Supporting information

A self-replicating linear DNA

Pei Sheng Liew [1], Tze Hao Tan [2], Yin Cheng Wong [1], Edmund Ui Hang Sim [3], Choon Weng Lee [4], Kumaran Narayanan [1]*

[1] School of Science, Monash University Malaysia, Bandar Sunway, Malaysia
[2] Faculty of Science, Kyushu University, Ito campus, Fukuoka, Japan
[3] Faculty of Resource Sciences and Technology, University Malaysia Sarawak, Malaysia
[4] Institute of Biological Science, Faculty of Science, University of Malaya, Malaysia

* To whom correspondence should be addressed:
Kumaran Narayanan
Address: Monash University Malaysia
School of Science
Bandar Sunway, 47500
Selangor Darul Ehsan
Malaysia
Fax: 6 03 5514 6364
E-mail: kumaran.narayanan@monash.edu

Content:

Extended discussion

TelN-expressing cells have lower transfection efficiency than parental cells

Supporting methods

Construction of S/MAR-containing *tos*-BAC

In vivo tos-DNA linearization assay

Flow cytometry analysis for quantification of transfection efficiency of mammalian cells

Supporting references

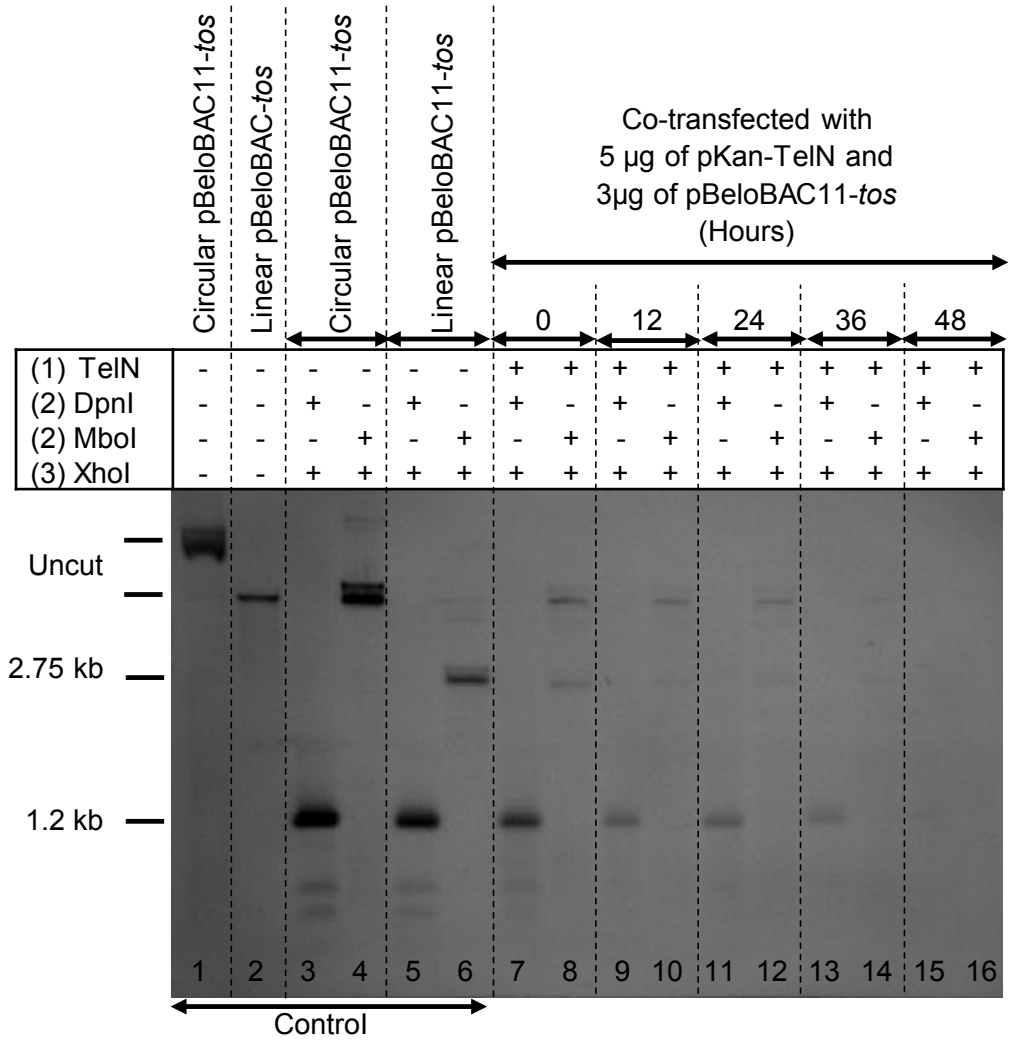


Figure S1. TelN did not help *tos*-DNA containing bacterial *ori* replicated in mammalian cells.

After isolation from co-transfected TelN-expressing HeLa cells, *tos*-DNA (pBeloBAC11-*tos*) was digested by either DpnI or MboI, prior to XhoI enzyme treatment. Absence of 2.75 kb or 1.2 kb fragments under respective DpnI or MboI treatment showed no replication of *tos*-DNA containing only bacterial *ori*. Representation of non-replicated *tos*-DNA in mammalian cells detected by Cm^R hybridization probe in Southern blotting.

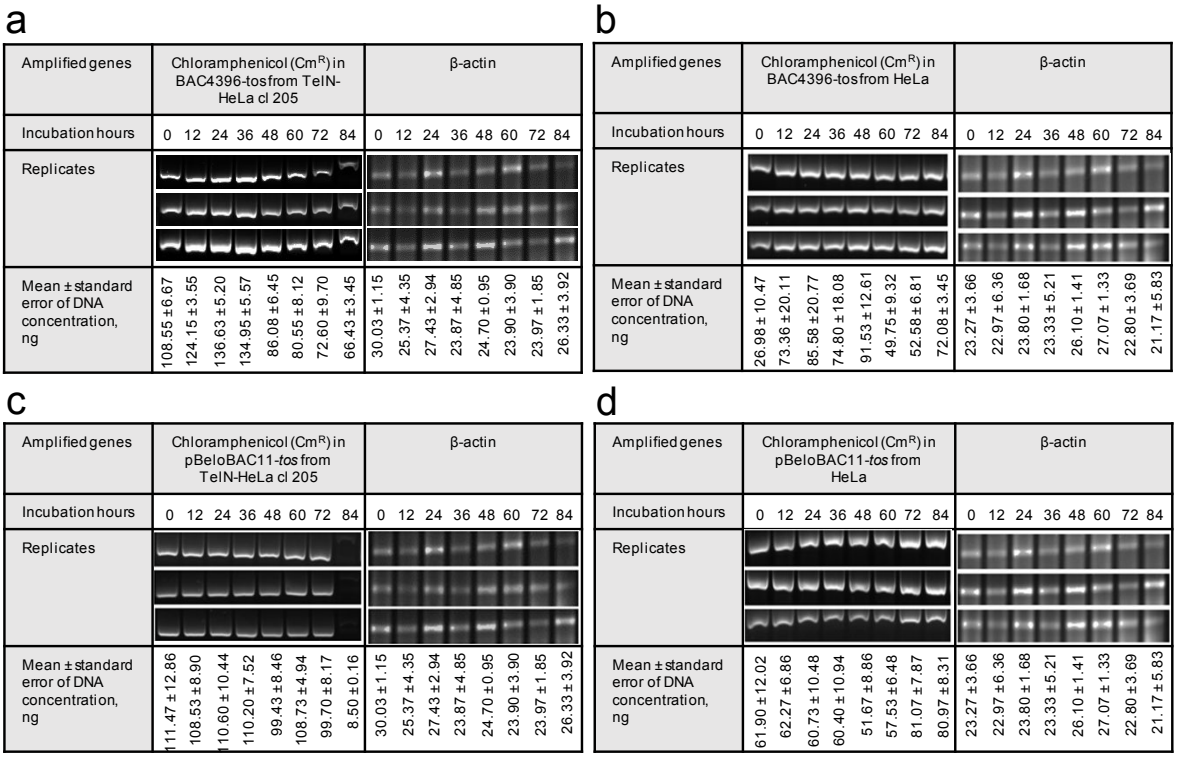


Figure S2. Mammalian-initiation-site, TelN or bacterial *ori* alone did not initiate linear *tos*-DNA replication event.

Gel-analyzed elevation of the *tos*-DNA concentration in the presence of (a) mammalian-initiation-site and TelN expression, (b) mammalian-initiation-site only, without TelN expression in HeLa cells, (c) bacterial *ori* with TelN expression in HeLa cells, and (d) bacterial *ori* only, without TelN expression in HeLa cells. Total DNAs were isolated at different time-points (as shown in incubation hours) after the transient transfection of *tos*-DNA. Relative intensities of amplified Cm^R bands were normalized to housekeeping β-actin gene and quantitated using agarose gel electrophoresis. Band intensities and relative DNA concentration were calculated using program ImageLab software version 5.1.

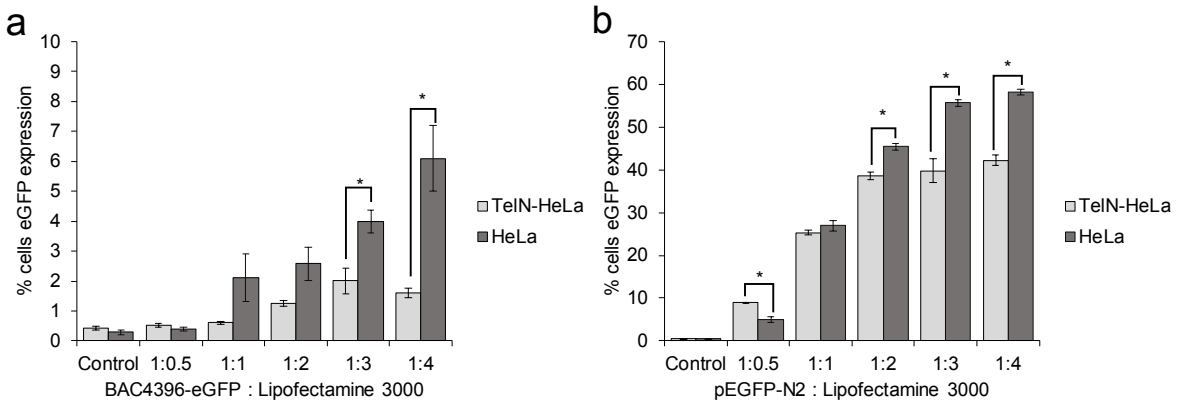


Figure S3. TelN does not increase transfection efficiency of mammalian cells.
TelN-expressing HeLa cells has lower transfection efficiency than parental HeLa cells. **(a)** BAC4396-eGFP, or **(b)** pEGFP-N2 was transfected into TelN-HeLa cl 205 and HeLa cells, respectively, at DNA: Lipofectamine 3000 complex ratio of 1:0.5, 1:1, 1:2, 1:3, and 1:4. Percentage of green fluorescent expression in the cells was determined by flow cytometry analysis. Significance was determined using Student *t*-test; asterisk denotes significant different between two groups of sample studies, data represent mean ± standard error, $p \leq 0.05$, $n = 3$.

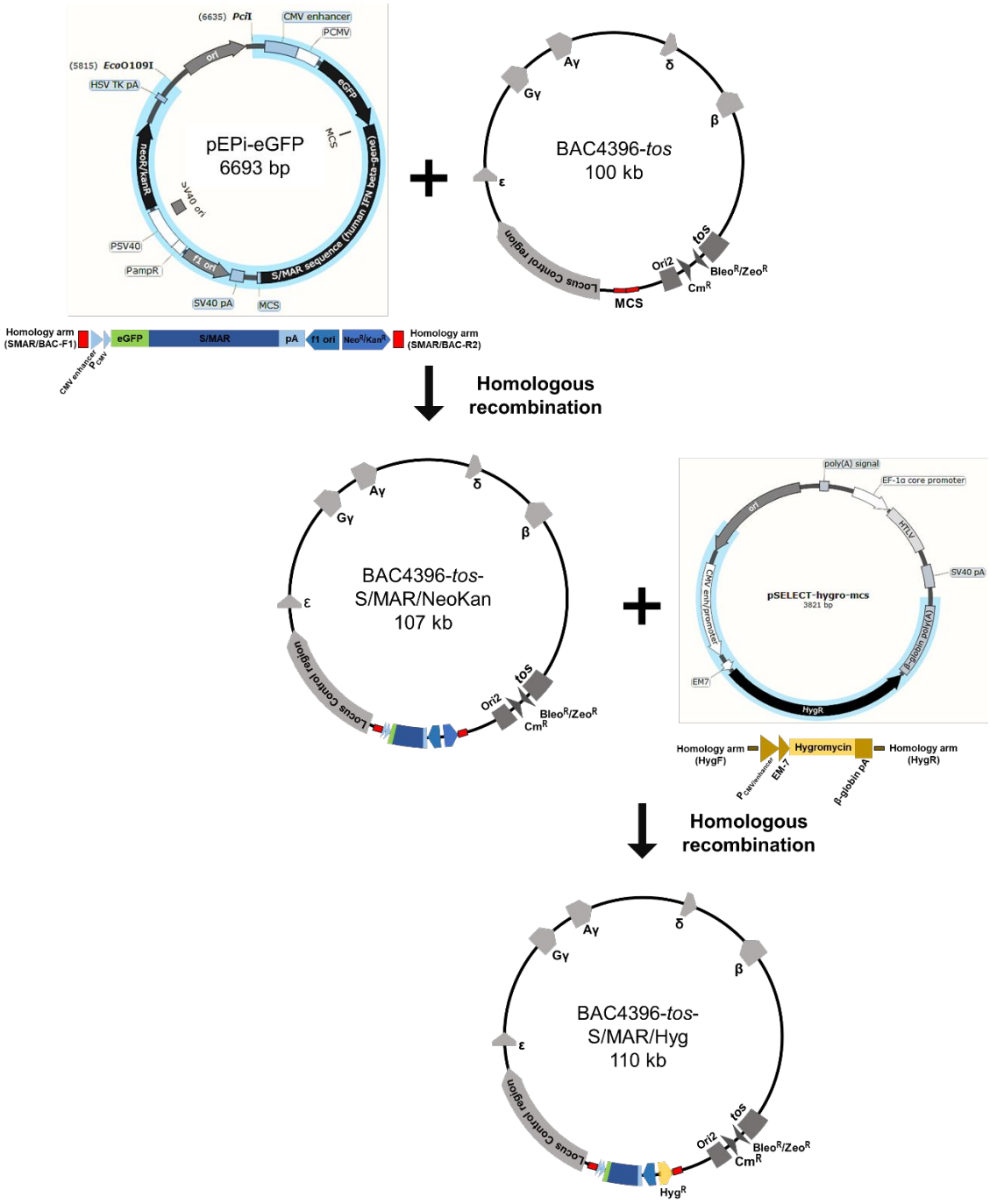
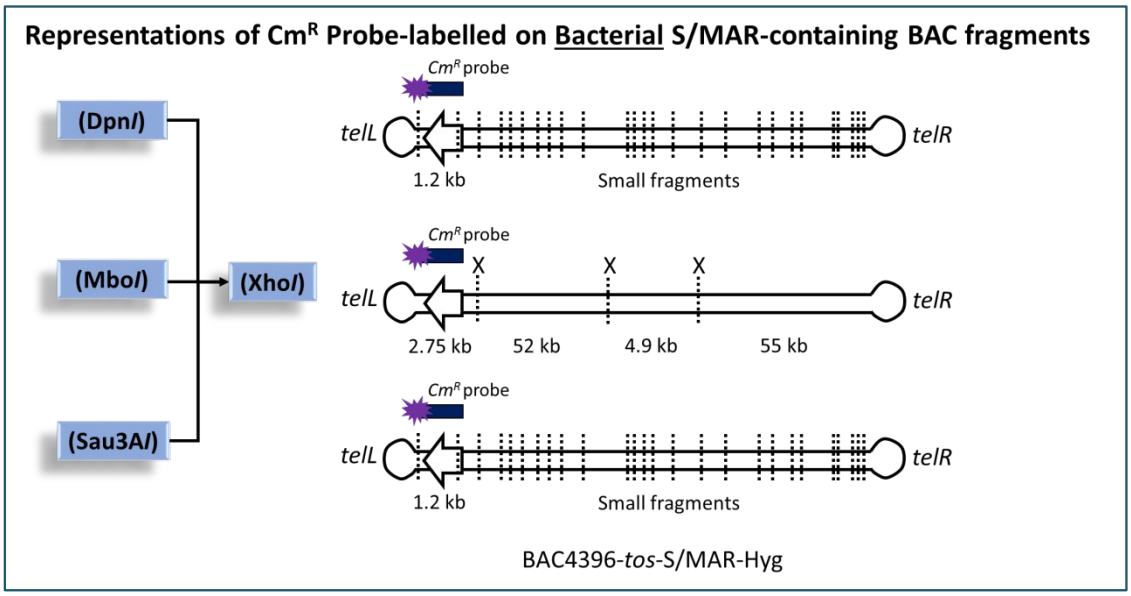


Figure S4. Construction of S/MAR-containing BAC4396-*tos*.

Two-steps homologous recombination are involved to construct S/MAR-contaning BAC4396-*tos*. The first step is to construct an intermediate BAC4396-*tos*-S/MAR/NeoKan from pEPi-eGFP and BAC4396-*tos*. The 6002 bp amplified product with 60 bp homology arms at 3' and 5' ends, consisted of eGFP gene, S/MAR and neo^R/kan^R cassette from pEPi-eGFP was recombined into the homology region of MCS of 100 kb BAC4396-*tos*, to form the intermediate 107 kb BAC4396-*tos*-S/MAR/NeoKan. The second step is to construct the final product, BAC4396-*tos*-S/MAR/Hyg by replacing the selection antibiotic. The 2340 bp hygromycin cassette from pSELECT-hygro-mcs replaced the 1269 bp neomycin/kanamycin cassette from BAC4396-*tos*-S/MAR/NeoKan by recombination to give the 110 kb BAC4396-*tos*-S/MAR/Hyg, which consists necessary elements for retention and selection.

a



b

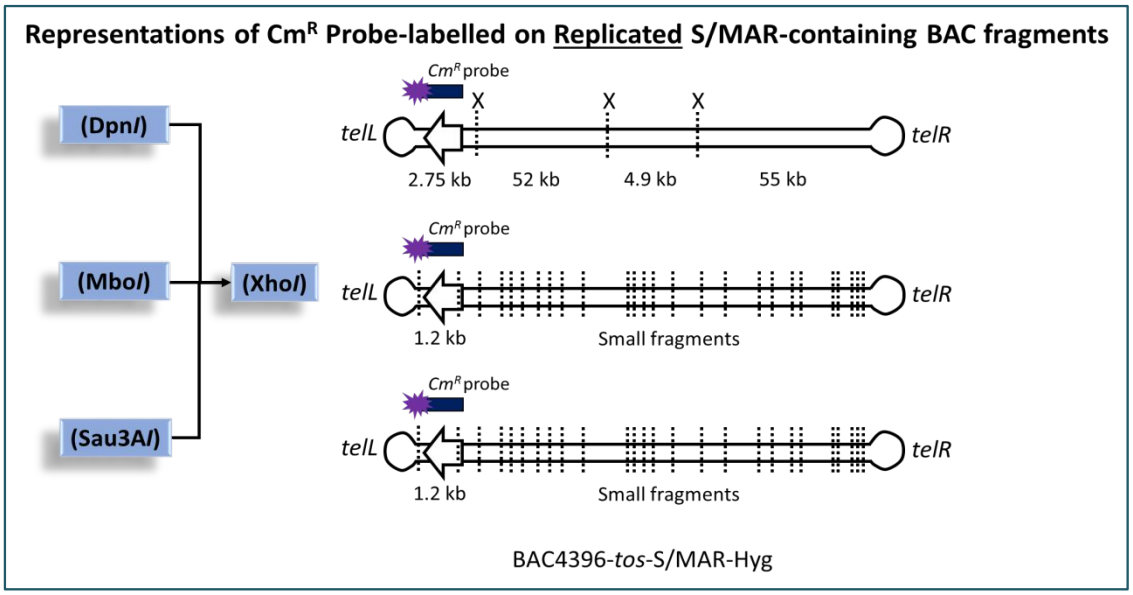


Figure S5. Representation of non-replicated and replicated *tos*-BAC in mammalian cells and labelled by Southern Cm^R probe.

Total DNA were isolated at different time-points after the transient transfection of linear BAC4396-*tos*-S/MAR/Hyg into TelN-NIH3T3 clone 69, followed *by* DpnI / MboI / Sau3AI and XhoI treatments. Southern hybridization using Cm^R probe to detect fragments that represent the (a) non-replicated and (b) replicated BAC4396-*tos*-S/MAR-Hyg.

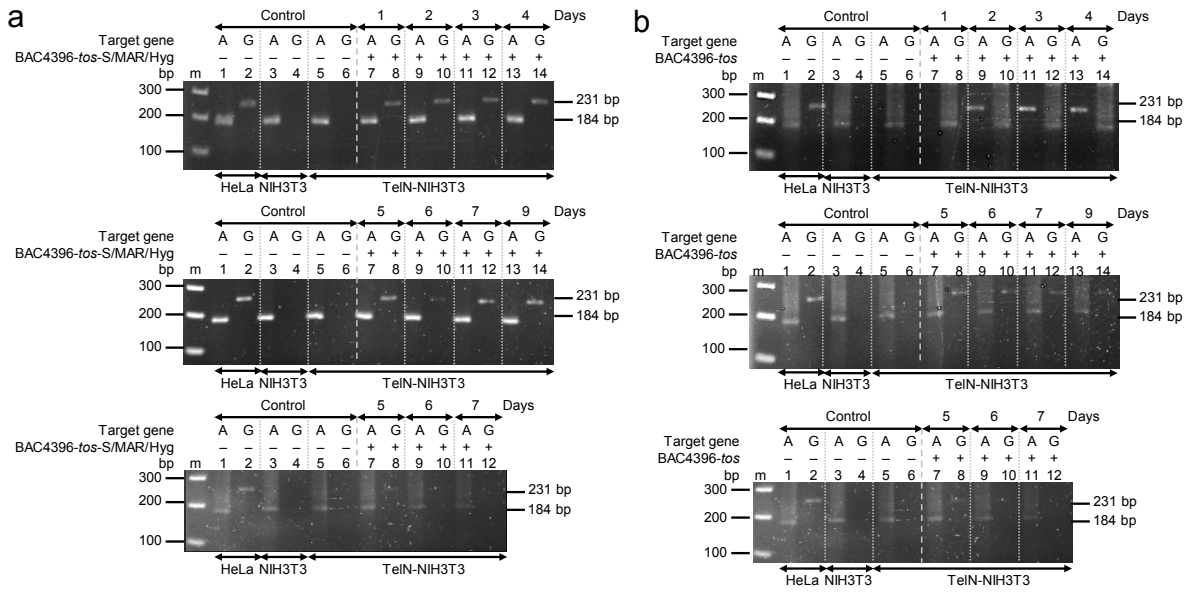


Figure S6. Human S/MAR enhances transient β-globin expression for longer period.

Accurately spliced β-globin mRNA was expressed for longer period by (a) a 110 kb S/MAR-containing BAC4396-*tos* compared to (b) 100 kb BAC4396-*tos* in TelN-expressing NIH3T3 mouse cells. Total RNAs were isolated at different time-points (as shown in days) after the transient transfection of BAC4396-*tos*-S/MAR/Hyg into TelN-NIH3T3 clone 69, reversed-transcribed and PCR amplified with specific primers HBB-e2F and HBB-e3R for β-globin (G) region in BAC4396-*tos*-S/MAR/Hyg, and primers ACTB-FWD and ACTB-REV for β-actin (A) region in chromosomal cDNA. Lanes marked "+" denote cells were transfected with BAC4396-*tos*-S/MAR/Hyg, while lanes marked "-" denote cells were not transfected with BAC4396-*tos*-S/MAR/Hyg. Control RT-PCR of β-globin transcript was carried out on RNA isolated from endogenous β-globin expressive cell (HeLa) and non-endogenous β-globin expressive cells (NIH3T3 and TelN-NIH3T3). Gene expression of β-globin and β-actin in TelN-expressing mouse cell were determined by 2.5% gel stained with EtBr.

Table S1. List of primers

Primer Name	Sequence
Cm^R-FWD	5'- ATGGAGAAAAAAATCACTGGAT
Cm^R-REV	5'- CTGCCACTCATCGCAGTACT
SMAR/BAC-F1	5'- GCATGCGGATCCCCGGGTACCGAGCTCGAATTCGCCCTATAGTGAGTCGTATTACAATTCTTCTTTCCTGCGTTATCCCC
SMAR/BAC-R2	5'- CGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCTATGGCAGGGCCTGCCGC
BAC/SMAR-AF	5'- GGCGAAAGGGGGATGTGCTGCAAGGCG
BAC/SMAR-AR2	5'- GCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGG
HygF	5'- CCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCGGAGCCTATGGAAAAACGC
HygR	5'- CAGAGTGCCAGCCCTGGGACCGAACCCCGCGTTTATGAACAAACGACCCAACACCGTGCGGGTTCAGGGGGAGGTGTGGG
S-HygF2	5'- GCGGAAAGAACCAGCTGTGG
S-HygR2	5'- GGGCCTTCACCCGAACTTGG
HBB-e2F	5'- GGACCCAGAGGTTCTTTGAGTCC
HBB-e3R	5'- GCACACAGACCAGCACGTTGCCC
ACTB-FWD	5'- AGAGCTACGAGCTGCCTGAC
ACTB-REV	5'- AGCACTGTGTTGGCGTACAG

*Underlined characters indicate homologous nucleotide sequence to either 5' or 3' end of a targeted site on a DNA vector.

Incubation hours	**0**	**12**	**24**	**36**	**48**	**60**	**72**	**84**
Replicate 1	2.549	4.546	4.887	4.380	4.421	2.675	3.303	3.921
Replicate 2	1.461	4.929	5.202	4.791	4.521	2.453	2.395	3.506
Replicate 3	0.628	3.144	3.769	3.073	3.031	1.116	1.855	3.019
Replicate 4	1.707	3.842	1.691	3.144	2.111	1.986	2.502	3.840
Replicate 5	2.661	5.326	2.351	4.188	2.535	2.282	2.820	4.029
Replicate 6	2.690	5.578	2.433	4.401	2.601	2.383	2.844	4.143
Replicate 7	1.319	2.927	1.331	2.311	1.386	1.165	1.502	2.283
Replicate 8	0.958	1.716	2.063	2.747	0.341	1.108	1.671	3.175
Replicate 9	1.281	2.160	2.601	2.974	0.954	0.399	0.671	2.806
Replicate 10	1.573	1.890	2.113	3.214	0.398	0.222	0.351	1.583

Table S2. The table tabulated Cm^R gene concentration of isolated total DNA from HeLa cells with BAC4396-*tos*. One-way ANOVA with single factor statistical analyses was performed to determine Cm^R gene concentration (Equation 1) elevation over eight time-points (Incubation hours) in HeLa cells (no TelN) with linear BAC4396-*tos*. The first three rows (Replicates 1 – 3) were previous data in Figure 2b that resulted in high standard error plot in the original manuscript. Extra seven replicates (Replicates 4 – 10) of the semi-quantitative PCR experiment were performed to support the finding in the revised Figure 3b.

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
0	10	16.82618	1.682618	0.523961
12	10	36.05566	3.605566	2.091178
24	10	28.44056	2.844056	1.757642
36	10	35.22461	3.522461	0.707732
48	10	22.29971	2.229971	2.244453
60	10	15.78758	1.578758	0.790322
72	10	19.91467	1.991467	0.928078
84	10	32.30463	3.230463	0.692669

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	47.0926	7	6.727515	5.527929	4.07125E-05	2.139656
Within Groups	87.6243	72	1.217005			
Total	134.7169	79				

Table S3. Statistical data showed there was no significant replication of BAC4396-*tos* in HeLa cells (no TelN). One-way ANOVA with single factor statistical analysis (p value < 0.05) showed that there was no significant elevation in Cm^R gene concentration within the eight tested incubation time-points.

Extended discussion

TelN-expressing cells have lower transfection efficiency than parental cells

In the PCR amplification study, we observed that Cm^R concentration of *tos*-DNA normalized by β-actin (labeled as Cm^R/β-actin ratio in Figure 3) differed at 0-hour in the TelN-HeLa (Figure 3a and c) and HeLa (no TelN) (Figure 3b and d) cells.

To investigate whether TelN expression in mammalian cells affects the transfer rate of *tos*-DNA, we compared the transfection efficiency of TelN-HeLa and HeLa (no TelN) by flow cytometry. To obtain a comparable transfection rate to the *tos*-DNA substrate (100 kb BAC4396-*tos*) in this study, a ≃100 kb DNA substrate with green fluorescence protein gene (*eGFP*), BAC4396-eGFP was used. BAC4396-eGFP was transfected into TelN-HeLa and HeLa (no TelN) and incubated for 24-hours before measuring eGFP-expressing cells. Higher percentage of eGFP-expressing cells was observed in HeLa (no TelN) compared to TelN-HeLa at all tested DNA: transfection-reagent ratios (Figure S3a), showing that TelN reduced the transfection rate of cells.

A second confirmation test on TelN-HeLa and HeLa (no TelN) transfection efficiency was demonstrated by transfecting a smaller eGFP-expressing DNA, the 4.7 kb plasmid pEGFP-N2. A similar fluorescent trend was obtained at all tested DNA: transfection-reagent ratios, where higher percentage of eGFP-expression was seen in HeLa (no TelN) than TelN-HeLa (Figure S3b). The flow cytometry data showed that TelN reduced transfection efficiency of HeLa cells. Thus, the higher Cm^R concentration at 0-hour in the TelN-HeLa (Figure 3a and c) was not due to higher uptake of *tos*-DNA.

Nevertheless, we believe that the phenomena of higher Cm^R/β-actin ratio seen at 0-hour in TelN-HeLa cells is more likely the result of early replication of *tos*-DNA in mammalian cells by TelN. It suggests that the higher Cm^R/β-actin ratio in TelN-HeLa is

not likely caused by the inherent increase in transfer efficiency of *tos*-DNA under TelN expression.

A study by Chandok et al. reported an unusual plasmid replication mode that occurred only at the first replication cycle, where it was started at random positions on the DNA introduced into mammalian cells [1]. Specifically, the process of the DNA replication is initiated at the G1 phase when an origin recognition complex (ORC) is bound to *ori* [2,3]. Accumulation of pre-replication complex (pre-RC), mini-chromosome-maintenance-complex (MCM) and other proteins to ORC form an initiation complex during cell cycle transition from G1 to S phase [4,5]. Upon activation of cyclin-dependent kinase (CDK) in S phase, the initiation complex unwinds DNA double helix and allows the first primer synthesis by DNA polymerase α – primase complex [6]. Elongation of the leading and lagging strands are then facilitated by DNA polymerase δ and ε [7]. More importantly, such replication mode particularly favored at single-stranded regions and assumed to have initiated chromosomal replication within immature chromatin structure.

We assume that the accumulation of pre-replication complex (pre-PC) activates checkpoint mechanism in cell cycle to coordinate S-phase for *tos*-DNA to undergo the alternative replication before entering mitosis. Therefore, during the first cycle replication, the circular head-to-head and tail-to-tail intermediates which are TelN-resolvable, are formed. Such replicated *tos*-DNA in TelN-expressing cells then resulted in higher initial Cm^R/β-actin ratio that was recorded at 0-hour of incubation (Figure 3a and c).

Supporting methods:

Construction of S/MAR-containing *tos*-BAC

S/MAR-containing BAC termed BAC4396-*tos*-S/MAR/Hyg is specifically constructed to provide long-term retention to the *tos*-DNA substrate in mammalian cells for this study. It is a 110 kb BAC and contains mammalian CMV enhancer and promoter to drive eGFP gene upstream of the S/MAR to provide an active upstream transcription running into S/MAR, a requirement for sufficient episomal vector replication in mammalian cells. BAC4396-*tos*-S/MAR/Hyg was constructed on the backbone of BAC4396-*tos* [8], via a two-steps homologous recombination strategy [9,10]. Firstly, the eGFP gene running upstream of S/MAR and neomycin/kanamycin resistance cassette from plasmid pEPi-eGFP [11] were recombined into multiple cloning site of BAC4396-*tos* to generate a BAC4396-*tos*-S/MAR/NeoKan intermediate BAC. The neomycin/kanamycin gene in BAC4396-*tos*-S/MAR/NeoKan was then replaced with hygromycin gene cassette from pSELECT-hygro-mcs (Invivogen). Details of primers described here are listed in Table S1.

To construct the intermediate BAC4396-*tos*-S/MAR/NeoKan (Figure S4), the 6002 bp DNA sequence consisted of *eGFP* gene, S/MAR and neomycin/kanamycin resistance cassette were amplified from pEPi-eGFP using primers SMAR/BAC-F1 and SMAR/BAC-R2, to include 60bp homologous regions of MCS of BAC4396-*tos* that flanked at 3' and 5' ends of the amplified sequence. The PCR product was gel-purified and transformed into fresh electro-competent *E. coli* strain DH10B (pGETrec, BAC4396-*tos*) and selected by the antibiotics zeocin, chloramphenicol and kanamycin. Positive clones were confirmed of their identities by PCR screening using primers BAC/SMAR-AF and BAC/SMAR-AR2 to check for the inserted DNA sequence.

Next, to construct BAC4396-*tos*-S/MAR/Hyg (Figure S4), a 2340 bp hygromycin cassette was amplified from pSELECT-hygro-mcs (Invivogen) using primers HygF and HygR to include additional 60 bp homologous region of neomycin/kanamycin cassette of BAC4396-*tos*-S/MAR/NeoKan, flanking at 3' and 5' ends of amplified product. The amplified product was gel-purified and electroporated into electro-competent *E. coli* strain DH10B (pGETrec, BAC4396-*tos*-S/MAR/NeoKan) and selected by using antibiotics - zeocin, chloramphenicol and hygromycin. Identities of positive clones were confirmed by PCR screening using primers S-HygF2 and S-HygR2 to confirm successful replacement of neomycin/kanamycin by hygromycin cassette.

To remove plasmid pGETrec from DH10B (pGETrec, BAC4396-*tos*-S/MAR/Hyg), the DH10B (pGETrec, BAC4396-*tos*-S/MAR/Hyg) was grown for 48 hours with 150rpm shaking at 30 ˚C in BHI media supplemented with zeocin, chloramphenicol and hygromycin. Plasmid and BAC were extracted and purified using NucleoBond® Xtra Midiprep kit (Macherey-Nagel). Purified DNA was re-transformed into fresh electro-competent *E. coli* DH10B cells and selected by antibiotics zeocin, chloramphenicol and hygromycin. Verification of BAC4396-*tos*-S/MAR/Hyg was determined by restriction digestion analysis and DNA sequencing (result not shown).

In vivo tos-DNA linearization assay

TelN-expressing NIH3T3 (TelN-NIH3T3 clone 69) cells (1.8×10^6 cells/10 cm dish) were seeded and incubated overnight at 37 °C with 5 % CO_2 condition. Twenty micrograms of circular BAC4396-*tos*-S/MAR-Hyg were transfected into cells using Lipofectamine 3000 (Invitrogen) at ratio 1:3. Total DNA was isolated at 24 hours post-transfection using Qiagen Plasmid Miniprep Kit to determine the structure of the BAC. Isolated DNA was subjected to RecBCD and XhoI enzyme digestions. Presence of the linearized DNA was

verified by 1% agarose gel electrophoresis and Southern-hybridized using Cm^R-probe (as described in **MATERIALS AND METHODS: Enzyme DpnI / MboI / Sau3AI assays for qualitative replication analysis**). Triplicates were included in this experiment.

Flow cytometry analysis for quantification of transfection efficiency of mammalian cells

TelN-expressing HeLa (TelN-HeLa) or HeLa (no TelN) cells (2×10^5 cells/well of 6-well plate) were seeded and incubated overnight at 37 °C with 5 % CO_2 condition. The cells were transfected with of 2 µg circular BAC4396-eGFP or pEGFP-N_2 by using Lipofectamine 3000 (Invitrogen) at ratio 1:0.5, 1:1, 1:2, 1:3 and 1:4 for an hour, following the manufacturer protocol. After a 1×PBS washing step, cells were incubated in RPMI supplemented with 10 % fetal bovine serum and 200 µg/ml G418 (for TelN-expressing cells only). After 24 hours post-transfection, cells were trypsinized and resuspended in 1 ml of double-filtered 1×PBS. Cell counting was performed using a BD FACSCalibur™ flow cytometer with CellQuest Pro software. Data significance was determined using Student *t*-test; asterisk is used to denote significant difference between two groups of sample studies, $p \leq 0.05$, $n = 3$.

Supporting references:

(1) Chandok, G. S., Kapoor, K. K., Brick, R. M., Sidorova, J. M., and Krasilnikova, M. M. (2011) A distinct first replication cycle of DNA introduced in mammalian cells. *Nucleic Acids Res. 39*, 2103–2115.
(2) Clark, D. P., and Pazdernik, N. J. (2013) Chapter 10 - Cell Division and DNA Replication. In *Molecular Biology* (Clark, D., Ed. and Pazdernik, N. J., Ed.) 2nd ed., pp 274–308. Academic Press, Boston.
(3) Bhagavan, N. V., and Ha, C.-E. (2015) Chapter 22 - DNA Replication, Repair, and Mutagenesis. In *Essentials of Medical Biochemistry* (Bhagavan, N. V., and Ha, C.E., Ed.) 2nd ed., pp 401–417. Academic Press, San Diego.

(4) Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T. J., and Walter, J. C. (2003) Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev. 17*, 1894–1908.
(5) Sclafani, R. A., and Holzen, T. M. (2007) Cell Cycle Regulation of DNA Replication. *Annu. Rev. Genet. 41*, 237–280.
(6) Maiorano, D., Lutzmann, M., and Méchali, M. (2006) MCM proteins and DNA replication. *Curr. Opin. Cell Biol. 18*, 130–136.
(7) Burgers, P. M. J. (2009) Polymerase Dynamics at the Eukaryotic DNA Replication Fork. *J. Biol. Chem. 284*, 4041–4045.
(8) Ooi, Y.-S., Warburton, P. E., Ravin, N. V., and Narayanan, K. (2008) Recombineering linear DNA that replicate stably in *E. coli*. *Plasmid 59*, 63–71.
(9) Narayanan, K., Williamson, R., Zhang, Y., Stewart, A. F., and Ioannou, P. A. (1999) Efficient and precise engineering of a 200 kb beta-globin human/bacterial artificial chromosome in *E. coli* DH10B using an inducible homologous recombination system. *Gene Ther. 6*, 442–447.
(10) Narayanan, K., and Warburton, P. E. (2003) DNA modification and functional delivery into human cells using Escherichia coli DH10B. *Nucleic Acids Res. 31*, e51.
(11) Stehle, I. M., Scinteie, M. F., Baiker, A., Jenke, A. C. W., and Lipps, H. J. (2003) Exploiting a minimal system to study the epigenetic control of DNA replication: the interplay between transcription and replication. *Chromosome Res. 11*, 413–421.