Published in final edited form as:

Stem Cells. 2009 May ; 27(5): 1032–1041. doi:10.1002/stem.38.

Cre Recombination Mediated Cassette Exchange For Building Versatile Transgenic Human ESC lines

Zhong-Wei Du¹, Bao-Yang Hu¹, Melvin Ayala¹, Brian Sauer², and Su-Chun Zhang^{1,*}

¹Departments of Anatomy and Neurology, School of Medicine and Public Health, Waisman Center, WiCell Institute, University of Wisconsin, Madison, WI, USA

²Stowers Institute for Medical Research, Kansas City, MO, USA

Abstract

To circumvent the silencing effect of transgene expression in human embryonic stem cells (hESC), we employed the Cre recombination mediated cassette exchange strategy to target the silencing-resistant site in the genome. We have identified new loci that sustain transgene expression during stem cell expansion and differentiation to cells representing the three germ layers in vitro and in vivo. The built-in double loxP cassette in the established master hESC lines was specifically replaced by a targeting vector containing the same loxP sites, using the cell-permeable Cre protein transduction method, resulting in successful generation of new hESC lines with constitutive functional gene expression, inducible transgene expression and lineage-specific reporter gene expression. This strategy and the master cell lines allow for rapid production of transgenic hESC lines in ordinary laboratories.

Keywords

Cre-loxP system; human embryonic stem cells; transgene expression; differentiation; RMCE

Introduction

Human embryonic stem cells (hESCs) offer an invaluable tool for revealing human biology and a potential source of functional cells/tissues for regenerative medicine [1,2]. Like their mouse counterparts which revolutionize biomedical research through transgenesis for the past decades [3,4], the utility of hESCs including the potential medical application, will likely be significantly enhanced and broadened by our ability to build versatile genetically modified hESC lines [5,6].

Building stable transgenic hESC lines remains a challenging and laborious process. The chief reasons include low transfection and cloning efficiency [7,8] as well as high incidence of transgene silencing caused by the integration site and following cellular differentiation [9, 10]. At present, lentivirus mediated approach can achieve as high as 40–70% transfection efficiency in hESCs [11,12]. However, its preference of integrating into the gene coding region,

©AlphaMed Press

*Corresponding author: Su-Chun Zhang, MD, PhD, Waisman Center, Rm 651, University of Wisconsin, 1500 Highland Avenue, Madison, WI 53705, Phone: (608) 265-2543, Fax: (608) 263-5267, zhang@waisman.wisc.edu.

Author contributions: Z.-W. D.: conception and design, collection and/or assembly of data, data analysis and interpretation, manuscript writing; B.-Y. H. and M. A.: collection and/or assembly of data; B. S.: provision of study material; S.-C. Z.: financial support, conception and design, manuscript writing, final approval of manuscript.

See www.StemCells.com for supporting information available online.

which poses a risk of insertional mutagenesis [13,14], and the limitation of inserting DNA size make the method less suitable for generating stable hESC clones.

Several strategies have been explored to circumvent the silencing effect of the integration site. The virus polynucleotide enhancer sequence PyF101 has been found to be resistant to silencing when it is located 5' to the CAG promoter and the transgene (GFP) retains expression in hESCs for over 120 passages [9]. Nevertheless, it is not shown if the GFP expression is sustained following cellular differentiation and if the PyF101 sequence is also effective for other promoters. Transgenes are introduced to some unique sites, such as the AAVS1 locus by adeno-associated virus type 2 (AAV2) [15] and the pseudo-attP sites by phiC31 integrase [16]. However, the silence-resistant effect of these sites has not been studied in detail. The targeting efficiency of AAV2 is low (4.16%) and hESCs possess 23 different pseudo-attP sites, making it extremely difficult to screen the right and stable cell clones. Several silence-resistant sites have been identified in the mouse genome, including ROSA26 [17], HPRT1 [18] and ColA1 [19]. Incorporation of a single copy of transgene expression cassette into these sites by a precise recombination mediated cassette exchange (RMCE) method has been demonstrated to be the best way in establishing transgenic cell lines and mice [20–23]. Irion and colleagues inserted GFP under the ROSA26 promoter in hESCs through homologous recombination although they did not show if the ROSA26 promoter sustains GFP expression in terminally differentiated cells [24]. In rodents, ROSA26 promoter is insufficient to drive transgene (GFP) expression for direct visualization in brain tissue [25]. Neurons derived from transgenic mice and rats or mouse ESCs with transgene (GFP or alkaline phosphatase) inserted into the Rosa26 locus often do not exhibit transgene expression (Zhang, unpublished studies). Therefore, identification of an appropriate site for stable transgene expression not only in hESCs but also in their differentiated progenies remains to be solved.

We sought to screen integration sites that are resistant to transgene silencing during hESC expansion and differentiation, especially in neural differentiation. In the screening vector, however, we built in a double loxP recombination exchange cassette, described by Sauer et al. [20], so that the selected clone can serve as a master cell line. Replacement of the built-in loxP cassette with any targeting transgene cassette, possessing the same loxP sites through RMCE, allowed the generation of versatile transgenic hESC lines. Newly developed cell permeable Cre protein transduction method [26] improved the efficiency of Cre-mediated recombination in hESCs. This technology and the master cell lines will make the establishment of transgenic hESC lines a laboratory routine.

Materials and Methods

Maintenance and differentiation of hESCs

Human ESC lines, H9 (NIH Code WA09, passages 17 to 45) was cultured and passaged every 6 days on a feeder layer of irradiated embryonic mouse fibroblasts as described [1]. Differentiated colonies were physically removed before passaging and the undifferentiated state of ESCs was confirmed by typical morphology and uniform expression of Oct4 and SSEA4. The established master hESC lines were examined by karyotype analysis (WiCell Institute/ National Stem Cell Bank protocols, <http://www.wicell.org/>) to make sure that the genetically modified ESCs retain the normal genetic background. The procedure for differentiation to neural precursors and motor neurons from hESCs was essentially the same as described [27,28]. The procedures for endoderm and mesoderm differentiation were from the published protocol [29].

Vector construction

The double loxP containing vector pLox was constructed by replacing the lox511 site in pSS66 plasmid [20] with lox2272 site by PCR cloning. The antibiotics resistant cassette PGK-neo and PGK-hygromycin from pBS524 and pBS528 was inserted into the EcoRV site of pLox to make the pLox-neo and pLox-hyg vector. To construct the master vector, CAG-hrGFP cassette was inserted into the EcoRV site of pLox-neo vector. For RFP targeting vector, CAG-RFP cassette was inserted into the EcoRV site of pLox-hyg vector. For the Olig2 targeting vector, the Olig2-FLAG cDNA (gift from Prof. Nakafuku) was used to replace the hrGFP in CAG-hrGFP, and then the CAG-Olig2-FLAG was inserted into the EcoRV site of pLox-hyg vector. For Syn-GFP targeting vector, Syn-GFP cassette from pMH4-I-SYN-EGFP (gift from Prof. Kugler) was cloned into the EcoRV site of pLox-hyg vector. For the inducible GFP targeting vector, the tetO sequence from pLVTHM (Addgene plasmid 12247) [30] was inserted into the EcoRI and BamHI sites of pLox-hyg vector, and the GFP-IRES-tTRKRAB fragment from pLVCT-tTRKRAB (Addgene plasmid 11643) [31] replaced hrGFP in CAG-hrGFP cassette, and then was inserted into the EcoRV site. The direction of the insertion was determined by restriction enzyme analysis.

Transfection of hESCs by electroporation

Human ESC colonies were detached with dispase (1 mg/ml; Invitrogen) treatment for 3 min, washed with the ESC culture medium, and resuspended in 0.6 ml cold culture medium ($1-2 \times 10^7$ cells/one well of a 6-well plate). Linearized master vector DNA (30 ug in 0.1 ml of PBS) was mixed with ESCs using a 1-ml pipette tip. Cells were then exposed to a single pulse (320 V, 200 μ F) using the BioRad Gene Pulser Xcell (0.4 cm gap cuvette; BioRad, Hercules, CA). The electroporated cells were incubated at room temperature for 5 min before they were plated and cultured under the regular hESC growth medium in three 6-well plates with MEF feeder layer. G418 selection (50 ug/ml, Invitrogen) was started 3 days after electroporation and the G418 concentration was increased to 100 ug/ml after one week. After two weeks, surviving colonies were picked out and expanded in each well of 24-well plates.

Flow cytometry analysis

ESCs were dissociated into single cells with trypsin-EDTA for 5 min. The cells were incubated with SSEA-4 antibody (mIgG3, Chemicon) or mIgG (as a negative control for SSEA-4 staining) for 45 min at 37°C, followed by washing with PBS three times and incubation with the secondary antibody Phycoerythrin (PE)-conjugated goat anti-mouse IgG (BD Bioscience). After washing the secondary antibody reaction, cell samples were analyzed using a FACScan flow cytometer (BD Bioscience). Dead cells were excluded from analysis by forward- and side-scatter gating. The mIgG stained cells were set up as a negative control for SSEA-4 analysis and the parental H9 cells were used as a negative control for GFP analysis. A minimum of 50,000 events was acquired for each sample. The data were analyzed with Cellquest software (BD Bioscience).

Quantification of transgene copy numbers by real-time PCR analysis

The genomic DNA was extracted from every hESC clone by MasterPure™ DNA purification kit (Epicentre). Real-time PCR was performed using the Bio-Rad MyiQ real-time PCR detection system. The reaction was conducted under the following condition: template denaturation at 95 °C for 3 min, followed by 40 cycles of denaturation at 95°C for 15s, annealing at 55°C for 30s, and extension at 72°C for 1 min. The primers were used for hrGFP: Forward CGGCTCTGCTTCCCTTAGACT; Reverse TCACAGCCAAGCATTCTACAAAC, for GAPDH: Forward AACGTGTCAGTGGTGGACCTG; Reverse AGTGGGTGTCGCTGTTGAAGT. The PCR efficiency was examined with five dilutions of hESC genomic DNA and the specificity of individual gene primers was validated by the

melting curve at the end of each PCR assay. To determine the integrating transgene copy number, the copy number of hrGFP gene was first determined as one gene copy by relating the C_T value to the standard curve in the control cell line. The copy number of the hrGFP transgene in different cell clones was then compared with the control cell line and quantified by normalization against the reference GAPDH gene using the $2^{-\Delta\Delta C_T}$ method.

Immunostaining and microscopy

Immunohistochemical staining was performed according to Zhang et al [27]. Coverslip cultures were fixed in 4% paraformaldehyde in PBS for 15 min at room temperature. After washing with PBS, cells were incubated for 30 minutes in 10% normal goat serum to block non-specific antibody binding. They were then incubated for 1 hour in the primary antibody, washed in PBS and incubated for 1 hour in Alexa Fluor 488 goat anti-mouse IgG or Alexa Fluor 594 goat anti-rabbit IgG second antibodies (1:1000, Molecular Probes) and for 5 min in bisbenzamide (5 ng/ml; Hoechst No. 33342; Sigma). Coverslips were mounted in a mounting medium (Immunotech) and examined with a Nikon Eclipse E600 fluorescence microscope. The following primary antibodies were used: nestin (mIgG 1:200, Chemicon), Olig2 (rIgG 1:4000), β III tubulin (TuJ1, mIgG 1:500, Sigma), GFAP (mIgG 1:500, Chemicon), HB9 (mIgG 1:50, DSHB), synapsin I (rIgG 1:1000, Calbiochem), Sox17 (mIgG 1:100, R&D), Pdx1 (rIgG 1:500, Abcam), Brachyury T (gIgG 1:50, R&D), cTnT (mIgG 1:500, Abcam). Quantification of HB9⁺ motoneuron was performed by averaging counts of fifteen neural cell aggregates from three independent experiments using Metamorph software (Universal Imaging Corporation).

Formation of Teratomas hESCs were injected subcutaneously into severe combined immunodeficient (SCID) mice (Jackson Laboratory, Bar Harbor, ME) on the back. Two months after injection, mice were sacrificed and teratomas were removed. The teratomas were then postfixed with 4% paraformaldehyde in PBS, followed by cryoprotection in 30% sucrose overnight and cryosectioned to 25 μ m. Histological examination of the sections was performed by hematoxylin and eosin staining. All animal experiments were performed following protocols approved by Institutional Animal Care and Use Committee.

Site-finding PCR

The transgene integration site was determined by site-finding PCR analysis [32]. The site-finding PCR mixture included 2 μ l of 10 $\times$ long Taq DNA polymerase buffer, 2 μ l of 10mM dNTP solution, 0.5 U of long Taq DNA polymerase (Epicenter), 10 pmol of SiteFinder primer (5' - CAC GAC ACG CTA CTC AAC ACA CCA CCT CGC ACA GCG TCCT CAA GCG GCC GCN NNN NNG CCT - 3') and 100 ng of genomic DNA extracted from hESC samples. The final volume was brought to 20 μ l with water, and then a single PCR cycle was run (Table 1). This was followed by two rounds of nested PCR (Table1) using the following primers, respectively, SFP1 (5' - CAC GAC ACG CTA CTC AAC AC - 3'), GSP1 (5' - TAT CCC GTA TTG ACG CCG GGC AAG - 3'), SFP2 (5' - ACT CAA CAC ACC ACC TCG CAC AGC - 3'), GSP2 (5' - AAG AGC AAC TCG GTC GCC GCA TAC - 3'). SFP1 and SFP2 were nested primer from SiteFinder primer. GSP1 and GSP2 were designed from the backbone of the master vector. The products were separated on 1.0% agarose gels. The PCR products were isolated and cloned into the pBluescript SK vector between the restriction sites of NotI and EcoRV. The positive clones were selected for sequencing. DNA sequence analyses were carried out using the BLAST program (<http://ncbi.nlm.nih.gov>) in the human genome sequence.

Purification of TAT-Cre protein

TAT-Cre fusion proteins were expressed from the plasmid pTAT-Cre (gift from Prof. Dowdy) [33]. The proteins were expressed in E. coli strain BL21 (DE3). Bacterial cultures (1L), grown to an A600 of 0.6–1.0, and induced with 0.5 mM IPTG. After harvesting, the bacterial cells were lysed in 50ml Ni-NTA lysis buffer (50 mM sodium phosphate and 300 mM NaCl, pH

8.0) containing 50ul Benzonase and centrifuged. The clear lysate was incubated with Ni-NTA His-Bind Resins (Novagen) for 1 hr and packed into a gravity-flow column. After affinity chromatography, recombinant proteins were eluted and dialyzed in PBS supplemented with 0.3 M NaCl and 8% glycerol. The protein concentration was determined with BCA reagent (Pierce).

Cre RMCE with the targeting vector

The hESCs were dissociated with dispase and split 1:3 in the 6-well plate (3×10^6 cells/well). After culturing for 24 hrs, the ESC medium was replaced with fresh medium. Targeting vector DNA (2 ug) and Eugene®HD reagent (6 ul) were mixed in 100-ul OPTIMEM (Invitrogen) for 15 min and then applied directly into the cell culture. Five hrs later, 2uM of TAT-Cre protein was added in the culture. The cells were fed with fresh ESC medium the next day and cultured for additional three days before hygromycin B (25ug/ml) was added to select antibiotics-resistant colonies for two weeks. The surviving colonies were picked out and expanded in 24-well plates. The correct exchanged cell clones were identified by PCR.

Southern blot analysis

20ug genomic DNA was extracted and digested with BamHI, then run in 0.8% agarose gels and blotted onto the nylon membrane (Pierce). ^{32}P -CTP labeled CAG promoter probe was used to detect the integration of the targeting vector.

Results

1. Silencing-resistant hESC lines are selected with a built-in double loxP exchange cassette

We have shown previously that transgenes are severely suppressed immediately following infection by lentivirus in hESCs [10]. In this study, transfecting hESCs with the plasmid containing hrGFP (humanized renilla green fluorescent protein) driven by the CAG (cytomegalovirus immediate early enhancer / chicken beta-actin promoter) promoter and neomycin resistance gene driven by PGK promoter resulted in dozens of stable neo-resistant cell clones. However, only 17 of the total 98 (17.3%) clones expressed GFP (three independent experiments), and some clones showed variegated GFP expression (Supplementary Fig. 1). Parallel experiments using the same construct to transfect mouse ESCs indicated that 211 of 254 (83.1%) cell clones expressed GFP (Table 1). This result suggests that the expression of transgenes randomly integrated into the human ESC genome will be largely silenced.

To circumvent the silencing effect of integration sites on transgene expression in hESCs, we screened silencing-resistant sites using a vector with a built-in double loxP sites. The strategy consists of two steps (Fig. 1A). In the first step, the expression cassettes of hrGFP gene driven by CAG promoter and the neomycin resistance gene driven by PGK promoter are flanked by loxP and lox2272 sites to make the master vector. The wild type loxP and mutant lox2272 sites are designed not to recombine with each other, but they can undergo recombination with the same site in the targeting vector. The master vector is transfected into hESCs, and the hESC clones, which show ubiquitous and stable GFP expression during expansion and differentiation, will be selected as a master cell line. In the second step, a new transgene driven by CAG (or any other) promoter and the hygromycin resistance gene driven by the PGK promoter are flanked by loxP and lox2272 sites to make the targeting vector. When co-transfecting with Cre recombinase into the master hESCs, the targeting vector can replace the master vector at the same integration site, which guarantees the silence-resistant expression of the new transgene as that of GFP.

Linearized master vector was transfected into the hESCs (H9 cell line) by electroporation. After selecting with neomycin for two weeks, 228 resistant cell clones were produced from six

transfection experiments. 17 clones were selected with uniform GFP expression in hESCs through regular passaging (Fig. 1B, C). FACS analyses with the hESC marker SSEA-4 indicated that 97%–99% of the SSEA-4 positive hESCs showed GFP expression (Fig. 1D, E), confirming the uniform expression pattern of GFP and the stem cell state of the transgenic cells.

For effective Cre-RMCE, it is ideal to have only one integration copy of the master vector in the genome of the master hESC line. We applied the qPCR method to determine the copy numbers of the hrGFP gene relative to the control GAPDH gene. Among the 17 clones, 8 clones had one integration copy, another 8 clones had two integration copies, and one clone had about one and half integrating copies (Supplementary Table 1).

2. The master hESC line is selected to sustain GFP expression during differentiation

A master hESC line should possess stable transgene (GFP) expression not only during stem cell expansion but also along differentiation into cells of the three germ layers. The above 8 clones with one integration copy were first differentiated to neuroectodermal cells in our chemically defined system [27]. Two clones showed uniform GFP expression in the hESC aggregates at day 6 and in the neuroepithelial cells in the form of rosettes at day 12 following hESC differentiation (Fig. 2A, B). In additional 4 and 8 weeks of differentiation, the majority of TuJ1 positive neurons and GFAP positive astrocytes retained GFP expression (Fig. 2C–F). When the two hESC clones were differentiated toward the mesoderm cell lineages, GFP expression was sustained in Brachyury T⁺ progenitor cells at day 3 and cTnT⁺ (cardiac troponin T) beating cardiac muscle cells at day 17 (Fig. 2I, J). Similarly, GFP expression was sustained in Sox17⁺ endoderm progenitor cells and Pdx1⁺ endoderm-derived pancreatic cells following 6 days and 17 days of hESC differentiation, respectively (Fig. 2M, N). Analysis of the teratomas formed from the two cell lines 2 months after injection showed ubiquitous and stable GFP expression in the neuroectoderm (Fig. 2G, H), mesoderm derived muscle and cartilage (Fig. 2K, L), and endoderm derived epithelial tube (Fig. 2O, P). These results confirm that GFP expression in the two hESC lines is not silenced following differentiation to cell types representing the three germ layers both in vitro and in vivo.

Besides resistance to gene silencing, the integration site of the master vector should not interfere with the expression of endogenous genes. Site-finding-PCR analysis [32] indicated that in one clone the master vector was integrated in chromosome 4, between the gene loci IGFBP7 and LOC255130, and in the other clone the master vector was integrated in chromosome 7, between the gene loci CLDN3 and CLDN4. Based on DNA sequence blast analyses, these two loci are not in the gene coding region or known gene regulatory elements. These two hESC lines sustained GFP expression and retained normal karyotype after expanding for more than 40 passages (Supplementary Fig. 2). Therefore, they were finally selected as the master hESC lines for later RMCE experiments.

3. Specific cassette exchange is mediated by cell permeable Cre protein

To test whether the double loxP cassette in the master hESCs can be exchanged specifically, we constructed a double loxP targeting vector with RFP (red fluorescent protein) expression driven by CAG promoter and hygromycin resistant gene driven by PGK promoter (Fig. 3A). The targeting vector and the Cre expression vector were co-transfected into the master hESC lines by the lipofection or electroporation method, both of which were successfully used in the RMCE experiments of mouse ESCs [20,34]. After selecting with hygromycin B for two weeks, most of the produced clones were still GFP positive. Few of the clones (2 out of 92 clones from three transfections) showed loss of GFP, indicating that site-specific recombination to replace the master GFP cassette with the RFP expression cassette is not efficient. This is likely due to

the overall low transfection efficiency for both vectors in hESCs, leading to the rare coexistence of the Cre recombinase and the targeting vector.

RMCE in hESCs may be significantly enhanced by applying the cell permeable Cre protein directly to the cells [26]. TAT-Cre, a fusion protein combining the protein translocation peptide derived from HIV-TAT with Cre recombinase, was expressed and purified in a bacteria expression system (Supplementary Fig. 3). The targeting vector for RFP expression was first transfected into the master hESCs by Eugene®HD reagent, and then the TAT-Cre protein was added into the culture medium 5 hours later. Under this condition, an average of 18 hygromycin resistant clones in one transfection experiment (three 30mm-wells) were produced after 2 weeks of selection. Six clones showed loss of GFP and gain of RFP expression (Fig. 3B, C), a sign of correct RMCE in the master hESCs. Two clones showed the loss of GFP but without the RFP, suggesting non-specific recombination. The remaining clones showed co-expression of GFP and RFP, suggesting that the cassette was un-exchanged and the RFP targeting vector randomly integrated in the master hESCs (Fig. 3D, E). PCR analysis using primers located on the targeting vector and the integrating locus showed the specific band for the exchanged cell clones (Fig. 3F), which will be used to identify the correctly exchanged cell clones in the later RMCE experiments. Southern blotting analysis indicated that the exchanged cell clones and the master cell line exhibited the same band whereas the un-exchanged cell clones displayed extra restriction fragments (Fig. 3G), confirming the specificity of the RMCE.

4. Versatile transgenic hESC lines are established via RMCE

(1) Human ESC lines with constitutive expression of a functional gene—To test whether a transgene in the targeted hESCs will be functional, we chose to express Olig2, a transcription factor that is critical for the differentiation of spinal motor neurons from ESCs [28,35]. The Olig2 targeting vector (Fig. 4A) was transfected into the master cell line to replace the master vector. Two hESC clones were established following the procedure described above. Following 14 days of neural differentiation, all the differentiated neural precursor cells in the rosette structure were positively stained for FLAG (Fig. 4B). Immunostaining for Olig2 confirmed that all cells, including the nestin⁺ neural precursor cells, expressed Olig2. In contrast, neural precursors differentiated from the parental H9 hESCs did not express Olig2 under this condition (Fig. 4C, D).

To determine if the transgenic Olig2 is functional, we differentiated the transgenic and parental H9 hESCs to spinal motor neurons using our established protocol [28]. Transcription of Olig2 is essential for specification of motor neuron progenitors, but sustained high-level expression of Olig2 represses the generation of post-mitotic motor neurons marked by HB9 expression. Similar to our previous finding, HB9-expressing postmitotic motor neurons began to appear after 4 weeks of differentiation in the presence of retinoic acid and sonic hedgehog in the parental hESCs, and reached a peak at the 5th week in which about 40% of total cells were positive for HB9 (Fig. 4E). However, in the Olig2 transgenic cell lines, less than 5% of the differentiated cells showed expression of HB9 and these HB9 positive cells also showed the down-regulation or the loss of the expression of transgenic Olig2 (Fig. 4F, G). This result is consistent with the finding in mouse experiments that sustained high-level expression of Olig2 blocks the generation of the HB9⁺ post-mitotic motor neurons [36]. This result suggests that forced Olig2 expression in the hESCs, established through RMCE from the master cell line, bears functional consequence.

(2) Human ESC lines with inducible GFP expression—We explored the possibility to build hESCs with inducible transgene (GFP) expression by exchanging the constitutive expression vector in the master line with an inducible GFP expression vector. The single-vector strategy described by Szulc et al. [31] was applied, which takes advantage of the promiscuous

repression activity of tTRKRAB, a fusion protein of the Kruppel-related box (KRAB) domain and the tetracycline repressor (tetR), which represses the activity of promoters located within the 2–3kb region of the tet operator (TetO) sequence and can be regulated by tetracycline. The inducible targeting vector (Fig. 5A) was transfected into the master hESC lines to replace the master vector with the treatment of TAT-Cre protein. After selection with hygromycin B for two weeks, we picked out two cell clones with negative GFP expression. When treated with doxycycline (Dox, 1ug/ml), GFP was observed uniformly in the hESC colonies 48 hours later (Fig. 5B, C), indicating inducible GFP expression.

To determine whether GFP expression level may be regulated in a dose-dependent manner, the inducible ESC lines were treated with different doxycycline concentrations (2, 1, 0.2, 0.04, 0.008, and 0.002 ug/ml). FACS analysis of the GFP intensity 3 days after Dox treatment indicated that the GFP expression was dose-dependent within the range of 0.002-1 ug/ml of Dox (Fig. 5D). To determine the kinetics of GFP expression in response to doxycycline, the hESC cultures were treated with the optimal doxycycline concentration (1 ug/ml) for 1, 2, 3, and 4 days. FACS analysis showed that the maximum level of GFP was reached in nearly all the hESCs by Day 2. At day 4, when the GFP expression was stable, doxycycline was then withdrawn from the cultures. Daily FACS measurements indicated that the GFP disappeared completely in the hESCs after 5 days of Dox withdrawal (Fig. 5E).

(3) Human ESC lines with neuron specific GFP expression—In practice, it is often necessary to restrict the transgene expression in a particular cell type. This may be achieved through the use of cell type-specific promoters, such as synapsin for neurons. In principle, the silence-resistant sites we identified may also function for other transgene expression such as cell type-specific promoters. To provide proof-of-principle, we constructed a targeting loxP vector by replacing the CAG promoter in the master vector with the synapsin promoter (Fig. 6A). In this way, the GFP will be turned on only when the hESCs differentiate to synapsin-expressing neurons. The hESCs can also serve simply as a neuronal reporter line.

Two different Syn-GFP cell lines were established from the master cell lines after Cre mediated exchange with the master vector described above. The newly established hESCs did not exhibit GFP. When induced for differentiation along the neural lineage, GFP was not observed in the ESC aggregate stage at day 6 or at the nestin-expressing neuroepithelial stage at day 14 (Fig. 6B, C). GFP began to appear after the expanded neuroepithelial cells were differentiated to neurons in the serum-free medium at day 28. Along further differentiation, GFP expression was seen in synapsin⁺ neurons (Fig. 6D, E). Thus, the GFP is specifically expressed in the maturing neurons.

Discussion

At the present study we have identified at least two sites that allow stable transgene expression not only during the expansion of hESCs, but also in their differentiated progenies representing the three germ layers in vitro and in vivo. To enable the making of transgenic hESC lines in an ordinary laboratory, we have built master hESC lines that carry a unique double loxP exchange cassette. By replacing the loxP cassette with a targeting vector through Cre RMCE, we have built transgenic hESC lines expressing a functional gene, conditionally expressing a transgene, or incorporating a reporter gene in a cell lineage specific pattern. These master hESC lines with a built-in double loxP cassette, together with the permeable Cre protein-mediated recombination demonstrated here, shall significantly facilitate the generation of transgenic hESC lines in a substantially shorter period.

Stability of transgene expression is highly dependent on the site of integration [18,37]. Several silence resistant sites have been identified in transgenic mouse studies, such as ROSA26 and

HPRT1. Although these loci are conserved between human and mouse, the silencing resistant effect on exogenous promoters inserted into the ROSA26 and HPRT1 loci has never been examined in human. Costa et al. identified one “Envy” site on chromosome 7 to retain the robust GFP expression in hESCs and differentiated progeny using a similar construct as our master vector [38], suggesting the possibility of identifying the new silence-resistant site in hESCs. However, they didn't build the recombination mediated cassette exchange system in the “Envy” site to express other transgenes. The two sites we identified through our laborious screening appear to sustain functional transgene expression along differentiation. Genomic blast analyses suggest that these two sites will unlikely affect endogenous gene expression. Our studies indeed show no observable changes in the growth, maintenance, directed in vitro differentiation, and teratoma formation in vivo. While we have shown that these sites sustain expression of GFP and some functional genes in differentiated neural cells, more extensive analysis is needed to verify whether these sites can serve as universal sites for transgene expression in functional human cells of other lineages.

With a known site for stable transgene expression, one may build transgenic hESC lines through homologous recombination. However, site-directed integration of transgenes into these sites via homologous DNA recombination in hESCs remains a laborious and challenging task for an ordinary laboratory, only a few successful homologous recombinations in hESCs have been reported [8,24,39,40]. Our master hESC lines with the Cre RMCE system eliminate the need for screening a large number of cell clones. We hope that versatile transgenic hESC lines may be built upon these master hESC lines by most laboratories.

Our system for establishing transgenic hESC lines depends on high efficient Cre RMCE, however, the efficiency of Cre-mediated recombination by traditional co-transfection remains low in hESCs. This is at least partly due to the low transfection efficiency, leading to the rare co-existence of the Cre recombinase and the targeting vector. By using the newly developed cell permeable Cre protein transduction method [26], we showed about 30% recombination efficiency. This is comparable to the RMCE efficiency in mouse ESCs [41]. Use of cell permeable Cre protein temporarily and at a low concentration is also beneficial to the hESC growth and karyotype stability, as Cre recombinase is known to have possible toxic effects that can compromise normal cell cycle and survival [42,43].

Together, the master hESC lines with the double loxP exchange cassette and RMCE via the cell permeable Cre protein transduction method offer a flexible and simple platform for genetic manipulation of hESCs. A transgenic cell line can be easily obtained by Cre recombination mediated exchange with a target gene of interest, and a series of different genes may be introduced into the same integration site to evaluate gene function without the variation in the level and pattern of gene expression. The master hESC lines were deposited to the WiCell Institute/ U.S. National Stem Cell Bank for distribution. The availability of these versatile tools will change the way we study human stem cells.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Prof. S.F. Dowdy for the plasmid pTAT-Cre, Prof. M. Nakafuku for mouse *Olig2* cDNA, Prof. S. Kugler for the plasmid pMH4-I-SYN-EGFP, and Prof. D.Trono for plasmid pLVTHM and pLVCT-tTRKRAB. This study was supported by NIH (NS061243) and National MS Society.

References

1. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. *Science* 1998;282:1145–1147. [PubMed: 9804556]
2. Reubinoff BE, Pera MF, Fong CY, et al. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat Biotechnol* 2000;18:399–404. [PubMed: 10748519]
3. Downing GJ, Battey JF Jr. Technical assessment of the first 20 years of research using mouse embryonic stem cell lines. *Stem Cells* 2004;22:1168–1180. [PubMed: 15579637]
4. Rossant J, Bernelot-Moens C, Nagy A. Genome manipulation in embryonic stem cells. *Philos Trans R Soc Lond B Biol Sci* 1993;339:207–215. [PubMed: 8097050]
5. Xia, X.; Zhang, SC. Genetic modification of human embryonic stem cells. In: Harding, SE., editor. *Biotechnology&Genetic Engineering Reviews*. Vol. Vol 24. Nottingham University Press; Nottingham: 2007. p. 297-310.
6. Zeng X, Rao MS. Controlled genetic modification of stem cells for developing drug discovery tools and novel therapeutic applications. *Curr Opin Mol Ther* 2008;10:207–213. [PubMed: 18535927]
7. Eiges R, Schuldiner M, Drukker M, et al. Establishment of human embryonic stem cell-transfected clones carrying a marker for undifferentiated cells. *Curr Biol* 2001;11:514–518. [PubMed: 11413002]
8. Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. *Nat Biotechnol* 2003;21:319–321. [PubMed: 12577066]
9. Liew CG, Draper JS, Walsh J, et al. Transient and stable transgene expression in human embryonic stem cells. *Stem Cells* 2007;25:1521–1528. [PubMed: 17379764]
10. Xia X, Zhang Y, Zieth CR, et al. Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. *Stem Cells Dev* 2007;16:167–176. [PubMed: 17348812]
11. Xiong C, Tang DQ, Xie CQ, et al. Genetic engineering of human embryonic stem cells with lentiviral vectors. *Stem Cells Dev* 2005;14:367–377. [PubMed: 16137225]
12. Zhou BY, Ye Z, Chen G, et al. Inducible and reversible transgene expression in human stem cells after efficient and stable gene transfer. *Stem Cells* 2007;25:779–789. [PubMed: 17158240]
13. Schroder AR, Shinn P, Chen H, et al. HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* 2002;110:521–529. [PubMed: 12202041]
14. Mitchell RS, Beitzel BF, Schroder AR, et al. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* 2004;2:E234. [PubMed: 15314653]
15. Smith JR, Maguire S, Davis LA, et al. Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells* 2008;26:496–504. [PubMed: 18024421]
16. Thyagarajan B, Liu Y, Shin S, et al. Creation of engineered human embryonic stem cell lines using phiC31 integrase. *Stem Cells* 2008;26:119–126. [PubMed: 17962703]
17. Zambrowicz BP, Imamoto A, Fiering S, et al. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc Natl Acad Sci U S A* 1997;94:3789–3794. [PubMed: 9108056]
18. Bronson SK, Plaehn EG, Kluckman KD, et al. Single-copy transgenic mice with chosen-site integration. *Proc Natl Acad Sci U S A* 1996;93:9067–9072. [PubMed: 8799155]
19. McCreath KJ, Howcroft J, Campbell KH, et al. Production of gene-targeted sheep by nuclear transfer from cultured somatic cells. *Nature* 2000;405:1066–1069. [PubMed: 10890449]
20. Soukharev S, Miller JL, Sauer B. Segmental genomic replacement in embryonic stem cells by double lox targeting. *Nucleic Acids Res* 1999;27:e21. [PubMed: 10471751]
21. Masui S, Shimosato D, Toyooka Y, et al. An efficient system to establish multiple embryonic stem cell lines carrying an inducible expression unit. *Nucleic Acids Res* 2005;33:e43. [PubMed: 15741176]
22. Beard C, Hochedlinger K, Plath K, et al. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* 2006;44:23–28. [PubMed: 16400644]
23. Yu J, McMahon AP. Reproducible and inducible knockdown of gene expression in mice. *Genesis* 2006;44:252–261. [PubMed: 16676321]

24. Irion S, Luche H, Gadue P, et al. Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat Biotechnol* 2007;25:1477–1482. [PubMed: 18037879]
25. Giel-Moloney M, Krause DS, Chen G, et al. Ubiquitous and uniform in vivo fluorescence in ROSA26-EGFP BAC transgenic mice. *Genesis* 2007;45:83–89. [PubMed: 17269129]
26. Nolden L, Edenhofer F, Haupt S, et al. Site-specific recombination in human embryonic stem cells induced by cell-permeant Cre recombinase. *Nat Methods* 2006;3:461–467. [PubMed: 16721380]
27. Zhang SC, Wernig M, Duncan ID, et al. In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nat Biotechnol* 2001;19:1129–1133. [PubMed: 11731781]
28. Li XJ, Du ZW, Zarnowska ED, et al. Specification of motoneurons from human embryonic stem cells. *Nat Biotechnol* 2005;23:215–221. [PubMed: 15685164]
29. D'Amour KA, Agulnick AD, Eliazer S, et al. Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol* 2005;23:1534–1541. [PubMed: 16258519]
30. Wiznerowicz M, Szulc J, Trono D. Tuning silence: conditional systems for RNA interference. *Nat Methods* 2006;3:682–688. [PubMed: 16929312]
31. Szulc J, Wiznerowicz M, Sauvain MO, et al. A versatile tool for conditional gene expression and knockdown. *Nat Methods* 2006;3:109–116. [PubMed: 16432520]
32. Tan G, Gao Y, Shi M, et al. SiteFinding-PCR: a simple and efficient PCR method for chromosome walking. *Nucleic Acids Res* 2005;33:e122. [PubMed: 16077029]
33. Wadia JS, Stan RV, Dowdy SF. Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat Med* 2004;10:310–315. [PubMed: 14770178]
34. Call LM, Moore CS, Stetten G, et al. A cre-lox recombination system for the targeted integration of circular yeast artificial chromosomes into embryonic stem cells. *Hum Mol Genet* 2000;9:1745–1751. [PubMed: 10915762]
35. Wichterle H, Lieberam I, Porter JA, et al. Directed differentiation of embryonic stem cells into motor neurons. *Cell* 2002;110:385–397. [PubMed: 12176325]
36. Lee SK, Lee B, Ruiz EC, et al. Olig2 and Ngn2 function in opposition to modulate gene expression in motor neuron progenitor cells. *Genes Dev* 2005;19:282–294. [PubMed: 15655114]
37. Grosveld F, van Assendelft GB, Greaves DR, et al. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 1987;51:975–985. [PubMed: 3690667]
38. Costa M, Dottori M, Ng E, et al. The hESC line Envy expresses high levels of GFP in all differentiated progeny. *Nat Methods* 2005;2:259–260. [PubMed: 15782217]
39. Urbach A, Schuldiner M, Benvenisty N. Modeling for Lesch-Nyhan disease by gene targeting in human embryonic stem cells. *Stem Cells* 2004;22:635–641. [PubMed: 15277709]
40. Costa M, Dottori M, Sourris K, et al. A method for genetic modification of human embryonic stem cells using electroporation. *Nat Protoc* 2007;2:792–796. [PubMed: 17446878]
41. Adams LD, Choi L, Xian HQ, et al. Double lox targeting for neural cell transgenesis. *Brain Res Mol Brain Res* 2003;110:220–233. [PubMed: 12591158]
42. Loonstra A, Vooijs M, Beverloo HB, et al. Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proc Natl Acad Sci U S A* 2001;98:9209–9214. [PubMed: 11481484]
43. Forni PE, Scuoppo C, Imayoshi I, et al. High levels of Cre expression in neuronal progenitors cause defects in brain development leading to microencephaly and hydrocephaly. *J Neurosci* 2006;26:9593–9602. [PubMed: 16971543]

Fig. 1. Silencing-resistant master hESC lines with a built-in double loxP cassette. (A) Scheme of Cre-double loxP exchange system. The master ESC line was examined by phase contrast (B) and fluorescent illumination for hrGFP expression at passage 40 (C). Bar = 50 μ m. (D, E) Flow cytometric quantification of the hrGFP expression population. 99.0% SSEA4⁺ ESCs showed hrGFP expression.

Fig. 2. Sustained GFP expression following differentiation of the master hESCs in vitro and in vivo. The hrGFP expression was detected in all the ESC aggregates at day 6 (A), neuroepithelial cells in the rosettes at day 14 (B), TuJ1⁺ neurons at 6 wks (C, D), and GFAP⁺ astrocytes at 10 wks (E, F) during neural differentiation. The hrGFP expression was also observed in Brachyury T⁺ precursors at day 3 (H) and cTnT⁺ cardiac muscle cells at day 17 (I) during mesodermal differentiation, and, in Sox17⁺ precursors at day 6 (M) and Pdx1⁺ pancreatic cells at day 17 (N) during endodermal differentiation. Teratoma analysis showed that GFP expression was sustained in neuroectodermal rosette cells (G, H), mesoderm derived muscle and cartilage (J, K), and endoderm derived epithelial tube (O, P). Bar = 50 μ m.

Fig. 3.

Replacement of the master GFP cassette with the targeting RFP cassette through RMCE. (A) Scheme of cassette exchange before and after RMCE. The sizes of DNA fragments from PCR and Southern blotting analysis were indicated. B, BamHI site. The correctly exchanged cell clones showed loss of GFP (B) and gain of RFP expression (C). Some un-exchange cell clones showed co-expression of GFP and RFP (D, E). The correctly exchanged clones were confirmed by PCR (F) and Southern blotting analysis (G). RMCE exchanged clones showed a specific 2.3 kb PCR-amplified band and a 5.4 kb BamHI-digested fragment which displayed the same restriction pattern as that in the master cells, whereas the un-exchanged clones showed no PCR-amplified band but an extra BamHI-digested fragment.

Fig. 4. hESC line with functional Olig2 gene expression. (A) Structure of the Olig2-FLAG targeting vector. (B) FLAG staining was present in all cells when the exchanged Olig2-FLAG expressing cells were differentiated to neuroepithelial cells (rosette) for 14 days. The Olig2 expression was detected in the nestin⁺ neural precursor cells (C), but not in the parental H9 cells differentiated at the same period (D). Following differentiation to motor neurons at the 5th week, the parental H9 cells generated many HB9⁺ motor neurons (E) whereas the Olig2-expressing cells produced few HB9⁺ neurons (F). Bar = 50 μ m. (G) Quantification of HB9⁺ motor neurons.

Fig. 5. hESC lines with inducible GFP expression. (A) Design of the inducible GFP targeting vector. Following the exchange of the loxP cassette, GFP was negative without Dox treatment (B), and was turned on when treated with 1 µg/ml Dox for 48 hrs (C). Bar = 50 µm. (D) Dox dose-dependent analysis of the inducible ESC line. (E) The kinetics of GFP induction and degradation in the inducible ESC line.

Fig. 6. hESC lines with neuron specific GFP expression (A) Schematic diagram of the targeting vector with GFP expression driven by the Synapsin promoter. The Syn-GFP ESC line did not show GFP expression in nestin⁺ neuroepithelial cells at day 14 (B, C), but displayed the GFP in synapsin⁺ maturing neurons after further differentiation at day 35 (D, E). Bar = 25 μ m. Blue represent Hoechst stained nuclei.

Table 1

Efficiency of GFP expression in the transfected human and mouse ESCs

	GFP positive clones/total clones				
	Human ESC			Mouse ESC	
Promoter	Exp 1	Exp 2	Exp 3	Total	
CAG	6/35	4/25	7/38	17/98 (17.3%)	211/254 (83.1%)