Shortened nuclear matrix attachment regions are sufficient for replication and maintenance of episomes in mammalian cells

Xiao-Yin Wang[a,b,†], Xi Zhang[b,†], Tian-Yun Wang[a,b,*], Yan-Long Jia[c], Dan-Hua Xu[a,b], and Dan-Dan Yi[a,b]

[a]Department of Biochemistry and Molecular Biology, [b]International Joint Research Laboratory for Recombinant Pharmaceutical Protein Expression System of Henan, and [c]Pharmacy Collage, Xinxiang Medical University, Xinxiang 453003, Henan, China

ABSTRACT Matrix attachment regions (MARs) can mediate the replication of vector episomes in mammalian cells; however, the molecular mode of action remains unclear. Here, we assessed the characteristics of MARs and the mechanism that mediates episomal vector replication in mammalian cells. Five shortened subfragments of β-interferon MAR fragments were cloned and transferred into CHO cells, and transgene expression levels, presence of the gene, and the episomal maintenance mechanism were determined. Three shortened MAR derivatives (position 781–1320, 1201–1740, and 1621–2201) retained full MAR activity and mediated episomal vector replication. Moreover, the three shortened MARs showed higher transgene expression levels, greater efficiency in colony formation, and more persistent transgene expression compared with those of the original pEPI-1 plasmid, and three functional truncated MARs can bind to SAF-A MAR-binding protein. These results suggest that shortened MARs are sufficient for replication and maintenance of episomes in CHO cells.

INTRODUCTION

Gene therapy is achieved by transferring vectors that deliver a gene of interest (GOI) into specific cells to treat disease. An ideal gene therapy vector must deliver therapeutic levels of GOI expression sustainability, without safety and reproduction concerns (Lin *et al.*, 2015; Ma *et al.*, 2017). Virus-based integrating vectors, which can integrate into the host genome and thus are stably retained during subsequent cell divisions to maintain the stability of transgene expression, are widely used in gene therapy. The vector that integrated into the genome DNA can cause insertional mutagenesis (Kohn *et al.*, 2003; Browning and Trobridge 2016) or safety issues because of the possibility of inactivating tumor suppressor genes or activating oncogenes during the process. Moreover, the integrated GOI are subject to position effects and silencing (Ellis, 2005), making their expression unpredictable. Therefore, novel vectors needed to be developed with the aim of maintaining stable expression of foreign genes and developing a novel strategy for the creation of safer vectors for gene therapy.

Nonintegrating vectors are maintained as nuclear episomes and therefore do not physically integrate into genomes; thus, they are a relatively safe alternative for use in gene therapy (Chen *et al.*, 2017; Verghese and Kurre 2017; Sánchez-Hernández *et al.*, 2018). One promising type of nonviral, nonintegrating gene therapy vector system is the construct based on matrix attachment regions (MARs) elements, which are designed to persist as extrachromosomal entities in host cells (Hagedorn *et al.*, 2017; Wang *et al.*, 2017). MAR has shown the ability to remain episomal and be mitotically stable during cell division (Baiker *et al.*, 2000; Jenke *et al.*, 2004a; Argyros *et al.*, 2011a). IS2 element, which harbors SARs and HS4 sequences, can increase the percentage and expression levels of the transgene in several cell lines (Sánchez-Hernández *et al.*, 2018). The original
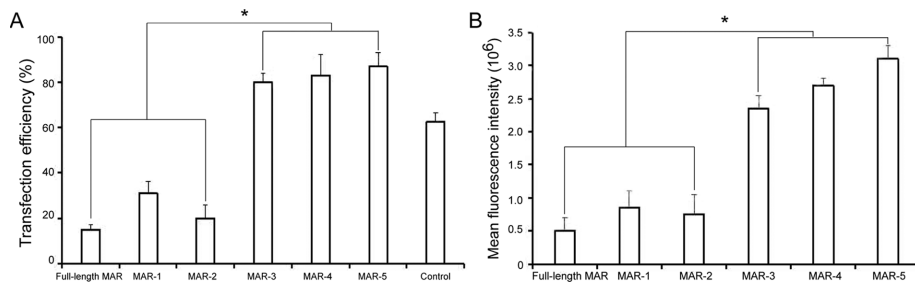
FIGURE 1: Transfection efficiency (A) and transient expression (B) in transfected CHO cells. Plasmids with shortened MAR 1–5 were transfected into CHO cells, respectively. Transfection efficiency and *eGFP* transient expression were determined using flow cytometry after transfection for 48 h. Three independent experiments were performed in this study. SEM is indicated (Student's *t* test,*P* < 0.05). The results indicated that all shortened MAR could enhance the transfection efficiency and transient expression. The transfection efficiency of plasmid with the 540-base-pair DNA random sequence were no difference compared with shortened MAR.

episomal vector pEPI, which contained a chromosomal MAR derived from the β-interferon (β-IFN) gene, was constructed by Piechaczek *et al.* (1999). pEPI can replicate episomally with a copy number of ~5–10 molecules per cell and is stably retained without selection pressure, thus allowing long-term expression of transgenes or small hairpin RNAs (Stehle *et al.*, 2007; Jenke *et al.*, 2008; Argyros *et al.*, 2011b; Wang *et al.*, 2017).

MARs are DNA sequences that bind chromatin to the nuclear matrix, which are ubiquitous and highly conserved in eukaryotic chromosomes (Wasag and Lenartowski, 2016). Stehle *et al.* (2003) found that MARs linked to an expression cassette is necessary to maintain episomes in mammalian cells. Whenever transcription of a transgene adjacent to the MARs sequence was abrogated, the constructs were either lost or integrated into the host genome (Stehle *et al.*, 2003). The mitotic stability of MAR episomal vectors is achieved via the direct interaction of MAR with scaffold attachment factor A (SAF-A) protein (Fackelmayer and Richter, 1994; Jenke *et al.*, 2002), which is a component of the nuclear matrix (Ferraro *et al.*, 1996; Kipp *et al.*, 2000).

The disadvantages of the original episomal vector were its variable expression, transgene silencing, and expression instability in cell lines. Therefore, studies have tried to improve transgene expression and stability. Jenke *et al.* (2004b) replaced the original 2-kb MAR with a tetramer of a 155-base-pair module comprising the core unwinding element of human β-IFN upstream of MAR, which is focused on improving transgene expression. Haase *et al.* (2010) reduced the CpG content of the pEPI backbone and omitted a second transcription unit, resulting in an episomal plasmid named pEPito. The pEPito vector is greatly reduced in size, contains only one transcription unit, and has 60% less CpG motives compared with that of pEPI. Moreover, MAR elements can promote episomal maintenance and replication, prevent epigenetic silencing (Jenke *et al.*, 2004a; Tessadori *et al.*, 2010; Hagedorn *et al.*, 2011; Stavrou *et al.*, 2017), and enhance transgene expression level (Rincón-Arano *et al.*, 2007; Buceta *et al.*, 2011).

In the present study, we first systematically studied the characteristics and mechanism by which truncated β-IFN MAR maintains episomal vectors. A deletion analysis of human β-IFN MAR was performed to characterize the regions that mediate episomal vector replication in CHO cells, with an aim to establish a novel strategy to develop safer vectors, particularly those that can remain episomal and therefore will be less genotoxic in gene therapies.

RESULTS

Transfection efficiency and transient expression

CHO cells were transfected with the constructed plasmids. The transfection efficiency, and transient *eGFP* expression after transfection for 48 h was detected by using flow cytometry. Transfection efficiency was found to be the highest for the plasmid with shortened MAR-5 (87.15%), followed by those for shortened MAR-4 (83.50%), MAR-3 (80.04%), MAR-1 (31.05%), and MAR-2 (20.37%) and full-length MAR (15.70%) (in this order) in CHO cells (Figure 1A). The statistical analysis revealed that shortened MAR-3, 4, and 5 resulted in significantly higher transfection efficiency than those of shortened MAR-1 and MAR-2 and full-length MAR (*P* < 0.05). However, transfection efficiency was not significantly different among MAR-1, MAR-2, and full-length MAR.

To explore the relationship between transfection efficiency and plasmid size, we tested a random sequence with an equal size as truncated S/MAR (a 540-base-pair DNA fragment). The transfection efficiency of the plasmid with the 540-base-pair DNA sequence was 62.62%, showing that the transfection efficiency was not related to the size of the plasmid (Figure 1B).

We also observed that shortened MAR-3, MAR-4, and MAR-5 resulted in significantly higher transient *eGFP* expression than that of shortened MAR-1, MAR-2, and full-length MAR (*P* < 0.05). This enhancement was greatest for shortened MAR-5, which increased transgenic *eGFP* expression 3.8-fold compared with that of full-length MAR, followed by 3.2-fold compared with that of MAR-4. MAR-1, and MAR-2, which resulted in a slight increase in transgene expression after transient transfection.

Establishment efficiency and *eGFP* expression in stably transfected cells

Diverse MARs were transfected into CHO cells, and the percentage G418-resistant clones was determined by colony-forming assay. The establishment efficiency of cells transfected with full-length MAR and shortened MAR-1 and MAR-2 was 8.5 ± 3.4%, 10.3 ± 3.2%, and 6.2 ± 2.5%, respectively. However, for MAR-3, MAR-4, and MAR-5, establishment efficiency was significantly increased to 16.2 ± 5.3%, 20.3 ± 2.6%, and 26.7 ±3.5%, respectively. The colony-forming efficiency with shortened MAR-5, MAR-4, MAR-3, MAR-2, and MAR-1 was 3.1-, 2.4-, 1.9-, 0.7-, and 1.2-fold, respectively, of the full-length MAR (Figure 2A). There were significant differences between MAR-3, MAR-4, MAR-5 and MAR-1, MAR-2, full-length MAR (Figure 2A, *P* < 0.05).

After exposure to G418 (800 μg/ml) for 2–3 wk, *eGFP* protein levels in stably transfected CHO cells were measured using flow cytometry. The results demonstrated that cells transfected with shortened MAR-5 exhibited the highest expression levels, followed by those transfected with shortened MAR-4, MAR-3, MAR-1, MAR- 2, and then by full-length MAR (in that order) (Figure 2B). The *eGFP* expression levels in the cells transfected with full-length MAR, shortened MAR-1, MAR-2, MAR-3, MAR-4, and MAR-5 were 0.32 ± 0.05, 0.51 ± 0.12, 0.45 ± 0.15, 1.41 ± 0.32, 1.62 ± 0.18, and 1.86 ± 0.24 × 10^6, respectively (Figure 2B). Next, quantitative PCR (qPCR) analysis was performed to determine the *eGFP* copy numbers and *eGFP* expression per plasmid copy was calculated (the amount of *eGFP* expression/DNA copy). Consistent with the result of *eGFP*
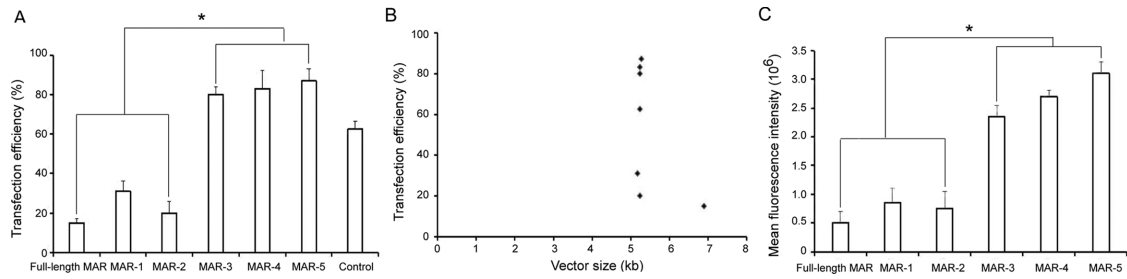
FIGURE 2: Establishment efficiency and *eGFP* expression in stably transfected CHO cells. (A) Establishment efficiency was determined by colony-forming assay. (B) The relation of transfection efficiency (%) and vector size (kb). There were significant differences among MAR-3,4,5, MAR-1,2, and full-length MAR. (C) Two weeks after transfection, CHO cells transfected with the shortened MAR 1–5 and full-length MAR were collected and the MFI of *eGFP* were examined by FACS Calibur. Mean values differed significantly between these plasmids (*$P < 0.05$).

expression levels, the cells transfected with shortened MAR-5, MAR-4, and MAR-3 exhibited significantly higher *eGFP* expression per plasmid copy than those transfected with shortened MAR-1 and MAR-2 and full-length MAR ($P < 0.05$, Table 1).

Furthermore, the analysis of the total population of *eGFP* protein-expressing cells suggested that cells transfected with shortened MAR-5, MAR-4, and MAR-3 resulted in significantly higher percentages of high-expression cells (median fluorescence intensity [MFI] > 10^3 or 10^4, designated %M3 and %M4) than those transfected with shortened MAR-1 and MAR-2, accompanying lower percentages of low-expression cells (MFI < 10^2 or 10 designated as %M2 and %M1, Supplemental Table S1).

Long-term expression
Aiming to evaluate the long-term expression of the trangene in this study, we chose the plasmids transfected with shortened MAR-3, MAR-4, and MAR-5 that showed relatively high sustained eGFP expression in the following experiments. pEPI (i.e., full-length MAR) was used as control.

CHO monoclonal cells stably transfected with MAR-3, MAR-4, MAR-5, and full-length MAR were cultured either in the presence or in the absence of G418 selection pressure, and MFI was examined to assess the stability of the *eGFP* expression 10, 18, 25, 33, 42, and 50 generations after transfection. The results showed that *eGFP* levels in all stably transfected CHO cells decreased gradually over time. At 50 generations after transfection, the cells transfected with MAR-5 showed the highest expression of *eGFP*, followed by cells transfected with MAR-4, MAR-3, and full-length MAR (Figure 3A). These findings indicated that there was no significance of *eGFP* expression levels between clones cultured in the presence or absence

of G418 (Figure 3A). Then, we evaluated the ability of MAR to sustain gene expression during long-term culture by measuring the percentage of cells still expressing *eGFP* in a monoclonal cell population after stability testing. MAR-5 and MAR-4 resulted in the maintenance of higher percentages of *eGFP*-expressing cells, with 81.5 ± 6.1%, 72.7 ± 5.8% (with exposure to G418) and 73.4 ± 8.2%, 60.9 ± 3.5% (without exposure to G418) *eGFP*-expressing cells in the stability tests, respectively. Full-length MAR resulted in the lowest percentage of *eGFP*-expressing cells, with only 28.5 ± 3.6% and 20.3 ± 3.4% of the cells expressing *eGFP* after exposure to no exposure to G418, respectively (Figure 3B).

In addition, the retention rates of MAR-5 were 50.34% and 40.15% (with and without G418, respectively); besides, the retention rates of MAR-4 were 37.62% and 32.64% (Figure 3C). However, the retention rates of MAR-3 were only 6.96% and 7.72%; similarly, the rates of full-length MAR were also only 5.86% and 5.21%. These results suggested that MAR-5 was most effective in maintaining transgene expression.

Consistent with the MFI results, fluorescence microscopy showed *eGFP* gene expression in the CHO cells transfected with full-length MAR, MAR-3, MAR-4, and MAR-5 at 50 generations after transfection when cells were cultured with or without selection pressure (Figure 3D).

Status of vectors and copy numbers
We evaluated the status of vector replication by plasmid rescue assay, Southern blot, and fluorescence in situ hybridization (FISH) analysis and analyzed copy numbers using FISH and qPCR at 50 generations posttransfection.

For the plasmid rescue assay, Hirt DNA was extracted from the CHO cells transfected with MAR-3, MAR-4, and MAR-5, and these rescued plasmids were identified by enzyme digestion. The results showed that the restriction pattern was the same as that of the original vector; that is, the sizes of MAR-3, MAR-4, and MAR-5 were 540, 540, and 581 base pairs, respectively (one example is shown in Figure 4A).

To further investigate the status of vector in transfected CHO cells, Southern blot analysis was carried out. Thirty colonies of CHO clones transfected with MAR-3, MAR-4, and MAR-5 (each 10 colonies) were randomly selected and subjected to Southern blot analysis. Hirt DNA was extracted and digested with *Bam*HI, and the results showed that the 30 colonies containing MAR-3, MAR-4, and MAR-5 were identical to those from the original plasmid DNA, that is, the sizes of MAR-3, MAR-4, and MAR-5 were 5271, 5271, and 5312 base pairs, respectively (one example is shown in Figure 4B, lanes 1–6). In contrast, in CHO cells transfected with the plasmid

MAR	Copy no.	*eGFP* expression per copy (10^5)
Full-length MAR	12.4 ± 3.25	0.26 ± 0.03
MAR-1	21.6 ± 6.17	0.23 ± 0.04
MAR-2	24.6 ± 8.33	0.19 ± 0.14
MAR-3	26.8 ± 9.02	0.56 ± 0.11*
MAR-4	23.2 ± 6.49	0.70 ± 0.09*
MAR-5	21.1 ± 7.26	0.88 ± 0.23*

Compared with full-length MAR, MAR-1, and MAR-2; *$P < 0.05$.

TABLE 1: *eGFP* expression per plasmid copy in stably transfected CHO cells.
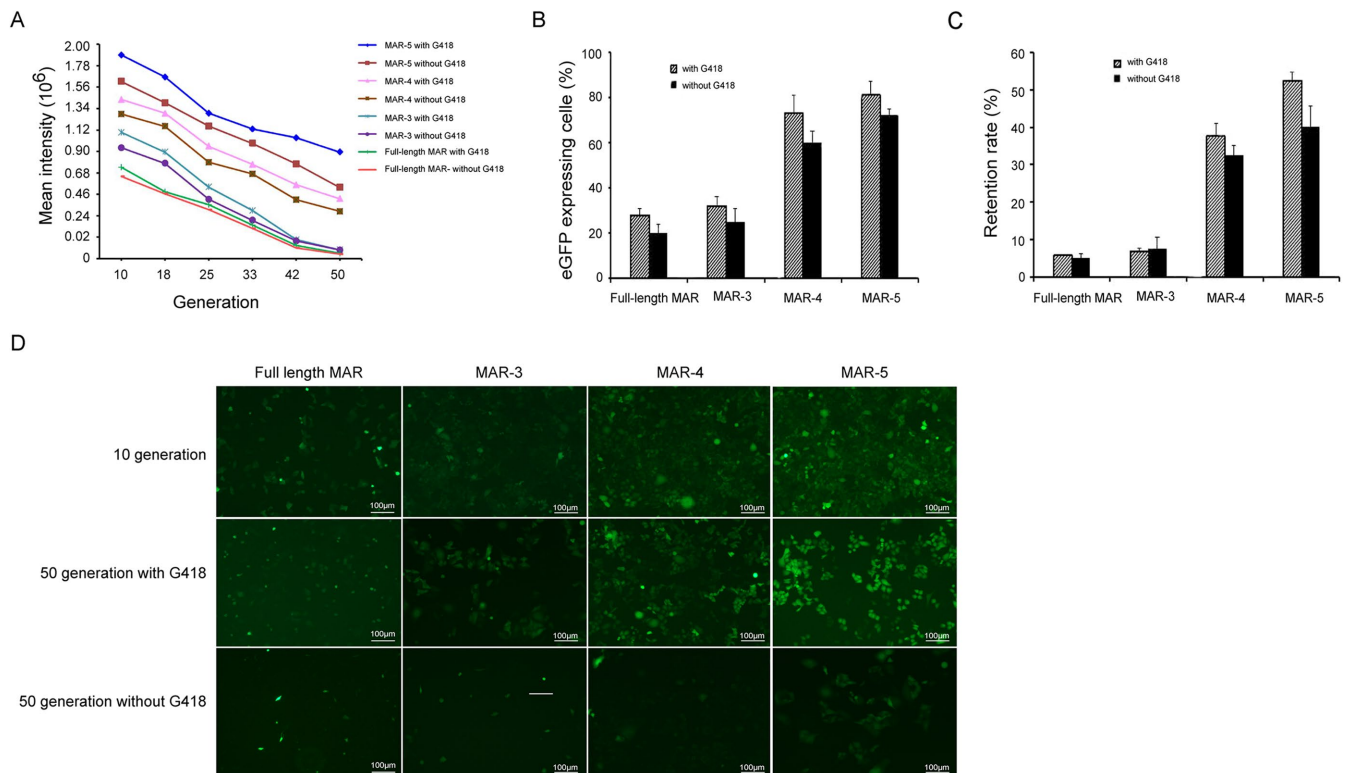
FIGURE 3: Long-term *eGFP* expression stability in transfected CHO cells. (A) The stably transfected CHO cells were cultured either with or without G418 selection pressure and MFI was detected to evaluate the stability of the expressed *eGFP* after 10, 18, 25, 33, 42, and 50 generations posttransfection, respectively. (B) The percentage of *eGFP* expressing cells in CHO cells transfected with full-length MAR, MAR-3, MAR-4, and MAR-5 vectors. (C) Retention of *eGFP* expression levels in cells transfected with full-length MAR, MAR-3, MAR-4, and MAR-5 vectors. (D) Fluorescence microscopy of *eGFP* gene in CHO cells transfected with full-length MAR, MAR-3, MAR-4, and MAR-5 grown with or without G418 selection pressure.

pEGFP-C1, its DNA was randomly integrated into the genome (Figure 4B, lanes 8 and 9), and bands with different sizes were present in lanes 8 and 9, whereas only one specific band was observed in lanes 1–6, indicating that the vectors mediated by the shortened MAR-3, MAR-4, and MAR-5 were present in an episomal form in CHO stable cells.

FISH analysis was performed on spread chromosomes of the CHO cells transfected with MAR-3, MAR-4, and MAR-5 either in the presence or in the absence of G418 selection pressure, with pECFP-C1 and full-length MAR vector as controls. The results demonstrated that the observed mitotic stabilities of the vectors with full-length MAR, MAR-3, MAR-4, and MAR-5 mainly resulted from the episomal state of the vector on metaphase chromosomes (Figure 4C). Combining plasmid rescue assay and Southern blot, the results showed that the plasmids containing MAR-3, MAR-4, and MAR-5 existed episomally in the CHO cells.

Next, each vector was analyzed by FISH with 50 metaphase plates, and the average copy numbers determined were 4.01 ± 0.23, 4.14 ± 0.58, 7.01 ± 1.53, and 8.26 ± 0.52 in the CHO cells transfected with full-length MAR, MAR-3, MAR-4, and MAR-5, respectively. At 50 generations posttransfection in the presence of G418 selection pressure, the average copy numbers were 3.22 ± 0.34, 4.13 ± 0.56, 6.17 ± 1.22, and 8.03 ± 1.42 in the absence of G418 selection pressure, respectively (range ~3–10 copies per cell; Figure 4D). There was no significant difference in the copy numbers between cells cultured in the presence or absence of selection pressure, suggesting that the vector can replicate episomally without significant copy number loss

during cell division and that this capacity is not compromised by the absence of selection pressure on mammalian cells.

qPCR was performed to determine the *eGFP* copy numbers of the vectors containing different MAR fragments in CHO cells at 50 generations posttransfection. The mean copy numbers in the cells transfected with full-length MAR, MAR-3, MAR-4, and MAR-5 were 2.26 ± 0.31, 5.31 ± 0.38, 10.33 ± 1.21, and 16.55 ± 0.79, respectively, in the presence of G418, and 2.30 ± 0.40, 4.59 ± 0.28, 11.57 ± 0.62, and 14.67 ± 1.34 in the absence of G418, respectively. The copy numbers in cells transfected with MAR-4 and MAR-5 were higher than those of MAR-3 and full-length MAR ($P < 0.05$) (Figure 4E).

Analysis of MAR binding to SAF-A in vivo

MAR-binding SAF-A was verified using chromatin immunoprecipitation (ChIP) experiment in vivo, and a PCR was conducted to test the elution DNA. As shown in Figure 5A, DNA fragments 177, 188, and 165 base pairs in length were obtained by PCR using the DNA template enriched with SAF-A antibody (lanes 2, 5, and 8). DNA fragments were amplified in the positive control group (lanes 1, 4, and 7), but amplified DNA fragment disappeared in the negative control (lanes 3, 6, and 9). These results showed that SAF-A antibody bound to the subfragments MAR-3, MAR-4, and MAR-5 of β-IFN S/MAR.

Then, the enriched DNA was used as a template for qPCR. It is found that the percentage of input value of subfragment MAR-5 was the largest, followed by the values for subfragment MAR-3 and subfragment MAR-4 (Figure 5B). Enrichment was 5.89-, 4.59-, and 4.38-fold compared with control (immunoglobulin G [IgG]).
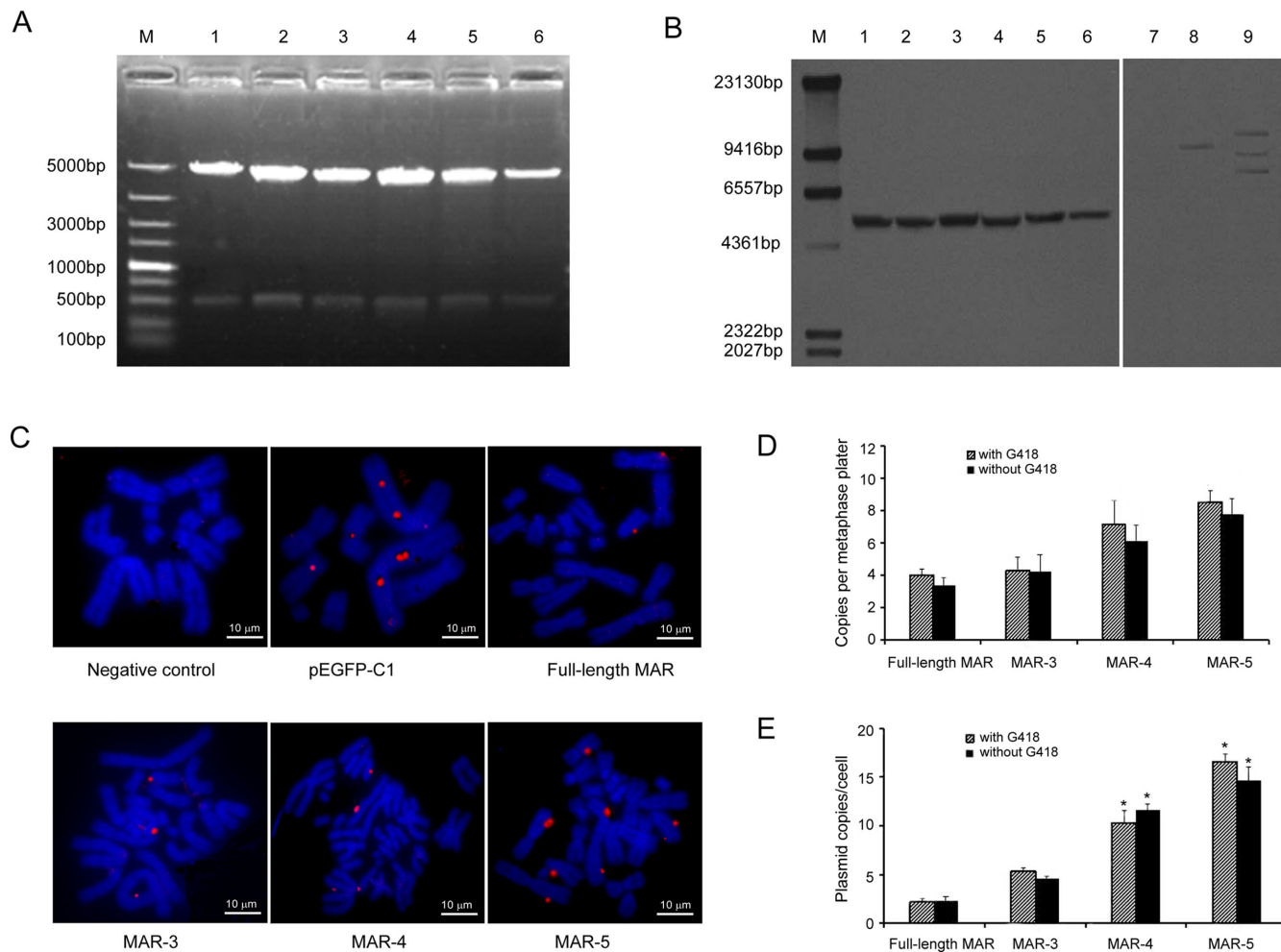
FIGURE 4: Analysis of state of existence of the transgene in CHO cells and copy number. (A) Rescue experiments in *E. coli* with Hirt extract from CHO cells transfected with MAR-3, MAR-4, and MAR-5. M: DNA marker; lanes 1, 3, 5: original plasmids with MAR-3, MAR-4, and MAR-5 double-digested with *Kpn*I and *Bam*HI, respectively; lanes 2, 4, 6: rescued plasmids with MAR-3, MAR-4, and MAR-5 double-digested with *Kpn*I and *Bam*HI 50 generations after transfection, respectively. (B) Southern analysis of DNA isolated from CHO cells transfected with MAR-3, MAR-4, and MAR-5. The hybridization pattern of one representative clone is shown for each construct. M: DNA marker; lane 1, 3, 5: original plasmids with MAR-3, MAR-4, and MAR-5 digested with *Bam*HI, respectively; lane 2, 4, 6: rescued plasmids with MAR-3, MAR-4, and MAR-5 digested with *Bam*HI 50 generations after transfection, respectively. Lane 7: *Eco*RI digestion of untransfected CHO cells; lanes 8, 9: CHO cells transfected with pEGFP-C (chromosomal DNA was isolated, digested with *Eco*RI, separated on a 0.8% agarose gel, blotted, and hybridized with pEGFP-C1 probe). (C) FISH analysis of *eGFP* served as a probe in CHO cells transfected with vector no *eGFP* gene (negative control), no MAR (pEGFP-C1), full-length MAR, MAR-3, MAR-4, and MAR-5, respectively. The vector without *eGFP* gene can only see blue metaphase chromosomes, integration vector pECFP-C1 can see vector insert into chromosomes, vector with full-length MAR, MAR-3, MAR-4 and MAR-5 was episomal state on metaphase chromosomes (blue: metaphase chromosomes; red: vectors). (D) The gene copies of each metaphase plate as determined by FISH analysis. Fifty metaphase spreads were analyzed by FISH for each clone, an average vector copy number was estimated, and SEM is indicated. (E) The copy number was assessed by qPCR analysis. A serial dilution with a plasmid containing the *eGFP* gene was used to determine the absolute copy number. Three independent experiments were performed in this study. SEM is indicated (Student's *t* test, *$P < 0.05$).

Analysis of MAR binding to SAF-A in vitro

To investigate whether SAF-A could bind to MAR DNA in vitro, an electrophoretic mobility shift assay (EMSA) was performed (MAR-4, as an example, is shown in Figure 5C). When >100 × unlabeled probe was used as competitor DNA, the DNA–protein complex band disappeared (Figure 5C, lane 2). When mutated unlabeled probes were used, the band showed little change compared with the band when probes for the MAR subfragments were used (Figure 5C, lane 3). A DNA–protein complex band appeared when the probes targeting the MAR subfragments interacted with cell nucleoprotein (Figure 5C, lane 4). A super-shift experiment showed that SAF-A antibody could bind to the DNA–protein complex, which caused a super-shift and retarded the migration of the DNA–protein complex. The DNA–protein complex band became significantly weaker (Figure 5C, lane 5). These results showed that SAF-A antibody can bind to subfragments of MAR-3, MAR-4, and MAR-5.
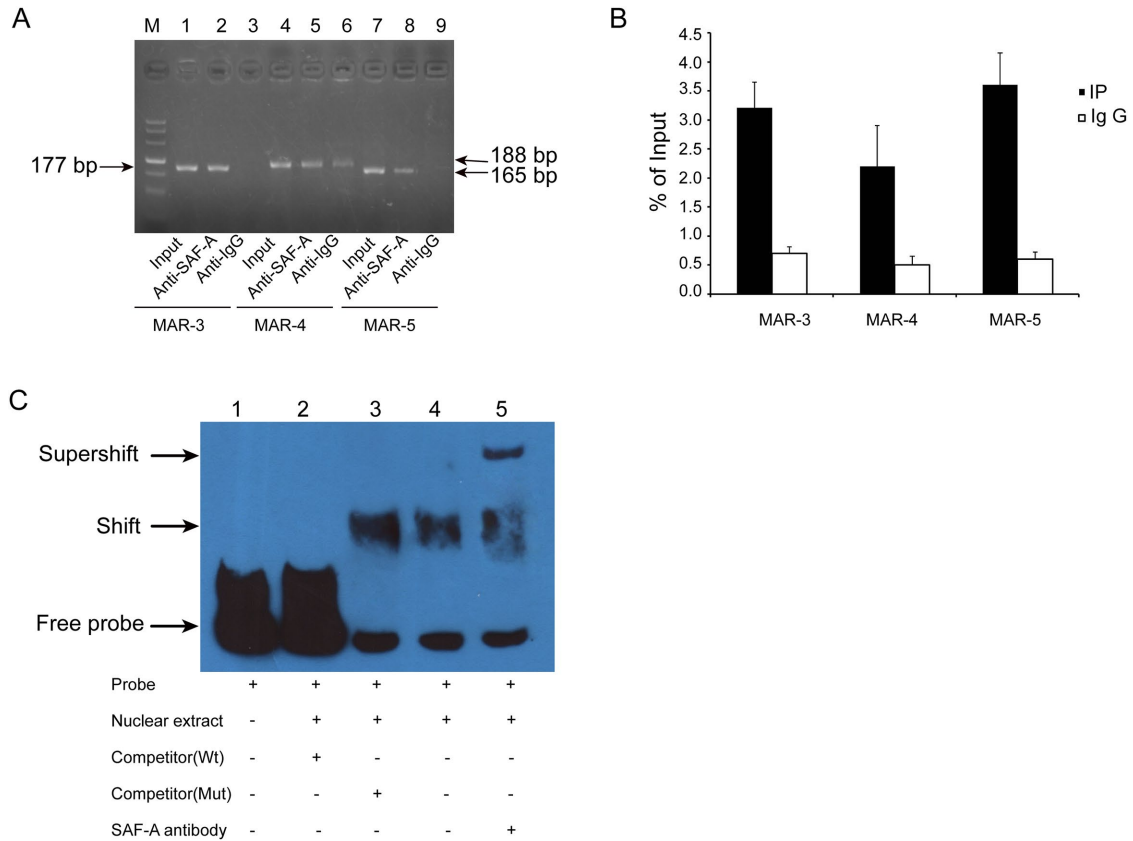
FIGURE 5: ChIP and EMSA experiment. (A) PCR analysis of chromatin immunoprecipitation reactions, M: 500-base-pair marker; (B) ChIP-qPCR enrichment is shown as the percentage of input DNA; ChIP was performed on CHO/ subfragment MAR-3,4, 5 using a specific antibody against SAF-A. Immunoprecipitated DNA was subjected to qPCR. (C) Representative of SAF-A binds specificity to subfragment MAR-4.

Analysis of transcription factor binding motifs

Transcription factor binding motifs can affect the activity of chromosomal elements such as introns and MAR (Running Deer and Allison, 2004; Arope et al., 2013). Therefore, we assessed the distribution of seven transcription factor binding motifs in the five β-IFN MAR subfragments. These motifs were those that allow binding of special AT-rich sequence binding protein 1 (SATB1); nuclear matrix protein 4 (NMP4); CCAAT/enhancer-binding protein (C/EBP); fast, octamer motif 1 (OCT1); and enhancer box (E-Box). As shown in Table 2, all of the MAR fragments contained transcription factors OCT1. In addition, we detected 0, 1, and 2 NMP4 motifs within subfragments MAR-3, MAR-4, and MAR-5, respectively, and 2 NFAT and 3 E-Box motifs in subfragment MAR-5. Moreover, no E-Box or NMP4 motifs were identified in subfragment MAR-1, MAR-2, or MAR-3. These findings suggested that the transcription factors NMP4 and E-Box may play roles in maintaining transgene expression.

DISCUSSION

In this study, we explored an alternative strategy for delivering transgenes using a nonviral, nonintegrating episomal construct. Our results showed that, of the five β-IFN MAR fragments tested (subfragments MAR1–5), MAR-5, followed by MAR-4 and then MAR-3, was the most effective for promoting episomal maintenance and replication of transgenes. The results obtained from the FISH analysis, Southern blot, and plasmid rescue experiments strongly suggested that plasmids containing MAR-3, MAR-4, and MAR-5 all replicated episomally and efficiently maintained transgene expression in CHO cells. Specifically, the episomal vector mediated by MAR-5 showed the highest transfection efficiency, highest expression level, and most stability of all of the other vectors.

We found that MAR-3, MAR-4, and MAR-5 increased the transfection efficiency of vector compared with that of the vectors with full-length MAR, but MAR-1 and MAR-2 did not. The size of the

MAR	SATB1 +	SATB1 −	NMP4 +	NMP4 −	C/EBP +	C/EBP −	Fast +	Fast −	OCT1 +	OCT1 −	GATA +	GATA −	E-Box +	E-Box −	CRE +	CRE −
MAR-1	0	0	0	0	0	2	0	0	3	0	2	0	0	0	0	2
MAR-2	1	0	0	0	1	0	2	0	4	1	1	1	0	0	0	2
MAR-3	0	1	0	0	0	2	1	0	1	3	1	1	0	0	0	3
MAR-4	0	0	1	0	0	0	0	0	2	4	0	0	1	0	0	1
MAR-5	0	0	0	2	1	0	1	0	2	0	1	2	2	1

TABLE 2: Locations of various transcription factor binding motifs within the five subfragments of β-IFN MAR.

vector affects transfection efficiency to some extent, but the structure and configuration of MARs also make significant effects on transfection efficiency. In this work, five shortened MARs of the same size showed different transfection efficiencies, and the results of transfecting a random sequence of the same size as the truncated S/MAR were consistent with the previous study (Zhang et al., 2017).

We also verified that shortened MAR-3, MAR-4, and MAR-5 increased establishment efficiency; transient and stable expression of eGFP in transfected CHO cells relative to those in cells transfected with MAR-1, MAR-2; and full-length MAR. Establishment efficiency is defined by the percentage of cells in which a particular S/MAR vector is stably retained as an episome after an initial selection period. The expected establishment efficiency for MAR vectors was ~1–5% (Jenke et al., 2004b; Wasag and Lenartowski, 2016) and, for MAR minicircles, it was ~10% (Broll et al., 2010; Hagedorn et al., 2012). Hagedorn et al. (2013) found that insulator sequence (cHS4) increased establishment efficiency of a MAR vector to 25.3 ± 13.98%. Our study showed that the highest establishment efficiency of MAR-5 could achieve to 26.7 ± 13.5%, which is a little higher than that reported for pEPIH-S4opp. The establishment efficiencies of MAR-3 and MAR-4 were 16.2 ± 10.3% and 20.3 ± 8.6%, higher than those of the MAR vectors alone, suggesting that shortened MAR-3, MAR-4, and MAR-5 can increase establishment efficiency. Moreover, all of the shortened MARs enhanced transgene expression levels compared with those of the full-length MAR in stably transfected CHO cells. Among them, shortened MAR-5, MAR-4, and MAR-3 exhibited significantly higher eGFP expression than shortened MAR-1, MAR-2, and full-length MAR. Jenke et al. (2004b) found that a tetramer of a 155-base-pair module comprising the core unwinding element of human IFN upstream of S/MAR was sufficient for replication and episomal maintenance, but they did not determine establishment efficiency or recombinant protein expression levels. pEPito, developed by Haase et al. (2010) showed higher transgene expression levels and increased colony-forming efficiency in vitro. Our shortened MAR-5, MAR-4, and MAR-3 also showed these advantages.

An ideal episomal vector for functional genetics or gene therapy applications should allow long-term expression of the delivered transgene at levels close to physiological ones (Ehrhardt et al., 2008; Xu et al., 2016). The delivery of vectors containing large DNA fragments by nonviral means has proven inefficient. Therefore, in this work, we cloned five shortened β-IFN MAR fragments into pEGFP-C1, and differences between the eGFP levels expressed in cells with these five plasmids suggested that vectors containing MAR-5 were more resistant to transcriptional silencing, remained episomal, and showed the greatest stability at 50 generations, irrespective of whether the cells were exposed or were not exposed to selection pressure.

In the present study, use of the shortened MAR sequence from the β-IFN MAR, instead of its natural longer sequence, reduced the vector size and greatly facilitated vector manipulation. Most importantly, small-sized vectors are likely to be packaged more efficiently, allowing high titers of viral particles to be generated and ultimately resulting in high gene transfer efficiency. The shortened S/MAR sequence also leaves more space for exogenous genes, which is a vital feature for efficient gene transfer vectors (Kurre et al., 2003).

Plasmid rescue assays were used to determine the status of plasmids and, in this study, plasmids containing MARs were found to be stably maintained in dividing cells. In addition, selection for a resistance marker present within the episome has been used to promote persistence (Sgourou et al., 2009; Argyros et al., 2011a; Dixon et al., 2012). In this study, plasmid rescue experiments clearly indicated that all of the rescued plasmids (MAR-3, MAR-4, and MAR-5) yielded the same band as the original plasmid, confirming the episomal state of the constructed vectors. The FISH analysis also revealed that the plasmids with MAR-3, MAR-4, and MAR-5 existed episomally on metaphase chromosomes in the presence or absence of 4 mo of selection pressure after transfection. The episomal state of plasmids with MAR-3, MAR-4, or MAR-5 in the transfected CHO cells was further investigated by Southern blot. The appearance of a single band in this Southern analysis shows that there is no detection of integration of plasmids MAR-3, MAR-4, and MAR-5 into the CHO cell genome. Theoretically, the free episomes would be lost during screening of stable transfectants with G418.

SAF-A is a prominent protein in the nuclear matrix that specifically binds to MAR through the highly conserved "SAF-box" (Romig et al., 1992; Krecic and Swanson, 1999). The episomal vector that originated from MAR is recognized by the nuclear matrix protein SAF-A, as shown for pEPI. In the ChIP and EMSA analysis, we showed that MAR-3, MAR-4, and MAR-5 specifically bind SAF-A in vivo and in vitro. Jenke et al. (2004b) demonstrated that a tetramer of a 155-base-pair minimal MAR DNA module was sufficient for replication and mitotic stability of a mammalian episome that associates with the nuclear matrix by means of specific interactions between MAR and SAF-A. Our results indicate that the shortened MAR derivatives mediate maintenance of episomal vectors through specific interactions between MAR and SAF-A.

Transcription factor binding motifs such as SATB1, NMP4, CEBP, Fast, OCT1, and E-Box binding sites are thought to play important roles in promoting transgene expression (Li et al., 2013; Wang et al., 2016). Consistent with these findings, in this study, bioinformatic analyses revealed MAR-like motifs in MAR-3, MAR-4, and MAR-5 and indicate that NMP4 and E-Box might play roles in maintaining transgene expression. Specifically, these results suggested that the mechanism by which MAR enhances transgene expression might be at least partially dependent on certain transcription factor binding sites in the MAR. In the previous studies, NMP4 and E-Box transcription factors have been demonstrated that they are positive transcription factors sites for enhancing transgene expression. Transcription factors can be divided into positive, neutral, and negative effectors according to their effects on gene expression (Brown et al., 2014). The strong promoter contains the highest ratio of positive transcription factors sites, such as E-Box (Brown et al., 2014). In addition, Girod et al. (2007) found that the most active MAR sequences are enriched in the binding motifs for the HoxF, CEBP, and NMP4 transcription factors as well as for FAST1, whereas sequences showing moderate or no activity are relatively devoid of such elements. In the present study, the shortened MAR-5 are rich in NMP4 and E-Box, which may be attribute to the enhancing function.

In conclusion, we explored five shortened β-IFN MAR fragments and found that vectors containing MAR-5, followed by those containing MAR-4 and MAR-3, can increase transient, stable transgene expression levels; colony-forming efficiencies; higher expression cell rates; transgene copy numbers; and stability compared with those of the original pEPI-1 plasmid in the episomal state in CHO cells. The reasons may be due to the type and number of transcription factors in MAR-3, MAR-4, and MAR-5. The development of such expression systems remains a major strategic task for nonviral gene therapy research and is a prerequisite for any future long-term clinical applications.
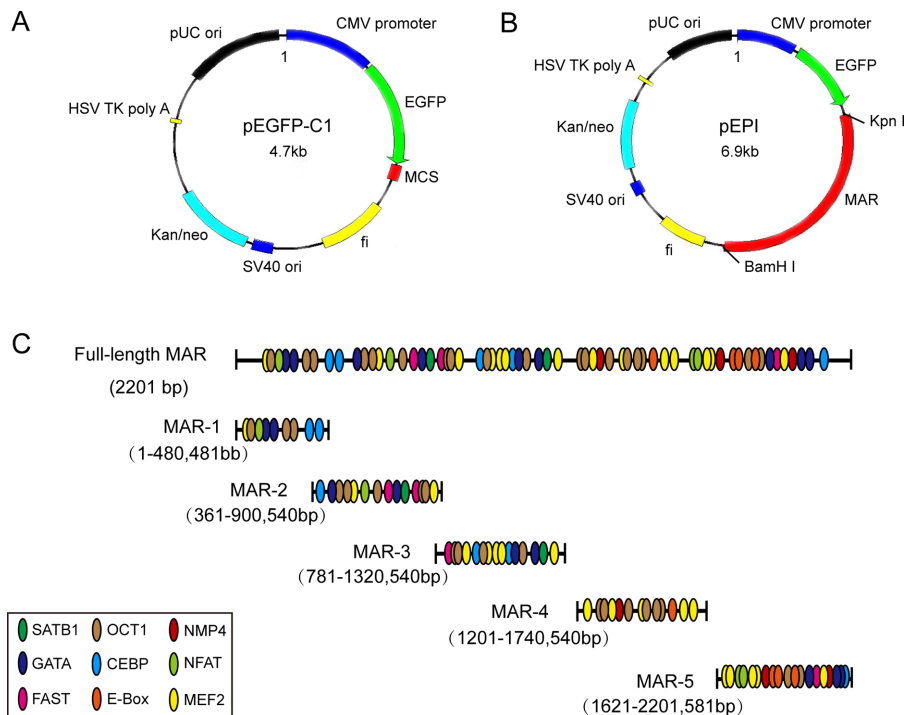
FIGURE 6: Schematic representation of vectors, β-IFN MAR subfragment and illustration of its transcription factor binding motifs. (A) pEGFP-C1 (Clontech, Mountain View, CA) was used as the original vector. (B) β-IFN MAR (GenBank No: M83137.1) inserts into the downstream of *eGFP* reporter gene of the pEGFP-C1 plasmid to generate pEPI. (C) MAR-1, MAR-2, MAR-3, MAR-4, and MAR-5 represent the DNA fragments from the positions 1–480, 361–900, 781–1320, 1201–1740, and 1621–2201 of human β-IFN MAR; schematic illustration of various transcription factor binding motifs within the full-length human β-IFN MAR and five shortened β-IFN MAR.

MATERIALS AND METHODS

Design and constructs

Five β-IFN MAR subfragments (nucleotides 1–480, 361–900, 781–1320, 1201–1740, and 1621–2201) were amplified from the flanking regions of the β-IFN MAR sequence using the primers listed in Supplemental Table S2. The fragments were each designed to overlap with adjacent subfragments by 120 nucleotides. Bidirectional cloning was achieved by digesting the resulting PCR fragments with *Kpn*I and *Bam*HI restriction enzymes (Takara Biotechnology, Dalian, China) and inserting these sequences downstream of the *eGFP* expression cassette region in pEGFP-C1 (Figure 6A) to generate plasmids containing different subfragments of β-IFN MAR (MAR1–5). pEPI (i.e., full-length MAR) was used as a control (Figure 6B). Various transcription factors that included full-length MAR and MAR 1–5 are shown in Figure 6C.

Cell culture and transfection

CHO cells were maintained in DMEM (Invitrogen) containing 10% fetal bovine serum and 1% penicillin-streptomycin. A total of 8×10^4 CHO cells were cultured in 24-well plates for each well. At 80% confluence, plasmids containing full-length MAR or shortened MAR (nucleotides 1–480, 361–900, 781–1320, 1201–1740, or 1621–2201 of the shortened MAR1–5) were transfected into cells by Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions.

Flow cytometry

After transfection for 48 h, 1×10^6 CHO cells were harvested to determine transfection efficiency and *eGFP* transient expression by flow cytometry using a FACSCalibur instrument (BD Accuri C6; BD Biosciences). Another 1×10^6 CHO cells were screened under Geneticin pressure (800 μg/ml G418, Invitrogen). Stably transfected cells pools were obtained after 2–3 wk. The stability expression of *eGFP* was again analyzed by flow cytometry. More than 100,000 events were recorded for each experiment. *eGFP* signals were measured in duplicate at 488 and 567 nm to exclude fluorescence of different cells. Mean fluorescence values greater than 10^3 or 10^4 were designated %M3 and %M4, smaller than 10^2 or 10 designated as %M2, %M1, respectively.

Colony formation assay

To determine the establishment efficiency of episomal vectors containing different MARs, a colony-forming assay was performed as previously described (Hagedorn *et al.*, 2012). Briefly, cells were serially diluted (1:1, 1:10, and 1:100) at 48 h posttransfection in the presence of 800 μg/ml G418. After 2 wk of selection, G418-resistant colonies formed and were fixed with 1% formaldehyde in phosphate-buffered saline, counterstained with methylene blue, and counted. Each sample was performed in triplicate.

Long-term stability testing

Once the stably transfected colonies formed, cells were diluted in 96-well plates to produce monoclonal cells using a limited dilution method, and stable clones were cultured with or without G418 (400 μg/ml), respectively. The stably transfected CHO cells were passaged in six-well plates. Next, the MFI for each plasmid was measured at 10, 18, 25, 33, 42, and 50 generations after transfection, and the percentage of *eGFP*-expressing cells was determined by a FACSCalibur cytometer. Retention of *eGFP* expression of each vector was calculated as the ratio of MFI at 50 generations to MFI at 10 generations.

Plasmid rescue experiments

To verify the status of plasmids in the transfected and stably selected CHO cell lines, plasmid rescue experiments were performed by chemical transformation of 10 μl of isolated Hirt DNA (~500 ng) into chemically competent *Escherichia coli* 5α cells (Sangon, Shanghai, China) as previously described (Zhang *et al.*, 2017). The transformed colonies were selected on agar plates containing 50 mg/ml kanamycin. DNA was isolated from individual resistant clones and then subjected to digestion with restriction enzymes *Kpn*I and *Bam*HI. The products were analyzed by electrophoresis on 1.0% agarose gels to verify the status of vectors in CHO cells.

Southern blot

For Southern blot analysis, extrachromosomal DNA was extracted from CHO cells transfected with the shortened MAR, followed by digestion of 3 μl DNA with *Bam*HI restriction enzyme and electrophoresis on a 1% agarose gel. Gels were transferred to a HyBond N+ positively charged nylon membrane (Amersham, Beijing, China; cat. no. RPN303B) and hybridized with *eGFP* probe labeled with

digoxin (Roche, Basel, Switzerland; cat. no. 11585614910), the concentration of the probe was 15 ng/µl, the volume of hybrids was 3.5 ml/100 cm^2. The *eGFP* probe sequence was 5′-GGCGTGCAGT-GCTTCAGCCGCTAC CCCGACCACATGAAGCA-3′, and hybridization was performed in Church buffer (0.25 M sodium phosphate buffer, pH 7.2, 1 mM EDTA, 1% bovine serum albumin, 7% SDS) at 65°C for 16 h.

FISH analysis
FISH analysis was carried out as described previously at 50 generations posttransfection (Zhang *et al.*, 2017). *eGFP* served as a probe (plasmid DNA was 2 ng, genomic DNA was 20 µg) and was labeled using a digoxigenin-nick translation kit (Roche, Mannheim, Germany). Before microscopy, the samples were counterstained with 1 µg/ml 4′,6′-diamidino-2-phenylindole and analyzed under a Leica DMRB fluorescence microscope with a Leica DC300 F camera. Approximately 50 fields were observed, and the mean copy number was calculated.

Real-time qPCR
qPCR was performed to measure the copy numbers of the *eGFP* gene at 2–3 wk after transfection and 50 generations posttransfection. Episomal DNA was extracted from CHO cells. The primers used were as follows: *eGFP* forward: 5′-GCT GGT TTA GTG AAC CGT CAG-3′; *eGFP* reverse: 5′-AGG TGG CAT CGC CCT CGC CC-3′. G6PDH forward: 5′-GAT GGG GTA CCC TTC ATC C-3′; G6PDH reverse: 5′-GCT CTG ACT CCT CAG GGT TG-3′. G6PDH was used for an internal control. The reaction was performed on an ABI 7500 SYBR fluorescence qPCR instrument (Applied Biosystems, Foster City, CA) and the results were analyzed by PikoReal Software 2.2 (ThermoFisher Scientific). The program consisted of 40 cycles of 95°C for 3 min, 94°C for 30 s, 50°C for 30 s, and 72°C for 30 s. A serial dilution with a plasmid containing the *eGFP* gene was used to determine the absolute copy number.

Quantitative ChIP (qChIP) assay
ChIP was carried out according to manufacturer's instructions (Bersin-Bio, Guangzhou, China). Cells were fixed with 1% formaldehyde to allow the DNA to cross-link with the proteins and then disrupted for six cycles of 30 s using an automatic ultrasonic homogenizer to disrupt the genomic DNA into 200- to 600-base-pair fragments. Specific antibodies (1 µg) against SAF-A (Abcam, Cambridge, UK; cat. no. EPR13588) were used in the ChIP assay, and 1 µg of mouse IgG was used as a negative control. RNase (1 µg) and 1 µl of protease K were added to the immune precipitation complex and held at 45°C for 2 h. DNA was purified for PCR and qPCR analyses.

The immunoprecipitated DNA was subjected to PCR. The input (positive control), DNA fragment precipitated with IgG antibody (negative control), and DNA fragment precipitated with SAF-A antibody were used as PCR templates, and the primers targeted the shortened MAR-3, MAR-4, and MAR-5 vectors. The PCR was performed using the standard procedure: 95°C for 3 min, 95°C for 10 s, 56°C for 10 s, and 72°C for 10 s, followed by 30 cycles at 72°C. Primers are listed in Supplemental Table S2.

Real-time qPCR was carried out using a SYBR Green system according to the manufacturer's instructions (Roche Applied Science). Primers were the same as those used for the PCR. The cycling parameters were as follows: 95°C for 10 min, 95°C for 30 s, 60°C for 20 s, and 72°C for 30 s. The qPCR was performed on an Bio-Rad CFX96 instrument, and the data were analyzed using Bio-Rad CFX Manager Software 1.6. The $2^{-\Delta\Delta Ct}$ method was used to calculate the percentage of DNA target fragment relative to the input control in the ChIP reaction.

EMSA
An oligonucleotide probe labeled with biotin was synthesized by Sangon Biotech (Shanghai, China). The sequences of MAR-3, MAR-4, MAR-5, and the competitor (wt) and competitor (mut) probes were 5′-AAGTCTTACAAATCAGTAA AAAATAAAACTAGACAAAA-3′, 5′-AAAAACTAATGAAATATAAATAAAGTTTGAACTTTAGTT-3′, 5′-AACCTGTTGATGGGACAAATTACGTTAATTGATTTTCAA-3′, 5′-AAAAACTAATGAAATATAAATAAAGTTTGAACTTTAGTT-3′, and 5′-GCCGGCCCGGCGGCTCCGCAGCGTGCCACAGGGTCC-GCG-3′, respectively. Cells were collected and nuclear proteins were extracted, and then the 0.5-µl probe was allowed to bind to the 5-µl nuclear proteins. More than 100× excess unlabeled probe was used as a competitive control to confirm specificity. SAF-A (1 µg) antibody was used in a super-shift analysis, followed by electrophoresis separation on a 6.5% nondenatured polyacrylamide gel. The gel was transferred for 254-nm ultraviolet cross-linking, and the DNA was labeled with biotin for detection using a chemiluminescence method.

Bioinformatic analysis
Bioinformatics analyses were performed according to previously described methods (Sun *et al.*, 2016). Allele-specific transcription factor binding sites were identified using MatInspector software (www.genomatix.de/products/ index.html) (Quandt *et al.*, 1995). Structural motifs were identified using GeneExpress.

Statistical analysis
All experiments were performed thrice, and samples were tested in triplicate. Experimental data ware analyzed using SPSS19.0 software (SPSS, Chicago, IL). *eGFP* expression from these constructs was analyzed by one-way analysis of variance to assess statistical significance. A postanalysis of variance using a multiple comparison procedure was used to assess pairwise differences in expression, which were confirmed by an analysis of variance. *P* values < 0.05 were considered statistically significant.

REFERENCES
Argyros O, Wong SP, Fedonidis C, Tolmachov O, Waddington SN, Howe SJ, Niceta M, Coutelle C, Harbottle RP (2011a). Development of S/MAR minicircles for enhanced and persistent transgene expression in the mouse liver. J Mol Med (Berl) 89, 515–529.

Argyros O, Wong SP, Harbottle RP (2011b). Non-viral episomal modification of cells using S/MAR elements. Expert Opin Biol Ther 11, 1177–1191.

Arope S, Harraghy N, Pjanic M, Mermod N (2013). Molecular characterization of a human matrix attachment region epigenetic regulator. PLoS One 8, e79262.

Baiker A, Maercker C, Piechaczek C, Schmidt SB, Bode J, Benham C, Lipps HJ (2000). Mitotic stability of an episomal vector containing a human scaffold/matrix-attached region is provided by association with nuclear matrix. Nat Cell Biol 2, 182–184.

Broll S, Oumard A, Hahn K, Schambach A, Bode J (2010). Minicircle performance depending on S/MAR-nuclear matrix interactions. J Mol Biol 395, 950–965.

Brown AJ, Sweeney B, Mainwaring DO, James C (2014). Synthetic promoters for CHO cell engineering. Biotechnol Bioeng 111, 1638–1647.

Browning DL, Trobridge GD (2016). Insulators to improve the safety of retroviral vectors for HIV gene therapy. Biomedicines 4, 1–14.

Buceta M, Galbete JL, Kostic C, Arsenijevic Y, Mermod N (2011). Use of human MAR elements to improve retroviral vector production. Gene Ther 18, 7–13.

Chen F, Qi X, Zhang R, Wu ZY, Yan CE, Li J, Liu QY, Qi J (2017). Episomal lentiviral vectors confer erythropoietin expression in dividing cells. Plasmid 90, 15–19.

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380.

Ehrhardt A, Haase R, Schepers A, Deutsch MJ, Lipps HJ, Baiker A (2008). Episomal vectors for gene therapy. Curr Gene Ther 8, 147–161.

Ellis J (2005). Silencing and variegation of gamma retrovirus and lentivirus vectors. Hum Gene Ther 16, 1241–1246.

Fackelmayer FO, Richter A (1994). Purification of two isoforms of hnRNP-U and characterization of their nucleic acid binding activity. Biochemistry 33, 10416–10422.

Ferraro A, Cervoni L, Eufemi M, Altieri F, Turano C (1996). Comparison of DNA-protein interactions in intact nuclei from avian liver and erythrocytes: a cross-linking study. J Cell Biochem 62, 495–505.

Girod PA, Nguyen DQ, Calabrese D, Puttini S, Grandjean M, Martinet D, Regamey A, Saugy D, Beckmann JS, Bucher P, et al. (2007). Mermod, Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. Nat Methods 4, 747–753.

Haase R, Argyros O, Wong SP, Harbottle RP, Lipps HJ (2010). pEPito: a significantly improved non-viral episomal expression vector for mammalian cells. BMC Biotechnology 10, 20.

Hagedorn C, Antoniou MN, Lipps HJ (2013). Genomic cis-acting sequences improve expression and establishment of a nonviral vector. Mol Ther Nucleic Acids 2, e118.

Hagedorn C, Baiker A, Postberg J, Ehrhardt A, Lipps HJ (2012). A colon forming assay for determining the establishment efficiency of S/MAR containing nonviral episomal expression vectors. Cold Spring Harb Protoc 2012, 706–708.

Hagedorn C, Schnödt-Fuchs M, Boehme P, Abdelrazik H, Lipps HJ, Büning H (2017). S/MAR element facilitates episomal long-term persistence of adeno-associated virus vector genomes in proliferating cells. Hum Gene Ther 28, 1169–1179.

Hagedorn C, Wong SP, Harbottle R, Lipps HJ (2011). Scaffold/matrix attached region-based nonviral episomal vectors. Hum Gene Ther 22, 915–923.

Jenke AC, Scinteie MF, Stehle IM, Lipps HJ (2004a). Expression of a transgene encoded on a non-viral episomal vector is not subject to epigenetic silencing by cytosine methylation. Mol Biol Rep 31, 85–90.

Jenke AC, Stehle IM, Herrmann S, Eisenberger T, Baiker A, Bode J, Fackelmayer FO, Lipps HJ (2004b). Nuclear scaffold/matrix attached region modules linked to a transcription unit are sufficient for replication and maintenance of a mammalian episome. Proc Natl Acad Sci USA 101, 11322–11327.

Jenke AC, Wilhelm AD, Orth V, Lipps HJ, Protzer U, Wirth S (2008). Long-term suppression of hepatitis B virus replication by short hairpin RNA expression using the scaffold/matrix attachment region-based replicating vector system pEPI-1. Antimicrob Agents Chemother 52, 2355–2359.

Jenke BH, Fetzer CP, Stehle IM, Jönsson F, Fackelmayer FO, Conradt H, Bode J, Lipps HJ (2002). An episomally replicating vector binds to the nuclear matrix protein SAF-A in vivo. EMBO Rep 3, 349–354.

Kipp M, Göhring F, Ostendorp T, van Drunen CM, van Driel R, Przybylski M, Fackelmayer FO (2000). SAF-Box, a conserved protein domain that specifically recognizes scaffold attachment region DNA. Mol Cell Biol 20, 7480–7489.

Kohn DB, Sadelain M, Glorioso JC (2003). Occurrence of leukaemia following gene therapy of X-linked SCID. Nat Rev Cancer 3, 477–488.

Krecic AM, Swanson MS (1999). hnRNP complexes: composition, structure, and function. Curr Opin Cell Biol 11, 363–371.

Kurre P, Morris J, Thomasson B, Kohn DB, Kiem HP (2003). Scaffold attachment region-containing retrovirus vectors improve long-term proviral expression after transplantation of GFP-modified CD34+ baboon repopulating cells. Blood 102, 3117–3119.

Li Y, Huang W, Niu L, Umbach DM, Covo S, Li L (2013). Characterization of constitutive CTCF/cohesin loci: a possible role in establishing topological domains in mammalian genomes. BMC Genomics 14, 553.

Lin Y, Li Z, Wang T, Wang X, Wang L, Dong W, Jing C, Yang X (2015). MAR characteristic motifs mediate episomal vector in CHO cells. Gene 559, 137–143.

Ma C, Fan Z, Gao Z, Wang S, Shan Z (2017). Delivery of human erythropoietin gene with an adeno-associated virus vector through parotid glands to treat renal anaemia in a swine model. Gene Ther 24, 692–698.

Piechaczek C, Fetzer C, Baiker A, Bode J, Lipps HJ (1999). A vector based on the SV40 origin of replication and chromosomal S/MARs replicates episomally in CHO cells. Nucleic Acids Res 27, 426–428.

Quandt K, Frech K, Karas H, Wingender E, Werner T (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res 23, 4878–4884.

Rincón-Arano H, Furlan-Magaril M, Recillas-Targa F (2007). Protection against telomeric position effects by the chicken cHS4 beta-globin insulator. Proc Natl Acad Sci USA 104, 14044–14049.

Romig H, Fackelmayer FO, Renz A, Ramsperger U, Richter A (1992). Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. EMBO J 11, 3431–3440.

Running Deer J, Allison DS (2004). High-level expression of proteins in mammalian cells using transcription regulatory sequences from the Chinese hamster EF-1alpha gene. Biotechnol Prog 20, 880–889.

Sánchez-Hernández S, Gutierrez-Guerrero A, Martín-Guerra R, Cortijo-Gutierrez M, Tristán-Manzano M, Rodriguez-Perales S, Sanchez L, Garcia-Perez JL, Chato-Astrain J, Fernandez-Valades R, et al. (2018). The IS2 element improves transcription efficiency of ntegration-deficient lentiviral vector episomes. Mol Ther Nucleic Acids 7,16–28.

Sgourou A, Routledge S, Spathas D, Athanassiadou A, Antoniou MN (2009). Physiological levels of HBB transgene expression from S/MAR element-based replicating episomal vectors. J Biotechnol 143, 85–94.

Stavrou EF, Lazaris VM, Giannakopoulos A, Papapetrou E, Spyridonidis A, Zoumbos NC, Gkountis A, Athanassiadou A (2017). The β-globin Replicator greatly enhances the potential of S/MAR based episomal vectors for gene transfer into human haematopoietic progenitor cells. Sci Rep 7, 40673.

Stehle IM, Postberg J, Rupprecht S, Cremer T, Jackson DA, Lipps HJ (2007). Establishment and mitotic stability of an extra-chromosomal mammalian replicon. BMC Cell Biol 8, 33.

Stehle IM, Scinteie MF, Baiker A, Jenke AC, Lipps HJ (2003). Exploiting a minimal system to study the epigenetic control of DNA replication: The interplay between transcription and replication. Chromosome Res 11, 413–421.

Sun QL, Zhao CP, Chen SN, Wang L, Wang TY (2016). Molecular characterization of a human matrix attachment region that improves transgene expression in CHO cells. Gene 582, 168–172.

Tessadori F, Zeng K, Manders E, Riool M, Jackson D, van Driel R (2010). Stable S/MAR-based episomal vectors are regulated at the chromatin level. Chromosome Res 18, 757–775.

Verghese SC, Kurre P (2017). Anchored lentiviral vector episomes for stem cell gene therapy in Fanconi anemia. Curr Gene Ther 16, 329–337.

Wang X, Xu Z, Tian Z, Zhang X, Xu D (2017). The EF-1α promoter maintains high-level transgene expression from episomal vectors in transfected CHO-K1 cells. J Cell Mol Med 21, 3044–3054.

Wang XY, Zhang JH, Zhang X, Sun QL, Zhao CP, Wang TY (2016). Impact of different promoters on episomal vectors harbouring characteristic motifs of matrix attachment regions. Sci Rep 6, 26446.

Wasag P, Lenartowski R (2016). Nuclear matrix-structure, function and pathogenesis. Postepy Hig Med Dosw (Online) 70, 1206–1219.

Xu Z, Chen F, Zhang L, Lu J, Xu P, Liu G, Xie X, Mu W, Wang Y, Liu D (2016). Non-integrating lentiviral vectors based on the minimal S/MAR sequence retain transgene expression in dividing cells. Sci China Life Sci 59, 1024–1033.

Zhang X, Wang XY, Jia YL, Guo X, Wang YF, Wang TY (2017). A vector based on the chicken hypersensitive site 4 insulator element replicates episomally in mammalian cell. Curr Gene Ther 16, 410–418.