
11. Neale, B. M. *et al.* Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).
12. Iossifov, I. *et al.* *De novo* gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285–299 (2012).
13. Plaschkes, I., Silverman, F. W. & Priel, E. DNA topoisomerase I in the mouse central nervous system: age and sex dependence. *J. Comp. Neurol.* **493**, 357–369 (2005).
14. Lyu, Y. L. *et al.* Role of topoisomerase II $\beta$  in the expression of developmentally regulated genes. *Mol. Cell Biol.* **26**, 7929–7941 (2006).
15. Vos, S. M., Tretter, E. M., Schmidt, B. H. & Berger, J. M. All tangled up: how cells direct, manage and exploit topoisomerase function. *Nature Rev. Mol. Cell Biol.* **12**, 827–841 (2011).
16. Baranello, L., Levens, D., Gupta, A. & Kouzine, F. The importance of being supercoiled: how DNA mechanics regulate dynamic processes. *Biochim. Biophys. Acta* **1819**, 632–638 (2012).
17. Capranico, G., Marinello, J. & Baranello, L. Dissecting the transcriptional functions of human DNA topoisomerase I by selective inhibitors: implications for physiological and therapeutic modulation of enzyme activity. *Biochim. Biophys. Acta* **1806**, 240–250 (2010).
18. Pommier, Y. Drugging topoisomerases: lessons and challenges. *ACS Chem. Biol.* **8**, 82–95 (2013).
19. Babak, T. *et al.* Global survey of genomic imprinting by transcriptome sequencing. *Curr. Biol.* **18**, 1735–1741 (2008).
20. Meng, L., Person, R. E. & Beaudet, A. L. *Ube3a-ATS* is an atypical RNA polymerase II transcript that represses the paternal expression of *Ube3a*. *Hum. Mol. Genet.* **21**, 3001–3012 (2012).
21. Landers, M. *et al.* Regulation of the large (~1000 kb) imprinted murine *Ube3a* antisense transcript by alternative exons upstream of *Snurf/Snrpn*. *Nucleic Acids Res.* **32**, 3480–3492 (2004).
22. Ljungman, M. & Hanawalt, P. C. The anti-cancer drug camptothecin inhibits elongation but stimulates initiation of RNA polymerase II transcription. *Carcinogenesis* **17**, 31–36 (1996).
23. Collins, I., Weber, A. & Levens, D. Transcriptional consequences of topoisomerase inhibition. *Mol. Cell Biol.* **21**, 8437–8451 (2001).
24. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl Acad. Sci. USA* **84**, 7024–7027 (1987).
25. Tiwari, V. K. *et al.* Target genes of topoisomerase II $\beta$  regulate neuronal survival and are defined by their chromatin state. *Proc. Natl Acad. Sci. USA* **109**, E934–E943 (2012).
26. Troester, M. A., Hoadley, K. A., Parker, J. S. & Perou, C. M. Prediction of toxicant-specific gene expression signatures after chemotherapeutic treatment of breast cell lines. *Environ. Health Perspect.* **112**, 1607–1613 (2004).
27. Joshi, R. S., Piña, B. & Roca, J. Topoisomerase II is required for the production of long Pol II gene transcripts in yeast. *Nucleic Acids Res.* <http://dx.doi.org/10.1093/nar/gks626> (2012).
28. Wu, H.-Y., Shyy, S., Wang, J. C. & Liu, L. F. Transcription generates positively and negatively supercoiled domains in the template. *Cell* **53**, 433–440 (1988).
29. Stewart, A. F., Herrera, R. E. & Nordheim, A. Rapid induction of c-fos transcription reveals quantitative linkage of RNA polymerase II and DNA topoisomerase I enzyme activities. *Cell* **60**, 141–149 (1990).
30. Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
31. Muse, G. W. *et al.* RNA polymerase is poised for activation across the genome. *Nature Genet.* **39**, 1507–1511 (2007).
32. Zeitlinger, J. *et al.* RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nature Genet.* **39**, 1512–1516 (2007).
33. Sanders, S. J. *et al.* *De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).
34. O’Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* **485**, 246–250 (2012).
35. Basu, S. N., Kollu, R. & Banerjee-Basu, S. AutDB: a gene reference resource for autism research. *Nucleic Acids Res.* **37**, D832–D836 (2009).
36. Hammer, E., Heilbronn, R. & Weger, S. The E3 ligase Topors induces the accumulation of polysumoylated forms of DNA topoisomerase I *in vitro* and *in vivo*. *FEBS Lett.* **581**, 5418–5424 (2007).
37. Zhao, X. *et al.* A unified genetic theory for sporadic and inherited autism. *Proc. Natl Acad. Sci. USA* **104**, 12831–12836 (2007).
38. Naughton, C. *et al.* Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nature Struct. Mol. Biol.* **20**, 387–395 (2013).
39. Sano, K., Miyaji-Yamaguchi, M., Tsutsui, K. M., Tsutsui, K. & Topoisomerase, I. I. Activates a subset of neuronal genes that are repressed in AT-rich genomic environment. *PLoS ONE* **3**, e4103 (2008).
40. Ben-David, E. & Shifman, S. Combined analysis of exome sequencing points toward a major role for transcription regulation during brain development in autism. *Mol. Psychiatry* <http://dx.doi.org/10.1038/mp.2012.148> (2012).
41. Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition: unexpected roles for chromatin. *Nature Rev. Genet.* **14**, 347–359 (2013).
42. Forsberg, E. C., Downs, K. M. & Bresnick, E. H. Direct interaction of NF-E2 with hypersensitive site 2 of the  $\beta$ -globin locus control region in living cells. *Blood* **96**, 334–339 (2000).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank M. Vernon at the UNC Expression Profiling Core for assistance with microarray experiments and with data analysis, H. Kelkar for pilot bioinformatics support, P. Mieczkowski and A. Brandt at the UNC High Throughput Sequencing Facility for advice and assistance with Illumina library preparation and sequencing and T. Kafri and P. Zhang at the UNC Lenti-shRNA Core for assistance with preparation of lentiviral vectors. This work was supported by grants to M.J.Z. and B.D.P. from The Angelman Syndrome Foundation, The Simons Foundation (SFARI 10-3625) and The National Institute of Mental Health (R01MH093372). I.F.K. and A.M.M. were supported by Joseph E. Wagstaff Postdoctoral Research Fellowships from the Angelman Syndrome Foundation. J.M.C. was supported by a grant from the American Cancer Society (117571-PF-09-124-01-DDC). J.M.C., J.S. and T.M. were supported by a grant from the National Institutes of Health (NIH) (R01GM101974). B.L.P. was supported by a NIH postdoctoral training grant (T32HD040127). S.J.C. was supported by a grant from National Institute of Child Health and Human Development (NICHD) (R01HD068730). The expression profiling and bioinformatics cores were funded by grants from the National Institute of Neurological Disorders and Stroke (P30NS045892) and NICHD (P30HD03110).

**Author Contributions** I.F.K., H.-S.H., A.M.M., J.S.H., S.J.C., B.D.P. and M.J.Z. conceived and designed experiments. I.F.K. performed RNA-seq and ChIP-seq experiments with mouse neurons. I.F.K., C.N.Y., J.M.C., J.S. and J.S.P. analysed data from genome-wide experiments and from published datasets. J.M.C. performed SNP filtering of RNA-seq data and J.S. performed statistical analysis of RNA-seq data. A.M.M. performed lentiviral shRNA knockdown experiments. J.S.H. and S.J.C. performed all experiments with iPSC-derived human neurons. B.L.P. assessed propensity of compounds to kill neurons and damage DNA. T.M. provided bioinformatics support. I.F.K., H.-S.H. and A.M.M. performed microarray experiments. I.F.K. and M.J.Z. wrote the manuscript.

**Author Information** Data from Microarray, RNA-seq and ChIP-seq experiments have been deposited in the Gene Expression Omnibus under accession number (GSE43900). Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.D.P. ([bphilpot@med.unc.edu](mailto:bphilpot@med.unc.edu)) or M.J.Z. ([zylka@med.unc.edu](mailto:zylka@med.unc.edu)).

## METHODS

**Mouse cell culture.** Cortical neurons were cultured from E13.5–E15.5 mouse embryos as described<sup>7</sup>. For RNA-seq and ChIP-seq, neurons were seeded on 10-cm diameter poly-D-lysine-treated culture plates at a density of  $5\text{--}10 \times 10^6$  per plate. Microarray experiments used 6-well plates seeded at  $1 \times 10^6$  cells per well. After 7 days in culture, drugs or an equivalent amount of vehicle were added and left in the culture medium for 24 or 72 h. For topotecan, irinotecan, DRB and paraquat, vehicle was 0.1% dimethylsulphoxide (DMSO). For ICRF-193, vehicle was 0.02% DMSO.

Lentiviral shRNA experiments used viruses from The RNAi Consortium (acquired from Sigma-Aldrich and from the UNC Lenti-shRNA core): *Top1* knockdown used clone TRCN0000011884 (5'-CCGCCAGCGAAGATTCTATCTTATCTCGAGATAAGATAGAAATCTTCGCTGGTTTTT-3') and *Top2b* knockdown used clone TRCN0000070988 (5'-CCGCCCTTGTTGTCCTTTGTCTTCTCGAGAAGACAAAGGACAACACAAGGTTTTT-3'). Virus expressing non-targeting hairpin RNA (SHC002, Sigma; 5'-CCGGCGTGATCTTCACCGACAAGATCTCGAGATCTTGTGCGGTGAAGATCACGTTTTT-3') was used as a control. Neurons were seeded on 24-well plates at  $2 \times 10^5$  per well. After 3 days in culture, cells were treated with lentivirus at a multiplicity of infection of at least 1. Virus was removed after 24 h, and RNA collected after a further 6 days in culture. Western blotting to assess knockdown was performed with anti-TOP1 (Santa Cruz Biotechnology, H5) or anti-TOP2B (Santa Cruz, H286) antibodies, with signal normalized to  $\beta$ -actin (Millipore, C4).

Cell death was assayed using Sytox Green (Invitrogen Molecular Probes).  $\gamma$ -H2AX foci were measured by immunohistochemistry. Primary antibodies were anti- $\gamma$ -H2AX (Millipore, 1:500 dilution) and anti-NeuN (Millipore, 1:500), used to mark neurons.

**iPSC culture and neuronal differentiation.** Human iPSC work was approved by the University of Connecticut Stem Cell Research Oversight Committee. iPSCs that carry a large deletion of maternal 15q11–q13 (AGdel1-0; see also ref. 43), this cell line was deemed exempt from IRB approval at the University of Connecticut due to its establishment in 1995 and lack of identifying information) were cultured on irradiated mouse embryonic fibroblasts and manually passaged as described<sup>43</sup>. iPSCs were differentiated into forebrain cortical neurons as described<sup>43</sup> with the following modifications: neural progenitors were generated by culturing iPSCs on feeders in N2B27 medium supplemented with noggin ( $500 \text{ ng ml}^{-1}$ ) for 8 days and then manually picking neural rosettes for two additional passages using trypsin and standard cell culture protocols. Topotecan was applied to mature neurons and RNA was collected by standard protocols 6 days after the addition of drug or vehicle. **qPCR.** qRT-PCR was carried out as described<sup>43</sup> using Taqman (Life Technologies) gene expression assays for *UBE3A* (Hs00166580\_m1) and *UBE3A-ATS* (Hs03454279\_m1) according to the manufacturer's instructions. The Taqman assay for GAPDH was used as a control.

**RNA-seq.** Total RNA was collected using Trizol reagent (Invitrogen). Mouse polyA-selected messenger RNA libraries were then prepared using the Illumina True-Seq kit for RNA. For RNA-seq on human-iPSC-derived neuronal samples, stranded multiplexed mRNA libraries were prepared using Illumina kits. Cluster generation and sequencing were performed using the Illumina HiSeq 2000 platform. For allele-specific expression analysis, equal amounts of total RNA from 3–6 biological replicates were pooled before polyA mRNA purification and library preparation.

For non-allelic expression analysis, data from five biological replicates were included. mRNA was isolated and libraries prepared independently for each replicate sample. Library preparation incorporated barcoded adapters and all samples were sequenced in one lane, using 50-base-pair (bp) reads.

**RNA-seq expression analysis.** For allelic expression analysis, informative CAST/B6 SNPs were downloaded from <http://www.sanger.ac.uk/resources/mouse/genomes/>.

CAST alleles were then substituted into their corresponding mouse reference genome (mm9) positions and sequence reads were aligned to the mm9 and the CAST version of mm9 using Bowtie, selecting for unique matches. Filtered read counts for autosomal genes were tested for allelic bias using Fisher's exact test against a background model derived from autosomes, and *P* values adjusted for multiple comparisons using the Benjamini–Hochberg procedure. Statistical analysis was performed using R. Genomic intervals were derived from UCSC known genes or created manually where annotation was absent, namely for *Ube3a-ATS*.

For non-allelic analysis, reads were aligned to mm9 using Bowtie. Read counts were obtained using DEGseq, and normalization and analysis of differential gene expression was performed using the R package, edgeR, using a negative binomial model.

**ChIP-seq.** ChIP-seq against RNA Pol II was performed as described previously<sup>30</sup>. Cultures ( $n = 4$  per condition) totalling approximately  $2 \times 10^7$  neurons were treated with vehicle or 300 nM topotecan as described above and pooled. Nuclear lysates were sheared to an average fragment size of approximately 200 bp. 2  $\mu\text{g}$  anti-RNA Pol II N20 (Santa Cruz Biotechnology) was added, and the sample incubated at 4 °C for 16 h. ChIP was performed as described previously<sup>30,42</sup>.

ChIP-seq libraries were prepared from immunoprecipitated samples and their corresponding inputs using the Illumina Tru-Seq kit for ChIP-seq. Ligation products were size-selected by purification on 2% PippinPrep gels (Sage Science). Samples from vehicle- and drug-treated cells and their inputs were sequenced using the HiSeq 2000 platform with single-end reads of 50 bp.

**ChIP-seq analysis.** Short read sequences were aligned to the mouse reference genome (mm9) with Bowtie. Duplicate reads were removed. The quality of the experiment and false discovery rate for enriched peaks was assessed using MACS 1.4.2. CoverageBED was used to obtain read counts covering the promoter region (from –30 to +300 bp, relative to TSS) and gene bodies (from +300 after TSS to 3,000 bp after the annotated TTS), and to count reads in intervals across genes. Read counts were normalized to the number of mapped unique reads per sample per base.

**Affymetrix microarrays.** For single-dose microarray experiments, cultured cortical neurons were treated with 300 nM or 1  $\mu\text{M}$  topotecan, 10  $\mu\text{M}$  irinotecan (Sigma), 3  $\mu\text{M}$  ICRF-193 (Santa Cruz Biotechnology), 100  $\mu\text{M}$  DRB (Sigma), 100  $\mu\text{M}$  H<sub>2</sub>O<sub>2</sub> (Fisher Scientific) or 10  $\mu\text{M}$  paraquat (Sigma) for 24 or 72 h. For topotecan dose response, cells were treated with 3 nM, 30 nM, 150 nM, 300 nM, 500 nM and 1,000 nM topotecan or vehicle for 72 h. Total RNA was used for all Affymetrix microarray experiments. Comparative expression with topotecan and ICRF-193 was measured with Affymetrix mouse genome 430 2.0 arrays. All other microarray experiments used Affymetrix Mouse Gene 1.0 ST 24-array plates. Linear RMA background correction and normalization was used for all microarray data.

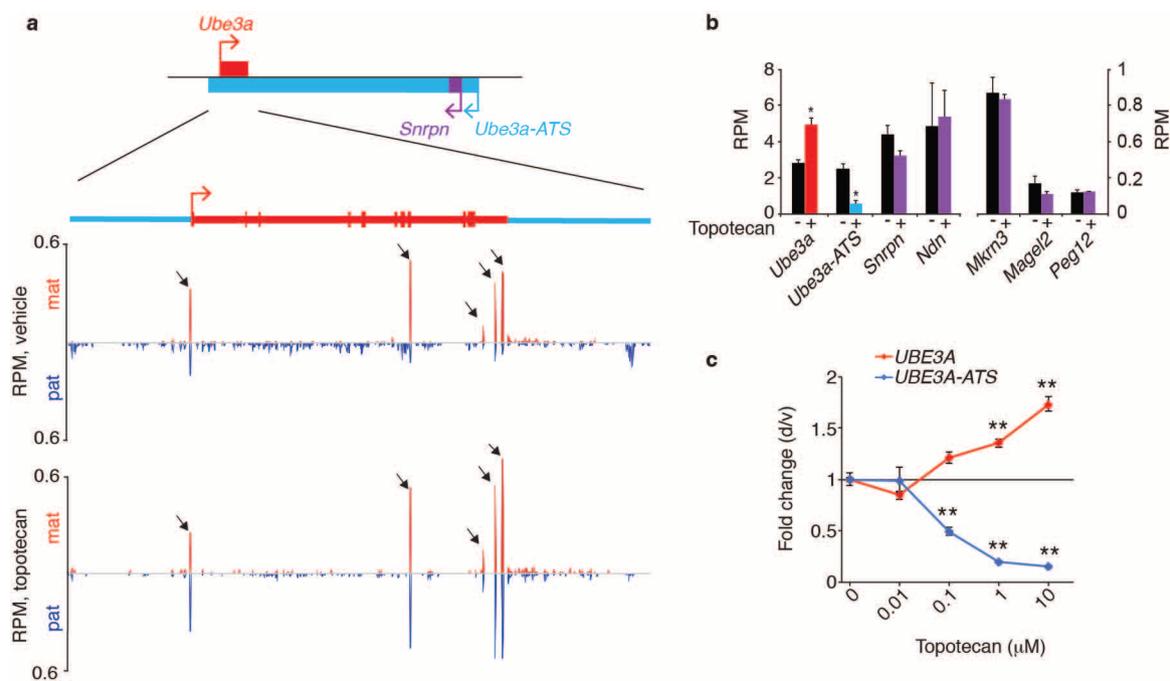
43. Chamberlain, S. J. *et al.* Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader–Willi syndromes. *Proc. Natl Acad. Sci. USA* **107**, 17668–17673 (2010).
44. Lamb, J. *et al.* The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313**, 1929–1935 (2006).
45. Iorio, F. *et al.* Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl Acad. Sci. USA* **107**, 14621–14626 (2010).
46. Groschel, B. & Bushman, F. Cell cycle arrest in G2/M promotes early steps of infection by human immunodeficiency virus. *J. Virol.* **79**, 5695–5704 (2005).
47. Carson, J. P. *et al.* Pharmacogenomic identification of targets for adjuvant therapy with the topoisomerase poison camptothecin. *Cancer Res.* **64**, 2096–2104 (2004).
48. Maclsaac, J. L., Bogutz, A. B., Morrissy, A. S. & Lefebvre, L. Tissue-specific alternative polyadenylation at the imprinted gene *Mest* regulates allelic usage at *Copg2*. *Nucleic Acids Res.* **40**, 1523–1535 (2012).

Extended Data Table 1 | Topotecan does not alter parent-of-origin bias of imprinted genes.

	Parent of origin bias (adjusted p value)		Change in bias, topotecan vs. vehicle (adjusted p value)	
	<i>B6 x CAST</i>	<i>CAST x B6</i>	<i>B6 x CAST</i>	<i>CAST x B6</i>
<b>Maternal Bias</b>				
<i>Ube3a*</i>	6.1x10 <sup>-29</sup>	1.2x10 <sup>-20</sup>	<b>3.9x10<sup>-7</sup></b>	<b>3.3x10<sup>-14</sup></b>
<i>Meg3</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Rian</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Kcnk9</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Mirg</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>H13</i>	3.0x10 <sup>-56</sup>	3.3x10 <sup>-67</sup>	ns	ns
<i>Eif2c2</i>	2.7x10 <sup>-39</sup>	6.3x10 <sup>-28</sup>	ns	ns
<i>H19</i>	1.3x10 <sup>-29</sup>	6.9x10 <sup>-56</sup>	ns	ns
<i>Rtl1</i>	1.2x10 <sup>-14</sup>	5.3x10 <sup>-30</sup>	ns	ns
<i>Cdkn1c</i>	0.00035	9.1x10 <sup>-5</sup>	ns	ns
<i>Igf2r</i>	0.0056	0.00019	ns	ns
<b>Paternal Bias</b>				
<i>Ndn*</i>	<1x10 <sup>-100</sup>	7.1x10 <sup>-61</sup>	ns	ns
<i>AK045535*</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	7.4x10 <sup>-9</sup>
<i>Snurf*</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Snrpn*</i>	2.3x10 <sup>-98</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>A330076H08Rik*</i>	2.2x10 <sup>-88</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>A230057D06Rik*</i>	6.9x10 <sup>-77</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Ube3a-ATS*</i>	2.1x10 <sup>-59</sup>	1.2x10 <sup>-66</sup>	ns	0.00074
<i>Dokist4*</i>	1.3x10 <sup>-46</sup>	3.4x10 <sup>-80</sup>	ns	ns
<i>Mkrn3*</i>	6.4x10 <sup>-18</sup>	1.3x10 <sup>-19</sup>	ns	ns
<i>AK086712*</i>	4.2x10 <sup>-13</sup>	1.9x10 <sup>-40</sup>	ns	ns
<i>U80893*</i>	1.7x10 <sup>-9</sup>	6.9x10 <sup>-10</sup>	ns	ns
<i>Magel2*</i>	1.0x10 <sup>-6</sup>	0.00065	ns	ns
<i>Peg12*</i>	0.00032	0.0040	ns	ns
<i>AK039108*</i>	0.0011	1.0x10 <sup>-10</sup>	ns	ns
<i>Peg13</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Peg3</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Peg3as</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Plagl1</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Impact</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Rasgrf1</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Zim3</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Inpp5f</i>	<1x10 <sup>-100</sup>	1.3x10 <sup>-58</sup>	ns	ns
<i>Usp29</i>	4.4x10 <sup>-99</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Sgce</i>	4.2x10 <sup>-95</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Mest</i>	2.5x10 <sup>-78</sup>	1.2x10 <sup>-36</sup>	ns	ns
<i>Peg10</i>	6.2x10 <sup>-77</sup>	3.0x10 <sup>-99</sup>	ns	ns
<i>Zdbf2</i>	9.1x10 <sup>-52</sup>	1.8x10 <sup>-43</sup>	ns	ns
<i>Zrsr1</i>	3.2x10 <sup>-48</sup>	1.7x10 <sup>-42</sup>	ns	ns
<i>Nap15</i>	4.1x10 <sup>-44</sup>	9.2x10 <sup>-28</sup>	ns	ns
<i>Kcnq1ot1</i>	1.2x10 <sup>-38</sup>	1.0x10 <sup>-73</sup>	ns	ns
<i>Kcnq1</i>	2.1x10 <sup>-36</sup>	1.3x10 <sup>-74</sup>	ns	ns
<i>Adam23</i>	1.8x10 <sup>-13</sup>	0.00010	ns	ns
<i>Grb10</i>	4.0x10 <sup>-13</sup>	4.3x10 <sup>-27</sup>	ns	ns
<i>Dlk1</i>	6.8x10 <sup>-5</sup>	4.3x10 <sup>-6</sup>	ns	ns
<i>Copg2†</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	0.014
<i>AF217545†</i>	<1x10 <sup>-100</sup>	<1x10 <sup>-100</sup>	ns	ns
<i>Cdh15</i>	2.6x10 <sup>-9</sup>	3.7x10 <sup>-12</sup>	ns	ns
<i>Bcl2l1</i>	0.028	0.0021	ns	ns

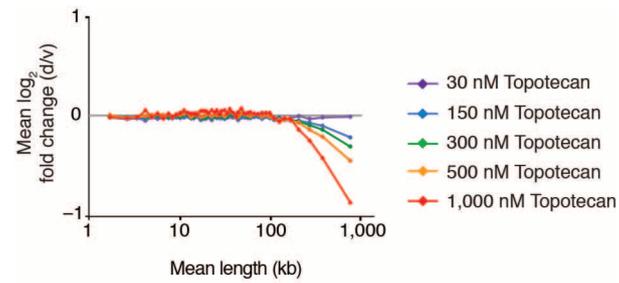
\*Genes in the *Ube3a* imprinted cluster†*AF217545* is contained within *Copg2*, which has also been reported to be maternally expressed<sup>48</sup>

Genes with significant parent-of-origin expression bias in cultured cortical neurons. Shown are *P* values for allelic bias in baseline expression, and *P* values for change in expression bias with topotecan treatment (Fisher's exact test, adjusted for multiple comparisons). NS = *P* > 0.05. Embryos from CAST male × B6 female crosses (*n* = 3) were pooled and B6 male × CAST female crosses (*n* = 6) were pooled for sequencing.

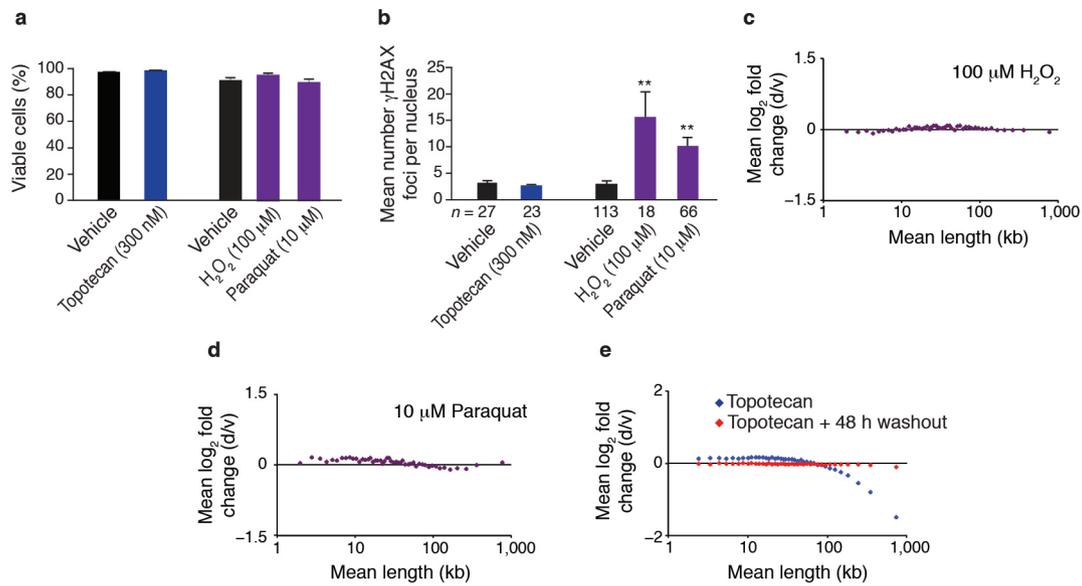


**Extended Data Figure 1 | Topotecan affects allelic expression of *Ube3a* and *Ube3a-ATS* but not expression of nearby genes.** **a**, Parent-of-origin-specific RNA-seq reads for *Ube3a* and *Ube3a-ATS* in vehicle- and topotecan-treated (300 nM for 3 days) mouse cortical neurons ( $n = 5$  biological replicates). Shown are SNP-filtered read counts per million mapped reads (RPM) from the maternally (mat) and paternally (pat) inherited chromosomes. Reads from exons and 3'-untranslated region of *Ube3a* are indicated by arrows. **b**, Expression of *Ube3a* and imprinted genes near *Ube3a*

neurons  $\pm$  300 nM topotecan for 3 days.  $*P < 0.05$ , Student's *t*-test. Error bars represent s.e.m. **c**, Expression of *UBE3A* and *UBE3A-ATS* in iPSC-derived neurons from an Angelman syndrome patient carrying a maternal deletion of the 15q11-q13 region. Differentiated neuronal cultures were treated with 10 nM–10  $\mu\text{M}$  topotecan or vehicle for 6 days. Expression quantified by qPCR.  $**P < 0.01$ , one-way analysis of variance with Dunnett's post-hoc test.  $n = 4$ . Error bars represent s.e.m.

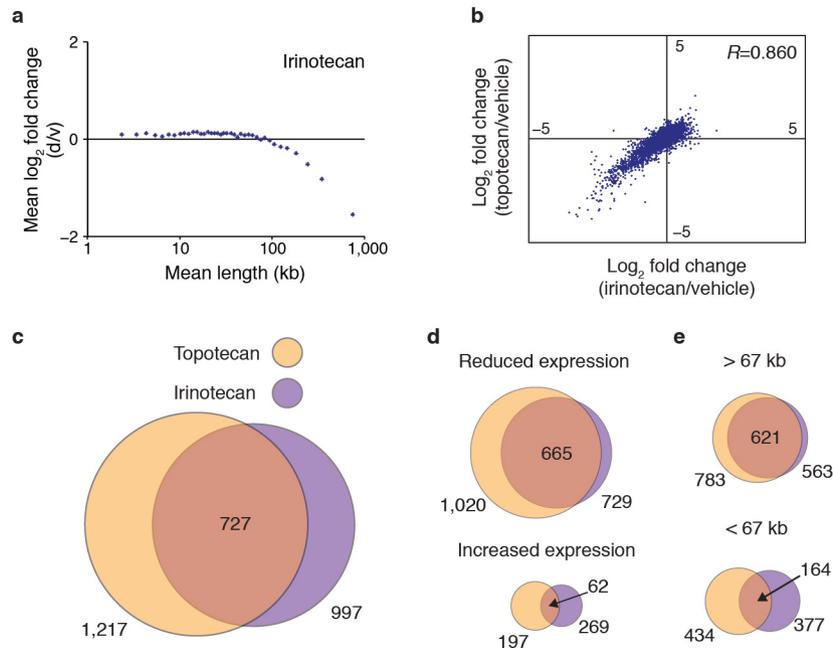


**Extended Data Figure 2 | Topotecan dose–response.** Mouse cortical neurons were treated with 3 nM, 30 nM, 150 nM, 300 nM and 1,000 nM topotecan for 3 days ( $n = 3$  for 300 nM topotecan, all other doses  $n = 1$ ). Gene expression was analysed by Affymetrix microarrays, plotted as mean expression change in bins of 200 genes.



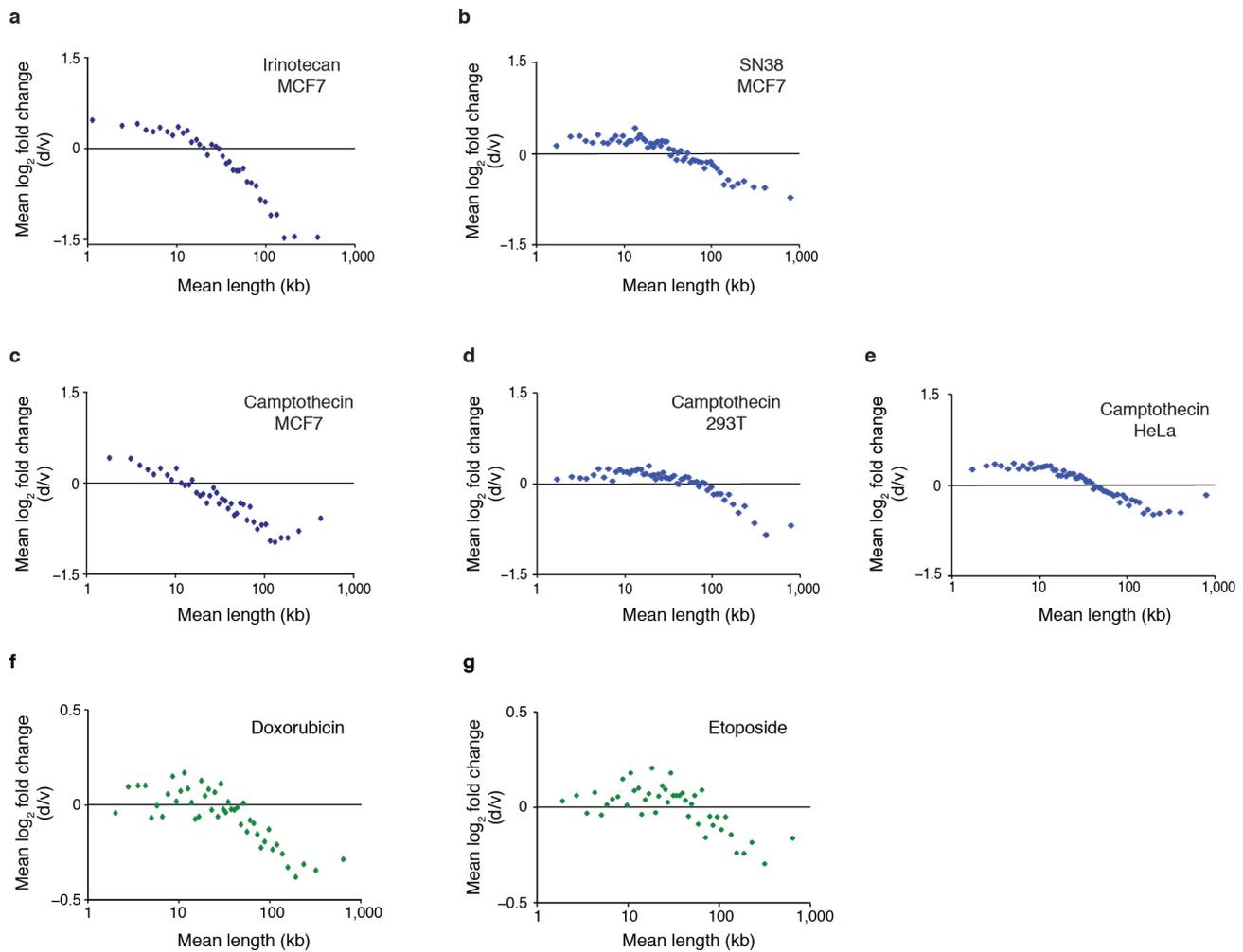
**Extended Data Figure 3 | DNA damage does not inhibit gene expression in a length-dependent manner.** **a**, Cultured mouse cortical neuron viability, assayed by Sytox Green staining after 72 h treatment with 300 nM topotecan, and after 24 h with 100 μM H<sub>2</sub>O<sub>2</sub> or 10 μM paraquat, compared to vehicle-treated controls. Error bars represent s.e.m.  $n = 4$ . **b**, Average number of  $\gamma$ -H2AX foci per nucleus for cultured cortical neurons treated with 300 nM topotecan for 72 h, 100 μM H<sub>2</sub>O<sub>2</sub> for 24 h and 10 μM paraquat for 24 h, compared to vehicle-treated controls. \*\* $P < 0.01$ , Student's  $t$ -test. Number of

cells counted is indicated for each sample. **c**, Gene expression compared to vehicle controls in bins of 200 genes by length, for cultured cortical neurons treated with 100 μM H<sub>2</sub>O<sub>2</sub> for 24 h. **d**, Gene expression compared to vehicle controls in bins of 200 genes by length, for cultured cortical neurons treated with 10 μM paraquat for 24 h. **e**, Gene expression in cultured cortical neurons treated with 300 nM topotecan for 24 h, or treated for 24 h followed by 48 h without drug (washout). Average change in expression for bins of 200 genes by length.



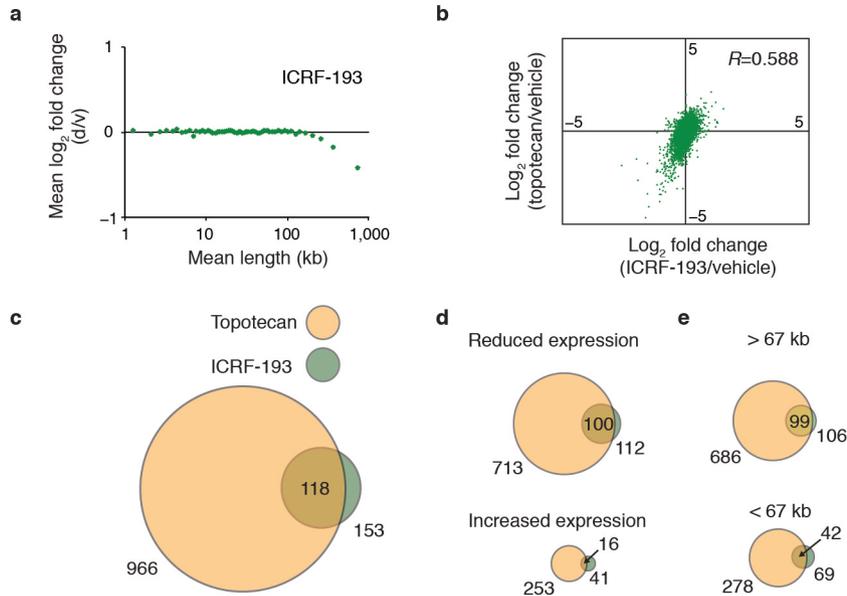
**Extended Data Figure 4 | Topotecan and irinotecan have highly similar effects on gene expression.** **a**, Affymetrix microarray analysis of gene expression in cultured mouse cortical neurons treated with vehicle or 10  $\mu$ M irinotecan ( $n = 3$  biological replicates), an inhibitor of TOP1, for 3 days. Mean expression fold change in bins of 200 genes, plotted by average gene length. **b**, Scatterplot of fold change with 1  $\mu$ M topotecan ( $n = 6$  biological replicates)

versus fold change with 10  $\mu$ M irinotecan for all expressed genes, measured by Affymetrix microarray. Pearson's  $R = 0.860$ . **c**, Overlap between genes showing positive or negative fold change of  $\log_2 = 0.5$  or greater with topotecan and irinotecan treatment. **d**, Overlap between genes reduced or increased in expression. **e**, Overlap between differentially expressed genes that are greater or less than 67 kb.



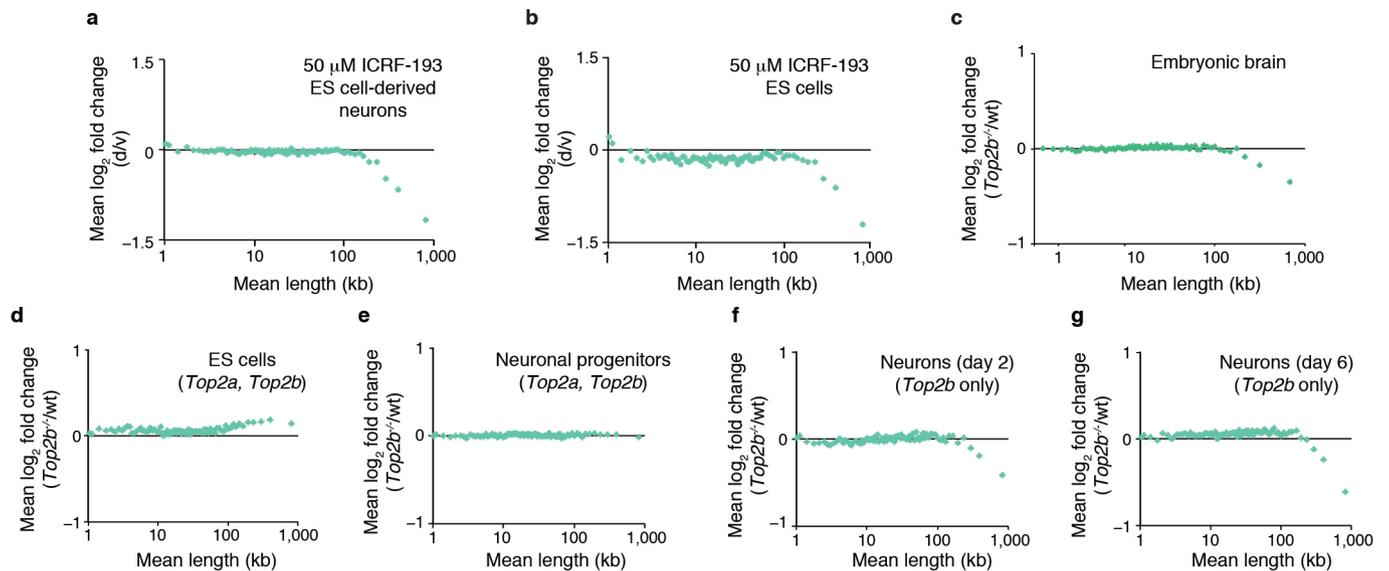
**Extended Data Figure 5 | TOP1 and TOP2 inhibitors reduce expression of long genes in human cell lines.** Re-analysis of microarray gene expression data sets from other laboratories. All plots are mean fold change in expression compared to vehicle controls in bins of 200 genes, plotted by average gene length. **a**, MCF7 cells treated with 10  $\mu$ M irinotecan for 24 h, from the CMAP2 project<sup>44</sup>. **b**, MCF7 cells treated with 165 nM SN38, the active metabolite of irinotecan, for 6 h<sup>45</sup>. **c–e**, Gene expression in three human cell lines treated with

camptothecin. **c**, MCF7 cells treated for 24 h with 10  $\mu$ M camptothecin, from CMAP2. **d**, 293T cells treated with 2  $\mu$ M camptothecin for 48 h<sup>46</sup>. **e**, HeLa cells treated with 10  $\mu$ M camptothecin for 8 h<sup>47</sup>. **f, g**, Re-analysis of microarray data from ref. 26, comparing gene expression in TOP2 inhibitor- and vehicle-treated ME16C cells. **f**, ME16C cells treated with 0.5  $\mu$ M doxorubicin for 36 h. **g**, ME16C cells treated with 50  $\mu$ M etoposide for 36 h.



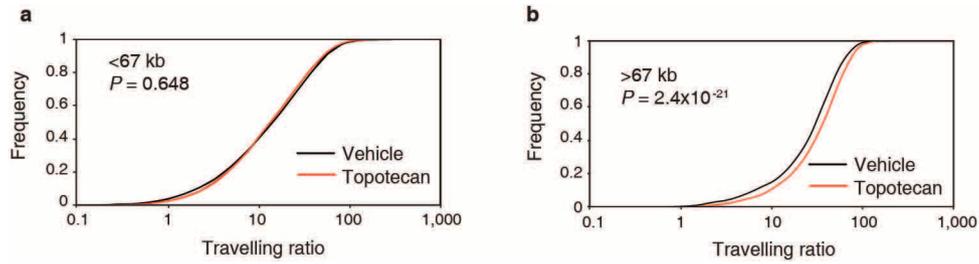
**Extended Data Figure 6 | Topotecan and the TOP2 inhibitor ICRF-193 have similar effects on gene expression.** **a**, Affymetrix microarray analysis of gene expression in cultured mouse cortical neurons treated with vehicle or 3  $\mu$ M ICRF-193 ( $n = 3$  biological replicates), an inhibitor of TOP2 enzymes, for 3 days. Mean expression fold change in bins of 200 genes, plotted by average gene length. **b**, Scatterplot of fold change with 300 nM topotecan ( $n = 3$

biological replicates) versus fold change with 3  $\mu$ M ICRF-193 for all expressed genes, measured by Affymetrix microarray. Pearson's  $R = 0.588$ . **c**, Overlap between genes showing positive or negative fold change of  $\log_2 = 0.5$  or greater with topotecan and ICRF-193 treatment. **d**, Overlap between genes reduced or increased in expression. **e**, Overlap between differentially expressed genes that are greater or less than 67 kb in length.



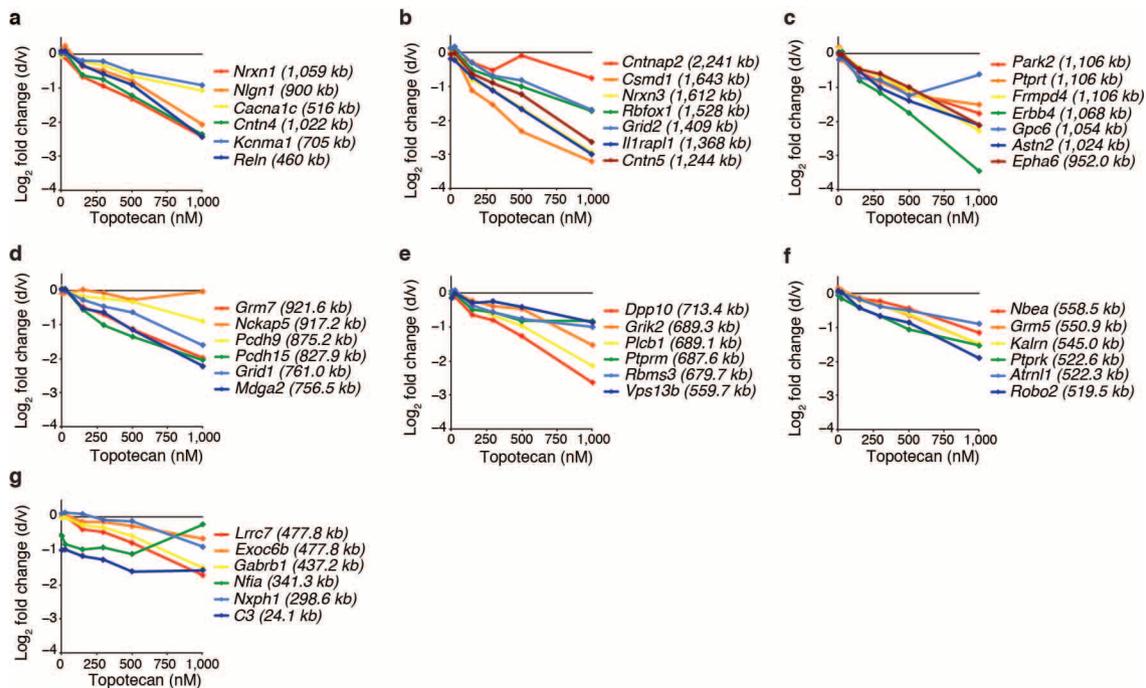
**Extended Data Figure 7 | Pharmacological inhibition of TOP2 or genetic deletion of *Top2b* reduces expression of long genes in ES-cell-derived neurons.** **a, b**, Re-analysis of microarray expression data from ref. 25. Mean fold change in expression in bins of 200 genes, plotted by average gene length. **a**, Gene expression in ES-cell-derived neurons treated with vehicle or 50  $\mu$ M ICRF-193 for 3 days. **b**, Gene expression in ES cells treated with vehicle or 50  $\mu$ M ICRF-193 for 3 days. **c**, Re-analysis of microarray expression data from ref. 14, comparing gene expression in whole brain from wild-type (WT) and *Top2b*<sup>-/-</sup> embryonic mice. Expression data from three developmental time points (E16.5, E17.5, E18.5) were averaged for each gene then plotted as mean fold change in expression between wild-type and *Top2b*<sup>-/-</sup> mice (in bins of

100 genes, by average gene length). **d–g**, Re-analysis of microarray expression data from ref. 25. Mean fold change in gene expression between wild-type and *Top2b*<sup>-/-</sup> cells, for bins of 200 genes, plotted by average gene length. Expression data from ES cells (**d**), neuronal progenitors (**e**), ES-cell-derived neurons, 2 days after plating of neuronal progenitors (**f**) and ES-cell-derived neurons, 6 days after plating of neuronal progenitors (**g**). ES cells and neuronal progenitors express *Top2a* and *Top2b* (indicated within parentheses) and do not show reduced expression of long genes when *Top2b* is knocked out, suggesting that *Top2a* and *Top2b* redundantly regulate the expression of long genes.



**Extended Data Figure 8 | Topotecan increases travelling ratio in genes larger than 67 kb.** **a, b,** Travelling ratio was calculated for all genes bound by Pol II in vehicle- and topotecan-treated (300 nM for 3 days) cultured cortical neurons. 92.1% of all genes had travelling ratio values greater than 2, consistent with a previous report using mouse ES cells<sup>30</sup>. Frequency of travelling ratios for

genes less than 67 kb in length (**a**) and for genes greater than 67 kb in length (**b**). Mean travelling ratio of vehicle- and topotecan-treated samples was significantly different for genes greater than 67 kb ( $P = 2.4 \times 10^{-21}$ , Student's *t*-test) but not for genes less than 67 kb ( $P = 0.648$ , Student's *t*-test).



**Extended Data Figure 9 | Topotecan dose-dependently reduces expression of ASD candidate genes in cortical neurons.** **a**, Topotecan dose-response for high-confidence autism candidate genes ( $n = 3$  for 300 nM topotecan, all other doses  $n = 1$ ). **b–g**, Topotecan dose-response for other ASD candidate genes organized by length from longest to shortest. Mouse cortical neurons were

treated with vehicle or the indicated doses of topotecan for 3 days. Gene expression was quantified using Affymetrix microarrays. Dose-responses are from all topotecan-downregulated ASD candidate genes that were identified by RNA-seq (Table 1) and that were present on the Affymetrix microarrays.