Efficient genome-wide first-generation phenotypic screening system in mice using the *piggyBac* transposon

Hao Chang[a,b,c,1], Yukun Pan[a,b,1,2], Sean Landrette[a,b], Sheng Ding[a,b], Dong Yang[a,c], Lufang Liu[a,b], Lei Tian[a,b], Hongyan Chai[d], Peining Li[d], Da-Ming Li[a,b,c], and Tian Xu[a,b,c,3]

[a]Department of Genetics, Boyer Center for Molecular Medicine, Yale University School of Medicine, New Haven, CT 06536; [b]Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT 06536; [c]School of Life Sciences, Westlake Institute for Advanced Study, Westlake University, 310024 Hangzhou, China; and [d]Laboratory of Clinical Cytogenetics and Genomics, Department of Genetics, Yale School of Medicine, New Haven, CT 06536

Edited by Norbert Perrimon, Howard Hughes Medical Institute, Boston, MA, and approved August 1, 2019 (received for review April 18, 2019)

Genome-wide phenotypic screens provide an unbiased way to identify genes involved in particular biological traits, and have been widely used in lower model organisms. However, cost and time have limited the utility of such screens to address biological and disease questions in mammals. Here we report a highly efficient *piggyBac* (*PB*) transposon-based first-generation (F1) dominant screening system in mice that enables an individual investigator to conduct a genome-wide phenotypic screen within a year with fewer than 300 cages. The *PB* screening system uses visually trackable transposons to induce both gain- and loss-of-function mutations and generates genome-wide distributed new insertions in more than 55% of F1 progeny. Using this system, we successfully conducted a pilot F1 screen and identified 5 growth retardation mutations. One of these mutants, a Six1/4$^{PB/+}$ mutant, revealed a role in milk intake behavior. The mutant animals exhibit abnormalities in nipple recognition and milk ingestion, as well as developmental defects in cranial nerves V, IX, and X. This *PB* F1 screening system offers individual laboratories unprecedented opportunities to conduct affordable genome-wide phenotypic screens for deciphering the genetic basis of mammalian biology and disease pathogenesis.

piggyBac transposon | genome-wide screen | growth retardation | *six1* | *six4*

Genome-wide phenotypic screens in mammals are critical for identifying the genes and pathways that are essential for human health and altered in diseases, but costs associated with the generation of mutant animals and mapping of mutations limit the practicality of current approaches. Traditionally, mouse mutants have been generated from embryonic stem (ES) cells with specific genes knocked out by homologous recombination (1, 2). CRISPR/Cas9 is a faster and more efficient technology to generate mutations in individual genes (3). However, both approaches remain prohibitively expensive for generating and maintaining genome-wide mouse mutant libraries for phenotypic screening. In contrast to gene-by-gene approaches, chemical mutagens, such as ENU, randomly generate point mutations and are highly potent in mice (4). While recessive screens require a costly 3 generations of breeding, first-generation (F1) dominant screens directly assay the progeny of mutagenized animals, reducing cost and time. Genome-scale ENU F1 screens have been conducted for gross morphological, neurological, and immunological defects. Key regulatory genes involved in colorectal cancer and circadian rhythm have also been identified in dominant screens in mice (5–8). However, ENU F1 screens are still limited by the costs associated with the back-crosses and sequencing required for mapping mutations. Alternative approaches that make genome-wide screens affordable to individual investigators are highly desirable.

Transposon insertional mutagenesis is an attractive alternative approach for genome-wide phenotypic screens in mammals. The DNA "cut and paste" *piggyBac* (*PB*) transposon from the cabbage looper moth *Trichoplusia ni*, which favorably inserts at TTAA tetranucleotide target sites, has been modified to generate a binary transposon system for mammalian cells and mammals. The system contains a nonautonomous *PB* transposon cassette to deliver the exogenous genes of interest flanked by the *PB* inverted repeat sequences, and a transgene expressing the *PB* transposase enzyme (*PBase*) for inducing transposition in the germline. Transposition occurs when *PBase* binds to the inverted repeat sequences of the *PB* transposon, initiating excision and reintegration into another locus. The binary *PB* system can efficiently transpose in mice and human cells (9). Somatic mutagenesis using multiple copies of Sleeping beauty (*SB*) or *PB* transposons to generate dominant mutations have facilitated cancer gene discovery (10–13), highlighting the potential to modify this strategy for germline mutagenesis. Indeed, an *SB* germline system has been developed for region-specific saturation screens (14). A comparable system for genome-wide mutagenesis is needed. The *PB* transposon has been shown to have higher jumping efficiency and less-severe local hopping than *SB* in cultured mouse ES cells (15), thus holding promise for affordable genome-wide in vivo germline mutagenesis screens in mice.

Significance

Genome-wide, phenotype-driven mutagenesis in animal models could provide an unbiased way to decode a gene's functions and its role in diseases. Here we have generated a *piggyBac* (*PB*) transposon-based first-generation (F1) dominant screening system in mice, which provides unprecedented opportunities to conduct a highly efficient and affordable genome-wide phenotypic screen for an individual investigator in a single laboratory. Using this system, we carried out an F1 dominant screen for growth retardation and discovered 5 isolated mutants that carry transposon insertions hitting the genes *Rin2*, *Rbm39*, *Mll*, *Zeb2*, and *Six1/4*. The Six1/4$^{PB/+}$ mutant animals exhibit abnormalities in nipple recognition and milk ingestion during the breastfeeding period and also exhibit cranial nerve defects.

Author contributions: H. Chang, Y.P., D.-M.L., and T.X. designed research; H. Chang, Y.P., S.D., L.L., L.T., and D.-M.L. performed research; H. Chang, Y.P., S.L., H. Chai, P.L., and D.-M.L. contributed new reagents/analytic tools; H. Chang, Y.P., D.Y., and T.X. analyzed data; and H. Chang, Y.P., and T.X. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]H. Chang and Y.P. contributed equally to this work.

[2]Present address: Yeda Research Center for Model Organisms, Taizhou, 318000 Zhejiang, China.

[3]To whom correspondence may be addressed. Email: xutian@westlake.edu.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1906354116/-/DCSupplemental.

Published online August 26, 2019.

Here, we developed a highly efficient F1 dominant screening system in mice. The new *PB* system utilizes multiple copies of visually trackable transposons to generate genome-wide mutations in 55.2% of F1 progeny. Therefore, we estimate that this system could allow individual laboratories to perform a genome-wide phenotypic F1 screen in mice within a year and with 300 cages in total. Using this system, we carried out an F1 dominant screen for growth retardation and identified both gain- and loss-of-function alleles. Both intrauterine growth retardation (IUGR) and postnatal growth retardation (PGR) phenotypes were observed in the 5 isolated mutants that carry transposon insertions hitting the endocytosis regulator *Rin2*, the steroid hormone receptor coactivator *Rbm39*, the histone methyltransferase *Mll*, and the homeobox transcriptional factors *Zeb2* and *Six1/4*. Altogether, our work demonstrates the applicability of our system for cost and time-effective phenotypic screens.

Results

Development of a Highly Efficient F1 Mutagenesis Screening System in Mice. The F1 mutagenesis system was designed as a binary *PB* system (Fig. 1): A founder line carrying nonautonomous *PB* transposon constructs and a Jump-Starter line carrying a *PBase* transgene under the control of a hybrid *actin* promoter, which exhibits better *PB* remobilization efficiency than the spermiogenesis-specific *protamine 1* (*prm1*) promoter (~8-fold higher) (*SI Appendix*, Fig. S1). To deploy *PB* for optimized F1 germline screens, we developed a mutagenic *PB* construct that includes 4 important features: 1) A strong promoter for gene overexpression that can be conditionally regulated; 2) a terminator for disrupting transcription; 3) a Katushka red fluorescent protein (Kat) transgene (16); and 4) a codon-optimized luciferase gene (*luc2*) under the control of a Ubiquitin C (UBC) promoter (Fig. 1*A*) (12), which is disrupted by the insertion of the *PB* construct. When *PBase* is expressed to mobilize *PB* for mutagenesis, the *luc2* gene activity is restored as the *PB* insertion is excised (*SI Appendix*, Fig. S2*A*). This construct, named *PB[mut-Kat]*, can induce both gain- and loss-of-function mutations, visually track mutants, and report mutagenesis activity.

High mutagenesis efficiency is required for a cost-effective screening system. Mobilizing a single *PB* transposon in the mouse germline is not sufficient to support a genome-wide screen (9). Increasing the copy number of the mutagenic transposon could improve efficiency, while too many copies of the transposon could cause lethality and tumorigenesis (11, 12). We therefore generated six different mouse lines with copy number of the *PB[mut-Kat]* transposon ranging from 1 to 22 (*SI Appendix*, Table S1). In order to identify a transgenic line able to induce mutations in the germline without causing too high a level of lethality or tumorigenesis, we assayed the mutagenesis activity (by *luc2* expression), lethality, and tumorigenesis in mice carrying both the transposon and the transposase (*PBase*). A transgenic line carrying 10 copies of the mutagenic transposons (*PB[mut-Kat]10,* determined by

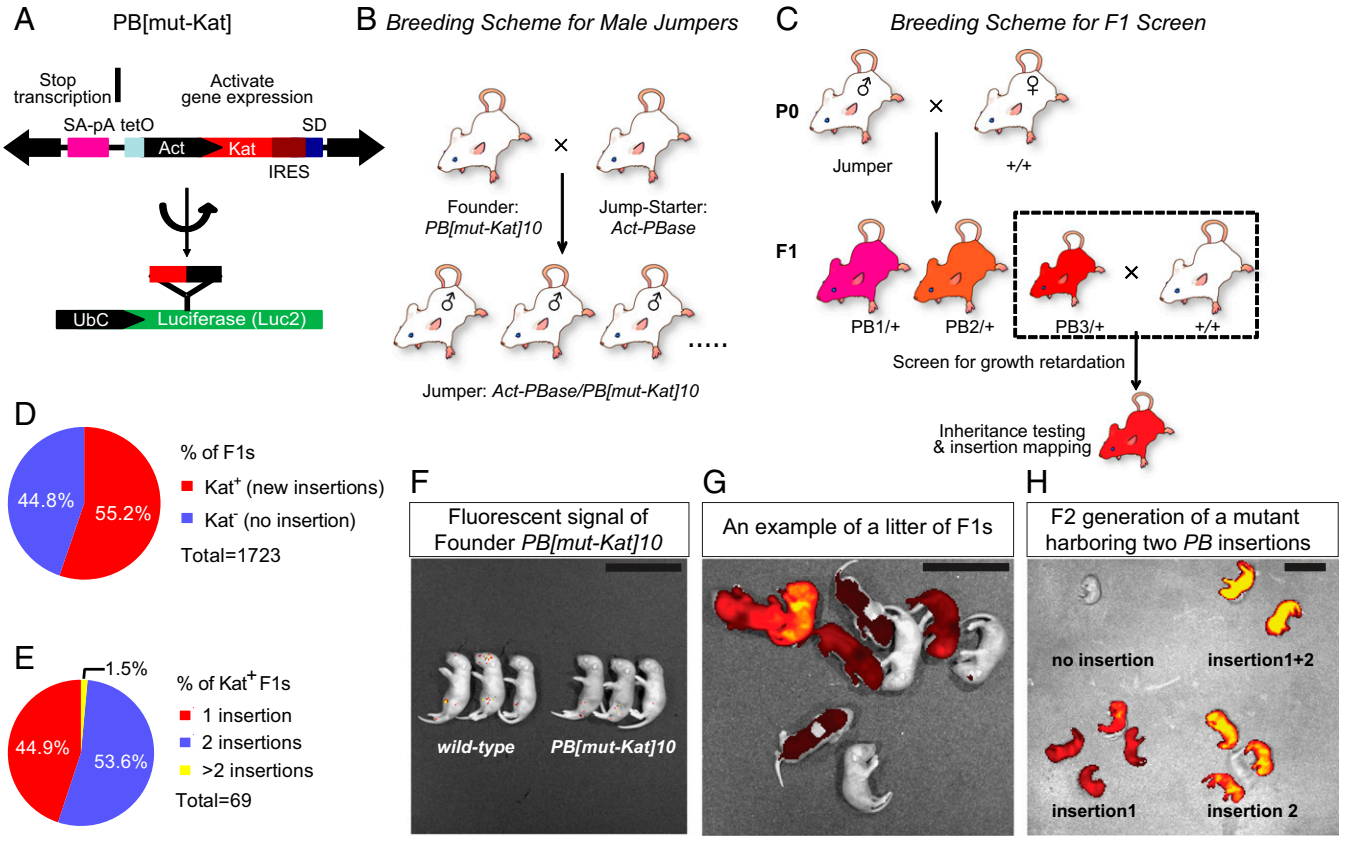

Fig. 1. An efficient and visually trackable *piggyBac* transposon system for an F1 dominant screen in mice. (*A*) *PB[mut-Kat]* construct. The transposon contains 2 parts: 1) An overexpression cassette including tet operon (tetO), CAG promoter, Katushka red fluorescent protein (Kat), IRES, and a splice donor (SD); and 2) a stop cassette including a splice acceptor (SA) and a poly-A signal (PA). The transposon was cloned into an internal TTAA site (the *PB* integration site) of the luciferase reporter (*luc2*) in the opposite orientation. (*B* and *C*) The breeding scheme for the F1 screen. (*D*) Percentage of the F1 animals that are Kat$^+$. (*E*) Percentage of F1 animals bearing 1 copy, 2 copies, or more than 2 copies of Kat$^+$ insertions. (*F*) Red fluorescent signals of a litter of wild-type and a litter of *PB[mut-Kat]10* founders. They have similar autofluorescent signals. (Scale bar, 2 cm.) (*G*) An example of a litter of F1s carrying Kat-labeled insertions. Individual Kat$^+$ F1s carry independent insertions. Due to position effects, each Kat$^+$ F1 displays distinct signal intensity. (Scale bar, 2 cm.) (*H*) An example of a mutant harboring 2 Kat$^+$ insertions (insertion 1 and 2), which are segregated in the F2 generation. (Scale bar, 2 cm.)

quantitative PCR) was selected due to having the highest *luc2* activity and a lack of lethality and tumorigenesis (*SI Appendix*, Figs. S2*B* and S3).

We next scored germline insertion mutagenesis efficiency by quantifying the percentage of Kat$^+$ F1 progeny from the *Act-PBase*/*PB[mut-Kat]10* males (jumpers) (Fig. 1 *B* and *C*). The *Katushka* gene found in the concatemer regions of the founder mice are silent (Fig. 1*F*), possibly due to an epigenetic mechanism for the maintenance of the repressed state provided by the plasmid backbone (14, 17, 18). The jumpers produced an efficiency of 55.2% (952 of 1,723) (Fig. 1 *D* and *G*). Moreover, 55.1% (38 of 69) of Kat$^+$ F1s carry more than 1 insertion, which can be segregated in the F2 generation (Fig. 1 *E* and *H*). Therefore, each Kat$^+$ F1 mouse carries an average of 1.56 insertions, and 8.2 insertions can be screened per litter, which is 10 times higher than the efficiency achieved by *SB* germline mutagenesis (14, 17). However, although a ubiquitous promoter has been used to control the expression of *PBase* in jumpers, somatic hopping has little influence on this *PB* F1 screen system for 2 reasons. First, about half of the *PB*$^+$ F1 mutant mice do not carry the *Act-PBase* transgene; therefore, they will not exhibit somatic jumping. Second, for *PB*$^+$ F1 mutant mice with the *Act-PBase* transgene, each *PB*$^+$ pup would acquire 1.56 copies of *PB* on average, such that the somatic jumping activity is significantly lower than in jumpers carrying 10 copies of *PB* and would most likely not affect the detection of phenotypes in F1 animals. Furthermore, the *Act-PBase* transgene would be segregated in the heritability test in the next generation (F2).

The germline insertion mutagenesis efficiency does not undergo detectable change as the jumper's age (*SI Appendix*, Fig. S4), indicating that mutations do not accumulate in spermatogonial stem cells. We applied FISH analysis of *PB[mut-Kat]10* concatemer (*PB*-FISH) to isolated single cells from adult testes of both *PB[mut-Kat]10* homozygous founders and *Act-PBase*/*PB[mut-Kat]10* jumpers (Fig. 2). The *PB*-FISH signal was observed inside the nucleoplasm and used to evaluate the *PB* exclusion and remobilization efficiency. In addition, anti-SALL4A immunostaining was applied to the same cells to label spermatogonial stem cells (and undifferentiated spermatogonia). In SALL4A$^+$ cells from either *PB[mut-Kat]10* founders or *Act-PBase*/*PB[mut-Kat]10* jumpers, the intensities of the *PB*-FISH signal were similar (Fig. 2 *A*–*H* and *O*). In contrast, *PB*-FISH signals were detected in mature sperm cells from *PB[mut-Kat]10* founders, but not from *Act-PBase*/*PB[mut-Kat]10* jumpers (Fig. 2 *I*–*N* and *P*). These results indicate that the *PB* transposition activity is low in the spermatogonial stem cells and provides an explanation for the diverse *PB* insertion distribution observed in sperm and F1 offspring, a new perspective that is important for the design and manipulation of *PB* germline mutagenesis.

Genome-Wide Mutagenesis via *PB* Transposition from a Concatemer. Local hopping is a concern for transposon insertional mutagenesis. When *SB* transposons are mobilized from a chromosomal concatemer, 63 to 75% of the new insertions are located on the donor chromosome, with the majority clustered near the concatemer (14, 17). To characterize *PB* transposition from a concatemer, we mapped the genome distribution of *PB* insertions in jumper germ cells. Of the 572 new insertions, 25.9% (148 of 572) mapped to chromosome 6 and 17.0% (97 of 572) were located in the same 8.5-Mb region (Fig. 3*A*). Karyotype analysis costained with the *PB*-FISH probe and a mouse chromosome 6 painting probe in *PB[mut-Kat]10* heterozygous mouse revealed the concatemer in the region of chromosome 6E1-3, consistent with the local hopping region (Fig. 3*A* and *SI Appendix*, Fig. S5). The remaining 74.1% of insertions were distributed randomly across the genome (Fig. 3*A*, inside). Thus, our *PB* system exhibits a 3-fold decrease in local hopping compared to *SB* germline mutagenesis (14). Furthermore, of the 572 insertions, 45.5% (260) sit in coding genes, with 1.8% (10) in exons and 43.7% (250) in introns (Fig. 3*A*, inside).

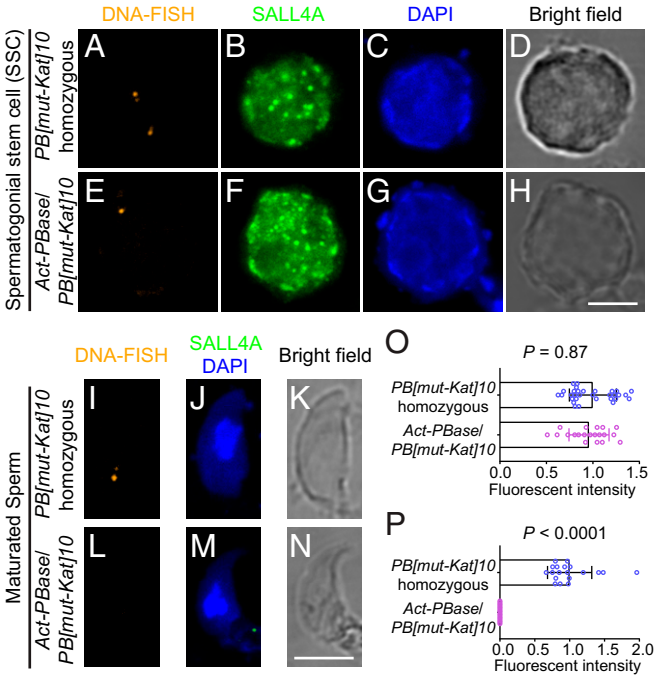

Fig. 2. Dynamics of *PB* concatemer FISH signals in spermatogonial stem cells and mature sperm cells from either *PB[mut-Kat]10* homozygous or *Act-PBase/PB[mut-Kat]10* animals. (*A*–*H*) Immunofluorescence (SALL4A$^+$, green) and PB-FISH (*PB* concatemer, orange) on spermatogonial stem cell nuclei from *PB[mut-Kat]10* homozygous animals (*A*–*D*) and *Act-PBase/PB[mut-Kat]10* animals (*E*–*H*). (Scale bar, 10 μm.) (*I*–*N*) PB-FISH (*PB* concatemer, orange) on mature sperm nuclei from *PB[mut-Kat]10* homozygous lines (*I*–*K*) and *Act-PBase/PB[mut-Kat]10* lines (*L*–*N*). (Scale bar, 10 μm.) (*O* and *P*) Statistical analysis of *PB* concatemer accumulation puncta in spermatogonial stem cells (*O*) and mature sperms (*P*) from either *PB[mut-Kat]10* homozygous or *Act-PBase/PB[mut-Kat]10*.

To characterize *PB*-based gene-targeting behavior, we carefully analyzed the 260 coding genes hit by *PB* in the germline. Based on the position and orientation of the transposon in the genes, we found that in 114 insertions, *PB* was inserted in the same orientation as the transcription unit and was predicted to cause gain-of-function mutations. In 146 insertions, *PB* was inserted in the opposite orientation of the transcription unit and was predicted to cause loss-of-function (hypermorphic) mutations (*SI Appendix*, Fig. S6*A*). While potential gain-of-function mutations were fewer than loss-of-function mutations, statistical analysis did not reveal a significant difference ($P = 0.16$), indicating that the majority of the grain-of-function mutations were recovered from the screen. Next, we analyzed the expression of the 260 genes hit by the transposon using expression information from the ENCODE project (19). We found that 250 of the genes (96.15%, fragments per kilobase of transcript per million mapped reads [FPKM] ≥ 0.1) were expressed in the mouse germline, while 10 of them (3.84%) were nonexpressed genes (Fig. 3*B*). Since the percentage of coding genes expressed in the mouse germline is 79.84%, there is an enrichment of *PB* insertions in expressed genes (Fig. 3*B*). Consistent with this, we plotted the distribution of gene expression for the 260 *PB* hit genes (the FPKM values) against the total coding genes and found that nonexpressed genes had a lower probability of being hit by *PB* than expressed genes (Fig. 3*F*). Interestingly, when we removed the nonexpressed genes and only considered expressed genes, we found that the distribution of *PB* hit genes was similar to that of total expressed genes (Fig. 3*G*). This reveals that while gene expression is required for *PB* insertion, it does not correlate with the level of gene expression.

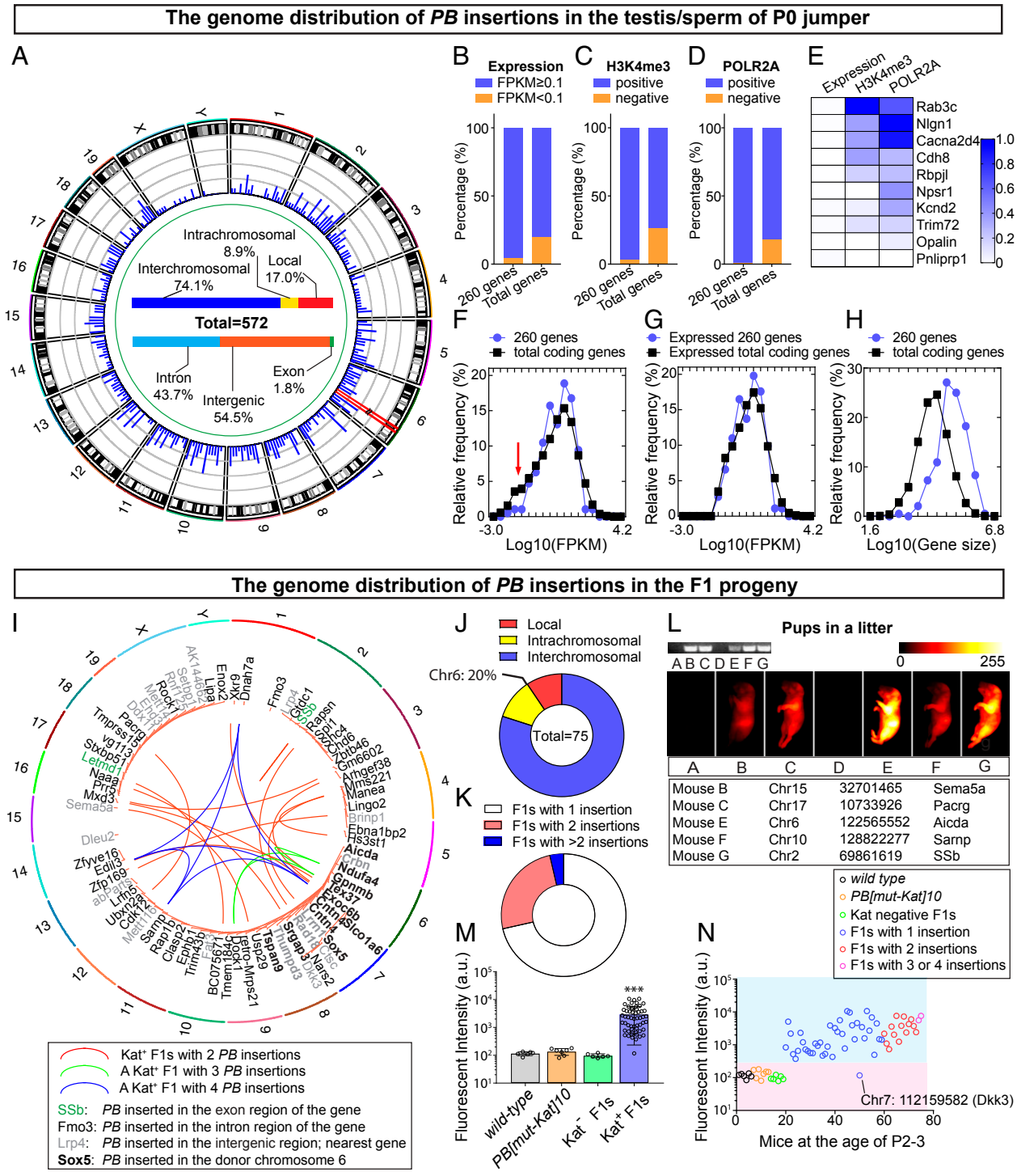

Fig. 3. Genome-wide distribution of *PB* insertion in the testis/sperm of the P0 jumper and in the F1 progeny. (*A*) A majority of the *PB* insertions mapped from testis and sperm samples of P0 jumpers exhibited genome-wide distribution. The red bar indicates insertions in the local region of the concatemer. The percentage of insertions located in introns, exons, or intergenic regions of transcriptional units are indicated as well. (*B*) The ratios of the expressed and nonexpressed genes for the 260 *PB* hits and of 21,528 coding genes ($P < 0.05$). (*C* and *D*) The ratios of H3K4me3 or POLR2A$^+$ and POLR2A$^-$ genes for the 260 *PB* hits and the 21,528 coding genes ($P < 0.05$). (*E*) Nine of 10 of the nonexpressed genes hit by *PB* have H3K4me3 modification and/or POLR2A binding sites. The values of H3K4me3 and POLR2A were normalized. (*F*) Distributions of the 260 *PB* hit genes (blue) and the total 21,528 coding genes (black) versus gene expression FPKM values (ENCODE). The red arrow marks the nonexpressed gene region. (*G*) Distributions of the expressed genes (FPKM ≥ 0.1) for the 260 *PB* hits (blue) and the 21,528 coding genes (black). (*H*) Distributions of the 260 *PB* hits (blue) and the 21,528 coding genes (black) versus gene length. (*I*) Genome distribution of insertion sites and the neighbor genes mapped from 56 *PB*$^+$ F1s, validated by genotyping PCR. The percentage of insertions located in introns, exons, or intergenic regions of transcriptional units are in *SI Appendix*, Figs. S6*B*. (*J*) Percentage of insertions located in different chromosomes. (*K*) Percentage of F1s with a single insertion, 2 insertions, and 3 to 4 insertions. (*L*) An example of a litter of F1s with both Kat$^+$ and Kat$^-$ pups, and the PB insertion sites of the Kat$^+$ pups. The table below lists the details of *PB* insertion locations and their neighboring genes. (*M* and *N*) Fluorescent distributions of wild-type pups, *PB[mut-Kat]10* founder pups, Kat$^-$ (Kat nonexpressing) F1s, and Kat$^+$ (Kat expressing) F1s at ages P2 to P3. Of all 56 Kat$^+$ F1s, only 1 was found with a weak Katushka fluorescent intensity, which could not be distinguished from Kat$^-$ F1s. ***$P < 0.001$.

On the other hand, when the insertion distribution is plotted against the length of the genes, the distribution of 260 *PB* hit genes has a clear shift toward longer genes (Fig. 3*H*), indicating that the preference of *PB* insertion correlates with the gene size or the frequency of the PB target sequence, TTAA. A possible explanation for these observations is that *PB* transposition might require an open chromatin region (OCR) in the genome in order to recognize its target site, TTAA. By analyzing 2 OCR-associated biomarkers, H3K4me3 modification and the POLR2A binding sites from the ENCODE project, we found that 259 of the 260 *PB* hit genes were located in an OCR (252 of 260 for H3K4me3 and 258 of 260 for POLR2A) (Fig. 3 *C* and *D*). This also provides an explanation for the *PB* insertions found in nonexpressed genes in the germline, since 9 of the 10 nonexpressed genes were associated with open chromatin biomarkers (Fig. 3*E*). This analysis predicts that genes located in OCRs in the germline (at least 89.10% of coding genes, or 19,182 of 21,528) can be targeted by *PB*.

Genome-Wide Mutagenesis via *PB* Transposition in the F1 Progeny. To examine whether a comparable genome-wide distribution of *PB* insertions occurred in the F1 generation, we next bred and identified 56 F1s with *PB* insertions using genotyping PCR (*Methods*). Then, we mapped them by sequencing and discovered 40 F1s with single insertions, 14 F1s with 2 insertions, 1 F1 with 3 insertions, and 1 F1 with 4 insertions (Fig. 3 *I*–*K*). Among the 75 insertions, 80.0% (60 of 75) of them were distributed randomly across the genome, while 20.0% (15 of 75) mapped to chromosome 6, where the *PB[mut-Kat]10* concatemer is located (Fig. 3*J*). Furthermore, 42.6% of them (32 of 75) resided in intergenic regions of the genome, and 57.4% of them (43 of 75) resided in coding genes, with 2.7% in exons and 54.7% in introns, which is quite similar to that in testis/sperm postnatal day (P)0 jumpers (*SI Appendix*, Fig. S6*B*). These results confirm a genome-wide and nonbiased distribution of *PB* insertions in the F1s. Finally, all but 1 of 56 F1s exhibited positive Katushka fluorescent signals, which were at least 2-fold higher than the autofluorescent signals from wild-type, *PB[mut-Kat]10* founders or F1s with no new *PB* insertions (Fig. 3 *L*–*N* and *SI Appendix*, Fig. S6*C*). These results indicate that the Katushka cassette and detection of fluorescence in neonatal mice using a handheld flashlight is a robust and convenient approach for detecting animals with new insertions.

A Pilot *PB* F1 Dominant Screen for Growth Retardation. We have previously used both ES-based knockout technology and the *PB*-based F3 recessive screening system in mice to study mechanisms underlying body size control (20, 21). To evaluate the power of our F1 screening system, we conducted a pilot screen for a growth retardation phenotype, an important phenotype that serves as an indicator of many systematic defects: 2,036 F1 animals were screened before weaning age and compared to their wild-type siblings from the same litter. Of these, 18 candidates were identified and tested for heritability in the next generations (F2 and F3). Five heritable mutants with a growth retardation phenotype were isolated and mapped using the Splinkerette PCR method (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S7). In 2 of the mutants, which carry insertions in the *Rab5*-binding protein *Rin2* (22) and the zinc finger E-box binding homeobox transcriptional factor *Zeb2* (23) genes, the transposons reside in the early introns in the coding direction, causing overexpression of the full-length genes (*SI Appendix*, Fig. S7). The integration patterns of splicing for transcripts of the 2 gain-of-function mutants were verified by allele-specific PCR and sequencing (*SI Appendix*, Fig. S8). The causative nature of the loss-of-function mutation in *Rin2* was further confirmed by phenocopying growth retardation using an independent conditional overexpression transgenic mouse model (*LSL-Rin2*) (*SI Appendix*, Fig. S9). In the other 3 mutants, which carry insertions in the steroid hormone receptor coactivator *Rbm39* (24), the histone methyltransferase *Mll* (25), and sine oculis homeobox transcription factors *Six1* and *Six4* genes (26–28), the transposons inserted in the intron, the 3′ UTR, and the upstream regulatory region, respectively, in the opposite orientation. These *PB* insertion sites decreased the expression levels of the *Rbm39* gene in $Rbm39^{PB/+}$ mice, the *Mll* gene in $Mll^{PB/+}$ mice, and both *Six1 and Six4* genes in $Six1/4^{PB/+}$ mice (*SI Appendix*, Figs. S7 and S10). Thus, our system has the power to generate both gain-of-function and loss-of-function mutations.

Growth retardation, which includes IUGR and PGR, is a very common syndrome with a broad influence on child development from maternal pregnancy until childhood and adolescence. IUGR affects 3 to 10% of first-time pregnancies and increases infants' perinatal mortality rates by 4 to 8 times, while PGR is correlated with many serious diseases, such as Costello syndrome and human dwarfism. $Rbm39^{PB/+}$ mutants exhibit IUGR, but the mutants recover gradually after birth (Fig. 4*D*). $Rin2^{PB/+}$ and $Mll^{PB/+}$ mutants exhibit IUGR and PGR, remaining small throughout postnatal development (Fig. 4 *C* and *E*). Finally, $Zeb2^{PB/+}$ and $Six1/4^{PB/+}$ mutants are born a normal size and exhibit PGR only (Fig. 4 *F* and *G*). The IUGR and PGR phenotypes have been previously reported in heterozygous *Mll* knockout mice, and Wiedemann–Steriner syndrome human patients carry de novo *MLL* mutations (25, 29), while the other 4 genes have not been previously reported to affect growth. Therefore, our F1 system can indeed efficiently identify novel genes involved in particular biological traits.

Defects in Milk Intake Behavior and Cranial Nerve Development in $Six1/4^{PB/+}$ Mutants. We were especially interested in the growth retardation phenotype of the $Six1/4^{PB/+}$ mutants, because growth reduction occurred during the period pups relied on breastfeeding for growth and recovered as the pups weaned (Fig. 4*G*), suggesting that the problem may be in breastfeeding itself. The *PB* hit a noncoding region ∼70 kb upstream of the *Six1* coding region and 550 bp upstream of the *Six4* coding region in the opposite orientation of the transcriptional units, causing the down-regulation of both genes. This *PB* insertion locus is associated with Branchio-oto-renal (BOR) syndrome (30), and BOR syndrome patients were previously reported to have feeding difficulties (31). Indeed, $Six1/4^{PB/+}$ mutants contained less milk in their stomachs 6 h after birth compared to control animals (Fig. 5*A*). These results indicate that $Six1/4^{PB/+}$ mutants have breastfeeding defects. There are a number of possible reasons for this. For example, olfactory/mechanical sensory and craniofacial defects in mouse mutants are known to be associated with abnormal feeding (32). Other milk intake behaviors, such as sibling competition, milk ingestion, and hunger/satiety response, could also have an effect, but have not been extensively studied. To determine the cause for the observed abnormal feeding, we studied each of these possible contributions. First, we tested the ability of the mutants to compete for nipples. Unlike piglets (33), mouse pups do not display a "teat order." The $Six1/4^{PB/+}$ mutants remained growth-delayed even when competition was lowered by reducing the litter size to 4, suggesting that sibling rivalry is not the cause of the observed feeding difficulties (*SI Appendix*, Fig. S11*A*). The $Six1/4^{PB/+}$ mutants display a longer latency for nipple attachment than controls (Fig. 5 *D*–*F* and Movie S1) and intake milk more slowly during 15-min refeeding after starvation (Fig. 5 *B* and *C* and *SI Appendix*, Fig. S11*B*). Latching on and tongue movement were normal, leaving ingestion defects as the final potential explanation for the slow intake speed. To directly assess ingestion, we examined the stomach contents 1 min after the administration of a charcoal suspension inside the mouths; this assay revealed that the $Six1/4^{PB/+}$ mutants have ingestion difficulties (Fig. 5 *G* and *H*). Finally, the $Six1/4^{PB/+}$ mutants exhibit normal nutritive suckling, which is characterized by the excitement of deprived pups upon nipple attachment, suggesting a normal hunger/satiety response (Movie S2). In sum, the $Six1/4^{PB/+}$

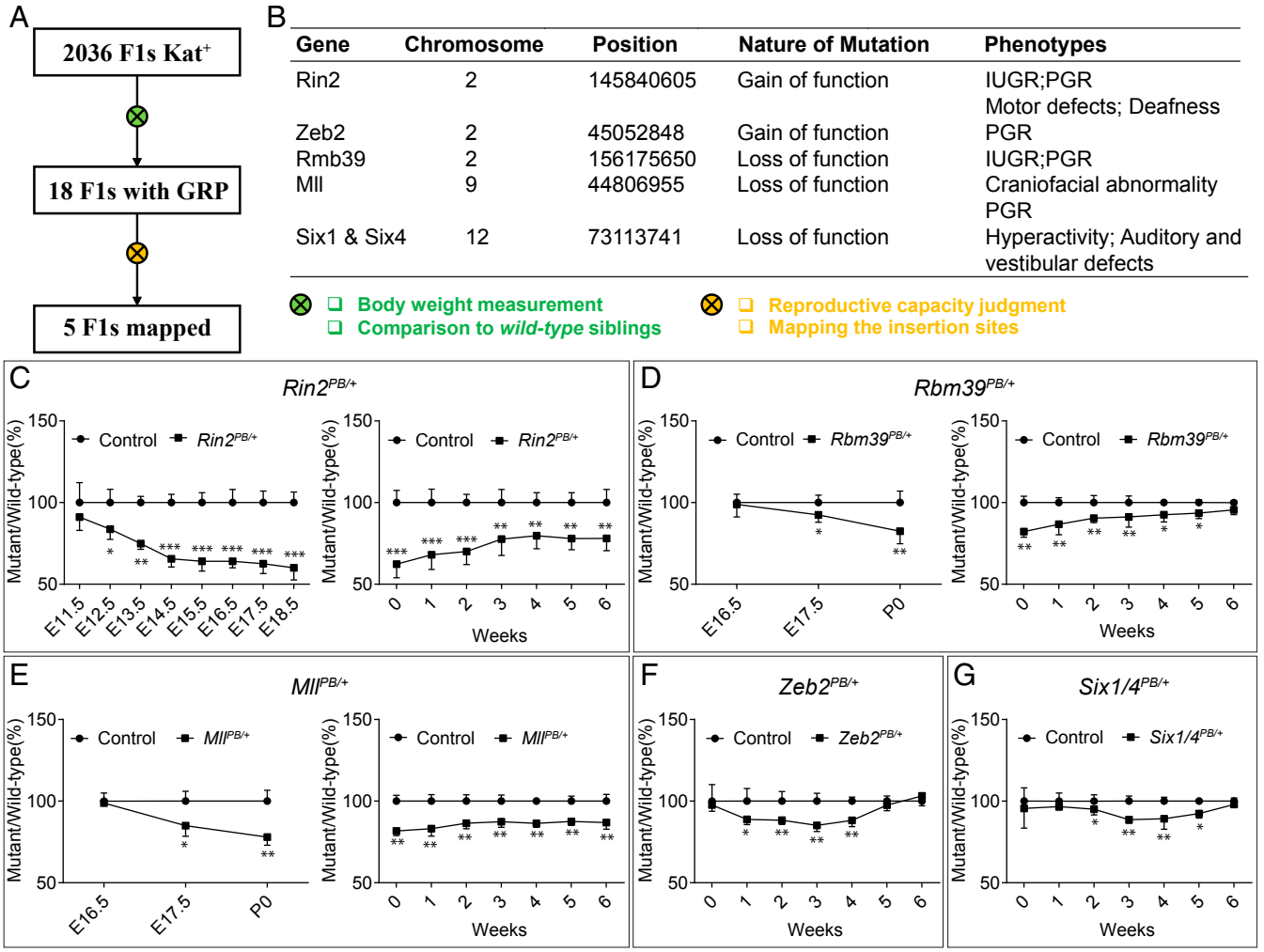

Fig. 4. Pilot screen of 2036 F1 mice identifies 5 different genes that affect growth. (*A*) Experimental outline of growth retardation screen. In this screen, 2,036 Kat$^+$ F1s were screened before weaning. Eighteen F1 mice with growth retardation phenotype (GRP) were isolated. Of these, 5 were validated to be heritable and finely mapped in the genome. (*B*) List of mutants isolated from the screen with *PB* insertion sites and growth retardation phenotypes. Three mutants exhibit both IUGR and PGR. Two mutants exhibit only PGR. Meanwhile, 3 mutants also exhibit other growth retardation-unrelated phenotypes or behaviors, including motor defects, deafness, craniofacial abnormality, hyperactivity, auditory loss, and vestibular defects (*SI Appendix*, Fig. S12). (*C–E*) Intrauterine growth and postnatal growth curves of $Rin2^{PB/+}$, $Rbm39^{PB/+}$, and $Mll^{PB/+}$ mutants ($n = 10$, 2-way ANOVA followed by post hoc Sidak's test). (*F* and *G*) Postnatal growth curves of $Rin2^{PB/+}$ and $Zeb^{PB/+}$ mutants ($n = 10$, 2-way ANOVA with repeated measurement followed by post hoc Sidak's test). Error bar: SD. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

mutants exhibit diminished milk intake, which is associated with defects in nipple recognition and milk ingestion.

We wondered if we could pinpoint the nervous system defect underlying this altered feeding behavior. The trigeminal nerve (cranial nerve [CN] V) is important for facial sensation, while the glossopharyngeal and vagus nerves (CN IX and X) are involved in swallowing and ingestion (34); these nerves are therefore good candidates for explaining the altered feeding behavior. To systematically assess CN axons, we performed antineurofilament immunostaining in the embryonic day (E)10.5 embryos of wild-type and $Six1/4^{PB/+}$ mutants. The lengths of the maxillary branches of CN V were diminished (Fig. 5 *J–L*) and the sensory ganglia of CN IX/X were fused (Fig. 5 *M–O*), suggesting developmental defects in these nerves. Growth retardation, milk intake difficulties, and CN defects have never been reported in heterozygous *Six1* or *Six4* mutants or in their double knockouts (26, 29, 30, 35, 36). This indicates that there is likely an ascertainment bias in past assays, and also illustrates the value of a phenotype-driven screen.

Discussion

A Highly Efficient *PB*-Based F1 Dominant Screening System. In this study, we developed a *PB*-transposon-based F1 screening system in mice, with the following advantages: 1) High efficiency, 2) genome-wide mutagenesis, 3) cage-effective phenotypic screening, 4) capability of inducing both gain- and loss-of-function mutations, and 5) a visual marker for genotype tracking. Based on the mutagenesis efficiency and insertion distribution, we estimate that our F1 system would enable virtually any investigator to conduct genome-wide dominant phenotypic screens for gain- and loss-of-function mutations in mice. The screen could be completed within a year with total animal housing below 300 cages. Although the mutation frequency of our system is lower than ENU mutagenesis (7) (the current standard for phenotypic screens, which produces 30 coding mutations per gamete), this shortage is compensated for by low mapping cost and high mutagenicity. Using our protocol and developed software, mapping an insertion can be accomplished in the F2 or F3 generation by Splinkerette PCR or the newly developed Tn5 transposase-based tagmentation assay, both of which require less time and labor than mapping an

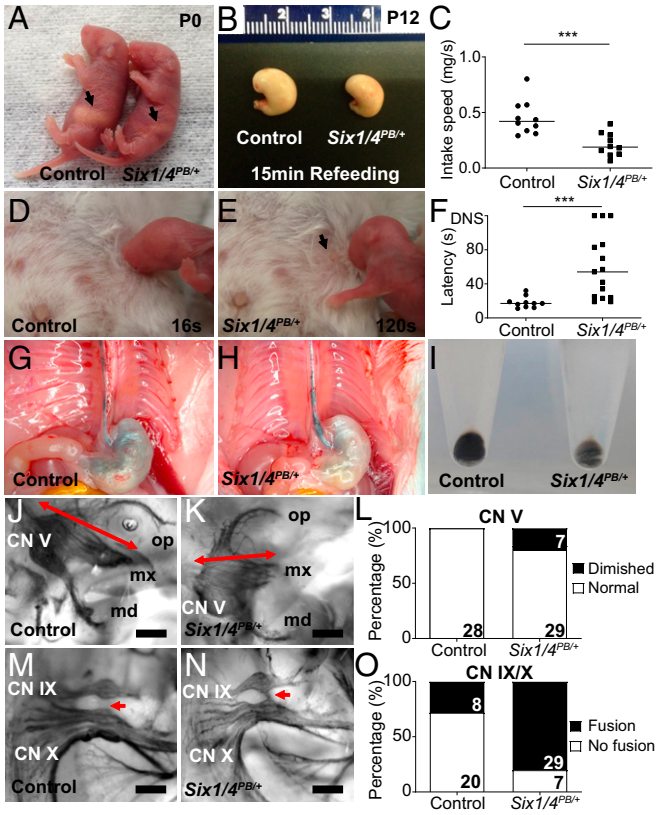

Fig. 5. *Six1/4*$^{PB/+}$ mutants exhibit abnormalities in milk intake behavior and defects in CNs. (*A*) Comparison between P0 wild-type and *Six1/4* mutant pups of the amount of milk in the stomach of a newborn. (*B*) Amount of milk in the stomach after 15-min refeeding. (*C*) Milk intake speed of the P12 wild-type and *Six1/4*$^{PB/+}$ mutant pups. ***$P < 0.001$. (*D–F*) The latency for nipple attachment of the P0 *Six1/4*$^{PB/+}$ newborns is increased (also see Movie S1). The arrow in *E* indicates an unattached nipple. ***$P < 0.001$. (*G–I*) Defects of the P12 *Six1/4*$^{PB/+}$ mutant pups in ingesting a charcoal suspension. (*J–L*) The length of the maxillary branch of CN V is diminished in 19.4% (7 of 36) of the E10.5 *Six1/4*$^{PB/+}$ embryos ($P = 0.0153$, Fisher's extract analysis).md, mandibular; mx, maxillary; op. ophthalmic. (*M–O*) The ganglia of CN IX and CN X are more frequently connected by fascicles or fused (arrows) in the E10.5 *Six1/4*$^{PB/+}$embryos than in control littermates ($P < 0.0001$, Fisher's exact analysis). (Scale bar, 100 μm.)

ENU-induced mutation. ENU is a powerful mutagen for efficiently inducing point mutations in a small indel in the mammalian genome, while *PB* transposon can serve as an excellent mutagen for interrogating the genome. The mutagens produced different types of alleles in both ENU and *PB* screening approaches, including null, hypermorph, hypomorph, antimorph, and neomorph alleles. However, *PB* insertion generates gain-of-function alleles at a much higher rate (*SI Appendix*, Fig. S6*A*). Based on the data obtained from our growth retardation screen and previously reported ENU screens (7, 8), we estimate that our engineered transposon is 5-fold more potent than ENU in its ability to induce functional mutations (*SI Appendix*, Table S2). Moreover, we previously conducted a *PB* insertional recessive screen for obesity (21). We recorded the body weight of all animals in the obesity screen and have now analyzed the original data from the screen for growth retardation. Through this analysis, we have discovered that 22 homozygous lines (20 genes) have a body size at least 10% smaller than wild-type (5.37%, total 408 lines) (*SI Appendix*, Table S3). This result indicates that a *PB* F3 insertional recessive screen provides a higher hit rate than a *PB* F1 dominant screen. Nevertheless, this system could also be used to perform recessive phenotypic screens in 3 generations, given that the *PB* construct can induce loss-of-function mutations.

The *PB* F1 screening system developed here should make it possible for individual laboratories to genetically decipher a variety of biological and disease traits in mammals. The CRISPR/Cas9 technology has demonstrated its power in mammalian reverse genetics, allowing efficient investigation of specific genes and their functions. Our *PB* F1 screening system is a similarly powerful tool in mammalian forward genetics, enabling investigators to quickly and cost-effectively identify genes and address biological questions in a nonhypothesis-driven manner. The International Mouse Phenotyping Consortium has generated and phenotypically characterized more than 5,000 knockout mouse lines. Their recent work has identified 410 (24%) lethal genes and 198 (10%) "subviable" genes from the first 1,751 unique gene knockouts (37). However, this knockout mouse database still does not cover all coding genes (*SI Appendix*, Table S4). Furthermore, unlike such F3 recessive screens, which are helpful to identify relationships between common phenotypes and genes, the *PB*-based F1 dominant screen provides a more flexible and highly efficient screening strategy to analyze any phenotypically driven screen, which could be used to complement the worldwide mouse knockout projects.

A Pilot Screen Identified Genes and Mechanisms of Growth Retardation. In this study, we successfully achieved a pilot screen for growth retardation and identified genes not previously known to be involved in this phenotype. Interestingly, in an effort to determine which tissue or organ was contributing to the growth retardation phenotypes in the *Rin2*$^{PB/+}$ mutant, we crossed the *LSL-Rin2* transgenic mice with different tissue-specific *Cre* lines and found that IUGR and PGR could be traced to different tissue origins (*SI Appendix*, Fig. S9). IUGR was partially caused by blood defects due to *Rin2* activation, while *Rin2* overexpression in neurons led to PGR (*SI Appendix*, Fig. S9). These findings exemplify how our F1 system can lead to the investigation of novel mechanisms underlying specific biological process. The *Six1/4*$^{PB/+}$ mutant provides a new animal model for BOR syndrome and displays abnormalities in nipple recognition and milk ingestion, a group of mammalian-specific phenotypes that has not been extensively studied. The milk intake defects led to PGR in our mouse model, while in humans, the corresponding defects could be compensated for by intensive maternal care. Our animal model argued for the significance of feeding difficulties as a clinical manifestation for BOR syndrome diagnosis. Transposons hit both coding and noncoding regions. In *Six1/4*$^{PB/+}$ mutants, the *PB* transposon inserted in a noncoding region about 70 kb upstream of the *Six1* coding region and 550 bp upstream of the *Six4* coding region, causing down-regulation of both *Six1* and *Six4 genes*. As *Six4* knockout mice have been reported to be free of any disease phenotypes[23], the BOR phenotypes in the *Six1/4*$^{PB/+}$ mutants should be explained by the loss of *Six1*, and the *PB* transposon might hit a long-range regulatory element of the *Six1* gene. This highlights that sequencing the corresponding noncoding genomic region might be a future direction for human BOR patients.

Future Applications of Our F1 System. Given its power, our *PB*-based F1 dominant screening system can be adapted to a variety of additional applications, such as modifier screens and in vivo enhancer/promoter functional analysis (Fig. 6). As many disease mouse models have been generated using traditional transgenic methods and advanced CRISPR/Cas9 technology, an F1 modifier screen in mice using our system will offer opportunities to identify enhancers and suppressors of particular disease processes (Fig. 6 *C* and *D*). These modifiers, once identified, could be directly translated into therapeutic targets for ameliorating human diseases that currently lack effective treatments. Furthermore, together with the *Cre-loxP* system and conditional reporters, activators, and inhibitors, our *PB* F1 system could be used to directly report and investigate the regulatory elements within the

mammalian genome in vivo (Fig. 6*E*). In summary, our *PB* F1 mutagenesis system and its future applications could enrich our understanding of the mammalian genome from multiple angles.

Methods

Ethics Statements. All procedures were approved by the Yale Animal Resources Center and the Institutional Animal Care and Use Committee under protocol nos. 10230, 11050, and 11250.

Mouse Lines. To generate the *PB[mut-Kat]* construct, the UbC promoter (pUB-GFP, Addgene) and luciferase (*luc2* from pGL4, Promega) were introduced outside the *PB[mut-tetO-Kat]* transposon (previously described in ref. 16) by overlapping PCR. The linearized *PB[mut-Kat]* construct was subsequently injected into FVB 1-cell embryos. The *PB[mut-Kat]* founders (FVB-Tg(PB[mut-Kat])Xut) were screened for the presence or absence of the transgene by PCR analysis. Primers are detailed in *SI Appendix*, Table S5. The *Luc-PB[mut]7* (FVB-Tg(Luc-PB[mut]7)Xut) and *Act-PBase* [FVB-Tg(Act-PBase)Xut] lines were previously reported (9, 12). Female FVB/NJ mice (Stock no: 001800) were purchased from the Jackson Laboratory.

Genotyping with PCR. We used 2 pairs of primers to validate new *PB* insertion mutants (*SI Appendix*, Table S5). Both DSP5 and DSP6 primers are within the *PB* element for detecting *PB* insertions, while the DSP7 primer is located in the *Luc* gene, which is used together with DSP5 to detect the original concatemer in genotyping PCR.

Ex Vivo Detection of Katushka Fluorescence. Katushka red fluorescence in neonatal mice was visualized in a dark room using a flashlight (Nightsea, Electronic Microscopy Sciences) on P0. Neonatal mice were imaged for Katushka fluorescence in the IVIS Spectrum (Xenogen) by epifluorescence (excitation: 570 nm; emission: 640 nm).

***PB* Construct DNA-FISH Probe Preparation Using Nick Translation.** The plasmid preparation was carried out following a protocol described previously (38). In detail, 1 μg *PB[mut-Kat]* construct without the *luc2* gene sequence was first incubated for 8 to 20 h at 15 °C with nick translation enzyme, 10× nick translation buffer, dNTP mix, and SpectrumGreen-dUTP/SpectrumOrange-dUTP from the nick translation kit (Abbott Molecular). The nick translation reaction was then stopped by heating in a 70 °C heat block for 10 min. The probe size was determined on a 1.5% agarose gel, and the smear signal should be between 100 and 1,000 bp. Probe concentration was measured on a Nanodrop microvolume spectrophotometer (Thermo Fisher). Next, 1 μg Cot1-DNA was added to 100 to 200 ng of each probe to form a mixture. The mixture was purified using isopropanol to precipitate, the pellet was resuspended into a DNA hydration solution, and the resulting *PB*-FISH probe solution was stored at −20 °C.

***PB*-FISH and Immunostaining of Tubular Cells in Testes.** Combinatorial immunostaining and *PB*-FISH were carried out following the protocol described previously. The primary antibody used to label spermatogonial stem cells was mouse monoclonal anti-SALL4 antibody (1:100; Abcam ab57577). The secondary antibody used was goat anti-mouse IgG with Alexa Fluor 488 (1:200; Thermo Fisher Cat # A32723). The *PB*-FISH probe used was the 1 labeled by SpectrumOrange-dUTP. Testes from 3-mo old *Act-PBase/PB[mut-Kat]10* mice and *PB[mut-Kat]10* homozygous mice were first removed and placed in a 60-mm Petri dish containing PBS. The tunica albuginea outside the testes was then removed, and the testes were cut into small pieces. The seminiferous tubule fragments were then dispersed by pipetting and trypsin/EDTA (1:1) solution treatment at 37 °C for 5 min. Isolated tubule cells were collected after digestion, dropped on the center of a poly-L-lysine–coated coverslip, and then fixed using 2% paraformaldehyde/1× PBS, pH 7 to 7.4, for 10 min at room temperature. The tubular cells were first incubated in primary antibody ant-SALL4 solution and secondary antibody solution for immunostaining, and then were postfixed for 10 min with 2% paraformaldehyde/1× PBS solution. After 3 rinses in 1× PBS, the tubular cells were denatured in 1.9 M HCl for 30 min at room temperature. After 3 rinses in ice-cold 1× PBS, cells were hybridized with *PB*-FISH probe overnight at 37 °C in a dark and humid chamber. Finally, cells were washed 3 times using the saline-sodium citrate buffer and mounted in ProLong Gold mounting medium containing 1.5 μg/mL DAPI. Slides could be stored in a 4 °C refrigerator for long time. The fluorescent images in Fig. 2 were taken by Leica TSC SP8 confocal microscopy. In detail, the 405-, 488-, and 543-nm excitation lasers were used to image DAPI (blue), Alexa Fluor 488 (green), and SpectrumOrange (orange), respectively. NIH ImageJ software was used to quantitate the fluorescence intensity of the *PB* FISH signals in Fig. 2 *O* and *P*. We analyzed 2 independent animals for each genotype, and examined 16 spermatogonial stem cells (32 puncta in total) and 20 mature sperms from *PB[mut-Kat]10* homozygous mice along with 20 spermatogonial stem cells and 20 mature sperms from *Act-PBase/PB[mut-Kat]10* mice.

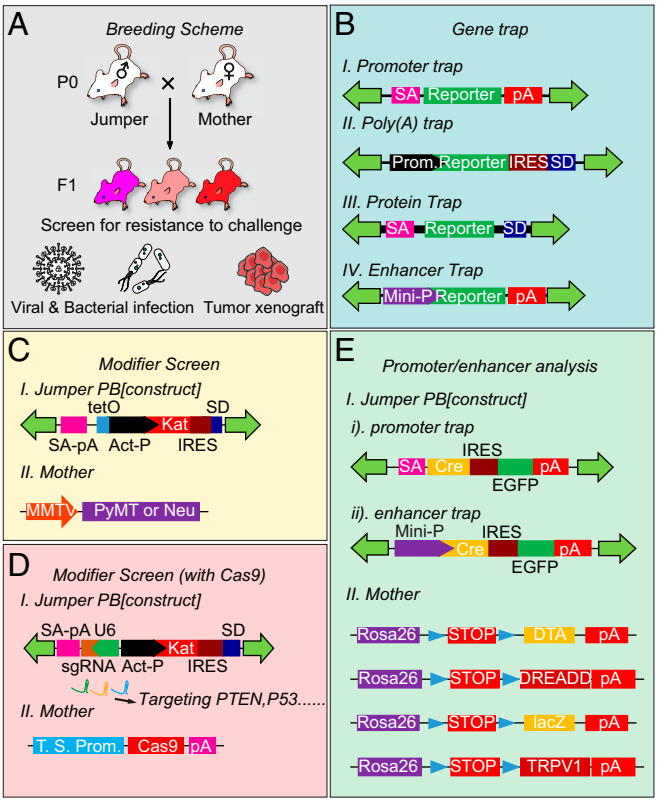

Fig. 6. Versatile applications of the efficient genome-wide *PB* F1 screening system for functional genome analysis. (*A*) The breeding scheme for positive selection. F1 progeny are challenged with deadly pathogens, such as viruses, bacteria, and tumor xenografts. Surviving mutants are bred for inheritance test and transposon mapping. (*B*) Gene-trap mutagenesis transposon constructs. Transposons can be modified for purposes such as promoter trap, poly-A trap, enhancer trap, and protein trap. Mini-P, minipromoter; PA, poly-A signal; Prom, promoter; SA, splice acceptor; SD, splice donor. (*C* and *D*) *PB*-based F1 modifier screens. (*C*) P0 *MMTV-PyMT* transgenic mothers are used in the screen for identifying enhancers or suppressors of breast tumor growth and metastasis. Act, CAG promoter; tetO, tet operon. (*D*) *CRISPR/Cas9* technology is used to knockout single or multiple tumor-suppressor genes in a tissue-specific manner. This genetic background is incorporated into our *PB* F1 screening system for identifying enhancers or suppressors of the tumor suppressors. sgRNA, small guide RNA; T.S. Prom, tissue-specific promoter; U6-P, U6 promoter. (*E*) Schematic design for promoter/enhancer analysis. Mutagenic transposon carries a promoter or enhancer trap using *Cre* as 1 of the reporters. P0 mothers carry transgenes conditionally controlled by the *Cre-loxP* system for expression pattern labeling, neuron activity regulation, and cell ablation. DREADD, designer receptors exclusively activated by designer drugs; DTA, diphtheria toxin; TRPV1, transient receptor potential cation channel subfamily V member 1.

Preparation of Mouse Embryonic Fibroblast Cells. E13 to E14 embryos were dissected sterilely in a 100-mm tissue culture dish. The gut was removed, the head was cut off for genotyping, and the bodies were minced with scissors and washed with sterile 1× PBS, then centrifuged at 800 × rpm for 5 min. This was repeated 3 times. Next, 1 mL of 0.25% trypsin was then added, and the mixture was transferred to a 1.5-mL Eppendorf tube and incubated at 37 °C for 15 min. The mixture was pipetted up and down to dissociate the tissue. The cell mixture was transferred to a 15-mL Falcon tube with 10 mL of DMEM with 10% FBS and 1× penicillin/streptomycin, mixed well, and the cells were allowed to sit vertically for 1 min. The supernatant with single cells was transferred to a 100-mm tissue culture dish and incubated in a 37 °C incubator

with 5% CO_2 overnight. The cells received the same fresh medium the next morning and continued to grow.

PB-FISH, Chromosome Painting in Mouse Embryonic Fibroblast Cells. The mouse chromosome 6 painting probe was purchased from MetaSystems (MP 6 Orange, D-1406-050-OR). The *PB*-FISH probe used was the 1 labeled by SpectrumGreen-dUTP.

Identification of *PB* Insertion Sites Using Splinkerette PCR. Splinkerette PCR for mapping *PB* transposon insertion sites was performed as described previously (11, 15) with the following modifications. Mouse genomic DNA was digested with Sau3AI or MspI (New England Biolabs). Digested DNA was then ligated with splinkerette adaptors. A primary PCR followed by a nested PCR was performed to amplify the flanking genomic sequence. The adaptor sequences and PCR primer sequences are listed in ref. 11 and *SI Appendix*, Table S4. The nested PCR product was gel-extracted for sequencing. The sequencing data were analyzed using the University of California, Santa Cruz genome browser and the National Center for Biotechnology Informaton GenBank database.

Identification of *PB* Insertion Sites in the Germ Cells of Jumpers. The insertion sites within the genomes of testes and sperm from all 3 male jumpers were mapped using the left arm of the *PB* transposon by splinkerette PCR and multiplexed Illumina sequencing. Genomic DNA was extracted from mouse testes and sperm separately according to the manufacturer's protocol (Qiagen). Splinkerette PCR was performed as described above. After adaptor ligation, the ligated genomic DNA library was digested with BspHI to destroy the concatemer flanking sequences. Illumina adaptor sequences and individually assigned bar-code sequences were incorporated into the nested PCR primers (*SI Appendix*, Table S4). PCR products were mixed for Illumina high-throughput sequencing (HiSeq 2000, Illumina). Sites with sequencing reads over 100 were selected for analyzing the genomic distribution of transposon insertions. The sequencing data were aligned onto the GRCm38/mm10 assembly database and the insertion distribution was plotted using the Ensembl Genome Browser.

Tn5 Tagmentation-Based Mapping Assay of *PB* Genomic Insertion Sites. Tn5 tagmentation and *PB* genomic insertions site mapping were carried out following a previously described protocol (39). Genotyping genomic DNA was purified from the mouse tail using Proteinase K (Sigma, 3115887001) digestion and the Quick-gDNA Micro Prep kit (Zymo, D3020). Fifty-six F1s' DNA were then diluted to a concentration of 50 ng/μL and transferred to a 96-well plate (Thermo Scientific, SKU # AB-0700). Tn5 transposase was purified from the plasmid psfTn5 (Addgene plasmid #79107, a gift from Sten Linnarsson, Karolinska Institutet, Stockholm, Sweden), which included a hyperactive Tn5 transposase carrying E54K and L372P mutations. The Tn5 transposon forward primer (Tn5F) and reverse primer (Tn5R) were synthesized by the Yale Keck Oligonucleotide Synthesis facility, and annealed at 95 °C in a thermal cycler (Bio-Rad). The Tn5 transposase as then precharged by Tn5 transposon at 37 °C for 30 min. The 56 F1's diluted DNA were transferred to a new 96-well plate, mixed with precharged Tn5 transposase and incubated at 55 °C for 7 min. Because of Tn5 transposase's random cutting and insertions, the genomic DNA were cut into pieces with a Tn5 transposon sequence at both ends. We then use a primer in the *PB* sequencing (PBL121) and a primer (Tn5_i7) in Tn5 transposon sequence carrying the index barcode and i7 sequence to amplify the *PB* genomic insertion sites with more or less 10 rounds of PCR. In the second round of PCR, a primer (PBL_i5) carrying the index barcode and the i5 sequence and the i7 primer (i7) were used to amplify the *PB* insertions sites with another 10 rounds of PCR. The PCR products from 56 F1s were then mixed together, and purified using DNA Clean & Concentrator (Zymo, D4013) for Illumina high-throughput sequencing (HiSeq 2000, Illumina). We used the 75-bp paired-end model to sequence the index barcodes in both 5′ and 3′ ends.

Sequencing Data Analysis. A python script was assembled to achieve the following bioinformatics steps. First, reads are barcode-split, trimmed, and merged using *fastx-toolbox* and *Biopython* (40). Second, reads are mapped and deduplicated to the mouse genome sequence (mm10) using *Bowtie2* (41). Third, regions of the genome are identified using *samtools* and some *mathematica* scripts (WOLFRAM, Mathematica 11). The python script is available in github (https://github.com/superresolution/PBMapper_python).

Growth Retardation Screen. For the growth retardation screen, 2,036 F1 mice were screened for small body size at the weaning age (3 to 4 wk old). We compared the body size of the Kat^+ F1s with its siblings of the same sex. The candidates with 10% smaller in body weight would be selected out. We then test the heritability by checking if the growth retardation phenotype is inheritable in the next generation (F2). The candidates were mapped by Splinkerette PCR if it could pass the heritability test. To draw the growth curves in Fig. 4 *C*–*G*, we performed a cross between wild-type females and PB^+ males to avoid the influence of the genotype of the mother on fetus growth retardation.

Detection of Transposon-*Rin2* and Transposon-*Zeb2* Fusion Transcripts. Fusion transcripts were detected in tissues by RT-PCR using an internal ribosome entry site sense primer (IRESF) and a Rin2 exon 3 specific antisense primer (RIN2R) or a Zeb2 exon 2 specific antisense primer (ZEB2R) (*SI Appendix*, Table S5). Sanger sequencing of PCR products was performed to analyze whether fusion transcripts in gain-of-function mutants indeed spliced into the downstream exons through the transposon's splice donor.

Nipple Recognition Assay. The assay to test the ability of pups to recognize nipples was described previously (42). Pups aged between 6 and 12 h after birth were fasted for 3 h by separating them from their mother and placing their cage on a 30 °C heating pad. Their lactating mother was anesthetized by intraperitoneal injection of ketamine (100 mg/kg) and xylazine (10 mg/kg). Pups were then individually removed from their nest and vertically supported toward the mother's nipple with their noses touching hairs around the nipple for 120 s. The nipple attachment latency of individual pups was recorded. To eliminate the bias introduced by operator influence, the operator was blinded to the genotypes of pups during the assay. When attachment did not occur within 120 s, the pups were recorded as belonging to the did not suckle group and a conservative latency value of 120 s was recorded.

Milk Intake Speed. P12 pups were separated from their mother for 5.5-h starvation to empty their stomachs, with their body temperature maintained by placing their cage on a 30 °C heating pad. Two pups (1 control and 1 mutant) were returned to the mother for 15-min refeeding. During the refeeding period, the suckling time for each pup was recorded. Immediately after the refeeding, the pups were killed and the amount of milk in the stomachs was weighed. Milk intake speed was calculated by dividing the amount of milk by the suckling time.

Milk Ingestion. Activated charcoal powder (Sigma) was suspended in drinking water with 10% (wt/vol). Next, 10 μL of the charcoal suspension was administrated into the mouths of P12 pups directly every 15 s for 1 min. After 1 min, pups were killed, and their stomach contents were imaged. To better visualize the ingested charcoal, the stomach contents were suspended in water and the suspension was then centrifuged.

Whole-Mount Staining. E10.5 embryos were fixed in 4% PFA at 4 °C and dehydrated through a graded methanol/PBS series. Embryos were subsequently incubated in 5:1 methanol:H_2O_2 for 4 h to quench endogenous peroxidases, and then rehydrated and blocked overnight in PBS with 3% milk and 0.1% Triton X-100 (PBSMT). Next, embryos were incubated in mouse monoclonal anti–165-kDa neurofilament protein (2H3, Developmental Studies Hybridoma Bank; 1:200) at 4 °C for 2 d. Embryos were then washed extensively in PBSMT and incubated overnight in a 1:500 dilution of HRP-conjugated goat anti-mouse secondary antibody (Jackson Immuno Research). Following Diaminobenzidine (DAB, Sigma) visualization of HRP, embryos were dehydrated, cleared with benzyl alcohol:benzyl benzoate (BABB), and imaged using a Leica MZ FLIII microscope.

Ex Vivo Imaging for Luciferase Activity. Mice were injected with D-luciferin (150-mg/kg body weight, intraperitoneally) and then anesthetized using isoflurane vaporizer. The bioluminescent signal was quantified using the IVIS Spectrum (Xenogen).

Cell Culture. HEK293 cells were cultured in DMEM containing 10% FBS under standard conditions. Cells were transfected with plasmids using Lipofectamine 2000 according to the manufacturer's instructions (Invitrogen). Doxycycline (Sigma) was added to the culture medium at a concentration of 1 μg/mL

Quantitative PCR for Gene Expression. Total RNA was extracted from the tissues of interest using TRIzol Reagent and a standard protocol (Life Technologies). cDNA was synthesized from mRNA by reverse transcription (Bio-Rad). Specific primer pairs (*SI Appendix*, Table S3) were designed to amplify the cDNA of target genes. Quantitative PCR with SYBR Green Mix (Bio-Rad) was performed in the StepOne Real-time PCR machine (Applied Biosystems). Relative transcript expression was determined by the relative standard curve method.

Mouse Lines Used in *SI Appendix*, Fig. S8. *ACTB-cre* [also known as B6N.FVB-Tmem163Tg(ACTB-cre)2Mrt/CjDswJ; Stock No: 019099], *Sox2-cre* [also known as B6.Cg-Edil3Tg(Sox2-cre)1Amc/J; Stock No: 008454], *Nestin-cre* [also known as B6.Cg-Tg(Nes-cre)1Kln/J; Stock No: 003771], *Alb-cre* [also known as B6.Cg-Speer6-ps1Tg(Alb-cre)21Mgn/J; Stock No: 003574], *Pdx1-cre* [also known as B6.FVB-Tg(Pdx1-cre)6Tuv/J; Stock No: 014647], *Vav-iCre* [also known as B6.Cg-Commd10Tg(Vav1-icre)A2Kio/J; Stock No: 008610], and female FVB/NJ mice were purchased from the Jackson Laboratory. The *Myogenin-Cre* transgenic line was a kind gift from Eric N. Olson, University of Texas Southwestern Medical Center, Dallas, TX. To create the LSL-Rin2 plasmid, a mouse *Rin2* cDNA (Open Biosystems) sequence was cloned downstream of a β-chicken actin promoter (pCAG-GFP, Addgene: 11150). A transcriptional termination sequence flanked by 2 loxP sites (pBS302, Addgene: 11925) was inserted between the promoter and the Rin2 coding sequence. The linearized LSL-Rin2 construct was subsequently injected into FVB 1-cell embryos. The *LSL-Rin2* founders [FVB-Tg(Act-LSL-Rin2)Xut] were screened for the presence or absence of the transgene by PCR analysis.

Statistical Analysis. To compare growth curves, milk intake behavior, and phenotypic frequency of CN defects, *P* values were calculated using appropriate analysis of variance, 2-tailed Mann–Whitney test, and Fisher's exact test, respectively (GraphPad Prism). $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

ACKNOWLEDGMENTS. We thank F. Qian, D. Manry, M. O. Dietrich, and C. Chabu for discussion; the Yale University Animal Genomics Service for pronuclear injections; the Yale Center for Genomic Analysis for sequencing and data analysis; and J. Li for technical assistance. This work was supported in part by funding from the Team for Growth Control and Size Innovative Research of Zhejiang (Grant 2018R01003). Y.P. was a China Scholarship Council and Yale predoctoral fellow. T.X. was a Howard Hughes Medical Institute Investigator.

1. A. Bradley, M. Evans, M. H. Kaufman, E. Robertson, Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**, 255–256 (1984).
2. K. R. Thomas, M. R. Capecchi, Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503–512 (1987).
3. H. Wang *et al*., One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
4. S. Hitotsumachi, D. A. Carpenter, W. L. Russell, Dose-repetition increases the mutagenic effectiveness of N-ethyl-N-nitrosourea in mouse spermatogonia. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6619–6621 (1985).
5. A. R. Moser, H. C. Pitot, W. F. Dove, A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* **247**, 322–324 (1990).
6. M. H. Vitaterna *et al*., Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science* **264**, 719–725 (1994).
7. M. H. Hrabé de Angelis *et al*., Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nat. Genet.* **25**, 444–447 (2000).
8. P. M. Nolan *et al*., A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat. Genet.* **25**, 440–443 (2000).
9. S. Ding *et al*., Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122**, 473–483 (2005).
10. L. S. Collier, C. M. Carlson, S. Ravimohan, A. J. Dupuy, D. A. Largaespada, Cancer gene discovery in solid tumours using transposon-based somatic mutagenesis in the mouse. *Nature* **436**, 272–276 (2005).
11. R. Rad *et al*., PiggyBac transposon mutagenesis: A tool for cancer gene discovery in mice. *Science* **330**, 1104–1107 (2010).
12. S. F. Landrette, J. C. Cornett, T. K. Ni, M. W. Bosenberg, T. Xu, piggyBac transposon somatic mutagenesis with an activated reporter and tracker (PB-SMART) for genetic screens in mice. *PLoS One* **6**, e26650 (2011).
13. Z. Ivics, P. B. Hackett, R. H. Plasterk, Z. Izsvák, Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510 (1997).
14. V. W. Keng *et al*., Region-specific saturation germline mutagenesis in mice using the Sleeping Beauty transposon system. *Nat. Methods* **2**, 763–769 (2005).
15. W. Wang *et al*., Chromosomal transposition of PiggyBac in mouse embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9290–9295 (2008).
16. J. Choi *et al*., Identification of PLX4032-resistance mechanisms and implications for novel RAF inhibitors. *Pigment Cell Melanoma Res.* **27**, 253–262 (2014).
17. K. Horie *et al*., Efficient chromosomal transposition of a Tc1/mariner- like transposon Sleeping Beauty in mice. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 9191–9196 (2001).
18. K. Horie *et al*., Characterization of Sleeping Beauty transposition and its application to genetic screening in mice. *Mol. Cell. Biol.* **23**, 9189–9207 (2003).
19. S. Lin *et al*., Comparison of the transcriptional landscapes between human and mouse tissues. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17224–17229 (2014).
20. M. A. St John *et al*., Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian tumours and pituitary dysfunction. *Nat. Genet.* **21**, 182–186 (1999).
21. J. Cui *et al*., Disruption of Gpr45 causes reduced hypothalamic POMC expression and obesity. *J. Clin. Invest.* **126**, 3192–3206 (2016).
22. K. Saito *et al*., A novel binding protein composed of homophilic tetramer exhibits unique properties for the small GTPase Rab5. *J. Biol. Chem.* **277**, 3412–3418 (2002).
23. K. Verschueren *et al*., SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5′-CACCT sequences in candidate target genes. *J. Biol. Chem.* **274**, 20489–20498 (1999).
24. D. J. Jung, S. Y. Na, D. S. Na, J. W. Lee, Molecular cloning and characterization of CAPER, a novel coactivator of activating protein-1 and estrogen receptors. *J. Biol. Chem.* **277**, 1229–1234 (2002).
25. B. D. Yu, J. L. Hess, S. E. Horning, G. A. Brown, S. J. Korsmeyer, Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* **378**, 505–508 (1995).
26. H. Ozaki *et al*., Six4, a putative myogenin gene regulator, is not essential for mouse embryonal development. *Mol. Cell. Biol.* **21**, 3343–3350 (2001).
27. C. Laclef *et al*., Altered myogenesis in Six1-deficient mice. *Development* **130**, 2239–2252 (2003).
28. R. Grifone *et al*., Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. *Development* **132**, 2235–2249 (2005).
29. S. P. Strom *et al*., De novo variants in the KMT2A (MLL) gene causing atypical Wiedemann-Steiner syndrome in two unrelated individuals identified by clinical exome sequencing. *BMC Med. Genet.* **15**, 49 (2014).
30. A. Kochhar *et al*., SIX1 mutation screening in 247 branchio-oto-renal syndrome families: A recurrent missense mutation associated with BOR. *Hum. Mutat.* **29**, 565 (2008).
31. A. Sanchez-Valle *et al*., HERV-mediated genomic rearrangement of EYA1 in an individual with branchio-oto-renal syndrome. *Am. J. Med. Genet. A.* **152A**, 2854–2860 (2010).
32. B. Turgeon, S. Meloche, Interpreting neonatal lethal phenotypes in mouse mutants: Insights into gene function and human diseases. *Physiol. Rev.* **89**, 1–26 (2009).
33. H. P. Donald, Suckling and suckling preference in pigs. *Emp. J. Exp. Agric.* **5**, 361–368 (1937).
34. S. P. Cordes, Molecular genetics of cranial nerve development in mouse. *Nat. Rev. Neurosci.* **2**, 611–623 (2001).
35. D. Zou, D. Silvius, B. Fritzsch, P. X. Xu, Eya1 and Six1 are essential for early steps of sensory neurogenesis in mammalian cranial placodes. *Development* **131**, 5561–5572 (2004).
36. Y. Konishi, K. Ikeda, Y. Iwakura, K. Kawakami, Six1 and Six4 promote survival of sensory neurons during early trigeminal gangliogenesis. *Brain Res.* **1116**, 93–102 (2006).
37. M. E. Dickinson *et al*.; International Mouse Phenotyping Consortium; Jackson Laboratory; Infrastructure Nationale PHENOMIN, Institut Clinique de la Souris (ICS); Charles River Laboratories; MRC Harwell; Toronto Centre for Phenogenomics; Wellcome Trust Sanger Institute; RIKEN BioResource Center, High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514 (2016).
38. D. J. Bolland, M. R. King, W. Reik, A. E. Corcoran, C. Krueger, Robust 3D DNA FISH using directly labeled probes. *J. Vis. Exp.*, 50587 (2013).
39. D. L. Stern, Tagmentation-based mapping (TagMap) of mobile DNA genomic insertion sites. *bioRxiv*:10.1101/037762 (31 August 2017).
40. P. J. Cock *et al*., Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
41. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
42. D. W. Logan *et al*., Learned recognition of maternal signature odors mediates the first suckling episode in mice. *Curr. Biol.* **22**, 1998–2007 (2012).