

## Video Article

# Genome-wide Purification of Extrachromosomal Circular DNA from Eukaryotic Cells

Henrik D. Møller<sup>1</sup>, Rasmus K. Bojsen<sup>2</sup>, Chris Tachibana<sup>3</sup>, Lance Parsons<sup>4</sup>, David Botstein<sup>5</sup>, Birgitte Regenberg<sup>1</sup><sup>1</sup>Department of Biology, University of Copenhagen<sup>2</sup>National Veterinary Institute, Technical University of Denmark<sup>3</sup>Group Health Research Institute<sup>4</sup>Lewis-Sigler Institute for Integrative Genomics, Princeton University<sup>5</sup>Calico Life Sciences LLCCorrespondence to: Birgitte Regenberg at [bregenberg@bio.ku.dk](mailto:bregenberg@bio.ku.dk)URL: <https://www.jove.com/video/54239>DOI: [doi:10.3791/54239](https://doi.org/10.3791/54239)

Keywords: Molecular Biology, Issue 110, Circle-Seq, deletion, eccDNA, rDNA, ERC, ECE, microDNA, minichromosomes, small polydispersed circular DNA, spcDNA, double minute, amplification

Date Published: 4/4/2016

Citation: Møller, H.D., Bojsen, R.K., Tachibana, C., Parsons, L., Botstein, D., Regenberg, B. Genome-wide Purification of Extrachromosomal Circular DNA from Eukaryotic Cells. *J. Vis. Exp.* (110), e54239, doi:10.3791/54239 (2016).

## Abstract

Extrachromosomal circular DNAs (eccDNAs) are common genetic elements in *Saccharomyces cerevisiae* and are reported in other eukaryotes as well. EccDNAs contribute to genetic variation among somatic cells in multicellular organisms and to evolution of unicellular eukaryotes. Sensitive methods for detecting eccDNA are needed to clarify how these elements affect genome stability and how environmental and biological factors induce their formation in eukaryotic cells. This video presents a sensitive eccDNA-purification method called Circle-Seq. The method encompasses column purification of circular DNA, removal of remaining linear chromosomal DNA, rolling-circle amplification of eccDNA, deep sequencing, and mapping. Extensive exonuclease treatment was required for sufficient linear chromosomal DNA degradation. The rolling-circle amplification step by  $\phi 29$  polymerase enriched for circular DNA over linear DNA. Validation of the Circle-Seq method on three *S. cerevisiae* CEN.PK populations of  $10^{10}$  cells detected hundreds of eccDNA profiles in sizes larger than 1 kilobase. Repeated findings of *ASP3-1*, *COS111*, *CUP1*, *RSC30*, *HXT6*, *HXT7* genes on circular DNA in both S288c and CEN.PK suggests that DNA circularization is conserved between strains at these loci. In sum, the Circle-Seq method has broad applicability for genome-scale screening for eccDNA in eukaryotes as well as for detecting specific eccDNA types.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/54239/>

## Introduction

Detecting early or transient chromosomal amplification is difficult because it requires identifying alterations in single DNA molecules in large populations of cells. Chromosomal copy-number variations (CNVs) are generally detected well after their establishment, leaving only the final CNV structure as evidence of the mechanism that generated the variation<sup>1,2</sup>. Detecting and recovering extrachromosomal circular DNA (eccDNA) in earlier stages of CNV formation might elucidate ongoing processes in genomic rearrangements.

Previously, *de novo* discovery of eccDNA was by electron micrographs<sup>3</sup>, Giemsa staining of metaphase chromosomes<sup>4</sup>, or two-dimensional gel electrophoresis<sup>5</sup>. These methods provide little or no information about the sequence of the circular DNA. Targeted techniques such as Southern blotting<sup>6,7</sup>, inverse PCR<sup>8</sup>, or fluorescence *in situ* hybridization<sup>9</sup> provide evidence only about specific eccDNA elements. None of these methods provide the sequence of all existing eccDNA types in a cell population.

Genomic divergence in a pool of cells can be characterized by genome sequencing and/or tiling arrays<sup>10,11</sup>. Detecting a deletion or amplification by conventional DNA purification methods usually requires that a mutated allele represent at least 0.1-1% of the cell population<sup>12,13</sup>. Acentric eccDNAs are expected to be even more transient in a cell culture due to their lack of centromeres and potential absence of DNA synthesis at replication. Thus, since most eccDNAs presumably are in low amounts and their sequences resemble the genome, alternative DNA extraction methods are needed to detect eccDNAs.

Several circular DNA purification techniques exploit the structural differences between chromosomes and circular DNA. For instance, high-speed ultracentrifugation in cesium-chloride gradients is used to isolate 350-3000 basepairs (bp) large eccDNAs from the human HeLa cancer cell line<sup>14</sup>. However, high-velocity can break or nick the backbone of supercoiled circular DNA structures, altering the sedimentation velocity<sup>15</sup> and the eccDNA yield. Dutta and coworkers developed a method for *de novo*, genome-scale identification of circular DNA from mouse tissues as well as from cultures of chicken and human cells<sup>16,17</sup>. Their method is extraction of nuclei from homogenized tissue by sucrose ultracentrifugation followed by plasmid purification and several rounds of enzymatic reactions and DNA extractions. Their protocol primarily identifies 200-400 bp

eccDNAs, called microDNAs. Dutta and coworkers also attempted purification of microDNAs from *Saccharomyces cerevisiae* but were unable to record microDNA from this yeast species<sup>16</sup>.

We have developed a novel method for *de novo* detection of eccDNA from yeast called Circle-Seq. This method enables genome-scale surveys for circular DNA molecules large enough to carry whole genes and as large as the 86 kilobase (kb) mitochondrial DNA (mtDNA). The Circle-Seq method was developed from a well-established prokaryotic plasmid purification method<sup>18,19</sup>, optimized for eukaryotic yeast cells and combined with deep sequencing. Using the Circle-Seq approach, 1756 different eccDNAs, all larger than 1 kb, were detected from ten *S. cerevisiae* S288c populations<sup>20</sup>. A size cut-off was chosen to focus on eccDNA that were large enough to carry whole genes. Circle-Seq was highly sensitive; it detected a single eccDNA within thousands of cells<sup>20</sup>. In the current study, Circle-Seq was used to isolate and identify 294 eccDNAs from three biological replicates of another *S. cerevisiae* yeast strain, CEN.PK. The data reveals that eccDNA is a common genetic element in *S. cerevisiae* strains.

## Protocol

NOTE: An overview of the circular DNA purification and sequencing method (Circle-Seq) is illustrated in **Figure 1**.

### 1. Cultivation, Cell Harvest and Plasma Membrane Disruption

- Inoculate yeast cells (for example *Saccharomyces cerevisiae*) from an O/N culture into 50 ml complete nutrient medium of yeast peptone dextrose (YPD). Inoculate at a low initial cell density of  $1-3 \times 10^5$  cells/ml or an optical density of approximately 0.01 OD<sub>600</sub>.
  - Incubate the cells at 30°C with agitation at 150 rounds per minutes (rpm) until cells reach maximum cell density of approximately  $1 \times 10^{10}$  cells, approximately after 24 to 48 hr or a optical density at OD<sub>600</sub> > 10.0.  
NOTE: The cultivation time is not crucial as lower cell concentrations can be used.
- Transfer the outgrown culture to a 50 ml conical tube, pellet the cells by centrifugation at 800 x g for 3 min and discard the supernatant.
- Wash the pellet with 25 ml buffer solution of 10 mM Tris-Cl, 1 mM EDTA, pH 8.0, re-pellet the cells by centrifugation at 800 x g for 3 min and discard the supernatant.
- Resuspend the cell pellet in 1.2 ml resuspension buffer supplied from a plasmid column-purification kit.
- Optional step: Add highly diluted plasmids as controls for purification of circular DNA elements<sup>20</sup>.  
NOTE: In the current dataset, a 7.7 µl plasmid mixture was applied for each sample containing  $10^{10}$  cells. The plasmid stock mixture consisted of three plasmids in different concentrations; pBR322 at 38 ng/sample, pUC19 at 0.5 ng/sample, and pUG72 at 0.01 ng/sample.
- Transfer the cell suspension into two 2 ml micro-centrifuge tubes, each supplemented with 0.5 mm glass beads at a 1:3 ratio of the total suspension volume.
- Vortex each tube at maximum speed for 10 min to disrupt plasma cell membranes. Pellet the beads by centrifugation at 268 x g for 30 sec and transfer the 1.2 ml combined supernatant from the two microcentrifuge tubes to a new tube.  
NOTE: Alternative to step 1.6-1.7, use zymolyase to disrupt cells in 0.6 ml resuspension buffer solution. Ten units of zymolyase can disrupt  $5 \times 10^7$  cells within 1.5 hr at 35°C.

### 2. EccDNA Enrichment by Column Chromatography

- Follow the protocol from a kit for column purification of plasmids. In brief, treat each sample with 1.2 ml alkaline solution, mix gently and incubate 3 min at RT.
- Add 1.2 ml neutralization buffer, mix gently and centrifuge at 9650 x g for 5 min.
- Load the solution onto a column equilibrated with 1 ml equilibration solution and allow the liquid to flow through the column by gravity.
- Wash the column with 4 ml washing solution. When the solution has passed through the resin, carefully add 0.3 ml elution solution to replace most of the 0.35 ml column void volume.
- Elute DNA into a new collection tube with 1 ml elution solution and precipitate the DNA by adding 0.8 ml precipitation mixture. Centrifuge at 9650 x g for 10 min.
- Wash the DNA pellet with 0.5 ml 70% ethanol, centrifuge at 9650 x g for 5 min, air dry for 5 to 15 min and dissolve the purified DNA in 25 µl sterile water.  
NOTE: Only short term storage of DNA in water is recommended. Preferentially, proceed directly to step 3.

### 3. Digestion of Remaining Linear Chromosomal DNA

- Optional step: To facilitate specific digestion of linear DNA by exonuclease, treat the purified DNA with a rare-cutting endonuclease such as NotI. For 5 µg DNA, use 1 unit NotI, 5 µl 10x digestion buffer and sterile water to a total volume of 50 µl. Incubate the reaction at 37°C for 16 hr and heat inactivate the endonuclease at 80°C for 5 min.
- Add 20 units exonuclease (2 µl), 4 µl ATP (25 mM), 34 µl sterile water and 10 µl 10x reaction buffer directly to the 50 µl endonuclease-cleaved DNA to reach a 1x reaction volume of 100 µl, using the ATP-dependent exonuclease kit.
- Perform hydrolysis of linear single-stranded and double-stranded DNA at 37°C for 5 days or more. Add an additional 4 µl ATP (25 mM), 0.6 µl 10x reaction buffer and 20 units exonuclease every 24 hr to continue the enzymatic DNA digestion at a 1x reaction volume.
- After removal of linear DNA, sample 2 µl from the exonuclease treated solution to confirm elimination of chromosomal linear DNA by quantitative polymerase chain reaction (qPCR), using a chromosomal marker such as the actin gene *ACT1*<sup>20</sup>.
  - Each 20 µl qPCR reaction volume contains 2 µl exonuclease-treated sample, 150 nM *ACT1* primers 5'-TCCGTCTGGATTGGTGGTTCTA-3' and 5'-TGGACCACTTTCGTCGTATTC-3', 2% (volume/volume) dimethyl sulfoxide, and 10 µl green fluorescent master mix.
  - Use the reaction condition; 3 min at 95°C, followed by 45 cycles of 15 sec at 95°C and 30 sec at 60°C.

NOTE: *ACT1* is a particularly suitable marker for linear DNA since copy number variations in this gene are deleterious<sup>21-23</sup> so eccDNA should not carry *ACT1*.

3. Alternatives to analysis of DNA digestion by qPCR are standard PCR (4.3) or propidium iodide staining (4.4).
  1. Use 2  $\mu$ l exonuclease-treated sample as PCR template with *ACT1* primers 5'-TGGATTCTGGTATGTTCTAGC-3' and 5'-GAACGACGTGAGTAACACC-3'. As positive *ACT1* control, use 50-100 ng genomic *S. cerevisiae* DNA as template. PCR reaction conditions; 3 min at 95°C, followed by 35 cycles of 30 sec at 95°C, 30 sec at 56°C and 1 min at 72°C.
  2. Run PCR reactions by gel electrophoresis on 1% agarose with 0.5  $\mu$ g/ml ethidium bromide. Look for a 0.8 kb *ACT1* band.
4. The absence or presence of linear DNA can also be examined by propidium iodide staining before and after DNA amplification.
  1. Mix each DNA sample in a 1:1 volume with a 1:1,000 H<sub>2</sub>O-diluted solution of 20 mM propidium iodide stock. Leave solution in darkness for 10-20 min at RT and analyze DNA staining by fluorescence microscopy at 100x magnification using a red excitation fluorescence filter at 663-738 nm and an exposure-time of 5 to 30 sec. As DNA-staining control, use  $\phi$ 29-amplified genomic DNA from yeast and/or  $\phi$ 29-amplified plasmid.
5. Heat inactivate the exonuclease solution at 70°C for 30 min.

## 4. DNA Amplification

1. Amplify the purified and enriched eccDNA from step 3.5) with  $\phi$ 29 DNA polymerase<sup>24-26</sup> according to the protocol of the polymerase manufacturer.
  1. In brief, mix 5  $\mu$ l enriched eccDNA with 5  $\mu$ l denaturation buffer.
  2. After 3 min at RT, add 10  $\mu$ l neutralization buffer. Mix gently and add 30  $\mu$ l master mix containing 29  $\mu$ l reaction buffer and 1  $\mu$ l  $\phi$ 29 DNA polymerase. Incubate the reaction at 30°C for 16 hr or more (up to 72 hr). Heat inactivate the  $\phi$ 29 DNA polymerase at 65°C for 3 min.

## 5. Sequencing and Data Analysis

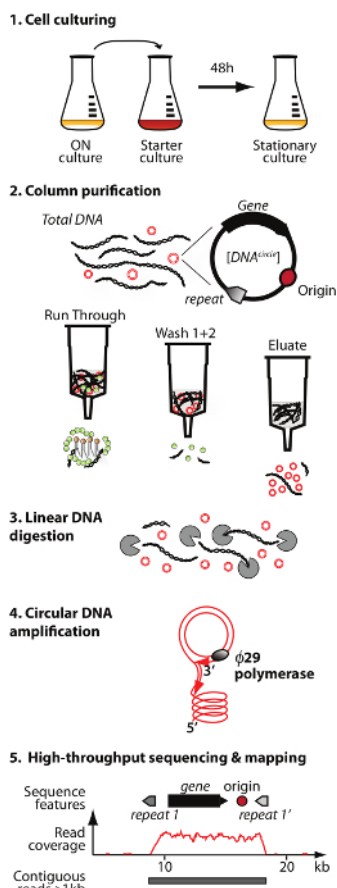
1. Shear the amplified eccDNA with a focused ultrasonicator to an average target peak size of 300 bp. Use the following settings for a 130  $\mu$ l DNA sample: 450W peak intensity power, 60 sec treatment, 30% duty factor, 200 cycles per burst, temperature 7°C.
  2. Add barcode index labels and adapters to the fragmented reads for synthesis of libraries for sequencing, using an appropriate method for library preparation.
  3. Run deep sequencing, for example as 141-nucleotide single-end reads on a high-throughput sequencing platform.
  4. Map reads to the yeast reference genome under investigation and allow reads to map to multiple regions. For example, use a freely available workflow system<sup>27,28</sup> and short-read aligner mapping software<sup>29</sup>.
  5. Identify reads from regions of putative eccDNAs using contiguous reads, for example, more than seven contiguous reads (>1 kb) without gaps<sup>20</sup>.
- NOTE: Software is available<sup>27,28</sup> for exploring mapped reads at genomic regions of interest.

## Representative Results

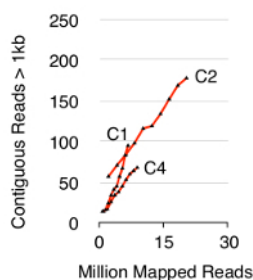
To validate the Circle-Seq method, three *S. cerevisiae* CEN.PK populations of  $1 \times 10^{10}$  cells were screened after cells were grown separately in YPD for ten generations. Chromosomal linear DNA elimination was confirmed by the absence of a qPCR *ACT1* signal as described previously<sup>20</sup> (data not shown). Purified and enriched eccDNA was sequenced up to 68 million reads (141-nucleotide single-end reads) and mapped to the CEN.PK113-7D reference genome (version 19 June 2012). Recordings of putative eccDNAs from the three samples named C1, C2 and C4 were assigned to genomic regions mapped by contiguous reads longer than 1 kb. Based on 10,000 Monte Carlo simulations, the significance of each region mapped by contiguous reads longer than 1 kb was estimated. From this 79, 159 and 56 regions were annotated as likely eccDNA sequences ( $p < 0.1$ , **Dataset 1**). The number of recorded contiguous reads > 1 kb increased as a function of sequence depth suggesting that even more eccDNA elements would have been recorded if samples had been sequenced further (**Figure 2**). As expected, the Circle-Seq method extracted numerous reads from a number of known circular DNA elements including the 2 $\mu$  plasmid, mitochondrial DNA, ribosomal RNA genes on chromosome XII, and the three internal control plasmids pBR322, pUC19 and pUG72 that were spiked into samples just before column purification (**Figure 3**).

The video shows an example of contiguous reads that mapped to the *HXT7\_ARS432\_HXT6* locus on chromosome IV. Previously, the [*HXT6/7*<sup>circle</sup>] was detected by Circle-Seq in ten S288c populations (each with  $1 \times 10^{10}$  cells) and the circular DNA structure was confirmed by inverse PCR analysis<sup>20</sup>. The [*HXT6/7*<sup>circle</sup>] was also recorded in each of the three CEN.PK populations (**Figure 4A**). Moreover, most of the common eccDNA genes among replicate samples of CEN.PK overlapped eccDNA genes from the S288c datasets (**Figure 4B**).

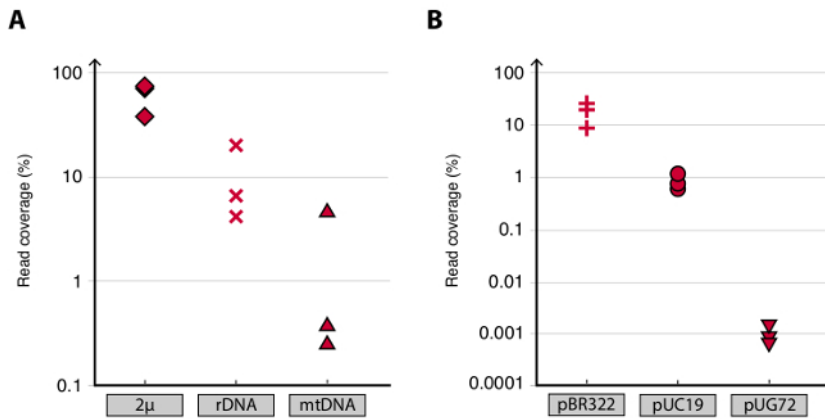
To test the specificity of the Circle-Seq protocol for circular DNA purification, two samples, each with 30  $\mu$ g genomic DNA, were tested. One sample was supplemented with 100 ng plasmid DNA and eccDNA from both samples were purified by the Circle-Seq protocol. After column separation, the DNA yield was 1.27% (380 ng) for the sample without plasmid (GD) and 1.60% (480 ng) for the sample with plasmid (GD +P). The efficiency of exonuclease treatment was tested for linear DNA content after 29 hr and 72 hr using PCR against *ACT1*. No samples contained amplified *ACT1* (data not shown). A fraction of each exonuclease-treated sample was further amplified by the  $\phi$ 29 polymerase and the products of enzymatic reactions were analyzed by propidium iodide staining (Figure 5A-F) and agarose gel electrophoresis (Figure 5G). Samples after exonuclease treatment showed minimal propidium iodide-stain (Figure 5A-B). The  $\phi$ 29-amplified sample with only genomic DNA revealed thread-like structures (Figure 5C) similar to the control sample (Figure 5E). The  $\phi$ 29-amplified sample that had added plasmid revealed foci (Figure 5D) resembling the plasmid control (Figure 5F). The images indicated that  $\phi$ 29 polymerase enriched for circular DNA over linear DNA. Most linear chromosomal DNA was removed from samples after 29 hr exonuclease treatment (Figure 5A-B, G). However, extensive exonuclease treatment for more than 100 hr and using more than 100 units was needed to remove all chromosomal linear DNA, as  $\phi$ 29-amplified samples still showed a background of thread-like structures after 72 hr exonuclease treatment (Figure 5C-D).



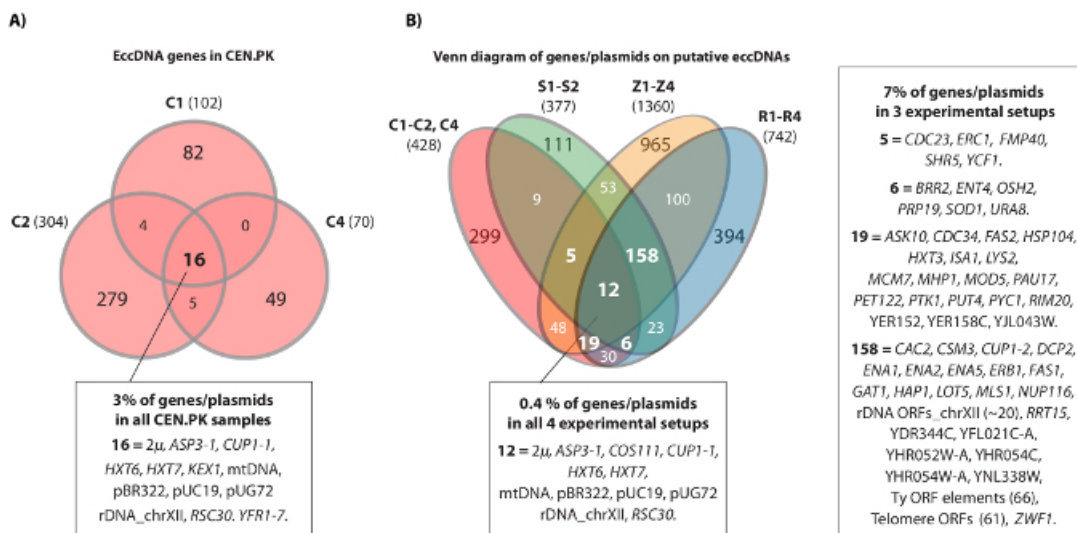
**Figure 1. Outline of the Circle-Seq method.** The protocol has 5 steps: 1) cell culturing, 2) purification and enrichment of eccDNA by column chromatography, 3) digestion of remaining linear chromosomal DNA in the eluate fraction, 4) amplification of DNA by  $\phi$ 29 DNA polymerase, and 5) sequencing of highly enriched eccDNA and mapping of reads to the *S. cerevisiae* reference genome. [Please click here to view a larger version of this figure.](#)



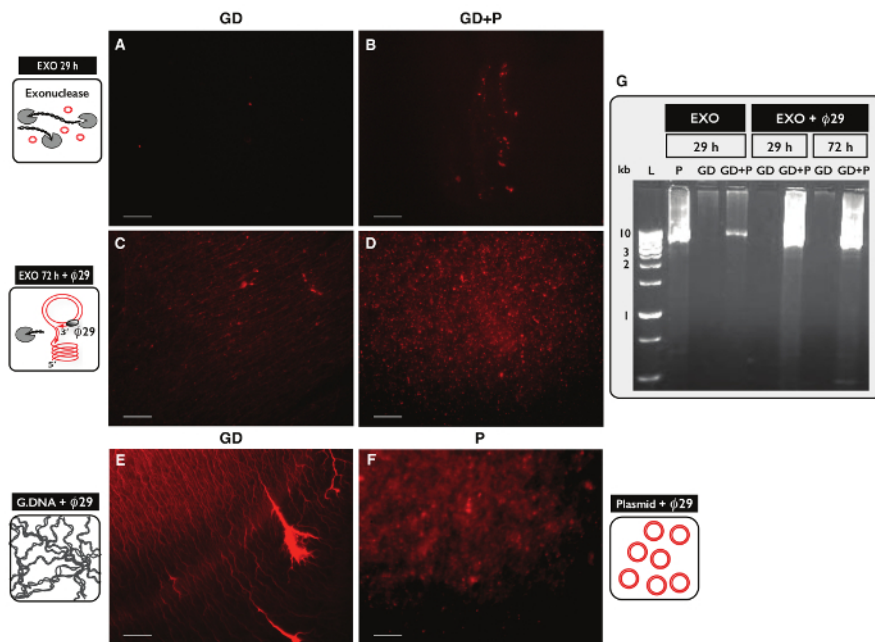
**Figure 2. Contiguous reads > 1 kb as function of sequence depth.** EccDNA from  $1 \times 10^{10}$  cells increase as a function of sequence depth (in millions of mapped reads). Shown: biological triplicates from haploid CEN.PK *S. cerevisiae* populations (C1, C2, C4) separated by  $10^{10}$  cell divisions. [Please click here to view a larger version of this figure.](#)



**Figure 3. Detection of known circular DNA elements.** (A-B) Scatter plots of read coverage (read density) in percent for plasmids in CEN.PK biological replicates C1, C2 and C4. (A) Mapped reads to the endogenous yeast plasmids were: 2μ; [*rDNA<sup>circle</sup>*] (ribosomal RNA genes from chromosome XII); and mtDNA (the mitochondrial DNA). (B) Unique reads mapped to control plasmids. Control plasmids were spiked into samples before column purification. Plasmid ratios per cell were: pBR322 (plus signs) 1:1, pUC19 (circles) 1:50, and pUG72 (triangles) 1:2,500.



**Figure 4. Common eccDNA elements in CEN.PK and S288c.** (A) Venn diagram displaying overlap among the 476 genes on 294 eccDNA elements in the three CEN.PK samples (C1, C2, C4). The 16 common overlapping eccDNA genes/plasmids are annotated (all gene names are in Dataset 1). (B) Venn diagram of all recorded genes on putative eccDNAs from the three CEN.PK samples (C1, C2, C4), compared to all recorded genes on putative eccDNAs from 10 S288c samples: S1-S2, R1-R4, Z1-Z4 (see reference<sup>20</sup>). Shown are 13 biological replicates (S1-S2, R1-R4, Z1-Z4, C1-C3) with genes/plasmids and putative eccDNA regions that overlapped a minimum of 2 strain backgrounds and either 3 or more experimental setups. C samples, CEN.PK; R and Z samples, S288c BY4741; S samples, S288c M3750. [Please click here to view a larger version of this figure.](#)



**Figure 5. Visualization of DNA samples after exonuclease and  $\phi 29$  treatment. (A-F)** Propidium iodide staining of DNA. Scale bar, 10  $\mu\text{m}$ . **(A, C and E)** Samples with genomic DNA (GD); **(B and D)** samples with GD plus plasmid (GD+P). **(A-B)** After 29 hr exonuclease treatment (EXO 29 h); **(C-D)** after 72 hr exonuclease treatment followed by  $\phi 29$  polymerase amplification (EXO 72 h +  $\phi 29$ ). **(E)** Genomic DNA control after  $\phi 29$  polymerase amplification; **(F)** plasmid control (5.5 kb) after 29 polymerase amplification; **(G)** agarose gel-electrophoresis. From left: L, 1 kb markers; P, plasmid control (5.5 kb) after EXO 29 hr; GD, after EXO 29 hr (sample as in **A**); GD+P, after EXO 29 hr (sample as **B**); GD and GD+P, after EXO 29 hr +  $\phi 29$ ; GD and GD+P, after EXO 72 hr +  $\phi 29$  (sample as in **C-D**). See Table S1 for extra details. [Please click here to view a larger version of this figure.](#)

Sample	eccDNA mapping				eccDNA content			eccDNA coverage and p-values							
	CHROM	START	END	ARS	complete_gp	partial_gene	BLASTN Ident	UNIQUE_BP	ALL_READ	C_ALL_FPKM	ALL_PVAL	UNIQUE_READ	UNIQUE_FPKM	UNIQUE_PVAL	
C4	2-micron	0	6318	0				1265	25324816	64217.5285	8.57E-05	3335	299.481788	0.00086493	
C7	2-micron	0	6318	0				1087	11054263	33970.7784	0.0001517	1458	65.9442422	0.00029148	
C1	2-micron	0	6318	0				1292	28351617	68252.1847	0.00018474	4624	540.578271	0.00087263	
C3	chrI	149318	150702	0	YAI001C			1306	42	0.58708548	0.76040679	39	1.46814967	0.05840066	
C2	chrI	152418	154262	0	YAR003C-A, YAR003W			1844	56	0.58760798	0.76048406	56	1.49305571	0.05826482	
C2	chrI	161460	163807	0	YAR074C, YAR015W			2377	96	0.79810772	0.04316885	91	1.92362003	0.03312725	
C2	chrII	277504	279228	0	YBR021W, YBR023C			1324	30	0.43634891	0.05982667	30	1.1399226	0.04648861	
C2	chrIII	279384	280498	0	YBR013C			1208	31	0.49400309	0.05586588	28	1.19554735	0.04527384	
C2	chrIII	292835	293989	0	YBR030W			1081	1288	21.5021312	0.00530511	959	43.6155964	0.00291363	
C1	chrIII	361914	363138	0	YBR063C, YBR065C			1224	27	0.33550631	0.02836102	27	3.3183904	0.02152419	
C1	chrIII	363831	365967	0	YBR066C, YBR067C			2136	39	0.2770372	0.03493575	39	2.75781434	0.02737681	
C4	chrIII	460474	463012	0	YBR114W, YBR115C			2538	80	0.50515097	0.01404933	79	3.53590482	0.00964674	
C1	chrIII	620603	622082	0	YBR202W			1424	30	0.30651155	0.03115049	27	2.8841379	0.02583471	
C1	chrIII	622118	623327	1	YBR203W			1209	31	0.3899924	0.02418203	28	3.4983948	0.02015998	
C2	chrIII	639220	640260	0	YBR210W			1002	36	0.60666236	0.04720079	30	1.47198178	0.03857555	
C1	chrIII	651834	652997	0	YBR217W			1163	779	12.9281818	0.00762556	744	31.451555	0.00415111	
C2	chrIII	749141	750412	0	YBR275C			1271	21	0.31963947	0.07159087	21	0.81223158	0.05740549	
C1	chrIII	787465	789500	0	YBR295W			1860	43	0.34049683	0.02739366	39	3.1670284	0.02304734	
C1	chrIII	51234	52512	0	YCL029W, YCL040W			1054	17	0.21949366	0.04474627	15	2.14957357	0.03705891	
C2	chrIII	73063	76309	1	YCL025C, YCL026C-B			3046	170	1.07970642	0.03725239	167	2.68547603	0.02766226	
C4	chrIII	82617	83644	0	YCL021W-A			926	54	0.84264789	0.01021817	41	5.02965243	0.00719739	
C1	chrIII	91033	92287	0	YCL018W	YCL018W		1254	31	0.37599537	0.02515787	30	3.61347775	0.01939156	
C2	chrIII	91578	92741	0	YCL018W	YCL018W		1163	28	0.46176799	0.05710963	27	1.14138768	0.04323993	
C1	chrIII	164680	166537	1	YCR078C			1852	48	0.29564125	0.02372733	36	2.39400823	0.02545153	
C2	chrIII	165746	167890	1	YCR078C	YCR078C		2015	59	0.53232399	0.0585102	54	1.31758153	0.04184299	

**Dataset 1. Potential DNA circularization regions in CEN.PK. [Please click here to download this file.](#)**

Shown are sequence data and analyses for 348 regions. Columns are A-D, eccDNA mapping. A (first column from left), sample from which putative eccDNA was identified; B, chromosome; C-D, start and end coordinates of putative eccDNAs. E-H, eccDNA content. E, autonomously replicating sequence (ARS) in the region; F, complete gene in the region; G, part of gene included in the region; H, BLASTN-identified gene. I-O, eccDNA coverage and p-values. I, longest region with a uniquely annotated sequence in bp; J, number of all mapped reads; K, coverage of all mapped reads by fragments per kb from a million mapped reads (FPKM); L, p-value for putative eccDNA compared to occurrence by chance from Monte Carlo simulations; M, number of uniquely mapped reads; N-O; as K and L using only uniquely mapped reads (UFPKM). Parameters for mapping of reads and Monte Carlo simulations were as described<sup>20</sup>.

**Discussion**

The Circle-Seq method allows genome-scale detection of eccDNA from yeast cells with sequence-level resolution. The method is a mild eccDNA purification that does not require intensive vortex or pipetting and uses column separation by gravity to limit eccDNA breakage that would lead to exonuclease digestion in the subsequent step. These features of the method may be crucial for detecting large eccDNAs that contain gene sequences. Circle-Seq detected numerous eccDNAs including full genes (**Dataset 1**). It also detected the 86-kb yeast mitochondrial DNA. Thus, this protocol facilitates purification of large circular DNA elements. Keeping the number of DNA extraction steps to a minimum reduces the risk of eccDNA loss and maximizes yield. Based on results for control, spiked-in plasmids, Circle-Seq is highly sensitive, detecting a single circular DNA from 2,500 cells. Furthermore, removing abundant endogenous plasmids such as 2 $\mu$  plasmid or mitochondrial DNA might significantly enhance sensitivity. Curing of 2 $\mu$  from yeast cultures has been described<sup>30</sup>. Alternatively, 2 $\mu$  and mitochondrial DNA removal might be achieved with a

rare-cutting endonuclease, such as *SwaI*. However, the restriction enzyme step could target other eccDNAs of interest and limit the total eccDNA yield.

Critical steps for eccDNA detection were removal of linear DNA (step 3) and DNA sequencing (step 5) to a proper depth. To record the majority of eccDNAs from a cell population, deep sequencing might be required<sup>20</sup>. Paired-end sequencing should provide even greater confidence of eccDNA detection, as circular DNA junctions are expected to yield paired-end reads that map discordantly. These discrepancies support the discovery of circular DNA structures and can potentially be used as an additional eccDNA-detection filter.

The Circle-Seq method was validated using three independent *S. cerevisiae* CEN.PK populations. Detected sequences included previously reported eccDNAs, endogenous plasmids and spiked-in plasmids and hundreds of putative eccDNAs (**Dataset 1**). These findings support previous Circle-Seq datasets from *S. cerevisiae* S288c<sup>20</sup>. The discovery of several eccDNAs common to CEN.PK and S288c populations indicates that these loci have a propensity to exist as circular elements (**Figure 4**). We have previously shown that the [*GAP1*<sup>circle</sup>] is enriched under nitrogen limited conditions in the CEN.PK background<sup>8</sup>, though evidence of [*GAP1*<sup>circle</sup>] in other strain backgrounds has not been found. Finding of eccDNA from the *CUP1-1*, *RSC30*, *ASP3-1*, *COS111*, and *HXT6* *HXT7* loci in both S288c and CEN.PK suggests that a predisposition for DNA circularization is conserved between the yeast strains. It remains to be shown if [*HXT6/7*<sup>circle</sup>], [*ASP3-1*<sup>circle</sup>], [*COS111*<sup>circle</sup>], and [*CUP1-1* *RSC30*<sup>circle</sup>] confer selective advantages to cells or if their existence is merely an effect of high rates of DNA circularization.

Taken together, the results indicate that Circle-Seq is well suited for detecting kilobase-sized eccDNAs and has advantages for identifying eccDNAs with complete genes. Circle-Seq is a highly sensitive method that enables whole genome-scale screens of eccDNAs from yeast. The Circle-Seq method could open a new field of research aimed at elucidating the role of eccDNA in generating gene deletions and amplifications. Given that DNA architecture and structure are largely conserved from eukaryotic yeast to higher eukaryotes, the Circle-Seq method should, in principle, be applicable to all eukaryotic cells, with slight modifications. At present, the method does not appear to have any limitations, although its ability to purify megabase-sized eccDNAs has yet to be shown. In addition, the use of  $\phi$ 29 DNA polymerase, which uses a rolling-circle amplification method<sup>31</sup>, creates a bias towards smaller eccDNAs making eccDNA quantification more difficult. Circle-Seq detects eccDNAs large enough to carry full genes, making it suitable for studies on double minutes—circular DNA from human somatic cells. Double minutes can contribute to cancer when proto-oncogenes are amplified on these elements<sup>32-37</sup>. Studies of eccDNAs in germline cells could be used to measure germline mutation rates and assess sperm quality, for example in livestock. Thus, Circle-Seq has the potential to yield insights into the rate at which genetic variation arises in the form of copy number variation, and lead to a novel understanding of diseases that involve gene copy number variation<sup>38-40</sup>.

## Disclosures

The authors declare that they have no competing financial interests.

## Acknowledgements

Thanks to Kenn D. Møller and Claus Sternberg (DTU) for technical assistance and to Tue S. Jørgensen for quantitative PCR analysis.

## References

1. Kugelberg, E., Kofoid, E., *et al.* The Tandem Inversion Duplication in *Salmonella enterica*: Selection Drives Unstable Precursors to Final Mutation Types. *Genetics*. **185** (1), 65-80, (2010).
2. Reams, A. B., Kofoid, E., Savageau, M., & Roth, J. R. Duplication Frequency in a Population of *Salmonella enterica*. Rapidly Approaches Steady State With or Without Recombination. *Genetics*. **184** (4), 1077-1094, (2010).
3. Smith, C. A., & Vinograd, J. Small polydisperse circular DNA of HeLa cells. *Journal of Molecular Biology*. **69** (2), 163-178, (1972).
4. Carroll, S. M., DeRose, M. L., *et al.* Double Minute Chromosomes Can Be Produced from Precursors Derived from a Chromosomal Deletion. *Molecular and cellular biology*. **8** (4), 1525-1533, (1988).
5. Cohen, S., Yacobi, K., & Segal, D. Extrachromosomal Circular DNA of Tandemly Repeated Genomic Sequences in *Drosophila*. *Genome research*. **13** (6A), 1133-1145, (2003).
6. Horowitz, H., & Haber, J. E. Identification of Autonomously Replicating Circular Subtelomeric Y' Elements in *Saccharomyces cerevisiae*. *Molecular and cellular biology*. **5** (9), 2369-2380, (1985).
7. Moore, I. K., Martin, M. P., Dorsey, M. J., & Paquin, C. E. Formation of Circular Amplifications in *Saccharomyces cerevisiae*. by a Breakage-Fusion-Bridge Mechanism. *Environmental and molecular mutagenesis*. **36** (2), 113-120, (2000).
8. Gresham, D., Usaite, R., Germann, S. M., Lisby, M., Botstein, D., & Regenbreg, B. Adaptation to diverse nitrogen-limited environments by deletion or extrachromosomal element formation of the *GAP1* locus. *Proceedings of the National Academy of Sciences of the United States of America*. **107** (43), 18551-18556, (2010).
9. Windle, B., Draper, B. W., Yin, Y. X., O'Gorman, S., & Wahl, G. M. A central role for chromosome breakage in gene amplification, deletion formation, and amplicon integration. *Genes & development*. **5** (2), 160-174, (1991).
10. Gresham, D., Ruderfer, D. M., *et al.* Genome-Wide Detection of Polymorphisms at Nucleotide Resolution with a Single DNA Microarray. *Science (New York, NY)*. **311** (5769), 1932-1936, (2006).
11. Kidd, J. M., Cooper, G. M., *et al.* Mapping and sequencing of structural variation from eight human genomes. *Nature*. **453** (7191), 56-64, (2008).
12. Gresham, D., Desai, M. M., Botstein, D., & Dunham, M. J. The Repertoire and Dynamics of Evolutionary Adaptations to Controlled Nutrient-Limited Environments in Yeast. *PLoS Genetics*. **4** (12), 1-19, (2008).
13. Lang, G. I., Botstein, D., & Desai, M. M. Genetic Variation and the Fate of Beneficial Mutations in Asexual Populations. *Genetics*. **188** (3), 647-661, (2011).

14. van Loon, N., Miller, D., & Murnane, J. P. Formation of extrachromosomal circular DNA in HeLa cells by nonhomologous recombination. *Nucleic Acids Research*. **22** (13), 2447-2452, (1994).
15. Vinograd, J., & Lebowitz, J. Physical and Topological Properties of Circular Dna. *Journal of General Physiology*. **49** (6P2), 103-125, (1966).
16. Shibata, Y., Kumar, P., *et al.* Extrachromosomal MicroDNAs and Chromosomal Microdeletions in Normal Tissues. *Science (New York, NY)*. **336** (6077), 82-86, (2012).
17. Dillon, L. W., Kumar, P., *et al.* Production of Extrachromosomal MicroDNAs Is Linked to Mismatch Repair Pathways and Transcriptional Activity. *Cell Reports*. **11** (11), 1749-1759, (2015).
18. Li, L. L., Norman, A., Hansen, L. H., & Sørensen, S. J. Metamobilomics - expanding our knowledge on the pool of plasmid encoded traits in natural environments using high-throughput sequencing. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*. **18 Suppl 4**, 8-11, (2012).
19. Brown Kav, A., Sasson, G., Jami, E., Doron-Faigenboim, A., Benhar, I., & Mizrahi, I. Insights into the bovine rumen plasmidome. *Proceedings of the National Academy of Sciences of the United States of America*. **109** (14), 5452-5457, (2012).
20. Møller, H. D., Parsons, L., Jørgensen, T. S., Botstein, D., & Regenberg, B. Extrachromosomal circular DNA is common in yeast. *Proceedings of the National Academy of Sciences of the United States of America*. **112** (24), E3114-E3122, (2015).
21. Drubin, D. G., Miller, K. G., & Botstein, D. Yeast Actin-Binding Proteins - Evidence for a Role in Morphogenesis. *The Journal of cell biology*. **107** (6), 2551-2561, (1988).
22. Magdolen, V., Drubin, D. G., Mages, G., & Bandlow, W. High levels of profilin suppress the lethality caused by overproduction of actin in yeast cells. *FEBS letters*. **316** (1), 41-47, (1993).
23. Sandrock, T. M., Brower, S. M., Toenjes, K. A., & Adams, A. Suppressor analysis of fimbrin (Sac6p) overexpression in yeast. *Genetics*. **151** (4), 1287-1297, (1999).
24. Blanco, L., Bernad, A., Lázaro, J. M., Martín, G., Garmendia, C., & Salas, M. Highly Efficient DNA Synthesis by the Phage  $\phi$ 29 DNA Polymerase. *The Journal of biological chemistry*. **264** (15), 8935-8940 (1989).
25. Dean, F. B. Rapid Amplification of Plasmid and Phage DNA Using Phi29 DNA Polymerase and Multiply-Primed Rolling Circle Amplification. *Genome research*. **11** (6), 1095-1099, (2001).
26. Hutchison, C. A., Smith, H. O., Pfannkoch, C., & Venter, J. C. Cell-free cloning using  $\phi$ 29 DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*. **102** (48), 17332-17336, (2005).
27. Goecks, J., Nekrutenko, A., Taylor, J., & Galaxy Team, T. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology*. **11** (8), R86, (2010).
28. Giardine, B., Riemer, C., *et al.* Galaxy: A platform for interactive large-scale genome analysis. *Genome research*. **15** (10), 1451-1455, (2005).
29. Langmead, B., & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. **9** (4), 357-359, (2012).
30. Tsalik, E. L., & Gartenberg, M. R. Curing *Saccharomyces cerevisiae* of the 2 micron plasmid by targeted DNA damage. *Yeast (Chichester, England)*. **14** (9), 847-852, (1998).
31. Norman, A., Riber, L., Luo, W., Li, L. L., Hansen, L. H., & Sørensen, S. J. An Improved Method for Including Upper Size Range Plasmids in Metamobilomes. *PLoS ONE*. **9** (8), e104405, (2014).
32. Storlazzi, C. T., Lonoce, A., *et al.* Gene amplification as double minutes or homogeneously staining regions in solid tumors: Origin and structure. *Genome research*. **20** (9), 1198-1206, (2010).
33. Hoff, Von, D. D., Needham-VanDevanter, D. R., Yucel, J., Windle, B. E., & Wahl, G. M. Amplified human MYC oncogenes localized to replicating submicroscopic circular DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America*. **85** (13), 4804-4808 (1988).
34. Raymond, E., Faivre, S., *et al.* Effects of hydroxyurea on extrachromosomal DNA in patients with advanced ovarian carcinomas. *Clinical cancer research : an official journal of the American Association for Cancer Research*. **7** (5), 1171-1180, (2001).
35. Shimizu, N. Extrachromosomal Double Minutes and Chromosomal Homogeneously Staining Regions as Probes for Chromosome Research. *Cytogenetic and genome research*. **124** (3-4), 312-326, (2009).
36. Eckhardt, S. G., Dai, A., Davidson, K. K., Forseth, B. J., Wahl, G. M., & Hoff, Von, D. D. Induction of differentiation in HL60 cells by the reduction of extrachromosomally amplified *c-myc*. *Proceedings of the National Academy of Sciences of the United States of America*. **91** (14), 6674-6678 (1994).
37. Vogt, N., Lefèvre, S.-H., *et al.* Molecular structure of double-minute chromosomes bearing amplified copies of the epidermal growth factor receptor gene in gliomas. *Proceedings of the National Academy of Sciences of the United States of America*. **101** (31), 11368-11373, (2004).
38. Ahn, K., Gotay, N., *et al.* High rate of disease-related copy number variations in childhood onset schizophrenia. *Molecular psychiatry*. **19** (5), 568-572, (2013).
39. Girirajan, S., Johnson, R. L., *et al.* Global increases in both common and rare copy number load associated with autism. *Human molecular genetics*. **22** (14), 2870-2880, (2013).
40. Vogt, N., Gibaud, A., Lemoine, F., la Grange, de, P., Debatisse, M., & Malfoy, B. Amplicon rearrangements during the extrachromosomal and intrachromosomal amplification process in a glioma. *Nucleic Acids Research*. **42** (21), 13194-13205, (2014).