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

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

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 1 (TOPI) 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 \textit{Top1} or \textit{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 TOPI inhibition. Our findings suggest that chemicals and genetic mutations that impair topoisomerases could commonly contribute to ASD and other neurodevelopmental disorders.

### Gene length effects
To determine whether topotecan, a TOPI inhibitor, altered the expression of imprinted genes, we treated cultured cortical neurons from C57BL/6J (B6) × CAST/Ei (CAST) F\textsubscript{1} 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\textsuperscript{19}. 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 \textit{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 \textit{Ube3a}, driving \textit{Ube3a} levels significantly above wild–type levels (Extended Data Fig. 1a, b).

As we found previously\textsuperscript{7}, topotecan reduced the expression of an extremely long (>1 megabase), paternally expressed antisense transcript that overlaps \textit{Ube3a} (\textit{Ube3a-ATS}; Extended Data Fig. 1a, b). The \textit{Ube3a-ATS} transcript is required for paternal \textit{Ube3a} silencing\textsuperscript{20,21}. Other imprinted genes in the same genomic region as \textit{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 \textit{UBE3A-ATS} and increased the expression of \textit{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 \textit{Ube3a} locus in mouse and human neurons. Because \textit{Ube3a-ATS} is extremely long and was strongly downregulated, we hypothesized that topotecan might reduce the expression of
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 $\sim$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 $Ubc3a$, or it might reflect other stimulatory effects of topoisomerase inhibitors.

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$_2$O$_2$) 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 (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, 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. 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, 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 (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. 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$^{-/-}$ mice, which were expressed normally in Top2b$^{-/-}$ ES cells and neuronal progenitors (Extended Data Fig. 7d, e), presumably because these cell types express both Top2a and Top2b. Last, two additional TOP2 inhibitors (doxorubicin and etoposide) reduced the expression of long genes in a human cancer cell line (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 which the topoisomerase gene is not essential for viability but is required for proper expression of long genes.
TOP1 inhibitors can stimulate the release of the positive elongation factor P-TEFb.72 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-β-D-ribofuranosylbenzimidazole (DRB, a P-TEFb inhibitor that can block the transition to elongation73,76). We found that DRB (100 µM) 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 µM and 10 µM 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 studies11,12,33–35 (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 ligase36) and several other genes that directly connect to TOP111,12.

**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 isolation1,2,37. 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 domains48, and the effects of topoisomerases on gene expression are strongly influenced by genomic structure and context23,39. 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.
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\(^{11,40-41}\), 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 \(\times\) 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\(^3\). 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\(^{32,33}\) using RNA Pol II N20 antibody (sc-499, 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. & Buixaum, 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 unsilence 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 maternity 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 mutations underscore the role of the humanльтective nervous system: age and sex dependence. Proc. Natl Acad. Sci. USA 104, E934–E943 (2008).

13. 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).

14. Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition: unexpected roles for chromatin. Nature Rev. Genet. 14, 347–359 (2013).

15. Forberg, E. C., Downs, K. M. & breast. E. H. Direct interaction of NF-E2 with hypersensitive site 2 of the β-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. Kefri 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-DCD), J.S. and T.M. were supported by a grant from the National Institutes of Health (R01GM101974). BL.P was supported by a NIH postdoctoral training grant (T32HD040127). S.J.C. was supported by a Joseph E. Wagstaff Postdoctoral Research Fellowships from the National Institute of Mental Health (R01MH093372). I.F.K. and A.M.M. were supported by the National Institute of Neurological Disorders and Stroke (P30NS045892) and NiCHD (P30HD03110). The manuscript was returned to the authors.

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 iPSC-derived human neurons. B.L.P. assessed propensity of compounds to kill cells. I.F.K., H.-S.H., A.M.M., J.S.H., S.J.C., B.D.P. and M.J.Z. analysed data from genome-wide and designed experiments. I.F.K. performed RNA-seq and ChIP-seq experiments with iPSC-derived human neurons. B.L.P. assessed propensity of compounds to kill cells. I.F.K., H.-S.H., A.M.M., J.S.H., S.J.C., B.D.P. 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. 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) or M.J.Z. (zylka@med.unc.edu).
**METHODS**

**Mouse cell culture.** Cortical neurons were cultured from E13.5–E15.5 mouse embryos as described. For RNA-seq and ChIP-seq, neurons were seeded on 10-cm diameter poly-D-lysine-treated culture plates at a density of 5–10 × 10^5 per plate. Microarray experiments used 6-well plates seeded at 1 × 10^5 per well. After 7 days in culture, drugs or an equivalent amount of vehicle were added and left in the culture media for 24 or 72 h. For topotecan, irinotecan, DRB and paraquat, vehicle was 0.1% dimethylsulfoxide (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 TRCN0000011894 (5’-CCGCGCTTGTGGTTCCTTTT-3’) and Top2β knockdown used clone TRCN0000070988 (5’-CCGGCTTGTGGTTCCTTTT-3’). GTCTTCTCGAGAAGACAAAGGACAACACAAGGTTTTTG-3’ was used as a control. Neurons were seeded on 24-well plates at 2 × 10^5 per well. After 3 days in culture, cells were treated with lentivirus at a multiplicity of infection of at least 1. Virus expression was assayed using Sytox Green (Invitrogen Molecular Probes). γ-H2AX foci were measured by immunohistochernistry. Primary antibodies were anti-γ-H2AX (Millipore, 1:500 dilution) and anti-NeuN (Millipore, 1:500), used to mark neurons.

**ipSC culture and neuronal differentiation.** Human iPSC was appoved 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. iPSCs were differentiated into forebrain cortical neurons as described with the following modifications: neural progenitors were generated by culturing iPSCs on feeders in N2B27 medium supplemented with noggin (500 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 using Taqman (Life Technologies) gene expression assays for UBE3A (Hs00166580_m1) and UBE3A-ATS (Hs003454279_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 allele expression analysis, informative CAST/B6 SNPs were downloaded from [http://www.sanger.ac.uk/resources/mouse/genomes/](http://www.sanger.ac.uk/resources/mouse/). 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. Cultures (n = 4 per condition) totalling approximately 2 × 10^5 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 μg 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.

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 μM topotecan, 10 μM irinotecan (Sigma), 3 μM ICRF-193 (Santa Cruz Biotechnology), 100 μM DRB (Sigma), 100 μM H2O2 (Fisher Scientific) or 10 μM 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 RNA 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, drugs, 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. Grosche1, B. & Bushman, F. Cell cycle arrest in G2/M promotes early steps of 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 μM topotecan, 10 μM irinotecan (Sigma), 3 μM ICRF-193 (Santa Cruz Biotechnology), 100 μM DRB (Sigma), 100 μM H2O2 (Fisher Scientific) or 10 μM 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 RNA 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, drugs, 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. Grosche1, 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. MacIsaac, J. L., Bogutz, A. B., Morrissy, A. S. & Lefebvre, L. Tissue-specific alternative polyadenylation at the imprinted gene Ube3a regulates allelic usage at Cpg2. Nucleic Acids Res. 40, 1523–1535 (2012).
Extended Data Table 1 | Topotecan does not alter parent-of-origin bias of imprinted genes.

| 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. |

| Maternal Bias | Parent of origin bias (adjusted p value) | Change in bias, topotecan vs. vehicle (adjusted p value) |
|---------------|----------------------------------------|--------------------------------------------------------|
| Ube3a*        | 6.1x10^{-29}                           | 1.2x10^{-20}                                         |
| Mag3          | <1x10^{-100}                           | <1x10^{-100}                                         |
| Rian          | <1x10^{-100}                           | <1x10^{-100}                                         |
| Kcnk10        | <1x10^{-100}                           | <1x10^{-100}                                         |
| Mirg          | <1x10^{-100}                           | <1x10^{-100}                                         |
| H13           | 3.0x10^{-06}                           | 3.3x10^{-07}                                         |
| Efllc2        | 2.7x10^{-08}                           | 6.3x10^{-08}                                         |
| H19           | 1.3x10^{-09}                           | 6.9x10^{-09}                                         |
| Hft1          | 1.2x10^{-14}                           | 5.3x10^{-13}                                         |
| Cdkn1c        | 0.00035                                | 9.1x10^{-6}                                          |
| Igf2r         | 0.0056                                 | 0.00019                                               |

| Paternal Bias | Parent of origin bias (adjusted p value) | Change in bias, topotecan vs. vehicle (adjusted p value) |
|---------------|----------------------------------------|--------------------------------------------------------|
| Ncln*         | <1x10^{-100}                           | 7.1x10^{-51}                                         |
| AK0485535*    | <1x10^{-100}                           | <1x10^{-100}                                         |
| Snurf*        | <1x10^{-100}                           | <1x10^{-100}                                         |
| Smpr*         | 2.3x10^{-09}                           | <1x10^{-100}                                         |
| A330076H08Rik*| 2.2x10^{-08}                           | <1x10^{-100}                                         |
| A230057D06Rik*| 6.9x10^{-07}                           | <1x10^{-100}                                         |
| Ube3a-ATS*    | 2.1x10^{-03}                           | 1.2x10^{-56}                                         |
| Doks4*        | 1.3x10^{-45}                           | 3.4x10^{-40}                                         |
| Mmm3*         | 6.4x10^{-18}                           | 1.3x10^{-19}                                         |
| AK086712*     | 4.2x10^{-13}                           | 1.9x10^{-10}                                         |
| U08083*       | 1.7x10^{-9}                            | 6.9x10^{-10}                                         |
| Mageo2*       | 1.0x10^{-9}                            | 0.00065                                               |
| Peg12*        | 0.00032                                | 0.0040                                                |
| AK039108*     | 0.0011                                 | 1.0x10^{-10}                                         |
| Peg13         | <1x10^{-100}                           | <1x10^{-100}                                         |
| Peg3          | <1x10^{-100}                           | <1x10^{-100}                                         |
| Peg3as        | <1x10^{-100}                           | <1x10^{-100}                                         |
| Plagl1        | <1x10^{-100}                           | <1x10^{-100}                                         |
| Impact        | <1x10^{-100}                           | <1x10^{-100}                                         |
| Rsgn1*        | <1x10^{-100}                           | <1x10^{-100}                                         |
| Zim3          | <1x10^{-100}                           | <1x10^{-100}                                         |
| Innp5f        | <1x10^{-100}                           | 1.3x10^{-35}                                         |
| Upp29         | 4.4x10^{-09}                           | <1x10^{-100}                                         |
| Spice         | 4.2x10^{-95}                           | <1x10^{-100}                                         |
| Mest          | 2.5x10^{-78}                           | 1.2x10^{-35}                                         |
| Peg10         | 6.2x10^{-77}                           | 3.0x10^{-39}                                         |
| Zdbf2         | 9.1x10^{-52}                           | 1.8x10^{-45}                                         |
| Zrat1         | 3.2x10^{-46}                           | 1.7x10^{-42}                                         |
| Nap115        | 4.1x10^{-64}                           | 9.2x10^{-78}                                         |
| Kcnq1at1      | 1.2x10^{-38}                           | 1.0x10^{-73}                                         |
| Kcnq1         | 2.1x10^{-06}                           | 1.3x10^{-74}                                         |
| Adam23        | 1.8x10^{-13}                           | 0.00010                                              |
| Grntr10       | 4.0x10^{-13}                           | 4.3x10^{-37}                                         |
| Dlk1          | 6.8x10^{-5}                            | 4.3x10^{-6}                                          |
| Cog2p†        | <1x10^{-100}                           | <1x10^{-100}                                         |
| AF217545†     | <1x10^{-100}                           | <1x10^{-100}                                         |
| Cdh15         | 2.6x10^{-09}                           | 3.7x10^{-13}                                         |
| Boc2f1†       | 0.028                                  | 0.0021                                               |

*Genes in the Ube3a imprinted cluster
†AF217545 is contained within Cog2p, which has also been reported to be maternally expressed.*
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 in mouse cortical neurons ± 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 μM topotecan or vehicle for 6 days. Expression quantified by qPCR. **P < 0.01, one-way analysis of variance with Dunnet’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 H₂O₂ or 10 μM paraquat, compared to vehicle-treated controls. Error bars represent s.e.m. n = 4. b, Average number of γ-H2AX foci per nucleus for cultured cortical neurons treated with 300 nM topotecan for 72 h, 100 μM H₂O₂ 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 H₂O₂ 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 μM 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 μM topotecan (n = 6 biological replicates) versus fold change with 10 μM 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 log2 = 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 μM irinotecan for 24 h, from the CMAP2 project.
- **b**, MCF7 cells treated with 165 nM SN38, the active metabolite of irinotecan, for 6 h.
- **c–e**, Gene expression in three human cell lines treated with camptothecin:
  - **c**, MCF7 cells treated for 24 h with 10 μM camptothecin, from CMAP2.
  - **d**, 293T cells treated with 2 μM camptothecin for 48 h.
  - **e**, HeLa cells treated with 10 μM camptothecin for 8 h.
- **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 μM doxorubicin for 36 h.
  - **g**, ME16C cells treated with 50 μM 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 μM 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 μM 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 log₂ = 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 µM ICRF-193 for 3 days. b, Gene expression in ES cells treated with vehicle or 50 µM 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^{−/−} 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^{−/−} 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^{−/−} 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 cells30. 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.