Oncogene Silencing via ecDNA Micronucleation

Lotte Brückner[1,2,3*], Robin Xu[2,3*], Jun Tang[4,5*], Alexander Herrmann[2], Ivy Tsz-Lo Wong[4,5], Shu Zhang[4,5], Fengyu Tu[6], Madison Pilon[2], Alexander Kukalev[1,7], Katharina Pardon[2,3], Olga Sidorova[2,3], Joshua Atta[2,3], Qinghao Yu[2,3], Davide Pradella[8], Mila Ilić[9], Marco-Novais-Cruz[9], Sarah Kaltenbach[10], Denise Treue[11], Madalina Giurgiu[2,3], Sergej Herzog[1,7], Ann-Sophie Hollinger[2,3], Martina Fernandez[2,3], Finnja Becker[2,3], Varvara-Rigina Louma[2,3], Rachel Schmargon[2,3], Jan Dörr[2,3], David Gamlin[2,3], Annika Lehmann[11], Dennis Gürgen[12], Matthias Richter[13], Frank Dubois[11], Fabrizio Simeoni[14], Betheney R. Pennycook[14], Alastair Hamilton[14], Ralph K. Lindemann[14], Matthias Fischer[15,16], Vineet Bafna[17], Geoffrey Wahl[18], Richard P. Koche[19], Howard Y. Chang[20], Stamatis Papathanasiou[10], René Medema[9], Bastiaan Spanjaard[2,3], Andrea Ventura[8], Ana Pombo[1,7,21,22], Weini Huang[23], Benjamin Werner[24], Paul S. Mischel[4,5#], Anton G. Henssen[1,2,3,25#]

[1]Max Delbrück Center for Molecular Medicine, Berlin, Germany.

[2]Department of Pediatric Oncology/Hematology, Charité-Universitätsmedizin Berlin, Germany.

[3]Experimental and Clinical Research Center (ECRC) of the MDC and Charité Berlin, Berlin, Germany.

[4] Department of Pathology, Stanford University School of Medicine

[5] Sarafan ChEM-H, Stanford University

[6]Group of Theoretical Biology, The State Key Laboratory of Bio-control, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, People's Republic of China.

[7]Max-Delbrück Centre for Molecular Medicine, Berlin Institute for Medical Systems Biology, Epigenetic Regulation and Chromatin Architecture Group, Berlin, Germany

[8]Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[9]Princess Maxima Center for Pediatric Oncology

[10]Institute of Molecular Biology, Mainz, Germany.

[11]Institute of pathology, Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin Germany.

[12]Experimental Pharmacology and Oncology (EPO), Berlin, Germany.

[13]Advanced Light Microscopy, Max Delbrück Center for Molecular Medicine in the Helmholtz Association, Robert-Rössle-Straße 10, 13125, Berlin, Germany.

[14]Econic Biosciences, London, United Kingdom.

[15]Department of Experimental Pediatric Oncology, University Children's Hospital of Cologne, Medical Faculty, Cologne, Germany.

[16]Center for Molecular Medicine Cologne (CMMC), Medical Faculty, University of Cologne, Cologne, Germany
[17]University of San Diego, California, USA.
[18]Salk Institute for Biological Studies, San Diego, USA.
[19]Center for Epigenetics Research, Memorial Sloan Kettering Cancer Center, New York, NY, USA.
[20]Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA
[21]Institute of Biology, Humboldt-Universität zu Berlin, Berlin, Germany
[22]Berlin Institute of Health, Berlin, Germany
[23]Queen Mary University of London, London, UK.
[24]Evolutionary Dynamics Group, Centre for Cancer Evolution, Barts Cancer Institute, Queen Mary University London, UK.
[25]German Cancer Consortium (DKTK), partner site Berlin, and German Cancer Research Center (DKFZ), Heidelberg, Germany.
*These authors contributed equally to this work.
[#] These authors co-supervised this work.
Correspondence should be addressed to Anton G. Henssen (henssenlab@gmail.com, +49(0)30450666118) and Paul S. Mischel (pmischel@stanford.edu)

Conflict of interest

A.G.H and R.P.K are founders of Econic Biosciences. F.S., B.P., A.H., and R.L. are employees of Econic Bioscience. P.S.M. is a co-founder of Boundless Bio. He has equity and chairs the scientific advisory board, for which he is compensated. H.Y.C. is a co-founder of Accent Therapeutics, Boundless Bio, Cartography Biosciences, and Orbital Therapeutics and was an advisor to 10x Genomics, Arsenal Biosciences, Chroma Medicine, and Exai Bio until Dec 15, 2024. H.Y.C. is an employee and stockholder of Amgen as of Dec. 16, 2024. The other authors declare no potential conflicts of interest.

Abstract

Extrachromosomal DNA (ecDNA) is a common source of oncogene amplification across many types of cancer. The non-Mendelian inheritance of ecDNA contributes to heterogeneous tumour genomes that rapidly evolve to resist treatment. Here, using single-cell and live-cell imaging, single-micronucleus sequencing, and computational modelling, we demonstrate that elevated levels of ecDNA predisposes cells to micronucleation. Damage on ecDNA, commonly arising from replication stress, detaches ecDNA from the chromosomes upon which they hitchhike during cell division, thereby causing micronucleus formation in daughter cells. Clusters of oncogene-containing, CIP2A-TOPBP1-associated ecDNA molecules form, and asymmetrically segregate into daughter cell micronuclei during cell division. ecDNA chromatin remains highly active during mitosis, but upon micronucleation, it undergoes suppressive chromatin remodeling, largely ceasing oncogene transcription. These studies provide insight into the fate of damaged ecDNA during cell division.

Introduction

EcDNAs have emerged as a common driver of cancer pathogenesis, promoting high copy oncogene amplification, intratumoural genetic heterogeneity, rapid tumour evolution and treatment resistance, leading to shorter survival for patients with a wide variety of cancers[1-3]. EcDNAs are circular DNA structures that contain oncogenes and other tumour-relevant genetic elements, having separated from chromosomes and existing independently in the cell. Unlike chromosomes, ecDNAs lack centromeres, allowing them to segregate randomly during mitosis, contributing to tumour heterogeneity. This feature enables ecDNA to accumulate rapidly across successive cell divisions, amplifying oncogenes in the process[4-6]. Additionally, ecDNAs possess highly accessible chromatin, altered cis- and trans- gene regulation[7-10] and exhibit increased levels of active histone modifications, leading to elevated transcriptional activity compared to linear chromosomal amplifications, such as homogeneously staining regions (HSRs)[8,9,11]. These molecular characteristics contribute to the aggressive nature of ecDNA-driven cancers and correlate with poor patient survival outcomes[1,2,10]. As such, ecDNA represents an attractive target for therapeutic intervention in cancer.

Despite their acentric nature, ecDNAs are generally thought to be replicated once per cell cycle and are remarkably faithful in their mitotic segregation to daughter cells[12,13]. This reliable transmission is believed to result from ecDNA "hitchhiking" with chromatids during

anaphase[14-16]. However, ecDNA can also be sequestered into micronuclei—abnormal, small nuclear structures that arise from mis-segregated chromosomes or acentric fragments[14,16]. DNA-damaging agents further promote such micronucleation[17-19]. Micronucleation can play a role in promoting ecDNA formation when lagging chromosomes undergo chromothripsis and critical fragments circularize[20,21]. The impact of micronucleation once ecDNAs have been formed is only partially understood.

Hydroxyurea treatment can promote the transit of ecDNA into micronuclei [18,19,22]. The impact of endogenous DNA damage is less well understood. Recent work demonstrates high levels of DNA damage on ecDNA, deriving largely from transcription-replication collisions[23], but the impact of this on micronucleation is unknown. The diverse potential fates for ecDNA during mitosis raise important questions: 1) How faithful is ecDNA segregation during mitosis? 2) Is mis-segregation increased when ecDNA is damaged? 3) How does ecDNA micronucleation affect its mitotic inheritance? 4) How does mis-segregation of ecDNA into micronuclei influence oncogene regulation and transcription? Here, we seek to address these fundamental questions regarding the unique mitotic dynamics of ecDNA.

Results

Micronucleation is a hallmark of cancers with high ecDNA load

Micronuclei are a well-known feature of malignant cells, yet their frequency varies across cancers[24]. We hypothesised that ecDNA, as acentric chromatin, increases micronucleation propensity. To test this, we performed interphase fluorescence *in situ* hybridisation (FISH) in cancer cell lines (n=15) including near-isogenic cell line pairs (n=4), drug-resistant cell lines (n=2), and patient tumour-derived samples (n=236). Semi-automated image analysis revealed that tumour cells harbouring ecDNA exhibited significantly higher micronucleation rates than those with intrachromosomal HSR or no amplifications (Fig. 1a-d, Extended Data Fig. 1a-e). Micronuclei observed in ecDNA+ cells were in general smaller than those in HSR cells, suggesting the presence of only one or few ecDNA copies (Extended Data Fig. 1f). A binominal segregation model estimated an ecDNA mis-segregation probability of 1.7% (range between 0 and 7.8%, Extended Data Fig. 1g), indicating efficient retention during mitosis. To test whether ecDNA alone is sufficient to increase micronucleation, we examined Cre-expressing $p53^{fl/fl};Myc^{ec/+}$ mouse adult neural stem cells (aNSCs) harbouring engineered *Myc*-containing ecDNAs. $p53fl/fl;Myc^{OE}$ control cells with chromosomally integrated *Myc* showed comparable Myc expression levels (Fig. 1e), yet only ecDNA+ cells exhibited increased micronucleation (Fig. 1f), confirming that ecDNA but not its expression drives micronucleation.

We next assessed whether micronucleation frequency correlates with ecDNA abundance. Quantification via FISH revealed that cells with micronuclei contained significantly more ecDNA copies than those without (Fig. 1g-j, Extended Data Fig. g-j). Fixed-cell staining confirmed that lagging ecDNA in anaphase correlated with ecDNA copy number (Fig. 1k-l). Live-cell imaging similarly showed that ecDNA decorated anaphase bridges (Extended Data Fig. 2a). This correlation persisted across EdU-positive and -negative cells (Extended Data Fig. 2b-d), ruling out cell cycle arrest as a confounder. Consistent with prior findings[23], high levels of replication stress and DNA damage on ecDNA were observed (Extended Data Fig. 2e), raising the question of whether it contributes directly to mis-segregation and subsequent micronucleation. Collectively, these results demonstrate that ecDNA drives high micronucleation rates, which may be due to their detachment from chromosomes in anaphase, a process that could be driven by replication stress.

Replication stress and DNA damage promote ecDNA micronucleation

We induced replication stress with hydroxyurea to test its effect on ecDNA micronucleation (Fig. 2a). In immunofluorescence and DNA FISH assays we observed that hydroxyurea increased replication stress and DNA damage in both ecDNA and HSR isogenic cell lines, but micronucleation frequency was significantly higher in ecDNA+ cells (Fig. 2b-c, Extended Data Fig. 3a-b). dNTP supplementation reduced this effect (Extended Data Fig. 3c), confirming replication stress as the cause of ecDNA mis-segregation. Micronucleated cells contained greater ecDNA content than intact cells (Extended Data Fig. 3d-f). This effect required long-term hydroxyurea exposure, suggesting that cell division is required for ecDNA micronucleation (Fig. 2b-c). Hydroxyurea also increased micronucleus size and amplicon content (Fig. 2d), suggesting that damaged ecDNAs may collectively mis-segregate. Consistent with earlier findings[7,25], co-clustering of ecDNA and γH2AX foci were observed at the nuclear periphery within 24 hours and persisted through mitosis, correlating with ecDNA detachment from chromatids (Extended Data Fig. 3g-h). In contrast, HSR cells exhibited minimal lagging acentric fragments (Extended Data Fig. 3i-j).

If DNA damage drives ecDNA micronucleation, it should also increase in patients undergoing genotoxic therapy. Indeed, micronucleation frequency was higher in *MYCN*-amplified ecDNA+ neuroblastomas post-treatment (Fig. 2e, Extended Data Fig. 4b). To directly test whether DNA damage on ecDNAs is sufficient to drive micronucleation, we used CRISPR-Cas9 to introduce targeted DNA damage at amplicon loci (Fig. 2f-g). CRISPR-mediated cleavage increased damaged *MYC* with partial DNA damage recovery after 48-72 hours

(Extended Data Fig. 4a). Targeted DNA damage promoted ecDNA micronucleation (Fig. 2h-i). These effects became more evident after multiple cell cycles, pointing at damage-induced ecDNA detachment from chromosomes as a possible cause of mis-segregation. Indeed, CRISPR-targeted cells exhibited significantly more lagging ecDNAs during anaphase than controls (Fig. 2j-k). This demonstrates that DNA damage is sufficient to drive ecDNA micronucleation by increasing lagging ecDNA during mitosis. The consistency of this observation across cancer cell lines and patient tumour underscores its general relevance.

Collective mis-segregation of ecDNA into micronuclei

The correlation between ecDNA copy number and micronucleation prompted us to explore whether ecDNA mis-segregation is stochastic or directed. To address this, we developed a laser microdissection-based technique to isolate single micronuclei and primary nuclei followed by paired-end short-read sequencing (Fig. 3a). This was performed in TR-14, a cell line containing multiple ecDNA species (*MYCN*, *MDM2*, *CDK4*, *ODC1*, *SMC6*) following hydroxyurea treatment. Our analysis aimed to determine the abundance and stoichiometry of different ecDNA species in micronuclei. All five ecDNA species were detected in most single micronuclei and primary nuclei (Fig. 3b). Micronuclei showed significant enrichment of ecDNA compared to flanking regions and the linear genome (Fig. 3b,c, Extended Data Fig. 5a), suggesting that ecDNA mis-segregation occurs independently of chromosomal origins and at higher frequencies than linear intrachromosomal elements. While most micronuclei and primary nuclei contained all ecDNA species, some contained only subsets (Extended Data Fig. 5b). Notably, stoichiometry varied between micronuclei and primary nuclei (Fig. 3d), indicating ecDNA species-specific mis-segregation patterns influenced by unknown retention factors. To validate this, FISH in two additional cancer cell lines confirmed significant ecDNA enrichment in micronuclei (Fig. 3e). These findings suggest that ecDNA mis-segregation occurs *en masse* in cells with elevated ecDNA levels, leading to frequent ecDNA transit into micronuclei.

CIP2A-TOPBP1 facilitates collective ecDNA mis-segregation

En masse micronucleation of ecDNA may result from its aggregation via DNA damage, as previously proposed[3,25]. To test this, we analysed the sequence of events leading to micronucleation. DNA damage accumulation in interphase preceded ecDNA detachment during mitosis, which was followed by micronucleation (Extended Data Fig. 6a). However,

γH2AX was absent in some detached ecDNA clusters and micronuclei, suggesting that damage persistence is not required for *en masse* mis-segregation (Extended Data Fig. 6b).

We hypothesized that ecDNA clustering could involve homologous recombination intermediates. RAD51, a key homologous recombination factor, promotes DNA strand exchange between DNA sequences, potentially aiding cluster formation. If this was the case, RAD51 inhibition should reduce ecDNA clustering and micronucleation. Indeed, pharmacologic RAD51 inhibition with B02 reduced hydroxyurea-induced micronucleation in ecDNA+ cells to control levels (Extended Data Fig. 6c,d), whereas non-homologous end joining inhibition (AZD7648) had no effect (Extended Data Fig. 6c), confirming the specific role of homologous recombination.

Cells typically stabilize mitotic double-strand breaks (DSBs) for repair in the next cell cycle[26-28]. The CIP2A-TOPBP1 complex is a mediator of this stabilization of chromosomal fragments[22,29-31], leading us to test its role in ecDNA clustering. Immunofluorescence staining revealed that detached ecDNA clusters were associated with CIP2A and TOPBP1 (Fig. 3f), which was significantly enhanced following hydroxyurea exposure (Fig. 3g). Inhibiting CIP2A stability with TD19 reduced CIP2A expression, dispersed ecDNA clusters into single detached ecDNA and increased the number of small micronuclei per cell without affecting the fraction of micronucleated cells (Fig. 3h-i, Extended Data Fig. 6f-j). These findings suggest that CIP2A-TOPBP1 stabilizes ecDNA clusters during mitosis, facilitating *en masse* mis-segregation of ecDNA into micronuclei.

Micronucleation leads to asymmetric mitotic ecDNA inheritance

To assess the impact of *en masse* ecDNA segregation on inheritance, we performed live-cell imaging of ecDNA during mitosis. EcDNA clusters persisted throughout mitosis, positioned between daughter cells in anaphase, and ultimately segregated into a single daughter cell, leading to asymmetric inheritance (Fig. 3j). This suggests that ecDNA micronucleation skews previously observed random inheritance patterns[4]. To quantify this, we measured ecDNA copy number in paired daughter cells during cytokinesis using DNA FISH following hydroxyurea treatment. Hydroxyurea-induced *en masse* segregation led to significantly more biased ecDNA distribution between daughter cells compared to controls (Fig. 3k).

To further analyse asymmetric segregation, we developed a computational model (Fig. 3l), assuming a probability q of asymmetric segregation and probability 1-q of random segregation. Random segregation follows a binomial distribution with a segregation probability $p_2 = 0.5$, while biased segregation exhibited a cluster-based partitioning with $p_1 < 0.5$, reflecting inherent

asymmetry. The resulting ecDNA distribution in daughter cells shifted from unimodal (random segregation) to bimodal or trimodal patterns as q and p_1 varied (Extended Data Fig. 7a-e). Using an approximate Bayesian computation framework (ABC), we inferred optimal parameter estimates for p_1 and q (Extended Data Fig. 7f). Applying this to our experimental data from 50 pairs of daughter cells with and without hydroxyurea treatment, we observed increased biased segregation post-treatment (Fig. 3m-o). Thus, replication stress enhances ecDNA segregation asymmetry, increasing intercellular variability in ecDNA content.

Micronucleation of ecDNA reduces oncogenic transcription via epigenetic remodelling

Oncogene amplification via ecDNA drives high oncogene expression[6,7,32]. Consistent with previous reports[17,18], hydroxyurea-treated ecDNA+ cells exhibited significantly lower oncogene expression (Extended Data Fig. 8a-b). Since ecDNA copy number did not decrease after short term hydroxyurea exposure (Extended Data Fig. 8c), we hypothesised that ecDNA micronucleation represses transcription through epigenetic remodelling, similar to mis-segregated linear chromatin[33,34].

EcDNA transcription is closely linked to its chromatin landscape, characterized by high H3K27ac and low H3K27me3[4,8,9]. To determine whether micronucleation disrupts this, we analysed histone modifications in ecDNA+ and HSR+ cells (Fig. 4a). During mitosis, ecDNA retained high H3K27Ac and low H3K27me3, much unlike chromosomal DNA, which showed the expected inactive chromatin patterns (Fig. 4b-c). Lagging ecDNA clusters exhibited no major differences in H3K27Ac and H3K27me3, indicating that clustering, detachment, or DNA damage alone does not alter ecDNA chromatin (Extended Data Fig. 9a). However, micronucleated ecDNA showed significantly reduced H3K27Ac and stably low H3K27me3 compared to ecDNA in mitosis and the primary nuclei following hydroxyurea treatment (Fig. 4d-e, Extended Data Fig. 10a). Chromatin immunoprecipitation sequencing confirmed reduced H3K27ac at ecDNA loci, including *MYC* (Fig. 4f, Extended Data Fig. 10b-d), suggesting that micronucleation drives epigenetic ecDNA remodeling and transcriptional repression.

To assess transcriptional consequences, we measured active RNA polymerase II (pRNAPII-S5) via immunofluorescence staining. Micronuclei exhibited significantly lower levels of pRNAPII-S5, indicating reduced transcriptional activity in micronuclei (Fig. 4g-h). Intron RNA-FISH further confirmed that most ecDNA-containing micronuclei lacked detectable intron RNA-FISH signals, independent of the oncogene identity (Fig. 4i). RNA sequencing following hydroxyurea treatment showed decreased expression of ecDNA-encoded genes, including *MYC*, alongside reduced MYC-driven transcriptional programs (Fig. 4j-k, Extended

Data Fig. 10e-f). Together, these findings demonstrate that ecDNA micronucleation leads to epigenetic remodelling and oncogene silencing.

Discussion

This study elucidates the role of ecDNA mis-segregation into micronuclei and its implications for oncogene regulation. We reveal that high ecDNA content correlates with increased micronucleation, primarily driven by replication stress and DNA damage. This process is not only a hallmark of aggressive cancers but also silences oncogene expression (Fig. 4l), suggesting a role in enabling genetic and epigenetic plasticity that may allow cancer cells to dynamically regulate oncogenic programs.

The enrichment of ecDNA in micronuclei suggests that its acentric nature predisposes it to mis-segregation, especially under replication stress. The involvement of homologous recombination repair in ecDNA detachment hints at additional, yet undefined molecular pathways influencing this process, warranting future investigation. Our findings further demonstrate that ecDNA mis-segregation largely occurs *en masse* rather than stochastically, thus introduce a novel mechanism by which DNA-damaging therapies alter oncogene copy number heterogeneity.

The transcriptional repression of ecDNA in micronuclei suggests a fundamental regulatory shift. While ecDNA remains transcriptionally active in mitosis[35], micronucleation triggers epigenetic remodelling primarily through H3K27ac loss, while H3K27me3 levels remain unchanged. This distinguishes ecDNA from linear acentric fragments, which undergo distinct chromatin changes upon missegregation[33,34]. These findings suggest that ecDNA micronucleation may not only offer a novel route for oncogene silencing in tumors with high ecDNA content but could also reflect a broader mechanism through which cancer cells achieve epigenetic plasticity. Understanding the mechanism sustaining ecDNA transcription until micronucleation could identify new therapeutic opportunities, particularly those targeting replication stress.

Our work also implicates the CIP2A-TOPBP1 complex in stabilizing damaged ecDNA clusters, revealing a potential therapeutic target. Disrupting this complex could enhance ecDNA mis-segregation or expose it to cytoplasmic DNA sensing[36], increasing its vulnerability to genotoxic therapies. Further studies are needed to dissect the molecular mechanisms governing ecDNA detachment from linear chromosomes during mitosis, which could uncover new intervention specific to ecDNA+ cancers.

Beyond fundamental cancer biology, our findings suggest that micronucleation is a hallmark of ecDNA-harbouring cancers and might serve as a biomarker for identifying patients most likely benefiting from therapies that exploit ecDNA vulnerabilities.

Methods

Cell lines

Human tumour cell lines were obtained from the American Type Culture Collection (ATCC) or a gift from collaborative laboratories. The identity of all cell lines was verified by short tandem repeat genotyping (Eurofins Genomics). The absence of *Mycoplasma* sp. contamination was determined using a Lonza MycoAlert system (Lonza). For this study, 10 ecDNA amplified cell lines (STA-NB10DM, COLO320DM, TR-14, CHP-212, Lan5, UKF-NB6, KP-N-YN, PC3, SiMA, SMS-KAN), 4 HSR amplified cell lines (STA-NB10HSR, COLO320HSR, NGP, IMR5/75), 2 non amplified cancer cell line (ShEP, HeLa) and 1 non-cancerous cell line (RPE-hTERT) were used. HEK293T cells were used for virus production. Near isogenic cell line pairs STA-NB10DM, STA-NB10HSR[37], GBM39ec, GBM39HSR, COLO320DM, COLO320HSR, PC3 DM, PC3 HSR cells were extensively characterized before[23].

Cell culture

Most neuroblastoma and colon cancer cell lines were cultured in RPMI-1640 medium (Gibco) supplemented with 1 % of penicillin, streptomycin, and 10 % of fetal calf serum (FCS) (Thermo Fisher). STA-NB10DM and HSR cells were cultured in RPMI-1640 medium (Gibco) supplemented with 1 % of penicillin, streptomycin, 10 % of fetal calf serum (FCS) (Thermo Fisher), 1 % Glutamax (Thermo Fisher), 1 % Sodium Pyruvat (Thermo Fisher) and 10 mM HEPES (Sigma). UKF-NB6 was cultured in IMDM medium (Gibco) supplemented with 1 % of penicillin, streptomycin, and 10 % of FCS. RPE, PC3 and HEK293T cells were cultured in DMEM (Gibco) supplemented with 1 % of penicillin, streptomycin, and 10 % of FCS. GBM39ec and GBM39HSR cell lines were cultured in Dulbecco's Modified Eagle's Medium/F12 (Gibco, catalogue no. 11320-033) supplemented with 1× B27 (Gibco, catalogue no. 17504-01), 20 ng/ml epidermal growth factor (Sigma, catalogue no. E9644), 20 ng/ml fibroblast growth factor (Peprotech, catalogue no. AF-100-18B), 5 µg/ml heparin (Sigma, catalogue no. H3149) and 1× GlutaMAX (Gibco, catalogue no. 35050-061). COLO320DM, COLO320HSR, PC3-DM, PC3-HSR and COLO320DM-Tg19 (live-cell imaging line) cells presented in figure 2, 3j and 4i were cultured in DMEM (Corning, #10-013-CV) with 1% PSQ and 10% FBS. To assess the number of viable cells, cells were trypsinized (Thermo Fisher), resuspended in medium, and sedimented at 300 *g* for 5 minutes. STA-NB10 cells were split using Accutase (Thermo Fisher). HeLa cells were cultured in Advanced DMEM-F12 (Gibco) with addition of 10% FCS, 1% Penicillin-Streptomycin (Gibco) and 1% GlutaMax (Gibco).

Cells were then resuspended in medium, mixed in a 1:1 ratio with 0.02 % trypan blue (Thermo Fisher), and counted with a Bio-Rad TC20 cell counter. In drug experiments, cells were treated for 24 h with 80 μM hydroxyurea or a combination of 80 μM hydroxyurea and one of the following: 1 μM TD-19 (Merck Millipore, 532912), 15 μM B02 (Merck Millipore, SML0364-5MG), 10 μM mirin (MedChemExpress, HY-117693), 5 μM AZD7648 (MedChemExpress, HY-111783), 10 μM NSC16168 (Adooq Bioscience, A13260), 10 μM LB-100 (Selleck Chemicals GmbH, S7537). Control cells were treated with an equivalent volume of DMSO or H2O. Drug concentrations were selected based on cell viability as assessed using the CellTiter-Glo Assay (Promega G7571). In time-course experiments, cells were treated with 80 μM hydroxyurea (Sigma-Aldrich, H8627-5G) for 6 h, 12 h, 24 h, 36 h, 48 h, 72 h, 96 h or with DMSO for 96 h. GBM39ec/GBM39HSR cells were treated with 100 μM hydroxyurea for 5 and 72 hrs, EdU (10 μM) was added for 30 mins before fixing.

Drug induced ecDNA cell lines

Drug-resistant cell lines were derived by treating HeLa parental cells with increasing concentration of either paclitaxel (T1912, Sigma-Aldrich) or methotrexate (A6770, Sigma-Aldrich). Then, the stable cell lines are propagated with addition of 100nM of paclitaxel or 600nM of methotrexate. To score micronuclei, asynchronous cells were fixed with 4% paraformaldehyde (P6148, Sigma Aldrich) and permeabilized with 0.5% Triton X-100(9036-19-5, Sigma Aldrich). After a brief wash with PBS, cells were mounted with ProLong™ Gold Antifade Mountant with DAPI (P3693, Thermo Fisher). Images were acquired using a THUNDER Imager equipped with a 63×/1.40–0.60 OIL Obj. HC PL APO objective and a deep-cooled 4.2 MP sCMOS camera (Leica Microsystems). Ten z-planes with a 0.75 μm step were collected. Micronuclei were scored manually.

Genetically induced ecDNA cell lines

Adult Neuronal Stem Cells (aNSCs) were isolated from Mycec/+; p53fl/fl and p53fl/fl animals described in Pradella et al.[37] and cultured following the protocol described by Ahmed et al.[38] aNSC were maintained in laminin-coated (Sigma-Aldrich, L2020) dishes in NeuroCult Stem Cell Basal Media with NeuroCult Proliferation Supplement (Mouse & Rat) (Stem Cells Technologies, 05702), 20 ng ml−1 EGF (Stem Cells Technologies, 78006), 10 ng ml−1 bFGF (Stem Cells Technologies, 78003), and 2 μg ml−1 heparin (Stem Cell Technologies, 07980). Cells were tested for mycoplasma contamination and maintained in a humidified, 5% CO2 atmosphere at 37 °C. aNSC overexpressing Myc transgene were generated by infecting p53fl/fl

aNSC with pBABE-Puro-Myc retrovirus. Mycec/+; p53fl/fl and p53fl/fl aNSC were also infected with pBABE-Puro retrovirus. Upon infection, cells were selected with puromycin (1 μg ml−1; Sigma-Aldrich, P9620) for 3-4 days. Selected cells were then infected with AdCre (ViraQuest, VQ-Ad-CMV-Cre; 1 × 1012particles per milliliter; 091317) to promote p53 loss and Myc-containing ecDNA accumulation as described in Pradella et al.[38]

Retrovirus production

pBABE-Puro-C-Myc plasmid was generated by site-directed mutagenesis from pBABE-Puro-C-Myc T58A (Addgene #53178) by removing the T to A mutation in the mouse coding region. For retrovirus production, HEK-293T cells were seeded in high-glucose Dulbecco's modified Eagle medium (DMEM, 4.5 g l−1; Thermo Fisher Scientific, 11995065) supplemented with 10% heat-inactivated fetal bovine serum (FBS, Sigma-Aldrich, F2442) without antibiotics in 100 mm petri dishes. The next day, cells at approximately 60–70% confluence were transfected with 20 μg of retroviral vector carrying the mouse C-Myc cDNA or a control vector, 5 μg of packaging plasmid and 1 μg of envelope plasmid. After 24 h, the medium was replaced with DMEM with antibiotics. Cells were incubated for 48 h with two consecutive collections of the medium containing the retroviral particles at 24 and 48 h. Medium collected was filtered using a 0.45 μm filter unit and used to transduce aNSC cells by incubating cells with viral supernatant supplemented with polybrene (0.2 μl ml−1; Sigma, TR-1003G).

Antibodies and immunoblots of aNSC

aNSC cells were lysed in Laemmli buffer, supplemented with protease (cOmplete; Roche, COEDTAF-RO) and phosphatase (EDTA-free Protease Inhibitor Cocktail; Roche, PHOSS-RO) inhibitors. Proteins were separated by SDS polyacrylamide gel electrophoresis and analyzed by western blotting using standard procedures. After protein transfer, the nitrocellulose membranes (Bio-Rad, catalogue no. 1704271) were blocked by incubation with LICOR Intercept (TBS) blocking buffer. The following primary antibodies were used: anti-vinculin (1:5,000; Millipore, MAB3574), anti-c-MYC (1:1,000; Cell Signaling, D84C12). The following secondary antibodies were used: IRDye 800 anti-rabbit (LICOR, 926-32213) and IRDye 680 anti-mouse (LICOR, 926-68072). Images were acquired using an Odyssey Imaging System (LICOR).

Imaging of aNSCs

aNSC were fixed with ice-cold methanol for 20 min. Briefly, fixed slides were pretreated with Pepsin (Sigma P-6887), dehydrated and counterstained with DAPI (Thermo, D1306) and mounted with Prolong Diamond Antifade Mounting medium (Thermo, P36965). Images were acquired with a REVOLVE R4 (Echo Laboratories) microscope through a ×63 objective lens. Images were analyzed in the built-in Echo app. Each nucleus was evaluated for the presence of micronuclei and Myc amplification.

Western immunoblotting

Whole-cell protein lysates were prepared by lysing cells in 10x Cell lysis buffer (Cell signaling) supplemented with cOmplete (Roche) and PhosStop (Roche) phosphatase inhibitors. Protein concentrations were assessed by bicinchoninic acid assay (BCA, Santa Cruz Biotechnology). For 5 minutes, 10 µg of protein was denatured in Laemmli buffer at 90 °C. Samples were run on NuPage 10 % polyacrylamide, 1 mm Tris-Glycine Protein Gels (Thermo Fisher Scientific) and transferred to PVDF membranes (Roche). Membranes were blocked with 5 % dry milk in TBS with 0.1 % (v/v) Tween-20 (Carl Roth). Membranes were probed with primary antibodies overnight at 4 °C and then with secondary antibodies conjugated to horseradish peroxidase for 1 hour at room temperature. Chemiluminescent detection of proteins was carried out using Pierce ECL Western blotting substrate (Thermo FIisher) and the ChemiDoc (Bio-Rad). Densitometry was performed using ImageJ (NIH).

Targeted DNA damage induction by CRISPR-cas9

To induce targeted DNA damage on amplicon, we used CRISPR-cas9. crRNA was designed to target an intergenic region amplified both in COLO320DM and COLO320HSR cells (sequence: chr8:+127572615: TACAACGAACTATTTAACCG, TGG(pam)). Non-targeting control was also included (IDT, cat: 1072544). Equal amounts of crRNA and tracerRNA were mixed and annealed before mixing with Alt-R™ S.p. Cas9 Nuclease V3, (IDT, cat:1081059) to make the RNP complex. RNP complex was stabilized by incubating at room temperature for 15 minutes, after which it was delivered into COLO320DM and COLO320HSR cells using electroporation with the Nucleofector system from Lonza, using program CM-138 and solution SF (Lonza, Catalog #: V4XC-2012). After electroporation, cells were immediately seeded onto coverslips previously coated with fibronectin (10 µg/ml). Cells were fixed at indicated time points and subjected to multiplex immunofluorescence and DNA-FISH staining.

Preparation of metaphase spreads

Cells were grown to 80% confluency in a 15 cm dish and metaphase-arrested by adding KaryoMAX™ Colcemid™ (10 µL/mL, Gibco) for 1-2 hours. Cells were washed with PBS, trypsinized and centrifuged at 200 g for 10 min. We added 10 mL of 0.075 M KCl preheated at 37 °C, one mL at a time and vortexing at maximum speed in between. Afterwards, cells were incubated for 20 min at 37 °C. 5 mL of ice-cold 3:1 MeOH/acetic acid (kept at -20 °C) were added, one mL at a time followed by resuspension of the cells by flicking the tube. The sample was centrifuged at 200 g for 5 min. Addition of the fixative followed by centrifugation was repeated four times. Two drops of cells within 200 µL of MeOH/acetic acid were dropped onto prewarmed slides from a height of 15cm. Slides were incubated overnight.

Fluorescence in situ hybridisation (FISH)

Metaphase spreads or cells seeded on coverslips for interphase FISH were fixed in MeOH/acetic acid for 20 min at -20 °C followed by a PBS wash for 5 min at RT. The wells were removed, and slides were digested in Pepsin solution (0.001 N HCl) with the addition of 10 µl pepsin (1 gr/50 mL) at 37°C for 10 min. After a wash in 0.5x SSC for 5 min, slides were dehydrated by washing in 70 %, 90 % and 100 % cold ethanol stored at -20 °C (3 min in each solution). Dried slides were stained with either a 5 µl of Vysis LSI N-MYC SpectrumGreen/CEP 2 SpectrumOrange Probes (Abbott), ZytoLight ® Spec CDK4/CEN12 Dual Color Probe (ZytoVision) or ZytoLight ® SPEC MDM2/CEN 12 Dual Color Probe (Zytovision), covered with a coverslip and sealed with rubber cement. Denaturing occurred in a Thermobrite (Abbott) for 5min at 72 °C followed by 37 °C overnight. The slides were washed for 5 min at RT within 2× SSC/0.1 % IGEPAL, followed by 3 min at 60 in 0.4× SSC/0.3 % IGEPAL (Sigma-Aldrich Inc.) and further 3 min in 2× SSC/0.1 % IGEPAL at RT. Dried slides were stained with 12 µl Hoechst 33342 (10 µM, Thermo Fisher) for 10 min and washed with PBS for 5 min. After drying, a coverslip was mounted on the slide and sealed with nail polish. Images were taken using a Leica Stellaris Confocal microscope. For ecDNA copy number estimation, we counted foci using FIJI's find maxima function. Nuclear boundaries were marked as regions of interest. The threshold for signal detection within the regions of interest was determined manually and used for all images analyzed within one group.

Indirect Immunofluorescence (IF)

Cells were fixed in 4% paraformaldehyde (Electron Microscopy Science) prepared in PHEM buffer (Sigma-Aldrich) for 10 minutes at room temperature, with fresh fixative applied after the first 2 minutes. Fixed cells were washed three times in 125 mM glycine/PHEM for 5

minutes each, followed by permeabilization with 0.5% Triton X-100 (Sigma-Aldrich) in PHEM for 10 minutes at room temperature. After a single 5-minute wash in PHEM, cells were blocked in 5% fetal calf serum (FCS, Sigma-Aldrich) containing 0.1% Tween-20 (Bio-Rad) in PHEM for 30 minutes at room temperature. Primary antibodies were diluted in 5% FCS/0.1% Tween-20/PHEM and applied to cells overnight at 4 °C. The following day, cells were washed three times in 0.1% Tween-20/PHEM for 10 minutes each and incubated with secondary antibodies diluted in 5% FCS/0.1% Tween-20/PHEM for 2 hours at room temperature. After incubation, cells were washed three additional times in 0.1% Tween-20/PHEM for 10 minutes each. Nuclei were counterstained with 5 µM Hoechst 33342 (Thermo Fisher Scientific) in PHEM for 10 minutes, followed by three washes in PHEM for 5 minutes each. Cells were air-dried for 5 minutes and mounted onto SuperFrost glass slides (Epredia) using 5 µl of ProLong Glass Antifade Mountant (Thermo Fisher Scientific). For double immunofluorescence, both primary and secondary antibodies were applied simultaneously.

Combined FISH and immunofluorescence

Interphase cells were plated on Poly-D-lysine-coated coverslips (Neuvitro Corporation) in 24-well plates. Cells were fixed with 4% paraformaldehyde for 10 minutes at room temperature, followed by three washes with 125 mM glycine in PBS (5 minutes each). Permeabilization was performed using 0.5% Triton X-100 in PBS for 10 minutes at room temperature, followed by three washes with PBS. Cells were then blocked in 3% bovine serum albumin (BSA) for 30 minutes and incubated overnight at 4 °C with the primary antibody diluted in 3% BSA. After three additional PBS washes, cells were incubated with the secondary antibody (diluted in 3% BSA) for 2 hours at room temperature. Post-incubation, cells were re-fixed in 4% paraformaldehyde for 20 minutes at room temperature. DNA was subsequently denatured by treating cells with 0.7% Triton X-100 and 0.1 M HCl on ice, followed by additional chemical denaturation in 1.9 M HCl for 30 minutes at room temperature. The slides were then washed with 125 mM glycine in PBS, followed by a 5-minute wash in 2× saline-sodium citrate (SSC) buffer. Dehydration and probe hybridization was conducted following the standard FISH protocol described above.

Intron RNA FISH

The Stellaris RNA FISH system (LGC Biosearch Technologies) was used following the manufacturer's protocol for adherent cells. Intron probe design was generated using the

Stellaris Probe Designer tool under these settings: maximum number of probes = 48, oligo length = 20, minimum spacing length = 2. Sequences of *MYC* intron 2 and *EGFRvIII* intron between exon 11 and 12 were used for probe design, generating a final probe design consisting of 31 probes (*MYC*) and 47 probes (*EGFRvIII*) respectively.

Live-cell imaging

Live-cell imaging was performed as previously described (cite co-segregation). The stably engineered COLO320DM cell line (cite co-segregation) was used to conduct live-cell imaging. Before cell seeding, 96-well glass-bottom plates were coated using poly-D-lysine. The medium was replenished with FluoroBrite DMEM (Gibco, A1896701), with 10% FBS, 1x GlutaMax and prolong live antifade reagent diluted in 1:200 ratio (Invitrogen, P36975) 30 minutes prior to imaging. The 96-well plate was fitted onto a top stage incubator (Okolab) and imaged on a Leica DMi8 widefield microscope using a x63 oil objective. Temperature (37˚C), humidity and CO2 (5%) were stably maintained throughout the experiment.

Microscopy, image acquisition and analysis

All images were acquired using a Leica Stellaris 8 confocal microscope at the Advanced Light Microscopy Technology Platform (Max Delbrück Center for Molecular Medicine). The system was equipped with an HC PL APO CS2 63×/1.40 OIL objective lens and used an immersion oil medium with a refractive index of 1.5185.

For imaging individual anaphase cells, the acquisition settings were as follows: 512 × 512 pixels frame size, 6× zoom, 0.79 µs pixel dwell time, and a final pixel size of 0.06 µm × 0.06 µm. Z-sampling was set to 0.23 µm to capture the entire cell in three dimensions.

For groups of interphase cells, images were acquired with a 2048 × 2048 pixels frame size, 2× zoom, 3.16 µs pixel dwell time, and a final pixel size of 0.05 µm × 0.05 µm.

Fluorophores were excited using diode laser lines at 561 nm (Alexa Fluor 568), 488 nm (Alexa Fluor 488), and 405 nm (Hoechst). Fluorescence emissions were detected at 566–620 nm (Alexa Fluor 568), 500–550 nm (Alexa 488), and 420–480 nm (Hoechst) using a pinhole set to 1 Airy unit (AU).

To ensure comparability across samples, all images were acquired under identical instrument settings.

The “Find Maxima” function was employed to quantify extrachromosomal DNA (ecDNA) copy numbers in FISH images. In anaphase cells, lagging DNA was defined as chromosomal material visibly separated from the primary segregating chromosomes by a minimum distance

of 0.3 µm across the Z-stack, consistent with the lateral resolution limit of the imaging system. Images for Fig 2 b-e were taken by a ZEISS LSM 880 inverted confocal microscope using ZEN (black v.2.3), and images for Fig 2 g-k were taken by a Leica DMi8 widefield microscope by Las X software (v.3.8.2.27713) using a ×63 oil objective. Z-stacks were taken for each field of view and a max projection was performed by ImageJ (v.1.53t) before proceeding for further image analysis. Individual splitted channel images were semi-automatically segmented and analyzed by CellProfiler (v.4.2.1).

Images presented in Figures 3J and 4I were acquired on a Leica DMi8 widefield microscope using a x63 oil objective. Z-stack images were acquired, which were then computationally deconvoluted using Small Volume Computational Clearing on the LAS X software prior to outputting as maximum intensity projections. RNA FISH foci were manually called from all micronuclei.

Laser-microdissection and Whole Genome Amplification (WGA) of single (micro-)nuclei

TR-14 cells were directly grown on Poly-D-Lysine coated (Gibco, A3890401) PEN-membrane metal frame slides (Leica, 11600289). Immunofluorescence was performed as described above. Samples were incubated with mouse anti-pan-histone H11-4 (1:500; Merck, MAB3422), followed by incubation with Alexa555 (1:1000; Invitrogen, A-31570) or Alexa488 anti-mouse secondary antibody (1:1000; Invitrogen, A-11001). Slides were stored in 1x PHEM in the dark at 4°C until dissection. Single micronuclei, pooled micronuclei (5-10 micronuclei) and single primary nuclei were manually visualised and identified on a Leica microdissection microscope (LMD7 or LMD7000, Leica microsystems). Shapes were collected into adhesive PCR caps (MicroDissect GmbH, MDCA8W1K20). WGA was performed as in the ImmunoGAM method, with minor modifications[39,40]. Sequencing libraries were prepared using the NEBNext Ultra II FS (New England Biolabs, E7805) kit according to the manufacturer's protocol with reduced reaction volumes (1/4). Samples were barcoded for Illumina sequencing using the NEBNext Multiplex Oligos for Illumina kit (New England Biolabs, E7874) and sequenced on a NextSeq2000 sequencer (Illumina) in 2×150-nt paired-end mode.

Amplicon reconstruction in TR-14

Four distinct ecDNA elements were reconstructed for TR14 using Decoil 1.1.2 with parameters –min-vaf 0.1 –min-cov-alt 10 –min-cov 8 –fragmentmin-cov 10 –fragment-min-size 1000 –filter-score 35 or –min-vaf 0.01 –min-cov-alt 10 –min-cov 10 –max-explog-threshold 0.01 –fragment-min-cov 10 –fragment-min-size 500, with the reference genome GRCh38/hg38 and

annotation GENCODE V42, (as in[41]). Additionally, a de novo assembly contig containing CDK4 for TR14 was assembled using Shasta[42] (v0.6.0) with adjusted parameters minReadLength = 10000, k = 14, consensus caller Bayesian:SimpleBayesianConsensusCaller-10.csv. The contig was aligned to GRCh38/hg38 using minimap2[43] with paramaters -x asm10. From the PAF alignment the overlapping genomic coordinates where extracted as BED file. Using the genomic regions of all amplicons found in TR14 a BED file and a custom reference genome was computed, which includes (1) a masked GRCh38/hg38 genome of the regions overlapping the amplicons and (2) the FASTA sequence of all the ecDNA elements as additional contigs.

Identification of micronucleated sequences

Reads were quality and adapter-trimmed using Trim Galore[44] (v0.6.10) and aligned to the human reference assembly GRCh38 with added EBV contig (NC_007605.1) using BWA-MEM[45] (v 0.7.18) with default parameters. Duplicate reads were removed using Picard[46] (v3.3.0) MarkDuplicates. BigWig files were generated using Deeptools2[47] (v3.5.5) bamCoverage with a binsize of 25kb and CPM normalization.

Candidate window identification

For each sample, we identify candidate windows, which represent genomic regions covered by reads significantly over background, detected by performing the following 4-steps: (1) The genome was divided into 25kb bins and the number of reads and number of bases covered by at least one read were counted with bedtools coverage. (2) For each sample, we fitted a negative binomial distribution on the per-bin read counts, similar to[48]. Distribution parameters were estimated using Maximum Likelihood Estimation as implemented in the fitdistrplus[49] (v1.2.1) R-Package. The threshold RC_{th} for per-bin read counts was derived from the first read count value of the cumulative mass function, where the probability of it being larger than threshold RC_{th} was <0.001, as defined as >1-P, where P denotes the probability of a value being less or equal than threshold RC_{th}.

(3) To remove bins with focal read accumulation, thus being bins with high read counts but low coverage, we further sorted all bins based on the percentage of bases covered. We defined a coverage threshold C_{th} as the 99th percentile. (4) Only bins with an read count > RC_{th} and coverage > C_{th} where retained as candidate windows.

Local window classification

To identify enriched sequences under the assumption that DNA sequestered into micronuclei, which are detectable by light microscopy, derive from chromosomes, chromosomal fragments or other longer chromosomal units (such as ecDNA) of a certain length, we further refine the candidate window set by removal of solidary windows, which adhere to the candidate window definition as described above, but do not have any adjacent candidate windows or are of small size. With this approach we limit the smallest detectable unit size, but make sure to further remove regions, which are most likely derived from amplification and/or mapping noise.

We pivot through each candidate window and search for adjacent candidate windows up and downstream within a radius r. A radius of r=4 was used in this study. If a candidate window has at least one other candidate window upstream and downstream within the radius r, this candidate window is further classified as a body segment, while candidate windows with only at least one other candidate window within the radius r on either side, are classified as edge segment. If a candidate window does not have any other candidate windows within its radius, this window is removed from analysis. We further align edge and body segments into closed segments, where body segments must be flanked by edge segments to be retained, thus making an edge-body-edge alignment the smallest detectable unit, which for this experiment is a fragment size of 75kb. If an edge segment is flanked by another edge segment, both windows are removed from further analysis.

Quality control and filtering

Single/pooled micronuclei samples with more than 20 called windows and >100000 uniquely mapping reads were retained for analysis. Micronuclei samples were further excluded by visual inspection of copy number profiles. Primary nuclei samples were retained if they had more than 50 called windows.

Copy number calling

Copy number profiles were computed similarly to CNVkit[50] using a custom python script utilizing the Pysam API (https://github.com/pysam-developers/pysam). First, the genome was tiled into 500kb bins. For each bin, per base coverages were summed up and divided by the bin size to get a per bin mean coverage, which was then log2 transformed.

Visualization and segmentation of copy number profiles were computed using a custom R script. First, only canonical chromosomes (autosomes and sex chromosomes) were retained and telomeric and centromeric regions filtered out. GC content for each bin was calculated using the R package Biostrings[51]. We performed GC correction for the log2 per bin mean

coverage values using the loess function from the stats R-package with a smoothing span of 0.3 and performed median centering based on all autosomal bins. Segmentation was performed using circular binary segmentation, as implemented in the DNAcopy[52] R-package. After smoothing and outlier removal using the smooth.CNA() function, segments were called using the segment() function with following parameters: undo.splits = "sdundo", undo.SD = 1.5. Log2 per bin coverages and segments were visualized with ggplot2[53].

(Micro-)nuclear sequence analysis

EcDNA consensus sequence calling in single micronuclei

After visual inspection of called regions in single micronuclei, we observed an enrichment of sequences overlapping the amplicon regions, as described above. To identify the consensus region between enriched sequences in single micronuclei and ecDNA reconstructions for further quantification, we generated a set of all called windows from single micronuclei passing quality control. We retained each genomic region, which was called in at least 2 samples. We overlapped these regions with the custom reconstructions and were able to retain the majority of ecDNA regions in single micronuclei. The only region we were not able to capture was a short region in chromosome 1, which length fell below the smallest detectable unit size of 75kb as described above. The consensus ecDNA regions were used for subsequent analysis.

Circular read enrichment estimation (log2 coverage)

To quantify read enrichment over the consensus ecDNA regions, we further normalized the absolute per-bin read counts to counts per million to account for different library size. To estimate the enrichment of circular reads, we divided the mean coverage of all consensus ecDNA regions by the winsorized mean coverage (top and bottom 5% values replaced) of all non-ecDNA regions, as implemented in the Winsorize function from the R-Package DescTools[54] (v0.99.58). The coverage ratio was further log2 transformed.

To quantify the signal over each connected consensus ecDNA region over their flanking region, we obtained region-scaled coverage values using Deeptools2 computeMatrix scale-regions with parameters: -m 250000 -b 125000 -a 125000 -bs 25000 –missingDataAsZero –outFileNameMatrix.

Co-segregation and relative ecDNA fraction analysis

To assess ecDNA co-segregation heterogeneity, we counted the presence or absence for each amplicon in all single micronuclei and primary nuclei samples. For each sample, we overlapped

all called windows with the amplicon reconstructions. An amplicon was deemed present, if it had at least 1 overlapping called window.

To evaluate the relative copy number differences between different amplicons within one sample we computed amplicon fractions within all circular reads. Read counts for each amplicon were first divided by its respective amplicon length to normalize counts for amplicon size. The length normalized counts for each amplicon were further divided by the sum of the length normalized counts of all amplicons to obtain a value, which represents the fraction of reads belonging to each respective amplicon within all circular reads.

Chromatin immunoprecipitation (ChIP)-sequencing

Hydroxyurea- and DMSO-treated COLO320DM cells were washed 2x with prewarmed 1X PBS. Cells were fixed by addition of 10mL ice-cold 1% FA in 10% FCS-RPMI-1640 and incubated for 10 min on a shaker. The fixation reaction was quenched by addition of 550uL ice-cold 2.5M Glycine in PBS. Cells were washed 1x with ice-cold 1X PBS before scraping down in 10 mL ice-cold 1X PBS. Fixed cells were centrifuged at 500 x g for 5 min at 4°C and resuspended in ice-cold 1X PBS for a total of 2 washes. Cells were pelleted and snap frozen in liquid nitrogen and stored at -80 C for no more than 1 week.

Cell pellets were resuspended in 5 mL ice-cold lysis buffer 1 (140 mM NaCl; 50 mM HEPES-KOH pH 7.5; 1 mM EDTA; 10% Glycerol (v/v); 0.5% IGEPAL CA-630 (v/v); 0.25% Triton X-100 (v/v); 1x Protease inhibitor (Roche, 04693132001)) and incubated on a roller at 4°C for 10 min. Lysed cells were centrifuged at 1357 x g for 5 min, resuspended in 4 mL lysis buffer 2 (20 mM NaCl; 10 mM Tris-HCl pH 8.0; 2 mM EDTA; 0.5M EGTA; 1x Protease inhibitor) and incubated on a roller for 10 min at room temperature. The lysate was centrifuged at 1357 x g for 5 min at 4°C and resuspended in 900 µL sonication buffer (100 mM NaCl; 10 mM Tris-HCl pH 8; 1 mM EDTA; 0.5 EGTA; 0.1% Na-deoxycholat (w/v); 0.5% N-Laroylsarcosine (w/v); 1x Protease inhibitor). Lysed chromatin was sheared to a fragment size of 200-500 base pairs (bp) on a Covaris S220 (PIP 140, DF 5, CPB 200) for 12 min. Sheared chromatin was clarified by addition of 1/10 volume 10% Triton X-100 (v/v) and centrifugation at maximum speed at 4°C.

Protein-DNA-complexes were immunoprecipitated by overnight incubation with the respective primary antibody on a rotator at 4°C. 10-15 µg total chromatin was used per sample. 30 µL Protein G beads (Invitrogen, 10003D) were washed 3x in 0.25% BSA in PBS (w/v) and added to each sample. Samples were incubated overnight on a rotator at 4°C.

Antibody-bead - complexes were washed in RIPA buffer (50 mM HEPES-KOH pH 7.5; 1mM EDTA; 1% IGEPAL CA-630 (v/v); 0.7% Na-deoxycholat (w/v); 500 mM LiCl; 1x Protease inhibitor) for a total of 7 times followed by 1 wash in TE buffer (10 mM Tris-HCl pH 8.0; 1mM EDTA; 50 mM NaCl; 1x Protease inhibitor). Beads were eluted in 200 µL Elution Buffer (50 mM Tris-HCl pH 8.0; 10 mM EDTA, 1% SDS (v/v)). 20 µL 5M NaCl and 5 µL Proteinase K (NEB, P8107S) were added to each sample and DNA was de-crosslinked at 65°C overnight, followed by RNA digestion with 2 uL RNase A (Invitrogen, R1253) at 37°C for 30 minutes. DNA was purified using the QIAquick PCR Purification Kit (Qiagen, 28104) according to the manufacturer's instructions. ChIP DNA libraries were constructed using the NEBNext Ultra II DNA Library Prep Kit (NEB, E7645S) and barcoded for Illumina sequencing using the NEBNext Multiplex Oligos for Illumina kit (NEB, E7780S). Libraries were sequenced on a NovaSeq X Plus sequencer (Illumina) in 100 nt single-end mode.

ChIP-sequencing analysis

The raw reads were quality and adapter-trimmed using Trim Galore and aligned to the human reference assembly GRCh38 with an added EBV contig (NC_007605.1) using BWA-MEM with default parameters. Duplicate reads were removed using Picard MarkDuplicates. Library quality was assessed using strand cross-correlation metrics (RSC and NSC) from Phamtompeakqualtools[55,56] (v1.2.2).

Using the deduplicated BAM file, read densities per 10 bp window were generated DeepTools2 bamCoverage. Genomic bins overlapping the ENCODE DAC blacklist regions were filtered out using the parameters: --ignoreForNormalization chrX chrM NC_007605.1 --scaleFactor 10 --effectiveGenomeSize 2805636231 --exactScaling --extendReads 200 --normalizeUsing CPM and input-substracted using the bigwigCompare function. Peak calling was performed with MACS2 with parameters -g 2805636231 --nomodel --extsize 200 -q 0.05. H3K27me3 peaks were called in broad peak mode and with a q-value cut-off of 0.1. Peaks in the blacklisted regions were removed using Bedtools[57] (v2.31.1) intersect.

Identification of differentially enriched histone modifications on ecDNA

We first generated separate consensus peak atlases for each histone ChIP by merging the respective MACS2[58] (v2.2.9.1) peak calls from the treated and untreated samples. Peaks within a distance of 2kb were merged using bedtools merge. Only merged peaks overlapping the amplicon regions were retained. Genomic coordinates for the COLO320DM amplicon were downloaded from AmpliconRepository (ampliconrepository.org). We then divided the

consensus peak atlases into 10 bp bins using bedtools makewindows, for which the scaled and normalized read density was quantified using the Deeptools2 multiBigwigSummary BED-file function for the input-substracted BigWig files. The same approach was applied to generate an additional read density profile by subtracting the treated sample from the untreated sample. To identify differential peaks between the treated and untreated samples, we used the treated-untreated subtracted peak read density profile and measured read density by (1) subtracting the median over all peak regions from each bin (2) and taking the absolute value. We then fitted a 3-component gaussian mixture model (GMM) on the median-subtracted read density values using the normalmixEM function from the R-package mixtools[59] (v2.0.0) and classified bins as differentially enriched/depleted if the respective bin had a probability of >0.999 belonging to the third component of the fitted GMM. Differential regions were visualized for ecDNA genes using Gviz[60] (v.1.42.1).

To compare peak heights between marks on ecDNA and the linear genome, we manually curated a set of housekeeping genes from[61,62] and created a housekeeping genes peak atlas as described above. For these regions, we used the treated-untreated subtracted peak read density profile and subtracted the median over all regions to account for background. For each gene we quantified the largest change by comparing the absolute values of their respective bins and assigned the signed value.

RNA-sequencing

Three replicates of Hydroxyurea- and DMSO-treated COLO320DM cells were washed with 1x PBS, trypsinized (Gibco) and centrifuged at 300xg for 5 min. Total RNA was immediately extracted from the cell pellets using the RNeasy Mini Kit (Qiagen, 74104) following the manufacturer's instructions. Genomic DNA was removed by DNase treatment using the RNase-Free DNase Set (Qiagen,79254). The mRNA library was constructed using the TruSeq Stranded mRNA Library Prep kit (Illumina) at the Berlin Institute of Health (BIH) Genomics Technology Platform. Libraries were sequenced on a NovaSeq X Plus sequencer (Illumina) in 100 nt single-end mode.

RNA-sequencing analysis

Reads were quality and adapter-trimmed using Trim Galore and aligned to the human reference assembly GRCh38 with an added EBV contig (NC_007605.1) using STAR[63] (v2.7.11b) with parameters --twopassMode Basic --outSAMtype BAM SortedByCoordinate. Reads were counted per gene against GENCODE gene annotation (Release 47) using htseq[64] (v2.0.5)

htseq-count with parameters -r pos -s no. Counts were normalized for library size and composition using sizefactors from DEseq2[65] (v1.38.3). Counts were z-score normalized for visualization using the scale function in R-base (v4.2.2). Differentially expressed genes between the untreated and treated samples were identified using DEseq2 and p-values were corrected for multiple testing using the Benjamini-Hichberg procedure. Gene set enrichment analysis was performed on differentially expressed genes with an adjusted p-value <0.05 against MSigDB hallmark gene sets using the GSEA function in Clusterprofiler[66] (v4.6.2) and visualized with DOSE[67] (v3.24.2).

Patient samples and clinical data access

This study comprised the analyses of tumour and blood samples of patients diagnosed with neuroblastoma between 1991 and 2022. Patients were registered and treated according to the trial protocols of the German Society of Pediatric Oncology and Hematology (GPOH). This study was conducted in accordance with the World Medical Association Declaration of Helsinki (2013) and good clinical practice; informed consent was obtained from all patients or their guardians. The collection and use of patient specimens was approved by the institutional review boards of Charité-Universitätsmedizin Berlin and the Medical Faculty, University of Cologne. Specimens and clinical data were archived and made available by Charité-Universitätsmedizin Berlin or the National Neuroblastoma Biobank and Neuroblastoma Trial Registry (University Children's Hospital Cologne) of the GPOH. The *MYCN* copy number was determined as a routine diagnostic method using FISH. DNA and total RNA were isolated from tumor samples with at least 60% tumor cell content as evaluated by a pathologist.

Estimation of ecDNA mis-segregation probability

To estimate the mis-segregation probability of ecDNA in a patient or cell line, we adapted a binomial model of micronucleation[20]. According to this model, the frequency of cells without micronuclei is given by $P_0(p_{MN}, C) = (1 - \frac{1}{2}p_{MN})^C$, where pMN is the missegregation probability of an ecDNA element during cell division and C is the amount of ecDNA elements in the cells. Per patient and cell line where we have more than five cells imaged, we then use the average observed ecDNA copy number per cell and frequency of non-micronucleated cells to calculate p_{MN}.

Biased Random Segregation Model

We developed a model to simulate ecDNA segregation during cell division. A single cell begins with a random ecDNA copy number n, sampled from a uniform distribution. Prior to cell division, ecDNA is amplified to $2n$ copies, which are then partitioned between two daughter cells. For cells undergoing random segregation, ecDNA partitioning follows a binomial distribution $B(2n, p_2)$, where $p_2 = 0.5$. In contrast, for cells exhibiting biased random segregation, partitioning follows a binomial distribution $B(2n, p_1)$, where $p_1 < 0.5$ reflects asymmetric segregation. The probability of biased segregation is denoted as q, while $1 - q$ represents the probability of random segregation. The ecDNA copy numbers in the two daughter cells are denoted as n_1 and $n_2 = 2n - n_1$, respectively. The ecDNA fractions f_1 and f_2 of the two daughter cells can be calculated as:$f_1 = \frac{n_1}{2n}$, $f_2 = \frac{n_2}{2n}$. By repeating this process, we generate the distribution of ecDNA fractions across daughter cells.

Reconstructing ecDNA Segregation Distribution under Finite Sampling

We first simulated ecDNA segregation in 10^7 cells to establish the "true" or "full" distribution. We then sampled subsets of cells at sample sizes of 10, 25, 50, 100, and 500, repeating each sampling 10^5 times. To quantify the deviation between sampled and true distributions, we calculated the Kolmogorov distances.

Parameter Estimation Using Approximate Bayesian Computation (ABC)

We applied Approximate Bayesian Computation (ABC) to infer the parameters p_1 and q in the model. For the simulated data, 50 pairs of daughter cells were randomly drawn using predefined parameters. During inference, p_1 and q were sampled from prior uniform distributions: $p_1 \sim U(0, 0.5)$ and $q \sim U(0, 1)$. 50 parent cells were randomly selected, with their ecDNA copy numbers drawn from a uniform distribution $U(10, 100)$. For each set of parameters, we ran the model to obtain the ecDNA segregation fraction distribution. To compare the similarity between simulated and target distributions, we computed the Wasserstein distance. A threshold was applied to determine acceptance, generating the posterior distributions of p_1 and q. For experimental data, the same procedure was followed, with the sample size adjusted according to the specific experimental conditions.

Statistical analysis

All experiments were performed a minimum of three times with a minimum of three independent measurements. All statistical analysis was performed with R 3.6, R 4.0 or Python 3.7. All data are represented as mean ± standard error. Statistical significance was defined as *, $P < 0.05$; **, $P < 0.01$, ***, $P < 0.001$.

Data availability

ChIP-seq, RNA-seq and micronuclear sequencing data will be deposited in the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/submit). Custom TR14 amplicon reconstructions and FASTA will be available under github.com/henssen-lab. The data that support the findings of this study are available from the corresponding author upon request.

Code availability

Scripts for data analysis within the scope of this publication will be available under github.com/henssen-lab.

AI assisted copy editing

All text was written by the authors. Parts of the text were copy edited using ChatGPT to improve readability and clarity.

Materials & Correspondence

Correspondence and requests for materials should be addressed to henssenlab@gmail.com

Author Contributions

L.B., R.X., J.T., and A.H. contributed to the study design and collection and interpretation of the data and wrote the manuscript. F.T, M.P, K.P., O.S., J.S., Q.Y., D.P., M.I., A.K., S.H., M.G., A.S.H., M.F., F.B., V.R.L., D.G., S.K., D.G., M.R., performed the experiments, analyzed data and reviewed this manuscript. J.T., I.T.L.W., S.Z., performed the experiments, analyzed and interpreted the data, and edited the manuscript. D.T., A.L., F.D. and M.F. provided data. B.P. and A.H. established essential methods. R.L., V.B., G.W., R.M., S.P., B.S., A.P., R.P.K., H.Y.C., A.V., W.H., and B.W. contributed to study design. A.G.H. and P.S.M. led the study design, to which all authors contributed.

Acknowledgements

This work was delivered as part of the eDyNAmiC team supported by the Cancer Grand Challenges partnership funded by Cancer Research UK and led by P.S.M. (CGCATF-2021/100017) (A.G.H.), (CGCATF-2021/100012) (P.S.M. and H.Y.C.), (CGCATF-2021/100020) (B.W and W.H), (CGCATF-2021/100025) (V.B.), and the National Cancer Institute (OT2CA278644) (A.G.H.), (OT2CA278688) (P.S.M. and H.Y.C.), (OT2CA278670) (B.W and W.H), (OT2CA278635) (V.B.), A.G.H. is supported by the *Deutsche Krebshilfe* (German Cancer Aid) Mildred Scheel Professorship program – 70114107. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 949172). This project was supported by Cancer Research UK and the National Institute of Health (398299703, the eDynamic Cancer Grand Challenge). Work in the A.V. group was supported by grants from NIH-NCI (P30 CA008748) and (R01 CA282913 to AV), the American Cancer Society Discovery Boost grant (AV), the The Mark Foundation for Cancer Research ASPIRE grant (AV) and by a rapid response grant from the Functional Genomics Initiative (AV). D.P. was supported by an AIRC fellowship for Abroad. The authors want to thank Steven Whittaker for providing important conceptual input. B.W. is supported by a Barts Charity Lectureship (grant no. MGU045) and a UKRI Future Leaders Fellowship (grant no. MR/V02342X/1). W.H is also supported by NNSF General Program (grant no. 3217024. We thank B. Hero, H. Düren, N. Hemstedt of the Neuroblastoma Biobank and Neuroblastoma Trial Registry (University Children's Hospital Cologne) of the German Society of Pediatric Oncology and Hematology (GPOH) for providing samples and clinical data. We thank Sabine Taschner at the St. Anna Kinderkrebsforschungsinstitut for providing cell lines. Computation has been performed on the HPC for Research cluster of the Berlin Institute of Health. R.P.K. is a Visiting Professor funded by the Stiftung Charité. Images were acquired at the Advanced Light Microscopy facility of the Max Delbrück Center for Molecular Medicine and the Systems Imaging Facility of the Berlin Institute for Medical Systems Biology.

References

1 Kim, H. *et al.* Mapping extrachromosomal DNA amplifications during cancer progression. *Nat Genet* (2024).

2 Kim, H. *et al.* Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. *Nat Genet* **52**, 891-897 (2020).

3 Yi, E., Chamorro Gonzalez, R., Henssen, A. G. & Verhaak, R. G. W. Extrachromosomal DNA amplifications in cancer. *Nat Rev Genet* **23**, 760-771 (2022).

4 Lange, J. T. *et al.* The evolutionary dynamics of extrachromosomal DNA in human cancers. *Nat Genet* **54**, 1527-1533 (2022).

5 Kanda, T., Sullivan, K. F. & Wahl, G. M. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr Biol* **8**, 377-385 (1998).

6 Chamorro González, R. *et al.* Parallel sequencing of extrachromosomal circular DNAs and transcriptomes in single cancer cells. *Nature Genetics* **55**, 880-890 (2023).
7 Hung, K. L. *et al.* ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature* **600**, 731-736 (2021).
8 Wu, S. *et al.* Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **575**, 699-703 (2019).
9 Helmsauer, K. *et al.* Enhancer hijacking determines extrachromosomal circular MYCN amplicon architecture in neuroblastoma. *Nat Commun* **11**, 5823 (2020).
10 Koche, R. P. *et al.* Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. *Nat Genet* **52**, 29-34 (2020).
11 Morton, A. R. *et al.* Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell* **179**, 1330-1341 e1313 (2019).
12 Barker, P. E., Drwinga, H. L., Hittelman, W. N. & Maddox, A. M. Double minutes replicate once during S phase of the cell cycle. *Exp Cell Res* **130**, 353-360 (1980).
13 Takayama, S. & Uwaike, Y. Analysis of the replication mode of double minutes using the PCC technique combined with BrdUrd labeling. *Chromosoma* **97**, 198-203 (1988).
14 Levan, A. & Levan, G. Have double minutes functioning centromeres? *Hereditas* **88**, 81-92 (1978).
15 Hamkalo, B. A., Farnham, P. J., Johnston, R. & Schimke, R. T. Ultrastructural features of minute chromosomes in a methotrexate-resistant mouse 3T3 cell line. *Proc Natl Acad Sci U S A* **82**, 1126-1130 (1985).
16 Tanaka, T. & Shimizu, N. Induced detachment of acentric chromatin from mitotic chromosomes leads to their cytoplasmic localization at G(1) and the micronucleation by lamin reorganization at S phase. *J Cell Sci* **113 (Pt 4)**, 697-707 (2000).
17 Eckhardt, S. G. *et al.* Induction of differentiation in HL60 cells by the reduction of extrachromosomally amplified c-myc. *Proc Natl Acad Sci U S A* **91**, 6674-6678 (1994).
18 Von Hoff, D. D. *et al.* Hydroxyurea accelerates loss of extrachromosomally amplified genes from tumor cells. *Cancer Res* **51**, 6273-6279 (1991).
19 Shimizu, N., Itoh, N., Utiyama, H. & Wahl, G. M. Selective entrapment of extrachromosomally amplified DNA by nuclear budding and micronucleation during S phase. *J Cell Biol* **140**, 1307-1320 (1998).
20 Shoshani, O. *et al.* Chromothripsis drives the evolution of gene amplification in cancer. *Nature* **591**, 137-141 (2021).
21 Engel, J. L. *et al.* The Fanconi anemia pathway induces chromothripsis and ecDNA-driven cancer drug resistance. *Cell* **187**, 6055-6070.e6022 (2024).
22 Von Hoff, D. D. *et al.* Elimination of extrachromosomally amplified MYC genes from human tumor cells reduces their tumorigenicity. *Proc Natl Acad Sci U S A* **89**, 8165-8169 (1992).
23 Tang, J. *et al.* Enhancing transcription–replication conflict targets ecDNA-positive cancers. *Nature* **635**, 210-218 (2024).
24 Savage, J. R. A comment on the quantitative relationship between micronuclei and chromosomal aberrations. *Mutat Res* **207**, 33-36 (1988).
25 Oobatake, Y. & Shimizu, N. Double-strand breakage in the extrachromosomal double minutes triggers their aggregation in the nucleus, micronucleation, and morphological transformation. *Genes Chromosomes Cancer* **59**, 133-143 (2020).
26 Blackford, A. N. & Stucki, M. How Cells Respond to DNA Breaks in Mitosis. *Trends Biochem Sci* **45**, 321-331 (2020).
27 Lukas, C. *et al.* 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat Cell Biol* **13**, 243-253 (2011).
28 Giunta, S., Belotserkovskaya, R. & Jackson, S. P. DNA damage signaling in response to double-strand breaks during mitosis. *J Cell Biol* **190**, 197-207 (2010).
29 Trivedi, P., Steele, C. D., Au, F. K. C., Alexandrov, L. B. & Cleveland, D. W. Mitotic tethering enables inheritance of shattered micronuclear chromosomes. *Nature* **618**, 1049-1056 (2023).
30 van Leen, E., Bruckner, L. & Henssen, A. G. The genomic and spatial mobility of extrachromosomal DNA and its implications for cancer therapy. *Nat Genet* **54**, 107-114 (2022).
31 Lin, Y. F. *et al.* Mitotic clustering of pulverized chromosomes from micronuclei. *Nature* **618**, 1041-1048 (2023).
32 Stöber, M. C. *et al.* Intercellular extrachromosomal DNA copy-number heterogeneity drives neuroblastoma cell state diversity. *Cell Rep* **43**, 114711 (2024).
33 Papathanasiou, S. *et al.* Heritable transcriptional defects from aberrations of nuclear architecture. *Nature* **619**, 184-192 (2023).
34 Agustinus, A. S. *et al.* Epigenetic dysregulation from chromosomal transit in micronuclei. *Nature* **619**, 176-183 (2023).
35 Hung, K. L. *et al.* Coordinated inheritance of extrachromosomal DNAs in cancer cells. *Nature* **635**, 201-209 (2024).
36 Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786-791 (2013).
37 Storlazzi, C. T. *et al.* Gene amplification as double minutes or homogeneously staining regions in solid tumors: origin and structure. *Genome Res* **20**, 1198-1206 (2010).
38 Pradella, D. *et al.* Engineered extrachromosomal oncogene amplifications promote tumorigenesis. *Nature* **637**, 955-964 (2025).

39 Irastorza-Azcarate, I. *et al.* Extensive folding variability between homologous chromosomes in mammalian cells. *bioRxiv*, 2024.2005.2008.591087 (2024).
40 Winick-Ng, W. *et al.* Cell-type specialization is encoded by specific chromatin topologies. *Nature* **599**, 684-691 (2021).
41 Giurgiu, M. *et al.* Reconstructing extrachromosomal DNA structural heterogeneity from long-read sequencing data using Decoil. *Genome Res* **34**, 1355-1364 (2024).
42 Shafin, K. *et al.* Nanopore sequencing and the Shasta toolkit enable efficient de novo assembly of eleven human genomes. *Nat Biotechnol* **38**, 1044-1053 (2020).
43 Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
44 Krueger, F., James, F., Ewels, P., Afyounian, E. & Schuster-Boeckler, B. TrimGalore v.0.6.10. *https://zenodo.org/records/7598955* (2021).
45 Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-bio.GN]* (2013).
46 Picard toolkit. *Broad Institute, GitHub repository* **https://broadinstitute.github.io/picard/** (2019).
47 Ramirez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160-165 (2016).
48 Beagrie, R. A. *et al.* Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519-524 (2017).
49 Delignette-Muller, M. L. & Dutang, C. fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software* **64**, 1 - 34 (2015).
50 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).
51 Pagès H, A. P., Gentleman R, DebRoy S. Biostrings: Efficient manipulation of biological strings version 2.66.0. *https://bioconductor.org/packages/Biostrings* (2022).
52 Seshan VE, O. A. DNAcopy: DNA copy number data analysis version 1.72.3. *https://bioconductor.org/packages/DNAcopy* (2023).
53 H, W. *ggplot2: Elegant Graphics for Data Analysis.*, (Springer-Verlag New York., 2016).
54 Signorell, A. DescTools: Tools for Descriptive Statistics. *https://CRAN.R-project.org/package=DescTools* (2024).
55 Kharchenko, P. V., Tolstorukov, M. Y. & Park, P. J. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat Biotechnol* **26**, 1351-1359 (2008).
56 Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22**, 1813-1831 (2012).
57 Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
58 Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
59 Benaglia, T., Chauveau, D., Hunter, D. R. & Young, D. S. mixtools: An R Package for Analyzing Mixture Models. *Journal of Statistical Software* **32**, 1 - 29 (2009).
60 Hahne, F. & Ivanek, R. in *Statistical Genomics: Methods and Protocols* (eds Ewy Mathé & Sean Davis) 335-351 (Springer New York, 2016).
61 Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends Genet* **29**, 569-574 (2013).
62 Hounkpe, B. W., Chenou, F., de Lima, F. & De Paula, E. V. HRT Atlas v1.0 database: redefining human and mouse housekeeping genes and candidate reference transcripts by mining massive RNA-seq datasets. *Nucleic Acids Res* **49**, D947-d955 (2021).
63 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
64 Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).
65 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
66 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012).
67 Yu, G., Wang, L. G., Yan, G. R. & He, Q. Y. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **31**, 608-609 (2015).

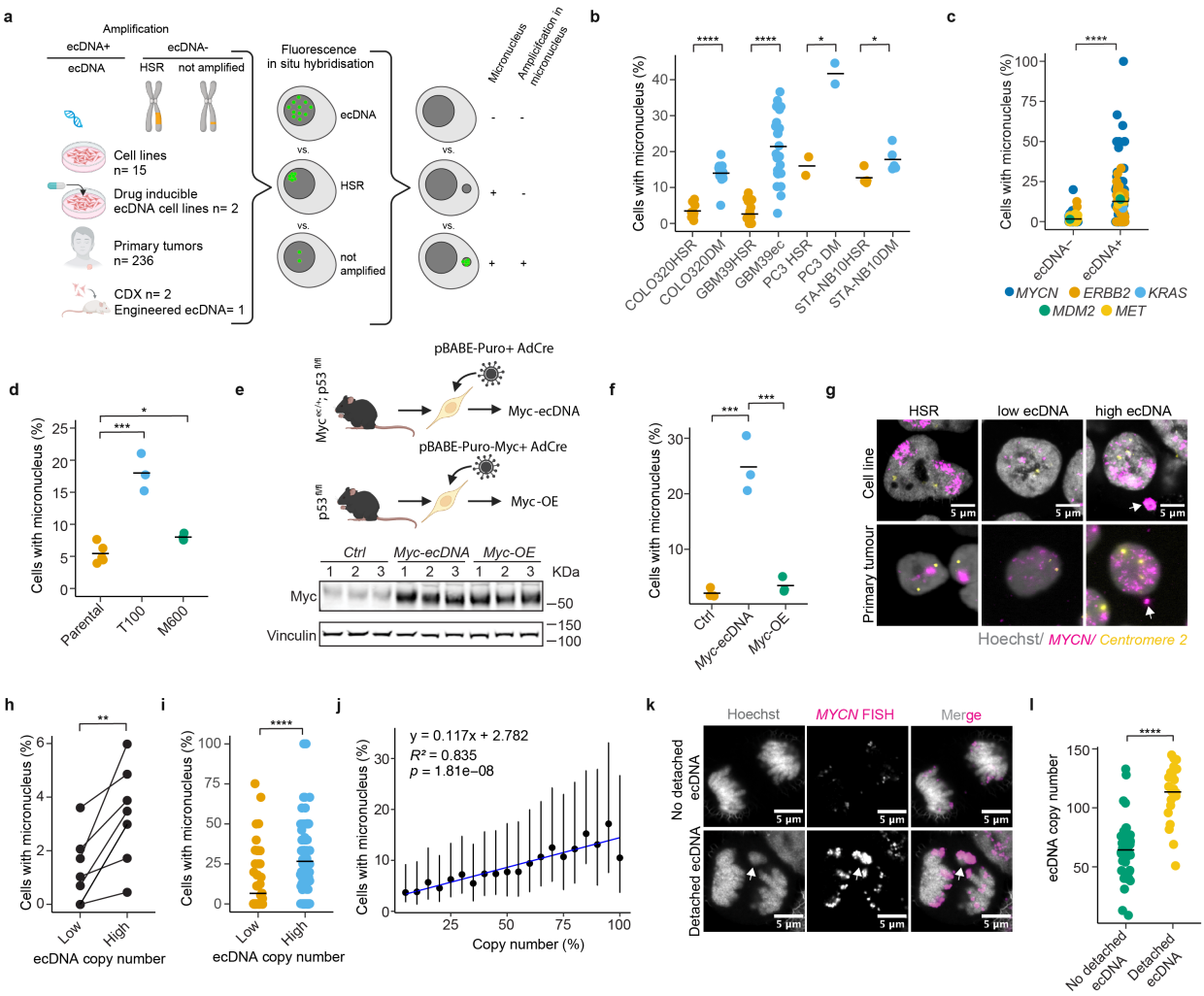

Fig 1. Cancer cells with ecDNA micronucleate more frequently.

a. Schematic of experimental setup to quantify micronuclei in cancer cell lines and tumours

b. Fraction of cells with micronuclei measured in near isogenic cell line pairs (Student's t-test)

c. or primary tumours with or without ecDNA (n = 236, Student's t-test).

d. Fraction of cells with micronuclei in drug-induced ecDNA cell lines treated with either Taxol (T100) or Methotrexate (M600) (One-Way ANOVA).

e. Schematic of experimental setup to induce ecDNA in aNSCs and immunoblot of control, induced Myc-ecDNA and control Myc-overexpressing (Myc-OE) aNSCs.

f. Fraction of cells with micronuclei control, induced Myc-ecDNA or Myc-OE aNSCs (One-Way ANOVA).

g. Exemplary photomicrographs of cells with different ecDNA content and micronuclei (white arrowhead)

h-i. Fraction of cells with micronuclei and low (bottom 30%) vs. high (top 30%) ecDNA content measured in (h) cell lines (i) and primary tumours (paired Student's t-test).

j. Fraction of cells with micronuclei combining all cell line and primary tumour data within different ecDNA copy number bins. Percentages of overall copy number were calculated for each dataset and micronucleation rate determined for each copy number bin from lowest (0-5% of overall copy number) to highest (95-100% of overall copy number).

k. Exemplary photomicrographs of cells in anaphase with and without detached ecDNA and stained with Hoechst (grey) and MYCN DNA FISH (magenta).

l. Number of ecDNA copies per cell with and without detached ecDNA clusters in anaphase (Student's t-test).

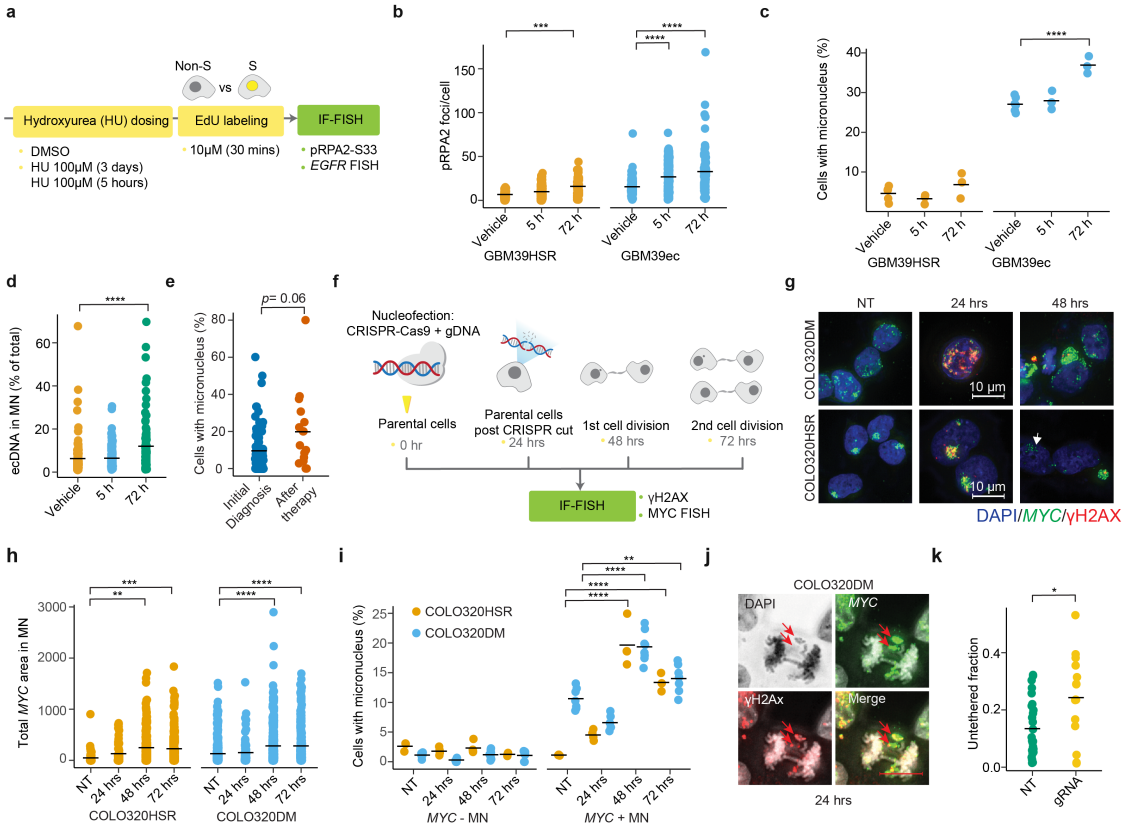

Fig 2. High ecDNA copy number and targeted DNA damage promote ecDNA micronucleation via detachment in mitosis

a. Schematics to depict replication stress by HU at different time points.

b. Quantification of pRPA2 foci per cell in isogenic glioblastoma cell lines at different timepoints of HU treatment in cycling cells (Two-way ANOVA).

c. Fractions of cells with micronuclei measured in GBM39ec and GBM39HSR cells (Two-way ANOVA)

d. Fraction of total ecDNA in micronuclei (One-way ANOVA)

e. Fraction of cells with micronuclei in tumour samples at initial diagnoses and after standard-of care chemotherapy (Student's t-test).

f. Targeted DNA damage inducing by CRISPR on amplicon. Schematics to depict samples collected at different time points.

g. Representative images in COLO320DM/COLO320HSR cells to show DNA damage induced at amplicon and amplicon+ micronuclei formation. White arrows indicate micronuclei.

h. Total *MYC* in micronuclei at different time points after nucleofection with CRISPR targeting the amplicon in isogenic cell lines (One way ANOVA)

i. Fraction of cells with *MYC*+ or *MYC*- micronuclei at different timepoints after nucleofection with CRISPR targeting the amplicon in isogenic cell lines (One-way ANOVA)

j. Exemplary photomicrograph of cells in anaphase with detached *MYC* ecDNA positive for the DNA damage marker γH2AX (Red arrows). Scale bar represents 10 μm

k. Untethered fraction of ecDNA in anaphase 24 hrs after nucleofection (Student's t-test)

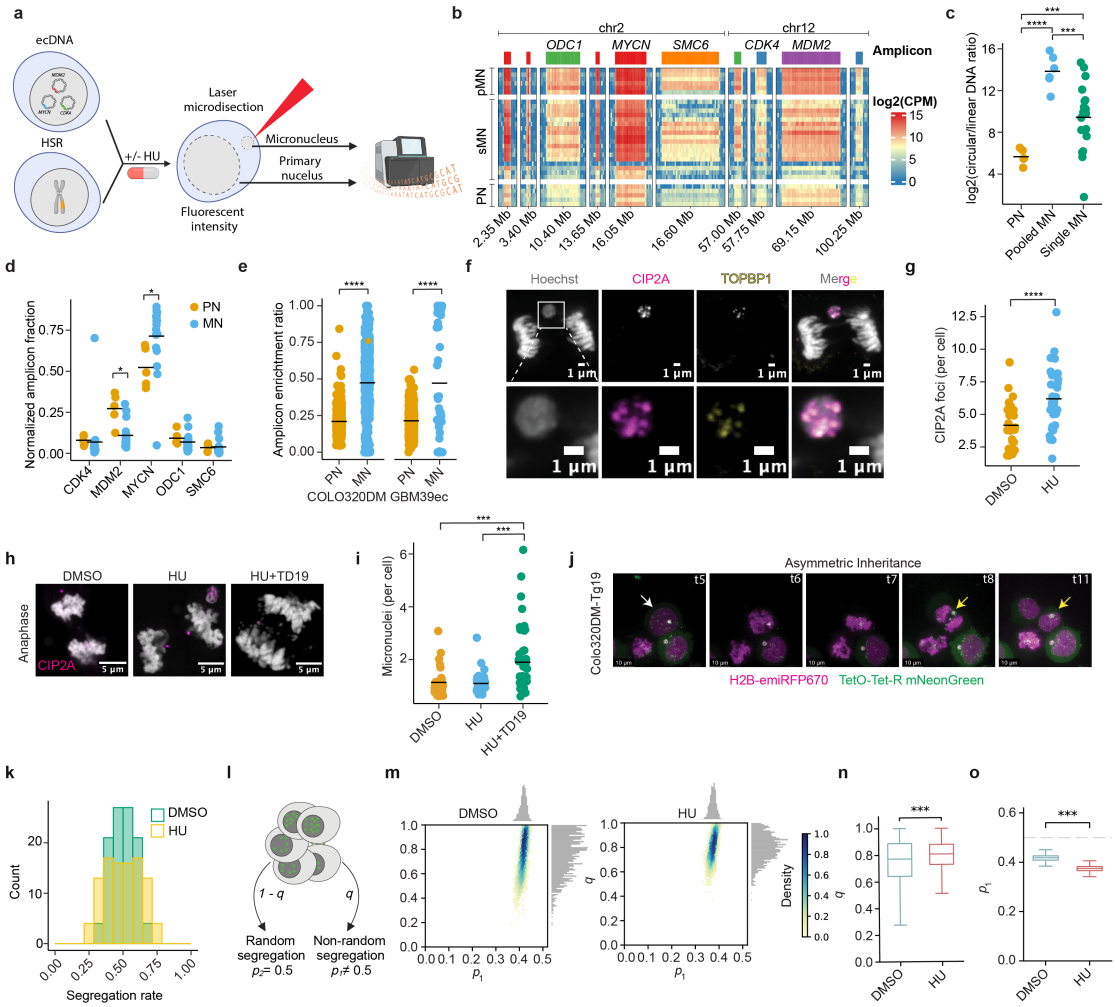

Fig. 3: Clustered ecDNA mis-segregation induces asymmetric mitotic inheritance.

a. Schematics of the experimental setup

b. Laser microdissection directed sequencing of primary nuclei (MN) and micronuclei (MN). Counts per million (CPM) – normalized and log2 transformed read coverage (25kb bins) across detected enriched regions and their flanking regions (100 kb). Top annotation: EcDNA reconstruction from long-read sequencing. Regions falling below the detection size limit (75 kb) were excluded (PN: Primary nucleus (n = 5), sMN: single micronucleus (n = 19), pMN: pooled micronuclei (n = 7))

c. Log2-transformed mean coverage ratio of circular reads over the winsorized mean of all linear reads (Welch's t-test).

d. Amplicon-size normalized fraction of circular reads mapping to specific ecDNA species in primary nuclei and micronuclei (Welch's t-test).

e. Amplicon enrichment as determined by FISH in two ecDNA+ cell lines (Student's t-test).

f. Exemplary photomicrographs of a cell with ecDNA in anaphase stained using immunofluorescence against TOPBP1 (yellow) or CIP2A (magenta).

g. CIP2a foci increase after 24 hrs HU treatment.

h. Exemplary photomicrographs of cells treated for 24 hrs with HU or HU+TD19.

i. HU+TD19 treatment increases the number of micronuclei per cell compared to HU only (One-way ANOVA).

j. Live-cell imaging of the COLO320DM-Tg19 cell line engineered to visualize *MYC* ecDNA (TetO-TetR-mNeonGreen) during mitotic transition. Representative snapshots of a mitotic cell (white arrow) carrying an ecDNA-containing micronucleus (yellow arrow) getting asymmetrically inherited into one of the daughter cells upon mitotic exit. Scale bar: 10µm.

k. Histograms of ecDNA fractions in the CHP212 cell line in DMSO treated cells without visible lagging cluster (left), DMSO treated cells (middle) and HU treated cells (HU), showing the distribution of ecDNA segregation across daughter cells. l. Overview and simulation results of the biased random segregation model.

m. Density scatter plot of inferred parameters p_1 (extent of biased segregation) and q (proportion of biased segregation). The marginal distributions of p_1 and q are shown above and to the right, respectively. Density values are normalized to 1.

n. The inferred proportion of cells exhibiting biased segregation is significantly increased in drug-treated cells compared to untreated cells. Statistical significance was determined using the Mann–Whitney U test ($p < 0.001$).

o. The inferred p_1 is significantly reduced in drug-treated cells compared to untreated cells. Statistical significance was determined using the Mann–Whitney U test ($p < 0.001$).

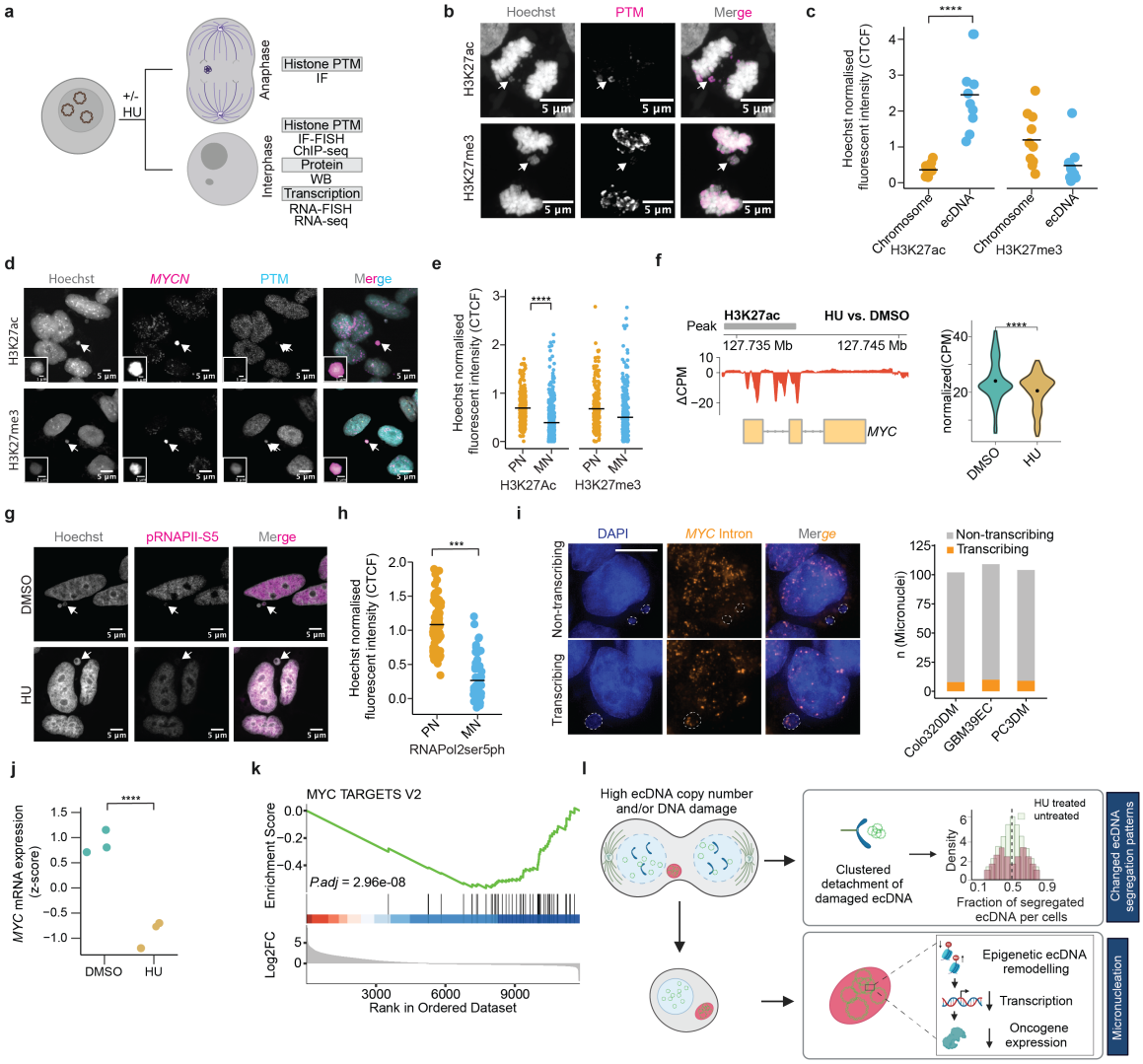

Fig. 4 Reduced oncogenic transcription of ecDNA in micronuclei.

a. Schematic of experimental setup

b. Exemplary photomicrographs of cells with ecDNA in anaphase stained using immunofluorescence against H3K27Ac or H3K27me3 (ecDNA is indicated with white arrowheads).

c. Ratio of H3K27Ac and H3K27me3 intensity on ecDNA vs. linear chromosomes.

d. Exemplary photomicrographs of a cell with ecDNA stained using immunofluorescence against H3K27Ac or H3K27me3 (micronuclei are indicated with arrowheads).

e. Hoechst normalized H3K27Ac and H3K27me3 intensity in micronuclei vs. primary nuclei.

f. ChIP-sequencing of H3K27ac in COLO320DM cells. Left panel: from top to bottom: peak call annotation. DMSO-subtracted read density as counts per million (CPM). Gene annotation. Right panel: Input-subtracted read density as CPM.

g. Exemplary photomicrographs of cells with ecDNA incubated with hydroxyurea (80 µM) or DMSO (vehicle control) and stained using immunofluorescence against RNA polymerase II phosphorylation at serine 5 (micronuclei are indicated with arrowheads).

h. Hoechst normalized intensity in micronuclei vs. primary nuclei from cells stained using immunofluorescence against RNA polymerase II phosphorylation at serine 5.

i. Exemplary photomicrographs of COLO320DM cells and their micronuclei (white dashed line) labelled with intron *MYC* RNA FISH signal (left), showing a non-transcribing micronucleus (top) and an actively transcribing micronucleus (bottom). A quantification of the number of transcribing and non-transcribing micronuclei in COLO320DM (n=102), GBM39EC (n=109), and PC3DM (n=104) (right). Scale bar: 10µm.

j. Z-score normalized transcript counts of the circularly amplified *MYC* oncogene in COLO320DM cells. Three replicates per treatment condition. Benjamini-Hochberg procedure-corrected p-value is shown (Wald test).

k. Gene set enrichment analysis (GSEA) of MYC target genes. Benjamini-Hochberg procedure-corrected p-value is shown.

l. Schematic highlighting the main findings of this publication