Article

Circular ecDNA promotes accessible chromatin and high oncogene expression

https://doi.org/10.1038/s41586-019-1763-5

Received: 6 November 2018

Accepted: 26 September 2019

Published online: 20 November 2019

Sihan Wu[1,18], Kristen M. Turner[1,14,18], Nam Nguyen[2,14,18], Ramya Raviram[1], Marcella Erb[3], Jennifer Santini[3], Jens Luebeck[4], Utkrisht Rajkumar[2], Yarui Diao[1,15,16], Bin Li[1], Wenjing Zhang[1], Nathan Jameson[1], M. Ryan Corces[5], Jeffrey M. Granja[5], Xingqi Chen[5,17], Ceyda Coruh[6], Armen Abnousi[7], Jack Houston[1], Zhen Ye[1], Rong Hu[1], Miao Yu[1], Hoon Kim[8], Julie A. Law[6], Roel G. W. Verhaak[8], Ming Hu[7], Frank B. Furnari[1], Howard Y. Chang[5,9,19*], Bing Ren[1,10,11,19*], Vineet Bafna[2,19*] & Paul S. Mischel[1,12,13,19*]

Oncogenes are commonly amplified on particles of extrachromosomal DNA (ecDNA) in cancer[1,2], but our understanding of the structure of ecDNA and its effect on gene regulation is limited. Here, by integrating ultrastructural imaging, long-range optical mapping and computational analysis of whole-genome sequencing, we demonstrate the structure of circular ecDNA. Pan-cancer analyses reveal that oncogenes encoded on ecDNA are among the most highly expressed genes in the transcriptome of the tumours, linking increased copy number with high transcription levels. Quantitative assessment of the chromatin state reveals that although ecDNA is packaged into chromatin with intact domain structure, it lacks higher-order compaction that is typical of chromosomes and displays significantly enhanced chromatin accessibility. Furthermore, ecDNA is shown to have a significantly greater number of ultra-long-range interactions with active chromatin, which provides insight into how the structure of circular ecDNA affects oncogene function, and connects ecDNA biology with modern cancer genomics and epigenetics.

DNA encodes information not only in its sequence, but also in its shape. The human genome is segmented into chromosomes that are made of chromatin fibres folded into dynamic, hierarchical structures[3,4]. This spatial architecture, including numerous loops of chromatin, brings distant elements into proximity and organizes transcriptional activities into distinct compartments, restricting the accessibility of DNA to the regulatory and transcriptional machinery. In cancer, this chromatin landscape is markedly altered[5,6]. ecDNA with amplified oncogenes was recently shown to be widespread in cancer[1], complementing the diversity of non-chromosomal DNA elements[7,8]. ecDNA differs from the kilobase-size circular DNA found in healthy somatic tissues[2,7,8], because ecDNA is 100–1,000 times larger and highly amplified, raising challenging questions about ecDNA topology and how it might affect transcriptional and epigenetic regulation in cancer.

ecDNA is circular

To understand ecDNA structure, transcription and chromatin organization, we studied three human cancer cell lines (Extended Data Fig. 1a) and clinical tumour samples from The Cancer Genome Atlas (TCGA), by integrating imaging and sequencing approaches (Fig. 1a). Whole-genome sequencing (WGS) analysis has previously been used to resolve ecDNA structure, using a computational tool—AmpliconArchitect[1,9]—that classifies amplicons as circular or linear (Supplementary Table 1). Circular amplicons in GBM39 cells detected by this approach were confirmed to be extrachromosomal by fluorescence in situ hybridization (FISH) of tumour cells in metaphase (Fig. 1b, Extended Data Fig. 1b–d). The reconstructed circular amplicon structure was supported by many paired-end discordant junctional reads and validated by Sanger sequencing (Extended Data Fig. 1e, f). Genes detected on linear amplicons were found on chromosomal DNA (chrDNA) (Extended Data Fig. 1g). Reconstruction of 41 circular amplicons from 37 human cancer cell lines[1] revealed amplicon sizes ranging from 168 kb to 5 Mb, with a median size of 1.26 Mb (Extended Data Fig. 1h).

AmpliconArchitect infers a shape on the basis of computational reconstruction of short, paired-end reads (100–200 bp), but does not unambiguously place large duplications in the structure. To augment our understanding of ecDNA shape based on its sequence,

[1]Ludwig Institute for Cancer Research, University of California at San Diego, La Jolla, CA, USA. [2]Department of Computer Science and Engineering, University of California at San Diego, La Jolla, CA, USA. [3]UCSD Light Microscopy Core Facility, Department of Neurosciences, University of California at San Diego, La Jolla, CA, USA. [4]Bioinformatics & Systems Biology Graduate Program, University of California at San Diego, La Jolla, CA, USA. [5]Center for Personal Dynamic Regulomes, Stanford University, Stanford, CA, USA. [6]Plant Molecular and Cellular Biology Laboratory, Salk Institute for Biological Studies, La Jolla, CA, USA. [7]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH, USA. [8]The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA. [9]Howard Hughes Medical Institute, Stanford University, Stanford, CA, USA. [10]Department of Cellular and Molecular Medicine, Center for Epigenomics, University of California at San Diego, La Jolla, CA, USA. [11]Institute of Genomic Medicine, Moores Cancer Center, University of California at San Diego, La Jolla, CA, USA. [12]Moores Cancer Center, University of California at San Diego, La Jolla, CA, USA. [13]Department of Pathology, University of California at San Diego, La Jolla, CA, USA. [14]Present address: Boundless Bio, Inc., La Jolla, CA, USA. [15]Present address: Department of Cell Biology, Regeneration Next Initiative, Duke University School of Medicine, Durham, NC, USA. [16]Present address: Department of Orthopaedic Surgery, Regeneration Next Initiative, Duke University School of Medicine, Durham, NC, USA. [17]Present address: Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [18]These authors contributed equally: Sihan Wu, Kristen M. Turner, Nam Nguyen. [19]These authors jointly supervised this work: Howard Y. Chang, Bing Ren, Vineet Bafna, Paul S. Mischel. *e-mail: howchang@stanford.edu; biren@ucsd.edu; vbafna@cs.ucsd.edu; pmischel@ucsd.edu

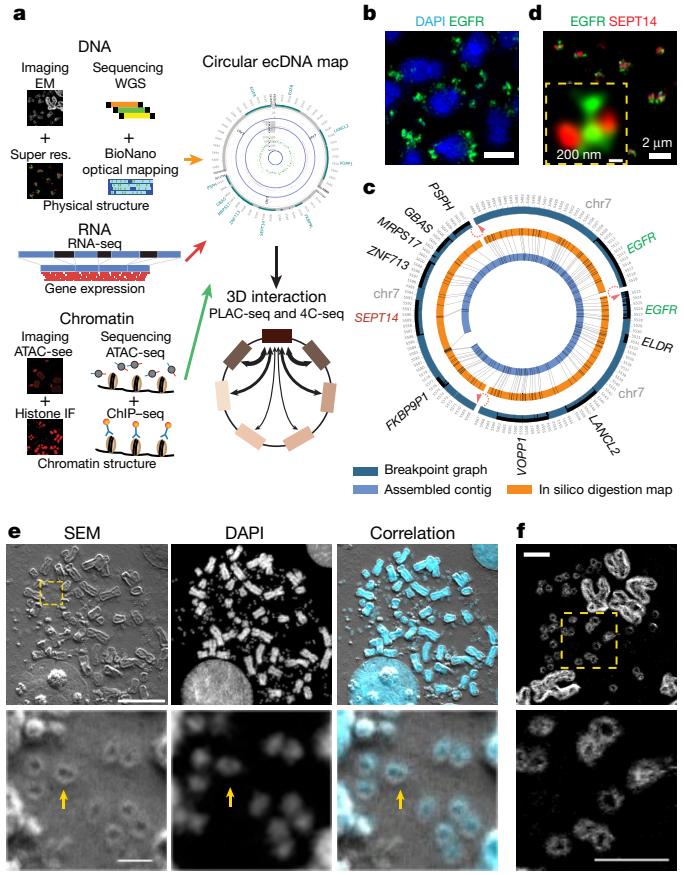

Fig. 1 | ecDNA physical structure is circular. a, Global workflow to characterize the structure and function of ecDNA. IF, immunofluorescence. **b**, Representative EGFR FISH in GBM39 cells. Scale bar, 5 μm. **c**, Composite breakpoint graph generated by AmpliconArchitect, in silico digestion map and the assembled contig from BioNano optical mapping of GBM39 ecDNA. Red arrows indicate breakpoints connected by discordant paired-end WGS reads. **d**, Double FISH of EGFR and SEPT14 identified from **c**. **e**, Correlated SEM and confocal light microscopy of chromosomal and ecDNA in COLO320DM cells. Scale bars, 10 μm (top) and 1 μm (bottom). **f**, SEM back-scatter in COLO320DM cells. Scale bars, 2 μm. All imaging experiments were repeated at least three times, with similar results.

we integrated optical mapping of long-range reads (approximately 160,000 bp) of DNA, using the BioNano technology platform, which permits the development of a physical map based on long contiguous pieces of DNA[10,11]. We developed a tool, AmpliconReconstructor, to integrate the optical mapping contigs with AmpliconArchitect-based WGS reconstructions, resolving a 1.3-Mb circular, contiguous ecDNA molecule in GBM39 cells (Fig. 1c, Extended Data Fig. 2a). Individual genes on the amplicon were visualized by super-resolution confocal microscopy (Fig. 1d, Extended Data Fig. 2b).

To visualize ecDNA architecture directly, we captured images of human COLO320DM cells containing *MYC* ecDNA (Extended Data Fig. 2c) using super-resolution three-dimensional structured illumination microscopy (3D-SIM)[12], which revealed circular ecDNA particles (Extended Data Fig. 2d). To obtain more definitive evidence, we performed scanning and transmission electron microscopy (SEM and TEM). Correlative light and electron microscopy analysis of COLO320DM cells—which contain larger-size ecDNA than GBM39 cells, making them advantageous for visualization (Extended Data Fig. 1h)—demonstrated that ecDNAs stained by the fluorescent dye 4′,6-diamidino-2-phenylindole (DAPI) are circular (Fig. 1e, f). TEM analysis of GBM39 cells independently confirmed the presence of circular ecDNAs, including classical double minutes[13,14] (Extended Data Fig. 2e). Together, these results combining DNA sequencing, optical mapping, super resolution 3D-SIM, SEM and TEM analysis demonstrate that the ecDNAs studied here are circular.

ecDNA drives massive oncogene expression

To determine the effect on transcription, we integrated RNA sequencing (RNA-seq) with WGS from cancer cell lines and from TCGA clinical tumour samples of diverse histological types, revealing that genes encoded on ecDNA—particularly bona fide oncogenes—are among the most highly expressed genes in cancer genomes (Fig. 2a, b, Extended Data Fig. 3a, b). Using our AmpliconArchitect-based approach to determine whether specific genes are amplified on circular ecDNA, we found that in cancer cell lines and clinical tumour samples, oncogenes amplified on ecDNA have markedly increased numbers of transcripts compared with the same genes when they are not amplified by circularization (Fig. 2c, d, Extended Data Fig. 3c–g). We searched for single nucleotide polymorphisms in the WGS and RNA-seq data that permitted us to distinguish between transcription from genes on ecDNA and their native chromosomal loci, revealing massively increased transcription from genes encoded on ecDNAs (Fig. 2e). In fact, oncogenes encoded on ecDNA—including *EGFR*, *MYC*, *CDK4* and *MDM2*—are among the top 1% of genes expressed in the cancer genomes (Fig. 2b, Supplementary Table 2).

The amount of RNA transcribed can be related to the amount of available DNA template. We hypothesized that the massively increased oncogene transcription on ecDNA is likely to be driven by their increased DNA copy number[15] (Extended Data Fig. 3g, h). Accordingly, oncogenes amplified on ecDNA were shown to achieve far higher copy numbers than the same genes amplified on linear structures (Fig. 2f, g). However, the amount of DNA template is not the only factor that determines gene transcription. Chromatin organization influences the accessibility of DNA to the regulatory machinery of transcription[4,16]. In some cases, oncogenes on ecDNA produced more transcripts, even when normalized to gene copy number (Extended Data Fig. 3g, h). We initiated a deeper examination of other chromatin structural features that may contribute to the massively increased expression of oncogenes amplified on ecDNA.

ecDNA contains highly accessible chromatin

Most of the human genome is not transcribed in a given cell because it is tightly wound around histone octamers that in turn are packed into complex hierarchical structures, rendering the DNA inaccessible to transcription factors and the transcription machinery[17,18]. We used complementary approaches to resolve the ecDNA chromatin landscape. First, we analysed active and repressive histone marks by immunofluorescence analysis of cancer cells in metaphase and also performed H3K4me1 and H3K27ac chromatin immunoprecipitation followed by high-throughput sequencing (ChIP–seq) analyses of actively cycling GBM39 cells, which revealed the presence of active histone marks on ecDNA[19] (Extended Data Fig. 4a–c), and a concomitant paucity of repressive histone mark on GBM39 ecDNA (Extended Data Fig. 4d, e). Second, we used the assay for transposase-accessible chromatin using sequencing (ATAC-seq) and micrococcal nuclease digestion and sequencing (MNase-seq) to assess chromatin accessibility and to map nucleosome positions. Finally, we used the assay of transposase-accessible chromatin with visualization (ATAC-see) to visualize accessible chromatin directly[20] (Extended Data Fig. 5a). The periodic length distributions of DNA fragments generated by ATAC-seq and MNase-seq demonstrated that ecDNA is packaged into chromatin, and consists of nucleosome units (Fig. 3a, Extended Data Fig. 5b, c). However, ecDNA displayed a significant deficit in the number of long fragments (more than 1,200 bp) from ATAC-seq and MNase-seq, indicative of compacted nucleosomal arrays (Fig. 3a, Extended Data Fig. 5b, c), and a significantly increased number of ATAC-seq peaks (Fig. 3b,

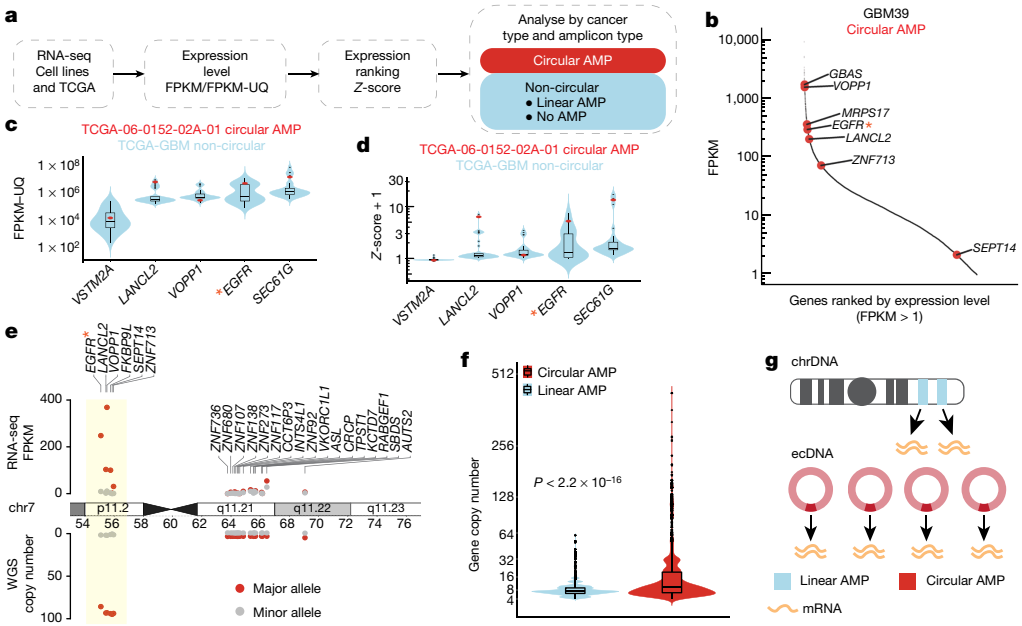

Fig. 2 | ecDNA drives high levels of RNA expression. a, Workflow of RNA-seq data analysis. FPKM, fragments per kilobase of transcript per million mapped reads; FPKM-UQ, FPKM upper quartile. **b**, ecDNA gene expression within the transcriptome of GBM39 cells. Red dots denote genes on ecDNA (circular amplification). *GBAS* is also known as *NIPSNAP2*; *SEPT14* is also known as *SEPTIN14*. **c**, ecDNA gene expression in one TCGA-GBM sample (red data points) compared to non-circular genes in the TCGA-GBM cohort (blue violin and box plot distribution) ($n = 36$ biologically independent samples). **d**, *Z*-score of the gene expression plotted in **c**. *Z*-scores were plotted as +1 to avoid negative values during $\log_{10}$ transformation. **e**, Allele-specific gene copy number and mRNA expression levels in GBM39 cells. Circular amplified region (ecDNA) was highlighted. **f**, Gene copy number comparing circular and linear amplifications (8,068 circular and 6,247 linear amplified (AMP) genes from 77 samples). *P* value determined by two-sided Wilcoxon test. **g**, Depiction of the mechanism of massive transcript levels from ecDNA. Asterisks in **b**–**e** indicate key oncogenes. Violin plots show the overall distribution of data points. Box plots show median, upper and lower quartiles; whiskers indicate 1.5 times the interquartile range, and black points are the outliers.

Extended Data Fig. 5d), which suggests that the ecDNA chromatin landscape is more accessible than chrDNA, because its nucleosomal organization is less compacted.

The recent landmark study deciphering the chromatin accessibility landscape in primary cancer samples[5] enabled us to examine chromatin accessibility in authentic clinical samples. By integrating ATAC-seq profiles with WGS data analysed by AmpliconArchitect, we found a significantly higher ATAC-seq signal in the DNA with predicted circular amplicons in clinical tumour samples, even after normalizing for DNA copy number (Fig. 3c, Extended Data Fig. 5e). Even in isogenic cell lines, ecDNA is more accessible than the same locus amplified as a homogeneous staining region (HSR)[21] on chromosomes (Extended Data Fig. 5f–h). Notably, the HSR region did not show a deficit in the number of long ATAC-seq fragments as compared to ecDNA (Extended Data Fig. 5i). We further validated that both the enhanced chromatin accessibility and active chromatin states are linked to the increased rates of transcription from the allele contained on highly amplified ecDNA (Extended Data Fig. 5j).

We then applied the ATAC-see technology to analyse accessible chromatin in actively cycling cells in interphase by staining COLO320DM cells with ATAC-see and DAPI to label accessible chromatin and DNA, respectively, and to permit the sorting of tumour cells in early G1 phase[20], followed by MYC fluorescence in situ hybridization (FISH) to label ecDNAs. A notable positive correlation between the ecDNA-containing MYC FISH signal and the ATAC-see signal was seen, which demonstrates highly accessible chromatin of ecDNA at single-cell resolution (Fig. 3d, e, Extended Data Fig. 6a–c). ecDNA remained similarly accessible during metaphase (Extended Data Fig. 7a–d). Together, these data demonstrate that some of the most accessible chromatin in the genome of cancer cells resides on ecDNA, possibly owing to the lower level of chromatin compaction (Fig. 3f). In fact, ATAC-see enabled us to identify unanticipated MYC ecDNAs in GBM39 cells because of their high signal, which was subsequently confirmed by ATAC-seq and WGS (Extended Data Fig. 7c, Supplementary Table 1).

To contextualize these genetic, transcriptional and epigenetic features, we generated circular maps of ecDNA in cancer cell lines and primary tumour samples (Fig. 4a, Extended Data Fig. 8). These topologically informed maps highlighted the high DNA copy number, high levels of transcription particularly of its constituent oncogenes, and high accessibility of its chromatin, bridging ecDNA circular structure with biological function. ecDNAs within a tumour can also vary in size and composition (that is, sequence), even when they contain the same oncogene. In GBM39 cells, the structures of EGFR-containing ecDNAs are uniform (Extended Data Fig. 9). Consequently, the WGS trace in its circular map is relatively uniform (Fig. 4a). By contrast, COLO320DM and PC3 cells contain diverse MYC-containing ecDNA populations, which results in a more heterogenous WGS trace in the circular ecDNA plots (Extended Data Figs. 8a, b and 9).

ecDNA enables ultra-long-range chromatin contacts

We performed proximity ligation-assisted ChIP–seq[22] (PLAC-seq, similar to HiChIP[23]) to map the 3D chromatin interactions genome-wide anchored at DNA bound by histone with H3K27ac modification in GBM39 cells. We also conducted circular chromosome conformation capture combined with high-throughput sequencing (4C-seq) to provide an independent assessment of chromatin contacts in GBM39 cells. Together with ChIP–seq of CTCF and cohesin subunit protein SMC3 to examine the locations of factors that are important for the organization of chromatin domains[24], these data revealed a massive increase in diagonal corner reads in the GBM39 ecDNA junctional region (Fig. 4b, Extended Data Fig. 10a), and rebound of the virtual 4C signal in the distal region (Extended Data Fig. 10b), providing further orthogonal evidence that ecDNA is circular (Fig. 4c). In addition, the binding of CTCF and cohesin demonstrate that ecDNA chromatin

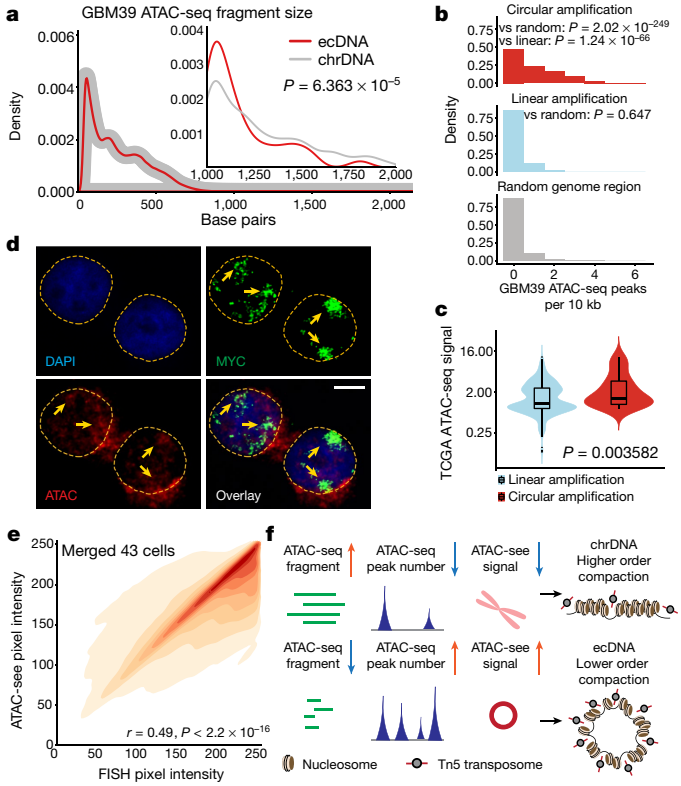

Fig. 3 | The chromatin landscape of ecDNA. a, Global and long (>1 kb) ATAC-seq fragment size distribution of ecDNA and chrDNA (110 ecDNA and 1,571 chrDNA long fragments; $n = 2$ biologically independent samples, showing one of the representative results). *P* value determined by two-sided Kolmogorov–Smirnov test. **b**, ATAC-seq peak number per 10 kb in GBM39 cells (circular, 714 windows; linear, 268 windows; random, 313,762 windows; $n = 2$ biologically independent samples). *P* values determined by Kruskal–Wallis test. **c**, TCGA ATAC-seq read counts normalized by copy number (circular: 8 samples, 33 amplicons; linear: 7 samples, 476 amplicons). Violin plots show the overall distribution of data points. Box plots as in Fig. 2g. *P* value determined by *Z*-test. **d**, Co-localization of the ATAC-see and FISH signal in interphase cell nuclei from COLO320DM cells. Scale bar, 5 µm. **e**, Pearson correlation of FISH and ATAC-see signal pixel intensity, merged from 43 COLO320DM single cells in interphase. **f**, Depiction and interpretation of integrated technologies to assess chromatin compaction.

is well organized, indicative of topologically associating domains (Fig. 4b). Furthermore, downsampling the PLAC-seq reads from the GBM39 ecDNA region to a level comparable to the same region in U87 cells that lack ecDNA demonstrated notably increased distal interactions in active chromatin on ecDNA (Fig. 4c, Extended Data Fig. 10a, b). Using the *EGFR* promoter as bait, the virtual 4C and actual 4C-seq independently demonstrated ultra-long-range chromatin contacts that can occur on ecDNA (Extended Data Fig. 10c, d), which could potentially have some effect on distal gene expression, as suggested by CRISPR interference targeting catalytically inactive Cas9 (dCas9) fused to the KRAB transcriptional repressor domain to mask the *EGFR* promoter (Extended Data Fig. 10e–j).

Amplification of oncogenes on ecDNA is surprisingly prevalent in cancer[1,25], and it can markedly increase oncogene copy number and drive intratumoural genetic heterogeneity because it lacks centromeres and is subject to unequal segregation[1,26]. These results demonstrate that ecDNA promotes massively increased transcription of the oncogenes studied here, owing to its increased DNA copy numbers and in association with enhanced chromatin accessibility, highlighting a mechanism by which ecDNA contributes to cancer pathogenesis by altering the shape of its chromatin.

In bacteria, small circular plasmids represent a prevalent and powerful mechanism for rapidly gaining selective advantage[27]. We speculate that oncogene-containing circular ecDNA in human cancers represents the conceptual equivalent, highlighting crucial gene variants and mechanisms for oncogenesis and therapeutic resistance[28–30].

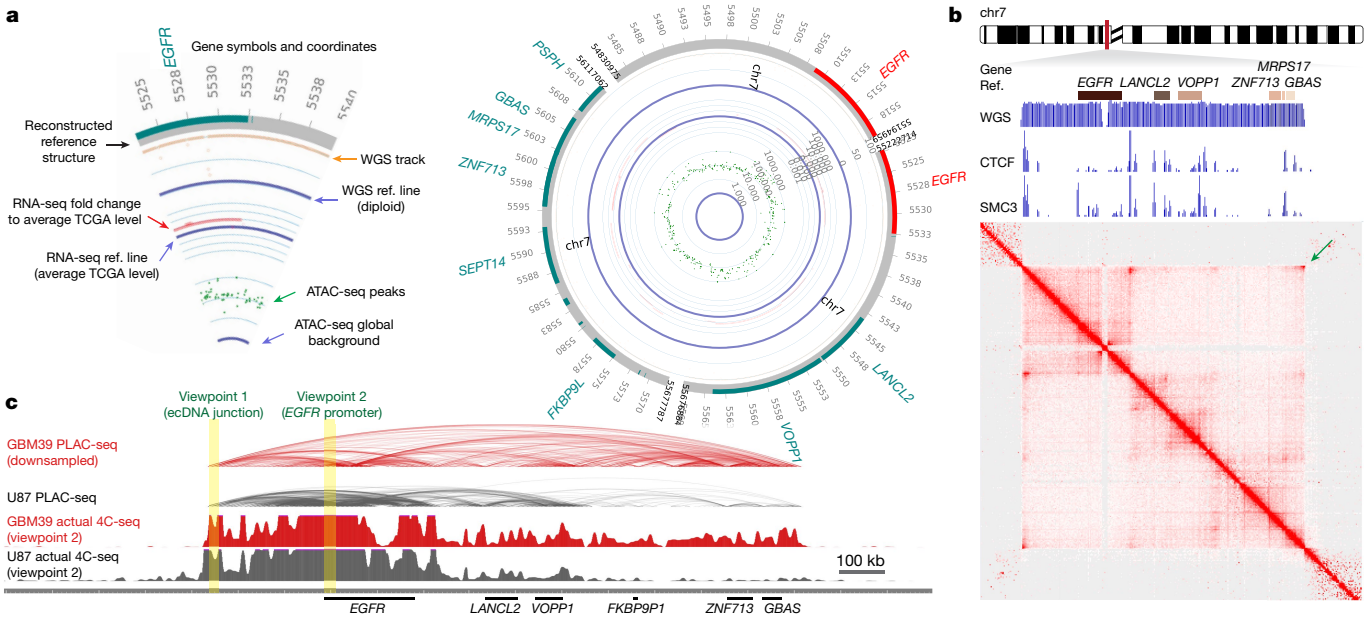

Fig. 4 | Circularization of ecDNA enables distal DNA interaction. a, Circular plot of ecDNA structure in GBM39 cells. Plot legend is shown on the left. **b**, MAGIC CollaborationH3K27ac anchored active chromatin interaction heat map by PLAC-seq and HiChIP in GBM39 cells. WGS depicts the ecDNA amplicon. ChIP–seq demonstrates CTCF and SMC3 binding to ecDNA. Arrow indicates the increased corner reads in ecDNA junction. **c**, Composite view of PLAC-seq, HiChIP and actual 4C-seq. Virtual and actual 4C-seq viewpoints are highlighted.

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/s41586-019-1763-5.

1. Turner, K. M. et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122–125 (2017).
2. Verhaak, R. G. W., Bafna, V. & Mischel, P. S. Extrachromosomal oncogene amplification in tumour pathogenesis and evolution. *Nat. Rev. Cancer* **19**, 283–288 (2019).
3. Gibcus, J. H. & Dekker, J. The hierarchy of the 3D genome. *Mol. Cell* **49**, 773–782 (2013).
4. Dixon, J. R., Gorkin, D. U. & Ren, B. Chromatin domains: the unit of chromosome organization. *Mol. Cell* **62**, 668–680 (2016).
5. Corces, M. R. et al. The chromatin accessibility landscape of primary human cancers. *Science* **362**, eaav1898 (2018).
6. Hnisz, D. et al. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
7. Møller, H. D. et al. Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. *Nat. Commun*. **9**, 1069 (2018).
8. Shibata, Y. et al. Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. *Science* **336**, 82–86 (2012).
9. Deshpande, V. et al. Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. *Nat. Commun*. **10**, 392 (2019).
10. Mendelowitz, L. & Pop, M. Computational methods for optical mapping. *Gigascience* **3**, 33 (2014).
11. Mak, A. C. et al. Genome-wide structural variation detection by genome mapping on nanochannel arrays. *Genetics* **202**, 351–362 (2016).
12. Demmerle, J. et al. Strategic and practical guidelines for successful structured illumination microscopy. *Nat. Protocols* **12**, 988–1010 (2017).
13. Schimke, R. T. Gene amplification in cultured animal cells. *Cell* **37**, 705–713 (1984).
14. Storlazzi, C. T. et al. Gene amplification as double minutes or homogeneously staining regions in solid tumors: origin and structure. *Genome Res*. **20**, 1198–1206 (2010).
15. L'Abbate, A. et al. MYC-containing amplicons in acute myeloid leukemia: genomic structures, evolution, and transcriptional consequences. *Leukemia* **32**, 2152–2166 (2018).
16. Baylin, S. B. & Jones, P. A. Epigenetic determinants of cancer. *Cold Spring Harb. Perspect. Biol*. **8**, a019505 (2016).
17. Lee, D. Y., Hayes, J. J., Pruss, D. & Wolffe, A. P. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**, 73–84 (1993).
18. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251–260 (1997).
19. Smith, G. et al. c-Myc-induced extrachromosomal elements carry active chromatin. *Neoplasia* **5**, 110–120 (2003).
20. Chen, X. et al. ATAC-see reveals the accessible genome by transposase-mediated imaging and sequencing. *Nat. Methods* **13**, 1013–1020 (2016).
21. Solovei, I. et al. Topology of double minutes (dmins) and homogeneously staining regions (HSRs) in nuclei of human neuroblastoma cell lines. *Genes Chromosom. Cancer* **29**, 297–308 (2000).
22. Fang, R. et al. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res*. **26**, 1345–1348 (2016).
23. Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
24. Rowley, M. J. & Corces, V. G. Organizational principles of 3D genome architecture. *Nat. Rev. Genet*. **19**, 789–800 (2018).
25. Bailey, M. H. et al. Comprehensive characterization of cancer driver genes and mutations. *Cell* **173**, 371–385.e318 (2018).
26. deCarvalho, A. C. et al. Discordant inheritance of chromosomal and extrachromosomal DNA elements contributes to dynamic disease evolution in glioblastoma. *Nat. Genet*. **50**, 708–717 (2018).
27. Lederberg, J. Cell genetics and hereditary symbiosis. *Physiol. Rev*. **32**, 403–430 (1952).
28. Nathanson, D. A. et al. Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* **343**, 72–76 (2014).
29. McGranahan, N. & Swanton, C. Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* **168**, 613–628 (2017).
30. Xu, K. et al. Structure and evolution of double minutes in diagnosis and relapse brain tumors. *Acta Neuropathol*. **137**, 123–137 (2019).