High-throughput and genome-scale targeted mutagenesis using CRISPR in a nonmodel multicellular organism, *Bombyx mori*

Sanyuan Ma, Tong Zhang, Ruolin Wang, et al.

Genome Res. published online January 8, 2024
Access the most recent version at doi:10.1101/gr.278297.123

Method

High-throughput and genome-scale targeted mutagenesis using CRISPR in a nonmodel multicellular organism, *Bombyx mori*

Sanyuan Ma,[1] Tong Zhang,[1] Ruolin Wang,[1] Pan Wang,[1] Yue Liu,[1,2] Jiasong Chang,[1,3] Aoming Wang,[1] Xinhui Lan,[1] Le Sun,[1] Hao Sun,[1] Run Shi,[1] Wei Lu,[1] Dan Liu,[1] Na Zhang,[1] Wenbo Hu,[1] Xiaogang Wang,[1,4] Weiqing Xing,[1] Ling Jia,[1] and Qingyou Xia[1]

[1]*Integrative Science Center of Germplasm Creation in Western China (CHONGQING) Science City, Biological Science Research Center, Southwest University, Chongqing 400716, China;* [2]*Medical Center of Hematology, Xinqiao Hospital, Army Medical University, Chongqing, 400037, China;* [3]*Key Laboratory of Cellular Physiology, Ministry of Education, Shanxi Medical University, Taiyuan, 030001, China;* [4]*China Chongqing Key Laboratory of Chinese Medicine & Health Science, Chongqing Academy of Chinese Materia Medica, Chongqing 400065, China*

Large-scale genetic mutant libraries are powerful approaches to interrogating genotype–phenotype correlations and identifying genes responsible for certain environmental stimuli, both of which are the central goal of life science study. We produced the first large-scale CRISPR-Cas9-induced library in a nonmodel multicellular organism, *Bombyx mori*. We developed a *piggyBac*-delivered binary genome editing strategy, which can simultaneously meet the requirements of mixed microinjection, efficient multipurpose genetic operation, and preservation of growth-defect lines. We constructed a single-guide RNA (sgRNA) plasmid library containing 92,917 sgRNAs targeting promoters and exons of 14,645 protein-coding genes, established 1726 transgenic sgRNA lines following microinjection of 66,650 embryos, and generated 300 mutant lines with diverse phenotypic changes. Phenomic characterization of mutant lines identified a large set of genes responsible for visual phenotypic or economically valuable trait changes. Next, we performed pooled context-specific positive screens for tolerance to environmental pollutant cadmium exposure, and identified KWMTBOMO12902 as a strong candidate gene for breeding applications in sericulture industry. Collectively, our results provide a novel and versatile approach for functional *B. mori* genomics, as well as a powerful resource for identifying the potential of key candidate genes for improving various economic traits. This study also shows the effectiveness, practicality, and convenience of large-scale mutant libraries in other nonmodel organisms.

[Supplemental material is available for this article.]

In the past centuries, the widely application of model organisms has helped to reveal the basics of molecular biology. However, some current model organisms are too simple and thus cannot answer more complex questions nor explain some specific phenomena (Fields and Johnston 2005). With the transition in biology from descriptive to mechanistic understanding, especially with the completion of genome sequencing of massive organisms, more nonmodel systems are emerging for tackling questions across the whole spectrum of biology or for exploiting the unique biological features of a special organism to address questions of general importance (Goldstein and King 2016). For example, axolotl (Echeverri et al. 2022), killifish (Beck et al. 2022), volvox (Umen and Herron 2021), ashbya (Wendland and Walther 2005), and diatoms (Jang et al. 2013) have been used to investigate regeneration, aging, evolution, phase transition, and nanobiotechnology, respectively. The silkworm, *Bombyx mori*, is another example, with advantages in showing the biology of Lepidoptera insects, which comprise the vast majority of agricultural and forest pests (Xia et al. 2014; Meng et al. 2017a,b). *B. mori* is also an economically important insect, with sericulture forming a pillar industry in the rural economy of many developing countries.

Currently, a major gap in methodologies and resources between the selected few models and other organisms greatly hinder researchers in choosing and retaining the use of nonmodel organisms (Russell et al. 2017). In *B. mori*, only a few hundred mutant lines with limited genetic background information have been preserved (Goldsmith et al. 2005; Tong et al. 2022). Functional genomic approaches are lagging, with the first transgenic silkworm obtained in 2000 (Tamura et al. 2000), 17 years after the first transgenic plant (Herrera-Estrella et al. 1983). Widely used RNAi technology has been shown to be extremely inefficient in *B. mori* and many other lepidoptera insects (Terenius et al. 2011; Kolliopoulou and Swevers 2014).

The emergence of CRISPR-Cas9 has expanded genetic toolboxes in a wide range of systems, including *B. mori* (Cong et al.

Corresponding authors: masy@swu.edu.cn, xiaqy@swu.edu.cn
Article published online before print. Article, supplemental material, and publication date are at https://www.genome.org/cgi/doi/10.1101/gr.278297.123.

2013; Jinek et al. 2013; Mali et al. 2013; Ma et al. 2019). Its simplicity and modularity have led to the successful construction of large CRISPR libraries for genome-wide functional genomics and drug target discoveries in mammalian cell systems (Shalem et al. 2014; Wang et al. 2014; Qi et al. 2017; Ford et al. 2019). Large-scale mutant libraries have also been constructed for four model organisms: rice (*Oryza sativa* L.) (Lu et al. 2017; Meng et al. 2017a,b), rapeseed (*Brassica napus*) (He et al. 2023), zebrafish (*Danio rerio*) (Sun et al. 2020), and fruit fly (*Drosophila melanogaster*) (Meltzer et al. 2019; Port et al. 2020; Zirin et al. 2020). CRISPR-based mutants have simple and clear genotype–phenotype correlations, and mutant libraries provide better opportunities to interrogate gene functions than previous libraries, such as those generated through radiation mutagenesis, chemical mutagenesis, random transgenic insertion, or transgenic RNAi (Venken et al. 2014).

We established a genome-wide CRISPR screen platform in cultured *B. mori* cells and showed its simplicity and efficiency in identifying genes essential for cell viability under normal conditions (Chang et al. 2020a), as well as for screening resistance genes for both biotic and nonbiotic stresses (Liu et al. 2021). However, the application of such powerful genome-scale mutagenesis systems at the whole-animal level in nonmodel organisms has been unsuccessful to date. Here, we report the construction of a high-throughput CRISPR-Cas9 mutant library for *B. mori*, a useful resource for silkworm research and breeding. We show its application for gene function interrogation, as well as its potential use for genetic improvement of economic traits. In silkworm, the most efficient germline transformation platform is *piggyBac* system, which has been proved and widely used for over 20 years (Ma et al. 2014a; Chang et al. 2020a). *piggyBac* transposon can drive gene delivery in a wide range of eukaryotes, so our strategy and results also lay the foundation for applying genome-scale mutagenesis to other nonmodel organisms.

Results

Design and construction of the CRISPR single-guide RNA library

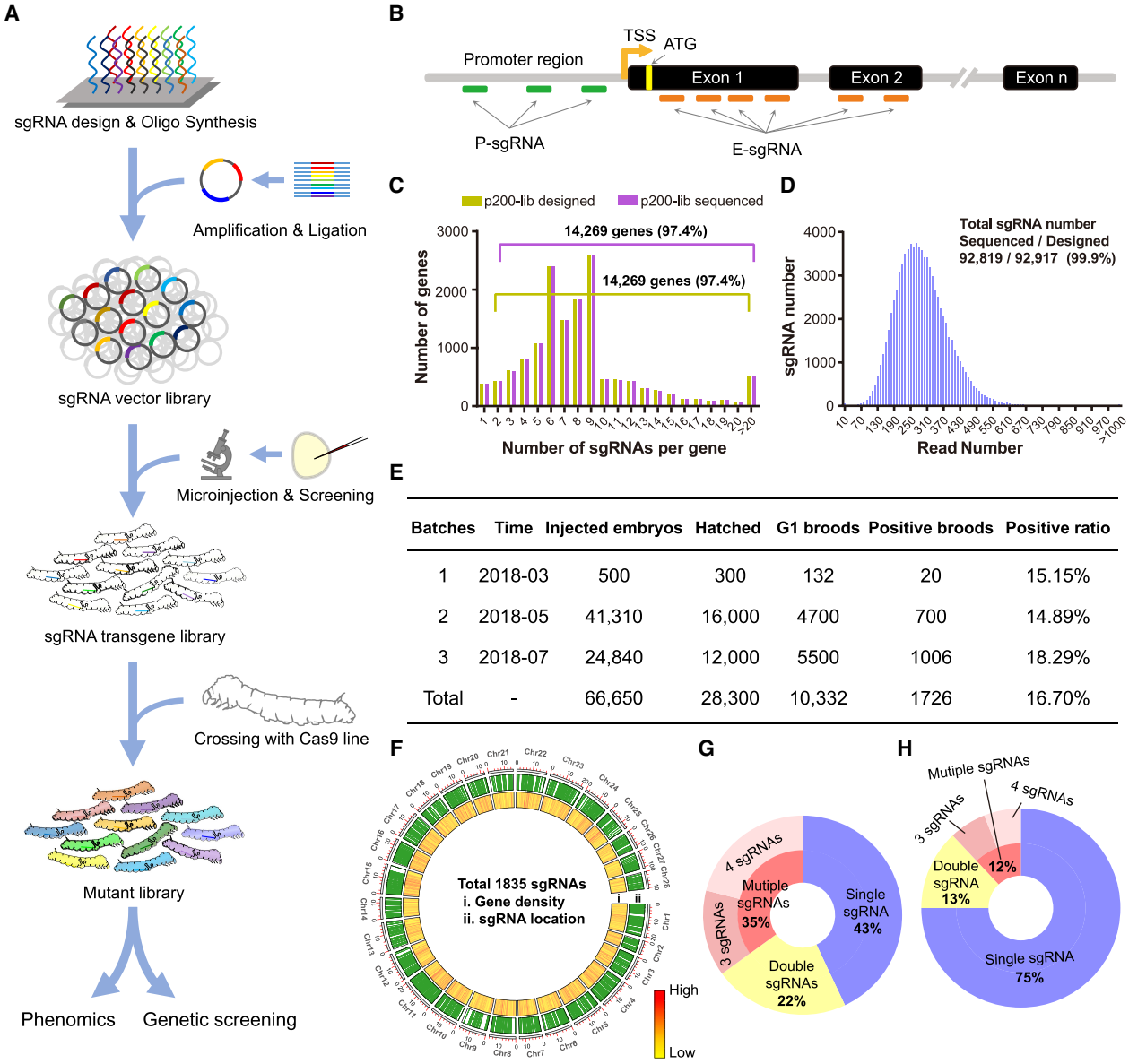
Based on our previously reported, highly efficient, CRISPR-Cas9-mediated *B. mori* targeted mutagenesis in the whole animal and the pooled library in cultured cells (Ma et al. 2014a; Chang et al. 2020a), we sought to explore the ability of the CRISPR system to generate a library of targeted loss-of-function mutants at the whole-animal level. In *B. mori*, genetic transformation is primarily achieved through embryonic microinjection (Ma et al. 2014b). This process is extremely labor intensive and equipment dependent and requires highly skilled technicians; thus, to maximize the efficiency, we designed a high-input/low-output pooled strategy (Fig. 1A). As in the pooled strategy, the target gene of an individual mutant is indicated by the integrated single-guide RNA (sgRNA) sequence itself, we used a binary CRISPR-Cas9 system (Xu et al. 2017) to avoid misediting of nontarget genes during the early stage of injection. This system expressed Cas9 protein and sgRNA in two separate transgenic lines that use the ubiquitous IE2 and eye-specific 3XP3 promoter-driven EGFP as transformation markers, respectively (Ma et al. 2014a; Xu et al. 2017), and knockout was achieved by crossing them (Fig. 1A).

To construct the mutant library, we designed a strategy shown in Figure 1A. The whole genome of *B. mori* was searched for all exons and ~500 bp upstream of transcription starting site (Fig. 1B), and 3,489,264 sgRNAs were found (Supplemental Fig. S1). sgRNAs targeting exons were ranked using criteria described previously (Chang et al. 2020a). sgRNAs targeting promoters were ranked using an additional criterion; only sgRNAs targeting the sense strand were selected owing to previous findings that the DSBs were repaired in a direction-biased manner (Chang et al. 2020b; Ma et al. 2021). A total of 92,917 sgRNAs were chosen (about eight E-sgRNAs and about two P-sgRNAs per gene) and synthesized on a microarray chip (Supplemental Table S1). These sgRNAs are designed to target 14,645 genes, and the distribution of these targeted genes across the genome is random (Supplemental Fig. S1), with these genes accounting for ~86.76% of all genes in the genome. The rest of the genes, 13.24%, was not included because we failed to design proper sgRNAs. The synthesized sgRNA oligonucleotide pool was elongated by polymerase chain reaction (PCR) and ligated into p200 to form a plasmid library (p200-lib). Deep sequencing of the library showed that 99.9% of the designed sgRNAs were successfully cloned; >97.4% of the genes had more than one sgRNA (Fig. 1C), and 99.9% of sgRNAs had 36–4109 reads (Fig. 1D). These results indicate sufficient coverage, accuracy, and diversity of p200-lib for subsequent experiments.

Pooled microinjection and generation of large-scale sgRNA transgenic lines

To test the usability and efficiency, a small fraction of p200-lib was coinjected with a *piggyBac* transposase expression vector (A3-Helper) into 500 G_0 embryos. Approximately 300 embryos hatched and were reared to adulthood. Twenty broods positive for 3Xp3-EGFP expression were obtained from 132 G_1 broods. The transformation rate was 15.15% (Fig. 1E), which was the average of our previous individual transgenic experiments (Ma et al. 2013). We crossed the 20 transgenic lines with the Cas9-expressing transgenic line (Fig. 1A), and observed clear phenotypic changes in the progenies of the two lines among the 20 F_1 lines (Supplemental Fig. S2). We tested the sgRNA of the two lines through simple genomic PCR and found that both phenotypes were theoretically related to their predicted functions (Paine et al. 2019; Schimmel et al. 2021). The target regions of *Peanuts* (KWMTBOMO05973) and *DNA primase large subunit* (KWMTBOMO01805) were amplified and subjected to deep sequencing. Sequencing results showed that 46% of 12,149 reads and 82% of 22,422 reads were edited (Supplemental Fig. S2). We concluded that the strategy is suitable for the construction of a genome-wide mutant library, and the established p200-lib plasmid library is suitable for efficient large-scale germline transformation.

Next, two batches of large-scale microinjections were performed. A total of 66,150 embryos were injected, and 28,000 injected embryos were hatched and were reared to adult G_0 moths. By large-scale fluorescent screening the obtained 10,200 G_1 broods (about 2,000,000 embryos), 1706 3XP3-EGFP-positive broods, containing one to 100 positive transgenic individuals per brood, were obtained (Fig. 1E). Average hatchability and transformation ratios were 42.46% and 16.70%, respectively. Given that individuals from the same G_1 brood might harbor the same sgRNA, we chose one positive individual from each positive G_1 brood and reared them in a pooled manner. Adult moths were crossed back with wild types to generate sgRNA transgenic lines, and an sgRNA library containing 1726 lines was established. To determine the sgRNAs in each line, we performed barcoded deep sequencing and revealed that 1726 lines contain a total of 1835 sgRNAs targeting 1615 genes (Fig. 1F). Over half of the individuals harbored more than one sgRNA, indicating multicopy *piggyBac*-mediated



Batches	Time	Injected embryos	Hatched	G1 broods	Positive broods	Positive ratio
1	2018-03	500	300	132	20	15.15%
2	2018-05	41,310	16,000	4700	700	14.89%
3	2018-07	24,840	12,000	5500	1006	18.29%
Total	-	66,650	28,300	10,332	1726	16.70%

Figure 1. Design and construction of the CRISPR single-guide RNA (sgRNA) library. (*A*) The strategy of the mutant library construction. (*B*) The schematic diagram of sgRNA design that targets the promoter region and exon around whole genome of *B. mori.* The sgRNAs targeting the promoter region were named by P-sgRNA, whereas sgRNAs targeting the exon1 or exon2 were named by E-sgRNA. (*C*) The sgRNA distribution for the p200-lib of designed and deep sequenced. (*D*) Distribution of the deep sequenced reads and sgRNA numbers. (*E*) Summary of microinjection information for *B. mori.* (*F*) Distribution of sgRNAs detected by deep sequencing over the *B. mori* whole genome. (*G*,*H*) The number of sgRNAs detected in individuals.

transformation (Fig. 1G). To generate as many single sgRNA lines as possible, we performed two continuous rounds of backcrossing of these sgRNA lines with wild types and increased the percentage of single sgRNA lines from 43% to 75% (Fig. 1H).

Genotypic characterization and editing outcome of the mutant library

After generating 1726 sgRNA lines, we investigated the potential of these lines to generate a genome-edited mutant library. As a proof-of-principle demonstration, we randomly selected 300 sgRNA lines, which contained 243 gRNAs targeting 188 genes, and crossed them with Cas9-expressing lines. A total of 150,000 hybrid progenies were reared on fresh food and screened under fluorescent microscopy for mutants. Individuals with both 3Xp3-EGFP- and IE2-EGFP-positive signals were considered as mutants. To determine editing efficiencies and patterns, we randomly chose three to six mutant silkworms from each line and performed deep sequencing of the amplified target regions. Of the 1436 tested silkworms, 89.8% showed detectable editing events (Supplemental Fig. S3), with >64% showing editing efficiency >50% (Fig. 2A). Such efficiency is sufficient for functional investigation of target genes, and the results encouraged us to improve it. We analyzed the potential parameters that may affect editing efficiency. The

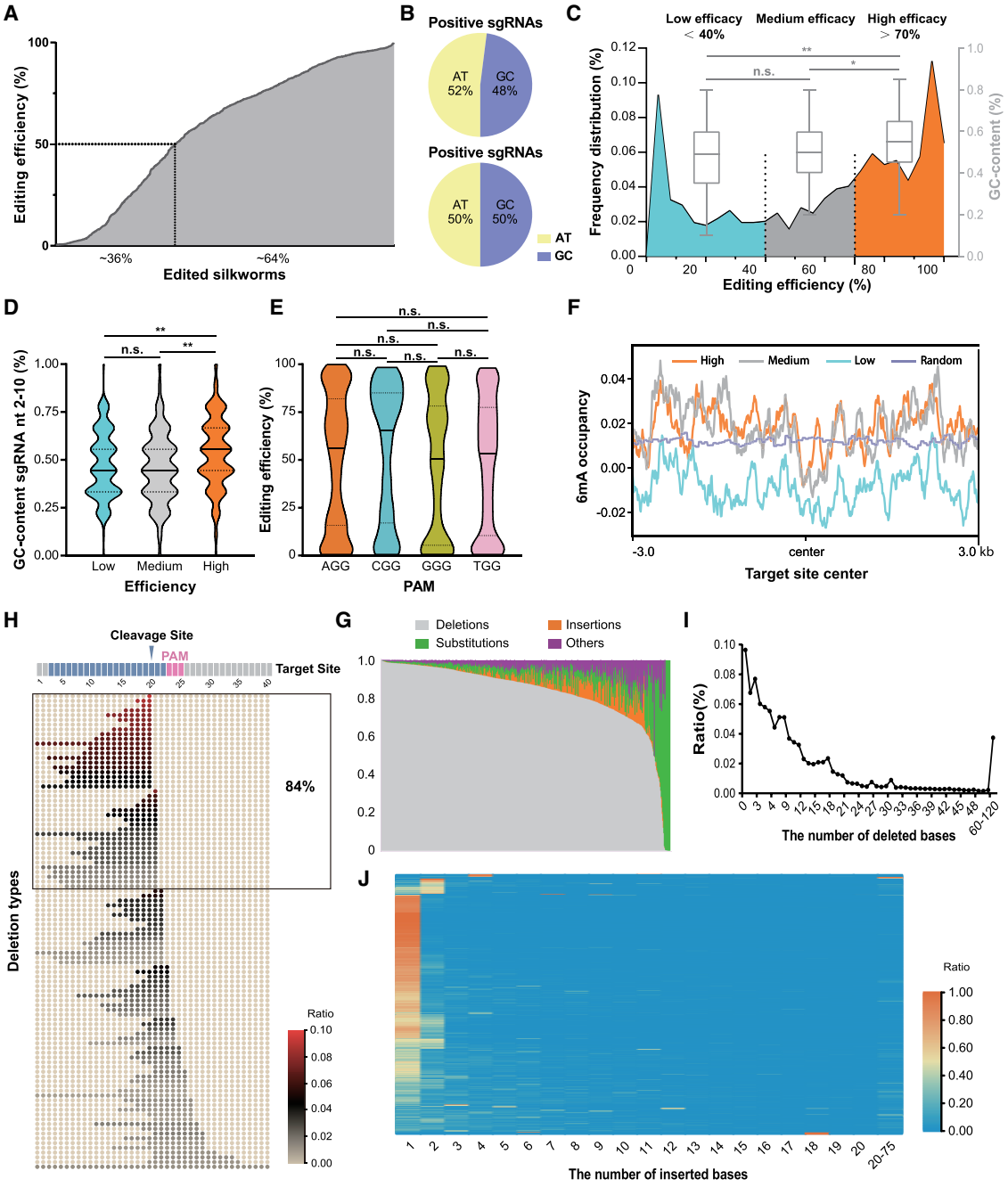


Figure 2. Genotypic characterization and editing outcome of the mutant library. (*A*) The overall editing efficiency of the edited silkworms. (*B*) The GC content in the positive and negative sgRNAs. (*C*) Frequency of distribution of editing efficiency and the GC content in high (editing efficiency > 70%)-efficiency, medium-efficiency, and low-efficiency (editing efficiency < 40%) sgRNA groups: (*) $P < 0.05$, (**) $P < 0.01$, (n.s.) not significant; one-way ANOVA with Benjamini–Hochberg post-test multiple comparison. (*D*) The GC content of different editing efficiency: (*) $P < 0.05$, (**) $P < 0.01$, (n.s.), not significant; one-way ANOVA with Benjamini–Hochberg post-test multiple comparison. (*E*) The editing efficiency of different PAM sequences: (*) $P < 0.05$, (**) $P < 0.01$, (n.s.), not significant; one-way ANOVA with Benjamini–Hochberg post-test multiple comparison. (*F*) The DNA 6mA methylation occupancy around the target center with different editing efficiencies. (*G*) The overall editing patterns of the edited silkworms. (*H*) The top 100 deletion types and their percentages. (*I*) The size of the deletions and their occurrences. (*J*) The size of the insertions and their percentage.

GC content was higher in positive than in negative sgRNAs (Fig. 2B). In positive sgRNAs, GC content was 54%, 49%, and 48% in the high-, medium-, and low-efficiency sgRNA groups (Fig. 2C), respectively, with this tendency more significant in the 2–10 nt of the sgRNA (distal to the PAM) region (Fig. 2D), indicating a positive contribution of GC content to editing efficiency (Sun et al. 2020). The first base of the PAM sequence has no effect on editing efficiency (Fig. 2E). We also found that G at the 3′ end (Supplemental Fig. S4A), G at the 5′ end (Supplemental Fig. S4B), G/CG/C dinucleotides (Supplemental Fig. S5) within the seed region, and 6mA methylation at the target site (Fig. 2F) are beneficial for efficient editing, whereas A at the 3′ end (Supplemental Fig.

S4A), C at the 5′ end (Supplemental Fig. S4B), and A/TA/T dinucleotides within the seed region (Supplemental Fig. S5) had a negative effect. We analyzed editing efficiencies of multiple sgRNAs in a single animal and found that the integrated multiple sgRNAs did not affect each other (Supplemental Fig. S6).

Large-scale deep sequencing of target sites provided the opportunity to investigate editing outcomes in vivo, the patterns of which were previously found to display distinct differences between cultured cells of *B. mori* and that of other species (Chang et al. 2020b; Ma et al. 2021). Of the total 61,720,266 reads, we detected 28,927,046 mutated reads containing 1,387,421 mutation types. Most target sites showed editing patterns mainly composed of deletions (86.50%) with a small fraction of insertions (7.39%) and substitutions (3.28%), similar to that of previous observations (Fig. 2G). A few target sites showed a preponderance of insertion or substitution, whereas deletion accounted for only a small part (Fig. 2G). More than 84% of the deletions occurred upstream of the cleavage sites (Fig. 2H), which further confirmed the distinct unidirectional feature of DSB repair in *B. mori* (Chang et al. 2020b; Ma et al. 2021). Deletion sizes were generally negatively correlated with their occurrence, with significant occurrence peaks at each 3X base pair (Fig. 2I), indicating a potential beneficial selection mechanism during DSB repair in *B. mori*. Within the insertions, 56.66% and 14.14% were 1-bp and 2-bp insertions, respectively (Fig. 2J). Most insertions were from templated repair (Chang et al. 2020b); 75% of the 1-bp insertions were duplications of −4 nt of the sgRNA sequence; and 80% of the 2-bp insertions were duplications of −5/−4 dinucleotides of the sgRNA sequence (Supplemental Fig. S7).

Phenotypic characterization of the mutant library

A major application of mutant libraries is to discover new phenotypes and their causal genotypes. To evaluate this potential in the CRISPR-Cas9 mutant library, we harvested all F_1 embryos from 300 randomly selected lines for phenotyping by rearing three broods per line. Thirty-three out of 300 F_1 strains showed significant morphological differences compared with the wild type in terms of body size, skin coloration, developmental duration, spot patterns, appendage development, and metamorphic arrest (Fig. 3A,B). We correlated these phenotypes with the corresponding sgRNA and target site sequencing results, finding individuals from all 33 F_1 lines had correct sgRNAs and high editing efficiency at target sites (Supplemental Fig. S8). Many mutant phenotypes were consistent with or theoretically reasonable to previously described mutants, either *B. mori* or other related species, such as *Drosophila*. For example, genes (*VPS11*, KWMTBOMO01164; *VPS16*, KWMTBOMO03319) of two mutants showed an oily skin phenotype and were previously shown to be involved in the transport and accumulation of uric acid in the epidermis (Fujii and Banno 2019; Segala et al. 2019), whereas the developmental arrest mutant gene, *neverland* (KWMTBOMO06146), was previously shown to be a key enzyme in ecdysone biosynthesis and caused the arrest of both molting and growth during *Drosophila* development (Yoshiyama et al. 2006; Lang et al. 2012), The wing-deficient mutant gene, *Tout-velu* (KWMTBOMO02722), was shown to be required for the diffusion of Hedgehog, a key regulator of limb development (Bellaiche et al. 1998). These results indicated that the current library can be used to identify mutants.

As *B. mori* is an economically valuable insect in addition to being a research system, we also investigated the economic traits of the mutant lines. Cocoon weight and shell ratio are the two most obvious traits used to evaluate quality of a commercial strain; we measured the cocoon weight and shell ratio of all 300 mutant lines. Most of these cocoons were average; however, we found a few lines that showed a notable higher or lower cocoon weight (Fig. 3C) and cocoon shell ratio (Fig. 3D), including a major component of silk protein (*BmFib-L*, KWMTBOMO08464) (Fig. 3D). The multibillion-dollar sericulture industry has long sought an easy way to breed only males, because they produce more silk of a higher quality (Kiuchi et al. 2014; Zhang et al. 2018). We counted the males and females in each line and found that progenies of four lines contained only males whereas one line only contained females (Fig. 3E), indicating that knocking out these genes might affect the survival of males or females. We also investigated cocoon shapes of these lines (Supplemental Fig. S9), an important trait that affects postharvesting industrial processes, and found several genes that may regulate spinning behaviors (Fig. 3F). These investigations provide targets for strains with improved traits and further illustrate the viability of our library.

Use of the CRISPR mutant library for pooled screens

Phenotypes of most mutant lines are not visible to the naked eye but may be responsible for certain physiological changes and exogenous stresses. An important application of the mutant library is pooled screening, which can identify genes responsible for certain fine-designed biotic or abiotic stresses. To explore this potential, we performed a proof-of-principle abiotic stress screening against a common environmental pollutant, cadmium (Cd), exposure to which causes significant economic losses in sericulture areas each year (Liu et al. 2021). We selected approximately 100 embryos (one-fourth of each brood) from each of the 300 mutant lines and pooled them. After hatching, the larvae were reared on a normal artificial diet in a pooled manner until the third instar. They were then fed an artificial diet containing 30 mg/kg $CdCl_2$, which is lethal to the wild-type *B. mori*. Surviving silkworms were moved to a fresh artificial diet containing $CdCl_2$ to ensure continuous exposure (Fig. 4A). Only 53 individuals survived to the fifth instar. sgRNA and their sequencing revealed that 11 larvae harbored the same sgRNA, targeting *Calreticulin* (KWMTBOMO12902) (Fig. 4B). Sequencing the target region of this sgRNA confirmed highly efficient mutagenesis of KWMTBOMO12902 (Supplemental Fig. S10). Subcellular localization showed that the gene was located in the endoplasmic reticulum and was involved in calcium ion transport (Lu et al. 2020). Given that the proportion of the sgRNAs in the input pool was principally equal before exposure, we speculated that KWMTBOMO12902 is a Cd stress response gene and that its knockout might cause Cd resistance in silkworms. We checked the performance of this gene in our previous cell line–based genome editing screening results and found that KWMTBOMO12902 was not among the top-ranked positive genes, possibly owing to the diversity of Cd ion exposure and responses between individuals and cells, which emphasized the importance of individual mutant pools for finding new strains. We further challenged silkworm individuals and cultured cells with $CdCl_2$; we observed that KWMTBOMO12902 was significantly up-regulated in the midgut, fat body, and *B. mori* embryonic cell line BmE cells of silkworm larvae after Cd exposure (Fig. 4C). An independent $CdCl_2$ tolerance test using third instar larvae on a larger scale showed that the mutant line KWMTBOMO12902 had a significantly higher tolerance to $CdCl_2$ than the wild type (Fig. 4D). These results indicate that the mutant library in this

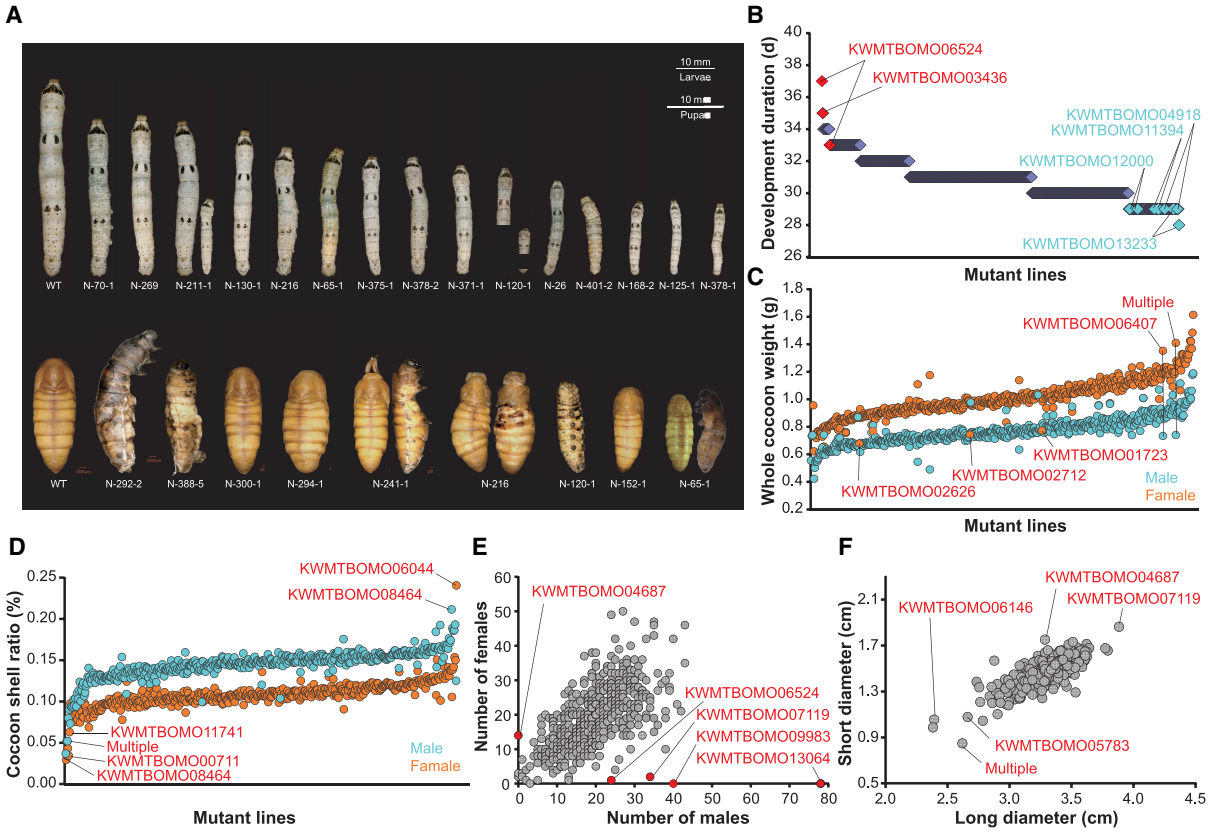


Figure 3. Phenotypic characterization of the mutant library. (*A*) The body sizes, epidermal coloring, developmental duration, speckle patterns, appendage development, and metamorphosis arrest of 33 F_1 lines. These 33 F_1 strains showed significant morphological differences among the 300 F_1 strains. (*B*) The development duration of different mutant lines. The genes that may be associated with development extending are highlighted by red; those in light blue are more possibly associated with a shortened development. (*C*) The whole cocoon weight of different mutant lines. The genes possibly associated with whole cocoon weight improvement are highlighted by red. Each blot represents the average cocoon weight of 50 individuals. (*D*) The cocoon shell ratio of different mutant lines. The genes possibly associated with cocoon shell ratio improvement are highlighted by red. Each blot represents the average cocoon shell ratio of 50 individuals. (*E*) The number of males and females in each line. The genes possibly associated with sex differentiation of silkworm are highlighted by red. (*F*) The short and long diameter of cocoon in each line. The genes that may regulate spinning behaviors are highlighted by red. Each blot represents the average diameters of 30 individuals.

study could be a powerful resource for identifying new genes responsible for specific physiological functions.

Discussion

Since the early twentieth century, the mutant library has been the main source of major breakthroughs in life science innovation and discovery. Several methods, including physical radiation (X-rays, fast neutrons, and γ-rays), chemical induction (ethyl methane sulphonate and 1,2:3,4-diepoxybutane), random DNA insertion (P-element, *piggyBac*, and T-DNA), and site-specific mutations (homologous recombination, integrase, RNAi, ZFN, TALEN, and CRISPR-Cas9) have been developed to generate large-scale mutant libraries in a few model organisms such as *D. melanogaster*, *Mus musculus*, *D. rerio*, *Arabidopsis thaliana*, *Nicotiana tabacum*, *O. sativa* L, and *B. napus*. (Amsterdam et al. 1999; Hrabé de Angelis et al. 2000; Nolan et al. 2000; Gondo 2008; Lein et al. 2008; Xu et al. 2008; Chang et al. 2012; Haelterman et al. 2014; Lu et al. 2017; Meng et al. 2017a,b; Port et al. 2020; He et al. 2023). Here we used the CRISPR-Cas9 system to construct a large-scale mutant library in a nonmodel organism, *B. mori*, and showed its application in generating novel mutants with important phenotypic changes and in identifying new genes with important physiological functions and economic potential from pooled screening. Research on nonmodel organisms is crucial for solving complex or specialized biological questions; however, they are severely hindered by the lack of efficient genetic tools and mutant resources. We showed the feasibility of generating a large-scale mutant library using a nonmodel organism. Furthermore, because we used *piggyBac* transposon, which can drive gene delivery in a wide range of eukaryotes and has the largest reported cargo capacity (Wu et al. 2007; Chang et al. 2019; Li et al. 2020) as the delivery strategy, the results in this study also provide a straightforward reference for genome-scale mutagenesis in other nonmodel organisms.

The strategy used in this study has several advantages over previous methods. Traditional mutant *B. mori* variates were collected from either naturally occurring local or chemically/physically induced mutations (Goldsmith et al. 2005). These have extremely poor phenotype–genotype correlations as the mutations occur randomly across the genomic region, and different mutations have diverse genetic backgrounds (Tong et al. 2022). It usually takes years to map a certain phenotype to its causal genetic alterations (Liu et al. 2010; Atsumi et al. 2012; Yamaguchi et al. 2013). In *Drosophila* spp. And mice, large-scale mutagenesis can

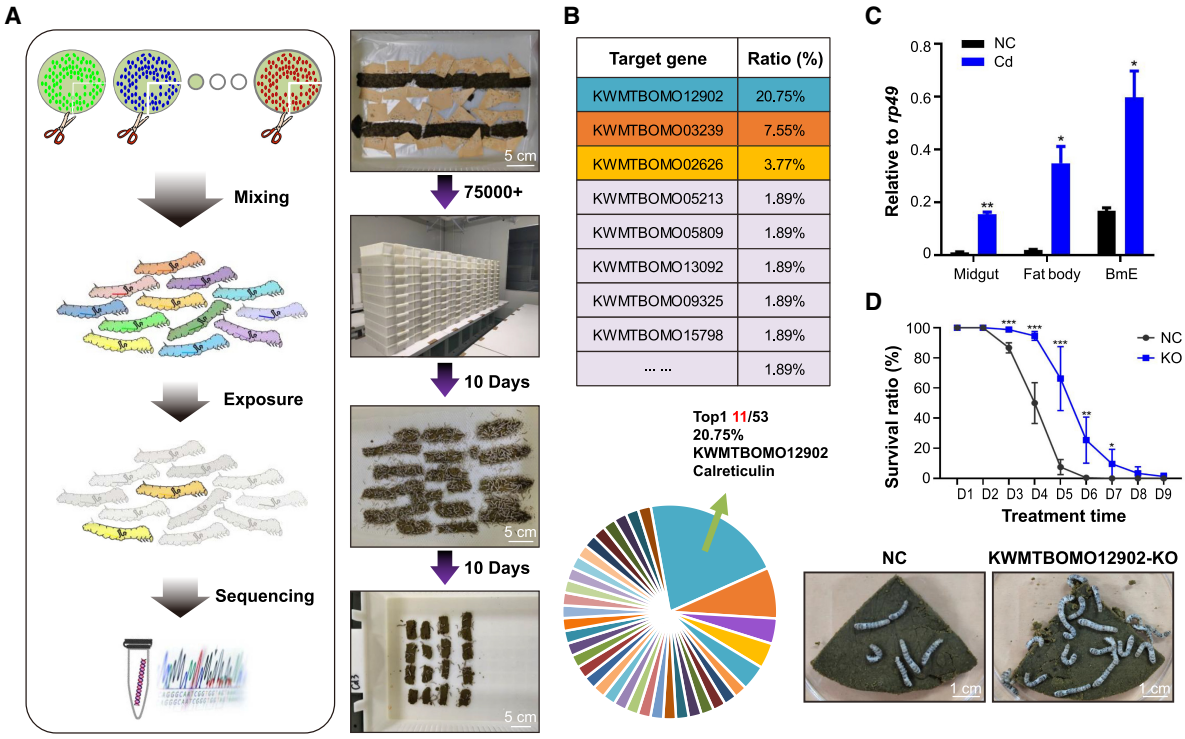


Figure 4. Cadmium (Cd) exposure screening in the mutant library. (*A*) The schematic diagram of abiotic stress screening in the mutant library with Cd exposure (*left*) and several scenes during the process (*right*). (*B*) The target genes survived from persistent Cd exposure. Individuals with sgRNA targeting KWMTBOMO12902 accounted for the highest proportion. (*C*) The expression level of KWMTBOMO12902 in midgut, fat body, and BmE cells with or without Cd exposure. (*D*) The survival rate of normal strain and KWMTBOMO12902-KO strain in Cd stress. The photographs were taken on the fifth day after Cd exposure. Only surviving silkworms were shown in the photograph, as the dead individuals were discarded to avoid cross-infection.

also be achieved through subsequently developed targeted methods such as homologous recombination and RNAi, but both are extremely inefficient in *B. mori* and many other nonmodel organisms. The CRISPR technology used in our study overcame all the shortcomings of previous methods. First, all mutagenesis was guided by the sgRNA sequences, which were designed to target only the coding regions of the genome, and all mutant lines in the library were truly functionally mutated. Second, the mutated gene and mutant sequences in each line can be tracked through integrated sgRNA sequences by simple targeted PCR amplification and sequencing, meaning all mutant lines in the library have clearer and closer genotype–phenotype correlations than do previous resources. Third, as shown, the efficiency of the CRISPR library for generating large-scale mutants is sufficient for any laboratory with common equipment. Our study also provided many parameters that may further improve this efficiency, such as the high input/low input strategy to avoid repeated sgRNA lines, the backcrossing strategy to reduce multiple sgRNA insertions, and the binary vector system to avoid the loss of growth-deficient lines with lethal genes.

B. mori is one of the most economically beneficial insects and a representative lepidopteran insect because of its crucial role in producing the best-known animal fibers for thousands of years, as well as its finely decoded genomic sequences (Xia et al. 2014; Tong et al. 2022). Systematic functional investigation of *B. mori* genes is required for both fundamental research and agricultural applications. To achieve this goal, we established multiple genetic manipulation tools, including transgenesis, site-specific recombinases, ZFN, TALEN, and CRISPR-Cas9, as well as a high-throughput CRISPR library screening platform in cultured *B. mori* cells (Ma et al. 2012, 2013, 2014b; Duan et al. 2013; Liu et al. 2014, 2018; Chang et al. 2020a). However, because the accumulation of genetic resources and research in *B. mori* relies heavily on the study of orthologous genes in related model organisms, the function of >80% of the *B. mori* genes remains unclear. We herein present the first high-throughput platform to generate large-scale whole-animal-level mutant libraries and provide a hypothesis-free, cost-effective pipeline to rapidly investigate the function of *B. mori* genes on a genome-wide scale. Considering the limitations of resources and time, we have decided to focus on protein-coding genes associated with specific mutant phenotypes to gain a better understanding of their functions. As a proof-of-principle demonstration, we designed and constructed a plasmid library containing 92,917 sgRNAs targeting 14,645 protein-coding genes, with high-quality coverage, distribution, and accuracy. A total of 1726 sgRNA lines were generated by microinjection of more than 60,000 embryos, and 300 mutant lines with deciphered mutant sequences containing more than 30 obvious phenotypic changes were obtained. A shortcoming of this study is that we did not identify a sgRNA line that targeted the promoter regions. As more transgenic lines are generated in our future work, it is promising to investigate the phenotypes of promoter-targeting sgRNA lines. These resources provide a powerful platform for functional annotation of *B. mori* genomes.

We previously established a pooled loss-of-function screening platform in cultured *B. mori* cells and showed its effectiveness in identifying genes involved in multiple biotic and abiotic stresses (Chang et al. 2020a; Liu et al. 2021). It remains a significant

challenge to perform pooled genetic screening at the whole-animal level owing to the lack of mutant resources and an efficient screening method. In large-scale functional screening experiments, it is generally required a high coverage and homogeneity of a library. However, most previous screenings were performed in cultured cells or single-cell systems, whereas this study used a mutant library in the whole organism of higher-order Eukaryote. In the whole-organism-level mutant library construction and screening, it is impossible and also unnecessary to reach a high coverage of sgRNAs. Thus, we performed the first in vivo pooled genetic screening of *B. mori* using the obtained 300 mutant lines. We showed that the CRISPR mutant library was highly efficient in identifying the genes responsible for stress. The screening results were mutually complemented genome-wide CRISPR screenings in cultured cells, and the selected mutants showed high performance when checked individually. We further identified the KWMTBOMO12902 gene, the absence of which significantly increased the tolerance to Cd exposure. Cd exists extensively in the ecosystem as a typical pollutant. It has been identified as a global threat because it cannot be degraded or resolved, is gradually enriched up the biological chain, and is severely harmful in various animals and plants, including humans (Yang et al. 2022). Studies aiming to decipher signaling and metabolic pathways involved in the pathological effects associated with Cd exposure are urgently needed, but it is difficult to perform large-scale toxicant exposure experiments in mammals. Here we used *B. mori* for large-scale pooled screens, and with the nature of *B. mori* being sensitive to environmental pollutants, it provided a potential target to explore organismal detoxification mechanisms (Mao et al. 2019). Environmental Cd exposure causes significant economic damage to the sericulture industry. Results showed the KWMTBOMO12902 gene mutant showed a high tolerance to Cd, even at a dosage much higher than that of environmental Cd, suggesting that KWMTBOMO12902 can be used to breed a Cd-tolerant silkworm strain.

Although the CRISPR-Cas9 system has been a powerful genome editing tool for more than 10 years, designing sgRNAs with high activity remains a major challenge (Chen and Wang 2022). As has been recognized in a number of previous reports and in the current study, not all sgRNAs can cleave their target DNA, and the cleavage efficiencies of different sgRNAs vary significantly (Li et al. 2023). A key parameter that determines the efficiency of sgRNA is its sequence and structural properties. We systematically investigated the effects of sgRNA sequence composition by sequencing 1436 mutant silkworms and found that the GC content in the 2–10 nt of the sgRNA (distal to the PAM) region, C content within the PAM sequences, G at the 3′ end, G at the 5′ end, and G/CG/C dinucleotides within the seed region are beneficial for efficient editing. We also investigated the effect of the local chromosome state on editing efficiency and found that 6mA methylation at the target site has a positive effect. This is an interesting observation, but further investigation in a future study is needed to confirm it and reveal its mechanism. These results provide practical guidance for the design of sgRNA for a successful genome editing experiment in *B. mori* and increase the understanding of the heterogeneity of sgRNA functions.

In conclusion, we have established the first genome-scale CRISPR mutagenesis platform at the whole-animal level in a nonmodel multicellular organism. A plasmid library containing 92,917 sgRNA-expressing vectors and a collection of 1726 sgRNA transgenic lines was established. By generating and analyzing 300 CRISPR mutant lines, we showed that the CRISPR library is a powerful resource for the rapid discovery of unknown gene functions and for investigating long-standing questions in silkworms and entomology. Applying the established mutant lines, we discovered and highlighted a large set of new genes that can cause divergent visible phenotypic changes, as well as several genes that have promising potential applications in breeding strains, with higher yield or greater tolerance for the sericulture industry. These resources are preserved in the Biological Research Center of Southwest University and are shared worldwide through the silk genome database, SilkDB 3.0 (https://silkdb.bioinfotoolkits.net) (Supplemental Fig. S11; Lu et al. 2020). Furthermore, the number of sgRNA transgenic and mutant lines will continue to increase. We believe that this study lays the foundation for large-scale functional genomics of *B. mori* and the rapid identification of novel genes involved in diverse biological processes. More importantly, the present study also sheds light on the application of genome-scale mutagenesis to other nonmodel organisms.

Methods

Silkworm strains and rearing conditions

Two silkworm strains, Dazao and Nistari, were used in the present study. Dazao was used to inject the sgRNA library and Nistari to express Cas9. Dazao was reared in our laboratory, whereas Nistari was provided by Prof. Anjiang Tan from Jiangsu University of Science and Technology, Zhenjiang, China. The silkworms were fed fresh mulberry leaves or artificial diet at 25°C and 80% relative humidity. BmE is a silkworm embryonic cell line preserved in our laboratory. These cells were grown in Grace medium, which was supplemented with 10% fetal bovine serum (FBS) as well as penicillin/streptomycin. The incubation temperature for the cultures was set at 27°C.

Design and construction of the vectors

The *piggyBac* transposase expression vector A3-helper was obtained from stocks in our laboratory. pUC57-3Xp3-EGFP-sv40 was constructed by replacing DsRed in pUC57-3Xp3-DsRed-sv40 (3DS) (Ma et al. 2014a) with the EGFP sequence. pUC57-U6-gRNA(AarI) was constructed by replacing the BbsI restriction site in pUC57-gRNA (Ma et al. 2014a) with an AarI restriction site. PUC57-3Xp3-EGFP-sv40 was cloned into the AgeI/AscI sites of the pBac-modified vector to generate pB-modified {3Xp3-EGFP-sv40}. pUC57-U6-gRNA(AarI) was cloned into the NheI/AscI site of pB-modified {3Xp3-EGFP-sv40} to generate the vector pB-modified {3Xp3-EGFP-sv40}{U6-gRNA(AarI)}, named p200.

sgRNA library design and synthesis

The silkworm genome assembly and gene models are obtained from SilkBase (http://silkbase.ab.a.u-tokyo.ac.jp). sgRNAs were designed using the CCTop method (Stemmer et al. 2015), which uses CRISPRater values to select gRNAs with potentially higher activity. All sgRNAs selected had 5′ G added to improve the U6 transcription efficiency. The sgRNA library was encoded within 70-nt oligonucleotides and synthesized on 92,917 arrays using the services of the Beijing Genomics Institute (China). The oligonucleotide sequence used was 5′-TACAAAATATCGTGCTCTACAAGTGNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTT-3′.

Construction of sgRNA library

The DGP-1 fragment (U6 promoter) was PCR-amplified from p200 using the primers KU-1R and DG-1R. The pool of sgRNA

oligonucleotides was amplified by PCR using primers DP-2F and DP-2R and named DGP-2 fragments (multiple sgRNAs). The DGP-3 fragment (sgRNA scaffolds) was amplified by PCR from p200 using the primers DG-3F and KU-1F. These three fragments were linked using overlap PCR to generate DGP-4 fragments (U6-sgRNA library). All PCR were performed using PrimeSTAR max DNA polymerase (Takara). The DGP-4 fragments were then cloned into the AscI/NheI (NEB) site of the p200 vector to construct the sgRNA library (p200 library) according to the following protocol: The gel-purified DGP-4 fragments (200 ng) and p200 vector (200 ng) were mixed and ligated in a 50 µL T4 ligation reaction (NEB) according to the manufacturer's instructions. The ligation reaction (1 µL) was transformed into 50 µL of *Escherichia coli* HST08 premium electro-cells (Takara) using a gene pulser Xcell (Bio-Rad) according to the manufacturer's instructions. To ensure the diversity of the sgRNA library, 15 parallel electro-transformations were performed with ampicillin selection, which resulted in more than 1000× library coverage. All clones were collected, and the plasmid library was extracted using a QIAprep spin miniprep kit (Qiagen).

Microinjection and positive transgenic screening

After mating, the female moths were placed on paper starched with paste to lay eggs in the dark. Eggs were collected every 0.5 h and arranged on slides (approximately 100 eggs per slide). After the eggs were fixed on slides, plasmids were injected using a microinjection apparatus (Eppendorf). The injection was completed within 60–220 min after laying. The injected plasmid was mixed with sgRNA plasmid:helper plasmid = 1 × 1 (molar ratio of the substance), and the final concentration was 400 ng/µL. Plasmids (10–15 nL) were injected into each egg. Immediately after injection, the eggs were sealed with nontoxic glue (instant strong glue mini). The eggs were maintained at a relative humidity >85% until they hatched. After hatching, the larvae were reared according to the type of vector injected. The G_0 generation obtained by self-cross or backcross seed production were detected and screened using a macrosteric fluorescence microscope (Olympus DP73) and fed separately.

Pooled genetic screening

After collecting the eggs from the silkworm mutant library, approximately one-fourth of each brood was gathered from every mutant line, totaling approximately 30,000–50,000 eggs. Subsequently, these larvae were uniformly hatched and fed a normal artificial diet until the third instar. The larvae were fed an artificial diet containing 30 mg/kg $CdCl_2$. Subsequently, dead larvae were removed, and survival was supported with a fresh diet. Ultimately only 53 individuals survived until the fifth instar. They were used as donors for the extraction of genomic DNA and sgRNA sequences.

sgRNA sequencing, target site sequencing, and data analysis

The genome was amplified using primers with distinct barcodes and sequenced using Illumina sequencing technology. Low-quality raw sequencing data of amplicons were filtered out using Trimmomatic (v.0.39; LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50) (Bolger et al. 2014). Paired-ended clean reads were merged using FLASH (v.1.2.11) (Magoč and Salzberg 2011). cutadapt (v.1.18) (Marcel 2011) was used to split the data based on the unique barcode of each sample. Bowtie 2 (v.2.2.5) (Langmead and Salzberg 2012) was used for aligning to identify the sgRNA sequences in each sequenced sample. Based on the detected sgRNA, Primer3 (v.2.4.0) (Untergasser et al. 2012) was used to design primers for each target site, amplify the genomes of different samples using diverse primers with specific barcodes, and conduct second-generation sequencing. The analysis methods for filtering out, merging, and splitting of target site sequences were the same as those used in the sgRNA data analysis, whereas mutation detection of target sites was performed using CRISPResso (v.2.0.45) (Clement et al. 2019). Barcode sequences associated with each sample are listed in Supplemental Tables S2 and S3. The m6A data set of silkworm was downloaded from the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA480541.

Cd exposure experiment

To further determine the function of KWMTBOMO12902 in Cd tolerance, we separately treated the KWMTBOMO12902-KO silkworm strain with Cd exposure. Embryos were hatched and fed a fresh artificial diet. Most toxicological studies focus on the third to fifth instar larvae, during which the organs related to detoxification such as the fat body and midgut are fully developed, and the larvae enter a stage of rapid feeding, which facilitates observation and toxicological analysis. Therefore, we choose to conduct exposure experiments starting from the third instar. In the third instar, the artificial diet was supplemented with 30 mg/kg $CdCl_2$. The survival rate was calculated daily. The gRNAs of the surviving individuals were detected from the extracted DNA after Cd exposure. To further investigate the role of KWMTBOMO12902 in Cd tolerance, we exposed KWMTBOMO12902-KO strain silkworms to Cd. Embryos were hatched and fed a fresh artificial diet. In the third instar, the artificial diet was supplemented with 30 mg/kg $CdCl_2$. The survival rate was calculated daily.

Phenotypic characterization

Fourth-instar larvae of 300 strains of the F_1 generation were screened using a stereo-fluorescence microscope (Olympus DP73), and all double-positive individuals (the eye with green light and the body with green light) in each brood were selected for subsequent research. Mutants with significant developmental abnormalities (larval development speed and individual size, larval pattern, pupation process, cocoon color and type, and moth-changing process) were selected. The cocoon weight and pupal weight of all double-positive males and females in each brood were measured separately using an analytical balance, and the sex ratio was counted simultaneously. The cocoon shape was measured for length and breadth using ImageJ software (https://imagej.nih.gov/ij/index.html), and 10 female and 10 male cocoons were randomly measured.

Quantitative real-time PCR

The larvae were dissected to obtain different tissues, and the total RNA from BmE cells or different tissues was extracted. For each sample, 1 µg of RNA was subjected to reverse transcription using the one-step gDNA removal and cDNA synthesis SuperMix kit (TransGen Biotech). Quantitative real-time PCR (qPCR) was performed using SYBR Premix Ex Taq II (Takara), and target gene expression was normalized to the reference gene *rp49*. The test was performed independently at least three times. The primer sequences for the target genes were as follows: KWMTBOMO12902-qPCR-F, TGAAGAGCGAACAAGACGAG; KWMTBOMO12902-qPCR-R, TTCCTCATCGTCAGGCTTTT; Rp49-qPCR-F, CAGGCGGTTCAAGGGTCAATAC; and Rp49-qPCR-R, TGCTGGGCTCTTTCCACGA.

Data access

The raw data generated in this study were submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject)

under accession number PRJNA901072. The mutant library resources are preserved in the Biological Research Center of Southwest University and are shared worldwide through the Silkworm Knowledgebase (SilkDB 3.0; https://silkdb.bioinfotoolkits.net/). The mutant module in the database contains comprehensive phenotype and genotype data for the mutants (Supplemental Fig. S11). The plasmid library can be obtained through Addgene (https://www.addgene.org/209615/).

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank Dr. Feng Wang, Dr. Zhangchuan Peng, Dr. Wenliang Qian, Dr. Kaiyu Guo, Dr. Qingsong Liu, Xia Zhang, and Xiaoxu Chen for their kind help with the microinjection experiments. This work was supported by grants from the National Natural Science Foundation of China (no. 32122084), the Chongqing Natural Science Foundation (no. cstc2021ycjh-bgzxm0005), the PhD Start-up Foundation of Southwest University (no. SWU120012), and the Fundamental Research Funds for Central Universities (no. SWU-KT22042). None of these funding sources played a role in the design of the study, collection, analysis, and interpretation of data, or in writing the manuscript.

Author contributions: S.M. conceived of the study and designed the experiments. T.Z. designed the sgRNA and performed all bioinformatic analyses. R.W. constructed the plasmid library. Y.L. performed the pooled screening experiments. P.W., L.J., and A.W. generated and maintained the transgenic lines with the help of J.C., L.S., H.S., R.S., W.L., D.L., N.Z., W.H., X.W., and W.X. X.L. performed the genotyping experiments. L.J. and S.M. wrote the manuscript with the help of Q.X. All authors have read and approved the final manuscript.

References

Amsterdam A, Burgess S, Golling G, Chen W, Sun Z, Townsend K, Farrington S, Haldi M, Hopkins N. 1999. A large-scale insertional mutagenesis screen in zebrafish. *Genes Dev* **13:** 2713–2724. doi:10.1101/gad.13.20.2713

Atsumi S, Miyamoto K, Yamamoto K, Narukawa J, Kawai S, Sezutsu H, Kobayashi I, Uchino K, Tamura T, Mita K, et al. 2012. Single amino acid mutation in an ATP-binding cassette transporter gene causes resistance to Bt toxin Cry1Ab in the silkworm, *Bombyx mori*. *Proc Natl Acad Sci* **109:** E1591–E1598. doi:10.1073/pnas.1120698109

Beck EA, Healey HM, Small CM, Currey MC, Desvignes T, Cresko WA, Postlethwait JH. 2022. Advancing human disease research with fish evolutionary mutant models. *Trends Genet* **38:** 22–44. doi:10.1016/j.tig.2021.07.002

Bellaiche Y, The I, Perrimon N. 1998. *Tout-velu* is a *Drosophila* homologue of the putative tumour suppressor *EXT-1* and is needed for Hh diffusion. *Nature* **394:** 85–88. doi:10.1038/27932

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* **30:** 2114–2120. doi:10.1093/bioinformatics/btu170

Chang Y, Long T, Wu C. 2012. Effort and contribution of T-DNA insertion mutant library for rice functional genomics research in China: review and perspective. *J Integr Plant Biol* **54:** 953–966. doi:10.1111/j.1744-7909.2012.01171.x

Chang H, Pan Y, Landrette S, Ding S, Yang D, Liu L, Tian L, Chai H, Li P, Li DM, et al. 2019. Efficient genome-wide first-generation phenotypic screening system in mice using the *piggyBac* transposon. *Proc Natl Acad Sci* **116:** 18507–18516. doi:10.1073/pnas.1906354116

Chang J, Wang R, Yu K, Zhang T, Chen X, Liu Y, Shi R, Wang X, Xia Q, Ma S. 2020a. Genome-wide CRISPR screening reveals genes essential for cell viability and resistance to abiotic and biotic stresses in *Bombyx mori*. *Genome Res* **30:** 757–767. doi:10.1101/gr.249045.119

Chang J, Chen X, Zhang T, Wang R, Wang A, Lan X, Zhou Y, Ma S, Xia Q. 2020b. The novel insight into the outcomes of CRISPR/Cas9 editing intra- and inter-species. *Int J Biol Macromol* **163:** 711–717. doi:10.1016/j.ijbiomac.2020.07.039

Chen Y, Wang X. 2022. Evaluation of efficiency prediction algorithms and development of ensemble model for CRISPR/Cas9 gRNA selection. *Bioinformatics* **38:** 5175–5181. doi:10.1093/bioinformatics/btac681

Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, Cole MA, Liu DR, Joung JK, Bauer DE, et al. 2019. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol* **37:** 224–226. doi:10.1038/s41587-019-0032-3

Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339:** 819–823. doi:10.1126/science.1231143

Duan J, Xu H, Ma S, Guo H, Wang F, Zhao P, Xia Q. 2013. Cre-mediated targeted gene activation in the middle silk glands of transgenic silkworms (*Bombyx mori*). *Transgenic Res* **22:** 607–619. doi:10.1007/s11248-012-9677-0

Echeverri K, Fei J, Tanaka EM. 2022. The Axolotl's journey to the modern molecular era. *Curr Top Dev Biol* **147:** 631–658. doi:10.1016/bs.ctdb.2021.12.010

Fields S, Johnston M. 2005. Cell biology. Whither model organism research? *Science* **307:** 1885–1886. doi:10.1126/science.1108872

Ford K, McDonald D, Mali P. 2019. Functional genomics via CRISPR-Cas. *J Mol Biol* **431:** 48–65. doi:10.1016/j.jmb.2018.06.034

Fujii T, Banno Y. 2019. Identification of a novel function of the silkworm integument in nitrogen metabolism: Uric acid is synthesized within the epidermal cells in *B. mori*. *Insect Biochem Mol Biol* **105:** 43–50. doi:10.1016/j.ibmb.2018.12.014

Goldsmith MR, Shimada T, Abe H. 2005. The genetics and genomics of the silkworm, *Bombyx mori*. *Annu Rev Entomol* **50:** 71–100. doi:10.1146/annurev.ento.50.071803.130456

Goldstein B, King N. 2016. The future of cell biology: emerging model organisms. *Trends Cell Biol* **26:** 818–824. doi:10.1016/j.tcb.2016.08.005

Gondo Y. 2008. Trends in large-scale mouse mutagenesis: from genetics to functional genomics. *Nat Rev Genet* **9:** 803–810. doi:10.1038/nrg2431

Haelterman NA, Jiang L, Li Y, Bayat V, Sandoval H, Ugur B, Tan KL, Zhang K, Bei D, Xiong B, et al. 2014. Large-scale identification of chemically induced mutations in *Drosophila melanogaster*. *Genome Res* **24:** 1707–1718. doi:10.1101/gr.174615.114

He J, Zhang K, Yan S, Tang M, Zhou W, Yin Y, Chen K, Zhang C, Li M. 2023. Genome-scale targeted mutagenesis in *Brassica napus* using a pooled CRISPR library. *Genome Res* **33:** 798–809. doi:10.1101/gr.277650.123

Herrera-Estrella L, Depicker A, Van Montagu M, Schell J. 1983. Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature* **303:** 209–213. doi:10.1038/303209a0

Hrabé de Angelis MH, Flaswinkel H, Fuchs H, Rathkolb B, Soewarto D, Marschall S, Heffner S, Pargent W, Wuensch K, Jung M, et al. 2000. Genome-wide, large-scale production of mutant mice by ENU mutagenesis. *Nat Genet* **25:** 444–447. doi:10.1038/78146

Jang D, Meza LR, Greer F, Greer JR. 2013. Fabrication and deformation of three-dimensional hollow ceramic nanostructures. *Nat Mater* **12:** 893–898. doi:10.1038/nmat3738

Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. 2013. RNA-programmed genome editing in human cells. *eLife* **2:** e00471. doi:10.7554/eLife.00471

Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, Ishihara G, Kawaoka S, Sugano S, Shimada T, et al. 2014. A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature* **509:** 633–636. doi:10.1038/nature13315

Kolliopoulou A, Swevers L. 2014. Recent progress in RNAi research in lepidoptera: intracellular machinery, antiviral immune response and prospects for insect pest control. *Curr Opin Insect Sci* **6:** 28–34. doi:10.1016/j.cois.2014.09.019

Lang M, Murat S, Clark AG, Gouppil G, Blais C, Matzkin LM, Guittard E, Yoshiyama-Yanagawa T, Kataoka H, Niwa R, et al. 2012. Mutations in the *neverland* gene turned *Drosophila pachea* into an obligate specialist species. *Science* **337:** 1658–1661. doi:10.1126/science.1224829

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9:** 357–359. doi:10.1038/nmeth.1923

Lein W, Usadel B, Stitt M, Reindl A, Ehrhardt T, Sonnewald U, Börnke F. 2008. Large-scale phenotyping of transgenic tobacco plants (*Nicotiana tabacum*) to identify essential leaf functions. *Plant Biotechnol J* **6:** 246–263. doi:10.1111/j.1467-7652.2007.00313.x

Li Y, Ma S, Sun L, Zhang T, Chang J, Lu W, Chen X, Liu Y, Wang X, Shi R, et al. 2018. Programmable single and multiplex base-editing in *Bombyx mori* using RNA-guided cytidine deaminases. *G3 (Bethesda)* **8:** 1701–1709. doi:10.1534/g3.118.200134

Li Z, Wang H, Cai C, Wong AH, Wang J, Gao J, Wang Y. 2020. Genome-wide *piggyBac* transposon-based mutagenesis and quantitative insertion-site analysis in haploid *Candida* species. *Nat Protoc* **15:** 2705–2727. doi:10.1038/s41596-020-0351-3

Li C, Chu W, Gill RA, Sang S, Shi Y, Hu X, Yang Y, Zaman QU, Zhang B. 2023. Computational tools and resources for CRISPR/Cas genome editing. *Genomics Proteomics Bioinformatics* **21:** 108–126. doi:10.1016/j.gpb.2022.02.006

Liu C, Yamamoto K, Cheng TC, Kadono-Okuda K, Narukawa J, Liu SP, Han Y, Futahashi R, Kidokoro K, Noda H, et al. 2010. Repression of tyrosine hydroxylase is responsible for the sex-linked chocolate mutation of the silkworm, *Bombyx mori*. *Proc Natl Acad Sci* **107:** 12980–12985. doi:10.1073/pnas.1001725107

Liu Y, Ma S, Wang X, Chang J, Gao J, Shi R, Zhang J, Lu W, Liu Y, Zhao P, et al. 2014. Highly efficient multiplex targeted mutagenesis and genomic structure variation in Bombyx mori cells using CRISPR/Cas9. *Insect Biochem Mol Biol* **49:** 35–42. doi:10.1016/j.ibmb.2014.03.010

Liu Y, Chang J, Yang C, Zhang T, Chen X, Shi R, Liang Y, Xia Q, Ma S. 2021. Genome-wide CRISPR-Cas9 screening in Bombyx mori reveals the toxicological mechanisms of environmental pollutants, fluoride and cadmium. *J Hazard Mater* **410:** 124666. doi:10.1016/j.jhazmat.2020.124666

Lu Y, Ye X, Guo R, Huang J, Wang W, Tang J, Tan L, Zhu JK, Chu C, Qian Y. 2017. Genome-wide targeted mutagenesis in rice using the CRISPR/Cas9 system. *Mol Plant* **10:** 1242–1245. doi:10.1016/j.molp.2017.06.007

Lu F, Wei Z, Luo Y, Guo H, Zhang G, Xia Q, Wang Y. 2020. SilkDB 3.0: visualizing and exploring multiple levels of data for silkworm. *Nucleic Acids Res* **48:** D749–D755. doi:10.1093/nar/gkz919

Ma S, Zhang S, Wang F, Liu Y, Liu Y, Xu H, Liu C, Lin Y, Zhao P, Xia Q. 2012. Highly efficient and specific genome editing in silkworm using custom TALENs. *PLoS One* **7:** e45035. doi:10.1371/journal.pone.0045035

Ma S, Wang X, Fei J, Liu Y, Duan J, Wang F, Xu H, Zhao P, Xia Q. 2013. Genetic marking of sex using a W chromosome-linked transgene. *Insect Biochem Mol Biol* **43:** 1079–1086. doi:10.1016/j.ibmb.2013.08.009

Ma S, Chang J, Wang X, Liu Y, Zhang J, Lu W, Gao J, Shi R, Zhao P, Xia Q. 2014a. CRISPR/Cas9 mediated multiplex genome editing and heritable mutagenesis of *BmKu70* in *Bombyx mori*. *Sci Rep* **4:** 4489. doi:10.1038/srep04489

Ma S, Shi R, Wang X, Liu Y, Chang J, Gao J, Lu W, Zhang J, Zhao P, Xia Q. 2014b. Genome editing of BmFib-H gene provides an empty Bombyx mori silk gland for a highly efficient bioreactor. *Sci Rep* **4:** 6867. doi:10.1038/srep06867

Ma SY, Smagghe G, Xia QY. 2019. Genome editing in *Bombyx mori*: new opportunities for silkworm functional genomics and the sericulture industry. *Insect Sci* **26:** 964–972. doi:10.1111/1744-7917.12609

Ma S, Wang A, Chen X, Zhang T, Xing W, Xia Q. 2021. Deep sequencing reveals the comprehensive CRISPR-Cas9 editing spectrum in *Bombyx mori*. *CRISPR J* **4:** 371–380. doi:10.1089/crispr.2021.0003

Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27:** 2957–2963. doi:10.1093/bioinformatics/btr507

Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. 2013. RNA-guided human genome engineering via Cas9. *Science* **339:** 823–826. doi:10.1126/science.1232033

Mao T, Li F, Fang Y, Wang H, Chen J, Li M, Lu Z, Qu J, Li J, Hu J, et al. 2019. Effects of chlorantraniliprole exposure on detoxification enzyme activities and detoxification-related gene expression in the fat body of the silkworm, Bombyx mori. *Ecotoxicol Environ Saf* **176:** 58–63. doi:10.1016/j.ecoenv.2019.03.074

Marcel M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet* **17:** 10. doi:10.14806/ej.17.1.200

Meltzer H, Marom E, Alyagor I, Mayseless O, Berkun V, Segal-Gilboa N, Unger T, Luginbuhl D, Schuldiner O. 2019. Tissue-specific (ts)CRISPR as an efficient strategy for in vivo screening in *Drosophila*. *Nat Commun* **10:** 2113. doi:10.1038/s41467-019-10140-0

Meng X, Yu H, Zhang Y, Zhuang F, Song X, Gao S, Gao C, Li J. 2017a. Construction of a genome-wide mutant library in rice using CRISPR/Cas9. *Mol Plant* **10:** 1238–1241. doi:10.1016/j.molp.2017.06.006

Meng X, Zhu F, Chen K. 2017b. Silkworm: a promising model organism in life science. *J Insect Sci* **17:** 97. doi:10.1093/jisesa/iex064

Nolan PM, Peters J, Strivens M, Rogers D, Hagan J, Spurr N, Gray IC, Vizor L, Brooker D, Whitehill E, et al. 2000. A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse. *Nat Genet* **25:** 440–443. doi:10.1038/78140

Paine I, Posey JE, Grochowski CM, Jhangiani SN, Rosenheck S, Kleyner R, Marmorale T, Yoon M, Wang K, Robison R, et al. 2019. Paralog studies augment gene discovery: *DDX* and *DHX* genes. *Am J Hum Genet* **105:** 302–316. doi:10.1016/j.ajhg.2019.06.001

Port F, Strein C, Stricker M, Rauscher B, Heigwer F, Zhou J, Beyersdörffer C, Frei J, Hess A, Kern K, et al. 2020. A large-scale resource for tissue-specific CRISPR mutagenesis in *Drosophila*. *eLife* **9:** e53865. doi:10.7554/eLife.53865

Qi X, Zhang J, Zhao Y, Chen T, Xiang Y, Hui J, Cai D, Liu Y, Xia L, Yu T, et al. 2017. The applications of CRISPR screen in functional genomics. *Brief Funct Genomics* **16:** 34–37. doi:10.1093/bfgp/elw020

Russell JJ, Theriot JA, Sood P, Marshall WF, Landweber LF, Fritz-Laylin L, Polka JK, Oliferenko S, Gerbich T, Gladfelter A, et al. 2017. Non-model model organisms. *BMC Biol* **15:** 55. doi:10.1186/s12915-017-0391-5

Schimmel J, Muñoz-Subirana N, Kool H, van Schendel R, Tijsterman M. 2021. Small tandem DNA duplications result from CST-guided Pol α-primase action at DNA break termini. *Nat Commun* **12:** 4843. doi:10.1038/s41467-021-25154-w

Segala G, Bennesch MA, Ghahhari NM, Pandey DP, Echeverria PC, Karch F, Maeda RK, Picard D. 2019. Vps11 and Vps18 of Vps-C membrane traffic complexes are E3 ubiquitin ligases and fine-tune signalling. *Nat Commun* **10:** 1833. doi:10.1038/s41467-019-09800-y

Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, et al. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* **343:** 84–87. doi:10.1126/science.1247005

Stemmer M, Thumberger T, Del Sol Keyer M, Wittbrodt J, Mateo JL. 2015. CCTop: an intuitive, flexible and reliable CRISPR/Cas9 target prediction tool. *PLoS One* **10:** e0124633. doi:10.1371/journal.pone.0124633

Sun Y, Zhang B, Luo L, Shi DL, Wang H, Cui Z, Huang H, Cao Y, Shu X, Zhang W, et al. 2020. Systematic gene editing of the genes on zebrafish Chromosome 1 by CRISPR/Cas9. *Genome Res* **30:** 118–126. doi:10.1101/gr.248559.119

Tamura T, Thibert C, Royer C, Kanda T, Abraham E, Kamba M, Komoto N, Thomas JL, Mauchamp B, Chavancy G, et al. 2000. Germline transformation of the silkworm *Bombyx mori* L. using a *piggyBac* transposon-derived vector. *Nat Biotechnol* **18:** 81–84. doi:10.1038/71978

Terenius O, Papanicolaou A, Garbutt JS, Eleftherianos I, Huvenne H, Kanginakudru S, Albrechtsen M, An C, Aymeric JL, Barthel A, et al. 2011. RNA interference in lepidoptera: an overview of successful and unsuccessful studies and implications for experimental design. *J Insect Physiol* **57:** 231–245. doi:10.1016/j.jinsphys.2010.11.006

Tong X, Han MJ, Lu K, Tai S, Liang S, Liu Y, Hu H, Shen J, Long A, Zhan C, et al. 2022. High-resolution silkworm pan-genome provides genetic insights into artificial selection and ecological adaptation. *Nat Commun* **13:** 5619. doi:10.1038/s41467-022-33366-x

Umen J, Herron MD. 2021. Green algal models for multicellularity. *Annu Rev Genet* **55:** 603–632. doi:10.1146/annurev-genet-032321-091533

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3—new capabilities and interfaces. *Nucleic Acids Res* **40:** e115. doi:10.1093/nar/gks596

Venken KJ, Bellen HJ. 2014. Chemical mutagens, transposons, and transgenes to interrogate gene function in *Drosophila melanogaster*. *Methods* **68:** 15–28. doi:10.1016/j.ymeth.2014.02.025

Wang T, Wei JJ, Sabatini DM, Lander ES. 2014. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* **343:** 80–84. doi:10.1126/science.1246981

Wendland J, Walther A. 2005. *Ashbya gossypii*: a model for fungal developmental biology. *Nat Rev Microbiol* **3:** 421–429. doi:10.1038/nrmicro1148

Wu S, Ying G, Wu Q, Capecchi MR. 2007. Toward simpler and faster genome-wide mutagenesis in mice. *Nat Genet* **39:** 922–930. doi:10.1038/ng2060

Xia Q, Li S, Feng Q. 2014. Advances in silkworm studies accelerated by the genome sequencing of *Bombyx mori*. *Annu Rev Entomol* **59:** 513–536. doi:10.1146/annurev-ento-011613-161940

Xu R, Deng K, Zhu Y, Wu Y, Ren J, Wan M, Zhao S, Wu X, Han M, Zhuang Y, et al. 2008. A large-scale functional approach to uncover human genes and pathways in *Drosophila*. *Cell Res* **18:** 1114–1127. doi:10.1038/cr.2008.295

Xu J, Chen S, Zeng B, James AA, Tan A, Huang Y. 2017. Bombyx mori P-element somatic inhibitor (BmPSI) is a key auxiliary factor for silkworm male sex determination. *PLoS Genet* **13:** e1006576. doi:10.1371/journal.pgen.1006576

Yamaguchi J, Banno Y, Mita K, Yamamoto K, Ando T, Fujiwara H. 2013. Periodic Wnt1 expression in response to ecdysteroid generates twin-spot markings on caterpillars. *Nat Commun* **4:** 1857. doi:10.1038/ncomms2778

Yang L, Wang J, Yang Y, Li S, Wang T, Oleksak P, Chrienova Z, Wu Q, Nepovimova E, Zhang X, et al. 2022. Phytoremediation of heavy metal pollution: hotspots and future prospects. *Ecotoxicol Environ Saf* **234:** 113403. doi:10.1016/j.ecoenv.2022.113403

Yoshiyama T, Namiki T, Mita K, Kataoka H, Niwa R. 2006. Neverland is an evolutionally conserved Rieske-domain protein that is essential for ecdysone synthesis and insect growth. *Development* **133:** 2565–2574. doi:10.1242/dev.02428

Zhang Z, Niu B, Ji D, Li M, Li K, James AA, Tan A, Huang Y. 2018. Silkworm genetic sexing through W chromosome-linked, targeted gene integration. *Proc Natl Acad Sci* **115:** 8752–8756. doi:10.1073/pnas.1810945115

Zirin J, Hu Y, Liu L, Yang-Zhou D, Colbeth R, Yan D, Ewen-Campen B, Tao R, Vogt E, VanNest S, et al. 2020. Large-scale transgenic *Drosophila* resource collections for loss- and gain-of-function studies. *Genetics* **214:** 755–767. doi:10.1534/genetics.119.302964

Received August 2, 2023; accepted in revised form November 29, 2023.