piggyBac as a high-capacity transgenesis and gene-therapy vector in human cells and mice

Rongbo Li1, Yuan Zhuang1,2, Min Han1,3, Tian Xu1,4,* and Xiaohui Wu1,*

SUMMARY
The stable genomic integration and expression of a large transgene is a major hurdle in gene therapy. We show that the modified piggyBac (PB) transposon system can be used to introduce a 207 kb genomic DNA fragment containing the RORγ/γt locus into human cells and mice. PB-mediated transgenesis results in a single copy of a stably inherited and expressed transgene. These results indicate that PB could serve as an effective high-capacity vector for functional analysis of the mammalian genome and for gene therapy in human cells.

INTRODUCTION
Gene therapy requires a safe and efficient way to introduce target DNA fragments with a defined orientation into cells to compensate for the deleterious effects of a mutated gene. A number of viral and non-viral gene delivery systems have been developed over the years (Atkinson and Chalmers, 2010; Thomas et al., 2003). Retroviral, lentiviral and adeno-associated viral vectors have been shown to be effective for gene introduction. In addition, stable integration of retroviral and lentiviral vectors benefit long-term expression of the transgene. However, limited cargo capacity prevents these vectors from being used to introduce large transgenes and genomic sequences. Adenoviral vectors are able to carry inserts of up to 30 kb, but they remain episomal after transduction and trigger immune responses easily, both leading to transient expression of the therapeutic DNA (Atkinson and Chalmers, 2010). In addition, virus-related risks are common concerns of using viral vectors because death and leukemia cases have been reported in clinical trials using adenoviral and retroviral vectors for gene therapy, respectively (Hacein-Bey-Abina et al., 2003; Raper et al., 2003).

As non-viral transgenic vectors, DNA transposons are attractive choices for gene therapy. We have previously shown that the modified piggyBac (PB) DNA transposon, which was originally identified in the cabbage looper moth Trichoplusia ni (Fraser et al., 1996), is highly efficient in mediating stable integration and expression of transgenes in human cells and in mice (Ding et al., 2005). Recently, PB has been shown to be able to carry fragments of more than 100 kb to hop in human and mouse embryonic stem (ES) cells (Li et al., 2011; Rostovskaya et al., 2012). Here we analyzed the in vivo expression of an extra endogenous locus carried by PB-mediated BAC transgenes. Our data suggest PB as a potential candidate of high-capacity gene-therapy vector.

RESULTS
Construction of PB/shuttle and PB[BAC]
To test whether PB can mediate the integration of a large DNA fragment in human cells, we built a PB element carrying BAC genomic DNA based on a new PB transposon shuttle vector (PB/shuttle). PB/shuttle was developed by incorporating BAC homologous sequences (Box A and Box B) and a PGK promoter-driven neomycin resistance (Neo) gene into the transposable element (Sparwasser et al., 2004) (Fig. 1A). Any BAC clones can be introduced into PB/shuttle via homologous recombination resulting a transposable unit containing the entire BAC sequence and the Neo selection marker. A diphtheria toxin A expression cassette (Dta) was placed in the PB/shuttle vector outside of the PB end termini (PBL and PBR) for selection against random integration events of the PB[BAC] construct (Fig. 1A). We then generated a PB[BAC] clone with a 207.7 kb BAC, RP23-263k17, containing the RORγ/yt locus of C57BL/6J mice (Fig. 1B). Correct recombinants were confirmed by PCR analysis with primers flanking each of the homologous sequences before further experiments (supplementary material Fig. S1).

PB-mediated BAC gene transfer in human 293 cells
We transfected PB[BAC] into human 293 cells with or without a CMV-PBase helper plasmid that ubiquitously expresses piggyBac transposase (PBase). Co-transfection with PBase generated on average sixfold more neomycin-resistant clones than the PBase negative control (Fig. 1C,D). Inverse PCR with primers on one of the PB termini was successfully used to isolate 15 integration sites from 13 clones derived from the co-transfection experiment (Fig. 2B and supplementary material Table S1). The TTAA target sequence of the PB transposition was found to be present in all 15 sites through inverse PCR. TTAA duplications were further confirmed in all five attempts by PCR analysis of the other PB termini (Fig. 2A-C), demonstrating that these integrations occur through a clean PB transposition.

To examine whether the entire cargo is delivered into the genome after transposition, we examined the integrity of the...
High-capacity piggyBac transgenesis

**Clinical issue**
In gene therapy, DNA fragments are introduced into cells to override the deleterious effects of a disease-causing mutation, most commonly by replacing a dysfunctional or abnormal gene with a functional version. A number of viral and non-viral gene-transfer systems have been developed, but their clinical application is currently limited owing to associated risks and shortcomings. A major hurdle is the low ‘cargo capacity’ of existing vectors, which prevents them from being used to deliver full-length genes and genomic sequences. Adenoviral vectors can carry inserts of a large size; however, viral vectors are associated with a lack of copy number control, a propensity for deleterious integration and the elicitation of unwanted immunological responses in target cells. As non-viral transgenic vectors, DNA transposons represent a more reliable and safe agent for the delivery of large transgenes into human cells.

**Results**
It was recently shown that the piggyBac (PB) transposon can carry DNA fragments that are greater than 100 kb in size. In the present study, the authors demonstrate that a modified PB transposon system can efficiently introduce an intact mammalian gene of more than 200 kb into human and mouse cells in vitro. Moreover, they show that the gene is integrated with a defined orientation, and is stably integrated and inherited in mice in vivo. Finally, the authors report that expression of the transgene mimics that of the endogenous gene.

**Implications and future directions**
The modified PB transposon system facilitates the stable integration and inheritance of large DNA fragments, and allows controlled expression of transgenes in situ. The vector has the capacity to carry regulatory elements associated with a gene to overcome positional effects of expression, and to deliver genomic sequences encompassing wild-type alleles. Thereby, the system provides an efficient and reliable transgenic and gene-therapy tool that is particularly promising for the treatment of cystic fibrosis and Huntington’s disease, where the affected gene and flanking regions cover long stretches of DNA.

**PB-mediated BAC gene transfer in mice**
We then examined whether PB-mediated BAC gene transfer could be stably inherited in vivo. To achieve this goal, we first introduced PB[BAC] into mouse W4/129S6 embryonic stem (ES) cells. Co-transfection with the Act-PBase helper plasmid produced 25 times more neomycin-resistant clones than the PBase negative control (Fig. 1E,F) (Ding et al., 2005). Inverse PCR detected a single band in 84% (31/37) of the clones tested, suggesting that most of the clones carried a single insertion of PB[BAC] (supplementary material Table S1). This outcome was consistent with the result from some clones tested by quantitative real-time PCR (data not shown). To further evaluate the integration events, we performed fluorescence in situ hybridization (FISH) in two ES clones expected to carry a single insertion of PB[BAC]. Full-length RP23-263k17 BAC DNA was used as the probe to detect both endogenous loci (paternal and maternal alleles) and any transgenic copies present in the genome. FISH analysis of mitotically arrested cells identified three foci in each of the ES clones (Fig. 3A). In addition, analysis of an ES clone that gave two bands in inverse PCR (ES37) showed four hybridization signals in FISH (Fig. 3A).

We then performed blastocyst injection to produce transgenic animals. We generated three transgenic lines in total. Two of them were derived from the ES clone ES37 (supplementary material Table S2). We first confirmed the integrity of PB[BAC] in the progenies by genotyping four SNP markers in the BAC DNA (Fig. 3B and supplementary material Table S3). The transmission of PB[BAC] was then followed by PCR using PB and flanking genomic primers. In all cases the transgene was stably inherited.

**PB-mediated BAC transgene expression in mice**
Generation of transgenic animals allowed us to further study the expression pattern of the transgene. The RORγt locus encodes two alternatively spliced transcripts differing in their 5’-terminal sequences. RORγ transcripts are expressed broadly in adult tissues. In contrast, RORγt mRNA can only be detected in CD4+CD8+ double-positive thymocytes (He et al., 1998; Kurebayashi et al., 2000). In transgenic animals, RORγ transcription is readily detected in kidney, liver and muscle by RT-PCR (Fig. 4A,D,E and supplementary material Fig. S3). In contrast, RORγt mRNA can only be detected in the thymus, but not the above tissues (Fig. 4A,C and supplementary material Fig. S3). These results suggest that PB[BAC] expression follows the same tissue specificity as the expression of endogenous alleles.

Transgenic expression is often affected by the surrounding genomic context. To evaluate the positional effects of PB-mediated transgene expression, we compared the expression of an Act-RFP transgenes in 101 PB[Act-RFP] lines. These randomly chosen transgenic lines were generated from mobilization of the same PB[Act-RFP] construct. Each line carries a single insertion at a different location in the genome. The Act-RFP transgene allows us to visually assess the level of gene expression in the transgenic animals. We have observed that the majority of the transgenic lines exhibit a comparable level of red fluorescence (68/101=67.3%), whereas 15.8% (16/101) and 16.9% (17/101) of the transgenic lines have either stronger or weaker signals, respectively (supplementary material Fig. S4A,B). Real-time RT-PCR analysis confirmed similar distribution of RFP expression levels in 18 randomly picked transgenic lines from this pool (supplementary material Fig. S4C). These data show that, similar to transposon-mediated transgenesis in Drosophila, PB can serve as a good transgenic tool in mammals despite some positional effects.

We next studied whether increasing the copy number of the PB[BAC] transgene would correlate with elevated of gene expression, by examining the expression level of the RORγ and RORγt transcripts in different PB[BAC] transgenic lines. We first measure the copy number of each transgenic integration by quantitative real-time PCR. Two of the transgenic lines carried a single copy transgene, while the other line (37-1) carried two closely linked PB[BAC] insertions (Fig. 4B). Real-time RT-PCR indeed detected a 66% increase in RORγt expression in line 37-1, and about...
33% increase in expression in lines 37-2 and 27 (Fig. 4C). Similarly, we measured the RORγ transcript in kidney and liver samples in two lines. Transgenic expression of the RORγ transcript of line 37-1 is approximately twofold of that in line 37-2 (Fig. 4D,E). These data are consistent with the expectation that PB-mediated BAC transgene expression correlates with the copy number of the PB[BAC] transgene.

Our previous work suggested that the PB element allows transgene expression without significant silencing or epigenetic modification among generations (Ding et al., 2005). This feature also applied for PB-mediated BAC transgenesis. We analyzed RORγt expression in different generations of lines 37-1 and 37-2 by real-time RT-PCR. In both lines, the expression level remained stable among generations (supplementary material Fig. S5).

Fig. 1. PB-mediated BAC gene transfer in human and mouse ES cells. (A) Diagram of the PB/shuttle vector. The head-to-head PB termini (PBR and PBL) are separated by the negative selection marker Dta. This cassette, along with the positive selection marker Neo, is placed between the BAC homologous recombination sequences Box A and Box B. The vector also carries the R6Kγ origin of replication, RecA, AmpR (ampicilin-resistance) and SacB genes are required for proper recombination and selection in E. coli cells (Sparwasser et al., 2004). (B) Diagram of PB[BAC]. Primers used to confirm the recombination events are marked with arrows. (C,D) An example of PB[BAC] integration in human 293 cells with CMV-PBase. Clones were stained with methylene blue after G418 selection (C) and the statistical results of three transfections plotted (D). (E,F) Similar experiments were repeated in murine W4 ES cells with Act-PBase. ***P<0.001.

Fig. 2. Molecular analysis of PB[BAC] insertions in human 293 cells. (A-F) Inverse PCR with primers PBLinvB1 and PBLinvF1 (A) recovered single insertions in most of the clones (B). The insertion sites were confirmed by genotyping PCR using pairs of primers targeting both flanking genomic sequences (G-RL or G-RB) and PB terminal sequences (PBL-B or PBR-F), as shown in panel C. Integrity of PB[BAC] transgenes was analyzed by PCR using ten primer pairs evenly spaced along the BAC insert (fragments 1-10 in panel D). Two clones showed all positive results (E) and were further examined by long-range PCR using 21 pairs of overlapping primers covering the full length (fragments A-U in panel D), as shown in panel F and supplementary material Fig. S2. M, 1 kb ladder; +, positive control with PB[BAC] as the template, −, negative control with 293 genomic DNA as the template; stars mark location of SNPs used to verify BAC integrity in transgenic mice.
DISCUSSION
A large cargo capacity is always desired for gene-therapy vectors. It allows the transgene to carry more regulatory elements to overcome the positional effect on expression. It also permits the transgene to carry more regulatory elements to allow the transgene to carry more regulatory elements to the genome, resulting in single copy integrations with defined orientation of the transgene. These features are attractive for gene-therapy vectors because the insertion site could be easily mapped to recognize hazardous integration events before further use. Such a process is essential for BAC and other large transgenes, because rearrangements might occur during transgenic manipulations (Giraldo and Montoliu, 2001). In fact, we observed that 45% of the PB[BAC] 293 clones had deletions in the transgene.

We developed a shuttle recombineering vector that allows fast construction of PB-based BAC transgenes. The vector has both positive and negative selection markers that lead to fast enrichment of the clones carrying transposition events. It provides a convenient tool for PB-based transgenesis with commonly used pBACe3.6 BACs. With properly modified homologous arms, the design should also be applicable for other BACs.

An ideal gene-therapy vector allows stable expression of the transgene that preferably has the same tissue specificity of the endogenous allele. In our study, the RORγ/yt locus carried by PB[BAC] expressed RORγt thymus specifically in transgenic animals. The expression level of both RORγt and RORγ are position independent but copy number dependent. However, addition of each copy of the 129/SvJ genome produced a 66% increase in RORγt transcripts of an endogenous allele. This could due to differential allelic expression (DAE), which is commonly observed in mammals (Cowles et al., 2002; Knight, 2004; Tang et al., 2011). Alternatively, the 200-kb BAC genomic sequences might not fully cover unknown remote regulatory sequences of the RORγt/yt locus. Nevertheless, the stable integration and expression of PB-mediated BAC gene transfer for DNA of more than 200 kb renders PB an attractive candidate vector for gene therapy of many known disease genes. It also provides a powerful tool for transgenesis in other mammals and for the functional dissection of the mammalian genome.

MATERIALS AND METHODS
PB/Shuttle construction
All primer sequences are listed in supplementary material Table S4. The SmaI fragment of pBac-AB (Ding et al., 2005) containing PB was cut and cloned into the EcoRV site to form pBac-BA. Homologous arms Box A and Box B was PCR-amplified with primers BXA1/BXA2 and BXB1/BXB2, and cloned into T vectors to form T-AB and T-CD, respectively. The EcoRI-NdeI fragment of T-AB was then replaced by the EcoRI-NdeI fragment of T-CD to form T-ABCD. The BglII-EcoRI fragment of T-ABCD was further replaced by the BglII-EcoRI fragment of pBac-BA to form T-ABCD-PB. Dta and neomycin cassettes of pGKNEO-F2L2-DTA were cloned into the XhoI and SmaI sites of T-ABCD-PB, respectively, to form T-ABCD-PB-Dta-Neo. The NotI fragment of the shuttle vector pDelsac (Sparwasser et al., 2004) was finally replaced by the NotI fragment of T-ABCD-PB-Dta-Neo to form PB/Shuttle.

PB[BAC] construction
PB/Shuttle was used in a two-step RecA strategy for PB[BAC] construction as previously described (Sparwasser et al., 2004). Primer pairs BAC-14/PB-LF1 and BAC-15/Neo-F were used to identify the positive clones. BAC DNA was purified by using NucleoBond BAC 100 (Macherey-Nagel, cat. no. 740579).

Cell transfections
293 cells were cultured in DMEM (GIBCO-BRL) medium containing 10% fetal bovine serum at 37°C and 5% CO2. For each well of a 24-well-plate, 2x10⁵ cells were seeded one day before transfection, transfected with 0.8 μg of circular PB[BAC] and 0.2 μg of circular CMV-PBase DNA as the test group, or the same amount of circular PB[BAC] with 0.2 μg of circular pcDNA4/HisA as the control. Transfection was performed with Lipofectamine 2000 following the standard protocol (Invitrogen). One day after transfection, cells were trypsinized and seeded onto one 10-cm plate. Drug selection with 500 μg/ml G418 (Gibco-BRL) was then continued for two weeks.
ES cells were cultured on 10-cm plates containing mitomycin-C-treated mouse embryonic fibroblast cells. 20 μg of PB[BAC] and 5 μg of Actin-PBase or pCX in the control group were used for electroporation of 10×10⁶ cells. Immediately after electroporation, cells were divided onto three 10-cm plates. Selection with 200 μg/ml G418 was initiated 24 hours after electroporation and lasted for one week. Cells were then fixed at room temperature with PBS containing 4% paraformaldehyde for 10 minutes and stained with 0.2% methylene blue for 1 hour.

**PCR**

To recover the flanking genomic sequences of PB insertions, HaeIII digests of the genomic DNA were self-ligated to serve as the template for inverse PCR. Primers used to recover the genomic sequences flanking the PBR terminus were PBRinvF3 and PBRinvB3.

To measure the copy number of BAC, genomic DNA extracted from mouse tails were used as the template for real-time PCR. Primers RT-BAC-F and RT-BAC-B were used for the amplification.

To examine the integrity of BAC transgenes, long-range PCR was carried out using the LongAmp Taq PCR kit (NEB).

**Fluorescence in situ hybridization**

We prepared metaphase chromosome spreads from ES cells. BAC DNA was probed with biotin by nick translation (Roche) and stained with Avidin-FITC (Roche). Chromosomes were identified by the banding patterns with DAPI staining. Photographs were taken with a Leica DMRXA2 camera and images prepared with Leica FW4000 and Adobe Photoshop.

**Statistics**

P-values were calculated with two-tailed unpaired Student’s t-test on Prism 5 (GraphPad Software).

**ACKNOWLEDGEMENTS**

We thank Yanling Yang, Fang Wang, Chunyan Gao, and Yuping Qiao for technical assistance.

**COMPETING INTERESTS**

The authors declare that they do not have any competing or financial interests.

**AUTHOR CONTRIBUTIONS**

X.W. and T.X. designed the project. R.L. did the experiments and collected data; Y.Z., M.H., T.X. and X.W. analyzed the data. R.L., T.X. and X.W. prepared the manuscript. X.W. supervised all aspects of the work.
FUNDING
This work was supported by Chinese Key Projects for Basic Research (973) [grant number 2006CB806701], Hi-tech Research and Development Project (863) [grant number 2007AA022101], 211 and 985 projects of Chinese Ministry of Education, Shanghai Rising-Star Program [grant number 10QH1400100] and the ‘Shu Guang’ project of Shanghai Municipal Education Commission and Shanghai Education Development Foundation.

SUPPLEMENTARY MATERIAL
Supplementary material for this article is available at http://dmm.biologists.org/lookup/suppl/doi:10.1242/dmm.010827/-/DC1

REFERENCES
Atkinson, H. and Chalmers, R. (2010). Delivering the goods: viral and non-viral gene therapy systems and the inherent limits on cargo DNA and internal sequences. Genetica 138, 485-498.

Cowles, C. R., Hirschhorn, J. N., Altshuler, D. and Lander, E. S. (2002). Detection of regulatory variation in mouse genes. Nat. Genet. 32, 432-437.

Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y. and Xu, T. (2005). Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell 122, 473-483.

Fraser, M. J., Ciszczon, T., Elick, T. and Bauser, C. (1996). Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. Insect Mol. Biol. 5, 141-151.

Giraldo, P. and Montoliu, L. (2001). Size matters: use of YACs, BACs and PACs in transgenic animals. Transgenic Res. 10, 83-103.

Gusella, J. F., MacDonald, M. E., Ambrose, C. M. and Duyao, M. P. (1993). Molecular genetics of Huntington's disease. Arch. Neurol. 50, 1157-1163.

Haecin-Bey-Abina, S., von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntryre, E., Radford, L., Villeval, J. L., Fraser, C. C., Cavazzana-Calvo, M. et al. (2003). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. N. Engl. J. Med. 348, 255-256.

He, Y. W., Deftos, M. L., Ojala, E. W. and Bevan, M. J. (1998). RORgamma t, a novel isoform of an orphan receptor, negatively regulates Fas ligand expression and IL-2 production in T cells. Immunity 9, 797-806.

Kerem, B., Rommens, J. M., Buchanan, J. A., Markiewicz, D., Cox, T. K., Chakravorti, A., Buchwald, M. and Tsui, L. C. (1989). Identification of the cystic fibrosis gene: genetic analysis. Science 245, 1073-1080.

Knight, J. C. (2004). Allele-specific gene expression uncovered. Trends Genet. 20, 113-116.

Kurebayashi, S., Ueda, E., Sakaue, M., Patel, D. D., Medvedev, A., Zhang, F. and Jetten, A. M. (2000). Retinoid-related orphan receptor gamma (RORgamma) is essential for lymphoid organogenesis and controls apoptosis during thymopoiesis. Proc. Natl. Acad. Sci. USA 97, 10132-10137.

Li, M. A., Turner, D. J., Ning, Z., Yusa, K., Liang, Q., Eckert, S., Rad, L., Fitzgerald, T. W., Craig, N. L. and Bradley, A. (2011). Mobilization of giant piggyBac transposons in the mouse genome. Nucleic Acids Res. 39, e148.

Raper, S. E., Chirmule, N., Lee, F. S., Wivel, N. A., Bagg, A., Gao, G. P., Wilson, J. M. and Batshaw, M. L. (2003). Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. Mol. Genet. Metab. 80, 148-158.

Rostovskaya, M., Fu, J., Obst, M., Baer, I., Weidlich, S., Wang, H., Smith, A. J., Anastassiadis, K. and Steward, A. F. (2012). Transposon-mediated BAC transgenesis in human ES cells. Nucleic Acids Res. 40, e150.

Sparwasser, T., Gong, S., Li, J. Y. and Eberl, G. (2004). General method for the modification of different BAC types and the rapid generation of BAC transgenic mice. Genesis 38, 39-50.

Tang, F., Barbacioru, C., Nordman, E., Bao, S., Lee, C., Wang, X., Tuch, B. B., Heard, E., Lao, K. and Surani, M. A. (2011). Deterministic and stochastic allele specific gene expression in single mouse blastomeres. PLoS ONE 6, e21208.

Thomas, C. E., Ehrhardt, A. and Kay, M. A. (2003). Progress and problems with the use of viral vectors for gene therapy. Nat. Rev. Genet. 4, 346-358.