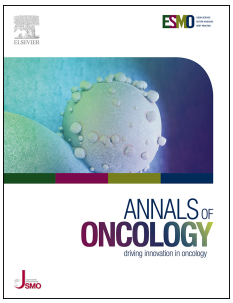
Journal Pre-proof



Extrachromosomal DNA – relieving heredity constraints, accelerating tumour evolution

Chris Bailey, Massa J. Shoura, Paul S. Mischel, Charles Swanton

PII: S0923-7534(20)36392-4

DOI: https://doi.org/10.1016/j.annonc.2020.03.303

Reference: ANNONC 144

To appear in: *Annals of Oncology*

Received Date: 21 February 2020

Accepted Date: 26 March 2020

Please cite this article as: Bailey C, Shoura MJ, Mischel PS, Swanton C, Extrachromosomal DNA – relieving heredity constraints, accelerating tumour evolution, *Annals of Oncology* (2020), doi: https://doi.org/10.1016/j.annonc.2020.03.303.

Extrachromosomal DNA – relieving heredity constraints, accelerating tumour evolution

Chris Bailey (1), Massa J. Shoura (2), Paul S. Mischel (3, 4, 5), Charles Swanton (1,6)

(1) Cancer Evolution and Genome Instability Laboratory, the Francis Crick Institute and University College London Cancer Institute, London, UK

(2) Departments of Pathology and Genetics, Stanford University School of Medicine, California, USA

(3) Ludwig Institute for Cancer Research, University of California at San Diego, USA

(4) San Diego Moores Cancer Center, University of California, La Jolla, CA, USA

(5) Department of Pathology, University of California San Diego, La Jolla, CA, USA

(6) Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK

Abstract

Oncogene amplification on extrachromosomal DNA (ecDNA) provides a mechanism by which cancer cells can rapidly adapt to changes in the tumour microenvironment. These circular structures contain oncogenes and their regulatory elements, and, lacking centromeres, they are subject to unequal segregation during mitosis. This non-Mendelian mechanism of inheritance results in increased tumour heterogeneity with daughter cells that can contain increasingly amplified oncogene copy number. These structures also contain favourable epigenetic modifications including transcriptionally active chromatin, further fuelling positive selection. ecDNA drives aggressive tumour behaviour, is related to poorer survival outcomes and provides mechanisms of drug resistance. Recent evidence suggests one in four solid tumours contain cells with ecDNA structures. The concept of tumour evolution is one in which cancer cells compete to survive in a diverse tumour microenvironment under the Darwinian principles of variation and fitness heritability. Unconstrained by conventional segregation constraints, ecDNA can accelerate intratumoural heterogeneity and cellular fitness. In this review, we highlight some of the recent discoveries underpinning this process.

Introduction

Extrachromosomal DNA (ecDNA) as a feature of cancer was first described over 50 years ago, but only recently is evidence emerging of its prevalence amongst tumour types and the crucial role it plays in oncogene amplification and tumour evolution[1]. ecDNAs have also been traditionally known as extrachromosomal double-minutes (DM), so called due to their original description as paired structures[2]. Spriggs and colleagues first described these double-minute chromatin bodies in 1965 whilst examining metaphase spreads of a bronchial carcinoma and 5 neuroblastoma samples[2]. In 1981, Schimke and colleagues reported that methotrexate resistance can occur as a consequence of dihydrofolate reductase gene amplification on double-minutes and homogenous staining regions (HSR) in the 3T6 mouse fibroblast cell line[3], and in 1983, Alt et al described the amplification of *MYCN* sequences in ecDNA in neuroblastoma cell lines[4]. This, alongside the observation of *MYC* amplification as homogenous staining regions in neuroendocrine cell tumour lines by Alitalo et al[5], were the first descriptions of ecDNA and HSR oncogene amplification.

Until the development of high throughput sequencing, ecDNA was traditionally detected through cytogenetic methods. In 1984, Gebhart et al. examined over 4,000 metaphase spreads of 22 primary breast cancers and 55 malignant effusions, identifying double-minutes in over 73% of samples[6], and in 1991, Benner et al. conducted cytogenetic analysis of 200 primary tumour samples, identifying DMs in 93.5% of samples[7]. Aggregating the studies conducted at the time, Gebhart found that DMs were identified in approximately 30% of ovarian cancers and 20% of gliomas, osteosarcomas, lung and breast cancers[8]. With the advent of high throughput sequencing, further characterisation of these structures has been facilitated through bioinformatic analysis. This has confirmed that oncogene amplification on ecDNA is a common event, accelerating intratumoural heterogeneity and causing extreme copy number amplification due to the unequal segregation of circular amplicons[1,9]. This complex inheritance pattern is coupled with a favourable transcriptional profile to drive massive oncogene expression[10]. Furthermore, it is becoming increasingly clear that ecDNA promotes aggressive tumour behaviour, is related to poorer survival outcomes, and can lead to drug resistance[9,11]. Consequently, there is an urgent need to better detect and describe evolutionary processes in the context of ecDNA.

In this review, we highlight the recent discoveries that have contributed to this emerging field.

ecDNA and homogenous staining regions

ecDNAs are a class of circular DNA structures found outside the normal chromosome architecture (they are also commonly referred under the umbrella term of extrachromosomal *circular* DNA, or eccDNA)[12,13]. They are distinguished from other DNA structures by their size and constituents; ranging between 0.2-5 Mb in length[10,14] and frequently contain one or more oncogenes and regulatory elements[10] (fig. 1). Typically, circular structures as large as ecDNA are rarely detected in healthy human cells[11]; however, smaller eccDNAs, structures that range from 100bp to 100kb, can be detected in both healthy and cancer cells (such as in HeLa cells and myeloproliferative neoplasms)[15,16]. Nevertheless, the repertoire of these eccDNAs is determined by cell-type and pathology[16]. Additionally, topological constraints of very small eccDNAs (<1Kb) (or microDNAs), allow for transcription of regulatory RNAs off the circular template, providing a pathway to regulation of gene expression[17]. Neochromosomes, which can be circular or linear, are also seen in the context of cancer. They can form as a result of chromothripsis and occasionally exceed 600 Mb in size[18]. Unlike ecDNAs, neochromosomes contain centromeres and therefore have a very different impact on heterogeneity and copy number.

Oncogenes expressed on ecDNA significantly amplify copy number through random segregation of amplicons during mitosis (fig. 2). This is due to a lack of centromeres, a phenomenon described by Levan and Levan in 1978[19]. Pichugin et al constructed a robust theoretical model of the evolutionary dynamics of ecDNA, demonstrating patterns of intratumour heterogeneity and cellular fitness under neutral and positive selection pressures[20]. Validating this model with experiments involving EGFRvIII ecDNA inheritance in the GBM39 cell line, they demonstrate that cells develop a 300% fitness increase where ecDNA is present[20].

The numbers of amplicons per cell can vary from 1 to over 100[11]. Koche et al reported that amongst a cohort of 93 neuroblastoma tumour samples and 12 cell lines, the

mean ecDNA count per cell was 0.82[14]. Kim et al analysed whole genome sequencing data from 34 cancer cell lines; amongst these samples, the median ecDNA count per cell was 14.5[11]. Using TCGA (The Cancer Genome Atlas) whole genome sequencing data from 1,979 patients across 25 cancer types, the same group reported that ecDNA was present in 26% of all cancer samples and was most commonly found in glioblastoma, oesophageal carcinoma and squamous cell lung cancer. EGFR, MYC, CCND1, CDK4 and MDM2 were the most commonly amplified oncogenes[11].

ecDNA have been shown to dynamically re-integrate onto aberrant genomic locations known as homogenous staining regions (HSR, fig. 2)[21–24]. These are highly duplicated intrachromosomal regions that stain uniformly on cytogenetic giemsa-banding[25]. Consequently, gene amplification has been classically divided by these cytogenetic techniques; intrachromosomal HSRs and extrachromosomal DMs[26]. HSRs also form as consequence of complex genomic rearrangements such tandem duplications and chromosome breakage fusion cycles[27,28]. The reversible relocation of EGFR mutant ecDNA onto a HSR was demonstrated by Nathanson et al to be a mechanism of drug resistance in GBM39 cell lines[29]. Koche et al recently demonstrated in neuroblastoma samples that circle reintegration contributes not only to oncogene amplification as HSRs, but also general oncogenic genome modelling[14].

The biology of ecDNA

The biological causes underpinning ecDNA origin and evolution are yet to be fully elucidated. ecDNA and HSR formation are traditionally characterised as distinct mechanisms of gene amplification, yet they both originate as a consequence of multiple double-stranded breaks (DSB)[30,31]. It has been shown that specific DSB repair pathways are implicated in the formation of ecDNA over HSRs. For example, Cai et al demonstrated that inhibition of homologous recombination through BRCA1 silencing of MTX-resistant HT-29 colon cancer cells resulted in decreased numbers of ecDNA, but no effect on HSR formation was detected.[32] The authors had previously shown in the same cell line increased expression of non-homologous end joining (NHEJ) proteins and reduction of ecDNA formation with inhibition of the key DNA-PKcs protein[33]. Vogt et al. resolved the structure of a co-amplified

MYC and EGFR ecDNA amplicon, present in successive passages of a xenografted human oligodendroglioma[23]. The structure comprised of both large and small fragments from chromosomes 7 and 8, with the smaller fragments originating from clusters of breakpoints at distal genomic locations[23]. The ecDNA relocated as an HSR in successive passages with additional breakpoint clusters accumulated during this transition. The data suggest a coordinated process of repair involving a complex that brings together distal genomic segments, not unlike that seen in V(D)J recombination[34]. Moreover, the junctions contained microhomologies suggestive of NHEJ or microhomologous end joining (MMEJ)[23].

It is still unknown how ecDNA forms; however, the range of structural complexities suggests multiple processes, including chromothripsis and episome formation, can result in their genesis. Chromothripsis is characterised by multiple rearrangements of one or a few chromosomes as a result of a shattering event[35]. The "episome model" is one where extrachromosomal DNA segments from an otherwise intact chromosome[22,28]. L'Abatte et al. characterised the structure and function of *MYC*-containing amplicons in acute myeloid leukaemia with 8q24 amplifications[36]. ecDNA formation in leukaemia is rare, but correlates with age and is associated with a poor prognosis[37]. They were able to determine that heterogeneous amplicon structures co-existed within the same leukemic cell population and the organisation of genomic segments within these structures favoured a model whereby amplicons were formed as precursors of episomes and not as a consequence of chromothripsis[36]. By inducing chromosomal segregation errors in the DLD-1 cell line, Ly et al were able to observe a range of simple and complex inter- and intrachromosomal genomic rearrangements[38]. In this study, there were examples of ecDNA formation as direct consequence of chromothripsis, resulting in amplification of the genes contained on the circular ecDNA[38]. Therefore, formation of ecDNA amplicons are likely to originate from a range of mechanisms, including those related to the generation of chromosomal instability and deficiencies in DNA repair. By introducing a plasmid that contains a mammalian replication and initiation region (IR) and a nuclear matrix attachment region (MAR), Shimizu et al were able to successfully amplify genes as DMs or HSRs[39]. They describe the process of DM formation whereby IR/MAR plasmids that initially replicate as small extrachromosomal circles multimerise to form larger structures, and HSR formation whereby a collision

between replication and transcription machinery due to the MAR sequences result in strand-breaks that initiate breakage-fusion-bridge cycles[26,40].

The loss of amplified oncogenes through ecDNA elimination has also been observed[26]. Shimizu et al noted that the spontaneous differentiation of the human promyelocytic leukaemia cell line HL-60 was a consequence of an active elimination process leading to c-MYC copy number loss[41]. It was demonstrated that DM loss was accelerated in the presence of hydroxyurea[42], and that this elimination had occurred following the formation of DM-enriched micronuclei after mitosis[43]. This demonstration of copy number loss further highlights the ability of tumour cells containing ecDNA to rapidly adjust to changing environments.

There are a number of mechanisms that underlie the means by which cells with oncogene-containing eDNA amplicons gain a competitive advantage. The TCGA work by Kim et al. explored the genomic locations of ecDNA[11]. Not only was there a significant enrichment for oncogenes in the constructed amplicons of 1,979 patients, but breakpoints were randomly distributed proximal to the oncogene locus, suggesting that ecDNA forms at random loci. A selective advantage is gained in those loci that contain oncogenes and regulatory elements. Further selective advantages are gained through unequal segregation of the circular amplicons that consequently increase daughter cell copy-number, and through increased expression of ecDNA independent of gene dosage; consequently, selective advantages are gained in cells through increased oncogene expression and copy number. Further selective advantages may be gained through circle re-integration. Koche et al reported an example of re-integration of amplicon sequences into the DCLK1 gene locus, resulting in loss of heterozygosity, and subsequent loss of tumour suppressor function[14]. Determining where and whether ecDNA relocates as HSRs will be essential if we are to target this phenomenon.

The epigenomic landscape

The role of chromatin modifying enzymes and epigenetic states in ecDNA oncogene copy number amplification is becoming apparent (fig. 3). Clarke et al. have elegantly

demonstrated the role of histone lysine methyltransferases (KMT) and demethylases (KDM) in modulating histone methylation balance and subsequent transient site-specific EGFR copy gains including high-level EGFR amplification[44]. In this work, combining public datasets with functional work in cell lines, including immortalised retinal pigment epithelial cells, they demonstrate that interference of H3K9 and H3K27 methylation states result in site-specific EGFR amplification and expression. The role of modulating H3K9 methylation was demonstrated through overexpression of KDM4A and knockdowns of K9 KMTs resulting in EGFR copy gains. The KDM4A-driven, EGFR-driven copy gains were dependent on the opposing balance of H3K4 and H3K27 methylation states. Specifically, H3K27 interference through KDM6A/B overexpression and EZH2 inhibition requires H3K4 methylation (fig. 3). This process demonstrated by overexpression of the H3K4 KMTs (KMT2A, SETD1A and SETD1B), and inhibition of KDM5A[44]. Furthermore, they build on previous work to demonstrate that hypoxia and EGF promote transient site-specific EGFR copy gains through targeting H3K4 methylation via different epigenetic means[44,45]. These data suggest that pharmacological targeting of epigenetic modifiers may attenuate extrachromosomal amplification of EGFR.

The chromatin topology of ecDNA can contribute towards positive selection and cell fitness through proximity to regulatory elements and accessibility to transcription machinery. Morton et al. demonstrated in glioblastoma cell lines that oncogenes are not only amplified with endogenous enhancers on ecDNA, but they can co-opt regulatory elements from other topologically-associated domains[46]. In this study ecDNA EGFR amplicons undergo enhancer rewiring, whereby the endogenous and ectopic enhancers that are incorporated into the amplicon each contribute to cellular fitness (fig. 3). They found further evidence for this phenomenon interrogating public datasets in other tumour types, including *MYC* in group 3 medulloblastoma and *MYCN* in neuroblastoma and Wilms tumours[46]. Interestingly, functional endogenous enhancer activity was maintained across all samples tested. This is in contrast to Helmsauer's et al characterisation of *MYCN* amplicons in neuroblastoma cell lines[47]. They describe a structurally complex subset that lacks endogenous enhancers and thus hijack ectopic enhancers from distal regions of the same chromosome into the ecDNA structure. Along with their comprehensive structural analysis of ecDNA amplicons, Wu et al. demonstrated that the nucleosomal organisation of ecDNA is

less compacted compared with linear DNA, allowing for highly accessible chromatin that remains organised[10] (fig 3). Furthermore, they show ecDNA enables ultra-long-range chromatin contacts, permitting distant interactions with regulatory elements[10]. Taken together, these studies indicate that the circular structure of ecDNA allows the oncogene to utilise additional enhancers on the amplicon and increase its transcriptional capacity.

It is becoming apparent is that the chromatin landscape of ecDNA is dramatically altered compared with homologous regions on linear DNA. This process of co-opting regulatory elements is a mechanism by which tumour cells can exploit epigenomic regulatory control. This enhancement of oncogenic drive is likely to further fuel positive selection.

Tools to identify and explore ecDNA

ecDNA contributes to focal amplification of oncogene copy number and is comprised of multiple breakpoints and genomic segments. This unique structural variant poses a mechanistic challenge beyond the conventional difficulties of detecting complex structural variants[48,49]. Conventional tools can infer copy number through sequence read density and de-lineate paths and cycles by connecting potential breakpoints from chimeric reads; however, it is very difficult to infer circularity. Moreover, breakpoints occur far more commonly in intergenic regions and, as such, whole genome sequencing has been the basis by which these amplicons are constructed. Bioinformatic tools such as Amplicon Architect use short-read whole genome sequencing data to detect breakpoints and use chimeric reads to infer circularity[50].

Due to the complexity and uncertainty of computational methods, a combined approach to detection, incorporating circle enrichment, new sequencing methods and cytogenetics, has been essential. In 2012, Shibata et. al described a method of purifying eccDNAs from mouse tissues using alkaline lysis followed by circular DNA enrichment through exonuclease depletion of linear DNA[51]. In 2015, Moller and colleagues described a similar method of purifying eccDNAs from the *Saccharomyces cerevisiae* genome, whereby circular DNA was also enriched through exonuclease depletion of linear DNA[52]. Following this, in both methods, samples are subjected to rolling circle amplification and high throughput sequencing. The method was adapted by Koche et al in 2019, who incorporated

magnetic bead separation of high molecular weight DNA, enriching both ecDNA and smaller eccDNA structures in neuroblastoma samples[14]. This was compared with computational reconstruction using whole genome short-read sequencing data, and demonstrated 100% concordance with ecDNA and 30% concordance with eccDNA. Combining this approach with long-read sequencing provided a detailed genomic structural map of the ecDNA amplicons[14]. In 2017, Shoura et al developed circulome-seq, which incorporated two (biophysical and biochemical) enrichment methods to characterise the landscape of small eccDNAs in *C. elegans* and human fibroblasts and granulocytes. Their method uses centrifugation in CsCl/ethidium-bromide gradients and exonuclease V digestion of linear DNA, enriching for circular DNAs up to 100Kb. Avoiding rolling circle amplification, they utilised a Tn5 transposition-based fragmentation and tagging system to directly target the circular DNAs, preserving copy number and providing a bioinformatic signature for identifying full assemblies of eccDNA[16]. In addition, Tn5-based ATAC-seq has been used to identify circular DNA structures, including both the small, non-coding eccDNAs in normal and cancer cells[53] and in cancer-specific ecDNA[10].

Homogenously staining regions and ecDNA have been conventionally described using cytogenetic methods such as karyotyping and fluorescence in-situ hybridization (FISH). More recently, electron microscopy has been used further understand ecDNA structure. In 2019, Wu et al. integrated super-resolution three-dimensional structured illumination microscopy (3D-SIM) with scanning and transmission electron microscopy to confirm circularity[10]. Moreover, visualisation tools have been developed to quantify ecDNA from DAPI (4',6-diamidino-2-phenylindole) stained metaphases[54]. This exciting development will aid understanding into how and where ecDNA relocates as HSR.

Clinical implications and drug resistance

The mechanisms underpinning drug resistance in tumours are yet to be fully described[55]. Despite examples of ecDNA-derived drug resistance, it is not a mechanism that has not been fully explored. Turner et al performed fine structure analysis of EGFRvIII amplification in GBM39 cells, proving that in response to erlotinib treatment, the ecDNA reintegrated as HSRs[9]. On removal of treatment, the ecDNA amplicons re-emerged, confirming previous work carried out in the same lab in 2014[29]. Schimke's work in the 1980s

showed that methotrexate resistance can be a consequence of ecDNA-derived dihydrofolate reductase gene amplification[3]. It must also be considered that accelerated intratumoural heterogeneity through ecDNA may in itself increase the likelihood of tumour cells developing drug resistance by increasing the diversity of genetically distinct subclones[56].

ecDNA-driven copy number amplification provides a mechanism by which tumours can accelerate intratumoural genetic heterogeneity, generating cellular diversity through unequal segregation of ecDNA to daughter cells[9]. This heterogeneity provides the substrate for which tumours can adapt quickly to environmental and treatment pressures[56]. Patients with ecDNA-based oncogene amplification show increase cell proliferation activity, increased likelihood of lymph node spread at diagnosis and decreased overall survival[11]. Understanding the precise role that ecDNA plays will not only enhance our understanding of cancer evolution as a whole, but will further inform our treatment strategies.

ecDNA parallels in nature

Circular DNA formation is widespread in nature, with many examples of eukaryotes and prokaryotes utilising this phenomenon to rapidly generate genetic variation and amplify copy number. The genome of the yeast *Saccharomyces cerevisiae* contains 16 chromosomes and is approximately 12 Mb in length[57]. By purifying eccDNA from the yeast *Saccharomyces cerevisiae,* Moller et al discovered that up to 23% of the genome was contained on eccDNA of sizes 1kb to 38kb[52]. The eccDNAs contained an abundance of autonomously replicating consensus sequences, and were enriched for ribosomal genes, transposon remnants and tandemly repeated genes. Despite the formation of eccDNA throughout the genome, the increased abundance of certain genes and autonomous replication sequences suggests cell selection through replicative fitness[52]. Interestingly, from the eccDNA sequencing reads, there was a wide variety of sequence homology with chromosomal loci at eccDNA junctions, suggesting that circle formation is mediated by homologous recombination, non-homologous end joining and micro homology-mediated DNA repair[52].

Glycophosphate resistance has become a considerable problem for weed control in major agronomic crops, with confirmation in thirty-eight weed species across thirty-seven countries reported in 2017[58]. In 2018, Koo et al. discovered that amplification of the 5-enolpyruvlyshikimate-3-phosphate synthase (EPSPS) gene can reside on eccDNA as a mechanism of glycophosphate resistance in *Amaranthus palmeri*[59]. During meiosis, eccDNAs tether to chromosomes, facilitating transmission to gametophytes. Moreover, the progeny plants of *Amaranthus palmeri* crosses (where one parent lacked eccDNA) displayed EPSPS copy number variation between individual offspring and within different tissues of the same offspring[59].

In bacteria, antibiotic resistance genes can be rapidly amplified on plasmids, circular double-stranded DNA molecules that replicate autonomously and can be transferred horizontally across genera[60]. In 2015, Liu et al reported the emergence of plasmid-mediated polymyxin E (colistin) resistance gene MCR-1 in Enterobacteriaceae amongst livestock and humans in China. They demonstrated the horizontal transfer of plasmids between different *Escherichia coli* strains and to strains of *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*[61]. The subsequent isolation of mcr-1 amongst gram negative strains in Europe and Africa have worrying implications for the worldwide threat of antibiotic resistance[62].

There are examples of circular DNA formation in viral latency. The gammaherpesviruses Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV, HHV-8) have linear genomes that are ligated to extrachromosomal circular DNA and maintained in latency as plasmids upon infection of B-lymphocytes[63,64]. Viral copy number is in part maintained through tethering of the viral genome to host chromosomes during mitosis. It is thought that a similar tethering process occurs in ecDNA during mitosis[19]. The EBV genome is replicated and partitioned in a pairwise association with host sister chromatids, resulting in non-random segregation of daughter plasmids[65], whereas KSHV genomes cluster at tethering sites resulting in random segregation[46]. Human papillomavirus (HPV) also maintains its genome as extrachromosomal DNA in the basal epithelial layer, tethering to host chromosomes and enabling the viral DNA to persist in daughter cells[64]. During the lytic phase of infection, herpesviridae rapidly amplify their genome through DNA circularisation and subsequent rolling circle amplification or via a theta structure[67,68].

Future perspectives

2020 is likely to be an exciting year in the field of ecDNA biology. There are a number of areas that can be addressed with current tools. The first centres around basic ecDNA biology. There is a need to detail the mechanisms generating chromosomal instability in ecDNA formation, and further describe the role of double strand break and NHEJ pathways in this process. Furthermore, the role of ecDNA re-integration in genomic remodelling and its relocation as HSR are still to be determined. It has been shown that HSR formation is implicated in drug resistance; how, why and where this occurs is still largely unknown.

We are starting to understand the role of ecDNA in tumour evolution. Nevertheless, there are many areas to address in this field, such as the impact of ecDNA on clonal competition (or cooperation), the timing of ecDNA formation (in the context, for example, of whole genome doubling), and whether events such as aneuploidy, or whole genome doubling predispose to ecDNA formation.

As we develop our understanding of ecDNA biology, we will be able to address some of the more clinical aspects of ecDNA pathogenesis. The detection of ecDNA in circulating tumour DNA may yield valuable information towards prognosis and treatment strategy. This will be further informed if we can describe the landscape of ecDNA-mediated drug resistance and determine if it is possible to therapeutically target ecDNA formation. Describing the role ecDNA in neoantigen formation and the immune microenvironment may also open therapeutic avenues.

Robust detection methods are needed. This includes refining current sequencing and purification methods to isolate circular amplicons, and the development of computational tools that utilise technologies such as long-read and single-cell sequencing to better construct circular amplicons. Using these technologies to address the temporal dimension of ecDNA formation and integration will prove challenging.

Recent evidence has shown ecDNA as an important player in the field of tumour evolution. Relieved of hereditary constraints and conventional segregation laws, its presence leads to rapidly gained selection advantages that enhance aggressive tumour

behaviour. Further understanding of this process can only serve to advance our treatment strategies.

Conflicts

C.S. receives grant support from Pfizer, AstraZeneca, BMS, Roche-Ventana, Boehringer-Ingelheim, and Ono. C.S. has consulted for Pfizer, Novartis, GlaxoSmithKline, MSD, BMS, Celgene, AstraZeneca, Illumina, Genentech, Roche-Ventana, GRAIL, Medicxi, and the Sarah Cannon Research Institute. C.S. is a shareholder of Apogen Biotechnologies, Epic Bioscience, GRAIL, and has stock options in and is co-founder of Achilles Therapeutics.

Acknowledgements

C.S. is Royal Society Napier Research Professor. This work was supported by the Francis Crick Institute that receives its core funding from Cancer Research UK (FC001169,FC001202), the UK Medical Research Council (FC001169, FC001202), and the Wellcome Trust (FC001169, FC001202). C.S. is funded by Cancer Research UK (TRACERx, PEACE and CRUK Cancer Immunotherapy Catalyst Network), the CRUK Lung Cancer Centre of Excellence, the Rosetrees Trust, NovoNordisk Foundation (ID16584) and the Breast Cancer Research Foundation (BCRF). This research is supported by a Stand Up To Cancer-LUNGevity-American Lung Association Lung Cancer Interception Dream Team Translational Research Grant (Grant Number: SU2C-AACR-DT23-17). Stand Up To Cancer is a program of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of SU2C. CS receives funding from the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) Consolidator Grant (FP7-THESEUS-617844), European Commission ITN (FP7-PloidyNet 607722), an ERC Advanced Grant (PROTEUS) from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 835297), and Chromavision from the European Union's Horizon 2020 research and innovation programme (grant agreement 665233). M.J.S. is funded by the Arnold O. Beckman Postdoctoral Independence Award (UAJAL).

References

1. Verhaak, R. G. W., Bafna, V. & Mischel, P. S. Extrachromosomal oncogene amplification in tumour pathogenesis and evolution. *Nat. Rev. Cancer* **19**, 283–288 (2019).

2. Cox, D., Yuncken, C. & Spriggs, A. Minute chromatin bodies in malignant tumours of childhood. *Lancet* **286**, 55–58 (1965).

3. Haber, D. A. & Schimke, R. T. Unstable amplification of an altered dihydrofolate reductase gene associated with double-minute chromosomes. *Cell* **26**, 355–362 (1981).

4. Kohl, N. E. *et al.* Transposition and amplification of oncogene-related sequences in human neuroblastomas. *Cell* **35**, 359–367 (1983).

5. Alitalo, K., Schwab, M., Lin, C. C., Varmus, H. E. & Bishop, J. M. Homogeneously staining chromosomal regions contain amplified copies of an abundantly expressed cellular oncogene (c-myc) in malignant neuroendocrine cells from a human colon carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 1707–1711 (1983).

6. Gebhart, E., Bruderlein, S., Tulusan, A. H., Maillot, K. V. & Birkmann, J. Incidence of double minutes, cytogenetic equivalents of gene amplification, in human carcinoma cells. *Int. J. Cancer* (1984). doi:10.1002/ijc.2910340313

7. Benner, S. E., Wahl, G. M. & Von Hoff, D. D. Double minute chromosomes and homogeneously staining regions in tumors taken directly from patients versus in human tumor cell lines. *Anti-cancer drugs* (1991). doi:10.1097/00001813-199102000-00002

8. Gebhart, E. Double minutes, cytogenetic equivalents of gene amplification, in human neoplasia - a review. *Clin. Transl. Oncol.* **7**, 477–485 (2005).

9. Turner, K. M. *et al.* Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122–125 (2017).

10. Wu, S. *et al.* Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **575**, 699–703 (2019).

11. Kim, H., Nguyen, N., Turner, K., Wu, S. & Liu, J. Frequent extrachromosomal oncogene amplification drives aggressive tumors. *bioRxiv* (2019). doi:10.1101/859306

12. Vinograd, J., Lebowitz, J. & Watson, R. Early and late helix-coil transitions in closed circular DNA the number of superhelical turns in polyoma DNA. *J. Mol. Biol.* **33**, 173–197 (1968).

13. Paulsen, T., Kumar, P., Koseoglu, M. M. & Dutta, A. Discoveries of Extrachromosomal Circles of DNA in Normal and Tumor Cells. *Trends Genet.* **34**, 270–278 (2018).

14. Koche, R. P. *et al.* Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. *Nat. Genet.* **52**, 29–34 (2020).

15. Smith, C. A. & Vinograd, J. Small polydisperse circular DNA of HeLa cells. *J. Mol. Biol.* **69**, 163–178 (1972).

16. Shoura, M. J. *et al.* Intricate and cell type-specific populations of endogenous circular DNA (eccDNA) in Caenorhabditis elegans and Homo sapiens. *G3 Genes, Genomes, Genet.* **7**, 3295–3303 (2017).

17. Paulsen, T., Shibata, Y., Kumar, P., Dillon, L. & Dutta, A. Small extrachromosomal circular DNAs, microDNA, produce short regulatory RNAs that suppress gene expression independent of canonical promoters. *Nucleic Acids Res.* **47**, 4586–4596 (2019).

18. Garsed, D. W. *et al.* The Architecture and Evolution of Cancer Neochromosomes. *Cancer Cell* **26**, 653–667 (2014).

19. Levan, A. & Levan, G. Have double minutes functioning centromeres? *Hereditas* **88**, 81–92 (1978).

20. Pichugin, Y., Huang, W. & Werner, B. Stochastic dynamics of extra-chromosomal DNA. *bioRxiv* 2019.12.15.876714 (2019). doi:10.1101/2019.12.15.876714

21. Turner, K. M. *et al.* Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122–125 (2017).

22. Carroll, S. M. *et al.* Double minute chromosomes can be produced from precursors

derived from a chromosomal deletion. *Mol. Cell. Biol.* **8**, 1525–1533 (1988).

23. Vogt, N. *et al.* Amplicon rearrangements during the extrachromosomal and intrachromosomal amplification process in a glioma. *Nucleic Acids Res.* **42**, 13194–13205 (2014).

24. L'Abbate, A. *et al.* Genomic organization and evolution of double minutes/homogeneously staining regions with MYC amplification in human cancer. *Nucleic Acids Res.* **42**, 9131–9145 (2014).

25. Biedler, J. & Spengler, B. Metaphase chromosome anomaly: association with drug resistance and cell-specific products. *Science (80-.).* **191**, 185–187 (1976).

26. Shimizu, N. Extrachromosomal double minutes and chromosomal homogeneously staining regions as probes for chromosome research. *Cytogenet. Genome Res.* **124**, 312–326 (2009).

27. Menghi, F. *et al.* The Tandem Duplicator Phenotype Is a Prevalent Genome-Wide Cancer Configuration Driven by Distinct Gene Mutations. *Cancer Cell* **34**, 197-210.e5 (2018).

28. Storlazzi, C. T. *et al.* Gene amplification as doubleminutes or homogeneously staining regions in solid tumors: Origin and structure. *Genome Res.* **20**, 1198–1206 (2010).

29. Nathanson, D. A. *et al.* Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science (80-.).* **343**, 72–76 (2014).

30. Singer, M. J., Mesner, L. D., Friedman, C. L., Trask, B. J. & Hamlin, J. L. Amplification of the human dihydrofolate reductase gene via double minutes is initiated by chromosome breaks. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 7921–7926 (2000).

31. Reams, A. B. & Roth, J. R. Mechanisms of Gene Duplication and Amplification. *Cold Spring Harb. Perspect. Biol.* **7**, a016592 (2015).

32. Cai, M. *et al.* Inhibiting homologous recombination decreases extrachromosomal amplification but has no effect on intrachromosomal amplification in methotrexate-resistant colon cancer cells. *Int. J. Cancer* **144**, 1037–1048 (2019).

33. Meng, X. *et al.* Novel role for non-homologous end joining in the formation of double minutes in methotrexate-resistant colon cancer cells. *J. Med. Genet.* **52**, 135–144 (2015).

34. Lucas, J. S., Zhang, Y., Dudko, O. K. & Murre, C. 3D trajectories adopted by coding and regulatory DNA elements: First-passage times for genomic interactions. *Cell* **158**, 339–352 (2014).

35. Stephens, P. J. *et al.* Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell* **144**, 27–40 (2011).

36. L'Abbate, A. *et al.* MYC-containing amplicons in acute myeloid leukemia: genomic structures, evolution, and transcriptional consequences. *Leukemia* **32**, 2152–2166 (2018).

37. Huh, Y. O. *et al.* Double minute chromosomes in acute myeloid leukemia, myelodysplastic syndromes, and chronic myelomonocytic leukemia are associated with micronuclei, MYC or MLL amplification, and complex karyotype. *Cancer Genet.* **209**, 313–320 (2016).

38. Ly, P. *et al.* Chromosome segregation errors generate a diverse spectrum of simple and complex genomic rearrangements. *Nat. Genet.* **51**, 705–715 (2019).

39. Shimizu, N., Miura, Y., Sakamoto, Y. & Tsutsui, K. Plasmids with a mammalian replication origin and a matrix attachment region initiate the event similar to gene amplification. *Cancer Res.* **61**, 6987–6990 (2001).

40. Shimizu, N., Misaka, N. & Utani, K. Nonselective DNA damage induced by a replication inhibitor results in the selective elimination of extrachromosomal double minutes from human cancer cells. *Genes, Chromosom. Cancer* **46**, 865–874 (2007).

41. Shimizu, N. *et al.* Loss of amplified c-myc genes in the spontaneously differentiated HL-60 cells. *Cancer Res.* **54**, 3561–7 (1994).

42. 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).

43. Valent, A. *et al.* In vivo elimination of acentric double minutes containing amplified MYCN from neuroblastoma tumor cells through the formation of micronuclei. *Am. J. Pathol.* **158**, 1579–1584 (2001).

44. Clarke, T. L. *et al.* Histone Lysine Methylation Dynamics Control EGFR DNA Copy-Number Amplification. *Cancer Discov.* **10**, 306–325 (2020).

45. Black, J. C. *et al.* Hypoxia drives transient site-specific copy gain and drug-resistant gene expression. *Genes Dev.* **29**, 1018–1031 (2015).

46. Morton, A. R. *et al.* Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell* **179**, 1330-1341.e13 (2019).

47. Helmsauer, K. *et al.* Enhancer hijacking determines intra- and extrachromosomal circular MYCN amplicon architecture in neuroblastoma. *bioRxiv* (2019).

48. Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol.* **15**, R84 (2014).

49. Rausch, T. *et al.* DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).

50. Deshpande, V. *et al.* Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. *Nat. Commun.* **10**, (2019).

51. Shibata, Y. *et al.* Extrachromosomal MicroDNAs and Chromosomal Microdeletions in Normal Tissues. *Science (80-.).* **336**, 82–86 (2012).

52. Møller, H. D., Parsons, L., Jørgensen, T. S., Botstein, D. & Regenberg, B. Extrachromosomal circular DNA is common in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3114–E3122 (2015).

53. Kumar, P. *et al.* ATAC-seq identifies thousands of extrachromosomal circular DNA in cancers and cell lines. *bioRxiv Genomics* doi:10.1101/845040

54. Rajkumar, U. *et al.* EcSeg: Semantic Segmentation of Metaphase Images Containing Extrachromosomal DNA. *iScience* **21**, 428–435 (2019).

55. Neel, D. S. & Bivona, T. G. Resistance is futile: overcoming resistance to targeted therapies in lung adenocarcinoma. *npj Precis. Oncol.* **1**, 1–5 (2017).

56. Dagogo-Jack, I. & Shaw, A. T. Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* **15**, 81–94 (2018).

57. Engel, S. R. *et al.* The Reference Genome Sequence of Saccharomyces cerevisiae: Then and Now. *G3 Genes, Genomes, Genet.* **4**, 389–398 (2014).

58. Heap, I. & Duke, S. O. Overview of glyphosate-resistant weeds worldwide. *Pest Manag. Sci.* **74**, 1040–1049 (2018).

59. Koo, D.-H. *et al.* Extrachromosomal circular DNA-based amplification and transmission of herbicide resistance in crop weed Amaranthus palmeri. *Proc. Natl. Acad. Sci.* **115**, 3332–3337 (2018).

60. Carattoli, A. Resistance plasmid families in Enterobacteriaceae. *Antimicrob. Agents Chemother.* **53**, 2227–2238 (2009).

61. Liu, Y. Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: A microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).

62. Zhi, C., Lv, L., Yu, L. F., Doi, Y. & Liu, J. H. Dissemination of the mcr-1 colistin resistance gene. *Lancet Infect. Dis.* **16**, 292–293 (2016).

63. Zimmermann, J. & Hammerschmidt, W. Structure and role of the terminal repeats of Epstein-Barr virus in processing and packaging of virion DNA. *J. Virol.* **69**, 3147–3155 (1995).

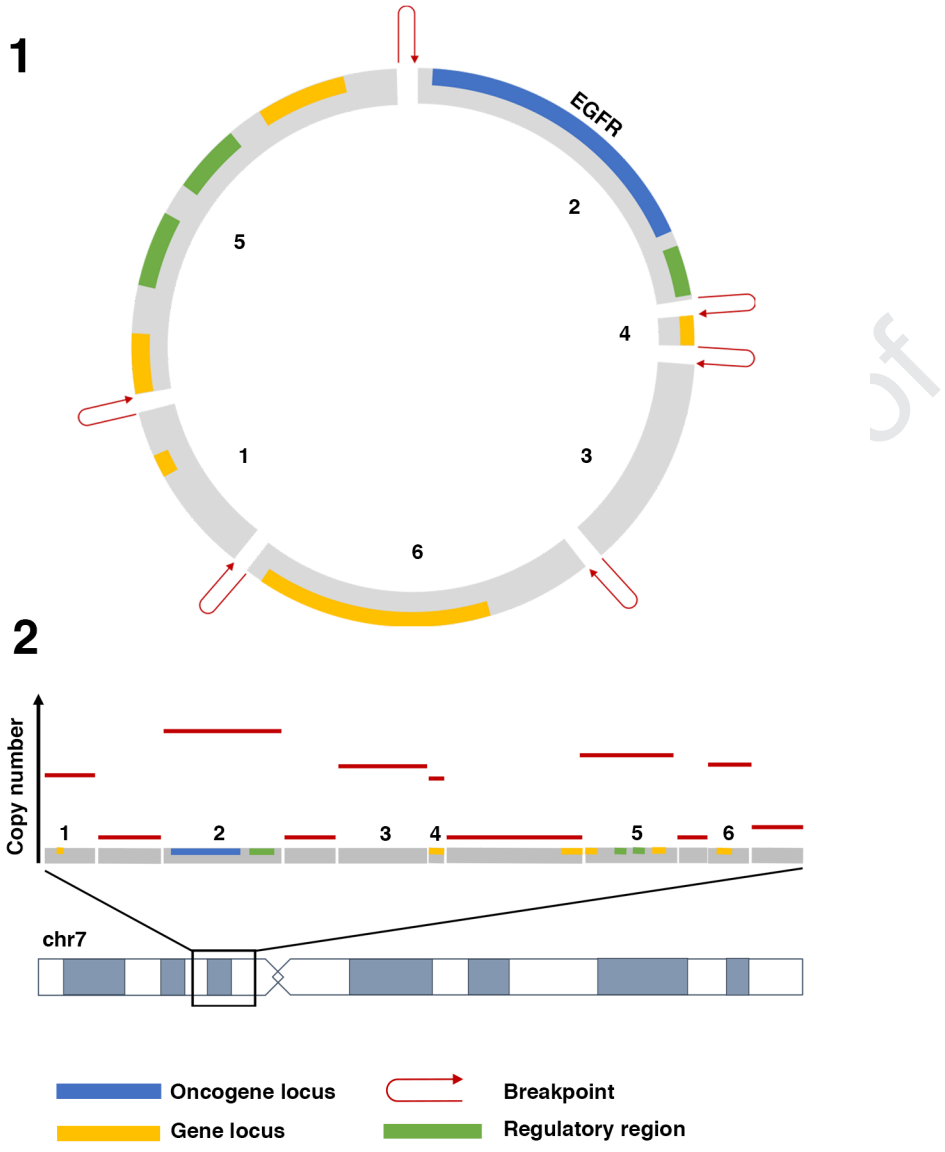
64. Coursey, T. L. & McBride, A. A. Hitchhiking of Viral Genomes on Cellular Chromosomes. *Annu. Rev. Virol.* **6**, 275–296 (2019).

65. Nanbo, A., Sugden, A. & Sugden, B. The coupling of synthesis and partitioning of EBV's plasmid replicon is revealed in live cells. *EMBO J.* **26**, 4252–4262 (2007).

66. Chiu, Y. & Sugden, B. Plasmid Partitioning by Human Tumor Viruses. *J. Virol.* **92**, 1–8 (2018).

67. Lehman, I. R. & Boehmer, P. E. Replication of Herpes Simplex Virus DNA. *J. Biol. Chem.* **274**, 28059–28062 (1999).

68. Pfüller, R. & Hammerschmidt, W. Plasmid-like replicative intermediates of the Epstein-Barr virus lytic origin of DNA replication. *J. Virol.* **70**, 3423–3431 (1996).


1
EGFR
2
5
4
1
3
6
2
Copy number
1
2
3
4
5
6
chr7
Oncogene locus
Gene locus
Breakpoint
Regulatory region