Chd1 protects genome integrity at promoters to sustain hypertranscription in embryonic stem cells

Aydan Bulut-Karslioglu1,2✉, Hu Jin3,4,7, Yun-Kyo Kim5, Brandon Cho5, Marcela Guzman-Ayala1,8, Andrew J. K. Williamson6,9, Miroslav Hejna3,4, Maximilian Stötzel2, Anthony D. Whetton6, Jun S. Song3,4 & Miguel Ramalho-Santos1,5✉

Stem and progenitor cells undergo a global elevation of nascent transcription, or hypertranscription, during key developmental transitions involving rapid cell proliferation. The chromatin remodeler Chd1 mediates hypertranscription in pluripotent cells but its mechanism of action remains poorly understood. Here we report a novel role for Chd1 in protecting genome integrity at promoter regions by preventing DNA double-stranded break (DSB) accumulation in ES cells. Chd1 interacts with several DNA repair factors including Atm, Parp1, Kap1 and Topoisomerase 2β and its absence leads to an accumulation of DSBs at Chd1-bound Pol II-transcribed genes and rDNA. Genes prone to DNA breaks in Chd1 KO ES cells are longer genes with GC-rich promoters, a more labile nucleosomal structure and roles in chromatin regulation, transcription and signaling. These results reveal a vulnerability of hypertranscribing stem cells to accumulation of endogenous DNA breaks, with important implications for developmental and cancer biology.
Proliferating stem and progenitor cells are net generators of new cellular biomass and therefore have high biosynthetic demand. One way that stem/progenitor cells cope with this demand is to enter a state of hypertranscription, which involves a global elevation of nascent transcriptional output. Hypertranscription is masked by most transcriptional profiling approaches, but has attracted renewed interest recently. Hypertranscription has been documented to occur and play critical roles in embryonic stem (ES) cells, the post-implantation epiblast, the post-implantation epiblast–implantation epiblast (E5.5–E6.5) and more abundantly at E6.5, and is mainly localized to the nucleolus. The surprising presence of yH2A.X in undamaged ES cells and its dependence on Chd1 led us to investigate the status of this histone mark in vivo. We found that yH2A.X is detected in vivo in control embryos at E5.5 and more abundantly at E6.5, and is mainly localized to the nucleolus along with a diffuse nuclear pattern (Fig. 2c, d and Supplementary Fig. 2a). Strikingly, Chd1 KO embryos entirely lack this yH2A.X signal at both developmental stages (Fig. 2e), while global H2A.X is retained (Supplementary Fig. 2b). These findings in vivo are in agreement with the ES cell data above, showing they are not an artifact of cell culture. The early post-implantation epiblast is one of the fastest proliferating cell types in mammals, with doubling times between 2 and 8 h. Interestingly, the fastest rates of proliferation were recorded at E6.5, where we find high levels of nucleolar yH2A.X (Fig. 2 and Supplementary Fig. 2). Overall, the results indicate that nucleolar accumulation of yH2A.X in vivo correlates with RNA synthesis and proliferation rate, and all three of these are dependent on Chd1 (this study and Guzman-Ayala et al.). The loss of yH2A.X in Chd1 KO cells could simply be a consequence of the reduced global transcriptional output and therefore a lower occurrence of transcription-induced DNA breaks. We therefore set out to determine the levels and genomic location of DSBs in control vs Chd1 KO ES cells. We performed

Results

In order to probe the function of Chd1 in the regulation of hypertranscription, we identified its interacting proteins by immunoprecipitation followed by mass spectrometry (IP-MS) using a Chd1-Flag knock-in mouse ES cell line. IPs were performed under two different salt concentrations (150 vs 250 mM) to gauge the interaction strengths (Fig. 1a and Supplementary Fig. 1a–b). In physiological salt concentration (150 mM) we detected 314 proteins, the majority of which were lost in higher salt (Fig. 1a, Supplementary Fig. 1b, and Supplementary Data 1). Previously described Chd1-interacting proteins such as Ssrp1 and Bptf were recovered at high confidence in physiological salt, therefore we focused further analyses on this dataset (Fig. 1a). As expected, putative Chd1-interacting proteins are enriched for factors involved in chromatin and transcriptional regulation (Fig. 1b and Supplementary Data 2). In addition, there is an unexpected enrichment for DNA repair factors among Chd1 interactors, such as the protein kinase Atm, histone variant H2A.X, MRN complex member Mre11, topoisomerase 2β (Top2B), Xrcc factors (Fig. 1a, c). Another repair-associated protein, Kap1, was retained as a Chd1 interactor at high salt immunoprecipitation (Supplementary Fig. 1b). We confirmed interactions of Chd1 with activated Atm (Atm phospho-S1981), Kap1, Parp1 and Top2β in wild-type ES cells via co-immunoprecipitation (Co-IP) (Fig. 1d and Supplementary Fig. 1c). Interaction of Chd1 with the single-stranded repair factor Xrcc1 was not detected by Co-IP (Fig. 1d). Therefore, we focused on the double-stranded DNA (DSB) repair components for the remainder of this study.

We next assessed how genetic deletion of Chd1 affects the levels and localization of its DSB repair interactors. Loss of Chd1 leads to reduced global levels of S139 phosphorylated H2A.X (yH2A.X), a mark of DSB repair, despite slight increases in the levels of H2A.X and Atm kinase, which phosphorylates H2A.X (Supplementary Fig. 1d). Immunofluorescence confirmed that Atm and Top2β levels increase upon Chd1 loss, with the unexpected observation that they accumulate in the nucleolus in ES cells (Fig. 2a). To our knowledge, this pattern of accumulation of DNA repair factors at the nucleolus of undamaged cells has not previously been described. We have previously shown that Chd1 binds directly to rDNA and that loss of Chd1 leads to reduced nascent synthesis of rRNA and fragmentation of nucleoli (Guzman-Ayala et al. and Fig. 2a). Chromatin immunoprecipitation-qPCR (ChIP-qPCR) revealed an accumulation of Atm, pKap1 and yH2A.X at rDNA in control cells, and a decrease in pKap1 and yH2A.X despite higher levels of Atm. H2A.X and Kap1 at rDNA in Chd1 KO ES cells (Fig. 2b and Supplementary Fig. 1c). Taken together, these results suggest that Chd1 modulates the function of interacting factors involved in DSB repair in undamaged ES cells, particularly at rDNA in nucleoli. We previously used mouse genetics to show that Chd1 is essential for rapid growth of the early post-implantation mouse epiblast (E5.5–6.5) by promoting a high transcriptional output, notably of nascent rRNA at the nucleolus. The surprising presence of yH2A.X in undamaged ES cells and its dependence on Chd1 led us to investigate the status of this histone mark in vivo. We found that yH2A.X is detected in vivo in control embryos at E5.5 and more abundantly at E6.5, and is mainly localized to the nucleolus along with a diffuse nuclear pattern (Fig. 2c, d and Supplementary Fig. 2a). Interestingly, the fastest rates of proliferation were recorded at E6.5, where we find high levels of nucleolar yH2A.X (Fig. 2 and Supplementary Fig. 2). Overall, the results indicate that nucleolar accumulation of yH2A.X in vivo correlates with rRNA synthesis and proliferation rate, and all three of these are dependent on Chd1 (this study and Guzman-Ayala et al.). The loss of yH2A.X in Chd1 KO cells could simply be a consequence of the reduced global transcriptional output and therefore a lower occurrence of transcription-induced DNA breaks. We therefore set out to determine the levels and genomic location of DSBs in control vs Chd1 KO ES cells. We performed
We previously reported that Chd1 deletion leads to global hypotranscription of both RNA Pol I and Pol II-transcribed genes. The unexpected increase in DSBs at rDNA (Pol I-transcribed) in Chd1 KO ES cells (Fig. 3c) led us to explore the status of DSBs at Pol II-transcribed genes, using DSB-seq (see Methods section for procedure). For comparative analyses, we also performed Chd1 and RNA Pol II ChIP-seq in Chd1-Flag knock-in ES cells.

DSB-seq confirmed the gain of breaks at rDNA (Supplementary Fig. 3a) and revealed a remarkable widespread accumulation of DSBs at promoter regions of RNA Pol II-transcribed genes in Chd1 KO cells, relative to controls (Fig. 4a–c). DSBs occur immediately downstream of the transcription start site (TSS), where Chd1 binding peaks in wild-type (WT) conditions (Fig. 4a–c). At promoter regions (±1kb of TSS), DSBs in Chd1 KO cells positively correlate with GC content (Spearman $\rho = 0.80$, $p < 10^{-300}$), Chd1 (Spearman $\rho = 0.71$, $p < 10^{-300}$), and RNA Pol II (Spearman $\rho = 0.63$, $p < 10^{-300}$) binding in WT cells, and negatively correlate with nucleosome occupancy (Spearman $\rho = -0.29$, $p < 5 \times 10^{-280}$). The propensity to accumulate DSBs in Chd1 KO ES cells does not correlate with wild-type gene expression levels (Fig. 4b–d, Spearman $\rho = 0.012$, $p > 0.306$) or reduced expression upon Chd1 loss (Supplementary Fig. 3a).
We detected 5671 DSB peaks in Chd1 KO ES cells (vs. only 54 in control), among which 1825 peaks mapped around the TSSs (−1kb to +100 bp) of 1785 genes (DSB-prone genes, Supplementary Data 3). DSB peaks are enriched at TSSs, 5′ untranslated regions and exons (Supplementary Fig. 3b). GO analysis predicts that genes involved in transcription, chromatin modification, and signaling are particularly prone to DSB in Chd1 KO ES cells (Supplementary Fig. 3c). To understand why these genes might be especially susceptible to DNA breaks, we analyzed the chromatin structure at DSB-prone TSSs in comparison to non-DSB-prone TSSs. We utilized ES cell MNase-seq datasets generated by Voong et al.19 (Fig. 4d). DSB-prone genes display a more open chromatin structure at TSSs, with less regular nucleosomes surrounding the start site. In particular, the +1 nucleosome is less...
abundant in DSB-prone genes (Fig. 4d). Taken into consideration that this graph depicts a population average of many cells going through the transcription cycle, the data suggest that the +1 nucleosome in DSB-prone genes is frequently displaced to expose naked DNA. Interestingly, Chd1 peaks at this +1 position (Fig. 4a), suggesting that it promotes eviction of this nucleosome, as previously reported in embryonic fibroblasts[^10], and in doing so facilitates DNA repair and transcriptional elongation.

To probe why these 1785 genes have a more exposed TSS region and are DSB-prone upon Chd1 loss, we investigated genomic features of these genes. DSB-prone genes have a significantly higher GC content in their promoter-proximal region compared to non-DSB-prone genes (Fig. 4a, e, Wilcoxon rank-sum test $p < 10^{-300}$). Moreover, DSB-prone genes are on average longer than non-DSB-prone genes (Fig. 4f, Wilcoxon rank-sum test $p < 1.62 \times 10^{-99}$). Previous studies showed that longer genes
are more DSB-prone in neurons, due to the increased Topoisomerase activity required to relieve the torsional stress that accumulates during DNA unwinding (reviewed in Teves et al.16). Taken together, our results suggest that Chd1 remodels nucleosomes at GC-rich promoters of long genes to facilitate DNA repair and repeated cycles of RNA Pol II elongation in hypertranscribing ES cells.

Finally, we explored the relationship between DSBs induced by Chd1 loss in ES cells and the activities of Topoisomerase, transcription initiation, and DNA repair. Topoisomerase creates then ligates DSBs, unless the cells are treated with etoposide which blocks its ligation activity. Treatment of control cells with etoposide increases DSB formation, indicating that the DSBs do occur in wt cells but are promptly repaired by Topoisomerases in...
the presence of Chd1 (Fig. 5a and Supplementary Fig. 4a). Etoposide treatment further increases DSB levels in Chd1 KO cells, suggesting that the DSBs observed in Chd1 KO cells are not simply due to defective Topoisomerase activity, a finding that requires further investigation. Overall, inhibition of topoisomerase activity can induce DSBs at active promoters, as has been shown previously17, and this is the case even in the absence of Chd1.
Although DSB occurrence in Chd1 KO cells does not correlate with gene expression levels in wt cells or with differential expression in Chd1 KO cells, DSBs do correlate to some extent with RNA Pol II occupancy at promoters (Fig. 4a). We note that this correlation is not linear, as the 1st and 2nd quartiles of highest DSB-accumulating genes have a greater difference in DSB levels than Pol II occupancy in control cells. Considering the fact that Chd1 associates with Pol II and facilitates transcription just downstream of Pol II through pumping DNA towards it20, Chd1 KO could compromise transcription via accumulation of initiated Pol II at TSSs. To probe the role of transcription initiation on DSB occurrence, we tested DSB levels upon inhibition of transcription initiation via triptolide treatment (Fig. 5b and Supplementary Fig. 4b). Inhibition of transcription initiation does increase DSBs in wt cells to levels comparable to Chd1 KO cells but does not further increase DSBs in Chd1 KO cells. These data suggest that DSBs that form in Chd1 KO ES cells may occur during Pol II loading and assembly, while reduced nascent transcription may further contribute to DSBs accumulation.

Discussion

This work describes a novel role for the chromatin remodeler Chd1 in protecting genome integrity at promoter regions by preventing DSB accumulation in pluripotent stem cells (Fig. 5c). Our results point to a central role for rDNA at the nucleolus in the coordination between hypertranscription and DNA integrity in hypertranscribing ES cells. rRNA comprises ~80% of the RNA being synthesized in ES cells and therefore represents both a major focal point of hypertranscription as well as a vulnerability to DNA breaks. We had previously implicated Kap1, generally thought to be a repressor, in rRNA transcription21, and this study further contributes to clarifying that role.

We propose that the accumulation of unrepaired DNA breaks in Chd1 KO cells compromises nascent transcriptional output and leads to the proliferation defects and ultimate developmental arrest of the rapidly expanding post-implantation epiblast. We note that the loss of Chd1 is not catastrophic to ES cells, despite the widespread occurrence of DSBs at GC-rich promoters of transcribed genes: Chd1 KO ES cells remain undifferentiated and capable of gene transcription, albeit with a lower transcriptional output and a self-renewal deficit. These results suggest that promoter DSBs are still eventually repaired in Chd1 KO ES cells, albeit at a lower rate, which compromises optimal nascent transcription and proliferation. Although we did not observe a direct correlation between DSB formation and expression levels in wt or expression changes in Chd1 KO cells, a role for reduced transcriptional output on DSB formation via RNA Pol II accumulation or pausing is a possibility that warrants further studies. Similarly, understanding to what extent Chd1 specifically suppresses DSB formation in the first place versus promotes their rapid repair will require more detailed biochemical studies. These functions may not be mutually exclusive. For example, it is possible that Chd1 facilitates the relocalization activity of Top2a, and in parallel recruits DNA repair factors as a backup mechanism that may be particularly important in hypertranscribing cells. In either scenario, Chd1 activity at transcribed promoters protects genome integrity, a finding with significant implications in stem cell biology and cancer.

Repair of DSBs requires efficient signaling spearheaded by the Atm kinase, which we show is defective in Chd1 KO ES cells. Interestingly, Chd1-deficient human cancer cells have defective DNA repair when exposed to ionizing radiation or radiomimetic chemicals22,23. These cells show reduced H2AX phosphorylation and are more sensitive to PARP inhibitors. The present study is the first report of Chd1-mediated DNA repair activity under native conditions without any external insult. The TSS sequence and chromatin landscape appear to be the main determinants of DSB generation in ES cells. DSBs were shown to be conducive to transcription in several cell types, in the absence of exogenous DNA damage11–13,17. Such studies point to a correlation between gene transcription and the occurrence of DNA breaks. Chd1 KO ES cells, on the other hand, have lower nascent transcription levels and higher incidence of DNA breaks. Thus, the DNA breaks we observe in Chd1 KO ES cells are not caused by increased transcription but rather by defective repair, a notion further supported by the loss of γH2AX. We speculate that hypertranscribing, proliferating cells may have an increased dependence on Chd1 to balance high transcriptional output with DNA integrity. It will be of interest to explore the potential role of Chd1 in transcription-associated DNA break repair in other cell types, both in physiological stem/progenitor cells as well as in proliferating tumor cells. Of note, Chd1 is the second most frequently mutated gene in prostate cancer, after PTEN2425. In agreement with our data, Dellino et al.26 showed that release of paused RNA Pol II in breast cancer cells induces DSBs preferentially at long genes and can lead to chromosomal translocations.

In highly proliferative cells, replication stress can occur due to the collision of replication and transcription complexes, resulting in DNA breaks that are marked by γH2AX27. It is unlikely that the high γH2AX signal in wild-type ES cells and epiblast is strictly due to replicative stress because high γH2AX is observed in all cells regardless of cell cycle stage. Moreover, Chd1 KO ES cells have higher incidence of DSBs and reduced γH2AX (this study) without significant deviations in cell cycle stage proportions23. In agreement, γH2AX staining is not limited to S phase cells, as might be expected from replication stress, but is found across all stages of the cell cycle, in both control and Chd1 KO ES cells (Supplementary Fig. 4c–e). Taken together, our findings suggest that hypertranscription puts ES cells, and potentially other stem/progenitor cells, at risk of DNA breaks and genomic instability, but this is countered by Chd1-dependent DNA repair.

Why Atm signaling is defective in Chd1 KO ES cells is a central question that remains elusive. In Chd1 KO cancer cells, γH2AX is reduced due to reduced incorporation or retention of H2A.X at damage sites23. Decreased γH2AX despite DSB accumulation could stem from defective γH2AX spreading due to compromised H2A.X deposition in Chd1 KO ES cells. H2A.X is deposited by the FACT complex28, which interacts with Chd1 to spread through transcribed regions29. Therefore it is conceivable that the amplification of the DNA damage signal through γH2AX spreading is defective in Chd1 KO cells. Alternatively, it is possible that the chromatin remodeling activity of Chd1 facilitates access of ATM to its target site on H2A.X, or that other interactors of Chd1 promote Atm activity towards H2AX. For example, Chd1 may also interact with PARP1 and the histone acetyltransferase Tip60, and both histone ADP ribosylation and acetylation are involved in chromatin relaxation at DNA repair sites30,31. Other major transcriptional regulators such as Myc and Paflc mediate repair of transcription-associated DNA breaks through H2B ubiquitination32, which may be a mechanism shared by Chd1, as global H2Bub reduction was observed in Chd1 KO cells33. Further biochemical studies of how the activity of Atm and other aspects of DNA repair are modulated by Chd1 will shed light on the mechanisms by which stem and progenitor cells can undergo hypertranscription while preserving DNA integrity.

Methods

Mice. Chd1−/− females (6–12-week-old) and males (6 week- to 6-month-old) were used to recover the embryos. Animals were maintained on 12 h light/dark
cycle and provided with food and water ad libitum in individually ventilated units (TecniIntat at TCT, Lab Products at UCSF) in the specific pathogen-free facilities at UCSF. All procedures involving animals were performed in compliance with the protocol approved by the IACUC at UCSF, as part of an AAALAC-accredited care and use program (protocol AN091331-03). Chd1 heterozygous mice were mated, embryos were collected at embryonic day 5 (E5.5) or day 6 (E6.5) after detection of the copulatory plug by dissecting uteri of pregnant females following standard practices.

ES cell culture. Chd1-Flag knock-in, Chd1<sup>ΔΔ</sup> and Chd1<sup>Δ/Δ</sup> ES cells were used as described<sup>3</sup>. Cells were grown in DMEM GlutaMAX with 15% FBS (Atlanta Biologicals), 0.1 mM non-essential amino acids, 50 U/ml Penicillin/Streptomycin (UCSF Cell Culture Facility), 0.1 mM EmbryoMax 2-Mercaptoethanol (Millipore), and 2000 U/ml ESGRO supplement (LIF, Millipore or Genmesi) under ambient air with 5% CO<sub>2</sub>. Cells tested negative for mycoplasma contamination.

Immunoprecipitation. Wild-type E14 or Chd1-Flag knock-in were used. Cells were fractionated into cytoplasmic and nuclear compartments prior to immunoprecipitation. In brief, cells were collected in Buffer A (10 mM Tris pH 7.9, 5 mM MgCl<sub>2</sub>, 0.25 M Sucrose and 0.1% NP-40 supplemented with 1X Halt protease inhibitor cocktail (Thermo Fisher Scientific, 78425), 1 mM PMSF, 5 mM Na<sub>3</sub>PO<sub>4</sub> and 1 mM Na<sub>2</sub>VO<sub>4</sub>) and were centrifuged for 10 min, 1800×g at 4 °C. For nuclear extracts, the resulting pellets were resuspended in buffer B (10 mM HEPES pH 7.9, 150 mM Na<sub>2</sub>EDTA, 25% glycerol, 0.1% Triton X-100, 10 mM Na<sub>3</sub>PO<sub>4</sub> supplemented with 1X Halt protease inhibitor cocktail (Thermo Fisher Scientific, 78425), 1 mM PMSF, 5 mM Na<sub>3</sub>PO<sub>4</sub> and 1 mM Na<sub>2</sub>VO<sub>4</sub>) and homogenized by passing through 18Gx1/2” size needles. Nuclear extracts were quantified using Pierce BCA Protein Assay kit (23225). For immunoprecipitation, 100 µg of extract was adjusted to 500 µl total volume and 150 mM final NaCl concentration, and incubated in the presence of 20 µl pre-washed Protein A or G Dynabeads (Thermo Fisher Scientific, 1002D or 1004D) and the following antibodies: Flag (Sigma, F1804), Kap1 (Abcam, ab22533), Atm (Abcam, ab22758), Polr1a (1:1000, Cell Signaling, D6S6S), H2A.X (1:2500, Santa Cruz Biotechnology, sc-8389), Atm (Genetex, GTX70103), Top2α (1:100, Cell Signaling, sc-25780), Top2β (1:100, Cell Signaling, sc-10105), and 2x Laemmli Buffer with 5% β-mercaptoethanol. Western blot was performed as described below.

IP- mass spectrometry. Chd1-Flag ES cells were lysed and processed as above. Nuclear extracts were used for immunoprecipitation at 150 mM and 250 mM salt concentrations. For mass spectrometry, Flag IP and IgG control were run on a denaturing gel. Gels were stained using Coomassie Brilliant Blue IP (10 mM Tris-HCl pH 6.8, 400 with an 8% acrylamide gel) and destained in 200 mM Na<sub>3</sub>PO<sub>4</sub> and then trypsinized (Trypsin resuspended in 100 mM ammonium bicarbonate, 5% [v/v] acetonitrile) overnight at 37 °C. Peptides were extracted from the gel pieces by incubation in 50% [v/v] acetonitrile, 0.1% [v/v] formic acid to 40% [v/v] acetonitrile, and 0.1% [v/v] formic acid in acetonitrile on-line to a LTQ Orbitrap Velos (Thermo). Data was acquired using an data dependent acquisition (DDA) method where, for each cycle one full MS scan of acid on-line to a LTQ Orbitrap Velos (Thermo). Data was acquired using an data dependent acquisition (DDA) method where, for each cycle one full MS scan of

Western blot analysis. Chd1<sup>ΔΔ</sup> and Chd1<sup>Δ/Δ</sup> (7 days post-induction of KO) ES cells were used. For analysis of whole cell extracts, cells were lysed in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris pH 8.0, 1X Halt protease inhibitor cocktail (Thermo Fisher Scientific), 1 mM PMSF, 5 mM Na<sub>3</sub>PO<sub>4</sub> and 1 mM Na<sub>2</sub>VO<sub>4</sub>). Cells were incubated for 30 min on ice, then sonicated using a Bioruptor (Diagenode) for 5 min with settings high, 30 s on, 30 s off. Lamin B1 buffer with 5% FE buffer (10 mM Tris pH 7.5, 0.1 M NaCl) and were boiled at 95 °C for 5 min. Extracts were loaded into 4–15% Mini-Protein TGX SDS Page gels (Bio-Rad). Proteins were transferred to PVDF membranes. Membranes were blocked in 5% milk/PBS-T buffer for 30 min and incubated either overnight at 4 °C or for 1 h at room temperature with the following antibodies: Chd1 (1:1000, Cell Signaling, #3351), Top2α (1:500, Santa Cruz Biotechnology, sc-8389), Atm (1:250, Genetex, GTX70103 and Abcam, ab78), p-Atm (1:250, Thermo, MA1-2020), Kap1 (Abcam, ab25443), Nucleolin (1:1000, Abcam, ab228758), Polr1a (1:1000, Cell Signaling, D6S6S), H2A.X (1:2500, Abcam, ab11775), yH2A.X (1:1000, Abcam, ab2893), Parp1 (1:500, Santa Cruz Biotechnology, sc-25780), γH2A.X (1:1000, Millipore, Ab5100), and anti-rabbit antibodies (1:2000, Jackson Labs, 115-035-062, 111-035-144). Membranes were incubated with ECL or ECL Plus reagents and exposed to X-ray films (Thermo Fisher Scientific).

Immunofluorescent staining and imaging. Chd1<sup>ΔΔ</sup> and Chd1<sup>Δ/Δ</sup> ES cells were used. Cells were plated on matrigel in 8-chamber polystyrene vessels. Cells were fixed in 4% paraformaldehyde for 10 min, washed with PBS, and permeabilized with 0.2% Triton X-100 in PBS for 5 min. Cells were blocked in 5% donkey serum for 1 h, cells were incubated overnight at 4 °C with the following antibodies: yH2A.X (1:500, Abcam, ab2893), Nucleolin (1:1000, Abcam, ab22758), Atm (Genetex, GTX70103), Top2α (1:100, Abcam, GTX70130), β3RP1 (1:300, Abcam, ab175933), and Fibrillarin (1:200, Abcam, ab4566). Cells were washed in PBS--TBS membrane. Cells were incubated in a Leica-2B microscope. Staining and imaging of E5.5 and E6.5 embryos were performed as described in Guzman--Ayala et al. Briefly, embryos were fixed in 4% paraformaldehyde (in PBS) and sonicated using a Covaris S2 sonicator with settings 5% duty cycle, intensity 4, cycles per burst 200, frequency sweeping. In all, 20 µl chromatin was incubated sequentially with 1 µl RNaseA and 5 µl proteinase K in 100 µl total volume at 37 °C for 30 min and 30 °C for 1 h, purified modification was oxidized methionine, phosphorylated serine, threonine, and tyrosine and the peptide tolerance of 0.025 Da and 0.03 Da for MS/MS tolerance. The false discovery rate (FDR) of protein identification was controlled using a target-decoy searching strategy<sup>34</sup>. The maximum allowed FDR for protein identification was set to 1%. The posterior error probability (PEP) score denotes the probability that the identified peptide is correct.

For gene ontology (GO) and protein network analysis, proteins with fewer than 4 peptides in Chd1 IP were filtered out and only those hits passing the log2FC (Chd1 IP/negative control IP) > 2 and PEP score < 0.05 were retained. Gene ontology (GO) pathway enrichment analysis was performed using the commonCluster function of clusterProfiler<sup>55</sup> (version 3.18.1.1). The “enrichGO” function for biological process with a significance cutoff p-value < 0.05 and q-value < 0.1 was used. Benjamini–Hochberg method was used to correct for multiple statistical testing (adjusted p-values).

For gene network analysis, STRING was performed using the STRING software<sup>36</sup> (version 11.0) with the following parameters: Network type: Full STRING network, meaning of network edges: confidence (line thickness indicates the strength of data support), active interaction sources: experiments, textmining, databases, co-expression, neighborhood, gene fusion, co-occurrence, minimum required interaction score = 0.4.
30 min and 65 °C for 2 h to overnight. DNA was purified using a QiaGen PCR purification kit. qPCR was performed with KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems) and amplified on a 7900HT Real-time PCR machine (Applied Biosystems). For sequencing, libraries were prepared using the NEBNext ChIP-seq Library Prep for Illumina kit. Library quality and quantity were analyzed using Bioanalyzer (Agilent). Samples were sequenced on a HiSeq 2500 using single-end 50 bp sequencing reads in rapid mode.

Detection of double-stranded DNA breaks. Chd1Δ fla and Chd1Δ flaΔ established ES cells were used. For detection of double-stranded DNA breaks, we combined the Baranello et al.17 and BLESS protocols38 to achieve end labeling in situ. 1.5 × 10⁸ cells were cultured in 2% formaldehyde suspension for 30 min at room temperature, followed by quenching in 0.25 M glycine for 5 min. Cells were centrifuged, washed twice in PBS, and incubated in 25 μl buffer: (10 mM Tris pH 8.0, 10 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.2% NP-40) for 1 h at 4 °C. Pellets were resuspended in nucleus break buffer (10 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.3% SDS, 1 mM DTT) and incubated for 45 min at 37 °C. After centrifugation, nuclei were resuspended in 1× NEB buffer 2 on ice. In all, 10 μl Proteinase K (20 mg/ml) was added, digestion was performed for 4 min at 37 °C, and samples were immediately returned to ice and 25 μl PMSF was added. Nuclei were centrifuged, and resuspended in 5 ml 1× NEB Buffer 2 + 15 μl Triton X-100, and centrifuged again at 200×g for 10 min at 4 °C. Nuclei were washed once in water, and divided into two tubes for end labeling and control reactions. Nuclei were resuspended in 625 μl 1× TdT buffer with 2.5 μl TdT (Promega, M1871) and 2 μl biotin-16-dUTP (Roche 11093070910), and were incubated at 37 °C for 1 h.

Control reaction was assembled the same way without TdT. 1:50 volume of 0.5 M K (20 mg/ml) at 2 h to overnight at 65 °C. Labeled DNA was precipitated using formamide, 10 mM EDTA. After incubation at 65 °C for 5 min, DNA was purified on a 1% agarose gel. For biotinylated DNA pull-down, 20 μl Proteinase K (20 mg/ml) at 2 h to overnight at 65 °C. Labeled DNA was adjusted to 600 μl final volume in W&B buffer (10 mM Tris pH 7.5, 1 mM EDTA, 1 mM NaCl). In all, 10 μl Dynabeads MyOne C1 was washed twice with W&B buffer and added to DNA. Sample was incubated for 1 h at 4 °C. Beads were washed twice in W&B buffer, and resuspended in 100 μl elution buffer (95% v/v formamide, 10 mM EDTA). After incubation at 65 °C for 5 min, DNA was purified using a QiaGen PCR purification column. For sequencing, biotinylated overhangs were removed using S1 nuclease. Sequencing libraries were prepared using the NEBNext ChIP-seq Library Prep for Illumina kit. Library quality and quantity were analyzed using Bioanalyzer (Agilent). Samples were sequenced on a HiSeq 4000 using single-end 50 bp sequencing reads. A list of primers used for qRT-PCR analysis is available in Supplementary Table 1.

Nascent RNA capture followed by qRT-PCR. To measure nascent transcriptional changes at specific loci, ES cells were analyzed using the Click-it Nascent RNA Capture Kit (Life Technologies). Chd1Δ fla and Chd1Δ flaΔ cells were incubated with 0.2 μM 5-ethylidene uridine (EU) for 30 min to label nascent transcripts. Cells were washed, harvested by trypsinization and counted. Total RNA was isolated from the 10³ Chd1Δ fla or Chd1Δ flaΔ cells using the QiaGen RNeasy Mini Kit (Qiagen) and processed according to manufacturer’s instructions. qPCR was performed with KAPA SYBR FAST qPCR Master Mix (Kapa Biosystems) and amplified on a 7900HT Real-time PCR machine (Applied Biosystems).

Inhibitor treatments. Chd1Δ fla and Chd1Δ flaΔ (established) ES cells were used. Cells were incubated for indicated durations and concentrations with Trinitolide (Sigma), Etoposide (Sigma), or CX-5461 (Sigma). Control cells were treated with DMSO. Inhibitors were withdrawn and cells were washed immediately prior to detection of double-stranded DNA breaks and qPCR.

Cell cycle analysis. Chd1Δ flaΔ and Chd1Δ flaΔ (established) ES cells were used. Overnight cultured cells were incubated with 10 μM 5-ethylidene-2'-deoxyuridine (EdU) for 1 h. Cells were harvested, fixed in 4% PFA for 15 min, and permeabilized in 0.5% Triton X-100 in PBS for 5 min on ice. Cells were washed twice in PBS and incubated in yH2AX primary antibody (Abcam, ab2893) for 20 min at room temperature. Following a second wash, cells were incubated with fluorescein-conjugated secondary antibody (Life Technologies) for 20 min at room temperature. Subsequent EdU labeling was conducted according to manufacturer’s instructions of the Click-iT EdU Alexa Fluor 488 Flow Cytometry Assay Kit (Life Technologies). SYTOX Blue (Invitrogen) was used to detect DNA content at 1:1000 dilution. Data was collected on a Sony MA900 Multi-Application Cell Sorter, analyzed using FCS Express 7, and plotted using Prism 7.

Bioinformatic analyses

ChIP-seq, DSB-seq, and MNase-seq data processing. ChIP-seq and DSB-seq reads were mapped to the mm9 genome using Bowtie2 with the options --end-to-end -s --score-min L-1,5,0.3. Reads with mapping quality <13 were filtered out. PCR duplicates were removed by keeping at most one mapped read at each genomic position. Biological replicates were then combined. Read coverage profiles were generated using bedtools with the options --shift-0 --nomodel --extsize 200q mm and default q-value cutoff 0.05. In total, 5903 and 183 peaks were identified for Chd1 KO and WT cells, respectively. After filtering out peaks overlapping with blacklisted regions (http://mitra.stanford.edu/kundaje/akundaje/release/blacklist/mm9-mouse/mm9-blacklist.bedgs), 5671 and 54 peaks were retained for Chd1 KO and WT cells, respectively. The identified peaks were associated with genomic annotations using HOMER annotatePeaks.pl42. In particular, 1825 peaks in Chd1 KO cells were mapped around TSSs (~1kb to -100 bp) of 1785 genes (DSB-prone genes).

Gene annotations. ReSeq gene annotations were downloaded from UCSC Table Browser. In order to avoid ambiguity, only protein-coding genes with unique TSSs in autosomes were used. Genes whose TSSs ±10 kb overlap with blacklisted regions were further filtered out. As a result, 14,109 genes were retained and used throughout the paper. In Fig. 4B, we further required that the gene has a unique expression value provided by Guzman-Ayala et al.35 (GSE57699), resulting in 7524 genes. In Fig. 4D, 1,107 DSB prone genes (intersection between the 1785 genes identified by HOMER and the 14109 filtered genes) were compared with 13002 (=14,109–1107) non-DSB-prone genes. In Fig. 4F, the mean gene length of different isoforms was used for genes with different transcription termination sites (TTS). Gene ontology analysis was performed using the DAVID software44.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The ChIP-seq and DSB-seq data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under accession code GSE132137. The mass spectrometry data generated in this study have been deposited in the ProteomeXchange Consortium database via the PRIDE44 partner repository with the dataset identifier PXD024604. The MNase-seq data used in this study are available in the GEO database under the accession code GSE82127. Source data are provided with this paper.

Received: 12 June 2019; Accepted: 20 July 2021; Published online: 11 August 2021

References

1. Percharde, M., Bulut-Karslioglu, A. & Ramalho-Santos, M. Hypertranscription in development, stem cells, and regeneration. Dev. Cell 40, 9–21 (2017).
2. Bulut-Karslioglu, A. et al. Inhibition of mTOR induces a paused pluripotent state. Nature 540, 119–123 (2016).
3. Guzman-Ayala, M. et al. Chd1 is essential for the high transcriptional output and rapid growth of the mouse epiblast. Development 142, 118–127 (2014).
4. Jer, F. M. et al. Emergence of hematopoietic stem and progenitor cells involves a Chd1-dependent increase in total nascent transcription. Proc. Natl Acad. Sci. USA 112, E1734–E1743 (2015).
5. Percharde, M., Wong, P. & Ramalho-Santos, M. Global hypertranscription in the mouse embryonic germine. Cell Rep. 19, 1987–1996 (2017).
6. Lavado, A. et al. The Hippo pathway prevents YAP/TAZ-driven hypertranscription and controls neural progenitor number. Dev. Cell 47, 576–591.e8 (2018).
7. Li, Y. et al. Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. Cell Stem Cell 13, 446–458 (2013).
8. Flanagan, J. F. et al. Double chromodomains cooperate to recognize the methylated histone H3 tail. Nature 438, 1181–1185 (2005).
9. Gaspar-Maia, A. et al. Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature 460, 863–868 (2009).
10. Skene, P. J., Hernandez, A. E., Groudine, M. & Henikoff, S. The nucleosomal barrier to promoter escape by RNA polymerase II is overcome by the chromatin remodeling complex CHD1. Genes Dev. 25, 2042–2047 (2011).
13. Bunch, H. et al. Transcriptional elongation requires DNA break-induced signalling. Nat. Commun. 6, 10191 (2015).

14. King, I. F. et al. Topoisomerases facilitate transcription of long genes linked to autism. Nature 501, 58–62 (2013).

15. Wong, M. M., Belew, M. D., Kwieraga, A., Nhan, J. D. & Michael, W. M. Programmed DNA breaks activate the germline genome in Caenorhabditis elegans. Dev. Cell 46, 302–315.e5 (2018).

16. Teves, S. S. & Henikoff, S. DNA torsion as a feedback mediator of transcription and chromatin dynamics. Nature 55, 211–218 (2014).

17. Baranello, L. et al. DNA break mapping reveals topoisomerases II activity genome-wide. Int. J. Mol. Sci. 15, 13111–13122 (2014).

18. Snow, M. H. L. Gastrulation in the mouse: growth and regionalization of the embryonic stem cells through chemical mapping. Cell 167, 1555–1570.e15 (2016).

19. Farnung, L., Ochmann, M., Engeholm, M. & Cramer, P. Structural basis of nucleosome transcription mediated by Chd1 and FACT. BioRxiv https://doi.org/10.1101/2020.11.30.403857 (2020).

20. Percherie, M. et al. LINE-1 nucleosome partnership regulates early development and ESC identity. Cell 174, 391–405.e19 (2018).

21. Kari, V. et al. Loss of CHD1 causes DNA repair defects and enhances prostate cancer therapeutic responsiveness. EMBO Rep. 17, 1609–1623 (2016).

22. Zhou, J. et al. Human CHD1 is required for early DNA-damage signaling and is uniquely regulated by its N terminus. Nucleic Acids Res. 46, 3891–3905 (2018).

23. Huang, S. et al. Recurrent deletion of CHD1 in prostate cancer with relevance to cell invasiveness. Oncogene 31, 4164–4170 (2012).

24. Zhao, D. et al. Synthetic essentiality of chromatin remodelling factor CHD1 in PTEN-deficient cancer. Nature 542, 484–488 (2017).

25. Dellino, G. I. et al. Release of paused RNA polymerase II at specific loci favors DNA double-strand break formation and promotes cancer translocations. Nat. Genet. 47, 1 (2019).

26. Garcia-Muse, T. & Aguilera, A. Transcription-replication conflicts: how they occur and how they are resolved. Nat. Rev. Mol. Cell Biol. 17, 553–563 (2016).

27. Piquet, S. et al. The histone chaperone FACT coordinates H2A.X-dependent signaling and repair of DNA damage. Mol. Cell 72, 888–901.e7 (2018).

28. Zhao, D. et al. FACT is recruited to the +1 nucleosome of transcribed genes and spreads in a Chd1-dependent manner. BioRxiv https://doi.org/10.1101/2020.08.20.259960 (2020).

29. Chaudhuri, A. R. & Nussenzweig, A. The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. Nat. Rev. Mol. Cell Biol. 18, 610–621 (2017).

30. Dhar, S., Gursoy-Yuzugullu, O., Parasuram, R. & Price, B. D. The tale of a tail: histone H4 acetylation and the repair of DNA breaks. Philos. Trans. R Soc. Lond. B Biol. Sci. 372, 20160284 (2017).

31. Endres, T. et al. Ubiquitylation of MYC couples transcription elongation with double-strand break repair at active promoters. Mol. Cell 81, 830–844.e13 (2021).

32. Lee, J.-S. et al. Codependency of H2B monoubiquitination and nucleosome reassembly on Chd1. Gene Dev. 26, 914–919 (2012).

33. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods 4, 207–214 (2007).

34. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS 16, 284–287 (2012).

35. Szklarczyk, D. et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 47, gky1311 (2018).

36. Brookes, E. et al. Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. Cell Stem Cell 10, 157–170 (2012).

37. Crosetto, N. et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. Nat. Methods 10, 361–365 (2013).

38. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359 (2012).

39. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010).

40. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).

41. Heinz, S. et al. Simple combinations of lineage-determination transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589 (2010).

42. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: pathways toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13 (2009).

43. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 47, gky1108 (2018).

Acknowledgements

We thank members of the Santos Lab for input and critical reading of the manuscript, Richard Luo and members of the UCSF Institute of Human Genetics for assistance with sequencing and sonication. Samples were sequenced at UCSF Institute of Human Genetics Core Facility and Center for Advanced Technology, which is supported by the NIH (5S06CA082103). This research was supported by grants from Bloodwise and Medical Research Council plus the CRUK Manchester Centre award (C5759/AS254) to A.D.W. who is supported by the NIHRR Manchester Biomedical Research Centre, 2R01CA163363 to J.S.S., the Sophja Koervakskaia Award (Humboldt Foundation) to A.-K.B., and by NIH R01GM131041, R01GM123556, CIHR Project Grant 420231, and Canada 150 Research Chair in Developmental Epigenetics to M.R.-S.

Author contributions

A.B.-K. and M.R.-S. conceived the project. A.B.-K. designed and performed the majority of experiments, with the following exceptions: M.G.-A. isolated, stained, and imaged the embryos, A.J.K.W performed the mass spectrometry runs, H.J. and M.H. performed bioinformatics analyses, under the supervision of J.S.S. A.D.W supervised the mass spectrometry experiments. B.C., Y.K.K., and M.S. performed follow-up validation experiments, M.R.-S. supervised the project. A.B.-K., and M.R.-S. wrote the manuscript with feedback from all authors.

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/s41467-021-25088-3.

Correspondence and requests for materials should be addressed to A.B.-K. or M.R-S.

Peer review information Nature Communications thanks Alexander Schmidt and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

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