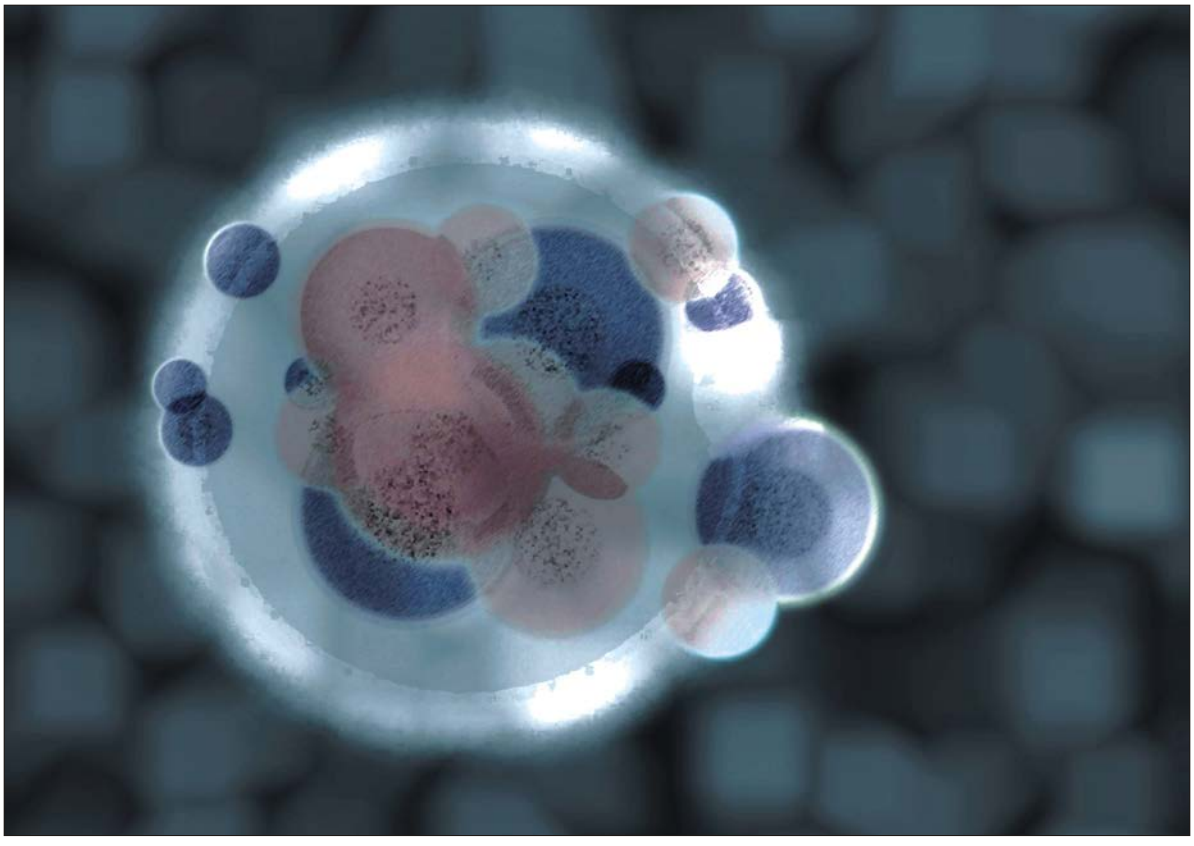

Journal of Biological Sciences

ISSN 1727-3048

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Journal of Biological Sciences

ISSN 1727-3048
DOI: 10.3923/jbs.2016.178.187

Research Article
New Improved Mammalian Expression System

[1]Shahrul Hisham Zainal Ariffin, [1]Nurul Atikah Ahmad, [2]Zafirah Azween Zainal Alam, [1]Intan Zarina Zainol Abidin, [1]Sahidan Senafi, [2]Zulkiflie Zamrod, [3]Rohaya Megat Abdul Wahab and [1]Nurul Yuziana Mohd Yusof

[1]School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
[2]Department of Research and Development Inno Biologics Sdn Bhd, Persiaran Negeri BBN, Putra Nilai, Nilai, 71800 Negeri Sembilan, Malaysia
[3]Department of Orthodontics, Faculty of Dentistry, Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia

Abstract

Background and Objective: Mammalian cells are often used as a host for recombinant protein production although there are some other host such as yeast and bacteria. Nevertheless, more efficient mammalian expression system is needed due to the lack of potential mammalian expression system in producing recombinant protein. Hence, this study evaluates the efficiency of expression vectors that contain Matrix Attachment Region (MAR) and Integration Element (IE). **Methodology:** The MAR and IE sequences were identified using online database; MARFinder, MARScan, SMARTest and NCBI. Transfection on Chinese Hamster Ovary (CHO) cell was carried out using lipofectamine® LTX reagent. Expression analysis for transient transfection was carried out 24 h after transfection while expression analysis for stable transfection was performed from passage 1-4. Expression analysis was done using confocal fluorescence microscope and spectrofluorometer. **Results:** Bioinformatics analysis of these sequences resulted in selecting MAR_7_1_956 and H-1C. The new construct vectors are named pZAAM956 (MAR) and pZAAH1C (IE) and the vector without MAR and IE; pZAAGFP and commercial vector (phrGFP) were act as controls. All expression vectors carried Green Flouresent Protein (GFP) that emit green fluoresence light. Transfection analysis using microscopy approach shows the number of transfected cells for pZAAGFP, pZAH1C and pZAAM956 is more than 80% and there is significant difference ($p<0.05$) between pZAAH1C and phrGFP at stable transfection where the percentage of transfected cells of pZAAH1C is higher than phrGFP. This value is similar with the number of transfected cells for phrGFP. The GFP intensity of pZAAM956 showed the highest intensity measured by spectrofluorometer at emission wavelength 506 nm and excitation 500 nm at cell number 1×10^5 cell mL^{-1}. Paired t-test showed significant difference ($p<0.05$) between pZAAM956 and phrGFP where the GFP intensity of pZAAM956 is higher than phrGFP. Expression analysis of stable transfection shows that the intensity of GFP produced by CHO cells which transfected with pZAAH1C is more stable with low decreasing intensity of GFP when compared to other expression vectors. **Conclusion:** These three constructed expression vectors were functioning; where pZAAM956 shows the highest GFP intensity while pZAAH1C is the most stable.

Key words: Vector construction, mammalian expression system, matrix attachment regions element, integration element

Received: May 23, 2016 Accepted: June 02, 2016 Published: June 15, 2016

Citation: Shahrul Hisham Zainal Ariffin, Nurul Atikah Ahmad, Zafirah Azween Zainal Alam, Intan Zarina Zainol Abidin, Sahidan Senafi, Zulkiflie Zamrod, Rohaya Megat Abdul Wahab and Nurul Yuziana Mohd Yusof, 2016. New improved mammalian expression system. J. Biol. Sci., 16: 178-187.

Corresponding Author: Shahrul Hisham Zainal Ariffin, School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Efficient mammalian expression system is still lacking despite of many studies on constructing mammalian expression system in order to improve the production of recombinant protein have been carried out[1]. Even though there are other potential eukaryotes such as plant, yeast and insect cells; mammalian cells still become the best option to be used as host for recombinant protein production[2-4]. This is due to the ability of mammalian cells to carry out a complete protein folding, post translational modification and can secrete proteins[5]. These characteristics make protein expressed by mammalian cells are more qualified and efficient when compared with other host systems such as bacteria, yeast and plants. Mammalian cells that are frequently used in the production of recombinant proteins are CHO cell[6], NS0 cell[7], baby hamster kidney, BHK cell[8], human embryonic kidney, HEK cell[9] and human retinal, PER-C6 cell[10]. The process of pre-translational modifications in mammalian cells involves glycosylation, phosphorylation of tyrosine, serine and threonine or addition of weak acid chain[11].

Although, mammalian expression system is suitable in the production of recombinant proteins, the production of complex recombinant proteins via mammalian host cell is still unstable. This problem cause low production of recombinant proteins by transfected cell lines[12]. As a result, only a small fraction of clones can produce high amount of targeted protein. Therefore, long term isolation and purification procedures are usually required to identify clones that produce high yield, favourable growth characteristics and product specifications to produce long period of targeted protein[12].

In order to increase the production of these complex recombinant proteins, the efficiency of mammalian cell expression system should be further improved. Most recombinant protein production are low and unstable. This is because the transfered genes are blocked by mammalian cells host due to the epigenetic effects. This problem can be solved with the use of transposal elements. Transposal element is an agent that can lead to genetic changes in eukaryotic genomes. The mobile element containing a transcription regulatory sequence is capable of modifying the genetic program of cells by the entry of mutagenesis or by controlling the transcription of the adjacent gene[13]. Using transposal elements such as MAR can reduce this effect because MAR can regulate the interaction between genes with complex chromatin activation and protect transgenes expression[14]. Retroviruses have long repeated sequence ends Long Terminal Repeat (LTR) which are known as Integration Element (IE) that serves as an intermediary for the integration into the host cell. The presence of these two transposal elements in the expression vector helps to increase and stabilize the mammalian expression system.

MATERIALS AND METHODS

MAR and IE identification: A total of 21 consecutive elements of MAR were obtained from the Inno Biologics Company Sdn. Bhd. These sequences are overlapped sequences determined from three different software; MARFINDER (http://www.futuresoft.org/MarFinder/)[15], the smartest (http://www.genomatix.de/cgi-bin/smartest_pd/smartest.pl)[16] and MARSCAN (http://www.hgmp.mrc.ac.uk/Software/EMBOSS/)[17]. *Mus musculus* genome sequence was used in this study. The given sequences were then analyzed using CLC Genomics Workbench (Denmark) to identify the position of MAR and their distance from the nearest genes. On the other hand, potential IE were obtained from the National Center for Biotechnology Information database; NCBI (http://www.ncbi.nlm.nih.gov). Required IE sequence is a repeated sequence that located at the end of retroviruses region Long Terminal Repeat (LTR).

Cloning of MAR and IE: The NS0 cells were cultured in Dulbecco Modified Eagle Medium (Gibco, USA) supplemented with 5% (v/v) Fetal Bovine Serum (Gibco, USA) incubated at 37°C in humidified incubator containing 5% CO_2. Approximately 1×10^7 cell mL^{-1} cells were harvested and Hi Yield™ Genomic DNA Mini Kit (Real Biotech Corporation) was used to extract the genome of the cells. MAR sequence was prepared by using genomic PCR (genome of NS0 cells as template) and cloned into the multiple cloning site of pGEMT-easy using the appropriate restriction enzyme (*Eco*RI). IE sequence was synthesized and cloned into pJET1.2.

Plasmids and constructs: The Green Fluorescence Protein (GFP) was used to test the MAR and IE based on the pZAA vector (InnoBiologic, Malaysia). The GFP sequence was isolated from phrGFP (Invitrogen, USA) through restriction enzymes digestion, *Sma*I and *Not*I (Thermo Scientific, German). The GFP sequence was cloned into the multiple cloning site of pZAA using the same restriction enzymes (*Sma*I and *Not*I). For cloning of MAR and IE into pZAA vector, a multiple cloning sites (*Age*I-*Eco*RV and *Psp*OMI-*Bsr*GI) were inserted into the upstream region of promoter and downstream region of GFP. Vector which contained MAR is named as pZAAM956 while vector contained IE is named as pZAAH1C. The pZAAGFP and commercial vector (phrGFP) were act as controls.

Cell culture and transfection: Chinese Hamster Ovary (CHO) DG44 cells were cultured in Alpha Modified Eagle Medium; AMEM (Gibco) supplemented with 5% (v/v) Fetal Bovine Serum; FBS (Gibco, USA) incubated at 37°C in humidified incubator containing 5% CO_2. A day before the transfection, CHO cells were seeded in 6 well plates at cell density 1.6×10^5 cell per well in a final volume of 2 mL. When the cells reached 70-80% confluent, the medium was removed and 0.6 mL serum-free medium, Opti-Mem® (Invitrogen, USA) was added. About 2000 ng plasmid was dissolved in 100 µL Opti-Mem® and 5 µL Plus Reagent; this mixture was labeled as mixture A. For mixture B, 3 µL Lipofectamine LTX and 100 µL Opti-Mem® were mixed together. Mixture A was incubated at room temperature for 15 min before added into mixture B. Then, this cocktail was incubated again for another 15 min at room temperature. This step is to allow the DNA-reagen PLUS-Reagen Lipofectamine® complexes to form. After 15 min of incubation, these complexes were added into the culture and incubated for 3 h before added with 1.2 mL of AMEM supplemented with 10% (v/v) FBS. For transient transfection, all vectors were in circular form. Meanwhile, for stable transfection, all constructed vectors; pZAAH1C, pZAAM956 and pZAAGFP were linearized using restriction enzyme, *Nde*II. On the other hand, commercial vector, phrGFP (Invitrogen) was linearized using *Hind*III.

Expression and stability analysis: Two approaches were used to analyze expression and stability of constructed vectors; microscopy and spectrometry analysis. These analysis were done for both transient and stable transfection. Expression analysis for transient transfection was done 24 h after transfection meanwhile for stable transfection, analysis was carried out at every passage starting from passage one until four. For microscopy analysis, the images were captured and numbers of transfected and non-transfected cells were determined to calculate the transfection efficiency. The cells were visualized under confocal fluorescence microscope LSM 5 PASCAL (Laser Scanning Microscope-Ziess, Germany). On the other hand, spectrometry analysis was done to analyze the fluorescence intensity of cells in the population based on the OD reading that expressed in Relative Fluorescence Unit (RFU). The emission wavelength used to detect the fluorescence is at 506 nm and the excitation wavelength is at 500 nm.

Statistical analysis: Paired t-test was used to determine the difference between the data using SPSS version 15.0 (SPSS Inc., USA). Significant value is the data showed $p<0.05$.

RESULTS

Identification of MAR and IE: A total of 21 MAR sequences received from Innobiologics Company Sdn. Bhd. were analyzed using CLC-BIO Genomic Workbench (Denmark) to identify positions and distances of the MAR from the nearest genes. This analysis was based on the MAR on 3 different contig (NT_039433, NT_039437 and NT_166306) in the chromosome 7 of mouse genome. From our analysis, 21 MAR position can be identified either located at intragene or intergene. Among these 21 MAR sequences; 12 are located at intergene and the rest are located at intragen. Eleven MAR sequences that located at intergene were used in the next analysis, while MAR sequences at intragene were eliminated from the analysis.

Analysis using software Genomic Workbench CLC-BIO can also determine MAR position; whether at the 5' or 3' from the nearest genes. From 12 intergenic MAR sequences, 8 were located at 5' of the nearest genes. Only 4 MAR sequences were located at 3' position. The MAR located at 5' of the nearest genes were chosen for the next analysis. The analysis was resulted in choosing MAR_7_1_956 which located at intergene and 5' from nearest gene. The identification of IE was focused on retroviruses. Long Terminal Repeat (LTR) is used as the search key. Database National Center of Biotechnology Information, NCBI (http://www.ncbi.nlm.nih.gov) was used in order to search the potential IE sequence. The searching was resulted in choosing 230 bp MT4 integration site 3' LTR HTLV-1 (S80213). The name given to the sequence in this study is HTLV-1C.

Construction of expression vectors: Expression vector is a plasmid that designed for expression of a protein in the cell. The vector is used to carry genes of interest into chosen host cells. By using this method, the mechanism of the cell can be controlled in order to produce the desired protein encoded by the gene. Green fluorescent protein (GFP) was isolated from commercial vector, phrGFP (Fig. 1a). The *Sma*I and *Not*I were used to digest the GFP sequence in phrGFP. These restriction enzymes were found in both pZAA and phrGFP vector area. After the GFP was subcloned into pZAA vector, sequencing is carried out to ensure that the GFP sequence is in a correct orientation. The pZAA vector that contains GFP was named as pZAAGFP (Fig. 1b). The MAR_7_1_956 and HTLV-1C were selected to be ligated into a pZAA vector to study the role of MAR and IE in improving mammalian cell expression system. Vector containing MAR_7_1_956 was named as pZAAM956 while vector containing HTLV-1C was named as pZAAH1C.

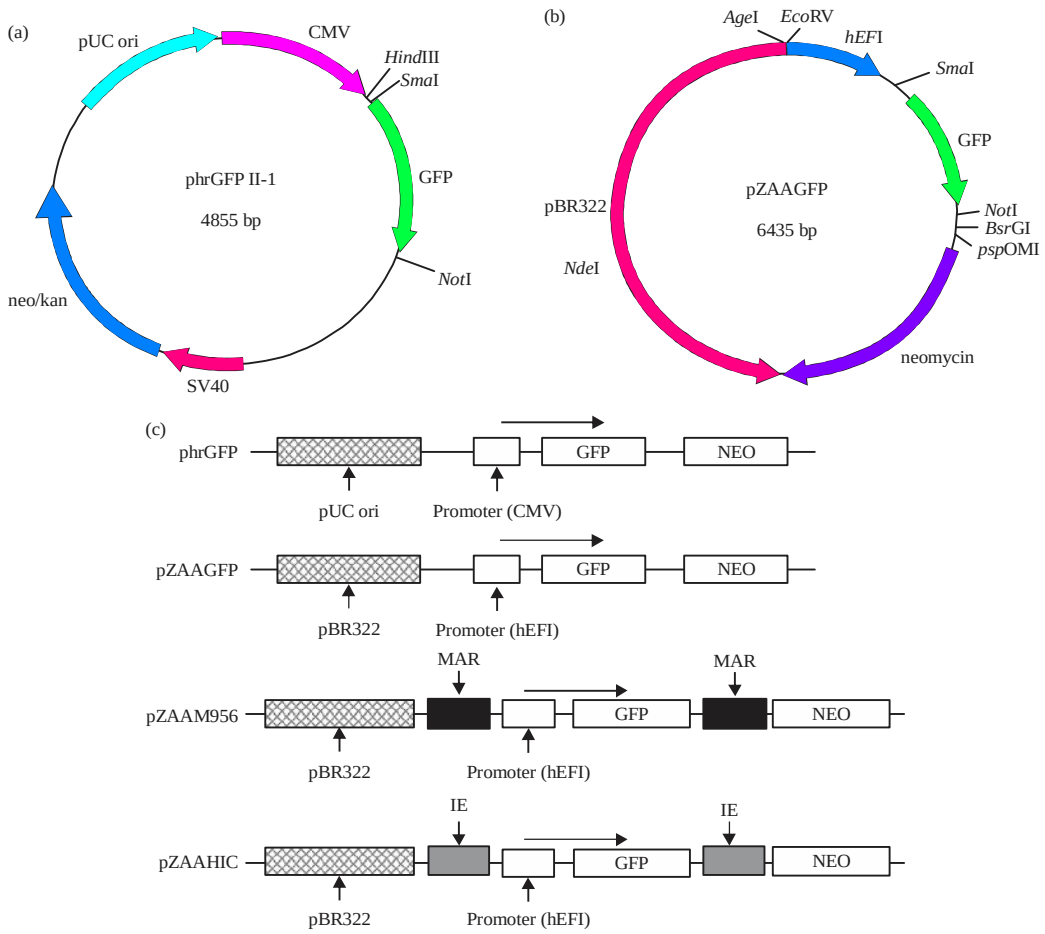

Fig. 1(a-c): Schematic diagrams of constructed expression vector, MAR and IE were flanking ligated between promoter and GFP. (a) Commercial vector, phrGFP, (b) pZAAGFP, pZAA vector that already ligated with GFP and (c) Mapping for all expression vectors

Subcloning was carried out step by step, starting with subcloning of MAR and IE at *Age*I-*Eco*RV site, followed by at *Psp*OMI-*Bsr*GI site and ends with the subcloning of GFP at *Sma*I-*Not*I site. The pZAA vector that already contained MAR, IE and GFP were sequenced to determine appropriate orientation of each element. Therefore, from the sequencing result, all constructed vectors; pZAAGFP, pZAAM956 dan pZAAH1C was successfully developed. These vectors were transfected into CHO cells in order to identify their efficiency.

Transfection into mammalian cell: The aim of transfection is to study the function of genes or gene products as well as to produce recombinant proteins in mammalian host cells[18]. The selection of either stable or transient is dependent on the objectives of the experiment. The expression vectors were successfully constructed and transfected into CHO cells. Two approaches were used to analyze the transfection efficiency of constructed vectors and the intensity of transgene (GFP). The approaches used were microscopy and spectrometry approaches. For analysis using microscopy approach, successful transfected cells were observed under confocal fluorescent microscope under magnification X400 with consistent parameters for all samples. Another approach used was spectrometry approach that applies spectrofluorometer with emission wavelength at 506 nm and the excitation wavelength at 500 nm to identify the intensity of GFP.

For transient transfection, transfected cells were usually analyzed after 24-72 h of transfection. In this study, the expression vectors used in transient transfection (phrGFP, pZAAGFP, pZAAM956 and pZAAH1C) were in full circle form and were analyzed 24 h after transfection. The advantage of choosing transient transfection is the expression analysis can be done in short duration because antibiotic selection is not required. The selection of stable transfected cells and expression analyses of the transfected cells were carried out in this study. Selection on stable transfected cells was done using selection medium which contain geneticin. For stable transfection, the transfection vectors were in linear form.

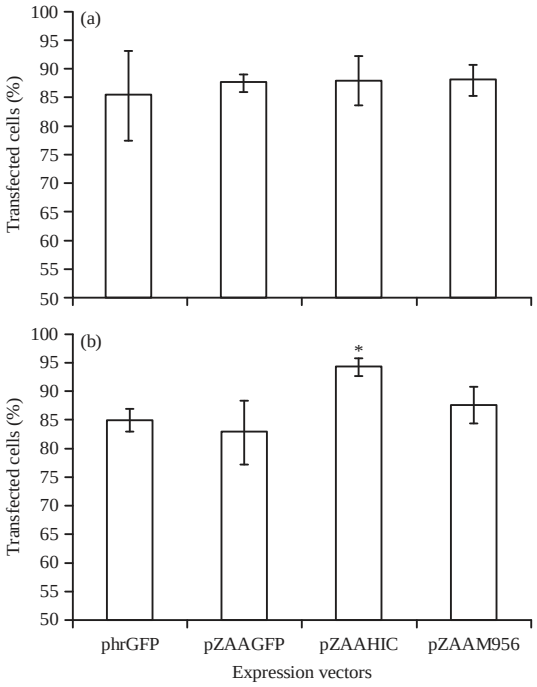

Fig. 2(a-b): Percentage of transfected cells for transient and stable transfection. Both transfection methods, (a) Transient and (b) Stable showed that the percentage of transfection cells are above 80%. *Paired t-test showed a significant difference ($p<0.05$) between the percentage of cells for vector pZAAH1C with phrGFP. Data are based on the average value of three different experiments (standard deviation)

Protein expression analysis: For transient transfection analyses, analyses were carried out after 24 h of transfection. The numbers of cells that are successful and unsuccessful transfected were then calculated to obtain the percentage of cells which is equivalent to transfection efficiency. The percentage of cells that successful transfected is shows in Fig. 2a. The transfection efficiency for constructed vectors are more than 80%. Paired t-test analysis showed that there is no significant difference ($p<0.05$) between the constructed vectors with phrGFP for transient transfection. This result indicates that the expression efficiency of constructed vectors are similar to the commercial vector.

For stable transfection analysis, the expression analyses were carried out on each passage starting from passage 1-4. Similar to transient transfection, after 24 h of transfection, the images of transfected cells were captured (Fig. 3a) and the percentage of successful transfected cells was counted. Figure 2b shows that more than 80% of cells were successful transfected with constructed expression vectors and there is significant difference ($p<0.05$) between pZAAMH1C with phrGFP. This indicates that the transfection efficiency of pZAAH1C into host cells is better than phrGFP. Through microscopy analysis, it can clearly see that the constructed expression vectors (pZAAGFP, pZAAH1C and pZAAM956) showed similar transfection efficiency with a commercial vector (phrGFP) for transient transfection (Fig. 3a) and higher (pZAAH1C) for stable transfection (Fig. 3b). Protein expression analysis using spectrometry approach in transient transfection found that cell culture transfected with pZAAM956 showed the highest GFP intensity compared to other expression vectors (Fig. 4a). There was a significant difference ($p<0.05$) between pZAAM956 and phrGFP.

DISCUSSION

Mouse genome was also used by Purbowasito and coworkers to identify MAR[19]. They were successfully identified 52 MAR when conducted *in silico* analysis of chromosome 7 in mouse. MAR sequence were analyzed and from 21 MAR sequences, 11 intergenes MAR sequences were used for futher analysis, while MAR sequences at intragene were eliminated from the analysis. Study by Martins *et al.*[20] showed that there was negative correlations between intragene MAR and the level of transcription due to their down-regulation function[21]. The study by Linnemann *et al.*[22] also stated that the intragene MAR has strong correlation to gene silencing. Gene silencing may suppress gene expression that eventually make the desired protein unable to produce.

The study by Rudd *et al.*[21] also proved there was a similar relationship between the numbers of intergene MAR with the numbers of genes. There were 21,707 MAR located at intergene and the number of genes found near MAR were 26,027. Study by Tetko *et al.*[23] has concluded that the presence of intergenic MAR correlated with gene expression level. Intergenic MAR not only organized the structure of chromosomes but also act as a dynamic DNA element that has important regulatory functions in gene expression. This selection is based on the study by Linnemann *et al.*[24] which indicated that MAR are mostly located at intergene regions. In their study, they found about 775 MAR sequences and 439 of them were located at intergene. The main purpose of this study is to improve mammalian cell expression system; therefore factors that can inhibit protein target production should be eliminated. Consequently, only 12 MAR sequences that located in the intergenic regions meet these criteria and used for further analysis.

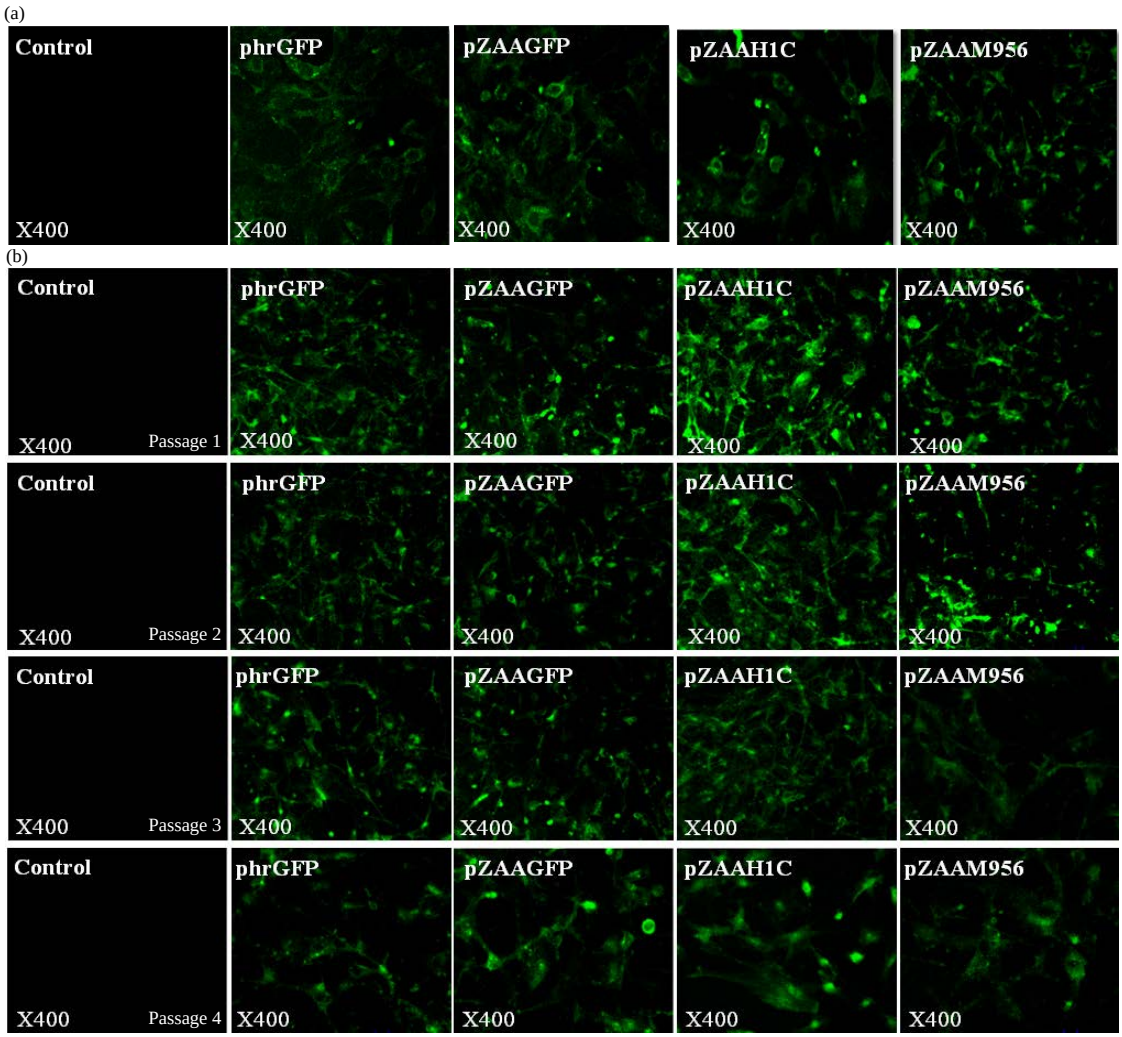

Fig. 3(a-b): Microscopy analysis, (a) For transient transfection, images were taken 24 h after transfection and (b) For stable transfection, images were taken at every passage start from passage 1 until 4 using confocal fluorescence microscope at magnification X400

Selection of MAR sequences that located at 5' gene were based on study conducted by Van der Geest and Hall[25] which proved that the presence of MAR at the 5' regions can lead to the expression of the gene eventhough inhibitor for enhancer was added into the expression system. Glazko *et al.*[26] reported that the MAR found in vertebrates were identified within intergenic regions before the end of the 5' of the nearest genes. MAR located at the 5' genes were directly proportional to the level of transcription and can mediate the chromosome arrangement and also facilitated the expression[24]. This strengthen that the MAR located at 5' of the nearest gene involved in regulating transcription. Therefore, MAR_7_1_956 was selected to be isolated due to its location at 5' end and near to the gene.

The IE was identified from the sequence of the virus is from Human T-cell lymphotropic virus type-I, HTLV-1. The HTLV-1 is a human retrovirus group which cause adult T-cell leukemia[27]. Study by Doi *et al.*[28] find and classify more than 56 integration site of HTLV-1. Selection of HTLV-1C is reinforced by the findings of the 541 integration sites of HTLV-1 in HeLa cells[29]. Based on previous studies, LTR in HTLV-1 have potential to be used in this study as an IE.

Both MAR and IE were located at the upstream of promoter and downstream of the GFP sequence (Fig. 1c). Based on study by Allen *et al.*[30], the transgene that placed between two MAR, will be highly expressed and more stable. This is related to the characteristic of MAR itself which can form a loop that can be found in the nuclei of

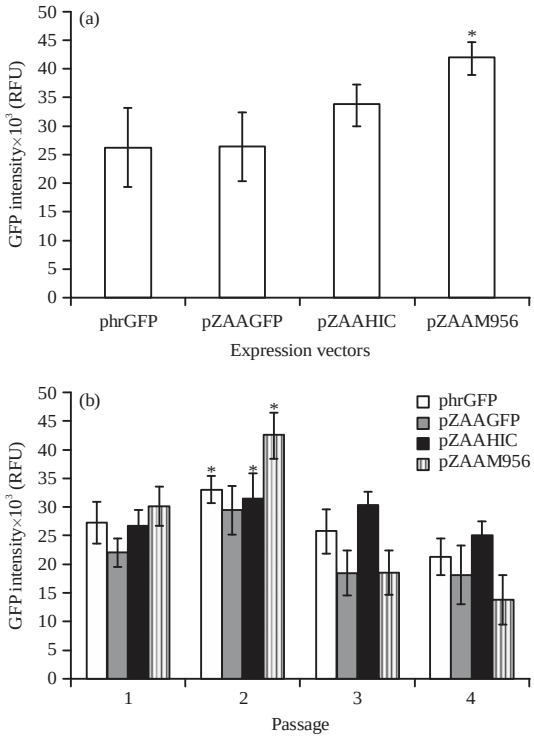

Fig. 4(a-b): Intensity of the Green Fluorescence Protein (GFP) for transient and stable transfection, (a) For transient transfection, *Paired T-test showed a significant difference ($p<0.05$) between the intensity of GFP culture and pZAAM956 pZAAH1C compared phrGFP and (b) For stable transfection, pZAAH1C vector showed a low and stable GFP intensity compared to other vectors. *Paired t-test showed a significant difference ($p<0.05$) between the GFP intensity for the culture at passage 2 as compared to passage 1. No significant difference was shown by the intensity of GFP in the passage 3 and 4 against the passage 1. The data are the average value based on three different experiments (standard deviation)

histone-depleted nuclei and chromosomes at metaphase stage when observed under the electron microscope[31,32]. Study by Iarovaia *et al.*[33] found that MAR can promote expression by forming loops. Similar to MAR, IE sequences were also ligated between GFP. Study by Ley *et al.*[34] found a plasmid that contains LTR; which is an IE that can help in gene expression when the IE were placed between the transgene. The GFP is widely used in studies involving protein expression, for example, a study by Gindullis *et al.*[35] has used GFP as a marker to indicate the position of expressed protein. The GFP was used in this study to determine the number of transfected cells and also to identify the expression intensity. The purpose of using the GFP was also used by Girod *et al.*[1] that observed the involvement of MAR from human genome in the expression of targeted protein.

Transient transfection requires expression analysis in a specific time period only and do not integrate into the genome of host cells[36]. Pereira and Morrell[37] used this transient transfection method to identify the appropriate medium that can be used to induce transfection. By using this method, the results can be obtained quickly. On the other hand, for stable transfection the expression vector is typically contains a marker gene for the selection of the transgene that integrated into the host genome[38]. Stable transfection is carried out to determine the efficiency and stability of the constructed expression vectors.

The efficiency of gene delivery into host cells can be identified using a reporter gene and visualized the expression product. In this study, GFP acts as reporter gene that can help to identify successful transfected cells. Study by Halweg *et al.*[39] was used confocal microscopy to capture the images of successful transfected tobacco cell lines with a vector containing GFP gene. Expressed GFP which emits green light were viewed under a confocal microscope. The same approach was used in this study. Significant difference ($p<0.05$) between pZAAM956 and phrGFP in protein expression analysis using spectrometry approach indicates that the presence of MAR can assist and improve protein expression. Other studies also found that MAR has similar ability to increase the level of transgene expression and degrade the variation between transgenes when MAR exist in the area between the transgene[33,40,41]. In addition, other studies also prove MAR is a repetitive sequence found in the genome of mammalians and MAR located between genes can increase the expression *in vivo* and *in vitro*[42,43].

This study was supported by other studies which proved that the presence of MAR can increase gene expression. The MAR can cause chromatin loop formation and intermediate gene will come off and become a free domain. This allows transcription to occur and increase the transgene expression[44]. The effects of MAR on increasing transgene expression has been shown by various studies. Some studies Gill *et al.*[45], Kwaks and Otte[46] and Wang *et al.*[47] showed that the effects of MAR are varies due to MAR origin, type of transformation approached and expression system used.

Spectrometry analysis was also done to determine the intensity of each cell in each passage at stable transfection. Results shows that the intensity of GFP for phrGFP, pZAAH1C and pZAAM956 at passage 2 have significant difference

($p<0.05$) as compared to the same vector in passage 1. This showed there was an increase in the intensity of GFP from passage 1-2. The intensity of the GFP then start to decrease at passage 3 and 4. The analysis shows that cells transfected with pZAAH1C is more stable with a slight decline in expression compared to the other expression vectors (Fig. 4b).

Factor that could influence the selection of such integration is the structure of nucleosomes and DNA-binding proteins. The pZAAH1C contain LTR sequences (HTLV-1C) which are repeated sequences used by viruses to introduce genetic material into host cells. The LTR also act as a center for gene expression[48-50]. Integration of this element and genes into host cells will make gene expression become more stable. From this spectrometry analysis, MAR and IE have a good function in improving the mammalian expression system. The MAR helps in increasing the level of protein expression, while IE help in stabilizing protein expression.

CONCLUSION

The CHO cells transfected with novel expression vector, pZAAGFP, pZAAM956 and pZAAH1C were able to emit the green fluorescence light which signifies that all constructed expression vectors are functional. The MAR contributes to the increment of GFP expression, while IE contributes in stabilizing GFP expression which resulting in more efficient mammalian expression system.

ACKNOWLEDGMENTS

We would like to thank Dr. Zulkifli Mustafa from Universiti Sains Malaysia for practical assistance with the used of confocal microscope and Prof. Dr. Gires Usop for the lab facility in Universiti Kebangsaan Malaysia. The authors also would like to thanks Inno Biologics Company Sdn. Bhd. for the expression vector (pZAA) and cells and Malaysia Genome Institute for the clean room facility.

REFERENCES

1. Girod, P.A., D.Q. Nguyen, D. Calabrese, S. Puttini and M. Grandjean *et al*., 2007. Genome-wide prediction of matrix attachment regions that increase gene expression in mammalian cells. Nat. Methods, 49: 747-753.
2. Bortesi, L., M. Rossato, F. Schuster, N. Raven and J. Stadlmann *et al*., 2009. Viral and murine interleukin-10 are correctly processed and retain their biological activity when produced in tobacco. BMC Biotechnol., Vol. 9. 10.1186/1472-6750-9-22.
3. Stockmann, C., M. Scheidle, B. Dittrich, A. Merckelbach and G. Hehmann *et al*., 2009. Process development in *Hansenula polymorpha* and *Arxula adeninivorans*, a re-assessment. Microb. Cell Fact., Vol. 8. 10.1186/1475-2859-8-22.
4. Hitchman, R.B., R.D. Possee, A.T. Crombie, A. Chambers and K. Ho *et al*., 2010. Genetic modification of a baculovirus vector for increased expression in insect cells. Cell Biol. Toxicol., 26: 57-68.
5. Zhu, J., 2012. Mammalian cell protein expression for biopharmaceutical production. Biotechnol. Adv., 30: 1158-1170.
6. Kim, J.Y., Y.G. Kim and G.M. Lee, 2012. CHO cells in biotechnology for production of recombinant proteins: Current state and further potential. Appllied Microbiol. Biotechnol., 93: 917-930.
7. Li, J., W. Gu, D.G. Edmondson, C. Lu and N. Vijayasankaran *et al*., 2012. Generation of a cholesterol-independent, non-GS NS0 cell line through chemical treatment and application for high titer antibody production. Biotechnol. Bioeng., 109: 1685-1692.
8. Moi, M.L., C.K. Lim, A. Kotaki, T. Takasaki and I. Kurane, 2010. Development of an antibody-dependent enhancement assay for dengue virus using stable BHK-21 cell lines expressing FcγRIIA. J. Virol. Methods, 163: 205-209.
9. Dietmair, S., M.P. Hodson, L.E. Quek, N.E. Timmins, P. Gray and L.K. Nielsen, 2012. A multi-omics analysis of recombinant protein production in Hek293 cells. PLoS One, Vol. 7. 10.1371/journal.pone.0043394.
10. Butler, M., 2005. Animal cell cultures: Recent achievements and perspectives in the production of biopharmaceuticals. Applied Microbiol. Biotechnol., 68: 283-291.
11. Walsh, G. and R. Jefferis, 2006. Post-translational modifications in the context of therapeutic proteins. Nat. Biotechnol., 24: 1241-1252.
12. Girod, P.A., M. Zahn-Zabal and N. Mermod, 2005. Use of the chicken lysozyme 5' matrix attachment region to generate high producer CHO cell lines. Biotechnol. Bioeng., 91: 1-11.
13. Lisch, D. and J.L. Bennetzen, 2011. Transposable element origins of epigenetic gene regulation. Curr. Opin. Plant Biol., 14: 156-161.
14. Gorman, C., S. Arope, M. Grandjean, P.A. Girod and N. Mermod, 2009. Use of MAR elements to increase the production of recombinant proteins. Cell Line Dev., 6: 1-32.
15. Singh, G.B., J.A. Kramer and S.A. Krawetz, 1997. Mathematical model to predict regions of chromatin attachment to the nuclear matrix. Nucleic Acids Res., 25: 1419-1425.
16. Frisch, M., K. Frech, A. Klingenhoff, K. Cartharius, I. Liebich and T. Werner, 2002. In silico prediction of scaffold/matrix attachment regions in large genomic sequences. Genome Res., 12: 349-354.

17. Van Drunen, C.M., R.G.A.B. Sewalt, R.W. Oosterling, P.J. Weisbeek, S.C.M. Smeekens and R. van Driel, 1999. A bipartite sequence element associated with matrix/scaffold attachment regions. Nucleic Acids Res., 27: 2924-2930.
18. Wurm, F.M., 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. Nat. Biotechnol., 22: 1393-1398.
19. Purbowasito, W., C. Suda, T. Yokomine, M. Zubair, T. Sado, K. Tsutsui and H. Sasaki, 2004. Large-scale identification and mapping of nuclear matrix-attachment regions in the distal imprinted domain of mouse chromosome 7. DNA Res., 11: 391-407.
20. Martins, R.P., R.E. Leach and S.A. Krawetz, 2001. Whole-body gene expression by data mining. Genomics, 72: 34-42.
21. Rudd, S., M. Frisch, K. Grote, B.C. Meyers, K. Mayer and T. Werner, 2004. Genome-wide *in silico* mapping of scaffold/matrix attachment regions in arabidopsis suggests correlation of intragenic scaffold/matrix attachment regions with gene expression. Plant Physiol., 135: 715-722.
22. Linnemann, A.K., A.E. Platts, N. Doggett, A. Gluch, J. Bode and S.A. Krawetz, 2007. Genomewide identification of nuclear matrix attachment regions: An analysis of methods. Biochem. Soc. Trans., 35: 612-617.
23. Tetko, I.V., G. Haberer, S. Rudd, B. Meyers, H.W. Mewes and K.F.X. Mayer, 2006. Spatiotemporal expression control correlates with intragenic scaffold matrix attachment regions (S/MARs) in *Arabidopsis thaliana*. PLoS Comput. Biol., Vol. 2. 10.1371/journal.pcbi.0020067
24. Linnemann, A.K., A.E. Platts and S.A. Krawetz, 2009. Differential nuclear scaffold/matrix attachment marks expressed genes. Hum. Mol. Genet., 18: 645-654.
25. Van der Geest, A.H.H. and T.C. Hall, 1997. The β-phaseolin 5 matrix attachment region acts as an enhancer facilitator. Plant Mol. Biol., 33: 553-557.
26. Glazko, G.V., E.V. Koonin, I.B. Rogozin and S.A. Shabalina, 2003. A significant fraction of conserved noncoding DNA in human and mouse consists of predicted matrix attachment regions. Trends Genet., 19: 119-124.
27. Matsuoka, M. and K.T. Jeang, 2005. Human T-cell leukemia virus type I at age 25: A progress report. Cancer Res., 65: 4467-4470.
28. Doi, K., X. Wu, Y. Taniguchi, J.I. Yasunaga and Y. Satou *et al*., 2005. Preferential selection of human T-cell leukemia virus type I provirus integration sites in leukemic versus carrier states. Blood, 106: 1048-1053.
29. Derse, D., B. Crise, Y. Li, G. Princler and N. Lum *et al* ., 2007. Human T-cell leukemia virus type 1 integration target sites in the human genome: Comparison with those of other retroviruses. J. Virol., 81: 6731-6741.
30. Allen, G., S. Spiker and W.F. Thompson, 2000. Use of matrix attachment regions (MARs) to minimize transgene silencing. Plant Mol. Biol., 43: 361-376.
31. Kumar, P.P., O. Bischof, P.K. Purbey, D. Notani, H. Urlaub, A. Dejean and S. Galande, 2007. Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. Nat. Cell Biol., 9: 45-56.
32. Paulson, J.R. and U.K. Laemmli, 1977. The structure of histone-depleted metaphase chromosomes. Cell, 12: 817-828.
33. Iarovaia, O.V., S.B. Akopov, L.G. Nikolaev, E.D. Sverdlov and S.V. Razin, 2005. Induction of transcription within chromosomal DNA loops flanked by MAR elements causes an association of loop DNA with the nuclear matrix. Nucleic Acids Res., 33: 4157-4163.
34. Ley, D., N. Harraghy, V. Le Fourn, S. Bire and P.A. Girod *et al*., 2013. MAR elements and transposons for improved transgene integration and expression. PloS One, Vol. 8. 10.1371/journal.pone.0062784.
35. Gindullis, F., N.J. Peffer and I. Meier, 1999. MAF1, a novel plant protein interacting with matrix attachment region binding protein MFP1, is located at the nuclear envelope. Plant Cell, 11: 1755-1767.
36. Recillas-Targa, F., 2006. Multiple strategies for gene transfer, expression, knockdown and chromatin influence in mammalian cell lines and transgenic animals. Mol. Biotechnol., 34: 337-354.
37. Pereira, M. and J.I. Morrell, 2011. Functional mapping of the neural circuitry of rat maternal motivation: Effects of site-specific transient neural inactivation. J. Neuroendocrinol., 23: 1020-1035.
38. Glover, D.J., H.J. Lipps and D.A. Jans, 2005. Towards safe, non-viral therapeutic gene expression in humans. Nat. Rev. Genet., 6: 299-310.
39. Halweg, C., W.F. Thompson and W. Spikera, 2005. The Rb7 matrix attachment region increases the likelihood and magnitude of transgene expression in tobacco cells: A flow cytometric study. Plant Cell, 17: 418-429.
40. Liu, S., Y. Hu, X. Wang, J. Zhong and Z. Lin, 2006. High content of resveratrol in lettuce transformed with a stilbene synthase gene of *Parthenocissus henryana*. J. Agric. Food Chem., 54: 8082-8085.
41. Wang, F., T.Y. Wang, Y.Y. Tang, J.H. Zhang and X.J. Yang, 2012. Different matrix attachment regions flanking a transgene effectively enhance gene expression in stably transfected Chinese hamster ovary cells. Gene, 500: 59-62.
42. Moreira, P.N., M. Perez-Crespo, M.A. Ramirez, J. Pozueta, L. Montoliu and A. Gutierrez-Adan, 2007. Effect of transgene concentration, flanking matrix attachment regions and RecA-coating on the efficiency of mouse transgenesis mediated by intracytoplasmic sperm injection. Biol. Reprod., 76: 336-343.
43. Wang, T.Y., R. Yang, C. Qin, L. Wang and X.J. Yang, 2008. Enhanced expression of transgene in CHO cells using matrix attachment region. Cell Biol. Int., 32: 1279-1283.

44. Zhang, Y. and L.C. Yu, 2008. Single-cell microinjection technology in cell biology. BioEssays, 30: 606-610.
45. Gill, D.R., I.A. Pringle and S.C. Hyde, 2009. Progress and Prospects: The design and production of plasmid vectors. Gene Ther., 16: 165-117.
46. Kwaks, T.H.J. and A.P. Otte, 2006. Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. Trends Biotechnol., 24: 137-142.
47. Wang, N., J.D. Tytell and D.E. Ingber, 2009. Mechanotransduction at a distance: Mechanically coupling the extracellular matrix with the nucleus. Nat. Rev. Mol. Cell Biol., 10: 75-82.
48. Holman, A.G. and J.M. Coffin, 2005. Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus and murine leukemia virus integration sites. Proc. Natl. Acad. Sci. USA., 102: 6103-6107.
49. Nowrouzi, A., M. Dittrich, C. Klanke, M. Heinkelein and M. Rammling *et al*., 2006. Genome-wide mapping of foamy virus vector integrations into a human cell line. J. Gen. Virol., 87: 1339-1347.
50. Wu, X., Y. Li, B. Crise, S.M. Burgess and D.J. Munroe, 2005. Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. J. Virol., 79: 5211-5214.