Genome-wide purification of extrachromosomal circular DNA from eukaryotic cells

Møller, Henrik D.; Bojsen, Rasmus Kenneth; Tachibana, Chris; Parsons, Lance; Botstein, David; Regenberg, Birgitte

Published in:
Journal of Visualized Experiments

DOI:
10.3791/54239

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY-NC-ND

Citation for published version (APA):

Download date: 28. Mar. 2022
Genome-wide Purification of Extrachromosomal Circular DNA from Eukaryotic Cells

Henrik D. Møller1, Rasmus K. Bojsen2, Chris Tachibana3, Lance Parsons4, David Botstein5, Birgitte Regenberg1

1Department of Biology, University of Copenhagen
2National Veterinary Institute, Technical University of Denmark
3Group Health Research Institute
4Lewis-Sigler Institute for Integrative Genomics, Princeton University
5Calico Life Sciences LLC

Correspondence to: Birgitte Regenberg at bregenberg@bio.ku.dk

Video Article

The video component of this article can be found at http://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 variation1,2. 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 micrographs3, Giemsa staining of metaphase chromosomes4, or two-dimensional gel electrophoresis5. These methods provide little or no information about the sequence of the circular DNA. Targeted techniques such as Southern blotting6,7, inverse PCR8, or fluorescence in situ hybridization9 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 arrays10,11. 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 population12,13. 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 line14. However, high-velocity can break or nick the backbone of supercoiled circular DNA structures, altering the sedimentation velocity15 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 cells16,17. Their method is extraction of nuclei from homogenized tissue by sucrose ultracentrifugation 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.  

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, 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. 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. 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

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

1. Inoculate yeast cells (for example Saccharomyces cerevisiae) from an O/N culture into 50 ml complete nutrient medium of yeast peptone dextrose (YPD). Incubate at a low initial cell density of 1-3 x 10^7 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 x 10^9 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.

2. 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.

3. Wash the pelleted 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.

4. Resuspend the cell pellet in 1.2 ml resuspension buffer supplied from a plasmid column-purification kit.

5. Optional step: Add highly diluted plasmids as controls for purification of circular DNA elements.
   - NOTE: In the current dataset, a 7.7 µl plasmid mixture was applied for each sample containing 10<sup>10</sup> 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.

6. 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.

7. 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 x 10<sup>9</sup> cells within 1.5 hr at 35°C.

#### 2. EccDNA Enrichment by Column Chromatography

1. 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.

2. Add 1.2 ml neutralization buffer, mix gently and centrifuge at 9650 x g for 5 min.

3. Load the solution onto a column equilibrated with 1 ml equilibration solution and allow the liquid to flow through the column by gravity.

4. 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.

5. 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.

6. 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

1. 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.

2. 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.

3. 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 1 x reaction volume.

4. 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.  
   - Each 20 µl qPCR reaction volume contains 2 µl exonuclease-treated sample, 150 nM ACT1 primers 5’TCCGGTCAGTTTTACCGCTTC’-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\(^{21-23}\) 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 µl 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 \textit{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 µg/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\(_2\)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 ø29-amplified genomic DNA from yeast and/or ø29-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 ø29 DNA polymerase\(^{24-26}\) according to the protocol of the polymerase manufacturer.
      1. In brief, mix 5 µl enriched eccDNA with 5 µl denaturation buffer.
      2. After 3 min at RT, add 10 µl neutralization buffer. Mix gently and add 30 µl master mix containing 29 µl reaction buffer and 1 µl ø29 DNA polymerase. Incubate the reaction at 30°C for 16 hr or more (up to 72 hr). Heat inactivate the ø29 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 µl 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\(^{27,28}\) and short-read aligner mapping software\(^{29}\).
   5. Identify reads from regions of putative eccDNAs using contiguous reads, for example, more than seven contiguous reads (>1 kb) without gaps\(^{20}\).

   \textbf{NOTE:} Software is available\(^{27,28}\) for exploring mapped reads at genomic regions of interest.

\textbf{Representative Results}

To validate the Circle-Seq method, three \textit{S. cerevisiae} CEN.PK populations of 1 x 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\(^{20}\) (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, \textit{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 (\textit{Figure 2}). As expected, the Circle-Seq method extracted numerous reads from a number of known circular DNA elements including the 2µ 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 (\textit{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](\text{circ}) was detected by Circle-Seq in ten S288c populations (each with 1 x 10\(^{10}\) cells) and the circular DNA structure was confirmed by inverse PCR analysis\(^{30}\). The [HXT6/7](\text{circ}) was also recorded in each of the three CEN.PK populations (\textit{Figure 4A}). Moreover, most of the common eccDNA genes among replicate samples of CEN.PK overlapped eccDNA genes from the S288c datasets (\textit{Figure 4B}).
To test the specificity of the Circle-Seq protocol for circular DNA purification, two samples, each with 30 µg 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 ø29 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 iodine-stain (Figure 5A-B). The ø29-amplified sample with only genomic DNA revealed thread-like structures (Figure 5C) similar to the control sample (Figure 5E). The ø29-amplified sample that had added plasmid revealed foci (Figure 5D) resembling the plasmid control (Figure 5F). The images indicated that ø29 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 ø29-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 ø29 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 x 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\textsuperscript{circle} (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 \textsuperscript{20}). 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 ø29 treatment. (A-F) Propidium iodide staining of DNA. Scale bar, 10 µm. (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 ø29 polymerase amplification (EXO 72 h + ø29). (E) Genomic DNA control after ø: ø29 polymerase amplification; (F) plasmid control (5.5 kb) after ø29 polymerase amplification; (G) agarose gel-eletrophoresis. 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 in B); GD and GD+P, after EXO 29 hr + ø29; GD and GD+P, after EXO 72 hr + ø29 (sample as in C-D). See Table S1 for extra details. Please click here to view a larger version of this figure.

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.

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 2µ plasmid or mitochondrial DNA might significantly enhance sensitivity. Curing of 2µ from yeast cultures has been described. Alternatively, 2µ 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. 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. 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] circle is enriched under nitrogen limited conditions in the CEN.PK background, though evidence of [GAP1] 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(circle)], [ASP3-1(circle)], [COS11(circle)], and [CUP1-1(RSC30)] 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 a29 DNA polymerase, which uses a rolling-circle amplification method\(^{32-37}\), 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\(^{32-37}\). 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\(^{38-40}\).

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


