Identification of a Specific Pseudo $\text{attP}$ Site for Phage PhiC31 Integrase in Bovine Genome

Mohammad Hadi Sekhavati 1, Mojtaba Tahmoorespur 1, Farnoosh Jafarpour 3, Kianoush Dormiani 2, Yahya Khazaie 2, Kamran Ghaedi 2,4,*, Sayyed Morteza Hosseini 3, Mahboobeh Forouzanfar 2, Mohammad Hossein Nasr Esfahani 2,3,**

1Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, I.R. IRAN
2Department of Molecular Biotechnology at Cell Science Research Center, Royan Institute for Biotechnology, ACECR, Isfahan, I.R. IRAN
3Department of Reproduction and Development at Reproductive Biomedicine Center, Royan Institute for Biotechnology, ACECR, Isfahan, I.R. IRAN
4Department of Biology, School of Sciences, University of Isfahan, Isfahan, I.R. IRAN

*Corresponding author: Kamran Ghaedi, Department of Molecular Biotechnology at Cell Science research Center, Royan Institute for Biotechnology, ACECR, P.O. Code 816513-1378, Isfahan, I.R. IRAN. Tel: +98-3195015694, Fax: +98-3195015687, E-mail: kamranghaedi@royaninstitute.org
**Co-Corresponding author: Mohammad Hossein Nasr-Esfahani, Department of Molecular Biotechnology at Cell Science research Center, Royan Institute for Biotechnology, ACECR, P.O. Code 816513-1378, Isfahan, I.R. IRAN. Tel: +98-3195015694, Fax: +98-3195015687, E-mail: mh_nasr@royaninstitute.org

Received: July 29, 2013; Revised: September 29, 2014; Accepted: October 08, 2014

Background: PhiC31 integrase system provides a new platform in various fields of research, mainly in gene therapy and creation of transgenic animals. This system enables integration of exogenous DNA into preferred locations in mammalian genomes, which results in robust, long-term expression of the integrated transgene.

Objectives: Identification of a novel pseudo $\text{attP}$ site.

Materials and Methods: Genomic DNA was extracted from primary bovine fetal fibroblast cells, which were stably transfected with EGFP and phiC31 integrase cDNAs carrying vectors. An inverse PCR was carried out for production of minicircle DNAs and followed by sequencing.

Results: A new specific pseudo $\text{attP}$ site termed BF5 was identified in bovine genome. This site is located in an intergenic AT rich region on chromosome 5 with similar features of other mammalian $\text{attP}$ pseudo sites. Furthermore, direct sequencing of generated attL site confirmed that site-specific transgene recombination was occurred at this site.

Conclusions: This finding confirmed that phiC31 integrase could be feasible for production of transgenic animals for biotechnological applications.

Keywords: PhiC31 intagrase; Pseudo $\text{attP}$ site; Protein expression; Recombination

1. Background

Advanced biotechnology based on genetic engineering techniques open a window of opportunity in transgenic animals and livestock improvement. One of the main aspects in the creation of transgenic animals is gene targeting technology, which allows inserting the exogenous genes in the identified sites in mammalian genome. PhiC31 integrase system has been developed for site specific gene integration and used as a tool for gene targeting in mammalian genome modification (1). This system mediated by a member of serine-catalyzed enzymes encoded by phiC31 phage. This phase initiates its life cycle by integrating its whole genome into specific sites named $\text{attB}$ in Streptomyces bacteria genome (2). This recombination accrued when phage DNA sequence named $\text{attP}$ annealed to $\text{attB}$ site in bacterial genome. These sites have often ~30 base pairs (bp) in length (2). In nature, phiC31 integrase pairs a phage $\text{attP}$ site with an $\text{attB}$ site in the host Streptomyces genome (2). Interestingly, Groth et al., (2000) showed that these sites existed in mammalian genome with a diverse homology to native $\text{attB}$ and $\text{attP}$ sites and could be recognized successfully by phiC31 integrase system (1). These sites located in different regions of mammalian chromosome and named pseudo $\text{attP}$ sites (1). PhiC31 integrase system as a
procedure in genome modification has two unique features. The system is transcriptionally active and leads to a long-term expression of integrated gene. Moreover, phiC31 integrase has limited target site in the host genome that helps site specific recombination (3, 4). PhiC31 integrase has proven to work well in a wide range of species, including plants and mammalian cells (1).

Several studies reported recognized pseudo attP sites with diverse homology to wild type of attP and attB sequence in human, mouse, rat and bovine genome (3, 5-6). Due to the attP length, the frequency for existence of this site is very low in mammalian genome (8, 9). These sites usually were located in repetitive elements, such as L1, HERV-L repeats and integration sites (6). A few studies have shown the application of phiC31 integrase system with the aim of creating transgenic cattle by somatic cell nuclear transfer (7, 10). In these studies, four pseudo attP sites found on chromosomes 4, 10, 19 and 28 in bovine genome.

2. Objectives
New pseudo attP site, namely BF5 on bovine fibroblast chromosome, was found that was distinct from earlier reports. For this, an inverse PCR strategy on stably EGFP transformed bovine cell line was implemented. Further direct sequencing and a BLAST search indicated that BF5 site was located on chromosome 5 and exhibited 39% identity to wild type of attP site. BF5 site could be considered for gene targeting towards producing transgenic bovine lines.

3. Materials and Methods
Unless otherwise indicated, all chemicals and media were obtained from Sigma Chemical Company (St. Louis, MO, USA) and Gibco-BRL (BRL, Grand Island, NY, USA).

3.1. Plasmids
Plasmids pCMVInt, pBCPB+ and PBD2 (1, 11) were kindly obtained from Prof. M. P. Calos (Stanford University, USA) (Figure 1). Briefly, pCMVInt

Figure 1. Schematic description of vectors. A: pDB2, donor plasmid containing the phiC31 attB site, the EGFP marker gene transcribed from the CMV promoter, and the neomycin resistance gene (G418) under the control of the SV40 promoter. This vector could be site specific integrated in bovine pseudo attP when co-transfected whit pCMVInt vector. B: pCMVInt, plasmid vector expressing phiC31 integrase under CMV promoter control and carrying the ampicillin resistant gene. C: pBCPB+, interamolecular integration vector assay containing phiC31 attP and attB sites in direct orientation flanking lacZ gene, and chloramphenicol resistant gene as a selectable marker. This vector was used for verifying phiC31 integrase functionality in bovine cells.
encodes phiC31 integrase cDNA and pBCPB+ contains attP and attB sites, which are flanked at both sides of LacZ encoding sequence specific for intramolecular assay. Furthermore, PBD2 encompasses coding regions for EGFP cDNA and one attB site. All plasmids were kept in DH5α strain of E. coli (Invitrogen, USA). Plasmids were purified using QIAprep Spin Miniprep Kit (Qiagen, Germany).

3.2. Primary Cell Culture

Primary bovine fetal fibroblast cell culture was established by isolating the cell from a 65 day old fetus derived from a natural mating. The skin of the fetus was washed extensively in Ca2+ and Mg2+ free phosphate buffer solution (PBS-) containing 1% (v/v) of a cocktail of penicillin-streptomycin and amphotericin B. The skin was cut into 2-3 mm² pieces. The explants were cultured in Dulbecco’s modified Eagle medium F-12 (DMEM/F-12) containing 10% (v/v) fetal calf serum (FCS), 1% (v/v) penicillin-streptomycin and amphotericin B at 37°C in a humidified atmosphere of 5% CO2 until reaching 90-95% confluency. Cells were passaged twice, and were frozen in liquid nitrogen. Cells were thawed and passaged prior to transfection.

3.3. Transfection and Colony Selection

Approximately 2×10⁵ fetal bovine fibroblast cells were cultured in 6-well tissue culture test plate (Orange Scientific, Switzerland) containing DMEM/F-12 enriched with 10% (v/v) FCS until reaching to 60-70% of confluency. At this stage, cells were co-transfected by either 1 μg pBCPB+ or 1 μg PDB2 with 3 μg pCMVInt (1:3 ratio) using Lipofectamine 2000 (Invitrogen, USA) according to the manufacturer’s instruction. The media were changed 6 h post transfection and cells were incubated for one more day. Cells were divided into ten 100 mm culture dishes (Falcon, USA) and cultured for 2 days. Selection was performed with 400 μg.ml⁻¹ of G418 for three weeks. Finally, G418 resistant colonies were isolated, cultured and were expanded for cryopreservation using standard procedures.

3.4. Fluorescence Microscopic Observation and PCR Screening for EGFP Integration

Cryopreserved transfected colonies were thawed and cultured in DMEM/F-12 supplemented with 10% FCS (v/v) and G418. Antibiotic resistant colonies (4) with pronounced growth were cultured on coverslips for microscopic analysis. In order to assess the expression of EGFP in G418 resistant colonies, cells were washed with PBS⁻ and fixed with 4% (w/v) paraformaldehyde for 20 min. After extensive wash with PBS⁺, cells were stained with Hoechst 33342 (bis Benzimide H33342 trihydrochloride). Samples of antibiotic resistant cells were visualized under a fluorescent microscope (Olympus, Japan) and digital images were taken with camera (DP-70, Japan). To ensure for genomic integration of EGFP, PCR was performed using EGFP-F and EGFP-R primers (Table 1). Approximately 10³ cells of each colony were snap frozen in liquid nitrogen and thawed in 37°C water bath and used in PCR. The PCR steps were carried out as follows: 94°C for 5 min as an initial denaturation step, followed by 38 cycles at 94°C for 30 s, 60°C for 45 s, and 72°C for 45 s. Final extension of 72°C for 5 min was performed. PCR products were subjected to 1% (w/v) agarose (CinnaGen, Iran) in TBE buffer (90 mM Tris-HCl, 64.6 mM boric acid, and 2.5 mM EDTA at pH 8.3).

3.5. Analysis of PhiC31Integrase Functionality in Bovine Fibroblast Cells

In order to test the phiC31 integrase activity in bovine fibroblast cells, pBCPB⁺ plasmid was used as an intra-molecular assay vector as it contains both attP and attB sites. In case of phiC31 integrase functionality, two att sites recombine and generate one attL site on pBCPB⁺ plasmid. The recombined site could be recognized by a PCR amplicon with a length of 400 bp (1) (Figure 2A). To do this, bovine fibroblast cells were co-transfected with 3 μg pCMVInt and 1 μg pBCPB⁺ using lipofectamine 2000 (Invitrogen, USA). Simultaneously, cells were transfected with either 3 μg pBCPB⁺ or pCMVInt as control. For plasmid removal from the cell surface, the cells were treated with 50 U/ml DNasel, one day post transfection.

### Table 1. The list of primers used in this study

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence 5’ → 3’</th>
<th>Tm (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>attR-F</td>
<td>GGCAGAAGAAGGAAGGGAAGA</td>
<td>62</td>
</tr>
<tr>
<td>attR-R</td>
<td>ATTAACCCCTCACTAAGGG</td>
<td></td>
</tr>
<tr>
<td>attB-F3</td>
<td>GTAGGTCACGGTCTGAGCG</td>
<td>56</td>
</tr>
<tr>
<td>attRR</td>
<td>GGATCAACTACCGCCACCT</td>
<td></td>
</tr>
<tr>
<td>EGFP-F</td>
<td>ATGGTGAGCAAGGGCGAGGAG</td>
<td>56</td>
</tr>
<tr>
<td>EGFP-R</td>
<td>ATTACTTGTAGCGGCTGGCCTAG</td>
<td>60</td>
</tr>
<tr>
<td>EGFP-F(nested)</td>
<td>CGCCACCATCTTCTTTCAAGGACG</td>
<td></td>
</tr>
<tr>
<td>BoFib-R1</td>
<td>GGTGCTAGGCATTGCGTATAG</td>
<td>56</td>
</tr>
</tbody>
</table>
hours of transfection, cells were harvested and extra-
chromosomal plasmids were extracted using QIAprep
Spin Miniprep Kit (Qiagen, Germany) (12). As a pos-
itive control, pBCPB\(^+\) was recombined in the presence
of \textit{in vitro} translation product of phiC31 integrase as
described previously (13). PCR was performed on the
extracted plasmids using attF and attR primers (Table
1) as previously described (1).

3.6. Identification of Pseudo \textit{attP} Sites in Bovine
Genome

In order to find possible new pseudo \textit{attP} sites rec-
ognizable by phiC31 integrase, an IPCR approach was
implemented as described previously (7). Stably trans-
formed bovine fibroblast cells with pCMVInt and
PDB2 were harvested and genomic DNA was extract-
ed by DNeasy Blood & Tissue Kit (Qiagen, Germany).
Genomic DNA (5 \(\mu\)g) from each sample was digested
with a couple of compatible enzymes, \textit{BgIII} & \textit{BamHI},
(Fermentas, Lithuania). These enzymes recognize two
different sites, but produce similar cohesive ends.
Moreover, both enzymes cut at least one site in PDB2.
The digested fragments were extracted with
phenol/chloroform and precipitated with ethanol. It
was important to use low amounts of DNA for appro-
priate self-circulation of digested DNA in the ligation

![Figure 2](image-url)

\textbf{Figure 2.} PCR screening for phiC31 integrase activity in bovine fibroblasts. A: Schematic representation of \textit{in vitro} phiC31 integrase activity on intra-molecular assay vector B: PCR analysis of recombination in bovine fibroblast cells post co-trans-
fection of intra-molecular assay vector and phiC31 integrase encoding vector. A 400-bp product is indicative of site-specif-
ic recombination. Lane 1, bovine fibroblast cells co-transfected by pBCPB\(^+\) (intra-molecular assay vector) and pCMVInt
(integrate encoding vector); lane 2, bovine fibroblast cells transfected by pBCPB\(^+\) alone; lane 3, untransfected bovine
fibroblast cells; lane 4, bovine fibroblast cells transfected by pCMVInt alone; lane 5, \textit{in vitro}-product of intra-molecular
assay vector as a positive control and lane 6, no template PCR reaction as a negative control.
reaction for efficient inverse PCR. Thus, various amounts of DNA (0.5 to 5 ng) were prepared and used for ligation using DNA Ligation Kit (TaKaRa, Japan) as described in manufacturer’s protocol. Semi-nested PCR was performed across the left junction of assumed recombination site according to our earlier work (14). The circulated DNA was used as a template for the first round of PCR utilizing EGFP-F and attB-F3 primers (Table 1). The first round PCR product (1 μl) with no dilution was used as template for the second round PCR; using attB-F3 and EGFP-F (nested) primers (Table 1). The obtained bands from IPCR were purified and ligated into T-vector (InsTAclone™ PCR Cloning Kit, Thermo, USA), and sequenced via Fazapajuh Co. (Iran). To determine genomic location of pseudo attP, the obtained sequences were analyzed by BLAST and BLAT (www.genome.ucsc.edu/cgi-bin/hgBlat) search engines against bovine genome in databases at UCSC.

3.7. PCR Screening for Site Specific Integration

For screening of site-specific recombination junctions (attL), conventional PCR was performed on cell lysates produced via freeze-thaw using BoFib-R1 and attB-F3 primers. Furthermore, to determine whether pDB2 vector was integrated into specific sites of phiC31, PCR was carried out using attRR and attB-F3 primers. In the case of genomic random integration, attB arm is intact and the PCR generated a 145 bp fragment. PCR generated a 167 bp band due to disruption of the attB sequence in the genome of target cells. PCR was carried out as follows: 94°C for 5 min as initial step of denaturation, followed by 35 cycles of 94°C for 30 s; 55°C for 30 s; and 72°C for 20 s, and a final extension period of 72°C for 5 min. Amplification of a 167 bp PCR product was an indicator of site specific recombination by phiC31 integrase.

4. Results

To examine the phiC31 integrase activity in vivo, pBCPB+, intra-molecular assay vector, was co-transfected with pCMVInt encoding phiC31 integrase in the primary bovine fibroblasts. PCR on transfected cell lysate extracts amplified a 400-bp PCR product (Figure 2A). This result indicated the phiC31 integrase proceeded appropriate site specific recombination on pBCPB+ in bovine fibroblast cytoplasm. As a positive control, in vitro product of phiC31 integrase activity was used according to our previous report (13; Figure 2B, Lane 5). The 400 bp band was evident in co-transfected lysate cells indicating phiC31 integrase activity (Figure 2B, Lane 1). Following verification of phiC31 activity in an intra-molecular assay in a bovine fibroblast cell line, the study was conducted by using cells
which stably were co-transformed with pCMVInt and pDB2 encoding EGFP. During which (21 d), the cells were incubated with G418. IPCR was performed on isolated whole genomic DNA from resistant cell lines. Amongst several bands which were amplified in this procedure, two were subsequently extracted from the gel, cloned in T-vector and sequenced. BLAST of one amplified IPCR sequence showed 39% identity in a 28 bp-length with wild type attP (Figure 3). Interestingly, the identity was evident in cutting site of phiC31 integrase (dinucleotide: TT) as seen in identified pseudo attP (Figure 3). BLAST analysis also indicated that the identified pseudo attP site was located on chromosome 5 (NCBI accession No. NW 001495037.1) and named BF5, accordingly (Figure 4). The results of sequencing showed that the phiC31 integrase, integrated PDB2 plasmid into the host genome in a site specific manner without any deletion or insertion in BF5 site. Furthermore, BLAT analysis showed that BF5 site was among long interspersed repeat elements (LINE) in a region between 53025424-53025451 (Figure 3).

After identification of BF5 site in bovine genome, mixed resistant cells were cultured again in presence of G418. Resistant colonies (4) with adequate proliferation rate were picked for EGFP screening (cell lines A, B, C & D). PCR reaction on genomic DNA of all selected colonies amplified a specific 723 bp fragment of EGFP (Figure 5B, upper panel). Data indicated that only one cell line out of four clones demonstrated considerable EGFP expression as observed under a fluorescent microscope (Figure 5A, Line A). While the rest of colonies showed a faint level of fluorescence (Figure 5A, Line B) or no detectable EGFP expression (Figure 5A, Lines C and D). Moreover, PCR screening on BF5 was set using a primer pair (attB-F3 and BoFib-R1; Table 1) amplifying left recombinant junction sequence (attL), which annealed to attB arm and a partial sequence of bovine chromosome 5, respectively (Figure 5C). A 167 bp PCR product was generated due to site specific recombination on BF5 site as observed only in cell line A (Figure 5B, lower panel, Line A). In order to determine whether random integration has been occurred, we screened the genomic presence of intact attB arms using attB-F3 and attRR primers (Table 1) in a PCR generating 145 bp product (Figure 5C). As observed, PCR on cell lines A and D did not yield the expected product, while random insertion was confirmed to occur in cell lines B and C (Figure 5B, middle panel).

5. Discussion

Zinc-finger nucleases (ZFNs), transcription activator like effector nucleases (TALENs), as well as clustered regulatory interspaced short palindromic repeat (CRISPR)/Cas-based RNA-guided are newly introduced systems which have been successfully used for gene targeting via homologous recombination technology (20, 21). Recombination stimulated by these tech-

"Figure 4. Site-specific recombination on bovine chromosome 5 (BF5). The sequence of left handed recombinant junction showed accurate activity of phiC31 integrase on pseudo attP site (gene bank accession number: NW 001495037.1). The relative location of BF5 on bovine chromosome 5 is indicated among the repeated sequences"
technologies increases recombination efficiency, although off-targeting effects, introduction of mutations near the target site, as well as the required high technical support level, are several potential drawbacks of this system (22). These techniques could be replaced or at least combined in many stem cell research areas with the simpler and less technology demanding phiC31 integrase system. The phiC31 integrase system provides a relatively quick, simple, economical and reliable method for stem cell genome modification.

PhiC31 integrase system is mainly used in gene therapy, and in creation of transgenic animals (5, 7 and 15). The most efficient feature of this system is site-specific recombination in host genome (16), which

Figure 5. PCR screening for detection of site specific integration on four different bovine fibroblast cells lines, which co-transfected by pCMVInt and pDB2. A: Morphologic observation of four different cells lines obtained by co-transfected with pCMVInt and pDB2 (A-D) and non-transfected bovine fibroblast (UT). B: PCR screening for detection of integrated EGFP (expected size: 723 bp), site specific integration (no amplification) and amplifying attL on bovine chromosome 5 (expected size: 167 bp). C: Schematic representation of PCR screening. Transfected bovine fibroblast cell lines were cultured for 3 weeks in G418. Scale bar is equal to 5 μm
causes the long-term expression of exogenous transgene with low copy number in host genome (3). In mammalian genome, the insertion sites, named pseudo attP, are subjective by minimal position effect of adjacent sequences. Additionally the random integration usually leads to gene silencing (17). In this paper, we have reported a new intrinsic pseudo attP site on chromosome 5 of cow, named BF5. BF5 with reasonable exogenous gene expression is located at an intergenic region, between two microRNAs, namely MIR763 and MIR2429, BF5 surrounding sequences were AT-rich (63%), suggesting the site to be appropriate for gene expression (10), similar to earlier reports. Furthermore, BF5 is noted to be among repeated elements similar to what was explained by Chalberg and coworkers. (6). It was shown that phiC31 integrase preferably integrates the DNA fragments in repetitive elements such as L1 and HERV-L repeats (18, 19).

Several reports indicated that most integration events by phiC31 integrase in the mammalian genome result small deletion or insertions of fragments from the vector or the genome (3, 6). The results of attL sequencing in the present study showed that integration of BF5 reacted exactly in TT core without any deletion or insertion (Figure 4). The percentage of identity in core region of 28 bp between BF5 and attP was 39% (Figure 3), which was in the adequate range to be recognized by phiC31 integrase (9, 10). Among bovine fibroblast cell clones transfected by EGFP and phiC31 integrase encoding vectors, a cell line (LineA) was detected with an integration of exogenous DNA at BF5 site. In this cell line, a persistent expressed EGFP was observed more than three weeks under treatment with G418 (Figure 5).

Primer pairs specific for BF4, BF10, BpsF1 and BpsM1 pseudo attP sites amplifications were used (10). However, no product was yielded (data was not shown). Accordingly, it can be concluded that the integration in cell Line D may be occurred in a randomized manner, causing a disruption into attB arm either via double strand break or insertional mutagenesis. To address this hypothesis, more experiments are needed to be performed.

In conclusion, we found a new pseudo attP in bovine genome, which recognized by phiC31 integrase at the precision of integration and could be a proper site for expression of exogenous transgene in bovine fibroblast cells. Moreover, these findings could be used in future animal transgenic studies in the aim of creation of transgenic bovine.

Acknowledgements

The authors send their gratitude to Prof Michelle Calos (Stanford University, USA) for his kind gifts: the vectors carrying phiC31 integrase coding sequence and expression recombinant vector transferring attB sites for efficient integration.

Funding/Support

This study was supported fully by a grant in aid of research from Royan Institute.

Conflict of interest

None of the authors has any conflicts of interest to disclose and all authors support submission to this journal.

Implication for health policy/practice/research/medical education

This project provides feasible application of PhiC31 integrase for generation of transgenic animals.

References


