Molecular Therapy
Nucleic Acids
Original Article

A novel hyperactive variant of the *Sleeping Beauty* transposase facilitates non-viral genome engineering

Matthias Thomas Ochmann,[1] Csaba Miskey,[1] Lacramioara Botezatu,[1] Nicolás Sandoval-Villegas,[1] Tanja Diem,[1] and Zoltán Ivics[1,2,3]

[1]Research Center, Division of Hematology, Cell and Gene Therapy, Paul-Ehrlich-Institut, 63225 Langen, Germany; [2]Institute of Clinical Immunology, University of Leipzig, 04103 Leipzig, Germany; [3]Fraunhofer Institute for Cell Therapy and Immunology, 04103 Leipzig, Germany

The *Sleeping Beauty* (SB) transposon system is a useful tool for genetic applications, including gene therapy. We discovered a hyperactive variant of the SB100X transposase, called SB200X. This mutant, resulting from a specific amino acid replacement (Q124C), showed an ∼2-fold increase in transposition activity in various human and murine cells. Other amino acid replacements in position 124 also led to a hyperactive phenotype. Position 124 is located at the very edge of the linker region that connects the DNA-binding and catalytic domains of the transposase. Consistent with a role of the linker in an autoregulatory mechanism called overproduction inhibition (OPI) in the monophyletic group of *mariner* transposases, we show that the hyperactivity of Q124C manifests at high concentrations of the transposase, suggesting a partial resistance of SB200X to OPI. We demonstrate that the hyperactive phenotype of Q124C can be combined with features of other useful mutations in the SB transposase. Namely, Q124C improves the transposition efficiency of the previously described K248R variant, while maintaining or even slightly improving its safer genome-wide integration profile. The SB200X transposase could enhance the utility of SB transposon-mediated genome engineering in preclinical and clinical applications.

INTRODUCTION

Barbara McClintock's initial discovery of mobile genetic elements in maize revealed their ability to transfer their genetic material from one locus to another within the genome.[1] Transposable elements (TEs) are classified into two groups. Class I TEs are retrotransposons that implement a copy and paste mechanism, employing an RNA intermediate for this particular process.[2] In contrast, the transposition process of class II transposons is exclusively reliant on DNA intermediates. In this category, the vast majority of transposons utilize a cut and paste mechanism, whereby the transposon is excised from one genomic location and subsequently reintegrated at a different location.[2] The members of this particular transposon superfamily are characterized by the presence of terminal inverted repeats (TIRs) and a gene that encodes a transposase, an enzyme that facilitates the transposition reaction (recently reviewed in Ochmann, Amberger and co-workers[3,4]). During cut and paste transposition, the transposase binds to the TIRs, excises the transposon from the donor site, and reintegrates it into a new location (recently reviewed by Ochmann and Ivics[3]) (Figure 1). Importantly, the transposase and specific TIRs exhibit *trans*-complementarity, which is fundamental in transforming transposons into genetic vector systems capable of transferring any desired gene into a host cell's genome.[5]

A member of the Tc1/*mariner* superfamily of class II TEs is the *Sleeping Beauty* (SB) transposon.[6] It was reconstructed from fossil DNA sequences within fish genomes, providing the first DNA transposon active in vertebrates.[6] It is widely used as a genetic engineering tool in germline transgenesis and generation of animal models for disease, insertional mutagenesis screens, and in various preclinical studies and clinical trials (recently reviewed in Amberger and Ivics[4]). The transposase is composed of two distinct domains[3,7]: an N-terminal DNA binding domain (DBD) spanning amino acids (aa) 1–109 (coordinates according to Voigt et al.[8]), and a C-terminal catalytic domain spanning aa 128–340 (Figure 2A). The two domains are connected by a flexible linker region (aa 110–127)[8] that harbors a nuclear localization signal (NLS)[7] (Figure 2A). The DBD is comprised of two subdomains, namely the PAI and RED subdomains (Figure 2A). Each subdomain features a helix-turn-helix motif, which is responsible for the recognition of TIR DNA.[7,9–11] The DNA cleavage and joining reactions are executed by the catalytic domain, which contains three conserved amino acids (D153, D244, E279 [DDE]) that form the catalytic center and the so-called clamp loop (aa 159–190)[7–9] (Figure 2A).

Understanding the structural and biochemical principles of the SB transposase is of great interest, as optimizing the enzymatic activity of the transposase can enhance the integration of SB into target cell genomes. Following the discovery of SB, several mutations have

Received 2 July 2024; accepted 31 October 2024;
https://doi.org/10.1016/j.omtn.2024.102381.

Correspondence: Zoltán Ivics, Fraunhofer Institute for Cell Therapy and Immunology IZI, Perlickstrasse 1, 04103 Leipzig, Germany.
E-mail: zoltan.ivics@izi.fraunhofer.de

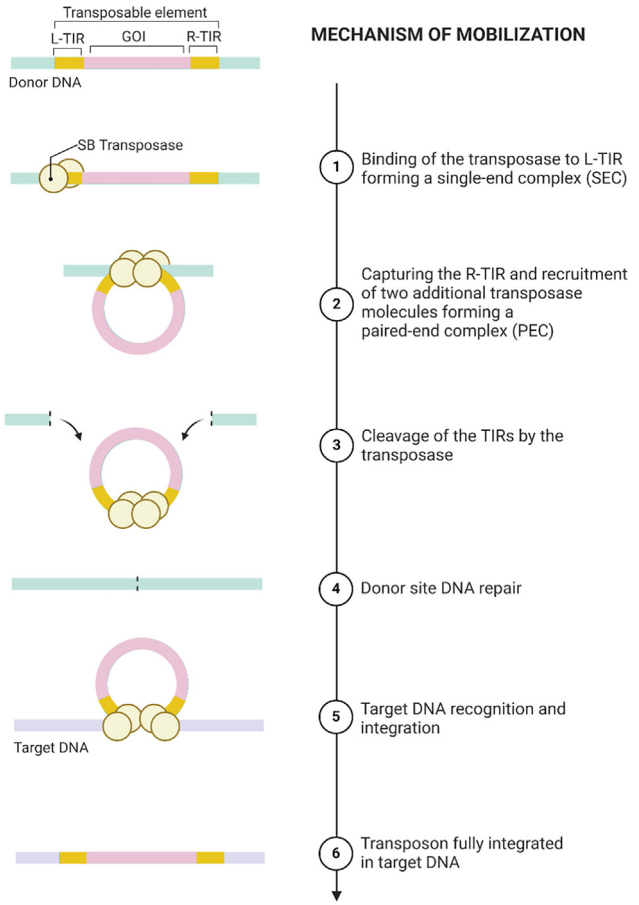

Figure 1. The mechanism of *Sleeping Beauty* transposition

When used as a gene vector, the *Sleeping Beauty* (SB) transposable element is flanked by the left and right terminal inverted repeats (L-TIR and R-TIR in yellow) and harbors a gene of interest (GOI) (in purple). The GOI can be any DNA, but in this study it was invariably a transcription unit consisting of a promoter, the coding region of a protein, and a polyadenylation site. (1) The SB transposase (light yellow sphere) binds to the L-TIR of SB, thereby forming a single-end complex (SEC). (2) The SEC captures the naked R-TIR to form a paired-end complex (PEC). (3) Excision of the SB transposon from the donor DNA (light green). (4) The donor site is repaired. (5 and 6) The excised transposon integrates into the target DNA (light purple).

been identified that result in an elevated integration efficiency.[8,14–18] One particular combination of such mutations has ultimately led to the development of the most widely used transposase variant of the SB system, known as SB100X.[19] Other modifications broadened the applications of the SB transposase, including the K248T mutation leading to an excision-proficient but integration-deficient SB transposase,[20] the K248R mutation leading to a safer genome-wide integration profile[21] and the hsSB variant that was found to spontaneously penetrate cells, enabling genetic modification of cells without the use of transfection protocols.[22]

Structural studies with the *Mos1 mariner* transposase highlighted a prominent dimerization interface between two transposase subunits, involving the clamp loop and the linker regions, in a manner that the clamp loop that extends from the catalytic domain of one subunit interacts with a WVPHEL motif within the linker region of the other.[13] The WVPHEL motif is highly conserved in *mariner* transposases (Figure 2B), and was found to be involved in allosterically downregulating the transposition reaction of the human *Hsmar1* element,[23] which manifests in a phenomenon called overproduction inhibition (OPI). OPI means that an increase in transposase concentration beyond a certain point reduces the rate of transposition rather than increasing it, and it was also described to affect SB transposition.[24] Mutations in the WVPHEL motif have resulted in the development of *mariner* transposase variants that exhibit hyperactivity due to lower vulnerability of the transposase to OPI.[23] We discovered a single amino acid replacement within the linker region of the SB100X transposase near the equivalent position to the *mariner* WVPHEL motif, which involves the substitution of glutamine (Q) with cysteine (C) at aa position 124 (Q124C). This alteration results in a 2-fold increase in transposition activity in the human HepG2 and HeLa cell lines, in human induced pluripotent stem cells (hiPSCs), and in primary murine hematopoietic stem and progenitor cells (mHSPCs). Q124C displays a significant reduction of OPI. We show that the hyperactive phenotype of Q124C can be combined with other clinically relevant transposase phenotypes, such as a safer genomic integration profile of the K248R variant. In general, these results will help enhancing the efficacy of SB transposon-mediated genome engineering in preclinical and clinical applications.

RESULTS

The Q124C mutation leads to hyperactivity in human cell lines and in hiPSCs

To explore *in vivo* control of SB transposition, we set out to tag the SB transposase with inteins that mediate post-translational protein splicing, whereby an intein domain excises itself from a larger precursor polypeptide.[25] There are two important requirements for ensuring both efficient intein splicing and functionality of the resulting spliced product: the intein needs to be situated next to a native C amino acid residue and in a position within the target protein so that the integrity of critical protein domains is maintained through proper folding.[26] We therefore hypothesized that the best region to split the SB transposase with inteins is the linker region that connects the DBD and catalytic domains (Figure 2A). Because the linker region of the SB transposase does not contain a C residue (Figure 2B), and overlaps with an NLS (Figure 2A) that is strongly conserved in the Tc1 clade of transposases[7] and whose alteration would likely result in compromising transposase activity, we elected to introduce a C residue in position 124 of the SB transposase. Position 124 is right at the edge of the linker and is a Q in the SB transposase, which is not conserved in the Tc1 clade of transposases (Figure 2B), suggesting that its conversion to C is unlikely to have a negative impact on transposition.

To evaluate the transposition activity of the Q124C mutation within the SB100X coding sequence, we conducted a transfection experiment using expression plasmids encoding either SB100X or its Q124C variant in human HepG2 cells. In addition, we included a puromycin

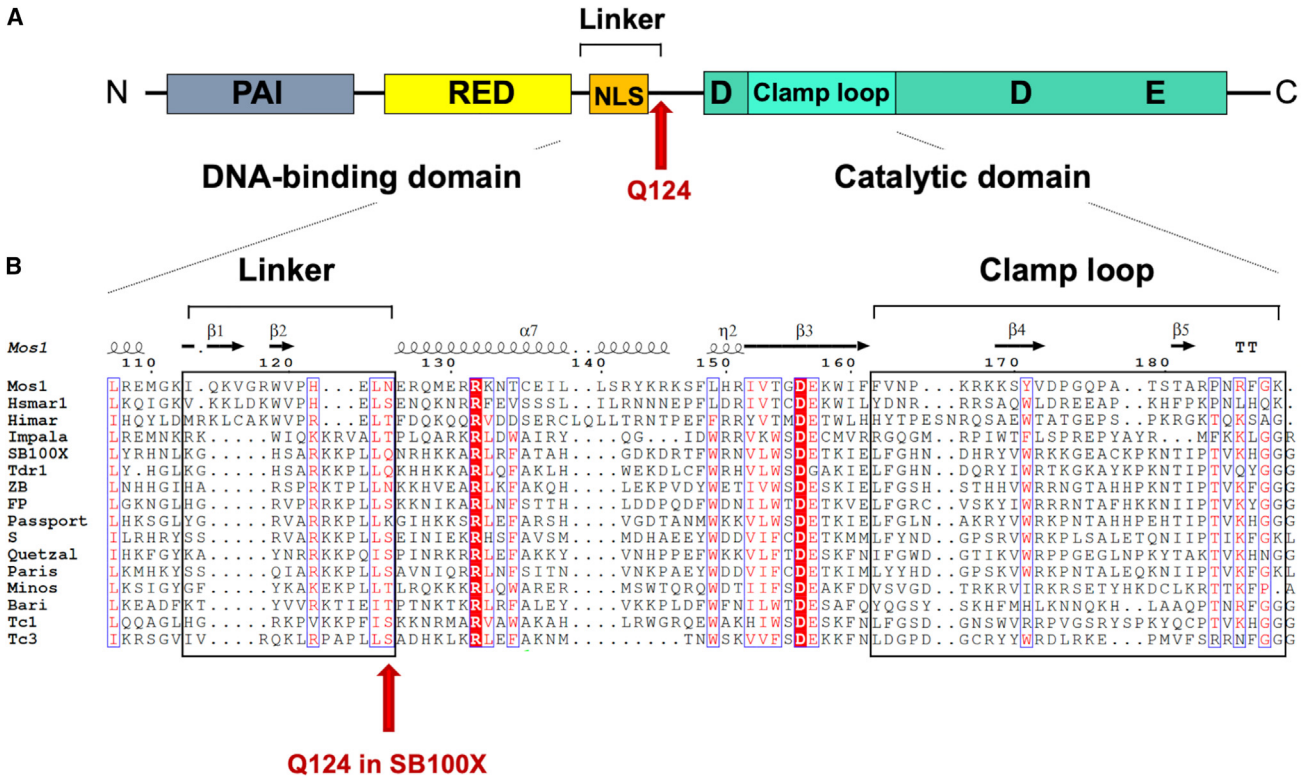

Figure 2. The *Sleeping Beauty* transposase

(A) Schematic representation of the domain structure of the *Sleeping Beauty* (SB) transposase. The SB transposase has an N-terminal, bipartite, paired-like DNA binding domain with the helix-turn-helix PAI subdomain (gray box) and RED subdomain (yellow box), followed by a linker region harboring a nuclear localization signal (NLS) (orange box) and a C-terminal catalytic domain (green box), with the DDE amino acid triad catalyzing the DNA cleavage and joining reactions. The clamp loop (light green box) important for protein-protein interactions is situated between the first and second D of the DDE triad. A red arrow indicates the position of Q124 at the end of the linker region. (B) Amino acid sequence alignment of segments of Tc1 family transposases, including SB100X, and *mariner* family transposases. The alignment was generated by CLUSTALW and then used to highlight secondary structures of the *Mos1* transposase (based on PDB: 4U7B) with ESPript 3.0.[12] Accordingly, the numbering of α helices and beta strands corresponds to *Mos1*. The linker between the DNA-binding and catalytic domains has two short β strands: β1 (residues 113–116) and β2 (residues 118–120). The clamp loop, situated between residues 162 and 189, also contains two short β strands: β4 (residues 169–172) and β5 (residues 180–182).[13] Red box, white character denotes strict identity within the group; red characters signify similarity defined by a similarity score calculated by ESPript 3.0 based on the physico-chemical properties of the amino acids. A red arrow indicates the position of the Q124 residue in the SB100X transposase at the end of the linker region.

resistance (*puro*) gene-tagged SB transposon vector and determined the relative efficiencies of transposition by quantifying the numbers of puro-resistant cell colonies following antibiotic selection in the presence of the transposase variants or green fluorescent protein (GFP) as negative control (Figures 3A and 3B). Unexpectedly, the transposition rate turned out to be about 2-fold increased by the Q124C mutant as compared with SB100X (Figure 3B). No difference between SB100X and Q124C regarding relative levels of protein expression was observed by western blot analysis (Figure S1A), indicating that the observed hyperactivity of Q124C is a genuine phenotype associated with the mutation itself in this particular transposase variant. Furthermore, the hyperactive behavior of the Q124C mutation was also apparent in hiPSCs that are of greater clinical significance. We performed transfection with expression plasmids that encoded for SB100X, Q124C, or the catalytically inactive variant DAE (containing an alanine substitution in the DDE catalytic triad) along with an SB transposon minicircle (MC) encoding a Venus fluorescent protein[27] (MC.T2-CAGGS-Venus, in short MC-Venus), as depicted in Figure 3C. The frequency of Venus$^+$ cells was more than 2-fold higher in the cell population transfected with the Q124C variant as compared with SB100X at day 14 post-nucleofection (Figures 3D and S1B).

Upon conducting saturating mutagenesis of Q124, we found that several other amino acid substitutions also resulted in a notable increase in transposition efficiency in HeLa cells (Figure 3E). Notably, there was no discernible trend in terms of the impact on activity depending on whether the amino acid is polar, hydrophobic, or electrically charged, regardless of the size of the side chain. Collectively, these results establish a hyperactivity of several mutants in position 124, including Q124C, when compared with the state-of-the-art SB100X transposase.

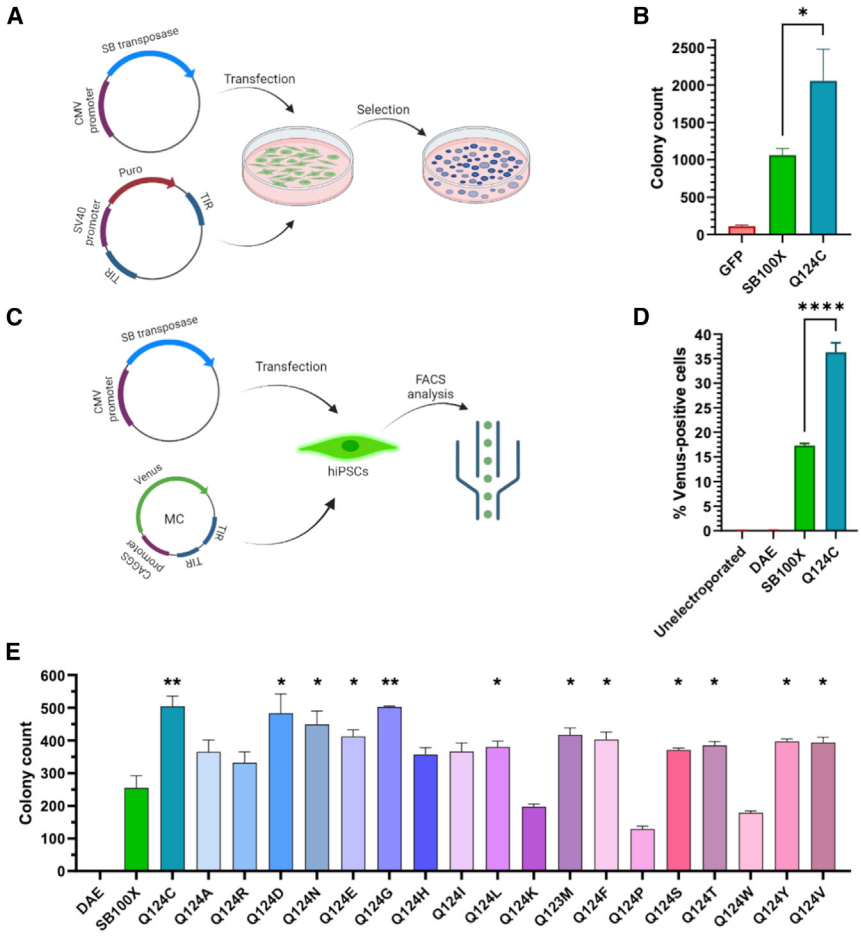

Figure 3. Hyperactivity of the Q124C mutant of the SB100X transposase in human HepG2 cells and hiPSCs

(A) Schematic representation of a colony formation-based transposition assay after antibiotic selection. The SB transposase expression plasmid is co-transfected with a puromycin resistance (*puro*) gene-tagged SB transposon plasmid and the puromycin-resistant colonies are counted after 2 weeks. (B) The Q124C mutant of SB100X displays a hyperactive phenotype, as assayed by a colony-forming transposition assay in human HepG2 cells (n = 3; $^{*}p < 0.05$). (C) Schematic representation of a fluorescence-based transposition assay. The SB transposase expression plasmid is co-transfected with a *Venus* gene-tagged SB transposon minicircle (MC). Transposition efficiency is determined by the frequency of Venus^{+} cells quantified by flow cytometry analysis. (D) The frequency of Venus^{+} hiPSCs determined by flow cytometry 14 days post-transfection (n = 3; $^{****}p < 0.0001$). (E) Colony-forming transposition assay in human HeLa cells after nucleofection of a *puro*-tagged SB transposon plasmid together with SB100X harboring other amino acid replacements at position Q124 (n = 3; significance assessed by t test always vs. SB100X: $^{*}p < 0.05$, $^{**}p < 0.01$).

The Q124C mutation leads to hyperactivity in mHSPCs

HSPC gene therapy has emerged as an effective treatment option for monogenic disorders of the blood system.[28] Because HSPCs are hard-to-transfect cells and SB transposition in this cell type is relatively inefficient,[27] increasing the efficiency of gene transfer in this cell type by improved vector components would be highly relevant. Thus, we next compared the transposition activity of the Q124C mutation in murine bone marrow-derived lineage-negative (Lin^{-}) cells, a heterogeneous cell population of HSPCs. Lin^{-} cells were co-transfected by nucleofection with the MC-Venus transposon and expression constructs encoding the SB transposase variants SB100X and Q124C, as well as the catalytically inactive variant DAE as negative control. The frequencies of Venus^{+} cells (Figure 4A) and the percentage of viable cells (Figure 4B) were monitored over time in liquid culture and stable Venus expression was achieved by day 6. At a dose of 1 μg expression plasmid encoding the transposases, no difference in the numbers of Venus^{+} cells was observed. However, at a higher dose of 2 μg expression plasmid, the Q124C mutation resulted in a 2-fold increase in the numbers of stably expressing Venus^{+} cells (Figure 4A). The difference between numbers of Venus^{+} cells obtained at days 2 and 6 post-nucleofection is due to transient expression of the reporter caused by unintegrated MC DNA present in the transfected cells at early time points, which then decreases due to cell division and gradual loss of episomal vector DNA. There was a complete lack of Venus expression in the negative control obtained with the catalytically dead mutant of SB100X (DAE) at day 9, suggesting that any reporter expression observed in the samples transfected with catalytically active transposases after this time point must derive from genomically integrated transgene copies. Cell viability 2 days post-nucleofection was below 25% in all samples, but the cells recovered quickly and reached a viability of over 80% by day 9 due to the increased proliferation of viable Lin^{-} cells (Figure 4B). These results demonstrate that the Q124C hyperactive SB transposase increases the efficiency of gene transfer in primary murine Lin^{-} cells, and could therefore be employed in future preclinical gene therapy protocols.

The Q124C mutation leads to partial resistance to OPI and manifests itself already at the excision step during SB transposition

There are no direct biochemical data available for the SB transposase with respect to the mechanism of OPI, and the WVPHEL motif is not conserved in Tc1 transposases, including SB (Figure 2B). However, the linker regions of *mariner* and SB transposases are structurally conserved and have been shown to contribute to the dimer interface by interacting with the clamp loop of another transposase monomer.[8,13] Thus, given the proximity of the Q124 mutation site to the interface responsible for protein-protein

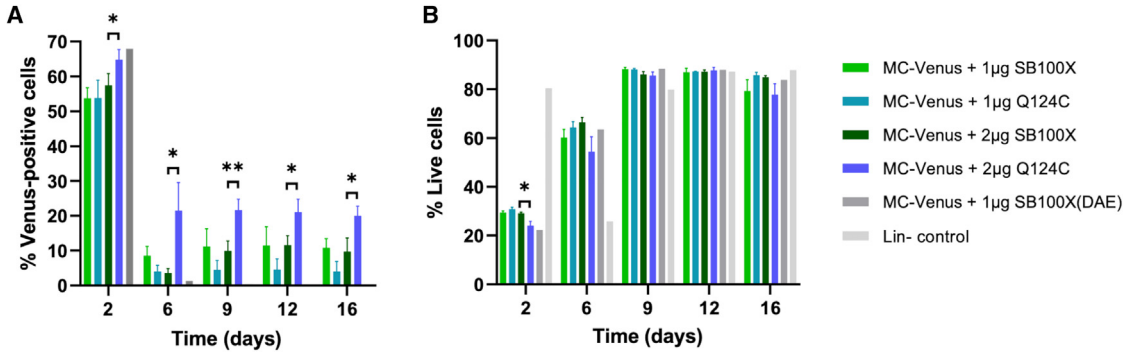

Figure 4. Hyperactivity of the Q124C mutant of the SB100X transposase in murine bone marrow-derived HSPCs

(A) Transposition activity of the Q124C mutant in mouse primary lineage-negative (Lin^-) cells. Lin^- cells were nucleofected with the MC-Venus transposon and expression vectors encoding different (inactive, SB100X and SB100X harboring the Q124C mutation) SB transposase versions at day 2 after isolation. Venus expression was measured by flow cytometry at days 2–16 post-nucleofection ($n = 3$; significance assessed by t test for 2 µg SB100X vs. 2 µg Q124C, $^*p < 0.05$, $^{**}p < 0.01$). (B) Viability of nucleofected Lin^- cells over time was assessed using the Zombie NIR fixable viability kit followed by flow cytometry analysis ($n = 3$; significance assessed by t test for 2 µg SB100X vs. 2 µg Q124C, $^*p < 0.05$). Lin^- cells transfected with the Q124C variant displayed a slightly reduced viability on day 2, possibly due to its increased transposon excision activity, resulting in a transient increase in the production of linear DNA fragments that may trigger the DNA damage response pathway. However, by day 6 the cells recovered, and the decrease in viability was not significant anymore. The decrease in viability of Lin^- control cells at day 6 is likely due to their faster expansion compared with nucleofected cells and resulting slight overgrowth of the culture. After this time point, both the expansion rates and the viabilities of all samples stabilized at day 9.

interactions involved in OPI in *mariner* elements (Figure 2B), we asked if the hyperactivity of the Q124C mutant may be linked to an alteration in OPI. A model of the SB transposase structure in complex with transposon ends was previously generated[20] by incorporating the crystal structure of the catalytic domain[8] into the paired-end complex (PEC) homology model described by Abrusan et al.[29] Figures 5A and 5B highlight the conserved structural positioning of the linker regions of the *Mos1* and SB100X transposases, including the WVPHEL motif in *Mos1* (Figure 5A) and Q124 in SB100X (Figure 5B).

To test if the hyperactivity of the Q124C variant is particularly evident at elevated levels of the transposase, we conducted a series of transposition assays with an output of puro-resistant colonies at varying amounts of SB transposase expression plasmids. We have previously shown that transposase protein concentration in the cell population correlates with transfected plasmid amounts in these assays.[24] The data revealed that, although Q124C is apparently subject to OPI (less colonies obtained with 2,500 ng transposase expression plasmid than with 500 ng, Figure 5C), its hyperactivity over SB100X is more evident at higher doses of expression plasmid, namely 500 and 2,500 ng (Figure 5C). Based on these observations, it can be inferred that the Q124C variant exhibits a degree of resistance to transposase inhibition under OPI conditions, thereby allowing utilization of the SB transposase in a wider range of concentration.

We have established a transgenic cell line derived from human HepG2 cells that serves as a reporter cell line that contains a single copy of a neomycin resistance gene (*neo*)-tagged SB transposon, which disrupts the open reading frame (ORF) of a puromycin resistance gene (*puro*). Consequently, the cells are resistant to G418 but sensitive to puromycin. When the SB transposase is expressed in these cells, it catalyzes excision of the SB transposon, followed by repair of the broken DNA ends through the non-homologous end joining double-strand DNA repair pathway.[30] This process results in the reconstitution of the *puro* ORF, leading to a selectable, puromycin-resistant phenotype. By selecting for puromycin resistance, the excision rate of different transposase variants can be measured, while double selection with G418 and puromycin allows for the determination of the transposition rate (Figure 5D). Upon transfecting the reporter cell line with the SB100X and Q124C variants, we detected apparent hyperactivity of Q124C already during the excision step that translates into a significant increase of the overall transposition rate (Figure 5E). This finding is consistent with our observation of a partial alleviation of OPI, as the previously proposed allosteric inhibition mechanism primarily affects the formation of the PEC that is required for transposon excision.[31] Taken together, these results demonstrate that the Q124C variant has a significantly increased transposase activity that manifests itself early in the transposition reaction, likely because it is less sensitive to OPI than SB100X.

The Q124C mutation leads to a partial rescue of K248R's transposition activity while maintaining its genome-wide integration pattern

Alternative SB transposase variants present certain benefits compared with SB100X. For example, the K248R mutation confers a safer integration profile with a lower tendency to integrate into genes and promoter regions, while displaying a higher propensity to integrate into genomic locations deemed as genomic safe harbors.[21] However, this variant's transposition activity is somewhat reduced when compared with SB100X,[21] so safety comes with an expense.

To rescue this activity loss, and to simultaneously test the hyperactive phenotype of Q124C in different polypeptide backgrounds, we

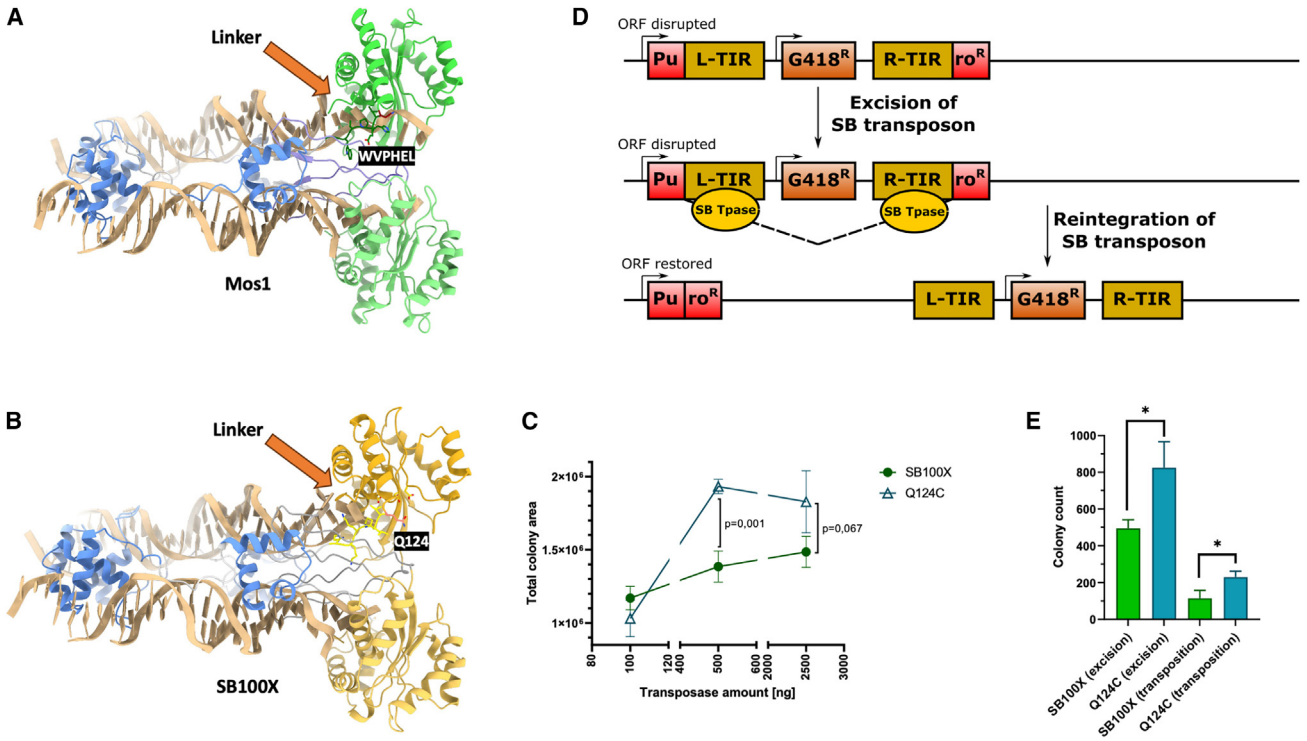

Figure 5. Mechanistic basis for hyperactivity of the Q124C mutation in the SB transposase

(A) Model of the *Mos1* transposase (green and blue), with the WVPHEL motif and the linker region (dark green) and the clamp loop protein-protein interaction interface (purple), bound to the transposon ends (brown). (B) Model of the SB100X transposase (orange and blue), with the Q124 side chain highlighted in red and the linker region (yellow) and the clamp loop protein-protein interaction interface (gray), bound to the transposon ends (brown). (C) A colony-forming transposition assay in human HeLa cells highlighting the different, transposase dose-dependent behaviors of the SB100X and Q124C mutant transposases as measured at 14 days after co-transfection with a puro-tagged SB transposon plasmid ($n = 3$; t test with p values indicating significance). Total colony area is depicted instead of colony count, because colony densities were too high for colonies to be separately detectable and countable. (D) Schematic representation of a transposon remobilization assay. The open reading frame (ORF) of a *puro* resistance cassette is disrupted by an SB transposon harboring a G418 resistance cassette. In the presence of SB transposase, the SB transposon can be excised, followed by DNA double-strand break repair restoring the *puro* ORF, which leads to resistance to puromycin. In the case of reintegration of the SB transposon encoding the *neo* resistance cassette, the cell remains resistant to G418. (E) A remobilization colony-forming assay in the human HepG2 cell line that contains a single copy of the reporter transposon as shown in (D). Upon transfection of these cells with SB100X or the Q124C mutant, the excision of the *neo* cassette was determined with puromycin-selection while transposition was assessed with G418 and puromycin double selection ($n = 3$; $^*p < 0.05$).

have crafted the Q124C mutation into an SB100X transposase variant containing K248R and performed standard colony-forming assays. We found that the inclusion of Q124C partially restores transposition activity of the K248R variant (Figure 6A). Vector copy number (VCN) analysis by droplet digital PCR (ddPCR) revealed that Q124C led to a slightly increased VCN as compared with SB100X-generated cells, whereas the incorporation of the Q124C mutation into the K248R variant resulted in slightly reduced VCNs (Figure 6B).

The integration of therapeutic gene constructs through SB transposition into safe sites in the human genome is crucial in minimizing the risk of insertional oncogenesis. These safe sites, also known as genomic "safe harbors" (GSHs), are specific regions within the human genome that can accommodate new DNA integrations without causing any adverse effects to the host cell. To qualify as a GSH, a chromosomal site or region must meet certain criteria, including (1) no overlap with transcription units, (2) a minimum distance of 50 kb from the 5′ end of any gene, a minimum distance of 300 kb from (3) cancer-related genes and (4) microRNA genes, and (5) exclusion from ultra-conserved elements.[32] The introduction of the Q124C mutation into the coding sequence of the SB100X transposase led to a minor elevation in the frequency of insertion into GSHs (Figures 6C and S2A). Assaying the representation of insertions into certain genetic features showed that, especially in exons the incorporation of the Q124C mutation had a lower insertion frequency (Figures 6D and S2B). In summary, Q124C increases the activity of the K248R variant without compromising the safety of its insertion profile.

DISCUSSION

Gene therapy is a promising solution for treating genetic diseases by offering lifelong and curative benefits. Various gene therapy methods involve permanent insertion of foreign genetic material into the

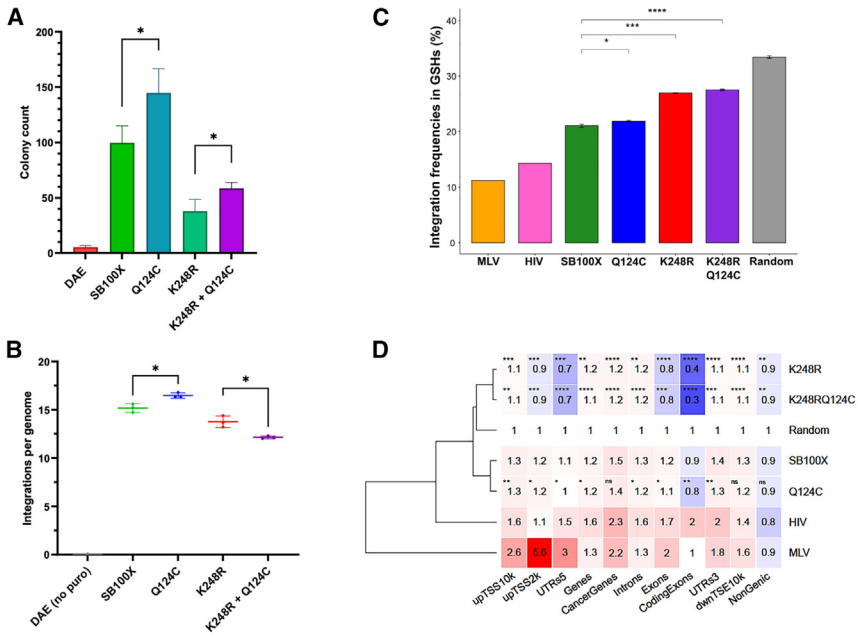

Figure 6. Transposition activity and genome-wide target site distribution of a Q124C/K248R double mutant of the SB100X transposase

(A) The introduction of the Q124C mutant to the K248R mutant partially rescues its transposition activity, as assayed by a colony-forming transposition assay in human HeLa cells ($n = 3$; $^{*}p < 0.05$). (B) VCN analysis of SB100X, Q124C, K248R, and Q124C/K248R double mutant ($n = 3$; $^{*}p < 0.05$). (C) Insertion frequency into genomic safe harbors ($n = 3$; $^{*}p < 0.05$, $^{***}p < 0.001$, $^{***}p < 0.0001$). (D) Fold change insertion frequency into genetic features ($n = 3$; ns, not significant; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{***}p < 0.0001$).

genome. While most gene delivery technologies rely on viruses to transduce cells, non-viral gene delivery systems such as the SB system have been gaining popularity. These systems have several advantages, including lower manufacturing costs due to the use of naked nucleic acids as vector components, a larger cargo capacity than viral systems, and a safer integration pattern across the host cell genome. The current state-of-the-art SB transposase, known as SB100X, is a hyperactive variant that has been clinically developed for CAR-T cell therapies.[33]

Here, we characterized a newly discovered and highly active variant of the SB100X transposase, which we name SB200X. This variant is the result of a specific amino acid substitution (Q124C), and it shows a ~2-fold hyperactivity (as compared with SB100X) in various human cell types, including standard tissue culture cell lines (such as HeLa and HepG2) as well as more clinically relevant cell types, such as hiPSCs and mHSPCs (or Lin^{-} cells) (Figures 3, 4, and 5). Importantly, SB200X displayed a hyperactive phenotype regardless of whether the DNA vector that carried the transposon to be mobilized was a plasmid or a MC (Figures 3 and 4). We have also observed that other amino acid substitutions at this position can similarly increase transposition efficiency (Figure 3).

Previous work has highlighted the potential to increase the transposition activity of the SB transposase by incorporating amino acid substitutions along the transposase polypeptide (reviewed by Hudecek et al.[34]). These hyperactive SB transposases were engineered by systematic alanine scanning,[14,15] by incorporating single amino acids or small (aa 2–7) blocks of amino acids from related transposases,[16–18] and by replacement of selected amino acid residues based on charge.[17] From these, only a few have been connected to an alteration in the biochemistry of transposition, including K33A, which displays increased DNA binding to the transposon TIRs,[14] and the RKEN214-217DAVQ mutations, which have an impact on the flexibility/rigidity of the SB transposase.[35] An important clue for the mechanism through which Q124C results in hyperactivity of the transposase comes from our observation that higher rates of transposition were observed at high transposase concentrations (Figure 5C), suggesting that the mutation affects, and partially alleviates, OPI.

OPI has been originally described as an autoregulatory mechanism controlling the activity of the *Mos1 mariner* transposon[36] and later extended to be applicable to other transposon families as well, including SB, *piggyBac*, and *Tol2*.[24] Biochemical experiments with the *Hsmar1* transposase revealed that *mariner* transposition is initiated by a transposase dimer bound to one transposon end (single-end complex [SEC] in Figure 1) followed by recruitment of the second, naked transposon end to form a transpososome (paired-end complex [PEC] in Figure 1) (as opposed to binding of transposase monomers at both ends of the transposon followed by synapsis).[31] The finding that the *Hsmar1* transposase binds fast and tight to the first transposon end, but has a much lower affinity for the second transposon end due to an allosteric change between the transposase subunits, led to the assembly-site-occlusion (ASO) model for OPI.[31] ASO posits that, as the concentration of transposase increases, unbound transposase dimers will outcompete the SEC in capturing the second end of the transposon; a process that poisons the transposition reaction because it progressively reduces the number of unbound ends available for synapsis.[31] Importantly, subsequent studies have shown that mutations in the conserved WVPHEL motif within the linker region of the *Hsmar1* transposase can lead to hyperactive variants, likely because they counteract the negative allosteric change, assemble the transpososome more quickly, and thus are less sensitive to OPI.[23]

Several observations allow us to propose that the mechanism through which the Q124C mutation confers a hyperactive phenotype to the SB transposase might be very similar to the ASO model described above.

First, as with the *mariner* transposons, SB transposition also proceeds through the formation of a SEC on one transposon end that subsequently captures the other transposon end.[15] Second, structural information revealed that the clamp loop of an SB transposase monomer interacts with the inter-domain linker region of a partner molecule,[8] resembling the clamp loop-linker interaction observed in the *Mos1* PEC.[13] Importantly, although SB and other Tc1 family transposases do not have the WVPHEL motif (Figure 2B), the part of the linker region in SB responsible for this interaction is structurally equivalent to the WVPHEL regulatory motif of *mariner* transposases.[8] Finally, the Q124C amino acid substitution responsible for hyperactivity is located right at the edge of the inter-domain linker region of the SB transposase (Figures 2B and 5B). Consistent with the hypothesis that the hyperactivity of Q124C is linked to interference with the process of OPI, our new variant exhibits hyperactivity during the excision step (Figure 5E), whose rate is directly influenced by the rate of synapsis.

We have shown that the Q124C mutation can be combined with other relevant SB transposase variants, such as the K248R variant. The K248R version has a safer integration profile than SB100X, due to redirecting SB-mediated integrations away from coding exons and transcriptional regulatory elements, thereby leading to an enrichment of insertions into genomic safe harbors.[21] However, this variant is less efficient in transposition than SB100X. Importantly, we observed a positive impact of Q124C on the transposition rate of K248R, while fully preserving or even slightly improving its safer integration profile (Figures 6C and 6D). It is plausible that other SB transposase variants may also be augmented with the Q124C mutation to enhance their transposition activity. In summary, our data demonstrate the efficacy and safety of the novel SB200X hyperactive variant of the SB transposase in genetic applications, such as gene therapy.

MATERIALS AND METHODS

Site-directed mutagenesis of the SB100X transposase

Q124C mutation were generated based on the template plasmid pcGlobin2-SB100X (containing the CMV promoter) using the Q5 polymerase (NEB, Ipswich, MA). 5′-Phosphorylated primers for the particular positions were designed with 5′ ends annealing back-to-back. To facilitate an intramolecular ligation reaction downstream, primers were synthesized with a 5′-phosphate and were procured from Eurofins (Eurofins/MWG, Luxembourg). Primer sequences (SB100X-Q124_to_C-rev & SB100X-Q124_to_C124-fwd) are listed in Table S1. The PCR cycling parameters were established in accordance with the guidelines provided by the manufacturer. The annealing temperatures of the mutagenic primers were calculated with the "NEB Tm calculator" software (https://www.neb.com/tools-and-resources/interactive-tools/tm-calculator). The QIAquick PCR Purification Kit (QIAGEN, Venlo, the Netherlands) was utilized to purify the PCR products that were eluted in a volume of 30 μL. The products were subjected to digestion with 2 μL *Dpn*I (NEB) for 2 h at 37°C, followed by heat-inactivation at 80°C for 20 min. To circularize the linear, double-stranded PCR products, T4 DNA ligase (NEB) was added overnight at a temperature of 16°C. The resulting circularized PCR products were transformed into chemically competent *E. coli* (Invitrogen/Life Technologies, Carlsbad, CA) and cultured in Luria-Bertani (LB) medium for 1 h. The transformed cells were subsequently selected for ampicillin resistance by plating them on LB agar plates containing 100 μg/mL ampicillin. To verify the existence of the intended mutations and the lack of any unintended mutations, plasmid DNA isolated from multiple colonies was purified using the QIAprep spin miniprep kit (QIAGEN) and subjected to Sanger sequencing analysis performed by Eurofins (Eurofins/MWG). The additional 18 mutants for the experiment displayed in Figure 3E were ordered from BioCat (Heidelberg, Germany).

Transposition assay

For the transposition assays in HepG2 and HeLa cells, 3×10^5 cells were seeded into 6-well plates 1 day before transfection in DMEM medium (Thermo Fischer Scientific, Waltham, MA). Transfection was carried out using TransIT-LT1 transfection reagent (Mirus Bio, Madison, WI), following the manufacturer's protocol. To ensure equal transfection efficiencies in transfection reactions with different amounts of transposase expression plasmids, each plasmid mixture was supplemented with different amounts of the nonrelevant plasmid pmaxGFP to reach the same total DNA amount. The transposase expression plasmids (pcGlobin2) expressing SB100X or mutants were co-transfected with a transposon donor plasmid harboring the puromycin resistance gene (pT2/Puro, containing the SV40 promoter). Cells were trypsinized 48-h post-transfection, and 1%–10% of the cells were replated into 10-cm dishes that were cultured in medium containing 1 μg/mL puromycin (InvivoGen, San Diego, CA) for selection of cells with transposon integrations. Medium was changed every 2 days. At day 14, the colonies were fixed with a 10% (vol/vol) solution of formaldehyde in phosphate-buffered saline (PBS), stained with methylene blue in PBS and counted. For the purpose of *in vitro* assessments of transposition efficiencies, at least three independent experiments were conducted.

The colony-forming assay depicted in Figure 3A was performed by transfecting 50 ng of pT2/Puro and 500 ng of the transposase expression plasmids (pcGlobin2) and plating 10% of the HepG2 cells for selection in 10-cm dishes. The colony-forming assay shown in Figure 3E was performed by transfecting 500 ng of pT2/Puro and 500 ng of the transposase expression plasmids (pcGlobin2) and plating 1% of the HeLa cells for selection in 10-cm plates. The colony-forming assay depicted in Figure 5C addressing OPI was performed by transfecting 500 ng of pT2/Puro and the respective amounts of transposase expression plasmids (pcGlobin2) (100, 500, 2,500 ng). For selection, 1% of the HeLa cells were seeded in 10-cm dishes. The transposition assay shown in Figure 6A was conducted by transfecting 20 ng of pT2/Puro with 200 ng of the transposase expression plasmids (pcGlobin2) and plating of 5% of the HeLa cells for selection on 10-cm plates.

For the analyses of VCN and target site selection properties of the SB100X transposase mutants, 4×10^5 HepG2 cells were seeded onto 6-well plates 1 day before transfection. Transfections were performed with QIAGEN-purified plasmid DNA using TransIT-LT1

transfection reagent according to the manufacturer's protocol. One microliter of TransIT-LT1 transfection reagent was used to transfect 550 ng of DNA, including 500 ng pT2Bpuro, 50 ng helper plasmid expressing the mutant transposases, or SB100X. The cells were trypsinized 48 h after transfection, diluted to multiple 10-cm dishes containing DMEM supplemented with 1 μg/mL puromycin and selected for growth over a period for 2 weeks. At least 10,000 puromycin-resistant HepG2 cell colonies were trypsinized and centrifuged at 1,000 rpm for 5 min. The pellet was washed with PBS and genomic DNA (gDNA) was extracted from cells using the QIAGEN DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's protocol.

ddPCR for VCN analysis

Puromycin-resistant HepG2 colonies were trypsinized and pelleted 14 days post-transfection. The pellet was treated with 500 μL medium containing 100 μg/mL DNaseI for 10 min at room temperature to degrade extracellular plasmid DNA eventually remaining in the culture. An equal volume of PBS containing 0.1 M EDTA was then added, the mixture was centrifuged for 5 min at 1,000 rpm and washed three times with PBS containing 0.1 M EDTA. Ten microliters of 0.5 M EDTA was added to the final pellet to fully inactivate the DNaseI. gDNA was then isolated and purified using the DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instructions. gDNA (50 ng) were digested overnight with 5 units of *Dpn*I and 2.5 units of *Mlu*CI (NEB) to degrade non-integrated transposon plasmid DNA and to fragment the gDNA, respectively. Fragmented gDNA (10 ng) were mixed with 1X ddPCR Supermix for Probes (No dUTP) (Bio-Rad, Hercules, CA, USA), 511.4 nM of puromycin- and RPP30-specific primers, and 255.7 nM of puromycin- and RPP30-specific probes (Table S1). Droplets were made with the QX100 Droplet Generator (Bio-Rad) following the manufacturer's instructions, and a PCR was performed in the C1000 Touch Thermal Cycler (Bio-Rad) using the following program: 95°C for 10 min; 40 cycles at 94°C for 30 s, 50°C for 30 s, and 60°C for 1 min; and 98°C for 10 min. The fluorescence of each droplet was then analyzed in the QX100 Droplet Reader (Bio-Rad) and the copy numbers were calculated with the QuantaSoft software (Bio-Rad).

Western blotting

One day before to transfection, 3×10^5 HeLa cells were seeded onto 6-well plates. Transfection was carried out using TransIT-LT1 transfection reagent (Mirus Bio), according to the manufacturer's protocol. Cells were transfected with 1,000 ng of DNA expressing the SB100X transposase, the Q124C mutant (both in pcGlobin2) or a GFP expression plasmid (pmaxGFP; Lonza, Basel, Switzerland). After 48 h, the cells were lysed with RIPA buffer (150 mM NaCl, 1.0% Triton X-100, 1.0% Na-deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 25 mM Tris [pH 8.0]) containing a protease inhibitor cocktail (Complete Mini, Roche, Basel, Switzerland). Protein was extracted using BioruptorPlus (Diagenode, Denville, NJ), 2 cycles, high power, 30 s ON/30 s OFF at 4°C, and total protein amount was determined with the BCA Protein Assay Kit (Pierce, Rockford, IL). Protein lysate (10 μg) was loaded per lane onto 10% polyacrylamide gels and subjected to SDS-polyacrylamide gel electrophoresis. Proteins were transferred to nitrocellulose membranes (Hybond ECL, Amersham Bioscience, Little Chalfont, UK) and immunoblotting was performed using standard procedures. The proteins were detected with a goat monoclonal anti-SB transposase antibody (R&D Systems, Minneapolis, MN) at dilution 1:1,000, or a mouse monoclonal anti-Actin antibody (Thermo Fisher Scientific) at dilution 1:10,000, and chemiluminescence using ECL Prime Western Blotting Detection Kit (Amersham Bioscience). Signals were captured on an ECL-imager (ChemoCam Imager ECL, Intas, Göttingen, Germany).

Genetic modification of hiPSCs with the SB system

hiPSCs were cultured and handled following the protocol of Skarnes et al.[37] Transfection was done with TransIT-LT1 transfection reagent (Mirus Bio) following the manufacturer's protocol for hiPSCs. hiPSCs (2×10^6) were co-transfected with 2 μg of the expression plasmid (pcGlobin2) expressing either SB100X, Q124C mutant hyperactive transposase, or catalytically inactive SB100X(DAE) transposase together with1 μL of the transposon donor MC expressing Venus fluorescent protein (MC.T2-CAGGS-Venus, shortly MC-Venus[27]). Venus expression was assessed every 3–4 days by cytofluorimetric analysis. For the purpose of *in vitro* assessments of transposition efficiencies, at least three independent experiments were conducted.

Structural modeling

The model of the *Mos1* and SB100X transposase in complex with transposon DNA was generated as in Kesselring et al.[20] In brief, it was assembled based on the SB PEC homology model described in Abrusan et al.,[29] where the transposase catalytic domain was replaced with the crystal structure.[8] Finally, the assembled model was refined in HADDOCK.[38] Structural models were visualized using ChimeraX.

Transposon remobilization assay

For the remobilization assay in the modified HepG2 reporter cell line (HepG2-PB(SB2#23)), 4×10^5 cells were seeded onto 6-well plates 1 day before transfection. Transfection was done with Lipofectamine 3000 transfection reagent (Invitrogen, Waltham, MA) following the manufacturer's protocol. One microgram of the plasmids (pcGlobin2) expressing SB100X or the Q124C transposases was transfected to the HepG2-PB(SB2#23) cells. Seventy-two hours post-transfection, cells were trypsinized and 100% of the cells were seeded on 10-cm dishes, and selected in DMEM supplemented with 1 μg/mL puromycin (InvivoGen, San Diago, CA) for transposon excision, or double selected for transposition using DMEM supplemented with 1 μg/mL puromycin and 1 mg/mL G418 (InvivoGen). After 3 weeks of selection, resistant cell colonies were fixed with 10% (vol/vol) formaldehyde in PBS, stained with methylene blue in PBS, and counted. For the purpose of *in vitro* assessments of transposition efficiencies, at least three autonomous experiments were conducted.

Nucleofection of primary murine Lin$^-$ cells

Lin$^-$ cells isolated from the bone marrow of C57BL/6-Ly5.1 mice were cultured for 2 days in StemSpan SFEM Stem Span medium

(STEMCELL Technologies, Vancouver, Canada) supplemented with 1% Pen/Strep (Thermo Fischer Scientific) and 10 ng/mL Recombinant Murine GM-CS, 25 ng/mL Recombinant Murine SCF, 25 ng/mL Recombinant Murine IL-3, 25 ng/mL Recombinant Murine IL-11, and 25 ng/mL Recombinant Murine TPO (all from PeproTech, Hamburg, Germany). At day 2 after isolation, approximately 1×10^6 live Lin^- cells were nucleofected with 0.5 μg MC-Venus (containing the CAGGS promoter) together with either 1 or 2 μg of the plasmid expressing SB100X or the Q124C mutant or the SB100X(DAE) inactive transposase using the 4D Nucleofector (Lonza) and P3 Primary Cell 4D-Nucleofector Kit L and DK-100 program. After 10 min incubation at room temperature, 500 μL StemSpan SFEM Stem Span medium supplemented with cytokines was added into each cuvette. After another 15 min incubation at 37°C, the cells were transferred into a 24-well plate. Venus expression was monitored every 3–4 days by flow cytometry.

Flow cytometry analysis

The efficiency of transient and stable Venus expression in Lin^- cells was assessed by flow cytometry at days 2, 6, 9, 12, and 16 post-nucleofection. Cell viability was determined using the Zombie NIR Fixable Viability Kit (BioLegend, San Diego, CA) according the manufacturer's protocol. For this, approximately 10% of the cell culture volume was collected, washed with PBS, and stained with Zombie NIR followed by flow cytometric analysis with the BD LSR II SORP Flow Cytometer (BD FACS, East Rutherford, NJ). The results were analyzed using FlowJo Software (FlowJo, Asheland, OR). $Venus^+$ cells were gated within the Zombie-negative population.

Generation of SB insertion libraries

The insertion site libraries were prepared as described earlier.[39] Two μg DNA was sheared with a Covaris M220 ultra-sonicator device to an average fragment size of 600 bp in Screw-Cap microTUBEs in 50 μL, using the following settings: peak incident power 50 W, duty factor 20%, cycles per burst 200, treatment 28 s; 1.2 μg of the sheared DNA was blunted and 5′-phosphorylated using the NEBNext End Repair Module (NEB), and 3′-A-tailed with NEBNext dA-Tailing Module (NEB) following the recommendations of the manufacturer. The DNA was purified with the Clean and Concentrator Kit Zymo Research Europe, Freiburg, Germany), and eluted in 8 μL 10 mM Tris (pH 8.0) (EB) for ligation with 50 pmol of T-linker (see below) with T4 ligase (NEB) in 20 μL volume, at 16°C, overnight. T-linkers were created by annealing 100 pmol of each of the oligonucleotides Linker_TruSeq_T+ and Linker_TruSeq_T– in 10 mM Tris-Cl (pH 8), 50 mM NaCl, 0.5 mM EDTA (Table S1). After heat-inactivation, ligation products enclosing fragments of non-integrated transposon donor plasmid DNA were digested with *Dpn*I (NEB) in 50 μL for 3 h, and the DNA was column-purified and eluted in 20 μL elution buffer (EB). Six microliters of elute was used for PCR I with 25 pmol of the primers specific for the linker and for the transposon inverted repeat: Linker and T-Bal-Long (Table S1), respectively, with the conditions: 98°C for 30 s; 10 cycles of 98°C for 10 s, 72°C for 30 s; 15 cycles of 98°C for 10 s, ramp to 62°C (1°C/s) for 30 s, 72°C for 30 s and 72°C for 5 min. All PCR reactions were performed with NEBNext High-Fidelity 2 × PCR Master Mix. The PCR was column purified and the DNA was eluted in 20 μL EB; 10 μL was used for PCR II with primers Nested and LAM-SB-50 (Table S1), with the following program: 98°C for 30 s, then 20 cycles of 64°C for 30 s, 72°C for 30 s, then 72°C for 5 min. PCR III was performed with primers PE-nest-ind-N and SB-20-bc-ill-N (Table S1) (where N is the number of the Illumina TrueSeq indexes) for barcoding the samples using the following PCR program: 98°C for 30 s; 12 cycles of 98°C for 10 s, ramp to 64°C (1°C/s) for 30 s, 72°C for 30 s, and 72°C for 5 min. The final PCR products were separated on a 1% agarose gel and the smears of 200–500 bp were isolated, purified, and sequenced on an Illumina NextSeq550 instrument using 1 × 86 nucleotide setting.

Sequencing and analyses of the insertion sites

After adapter and quality trimming (Phred score $\geq$20) with fastp,[40] the reads were tested for the transposon sequences downstream of the SB-specific primer and were filtered for the presence of the remaining part of the transposon TIR and a minimum length of 28 bases of genomic sequence for mapping with bowtie2[41] with –sensitive and –end-to-end settings to the hg38 human assembly. A mapped locus was considered valid if the mapping quality for the read supporting it was $\geq$20. Any insertion site needed to be supported by at least 10 independent reads at a TA target site of the human genome (hg38). If multiple insertions were detected within 10 bases, the insertion site supported by the largest number of independent reads was considered as the valid one. Integration sites can be found in Data S1. Representation of the insertion sites in various genic categories were investigated using the Genomation package[42] in the *R* environment (version 3.5). A computationally generated random set of 100,000 loci of the human hg38 genome assembly was used as reference for investigations of insertion site representation in various genomic bins. The genomic safe harbor coordinates and the cancer gene annotations (using the allOnco file of the Bushman laboratory) for the hg38 assembly were created following the earlier defined criteria.[32] The gene-related annotations for the human genome were downloaded from the Table Browser of the UCSC Genome Browser (https://genome.ucsc.edu/).

DATA AND CODE AVAILABILITY

All primary experimental data are available from the authors. The lists of insertion sites are available as supplemental files.

ACKNOWLEDGMENTS

The lists of insertion sites of HIV and MLV vectors in the human genome were generous gifts from Fulvio Mavilio (University of Modena). We thank Vladimir Arinkin and Orsolya Barabas (Department of Molecular and Cellular Biology, University of Geneva, Switzerland) for their valuable input in structural modeling of transposon/transposase complexes, preparation of figures, and critically reading the manuscript. This project received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 754658 (CARAMBA).

AUTHOR CONTRIBUTIONS

M.T.O. generated the Q124C SB transposase mutant, tested it for expression and excision/transposition activity in HeLa, HepG2, and hiPSCs, generated PCR libraries representing transposon insertions, prepared the figures, and wrote the first draft of the paper.

C.M. sequenced the PCR libraries representing transposon insertions and bioinformatically analyzed the insertion sites. L.B. assessed the transposition activity of Q124C in mouse Lin$^-$ cells. N.S.-V. did VCN analysis. T.D. carried out the experiment reported in Figure 3E. Z.I. conceived the study, supervised the experimental work, and wrote the paper.

DECLARATION OF INTERESTS

M.T.O. and Z.I. are co-inventors on a patent relating to hyperactive SB transposases described in this paper. M.T.O. is now an employee of BioNTech, Mainz, Germany, but conducted the work presented in this paper entirely at the Paul-Ehrlich-Institut.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2024.102381.

REFERENCES

1. McClintock, B. (1953). Induction of Instability at Selected Loci in Maize. Genetics *38*, 579–599.
2. Wicker, T., Sabot, F., Hua-Van, A., Bennetzen, J.L., Capy, P., Chalhoub, B., Flavell, A., Leroy, P., Morgante, M., Panaud, O., et al. (2007). A unified classification system for eukaryotic transposable elements. Nat. Rev. Genet. *8*, 973–982.
3. Ochmann, M.T., and Ivics, Z. (2021). Jumping Ahead with Sleeping Beauty: Mechanistic Insights into Cut-and-Paste Transposition. Viruses *13*, 76.
4. Amberger, M., and Ivics, Z. (2020). Latest Advances for the Sleeping Beauty Transposon System: 23 Years of Insomnia but Prettier than Ever: Refinement and Recent Innovations of the Sleeping Beauty Transposon System Enabling Novel, Nonviral Genetic Engineering Applications. Bioessays *42*, e2000136.
5. Sandoval-Villegas, N., Nurieva, W., Amberger, M., and Ivics, Z. (2021). Contemporary Transposon Tools: A Review and Guide through Mechanisms and Applications of Sleeping Beauty, piggyBac and Tol2 for Genome Engineering. Int. J. Mol. Sci. *22*, 5084.
6. Ivics, Z., Hackett, P.B., Plasterk, R.H., and Izsvák, Z. (1997). Molecular reconstruction of *Sleeping Beauty*, a Tc1-like transposon from fish, and its transposition in human cells. Cell *91*, 501–510.
7. Ivics, Z., Izsvák, Z., Minter, A., and Hackett, P.B. (1996). Identification of functional domains and evolution of Tc1-like transposable elements. Proc. Natl. Acad. Sci. USA *93*, 5008–5013.
8. Voigt, F., Wiedemann, L., Zuliani, C., Querques, I., Sebe, A., Mátés, L., Izsvák, Z., Ivics, Z., and Barabas, O. (2016). Sleeping Beauty transposase structure allows rational design of hyperactive variants for genetic engineering. Nat. Commun. *7*, 11126.
9. Izsvák, Z., Khare, D., Behlke, J., Heinemann, U., Plasterk, R.H., and 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.
10. Carpentier, C.E., Schreifels, J.M., Aronovich, E.L., Carlson, D.F., Hackett, P.B., and Nesmelova, I.V. (2014). NMR structural analysis of Sleeping Beauty transposase binding to DNA. Protein Sci. *23*, 23–33.
11. Konnova, T.A., Singer, C.M., and Nesmelova, I.V. (2017). NMR solution structure of the RED subdomain of the Sleeping Beauty transposase. Protein Sci. *26*, 1171–1181.
12. Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. *42*, W320–W324.
13. Richardson, J.M., Colloms, S.D., Finnegan, D.J., and Walkinshaw, M.D. (2009). Molecular architecture of the Mos1 paired-end complex: the structural basis of DNA transposition in a eukaryote. Cell *138*, 1096–1108.
14. Yant, S.R., Park, J., Huang, Y., Mikkelsen, J.G., and 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.
15. Wang, Y., Pryputniewicz-Dobrinska, D., Nagy, E.É., Kaufman, C.D., Singh, M., Yant, S., Wang, J., Dalda, A., Kay, M.A., Ivics, Z., and Izsvák, Z. (2017). Regulated complex assembly safeguards the fidelity of Sleeping Beauty transposition. Nucleic Acids Res. *45*, 311–326.
16. Geurts, A.M., Yang, Y., Clark, K.J., Liu, G., Cui, Z., Dupuy, A.J., Bell, J.B., Largaespada, D.A., and Hackett, P.B. (2003). Gene transfer into genomes of human cells by the sleeping beauty transposon system. Mol. Ther. *8*, 108–117.
17. Zayed, H., Izsvák, Z., Walisko, O., and Ivics, Z. (2004). Development of hyperactive sleeping beauty transposon vectors by mutational analysis. Mol. Ther. *9*, 292–304.
18. Baus, J., Liu, L., Heggestad, A.D., Sanz, S., and Fletcher, B.S. (2005). Hyperactive transposase mutants of the Sleeping Beauty transposon. Mol. Ther. *12*, 1148–1156.
19. Mates, L., Chuah, M.K., Belay, E., Jerchow, B., Manoj, N., Acosta-Sanchez, A., Grzela, D.P., Schmitt, A., Becker, K., Matrai, J., et al. (2009). Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nature Genet. *41*, 753–761.
20. Kesselring, L., Miskey, C., Zuliani, C., Querques, I., Kapitonov, V., Laukó, A., Fehér, A., Palazzo, A., Diem, T., Lustig, J., et al. (2020). A single amino acid switch converts the Sleeping Beauty transposase into an efficient unidirectional excisionase with utility in stem cell reprogramming. Nucleic Acids Res. *48*, 316–331.
21. Miskey, C., Kesselring, L., Querques, I., Abrusán, G., Barabas, O., and Ivics, Z. (2022). Engineered Sleeping Beauty transposase redirects transposon integration away from genes. Nucleic Acids Res. *50*, 2807–2825.
22. Querques, I., Mades, A., Zuliani, C., Miskey, C., Alb, M., Grueso, E., Machwirth, M., Rausch, T., Einsele, H., Ivics, Z., et al. (2019). A highly soluble Sleeping Beauty transposase improves control of gene insertion. Nat. Biotechnol. *37*, 1502–1512.
23. Liu, D., and Chalmers, R. (2014). Hyperactive mariner transposons are created by mutations that disrupt allosterism and increase the rate of transposon end synapsis. Nucleic Acids Res. *42*, 2637–2645.
24. Grabundzija, I., Irgang, M., Mátés, L., Belay, E., Matrai, J., Gogol-Döring, A., Kawakami, K., Chen, W., Ruiz, P., Chuah, M.K.L., et al. (2010). Comparative analysis of transposable element vector systems in human cells. Mol. Ther. *18*, 1200–1209.
25. Novikova, O., Topilina, N., and Belfort, M. (2014). Enigmatic distribution, evolution, and function of inteins. J. Biol. Chem. *289*, 14490–14497.
26. Tornabene, P., Trapani, I., Minopoli, R., Centrulo, M., Lupo, M., de Simone, S., Tiberi, P., Dell'Aquila, F., Marrocco, E., Iodice, C., et al. (2019). Intein-mediated protein trans-splicing expands adeno-associated virus transfer capacity in the retina. Sci. Transl. Med. *11*, eaav4523.
27. Holstein, M., Mesa-Nuñez, C., Miskey, C., Almarza, E., Poletti, V., Schmeer, M., Grueso, E., Ordóñez Flores, J.C., Kobelt, D., Walther, W., et al. (2018). Efficient Non-viral Gene Delivery into Human Hematopoietic Stem Cells by Minicircle Sleeping Beauty Transposon Vectors. Mol. Ther. *26*, 1137–1153.
28. Ferrari, G., Thrasher, A.J., and Aiuti, A. (2021). Gene therapy using haematopoietic stem and progenitor cells. Nat. Rev. Genet. *22*, 216–234.
29. Abrusan, G., Yant, S.R., Szilagyi, A., Marsh, J.A., Mates, L., Izsvak, Z., Barabas, O., and Ivics, Z. (2016). Structural Determinants of Sleeping Beauty Transposase Activity. Mole. Ther. J. Am. Soc. Gene Ther. *24*, 1369–1377.
30. Izsvák, Z., Stüwe, E.E., Fiedler, D., Katzer, A., Jeggo, P.A., and Ivics, Z. (2004). Healing the wounds inflicted by *Sleeping Beauty* transposition by double-strand break repair in mammalian somatic cells. Mol. Cell *13*, 279–290.
31. Claeys Bouuaert, C., Lipkow, K., Andrews, S.S., Liu, D., and Chalmers, R. (2013). The autoregulation of a eukaryotic DNA transposon. Elife *2*, e00668.
32. Sadelain, M., Papapetrou, E.P., and Bushman, F.D. (2011). Safe harbours for the integration of new DNA in the human genome. Nat. Rev. Cancer *12*, 51–58.
33. Prommersberger, S., Reiser, M., Beckmann, J., Danhof, S., Amberger, M., Quade-Lyssy, P., Einsele, H., Hudecek, M., Bonig, H., and Ivics, Z. (2021). CARAMBA: a first-in-human clinical trial with SLAMF7 CAR-T cells prepared by virus-free Sleeping Beauty gene transfer to treat multiple myeloma. Gene Ther. *28*, 560–571.
34. Hudecek, M., Izsvák, Z., Johnen, S., Renner, M., Thumann, G., and Ivics, Z. (2017). Going non-viral: the Sleeping Beauty transposon system breaks on through to the clinical side. Crit. Rev. Biochem. Mol. Biol. *52*, 355–380.

35. Singer, C.M., Joy, D., Jacobs, D.J., and Nesmelova, I.V. (2019). Rigidity and flexibility characteristics of DD[E/D]-transposases Mos1 and Sleeping Beauty. Proteins *87*, 313–325.

36. Lohe, A.R., and Hartl, D.L. (1996). Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. Mol. Biol. Evol. *13*, 549–555.

37. Skarnes, W.C., Pellegrino, E., and McDonough, J.A. (2019). Improving homology-directed repair efficiency in human stem cells. Methods *164–165*, 18–28.

38. Dominguez, C., Boelens, R., and Bonvin, A.M.J.J. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J. Am. Chem. Soc. *125*, 1731–1737.

39. Kovač, A., Miskey, C., and Ivics, Z. (2023). Sleeping Beauty Transposon Insertions into Nucleolar DNA by an Engineered Transposase Localized in the Nucleolus. Int. J. Mol. Sci. *24*, 14978.

40. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890.

41. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

42. Akalin, A., Franke, V., Vlahovíček, K., Mason, C.E., and Schübeler, D. (2015). Genomation: a toolkit to summarize, annotate and visualize genomic intervals. Bioinformatics *31*, 1127–1129.