3-31-2022

Efficient targeted transgenesis of large donor DNA into multiple mouse genetic backgrounds using bacteriophage Bxb1 integrase.

Benjamin E. Low

Vishnu Hosur

Simon Lesbirel

Michael V. Wiles

scientific reports

Check for updates

OPEN

Efficient targeted transgenesis of large donor DNA into multiple mouse genetic backgrounds using bacteriophage Bxb1 integrase

Benjamin E. Low, Vishnu Hosur, Simon Lesbirel & Michael V. Wiles✉

The development of mouse models of human disease and synthetic biology research by targeted transgenesis of large DNA constructs represent a significant genetic engineering hurdle. We developed an efficient, precise, single-copy integration of large transgenes directly into zygotes using multiple mouse genetic backgrounds. We used in vivo Bxb1 mediated recombinase-mediated cassette exchange (RMCE) with a transgene "landing pad" composed of dual heterologous Bxb1 attachment (att) sites in cis, within the Gt(ROSA)26Sor safe harbor locus. RMCE of donor was achieved by microinjection of vector DNA carrying cognate attachment sites flanking the donor transgene with Bxb1-integrase mRNA. This approach achieves perfect vector-free integration of donor constructs at efficiencies > 40% with up to ~ 43 kb transgenes. Coupled with a nanopore-based Cas9-targeted sequencing (nCATS), complete verification of precise insertion sequence was achieved. As a proof-of-concept we describe the development of C57BL/6J and NSG Krt18-ACE2 models for SARS-CoV2 research with verified heterozygous N1 animals within ~ 4 months. Additionally, we created a series of mice with diverse backgrounds carrying a single att site including FVB/NJ, PWK/PhJ, NOD/ShiLtJ, CAST/EiJ and DBA/2J allowing for rapid transgene insertion. Combined, this system enables predictable, rapid development with simplified characterization of precisely targeted transgenic animals across multiple genetic backgrounds.

The mouse is a powerful and versatile research tool to gain a greater understanding of the human condition. A key capability is the ease and speed with which mice can be genetically modified, fueling the development of preclinical mouse models for drug development, human disease modeling and synthetic biology[1–5]. However, it remains technically challenging to integrate DNA constructs into the mouse genome in vivo in a precise and defined manner, which is essential for appropriate transgene control and expression[6]. Although targeting nucleases, e.g., Zinc Finger Nucleases (ZFNs), Transcription Activator-Like Effector nucleases (TALENs), and CRISPR/Cas9 can mediate integration of exogenous DNA into the mouse genome in the zygote via homology-directed repair (HDR), the success of this approach is fraught with poorly defined variables and an inherent low efficiency of the intended event occurring, especially when the exogenous DNA exceeds a few kilobases[6–9].

Often where larger constructs are required to be added, researchers resort to the deceptively expedient solution of random (integration) transgenesis[10]. Although this approach is well established and rapid, the resulting transgene insertions are haphazard and inefficient, often leading to unintentional genetic damage with potentially muddling downstream outcomes. For example, unpredicted local effects on host gene expression and silencing can occur[11,12]. Random transgenesis is also a misnomer. It has been shown that ~ 50% of "random" transgene insertions disrupt coding sequences of endogenous genes, often invoking large deletions and structural variations at the integration regions[13,14]. Random transgenesis often results in multiple copy integrations, including concatemers leading to aberrant gene expression or silencing[12,15–18]. Although the resulting serendipitous heterogeneity of the transgene expression can be of use, a significant amount of time is required to fully characterize the new strains, including multiple back-crosses to allow allelic segregation to fully stabilize the phenotype.

To overcome these limitations we developed a precision transgenesis platform for integration of DNA constructs with minimal extraneous sequences directly into mouse zygotes. To minimize potential damage to the mouse genome the strategy uses the well-characterized safe harbor locus, Gt(ROSA)26Sor (ROSA26) with the pre-positioning of an integrase "landing pad" (attachment sites) for subsequent recombinant transgene integration. ROSA26 encodes a long non-coding RNA (lncRNA) under the control of a constitutive promoter. The

The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA. ✉email: Michael.Wiles@jax.org

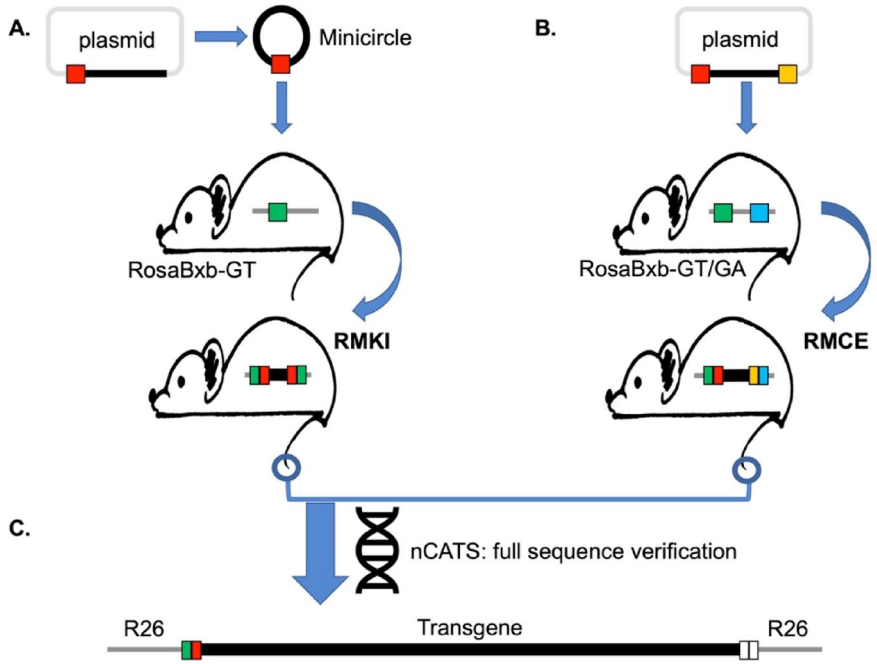

Figure 1. Overview of the Bxb1 Int-mediated transgenesis strategy. (**A**) Schematic showing RMKI allele generated by in vivo recombination between vector free minicircle donor DNA carrying a single attB site and host mouse containing a single cognate attP site in the ROSA26 locus (RosaBxb-GT). (**B**) Recombinase-Mediated Cassette Exchange (RMCE) allele generated by in vivo recombination between plasmid DNA, carrying two flanking heterologous attB sites and host mouse carrying dual cognate attP sites (RosaBxb-GT/GA) in the ROSA26 locus. (**C**) Nanopore Cas9-targeted sequencing (nCATS) enables complete verification at the sequence level of the entire transgene including the adjacent genomic DNA (R26).

region is transcriptionally active, with an open chromatin configuration leading to expression in all cell types examined. Further, its disruption does not lead to a reduction in fertility or other notable phenotypes[19–23]. To facilitate the integration of large transgenes into ROSA26 we utilized the highly efficient serine site-specific recombinase derived from the bacteriophage Bxb1[24–26].

The Bxb1 mycobacteriophage was first isolated from *Mycobacterium smegmatis* in 1990, by Dr. William R. Jacobs, Jr., (see https://phagesdb.org/phages/Bxb1/)[25]. Subsequently, it was determined that Bxb1 integrase (Int) is exquisitely suited to integrate DNA donor constructs into mammalian cells[27–30]. Bxb1 Int-mediated recombination uses a minimal attachment site of 48 bp attP (attachment *Phage*) and a cognate 38 bp site attB (attachment *Bacterial*) between the donor and recipient DNAs. The attachment sites B and P represent four half-sites. After Bxb1 recombination, a half-site from each remains, yielding two 43 bp sites (attL and attR, *Left* and *Right*) flanking the inserted DNA. Bxb1 Int requires no other secondary factors and under these conditions provides irreversible unidirectional integration of donor DNAs[31–33]. Importantly, pseudo or intrinsic Bxb1 attachment sites that could lead to background off-target integrations (OTI) of donor DNA have not been detected in the mouse genome[27,30,34–36]. Lastly, upon the comparison of fifteen integrases using cell lines, Bxb1 Int yielded approximately two-fold greater and accurate recombinants when compared to the most widely studied recombinase, φC31[27,35,37].

Here we demonstrate that Bxb1 Int can efficiently and precisely introduce donor DNA constructs into the ROSA26 locus as single-copy transgenes in the desired orientation, directly in zygotes. In the first phase of this work, we used a *single* attachment site (attP-GT) in the mouse lines C57BL/6J (B6), FVB/NJ (FVB), and NOD. Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (NSG), successfully generating RMKI alleles using DNA minicircle donors to avoid integration of the prokaryotic backbone. In a second phase, we added a *second* heterologous attachment site (attP-GA) to the ROSA26 landing pad and show that this improved transgene insertion efficiency by more than threefold. Also, using dual sites enables larger donor DNAs to be inserted via recombinase mediated cassette exchange (RMCE), excluding the vector backbone and eliminating any need for a minicircle conversion step (outline of approach in Fig. 1).

Finally, combining this strategy with nanopore Cas9-targeted sequencing (nCATS) enables us to completely sequence-verify the resulting transgenic allele rapidly *in its genomic context*. This complete strategy lends speed and confidence to the creation and characterization of genetically modified animal production with a level of precision and speed not previously obtainable.

Results

The dual attP site landing pad alleles were constructed over two phases enabling testing of multiple CRISPR/Cas9 based strategies for optimal knock-in allele generation. After establishing the single site attP allele on various backgrounds, the B6 and NSG versions were re-targeted adding a second attP site in *cis*. Both single and dual-site mouse strains were used to generate precision transgenics via Bxb1-mediated recombination. Three single site strains (B6, FVB, NSG) used donor DNA as minicircles to generate vector-free RMKI alleles. In contrast, the two (B6 and NSG) dual-site strains used plasmids or modified BACs to generate vector-free RMCE alleles.

Generation of single site Bxb1 attP-GT landing pad mouse strains (RosaBxb-GT). A single Bxb1 attP-GT site was introduced 3' of the *XbaI* site in intron 1 of the ROSA26 locus (Fig. 2). To reduce potential off-target cutting, we used an 18-mer TRU-guide gRNA[38] which consistently cuts at ≥ 85% at this locus. For the mouse strain B6 only, we used a GUIDE-seq strategy[39] introducing the attP site using Cas9 mRNA, gRNA#1888 and a 48 bp dsDNA oligonucleotide donor (duplexed oligos #1927/#1928) delivered by microinjection (MIJ). Analysis of the resulting allele in B6 showed successful integration of the complete attP-GT site into the ROSA26 locus. However, this approach was inefficient (1/15) and in this instance led to a lesion at the integration site, with a +6 bp/−3 bp INDEL preceding the 5' end of the attP-GT site; Fig. 2. As a consequence this strategy was abandoned and for all other attP integrations into other strains of mice strains we used Cas9 mediated HDR with a single-stranded oligo (#1969 or #1970) which included homology arms[13]. These data are outlined in Table 1.

As shown in Table 1, all strain backgrounds attempted resulted in successful generation of the ROSA26 attP-GT allele, with the exception of WSB. The primary challenge for CAST and WSB strains was low embryo production, due to super-ovulation resistance combined with poor embryo survival post-manipulation. For CAST this was overcome using timed matings followed by electroporation of zygotes to introduce duplexed Cas9 protein + guide and donor oligo, producing two carrier Founders. Unfortunately, while CRISPR/Cas9 targeted NHEJ was detected in WSB Founders, we did not detect knock-in of the attP site in the very limited number of offspring obtained. For all other strains listed, the attP knock-in allele was detected, transmitted and confirmed as described. After a minimum of one back-cross to the appropriate strain background, each new strain was bred to homozygosity. Ultimately a single line was selected, Supplemental Table S1.

Use of single site RosaBxb-GT strains to generate RMKI alleles. The generation of vector-free RMKI alleles was achieved as shown in Fig. 3, with the screening strategy as outlined in Fig. 4. A range of conditions were attempted to optimize the system for efficient generation of RMKI alleles with three single-site strains, B6.RosaBxb-GT, FVB.RosaBxb-GT, and NSG.RosaBxb-GT. Results from all conditions tested for each donor minicircle were combined and summarized in Table 2. This table highlights our attempts to integrate twelve unique minicircle plasmid donors ranging in size from 3.2 to 9.9 kb. The highly variable results were expected due to the multiple conditions tested. Overall, we demonstrated that more than half (22/39) of the individual conditions resulted in the successful generation of an RMKI allele, with rates as high as 33%.

In 14 independent tests we attempted to integrate twelve unique minicircles into the RosaBxb-GT single site allele across three different genetic backgrounds (B6, FVB, NSG). Minicircle DNA donor size, the number of microinjected embryos transferred and the number of positive Founders over the number of animals screened is shown. Recommended conditions are given in Material and Methods. Note, results from multiple conditions using the same donor were combined. Individual conditions tested ranged from 0 to 33% correct integrations. Aberrant alleles resulting from the minicircle production step occasionally resulted in larger inserts than intended (indicated by*). All projects taken to liveborn and established colonies, except for MC-12, which was used solely to optimize conditions and were collected as embryos for rapid analysis.

In < 5% of identified integrations, we detected aberrant integration events when analyzing RMKI alleles. These appear to be the result of unintended rearrangement occurring during donor DNA minicircle production. Analysis of these alleles suggests that in vitro φC31-based recombination designed to aid minicircle production has combined multiple plasmid DNAs producing some minicircles with 2–3 copies in tandem of the intended transgene. Where identified, these aberrant alleles are shown in Table 2. In the case of MC-1, which resulted in 3 copies of a double-inverted open-reading frame (DIO) allele designed to express a toxin only in cells where Cre is expressed, the allele was maintained and shown to function as intended (data not shown). As a general rule however, such aberrant allele lines were detected and discarded early in the screening process, or after sequencing the region by nCATS.

Addition of second site, attP-GA. Although the RMKI strategy was successful with inserts up to ~ 15 kb, the overall efficiency of donor construct integration, plus operational constraints in making minicircle donor DNA led us to engineer RMCE enabled host mice. This was achieved through the addition of a second (heterologous) attP site, with "GA" as the central dinucleotide as opposed to the wild type "GT" used for the first site. This second site (attP-GA) was placed in *cis*, 240 bp 3' of the first attP-GT site, in B6.RosaBxb-GT and NSG. RosaBxb-GT strains (Fig. 2). This sequential modification was achieved using CRISPR/Cas9 and sgRNA#2254 to mediate the HDR integration of donor oligo #2255 into zygotes homozygous for the initial attP-GT insertion, successfully generating the RosaBxb-GT/GA allele in B6 and NSG; Table 1. After a minimum of one back-cross to the appropriate strain background these were bred to homozygosity; see Supplemental Table S1.

Use of dual site RosaBxb-GT/GA strains to generate RMCE Alleles. Typically, homozygous males carrying the dual landing pad are mated to super-ovulated wild-type females, generating heterozygous embryos for MIJ with Bxb1-Int mRNA and various DNA donor plasmids. *Highly* purified donor plasmid carrying dual cognate attB sites flanking the donor transgene were MIJ'd into zygotes as outlined in Fig. 5, and results sum-

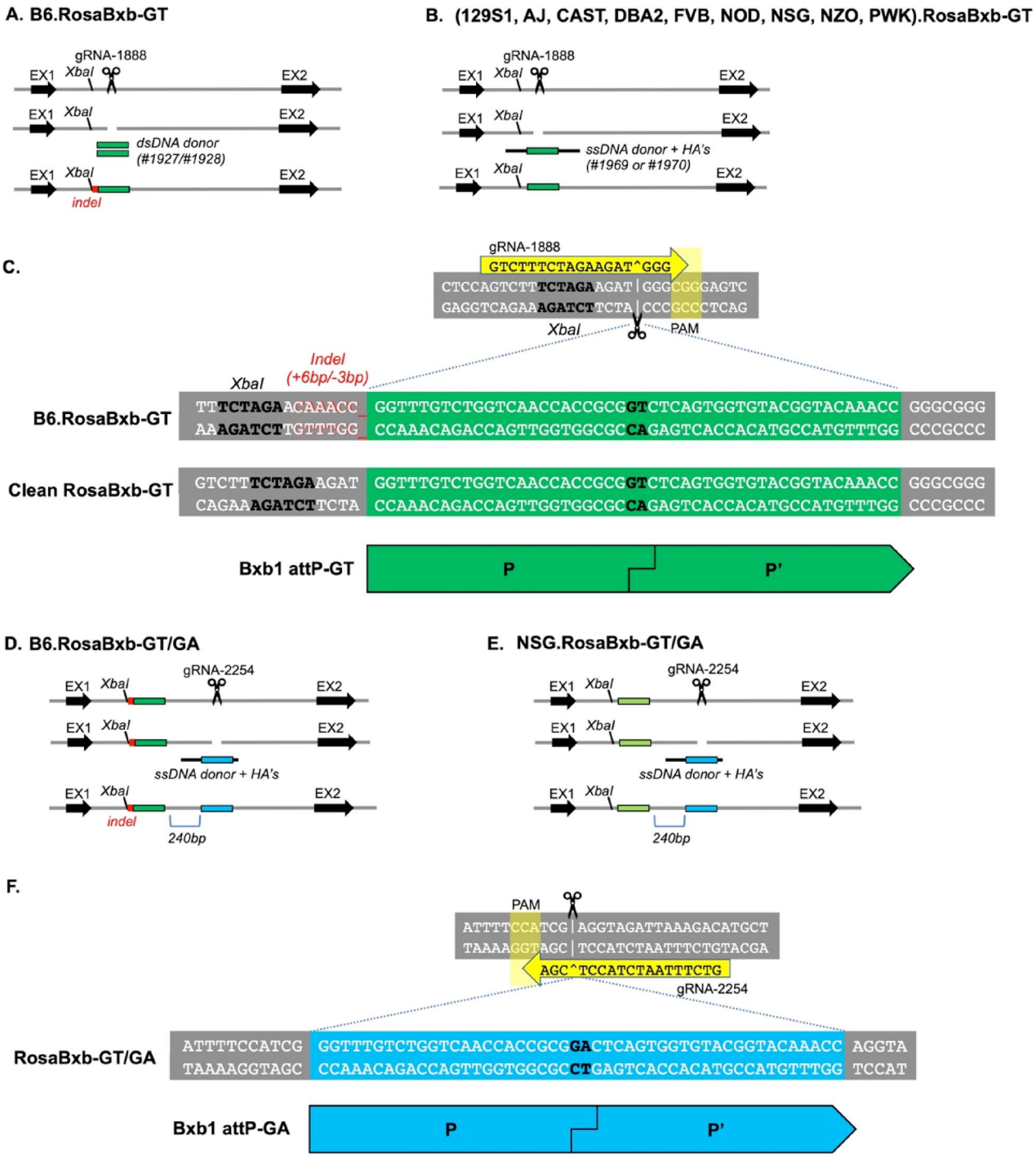

Figure 2. Construction of RosaBxb-GT alleles on multiple genetic backgrounds. (**A**) For the mouse strain B6 only, a CRISPR/Cas9 GUIDE-seq strategy was used resulting in the insertion of the attP-GT site with an indel. (**B**) For all other backgrounds a CRISPR/Cas9 mediated HDR strategy was used, resulting in scarless incorporation of the attP site to generate the RosaBxb-GT alleles. (**C**) Sequence of Bxb1 attP-GT site for B6.RosaBxb1-GT, showed that an indel occurred in B6 (only) and the correct RosaBxb-GT allele with perfect integration occurred in mouse strains 129S1.RosaBxb-GT, AJ.RosaBxb-GT, CAST.RosaBxb-GT, DBA2. RosaBxb-GT, FVB.RosaBxb-GT, NOD.RosaBxb-GT, NSG.RosaBxb-GT, NZO.RosaBxb-GT, and PWK. RosaBxb-GT. Generation of RosaBxb-GT/GA alleles on B6 (**D**) and NSG (**E**) genetic backgrounds using a CRISPR/Cas9 HDR strategy wherein the attP-GA site was placed 240 bp 3' of the attP-GT site, using gRNA #2254 and asymmetric donor oligo #2255. (**F**) Sequence detail of the Bxb1 attP-GA site region in strains B6.RosaBxb1-GT/GA and NSG.RosaBxb1-GT/GA.

Target strain (JAX Stock #)	Method	Cas9 (ng/μl)	sgRNA #1888 (ng/μl)	Donor (ng/μl)	Donor ID and length (nt)	Liveborn	Knock-in efficiency
(a) Generation of RosaBxb1-GT alleles (single site attP-GT)							
129S1 (2448)	EP	125 (P)	152	138	#1970 (200)	21/88 (24%)	7/20 (35%)
A/J (646)	EP	125 (P)	154	113	#1970 (200)	9/14 (9%)	6/9 (67%)
B6 (664)	MIJ	60	30	3	#1925/#1926 (48)	19/105 (18%)	0/19 (0%)
	MIJ	60	30	3	#1927*/#1928* (48)	15/79 (19%)	1/15 (7%)
CAST (928)	MIJ	60	30	1	#1970 (200)	0/17 (0%)	N/A
	MIJ	60	30	1	#1970 (200)	0/19 (0%)	N/A
	EP	208 (P)	139	1106	#1970 (200)	2/24 (8%)	2/2 (100%)
DBA2 (671)	MIJ	60	30	1	#1970 (200)	5/81 (6%)	1/5 (20%)
FVB (1800)	MIJ	60	30	1	#1970 (200)	70/169 (41%)	13/70 (19%)
NOD (1976)	EP	208 (P)	146	70	#1970 (200)	47/144 (33%)	27/47 (57%)
NSG (5557)	MIJ	60	30	3	#1970 (200)	6/57 (11%)	6/6 (100%)
	MIJ	60	30	3	#1969* (200)	15/62 (24%)	11/15 (73%)
NZO (2105)	EP	250 (P)	150	100	#1970 (200)	26/144 (18%)	14/26 (54%)
PWK (3715)	MIJ	60	30	1	#1970 (200)	3/16 (19%)	0/3 (0%)
	EP	208 (P)	139	1106	#1970 (200)	14/120 (12%)	3/14 (21%)
WSB (1145)	EP	250 (P)	100	11	#1970 (200)	4/19 (21%)	0/4 (0%)
	EP	250 (P)	100	11	#1970 (200)	1/6 (17%)	0/1 (0%)
	EP	250 (P)	108	56	#1970 (200)	0/15 (0%)	N/A
	EP	250 (P)	102	58	#1970 (200)	0/10 (0%)	N/A
					Totals:	257/1189 (22%)	91/256 (36%)
Target strain (JAX stock #)	**Method**	**Cas9 (ng/μl)**	**sgRNA #2254 (ng/μl)**	**Donor (ng/μl)**	**Donor ID and length (nt)**	**Liveborn**	**Knock-in efficiency**
(b) Generation of RosaBxb1-GT/GA alleles (dual site attP-GT, attP-GA)							
B6.RosaBxb-GT (28,573)	MIJ	100, 30 (P)	50	3	#2255** (131)	18/105 (17%)	2/18 (11%)
NSG.RosaBxb-GT (29,294)	MIJ	100, 30 (P)	50	3	#2255** (131)	13/99 (13%)	2/13 (15%)
					Totals:	31/204 (15%)	4/31 (13%)

Table 1. Summary of CRISPR/Cas9-mediated Knock-In efficiency of Bxb1 attP site(s) into the ROSA26 locus. A range of CRISPR/Cas9 conditions were used to generate the RosaBxb-GT and RosaBxb-GT/GA alleles on multiple genetic backgrounds. MIJ, microinjection; EP, electroporation are listed and all oligonucleotides were phosphothioated unless noted (*). Cas9 was delivered as mRNA, except where indicated as Protein, (P). All 200nt oligo donors contained the attP site at the center, i.e., 76nt of homology to the ROSA26 locus on either side. Liveborn is the number of animals available to screen at wean over the number of embryos transferred following EP/MIJ. KI efficiency is correct integration of the attP site into the ROSA26 locus. For B6 only we used a GUIDE-seq strategy with a dsDNA donor. The attP-GT was inserted into the ROSA26 locus into the same place and orientation in all genetic backgrounds using Cas9 guide #1888. The attP-GA site was inserted 240 bp 3' of the attP-GT site in previously-modified B6.RosaBxb-GT and NSG.RosaBxb-GT strains with guide #2255.

marized in Table 3. Twenty-five unique transgenic alleles were created at efficiencies ranging from 3 to 43%, with a combined average efficiency across all projects of 15% RMCE positive alleles, with constructs ranging in size from 1.5 to ~ 43 kb. As highlighted in Table 3, a few projects initially performed poorly (P-7, P-12, P-15 and P21). However, after repeating the MIJ with a newly purified donor DNA, these generally succeeded, often with dramatically improved results.

Founder offspring carrying the donor insertion were backcrossed to their respective genetic background generating N1 animals for germline transmission and complete validation of the transgenic allele. Screening was performed using a combination of PCR strategies and Sanger sequencing of product as described. A limited number of alleles were subject to nCATS to completely verify the precise nature of the transgene in its genomic context. On rare occasions (> 1%), we did identify instances where either a mis-matched recombination event occurred (i.e., attB-GA + attP-GT) or only one of the two sites successfully recombined. It should be emphasized that these were rare events, see[40], and these alleles were easily identified by our screening strategy allowing for the aberrant lines to be detected and discarded. We did detect off-target, random integrations of the donor DNA at expected frequencies for random transgenesis (~ 5% or less). Typically, Founders carrying an off-target allele were discarded. On the rare occasion where a positive founder also carried a random integration these, mice were backcrossed to segregate and select for the correct integration event.

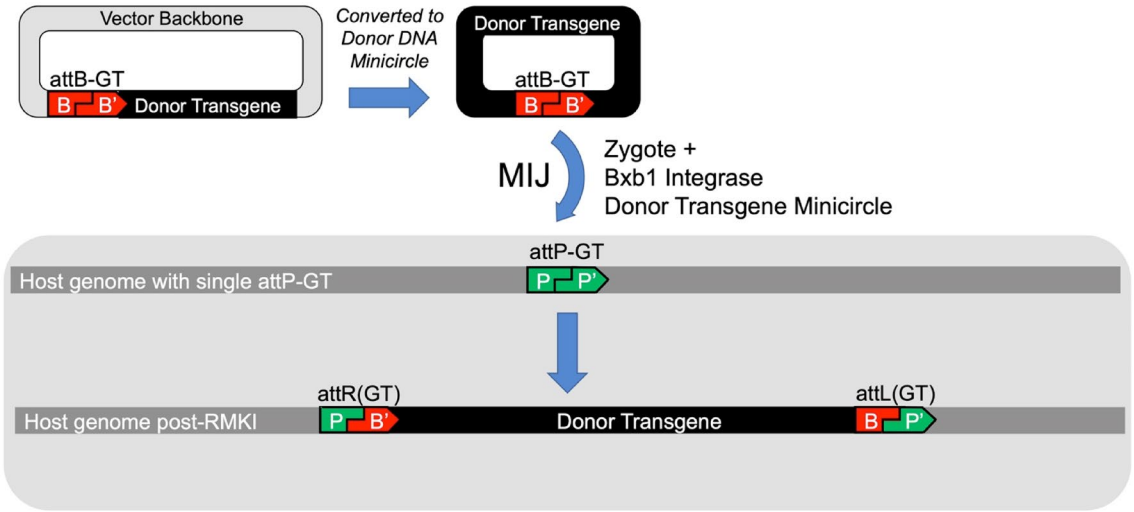

Figure 3. Schema for using single attP-GT site mouse strains targeted by RMKI. Before microinjection, the plasmid vector containing the donor transgene and a single cognate attB-GT site are converted into (vector-less) DNA minicircles and purified. Along with Bxb1 Int, the minicircles are delivered into zygotes carrying the attP-GT landing pad allele, enabling precision transgene integration as the attP (host genome), and attB (donor DNA) sites recombine to form attR and attL sites.

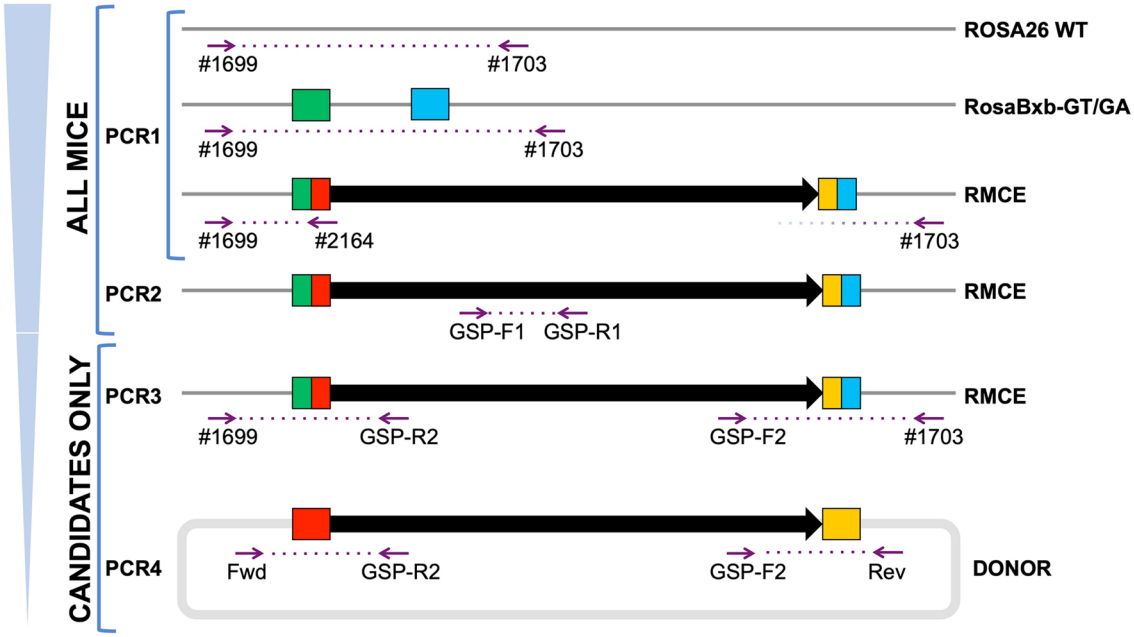

Figure 4. RMCE screening strategy. Screening offspring generated by microinjection follows a reductive series of PCRs to identify candidate animals. Both PCR1 and PCR2 assays are used to screen all potential Founder animals, while PCR3 and PCR4 assays are performed on only those positive by PCR1 and/or PCR2. A "generic" multiplex PCR1 was used to identify the presence of the RMCE allele resulting from any DNA donor, targeting both the ROSA26 locus and the recombined attR-GT site (green/red box). PCR1 serves as a DNA template quality control and used later to determine zygosity when verified animals are inbred. A project-specific PCR2, using transgene-specific forward and reverse primers (GSP), verifies that donor DNA is present in the sample (note—this is not designed to distinguish between random insertion events and the RMCE allele). PCR3 assays verify that the correct (ROSA26) integration event has occurred utilizing a PCR In/Out strategy, where the PCR span each newly created unique junction sites. This is followed by Sanger Sequencing verification of the PCR product. For transgenes < 10 Kb, re-positioning of GSP-R2 and GSP-F2 can be used to generate overlapping PCR products that allow sequence verification of the entire allele from the two long-range amplicons. Assay PCR4 is performed to screen for the presence of aberrant/random insertion using vector-backbone specific primers combined with transgene specific primers (GSP) designed to span the attB sites in the DNA donor.

Project ID	Donor size (kb)	# embryos	RMKI POS	Note
B6 background (B6.RosaBxb-GT)				
MC-1	3.2*	243	2/67 (3%)	**Aberrant (3X tandem), actually 9.6 kb*
MC-2	3.8	278	2/45 (4%)	
MC-3	4.1	200	2/58 (3%)	
MC-4	7.5	99	3/19 (16%)	
MC-5	7.5	118	2/38 (5%)	
MC-6	7.7*	125	1/42 (2%)	**Aberrant (2X tandem), actually 15.3 kb*
MC-7	9.9	594	2/135 (1%)	
Total	3.2–9.9 kb (15.3 kb*)	1657	14/404 (3%)	**Including aberrant donor insertions*
FVB background (FVB.RosaBxb-GT)				
MC-8	7.2	105	2/18 (11%)	*One test = 2/6 (33%)*
MC-9	7.3	136	3/38 (8%)	
MC-10	7.5	119	2/43 (5%)	
MC-11	7.6	99	4/33 (12%)	
MC-6	7.7	287	2/113 (2%)	
Total	7.2–7.7 kb	746	13/245 (5%)	
NSG background (NSG.RosaBxb-GT)				
MC-12	3.3	711	16/187 (9%)	*Fetal harvests. One test = 4/15 (27%)*
MC-6	7.7	114	0/20 (0%)	
Total	3.3–7.7 kb	825	16/207 (8%)	
All backgrounds combined (B6, FVB, NSG)				
Total	3.2–15.3 kb*	3228	43/856 (5%)	**Including aberrant donor insertions*

Table 2. Summary of RMKI attempts.

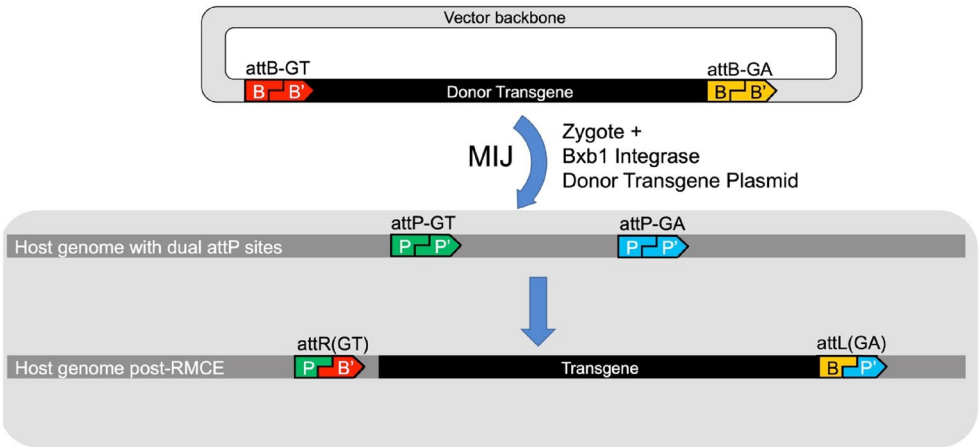

Figure 5. Schema for the use of dual attP-GT/attP-GA mouse strains targeted for transgenesis by RMCE. Bxb1 integrase mRNA and the vector carrying the donor transgene (flanked by dual attB-GT, attB-GA sites) are microinjected into zygotes carrying the cognate attP-GT, attP-GA sites. Upon integration, RMCE excludes the vector backbone and results in the precise integration of the desired donor transgene into the ROSA26 locus.

The rapid development of a Krt18 driven human ACE2 transgenic model in B6 and NSG backgrounds. Based on a construct used to make the random transgenic mouse line B6.Cg-Tg(K18-ACE2)2Prlmn/J JAX# 034860[41], a near identical 6.84 kb construct ("P-11") was assembled to express human ACE2 under the direction of a human Krt18 promoter. Using Bxb1 Int-mediated transgenesis this construct was introduced into the B6.RosaBxb1-GT/GA and NSG.RosaBxb1-GT/GA strains. Initial characterization of offspring detected the correctly targeted transgene in Founder animals in the B6 background at 19% (10/54) and in the NSG background, 29% (18/63)—these data are summarized in Table 3. Carrier Founders were backcross to their respective background and confirmed faithful germline transmission. The resulting heterozygous targeted offspring were identified and the transgene confirmed by PCR and Sanger sequencing. These N1 animals were determined to be free of random transgenes and were also verified by nCATS. For both strains, heterozygous

Donor ID	Donor size (kb)	# embryos	RMCE POS	Note
B6 background (B6.RosaBxb-GT/GA)				
P-1	1.5	78	9/23 (39%)	
P-2	3.2	128	1/40 (3%)	
P-3	3.3	130	2/36 (6%)	
P-4	3.3	130	5/36 (14%)	
P-5	3.7	129	16/37 (43%)	
P-6	3.7	149	8/48 (17%)	
P-7	4.5	63	0/11 (0%)	Rescued by upon repeat
P-7 (RPT)	4.5	106	3/22 (14%)	
P-8	4.6	158	10/43 (23%)	
P-9	5.7	165	6/50 (12%)	
P-10	5.8	140	7/39 (18%)	
P-11	**6.8**	**182**	**10/54 (19%)**	**Krt18-ACE2**
P-12	7.1	150	1/31 (3%)	Improved by upon repeat
P-12 (RPT)	7.1	190	12/63 (19%)	
P-13	8.5	105	1/33 (3%)	
P-14	8.8	62	1/14 (7%)	
P-15	9.4	73	0/22 (0%)	Rescued by upon repeat
P-15 (RPT)	9.4	143	8/44 (18%)	
P-16	10.3	67	2/11 (18%)	
P-17	11.5	102	1/37 (3%)	
P-18	12.3	137	4/36 (11%)	
P-19	13.0	115	2/21 (10%)	
P-20	13.0	104	1/33 (3%)	
P-21	23.6	88	0/28 (0%)	Rescued by upon repeat
P-21 (RPT)	23.6	107	2/26 (8%)	
P-22	25.9	112	2/26 (8%)	
P-23	30.6	139	4/35 (11%)	
P-24	42.9	129	3/21 (14%)	
Total	1.5–42.9 kb	3381	121/890 (14%)	
NSG background (NSG.RosaBxb-GT/GA)				
P-11	**6.8**	**260**	**18/63 (29%)**	**Krt18-ACE2**
P-25	9.1	162	8/20 (40%)	
Total	6.8–9.1 kb	422	26/83 (31%)	
All backgrounds combined (B6 and NSG)				
Total	1.5–42.9 kb	3803	147/973 (15%)	

Table 3. Successful targeted transgenesis in dual site RosaBxb-GT/GA strains. Summary of 29 independent attempts to integrate 25 unique donor transgenes into the RosaBxb-GT/GA dual-site allele in B6 and/or NSG genetic backgrounds. Shown are DNA donor size, the number of microinjected embryos transferred, and the number of positive Founders animals which carry the transgene, expressed as a percentage of the number of animals screened. The Krt18-ACE2 construct (P-11) was delivered to both B6. RosaBxb-GT/GA and NSG. RosaBxb-GT/GA hosts, highlighted in bold. For a few constructs (P-7, P12, P-15, P-21), the microinjection was repeated following re-purification of the donor vector, which resulted in improved RMCE rate.

animals were available for experiments ~ 4 months from the date of injection. Both lines were bred to homozygosity and deposited at the Jackson Laboratory Repository (see Supplemental Table S1).

Comparative transgene sequence verification of the B6.RMCE[Krt18-ACE2] allele versus a random ACE2 transgenic strain B6.Cg-Tg(K18-ACE2)2Prlmn/J. The development of a targeted transgenesis strategy was driven by the ever-increasing need for efficient and precise genetic engineering of animals. To fully verify insertions of increasingly larger DNA constructs we developed a DNA Nanopore sequencing workflow, based on nCATS, which provides enrichment of reads spanning the transgenic insertion[42]. Currently, we have applied this targeted sequencing workflow to Bxb1-int mediated transgenic strains with insertions of 5–43 kb. In Fig. 6a we also present an example of the nCATS workflow and the resulting characterization of the B6.RMCE[K18-HuACE2] mouse. This allele was validated by gRNA targeting and subsequent enrichment of an 8.5 kb region, generating an average per-base coverage of 195X over the transgene and surrounding

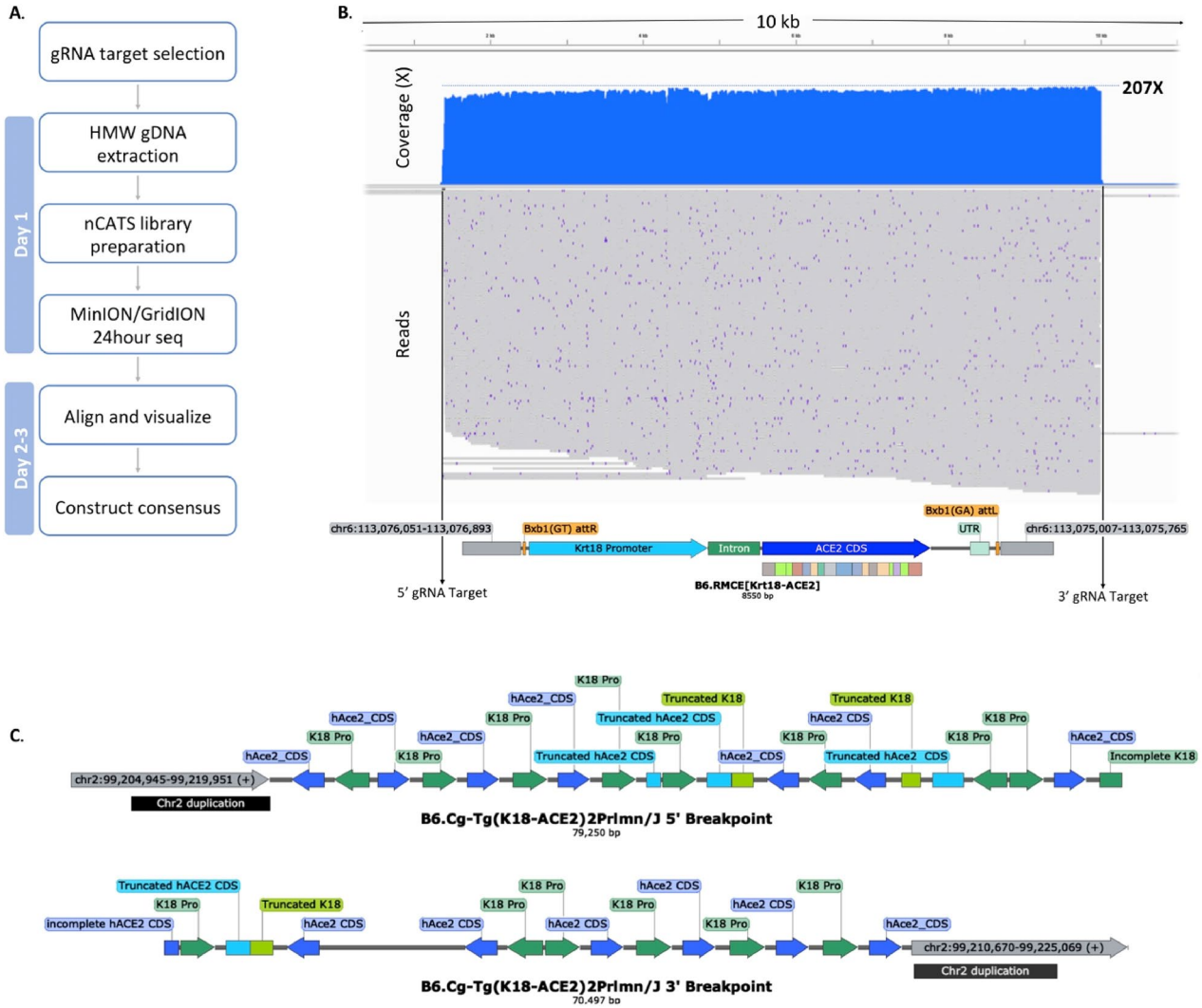

Figure 6. Characterization of transgenic mice using nCATS. (**A**) Post sequencing workflow to construct insert regions and identify genomic locations. (**B**) nCATS characterization of B6.Cg-Tg(Krt18-ACE2) allele generated using the Bxb1 Int to achieve single copy integration, 215 reads on-target generating 195X coverage across the entire 8.5 kb insertion site and 900/800 bp genomic breakpoints. (**C**) nCATS characterization B6.Cg-Tg(K18-ACE2)2Prlmn allele, 5' and 3' breakpoints identified at least 11 copies of K18-ACE2 in varying orientations.

chromosomal region (Fig. 6b). These nCATS-derived data, which include 800 bp and 900 bp flanking genomic (ROSA26) sequence, fully support that the transgenic sequence integrated precisely as expected.

To illustrate the precision of Bxb1 Int-mediated transgene integration compared to conventional random integration, we applied nCATS to the B6.Cg-Tg(K18-ACE2)2Prlmn/J mouse (JAX Stock #034860). Guides (Supplemental Table S2) targeting the suspected integration site on Chromosome 2 (chr2:99204917–99225067) as well as against the CDS of ACE2 revealed the complex architectural nature of this random transgenic insertion. These data lead to a reconstruction of a 79.3 kb 5' and 70.5 kb 3' breakpoint sequences (Fig. 6c) and identified K18-ACE2 concatemers consisting of at least 11 copies in both orientations. In addition, we also detected a 10.5 kb duplication of Chromosome 2 at the periphery of each breakpoint along with K18-ACE2 cassette truncations at sites of inversion throughout the insertion. Due to the repeat-rich complexity of the insertion we were unable to generate a full consensus for the entire integration site.

Discussion

There is a need to create genetically modified animals where large transgenes are integrated precisely, in a defined genomic context efficiently to further advance biological research and transgenic development[3,6,43]. To address this we developed a strategy enabling efficient precision integration of large DNA constructs into the ROSA6 safe harbor locus directly via the zygote, across multiple mouse genetic backgrounds. Our aim was high efficiency, site-specific recombination which is not significantly impacted by donor constructs of up to at least 10 kb. For this we selected from the limited number of lysogenic bacteriophages tested for targeted integration into the mammalian genome, Bxb1 integrase.

To test the Bxb1 Int approach, we used CRISPR/Cas9-mediated insertion of the Bxb1 attP site with the wild type "GT" central dinucleotide into the ROSA26 locus in mouse background B6, FVB and the immunocompromised line NSG. Once established, these lines B6.RosaBxb1-GT, FVB.RosaBxb1-GT and NSG.RosaBxb1-GT) were used to test Bxb1 Int-mediated integration. We focused on establishing MIJ conditions (Fig. 3). To speed the process we developed a PCR screening and sequencing strategy which also covered the flanking regions of the ROSA26 integration attP-GT site. This provided sequence verification of Bxb1 Int-mediated insertions in the context of the ROSA26 locus. Our tests, summarized in Table 2, show clearly that targeted integration into ROSA26 attP-GT occurs at levels at or more often, considerably higher than random transgenesis.

A key driver of this work was to develop improved access to efficient transgenic targeting using multiple genetic backgrounds to facilitate comparative genetic background studies. Our success in the construction of and use the ROSA26 Bxb1-GT allele in three backgrounds inspired us to add the same attP site/location to a host of novel genetic backgrounds. The choice of these based on the Collaborative Cross and also included DBA/2J due to its importance in developing the BXD RI strains[44,45].

The genetic modification of many genetic backgrounds has been hindered as many of these "non-conventional" mouse strains exhibit poor zygote yield and/or survivability post modification attempts. Although mouse embryonic stem cell (mESC) lines exist for many backgrounds and could be used to circumvent this, the introduction of constructs into these is often encumbered by drug selection markers, requires additional time to produce the chimeras and obtain germline transmission, verses direct modification of zygotes from the required background[46–48]. Also, by using extant mouse strains the actual genetic background of the resulting models will represent the present genome of the background and not that of a long-established mESC line[49].

Although CRISPR/Cas9 mediated HDR with ssDNA oligonucleotides was relatively efficient, embryo yield and survivability from some genetic background remained problematic. Our initial attempts to introduce the attP-GT site used MIJ of zygotes. However as shown in Table 1, the results were mixed on both success to live born and the efficiency of introduction of the attP-GT site into the ROSA26 locus. This difficulty was especially evident in the wild inbred strains CAST, WSB and PWK, which failed due to a low embryo yield/female, compounded by subsequent poor survivability to term following MIJ. The use of electroporation as an approach to introduce CRISPR/Cas9 and *small* ssDNA donors has been shown by others to be efficient in B6 mice[50,51]. Using conditions previously established for B6 zygote electroporation we introduced the attP-GT site into a series of different genetic backgrounds. Although multiple direct comparative experiments of the overall efficiency of MIJ vs electroporation were not part of this study, we did succeed in obtaining attP-GT site integration in multiple strains, including CAST and PWK (Table 1), using electroporation. A review of these data suggests that the success is linked to the higher survivability of zygotes to term post electroporation as compared to MIJ, although further experiments would be needed to fully confirm this hypothesis. Combined, these efforts generated a diverse genetic panel of ten strains (listed in Tables 1 and 2) carrying the attP-GT site in the ROSA26 locus. However, even with our best efforts we failed at this time to generate the desired allele in the wild inbred mouse strain WSB due to poor availability of zygotes combined with low survivability post electroporation.

Out of concern for potential off-target integration of the donor attP-GT sequence, we modulated the concentration of the oligonucleotide delivered by EP in some of our tests. Although we did not screen for off-target integration, we did find that greatly reducing the level of donor oligo in the electroporation does not necessarily result in reduced knock-in efficiency. For example, despite reducing the concentration of the donor by more than 90% (to 70 ng/µl) when targeted NOD/ShiLtJ embryos, we still successfully generated the intended HDR allele with 57% efficiency (Table 1). This parameter should be further investigated, especially in the light of possible off-target integrations at the higher concentrations.

Precise Bxb1 Int-mediated integration of donor constructs into the ROSA attP-GT site was demonstrated in three genetic backgrounds. However, we sought to further improve efficiency and easy of use by exploring the use of a second Bxb1 attachment site (Fig. 2). We reasoned that the second site in close proximity to the initial attP-GT site would increase stochastic interaction and hence integration. A key attribute of using a dual heterologous attachment site strategy is that the donor vector can now be designed to integrate by RMCE, leading to excision and loss of the plasmid backbone and directional integration of the construct (Fig. 5)[52]. We selected the GA dinucleotide for our second site based on Jusiak et al.'s work which compared wild type Bxb1-GT to the Bxb1-GA variant in cell lines and noted that the GA variant resulted in more than double the integration efficiency[40]. They also reported minimal crosstalk between Bxb1 GA and GT att sites, an essential requirement when using dual heterologous sites[40]. We added the Bxb1 attP-GA site *in cis* to the attP-GT site in B6.RosaBxb-GT and NSG.RosaBxb-GT mouse, creating the strains B6.RosaBxb-GT/GA and NSG.RosaBxb-GT/GA respectively. Using these lines tests were conducted with a range of constructs (Fig. 5), and data summarized in Table 3. This process improvement also lead to operational advantages, including avoiding troublesome minicircle preparation, use and artifacts. Surprisingly, when using these dual attP site mice with cognate constructs we saw a more than three-fold increase in overall efficiency over similar tests with animals carrying a single attP-GT site. This increased efficiency is intriguing, perhaps reflecting synergy or simply an additive effect of the combined Bxb1-GT and the more efficient Bxb1-GA sites. RMCE may also improve operational efficiency, especially as transgene plasmid preparation is considerably simpler and potentially of higher quality[52–56]. The use of RMCE also opens the possibility of constructs of tens, if not hundreds of kilobases, which would be difficult to prepare as minicircles at the concentration and purity required.

Our data are insufficient to ascertain if there is a strong correlation between construct size and efficiency of integration. However, it would be logical to assume that increasing the size of the insert is likely to decrease the integration efficiency. Using identical DNA concentrations for MIJ (e.g., 30 ng/µl), larger constructs are actually delivered at a lower *copy* number per zygote, and the reduced molar concentration of donor attB sites would be expected to impact integration efficiency due to a simple reduction in stochastic interactions[57]. It also should be considered that larger constructs are more susceptible to DNA breakage during handling, including during MIJ,

which would be expected to reduce their integration efficiency. To counter this, it may be worthwhile to incorporate some of the strategies previously described for BAC transgenics, notably modifying the MIJ buffer[58,59]. Preliminary work on this front shows promise, however an increase in aberrant insertion events suggests the need for greater optimization.

To demonstrate the Bxb1-int approach's speed and to aid SARS-CoV research, two novel mouse models were created based on a previous random Krt18 driven human ACE2 transgenic mouse line[41]. Using dual aatP sites we produced a large number of ROSA26 targeted transgenics enabling the development of sequence-verified, single copy, N1 heterozygous transgenic alleles in the B6 and NSG backgrounds in under five months. Bxb1-Int mediated transgenesis is directional, for the Krt18-ACE2 lines developed here the transgene was inserted in the same orientation as Gt(Rosa26)Sor lncRNA. Ideas concerning which direction is optimal for transcription of inserted transgenes into the ROSA26 remain unresolved and are undoubtedly complex involving the particular integration site and the construct's promoter[21,23]. Preliminary work on these two strains indicated that the ROSA26 transgene human ACE2 under the Krt18 promoter is transcribed at high levels in both B6 and NSG backgrounds (data not shown).

A central tenet of the approach outlined here is to provide precision and definition of DNA integration. In pursuit we also developed methods to deliver complete sequencing of integrated DNAs. using amplification-free targeted long read sequencing by nCATS. This allowed for example here, validation of > 43 kb insertions including ~ 1 kb of the flanking regions. As expected, nCATS verification of the Bxb1 Int-generated Krt18-ACE2 strains showed a single on-target integration with no disruption to the surrounding genomic sequence. These data demonstrate that the Bxb1 Int system can precisely integrate transgenes intact into a known genomic location and this can be rapidly validated including the surrounding genomic region. Combining the Bxb1 system with nCATS as demonstrated here, allows us to construct models with ever increasing confidence with reduced/no host chromosomal deformation. This is in sharp contrast, to the random transgenic B6.Cg-Tg(K18-ACE2)2Prlmn/J mouse where we clearly saw highly abberant integration. This illustrates how random integration can adversely impact the host genomic environment, for example here we identified a partial duplication of Chromosome 2 (Fig. 6c). Additional models generated via random integration transgenesis have been assayed with nCATS, and every model assayed showed concatemerization and often deletion of large sections of the host chromosome.

Conclusion and future directions

The strategy devised and presented here was envisioned to be used in any laboratory with access to zygote MIJ and general assisted reproductive technologies. Operational advantages include: scalability, high efficiency, rapid screening and defined precision integration of constructs to at least 43 kb. The approach can be used on multiple genetic backgrounds extending ideas of genetic diversity and crucially providing a uniform integration site for expression of transgenes. The attP/B approach also makes feasible ideas of subsequent sequential gene addition or "gene-stacking" by embedding other heterologous att sites in introduced constructs or by multipart assembly in vivo[33,60,61]. The advent of advanced sequencing protocol/methods exemplified here in the single molecule reading of ~ 80 kb further advances the reality of precision in the genetic modification of animals. Future directions with site specific recombinases include the potential to isolate novel integrases and/or manipulate targeting via directed evolution, enhancing efficiency through integrase site engineering, reprogramming specificity, or fusion to, e.g., Cas9 + gRNA, enabling targeting insertions to specific naturally occurring sequences within the genome[28,30,37,62,63]. Such designer site specific recombinases would have use in synthetic biology, biotechnology and genetic engineering, including human gene therapy.

Materials and methods

Zygote Microinjection. Microinjection was performed using a Zeiss AxioObserver.D1 microscope with Eppendorf NK2 micromanipulators and Narashige IM-5A injectors. Needles for microinjection were pulled fresh daily using WPI TW100F-4 capillary glass and a Sutter P97 horizontal puller. Zygotes were removed from culture and placed onto a slide containing 150µL of fresh room temperature M2 media. Standard zygote microinjection procedure was followed with special care taken to deposit material into both the pronucleus and the cytoplasm of each zygote. Before same-day embryo transfer was initiated, injected zygotes were removed from the slide and rinsed through three 30µL drops of equilibrated (under oil) K-RVCL and pooled in a separate 30µL microdrop of equilibrated K-RVCL.

Zygote electroporation. After collection, zygotes were placed into a 100µL drop of room temperature M2 media, then placed in Acidified Tyrode's solution (Sigma, T1788) for 10 s followed by three washes through 100uL drops of room temperature M2 media. Zygotes were transferred into a 100µL drop of Opti-MEM media. Zygotes were then removed in a 10 µl volume, combined with 10 µl of the CRISPR reagents and deposited in a 1 mm electroporation cuvette (BTX, P/N 610) and electroporated in a Square Wave Electroporation System (BTX, ECM 830) using a voltage of 30 V, 100 ms pulse interval, 1 ms pulse duration, two pulses with 5 repeats (total of 6). Repeated aliquots of room temperature M2 media (100µL) were used to flush the zygotes from the cuvette to ensure all zygotes were recovered. Electroporated zygotes were then rinsed through three 30µL drops of equilibrated K-RVCL before being placed into a separate 30µL microdrop of equilibrated K-RVCL where they were allowed to recover before being processed for same-day embryo transfer.

Embryo transfer. Zygotes processed for same-day transfer were removed from culture and placed in a 1.8 mL screw-top tube (Thermo Scientific, 363,401) containing 900 µL of pre-warmed/equilibrated K-RVCL media for transport to the surgical station. Zygotes were removed from the tube and transferred via the oviduct into d0.5 pseudopregnant CByB6F1/J females (age 9–11 weeks).

Generation of RosaBxb-GT and RosaBxb-GT/GA landing Pad alleles in host strains. Bxb1 attachment sites (attP) were inserted into the ROSA26 locus by CRISPR/Cas9 using TruGuide gRNAs[38] with donor oligonucleotides directly into zygotes using either microinjection (MIJ) or electroporation (EP), as detailed above. Project-specific parameters are indicated in Table 1, and all donor oligonucleotides, CRISPR gRNA targeting sequences and PCR screening primers are listed in Supplemental Table S2. Microinjection reagents were assembled as described previously[64]. Briefly, a 25 µl solution containing the Cas9 mRNA (60–100 ng/µl, Trilink), sgRNA (30–50 ng/µl), donor oligonucleotide (1–3 ng/µl, IDT) and RNasin® (0.2U/µl, Promega, P/N 2515) was assembled in MIJ TE buffer (10 mM Tris/0.1 mM EDTA/pH7.5; IDT). For some projects, Cas9 protein (100 ng/µl, PNABio P/N CP01), was also included (Table 1). Some strains were generated using electroporation as described by[50,51] and detailed above. Briefly, a 10 µl solution containing the Cas9 protein (125–250 ng/µl, PNABio P/N CP01), sgRNA (100–154 ng/µl), and donor oligonucleotide (11–1106 ng/µl, IDT) was prepared in MIJ TE buffer, incubated for 30 min at 37 °C and then placed on ice before combining with zygotes.

For all strains created here, insertion of the initial Bxb1 attP-GT site used CRISPR/Cas9 with gRNA #1888, targeting the ROSA26 locus. For only mouse strain B6, a CRISPR/Cas9 GUIDE-seq strategy using duplexed donor oligos #1927/#1928 was applied[39]. For all other strains insertion of the initial Bxb1 attP-GT site used CRISPR/Cas9 mediated HDR with single-stranded donor oligo #1969 or #1970. The second Bxb1 site attP-GA, was inserted into zygotes isolated from B6 or NSG RosaBxb-GT mice homozygous for the initial Bxb1 attP-GT site and used CRISPR/Cas9 mediated HDR with gRNA #2254 and donor oligo #2255. The resulting two 48 bp attP sites are positioned 240 bp apart, with the initial attP-GT site placed 4 bp from the *XbaI* site in intron 1 of the ROSA26 locus. In all cases following MIJ or electroporation zygotes were transferred to pseudopregnant females and brought to term.

Tissue from tail tip or ear punch was used to isolate DNA for PCR analysis as outlined in[64]. Specifically for the first attP-GT site we used PCR primers #1699 and #2162, yielding a 254 bp product if the RosaBxb-GT allele was present. For the detection of the second site attP-GA, PCR using primers #2161 and #2162 yielded a 336 bp product if the RosaBxb-GT/GA allele was present. To fully validate the modified ROSA26 locus, PCR using primers #1699 & #1703 flanking the entire modified region were used followed by Sanger sequencing. After one back-cross to the unmodified parental strain, lines confirmed to be carrying the attP site/s were inbred to establish homozygous colonies as listed in the Results. The RosaBxb-GT/GA alleles on B6-albino and NOD.ShiLtJ backgrounds were produced by selective back-crossing of B6.RosaBxb-GT/GA mice with B6(Cg)-*Tyr*$^{c-2J}$/J (stock#58), and NSG.RosaBxb-GT/GA with NOD/ShiLtJ (stock#1976), respectively.

RMKI host donor plasmid construction. To prepare minicircle DNA, desired transgenes are first cloned into a donor host plasmid (p5087). This plasmid was derived from the MN530A-1 (System Biosciences) by replacing the GFP Reporter sequence with a multiple cloning site using the restriction enzymes *XmaI* and *StuI* and duplexed donor oligonucleotides #2020 and #2021 (Supplemental Table S2). This plasmid was further modified to include the Bxb1 attB-GT site using restriction enzymes *XbaI* and *EcoRI* and duplexed oligonucleotides #2025 and #2026. Donor plasmid DNA (containing the attB-GT and the desired transgene) is prepared for injection by conversion into a vector-free minicircle using the MC-Easy™ Minicircle DNA Production Kit (SystemBio). Minicircle DNA is then isolated using the Purelink HiPure Plasmid Filter Midiprep kit (Invitrogen). To ensure complete removal of the parental plasmid, a restriction digest is performed to exclusively linearize the parental plasmid followed by an overnight digestion with Plasmid-Safe™ ATP-Dependent DNase (Lucigen). Prior to MIJ, the resultant minicircle DNA is *extensively* purified by phenol–chloroform extraction, precipitated with ethanol-sodium acetate and reconstituted in nuclease-free 10 mM Tris/0.1 mM EDTA/pH7.5.

Bxb1 Int mRNA. The Bxb1 Integrase mRNA used in RMCE and RMKI microinjections was generated by Trilink Biotechnologies (San Diego, CA USA), as a custom order. The mRNA was fully substituted with Pseudo-U, enzymatically capped and 2'OMethylated (Cap1), and delivered at ~ 1 mg/ml in 1 mM Sodium Citrate pH6.4. Batches of 10 µg single-use aliquots were prepared following phenol–chloroform purification with ethanol-ammonium acetate precipitation and reconstituted in MIJ TE and then stored at -80C until used. The full sequence provided to Trilink is shown in the Supplement, and was derived from pCAG-NLS-HA-Bxb1, a gift from Pawel Pelczar (Addgene plasmid # 51,271; http://n2t.net/addgene:51271; RRID:Addgene_51271)[65].

RMKI by Bxb1 Int using RosaBxb-GT Host Mice. RMKI of donor DNA into the single-site RosaBxb-GT host strains was achieved by pronuclear microinjection of 1–30 ng/µl donor minicircle DNA with 100 ng/µl Bxb1 mRNA (Trilink) into zygotes of the desired attP-GT carrier strains. Typically, host zygotes heterozygous for Bxb1-attP were used, as these can be readily generated by super-ovulating wild-type dams, using for example C57BL/6J females mated to B6.RosaBxb-GT homozygous studs. Microinjection preps were prepared in MIJ TE supplemented with RNasin® (0.2U/µl, Promega, P/N 2515).

RMCE donor plasmid construction. For RMCE alleles, DNA vector requires the Bxb1 attB-GT and Bxb1 attB-GA sites are positioned in the correct orientation and flanking the desired sequence. Our donor plasmid p5154 (Addgene reference 175,390) is intended for small-medium sized donor DNAs (< 15 kb) while the p5155 (Addgene reference 175,391) is a low-copy version for use with larger donor DNAs (> 15 kb). Prior to microinjection, plasmid DNA containing the requisite attB-GT and attB-GA sites flanking the transgenic donor DNAs are isolated using the Purelink HiPure Plasmid Filter Midiprep kit (Invitrogen) followed by *extensive* purification via phenol–chloroform extraction, ethanol-sodium acetate precipitation with a final reconstitution in nuclease-free 10 mM Tris/0.1 mM EDTA/pH7.5.

RMCE in RosaBxb-GT/GA host mice. Recombinase-Mediated Cassette Exchange (RMCE) of donor DNA into the dual-site RosaBxb-GT/GA host strains is achieved by MIJ of 1–30 ng/µl donor DNA with 100 ng/µl Bxb1 mRNA (Trilink) into zygotes of the desired strain. Generally, host zygotes heterozygous for the dual Bxb1-attP sites were generated by super-ovulating wild-type dams, e.g., B6 mated to B6.RosaBxb-GT/GA homozygous studs. MIJ preps were prepared in nuclease-free 10 mM Tris/0.1 mM EDTA/pH7.5 (IDT).

Screening offspring for RMCE and RMKI donor integration. A generalized reductive sequential screening strategy for identifying all Bxb1 Int-mediated alleles was developed using a series of four to six simple PCR assays on crude DNA lysates prepared from tail-tip or ear punch biopsies[64]. As outlined in Fig. 4, these assays are designed to rapidly identify the correctly integrated alleles as well as screen for founder animals that also carry random insertions of donor DNA. The PCR primers used are defined in Supplemental Table S2. It is strongly recommended that time be taken to optimize these screening primers, see[64].

PCR1 Assay is designed to rapidly identify candidate Founders using a generic multiplex PCR with primers #1699, #1703 and #2164. PCR1 also verifies the integrity of the genomic DNA template, yielding a wild type band of 776 bp, simultaneously generating a unique band in samples where the recombination event between attP-GT and attB-GT has occurred (~252 bp); Fig. 4. This PCR is performed on all potential Founder animals.

PCR2 Assay is a transgene-specific PCR using gene sequence-specific primers (GSP) designed to target a unique portion of the donor and is performed on all candidate offspring. Note, while this assay does not differentiate between random transgene events and correctly integrated alleles, its increased sensitivity will detect candidates that the multiplex PCR1 may miss due to low copy number.

PCR3 Assays Founder animals that show a positive result by PCR1 or PCR2 are next screened using two project-specific In/Out PCR assays designed to amplify the allele across both left and right ROSA26/Transgene junctions. As shown in Fig. 4, the In/Out-Left (IOL) PCR uses a forward primer in ROSA26 5' of the attP-GT site (e.g., primer #1699) with a transgene-specific reverse primer (GSP) that is 3' of the attB-GT site. Similarly, the In/Out-Right (IOR) PCR assay uses a reverse primer in ROSA26 3' of the attP site (e.g., primer #1703) with a matching donor-specific forward primer. Resulting PCR products are sequenced to confirm the correct junction between host and donor DNA resulted[64]. Where possible (~ < 10 kb), these assays are modified to enable full-sequence verification using long-range PCR. In these instances, the In/Out PCRs are re-designed so as to overlap, i.e., the forward primer for the IOR PCR should be 5' of the reverse primer for the IOL PCR.

PCR4 Assays To identify random transgenesis in putative candidate Founders, project-specific PCR for OTI's are performed. These assays are designed to target across the attB site(s) in the *donor* DNA using vector backbone-specific primers with gene matching transgene-specific primers (GSP). Products from one or both of these assays indicate either a random insertion or an aberrant integration at the landing pad site has occurred.

RMKI alleles generated by minicircle integration into a single attP-GT landing pad in ROSA26 are screened in the same manner as RMCE alleles, except only one PCR4 Assay is needed to span the single attB site in the minicircle donor. However, an additional PCR is recommended to screen for the presence of any parental plasmid backbone that may have eluded the minicircle production process. These assays are project-specific and should use primers spanning the φC31 and attB sites in the parental plasmid (not shown).

Following the RMKI/RMCE strategy outlined here, Founder animals are expected be heterozygous for any Bxb1 Int-mediated integration. However, donor DNA integration may not occur with 100% efficiency in the single cell zygote, leading to a mosaic founder animal whose actual zygosity will be less than 50%. To confirm germline transmission and establish the integrity of the allele, one to four carrier Founders are back-crossed to the wild type parental strain (*never* other Founders). The resulting N1 animals are then subjected to the same PCR screening process (including PCR4), regardless of the genotype of the mosaic Founder animal. We recommend that a single verified N1 animal then be used to establish the new strain. Consequently, subsequent genotyping requires only a single multiplex PCR (e.g., IOL PCR3 with primer #1703 included) to determine the presence and zygosity of the new allele.

nCATS and data visualization. Unique gRNA's for nanopore Cas9-targeted sequencing (nCATS) are designed 5' and 3' of the region to be sequenced (see Supplemental Table S2). For genomic DNA extraction, tissue is pulverized on dry ice in a Bessman Tissue Pulverizer (Spectrum™, 189,476). High molecular weight gDNA extraction using a Monarch® HMW DNA Extraction Kit for Tissue (New England Biolabs, T3060L) is performed on tissues isolated from N1 or higher generation mice. We do not recommend this process on founder animals as they are both often mosaic and potentially heterogeneous for the modification. nCATS targeted libraries are constructed as detailed in[42] with the exception of ACE2-CDS of B6.Cg-Tg(K18-ACE2)2Prlmn/J JAX# 034860, where a single cut and read library was made per ACE2-CDS targeting gRNA, and pooled prior to Ampure bead purification. nCATS libraries are routinely sequenced for 24 h on a GridION using a R9.4.1 flow cell (Oxford Nanopore Technologies). After 24 h of sequencing, flow cells are nuclease flushed (Flow Cell Wash Kit EXP-WSH004, Oxford Nanopore Technologies) and reloaded with new nCATS libraries up to three times.

Base calling is done using GUPPY (v3.2.10) and the resulting FASTQ files are aligned to a reference sequence using minimap2(v2.17). Custom reference sequences were constructed for transgene insertion sites using the Mus musculus C57BL/6J reference genome (MM10) and corresponding insert sequence. Alignment results are subject to MapQ score filtering using Samtools (v1.11). Subsequent coverage depth for on-target reads are generated using Samtools (v1.11) and Bedtools (v2.29.2). On-target reads are visualized using the Integrative Genomics Viewer (IGV) and consensus sequences were generated using Medaka (v1.0.3) and annotated in Snapgene (v5.3).

Ethics statement. The Institutional Animal Care and Use Committee of The Jackson Laboratory approved all procedures used in this study and all mice were maintained at The Jackson Laboratory (Bar Harbor, ME,

USA) in strict accordance with all institutional protocols and the Guide for the Care and Use of Laboratory Animals. Data and experiments reported are in accordance with ARRIVE guidelines (https://arriveguidelines.org) where applicable.

Received: 22 October 2021; Accepted: 23 March 2022
Published online: 31 March 2022

References

1. Nelson, A. L., Dhimolea, E. & Reichert, J. M. Development trends for human monoclonal antibody therapeutics. *Nat. Rev. Drug Discovery* **9**, 767–774 (2010).
2. Shultz, L. D., Ishikawa, F. & Greiner, D. L. Humanized mice in translational biomedical research. *Nat. Rev. Immunol.* **7**, 118–130 (2007).
3. Zhu, F., Nair, R.R., Fisher, E.M.C. & Cunningham, T.J. Humanising the mouse genome piece by piece. *Nat. Commun.* **10** (2019).
4. Douam, F. *et al.* Selective expansion of myeloid and NK cells in humanized mice yields human-like vaccine responses. *Nat. Commun.* **9** (2018).
5. Masemann, D., Ludwig, S. & Boergeling, Y. Advances in transgenic mouse models to study infections by human pathogenic viruses. *Int. J. Mol. Sci.* **21**, 9289 (2020).
6. Erwood, S. & Gu, B. Embryo-based large fragment knock-in in mammals: Why, how and what's next. *Genes* **11**(2020).
7. Li, K., Wang, G., Andersen, T., Zhou, P. & Pu, W.T. Optimization of genome engineering approaches with the CRISPR/Cas9 system. *PLoS ONE* **9**(2014).
8. Gurumurthy, C.B. *et al.* Reproducibility of CRISPR-Cas9 methods for generation of conditional mouse alleles: A multi-center evaluation. *Genome Biol.* **20**(2019).
9. Zhang, X., Li, T., Ou, J., Huang, J. & Liang, P. Homology-based repair induced by CRISPR-Cas nucleases in mammalian embryo genome editing. *Protein Cell* (2021).
10. Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Yagle, M. K. & Palmiter, R. D. Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 4438–4442 (1985).
11. Garrick, D., Sutherland, H., Robertson, G. & Whitelaw, E. Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. *Nucleic Acids Res.* **24**, 4902–4909 (1996).
12. Henikoff, S. Conspiracy of silence among repeated transgenes. *BioEssays* **20**, 532–535 (1998).
13. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
14. Goodwin, L. O. *et al.* Large-scale discovery of mouse transgenic integration sites reveals frequent structural variation and insertional mutagenesis. *Genome Res.* **29**, 494–505 (2019).
15. Mehtali, M., LeMeur, M. & Lathe, R. The methylation-free status of a housekeeping transgene is lost at high copy number. *Gene* **91**, 179–184 (1990).
16. Calero-Nieto, F.J., Bert, A.G. & Cockerill, P.N. Transcription-dependent silencing of inducible convergent transgenes in transgenic mice. *Epigenetics and Chromatin* **3**(2010).
17. Smirnov, A. *et al.* DNA barcoding reveals that injected transgenes are predominantly processed by homologous recombination in mouse zygote. *Nucleic Acids Res.* **48**, 719–735 (2020).
18. McBurney, M. W., Mai, T., Yang, X. & Jardine, K. Evidence for repeat-induced gene silencing in cultured mammalian cells: Inactivation of tandem repeats of transfected genes. *Exp. Cell Res.* **274**, 1–8 (2002).
19. Zambrowicz, B. P. *et al.* Disruption of overlapping transcripts in the rosa beta-geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 3789–3794 (1997).
20. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat. Genet.* **21**, 70–71 (1999).
21. Strathdee, D., Ibbotson, H. & Grant, S.G. Expression of transgenes targeted to the Gt(ROSA)26Sor locus is orientation dependent. *PLoS One* **1**, e4 (2006).
22. Giel-Moloney, M., Krause, D. S., Chen, G., Van Etten, R. A. & Leiter, A. B. Ubiquitous and uniform in vivo fluorescence in ROSA26-EGFP BAC transgenic mice. *Genesis* **45**, 83–89 (2007).
23. Chen, C.M., Krohn, J., Bhattacharya, S. & Davies, B. A comparison of exogenous promoter activity at the ROSA26 locus using a PhiC31 integrase mediated cassette exchange approach in mouse es cells. *PLoS ONE* **6**(2011).
24. Barletta, R. G., Kim, D. D., Snapper, S. B., Bloom, B. R. & Jacobs, W. R. Jr. Identification of expression signals of the mycobacteriophages Bxb 1, L1 and TM4 using the Escherichia-Mycobacterium shuttle plasmids pYUB75 and pYUB76 designed to create translational fusions to the lacZ gene. *J. Gen. Microbiol.* **138**, 23–30 (1992).
25. Mediavilla, J. *et al.* Genome organization and characterization of mycobacteriophage Bxb1. *Mol. Microbiol.* **38**, 955–970 (2000).
26. Bai, H. *et al.* Single-molecule analysis reveals the molecular bearing mechanism of DNA strand exchange by a serine recombinase. *Proc. Natl. Acad. Sci. USA.* **108**, 7419–7424 (2011).
27. Russell, J. P., Chang, D. W., Tretiakova, A. & Padidam, M. Phage Bxb1 integrase mediates highly efficient site-specific recombination in mammalian cells. *Biotechniques* **40**, 460–464 (2006).
28. Keravala, A. *et al.* A diversity of serine phage integrases mediate site-specific recombination in mammalian cells. *Mol. Genet. Genomics* **276**, 135–146 (2006).
29. Fogg, P. C. M., Colloms, S., Rosser, S., Stark, M. & Smith, M. C. M. New applications for phage integrases. *J. Mol. Biol.* **426**, 2703–2716 (2014).
30. Olorunniji, F. J., Rosser, S. J. & Stark, W. M. Site-specific recombinases: Molecular machines for the Genetic Revolution. *Biochem. J.* **473**, 673–684 (2016).
31. Ghosh, P., Kim, A. I. & Hatfull, G. F. The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of attP and attB. *Mol. Cell* **12**, 1101–1111 (2003).
32. Ghosh, P., Wasil, L. R. & Hatfull, G. F. Control of phage Bxb1 excision by a novel recombination directionality factor. *PLoS Biol.* **4**, 0964–0974 (2006).
33. Merrick, C. A., Zhao, J. & Rosser, S. J. Serine Integrases: Advancing Synthetic Biology. *ACS Synth. Biol.* **7**, 299–310 (2018).
34. Thomson, J. G. & Ow, D. W. Site-specific recombination systems for the genetic manipulation of eukaryotic genomes. *Genesis* **44**, 465–476 (2006).
35. Xu, Z. *et al.* Accuracy and efficiency define Bxb1 integrase as the best of fifteen candidate serine recombinases for the integration of DNA into the human genome. *BMC Biotechnol.* **13**(2013).
36. Carroll, D. Genome editing by targeted chromosomal mutagenesis. in *Chromosomal Mutagenesis: Second Edition* 1–13 (2014).
37. Chao, G., Travis, C. & Church, G. Measurement of large serine integrase enzymatic characteristics in HEK293 cells reveals variability and influence on downstream reporter expression. *Febs j* (2021).

38. Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).
39. Tsai, S. Q. *et al.* GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015).
40. Jusiak, B. *et al.* Comparison of integrases identifies Bxb1-GA mutant as the most efficient site-specific integrase system in mammalian cells. *ACS Synth. Biol.* **8**, 16–24 (2019).
41. McCray, P. B. Jr. *et al.* Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. *J. Virol.* **81**, 813–821 (2007).
42. Gilpatrick, T. *et al.* Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nat. Biotechnol.* **38**, 433–438 (2020).
43. Chow, K.-H.K. *et al.* Imaging cell lineage with a synthetic digital recording system. *Science* **372**, eabb3099 (2021).
44. Peirce, J.L., Lu, L., Gu, J., Silver, L.M. & Williams, R.W. A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genet.* **5**(2004).
45. Churchill, G. *et al.* The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* **36**, 1133–1137 (2004).
46. Babinet, C. & Cohen-Tannoudji, M. Genome engineering via homologous recombination in mouse embryonic stem (ES) cells: An amazingly versatile tool for the study of mammalian biology. *An Acad. Bras. Cienc* **73**, 365–383 (2001).
47. Brosh, R. *et al.* A versatile platform for locus-scale genome rewriting and verification. *Proc. Natl. Acad. Sci. USA* **118** (2021).
48. Czechanski, A. *et al.* Derivation and characterization of mouse embryonic stem cells from permissive and nonpermissive strains. *Nat. Protoc.* **9**, 559–574 (2014).
49. Taft, R. A., Davisson, M. & Wiles, M. V. Know thy mouse. *Trends Genet.* **22**, 649–653 (2006).
50. Qin, W. *et al.* Efficient CRISPR/cas9-mediated genome editing in mice by zygote electroporation of nuclease. *Genetics* **200**, 423–430 (2015).
51. Wang, W. *et al.* Delivery of Cas9 protein into mouse zygotes through a series of electroporation dramatically increases the efficiency of model creation. *J. Genet. Genomics* **43**, 319–327 (2016).
52. Baer, A. & Bode, J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr. Opin. Biotechnol.* **12**, 473–480 (2001).
53. Seibler, J. & Bode, J. Double-reciprocal crossover mediated by FLP-recombinase: A concept and an assay. *Biochemistry* **36**, 1740–1747 (1997).
54. Wallace, H. A. C. *et al.* Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. *Cell* **128**, 197–209 (2007).
55. Turan, S., Zehe, C., Kuehle, J., Qiao, J. & Bode, J. Recombinase-mediated cassette exchange (RMCE)—A rapidly-expanding toolbox for targeted genomic modifications. *Gene* **515**, 1–27 (2013).
56. Inniss, M.C. *et al.* A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO Cells. *Biotechnol. Bioeng.* (2017).
57. Nottle, M. B. *et al.* Effect of DNA concentration on transgenesis rates in mice and pigs. *Transgenic Res.* **10**, 523–531 (2001).
58. Schedl, A. *et al.* A method for the generation of YAC transgenic mice by pronuclear microinjection. *Nucleic Acids Res.* **21**, 4783–4787 (1993).
59. Montoliu, L., Bock, C. T., Schütz, G. & Zentgraf, H. Visualization of large DNA molecules by electron microscopy with polyamines: Application to the analysis of yeast endogenous and artificial chromosomes. *J. Mol. Biol.* **246**, 486–492 (1995).
60. Ow, D. W. Recombinase-mediated gene stacking as a transformation operating system. *J. Integr. Plant Biol.* **53**, 512–519 (2011).
61. Olorunniji, F. J. *et al.* Multipart DNA assembly using site-specific recombinases from the large serine integrase family. *Methods Mol. Biol.* **1642**, 303–323 (2017).
62. Sclimenti, C. R., Thyagarajan, B. & Calos, M. P. Directed evolution of a recombinase for improved genomic integration at a native human sequence. *Nucleic Acids Res.* **29**, 5044–5051 (2001).
63. Li, H., Sharp, R., Rutherford, K., Gupta, K. & Van Duyne, G. D. Serine integrase attP binding and specificity. *J. Mol. Biol.* **430**, 4401–4418 (2018).
64. Low, B. E., Kutny, P. M. & Wiles, M. V. Simple, efficient CRISPR-cas9-mediated gene editing in mice: Strategies and methods. *Methods Mol. Biol.* **1438**, 19–53 (2016).
65. Hermann, M. *et al.* Binary recombinase systems for high-resolution conditional mutagenesis. *Nucleic Acids Res.* **42**, 3894–3907 (2014).

Acknowledgements

We thank Cindy Avery and Todd Nason for excellent maintenance of mouse colonies, Richard S. Maser for building the Krt18-ACE2 construct, Peter Kutny and group at Genetic Engineering Technologies for MIJs. We also thank various groups who assisted us, especially the Reinholdt laboratory (Dr David Bergstrom and Tiffany Leidy-Davis at JAX), and the Steve Murray laboratory (Dr. Kathy Snow at JAX) for providing constructs. We are grateful to the Genome Technologies (GT) group at JAX for Sanger and Nanopore sequencing services. We thank Cindy Avery for her assistance in animal maintenance, Melissa Berry for her careful reading of the manuscript and assistance in generating the mouse strain nomenclature. This work was funded by National Institutes of Health grants P30CA034196, R21 OD027052, R24 OD 021325 (Drs. Reinholdt and Bergstrom), U42 OD026635 (Drs the Steve Murray and Kathy Snow). The Krt18-ACE2 related work was partially funded by the Rosenthal laboratory at JAX. We also gratefully acknowledge funding from JAX.

Author contributions

M.V.W. and B.E.L. originated and conceptualized the project. B.E.L., V.H., and M.V.W. designed the experiments. B.E.L prepared the reagents for MIJs and performed the DNA analysis experiments. S.L. developed and prepared nCATS. B.E.L., S.L., V.H., and M.V.W. analyzed these data and wrote the manuscript. All the authors contributed to data interpretation, critical review and approval of the manuscript.

Funding

This article was funded by Jackson Laboratory, National Institute for Health (Grant nos. P30CA034196, R21 OD027052, R24 OD 021325, U42 OD026635).

Competing interests

MVW and BEL are co-inventors on a patent application "High Frequency Targeted Animal Transgenesis", International Application No. PCT/US2020/054745, published as WO 2021/072049. This application covers the use

of dual heterologous att sites to provide efficient and precise integration of donor DNAs into organisms. VH and SL have no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-09445-w.

Correspondence and requests for materials should be addressed to M.V.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.