



## A DNA transposon-based approach to functional screening in neural stem cells

Ilaria Albieri<sup>a,c,1</sup>, Marco Onorati<sup>b,1</sup>, Giovanna Calabrese<sup>a</sup>, Alessia Moiana<sup>b</sup>, Daniele Biasci<sup>e</sup>, Aurora Badaloni<sup>a</sup>, Stefano Cannasio<sup>b</sup>, Dimitrios Spiliotopoulos<sup>b</sup>, Zoltán Ivics<sup>d</sup>, Elena Cattaneo<sup>b,\*,2</sup>, G. Giacomo Consalez<sup>a,\*\*,2</sup>

<sup>a</sup> Division of Neuroscience, San Raffaele Scientific Institute, Via Olgettina 58, 20132 Milano, Italy

<sup>b</sup> Department of Pharmacological Sciences and Centre for Stem Cell Research, Università degli Studi di Milano, 20133 Milano, Italy

<sup>c</sup> Università Vita-Salute San Raffaele, Via Olgettina 58, 20132 Milano, Italy

<sup>d</sup> Max Delbrück Center for Molecular Medicine, Berlin, Germany

<sup>e</sup> University of Pisa, Department of Biology, Pisa, Italy

### ARTICLE INFO

#### Article history:

Received 17 May 2010

Accepted 19 July 2010

#### Keywords:

Neural stem cells

Transposons

Screenings

Sleeping beauty

### ABSTRACT

We describe the use of DNA transposons as tools for carrying out functional screenings in murine embryonic stem (ES) cell-derived neural stem (NS) cells. NS cells are a new type of stem cells featuring radial glial properties, that undergoes symmetric cell division for an indefinite number of passages, expanding as a monolayer. In this model, the previously unreported *Sleeping Beauty* transposase M3A achieves an optimal blend of clone generation efficiency and low redundancy of integrations per clone, compared to the SB100X *Sleeping Beauty* variant and to the *piggyBac* transposon. The technology described here makes it possible to randomly trap genes in the NS cell genome and modify their expression or tag them with fluorescent markers and selectable genes, allowing recombinant cells to be isolated and expanded clonally. This approach will facilitate the identification of novel determinants of stem cell biology and neural cell fate specification in NS cells.

© 2010 Elsevier B.V. All rights reserved.

### 1. Introduction

The wealth of genetic information and functional annotation contained in genomic databases makes it possible to conduct genetic analysis on a scale that was previously unimaginable. Powerful technologies have been made available to the scientific community to perform high throughput transcriptome analysis, exploiting the huge power of gene profiling platforms. Despite the invaluable use of these studies, functional screenings remain essential for the isolation and characterization of genes that play key roles in development, differentiation, survival and neoplastic transformation.

Expression cloning screens have left an indelible mark on our understanding of several key developmental processes, among which neural induction in *Xenopus* embryos (Piccolo et al., 1996; Smith and Harland, 1992) and stemness/self-renewal in embryonic

stem (ES) cells (Chambers et al., 2003). In these screens, collections of *in vitro* synthesized mRNAs or pools of cDNA library clones were used to isolate genes producing selectable phenotypes.

As a conceptually different approach, retro- and lentiviral vectors can be used as sources of dominant enhancers or other biologically active sequences integrating into the host genome, to randomly perturb gene expression or function. While the power of viral vectors is undisputable, their preparation can be problematic as their non-random genomic integration pattern (Cattoglio et al., 2007; Kustikova et al., 2005, 2007). An alternative to viral vectors is offered by DNA transposons. DNA transposons are non-viral molecules that have been widely used to manipulate the genome of cell lines and living organisms.

Unlike retroviral and lentiviral vectors, DNA transposons show a significantly lower level of insertion bias in the mouse and human genomes, and can integrate both within genes and in intergenic regions, either transcribed or functionally inert (Liang et al., 2009; Yant et al., 2005). Transposons are nucleotide sequences found in the genome of many species as active (Ding et al., 2005; Fraser et al., 1996) or fossil (Goodier and Davidson, 1994; Ivics et al., 1997) elements. They are flanked on each side by a DNA sequence recognized and excised by a specific transposase. This enzyme excises a transposable element from its primitive location, genomic or episomal, and promotes its random single-copy integration elsewhere in the genome.

\* Corresponding author at: Department of Pharmacological Sciences, University of Milano, Via Balzaretti 9, 20133 Milano, Italy. Tel.: +39 02 5032 5840; fax: +39 02 5031 8284.

\*\* Corresponding author. Tel.: +39 02 2643 4838; fax: +39 02 2643 5283.

E-mail addresses: [elena.cattaneo@unimi.it](mailto:elena.cattaneo@unimi.it) (E. Cattaneo), [g.consalez@hsr.it](mailto:g.consalez@hsr.it) (G.G. Consalez).

<sup>1</sup> Equal first authors.

<sup>2</sup> Equal last authors.

In particular, the *Sleeping Beauty* (SB) family of DNA transposons was reconstructed from inactive fossil elements identified in salmonid fishes (Ivics et al., 1997). They consist of two almost identical inverted repeat/direct repeats (IR/DRs) flanking a transposase gene. Several generations of SB transposons have been produced through successive modifications of the IR/DRs (Cui et al., 2002; Liu et al., 2004) and of the SB transposase coding sequence (Baus et al., 2005; Izsvák et al., 2002; Zayed et al., 2004). In nature, the transposase-encoding sequence is built into the transposable element. However, in certain experimental paradigms, the transposase construct can be provided separately, *in trans*, while the transposable element can be modified to contain recombinant genes, *cis*-acting regulatory sequences, gene traps, etc. Thus, DNA transposons lend themselves to functional expression cloning, as shown by the success of some ingenious *in vivo* applications adopted in cancer research (Dupuy et al., 2005; Starr et al., 2009).

Our group is interested in exploiting the power of functional screening to address some key questions regarding the chain of events leading to the development of mature neurons from multipotent neural progenitors. Open questions of great fundamental relevance and practical consequence in developmental neurobiology concern the mechanisms controlling self-renewal of neural stem cells, the switch between symmetric and asymmetric (neurogenic) cell division, the choice between neuronal and glial fates, and the mechanisms regulating neuronal fate determination and differentiation.

Until recently, progress in unraveling the biochemical and molecular nature of these processes was partially hampered by the unavailability of a suitable cellular model. To overcome this limitation, a protocol was developed (Conti et al., 2005) allowing neural stem cells to be cultured for the first time as an adherent monolayer, and to be expanded indefinitely as a homogeneous population of radial glial-like progenitors capable of self-renewal and symmetric cell division. These cells were termed neural stem (NS) cells, by virtue of their properties analogous to those of ES cells. NS cells were derived from mouse ES cells, and from mouse and human fetal or adult brain (Conti et al., 2005; Pollard et al., 2006; Sun et al., 2008). NS cells can be also efficiently obtained from induced pluripotent stem (iPS) cells (Kim et al., 2009; Onorati et al., 2009). Here we describe the development and proof of feasibility of a transposon-based expression cloning approach developed in NS cells, which is based on DNA transposable elements of the SB family. We show that these molecules integrate at extremely high frequencies and in a stable fashion into the genome of these cells, where they can be used to manipulate and/or trace the expression of genes flanking their integration site, and to select for, and expand clones carrying integrations that produce a cellular response of interest. Our results suggest that transposon tagging provides a suitable approach for the isolation of coding and noncoding regulatory sequences in NS cells.

## 2. Material and methods

### 2.1. Constructs

The M3A transposase plasmid was generated introducing the following amino acid replacements in the *Sleeping Beauty* 11 sequence: K13A, K33A, T83A, R214D, K215A, E216V, N217Q and S270A. The replacements were introduced by using the QuickChange mutagenesis kit (Stratagene). The transposable elements described in this paper were prepared as follows: a cyan fluorescent protein (CFP) sequence was amplified from the pECFP-N3 (Clontech) plasmid using the following primers: CFP.fus.F: 5'-TCGCCACCATGGTGTAGCAAGGGCGAG-3'; CFP.fus.R:

5'-CCGGAATT CCTGTACAGCTCGTC-3', and cloned into a pBlue-script plasmid cut with *SmaI* and *EcoRI*. To generate a fusion sequence between CFP and a neomycin resistance gene (Neo), the Neo sequence was amplified from the pcDNA 3.1 plasmid (Clontech) and cloned in frame with the CFP sequence into the *EcoRI* restriction site. The primers for Neo amplification are listed as follows: Neo.fus.F: 5'- CCGGAATTCATGATTGAACTCGTC-3' and Neo.fus.R: 5'- CGCGGATCCTCAGAAGAAGCTCGTC-3'. The CFPNeo coding sequence was excised with *HindIII* and *NotI* and cloned into an *EcoRI*-cut pCAGGs plasmid. The positive control transposon (C+) was obtained cloning the CAG-CFPNeo-poly(A) (poly A is the rabbit  $\beta$ -globin polyadenylation site) into an *EcoRV*-cut pBlue-script plasmid containing the *Sleeping Beauty* inverted/direct repeat sequences (left and right IR/DR). The IR/DR sequences were derived from the pMSZ-Neo plasmid, a kind gift of Bradley Fletcher. The hybrid transposon (Hyb) was generated cloning the IR/DRCAG-CFPNeo-poly(A)-IR/DR fragment into the pGG85 plasmid (a kind gift of Austin Smith) linearized with *HindIII*. The poly(A) trap transposon contains the CAG promoter, the CFPNeo fusion gene, followed by an internal ribosome entry site (IRES) sequence and a splicing donor site (SD). This construct was obtained replacing the polyadenylation sequence of IR/DR-CAGCFPNeo-poly(A)-IR/DR plasmid with an IRES-SD fragment, previously cloned into pBlue-script. The M3A *Sleeping Beauty* transposase variant was generated on the pCMV-SB10 background by the QuikChange site-directed mutagenesis kit (Stratagene), following the protocol provided by the manufacturer. The M3A coding sequence was recovered from the parental plasmid by *SacII*, and cloned into pCAG-Cre after removal of the Cre coding region by *PstI/MluI* and blunting the ends by T4 polymerase. SB100X was kindly provided by Zsuzsanna Izsvák. The *piggyBac* transposase plasmid (pCAGG-PBase) was a kind gift of Austin Smith.

As regards the gene trap construct, a green fluorescent protein (GFP) sequence was fused to a neomycin resistance gene (Neo) using the same set of primers described for CFPNeo fragment. The GFP sequence was amplified from the pEGFP-N3 plasmid (Clontech). A pBlue-script plasmid containing the GFPNeo sequence was obtained (GFPNeo-pBlue-script). A SV40 polyadenylation sequence was excised from pDS-Red plasmid (Clontech) with *HindIII* and *DraIII* and cloned into the *HindIII* site of GFPNeo-pBlue-script plasmid. A splicing acceptor site (SA), a triple stop signal in 3 different frames (3xSTOP) and an IRES sequence were cloned into GFPNeo-poly(A)- pBlue-script plasmid into the *NotI* site. The 3xSTOP sequence was generated by denaturation and successive annealing of the following primers: F: 5'- GGGTACTGACTGAGCGGCCCTACCTACCTAGGG, R 5'-CCTAGGTAGGTAGGCGGCCGCTCAGTCAGTCACCC. The entire SA-3XSTOP-GFPNeo-poly(A) fragment was excised with *XhoI* and *SacI* and cloned into an *EcoRV*-cut pBlue-script plasmid containing the *Sleeping Beauty* inverted/direct repeat sequences (left and right IR/DR).

### 2.2. NS cells and transfection

NS cells were cultured as described (Conti et al., 2005). Transfection of NS cells was performed according to an optimized nucleofection protocol (AMAXA) using Mouse NSC nucleofector kit. Briefly, 4 million cells were resuspended in 100  $\mu$ l of nucleofector solution and transfected with a transposon plasmid in presence or absence of a transposase helper plasmid. Two different transposon/transposase molar ratios have been tested (1:1 and 10:1 for M3A and SB100X transposases; 1:1 and 1:40 for *piggyBac* transposase), transfecting 3  $\mu$ g of total DNA. Transfected cells were plated at different densities (2500, 5000, 10,000 cells/plate) and maintained in culture with G418 (Sigma, A1720) selection (400  $\mu$ g/ml). After 12 days of G418 selection, neomycin resistant

clones on the plate were analyzed by fluorescent microscopy and stained by cresyl violet (Sigma).

### 2.3. Southern blotting

Genomic DNA was extracted from cell clones after an overnight digestion with protein kinase, using the phenol–chloroform extraction method. After 2-propanol precipitation, DNAs were rinsed with 70% ethanol and resuspended in water. After quantification, DNAs were digested with *EcoRI* or *HindIII*, run out on a 0.9% agarose gel and transferred to the UltraBind transfer membrane (PALL, Life Science). The neomycin probe used was an 807 bp fragment from a pBlue-Neo plasmid. The probe was labeled with  $\alpha$ -dCTP 32P (Perkin Elmer) using the Prime-a-gene Kit (Promega). Overnight hybridizations were performed at 65 °C.

### 2.4. LM-PCR

Genomic DNAs were digested with the *MspI* restriction enzyme. Digested DNAs were then ligated to an *MspI*-linker. The first round of LM-PCR was performed with one primer specific for the left IR/DR sequence and the other primer complementary to the linker sequence, under the following conditions: denaturation at 95 °C for 2 min, then 25 cycles as follows: 95 °C for 15 s, 56 °C for 30 s, 72 °C for 1 min. PCR products were diluted 1:50 and nested PCR was performed using a second set of nested primers, under these conditions: denaturation step 95 °C for 2 min, then 25 cycles as follows: 95 °C for 30 s, 56 °C for 45 s, 72 °C for 1 min. Nested PCR products were cloned into TOPO cloning Kit (Invitrogen) and sequenced using an M13 Reverse primer. Oligos used in the experiments are listed as follows (5'–3'): Linker+: 5'-GTAATACGACTCACTATAGGG-CTCCGCTTAAGGGAC; Linker–: 5'-PO4-CGGTCCCTTAAGCGGAG-NH<sub>2</sub>; IR/DR F: 5'-ACAAGAAATTTGTGGAGTAGTTG; Nested IR/DR F: 5'-AGTTTAAATGACTCCAACCTAAG; Linker R: 5'-GTAATACGACTCACTATAGGGC; Nested Linker R: 5'-AGGGC-TCCGCTTAAGGGAC.

### 2.5. Reverse transcription and polymerase chain reaction

Total RNA was prepared from wt NS cells and clone 10 cells by using the Trizol reagent (Roche), following manufacturer's recommendations. RT-PCR amplification was performed on 1–2  $\mu$ g total RNA. For first-strand synthesis we used the MMLV first-strand synthesis system for reverse transcription PCR (Invitrogen). For PCR on cDNAs derived from fusion transcripts, a forward primer was designed on the IRES sequence (5'-CTCGGTGCACATGCTTTACATG-3') and a set of reverse primers were designed on sequences of trapped genes. Reverse oligos used in the experiments are listed as follows: chr11.fus R: 5'-AGGACACAACAGGGCATGCTG; chr14.fus.R: 5'-GTTCTCTAGACAGCCTGATTATG; chr18.fus R: 5'-CTCCAATCTTCCCAGTTCCTC. cDNA amplification conditions were: denaturation step at 95 °C for 2 min, then 30 cycles as follows: 95 °C for 30 s, 60 °C for 45 s, 72 °C for 30 s. To test the over-expression of transcripts and of the hairpin containing sequence (hcs, or pri-miRNA) downstream of the transposon integration site, two primers were designed for each trapped gene. Specific oligos are listed as follows: chr11 F: 5'-CAAGAACAGCTGAACTGGCTG and R: 5'-CTTCAAGGTTGACATGGTGCAC; chr14 F: 5'-CATAATCAGGCTGTCTAGAGAAC and R: 5'- ATAGTGACAC-TACCACCTCATG; chr18 F: 5'- CAGATGAGCTCACAGCAGAG and R: 5'- CTCCAATCTTCCCAGTTCCTC; hcs F: 5'- TGAGCCACCCAC-TATGTGTG and R: 5'-GAGACAGGTTTCTGTATAGCC. Amplification conditions: denaturation step 95 °C for 2 min, then 30 cycles as follows: 95 °C for 30 s, 60 °C for 45 s, 72 °C for 30 s. Amplification products were analyzed on 2% agarose gels.

### 2.6. Quantitative real time RT-PCR

For RT-qPCR, first-strand synthesis was followed by real time amplification using ready-to-use SYBR Green PCR (Roche). The threshold cycles for the control (GAPDH) and gene of interest were determined with the "Second Derivative Methods" according to the LightCycler 480 Software. The following forward or reverse oligonucleotides were used: GAPDH F: 5'-TGAAGCAGGCATCTGAGGG-3' and GAPDH R: 5'-CGAAGGTGGAAGAGTGGGAC-3'; Chr11 RT F: 5'- CAAGAACAGCTGAACTGGCTG-3' and Chr11 RT R: 5'-CTTCAAGGTTGACATGGTGCAC-3'; hcs F: 5'- TGAGCCACCCAC-TATGTGTG and R: 5'- GAGACAGGTTTCTGTATAGCC.

### 2.7. In silico analysis

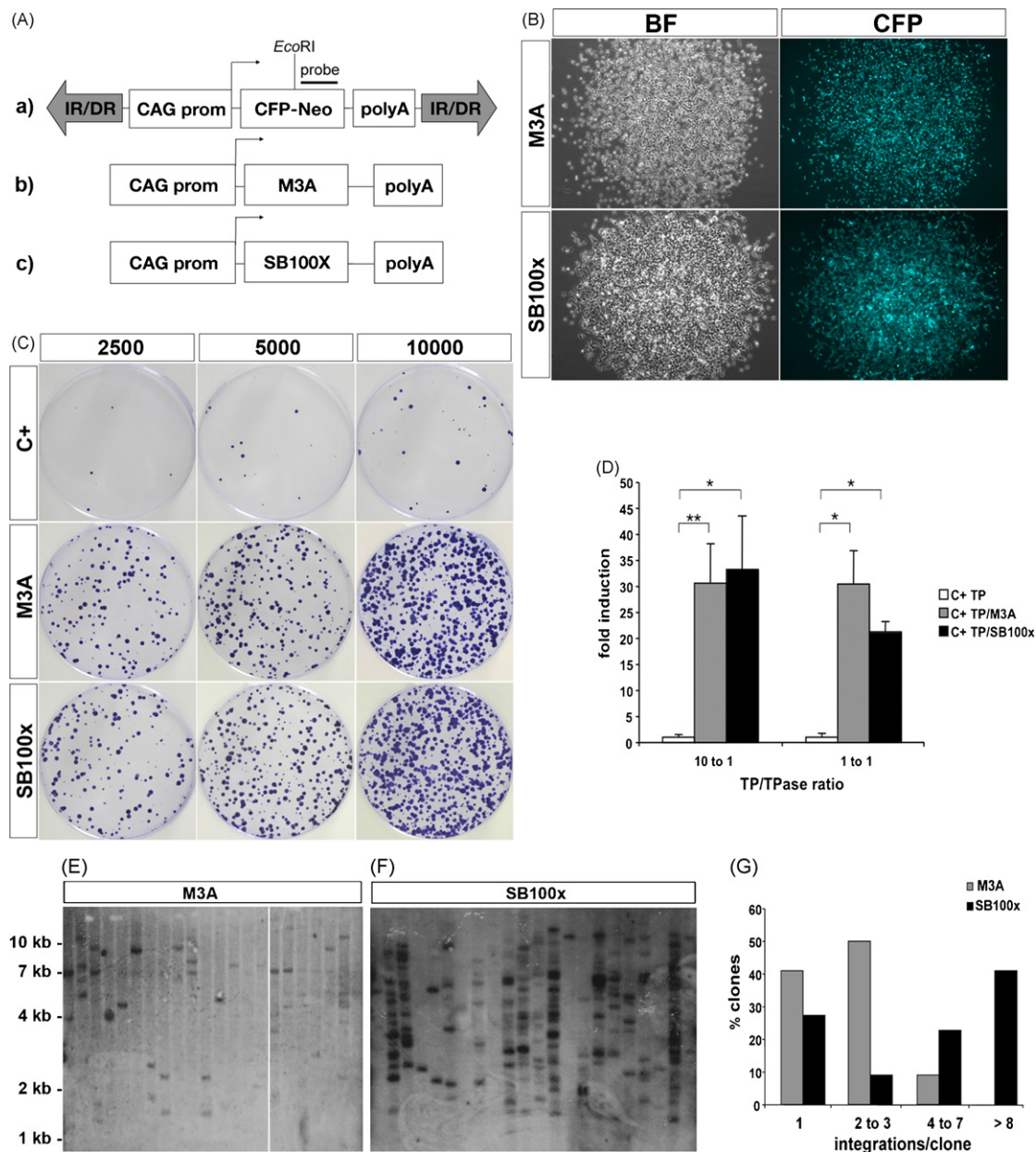
Putative microRNA precursors were annotated according to criteria established in Ambros et al. (2003). Precursor secondary structures were predicted using mfold ver. 3.5 (Zuker, 2003). Alignment between putative pre-miRNAs and mouse transcript databases were performed using BLAT at UCSC Genome Browser (Rhead et al., 2010). Annotated pre-miRNAs were downloaded from miRBase ver. 14 (Griffiths-Jones et al., 2008) and similarity to novel putative pre-miRNAs was computed using Smith–Waterman algorithm as implemented in the fasta35 package (Pearson and Lipman, 1988). 454 sequencing reads from smallRNA libraries were downloaded from the Gene Expression Omnibus (Barrett et al., 2009) and aligned to putative pre-miRNAs using the blast-n program from the blastall package (Strahm et al., 2006) (-F F -p blastn -W 18 -S 1 -m 8). All data analysis was automated using the in-house developed miRNADiscovery pipeline.

## 3. Results

### 3.1. SB transposons integrate at high efficiencies into the NS cell genome

A synthetic SB transposon (Fig. 1A) comprising a CAG promoter (Niwa et al., 1991), a cyan fluorescent protein (CFP) gene fused in frame to a neomycin phosphotransferase (Neo), hereon referred to as CFPNeo, followed by a rabbit  $\beta$ -globin polyadenylation sequence (poly(A)) was generated. The CFPNeo selectable gene was successfully tested in the HEK293 cell line for its ability to fluoresce and confer neomycin resistance to cells (data not shown). This transposon was dubbed C+, for positive control, and contains modified IR/DRs, as described (Liu et al., 2004; Zayed et al., 2004). This element does not contain a transposase coding sequence (CDS), thus a helper construct was provided in *trans*, consisting of a pCAGgs plasmid (Niwa et al., 1991) encoding an hyperactive transposase: M3A (for aminoacid substitutions see Section 2.1) or SB100X (Mates et al., 2009). The M3A transposase variant was generated by combining the following hyperactive amino acid replacements: R214D, K215A, E216V, N217Q (Zayed et al., 2004). Furthermore, it was engineered to contain the following mutations: K13A, K33A, T83A and S270A (Yant et al., 2004). M3A represents the most active composite variant obtained thus far by bringing the individual mutations into the same protein in different combinations. M3A catalyzes SB transposition at an efficiency of 10–15-fold higher than the originally reconstructed SB transposase in human HeLa cells (data not shown).

To optimize the efficiency of transposon-mediated mutagenesis in NS cells, we titrated transposon/transposase molar ratios, in order to achieve the ideal combination between the highest possible number of generated stable clones per plate and a manageable number of integrations per clone. To this end, we nucleofected NS

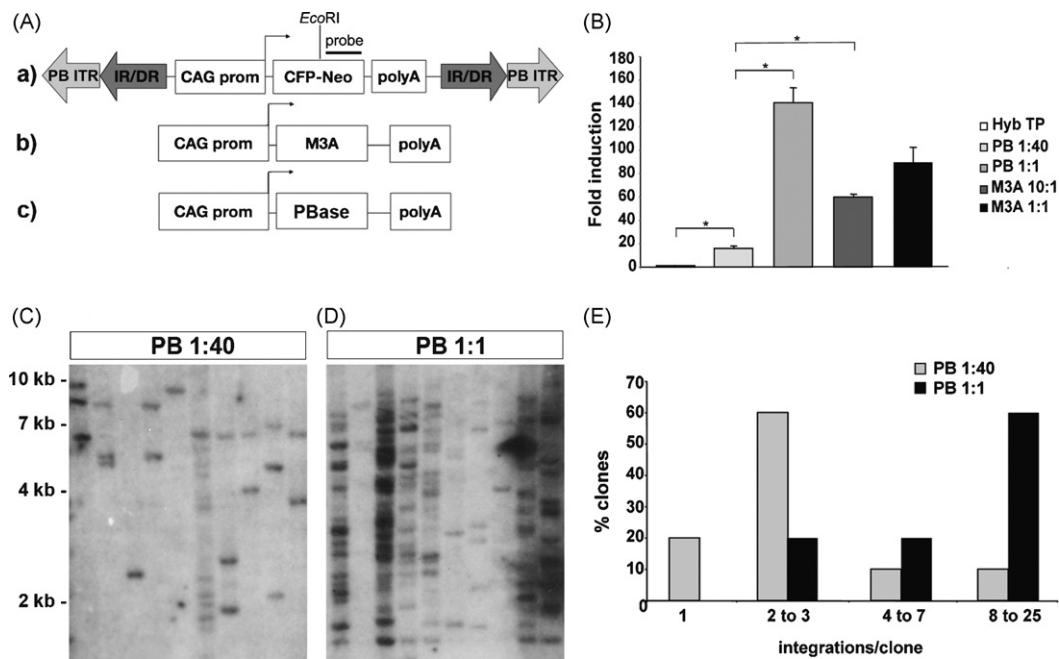


**Fig. 1.** Hyperactive SB transposases promote extremely high DNA integration efficiencies in NS cells. (A) Schematic representation of C+ transposon (positive control) (a) and helper plasmids encoding *Sleeping Beauty* (SB) transposase constructs (M3A, b; SB100X, c; see text for details). (B) Bright field (BF) and fluorescent (cyan fluorescent protein, CFP) images of clones obtained after neomycin selection of NS cells nucleofected in the presence of SB helper plasmids M3A and SB100X, as indicated. (C) Cresyl violet staining of clones obtained after neomycin selection of NS cells. In this experiment, the molar ratio of transposable element to helper plasmid was 10:1. (D) Histogram representing the fold increase in the number of clones/plate obtained plating 2500 cells/plate nucleofected with the C+ transposon plus either SB helper plasmid, as indicated (grey and black columns) vs. cells nucleofected with the transposable element alone (C+, empty column, equated to 1) ( $n=3$ , \* $p<0.005$ , \*\* $p<0.0005$ ). (E and F) Southern blots obtained digesting DNAs from NS cell clones with *EcoRI* and hybridizing them with the probe sketched in Aa to reveal a single junction fragment for each transposon integration. Clones in (E) were obtained nucleofecting cells with the M3A transposase; clones in (F) nucleofecting cells with the SB100X transposase. (G) Distribution of the number of integrations per clone obtained in (E) (grey columns) and in (F) (black columns).

cells with the C+ transposon (2  $\mu$ g) either without helper plasmid, or using two different transposon/transposase molar ratios for both M3A and SB100X, i.e. 10:1 and 1:1. Transfected cells were plated at three different densities (2500, 5000 and 10,000 cells/plate) and cultured under G418 selection for 12 days. Ensuing clones showed endogenous CFP fluorescence (Fig. 1B). Clones were stained with cresyl violet and counted. At both molar ratios tested, cresyl violet staining (Fig. 1C and Supplemental Fig. 1) revealed a dramatic increase in the number of Neo-resistant clones when NS cells were transfected in the presence of the transposase-encoding helper plasmid. Fig. 1D illustrates the numbers of clones per plate scored after neomycin selection. Our statistical analysis was based on

observations made at the lowest cell plating density used (2500 cells per 10 cm dish), which facilitates an accurate quantitation of clone numbers. At this cell density, both transposases increased clone formation by about 30 times compared to cells nucleofected with the transposable element alone. In other words, we isolated one clone out of approximately 12.5 cells plated.

The above results indicate that both M3A and SB100X are highly efficient as regards clone formation. No statistically significant differences were observed between the two, in our experimental conditions. The 10:1 molar ratio can be considered optimal because it reduces the amount of transfected helper plasmid (from 1.5  $\mu$ g to 150 ng) without significantly decreasing the number of



**Fig. 2.** *piggyBac* transposase increases the number of neomycin resistant clones producing a high number of integrations per clone. (A) Schematic representation of (a) hybrid transposon (Hyb; see text for details) and (b and c) two helper plasmids encoding hyperactive SB transposase (M3A) and *piggyBac* (PB) transposase (PBBase). (B) Histogram representing the fold increase in the number of clones/plate obtained plating 2500 cells per plate nucleofected with the Hyb transposon plus either helper plasmid (M3A or PB) vs. cells nucleofected with the transposable element alone (Hyb, empty column, equated to 1) ( $n = 3$ ,  $^* p < 0.005$ ). Two different transposon/transposase molar ratios were tested for each co-transfection conditions (PB: 1:40, and 1:1; M3A: 10:1 and 1:1). (C and D) Southern analysis of integrations promoted by the PB transposase. Genomic DNAs from neo-resistant clones obtained by Hyb/PBBase co-transfection experiments (molar ratio 1:40 and 1:1) were digested with *EcoRI* and hybridized with a Neo probe (see sketch in Fig. 2Aa) to reveal a single junction fragment for each transposon integration. (E) Distribution of the number of PB transposon integrations per clone observed in clones shown in (C) (grey column) and (D) (black column).

stable clones generated. Low helper plasmid concentrations permit a rapid loss of the episomal SB transposase, thus favoring integration stability and reducing the occurrence of multiple rounds of transposition.

Next, we analyzed the number of integrations per clone obtained in the experimental conditions described above. In fact, while high cloning efficiencies are highly desirable in a large scale screening, an excessive number of integrations per clone makes mapping of each integration problematic or unmanageable. By Southern analysis of *EcoRI*-digested DNAs from 44 cellular clones using a probe (Fig. 1Aa) that reveals one restriction fragment per integration (Fig. 1E and F), we observed that M3A (Fig. 1E) produces a low number of integrations per clone, in a range from 1 to 6, whereas SB100X (Fig. 1F) produces integrations in a range from 1 to 15 per clone. Further analysis showed that 90% of clones generated by the M3A transposase contained 1 to 3 integrations per clone, whereas about 40% of SB100X-produced clones showed more than 8 integrations/clone (Fig. 1G). The observed differences were statistically significant ( $p < 0.001$ , Mann–Whitney test).

### 3.2. Comparative analysis of SB and *piggyBac* transposons in NS cells

In addition to *Sleeping Beauty*, an unrelated family of mobile DNA elements has been successfully employed in many research applications (Wang et al., 2009; Wilson et al., 2007; Woltjen et al., 2009), the *piggyBac* (PB) transposon. We set out to compare the performance of the two systems in NS cells. Sketched in Fig. 2A is a hybrid transposable element that can be excised and reintegrated by the SB and PB transposase alike. For SB, we chose to employ 10:1 and 1:1 transposon/transposase molar ratios. As regards the PB element, we employed a 1:1 molar ratio, in order to maximize the number of neomycin resistant clones, as well as a 1:40 transposon/transposase molar ratio, along the lines of experiments

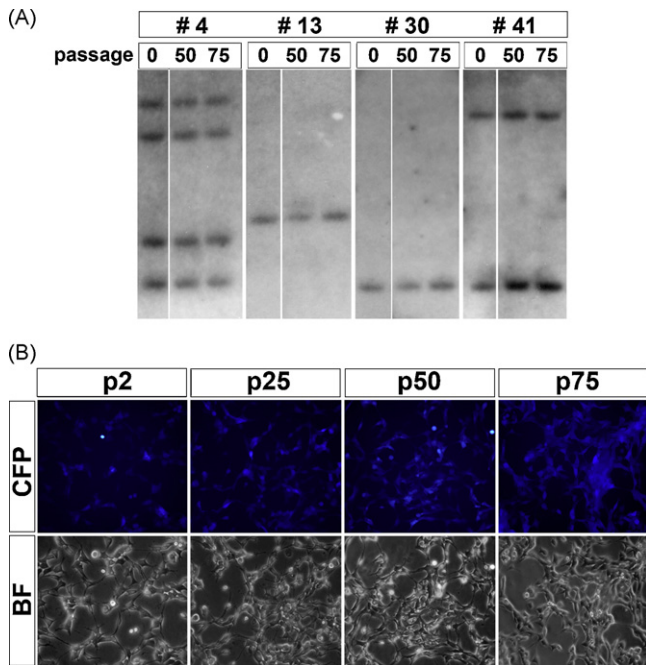
described by other authors (Wang et al., 2009) to minimize the number of integrations per clone.

NS cells were nucleofected with the hybrid transposon (2  $\mu\text{g}$ ), either with or without SB or PB helper plasmid. Transfected cells were plated at two different densities (2500, 5000 cells/plate) and cultured under G418 selection. Ensuing clones were tested for CFP fluorescence (Supplemental Fig. 2A). After 12 days of selection, clones were fixed, stained with cresyl violet (Supplemental Fig. 2B and C) and counted (Fig. 2B). At the 1:1 molar ratio, the PB transposase increases the number of generated stable clones by about 140-fold compared to single transfected cells. Reducing the amount of transposon plasmid (from 2  $\mu\text{g}$  to 50 ng, to obtain a 1:40 transposon/transposase molar ratio) results in a substantial decrease in the number of neomycin resistant clones ( $47 \pm 6.66$  clones vs.  $421 \pm 13.05$ ). For the SB system, the 10:1 and 1:1 molar ratios yield a 60- and 80-fold increase, respectively, compared to transfection with the hybrid transposon alone; the histograms in Fig. 2B illustrate these findings.

Finally, we estimated by Southern blotting the number of integrations per clone obtained using the PB transposon under the conditions described above (Fig. 2C and D). Our results indicate that, at the 1:40 transposon/transposase molar ratio, 80% of stable clones exhibit integrations in a range between 1 and 3 per clone. Conversely, at the 1:1 molar ratio, 60% of stable clones show greater than 8 integrations/clone. While at the 1:1 molar ratio the PB system produces higher total numbers of neomycin resistant clones, the flip side of the coin is that the number of integrations/clone obtained with this reagent may be unmanageably high for certain applications.

### 3.3. Transposon integration stability over 75 passages in culture

The next step was to estimate the ability of the SB transposon to integrate stably and express a fluorescent marker for a pro-



**Fig. 3.** Integration stability over 75 passages in culture. (A) Southern blot analysis of DNAs extracted from four neomycin resistant clones (#4, #13, #30, #41) expanded in culture for 75 passages. DNAs were digested with *EcoRI* and hybridized with a Neo probe (see sketch in Fig. 1Aa). Pattern of bands of each clone were checked at passages 2, 50 and 75. (B) Endogenous cyan fluorescent protein (CFP) and bright field (BF) images of clone 41 (#41) expanded in culture for 75 passages.

longed period in culture. To this end, we expanded and analyzed four neomycin resistant NS cell clones, containing eight overall SB transposon integrations, over 75 passages *in vitro*. The banding patterns relative to all four clones remained unmodified throughout 74 passages in culture (from +2 through +75), indicating that our transposable element can be integrated stably for a long period of time in the NS cell genome (Fig. 3A). To assess whether stably integrated transposons remained transcriptionally active, we analyzed endogenous CFP fluorescence at 2, 25, 50 and 75 passages *in vitro* in the absence of antibiotic selection (Fig. 3B). Our results confirmed that stably nucleofected clones maintained fluorescent reporter gene expression over 74 passages in culture, likely escaping epigenetic silencing.

#### 3.4. Expression cloning in NS cells: poly(A) trapping

Several strategies may be adopted to exploit the features of SB transposons. For example, a poly(A) trap transposon can be used to overexpress sequences (importantly, including noncoding RNAs) downstream of the integration site by generating fusion transcripts. To this end, a poly(A) trap transposon (sketched in Fig. 4A) was generated as described in Section 2.1.

This construct (poly(A)trap) is inactive as an episome. However, upon co-transfection with the helper plasmid encoding the SB transposase, the transposable element is excised from its donor episomal vector, and integrates randomly into the NS cell genome. By integrating upstream of any given gene, the transposon will create a fusion transcript, recruiting the host gene's poly(A) sequence and overexpressing the gene or a part of it. At the same time, both the fluorescent reporter and neomycin resistance will be activated. The expected mode of action of this construct is sketched in Supplemental Fig. 3.

To test the system, NS cells were nucleofected with the poly(A) trap plasmid with or without the helper plasmid expressing the M3A transposase. After 12 days of G418 selection, neomycin resis-

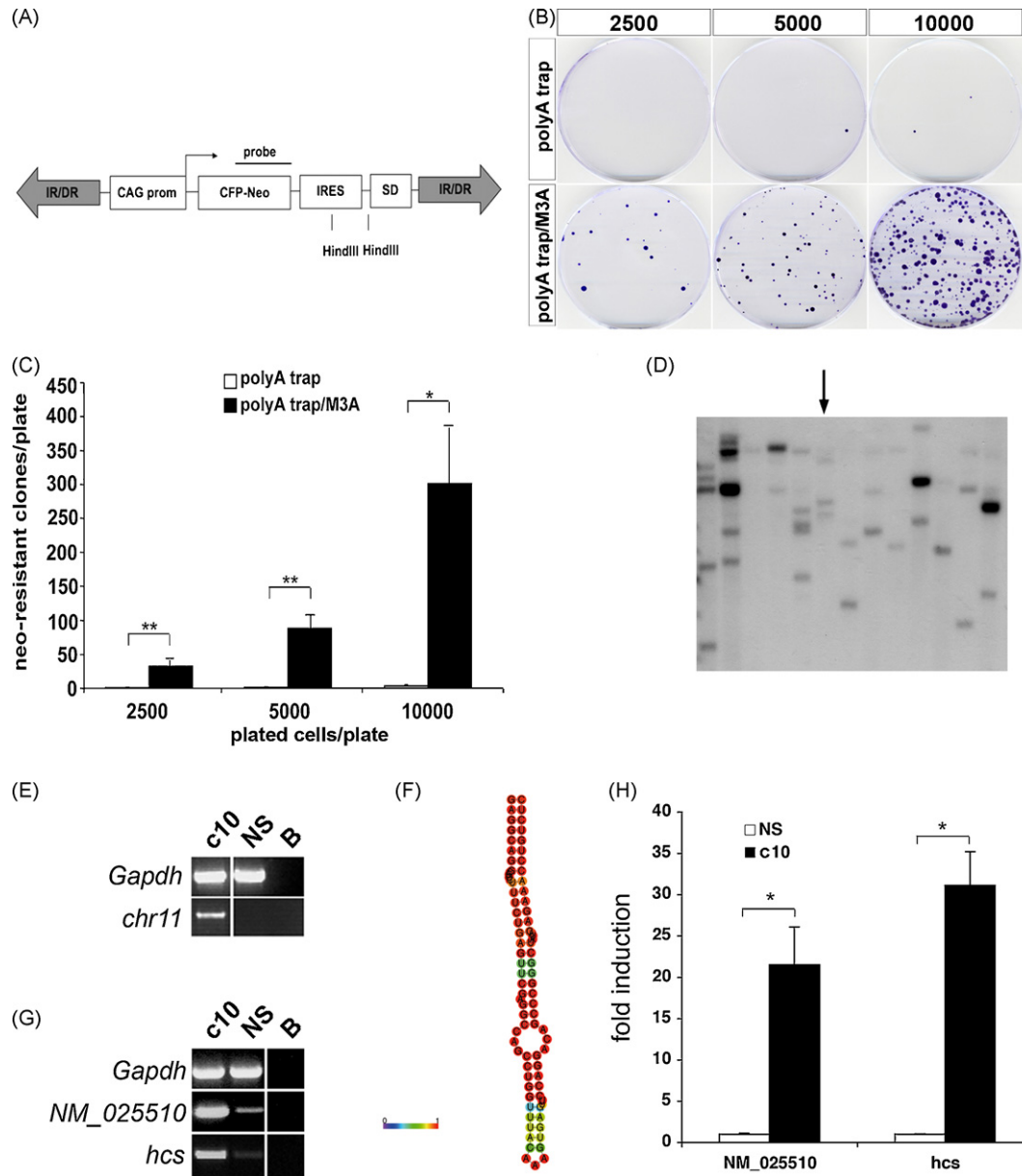
tant clones were analyzed for expression of CFP (Supplemental Fig. 4), and stained by cresyl violet (Fig. 4B). In this experiment the M3A helper plasmid increased the number of neomycin resistant clones by about 72-fold in cells plated at a density of 10,000/plate, compared to the circular transposon plasmid alone (Fig. 4C); in other words, 3.17% of all plated cells gave rise to a neomycin resistant clone. In this experiment, from each nucleofection (4 million cells transfected) we obtained about 50,000 neomycin resistant clones, roughly corresponding to 127,000 total integration sites, based on an average of 2.54 integrations/clone (Fig. 4D).

#### 3.5. Genomic distribution of poly(A) trap transposon integrations

As a proof of principle of this approach, we characterized by LM-PCR 30 sample clones containing poly(A) trap transposon integrations. 51 out of 74 integrations were mapped and their distribution analyzed in the NS cell genome (Table 1). 17/26 poly(A) trap transposon integrations occurred intragenically and in a sense orientation, suggesting a preference for those integrations that allow the recruitment of a poly(A) signal by the trap construct. Included in this class were all integrations occurring within transcribed sequences, or up to 10 kb upstream of the transcriptional start site (TSS). The other integrations (nonsense) were distributed as follows: 9 intragenic integrations in the antisense orientation, 22 intergenic integrations and 3 integrations within repeated sequences. 33/34 nonsense integrations were found in clones carrying multiple insertions of the gene trap construct. Since only one polyA trap event is required to confer antibiotic resistance to the host cell, the nonsense integrations should not necessarily be regarded as trap events.

One clone (c10) contained 3 transposon integrations (arrow in Fig. 4D), one of which was located upstream of the first coding exon of a mouse chromosome (chr.) 11 gene (NM.025510). The other two integrations occurred on chr. 14 and 18, within unannotated sequences (one of which contains at least one intron–exon junction) (Supplemental Fig. 5A). By RT-PCR, we analyzed the levels of fusion transcripts (Fig. 4E and Supplemental Fig. 5B) and of the corresponding endogenous transcripts, if present (Fig. 4G and Supplemental Fig. 5C). While NM.025510 is expressed at low levels in wt NS cells, the other two transcripts are undetectable in the parental line. These results demonstrate the ability of the CAG exogenous promoter to activate the transcription of genes that are normally silenced in proliferating NS cells. The trapped sequences were further examined: NM.025510 encodes a Mn<sup>2+</sup>-dependent ADP-ribose/CDP-alcohol pyrophosphatase (Canales et al., 2008). Sequencing of the fusion transcript showed that the splicing event occurred efficiently across the 3' IR/DR sequence, generating a dicistronic mRNA that contains the fusion neomycin-CFP transcript, the IRES sequence, and the NM.025510 mRNA. By quantitative real time RT-PCR, we observed a 21-fold overexpression of the trapped gene compared to the corresponding levels measured in the parental line (Fig. 4H), confirming the efficacy of our gain-of-function construct in modifying the expression of trapped genes. Importantly, this demonstrates that the presence of the IR/DR sequences that delimit the SB transposable element does not interfere in any relevant way with a robust activation of downstream gene expression, or with an efficient recruitment of a downstream poly(A) sequence.

Besides overexpressing coding regions, a poly(A) trapping approach may lead to the upregulation of noncoding sequences of importance in neural stem cell biology. Through a computerized prediction method, we searched c10 integrations for the possible presence of putative microRNA (miRNA) sequences. According to Ambros et al. (2003) we set 2 main inclusion criteria for the analyzed sequences: (i) formation of a ~80-nt stem-loop secondary structure, as predicted by the RNAfold software (Gruber et al.,



**Fig. 4.** Expression cloning in NS cells by a poly(A) trap-mediated gain-of-function approach. (A) Schematic representation of the poly(A) trap transposon. This construct contains a CAG promoter (CAG prom), a neomycin resistance gene (Neo) fused to a fluorescent reporter gene (Cyan Fluorescent Protein, CFP), an internal ribosome entry site (IRES) followed by a splicing donor site (SD). (B) Cresyl violet staining of neomycin resistant clones obtained by nucleofecting NS cells with the poly(A) trap transposable element (poly(A) trap) in the presence or absence of a helper plasmid encoding M3A transposase. Transfected cells were plated at the indicated cell densities (2500, 5000 and 10,000 cells/plate). In this experiment, the transposon to transposase molar ratio was 10:1. (C) Histogram showing the average number of neomycin resistant clones ( $n = 3$ , \*  $p < 0.005$ , \*\*  $p < 0.0005$ ) obtained under the transfection conditions described in (B). (D) A representative Southern blot of NS cell clones deriving from co-transfection experiments (poly(A) trap + M3A). DNAs were digested with *HindIII* and hybridized with the probe sketched in (A) to reveal a single junction fragment for each transposon integration. The arrow indicates bands derived from transposon integrations in the genome of clone 10 (c10, see text for details). (E) RT-PCR analysis on cDNA derived by retrotranscription of RNA from the parental cell line (NS) and c10; B = blank (H<sub>2</sub>O). PCR primers were designed on the IRES sequence (forward primer) and on the trapped transcript sequence on chromosome 11 (reverse primer) (see Section 2). The chr. 11 integration (chr. 11) gives rise to fusion transcript between transposon sequence and the trapped gene; (F) Stem-loop secondary structure predicted by the RNAfold software of the miRNA hairpin identified on chr. 11, downstream of the transposon integration site. Different colors indicate the base-pair probabilities as shown in legend (blue = low probability; red = high probability). (G) RT-PCR analysis on cDNA derived by retrotranscription of RNA from the parental cell line (NS) and c10, B = blank (H<sub>2</sub>O). Forward and reverse primers were designed on the NM.025510 endogenous gene and on the hairpin-containing sequence (hcs) localized in intron 2 of the NM.025510 transcript. The CAG promoter leads to the over-expression of the downstream sequence. (H) RT-qPCR analysis quantifying the fold increase of hcs and NM.025510 expression in c10 (black columns) compared to the endogenous levels in the parental line (white columns); \*  $p < 0.05$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

2008); (ii) detection of distinct ~22-nt sequence in mouse small RNA libraries. Furthermore, we imposed four additional features: (iii) absence of the putative mature ~22-nt sequence in small RNA libraries from *Dicer*-deleted ES cell line; (iv) presence of the putative pre-miRNA sequence in mouse ESTs databases; (v) similarity to an annotated pre-miRNA sequence aligned using the Smith–Waterman algorithm; and (vi) classification of the hairpin

sequence as an actual miRNA precursor using miPred (Jiang et al., 2007), an *ab initio* hairpin classifier that help to distinguish *bona fide* pre-miRNAs from randomly occurring hairpins. Our *in silico* analysis predicted the presence of a hairpin-containing sequence downstream of two transposon integrations (Supplemental Fig. 5A). The hairpin-containing sequence on chr. 14 (Supplemental Fig. 5A) forms a stable secondary structure (Supplemental Fig. 5D)

**Table 1**  
Genomic distribution and analysis of polyA trap transposon integrations.

Genomic region	No.	Chr.	Trapped gene	ID number	Integration site	Direction	Function
Intragenic <sup>a</sup>	26	19	Cuedc2	NM.024192	ex5	Sense	Ubiquitination
		15	Nell2	NM.016743	int 3–4	Sense	Cell adhesion
		19	RIKEN cDNA	AK006056	int 5–6	Sense	Unknown function
		5	Magi2	NM.015823	int 2–3	Sense	Signal transduction
		1	Ipo9	NM.153774	int 5–6	Antisense	Intracellular transport
		11	alcohol pyrophosphatase	NM.025510.3	int 2–3	Sense	Unknown function
		17	RIKEN cDNA	NM.028576	int 3–4	Sense	Unknown function
		10	RIKEN cDNA	NM.027878	3'-UTR	Antisense	Unknown function
		7	Mical2	NM.177282	int 1–2	Sense	Metal ion binding protein
		15	Cacng2	NM.007583	int 1–2	Sense	Calcium channel activity
		8	Rasa	NM.009025	int 7–8	Sense	Intracellular signal cascade
		16	Tiami1	NM.009384	ex 25	Antisense	T-cell lymphoma invasion and metastasis 1
		14	EST sequence	ENSMUSESTG00000003987	int 3–4	Antisense	Unknown function
		3	Dnajb14 (Hsp40)	AK160961	Int 3–4	Sense	Heat shock protein
		10	RIKEN cDNA	NM.027878	5'-UTR	Antisense	Unknown function
		11	Flt4	NM.008029	Int 27/28	Sense	Signal transduction
		4	Tceb3	NM.013736	Int 5–6	Sense	Regulation of transcription
		4	Mdn1	NM.001081392	Ex 93	Antisense	Cytoplasmatic protein
		1	Tmeff2	NM.019790	int 5–6	Sense	Transmembrane protein
		10	EST sequence	ENSMUSESTG00000011782	Int 1–2	Antisense	Unknown function
		3	Alg5	NM.025442.2	Int 7–8	Sense	glucosyltransferase
		3	EST sequence	ENSMUSETT00000035836	5'-UTR	Antisense	Unknown function
		5	Pmpcb	NM.028431	Int 4–5	Sense	Catalytic activity in mitochondria
		X	ll1rap11	AK081272	Int 1–2	Antisense	Transmembrane receptor activity
		8	Fcho1	NM.028715	–7.3 kb	Sense	FCH domain only 1
		4	Fam167b	NM.182783	–4 kb	Sense	Seq. similarity 167, member B
Intergenic <sup>b</sup>	22						
Repeated	3						

<sup>a</sup> Within transcribed region or up to 10 kb upstream of TSS.

<sup>b</sup> >10 kb upstream of TSS.

and is similar to *mmu-mir-1196*. However, the putative mature sequence was not found in the small RNA libraries that we analyzed. On chr. 18 we did not find any annotated sequence or putative non-coding gene. On the other hand, the hairpin-containing sequence identified on chr. 11 (Supplemental Fig. 6A) satisfies all requested criteria (Supplemental Fig. 6B–F). In particular, the putative pre-miRNA sequence forms a stable secondary structure (Fig. 5F and Supplemental Fig. 6B) and is represented several times in two *Dicer*<sup>+/+</sup> small RNA libraries (GEO: GSM237108; GSM237110) but not in a *Dicer*<sup>-/-</sup> library (GEO: GSM237109). The predicted miRNA is contained in four expressed sequence tags from visual cortex (2) and ES cell (1) library (Supplemental Fig. 6C). Furthermore, the 18 nt hairpin has been found in multiple copies in a library of Microprocessor-independent miRNAs (Babiarz et al., 2008). Taken together, these findings strongly suggest that the chr. 11 trapped sequence contains an actual miRNA of as yet uncharacterized function. The expression of the putative chr. 11 miRNA was analyzed by RT-PCR (Fig. 4G) and RT-qPCR in c10 cells and the parental NS cell line (Fig. 4H). Our data confirm a significant over-expression of the putative miRNA gene in c10 cells compared to the parental NS line ( $p < 0.05$ ). Our results suggest that transposon tagging provides a suitable approach for trapping and identification of coding and noncoding sequences in NS cells. By simply abolishing the IRES sequence in our PolyA vector, our transposon could be used even more stringently for expression cloning of noncoding small RNAs.

### 3.6. Transposon-mediated gene trapping in NS cells

Gene trapping has been largely used *in vivo* to ablate gene function and to analyze promoter activity of the trapped gene in various animal models. To this end, we also generated a gene trap transposon (Fig. 5A) containing a splicing acceptor site (SA) followed by a stop signal in 3 different frames, an IRES sequence, a GFP-Neo fusion gene and a polyadenylation site. By integrating within a given gene, this transposon will interrupt the coding sequence and

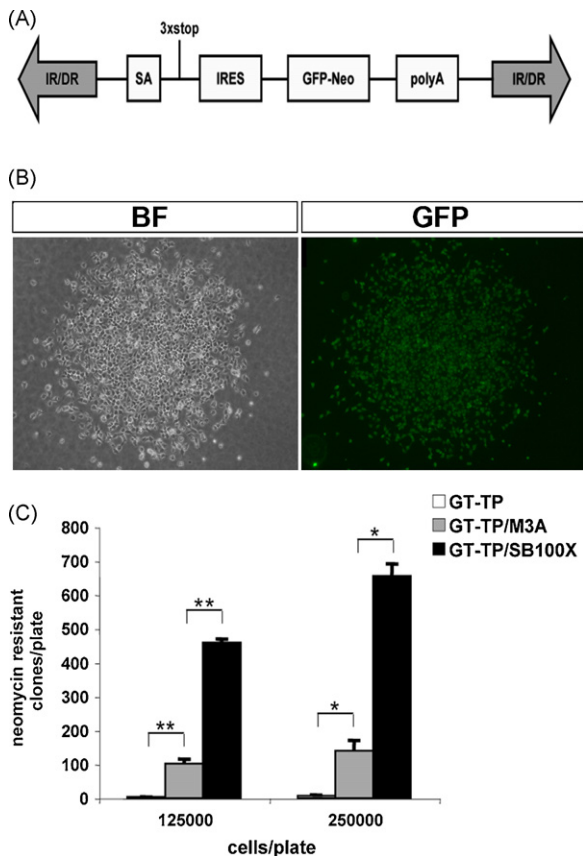
express the fluorescent and selectable marker under control of the trapped promoter (sketched in Supplemental Fig. 7). First, to test the efficiency of our gene trap approach in NS cells, we nucleofected them with the gene trap transposon either with or without SB transposase helper plasmid (M3A or SB100X). Transfected cells were plated at two different densities (125,000 and 250,000 cells/plate) and cultured under G418 selection. GFP expression was analyzed in neomycin resistant clones (Fig. 5B). After 12 days of selection, clones were stained with cresyl violet and counted (Fig. 5C). At a cell density of 125,000 cells/plate, the M3A transposase generated about 100 trapped clones (0.08% of plated cells), while, the SB100X transposase produced about 460 clones (0.38% of plated cells). The greater efficiency of SB100X transposase in this experimental scheme could be related to its ability to generate more transposon integrations per clone, thus increasing the probability of trapping at least one active gene promoter. These results indicate that SB transposons are appropriate tools for the execution of gene trapping screens in NS cells.

## 4. Discussion

NS cells provide an easily accessible model for the *in vitro* analysis of stem cell biology and neurodevelopmental pathways. Our results indicate that transposon-based constructs may provide an efficient tool to conduct functional screenings in NS cells. DNA transposons are safe to handle and can be manipulated as conveniently as any plasmid vectors. They can be easily upgraded by standard molecular techniques by adding new or modified features to them. Most importantly, the cargo capacity of SB transposons can be extended to integrate large (>10 kb) inserts (Ivics et al., 2009). NS cells can be easily and efficiently nucleofected, permitting the use of non-viral vectors to manipulate their biology.

Transposons may be engineered to contain promoters and enhancers, splicing donors and acceptors, and polyadenylation sequences, since none of these signals affect their basic biology. Our





**Fig. 5.** Gene trapping in NS cells. (A) Schematic representation of the gene trap transposon. This construct contains a splicing acceptor site (SA) followed by a stop signal in 3 different frames, a green fluorescence protein sequence (GFP) fused to a neomycin resistance gene (Neo) and a polyadenylation site; (B) Bright field (BF) and green fluorescent protein (GFP) images of neomycin resistant clones obtained by transfecting NS cells with gene trap transposon plus M3A transposase or SB100X transposase. In this experiment, the transposon to transposase molar ratio was 10:1; (C) Histogram showing the average number of neomycin resistant clones ( $n=3$ , \*  $p < 0.05$ , \*\*  $p < 0.005$ ) obtained transfecting NS cells with the gene trap transposon alone or in presence of M3A or SB100X transposase. Transfected cells were plated at 125,000 and 250,000 cells per plate.

results also indicate that, after integration, these signals will work efficiently across the IR/DR repeats that delimit SB transposons on each side.

#### 4.1. Sleeping beauty transposons in NS cells

Hyperactive *Sleeping Beauty* transposons (Baus et al., 2005; Yant et al., 2004; Zayed et al., 2004) represent highly effectual tools for the manipulation of the NS cell genome. In this paper, their power is measured focusing on two parameters. The first one is the increase in clone formation obtained by co-transfecting the transposable element and transposase vs. the number of clones obtained in the absence of helper plasmid. This gauges the potency of active transposition vs. naked DNA integration. In this respect, SB transposons perform very effectively in NS cells, allowing 8% of all transfected cells to be expanded clonally in G418 selection. The second parameter examined in this study is the multiplicity of integrations in each transfected cell. SB transposons, transfected in the presence of optimized concentrations of the M3A transposase, produce a reasonably low number of integrations/clone, facilitating the isolation of flanking sequences. In this respect, the hyperactive SB100X transposase reveals its greater activity by promoting a higher average number of integrations per clone in comparison with M3A. However, this may make it cumbersome to isolate a functionally

relevant integration from a background of silent ones. This problem could be circumvented by removing transposon integrations by Cre- or Flp-mediated recombination. After subcloning recombinase treated cells, phenotypes could be correlated with the pattern of residual transposon integrations.

#### 4.2. Piggybac vs. Sleeping Beauty: efficiency and multiple integrations

The efficiency of SB transposons was also compared to that of another powerful reagent, the *piggyBac* (PB) transposon (Ding et al., 2005; Fraser et al., 1996). The latter system was developed in parallel to SB transposons, and features a unique characteristic: the fact that it can be removed flush from its genomic location, leaving no molecular signature of its previous integration (Elick et al., 1996; Fraser et al., 1996), unlike the SB element (Fischer et al., 2001). PB transposons are equally or more potent than the SB element mobilized by the M3A transposase when used in NS cells at an optimal transposon to transposase molar ratio. However, under those conditions, PB transposons exhibit a tendency to produce a multiplicity of integrations in each cell.

#### 4.3. Stability of integrations over time

Our results also indicate that SB transposons produce stable integrations in the NS cell genome, and that expression of genes encoded by the integrated molecule can be detected after numerous passages *in vitro*. This property suggests that SB transposons could be profitably employed as tools for sustained expression of transgenes and reporters in this cell line, and not just as random mutagenesis tools in the context of functional screening approaches.

#### 4.4. Use in gain-of-function screens

In the present report we provide a proof of principle of the suitability of SB transposons as tools for poly(A) trapping in NS cells. Poly(A) trapping is a well established approach to random mutagenesis in various systems, and can be applied to the over-expression of coding and noncoding sequences. Other potential approaches are available, including the use of transposons containing selectable markers and dominant enhancers promoting the expression of genes flanking the genomic site of transposon integration. Such transposable elements can be engineered to promote RNA polymerase II or III-dependent transcription, and can be adapted to a variety of applications. In NS cells, gain-of-function tools can be used to identify determinants of multipotency and GF-independent self-renewal, and in many other developmental neurobiology applications. These may include coding and noncoding genes; in fact, in our survey of poly(A) trap transposon integrations, we encountered two overexpressed noncoding sequences, one of these exhibiting *in silico* thermodynamic features compatible with a novel putative pre-miRNA elements.

#### 4.5. Use in gene trapping and/or loss-of-function screens

In the present paper we also provide evidence of the feasibility of transposon-mediated gene trapping in the same cell line. This type of approach, based on a reporter gene trap, could be applied usefully to a search for genes activated in NS cells by pharmacological compounds or extracellular molecules that modulate neurogenesis, or again to create a collection of reporter lines for studies of NS cell differentiation. Furthermore, gene traps could be used to generate homozygous mutants, employing NS cells derived from specific genetic backgrounds, e.g. *Blm1* null mutants, in which heterozygous integrations have a high probability of being reduced

to homozygosity due to sister chromatid exchange (Wang et al., 2009). Finally, given the feasibility of producing mouse iPSCs from NS cells (Kim et al., 2009; Kim et al., 2008; Silva et al., 2008), the recombinant NS cell clones isolated by gene trapping could provide a source of mouse iPSCs and be used to produce reporter knock-in and knock-out animal models.

## 5. Conclusions

The present report illustrates the power and feasibility of transposon-mediated mutagenesis in a novel and highly plastic model system, adherently expandable Neural Stem cells, proposing SB and other transposons as optimal tools for the execution of high throughput functional screens for determinants of neural stem cell biology and neural cell type specification. The large cargos hosted by transposons, their ease of manipulation and the high efficiency of DNA-mediated gene transfer (nucleofection) in NS cells make DNA transposons a tool of choice for the application of high throughput screening approaches in this innovative cellular model.

## Acknowledgements

We thank Maurizio Binetti for cell counts and Giulio Simonutti for image analysis. This work was supported by Fondazione CARIPO (Network Operativo per la Biomedicina di Eccellenza in Lombardia) project entitled “A genetic toolkit for the analysis of Neural Stem cells – acronym: Mouse NS-toolkit”. Additional support came from the European Union Framework Program VII through the EuroSyStem consortium contract to E.C. and G.G.C. The development of hyperactive SB transposases was supported by grant 0313435A from the Bundesministerium für Bildung und Forschung (NGFN-2) to Z.I.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jbiotec.2010.07.027.

## References

- Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., Tuschl, T., 2003. A uniform system for microRNA annotation. *RNA* 9, 277–279.
- Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., Belloch, R., 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other microprocessor-independent, dicer-dependent small RNAs. *Genes Dev.* 22, 2773–2785.
- Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Rudnev, D., Evangelista, C., Kim, I.F., Soboleva, A., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Muerter, R.N., Edgar, R., 2009. NCBI GEO: archive for high throughput functional genomic data. *Nucleic Acids Res.* 37, D885–890.
- Baus, J., Liu, L., Heggstad, A.D., Sanz, S., Fletcher, B.S., 2005. Hyperactive transposase mutants of the sleeping beauty transposon. *Mol. Ther.* 12, 1148–1156.
- Canales, J., Fernandez, A., Ribeiro, J.M., Cabezas, A., Rodrigues, J.R., Cameselle, J.C., Costas, M.J., 2008. Mn<sup>2+</sup>-dependent ADP-ribose/CDP-alcohol pyrophosphatase: a novel metallophosphoesterase family preferentially expressed in rodent immune cells. *Biochem. J.* 413, 103–113.
- Cattoglio, C., Facchini, G., Sartori, D., Antonelli, A., Miccio, A., Cassani, B., Schmidt, M., von Kalle, C., Howe, S., Thrasher, A.J., Aiuti, A., Ferrari, G., Recchia, A., Mavilio, F., 2007. Hot spots of retroviral integration in human CD34<sup>+</sup> hematopoietic cells. *Blood* 110, 1770–1778.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., Smith, A., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 113, 643–655.
- Conti, L., Pollard, S.M., Gorb, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.L., Cattaneo, E., Smith, A., 2005. Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biol.* 3, e283.
- Cui, Z., Geurts, A.M., Liu, G., Kaufman, C.D., Hackett, P.B., 2002. Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon. *J. Mol. Biol.* 318, 1221–1235.
- Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., Xu, T., 2005. Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* 122, 473–483.
- Dupuy, A.J., Akagi, K., Largaespada, D.A., Copeland, N.G., Jenkins, N.A., 2005. Mammalian mutagenesis using a highly mobile somatic sleeping beauty transposon system. *Nature* 436, 221–226.
- Elick, T.A., Bauser, C.A., Fraser, M.J., 1996. Excision of the piggyBac transposable element in vitro is a precise event that is enhanced by the expression of its encoded transposase. *Genetica* 98, 33–41.
- Fischer, S.E., Wienholds, E., Plasterk, R.H., 2001. Regulated transposition of a fish transposon in the mouse germ line. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6759–6764.
- Fraser, M.J., Ciszczon, T., Elick, T., Bauser, C., 1996. Precise excision of TTA-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.
- Goodier, J.L., Davidson, W.S., 1994. Tc1 transposon-like sequences are widely distributed in salmonids. *J. Mol. Biol.* 241, 26–34.
- Griffiths-Jones, S., Saini, H.K., van Dongen, S., Enright, A.J., 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 36, D154–158.
- Gruber, A.R., Lorenz, R., Bernhart, S.H., Neubock, R., Hofacker, I.L., 2008. The Vienna RNA websuite. *Nucleic Acids Res.* 36, W70–74.
- Ivics, Z., Hackett, P.B., Plasterk, R.H., Izsvak, Z., 1997. Molecular reconstruction of sleeping beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* 91, 501–510.
- Ivics, Z., Li, M.A., Mates, L., Boeke, J.D., Nagy, A., Bradley, A., Izsvak, Z., 2009. Transposon-mediated genome manipulation in vertebrates. *Nat. Methods* 6, 415–422.
- Izsvak, Z., Khare, D., Behlke, J., Heinemann, U., Plasterk, R.H., Ivics, Z., 2002. Involvement of a bifunctional, paired-like DNA-binding domain and a transpositional enhancer in Sleeping Beauty transposition. *J. Biol. Chem.* 277, 34581–34588.
- Jiang, P., Wu, H., Wang, W., Ma, W., Sun, X., Lu, Z., 2007. MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. *Nucleic Acids Res.* 35, W339–344.
- Kim, J.B., Sebastiano, V., Wu, G., Arauzo-Bravo, M.J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrlich, M., van den Boom, D., Meyer, J., Hubner, K., Bernemann, C., Ortmeier, C., Zenke, M., Fleischmann, B.K., Zaehres, H., Scholer, H.R., 2009. Oct4-induced pluripotency in adult neural stem cells. *Cell* 136, 411–419.
- Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Arauzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M., Scholer, H.R., 2008. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 454, 646–650.
- Kustikova, O., Fehse, B., Modlich, U., Yang, M., Dullmann, J., Kamino, K., von Neuhoff, N., Schlegelberger, B., Li, Z., Baum, C., 2005. Clonal dominance of hematopoietic stem cells triggered by retroviral gene marking. *Science* 308, 1171–1174.
- Kustikova, O.S., Geiger, H., Li, Z., Brugman, M.H., Chambers, S.M., Shaw, C.A., Pike-Overzet, K., de Ridder, D., Staal, F.J., von Keudell, G., Cornils, K., Nattamai, K.J., Modlich, U., Wagemaker, G., Goodell, M.A., Fehse, B., Baum, C., 2007. Retroviral vector insertion sites associated with dominant hematopoietic clones mark “stemness” pathways. *Blood* 109, 1897–1907.
- Liang, Q., Kong, J., Stalker, J., Bradley, A., 2009. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. *Genesis* 47, 404–408.
- Liu, L., Sanz, S., Heggstad, A.D., Antharam, V., Notterpek, L., Fletcher, B.S., 2004. Endothelial targeting of the sleeping beauty transposon within lung. *Mol. Ther.* 10, 97–105.
- Mates, L., Chuah, M.K., Belay, E., Jerchow, B., Manoj, N., Acosta-Sanchez, A., Grzela, D.P., Schmitt, A., Becker, K., Matrai, J., Ma, L., Samara-Kuko, E., Gysemans, C., Pryputniowicz, D., Miskey, C., Fletcher, B., Vandendriessche, T., Ivics, Z., Izsvak, Z., 2009. Molecular evolution of a novel hyperactive sleeping beauty transposase enables robust stable gene transfer in vertebrates. *Nat. Genet.* 41, 753–761.
- Niwa, H., Yamamura, K., Miyazaki, J., 1991. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* 108, 193–199.
- Onorati, M., Camnasio, S., Binetti, M., Jung, C.B., Mioretti, A., Cattaneo, E., 2009. Neuropotent self-renewing neural stem (NS) cells derived from mouse induced pluripotent stem (iPS) cells. *Mol. Cell. Neurosci.*
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* 85, 2444–2448.
- Piccolo, S., Sasai, Y., Lu, B., De Robertis, E.M., 1996. Dorsal-ventral patterning in *Xenopus*: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* 86, 589–598.
- Pollard, S.M., Conti, L., Sun, Y., Goffredo, D., Smith, A., 2006. Adherent neural stem (NS) cells from fetal and adult forebrain. *Cereb. Cortex* 1 (Suppl. 16), i112–120.
- Rhead, B., Karolchik, D., Kuhn, R.M., Hinrichs, A.S., Zwiweig, A.S., Fujita, P.A., Diekhans, M., Smith, K.E., Rosenbloom, K.R., Raney, B.J., Pohl, A., Pheasant, M., Meyer, L.R., Learned, K., Hsu, F., Hillman-Jackson, J., Harte, R.A., Giardine, B., Dreszer, T.R., Clawson, H., Barber, G.P., Haussler, D., Kent, W.J., 2010. The UCSC Genome Browser database: update. *Nucleic Acids Res.* 38, D613–619.
- Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., Smith, A., 2008. Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol.* 6, e253.
- Smith, W.C., Harland, R.M., 1992. Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in *Xenopus* embryos. *Cell* 70, 829–840.
- Starr, T.K., Allaei, R., Silverstein, K.A., Staggs, R.A., Sarver, A.L., Bergemann, T.L., Gupta, M., O’Sullivan, M.G., Matise, I., Dupuy, A.J., Collier, L.S., Powers, S., Oberg, A.L., Asmann, Y.W., Thibodeau, S.N., Tessarollo, L., Copeland, N.G., Jenkins, N.A., Cormier, R.T., Largaespada, D.A., 2009. A transposon-based genetic screen in mice identifies genes altered in colorectal cancer. *Science* 323, 1747–1750.
- Strahm, Y., Powell, D., Lefevre, C., 2006. EST-PAC a web package for EST annotation and protein sequence prediction. *Source Code Biol. Med.* 1, 2.

- Sun, Y., Pollard, S., Conti, L., Toselli, M., Biella, G., Parkin, G., Willatt, L., Falk, A., Cattaneo, E., Smith, A., 2008. Long-term tripotent differentiation capacity of human neural stem (NS) cells in adherent culture. *Mol. Cell. Neurosci.* 38, 245–258.
- Wang, W., Bradley, A., Huang, Y., 2009. A piggyBac transposon-based genome-wide library of insertionally mutated Blm-deficient murine ES cells. *Genome Res.* 19, 667–673.
- Wilson, M.H., Coates, C.J., George Jr., A.L., 2007. PiggyBac transposon-mediated gene transfer in human cells. *Mol. Ther.* 15, 139–145.
- Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M., Kaji, K., Sung, H.K., Nagy, A., 2009. PiggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 458, 766–770.
- Yant, S.R., Park, J., Huang, Y., Mikkelsen, J.G., Kay, M.A., 2004. Mutational analysis of the N-Terminal DNA-binding domain of sleeping beauty transposase: critical residues for DNA-binding and hyperactivity in mammalian cells. *Mol. Cell. Biol.* 24, 9239–9247.
- Yant, S.R., Wu, X., Huang, Y., Garrison, B., Burgess, S.M., Kay, M.A., 2005. High-resolution genome-wide mapping of transposon integration in mammals. *Mol. Cell. Biol.* 25, 2085–2094.
- Zayed, H., Izsvak, Z., Walisko, O., Ivics, Z., 2004. Development of hyperactive sleeping beauty transposon vectors by mutational analysis. *Mol. Ther.* 9, 292–304.
- Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31, 3406–3415.