

HiChIRP reveals RNA-associated chromosome conformation

Maxwell R. Mumbach^{1,2,13}, Jeffrey M. Granja^{1,3,13}, Ryan A. Flynn¹, Caitlin M. Roake^{4,5}, Ansuman T. Satpathy^{1,6}, Adam J. Rubin⁷, Yanyan Qi¹, Zhaozhao Jiang⁸, Shadi Shams¹, Bryan H. Louie¹, Jimmy K. Guo¹, David G. Gennert^{1,2}, M. Ryan Corces¹, Paul A. Khavari^{1,7}, Maninjay K. Atianand^{1,9}, Steven E. Artandi⁴, Katherine A. Fitzgerald⁸, William J. Greenleaf^{1,2,10,11,14*} and Howard Y. Chang^{1,2,7,12,14*}

Modular domains of long non-coding RNAs can serve as scaffolds to bring distant regions of the linear genome into spatial proximity. Here, we present HiChIRP, a method leveraging bio-orthogonal chemistry and optimized chromosome conformation capture conditions, which enables interrogation of chromatin architecture focused around a specific RNA of interest down to approximately ten copies per cell. HiChIRP of three nuclear RNAs reveals insights into promoter interactions (7SK), telomere biology (telomerase RNA component) and inflammatory gene regulation (lincRNA-EPS).

Long non-coding RNAs (lncRNAs) play important roles in diverse biological processes, including dosage compensation, cell differentiation and cell growth control¹. Many lncRNAs have been reported to function by using their modular domains to bind proteins, DNA or other RNAs, allowing them to engage in long-range chromatin interactions^{2–4}. Despite recent advances in all-to-all mapping of RNA–chromatin interactions^{5–7}, the dynamic range of RNA expression levels (over 1 million fold) means that these methods can report only on abundant RNAs, despite exhaustive sequencing. Therefore, targeted technologies that comprehensively map spatial interactions between an RNA of interest and chromatin are needed. Since Hi-C maps all possible proximity ligations in the genome, enrichment-based strategies to target locus- or protein-specific interactions allow for increased sequencing power at specific chromatin loops of interest in the biological system^{8–12}. We reasoned that developing a similar strategy to target loops associated with specific RNAs would provide insights into the landscape of RNA-centric chromosome conformation.

We designed HiChIRP to identify RNA-associated chromatin contacts by replacing the chromatin immunoprecipitation step in the HiChIP protocol with chromatin isolation by RNA purification (ChIRP) enrichment¹³ (Fig. 1a and Methods). Since ChIRP (and similar methods) requires RNA capture by biotinylated probes, we incorporated an azido-modified nucleotide into chromatin contacts during chromosome conformation capture (3C), and, following RNA enrichment with biotinylated probes, the azido-containing

chromatin contacts were subjected to copper-free dibenzocyclooctyne (DIBO) ‘CLICK’ chemistry to covalently conjugate a biotin for subsequent contact enrichment (Supplementary Fig. 1a and Supplementary Note). An additional consideration is that ChIRP is generally performed in glutaraldehyde cross-linked cells; however, this has been previously characterized as a cross-linker for 3C¹⁴. Finally, the biochemical steps required to generate chromatin contact libraries revealed that standard conditions are not suitable for maintaining intact RNA (Fig. 1b). To address this challenge, we optimized a 3C protocol to reduce the preparation time from ~9 h to ~4 h while preserving RNA integrity and maintaining a similar percentage of long-range ligation efficiency relative to standard conditions (Fig. 1b)⁸. We then performed HiChIP of Oct4 and CTCF in mouse embryonic stem cells (mESCs) and a human B lymphoblastoid cell line (GM12878), respectively, using optimized HiChIRP 3C conditions, and observed increased signal-to-background relative to that observed under standard 3C conditions (Fig. 1c, Supplementary Fig. 1b–e and Supplementary Table 1).

We first performed HiChIRP on the 7SK small nuclear RNA (snRNA) (~200,000 copies per cell, 332 nucleotides), which functions to promote transcriptional pause release by regulating chromatin structure (Supplementary Fig. 2)¹⁵. To ensure that the HiChIRP signal is RNA-dependent, we performed a ‘split-probe’ approach using ‘even’ and ‘odd’ 7SK HiChIRP probe datasets, and found high concordance (Supplementary Fig. 3a; $R = 0.87$). Non-targeting lacZ and RNase control experiments identified 55- and 42-fold fewer unique contacts than in the HiChIRP sample, respectively (Supplementary Figs. 2 and 3b and Supplementary Note). While the depletion of DNA in the controls indicates that the majority of the signal is RNA-dependent, we conservatively removed 7SK HiChIRP loops that had significantly high signal (false discovery rate (FDR) < 0.1) in the RNase control (Supplementary Fig. 2 and Methods). These data demonstrate that 7SK HiChIRP recovers DNA contacts similarly to other conformation technologies that report architectural features of the genome.

¹Center for Personal Dynamic Regulomes, Stanford University School of Medicine, Stanford, CA, USA. ²Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. ³Program in Biophysics, Stanford University School of Medicine, Stanford, CA, USA. ⁴Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA. ⁵Cancer Biology Program, Stanford University School of Medicine, Stanford, CA, USA. ⁶Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA. ⁷Program in Epithelial Biology, Stanford University School of Medicine, Stanford, CA, USA. ⁸Program in Innate Immunity, University of Massachusetts Medical School, Worcester, MA, USA. ⁹Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. ¹⁰Department of Applied Physics, Stanford University, Stanford, CA, USA. ¹¹Chan Zuckerberg Biohub, San Francisco, CA, USA. ¹²Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA. ¹³These authors contributed equally: M. R. Mumbach, J. M. Granja. ¹⁴These authors jointly supervised this work: W. J. Greenleaf, H. Y. Chang. *e-mail: wjg@stanford.edu; howchang@stanford.edu

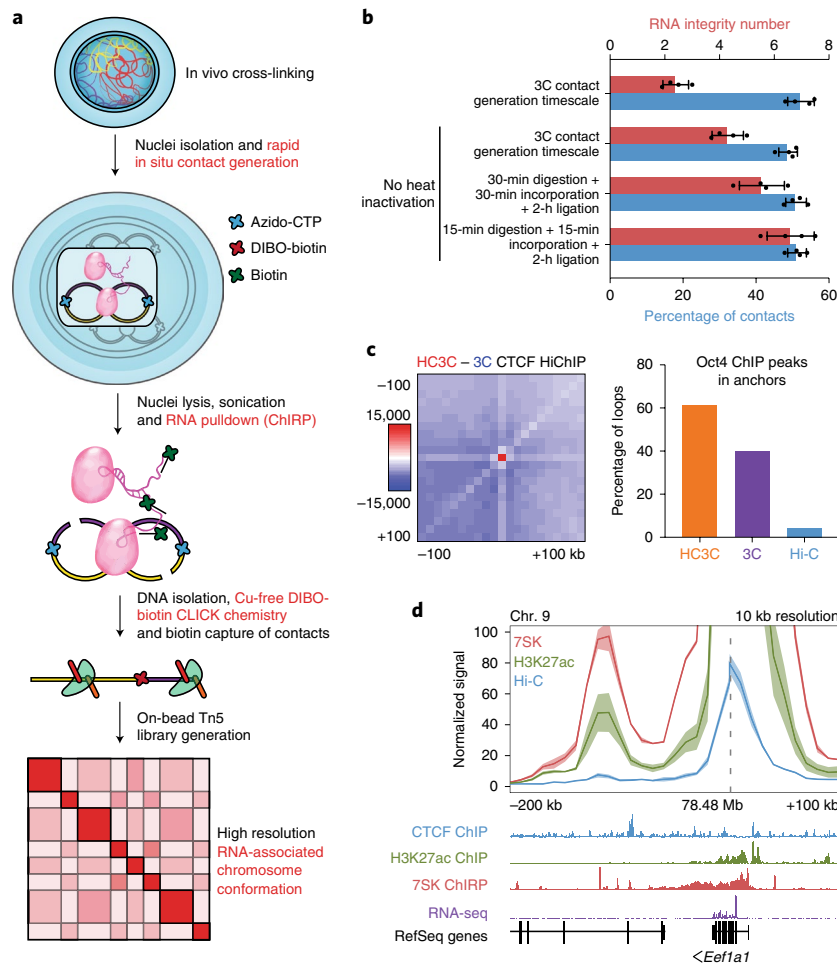


Fig. 1 | HiChIRP enables RNA-centric chromosome conformation associated with 7SK. **a**, Schematic of the HiChIRP method. **b**, RNA integrity and 3C quality after 3C library construction in the standard 3C protocol and modified HiChIRP (HC) 3C conditions ($n = 4$; mean with s.d.). **c**, Left, differential APA on the union set of loops between CTCF HiChIP performed with HC 3C and standard 3C protocols. Right, percentage of Oct4 HiChIP and Hi-C biased loop anchors containing Oct4 ChIP-sequencing peaks using the HC 3C and standard 3C protocols. **d**, Virtual 4C visualization of 7SK HiChIRP loop signal bias at the *Eef1a1* promoter ($n = 2$; mean and s.d. shading; RefSeq, NCBI Reference Sequence Database).

Considering the low complexity of the RNase control samples, we made use of mESC Hi-C as a non-enriched ‘input’ dataset to interrogate the enrichment and specificity of the chromatin loops identified by 7SK HiChIRP (Supplementary Note)¹⁶. We identified 2,421 loops enriched and 3,794 loops depleted in 7SK HiChIRP signal (Supplementary Fig. 3c and Supplementary Tables 2 and 3). We then verified that the 7SK HiChIRP-enriched loops were supported by Hi-C as bona fide chromatin loops (Supplementary Fig. 3d–f). Furthermore, we performed loop calling using an orthogonal algorithm and observed high concordance, demonstrating the robustness of our enriched loops (Supplementary Fig. 3g and Supplementary Table 4)¹⁷. To validate our loops, we observed increased 7SK ChIRP-sequencing (ChIRP-seq) signal at enriched loop anchors compared with depleted (Supplementary Fig. 3h; Kolmogorov–Smirnov test $P < 2.2 \times 10^{-16}$). The enriched loops also spanned shorter distances and were annotated for active regulatory elements, in agreement with known functions of 7SK at promoters and the concept of shorter gene regulatory loops within topological domains (Supplementary Fig. 3i,j)¹⁵. To visualize the enrichment of 7SK HiChIRP over Hi-C, we performed virtual circular chromosome conformation capture (4C) analysis of the *Eef1a1* promoter, and observed 7SK HiChIRP enrichment (Fig. 1d). Conversely, virtual 4C analysis of a region containing no gene

activity demonstrated a depletion of 7SK HiChIRP signal (Supplementary Fig. 3k). To observe this globally, we performed meta-virtual 4C (metav4C) analysis of all chromatin loops containing a 7SK ChIRP-seq peak or an active promoter, as well as loops exhibiting no gene activity, and identified increased and depleted 7SK HiChIRP signal, respectively (Supplementary Fig. 3l).

Telomerase RNA component (TERC; ~750 copies per cell, 541 nucleotides) is part of the telomerase holoenzyme, and it provided an opportunity to probe a medium-expression non-coding RNA with HiChIRP (Supplementary Fig. 2)¹³. We generated TERC HiChIRP libraries, as well as RNase and TERC-knockout cell negative controls, and observed 24-fold and 97-fold fewer unique contacts, respectively (Supplementary Figs. 2 and 4a,b). Given the specialized function of TERC at telomeres, we first examined telomeric DNA sequence repeats in the HiChIRP data. TERC HiChIRP libraries were enriched for telomeric DNA relative to TERC-knockout cells, the RNase control and Hi-C, indicating high specificity (Fig. 2a and Supplementary Fig. 4c). Strikingly, the majority of reads containing telomeric sequences (~92%) formed interactions with other telomeric reads (T-T), supporting previous observations that telomeres are highly compacted (Fig. 2b)¹⁸. Furthermore, the majority of the ~8% of telomere-to-nontelomere (T-X) reads map to subtelomeric regions within 1 Mb of a chromosome end, supported

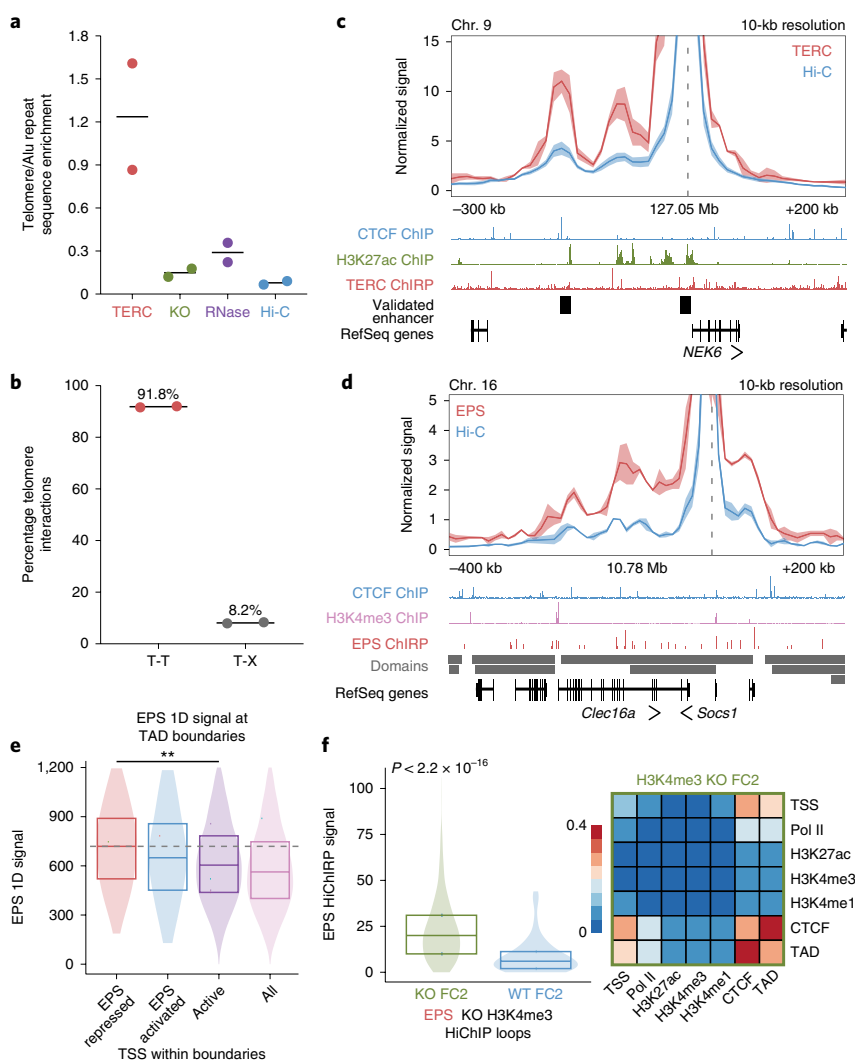


Fig. 2 | HiChIRP of TERC and low-abundance lincRNA-EPS. **a**, Telomere versus Alu repeat enrichment in TERC HiChIRP relative to its knockout (KO) and RNase controls, as well as GM12878 Hi-C ($n=2$; mean is denoted by the line). **b**, Percentage of TERC HiChIRP telomere reads interacting with other telomere reads (T-T) or with DNA regions outside of the telomere (T-X; $n=2$; mean is denoted by the line). **c**, Virtual 4C visualization of TERC HiChIRP loop signal bias at the *NEK6* promoter, with annotation of clustered regularly interspaced short palindromic repeats (CRISPR) editing-validated enhancers regulating *NEK6* expression ($n=2$; mean and s.d. shading; RefSeq, Reference Sequence Database). **d**, Virtual 4C visualizations of lincRNA-EPS loop signal bias at the *Socs1* promoter with nearby domain boundaries ($n=2$; mean and s.d. shading). **e**, lincRNA-EPS 1D signal at TAD boundaries insulating genes repressed by lincRNA-EPS ($n=77$), boundaries insulating genes activated by lincRNA-EPS ($n=46$), boundaries insulating active Ensembl TSS ($n=4,965$) and all boundaries ($n=6,639$). Box is median with quartiles, and significance ($P=0.014$ between repressed and active) was assessed using a Kolmogorov-Smirnov test. **f**, Left, lincRNA-EPS HiChIRP signal supporting H3K4me3 lincRNA-EPS knockout (KO; $n=814$) and wild-type (WT; $n=84$) biased (FC2 = fold change > 2) loop sets. Box is median with quartiles, and significance was assessed using a Kolmogorov-Smirnov test. Right, BMDM ENCODE ChIP annotation of lincRNA-EPS KO-biased H3K4me3 HiChIP loops.

by Hi-C (Supplementary Fig 4d). Finally, examination of the reads not associated with telomeres revealed that TERC HiChIRP was enriched over Hi-C at the subtelomeric ends of chromosomes, analogous to its telomeric enrichment, and broadly depleted versus Hi-C for internal regions (Supplementary Fig. 4e). Because TERC HiChIRP enriches for subtelomeric interactions, we next considered the *IGH* locus, which has a high incidence of translocations in B-cell neoplasms and is within the subtelomeric region of chromosome 14. We tested whether TERC HiChIRP in GM12878 cells could identify known *IGH* translocations on the basis of spatial proximity. Multiple *IGH* translocations were found to have strong and reproducible ($R=0.89$) TERC HiChIRP signal, supported by Hi-C ($R=0.96$), which agrees with reports that certain classes of translocations can be predicted by chromosome conformation (Supplementary Fig. 4f)¹⁹.

While TERC HiChIRP is generally depleted for internal chromosome regions, we found specific internal chromatin contacts in TERC HiChIRP, consistent with our previous one-dimensional (1D) ChIRP-seq of TERC¹³. First, we conservatively removed TERC HiChIRP loops with significant signal in the RNase control (Supplementary Fig. 2). Loop overlap and differential analysis with published GM12878 Hi-C as ‘input’ maps identified 794 loops enriched and 770 loops depleted in TERC signal (Supplementary Fig. 5a and Supplementary Table 3)⁸. These TERC HiChIRP-enriched loops were supported by Hi-C signal as well as a reciprocal loop-calling algorithm (Supplementary Fig. 5b–d and Supplementary Table 4)¹⁷. Annotation of TERC HiChIRP-enriched loops identified shorter-length enhancer–promoter loops at select oncogenes, supported by ChIRP-seq (Supplementary Fig. 5e,f; Kolmogorov–Smirnov test $P < 2.2 \times 10^{-16}$)¹³. To demonstrate this enrichment, we performed virtual 4C analysis on *NEK6*, a gene

commonly overexpressed in B-cell lymphomas, and observed enrichment of TERC HiChIRP signal, while loops containing no gene regulatory activity are depleted for TERC HiChIRP signal (Fig. 2c and Supplementary Fig. 5g)²⁰. Interestingly, TERC HiChIRP signal was found at a recently characterized enhancer–promoter interaction regulating *NEK6* expression, which suggests that the effects of TERC knockout on certain oncogene expression levels could be of interest²⁰. HiChIRP of TERC adds support to the theory that telomerase subunits may have potential gene regulatory function in regions other than telomeres²¹.

Finally, we assessed the applicability of HiChIRP on the low-abundance lincRNA-EPS (~10 copies per cell, 2,531 nucleotides), which is expressed in resting macrophages to restrain inflammatory gene expression programs²². We performed HiChIRP on lincRNA-EPS in bone-marrow-derived macrophages (BMDMs), as well as on RNase-treated and lincRNA-EPS-knockout cell negative controls, and observed 192-fold and 42-fold fewer unique contacts, respectively (Supplementary Fig. 2). First, we conservatively removed lincRNA-EPS HiChIRP loops with significant signal in the RNase control (Supplementary Fig. 2). Loop overlap and differential analysis with low-depth ‘input’ Hi-C maps that we generated identified 543 loops enriched and 650 loops depleted in lincRNA-EPS signal (Supplementary Fig. 6a and Supplementary Table 3). These lincRNA-EPS HiChIRP-enriched loops were supported by Hi-C signal and an orthogonal loop-calling algorithm (Supplementary Fig. 6b–e and Supplementary Table 4)¹⁷. Annotation of lincRNA-EPS-enriched loops revealed that they were not most enriched for enhancer–promoter interactions, as 7SK and TERC were, but rather for interactions linking a CTCF boundary to a gene promoter, supported by ChIRP-seq (Supplementary Fig. 6f,g; Kolmogorov–Smirnov test $P < 2.2 \times 10^{-16}$).

Since lincRNA-EPS is highly associated with CTCF–transcriptional start site (TSS) interactions, we next examined whether this association increased at topological domain boundaries containing immune response genes (IRGs) that are repressed by lincRNA-EPS identified by RNA-sequencing²². We performed virtual 4C analysis on *Socs1*, a well-known upstream regulator of cytokine signaling that is repressed by lincRNA-EPS, and observed TAD boundary–promoter loops with enriched signal in lincRNA-EPS HiChIRP (Fig. 2d). To examine this globally, we found increased lincRNA-EPS 1D signal at both repressed gene promoters (Kolmogorov–Smirnov test $P = 2.0 \times 10^{-4}$ for repressed versus all active promoters) and the domain boundaries that insulate them (Fig. 2e and Supplementary Fig. 6h,j; Kolmogorov–Smirnov test $P = 1.4 \times 10^{-3}$ for boundaries insulating repressed versus all active promoters). Conversely, virtual 4C analysis of an enhancer–enhancer interaction not associated with a CTCF TAD boundary or gene promoters demonstrated depletion of signal compared with Hi-C (Supplementary Fig. 6j). Therefore, HiChIRP of lincRNA-EPS suggests that lincRNA-EPS uses chromatin interactions between topologically associated domain (TAD) boundaries and the promoters of certain IRGs to mediate gene repression.

Since lincRNA-EPS is reported to bind to IRG promoters and repress their expression, we next determined whether this regulation involved changes in the nearby topological boundary chromatin architecture. We therefore used *lincRNA-EPS*^{-/-} (knockout) BMDM cells and performed HiChIP of SMC1a and H3K4me3 to identify protein-associated chromatin looping changes at TADs (SMC1a) and promoters (H3K4me3) resulting from a loss of lincRNA-EPS (Supplementary Fig. 6k)²². lincRNA-EPS knockout did not alter global TAD boundaries, which is expected owing to the low copy number of the lincRNA. Interestingly, we found that knockout-biased H3K4me3 loops were enriched at TAD boundaries and contained high lincRNA-EPS HiChIRP signal (Fig. 2f; Kolmogorov–Smirnov test $P < 2.2 \times 10^{-16}$). These results further support the lincRNA-EPS HiChIRP results and suggest that lincRNA-EPS may impact target promoter interactions, perhaps as an indirect effect of gene derepression (Supplementary Fig. 6l).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at <https://doi.org/10.1038/s41592-019-0407-x>.

Received: 5 October 2018; Accepted: 5 April 2019;
Published online: 27 May 2019

References

- Batista, P. J. & Chang, H. Y. *Cell* **152**, 1298–1307 (2013).
- Guttman, M. & Rinn, J. L. *Nature* **482**, 339–346 (2012).
- Wang, K. C. et al. *Nature* **472**, 120–124 (2011).
- Engreitz, J. M. et al. *Science* **341**, 1237973 (2013).
- Li, X. et al. *Nat. Biotechnol.* **35**, 940–950 (2017).
- Quinodoz, S. A. et al. *Cell* **174**, 744–757 (2018).
- Bell, J. C. et al. *eLife Sci.* **7**, e27024 (2018).
- Rao, S. S. P. et al. *Cell* **159**, 1665–1680 (2014).
- Mumbach, M. R. et al. *Nat. Methods* **13**, 919–922 (2016).
- Mumbach, M. R. et al. *Nat. Genet.* **49**, 1602–1612 (2017).
- Mifsud, B. et al. *Nat. Genet.* **47**, 598–606 (2015).
- Fullwood, M. J. et al. *Nature* **462**, 58–64 (2009).
- Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. *Mol. Cell* **44**, 667–678 (2011).
- Hsieh, T.-H. S., Fudenberg, G., Goloborodko, A. & Rando, O. J. *Nat. Methods* **13**, 1009–1011 (2016).
- Flynn, R. A. et al. *Nat. Struct. Mol. Biol.* **23**, 231–238 (2016).
- Bonev, B. et al. *Cell* **171**, 557–572.e24 (2017).
- Bhattacharyya, S., Chandra, V., Vijayanand, P. & Ay, F. Preprint at <https://doi.org/10.1101/412833> (2018).
- Bandaria, J. N., Qin, P., Berk, V., Chu, S. & Yildiz, A. *Cell* **164**, 735–746 (2016).
- Robbiani, D. F., Nussenzweig, M. C. & Chromosome Translocation, B. *Annu. Rev. Pathol.* **8**, 79–103 (2013).
- Huang, Y. et al. *Cell Rep.* **18**, 2918–2931 (2017).
- Park, J.-I. et al. *Nature* **460**, 66–72 (2009).
- Atianand, M. K. et al. *Cell* **165**, 1672–1685 (2016).

Acknowledgements

We thank members of the Chang and Greenleaf laboratories for helpful discussions, and J. Tumej for artwork. We thank Jena Biosciences for design and synthesis of the 5-azido-PEG4-dCTP nucleotide. We thank X. Ji and J. Collier at the Stanford Functional Genomics Facility. This work was supported by National Institutes of Health (NIH) grants (no. P50-HG007735 to H.Y.C. and W.J.G.; no. R35-CA209919 to H.Y.C.; no. U19-AI057266 to W.J.G.; no. K08CA23188-01 to A.T.S.); the Human Frontier Science Program (to W.J.G.); the Rita Allen Foundation (W.J.G.); and the Scleroderma Research Foundation (H.Y.C.). M.R.M. acknowledges support from the National Science Foundation Graduate Research Fellowship. A.T.S. was supported by a Parker Bridge Scholar Award from the Parker Institute for Cancer Immunotherapy, a Career Award for Medical Scientists from the Burroughs Wellcome Fund and the Human Vaccines Project Michelson Prize for Human Immunology and Vaccine Research. M.R.C. is supported by a grant from the Leukemia & Lymphoma Society Career Development Program. W.J.G. is a Chan Zuckerberg Biohub investigator. Sequencing was performed by the Stanford Functional Genomics Facility (NIH grant no. S10OD018220). H.Y.C. is an Investigator of the Howard Hughes Medical Institute.

Author contributions

M.R.M., J.M.G., R.A.F. and H.Y.C. developed the method. M.R.M., J.M.G., R.A.F., S.S., B.H.L., J.K.G., D.G.G. and M.R.C. performed experiments. M.K.A., A.T.S., Y.Q. and Z.J. cultured bone-marrow-derived macrophages. C.M.R. generated TERC knockout cell lines. J.M.G., M.R.M. and A.J.R. analyzed HiChIRP and HiChIP datasets. M.R.M., J.M.G., W.J.G. and H.Y.C. wrote the manuscript with input from all authors.

Competing interests

H.Y.C. has affiliation with Accent Therapeutics (Founder, SAB), 10X Genomics (SAB) and Spring Discovery (SAB). W.J.G. has affiliation with 10X Genomics (SAB) and Guardant Health (Consultant).

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41592-019-0407-x>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to W.J.G. or H.Y.C.

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 America, Inc. 2019

Methods

Cell culture and fixation. The University of Massachusetts Medical School Animal Care and Use Committees approved all animal work.

mESCs (v.6.5, Novus Biologicals; NBPI-41162) were cultured in Knockout DMEM (Gibco) + 15% fetal bovine serum and leukemia inhibitory factor (Millipore) to 80% confluence. GM12878 (Coriell) cells were grown in RPMI 1640 (Gibco) with 15% fetal bovine serum to a concentration of 500,000 to 1 million cells per ml. lincRNA-EPS-knockout mice were obtained from a previous study²². BMDMs were differentiated from bone marrow cells with 20% L929 supernatant for 7 d.

Detached cell lines were pelleted and resuspended in fresh 1% formaldehyde (Thermo Fisher) or 1% glutaraldehyde (Sigma) at a volume of 1 ml cross-linker for 1 million cells. Cells were incubated at room temperature for 10 min with rotation. Glycine was added at a final concentration of 125 mM to quench the cross-linker, and cells were incubated at room temperature for 5 min with rotation. Finally, cells were pelleted and washed with PBS, pelleted again and stored at -80°C or immediately taken into the HiChIP or HiChIRP protocols⁹.

HiChIRP. Optimized *in situ* contact generation. Contact generation should be done on no more than 15–20 million cells per tube, and therefore for lower abundance lincRNAs, multiple tubes should be processed in parallel through contact generation, sonication and ChIRP, and then pooled after the DNA elution step. An RNase negative control is recommended, similar to ChIRP-seq and ChIRP-MS, to remove loops identified by both HiChIRP and the RNase control as potential artifacts of direct probe-DNA binding^{13,23}. ChIRP probes were designed using the same parameters as ChIRP-seq and ChIRP-MS protocols, and it is also recommended to use non-overlapping probe sets ('even' and 'odd') in separate ChIRP experiments to verify that the signal is probe-independent^{13,23}. For the abundant 7SK RNA, we used 15 million cells per replicate. For the less abundant lincRNA-EPS and TERC RNAs, we used ~90 million cells per replicate (6 tubes); after contact generation these 6 tubes were pooled into 3 Covaris tubes for sonication, and then after the ChIRP DNA elution the 3 ChIRP DNA samples were pooled together on DNA Clean and Concentrator columns (Zymo Research). Cross-linked cells were resuspended in 500 μl of ice-cold Hi-C lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 0.2% NP-40, 1 \times Roche protease inhibitors: 11697498001) and rotated at 4°C for 15 min. Nuclei were pelleted at 4°C for 5 min at 2,500g and the supernatant was discarded. Pelleted nuclei were washed once with 500 μl of ice-cold Hi-C lysis buffer. Supernatant was removed again and the pellet was resuspended in 100 μl of 0.5% SDS and incubated at 62°C for 5 min with no shaking or rotation. Then, 285 μl of water and 50 μl of 10% Triton X-100 were added and samples were rotated at 37°C for 15 min to quench the SDS. NEBuffer 2 (50 μl) and 15 μl of 25 U μl^{-1} MboI restriction enzyme (NEB, R0147) were then added and the sample was rotated at 37°C for 15 min. For lower starting material, the amount of restriction enzyme used was decreased linearly. Rather than heat-inactivating the MboI, we pelleted nuclei at 4°C for 5 min at 2,500g and discarded the supernatant. Pelleted nuclei were washed twice with 500 μl of 1 \times NEBuffer 2 and then resuspended in 536 μl of 1 \times NEBuffer 2. To fill in the restriction fragment overhangs and mark the DNA ends with biotin, we then added 16 μl of incorporation master mix: 1.5 μl of 10 mM azido-dCTP (Jena Biosciences, CLK-070); 1.5 μl of dATP, dGTP and dTTP at 10 mM each; and 10 μl of 5 U μl^{-1} DNA Polymerase I, Large (Klenow) Fragment (NEB, M0210). The reactions were then rotated at 37°C for 15 min. Next, 948 μl of ligation master mix was added: 150 μl of 10 \times NEB T4 DNA ligase buffer with 10 mM ATP (NEB, B0202), 125 μl of 10% Triton X-100, 3 μl of 50 mg ml^{-1} BSA (Thermo Fisher AM2616), 10 μl of 400 U μl^{-1} T4 DNA Ligase (NEB, M0202) and 660 μl of water. The reactions were then rotated at room temperature for 2 h. After proximity ligation, the nuclei within *in situ*-generated contacts were pelleted at 2,500g for 5 min at room temperature and the supernatant was removed.

Sonication and RNA pull-down. The nuclear pellet was brought up to 880 μl in nuclear lysis buffer (50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1% SDS, 1 \times Roche protease inhibitors: 11697498001) and transferred to a Covaris milliTUBE. Samples were sheared using a Covaris E220 using the following parameters: Fill Level = 10, Duty Cycle = 5, PIP = 140, Cycles/Burst = 200, Time = 40 min, and then clarified by centrifugation for 15 min at 16,100g at 4°C . We kept the sonication constant at 40 min for different amounts of cell starting material, although sonication time may need to be adjusted for different cell types. The ideal sonication time will be as short as possible to allow for efficient ChIRP signal over background. Too long a sonication can lead to separation of the lincRNA from the azido contact, which will increase the likelihood of a DNA fragment not making it through both enrichments and a loss in sample complexity. Clarified samples were transferred to Eppendorf tubes and a 2 \times volume of ChIRP hybridization buffer (1% SDS, 15% formamide, 1 mM EDTA, 50 mM Tris-HCl, pH 7, 750 mM NaCl, 1 \times Roche protease inhibitors: 11697498001) was added. Samples were split between 2 Eppendorf tubes to allow for the volume to fit at approximately 1.2 ml in each tube with a total of 2.4 ml for the entire ChIRP. For low-abundance RNAs starting with more clarified material, lysates can be pooled so that the ChIRP is done in a single 15-ml Falcon tube. For every 15 million cells, 30 μl of MyOne streptavidin C1 beads were washed and resuspended (30 μl) in ChIRP hybridization buffer and

then added to the clarified lysate along with 5 μl of RNaseOUT (Thermo). For the RNase control samples, instead of RNaseOUT, 30 μg of RNase A and RNase H for every 15 million cells were added. Samples were incubated at 37°C for 45 min with rotation and then placed on a magnetic rack. Supernatants were eluted into new tubes, 100 pmol of probe for every 15 million cells was added and samples were incubated at 37°C for 2 h (or overnight as a stopping point for the day) with rotation. For every 15 million cells, 100 μl of MyOne streptavidin C1 beads were washed and resuspended (50 μl) in nuclear lysis buffer and then added to the sample at 37°C for 45 min with rotation. After bead capture, beads were washed 5 times with ChIRP wash buffer (2 \times SSC, 0.5% SDS). Washing was performed at 37°C by shaking with 500 μl of a wash buffer, then placing the sample on the magnet and removing the supernatant.

DNA elution and copper-free CLICK chemistry. Sample beads were washed in 200 μl of ChIRP elution buffer (150 mM NaCl, 50 mM Tris-HCl, pH 8.0, 3 mM MgCl_2 , 10 mM dithiothreitol, 0.1% NP-40) and then resuspended in 200 μl of ChIRP elution buffer containing 4 μl of RNase A (5 mg ml^{-1} , Thermo) and 4 μl of RNase H (5 U μl^{-1} , Thermo)²⁴. Samples were then incubated at 37°C with shaking for 30 min, placed on a magnet and eluted into new tubes. Elution was repeated with another 200 μl of ChIRP elution buffer containing RNases. SDS to a final concentration of 0.5% and 20 μl of Proteinase K (20 mg ml^{-1} , Thermo Fisher) were then added to the 400- μl reaction. Samples were incubated at 55°C for 45 min with shaking. Samples were purified with DNA Clean and Concentrator columns (Zymo Research) and eluted in 50 μl of water. To incorporate a biotin into the DNA contacts for capture, we carried out copper-free CLICK chemistry using 1 μl of DIBO-Biotin (Thermo) added to the sample and incubating at 37°C for 1 h with shaking. Samples were purified again with DNA Clean and Concentrator columns (Zymo Research) and eluted in 10 μl of water.

Biotin capture and preparation for Illumina sequencing. DNA was quantified by Qubit (Thermo Fisher) to estimate the amount of Tn5 (Illumina) needed to generate libraries at the correct size distribution (this is assuming that contact libraries were generated properly, samples were not oversonicated, CLICK chemistry incorporated biotin efficiently and material will robustly capture on streptavidin beads). For libraries with greater than 150 ng of post-ChIRP DNA, material was set aside and a maximum of 150 ng was taken into the biotin capture step. For 7SK HiChIRP (an abundant RNA) with 15 million cells, the yield of post-ChIRP DNA was ~20 ng; for lincRNA-EPS and TERC of 90 million cells, the yield of post-ChIRP DNA was ~15 ng and ~8 ng, respectively. Streptavidin C1 beads (5 μl , Thermo Fisher) were washed with Tween wash buffer (5 mM Tris-HCl, pH 7.5, 0.5 mM EDTA, 1 M NaCl, 0.05% Tween-20) and then resuspended in 10 μl of 2 \times biotin binding buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 2 M NaCl). Beads were added to the samples and incubated at room temperature for 15 min with shaking. After capture, beads were placed on a magnet and supernatant was discarded. Samples were washed twice with 500 μl of Tween wash buffer and incubated at 55°C for 2 min with shaking. Samples were then washed in 100 μl of 1 \times TD Buffer (2 \times TD Buffer is 20 mM Tris-HCl, pH 7.5, 10 mM magnesium chloride, 20% dimethylformamide). After washes, beads were resuspended in 25 μl of 2 \times TD Buffer, Tn5 (for 10 ng of post-ChIRP DNA we used 0.5 μl of Tn5) and water to 50 μl . The Tn5 amount was adjusted linearly for different amounts of post-ChIRP DNA, with a maximum amount of 4 μl of Tn5. For example, 25 ng of DNA was transposed using 1.25 μl of Tn5, while 125 ng of DNA was transposed with 4 μl of Tn5. For low amounts, Tn5 can be diluted in 1 \times TD Buffer to help with pipetting. Using the correct amount of Tn5 is critical to the HiChIRP protocol to achieve an ideal size distribution. An overtransposed sample will have shorter fragments and will exhibit lower alignment rates (when the junction is close to a fragment end). An undertransposed sample will have fragments that are too large to cluster properly on an Illumina sequencer. A maximum amount of Tn5 is used to save on Tn5 costs, and considering that a library with this much material will be amplified in five cycles and have enough complexity to be sequenced deeply regardless of how fully transposed the library is to achieve an ideal size distribution. Samples were incubated at 55°C with interval shaking for 10 min. Beads were then placed on a magnet and supernatant was removed. EDTA (50 mM) was added to samples and incubated with interval shaking at 50°C for 30 min. Samples were then placed on a magnet and supernatant was removed. Samples were washed twice each in Tween wash buffer at 55°C for 2 min. Beads were lastly washed in 10 mM Tris before PCR amplification.

PCR and size selection. Beads were resuspended in 25 μl of Phusion HF 2 \times (New England Biosciences), 1 μl each of Nextera Ad1_noMX and Nextera Ad2.X at 12.5 μM (Supplementary Table 5) and 23 μl of water. The following PCR program was performed: 72°C for 5 min, 98°C for 1 min, then cycle at 98°C for 15 s, 63°C for 30 s and 72°C for 1 min. Cycle number was estimated using one of two different methods: (1) Reactions were first run for five cycles on a regular PCR and then removed from beads. 0.25 \times SYBR green was added and then run on a quantitative PCR (qPCR) where samples were pulled out at the beginning of exponential amplification. (2) Reactions were run on a PCR, and cycle number was estimated on the basis of the amount of material from the post-ChIRP Qubit (approximately 50 ng was run in 6 cycles, while 25 ng was run in 7, 12.5 ng was run in 8, etc.).

Size selection was performed using one of two different methods: (1) PAGE size selection of the final library. After PCR, libraries were placed on a magnet and eluted into new tubes, then purified with DNA Clean and Concentrator columns to a volume of 10 μ l. Amplified DNA was run on a 6% PAGE gel, stained with SYBR and cut to a size range of 300–700 bp. Note that if the bulk of the material is smaller, those sizes can be included, but the paired-end libraries will have a lower alignment rate. In the future, the Tn5 amount should be adjusted accordingly. (2) Two-sided size selection with the Ampure XP beads. After PCR, libraries were placed on a magnet and eluted into new tubes. Ampure XP beads (25 μ l) were added and the supernatant was kept to capture fragments less than 700 bp. Supernatant was transferred to a new tube and 15 μ l of fresh beads were added to capture fragments greater than 300 bp. After size selection, libraries were quantified with qPCR against Illumina primers and/or Bioanalyzer. Libraries were paired-end sequenced with read lengths of 75 bp.

HiChIP. HiChIP was performed on the same cross-linked cells using either the standard 3C protocol or the optimized HiChIRP (HC) 3C protocol using CTCF (Cell Signaling Technologies, 3418S) and Oct4 (Abcam, ab19857) antibodies^{8,9} using a 1:10 dilution with ChIP dilution buffer.

HiChIP for the lincRNA-EPS knockout experiments was performed as previously described using SMC1a (Bethyl, A300–055A) and H3K4me3 (Abcam, ab8580) antibodies^{8,9}.

HiChIRP RNA recovery qPCR with reverse transcription (RT–qPCR). To test recovery of the RNA of interest by HiChIRP, 1% of the precleared lysate was set aside as the input control during the HiChIRP protocol. ChIRP was performed as detailed above, and after the ChIRP washes 1% of the C1 streptavidin beads were set aside as the ChIRP test. The input and ChIRP test samples were brought up to 95 μ l with Proteinase K buffer (10 mM Tris–HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.5% SDS) and 5 μ l of Proteinase K (20 mg ml⁻¹, Thermo), and samples were incubated at 55 °C with shaking for 45 min. Samples were then boiled at 95 °C for 10 min, chilled on ice, and then RNA was Trizol extracted (Thermo Fisher, 15596026) and purified using the Zymo RNA Clean and Concentrator kit (Zymo Research, R1016). RT–qPCR was performed with Brilliant RT–qPCR mastermix (Agilent, 600825). Cycle threshold values were measured with Lightcycler 480 (Roche), and relative expression levels for the RNA of interest and negative control RNAs in the ChIRP and input samples were calculated by the double-delta cycle threshold method in comparison with a *GAPDH* control. Primer sequences are listed in Supplementary Table 4.

3C azido incorporation streptavidin dot blot. To test the incorporation efficiency of different biotin and azido nucleotides into the 3C library, 10 million cells were taken through the HiChIRP 3C protocol until incorporation. The digested and washed pellet was then resuspended in 1 \times NEBuffer2 and split into multiple tubes to test incorporation. Incorporation reactions were then set up with replicates for different nucleotides and incubated at 37 °C for 15 min. After incorporation, samples were treated with 0.5% SDS and 1 mg ml⁻¹ Proteinase K, Zymo purified, subjected to DIBO-CLICK, and then Zymo purified again as described above. DNA was then blotted on nitrocellulose using a pipette tip insert as a guide. DNA was ultraviolet cross-linked to the membrane using the 'Auto' setting of an ultraviolet Stratilinker. Membrane was then incubated with PBST containing a 1:15,000 dilution of 800CW Streptavidin IRDye (Licor), washed 3 times in PBST for 5 min with shaking and then scanned on a Licor machine.

HiChIRP 3C RNA integrity and 3C quality. HiChIRP 3C libraries were made as described above. To assess RNA integrity and 3C quality, the 3C ligated pellets were split in half. Half of the pellet was Trizol extracted and purified using the Zymo RNA Clean and Concentrator (Zymo Research, R1016). The RNA was then run on a Bioanalyzer (Agilent) to obtain RNA integrity numbers. The other half of the pellets were treated with 0.5% SDS and 1 mg ml⁻¹ Proteinase K, and then taken through the remainder of the HiChIRP library preparation protocol. Libraries were then shallowly sequenced to obtain 3C library quality information from Hi-C Pro (see below)²⁵.

Generation of TERC knockout cell lines. To achieve TERC deletion, guides were designed in the 5' and 3' regions flanking the TERC locus. TERC guides were cloned into px459 (Addgene 62988) (5' guides) or px458 (Addgene 48138) (3' guides) by restriction cloning. A 5' guide plasmid together with a 3' guide plasmid was transfected (2 μ g) into HeLa cell cultures using Lipofectamine 2000. Forty-eight hours post-transfection, 2 μ g ml⁻¹ puromycin was added to the culture medium. Forty-eight hours after antibiotic selection, cells were FACS-sorted for green fluorescent protein–positive cells into 96-well plates (1 cell per well). Clones were grown for 4 weeks and then genomic DNA was isolated using Lucigen QuickExtract (QE09050). RNA was isolated using Trizol. Clones were screened by PCR for deletion of the TERC locus.

HiChIRP and HiChIP data processing. HiChIP paired-end reads were aligned to hg19 or mm10 genomes using the HiC-Pro pipeline²⁵. Default settings were used to remove duplicate reads, assign reads to MboI restriction fragments, filter for

valid interactions and generate binned interaction matrices. HiC-Pro filtered reads were then processed into a .hic file using the hicpro2juicebox function²⁵. The Juicer pipeline HiCCUPS tool was used to identify high-confidence loops using the same parameters as for the GM12878 in situ Hi-C map: hiccup -m 500 -r 5000,10000 -f 0.1,0.1 -p 4,2 -i 7,5 -d 20000,20000 .hic_input HiCCUPS_output (refs. ^{8,26,27}). The FitHiChIP interaction caller was used as a reciprocal algorithm to confirm our HiCCUPS loops¹⁷. FitHiChIP input files included 1D peak calls by MACS2 using the FitHiChIP suggested parameters and HiC-Pro valid interactions. Peak-to-peak interactions were identified with depth normalization.

Reproducibility and RNase filtering of HiChIRP loops. We called high-confidence loops with HiCCUPS on the matrix balanced merged .hic file across replicates. We wanted to include in our dataset loops that were (1) significant in both counts and matrix balanced space, (2) reproducible in both replicates and (3) not significantly represented within the RNase control. To do this, first we took each 10-kb anchor from our loop set and then created all possible interactions greater than 30 kb and less than 2 Mb. We then randomly selected up to 50,000 of these interactions that were not in our loop set as a conservative (same anchors as our loop set) null set. Next, we compiled the counts matrix for our loop set and null set. We then took the minimum signal across replicates in each loop set to be conservative and ensure that we were not being biased by a higher signal replicate. We then fitted a local regression curve to the background signal and the log₁₀(loop width) using 'loess' in R. Next, we predicted the expected signal for each loop for our HiChIRP experiments and selected only loops that were greater in signal than the expected value, filtering less-confident loops. We did not use a significance threshold because this filtering is very conservative owing to the strong ChIRP signal at the anchors.

To filter loops that exhibited significantly higher RNase signal, we adopted a similar approach to above. First, we took each 10-kb anchor from our filtered loop set and then created all possible interactions greater than 30 kb and less than 2 Mb. We then randomly selected up to 50,000 of these interactions that were not in our filtered loop set as a conservative (same anchors as our loop set) null set. Next, we compiled the counts matrix for our filtered loop set and null set. We then took the sum of signal across replicates in each loop set to ensure that all biased loops would be identified. We then fitted a local regression curve to the background signal and the log₁₀(loop width) using 'loess' in R. Next, we predicted the expected signal for each loop for our RNase experiments and filtered loops that were significantly greater in signal (FDR < 0.1) than the expected value, by using a Poisson test in R and then adjusting the *P* values using the Benjamini–Hochberg correction. All loops called in the RNase-treated controls were then removed from the corresponding HiChIRP output.

Interaction matrices, virtual 4C and metav4C visualizations. HiChIP interaction maps were generated with Juicebox using Knight–Ruiz matrix balancing and visualized using Juicebox software at different resolutions as indicated in each analysis^{8,27}.

Virtual 4C plots were generated from dumped matrices generated with Juicebox. The Juicebox tools dump command was used to extract the chromosome of interest from the .hic file^{8,27}. The interaction profile of a specific 5-kb or 10-kb bin containing the anchor was then plotted in R. Replicate reproducibility was visualized with the mean profile shown as a line and the shading surrounding the mean representing the s.d. between replicates.

We made distance-scaled metav4C plots as previously described²⁸. Briefly, each chromosome was dumped (10 kb resolution) from the .hic file using Juicer 'dump' and read into a 'sparseMatrix' in R. There, for each loop (distance filtered 150 kb unless otherwise stated), the upstream and downstream relative components were averaged and interpolated linearly using 'approx' to get the value at each 0.1%. This was then summed for each loop and divided by the number of loops. Replicate reproducibility was visualized with the mean profile shown as a line and the shading surrounding the mean representing the s.d. between replicates.

High-confidence Juicer loop calls were loaded into the WashU Epigenome Browser to be visualized with publicly available ChIP-seq and RNA-seq data from the ENCODE Project²⁹. Browser shots from WashU track sessions were then included in virtual 4C and interaction map anecdotes.

Reproducibility scatter plots and aggregate peak analysis. To assess reproducibility for each lincRNA, reads were obtained for each of the lincRNA HiCCUPS and then the reads were converted to log₂ counts per million and a reproducibility plot was made with a Pearson correlation. Aggregate peak analysis (APA) was performed using Juicebox tools apa with default input, yielding APA at 25-kb, 10-kb and 5-kb resolution.

Differential analysis and annotation of HiCCUPS loop calls. A counts matrix was constructed on the union HiCCUPS loop calls. For pairwise comparisons, edgeR's glm model (glmQLFTest) was used and differential loops were called. Unless otherwise stated, differential loops were called with threshold of log₂ fold change of 1 and FDR of 0.1.

For annotating enriched and depleted loop sets with .bed files such as chromHMM and ChIP peaks^{29,30}, the .bed file was intersected with each anchor

and if it was present in anchor 1, anchor 2 or both, it was denoted as 1, 2 or 3, respectively. To assess whether any biases were due to random chance, the .bed file was shuffled genome-wide 100 times to generate a null annotation using ChIPseeker 'shuffle'³¹.

For ChIRP global signal at enriched and depleted loop sets, enriched and depleted anchors were overlapped with ChIRP .bam files using bedtools intersect counting. The counts within each anchor was then plotted as a distribution in the different loop sets. This same strategy was used to assess lincRNA-EPS 1D HiChIRP signal at IRG gene promoters and insulating TAD boundary regions.

Telomere and subtelomere enrichment in TERC HiChIRP. It is important to note that because MboI does not cut inside telomere repeat sequences, we are obtaining a far smaller percentage of telomere-associated TERC HiChIRP reads than we should. The only telomere-associated reads we are likely to get are those that are at the ends of the telomere, bordering the subtelomeric regions, as those are the only regions where MboI would cut and the azido-dCTP nucleotide would be incorporated. Therefore, both the telomere enrichment and the presence of intratelomere interactions would probably be far stronger if we used a restriction enzyme (or DNase) that could cut inside of telomeres.

Telomere enrichment analysis was performed similar to the TERC ChIRP study¹³. Briefly, fastq reads were searched for the telomere sequence (CCCTAA x3) and Alu sequence (GTGATCCGCCCGCCTCGGCCTCCCAAAGTG). In addition, we created a null set of 500 sequences by shuffling the nucleotides of the telomere sequence (CCCTAA x3). We then read the fastq pairs into R with the package ShortRead and then searched read1 and read2 for each sequence (Telomere, Alu and shuffled sequence) and reverse complement. To be conservative for telomere enrichment over the shuffled nucleotide set, we used the 95th percentile of the shuffled sequences for telomere enrichment. For the TERC knockout telomere enrichment, it should be noted that these experiments were performed in a different cell line, and also that the knockout cells have shorter telomere lengths.

To examine TERC HiChIRP interactions between a telomere and a non-telomeric sequence, we first identified these contacts (CCCTAA x3 in 1 read but not the other) and then made a single-ended fastq file of the non-telomeric read. We then aligned this fastq to the hg19 genome using bowtie2. We converted the .bam file to a .bed file using bedtools bamToBed and read into a genomic ranges object in R. Next, we filtered the genomic ranges by the hg19 ENCODE blacklist³². To compute the distance to chromosome ends we created a genomic ranges object of the chromosome start and end and then used distanceToNearest in R. We then computed 1-kb windows genome-wide using 'slidingWindows(chromSizes, 1000, 1000)' and then counted the overlaps of the aligned fragments. Finally, we filtered regions that did not have any counts in either replicate.

To determine the signal distribution across each chromosome, each chromosome was dumped using Juicer (10 kb resolution) and read into a sparseMatrix for each chromosome in R. Then, we determined the signal distribution by taking the average of the rowSums and colSums for each bin. To better compare multiple chromosomes, the signal by percentage was generated by approximating the value every 1% using 'approx' in R.

We wanted to see whether TERC HiChIRP could capture known translocations just as in Hi-C. To do this, we dumped all intrachromosomal interactions using Juicer 'dump' with no normalization and at 2.5-Mb resolution. We then gathered all interactions corresponding to the IGH locus at the end of chr14 (106,032,614–107,288,051, bin = 105,000,000). We then normalized both TERC and replicates by the total unique valid interactions and then averaged each interaction for both replicates. We then took the top 50 *trans* interactions to IGH and then plotted them in a circos plot using the cran package 'circlize' in R. We repeated this procedure for published Hi-C datasets to observe concordance and demonstrate that the TERC HiChIRP interactions are not technical artifacts.

lincRNA-EPS ChIRP-seq analysis. To create a bigwig track for lincRNA-EPS ChIRP-seq, both ChIRP replicates and input replicates were read into R using RSamtools. Then, the genome (mm10) was tiled into 250-bp bins using 'tile' in R. The reads overlapping each of these bins were quantified using countOverlaps and then normalized to 25 million reads. These were then converted to a run-length encoding using the coverage function where the weights were the reads overlapping each window. We wanted to be conservative with our bigwig track; thus, we took the minimum at each position using pmin(Cov^{EPSevent}, Cov^{EPSoDd}) and subtracted the mean coverage of the input control. This generated a conservative bigwig track that could then be visualized using WashU Epigenome Browser.

A step-by-step protocol is available as a Supplementary Protocol and an open resource in Protocol Exchange³³.

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

Data availability

Raw and processed data are available at NCBI Gene Expression Omnibus, accession number [GSE115524](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115524).

References

23. Chu, C. et al. *Cell* **161**, 404–416 (2015).
24. Chu, H.-P. et al. *Cell* **170**, 86–101.e16 (2017).
25. Servant, N. et al. *Genome Biol.* **16**, 259 (2015).
26. Durand, N. C. et al. *Cell Syst.* **3**, 95–98 (2016).
27. Durand, N. C. et al. *Cell Syst.* **3**, 99–101 (2016).
28. Corces, M. R. et al. *Science* **362**, eaav1898 (2018).
29. Yue, F. et al. *Nature* **515**, 355–364 (2014).
30. Ernst, J. & Kellis, M. *Nat. Methods* **9**, 215–216 (2012).
31. Yu, G., Wang, L.-G. & He, Q.-Y. *Bioinformatics* **31**, 2382–2383 (2015).
32. ENCODE Project Consortium. *Nature* **489**, 57–74 (2012).
33. Mumbach, M. R. et al. *Protocol Exchange* <https://doi.org/10.1038/protex.2019.032> (2019).

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 [Authors & Referees](#) 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 (n) 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
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- 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. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- 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 d , Pearson's r), 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

No software was used for data collection.

Data analysis

Published software for data analysis used throughout the study includes HiC-Pro, Juicer Tools, Juicebox, Juicer, and FitHiChIP. All custom software is available upon request.

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/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

Raw and processed data available at NCBI Gene Expression Omnibus, accession number GSE115524.

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

Life sciences study design

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

Sample size	Sample sizes were chosen to provide sufficient confidence to validate methodological conclusions of the applicability of HiChIRP. We performed HiChIRP on three orthogonal RNAs of varying expression levels across three different cell types with biological replicates for each.
Data exclusions	No data were excluded from analysis.
Replication	All experiments were performed with two or three biological replicates and technical duplicates. The reproducibility analysis for HiChIRP is presented in Supplementary Figure 2, and reproducibility of specific analyses is shown in the relevant sections of the main figures. All attempts at replication were successful.
Randomization	No randomization was used as 7SK and TERC experiments were performed in cell lines and not animals or human research participants. lincRNA-EPS was performed on bone-marrow derived macrophages, however randomization is not needed to validate the methodological conclusions of the applicability of HiChIRP on this lincRNA.
Blinding	No blinding was used as 7SK and TERC experiments were performed in cell lines and not animals or human research participants. lincRNA-EPS was performed on bone-marrow derived macrophages, however blinding is not needed to validate the methodological conclusions of the applicability of HiChIRP on this lincRNA.

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

Methods

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

n/a	Involvement in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	Antibodies used for HiChIP included: CTCF (Cell Signaling Technologies 2418S), Oct4 (Abcam ab19857), SMC1a (Bethyl A300-055A) and H3K4me3 (abcam ab8580).
Validation	All of these antibodies have been validated by ChIP in published research.

Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	Cell lines used in this study include mouse embryonic stem cells (Novus Biologicals NBP1-41162), GM12878 (Coriell Institute), and HeLa (ATCC).
Authentication	Where possible, cell lines were validated by comparison to published sequencing data or by in-house genotyping.
Mycoplasma contamination	Cell lines were tested for mycoplasma contamination upon receipt and periodically thereafter but not prior to each experiment. All cell lines used in this study tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

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

Laboratory animals	lincRNA-EPS KO mouse was generated by replacing the lincRNA-EPS genomic locus with a neomycin cassette under control of a
--------------------	---

Laboratory animals	Pgk1 promoter. lincRNA-EPS targeting vector was electroporated into C57BL/6 mES cells. Positive ES cells were injected into blastocytes to generate chimeric mice. lincRNA-EPS heterozygous mice were obtained by gamete line transmission from mating the chimeric mice with WT C57BL/6 mice. Wild-type and KO mice used for HiChIRP and HiChIP were female and 6-8 weeks old.
Wild animals	Study did not involve wild animals.
Field-collected samples	Study did not involve samples collected from the field.
Ethics oversight	The University of Massachusetts Medical School Animal Care and Use Committees approved all animal work.

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

ChIP-seq

Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links <i>May remain private before publication.</i>	GEO: GSE115524
Files in database submission	Fastq files, validPair interaction files, HiC matrix files have all been deposited.
Genome browser session (e.g. UCSC)	Loop calls are available in Supplementary Tables.

Methodology

Replicates	HiChIRP and HiChIP experiments performed in biological duplicate. Reproducibility is shown in Supplementary Figure 2.
Sequencing depth	Sequencing depth and read processing metrics are available in Supplementary Table 1.
Antibodies	HiChIP was performed with CTCF (Cell Signaling Technologies 2418S), Oct4 (Abcam ab19857), SMC1a (Bethyl A300-055A), and H3K4me3 (Abcam ab8580).
Peak calling parameters	Loops were called with HiCCUPS and FitHiChIP using standard parameters (details in Methods).
Data quality	HiCCUPS loops were assessed by Aggregate Peak Analysis to ensure high signal to background.
Software	HiC-Pro was used to process reads. HiC matrix files were generated using Juicebox. HiCCUPS and FitHiChIP were used to call loops on the interaction matrices.