

A modified sleeping beauty transposon vector system integrates and settles genes in the genome

Kazuyoshi Kaminaka^{1*} and Chikateru Nozaki²

¹Research and Development Division, Kikuchi Research Center, KM Biologics Co., Ltd., 1314-1 Kyokushi Kawabe, Kikuchi-city, Kumamoto 869-1298, Japan

²Department of Medical Technology, Kumamoto Health Science University, 325 Izumi-machi, Kita-ku, Kumamoto-city, Kumamoto 861-5598, Japan

Abstract

Sleeping Beauty (SB) exhibits efficient cut-and-paste transposition in a wide range of mammalian cells. Using this characteristic, a new vector system (the SB system) was previously developed for integration of target genes into the host genome and is now used as a research tool and for gene therapy. However, there are concerns about the retransposition and the size restriction to be integrated into this system. To overcome these issues, we constructed a modified transposon vector system with lox sequences and estimated its transposable activity and Cre/lox recombination ability in HeLa cells. We also estimated the possibility of the retransposition by plasmids with defective inverted repeat/direct repeat structures presumed to be taken after Cre/lox recombination. We demonstrated that our modified transposon vector system could efficiently cause recombination mediated by the Cre/lox system following the transposition of the target genes into the genome and eliminate the possibility of retransposition of the genes. We believe our modified transposon vector system could be used not only as a research tool but also as a safer and more versatile gene integration system in the medical field.

Introduction

Gene integration technology has become essential for the production of recombinant proteins and organisms, as well as gene therapy application [1-3]. The most critical point of this technology is to integrate target genes into the host genome safely while yielding high integration frequency.

Sleeping Beauty (SB) is a transposon belonging to the Tc1/mariner superfamily, which was reconstructed from transposon fossils found in fish genomes and exhibited efficient cut-and-paste transposition in mammalian cells [4]. These peculiar characteristics were exploited by researchers to develop a new vector system (the SB system) for the integration of target genes into the host genome. The system encompasses two plasmids: one expresses a transposase gene, which encodes an enzyme to carry the transposable activity, and the other is a transposon vector (the SB vector) having inverted repeat (IR) sequences with two direct repeats (DR) sequences (termed IR/DR) at both ends. The gene of interest is placed between the IR/DRs. This system has a broad spectrum and high integration efficiency in mammalian cells [5]. In addition, the possibility of retransposition of the target genes can be reduced by segregating the transposase expression cassette from the transposon.

Although the SB system is useful for the integration of target genes, several issues exist. First, there is a possibility that the target genes integrated by the SB system can be retransposed to another site in the genome. Because the transposon vector utilizes the transposable activity on the genome that triggered its discovery, the target genes integrated by the transposon vector must be transposition competent on the genome. In the SB system, the possibility of retransposition is minimized using a two-component plasmid system. However, the target genes between both IR/DRs in a transposon vector (transposable substrate DNA) can readily retranspose to another site when a transposase is supplied [6,7]. Additionally, a few reports have demonstrated that the occasional, unintended retransposition occurred in fish genomes [8,9].

Second, the target gene size integrated by the SB system is limited. The transposition efficacy of the SB system decreases as the size of the target gene increases. The efficiency decreased to 10% of the original transposon size when the transposon exceeds 6 kbp in size [5]. In general, the addition of a promoter and a polyA addition signal sequences are required for gene expression within the host, further limiting the size of the target gene to be loaded in the SB vector. Moreover, to select hosts with integrated target genes effectively based on drug resistance, the expression cassette of a drug-resistance gene is also introduced into the vector, additionally shortening the size of the target genes to be introduced. Thus, genes with a large size [10,11] or those requiring the expression of multiple components [12] would not be suitable for integration using this system.

In this study, to overcome these issues with the SB system, we constructed a modified transposon vector system with lox sequences that prevented retransposition and enabled insertion of genes without size limitations by Cre/lox recombination.

Results

Transposition activity of the modified transposon vector

A modified transposon vector containing the lox sequence may allow us to overcome issues in the SB system by using Cre/lox recombination. Izsvák et al. (2002) [13] showed that disrupting the

**Correspondence to:* Kazuyoshi Kaminaka, Development Department, Research and Development Division, Kikuchi Research Center, KM Biologics Co., Ltd., 1314-1 Kyokushi Kawabe, Kikuchi-city, Kumamoto 869-1298, Japan, Tel: +81-968-37-4064; Fax: +81-968-37-3108; E-mail: kaminaka@kmbiologics.com

Key words: sleeping beauty, transposon, retransposition, cre/lox recombination, inverted repeat/direct repeat structures, gene integration

Received: July 15, 2019; **Accepted:** July 22, 2019; **Published:** July 26, 2019

IR/DR structure, e.g., by deletion or mutation, leads to a significant reduction in transposition activity. As shown in Figure 1, in the modified transposon vector with lox sequences between two DR sequences in the IR/DRs, the IR/DR structure was disrupted, and its transposable activity seemed to disappear by Cre/lox recombination. Based on this idea, we constructed modified transposon vectors with lox sequence(s) between two DR sequences present in the IR/DRs (Figure 2).

Our modified vectors, IR/DR-puro/5'Lxp, IR/DR-puro/LxpADb and a positive control vector IR/DR-puro (SB vector) were constructed from a basal plasmid IR/DR-NTA-Ad, which had IR/DRs at both ends (Figure 2A). IR/DR-puro (SB vector) had the same IR/DRs elements previously reported by Ivics, Hackett, Plasterk, and Izsák (1997) [4]. The modified transposon vector IR/DR-puro/5'Lxp had the lox71 sequence, which is a left element mutant of loxP [14], between two DR sequences in the 5'-IR/DR. Another modified transposon vector, IR/

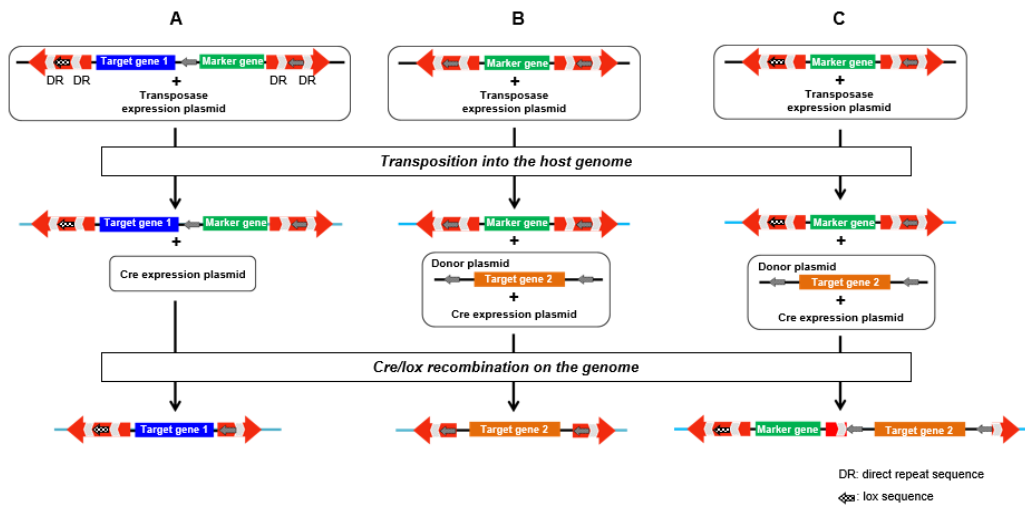
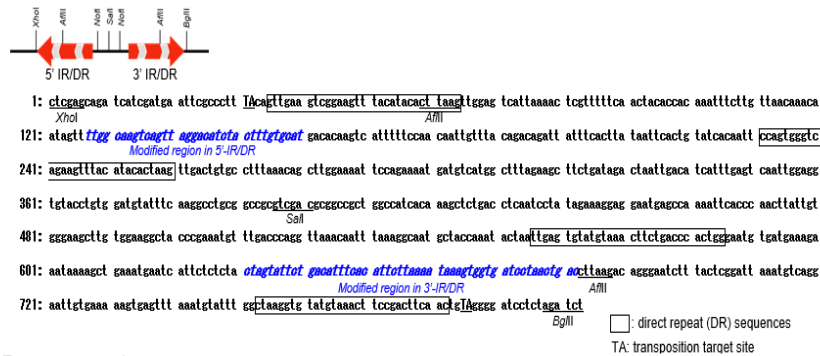


Figure 1. Modified transposon vector system. In the modified transposon vector system, the modified transposon vector with lox sequence(s) in the IR/DR(s) is used instead of the SB vector. The target gene and/or a marker gene are first introduced into the genome of the host by the transposition, similar to the SB system. Excision of the marker gene (A), exchange with another gene (B), and insertion of another gene (C) can be performed by Cre/lox recombination after integration into the genome by the transposition

A: IR/DR-NTA-Ad (Basal transposon plasmid)



B: Transposon vectors

Vector	Sequence of modified region	
	5'-IR/DR	3'-IR/DR
IR/DR-puro (SB vector)	Same as IR/DR-NTA-Ad	Same as IR/DR-NTA-Ad
IR/DR-puro/5'Lxp (modified transposon vector)	5'-taoogttagt atagataaa ttataoang ttat (lox71)	Same as IR/DR-NTA-Ad
IR/DR-puro/LxpADb (modified transposon vector)	5'-taoogttagt atagataaa ttataoang ttat (lox71)	5'-atanottagt atagataaa ttataoang ttataaaa atgagaaat (loxP + polyA addition signal)

Figure 2. Constructed transposon vectors. (A) A basal plasmid IR/DR-NTA-Ad for the transposon vector construction. The plasmid was repaired from a transposon in the salmon genome and constructed, as shown in the experimental procedures. The nucleotide sequences of the transposable element in the plasmid were also shown. (B) Transposon vectors investigated in this study. SB vector (IR/DR-puro) was constructed as a positive control. The modified transposon vector IR/DR-puro/5'Lxp had a mutant lox sequence (lox71) between two DR sequences in the 5'-IR/DR. Another vector, IR/DR-puro/LxpADb, had the same 5'-IR/DR with IR/DR-puro/5'Lxp and the loxP and polyA addition signal sequences between two DR sequences in the 3'-IR/DR. All three vectors had Puro between both IR/DRs as a marker gene

DR-puro/LxpADb, had the lox71 sequence in the 5'-IR/DR and the loxP sequence flanked with the polyA addition signal sequence derived from bovine growth hormone in the 3'-IR/DR. On IR/DR-puro/LxpADb, we specifically chose the lox71 and loxP combination to generate the three expected types of Cre/lox recombination shown in Figure 1.

After cotransfection of IR/DR-puro (positive control SB vector) with the transposase expression plasmid pCAGGS/SB into HeLa cells, puromycin-resistant HeLa cell clones (HeLa/puro) were obtained 24.5 times more frequently than those obtained from transfection with the vector alone (C_{SB}/C_N index of 24.5). In contrast, the modified transposon vectors IR/DR-puro/5'Lxp and IR/DR-puro/LxpADb had C_{SB}/C_N indexes of 49.3 and 19.9, respectively (Table 1).

This result indicated that our modified transposon vectors retained transposable activity like as the SB vector, even though the lox sequences and polyA addition signal sequence were inserted between two DR sequences.

Gene replacement by the Cre/lox system after transposition of the transposon vector

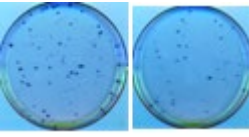
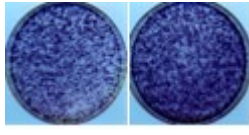
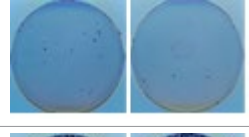

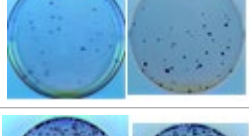

In order to confirm that our modified transposon vectors also retained the recombination ability mediated by the Cre/lox system after transposition into the genome, the donor plasmid pLx/GFP/neo/pA(-) was introduced with pCAGGS/Cre into HeLa/puro, which was obtained from the transposition of IR/DR-puro/LxpADb, to replace the puromycin-resistance gene expression cassette (*Puro*) in the genome with the neomycin-resistance gene and the green fluorescent gene

expression cassettes (*Neo* and *GFP*). The genomic DNAs from three HeLa cell clones sensitive to puromycin, resistant to neomycin, and emitting green fluorescence (HeLa/neo) were analyzed by Southern blotting (Figure 3).

If HeLa/neo acquired *Neo* and *GFP* by transposition of IR/DR-puro/LxpADb followed by Cre/lox recombination with pLx/GFP/neo/pA(-), a 4.3-kbp fragment would be detected from the genomic DNA digested with *Afl*II by a *Neo*-derived probe, and no DNA fragment would be detected using a *Puro*-derived probe (Figure 3A). Instead, a 2.1-kbp fragment would be detected using a *Puro*-derived probe when the random integration of pLx/GFP/neo/pA(-) occurred (Figure 3A). As shown in Figure 3B, clone nos. 1 and 2 were considered to be obtained by the transposition of IR/DR-puro/LxpADb followed by Cre/lox recombination with pLx/GFP/neo/pA(-) since a 4.3-kbp fragment was detected using the *Neo*-derived probe. In clone no. 3, a fragment over 4.3 kbp was detected by the *Neo*-derived probe. This clone was considered not to be obtained by gene replacement but as a result of insertion of pLx/GFP/neo/pA(-) at the lox71 site and subsequent excision of DNA between loxP sequences in the 3'-IR/DR of IR/DR-puro/LxpADb and at the 3'-end of pLx/GFP/neo/pA(-) by Cre/lox recombination.

In order to further confirm that HeLa/neo was obtained by transposition and subsequent gene replacement, we also analyzed both terminal regions of the inserted gene in clone no. 1 using a genome walking method. As shown in Figure 4A, the lox71 sequence in the 5' region was exchanged for the double mutant lox sequence (mutant

Table 1. Transposition activity of transposon vectors

Construct	Lox sequences		pCAGGS/SB (Transposase)	Colonies	C_{SB}/C_N
	5'-IR/DR	3'-IR/DR			
IR/DR-puro (SB vector)	0	0	No (C_N)		24.5
			Yes (C_{SB})		
IR/DR-puro/5'Lxp	1	0	No (C_N)		49.3
			Yes (C_{SB})		
IR/DR-puro/LxpADb	1	1	No (C_N)		19.9
			Yes (C_{SB})		

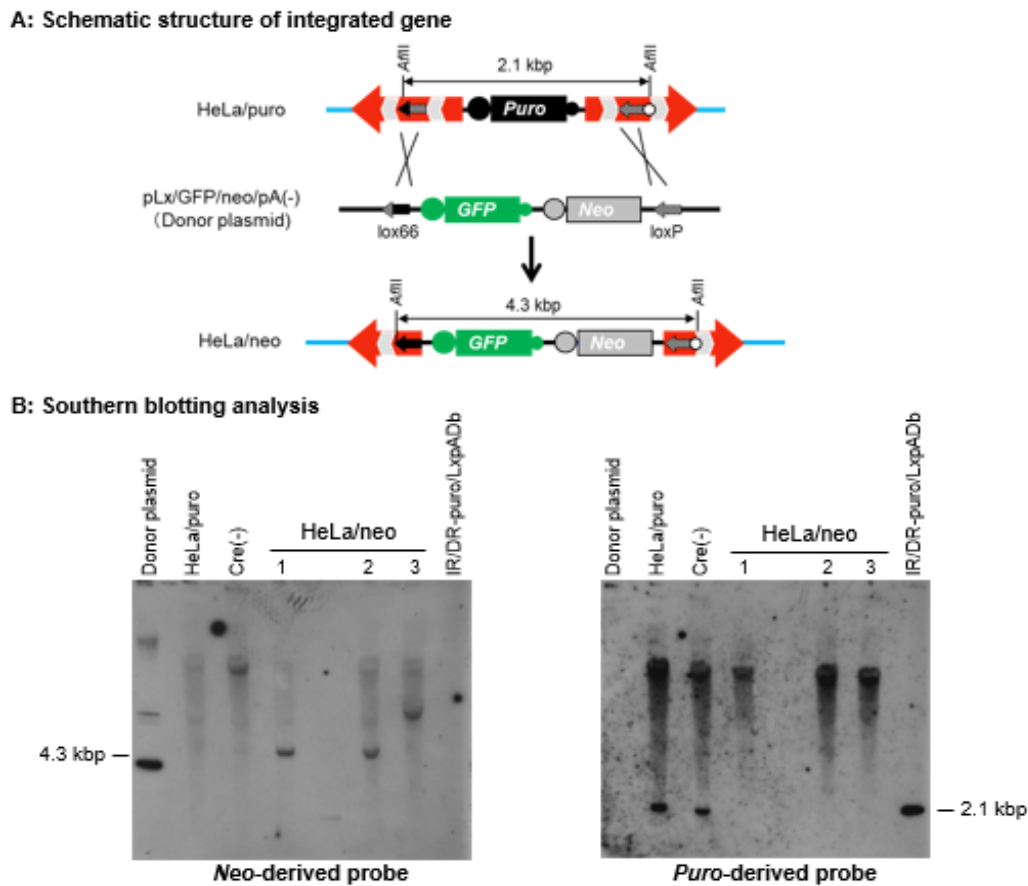


Figure 3. Gene exchange in HeLa cells. (A) Schematic structure of integrated genes in the genome of HeLa/puro and HeLa/neo cells and the donor plasmid pLx/GFP/neo/pA(-). (B) Southern blotting analysis of *AflII*-digested genomic DNAs of HeLa/neo clones. HeLa/puro: *AflII*-digested genomic DNA of the clone before Cre/lox-mediated gene exchange was performed; Cre(-): *AflII*-digested genomic DNA of the clone obtained by transfection of pLx/GFP/neo/pA(-) without pCAGGS/Cre; IR/DR-puro/LxpADb: the modified transposon vector introduced to obtain HeLa/puro cells

sequences of lox71 and lox66), and the loxP sequence in the 3' region remained. Furthermore, due to gene replacement, the inner two DR sequences in the four DRs of the transposon vector were removed, and the region between the 5' and 3' lox sequences was the same as the sequence of the donor plasmid pLx/GFP/neo/pA(-). The region outside of the lox sequences was the same as the sequence of IR/DR-puro/LxpADb (Figure 4A). Besides, the region outside of both ends of IR/DR-puro/LxpADb (TA dinucleotide) matched with the sequence of human chromosome 3 by a search in the GeneBank database (Figure 4B).

From these results, we concluded that our modified transposon vector was able to efficiently insert the target genes into the genome based on transposable activity and to cause the subsequent recombination mediated by Cre/lox recombination.

Transposable activity after Cre/lox recombination of the integrated transposon vector

As described above, the modified transposon vector was capable of recombination by the Cre/lox system after transposition into the genome. Next, in order to confirm that the gene integrated into the genome could not retranspose after recombination by the Cre/lox system (Figure 1), we constructed two types of plasmids with possible terminal structures mimicking those after the Cre/lox recombination (Figure 5) and compared their transposable activities with IR/DR-puro/

LxpADb. One (IR/DR-3'IR/puro) was the deleted construct of the inner DR sequence in the 3'-IR/DR, which would be formed after the excision between loxPs existing inside and outside of the IR/DR by Cre/lox recombination (Figure 1A). Another (5'+3'IR/puro) was the deleted construct of the two inner DR sequences in the 5'- and 3'-IR/DR, which would be formed after gene replacement at lox sequences in the 5'- and 3'-IR/DR (Figure 1B).

IR/DR-3'IR/puro, in which only an inner DR sequence in the 3'-IR/DR had been removed, had no transposable activity, similar to 5'+3'IR/puro, in which both of the inner DR sequences had been removed (Table 2). From this result, we concluded that the target genes could be integrated into the genome with no retransposition by utilizing the Cre/lox system following the transposition of our modified vectors.

Discussion

By inserting a lox sequence between two DR sequences present in the IR/DR of the SB vector, we believed that we could excise the inner DR in the IR/DR and eliminate transposable activity via Cre/lox recombination. However, in a previous study, it was found that mutating several sequences or changing the distance between the two DR sequences in the IR/DR could cause a significant reduction in transposition efficiency [13]. Therefore, to overcome these contradictory problems, we decided to not change the distance (number of bases) between the two DR sequences in the IR/DR and to avoid

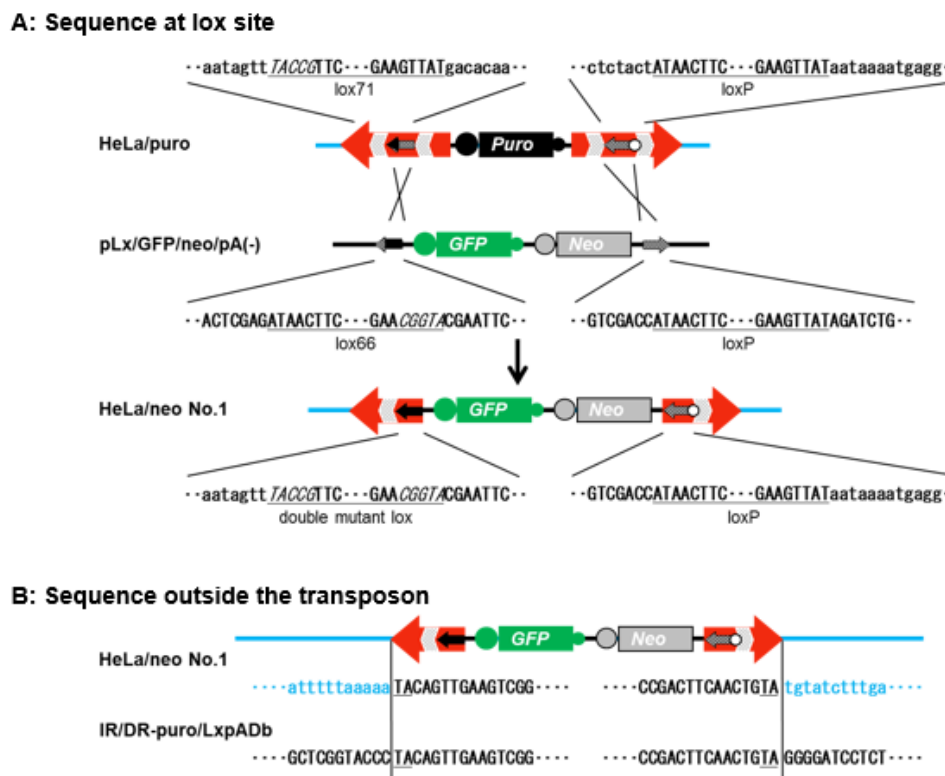


Figure 4. Sequences at lox sites and outside the transposon. (A) Sequence at the lox site of the genome or plasmids at each step of the gene exchange. Underlining shows lox sequences, and italic letters indicate the mutated site of the left element in lox71 and the right element in lox66. (B) The sequence outside of the transposon. The sequences of both ends and outside of the modified transposon vector IR/DR-puro/LxpADb are indicated by black letters, and the sequences outside of the integrated gene in HeLa/neo cells are indicated by light blue lower-case letters. The underlined dinucleotide TA indicates the transposition target sites

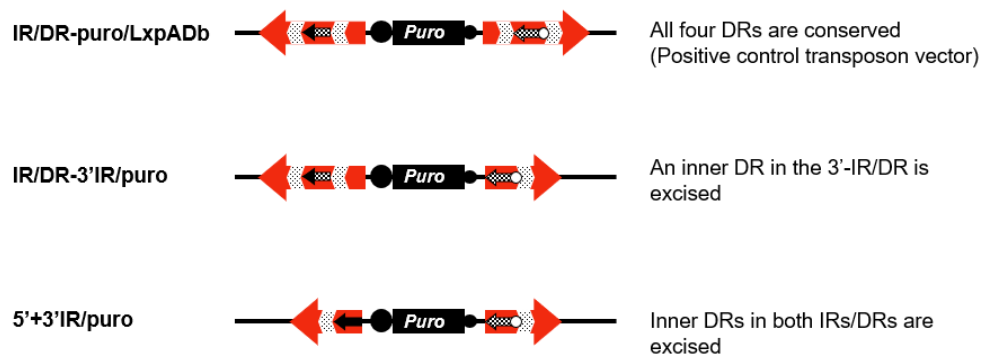


Figure 5. Plasmids with defective IRs/DRs. IR/DR-puro/LxpADb was the modified transposon vector with the native IR/DR and was used as a positive control. IR/DR-3'IR/puro is a plasmid with a defective 3'-IR/DR, in which the inner DR was excised, and with the native 5'-IR/DR. 5'+3'IR/puro is a plasmid with defective IR/DRs, in which the inner DRs were excised

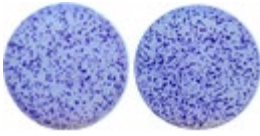

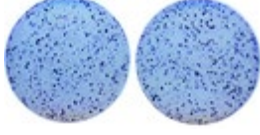
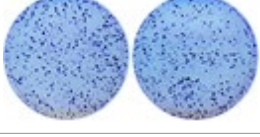
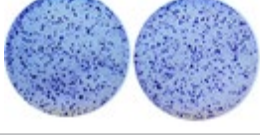
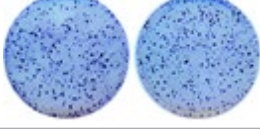
the hotspot (3'-half sequences of DR) when inserting the lox sequences between two DR sequences in the IR/DR of the SB vector. As a result, our modified transposon vectors were able to efficiently introduce a gene into the genome via a plasmid vector, similar to the SB system. In addition, plasmids having a defective IR/DR structure predicted by Cre/lox recombination after integration into the genome with our modified transposon vector could not transpose, even when a transposase was supplied. Therefore, we demonstrated that our modified transposon vectors could efficiently introduce the target genes into the genome and settle these genes.

In addition to gene replacement and excision, our modified transposon vector system could be utilized for Cre/lox recombination, a gene insertion reaction, to eliminate the possibility of retransposition.

When this reaction occurred at the lox site present between two DR sequences in the IR/DR, the distance between the two DRs changed, resulting in structural changes in the IR/DR. This change could be easily applied to lose transposable activity, as previously demonstrated [13]. A transposon vector with additional loxP sequences between the two DRs in the 5'-IR/DR had no transposable activity (data not shown).

In this study, the efficiency of Cre/lox recombination after the introduction of our modified transposon vector into the genome was not clarified. However, because we could easily select gene-replaced clones by the expression of marker genes (*Neo* and *GFP*), we assumed this could not affect Cre/lox recombination, even if a lox sequence exists within the IR/DR. Furthermore, we chose the combination of lox71 and loxP combination in IR/DR-puro/LxpADb as we expected

Table 2. Transposition activity of plasmids with defective IR/DR

Construct	pCAGGS/SB (Transposase)	Colonies	C_{SB}/C_N
IR/DR-puro/LxpADB (All four DRs conserved)	No (C_N)		22.9
	Yes (C_{SB})		
IR/DR-3'IR/puro (No inner DR in 3'-IR/DR)	No (C_N)		0.8 (<1)
	Yes (C_{SB})		
5'+3'IR/puro (No inner DRs in both IR/DRs)	No (C_N)		0.5 (<1)
	Yes (C_{SB})		

the three types of Cre/lox recombination that are shown in Figure 1. When one expects more effective gene replacement after transposition, the combination of lox2272 and lox5171 can be used instead of that of lox71 and loxP [15]. However, the lox sequence inserted in the modified transposon vector could be modified accordingly to the expected Cre/lox recombination.

It is challenging to introduce genes larger than 6 kbp efficiently using the SB system [5]. The replaced gene after transposition of our modified transposon vector system was about 4.3 kbp in size. Therefore, we did not show that our vector system could introduce genes over 6 kbp in size in this study. However, gene replacement by the Cre/lox system can substitute genes of over 6 kbp [16] efficiently, and DNA up to 230 kbp can be inserted into the genome in plants using this system [17]. Therefore, we expect that genes can be efficiently introduced into the genome using our modified transposon vector system and the Cre/lox system, without size limitations.

To overcome the size restriction to be transposed in the SB system, a next-generation transposon system was previously developed [18]. This system combined a hyperactive SB transposase with a transposon vector having two complete pairs of IR/DRs in inverted orientations (sandwich configuration). Consequently, their new transposon system could load genes over 10 kbp while maintaining relatively efficient transposition. Despite this improvement in the SB system, our modification of the SB vector shown in this study is useful because it could also be applied to their new transposon system and could eliminate retransposition after transposition into the genome.

The SB system containing our modified transposon vector system utilized transposons belonging to the Tc1/mariner superfamily. Transposons in this superfamily are less likely to be inserted into and disrupt functional genes compared with other transposons or viral vectors [19-21]. However, once a transposase is supplied, although it may not have a high incidence rate, transposition would occur at a probability of 3.5×10^{-5} events per cell generation in the genome [6]. Furthermore, even if the probability is low, transpositions may be caused by invasion from other active transposons [22]. These possibilities may limit the application of the SB system in gene therapy owing to increased risk of oncogenesis [23,24].

To safely carry out gene integration into the genome with our modified transposon vector, multiple insertions of modified transposon vectors must also be avoided because the region between those vectors might be deleted by Cre/lox recombination events, especially when multiple vectors are inserted into the same chromosome. Under our experimental condition, we predicted that our modified transposon vectors would be inserted into multiple genomic loci. According to a previous report [25], a single insertion event could be ensured at the "low" transposon dosage condition in the SB system. Therefore, by optimizing the "low" transposon dosage condition for our modified transposon vector, we consider that the insertion of a single copy of the transposon into the genome could be achieved.

In summary, our modified transposon vector system eliminated concerns regarding potential retransposition after the efficient introduction of genes into the genome. By studying optimal conditions

for a single copy insertion and *in vivo* application, we believe that our modified transposon vector system could be used not only as a research tool but also as a safer and more versatile gene integration system in the medical field.

Experimental procedures

Restriction enzymes, genes, and primers

Unless otherwise stated, restriction enzymes were purchased from Takara Bio (Shiga, Japan). The neomycin-resistance gene in pMC1neo (Agilent, CA, USA) and the green fluorescent gene in pEGFP-N1 (BD Biosciences, NJ, USA) were used for the construction. *Puro* in pPGKpuro kindly provided by Dr. Watanabe (NARO) was also used. Furthermore, the Cre recombinase expression plasmid (pCAGGS/Cre) and the lox sequences loxP, lox66, and lox71 were kindly provided by Dr. Araki (Kumamoto University).

Primers used for the construction of vectors or plasmids were synthesized at Eurofins Genomics (Tokyo, Japan) and are listed in Table 3.

Construction of plasmids

SB system

The two components of the SB system, a transposase expression plasmid that carries transposable activity and an SB vector that becomes the substrate DNA of the transposition, were constructed as follows.

Polymerase chain reaction (PCR) was performed with primers TCE/In and TCE/St based on the sequences of GenBank accession nos. L12206 and L12209, respectively, and with salmon sperm DNA (Nippon Gene, Tokyo, Japan) as a template. An amplified DNA fragment of about 1 kbp, which encoded the transposase gene, was repaired to the same amino acid sequence as the transposase in the SB system by site-directed mutagenesis according to the method reported by Ivics, Hackett, Plasterk, and Izsák (1997) [4]. Subsequently, the repaired transposase gene was inserted downstream of the chicken β -actin promoter of the expression plasmid pCAGGS [26], resulting in the transposase expression plasmid pCAGGS/SB.

In order to construct an SB vector as a substrate for the transposase, first, a transposon of approximately 1.6 kbp with IR/DRs at the 5'- and 3'-ends was amplified by PCR with a primer Tc1 [27] and salmon sperm DNA (Nippon Gene) as a template. Next, the sequences of the isolated IR/DRs were repaired to those of GenBank accession no. L48685 [4], using their site-directed mutagenesis method. Then, unique restriction

enzyme recognition sites (*NotI* and *SalI*) were introduced between the IR/DRs of the repaired transposon (IR/DR-NTA-Ad, Figure 2A).

Transposon vectors

The transposon vectors we investigated in this study were constructed as below.

The SB vector IR/DR-puro was constructed as positive control by inserting *Puro* into the *SalI* site of IR/DR-NTA-Ad, which had the same IR/DRs as the SB vector (Figure 2B).

The modified transposon vector IR/DR-puro/5'Lxp was constructed as follows. First, PCR was performed using IR/DR-NTA-Ad as a template and primers SP6 and Ps/Lx71R to amplify a DNA fragment of about 200 bp in which the lox71 sequence was inserted between two DRs in the 5'-IR/DR. After digestion of the fragment with *PshAI* and *XhoI*, the fragment was swapped with a fragment of IR/DR-NTA-Ad digested with the same enzymes, yielding the transposon vector IR/DR-Ad/5'Lxp having the lox71 sequence. Subsequently, *Puro* was inserted into the *SalI* site of this vector to construct IR/DR-puro/5'Lxp (Figure 2B).

Another modified transposon vector IR/DR-puro/LxpADb was constructed as follows. First, the 3'-IR/DR fragment from IR/DR-NTA-Ad digested with *SalI* and *BglII* was subcloned into pSP72 (Promega, WI, USA), resulting in 3'IR/DR-Ad. Next, a DNA fragment amplified by PCR using this 3'IR/DR-Ad as a template and primers SP6 and Af/LxpAR was digested with *AflIII* and *SalI* and swapped with the fragment of 3'IR/DR-Ad digested with the same restriction enzymes. This resulting plasmid 3'IR/DR-LxpA had a loxP sequence flanked with a polyA addition signal sequence between two DRs in the 3'-IR/DR. Subsequently, the 3'IR/DR-LxpA was digested with *BglII* and *SalI* and swapped with the DNA fragment of IR/DR-Ad/5'Lxp digested with the same restriction enzymes. The resulting plasmid IR/DR-Ad/LxpADb had a lox71 sequence in the 5'-IR/DR and a loxP flanked with a polyA addition signal sequence in the 3'-IR/DR. *Puro* was inserted into the *SalI* site of this plasmid to construct IR/DR-puro/LxpADb (Figure 2B).

Plasmid for gene replacement by the Cre/lox system

In the Cre/lox system, gene excision, insertion, and replacement reactions can occur. In order to confirm the Cre/lox recombination reaction after transposition of the modified transposon vectors into the genome, we constructed a plasmid for gene replacement (a donor plasmid) having a mutant lox sequence and *Neo* without a polyA addition signal sequence to efficiently induce the replacement reaction [28].

Table 3. Primers used for vector construction

Primer	Sequence	Sequence origin
TCE/In	atgggaaaaat caaaagaaat cagccaaga	GenBank Accession. No. L12206
TCE/St	ttagtatttg gtagcattgc ctttaaat	GenBank Accession. No. L12209
Tc1	tacagttgaa gtgtaagttt a	A.D. Radice et al., 1994
SP6	atttaggtga cactatagaa c	SP6 promoter
Ps/Lx71R	atgactgtg tcataactc gtataatgta tgctatagca acggtaaact attgtttgtt aacaagaat	PshA I site + lox71
Af/LxpAR	ccgtctctaa gatttcctca ttttattata acctcgata atgtatgcta tacgaagtta tagtagagag aatgattcat ttcagc	<i>AflIII</i> site + bovine growth hormone-derived poly A addition signal + loxP
Lx66/Lxp-F	tccgctcgag ataactcgt atagcataca ttatagcaac ggtacgaatt caggcctgtc gacggatcct	<i>XhoI</i> site + lox66 + <i>EcoRI</i> and <i>StuI</i> sites
Lx66/Lxp-R	tccagatcta taacttgta taatgatgc tatacgaagt tatggccagg atccctcgac aggcctgaat tc	<i>BglII</i> site + loxP + <i>MscI</i> , <i>StuI</i> , and <i>EcoRI</i> sites
AP1	gtaatacgac tcactatagg gc	Adaptor in BD GenomeWalker Universal Kit
AP2	actatagggc acgcgtggt	Adaptor in BD GenomeWalker Universal Kit
CAG/GSP2	gcatatgata cacttgatgt actgcca	Promoter in pCAGGS
CAG/GSP4	gcgggccatt taccgtaagt tatgta	Promoter in pCAGGS
neo/1306F	gcgatgctg cttgccgaat atcat	Neomycin resistance gene in pMC1neo
neo/1389F	cgctatcagg acatagcgtt ggcta	Neomycin resistance gene in pMC1neo

After mixing Lx66/Lxp-F and Lx66/Lxp-R with a complementary sequence of 25 bases at the 3'-ends, a DNA fragment consisting of lox66 and loxP sequences was amplified by PCR with no template DNA and primers and then cloned into the *XhoI* and *BglIII* sites of pSP72 (Lx66/Lxp/pSP). Next, after *EcoRI*- and *MscI*-mediated digestion (New England Biolabs, MA, USA) of the plasmid, *GFP* with the promoter derived from pCAGGS was inserted at these sites between the lox66 and loxP sequences. Furthermore, downstream of *GFP*, *Neo* derived from pMCneo was inserted as a marker gene, resulting in a donor plasmid pLx/GFP/neo/pA (-) with lox66 and loxP sequences at both ends (Figure 3A).

Estimation of transposable activity in HeLa cells

The transposable activities of the constructed transposon vectors (Figure 2) were evaluated according to the method reported by Izsvák et al. (2002) [13] in human cervical cancer-derived HeLa cells (Sumitomo Dainippon Pharma, Tokyo, Japan). Briefly, HeLa cells seeded into 6-well plates (2×10^5 cells/2 mL/well) the day before were transfected with the complex of Trans-IT LT1 (Takara Bio) and the DNA mixture of pCAGGS/SB and one of the constructed transposon vectors (1.5 µg each), according to the manufacturer's protocol. Simultaneously, cells transfected with the complex with each of the constructed transposon vectors alone were prepared. Two days after the transfection, the cells were passaged into 10-cm dishes and cultured in Dulbecco's modified Eagle medium (Sigma-Aldrich, MO, USA) supplemented with 10% fetal bovine serum (Hyclone, UT, USA) and 1 µg/mL puromycin for approximately 2 weeks. Cells were then stained with 20% methanol (Wako Pure Chemical, Osaka, Japan) containing 0.2% crystal violet (Kishida Chemical, Osaka, Japan). All procedures were performed duplicate for each transposon vector, and the transposable activity of each vector was estimated by the C_{SB}/C_N ratio, where C_{SB} indicates the mean colony number obtained by cotransfection of pCAGGS/SB and each transposon vector, and C_N indicates the mean number obtained by the transfection of each transposon vector alone.

Gene replacement by the Cre/lox system

Puro in HeLa/puro obtained by the transposition of IR/DR-puro/LxpADb was replaced as follows. HeLa/puro cells seeded into 6-well plates (2×10^5 cells/well) the day before were transfected with the complex of Trans-IT LT1 and the DNA mixture of pCAGGS/Cre and pLx/GFP/neo/pA(-) (1.5 µg each) and then cultured in Dulbecco's modified Eagle medium supplemented 10% fetal bovine serum. After 2 days, the cells were passaged into 10-cm dishes and cultured for approximately 2 weeks in the presence of 750 µg/mL G418 (Takara Bio). Subsequently, clones emitting green fluorescence and showing puromycin sensitivity (HeLa/neo cells) were selected.

Genome analysis in HeLa/puro and HeLa/neo cells

Genomic DNAs in HeLa/puro and HeLa/neo cells ($5-10 \times 10^6$ cells each) were prepared according to a previously described method [29], and their concentrations were determined by measuring the absorbance at 260 nm.

Southern blotting analysis was performed as follows. Briefly, genomic DNAs (20 µg each) and plasmids (IR/DR-puro/LxpADb and pLx/GFP/neo/pA (-), 2 ng each) were digested with *AflIII* and separated by electrophoresis on 0.7% agarose gels (Bio-Rad, CA, USA). After separation, DNAs were transferred to nylon filters (Hybond-N+; GE Healthcare Japan, Tokyo, Japan) by alkaline transfer. The DNA fragment with *Puro* or *Neo* was detected using probe DNA prepared from *Puro* or *Neo* and labelled with [$\alpha^{32}P$] dCTP (PerkinElmer, MA, USA).

In order to confirm that the integrated genes in HeLa/neo cells were transposed and subsequently replaced from *Puro*, both terminal regions of the integrated genes were analyzed by the genome walking method. Briefly, the genomic DNA from HeLa/neo cells was digested independently with restriction enzymes *EcoRV*, *PvuII*, *SspI*, and *NaeI*, and an adaptor attached to a BD GenomeWalker Universal Kit (Takara Bio) was ligated. The resulting libraries were amplified by nested PCR with the primer combination of AP1 and CAG/GSP2 followed by AP2 and CAG/GSP4 for analysis of the 5' terminal region. Nested PCR was also performed with the primer combination of AP1 and neo/1306F followed by AP2 and neo/1389F for analysis of the 3' terminal region. The amplified DNA fragments were cloned into plasmid pCR2.1 using a TOPO TA Cloning kit (Thermo Fisher Scientific, MA, USA), and the nucleotide sequences were analyzed. Furthermore, the sequence data were compared with data from the GeneBank database in NCBI to identify the integrated site of the gene in the genome.

Estimation of transposable activity after Cre/lox recombination

Instead of estimating the transposable activity of the gene after Cre/lox recombination following the transposition from the modified transposon vector to the genome, we constructed and estimated plasmids having defective IR/DR structures predicted after Cre/lox recombination as follows.

HeLa/puro/Sg cells obtained by transfection of IR/DR-puro/LxpADb without pCAGGS/SB were transfected with pCAGGS/Cre and pLx/GFP/neo/pA(-), and HeLa/neo/Sg cells were then selected as the same manner described above. The HeLa/neo/Sg cells were expected to have the sequences of IR/DR-puro/LxpADb in the 5' and 3' ends and those of pLx/GFP/neo/pA(-) between lox sequences. The genomic DNA of HeLa/neo/Sg cells was subjected to genome walking to obtain DNA fragments with defective 5'- or 3'-IR/DR structures. Using these fragments, a plasmid (IR/DR-3'IR/pSP) in which only the 5'-IR/DR was defective and a plasmid (5'+3'IR/pSP) in which both the 5'- and 3'-IR/DRs were defective were constructed. Furthermore, *Puro* was inserted into the *SallI* site of these plasmids, resulting in IR/DR-3'IR/puro and 5'+3'IR/puro, respectively.

IR/DR-3'IR/puro with the defective 3'-IR/DR structure, 5'+3'IR/puro with both defective IR/DR structures, and IR/DR-puro/LxpADb with the complete IR/DRs (Figure 5) were evaluated to estimate their transposable activities as described above.

Sequence analysis

Nucleotide sequences of DNA fragments amplified by PCR during construction, clones obtained by the genome walking method, and constructed plasmids were determined using a Big Dye Terminator Cycle Sequencing FS Ready Reaction Kit and ABI PRISM 310 Genetic Analyzer (Thermo Fisher Scientific).

Acknowledgements

This work was supported by KM Biologics; potential conflict of interest is managed by KM Biologics. We are grateful to Dr. M. Araki for providing us with pCAGGS/Cre and lox sequences and for helpful advice regarding the Cre/lox system. We also thank Dr. S. Watanabe for providing us with pPGKpuro and Ms. R. Kawamura for her experimental assistance in this study.

References

1. Pettit DK, Rogers RS, Arthur K, Brodsky Y, Clark RH, et al. (2016) CHO cell production and sequence improvement in the 13C6FR1 anti-Ebola antibody. *Mabs* 8: 347–357. [[Crossref](#)]
2. Wu G, Bazer FW (2019) Application of new biotechnologies for improvements in swine nutrition and pork production. *J Anim Sci Biotechnol* 10: 28. [[Crossref](#)]
3. Mavilio F (2017) Developing gene and cell therapies for rare diseases: an opportunity for synergy between academia and industry. *Gene Ther* 24: 590–592. [[Crossref](#)]
4. Ivics Z, Hackett PB, Plasterk RH, Izsvák Z (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 91: 501–510. [[Crossref](#)]
5. Izsvák Z, Ivics Z, Plasterk RH (2000) Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J Mol Biol* 302: 93–102. [[Crossref](#)]
6. Luo GL, Ivics Z, Izsvák Z, Bradley A (1998) Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. *Proc Natl Acad Sci U S A* 95: 10769–10773. [[Crossref](#)]
7. Loeb KR, Hughes BT, Fissel BM, Osteen NJ, Knoblaugh SE, et al. (2019) Insertional mutagenesis using the Sleeping Beauty transposon system identifies drivers of erythroleukemia in mice. *Sci Rep* 9: 5488. [[Crossref](#)]
8. Lam WL, Lee TS, Gilbert W (1996) Active transposition in zebrafish. *Proc Natl Acad Sci U S A* 93: 10870–10875. [[Crossref](#)]
9. Koga A, Hori H (2001) The Tol2 transposable element of the medaka fish: an active DNA-based element naturally occurring in a vertebrate genome. *Genes Genet Syst* 76: 1–8. [[Crossref](#)]
10. Toole JJ, Knopf JL, Wozney JM, Sultzman LA, Buecker JL, et al. (1984) Molecular cloning of a cDNA encoding human antihemophilic factor. *Nature* 312: 342–347. [[Crossref](#)]
11. Ginsburg D, Handin RI, Bonthron DT, Donlon TA, Bruns GA, et al. (1985) Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) clones and chromosomal localization. *Science* 228: 1401–1406. [[Crossref](#)]
12. Doolittle RF (1984) Fibrinogen and fibrin. *Ann Rev Biochem* 53: 195–229.
13. Izsvák Z, Khare D, Behlke J, Heinemann U, Plasterk RH, et al. (2002) Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in Sleeping Beauty transposition. *J Biol Chem* 277: 34581–34588. [[Crossref](#)]
14. Albert H, Dale EC, Lee E, Ow DW (1995) Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant J* 7: 649–659. [[Crossref](#)]
15. Lee G, Saito I (1998) Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination. *Gene* 216: 55–65. [[Crossref](#)]
16. Liu K, Hipkens S, Yang T, Abraham R, Zhang W, et al. (2006) Recombinase-mediated cassette exchange to rapidly and efficiently generate mice with human cardiac sodium channels. *Genesis* 44: 556–564. [[Crossref](#)]
17. Choi S, Begum D, Koshinsky H, Ow DW, Wing RA (2000) A new approach for the identification and cloning of genes: the pBACwich system using Cre/lox site-specific recombination. *Nucleic Acids Res* 28: e19. [[Crossref](#)]
18. Zayed H, Izsvák Z, Walisko O, Ivics Z (2004) Development of hyperactive sleeping beauty transposon vectors by mutational analysis. *Mol Ther* 9: 292–304. [[Crossref](#)]
19. Huang X, Guo H, Tammana S, Jung YC, Mellgren E, et al. (2010) Gene transfer efficiency and genome-wide integration profiling of Sleeping Beauty, Tol2, and piggyBac transposons in human primary T cells. *Mol Ther* 18: 1803–1813. [[Crossref](#)]
20. Izsvák Z, Hackett PB, Cooper LJ, Ivics Z (2010) Translating Sleeping Beauty transposition into cellular therapies: victories and challenges. *Bioessays* 32: 756–767. [[Crossref](#)]
21. Ammar I, Gogol-Döring A, Miskey C, Chen W, Cathomen T, et al. (2012) Retargeting transposon insertions by the adeno-associated virus Rep protein. *Nucleic Acids Res* 40: 6693–6712. [[Crossref](#)]
22. Sundararajan P, Atkinson PW, O'Brochta DA (1999) Transposable element interactions in insects: cross mobilization of hobo and Hermes. *Insect Mol Biol* 8: 359–368. [[Crossref](#)]
23. Tipanee J, VandenDriessche T, Chuah MK (2017) Transposons: moving forward from preclinical studies to clinical trials. *Hum Gene Ther* 28: 1078–1104. [[Crossref](#)]
24. Tipanee J, Chai YC, VandenDriessche T, Chuah MK (2017) Preclinical and clinical advances in transposon-based gene therapy. *Bioscience Reports* 37: BSR20160614. [[Crossref](#)]
25. Grabundzija I, Irgang M, Mátés L, Belay E, Matrai J, et al. (2010) Comparative analysis of transposable element vector systems in human cells. *Mol Ther* 18: 1200–1209. [[Crossref](#)]
26. Niwa H, Yamamura K, Miyazaki J (1991) Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* 108: 193–199. [[Crossref](#)]
27. Radice AD, Bugaj B, Fitch DH, Emmons SW (1994) Widespread occurrence of the Tc1 transposon family: Tc1-like transposons from teleost fish. *Mol Gen Genet* 244: 606–612. [[Crossref](#)]
28. Araki K, Okada Y, Araki M, Yamamura K (2010) Comparative analysis of right element mutant lox sites on recombination efficiency in embryonic stem cells. *BMC Biotechnol* 10: 29. [[Crossref](#)]
29. Sambrook J, Russel DW (2001) Molecular cloning: a laboratory manual. (3rd ed.). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.