

Enhancing transcription–replication conflict targets ecDNA-positive cancers

<https://doi.org/10.1038/s41586-024-07802-5>

Received: 15 November 2023

Accepted: 9 July 2024

Published online: 6 November 2024

Open access

 Check for updates

Jun Tang^{1,2,13}, Natasha E. Weiser^{1,3,13}, Guiping Wang^{3,4,13}, Sudhir Chowdhry⁵, Ellis J. Curtis^{1,2,6}, Yanding Zhao^{3,4,7}, Ivy Tsz-Lo Wong^{1,2}, Georgi K. Marinov⁴, Rui Li³, Philip Hanoian⁸, Edison Tse⁵, Salvador Garcia Mojica⁵, Ryan Hansen⁵, Joshua Plum⁵, Auzon Steffy⁵, Snezana Milutinovic⁵, S. Todd Meyer⁵, Jens Luebeck⁹, Yanbo Wang^{1,2,3}, Shu Zhang^{1,2,3}, Nicolas Altemose⁴, Christina Curtis^{4,10,11}, William J. Greenleaf⁴, Vineet Bafna⁹, Stephen J. Benkovic⁸, Anthony B. Pinkerton⁵, Shailaja Kasibhatla⁵, Christian A. Hassig^{5,14}✉, Paul S. Mischel^{1,2,14}✉ & Howard Y. Chang^{3,4,7,12,14}✉

Extrachromosomal DNA (ecDNA) presents a major challenge for cancer patients. ecDNA renders tumours treatment resistant by facilitating massive oncogene transcription and rapid genome evolution, contributing to poor patient survival^{1–7}. At present, there are no ecDNA-specific treatments. Here we show that enhancing transcription–replication conflict enables targeted elimination of ecDNA-containing cancers. Stepwise analyses of ecDNA transcription reveal pervasive RNA transcription and associated single-stranded DNA, leading to excessive transcription–replication conflicts and replication stress compared with chromosomal loci. Nucleotide incorporation on ecDNA is markedly slower, and replication stress is significantly higher in ecDNA-containing tumours regardless of cancer type or oncogene cargo. pRPA2-S33, a mediator of DNA damage repair that binds single-stranded DNA, shows elevated localization on ecDNA in a transcription-dependent manner, along with increased DNA double strand breaks, and activation of the S-phase checkpoint kinase, CHK1. Genetic or pharmacological CHK1 inhibition causes extensive and preferential tumour cell death in ecDNA-containing tumours. We advance a highly selective, potent and bioavailable oral CHK1 inhibitor, BBI-2779, that preferentially kills ecDNA-containing tumour cells. In a gastric cancer model containing *FGFR2* amplified on ecDNA, BBI-2779 suppresses tumour growth and prevents ecDNA-mediated acquired resistance to the pan-FGFR inhibitor infigratinib, resulting in potent and sustained tumour regression in mice. Transcription–replication conflict emerges as a target for ecDNA-directed therapy, exploiting a synthetic lethality of excess to treat cancer.

Extrachromosomal DNAs (ecDNAs) are a frequent mechanism for oncogene amplification in diverse cancer types and are associated with worse patient outcomes than other kinds of focal amplification^{1,2}. ecDNAs can arise during the transition to, development and progression of cancers, and they exhibit unique biological features that provide fitness advantages to malignant cells³. The acentric structure of ecDNA facilitates random segregation, highly elevated copy number, intratumoural genetic heterogeneity and rapid tumour evolution^{4,5}, contributing to aggressive tumour growth and therapeutic resistance^{6,7}. The circular topology of ecDNAs also profoundly alters transcription^{8,9}. ecDNAs exhibit highly accessible chromatin and increased oncogene expression

compared to non-circular amplifications, even after controlling for DNA copy number^{10–12}. Further, ecDNAs can cluster in the nucleus to generate new, functional enhancer–promoter interactions both in *cis* and in *trans*^{10,12}. Earlier studies showed that ecDNAs highly transcribe annotated protein-coding genes¹¹, but it is unclear whether the full landscape of RNA transcription—such as intergenic, antisense or other long non-coding RNAs—is altered. ecDNA exhibits open chromatin and is marked by active histone modifications such as H3K27ac and H3K4me3 (refs. 1,11,13,14), raising the possibility of a more permissive transcriptional environment. We hypothesized that the highly accessible chromatin of ecDNA could generate a therapeutically exploitable vulnerability.

¹Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA. ²Sarafan ChEM-H, Stanford University, Stanford, CA, USA. ³Center for Personal Dynamic Regulomes, Stanford University, Stanford, CA, USA. ⁴Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. ⁵Boundless Bio, San Diego, CA, USA. ⁶Medical Scientist Training Program, University of California, San Diego, La Jolla, CA, USA. ⁷Department of Dermatology, Stanford University School of Medicine, Stanford, CA, USA. ⁸Department of Chemistry, Pennsylvania State University, University Park, PA, USA. ⁹Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, USA. ¹⁰Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA. ¹¹Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA. ¹²Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA. ¹³These authors contributed equally: Jun Tang, Natasha E. Weiser, Guiping Wang. ¹⁴These authors jointly supervised this work: Christian A. Hassig, Paul S. Mischel, Howard Y. Chang. ✉e-mail: chassig@boundlessbio.com; pmischel@stanford.edu; howchang@stanford.edu

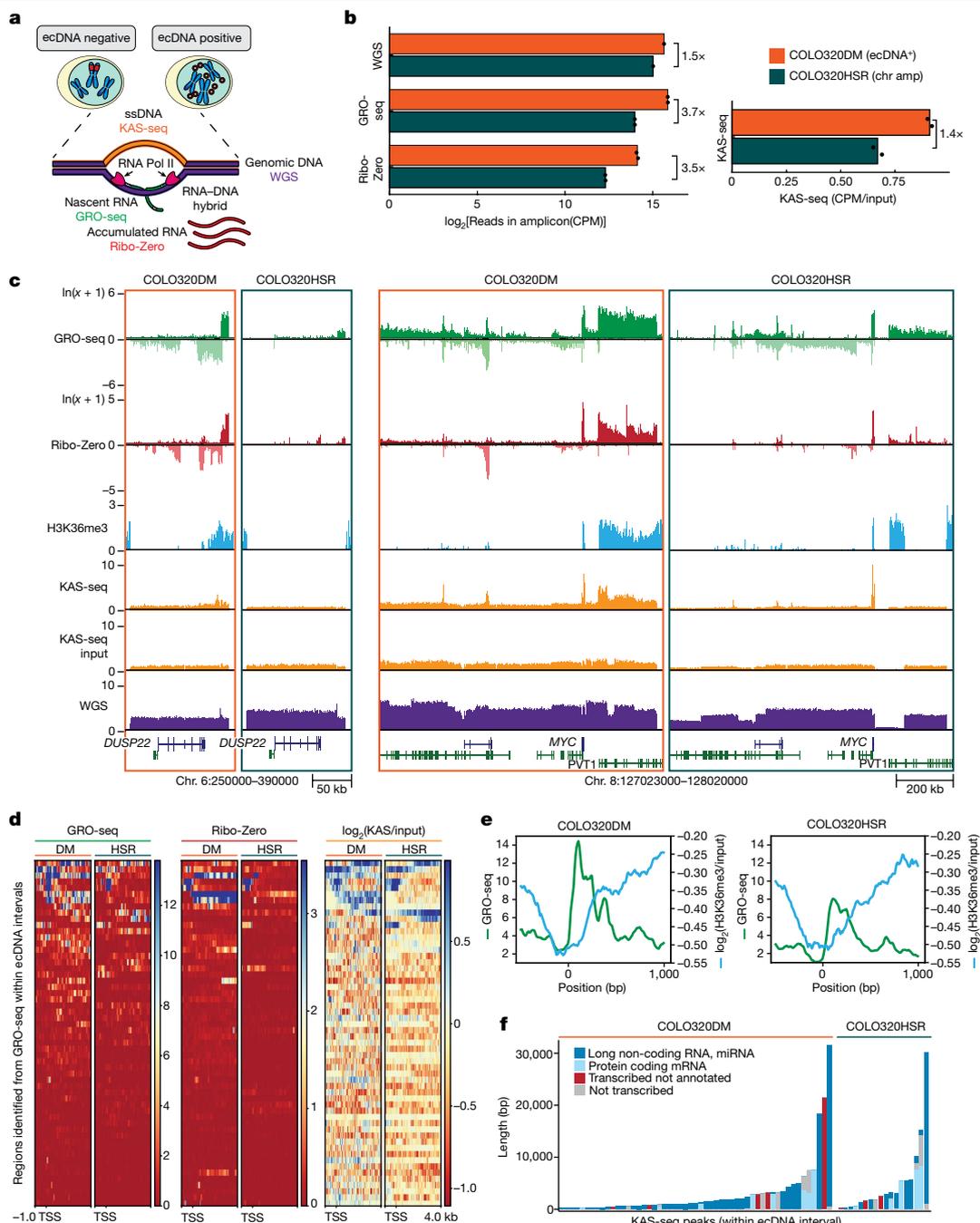


Fig. 1 | Pervasive transcription on ecDNA drives ssDNA accumulation.

a, Schematic of relevant genomic assays. **b**, Read density of genomic assays in COLO320DM and COLO320HSR in total counts per million (CPM) within the ecDNA intervals (amplicon boundaries defined in Extended Data Fig. 1). KAS-seq read density is shown as CPM of the KAS-seq relative to CPM of the input of total DNA after fragmentation but before biotin enrichment for ssDNA signals. The mean of two biological replicates is shown for GRO-seq, Ribo-Zero and KAS-seq; a single replicate is shown for WGS. **c**, Genome tracks highlighting two regions within the ecDNA interval. H3K36me3 chromatin immunoprecipitation followed by sequencing (ChIP-seq) is displayed as \log_2 of input-normalized coverage. **d**, Metagene heatmap plot visualization of GRO-seq, Ribo-Zero RNA sequencing (RNA-seq) and \log_2 of input-normalized coverage of KAS-seq within

the ecDNA interval. All plots are anchored at the transcription start site (TSS) of combined transcribed regions as identified by HOMER using both biological replicates of GRO-seq in COLO320DM and COLO320HSR. **e**, Metagene plot showing GRO-seq and H3K36me3 ChIP-seq coverage within the ecDNA interval. All plots are anchored at the GRO-seq TSS as identified by HOMER using both biological replicates. H3K36me3 ChIP-seq coverage is displayed as \log_2 (H3K36me3/input). **f**, KAS-seq peaks from two biological replicates in the ecDNA interval annotated by transcription status according to GRO-seq data and annotation status according to Gencode v.43. One representative biological replicate for each condition is visualized for **c**, **d** and **e**. chr amp, chromosomal amplification; RNA PolII, RNA polymerase II.

Rampant transcription on ecDNA

To test this hypothesis, we performed global run-on sequencing (GRO-seq)¹⁵ and ribosomal RNA (rRNA)-depleted RNA sequencing

(Ribo-Zero) to profile nascent transcription and accumulated RNAs, respectively (Fig. 1a), providing a comprehensive landscape of RNA biogenesis from ecDNAs. To control for the effects of focal amplification and assess ecDNA-specific transcriptional changes, we focused

on a pair of isogenic colorectal cancer cell lines derived from the same patient: COLO320DM (*MYC* amplification on ecDNA, also called double minute (DM)) and COLO320HSR (chromosomal *MYC* amplification on homogeneously staining region (HSR)), which are nearly matched for amplicon copy number as revealed by whole-genome sequencing (WGS) (Extended Data Fig. 1a,d)¹¹. Notably, COLO320DM showed a nearly 4-fold increase in nascent RNA and accumulated RNA read density from ecDNA, beyond the level expected from differences in amplicon copy number compared to COLO320HSR (Fig. 1b).

The increase in transcription was not limited to the *MYC* oncogene but was pervasive across the entire ecDNA, including non-coding, antisense and numerous previously unannotated transcripts (Fig. 1c,d). This widespread increase in transcription is specific to the ecDNA, as GRO-seq and Ribo-Zero read densities on chromosomes were comparable between COLO320DM and COLO320HSR (Extended Data Fig. 2). We performed de novo transcript identification within the amplicon intervals using GRO-seq data and compared the same regions in COLO320DM versus COLO320HSR. We observed increases in both nascent and accumulated transcripts in COLO320DM compared to COLO320HSR, confirming that the increased transcription from ecDNA is amplicon-wide and not driven by a small number of differentially expressed transcripts (Fig. 1d). ecDNA-transcribed regions, including those not previously annotated, were also marked by the H3K36me3 histone mark, which is associated with RNA polymerase II elongation, providing orthogonal validation of rampant transcription (Fig. 1c,e and Extended Data Fig. 3a).

Elevated transcription is associated with single-stranded DNA (ssDNA) accumulation, due to the process of transcription itself, R loop formation from RNA:DNA hybrids and transcription–replication conflict¹⁶. To assess the influence of pervasive transcription on ecDNA structure, we performed kethoxal-assisted ssDNA sequencing (KAS-seq)^{17,18} to map ssDNA genome-wide. After normalizing to input to account for copy number differences, we observed a 1.4-fold increase in KAS-seq read density within the ecDNA amplicon in COLO320DM compared to COLO320HSR (Fig. 1b,c). The ssDNA regions on ecDNA extend from hundreds to over 20,000 basepairs (bp), and the majority of KAS-seq peaks overlap with transcribed regions, such as annotated non-coding transcripts (long non-coding RNAs, microRNAs, 60%) and novel transcripts identified in GRO-seq (18%; Fig. 1d,f and Extended Data Fig. 3b). Taken together, these results suggest that ecDNAs provide a permissive chromatin environment for pervasive transcription initiation, leading to accumulated RNA species and ssDNA.

Transcription-driven RS on ecDNA

Pervasive transcription on ecDNA increases the possibility of transcription–replication conflict. When RNA polymerase II collides with the DNA replication machinery, progression of the replication fork is stalled, incorporation of new nucleotides is slowed, ssDNA behind the replication fork is exposed and bound by phosphorylated RPA2 protein (pRPA2-S33) and the cell experiences replication stress (RS)¹⁹ (Fig. 2a). This hypothesis predicts that ecDNA-containing cancer cells should have elevated DNA RS, and that the RS will be relieved by limiting transcription. First examining ecDNA-containing primary tumours, we grouped tumours from The Cancer Genome Atlas (TCGA) tumour patients into ecDNA-positive versus ecDNA-negative cohorts based on WGS data analysed by AmpliconArchitect¹. We computed the RS score through two gene expression signatures identified in ref. 20 (RS score 1) and ref. 21 (RS score 2), and found a significantly higher RS score in ecDNA-containing tumours using both methods (Fig. 2b and Extended Data Fig. 4a). This result indicates that increased RS may be a common feature shared by ecDNA⁺ cancers. Next, conflicts between transcriptional and replicative machinery should lead to slower replication fork progression. We combined a DNA-fibre assay with DNA fluorescence in situ hybridization (FISH) to analyse

replication fork dynamics in *MYC*-amplified isogenic COLO320DM versus COLO320HSR cells. Nascent DNA synthesis was labelled by sequential incubations with thymidine analogues IdU and CldU. The velocity of the replication fork was then calculated by the length of each IdU/CldU track. We observed a slower replication fork progression rate in COLO320DM compared with COLO320HSR cells; importantly, double labelling of thymidine analogue incorporation and *MYC* DNA FISH showed that ecDNA had significantly slower replication fork progression compared to the same sequence on the chromosome (Fig. 2c).

To directly visualize RS in individual tumour cells and to determine its subnuclear localization, we used immunofluorescence (IF) to detect RPA2 protein phosphorylation on serine 33 (pRPA2-S33), a marker of RS. We analysed pRPA2-S33 in a panel of cell lines, including three near-isogenic cell line pairs: COLO320DM/COLO320HSR (*MYC*-amplified colorectal cancer), GBM39ec/GBM39HSR (ref. 2) (*EGFR*-amplified glioblastoma) and PC3-DM/PC3-HSR (*MYC*-amplified prostate cancer) (Extended Data Fig. 1a–h), along with several other cell lines with or without ecDNA. Within each isogenic cell line pair, the amplified oncogene is shared but differs in its location on ecDNA or on a chromosome/HSR. We detected 2- to 3-fold higher pRPA2-S33 foci in ecDNA⁺ compared with ecDNA⁻ tumour cells, indicating increased RS in ecDNA-containing tumour cells (Extended Data Fig. 4b).

To determine whether RS is preferentially elevated on ecDNA, we performed concurrent DNA FISH to detect the ecDNA-amplified oncogene (*EGFR*) and IF to detect RS (pRPA2-S33) in GBM39ec cells. We added EdU labelling to detect actively replicating cells. We also examined these features in the isogenic counterpart, GBM39HSR, in which amplified *EGFR* has a similar copy number on chromosomes (Extended Data Fig. 1d)²¹. As hypothesized, we detected significantly higher RS on ecDNA in GBM39ec tumour cells, as measured by colocalization of pRPA2-S33 and *EGFR* FISH signal compared to GBM39HSR tumour cells, especially in EdU-positive cells. Notably, in pixels with increasing pRPA2-S33 intensity, a higher colocalization ratio was observed with a total of more than a 3-fold higher ratio on ecDNA as opposed to HSR, which suggests specific molecular interactions rather than just spatial organization differences between oncogenes on ecDNA and HSR (Fig. 2e). We continued to observe an increased pRPA2-S33 signal on *EGFR* in ecDNA⁺ cells after accounting for the total *EGFR* FISH signal, confirming that the higher RS on ecDNA compared to chromosomal amplification is not driven by differences in copy number (Extended Data Fig. 4c). To establish whether ecDNAs experience higher RS than the rest of the genome, we quantified the percentage of total nuclear pRPA2-S33 signal colocalized with *EGFR* in GBM39ec and GBM39HSR tumour cells. Based on WGS, ecDNA accounts for approximately 2% of the genomic content of GBM39ec cells. Therefore, if ecDNAs experienced a comparable level of RS compared to the rest of the genome, we would expect that they would account for a similar proportion of the total nuclear pRPA2-S33 signal. However, we found that a median of 14.5% of the total nuclear pRPA2-S33 signal is found on ecDNA, a 7-fold enrichment. In contrast, the proportion of pRPA2-S33 colocalized with *EGFR* is comparable to the relative genomic content of the amplicon in GBM39HSR (Fig. 2e). These findings indicate that RS is preferentially increased on ecDNA compared to the rest of the genome. Moreover, pRPA2-S33 IF combined with DNA FISH staining in two other near-isogenic cell line pairs containing *MYC* amplifications, COLO320 and PC3, also showed higher RS on ecDNA compared with HSR (Fig. 2d and Extended Data Fig. 4c–e). To confirm that ecDNAs are, in fact, drivers of RS, we binned individual cells by ecDNA copy number based on oncogene FISH intensity and compared the intensity of pRPA2S33 staining in cells with the highest 30%, lowest 30% and middle 40% of oncogene copy number in COLO320DM, GBM39ec and PC3-DM cells. We found that in all three cell lines, cells with the top 30% of ecDNA content have significantly higher RS than those with the bottom 30% of ecDNA content (Fig. 2f). Our results across multiple cancer cell types

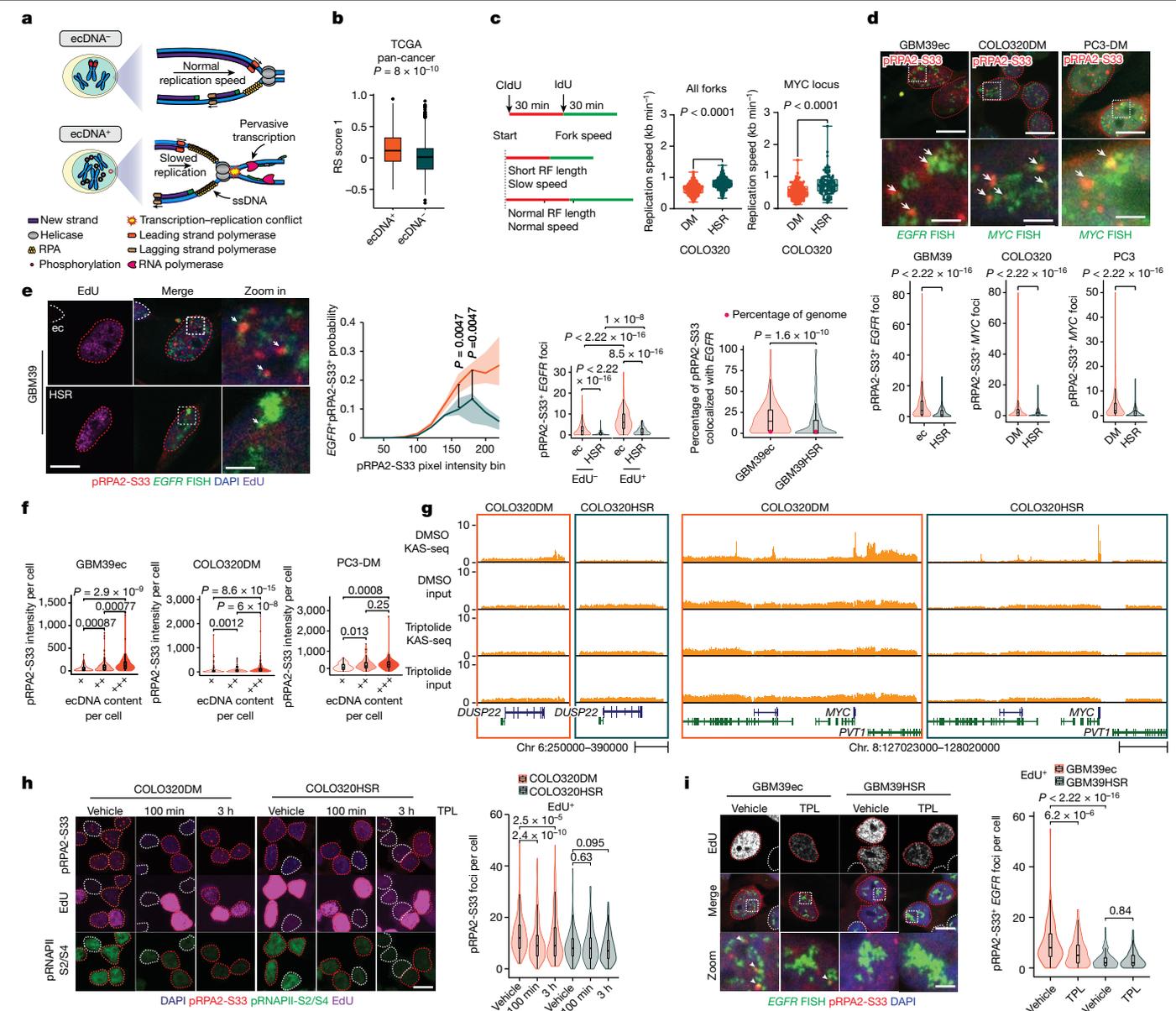


Fig. 2 | Transcription–replication conflict creates RS on ecDNAs.

a, Schematics depicting transcription–replication conflict and RS. **b**, RS score 1 computed in TCGA patients grouped by ecDNA amplification status (n : 232, 582). **c**, DNA-fibre assay combined with *MYC* FISH in COLO320DM and COLO320HSR cells. Replication fork (RF) progression rate was measured globally (middle) or at the *MYC* locus (right) (box whiskers indicate min. to max.; n : 348, 317, 143, 101). **d**, Replication protein A phosphorylation: pRPA2-S33 IF combined with *EGFR* or *MYC* DNA FISH to show higher RS on ecDNA (n : 370, 274, 939, 568, 209, 244). **e**, pRPA2-S33 IF combined with *EGFR* FISH with 5-ethynyl-2'-deoxyuridine (EdU) added for 30 min; nuclei were co-stained by DAPI. Left, representative images. Second left, proportion of pixels with colocalization within each pRPA2-S33 pixel intensity bin (shade indicates median \pm 25% quantile range, n : 10, 6). Second right, colocalization foci number (n : 267, 194, 104, 75). Right, percentage of pRPA2-S33 colocalized with *EGFR* (n : 371, 269). Red dot indicates

percent of genome taken up by amplicon as calculated by WGS counts. **f**, Comparison of RS in tumour cells with different ecDNA content grouped by total DNA FISH intensity (GBM39ec, n : 111, 148, 111; COLO320DM, n : 282, 375, 282; PC3: 63, 83, 63). **g**, Genome tracks highlighting two regions with in the ecDNA interval in COLO320 cells treated with triptolide or vehicle. **h**, **i**, Triptolide (TPL) treatment decreased RS on ecDNA in COLO320 (**h**) and GBM39 (**i**) cells (COLO320DM/HSR cells, n : 354, 350, 269, 130, 185, 161; GBM39ec/HSR cells, n : 139, 191, 264, 222). Boxplots **b–i** indicate centre line, median; limits, 25–75 quartiles; whiskers, 1.5 \times interquartile range or as otherwise specified. **d–f**, **h–i** were presented as violin plot and boxplot. Violin plot outlines kernel probability density. P determined by two-sided Wilcoxon test except unpaired Kolmogorov–Smirnov test in Fig. 2c. Scale bar, 10 μ m (**d** (top row), **e** (left), **h**, **i** (top)), 2 μ m (**d** (bottom row), **e** (right), **i** (bottom)), 50 kb (**g** (left)), 200 kb (**g** (right))).

agnostic to the identity of the amplified oncogene collectively suggest that higher RS is a common feature of ecDNAs (Extended Data Fig. 4b–e).

Having shown that ecDNAs have more open chromatin¹¹, increased transcription and elevated RS, we set out to determine whether the elevated RS on ecDNA is a direct and potentially actionable consequence of pervasive transcription generated by ecDNA's topology. We treated COLO320DM and COLO320HSR cells with triptolide, which

inhibits transcription initiation through binding to the XPB subunit of the transcription factor complex TFIIH²². Active RNA polymerase II detected by IF showed that triptolide treatment significantly decreased transcriptional activity (Extended Data Fig. 5a). KAS-seq analysis in COLO320DM and COLO320HSR cells treated with triptolide revealed drastic reduction in ssDNA signals across the ecDNA amplicon (Fig. 2g). We found that triptolide treatment significantly decreased pRPA2-S33

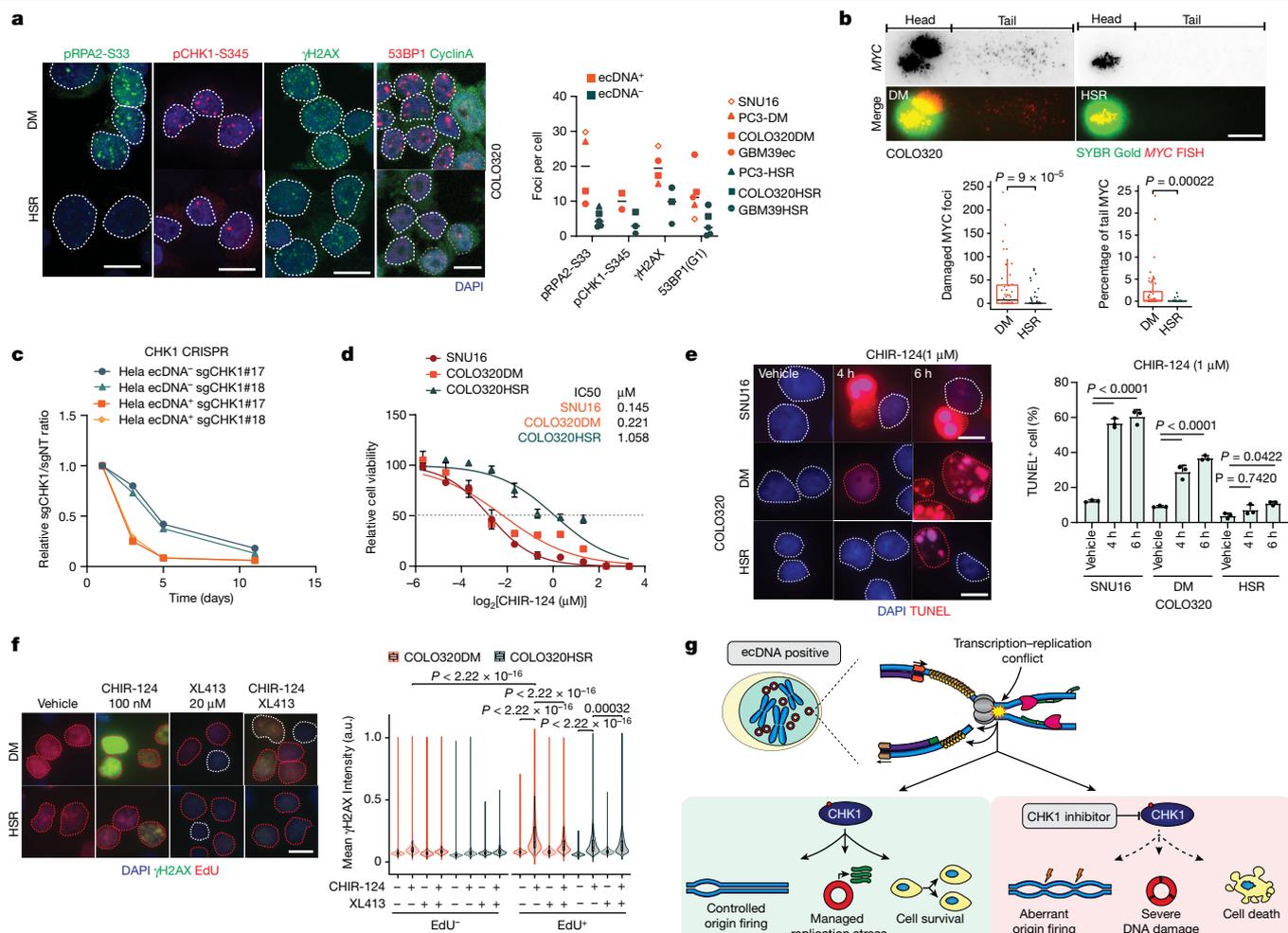


Fig. 3 | RS activated S-phase checkpoint and generated vulnerability to CHK1 inhibition in ecDNA-containing tumour cells. **a**, Detection of pRPA2-S33, γ H2AX, pCHK1-S345 and 53BP1/cyclin A in multiple cancer cell lines with different ecDNA amplification status. Left, representative images in COLO320DM and COLO320HSR cells. Right, mean foci number in individual cell lines. Line indicates median; every dot indicates mean foci number in each cell line. **b**, Comet-FISH assay in COLO320DM and COLO320HSR cells. Top, representative images. Bottom left, MYC foci number in tail. Bottom right, percentage of MYC in comet tail (two-sided Wilcoxon test, n : 47, 60, 49, 33). **c**, Relative cell number of HeLa ecDNA⁺ and HeLa ecDNA⁻ cells transduced with sgRNAs targeting CHK1 normalized to cells transduced with non-targeted (NT) sgRNA over time. **d**, Cell viability curves of SNU16, COLO320DM and COLO320HSR in response to CHIR-124 for three days ($n = 4$, mean \pm s.d.).

foci in COLO320DM cells, with negligible effect in COLO320HSR cells (Fig. 2h), suggesting that transcription contributes to the elevated RS in COLO320DM cells. In the GBM39 isogenic model, although amplicon-wide nascent transcription is similar between ecDNA and HSR cells, we observed specific regions that are induced in GBM39ec compared to GBM39HSR, including the intragenic antisense transcript *EGFR-AS1* within the *EGFR* oncogene locus, resulting in convergent transcription (Extended Data Fig. 6a–c). Triptolide treatment of GBM39ec cells significantly decreased RS on ecDNA as detected by combined pRPA2-S33 IF and *EGFR* FISH, whereas no obvious difference was observed on HSR (Fig. 2i and Extended Data Fig. 5b). Furthermore, triptolide treatment only reduced the pRPA2S33 signal in actively replicating Edu⁺ GBM39ec cells (Extended Data Fig. 5b). Taken together, our results demonstrate that ecDNAs exhibit higher levels of RS than chromosomal loci, and that this increased RS is driven in large part by concurrent transcription and replication (Extended Data Fig. 5c,d).

e, TUNEL assay in cells subjected to CHIR-124 for indicated time (mean \pm s.d., ordinary one-way ANOVA with multiple comparison test, $n = 3$). **f**, γ H2AX IF in COLO320DM and COLO320HSR cells treated with CHIR-124 with or without the combination of CDC7i (XL413), with Edu added for 30 min. Left, representative images; red lines mark Edu⁺ and white lines mark Edu⁻ nuclei. Right, mean γ H2AX intensity (arbitrary units). Edu⁻, n : 3,074, 4,246, 3,291, 4,742, 3,101, 3,770, 2,608, 2,091; Edu⁺, n : 2,428, 2,859, 2,909, 2,890, 3,346, 3,491, 3,232, 2,060; two-tailed Student's *t*-test. **g**, Schematics depicting CHK1 activation in response to RS, which sensitizes ecDNA-containing tumour cells to targeted CHK1i through unscheduled replication origin firing and accumulation of excessive DNA damage, leading to cell death. Parameters for boxplots **b**, **f** and violin plot **f** are the same as Fig. 2 or as otherwise specified. Scale bar, 10 μ m (**a**, **e**, **f**), 20 μ m (**b**). a.u., arbitrary units.

RS induces DNA damage on ecDNA

RS contributes to endogenous DNA damage because stalled replication forks are unstable and prone to breakage, generating DNA lesions²³. Therefore, we hypothesized that ecDNA-containing tumour cells may have higher baseline levels of DNA damage. We tested this hypothesis using two markers for DNA damage: γ H2AX marks all double-stranded DNA breaks and 53BP1 marks unrepaired DNA damage that arises from DNA replication during the previous cell cycle specifically in G1 daughter cells. In a panel of ecDNA⁺ and ecDNA⁻ cancer cell lines, including three near-isogenic cell line pairs, we found that in addition to having more pRPA2-S33 foci, ecDNA⁺ cells showed an average increased number of γ H2AX and 53BP1 foci than the corresponding isogenic HSR and/or other ecDNA⁻ cell lines (Fig. 3a and Extended Data Fig. 7b,c). Combined γ H2AX IF with DNA FISH staining in isogenic cell line pairs confirmed enhanced DNA damage on ecDNAs, compared to chromosomal

Table 1 | In vitro and cellular potency of BBI-2779 and of reference compounds

	BBI-2779	GDC-0575	SRA737
CHK1 biochemical potency (IC ₅₀), nM	0.3	12	85
CHK1/CHK2 selectivity	160×	1.5×	2,000×
CHK1 cellular potency AlphaLisa (IC ₅₀), nM	3	122	1,500
CTG proliferation ecDNA ⁺ (IC ₅₀), nM	6	105	1,010

AlphaLisa pCHK1-S345 activity was assessed in HT29 cells, while antiproliferation potency was evaluated in COLO320DM cells.

amplicons (Extended Data Fig. 7d,e). To further confirm the presence of DNA damage on ecDNA itself, we performed an alkaline comet assay combined with *MYC* FISH staining in COLO320DM and COLO320HSR cells, where damaged DNA appears in the tail region of the comet. We observed significantly more *MYC* foci in the tail region of COLO320DM cells compared to COLO320HSR cells (Fig. 3b), which have a comparable amplicon copy number. These data demonstrate elevated DNA damage on ecDNAs, relative to the same loci amplified on chromosomes. Thus, ecDNA-containing cancer cells may be hyperreliant on the RS regulation machinery to cope with the elevated levels of baseline DNA damage driven by transcription–replication conflicts.

ecDNA sensitizes cells to CHK1i

We reasoned that this hyperreliance on the RS regulation machinery in ecDNA-bearing tumour cells might generate an actionable therapeutic vulnerability. To cope with stalled replication forks, cells employ a signalling cascade known as the S-phase checkpoint to ensure that they do not progress to mitosis when the DNA is incompletely replicated. Checkpoint kinase 1 (CHK1), which is phosphorylated when the checkpoint is activated, is a central node for this checkpoint pathway. We detected more pCHK1-S345 by IF in ecDNA-containing tumour cells compared with the corresponding isogenic HSR cells (Fig. 3a and Extended Data Fig. 7a), indicating that transcription–replication conflict on ecDNA leads to S-phase checkpoint activation in ecDNA-containing tumour cells. In the absence of a functioning checkpoint, cells with highly damaged DNA proceed through the cell cycle, leading to cell death²⁴. We therefore hypothesized that ecDNA-containing tumour cells, due to their intrinsic heightened RS, would be hyperreliant on CHK1 to manage DNA damage, and that CHK1 inhibition (CHK1i) could trigger preferential cell death in ecDNA-containing tumour cells.

To test this hypothesis, we used clustered regularly interspaced short palindromic repeats (CRISPR) to knock out the gene encoding CHK1 in a pair of HeLa cell lines with or without *DHFR* amplification on ecDNA. Two different single guide RNAs (sgRNAs) targeting CHK1 induced 2- to 3-fold higher growth inhibition in ecDNA⁺ compared with ecDNA⁻ HeLa cells across different time points (Fig. 3c). We next inhibited CHK1 pharmacologically using CHIR-124 (ref. 25) and found that ecDNA-containing tumour cells were more sensitive to CHK1i than their corresponding isogenic HSR cells, with a half-maximal inhibitory concentration (IC₅₀) approximately 4-fold higher in COLO320HSR compared to COLO320DM cells (Fig. 3d). The susceptibility of ecDNA-containing tumour cells to CHK1 inhibition was confirmed with three structurally different CHK1 inhibitors—GDC-0575, SRA737 and CHIR-124—whereas the checkpoint kinase 2 (CHK2) inhibitor CCT241533 showed no differential inhibitory effect between ecDNA⁺ and ecDNA⁻ isogenic cell lines (Extended Data Fig. 8a–e). More importantly, suppression of cell growth by CHK1i was mediated through induction of cell death, as a more rapid and higher degree of cell apoptosis was observed in ecDNA-containing tumour cells treated with CHIR-124, as detected by terminal deoxynucleotidyl transferase dUTP nick end labelling (TUNEL; Fig. 3e) and PI-Annexin V staining (Extended Data Fig. 8f).

Table 2 | Pharmacokinetic parameters of BBI-2779 indicate that it is well tolerated in mice

Pharmacokinetic parameter	BBI-2779
In vivo cl (mlmin ⁻¹ kg ⁻¹)	229
t _{1/2} (h)	1.11
t _{max} (h)	0.5
C _{max} (ng ml ⁻¹)	713
AUC _{inf} (h ng ml ⁻¹)	1,568
%F	72

Oral bioavailability was determined in fasted male CD-1 mice dosed at 30 mg kg⁻¹ (n = 3). AUC_{inf}, area under concentration–time curve; cl, plasma clearance; C_{max}, maximum observed concentration; %F, % oral bioavailability; t_{1/2}, terminal half life; t_{max}, time maximum concentration.

As a master effector of S-phase checkpoint, CHK1 activation maintains cell viability by restricting cell cycle progression^{24,26}, limiting late replication origin firing to prevent excessive DNA damage accumulation, and protecting stalled replication forks^{27,28}. γH2AX IF combined with EdU labelling in COLO320DM and COLO320HSR cells treated with CHIR-124 showed that CHK1i induced significantly higher DNA damage in COLO320DM compared with COLO320HSR cells, especially in S-phase cells as indicated by EdU⁺ staining, consistent with the function of CHK1 in replication (Fig. 3f). Furthermore, inhibition of replication origin firing by CDC7i (XL413), indicated by the decreased EdU⁻ staining intensity (Extended Data Fig. 8g), partially blocked DNA damage induced by CHK1i (Fig. 3f), suggesting that CHK1i leads to extensive RS and DNA damage partially through unscheduled replication origin firing. Furthermore, by combining pRPA2-S33 IF with *MYC* FISH, we found that the increased sensitivity of COLO320DM cells to CHK1i was consistent, regardless of ecDNA copy number (Extended Data Fig. 8h).

Taken together, our findings demonstrate that transcription–replication conflict, RS and increased baseline DNA damage are common features of ecDNAs and drive activation of the S-phase checkpoint. Targeted CHK1i in ecDNA⁺ cells leads to unscheduled replication origin firing and accumulation of DNA damage. Furthermore, the high levels of transcription–replication conflict and RS drive a selective vulnerability to CHK1i in ecDNA⁺ cells compared to ecDNA⁻ cells, raising the possibility for an effective ecDNA-directed therapy (Fig. 3g).

Oral CHK1i stops ecDNA⁺ tumours

Despite convincing preclinical data and preliminary evidence of single-agent clinical activity for CHK1i, there are currently no approved CHK1 inhibitors for any cancer indication. Several limitations of prior CHK1 inhibitors include insufficient potency, potential off-target liabilities (for example, CHK2), and overlapping toxicity in combination with DNA-damaging chemotherapy²⁹. To further interrogate the potential of CHK1i as a treatment for ecDNA⁺ cancers, we advanced BBI-2779, an orally bioavailable, potent and selective small molecule inhibitor of CHK1 (Fig. 4a). The potency of BBI-2779 against CHK1 was confirmed in vitro using biochemical enzyme inhibition and cellular biomarker assays. The biochemical inhibition IC₅₀ of BBI-2779 against CHK1 was found to be 0.3 nM, while cellular induction of RS (as judged by pCHK1-S345, due to CHK1 phosphorylation by upstream kinases) in tumour cells was observed to be 3 nM. BBI-2779 has superior biochemical and selective cell growth inhibition compared to other orally bioavailable CHK1 inhibitors tested (IC₅₀ of ecDNA⁺ CellTiter-Glo proliferation is around 18–168-fold more potent) (Table 1). The inhibitor was observed to be greater than 160-fold selective for CHK1 over CHK2, suggestive of high pharmacological specificity (Table 1). BBI-2779 also displays excellent bioavailability (%F = 71) and good exposure in rodents, allowing for robust CHK1 target coverage after oral administration (Table 2 and Extended Data Fig. 9).

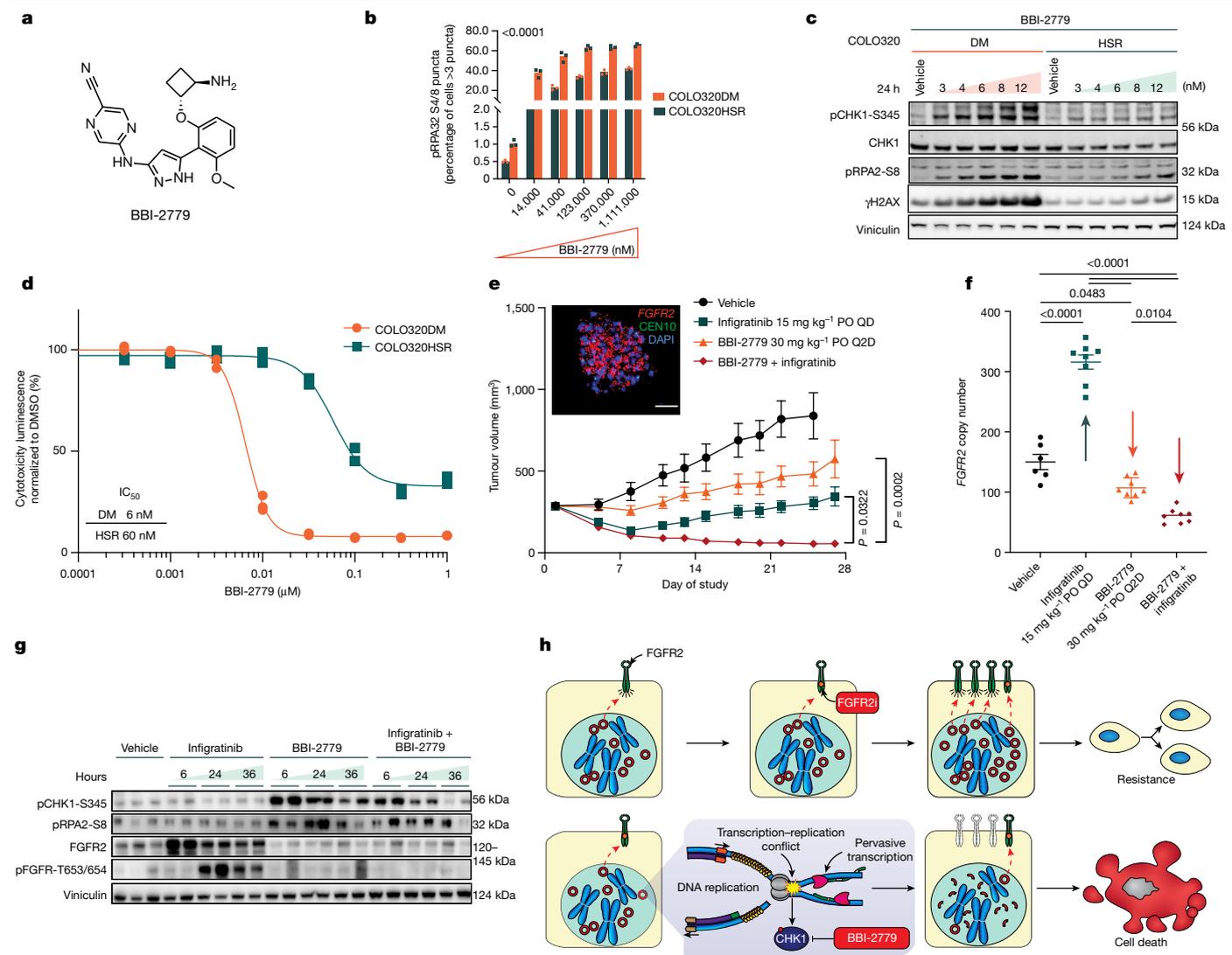


Fig. 4 | Oral CHK1i in combination with a pan-FGFRi demonstrates synergistic antitumour activity and inhibits acquired resistance to targeted therapy manifested by ecDNA. **a**, Chemical structure of BBI-2779. **b, c**, Dose-dependent induction of RS and associated biomarkers measured by phosphorylated RPA32 Ser8 level using IF (**b**) and immunoblotting (**c**). For **b**, significance determined using ordinary two-way ANOVA, $n = 3$. **d**, Differential tumour cell antiproliferation activity of BBI-2779 in COLO320DM and HSR cells ($n = 3$). **e**, Embedded FISH image of SNU16 cells demonstrating *FGFR2*⁺ ecDNA. SNU16 cells were grown as tumour xenografts in mice. After tumour establishment (approximately 285 mm³), mice were treated with vehicle, BBI-2779 (30 mg kg⁻¹), infigratinib (15 mg kg⁻¹) or BBI-2779 (30 mg kg⁻¹) plus infigratinib (15 mg kg⁻¹) for 25 days (vehicle) or 27 days (other arms). Mean tumour volumes \pm s.e.m.

are shown ($n = 8$ mice per group). **f**, *FGFR2* copy number was evaluated by quantitative polymerase chain reaction (qPCR) on tumour DNA. Significance was determined by one-way ANOVA with Tukey's multiple comparisons. **g**, Immunoblots of tumour lysates measuring elevated RS, DNA damage and abrogation of oncoprotein *FGFR2* expression ($n = 3/8$ mice per group). **h**, ecDNA-amplified oncogenes are hypertranscribed, resulting in elevated RS and reliance on CHK1 to manage DNA replication to maintain oncoprotein overexpression and proliferation. CHK1i results in uncontrolled origin firing and failed cell cycle checkpoints, exacerbating RS in ecDNA-enabled tumour cells. Synthetic lethality to CHK1i in ecDNA⁺ oncogene-amplified tumour cells is synergistic with targeted therapy resulting in enhanced cytotoxicity. Scale bar, 10 μ m. PO, oral; QD, once-daily; Q2D, every other day.

As ecDNA⁺ oncogene-amplified tumour cells harbour elevated intrinsic RS and are sensitive to other CHK1 inhibitors (Fig. 3d), we hypothesized that they would also be hypersensitive to BBI-2779. Consistent with this notion, BBI-2779 treatment of COLO320DM cells resulted in a significantly greater dose-dependent increase in the expression of the RS biomarker pRPA2-S8 compared to COLO320HSR (Fig. 4b). COLO320DM cells also showed a greater dose-dependent increase of pCHK1-S345 and γ H2AX, as determined by Western blotting (Fig. 4c). The concentration-dependent induction of RS induced by BBI-2779 directly correlated with enhanced cytotoxicity in the COLO320DM cells as compared to COLO320HSR cells, with an approximately 10-fold difference in IC₅₀ between COLO320DM and COLO320HSR cells, demonstrating synthetic lethality in the ecDNA⁺ context (Fig. 4d).

Applying targeted therapy pressure to the protein products of oncogenes amplified on ecDNA induces cancer cells to evade such pressures, either by increasing ecDNA amplification of the dominant oncodriver (Extended Data Fig. 10a–d), or by ecDNA amplification of a new bypass oncogene³⁰. We therefore investigated whether combining targeted therapy with CHK1i in ecDNA-amplified tumour cells provides a synergistic therapeutic effect resulting in cancer cell death and tumour regression. The synergistic antitumour activity and pharmacodynamics of BBI-2779 was evaluated in combination with the pan-FGFR tyrosine kinase inhibitor infigratinib in the *FGFR2*-amplified ecDNA⁺ gastric cancer SNU16 xenograft tumour model.

Single-agent BBI-2779 or infigratinib resulted in significant tumour growth delay with mean per cent tumour growth inhibition of 64% and

97% compared to the vehicle arm on day 25 ($P < 0.05$ and $P < 0.0005$, respectively) (Fig. 4e). Prolonged treatment of SNU16 tumour cells in vitro and SNU16 xenograft tumours in vivo with infigratinib resulted in tumour cell stasis for a period of 1–2 weeks, followed by acquired resistance to infigratinib and reinitiation of tumour growth concomitant with increased *FGFR2* amplifications on ecDNA (Fig. 4f and Extended Data Fig. 10a–c). The lack of robust or sustained antitumour activity observed with infigratinib alone is consistent with the absence of compelling clinical efficacy reported for pan-FGFR inhibitors in *FGFR1/2/3*-amplified settings³¹. Increased *FGFR2* gene amplification correlated with *FGFR2* protein levels that likely out-titrate the exposure of infigratinib at its maximally tolerated dose in mice (Fig. 4g and Extended Data Fig. 10d). The combination of BBI-2779 plus infigratinib resulted in significant tumour growth inhibition compared to vehicle-treated animals ($P < 0.0001$), with tumour regressions observed over the duration of the study, which was directly correlated with the suppression of further (adaptive) *FGFR2* oncogene copy number amplification on ecDNA, otherwise induced by single-agent infigratinib (Fig. 4e,f). As expected, both single-agent BBI-2779 and combination of BBI-2779 plus infigratinib treatment resulted in a heightened tumour expression of RS biomarkers pCHK1-S345 and pRPA2-S8 compared to vehicle-treated tumours (Fig. 4g). Taken together, these findings demonstrate synergistic antitumour activity by combining a selective CHK1 inhibitor with a targeted therapy against the protein product of the amplified driver oncogene to attenuate ecDNA-mediated resistance. Uncontrolled origin firing caused by selective CHK1i severely disrupts oncogene expression on hypertranscribed ecDNA templates, thereby rendering the oncogene-addicted tumour cells highly vulnerable to FGFR inhibition (Fig. 4h).

Discussion

ecDNA is a pernicious driver of tumour evolution because it is a platform for massive oncogene expression and rapid genome adaptation. Here we show that the transcriptional advantage of ecDNA can be turned on its head to selectively target ecDNA-containing tumours. The increased transcription of ecDNA is not limited to the protein-coding oncogene loci, but also extends to multiple non-coding intergenic and antisense regions throughout ecDNAs, implying violation of evolved configurations of gene directionality and replication origins in the genome. The pervasive transcription initiation is consistent with increased chromatin accessibility and promiscuous enhancer–promoter contacts on ecDNA^{9,10}. Thus, rampant ecDNA transcription comes at the cost of increased transcription–replication conflict that cancer cells must manage. DNA damage has been previously associated with ecDNA-containing cancers principally as a source of ecDNA generation⁸. Our results show that, once formed, ecDNAs themselves become a major driver of DNA damage. The RNA transcription and DNA replication machineries are two processive holoenzymes that both run along DNA; they must take turns or risk collision. Our findings demonstrate that concurrent transcription and replication on ecDNA drives a significant increase in DNA damage, and cancer cells become heavily reliant on the S-phase CHK1 to limit origin firing. The alternative for the cancer cell is to limit ecDNA transcription and lose oncogene overexpression, undermining their unique oncogenic and adaptive growth advantage. Differences in replication rate or origin firing at ecDNAs may also contribute to their increased RS, which should be addressed in future studies. Notably, the elevated levels of DNA damage on ecDNA due to RS may drive further evolution of the ecDNAs themselves. However, there are likely many mechanisms functioning in parallel that drive ecDNA mutations over time, including differential expression of some DNA damage repair pathways in ecDNA⁺ cancers³² and increased APOBEC3-mediated mutagenesis³³. Our study examined a limited number of cell line models for mechanistic studies and potential indirect effects. As oncogene encoded on ecDNA (*MYC*, *EGFR*) drives transcription and cell replication, ecDNAs

may promote transcription–replication conflicts indirectly throughout the genome. Nonetheless, both direct and indirect effects of ecDNA highlight transcription–replication conflict as a therapeutic opportunity in ecDNA⁺ cancers.

We tested the concept that enhancing transcription–replication conflict will cause ecDNA-containing tumour cells to self-destruct. Inhibition of CHK1 substantially increases ecDNA damage during DNA replication and leads to preferential killing of ecDNA-containing cancer cells. There are currently no approved CHK1 inhibitors for use in cancer patients. Despite convincing preclinical data and preliminary evidence of single-agent clinical activity for CHK1i, a predictive biomarker(s) and an optimal clinical development strategy have been lacking. Furthermore, a major challenge to the successful clinical development of CHK1 inhibitors has been the lack of reliable methods to identify high-RS tumours that are predicted to be hypersensitive to CHK1i^{34–41}. Long durability of CHK1i in vivo is likely required to exploit unscheduled DNA replication to ensure cancer cell death. The results presented here suggest a promising strategy for a next-generation CHK1 inhibitor to target ecDNA-containing cancers. Notably, CHK1i showed synergy with a targeted therapy blocking the ecDNA oncogene-encoded protein product and prevented the adaptive elevation of ecDNA copy number that previously foiled single-agent therapies targeting oncogene-amplified protein products. Previous successes in cancer therapy have exploited the synthetic lethality of cancer-specific cellular deficiencies, for example PARP inhibition in BRCA2-deficient cancer cells⁴². Our work demonstrates the feasibility of a synthetic lethality of cancer-specific cellular excess to turn the molecular advantages of ecDNA in cancer against itself.

Online content

Any methods, additional references, Nature Portfolio 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/s41586-024-07802-5>.

- Kim, H. et al. Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. *Nat. Genet.* **52**, 891–897 (2020).
- Turner, K. M. et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122–125 (2017).
- Luebeck, J. et al. Extrachromosomal DNA in the cancerous transformation of Barrett's oesophagus. *Nature* **616**, 798–805 (2023).
- Levan, A. & Levan, G. Have double minutes functioning centromeres? *Hereditas* **88**, 81–92 (1978).
- Lange, J. T. et al. The evolutionary dynamics of extrachromosomal DNA in human cancers. *Nat. Genet.* **54**, 1527–1533 (2022).
- Song, K. et al. Plasticity of extrachromosomal and intrachromosomal BRAF amplifications in overcoming targeted therapy dosage challenges. *Cancer Discov.* **12**, 1046–1069 (2022).
- Nathanson, D. A. et al. Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* **343**, 72–76 (2014).
- Hung, K. L., Mischel, P. S. & Chang, H. Y. Gene regulation on extrachromosomal DNA. *Nat. Struct. Mol. Biol.* **29**, 736–744 (2022).
- Wu, S., Bafna, V., Chang, H. Y. & Mischel, P. S. Extrachromosomal DNA: an emerging hallmark in human cancer. *Annu. Rev. Pathol.* **17**, 367–386 (2022).
- Hung, K. L. et al. ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature* **600**, 731–736 (2021).
- Wu, S. et al. Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **575**, 699–703 (2019).
- Yi, E. et al. Live-cell imaging shows uneven segregation of extrachromosomal DNA elements and transcriptionally active extrachromosomal DNA hubs in cancer. *Cancer Discov.* **12**, 468–483 (2022).
- Zhu, Y. et al. Oncogenic extrachromosomal DNA functions as mobile enhancers to globally amplify chromosomal transcription. *Cancer Cell* **39**, 694–707.e7 (2021).
- Koche, R. P. et al. Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. *Nat. Genet.* **52**, 29–34 (2020).
- Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
- Stork, C. T. et al. Co-transcriptional R-loops are the main cause of estrogen-induced DNA damage. *eLife* **5**, e17548 (2016).
- Wu, T., Lyu, R., You, Q. & He, C. Kethoxal-assisted single-stranded DNA sequencing captures global transcription dynamics and enhancer activity in situ. *Nat. Methods* **17**, 515–523 (2020).

18. Marinov, G. K. et al. CasKAS: direct profiling of genome-wide dCas9 and Cas9 specificity using ssDNA mapping. *Genome Biol.* **24**, 85 (2023).
19. Garcia-Muse, T. & Aguilera, A. Transcription–replication conflicts: how they occur and how they are resolved. *Nat. Rev. Mol. Cell Biol.* **17**, 553–563 (2016).
20. McGrail, D. J. et al. Defective replication stress response is inherently linked to the cancer stem cell phenotype. *Cell Rep.* **23**, 2095–2106 (2018).
21. Guerrero Llobet, S. et al. An mRNA expression-based signature for oncogene-induced replication-stress. *Oncogene* **41**, 1216–1224 (2022).
22. Titov, D. V. et al. XPB, a subunit of TFIIH, is a target of the natural product triptolide. *Nat. Chem. Biol.* **7**, 182–188 (2011).
23. Gaillard, H., Garcia-Muse, T. & Aguilera, A. Replication stress and cancer. *Nat. Rev. Cancer* **15**, 276–289 (2015).
24. Lecona, E. & Fernandez-Capetillo, O. Targeting ATR in cancer. *Nat. Rev. Cancer* **18**, 586–595 (2018).
25. Tse, A. N. et al. CHIR-124, a novel potent inhibitor of CHK1, potentiates the cytotoxicity of topoisomerase I poisons in vitro and in vivo. *Clin. Cancer Res.* **13**, 591–602 (2007).
26. Zhang, Y. & Hunter, T. Roles of CHK1 in cell biology and cancer therapy. *Int. J. Cancer* **134**, 1013–1023 (2014).
27. Guo, C. et al. Interaction of CHK1 with treslin negatively regulates the initiation of chromosomal DNA replication. *Mol. Cell* **57**, 492–505 (2015).
28. Técher, H. et al. Signaling from Mus81-Eme2-dependent DNA damage elicited by CHK1 deficiency modulates replication fork speed and origin usage. *Cell Rep.* **14**, 1114–1127 (2016).
29. da Costa, A. A. B. A., Chowdhury, D., Shapiro, G. I., D'Andrea, A. D. & Konstantinopoulos, P. A. Targeting replication stress in cancer therapy. *Nat. Rev. Drug Discov.* **22**, 38–58 (2023).
30. Turner, K. M. et al. Abstract 1089: extrachromosomal DNA (ecDNA)-driven switching of oncogene dependency facilitates resistance to targeted therapy. *Cancer Res.* **81**, 1089 (2021).
31. Subbiah, V. et al. FIGHT-101, a first-in-human study of potent and selective FGFR 1-3 inhibitor pemigatinib in pan-cancer patients with FGF/FGFR alterations and advanced malignancies. *Ann. Oncol.* **33**, 522–533 (2022).
32. Lin, M. S. et al. Transcriptional immune suppression and upregulation of double stranded DNA damage and repair repertoires in ecDNA-containing tumors. *eLife* **12**, RP88895 (2024).
33. Bergstrom, E. N. et al. Mapping clustered mutations in cancer reveals APOBEC3 mutagenesis of ecDNA. *Nature* **602**, 510–517 (2022).
34. Cash, T. et al. A phase 1 study of prexasertib (LY2606368), a CHK1/2 inhibitor, in pediatric patients with recurrent or refractory solid tumors, including CNS tumors: a report from the Children's Oncology Group Pediatric Early Phase Clinical Trials Network (ADVL1515). *Pediatr. Blood Cancer* **68**, e29065 (2021).
35. Byers, L. A. et al. A phase II trial of prexasertib (LY2606368) in patients with extensive-stage small-cell lung cancer. *Clin. Lung Cancer* **22**, 531–540 (2021).
36. Hong, D. S. et al. Evaluation of prexasertib, a checkpoint kinase 1 inhibitor, in a phase Ib study of patients with squamous cell carcinoma. *Clin. Cancer Res.* **24**, 3263–3272 (2018).
37. Lee, J.-M. et al. Prexasertib, a cell cycle checkpoint kinase 1 and 2 inhibitor, in BRCA wild-type recurrent high-grade serous ovarian cancer: a first-in-class proof-of-concept phase 2 study. *Lancet Oncol.* **19**, 207–215 (2018).
38. Calvo, E. et al. Phase I study of CHK1 inhibitor LY2603618 in combination with gemcitabine in patients with solid tumors. *Oncology* **91**, 251–260 (2016).
39. Hong, D. et al. Phase I study of LY2606368, a checkpoint kinase 1 inhibitor, in patients with advanced cancer. *J. Clin. Oncol.* **34**, 1764–1771 (2016).
40. Sen, T. et al. CHK1 inhibition in small-cell lung cancer produces single-agent activity in biomarker-defined disease subsets and combination activity with cisplatin or olaparib. *Cancer Res.* **77**, 3870–3884 (2017).
41. Do, K. T. et al. Phase 1 combination study of the CHK1 inhibitor prexasertib and the PARP inhibitor olaparib in high-grade serous ovarian cancer and other solid tumors. *Clin. Cancer Res.* **27**, 4710–4716 (2021).
42. Lord, C. J. & Ashworth, A. PARP inhibitors: synthetic lethality in the clinic. *Science* **355**, 1152–1158 (2017).

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



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

© The Author(s) 2024

Methods

Antibodies and reagents

Antibodies. Antibodies were procured from the following: H3K36me3 (Abcam, catalogue no. ab9050), γ H2AX (Millipore, catalogue no. 05-636 for IF), γ H2AX (Cell Signaling Technology, catalogue no. CST9718 for western blot), pRPA2S33 (Novus Biological, catalogue no. NB100-544), pCHK1S345 (Invitrogen, catalogue no. PA5-34625), 53BP1 (Novus Biological, catalogue no. NB100-304), cyclin A (BD Biosciences, catalogue no. 611268), pRNAPII S2/S4 (Abcam, catalogue no. ab252855), pCHK1-S345 (Cell Signaling Technology, catalogue no. CST2348), CHK1 (Abcam, catalogue no. ab32531), pRPA32/RPA2-Ser8 (Cell Signaling Technology, catalogue no. 54762 S), Vinculin (Cell Signaling Technology, catalogue no. CST13901), pFGFR2-Tyr653/654 (Cell Signaling Technology, catalogue no. CST3476S) and FGFR2 (Cell Signaling Technology, catalogue no. CST11835S).

Chemicals. Chemicals were procured from the following: CHIR-124 (Selleckchem, catalogue no. S2683), XL413 (Selleckchem, catalogue no. S7547) and triptolide (Millipore, catalogue no. 645900-5MG).

Cell culture

GBM39ec, GBM39HSR and HK296 were patient-derived neurosphere cell lines and were established as previously described²⁷. The parental PC3 line was obtained from ATCC. PC3 DM and PC3 HSR lines were isolated by the Mischel Lab through single-cell expansions of the parental PC3 line and are available from the Mischel Lab upon request. All the other cell lines were purchased from ATCC. Human prostate cancer cell line PC3 DM, PC3 HSR; colorectal cancer cell line COLO320DM, COLO320HSR; gastric cancer cell line SNU16; lung cancer cell line PC9 and hTERT-immortalized retinal pigment epithelial cell line RPE1 were cultured in 4.5 g l⁻¹ glucose-formulated Dulbecco's Modified Eagle's Medium (Corning) supplemented with 10% fetal bovine serum (FBS; Gibco). For GRO-seq and CHIP-seq, COLO320DM and COLO320HSR were grown in Roswell Park Memorial Institute 1640 with GlutaMAX (Gibco) with 10% FBS. GBM39ec, GBM39HSR and HK296 cell lines were cultured in Dulbecco's Modified Eagle's Medium/F12 (Gibco, catalogue no. 11320-033) supplemented with 1 \times B27 (Gibco, catalogue no. 17504-01), 20 ng ml⁻¹ epidermal growth factor (Sigma, catalogue no. E9644), 20 ng ml⁻¹ fibroblast growth factor (Peprotech, catalogue no. AF-100-18B), 1–5 μ g ml⁻¹ heparin (Sigma, catalogue no. H3149) and 1 \times GlutaMAX (Gibco, catalogue no. 35050-061). GBM39 cells used in sequencing assays were cultured without additional GlutaMAX. All the cells were maintained at 37 °C in a humidified incubator with 5% CO₂. Cell lines routinely tested negative for mycoplasma contamination.

GRO-seq

COLO320DM and COLO320HSR RNA was prepared by washing cells with ice-cold phosphate-buffered saline (PBS), then adding ice-cold LB (10 mM Tris-HCl pH 7.4, 2 mM MgCl₂, 3 mM CaCl₂, 0.5% IGEPAL-CA630, 10% glycerol, 1 mM DTT, protease inhibitors (Roche, catalogue no. 11836170001), RNase inhibitor (Ambion, catalogue no. AM2696)) and scraping cells into a 15 ml conical tube. Cells were spun at 1,000g for 10 min at 4 °C. Supernatant was removed and pellet was thoroughly resuspended in 1 ml LB using a wide bore tip. An additional 9 ml LB was added and then cells were spun at 1,000g for 10 min at 4 °C. Cells were resuspended in LB and spun down. Pellets were resuspended in ice-cold freezing buffer (50 mM Tris-HCl pH 8.3, 5 mM MgCl₂, 40% glycerol, 0.1 mM EDTA, 0.2 μ l RNase inhibitor per ml of freezing buffer) and spun at 2,000g for 2 min at 4 °C. Nuclei were resuspended in 100 μ l freezing buffer per 5 million cells. A nuclear run-on master mix was prepared (10 mM Tris-HCl pH 8.0, 5 mM MgCl₂, 1 mM DTT, 300 mM KCl, 0.5 mM ATP, 0.5 mM GTP, 0.003 mM CTP (unlabelled ribonucleotide triphosphates from Roche, catalogue no. 11277057001), 0.5 mM Bromo-UTP (Sigma, catalogue no. B7166), 1% Na-laurylsarcosine, 1 μ l

RNase inhibitor per 100 μ l) and preheated to 30 °C. An equal volume of master mix was added to aliquoted nuclei (5 million nuclei per replicate) and incubated at 30 °C for 5 min with gentle shaking. DNase digestion was performed using RQ1 DNase I and RQ1 buffer (Promega, catalogue no. M610A) for 30 min at 37 °C; the reaction was stopped with the addition of stop buffer to a final concentration of 10 mM Tris-HCl pH 7.4, 1% sodium dodecyl sulfate (SDS), 5 mM EDTA, 1 mg ml⁻¹ proteinase K. Samples were incubated for 1 h at 55 °C. NaCl was added to final concentration of 225 mM. Two phenolchloroform extractions were done, followed by one extraction with chloroform. RNA was precipitated in 75% EtOH with 1 μ l glycoblue (Ambion, catalogue no. 9516) overnight at -20 °C.

For GBM39ec and GBMHSR, cells were washed with ice-cold PBS and then spun for 5 min at 500g at 4 °C. Cells were then resuspended in ice-cold 10 ml swelling buffer (10 mM Tris-HCl pH 7.5, 2 mM MgCl₂, 3 mM CaCl₂, protease inhibitor, RNase inhibitor) and incubated on ice for 5 min. Cells were spun at 400g for 10 min at 4 °C and resuspended in 10 ml ice-cold glycerol swelling buffer (0.9 \times swelling buffer, 10% glycerol). While agitating the tube, 10 ml ice-cold lysis buffer (glycerol swelling buffer, 1% IGEPAL-CA630) was slowly added. Samples were incubated on ice for 5 min, then another 25 ml lysis buffer was added and samples were spun for 5 min at 600g at 4 °C. Samples were resuspended in ice-cold freezing buffer (50 mM Tris-HCl pH 8.0, 5 mM MgCl₂, 40% glycerol, 0.1 mM EDTA, RNase inhibitor) and spun at 900g for 6 min at 4 °C. An equal volume of pre-warmed nuclear run-on master mix was added to aliquoted nuclei (10 million nuclei per replicate) and incubated at 30 °C for 7 min with gentle shaking. Samples were then mixed thoroughly with 600 μ l Trizol LS and incubated at room temperature for 5 min. Next, 160 μ l chloroform was added to each sample, shaken vigorously, then incubated at room temperature for 3 min and centrifuged at 12,000g at 4 °C for 30 min. NaCl was added to the aqueous phase to a final concentration of 300 mM and RNA was precipitated in 75% EtOH with 1 μ l glycoblue overnight at -20 °C.

For all cell types, after overnight RNA precipitation, RNA was spun for 20 min at 21,130g at 4 °C. RNA pellets were washed in fresh 75% EtOH, briefly air-dried and then resuspended in 20 μ l water. Base hydrolysis was performed using 5 μ l 1 N NaOH for 10 min and then neutralized with 25 μ l 1 M Tris-HCl pH 6.8. Buffer exchange was performed using P30 Micro columns (Bio-Rad, catalogue no. 7326250), then treated with RQ1 DNase I and RQ1 buffer and incubated at 37 °C (10 min for COLO320 and 30 min for GBM39). Buffer exchange was performed again. Samples were treated with 3 μ l T4 polynucleotide kinase (PNK; New England Biolabs, catalogue no. M0201), 1 \times PNK buffer, 2 μ l 10 mM ATP and 2 μ l RNase inhibitor and incubated for 1 h at 37 °C. Another 2 μ l PNK was added per sample and incubation was continued for 30–60 min. RNA decapping was performed by adding ammonium chloride (final concentration 50 mM), poloaxamer 188 (final concentration 0.1%), 2 μ l messenger RNA decapping enzyme (New England Biolabs, catalogue no. M0608S) and 1 μ l RNase inhibitor and incubated at 37 °C for 30 min. EDTA was then added to the final concentration of 25 mM and samples were incubated at 75 °C for 5 min. Samples were then incubated on ice for at least 2 min. Sample volume was then brought to 100 μ l with binding buffer (0.25 \times SSPE, 1 mM EDTA, 0.05% Tween 20, 37.5 mM NaCl, RNase inhibitor). During T4 PNK treatment, 60 μ l anti-BrdU agarose beads (Santa Cruz Biotechnology, catalogue no. sc-32323ac) per sample were equilibrated in 500 μ l binding buffer by rotating for 5 min at room temperature, spun and washed again in binding buffer. Beads were then blocked in blocking buffer (1 \times binding buffer, 0.1% polyvinylpyrrolidone, 1 μ g ml⁻¹ ultrapure bovine serum albumin (BSA), RNase inhibitor) by rotating for 1 h at room temperature. Beads were then washed twice in binding buffer and resuspended in 400 μ l binding buffer. Decapped RNA was then added to the blocked beads and rotated for 1 h at room temperature. Beads were then washed once in binding buffer, once in low-salt buffer (0.2 \times SSPE, 1 mM EDTA, 0.05% Tween 20, RNase inhibitor), once in high-salt buffer (0.2 \times SSPE, 1 mM EDTA,

Article

0.05% Tween 20, 137.5 mM NaCl, RNase inhibitor) with 3 min of rotation, and twice in Tris-EDTA-Tween20 buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.05% Tween 20, RNase inhibitor). All spins with agarose beads were performed for 2 min at 1000g at room temperature and all washes were performed in 500 μ l buffer rotating for 5 min at room temperature unless otherwise noted. Samples were then eluted in elution buffer (50 mM Tris-HCl pH 7.5, 20 mM DTT, 1 mM EDTA, 150 mM NaCl, 0.1% SDS, RNase inhibitor) pre-warmed to 42 °C; four 10-min elutions were performed at 42 °C with periodic vortexing. The eluates for each replicate were pooled and RNA was then purified by phenolchloroform and chloroform with EtOH precipitation (COLO320) or by column purification using New England Biolabs Monarch RNA Cleanup Kit T2030 (GBM39). Sequencing libraries were prepared using the NEBNext Small RNA Library Prep Kit (New England Biolabs, catalogue no. E7330) and sequenced by Novaseq PE150. The sequence data were mapped to human reference genome (hg38) using STAR, v.2.7.10b (ref. 17). HOMER (v.4.11.1) was used for de novo transcript identification on each strand separately using the default GRO-seq setting. Reads with MAPQ values less than 10 were filtered using SAMtools (v.1.8). Duplicate reads were removed using picard-tools. GRO-seq signal was converted to the bigwig format for visualization using deepTools bamCoverage¹⁸ (v.3.3.1) with the following parameters: --binSize 10 --normalizeUsing CPM --effectiveGenomeSize 3209286105 --exactScaling.

Total RNA library preparation

Total RNA from each sample was isolated with Quick-RNA Miniprep Kit (Zymo Research, catalogue no. R1054) with input of 1–2 million cells. RNA libraries were constructed using TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero (Illumina, catalogue no. 20020596). Nextseq 550 sequencing system (Illumina) produced 20–30 million of $\times 2$, 75 bp paired-end reads per sample. The sequence data were mapped to human reference genome hg38 using STAR, v.2.7.10b (ref. 17), following the ENCODE RNA-seq pipeline. Reads with MAPQ values less than ten were filtered using SAMtools (v.1.8). Ribo-Zero signal was converted to the bigwig format for visualization using deepTools bamCoverage¹⁸ (v.3.3.1) with the following parameters: --binSize 10 --normalizeUsing CPM --effectiveGenomeSize 3209286105 --exactScaling.

KAS-seq library preparation

KAS-seq experiments were carried out as previously described¹² with modifications¹³. Briefly, cell culture media was supplemented with 5 mM N₃-kethoxal (final concentration), and cells were incubated for 10 min at 37 °C in a six-well plate. Genomic DNA was then extracted using the Monarch gDNA Purification Kit (NEB T3010S) following the standard protocol but with elution using 50 μ l 25 mM K₃BO₃ at pH 7.0. Click reaction was carried out by mixing 87.5 μ l purified DNA, 2.5 μ l 20 mM DBCO-PEG4-biotin (dimethylsulfoxide (DMSO) solution, Sigma, catalogue no. 760749) and 10 μ l 10 \times PBS in a final volume of 100 μ l. The reaction was then incubated at 37 °C for 90 min. DNA was purified using AMPure XP beads by adding 50 μ l beads per 100 μ l reaction, washing beads on a magnetic stand twice with 80% EtOH and eluting in 130 μ l 25 mM K₃BO₃. Purified DNA was then sheared using a Covaris E220 instrument down to around 200–400 bp size. Pulldown of biotin-labelled DNA was initiated by separating 10 μ l of 10 mg ml⁻¹ Dynabeads MyOne Streptavidin T1 beads (Life Technologies, catalogue no. 65602) on a magnetic stand, then washing with 180 μ l of 1 \times Tween Washing Buffer (TWB; 5 mM Tris-HCl pH 7.5; 0.5 mM EDTA; 1 M NaCl; 0.05% Tween 20). Beads were then resuspended in 300 μ l of 2 \times binding buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 2 M NaCl), sonicated DNA was added (diluted to a final volume of 300 μ l if necessary) and beads were incubated for at least 15 min at room temperature on a rotator. Beads were separated on a magnetic stand and washed with 300 μ l of 1 \times TWB and heated at 55 °C in a Thermomixer with shaking at 1,000 rpm for 2 min. The supernatant was removed on a magnetic stand and the TWB wash and 55 °C incubation were repeated.

Libraries were prepared on beads using the NEBNext Ultra II DNA Library Prep Kit (NEB, catalogue no. E7645). First, end repair was carried out by incubating beads for 30 min at 20 °C in a Thermomixer with shaking at 1,000 rpm in 50 μ l 1 \times EB buffer plus 3 μ l NEB Ultra End Repair Enzyme and 7 μ l NEB Ultra End Repair Enzyme. This was followed by incubation at 65 °C for 30 min. Second, adaptors were ligated by adding 2.5 μ l NEB adaptor, 1 μ l ligation enhancer and 30 μ l blunt ligation mix, incubating at 20 °C for 20 min, then adding 3 μ l USER enzyme and incubating at 37 °C for 15 min (in a Thermomixer, with shaking at 1,000 rpm). Beads were separated on a magnetic stand and washed with 180 μ l TWB for 2 min at 55 °C and 1,000 rpm in a Thermomixer. After magnetic separation, beads were washed in 100 μ l 0.1 \times TE buffer, resuspended in 15 μ l 0.1 \times TE buffer and heated at 98 °C for 10 min. PCR was carried out by adding 5 μ l of each of the i5 and i7 NEBNext sequencing adaptors together with 25 μ l 2 \times NEB Ultra PCR Mater Mix, with a 98 °C incubation for 30 s and 15 cycles of 98 °C for 10 s, 65 °C for 30 s and 72 °C for 1 min, followed by incubation at 72 °C for 5 min. Beads were separated on a magnetic stand and the supernatant was cleaned up using 1.8 \times AMPure XP beads.

Libraries were sequenced in a paired-end format on an Illumina NextSeq instrument using NextSeq 550 High-Output Kits (2 \times 36 cycles). The sequence data were mapped to the hg38 assembly of the human genome using Bowtie^{19,20} with the following settings: -v 2 -k 2 -m 1 -best --strata-X 1000. Duplicate reads were removed using picard-tools (v.1.99). MACS2 (ref. 21) (v.2.1.1) was used for peak-calling with the following parameters: --broad -g hs --broad-cutoff 0.01 -q 0.01. Browser tracks are generated after normalizing to input using bamCompare default setting.

ChIP-seq library preparation

Three million cells per replicate were fixed in 1% formaldehyde for 15 min at room temperature with rotation and then quenched with 0.125 M glycine for 10 min at room temperature with rotation. Fixed cells were pelleted at 1,300g for 5 min at 4 °C and washed twice with cold PBS before storing at -80 °C. Membrane lysis was performed in 5 ml LB1 (50 mM HEPES pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% IPEGAL-CA630, 0.25% Triton X-100, Roche protease inhibitors 11836170001) for 10 min at 4 °C with rotation. Nuclei were pelleted at 1,400g for 5 min at 4 °C and lysed in 5 ml LB2 (10 mM Tris-Cl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, Roche protease inhibitors) for 10 min at room temperature with rotation. Chromatin was pelleted at 1,400g for 5 min at 4 °C and resuspended in 1 ml of TE buffer plus 0.1% SDS before sonication on a Covaris E220 with the following settings: 140 W, 10% duty, 200 cycles per burst, 600 s per sample. Samples were clarified by spinning at 16,000g for 10 min at 4 °C. Supernatant was transferred to a new tube and diluted with two volumes of IP dilution buffer (10 mM Tris pH 8.0, 1 mM EDTA, 200 mM NaCl, 1 mM EGTA, 0.2% Na-DOC, 1% Na-laurylsarcosine, 2% Triton X-100). Then, 50 μ l of sheared chromatin was reserved as input and ChIP was performed overnight at 4 °C with rotation with 7.5 μ g of H3K36me3 antibody (ab9050) (1:300 dilution). Per sample, 100 μ l protein A dynabeads were washed three times with 1 ml chilled block buffer (0.5% BSA in PBS) and then added to the chromatin after overnight incubation with antibody and rotated for 4 h at 4 °C. Samples were washed five times in 1 ml pre-chilled wash buffer (50 mM HEPES pH 7.5, 500 mM LiCl, 1 mM EDTA, 1% IPEGAL-CA630, 0.7% Na-DOC) and then 1 ml pre-chilled TE + 50 mM NaCl. Samples were eluted in elution buffer (50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS) at 65 °C. NaCl was added to a final concentration of 455 mM. Samples were incubated with 0.2 mg ml⁻¹ proteinase K for 1 h at 55 °C and then decross-linked overnight at 65 °C. Samples were treated with 0.2 mg ml⁻¹ RNAase for 2 h at 37 °C and then purified with the Zymo ChIP DNA Clean & Concentrator Kit (D2505). Libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (E7645) and sequenced by NovaSeq PE150. The sequence data were trimmed by Trimmomatic²² (v.0.36) to remove adaptor and then mapped to

the hg38 assembly of the human genome using Bowtie2 (refs. 19,20) with the following settings: --local --very-sensitive --phred33 -X 1000. Reads with MAPQ values less than ten were filtered using SAMtools (v.1.8). Duplicate reads were removed using picard-tools. CHIP-seq signal was converted to the bigwig format for visualization using deepTools bamCoverage¹⁸ (v.3.3.1) with the following parameters: --binSize 10 --normalizeUsing CPM --effectiveGenomeSize 3209286105 --exactScaling.

IF and DNA FISH staining

Coverslips were coated with 100 $\mu\text{g ml}^{-1}$ poly-L-lysine overnight or 10 $\mu\text{g ml}^{-1}$ laminin for 1 h at 37 °C before seeding cells. Asynchronous cells were seeded onto slides and subject to different treatment. Where indicated, EdU was added to each well at 10 $\mu\text{g ml}^{-1}$ 30 min before collecting samples. IF and dual-IF DNA FISH staining were performed as described before. Briefly, slides were fixed with ice-cold 4% paraformaldehyde (PFA) for 15 min, followed by permeabilization with 0.5% Triton X-100 in PBS for 15 min at room temperature. Samples were blocked with 3% BSA in PBS for 1 h at room temperature before incubation with primary antibody diluted in blocking buffer at 4 °C overnight. Dilution ratio for first antibodies was as follows: γH2Ax , 1:500; pRPA2-S33, 1:1,000; pCHK1S345, 1:250; 53BP1, 1:500; cyclin A, 1:100; pRNAPII S2/S4, 1:1,000. After washing with PBS a total of three times for 5 min each, slides were incubated with secondary antibody diluted in blocking buffer at room temperature for 1 h. Samples were fixed with ice-cold 4% PFA for 20 min after washing with PBS. If combined with DNA FISH staining, fixed samples were further permeabilized with ice-cold 0.7% Triton X-100 per 0.1 M HCl (diluted in PBS) for 10 min on ice. DNA was denatured by 1.5 M HCl for 30 min at room temperature, followed by dehydration in ascending ethanol concentration. Diluted FISH probes (Empire Genomics) were pre-denatured at 75 °C for 3 min and added onto air-dried slides. After incubation at 37 °C overnight, slides were washed with 2 \times SSC to get rid of non-specific binding, followed by DAPI staining. Where indicated, EdU staining was performed with the Click-iT Plua EdU Alexa Fluor 647 Imaging Kit (Invitrogen, catalogue no. C10640).

Validation of PC3-DM and PC3-HSR cell lines

Genomic DNA was extracted from a confluent six-well dish using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's protocol. Briefly, single cells were collected and resuspended in 200 μl 1 \times PBS, followed by the addition of 20 μl QIAGEN proteinase K and 200 μl buffer AL. The mixture was pulse-vortexed for 15 s and incubated at 56 °C for 10 min. A volume of 200 μl absolute ethanol was added to the sample and pulse-vortexed for 15 s. The entire mixture was pipetted into a QIAamp Mini spin column and centrifuged at 6,000g for 1 min. Filtrate was discarded and 500 μl buffer AW1 was added to the column. After centrifugation at 6,000g for 1 min, the column was subjected to another round of wash with 500 μl buffer AW2. The filtrate was discarded after centrifugation at full speed for 3 min. The column was then placed in a clean 1.5 ml microfuge tube and 50 μl of buffer AE was added to reconstitute genomic DNA after centrifugation at 6,000g for 1 min.

WGS library preparation was performed with the FS DNA Library Prep Kit from NEB according to the manufacturer's protocol, with these parameters in place: (1) 250 ng gDNA was used as input; (2) fragmentation was done with an incubation time of 18 min to yield 200–450 bp fragments; (3) the final library size distribution was between 320–470 bp (that is, first bead selection was done with a bead volume of 30 μl and second bead selection was done with a bead volume of 15 μl); (4) the final PCR amplification was performed for four cycles. PE150 sequencing was performed on NovaSeq to yield at least 10 \times coverage at Novogene. Adaptor sequences were removed from raw fastq files using Trim Galore at default settings, followed by alignment to the hg38 reference genome using Map with BWA-MEM to generate the BAM

files. BAM files were then uploaded to the GenePattern Notebook for AmpliconArchitect analysis under default settings.

ecDNA structure analysis

We utilized the AmpliconSuite-pipeline (v.1.2.2, <https://github.com/AmpliconSuite/AmpliconSuite-pipeline>), which invoked CNVKit (v.0.9.9)⁴³, AmpliconArchitect⁴⁴ (AA; v.1.3.r8) and AmpliconClassifier³ (AC; v.1.1.2). In brief, the analysis pipeline first identifies seed regions of focal amplification from whole-genome copy number calls, then among the seed regions AA analyses copy number and structural variation jointly to construct a local genome graph encoding structural rearrangements and copy numbers. AA then extracts genome paths and cycles from the genome graph that explain the observed changes in copy number and structural variation. The outputs of AA are passed to AC, which applies a rule-based method to match the patterns of copy number, structural variation and structures extracted from the genome graph to known types of focal amplifications, such as ecDNA. To minimize sequencing artefacts derived from insert size distribution variance, we set the AmpliconSuite-pipeline argument --AA_insert_sdevs 9. For PC3 samples, --downsample 1 was also set to reduce additional sequencing artefacts. Default parameters were used otherwise.

For COLO320DM/HSR, we utilized the general ecDNA regions and the candidate ecDNA structure from ref. 10, after lifting over coordinates to hg38. For GBM39ec/HSR and PC3-DM/HSR, ecDNA regions were derived from AA output files. From the DM samples, regions with copy number greater than ten in the AA amplicon containing the ecDNA of interest were defined as the ecDNA regions. In GBM39ec/HSR, we also included the VIII deletion in the ecDNA region. Candidate ecDNA structures were derived from the AA cycle with highest assigned copy count containing the oncogene of interest (GBM39: *EGFR* and PC3: *MYC*). For GBM39, the ecDNA structure was consistent with a previously published reconstruction¹¹. Circular ecDNA visualizations were generated with CycleViz (<https://github.com/AmpliconSuite/CycleViz>). Gene and focal amplification copy numbers were derived from the AA graph file and the AC feature basic properties file, respectively. Structural similarity scores of the focal amplifications were computed using the feature_similarity.py script in AC, which computes a similarity score based on the overlapping genomic boundaries and shared structural variants between two focal amplifications. For the PC3 samples, we utilized the related amplicon_similarity.py script to obtain similarity scores, as the exact boundaries of the ecDNA could not be easily resolved with AC.

Replication combing assay

Replication fork speed in ecDNA was evaluated using the molecular combing assay. COLO320DM and COLO320HSR cells were seeded into plates and allowed to grow into log phase, nascent DNA synthesis was pulse labelled with thymidine analogues: CldU and IdU sequentially for equal amount of time. Following pulse labelling, cells were harvested and embedded into agarose plugs using the Genomic Vision FiberPrep Kit (Genomic Vision). DNA extraction, combing and immunostaining was performed according to the EasyComb service procedures (Genomic Vision). Coverslips were scanned with a FiberVision scanner and images were analysed using FiberStudio software (Genomic Vision). Fork speed was calculated using replication signals with contiguous CldU–IdU tracks. Only intact signals, flanked by counterstaining of the DNA fibre, were selected for analysis.

Locus-replication combing assay

DNA replication activity at the *MYC* loci was assessed using molecular combing assay. COLO320DM and COLO320HSR cells were seeded into plates and allowed to grow into log phase, nascent DNA synthesis was pulse labelled with thymidine analogues: CldU and IdU for equal amount of time. Following pulse labelling, cells were harvested and

Article

embedded into agarose plugs using the Genomic Vision FiberPrep kit (Genomic Vision). DNA extraction and combing was performed according to the EasyComb service procedures (Genomic Vision). DNA-labelled FiberProbes (Genomic Vision) targeting *MYC* loci were produced and hybridized to combed DNA. Correspondence between theoretical and experimental probe coverage patterns was validated by measuring hybridized probe length in control samples. After immunostaining of replication signals and DNA probes, coverslips were scanned with a FiberVision scanner. Image analysis and measurements were performed using FiberStudio software (Genomic Vision). Fork speed was calculated using replication signals with contiguous CldU–IdU tracks.

Comet-FISH

Alkaline comet-FISH assays were performed according to the literature, with minor modifications^{45,46}. Cells were harvested by trypsinization, washed with PBS and placed on ice. Cells were diluted in 37 °C low melting point (LMP) agarose (IBI Scientific) in PBS to a final concentration of 0.7% and spread on precoated glass slides with a coverslip. Overnight lysis was performed at 4 °C in alkaline lysis solution (2.5 M NaCl, 100 mM EDTA, 10 mM Tris pH 10, 1% Triton X-100, 10% DMSO) protected from light. The following day, slides were equilibrated for 30 min in alkaline electrophoresis buffer (200 mM NaOH, 1 mM EDTA, pH less than 13) in a Coplin jar and subsequently electrophoresed at 25 V for 30 min. Slides were then neutralized with Tris, dehydrated in 70% ethanol and dried at room temperature.

To detect ecDNA through FISH, Cy5-labelled probes were generated from RP11-440N18 BAC DNA sonicated to 150 bp and labelled using a DNA labelling kit (Mirus Bio). Slides were denatured with 0.5 M NaOH for 30 min at room temperature, dehydrated in an ethanol series (70%, 85%, 95%) and allowed to dry at room temperature. The hybridization mixture containing probe DNA (200 ng per slide) and Cot-1DNA (8 µg per slide) was denatured separately at 75 °C for 10 min and pre-annealed at 37 °C for 1 h. Probe was added to the slides and spread with a glass coverslip and incubated at 37 °C overnight in a humidified chamber. The following day, slides were washed four times in 2× SSC, 50% formamide at 42 °C and subsequently washed twice in 2× SSC at 42 °C. Slides were dipped briefly in 70% ethanol and air-dried. Slides were mounted with Everbrite (Biotium) containing SYBR Gold (Invitrogen) diluted 1:10,000 and sealed with nail polish. Images were collected on a Nikon Eclipse TE2000-E using a ×60 oil objective.

Cell viability assay

Cell viability assay was performed using CellTiter-Glo (Promega, catalogue no. G8461) as previously described⁴⁷. Briefly, cells were seeded into a 384-well plate one day before adding inhibitors. Equal volumes of vehicles or drugs diluted at indicated concentration were added into each well the next day, and the cells were incubated for three days. On the third day, after equilibrating plate and CellTiter-Glo reagent at room temperature for 30 min, reagent was added into each well and incubated for 15 min at room temperature. Luminescence was measured using a Synergy 2 microplate reader. Four biological replicates were performed for each condition. Data analysis was performed with GraphPad Prism (v.9.1.0).

TUNEL

TUNEL assay (Invitrogen, catalogue no. C10617) was performed to detect DNA fragmentation during apoptosis. COLO320DM, COLO320HSR and SNU16 cells were treated with 1 µM CHIR-124 for indicated times. All cells including floating cells were collected and spun down onto slides using a cytospin (Thermo Scientific Cytospin 4 Centrifuge). Slides were fixed with 4% PFA and permeabilized with 0.25% Triton X-100, followed by labelling of free double strand end with EdUTP by reaction catalysed by TdT enzyme in a humidified chamber at 37 °C for 60 min. Incorporated EdUTP was detected through Click-iT

Plus TUNEL reaction according to the manufacturer's manual at 37 °C for 30 min. Slides were counterstained with DAPI and mounted with ProLong Diamond Antifade.

Annexin V staining

Cell apoptosis was detected through flow cytometry using a FITC Annexin V Apoptosis Detection Kit (BD Biosciences, catalogue no. 556547). Cells were treated with 1 µM of CHIR-124 for the indicated time, and all the cells including floating cells were collected. After washing with PBS twice and cell number counting, cells were resuspended in 1× binding buffer provided by the kit at a concentration of 1 × 10⁶ cells per millilitre. One hundred microlitres of the cell suspension was transferred to a FACS tube and stained with FITC Annexin V and propidium iodide. After incubation at room temperature for 15 min, all the samples were analysed with BD LSR II flow cytometry (BD Biosciences) within 1 h. Flow cytometry data were analysed through Beckman Coulter Kaluza software (v.2.1).

Microscope and image analysis

Images were taken by conventional fluorescence microscopy or confocal microscopy. Conventional fluorescence microscopy was performed on a Leica DMi8 widefield microscope by Las X software (v.3.8.2.27713) using a ×63 oil objective. Confocal microscopy was performed on a ZEISS LSM 880 inverted confocal microscope using ZEN (black v.2.3) (Stanford CSIF Facility). Z-stacks were taken for each field of view and a best-in-focus stack was identified for downstream image analysis, except for Fig. 3a, where a max projection was performed by ImageJ (v.1.53t).

Image analysis and quantification were performed using the open-source software CellProfiler (v.4.2.1). For foci number analysis, DAPI staining, IF staining and DNA FISH channel were analysed through automatic thresholding and segmentation to cell nuclei, pRPA2S33/γH2AX foci and DNA FISH foci respectively. Colocalization was performed using an object-based colocalization method. For fluorescence intensity measurement, nuclei were called based on DAPI channel through automatic thresholding and segmentation; mean fluorescence intensity was retrieved by measuring mean fluorescence intensity within each nucleus.

RS score computation

RS score 1. The gene set variation analysis⁴⁸ was utilized to assess the enrichment of the DNA RS response (RSR) gene set²⁰ in TCGA samples using RNA-seq data⁴⁹. The RSR gene set was curated based on genes affected by defects in the DNA RS response. RNA-seq transcripts per kilobase million values for TCGA samples were retrieved from the GDC data portal⁴⁹. Gene set variation analysis generated enrichment scores for both up- and down-regulated RSR genes. The final RSR score was determined as the difference between the up and down enrichment scores.

RS score 2. The RS signature score of each sample from TCGA was retrieved from the literature from ref. 21, which was transformed linearly between zero and one by subtracting the minimum score and dividing by the maximum score. TCGA sample ecDNA status classification was performed as stated in a previous publication¹.

Both methods. Briefly, 1,921 TCGA samples were grouped into five subtypes by AC (<https://github.com/AmpliconSuite>): ecDNA, breakage–fusion–bridge, complex non-cyclic, linear and no amplification. Samples with a break–fusion–bridge or complex non-cyclic status were removed from the analysis due to the challenges of detecting ecDNA from short-read data. Samples with linear amplification and no amplification were classified as ecDNA[−]. After removing metastasis sample and ecDNA[−] samples without matching ecDNA⁺ samples of the same tissue origin, a total of 232 ecDNA⁺ and 582 ecDNA[−] samples were included in the analysis.

CRISPR experiment

sgRNA template oligos targeting the gene encoding CHK1 was synthesized (Integrated DNA Technologies) and was ligated into a CRISPR expression vector with red fluorescent protein (RFP) (Cellecra-pRSG1 6-U6-sg-HTS6C-UbiC-TagRFP-2A-Puro). Non-targeting green fluorescent protein (GFP) (sgNT-GFP) plasmid was purchased.

ecDNA⁺ and ecDNA⁻ HeLa cells were transduced with sgCHK1-RFP or sgNT-GFP virus, and puromycin (Sigma) was added at 2.5 $\mu\text{g ml}^{-1}$ for selection for 48 h. After 48 h of puromycin selection (day 0), an equal number of cells expressing either sgCHK1-RFP or sgNT-GFP were mixed to obtain the RFP to GFP population ratio. In the following days, flow cytometry analysis was performed to determine the sgCHK1-RFP to sgNT-GFP ratio. The mixed cell population cultures were maintained at subconfluency. The sgRNA sequences targeting CHK1 were as follows:

No. 17: CCTGACAGCTGCTCACTGGGT

No. 18: GCTGTCAGGAGTATTCTGAC

Western blotting

Samples were lysed in radioimmunoprecipitation assay buffer (Boston BioProducts, catalogue no. BP-115) supplemented with protease/phosphatase inhibitors (Fisher Scientific, catalogue no. 78444). Protein concentration was quantified with bicinchoninic acid assay (Fisher Scientific, catalogue no. 23225) and samples were prepared in 4× sample buffer (Bio-Rad, catalogue no. 1610747). Samples were loaded and run on 4–12% Bis-Tris Gradient Gel (Fisher Scientific, catalogue no. WG1403BOX) and transferred onto a nitrocellulose membrane (Bio-Rad, catalogue no. 1704271). The membrane was blocked with 5% BSA in Tris-buffered saline with Tween (Fisher Scientific, catalogue no. 28360) for an hour, and then primary antibody (1:1,000 dilution) was added and incubated overnight at 4 °C. Following primary antibody incubation, the membrane was washed with Tris-buffered saline with Tween and incubated with secondary antibody for 1 h. The membrane was then incubated with enhanced chemiluminescence reagent (Fisher Scientific, catalogue no. 32106) and image acquisition was performed on ProteinSimple FluorChemE.

Detection of phosphorylated CHK1 Ser345 using the AlphaLisa SureFire assay

Compound activity in cells was measured using an AlphaLISA SureFire Ultra p-CHK1 (Ser345) assay (Perkin Elmer, catalogue no. ALSU-PCHK1-A10K). HT29 cells were cultured in McCoy 5 A medium with 10% FBS and 1% penicillin-streptomycin and seeded to 96-well plates (Corning, catalogue no. 3599). Compounds were serially diluted in DMSO over a 10-point dose range with 3-fold dilution, and compound solution was added to each well containing cells. Plates were centrifuged at 1,000 rpm for 30 s. Plates were incubated at 37 °C for 16 h. Supernatant was removed by flicking the plate against a paper towel. Wells were washed once with PBS solution. To each well was added freshly prepared lysis buffer and plates were agitated on a plate shaker at 400 rpm for 30 min. The 96-well cell plates were centrifuged at 1,500 rpm for 1 min. From each well was transferred 10 μl of the lysates to a 384-well Optiplate (Perkin Elmer, catalogue no. 6007290). To each well was added Acceptor Mix (5 μl) and the plates were sealed and wrapped in foil. Plates were agitated on a plate shaker for 2 min, then incubated at room temperature for 1 h. To each well was added Donor Mix (5 μl) and the plates were sealed and wrapped in foil. Plates were agitated on a plate shaker for 2 min, then incubated at room temperature for 1 h. AlphaLisa signal was read on an EnVision multimode plate reader (Perkin Elmer). Data were fitted to dose–response curves using XLfit (IDBS) or GraphPad Prism (GraphPad software) to calculate IC₅₀ values for each compound tested.

Kinase HTRF biochemical assay

CHK1 enzyme activity was measured using a homogeneous time resolved fluorescence (HTRF) KinEASE assay (Cisbio, catalogue no.

62STIPEC). Full-length human CHK1 protein (GenBank accession number NP_001265.1) was obtained from Carna Biosciences, Inc. (catalogue no. 02-117). The enzyme reaction was carried out in assay buffer containing (final concentrations): CHK1 enzyme (0.012 $\text{ng } \mu\text{l}^{-1}$), MgCl₂ (5 mM) and DTT (1 mM). To determine compound dose response, DMSO stock solutions were serially diluted in a ten-point concentration series in duplicate. Compound solution (50 nL) was added to 384-well assay plates (Greiner, catalogue no. 784075). To each well containing compound solution was added assay buffer solution (5 μl). Plates were centrifuged at 1,000 rpm for 1 min, then incubated at room temperature for 10 min. The reaction was started by addition of substrate buffer (5 μl per well) containing (final concentrations): STK substrate 1-biotin (120 nM) and ATP (1 mM). Assay plates were centrifuged at 1,000 rpm for 1 min, then incubated at room temperature for 60 min. The reaction was stopped by the addition of detection buffer (Cisbio, 10 μl) containing (final concentrations): STK antibody cryptate (0.25 nM) and streptavidin-XL665 (7.5 nM). Plates were centrifuged at 1,000 rpm for 1 min, then incubated at 25 °C for 2 h. HTRF signal was read on an EnVision multimode plate reader (Cisbio) in HTRF mode. Data were fit to dose–response curves using XLfit (IDBS) or Prism (GraphPad Software) to calculate IC₅₀ values for each compound tested.

Phospho-RPA32 S8 IF high content imaging

Optical-bottom 96-well plates (Thermo Scientific, catalogue no. 165305) were coated with 50 μl of 1:1 poly-L-lysine (R&D Systems, catalogue no. 3438-100-01) and poly-D-lysine (R&D Systems, catalogue no. 3439-100-01) for 3 h at room temperature. The wells were washed once with 100 μl of PBS (Gibco, catalogue no. 10010-023) and all liquid was removed from the wells and allowed to dry fully at room temperature. COLO320 ecDNA⁺ cells were seeded at 15,000 cells per well in 100 μl of Roswell Park Memorial Institute media (Thermo Fisher, catalogue no. 22400089) supplemented with 10% FBS (Omega Scientific, catalogue no. FB-01). Cells were left to attach in a 37 °C incubator with 5% CO₂ overnight. The following day, cells were treated with BBI-825 for 16 h. Following treatment, all culture media was removed, and cells were fixed with 4% PFA (Boston BioProducts, catalogue no. BM-155) for 15 min at room temperature. After fixation, the 4% PFA was removed and wells were washed twice with 100 μl of PBS. The cells were then permeabilized with 100 μl of 0.5% Triton X-100 (Sigma-Aldrich, catalogue no. T8787) in PBS for 15 min at room temperature. After permeabilization, wells were washed twice with 100 μl of PBS and then blocked with 5% goat serum (Abcam, catalogue no. ab7481) and 1 mg ml⁻¹ of BSA (GeminiBio, catalogue no. 700-100 P) for 1 h at room temperature. The primary antibody (phospho-RPA32 (S8); Cell Signaling, catalogue no. 54762) was diluted at 1:200 in blocking buffer and 50 μl was added to all wells and incubated at 4 °C overnight. Plates were then washed three times with 100 μl of PBS and then incubated with 1:1,000 dilution of secondary antibody (Goat anti-Rabbit IgG Alexa Fluor Plus 594; Thermo Fisher, catalogue no. A32740s) and 1:1,000 dilution of Hoechst 33342 (Biotium, catalogue no. 40046) in blocking buffer for 1 h at room temperature. Plates were then washed three times with 100 μl of PBS; 100 μl of PBS was left in the wells following the final wash. The plate was imaged using a CellInsight CX7 LZR Pro High Content imager (Thermo Fisher Scientific) and data analysed using the Spot Detector BioApplication module on the HCS Studio Cell Analysis software (Thermo Fisher Scientific). Puncta were detected using a pixel thresholding method within a nucleus, and cells that contained three or more puncta of phosphorylated RPA32 Ser8 staining were considered as a positive signal.

Xenograft

Animal experiments were performed in accordance with protocols approved by the Charles River Accelerator and Development Lab (CRADL) Institutional Animal Care and Use Committee (protocol no. EB17-010-066). Mice were socially housed in individually ventilated cages on a 12/12 h light/dark cycle with temperatures between 65 and

Article

75 °F and 30–50% humidity. The SNU16 gastric cancer cell line was purchased from ATCC (catalogue no. CRL5974) and maintained in Roswell Park Memorial Institute growth medium (Gibco, catalogue no. 22400-089) supplemented with 10% FBS (Omega Scientific, catalogue no. FB-02). To establish tumours, 1×10^6 SNU16 cells in 200 μ l of a 1:1 mixture of PBS and Matrigel (Corning, catalogue no. 354234) were given by subcutaneous injection into the right flank of 9-week-old female severe combined immunodeficient beige mice (Envigo, strain code 186). Tumour measurements were taken two to three times per week and body weights were taken daily. Tumour volume measurements were obtained using digital calipers and tumour volumes (mm^3) were determined using the formula: tumour volume = $(L \times W^2)/2$, where L is the length/largest tumour diameter and W is the width/shortest tumour diameter, with all tumours collected before reaching 1,500 mm^3 . Animals (eight mice per group, which historically allowed for significance determination between vehicle and infigratinib) were randomly assigned to unblinded treatment with vehicle, infigratinib (15 mg kg^{-1} oral (PO) once-daily (QD)), BBI-2779 (30 mg kg^{-1} PO every other day (Q2D)) or the combination of BBI-2779 and infigratinib once average tumour volume was 285 (± 10)/mean (\pm s.e.m.) mm^3 . One vehicle tumour was taken down on day 22; the mouse was sacrificed due to large tumour volume. Infigratinib was formulated in a 1:1 mixture of sodium acetate buffer, pH 4.6 and polyethylene glycol 300. BBI-2779 was formulated in 0.5% methylcellulose (Sigma-Aldrich, catalogue no. M0512) and 0.2% Tween 80 (AG Scientific, catalogue no. T-2835) in HyPure Molecular Biology Grade Water (HyClone, catalogue no. SH30538.02). Dose holidays were provided to individual animals that demonstrated greater than –10% body-weight change from baseline, and Nutra-Gel was provided to the entire treatment group. Animals were sacrificed 6 h, 24 h or 36 h after the last dose, and tumours were collected for western blot or copy number analysis.

Copy number analysis from xenograft samples

For copy number analysis, tumours were cut into 10–20 mg pieces and flash-frozen in liquid nitrogen. DNA was extracted using the QIAcube DNA Extraction Kit (Qiagen, no. 51331). Briefly, a mixture of buffer ATL and proteinase K was added to the frozen tumour pieces, and they were set out to equilibrate to room temperature. Tumours were then vortexed for 30 s and placed into an incubator at 56 °C to digest overnight. The next morning, an additional 150 μ l of buffer ATL was added and samples vortexed for an additional 30 s to reduce the viscosity of the samples before transfer to the S block. Qiagen protocol for the 96 QIAcube HT was followed for the remainder of the DNA isolation. Purified DNA was quantified for the presence of double-stranded DNA on the QIAxpert (Qiagen, catalogue no. 9002340). The DNA was diluted to 5 $\text{ng } \mu\text{l}^{-1}$ ($5\times$ working stock) in RNase/DNase free water (Thermo Fisher Scientific, catalogue no. 10977015) and 2 μ l was loaded into a 384-well plate. Master mix recipe (Master Mix ($2\times$), 5.5 μ l; CNA (Target Gene) $20\times$, 0.55 μ l; CNR telomerase reverse transcriptase (TERT) $20\times$, 0.55 μ l; nuclease-free water, 2.2 μ l) was made containing TaqPath Pro Master Mix $2\times$ (Thermo Fisher Scientific, catalogue no. A30866) human female genomic DNA (Promega, catalogue no. G1521) as a reference, internal controls (human TERT) and FGFR2 or MYC target gene probe (Thermo Fisher Scientific, catalogue no. 4400292). Reactions were run on the QuantStudio 6 or 7 (Thermo Fisher Scientific) using the qPCR reaction settings as follows: denature/enzyme activation: 95 °C, 10 min; 40 cycles of denature 95 °C, 15 s; anneal/extend 60 °C, 60 s.

Quantifications and statistical analysis

All statistical methods and sample size have been stated in figure legends or the Methods section. No statistical methods were used to predetermine the sample size. The default test type was a two-sided statistic test, unless indicated in the text. The investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

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

Data availability

GRO-seq, Ribo-Zero RNA-seq, CHIP-seq, WGS and KAS-seq data generated in this study can be accessed by GEO under accession number <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE249657>. AA outputs and copy numbers of COLO320, GBM39 and PC3 can be accessed at <https://ampliconrepository.org/project/6639560c48cbf4a5ccffad4d>. WGS data of COLO320 and GBM39 used in this study are publicly available under SRA: <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA506071>¹¹. BBI-2779 is available upon request to C.A.H. at Boundless Bio. Source data are provided with this paper.

43. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).
44. Deshpande, V. et al. Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. *Nat. Commun.* **10**, 392 (2019).
45. Shaposhnikov, S., El Yamani, N. & Collins, A. R. Fluorescent in situ hybridization on comets: FISH comet. *Methods Mol. Biol.* **1288**, 363–373 (2015).
46. Laubenthal, J. & Anderson, D. in *DNA Electrophoresis: Methods and Protocols* (ed. Makovets, S.) 219–235 (Humana, 2013); https://doi.org/10.1007/978-1-62703-565-1_15.
47. Bi, J. et al. Targeting glioblastoma signaling and metabolism with a re-purposed brain-penetrant drug. *Cell Rep.* **37**, 109957 (2021).
48. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* **14**, 7 (2013).
49. Grossman, R. L. et al. Toward a shared vision for cancer genomic data. *N. Engl. J. Med.* **375**, 1109–1112 (2016).

Acknowledgements We thank members of the Chang and Mischel Labs for discussion. The work from the Mischel, Chang and Bafna Labs was delivered as part of the eDNAmiC team, supported by the Cancer Grand Challenges partnership funded by Cancer Research UK (CGCATF-2021/100012 and CGCATF-2021/100025) and the National Cancer Institute (OT2CA278688 and OT2CA278635) to P.S.M., H.Y.C. and V.B. This project was supported by National Institutes of Health RM1-HG007735 (H.Y.C., W.J.G. and C.C.). Boundless Bio provided its proprietary compound BBI-2779 (PCT no. WO2022251502), and the CRISPR and chemoproteomic synthetic lethality data that first identified CHK1 as a potential ecDNA target was performed in the Boundless Bio laboratory. Confocal microscopy was performed on instruments in the Stanford Cell Science Imaging Facility core facility (supported by RRID: SCR_017787). FACS analysis was performed on instruments in the Stanford Shared FACS Facility. We thank the Stanford Genomics Core for short tandem repeat profiling of PC3 cell lines. H.Y.C. is an investigator of the Howard Hughes Medical Institute.

Author contributions H.Y.C., P.S.M. and C.A.H. conceived the study. J.T., E.J.C., I.T.-L.W., S.C. and G.W. conducted studies of ecDNA RS and CHK1 activation. S.C. and S.M. conducted CRISPR and chemoproteomic synthetic lethality screens that first identified CHK1 as a potential ecDNA target. N.E.W., G.W., G.K.M. and R.L. conducted genomic studies of ecDNA transcription, ssDNA accumulation and response to triptolide. J.L., Y.W., S.Z., N.A., C.C., W.J.G. and V.B. supported analyses. S.C. mapped ecDNA replication speed. P.H. and S.J.B. conducted comet assays of ecDNA breakage. S.C., E.T., S.G.M., R.H., J.P., A.S., S.M., S.T.M., A.B.P., S.K. and C.A.H. conducted CHK1 knock-out experiments and designed, synthesized and evaluated BBI-2779. J.T., N.E.W., G.W., S.C., P.S.M. and H.Y.C. wrote the manuscript with input from all co-authors. C.A.H., P.S.M. and H.Y.C. supervised their respective teams for the study.

Competing interests H.Y.C., P.S.M. and V.B. are each a co-founder, advisor and have an equity interest in Boundless Bio. H.Y.C. is also a co-founder of Accent Therapeutics, Cartography Biosciences and Orbital Therapeutics, and an advisor of 10x Genomics, Arsenal Biosciences, Chroma Medicine, Exai Bio and Spring Discovery. S.J.B. is also an advisor and has an equity interest in Boundless Bio. S.C., E.T., S.G.M., R.H., J.P., A.S., S.M., S.T.M., A.B.P., S.K. and C.A.H. are employees of Boundless Bio. V.B. is a co-founder, consultant, SAB member and has equity interest in Abterra and Boundless Bio, and the terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict-of-interest policies. J.L. previously provided consulting services to Boundless Bio. The remaining authors declare no competing interests.

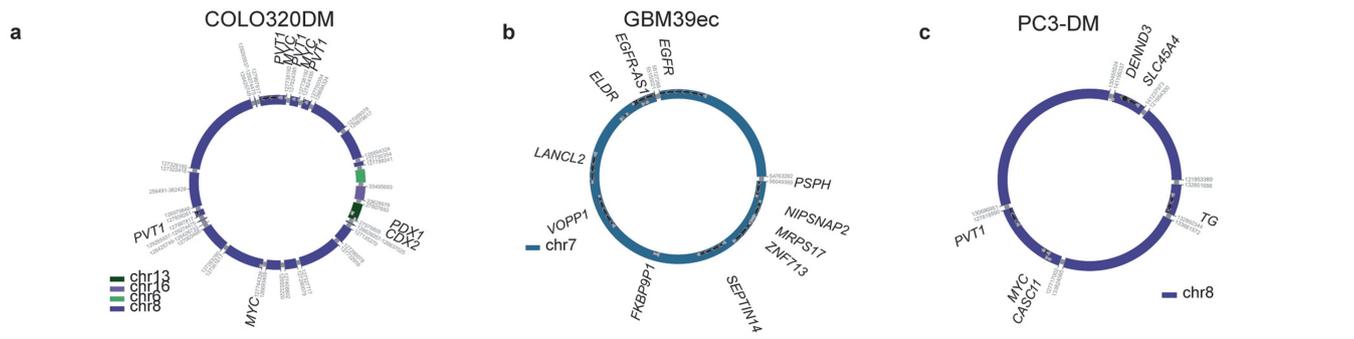
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-024-07802-5>.

Correspondence and requests for materials should be addressed to Christian A. Hassig, Paul S. Mischel or Howard Y. Chang.

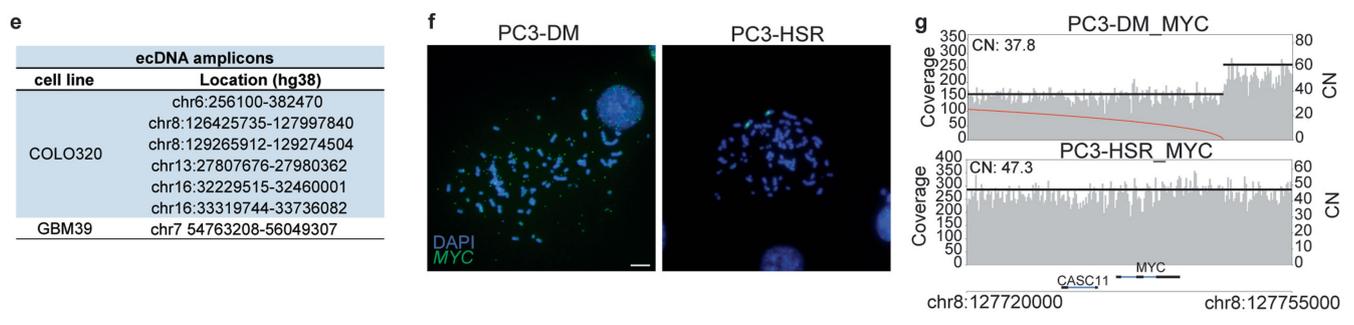
Peer review information Nature thanks Maite Huarte and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



d

Isogenic pair	Similarity Score	Similarity Score p-value	DM size	HSR size	Genomic overlap	DM oncogene CN	HSR oncogene CN
COLO320DM/HSR pair	0.583	0.0024	1580615	1786777	1507628	195.5	133
GBM39ec/HSR pair	1	$P < 1e-16$	1258334	1258334	1258334	97.8	93.9
PC3DM/HSR pair	0.522	0.0061	3482318	3201347	2382668	37.8	47.3

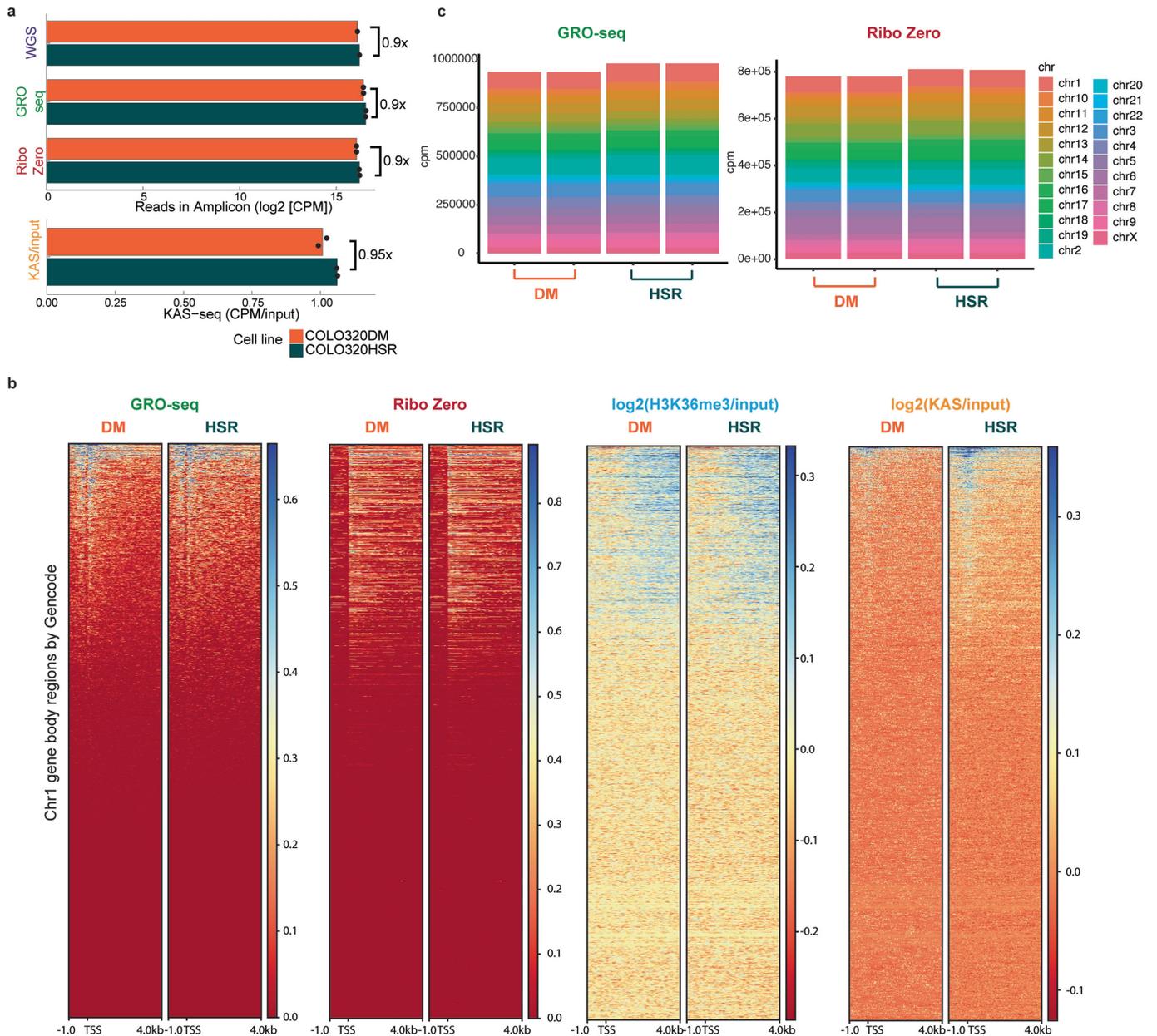


h

STR markers	TH01	D21S11	D5S818	D13S317	D7S820	D16S539	CSF1PO	AMEL	vWA	TPOX	Notes	
PC-3 (ATCC® CRL-1435)	6, 7	-	13, 13	11, 11	8, 11	11, 11	11, 11	X, X	17, 17	8, 9	Reference cell line	
Genotype	PC3-DM	6, 7	29, 31.2	13, 13	11, 11	8, 11	11, 11	11, 11	X, X	17, 17	8, 9	Identical to PC-3 (ATCC® CRL-1435)
	PC3-HSR	6, 7	29, 31.2	13, 13	11, 11	8, 11	11, 11	11, 11	X, X	17, 17	8, 9	Identical to PC-3 (ATCC® CRL-1435)

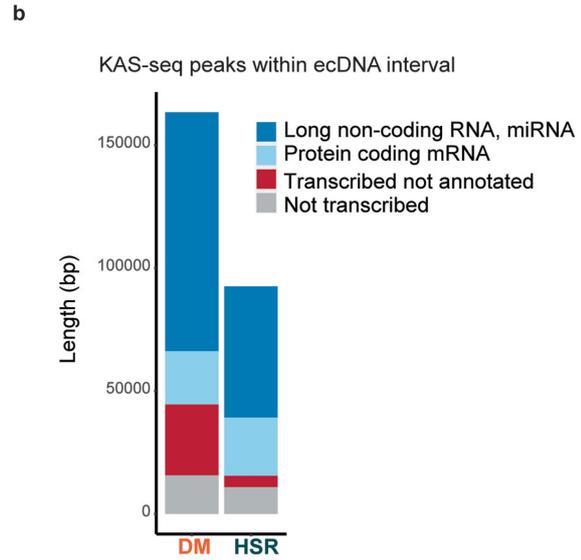
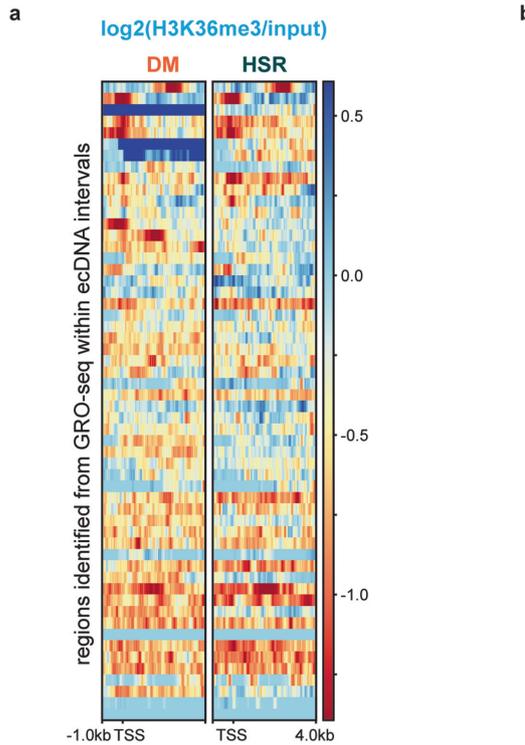
Extended Data Fig. 1 | Characterization of isogenic cell line pairs. (a-c) The structure of one dominant ecDNA detected in COLO320DM (a), GBM39ec (b) and PC3-DM (c). COLO320DM ecDNA reconstruction is adapted from Hung, et al. Nature 2021¹⁰ using hg38 coordinates; GBM39ec ecDNA reconstruction is adapted from Turner, et al. Nature 2017⁷ using hg38 coordinates; PC3-DM ecDNA was reconstructed by short-read WGS aligned to hg38, capturing the MYC-containing genome cycle identified by AmpliconArchitect. Some complexities of the PC3 DM ecDNA may not be captured from short read sequencing alone. (d) Amplicon similarity analysis of 3 near-isogenic cell line pairs as computed from whole genome sequencing data. Major oncogene copy number was

extracted for individual cell line: COLO320 pairs: *MYC*; GBM39 pairs: *EGFR*; PC3: *MYC*. (e) Genomic intervals in COLO320 and GBM39 cells used in this study for ecDNA amplicons. (f) Representative metaphase-FISH images of PC3-DM and PC3-HSR cells confirming *MYC* gene amplification on ecDNA and chromosome respectively. For each clone, about 6–15 metaphase spread images were collected from a one-off validation experiment to check amplicon status. Scale bar: 10 μ m (g) Amplicon structure and absolute gene copy of *MYC* between PC3-DM and PC3-HSR based on WGS. (h) Comparison of STR profiles between PC-3 from ATCC, PC3-DM and PC3-HSR.



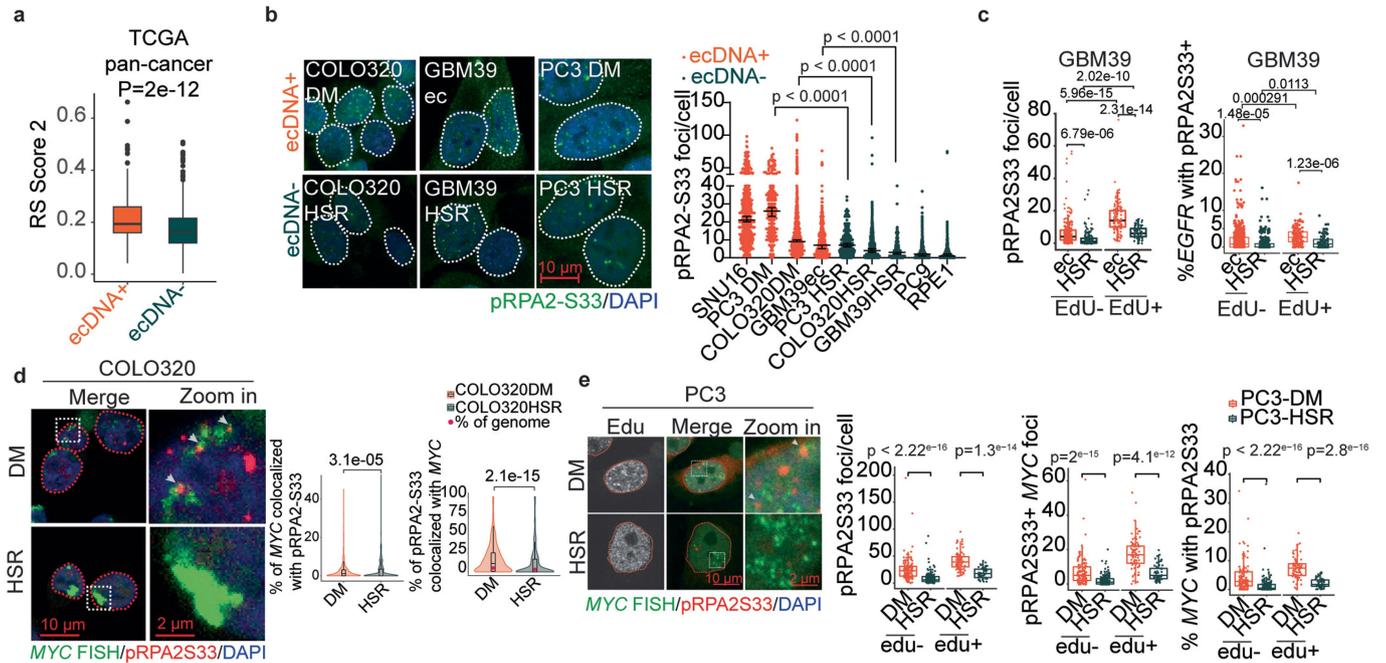
Extended Data Fig. 2 | Similar levels of transcription and ssDNA accumulation on normal chromosome 1 of COLO320 cell lines. (a) Read density of genomic assays in COLO320DM and COLO320HSR in total counts per million (CPM) on chromosome 1, which is outside of ecDNA intervals. KAS-seq read density is shown as (CPM) of the KAS-seq relative to CPM of the input. The mean of two biological replicates is shown for GRO-seq, Ribo-Zero and KAS-seq; a single replicate is shown for WGS. (b) Metagenome plot

visualization of GRO-seq, Ribo Zero RNA-seq, log₂ of input-normalized H3K36me3 ChIP-seq, and log₂ of input-normalized coverage of KAS-seq within chromosome 1. All plots are anchored at the GRO-seq TSS as identified by HOMER. One representative biological replicate for each condition is visualized. (c) Stacked bar charts of read density in total counts per million (CPM) of GRO-seq and Ribo-Zero in COLO320DM and COLO320HSR across all chromosomes.



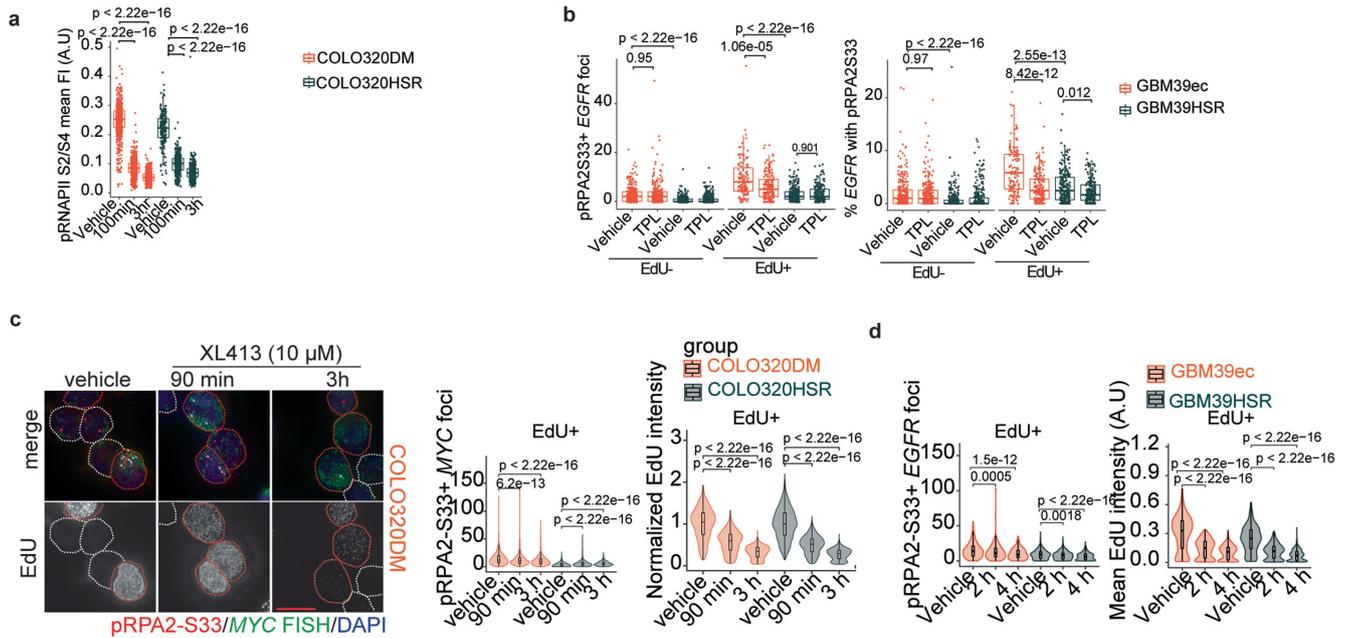
Extended Data Fig. 3 | H3K36me3 and KAS-seq signals within the ecDNA interval of COLO320 cell lines. (a) Metagene heatmap plot visualization of \log_2 of input-normalized H3K36me3 ChIP-seq within the ecDNA interval. Plots are anchored at the GRO-seq TSS as identified by HOMER. One representative

biological replicate for each cell line is visualized. (b) Accumulative bar plots of length distributions of all KAS-seq peaks identified within the ecDNA interval classified by GRO-seq transcription status and Gencode v43 annotation. Two biological replicates were used per cell line.



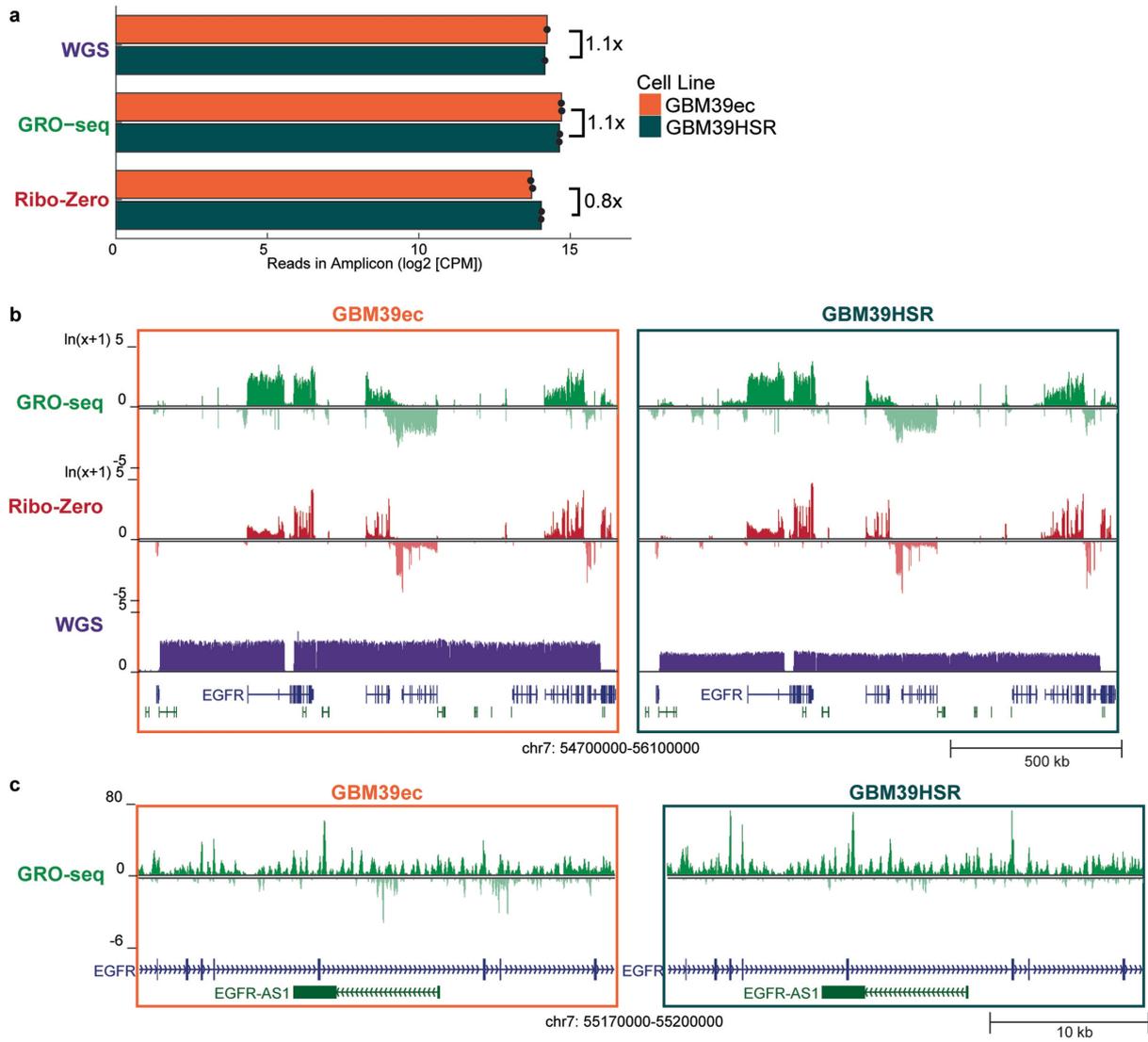
Extended Data Fig. 4 | RS on ecDNA with different amplification sequence in different tumour cells. (a) RS score 2 computed in TCGA patients grouped by ecDNA amplification status (232 ecDNA positive, 582 ecDNA negative, two-sided Wilcoxon test. Boxplot parameter same as in Fig. 2b). (b) Left: Representative images of pRPA2-S33 immunofluorescence staining in 3 isogenic tumour cell line pairs. Right, pRPA2-S33 foci number in individual cells, each dot indicated one cell. (Median with 95% CI, ordinary one-way ANOVA with multiple comparisons test, n, 666, 205, 1191, 369, 244, 655, 244, 1505, 775). (c) Quantification of images in Fig. 2e, left: total pRPA2-S33 foci/cell; right: % of *EGFR* co-localized with pRPA2-S33. (Box plot parameters same as in Fig. 2b, Two-tailed Wilcoxon test, EdU- group: GBM39ec, n = 244, GBM39HSR, n = 143; EdU+ group: GBM39ec, n = 104, GBM39HSR, n = 72). (d) pRPA2-S33 immunofluorescence combined with *MYC* FISH staining in COLO320DM and COLO320HSR cells. Left: representative images. Middle: % of *MYC* colocalized

with pRPA2-S33. Right: % of pRPA2-S33 that colocalized with *MYC*, red dot indicated % of genome that taken by amplicon as calculated by WGS counts in each cell line. (Data was presented as violin plot with the addition of boxplot. Violin plot outlines kernel probability density and boxplot parameters same as in Fig. 2b, Two-tailed Wilcoxon test COLO320DM, n = 939; COLO320HSR, n = 568). (e) pRPA2-S33 immunofluorescence combined with *MYC* FISH staining in PC3-DM and PC3-HSR cells, with EdU added 30 min before fixation. Left: representative images. 2nd left: pRPA2-S33 foci number. 2nd right: colocalized foci number between pRPA2-S33 and *MYC* FISH. Right: percentage of *MYC* co-localized with pRPA2-S33. (Box plot parameters were the same as in Fig. 2b, two-tailed Wilcoxon test, EdU+ group: PC3-DM, n = 81, PC3-HSR, n = 58; EdU- group: PC3-DM, n = 128, PC3-HSR, n = 184). (Scale bar represents 10 μ m or as otherwise specified).



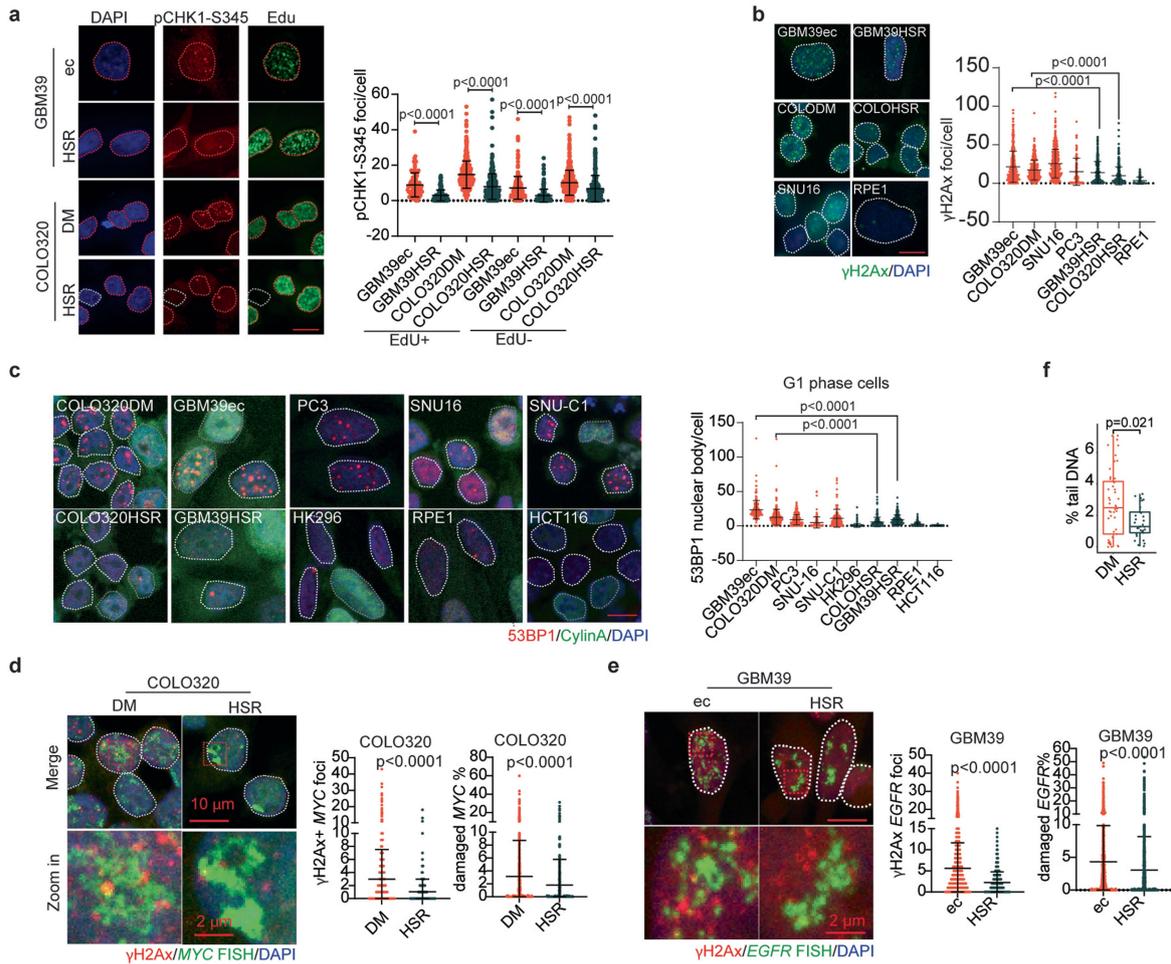
Extended Data Fig. 5 | Transcription replication conflict drives replication stress in ecDNA containing tumour cells. (a) Mean pRNAPII S2/S4 fluorescence intensity (arbitrary units) was measured in datasets shown in Fig. 2h. (Box plot parameters were the same as in Fig. 2b, two-tailed Wilcoxon test, n : 354, 350, 269, 130, 185, 161). (b) Quantification of dataset shown in Fig. 2i. Left, pRPA2-S33 and *EGFR* colocalized foci number; right, percentage of *EGFR* colocalized with pRPA2-S33. (Box plot parameters were the same as in Fig. 2b, two-tailed Wilcoxon test, sample size from left to right: n = 280, 256, 332, 348, 138, 189, 253, 208) (c) pRPA2-S33 IF combined with *MYC* FISH staining in COLO320DM and COLO320HSR cells treated with XL413 (10 μ M) for indicated time, EdU was added 30 min before fixing. Left, representative images in COLO320DM cells, middle, quantification of pRPA2-S33 and *MYC*

DNA FISH co-localized foci in EdU+ cells; right, Mean EdU intensity normalized to vehicle treated cells in each group. (Two-tailed Wilcoxon test, sample number from left to right: 1835, 1241, 1195, 2578, 990, 1068). (d) pRPA2-S33 IF combined with *EGFR* FISH staining in GBM39ec and GBM39HSR cells treated with XL413 (10 μ M) for indicated time, EdU was added 30 min before fixing. Left, quantification of pRPA2-S33 and *MYC* DNA FISH co-localized foci in EdU+ cells; right, Mean EdU intensity (arbitrary units) in EdU+ cells. (Two-tailed Wilcoxon test, sample number from left to right: 431, 494, 603, 674, 820, 829). (plots c-d were presented as violin plot with the addition of boxplot. Violin plot outlines kernel probability density and boxplot parameters same as in Fig. 2b. Scale bar represents 10 μ m or as otherwise specified).



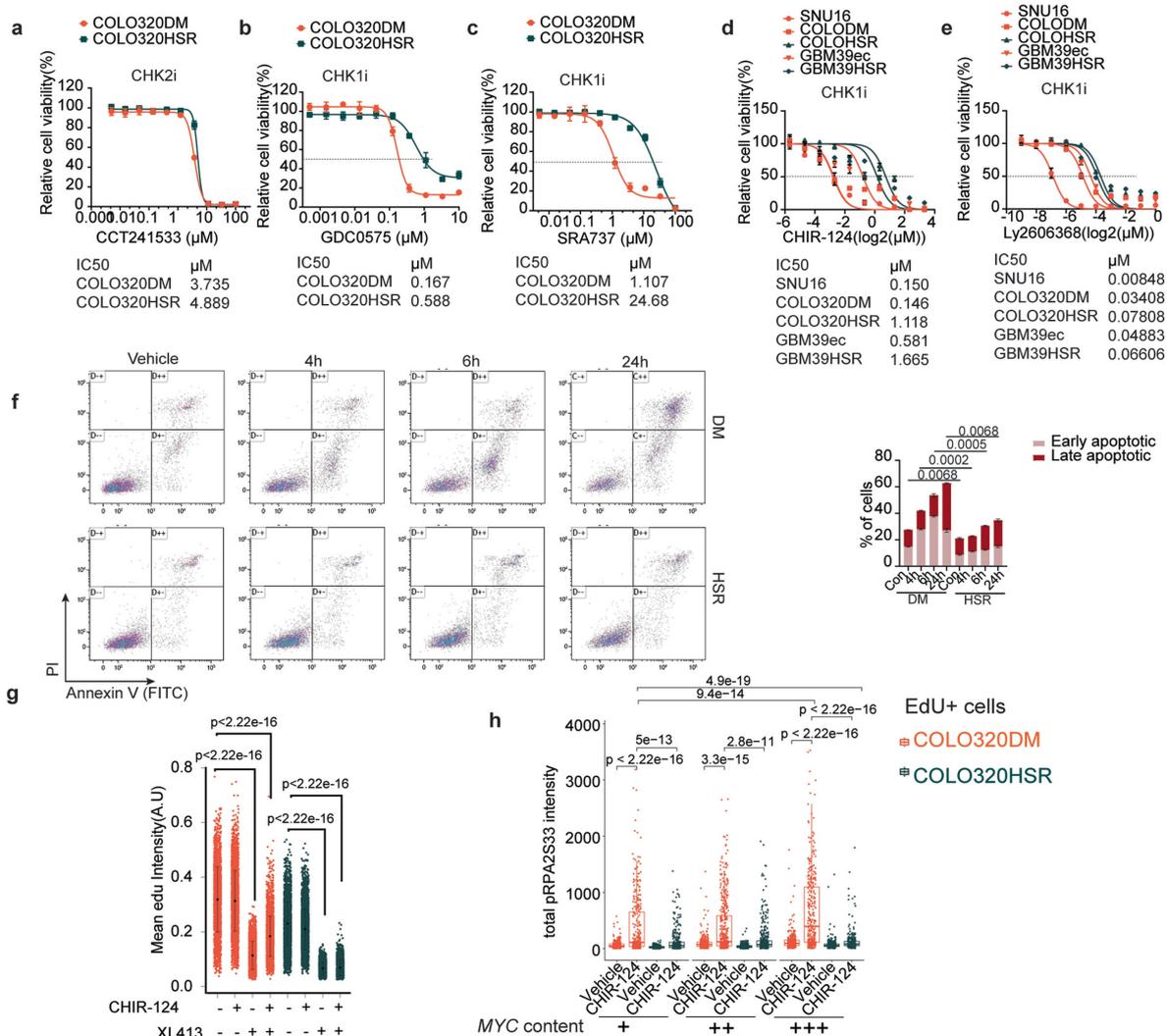
Extended Data Fig. 6 | Transcription within the ecDNA interval of GBM39 cell lines. (a) Read density of genomic assays in GBM39ec and GBM39HSR within the ecDNA interval in total counts per million (CPM). The mean of two biological replicates is shown for GRO-seq, Ribo-Zero and KAS-seq; a single replicate is shown for WGS. (b) Genome tracks highlighting the GBM39 ecDNA

interval. One representative biological replication for each condition is visualized. (c) Genome tracks highlighting the increased transcription at EGFR-AS1 in GBM39ec compared to GBM39HSR. One representative biological replication for each condition is visualized.



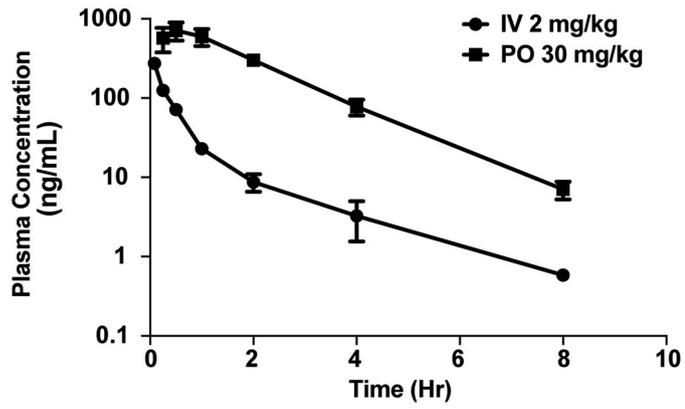
Extended Data Fig. 7 | RS induces DNA lesions and activates S phase check point in ecDNA containing tumour cells. (a) pCHK1-S345 staining in 2 isogenic cell line pairs, with EdU added for 30 min. Left, representative images, red dotted lines mark EdU+ and white dotted lines mark EdU- nuclei; right, quantification of pCHK1 foci number. (mean ± SD, two tailed Mann-Whitney test, n: 97, 192, 566, 339, 267, 252, 610). For COLO320HSR the same image is shown in Fig. 3a; they are used here to compare against multiple cell lines. (b) γH2AX staining in cell line panels with or without ecDNA. Left, representative image; right, quantification of γH2AX foci number per cell. (mean ± SD, Ordinary one-way ANOVA with multiple comparison test, n: 402, 362, 499, 101, 388, 418, 80) (c) 53BP1 combined with CyclinA staining in cell line panels with or without ecDNA. Left, representative image. White dotted lines mark G1 phase cells (CyclinA-). Right, quantification of 53BP1 in G1 phase cells. (mean ± SD, Ordinary one-way ANOVA with multiple comparison test, n, 180, 277, 240, 146,

126, 211, 188, 163, 221, 494). For COLO320DM and COLO320HSR the same images are shown in Fig. 3a; they are used here to compare against multiple cell lines. (d) γH2AX IF combined with MYC FISH staining in COLO320DM and COLO320HSR cells. Left: representative images. Middle, γH2AX and MYC colocalized foci number, right, percentage of MYC colocalized with γH2AX. (mean ± SD, two-tailed Mann-Whitney test, sample size, n, 804, 411). (e) γH2AX IF combined with EGFR FISH staining in GBM39ec and GBM39HSR cells. Left, representative image. Middle, γH2AX and EGFR colocalized foci number, right, percentage of EGFR colocalized with γH2AX. (mean ± SD, two tailed Mann-Whitney test, sample size, GBM39ec, n = 1638; GBM39HSR, n = 1863). (f) Quantification of percentage of tail DNA content with the dataset shown in Fig. 3b. (Box plot parameters were same as in Fig. 2b, two-tailed Wilcoxon test, COLO320DM, n = 49, COLO320HSR, n = 33). (Scale bar represents 10 μm or specified).



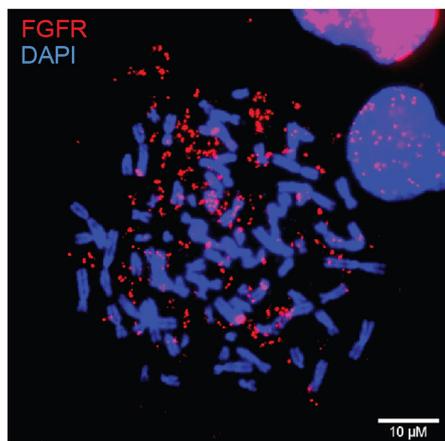
Extended Data Fig. 8 | ecDNA containing tumour cells are sensitive to targeted CHK1 inhibition. (a-e) Cell viability curve of COLO320DM, COLO320HSR, GBM39ec, GBM39HSR and SNU16 in response to different chemicals targeting CHK1 or CHK2. a. CHK2i, CCT241533; b. CHK1i, GDC0575; c. CHK1i, SRA737; d. CHK1i, CHIR-124; e. CHK1i, Ly2606368. Half maximal inhibitory concentrations (IC50) for each inhibitor in individual cell lines were listed on the bottom. (sample size in a-c, n = 2; d-e, n = 4, mean \pm SD) (f) FACS analysis of Annexin V staining in COLO320DM and COLO320HSR cells subjected to 1 μM CHIR-124 for indicated time. Left, gating setting and representative plots, Early apoptotic cells: Annexin V + and PI-; late apoptotic cells: Annexin V + and PI+. Right, % of apoptotic cells (mean \pm SEM, P values

quantified by two-tailed students' t test, n = 2.) (g) Quantification of mean EdU intensity (arbitrary units) in dataset shown in Fig. 3f. (mean \pm SD, P values quantified by two-tailed students' t test. Sample size from left to right: n = 419, 284, 1085, 596, 209, 242). (h) pRPA2-S33 IF combined with MYC FISH staining in COLO320DM and COLO320HSR cells treated with 100 nM CHIR-124 for 2 h, EdU was added 30 min before fixing. Accumulation of further RS upon CHK1i was quantified by total pRPA2-S33 fluorescence intensity in EdU+ cells. COLO320DM and COLO320HSR cells were grouped into 3 subgroups with different amplicon content based on MYC DNA FISH staining. (Box plot parameters were same as in Fig. 2b, two-tailed Wilcoxon test, Two-tailed Wilcoxon test, sample number from left to right: 355, 337, 466, 315; 472, 448, 622, 420; 354, 337, 466, 315).



Extended Data Fig. 9 | BBI-2779 has optimal PK exposure in mice.
Plasma concentration time curve of BBI-2779 in mouse administered either intravenously (IV) at 2 mg kg⁻¹ or orally (PO) at 30 mg/kg⁻¹. Data are mean ± s.e.m., n = 3 per group.

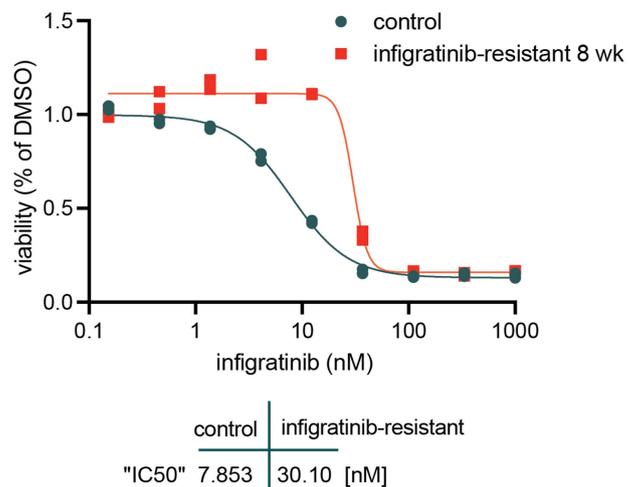
a.



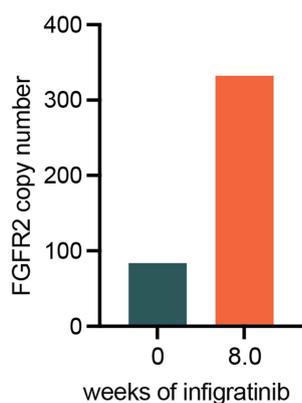
infigratinib resistance
(IC50, 25nM)

0 2 4 6 8
weeks of treatment

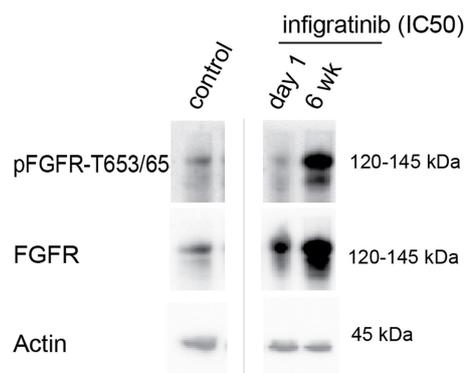
b.



c.



d.



Extended Data Fig. 10 | Targeted therapeutic resistance shaped by intracellular ecDNA-driven oncogene amplification. (a) *FGFR2* (red) FISH imaging of cells in metaphase demonstrate amplification of *FGFR2* oncogene on ecDNA in SNU16 cells. Nuclear staining is illustrated using DAPI (blue). Representative FISH image from multiple independent cytogenetic analysis with similar results. Scale bar represents 10 μ m. (b) Timeline of experimental overview. After 8 weeks of infigratinib treatment (EC50 dose of 25 nM), cells were assessed for infigratinib resistance. A 3-day Cell Titer Glo reveals resistance in SNU16 cells treated with infigratinib (n = 2 biologically

independent samples). (c) qPCR based quantification of *FGFR2* oncogene numbers after 8 weeks of infigratinib treatment showing SNU16 cells resistant to infigratinib with significant increase in *FGFR2* target selection/amplification (qPCR data represents n = 1 biological sample. Multiple repeat analysis reveal similar results). (d) Western blotting illustrating enhanced expression of FGFR signaling pathways involved in therapeutic resistance (Protein data represents n = 1 biological sample. Multiple repeat analysis reveal similar results). Lane 1 (control) is from a non-contiguous portion of the gel from lanes 2 and 3 (infigratinib treated samples).

Reporting Summary

Nature Portfolio 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 Portfolio 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 (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 Confocal images were taken by ZEN (black version 2.3), wide field fluorescence images were taken by Las X software (version: 3.8.2.27713). Western Blot images were taken by ProteinSample FluorChem E system.

Data analysis Interphase image analysis:
Foci counting analysis was performed by CellProfiler (4.2.1). Images were split into single channel images, and segmentation was performed using IdentifyPrimaryObjects module with automatic thresholding and declamping. To analyze the foci structure, IF channel images were enhanced by calculating difference of two images applied with different Gaussian filter ($\sigma=1$, $\sigma=2$).
For colocalization analysis between IF and DNA FISH images, images were split into single channel image, and IF, DNA FISH objects were called using IdentifyPrimaryObjects module in CellProfiler (4.2.1) independently. The colocalization between IF and DNA FISH objects were quantified through their presence at same location (x , y) as an object based colocalization analysis. Max projection was performed by FIJI-ImageJ (version 1.53t) before foci counting in Fig 3a.
Flow Cytometry data analysis:
The image in the schema in Extended Data Fig. 8f was generated by Beckman Coulter Kaluza software(version 2.1).
Cell viability assay data analysis was performed with GraphPad Prism (Version 9.1.0)
GRO-seq Data analysis:
The sequence data were mapped to human reference genome (hg38) using STAR (version 2.7.10b), HOMER (v4.11.1) was used for de novo transcript identification on each strand separately using the default GRO-seq setting. Reads with MAPQ values less than 10 were filtered using SAMtools (v1.8). Duplicate reads were removed using picard-tools. GRO-seq signal was converted to the bigwig format for visualization using deepTools bamCoverage (v3.3.1) with the following parameters: --binSize 10 --normalizeUsing CPM --effectiveGenomeSize 3209286105 --exactScaling.

KAS-seq Data analysis:

The sequence data were mapped to the hg38 assembly of the human genome using Bowtie with the following settings: -v 2-k 2-m 1--best--strata-X 1000. Duplicate reads were removed using picard-tools (version 1.99). MACS2 (v.2.1.1) was used for peak-calling with the following parameters: --broad -g hs --broad-cutoff 0.01 -q 0.01. Browser tracks are generated after normalizing to input using bamCompare default setting.

ChIP-seq Data analysis:

The sequence data were trimmed by Trimmomatic (v0.36) to remove adapter and then mapped to the hg38 assembly of the human genome using Bowtie2 with the following settings: --local --very-sensitive --phred33 -X 1000. Reads with MAPQ values less than 10 were filtered using SAMtools (v1.8). Duplicate reads were removed using picard-tools. CHIP-seq signal was converted to the bigwig format for visualization using deepTools bamCoverage (v3.3.1) with the following parameters: --binSize 10 --normalizeUsing CPM --effectiveGenomeSize 3209286105 --exactScaling.

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 Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

GRO-seq, Ribo-Zero RNA-seq CHIP-seq, WGS and KAS-seq data generated in this study can be accessed by GEO under accession number GSE249657. AmpliconArchitect (AA) outputs and copy numbers of COLO320, GBM39 and PC3 can be accessed at <https://ampliconrepository.org/project/6639560c48cbf4a5ccffad4d>.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	<input type="text" value="not applicable"/>
Reporting on race, ethnicity, or other socially relevant groupings	<input type="text" value="not applicable"/>
Population characteristics	<input type="text" value="not applicable"/>
Recruitment	<input type="text" value="not applicable"/>
Ethics oversight	<input type="text" value="not applicable"/>

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

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	<input type="text" value="No statistical methods were used to predetermine the sample size. For imaging, all the cells within 6-50 images were included into the analysis, except for cells on the edge, and experiments were repeated for 2 or more times independently. For western blot, experiments were performed at least 3 times with biologically independent replicates. For animal experiment, 8 mice was included in each group."/>
Data exclusions	<input type="text" value="No data we generated were excluded."/>
Replication	<input type="text" value="Experiments were repeated 2 or more times with similar outcome to ensure reproductivity. Animal experiment in Fig 4f was conducted once, with sample size of 8 within each group to ensure data reproductivity."/>
Randomization	<input type="text" value="To establish tumors, 1 million SNU-16 cells in 200 µL of a 1:1 mixture of PBS and Matrigel by subcutaneous injection into the right flank of 9-week-old female SCID beige mice. Animals (eight mice per group) were randomly assigned to treatment with vehicle, infigratinib (15 mg kg-1"/>

PO QD), BBI-2779 (30 mg kg⁻¹ PO Q2D), or the combination of BBI-2779 and infigratinib once average tumor volume was 286 (+/- 10) / mean (+/-SEM) mm³

Blinding

All data were collected using instruments without bias. Data within different groups were analyzed through same pipeline to ensure unbiased analysis. Blinding is not relevant to this study.

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
<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 and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

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

Antibodies

Antibodies used

ChIP-seq:
H3K36me3: Abcam, Cat# 9050, Lot # GR3459586-1

Western Blot:
pCHK1-S345 : Cell Signaling Technology, Cat# CST2348;
CHK1, Abcam, Cat# ab32531;
pRPA32/RPA2-Ser8 , Cell Signaling Technology, Cat# 54762S
γH2AX, Cell Signaling Technology, Cat# CST9718;
Vinculin , Cell Signaling Technology, Cat# CST13901
pFGFR2-Tyr653/654, Cell Signaling Technology, Cat# CST3476S
FGFR2 , Cell Signaling Technology, Cat# CST11835S;

Immunofluorescence:
γH2Ax, Millipore, Cat# 05-636,
pRPA2-S33, Novus Biological, Cat# NB100-544,
pCHK1S345, Invitrogen, Cat# PA5-34625,
53BP1, Novus Biological, Cat# NB100-304,
cyclin A, BD Bioscience, Cat# 611268,
pRNAPII S2/S4, Abcam, Cat# ab252855
Goat anti-rabbit IgG Alexa Fluor Plus 594; Thermo Fisher, #A32740s

Validation

All antibodies were validated by manufacture, and their validation statements are as follows:
H3K36me3: ChIP-grade antibody with over 900 references on abcam website and 760 reviews on Biocompare; validation based on overlap with gene body regions and depletion in non-coding regions of chromosomes; abcam validation by ChIP-qPCR

Western Blot:
pCHK1-S345, Cell Signaling Technology, Cat# CST2348; 841 citations.
CHK1, Abcam, Cat# ab32531; 11 citations. This antibody was validated by CHEK1 KO in A549 cells, using normal A549 as a control through Western Blot assay.
pRPA32/RPA2-Ser8 , Cell Signaling Technology, Cat# 54762S, 8 citations. Validated in HeLa and 293 cells, untreated or treated with UV (100 mJ/cm², 2 hr recovery) through Western Blot.
γH2AX, Cell Signaling Technology, Cat# CST9718; 2071 citations. Validated in untreated or UV-treated 293 cells through Western Blot.
Vinculin , Cell Signaling Technology, Cat# CST13901, 426 citations.
pFGFR2-Tyr653/654, Cell Signaling Technology, Cat# CST3476S, 64 citations, validated in COS cells overexpressing human FGF receptor-1, untreated or calf intestine phosphatase (CIP)-treated, using Phospho-FGF Receptor (Tyr653/654) (55H2) Mouse mAb. Overexpression of human FGF receptor-1 results in constitutive activation of the receptors.
FGFR2 , Cell Signaling Technology, Cat# CST11835S; 36 citations. Validated by Western blot analysis of extracts from KATO III cells, transfected with 100 nM SignalSilence® Control siRNA (Unconjugated) #6568 (-) or SignalSilence®FGF Receptor 2 siRNA I #12600 (+).

Immunofluorescence:
γH2Ax, Millipore, Cat# 05-636, 3402 citations. 2 μg/ml of this antibody detected phosphorylated histone H2A.X in HeLa cells treated with 0.5 μM staurosporine for 4-6 hours.
pRPA2-S33, Novus Biological, Cat# NB100-544, 26 citations. 36 h post-IR treated OVCAR-8 cells with pro-resection genes (MRE1,

BRCA1 and WDR70) knockdown by specific siRNA, were subject to immunofluorescence detection of pRPA2-S33 detection in micronuclei and primary nuclei.
 pCHK1S345, Invitrogen, Cat# PA5-34625, 5 citations. HeLa cells mock and treated with 100 J/m² UVC and recover for 8 hrs were fixed in 4% paraformaldehyde at RT for 15 min and detected Phospho-CHK1-S345 by immunofluorescence.
 53BP1, Novus Biological, Cat# NB100-304, 631 citations. Validated in 53BP1 knock out Hela cells by immunofluorescence.
 cyclin A, BD Bioscience, Cat# 611268, 34 citations.
 pRNAPII S2/S4, Abcam, Cat# ab252855, 5 citations. Validated in Hela cells treated with or without phosphatase at 37°C for 2h by immunofluorescence.

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	GBM39ec, GBM39HSR and HK296 were patient derived neurosphere cell lines and were established as previously described. COLO320DM, COLO320HSR, SNU16, PC9, RPE1 were purchased from ATCC. The parental PC3 line was obtained from ATCC. PC3 DM and PC3 HSR lines were isolated by the Mischel lab through single cell expansions of the parental PC3 line.
Authentication	Cell lines obtained from ATCC were not authenticated. PC3-DM and PC3-HSR lines were isolated by the Mischel lab through single cell expansions of the parental PC3 line from ATCC and cell identity has been authenticated by STR profiling.
Mycoplasma contamination	Cells were tested negative for mycoplasma.
Commonly misidentified lines (See ICLAC register)	None of the cell line is listed in ICLAC register of Misidentified Cell lines.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	SCID beige mice (strain code 186; Envigo, Livermore, CA), 9 week-old
Wild animals	No wild animals were used in the study
Reporting on sex	This study does not imply any sex preference. Female SCID beige mice were utilized to minimize in fighting often found with non-liter mate male mice.
Field-collected samples	No field collected samples were used in the study
Ethics oversight	Animal experiments were performed in accordance with protocols approved by the CRADL Institutional Animal Care and Use Committee (Protocol #EB17-010-066)

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

Plants

Seed stocks	<i>Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.</i>
Novel plant genotypes	<i>Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.</i>
Authentication	<i>Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.</i>

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
 May remain private before publication.

CHIP-seq data generated in this study can be accessed by GEO under accession number GSE249657

Files in database submission	COLO320, DM, input, rep1 COLO320, DM, input, rep2 COLO320, HSR, input, rep1 COLO320, HSR, input, rep2 COLO320, DM, H3K36me3, rep1 COLO320, DM, H3K36me3, rep2 COLO320, HSR, H3K36me3, rep1 COLO320, HSR, H3K36me3, rep2
Genome browser session (e.g. UCSC)	https://genome.ucsc.edu/s/gpw/hg38_ecDNA_TRC

Methodology

Replicates	Two biological replicates per sample
Sequencing depth	All CHIP-seq libraries were sequenced with paired end 75 bp reads COLO320, DM, input, rep1, 23720427 total read pairs, 15028417 uniquely mapped reads, 99.79% overall alignment rate. COLO320, DM, input, rep2, 16966012 total read pairs, 10405587 uniquely mapped reads, 99.83% overall alignment rate. COLO320, HSR, input, rep1, 22223258 total read pairs, 13514979 uniquely mapped reads, 99.86% overall alignment rate. COLO320, HSR, input, rep2, 17240809 total read pairs, 10740902 uniquely mapped reads, 99.81% overall alignment rate. COLO320, DM, H3K36me3, rep1, 21404356 total read pairs, 13757242 uniquely mapped reads, 99.81% overall alignment rate. COLO320, DM, H3K36me3, rep2, 27533870 total read pairs, 17330494 uniquely mapped reads, 99.83% overall alignment rate. COLO320, HSR, H3K36me3, rep1, 19218975 total read pairs, 12114002 uniquely mapped reads, 99.81% overall alignment rate. COLO320, HSR, H3K36me3, rep2, 16473728 total read pairs, 10468474 uniquely mapped reads, 99.80% overall alignment rate.
Antibodies	H3K36me3: Abcam, Cat# 9050, Lot # GR3459586-1
Peak calling parameters	no peak calling was needed
Data quality	no peak calling was needed
Software	Libraries were prepared using the NEBNext Ultra II DNA library prep kit (E7645) and sequenced by NovaSeq PE150. The sequence data were trimmed by Trimmomatic22 (v0.36) to remove adapter and then mapped to the hg38 assembly of the human genome using Bowtie219,20 with the following settings: --local --very-sensitive --phred33 -X 1000. Reads with MAPQ values less than 10 were filtered using SAMtools (v1.8). Duplicate reads were removed using picard-tools. CHIP-seq signal was converted to the bigwig format for visualization using deepTools bamCoverage18 (v3.3.1) with the following parameters: --binSize 10 --normalizeUsing CPM --effectiveGenomeSize 3209286105 --exactScaling.

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	Cell apoptosis was detected through flow cytometry using a FITC Annexin V Apoptosis Detection kit (BD bioscience, 556547). Cells were treated with inhibitor for the indicated time, and all the cells including floating cells were collected. After washing with PBS twice and cell number counting, cells were resuspended in 1X binding buffer, and stained with FITC Annexin V and PI for 15 min RT.
Instrument	BD LSRII flow cytometry (BD Biosciences)
Software	The image in the schema in Extended Data Fig. 8f was generated by Beckman Coulter Kaluza software
Cell population abundance	Cells were not sorted in this experiment.

Gating strategy

Major population cells were gated by FSC-A and SSC-A, and singlets were gated by FSC-A/FSC-H. Compensation degree was determined through single dye stained samples and non-stained control sample. All the singlets were further gated into Annexin V-FITC-/PI-; Annexin V+/PI-; Annexin V-/PI+; Annexin V+/PI+ populations based on non-stained control and single dye-stained samples.

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