

MeCP2 links heterochromatin condensates and neurodevelopmental disease

<https://doi.org/10.1038/s41586-020-2574-4>

Received: 31 July 2019

Accepted: 15 July 2020

Published online: 22 July 2020

 Check for updates

Charles H. Li^{1,2,7}, Eliot L. Coffey^{1,2,7}, Alessandra Dall'Agnese¹, Nancy M. Hannett¹, Xin Tang¹, Jonathan E. Henninger¹, Jesse M. Platt^{1,3}, Ozgur Oksuz¹, Alicia V. Zamudio^{1,2}, Lena K. Afeyan^{1,2}, Jurian Schuijers^{1,5}, X. Shawn Liu^{1,6}, Styliani Markoulaki¹, Tenzin Lungjangwa¹, Gary LeRoy⁴, Devon S. Svoboda¹, Emile Wogram¹, Tong Ihn Lee¹, Rudolf Jaenisch^{1,2,8} & Richard A. Young^{1,2,8}

Methyl CpG binding protein 2 (MeCP2) is a key component of constitutive heterochromatin, which is crucial for chromosome maintenance and transcriptional silencing^{1–3}. Mutations in the *MECP2* gene cause the progressive neurodevelopmental disorder Rett syndrome^{3–5}, which is associated with severe mental disability and autism-like symptoms that affect girls during early childhood. Although previously thought to be a dense and relatively static structure^{1,2}, heterochromatin is now understood to exhibit properties consistent with a liquid-like condensate^{6,7}. Here we show that MeCP2 is a dynamic component of heterochromatin condensates in cells, and is stimulated by DNA to form liquid-like condensates. MeCP2 contains several domains that contribute to the formation of condensates, and mutations in *MECP2* that lead to Rett syndrome disrupt the ability of MeCP2 to form condensates. Condensates formed by MeCP2 selectively incorporate and concentrate heterochromatin cofactors rather than components of euchromatic transcriptionally active condensates. We propose that MeCP2 enhances the separation of heterochromatin and euchromatin through its condensate partitioning properties, and that disruption of condensates may be a common consequence of mutations in MeCP2 that cause Rett syndrome.

MeCP2 and HP1 proteins are key regulators of heterochromatin^{1–4}. Recent studies have shown that HP1 proteins are dynamic components of heterochromatin in vivo and can form phase-separated condensates in vitro, which suggests that heterochromatin is a dynamic liquid-like condensate^{6,7}. To confirm that MeCP2 is also a dynamic component of heterochromatin, we used live-cell fluorescence microscopy to image both MeCP2 and HP1 α , endogenously tagged with fluorescent proteins, in mouse embryonic stem (ES) cells (Fig. 1a–c, Extended Data Fig. 1). The results showed that green fluorescent protein (GFP)-tagged MeCP2 and mCherry-tagged HP1 α occur in nuclear bodies that overlap Hoechst-dense heterochromatin foci (Fig. 1a, Extended Data Fig. 1a) and that the two proteins occur in the same heterochromatin condensates (Extended Data Fig. 1b). Fluorescence recovery after photobleaching (FRAP) of MeCP2–GFP and HP1 α –mCherry puncta revealed recovery on the timescale of seconds (Fig. 1b, c, Extended Data Fig. 1c–f), consistent with characteristics of liquid-like condensates. These results show that MeCP2 is a dynamic component of heterochromatin condensates in live mouse ES cells.

To determine whether MeCP2 is a dynamic component of heterochromatin in mammalian tissues, we generated mice expressing MeCP2–GFP protein from the endogenous locus (Extended Data Fig. 2a–c). MeCP2 is reported to be expressed in all cell types (Extended Data Fig. 2d); we studied neurons because MeCP2 is highly abundant in these cells⁸ and mutations in *MECP2* cause neurodevelopmental disorders^{3–5}.

Imaging of MAP2-positive neurons revealed that MeCP2–GFP occurs in Hoechst-dense heterochromatin foci (Fig. 1d). FRAP analysis of MeCP2–GFP puncta revealed rapid and complete recovery on the timescale of seconds (Fig. 1e, f). These results indicate that MeCP2 is a dynamic component of liquid-like heterochromatin condensates in mouse brain cells.

To investigate whether MeCP2 has physicochemical properties that may contribute to heterochromatin condensates in cells, we examined purified MeCP2–GFP fusion protein using in vitro droplet assays. MeCP2–GFP formed spherical droplets that displayed properties consistent with phase-separated liquid condensates, including sensitivity to protein and salt concentration, droplet fusion behaviour, and dynamic rearrangement of molecules measured using FRAP (Extended Data Fig. 3a–i). Purified HP1 α –mCherry also formed droplets (Extended Data Fig. 3j), consistent with previous findings^{6,7}. MeCP2 binds to DNA^{3,4}, so we studied the effects of DNA on MeCP2 droplet formation. When DNA was added to physiologically relevant concentrations of MeCP2–GFP, MeCP2–GFP formed droplets (Fig. 1g). After the addition of methylated DNA (which MeCP2 binds with higher affinity^{3,4}), larger droplets were formed (Fig. 1g, h), and these droplets contained a larger fraction of MeCP2 (Fig. 1i) and increased levels of DNA (Extended Data Fig. 3k). These results were observed across a range of MeCP2 concentrations (Fig. 1j). These observations suggest that DNA can cause crowding of MeCP2 and thus lower the threshold

¹Whitehead Institute for Biomedical Research, Cambridge, MA, USA. ²Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA. ³Division of Gastroenterology, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA. ⁴Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY, USA. ⁵Present address: Center for Molecular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands. ⁶Present address: Department of Physiology and Cellular Biophysics, Columbia University Medical Center, New York, NY, USA. ⁷These authors contributed equally: Charles H. Li, Eliot L. Coffey. ⁸e-mail: jaenisch@wi.mit.edu; young@wi.mit.edu

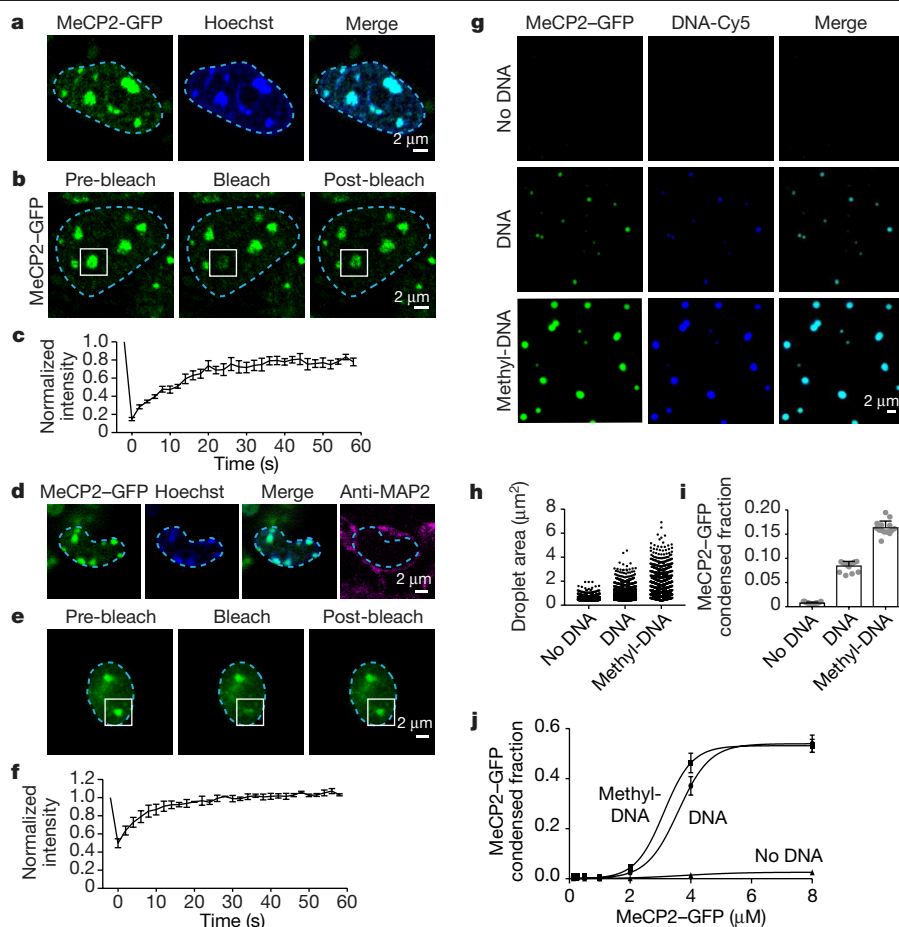


Fig. 1 | MeCP2 forms condensates in vivo and in vitro. **a**, Live-cell images of endogenous-tagged MeCP2-GFP and Hoechst staining in mouse ES cells. **b**, Live-cell images of FRAP experiments with endogenous-tagged MeCP2-GFP mouse ES cells. **c**, FRAP curves for experiments in **b**. Photobleaching occurs at $t = 0$ s. $n = 7$ cells. Data are mean $\pm$ s.e.m. **d**, Fixed-cell images of endogenous-tagged MeCP2-GFP brain sections from chimeric mice. **e**, Images of FRAP experiments performed on acute brain slices from endogenous-tagged MeCP2-GFP chimeric mice. **f**, FRAP curves for experiments in **e**. Photobleaching occurs at $t = 0$ s. $n = 3$ cells. Data are mean $\pm$ s.e.m. **g**, Droplet experiments examining

MeCP2 droplet formation with DNA. MeCP2-GFP at 2 μ M was mixed with 160 nM unmethylated DNA (DNA), methylated DNA (methyl-DNA), or no DNA in droplet formation buffers with 100 mM NaCl. **h**, Droplet areas for experiments in **g**. $n = 15$ fields per condition. **i**, MeCP2-GFP condensed fraction for experiments in **g**. $n = 15$ fields per condition. Data are mean $\pm$ s.d. **j**, MeCP2-GFP condensed fraction curves for experiments examining MeCP2 droplet formation with DNA. MeCP2-GFP was mixed with 160 nM DNA, methyl-DNA, or no DNA in droplet formation buffers with 100 mM NaCl. $n = 15$ fields per condition. Data are mean $\pm$ s.d.

for condensate formation, analogous to the manner in which enhancer DNA elements crowd transcription factors to lower the threshold for formation of transcriptional condensates⁹.

We used droplet assays to identify domains of MeCP2 that contribute to condensate formation. Intrinsically disordered regions (IDRs) can participate in condensate formation¹⁰, and MeCP2 contains two conserved IDRs that flank its structured methyl-DNA binding domain (MBD) (Fig. 2a, Extended Data Fig. 3l). We conducted droplet formation assays using physiologically relevant concentrations of recombinant MeCP2-GFP domain deletion mutant proteins in the presence of DNA. Although mutant proteins that lack the N-terminal IDR (Δ IDR-1) formed droplets, those that lack the C-terminal IDR (Δ IDR-2) did not (Fig. 2b–d). Furthermore, IDR-1 alone did not form droplets, whereas IDR-2 alone did – albeit with diminished size and number relative to both full-length and mutant Δ IDR-1 proteins (Fig. 2b–d). Similar results were observed when MeCP2-GFP domain deletion mutant proteins were examined in droplet assays in the absence of DNA (Extended Data Fig. 3m–o). These results indicate that the MeCP2 C-terminal IDR, which has previously been implicated in various functions including heterochromatin association¹¹, chromatin compaction¹², co-repressor recruitment¹³, and transcriptional repression¹⁴, contributes to condensate formation. Furthermore, the results indicate that the MBD also contributes to

condensate formation because DNA binding lowers the threshold for condensate formation.

Specific sequence features within IDRs have previously been found to contribute to condensate formation¹⁰; several of these features occur within the MeCP2 IDR-2 (Fig. 2a, Extended Data Fig. 3l), leading us to investigate whether these contribute to MeCP2 condensate behaviours. We found that deletion of basic patches within IDR-2 disrupted MeCP2 droplet formation, whereas deletion mutants removing aromatic residues, a histidine-rich patch, and a proline-rich patch remained capable of droplet formation (Fig. 2e–g). Droplet formation correlated with the ability to repress transcription, a key MeCP2 function¹⁴, in a transcriptional repression reporter assay (Fig. 2h, i). These results suggest that the basic patches in the MeCP2 C-terminal IDR, some of which are disrupted in Rett syndrome¹⁵, have especially important roles in MeCP2 condensate formation.

Active transcriptional condensates^{16,17} do not overlap heterochromatin condensates (Extended Data Fig. 4a). Although separation of euchromatin and heterochromatin can be attributed to different DNA-binding factors and differential association with nuclear lamina¹⁸, it is possible that the condensate properties of specific proteins might also contribute to the separation of these distinct compartments¹⁹. To investigate this possibility, we tested whether condensates formed by

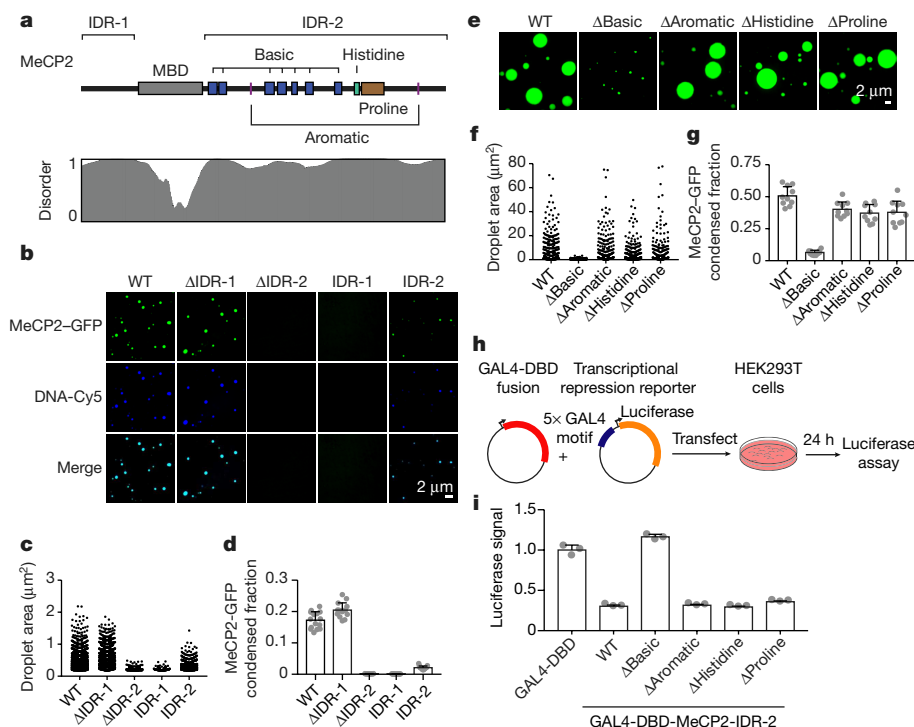


Fig. 2 | MeCP2 features that contribute to condensate formation.

a, Schematic of MeCP2 protein indicating the MBD, IDR-1, IDR-2, and sequence features within IDR-2 previously implicated in condensate formation for other proteins. Contribution of IDR-2 sequence features to condensate formation was examined using deletion mutants that remove the basic patches (ΔBasic), aromatic residues (ΔAromatic), histidine-rich patch (ΔHistidine), and proline-rich patch (ΔProline). Predicted protein disorder is displayed below. **b**, Droplet experiments examining the ability of MeCP2 deletion mutants to form droplets with DNA. MeCP2-GFP deletion mutants at 2 μM were mixed with 40 nM DNA in droplet formation buffers with 100 mM NaCl. WT, wild type. **c**, Droplet areas for experiments in **b**. $n = 15$ fields per condition. **d**, MeCP2-GFP condensed fraction for experiments in **b**. $n = 15$ fields per condition. **e**, Droplet

experiments examining ability of MeCP2 IDR-2 sequence feature deletion mutants to form droplets. MeCP2-GFP IDR-2 sequence feature deletion mutants at 10 μM were added to droplet formation buffers with 150 mM NaCl and 10% PEG-8000. **f**, Droplet areas for experiments in **e**. $n = 10$ fields per condition. **g**, MeCP2-GFP condensed fraction for experiments in **e**. $n = 10$ fields per condition. **h**, Schematic of transcriptional repression reporter assay used to examine the ability of MeCP2 IDR-2 sequence features to contribute to transcriptional repression. **i**, Normalized luciferase signals for reporter assay examining ability of MeCP2 IDR-2 sequence features to contribute to transcriptional repression. Luciferase signal was normalized to GAL4 DNA-binding domain (GAL4-DBD) alone. $n = 3$ biologically independent samples per condition. All data are mean $\pm$ s.d.

MeCP2 preferentially incorporate and concentrate HP1α compared to key components of euchromatic transcriptional condensates, such as MED1 and BRD4. We found that MeCP2-GFP droplets incorporated and concentrated HP1α-mCherry to a substantially greater extent than MED1 and BRD4 IDRs (Extended Data Fig. 4b–g). Similar results were obtained in the presence or absence of DNA (Extended Data Fig. 4b–e) and in the presence of nucleosomal DNA, albeit with less efficiency (Extended Data Fig. 4f, g). Nucleosomal DNA alone did not form droplets under these conditions (Extended Data Fig. 4h), although it enhanced MeCP2 droplet formation (Extended Data Fig. 4i–k). Notably, when combined with MeCP2-GFP in the presence or absence of DNA, BRD4-IDR-mCherry was enriched in a distinct droplet phase that did not coalesce with the MeCP2-GFP droplet phase, although the two phases appeared adjacent and touching (Extended Data Figs. 4b, d, 5). These results suggest that MeCP2 condensates may contribute to selective partitioning of components of heterochromatin and active euchromatin. There is some evidence for MeCP2 occupancy of euchromatin^{8,15,20}, but MeCP2 levels in active euchromatin may not be sufficient to form condensates that facilitate the partitioning of heterochromatin components.

In patients with Rett syndrome, mutations occur predominantly in the MeCP2 MBD and IDR-2 domains (Fig. 3a), which both contribute to condensate formation. To examine whether patient mutations in these domains disrupt the ability of MeCP2 to form condensates, we examined MeCP2-GFP proteins with Rett-syndrome-causing mutations using droplet formation assays (Fig. 3b–g, Extended Data Fig. 6). Patient

missense mutations that affect the MBD reduced the ability of MeCP2 to form droplets (Fig. 3d, e, Extended Data Fig. 6b). Similarly, patient mutations that truncate IDR-2 disrupted the ability of MeCP2 to form droplets, with mutations that truncated a greater portion of IDR-2 having a greater disruptive effect on droplet formation (Fig. 3b, c, Extended Data Fig. 6a). These results suggest that condensate disruption may be a common consequence of mutations in patients with Rett syndrome.

The observation that missense mutations that cause Rett syndrome occur frequently in the MBD, whereas truncation mutations occur frequently in IDR-2 (Fig. 3a), is consistent with a condensate model. Missense mutations in the structured MBD reduce condensate formation because DNA binding lowers the threshold for formation, whereas deletion mutations abrogate the multivalent interactions that contribute to IDR-mediated condensate formation. Nonetheless, in Rett syndrome there are missense mutations in IDR-2, so we investigated whether three of these mutations (P225R, R306C and P322L) disrupt condensate formation. All three mutations reduced the ability of MeCP2 to form droplets in vitro (Fig. 3f, g, Extended Data Fig. 6c–f). The R306C mutation was previously shown to disrupt an interaction between the MeCP2 NCoR-interaction domain (NID) and TBLR1—a subunit of the NCoR co-repressor complex^{13,21}. We therefore examined the ability of R306C mutant condensates to incorporate the C-terminal domain of TBLR1 (TBLR1-CTD), which directly interacts with the NID²¹. Wild-type MeCP2 droplets readily enriched TBLR1-CTD-mCherry, whereas R306C mutant droplets showed less enrichment (Extended Data Fig. 7), which suggests that MeCP2 condensates can contribute to NID-mediated

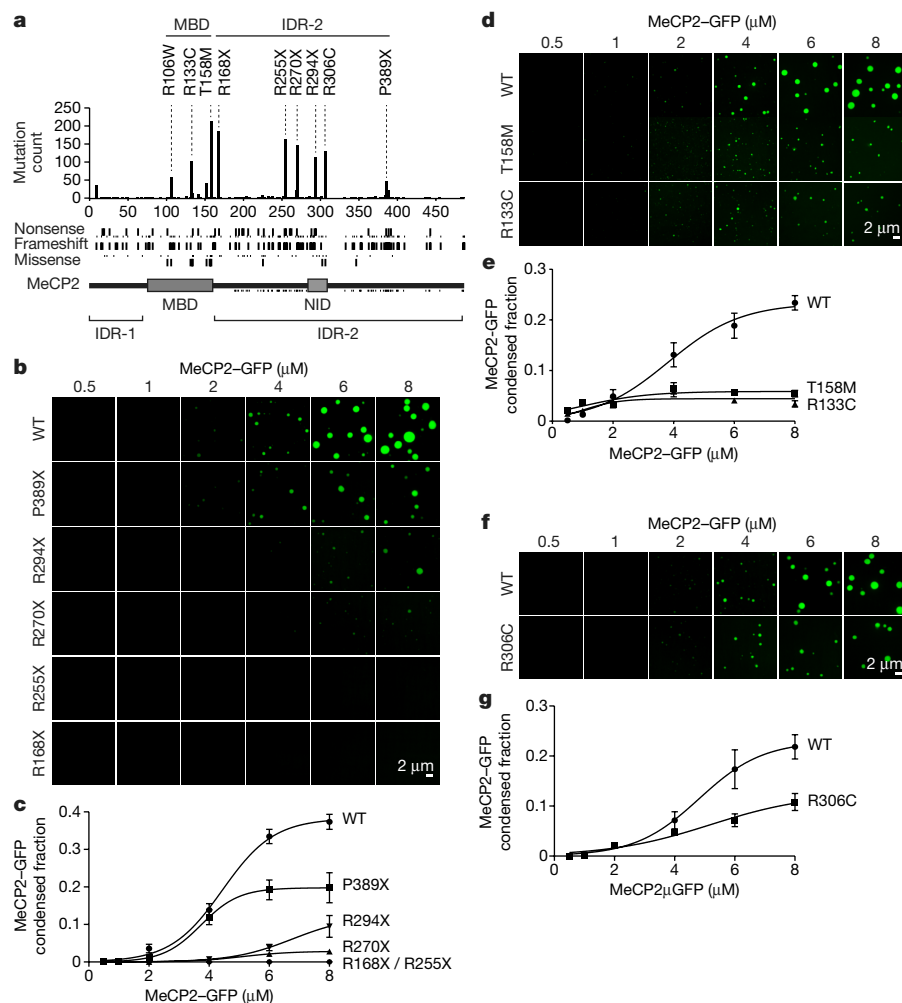


Fig. 3 | Mutations in patients with Rett syndrome disrupt MeCP2 condensate formation.

a, Schematic of MeCP2 protein with bar chart displaying the number of *MECP2* coding mutations in female patients with Rett syndrome found in the RettBASE database for each amino acid position. Positions of nonsense, frameshift, and missense mutations are shown below. **b**, Droplet experiments examining effects of Rett syndrome truncation mutations that disrupt IDR-2 on MeCP2 droplet formation. Wild-type MeCP2-GFP and Rett syndrome IDR-2 mutants (R168X, R255X, R270X, R294X and P389X) at the indicated concentrations were mixed with 40 nM methylated DNA in droplet formation buffers with 100 mM NaCl. **c**, MeCP2-GFP condensed fraction as a function of MeCP2-GFP concentration for experiments in **b**. $n = 15$ fields per condition. **d**, Droplet experiments examining the effects of Rett syndrome missense mutations that disrupt the MBD on MeCP2 droplet formation. Wild-type MeCP2-GFP and Rett syndrome MBD mutants (R133C and T158M) at indicated concentrations were mixed with 20 nM methylated DNA in droplet formation buffers with 100 mM NaCl. **e**, MeCP2-GFP condensed fraction as a function of MeCP2-GFP concentration for experiments in **d**. Fields per condition $n = 15$. **f**, Droplet experiments examining the effects of Rett syndrome missense mutation R306C on MeCP2 droplet formation. Wild-type and R306C mutant MeCP2-GFP at indicated concentrations were mixed with 20 nM methylated DNA in droplet formation buffers with 100 mM NaCl. **g**, MeCP2-GFP condensed fraction as a function of MeCP2-GFP concentration for experiments in **f**. $n = 15$ fields per condition. All data are mean $\pm$ s.d.

recruitment of NCoR, a key MeCP2 function previously shown to be disrupted in Rett syndrome^{13,21}. These results suggest that missense mutations in IDR-2 that occur in patients with Rett syndrome contribute to condensate disruption.

A minimal MeCP2 fragment (Mini), which removes most of IDR-2 but retains the NID (and thus R306) (Extended Data Fig. 8a), can partially rescue Rett syndrome phenotypes in a mouse model of the disorder²². This observation led us to investigate whether MeCP2 Mini protein can form droplets. MeCP2 Mini was capable of forming droplets (Extended Data Fig. 8b–d) that could enrich DNA and HP1 α -mCherry (Extended Data Fig. 8e–g), as well as TBLR1-CTD-mCherry (Extended Data Fig. 8h, i). Furthermore, live-cell imaging of mouse ES cells expressing endogenously tagged wild-type MeCP2-GFP and the Mini fragment showed that both proteins partitioned similarly into heterochromatin condensates (Extended Data Fig. 8j, k). These results show that MeCP2 Mini retains condensate formation capabilities and suggest that this ability to form condensates may contribute to the partial rescue of Rett syndrome phenotypes.

To explore the possibility that patient mutations that cause loss of IDR-2 lead to deficiencies in condensate incorporation in living cells, we focused on the common R168X patient mutation, which completely deletes IDR-2 and corresponds to the Δ IDR-2 deletion mutant used to examine condensate-forming ability of MeCP2 in vitro (Fig. 2b–d, Extended Data Fig. 3m–o). We examined mouse ES cells expressing endogenously tagged wild-type and R168X mutant MeCP2-GFP proteins (Extended Data Fig. 9). Live-cell imaging showed a marked reduction in the ability of mutant protein to partition into heterochromatin condensates (Extended Data Fig. 9a, b). Reduced partitioning was not a simple consequence of

decrease in the abundance of mutant protein (Extended Data Fig. 9c, d), as partitioning into heterochromatin condensates was not rescued by over-expression of the R168X mutant (Extended Data Fig. 10a–c). Reduced partitioning of MeCP2 into heterochromatin condensates was also observed in R168X mutant neurons (Fig. 4a, b, Extended Data Fig. 10d–f). These results indicate that mutations that occur in patients with Rett syndrome reduce the condensate interactions of MeCP2 in cells.

Rett syndrome is associated with various cellular phenotypes, including altered chromatin architecture²³, disrupted cofactor recruitment^{13,15}, and widespread transcriptional dysregulation^{8,20,24}. R168X mutant mouse ES cells showed evidence of each of these disease-associated cellular phenotypes. R168X mutant mouse ES cells displayed changes in chromatin architecture, as heterochromatin condensates increased in number (Extended Data Fig. 9e) but decreased in volume (Extended Data Fig. 9f). Mutant cells showed reduced ability to partition the HP1 α cofactor into heterochromatin condensates (Extended Data Fig. 9g, h), which was not due to reduced HP1 α abundance (Extended Data Fig. 9i), consistent with the ability of MeCP2 condensates to selectively partition and concentrate HP1 α in vitro (Extended Data Fig. 4b–g). R168X mutant mouse ES cells displayed evidence of widespread transcriptional dysregulation with loss of heterochromatin-associated repetitive element silencing (Extended Data Fig. 9j), reduced total RNA abundance (Extended Data Fig. 9k), and broad downregulation of euchromatic genes (Extended Data Fig. 9l). These cellular phenotypes associated with Rett syndrome were also observed in R168X mutant neurons (Fig. 4c–f, Extended Data Fig. 10g–i). Thus, the loss of the IDR-2 domain, which has a major role in condensate formation, produced a range of cellular phenotypes that are associated with Rett syndrome.

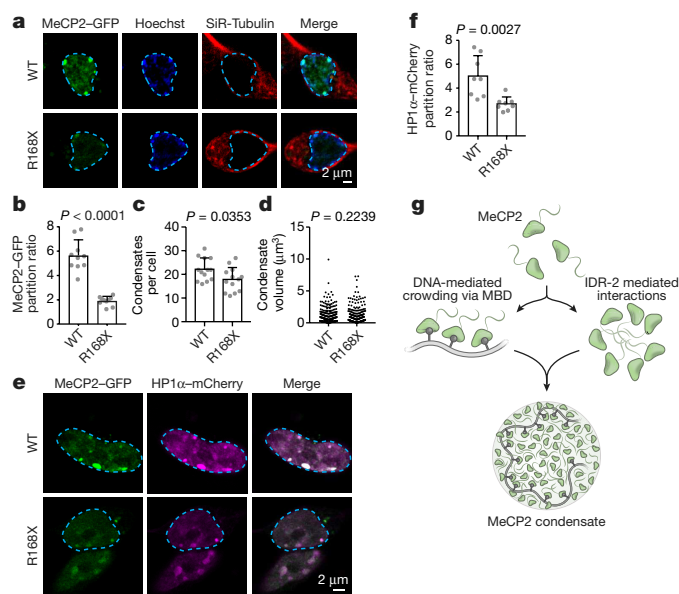


Fig. 4 | R168X mutant MeCP2 displays reduced partitioning into heterochromatin condensates and causes disease-relevant cellular phenotypes in neurons. **a**, Live-cell images of endogenous-tagged wild-type and R168X MeCP2–GFP mutant proteins with Hoechst and SiR-tubulin staining in neurons. **b**, Partition ratios of MeCP2–GFP proteins at heterochromatin condensates for experiments in **a**. $n = 10$ cells per condition. $P < 0.0001$, $t = 8.8921$, $df = 18$, two-tailed Student's t -test. **c**, Number of heterochromatin condensates per cell in endogenous-tagged wild-type and R168X mutant MeCP2–GFP neurons. $n = 13$ cells per condition. $P = 0.0353$, $t = 2.2314$, $df = 24$, two-tailed Student's t -test. **d**, Heterochromatin condensate volumes in endogenous-tagged wild-type MeCP2–GFP and R168X mutant neurons. Condensates per condition: WT ($n = 311$), R168X ($n = 252$). $P = 0.2239$, $t = 1.2176$, $df = 561$, two-tailed Student's t -test. **e**, Live-cell images of endogenous-tagged MeCP2–GFP (wild type or R168X mutant) and HP1 α –mCherry in neurons. **f**, Partition ratios of HP1 α –mCherry at heterochromatin condensates for experiments in **e**. $n = 8$ cells per condition. $P = 0.0027$, $t = 3.6444$, $df = 14$, two-tailed Student's t -test. **g**, Model of interactions contributing to MeCP2 condensate formation with DNA. All data are mean $\pm$ s.d.

We propose a condensate model for MeCP2 (Fig. 4g) that incorporates our conventional understanding of the mechanisms by which MeCP2 dysregulation contributes to cellular phenotypes, but adds the view that large numbers of MeCP2 molecules, using several weak and dynamic interactions, form membrane-less bodies that can concentrate and compartmentalize additional components engaged in heterochromatin function. A recent study also reported that MeCP2 exhibits condensate properties that may be relevant to its interaction with histone H1²⁵. Our results suggest a link between Rett syndrome mutations, altered MeCP2 condensate properties, and disease-associated cellular phenotypes. The MeCP2 MBD and IDR-2 domains are both required for efficient condensate formation, and because several mutations in these domains in Rett syndrome disrupt condensate formation, condensate disruption may be a common pathway for disease pathology caused by mutations in both domains. Rett syndrome mutations can also reduce levels of MeCP2 protein²⁶, which may contribute to condensate disruption, as condensates can be highly sensitive to protein concentration¹⁰. Rett syndrome mutations are a leading cause of intellectual disability in women and girls, yet evidence in animal models indicates that some symptoms may be reversible if a suitable therapy were to be developed^{22,27,28}. We suggest that new approaches

to the pharmacological modification of condensate behaviours^{29,30}, if developed to selectively affect heterochromatin condensates, might provide therapeutic benefits for patients with Rett syndrome.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2574-4>.

- Janssen, A., Colmenares, S. U. & Karpen, G. H. Heterochromatin: guardian of the genome. *Annu. Rev. Cell Dev. Biol.* **34**, 265–288 (2018).
- Allshire, R. C. & Madhani, H. D. Ten principles of heterochromatin formation and function. *Nat. Rev. Mol. Cell Biol.* **19**, 229–244 (2018).
- Lyst, M. J. & Bird, A. Rett syndrome: a complex disorder with simple roots. *Nat. Rev. Genet.* **16**, 261–275 (2015).
- Ip, J. P. K., Mellios, N. & Sur, M. Rett syndrome: insights into genetic, molecular and circuit mechanisms. *Nat. Rev. Neurosci.* **19**, 368–382 (2018).
- Amir, R. E. et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23**, 185–188 (1999).
- Larson, A. G. et al. Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* **547**, 236–240 (2017).
- Strom, A. R. et al. Phase separation drives heterochromatin domain formation. *Nature* **547**, 241–245 (2017).
- Skene, P. J. et al. Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol. Cell* **37**, 457–468 (2010).
- Shrinivas, K. et al. Enhancer features that drive formation of transcriptional condensates. *Mol. Cell* **75**, 549–561.e7 (2019).
- Shin, Y. & Brangwynne, C. P. Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382 (2017).
- Kumar, A. et al. Analysis of protein domains and Rett syndrome mutations indicate that multiple regions influence chromatin-binding dynamics of the chromatin-associated protein MECP2 in vivo. *J. Cell Sci.* **121**, 1128–1137 (2008).
- Georgel, P. T. et al. Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J. Biol. Chem.* **278**, 32181–32188 (2003).
- Lyst, M. J. et al. Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nat. Neurosci.* **16**, 898–902 (2013).
- Nan, X., Campoy, F. J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471–481 (1997).
- Baker, S. A. et al. An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. *Cell* **152**, 984–996 (2013).
- Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
- Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855.e16 (2018).
- van Steensel, B. & Belmont, A. S. Lamina-associated domains: links with chromosome architecture, heterochromatin, and gene repression. *Cell* **169**, 780–791 (2017).
- Gibson, B. A. et al. Organization of chromatin by intrinsic and regulated phase separation. *Cell* **179**, 470–484.e21 (2019).
- Chahrouh, M. et al. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**, 1224–1229 (2008).
- Kruusvee, V. et al. Structure of the MeCP2-TBLR1 complex reveals a molecular basis for Rett syndrome and related disorders. *Proc. Natl Acad. Sci. USA* **114**, E3243–E3250 (2017).
- Tillotson, R. et al. Radically truncated MeCP2 rescues Rett syndrome-like neurological defects. *Nature* **550**, 398–401 (2017).
- Linhoff, M. W., Garg, S. K. & Mandel, G. A high-resolution imaging approach to investigate chromatin architecture in complex tissues. *Cell* **163**, 246–255 (2015).
- 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).
- Wang, L. et al. Rett syndrome-causing mutations compromise MeCP2-mediated liquid-liquid phase separation of chromatin. *Cell Res.* **30**, 393–407 (2020).
- Brown, K. et al. The molecular basis of variable phenotypic severity among common missense mutations causing Rett syndrome. *Hum. Mol. Genet.* **25**, 558–570 (2016).
- Guy, J., Gan, J., Selfridge, J., Cobb, S. & Bird, A. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **315**, 1143–1147 (2007).
- Giacometti, E., Luikenhuis, S., Beard, C. & Jaenisch, R. Partial rescue of MeCP2 deficiency by postnatal activation of MeCP2. *Proc. Natl Acad. Sci. USA* **104**, 1931–1936 (2007).
- Wheeler, R. J. et al. Small molecules for modulating protein driven liquid-liquid phase separation in treating neurodegenerative disease. Preprint at <https://www.biorxiv.org/content/10.1101/721001v2> (2019).
- Klein, I. A. et al. Partitioning of cancer therapeutics in nuclear condensates. *Science* **368**, 1386–1392 (2020).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

Methods

Cell culture

V6.5 mouse ES cells were a gift from R. Jaenisch, and were authenticated by STR analysis compared to commercially acquired cells of the same line. MeCP2-GFP Mini (Δ NIC)²² mouse ES cells were a gift from A. P. Bird, and were not subject to authentication.

ES cells were cultured in 2i/LIF medium on tissue culture-treated plates coated with 0.2% gelatin (Sigma G1890). ES cells were grown in a humidified incubator with 5% CO₂ at 37 °C. Cells were passaged every 2–3 days by dissociation using TrypLE Express (Gibco 12604). The dissociation reaction was quenched using serum/LIF medium. Cells were tested regularly for mycoplasma using the MycoAlert Mycoplasma Detection Kit (Lonza LT07-218) and found to be negative.

The composition of N2B27 medium is as follows: DMEM/F12 (Gibco 11320) supplemented with 0.5× N2 supplement (Gibco 17502), 0.5× B27 supplement (Gibco 17504), 2 mM L-glutamine (Gibco 25030), 1× MEM non-essential amino acids (Gibco 11140), 100 U ml⁻¹ penicillin–streptomycin (Gibco 15140), 0.1 mM 2-mercaptoethanol (Sigma M7522).

The composition of 2i/LIF medium is as follows: N2B27 medium, 3 μ M CHIR99021 (Stemgent 04-0004), 1 μ M PD0325901 (Stemgent 04-0006), and 1000 U ml⁻¹ leukaemia inhibitor factor (LIF) (ESGRO ESG1107).

The composition of serum/LIF medium is as follows: KnockOut DMEM (Gibco 10829) supplemented with 15% fetal bovine serum (Sigma F4135), 2 mM L-glutamine (Gibco 25030), 1× MEM non-essential amino acids, 100 U ml⁻¹ penicillin–streptomycin (Gibco 15140), 0.1 mM 2-mercaptoethanol (Sigma M7522), and 1,000 U ml⁻¹ LIF (ESGRO ESG1107).

HEK293T cells were purchased from ATCC (ATCC CRL-3216) and cultured in DMEM (Gibco 11995-073) with 10% fetal bovine serum (Sigma F4135), 100 U ml⁻¹ penicillin–streptomycin (Gibco 15140), 2 mM L-glutamine (Gibco 25030). Cells were not subject to authentication. Cells were tested regularly for mycoplasma using the MycoAlert Mycoplasma Detection Kit (Lonza LT07-218) and found to be negative.

Genome editing

The CRISPR–Cas9 system was used to generate genetically modified ES cell lines. Target-specific sequences were cloned into a plasmid containing sgRNA backbone, a codon-optimized version of Cas9, and mCherry or BFP. For generation of the MeCP2-mEGFP and HP1 α -mCherry endogenously tagged lines, homology directed repair templates were cloned into pUC19 using NEBuilder HiFi DNA Master Mix (NEB E2621S). The homology repair template consisted of mEGFP or mCherry cDNA sequence flanked on either side by 800 bp homology arms amplified from genomic DNA using PCR. The following sgRNA sequences with PAM sequence in parentheses were used for CRISPR–Cas9 targeting:

sgRNA_Mecp2_C-term: GTAAAGTCAGCTAACTCTCT (CGG); sgRNA_Mecp2_R168: gAGGTGGTTTCTGCTCTCTCC (TGG); sgRNA_Cbx5_C-term: gAAGAAAGCGCGAAGAGCTAA (AGG).

To generate genetically modified cell lines, 750,000 cells were transfected with 833 ng Cas9 plasmid and 1,666 ng nonlinearized homology repair template using Lipofectamine 3000 (Invitrogen L3000). Cells were sorted 48 h after transfection for the presence of either mCherry or BFP fluorescence proteins encoded on the Cas9 plasmid to enrich for transfected cells. This population was allowed to expand for 1 week before sorting a second time for the presence of GFP or mCherry. Approximately 40,000 GFP- or mCherry-positive cells were plated in serial dilution in a 6-well plate and allowed to expand for a week before individual colonies were manually picked into a 96-well plate. Twenty-four colonies were screened for successful targeting using PCR genotyping to confirm insertion. PCR genotyping was performed using Phusion polymerase (Thermo Scientific F531S). Products were amplified according to kit recommendations and visualized on a 1% agarose gel. The following primers were used for PCR genotyping:

MeCP2-GFP_fwd: AGCAGCATCTGCAAAGAAGAG; MeCP2-GFP_rev: CAGAGCCCTACCCATAAGGAG; HP1 α -mCherry_fwd: AACGTGAAGT-GTCCACAGATTG; HP1 α -mCherry_rev: TTATGGATGCGTTTAGGATGG; MeCP2-GFP_R168X_fwd: AGACACCTCCTTGGACCCTAA; MeCP2-GFP_R168X_rev: ACCCTTTTCACCTGAACACCT.

Neuronal differentiation

Neurons were derived from mouse ES cells by expression of *NGN2* to induce neuronal differentiation^{31,32}. A doxycycline-inducible *NGN2* expression construct with a puromycin-resistance gene was integrated into mouse ES cells using the PiggyBac transposon system. ES cells with successful integration of the expression construct were selected with puromycin (Gibco A1113803). Before induction of neuronal differentiation, mouse ES cells were seeded in 2i/LIF medium onto a layer of mouse astrocytes grown on either tissue culture treated plates or 35 mm glass plates (MatTek P35G-1.5-20-C) coated with poly-L-ornithine (Sigma P4957) and laminin (Corning 354232). Twenty-four h after seeding mouse ES cells, *NGN2* expression was induced by changing to N2B27 medium with 2 μ g ml⁻¹ doxycycline (Sigma D9891). Medium was changed daily with 2 μ g ml⁻¹ doxycycline in N2B27 medium. Neurons were collected for experiments 5 days after induction of *NGN2* expression, and neuronal status was confirmed by immunofluorescence staining for TuJ1 (Covance MMS-435P).

Live-cell imaging

Cells were grown on 35 mm glass plates (MatTek P35G-1.5-20-C) coated with poly-L-ornithine (Sigma P4957) for 30 min at 37 °C followed by coating with laminin (Corning 354232) for 2 h at 37 °C, and imaged in 2i/LIF medium using an LSM880 confocal microscope with Airyscan detector (Zeiss). Cells were imaged on a 37 °C heated stage supplemented with 37 °C humidified air. In addition, the microscope was enclosed in an incubation chamber heated to 37 °C. ZEN Black Edition v.2.3 (Zeiss) software was used for acquisition. Images were acquired with the Airyscan detector in super-resolution (SR) mode with a Plan-Apochromat 63×/1.4 oil objective. Raw Airyscan images were processed using ZEN v. 2.3.

To quantify MeCP2 condensate volumes, Z-stack images were taken using the ZEN v.2.3 software. Cells were treated with SiR-Hoechst (also known as SiR-DNA dye) (Cytoskeleton CY-SC007) to stain DNA or SiR-Tubulin (Cytoskeleton CY-SC002) to stain tubulin to facilitate cell identification and microscope focusing. Far-red (SiR-DNA) signal was used to determine the upper- and lower-Z boundaries of the nucleus. Then, images were taken in both 488 nm channel (MeCP2-GFP) and the 643 nm channel (SiR-DNA) at 0.19- μ m steps up through the nucleoplasm. Images are the result of a single Airyscan image, processed using the ZEN v.2.3 software. Heterochromatin condensate volumes were calculated using a custom script (www.github.com/jehenninger/MECP2_neuron) in Python v.3.4.3. To calculate heterochromatin condensate volumes, the SiR-DNA signal was used to define nuclear boundaries for a given cell. Heterochromatin condensates were identified as signal dense objects within the nuclear boundary with an empirical cutoff of 2.35 s.d. above the mean signal. Once identified, the volume of each heterochromatin condensate was quantified.

Fluorescence recovery after photobleaching (FRAP) was used to investigate dynamic internal rearrangement and internal-external exchange of molecules within heterochromatin foci, which are properties expected for liquid-like condensates³³. FRAP was performed on LSM880 Airyscan microscope with 488 nm and 561 nm lasers. Bleaching was performed at 100% laser power and images were collected every two seconds. Each image uses the LSM880 Airyscan averaging capacity and is the averaged result of two images. The combined image was then processed using ZEN v.2.3. FIJI/ImageJ (v.2.0.0-rc-65) was used to calculate intensity values in images. Recovery after photobleaching was calculated by first subtracting background values, and then quantifying fluorescence intensity lost within the bleached condensate normalized to signal within a condensate in a separate, neighbouring

Article

cell to account for image acquisition photobleaching. Post-bleach image taken 12 s post-photobleaching.

Partition ratios were used to quantify the ability of a protein to partition into heterochromatin condensates relative to the nucleoplasm in live-cell imaging experiments. A partition ratio for each cell was calculated as the ratio of the average pixel intensity within heterochromatin condensates relative to the average pixel intensity within 8–12 non-heterochromatic nucleoplasmic regions. Heterochromatin condensates and the nucleoplasm were defined using Hoechst staining. A single focal plane was analysed for each cell and cells with two or more heterochromatin condensates were used for analysis.

Immunofluorescence microscopy

Mouse ES cells were plated onto glass coverslips in preparation for immunostaining. After 24 h, cells were fixed with 4% paraformaldehyde in PBS, washed 3 times with PBS, and then permeabilized with 0.5% Triton X-100 in PBS for 10 min at room temperature. Cells were then washed three times with PBS, blocked for 1 h in 4% IgG-free BSA (Jackson ImmunoResearch 001-000-162) in PBS, and then stained overnight with the indicated antibody in 4% IgG-free BSA at room temperature in a humidified chamber. Cells were then washed three times with PBS. Secondary antibodies were added to cells in 4% IgG-free BSA and incubated for 1 h at room temperature. Cells were then washed twice in PBS. Cells were stained with Hoechst dye (Molecular Probes H3570) in PBS or water for 5 min, and then mounted in Vectashield mounting medium (Vector Laboratories H-1000). Imaging was performed on an RPI spinning disk confocal at 100× magnification using the MetaMorph software v.7.10.3.279 (Molecular Devices). Primary antibodies: anti-MAP2 (Invitrogen MA5-12823) and anti-MED1 (Abcam ab64965). Secondary antibodies: anti-mouse Alexa Fluor 568 (Invitrogen A11031) and anti-rabbit Alexa Fluor 555 (Invitrogen A21428).

Overexpression

To examine the ability of wild-type MeCP2–GFP and R168X mutant proteins to partition into heterochromatin condensates when present at high levels in cells, the proteins were overexpressed in cells with N-terminal GFP fusions by transfection of expression constructs. Seven hundred and fifty thousand mouse ES cells were transfected with 2.5 µg wild-type MeCP2–GFP or R168X mutant expression plasmid using Lipofectamine 3000 (Invitrogen L3000). The following day, cells were dissociated and seeded onto poly-L-ornithine and laminin-coated 35 mm glass-bottom dishes (MatTek P35G-1.5-20-C) for live-cell imaging the next day.

Flow cytometry

Relative expression levels of endogenous-tagged proteins were examined using flow cytometry. Mouse ES cells were dissociated using TrypLE Express (Gibco 12604) and the dissociation reaction was quenched using serum/LIF medium. Cells were resuspended in single cell suspension in PBS and passed through a cell strainer (Corning 352235). Cells were analysed using a LSRII flow cytometer (BD) and data were analysed using FlowJo v.10 (BD). Standard forward and side scatter gating was used to exclude debris and isolate singlet cells. Example flow cytometry gating strategy is shown in Supplementary Fig. 2. Mean fluorescence intensity was quantified for the singlet population, to determine the relative levels of endogenous-tagged fluorescent fusion proteins (MeCP2–GFP and HP1α–mCherry).

Western blot

Western blot was used to confirm expression of wild-type MeCP2–GFP and R168X proteins. Cell lysates were prepared by resuspending cell pellets in 2× Laemmli buffer and incubating at room temperature for 20 min. Lysates were then sonicated using a probe sonicator and boiled at 95 °C for 10 min. Samples were run on a 4–12% Bis-Tris polyacrylamide gel (Bio-Rad 3450125) using XT MOPS running buffer (Bio-Rad

1610788) at 80 V for 20 min, followed by 150 V until dye front reached the end of the gel. Protein was wet transferred to a 0.45-µm PVDF membrane (Millipore IPVH00010) in ice-cold transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) at 250 mA for 2 h at 4 °C. After transfer, the membrane was blocked with 5% non-fat milk in TBS for 1 h at room temperature, then incubated with 1:1,000 anti-GFP (Takara Bio 632381) or 1:1,000 anti-histone H3 (Cell Signaling Technology 4499) antibody in 5% non-fat milk in TBST overnight at 4 °C. Primary antibodies were validated by their respective vendors. After washing with TBST, the membrane was incubated with 1:10,000 horseradish peroxidase (HRP)-conjugated anti-Mouse IgG (GE Healthcare NXA931V) or anti-Rabbit IgG (GE Healthcare NA934V) secondary antibody diluted in 5% non-fat milk in TBST for 2 h at room temperature. After washing with TBST, the membrane was developed with chemiluminescent HRP substrate (Millipore WBKLS0100) and imaged using a CCD camera. Images were captured and analysed using Image Lab software v.6.0.1 (Bio-Rad).

Generation of chimeric mice

To generate endogenous MeCP2–GFP tagged chimeric mice, we injected endogenous MeCP2–GFP-tagged mouse ES cells grown on irradiated mouse embryonic fibroblasts into 8-cell embryos or blastocysts and implanted into pseudo-pregnant female CD1-IGS mice. Chimeric mice were identified on the basis of the colour of the fur and female chimaeras were used for imaging experiments at 10-weeks of age. Mouse studies were observational in nature, and not subject to randomization or blinding. Imaging experiments using mouse brain tissues were performed with a sample size of three cells, which is in line with other studies of condensates in cells¹⁶. All experiments using mice were carried out with approval from the MIT Committee on Animal Care (CAC) under protocol number 1019-059-22. Experiments were carried out under the supervision of the Division of Comparative Medicine (DCM) at MIT, which provides centralized management of the animal facility at the Whitehead Institute for Biomedical Research. The mouse facility conforms to federal guidelines (Animal Welfare Assurance Number A3125-01), and MIT is accredited by the Assessment and Accreditation of Laboratory Animal Care (AAALAC). Routine bedding, food, and water changes were performed by DCM. Mice were housed in a centrally controlled environment with a 12-h light/12-h dark cycle, temperature of 20–22.2 °C, and humidity of 30–50%. No statistical methods were used to predetermine sample size.

Brain slice immunofluorescence microscopy

Adult female endogenous MeCP2–GFP chimeric mice were perfused with 10% formalin to fix brain tissues. After fixation, mouse brains were incubated at 4 °C in 30% sucrose for 3 days. Brains were then sectioned using cryostat (Leica CM3050 S). Brain sections were then placed on a slide and stored at –20 °C. For immunofluorescence, brain sections were allowed to warm to room temperature, fixed with 4% paraformaldehyde for 10 min, and immunofluorescence was performed as described in the above immunofluorescence microscopy methods section.

To determine size and number of heterochromatin condensates in mouse neurons, brain sections taken from endogenous-tagged MeCP2–GFP mice (described above) were stained with anti-MAP2 to indicate neuronal cells. Sections were then imaged with 0.2-µm Z-stacks using MetaMorph v.7.10.3.279 (Molecular Devices). In MAP2-positive cells, endogenous MeCP2–GFP signal was used to determine the size and number of heterochromatin condensates using FIJI/ImageJ v.2.0.0-rc-65 3D object counter with automatic threshold determination.

Brain slice FRAP

Organotypic brain slices were obtained from adult endogenous MeCP2–GFP chimeric mice, based on a slight modification of a previously described method³⁴. After decapitation, the brain was extracted and placed into ice-cold dissection medium composed of hibernate A

(BrainBits HA), 2% B27 supplement (Gibco 17504), 2 mM L-glutamine (Gibco 25030), and 1% penicillin–streptomycin (Gibco 15140). The cerebellum and midbrain were removed and the remaining cerebral hemispheres were separated and sliced coronally at 250 μ m thickness using a McIlwain tissue chopper (Ted Pella MTC/2E). The slices were gently separated from each other in chilled dissection medium and transferred onto glass-bottom dishes in culture medium containing Neurobasal A (Gibco 10888022) with 2% B27 supplement, 2 mM L-glutamine, and 1% penicillin–streptomycin. Imaging was performed immediately after brain slice preparation. FRAP experiment was performed using the Andor Revolution spinning disk confocal with the FRAPPA module (Andor Technology). Bleaching was performed using 5–7 pulses of 20- μ s dwell time and images were collected every second. Fluorescence intensity was measured using Fiji/ImageJ v.2.0.0-rc-65 and analysed as described above. Post-bleach image was taken 12 s after photobleaching.

Protein purification

Human cDNA was cloned into a modified version of a T7 pET expression vector. The base vector was engineered to include sequences encoding a N-terminal 6 $\times$ His followed by either mEGFP or mCherry and a 14-amino-acid linker sequence 'GAPGSAAGSG'. cDNA sequences, generated by PCR, were inserted in-frame after the linker sequence using NEBuilder HiFi DNA Assembly Master Mix (NEB E2621S). Mutant cDNA sequences were generated by PCR and inserted into the same base vector as described above. All expression constructs were subject to Sanger sequencing to confirm sequence identity. The following human proteins were used in experiments:

MeCP2 full length (WT): residues 1-486; MeCP2 Δ IDR-1: residues 78-486; MeCP2 Δ IDR-2 (R168X): residues 1-167; MeCP2 IDR-1: residues 1-77; MeCP2 IDR-2: residues 168-486; MeCP2 Δ Basic: residues 1-486, removing IDR-2 basic patches (residues 170-181, 184-194, 246-258, 263-274, 282-289, 301-310, and 340-348); MeCP2 Δ Aromatic: residues 1-486, removing IDR-2 aromatic residues (F226 and Y450); MeCP2 Δ Histidine: residues 1-486, removing IDR-2 histidine-rich domain (residues 366-372); MeCP2 Δ Proline: residues 1-486, removing IDR-2 proline-rich domain (residues 376-405); MeCP2 R133C: residues 1-486, R133C; MeCP2 T158M: residues 1-486, T158M; MeCP2 P225R: residues 1-486, P225R; MeCP2 R255X: residues 1-254; MeCP2 R270X: residues 1-269; MeCP2 R294X: residues 1-293; MeCP2 R306C: residues 1-486, R306C; MeCP2 P322L: residues 1-486, P322L; MeCP2 P389X: residues 1-288; MeCP2 Mini: as in the Δ NIC mutant from ref.²². HP1 α : residues 1-191; MED1 IDR: residues 948-1574; BRD4 IDR: residues 674-1351; BRD4 Bromo domain 1: residues 40-168; BRD4 ET domain: residues 600-683; TBLRI-CTD: residues 134-514.

For protein expression, plasmids were transformed into LOBSTR cells (gift from I. M. Cheeseman) and grown as follows. A fresh bacterial colony was inoculated into LB medium containing kanamycin and chloramphenicol and grown overnight at 37 °C. Cells were diluted 1:30 in 500 ml pre-warmed LB with freshly added kanamycin and chloramphenicol and grown 1.5 h at 37 °C. To induce expression, IPTG was added to the bacterial culture at 1 mM final concentration and growth continued for 4 h. Induced bacteria were then pelleted by centrifugation and bacterial pellets were stored at –80 °C until ready to use.

The 500-ml cell pellets were resuspended in 15 ml of lysis buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, and 1 $\times$ cOmplete protease inhibitors) followed by sonication of ten cycles of 15 s on, 60 s off. Lysates were cleared by centrifugation at 12,000g for 30 min at 4 °C, added to 1 ml of pre-equilibrated Ni-NTA agarose, and rotated at 4 °C for 1.5 h. The slurry was centrifuged at 3,000 rpm for 10 min, washed with 10 volumes of lysis buffer and proteins were eluted by incubation for 10 or more minutes rotating with lysis buffer containing 50 mM imidazole, 100 mM imidazole, or 3 $\times$ 250 mM imidazole followed by centrifugation and gel analysis. Fractions containing protein of the correct size were dialysed against two changes of buffer containing 50 mM Tris-HCl

pH 7.5, 125 mM or 500 mM NaCl, 10% glycerol and 1 mM DTT at 4 °C. Protein concentration of purified proteins was determined using the Pierce BCA Protein Assay Kit (Thermo Scientific 23225). Recombinant proteins were stored in 10% glycerol, 50 mM Tris-HCl pH 7.5, 125 mM or 500 mM NaCl, 1 mM DTT. Amicon Ultra Centrifugal filters (30K or 50K MWCO, Millipore) were used to concentrate proteins to desired working concentrations.

In vitro droplet assay

In vitro droplet assays were used to investigate the physicochemical properties of condensate-associated proteins^{33,35}. In vitro droplet assays containing DNA were performed by adding recombinant protein to buffer D (10% glycerol, 50 mM Tris-HCl pH 7.5, 1 mM DTT) containing DNA at the indicated concentration. In vitro droplet assays containing nucleosomal arrays were performed by diluting purified nucleosomes to desired concentration in buffer containing 6 mM MgCl₂, 2% glycerol, 50 mM Tris-HCl pH 7.5 and 1 mM DTT. Recombinant protein was mixed with buffer containing 2% glycerol, 50 mM Tris-HCl pH 7.5 and 1 mM DTT, and then combined with the diluted nucleosomes to initiate droplet formation. In vitro droplet assays containing PEG-8000 were induced by adding recombinant proteins to droplet formation buffer composed of 10% glycerol, 50 mM Tris-HCl pH 7.5, 1 mM DTT and NaCl ranging from 0 mM to 500 mM, with 10% PEG-8000 added. For phase diagram generation (Extended Data Fig. 3g) droplet formation buffer was modified to contain 5% PEG-8000. Droplet assays were performed in 8-tube PCR strip. The indicated protein amount was added to droplet formation buffers and the solution was mixed by pipetting. The reaction was incubated for 10 min at room temperature in the 8-well PCR strip, and then loaded onto either a custom slide chamber created from a glass coverslip mounted on two parallel strips of double-sided tape mounted on a glass microscopy slide, or a well of a glass-bottom 384-well plate (CellVis P384-1.5H-N). Reactions were incubated for 20 min in the imaging vessel to allow droplets in solution to settle on the glass imaging surface. The reaction was then imaged on an Andor Revolution spinning disk confocal microscope using an Andor iXion+ EM-CCD camera with a 100 $\times$ or 150 $\times$ objective using MetaMorph v.7.10.3.279 (Molecular Devices). Images presented are of droplets that have settled on the glass coverslip or the glass bottom of the 384-well plate.

To analyse in-vitro phase separation imaging experiments, custom Python v.3.4.3 scripts (www.github.com/jehenninger/in_vitro_droplet_assay) were used to identify droplets and characterize their size and shape. For any particular experimental condition, intensity thresholds based on the peak of the histogram and size thresholds (2 or 9 pixels per z-slice) were used to segment the image. Droplet identification was performed on the 488 nm channel (MeCP2–GFP) and areas and aspect ratios were determined. Hundreds of droplets, identified in between 5 and 15 independent fields of view from each reaction, were quantified. Exact number of visual fields and droplets used for visualization and quantification are reported in the associated figure legends of relevant panels or in the methods below. To calculate the condensed fraction, the sum total of the intensities in all droplets of a given field (I-in) and the sum total intensity in the bulk dilute phase outside the droplets were calculated for each channel. Condensed fraction was computed as (I-in)/((I-in) + (I-out)). To calculate the partition ratio, the average intensity of each droplet (C-in) and the average intensity of the bulk dilute phase outside the droplet (C-out) was calculated for each channel. The partition coefficient was computed as (C-in)/(C-out). In Figs. 1, 3 and Extended Data Fig. 6, the condensed fraction curves were fitted to the data using a logistic curve³⁶ in Prism v.7.0a (GraphPad).

For in vitro droplet FRAP, droplets were formed as described above. The experiment was performed using the Andor Revolution spinning disk confocal microscope with FRAPPA module (Andor Technology). Bleaching was performed using 1 pulse of 20 μ s dwell time and images were collected every second. Fluorescence intensity was measured using Fiji/ImageJ v.2.0.0-rc-65 and analysed as described above.

Article

To generate a phase diagram for MeCP2, MeCP2–GFP droplets formation reactions were performed in a range of NaCl and protein concentration. Ten independent fields of view were captured for each condition, and droplets were identified as described above. An average partition ratio threshold of >1.85 was used to determine whether a given condition formed droplets.

Droplet numbers

For relevant figure panels, the number of droplets analysed per condition are indicated below. Fig. 1h, Extended Data Fig. 3k: no DNA ($n=592$), DNA ($n=1,395$), methyl-DNA ($n=1,130$). Fig. 2c: WT ($n=1,419$), Δ IDR-1 ($n=1,084$), Δ IDR-2 ($n=208$), IDR-1 ($n=112$), IDR-2 ($n=626$). Fig. 2f: WT ($n=273$), Δ Basic ($n=538$), Δ Aromatic ($n=210$), Δ Histidine ($n=274$), Δ Proline ($n=193$). Extended Data Fig. 3b: MeCP2–GFP 1.25 μ M ($n=1,767$), 2.5 μ M ($n=1,041$), 5 μ M ($n=834$), 10 μ M ($n=483$). Extended Data Fig. 3e: NaCl 100 mM ($n=685$), 200 mM ($n=603$), 300 mM ($n=521$), 400 mM ($n=930$). Extended Data Fig. 3n: WT ($n=106$), Δ IDR-1 ($n=228$), Δ IDR-2 ($n=89$), IDR-1 ($n=51$), IDR-2 ($n=247$). Extended Data Fig. 4c: HP1 α –mCherry ($n=476$), MED1-IDR-mCherry ($n=561$), BRD4-IDR-mCherry ($n=462$), mCherry ($n=413$). Extended Data Fig. 4e: HP1 α –mCherry ($n=1,221$), MED1-IDR-mCherry ($n=1,156$), BRD4-IDR-mCherry ($n=1,124$), mCherry ($n=1,143$). Extended Data Fig. 4g: HP1 α –mCherry ($n=456$), MED1-IDR-mCherry ($n=331$), BRD4-IDR-mCherry ($n=338$), mCherry ($n=402$). Extended Data Fig. 4j: No poly-nucleosomes ($n=599$), poly-nucleosomes ($n=351$). Extended Data Fig. 5b: HP1 α –mCherry ($n=496$), BRD4-IDR-mCherry ($n=484$), BRD4-BD1-mCherry ($n=596$), BRD4-ET-mCherry ($n=451$), mCherry ($n=398$). Extended Data Fig. 6a: WT 0.5 μ M ($n=24$), WT 1 μ M ($n=35$), WT 2 μ M ($n=390$), WT 4 μ M ($n=752$), WT 6 μ M ($n=733$), WT 8 μ M ($n=508$), P389X 0.5 μ M ($n=36$), P389X 1 μ M ($n=49$), P389X 2 μ M ($n=315$), P389X 4 μ M ($n=680$), P389X 6 μ M ($n=578$), P389X 8 μ M ($n=509$), R294X 0.5 μ M ($n=30$), R294X 1 μ M ($n=47$), R294X 2 μ M ($n=14$), R294X 4 μ M ($n=200$), R294X 6 μ M ($n=545$), R294X 8 μ M ($n=516$), R270X 0.5 μ M ($n=58$), R270X 1 μ M ($n=44$), R270X 2 μ M ($n=12$), R270X 4 μ M ($n=158$), R270X 6 μ M ($n=549$), R270X 8 μ M ($n=541$), R255X 0.5 μ M ($n=39$), R255X 1 μ M ($n=53$), R255X 2 μ M ($n=21$), R255X 4 μ M ($n=7$), R255X 6 μ M ($n=1$), R255X 8 μ M ($n=7$), R168X 0.5 μ M ($n=42$), R168X 1 μ M ($n=19$), R168X 2 μ M ($n=3$), R168X 4 μ M ($n=1$), R168X 6 μ M ($n=1$), R168X 8 μ M ($n=1$). Extended Data Fig. 6b: WT 0.5 μ M ($n=346$), WT 1 μ M ($n=1,304$), WT 2 μ M ($n=1,442$), WT 4 μ M ($n=1,117$), WT 6 μ M ($n=1,027$), WT 8 μ M ($n=946$), T158M 0.5 μ M ($n=2,274$), T158M 1 μ M ($n=1,561$), T158M 2 μ M ($n=3,798$), T158M 4 μ M ($n=2,085$), T158M 6 μ M ($n=1,723$), T158M 8 μ M ($n=1,165$), R133C 0.5 μ M ($n=2,577$), R133C 1 μ M ($n=1,465$), R133C 2 μ M ($n=2,305$), R133C 4 μ M ($n=1,937$), R133C 6 μ M ($n=1,380$), R133C 8 μ M ($n=764$). Extended Data Fig. 6c: WT 0.5 μ M ($n=31$), WT 1 μ M ($n=90$), WT 2 μ M ($n=1,237$), WT 4 μ M ($n=672$), WT 6 μ M ($n=536$), WT 8 μ M ($n=537$), R306C 0.5 μ M ($n=23$), R306C 1 μ M ($n=221$), R306C 2 μ M ($n=1,236$), R306C 4 μ M ($n=520$), R306C 6 μ M ($n=507$), R306C 8 μ M ($n=465$). Extended Data Fig. 6f: WT 0.5 μ M ($n=1,580$), WT 1 μ M ($n=1,700$), WT 2 μ M ($n=1,042$), WT 4 μ M ($n=1,202$), WT 6 μ M ($n=1,293$), WT 8 μ M ($n=971$), P322L 0.5 μ M ($n=934$), P322L 1 μ M ($n=1,688$), P322L 2 μ M ($n=2,719$), P322L 4 μ M ($n=4,782$), P322L 6 μ M ($n=1,395$), P322L 8 μ M ($n=2,731$), P225R 0.5 μ M ($n=1,378$), P225R 1 μ M ($n=2,061$), P225R 2 μ M ($n=1,632$), P225R 4 μ M ($n=4,510$), P225R 6 μ M ($n=2,876$), P225R 8 μ M ($n=3,015$). Extended Data Fig. 7b: MeCP2–GFP WT ($n=719$), MeCP2–GFP R306C ($n=707$). Extended Data Fig. 7d: MeCP2–GFP WT ($n=1,103$), MeCP2–GFP R306C ($n=535$). Extended Data Fig. 8c: MeCP2–GFP WT ($n=459$), MeCP2–GFP Mini ($n=363$). Extended Data Fig. 8f, g: MeCP2–GFP WT ($n=288$), MeCP2–GFP Mini ($n=341$). Extended Data Fig. 8i: MeCP2–GFP WT ($n=1,109$), MeCP2–GFP Mini ($n=910$).

Fluorescent DNA production

Fluorescent DNA for droplet assays was produced by amplifying plasmid DNA using oligonucleotide primers with 5'-Cy5 fluorophore modifications (Integrated DNA Technologies). Amplification of plasmid

templates was performed using Phusion polymerase (Thermo Scientific F531S). Fluorescent PCR products were gel purified using the Monarch gel extraction kit (NEB T1020S). The 376-bp DNA sequence used in droplet assays is: TGTAACACGACGGCCAGTGGATCCTAGGCTTAATTTGCATTGTCAGTACATTTGCATGCATGATATTGCATTAAAGCTTGATTGTCATGTTTCAGAATTTGCATCGGCTAGCATTTGCATGGGCTAGAAATTTGCATGCGGATAATTTGCATGGCGATTTCATTTGCATGCCAAATCATTTGCATGCA TGAACATTTGCATGGCTTACAATTTGCATGAAACATAATTTGCATCGATCGAAATTTGCATGTAGCCGAATTTGCATGTAGCTAAATTTGCATGAAATCGGATTTGCATGTAGCAATATTTGCATCTAGCCTAATTTGCATACCCTAGCATTTGCATTAGATTTCGGCGGCCGCGTCATAGCTGTTTCCTG.

To generate methylated DNA template for in vitro droplet assays, Cy5-labelled fluorescent PCR product produced as described above was treated with M.SssI methyltransferase (NEB M0226L). The reaction was performed in 50 μ l and contained 160 μ M S-adenosylmethionine (SAM), 1 μ l of M.SssI, and 4 μ g of DNA. Contents were incubated for 4 h at 37 °C, and then M.SssI was heat-inactivated for 20 min at 65 °C. Resulting methylated templates were purified using the NEB Monarch DNA and PCR cleanup kit (NEB T1030S). Methylation of templates was verified by methyl-specific restriction digestion using ClaI (NEB R0197S).

Poly-nucleosome purification

Poly-nucleosome arrays were purified from mouse ES cells based on a previous protocol³⁷. In brief, nuclei were isolated from mouse ES cells by resuspending cells in a hypotonic buffer BC50 (HEPES pH 7.5, 50 mM NaCl) + 5 mM MgCl₂ + 0.05% NP-40 and douncing with a Kontes glass dounce (15 strokes with each pestle A then B). The nuclei were then digested with a limited amount of micrococcal nuclease and then the samples were centrifuged at maximum speed for 10 min. To purify poly-nucleosome arrays, the supernatant was loaded on a sucrose gradient and centrifuged for 20 h in a swinging bucket rotor (Sorvall SW28) at 18,000 rpm. The sucrose gradients (28 ml each) were 5–15% in a base buffer of HEPES pH 7.5 and 200 mM NaCl. Individual fractions corresponding to poly-nucleosome arrays were collected. To determine the length distribution of the poly-nucleosome arrays in each fraction, DNA was purified from each fraction and analysed on an agarose gel. Fractions containing nucleosomal arrays ranging between 7 and 20 nucleosomes in length were pooled and dialysed against buffer BC50 + 5 mM MgCl₂. Purified poly-nucleosomes were stored in liquid nitrogen until ready to use in droplet assays.

MeCP2 IDR-2 sequence features

Specific sequence features within protein IDRs have been found to contribute to condensate formation^{36,38–43}. Sequence features within MeCP2 IDR-2 were identified and deletion mutants were used to examine for their ability to contribute to droplet formation in vitro and transcriptional repression in a reporter assay. Basic patches in IDR-2 were defined as previously described⁴². In brief, net charge per residue (NCPR) along the MeCP2 protein sequence was computed using a sliding window of 5 residues and a step size of 1 residue using localCIDER (v.0.1.14)⁴⁴. Stretches of 4 or more consecutive windows having a NCPR > +0.35 per window were considered to be basic patches. MeCP2 IDR-2 contained seven basic patches corresponding to residues 170–181, 184–194, 246–258, 263–274, 282–289, 301–310 and 340–348. Two aromatic residues (residues F226 and Y450) are present in IDR-2. A histidine-rich domain (residues 366–372) and a proline-rich domain (residues 376–405) in IDR-2 were defined based on UniProt annotations.

Transcriptional repression reporter assay

A transcriptional repression reporter assay was used to examine the ability of MeCP2 IDR-2 sequence feature deletion mutants to repress transcription. Plasmids expressing MeCP2 IDR-2 sequence feature deletion mutants as fusions with the GAL4-DBD from a SV40 promoter were co-transfected with a transcriptional repression reporter plasmid, containing an array of five GAL4 DNA binding sequence motifs located

upstream of a chicken β -actin promoter driven Firefly luciferase gene. To control for transfection efficiency, a plasmid expressing *Renilla* luciferase under the control of the SV40 promoter was also co-transfected. HEK293T cells were transfected using Lipofectamine 3000 (Invitrogen L3000), 24 h after plating in a 96-well white flat bottom plate (Corning 3917). Twenty-four hours h after transfection, expression of the transcriptional repression reporter Firefly luciferase and control *Renilla* luciferase were assayed using the Dual-Glo Luciferase Assay System (Promega E2940) and measured using a plate reader. Luciferase activity was calculated for each condition by dividing the Firefly luciferase signal by the *Renilla* luciferase signal, and was normalized to the GAL4-DBD alone condition. Assay was performed with three biologically independent samples per condition.

Gene expression analysis

Quantitative PCR with reverse transcription (RT-qPCR) was used to quantify expression of heterochromatin-associated major satellite repeats. RNA was collected using the RNeasy Mini Plus kit (QIAGEN 74134). A reverse transcriptase reaction was then performed using SuperScript III (Invitrogen 18080). RT-qPCR reactions were performed using Power SYBR Green PCR Master Mix (Applied Biosystems 43676) and measured using a QuantStudio 5 Real-Time PCR System (Applied Biosystems). Major satellite expression level was calculated using the $\Delta\Delta C_t$ method using *Gapdh* as a control and normalized to expression level in the wild-type condition. The following primers were used: MajorSat_for: TGG AATATGCGAGAAACTG; MajorSat_rev: AGGTCC TTCAGTGGGCATTT; *Gapdh*_for: AACTTTGGCATTGTGGAAGG; *Gapdh*_rev: CACATTGGGGGTAGGAACAC.

RNA-seq was used to profile expression of genes. RNA was collected from 2 million cell aliquots using the RNeasy Mini Plus kit (QIAGEN 74134). Amount of RNA extracted was quantified using a Nanodrop spectrophotometer (Thermo Scientific). A fixed amount of ERCC RNA Spike-In (Invitrogen 4456740) was added to each sample for use in cell number normalization⁴⁵. Samples were treated with DNA-free DNA Removal Kit (Invitrogen AM1906) before library preparation using the KAPA RNA HyperPrep Kit with RiboErase (KAPA Biosystems K8562) and sequencing on a HiSeq2500 (Illumina).

RNA-seq reads were mapped using STAR aligner (v.2.6.1a)⁴⁶ to the murine RefSeq mm9 reference with ERCC spike-in reference sequences added. Alignment was performed using ENCODE long RNA-seq pipeline default parameters: `-outFilterType BySJout`, `-outSAMattributes NH HI AS NM MD`, `-outFilterMultimapNmax 20`, `-outFilterMismatchNmax 999`, `-outFilterMismatchNoverReadLmax 0.05`, `-alignIntronMin 20`, `-alignIntronMax 1000000`, `-alignMatesGapMax 1000000`, `-alignSJoverhangMin 8`, `-alignSJDBoverhangMin 1`, `-sjdbScore 1`. Gene expression values were quantified using RSEM (v.1.2.31) with default parameters⁴⁷. Differential expression analysis was performed using DESeq2 (v.1.24.0) with default parameters⁴⁸. Spike-in cell number normalization was performed by using ERCC spike-ins to estimate size factors used for DESeq2 library normalization. DESeq2 uses a two-tailed Wald test to identify differentially expressed genes, and the default multiple test adjusted *P* value cutoff of 0.1 was used to determine differentially expressed genes.

Bioinformatic analysis

MECP2 gene expression values in transcripts per million (TPM) from RNA-seq of human tissues were acquired from the Genotype-Tissue Expression (GTEx) project release v.7. In instances where multiple regions of the same tissue were assayed, the highest expression value was used to represent the tissue. TPM values greater than 1 were considered to be expressed. The GTEx Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH and NINDS.

Mutation data from patients with Rett syndrome were acquired from RettBASE⁴⁹. Coding mutations associated with female patients with Rett syndrome were used for analysis. A histogram of mutation count

along the length of the protein was plotted and the type of mutation (nonsense, frameshift, missense) was indicated.

Predicted disorder values along the length of human MeCP2 protein was determined using PONDR VLS2 algorithm. Higher values indicate greater predicted disorder.

Protein sequence conservation along the length of human MeCP2 protein was determined by extracting protein sequences in UniProt that shared at least 50% identity with human MeCP2 sequence (UniRef50_Q9Z2D6). Extracted sequences were subject to multiple sequence alignment using Clustal Omega (v.1.2.4). Alignments were scored for protein sequence conservation along the length of human MeCP2 using Jensen-Shannon divergence⁵⁰. Higher values indicated greater conservation.

Statistics and reproducibility

Relevant statistical information for each experiment are included in the associated figure legends. For *t*-tests, data were assumed to be normal. For RNA-seq analysis, a two-sided Wald test was used to identify differentially expressed genes and *P* values were adjusted for multiple comparison.

Experiments with representative images conducted in this study were repeated multiple times independently with similar results. Live-cell imaging of endogenously tagged mouse ES cells and neurons (Figs. 1a, 4a, e, Extended Data Figs. 1a, b, 9a, g) was performed at least three times, on different days with cells plated independently. Live-cell imaging of endogenously tagged MeCP2-Mini mouse ES cells (Extended Data Fig. 8j) was performed twice. Live-cell imaging of overexpressed MeCP2-GFP in mouse ES cells (Extended Data Fig. 10a) was performed at least three times. Immunofluorescence in mouse brain cells (Fig. 1d) was performed on two brain sections. Immunofluorescence in mouse ES cells (Extended Data Fig. 4a) was performed twice. Immunofluorescence in differentiated neurons (Extended Data Fig. 10e) was performed twice. FRAP experiments in mouse ES cells (Fig. 1b, Extended Data Fig. 1c) were performed independently at least twice. FRAP experiments on mouse brain sections (Fig. 1e) was performed on three brain slices. FRAP experiments on MeCP2 droplets (Extended Data Fig. 3i) were performed on 10 individual droplets. Droplet fusion (Extended Data Fig. 3h) was observed more than 10 times. PCR genotyping of MeCP2-GFP tagged ES cells (Extended Data Fig. 2c) was performed twice. Droplet experiments in Figs. 1g, 2b, e, 3b, d, f, Extended Data Figs. 3a, d, j, m, 4b, d, f, h, i, 5a, 6d, 7a, c, 8b, e, h were performed at least two times independently. Western blots (Extended Data Figs. 9d, 10c, f) were performed twice.

Reporting summary

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

Data availability

Relevant data supporting the findings of this study are available within the paper and its Supplementary Information. RNA-seq datasets generated in this study have been deposited in the Gene Expression Omnibus under accession code GSE139033. Uncropped gel images can be found in Supplementary Fig. 1. Additional data are available from the corresponding author upon reasonable request. The following publicly available data were used in this study: GTEx v. 7 RNA-seq Median Gene TPMs by Tissue (www.gtexportal.org), RettBASE *MECP2* Variant List (mecp2.chw.edu.au/mecp2/mecp2_home.php), and UniProt Cluster ID: UniRef50_Q9Z2D6 (www.uniprot.org/uniref/UniRef50_Q9Z2D6). Source data are provided with this paper.

Code availability

Custom code used for analysis of images from in vitro droplet assays is available at www.github.com/jehenninger/in_vitro_droplet_assay.

Custom code used for analysis of heterochromatin condensate volumes is available at www.github.com/jehenninger/MECP2_neuron.

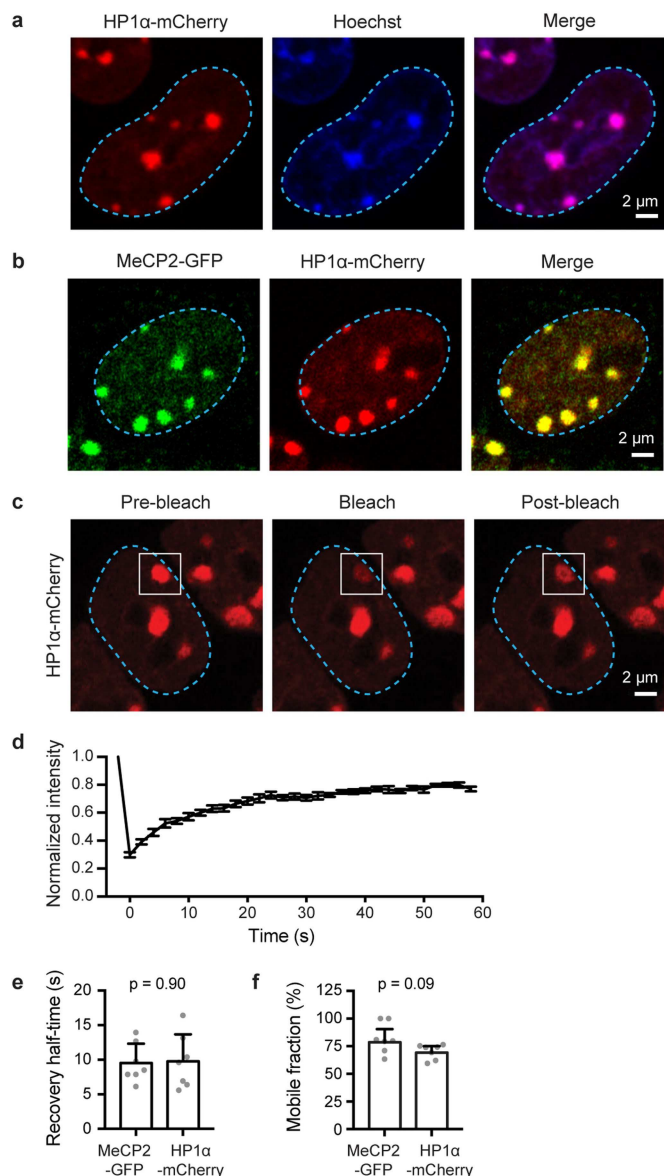
31. Thoma, E. C. et al. Ectopic expression of neurogenin 2 alone is sufficient to induce differentiation of embryonic stem cells into mature neurons. *PLoS ONE* **7**, e38651 (2012).
32. Zhang, Y. et al. Rapid single-step induction of functional neurons from human pluripotent stem cells. *Neuron* **78**, 785–798 (2013).
33. Alberti, S., Gladfelter, A. & Mittag, T. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 (2019).
34. Jang, S. et al. Long-term culture of organotypic hippocampal slice from old 3xTg-AD mouse: an ex vivo model of Alzheimer's disease. *Psychiatry Investig.* **15**, 205–213 (2018).
35. Alberti, S. et al. A user's guide for phase separation assays with purified proteins. *J. Mol. Biol.* **430**, 4806–4820 (2018).
36. Wang, J. et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699.e16 (2018).
37. Loyola, A., Bonaldi, T., Roche, D., Imhof, A. & Almouzni, G. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol. Cell* **24**, 309–316 (2006).
38. Li, P. et al. Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).
39. Martin, E. W. et al. Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).
40. Lu, H. et al. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**, 318–323 (2018).
41. Zamudio, A. V. et al. Mediator condensates localize signaling factors to key cell identity genes. *Mol. Cell* **76**, 753–766.e6 (2019).
42. Pak, C. W. et al. Sequence determinants of intracellular phase separation by complex coacervation of a disordered protein. *Mol. Cell* **63**, 72–85 (2016).
43. Dao, T. P. et al. ALS-linked mutations affect UBQLN2 oligomerization and phase separation in a position- and amino acid-dependent manner. *Structure* **27**, 937–951.e5 (2019).
44. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G. & Pappu, R. V. CIDR: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys. J.* **112**, 16–21 (2017).
45. Lovén, J. et al. Revisiting global gene expression analysis. *Cell* **151**, 476–482 (2012).
46. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
47. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
48. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
49. Christodoulou, J., Grimm, A., Maher, T. & Bennetts, B. RettBASE: The IRSA MECP2 variation database—a new mutation database in evolution. *Hum. Mutat.* **21**, 466–472 (2003).
50. Capra, J. A. & Singh, M. Predicting functionally important residues from sequence conservation. *Bioinformatics* **23**, 1875–1882 (2007).

Acknowledgements We thank A. P. Bird for sharing *Mecp2* mutant cell lines; D. Reinberg for sharing purified poly-nucleosomes; P. A. Sharp for discussions; D. Richardson and the Harvard Center for Biological Imaging; W. Salmon and the Whitehead W.M. Keck Microscopy Facility; and R. Flannery, J. Drotar, N. Rosenau and the Whitehead Genetically Engineered Models Center for technical support. The work was supported by NIH grant R01 GM123511 (R.A.Y.), NSF grant PHY1743900 (R.A.Y.), NIH grant 2 R01 MH104610-20 (R.J., R.A.Y.), NIH grant R37 CA084198 (R.J.), Hope Funds for Cancer Research Fellowship (A.D.), NIH grant T32 5T32DK007191-45 (J.M.P.), NSF Graduate Research Fellowship (A.V.Z.), and NIH grant K99/RO0 MH113813 (X.S.L.).

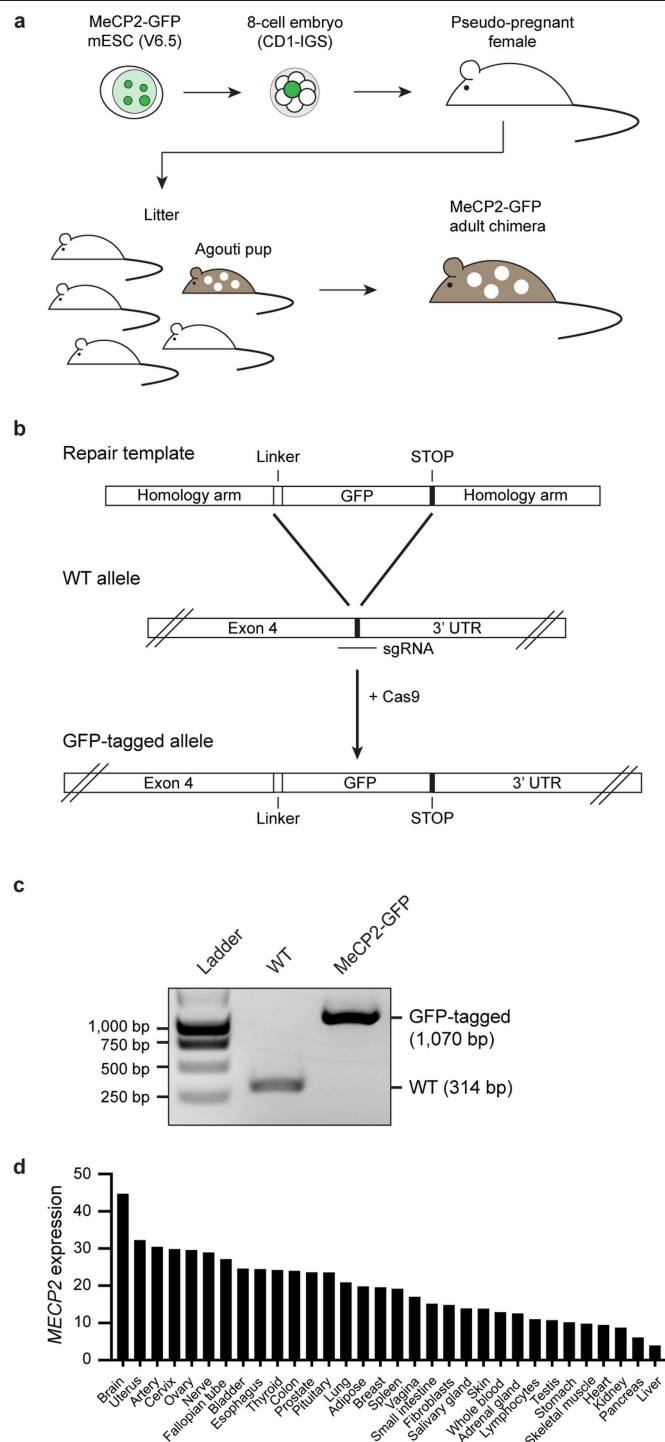
Author contributions C.H.L., E.L.C., R.J. and R.A.Y. conceived the project. C.H.L., E.L.C., T.I.L., R.J. and R.A.Y. organized the studies. C.H.L., E.L.C. and R.A.Y. wrote the manuscript. E.L.C., J.E.H., O.O., A.V.Z. and J.S. performed in vitro droplet assays. N.M.H. performed protein purification. G.L. performed purification of poly-nucleosomes. C.H.L., E.L.C. and J.E.H. developed and performed computational analyses. E.L.C. and A.D. performed cellular imaging experiments. L.K.A. performed transcriptional reporter assays. C.H.L. generated endogenous-tagged cell lines and gene expression analyses. X.T. and T.L. performed neuronal differentiation. A.D., X.S.L., S.M., D.S.S. and E.W. performed chimeric mouse generation. C.H.L., E.L.C. and J.M.P. generated constructs. R.J. and R.A.Y. supervised the project with help from T.I.L. All authors contributed to editing the manuscript.

Competing interests R.A.Y. is a founder and shareholder of Syros Pharmaceuticals, Camp4 Therapeutics, Omega Therapeutics, and Dewpoint Therapeutics. R.J. is an advisor/co-founder of Fate Therapeutics, Fulcrum Therapeutics, Omega Therapeutics, and Dewpoint Therapeutics. T.I.L. is a shareholder of Syros Pharmaceuticals and a consultant to Camp4 Therapeutics. All other authors declare no competing interests.

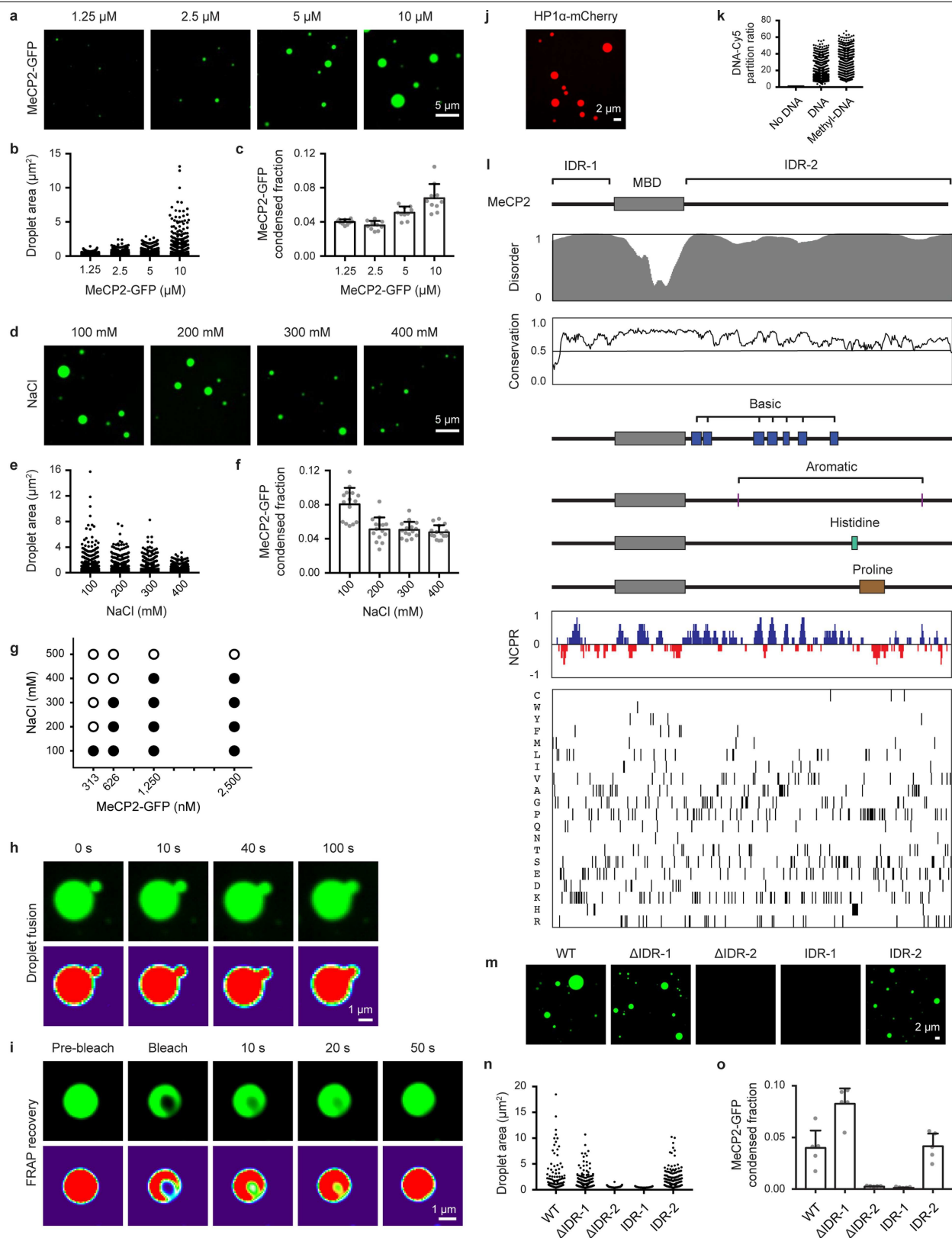
Additional information
Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-020-2574-4>.
Correspondence and requests for materials should be addressed to R.J. or R.A.Y.
Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | MeCP2 and HP1α are dynamic components of heterochromatin condensates. **a**, Live-cell images of endogenous-tagged HP1α-mCherry and Hoechst staining in mouse ES cells. **b**, Live-cell images of endogenous-tagged MeCP2-GFP and HP1α-mCherry in mouse ES cells. **c**, Live-cell images of FRAP experiments with endogenously tagged HP1α-mCherry mouse ES cells. **d**, FRAP curves for experiments in **c**. Photobleaching occurs at $t = 0$ s. $n = 7$ cells. **e**, Half-time of photobleaching recovery for MeCP2-GFP and HP1α-mCherry at heterochromatin condensates in imaging experiments in **c** and Fig. 1b. $n = 7$ cells per condition. $P = 0.90$, $t = 0.13$, $df = 12$, two-tailed Student's t -test. **f**, Mobile fractions of MeCP2-GFP and HP1α-mCherry within heterochromatin condensates in imaging experiments in **c** and Fig. 1b, determined by FRAP analysis. $n = 7$ cells per condition. $P = 0.09$, $t = 1.87$, $df = 12$, two-tailed Student's t -test. Data are mean $\pm$ s.e.m.



Extended Data Fig. 2 | Generation of endogenous-tagged MeCP2-GFP chimeric mice. a. Schematic of MeCP2-GFP chimeric mouse generation using endogenous-tagged MeCP2-GFP mouse ES cells. Endogenous-tagged MeCP2-GFP mouse ES cells derived from V6.5 background were injected into embryos from CD1-IGS mice and multiple embryos were implanted into pseudo-pregnant female mice. Chimeric pups were distinguished from non-chimeric littermates by agouti coat colour. MeCP2-GFP tagged adult chimeric mice were used for experiments. **b.** Schematic of strategy used to generate endogenous-tagged MeCP2-GFP mouse ES cells. **c.** PCR genotyping of endogenous-tagged MeCP2-GFP mouse ES cells. For gel source data, see Supplementary Fig. 1. **d.** *MECP2* expression values in transcripts per million (TPM) measured by RNA-seq for various human tissues surveyed by GTEx. Tissues are ordered based on expression level. TPM values greater than 1 are considered to be expressed.

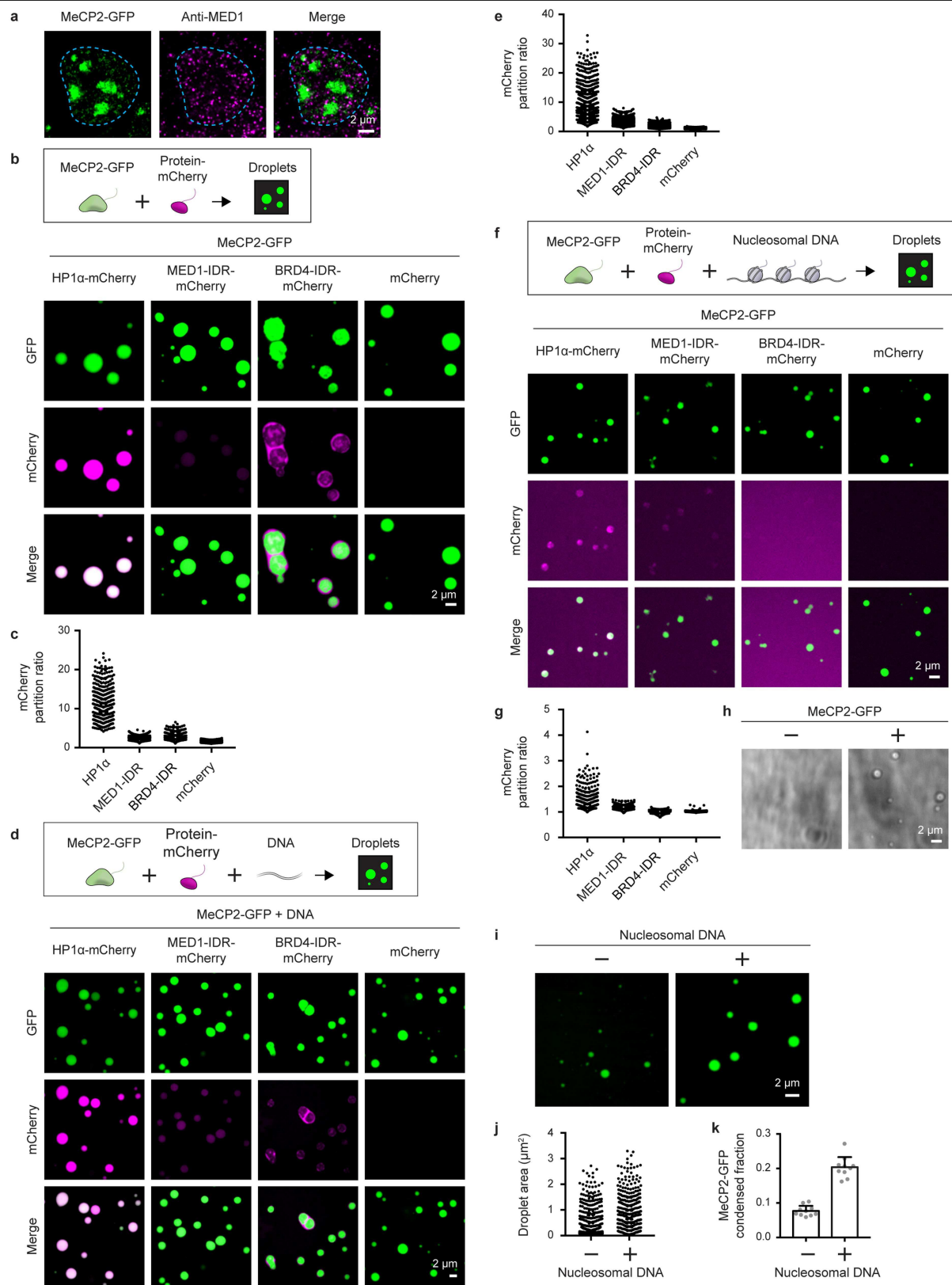


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | MeCP2 forms phase-separated droplets in vitro.

a, Droplet experiments examining effect of MeCP2 concentration. MeCP2–GFP was added to droplet formation buffers with 100 mM NaCl and 10% PEG-8000. **b**, Droplet areas for experiments in **a**. Fields per condition $n = 10$. **c**, MeCP2–GFP condensed fraction for experiments in **a**. Fields per condition $n = 10$. **d**, Droplet experiments examining effect of salt concentration. MeCP2–GFP at 10 μ M was added to droplet formation buffers with indicated NaCl concentrations and 10% PEG-8000. **e**, Droplet areas for experiments in **d**. Fields per condition $n = 15$. **f**, MeCP2–GFP condensed fraction for experiments in **d**. Fields per condition $n = 15$. **g**, Phase diagram of MeCP2 droplet formation. MeCP2–GFP was added to droplet formation buffers with indicated NaCl concentrations and 5% PEG-8000. Filled-in circles indicate conditions with droplets. Fields per condition $n = 10$. **h**, Droplet experiments showing MeCP2 droplet fusion.

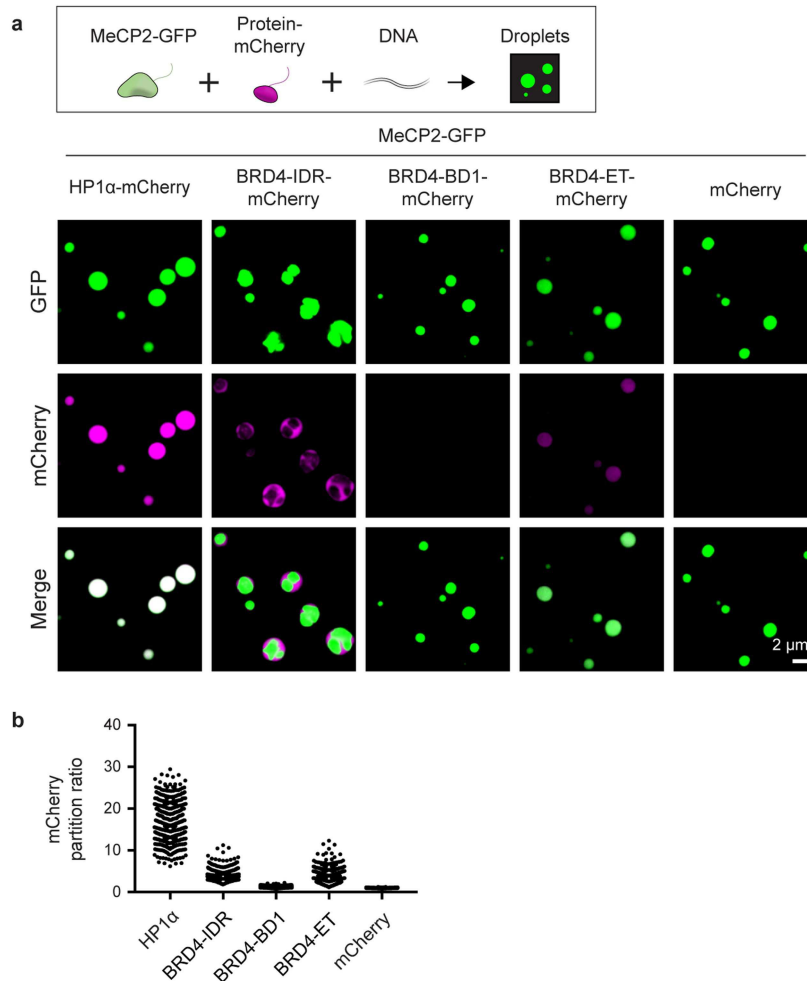
MeCP2–GFP at 10 μ M was added to droplet formation buffers with 100 mM NaCl and 10% PEG-8000. **i**, Droplet experiments showing MeCP2 droplet FRAP. Conditions as in **h**. Photobleaching at $t = 0$ s. **j**, Droplet experiments examining HP1 α . HP1 α –mCherry at 10 μ M was added to droplet formation buffers with 100 mM NaCl and 10% PEG-8000. **k**, DNA–Cy5 partition ratios in MeCP2–GFP droplets for experiments in Fig. 1g. Fields per condition $n = 15$. **l**, Expanded schematic of MeCP2 protein (Fig. 2a) with protein sequence conservation, net charge per residue (NCPR), and residue plots. **m**, Droplet experiments examining MeCP2 deletion mutants. MeCP2–GFP deletion mutants at 10 μ M were added to droplet formation buffers with 100 mM NaCl and 10% PEG-8000. **n**, Droplet areas for experiments in **m**. Fields per condition $n = 5$. **o**, MeCP2–GFP condensed fraction for experiments in **m**. Fields per condition $n = 5$. Data are mean $\pm$ s.d.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | MeCP2 condensates preferentially concentrate HP1α compared to components of transcriptional condensates. **a**, Immunofluorescence images of heterochromatin condensates (MeCP2-GFP) and transcriptional condensates (anti-MED1) in mouse ES cells. **b**, Droplet experiments examining ability of MeCP2 condensates to preferentially concentrate HP1α compared to components of transcriptional condensates. MeCP2-GFP at 7.5 μM was mixed with HP1α-mCherry, MED1-IDR-mCherry, BRD4-IDR-mCherry, or mCherry at 7.5 μM in droplet formation buffers with 150 mM NaCl and 10% PEG-8000. **c**, mCherry partition ratios in MeCP2-GFP droplets for experiments in **b**. Fields per condition $n = 15$. **d**, Droplet experiments with naked DNA examining ability of MeCP2 condensates to preferentially concentrate HP1α compared to components of transcriptional condensates. Conditions same as in **b**, but with the addition of 160 nM DNA. **e**, mCherry partition ratios in MeCP2-GFP droplets for experiments in **d**. Fields per condition $n = 15$. **f**, Droplet experiments

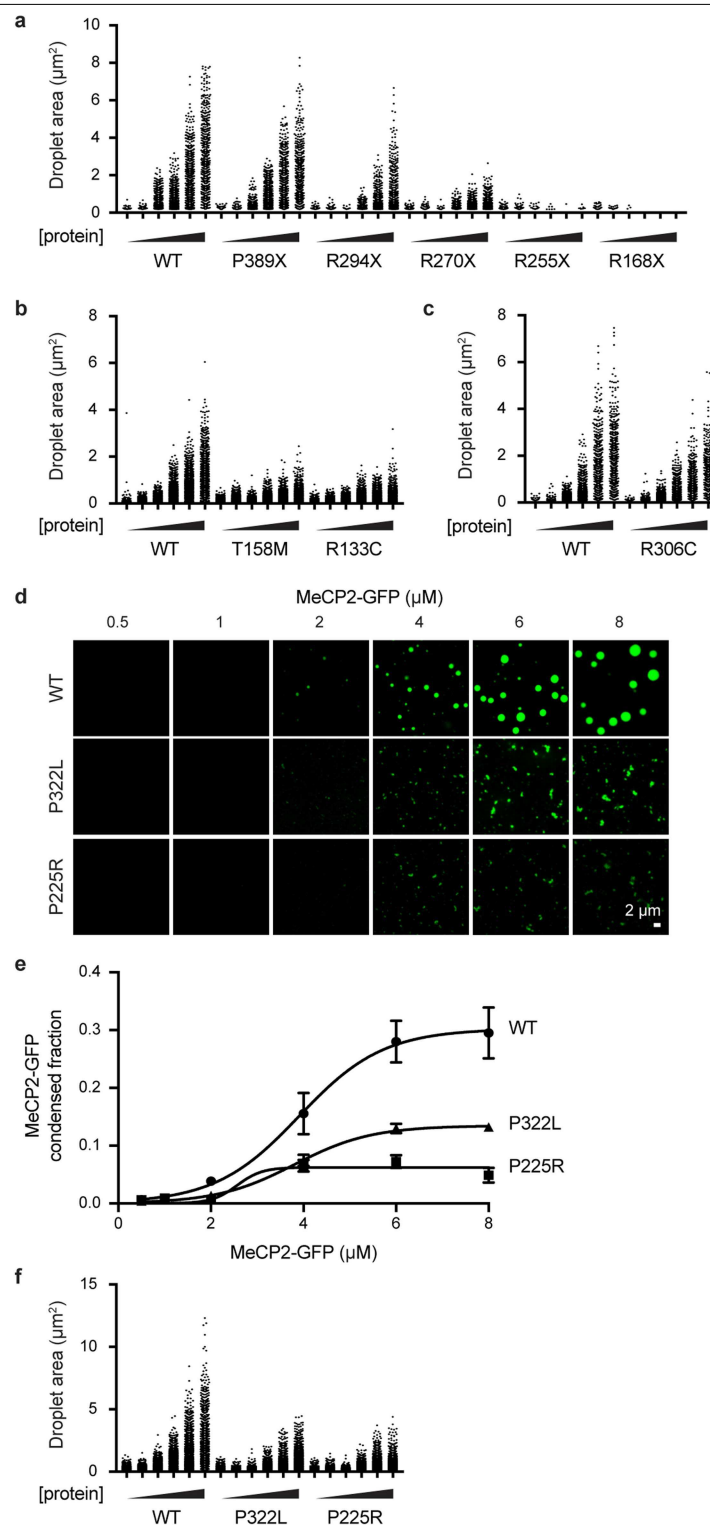
with nucleosomal DNA examining ability of MeCP2 condensates to preferentially concentrate HP1α compared to components of transcriptional condensates. MeCP2-GFP at 5 μM was mixed with HP1α-mCherry, MED1-IDR-mCherry, BRD4-IDR-mCherry, or mCherry at 5 μM and 6 nM poly-nucleosomes in droplet formation buffers with 100 mM NaCl and 3 mM MgCl₂. **g**, mCherry partition ratios in MeCP2-GFP droplets for experiments in **f**. Fields per condition $n = 10$. **h**, Brightfield images examining droplet formation with nucleosomal DNA alone and with MeCP2. Poly-nucleosomes at 6 nM were mixed with 5 μM MeCP2-GFP or no MeCP2-GFP in droplet formation buffers with 100 mM NaCl and 3 mM MgCl₂. **i**, Droplet experiments examining MeCP2 droplet formation with nucleosomal DNA. Conditions same as in **h**. **j**, Droplet areas for experiments in **i**. Fields per condition $n = 10$. **k**, MeCP2-GFP condensed fraction for experiments in **i**. Fields per condition $n = 10$. Data are mean $\pm$ s.d.



Extended Data Fig. 5 | MeCP2 condensate partitioning of BRD4 domains.

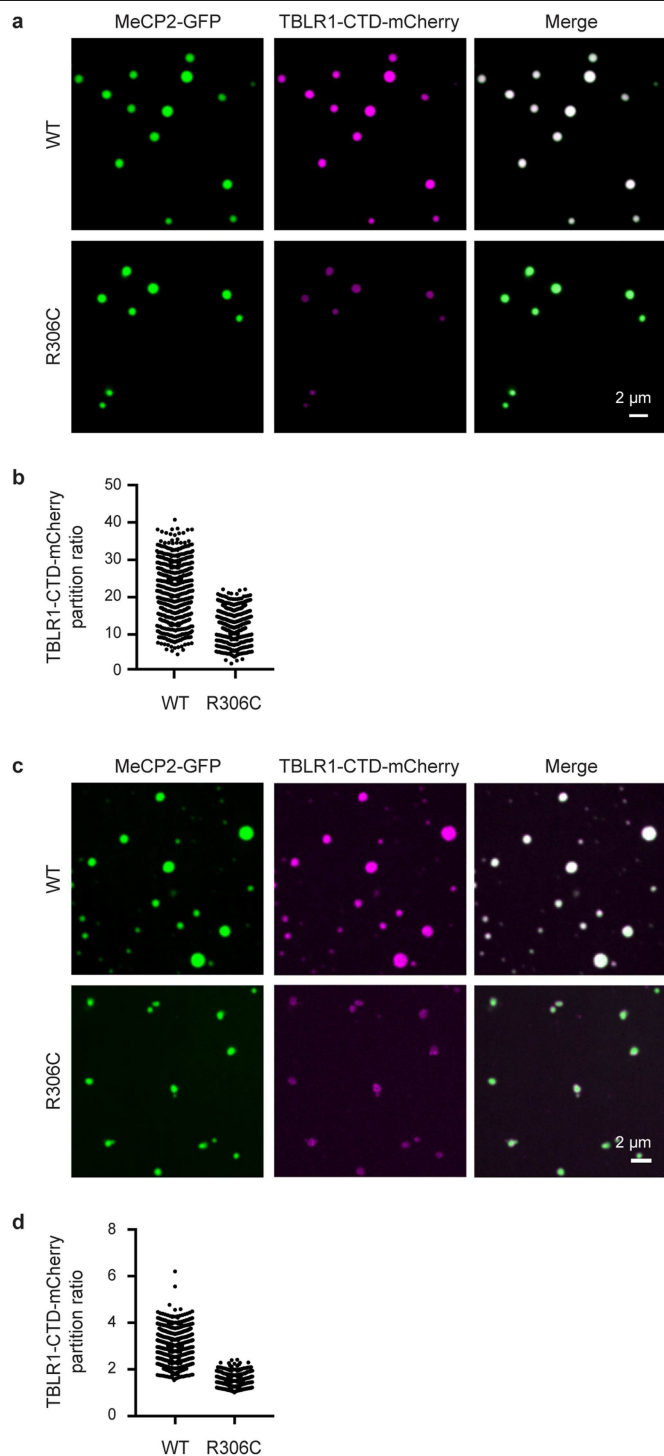
a, Droplet experiments examining ability of MeCP2 condensates to preferentially incorporate and concentrate HP1α compared to BRD4 domains in the presence of naked DNA. BRD4-IDR, bromodomain 1 (BD1), and extra-terminal (ET) domain were examined. MeCP2-GFP at 7.5 μM was mixed

with either HP1α-mCherry, BRD4-IDR-mCherry, BRD4-BD1-mCherry, BRD4-ET-mCherry, or mCherry each at 7.5 μM and 160 nM methylated DNA in droplet formation buffers with 150 mM NaCl and 10% PEG-8000. **b**, mCherry partition ratios in MeCP2-GFP droplets for experiments in **a**. Fields per condition $n = 15$.

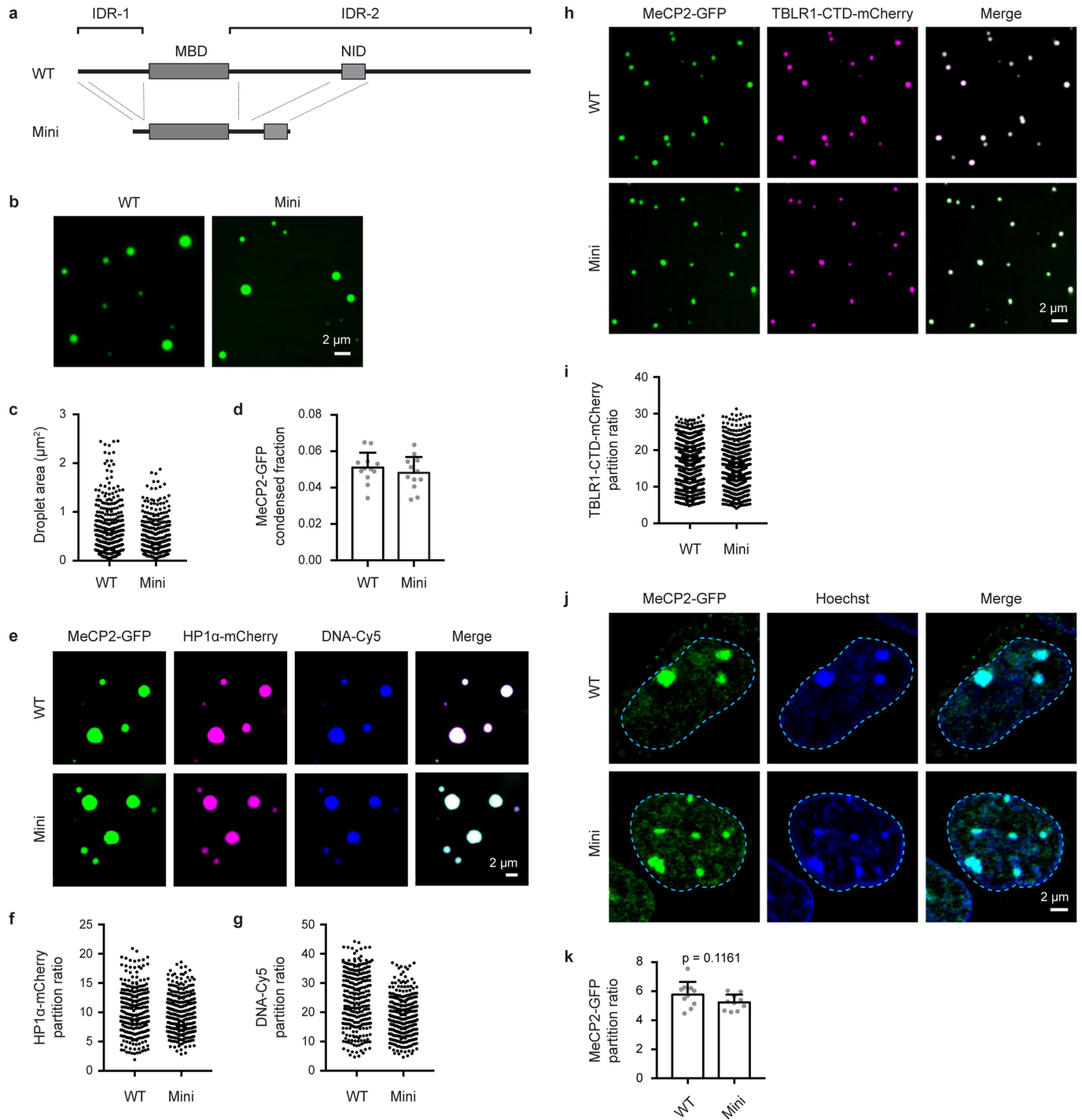


Extended Data Fig. 6 | Rett syndrome mutations disrupt MeCP2 condensate formation. **a**, Droplet areas for experiments in Fig. 3b. Fields per condition $n = 15$. **b**, Droplet areas for experiments in Fig. 3d. Fields per condition $n = 15$. **c**, Droplet areas for experiments in Fig. 3f. Fields per condition $n = 15$. **d**, Droplet experiments examining effects of Rett syndrome missense mutations that disrupt IDR-2 on MeCP2 droplet formation. Wild-type MeCP2-GFP and Rett

syndrome IDR-2 mutants (P225R and P322L) at indicated concentrations were mixed with 20 nM methylated DNA in droplet formation buffers with 100 mM NaCl. **e**, MeCP2-GFP condensed fraction as a function of MeCP2-GFP concentration for experiments in **d**. Data are mean $\pm$ s.d. Fields per condition $n = 15$. **f**, Droplet areas for experiments in **d**. Fields per condition $n = 15$.

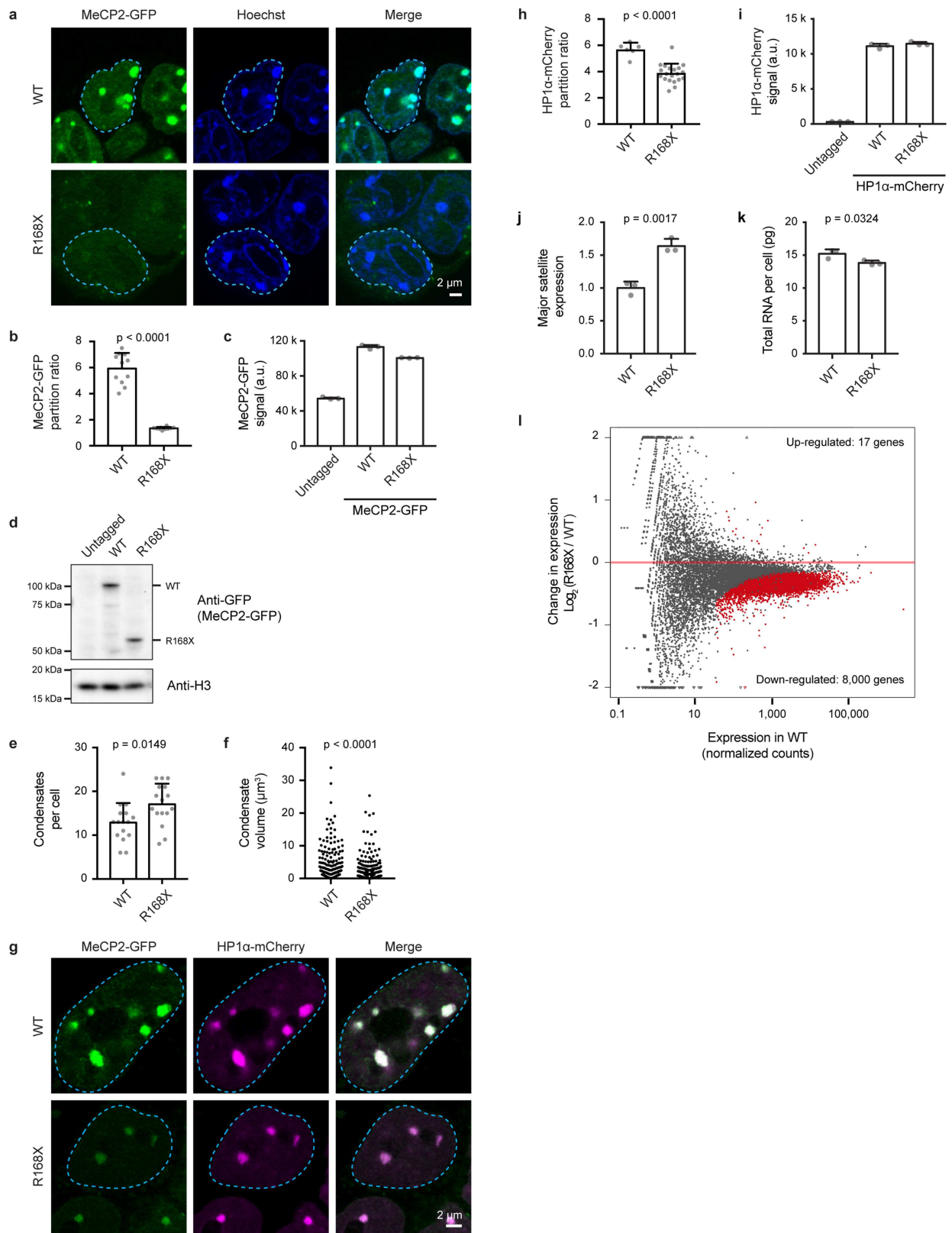


Extended Data Fig. 7 | TBLR1 partitioning into MeCP2 droplets is disrupted by Rett syndrome mutation R306C. **a**, Droplet experiments examining ability of wild-type MeCP2 and R306C mutant condensates to enrich TBLR1-CTD. Wild-type MeCP2-GFP or R306C mutant at 6 μ M was mixed with TBLR1-CTD-mCherry at 10 μ M in droplet formation buffers with 125 mM NaCl and 10% PEG-8000. **b**, TBLR1-CTD-mCherry partition ratios in MeCP2-GFP wild-type and R306C mutant droplets for experiments in **a**. Fields per condition $n = 15$. **c**, Droplet experiments examining ability of wild-type MeCP2 and R306C mutant condensates to enrich TBLR1-CTD. Wild-type MeCP2-GFP or R306C mutant at 10 μ M was mixed with TBLR1-CTD-mCherry at 4 μ M in droplet formation buffers with 125 mM NaCl. **d**, TBLR1-CTD-mCherry partition ratios in wild-type MeCP2-GFP and R306C mutant droplets for experiments in **c**. Fields per condition $n = 12$.



Extended Data Fig. 8 | MeCP2 Mini forms droplets in vitro and partitions into heterochromatin condensates in mouse ES cells. **a**, Schematic of MeCP2 protein with a minimal MeCP2 protein (Mini)²² that retains the MBD and NID displayed below. **b**, Droplet experiments examining ability of Mini MeCP2 to form droplets. Wild-type MeCP2-GFP and Mini at 4 μ M were added to droplet formation buffers with 125 mM NaCl and 10% PEG-8000. **c**, Droplet areas for experiments in **b**. Fields per condition $n = 12$. **d**, MeCP2-GFP condensed fraction for experiments in **b**. Fields per condition $n = 12$. **e**, Droplet experiments examining ability of wild-type MeCP2 and Mini to form droplets with HP1 α and DNA. Wild-type MeCP2-GFP or Mini at 7.5 μ M was mixed with 7.5 μ M HP1 α -mCherry and 160 nM DNA in droplet formation buffers with 150 mM NaCl and 10% PEG-8000. **f**, HP1 α -mCherry partition ratios in wild-type MeCP2-GFP and Mini droplets for experiments in **e**. Fields per condition $n = 15$.

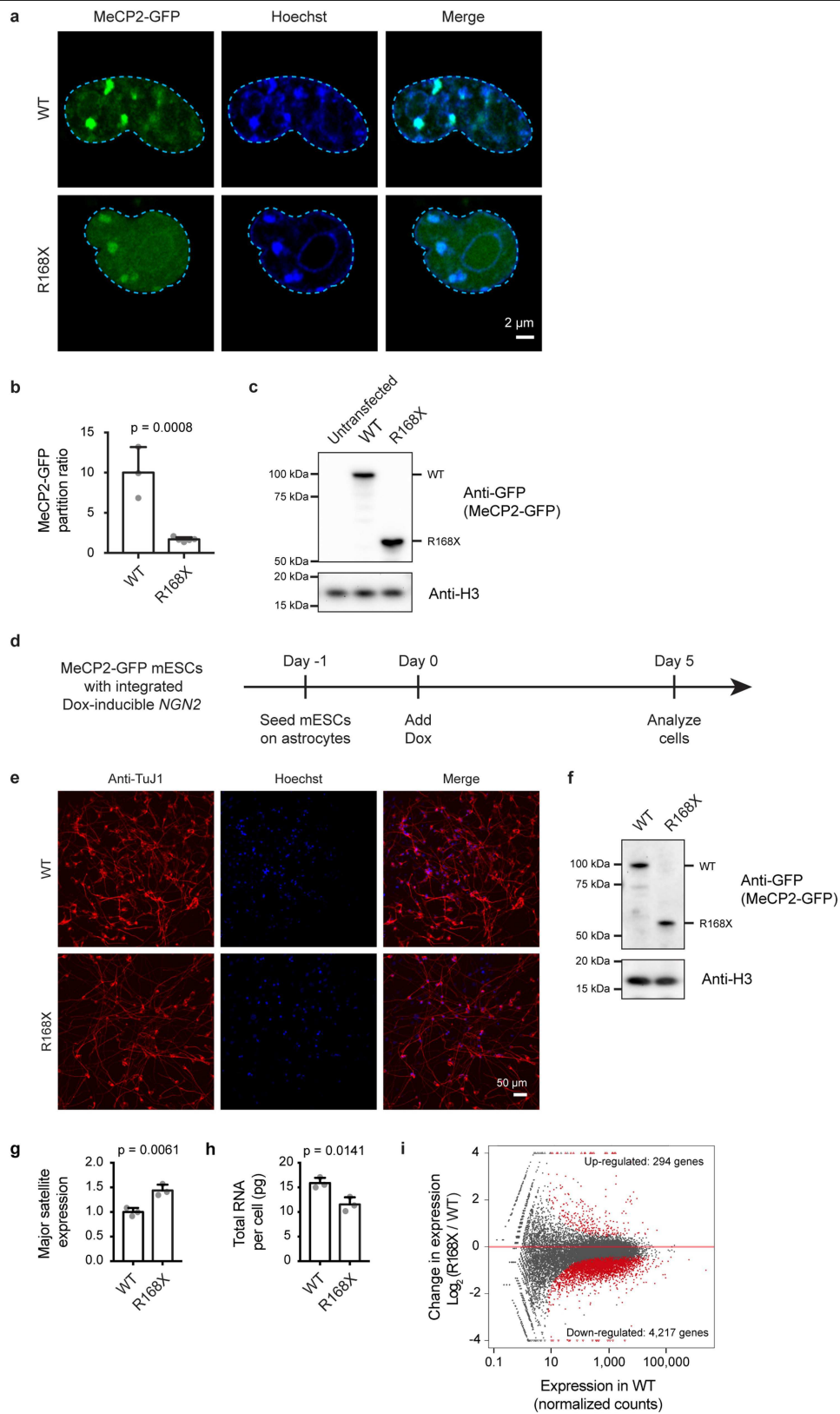
g, DNA-Cy5 partition ratios in wild-type MeCP2-GFP and Mini droplets for experiments in **e**. Fields per condition $n = 15$. **h**, Droplet experiments examining ability of MeCP2 wild-type and Mini condensates to enrich TBLR1-CTD. Wild-type MeCP2-GFP or Mini at 4 μ M was mixed with TBLR1-CTD-mCherry at 10 μ M in droplet formation buffers with 125 mM NaCl and 10% PEG-8000. **i**, TBLR1-CTD-mCherry partition ratios in wild-type MeCP2-GFP and Mini droplets for experiments in **h**. Fields per condition $n = 12$. **j**, Live-cell microscopy of endogenous-tagged wild-type MeCP2-GFP and Mini proteins with Hoechst staining in mouse ES cells. **k**, Partition ratios of MeCP2-GFP proteins at heterochromatin condensates for experiments in **j**. $n = 10$ cells per condition. $P = 0.1161$, $t = 1.6509$, $df = 18$, two-tailed Student's t -test. All data are mean $\pm$ s.d.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | R168X mutant MeCP2 displays reduced partitioning into heterochromatin condensates and causes disease-relevant cellular phenotypes in mouse ES cells. **a**, Live-cell images of endogenous-tagged wild-type MeCP2-GFP and R168X mutant proteins with Hoechst staining in mouse ES cells. **b**, Partition ratios of MeCP2-GFP proteins at heterochromatin condensates for experiments in **a**. Cells per condition: WT ($n=11$), R168X ($n=10$). $P<0.0001$, $t=12.13$, $df=19$, two-tailed Student's t -test. **c**, MeCP2-GFP signal in endogenous-tagged wild-type MeCP2-GFP and R168X mutant mouse ES cells measured by flow cytometry. $n=3$ biologically independent samples per condition. For example flow cytometry gating strategy, see Supplementary Fig. 2. **d**, Western blot of endogenous-tagged wild-type MeCP2-GFP and R168X mutant mouse ES cells. Anti-H3 was used as a processing control. For gel source data, see Supplementary Fig. 1. **e**, Number of heterochromatin condensates per cell in endogenous-tagged wild-type MeCP2-GFP and R168X mutant mouse ES cells. $n=16$ cells per condition. $P=0.0149$, $t=2.5832$, $df=30$, two-tailed Student's t -test. **f**, Heterochromatin condensate volumes in endogenous-tagged wild-type MeCP2-GFP and R168X mutant mouse ES cells. Condensates per condition: WT ($n=206$), R168X ($n=273$). $P<0.0001$, $t=4.2065$, $df=477$,

two-tailed Student's t -test. **g**, Live-cell images of endogenous-tagged MeCP2-GFP (wild type or R168X mutant) and HP1 α -mCherry in mouse ES cells. **h**, Partition ratios of HP1 α -mCherry at heterochromatin condensates for experiments in **g**. Cells per condition: WT ($n=6$), R168X ($n=20$). $P<0.0001$, $t=5.7136$, $df=24$, two-tailed Student's t -test. **i**, HP1 α -mCherry signal in endogenous-tagged MeCP2-GFP (wild-type or R168X mutant) and HP1 α -mCherry mouse ES cells measured by flow cytometry. $n=3$ biologically independent samples per condition. For example flow cytometry gating strategy, see Supplementary Fig. 2. **j**, Normalized major satellite repeat expression in endogenous-tagged wild-type MeCP2-GFP and R168X mutant mouse ES cells. $n=3$ biologically independent samples per condition. $P=0.0017$, $t=7.5436$, $df=4$, two-tailed Student's t -test. **k**, Total RNA per cell in endogenous-tagged wild-type MeCP2-GFP and R168X mutant mouse ES cells. $n=3$ biologically independent samples per condition. $P=0.0324$, $t=3.2154$, $df=4$, two-tailed Student's t -test. **l**, RNA-seq comparing endogenous-tagged wild-type MeCP2-GFP and R168X mutant mouse ES cells. Differentially expressed genes (red dots) were determined by two-tailed Wald test with multiple test adjusted $P<0.1$. For both conditions, $n=3$ biologically



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | R168X mutant in mouse ES cells and neurons. a, Live-cell images of mouse ES cells overexpressing either wild-type or R168X mutant MeCP2–GFP. **b,** Partition ratios of MeCP2–GFP proteins at heterochromatin condensates relative to the nucleoplasm for experiments in **a**. Cells per condition: WT ($n = 3$), R168X ($n = 5$). $P = 0.0008$, $t = 6.1529$, $df = 6$, two-tailed Student's t -test. **c,** Western blot of mouse ES cells overexpressing either wild-type or R168X mutant MeCP2–GFP. Anti-H3 was used as a processing control. For gel source data, see Supplementary Fig. 1. **d,** Schematic of generation of mouse ES-cell-derived neurons. Endogenous-tagged wild-type MeCP2–GFP and R168X mutant mouse ES cells were modified for Dox-inducible *NGN2* expression using the PiggyBac system. Prior to neuronal differentiation, mouse ES cells were seeded on astrocytes. Neuronal differentiation was induced by adding doxycycline to drive *NGN2* expression. Five days after induction of *NGN2* expression, neurons were analysed. **e,** Fixed-cell

immunofluorescence images of neurons derived from wild-type MeCP2–GFP and R168X mutant mouse ES cells. Anti-Tuj1 staining was used to distinguish neurons. **f,** Western blot of endogenous-tagged wild-type MeCP2–GFP and R168X mutant neurons. Anti-H3 was used as a loading control. For gel source data, see Supplementary Fig. 1. **g,** Normalized major satellite repeat expression in endogenous-tagged wild-type MeCP2–GFP and R168X mutant neurons. $n = 3$ biologically independent samples per condition. $P = 0.0061$, $t = 5.3004$, $df = 4$, two-tailed Student's t -test. **h,** Total RNA per cell in endogenous-tagged wild-type MeCP2–GFP and R168X neurons. $n = 3$ biologically independent samples per condition. $P = 0.0141$, $t = 4.1676$, $df = 4$, two-tailed Student's t -test. **i,** RNA-seq comparing endogenous-tagged wild-type MeCP2–GFP and R168X mutant neurons. Differentially expressed genes (red dots) were identified using a two-tailed Wald test with multiple test adjusted $P < 0.1$. For both conditions, $n = 3$ biologically independent samples. Data are mean $\pm$ s.d.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ A description of all covariates tested
☒	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Microscopy image collection was performed using Zeiss ZEN Black Edition (v. 2.3) or with MetaMorph (v. 7.10.3.279) depending on the microscope (see methods). Western blot image collection was performed using Bio-Rad Image Lab (v. 6.0.1).
Data analysis	Image analysis was performed using FIJI/ImageJ (v. 2.0.0-rc-65). In vitro droplet images and heterochromatin condensate volumes were analyzed with custom scripts using Python (v. 3.4.3). Custom code used for in vitro droplet analysis is available at www.github.com/jehenninger/in_vitro_droplet_assay . Custom code used for heterochromatin condensate volume analysis is available at www.github.com/jehenninger/MECP2_neuron . Western blot images were analyzed using Bio-Rad Image Lab (v. 6.0.1). Flow cytometry data was analyzed using FlowJo (v. 10). Data was plotted using GraphPad Prism (v. 7.0a). RNA-seq data was processed and analyzed using STAR aligner (v. 2.6.1a), RSEM (v. 1.2.31), DESeq2 (v. 1.24.0). Protein sequence analysis was performed using localCIDER (v. 0.1.14), PONDR VSL2, and Clustal Omega (v. 1.2.4).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Relevant data supporting the findings of this study are available within the paper and Supplementary Information files. RNA-seq datasets generated in this study

have been deposited in the Gene Expression Omnibus under accession code GSE139033. Uncropped gel images can be found in Supplementary Figure 1. Additional data are available from the corresponding author upon reasonable request. The following publicly available data were used in this study: GTEx v. 7 RNA-seq Median Gene TPMs by Tissue (www.gtexportal.org), RettBASE MECP2 Variant List (mecp2.chw.edu.au/mecp2/mecp2_home.php), and UniProt Cluster ID: UniRef50_Q9Z2D6 (www.uniprot.org/uniref/UniRef50_Q9Z2D6).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For in vitro droplet assays, we typically imaged between 5-15 independent fields of view (typically a few hundred droplets) per condition, which is consistent with current practices in this field (Lin et al., Mol Cell 2015; Sabari et al., Science 2018; Boija et al., Cell 2018; Guo et al., Nature 2019). For cell imaging studies, we examined between 3-20 cells per condition, which in line with current studies in this field (Sabari et al., Science 2018; Cho et al., Science 2018; Chong et al., Science 2018). For transcription reporter assay, we used 3 biologically independent samples per condition, which is in line with current standards (Chong et al., Science 2018; Boija et al., Cell 2018; Hnisz et al., Mol Cell 2015). For RNA-seq and RT-qPCR analyses, we used 3 biologically independent samples per condition, which is in line with current standards in this field (Weintraub et al., Cell 2017).
Data exclusions	No data were excluded from the analyses.
Replication	Experiments were performed multiple times with similar results.
Randomization	Randomization was not possible due to the nature of the study, as the experiments examined properties of samples of distinct identities. Instead, appropriate controls were performed and are included in the figures.
Blinding	Blinding was not performed for any experiments in this study due to feasibility of data collection and data analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Primary antibodies: anti-GFP (Takara Bio 632381), anti-Histone H3 (Cell Signaling Technology 4499), anti-MAP2 (Invitrogen MA5-12823), anti-MED1 (Abcam ab64965), anti-TuJ1 (Covance MMS-435P). Secondday antibodies: anti-mouse Alexa Fluor 568 (Invitrogen A11031), anti-rabbit Alexa Fluor 555 (Invitrogen A21428), anti-mouse-HRP (GE Healthcare NXA931V), anti-rabbit-HRP (GE Healthcare NA934V).
Validation	Anti-MAP2 antibody (Invitrogen MA5-12823) was validated by the vendor using relative immuno-staining of MAP2-positive and negative tissues. Anti-MED1 antibody (Abcam ab64965) was validated by immuno-staining after siRNA knockdown using Silencer Select siRNAs toward MED1 (s72006 and s72007) (Ambion 4390771); this antibody was previously used to detect MED1 in mouse and human cells (Sabari et al., Science 2018; Boija et al., Cell 2018; Guo et al., Nature 2019; Klein et al., Science 2020). Anti-GFP antibody (Takara Bio 632381) was validated by western blot of lysates expressing GFP and a negative control lysate. Anti-Histone H3 antibody (Cell Signaling Technology 4499) was validated by vendor using western blot of human, mouse, rat, and monkey cell lysates. Anti-TuJ1 antibody (Covance MMS-435P) was validated by vendor using immuno-staining of human neurons and western blot in human, mouse, and rat lysates.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	V6.5 murine ESCs were acquired from the Rudolf Jaenisch. MeCP2-GFP Mini (Δ NIC, Tillotson et al., Nature 2017) murine ESCs were acquired from Adrian Bird. HEK293T cells were commercially acquired (ATCC CRL-3216).
Authentication	V6.5 murine ESCs were authenticated by STR analysis compared to commercially acquired cells of the same line. MeCP2-GFP Mini murine ESCs and HEK293T cells were not subjected to authentication.
Mycoplasma contamination	All cell lines tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified lines were used in this study.

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	Female chimeric mice of CD1-IGD/V6.5 background were used for experiments at 10-weeks of age. Mice were housed in a facility with centrally controlled environment with a 12-hour light/12-hour dark cycle, temperature of 68-72 °F, humidity of 30-50%, and routine bedding, food, and water changes.
Wild animals	No wild animals were used in this study.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All experiments using mice were carried out with approval from the MIT Committee on Animal Care (CAC) under protocol number 1019-059-22. Experiments were carried out under the supervision of the Division of Comparative medicine (DCM) at MIT, which provides centralized management of the animal facility at the Whitehead Institute for Biomedical Research. The mouse facility conforms to federal guidelines (Animal Welfare Assurance Number A3125-01), and MIT is accredited by the Assessment and Accreditation of Laboratory Animal Care (AAALAC).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Murine ESCs were dissociated using TrypLE Express (Gibco 12604) and the dissociation reaction was quenched using serum/LIF media. Cells were resuspended in single cell suspension in PBS and passed through a cell strainer (Corning 352235).
Instrument	Cells were analyzed using a LSRII flow cytometer (BD).
Software	Data was analyzed using FlowJo v. 10 (BD).
Cell population abundance	Analysis of sorted cell populations was not performed in this study.
Gating strategy	Standard forward and side scatter gating was used to exclude debris and isolate singlet cells for analysis. See Supplementary Figure 2 for example gating strategy.

- ☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.