

A quantitative model for efficient construction of lentiviral vectors with a unique clone site

Gang Zhang (Ph. D), Anurag Tandon (Ph.D)

**Department of Medicine, Centre for Research in Neurodegenerative Diseases,
University of Toronto, 6 Queen's Park Crescent West, Toronto, ON, M5S 3H2, Canada**

Correspondence and requests for materials should be addressed to G. Z.

(gang.zhang@utoronto.ca) or A. T. (a.tandon@utoronto.ca).

Lentiviral vectors (LVs) were a powerful tool for transgene expression *in vivo* and *in vitro*. However, the construction of LVs is of low efficiency, due to the large sizes and lack of proper clone sites. Therefore, it is critical to develop efficient strategies for cloning LVs. Here, we reported a combinatorial strategy to efficiently construct LVs using EGFP, hPlk2 wild type (WT) and mutant genes as inserts. Firstly, site-directed mutagenesis (SDM) was performed to create BamH I site for the inserts; secondly, pWPI LV was dephosphorylated after BamH I digestion; finally, the amounts and ratios of the insert and vector DNA were optimized to increase monomeric ligation. Our results showed that the total percentage of positive clones was approximately $51.3\% \pm 15.2\%$. Using this model, almost all the vectors could be constructed through two or three minipreps, therefore, our study provided an efficient quantitative model for constructing large-size vectors.

Lentiviral vectors (LVs) are a powerful tool for gene transduction *in vitro* and *in vivo*, because of the following advantages. Firstly, LVs can transduce not only mitotically active cells, but also slowly dividing cells, and even nondividing terminally differentiated cells¹⁻⁴. Secondly, transgenes delivered by LVs are more resistant to transcriptional silencing, whereas, it is a common phenomenon associated with retroviral vector transduced cells after prolonged *in vitro* or *in vivo* transplantation²⁻⁴. In addition, LVs can accommodate the use of various ubiquitous or tissue-specific transcriptional promoters^{3,5}. Furthermore, the self-inactivating safety modification of LVs, which permanently disables the viral promoter within the viral long-terminal repeat after integration, enables transgene's expression in the targeted cells to be controlled solely by internal promoters^{3,6}. Finally, self-inactivating modification of LVs does not reduce viral titers significantly⁶. These advantages make LVs to be a powerful tool for stable gene transfer and expression⁷⁻⁹.

However, it is highly challenged to subclone interested genes into the third-generation bicistronic LVs, due to their large sizes and limited clone sites, therefore, it is important to develop efficient strategies for LV subcloning. Classically, two different matched sites are used to directly subclone the inserts into vectors by ligation and transformation¹⁰. Because of the limited clone sites, this strategy is almost not applicable for LV cloning. To create matched clone sites, the incorporation of restriction sites into the primers used for PCR is perhaps the most common strategy, but this method decreases ligation efficiencies due to the inability of some restriction endonucleases to cleave sites efficiently near the termini of DNA molecules¹⁰⁻¹³. Another strategy is to create blunt ends for the vectors and/or inserts with the Klenow fragment of DNA Polymerase I or T4 DNA Polymerase, but this method can

generate recessed ends due to their 3'-5' exonuclease activity¹⁴. As a result, the efficiencies of afterward ligation and transformation are also significantly decreased. SDM was first established at 1978 by Hutchison et al., and it is essential in gene functional studies, genetic engineering, protein engineering, and vector modifications¹⁵⁻¹⁷. Currently, the QuikChange™ SDM System developed by Stratagene is a commonly used kit for mutagenesis using plasmid double-stranded DNA as templates. The advantage of this strategy is that the products after mutagenesis are circular, double-stranded DNA. After restriction endonuclease digestion and purification by agarose gel electrophoresis, theoretically, 100% of the linearized DNA fragments are with correct-digested ends. Therefore, maximal efficiencies of ligation can be achieved.

Up to now, to our knowledge, there are still no reports on how to efficiently construct LVs with restriction-dependent cloning strategies. The aim of this study is to optimize a combinatorial method to efficiently construct LVs, and the strategy described herein is also suitable for constructing other expression vectors. In order to investigate the relationship between human polo-like kinas 2 (hPlk2) expression and α -synuclein phosphorylation *in vitro*, we adopted a modified pWPI LV, which carried an EF1 α -IRES-Neomycin (Neo) cassette for the expression of dual genes, to construct LVs of hPlk2 WT and mutants, K111M, T239D, T239V, and enhanced fluorescence protein gene (EGFP) as control, respectively. The size of the pWPI vector is about 11.4 kb, and there is a unique BamH I clone site within the vector. To quantitatively establish an efficient model for constructing LVs, we developed a combinatorial method, which included inserting a BamH I site at the 3'-ends of hPlk2 and EGFP through SDM to create compatible clone sites; dephosphorylating the vector DNA

after BamH I digestion by calf intestinal phosphatase (CIP) treatment to protect the self-ligation of vectors; optimizing the amounts and ratios of the insert and vector DNA to facilitate the monomeric ligation; and using Top10 competent cells as hosts to improve the transformation efficiencies. With our optimized methods, we successfully constructed five LVs, and the percentages of positive clones containing EGFP, hPlk2 WT, K111M, T239D, and T239V inserts, were $51\% \pm 16.5\%$, $37\% \pm 15.2\%$, $52\% \pm 33.6\%$, $45\% \pm 16.2\%$, and $75\% \pm 25\%$, respectively. This study provided an important quantitative model for efficient construction of LVs.

Results

BamH I site insertion and mutagenesis of hPlk2 mutants

SDM was performed by PCR, and the parental template DNA was digested with Dpn I. Different annealing temperatures were used for pcDNA3.1 BamH I insertion and hPlk2 mutagenesis, representing approximately $T_m - 5^\circ\text{C}$, $T_m^* + 2^\circ\text{C}$, and $T_m^* - 5^\circ\text{C}$ (Figure 1, Table 1, 2 and 3), respectively. After transformation and identification by sequencing, we found that, when lower annealing temperatures (approximately $T_m^* - 5^\circ\text{C}$) were adopted, more colonies could be obtained, and the mutagenesis efficiencies were approximately 90% for pcDNA3.1/hPlk2/BamH I insertion, 50% for K111M mutagenesis, 60% for T239D and 90% for T239V mutagenesis, respectively (Table 1, 2). Whereas, when higher annealing temperatures (approximately $T_m - 5^\circ\text{C}$) were used, no colonies were obtained (Table 3). Interestingly, when the mutagenesis reactions were annealed at $T_m^* + 2^\circ\text{C}$, a few colonies

were obtained after transformation, but the mutagenesis efficiencies were 100% (Table 3). These data suggested that lower annealing temperatures were more efficient both for the BamH I insertion and hPlk2 mutagenesis, however, higher annealing temperatures resulted in decreased transformation colonies (Table 1, 2 and 3). Our results were in accordance with other report that the annealing temperatures should be determined empirically case by case¹⁰. In addition, although the Statagene developed protocol suggested that the primer pairs should be of 25~45 bases in length with melting temperature (T_m) \geq 78°C, our experiments showed that correct mutagenesis was achieved with longer primers up to 56 bases, and the mutations ranged from 2 to 6 bases (Table 1). DH5 α competent cells were also used for transformations, but no positive clones were obtained after sequencing (Data not shown). Therefore, for low efficiency mutagenesis, Top10 cells were recommended as transformation hosts, because their transformation efficiency could reach up to 1×10^9 cfu/ μ g supercoiled DNA, whereas, DH5 α only about 10^6 cfu/ μ g supercoiled DNA (Invitrogen). In addition, we also performed mutagenesis for pEGFP-N1 BamH I site insertion, after annealed at 62°C, the mutation efficiency was about 83.5% (Figure 1, Table 2). These data demonstrated that the designed mutagenesis was successfully achieved.

CIP treatment improved the rates of recombinant vectors

A unique BamH I site of pWPI/Neo/BamH I LV was employed as clone site. To protect the vector self-circularization, we removed the 5'-phosphate groups of the vector DNA with CIP treatment following BamH I digestion. This treatment could diminish the background of transformed bacterial colonies that carried empty plasmids. At the same time, the colony number would be significantly decreased after transformation^{10, 18, 19}. Insert (EGFP, hPlk2 WT, K111M, T239D and T239V, respectively) and vector DNAs were pooled together in 10 μ l ligation reactions with total DNA concentrations around 19.0~25.3 ng/ μ l, and insert to

vector molar ratios about 1~3:1 (Table 4). To improve the transformation efficiencies, Top10 cells were used for transformation. Compared with DH5 α cells, more colonies were obtained from Top10 cell transformations, ranging from 19 to about 200 colonies from one transformation of each; whereas, totally 1 to 7 colonies were gained from 3 DH5 α cell transformations of each, respectively (Table 4). 14 to 20 of those colonies were selected (totally 83 colonies) for identification, and the results showed that only one colony of pWPI/hPlk2 WT/Neo was empty vector (Figure 3A), five colonies contained unknown DNA (Figure 3, A, B and D), and all the other colonies contained inserts (Figure 2 and 3; Table 4), therefore, the total percentage of recombinant vectors was up to 92.8% [(83-5-1)/83]. Among them, the individual percentages of inserted vectors of EGFP, hPlk2 WT, K111M, T239D, and T239V, were 100% \pm 0, 78% \pm 2.9%, 85% \pm 17.2%, 100% \pm 0, and 95% \pm 11.2%, respectively. Interestingly, the percentages of inserted vectors between EGFP/K111M, EGFP/T239D, EGFP/T239V, hPlk2 WT/K111M, K111M/T239D, K111M/T239V, and T239D/T239V, had no significant differences ($P>0.05$), but those between EGFP/hPlk2 WT, T239D/hPlk2 WT, and T239V/hPlk2 WT were significantly different ($P<0.05$) (Table 4), although the sizes of the hPlk2 WT and mutants were same, and among them, 2 or 3 bases were changed (WT to T239V: ACG to GTG; WT to T239D: ACG to GAC). These data demonstrated that the CIP treatment efficiently protected the self-ligation of the vectors, and increased the rates of recombinant vectors.

Efficient construction of LVs with monomeric, correct orientations

After miniprep, the insert numbers and orientations were primarily identified by Not I digestion. Because there was a unique Not I restriction endonuclease site inside pWPI/Neo/BamH I vector at 1145 nucleotide (nt) and the BamH I clone site at 3502 nt. There

was also a Not I site at the 3'-end of EGFP (about 0.75kb), therefore, after Not I digestion, if the vector was empty, one band was detected by Agarose Gel Electrophoresis; if the vector contained one copy of the insert and in correct orientation, two bands of ~3.1 kb and ~9 kb sizes were detected; and if the inserts were in opposite orientation, two bands of ~2.4 kb and ~9.75 kb sizes were detected (Figure 2). In addition, within the 5'-BamH I-hPlk2-3'-BamH I fragment (about 2.1kb), there was a Not I site at 118 nt as well. As a result, if the vector contained one copy of the insert and in correct orientation, two bands of ~2.5 kb and ~11 kb sizes were detected; and if the inserts were in opposite orientation, two bands of ~4.4 kb and ~9.2 kb sizes were detected (Figure 3). All the positive vectors were further confirmed by sequencing. Our data revealed that the total percentage of monomeric inserted vectors of EGFP, hPlk2 WT, K111M, T239D, and T239V, was approximately 88%, and the individual percentages of those were 95%±10%, 78%±2.9%, 85%±17.2%, 83%±15.6%, and 95%±11.2%, respectively. In addition, the percentages of monomeric inserted vectors between EGFP/K111M, EGFP/T239D, EGFP/T239V, hPlk2 WT/K111M, hPlk2 WT/T239D, K111M/T239D, K111M/T239V, and T239D/T239V, had no significant differences ($P>0.05$), but those between EGFP/hPlk2 WT, and hPlk2 WT/T239V were significantly different ($P<0.05$) (Figure 2 and 3; Table 4). Furthermore, the total percentage of positive clones of EGFP, hPlk2 WT, K111m, T239D, and T239V, in which monomeric, corrected-oriented inserts were carried, was about 51.3%±15.2%, and the individual percentages of those were about 51%±16.5%; 37%±15.2%; 52%±33.6%; 45%±16.2%; and 75%±25%, respectively (Table 4). Finally, the percentages of positive clones between hPlk2 WT/T239V had significant difference ($P<0.05$), but all the others were not significantly different ($P>0.05$).

As controls, pWPI vector DNA, which was not treated with CIP, was also used for ligation with the inserts, EGFP, hPlk2 WT, K111M, T239D and T239V, respectively. After transformation into Top10 cells, 10 clones of each were selected for identification. Our results revealed that, among the 50 clones identified, 5 clones were with monomeric insert, and one clone of pWPI/hPlk2T239V/Neo was with monomeric corrected-oriented insert (Table 5). The total percentage of positive clones was $2\% \pm 4.5\%$, compared with the CIP-treated experiment data, the difference between them was highly significant ($2\% \pm 4.5\%$ vs $51.3\% \pm 15.2\%$, $P < 0.01$).

Discussion

Creation of compatible ends between vectors and inserts is the first step for subcloning, and it is critical that the ends of DNA fragments are correctly generated after restriction digestion, particularly for large-size, low-efficiency subcloning. Regular vectors, such as pcDNA3 and pcDNA4 (Invitrogen), contain multiple clone sites, therefore, it is convenient to choose two different clone sites between the vectors and inserts for directional subcloning. Generally, LVs carry limited clone sites, and a unique clone site usually used for subcloning, such as BamH I or EcoR V^{5, 20} etc. Therefore, it is necessary to create compatible clone sites between the vectors and inserts. The incorporation of restriction endonuclease sites into the primers for PCR is a commonly used strategy, but the extra 3-4 bases at the terminal restriction sites are insufficient for stable association with and cutting by certain restriction endonucleases, therefore the cloning efficiencies were low, for example, as reported $\leq 0.05\%$ ^{11, 13}. Another strategy is to create blunt ends for the vectors and/or inserts with Klenow fragment of DNA Polymerase I or T4 DNA Polymerase. This method is also of low efficiency due to two reasons. One is the low efficiency of blunt end ligation compared with cohesive ends, and the other is the inefficiency of creation of correct blunt ends, for example, the filling-in

efficiency of Klenow was reported as $\leq 50\%$ ^{14, 21, 22}. In our previous LV vector subcloning experiments, blunt end was created for the inserts with Klenow and T4 DNA polymerase treatment, after dozens of ligations and transformations with Top10 cells, no positive clones were obtained (Data not shown). SDM is a powerful tool to change DNA sequences at specific positions in genetic engineering, including insertion of restriction sites^{16, 17, 23}. According to the working format of the QuikChangeTM SDM System, using double-stranded plasmid DNA as templates, the mutagenesis products after Dpn I digestion and transformation are circular double-stranded plasmid DNAs (Stratagene). Therefore, after restriction digestion, 100% of the purified linearized DNA fragments through Agarose Gel Electrophoresis were theoretically with correct ends. Whereas, the correct cutting ends could not be confirmed with the method of incorporating of restriction sites into PCR primers^{11, 13}. Our data demonstrated that the clone efficiencies were significantly improved with this strategy. The total percentage of recombinant clones was about 92.8%, and the percentage of positive clones with monomeric, corrected-oriented inserts was around 51.3% \pm 15.2% (Table 4, Figure 2, 3).

Design of the primers is crucial for successful mutagenesis. According to the guide of the QuikChangeTM SDM System of Stratagene, complementary primer pairs were employed in the same PCR. One disadvantage of using complementary primer pairs is the formation of “primer dimers” in PCR reactions, therefore the yield of successful transformants is reduced. This phenomenon is particularly severe in SDM, because the primers used include mismatched nucleotides for generating the desired mutations²³. Another disadvantage of the Stratagene strategy is that the newly synthesized DNA is “nicked”. Therefore, they cannot be used as templates for subsequent amplification compared with regular PCR. This constraint also leads to lower PCR efficiency¹⁷. To circumvent these problems, several modifications were developed, such as using primers containing extended non-overlapping sequences at the 3' end (significantly larger than suggested in Ref. 24) and primer-primer complementary

sequences at the 5' end¹⁷, running two PCR reactions in parallel with each one of the forward and reverse primers²³ and so on. At the present study, we adopted the classical Stratagene strategy, and encountered problems in transformation using DH5 α cells with Stratagene suggested annealing temperatures. After optimizing the annealing temperatures (approximately $T_m^* - 5^\circ\text{C}$), and employing Top10 cells as transformation hosts, the mutagenesis efficiencies were approximately 90% for pcDNA3.1/hPlk2/BamH I insertion, 50% for K111M mutagenesis, 60% for T239D and 90% for T239V mutagenesis, respectively (Table 1, 2). In addition, compared with DH5 α cells, Top10 cells were more applicable for low-efficiency transformation. This was further proved by our ligation and transformation experiments (Table 4).

Ligation and transformation are complicated procedures, and many factors could affect their efficiencies, such as the dephosphorylation of the vectors, the concentrations and ratios of the vector and insert DNA, the amount of DNA used for transformation, and so on. Removing the 5'-phosphate residues from both termini of the vector DNA can efficiently minimize the recircularization of vector DNA (Figure 1), and therefore, decrease the background with empty transformants. In our experiment, a unique BamH I site was used as clone site (Figure 1), and the total percentage of recombinant clones after CIP treatment of the vector DNA was up to about 92.8%. Our data were consistent with other reports^{18, 19}. As a control, without CIP treatment of the vector DNA, the clone efficiency was highly significantly lower than the CIP-treated experiment data ($2\% \pm 4.5\%$ vs $51.3\% \pm 15.2\%$, $P < 0.01$). Some reports recommended that DNA concentration of ligation was about 10ng/ μl , the ratio of insert to vector DNA could be around 1 to 1, and the transformation volume was less than 10% of the competent cells^{10, 25}. Considering these recommendations, we used around 19.0~25.3 ng/ μl DNA, low insert to vector molar ratios (1~3:1) for ligation due to the large size of the vector,

and 2 μ l (about 40-50 ng) volume of the ligation products for transformation (Table 4). We found that these optimizations resulted in high efficiencies for the construction of LVs with monomeric, corrected-oriented inserts (Table 4). When EGFP (about 0.75kb) and hPlk2 WT (about 2.1kb) were used as inserts, their recombinant rates were significantly different (100 ± 0 vs $78\%\pm 2.9\%$, $P<0.05$), this finding suggested that the sizes of the inserts significantly affected the ligation efficiencies. Furthermore, when the recombinant rates were compared among hPlk2 WT, T239D, and T239V mutants, their recombinant rates had significant differences ($78\%\pm 2.9\%$, $100\%\pm 0$, $95\%\pm 11.2\%$, respectively, $P<0.05$), whereas the recombinant rate of K111M had no significant difference with them ($85\%\pm 17.2\%$, $P>0.05$). These data suggested that two or three base pair changes could significantly affect the ligation efficiencies between the vector and inserts, although they were of the same length. This might be because T239V mutant gene was easier to be ligated into the vector with correct orientation than hPlk2 WT gene. In addition, the transformation data also indicated this viewpoint, in which much less transformation colonies of pWPI/hPlk2 WT/Neo were obtained than the others (Table 4).

Here we reported a quantitative model to construct LVs, which circumvented the barriers for efficient subcloning of large-size vectors. Firstly, creation of clone site by SDM could guarantee that 100% of the linearized DNA fragments were with correct cutting-ends. Secondly, dephosphorylation of vector DNA could confirm that most of the transformants (92.8%) were with recombinants. Thirdly, optimization of the amount and ratio of the insert and vector DNA could increase the rate of monomeric, corrected-oriented recombinants ($51.3\%\pm 15.2\%$). Finally, Top10 cells could improve the transformation efficiencies, and therefore facilitate to obtain more colonies for identification. As a result, five bicistronic LVs

were successfully constructed. These vectors were further used to infect HEK293 and SHSY5Y cell lines, neural progenitor cells²⁶, respectively (Data not shown). This study provided an important quantitative model for subcloning large-size, low-efficiency vectors, such as LVs, and furthermore, accelerated the establishment of *in vitro* gain-of-function models for gene function analyses.

Material and methods

Design of SDM primers. Two-step SDM strategies were performed to sequentially insert BamH I sites for plasmids pcDNA3.1/V5-His-Snk/hPlk2²⁷ (pcDNA3.1/hPlk2, Addgene plasmid 16015) and pEGFP-N1 (Clontech) at the 3'-end of hPlk2 WT and EGFP open reading frames, respectively, and then create hPlk2 mutants: K111M, T239D, and T239V with pcDNA3.1/V5-His-Snk/hPlk2-BamH I as template (Figure 1). All primers, including BamH I insertion and mutagenesis of hPlk2 mutants, were designed according to the guide of Stratagene's QuickChangeTM SDM kit, synthesized and purified by Integrated DNA Technologies. For all primers, mutagenized positions were denoted in lower case and underlined.

pcDNA3.1/hPlk2 WT/ BamH I insertion forward:

5'-CATCATCACCATCACCATTGA~~ggatcc~~GTTTAAACCCGCTGATCAGCC-3';

pcDNA3.1/hPlk2 WT/BamH I insertion complement:

5'-GGCTGATCAGCGGGTTTAAAC~~ggatcc~~TCAATGGTGATGGTGATGATG-3';

pcDNA3.1/hPlk2 K111M/BamH I forward:

5'-CAAAGTCTACGCCCAA~~tg~~ATTATTCCTCACAGCAG-3';

pcDNA3.1/hPlk2 K111M/BamH I complement:

5'-CTGCTGTGAGGAATAATcaTTGCGGCGTAGACTTTG-3';

pcDNA3.1/hPlk2 T239D/BamH I forward:

5'-GAACCCTTGGAACACAGAAGGAGAGgacATATGTGGTACCCCAAATTATCTC-3';

pcDNA3.1/hPlk2 T239D/BamH I complement:

5'-GAGATAATTTGGGGTACCACATATgtcTCTCCTTCTGTGTTCCAAGGGTTC-3';

pcDNA3.1/hPlk2 T239V/BamH I forward:

5'-CTAGAACCCTTGGAACACAGAAGGAGAGgtGATATGTGGTACCCCAAATTATCTC
TC-3';

pcDNA3.1/hPlk2 T239V/BamH I complement:

5'-GAGAGATAATTTGGGGTACCACATATCacTCTCCTTCTGTGTTCCAAGGGTTC
G-3';

pEGFP-N1/BamH I insertion forward:

5'-GTACAAGTAAAGCGGCCGCggatccGACTCTAGATCATAATCAG-3';

pEGFP-N1/BamH I insertion complement:

5'-CTGATTATGATCTAGAGTCggatccGCGGCCGCTTTACTTGTAC-3'.

The melting temperatures (T_m , primer-to-template annealing temperature) and primer-pair self-annealing temperatures (T_m^*) were calculated by Stratagene Quikchange Primer T_m Calculator (<http://www.stratagene.com/QPCR/tmCalc.aspx>), and Integrated DNA Technologies SciTools OligoAnalyzer 3.1 (<http://www.idtdna.com/analyzer/applications/oligoanalyzer/default.aspx>), respectively

(Table 1). Hairpin and self-dimer formation of the primers were analyzed by Integrated DNA Technologies SciTools OligoAnalyzer 3.1 as well (Table 1). All the DNA preparation kits, including Miniprep, Maxiprep, and Gel extraction kits, were purchased from QIAGEN. The DNA purifications used in experiments were tested by NanoDrop-1000 Spectrophotometer (NanoDrop Technologies), and all the A260/280 values were ≥ 1.80 .

Mutagenesis. The PCR reactions were carried out with GeneAmp® PCR System 2700 (AB Applied Biosystems). The 50 μ l PCR reaction was carried out with 50 ng templates, 125 ng of each forward and complement primers, 20 μ M of each dNTP (Invitrogen), 2.5 U of PfuUltra DNA polymerase in 1 \times reaction buffer (Stratagene). The thermal cycler program for amplifications was as follows: denaturation at 94°C for 2 min; 18 cycles at 94°C for 30 sec, annealing for 30 sec, at 59°C for pcDNA3.1 BamH I insertion and K111M mutagenesis, 63°C for T239D and T239V mutagenesis, 62°C for pEGFP-N1 BamH I insertion, respectively (Table 2), and at 72°C for 8 min for pcDNA3.1/hPlk2 template, 5 min for pEGFP-N1 template; followed by a final extension at 72°C for 10 min. When the amplifications were finished, 1 μ l (10 U) of Dpn I (Stratagene) was added into each reaction, and incubated at 37°C for 1 hour. Finally, 1 μ l reaction products were used for transformation into 50 μ l DH5 α or Top10 competent cells (Invitrogen), respectively, according to the manufacturer's guides. All the transformants were used to spread plates. Transformation colonies were selected and their plasmids were isolated by miniprep, and the positive mutants were identified by sequencing.

Preparation of vector and insert DNA. The bicistronic LV pWPI (Addgene plasmid 12254) was modified by creating a BamH I site at 3502nt and replacing EGFP sequence with Neo, to form pWPI/Neo/BamH I (a gift from Robert Strome). pWPI/Neo/BamH I DNA was digested with BamH I (NEW ENGLAND BioLabs), then divided into two aliquots, one aliquot of the vector DNA was used directly for ligation, and another aliquot was treated with CIP (NEW ENGLAND BioLabs) to remove the 5'-phosphate groups¹⁰ (Figure 1) as follows: in a 300µl reaction, containing digested pWPI/Neo DNA (about 15 µg), 50 U CIP in 1×NEBuffer 3, at 37°C water bath for 1 hours. Both aliquots of pWPI/Neo vector DNA (CIP-treated and not treated) were purified by 1% Agarose Gel Electrophoresis to remove CIP, and recovered by QIAGEN Gel Extraction Kit. pcDNA3.1/BamH I plasmids carrying hPlk2 WT and mutants K111M, T239D, and T239V, respectively, and plasmid pEGFP-N1/BamH I were digested with BamH I, then the inserts, hPlk2 WT, K111M, T239D, T239V and EGFP, were also recovered by 1% Agarose Gel Electrophoresis and gel extraction procedure, respectively.

Ligation. To achieve the molar ratios of inserts to vector about 1 to 3:1, 5 µl of pWPI/Neo vector DNA and 3.5 µl of each insert DNA were pooled together (Table 4, 5), warmed at 45°C for 5 min, then chilled on ice for 2 min, and then ligated with 0.5 µl (1000 U) high-concentration T4 DNA ligase (NEW ENGLAND BioLabs) in 1×T4 DNA ligase buffer (Figure 1). The 10 µl reaction mixtures were incubated in GeneAmp® PCR System 2700 at 16°C for 16 hours followed by inactivation at 65°C for 10 min, and then set at 4°C until transformation¹⁰.

Transformation and identification. 2 μ l (about 40~50 ng) volumes of the ligation products were used to transform 50 μ l of DH5 α and Top10 competent cells according to the manufacturer's instructions (Table 4). In order to obtain more colonies, all the transformation cells were used to spread plates. Positive colonies were primarily analyzed by Not I (NEW ENGLAND BioLabs) digestion, and further confirmed by DNA sequencing.

Data Statistics. Data were analyzed by mean \pm SD and Student's t-Test.

1. Naldini, L. *et al.* In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263-267 (1996).
2. Cui, Y. *et al.* Targeting transgene expression to antigen-presenting cells derived from lentivirus-transduced engrafting human hematopoietic stem/progenitor cells. *Blood* **99**, 399-408 (2002).
3. Lois, C., Hong, E. J., Pease, S., Brown, E. J. & Baltimore, D. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* **295**, 868-872 (2002).
4. Pfeifer, A., Ikawa, M., Dayn, Y. & Verma, I. M. Transgenesis by lentiviral vectors: lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. *Proc Natl Acad Sci USA* **99**, 2140-2145 (2002).

5. Yu, X. *et al.* Lentiviral vectors with two independent internal promoters transfer high-level expression of multiple transgenes to human hematopoietic stem-progenitor cells. *Mol Ther* **7**, 827-838 (2003).
6. Zufferey, R. *et al.* Self-inactivating lentivirus vector for safe and efficient *in vivo* gene delivery. *J Virol* **72**, 9873-9880 (1998).
7. Szulc, J., Wiznerowicz, M., Sauvain, M. O., Trono, D. & Aebischer, P. A versatile tool for conditional gene expression and knockdown. *Nat Methods* **3**, 109-116 (2006).
8. Hotta, A. *et al.* Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. *Nat Methods* **6**, 370-376 (2009).
9. Sommer, C. A. *et al.* Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* **27**, 543-549 (2009).
10. Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press: Plainview, NY (2001).
11. Jung, V., Pestka, S. B. & Pestka, S. Efficient cloning of PCR generated DNA containing terminal restriction endonuclease recognition sites. *Nucleic Acids Res* **18**, 6156 (1990).
12. Kaufman, D. L. & Evans, G. A. Restriction endonuclease cleavage at the termini of PCR products. *BioTechniques* **9**, 305-306 (1990).
13. Testori, A., Listowsky, I. & Sollitti, P. Direct cloning of unmodified PCR products by exploiting an engineered restriction site. *Gene* **143**, 151-152 (1994).
14. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press: Plainview, NY (1989).

15. Hutchison, C. A. III. *et al.* Mutagenesis at a specific position in a DNA sequence. *J Biol Chem* **253**, 6551-6560 (1978).
16. Norrander, J., Kempe, T. & Messing, J. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. *Gene* **26**, 101-106 (1983).
17. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotech* **8**, 91 (2008).
18. Seeburg, P., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. Nucleotide sequence and amplification in bacteria of structural gene for rat growth hormone. *Nature* **270**, 486-494 (1977).
19. Ullrich, A. *et al.* Rat insulin genes: construction of plasmids containing the coding sequences. *Science* **196**, 1313-1319 (1977).
20. Marumoto, T. *et al.* Development of a novel mouse glioma model using lentiviral vectors. *Nat Med* **15**, 110-116 (2009).
21. Yang, S. *et al.* A method for filling in the cohesive ends of double-stranded DNA using *Pfu* DNA polymerase. *Biotechnol Appl Biochem* **42**, 223-226 (2005).
22. Ben-Dor, I., Itsykson, P., Goldenberg, D., Galun, E. & Reubinoff, B. E. Lentiviral vectors harboring a dual-gene system allow high and homogeneous transgene expression in selected polyclonal human embryonic stem cells. *Mol Ther* **14**, 255-267 (2006).
23. Edelheit, O., hanukoglu, A. & Hanukoglu, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotech* **9**, 61 (2009).

24. Zheng, L., Baumann, U. & Reymond, J. L. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res* **32**, e115 (2004).
25. Bercovich, J. A., Grinstein, S. & Zorzopulos, J. Effect of DNA concentration of recombinant plasmid recovery of blunt-end ligation. *BioTechniques* **12**, 190-193 (1992).
26. Visanji, N. P. *et al.* The effect of S129 phosphorylation on the interaction of α -synuclein with synaptic and cellular membranes. *J Biol Chem* In press.
27. Burns, T. F., Fei, P., Scata, K. A., Dicker, D. T. & El-Deiry, W. S. Silencing of the novel p53 target gene Snk/Plk2 leads to mitotic catastrophe in paclitaxel (taxol)-exposed cells. *Mol Cell Biol* **23**, 5556-5571 (2003).

Acknowledgements

We thank Dr. Didier Trono for the original pWPI vector, Dr. Wafik S. El-Deiry for the original pcDNA3.1/V5-His Snk/hPlk2 vector, Robert Strome for the useful discussions, Yan Liang for the sequencing, Kyle (Hua) Long for the help in editing the figures. We particularly thank Dr. Paulo Lee Ho for critical reading of the manuscript and important suggestions on the manuscript writing. This work was supported by Canadian Institutes of Health Research (CIHR) grant MOP84501 to AT.

Author contributions

G.Z. and A.T. conceived and designed the experiments. G.Z. performed the experiments and carried out data analysis. G.Z. wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.

Tables

Table 1. Characteristics of mutagenesis primer pairs

Primer name	Length/mutation (bases)	T _m (°C)	T _m [*] (°C)	Hairpin formation	Self-Dimer formation
pcDNA3.1/hPlk2 WT/ BamH I insertion forward	48/6	87.9	68.6	No	Yes
pcDNA3.1/hPlk2 WT/BamH I insertion complement	48/6	87.9	68.6	No	Yes
pEGFP-N1/BamH I insertion forward	44/6	86.7	66.7	No	Yes
pEGFP-N1/BamH I insertion complement	44/6	86.7	66.7	No	Yes
pcDNA3.1/hPlk2 K111M forward	36/2	76.6	64.1	No	Yes
pcDNA3.1/hPlk2 K111M complement	36/2	76.6	64.1	No	Yes
pcDNA3.1/hPlk2 T239D forward	51/3	80.9	66.5	No	Yes
pcDNA3.1/hPlk2 T239D complement	51/3	80.9	66.5	No	Yes
pcDNA3.1/hPlk2 T239V forward	56/2	84.2	66.8	No	Yes
pcDNA3.1/hPlk2 T239V complement	56/2	84.2	66.8	No	Yes

Note: T_m: primer-to-template annealing temperature, which considered the mismatches of the bases; T_m^{*}: primer-pair self-annealing temperatures.

Table 2. Mutagenesis efficiency of BamH I insertion and hPlk2 mutants

mutagenesis	Annealing Temperature (°C)	Hosts of transformation	No. of obtained colonies	No. of sequenced colonies	Percentage of positive colonies
pcDNA3.1/BamH I	59	Top10	13 (n=1)	10	90% (9)
pEGFP-N1/BamH I	62	Top10	6 (n=1)	6	83.3% (5)
K111M	59	Top10	~200 (n=1)	10	50% (5)
T239D	63	Top10	~200 (n=1)	10	60% (6)
T239V	63	Top10	19 (n=1)	10	90% (9)

Note. pcDNA3.1/BamH I: pcDNA3.1/hPlk2WT/ BamH I insertion;

pEGFP-N1/BamH I: pEGFP-N1/BamH I insertion;

K111M: pcDNA3.1/hPlk2K111M/BamH I;

T239D: pcDNA3.1/hPlk2T239D/BamH I;

T239V: pcDNA3.1/hPlk2T239V/BamH I.

Table 3. Mutagenesis efficiencies with different annealing temperatures

Mutagenesis	Hosts of transformation	Annealing temperature (°C) & total No. of transformation colonies		No. of identified colonies	Percentages of positive mutagenesis
pcDNA3.1/BamH I	Top10	83/0 (n=1)	70/8 (n=1)	4	100% (4/4)
K111M	Top10	72/0 (n=1)	66/1 (n=1)	1	100% (1/1)
T239D	Top10	76/0 (n=1)	68/0 (n=1)	0	0/0
T239V	Top10	79/0 (n=1)	68/1 (n=1)	1	100% (1/1)

Table 4. Construction efficiencies of lentiviral vectors with CIP-treated vector DNA

Vector	Molar ratio & concentration of inserts/vector	Hosts of transformation	Total No. of transformed colonies	Total No. of identified colonies	Percentage of inserted vectors (Mean±SD)	Percentage of monomeric inserted vectors (Mean±SD)	Percentage of Corrected-oriented inserts (Mean±SD)
EGFP	1.7 : 1 (23.0ng/μl)	Top10/DH5α	~200 (n=1)	20 (n=4)	100%±0^{a,b}	95%±10%^{a,b}	51%±16.5%^a (10)
			/2 ^a (n=3)	/2 (n=1)	(20)/100% (2)	(19)/100% (2)	/0 (0)
hPlk2	1.8 : 1 (21.4 ng/μl)	Top10/DH5α	19 (n=1)	14 (n=3)	78%±2.9%^{a,b}	78%±2.9%^{a,b}	37%±15.2%^{a,b} (5)
			/1 ^a (n=3)	/1 (n=1)	(11)/100% (1)	(11)/100% (1)	/100% (1)
K111M	3.2 : 1 (25.3 ng/μl)	Top10/DH5α	~100 (n=1)	14 (n=4)	85%±17.2%^a	85%±17.2%^a	52%±33.6%^a (7)
			/3 ^a (n=3)	/3 (n=1)	(12)/100% (3)	(12)/100% (3)	/66.7% (2)
T239D	2 : 1 (21.8 ng/μl)	Top10/DH5α	~200 (n=1)	18 (n=5)	100%±0^{a,b}	83%±15.6%^a	45%±16.2%^a (8)
			/7 ^a (n=3)	/6 (n=1)	(18)/100% (6)	(15)/100% (6)	/83.3% (5)
T239V	1 : 1 (19.0 ng/μl)	Top10/DH5α	~100 (n=1)	17 (n=5)	95%±11.2%^{a,b}	95%±11.2%^{a,b}	75%±25%^{a,b} (13)
			/4 ^a (n=3)	/4 (n=1)	(16)/100% (4)	(16)/100% (4)	/100% (4)

Note. Data in boldfaces are obtained from Top10 cell transformation

Percentage of inserted vectors=No. of inserted vectors/total No. of identified colonies;

Percentage of monomeric inserted vectors=No. of monomeric inserted vectors/total No. of identified colonies;

Percentage of positive colonies=No. of vectors with monomeric correct-oriented insert/total No. of identified colonies.

^a Values in the same column indicates no significant difference ($P>0.05$);

^bValues in the same column have significant difference ($P < 0.05$).

EGFP: pWPI/EGFP/Neo; hPLK2WT: pWPI/hPlk2WT/Neo; K111M:

pWPI/hPlk2K111M/Neo; T239D: pWPI/hPlk2T239D/Neo; T239V: pWPI/hPlk2T239V/Neo.

Table 5. Construction efficiencies of lentiviral vectors with vector DNA untreated by CIP

Vector	Molar ratio & concentration of inserts/vector	Hosts of transformation	Total No. of identified colonies	Percentage of inserted vectors	Percentage of monomeric inserted vectors	Percentage of Corrected-oriented inserts
EGFP	3.7 : 1 (21.6 ng/μl)	Top10	10	0% (0/10)	0% (0/10)	0% (0/10)
hPlk2	1.1 : 1 (21.0 ng/μl)	Top10	10	10% (1/10)	10% (1/10)	0% (0/10)
WT						
K111M	1.1 : 1 (20.9 ng/μl)	Top10	10	10% (1/10)	10% (1/10)	0% (0/10)
T239D	2.1 : 1 (24.1 ng/μl)	Top10	10	0% (0/10)	0% (0/10)	0% (0/10)
T239V	2.7 : 1 (26.1 ng/μl)	Top10	10	30% (3/10)	30% (3/10)	10% (1/10)

Figure Legends

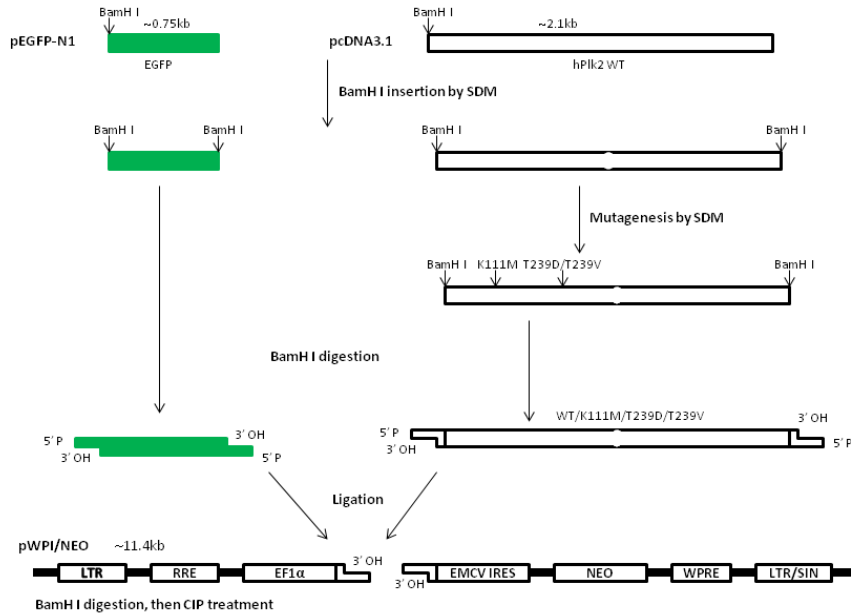


Figure 1. Schematic representation of the LV construction strategy. A BamH I clone site was inserted at the 3'-ends of EGFP and hPlk2 WT gene by SDM, respectively. Then, K111M, T239D and T239V mutants were created through SDM using hPlk2 WT gene as template. And then, the inserts were digested with BamH I, and purified, at the same time, pWPI vector was digested by BamH I, and treated with CIP to protect the self-circularization of the vector DNA. Finally, LVs were constructed through ligation and transformation.

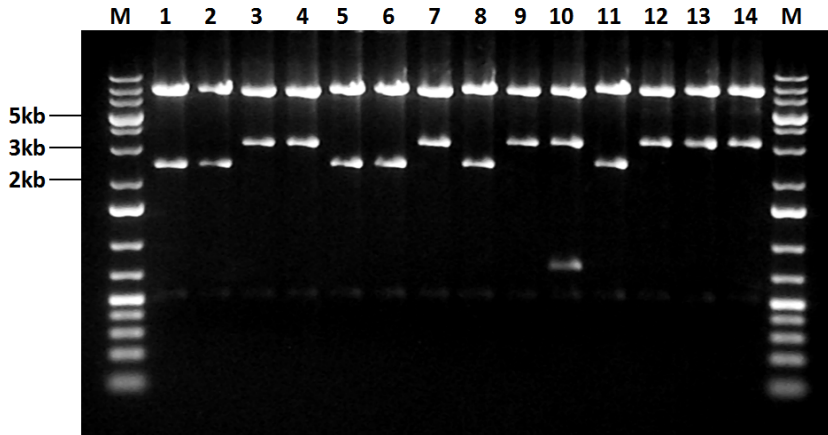


Figure 2. Agarose gel electrophoresis for identification of pWPI/EGFP/Neo digested by Not I. Lane M: GeneRuler™ 1kb DNA Ladder Plus (Fermentas). Lanes 1-14: Colonies 1-14 of pWPI/EGFP/Neo digested with Not I, among them colonies #3, 4, 7, 9, 12, 13, 14, were positive with correct orientation; colonies # 1, 2, 5, 6, 8, 11, were negative with opposite orientation; colony #10 was with two copies of EGFP gene.

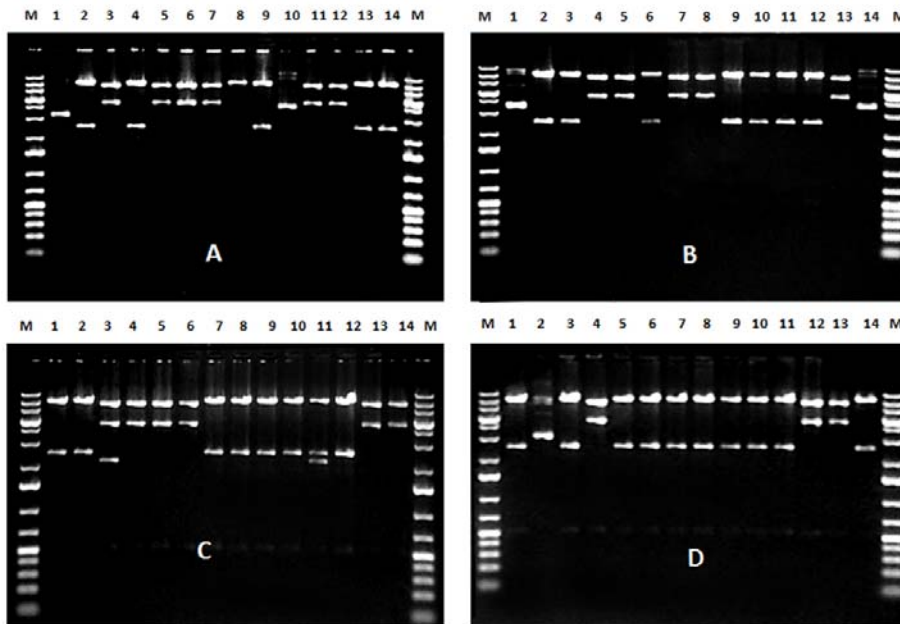


Figure 3. Agarose gel electrophoresis for identification of bicistronic lentiviral vectors carrying hPlk2 WT and mutants digested by Not I.

A: pWPI/hPlk2WT/Neo digested by Not I. Lane M: GeneRuler™ 1kb DNA Ladder Plus.

Lanes 1-14: Colonies 1-14 of pWPI/hPlk2WT/Neo digested with Not I, among them colonies #2, 4, 9, 13, 14, were positive with correct orientation; colonies #3, 5, 6, 7, 11, 12, were negative with opposite orientation; colony #8 was empty vector; colonies #1, 10 were unknown DNAs.

B: pWPI/hPlk2K111M/Neo digested by Not I. Lane M: GeneRuler™ 1kb DNA Ladder Plus.

Lanes 1-14: Colonies 1-14 of pWPI/hPlk2K111M/Neo digested with Not I, among them colonies #2, 3, 6, 9, 10, 11, 12, were positive with correct orientation; colonies #4, 5, 7, 8, 13, were negative with opposite orientation; colonies #1, 14 were unknown DNAs.

C: pWPI/hPlk2T239D/Neo digested by Not I. Lane M: GeneRuler™ 1kb DNA Ladder Plus.

Lanes 1-14: Colonies 1-14 of pWPI/hPlk2T239D/Neo digested with Not I, among them colonies #1, 2, 7, 8, 9, 10, 12, were positive with correct orientation; colonies #4, 5, 6, 13, 14, were negative with opposite orientation; colonies #3, 11 were with two copies of hPlk2T239D genes.

D: pWPI/hPlk2T239V/Neo digested by Not I. Lane M: GeneRuler™ 1kb DNA Ladder Plus.

Lanes 1-14: Colonies 1-14 of pWPI/hPlk2T239V/Neo digested with Not I, among them colonies #1, 3, 5, 6, 7, 8, 9, 10, 11, 14, were positive with correct orientation; colonies #4, 12, 13, were negative with opposite orientation; colonies #2 was unknown DNA.