eccDNAs are apoptotic products with high innate immunostimulatory activity

https://doi.org/10.1038/s41586-021-04009-w

Received: 16 February 2021

Accepted: 7 September 2021

Published online: 20 October 2021

Check for updates

Yuangao Wang[1,2], Meng Wang[1,2], Mohamed Nadhir Djekidel[1,2], Huan Chen[1,2], Di Liu[3,4], Frederick W. Alt[1,2,5] & Yi Zhang[1,2,5,6]✉

Extrachromosomal circular DNA elements (eccDNAs) have been described in the literature for several decades, and are known for their broad existence across different species[1,2]. However, their biogenesis and functions are largely unknown. By developing a new circular DNA enrichment method, here we purified and sequenced full-length eccDNAs with Nanopore sequencing. We found that eccDNAs map across the entire genome in a close to random manner, suggesting a biogenesis mechanism of random ligation of genomic DNA fragments. Consistent with this idea, we found that apoptosis inducers can increase eccDNA generation, which is dependent on apoptotic DNA fragmentation followed by ligation by DNA ligase 3. Importantly, we demonstrated that eccDNAs can function as potent innate immunostimulants in a manner that is independent of eccDNA sequence but dependent on eccDNA circularity and the cytosolic DNA sensor Sting. Collectively, our study not only revealed the origin, biogenesis and immunostimulant function of eccDNAs but also uncovered their sensing pathway and potential clinical implications in immune response.

Since its first description in wheat embryos and boar sperm in 1964 (ref. [3]), extrachromosomal circular DNA (eccDNA) has been reported in almost all cell lines and tissues[4] across different species, although its abundance is highly variable[1,2]. Unlike the independently existing circular DNA in organelles, such as mitochondrial DNA (mtDNA), eccDNAs are derived from genomic DNA and range in size from a few hundred bases to megabases[1]. Although some studies have suggested that eccDNA generation might be linked to DNA damage repair[5], hyper-transcription[5,6], homologous recombination[7] and replication stress[5], how exactly eccDNAs are generated is largely unknown. Similarly, it is also unclear whether eccDNA has any function, although some studies have suggested that eccDNAs might contribute to gene amplification in cancer[1] or might be linked to ageing[6–8].

To understand eccDNA biogenesis, efficient and robust methods that allow purification and sequencing of eccDNAs are needed. Most existing eccDNA purification procedures involve two sequential steps, isolation of crude extrachromosomal DNA followed by removal of contaminating linear DNA through exonuclease digestion[5,6,9] and rolling circle amplification (RCA) for profiling[5,9,10]. However, most eccDNA samples prepared in this way contain a high level of linear DNA before RCA, as revealed by electron microscopy[5,9], indicating that exonuclease digestion alone is not sufficient to eliminate all contaminating linear DNA.

An efficient eccDNA purification method

We have developed a new three-step eccDNA enrichment method that allows efficient eccDNA purification (Fig. 1a). In the first step, to minimize eccDNA loss, we replaced the conventional unbuffered sodium hydroxide lysis, which may cause irreversible denaturation or breakage of DNA circles[11], with a modified alkaline buffer at pH 11.8 to lyse the whole cells. In the second step, we used the rare cutter PacI restriction enzyme to linearize mtDNA before addition of an exonuclease (Plasmid-Safe ATP-dependent DNase) to digest linear DNA. In the third step, solution A, which could selectively recover circular DNA, but not linear DNA, on silica beads (Extended Data Fig. 1a), was used to exclude any linear DNA that escaped exonuclease digestion (Fig. 1a). Additionally, vertical agarose gel electrophoresis was used to increase the sensitivity of eccDNA detection (Extended Data Fig. 1b). Using this three-step purification procedure, we purified eccDNA from 10 million HeLa cells growing at confluence, a stress condition known to increase eccDNA abundance[12]. The purified eccDNAs exhibited a discrete banded pattern (Fig. 1b). Furthermore, mtDNA could be removed by PacI treatment (Fig. 1b, compare lanes 1 and 2). We further confirmed the circularity of the purified eccDNAs using scanning atomic-force microscopy (SAFM) (Fig. 1c). To determine whether eccDNAs occur in non-cancer cells, mouse embryonic stem cells (mESCs) were used, and mESC-derived eccDNAs exhibited a similar banded pattern (Fig. 1d), and their purity and circularity were also verified by SAFM (Fig. 1e).

eccDNAs map to the entire genome

To gain insights into the potential mechanism of eccDNA biogenesis, we determined the genomic source of eccDNAs. HeLa cells are notorious for their aberrant genome, including aneuploidy and numerous structural variations, such as deletions, duplications, inversions, translocations and rearrangements, etc.[13], making interpretation of sequencing

[1]Howard Hughes Medical Institute, Boston, MA, USA. [2]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA. [3]Wyss Institute for Biologically Inspired Engineering, Harvard University, Boston, MA, USA. [4]Department of Systems Biology, Harvard Medical School, Boston, MA, USA. [5]Department of Genetics, Harvard University, Boston, MA, USA. [6]Harvard Stem Cell Institute WAB-149G, Boston, MA, USA. ✉e-mail: yzhang@genetics.med.harvard.edu

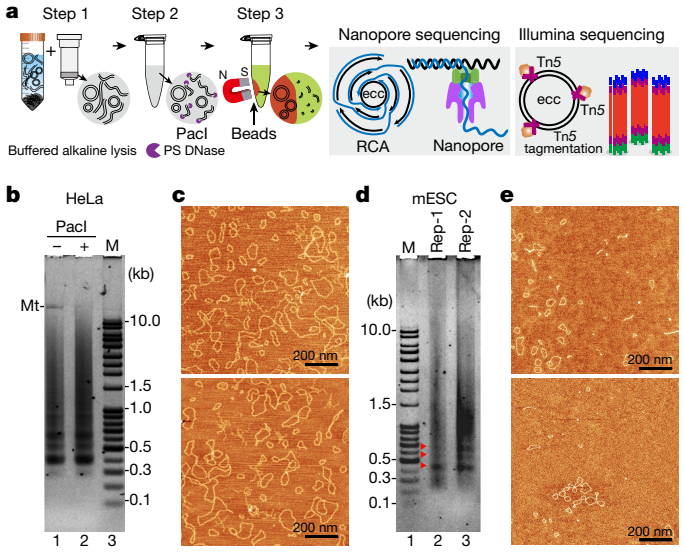


Fig. 1 | Development of a three-step eccDNA purification procedure. **a**, Schematic of the three-step eccDNA purification and sequencing procedure. Step 1, extract crude DNA circles from whole cells in a buffered alkaline lysis solution and bind them to a silica column; step 2, linearize mtDNA with PacI and reduce overall linear DNA levels with Plasmid-Safe (PS) DNase; step 3, selectively recover eccDNAs by excluding residual linear DNA in solution A. eccDNAs are then sequenced by Oxford Nanopore sequencing after RCA (left) or by Illumina sequencing after Tn*5* tagmentation on eccDNAs (right). **b**, Agarose gel showing eccDNAs purified from over-confluent HeLa cells without (–) or with (+) PacI treatment. M, linear DNA maker; Mt, mtDNA. **c**, eccDNAs in **b** (lane 2: top image; lane 3: bottom image) scanned with SAFM. **d**, Agarose gel showing eccDNAs purified from normal cultured mESCs. Red arrowheads indicate distinct DNA bands. **e**, eccDNAs in **d** (lane 2: top image; lane 3: bottom image) scanned with SAFM. In **b** and **d**, representative gels are shown from three independent experiments. In **c** and **e**, two representative fields are shown.

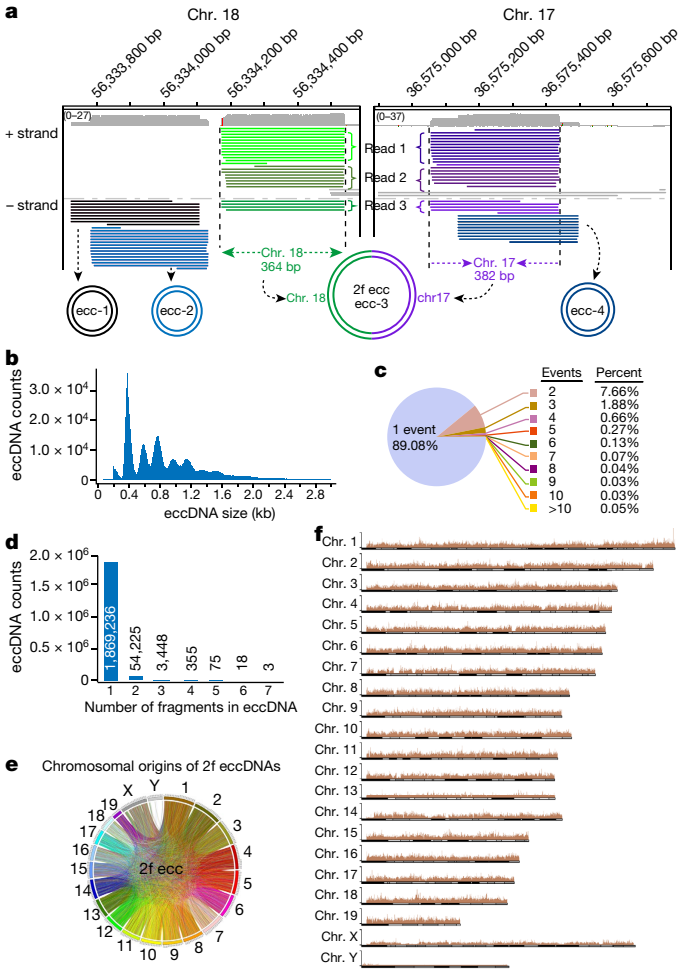


Fig. 2 | eccDNAs are circularized genomic DNA fragments that map across the genome. **a**, Integrative Genomics Viewer (IGV) alignments showing eccDNA examples from two genomic loci on chromosomes 17 and 18. Individual horizontal bars in the same colour represent subreads from a unique Nanopore long read that was repeatedly aligned to the same genomic locus or loci. ecc-1, ecc-2 and ecc-4 are single-fragment circles. ecc-1 partially overlaps ecc-2; ecc-4 partially overlaps one fragment of ecc-3, which is a two-fragment circle (2f eccDNA) aligned to two loci on chromosomes 17 and 18. **b**, Histogram showing eccDNA size distribution and relative abundance. **c**, Pie chart showing the percentages of eccDNAs with the indicated event numbers among the total unique eccDNAs identified. **d**, Bar graph showing eccDNA counts with the indicated number of fragments (1–7) in each circle. **e**, Circle plot showing the chromosomal origin of all two-fragments eccDNAs. Sub-reads from the same chromosome are in the same colour. **f**, Overall chromosomal distribution of eccDNAs across the genome.

data and dissection of the eccDNA biogenesis mechanism difficult. Thus, we performed eccDNA sequencing and mapping using mESCs, whose genetic integrity is maintained during culture[14]. To obtain full-length eccDNA sequences, we performed RCA and subsequent long-read Nanopore sequencing of the multiple tandem copies of individual eccDNA molecules (Fig. 1a and Extended Data Fig. 2a). Repeated sequencing of the same eccDNA with long reads allows generation of a consensus sequence matching the full-length sequence of the original eccDNA by a computational threading method (Fig. 2a and Extended Data Fig. 2c). We obtained 4 million long reads with a mean size of 3.7 kb (Extended Data Fig. 2b). To reduce false positives due to RCA artefacts and sequencing errors from the Nanopore technique[15], we only used the 1.9 million long reads that each contained at least two full passes covering their original eccDNA molecule to identify high-confidence eccDNAs, resulting in the identification of 1.6 million unique eccDNAs with a median size of 1 kb (Extended Data Fig. 2b). Interestingly, the eccDNAs exhibited a regular average size interval of 188 bp (Fig. 2b). The great majority (89%) of unique eccDNAs were sequenced from a single long read (single-event eccDNA), with less than 1.5% of unique eccDNAs sequenced from more than three unique long molecules (Fig. 2c); no dominant eccDNA was identified. Such large numbers of single-event eccDNAs coupled with the lack of dominant eccDNA molecules suggest that eccDNAs are unlikely to be derived from specific genome regions.

Genome mapping of full-length eccDNAs revealed their various genomic alignment patterns, including at adjacent, overlapped or nested positions on the same chromosome or even across different chromosomes (Fig. 2a). We found that a great majority of eccDNAs originated from single continuous genomic loci (continuous eccDNAs, self-circularization of a single genomic fragment), and only a relatively small number of eccDNAs were formed from multiple genomic fragments (non-continuous eccDNAs, circularization of multiple genomic fragments) (Fig. 2d and Extended Data Fig. 2c), including three eccDNAs each with seven genomic fragments joined together to form a circle (7f eccDNA) (Fig. 2d). To determine whether the physical distance between genomic fragments affects the frequency of eccDNA formation, we analysed the genomic origin of two-fragment eccDNAs (2f eccDNAs). A circle plot clearly showed that paired fragments of 2f eccDNAs are not restricted to the same chromosome (Fig. 2e), but rather are randomly bridged between chromosomes, indicating that eccDNAs can be formed by joining genomic fragments from different chromosomes. Consistent with this, genome mapping of all eccDNAs revealed that eccDNAs are widespread across the entire genome (Fig. 2f).

To rule out potential biases caused by uneven amplification by RCA[16], we purified another batch of eccDNAs and directly tagged them with Tn*5* transposase without RCA for Illumina sequencing (Fig. 1a and Extended Data Fig. 3a). eccDNA sequences obtained in this way should faithfully reveal their genomic location and relative abundance. Consistent with the Nanopore sequencing results, Illumina sequencing showed widespread alignment of eccDNAs across the entire genome (Extended Data Fig. 3b). We noticed that the eccDNA density on the X chromosome was about half that of the autosomes (Fig. 2f and Extended Data Fig. 3b), consistent with the fact that the diploid male genome of mESC/E14 cells carries one copy of the X chromosome but two copies of each autosome (Fig. 2f and Extended Data Fig. 3b). The lack of eccDNAs mapped to the Y chromosome is largely due to the many undetermined sequences and repeat sequences on the Y chromosome[17]. Collectively, these data suggest that eccDNAs are widespread across the entire genome and their abundance is correlated with genomic copy numbers.

DNase γ is required for eccDNA generation

The great diversity, randomness and nucleosome 'ladder' size (Fig. 2) suggest that eccDNAs might be generated by random ligation (including self-ligation) of oligonucleosomal DNA fragments, which can be visualized as 'ladders' in agarose gel and are a known feature of apoptosis[18]. To determine whether apoptotic cells are the source of eccDNAs, mESCs were treated with the apoptosis inducers staurosporine, etoposide or UV light. Successful induction of apoptosis was confirmed by the typical nucleosomal ladder pattern of genomic DNA (Fig. 3a). When equal amounts of control (DMSO-treated) cells and apoptosis inducer-treated cells were subjected to the three-step eccDNA purification procedure and visualized on an agarose gel, all three treatments induced eccDNA production, although UV treatment resulted in the strongest induction (Fig. 3b and Extended Data Fig. 4a).

We next determined whether eccDNA generation requires apoptotic DNA fragmentation (ADF), which is mediated by caspase-activated DNase (CAD)[19], endonuclease G (EndoG)[20] or DNase γ[21] in a cell-type-specific manner. Genetic manipulation (Extended Data Fig. 4b, c) indicated that DNase γ, but not EndoG, mediates ADF in mESCs, as indicated by the lack of a ladder pattern with DNase γ (encoded by *Dnase1l3*) knockout[21] (Fig. 3c). *Dnase1l3* knockout did not affect cell viability under either normal culture conditions or UV treatment (Extended Data Fig. 4d, e). Purification of the eccDNAs from UV-treated cells demonstrated that abrogation of ADF prevented eccDNA generation (Fig. 3d and Extended Data Fig. 4f). We intentionally skipped PacI digestion to retain mtDNA as an internal control for equal cell input and circular DNA recovery (Fig. 3d). These results demonstrate that ADF is a prerequisite for eccDNA generation.

Lig3 is required for eccDNA generation

Next, we attempted to identify the DNA ligase responsible for circularizing the fragmented DNA. Mammals have three DNA ligase genes (*Lig1*, *Lig3* and *Lig4*), each of which has a specific function, although they also function redundantly in DNA metabolism[22]. The functions of these ligases have been well studied in the CH12F3 mouse B-lymphocyte cell line[23]. To determine which of the three DNA ligases is responsible for ADF circularization, individual DNA ligases and their combinations were knocked out in CH12F3 cells by CRISPR–Cas9 with knockout confirmed by western blotting (Fig. 3e). Lig3 has both nuclear and mitochondrial isoforms, the later of which is essential for mitochondria maintenance and, consequently, cell viability[24]. Thus, the *Lig3*-knockout cell line was generated by specifically targeting the nuclear isoform (*NucLig3*$^{-/-}$) without interfering with the mitochondrial isoform (*MtLig3*; Fig. 3e, lower diagram). Equal numbers of wild-type (WT) and mutant cells were treated with staurosporine to induce ADF (Fig. 3f), and eccDNAs were purified and visualized in agarose gel (Fig. 3g). The results indicated that knockout of *Lig1* or *Lig4* alone or in combination did not significantly affect eccDNA generation. In contrast, knockout of *Lig3* greatly reduced eccDNA generation (Fig. 3g and Extended Data Fig. 4g). Because double knockout of *Lig1* and *Lig3* is lethal to cells[23,24], it is unknown whether double knockout could completely abrogate eccDNA generation. Nevertheless, these data support Lig3 as the main ligase for eccDNA generation in CH12F3 cells.

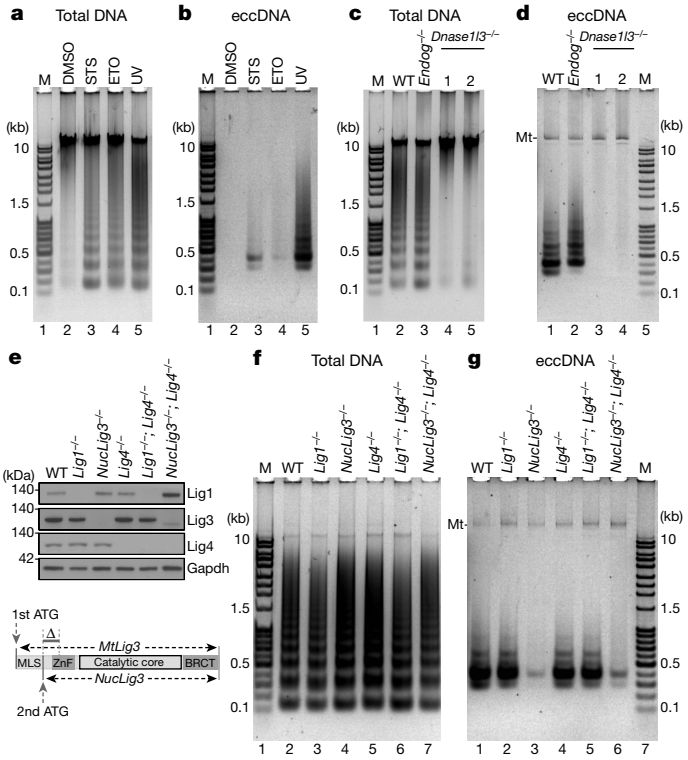


Fig. 3 | Apoptotic DNA fragmentation and subsequent ligation by Lig3 are required for eccDNA production in mESCs. a, **b**, eccDNA production is induced by apoptosis. mESCs were treated with the indicated apoptosis inducers, and total DNA (**a**) and eccDNAs (**b**) were purified. ETO, etoposide; STS, staurosporine. **c**, **d**, Deficiency of oligonucleosomal DNA fragmentation abolishes apoptosis-induced eccDNA production. Knockout of *Dnase1l3*, but not *Endog* (encoding endonuclease G), abolishes UV-induced oligonucleosomal DNA fragmentation (**c**) and eccDNA production (**d**) in mESCs. mtDNA was kept as an internal control. **e**, Confirmation of DNA ligase-deficient CH12F3 cell lines. Top, immunoblotting confirming knockout of DNA ligases in CH12F3 cell lines. Bottom, genomic structure of *Lig3* with CRISPR–Cas9 specifically targeting (Δ) *NucLig3* but retaining *MtLig3* for cell viability. **f**, **g**, Lig3 is the major DNA ligase for eccDNA generation. Shown are staurosporine-induced oligonucleosomal DNA fragmentation (**f**) and eccDNA (**g**) from the indicated CH12F3 cell lines. In **a**–**d**, **f** and **g**, the amount of input cells and elution and loading volumes were equal among the samples on each agarose gel. Shown are representatives of three independent experiments.

eccDNAs are potent innate immunostimulants

The above results demonstrate that eccDNAs are ligation products of fragmented genomic DNA from apoptotic cells. DNA released from dying cells has previously been reported to promote immune responses[18]. Two important mediators of immune response, Toll-like receptor 9 (TLR9)[25] and high-mobility-group box 1 (HMGB1)[26], have been reported to preferentially bind to DNA curvatures. These observations suggest that eccDNAs may serve as immunostimulants. To test this idea, we generated bone marrow-derived dendritic cells (BMDCs) (Extended Data Fig. 5a) and compared the immunostimulant activity of sheared linear genomic DNA, eccDNAs and the widely used potent DNA ligand

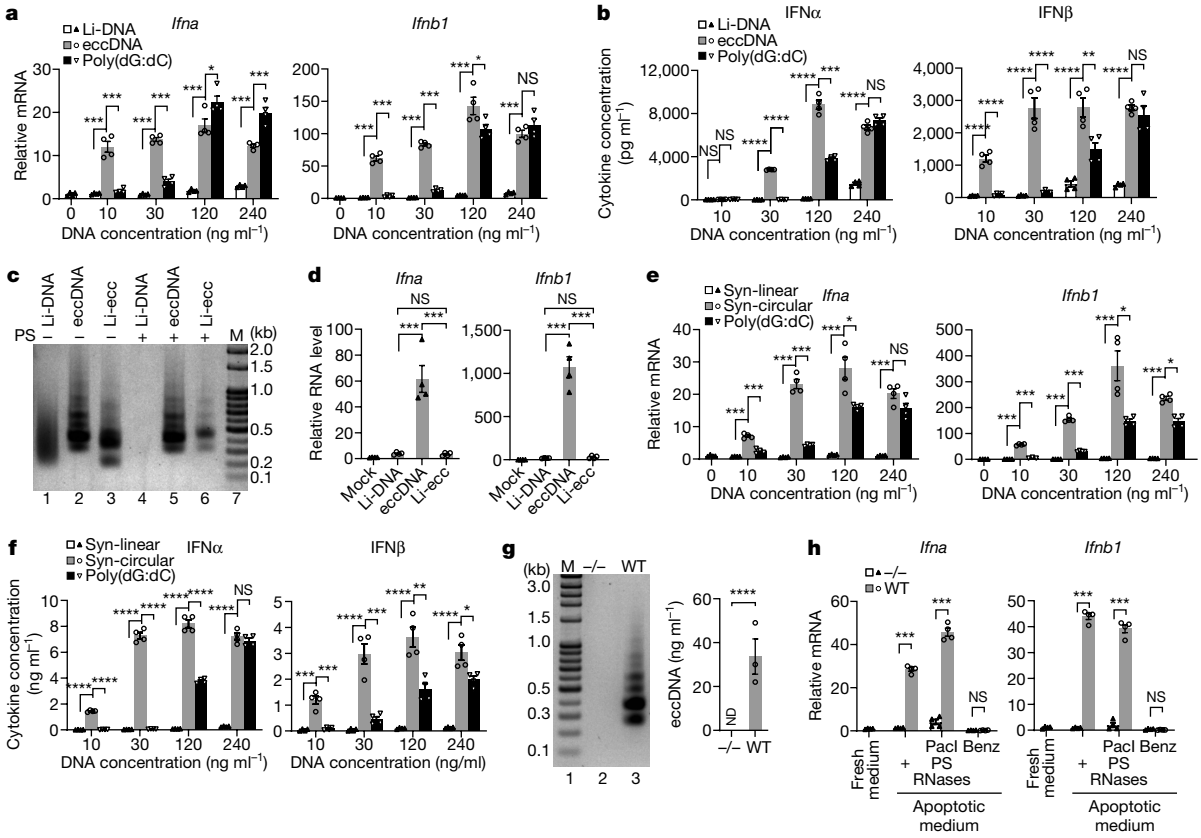


Fig. 4 | eccDNAs are potent immunostimulants. a, eccDNAs induce *Ifna* (encoding IFNα) and *Ifnb1* (encoding IFNβ) expression in BMDCs. Equal amounts of the indicated DNA were transfected into BMDCs at increasing concentrations for RT–qPCR analysis. Data are presented as relative mRNA fold change (*y* axis) with respect to that in mock transfected cells (without DNA). Li-DNA, sonicated genomic linear DNA with sizes similar to those of eccDNAs. **b**, ELISA analysis of IFNα and IFNβ production in medium from **a**. **c**, Confirmation of eccDNA linearization. A representative gel shows equal amounts of the indicated DNA digested with Plasmid-Safe DNase or left undigested. Li-ecc, linearized eccDNA. **d**, Linearized eccDNAs lose their immunostimulatory activities. The DNAs from lanes 1–3 in **c** were transfected into BMDCs at 30 ng ml^{-1} and mRNA levels were evaluated as in **a**. **e**, **f**, Synthetic small DNA circles are potent immunostimulants. The same experiments were performed as in **a** and **b**, except that the linear DNA and eccDNA were replaced by synthetic linear (Syn-linear) and circular (Syn-circular) DNAs with the same sequence. **g**, eccDNAs are present in apoptotic medium from WT cells but not *Dnase1l3*$^{-/-}$ (−/−) cells. A representative gel of eccDNA (left) and quantification (right; $n = 3$) are shown; ND, not detected. **h**, Exonuclease-resistant DNA (not mtDNA) in apoptotic medium activates *Ifna* and *Ifnb1* expression. Supernatants from apoptotic WT or *Dnase1l3*$^{-/-}$ (−/−) cells were treated as indicated, then incubated with BMDCs for RT–qPCR analysis as in **a**. Benz, benzonase. In **a**, **b**, **d**–**f** and **h**, data are shown as the mean ± s.e.m. of replicates ($n = 4$ per group) of a representative from three independent experiments. Statistics were calculated on biological replicates with ordinary one-way ANOVA with Tukey's multiple-comparison test (**a**, **b**, **d**–**f**, **h**) or two-tailed unpaired *t*-tests (**g**): *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, ****$P < 0.0001$; NS, not significant.

of cytosolic DNA sensors poly(dG:dC)[27] (Extended Data Fig. 5b). We transfected BMDCs with different amounts of the three forms of DNA and then collected cells for RT–qPCR assays. Compared with transfection with linear DNA, type I interferons (IFNα, IFNβ), interleukin-6 (IL-6) and tumour necrosis factor (TNF) were all significantly induced by eccDNAs at a wide range of concentrations (10–240 ng ml^{-1}) (Fig. 4a and Extended Data Fig. 5c). Surprisingly, the widely regarded 'potent' cytokine inducer poly(dG:dC) was not nearly as potent as eccDNAs at lower concentrations, and the linear DNA triggered only a mild response, even at the highest concentration, when compared with mock transfection, indicating that dendritic cells are much more sensitive to eccDNA treatment than that with linear DNA or poly(dG:dC). Consistent with this, enzyme-linked immunosorbent assays (ELISAs) confirmed the strong potency of eccDNAs in cytokine induction (Fig. 4b and Extended Data Fig. 5d).

In addition to dendritic cells, macrophages are also known to respond to immunostimulants[28]. To determine whether macrophages behave like dendritic cells upon eccDNA transfection, we generated bone marrow-derived macrophages (BMDMs) (Extended Data Fig. 6a). Similarly to the observations in BMDCs, eccDNAs also displayed much higher immunostimulant activities in BMDMs than linear DNA or poly(dG:dC), particularly at lower concentrations (10 and 30 ng ml^{-1}) (Extended Data Fig. 6b, c). These data indicate that eccDNAs are very potent immunostimulants in activating both BMDCs and BMDMs. Furthermore, pretreatment of eccDNA with DNase I before transfection completely abrogated the capacity of eccDNA to induce cytokine production (Extended Data Fig. 5e, f), demonstrating that eccDNA, rather than potential concomitants of eccDNA, is responsible for the immune activation.

Circularization confers eccDNA potency

To determine whether the circular nature of eccDNAs is critical for their strong immunostimulatory activity, purified eccDNAs were first treated with FnoCas12a (Cpf1), which introduces one nick per circular DNA in the presence of Mn^{2+} and the absence of guide RNA[29]. The nicked eccDNAs were then treated with the single-strand-specific endonuclease nuclease S1, which cleaved the intact circular strand at the site opposite to the nick to generate linearized eccDNA. Linearization of eccDNAs was confirmed by their sensitivity to exonuclease digestion while intact eccDNAs were resistant (Fig. 4c, compare lanes 5 and 6). When equal amounts of linear DNAs, eccDNAs and linear eccDNAs were transfected into BMDCs, linear eccDNAs behaved like linear DNAs and failed to activate IFNα, IFNβ,

IL-6 or TNF (Fig. 4d and Extended Data Fig. 7a). These results demonstrated that the circular nature of eccDNAs is critical for their strong immunostimulant activity. Because eccDNAs are derived from randomly ligated genomic fragments, their sequences are unlikely to significantly contribute to their potency. This notion was confirmed by the ability of a synthetic 200-bp circular DNA, but not its linear counterpart, to greatly induce cytokine genes transcription in BMDCs (Fig. 4e and Extended Data Fig. 7b). Similarly to native eccDNAs, synthetic circular DNA also showed higher potency in cytokine gene activation than poly(dG:dC) (Fig. 4e and Extended Data Fig. 7b). Consistent with these findings, ELISAs confirmed the strong cytokine induction capacity of synthetic circular DNA (Fig. 4f and Extended Data Fig. 7c).

To rule out the possibility that the increased immunostimulatory potency of circular DNA is due to increased stability and transfection efficiency, and to minimize the effects of exonuclease activity on linear DNA, we added phosphorothioate[30] bonds on both ends of the synthetic 200-bp linear DNA. To exclude potential effects from the phosphorothioate bonds on transfection and immune stimulation, equal numbers of phosphorothioate bonds were also added in the circular counterpart. Linear and circular DNAs were separately transfected into BMDCs, and cell lysates and culture media were collected for qPCR and ELISA to compare transfection efficiency (1 h after transfection), stability and cytokine induction (12 h after transfection) (Extended Data Fig. 7d). We found no significant difference in transfection efficiency or stability between the linear and circular 200-bp DNAs (Extended Data Fig. 7e). Yet, the circular DNA induced a high level of cytokine production while its linear counterpart did not (Extended Data Fig. 7f). Collectively, these data support the idea that the circularity but not the sequence of eccDNAs confers the high potency of their immunostimulant activity.

BMDCs sense eccDNAs in medium from apoptotic cells

Because eccDNAs are generated in apoptotic cells, they could be released into the culture medium. Indeed, a substantial amount of eccDNAs could be detected in the cell-free supernatant of UV-treated mESCs undergoing apoptosis (Fig. 4g). To determine whether eccDNAs from the supernatant of apoptotic cells can be actively sensed by BMDCs without transfection, BMDCs were co-incubated with cell-free apoptotic supernatant from WT or *Dnase1l3*$^{-/-}$ mESCs. RT–qPCR analysis indicated that the supernatant from WT, but not *Dnase1l3*$^{-/-}$, apoptotic cells stimulated IFNα and IFNβ expression (Fig. 4h). Importantly, this stimulation was not sensitive to pretreatment of the supernatant with Plasmid-Safe DNase, PacI restriction enzyme or RNases, which digest linear DNA, mtDNA and RNA, respectively (Fig. 4h), but was sensitive to pretreatment with benzonase, a nuclease that destroys all forms of DNA and RNA without proteolytic activity (Fig. 4h). These data indicate that eccDNAs, but not linear DNAs, mtDNAs or RNAs, in the supernatant of apoptotic cells are responsible for the induced immune response. Furthermore, these results also indicate that eccDNAs can be actively sensed by BMDCs without transfection. Collectively, our results indicate that eccDNAs are potent damage-associated molecular patterns of the innate immune system[28].

eccDNA-triggered immune response requires Sting

To assess the global transcriptional effect of eccDNA, we performed RNA-seq analysis of BMDCs transfected with purified eccDNAs or sonicated with genomic DNA of similar size (Extended Data Fig. 8a–c). Comparative analysis indicated that eccDNAs, but not the linear DNA control, significantly increased the expression of 290 genes (fold change ≥ 5, $P < 0.001$), including 34 cytokines and chemokines (Fig. 5a and Supplementary Table 4), under our experimental conditions (30 ng ml^{-1} DNA transfected). Importantly, 9 of the top 20 most upregulated genes belong to the family of type I interferons (Fig. 5b). Gene ontology (GO) enrichment analysis revealed that the upregulated genes were enriched in terms relevant to immune response and related signalling pathways (Fig. 5c), supporting our conclusion that eccDNA is a potent innate immunostimulant that can generally increase the innate immune response. Parallel experiments further demonstrated a similar effect of eccDNAs in BMDMs (Extended Data Fig. 9a–f and Supplementary Table 5). Collectively, these data support a higher capacity and potency of eccDNA in triggering a general immune response than linear genomic DNA fragments (Fig. 5b). Importantly, this eccDNA property depends on eccDNA circularization, but not sequence, as transfection of the 200-bp synthetic circular DNA into BMDCs triggered a similar transcriptional response as purified eccDNAs (Fig. 5d and Extended Data Fig. 10a).

To determine how eccDNA is sensed, two well-known mouse lines deficient in DNA sensing, with knockout of *Sting1* (stimulator of interferon genes)[31] or *Myd88* (myeloid differentiation primary response 88)[32], were used to generate BMDCs (Extended Data Fig. 10b), which were then subjected to eccDNA transfection. Comparative RNA-seq analysis demonstrated that, while loss of function of Myd88 did not affect BMDC responses to eccDNAs, loss of Sting function completely abrogated the capacity of BMDCs to respond to eccDNAs, as almost all the genes normally induced by eccDNA were not induced in the absence of Sting (Fig. 5e, f, Extended Data Fig. 10c and Supplementary Table 6). These data strongly suggest that the Sting pathway is responsible for sensing eccDNA to mediate its immune response.

Our data indicate that apoptotic oligonucleosomal DNA fragmentation (ODF) is directly linked to eccDNA generation, as blocking ADF abolishes eccDNA production (Fig. 3c, d). This notion challenges the assumption that eccDNAs isolated from a cell population or tissue are equally contributed by each cell[9]. On the contrary, our data suggest that eccDNAs are mostly derived from cells that undergo genomic DNA fragmentation. During apoptosis, genomic DNA is first broken into high-molecular-weight (HMW) fragments (>50 kb) and subsequently undergoes ODF to generate oligonucleosomal fragments[19,33]. Similarly to ODF in myoblasts[21] and neuroblastomas under differentiation conditions[34], we showed that DNase γ is required for ODF and subsequent eccDNA generation in mESCs (Fig. 3c, d). This suggests that the eccDNAs that we purified are generated in the late stage of apoptosis when ODF occurs[33], which is consistent with our observation that the great majority (99.5%) of eccDNAs are within 3 kb in size (Fig. 2b), although we identified eccDNAs as long as 10 kb. Although random ligation by Lig3 of nucleosome-sized DNA fragments in the late stage of apoptosis explains the predominantly oligonucleosomal sizes and the absence of abundant individual eccDNAs in our study and previous studies[5,9], it is possible that rare ligation of HMW fragments in the early stage of apoptosis might also occur. Our demonstration that Lig3 is responsible for nucleosome-sized eccDNA generation is consistent with the ability of Lig3 to circularize DNA fragments in vitro[35]. Whether Lig3 is also involved in the biogenesis of larger eccDNAs, double minutes[36] and extrachromosomal DNAs[37] remains to be determined. Although this study mainly focused on eccDNA generation under apoptotic conditions, we do not rule out the possibility that eccDNAs can also be generated under other conditions that cause genomic DNA fragmentation (for example, replication stress, double-strand DNA breaks, V(D)J recombination, etc.) with subsequent circularization.

We demonstrated that purified eccDNAs or synthetic circular DNA, but not their linear counterparts, have strong immunostimulatory activity (Figs. 4 and 5a–c). Importantly, Sting, but not Myd88, is required to mediate this process (Fig. 5e, f). cGAS–Sting is a well-known intracellular DNA sensing pathway[38], and DNA sensing by cGAS has been reported to be enhanced by the host factors HMGB1 and TFAM, which facilitate DNA bending or form U-shaped structures[39]. Whether these factors are involved in eccDNA-mediated immune stimulation remains to be determined. In addition to the immunostimulatory activity of eccDNA we showed in this study, it is not clear whether eccDNAs from

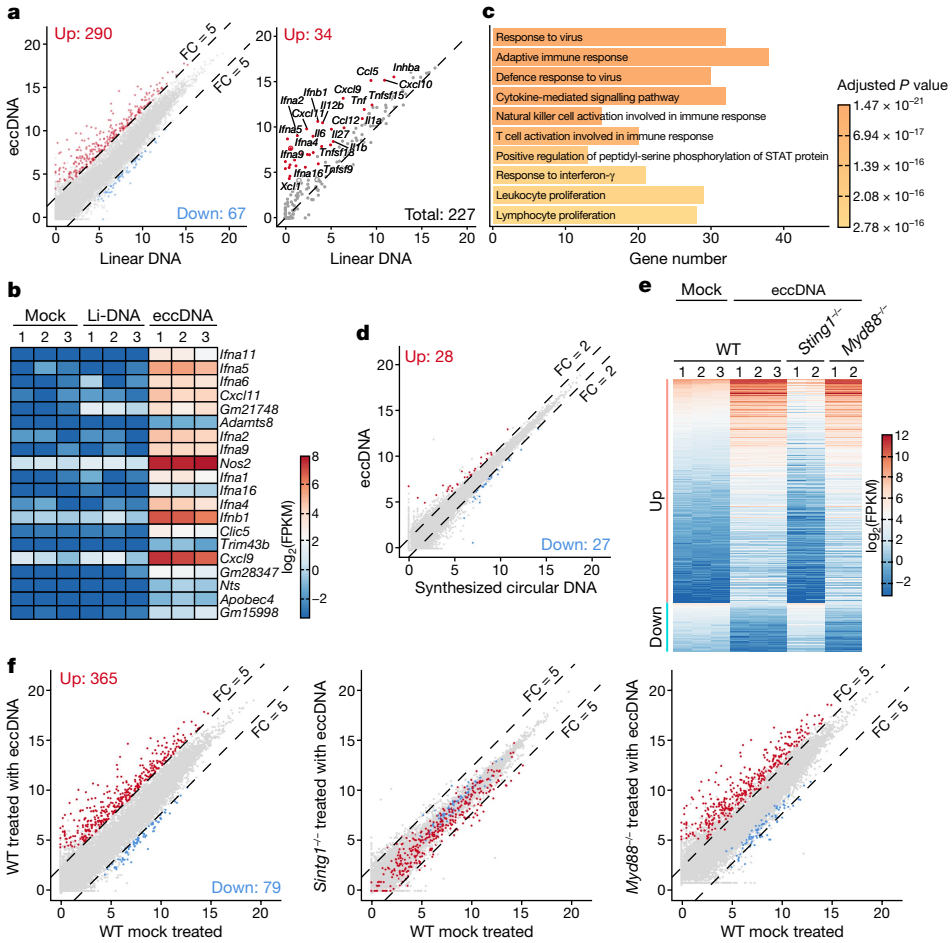


Fig. 5 | Sting is required for eccDNA-induced gene expression. a, Scatterplot showing 290 genes (left, red dots) that are significantly induced by eccDNA, but not linear DNA, in BMDCs. Thirty-four significantly induced cytokine genes are indicated (right, red dots). FC, fold change. **b**, Heatmap representation of the top 20 most strongly induced genes. **c**, GO terms enriched in the genes activated by eccDNA treatment in BMDCs. The number of genes in each term and the *P* values of enrichment are indicated. **d**, Scatterplot indicating that the transcriptomes of BMDCs treated with eccDNAs and BMDCs treated with synthetic circular DNA are highly similar. **e**, Heatmap representation of the eccDNA-responsive genes in control and eccDNA-treated BMDCs of the indicated genotype. **f**, Scatterplots comparing the transcriptome affected by eccDNA in BMDCs from WT, *Sting1*$^{-/-}$ and *Myd88*$^{-/-}$ mice. eccDNA-responsive genes in WT BMDCs are indicated by red dots (upregulated, $n = 365$) and blue dots (downregulated, $n = 79$). In **a**, **d** and **f**, the *x* and *y* axes are $\log_2$-transformed normalized read counts; *P* values were generated using DESeq2 (ref. [42]) and adjusted by IHW[43]. FC ≥ 5, adjusted $P < 0.001$ in **a** and **f**; FC ≥ 2, adjusted $P < 0.01$ in **d**.

apoptotic cells are linked to the oncogene amplification and tumour progression shown for large extrachromosomal DNAs[37].

Our demonstration that eccDNAs can dramatically induce type I interferon expression (Figs. 4 and 5), combined with previous observations that type I interferons possess adjuvant activity[40] and that DNA released from dying cells mediates aluminum adjuvant activity[18], prompt us to propose that eccDNAs possess high adjuvant activity. In addition, the existence of eccDNAs in plasma[10] suggests that eccDNA is highly mobile. Given that increased levels of cell-free DNA and serum IL-6 and TNF are good predictors of disease severity leading to a cytokine storm[41], we suspect that apoptosis-driven eccDNA generation and subsequent induction of cytokines might underlie the cytokine storm observed in diseases such as severe sepsis and coronavirus disease 2019 (COVID-19), as these diseases can involve massive cell death. If future studies confirm this notion, the eccDNA biogenesis and sensing pathway revealed in this study could serve as the basis for therapeutic interventions.

In summary, by providing answers to three key questions regarding the origin, biogenesis and biological function of eccDNAs, our study substantially advances understanding of eccDNAs. Further characterization of the molecular basis of eccDNA-mediated immune responses can provide new insight into innate immunity as well as vaccine design and immunotherapy.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-021-04009-w.

1. 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).
2. Gaubatz, J. W. Extrachromosomal circular DNAs and genomic sequence plasticity in eukaryotic cells. *Mutat. Res.* **237**, 271–292 (1990).
3. Hotta, Y. & Bassel, A. Molecular size and circularity of DNA in cells of mammals and higher plants. *Proc. Natl Acad. Sci. USA* **53**, 356–362 (1965).
4. Moller, H. D. et al. Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. *Nat. Commun.* **9**, 1069 (2018).
5. Dillon, L. W. et al. Production of extrachromosomal microDNAs is linked to mismatch repair pathways and transcriptional activity. *Cell Rep.* **11**, 1749–1759 (2015).
6. Hull, R. M. et al. Transcription-induced formation of extrachromosomal DNA during yeast ageing. *PLoS Biol.* **17**, e3000471 (2019).

7. Gresham, D. et al. Adaptation to diverse nitrogen-limited environments by deletion or extrachromosomal element formation of the *GAP1* locus. *Proc. Natl Acad. Sci. USA* **107**, 18551–18556 (2010).
8. Sinclair, D. A. & Guarente, L. Extrachromosomal rDNA circles—a cause of aging in yeast. *Cell* **91**, 1033–1042, (1997).
9. Shibata, Y. et al. Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. *Science* **336**, 82–86 (2012).
10. Zhu, J. et al. Molecular characterization of cell-free eccDNAs in human plasma. *Sci. Rep.* **7**, 10968 (2017).
11. Cloninger, C., Felton, M., Paul, B., Hirakawa, Y. & Metzenberg, S. Control of pH during plasmid preparation by alkaline lysis of *Escherichia coli*. *Anal. Biochem.* **378**, 224–225 (2008).
12. DeLap, R. J., Rush, M. G., Zouzias, D. & Khan, S. Isolation and preliminary characterization of the small circular DNA present in African green monkey kidney (BSC-1) cells. *Plasmid* **1**, 508–521 (1978).
13. Landry, J. J. et al. The genomic and transcriptomic landscape of a HeLa cell line. *G3* **3**, 1213–1224 (2013).
14. Tichy, E. D. Mechanisms maintaining genomic integrity in embryonic stem cells and induced pluripotent stem cells. *Exp. Biol. Med.* **236**, 987–996 (2011).
15. Rang, F. J., Kloosterman, W. P. & de Ridder, J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. *Genome Biol.* **19**, 90 (2018).
16. Joffroy, B., Uca, Y. O., Presern, D., Doye, J. P. K. & Schmidt, T. L. Rolling circle amplification shows a sinusoidal template length-dependent amplification bias. *Nucleic Acids Res.* **46**, 538–545 (2018).
17. Soh, Y. Q. et al. Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell* **159**, 800–813 (2014).
18. Marichal, T. et al. DNA released from dying host cells mediates aluminum adjuvant activity. *Nat. Med.* **17**, 996–1002 (2011).
19. Sakahira, H., Enari, M. & Nagata, S. Cleavage of CAD inhibitor in CAD activation and DNA degradation during apoptosis. *Nature* **391**, 96–99 (1998).
20. Li, L. Y., Luo, X. & Wang, X. Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature* **412**, 95–99 (2001).
21. Shiokawa, D., Kobayashi, T. & Tanuma, S. Involvement of DNase γ in apoptosis associated with myogenic differentiation of C2C12 cells. *J. Biol. Chem.* **277**, 31031–31037 (2002).
22. Arakawa, H. et al. Functional redundancy between DNA ligases I and III in DNA replication in vertebrate cells. *Nucleic Acids Res.* **40**, 2599–2610, (2012).
23. Lu, G. et al. Ligase I and ligase III mediate the DNA double-strand break ligation in alternative end-joining. *Proc. Natl Acad. Sci. USA* **113**, 1256–1260 (2016).
24. Simsek, D. et al. Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature* **471**, 245–248 (2011).
25. Li, Y., Berke, I. C. & Modis, Y. DNA binding to proteolytically activated TLR9 is sequence-independent and enhanced by DNA curvature. *EMBO J.* **31**, 919–931 (2012).
26. Lange, S. S. & Vasquez, K. M. HMGB1: the jack-of-all-trades protein is a master DNA repair mechanic. *Mol. Carcinog.* **48**, 571–580 (2009).
27. Wang, Z. et al. Regulation of innate immune responses by DAI (DLM-1/ZBP1) and other DNA-sensing molecules. *Proc. Natl Acad. Sci. USA* **105**, 5477–5482 (2008).
28. Green, D. R., Ferguson, T., Zitvogel, L. & Kroemer, G. Immunogenic and tolerogenic cell death. *Nat. Rev. Immunol.* **9**, 353–363 (2009).
29. Sundaresan, R., Parameshwaran, H. P., Yogesha, S. D., Keilbarth, M. W. & Rajan, R. RNA-independent DNA cleavage activities of Cas9 and Cas12a. *Cell Rep.* **21**, 3728–3739 (2017).
30. Putney, S. D., Benkovic, S. J. & Schimmel, P. R. A DNA fragment with an α-phosphorothioate nucleotide at one end is asymmetrically blocked from digestion by exonuclease III and can be replicated in vivo. *Proc. Natl Acad. Sci. USA* **78**, 7350–7354 (1981).
31. Sauer, J. D. et al. The *N*-ethyl-*N*-nitrosourea-induced Goldenticket mouse mutant reveals an essential function of Sting in the in vivo interferon response to *Listeria monocytogenes* and cyclic dinucleotides. *Infect. Immun.* **79**, 688–694 (2011).
32. Hou, B., Reizis, B. & DeFranco, A. L. Toll-like receptors activate innate and adaptive immunity by using dendritic cell-intrinsic and -extrinsic mechanisms. *Immunity* **29**, 272–282 (2008).
33. Oberhammer, F. et al. Apoptotic death in epithelial cells: cleavage of DNA to 300 and/or 50 kb fragments prior to or in the absence of internucleosomal fragmentation. *EMBO J.* **12**, 3679–3684 (1993).
34. Shiokawa, D. & Tanuma, S. Differential DNases are selectively used in neuronal apoptosis depending on the differentiation state. *Cell Death Differ.* **11**, 1112–1120 (2004).
35. Wang, H. et al. DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res.* **65**, 4020–4030 (2005).
36. Cox, D., Yuncken, C. & Spriggs, A. I. Minute chromatin bodies in malignant tumours of childhood. *Lancet* **1**, 55–58 (1965).
37. Turner, K. M. et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122–125 (2017).
38. Li, X. D. et al. Pivotal roles of cGAS–cGAMP signaling in antiviral defense and immune adjuvant effects. *Science* **341**, 1390–1394 (2013).
39. Andreeva, L. et al. cGAS senses long and HMGB/TFAM-bound U-turn DNA by forming protein–DNA ladders. *Nature* **549**, 394–398 (2017).
40. Tovey, M. G., Lallemand, C. & Thyphronitis, G. Adjuvant activity of type I interferons. *Biol. Chem.* **389**, 541–545 (2008).
41. Moore, J. B. & June, C. H. Cytokine release syndrome in severe COVID-19. *Science* **368**, 473–474 (2020).
42. 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).
43. Ignatiadis, N., Klaus, B., Zaugg, J. B. & Huber, W. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. *Nat. Methods* **13**, 577–580 (2016).

Methods

Cell culture and apoptosis induction

mESC/E14 cells were cultured on dishes coated with 0.1% gelatin in standard LIF/serum medium containing mouse LIF (1,000 U ml^{-1}), 15% foetal bovine serum (FBS), 0.1 mM non-essential amino acids, 0.055 mM β-mercaptoethanol (BME), 2 mM GlutaMAX, 1 mM sodium pyruvate and penicillin-streptomycin. HeLa S3 cells were grown in DMEM supplemented with 10% FBS and 100 U ml^{-1} penicillin-streptomycin. CH12F3 cells were cultured in RPMI 1640 supplemented with 10% heat-inactivated FBS,100 U ml^{-1} penicillin-streptomycin, 2 mM BME and 2 mM GlutaMAX. L929 cells were cultured in DMEM supplemented with 10% heat-inactivated FBS and 100 U ml^{-1} penicillin-streptomycin.

Apoptotic cell death of mESCs was induced with 0.5 μM etoposide (Selleck) or 2 μM staurosporine (Selleck) for 24 h or by irradiation at 3 mJ using UV-C light in a Stratagene Stratalinker 2400 without medium and continued culture for 16 h. Apoptotic cell death was induced in CH12F3 cells by treatment with 2 μM staurosporine for 16 h. Cell viability was analysed with a BD FACSCanto II instrument after staining with the Live/Dead Fixable Far Red Dead Cell Stain Kit. FSC-A and SSC-A gates were used to exclude debris, and FSC-A and FSC-H were used to gate on singlets and then gate on APC^+ cells as dead cells. Data were analysed with FlowJo v.10.8.0.

Knockout cell line generation

Dnase1l3- and *Endog*-knockout mESC cell lines were generated by CRISPR–Cas9 through transient transfection with px330-mCherry (Addgene, 98750). $mCherry^+$ cells were sorted by flow cytometry. Guide RNAs targeting sequences and PCR genotyping primers for *Dnase1l3* and *Endog* are listed in Supplementary Table 1. The *Lig1*$^{-/-}$ CH12F3 cell line was generated using CRISPR–Cas9 guide RNAs targeting introns 17 and 19 to delete exons 18 and 19, which encode the conserved ligase catalytic site. Deletion resulted in a premature stop codon. The *NucLig3*$^{-/-}$ CH12F3 cell line was generated by specifically deleting the sequences encoding the nucLig3 start codon and the two subsequent methionine residues (Met89-Met144) with CRISPR–Cas9 guide RNAs while keeping mtLig3 in frame and functional. *Lig4* is a single-exon gene, and two pairs of guide RNAs were used for two rounds of targeting to obtain homozygous deletion of the entire 2.7-kb exon. Guide RNA sequences are listed in Supplementary Table 1. Knockout cell lines were confirmed by immunoblotting.

eccDNA purification and visualization on agarose gels

To purify eccDNAs, cells were first dehydrated in >90% methanol before crude extrachromosomal DNA was extracted in an alkaline lysis buffer at pH 11.8. After neutralization and precipitation, crude extrachromosomal DNA was bound to a silica column (QIAGEN Plasmid Plus Midi Kit) in binding buffer (buffer BB from the QIAGEN Plasmid Plus Midi Kit). Bound DNA was eluted, digested with PacI (NEB) and Plasmid-Safe ATP-dependent DNase (Lucigen) for 4–16 h, then extracted with phenol:chloroform:isoamyl alcohol (PCI) solution (25:24:1) in a Phase Lock Gel tube (QuantaBio) to minimize DNA loss. After precipitation with carrier glycogen (Roche) and 1/10 volume of 3 M sodium acetate (pH 5.5), the precipitated crude eccDNAs were resuspended in solution A (One-Step Max Plasmid DNAout, TIANDZ). eccDNAs were selectively bound to magnetic silica beads in this solution and were then eluted with 0.1× elution buffer (1 mM Tris-HCl, pH 8.0), and the concentration was measured by Qubit dsDNA HS Assay kit (Thermo Fisher). A detailed eccDNA purification protocol is in preparation for publication.

For comparisons of eccDNA production among treatments or genotypes, both total DNA (Quick-DNA microPrep Plus Kit, Zymo) and eccDNA were purified from equal numbers of cells, eluted and loaded onto an agarose gel with equal volume. All DNA, except for in PCR genotyping, was resolved with vertical agarose gel electrophoresis and visualized by staining with SYBR Gold (Fisher Scientific; 1:10,000).

Synthetic small DNA circle preparation

Synthetic small DNA circles were prepared by the procedure of ligase-assisted minicircle accumulation (LAMA)[44]. Random DNA sequences were generated (https://faculty.ucr.edu/~mmaduro/random.htm) with 50% G+C content. The isomers of single-strand templates as well as their amplification primer sets were synthesized by IDT, and their sequences are listed in Supplementary Table 3. Products with a 5′-end phosphate were prepared with 2× Q5 DNA polymerase (NEB). Equal amounts of isomers were added to generate 100-μl HiFi Taq DNA ligase reaction mixtures and placed in thermocyclers using the following cycles: 95 °C for 3 min, 60 °C for 10 min and 37 °C for 5 min for at least 10 cycles. Circularized products were recovered with a PCR Purification Kit (Qiagen) and digested with Plasmid-Safe ATP-dependent DNase (Lucigen) before being recovered with a PCR Purification Kit.

SAFM imaging

SAFM imaging of DNA was performed in dry mode[45]. Briefly, 1/10 volume of 10× imaging buffer (100 mM $NiCl_2$ and 100 mM Tris-HCl, pH 8.0) was added to the sample to reach a final DNA concentration of 0.6–1.0 ng $μl^{-1}$, and 2–5 μl of the mixture was then spread on a freshly cleaved mica (Ted Pella) surface. After 2 min of incubation, the specimen was rinsed twice with 30 μl of 2 mM magnesium acetate, drying before and after the rinses with compressed air. Images were acquired by using tip C of an SNL-10 probe on a Veeco MultiMode atomic-force microscope with a Nanoscope V Controller in 'ScanAsyst in Air mode' and processed with Gwyddion 2.50.

Library preparation and eccDNA sequencing

The Nanopore sequencing library for eccDNA was prepared with the Ligation Sequencing Kit (Oxford Nanopore) according to the manufacturer's instructions after RCA and debranching. RCA was performed with phi29 DNA polymerase (NEB) with some modifications to ensure efficient amplification from 100 pg of template per reaction. Briefly, each 20-μl reaction mixture contained 2 μl of 10× phi29 DNA polymerase buffer (NEB), 2 μl of 25 mM dNTPs, 1 μl Exo-Resistant Random Primer (Thermo Fisher) and ≥100 pg of eccDNA, with ultra-pure water added to a maximum volume of 17.6 μl. Reactions were mixed and incubated at 95 °C for 5 min before ramping the temperature down to 30 °C at a 1% ramp rate. Then, 1 μl of phi29 DNA polymerase, 0.6 μl of inorganic pyrophosphatase (yeast, NEB), 0.4 μl of 0.1 M DTT (NEB) and 0.4 μl of 20 mg ml^{-1} BSA (NEB) were added. The reaction mixture was incubated at 30 °C for 10–16 h. Because high branch structure in RCA products can block nanopores and abolish sequencing, RCA products for eccDNAs were further debranched with T7 endonuclease I (NEB) before being used for sequencing library construction with the Ligation Sequencing Kit (Oxford Nanopore, SQK-LSK109). The library was sequenced in a flow cell (R9.4.1, FL-MIN106D) on a MinION instrument according to the manufacturer's instructions.

Illumina sequencing libraries for eccDNA were prepared by Tn*5*-transposon-based tagmentation with the Nextera XT DNA Sample Preparation Kit according to the manufacturer's instructions. Briefly, after validating the purity of the eccDNAs with SAFM imaging, 0.5 ng of pure eccDNAs was directly tagmented with Tn*5* transposase, followed by 12–14 cycles of PCR amplification with Illumina sequencing adaptors. Barcoded libraries were pooled and sequenced with the Illumina 2500 platform in 150-bp paired-end mode.

eccDNA sequencing data analyses

Nanopore base calling and reads mapping. The fast5 files generated by Nanopore MinION were fed to Guppy (version 3.5.2) for base calling. The parameters used for Guppy were as follows: -flowcell FLO-MIN106 --kit SQK-LSK109 --qscore_filtering --calib_detect --trim_barcodes --trim_strategy dna --disable_pings --device auto --num_callers 16. The generated reads in fastq format were further processed by porechop (version 0.2.4) to

remove adaptor sequences for each read with the following parameters: --extra_end_trim 0 --discard_middle. To reduce artefacts due to misalignment during read mapping, we compiled a customized reference mouse genome (mm10combine) based on mm10 reference sequences. Briefly, we downloaded all the nucleotide sequences from the NCBI NT database (30 October 2018), and R/Bioconductor genbankr (version 1.10.0) was then used to distinguish mouse contigs from the contigs of other species. On the basis of each contig's description and manual inspection, we removed all gene-related contigs, retaining only 15,984. The selected fasta sequences were extracted using the command 'blastdbcmd -db nt_db -entry_batch selected_ids.txt -out selected_ids.fa -outfmt %f'. The fasta sequences were mapped to the mm10 genome. Finally, we selected contigs that could not be mapped, contigs for which less than 50% of the sequence mapped uniquely and contigs that were uniquely mapped but with a sequencing quality score of <10. In total, 103 contig sequences were added to mm10 to build the mm10combine reference genome. Cleaned reads were then aligned to mm10combine using minimap2 (version 2.17)[46] with the following parameters: -x map-ont -c --secondary=no -t 16. The alignments for each read were stored in PAF format.

Consensus eccDNA generation. To obtain the consensus boundary and sequence of each eccDNA from the mapped RCA long reads, we developed a tool (https://github.com/YiZhang-lab/eccDNA_RCA_nanopore) that uses the alignments in PAF files as input and outputs eccDNA fragment composition (chromosome and genomic start and end positions of each fragment), successive fragment coverage (number of passes) and the consensus sequence derived from each RCA long read. The subreads of each RCA long read could be mapped to one genomic location or multiple locations. Subreads with mapping quality lower than 30 were discarded. The tool performed bootstrapping of successive subreads in each RCA long read to check whether the order of the mapped genomic locations for each subread was concordant with the order in the RCA long read. Because of the inaccuracy and gap-prone nature of Nanopore reads, we allowed a maximum of 20 bp offset of the mapped genomic positions (start and end positions) for two subreads to be considered as mapping to the same location. Reads with discordant subread order, location or strand were discarded. The exact boundaries of eccDNA fragments were determined by voting from the subreads' start and end positions. Boundary positions were further refined by threading the subreads to ensure no gaps or overlaps between any successive subreads. The number of passes for each eccDNA fragment was calculated as the number of concordant subreads mapped to that location. Only eccDNAs with at least two passes were kept for downstream analysis. Each eccDNA sequence was derived from the reference genome sequence to which it mapped, with sequence variants incorporated. Sequence variants were called from subreads mapped to the corresponding location, requiring a minimum depth of 4 and a minimum allele frequency of 0.75.

Genomic distribution of eccDNAs. The eccDNA fragments were piled up across the genome. To remove PCR duplicates, eccDNA fragments on the same chromosome with the same start and end positions were treated as duplicates and only one was retained. The coverage of unique eccDNA fragments at each base of the genome was obtained using bedtools (version 2.29.2)[47] and stored in bigwig files. The distribution of eccDNA fragments across each chromosome was plotted using karyoploteR (version 1.14.1)[48] with the bigwig file as input.

Mapping of Illumina sequencing reads. Raw Illumina sequence reads were first processed by Trimmomatic (version 0.39)[49] to remove sequencing adaptors and low-quality reads, using the following parameters: ILLUMINACLIP:adapters/NexteraPE-PE.fa:2:30:10:1:true LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:75 TOPHRED33. BWA[50] MEM (version 0.7.17) was used with default parameters. Then, the reads were mapped to our customized mm10combine reference genome. Duplicated reads were removed by Picard (version 2.23.4). Reads with a mapping quality score of at least 60 were considered as uniquely mapped and used for downstream analysis. The genomic coverage was calculated using bamCoverage from deeptools (version 3.5.0)[51] with binSize 1.

Western blot analysis

Equal numbers of cells were lysed in NuPAGE LDS sample buffer (Thermo Fisher), and protein extracts were resolved by SDS–PAGE and transferred to PVDF membrane. Antibodies against Lig1 (Proteintech; 1:1,000), Lig3 (BD Biosciences; 1:1,000), Lig4 (a gift from D. Schatz (Yale University); 1:1,000), Myd88 (ProSci; 1:1,000), Sting (Proteintech; 1:1,000) and Gapdh (Thermo Fisher; 1:20,000) were used. Uncropped and unprocessed scans of blots are provided in Supplementary Fig. 1.

BMDC and BMDM preparation and stimulation

Male mice, including WT C57BL/6, *Sting1*$^{-/-}$ (Tmem173gt, stock no. 017537) and *Myd88*$^{-/-}$ (Myd88tm1.1Defr, stock no. 009088) mice, were purchased from Jackson Labs and were housed on a 12-h light/dark cycle at 23 °C with 45–55% humidity. After at least 7 days of habituation, mice between 8 and 12 weeks old were used to collect bone marrow cells for BMDC and BMDM differentiation. BMDCs were differentiated in RPMI 1640 supplemented with 10% heat-inactivated FBS (Sigma), 10 mM HEPES, 1 mM sodium pyruvate, 100 U ml^{-1} penicillin-streptomycin, 2 mM GlutaMAX and 20 ng ml^{-1} mouse granulocyte and macrophage colony-stimulating factor (GM-CSF; Peprotech). BMDMs were differentiated in DMEM supplemented with 10% heat-inactivated FBS, 100 U ml^{-1} penicillin-streptomycin and 20% L929 conditioned medium. Half of the medium was replaced at day 3 and day 6. BMDCs and BMDMs were confirmed to be CD11c^{+}MHC II^{+} and F4/80^{+}CD11b^{+}, respectively, after excluding debris with FSC-A and SSC-A gates and subsequent FSC-A and FSC-H gates on singlets. Data were analysed with FlowJo v.10.8.0. For cell stimulation, cells at days 7–9 were seeded in 96-well plates at 3.5×10^4 cells per well. DNA was transfected into cells with FuGENE HD (Promega) in Opti-MEM (Gibco) according to the manufacturer's instructions after measuring its concentration with a Qubit dsDNA HS Assay Kit (Thermo Fisher). All transfections were performed for 12 h except as indicated, media were collected for ELISA and cells were lysed with TRIzol (Thermo Fisher) for RNA isolation.

Transfection efficiency assays

To determine the transfection efficiency for linear and circular DNA, a set of primers (for sequences, see Supplementary Table 2) with five phosphorothioate bonds at their 5′ end were used to prepare end-protected linear DNA by PCR. To balance the effects of the phosphorothioate bonds, the circular form was prepared with the same number of phosphorothioate bonds as the linear one. Then, DNA concentration was determined by Qubit dsDNA HS Assay Kit (Thermo Fisher), and DNA was transfected into BMDCs as described above with FuGENE HD (Promega). After transfection, cells were rinsed three times with PBS and lysed in 100 µl of lysis buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5% Tween-20, 3 U ml^{-1} thermolabile proteinase K (NEB, P8111S)) and incubated at 37 °C for 2 h followed by incubation for 15 min at 55 °C to inactivate proteinase K. Four microlitres of cell lysate was used for qPCR with a set of primers targeting both linear and circular DNA to determine the amount of transfected DNA.

Incubation of BMDCs with supernatant from apoptotic cells

WT and *Dnase1l3*$^{-/-}$ cells in 10-cm dishes that were 80–100% confluent were washed three times with PBS and irradiated with 3 mJ of UV-C light in a Stratagene Stratalinker 2400, 10 ml Opti-MEM (Gibco) was added and cells were cultured for another 48 h. Medium was centrifuged at 650*g* for 5 min, and the supernatant was filtered through a 0.45-µm filter. Four hundred microlitres of supernatant was left untreated or treated with enzyme (PacI, Plasmid-Safe ATP-dependent DNase, RNase A/T1 or benzonase, as indicated) in a 500-µl reaction volume at 37 °C for 2 h and then dialyzed (molecular weight cut-off (MWCO), 10 kDa)

with fresh Opti-MEM at 4 °C overnight to deplete ATP, which is required for the activity of Plasmid-Safe ATP-dependent DNase. Then, 100 µl per well of dialyzed supernatant was added to BMDCs in 96-well plates, and an equal volume of fresh Opti-MEM was added in parallel to separate wells as a mock control; 12 h later, cells were collected for RT–qPCR analysis, and data are presented as relative mRNA levels with respect to those of mock controls after normalizing to *Gapdh*.

RNA isolation, RT–qPCR, RNA-seq and ELISA analyses

Cellular RNA was isolated with a Zymo Direct-zol RNA Miniprep kit. cDNA was synthesized with SuperScript III, and qPCR was performed with Fast SYBR Green Master Mix (Thermo Fisher). The primer sequences for qPCR of each gene are listed in Supplementary Table 2. Gene induction levels are presented as the relative fold change with respect to mock treatment after normalizing to *Gapdh*. Bulk RNA-seq libraries were prepared by following the manufacturer's instructions for the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, E7420S). For ELISA analysis, ELISA kits for IFNβ, IL-6 and TNF were obtained from BioLegend, and IFNα ELISA kits were obtained from PBL Assay Science. Assays were performed according to the manufacturers' instructions. Appropriate volumes of culture medium were used to ensure that the readouts were within the range of the standard curve.

RNA-seq data analysis

For RNA-seq data, adaptors and low-quality reads were trimmed using Trimmomatic (version 0.39)[49] with the following parameters: ILLUMINACLIP:adapters/TruSeq3-PE.fa:2:30:10:1:true LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50 TOPHRED33. Cleaned paired-end reads were aligned to the mm10 reference genome with GENCODE[52] mouse gene set M24, using STAR (version 2.7.6a)[53] with the following parameters: --outSAMunmapped Within --outFilterType BySJout --outSAMattributes NH HI AS NM MD --outFilterMultimapNmax 20 --outFilterMismatchNmax 999 --outFilterMismatchNoverReadLmax 0.04 --alignIntronMin 20 --alignIntronMax 1000000 --alignMatesGapMax 1000000 --alignSJoverhangMin 8 --alignSJDBoverhangMin 1 --sjdbScore 1 --outSAMtype BAM SortedByCoordinate --quantMode TranscriptomeSAM. RSEM (version 1.3.3)[54] was used to quantify gene expression levels using the reads aligned to the transcriptome in the bam file as input, with the following parameters: --alignments --estimate-rspd --calc-ci --no-bam-output --seed 12345 --ci-memory 30000 --paired-end --strandedness reverse. Differentially expressed genes were identified using the DESeq2 package[42].

eccDNA linearization

eccDNA linearization was performed by sequential treatment of eccDNAs with the nickase fnCpf1 (Applied Biological Materials)[37] and single-strand DNA-specific nuclease. Fifty nanograms of eccDNAs was nicked in a 20-µl reaction that contained 1/8 volume of 8× fnCpf1 linearization buffer (160 mM HEPES pH 7.5, 1.2 M KCl, 4 mM DTT, 0.8 M EDTA, 80 mM $MnCl_2$) and 1 µl fnCpf1. After incubating at 37 °C for 1 h, the treated eccDNAs were extracted with PCI solution (25:24:1) in a Phase Lock Gel tube (QuantaBio) and precipitated at −80 °C with carrier glycogen (Roche) and 1/10 volume of 3 M sodium acetate (pH 5.5). Nicked eccDNAs were linearized in 10-µl reactions that contained 2 µl of 5× buffer (0.25 M sodium acetate pH 5.2, 1.4 M NaCl, 25 mM $ZnSO_4$) and 1 µl S1 nuclease (Thermo Fisher) and were incubated at 37 °C for 5 min. Reactions were stopped by adding 40 µl of 10 mM Tris-HCl (pH 8.0), and linear eccDNAs were immediately recovered with 75 µl of SPRIselect beads (Beckman Coulter). Successful linearization of eccDNAs was confirmed by efficient digestion with Plasmid-Safe ATP-dependent DNase (Lucigen).

Statistics

Ordinary one-way ANOVA and two-tailed unpaired *t*-tests were performed with GraphPad Prism 9.

Ethics statement

All procedures on animals involved in this study were conducted according to protocols approved by the Harvard Medical School IACUC.

Reporting summary

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

Data availability

The eccDNA sequencing data and RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) with accession number GSE165919.

Code availability

The code used to analyse eccDNA sequencing data is available at GitHub (https://github.com/YiZhang-lab/eccDNA_RCA_nanopore).

44. Du, Q., Kotlyar, A. & Vologodskii, A. Kinking the double helix by bending deformation. *Nucleic Acids Res.* **36**, 1120–1128 (2008).
45. Liu, D. et al. Branched kissing loops for the construction of diverse RNA homooligomeric nanostructures. *Nat. Chem.* **12**, 249–259 (2020).
46. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
47. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
48. Gel, B. & Serra, E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088–3090 (2017).
49. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
50. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
51. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
52. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).
53. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
54. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

Acknowledgements We thank D. Schatz (Yale University) for the antibody to mouse Lig4, R. Rajan (University of Oklahoma) for the initial gift of FnoCas protein, X. Zhang for help in BMDC preparation, P. Yin for allowing us to use the AFM instrument and J. Lieberman for critical reading of the manuscript. H.C. was supported by a Leukemia and Lymphoma Society Fellow Award. D.L. is a Merck Fellow of the Life Sciences Research Foundation. This work was supported by the Howard Hughes Medical Institute. F.W.A. and Y.Z. are Investigators of the Howard Hughes Medical Institute.

Author contributions Y.Z. conceived the project; Y.W. and Y.Z. designed the experiments; Y.W. performed the majority of the experiments; H.C. generated the ligase-knockout cell lines; M.W. developed the method to call eccDNA from Nanopore data and performed all bioinformatics analyses; M.N.D. performed pilot analysis of eccDNA sequencing data; D.L. helped with the SAFM imaging; F.W.A. supervised the work by H.C.; Y.W. and Y.Z. wrote the manuscript.

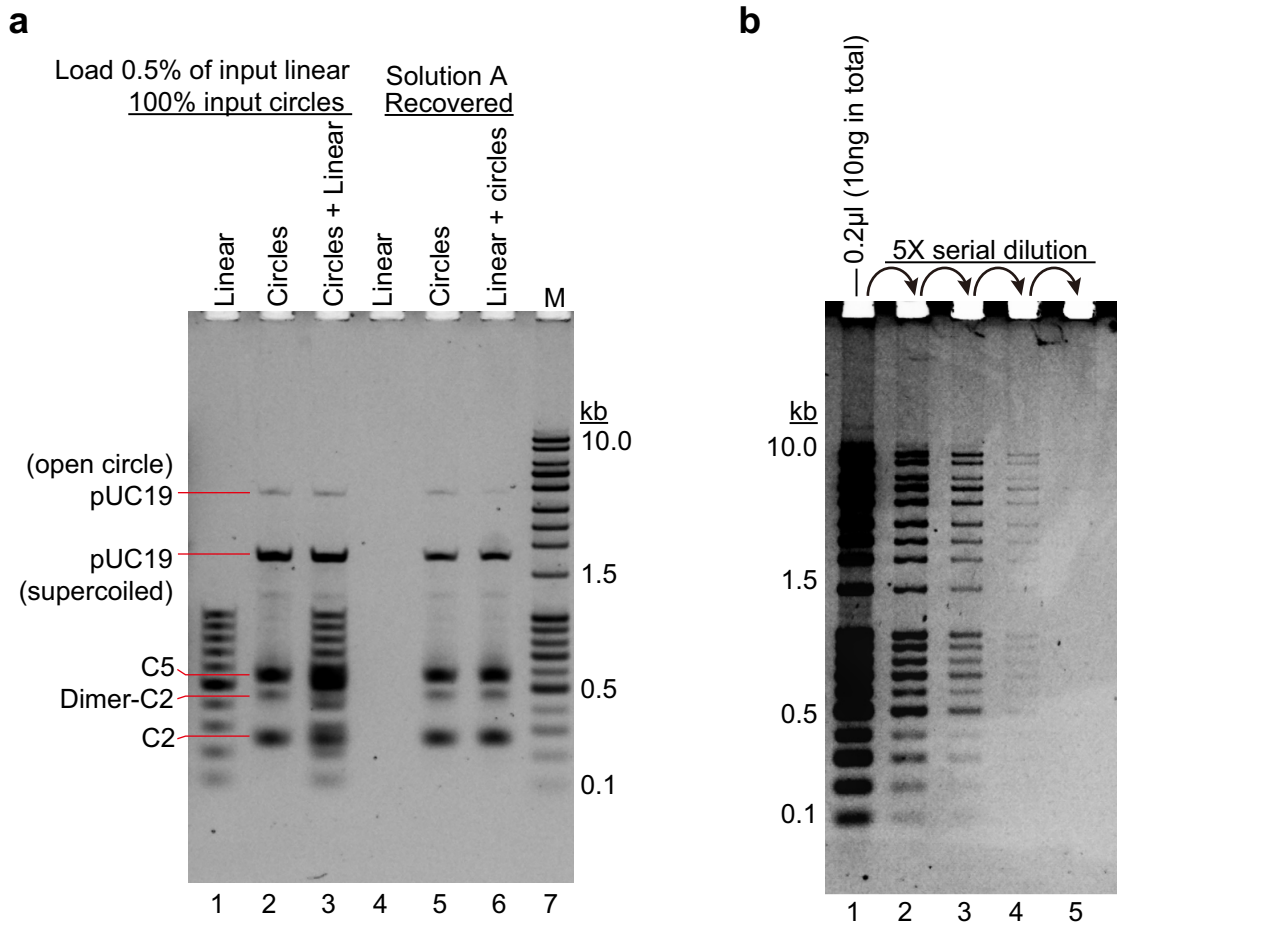
Competing interests A patent application (with the patent agency Wolf, Greenfield & Sacks) covering the eccDNA purification method and the application of eccDNA as an immunostimulatant has been filed by the Boston Children's Medical Center Corporation.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-021-04009-w.

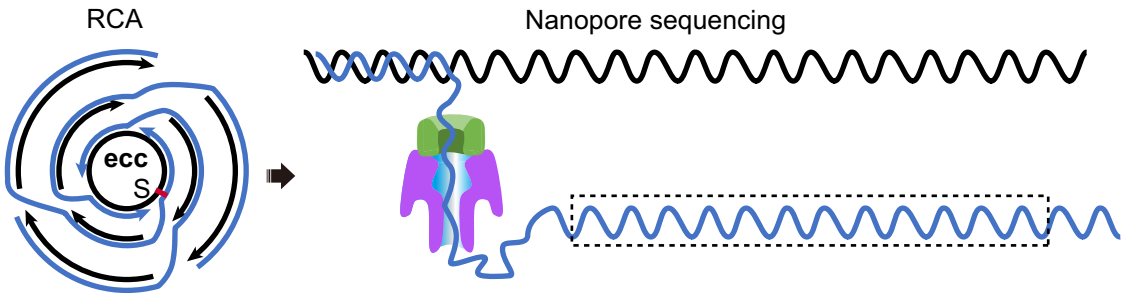
Correspondence and requests for materials should be addressed to Yi Zhang.

Peer review information *Nature* thanks Birgitte Regenberg and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | Selective binding of eccDNAs to beads and high sensitivity of DNA detection by SYBR Gold staining. a, Representative gel image showing selective binding of circular DNA to magnetic silica beads in solution A. The assay was tested with 200× (mass) linear DNA (lane 4 and 6) and 1X (mass) DNA circles (lane 5 and 6). Lanes 1-3 indicate the input DNAs. Lane 1, 0.5% of input linear DNA for lane 4 was loaded; lane 2, equal amount of input circular DNAs for lane 5 were loaded; lane 3, 0.5% of input linear DNA and equal amount of input circular DNAs for lane 6 were loaded. Lanes 4-6 were DNAs recovered from Solution A. lane 4, 200× (mass) linear DNA alone undergone purification by solution A; lane 5, 1× (mass) DNA circles undergone purification by solution A; lane 6 mixture of 200× (mass) linear DNA and 1× (mass) DNA circles undergone purification by solution A. Note, only circular DNAs were recovered. The recovery rate was very high, particularly for the smaller circular DNAs, C2: a 200 bp circle, C5: a 525 bp circle. **b**, High sensitivity of DNA detection using vertical agarose gel electrophoresis and SYBR Gold staining. 0.2 µl commercial DNA ladder (total of 10 ng) was undergone 5× series dilutions and fractionated by 1% vertical agarose gel electrophoresis, stained by SYBR Gold, DNA in lane-4 were 125× diluted of that in lane-1, and could still be visualized. The detection limit is estimated as 5-10 pg/band. Results in **a**–**b** are representative of three independent experiments.

a

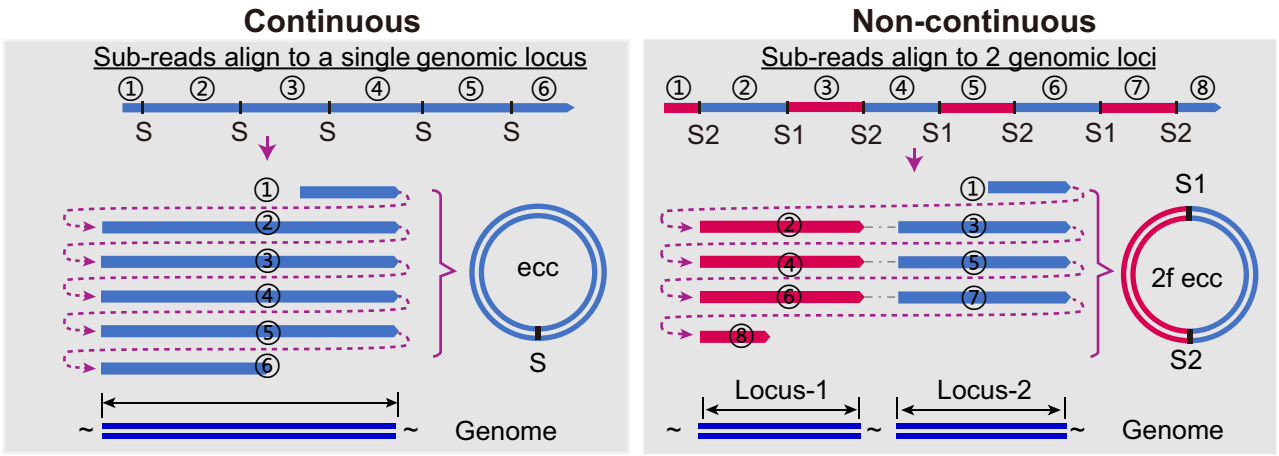


b

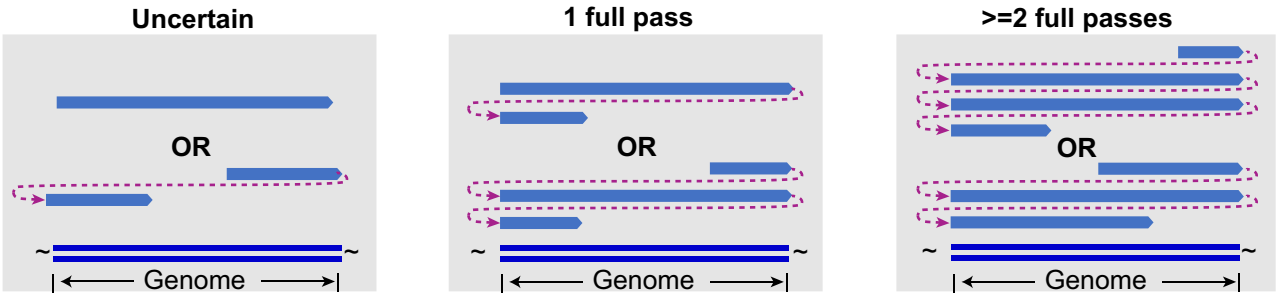
Summary of Nanopore sequencing

								Unique eccDNA (>= 2 full passes)				
	Total reads	Mean size (bp)	N50 Length (bp)	Longest read	Mapped reads	>= 1 coverage	>= 2 coverage	Unique eccDNA	Mean length (bp)	N50 length (bp)	Longest eccDNA (bp)	Shortest eccDNA (bp)
Number	4,046,970	3,746	4,280	47,905	4,020,393	2,443,525	1,927,361	1,646,531	879	1,089	9,990	70
Percentage	100%				99.34%	60.38%	47.62%	40.69%				

c



d



Extended Data Fig. 2 | See next page for caption.

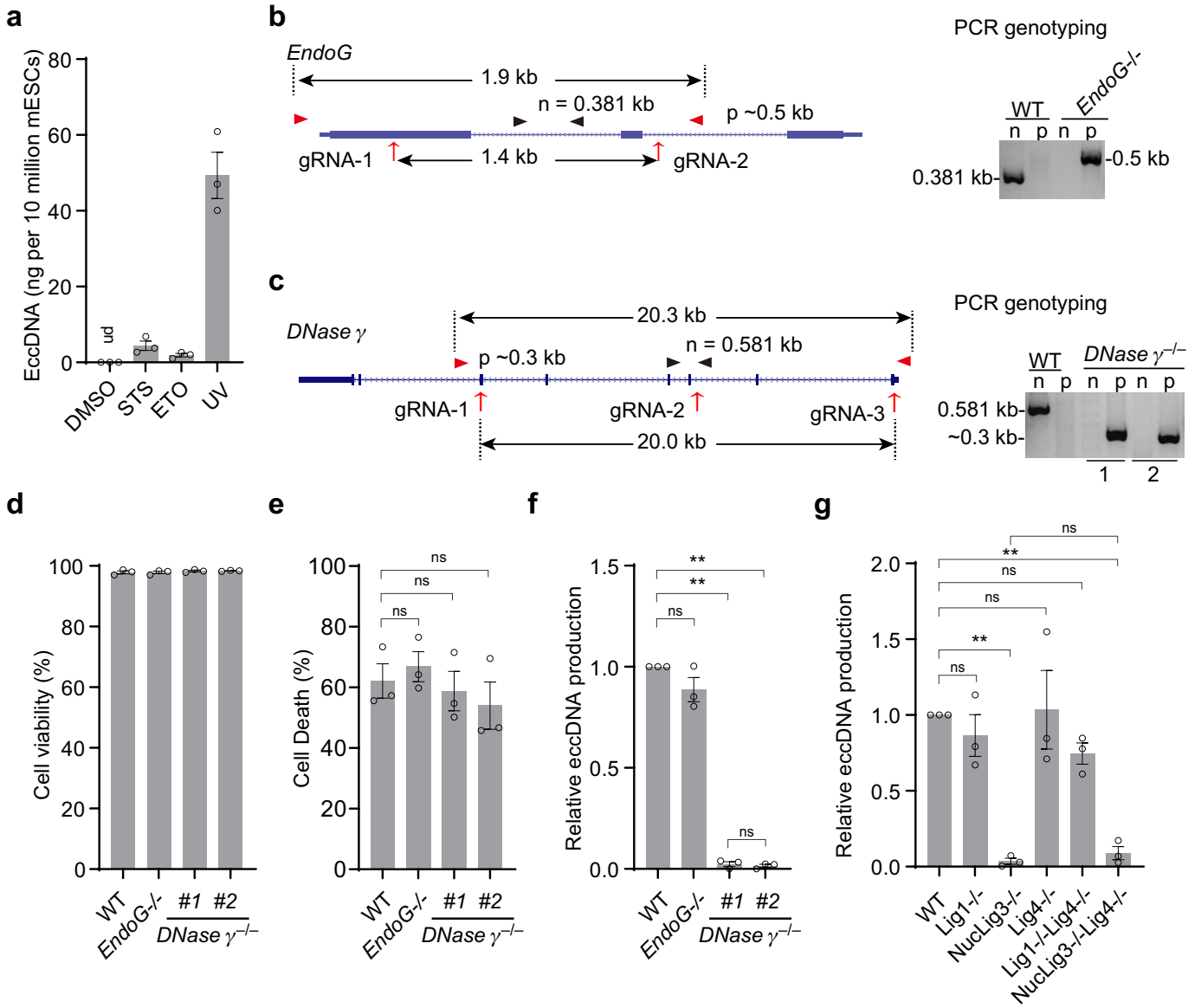
Extended Data Fig. 2 | Summary of the nanopore sequencing data. **a**, Diagram of nanopore long read sequencing of eccDNA. Tandem copies of eccDNAs were self-concatenated to long molecule by rolling cycle amplification (RCA), and directly read through by Oxford Nanopore. Each copy of eccDNA molecule in a single long molecule was sequenced multiple times. "S": split site. **b**, Summary of mESC eccDNA long reads from an Oxford Nanopore MinION flow cell. **c**, Diagram showing how the eccDNA full-length sequence is called and categorized. Nanopore reads from (**a**), the dashed box, were aligned to a single locus (Continuous) or to multiple loci (Non-continuous) in genome. Continuous: an example of six sub-regions of a long nanopore read repeatedly aligned to a single locus, where full length sequence of eccDNA was presented with only one split site (S); Non-continuous: an example of eight sub-regions of a single long read sequentially aligned to two separate loci (Locus-1 and Locus-2), representing an eccDNA ligated by two genomic fragments (2f ecc) with two split sites (S1 and S2). **d**, Criteria for eccDNA calling. EccDNAs were called based on their number of full passes aligned to the genome. Long reads with less than two full passes were discarded (left and middle panel). Left, Nanopore read that hits genome only once either fails to designate the genomic start and end site of eccDNA (up panel), or miss the middle region (bottom panel) that may or may not include in the original eccDNA molecule (Uncertain); Middle, because of potential sequencing error of Oxford Nanopore, reads that hit genome more than once but less than twice (1 full pass) were also discarded due to the lack of confirmation in eccDNA calling, particularly on designating the start and end site of eccDNA. Right, eccDNA molecules were called from long reads when covered at least twice (>=2 full pass) on their aligned loci.

a

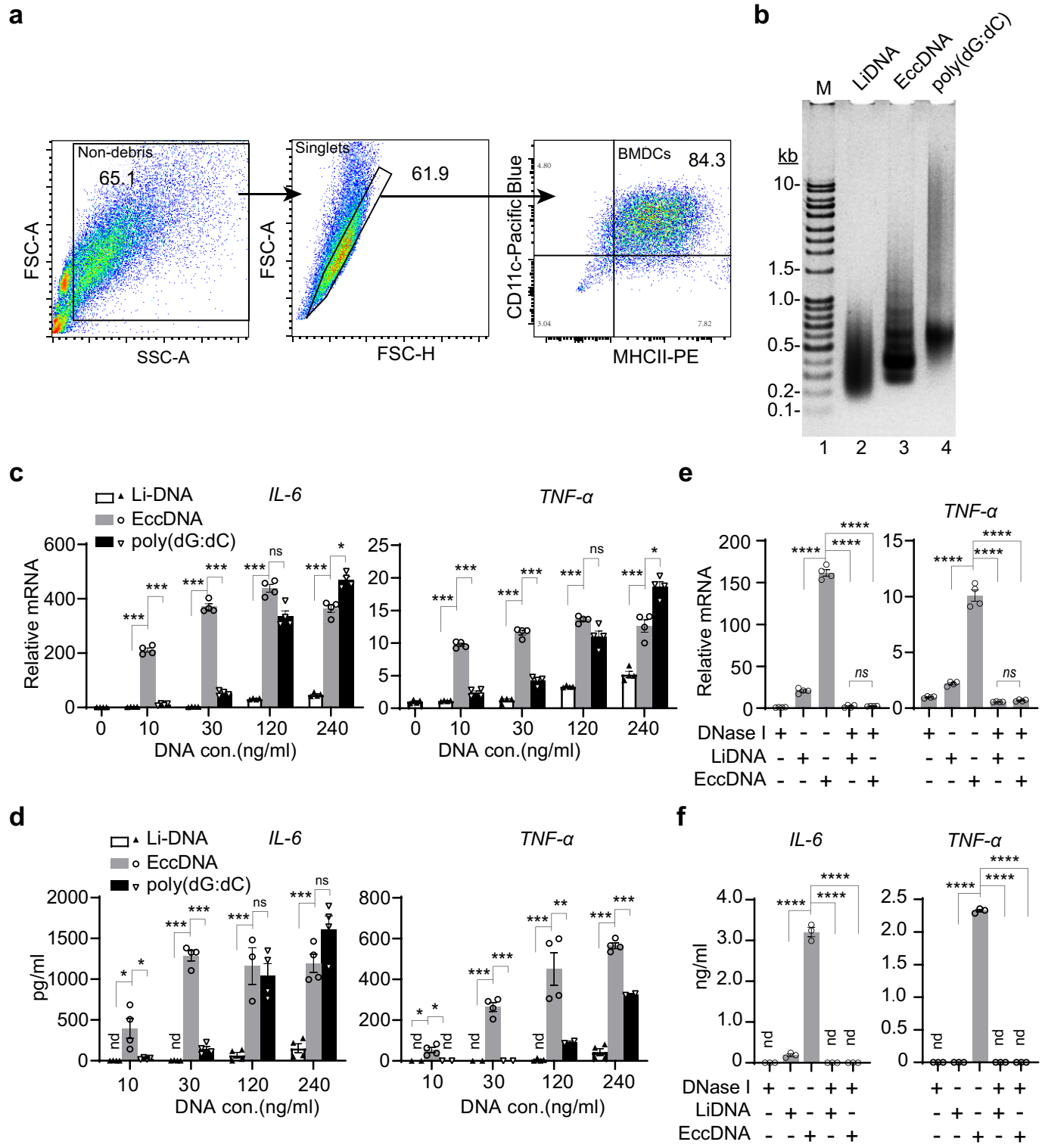
	Raw read pairs	Clean read pairs	Mapped reads			
				Removed Duplicates	Unique reads (MQ≥60)	Split reads
Number	118,137,295	104,182,680	104,162,136	77,407,461	65,514,518	16,866,752
Percentage	-	100.00%	99.98%	74.30%	62.88%	16.19%

b

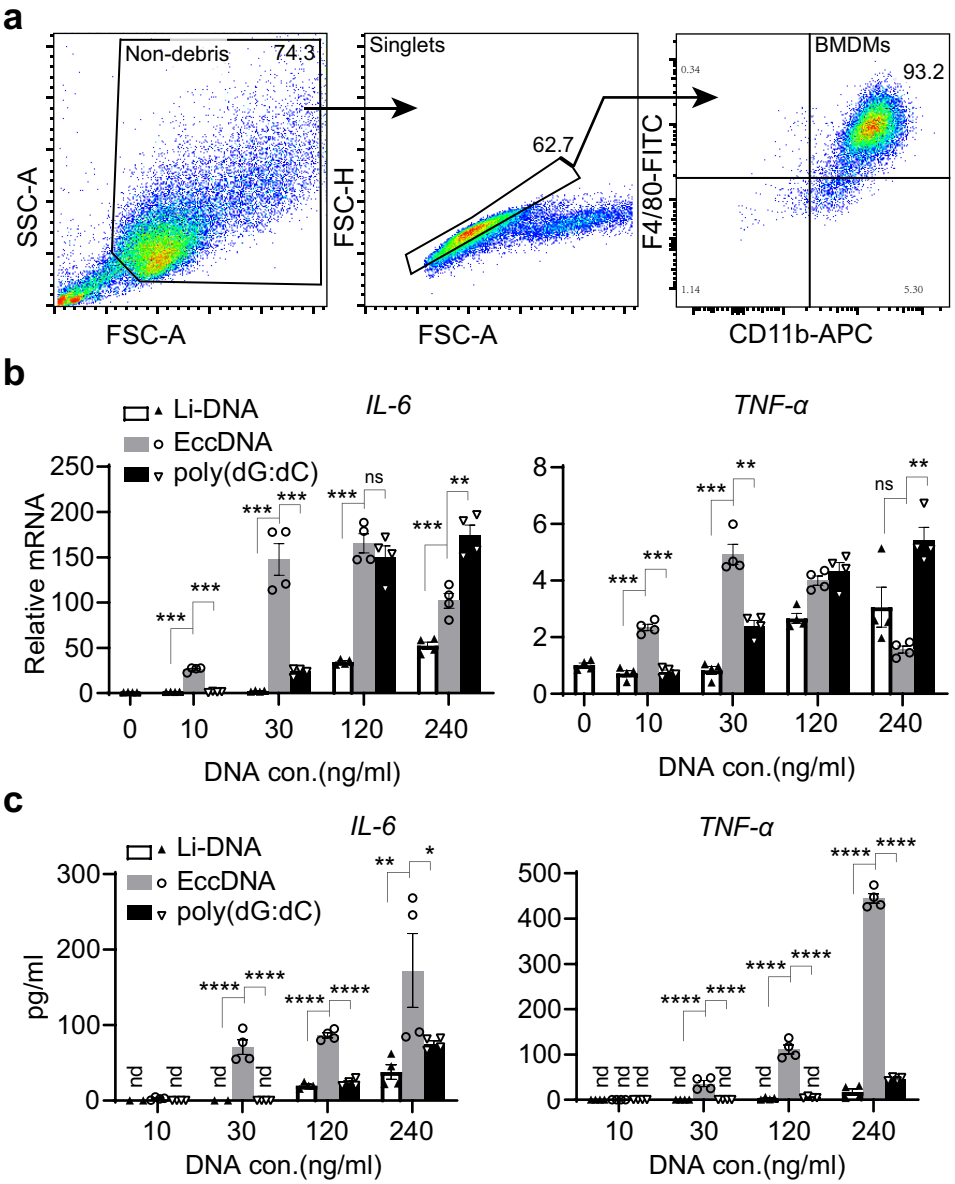
Extended Data Fig. 3 | Summary of Illumina short read sequencing and Chromosomal distribution of eccDNA fragments. a, Summary of eccDNA short read sequencing without RCA. Purified eccDNA was directly tagmented with Tn5 (Illumina Nextera), and sequenced with Illumina Hiseq 2500 in PE150 mode. **b**, Chromosomal distribution of eccDNA short reads.



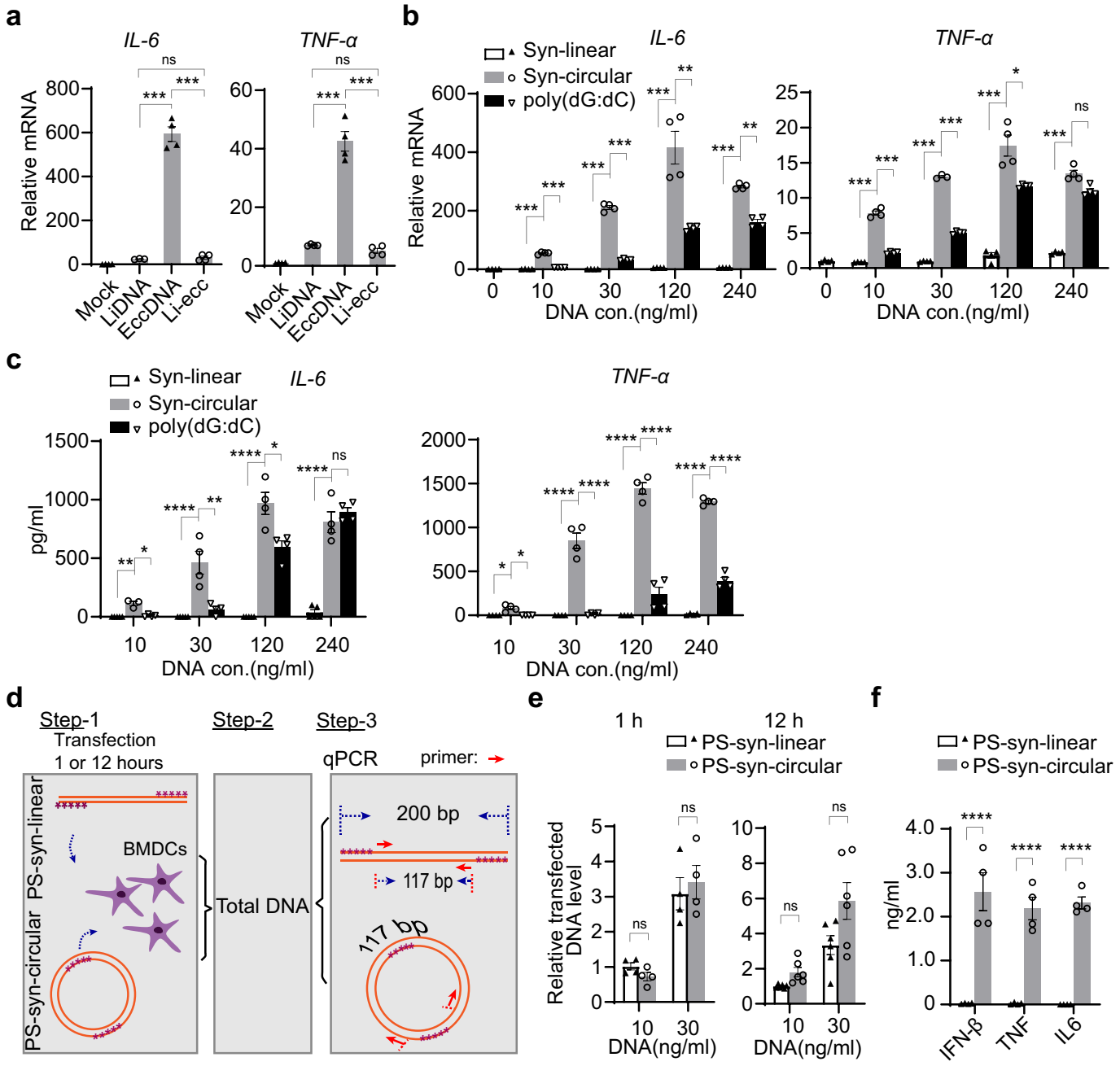
Extended Data Fig. 4 | Effects of *Endonuclease G*, *DNase γ*, and DNA ligases knockout on apoptosis and eccDNA generation. a, Quantification of ecc DNA from cells of the indicated treatment. EccDNA production was presented as nanogram per 10 million cells; ud, under detection limit. Bars indicate mean ± s.e.m. of three independent experiments. **b**, Diagram illustration and PCR confirmation of the *EndoG* knockout mESC line. Gene structure of *EndoG*, sgRNAs (red arrows), two sets of screen primers for internal sites (black arrowhead, "n" =negative knockout cell line) and external sites (red arrowhead, "p" =positive knockout cell line) were shown. **c**, Diagram illustration and PCR confirmation of the *DNase γ* knockout mESC lines. Gene structure of *DNase γ*, sgRNAs (red arrows), two sets of screen primers for internal sites (black arrowhead, "n" =negative knockout cell line) and external sites (red arrowhead, "p" =positive knockout cell line) were shown. **d**, Cell viability is not affected by *EndoG* or *DNase γ* KO. Cell viability was evaluated by flow cytometry, after staining with Far Red Live/Dead Cell Stain Kit, FSC-A/SSC-A and FSC-A/FSC-H were sequentially used to exclude debris and gate on singlets, respectively, then gate on APC+ as dead cells. Bars indicate mean ± s.e.m. of dead cell ratio from three independent experiments. **e**, Deficiency of *DNase γ* or *EndoG* in mESC do not significantly alter UV-induced cell death. Cell death was measured by flow cytometry, after staining with Far Red Live/Dead Cell Stain Kit, FSC-A/SSC-A and FSC-A/FSC-H were sequentially used to exclude debris and gate on singlets, respectively, then gate on APC+ as dead cells. Bar indicate mean ± s.e.m. of dead cell ratio of 3 independent experiments. **f**, Quantification of eccDNAs presented in Fig. 3d. EccDNAs below the mtDNA band were quantified by densitometry (by ImageJ 1.53e) and presented as relative levels to that in WT cells of parallel eccDNA purifications. Bars indicate mean ± s.e.m. of three independent experiments. **g**, Quantification of eccDNA below the mtDNA band by densitometry (by ImageJ 1.53e). Data are presented as relative levels to that purified from WT cells of parallel eccDNA purifications (Fig. 3g), bars indicate mean ± s.e.m. of three independent experiments. Ordinary one-way ANOVA with Tukey's multiple comparison test was used for multiple comparisons in **e**, **f**, **g**. **, $p<0.01$; ns, no significant.



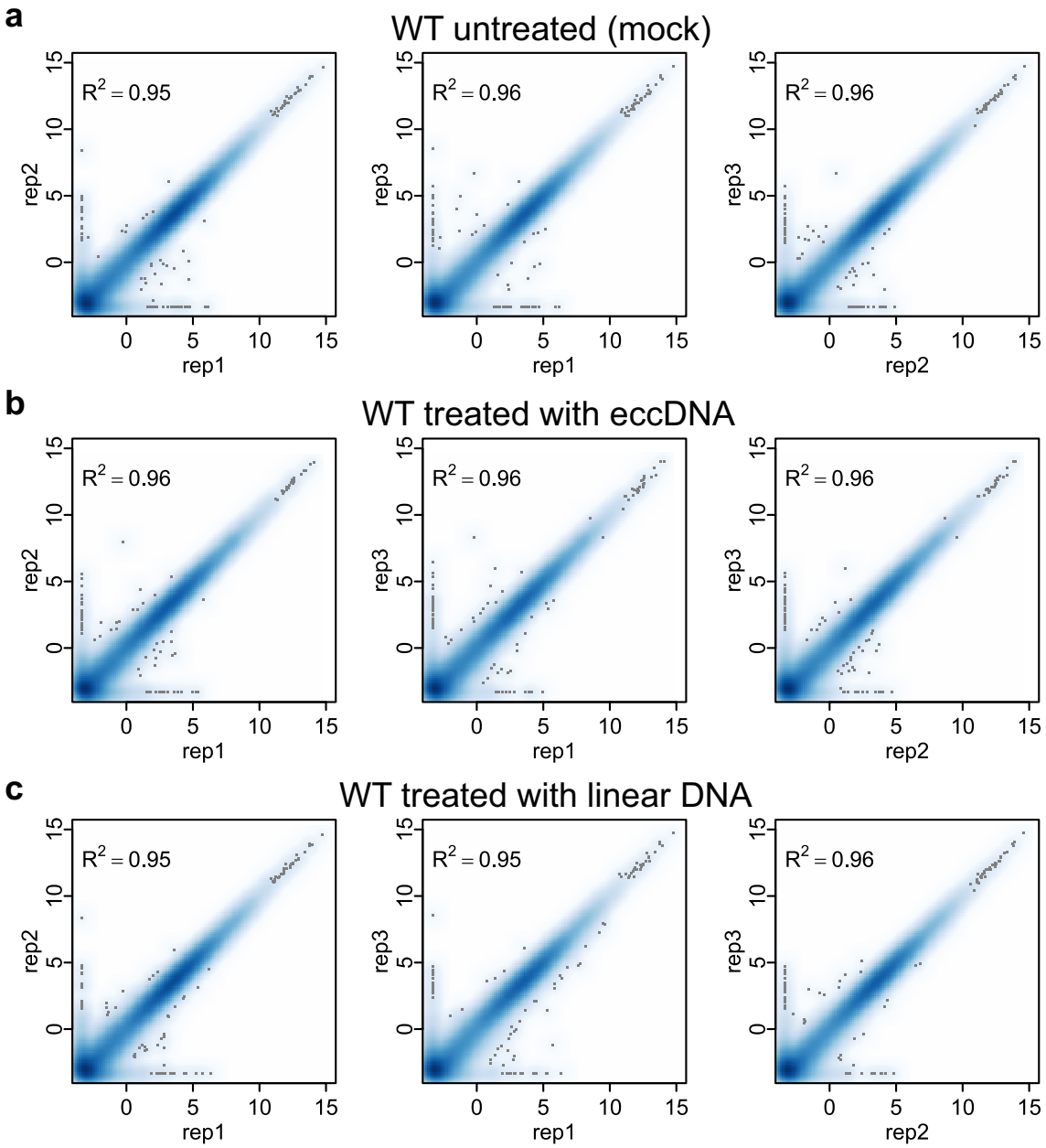
Extended Data Fig. 5 | EccDNAs are potent immunostimulants for BMDCs. **a**, Confirmation of BMDC identity by flow cytometry. After differentiating bone marrow cells with 20 ng/ml GM-CSF for 7 days, cell debris (left) were first excluded and singlets (middle) were then phenotyped on the basis of their CD11c and major histocompatibility complex II (MHCII) expression (right) to define BMDC, the numbers indicate the percentage of gated cells. Data were further analyzed with FlowJo. V10.8.0 and shown as representative of 2 independent experiments. **b**, Various DNAs were resolved by agarose gel electrophoresis. Li-DNA: sheared linear genomic DNA; poly(dG:dC), pol y(deoxyguanylic-deoxycytidylic). Representative gel of three independent experiments was shown. **c**, **d**, Bar graphs showing the relative mRNA (**c**) and protein (ELISA) (**d**) levels of *IL-6* and *TNF-α*. Data is shown as mean ± s.e.m. of replicates (n=4 per group) of a representative experiment in three independent experiments. **e**, **f**, DNase I pretreatment abolishes *IL-6* and *TNF-α* induction by eccDNA. Equal amount of DNA (120 ng/ml) used to prepare transfection mixtures were pre-treated with or without DNase I as indicated, then transfected to BMDCs. 12 h later, total RNA was extracted for RT-qPCR (**e**) and medium was collected for ELISA (**f**). Data are shown as mean ± s.e.m. of replicates (n=4 per group) of a representative experiment in three independent experiments. Comparisons in **c**–**f** were done on biological replicates by Ordinary one-way ANOVA with Tukey's multiple comparison test. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$; ns, no significant; nd, not detected.



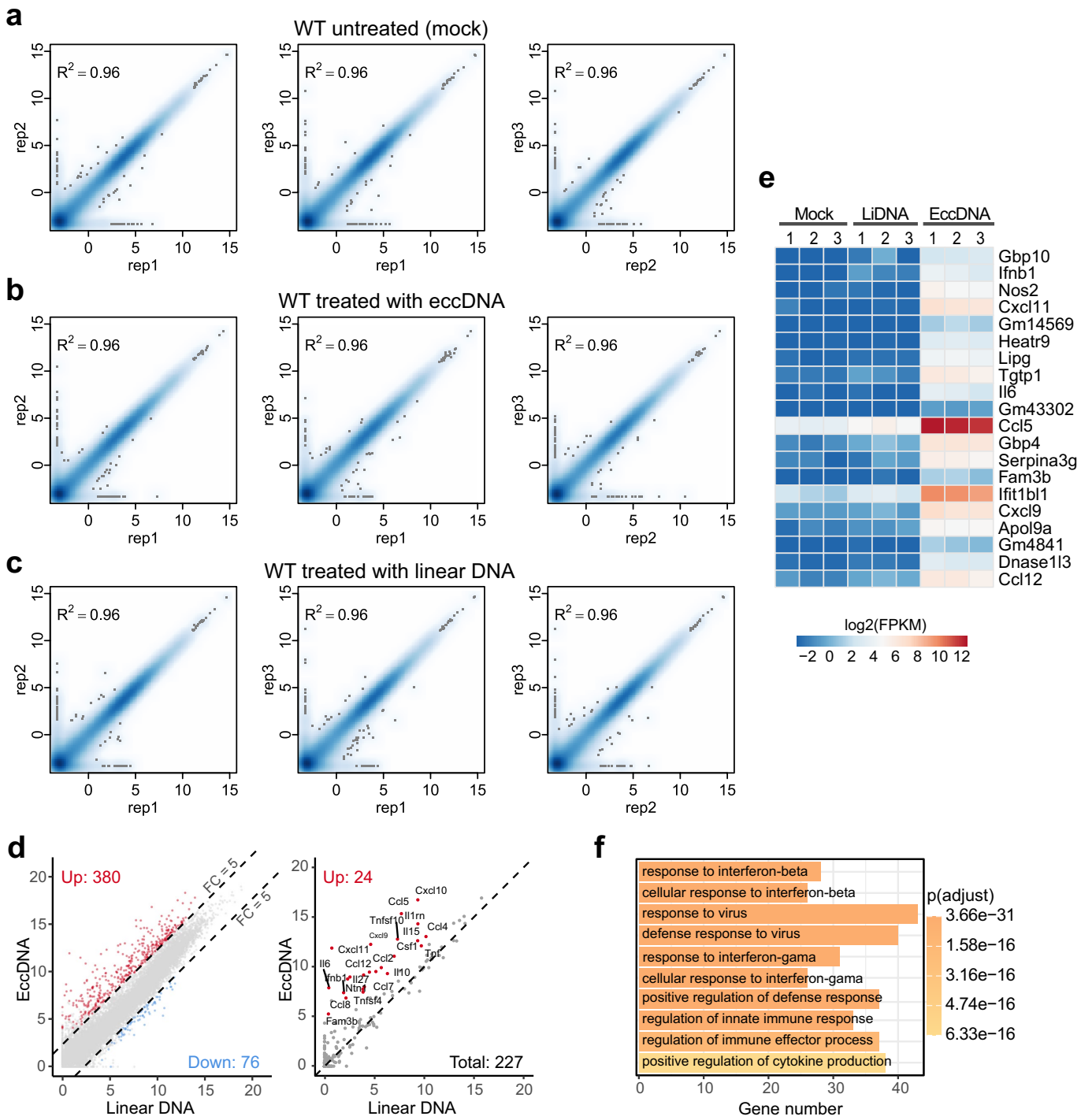
Extended Data Fig. 6 | EccDNAs are potent immunostimulants for BMDMs. **a**, Confirmation of BMDMs identify by flow cytometry. After differentiating bone marrow cells with L929 conditioned medium for 7 days, cells debris (left) were first excluded and singlets (middle) were phenotyped on the basis of F4/80 and CD11b expression (right) to define BMDM, the numbers indicate the percentage of gated cells. Data were further analyzed with FlowJo. V10.8.0 and shown as representative of 2 independent experiments. **b**, **c**, EccDNAs induce cytokine genes in BMDMs. Bar graphs showing the induction of mRNA (**b**) and protein (**c**) of *IL-6* and *TNF-α* in BMDMs that transfected with varying levels of eccDNA, compared with fragmented linear DNA, and poly(dG:dC). RT-qPCR were performed after 12-hours transfection. ELISA analysis (**c**) was performed after 24-hours transfection. Data are shown as mean ± s.e.m. of replicates (n=4 per group) of a representative experiment of three independent experiments. Comparisons were performed on biological replicates with equal amount of DNA transfection by Ordinary one-way ANOVA with Tukey's multiple comparison test. *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$; ns, no significant; nd, not detected.



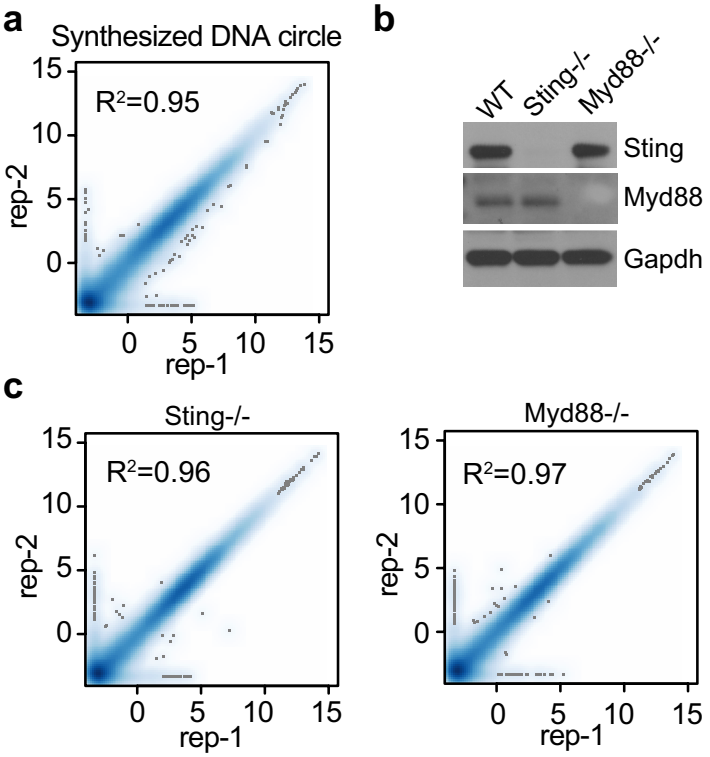
Extended Data Fig. 7 | The circularity of eccDNAs, but not the sequence, is critical for their immunostimulant activity. a, Bar graphs showing the relative *IL-6* and *TNF-α* mRNA levels. Data is shown as mean ± s.e.m. of replicates (n=4 per group) of a representative experiment in three independent experiments. **b**, **c**, Bar graphs showing the relative *IL-6* and *TNF-α* mRNA levels (**b**), and protein (ELISA) levels (**c**). Data is shown as mean ± s.e.m. of replicates (n=4 per group) of a representative experiment in three independent experiments. **d**, Diagram of DNA transfection efficiency and stability assay. Step-1: 5 Phosphorothioate (purple "*") end-protected synthetic linear DNA (PS-syn-linear) and its circular form with equal number of phosphorothioate bonds (PS-Syn-circular) were transfected to BMDCs in the same way as in Fig. 5 for either 1 or 12 h. Step-2: cells were lysed in 100 µl lysis buffer and treated with Thermoliable Proteinase K to prepare total DNA. Step-3: 4 µl total DNA was used for qPCR with a set of primers that amplify a 117 bp fragment in both linear and circular DNA. **e**, qPCR analysis of the samples prepared as described in (**d**). Transfected DNA level was normalized to that of 10 ng/ml PS-syn-linear transfection (n= 4 or 6, as indicated by dotes). Bars indicate mean ± s.e.m. of three independent experiments. **f**, 12 h after transfection of indicated DNA at 30 ng/ml, medium was collected for ELISA essay, Data are shown as mean ± s.e.m. of replicates (n=4 per group) of a representative experiment in three independent experiments. Comparisons in **a**–**c** were performed on biological replicates by Ordinary one-way ANOVA with Tukey's multiple comparison test; two-tailed unpaired *t* tests (**e**, **f**). *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$; ns, no significant.



Extended Data Fig. 8 | Reproducibility of the BMDC transcriptomes. a–**c**, Scatter plot of pair-wise comparison of the transcriptomes of three independent BMDC samples of untreated (**a**), treated with eccDNA (**b**) or linear DNA (**c**). The x and y-axis are log2(FPKM+0.1).



Extended Data Fig. 9 | Global transcriptional response to eccDNA in BMDMs. a–**c**, Scatter plot of pair-wise comparison of the transcriptomes of three independent BMDM samples of untreated (**a**), treated with eccDNA (**b**) or linear DNA (Extended Data Fig. 6). The x and y-axis are log2(FPKM+0.1). **d**, Scatter plot showing 380 genes (left panel, red dots) that are significantly induced (FC ≥ 5, adjusted p-value < 0.001; FC and p-value were generated using DESeq2[42] and p-values were adjusted by IHW[43]) by eccDNA, but not linear DNA, in BMDMs. 24 significantly induced cytokine genes are indicated (right panel, red dots). The x and y-axis are log2-transformed normalized read counts. **e**, Heatmap presentation of the top 20 induced genes. **f**, GO terms enriched in the genes activated by eccDNA treatment in BMDMs. The number of genes in each term and the p-value of the enrichment is indicated.



Extended Data Fig. 10 | Analysis of the eccDNA sensing pathway. a, Scatter plot indicates the transcriptomes of the two replicates of synthetic circular DNA treated BMDCs are highly similar. The x and y-axis are log2(FPKM+0.1). **b**, Western blot confirmation of the Sting-/- and Myd88-/- cells. Blots were performed once, but cell genotypes were further confirmed by mRNA sequences from RNA-seq data (data not shown). **c**, Correlation of the transcriptomes of 2 replicates eccDNA treated Sting-/- and Myd88-/- BMDC. The x and y-axis are log2(FPKM+0.1).

nature portfolio

Corresponding author(s): Yi Zhang

Last updated by author(s): Jul 28, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The single molecular long read sequencing data was collected from Oxford Nanopore MinION sequencer with flow cell R9.4.1 (FL-MIN106D). Guppy (version 3.5.2) was used for base calling from the fast5 files generated by the MinION sequencer. The short sequencing reads of eccDNA were collected from Illumina HiSeq 2500 in 150 PE mode. The RNA-seq data was sequenced by Illumina NextSeq 500 in PE mode.
Data analysis	Sequencing data was analyzed as described in the Methods section. The code used to generate consensus eccDNA sequences from Nanopore long reads was deposited in GitHub (https://github.com/YiZhang-lab/eccDNA_RCA_nanopore). The following public softwares were used for sequencing data analysis: porechop (version 0.2.4) minimap2 (version 2.17) trimmomatic (version 0.39) bwa (version 0.7.17) picard (version 2.23.4) bedtools (version 2.29.2) deeptools (version 3.5.0) IGV (version 2.7.2) circos (version 0.69-9) karyoploteR (version 1.14.1) STAR (version 2.7.6a) RSEM (version 1.3.3) DESeq2 (version 1.26.0) R (version 3.6.3) Python (version 3.7.4)

Atomic Force microscope images were processed with Gwyddion 2.50.
Quantification of densitometry was done by Image J 1.53e.
Flow cytometry analysis were done with FlowJo_v10.8.0
Ordinary One-way ANOVA and Two tailed unpaired T-test were done with GraphPad Prism 9.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The eccDNA sequencing data and RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) with accession number GSE165919.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	No statistical methods were used to predetermine sample size.
Data exclusions	For generating high-confidence consensus eccDNA from Nanopore reads, only sequences with at least 2 full passes were kept for downstream analysis. This step would remove potential linear DNA, partially amplified eccDNA and artefacts due to random jumping of Phi29 during amplification, because linear ones would have only 1 pass and random jumping ones would not repeat the same structure for twice or more times.
Replication	At least two or three independent experiments have been performed, and all repeat experiments generates similar results.
Randomization	Mice were randomly taken for collecting bone marrows. For comparison between WT and mutants, randomization was not feasible as genotypes of mice needed to be first determined therefore, randomization was not used.
Blinding	Most majority of the experiments and result analysis were carried out by the same person, who need to verify the control and experimental groups to correctly interpret the results, therefore blinding was not used. Furthermore, due to all our experiment measures were objective, no blinding should not affect interpretation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Anti-Lig1 (Proteintech, #18051-1-AP,1:1000), anti-Lig3 (BD Science, #611876, 1:1000), anti-Lig4 (1:1000, a kind gift from David Schatz at Yale, We couldn't find a competent commercial alternative for anti-Lig4 from David Schatz's lab.), anti-Gapdh (Thermo fisher, #AM4300, 1:20000), anti-CD11c Pacific Blue (Biolegend, Clone N418, #117321, 1:100), anti-F4/80 Alexa Fluor 488 (Biolegend, Clone BM8, #123119, 1:100), anti-CD11b APC (eBioscience, M1/70, #17-0112-81, 1:100), anti-MHC Class II PE (I-A/I-E) (M5/114.15.2, eBioscience, #12-5321-81, 1:100), anti-TMEM173/Sting(Proteintech, #19851-1-AP, 1:1000), Anti-Myd88(ProSci, #2127, 1:1000)

Validation

Antibodies for validation of cell identities are validated by previous publications:
(117321) Pacific Blue™ anti-mouse CD11c Antibody
https://www.citeab.com/antibodies/518339-117321-pacific-blue-anti-mouse-cd11c-antibody?des=e5ee8f77962bc3a5
(123119) Alexa Fluor® 488 anti-mouse F4/80 Antibody
https://www.citeab.com/antibodies/519048-123119-alexa-fluor-488-anti-mouse-f4-80-antibody?des=85185def3b09d650
(17011281)CD11b Monoclonal Antibody (M1/70), APC, eBioscience™
https://www.thermofisher.com/antibody/product/CD11b-Antibody-Monoclonal/17-0112-81
(12532181)MHC Class II (I-A/I-E) Monoclonal Antibody (M5/114.15.2), PE, eBioscience™
https://www.thermofisher.com/antibody/product/MHC-Class-II-I-A-I-E-Antibody-Monoclonal/12-5321-81

Anti-Lig1(Proteintech, #18051-1-AP), anti-Lig3 (BD Science, #611876) and antiLig4 (Dacid Schatz, Yale University) were validated by immunoblot results of this study, as well as in
Masani, S., Han, L., Meek, K., & Yu, K. (2016). Redundant function of DNA ligase 1 and 3 in alternative end-joining during immunoglobulin class switch recombination. Proceedings of the National Academy of Sciences, 113(5), 1261-1266.
Anti-Sting was validated by knockout in this study and previous studies (PMID: 30936875;PMID: 30943264;PMID: 33406424)
Anti-Myd88 was validated by knockout in this study and previous studies (PMID: 21233312;PMID: 28674182;PMID: 26671458)

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

HeLa cell lines was purchased from ATCC. mESC/E14 cell line was kindly provided by the laboratory of Beverly Koller. The parental CH12F3 line was originally from Dr. Tasuku Honjo, the other derivatives were generated in the Alt lab. L929 cell line was a kind gift of Drs. Judy Lieberman (Boston Children's Hospital), Commercial alternative of L929 cell lines is available at ATCC by Cat. ccl-1.

Authentication

HeLa Cell line obtained were authenticated by ATCC.
mESC/E14 cell line has been authenticated in
Fu, X., Wu, X., Djekidel, M.N. and Zhang, Y., 2019. Myc and Dnmt1 impede the pluripotent to totipotent state transition in embryonic stem cells. Nature cell biology, 21(7), pp.835-844.

Identities of BMDCs and BMDMs were confirmed by the canonical cell surface makers CD11c+/MHCII+ and CD11b+/F4/80+ double positive respectively, as shown in Extended Data Fig.S5a and S6a.
L929 was authenticated by its production of conditional medium to high efficiently differentiate Bone marrow cells to macrophages (BMDMs) , over 93% of differentiated BMDMs were positive for CD11b+/F4/80+, the canonical markers of macrophages (Extended Data Fig.S6a).

Mycoplasma contamination

Mycoplasma tests were negative in all cell lines by MycoAlert Mycoplasma detection kit from Lonza.

Commonly misidentified lines
(See ICLAC register)

None of the cell lines in this study is on the list of ICLAC.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

Male Wild Type C57BL/6 mice, Sting1/gt (Sting-/-, https://www.jax.org/strain/017537) Myd88/tm.1.1Defr (Myd88-/-, https://www.jax.org/strain/009088) mice were purchased from Jackson Laboratories for BMDCs and BMDMs preparation. Mice were randomly selected for bone marrows isolation.

Wild animals

No wild animals was involved in this study.

Field-collected samples

No animal in this study was collected from the field.

Ethics oversight

All procedures on animals involved in this study were conducted according to protocols approved by Harvard Medical School IACUC.

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

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For cell death analysis, mESCs were trypsinized and stained with LIVE/DEAD™ Fixable Far Red Dead Cell Stain Kit, then fixed with 90% methanol, cells were resuspended in PBS before cytometry analysis. BMDCs and BMDMs were detached in Versene Solution (Thermo Fisher, #A4239101), and stained with the indicated cell surface makers after blocking FcRs with Mouse TruStain FcX Plus (BioLegend, #156603).
Instrument	BD FACSCanto II
Software	Data were collected with BD FACSDiva software, and then analyzed by FlowJo software V.10.8.0
Cell population abundance	At least 20,000 events were quantified.
Gating strategy	Significant separations between negative and positive populations were used to define cell status and identities. FSC-A/SSC-A were used to exclude debris, and FSC/FSH gates were used to define singlets. Marched isotype control antibodies were included to define positive/negative gates for BMDCs and BMDMs.

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