Published in final edited form as:
*Nat Biotechnol*. 2019 December ; 37(12): 1502–1512. doi:10.1038/s41587-019-0291-z.

# A highly soluble Sleeping Beauty transposase improves control of gene insertion

**Irma Querques**[#1], **Andreas Mades**[#2], **Cecilia Zuliani**[#1], **Csaba Miskey**[3], **Miriam Alb**[2], **Esther Grueso**[3], **Markus Machwirth**[2], **Tobias Rausch**[4], **Hermann Einsele**[2], **Zoltán Ivics**[3], **Michael Hudecek**[2,*], **Orsolya Barabas**[1,*]

[1]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany

[2]Medizinische Klinik und Poliklinik II, Universitätsklinikum Würzburg, Würzburg, Germany

[3]Division of Medical Biotechnology, Paul Ehrlich Institute, Langen, Germany

[4]Genomics Core Facility, European Molecular Biology Laboratory, Heidelberg, Germany

[#] These authors contributed equally to this work.

## Abstract

The *Sleeping Beauty* (SB) transposon system is an efficient non-viral gene transfer tool in mammalian cells but its broad use has been hampered by uncontrolled transposase gene activity from DNA vectors, posing a risk for genome instability, and by the inability to use transposase protein directly. Here, we used rational protein design based on the crystal structure of the hyperactive SB100X variant to create an SB transposase (hsSB) with enhanced solubility and stability. We demonstrate that hsSB can be delivered with transposon DNA to genetically modify cell lines and embryonic, hematopoietic and induced pluripotent stem cells (iPSCs), overcoming uncontrolled transposase activity. We used hsSB to generate chimeric antigen receptor (CAR) T-



Michael Hudecek, Orsolya Barabas; Hudecek_M@ukw.de, barabas@embl.de.
[*]These authors jointly supervised this work

**Reporting Summary**
Further information on research design is available in the Life Sciences Reporting Summary linked to this article.

**Data availability**
Data generated or analyzed during this study and all unique materials can be made available by the corresponding authors.

**Author contributions**
I.Q., A.M., C.Z., M.H. and O.B. designed research. All authors contributed to analysis and discussion of the results, read and approved the manuscript. I.Q., A.M., C.Z., M.H. and O.B. wrote the manuscript with input from all authors. I.Q. designed, screened and characterized transposase variants. A.M., M.A., M.M. performed T cell engineering and characterization. C.Z. conducted transposition assays and engineering of HeLa, CHO, mESC and iPSC lines. C.M. performed integration site profiling in T cells and copy number analyses. E.G. performed HSPC engineering. T.R. analyzed integration sites in HeLa cells. H.E. provided expert advice and support during the project. Z.I. supervised integration profiling, copy number analyses, HSPC engineering and participated in coordinating the research. M.H. supervised T cell engineering and analyses and oversaw the project. O.B. supervised protein work, cell line engineering and oversaw the project.

**Competing interests**
Z.I. is inventor on patents concerning the development and use of the Sleeping Beauty technology (Proprietor: Max-Delbrück-Centrum für Molekulare Medizin; Patent Nos. US9228180B2 and EP2160461B1) and two patent applications have been filed concerning the hsSB transposase.


Europe PMC Funders Author Manuscripts

cells, which exhibit potent anti-tumor activity *in vitro* and in xenograft mice. We found that hsSB spontaneously penetrates cells, enabling modification of iPSCs and generation of CAR-T cells without the use of transfection reagents. Titration of hsSB to modulate genomic integration frequency achieved as few as two integrations per genome.

Introduction of desired transgenes in cells and organisms has emerged as a crucial technology for research and biotechnology, and clinical application of *ex vivo* engineered human cells has demonstrated their therapeutic potential in regenerative medicine and cancer therapy. For instance, the use of reprogrammed T cells that incorporate genetic information for a chimeric antigen receptor (CAR) has lately emerged as a new pillar in cancer treatment, showing remarkable response rates in the treatment of leukemia and lymphoma[1–3]. In these therapies, CARs serve as synthetic immune receptors that provide T cells with a new specificity against malignancy-associated antigens, thus directing the immune system to attack and eradicate tumor cells.

To introduce a CAR gene, current protocols rely on viral vectors, which provide efficient gene transfer, but their manufacturing and clinical use is lengthy and expensive. Viral vector-encoded epitopes also bear a risk for inflammatory responses[4], and preferential cargo integration in transcribed regions may lead to adverse genomic changes[5]. The use of non-viral vectors could improve safety and reduce cost, but has been constrained by moderate gene transfer efficiency, limited transgene size and cytotoxicity of vector DNA or RNA[6,7]. For instance, non-viral genome editing nucleases enable site-specific genome modifications with simplicity and low cost, but they depend on homology directed repair for DNA insertion, which is generally infrequent in primary cells and compromises insertion of large transgenes (such as a ~3 kb CAR gene)[8].

DNA transposons constitute a further non-viral alternative for gene delivery. They comprise two essential components: the transposase enzyme and the transposon DNA that contains a genetic cargo flanked by specific DNA end sequences. Conventionally, both components are provided as plasmid DNA vectors and the transposase is expressed in the target cells. After expression, the transposase protein specifically binds the transposon ends of the cargo vector, excises the transgene and integrates it in the genome of the target cell (transposition) (Fig. 1a). As transposons insert DNA self-sufficiently, they elicit comparable transgenesis rates to gammaretroviral and lentiviral vectors[9]. Simultaneously, they have favorable attributes regarding immunogenicity, insertion profile, cargo capacity (up to 20-150 kb), complexity and cost for clinical implementation[10–12]. Recent discoveries in targetable and RNA-guided transposition in bacteria also illustrate the potential of these systems in advanced genetic engineering[13–15].

A widely used transposon system is *Sleeping Beauty* (SB), a synthetic element reconstructed from inactive copies in fish genomes[16]. Due to its high insertion efficiency in vertebrates, SB has been applied for functional genomics, cancer gene discovery, transgenesis and gene therapy applications (reviewed elsewhere[17–20]). In the medical setting, SB is being used in several clinical phase I/II trials for *ex vivo* engineering of therapeutic cells, including production of CAR T cells for cancer immunotherapy[18,19,2122]. The first two completed clinical trials have shown that CAR T cells produced with SB elicit comparable efficacy to




Europe PMC Funders Author Manuscripts

cells generated by virus-based gene transfer, with the added benefit of reduced manufacturing complexity and vector procurement cost[11,22].

However, one important shortcoming of current transposon systems, including SB, is that the transposition events that occur in target cells are not well controlled. In particular, the use of transposase-coding DNA causes extended protein expression[22] and can even lead to transposase gene acquisition in the target cells[23]. This lack of control over timing and kinetics of transposase exposure bears the risk for ongoing and uncontrolled transposition, cytotoxicity[24,25] and transgene remobilization[26], thus adding to concerns about the potential for adverse transformation of the therapeutic cell products. In ongoing trials, the engineered T cells are cultured for 2-4 weeks after SB-mediated CAR gene delivery to ensure transposase clearance and avoid the infusion of aberrant or unstable cell products; yet, this extended culturing compromises cell fitness and therapeutic efficacy[11,22]. This constitutes a need for improving control and safety of SB, which are also critical requirements for cell and gene therapy in general.

Previous attempts to control transposase exposure have focused on mRNA-based approaches, which shortened the time of protein expression[24,27] and reduced toxicity in hematopoietic stem and progenitor cells (HSPCs)[28]; however, limitations in scalable mRNA production and stability have hampered large scale clinical implementation[29]. In order to overcome these challenges and achieve precise control of genetic engineering, the direct use of protein is desired[30,31], but efficient delivery of transposases has been hampered by challenges in recombinant protein production[32] and low solubility in physiological conditions. Transfection of mammalian cells with Mos1, Mboumar-9 and Mu transposase-DNA complexes has been reported[33,34], but for SB, protein aggregation, low stability and solubility have remained a major bottleneck for protein production and delivery[32].

We have previously determined the structure of the SB transposase catalytic domain, providing an opportunity to rationally engineer advanced SB variants[35]. Here, we used rational structure-based design to create a high-solubility SB (hsSB) transposase for recombinant production. With this, we developed a transposase protein-based genome engineering strategy and established efficient genetic modification of several mammalian cell lines, primary stem cells, and therapeutic CAR T cells. Our strategy overcomes limitations associated with the use of transposase-coding DNA and allows tight control of transposase activity, which augments safety and opens new possibilities for therapeutic genome engineering.

## Results

### Development of a high-solubility recombinant SB transposase variant

Efficient transfection of mammalian cells generally requires 2-50 μg of protein with ≥95% purity in a small volume (up to 50 mg/ml concentration)[36,37]. We therefore started by manufacturing the SB transposase at a suitable quality and quantity to enable protein delivery. First, we recombinantly expressed the SB100X transposase protein, a hyperactive SB derivate[38], which is functional *in vitro* (Querques, Voigt, Zuliani, Barabas manuscript in preparation). However, SB100X was difficult to purify, provided only modest yields (~2 mg

from 1 liter of *E. coli* culture) and was unstable in the low-salt containing buffers required for electroporation (solubility <7 mg/ml; Supplementary Fig. 1), prohibiting protein delivery and transposition. Other known SB variants were similarly difficult to produce at sufficient quantity and quality[32].

Therefore, we explored a rational design approach based on our recently determined crystal structure of the SB100X catalytic domain[35] and introduced specific amino acid substitutions into SB100X to increase its hydrophilic surface character and solubility (Fig. 1b). First, we selected isoleucine 212 in the target DNA binding surface and substituted it with serine. This mutation was previously shown to elicit hyperactivity in human cells[35] and based on its location in the SB100X structure we predicted that it would also increase protein solubility *in vitro*. We found that solubility indeed improved, but not sufficiently to enable efficient transfection. Therefore, we further analyzed the transposase structure to identify another target for enhancing solubility. We selected C176, a single surface-exposed cysteine that could promote protein aggregation[39,40]. Mutation of C176 into serine, in combination with the I212S substitution, generated a remarkably high-solubility SB (hsSB) variant. As control, we tested mutating other cysteines (C197S, C304S and C316S) that are located inside the core of the catalytic domain, but this compromised protein solubility, consistent with our structural prediction (Fig. 1b). HsSB could be recombinantly expressed and purified at high quality and quantity, yielding up to 6 mg of 99.9 % pure transposase protein from 1 liter of *E. coli* culture (Supplementary Fig. 1a), and it was stable at high concentrations (up to 80 mg/ml) in low-salt containing buffers (Supplementary Fig. 1b).

Next, we probed the stability and transposition activity of hsSB. In thermal melting assays, hsSB showed outstanding thermostability up to 95 °C (Fig. 1c and Supplementary Fig. 1c), which provides a distinct advantage for electroporation procedures. Functional analysis confirmed that the hsSB transposase cleaves and integrates transposon end DNA *in vitro* (Fig. 1d, e) and mediates transposition in HeLa cells, when provided from an expression plasmid, with practically identical efficiency as SB100X (Fig. 1f). We further found that hsSB withstood freezing and thawing and remained equally active in transposition assays following long-term storage (up to 12 months at -80 °C).

Taken together, these results show that recombinant production of the hsSB protein is straightforward and that hsSB retains the high transposition activity of SB100X with enhanced solubility and stability.

## hsSB protein transfection allows for efficient and controlled cell engineering

Next, we tested if the purified hsSB protein can be directly delivered to and promotes transposition in HeLa cells using a neomycin-based reporter system[16]. For this, we transfected cells with a transposon carrying a neomycin resistance ($neo^r$) cassette (conferring resistance to geneticin, G418), which can undergo genomic integration upon successful transposase delivery. Following drug selection, the number of G418-resistant clones provided a measure of the transposition efficiency (Fig. 2a). We first transfected the transposon donor plasmid using a cationic transfection reagent and then delivered the hsSB protein via electroporation (Fig. 2a). We obtained robust and dose-dependent transgene insertion, as demonstrated by the gradually increasing number of G418-resistant HeLa cell

clones with increasing transposase concentrations, with maximum activity achieved using 10-20 μg of protein (per ~$10^5$ cells, Fig. 2b, Supplementary Fig. 2a). Transposition efficiency was comparable to the levels obtained with plasmid-based delivery (>76%, Supplementary Fig. 2b). Sequence analysis of isolated $neo^r$ cells confirmed that the insertions occurred at TA dinucleotides, following the typical integration pattern of SB transposition (Fig. 2c and Supplementary Table 1).

We further investigated how long the transposase was present in target cells by Western blot analysis. We observed substantial loss of the hsSB protein 24 hours after electroporation (down to ~17%), whereas cells transfected with hsSB expression plasmids continued to express high protein levels for over 5 days (>73% of the maximum; Fig. 2d and Supplementary Fig. 3). This short retention time of hsSB implies that direct protein delivery reduces the period of active transposition as compared to DNA or mRNA-based protocols, thus allowing tighter control of transgenesis.

Analysis of the number of transposon insertions into the HeLa cell genome revealed that the integrated transgene copy numbers gradually increased by electroporating increasing amounts of hsSB protein (at constant transposon concentration) into the cells. This direct correlation was not observed when the protein was provided from an expression plasmid (Fig. 2e). Moreover, hsSB protein generated fewer insertions (2-12) per cell than the plasmid-based strategy (8-20) at comparable transgenesis rates (Fig. 2e, 2b and Supplementary Fig. 2b). As uncontrolled transgene integration can lead to disruption or mis-regulation of nearby genes, lowering the copy number is important to reduce the risk of genomic damage in the target cells. Notably, hsSB could be titrated to generate as few as two insertions per cell, going well below the limit recommended by U.S. and European regulators[41].

## Efficient transgenesis in diverse mammalian cells

To better quantify hsSB-mediated transposition, we applied a fluorescent reporter system[38] (Fig. 2a). We transfected HeLa cells with a transposon plasmid containing the Venus gene, followed by hsSB protein delivery by electroporation. Cells that acquired the transposon plasmid were selected by fluorescence activated cell sorting 2 days post-transfection. The transposition efficiency was then quantified three weeks later by flow cytometric analysis of green fluorescent cells that stably expressed the Venus reporter gene as a consequence of genomic insertion by hsSB (Fig. 3a). We detected a clear, dose-dependent increase in the percentage of fluorescent cells, with the maximum efficiency (41%) achieved with 20 μg of hsSB protein (Fig. 3a, upper panel, and Fig. 3b).

To explore the range of applicability of our technique, we delivered hsSB to Chinese hamster ovary (CHO) cells, which are widely used for the production of protein therapeutics. Using the fluorescence assay described above, we observed a dose-dependent increase of engineered CHO cells, up to a maximum of a 73% transgenesis rate with 20 μg of the hsSB protein (Fig. 3a, middle panel, and Fig. 3b).

As engineering of stem cells is instrumental for developmental biology research, establishment of animal models, and regenerative medicine, we tested hsSB protein delivery

in mouse embryonic stem cells (mESCs). We achieved up to 20% transgene insertion in mESCs using 20 μg of the hsSB protein (Fig. 3a, lower panel; Fig. 3b and Supplementary Fig. 4a). The engineered cells maintained pluripotency after treatment, as confirmed by immunofluorescence detection of the Oct4 self-renewal marker (Supplementary Fig. 4b). We further probed hsSB for genome engineering of primary human $CD34^+$ HSPCs. Using an adapted protocol consisting of two sequential electroporations, we achieved successful hsSB-mediated transgenesis (up to 3.33%; calculated as the percentage of stable integrants over all transfected cells, Fig. 3c).

Collectively, these data demonstrate that direct delivery of the hsSB protein leads to efficient genetic engineering of various mammalian cell types including medically relevant primary human cells. We named this protein activity-based cell engineering platform (using recombinant hsSB with any transgene vector) SBprotAct.

## Manufacturing of potent CD19 CAR T cells

To further probe our method in primary human cells and evaluate therapeutic utility, we next tested SBprotAct for CAR T cell manufacturing. We chose a CAR that targets CD19, a well-studied B cell antigen, which is used in most current CAR T therapies[42–44]. Minicircles (MCs) as transposon carriers allow better transfection rates and lower cytotoxicity than plasmids[21,28]. Accordingly, we electroporated CD19 CAR-encoding transposon MCs and the hsSB protein to generate human $CD4^+$ and $CD8^+$ CD19 CAR T cells. We observed a dose-dependent increase in CD19 CAR-expressing T cells (Fig. 4a), achieving overall integration rates up to 20–30%. This rate is similar to the approved CAR-T product tisagenlecleucel[3], which is produced with viral gene transfer and is highly effective in clinical use. Subsequently, before functional analysis, all CAR T cells were enriched to >90% purity by magnetic-activated cell sorting (MACS) using a truncated epidermal growth factor receptor (EGFRt) as a transfection marker[45] (Fig. 4b).

To assess CAR T cell function, we first analyzed the cytolytic activity of CD19 CAR T cells against Raji and JeKo-1 lymphoma cells, and K562 cells stably expressing CD19 as targets *in vitro*. CD19 CAR T cells showed strong lysis of $CD19^+$ targets (> 60% after 4 hours) with low background killing of $CD19^-$ controls (Fig. 4c). CAR T cells secreted high concentrations of the effector cytokines IL-2 and IFN-γ (> 2 pg/ml in 24 hours; Supplementary Fig. 5a) and exhibited productive proliferation in response to antigen binding (Supplementary Fig. 5b). Genomic DNA of EGFRt-enriched CD19 CAR T cells was used to analyze the copy number and integration site profile of the CD19 CAR transposon. The average copy number of the CAR transgene in SBprotAct transfected cells (0.6 μg of CD19 CAR MC and 10 μg of hsSB) was five, and therefore roughly half of that observed in MC-MC (0.6 μg of CAR MC and 0.5 μg of SB100X MC) DNA controls (Fig. 4d). Analysis of the hsSB-mediated CAR insertion sites revealed a close-to-random genomic integration pattern (Fig. 4e and Supplementary Fig. 5c,d). CAR insertions mediated by hsSB occurred 8.5 times more frequently in genomic safe harbors than integrations mediated by an HIV-derived lentiviral vector, causing fewer insertions in exons (3.8 times less), regulatory regions (1.1-1.4 times less) and oncogenes (2.9 times less) (Fig. 4e). This reduced copy number and favorable insertion profile help to minimize the risk of genotoxicity.

Finally, to assess CD19 CAR T cell function *in vivo*, we used a xenograft lymphoma model (NOD scid gamma (NSG) mice / human Raji). CD19 CAR T cells generated with hsSB caused rapid tumor regression within a week after CAR T cell transfer and complete survival of CAR T cell treatment cohorts until the end of the observation period (Fig. 4f and Supplementary Fig. 6). The efficacy of hsSB CAR T cells was similar to and in some mice even exceeded that of the MC-MC controls, showing faster and deeper tumor regression (Fig. 4f). Together, these data indicate that SBprotAct can be integrated into CAR T cell manufacturing and provides therapeutically effective cell products.

## hsSB has intrinsic cell penetrating properties

Engineering of primary cells is challenging because they are often recalcitrant to transfection and generally sensitive to *ex vivo* manipulation. In order to increase the scope of our method for these cells, we sought to make transposase delivery simpler and gentler within SBprotAct.

Remarkably, we observed that the engineered hsSB protein autonomously penetrates HeLa cells and enters the nucleus when simply added to the culture medium (Fig. 5a,b). To test if hsSB remains functional and can mediate transposition when delivered this way, we electroporated HeLa cells with a MC containing the Venus gene and then added hsSB to the culture medium without a further pulse or use of a transfection reagent (Fig. 5a). Longitudinal Western blot analysis showed hsSB uptake within 4 hours, followed by clearance already 24 hours after delivery (Fig. 5c and Supplementary Fig. 7a). Fluorescent cell sorting 3 weeks post transfection revealed up to 12 % Venus-positive cells (Fig. 5d), demonstrating that hsSB mediated effective transgene integration.

Next, we probed a similar procedure for engineering human iPSCs. iPSCs offer great potential for regenerative medicine but are among the most difficult cells to engineer due to their sensitivity to transfection procedures. We first transfected the iPSCs with a Venus-carrying MC using a stem cell specific transfection reagent and then incubated them with hsSB protein-containing medium to allow protein penetration in the cells. hsSB successfully penetrated iPSCs (Fig. 5e and Supplementary Fig. 7b) and flow cytometry of the treated cells after three weeks revealed transgenesis efficiencies up to 3.31% (calculated as the percentage of stable integrants at 3 weeks over all transfected cells, Fig. 5f). This indicates that hsSB's non-invasive cell penetration helps to genetically modify iPSCs.

Finally, we tested if the intrinsic cell penetration property of hsSB can be exploited for T cell engineering (Fig. 6a). We first analyzed hsSB penetration in primary T cells by immunofluorescence imaging, which demonstrated protein uptake in both stimulated and non-stimulated cells within 3 hours (Fig. 6b and Supplementary Fig. 8a). hsSB effectively entered the nucleus also in non-dividing cells, consistent with active transport using its intrinsic nuclear localization signal[46]. To probe transposition, we electroporated T cells with CD19 CAR MC and added hsSB to the cell culture media. This successfully generated human $CD8^+$ CD19 CAR T cells at an overall transgenesis frequency of 5-7 % (Fig. 6c). CAR T cells were then enriched up to 90% purity by MACS[45] and showed potent lysis of $CD19^+$ target cells, as well as high levels of effector cytokine secretion (Fig. 6c,d and Supplementary Fig. 8b). CAR T cells produced with this procedure showed an average

number of four insertions, which is similar to the numbers observed with protein electroporation (5 insertions; Fig. 4d), but lower compared to the CAR MC - SB MC DNA based protocol (6-8 insertions; Fig. 6e).

These data demonstrate that hsSB spontaneously enters diverse cells and retains its ability to mediate transposition, an unanticipated feature that can help promote gentle engineering of hard-to-transfect and transfection-sensitive cells. Future efforts will be directed at optimizing transgenesis rates to enable efficient clinical use, exploring possibilities for additional simplifications and analyzing how cell fitness and therapeutic efficacy are affected by delivery conditions.

## Discussion

Using rational protein design, we have created an SB transposase that allows efficient transgenesis in target cells by direct protein administration. hsSB achieves efficient modification of mammalian cells, as demonstrated herein by the engineering of HeLa, CHO, T cells and stem cells (mESCs, HSPCs and iPSCs), and promotes production of potent CAR T cells for medical use. Uniquely, hsSB penetrates cellular membranes by itself, opening new opportunities for engineering sensitive cell types.

Relying on hsSB, the SBprotAct platform outperforms current technologies in two important respects. As for all integrating vectors, transgene insertion by SB can cause insertional mutagenesis, activation of proto-oncogenes, and genome rearrangements. This risk is proportional to the number of transgene integrations per genome and can be alleviated by controlled reduction of the integration events. We have shown that SBprotAct permits controlled reduction of transposon copy numbers down to levels that are considered acceptable by U.S. and European regulators and ensures low variance between individual cells. This level of control could not be accomplished by conventional transfer vectors, which generate higher numbers of integrations and considerably higher variance (Figs. 2e, 4d, 6e)[47]. Additionally, it can be inferred that the short half-life of hsSB in cells adds to the fidelity of SBprotAct. The SB transposase can excise and integrate transposable elements multiple times, which creates a risk for mutations, as each remobilization leaves a genetic footprint[48]. Examination of CAR T cell products from a phase I clinical trial with a CD19-CAR has shown that the transposase plasmid was still present 21 days after transfection in 5 out of 8 autologous CAR T cell products, likely allowing for multiple transposition cycles[22,49]. In contrast, SBprotAct reduces transposase exposure to less than three days, which diminishes the risk for multiple transposition cycles and guarantees a cell product that is free from residual transposase.

Due to the special properties of the hsSB transposase, SBprotAct has a range of possible applications. The high transgenesis rates observed in CHO cells offer applications in the pharmaceutical industry, since CHO constitute the main manufacturing platform for protein therapeutics (or biologics, e.g. therapeutic antibodies). The ability to customize the number of inserted transgene copies per cell is an additional benefit in this regard, because the optimal gene dosage is directly linked to the yield of the therapeutic product[50]. Further, SBprotAct is well suited for all gene therapy applications that require *ex vivo* manipulation

of cells to correct a genetic defect by gene insertion. We have shown that the platform works reliably in embryonic, hematopoietic and induced pluripotent stem cells as well as in human T cells, all of which are important cell types for gene therapy. We further provided proof-of-concept that hsSB's unique cell penetrating property permits to engineer hard-to-transfect and transfection-sensitive cells. Whereas transgenesis rates may require further optimization for clinical use, hsSB penetration opens new possibilities for future advances. For instance, non-invasive transposase administration can help simplify cell manufacturing protocols and may be beneficial for preserving cell fitness for downstream clinical use. Due to its intrinsic nuclear localization signal[46], cytoplasmic hsSB can actively enter the nucleus, potentially enabling applications in non-dividing cells. Ultimately, hsSB penetration may provide new means to deliver diverse molecular cargos, including chemicals, dyes and macromolecules into cells on demand.

Furthermore, we have shown that SBprotAct allows convenient manufacturing of CAR T cells. Notably, the CD19 CAR T cells generated by SBprotAct exhibited equivalent effector functions to cells produced by SB100X MC, despite having fewer CAR gene copies. This indicates that the lower copy number was sufficient to reach a level of CAR surface expression at which a further increase did not enhance CAR T cell function.

Much effort is currently devoted to reducing *ex vivo* manipulation of CAR T cells[51] to preserve their longevity and fitness, which are important factors for their anti-tumor activity[52,53]. Given the direct onset and fast decline of transposition with a total absence of residual transposase 72 hours after transfection, SBprotAct is compatible with rapid cell manufacturing protocols. As primary T cells are sensitive to transfection procedures, hsSB's cell penetration property can further help to reduce the number of transfection steps, simplify production and preserve cell fitness. Moreover, immunogenicity, which is caused by the presence of foreign protein or DNA, can lead to rejection of engineered cell grafts by the patient. Patient immune responses against CAR T cells have been observed in clinical trials and were associated with reduced anti-tumor activity[54,55]. In addition to the CARs themselves, modified cells may also present vector-encoded immunogenic epitopes[4]. The fish-derived SB transposase is probably immunogenic in humans, but the short half-life of hsSB essentially reduces the immunogenic potential of engineered cells to that of the genetic cargo alone. From a regulatory point-of-view, the low immunogenic potential and the absence of transposase activity after 72 hours promotes an early release of the cell product[19]. Thus, SBprotAct enables the generation of genetically stable and transposase-free CAR T cells within a time window of 3 days, or even below.

Therefore, we anticipate that SBprotAct will be a valuable tool for CAR T and other gene therapies and elicit a significant impact on their cost and accessibility. The hsSB transposase is compatible with large-scale and automated production, and the quality and regulatory requirements are similar to other therapeutic proteins, such as monoclonal antibodies. Thus, the manufacturing and implementation of hsSB into clinical trials should be straightforward. Moreover, hsSB's profound stability and the ability to freeze and thaw the protein without compromising its activity allows shipping to point-of-care locations. As SBprotAct contains no infectious agents, it can be handled at the lowest biosafety level, which permits engineering of therapeutic cells directly at the point-of-care instead of in few specialized

centers. Of note, convenient production and stability of hsSB provide a particular asset over mRNA-based approaches, which also helped to shorten the timeline of genome engineering compared to DNA-based protocols[24,27] but limitations in production and low stability of mRNA have hampered clinical implementation. In particular, *in vitro* transcription or synthesis of suitably modified mRNA is cumbersome under clinical-grade conditions and its sensitivity to heat and degradation during transport, storage and cellular delivery make it challenging to ensure quality control standards for use in large patient cohorts. Direct delivery of hsSB can be leveraged to overcome these shortcomings.

Compared to other genetic engineering tools, transposon-based systems offer several advantages for clinical gene delivery, including high integration efficiency, simple and cost-effective vector manufacturing, large cargo capacity (up to 20-150 kb) and an unbiased genomic insertion profile[10,11,19,56]. However, an obvious constraint of transposon tools is the generally nonspecific nature of their target site selection, which leads to integration at diverse positions throughout the recipient genome. Although transposon vectors did not cause noticeable genotoxicity in clinical trials to date, minimizing the risk is highly desired for broad public use. Genome editing nucleases (such as zinc-finger nucleases, TALENs or CRISPR/Cas9) offer site-specific genome modifications, but their efficiency is limited for inserting large transgenes. A recent report has shown CRISPR-mediated insertion of up to 1.5-kb cargo in human T cells[57], but delivery of larger inserts (such as a ~3 kb CAR gene) could not be achieved. The induction of genomic double strand DNA breaks was also recently shown to create a risk for genomic rearrangements[58] and malignant transformation[59,60] in the target cells. Looking forward, combination of site-specific genome editing and transposon technologies could provide an attractive solution, merging the advantages of both methods to advance safety and efficiency of genetic engineering in research and biomedicine. Recent discoveries of targetable transposition in bacteria[13–15] and advances in the development of transposase hybrids[61–63] clearly demonstrate the feasibility of this approach, although more work is needed to establish utility in eukaryotic cells. By virtue of its high efficiency in mammalian cells, hsSB could be an attractive candidate for future developments in this context. The possibility of direct transposase delivery or penetration is of particular relevance, as nucleoprotein complex administration is emerging as the preferred method in genome editing clinical trials[64].

In conclusion, we expect that SBprotAct, due to its simplicity and versatility, will substantially lower manufacturing costs and help raise the accessibility of gene and cell therapies as compared to the current standard. Implementation of SBprotAct for engineering of further cell types, such as induced pluripotent stem cells, natural killer cells, retinal, epithelial and neuronal cells, could also offer new opportunities in research, molecular and regenerative medicine. Future efforts could help to further reduce cell manipulation by co-delivery or co-penetration of transposase and transgene; and development of a SB transposase with targeted genomic insertion, for example by structure-based design, would provide an important asset for precision genome editing. The platform could also be combined with the orthogonal CRISPR/Cas technology to link site-specific genome editing of programmable nucleases with efficient gene delivery of the SB transposase. As use of other transposases (e.g. the piggyBac transposase) is hampered by similar challenges in

protein production and delivery, applying our protein design approach could facilitate their implementation in manifold genetic engineering applications.

## Methods

### Protein expression and purification

The full-length SB100X transposase (hereafter named SB) was provided from vector pETM-22-SB[35] (containing a 6xHis-thioredoxin fusion tag) for over-expression in *E. coli*. To generate the hsSB transposase, I212S and C176S mutations were introduced via site-directed mutagenesis (protocol adapted from[65]; primers listed in Supplementary Table 2). All transposase variants were expressed in *E. coli* and purified according to the protocol described in[35]. In short, cell pellets were lysed by sonication and proteins were purified from the soluble fraction via Ni-affinity chromatography on a HisTrap column (GE Healthcare), followed by tag cleavage with the PreScission protease, tag removal on a Ni-affinity column, and size exclusion chromatography on a Superdex 200 16/60 column (GE Healthcare). Proteins were assessed for purity by SDS-PAGE (Supplementary Fig. 1a).

### Circular dichroism (CD) spectroscopy

CD spectra of SB100X and hsSB proteins (7.6 μM, in 1× phosphate buffered saline (PBS), pH 7.5 and 200 mM NaCl) were recorded on a Jasco J-715 spectropolarimeter using fused silica cuvettes (1 mm) thermostated with a PTC-348WI Peltier unit. Thermal stability was assessed by monitoring the ellipticity at 206 nm while heating the sample at a rate of 1 °C/min.

### *In vitro* activity analysis

*In vitro* activity analysis of SB100X and hsSB proteins was performed using cleavage and integration assays described in Voigt et al.[35] and in (Querques, Voigt, Zuliani, Barabas manuscript in preparation).

### Cell culture

HeLa Kyoto cells (RRID:CVCL_1922) were maintained in DMEM medium supplemented with 10% (v/v) foetal bovine serum (FBS, Sigma-Aldrich) and 2 mM L-glutamine (Life Technologies). CHO cells (Invitrogen; R758-07) were maintained in F12 K medium containing 10% (v/v) FBS. Mouse ESCs (AB2.2 ES cell line (*Mus musculus*), ATCC, courtesy of P. Neveu, EMBL Heidelberg) were maintained without feeders in "LIF+serum" medium composed of DMEM (high glucose, no glutamine, with sodium bicarbonate, Invitrogen) supplemented with 15% (v/v) ES-qualified foetal calf serum (EmbryoMax, Millipore), 10 ng/ml murine LIF (PepCore, EMBL Heidelberg, Germany), 1× non-essential amino acids, 2 mM L-glutamine, 1 mM sodium pyruvate, 100 U/ml penicillin, 100 μg/ml streptomycin, and 0.1 mM β-mercaptoethanol (all Invitrogen) on culture dishes (Thermo Fisher) coated with 0.1% (v/v) gelatine (Sigma-Aldrich) solution. The medium was changed daily, and cells were passaged every other day with 0.05% Trypsin-EDTA (Invitrogen) at passaging ratios of 1/6–1/8. Induced pluripotent stem cells (iPSC) derived from human peripheral blood mononuclear cells were kindly provided by Dr. Michael Snyder laboratory (Stanford University, USA). iPSCs were cultured in Essential 8 media (Life Technologies) in

recombinant human Vitonectin-coated surfaces (0.5 mg/mL; Life Technologies). The medium was changed every day and cells were passaged using Versene solution (Life Technologies) every 3 to 5 days, before they reached ~85% confluence. $CD34^+$ cells were purified from apheresis products collected after peripheral blood stem cell mobilization induced by means of injections of granulocyte-colony stimulating factor. Apheresis products were obtained from independent, healthy donors after informed consent and with approval of the responsible ethics committee (Goethe University, Permit #329/10). Harvested material was first subjected to red blood cell lysis using ACK Buffer (Invitrogen) and then to positive selection by magnetic cell separation using the MACS human CD34 MicroBead Kit, MACS separator and MACS LS columns (Miltenyi Biotec) according to the manufacturer's instructions. Purity was over 95% by staining with an APC mouse anti-human CD34 antibody (BD Pharmingen). All cell lines and primary cells were maintained at 37 °C in a humidified 5% $CO_2$ atmosphere.

## Transposition assays using SB coding plasmids in HeLa cells

Transposition efficiency of SB100X and hsSB proteins were quantified in HeLa cells using a modified form of the protocol described by Ivics et al.[16]. One day prior to transfection, cells were plated in a 6-well plate at a density of $3 \times 10^5$ cells per well. Transfection was performed with 0.5 μg of pT2/PGK-neo plasmid (containing a PGK promoter-driven neomycin resistance gene flanked by the SB transposon ends on a pBluescript II KS backbone)[66] and 0.5 μg of the pCMV(CAT)T7-SB[38] or the pCMV(CAT)T7-hsSB plasmid (encoding SB100X or hsSB protein, respectively) using JetPei transfection reagent (Polyplus-transfection). After a recovery period of 24 h, transfected cells were trypsinized in a total volume of 1 ml (in one well of a 6-well plate), and 10 μl aliquots of the cell suspension (Fig. 1f) or 16,500 cells (Supplementary Fig. 2b) were re-plated onto 10 cm dishes. The medium was exchanged for selection medium containing 1 mg/ml G-418 (Geneticin, Life Technologies) 24 h later. Cells were selected for two weeks by replacing the selection medium twice a week. After selection, surviving colonies were fixed, stained with methylene blue (Sigma-Aldrich) and counted using ImageJ (Colony Counter plugin) software.

## Transposition assays using hsSB protein electroporation in HeLa cells, CHO cells and mESCs

hsSB protein was prepared for electroporation as follows: purified hsSB was dialysed in two steps against buffer containing 20 mM HEPES, 300 mM NaCl, 1 mM $MgCl_2$ and 0.2 mM TCEP, pH 7.5 and concentrated to 20 mg/ml using Vivaspin 2 concentrators (Sartorius). Protein samples were filtered using 0.22 μm centrifugal filters (Merck Millipore) prior to electroporation into cells. The activity of hsSB was quantified by drug selection or FACS using modified protocols from Ivics et al.[16]. HeLa cells, CHO cells and mESCs were plated one day prior to transfection in a 6-well plate at a density of $3 \times 10^5$ cells per well. Cells were transfected with 0.5 μg of the pT2/PGK-neo plasmid or 1 μg of the pT2/Venus plasmid[38] (containing a neomycin resistance gene and a Venus gene, respectively, flanked by the SB transposon ends) per well. HeLa cells were transfected using JetPei (Polyplus-transfection) transfection reagent (observed transfection efficiency 20-30%). CHO and mESCs were transfected using FuGENE (Promega), resulting in transfection efficiencies of

6-14% for CHO and 5-9% for mESCs. After a recovery period of 4.5 h, cells were trypsinized, washed with PBS and resuspended in R buffer (Neon system) at a cell density of 3–3.5 × $10^7$ cells per ml. For each condition, 10 µl of the cell suspension (containing 3–3.5 × $10^5$ cells) was transfected with 2 µl protein solution containing increasing amounts of hsSB using a Neon transfection system (10 µl kit, Thermo Fisher Scientific; see Supplementary Table 3 for the settings). Following electroporation, cells were plated onto 6-well plates containing medium and incubated at 37 °C in a 5% $CO_2$ atmosphere.

In transposition assays by drug selection, cells were trypsinized after a recovery period of 24 h in a total volume of 1 ml, and 100 µl of the cell (Supplementary Fig. 4a) suspension or 16,500 cells (Fig. 2b) were plated onto 10 cm dishes containing medium. Selection, staining and quantification of the surviving colonies were performed as described for the plasmid-based assays. In transposition assays by FACS, cells were sorted 2 days after electroporation to obtain a pure population of Venus-expressing cells, which had acquired the pT2/Venus plasmid. Twenty-one days after electroporation, the percentage of Venus-positive cells (in whose genome the Venus cassette was stably integrated by hsSB) was quantified by FACS to determine the transposition efficiency.

### Transposition assays using hsSB protein electroporation in HSPCs

Prior to nucleofection, $CD34^+$ HSPCs were pre-stimulated for 2 days in StemSpan serum-free medium (StemCell Technologies) supplemented with 2 mM L-glutamine, 5 U/ml penicillin, 5 µg/ml streptomycin and the following cytokine cocktail: 100 ng/ml hSCF, 100 ng/ml hFlt3-ligand, 100 ng/ml TPO (PeproTech) and 25 µg/ml of human LDL (StemCell Technologies). Cells were maintained under standard liquid cell culture conditions at a cell density of 0.5–1 × $10^6$ cells per ml. 2 × $10^6$ cells per sample were nucleofected with 10 µg of a Venus-encoding SB transposon minicircle[28] (MC-Venus; observed transfection efficiency 67%) and one day later with 10 µg hsSB protein (prepared as described above). Nucleofection was performed using the 4D Nucleofector (Lonza) with P3 buffer (100 µl per condition) and the E0100 program. After nucleofection, cells were resuspended in 2 ml of complete StemSpan serum-free medium (StemCell Technologies) supplemented with 2 mM L-glutamine, 5 U/ml of Penicillin, 5 µg/ml Streptomycin and the following cytokine cocktail: 100 ng/ml hSCF, 100 ng/ml hFlt3-Ligand, 100 ng/ml TPO (PeproTech) and 25 µg/ml of human LDL (StemCell Technologies). Cells were maintained under liquid standard cell culture conditions at a cell density of 0.5-1x$10^6$ cells per ml until up to 3 weeks post-nucleofection. Transposition efficiency was calculated as the percentage of Venus positive cells 21 days post-delivery over the percentage of Venus positive cells 2 days post-delivery.

### Transposition assays using autonomous hsSB transposase penetration in HeLa cells and iPSCs

hsSB protein was prepared as for electroporation. HeLa cells were plated one day prior to transfection in a 6-well plate at a density of 3 × $10^5$ cells per well in DMEM medium supplemented with 10% (v/v) FBS (Sigma-Aldrich) and 2 mM L-glutamine (Life Technologies). Transposon transfection was performed with 0.5 µg of MC-Venus per well using JetPei (Polyplus-transfection) transfection reagent, resulting in transfection efficiencies of 73-85%. After a recovery period of 4.5 h, cells were washed in serum-free

media to remove residual transfection reagents. Cells were then incubated with 1 ml DMEM serum-free medium (SFM) containing 39 or 78 μg of hsSB protein (1-2 μmol/liter hsSB concentration) per well. After 1 h incubation, the media was replaced with fresh serum-containing media without hsSB and the cells were incubated at 37 °C in a 5% $CO_2$ atmosphere. Induced pluripotent stem cells (iPSC) were plated one day prior to transfection in a Vitonectin-coated 6-well plate at a density of $3 \times 10^5$ cells per well in Essential 8 media (Thermo Fisher). Transposon transfection was performed with 0.5 μg MC-Venus per well using Lipofectamine Stem (Thermo Fisher) transfection reagent, yielding 70-88% transfection efficiency. After a recovery period of 4.5 h, cells were incubated with 0.25 μmol/liter hsSB protein (9.75 μg of hsSB in 1 ml of media) in serum-free media (Opti-MEM I Medium; Thermo Fisher). After an incubation time of 1 h, the hsSB-containing media was replaced by the full media (Essential 8) without hsSB and the cells were incubated at 37 °C in a 5% $CO_2$ atmosphere. Three days post-transfection, iPSCs were analyzed by FACS to determine the percentage of cells that acquired the MC-Venus. Fluorescent cells were counted again after 21 days to quantify the cells that have integrated the Venus gene into the genome. Transposition efficiency was calculated as the percentage of Venus positive cells 21 days post-delivery over the percentage of Venus positive cells 3 days post-delivery.

## Flow cytometry

Two days post-transfection, HeLa cells, CHO cells and mESCs were dissociated to a single-cell suspension with 0.05% Trypsin-EDTA (Invitrogen), resuspended in D-PBS containing 2% BSA and 2.5 mM EDTA, strained through a 40 μm cell strainer (BD Biosciences) and sorted on a BD FACSMelody™ or a BD FACSAria Fusion™ flow cytometer (BD Bioscience). Only $Venus^+$ cells were collected for further culturing. After 21 days, the cells were measured for Venus expression and PI- or DAPI-staining was used to determine live/dead cells. Flow cytometry measurements were performed on a BD LSRFortessa, BD Accuri or Beckman Coulter Cyan flow cytometer. When necessary, GFP/PI fluorescent spillover was compensated with the according controls. Detection of $Venus^+$ human $CD34^+$ HSPCs was performed up to three weeks after nucleofection using a BD LSR II flow cytometer (BD Biosciences). iPSC were dissociated to a single-cell suspension with Versene Solution (Life Technologies), resuspended in D-PBS containing 2% BSA and 2.5 mM EDTA, strained through a 40 μm cell strainer (BD Biosciences) and the percentage of fluorescent cells was measured on a BD LSRFortessa™ flow cytometer. DAPI-staining was used to exclude dead cells. The results were analysed using FlowJo software or the FCS Express 4 Flow Cytometry software (De Novo Software).

## CAR T cell generation

Peripheral blood was obtained from healthy donors after written informed consent to participate in a research study. The study was approved by the Institutional Review Board of the University of Würzburg. Mononuclear cells were isolated by density centrifugation with Biocoll separating solution (Biochrom), and $CD4^+$ and $CD8^+$ T cells were sorted by negative isolation with immuno-magnetic associated cell sorting (MACS, Miltenyi). $CD8^+$ and $CD4^+$ primary human T cells were activated by anti-CD3/CD28 bead stimulation (Thermo Fisher). For assays in Fig. 4, two days after activation, CD19 CAR_EGFRt-encoding transposon MC and hsSB protein were transfected consecutively using a 4D-Nucleofector according to the

manufacturer's instructions (Lonza). 2 x $10^6$ T cells were pulsed with 10 μg hsSB in a volume of 20 μl. LV transduction of controls was performed on day one by spinoculation. CAR T cells were propagated in RPMI-1640 with 10% (v/v) human serum, 2 mM L-glutamine, 100 U/ml penicillin-streptomycin and 50 U/ml IL-2. CAR-modified (i.e. EGFRt-positive) T cells were enriched using biotin-conjugated anti-EGFR mAb and anti-biotin beads (Miltenyi) and expanded with irradiated $CD19^+$ feeder cells for 7 days. Functional analysis of CAR T cells was performed as described elsewhere[67–69]. For assays in Fig. 6, two days after activation, 1 x $10^6$ T cells were transfected with 0.5 μg CD19 CAR_EGFRt-encoding transposon MC in a volume of 20 μl using a 4D Nucleofector according to the manufacturer's instructions (Lonza). Directly after electroporation, cells were incubated with 20 μg hsSB in a volume of 500 μl RPMI-1640 serum-free medium for 4 hours. Then, the medium was replaced with fresh serum-containing medium without hsSB. $CD8^+$ T cells were cultured in RPMI-1640 with 10% (v/v) human serum, 2 mM L-glutamine, 100 U/ml penicillin-streptomycin and 50 U/ml IL-2 for additional five days. Then, CD3/CD28 beads were removed and CAR-modified (i.e. EGFRt-positive) T cells (~6.3% of the initial population, as determined by flow cytometry) were enriched using biotin-conjugated anti-EGFR mAb and anti-biotin beads (Miltenyi) and expanded using irradiated $CD19^+$ feeder cells.

### Xenograft mouse model

The Institutional Animal Care and Use Committee of the University of Würzburg approved all mouse experiments. The NSG/Raji model has previously been described[68]. In brief, six- to eight-week-old female NSG mice were obtained from Charles River Laboratories and inoculated with $5 \times 10^5$ firefly luciferase expressing Raji tumour cells by tail vein injection on day 0. On day 7, groups of n = 5 mice received intravenous injections of $5 \times 10^6$ CAR-modified or unmodified T-cells, containing a 1:1 ratio of $CD8^+$ and $CD4^+$ cells. Bioluminescence imaging was performed on an IVIS Lumina (PerkinElmer) following intraperitoneal injection of D luciferin (Biosynth) at 0.3 mg per g body weight, and the data were analysed using LivingImage software (PerkinElmer).

### Western blot analysis

Cells were harvested, washed twice with PBS and sonicated (eppi sonicator; amplitude 65%) in RIPA buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.1% SDS) for lysis. Total protein extracts (20 μg) were separated by gel electrophoresis (NuPage 4–12%; Life Technologies) and transferred to a nitrocellulose membrane using a semi-dry transfer system (Bio-Rad). Membranes were probed with Sleeping Beauty transposase antibody, polyclonal goat IgG (R&D). The internal loading controls were glyceraldehyde 3-phosphate dehydrogenase (GAPDH) or lamin B1/B2 and were detected with an anti-GAPDH (Biozol) or anti-lamin B1/B2 (X223, Invitrogen) antibody, respectively. The intensities of the bands were measured using Image Lab Software (Bio-Rad).

### Sequence analysis of SB insertions in the HeLa cell genome

Genomic DNA from neo-resistant homogenous HeLa clones was purified using the Genelute Mammalian Genomic DNA Kit (Sigma). For identification of insertion sites, we adapted the TRACER protocol[66] so that PCR amplicon libraries generated by the two-step PCR protocol

can be directly sequenced on the Illumina (ILMN) sequencer. Briefly, neomycin transgene insertions were PCR amplified using specific primers for the left end (Left_end) and right end (Right_end) and a random primer (Random_primer) (Supplementary Table 2). All the primers were extended by an ILMN sequencer compatible sequence 'handle' (underlined in the primer sequences) for further sequencing. The first PCR was run applying stringent annealing temperature to minimize the background caused by spurious priming[66]. After the 1st PCR, 1 μl of the PCR mixture was transferred as a template to the 2nd PCR; this step is used to incorporate sequences required for clustering on the ILMN sequencer flowcell and include sequence indices/barcodes enabling multiplexing of individual libraries. Equimolar amounts of purified, individually barcoded libraries were pooled and sequenced bi-directionally on ILMN sequencer MiSeq, each read 250 bases long. The data analysis was carried out using custom scripts. Briefly, we first aligned the data to the hg19 human reference genome using BWA[70]. We then filtered alignments for a mapping quality >=10 and for reads that contain the CTCAG recognition site or its reverse complement (CTGAG) at the expected location (i.e., after the random 7 bp start). In addition, we required that a small stretch of the transposon sequence is present in the respective read sequence. We then clustered all these reads by extending their mapping locations +/-300bp using BEDTools[71] and annotated each cluster with the number of supporting reads and whether the left end or right end of the insertion was observed. All predicted insertion sites were manually inspected and verified using IGV[72] and annotated with genomic features using the UCSC Genome Browser (http://genome.ucsc.edu) (Supplementary Table 1).

## Transposon copy number analysis

Genomic DNA from neo$^r$-positive HeLa cells was purified using the Genelute Mammalian Genomic DNA Kit (Sigma), and the PureLink Genomic DNA Mini Kit (Invitrogen) was used for CD19 CAR T cells. The average number of neomycin-resistance transgene insertions was measured by digital droplet PCR as follows: 200 ng of genomic DNA (gDNA) was digested with 20 units of DpnI restriction enzyme (NEB) in 30 μl final reaction volume overnight to eliminate non-integrated transposon plasmid DNA present in the samples. Next, the DNA was fragmented with CviQI (NEB) for 4 h at 25 °C. The samples were subjected to PCR amplification using Taqman probes and primers specific for the right end of the transposon and for the single-copy human RPP30 gene to measure the genome count. The PCR reactions were performed in 20 μl final volume with 10 ng of fragmented gDNA using the ddPCR Supermix for Probes (No dUTP) master mix (Bio-rad) with 11 pmol primers and 50 pmol of TaqMan probes (for sequences of primers and probes see Supplementary Table 2). The PCR droplets were generated in the QX100 device (Bio-rad). The program was 95 °C 10 min; 40 cycles of 94 °C 30 s, 50 °C 30 s, 60 °C 1 min; 98 °C 10 min. After thermal cycling, the fluorescent droplets were counted in the QX100 Droplet Reader (Bio-rad), and genomic copy numbers were calculated with the Quanta Soft program (Bio-Rad).

## Sequence analysis of SB insertions in the T cell genome

A modified version of the protocol from Monjezi et al.[21] was used. Briefly, 1 μg of the gDNA of T cells containing SB protein-mediated transposon insertions were sheared with a Covaris M220 ultrasonicator to an average fragment size of 600 bp. 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. DNA was purified with AMPure XP magnetic beads (Beckman Coulter) and ligated with 50 pmol of T-linker (see below) with Blunt/TA Ligase Master Mix (NEB) in 20 µl at room temperature for 15 min. T-linkers were created by annealing 100 pmol of the oligonucleotides Linker_TruSeq_T+ and Linker_TruSeq_T-. Non-integrated transposon donor plasmid DNA present in the ligation reactions was digested with DpnI (NEB). After bead-purification, 6 µl of eluate was used for PCR I with 25 pmol of the primers specific for the linker and for the transposon end (Linker and T-Bal-Long) with the following conditions: 98 °C 30 s; 10 cycles of 98 °C 10 s, 72 °C 30 s; 15 cycles of: 98 °C 10 s, ramp to 62 °C (1 °C/s) 30 s, 72 °C 30 s; 72 °C 5 min. All PCR reactions were performed with NEBNext High-Fidelity 2× PCR Master Mix (for PCR primer sequences see Supplementary Table 2). The PCR products were bead-purified, eluted in 20 µl of water and 10 µl was used for PCRII with the primers PE-nest-ind-N and SB-20-bc-ill-N, using the following PCR program: 98 °C 30 s; 15 cycles of 98 °C 10 s, ramp to 64 °C (1 °C/s) 30 s, 72 °C 30 s; 72 °C 5 min. The final PCR products were separated on a 1% agarose gel and the smears of 200–500 bp were gel-isolated and purified. The libraries were sequenced on an Illumina HiSeq2500 instrument using single end, 100 nucleotide, sequencing (Genewiz, USA).

The detailed bioinformatic analysis is described elsewhere[21]. Briefly, the sequences of the amplified transposon sequences and Illumina adaptors were trimmed, and the remaining reads were subjected to quality-trimming using R and various Bioconductor packages[73]. The reads were mapped to the human genome (hg19) with Bowtie[74] allowing unique matching. Only insertions supported by at least 10 independent mapped reads were considered valid. Human genomic features downloaded from the UCSC Genome Browser (http://genome.ucsc.edu) were used to annotate the transposon insertion sites and 10,000 random genomic positions as a reference set. The list of cancer-related genes was downloaded from http://www.bushmanlab.org/links/genelists. The list of ultra-conserved elements was described in[75]. The coordinates of miRNA genes were downloaded from http://www.mirbase.org/ftp.shtml. The coordinates of the category "genomic safe harbours" were derived from the intersection of the coordinates of all safe harbour sub-categories for the hg19 genome assembly.

## mESC immunostaining

To test the maintenance of pluripotency after transfection/electroporation, mESCs were tested for expression of the self-renewal marker Oct4. Cells were fixed for 10 min with 4% PFA in PBS and quenched with 150 mM glycine in PBS for 5 min at RT. Cells were blocked for 30 min in PBS buffer containing 5% BSA and incubated overnight at 4 °C with anti-Oct4 antibody (Oct-3/4, Santa Cruz). Cells were washed and incubated with an anti-mouse secondary antibody (F(ab')2-goat anti-mouse IgG (H+L) secondary antibody, Pacific blue conjugate from Invitrogen, for 2 h at RT. Cells were analysed on a LSRFortessa flow cytometer (BD BioScience).

### Immunofluorescence staining

Apheresis products were obtained from healthy donors after written informed consent to participate in a research study. The study was approved by the Institutional Review Board of the University of Würzburg. Mononuclear cells were isolated by density centrifugation with Biocoll separating solution (Biochrom), and $CD4^+$ and $CD8^+$ T cells were sorted by magnetic associated cell sorting (MACS) using untouched kits (Miltenyi). $CD8^+$ and $CD4^+$ primary human T cells were either left unstimulated or activated by anti-CD3/CD28 bead stimulation (Thermo Fisher). Nunc™ Lab Tek™ II 8-well Chamber Slides™ (Thermo Fisher) were coated with Poly-D-Lysine (33 μg/mL; Merck Millipore). T cells were seeded onto Poly-D-Lysine coated 8-well chamber slides ($1 \times 10^5$ cells per well in 200 μl RPMI-1640 supplemented with 10% (v/v) human serum, 2 mM L-glutamine, 100 U/ml penicillin-streptomycin) one day prior to immunofluorescence staining. HeLa cells were seeded onto un-coated 8-well chamber slides ($2 \times 10^4$ cells per well in 500 μl DMEM supplemented with 10% (v/v) human serum and 2 mM L-glutamine).

On the next day, cells were incubated with 2 μg hsSB in a volume of 200 μl/well serum-free RPMI-1640 (T cells) or 1.86 μg hsSB in a volume of 500 μl/well serum-free DMEM (HeLa cells) for 3 hours. Then, media was aspirated and cells were fixed with 300 μl/well Fixation/Permeabilization solution (BD Biosciences) and permeabilized with 300 μl/well 0.1% (v/v) Triton™ X-100 (Sigma-Aldrich). Slides were washed 3 times with 300 μl/well PBS and incubated with 300 μl/well 5% (v/v) BSA (Sigma-Aldrich) for 30 min at room temperature to block unspecific binding. Slides were washed 2 times with PBS for 5 minutes and incubated with anti-SB antibody (100 μl/well; diluted 1:100 in 1% (v/v) BSA; goat polyclonal IgG; R&D systems) for 1 hour at room temperature. Slides were washed 4 times with 200 μl/well 1% (v/v) BSA and incubated for 45 minutes in the dark with anti-goat Alexa Fluor® 488 antibody (100 μl/well; diluted 1:1000 in 1% (v/v) BSA; Thermo Fisher). Slides were washed 4 times with 500 μl/well PBS and SlowFade™ Gold Antifade Mountant with DAPI (Thermo Fisher) was applied. Slides were covered with a coverslip and stored at +4°C. Cells were imaged with a Zeiss LSM 780 confocal microscope (using a 63x oil submersion objective) in the ALMF core facility at EMBL Heidelberg. For imaging, the middle part of the nucleus was in the focus to detect nuclear localization of hsSB.

### Statistical methods

All analyses were performed with Excel software. Data are presented as means with error bars representing s.d. or s.e.m. as indicated in figure legends.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

The authors thank F. Dyda, K.R. Patil, M. Beck and members of the Barabas lab for helpful discussions and the Nature Editing Service for editing assistance. We also thank the Flow Cytometry Core Facility, the Genomics Core Facility, the Advanced Light Microscopy Facility, and the Protein Expression and Purification Core Facility at EMBL Heidelberg for materials and support. We thank Z. Izsvak (Max Delbrück Center, Berlin, Germany) for providing the CMV(CAT)T7-SB100X plasmid and F. Spitz (EMBL, Heidelberg, Germany) for the pT2/PGK-neo

construct; A. Aulehla for providing access to the Neon transfection system and reagents; S. Henkel for providing reagents and assistance for mESCs culture and immunostaining; H. Bönig (Blutspendedienst des Deutschen Roten Kreuz, Frankfurt, Germany) for providing human HSPCs; K.-M. Noh and N. Diaz for providing materials and advice with iPSC culturing; V. Rybin for assistance and advice with the circular dichroism spectroscopy experiments; and V. Benes for assistance and support with sequence analysis. This work was supported by the European Molecular Biology Laboratory (EMBL), the Paul Ehrlich Institute (PEI), the EMBL International PhD Programme (fellowship to I.Q.), the Deutsche Forschungsgemeinschaft (Projectnumber 324392634, TRR 221 to M.H. and H.E.), and the German Cancer Aid (Deutsche Krebshilfe e.V., Max Eder Program Award 70110313 to M.H.). M.H. was supported by the Young Scholar Program of the Bavarian Academy of Sciences (Junges Kolleg, Bayerische Akademie der Wissenschaften) and the $m^4$ Award in Personalized Medicine (Free State of Bavaria, BIO-1601-0002). Z.I. was supported by the Center for Cell and Gene Therapy of the LOEWE (Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz) program in Hessen, Germany, and by a grant from the Deutsche Forschungsgemeinschaft (IV 21/11-1).

## References

1. Turtle CJ, et al. Durable Molecular Remissions in Chronic Lymphocytic Leukemia Treated With CD19-Specific Chimeric Antigen Receptor-Modified T Cells After Failure of Ibrutinib. J Clin Oncol. 2017; 35:3010–3020. [PubMed: 28715249]
2. Klebanoff CA, Rosenberg SA, Restifo NP. Prospects for gene-engineered T cell immunotherapy for solid cancers. Nat Med. 2016; 22:26–36. [PubMed: 26735408]
3. Maude SL, et al. Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. N Engl J Med. 2018; 378:439–448. [PubMed: 29385370]
4. Lamers CH, et al. Immune responses to transgene and retroviral vector in patients treated with ex vivo-engineered T cells. Blood. 2011; 117:72–82. [PubMed: 20889925]
5. Bushman F, et al. Genome-wide analysis of retroviral DNA integration. Nat Rev Microbiol. 2005; 3:848–858. [PubMed: 16175173]
6. Clauss J, et al. Efficient Non-Viral T-Cell Engineering by Sleeping Beauty Minicircles Diminishing DNA Toxicity and miRNAs Silencing the Endogenous T-Cell Receptors. Human gene therapy. 2018; 29:569–584. [PubMed: 29562762]
7. Cornu TI, Mussolino C, Cathomen T. Refining strategies to translate genome editing to the clinic. Nat Med. 2017; 23:415–423. [PubMed: 28388605]
8. Li K, Wang G, Andersen T, Zhou P, Pu WT. Optimization of genome engineering approaches with the CRISPR/Cas9 system. PloS one. 2014; 9:e105779. [PubMed: 25166277]
9. Peng PD, et al. Efficient nonviral Sleeping Beauty transposon-based TCR gene transfer to peripheral blood lymphocytes confers antigen-specific antitumor reactivity. Gene therapy. 2009; 16:1042–1049. [PubMed: 19494842]
10. Ivics Z, et al. Transposon-mediated genome manipulation in vertebrates. Nature methods. 2009; 6:415–422. [PubMed: 19478801]
11. Tipanee J, Chai YC, VandenDriessche T, Chuah MK. Preclinical and clinical advances in transposon-based gene therapy. Biosci Rep. 2017
12. Vargas JE, et al. Retroviral vectors and transposons for stable gene therapy: advances, current challenges and perspectives. J Transl Med. 2016; 14:288. [PubMed: 27729044]
13. Klompe SE, Vo PLH, Halpin-Healy TS, Sternberg SH. Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. Nature. 2019; 571:219–225. [PubMed: 31189177]
14. Morero NR, et al. Targeting IS608 transposon integration to highly specific sequences by structure-based transposon engineering. Nucleic acids research. 2018; 46:4152–4163. [PubMed: 29635476]
15. Strecker J, et al. RNA-guided DNA insertion with CRISPR-associated transposases. Science. 2019; 365:48–53. [PubMed: 31171706]
16. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell. 1997; 91:501–510. [PubMed: 9390559]
17. Dupuy AJ, Jenkins NA, Copeland NG. Sleeping beauty: a novel cancer gene discovery tool. Human molecular genetics. 2006; 15(Spec No 1):R75–79. [PubMed: 16651372]

18. Hodge R, Narayanavari SA, Izsvak Z, Ivics Z. Wide Awake and Ready to Move: 20 Years of Non-Viral Therapeutic Genome Engineering with the Sleeping Beauty Transposon System. Human gene therapy. 2017; 28:842–855. [PubMed: 28870121]
19. Hudecek M, et al. Going non-viral: the Sleeping Beauty transposon system breaks on through to the clinical side. Crit Rev Biochem Mol Biol. 2017; 52:355–380. [PubMed: 28402189]
20. Kawakami K, Largaespad DA, Ivics Z. Transposons As Tools for Functional Genomics in Vertebrate Models. Trends in genetics : TIG. 2017; 33:784–801. [PubMed: 28888423]
21. Monjezi R, et al. Enhanced CAR T-cell engineering using non-viral Sleeping Beauty transposition from minicircle vectors. Leukemia. 2017; 31:186–194. [PubMed: 27491640]
22. Kebriaei P, et al. Phase I trials using Sleeping Beauty to generate CD19-specific CAR T cells. The Journal of clinical investigation. 2016; 126:3363–3376. [PubMed: 27482888]
23. Wang Z, et al. Detection of integration of plasmid DNA into host genomic DNA following intramuscular injection and electroporation. Gene therapy. 2004; 11:711–721. [PubMed: 14724672]
24. Galla M, et al. Avoiding cytotoxicity of transposases by dose-controlled mRNA delivery. Nucleic acids research. 2011; 39:7147–7160. [PubMed: 21609958]
25. Huang X, et al. Unexpectedly high copy number of random integration but low frequency of persistent expression of the Sleeping Beauty transposase after trans delivery in primary human T cells. Human gene therapy. 2010; 21:1577–1590. [PubMed: 20528476]
26. Liang Q, Kong J, Stalker J, Bradley A. Chromosomal mobilization and reintegration of Sleeping Beauty and PiggyBac transposons. Genesis. 2009; 47:404–408. [PubMed: 19391106]
27. Wilber A, et al. Messenger RNA as a source of transposase for sleeping beauty transposon-mediated correction of hereditary tyrosinemia type I. Molecular therapy : the journal of the American Society of Gene Therapy. 2007; 15:1280–1287. [PubMed: 17440442]
28. Holstein M, et al. Efficient Non-viral Gene Delivery into Human Hematopoietic Stem Cells by Minicircle Sleeping Beauty Transposon Vectors. Molecular therapy : the journal of the American Society of Gene Therapy. 2018; 26:1137–1153. [PubMed: 29503198]
29. Hackett, PB, Starr, TK, Cooper, LJN. Translating Gene Therapy to the Clinic. Laurence, J, Franklin, M, editors. 2015. 65–83.
30. Hendel A, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. Nat Biotechnol. 2015; 33:985–989. [PubMed: 26121415]
31. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF 3rd. Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nature methods. 2012; 9:805–807. [PubMed: 22751204]
32. Zayed H, Izsvak Z, Khare D, Heinemann U, Ivics Z. The DNA-bending protein HMGB1 is a cellular cofactor of Sleeping Beauty transposition. Nucleic acids research. 2003; 31:2313–2322. [PubMed: 12711676]
33. Paatero AO, et al. Bacteriophage Mu integration in yeast and mammalian genomes. Nucleic acids research. 2008; 36:e148. [PubMed: 18953026]
34. Trubitsyna M, et al. Use of mariner transposases for one-step delivery and integration of DNA in prokaryotes and eukaryotes by transfection. Nucleic acids research. 2017; 45:e89. [PubMed: 28204586]
35. Voigt F, et al. Sleeping Beauty transposase structure allows rational design of hyperactive variants for genetic engineering. Nat Commun. 2016; 7
36. Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res. 2014; 24:1012–1019. [PubMed: 24696461]
37. Zuris JA, et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nat Biotechnol. 2015; 33:73–80. [PubMed: 25357182]
38. Mates L, et al. Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. Nature genetics. 2009; 41:753–761. [PubMed: 19412179]
39. Avramopoulou V, Mamalaki A, Tzartos SJ. Soluble, oligomeric, and ligand-binding extracellular domain of the human alpha7 acetylcholine receptor expressed in yeast: replacement of the

Europe PMC Funders Author Manuscripts

hydrophobic cysteine loop by the hydrophilic loop of the ACh-binding protein enhances protein solubility. The Journal of biological chemistry. 2004; 279:38287–38293. [PubMed: 15226316]

40. Slusarczyk H, Felber S, Kula MR, Pohl M. Stabilization of NAD-dependent formate dehydrogenase from Candida boidinii by site-directed mutagenesis of cysteine residues. Eur J Biochem. 2000; 267:1280–1289. [PubMed: 10691964]
41. Zhao Y, Stepto H, Schneider CK. Development of the First World Health Organization Lentiviral Vector Standard: Toward the Production Control and Standardization of Lentivirus-Based Gene Therapy Products. Hum Gene Ther Methods. 2017; 28:205–214. [PubMed: 28747142]
42. Davila ML, et al. Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukemia. Sci Transl Med. 2014; 6
43. Maude SL, et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. N Engl J Med. 2014; 371:1507–1517. [PubMed: 25317870]
44. Lee DW, et al. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: a phase 1 dose-escalation trial. Lancet. 2015; 385:517–528. [PubMed: 25319501]
45. Wang X, et al. A transgene-encoded cell surface polypeptide for selection, in vivo tracking, and ablation of engineered cells. Blood. 2011; 118:1255–1263. [PubMed: 21653320]
46. Ivics Z, Izsvak Z, Minter A, Hackett PB. Identification of functional domains and evolution of Tc1-like transposable elements. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93:5008–5013. [PubMed: 8643520]
47. Charrier S, et al. Quantification of lentiviral vector copy numbers in individual hematopoietic colony-forming cells shows vector dose-dependent effects on the frequency and level of transduction. Gene therapy. 2011; 18:479–487. [PubMed: 21160533]
48. Luo G, Ivics Z, Izsvak Z, Bradley A. Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95:10769–10773. [PubMed: 9724779]
49. Hackett PB, Largaespada DA, Switzer KC, Cooper LJ. Evaluating risks of insertional mutagenesis by DNA transposons in gene therapy. Transl Res. 2013; 161:265–283. [PubMed: 23313630]
50. Chusainow J, et al. A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? Biotechnol Bioeng. 2009; 102:1182–1196. [PubMed: 18979540]
51. Lu TL, et al. A Rapid Cell Expansion Process for Production of Engineered Autologous CAR-T Cell Therapies. Hum Gene Ther Methods. 2016; 27:209–218. [PubMed: 27897048]
52. Gattinoni L, et al. A human memory T cell subset with stem cell-like properties. Nat Med. 2011; 17:1290–1297. [PubMed: 21926977]
53. Klebanoff CA, Gattinoni L, Restifo NP. Sorting through subsets: which T-cell populations mediate highly effective adoptive immunotherapy? J Immunother. 2012; 35:651–660. [PubMed: 23090074]
54. Turtle CJ, et al. CD19 CAR-T cells of defined CD4+:CD8+ composition in adult B cell ALL patients. The Journal of clinical investigation. 2016; 126:2123–2138. [PubMed: 27111235]
55. Maus MV, et al. T cells expressing chimeric antigen receptors can cause anaphylaxis in humans. Cancer Immunol Res. 2013; 1:26–31.
56. D'Aloia MM, Zizzari IG, Sacchetti B, Pierelli L, Alimandi M. CAR-T cells: the long and winding road to solid tumors. Cell Death Dis. 2018; 9:282. [PubMed: 29449531]
57. Roth TL, et al. Reprogramming human T cell function and specificity with non-viral genome targeting. Nature. 2018; 559:405–409. [PubMed: 29995861]
58. Kosicki M, Tomberg K, Bradley A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. Nat Biotechnol. 2018; 36:765–771. [PubMed: 30010673]
59. Haapaniemi E, Botla S, Persson J, Schmierer B, Taipale J. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. Nat Med. 2018; 24:927–930. [PubMed: 29892067]
60. Ihry RJ, et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. Nat Med. 2018; 24:939–946. [PubMed: 29892062]
61. Bhatt S, Chalmers R. Targeted DNA transposition in vitro using a dCas9-transposase fusion protein. Nucleic acids research. 2019

62. Hew BE, Sato R, Mauro D, Stoytchev I, Owens JB. RNA-guided piggyBac transposition in human cells. Synth Biol (Oxf). 2019; 4
63. Voigt K, et al. Retargeting sleeping beauty transposon insertions by engineered zinc finger DNA-binding domains. Molecular therapy : the journal of the American Society of Gene Therapy. 2012; 20:1852–1862. [PubMed: 22776959]
64. Wang M, Glass ZA, Xu Q. Non-viral delivery of genome-editing nucleases for gene therapy. Gene therapy. 2017; 24:144–150. [PubMed: 27797355]
65. Makarova O, Kamberov E, Margolis B. Generation of deletion and point mutations with one primer in a single cloning step. BioTechniques. 2000; 29:970–972. [PubMed: 11084856]
66. Ruf S, et al. Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor. Nature genetics. 2011; 43:379–386. [PubMed: 21423180]
67. Hudecek M, et al. The nonsignaling extracellular spacer domain of chimeric antigen receptors is decisive for in vivo antitumor activity. Cancer Immunol Res. 2015; 3:125–135. [PubMed: 25212991]
68. Hudecek M, et al. The B-cell tumor-associated antigen ROR1 can be targeted with T cells modified to express a ROR1-specific chimeric antigen receptor. Blood. 2010; 116:4532–4541. [PubMed: 20702778]
69. Hudecek M, et al. Receptor affinity and extracellular domain modifications affect tumor recognition by ROR1-specific chimeric antigen receptor T cells. Clin Cancer Res. 2013; 19:3153–3164. [PubMed: 23620405]
70. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25:1754–1760. [PubMed: 19451168]
71. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. [PubMed: 20110278]
72. Robinson JT, et al. Integrative genomics viewer. Nat Biotechnol. 2011; 29:24–26. [PubMed: 21221095]
73. Huber W, et al. Orchestrating high-throughput genomic analysis with Bioconductor. Nature methods. 2015; 12:115–121. [PubMed: 25633503]
74. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology. 2009; 10:R25. [PubMed: 19261174]
75. Chiang CW, et al. Ultraconserved elements in the human genome: association and transmission analyses of highly constrained single-nucleotide polymorphisms. Genetics. 2012; 192:253–266. [PubMed: 22714408]


Europe PMC Funders Author Manuscripts

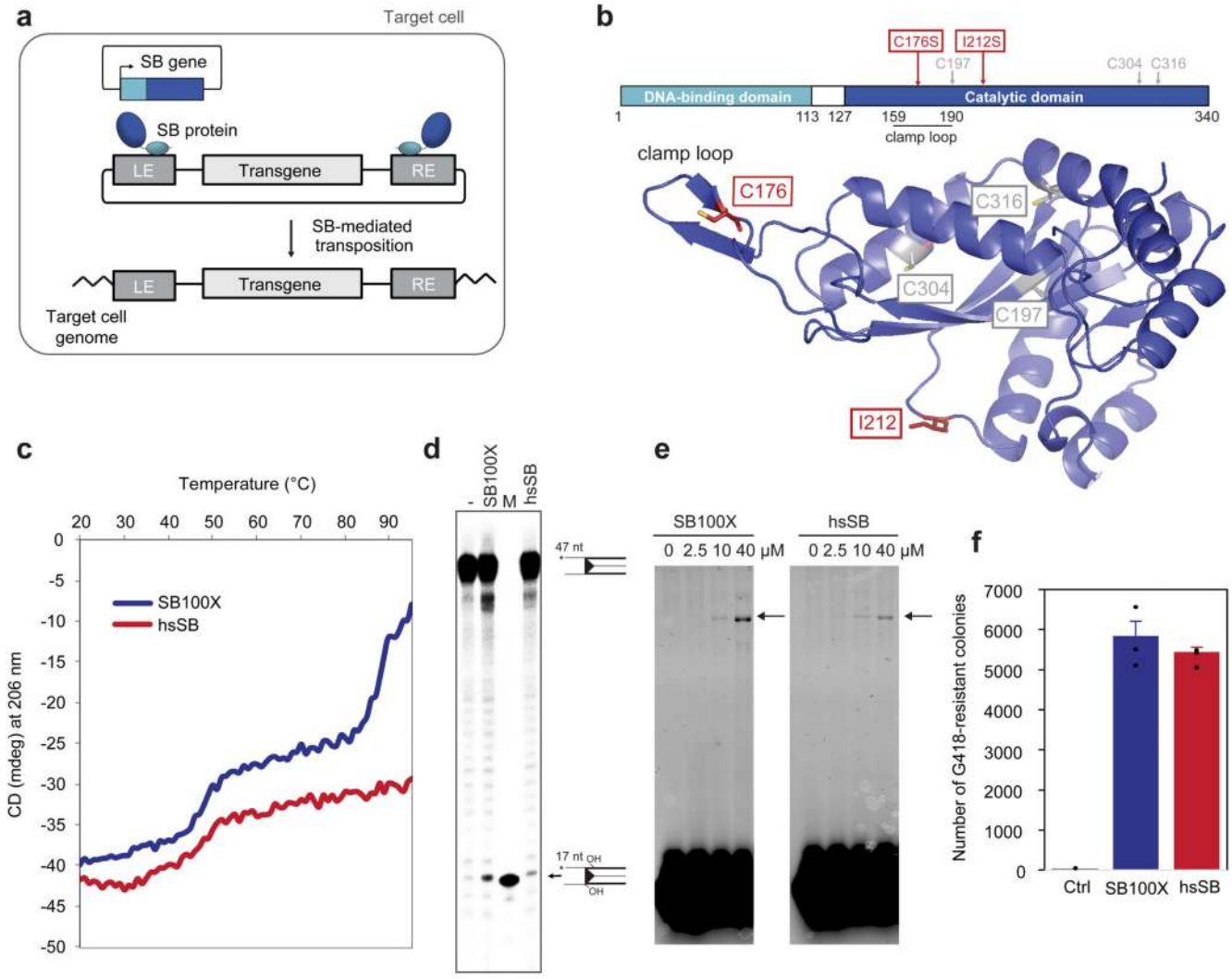


**Figure 1. Design and characterization of the hsSB protein variant.**

**a**, Schematic representation of genome engineering by SB transposase. LE and RE mark the left and right transposon end sequences, respectively. Cargo gene transfer in the target genome is executed by the transposase, expressed from a plasmid vector (bent arrow) in the target cells. **b**, Domain composition of the SB protein and crystal structure of the SB100X transposase catalytic domain (PDB 5CR4)[35] with the hsSB mutations marked (red). Structurally buried cysteines (grey) were mutated as control. **c**, Thermal melting curves of the SB100X and hsSB proteins followed by Circular Dichroism (CD) spectroscopy. Increased CD signal at 206 nm reflects unfolding of α-helices. Experiment was repeated independently two times with similar results. **d**, *In vitro* cleavage assays with SB100X or hsSB and radiolabeled transposon end DNA substrates. $^{32}$P-labeled substrates (asterisk) are separated from reaction products on denaturing PAGE. Arrow indicates specific cleavage product (17 nt). The marker (M) is 17 nt; (-) no protein. **e**, *In vitro* integration assays detecting insertion of transposon end DNA into a target plasmid. Expected integration products are marked (arrow) on a native agarose gel. **d**,**e**, Experiments were repeated




Europe PMC Funders Author Manuscripts

independently three times with similar results. **f**, Transposition assay demonstrating the *in cellulo* activity of the SB100X and hsSB proteins delivered on an expression plasmid in HeLa cells. Mean values; error bars indicate the s.e.m. ($n = 3$ independent experiments). Difference between SB100X and hsSB samples is not significant as estimated by an unpaired two-sided *t*-test (*P* value = 0.446). Colony counts were corrected for the dilution used. The control (Ctrl) cells did not contain SB expression plasmid.



Europe PMC Funders Author Manuscripts

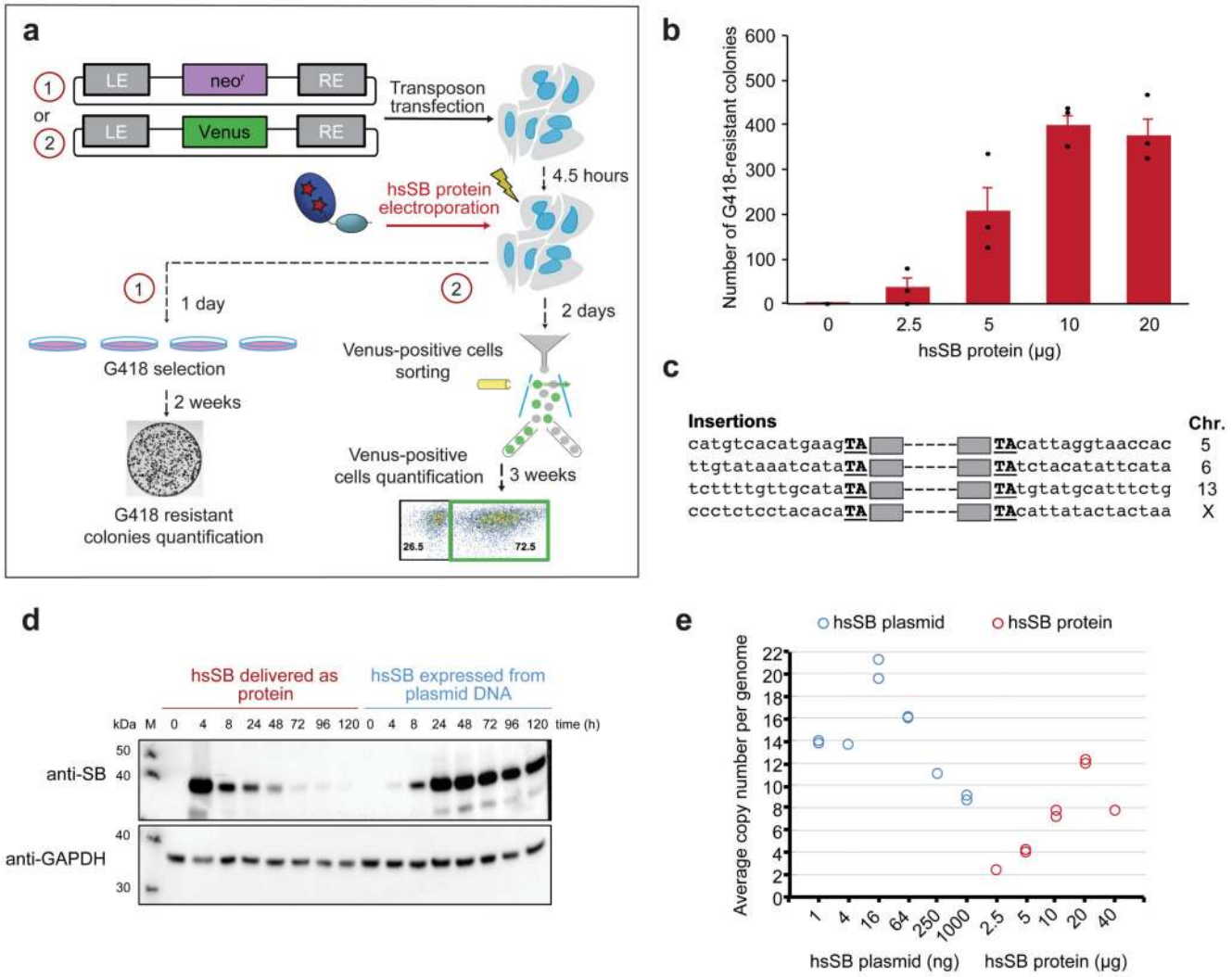


**Figure 2. Efficient cell engineering by direct hsSB delivery.**

**a**, Schematic representation of the hsSB-based engineering procedure, using two alternative reporter systems (1, with neomycin-resistance cassette [neo$^r$]; 2, with Venus gene). **b**, Number of colonies with neo$^r$ insertions following direct transfection of different amounts of hsSB transposase. Mean values; error bars represent s.e.m. ($n = 3$ independent experiments). **c**, Representative sites of hsSB-mediated insertions as derived by sequence analysis of the neo$^r$ locus from isolated neo$^r$ positive HeLa cell clones. Grey boxes represent the transposon ends; Chr., chromosome. **d**, Western blot analysis showing differential retention of hsSB delivered into HeLa cells as protein (10 µg) or expressed from plasmid DNA (500 ng). Samples were blotted with either anti-SB antibody or anti-GAPDH (glyceraldehyde 3-phosphate dehydrogenase) as internal loading control. Experiment was repeated independently two times with similar results. **e**, The average number of neo$^r$ insertions per HeLa cell genome, mediated by hsSB delivered as a protein or expressed from plasmid, as measured by two independent digital droplet PCR assays.




Europe PMC Funders Author Manuscripts

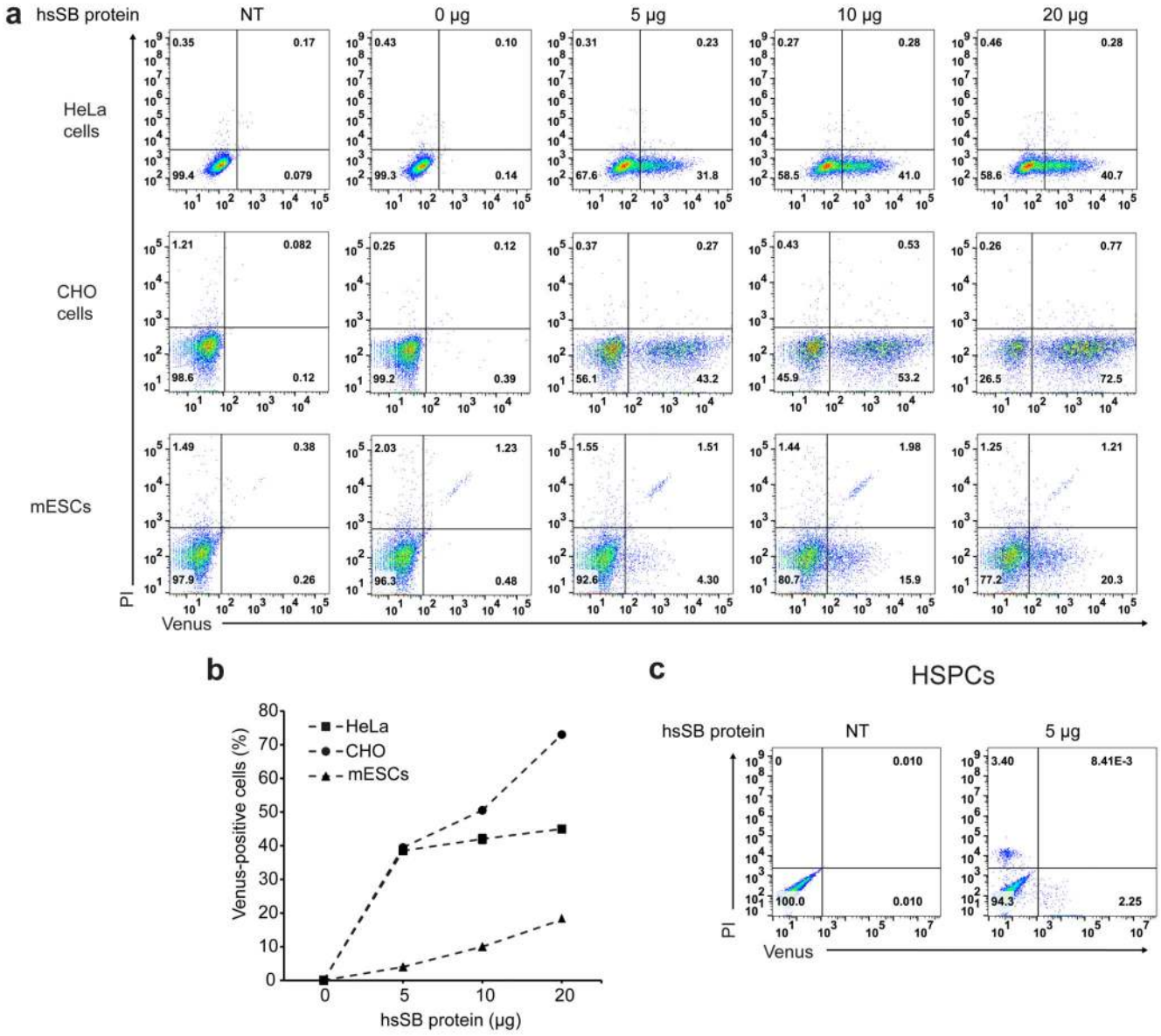


**Figure 3. Direct hsSB delivery allows for efficient transgenesis in diverse mammalian cells and stem cells.**

**a**, Representative flow cytometric analysis of HeLa cells (top), Chinese hamster ovary (CHO) cells (middle) and mouse embryonic stem cells (mESCs; bottom) transfected with Venus-carrying transposon plasmids and electroporated with hsSB transposase. Cells stably expressing an integrated Venus gene were identified 3 weeks post-transfection. The electroporated hsSB amounts are indicated above (NT: non-transfected). X-axis, green fluorescence from Venus; y-axis, propidium iodide (PI) staining intensity to exclude dead cells. **b**, Transposition efficiency of hsSB protein in different cell lines, quantified by flow cytometry. Mean values ($n = 2$ independent experiments). **c**, Representative flow cytometry plots from human hematopoietic stem and progenitor cells (HSPCs) engineered with Venus-




Europe PMC Funders Author Manuscripts

carrying transposon plasmids and electroporated hsSB transposase, analyzed 3 weeks post-delivery. Experiment was repeated independently two times with similar results.



Europe PMC Funders Author Manuscripts

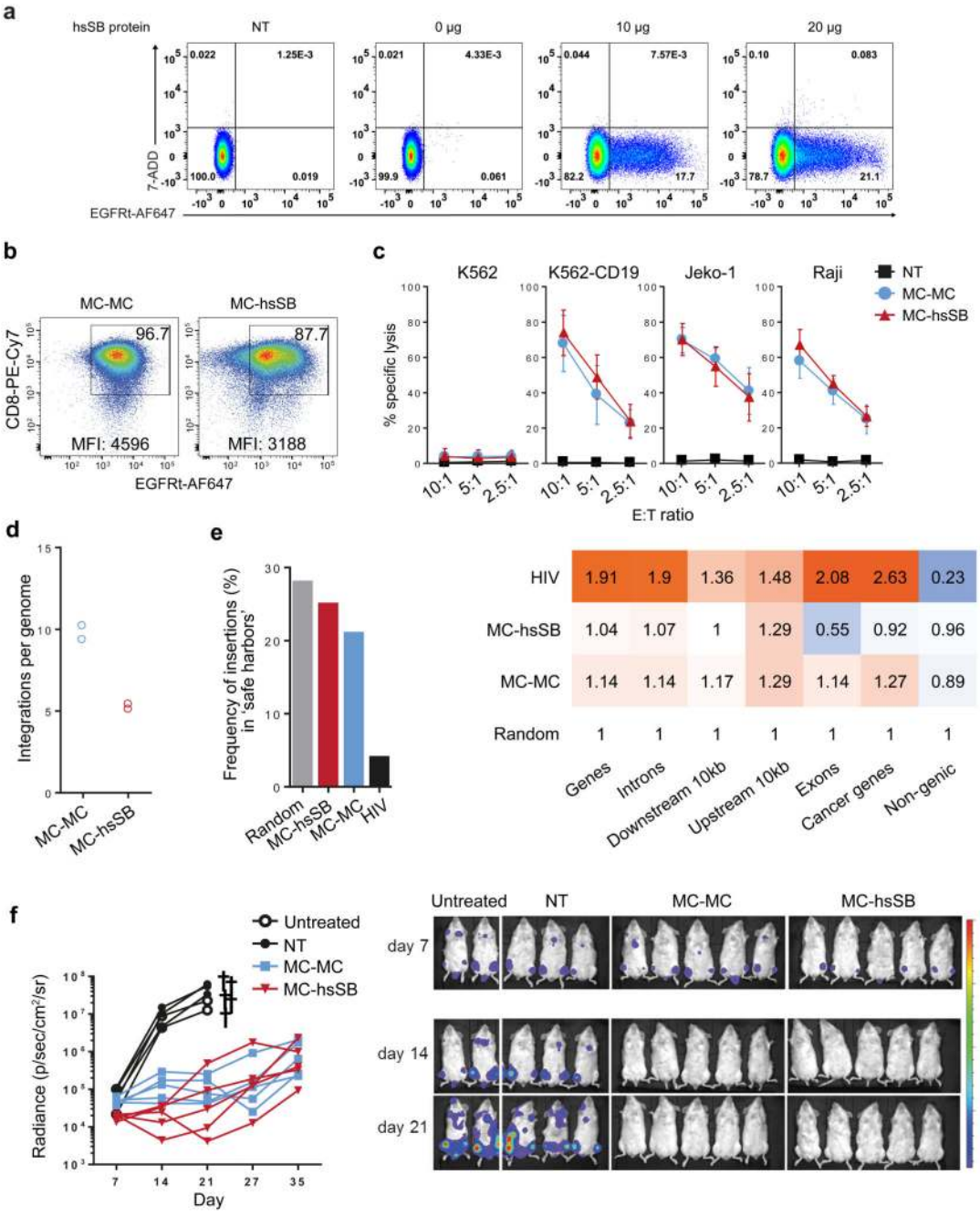


**Figure 4. CD19 CAR T cell generation by SBprotAct and functional analysis of the cell product.**

**a**, $CD8^+$ T cells were isolated from the peripheral blood of healthy donors and transfected with CD19 CAR minicircle and hsSB protein. Representative FACS plots are shown. The electroporated hsSB amounts are indicated above each plot (NT: non-transfected). X-axis, fluorescence from EGFRt specific antibody (EGFRt-AF647); y-axis, 7-Aminoactinomycin D (7-ADD) staining intensity to exclude dead cells. **b**, Phenotype of enriched CAR-positive cells (EGFRt serves as marker) generated with CAR MC and hsSB (MC-hsSB) or CAR MC and SB100X MC (MC-MC) following expansion with irradiated $CD19^+$ feeder cells prior to




Europe PMC Funders Author Manuscripts

functional assays. Representative FACS plots with fluorescence from CD8 and EGFRt specific antibodies (y-axis, CD8-PE-Cy7; x-axis, EGFRt-AF647). MFI, mean fluorescence intensity. **a**,**b**, Experiments were repeated independently three times using cells from three different donors showing similar results. **c**, Cytolytic activity of CD19 CAR T cells generated by SBprotAct or MC-MC. Cytolysis was calculated from the luminescence signals of ffLuc-expressing target cells in a 4 h co-culture assay in the presence of excess luciferin. Data were obtained from $n = 3$ independent experiments with $n = 3$ different T cell donors. Mean values; error bars represent s.d.; NT, non-transfected. E:T ratio, effector to target ratio. **d**, The average number of CAR transgene insertions per CAR T cell genome as measured by two independent digital droplet PCR assays. **e**, Analysis of hsSB protein, SB 100X MC and HIV mediated insertions in human T cells. The left panel shows the frequency of insertions (%) in genomic safe harbors. The occurrences of insertions in gene-related segments of the human genome are presented as fold-changes compared to the random expected frequency (set to 1) on the right. Red and blue shades signify over- and under-representation, respectively. **f**, NSG mice received intravenous injections of $5 \times 10^5$ Raji-ffLuc lymphoma cells (day 0). Seven days later, mice were treated with $5 \times 10^6$ CD19 CAR T cells or non-transfected control T cells (NT) and imaged at the indicated time points. A summary of luminescence signals (left) and images of individual mice (right) are displayed.

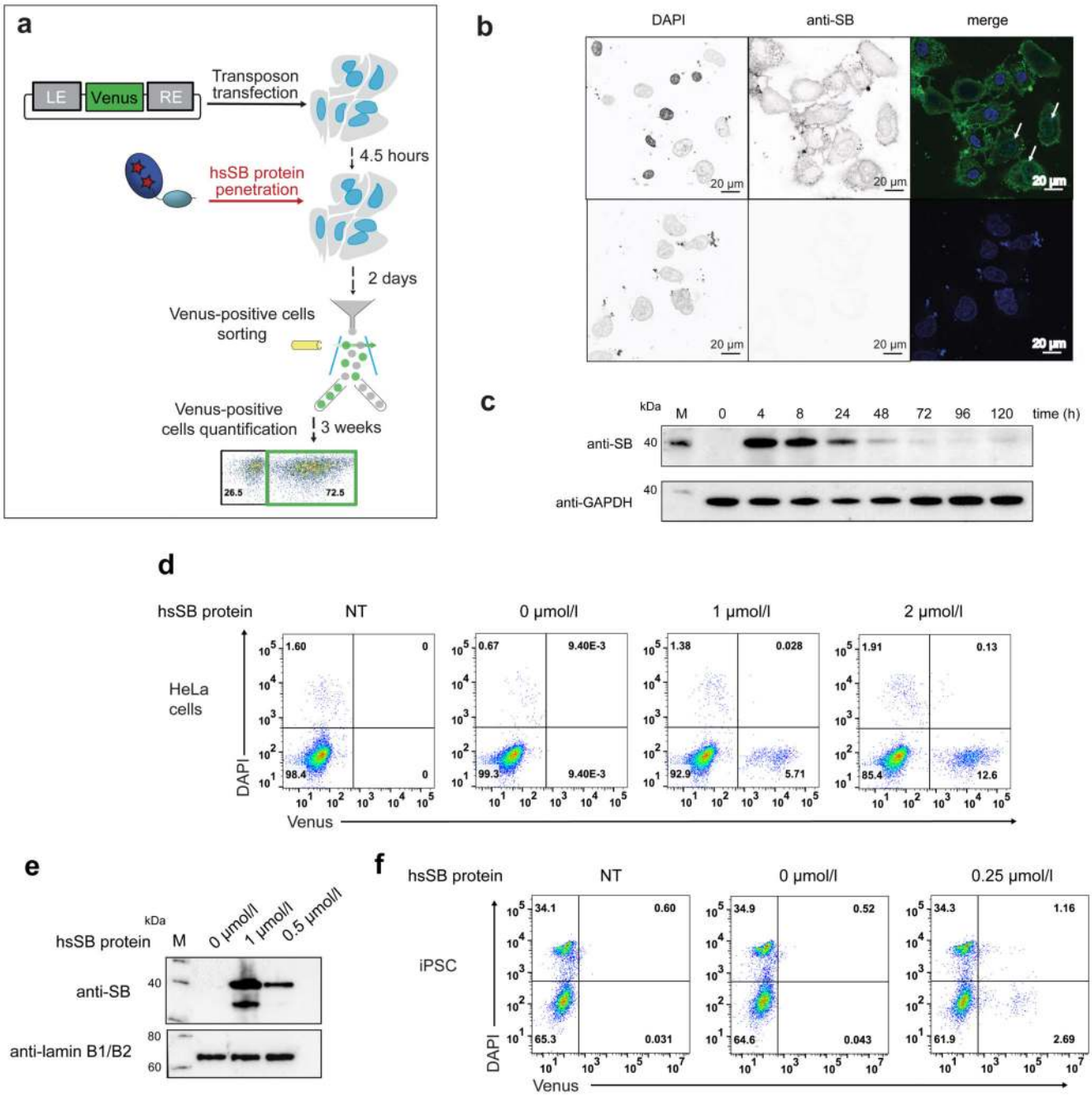


**Figure 5. hsSB transposase has intrinsic cell penetrating properties, allowing engineering of HeLa cells and iPSCs.**

**a**, Schematic representation of the cell engineering procedure, using spontaneous hsSB penetration. **b**, Immunofluorescence imaging of hsSB-treated (top) and non-treated (bottom) HeLa cells, showing DAPI-stained nuclei (left, blue), hsSB staining (middle, green) and the merge (right). Arrows mark cells with hsSB in the nucleus. **c**, Western blot analysis showing cellular uptake and retention of hsSB in HeLa cells upon addition to the culture media. Samples were blotted with either anti-SB antibody or anti-GAPDH (glyceraldehyde 3-

phosphate dehydrogenase) as internal loading control. **d**, Representative flow cytometric analysis of HeLa cells transfected with Venus-encoding transposon MC and incubated with hsSB. Venus positive cells were sorted after 2 days and analyzed 3 weeks post-delivery. X-axis: green fluorescence from Venus; y-axis: 4',6-diamidino-2-phenylindole (DAPI) staining to exclude dead cells. hsSB protein concentration in the culture media is indicated above each plot (NT: non-transfected). **e**, Western blot analysis of iPSCs following hsSB penetration from the culture media, blotted with anti-SB antibody or anti-lamin B1/B2 as loading control. **f**, Representative flow cytometric analysis of induced pluripotent stem cells (iPSCs) 3 weeks after transfection with Venus transposon MC and incubation with hsSB. **b-f**, Experiments were repeated independently two times with similar results.



Europe PMC Funders Author Manuscripts

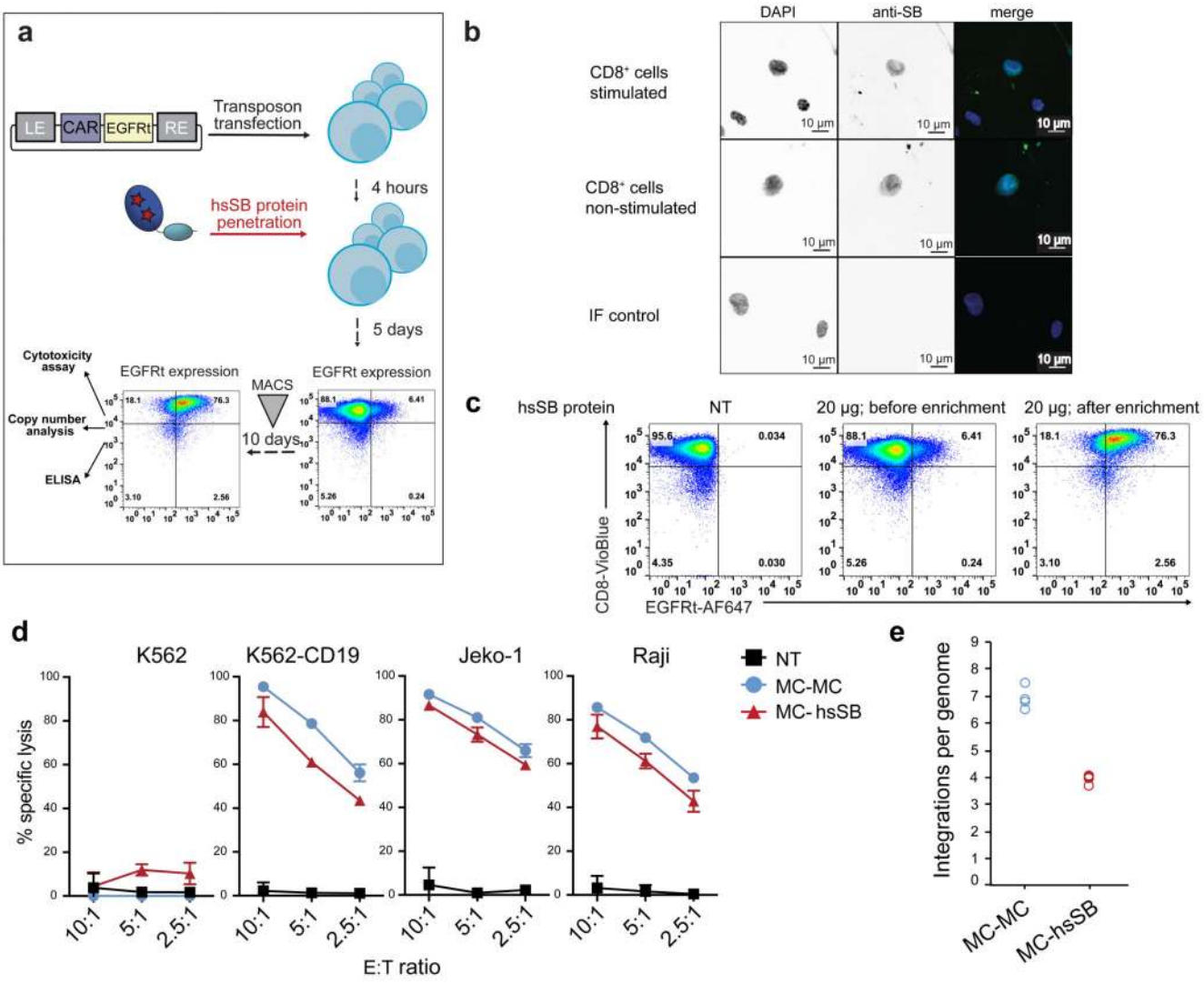


**Figure 6. Generation of CAR T cells by spontaneous hsSB penetration.**

**a**, Schematic representation of T cell engineering procedure using spontaneous hsSB penetration. **b**, Immunofluorescence imaging of T cells showing DAPI-stained nuclei (blue), hsSB staining (green) and the merge. Cells stained in absence of primary SB antibody are shown below (IF control). Experiments were repeated two times using cells from different donors with similar results. **c**, Flow cytometric analysis of $CD8^+$ T cells transfected with transposon minicircles (MC) and incubated with hsSB. $CD8^+$ T cells from healthy donors were transfected with CD19 CAR MC and enriched for CAR-positive cells (using EGFRt as marker) by magnetic associated cell sorting (MACS). Representative FACS plots are shown with fluorescence from CD8 and EGFRt specific antibodies (CD8-VioBlue and EGFRt-AF647, respectively). hsSB concentration in the culture media are indicated above each plot (NT: non-transfected). Experiments were repeated three times from three different T cell donors with similar results. **d**, Cytolytic activity of CD19 CAR T cells generated by hsSB penetration or MC-MC controls. Cytolysis was calculated from the luminescence signals of ffLuc-expressing target cells in a 5 h co-culture assay in the presence of excess luciferin




Europe PMC Funders Author Manuscripts

(NT: non-transfected; E:T ratio: effector to target ratio). Data were obtained from n = 3 independent experiments with one T cell donor. Mean values; error bars represent the s.d. **e**, The average number of CAR transgene insertions per CAR T cell genome as measured by four independent digital droplet PCR assays.



Europe PMC Funders Author Manuscripts