



OPEN

## Transcription-associated DNA DSBs activate p53 during hiPSC-based neurogenesis

Nadine Michel<sup>1,2</sup>, Heather M. Raimer Young<sup>2</sup>, Naomi D. Atkin<sup>2</sup>, Umar Arshad<sup>2</sup>, Reem Al-Humadi<sup>2</sup>, Sandeep Singh<sup>2</sup>, Arkadi Manukyan<sup>2</sup>, Lana Gore<sup>4</sup>, Ian E. Burbulis<sup>2,3</sup>, Yuh-Hwa Wang<sup>2</sup> & Michael J. McConnell<sup>4</sup>✉

Neurons are overproduced during cerebral cortical development. Neural progenitor cells (NPCs) divide rapidly and incur frequent DNA double-strand breaks (DSBs) throughout cortical neurogenesis. Although half of the neurons born during neurodevelopment die, many neurons with inaccurate DNA repair survive leading to brain somatic mosaicism. Recurrent DNA DSBs during neurodevelopment are associated with both gene expression level and gene length. We used imaging flow cytometry and a genome-wide DNA DSB capture approach to quantify and map DNA DSBs during human induced pluripotent stem cell (hiPSC)-based neurogenesis. Reduced p53 signaling was brought about by knockdown (p53<sup>KD</sup>); p53<sup>KD</sup> led to elevated DNA DSB burden in neurons that was associated with gene expression level but not gene length in neural progenitor cells (NPCs). Furthermore, DNA DSBs incurred from transcriptional, but not replicative, stress lead to p53 activation in neurotypical NPCs. In p53<sup>KD</sup> NPCs, DNA DSBs accumulate at transcription start sites of genes that are associated with neurological and psychiatric disorders. These findings add to a growing understanding of how neuronal genome dynamics are engaged by high transcriptional or replicative burden during neurodevelopment.

Evolution has devoted significant cellular resources to genome maintenance and the prevention of somatic mutations. Numerous DNA repair pathways exist to counteract the myriad of DNA lesions brought about by ongoing replication, transcription, and metabolism<sup>1–4</sup>. DNA damage signaling pathways promote DNA repair by engaging cell cycle checkpoints that stall cell cycle progression until repair is complete<sup>5</sup>. When DNA repair fails, persistent DNA damage signaling often initiates cell death<sup>6</sup>. Following from this general mechanistic model, DNA repair mechanisms have been thought of as cellular caretakers. DNA damage signaling mechanisms, in turn, play the role of cellular gatekeepers protecting an individual from potentially tumorigenic somatic mutations<sup>7–9</sup>. The tumor suppressor protein p53 is a prototypical gatekeeper, sometimes referred to as “the guardian of the genome”<sup>10,11</sup>.

Human development requires hundreds of billions of cell divisions to reach and maintain a steady-state level of ~ 10<sup>15</sup> cells in an adult. Human neocortical development is additionally constrained by an absence of neurogenesis after birth<sup>12,13</sup>. The human cerebral cortex is comprised of 15–20 billion neurons that are among the longest-lived and most diverse mammalian cells<sup>14</sup>. These arise from a discrete neural progenitor cell (NPC) pool in the ventricular zone of the developing cerebral cortex<sup>15</sup>. NPCs overproduce neurons by ~ twofold, thus requiring at least 40 billion NPC divisions during a few months of development<sup>16–20</sup>. This presents an enormous burden for genome maintenance, notably in the NPC response to DNA double strand breaks (DSBs).

DNA DSBs are potent somatic mutagens. Mouse mutants that lack nonhomologous end-joining (NHEJ) DNA DSB repair proteins exhibit extensive NPC death that precludes neurodevelopment and perinatal survival<sup>21–24</sup>. However, NHEJ-deficient mice survive in a p53-deficient background, albeit with an increased occurrence of medulloblastoma and complex karyotypes<sup>25,26</sup>. A more recent study further demonstrates a clear role for p53 in the removal of aneuploid cells during mouse cortical development<sup>27</sup>.

<sup>1</sup>Neuroscience Graduate Program, University of Virginia School of Medicine, 1340 Jefferson Park Avenue, Charlottesville, VA 22908, USA. <sup>2</sup>Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, 1340 Jefferson Park Avenue, Charlottesville, VA 22908, USA. <sup>3</sup>Sede de la Patagonia, Facultad de Medicina y Ciencias, Universidad San Sebastián, Puerto Montt, Chile. <sup>4</sup>Lieber Institute for Brain Development, 855 N. Wolfe St., Ste. 300, Baltimore, MD 21205, USA. ✉email: mikemc@libd.org

Genetic neurodegenerative diseases are brought about by deficits in genome maintenance<sup>28,29</sup>. Motivated in part by many new findings of somatic mutations in neurons<sup>30–35</sup>, novel approaches to identify recurrent sites of DNA damage in neurons have been developed<sup>36</sup>. In addition to primary, typically post-mortem, tissue, human induced pluripotent stem cell (hiPSC)-based neurogenesis has proven to be a concordant and tractable model system. One DNA DSB mapping approach, high-throughput, genome-wide, translocation sequencing (HTGTS), uses a genome maintenance-deficient genetic background (e.g., XRCC4<sup>-/-</sup>P53<sup>-/-</sup>) to slow DNA DSB resolution dynamics and create endogenous “prey” sites that are captured at engineered “bait” sites via chromosomal translocation<sup>37</sup>. Further pharmacological and patient hiPSC-based interrogation using HTGTS finds that both DNA replication and gene transcription are endogenous sources of recurrent DNA DSBs<sup>38–40</sup>. Gene expression level and gene length also predispose subsets of neural genes to recurrent DNA DSBs<sup>41–43</sup>. Neuronal genes with this predisposition for somatic mutation are enriched for associations with neurodevelopmental and neuropsychiatric disease<sup>44</sup>; thus, these observations point to distinct, unresolved aspects of neuronal genome dynamics that potentially mediate neurogenetic architecture through brain somatic mosaicism.

Previously, we developed imaging flow cytometry (IFC)<sup>45</sup> to quantify DNA DSBs with single cell resolution during hiPSC-based neurogenesis. This study revealed increasing levels of DNA DSBs during the first 2 weeks of hiPSC-based neurogenesis prompting us to further investigate neurogenic DNA damage signaling. Here, we sought to limit, rather than abrogate, p53 activity during this time and generated neurotypical NPCs that express short-hairpin RNA (shRNA) targeting p53 (p53<sup>KD</sup>) or a scrambled shRNA control. Marked increases in proliferation and survival accompanied increased DNA DSB burden during p53<sup>KD</sup> neurogenesis. Genome-wide DNA DSB mapping<sup>46</sup> identified an acute (24 h) DNA DSB increase in p53<sup>KD</sup> NPCs that was associated with gene transcription, but not gene length. Interestingly, the only additive effect of replication stress (via aphidicolin) was observed in control NPCs after 4 days. This effect was associated with both gene transcription and gene length, and was mitigated by p53<sup>KD</sup>. Collectively, our work demonstrates a role for p53 signaling in the rapid resolution of transcription-associated DNA DSBs that are incurred during neurotypical human neurogenesis.

## Results

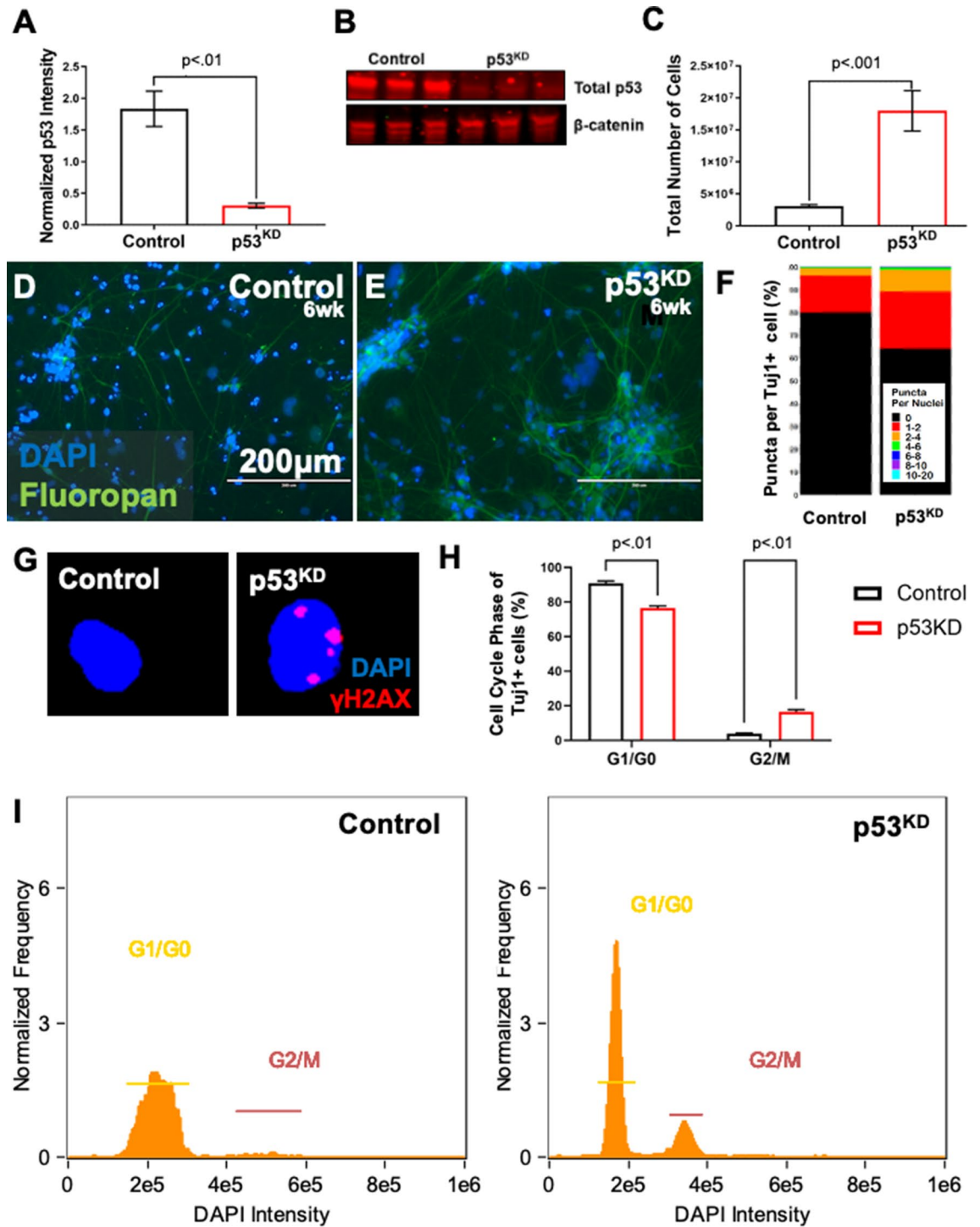
**p53<sup>KD</sup> elevates DNA DSB burden in hiPSC-derived neurons.** Genomic loci that frequently incur DNA DSBs, such as fragile sites, are prone to chromosomal recombination<sup>47,48</sup>. Neuronal genes that harbor sites of recurrent DNA DSBs have been put forward as a mechanistic link to the consequences of brain somatic mosaicism<sup>40</sup>. We probed the role of p53 during hiPSC-based neurogenesis using a knockdown (KD) approach. NPCs from two neurotypic hiPSC lines (BJ and 9429, Fig. S1) were transduced to express a scramble short hairpin RNA (shRNA) or shRNA targeting p53 transcripts; p53<sup>KD</sup> markedly diminished p53 protein relative to shRNA controls (Figs. 1A,B, S7A). When differentiated to cortical neuron fates, an abundance of cells was observed in p53<sup>KD</sup> cultures relative to controls after 6 weeks (Fig. 1C–E).

DNA DSB levels were quantified at 6 weeks of hiPSC-based neurogenesis using IFC<sup>45,49,50</sup>. Immunoreactive  $\gamma$ H2aX puncta were counted in single  $\beta$ -III tubulin-positive cells (i.e., neurons) using the TUJ1 antibody<sup>51</sup>. p53<sup>KD</sup> cultures had more neurons with DNA DSBs (Fig. 1F,G). Given the increase in cell number observed in p53<sup>KD</sup> cultures, we examined whether TUJ1-positive p53<sup>KD</sup> cells were still mitotic. In control cultures, neurons were predominately in G1/G0 as expected; however, p53<sup>KD</sup> cultures had a significant increase of TUJ1-positive cells in G2/M (Fig. 1H,I). This suggests that not all p53<sup>KD</sup> TUJ1-positive cells are fully post-mitotic neurons.

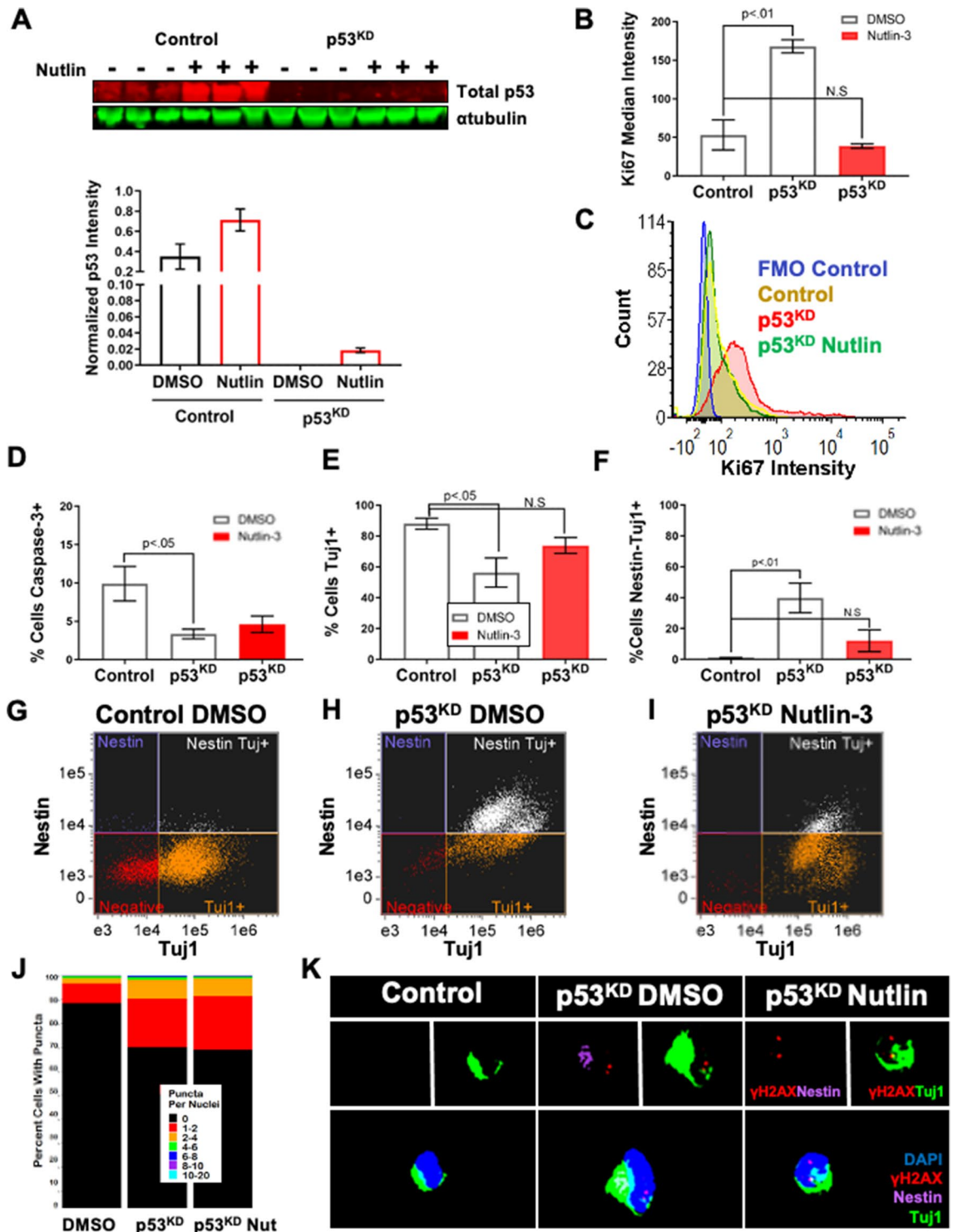
**DNA DSBs accumulate upstream of p53<sup>KD</sup> phenotypes.** Rescue studies were performed to identify p53<sup>KD</sup> phenotypes that were dependent on p53 signaling. Neural differentiation experiments were supplemented with the small molecule Nutlin-3 which promotes p53 activation by inhibiting MDM2<sup>52</sup>. Notably, Nutlin treatment was lethal to control NPCs (Fig. S2), as observed in other cell types<sup>53,54</sup>. A modest increase in total p53 levels was observed in Nutlin-treated p53<sup>KD</sup> NPCs (Figs. 2A, S7B). Concordant with an increased cell number observed in p53<sup>KD</sup> after 6 weeks of differentiation (Fig. 1), after 1 week p53<sup>KD</sup> showed higher median Ki67 staining (Fig. 2B,C) and reduced levels of active Caspase 3 (Fig. 2D) relative to control. Diminished levels of TUJ1-positive cells in p53<sup>KD</sup> were replaced by a large population of cells that expressed both TUJ1 and the NPC marker Nestin (Fig. 2E–H). Nestin is neurofilament specific to NPCs<sup>55,56</sup> and double positives indicate a transitioning population of NPCs that maintain proliferative potential<sup>45,57</sup>. Nutlin treatment restored Ki67 and TUJ1 phenotypes to control levels (Fig. 2B,C,E,I) and dramatically reduced the TUJ1/Nestin population (Fig. 2F,I). DNA DSB burden in p53<sup>KD</sup> is more prevalent after 1 week relative to 6 weeks (Fig. 1F) of differentiation, and is unaffected by Nutlin treatment (Fig. 2J,K).

**Elevated DNA DSB burden in p53<sup>KD</sup> NPCs is associated with gene expression level.** Additional DNA DSBs in p53<sup>KD</sup> NPCs likely arise due to endogenous stressors like transcription or replication. To address this question, we mapped unrepaired DNA DSBs with single nucleotide resolution using DNA DSB Capture with modifications<sup>46,58</sup>. Wang, et al. linked unrepaired DNA DSBs to NPC replicative stress and additional somatic mosaicism at autism risk genes using HTGTS to identify DNA DSB location after DNA DSBs were resolved via chromosomal translocation<sup>39</sup>. Their measurement was taken after 4 days of NPC expansion when treated with the DNA polymerase inhibitor aphidicolin (APH).

We reasoned that a p53 response may have faster kinetics, therefore we performed DSB mapping on control and p53<sup>KD</sup> NPCs after both 1 and 4 days APH treatment. Biological replicates showed strong concordance and were pooled for subsequent analysis (Tables S1, S3A). Captured reads were mapped and DNA DSB location was assessed in 5 categories of genomic regions (promoter, transcription start sites (TSS), gene body, transcription termination sites (TTS), and intergenic) (Fig. S4A,B). DSB density was similarly distributed among all categories and cell conditions. Notably, across all conditions the TSS had the highest percentage of DSB density (Fig. S4A,B).



**Figure 1.** Neural p53<sup>KD</sup> increases cell number, DNA DSBs, and mitotic activity. (**A,B**) Marked reduction in p53 protein in p53<sup>KD</sup> neurogenesis (Unpaired t-test,  $p < .01$ ). Blot is cropped, original blot presented in Supplementary Figure S7A. (**C**) After 6 weeks of neuronal differentiation, p53<sup>KD</sup> cultures have >threefold more cells than control (Error Bars = Mean  $\pm$  SEM, unpaired t-test,  $p < .001$ ). (**D,E**) Representative images of Control and p53<sup>KD</sup> neuronal cultures stained with DAPI and Fluoropan (Neuron Specific Cocktail antibody). (**F**) Pooled stacked histogram of  $\gamma$ H2AX puncta per nucleus from two neurotypic cell lines (BJ, 9429). p53<sup>KD</sup> neurons (Tuj1+) have significantly more DNA DSBs ( $N = 47,406$ , Chi-squared test,  $p < .0001$ ). (**G**) Representative IFC images of neurons with  $\gamma$ H2AX foci. (**H,I**) p53<sup>KD</sup> (Tuj1+) vs Control have fewer cells in G1/G0 (Two-way ANOVA, Tukey's Test  $p < .01$ ), and more cells in G2/M (Two-way ANOVA, Tukey's Test  $p < .01$ ).



**Figure 2.** Nutlin treatment rescues p53<sup>KD</sup> phenotypes in early neurogenesis. Nutlin treatment rescues delayed neurogenesis phenotypes in p53<sup>KD</sup> at 1 week of neuronal differentiation. (A) One week treatment with 15 μM Nutlin-3 partially restores total p53 levels in p53<sup>KD</sup> neural cells. Blot is cropped, original blot presented in Supplementary Figure S7B. (B,C) p53<sup>KD</sup> neuronal cultures have more Ki67+ cells and nutlin treated p53<sup>KD</sup> (One-way ANOVA, Dunnett’s Test  $p < .01$ ). (D) p53<sup>KD</sup> have less cell death vs controls (One-way ANOVA, Dunnett’s Test,  $p < .05$ ). (E) Control cells had the highest percentage of Tuj1+ cells vs p53<sup>KD</sup> cells (One-way ANOVA, Dunnett’s test,  $p < .05$ ). No significant difference between controls vs Nutlin p53<sup>KD</sup> (One-way ANOVA, Dunnett’s Test  $p = .10$ ). (F) Nestin+/Tuj1+ cells were strikingly enriched in p53<sup>KD</sup> cultures (One-way ANOVA, Dunnett’s Test  $p < .01$ ). These account for the Tuj1+ cells in G2/M from Fig. 1. p53<sup>KD</sup> cells treated with DMSO had significantly more double positive cells, nutlin treatment significantly reduced this population (One-way ANOVA, Dunnett’s Test,  $p = .51$ ). (G,H,I) Representative imaging flow cytometry plots of samples of control, p53<sup>KD</sup>, and p53<sup>KD</sup> treated with Nutlin. (J) Nutlin treatment did not reduce DNA DSB level in p53<sup>KD</sup> vs control ( $p < .0001$ , KS test) ( $p < .0001$ , KS test). (K) Representative images from imaging flow cytometry.



In other cell types the TSS of highly expressed genes incurs DNA DSBs<sup>46,59–61</sup>, therefore, we examined this in NPCs. Relative gene expression levels were binned into 10 subsets ( $n = 25,420$  total genes) from control scRNA seq data reprocessed as bulk (Fig. S5E). When we fine-mapped to TSS  $\pm 2$  kb, accumulation of DSBs was observed in the top 10% of expressed genes compared to the bottom 10% of expressed genes in both WT and p53<sup>KD</sup> across treatment conditions (Fig. 3A,B). Consistent with elevated DNA DSB burden in p53<sup>KD</sup> neurons (Figs. 1F,G and 2J,K), we captured significantly more DNA DSBs at TSSs in p53<sup>KD</sup> NPCs compared to control, regardless of APH treatment, after 24 h of NPC expansion (Fig. 3C). However, no difference was observed as a result of p53<sup>KD</sup> after 4 days (Fig. 3D).

Many neuronally-expressed genes are among the longest genes in the human genome<sup>43,44</sup>. Long genomic loci are transcribed into S-phase leading to potential collisions between RNA and DNA polymerases and ensuing DNA DSBs<sup>42</sup>. However, we found no effect of p53<sup>KD</sup> on DNA DSB location with respect to gene length at either time point (Fig. S3B, C). An association between DNA DSBs and gene length was observed in control cultures treated with APH for 4 days (Fig. S3B), as described in previous studies<sup>39</sup>. Importantly, our DSB mapping approach is both genome-wide and unbiased; it is not restricted to translocations with defined bait sites as in HTGTS. This finding strengthens accumulating evidence that NPC replication stress leads to recurrent DNA DSB sites at long genes. However, p53<sup>KD</sup> precludes the effect of 4 day APH treatment at both long and highly expressed genes (Figs. S3C and 3D).

Given no additive effect of APH treatment on DNA DSB levels in p53<sup>KD</sup> NPCs and no association between gene length and recurrent p53<sup>KD</sup> DNA DSBs, these data indicate that p53 may mediate an acute (24 h) response to DNA DSBs associated with high transcriptional activity. And that transcription-associated DNA DSBs in p53<sup>KD</sup> NPCs are resolved at the population level during 3 additional days of NPC expansion. Taken together, genome-wide DNA DSB mapping data suggest that elevated DNA DSB burden in p53<sup>KD</sup> NPCs is due to a diminished p53 response to transcriptional, but not replicative, stressors.

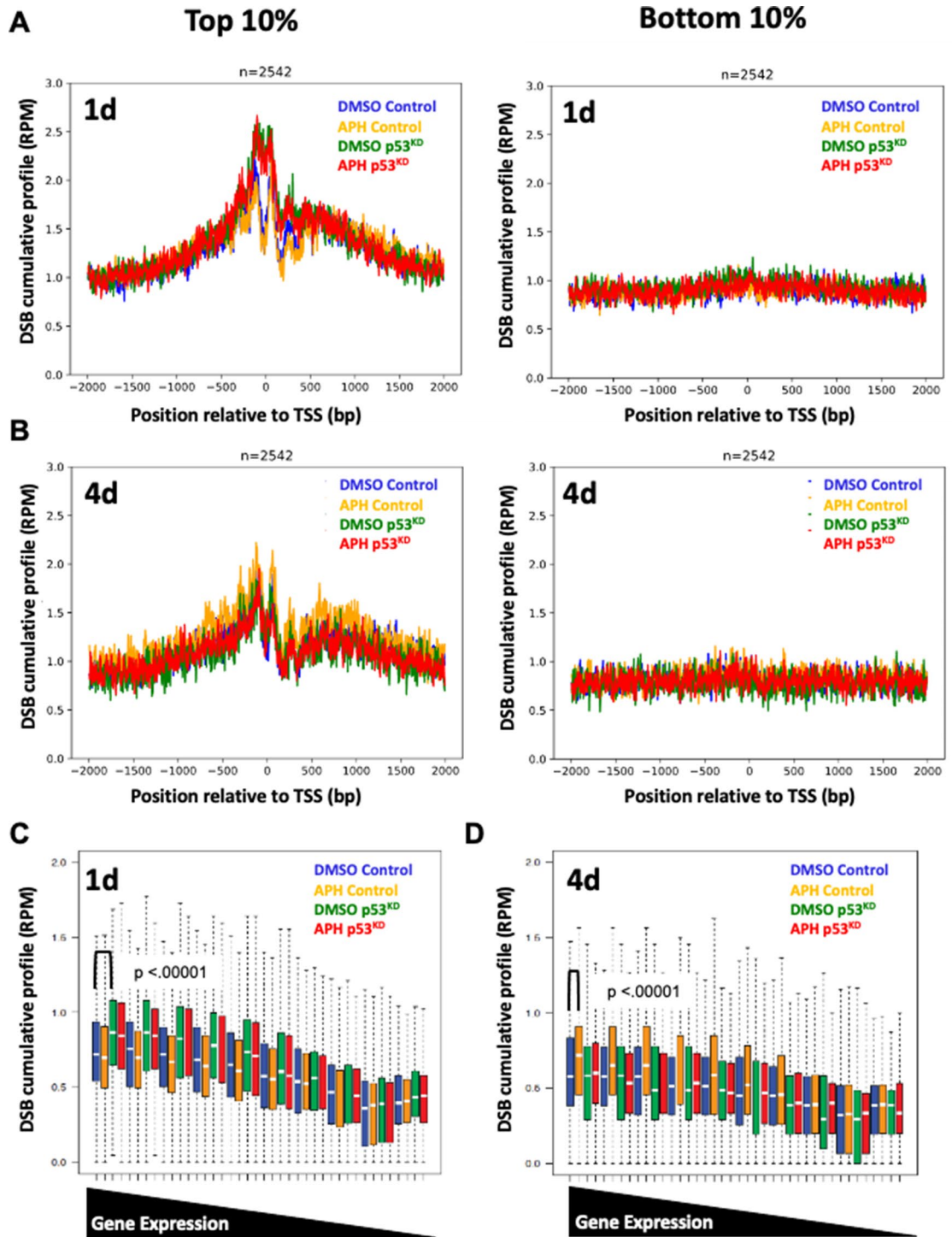
**p53 is activated during TOP1-mediated stress in hiPSC-derived NPCs.** In hyperproliferative autism spectrum disorder patient-derived NPCs, replicative stress activates ATR with minimal p53 activation<sup>39</sup>. Given that p53<sup>KD</sup> leads to an acute increase of recurrent DNA DSBs at TSSs, we sought to define p53 activation in response to replicative or transcriptional stress. We treated control NPCs with APH or with a topoisomerase I inhibitor camptothecin (CPT)<sup>62</sup>.

Topoisomerase I (TOP1) relieves DNA supercoiling due to ongoing transcription and replication by transiently binding to the DNA, forming a cleavage complex (TOP1cc), and then catalyzing single strand cleavage. Inhibition of TOP1 using camptothecin has revealed that TOP1 plays an important role in the elongation stage of transcription at long neural genes that could contribute to somatic mosaicism in neurons<sup>63</sup>. Neuronal deletion of *Top1* in mice leads to an increase in DNA DSBs, elevated somatic mosaicism, and early onset neurodegeneration<sup>64</sup>. Ataxia telangiectasia (A-T) patient-derived neurons are sensitive to CPT treatment<sup>63</sup>. We used identified concentrations of APH and CPT that increased the percentage of  $\gamma$ H2AX-positive NPCs (Fig. S5A). Relative to controls, both APH and CPT treated NPCs had more total DNA DSBs (Figs. 4A, S5B) and significant S-phase arrest (Figs. 4B,C, S5C).

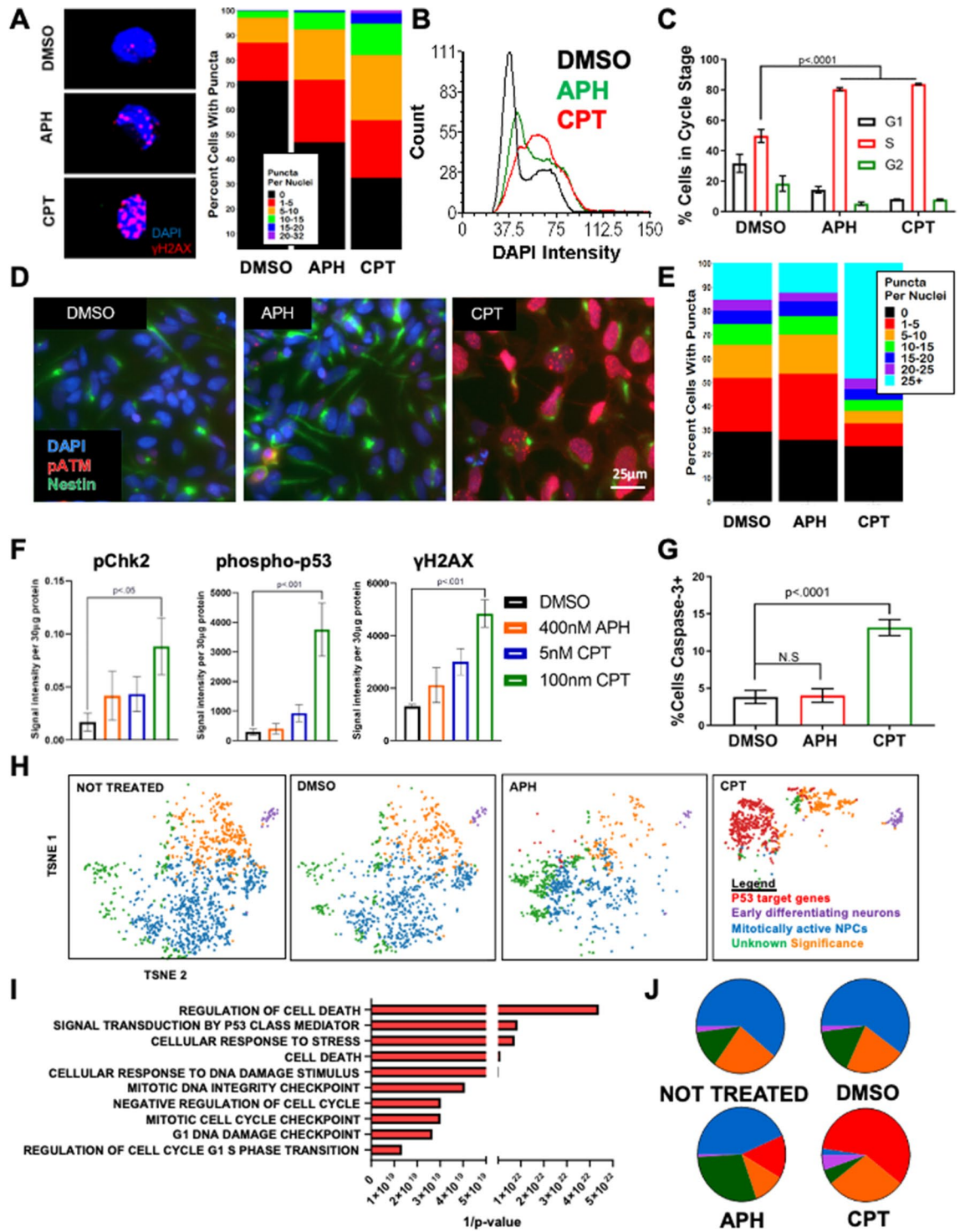
A canonical p53-dependent DNA DSB response includes phosphorylation of ATM, Chk2, and p53<sup>65–67</sup>. CPT-treated NPCs showed robust phosphorylation of ATM compared to controls and APH-treated NPCs (Fig. 4D,E). In addition, transcriptionally stressed NPCs had higher levels of phosphorylated p53, pChk2, and  $\gamma$ H2AX compared to controls and replication stressed NPCs (Figs. 4F, S5D, S7C). NPCs exposed to replication stress had similar levels of cell death (assessed by active Caspase 3) relative to controls, but this population was more prevalent in TOP1 stressed cells (Fig. 4G). This was also observed when CPT and APH treatment led to more similar distributions of  $\gamma$ H2AX puncta per nucleus (Fig. 4A, 9429, Figure S5B).

To further define the NPC response to TOP1 inhibition we performed single-cell RNA sequencing (scRNA-seq). We observed distinct phenotypic clusters present in all conditions (Fig. 4H, S5E). Actively proliferating NPCs (blue cluster) were abundant in the two controls (Not Treated, DMSO) and APH treated cells, but minimally present in the CPT treated NPCs (Figs. 4H,J, S5E). CPT treated cells had a significant emergence of NPCs upregulating genes involved in the DNA damage checkpoint, p53 mediators, and cell death (Fig. 4I). Consistent with increased cell numbers observed during p53<sup>KD</sup> neurogenesis (Fig. 1C), these data confirm that TOP1 inhibition, but not replicative genomic stress, leads to a robust p53-mediated NPC response to DNA DSBs associated with increased cell death.

**Enrichment of DNA DSBs in p53<sup>KD</sup> NPCs at the TSS of genes with clinical risk for neuropsychiatric disorders.** We sought to examine whether p53<sup>KD</sup> increased the prevalence of DNA DSBs in genes with neuronal function or causal links to brain disease. We performed gene ontology (GO) analysis of expressed genes  $> 1.5$  fold enrichment at TSSs in p53<sup>KD</sup> NPCs. Of the 61 genes with enriched breaks, 47% were associated with disorders involving the nervous system (Fig. 5A, left). We analyzed these genes further and observed that 33% were involved in psychiatric disorders, ~15% in neurodegeneration, and more than 40% in intellectual disability and brain development (Fig. 5A, right). Specifically, recurrent DNA DSBs in p53<sup>KD</sup> NPCs were enriched at STAT3 (Multiple Sclerosis<sup>68</sup>), MAN1B1<sup>69</sup>, and STARD9 (Bipolar Disorder<sup>70</sup>) (Fig. 5B, Table 1). We sought to determine whether the enrichment observed in the TSS of p53<sup>KD</sup> NPCs was simply a function of these genes being selectively transcribed during neurogenesis, however, we observed that expression of the identified neuronal genes was not driving DSB enrichment (Fig. S6A–C). Together these data show that a lack of p53 is associated with recurrent DNA DSBs in the TSS of genes that are important for neurotypical brain development.

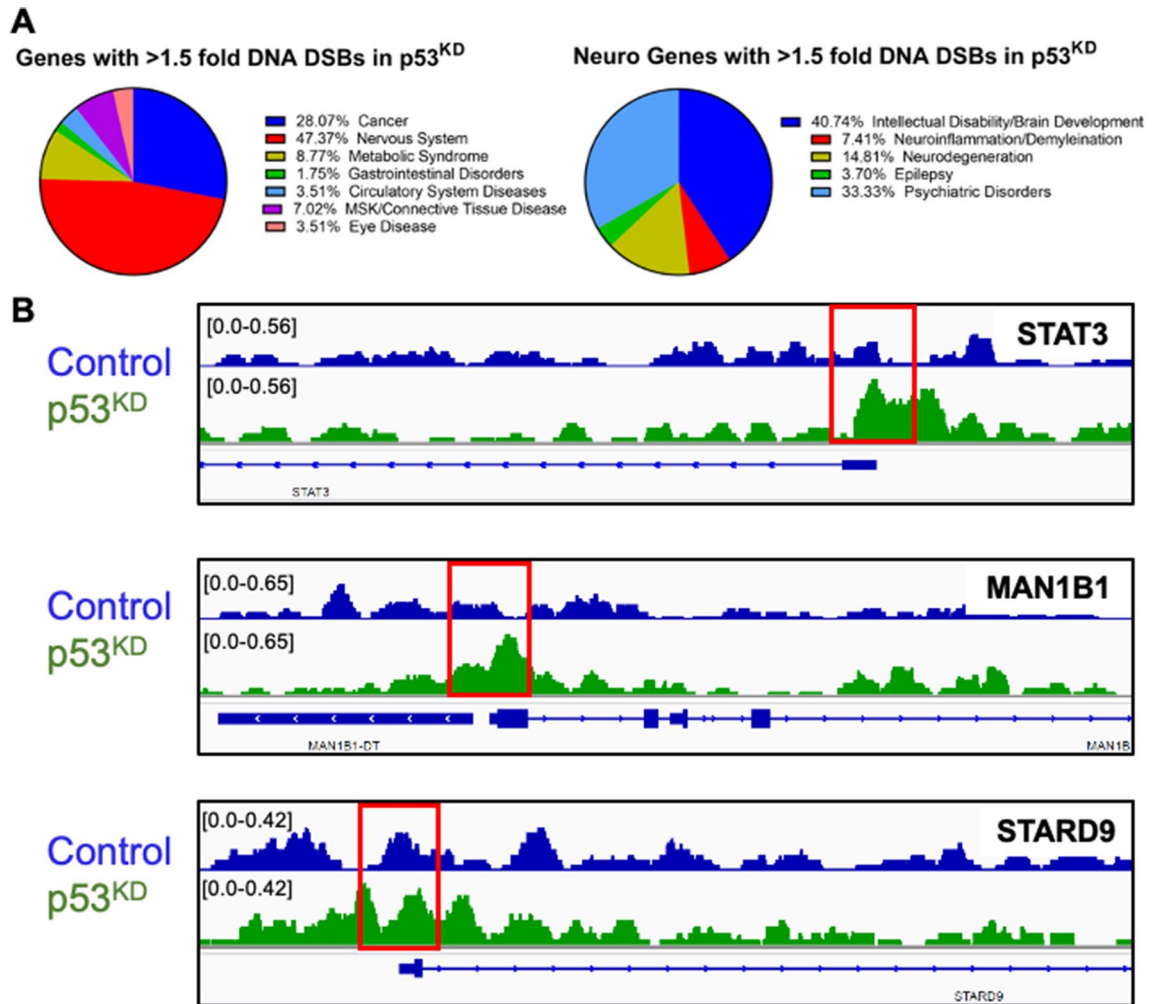


**Figure 3.** p53<sup>KD</sup> leads to an acute increase in TSS-associated DSBs. In highly transcribed genes ( $n = 2542$ ), DNA DSBs are enriched at the TSS of NPCs treated for 1 day (A) and 4 days (B). (C) Elevated DNA DSB levels in p53<sup>KD</sup> NPCs are associated with gene expression levels in 1 day cultures. Boxplot of DSB coverage (TSS  $\pm$  500 bp) in 10 equal bins based on the gene expression. p53<sup>KD</sup> NPCs have more DNA DSBs in highly transcribed genes compared to DMSO Control cells (Wilcoxon signed-rank test, two-sided, paired test,  $p < .00001$ ). (D) After 4 days APH treatment, control NPCs treated with APH now have more DNA DSBs compared to untreated, p53<sup>KD</sup>, and p53<sup>KD</sup> treated with APH in highly transcribed genes (Wilcoxon signed-rank test, two-sided, paired test,  $p < .00001$ ). The box plot shows the 25th and 75th percentile; the middle line is the median; the whiskers span 5% to 95%, and outliers are not shown.



**Figure 4.** TOP1 inhibition elicits a robust p53 response in human NPCs. (A) NPCs treated with either 400 nM APH or 100 nM Camptothecin (CPT) had significantly more DNA DSBs per nucleus (Chi-squared,  $p < .0001$ ,  $N = 63,436$  cells). CPT cells had more DSB/cell (KS test,  $p < .0001$ ,  $N = 63,436$  cells). (B,C) NPCs exposed to either APH or CPT show significant S-phase arrest (Two-way ANOVA, Dunnett’s Test  $p < .0001$ ). (D,E) NPCs have more pATM foci per nucleus after CPT treatment (One-way ANOVA, Dunnett’s Test,  $p < .0001$ ) and less pATM foci per nucleus in replication stress (APH) (One-way ANOVA, Dunnett’s Test,  $p < .01$ ). (F) Western blots identify elevated pChk2 (unpaired t-test,  $p < .05$ ), phosphorylated p53 (unpaired t-test,  $p < .001$ ), and  $\gamma$ H2AX (unpaired t-test,  $p < .001$ ). (G) Caspase-3 positive cells were elevated only in CPT treated NPCs (One-way ANOVA, Dunnett’s Test  $p < .0001$ ) (One-way ANOVA, Dunnett’s Test,  $p = .98$ ). (H) Single-cell RNA seq reveals that CPT-treated NPCs have differential expression of p53 target genes involved in DNA damage response, cell cycle arrest, and cell death ( $N = 4598$ ,  $n$  is number of cells). TSNE plots shown for each condition. TSNE 1 and TSNE 2 are the same for each condition. (I) Functional pathway analysis of the top 100 highest variant genes from the 592 cells in the red cluster (enriched in CPT treatment) are involved cellular processes related to DNA damage checkpoint ( $p < 1.37E-22$ ), cell death ( $p < 1.16E-22$ ), and p53 mediators ( $p < 1.16E-22$ ). (J) The percentage of cells in each cluster per condition is shown in a pie chart.





**Figure 5.** Recurrent sites of p53-susceptible DNA DSBs are associated with neurodevelopment and psychiatric disease. (A) 61 genes had >1.5 fold DNA DSB increase in p53<sup>KD</sup> NPCs vs control and were analyzed for known associations with disease. 47% of these genes are involved in the nervous system<sup>80</sup>. Among them, more than 40% play role in intellectual disability and 33% in psychiatric disorders (right). (B) An IGV plot of three genes associated with disorders of the nervous system showing an increase in DNA DSBs in p53<sup>KD</sup> vs controls are shown (from top to bottom: STAT3, MAN1B1, STARD9). Red boxes highlight the TSS where the differences in DNA DSBs were quantified.

## Discussion

A deeper understanding of human neuronal genome dynamics is motivated by increasing recognition that both neurodevelopmental and neurodegenerative human diseases are linked to genome maintenance. The abnormal accumulation of DNA damage is a common pathological feature among neurodegenerative diseases including Alzheimer's disease<sup>71,72</sup>, Parkinson's disease<sup>73,74</sup>, and Amyotrophic Lateral Sclerosis (ALS)<sup>75,76</sup> and is correlated with normal aging<sup>28,77</sup>. DNA repair is not perfect and somatic mutations accumulate in neurotypical human neurons<sup>78</sup>, higher frequencies of somatic mutations are observed in neurodegenerative disorders<sup>79</sup>. Indeed, genome maintenance is critical in neurons and enhanced DNA repair is a pharmacological target to ameliorate cognitive decline<sup>28,29</sup>. Moreover, mutations affecting DNA repair and damage signaling (e.g., *Atm*) lead to childhood onset neurodegenerative disease (e.g., ataxia-telangiectasia)<sup>80</sup> and have emerged as critical mediators of Huntington's Disease onset<sup>81</sup>.

The tumor suppressor p53 is a central mediator of genome maintenance. Our findings are consistent with observations from different cell lines in which p53 participates in DNA damage response by activation of ATM-pChk2 pathway, cell cycle arrest, and cell death<sup>82-84</sup>. We have shown that p53 is similarly activated when DNA damage occurs in hiPSC-derived NPCs. Reduced p53 levels during hiPSC-based neurogenesis led to elevated levels of DNA DSBs in differentiating neurons. Nutlin treatment rescued p53-dependent phenotypes, but elevated DNA DSB levels persisted. Our mapping of recurrent DNA DSBs resulted in three primary findings supporting the overall conclusion that transcription-associated DNA DSBs activate p53 during hiPSC-based neurogenesis. First, unresolved DNA DSBs map to the TSS of transcribed genes, and these are more abundant in p53<sup>KD</sup>. Second, replication stress had no additive effect on recurrent p53<sup>KD</sup> DNA DSB sites. Third, DNA DSBs brought about by transcriptional stress, but not replication stress, activate p53 during neurotypical hiPSC-based neurogenesis.



Rank	Gene name	DNA DSB fold change	Function	Disease association	Overlapping gene(s)
2	STAT3	3.2	Controls cell proliferation, migration, and apoptosis	Multiple Sclerosis	
3	MAN1B1	3.1	ER protein that disposes misfolded glycoproteins	Rafiq Syndrome (Intellectual Disability, impaired motor movement, facial dysmorphism)	
5	KAT6B	2.9	Histone acetyltransferase	Ohdo Syndrome (Intellectual Disability, ptosis, narrow eye openings)	SNORD172
7	STARD9	2.7	Regulates the assembly and stability of both interphase and spindle microtubules	Bipolar Disorder	
8	VPS35	2.7	Creates mitochondria-derived vesicles, directs degradation of mitochondrial proteins	Parkinson's Disease	
9	FIZ1	2.6	Zinc finger protein that regulates hematopoietic and lymphoid cells	Bardet-Biedl Syndrome (Intellectual Disability, polydactyly, renal anomalies, obesity, retinopathy)	
12	B9D1	2.5	Located in transition zone of cilia and prevents diffusion of proteins	Joubert Syndrome (Cerebellar and Brainstem Malformation)	
14	SGK3	2.3	Regulates ion channels, cell growth, proliferation, membrane transporters, survival, migration	Alzheimer's Disease	
15	SIX5	2.3	Transcription factor	Alzheimer's Disease	
18	CIPC	2.2	Regulates circadian rhythms	Cohen Syndrome (Intellectual Disability, microcephaly, hypotonia, developmental delay)	

**Table 1.** Top 10 genes with > 1.5-fold DNA DSB enrichment in p53<sup>KD</sup> treated for 24 h. The top 10 genes with > 1.5-fold DNA DSB enrichment related to brain disorders are listed from highest to lowest DNA DSB fold change.

Additional support for the veracity of our observations is reported when NPCs are stressed for 4 days with APH. As observed previously using other approaches<sup>39</sup>, we also report an association between gene length and recurrent DNA DSBs after long term replication stress.

Other recent studies suggest divergent roles for p53 in stem cell fate decisions during human pluripotent stem cell-based neurogenesis. Impaired neurogenesis is observed from p53 knockout (p53<sup>-/-</sup>) human embryonic stem cells (hESCs) and from p53KD hiPSCs<sup>85,86</sup>. Disorganized neuroepithelial rosettes were observed in neurodifferentiated p53<sup>-/-</sup> hESCs<sup>86</sup>, and likewise limited neurogenesis was observed in p53KD hiPSC-derived brain organoids<sup>85</sup>. By contrast, exacerbated neurogenesis was observed when Marin Navarro, et al. initiated p53KD after p53-competent hiPSCs were differentiated to neuroepithelial stem cells. Similar to our study, cortical organoid differentiation with p53KD in this stem cell population was found to promote neurogenesis. Despite different approaches (i.e., p53KD in established NPCs (passage 5 and 6) and 6 week 2D cortical neuron differentiation), we also find that p53KD leads to an abundance of TUJ1 positive neurons, a 2 -threefold increase in total cell numbers, and diminished cell death. However, we also observe that many TUJ1-positive cells remain Nestin-positive, and are not post-mitotic. This neural population has been observed during hiPSC-based neurogenesis before<sup>45,87</sup>, and here we show that Nutlin treatment during the first week of p53KD neurogenesis limits prolonged Nestin expression and mitosis in a subset (~20%) of TUJ1-positive cells. Although not conclusively established here, these data are consistent with a model wherein p53 activity promotes cell cycle exit and terminal differentiation of NPCs into post-mitotic neurons.

Collectively, our data indicate a distinct neurodevelopmental role for p53 in the response to transcription-associated DNA DSBs. It has been proposed, based on several lines of evidence, that collisions between DNA and RNA polymerases lead to DNA DSBs and subsequent somatic mosaicism at a subset of neuronal genes that are long and are also risk alleles for psychiatric disease<sup>37,40,42</sup>. Consistent with this mechanism, we find an association between gene length and DNA DSB prevalence in neurotypical NPCs. However, it is only when considered with regard to gene expression level that we observe a differential incidence of DNA DSBs in p53KD NPCs. Furthermore, our DNA DSB fine-mapping approach found increased levels of TSS-associated DNA DSBs in p53KD NPCs. The TSS has previously been shown to have enriched DSB density compared to other genomic regions<sup>46,59-61</sup>. Elevated DNA DSB levels in p53KD could be an indirect consequence of diminished p53-dependent transcriptional activation<sup>88</sup>, or these could arise as a direct consequence of impaired p53 recruitment to sites of DNA damage<sup>89</sup>. Further investigation is required to determine if elevated TSS-associated DNA DSBs are a direct or indirect consequence of p53KD.

Longitudinal molecular genetic and biochemical studies of human neurodevelopment are intractable. Cross-sectional familial and population-based genetic studies collectively conclude that most human neurological disease is polygenic and challenging to interpret. Studies of mouse mutants together with studies of post-mortem human brains have yielded many insights; yet the translation of mouse data to human disease is imperfect. This barrier to progress is brought about both by species-specific brain differences and by the end-stage nature of most human post-mortem tissue. The study of human neurodevelopment in-a-dish is enabled by hiPSC-based neurogenesis. In addition to providing mechanistic insight and drug development platforms for human neurological disease<sup>90-95</sup> hiPSC-based models offer insights into neurotypical neurodevelopment that are uniquely human. Given dramatically different neuronal production requirements and human-specific brain phenotypes, it is not surprising that brain somatic mosaicism is more abundant in human brains<sup>30,32</sup>, relative to mouse<sup>35,96</sup>. A human-specific understanding of the interplay between endogenous genomic stress and genome maintenance responses portends new drug targets for neuropsychiatric and neurodegenerative disease.

## Experimental procedures

**Cell culture.** Fibroblasts from neurotypical individuals were reprogrammed to induced pluripotent stem cells using lentiviral (9429) or non-integrating Sendai (Invitrogen) vectors<sup>12</sup>. NPCs and neurons were generated following methods outlined by Brennand and Livesey<sup>91,97</sup>. Briefly, hiPSC-derived NPCs were plated on Matrigel (Fisher) or Poly-Ornithine (Sigma)/ Laminin (Invitrogen) and passaged 1:6 every 4 days. NPCs (< passage 14) were cultured in NPC medium (DMEM/F12 + Glutamax (Invitrogen), 1 × N2 (Invitrogen), 1X B27-Vitamin A (Invitrogen), 1 µg/ml Laminin (Invitrogen), 20 ng/ml FGF-2 (Peprotech) and dissociated in Accutase (Fisher) for 5 min at 37 °C. hiPSC-derived neurons were cultured on Matrigel in Neuron Medium (DMEM/F12 + Glutamax, 1 × N2, 1X B27 with Vitamin A (Invitrogen), 20 ng/ml BDNF (Shenandoah Biotechnology), 20 ng/ml GDNF (Shenandoah Biotechnology), 1 mM dibutyl-cyclic AMP (Sigma), and 200 nM ascorbic acid (Stem Cell Technologies).

**Drug treatment.** NPCs were exposed to various doses of Aphidicolin (Sigma), Camptothecin (Sigma), and Nutlin-3 (Sigma). DMSO was used to solubilize all drugs and was used as a vehicle control. For 1 day APH treatments, cells were passaged and the next day treated with 0.4 µM. In 4 days APH treatments, cells were passaged and treated 2 days later. Treatment consisted of 3 days of 0.5 µM APH and 1 day of 0.25 µM APH as outlined by Wang et al.<sup>39</sup>. All treatments were added directly to the appropriate cell media. All drugs used in this study were stored appropriately at – 20 °C and aliquoted.

**Immunocytochemistry.** NPCs were cultured for 24 h with various drugs and rinsed with PBS. NPCs at ~ 80% confluency were fixed on 18 mm glass coverslips with 4% PFA for 15 min at room temperature and rinsed (3 times for 5-min washes) in PBS. Cells were permeabilized with 0.5% Triton-X 100 for 10 min and incubated in blocking buffer (2% BSA in PBS-T (0.1% Triton-X 100)) for one hour at room temperature. The cells were probed overnight at 4 °C with anti-phosphorylated ATM (pATM), DAPI, and Nestin followed by labeling with secondary staining and incubated for one hour at room temperature. pATM foci were quantified as described in our previous studies<sup>45</sup>. In short, NPCs were imaged on a fluorescent EVOS Microscope and analyzed through a Cell Profiler (Broad Institute) pipeline that (1) isolates channels from an RGB image, (2) identifies the nuclei<sup>98</sup>, identifies foci within the nucleus, and (4) quantifies the number of foci per nucleus.

**Western blots.** NPCs were dissociated with Accutase, pelleted, flash frozen with liquid nitrogen, and stored at – 80 °C. Cells were lysed with 2 × Laemmli buffer (4% SDS, 10% 2-mercaptoethanol, 20% glycerol, 0.004% bromophenol blue, 0.125 M Tris HCl, pH 6.8) and boiled for 10 min at 95 °C. For immunoblotting, total cell lysate was resolved by 4–15% Mini-PROTEAN precast gradient gel (Biorad). Proteins were transferred to a 0.2 µm nitrocellulose membrane for 90 min at 100 V and incubated at room temperature for 1 h in blocking buffer (1% Fish Gelatin (Sigma) in TBS-T (Tris-buffered Saline 0.1% Tween-20)). The blots were incubated at 4 °C overnight in blocking buffer in primary antibody and incubated for 45 min at room temperature in secondary antibody. Blots were visualized on a Licor Odyssey Imaging System. All gels were run with a positive control (High dose CPT treated HEK 293 T cells). Differences in abundance of protein in loading controls (αtubulin, actin, β-catenin) between samples exposed to genomic stress were observed; therefore, samples were normalized by loading equivalent protein (30 µg) in each lane (as determined through BCA assay). The following well characterized primary antibodies were used for studies included in this publication: Total p53<sup>99</sup>, phosphorylated-p53<sup>98</sup>, pChk2<sup>100</sup>, αtubulin<sup>101</sup>, β-catenin<sup>102</sup>, Total H2AX<sup>103</sup>, and γ-H2AX<sup>104</sup>.

**Single-cell RNA sequencing.** NPCs were exposed to either 400 nM APH, 300 nM CPT, DMSO (control), or no treatment for 24 h. High viability cell samples collected for scRNA-seq using the 10 × Genomics V2 Kit for reverse transcription and library preparation. Sequencing was performed on the Next-seq 500 and FASTQ files were run through the cellranger pipeline and further analyzed through loup browser software and software packages in R program. Gene ontology analysis was performed using Panther. scRNA-seq data can be accessed at the NCBI Sequence Read Archive (SRA), accession number: PRJNA591220.

**Wildtype bulk gene expression analysis.** To evaluate overall gene expression relationship with DSBs, scRNA-seq data from the wildtype cell line was aggregated and processed using the Tuxedo suite<sup>105</sup>. Genes with FPKM > 0 were evenly divided into 10 bins for further analysis with genome-wide DSB mapping and sequencing data.

**P53 knockdown NPCs.** hiPSC-derived NPCs were transfected with a lentivirus containing a puromycin resistant plasmid containing a vector backbone control (pLKO.1) or p53 shRNA containing plasmid (AdGene). After 24 h of exposure to the lentivirus, NPCs were changed to normal NPC media for 24 h and puromycin selection for 4 days at 0.4 µg/mL. Knockdown efficiency was confirmed with Western Blot.

**Cell counts.** Cell counts were performed for every sample to quantify the number of cells remaining in culture and accurately add the appropriate dilution of antibody for flow cytometry experiments. NPCs and Neurons were dissociated with accutase, spun down, and filtered through a 40 µm strainer into fixative solution (2% PFA in PBS) to obtain single cell suspensions. Cells were then counted using an automated Cell Countess system (Thermo Scientific). Live cell counts were also obtained using the Cell Countess system with Trypan Blue to distinguish between living and dead cells.

**Flow cytometry and imaging flow cytometry.** NPCs and Neurons were collected, fixed, and stained for flow and imaging cytometry as outlined in detail in our previous study<sup>45</sup>. Samples were stored at 4 °C in the dark until they were run through the Imagestream Mark II System (Amnis Inc/Luminex Corp.) or a Cytex modified FACSCalibur™ cytometer. Single stain controls were used for compensation and gating was determined through fluorescence minus one (FMOs) controls and isotype controls. For flow cytometry experiments gating and analysis were performed using FCS Express 6 (DeNovo Software). Imaging flow cytometry data were analyzed using IDEAS® software (version 6.2).  $\gamma$ H2AX foci were quantified using an efficient foci pipeline and masking strategy<sup>45,49</sup>. Cell cycle statistics were acquired using *Modfit* software. See supplemental information for more details on cytometry methods.

**Genome-wide break mapping and sequencing.** Detection of DNA breaks was performed as previously described<sup>58</sup>. Briefly, fixed nuclei of the p53<sup>KD</sup> and the control samples of NPC 9429 with and without APH treatment were subjected to blunting/A-tailing reactions, and Illumina P5 adaptor ligation to capture broken DNA ends. Genomic DNA was then purified and fragmented by sonication, and subsequently ligated to Illumina P7 adaptor, and the libraries were PCR-amplified for 15 cycles. Prepared libraries were then subjected to whole-genome, 75-bp and 150-bp paired-end sequencing with the Illumina NextSeq 500 and the HiSeq X Ten platforms, respectively. Two biological duplicates were performed for each sample treated for 1 day.

**DSB read processing.** Sequencing reads were aligned to the human genome (GRCh38/hg38) with bowtie2<sup>106</sup> (v.2.3.4.1) aligner running in high sensitivity mode. Restriction on the fragment length from 100 to 2000 nt (-X 2000 -I 100 options) was imposed. Unmapped, non-primary, supplementary and low-quality reads were filtered out with SAMtools<sup>107</sup> (v. 1.7) (-F 2820). Furthermore, PCR duplicates were marked with picard-tools (v. 1.95) MarkDuplicates, and finally, the first mate of non-duplicated pairs (-f 67 -F 1024) were filtered with SAMtools for continued analysis. For each detected break, the most 5' nucleotide of the first mate defined the DNA break position. Sequencing and alignment statistics for the DSB mapping/sequencing libraries are listed in Table S1. Downstream data analysis following DSB read processing has been performed with BEDtools<sup>108</sup> (v. 2.27.1) and standard Linux commands to compute coverages. Results were visualized in Python3 (v. 3.6.5) with matplotlib (v. 2.2.2), numpy<sup>109</sup> (v. 1.15.0), pandas<sup>110</sup> (v. 0.23.3), and R statistical software (v.3.4.4).

**DNA DSB gene analysis.** Sixty-one genes had >1.5-fold enrichment at TSSs +/- 250 bp in p53<sup>KD</sup> NPCs treated for 1 day. These genes were examined using Panther for gene function and GeneCard for clinical risk/association. The genes were further categorized based on clinical risk and disease type. These 61 genes were then coded as either neuronal or non-neuronal related and the gene expression in each class was compared to the bulk gene expression distribution from wildtype cells.

**Genomic regions annotation.** To assign genomic annotations BEDtools (v. 2.27.1) intersect was used to sequentially assign genomic features with each single-nucleotide break only being assigned to one genomic feature. The sequential feature assignment filters out breaks as they are assigned to a feature. The order for assigning genomic features was TSS, promoter, TTS, gene body, and those not assigned to any of these features are coded as intergenic. The GRCh38/hg38 build RefSeq genes were downloaded from the UCSC browser. The definitions used for each genomic feature is as follows: promoter region ranging from TSS - 1000 nt to - 250 nt, TSS region ranging from TSS - 250 to + 250 nt, gene body region ranging from TSS + 250 nt to TTS - 250 nt, and TTS regions ranging from TTS - 250 nt to + 250 nt. Counts in each genomic feature were then normalized to the megabase size of the genomic feature (breaks/Mb). The normalized density for breaks in each genomic feature were then plotted using Python (v. 3.8.10) and matplotlib.pyplot (v. 3.1.2).

**Statistical analyses.** For all experiments, data are shown as mean +/- standard error of the mean of three to six biological replicates per cell line. At least 2–3 neurotypic cells lines were used per experiment and pooled together for further analysis. Statistical significance was determined by using an unpaired t-test, one-way ANOVA and Dunnett post hoc test for differences, or two-way ANOVA when applicable with a Sidak Multiple Comparison Test. DNA DSBs were plotted as stacked histograms and analyzed using a contingency table and the Chi-squared test for two samples and nonparametric KS test for greater than three samples.

**Accession numbers.** DSB mapping and scRNA-seq datasets generated in this study can be accessed at the NCBI Sequence Read Archive (SRA) under the accession number: PRJNA591220. Flow cytometry data will be made available through Community Cytobank upon publication.

Received: 21 February 2022; Accepted: 11 July 2022  
Published online: 15 July 2022

## References

1. Lans, H. *et al.* The DNA damage response to transcription stress. *Nat. Rev. Mol. Cell Biol.* **20**(12), 766–784 (2019).
2. Liu, P. *et al.* Mechanisms for recurrent and complex human genomic rearrangements. *Curr. Opin. Genet. Dev.* **22**(3), 211–220 (2012).

3. Shrivastav, M., De Haro, L. P. & Nickoloff, J. A. Regulation of DNA double-strand break repair pathway choice. *Cell Res* **18**(1), 134–147 (2008).
4. Tubbs, A. & Nussenzweig, A. Endogenous DNA damage as a source of genomic instability in cancer. *Cell* **168**(4), 644–656 (2017).
5. Dasika, G. K. *et al.* DNA damage-induced cell cycle checkpoints and DNA strand break repair in development and tumorigenesis. *Oncogene* **18**(55), 7883–7899 (1999).
6. Roos, W. P. & Kaina, B. DNA damage-induced cell death by apoptosis. *Trends Mol. Med.* **12**(9), 440–450 (2006).
7. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**(5), 646–674 (2011).
8. Kinzler, K. W. & Vogelstein, B. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* **386**(6627), 761–763 (1997).
9. van Heemst, D., den Reijer, P. M. & Westendorp, R. G. Ageing or cancer: A review on the role of caretakers and gatekeepers. *Eur. J. Cancer* **43**(15), 2144–2152 (2007).
10. Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88**(3), 323–331 (1997).
11. Levine, A. J. & Oren, M. The first 30 years of p53: Growing ever more complex. *Nat. Rev. Cancer* **9**(10), 749–758 (2009).
12. Bhardwaj, R. D. *et al.* Neocortical neurogenesis in humans is restricted to development. *Proc. Natl. Acad. Sci. U.S.A.* **103**(33), 12564–12568 (2006).
13. Rakic, P. Neuroscience. No more cortical neurons for you. *Science* **313**(5789), 928–929 (2006).
14. von Bartheld, C. S., Bahnney, J. & Herculano-Houzel, S. The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *J. Comp. Neurol.* **524**(18), 3865–3895 (2016).
15. Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Development and evolution of the human neocortex. *Cell* **146**(1), 18–36 (2011).
16. Blaschke, A. J., Staley, K. & Chun, J. Widespread programmed cell death in proliferative and postmitotic regions of the fetal cerebral cortex. *Development* **122**(4), 1165–1174 (1996).
17. Kuan, C. Y. *et al.* Mechanisms of programmed cell death in the developing brain. *Trends Neurosci.* **23**(7), 291–297 (2000).
18. Rakic, S. & Zecevic, N. Programmed cell death in the developing human telencephalon. *Eur. J. Neurosci.* **12**(8), 2721–2734 (2000).
19. Thomaidou, D. *et al.* Apoptosis and its relation to the cell cycle in the developing cerebral cortex. *J. Neurosci.* **17**(3), 1075–1085 (1997).
20. Wong, F. K. & Marin, O. Developmental cell death in the cerebral cortex. *Annu. Rev. Cell Dev. Biol.* **35**, 523–542 (2019).
21. Chun, J. & Schatz, D. G. Rearranging views on neurogenesis: Neuronal death in the absence of DNA end-joining proteins. *Neuron* **22**(1), 7–10 (1999).
22. Ferguson, D. O. *et al.* The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. *Proc. Natl. Acad. Sci. U.S.A.* **97**(12), 6630–6633 (2000).
23. Ferguson, D. O. *et al.* The interplay between nonhomologous end-joining and cell cycle checkpoint factors in development, genomic stability, and tumorigenesis. *Cold Spring Harb. Symp. Quant. Biol.* **65**, 395–403 (2000).
24. Gu, Y. *et al.* Defective embryonic neurogenesis in Ku-deficient but not DNA-dependent protein kinase catalytic subunit-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **97**(6), 2668–2673 (2000).
25. Ratnaparkhe, M. *et al.* Defective DNA damage repair leads to frequent catastrophic genomic events in murine and human tumors. *Nat. Commun.* **9**(1), 4760 (2018).
26. Yan, C. T. *et al.* XRCC4 suppresses medulloblastomas with recurrent translocations in p53-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **103**(19), 7378–7383 (2006).
27. Shi, L. *et al.* Robust elimination of genome-damaged cells safeguards against brain somatic aneuploidy following Kln1 deletion. *Nat. Commun.* **10**(1), 2588 (2019).
28. Chow, H. M. & Herrup, K. Genomic integrity and the ageing brain. *Nat. Rev. Neurosci.* **16**(11), 672–684 (2015).
29. McKinnon, P. J. Genome integrity and disease prevention in the nervous system. *Genes Dev.* **31**(12), 1180–1194 (2017).
30. Chronister, W. D. *et al.* Neurons with complex karyotypes are rare in aged human neocortex. *Cell Rep.* **26**(4), 825–835 (2019).
31. Kaeser, G. & Chun, J. Brain cell somatic gene recombination and its phylogenetic foundations. *J. Biol. Chem.* **295**(36), 12786–12795 (2020).
32. McConnell, M. J. *et al.* Mosaic copy number variation in human neurons. *Science* **342**(6158), 632–637 (2013).
33. McConnell, M. J. *et al.* Intersection of diverse neuronal genomes and neuropsychiatric disease: The Brain Somatic Mosaicism Network. *Science* **356**(6336), eaal1641 (2017).
34. Rodin, R. E. *et al.* The landscape of somatic mutation in cerebral cortex of autistic and neurotypical individuals revealed by ultra-deep whole-genome sequencing. *Nat. Neurosci.* **24**(2), 176–185 (2021).
35. Rohrback, S. *et al.* Submegabase copy number variations arise during cerebral cortical neurogenesis as revealed by single-cell whole-genome sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **115**(42), 10804–10809 (2018).
36. Reid, D. A. *et al.* Incorporation of a nucleoside analog maps genome repair sites in postmitotic human neurons. *Science* **372**(6537), 91–94 (2021).
37. Wei, P. C. *et al.* Long neural genes harbor recurrent DNA Break clusters in neural stem/progenitor cells. *Cell* **164**(4), 644–655 (2016).
38. Schwer, B. *et al.* Transcription-associated processes cause DNA double-strand breaks and translocations in neural stem/progenitor cells. *Proc. Natl. Acad. Sci. U.S.A.* **113**(8), 2258–2263 (2016).
39. Wang, M. *et al.* Increased neural progenitor proliferation in a hiPSC model of autism induces replication stress-associated genome instability. *Cell Stem Cell* **26**(2), 221–233 (2020).
40. Weissman, I. L. & Gage, F. H. A mechanism for somatic brain mosaicism. *Cell* **164**(4), 593–595 (2016).
41. Glover, T. W. & Wilson, T. E. Molecular biology: Breaks in the brain. *Nature* **532**(7597), 46–47 (2016).
42. Wilson, T. E. *et al.* Large transcription units unify copy number variants and common fragile sites arising under replication stress. *Genome Res.* **25**(2), 189–200 (2015).
43. Zylka, M. J., Simon, J. M. & Philpot, B. D. Gene length matters in neurons. *Neuron* **86**(2), 353–355 (2015).
44. King, I. F. *et al.* Topoisomerases facilitate transcription of long genes linked to autism. *Nature* **501**(7465), 58–62 (2013).
45. Michel, N. *et al.* Imaging flow cytometry quantifies neural genome dynamics. *Cytometry A* **95**(8), 825–835 (2019).
46. Szlachta, K. *et al.* Topoisomerase II contributes to DNA secondary structure-mediated double-stranded breaks. *Nucleic Acids Res.* **48**(12), 6654–6671 (2020).
47. Le Tallec, B. *et al.* Updating the mechanisms of common fragile site instability: How to reconcile the different views?. *Cell Mol. Life Sci.* **71**(23), 4489–4494 (2014).
48. Lokanga, R. A., Kumari, D. & Usdin, K. Common threads: Aphidicolin-inducible and folate-sensitive fragile sites in the human genome. *Front. Genet.* **12**, 708860 (2021).
49. Bourton, E. C. *et al.* Multispectral imaging flow cytometry reveals distinct frequencies of gamma-H2AX foci induction in DNA double strand break repair defective human cell lines. *Cytometry A* **81**(2), 130–137 (2012).
50. Geric, M., Gajski, G. & Garaj-Vrhovac, V. gamma-H2AX as a biomarker for DNA double-strand breaks in ecotoxicology. *Ecotoxicol. Environ. Saf.* **105**, 13–21 (2014).
51. Caccamo, D. V. *et al.* An immunohistochemical study of neuropeptides and neuronal cytoskeletal proteins in the neuroepithelial component of a spontaneous murine ovarian teratoma. Primitive neuroepithelium displays immunoreactivity for neuropeptides and neuron-associated beta-tubulin isotype. *Am. J. Pathol.* **135**(5), 801–813 (1989).



52. Arva, N. C. *et al.* Disruption of the p53-Mdm2 complex by Nutlin-3 reveals different cancer cell phenotypes. *Ethn. Dis.* **18**(2 Suppl 2), S2-1-S2-8 (2008).
53. Roh, J. L. *et al.* p53-Reactivating small molecules induce apoptosis and enhance chemotherapeutic cytotoxicity in head and neck squamous cell carcinoma. *Oral Oncol.* **47**(1), 8–15 (2011).
54. Wang, B. *et al.* MDM2 inhibitor Nutlin-3a suppresses proliferation and promotes apoptosis in osteosarcoma cells. *Acta Biochim. Biophys. Sin. (Shanghai)* **44**(8), 685–691 (2012).
55. Lendahl, U., Zimmerman, L. B. & McKay, R. D. CNS stem cells express a new class of intermediate filament protein. *Cell* **60**(4), 585–595 (1990).
56. Park, D. *et al.* Nestin is required for the proper self-renewal of neural stem cells. *Stem Cells* **28**(12), 2162–2171 (2010).
57. Bott, C. J. *et al.* Nestin in immature embryonic neurons affects axon growth cone morphology and Semaphorin3a sensitivity. *Mol. Biol. Cell* **30**(10), 1214–1229 (2019).
58. Lensing, S. V. *et al.* DSBCapture: In situ capture and sequencing of DNA breaks. *Nat. Methods* **13**(10), 855–857 (2016).
59. Canela, A. *et al.* Topoisomerase II-induced chromosome breakage and translocation is determined by chromosome architecture and transcriptional activity. *Mol. Cell* **75**(2), 252–266 (2019).
60. Gothe, H. J. *et al.* Spatial chromosome folding and active transcription drive DNA fragility and formation of oncogenic MLL translocations. *Mol. Cell* **75**(2), 267–283 (2019).
61. Singh, S. *et al.* Pausing sites of RNA polymerase II on actively transcribed genes are enriched in DNA double-stranded breaks. *J. Biol. Chem.* **295**(12), 3990–4000 (2020).
62. Capranico, G. *et al.* The effects of camptothecin on RNA polymerase II transcription: Roles of DNA topoisomerase I. *Biochimica* **89**(4), 482–489 (2007).
63. Corti, A. *et al.* DNA damage and transcriptional regulation in iPSC-derived neurons from Ataxia Telangiectasia patients. *Sci. Rep.* **9**(1), 651 (2019).
64. Fragola, G. *et al.* Deletion of Topoisomerase 1 in excitatory neurons causes genomic instability and early onset neurodegeneration. *Nat. Commun.* **11**(1), 1962 (2020).
65. Marechal, A. & Zou, L. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb. Perspect. Biol.* **5**(9), a012716 (2013).
66. Matsuoka, S. *et al.* ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166 (2007).
67. Sancar, A. *et al.* Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* **73**, 39–85 (2004).
68. Lu, H. C. *et al.* STAT3 signaling in myeloid cells promotes pathogenic myelin-specific T cell differentiation and autoimmune demyelination. *Proc. Natl. Acad. Sci. U.S.A.* **117**(10), 5430–5441 (2020).
69. Rafiq, M. A. *et al.* Mutations in the alpha 1,2-mannosidase gene, MAN1B1, cause autosomal-recessive intellectual disability. *Am. J. Hum. Genet.* **89**(1), 176–182 (2011).
70. Bundo, M. *et al.* Decreased DNA methylation at promoters and gene-specific neuronal hypermethylation in the prefrontal cortex of patients with bipolar disorder. *Mol. Psychiatry* **26**(7), 3407–3418 (2021).
71. Coppede, F. & Migliore, L. DNA damage and repair in Alzheimer's disease. *Curr. Alzheimer Res.* **6**(1), 36–47 (2009).
72. Lovell, M. A. & Markesbery, W. R. Oxidative damage in mild cognitive impairment and early Alzheimer's disease. *J. Neurosci. Res.* **85**(14), 3036–3040 (2007).
73. Ogino, M. *et al.* Roles of PTEN with DNA repair in Parkinson's disease. *Int. J. Mol. Sci.* **17**(6), 954 (2016).
74. Sepe, S. *et al.* Inefficient DNA repair is an aging-related modifier of Parkinson's disease. *Cell Rep.* **15**(9), 1866–1875 (2016).
75. Coppede, F. An overview of DNA repair in amyotrophic lateral sclerosis. *ScientificWorldJournal* **11**, 1679–1691 (2011).
76. Martin, L. J. DNA damage and repair: Relevance to mechanisms of neurodegeneration. *J. Neuropathol. Exp. Neurol.* **67**(5), 377–387 (2008).
77. Liu, D. & Xu, Y. p53, oxidative stress, and aging. *Antioxid. Redox Signal* **15**(6), 1669–1678 (2011).
78. Alsina, F. C. & Silver, D. L. Damage control in the developing brain: Tradeoffs and consequences. *Trends Neurosci.* **42**(10), 661–663 (2019).
79. Lodato, M. A. *et al.* Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**(6375), 555–559 (2018).
80. Rothblum-Oviatt, C. *et al.* Ataxia telangiectasia: A review. *Orphanet. J. Rare Dis.* **11**(1), 159 (2016).
81. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. Identification of genetic factors that modify clinical onset of Huntington's disease. *Cell* **162**(3), 516–526 (2015).
82. Cao, L. *et al.* ATM-Chk2-p53 activation prevents tumorigenesis at an expense of organ homeostasis upon Brca1 deficiency. *EMBO J.* **25**(10), 2167–2177 (2006).
83. Chung, Y. M. *et al.* FOXO3 signalling links ATM to the p53 apoptotic pathway following DNA damage. *Nat. Commun.* **3**, 1000 (2012).
84. Squatrito, M. *et al.* Loss of ATM/Chk2/p53 pathway components accelerates tumor development and contributes to radiation resistance in gliomas. *Cancer Cell* **18**(6), 619–629 (2010).
85. Marin Navarro, A. *et al.* p53 controls genomic stability and temporal differentiation of human neural stem cells and affects neural organization in human brain organoids. *Cell Death Dis.* **11**(1), 52 (2020).
86. Sivakumar, S. *et al.* TP53 promotes lineage commitment of human embryonic stem cells through ciliogenesis and sonic hedgehog signaling. *Cell Rep.* **38**(7), 110395 (2022).
87. Bardy, C. *et al.* Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. *Mol. Psychiatry* **21**(11), 1573–1588 (2016).
88. Fischer, M. Census and evaluation of p53 target genes. *Oncogene* **36**(28), 3943–3956 (2017).
89. Wang, Y. H. *et al.* Rapid recruitment of p53 to DNA damage sites directs DNA repair choice and integrity. *Proc. Natl. Acad. Sci. U.S.A.* **119**(10), e2113233119 (2022).
90. Ardhanareswaran, K. *et al.* Human induced pluripotent stem cells for modelling neurodevelopmental disorders. *Nat. Rev. Neurol.* **13**(5), 265–278 (2017).
91. Brennand, K. J. *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* **473**(7346), 221–225 (2011).
92. Choi, S. H. *et al.* A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature* **515**(7526), 274–278 (2014).
93. Flaherty, E. *et al.* Patient-derived hiPSC neurons with heterozygous CNTNAP2 deletions display altered neuronal gene expression and network activity. *NPJ Schizophr.* **3**, 35 (2017).
94. Liu, Y. *et al.* Dravet syndrome patient-derived neurons suggest a novel epilepsy mechanism. *Ann. Neurol.* **74**(1), 128–139 (2013).
95. Marchetto, M. C. *et al.* Altered proliferation and networks in neural cells derived from idiopathic autistic individuals. *Mol. Psychiatry* **22**(6), 820–835 (2017).
96. Hazen, J. L. *et al.* The complete genome sequences, unique mutational spectra, and developmental potency of adult neurons revealed by cloning. *Neuron* **89**(6), 1223–1236 (2016).
97. Shi, Y., Kirwan, P. & Livesey, F. J. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat. Protoc.* **7**(10), 1836–1846 (2012).

98. Pan, M. *et al.* The chemotherapeutic CX-5461 primarily targets TOP2B and exhibits selective activity in high-risk neuroblastoma. *Nat. Commun.* **12**(1), 6468 (2021).
99. Jin, Y. *et al.* Inactivation of EGLN3 hydroxylase facilitates Erk3 degradation via autophagy and impedes lung cancer growth. *Oncogene* **41**(12), 1752–1766 (2022).
100. Zheng, Y. *et al.* X-ray repair cross-complementing protein 1 (XRCC1) loss promotes beta-lapachone-induced apoptosis in pancreatic cancer cells. *BMC Cancer* **21**(1), 1234 (2021).
101. Liu, X. *et al.* Human parainfluenza virus type 3 expressing the respiratory syncytial virus pre-fusion F protein modified for virion packaging yields protective intranasal vaccine candidates. *PLoS ONE* **15**(2), e0228572 (2020).
102. Andreeva, A. *et al.* PTK7-Src signaling at epithelial cell contacts mediates spatial organization of actomyosin and planar cell polarity. *Dev. Cell* **29**(1), 20–33 (2014).
103. Mu, J. *et al.* BRD9 inhibition promotes PUMA-dependent apoptosis and augments the effect of imatinib in gastrointestinal stromal tumors. *Cell Death Dis.* **12**(11), 962 (2021).
104. Ruiz, S. *et al.* Limiting replication stress during somatic cell reprogramming reduces genomic instability in induced pluripotent stem cells. *Nat. Commun.* **6**, 8036 (2015).
105. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**(3), 562–578 (2012).
106. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**(4), 357–359 (2012).
107. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**(16), 2078–2079 (2009).
108. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**(6), 841–842 (2010).
109. van der Walt, S., Colbert, S. C. & Varoquaux, G. The NumPy array: A structure for efficient numerical computation. *Comput. Sci. Eng.* **13**(2), 22–30 (2011).
110. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**(3), 90–95 (2007).

## Acknowledgements

We thank the Stem Cell Core, Genome Analysis and Technology Core, and the Flow Cytometry Core at the University of Virginia School of Medicine. This work was supported by a Wagner Fellowship (N.M., H.M.R.), F31AG062042 (N.M.), 5T32GM008136 (N.M., H.M.R.), 5T32GM007267 (N.M.), a Maddy Deal Pilot Grant (M.J.M.), the Owens (M.J.M.) and Familant (M.J.M.) Family Funds, McDonnell Foundation (I.E.B.), FONDECYT Regular #1191737 from the Agencia Nacional de Investigación y Desarrollo de Chile (I.E.B.) and R01GM101192 (Y.H.W.).

## Author contributions

N.M., Y.H.W., and M.J.M. designed the experiments and wrote the manuscript. N.M., U.A., L.G., I.E.B. and R.A.H. performed biochemical and pharmacological experiments on NPCs and neurons. Y.H.W., H.M.R., N.D.A., S.S., and A.M. processed and analyzed NPCs for DSB Mapping.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-16516-5>.

**Correspondence** and requests for materials should be addressed to M.J.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022