

Article

Identification of Extrachromosomal Linear microDNAs Interacted with microRNAs in the Cell Nuclei

Teng Sun ^{1,2,*}, Kun Wang ³, Cuiyun Liu ³, Yin Wang ³, Jianxun Wang ^{3,*} and Peifeng Li ^{3,*}

¹ Key Laboratory of Cellular Physiology, Ministry of Education, Shanxi Medical University, Taiyuan 030001, China

² Department of Physiology, Shanxi Medical University, Taiyuan 030001, China

³ Institute for Translational Medicine, Qingdao University, Qingdao 266071, China; wangk696@yeah.net (K.W.); zhuanhuacuuiyun@126.com (C.L.); wangying@sibs.ac.cn (Y.W.)

* Correspondence: tengsun@qdu.edu.cn (T.S.); wangjx@qdu.edu.cn (J.W.); peifliqdu@126.com (P.L.); Tel.: +86-(351)-41-351-17

Received: 18 December 2018; Accepted: 31 January 2019; Published: 1 February 2019



Abstract: Extrachromosomal DNA exists in two forms: Covalently closed circular and linear. While diverse types of circular extrachromosomal DNA have been identified with validated *in vivo* functions, little is known about linear extrachromosomal DNA. In this study, we identified small, single-stranded linear extrachromosomal DNAs (SSLmicroDNAs) in the nuclei of mouse hearts, mouse brains, HEK293, and HeLa cells. We used a pull-down system based on the single-stranded DNA binding protein RecA_f. We found that SSLmicroDNAs aligned predominantly to intergenic and intragenic regions of the genome, owned a variety of single nucleotide polymorphism sites, and strongly associated with H3K27Ac marks. The regions were tens to hundreds of nucleotides long, periodically separated by AT, TT, or AA dinucleotides. It has been demonstrated that SSLmicroDNAs in the nuclei of normal cells target microRNAs, which regulate biological processes. In summary, our present work identified a new form of extrachromosomal DNAs, which function inside nuclei and interact with microRNAs. This finding provides a possible research field into the function of extrachromosomal DNA.

Keywords: extrachromosomal DNAs; small single-stranded linear extrachromosomal DNAs (SSLmicroDNAs); sequence analysis; microRNAs; cell nucleus

1. Introduction

Extrachromosomal DNA—DNA molecules separated from chromosomes—exist in two main forms: Covalently closed circular and linear DNA [1–5]. Mitochondrial DNA are located in mitochondria and represent typical extrachromosomal circular DNA (eccDNA). This type of DNA encodes functional genes and contributes to the nuclear genomes' instability [6]. Small polydispersed circular DNA, double minute chromosomes, episomes, minichromosomes, autonomously replicating sequences, telomeric-circles, B- and T-cell receptor excision circles, and extrachromosomal elements induced by *c-myc* oncogene deregulation and resulting in genomic instability were discovered as extrachromosomal circular DNA molecules with unique properties and relevant *in vivo* functions, which are produced randomly or non-randomly [1,7–10]. In recent years, an increasing number of eccDNAs have been identified in many organisms. Preliminary research links eccDNAs to cancer [11,12]. Recently, much attention has focused on a class of small circular extrachromosomal DNAs (microDNAs), with short length and unique features. microDNAs have been found ubiquitously expressed in mammalian cells, including mouse brain, heart, kidney, liver, lung, skeletal muscle, sperm,

spleen, testis, and thymus cells as well as in cancer cell lines. Several hypotheses exist regarding the origin of microDNAs, including genome microdeletion, mismatch repair pathways, and transcriptional activity [7,13].

Studies have demonstrated wide distribution and function of diversiform eccDNAs in animals, plants, and microorganisms [1,7,8], while extrachromosomal linear DNAs only exist in certain organisms or at certain states [2,4,14–18]. Linear plasmid pAM1 was identified outside chromosomes in the *Actinoplanes* species [17]. Small linear extrachromosomal DNAs has been shown to exist in the mitochondria of plant and fungi [2,16]. *Creophilus maxillosus* and *Candida albicans* have been shown to contain genome-independent linear ribosomal DNA (rDNA) [15]. Some linear extrachromosomal DNAs appear during pathological states. Twelve linear extrachromosomal plasmids with sequence similarities to plasmids in *Borrelia burgdorferi* type strain isolate B31 were discovered in isolates of Lyme disease agent [5,18]. During retroviral infections, numerous un-integrated retroviral DNAs accumulate outside chromosomes in infected cells [14]. It is not yet clear whether new forms of extrachromosomal linear DNA exist in higher organisms, and whether they perform a certain function.

MicroRNAs (miRNAs) are a type of ~22 nt long non-coding RNAs, which play a critical role in regulating gene expression. They usually target the 3'-untranslated region (UTR) or mRNA coding sequences (CDS) to prevent mRNA translation or promote degradation [19]. miRNAs are now a fundamental part of the entire regulatory network of biological process, including cell proliferation, differentiation, and death both during physiological and pathological states [19–22]. Many miRNAs are found in the nucleus where they regulate multiple processes, such as chromatin remodeling [23], transcriptional silencing [24], mRNA alternative splicing [25,26], and microRNA maturation [20,27].

Our present work identified a new form of extrachromosomal linear DNA, single-stranded linear microDNAs (SSLmicroDNAs), which are found in the nuclei of multiple cell types, including adult mouse hearts, mouse brains, HEK293 cells, and HeLa cells. We analyzed the unique features of SSLmicroDNAs, and we proposed several hypotheses. Our results revealed that SSLmicroDNAs interact with microRNAs in nuclei, implying a potential role in microRNA regulatory pathways.

2. Materials and Methods

2.1. Isolation of Extrachromosomal Single-Stranded Linear DNAs and Related microRNAs

Nuclei were incubated with a weak lysis buffer (20 mM Tris-HCl (pH 7.5), 200 mM NaCl, 2.5 mM MgCl₂, 0.05% Igepal, 60 U RNase inhibitor, 1 mM DTT, and proteinase inhibitor) and shaken overnight at 4 °C. The lysates were pre-cleared by centrifugation at 3000 rpm for 10 min, followed by incubation with single-stranded DNA binding protein RecA_f (New England Biolabs@ Inc., Ipswich, MA, USA), and 2.4 mM ATP at 37 °C for 3 h. NTA-Ni agarose beads were incubated with RecA_f-DNA complexes at 4 °C with shaking for 4 h, then washed with Ni-washing buffer (20 mM hepes (pH 7.5), 10% glycerol, 0.3 M NaCl, 0.2% Triton X-100, 25 mM imidazole, 10 mM beta-mercaptoethanol, and 0.5 mM PMSF) three times. The beads were divided into two aliquots: One was eluted using Ni-Elution buffer (20 mM hepes (pH 7.5), 10% glycerol, 0.3 M NaCl, 0.35 M imedazole, 10 mM beta-mercaptoethanol, and 0.5 mM PMSF) to extract total single-stranded DNA, while the other was used to extract related RNAs using Trizol (Invitrogen™ life technology, Waltham, CA, USA).

2.2. SSLmicroDNA Library Construction and Sequencing

Total single-stranded linear DNAs were ligated with two specific double-stranded adaptors— adaptor A or adaptor B. The adaptor A forward sequence was: 5'-CACACTCTTCCCTACACGACGCTCTTCCGATCTTTGCNNNNNN-3', and the reverse sequence was 5'-GCAAAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTG-3'. The adaptor B forward sequence was 5'-NNNNNNGTTCAGAGTTCTGCGACAGGAGAGGTCGTATGCCGTCTTCTGCTTG-3', and the reverse sequence was 5'-CAAGCAGAAGACGGCATACGACCTCTCCTGTGCGAGAACTCTGAAC-3'. The adaptors could only ligate to single-stranded linear DNA through sticky ends and six random

bases pairing to single strands. The complementary strand was synthesized and amplified by a PCR reaction with the forward primer 5'-CAAGCAGAAGACGGCATAACGA-3' and the reverse primer 5'-ACACTCTTTCCCTACACGAC-3'. DNA mixtures ranging from less than 500 bp, 500–1000 bp, and 1000–2000 bp were collected and cloned into the pZeroback T vector (TianGen biotech co., LTD., Beijing, China). Sequencing was performed using the TSINGKE Biological Technology.

2.3. Atomic Force Microscopy

DNA was imaged using atomic force microscopy, the experiment was conducted following the protocol described in Reference [28]. Briefly, a drop of DNA (5 ng/ μ L) with 5 mM MgCl₂ was placed on the surface of freshly cleaved mica, and left for 2 min at room temperature. The mica was then rinsed with 1 mL water, blotted with filter paper, and dried by the flow of compressed nitrogen for 2 min. Samples in mica were scanned using a Digital Instrument's MultiMode scanning probe on a Nanoscope IIIa (Veeco, New York, NY, USA) microscope operating in Tapping Mode.

2.4. Fluorescence in situ Hybridization

Fluorescence in situ hybridization was performed following after modifying the method described in Reference [29]. Briefly, cells were grown on glass coverslips, coated with poly-L-lysine (Sigma-Aldrich, St. Louise, MO, USA), and fixed with 4% paraformaldehyde for 30 min at room temperature. They were treated consecutively with 4.5% sucrose buffer for 1 h four times, fresh 0.1 M phosphate buffered saline (PBS) for 5 min, 0.3% Triton X-100 for 5 min, 0.1 M PBS for 5 min two times, and 2 μ g/mL proteinase K (diluted by 50 mM Tris and 2 mM calcium chloride) at 37 °C for 20 min. The cells were then fixed again, and treated with 0.25% acetic anhydride and 0.1 M ethanolamine for 10 min, 50% formamide/2*SSC (0.3MNaCl, 30mM sodium citrate) for 10 min at room temperature, and then incubated at 37 °C for 15 min. Pre-hybridization was performed using 100 μ g/mL salmon sperm DNA and incubated at 37 °C for 2 h. Finally, the samples were incubated with 400 ng/mL of probe, in hybridization buffer, at 37 °C overnight. The 293mid-28 probe was modified using 6-FAM at the 5' end, and its sequence was 5'-AGATCGCACTACTGCACTCCAGCCTGGGTGACAGAGTAAGACTTCGTCTC-3'. The HeLamid-18 probe was modified using 6-TAMRA at the 5' end and its sequence was 5'-GATGGGGGTAAATATCCAGGCTTCCGCTCTGCTTCCATTGACACCCAAT-3'.

2.5. Pull-Down Assays and Quantitative Real-Time PCR

Pull-down assays were performed according to the protocol in [20]. Briefly, cells were transfected with biotinylated miDNAs (50 nM; Sangon Biotech, Shanghai, China), and harvested 72 h after transfection. The cells were incubated in lysis buffer (20 mM Tris (pH 7.5), 200 mM NaCl, 2.5 mM MgCl₂, 0.05% Igepal, 60 U/mL Superase-In (Ambion, Waltham, MA USA), 1 mM DTT, and protease inhibitors (Roche, Basel, Switzerland)), on ice for 30 min after washing with PBS two times. M-280 streptavidin magnetic beads (Sigma-Aldrich) were blocked by RNase-free BSA and salmon sperm DNA (both from Sigma-Aldrich), which were added to the lysates and incubated at 4 °C for 3 h. The beads were washed three times with ice-cold lysis buffer, three times with a low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, and 150 mM NaCl), and once with a high salt buffer (0.1%SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, and 500 mM NaCl). Bound RNA was extracted using Trizol (Invitrogen™ life technology) for downstream quantitative real-time PCR. The sequence of biotin-293mid-28 was 5'-TTTTTTTTTTTTTTGAGACGAAGTCTTACTCTGTACCCAGGCTGGAGTGCAGTAGTGCGATC TCAGCTCACTGCAACCTCCGCTCCAG-3'. The sequence of 293mid-54 was 5'-CAAGGTCATGA GTTTGAGATCAGCCTGGCTAATATGGTGAAACCCATCTCTACTAAAAATACAAAAATTAGCTG GGTGTGGTGGTGGGCACCTGTAGTCCATCTACTTGGGAGAC-3'. Quantitative real-time PCR was performed as previously described [20]. Briefly, miRNA levels were measured by qRT-PCR

using SYBR®Premix Ex Taq™ II (Takara Clontech, Kyoto, Japan). miRNAs expression levels were normalized to U6 before the pull-down assay.

2.6. Cell Culture and Isolation of Nuclei

HEK293 and HeLa cells were grown in Dulbecco's modified Eagle's medium (DMEM; Invitrogen™ life technology) supplemented with 10% Fetal bovine serum (FBS), 100 units/mL penicillin, and 100 µg/mL streptomycin and incubated at 37 °C and 5% CO₂. Nuclei were isolated according to a previously described method with minor modifications [7]. In this study, cells were isolated from mouse hearts (adult males, 8–12 weeks of age), mouse brains (adult males, 8–12 weeks of age), HEK293 cells, or HeLa cells. Nuclei were extracted from tissue or cell lines using sucrose gradient ultracentrifugation. Fresh tissue or cell lines were homogenized using a wheaton douncer, in buffer A (20 mM hepes pH 7.5, 10 mM KCl, 1.5 mM MgCl₂, 1 mM EGTA, 1 mM EDTA, 1 mM DTT, 0.1 mM PMSF, and 250 mM sucrose). Cell debris was removed and intact cells were pelleted after centrifugation at 800 rpm for 10 min. Nuclei were precipitated through centrifugation at 3000 rpm for 15 min.

2.7. Sequence Analysis

Genome mapping, H3K27Ac activity, single nucleotide polymorphism (SNP) site, and conservation level analysis were performed in the UCSC (University of California Santa Cruz) genome browser. The length distribution and GC content were analyzed by Mega software. Base-pairing analysis was performed using the bioinformatics program RNAhybrid. All statistical analyses were performed with the SPSS 16.0 software.

2.8. Detection of the Endogenous Linear microDNAs

HEK293 cells were transfected with biotinylated 293mid-62 antisense sequence 5'-TTGAGCATTTACCATGTTCTGGGAGTGTGGAAGGCCTTGGGTGTATAATGTGAGCATGTG CCTGACCTGTGCTGATGGAGGAGACAAATA-3', which was 20 nt shorter at the 5'-end than 293mid-62. The pull-down assay, based on interactions between biotin and streptavidin (described in the method—pull-down assay and quantitative real-time PCR), was performed to purify 293mid-62. The complementary strand was synthesized and amplified by a PCR reaction with the forward primer 5'-CTTTGATCAGCATTATTTATTTG-3' and the reverse primer 5'-TATTTGTCTCCTCCATCAGCACAGG-3'. The amplified product was cloned into the pLB Vector (TIANGEN biotech) and sequenced. HeLamid-1 was analyzed via similar experiments with 293mid-62. The biotinylated HeLamid-1 antisense sequence was 5'-TTTTGTGGGCAACGTGACTCACTATGCCAGGTCTGGCTCAACATCTCGGCGGAG-3'. The forward primer was 5'-CAGGCCAACGAGACTTTTGCTTTTG-3' and the reverse primer was 5'-CTCCGCCGAGATGTTGAGCCAGACC-3'.

2.9. Transfection and Detection of Exogenous SSLmicroDNAs and Reverse Strands

The 293mid-28, HeLamid-18, and their reverse strands were synthesized by the Sangon Biotech company. The 293mid-28 and its reverse strands were modified using 6-FAM at the 5' end. HeLamid-18 and its reverse strands were modified using 6-TAMRA at the 5' end. The sequence of 293mid-28 was 5'-TTTTTTTTTTTTTTTGGAGACGAAGTCTTACTCTGTCAACCAGGCTGGAGTGCAGTAGTGCGA TCTCAGCTCACTGCAACCTCCGCCTCCCAG-3', and its reverse strand sequence was 5'-CTGGGAGGCGGAGGTTGCAGTGAGCTGAGATCGCACTACTGCACTCCAGCCTGGGTGACA GAGTAAGACTTCGTCTCAAAAAAAAAAAAAAAAAA-3'. The sequence of HeLamid-18 was 5'-CACCCAGCAGCAACAGGGGTCCCCACCCATTGGGTGTCATGGAAGCAGAGCGGAAAG CCTGGATATTTACCCCATCTAGAAGTAACAAG-3', and its reverse strand sequence was 5'-CTTGTTACTTCTAGATGGGGGTAAATATCCAGGCTTTCCGCTCTGCTTCCATTGACACCCAAT GGGGTGGGGGACCCCTGTTGCTGCTGGGTG-3'. HEK293 and HeLa cells were grown on glass

coverslips coated with poly-L-lysine (Sigma-Aldrich). The transfection was performed using the transfection reagent Fugene (Promega, Fitchburg, WI, USA), according to the manufacturer’s instructions. Cells were fixed with cold methanol at $-20\text{ }^{\circ}\text{C}$ for 5–10 min, and stained with DAPI. They were analyzed using a laser confocal fluorescence microscope.

3. Results

3.1. Small, Single-Stranded Linear Extrachromosomal DNAs (named SSLmicroDNAs) were Identified in Cell Nucleus

We purified single-stranded linear DNA using the single-stranded DNA binding protein RecA_f. The DNA was extracted from the nuclei of adult mouse hearts, brains, HEK293 cells, and HeLa cells [30,31] (Figure 1a). The purified single-stranded linear DNAs were further identified using atomic force microscopy (AFM; Figure 2a). Smear-like band distributions with a broad range were observed in agarose gel electrophoresis analysis (Figure 2b).

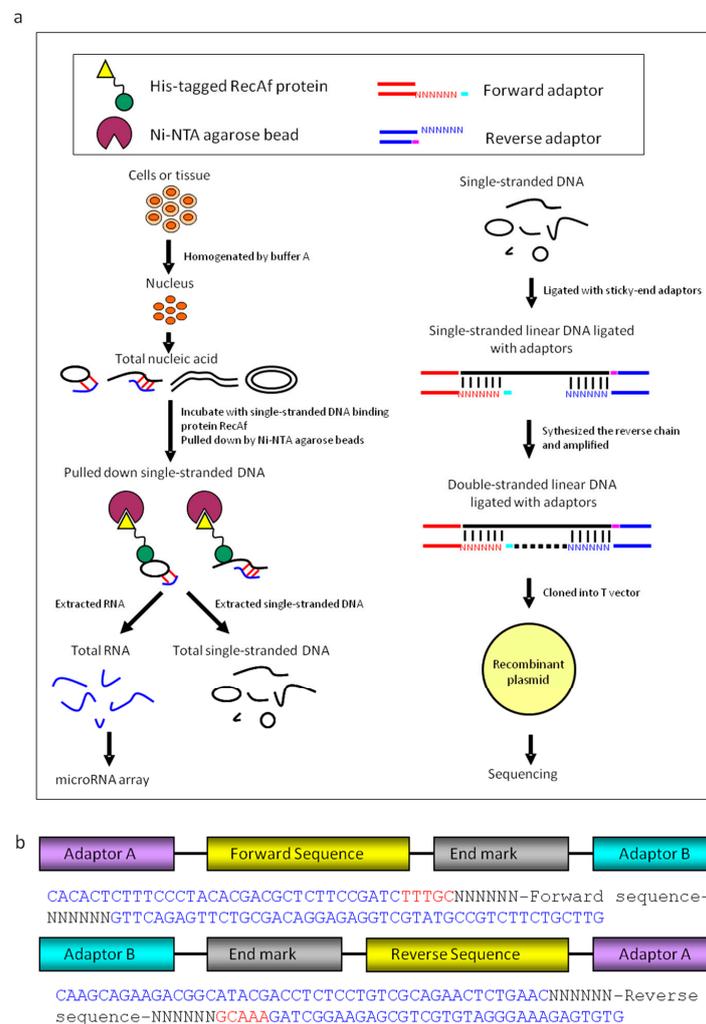


Figure 1. Preparation of extrachromosomal single-stranded linear microDNAs and microRNAs. (a) The single-stranded linear DNAs were pulled down by the RecA_f protein. RNAs and single-stranded linear (SSL) DNAs were isolated through pulldown system based on RecA_f protein. microRNA array analysis was performed with the RNA samples. Following ligation with two adaptors, the single-stranded linear DNAs libraries were constructed and sequenced. (b) The composition of adaptor A and adaptor B. Adaptors composed of end markers and random bases. N represents a random base. TTTGC and GCAAA were end markers in adaptor A or B. Ni-NTA.

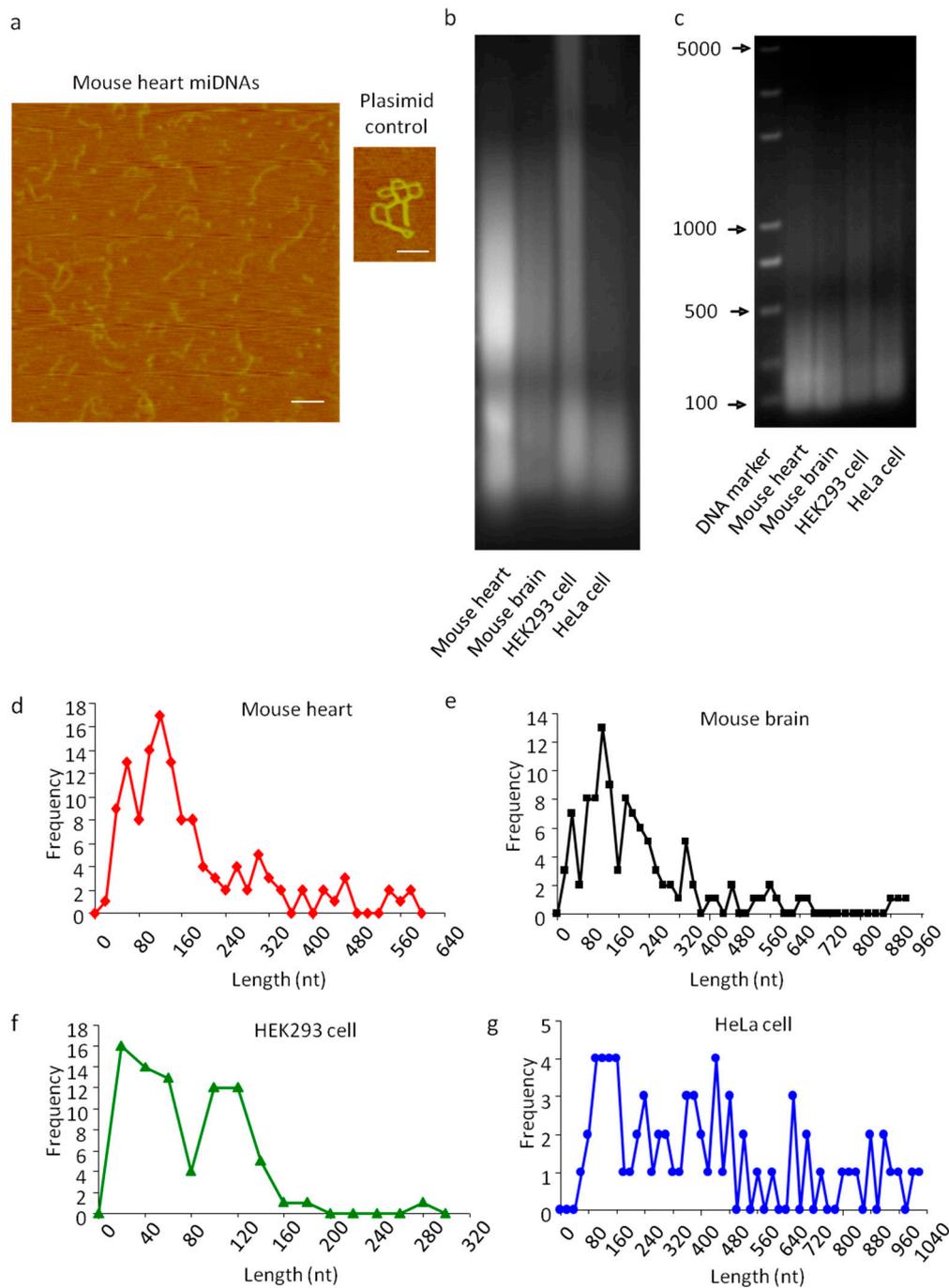


Figure 2. Identification of single-stranded linear extrachromosomal microDNAs. (a) Analysis of mouse heart nucleic single-stranded extrachromosomal DNA by atomic force microscopy. A 2982 bp double-stranded plasmid was used as a control. Bar = 200 nm. (b) Analysis of single-stranded extrachromosomal DNA by agarose electrophoresis. (c) Library amplification of single-stranded extrachromosomal DNA from mouse hearts (MHSSLmicroDNA), mouse brains (MBSSLmicroDNA), HeLaSSLmicroDNA and 293SSLmicroDNA, respectively. d-g, Analysis of the length distribution of SSLmicroDNAs from mouse hearts (d), mouse brain (e), HEK293 cells (f) and HeLa cells (g), respectively.

To characterize linear DNAs, four libraries of single-stranded linear DNAs from mouse hearts, mouse brains, HEK293 cells, and HeLa cells were constructed. The library construction method consisted of ligating single-stranded DNA-specific adaptors, and synthesizing complementary chains

(Figures 1c and 2c and Supplementary Figure S2). DNA sequencing analysis identified 128 single linear DNA sequences from mouse hearts (Supplementary data set 1), 107 from mouse brains (Supplementary data set 2), 78 from HEK293 cells (Supplementary data set 3), and 70 from HeLa cells (Supplementary data set 4). These data showed that mouse heart-derived linear DNA had a 16–579 nt length distribution with a peak at approximately 130 nt (Figure 2d); 50–400 nt length for mouse brain-derived linear DNA (Figure 2e); and 20–290 nt length for those derived from HEK293 cells, of which >90% were distributed between 40–150 nt (Figure 2f). However, compared to normal tissue, the length distribution of linear DNAs from HeLa cells was irregular, without a defined length peak (Figure 2d–g). Since linear DNAs have molecular weights ranging from tens to thousands, with peaks at the hundreds mark, we named them single-stranded linear microDNAs (SSLmicroDNAs). SSLmicroDNAs derived from mouse hearts, mouse brains, HEK293 cells, and HeLa cells were named MHSSLmid, MBSSLmid, 293SSLmid, and HeLaSSLmid, respectively.

3.2. SSLmicroDNAs Exhibited a Series of Unique Features and Their Origins Hypothesis was Put Forward

To understand the nature of SSLmicroDNAs, we performed detailed characterization using statistical and bioinformatic methods. We conducted a comparative analysis of all 383 SSLmicroDNA sequences in the UCSC genome browser (Figure S3). The results showed >90% of these sequences were located in non-coding regions, including intergenic and intragenic regions (Figure 3a). Analysis of microDNA loci and their neighboring regions showed that the 42.5% GC content of HeLaSSLmids was slightly higher than the 37.5% GC content of their 500 bp up or downstream neighboring regions (Figure 3b). MHSSLmids had a similar GC content to HeLaSSLmids, while MBSSLmids and 293SSLmids did not (Figure 3b). Remarkably, AT, TT, or AA dinucleotides appeared periodically, every 10 bp along the microDNA sequence, and were identified in approximately 76.64% of 293SSLmids, 64.84% of MHSSLmids, 57.01% of MBSSLmids, and 38.57% of HeLaSSLmids (Figure 3c). The periodicity of microDNAs isolated from normal tissues was much greater than those from tumor cells (Figure 3c). A previous study revealed that AT, TT, or AA dinucleotide periodicity usually appeared in sequences assembled into nucleosomes [32]. Therefore, we hypothesized that SSLmicroDNAs may originate from nucleosomes.

Single nucleotide polymorphisms (SNPs) refer to changes in single nucleotides of DNA, which lead to genome diversity and instability [33]. Our results showed that ~50% of microDNAs in mouse heart, mouse brain, and HEK293 cells, and >90% in HeLa cells were located in SNP-rich regions—2 kb upstream or downstream (Figure 3d,f). H3K27Ac chromatin marks, usually appear at the site of active regulatory elements, such as gene enhancers and promoters [34]. Approximately 30% of SSLmicroDNAs from HEK293 and HeLa cells were located near high-H3K27Ac-activity regions (Figure 3e,f), which led us to hypothesize that H3K27Ac-marked regions could be a source for SSLmicroDNA production. Taken together, these data indicate that SSLmicroDNAs are likely generated from regions with strong H3K27Ac marks, and frequent single nucleotide polymorphisms.

To further confirm the type of purified extrachromosomal DNA; single-stranded or linear, we designed a pull-down system based on the interaction between biotin and streptavidin (Supplementary Figure S1). The results showed that 293mid-62 was pulled down by a 293mid-62 probe modified with a biotin tag at the 5'-end, and HeLamid-1 was pulled down by its respective probe (Figure 4a).

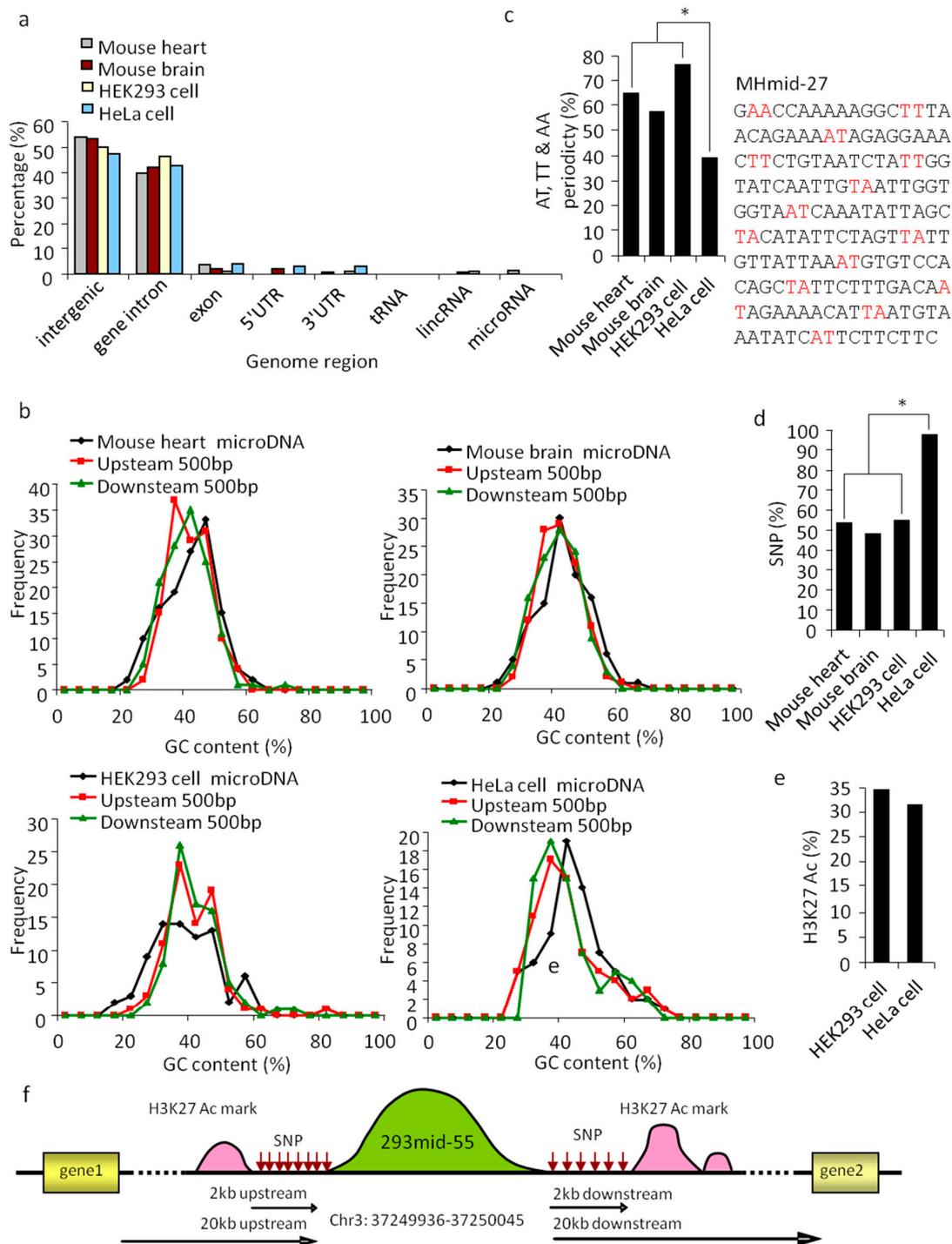


Figure 3. Properties of SSLmicroDNA sequences and their genomic loci. **(a)** Genome mapping of SSLmicroDNAs from mouse hearts, mouse brains, HEK293 cells and HeLa cells. **(b)** GC content of SSLmicroDNAs and their 500 bp-upstream and 500 bp-downstream regions. **(c)** Analysis of AT, TT or AA periodicity of SSLmicroDNA sequences from mouse hearts, mouse brains, HEK293 cells and HeLa cells. The MHmid-27 sequence periodically separated by AT, TT or AA is presented. * $p < 0.05$. **(d)** Analysis of single nucleotide polymorphism sites (SNPs). The data are the percentages of microDNA loci with SNPs within 2kb upstream or downstream. * $p < 0.05$. **(e)** Analysis of H3K27 Ac marks within 2 kb-upstream or -downstream regions of SSLmicroDNAs genomic loci. The data are the percentages of microDNA loci that have H3K27 Ac marks within the indicated regions. **(f)** A glance at the scenery near 293mid-55 in the genome.

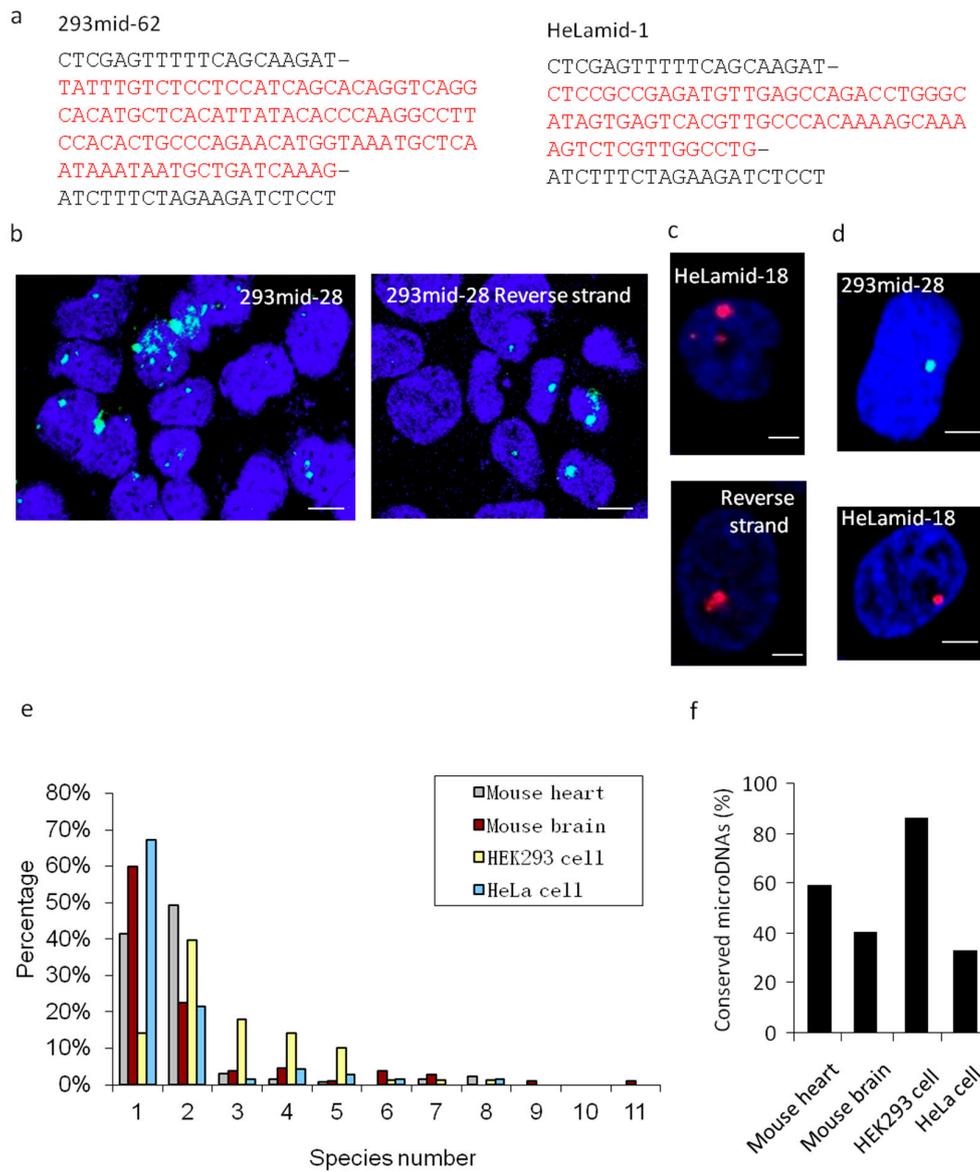


Figure 4. Determination of microDNA subcellular localization and conservation level. (a) The sequencing data of 293mid-62 and HeLamid-1 in the detection of endogenous linear microDNAs. The 293mid-62 probe with a biotin tag at the 5'-end was transfected into HEK293 cells, and HeLamid-1 probe with a biotin tag at the 5'-end was transfected into HeLa cells. Then the pull-down system based on interaction between biotin and streptavidin was performed and the 293mid-62 and HeLamid-1 were detected by sequencing. (b,c) Detection of SSLmicroDNA subcellular localization. HEK293 cells were transfected with 293mid-28 or its reverse strands labeled with a FAM at the 5'-end (b), and HeLa cells were transfected with HeLamid-18 or its reverse strands labeled with a TAMRA at the 5'-end (c). The nuclei were stained by DAPI (the dark blue area). The light blue dots represented 293mid-28 or its reverse strands. The red dots represented HeLamid-18 or its reverse strands. The Bar = 10 μm. (d) Fluorescence in situ hybridization (FISH) analysis of SSLmicroDNA localization in cell nucleus. 293mid-28 was detected by its probes labeled with FAM at the 5'-end in HEK293 cell nucleus. HeLamid-18 was detected by its probes labeled with TAMRA at the 5'-end in HeLa cell nuclei. The nuclei were stained by DAPI (the dark blue area). The light blue dots represented 293mid-28. The red dots represented HeLamid-18. Bar = 20 μm. (e) Analysis of the conservation level of SSLmicroDNA sequences. Species number represents the number of species in which SSLmicroDNAs were conserved. f, Summary of the conservation level of SSLmicroDNAs. The SSLmicroDNAs that were conserved in more than one species were counted.

3.3. Cell Nuclei- Located SSLmicroDNAs Owned a High Conservation Level

We then studied the subcellular localization of microDNAs, 293mid-28 and HeLamid-18 were chosen randomly. Twenty-four hours post transfection, both 293mid-28 and its reverse strand were observed in the nuclei of HEK293 cells (Figure 4b). HeLamid-18 and its reverse strand were transfected into HeLa cells, and were also found in the nucleus (Figure 4c). These results were further confirmed by fluorescence in situ hybridization (FISH), showing endogenous 293mid-28 and HeLamid-18 in the nuclei (Figure 4d).

After analyzing the conservation levels of SSLmicroDNAs, we found that >80% of HEK293SSLmicroDNAs, ~60% of MHSSLmicroDNAs, and ~40% of MBSSLmicroDNAs were conserved between two or more species, whereas only 32.86% of HeLaSSLmicroDNAs shared this feature (Figure 4e,f). These results indicated that SSLmicroDNAs in normal tissues or cells are slightly more conserved compared to their tumor-derived counterparts.

3.4. SSLmicroDNAs Interacted with microRNAs in the Cell Nucleus

Many miRNAs are distributed in nuclei, where they can regulate chromatin remodeling, transcriptional silencing, mRNA alternative splicing, and microRNA maturation [23–25,27]. The abundance of SSLmicroDNAs in the nuclei of normal tissues and cells suggested a potential participation in microRNA regulatory pathways. Therefore, we studied the interaction between microDNAs and microRNAs. MicroRNAs were extracted from the nuclear pull-down product via the single-stranded DNA binding protein RecA_f and analyzed using microRNA arrays. MicroRNA arrays identified 421 mouse heart (Table S1), 285 mouse brain (Table S2), and 1101 HEK293 microRNAs (Table S3). We analyzed all of the microRNA and SSLmicroDNA sequences using the bioinformatics program RNAhybrid. The results showed that microDNAs pair well to at least one microRNA (Figure 5a–c, and Supplementary Figures S4–S6). Among all the microDNA–microRNA sequence alignments, we detected ~45.76% hybridization with more than 20 kcal/mol free energy in the mouse brain library, and 36.38% in the mouse heart library (Figure 5d). Furthermore, 24.41% of the hybridizations between microDNAs and microRNAs in mouse hearts, 28.97% in mouse brains, and 27.27% in HEK293 cells exhibited more than 6 bp-continuous base pairing at microRNA seed regions. At any given location, 10 bp-continuous base pairing appeared in 56.69% of the hybridizations in the mouse heart, 58.88% in mouse brains, and 66.23% in HEK293 cells (Figure 5e). The microRNAs seed region is a critical recognition and functional site for targeting to mRNAs [35]. Therefore, we hypothesized that this was also a site for microDNAs interaction with microRNAs in the cell's nucleus. To validate this hypothesis, we performed pull-down experiments, and showed that mir-1273g-3p was pulled down by the 293mid-28, and mir-5096 was pulled down by the 293mid-54, which was both interrupted by the mutation in the predicted target region of microDNA (Figure 5f). Taken together, these data indicate that SSLmicroDNAs likely interact with microRNAs in nuclei.

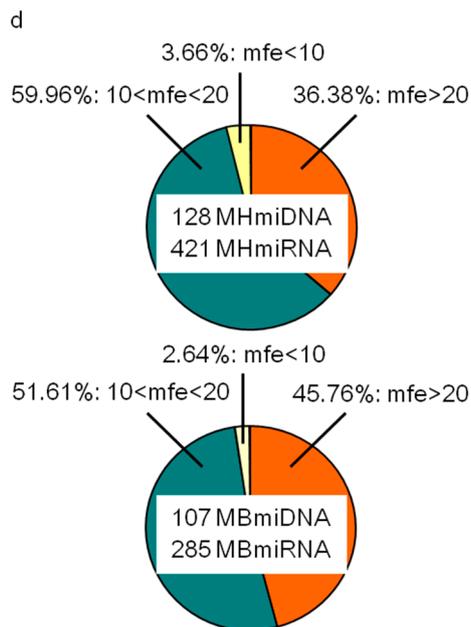
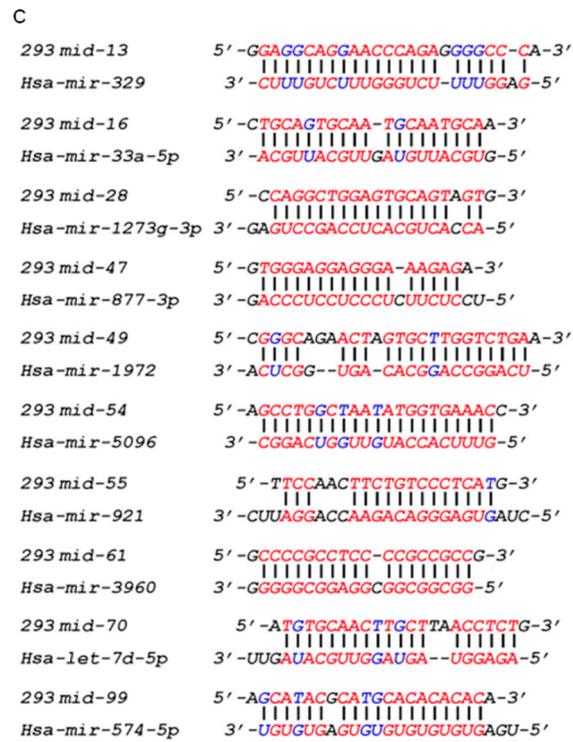


Figure 5. Cont.

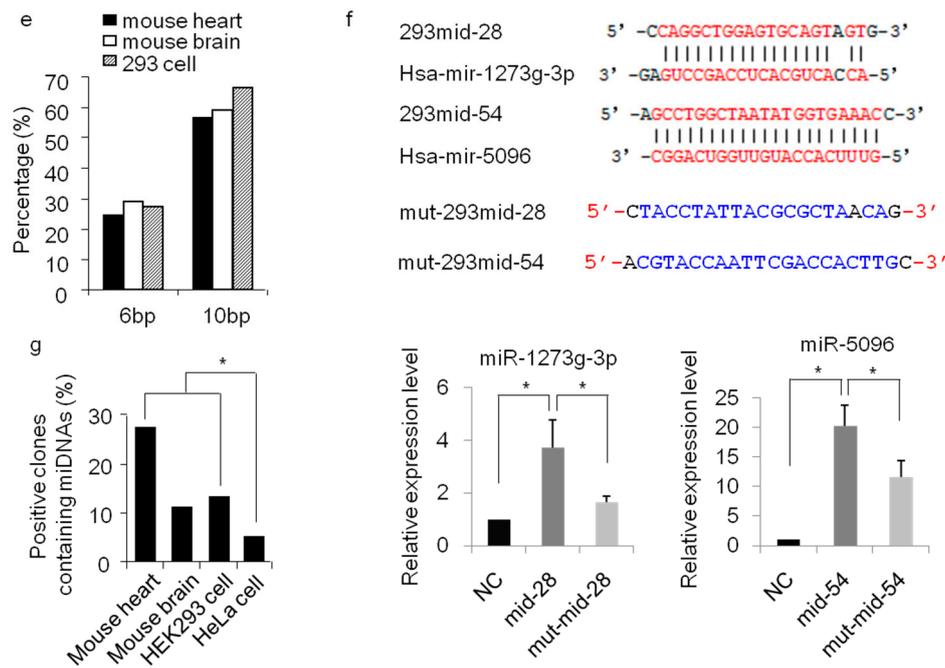


Figure 5. Interaction between microDNAs and microRNAs. (a–c) Sequence alignments between SSLmicroDNAs and microRNAs in HEK293 cells (a), mouse hearts (b) and mouse brains (c) were analyzed using the bioinformatics program RNAhybrid. (d) Analysis of the free energy (mfe) in microDNA-microRNA alignments. The unit of energy was $-kcal/mol$. (e) Analysis of the microDNA-microRNA sequence alignment. 6 bp represented 6 bp- or more than 6 bp-continuous pairing in the seed region of microRNAs. 10 bp represented 10 bp- or more than 10 bp-continuous pairing in microRNAs. (f) Detection of the interaction between microDNAs and microRNAs in cell nucleus. The upper panel showed the sequence alignment we predicted between 293mid-28 and hsa-mir-1273g-3p, as well as 293mid-54 and hsa-mir-5096. The middle panel showed the mutant sequence region in 293mid-28 and 293mid-54. HEK293 cells were transfected with biotinylated 293mid-28 or negative control RNA (NC) or mutant 293mid-28 and then performed with the pulldown system based on interaction between biotin and streptavidin. The levels of miR-1273g-3p were detected by qRT-PCR in the pulldown product. $n = 3$, * $p < 0.05$ vs. NC. The combination of 293mid-54 and miR-5096 was analyzed similarly to that of 293mid-28. (g) Analysis of SSLmicroDNA yield. The rates of SSLmicroDNA-positive clones relative to the total number of clones during purification were counted. * $p < 0.05$.

4. Discussion

Extrachromosomal DNAs exist in diversiform and are ubiquitously expressed *in vivo*, however, little is known about their function. Our present work identified a new form of extrachromosomal linear DNAs found in normal tissue and tumor cells. These DNAs were characterized by linear single strands, tens to hundreds of nucleotides length, a little higher than average GC content periodically intercepted by AT, TT, or AA dinucleotides. The single-stranded linear microDNAs were mainly localized at the non-coding regions of genome, where various SNP sites and strong H3K27Ac marks existed. SSLmicroDNAs, located in nuclei, were shown to interact with microRNAs *in vivo*. The difference in microDNAs content between normal tissues and tumor cells suggest that microDNAs may have a function in cellular defense against tumorigenesis. Briefly, our results revealed a new form of extrachromosomal DNAs, which could interact with microRNAs in the cell nucleus.

Circular microDNAs have been identified in normal tissues, and exhibit unique sequence features [7]. Compared to circular microDNAs, linear microDNAs own a similar length distribution and GC content; both forms shared AT, TT, or AA dinucleotide periodicity; circular microDNAs originate predominantly from functional regions such as exons, 5' UTRs, 3' UTRs, and CpG regions [7],

whereas linear microDNAs are mainly produced from non-coding regions, such as intergenic or intragenic regions of genome.

Attention has been focused on extrachromosomal DNAs' general mechanisms of action, with no clear understanding as of yet. Viral DNA likely fails to integrate into host genomes and remains outside host chromosomes during viral infection, of which a typical example is human immunodeficiency virus (HIV) [36]. Additionally, DNA of the adeno-associated virus (AAV) exists mainly as circular episomes in human tonsil-adenoid, spleen, and lung tissues after infection [37]. Non-integrated Moloney murine leukemia virus (M-MuLV) DNA has been identified in a human rhabdomyosarcoma cell line (TE671 subline) [38]. Therefore, a bold hypothesis about the origin of linear microDNAs is that they could evolve from infectious organisms.

Cell defense mainly relies on the immune system [39]. DNA-protein crosslinking and non-coding RNAs have been shown to function directly in regulating cell defense [40,41]. Interestingly, HeLa microDNAs demonstrated an irregular length distribution (Figure 2d–g) and lower yield during the extraction of SSL microDNAs (Figure 5g, and Table 1) from HeLa cells compared to normal tissues. This indicated that microDNAs may participate in maintaining the physiological environment against tumorigenesis.

Table 1. Summary of single-stranded microDNA library construction.

	The Number of Clones	The Number of Sequenced Clones	The Number of Clones Containing Sequences	The Number of Clones Containing SSL microDNAs	Rate
Mouse Hearts	3140	470	165	128	27.23%
Mouse Brains	3692	966	274	107	11.08%
HEK293 Cells	4900	702	642	78	12.15%
HeLa Cells	4700	2020	331	70	3.47%

In summary, we identified a new form of extrachromosomal microDNAs and analyzed their characteristics. SSL microDNAs are located in the cell nucleus, and interacted with microRNAs. Our present work revealed a new kind of regulatory molecule that functions during physiological processes in cells, through targeting small non-coding RNAs.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4409/8/2/111/s1>.

Author Contributions: Conceptualization, P.L., J.W. and T.S.; methodology, T.S., J.W., K.W., C.L. and Y.W.; software, T.S.; validation, P.L., J.W. and T.S.; formal analysis, P.L., J.W. and T.S.; investigation, T.S., K.W. and C.L.; resources, P.L., J.W., T.S., K.W., C.L. and Y.W.; data curation, P.L. and T.S.; writing—original draft preparation, T.S.; writing—review and editing, P.L.; visualization, P.L. and T.S.; supervision, P.L., J.W.; project administration, P.L., J.W. and T.S.; funding acquisition, P.L. and T.S.

Funding: This research was funded by the National Natural Science Foundation of China (81800268, 31430041), supported by Applied Basic Research Program of Shanxi Province (201801D221268), Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (STIP)(201802060), and China Postdoctoral Science Foundation (2016M592134).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kuttler, F.; Mai, S. Formation of non-random extrachromosomal elements during development, differentiation and oncogenesis. *Semin. Cancer Biol.* **2007**, *17*, 56–64. [CrossRef] [PubMed]
2. Handa, H. Linear plasmids in plant mitochondria: Peaceful coexistences or malicious invasions? *Mitochondrion* **2008**, *8*, 15–25. [CrossRef] [PubMed]
3. Sinclair, D.A.; Guarente, L. Extrachromosomal rDNA circles—A cause of aging in yeast. *Cell* **1997**, *91*, 1033–1042. [CrossRef]
4. Huber, D.; Rustchenko, E. Large circular and linear rDNA plasmids in candida albicans. *Yeast (Chichester Engl.)* **2001**, *18*, 261–272. [CrossRef]

5. Casjens, S.; Palmer, N.; van Vugt, R.; Huang, W.M.; Stevenson, B.; Rosa, P.; Lathigra, R.; Sutton, G.; Peterson, J.; Dodson, R.J.; et al. A bacterial genome in flux: The twelve linear and nine circular extrachromosomal dnas in an infectious isolate of the lyme disease spirochete borrelia burgdorferi. *Mol. Microbiol.* **2000**, *35*, 490–516. [[CrossRef](#)] [[PubMed](#)]
6. Shadel, G.S.; Clayton, D.A. Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem.* **1997**, *66*, 409–435. [[CrossRef](#)] [[PubMed](#)]
7. Shibata, Y.; Kumar, P.; Layer, R.; Willcox, S.; Gagan, J.R.; Griffith, J.D.; Dutta, A. Extrachromosomal microdnas and chromosomal microdeletions in normal tissues. *Science (New York NY)* **2012**, *336*, 82–86. [[CrossRef](#)]
8. Cox, D.; Yuncken, C.; Spriggs, A.I. Minute chromatin bodies in malignant tumours of childhood. *Lancet (Lond. Engl.)* **1965**, *1*, 55–58. [[CrossRef](#)]
9. Hahn, P.J. Molecular biology of double-minute chromosomes. *BioEssays* **1993**, *15*, 477–484. [[CrossRef](#)]
10. Shimizu, N.; Ochi, T.; Itonaga, K. Replication timing of amplified genetic regions relates to intranuclear localization but not to genetic activity or g/r band. *Exp. Cell Res.* **2001**, *268*, 201–210. [[CrossRef](#)]
11. Turner, K.M.; Deshpande, V.; Beyter, D.; Koga, T.; Rusert, J.; Lee, C.; Li, B.; Arden, K.; Ren, B.; Nathanson, D.A.; et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **2017**, *543*, 122–125. [[CrossRef](#)] [[PubMed](#)]
12. Starling, S. Cancer genomics: Ecdetect hunts extrachromosomal DNA. *Nat. Rev. Genet.* **2017**, *18*, 212. [[CrossRef](#)] [[PubMed](#)]
13. Dillon, L.W.; Kumar, P.; Shibata, Y.; Wang, Y.H.; Willcox, S.; Griffith, J.D.; Pommier, Y.; Takeda, S.; Dutta, A. Production of extrachromosomal microdnas is linked to mismatch repair pathways and transcriptional activity. *Cell Rep.* **2015**, *11*, 1749–1759. [[CrossRef](#)] [[PubMed](#)]
14. Cara, A.; Reitz, M.S., Jr. New insight on the role of extrachromosomal retroviral DNA. *Leukemia* **1997**, *11*, 1395–1399. [[CrossRef](#)] [[PubMed](#)]
15. Matuszewski, B.; Ciechomski, K.; Kloc, M. Extrachromosomal rDNA and polarity of pro-oocytes during ovary development in creophilus maxillosus (coleoptera, staphylinidae). *Folia Histochem. Cytobiol.* **1999**, *37*, 179–190. [[PubMed](#)]
16. McDermott, P.; Connolly, V.; Kavanagh, T.A. The mitochondrial genome of a cytoplasmic male sterile line of perennial ryegrass (*Lolium perenne* L.) contains an integrated linear plasmid-like element. *TAG. Theor. Appl. Genet. Theor. Angew. Genetik* **2008**, *117*, 459–470. [[CrossRef](#)]
17. Rose, K.; Fetzner, S. Identification of linear plasmid pam1 in the flavonoid degrading strain actinoplanes missouriensis(t) (dsm 43046). *Plasmid* **2006**, *55*, 249–254. [[CrossRef](#)] [[PubMed](#)]
18. Palmer, N.; Fraser, C.; Casjens, S. Distribution of twelve linear extrachromosomal dnas in natural isolates of lyme disease spirochetes. *J. Bacteriol.* **2000**, *182*, 2476–2480. [[CrossRef](#)]
19. Sun, T.; Dong, Y.H.; Du, W.; Shi, C.Y.; Wang, K.; Tariq, M.A.; Wang, J.X.; Li, P.F. The role of micrnas in myocardial infarction: From molecular mechanism to clinical application. *Int. J. Mol. Sci.* **2017**, *18*, 745. [[CrossRef](#)]
20. Wang, K.; Sun, T.; Li, N.; Wang, Y.; Wang, J.X.; Zhou, L.Y.; Long, B.; Liu, C.Y.; Liu, F.; Li, P.F. Mdr1 Incrna regulates the processing of mir-484 primary transcript by targeting mir-361. *PLoS Genet.* **2014**, *10*, e1004467. [[CrossRef](#)]
21. Thum, T.; Gross, C.; Fiedler, J.; Fischer, T.; Kissler, S.; Bussen, M.; Galuppo, P.; Just, S.; Rottbauer, W.; Frantz, S.; et al. Microrna-21 contributes to myocardial disease by stimulating map kinase signalling in fibroblasts. *Nature* **2008**, *456*, 980–984. [[CrossRef](#)] [[PubMed](#)]
22. Bartel, D.P. Micrnas: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 281–297. [[CrossRef](#)]
23. Zhang, X.; Li, H.; Burnett, J.C.; Rossi, J.J. The role of antisense long noncoding rna in small rna-triggered gene activation. *RNA (New York NY)* **2014**, *20*, 1916–1928. [[CrossRef](#)] [[PubMed](#)]
24. Kim, D.H.; Saetrom, P.; Snove, O., Jr.; Rossi, J.J. Microrna-directed transcriptional gene silencing in mammalian cells. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 16230–16235. [[CrossRef](#)] [[PubMed](#)]
25. Kalsotra, A.; Wang, K.; Li, P.F.; Cooper, T.A. Micrnas coordinate an alternative splicing network during mouse postnatal heart development. *Genes Dev.* **2010**, *24*, 653–658. [[CrossRef](#)] [[PubMed](#)]
26. Allo, M.; Buggiano, V.; Fededa, J.P.; Petrillo, E.; Schor, I.; de la Mata, M.; Agirre, E.; Plass, M.; Eyraes, E.; Elela, S.A.; et al. Control of alternative splicing through sirna-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* **2009**, *16*, 717–724. [[CrossRef](#)] [[PubMed](#)]

27. Zisoulis, D.G.; Kai, Z.S.; Chang, R.K.; Pasquinelli, A.E. Autoregulation of microRNA biogenesis by let-7 and argonaute. *Nature* **2012**, *486*, 541–544. [[CrossRef](#)]
28. Faas, F.G.; Rieger, B.; van Vliet, L.J.; Cherny, D.I. DNA deformations near charged surfaces: Electron and atomic force microscopy views. *Biophys. J.* **2009**, *97*, 1148–1157. [[CrossRef](#)]
29. Rowley, J.D.; Diaz, M.O.; Espinosa, R., 3rd; Patel, Y.D.; van Melle, E.; Ziemins, S.; Taillon-Miller, P.; Lichter, P.; Evans, G.A.; Kersey, J.H.; et al. Mapping chromosome band 11q23 in human acute leukemia with biotinylated probes: Identification of 11q23 translocation breakpoints with a yeast artificial chromosome. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9358–9362. [[CrossRef](#)]
30. Fu, H.; Le, S.; Chen, H.; Muniyappa, K.; Yan, J. Force and atp hydrolysis dependent regulation of reca nucleoprotein filament by single-stranded DNA binding protein. *Nucleic Acids Res.* **2013**, *41*, 924–932. [[CrossRef](#)]
31. Ha, T.; Kozlov, A.G.; Lohman, T.M. Single-molecule views of protein movement on single-stranded DNA. *Annu. Rev. Biophys.* **2012**, *41*, 295–319. [[CrossRef](#)] [[PubMed](#)]
32. Segal, E.; Fondufe-Mittendorf, Y.; Chen, L.; Thastrom, A.; Field, Y.; Moore, I.K.; Wang, J.P.; Widom, J. A genomic code for nucleosome positioning. *Nature* **2006**, *442*, 772–778. [[CrossRef](#)] [[PubMed](#)]
33. Gaasenbeek, M.; Howarth, K.; Rowan, A.J.; Gorman, P.A.; Jones, A.; Chaplin, T.; Liu, Y.; Bicknell, D.; Davison, E.J.; Fiegler, H.; et al. Combined array-comparative genomic hybridization and single-nucleotide polymorphism-loss of heterozygosity analysis reveals complex changes and multiple forms of chromosomal instability in colorectal cancers. *Cancer Res.* **2006**, *66*, 3471–3479. [[CrossRef](#)] [[PubMed](#)]
34. Creighton, M.P.; Cheng, A.W.; Welstead, G.G.; Kooistra, T.; Carey, B.W.; Steine, E.J.; Hanna, J.; Lodato, M.A.; Frampton, G.M.; Sharp, P.A.; et al. Histone h3k27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21931–21936. [[CrossRef](#)] [[PubMed](#)]
35. Johnnidis, J.B.; Harris, M.H.; Wheeler, R.T.; Stehling-Sun, S.; Lam, M.H.; Kirak, O.; Brummelkamp, T.R.; Fleming, M.D.; Camargo, F.D. Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. *Nature* **2008**, *451*, 1125–1129. [[CrossRef](#)] [[PubMed](#)]
36. Sloan, R.D.; Wainberg, M.A. The role of unintegrated DNA in HIV infection. *Retrovirology* **2011**, *8*, 52. [[CrossRef](#)] [[PubMed](#)]
37. Schnepf, B.C.; Jensen, R.L.; Chen, C.L.; Johnson, P.R.; Clark, K.R. Characterization of adeno-associated virus genomes isolated from human tissues. *J. Virol.* **2005**, *79*, 14793–14803. [[CrossRef](#)] [[PubMed](#)]
38. Schneider, W.M.; Wu, D.T.; Amin, V.; Aiyer, S.; Roth, M.J. Mulv in mutants responsive to hdac inhibitors enhance transcription from unintegrated retroviral DNA. *Virology* **2012**, *426*, 188–196. [[CrossRef](#)]
39. Weiss, G.; Schaible, U.E. Macrophage defense mechanisms against intracellular bacteria. *Immunol. Rev.* **2015**, *264*, 182–203. [[CrossRef](#)]
40. Lecellier, C.H.; Dunoyer, P.; Arar, K.; Lehmann-Che, J.; Eyquem, S.; Himber, C.; Saib, A.; Voinnet, O. A cellular microRNA mediates antiviral defense in human cells. *Science (New York NY)* **2005**, *308*, 557–560. [[CrossRef](#)]
41. Stinglele, J.; Schwarz, M.S.; Bloemeke, N.; Wolf, P.G.; Jentsch, S. A DNA-dependent protease involved in DNA-protein crosslink repair. *Cell* **2014**, *158*, 327–338. [[CrossRef](#)] [[PubMed](#)]

