

Corrected 6 April 2012; see below



[www.sciencemag.org/cgi/content/full/science.1213307/DC1](http://www.sciencemag.org/cgi/content/full/science.1213307/DC1)

## Supporting Online Material for

### **Extrachromosomal MicroDNAs and Chromosomal Microdeletions in Normal Tissues**

Yoshiyuki Shibata, Pankaj Kumar, Ryan Layer, Smaranda Willcox, Jeffrey R. Gagan, Jack D. Griffith, Anindya Dutta\*

\*To whom correspondence should be addressed. E-mail: [ad8q@virginia.edu](mailto:ad8q@virginia.edu)

Published 8 March 2012 on *Science Express*  
DOI: 10.1126/science.1213307

#### **This PDF file includes:**

Materials and Methods  
SOM Text  
Figs. S1 to S13  
Tables S1 and S2  
References (20–30)

**Correction:** A reference has been added in the main text, and an erroneous reference formerly in the supporting online material has been removed.

## **SUPPORTING ONLINE MATERIAL**

### **MATERIALS AND METHODS**

**Mice.** C57BL/6 were used for all experiments. All experiments were performed in strict accordance with the University of Virginia Guidelines regarding use of experimental animals.

**Extraction of nuclei from tissues (Fig. S1).** Nuclei were extracted from tissues and cell lines as described (20) with minor modifications. All the reagents used were pre-chilled and the entire procedure was performed on ice. Tissues and cells were homogenized by douncing 50 times in 10 ml of Nuclear Extraction Buffer. After homogenization, samples were layered on 18 ml of Sucrose Cushion in 40 ml ultracentrifuge tubes (Beckman, 25 × 89 mm, 344058) and spun at 30,000 × g for 45 min at 4°C (Beckman, L8-70M, SW28 rotor). Supernatant, including debris, was removed, and 200 µl of chilled 1 × PBS was added to the tube. After 20 min incubation on ice, the nuclei were dissociated by pipetting, and recovered by centrifugation at 500 × g for 5 min.

**MicroDNA library preparation and sequencing (Fig. S1 and Table S1).** Extrachromosomal DNA was prepared from isolated nuclei as described (21) with minor modifications (HiSpeed Plasmid Midi Kit, Qiagen, 12643). Nuclei were suspended in 10 ml Buffer P1 containing 10 ng/µl RNase A and lysed by adding 10 ml Buffer P2. Genomic DNA was precipitated on ice after the addition of 10 ml Buffer P3. The precipitate was removed by QIAfilter Cartridge and the clear lysate was loaded onto an equilibrated HiSpeed Midi Tip. After washing a tip with 20 ml Buffer QC, DNA was eluted with 5 ml Buffer QF. Eluted DNA was precipitated by adding 3.5 ml isopropanol and centrifuged immediately at 15,000 × g for 30 min at 4°C. DNA pellet was washed with 2 ml 70% ethanol and dissolved in 20 µl Buffer EB. The extrachromosomal DNA was treated sequentially with 10 ng/µl RNase A and 200 ng/µl proteinase K (reaction volume: 20 µl/µg DNA), followed by phenol-chloroform extraction and ethanol

precipitation. The fraction was then digested with 0.4 U/ $\mu$ l ATP-dependent DNase (reaction volume: 250  $\mu$ l/ $\mu$ g DNA) (Epicentre, E3101K) to remove linear double-stranded DNA (6). This treatment cycle was repeated twice for some preparations. ATP-dependent DNase-resistant DNA (eccDNA fraction) was purified by QIAquick PCR Purification Kit (Qiagen, 28104) to further remove oligonucleotides and DNA of >10 kb (including mitochondrial DNA). The yield of eccDNA from indicated amounts of starting tissue (or cells) is shown in Table S1. This was the material used for PCR with outward directed primers (Fig. 1a, b) and for electron microscopy (Fig. 1d, e).

The eccDNA was then amplified by REPLI-g Mini Kit using Multiple Displacement Amplification. 20 ng of eccDNA in 10  $\mu$ l Buffer EB was denatured by 10  $\mu$ l Buffer D1. After denaturation has been stopped by 20  $\mu$ l Buffer N1, 60  $\mu$ l master mix containing buffer and DNA polymerase was added. Reaction mixture was incubated at 30°C for 16 hours. After inactivating REPLI-g Mini DNA Polymerase by heating the sample for 3 min at 65°C, the buffer was exchanged using Microcon YM-100 column. The amplified DNA was then recovered with phenol-chloroform extraction and ethanol precipitation. 5  $\mu$ g of Amplified DNA was nebulized and used for paired-end (PE) library construction as described (15). Paired-end high-throughput sequencing was performed according to the manufacturer's protocol (Illumina).

**Chromosomal DNA library preparation and sequencing.** Genomic (chromosomal) DNA was prepared from embryonic mouse brain nuclei by DNazol (Invitrogen, 10978021). The 20 ng chromosomal DNA was amplified by REPLI-g Mini Kit, and 2  $\mu$ g of amplified DNA was used for paired-end (PE) library construction as done for MicroDNA library preparation.

**Outward directed PCR for the confirmation of circular form of DNA (Fig. 1a, b).** The eccDNA fraction was denatured by 0.2M NaOH and 0.2 mM EDTA for 5 min at room temperature. After

addition of 1/10 volume of 3M sodium acetate (pH 5.2), DNA was ethanol precipitated and dissolved in EB buffer. PCR was done with KOD DNA polymerase according to the manufacturer's instructions.

Outward PCR primer set for clone 1

forward: GTGGGTCACCTGACCGGGACTT

reverse: CGGAAGACTACTGTTCCCAGCAACC

expected product size: 113 bp

Inward PCR primer set for clone 1

forward: TTTGAATTGGGTTGCTGGGAACAGT

reverse: AGGCTCCCTGGGAAGCGTAGTTCT

expected product size: 90 bp

Outward PCR primer set for clone 2

forward: GCTTTCAAACAGCCCCATTCCA

reverse: GCCACTTCTACCAGAGTCCTCCTTGC

expected product size: 93 bp

Inward PCR primer set for clone 2

forward: TGGCAATCTGAAACTCCCTTGTC

reverse: GTTTGAAAGCAGCGCCATTCTCCTA

expected product size: 104 bp

**Identification of microDNA by paired-end sequencing of MDA products.** The algorithms for identification of microDNA from paired-end sequencing of MDA products are shown in Fig. S4. All sequence tags were mapped on the reference genome (mm9 for mouse and hg19 for human cell lines) by using the Novoalign software. Table S2 summarizes the number of reads in different libraries. Only those paired tags that could be mapped uniquely and in opposite orientation to each other were

considered for the identification of islands in the "island method". The sequence coverage of each base pair was profiled for each chromosome, and islands (potential circles) delineated where there were two consecutive sequenced bases. In other words an "island" is a contiguous portion of the genome where a cluster of PE-tags were mapped uniquely and in the opposite orientation to its pair. Paired End tags that had one tag mapping uniquely to an island, while the other was unmapped but passed the sequence quality filter, were considered separately for the validation of circles. We generated a database of hypothetical junctional tags that could be created by ligation of the ends of islands to generate circles. The unmapped tags described above were matched against these hypothetical junctional tags. If the mapped tag of a PE read falls in an island and the un-mapped tag matches with a hypothetical junctional tag generated by ligation of the ends of the island, we annotated that island as a circle. This approach underestimates the number of true microDNAs because further sequence coverage is likely to reveal more junctional tags ligating the ends of islands.

The "split read" method extracted the paired end tags where one end was mapped (M) and other end was un-mapped (UM) but had good base quality score. The UM tags were split into two equal halves and mapped individually by using Novoalign software. Those split pairs which were mapped on the same strand (F-F or R-R) with no mismatch were considered for the identification of circles. The split read pair had to flank the mapped tag and be on the opposite strand relative to the mapped tag. In addition, the distance between the mapped tag (M) and at least one part of the split read had to be within the estimated length of the sheared DNA fragments. The mean + 3SD of the distance between paired-end reads mapping normally to genomic DNA gave an estimate of the maximal length of sheared DNA fragments in the library. Thus we discarded all those pairs where the distances between the mapped tag (M) and both parts of the split read were greater than this maximal estimated length of the sheared DNA fragments. The split read method reports the unique circles on the basis of the start and end of the split read positions.

The "Island method" and the "split-read method" both are dependent on the junctional tags for the identification of circles. >80% of the circles were identified by both algorithms. The "Island method" failed to detect circles 1) with low base coverage, 2) with the start and the end of the circle contaminated by reads from contaminating linear DNA or lost due to sequencing errors and 3) that were overlapping with each other. On the other hand the "Split-read method" failed to detect circles if the two split reads do not map uniquely to the genome.

**Generation of random control for microDNAs (Fig. 2d).** To determine if the direct repeats observed at the starts and ends of circles could have arisen by chance, we randomized the co-ordinates of each identified circle in the EMB1 library 100 times to generate 2,492,000 random entities.

**Kleinschmidt method for visualization of dsDNA by EM.** All procedures were performed as described (18). 30-50 ng of eccDNA (after nuclease digestion, but without MDA) was mixed with a buffer containing 250 mM ammonium acetate in a total volume of 43.3  $\mu$ l. Just before spreading, 1.7  $\mu$ l of 200  $\mu$ g/ml cytochrome C solution was added, mixed briefly and the entire volume placed immediately on a piece of parafilm as a drop. After 5 to 10 min incubation, the surface of the drop was touched lightly by a copper grid previously coated with parlodion film. The grid was then dehydrated for 30 sec in 75% ethanol and 30 sec in 90% ethanol. Following quick air drying, the grid was placed in a high vacuum evaporator and a thin layer of platinum (80%): palladium (20%) evaporated on the sample at an angle of 8 degrees at  $2 \times 10^{-6}$  torr. Finally, the grid was covered with a thin layer of carbon to aid in stabilizing the parlodion film. Samples were examined in an FEI T12 TEM and a Philips CM12 TEM equipped with Gatan 2kx2k SC200 CCD cameras at 40 kV. Adobe Photoshop software was used to arrange images into panels for publication.

**Preparation of DNA for visualization of ssDNA by electron microscopy.** 100-150 ng of DNA was incubated with 100 ng T4gp32 protein in 20 mM Hepes pH 7.4, in a 20  $\mu$ l volume, on ice for 15 minutes, followed by fixation of DNA-protein complexes with 0.3% glutaraldehyde for 5 min at room temperature. The sample was then passed over a 2-ml column of 6% agarose beads (ABT inc, Burgos Spain) equilibrated with TE buffer (10 mM Tris-HCl pH 7.4, 0.1 mM EDTA-NaOH) to eliminate unwanted salts and unbound proteins, and fractions enriched for DNA-protein complexes were collected. The DNA was then prepared for EM by the direct mounting procedure.

**Electron microscopy direct mounting method.** All procedures were performed as described (19). Aliquots of the samples containing DNA-T4gp32 complexes were mixed with a buffer containing spermidine and adsorbed onto copper grids coated with a thin carbon film glow-charged shortly before sample application. After adsorption of the samples for 2–3 min, the grids were washed with EM grade water and dehydrated through a graded ethanol series from 25% to 95%, 5 min each. Following quick air drying the grids were rotary shadowcast with tungsten at  $10^{-7}$  torr. Samples were examined in an FEI T12 TEM and a Philips CM12 TEM equipped with Gatan 2kx2k SC200 CCD cameras at 40 kV. Adobe Photoshop software was used to arrange images into panels for publication.

**Paired End-Sequencing of KCNK3 genomic locus.** chr5:30,890,697-30,910,805 region was amplified by LA Taq with GC buffer (Takara, RR02AG) using 670 ng genomic DNA from 6 month old mouse brains as a template. Amplified DNA was nebulized and used for paired-end library construction as described (15). High-throughput sequencing was performed according to the manufacturer's protocol (Illumina).

**Paired End-Sequencing of microDNA-enriched region in chromosome 10.** Anchored ChromPET

(15) was used to capture and sequence chr10:80,213,587-80,372,454 region. Biotin labeled bait RNA library was prepared from BAC clone RP23-74G6. 2 µg Genomic DNA from 6 month old mouse brain was used to construct the sequencing library according to manufacturer's protocol (Illumina). 700 ng of heat-denatured chromPET library was hybridized to 700 ng of biotin-labeled unique single-stranded bait RNA. Captured DNA was amplified by PCR and recaptured on another aliquot of bait RNA. After converting the captured DNA to double-stranded DNA by PCR, high-throughput sequencing was performed according to the manufacturer's protocol (Illumina).

**Identification of deletions in the genome (Fig. 3a).** Two loci were selected on the basis of high density of microDNAs to find deletions in genomic DNA of adult mouse brain. To avoid artifacts specific to an enrichment technique, two different methods were used to enrich the two genomic loci for highthroughput sequencing. Following paired-end highthroughput sequencing of the enriched genomic DNA, the PE-tags were mapped using Novoalign. To find the deletions, each locus was binned into 1000 bp sliding windows with 500 bp overlap. Deletions in the genome would create a new junctional sequence by the ligation of two ends and hence we created a database of possible junctions created by deletions of 1-957 bp in each bin. Since the length of the tag was 43 bp, the junction tag could be any combination of any two fragments which add to 43 bp (for example 10+33, 11+32, 12+31 etc.). All these combinations were generated for each potential break-point. We then searched for PE-tags where one end maps inside the chromosomal bin and the other end is unmappable to normal genomic DNA but maps to a hypothetical deletion junction in that bin.

**Sequences.** All raw sequence data and lists of the microDNA sequences will be deposited to a public database at time of publication.



## **SUPPORTING TEXT**

**Previous evidence of DNA rearrangements and aneuploidy in neurogenesis.** DNA repair proteins are required for neurogenesis during late embryonic development and somatic DNA recombination and aneuploidy has been reported in the brain (22-30) .

**No microDNA in *S. cerevisiae*.** We also tried to identify short eccDNAs in *S.cerevisiae*, but failed to detect them when the protocol described for mouse or human microDNA was performed with the eccDNA fraction from yeast after removing mitochondrial and ribosomal DNA circles by restriction enzyme digestion. This result suggests that the short eccDNAs we observed in mouse and human cells are not artifactually produced from the simple extraction and purification of genomic DNA. The failure to find a significant number of eccDNA with chromosomal genomic DNA from mouse brain supports this assertion.

**Size of circular DNA in the embryonic mouse brain eccDNA fraction.** Both in Kleinschmidt grids (for double stranded DNA) and the T4gp32 stained samples (to see single stranded DNA), the majority of circles are small, <1000 bp. The ratio of microDNA circles to large eccDNA circles is about 50:1.

**MicroDNA copy number estimation.** 1.2 µg eccDNA fraction was purified from  $1 \times 10^9$  HeLa cells (Table S1). Electron microscopy without MDA, indicated 50% of this material is small circular DNA and the other 50% is either bigger circles or linear DNA. The average size of microDNA is 350 bp (MW = 230 kDa). Thus 0.6 µg eccDNA from  $1 \times 10^9$  HeLa cells contained  $1.6 \times 10^{12}$  microDNA molecules. We therefore estimate that on an average, a single HeLa cell yields about 1600 microDNAs.

About 40 ng of extrachromosomal DNA was produced per embryonic mouse brain in three

different library preparations (Table S1). Assuming a 80% contamination of this material with genomic linear DNA, we estimate that an embryonic mouse brain yielded roughly  $3 \times 10^{10}$  microDNAs.

**Properties of islands of chromosomal DNA from ED13.5 mouse brain are very different from microDNAs (Fig. S11).** To rule out the high GC composition of miDNA and their biased distribution in genic region are not an artifact of high throughput sequencing technology bias we did the MDA on genomic DNA isolated from 13.5 days old embryo mouse followed by high throughput sequencing. On the basis of sequencing coverage we selected 10000 most abundantly sequenced genomic regions (genomic islands) to study their GC composition and genomic distribution. The lengths of these negative control chromosomal islands peaked once at 100 bp (Fig. S11a) and the GC content was  $\sim 35\%$ , significantly below the average of the mouse genome (Fig. S11b). In addition these sequences were not enriched in 5'UTRs, CpG islands or genic regions (Fig. S11c). Thus the sequences obtained from the microDNA preparations have very unique features.

### **Frequency of chromosomal microdeletions**

Based on our yield of microdeletions relative to the sequence coverage, the frequency of microdeletions affecting a given locus appears to be  $\sim 1:2000$ . Another way to estimate the frequency of a specific microdeletion is to measure the frequency of the junctional tag created by the deletion relative to the frequency of the wild-type tags that span the deletion breakpoints. By this measure, individual microdeletions occurred at frequencies ranging from 1:300 to 1:4000.

**Properties of small germline deletions reported in the Thousand Genomes Project (Fig. 4).** The length distribution of the deletions of  $< 1000$  bp peak at 100 and 350 bp (Fig. 4a). The subset of these deletions that are in the vicinity of genes (genes and 200 bp upstream) are enriched in 5'UTRs, CpG

islands and exons (Fig. 4b), similar to the microDNAs we have observed. The GC content of the deletions is significantly enriched relative to the genomic average (Fig. 4c). The vast majority of the deletions are flanked by direct repeats  $\geq 2$  bp with one repeat at one end of the deletion interval, and the other in the flanking DNA near the opposite end (Fig. 4d) (16).

### **Abbreviations**

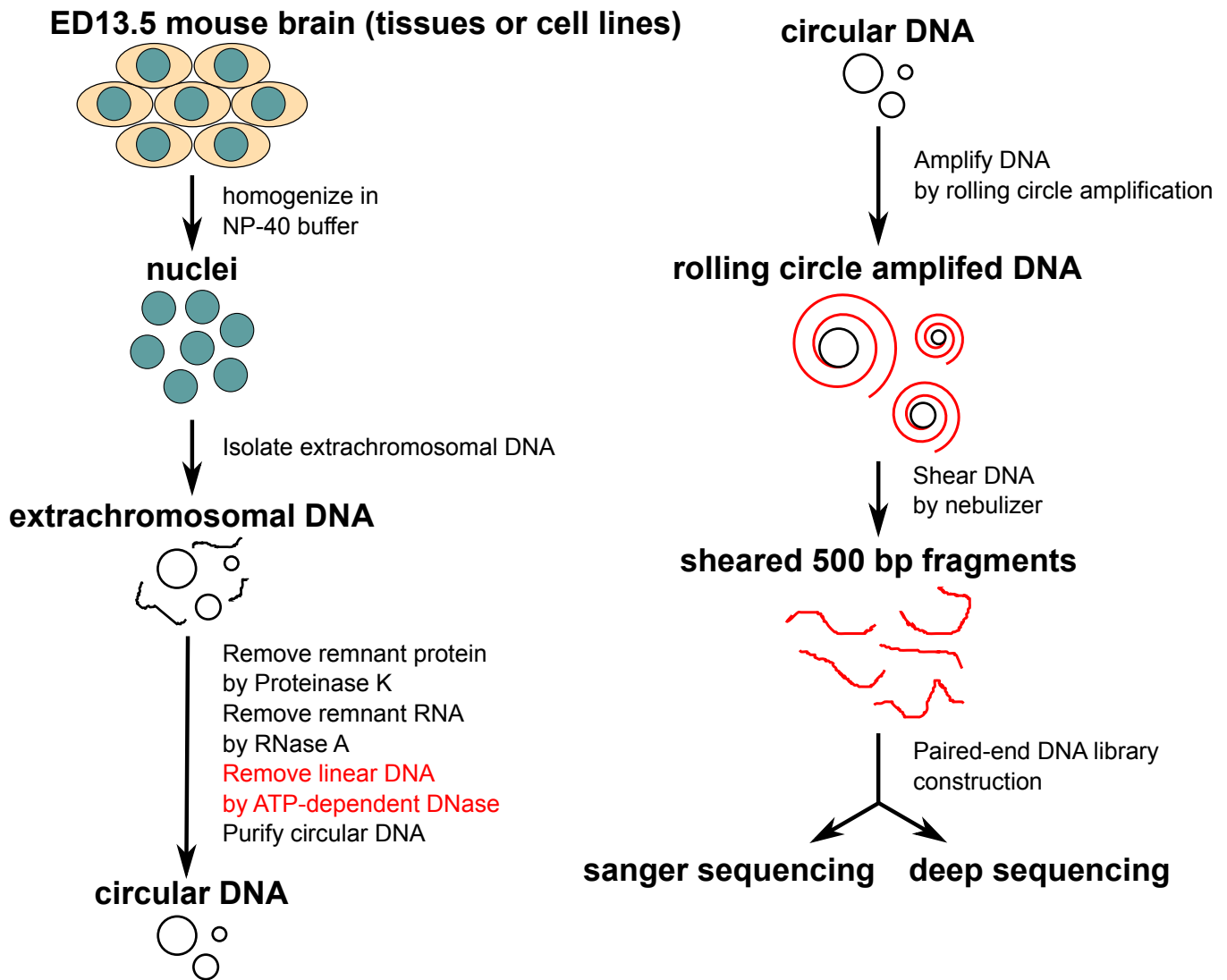
EMB1 = Embryonic Day (ED) 13.5 mouse brain library 1, EMB2= ED13.5 mouse brain library 2, EMB3 = ED13.5 mouse brain library 3, AMB = Adult mouse brain, EMH = ED13.5 mouse heart, EML= ED13.5 mouse liver, NIH3T3 = mouse embryonic fibroblast cell line, HeLaS3 = human cervical adenocarcinoma cell line and U937 = human histiocytic lymphoma cell line, CpG2kbU = sequences 2 kb upstream from the start sites of CpG islands, CpG2kbD = sequences 2 kb downstream from the end sites of CpG islands, Gene2kbU = sequences 2 kb upstream from the start of the gene, Gene2kbD = sequences 2 kb downstream from the end of the gene and Gene200bU = sequences 200bp upstream from the start of the gene.

## References and Notes

1. J. S. Beckmann, X. Estivill, S. E. Antonarakis, Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat. Rev. Genet.* **8**, 639 (2007).  
[doi:10.1038/nrg2149](https://doi.org/10.1038/nrg2149) [Medline](#)
2. M. Flores *et al.*, Recurrent DNA inversion rearrangements in the human genome. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6099 (2007). [doi:10.1073/pnas.0701631104](https://doi.org/10.1073/pnas.0701631104) [Medline](#)
3. K. A. Frazer, S. S. Murray, N. J. Schork, E. J. Topol, Human genetic variation and its contribution to complex traits. *Nat. Rev. Genet.* **10**, 241 (2009). [doi:10.1038/nrg2554](https://doi.org/10.1038/nrg2554) [Medline](#)
4. P. Stankiewicz, J. R. Lupski, Structural variation in the human genome and its role in disease. *Annu. Rev. Med.* **61**, 437 (2010). [doi:10.1146/annurev-med-100708-204735](https://doi.org/10.1146/annurev-med-100708-204735) [Medline](#)
5. J. R. Lupski, New mutations and intellectual function. *Nat. Genet.* **42**, 1036 (2010).  
[doi:10.1038/ng1210-1036](https://doi.org/10.1038/ng1210-1036) [Medline](#)
6. Y. Hideo *et al.*, Purification of small polydisperse circular DNA of eukaryotic cells by use of ATP-dependent deoxyribonuclease. *Gene* **26**, 317 (1983). [doi:10.1016/0378-1119\(83\)90205-6](https://doi.org/10.1016/0378-1119(83)90205-6)  
[Medline](#)
7. F. B. Dean *et al.*, Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5261 (2002). [doi:10.1073/pnas.082089499](https://doi.org/10.1073/pnas.082089499)  
[Medline](#)
8. L. Lovmar, A.-C. Syvänen, Multiple displacement amplification to create a long-lasting source of DNA for genetic studies. *Hum. Mutat.* **27**, 603 (2006). [doi:10.1002/humu.20341](https://doi.org/10.1002/humu.20341) [Medline](#)
9. T. Maeda *et al.*, Somatic DNA recombination yielding circular DNA and deletion of a genomic region in embryonic brain. *Biochem. Biophys. Res. Commun.* **319**, 1117 (2004).  
[doi:10.1016/j.bbrc.2004.05.093](https://doi.org/10.1016/j.bbrc.2004.05.093) [Medline](#)
10. S. Cohen, D. Segal, Extrachromosomal circular DNA in eukaryotes: possible involvement in the plasticity of tandem repeats. *Cytogenet. Genome Res.* **124**, 327 (2009). [doi:10.1159/000218136](https://doi.org/10.1159/000218136)  
[Medline](#)
11. C. A. Smith, J. Vinograd, Small polydisperse circular DNA of HeLa cells. *J. Mol. Biol.* **69**, 163 (1972). [doi:10.1016/0022-2836\(72\)90222-7](https://doi.org/10.1016/0022-2836(72)90222-7) [Medline](#)

12. E. Segal *et al.*, A genomic code for nucleosome positioning. *Nature* **442**, 772 (2006).  
[doi:10.1038/nature04979](https://doi.org/10.1038/nature04979) [Medline](#)
13. E. Segal, J. Widom, What controls nucleosome positions? *Trends Genet.* **25**, 335 (2009).  
[doi:10.1016/j.tig.2009.06.002](https://doi.org/10.1016/j.tig.2009.06.002) [Medline](#)
14. L. N. Changolkar *et al.*, Genome-wide distribution of macroH2A1 histone variants in mouse liver chromatin. *Mol. Cell. Biol.* **30**, 5473 (2010). [doi:10.1128/MCB.00518-10](https://doi.org/10.1128/MCB.00518-10) [Medline](#)
15. Y. Shibata, A. Malhotra, A. Dutta, Detection of DNA fusion junctions for BCR-ABL translocations by Anchored ChromPET. *Genome Med.* **2**, 70 (2010). [doi:10.1186/gm191](https://doi.org/10.1186/gm191) [Medline](#)
16. R. E. Mills *et al.*; 1000 Genomes Project, Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59 (2011). [doi:10.1038/nature09708](https://doi.org/10.1038/nature09708) [Medline](#)
17. S. J. Lolle, J. L. Victor, J. M. Young, R. E. Pruitt, Genome-wide non-mendelian inheritance of extra-genomic information in Arabidopsis. *Nature* **434**, 505 (2005). [doi:10.1038/nature03380](https://doi.org/10.1038/nature03380)  
[Medline](#)
18. R. Thresher, J. Griffith, Electron microscopic visualization of DNA and DNA-protein complexes as adjunct to biochemical studies. *Methods Enzymol.* **211**, 481 (1992). [doi:10.1016/0076-6879\(92\)11026-F](https://doi.org/10.1016/0076-6879(92)11026-F) [Medline](#)
19. J. D. Griffith, G. Christiansen, Electron microscope visualization of chromatin and other DNA-protein complexes. *Annu. Rev. Biophys. Bioeng.* **7**, 19 (1978).  
[doi:10.1146/annurev.bb.07.060178.000315](https://doi.org/10.1146/annurev.bb.07.060178.000315) [Medline](#)
20. Y. Jiang, A. Matevossian, H.-S. Huang, J. Straubhaar, S. Akbarian, Isolation of neuronal chromatin from brain tissue. *BMC Neurosci.* **9**, 42 (2008). [doi:10.1186/1471-2202-9-42](https://doi.org/10.1186/1471-2202-9-42) [Medline](#)
21. N. van Loon, D. Miller, J. P. Murnane, Formation of extrachromosomal circular DNA in HeLa cells by nonhomologous recombination. *Nucleic Acids Res.* **22**, 2447 (1994).  
[doi:10.1093/nar/22.13.2447](https://doi.org/10.1093/nar/22.13.2447) [Medline](#)
22. Y. Gao *et al.*, A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* **95**, 891 (1998). [doi:10.1016/S0092-8674\(00\)81714-6](https://doi.org/10.1016/S0092-8674(00)81714-6) [Medline](#)

23. K. M. Frank *et al.*, DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway. *Mol. Cell* **5**, 993 (2000). [doi:10.1016/S1097-2765\(00\)80264-6](https://doi.org/10.1016/S1097-2765(00)80264-6) [Medline](#)
24. Y. Gu *et al.*, Defective embryonic neurogenesis in Ku-deficient but not DNA-dependent protein kinase catalytic subunit-deficient mice. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2668 (2000). [doi:10.1073/pnas.97.6.2668](https://doi.org/10.1073/pnas.97.6.2668) [Medline](#)
25. S. K. Rehen *et al.*, Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13361 (2001). [doi:10.1073/pnas.231487398](https://doi.org/10.1073/pnas.231487398) [Medline](#)
26. A. H. Yang *et al.*, Chromosome segregation defects contribute to aneuploidy in normal neural progenitor cells. *J. Neurosci.* **23**, 10454 (2003). [Medline](#)
27. Y. B. Yurov *et al.*, The variation of aneuploidy frequency in the developing and adult human brain revealed by an interphase FISH study. *J. Histochem. Cytochem.* **53**, 385 (2005). [doi:10.1369/jhc.4A6430.2005](https://doi.org/10.1369/jhc.4A6430.2005) [Medline](#)
28. S. K. Rehen *et al.*, Constitutional aneuploidy in the normal human brain. *J. Neurosci.* **25**, 2176 (2005). [doi:10.1523/JNEUROSCI.4560-04.2005](https://doi.org/10.1523/JNEUROSCI.4560-04.2005) [Medline](#)
29. B. Mosch *et al.*, Aneuploidy and DNA replication in the normal human brain and Alzheimer's disease. *J. Neurosci.* **27**, 6859 (2007). [doi:10.1523/JNEUROSCI.0379-07.2007](https://doi.org/10.1523/JNEUROSCI.0379-07.2007) [Medline](#)
30. Y. B. Yurov *et al.*, Aneuploidy and confined chromosomal mosaicism in the developing human brain. *PLoS ONE* **2**, e558 (2007). [doi:10.1371/journal.pone.0000558](https://doi.org/10.1371/journal.pone.0000558) [Medline](#)



**Fig. S1**

MicroDNA library preparation. ATP-dependent DNase resistant-DNA from nuclei (eccDNA) was amplified by multiple-displacement amplification (MDA) with random primers. The amplified DNA was sheared to obtain 500 bp fragments and sequenced.

>clone 1

CACCTGACCGGGACTTCGGCGACGCTATGGCGGGCGCCATGCTGCGCTTGGGTTTCCCAAAGTTCCGGCAGCCTTTGAA
TTGGGTTGCTGGGAACAGTAGTCTTCCGGTCAAGAGCTACACGGGTTGTGGCGCATTGCCAGAACTACGCTTCCCAGGG
AGCCTCGCGGCAGAGGGGAGGGTGGGTACCTGACCGGGACTTCGGCGACGCTATGGCGGGCGCCATGCTGCGCTTGGGT
TTCCCAAAGTTCCGGCAGCCTTTGAATTGGGTTGCTGGGAACAGTAGTCTTCCGGTCAAGAGCTACACGGGTTGTGGCG
CATTGCCAGAACTACGCTTCCCAGGGAGCCTCGCGGCAGAGGGAGGGTGGGTACCTGACCGGGACTTCGGCGACGCTA
TGGCGGGCGCCATGCTGCGCTTG

repeated sequences mapped position
START END IDENTITY CHRO STRAND START END
1 58 100.0% 3 + 85378040 85378097
58 242 100.0% 3 + 85377913 85378097
242 418 100.0% 3 + 85377913 85378089

>clone 2

CTGAAACACTCCCTTGTGTCAGCACATCAGCAGGCCAAGGAAGCTTTCAGGGAATGCTCTGAGGGCCTGCACTGTAGGAGA
ATGGCGCTGCTTTCAAACAGCCCCATTTCCAACAGGCAAGAGTGCCTGAGAGTACAGTCTGCCCATAGCCCTCATGCAA
GGAGGACTCTGGTAGAAGTGGCTAGTGGCAATCTGAAACACTCCCTTGTGTCAGCACATCAGCAGGCCAAGGAAGCTTTCA
GGGAATGCTCTGAGGGCCTGCACTGTAGGAGAATGGCGCTGCTTTCAAACAGCCCCATTTCCAACAGGCAAGAGTGCCT
GAGAGTACAGTCTGCCCATAGCCCTCATGCAAGGAGGACTCTGGTAGAAGTGGCTAGTGGCAATCTGAAACACTCCCTT
GTCAGCACATCAGCAGGCCAAGGAAGCTTTCAGGGAATGCTCTGAGGGCCTGCACTGTAGGAGAATGGCGCTGCTTTCA
AACAGCCCCAT

START END IDENTITY CHRO STRAND START END
1 132 100.0% 15 + 38176284 38176415
133 322 100.0% 15 + 38176226 38176415
323 485 100.0% 15 + 38176226 38176388

>clone 3

GGATGCTGATGAACATCATAACATTTTCGCAATGTATTTGTTTGTACCCTGAGGGGATGTGACAGTGGCATCCAAGACAG
AGCTATGGATGCTGATGAACATCATAACATTTTCGCAATGTATTTGTTTGTACCCTGAGGGGATGTGACAGTGGCATCCA
AGACAGAGCTATGGATGCTGATGAACATCATAACATTTTCGCAATGTATTTGTTTGTACCCTGAGGGGATGTGACAGTGG
CATCCAAGACAGAGCTATGGATGCTGATGAACATCATAACATTTTCGCAATGTATTTGTTTGTACCCTGAGGGGATGTGA
CAGTGGCATCCAAGACAGAGCTATGGATGCTGATGAACATCATAACATTTTCGCAATGTATTTGTTTGTACCCTGAGGGG
ATGTGACAGTGGCATCCAAGACAGAGCTATGGATGCTGATGAACATCATAACATTTTCGCAATGTATTTGTTTGTACCCT
GAGGGGATGTGACAGTGGCATCCAAGACAGAGCTATGGATGCTGATGAACATCATAACATTTTCGCAATGTATTTGTTTGT
TACCCTGAGGGGATGTGACAGTGGCATC

START END IDENTITY CHRO STRAND START END
1 29 100.0% 2 + 74790702 74790730
28 114 100.0% 2 + 74790644 74790730
113 199 100.0% 2 + 74790644 74790730
198 284 100.0% 2 + 74790644 74790730
283 369 100.0% 2 + 74790644 74790730
368 454 100.0% 2 + 74790644 74790730
453 529 100.0% 2 + 74790644 74790730
528 571 100.0% 2 + 74790644 74790687

>clone 4

CGGAGCTGACAAGAGAAGAGCTGTCTTCATACCTGCTCCTCCAAAATTCCAACCTGCAGGGGGCTCTCCTGCAGTTTAC
CATGGCTCCTCTGCAAGATGGCCACCACCTGGTTCCACACTGGAGTTCACCAGGTGCTCTTGTGGCCAGCATGGGATC
CGTGGCTTGTCCATTCCGGAGGGACCATGCCCTGTGAGGAGGCATTGACCACCCCCACCCCATCCAGGTTACCCTTT
CCCTGGGAGGGCAGAGCTGACAAGAGAAGAGCTGTCTTCAAACCTGCTCCTCCGAAATTCCAACCTGCAGGGGGCTCTCC
TGCATTTTACCATGGCTCCTCTGCAAGATGGCCACCACCTGGTTCCACACTGGAGTTCACCAGGTGCTCTTGTGGCC



AGCATGGGATCCGTGGCTTGTCCATTCCGGAGGGACCATGCCCTGTGAGGAGGCACTTGACCACCCCCACCCCCATCCA  
GGTTACCCTTTCCCTGGGAGGGCAAAGCTGACAAGAGAAGAGCTGTCTTCAGACCTGCTCCTCCAAAATTCCAAC

START	END	IDENTITY	CHRO	STRAND	START	END
1	92	97.9%	3	+	89469427	89469518
93	340	98.8%	3	+	89469271	89469518
341	549	99.6%	3	+	89469271	89469479

>clone 5

GTGGGAGGGTCTCCTCTTATGGCTGTCTTCAGCCATGAGGAGTGAGGTCTCGTGGTGGGGTGGGGCCCTGGAGAGATG  
GCTCAGTGGTTAAAAGCACTGGCTGCTCTTCCAGAGGTCTTGAGTTCAATTTCCAGGAACCACATGGTGGCTCACAAAC  
TCTGTAATGGGATCCGATGCCCTCTTCTGGTGTGTCTGAGGACGTTAGGCCATTTCCCTTCCCATATAAAGTAGATGAT  
CAGGTACAATGCCAGGTCCCTGCCAGTGGGAGGGTCTCCTCTTATGGCTGTCTTCAGCCATGAGGAGTGAGGTCTCGTG  
GTGGGGTGGGGCCTGGAGAGATGGCTCAGTGGTTAAAAGCACTGGCTGCTCTTCCAGAGGTCTTGAGTTCAATTTCC  
AGGAACCACATGGTGGCTCACAACTCTGTAATGGGATCCGATGCCCTCTTCTGGTGTGTCTGAGGACGTTAGGCCATT  
TCCCTTCCCATATAAAGTAGATGATCAGGTACAATGCCAGGTCCCTGCCAGTGGGAGGGTCTCCTCTTATGGCTGTCTT  
CAGCCATGAGGAGTGAGGT

START	END	IDENTITY	CHRO	STRAND	START	END
1	39	100.0%	9	+	44821158	44821196
38	301	100.0%	9	+	44820933	44821196
300	563	100.0%	9	+	44820933	44821196

>clone 6

CCCAGCCTAGACCTAGGACATATTAGGAAGCTTCCCTGGCCTTCTGGTGGGCAGGGTATGTTGCCCTGAGACACACCCA  
ACCCCCTTCTGGCTTGGGACTCAGCAGGACTGAATGAGGCTTCCACTTGCAGACTTTCCCCTAGGCTGAGGATCCTAGC  
CTGTCTCGGGTAGACATCCGGCCAGGGTGCCTAGAAAAGTCTATGTGCAGGTTCCCTGGGTACCCAGTCAATGCCCC  
AGCCTAGACCTAGGACATATTAGGAAGCTTCCCTGGCCTTCTGGTGGGCAGGGTATGTTGCCCTGAGACACACCCAAC  
CCCCTTCTGGCTTGGGACTCAGCAGGACTGAATGAGGCTTCCACTTGCAGACTTTCCCCTAGGCTGAGGATCCTAGCCTG  
TCTCGGGTAGACATCCAGCCAGGGTGCCTAGAAAAGTCTATGTGCAGGTTCCCTGGGTACCCAGTCAATGCCCCAGC  
CTAGACCTAGGACATATTAGGAAGCTTCCCT

STAR	END	IDENTITY	CHRO	STRAND	START	END
1	85	100.0%	5	+	30830261	30830345
82	319	99.6%	5	+	30830108	30830345
316	505	100.0%	5	+	30830108	30830297

>clone 7

AGGAGGAGGAGGAGGACGACGACGACGAGGACGGCGGGCGGGCCACCACGGCGGGCGGGGGCGGCCACGGGCACGC  
CGGCCACCACCATCACCACCACCACCACCACCACCACCACCACCAGCCATGATCGCGCTGCAGCCGCTGGTGACGGACGAC  
CCGACCCAAGTGCACCACCACCAGGAGGTGATCCTGGTGCAGACGCGCGAGGAGGTGGTTCGGCGGGGACGACTCGGACG  
GGCTGCGCGCCGAGGACGGCTTTCGAGGACCAGATCCTCATCCCGGTGCCCGCGCCGGCCGGCGGGCAGCAGACTACAT  
AGAGCAGACGCTGGTCACCGTGGCGGGCGGCCGCAAGAGCGGCGGGCGGGCCTCGTCCGGCGGCATGAGATCGAGGT  
GGAGACCATCCCGGTGGAGACCATCGAGACCAGGTGGTGGGCGAGGAGGAGGAGGAGGACGACGACGACGAGGACGGC  
GGCGGGCGGACACGGCGGGCGGGCGGGGGCGGCCACGGGCACGCCGCCA

START	END	IDENTITY	CHRO	STRAND	START	END
1	381	100.0%	12	+	110031750	110032130
381	524	100.0%	12	+	110031691	110031834

>clone 8

TGGTGCTGTAGAAAAGTGGCCAGGCTGATGGTTGTGGAGGAGCCTGTGAACTTGATTAAGTGCACAGGGCCTTGATTT

CCCTACTTGCAAAGGAGTAGGGAAATATGAGCCTTGGCTCTGCCTACCTCACGTTGCTGGTGCTGTAGAAAAGTGGCCA  
 GGCTGATGGTTGTGGAGGAGCCTGTGAACTTGATTAAGTGC**CACAGGGCCTTGATTTCCCTACTTGCAAAGGAGTAGGG**  
**AAATATGAGCCTTGGCTCTGCCTACCTCACGTTGCTGGTGCTGTAGAAAAGTGGCCAGGCTGATGGTTGTGGAGGAGCC**  
**TGTGAACTTGATTAAGTGCACAGGGCCTTGATTTCCCTACTTGCAAAGGAGTAGGGAAATATGAGCCTTGGCTCTGCC**  
**TACCTCACGTTGCTGGTGCTG**

START	END	IDENTITY	CHRO	STRAND	START	END
1	64	100.0%	18	+	61078141	61078204
64	200	100.0%	18	+	61078068	61078204
200	336	100.0%	18	+	61078068	61078204
336	416	100.0%	18	+	61078068	61078148

>clone 9

**A**CA**C**AGAGCCCCAGCAGGTCATTCATGGGCTAGAAGTTGGTGTGGCTAACTCACAGCCCTGGATTTGGGGCCTGAGAAGT  
 ATCTGAGCTGGTCAGCCAGCCCGCTCTCCACATCCATAACCAGGGTGAGCTCTCCTGCACCTGTCCCAGCTAACTCA  
 TCCACCCGGTGCCCTAGCTGGCAAGGGACAGAAACAGCTCTATCACTCTTCTACCCTTGGGGCTTGCTCACTGTAACAG  
 GGCCAGCTTTGCTGTGCTGCCAGGCAAGGTGCAGGGCCTGCAGCAGGTGAGGGGCAGGGCCCGCTCTCCTGCTCTCAT  
 GACCCAGGGCCAGATCTCCACCTGTTGCAGGTAGCAGGACGGTGGACATCTCCCATTTGCCACACTGCCACAGGGC  
 AGACAAGCGA**CA**CAGAGCCCCAGCAGGTCATTCATGGGCTAGAAGTTGGTGTGGCTAACTCACAGCCCTGGATTTGGGGC  
 CTGAGAAGTATCTGAGCTGGTCAGCCAGCCCGCTCT

START	END	IDENTITY	CHRO	STRAND	START	END
2	407	100.0%	2	+	168285200	168285605
406	510	100.0%	2	+	168285200	168285304

>clone 10

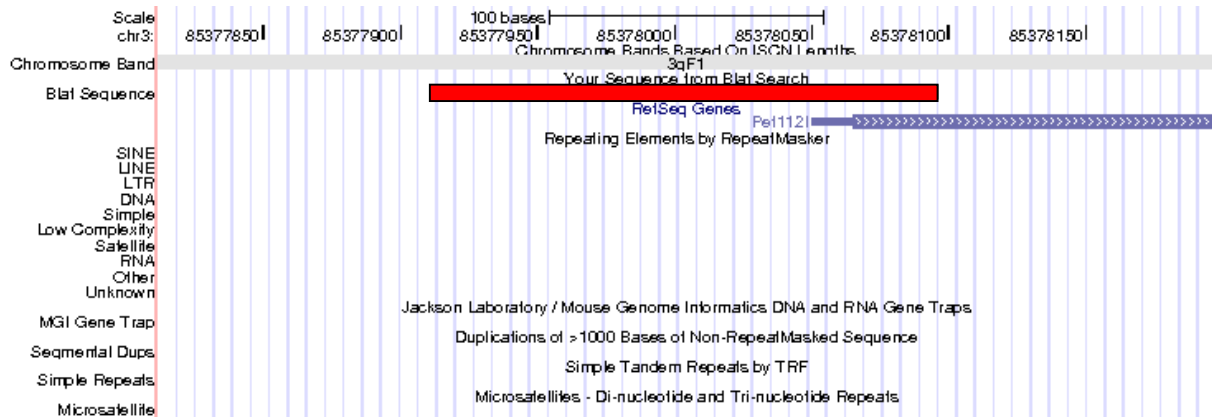
CTGTGGGTGGCCAGAAGCTTCTGGTCCTGAAGACACAGCTCCATCCTCACCTGGACCTATGACCCGGGCCTCGCTCCC  
 CTCTGCCATTG**AG**ATGGACAGACAGAGAGACAGGAGCCTCAGGCGCAATGGAGGCTCTTGCTGTGTCTCTTTCTTCA  
 TTGATTTTTAGGATCATTAGCTAGAGAGTCTAGCCAGCATCCCTCGGTTACATGTTTTAGCCAGGTTCTCAATCAAG  
 GCTGTCTCTCATTTCAGGGCTTGGCTGAGTAGGGTTTTAGACACCCAGCAGCAGCGTTGAGAGACACCTTCCCTGTG  
 GGTGGCCAGAAGCTTCTGGTCCTGAAGACACAGCTCCATCCTCACCTGGACCTATGACCCGGGCCTCGCTCCCCTCCT  
 GCCATTG**AG**ATGGACAGACAGAGAGACAGGAGCCTCAGGCGCAATGGAGGCTCTTGCTGTGTCTCTTTCTTCAATGAT  
 TTTAGGATCATTAGCTAGAGAGTCTAGCCAG

START	END	IDENTITY	CHRO	STRAND	START	END
1	93	100.0%	2	-	101612416	101612508
93	404	99.7%	2	-	101612416	101612727
404	507	100.0%	2	-	101612624	101612727

**Fig. S2**

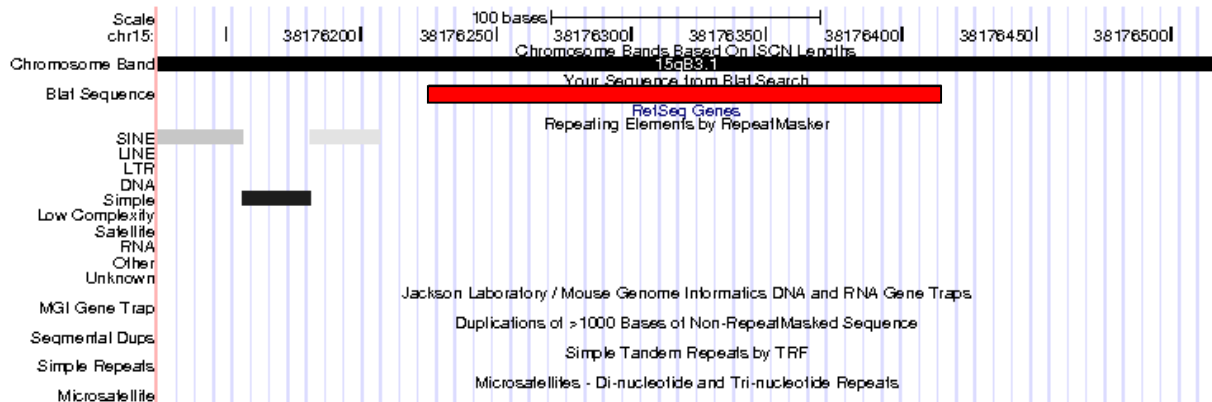
Sanger sequencing of cloned DNA fragments. The fragments contained repeated sequences (different color segments in the ten clones shown) consistent with rolling circle amplification of circular DNA. Black bold letters in sequences are direct repeats of micro-homology which were found at the starts and ends of the circles. Each segment was mapped to same chromosome, direction and position.

# clone1



CACACAAGCTCTACCGCCACAAAGCTCTCCAACCTCCCCGCTAGGAGGAGTCACACAGGAC  
 CCACCAACCTTCCCACCTTCTTTGGTTCCGCCCTCCCTAACAAAGTTCGGCAGCCTTTG  
 AATTGGGTTGCTGGGAACAGTAGTCTTCCGGTCAAGAGCTACACGGGTGTGGCGCATTG  
 CCAGAACTACGCTTCCCAGGGAGCCTCGCGGCAGAGGGAGGGTGGGTCACCTGACCGGGA  
 CTTGGCGACGCTATGGCGGCGCCCATGCTGCGCTTGGGTTTCCCTGGAAGACGCTGGGC  
 TTTAACATGGCTAGATGGCGGTTCCCGCCACCGAAGTGGCTCTCAGACTGGGCCACATC  
 CAACTGGGCTAGGAGACAGAGTTC

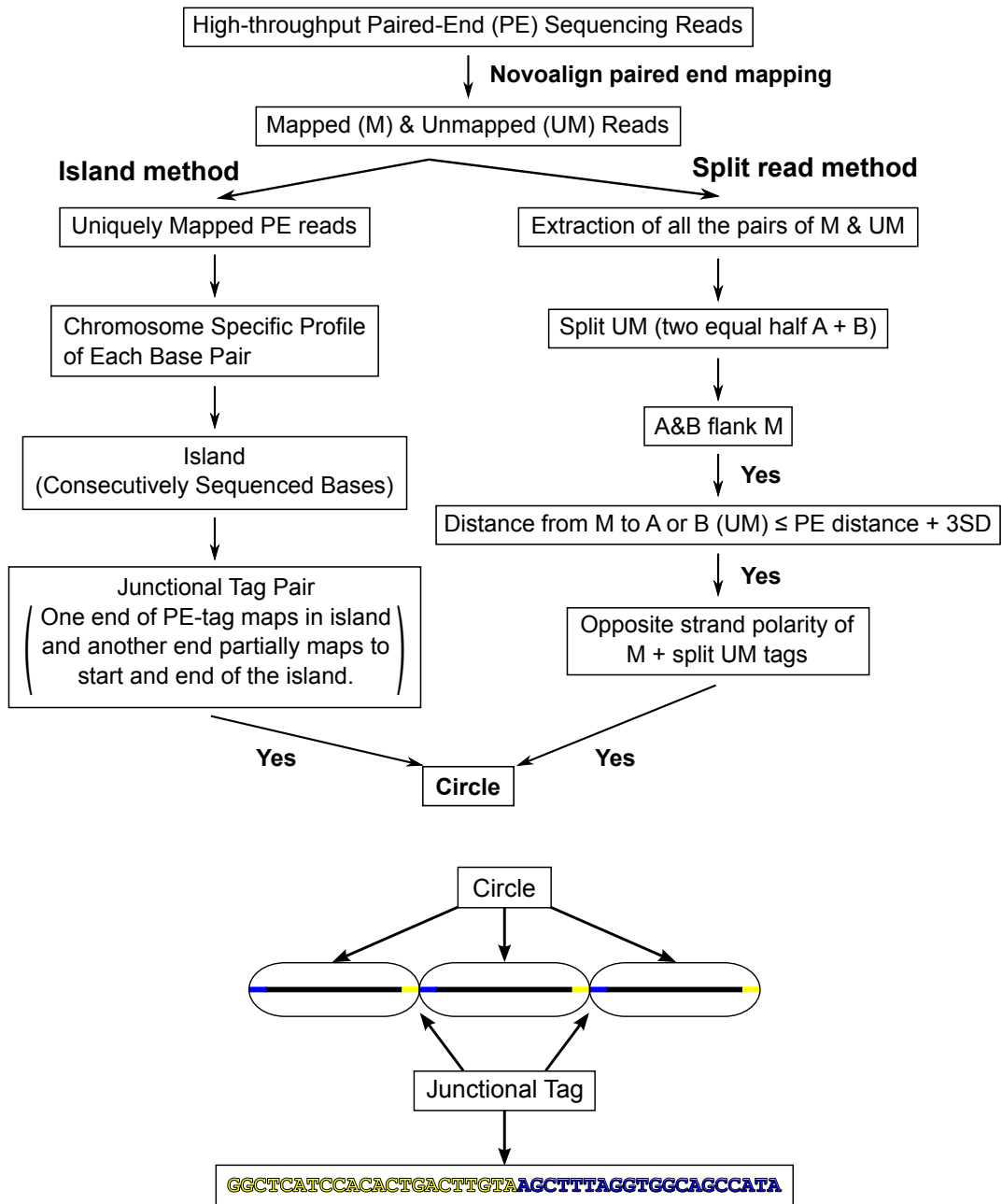
# clone 2



AAAAGGACCTGAATTCCCCAGGACATTGTTGTTGTTTTTTGTTTGTGTTTGTGTTTCAAG  
 CTGGAGCGGTGACACACATGTGGCTGAAAGACCAAGGAACACGTCTGCCCATAGCCCTC  
 ATGCAAGGAGGACTCTGGTAGAAGTGGCTAGTGGCAATCTGAAACACTCCCTTGTCAGCA  
 CATCAGCAGGCCAAGGAAGCTTTCAGGGAATGCTCTGAGGGCCTGCACTGTAGGAGAATG  
 GCGCTGCTTTCAAACAGCCCCATTTCCAACAGGCAAGAGTGCCTGAGAGTGCCTCCTTTC  
 TACCAGGGGCCAGAGGGTTAGGTAGTGTGTCTGCCAAGCAAGCCAGAACCATAGGATTC  
 AAGCAAGGCCTTTTCTGCCACCTATCACC

**Fig. S3**

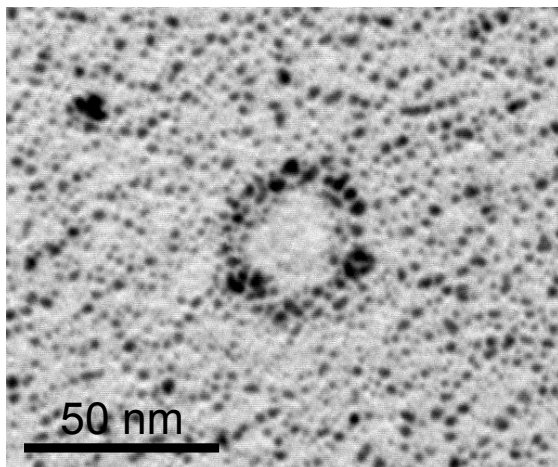
Clones 1 and 2 (Fig. S2) were mapped to the mouse genome in the UCSC genome browser. Only one copy of the direct repeats observed in clone 1 or 2 was uniquely mapped to chr3:85,377,913-85,378,137 or chr15:38,176,226-38,176,415, respectively (red bars and red sequences). The sequences were not repeated in the genomic locus, suggesting that the direct repeats observed in Fig. S2 were generated by rolling-circle amplification of circular DNA molecules containing at least one copy of each sequence.



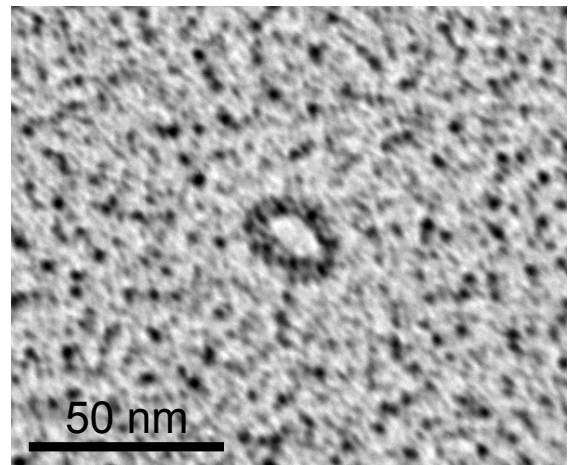
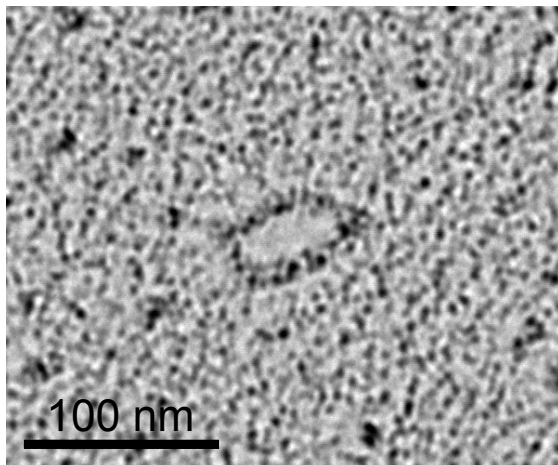
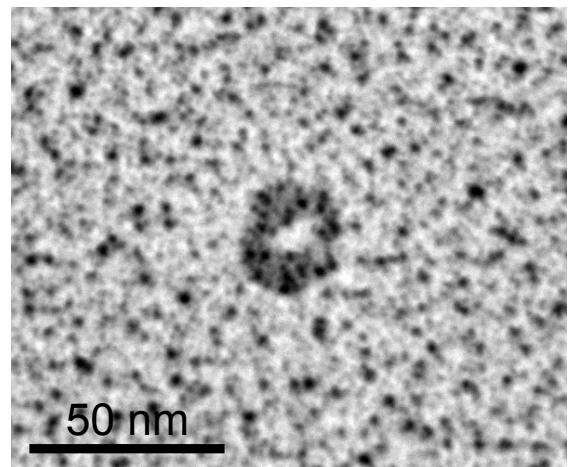
**Fig. S4**

Steps in the identification of microDNA and schematic representation of junctional tag (see methods in SOM for more details).

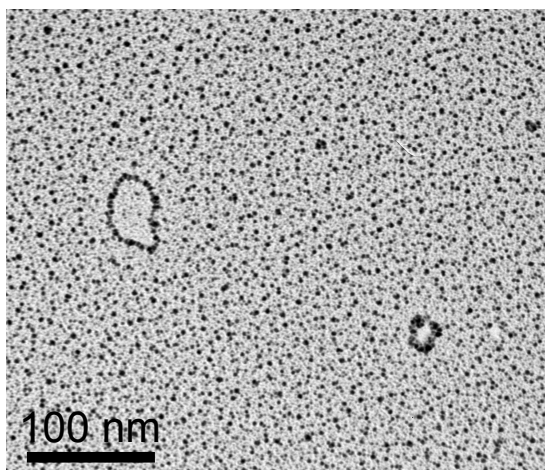
**a** double-stranded



single-stranded

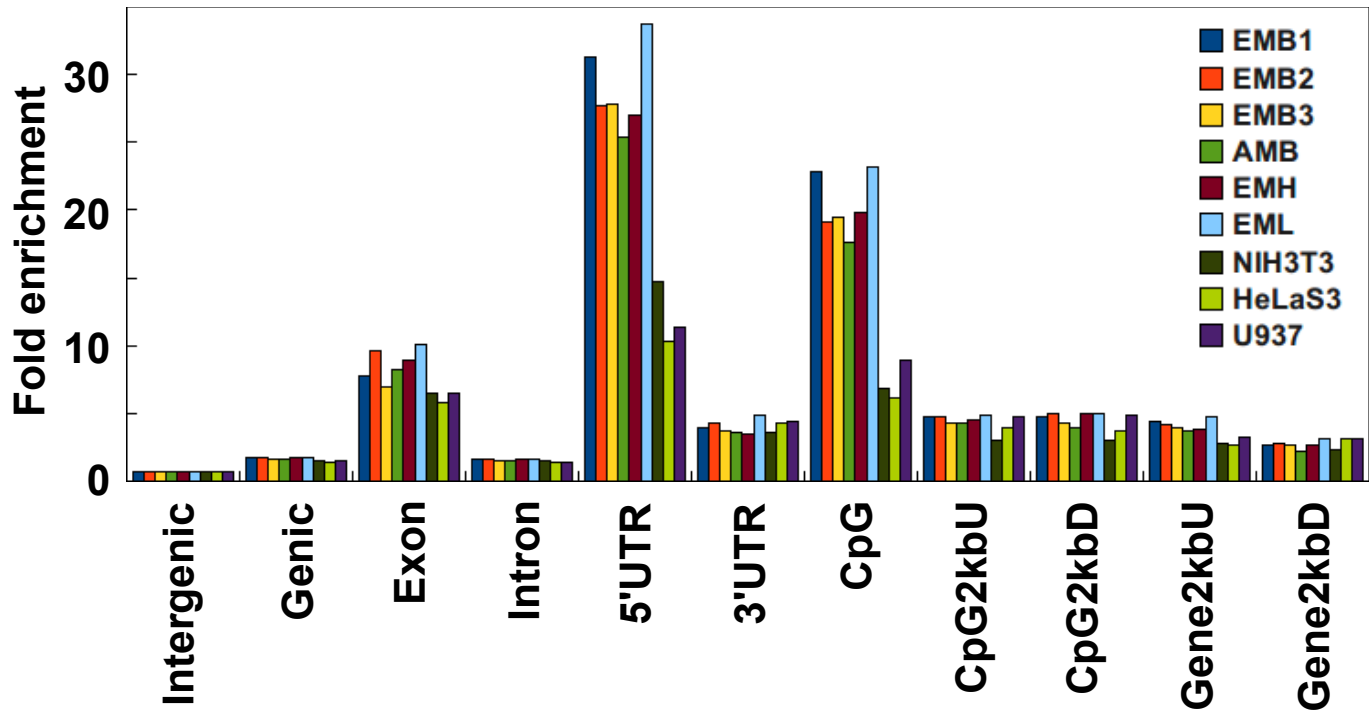


**b**



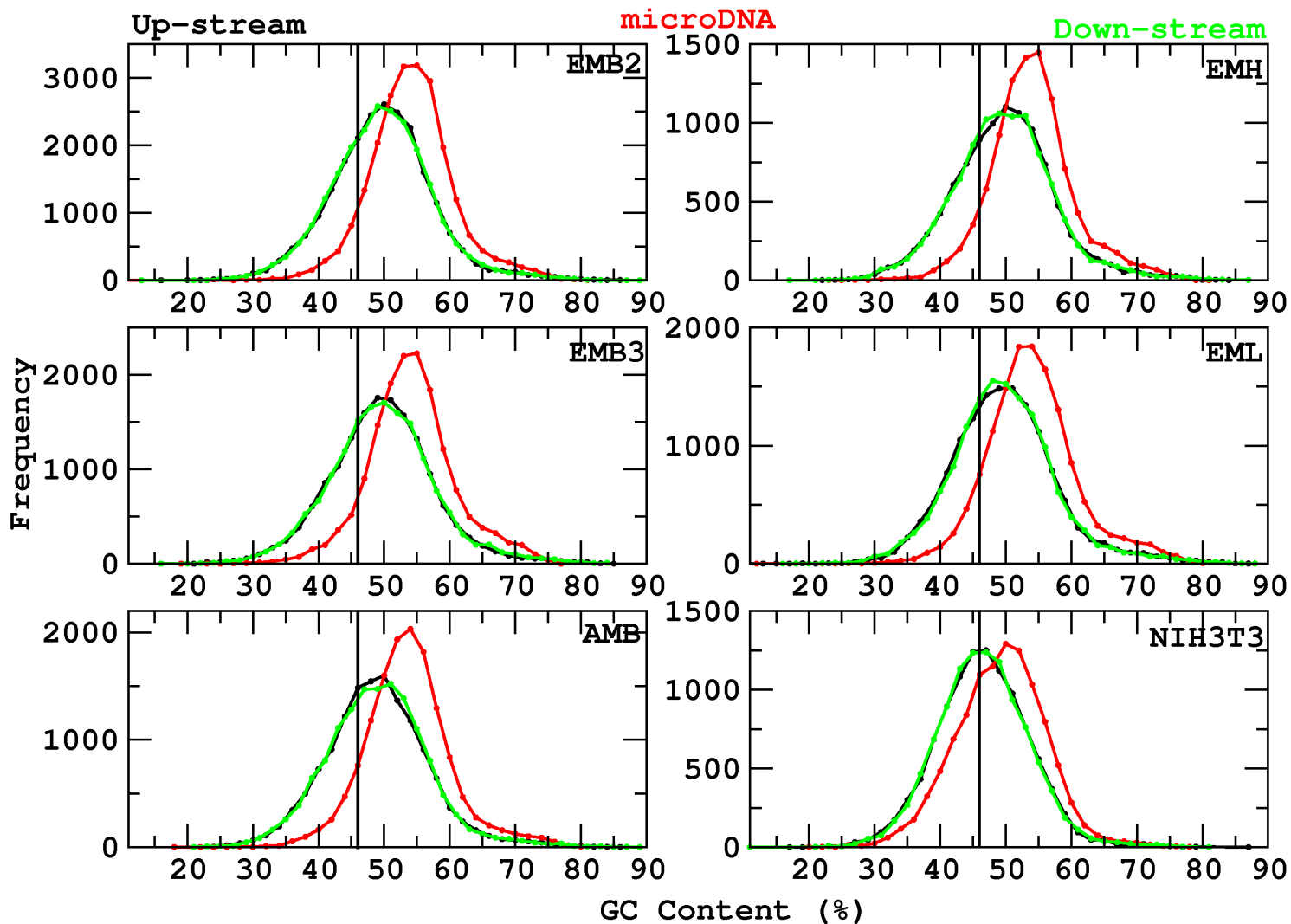
**Fig. S5**

a. Electron microscopic images of additional double-stranded and single-stranded microDNAs from ED13.5 mouse brain. b. EM of double-stranded (left circle) and single-stranded (right circle) microDNA from HeLa-S3. Rest as in Fig. 1d and e.



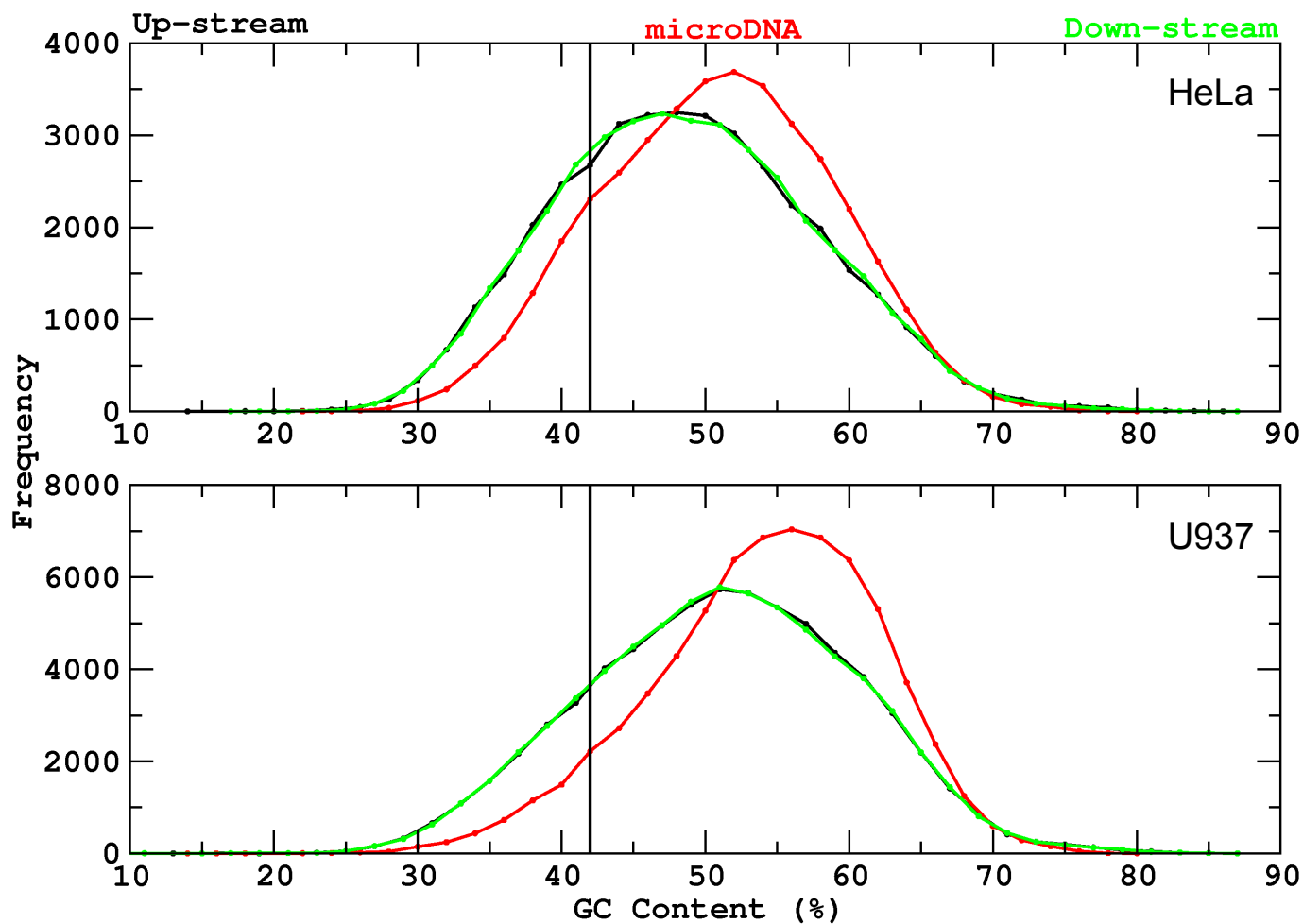
**Fig. S6**

Genomic sources of microDNAs of various tissues and cell lines: fold enrichment relative to random expectation. MicroDNAs are significantly enriched in all the regions except intergenic regions of the genome. The 5'UTR, exon and CpG islands are highly enriched. Gene2kbU or Gene2kbD stand for 2 kb up- or downstream of the starts and ends of genes. CpG2kbU or CpG2kbD are 2 kb up- or downstream from starts and ends of CpG islands.



**Fig. S7**

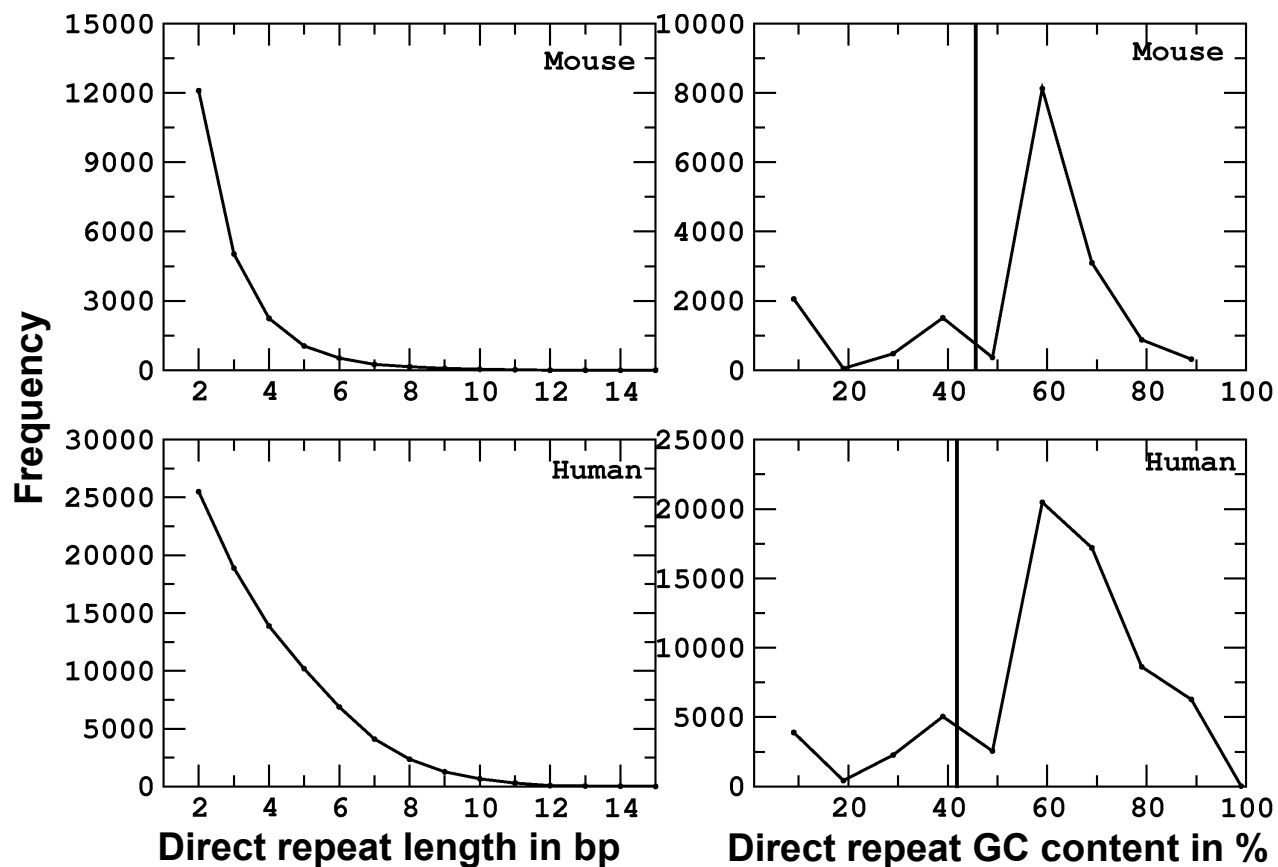
GC composition of microDNAs as well as the up- and downstream flanking genomic regions for microDNAs from various mouse tissues and mouse cell lines. The perpendicular black line is for average GC content of the mouse genome.



**Fig. S8**

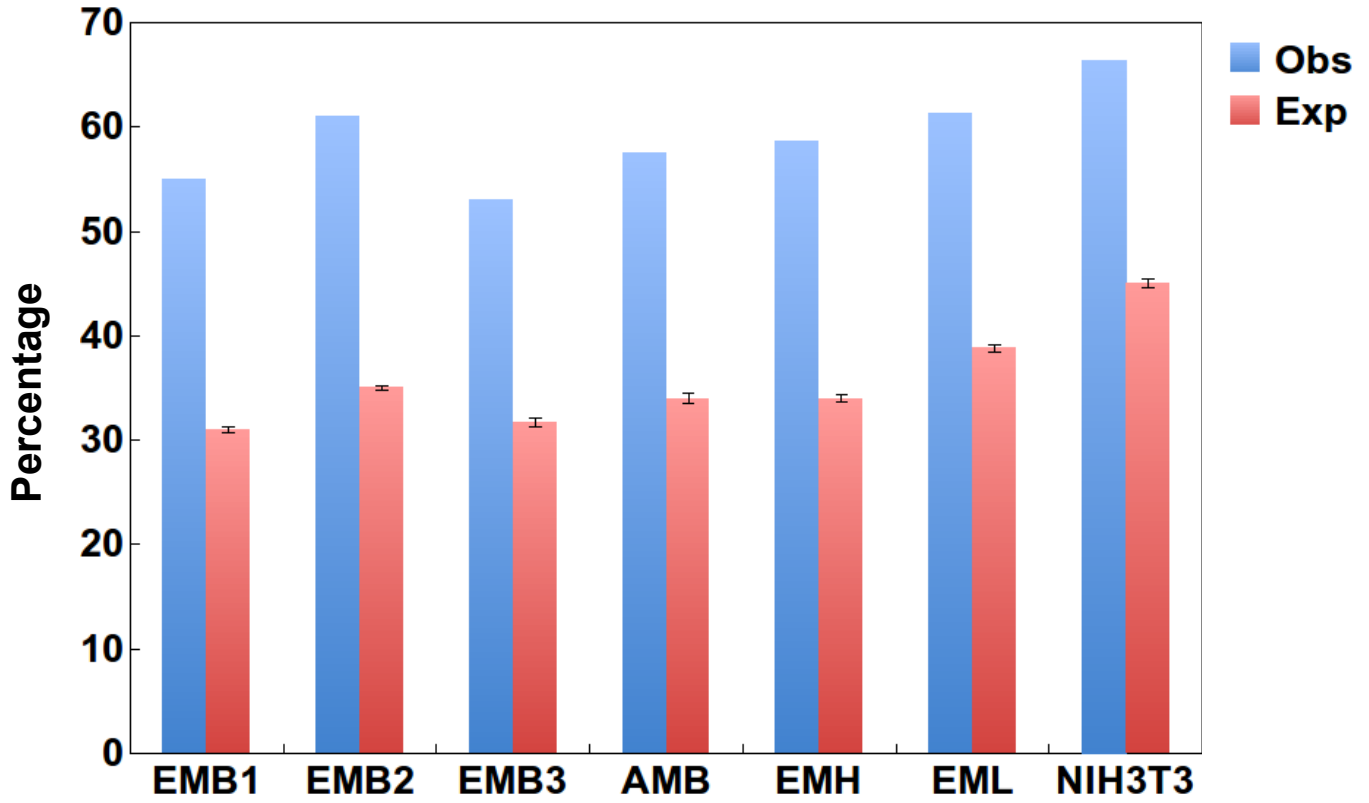
GC composition of microDNAs as well as the up- and downstream flanking genomic regions for microDNAs from two human cell lines. The perpendicular black line is for average GC content of the human genome.





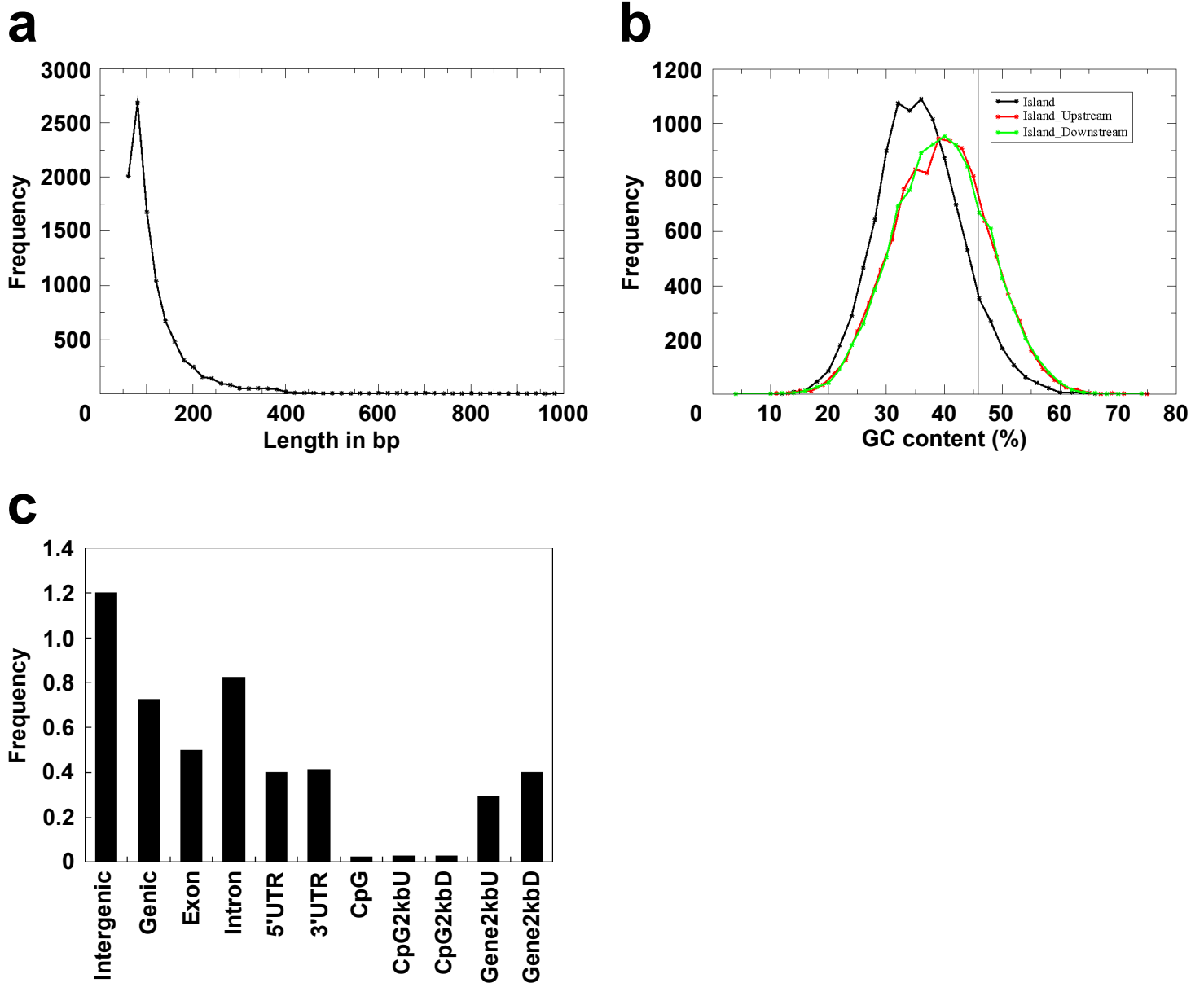
**Fig. S9**

Length distribution and GC content of direct repeats at the starts and ends of microDNAs in mouse and human samples. The sequences around the starts and ends of the microDNAs did not contain any conserved sequence, any specific change in GC/AT composition or any family of repetitive sequences. However the starts and ends of the circles revealed direct repeats of micro-homology with high GC content. The observed direct repeats are 2-15bp long and direct repeats are also rich in GC content.



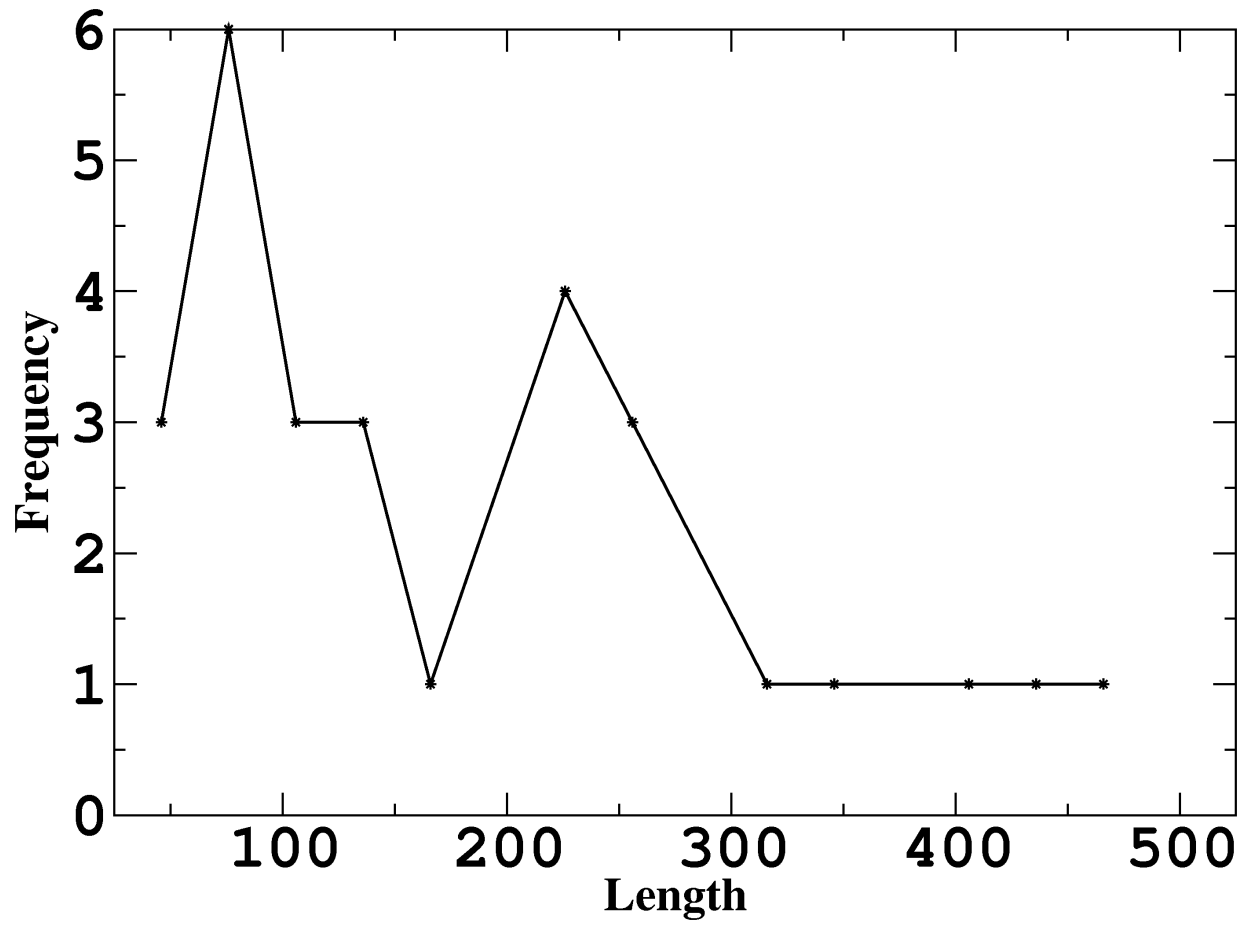
**Fig. S10**

MicroDNAs from different tissues and cell lines are enriched for nucleosome-occupied regions of the genome. Rest as in Fig. 2e.



**Fig. S11**

Genomic DNA was subjected to paired-end sequencing as a negative control to analyze the properties of the 10,000 islands with the highest frequency of mapped tags. The features of these genomic islands were completely different from the features of microDNA, suggesting that the features identified in microDNA are not an artifact of MDA or highthroughput sequencing. **a.** Length distribution shows no peak at ~200 and ~400 bp. **b.** GC content is below the genomic average. **c.** There is no significant enrichment of the genomic islands in different annotated portions of the genome relative to random expectation.



**Fig. S12**  
Length distribution of microdeletions identified in adult mouse brain.

1

chr5 30891274 (Mapped tag)GGCAGAATGTGCGCACGTTGGCTCCTTCTACTTCGCCATCACC

Junctional tag: GGCAGAATGTGCGCACGTTGGCTCCTTCTACTTCGCCATCACC

30890697 30890956:

GGCAGAATGTGCGCACGTTGGCTCCTC**AT**CGTGTGCACC**TT**CACCTACCTGCTGGTGGGCGCCGCGGTG**T**  
**T**CGACGCACTGGAGTCGGAGCCGGAG**AT**GATCGAGCGGCAGCGGCTGGAGCTGCGGCAGCTGGAGCTGC  
GGGCGCGCTAC**AA**CCTCAGCGAGGGCGGC**TA**CGAGGAGCTGGAGCGCGTCTGTGCTGCGCCTC**AA**GCCGC  
AC**AA**GGCCGGCGTGCAGTGGCGC**TT**CGCCGG**CTCCTTCTACTTCGCCATCACC**

2

chr5 30891083 (Mapped tag)GGCAGAATGTGCGCACGTTGGGCGCTTCGCCGGCTCCTTCTAC

Junctional tag: GGCAGAATGTGCGCACGTTGGGCGCTTCGCCGGCTCCTTCTAC

30890697-30890944:

GGCAGAATGTGCGCACGTTGGCTCCTC**AT**CGTGTGCACC**TT**CACCTACCTGCTGGTGGGCGCCGCGGTG**T**  
**T**CGACGCACTGGAGTCGGAGCCGGAG**AT**GATCGAGCGGCAGCGGCTGGAGCTGCGGCAGCTGGAGCTGC  
GGGCGCGCTAC**AA**CCTCAGCGAGGGCGGC**TA**CGAGGAGCTGGAGCGCGTCTGTGCTGCGCCTC**AA**GCCGC  
AC**AA**GGCCGGCGTGCAGTGCAGTGC**CGCTTCGCCGGCTCCTTCTAC**

3

chr5 30891659 (Mapped tag)CGGGAGGAGGGGTGTGGCCGAGAGCGCCACTCCTCCGTGTCC

Junctional tag: CGGGAGGAGGGGTGTGGCCGG\*AGAGCGCCACTCCTCCGTGTCC

30891244-30891326:

CGGGAGGAGGGGTGTGGCCGGC**TA**GGCCCTGGGGCTGGGC**AT**CCAGAGT**AA**GGGCGGAG**AGAGCGCC**  
**ACTCCTCCGTGTCC**

4

chr5 30891947 (Mapped tag)AGGGGGGTGGGTTGCGTGTGTCTGTTAGCCTTCTTCCCCAGT

Junctional tag: AGGGGGGTGGGTTGCGTGTGTCTGTTAGCCTTCTTCCCCAGT

30891453-30891615:

AGGGGGGTGGGTTGCGTGTGT**CC**GTGTGGGTGTGTGAGTGTGAGTGTGAGTGTGTGTCCACCTCCATGA  
G**TTT**GGGCTGCACAGACAGGGCCTGTAA**TT**CTCTGAGGGCAGCAG**AT**CAGGGGCTGGA**AT**CCAG**ATT**CT  
**GTCTT**GTTAGCCTTCTTCCCCAGT

5

chr5 30893082 (Mapped tag)GCCCCATCCCCCTTACCAAAGCCACCCCCTCACTAAGCCTC

Junctional tag: GCCCCATCCCCCTTACCAAAGCCACCCCCTCACTAAGCCTC

30892623-30892863:

GCCCCATCCCCCTTACCAAAG**CC**CGCACAGCTCCCCACTGCCCC**AT**GCTCCTCACCCCGACTCCACCT  
CC**AT**GCCTGACTTTA**TTT**ACTTACTTGG**TT**GACCTCCTTC**TA**CAGCCTCCA**AA**T**CA**CTCTCAGAGAGC  
ACTTTTAC**AT**GTTTATATCC**AA**CACATCTTTTC**TT**CTGACTCCCCT**TT**GCTATCCAAGC**TA**GCCCCCTG  
GGTCCAG**TTT**AA**GCCACCCCCTCACTAAGCCTC**

6

chr5 30893141 (Mapped tag) AGCTCCCCACTGCCCCATGCTCCAAATCACCTCTCAGAGAGCA

Junctional tag: AGCTCCCCACTGCCCCATGCT\*CCAAATCACCTCTCAGAGAGCA

30892652 30892761:

AGCTCCCCACTGCCCATGCT**CCTC**ACCCCGACTCCACCTCC**AT**GCCTGACTT**TA**TTTACTTAC**TT**GGT  
TGACCTCC**TT**CTACAG**CCT****C**CAAATCACCTCTCAGAGAGCA

7

chr5 30893071 (Mapped tag)ATGCTCCTCACCCGACTCCACCTCAAATCACCTCTCAGAGA

Junctional tag: **ATGCTCCTCACCCGACTCCA**\***CCTCAAATCACCTCTCAGAGA**

30892668-30892758:

**ATGCTCCTCACCCGACTCCA****CCTCC****AT**GCCTGACTTT**AT**TTACTTACTTGG**TT**GACCTCCTTC**TA**CAG  
**CCTCAAATCACCTCTCAGAGA**

8

chr5 30893938 (Mapped tag)AACTCAGCGTGGTTACCAGATCAGAGCCACCAGGATGATGGGC

Junctional tag: **AACTCAGCGTGGTTACCAGAT**\***CAGAGCCACCAGGATGATGGGC**

30893509-30893630:

**AACTCAGCGTGGTTACCAGAT****CAGGG****TT**ACCAGATTG**TT**AGCTGGGT**TA**GTCAGAGCAGTGGG**TA**CAGG  
TCAGGG**AA**GGGAAGGG**AA**GAGATCTCCTGCC**CAGAGCCACCAGGATGATGGGC**

9

chr5 30894924 (Mapped tag)AGACTGACCTTATTACCCCACTTTCTCTCTCTTTCTCTCT

Junctional tag: **AGACTGACCTTATTACCCCA**\***CTTTCTCTCTCTCTTTCTCTCT**

30894416-30894558:

**AGACTGACCTTATTACCCCA****CTTTC**GGC**TA**CGCAGTGG**AA**GACTCTAAGTCAGCCACTCACCC**AT**GGC  
CTCAGAGC**TT**GAGAGCA**AA**GCACCTGTGCAGACTC**AA**TGGGCCTAT**TT**GACT**CTTTC**TCTCTCTCTTTCTCTCTCTCTCT

10

chr5 30895149 (Mapped tag)TATCCAAGCTGGTGCATCCCTAGCCAGCCTGCTCTCTTCACAC

Junctional tag: **TATCCAAGCTGGTGCATCCCT**\***AGCCAGCCTGCTCTCTTCACAC**

30894640-30895127:

**TATCCAAGCTGGTGCATCCCT****AGCCAGGC****AT**CCCAGGAGACGGGGCCTCTGGTGACAGCCTGCTGCTCT  
CCCACCACCTA**AT**GACCCACAGAGGCAG**AT**CTGTGCCTGACACCTGAGTGTCCCGAGGAG**AT**GTCGGGCT  
CACAGGTGCTGG**TT**CTCCACAGCAGGCCTGGGAGAGACCAGG**TT**ATTGATCTGCCG**AT**GGAGCTCCAG  
CAGG**AA**GGCTGTGCTG**TT**GCTTTCCGGCTCC**TA**GCCCAACCCCGCTGAGTGTGCAGC**AT**GGTGACCCT  
GGGACTGAA**TT**CAAGCAGGTGGCCC**AA**GAGGGC**AA**CAGTGCCTTT**AT**GTCAGAAGTGGCCGGCCAC**TT**  
CCCAAATCACAGAGCCC**TA**TAATAAGAGAGTGAGGGG**TA**GCATTAGGGAC**AT**CCAGCCACTC**AA**AGAGC  
TCAG**TA**CTTGCCCAAGG**TT**ACAAAGCAAG**TT**GAGCACTTG**AT**GAAACCCAT**AGCCAG**CCTGCTCTCTTC  
ACAC

11

chr5 30896856 (Mapped tag)CCTGTAAGTGTGGGCATCTTTCTGTGGAGTCTCTGTGATTT

Junctional tag: **CCTGTAAGTGTGGGCATCT**\***TTCTGTGGAGTCTCTGTGATTT**

30896372-30896431:

**CCTGTAAGTGTGGGCATCT****TT**CCGAAAGCCA**TT****CTTT**CTGTGGAGTCTCTGTGATTT

12

chr5 30898490 (Mapped tag)CCCCAAGGCCCTGGGCTTGATCCCCACACTCACACACATAACC

Junctional tag: **CCCCAAGGCCCTGGGCTTGAT**\***CCCCACACTCACACACATAACC**

30897860-30898457:







TGTGTGTGTGTGTACGTGC**ATGAGT**GA**AA**TATTGAGGATGAGGCTCCCAC**AT**CTGCAGACTGGCTCAG  
GGG**AA**AGAAGGTTCCAGCAGCTGCCGAGGCC**AA**GTCCCTGTTGCC**TA**GAGTCCCTC**TT**CTCCCAGGC  
GCCACGCCAGCCTC**TA**CCATCTGAGCTCCTGAGCCCTGGGGAGGGTGTCTGAGCCT**AA**GAAAAGCCTCCC  
**TT**CACAGCTCCTGGC**AT**CTGCTCCCGAA**TT**CCCTGCCAGCCAGTGCTGTCTGGGCCAC**TA**GCCATGAC  
CAGG**AA**GGAAAG**TA**GAGAGCAGGTGATAGAGA**AA**GGATGCC**AA**GGCCTCTCCCCGTCCCCAGGCCAT  
CCATC**AA**GACTGGCCTCC**AT**GCATCTGCTGGG**AT**TCTGCC**AT**GGACCGTGCAGTGGACCGGC**ATGAGAA**  
GGACCTGCTCCCACC

23

chr5 30909982 (Mapped tag)GTGTGCCAGAAGCAGCTCCATATTAATAACATCTTCTGTGTCT  
Junctional tag: **GTGTGCCAGAAGCAGCTCCAT**\***ATTAATAACATCTTCTGTGTCT**  
30909232-30909588:

**GTGTGCCAGAAGCAGCTCCATAT**CT**AAT**CTGGGAGGACAGAGGCG**AT**CCAGGAGGG**AA**CTCCAGGCTG**T**  
**AA**CCACGCTGTGAGCGAGTGTCTCAGGG**AA**GAGCTGCGCTCC**AA**GGCAAAGCTGGCAGG**AAA**CGCGGGGT  
GCAGAC**TA**GCTTGGGGTTTT**AT**TGGCAGTCAGGA**AA**CGGAAGGTCCAGAGAGGAGTGTG**TT**GGATCAGAG  
AT**TT**ACAGCCTGTG**TA**CTGGACCTGGACCC**TT**TCACTGTATTGTGTAAATTTAAGGGATGAGTGGGGAA  
GATTTCTGAGTGGGGAGTGGCAACTCAGAGCTGTGCTGG**AA**AAAAAAGGT**AA**CAGAC**CATAT**TAATAACA  
TCTTCTGTGTCT

24

chr10 80327887 (Mapped tag)GAGGGTCTTCAGCATGTACAGGCAGAAGGCAATGCCAGGGTG  
Junctional tag: **AGTCAGACAGGAGA**\***GCCTAAATTCACAGTTGTGTGTTCTGGG**  
80327503-80327612:

**AGTCAGACAGGAGA**GTCGTGCCCCACAGAAGC**TT**GCCTCTGCC**TT**CACACTGGGCTGC**AA**GGCCCTG  
GGCAGGCC**AA**GCCTAAATTCACAGTTGTGTGTTCTGGG

25

chr10 80370546 (Mapped tag)CCAGAACTGGACTTCTGATCACAGCTGGGAAGGCTGCCGGCAC  
Junctional tag: **GGGACCCAGCAGCTCCAA**\***CTGCGGGCTGCCTGTCTGCTGTTT**  
80370104-80370384

**GGGACCCAGCAGCTCCAA**GACCCACC**TT**GGCCAGCTGTGCTCG**AA**GCTCTGCCACCTCTGTCTCCAGG  
CCCTGC**TT**GTCACTGAGGGCTGTGGCCAGCTCAGCCTCACTCCGGTG**AA**ACAGTG**AT**TCCAGATCC**TT**C  
ACTCGGCCCTGGGCCACTGTGAGCTCACC**TT**CCCGCTTC**TT**GGCACTGAAAGC**AA**GGAGGACGGGGTGC  
AG**AA**GGCCCAGGGAGGCC**AA**GCCCAAGG**TT**GCCGTGGGCCCTGGTAG**TT**G**CTGCGGGCTGCCTGTCTGC**  
**TGTTT**

26

chr10 80321886 (Mapped tag)CTATTGTTGGCAGGATCTCATGTAGCCCAGGCTGGCACCTGAC  
Junctional tag: **AGAGTTATGGTAGAGTCCTGAG**\***TCCTCAGGACTCACTTGGCTC**  
80321441-80321617

**AGAGTTATGGTAGAGTCCTGAG****CCT****TA**GCTACCTCCAGCCCCTGCCAGTC**TT**CCCTCCTCTGGGCC**AT**G  
CCCTGCGCC**TT**GGCCTGATGTAG**AT**ATGGTTGT**TTT**CTCTGGTGAGTGGCAGCC**AT**GGGGCCACTGTGC  
CACAGCAATAGTGAGCTC**TCCTCAGGACTCACTTGGCTC**

27

chr10 80256192 (Mapped tag)ACAGGTTAGGTGCATTCAATATGCCTGTCTCAGCTGAGCTGCC  
Junctional tag: **ATAGCAGTTCCTAATGATGAGGTT**\***TACTCTGGATGTTTCCAGCA**  
80255754-80256118:

ATAGCAGTTCCTAATGATGAGGTTGCATTCTCAGCCTCCACCCTGGAGTCGTGTGTGGAGGAAATCGGG  
G TAGTCAGGTGGGTGAGGTTCTAGGCTGGCCATAGGCAGAAAGCTGCCTGGGGTAGCATGGTATCCTGG  
AGCAGAGAGTTGGGCCACC TACCCTCTCCAGAACAGAGGTTGTGAAACCCTGCCTGGCTGTCAGCCCT  
GTAA GTGAAGAAGAGCCAGGTGGCAGCCCTCTGGCCTTCTCCATTGGGTCTGGCTATGTGGTGGCAGTA  
GCCTGGGCAGCGAGGGCAGCCTGGCCATTGCCATCTGGAGGTCC TAGAGCTAGCACCTCTCCAGCTGAC  
CTACTCTGGATGTTTCAGCAA

28

chr10 80258129 (Mapped tag) TAAAGGCAGCAACAATGTGTGAGATGCCTTGGGGGAACCAAGG  
Junctional tag: CGCCATCCTGGCAGCTGCTGCTCC\*CCAAGGTTGCTGCAGGGCT

80257664-80257931:

CGCCATCCTGGCAGCTGCTGCTCCAGCAGCCA TTGCCTCCACC TTA CTTGAGTGT TTAGGGGACAGACAC  
AGCATGTT CATGGCCAGTGGGCCACAGGTGCTCTGGGTACAACTGCAAGTCC TTT CCTGAGCCC ATAA  
GAAACAGG AACCCACAGTA AACTGACCATGGTGC ATAGTGGCCCCAGGTA CCCTGTGACTG ATGGGCCT  
AG TA CAGGGT AAGGCTGCCTCAGTC ATCCTTGCCACTCC AA TCC AAGGTTGCTGCAGGGCT

29

chr10 80260914 (Mapped tag) CCGCATCCCGCCGCACGGATCCGCGGAACCAACACGCGCCCT  
Junctional tag: CGGGAGCAAAGGCGTCGG\*GGCGGAGAGAAAGGAGCGTCGGGG

80261406-80261317:

CGGGAGCAAAGGCGTCGGCGTTTCGTGTGTGCGCAGGCCAGCGGATGGACTTCCGTCTTCGGTAGGCG  
GGAGAGAAAGGAGCGTCGGGG

30

chr10 80341911 (Mapped tag) GTGCATGTGTATATGTGTTATGCACACACATATCCACATGTGT  
Junctional tag: CAGGCATGCACCCAGCATTCCCTG\*AGGAGCCATGCTGCAGCTC

80342421-80342142:

CAGGCATGCACCCAGCATTCCCTGACACCCATGTGGACCC AAGTGCACCC TTT GCTAATCCTGG TATCC  
CAAGGCTGAGGCC AAGCCAATTTGTCCCATCC TTAAGGAGAC TTGATCTGCCTGA AATGGCCACTCAGA  
CCC TTGTGACCGCC ATGCACAGCAGAGGGGACAGGTGAGTGGCCCCACTA TGACATCACT TAAAGCTACA  
GAG ATTCTGGGTGAGACGC AACCATATCCCC AAA TCATCACTGAA GTGAGAGAC AGGAGCCATGCTGCA  
GCTC

### Fig. S13

All microdeletions identified in the adult mouse brain at the two loci tested. In each block the first line shows the uniquely mapped tag position and sequence. The second line is the junctional tag with the two parts highlighted in red and green. The fourth line shows the normal genomic sequence at this locus with the deleted sequence in black, the junctional tag halves in red and green and the direct repeats at the ends of the microdeletions are in bold. The AA/AT/TT/TA periodicity is highlighted in yellow.

**Table S1**

Yields of purified eccDNAs from mouse tissues, mouse and human cell lines\*

	EMB1	EMB2	EMB3	AMB	EMH	EML	NIH3T3	HeLa-S3	U937
number of tissues or cell numbers	10	7	9	2	7	6	5x10 <sup>8</sup>	1x10 <sup>9</sup>	1x10 <sup>9</sup>
extrachromosomal DNA (ng)	5930	4290	7610	4030	3220	4850	6960	19800	15600
ATP-dependent DNase resistant DNA: extrachromosomal circular DNA (ng)	237	411	304	290	269	545	384	1210	931

\*In normal tissues, the extrachromosomal DNA accounted for 0.1-0.2% weight of chromosomal DNA and in tumor cells lines for 0.01% weight. For comparison, telomeric DNA is about 0.4% of chromosomal DNA by weight.

## Table S2

Sequencing and Mapping numbers of various tissue samples and cell lines

	EMB1	EMB2	EMB3	AMB	EMH	EML	Genomic	NIH3T3	HelaS3	U937
paired reads	17667017	27782906	19314322	31196688	24068557	20468093	26993558	28009789	29056232	30065750
pairs aligned	8183161	16363457	10067149	11614968	12244947	8617404	15846611	19558292	13535751	13284489
read sequences	35334034	55565812	38628644	62393376	48137114	40936186	53987116	56019578	58112464	60131500
aligned	28227549	44942997	32990560	36524091	39350135	34026009	51686809	50454159	54325530	54992859
unique alignment	25634832	40792519	30556149	32931141	33837614	29666618	45368559	47764234	44250974	43666541
number of unique sites in the genome yielding circles	24921	22604	15722	14046	9689	13758	114	10517	38876	69483

Number of paired end reads obtained, total number of tags mapped to the genome, number of tags that were mapped uniquely in the genome and number of unique microDNAs identified in each library.