nature
genetics

Molecular evolution of a novel hyperactive *Sleeping Beauty* transposase enables robust stable gene transfer in vertebrates

Lajos Mátés[1,6], Marinee K L Chuah[2,6], Eyayu Belay[2], Boris Jerchow[1], Namitha Manoj[1], Abel Acosta-Sanchez[2], Dawid P Grzela[1], Andrea Schmitt[1], Katja Becker[1], Janka Matrai[2], Ling Ma[2], Ermira Samara-Kuko[2], Conny Gysemans[3], Diana Pryputniewicz[1], Csaba Miskey[1], Bradley Fletcher[4], Thierry VandenDriessche[2], Zoltán Ivics[1] & Zsuzsanna Izsvák[1,5]

The *Sleeping Beauty* (SB) transposon is a promising technology platform for gene transfer in vertebrates; however, its efficiency of gene insertion can be a bottleneck in primary cell types. A large-scale genetic screen in mammalian cells yielded a hyperactive transposase (SB100X) with ~100-fold enhancement in efficiency when compared to the first-generation transposase. SB100X supported 35–50% stable gene transfer in human CD34+ cells enriched in hematopoietic stem or progenitor cells. Transplantation of gene-marked CD34+ cells in immunodeficient mice resulted in long-term engraftment and hematopoietic reconstitution. In addition, SB100X supported sustained (>1 year) expression of physiological levels of factor IX upon transposition in the mouse liver *in vivo*. Finally, SB100X reproducibly resulted in 45% stable transgenesis frequencies by pronuclear microinjection into mouse zygotes. The newly developed transposase yields unprecedented stable gene transfer efficiencies following nonviral gene delivery that compare favorably to stable transduction efficiencies with integrating viral vectors and is expected to facilitate widespread applications in functional genomics and gene therapy.

DNA transposons can be viewed as natural gene delivery vehicles that integrate into the host genome through a 'cut-and-paste' mechanism. These mobile DNA elements encode a transposase flanked by inverted terminal repeats (ITRs) that contain the transposase binding sites necessary for transposition. Any gene of interest flanked by these ITRs can undergo transposition in the presence of the transposase supplied in *trans* (**Fig. 1**). These features make transposons ideal molecular tools for transgenesis, gene therapy and functional genomics. *Sleeping Beauty* (SB) is a synthetic transposable element that has been

generated by 'reverse engineering' from defective copies of an ancestral Tc1/*mariner*-like transposon in fish[1]. SB is active in different somatic tissues of a wide range of vertebrate species including humans, as well as in the germline of fish, frogs, mice and rats[2]. SB has been shown to provide long-term transgene expression and represents a promising technology platform for gene transfer in vertebrates[3].

In evolutionary terms, the SB transposon represents a successful element that was able to colonize several fish genomes millions of years ago[1]. However, even successful transposons have not been selected for the highest possible activity in nature, as they have to coexist with their hosts and, consequently, there is strong selective pressure to avoid insertional mutagenesis of essential genes. Thus, enhancing transpositional activity is one of the main targets for transposon vector development. To date, almost every single amino acid has been changed in the transposase in an attempt to increase its activity. Three main strategies have been applied to derive hyperactive mutants of the SB transposase: 'importing' amino acids and small blocks of amino acids from related transposases[4–6], systematic alanine-scanning[7], and rational replacement of selected amino acid residues[5]. Together, these studies have yielded about 16 amino-acid replacements, each resulting in a relatively modest increase in transpositional activities. Notably, the hyperactivity of the SB transposase mutants selected in immortalized cell lines does not necessarily translate to efficient stable gene transfer in primary cells *in vivo*[6,7].

Here we report a high-throughput genetic screen to derive hyperactive SB transposases generated by *in vitro* evolution. We demonstrate the utility of the newly developed transposase in stable gene transfer in hard-to-transfect CD34+ cells enriched in bona fide hematopoietic stem cells and in transgenesis in mouse embryos as well as in the liver of adult mice. The present study validates the use of hyperactive transposases for gene transfer applications in various primary target

[1]Max Delbrück Center for Molecular Medicine, Berlin, Germany. [2]Flanders Institute for Biotechnology (VIB), Vesalius Research Center, University of Leuven, Leuven, Belgium. [3]Department of Experimental Medicine, Laboratory for Experimental Transplantation, University of Leuven, Belgium. [4]Department of Pharmacology and Therapeutics, University of Florida, College of Medicine, Gainesville, Florida, USA. [5]Institute of Biochemistry, Biological Research Center of the Hungarian Academy of Sciences, Szeged, Hungary. [6]These authors contributed equally to this work. Correspondence should be addressed to Z.I. (zizsvak@mdc-berlin.de) or T.V. (thierry.vandendriessche@med.kuleuven.be).
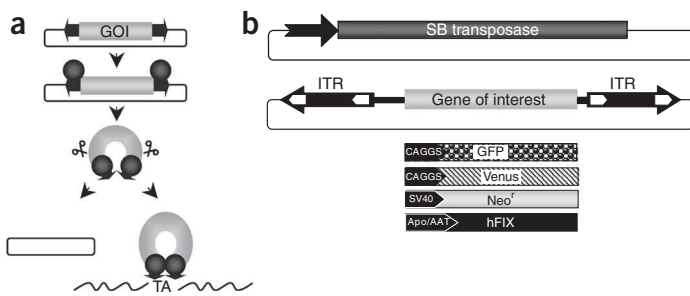
Figure 1 The *Sleeping Beauty* (SB) transposon system. (**a**) Schematic representation of SB transposition. In the two-component system, a gene of interest (GOI) can be positioned between inverted terminal repeats (ITRs, black arrows) of the transposon that is mobilized by the transposase protein (sphere) supplemented in *trans*. White arrows in the ITRs represent transposase binding sites. In a typical gene transfer experiment the transposon is mobilized from a plasmid to a genomic target site (TA). (**b**) GOIs used in this work.

cells following *ex vivo* or *in vivo* gene delivery with broad implications for functional genomics and gene therapy.

RESULTS

A phylogenetic approach to hyperactive transposases

We hypothesized that hyperactive SB transposases could be generated by incorporating phylogenetically conserved amino acids from related transposases, belonging to the Tc1 transposon family, into SB (see an example in **Fig. 2**). The corresponding mutations were engineered into the SB transposase gene, and 53 variants were tested for activity using a cell culture–based transposition assay[1] in human HeLa cells. We found that 25 of these substitutions resulted in hyperactivity as compared to the original SB transposase, underscoring the biological relevance of our phylogenetic approach. Subsequently, this collection of substitutions was supplemented with hyperactive mutations reported previously[4–7], resulting in a library of 41 clones, each containing a single hyperactive mutation (**Fig. 3a** and **Supplementary Table 1** online).

High-throughput transposase screening

None of the individual mutations described above resulted in hyperactivity higher than fourfold (**Supplementary Table 1**). However, combinations of individual mutations could potentially result in additive or synergistic effects[4–7]. The number of combinations can be calculated from $c = n!/(n − k)!k!$ (where n is the number of individual mutations and k is the desired number of combinations per transposase). Because of the large number of possible combinations of 41 variants, it was necessary to develop a high-throughput, PCR-based, DNA-shuffling strategy and screening in mammalian cells. A library of mutant transposase genes was established with an average number of two mutations per gene in the hope of identifying pairwise, synergistic combinations (**Fig. 3b**). The quality of the library and the clonal distribution of the variants were determined by sampling 45 clones, and the average number of mutations per clone was found to be 2.2 (**Supplementary Fig. 1** online). The best 38 clones isolated from 2,000 clones screened in total showed up to 25-fold higher activity as compared to the original SB transposase (**Fig. 3b**). On average, the best 8 of the 38 clones carried 3.6 mutations per gene, supporting our hypothesis that the key to achieve a higher degree of hyperactivity is finding the right combination of multiple hyperactive variants. We observed that 16 of the 41 hyperactive mutations were never recovered in combinations ('unfriendly' mutations), whereas 25 appeared repeatedly in the most hyperactive versions, indicating that

only a fraction of hyperactive combinations are compatible with others ('friendly' mutations) (**Fig. 3b**). Notably, the clones 6A5, 3D5, 12B1 and 2G6 had a similar level of hyperactivity, but different numbers of mutations (5, 4, 3 and 2, respectively, **Supplementary Fig. 2** online), indicating that the nature of combinations can be additive (for example, 6A5) or synergistic (for example, 2G6). These four clones were used as a base for further rounds of manual combinations with friendly mutations identified by the screen (**Supplementary Fig. 2**). In summary, our strategy yielded a series of mutants showing ~10- to 80-fold hyperactivity (**Fig. 3b**, **Supplementary Table 1** and **Supplementary Figs. 2** and **3** online). The most hyperactive version, hereafter referred to as SB100X, contained six combinatorial units that yield nearly 4,500,000 possible combinations ($k = 6$; $c6 = 4,496,388$) (**Supplementary Table 1**), underscoring the necessity of combining 'high-throughput' and 'analytical' strategies.

Characterization of the SB100X transposase in HeLa cells

We next characterized SB100X with respect to its performance in gene transfer experiments and biochemical properties. In terms of generating stable transgenic cells, as determined by colony-forming assays in HeLa cells transfected with varying transposon DNA dosages, SB100X gave rise to up to 35% transgenesis frequencies as opposed to 8% obtained with the original SB transposase at the highest transposon concentration (**Fig. 4a**). Notably, SB100X maintained its high transgenic frequencies at limiting amounts of transposon template, as
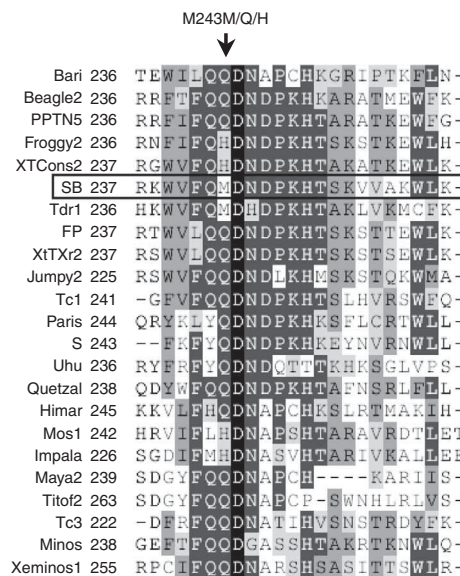

Figure 2 Selecting phylogenetic variations in the SB transposase. Inclusion of phylogenetically conserved amino acids from related transposases may generate hyperactive variants. The amino-acid position 243 in SB serves here as a typical example. The amino acid in this position varies in the Tc1 family of transposases (M, Q or H). Both of the variations (Q and H) distinct from the M present in the original SB transposase increased its transpositional activity (see **Supplementary Table 1**).

Figure 3 The high-throughput transposase screen. (**a**) The 41 hyperactive variants are distributed throughout the transposase, with three exceptions: no hyperactive mutations were recovered in the regions between amino-acid positions 148–186, 218–242 and 317–340, located in the conserved DDE catalytic domain and the C terminus of the transposase (boxes). (**b**) Flow-chart of the hyperactive screen. In order to decrease the overrepresentation of the 'wild-type' sequences in the library[46], the combinatorial units were isolated on short DNA fragments using restriction endonuclease (REN) digestions. These ~300-bp REN fragments were broken in a random fashion by DNaseI. Populations of 30- to 70-bp DNaseI-generated DNA fragments were used to recombine randomly in a PCR-like reaction in the presence of oligonucleotides that bridged the nonoverlapping REN fragments from the first step. The recombined gene pool was amplified by primers designed to the ends of the transposase gene. The resulting library was cloned into a eukaryotic expression vector, and 2,000 transposase variants were purified by automated plasmid DNA purification. The library was screened for transpositional activity by adapting the antibiotic resistance–based transposition assay to a 96-well format in HeLa cells. Variations repeatedly recovered from the screen (11 out of 25 shown) were combined further manually, yielding SB100X.
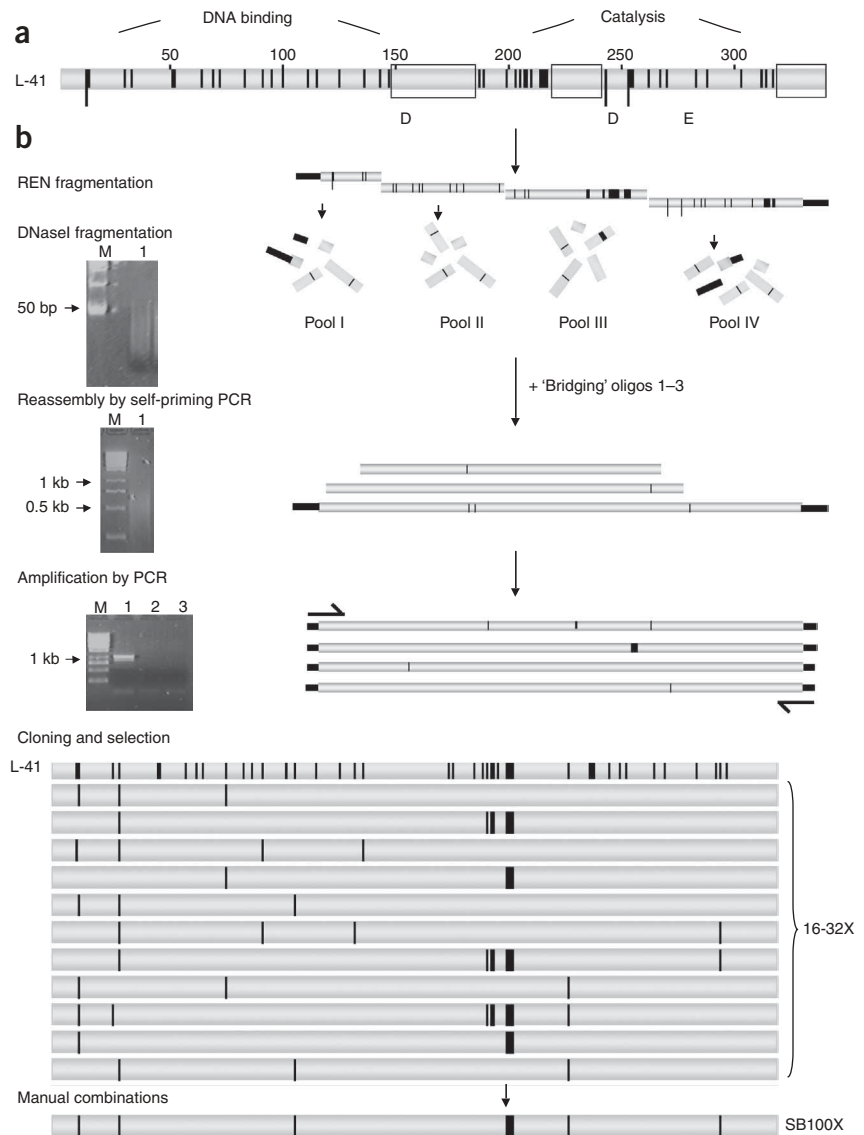
opposed to SB (**Fig. 4a**). Moreover, SB100X was more than 100-fold more potent compared to SB in mobilizing a chromosomally located element (**Fig. 4b**), a highly desirable feature for genetic screens. Thus, it seems that the difference in relative transpositional efficiencies obtained with varying substrate dosages is due to an effect elicited by cellular concentrations of transposon DNA, and that the hyperactivity of SB100X will be best manifested under conditions where the availability of the transposon DNA is limiting the transposition reaction (for example, in hard-to-transfect cell types and in chromosomal remobilization of single-unit transposons). In order to address whether the number of integrations per cell increases as a function of transposase activity, we determined copy numbers of integrated transposons obtained by SB and SB100X. SB100X generated approximately three- to fivefold more insertions than SB independent of DNA dosage, indicating that the increase in copy number was not proportional to the increase in activity (**Fig. 4c**). In comparison to the original transposase, SB100X was not significantly different in its stability, overproduction inhibition profile and affinity to the transposon ITRs (**Supplementary Fig. 4** online). However, SB100X showed markedly different subcellular localization than SB upon heat shock (**Supplementary Fig. 4e**), suggesting that the particular mutations in SB100X might affect protein folding and the complex interactions between the DNA-binding and catalytic domains that ultimately facilitate transposition. Notably, SB100X proved to be approximately sevenfold more active in colony-forming transposition assays in HeLa cells than a codon-optimized version of the *piggyBac* transposon[8], which is currently being developed as an alternate system for gene delivery in vertebrate cells[8–10] (**Supplementary Fig. 5a** online).

Generation of transgenic mice with hyperactive SB transposase
Highly efficient transgene insertion makes SB a potentially attractive tool for transgenesis. In order to investigate whether the hyperactivity established in HeLa cells is amenable to other cell types or *in vivo* conditions, we co-injected circular plasmid DNA of a Venus-tagged transposon (pT2/CAGGS-Venus) with SB100X transposase mRNA into fertilized mouse oocytes. **Figure 5** shows that when we used the SB100X transposase at the most optimal condition, 5 ng/µl transposase mRNA plus 0.4 ng/µl transposon plasmid DNA, on average 45% of the embryos scored reproducibly positive for fluorescence at day 7 postinjection. Upon transplantation into foster mice, 37% (7 out of 19 total newborns in one experiment) of the injected embryos gave rise to transgenic offspring. In general, the transgenic animals showed similarly ubiquitous overall fluorescence with no apparent sign of mosaicism (**Fig. 5** and **Supplementary Fig. 6** online). On average, transgenic animals showed 1–2 copies of integrated transposons per genome (data not shown). Thus, the hyperactive transposon system may offer a simple and highly efficient means for generating transgenic animal models.
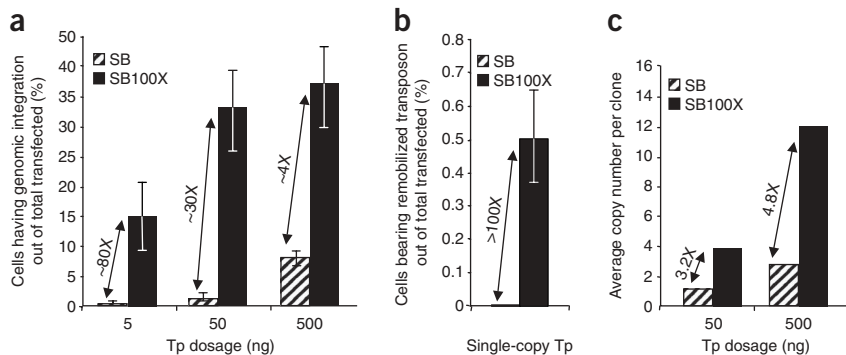
Figure 4 Characterization of the SB100X transposase in transfected human HeLa cells. (**a–c**) Comparison of SB and SB100X regarding efficacy of stable transgene integration at three different transposon (Tp) plasmid dosages (**a**); efficacy of remobilization of a single-copy chromosomal transposon by the SB and SB100X transposases (**b**); and copy numbers of transposon insertions per cell at two different transposon plasmid dosages by the SB and SB100X transposases (**c**). Equal numbers of colonies of different size were randomly picked from each experiment. The transfected amount of the transposase expression plasmids was 50 ng. Error bars, s.d.

maintained a high viability before injection ($82 \pm 7\%$, $n = 5$). We assessed stable gene transfer with respect to GFP-positive CFUs in clonogenic assays following *in vitro* differentiation into the erythroid (CFU-E), megakaryocytic (CFU-MK) and granulocyte/monocyte/macrophage (CFU-GM) lineages (**Fig. 6a**). The highest percentage of stable GFP-positive colonies was obtained with SB100X in all three lineages: $34 \pm 4\%$ of CFU-E, $45 \pm 17\%$ of CFU-MK and $33 \pm 7\%$ CFU-GM (**Fig. 6**). After correcting for an average transfection efficiency of $55 \pm 4\%$ in CD34+ cells ($n = 4$), we calculated that 70–90% of transfected cells had undergone transposition when the hyperactive SB100X transposase was used. Stable gene transfer with the early-generation transposases was at basal levels or only slightly higher (SB: 1% of CFU-E, 0.2% CFU-GM; SB11: $5 \pm 1\%$ of CFU-E, $10 \pm 4\%$ of CFU-MK and $1 \pm 1\%$

Transposition in primary human CD34+ hematopoietic stem cells

Transposition in many primary cells with previously available transposons is relatively inefficient, which compromises their usefulness for functional genomics and gene therapy. Clinically relevant human CD34+ cells enriched in hematopoietic stem or progenitor cells (HSCs) were co-transfected by nucleofection with a GFP-marked SB transposon (pT2/CAGGS-GFP) and expression constructs encoding the hyperactive SB16X, SB32X and SB100X transposases alongside with the original SB transposase and the second-generation, hyperactive mutant SB11 (ref. 4). The percentage of transfected CD34+ cells that expressed GFP two days post-transfection was similar across the panel of different transposases (**Supplementary Fig. 7a–c** online) and, in particular, was statistically not significantly different between hyperactive and inactive transposases (50–60%), implying that any difference in the percentages of GFP-positive colony forming units (CFUs) after *in vitro* differentiation would reflect bona fide differences in transposition efficiency. The transfected CD34+ cells retained the ability to differentiate into distinct lymphohematopoietic lineages *in vitro* (**Fig. 6** and **Supplementary Fig. 7**) and *in vivo* (**Fig. 7**), and

CFU-GM) than the controls transfected with a construct encoding catalytically inactive transposase (SB-3D: % GFP+ CFU-E, CFU-MK and CFU-GM: <0.5%) (**Fig. 6**). These data establish that hyperactive SB transposases result in a robust and significant (Student's t-test SB100X vs. SB11: $P < 0.002$) increase in transposition-mediated stable gene transfer in CD34+ cells. Finally, SB100X was significantly more efficient in human CD34+ cells than an optimized *piggyBac* transposon system[8] in both CFU-E and CFU-GM clonogenic assays (% GFP+ 33–35% for SB100X vs 7–9% for PB, **Supplementary Fig. 5b**).

Most of the GFP-positive colonies were uniformly green, suggesting that they were derived from progenitor/stem cell clones containing the GFP-tagged transposon integrated in their genomes, followed by clonal expansion (**Fig. 6**). Notably, the phenotype of the erythroid, megakaryocytic and granulocytic/monocytic cells derived from the SB100X-transfected CD34+ cells was indistinguishable from nontransfected controls judging from cytologic characterization and cytofluorimetric analysis using markers that are specifically expressed in these different lineages, including glycophorin A (CFU-E) and CD61

Figure 5 Production of transgenic mice using SB100X. (**a**) Circular pT2/CAGGS-Venus plasmid DNA was injected together with SB and SB100X transposase mRNAs in different ratios into mouse zygotes. The embryos were cultured *ex vivo*, and transferred onto feeder cells for culture at the blastocysts stage. Reporter gene expression was scored at day 7 post-injection, when the embryos were already hatched and the background fluorescence was negligible. The data are collected from two experiments (30–40 embryos per experiment), except injections with 5 ng transposase mRNA that were done six, three, or two times at the different transposon DNA concentrations (gray, black and striped, respectively). Under these conditions, the 80–90% survival rate of the embryos was similar to the noninjected controls. Error bars, s.d. (**b**) Uniform expression in a transgenic animal (see also **Supplementary Fig. 6**).
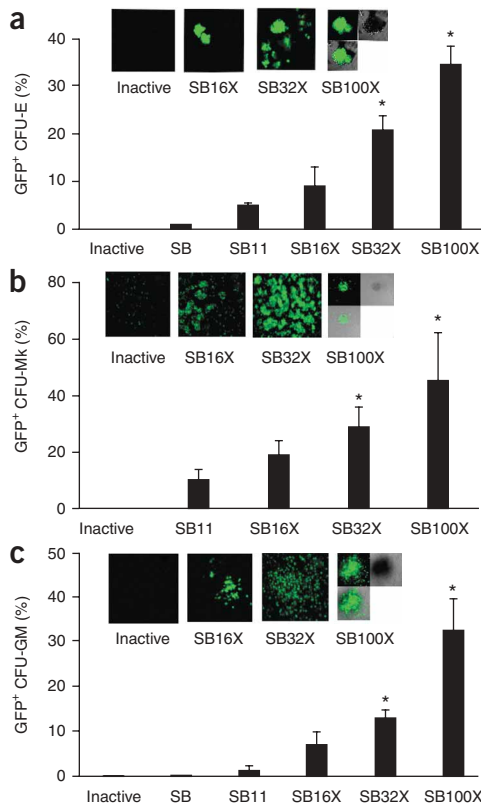
a

b

c
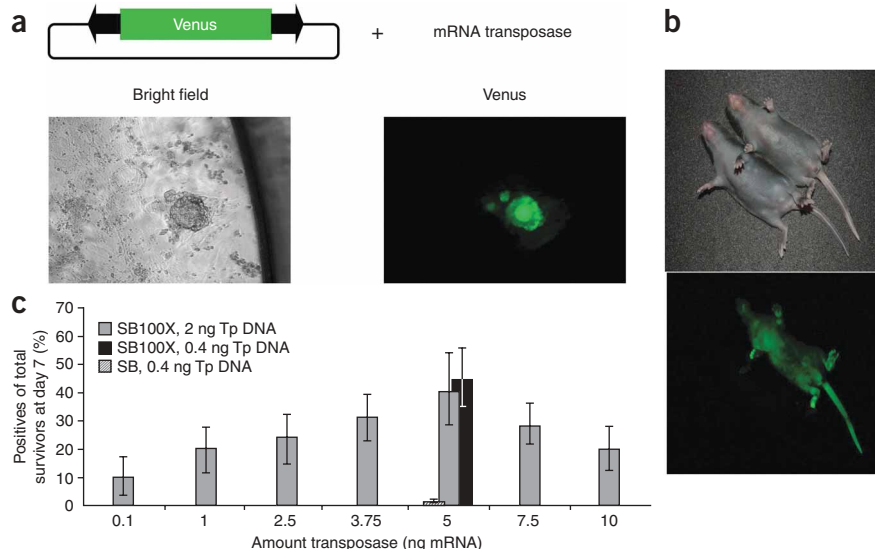
Figure 6 Stable transposition into human CD34[+] cells. Cord blood–derived CD34[+] cells were transfected by nucleofection with the pT2/CAGGS-GFP transposon and expression vectors encoding the different (inactive, active and hyperactive) SB versions. (**a**–**c**) Stable gene transfer was assessed by GFP-expressing colony forming units (CFUs) in clonogenic assays following *in vitro* differentiation into the erythroid (CFU-E) (**a**), megakaryocytic (CFU-MK) (**b**) and granulocyte/monocyte/macrophage (CFU-GM) (**c**) lineages ($n = 4$). Representative fluorescent microphotographs are shown for each of the novel hyperactive transposases compared to the inactive mutant SB. Individual colonies under normal and fluorescent light are shown to indicate lack of mosaicism. Experiments were repeated 3–4 times using cord blood–derived CD34[+] cells from independent donors. Magnification is ×10 except for individual colonies (×20). Error bars, s.d. Asterisks (*) indicate significant difference versus SB11 (Students's *t*-test $P < 0.01$).

To further confirm that bona fide hematopoietic stem or progenitor cells were transfected, we subsequently assessed the extent of gene marking by qPCR in bone marrow and spleen cells enriched for lymphoid (CD20[+]) and myeloid (CD33[+]) cells in an additional cohort of NOD-SCIDγc null mice. Five out of ten mice showed long-term engraftment, yielding an average of 33 ± 10% human CD45[+] engraftment in the peripheral blood. The results shown in **Figure 7** indicate long-term (>5 months) gene marking in both the lymphoid (CD20[+]) and myeloid (CD33[+]) compartments in bone marrow and spleen following transposition in CD34[+] cells with the hyperactive

(CFU-MK) (**Supplementary Fig. 7d–k**). This suggests that there is no overt toxicity associated with hyperactive transposase-mediated gene expression in CD34[+] cells.

To ascertain that transposon integration occurs in bona fide stem cells, we transfected cord blood–derived CD34[+] cells ($n = 6$ donors) with components of the transposon system and transplanted them into immunodeficient NOD-SCIDγc null mice. Long-term (4 months) engraftment of human CD45[+] cells was apparent in the peripheral blood or bone marrow (44–73% CD45[+]) of 50% of the recipient mice (**Fig. 7a**). On the basis of qPCR analysis, we detected 0.4–1.0 transposon copies per cell in the peripheral blood leukocytes or bone marrow from those recipient mice that showed robust human CD45[+] cell engraftment (**Fig. 7b**). We also found that gene-marked CD34[+] cells persisted in the bone marrow (0.2 transposon copies per cell; **Fig. 7c**).

a

b BM/PB - CD45

c BM-CD34

d BM-CD33

e Spleen-CD33

f BM-CD20

g Spleen-CD20

Figure 7 Long-term engraftment and lymphohematopoietic reconstitution with transposon-containing CD34[+] cells. (**a**) Cytofluorimetric analysis of peripheral blood leukocytes and bone marrow in NOD/SCIDγc null mice transplanted with SB100X-transfected CD34[+] cells versus nontransplanted controls. Percentage CD45[+] cells are shown in representative FACS profiles (SSC: side scatter). (**b**–**g**) Transposon copy number determination by real-time PCR analysis on CD45[+] (**b**), CD34[+] (**c**), CD33[+] (**d**, **e**) and CD20[+] (**f**,**g**) cells after transfection of CD34[+] cells with transposon and pCMV-SB100X constructs and subsequent transplantation into immunodeficient NOD/SCIDγc null mice. Cells were enriched by MACS from bone marrow (BM), peripheral blood (PB) or spleen (SP). The [C_t]-[log C_0] standard curves (insert) are shown, and threshold log (ΔR_n) levels are indicated. Transposon copy numbers were as follows: **b**, CD45[+]: 0.5–1 copies/cell ($n = 3$) in PB (filled circle) or BM (filled triangle); **c**, BM CD34[+]: 0.22 copies/cell ($n = 2$); **d**, BM CD33[+]: 0.031 ± 0.008 copies/cell ($n = 3$); **e**, SP CD33[+]: 0.12 copies/cell ($n = 1$); **f**, BM CD20[+]: 0.019 ± 0.015 copies/cell ($n = 3$); **g**, SP CD20[+]: 0.05 copies/cell ($n = 2$). Controls were either not injected with CD34[+] cells or injected with CD34[+] cells transfected with an inactive transposase encoding construct, as indicated.
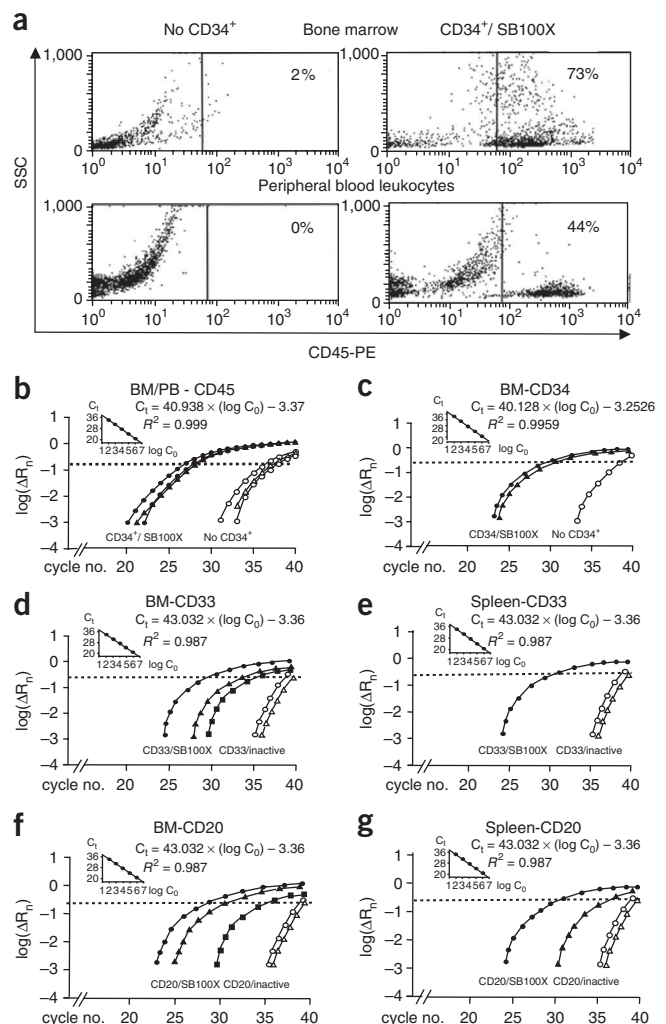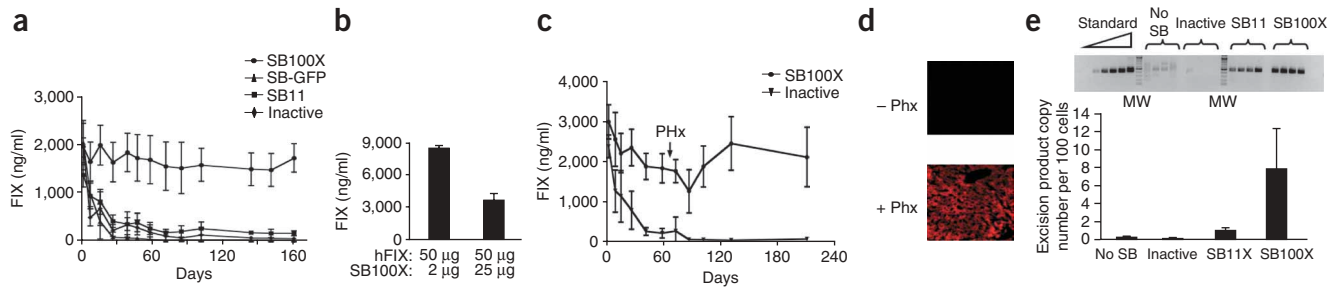
Figure 8 *In vivo* hepatic gene delivery using hyperactive SB. pT2-HB-Apo/AAT-FIX transposon-containing plasmid (50 µg) was co-transfected by hydrodynamic gene delivery with different SB transposase versions (25 µg or 2 µg) or an inactive SB control. Plasma hFIX concentrations were measured (*n* = 4 to 5 mice per group) by ELISA. Mean values of hFIX concentration ± s.d. are shown. (**a**) Comparison of plasma hFIX expression level using pCMV-SB100X (circle), pCMV-SB11 (square), inactive transposase control (diamond) or pCMV-GFP (triangle). (**b**) Overproduction inhibition in liver tissue following co-transfection with pT2-HB-Apo/AAT-FIX (50 µg) and increasing doses of SB100X (2 µg vs. 50 µg). The total amount of DNA injected was normalized. (**c**) hFIX expression levels in actively dividing mouse liver after partial hepatectomy using SB100X (circle) compared to inactive control (triangle). (**d**) BrdU staining of liver sections confirmed active hepatocyte proliferation in the non-hepatectomized vs. hepatectomized liver tissue. (**e**) Excision product (EP) assay on liver (*n* = 4). Hepatocytes were transfected with either SB100X or SB11 and inactive SB was used as negative control. Transposition efficiency was determined by quantifying EP accumulation. Serially diluted plasmid from which the transposon had been excised by restriction digestion/religation was used as standard.

SB100X transposase. No long-term gene marking was apparent when an inactive transposase was used as control. The transposon integration sites recovered from the CFUs *in vitro* (**Supplementary Table 2** online) and from the transplanted cells *in vivo* (**Supplementary Tables 2** and **3** online) confirmed a molecular signature consistent with transposition. Moreover, extensive analysis of the transposon integration sites revealed that identical integration sites were present in both the myeloid and lymphoid lineages, which provides independent confirmation for gene transfer in a common bona fide hematopoietic stem or progenitor cell (**Supplementary Table 3**). Our results ascertain that transposition is essential to achieve long-term and robust gene marking in lymphoid and myeloid compartments following hemato-poietic reconstitution. Thus, the newly developed hyperactive transposon system can be used for relatively efficient gene transfer in primary human stem or progenitor cells that retain their stem cell characteristics and engraft upon *in vivo* transplantation.

Efficient transposition in the mouse liver

To assess the activity of SB100X *in vivo*, we conducted a liver-directed gene transfer experiment in adult C57BL/6 mice. A plasmid containing a transposon that expresses human factor IX (hFIX) from a potent liver-specific promoter (pT2/Apo/AAT-FIX) was hydrodynamically injected along with pCMV-SB100X and pCMV-SB11. SB100X promoted robust and stable hFIX expression levels (2,000–3,000 ng/ml) when compared to control plasmids that expressed either GFP or a catalytically inactive transposase (Student *t*-test: $P < 0.0002$, **Fig. 8a**). SB11 transposase was far less efficient than SB100X (Student's *t*-test: $P < 0.0001$, **Fig. 8a**), and resulted in a low level of sustained hFIX expression that was slightly above the controls. Lower amounts of the transposase (2 µg vs. 25 µg pCMV-SB100X) at fixed amounts of the transposon (50 µg pT2/Apo/AAT-FIX) resulted in higher stable hFIX expression levels in the supraphysiologic range (>8.000 ng/ml, 160% of normal hFIX levels) (**Fig. 8b**). This implies that an excess of hyperactive transposase likely diminished the transposition reaction, consistent with overproduction inhibition[11]. Collectively, these results indicate that prolonged FIX expression in hepatocytes can be ascribed to efficient transposition by the SB100X transposase *in vivo*.

To ascertain that stable hFIX expression is the result of stable genomic integration, we subjected the recipient mice to a partial hepatectomy (Phx) (**Fig. 8c**). In the weeks following Phx, the liver regenerates by de novo cell proliferation (which was confirmed by BrDU incorporation, **Fig. 8d**), during which transfected episomal DNA is lost from the cells. In mice co-transfected with pT2/Apo/AAT-FIX and pCMV-SB100X, high levels of hFIX expression remained stable after Phx when compared to controls (**Fig. 8c**). In contrast, low hFIX levels declined to background levels after Phx in control mice (**Fig. 8c**), suggesting that episomally maintained expression vectors only marginally contributed to the total hFIX levels in the SB100X-treated animals. Indeed, cloned integration sites revealed a molecular signature consistent with transposition (**Supplementary Table 4** online) with no integration bias into genes (30% of insertions within genes vs. 70% intergenic, *n* = 60). We assessed relative transposition efficiencies using a PCR-based transposon excision assay[12] (**Fig. 8e**). The efficiency of transposon excision obtained 7 d post-transfection was approximately eightfold higher with SB100X than with SB11 (**Fig. 8e**), which correlates with the difference in hFIX expression (9- to 14-fold, **Fig. 8a**). The long-term expression of hFIX is consistent with the lack of an immune response against the transgene product, because no antibodies against hFIX could be detected in the experimental animals (**Supplementary Fig. 8** online). This suggests that hepatic expression of FIX following nonviral transposon-mediated gene delivery may be associated with the induction of FIX-specific immune tolerance[13]. Together, the results indicate that hyperactive transposons are ideally suited to achieve long-term and robust transgene expression *in vivo*, which has implications for gene therapy and functional genomics.

DISCUSSION

In the present study, we generated novel hyperactive transposases that showed robust transposition activity at a level up to ~100-fold higher than the original SB transposase. A library of mutant transposase genes obtained by *in vitro* evolution using DNA shuffling was screened. A key technical step was the recovery of the individual hyperactive mutations on ~300-bp restriction fragments that excluded approximately two-thirds of the wild-type, nonproductive sequence input (the SB transposase coding region is 1023 bp) from shuffling. This strategy enabled us to test a large numbers of variants, covering all of the possible pairwise combinations of 41 combinatorial units ($k = 2$: $c = 820$). We believe that a crucial technical advance was the opportunity to screen

for synergistic pairs of mutations, as well as for variants repeatedly recovered in higher-order combinations. This allowed us to analyze compatibility patterns and identify possible strategies for building highly active combinations. Thus, the protocol described here may prove useful in devising *in vitro* evolution screens for other transposases as well. Our results indicate that the increased transposition efficiency of SB100X cannot be merely explained by altered transposase stability, decreased overproduction inhibition or increased binding to the transposon ITRs; instead, the particular combination of mutations in SB100X might affect the folding properties of the transposase (**Supplementary Fig. 4**). However, a more detailed structural analysis of the protein–DNA complex would be necessary to fully understand the hyperactive nature of SB100X. In contrast to hyperactive transposases described earlier[6,7], a large fraction of hyperactivity of SB100X was maintained in primary cell types *ex vivo* as well as in mouse embryos and adult mouse liver *in vivo* (**Figs. 4–8**). The results suggest that the newly developed transposase will serve as a useful reagent for genetic manipulations in vertebrates, with applications including transgenesis, insertional mutagenesis and gene therapy.

For transgenesis, classical methods in vertebrates rely on microinjection of nucleic acids into oocytes or fertilized eggs, typically yielding ~10% transgenesis frequencies[14]. Both of the main drawbacks of these approaches—the low rates of genomic integration and the concatemeric structure of the integrating transgene construct—can be circumvented using transposition-mediated gene delivery[15]. Single units of expression cassettes, resulting from transposition-mediated integration, are presumably less prone to transgene silencing and recombination than the concatemeric insertions created by classical methods. We have shown here that the hyperactive SB100X transposase yields average transgenic frequencies of 45% in mouse embryos following a simple procedure of microinjection of transposon plasmid DNA and synthetic transposase mRNA (**Fig. 5**). The results indicate that the hyperactive transposon system can be developed into a simple and efficient tool for transgenesis in vertebrates.

Insertional mutagenesis with SB in the germline of mice and rats has been approached with mutator lines harboring transposon donor loci containing many (up to several hundred) copies of the transposon vector in the form of concatemeric arrays[16–19]. The reason for this is that transposition out of multicopy concatemers is far more efficient than out of single-copy donor sites[20]. However, transposition out of multicopy donor loci may be associated with multiple insertion events, and segregation of the insertions may lead to loss or change in the phenotype. In addition, recombination between newly transposed transposon copies and the donor concatemer could lead to unwanted genomic rearrangements[21]. The SB100X hyperactive transposase is ~120-fold more active in mobilizing single-copy, chromosomally resident transposons than the wild-type transposase (**Fig. 4b**), and may thus potentially eliminate the need for concatemeric donor sites in genetic screens.

For gene therapy to be effective it is necessary to (i) achieve robust delivery of the desired genes to the relevant target cells, (ii) express the genes long-term and (iii) minimize the risks of secondary effects. The use of viral vectors for gene therapy has been thwarted by immune responses to the viral vector particles and/or the transduced target cells, leading to transient and acute toxicity that precluded long-term expression of the potentially therapeutic gene product[22]. Moreover, production of viral vectors for clinical trials has been fraught with technical and regulatory hurdles. The development of efficient and safe nonviral vectors would therefore greatly facilitate clinical implementation of *ex vivo* and *in vivo* gene

therapies. However, nonviral gene transfer approaches typically result in only limited stable gene transfer efficiencies in most primary cells[23]. We demonstrated here that *ex vivo* transfection of human CD34$^+$ cells with hyperactive transposases resulted in robust and stable gene expression, reaching up to 35–50% GFP-positive CFUs (**Fig. 6**). Transposition events likely occurred in bona fide HSCs, as stable transfectants persisted long term and resulted in multilineage hematopoietic reconstitution after transplantation into immunodeficient mice (**Fig. 7**), which had not been shown previously. This was corroborated by the identification of transposon integration sites common to both the lymphoid and the myeloid lineages *in vivo*. To our knowledge, this is the first demonstration of such robust stable gene transfer in CD34$^+$ cells and their differentiated progeny *in vivo* using a nonviral gene transfer approach. Though the stable gene transfer efficiencies obtained here compare favorably to some of the earlier reports using retroviral or lentiviral vectors[24,25], those efficiencies have been improved over recent years. In particular, we and others have shown that lentiviral transduction efficiencies can be increased by incorporating more robust promoters[26] or *cis*-acting elements[27–29] into the vector backbone or by altering the transduction conditions[30]. Similarly, it may be possible to increase transgene expression levels following transposon-mediated gene transfer by using different promoters, incorporating *cis*-acting elements that enhance expression and/or by exploring alternative transfection methods and conditions. In contrast to retroviral delivery, stable transfection using transposon technology does not require active cell division because the transposon–transposase constructs can reach the nucleus following nucleofection. In addition, efficiency of stable gene transfer obtained by the novel SB100X transposase was found to be superior to transposition mediated by an optimized *piggyBac* transposase[8] (**Supplementary Fig. 5**). Notably, in a therapeutical setup (for example, β-thalassemia, hemophilia) under nonselective conditions in HSCs, the measured difference in transpositional efficiencies is expected to be decisively in favor of SB100X. Another advantage of using the SB system is that there is no apparent preference for transposon integration into genes, as opposed to when lentiviral, retroviral or *piggyBac* transposon–derived vectors are used[9,31,32]. Typically, 70% of the integrations with SB100X occurred in intergenic regions (**Supplementary Tables 2**, **3** and **4**), consistent with the transposon integration pattern seen with early-generation SB transposases[33,34]. This may reduce the risk of insertional oncogenesis in conjunction with the use of transposon vectors that are preferentially targeted to certain genomic regions[35,36] and/or that are transcriptionally shielded using insulators[37,38]. Consistent with their increased transposition efficiencies, the hyperactive transposases produced robust and sustained expression of blood clotting factor FIX at levels significantly higher than previously seen with early-generation SB transposases[39–41], following direct *in vivo* hepatic gene delivery in adult mice (**Fig. 8**). Our data suggest that the power of SB100X manifests in performing transposition in a higher percentage of a transfected cell population rather than pumping up copy numbers of integrated transposons in a single cell (**Fig. 4**). This observation is important, because it indicates that by using SB100X it is possible to titrate the transposon components to obtain predominantly low-copy integrations in large numbers of cells, as opposed to transposition by the original SB that is inefficient below a certain threshold concentration of the transposon substrate DNA.

In summary, the availability of the novel hyperactive transposases may contribute to the development of efficient and safe nonviral

vectors that would greatly facilitate clinical implementation of *ex vivo* and *in vivo* gene therapies and functional genomics studies.

METHODS

Transposase gene library. We recovered the 41 combinatorial units including the individual hyperactive mutations on ~300-bp restriction fragments. The isolated restriction fragments were broken in a random fashion by DNaseI, and 30- to 70-bp populations of fragments were isolated from 12% acrylamide gels. These isolated fragment populations, in combination with oligonucleotides bridging nonoverlapping fragments, were used in a DNA shuffling procedure to reassemble the SB transposase gene. The reassembly reaction was done similarly as described[42]. Briefly, the isolated 30- to 70-bp fragment populations were mixed at the same ratio at a concentration of ~20 ng/μl, and this was supplemented with 2 pmol of each bridging oligonucleotide. We used a high-fidelity PfuUltra polymerase (Stratagene) in a cycle program that consisted of 94 °C for 60 s and 40 rounds of 94 °C for 30 s, 50 °C for 30 s and 68 °C for 1 min, followed by 68 °C for 5 min. A 40× diluted assembly reaction was used as a template in a PCR reaction using primers designed to the ends of the transposase coding region, and the resulting pool of PCR products was cloned into a CMV expression vector using *Spe*I and *Apa*I. We carried out large-scale automated purification of about 2,000 plasmid DNAs representing individual clones of the library using a Genesis Workstation 150 (TECAN) pipetting robot and the Nucleospin Robot 96 plasmid kit (Macherey-Nagel).

Constructs. The pT2/CAGGS-GFP and pT2/CAGGS-Venus transposons contain the green fluorescent protein (GFP) and Venus reporter genes driven by the CAGGS promoter, respectively. The pT2-based transposons used for gene marking studies *in vivo* contain a marker gene (Neo[R] or MGMT). The pXL-BacII/CAGGS-GFP construct contains a *piggyBac* transposon that has the same CAGGS-GFP expression cassette as in pT2/CAGGS-GFP. pT2/Apo/AAT-FIX expresses a human factor IX (FIX) minigene from a chimeric promoter composed of the apolipoprotein E enhancer/α1-antitrypsin promoter and the hepatocyte control region that was obtained by cloning a 4,560-bp *Spe*I fragment from pBS-HCRHP-FIXIA[43] (provided by C. Miao, University of Washington) into the *Spe*I sites of pT2HB. All the SB transposases (active, inactive and hyperactive) and the codon-optimized *piggyBac* transposase[8] are encoded by a CMV expression plasmid.

Cells and transfection. Human HeLa cell tissue culture and transfection was done as described[1]. We obtained CD34[+] cells from umbilical cord blood or from the bone marrow of NOD-SCIDγc null mice transplanted with CD34[+] cells after Ficoll/Hypaque centrifugation followed by immunomagnetic separation using magnetic beads conjugated to antibodies to CD34 according to the manufacturer's instructions (Miltenyi Biotech). Similarly, myeloid (CD33[+]) and lymphoid (CD20[+]) cells were enriched from the bone marrow and spleen of recipient NOD-SCIDγc null mice transplanted with CD34[+] cells by immunomagnetic separation using magnetic beads conjugated to the respective antibodies (Miltenyi Biotech).

Transfection of CD34[+] cells was done according to the optimized protocol for human CD34[+] cells using nucleofection (Amaxa Biosystems). For *in vitro* comparisons of relative transposition efficiencies, we carried out at least three independent experiments using cells obtained from 3–5 independent donors. Enriched CD34[+] cells in PBS were centrifuged at 300*g* (4 °C) for 10 min and resuspended in Nucleofector buffer. Typically, 1.5×10^5 cells in 100 μl of human CD34 cell Nucleofector buffer per cuvette were subjected to electroporation with purified plasmids containing the transposon (10 μg) and transposase (5 μg). CD34[+] cells were kept in maintenance medium (Stemline medium, Sigma Aldrich, supplemented with 100 ng/ml SCF, 20 ng/ml IL-6, 100 ng/ml IL-3, 20 ng/ml Flt3-l and 100 ng/ml TPO), and GFP expression was assessed by cytofluorimetric analysis (FACSvantage). Clonogenic and cytologic assays were carried out as described in **Supplementary Methods** online. For transplantations in NOD-SCIDγc null mice, 10[6] CD34[+] cells (typically pooled from 2 to 3 donors) were co-transfected as described above with 10 μg transposon DNA and 5 μg pCMV-SB100X. Integration site analysis on transfected CD34[+] cells was done by splinkerette PCR as described in **Supplementary Methods**.

Mice and injections. NOD-SCIDγc null mice were kept and bred in a specific pathogen-free facility at the University of Leuven and were pre-conditioned with 100 cGy (Linear accelerator, University Hospital Gasthuisberg, Belgium) one day before injection of CD34[+] cells, as described previously[44]. One million transfected CD34[+] cells were resuspended in 50 μl PBS with 5% FBS and injected intravenously, 4–6 h postnucleofection and without any cytokine stimulation (Stemline medium). We determined persistence of engrafted human cells by FACS analysis, carried out integration site analysis on transfected CD34[+] cells by splinkerette PCR and determined transposon copy number by quantitative PCR as described in **Supplementary Methods**.

C57BL/6 mice were hydrodynamically transfected with 50 μg of transposon with 2 μg or 25 μg of transposase plasmid diluted in 2 ml of PBS and injected into the tail vein. Typically, the injection took less than 10 s for each mouse. Some animals were anesthetized and subjected to a partial hepatectomy whereby two-thirds of the liver was surgically removed to stimulate hepatocyte proliferation and liver regeneration. Hepatocyte proliferation was confirmed by BrdU incorporation using BrdU-specific antibodies. All animal procedures were approved by the Animal Ethics Committee of the University of Leuven. We carried out integration site analysis on transfected liver by splinkerette PCR and determined FIX expression by ELISA as described in **Supplementary Methods**.

Animal transgenesis. Three- to four-week-old B6D2F1 hybrid female mice were superovulated and mated to B6D2F1 hybrid males. Fertilized oocytes were harvested and injected as described[45]. The injected fertilized oocytes were either transferred to foster mice or cultured in KSOM medium until blastocysts stage (day 4 after injection). The embryos were transferred into a 96-well tissue culture plate covered with mitotically inactive mouse embryonic feeder cells, and cultured using standard ES cell culturing conditions. Venus fluorescence was monitored in hatched embryos at day 7 post-injection or in transgenic animals using an Olympus IX81 microscope or FHS/LS-1B macro-visualization technology (BLS Ltd), respectively.

Note: Supplementary information is available on the Nature Genetics website.

supervised the project and wrote the paper. Z. Izsvák designed the hyperactive screen and characterization of the hyperactive transposases, analyzed and interpreted data, supervised the project and wrote the paper.

1. Ivics, Z., Hackett, P.B., Plasterk, R.H. & Izsvák, Z. Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510 (1997).
2. Mates, L., Izsvak, Z. & Ivics, Z. Technology transfer from worms and flies to vertebrates: transposition-based genome manipulations and their future perspectives. *Genome Biol.* **8**(Suppl 1), S1 (2007).
3. Ivics, Z. & Izsvak, Z. Transposons for gene therapy! *Curr. Gene Ther.* **6**, 593–607 (2006).
4. Geurts, A.M. *et al.* Gene transfer into genomes of human cells by the *Sleeping Beauty* transposon system. *Mol. Ther.* **8**, 108–117 (2003).
5. Zayed, H., Izsvak, Z., Walisko, O. & Ivics, Z. Development of hyperactive *Sleeping Beauty* transposon vectors by mutational analysis. *Mol. Ther.* **9**, 292–304 (2004).
6. Baus, J., Liu, L., Heggestad, A.D., Sanz, S. & Fletcher, B.S. Hyperactive transposase mutants of the Sleeping Beauty transposon. *Mol. Ther.* **12**, 1148–1156 (2005).
7. Yant, S.R., Park, J., Huang, Y., Mikkelsen, J.G. & Kay, M.A. 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 (2004).
8. Cadinanos, J. & Bradley, A. Generation of an inducible and optimized piggyBac transposon system. *Nucleic Acids Res.* **35**, e87 (2007).
9. Wilson, M.H., Coates, C.J. & George, A.L. Jr. PiggyBac Transposon-mediated Gene Transfer in Human Cells. *Mol. Ther.* **15**, 139–145 (2007).
10. Wang, W. *et al.* Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. USA* **105**, 9290–9295 (2008).
11. Izsvak, Z. & Ivics, Z. Sleeping beauty transposition: biology and applications for molecular therapy. *Mol. Ther.* **9**, 147–156 (2004).
12. Liu, G., Aronovich, E.L., Cui, Z., Whitley, C.B. & Hackett, P.B. Excision of Sleeping Beauty transposons: parameters and applications to gene therapy. *J. Gene Med.* **6**, 574–583 (2004).
13. Mingozzi, F. *et al.* Induction of immune tolerance to coagulation factor IX antigen by in vivo hepatic gene transfer. *J. Clin. Invest.* **111**, 1347–1356 (2003).
14. Conner, D.A. Transgenic mouse production by zygote injection. *Curr. Protoc. Mol. Biol.* **Chapter 23**, Unit 23.9 (2004).
15. Dupuy, A.J. *et al.* Mammalian germ-line transgenesis by transposition. *Proc. Natl. Acad. Sci. USA* **99**, 4495–4499 (2002).
16. Horie, K. *et al.* Characterization of *Sleeping Beauty* transposition and its application to genetic screening in mice. *Mol. Cell. Biol.* **23**, 9189–9207 (2003).
17. Kitada, K. *et al.* Transposon-tagged mutagenesis in the rat. *Nat. Methods* **4**, 131–133 (2007).
18. Lu, B. *et al.* Generation of rat mutants using a coat color-tagged Sleeping Beauty transposon system. *Mamm. Genome* **18**, 338–346 (2007).
19. Dupuy, A.J., Fritz, S. & Largaespada, D.A. Transposition and gene disruption in the male germline of the mouse. *Genesis* **30**, 82–88 (2001).
20. Yusa, K., Takeda, J. & Horie, K. Enhancement of Sleeping Beauty transposition by CpG methylation: possible role of heterochromatin formation. *Mol. Cell. Biol.* **24**, 4004–4018 (2004).
21. Geurts, A.M. *et al.* Gene mutations and genomic rearrangements in the mouse as a result of transposon mobilization from chromosomal concatemers. *PLoS Genet.* **2**, e156 (2006).
22. Manno, C.S. *et al.* Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat. Med.* **12**, 342–347 (2006).
23. Hollis, R.P. *et al.* Stable gene transfer to human CD34(+) hematopoietic cells using the Sleeping Beauty transposon. *Exp. Hematol.* **34**, 1333–1343 (2006).
24. Conneally, E., Eaves, C.J. & Humphries, R.K. Efficient retroviral-mediated gene transfer to human cord blood stem cells with in vivo repopulating potential. *Blood* **91**, 3487–3493 (1998).
25. Miyoshi, H., Smith, K.A., Mosier, D.E., Verma, I.M. & Torbett, B.E. Transduction of human CD34+ cells that mediate long-term engraftment of NOD/SCID mice by HIV vectors. *Science* **283**, 682–686 (1999).
26. Salmon, P. *et al.* High-level transgene expression in human hematopoietic progenitors and differentiated blood lineages after transduction with improved lentiviral vectors. *Blood* **96**, 3392–3398 (2000).
27. Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M. & Naldini, L. Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat. Genet.* **25**, 217–222 (2000).
28. VandenDriessche, T. *et al.* Lentiviral vectors containing the human immunodeficiency virus type-1 central polypurine tract can efficiently transduce nondividing hepatocytes and antigen-presenting cells in vivo. *Blood* **100**, 813–822 (2002).
29. Zufferey, R., Donello, J.E., Trono, D. & Hope, T.J. Woodchuck hepatitis virus post-transcriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J. Virol.* **73**, 2886–2892 (1999).
30. Piacibello, W. *et al.* Lentiviral gene transfer and ex vivo expansion of human primitive stem cells capable of primary, secondary, and tertiary multilineage repopulation in NOD/SCID mice. Nonobese diabetic/severe combined immunodeficient. *Blood* **100**, 4391–4400 (2002).
31. Bushman, F. *et al.* Genome-wide analysis of retroviral DNA integration. *Nat. Rev. Microbiol.* **3**, 848–858 (2005).
32. Schroder, A.R. *et al.* HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**, 521–529 (2002).
33. Liu, G. *et al.* Target-site preferences of Sleeping Beauty transposons. *J. Mol. Biol.* **346**, 161–173 (2005).
34. Vigdal, T.J., Kaufman, C.D., Izsvák, Z., Voytas, D.F. & Ivics, Z. Common physical properties of DNA affecting target site selection of *Sleeping Beauty* and other Tc1/*mariner* transposable elements. *J. Mol. Biol.* **323**, 441–452 (2002).
35. Ivics, Z. *et al.* Targeted Sleeping Beauty transposition in human cells. *Mol. Ther.* **15**, 1137–1144 (2007).
36. Yant, S.R., Huang, Y., Akache, B. & Kay, M.A. Site-directed transposon integration in human cells. *Nucleic Acids Res.* **35**, e50 (2007).
37. Walisko, O. *et al.* Transcriptional activities of the Sleeping Beauty transposon and shielding its genetic cargo with insulators. *Mol. Ther.* **16**, 359–369 (2008).
38. Dalsgaard, T. *et al.* Shielding of sleeping beauty DNA transposon-delivered transgene cassettes by heterologous insulators in early embryonal cells. *Mol. Ther.* **17**, 121–130 (2009).
39. Mikkelsen, J.G. *et al.* Helper-Independent Sleeping Beauty transposon-transposase vectors for efficient nonviral gene delivery and persistent gene expression in vivo. *Mol. Ther.* **8**, 654–665 (2003).
40. Ohlfest, J.R. *et al.* Phenotypic correction and long-term expression of factor VIII in hemophilic mice by immunotolerization and nonviral gene transfer using the Sleeping Beauty transposon system. *Blood* **105**, 2691–2698 (2005).
41. Yant, S.R. *et al.* Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system. *Nat. Genet.* **25**, 35–41 (2000).
42. Stemmer, W.P. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc. Natl. Acad. Sci. USA* **91**, 10747–10751 (1994).
43. Miao, C.H. *et al.* Inclusion of the hepatic locus control region, an intron, and untranslated region increases and stabilizes hepatic factor IX gene expression in vivo but not in vitro. *Mol. Ther.* **1**, 522–532 (2000).
44. Ishikawa, F. *et al.* Development of functional human blood and immune systems in NOD/SCID/IL2 receptor {gamma} chain(null) mice. *Blood* **106**, 1565–1573 (2005).
45. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. *Manipulating the Mouse Embryo: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2003).
46. Moore, J.C., Jin, H.M., Kuchner, O. & Arnold, F.H. Strategies for the in vitro evolution of protein function: enzyme evolution by random recombination of improved sequences. *J. Mol. Biol.* **272**, 336–347 (1997).