

RESEARCH ARTICLE

Construction of a universal recombinant expression vector that regulates the expression of human lysozyme in milk

Shen LIU^{1*}, Shengzhe SHANG^{2*}, Xuezhen YANG¹, Huihua ZHANG¹, Dan LU (✉)^{2,3}, Ning LI (✉)²

¹ School of Life Science and Engineering, Foshan University, Foshan 528000, China

² State Key Laboratory of Agrobiotechnology, China Agricultural University, Beijing 100193, China

³ Shanghai Institute of Medical Genetics, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai 200040, China

Abstract The mammary gland provides a novel method for producing recombinant proteins in milk of transgenic animals. A key component in the technology is the construction of an efficient milk expression vector. Here, we established a simple method to construct a milk expression vector, by a combination of homologous recombination and digestion-ligation. Our methodology is expected to have the advantages of both plasmid and bacterial artificial chromosome (BAC) vectors. The BAC of mouse whey acidic protein gene (*mWAP*) was modified twice by homologous recombination to produce a universal expression vector, and the human lysozyme gene (*hLZ*) was then inserted into the vector by a digestion-ligation method. The final vector containing the 8.5 kb *mWAP* 5' promoter, 4.8 kb *hLZ* genomic DNA, and 8.0 kb *mWAP* 3' genomic DNA was microinjected into pronuclei of fertilized mouse embryos, to successfully generate two transgenic mouse lines that expressed recombinant human lysozyme (rhLZ) in milk. The highest expression level of rhLZ was 0.45 g·L⁻¹, and rhLZ exhibited the same antibacterial activity as native hLZ. Our results have provided a simple approach to construct a universal milk expression vector, and demonstrated that the resulting vector regulates the expression of hLZ in milk.

Keywords BAC recombinant methods, gene expression, human lysozyme, transgenic mice, milk expression vector

1 Introduction

Transgenesis provides a novel platform to express foreign protein in mammary glands used as bioreactors, and

the resulting mammary gland bioreactors have been considered better than other bioreactors, such as the blood, urine, seminal plasma, egg whites and silkworm cocoons^[1]. There are many advantages for the use of the mammary gland as a bioreactor. The mammary gland is a specialized organ for protein synthesis in which milk protein expression levels reach gram levels per liter of milk^[2]. In addition, milk can easily be collected in large quantities, resulting in low production costs. The purification of recombinant proteins from milk is also easy and the mammary gland tissue glycosylation patterns are usually similar to those of human native proteins^[3]. Thus far, many mammalian species (pig, goat, sheep, cow and rabbit) have been studied as bioreactors^[1], while the transgenic mouse model has been an irreplaceable tool for bioreactor studies^[4].

The construction of a milk expression vector is an essential step for protein generation in an animal system. In general, genomic DNA (gDNA) or complementary DNA (cDNA) from the foreign gene is designed to be driven by the 5'-flanking (promoter) and 3'-flanking regions of milk protein genes (such as α , β , γ and κ caseins, α -lactalbumin, β -lactoglobulin, and whey acidic protein)^[5]. Other regulatory elements, such as signal peptide and insulator sequences, can also be inserted into the vector to ensure high level and/or position-independent expression. The vector construction method most widely used involves using DNA ligases to repeatedly insert the gene of interest and the regulatory DNA fragment^[6], or to insert the gene of interest into commercial vectors such as pBC1^[7]. However, this method can only use short DNA fragments, usually < 20 kb. Statistical analyses have suggested that the plasmid vector transgene often leads to low levels and unstable expression because of position effects and incomplete regulatory sequences^[8].

With the development of homologous recombination methods, artificial chromosome type vectors have become a relatively novel method for gene expression. These

Received October 31, 2017; accepted November 24, 2017

Correspondences: danlu125@163.com, ninglcau@cau.edu.cn

*These authors contribute equally to the work

vectors have large cloning capacities that may include long distance regulatory elements used to achieve correct gene expression and to maximize position-independence, copy number-dependence and optimal levels of transgene expression. In comparison to yeast artificial chromosomes, bacterial-derived bacterial artificial chromosomes (BACs) and P1 artificial chromosome are more stable and suitable for genetic engineering and are often used to generate transgenic animals for gene expression^[9,10]. It has been difficult to regulate BAC foreign gene expression, however, before the development of homologous recombination methods. This method is used to direct cloning and subcloning, including ET recombination (recombination derived from the RecE and RecT proteins)^[11] and recombination-mediated genetic engineering (recombineering)^[12], and allows modifications of BACs or P1 artificial chromosomes with precise junctions without constraints being imposed by restriction enzyme site locations. It has been reported that such a modified BAC regulates the expression of foreign genes^[13] as large as itself^[9]. Presently, a problem involving BAC transgenesis is the length of the DNA sequence insert, making it difficult to operate and identify and makes the analyses of the locations and integration patterns of BACs difficult^[14].

In the present study, we combined the homologous recombination and digestion-ligation methods to take advantage of the plasmid and BAC vectors. The mouse whey acidic protein (*mWAP*) BAC was first subcloned into a plasmid with 8.5 kb 5' and 8.0 kb 3' flanking sequences of the *mWAP* gene, then the *mWAP* gene was deleted by directed cloning with two *NotI*-flanked sites for insertion of foreign genes, and a universal expression vector, pMWAP, was constructed. Finally, human lysozyme (*hLZ*) gDNA was inserted into the vector using the digestion-ligation method, and the final expression vector, pMWAP-*hLZ*, contained the 8.5 kb *mWAP* 5' promoter, 4.8 kb *hLZ* gDNA, and the 8.0 kb *mWAP* 3' gDNA. Transgenic mice were generated by microinjection of the pMWAP-*hLZ* expression vector, and recombinant human lysozyme (rhLZ) was expressed in the milk. Our results have provided a simple approach to constructing a universal milk expression vector, and demonstrate that the resulting vector regulated the expression of hLZ in milk.

2 Materials and methods

2.1 Construction of the pMWAP-*hLZ* expression vector by BAC recombineering

The construction of the pMWAP-*hLZ* expression vector consisted of two recombineering steps and one step for enzyme digestion and ligation. Recombineering was performed as previously described^[13]. Step 1 involved

subcloning the DNA fragments of the *mWAP* 5' and 3' flanking regions from the *mWAP* BAC (GenBank No. RP23-169J19). The *mWAP* BAC containing the entire *mWAP* gDNA and including the 173 kb 5' flanking region and 50 kb 3' flanking region was introduced into the SW102 *Escherichia coli* strain in advance, which contained the bacteriophage λ recombination system on the chromosome for BAC recombineering. Then the PCR product amplified from pBR322 plasmid containing the origin of replication and ampicillin resistance marker (Amp) with a 60 bp homologous arm (Table 1) was electroporated into bacterial cells. Recombineering occurred, the bacterial cells were screened using ampicillin and verified by PCR. The final plasmid from DNA subcloning contained the 8.5 kb *mWAP* 5' promoter, 8.0 kb *mWAP* 3' gDNA, and the *mWAP* gDNA. In step 2, The *mWAP* genomic gene was replaced by a zeocin resistance gene flanked by the *NotI* restriction sites. The PCR product containing the zeocin resistance marker (*Zeo*) with a 50 bp homologous arm (Table 1) was electroporated into the bacterial cells for the second recombineering step, and the positive bacterial clones containing pMWAP-*Zeo* were screened by zeocin and identified by PCR. In step 3, the *hLZ* gene was ligated into the *NotI* sites of the pMWAP-*Zeo* vector in DH5 α *Escherichia coli* strain, and pMWAP-*hLZ* construction was completed (Fig. 1). The 4.8 kb *hLZ* genomic sequence from ATG to TAA flanked by the *NotI* sites was amplified by PCR using ghLZ primers (Table 1) with the BAC clone RP11-1143G9 (Genome Systems Inc., Sunnyvale, CA, USA) as the template.

Step 1, subcloning DNA fragments of *mWAP* 5' and 3' flanking regions from BAC by the linear PCR product. *mWAP* BAC that contains entire *mWAP* gDNA and includes 173 kb 5' and 50 kb 3' flanking regions, was introduced into *Escherichia coli* strain SW102 for BAC recombineering, PCR product containing the origin of replication and ampicillin resistance marker (Amp) with 60 bp homologous arm was electroporated into the bacterial cells, then recombineering occurred and bacterial cells screened by ampicillin. After PCR identification and DNA sequencing, the plasmid contains 8.5 kb *mWAP* 5' promoter, 8.0 kb *mWAP* 3' gDNA and *mWAP* gDNA.

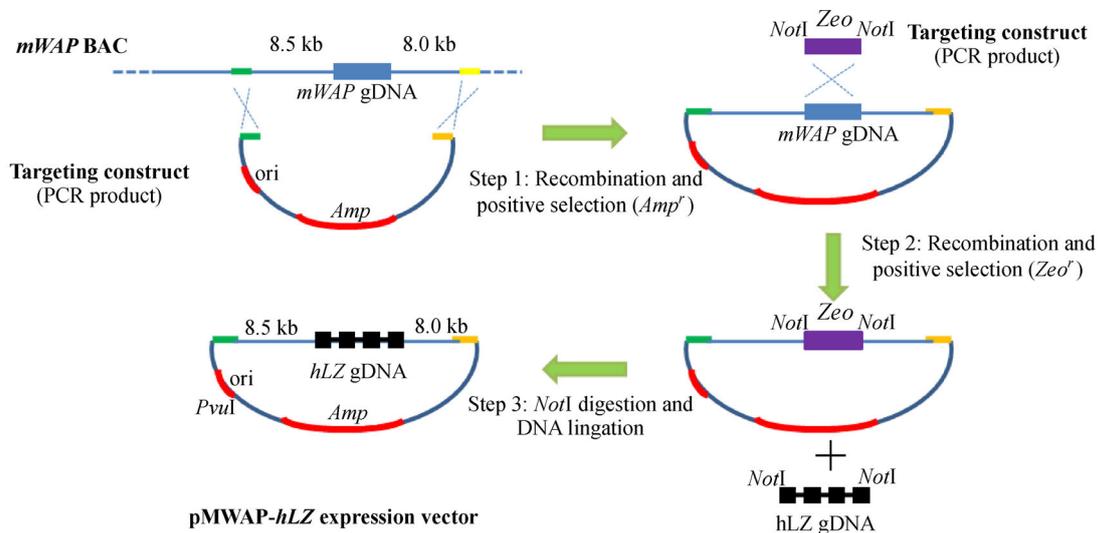
Step 2, *mWAP* genomic gene was replaced by zeocin resistance gene with two *NotI*-flanked. PCR product containing *Zeo* with 50 bp homologous arm was electroporated into the bacterial cells for the second recombineering, the positive bacterial cells screened by zeocin and identified by PCR.

Step 3, *hLZ* gene was ligated by *NotI* sites and pMWAP-*hLZ* expression vector was constructed. The 4.8 kb *hLZ* genomic sequence from ATG to TAA flanking by two *NotI* sites was amplified by PCR, and replaced the zeocin gene after *NotI* digestion and DNA ligation. The final expression vector contains 8.5 kb *mWAP* 5' promoter, 4.8 kb *hLZ* gDNA and 8.0 kb *mWAP* 3' gDNA.

Table 1 Primers for vector construction and PCR analysis

Primer	Primer sequence	Utility
PBR-F	5'-CACTAAGAATTTTATACATTCTCAGCTGGCTGGTGGTATGCCACCACTG-CAACCCTTcgatcgGATACGCGAGCGAACGTGA-3'	PCR amplification of the origin of replication (ori) and ampicillin resistance gene from pBR322 plasmid for subcloning (2763 bp)
PBR-R	5'-ACTACTTCACAACACTATAATTTTCTACTGTTGGTACTTTCAGGTTGCTGACAT-CAGATAATTCTTAGACGTCAGGTGGCAC-3'	
Zeo-F	5'-ATGCATCCCAGACACTCAGACAGCCATCAGTCACTGCTGACACCGG-TACCgcgccgcCAGACATGATAAGATACATTGATG-3'	PCR amplification of zeocin resistance gene from pBudCE4.1 plasmid (786 bp)
Zeo-R	5'-TAGACTTGGGCTGGTCACTCCCACAGGCAGGGATGC-CAGGGCCGGCTGcgcgccgcAGTGAGGAGGCTTTTTTGG-3'	
ghLZ-F	5'-AATgcgccgcATGAAGGCTCTCATTGTTCTG-3'	PCR amplification of human lysozyme gene (4.8 kb) with two <i>NotI</i> restriction sites
ghLZ-R	5'-AATgcgccgcTTACACTCCACAACCTTGAAC-3'	
mWAP-atg-F	5'-AC CCTTGGCACAGTATGG-3'	The PCR product is 932 bp + gene length (<i>WAP</i> is 3.0 kb, <i>Zeo</i> is 1.0 kb, and <i>hLZ</i> is 5.4 kb)
mWAP-taa-R	5'-ATGGAGTTTGGCTGTGA GCTC-3'	
P1-F	5'-GATCCACAGGACGGGTGT-3'	PCR detection of transgenic founders (538 bp)
P1-R	5'-CTCCAGCCCCTATTAGACA-3'	
P2-F	5'-CCGAGTGAATAAATTAGACA-3'	PCR detection of transgenic founders (568 bp)
P2-R	5'-ACGGAAATGTTGAATACTCAT-3'	
hLZ-F	5'-TTATACACACGGCTTAC-3'	Primers for PCR detection and DIG-labeled probe synthesis (637 bp)
hLZ-R	5'-CAGCATCAGCGATGTTATCT-3'	
Exon1-2-F	5'-ATCAGCCTAGCAAACCTGGAT-3'	RT-PCR for <i>hLZ</i> (322 bp)
Exon4-R	5'-CTCCACAACCTTGAACATAC-3'	
GAPDH-F	5'-AGGCCGGTGTGCTGAGTATGTC-3'	RT-PCR for GAPDH control (530 bp)
GAPDH-R	5'-TGCCTGCTTACCACCTTCT-3'	
hLZ-CP-F	5'-TGCTACCAGGGCTGGAGAAT-3'	Real-time PCR for <i>hLZ</i> (140 bp)
hLZ-CP-R	5'-AGCTCCTTCGCCTCTACCA-3'	
Fabpi110-F	5'-TGTTTCAGAGCCAGGAAATCCATA-3'	Real-time PCR for mouse Fabpi gene (110 bp) as control
Fabpi110-R	5'-CATAGGTGTCTCTTTCTTT GGTGTGT-3'	

Note: RT-PCR, reverse transcriptase PCR; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; Underlined, restriction sites, *PvuII*, cgatcg; *NotI*, gcgccgc; the italic, homologous sequence.

**Fig. 1** Strategy for construction of pMWAP-hLZ expression vector by BAC recombineering

2.2 Generation of transgenic mice

The pMWAP-*hLZ* expression vector was linearized with *PvuI* restriction enzyme, purified by agarose gel electrophoresis and diluted to 2–3 ng· μL^{-1} in TE buffer for microinjection. Linearized DNA was microinjected into the pronuclei of fertilized eggs of the mouse breed Kunming White. The injection of transgenic DNA was conducted according to the procedure described by Hogan et al.^[15] gDNA was extracted from tails (1 cm long) of F₀ mice, and P1, P2, *hLZ* primers (Table 1) were used to screen for positive transgenic lines, which were confirmed by Southern blotting. In brief, 10 μg of gDNA from transgenic and wild-type (negative control) mice was digested with *EcoRI*. After resolving by 0.8% agarose gel electrophoresis bands were transferred to nylon membrane (Roche Applied Science, Mannheim, Germany), the samples were hybridized with a digoxigenin-labeled probe amplified with the primer *hLZ* (Table 1; Fig. 1, 637 bp amplification products) to produce a 3.3 kb positive hybridization signal. Transgene copy number was estimated by real-time PCR, the primers (*hLZ*-CP-F and *hLZ*-R, 140 bp) for *hLZ* assay were designed by using Primer Express software version 3.0. The mouse fatty acid binding protein gene (*Fabpi110*-F and *Fabpi110*-R, 110 bp) was used as control (Table 1). Real-time PCR was performed on the LightCycler® 480 II System (Roche) and all reactions were performed in triplicate.

2.3 Reverse transcription PCR

Total RNA was extracted from the mammary glands and other tissues during lactation using Trizol (Tiangen Biotech, Beijing, China). First strand cDNA was synthesized with oligo-dT (Promega, Madison, WI, USA). Reverse transcription PCR (RT-PCR) primers were designed on the basis of the *hLZ* coding sequences, and the upstream primer (Exon1-2-F) was designed across one intron (Table 1). The predicted 322 bp fragment was amplified. Mouse glyceraldehyde-3-phosphate dehydrogenase gene was used as an RT-PCR internal control.

2.4 SDS-PAGE and Western blotting

Milk samples from transgenic mice were collected on postnatal days 9 and 16, diluted 10-fold with phosphate-buffered saline (PBS), and centrifuged (10000 g, 20 min, 4°C) to remove the whey from the fat layer and insoluble precipitates. The cleared fraction lacking whey was mixed with SDS-PAGE sample buffer and subjected to 15% SDS-PAGE under both reducing and non-reducing conditions. Protein bands were visualized by staining with Coomassie brilliant blue R-250.

For Western blot analysis, proteins resolved by SDS-PAGE were electrophoretically transferred to a nitrocellulose membrane (GE Healthcare, Amersham, UK) that was

then blocked overnight at 4°C with 3% bovine serum albumin in PBS containing 0.05% (w/v) Tween 20. Polyclonal rabbit anti-*hLZ* (1:1000) (US Biological Inc., Swampscott, MA, USA) and horseradish peroxidase-conjugated goat anti-rabbit IgG (1:5000) (Sino-American Co., Beijing, China) were used to detect rhLZ. Milk samples from wild-type mice served as negative controls. Blots were developed by enhanced chemiluminescence and autoradiography.

2.5 ELISA

The amount of rhLZ in the milk of transgenic mice was quantified using a Human Lysozyme EIA kit (Biomedical Technologies, Inc., MA). Each sample was analyzed at least three times, and the results represent mean \pm SD.

2.6 Lysozyme activity assays

Lysozyme activity was evaluated by two methods. For the lysoplate or gel diffusion assay, a suspension of *Micrococcus luteus* cells (China General Microbiological Culture Collection Center, Beijing, China) was prepared in nutrient broth medium containing agar. After solidification of the medium, 6-mm-diameter circles of quantitative filter paper were spotted with milk samples to be tested for lysozyme activity and placed on the agar plates. After 36 h incubation at 28°C, zones of transparency appeared around the filter paper circles due to lysis of *M. luteus* cells by lysozyme produced by the transgenic mice. The experiment was repeated three times.

For the turbidimetric assay, the enzymatic activity of lysozyme was determined as previously described^[13] by monitoring the reduction in turbidity of a suspension of *M. luteus* cells at 450 nm. First, 2.5 mL of *M. luteus* cell suspension (OD₄₅₀ 0.80 to 0.85) in 66 mmol·L⁻¹ potassium phosphate, pH 6.24, was placed in a 4-mL cuvette at room temperature. The reaction was initiated by adding 100 μL of 1:100 dilutions of mouse milk samples or 100 μL purified chicken egg lysozyme at concentrations of 1000, 2000, 3000, 4000, 5000, 6000 and 8000 U· μL^{-1} . Absorbance at 450 nm (A₄₅₀) was recorded every 15 s over a 3 min period. The activity of rhLZ was calculated from the standard curve. All samples were measured in triplicate.

3 Results

3.1 Construction of the pMWAP-*hLZ* expression vector

During step 1 of the pMWAP-*hLZ* construction, nine clones were selected from the LB plates containing ampicillin, for PCR analyses, with two being positive. The positive clones were whole genome sequenced, and the sequences were compared to the *mWAP* BAC sequence

using BLAST. The plasmid, pMWAP, contained the 8.5 kb *mWAP* 5' promoter, 8.0 kb *mWAP* 3' gDNA and the *mWAP* gDNA. During step 2, nine clones were selected from the LB plates with zeocin, for PCR analyses, with six of them testing positive. The positive PCR products were further subjected to *NotI* digestion, because of the added two *NotI* sites, with the plasmid being named pMWAP-Zeo. During step 3, the 4.8 kb *hLZ* genomic sequence from ATG to TAA, flanking the two *NotI* sites, was amplified by PCR with *ghLZ* primers (Table 1) and using *hLZ* BAC as the template, and was ligated to the pMD19-T simple vector (Sangon Biotech Inc., Shanghai, China) for sequencing. Both *hLZ* gDNA and pMWAP-Zeo were digested with *NotI* and ligated with the *NotI* sites, and PCR was used to identify the correct direction of connection. The final expression vector, pMWAP-*hLZ*, contained the 8.5 kb *mWAP* 5' promoter, the 4.8 kb *hLZ* gDNA and the 8.0 kb *mWAP* 3' gDNA (Fig. 2a).

3.2 Generation of transgenic mice using the pMWAP-*hLZ* expression vector

Two female (lines 2 and 18) and four male (lines 10, 13, 25 and 26) transgenic mouse lines were obtained from 27 mice analyzed by PCR (Fig. 2b), with an efficiency of 22.2%. The integration of transgenes and copy number was further confirmed by Southern blotting and real-time PCR. All six lines had low transgene copy numbers (ranging from 1 to 10; Fig. 2c). At sexual maturity, the transgenic lines were mated with wild-type mice, and the transgene transmission in the offspring are listed in Table 2.

3.3 Expression of the rhLZ in the milk of transgenic mice

To determine whether the transgene was transcribed and translated correctly in transgenic mice, RT-PCR and

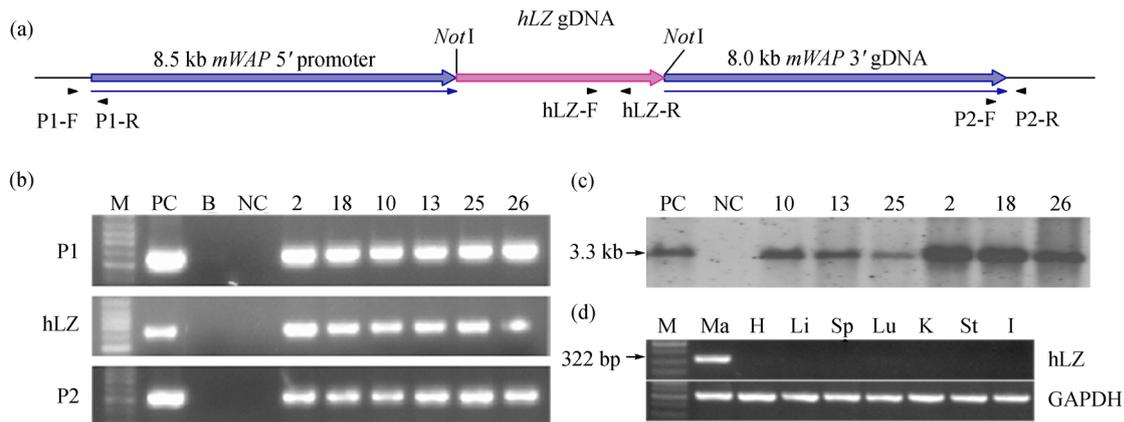


Fig. 2 Molecular characterization of transgenic mice. (a) Schematic presentation of the pMWAP-*hLZ* expression vector used for microinjection containing 8.5 kb *mWAP* 5' promoter, 4.8 kb *hLZ* gDNA and 8.0 kb *mWAP* 3' gDNA. The primers P1, *hLZ* and P2 were shown in arrow; (b) PCR detection of transgenic lines. M, 100 bp DNA ladder; PC, positive control using pMWAP-*hLZ* vector; B, blank control; NC, negative control; transgenic mouse lines 2, 18, 10, 13, 25 and 26; The amplified products of P1, *hLZ* and P2 are 538, 637 and 568 bp, respectively; (c) Southern blot analysis of transgenic lines. *EcoRI*-digested gDNA was hybridized with digoxigenin-labeled probe of amplified *hLZ* fragments. PC, positive control with two copies of pMWAP-*hLZ*; NC, gDNA of non-transgenic mice as a negative control; (d) reverse transcription PCR analysis transgene expression in transgenic mouse tissues. The 322 bp product was abundant in lactating mammary gland of transgenic mice, while not detected in six out of seven other tissues. M, 100 bp DNA ladder; Ma, mammary gland tissue of transgenic line on day 13 of lactation; H, heart; Li, liver; Sp, spleen; Lu, lung; K, kidney; St, stomach; I, intestine. The mouse glyceraldehyde-3-phosphate dehydrogenase gene was used as a positive control for mRNA extraction and loading.

Table 2 Expression of rhLZ in the milk of transgenic mouse lines

Line	Sex	Expression level ^a in F ₀ /(g·L ⁻¹)	Expression level in F ₁ /(g·L ⁻¹)	Antibacterial activity/(U·μL ⁻¹)	Copy number	Germ line transmission
2	Female		0	26±20	6	Yes (6/11)
10	Male		-	-	2	No
13	Male		-	-	2	No
18	Female	0.45±0.05		940±50	6	Yes (2/10)
25 ^b	Male		0.15±0.03 0.14±0.02	230±30 210±25	2	Yes (3/10)
26	Male		0	20±12	5	Yes (2/7)

Note: ^a The concentration of rhLZ was quantified by ELISA; ^b mouse line 25 gave birth to two female offspring (F₁: 25-19 and 25-22)

Western blotting was conducted. For RT-PCR, total RNA was isolated from lactating mammary gland and other tissues (heart, liver, spleen, lung, kidney, stomach and intestine) of the F₁ generation. As expected, hLZ mRNA was detected in the mammary gland of transgenic mice during the middle of the lactation period, but not in other tissues and wild-type mammary glands (Fig. 2d). For Western blotting milk samples from transgenic mice were collected (days 9–14) and diluted, natural hLZ standard was used as a positive control and milk from wild-type mice as a negative control. The mouse line 18 and F₁ offspring from line 25 (25-19 and 25-22) expressed rhLZ and showed a 14.7 kDa band with the same protein size as hLZ standard (Fig. 3). The concentration of rhLZ was further quantified using ELISA assay, the rhLZ expression level of line 18 and F₁ offspring was 0.45 ± 0.05 , 0.15 ± 0.03 and 0.14 ± 0.02 g·L⁻¹, respectively (Table 2).

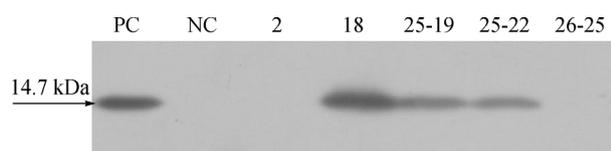


Fig. 3 Western blot detection of rhLZ expressed in milk of transgenic mice

Milk samples were diluted 1:10 in phosphate-buffered saline, and 3 μ L of each sample was separated by SDS-PAGE under reducing conditions. PC, 0.5 μ g natural hLZ standard (Sigma; 14.7 kDa); NC, milk of non-transgenic mice as a negative control; 2 and 18 are milk from female mouse lines; 25-19, 25-22 and 26-25 are milk from different female F₁ offspring of male mouse lines.

3.4 Assessment of rhLZ antibacterial activity

The lysoplate provided a convenient method to analyze the bactericidal activity and to estimate the expression levels of rhLZ in milk. Transparent zones (filled with 1 μ L samples of milk) of mouse lines 18 and 25, and the hLZ standard were clearly visible after an incubation for 24 h, while there was no transparent zone with milk from mouse lines 2 and 26, and from non-transgenic mice (Fig. 4). Milk samples were then quantitatively examined using the turbidimetric method. The antibacterial activities of the transgenic mice were 940 ± 50 and 230 ± 30 U· μ L⁻¹ (Table 2), whereas that of the non-transgenic mice and hLZ standard were 20 ± 12 and 1528 ± 275 U· μ L⁻¹, respectively.

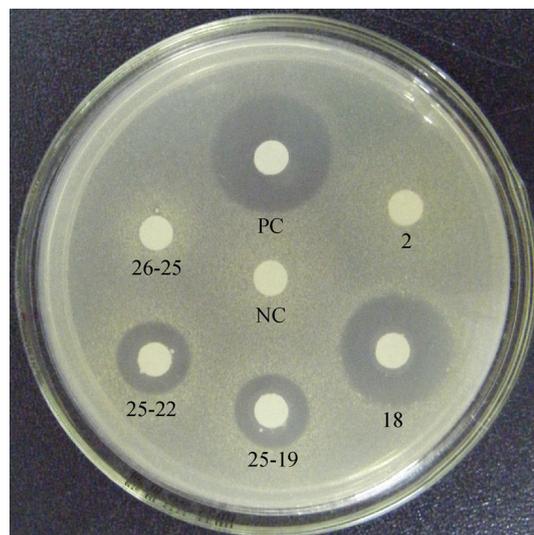


Fig. 4 Lysoplate assay for lytic activity of rhLZ against *M. luteus*. Small white circles are 6-mm quantitative filter papers spotted with 10 μ L of 1:10 diluted milk sample. Transparent zones around the quantitative filter papers indicate bacterial lysis. PC, 0.5 μ g natural hLZ standard; NC, milk of non-transgenic mice as a negative control; 2 and 18 are milk from female mouse lines; 25-19, 25-22 and 26-25 are milk from different female F₁ offspring of male mouse lines.

development of animal bioreactor systems. Usually, the 5'-flanking regions as promoters and the 3'-flanking regions of milk protein genes are cloned repeatedly and ligated together to drive foreign gene expression. The most well-known universal milk expression vector is pBC1 (Invitrogen, Carlsbad, CA, USA), designed to facilitate the expression of recombinant proteins in milk of transgenic animals. This vector contains two chicken β -globin insulator sequences, a 4.1 kb goat β -casein promoter, and a 5.5 kb β -casein 3' gDNA. Many milk expression vectors have been based on the pBC1 vector and/or have used some of its regulatory elements^[16].

Previously a pBC-hLZ vector based on the pBC1 for hLZ milk expression was constructed^[7]. This vector had mid-level expression in mice, but low-level expression in pigs^[17] and cows (unpublished data). The reasons for this are still being explored, but it may be due to species differences. Later, the pBC-hLZ vector was improved and had a relatively higher expression levels at concentrations about 0.12 mg·L⁻¹ in pigs^[18]. A new strategy involving the pBAC-hLF-hLZ-Neo BAC vector has been studied for increasing the expression level, the vector was modified from the human lactoferrin (hLF) BAC by recombineering, and the original hLF gDNA was replaced by hLZ gDNA, resulting in high expression in the milk of mice^[13], pigs^[19] and cow^[20]. The BAC vector solved the problem of low expression level among species, but there are several disadvantages that should be considered: (1) the length of vector is so long that it is difficult to be inserted into the

4 Discussion

Construction and *in vivo* evaluation of an efficient milk expression vector is one of the key steps in the

genome of cells or zygotes easily, especially for knock in mediated by CRISPR/Cas9^[21]; (2) the analyses of the locations and integration patterns of BACs are difficult^[14]; (3) the expression level is too high so that the lactation period is shortened in transgenic pigs (unpublished data). Therefore, it is necessary to construct a universal vector which has middle DNA length and middle expression level.

The *mWAP* gene locus is a good candidate for regulating foreign protein at different expression levels. It is reported that 2.6 kb of upstream DNA could direct the expression of human tissue plasminogen activator cDNA to the mammary gland in mice^[22], and regulate *mWAP* gDNA in the mammary gland in pig at concentration about 1.0 g·L⁻¹ milk^[23]. Recently, it was reported that 13 kb of upstream DNA could direct expression of human lactoferrin at between 16.7 to 29.8 g·L⁻¹^[24], human serum albumin at 11.9 g·L⁻¹^[25] and human lysozyme ranging from 18.4 to 35.0 g·L⁻¹^[26]. So, the *mWAP* gene locus could meet the requirements of middle DNA length and middle expression level by adjusting the length of its regulatory sequence.

In the present study, we constructed a universal vector, pMWAP, containing the 8.5 kb 5' promoter and 8.0 kb 3' gDNA of the *mWAP* gene and provided two *NotI* sites for easy foreign gene insertion. To check the expression efficiency of the pMWAP vector, we chose the 4.8 kb *hLZ* gene and obtained pMWAP-*hLZ*. Two of the six transgenic lines showed rhLZ expression, the highest expression was 0.45±0.05 g·L⁻¹, at a mid-level expression. The rhLZ had no ectopic expression and exhibited the same antibacterial activity as native hLZ. Compared to the BAC vector, the pMWAP vector has the advantages of being more convenient to use and allowing time saving in transgenic analyses.

Our final goal is to construct a universal vector with the expected expression level about 1.0 g·L⁻¹, but further work is needed to achieve this goal. The reasons that this was not achieved in the current study may be that the pBC-*hLZ* vector was not long enough to contain all essential regulatory elements^[8], or/and the foreign gene expression was affected by the chromosomal position^[27], or/and the transcription of hLZ was affected by the *NotI* sites near the ATG start codon and the TAA stop codon. Therefore, the length of *mWAP* gene locus needs to be readjusted and improved, and the method of gateway sites or Gibson assembly needs to be changed to avoid potential problems with *NotI* sites.

5 Conclusions

We have established a method to construct a universal milk expression vector, which regulated the expression of hLZ in milk.

Acknowledgements We would like to thank Dr. Hongxing Chen for

technical advice and his contribution of *mWAP* clone. This work was supported by High-Level Talent Project of Foshan University (GG040965), Science and Technology Project of Foshan City (GS54207) and Science Foundation of Jiangsu Province (BK20160206).

Compliance with ethics guidelines Shen Liu, Shengzhe Shang, Xuezheng Yang, Huihua Zhang, Dan Lu, and Ning Li declare that they have no conflicts of interest or financial conflicts to disclose.

All applicable institutional and national guidelines for the care and use of animals were followed.

References

- Bertolini L R, Meade H, Lazzarotto C R, Martins L T, Tavares K C, Bertolini M, Murray J D. The transgenic animal platform for biopharmaceutical production. *Transgenic Research*, 2016, **25**(3): 329–343
- Romagnolo D, DiAugustine R P. The mammary gland: protein factory of the future. *Environmental Health Perspectives*, 1994, **102** (8): 644–646
- Houdebine L M. Production of pharmaceutical proteins by transgenic animals. *Comparative Immunology, Microbiology and Infectious Diseases*, 2009, **32**(2): 107–121
- Babinet C. Transgenic mice: an irreplaceable tool for the study of mammalian development and biology. *Journal of the American Society of Nephrology*, 2000, **11**(S16): S88–S94
- Monzani P S, Adona P R, Ohashi O M, Meirelles F V, Wheeler M B. Transgenic bovine as bioreactors: challenges and perspectives. *Bioengineered*, 2016, **7**(3): 123–131
- Li G, Shi W, Chen G, Chen H, Jiao H, Yan H, Ji M, Sun H. Construction and *in vivo* evaluation of a mammary gland-specific expression vector for human lysozyme. *Plasmid*, 2014, **76**: 47–53
- Yu Z, Meng Q, Yu H, Fan B, Yu S, Fei J, Wang L, Dai Y, Li N. Expression and bioactivity of recombinant human lysozyme in the milk of transgenic mice. *Journal of Dairy Science*, 2006, **89**(8): 2911–2918
- Giraldo P, Montoliu L. Size matters: use of YACs, BACs and PACs in transgenic animals. *Transgenic Research*, 2001, **10**(2): 83–103
- Yang P, Wang J, Gong G, Sun X, Zhang R, Du Z, Liu Y, Li R, Ding F, Tang B, Dai Y, Li N. Cattle mammary bioreactor generated by a novel procedure of transgenic cloning for large-scale production of functional human lactoferrin. *PLoS One*, 2008, **3**(10): e3453
- Rival-Gervier S, Viglietta C, Maeder C, Attal J, Houdebine L M. Position-independent and tissue-specific expression of porcine whey acidic protein gene from a bacterial artificial chromosome in transgenic mice. *Molecular Reproduction and Development*, 2002, **63**(2): 161–167
- Zhang Y, Muylers J P, Testa G, Stewart A F. DNA cloning by homologous recombination in *Escherichia coli*. *Nature Biotechnology*, 2000, **18**(12): 1314–1317
- Sharan S K, Thomason L C, Kuznetsov S G, Court D L. Recombineering: a homologous recombination-based method of genetic engineering. *Nature Protocols*, 2009, **4**(2): 206–223
- Liu S, Li X, Lu D, Shang S, Wang M, Zheng M, Zhang R, Tang B, Li Q, Dai Y, Li N. High-level expression of bioactive recombinant human lysozyme in the milk of transgenic mice using a modified human lactoferrin BAC. *Transgenic Research*, 2012, **21**(2): 407–

414

14. Zhang R, Guo C, Sui S, Yu T, Wang J, Li N. Comprehensive assessment of milk composition in transgenic cloned cattle. *PLoS One*, 2012, **7**(11): e49697
15. Hogan B, Constantini F, Lacy E. Manipulating the mouse embryo. A laboratory manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor*, 1986
16. Yu H, Chen J, Liu S, Zhang A, Xu X, Wang X, Lu P, Cheng G. Large-scale production of functional human lysozyme in transgenic cloned goats. *Journal of Biotechnology*, 2013, **168**(4): 676–683
17. Tong J, Wei H, Liu X, Hu W, Bi M, Wang Y, Li Q, Li N. Production of recombinant human lysozyme in the milk of transgenic pigs. *Transgenic Research*, 2011, **20**(2): 417–419
18. Lu D, Li Q, Wu Z, Shang S, Liu S, Wen X, Li Z, Wu F, Li N. High-level recombinant human lysozyme expressed in milk of transgenic pigs can inhibit the growth of *Escherichia coli* in the duodenum and influence intestinal morphology of sucking pigs. *PLoS One*, 2014, **9**(2): e89130
19. Lu D, Liu S, Shang S, Wu F, Wen X, Li Z, Li Y, Hu X, Zhao Y, Li Q, Li N. Production of transgenic-cloned pigs expressing large quantities of recombinant human lysozyme in milk. *PLoS One*, 2015, **10**(5): e0123551
20. Lu D, Liu S, Ding F, Wang H, Li J, Li L, Dai Y, Li N. Large-scale production of functional human lysozyme from marker-free transgenic cloned cows. *Scientific Reports*, 2016, **6**(1): 22947
21. Jung C J, Ménoret S, Brusselle L, Tesson L, Usal C, Chenouard V, Remy S, Ouisse L H, Poirier N, Vanhove B, de Jong P J, Anegon I. Comparative analysis of piggyBac, CRISPR/Cas9 and TALEN mediated BAC transgenesis in the zygote for the generation of humanized SIRPA rats. *Scientific Reports*, 2016, **6**(1): 31455
22. Pittius C W, Hennighausen L, Lee E, Westphal H, Nicols E, Vitale J, Gordon K. A milk protein gene promoter directs the expression of human tissue plasminogen activator cDNA to the mammary gland in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America*, 1988, **85**(16): 5874–5878
23. Wall R J, Pursel V G, Shamay A, McKnight R A, Pittius C W, Hennighausen L. High-level synthesis of a heterologous milk protein in the mammary glands of transgenic swine. *Proceedings of the National Academy of Sciences of the United States of America*, 1991, **88**(5): 1696–1700
24. Shi G, Chen H, Wu X, Zhou Y, Liu Z, Zheng T, Huang P. A *mWAP-hLF* hybrid gene locus gave extremely high level expression of human lactoferrin in the milk of transgenic mice. *Transgenic Research*, 2009, **18**(4): 573–582
25. Wu X, Lin Y, Xiong F, Zhou Y, Yu F, Deng J, Huang P, Chen H. The extremely high level expression of human serum albumin in the milk of transgenic mice. *Transgenic Research*, 2012, **21**(6): 1359–1366
26. Wu X, Lin Y, Xi Y, Shao Z, Zhou Y, Liu F, Chen H. The development of transgenic mice for the expression of large amounts of human lysozyme in milk. *Biotechnology Letters*, 2014, **36**(6): 1197–1202
27. Clark A J, Bissinger P, Bullock D W, Damak S, Wallace R, Whitelaw C B, Yull F. Chromosomal position effects and the modulation of transgene expression. *Reproduction, Fertility, and Development*, 1994, **6**(5): 589–598