

Formation of extrachromosomal circular DNA in HeLa cells by nonhomologous recombination

Nanette van Loon⁺, Douglas Miller and John P. Murnane*

Laboratory of Radiobiology and Environmental Health, University of California, San Francisco, CA 94143-0750, USA

Received May 13, 1994; Accepted June 6, 1994

ABSTRACT

Extrachromosomal circular DNA (eccDNA) generated from chromosomal DNA is found in all mammalian cells and increases with cell stress or aging. Studies of eccDNA structure and mode of formation provide insight into mechanisms of instability of the mammalian genome. Previous studies have suggested that eccDNA is generated through a process involving recombination between repetitive sequences. However, we observed that approximately one half of the small eccDNA fragments cloned from HeLa S3 cells were composed entirely of nonrepetitive or low-copy DNA sequences. We analyzed four of these fragments by polymerase chain reaction and nucleotide sequencing and found that they were complete eccDNAs. We then screened a human genomic library with the eccDNAs to isolate the complementary chromosomal sequences. Comparing the recombination junctions within the eccDNAs with the chromosomal sequences from which they were derived revealed that nonhomologous recombination was involved in their formation. One of the eccDNAs was composed of two separate sequences from different parts of the genome. These results suggest that rejoining of ends of fragmented DNA is responsible for the generation of a substantial portion of the eccDNAs found in HeLa S3 cells.

INTRODUCTION

Extrachromosomal circular DNA (eccDNA) has been found in all mammalian cells studied, and most if not all eccDNA found in the nucleus is chromosomal in origin (1). Mammalian eccDNAs (also known as small polydisperse circular DNA, spcDNA) range in size from a few hundred base pairs to double minutes that can contain millions of base pairs. Most eccDNA has no known function, although double minutes in tumor cells often contain amplified copies of drug resistance genes or oncogenes and therefore play an important role in cancer (2, 3). Some eccDNAs show sequence homology to transposons (4–6),

and others are the product of developmentally controlled processes, such as excision products of immunoglobulin gene rearrangement (7).

There is no consensus about the mechanisms of formation or maintenance of eccDNAs in mammalian cells. Most smaller eccDNAs contain interspersed (8–14) or tandemly repeated (8, 14–17) DNA sequences. The prevalence of repetitive sequences suggests that homologous recombination is often involved (14–17, 23). However, a few eccDNAs have been found to be derived from recombination between short repeat sequences (6, 11, 18), and in one case, nonhomologous recombination (18). eccDNAs may also be generated by reverse transcription (9) and topoisomerases (24).

Whether smaller eccDNAs can replicate or are continuously turning over is also unknown. It has been proposed that some eccDNAs replicate and provide a selective advantage to the cell (19). Constructs incorporating putative eukaryotic autonomously replicating sequences have been stably maintained over many generations as circular DNAs in cell culture (25). The ability of naturally occurring eccDNAs to replicate independently has been established in eccDNAs containing elements homologous to viral retrotransposons (5, 6) and in stable double minute chromosomes. It has been proposed that double minutes arise through the increase in size of smaller circles (26, 27); however, other studies have found no evidence for an increase in the size of circular DNAs (3). Although double minutes appear to reintegrate into the genome (27), the frequency and mechanisms of integration of eccDNAs are unknown. Most transfected DNA is integrated into the mammalian genome through nonhomologous recombination (28).

Inhibitors of nuclear DNA or protein synthesis (cycloheximide, hydroxyurea, puromycin) and other stresses such as ionizing radiation, hypoxia, and maintenance in stationary phase can increase the number of eccDNAs in eukaryotic cells (8, 19, 20). In addition, changes in the size distribution of eccDNA populations can be observed during the aging process (21, 22). These results show that the distribution and size of eccDNAs reflect the stability of chromosomal DNA. The smaller size and lesser complexity of eccDNAs compared with chromosomal

*To whom correspondence should be addressed

⁺Present address: Department of Microbiology and Immunology, University of Rochester Medical Center, Rochester, NY 14642, USA

DNA makes them well suited for analyzing instability in specific genomic sequences of mammalian cells.

To address some of the unanswered questions concerning eccDNAs, we studied naturally occurring eccDNAs in the HeLa S3 cell line. We cloned several eccDNAs from this cell line and compared their sequences with complementary sequences within chromosomal DNA. We then characterized the recombination junctions by polymerase chain reaction (PCR) and nucleotide sequence analysis to further our understanding of the mechanisms involved in their formation.

MATERIALS AND METHODS

Cell culture

HeLa S3 cells were obtained from the American Type Culture Collection and grown as monolayers on 75-cm² flasks in minimal essential medium supplemented with 5% calf bovine serum in a humidified incubator with a mixture of 95% air and 5% CO₂ at 37°C.

Purification of eccDNAs

HeLa cells (4.3×10^8) were harvested by scraping into 300 ml of isotonic saline. Circular DNA was extracted by slightly modifying the Qiagen Inc. protocol for isolation of plasmid DNA. The cells were centrifuged for 15 min in a GS3 rotor at 2000 rpm ($675 \times g$) at 4°C. The cell pellet was resuspended in 15 ml of 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 100 µg/ml RNase I (Qiagen P1 buffer) and divided into three aliquots for easier handling in subsequent steps. After 5 min, each aliquot was lysed by adding 5 ml of 200 mM NaOH containing 1% sodium dodecyl sulfate (SDS) (Qiagen P2 buffer), and the nuclear DNA was precipitated on ice after the addition of 5 ml of 3 M potassium acetate, pH 5.5 (Qiagen P3 buffer). The lysate was centrifuged for 15 min in an SS34 rotor at 17,000 rpm ($34,500 \times g$) at 4°C. The clear supernatant was loaded onto an equilibrated Qiagen Tip 500 ion-exchange chromatography column, and circular DNA was eluted according to the manufacturer's instructions. The DNA was then further purified by ethidium bromide/cesium chloride density gradient centrifugation for 16 h at $300,000 \times g$. The lower band visible under black light contained circular DNA and was removed with a 21-gauge needle. The DNA was extracted twice with an equal volume of isoamyl alcohol to remove the ethidium bromide, and was concentrated and separated from the cesium chloride with a Centricon 30 concentrator.

Cloning of eccDNAs into Bluescript

A 100-µl sample containing 3 µg of heterogeneous eccDNAs was boiled for 10 min in a water bath and allowed to cool gradually to 40°C. The eccDNA was digested at 37°C with *EcoRI*, and after 2 h the DNA was phenol extracted, mixed with *EcoRI*-cut Bluescript (Stratagene), and ligated overnight at 16°C. The next day the ligation products were used to transform DH5α competent bacteria (Stratagene).

Purification of eccDNA inserts from Bluescript

Probes for colony screening and Southern hybridization were purified from agarose preparative gels. The eccDNA insert bands were cut out of the gels and placed in dialysis tubing in 0.5×buffer E (20 mM Tris-HCl, pH 7.4, 10 mM sodium acetate, 0.5 mM Na₂ EDTA). The DNA was electrophoresed

out of the gel fragments and loaded onto NACS chromatography columns (Bethesda Research Laboratories). Purified inserts were labeled with α³²P-dCTP or with digoxigenin-UTP (Boehringer Mannheim).

Primer construction

Primers (18 bases) with a 40–60% GC content were made on a PCR Mate synthesizer with small-scale (0.2 µm) synthesis columns (Cruachem). The primers were removed from the columns and activated by flushing with ammonium hydroxide for 1 h, followed by overnight incubation at 55°C in the same solution. The ammonium hydroxide was evaporated in a Speedvac vacuum centrifuge. The primers were resuspended in water and diluted to 10 pmol/µl for PCR experiments and to 2 pmol/µl for nucleotide sequencing.

Polymerase chain reaction

Amplification was carried out in a 50-µl reaction volume containing 0.5 µg of genomic DNA or 1 ng of plasmid DNA, 0.4–1.0 pmol/µl of each primer, 5 µl of 10×PCR buffer (450 mM Tris-HCl, pH 8.8, 40 mM MgCl₂, 166 mM (NH₄)₂SO₄, 7% (v/v) β-mercaptoethanol, 20% (v/v) formamide, in H₂O), 5 µl of 1 mM dNTPs, and 0.25 µl of Perkin-Elmer Taq polymerase (5 units/µl). Mineral oil was added to each 0.5-ml reaction tube to prevent evaporation. A Hybaid Omnigene temperature cycler was used at the following settings: 2 min at 93°C, followed by 35 cycles of 8 sec at 93°C, 4 sec at 55°C, and 60 sec at 76°C (29).

Nucleotide sequencing

Sequencing reactions were performed with a Sequenase version 2.0 sequencing kit (United States Biochemical Corporation), according to the manufacturer's instructions. Five micrograms of plasmid DNA and 4 pmol of 18-base primer were used for each reaction. The DNA was labeled with ³⁵S-dCTP. The reactions were electrophoresed on 6% polyacrylamide gels on a Sequencing Gel Electrophoresis System (Bethesda Research Laboratories) for 2 h at 70 watts. The gels were fixed for 1 h in a solution of 10% methanol and 10% acetic acid, dried, and exposed overnight to photographic film.

Lambda library screening and preparation of lambda DNA

A human primary fibroblast genomic library in the lambda Dash vector (Stratagene) was plated, with LE392 bacteria as host. Plaque lifts were taken with the use of Fisher Magnagraph or Hybond N 137-mm nylon filters, and hybridization was performed with ³²P-labeled eccDNA probes. For DNA preparation, lambda particles eluted from purified plaques were incubated for 15 min with log-phase LE392 bacteria and inoculated into 100 ml of NZCYM growth medium (Gibco). After growth overnight, lysis was stimulated by adding chloroform to 0.8%. The lysate was treated with DNase I and RNase A and centrifuged to remove the bacterial debris. Phage particles were precipitated by adding polyethylene glycol 8000, cooling on ice for 2 h, and centrifuging again. The phage pellet was resuspended in a small volume of SM medium (0.1 M NaCl, 8 mM MgSO₄·7H₂O, 2% gelatin, 0.05 M Tris, pH 7.5), treated with DNase I and RNase A, and lysed with SDS. Phage DNA was precipitated with potassium acetate, purified by sequential extractions with phenol-chloroform and chloroform, ethanol precipitated, and resuspended in TE buffer.

eccDNA 15

```

1 (G)AATTCACTCT CTTTAATACA GGCACTATA TTTTACTTTT GGAAGCCTTT ATATTGTGTA TTGATTCCAT GAAGGGCATT GCTCAAAAAA CATAGGCTCC
      15 INV K5a =>
      t gtatccgagg
      <= 15 KS
101 ATGGCCAOCT GTCTGTCTTG TTTAATGGGG CCATTCTGAG ATATCATAGC CTGGAGGAGA CATTCCTTAA CCAGTCCCAT CAGAAAAAA TCACCTCAGTG
      taccoggt
      gacctcctct gtaagaaa
      <= 15 K5b
201 TTTACCTTCC AGGATTTTCAG TCTAGGTTCT TCAAAGACAT ACACGAGATT TTTGTAAATA GCAATCATTT GCAACATGTT CTGGAAGACC COGGCACTGA
301 AACAAACCCA CCATACATAA GAACACTTCA TCAGCTAATT GGAAACATC AITGTGTGTC TATCTTTAAA TTGCTACCAA TCATCCCAAG CAAAAACTTA
401 TATGCCAATA TTATCAGAGA CAATTACTTA GCATGTCACA CTGGTAGTC TGAAAAAAA AACTAGCAAA TGCTTCTGCG AATAAATAAC CACATGTGCA
      cacgt
501 CTTAGGATTC AACCTAATTT TACTCAGTGA GGAAGAGGC ACTGTGTAAA TACATGCTGG CAGTTTTTTT TTTCTTTTTT TTTGTTTCTT CTTTTTAAAT
      15 SKa =>
      gaatcctaag ttg
      <= 15 INV SKa
601 GTATATTCAA GGACAACAGA TAAACTGAAA TTGGAAAAG TTTGACTGGA GAGAACTGCC AATTCAGTCA GAGGAGTTAG GTTACAAATC TCACTTG (AATTC)
      atataagtt cctgttgc
      <= 15 INV SKb
      tggact cctctgaagg ta
      <= 15 INV SK

```

eccDNA 24

```

1 (G)AATTCTTTC TGCTTTATGT TTCCAGGCCA AGGTTGGCCA CAAAAGAAC CTGTAGTGAA CTGGAAGGCA GAAGTGAAGG CACATCCAGT AGGCTCAGAA
101 GGGCTGGGG ATGTTGAAGC ACCTGCTCGG TGTGTCTGAT TATGGGCCAG GCCTGCAGCT TCCAGCTCCA GTGAGTCTCG GCCCAGGTGC GOGAGCAGAG
201 CAGCCTCCTG GCAGCTTCTT TCTGCAAGTC ACCTGTGACA GATCTGGGTT CTCTGTGTCG CTGTGCAITC CCAGGTTTTT CTTATCTTCC CACTACTTCT
301 GTCTCTCAGA CTCTCAGCCC TGCTGATCTA TGGTGAATTC AGGCCAACCA CCAGGTTTGG AGACAACAGC CTTCATAGA CTTCTTCCAC ATCTTCCATA
401 CTGTGATAAT GTATAACCCT AGAAGAAATC TCTAGCTCCA AATCACTCAC AGGGATTGAT CAGGGAAAGCA AAACTTTTTA TCAACACTGC CTTGGACTTC
501 ACACATCTCC AAAGCAGCGA TTAATTTTGG TAGTCACAGG TCAATGTGAA AATCACAATG GAATCTCTGG ATCTCTCACG ACAATTCGAA ACAITTTGCC
601 CAATGGATTC AATAGACCAG TGGGTCTCAA ACCTTAATGT ACACATGATC CACTGGGGGA TCGTGGTCTC TTTTGTAAAT GCAGATCTCG GTGCAGCAGG
701 CCTGGGGTGG GGGCTCAGAT TCTGCAITTG TAACAGGTTT COGGAGAATG CTGGTGTCTG TGTCCAGTGT ACCCACAGTT GAGTAGCAAG GTGTCAAAAG
801 G (AATTC)

```

eccDNA 80

```

1 (G)AATTCAAAG GTGTCTCTAA TCCTCTTCCA CTGAACCTCT TGCTGTAAACA GGCAAGGATC CTCTGCCAGG CAGCCTTGAG CCAGACAGAG GGGAAACCTG
101 ACTATACTCC GTGCATGATC TTCCCAATA GCCTTCTTTA TTGAGATAAT TTCATTGTTG TTGAAGGAGA GGTAGAGGTT TTTCTAAGCC CACTGAGGGA
201 AAGGGTTTGG CTTATTTTGG GGAACAATG TAGTGGAGAA TGTAGGGGGA GGAATGGGGA AGATGGNCTC CGAATTCAGG GCAGCCAGAC CAACCAAGAA
301 GGCCTTACTA TATGCCITGT TGAGTCTGGA GTTTATGCAAT GTGGGAAGTT AGGATCCAAA CTTCACAAAT TTTTACATGA TAGAAGTGG (AATTC)

```

eccDNA 82

```

1 (G)AATTCTCAG AITGATGGTT CATTAGAAAT TGTGCAATGCC CACCTACCAC CCAGTGTCTT TACTTCTTTA TGTTTTTTTG GAATCCGGAA GATCTGAGTT
101 TCAATCTTAG TTTTGCCTACT GTTTGACCTT GGGCAAGATA AITGACCTCT CTGAACITTA TTTCTCTAAT TGTAAAACCTG GCCTAATCAT GCTGATGGGG
201 TTTGTTGGGG GAATGAGAGA ATACATAATA AAGGTTTAGT AATGAGCCIA GCACATAGAA TGTACTCAAT AAATGATGGC TATGATATG ACAAGCATG
301 CCCACTTTTA AATAGTCCAT CTCCCTTAAA TAGTTAGGAA GCTCTCTTGT GCTTGACTTG CAGTGCCAC TAGTAGGCTC TOSGTTGTTG CTTGTTAGGT
401 TGAATCAATC CAAGCCITTT TCCGAGTACA CTTGAGGGCT TGAGGGATTC AGGAGCTTAA ACITTCAGATT GTAAATMAAA GTCTTTTTTA TTAAGCAGTA
501 AGTTTGAATT TTTAACAATA TACCAGAACT TAITCACCTA GATTGACTCC TTTCTTCCAA AGCAGTACAC GTTATATGGA TAATGCTGTT ACTACTTAAA
601 AGTCTTTGTA AITCTTCTTT GGAATGGCTA CTAAGCCGG TTTTGTAGAC ACACAAAACA GTAGTCTCAT TAITTGAGGG CAATATCTTA TATATGATAA
701 AATGAGTAT GTCAITTTAT CACCAGATA ATTAAATAAA CTGAAGGCTT ATAG (AATTC)

```

Figure 1. Nucleotide sequences of eccDNAs 15, 24, 80, and 82. The sequences begin and end at the *EcoRI* cloning sites. The sequences of oligonucleotide primers used for PCR and sequence analysis are underlined. The locations of the recombination junctions are shown by arrows, and multiple junctions are indicated where overlaps were observed.

Southern blot analysis

DNA was digested with the appropriate enzymes and separated on agarose gels. Thirteen micrograms of genomic DNA or 300 ng of cloned lambda DNA were added per well. The DNA was transferred overnight onto positively charged nylon filters (Boehringer Mannheim). Filters were baked at 80°C for 2 h in

a vacuum oven and prehybridized for at least 2 h in Genius prehybridization buffer (Boehringer Mannheim). Hybridization was carried out overnight at 65°C in a shaking water bath. Probes were labeled with either $\alpha^{32}\text{P}$ -dCTP or Genius digoxigenin-UTP (Boehringer Mannheim). Radioactively labeled filters were washed in 2×SSC (1×SSC = 0.15 M sodium chloride, 0.015 M sodium citrate), 0.1% SDS and 0.5×SSC, 0.1% SDS, and

exposed. Digoxigenin-UTP-labeled filters were incubated for 1 h in Genius buffer 2 (Boehringer Mannheim), incubated with alkaline phosphatase conjugate, rinsed, and activated with Genius buffer 3 (Boehringer Mannheim). Hybridization was detected either colorimetrically (NBT+ X phosphate; Boehringer Mannheim) or with chemoluminescence (Lumiphos 530; Boehringer Mannheim).

RESULTS

Nucleotide sequence analysis of cloned eccDNAs

Purified eccDNA from HeLa S3 cells was digested with *EcoRI* restriction enzyme and cloned into the Bluescript vector. Forty-five clones containing eccDNA fragments ranging in size from 350 to 3000 base pairs (bp) were isolated. To identify inserts containing repetitive DNA, we analyzed clones representing the full range of insert sizes on Southern blots, using total human genomic DNA as a probe. Of the clones analyzed, approximately one half were composed of repetitive sequences and one half were composed of nonrepetitive or low-copy sequences (data not shown). Nucleotide sequence analysis showed that most of the repetitive inserts were the same human mitochondrial DNA *EcoRI* restriction fragment. Inserts containing only nonrepetitive DNA were analyzed further because the chromosomal DNA from which they originated could be unambiguously identified. The four inserts described here (clones 15, 24, 80, and 82) were shown by nucleotide sequence analysis to have sizes of 696, 800, 388, and 753 bp, respectively (Figure 1). None had significant homology to known sequences in GenBank.

Polymerase chain reaction analysis of cloned eccDNAs

PCR analysis was performed to compare the internal arrangement of the eccDNA sequences with that of the chromosomal sequences from which they originated. Total genomic DNA from HeLa cells, total genomic DNA from LM217 cells, or the eccDNA clones themselves were used as templates. The primers were constructed with sequence information from the cloned eccDNAs, as shown for eccDNA 15 in Figure 1. The PCR products obtained are shown in Figure 2 (with eccDNA as template) and Figure 3 (with total genomic DNA as template). Primers pointing toward each other on eccDNA 15 always gave PCR products when the eccDNA was used as a template (Figure 2). However, with many primers (e.g., 15-inv-KSb and 15-inv-SKa) PCR products were not obtained from genomic DNA (Figure 3, lane 10), indicating that rearrangements had occurred during formation of the eccDNA. Other primers gave rise to the same product on both eccDNA (Figure 2, lane 6) and genomic DNA (Figure 3, lane 7), indicating that no rearrangement had occurred between these primers.

With primers pointing outward toward the *EcoRI* sites at the ends of the cloned eccDNA 15 (e.g., 15-SK and 15-KS) only the genomic DNA templates were amplified (Figure 3, lane 6). Similar results were obtained with eccDNA 82 (data not shown). PCR analysis indicated that both of these eccDNA fragments were complete eccDNAs, because the size of the PCR products from the genomic DNA was the sum of the distance between the primers and the *EcoRI* sites in the eccDNA. For example, on eccDNA 15 the distance between primer 15-KS and the 5' *EcoRI* site was 106 bp, and the distance between primer 15-SK and the 3' *EcoRI* site was 52 bp. The predicted size of the PCR product

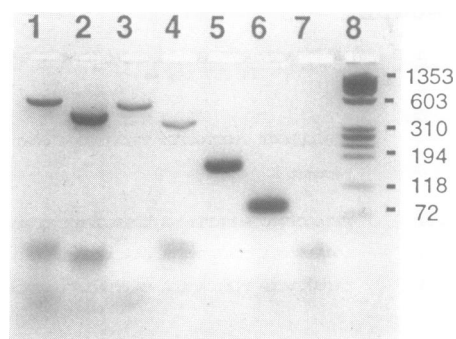


Figure 2. PCR products synthesized with eccDNA 15 as a template. The oligonucleotide primers used were: lane 1, 15-inv-KS and 15-inv-SK; lane 2, 15-inv-KS and 15-inv-SK; lane 3, 15-inv-KSb and 15-inv-SKa; lane 4, 15-inv-KSb and 15-inv-SKa; lane 5, 15-SKa and 15-inv-SK; lane 6, 15-inv-KS and 15-KSb; lane 7, 15-inv-KS; lane 8, ϕ X174 *HaeIII*-digested DNA size markers (indicated at right in bp).

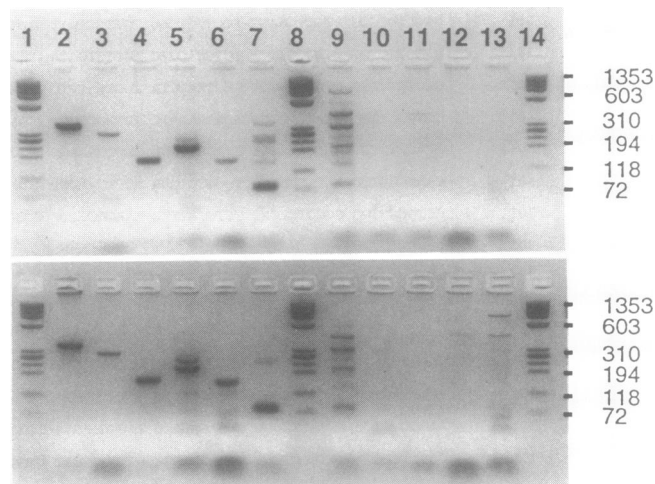


Figure 3. PCR products synthesized with HeLa S3 (upper) or LM217 (lower) genomic DNA as template. The primers used were: lane 2, 15-SKa and 15-KSb; lane 3, 15-SKa and 15-KS; lane 4, 15-SKa and 15-inv-SK; lane 5, 15-SK and 15-KSb; lane 6, 15-SK and 15-KS; lane 7, 15-inv-KS and 15-KSb; lane 9, 15-inv-KS and 15-inv-SKa; lane 10, 15-inv-KSb and 15-inv-SK; lane 11, 15-inv-KSb and 15-inv-SK; lane 12, 15-inv-KSb and 15-KS; lane 13, 15-inv-SKa and 15-SK. Size markers (lanes 1, 8, and 14) are the same as in Figure 2.

synthesized with these primers on the genomic DNA counterpart of eccDNA 15 would therefore be 158 bp, the size that was observed (Figure 3, lane 6). If an incomplete eccDNA had been cloned, the corresponding chromosomal DNA would have contained at least one more *EcoRI* fragment, which would lead to either a larger than expected PCR product or no product at all.

The results with eccDNAs 24 and 80 were much more difficult to interpret. We were unable to obtain PCR products from their corresponding genomic DNA with several sets of primers, suggesting that the recombination events involved in the formation of these eccDNAs were more complicated. This was confirmed by nucleotide sequence analysis (see below).

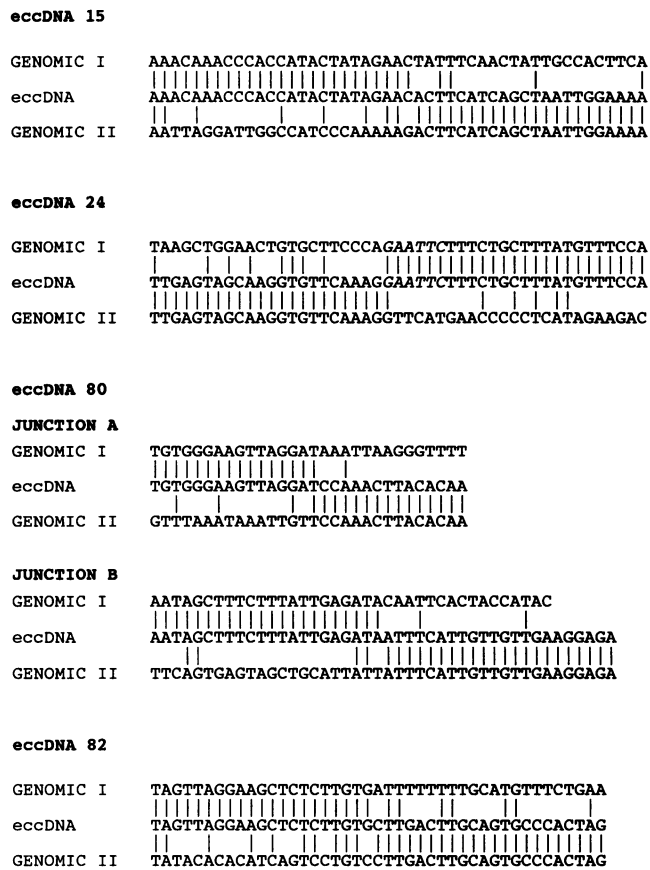


Figure 4. Comparison of nucleotide sequences of the eccDNA and chromosomal DNA at the recombination junctions in eccDNAs 15, 24, 80, and 82. Identical base pairs are indicated by vertical lines. EccDNA 80 is composed of two separate chromosomal fragments and thus has two recombination junctions (A and B). The *EcoRI* site near the junction in eccDNA 24 is indicated by italics.

Isolation and analysis of genomic DNA complementary to the eccDNAs

The chromosomal DNA containing sequences corresponding to four of the eccDNAs (clones 15, 24, 80, and 82) was isolated from a human genomic bacteriophage library, with the use of the purified eccDNA inserts as probes. The low frequency of positive plaques supported the conclusion that these clones were composed of nonrepetitive or low-copy sequences. Each of the genomic clones was digested with various restriction enzymes, and the restriction fragment containing the complete eccDNA sequence was identified by Southern blot analysis, with the use of the corresponding eccDNA as a probe (data not shown). The hybridization patterns obtained with eccDNAs 15 and 82 showed that for most restriction enzymes the probe hybridized to only one restriction fragment. However, with *EcoRI*, two hybridizing fragments were observed; therefore, the restriction patterns were consistent with the conclusion from PCR analysis that these cloned fragments were complete eccDNAs containing single *EcoRI* sites.

As with PCR, Southern blot analysis of eccDNA clones 24 and 80 gave results that were difficult to interpret (data not shown). The *EcoRI* restriction pattern of the genomic clone

homologous to eccDNA 24 had only one fragment that hybridized to the eccDNA probe, suggesting that the recombination junction was at or near the *EcoRI* cloning site, or that the eccDNA contained more than one *EcoRI* site. EccDNA 80 was also unusual in that it hybridized to two separate genomic lambda clones, indicating that it was composed of DNA from two different locations in the genome.

For each of the genomic bacteriophage clones, a restriction fragment hybridizing to the corresponding eccDNA probe (or, for eccDNA 80, fragments from two lambda clones) was subcloned into Bluescript, and the boundaries between the region homologous to the eccDNAs and the surrounding genomic DNA were sequenced with the use of the same primers used in the PCR analysis (Figure 4). This enabled us to compare the eccDNAs with the chromosomal DNA from which they were derived. In eccDNA 15 there was no sequence identity in the two regions that recombined to form the junction (Figure 4), although 2 bp of identity was found 1 bp from the recombination junction; there was also a 5-bp stretch of poly A/T 1 bp from the recombination junction. EccDNA 24 appeared to have a recombination junction with a 1-bp overlap; however, this recombination junction was located at or near the *EcoRI* cloning site. Although no *EcoRI* restriction site was originally found at this location in the genomic DNA, it is possible that a new *EcoRI* restriction site was created during cloning. EccDNA 24 could therefore be part of a larger eccDNA or a fragment of contaminating chromosomal DNA and may not actually contain a recombination junction. EccDNA 80 had two recombination junctions that joined the two chromosomal fragments from which it was formed. Junction A had a 1-bp overlap, whereas junction B had no overlap but had a 2-bp identity 1 bp from the recombination junction. In eccDNA 82 there was no overlap at the junction, but there was 3 bp of identity 1 bp from one side of the recombination junction and 2 bp of identity 1 bp from the other side. There was also an 8-bp stretch of poly A/T 1 bp from the recombination junction.

DISCUSSION

Previous studies have suggested that eccDNAs are generated by recombination between repetitive sequences. This conclusion was based primarily on the presence of interspersed and tandemly repetitive DNA sequences in the eccDNA population (1). One study in HeLa cells suggested that inverted repeat sequences are involved in the formation of eccDNAs, but the chromosomal DNA was not isolated and therefore the recombination junctions were not positively identified (14). In a few cases, the presence of nonrepetitive sequences in eccDNAs has permitted the identification and analysis of the recombination junction. Jones and Potter (11) cloned an eccDNA from HeLa cells that was composed of both a LINE element and unique DNA. The presence of the unique region enabled them to clone the genomic DNA that was complementary to the eccDNA. They found a 9-bp stretch of identity within the genomic DNA at each end of the sequences complementary to the eccDNA and suggested that recombination between these two sequences resulted in the excision of the intervening DNA as a circular molecule. Similarly, Misra *et al.* (6) analyzed an eccDNA containing both nonrepetitive DNA and a THE-1 element and found that recombination occurred within a short, imperfect repeat region. Stanfield and Helinski (18) suggested 'near-homologous

recombination' to explain excision of an eccDNA between imperfect 9-bp direct repeat sequences in a unique region of CHO DNA; however, a second eccDNA analyzed in the same study showed no homology at the site of the recombination junction.

Our results indicate that nonhomologous recombination within nonrepetitive DNA is a much more common mechanism in the generation of eccDNA than was previously thought. A previous study with HeLa cell eccDNAs found that nearly all contained repetitive sequences (14). However, approximately one half of the eccDNA fragments we cloned were composed entirely of nonrepetitive DNA, possibly because of the small size of the fragments that were isolated. Most of these fragments appeared to be complete eccDNAs, as shown by the analysis of the four eccDNAs described here. The five recombination junctions involved in their formation showed no direct or inverted repeat sequences at the site of recombination. In two eccDNAs (24 and 80A), there was a 1-bp overlap between sequences involved in the recombination (Figure 4); however, the recombination junction in eccDNA 24 is questionable because of its proximity to the *EcoRI* cloning site. The three other junctions showed no definitive overlapping sequences, although two could be considered to have imperfect 3-bp overlaps (eccDNAs 15 and 80B), and the third has an imperfect 7-bp overlap (eccDNA 82). The significance of the long stretches of poly A/T 1 bp from the recombination junction in eccDNAs 15 and 82 is unknown.

The lack of obvious homology or repeat sequences at the site of the recombination junctions in the eccDNAs suggests that they were circularized by nonhomologous recombination. Similar junctions formed by nonhomologous recombination have been observed in a number of experimental systems and chromosome rearrangements in mammalian cells (30), including the end rejoining (31) and integration (28) of transfected DNA. Short regions of identity such as those seen in some of the eccDNA junctions described here are a common feature of nonhomologous recombination (31) and are thought to stabilize the ends during rejoining, although they are not required. End rejoining as a mechanism of eccDNA formation would be consistent with the increase in eccDNA in cells treated with various types of stress that increase chromosome breaks (8, 19, 20). The fact that most eccDNAs in this study originated from single chromosome locations suggests that intramolecular end rejoining is favored; however, the presence of one eccDNA that was composed of two separate DNA sequences (eccDNA 80) shows that intermolecular recombination can also occur. This formation of composite eccDNAs would also argue against other possible mechanisms of eccDNA formation, such as replication errors (15), topoisomerase-mediated recombination (24), or reverse transcription (9).

It is unknown what fraction of eccDNA is continuously turning over and what fraction is stably maintained by its ability to replicate owing to the presence of origins of replication. Clearly, double minutes can replicate in mammalian cells (2, 3). However, the difficulty of purifying eccDNA has hampered a more comprehensive analysis of the behavior of most eccDNAs. The use of PCR to detect the recombination junctions for specific eccDNAs provides a sensitive method for detecting their presence in a relatively small number of cells. The PCR primers spanning the recombination junctions of the eccDNAs analyzed in this study mostly failed to amplify DNA from the total genomic DNA isolated from the HeLa S3 cell line from which they were derived (Figure 3; unpublished observation), demonstrating that these eccDNAs are not found in most cells in the population. eccDNAs

may therefore result primarily from spontaneous deletions. Further analysis by PCR should provide additional information on the behavior of individual eccDNAs.

ACKNOWLEDGEMENTS

We thank Mary McKenney for editing the manuscript. This work was supported by the Office of Health and Environmental Research, US Department of Energy, contract DE-A-C03-76-SF01012, and by a National Research Service Award (5 T32 ES07106) from the National Institute of Environmental Health Sciences.

REFERENCES

- Gaubatz, J. W. (1990) *Mutat. Res.* **237**, 271–292.
- Wahl, G. M. (1989) *Cancer Res.* **49**, 1333–1340.
- Hahn, P. J. (1993) *Bioessays* **15**, 477–484.
- Fujimoto, S., Tsuda, T., Toda, M. and Yamagishi, H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2072–2076.
- Misra, R., Shih, A., Rush, M., Wong, E. and Schmid, C.W. (1987) *J. Mol. Biol.* **196**, 233–243.
- Misra, R., Matera, A.G., Schmid, C.W. and Rush, M.G. (1989) *Nucleic Acids Res.* **17**, 8327–8341.
- Iwasato, T., Shimizu, A., Honjo, T. and Yamagishi, H. (1990) *Cell* **62**, 143–149.
- Bertelsen, A. H., Humayun, M. Z., Karfopoulos, S. G. and Rush, M. G. (1982) *Biochemistry* **21**, 2076–2085.
- Krolewski, J. J. and Rush, M. G. (1984) *J. Mol. Biol.* **174**, 31–40.
- Krolewski, J. J., Schindler, C. W. and Rush, M. G. (1984) *J. Mol. Biol.* **174**, 41–54.
- Jones, R. S. and Potter, S. S. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1989–1993.
- Schindler, C. W. and Rush, M. G. (1985) *J. Mol. Biol.* **181**, 161–173.
- Schindler, C. W. and Rush, M. G. (1985) *Nucleic Acids Res.* **13**, 8247–8258.
- Kunisada, T. and Yamagishi, H. (1987) *J. Mol. Biol.* **198**, 557–565.
- Jones, R. S. and Potter, S. S. (1985) *Nucleic Acids Res.* **13**, 1027–1042.
- Kiyama, R., Matsui, H., Okumura, K. and Oishi, M. (1987) *J. Mol. Biol.* **193**, 591–597.
- Okumura, K., Kiyama, R. and Oishi, M. (1987) *Nucleic Acids Res.* **15**, 7477–7489.
- Stanfield, S. W. and Helinski, D. R. (1986) *Nucleic Acids Res.* **14**, 3527–3538.
- Rush, M. G. and Misra, R. (1985) *Plasmid* **14**, 177–191.
- Sunnerhagen, P., Sjöberg, R.-M. and Bjursell, G. (1989) *Somatic Cell Mol. Genet.* **15**, 61–70.
- Kunisada, T., Yamagishi, H., Ogita, Z., Kirakawa, T. and Mitsui, Y. (1985) *Mech. Ageing Dev.* **29**, 89–99.
- Flores, S. C., Sunnerhagen, P., Moore, T. K. and Gaubatz, J.W. (1988) *Nucleic Acids Res.* **16**, 3889–3906.
- Smith, C. A. and Vinograd, J. (1972) *J. Mol. Biol.* **69**, 163–178.
- Kim, R. A. and Wang, J. C. (1989) *Cell* **57**, 975–985.
- McWhinney, C. and Leffak, M. (1990) *Nucleic Acids Res.* **18**, 1233–1242.
- Carroll, S. M., DeRose, M. L., Gaudray, P., Moore, C. M., Needham-Vandevanter, D.R., Von Hoff, D. D. and Wahl, G. M. (1988) *Mol. Cell. Biol.* **8**, 1525–1533.
- Von Hoff, D. D., Forseth, B., Clare, C. N., Hansen, K. L. and Vandevanter, D.R. (1990) *J. Clin. Invest.* **85**, 1887–1895.
- Murnane, J. P., Yezzi, M. J. and Young, B. R. (1990) *Nucleic Acids Res.* **18**, 2733–2738.
- Wittwer, C. T. and Garling, D. J. (1991) *Biotechniques* **10**, 76–83.
- Meuth, M. (1989) In D. E. Berg and M. M. Howe (eds) *Mobile DNA*, Amer. Soc. Molec. Biol., Washington, D. C., 833–860.
- Roth, D. B. and Wilson, J. H. (1986) *Mol. Cell. Biol.* **6**, 4295–4304.