CAMIO for deletion analysis of endogenous DNA sequences in multicellular organisms

Hui-Min Chen[1], Jorge Garcia Marques[1], Ken Sugino[1], Dingjun Wei[1], Rosa Linda Miyares[1] and Tzumin Lee[1,*]

[1] Howard Hughes Medical Institute, Janelia Research Campus, 19700 Helix Drive, Ashburn, VA 20147, USA

* To whom correspondence should be addressed.
Tel: 571-209-4198
Email: leet@janelia.hhmi.org

ABSTRACT

Interrogating the genome, especially locating critical cis-elements in non-protein coding regions, requires efficient deletion analysis. Conventional deletion analyses using synthetic constructs are cumbersome and nonphysiological. Here we invent CAMIO (Cas9-mediated Arrayed Mutagenesis of Individual Offspring), to achieve high-throughput organism-level structure-functional analysis of native DNA sequences. CAMIO permits creating arrays of targeted deletions directly in the genome. First, we assemble mutually exclusive gRNAs into a set. Combining two sets allows stochastic expression of many possible combinations of two gRNAs in germline stem cells. Second, we restrict Cas9 to male germ cells to elicit independent deletions around or between the gRNA-directed Cas9 cut sites in individual offspring. Thus, we can obtain from a single population cross a matrix of variable deletions covering a sizable DNA region at both coarse and fine levels. We exemplify the power of CAMIO in mapping the 5'UTR sequences crucial for graded neuronal expression of *chinmo*.

INTRODUCTION

It never ceases to amaze us how genomes, sequences of seemingly simple four nucleotides, serves as the fundamental blueprints to build up layers and layers of complexity and form full functioning living organisms. The central dogma depicts the basic information flow of biology from DNA to RNA and protein. Numerous mechanisms further diversify the outputs, such as epigenetic effects, promoters and enhancers, alternative splicing, post-transcriptional and post-translational modifications, to name just a few. Understanding how complex biology unfolds, advances, and evolves from genomic sequences requires development of precise and efficient tools for interrogating the genome. While spontaneous or induced mutations once begot our understanding of many biological processes,

we now desire sophisticated methods to directly manipulate genomes for systematic mechanistic studies.

Pioneering geneticists, like Dr. Barbara McClintock, enlightened us with fundamental understandings of genetic information and regulation. More recently, molecular biologists relied heavily on *in vitro* or cell culture-based assays to assess functions of DNA fragments and understand gene regulations. One example is promoter bashing that facilitates pinpointing critical or regulatory sequence elements in a promoter (1,2). Also, larger DNA fragments upstream or downstream of a target gene were fused to reporter genes in enhancer screening (3-5). This synthetic approach often fails to recapitulate full endogenous patterns, and thus requires a second-round rescue experiment to draw solid conclusions. Such indirect methods were sufficient but not ideal, prompting the search for tailor-made sequence-specific DNA nucleases that permit targeted mutagenesis of native sequences.

In the early 2000s, zinc-finger nuclease (ZFN) (6), the first of the tailored nucleases, was heavily adopted (7,8). Although it seemed a promising strategy for genome editing, major setbacks, especially off-target toxicity limited its utility/value (9). Therefore, the enthusiasm for ZFN declined after the arrival of transcription activator-like effector nuclease (TALEN) (10-12). The higher specificity of TALENs led to fewer off-target disruptions, and hence less cytotoxicity (13). However, constructing a TALEN is challenging because the homologous sequences encoding the TALEN repeats are prone to recombine with each other (9).

The newest technology, CRISPR (clustered regularly interspaced short palindromic repeats) was first exploited for targeted mutagenesis (14), and swiftly adopted to edit genomes of diverse organisms (15-20). CRISPR's popularity lies in its simplicity: a Cas9 nuclease plus an easily made gRNA to create a targeted double-strand break (DSB) through DNA-RNA base pairing (14,16). Cells repair the DSBs either through nonhomologous end joining (NHEJ) (21) incurring insertion or deletion (indel), or through homology-directed repair (HDR) (22) replacing discontinuous sequences with an available template. Such targeted DNA modifications with unlimited specificity empower the rapidly expanding CRISPR technology. Indeed, CRISPR has already transformed genome studies in all biomedical fields from stem cell research to cancer biology. (23).

Drosophila melanogaster has been a powerful model organism for decoding the genome to understand complex biology (24). However, it remains quite challenging to interrogate the entire fly genome, especially regions of non-protein coding sequences, with existing fly genetic tools. Enhancing fly genetics with CRISPR is particularly needed for large-scale genome-wide screens as well as focused detailed sequence analyses (25).

Here we describe a new high-throughput technology for deletion analysis called

CAMIO (Cas9-mediated Arrayed Mutagenesis of Individual Offspring). We built a CRISPR-based mutagenesis pipeline in *Drosophila* male germ cells, to achieve massive production of independent indels in targeted loci with germline transmission (26,27). We further created a transgenic system for targeting multiple sites simultaneously with an array of gRNAs. This way, we can readily generate a huge collection of organisms harboring diverse small localized or large defined deletions for efficient deletion analysis of sizable genomic regions in vivo. We exemplified the power of CAMIO by promptly ascribing an intriguing temporal control of neuronal fate diversification to a 300-bp sequence in the 2-kb 5'UTR of *chinmo* (28,29). Taken together, CAMIO enables high-throughput organism-level genome structure-functional studies.

MATERIAL AND METHODS

Fly strains.
We used the following fly strains in this work: (1) *bamP(898)-Cas9* in *attP2*; (2) *U6:3-gRNA-e* (19); (3) *Df(3R)ED10838/TM2* (BDSC #9485); (4) *dU6-3-gRNA-shi* in *attP40*; (5) *UAS-shibire* in *attP2*; (6) *GMR-Gal4*; (7) *nos-phiC31-nls #12*; (8) *10XUAS-mCD8::GFP* in *attP40* (30); (9) *act5C-Gal4/TM6B*; (10) *pCis-{4gRNAs_mCD8}* P element insertion line #3, #10 on II, and #9, #12 on III; (11) *13XLexAop2-5' UTR-smGFP-OLLAS* in *VK00027*, *13XLexAop2 –smGFP-cMyc-3' UTR* in *attP40*, and *13XLexAop2-5' UTR-smGFP-V5-3' UTR* in *su(Hw)attP5*; (12) *41A10-KD*; (13) *DpnEE-KO-LexAp65*; (14) *pCis-{4gRNAs_chinmo Exon1}* P element insertion line #23, #25 on II, and #4, #5 on III; *pCis-{6gRNAs_chinmo Exon2}* #7, #11 on II, and #14 on III; *pCis-{4gRNAs_chinmo Exon3}* #5 on II, and #2, #3 on III; (15) *FRT40A,UAS–mCD8::GFP,UAS–Cd2-Mir/CyO,Y* (31); (16) *OK107-Gal4*; (17) *UAS-Cas9* (19); (18) *dU6_g2+3* in *VK00027*; (19) *Dpn-Cas9* (unpublished reagent); (20) *act5C-Cas9* (19); (21) *13XLexAop2-53UTR-smGFP-V5-d*, *13XLexAop2-53UTR-smGFP-V5-bD*, and *13XLexAop2-53UTR-smGFP-V5-D23* in *attP40*.

Molecular biology.
The full *bam* promoter (-898) (26) was ordered from gBlocks, IDT to create *bamP(898)-Cas9*, which is also flanked by *bam* 3' UTR. Codon-optimized *shibire* coding sequence (CDS) that carries the gRNA-shi target site was ordered from GeneArt gene synthesis, and then cloned into pJFRC28 (32) to create UAS-shibire. To generate dU6-3-gRNAs, we replaced 10XUAS-IVS-GFP-p10 of pJFRC28 with dU6-3 promter and gRNA scaffold fragment from pTL2 (27). For dU6-3-gRNA-shi, GTATGGGGTATCAAGCCGAT was selected as the spacer. To create dU6_g2+3, we first generated dU6-3-g2_2 and dU6-3-g2_3 separately, and cloned dU6-3-g2_2 into the backbone of dU6-3-g2_3.

pCis-{4gRNAs_mCD8}: a U6 promoter-AttB fragment was synthetized by PCR amplification from pCFD3 (19) and cloned into pCaST-elav-VP16AD, which contained the p-Element inverted repeats (Addgene, #15308). Then, we inserted

a DNA fragment (Genscript) containing 4 different gRNAs targeting the mCD8 protein tag. These gRNAs were selected based on their ON and OFF target scores (Benchling). Each of these gRNAs was preceded by an AttP site and a HammerHead ribozyme (33). Finally, a 3Xp3-RFP-polyA(α-tubulin) fragment was synthetized by PCR amplification, using pure genomic DNA from a fly line in which this cassette was used as a marker (Bloomington, #54590). Then, this fragment was inserted upstream of this gRNA region. In the final construct, the AttB and AttP sites were separated by a 3.7 Kb region containing an ampicillin resistance gene, an origin of replication in bacteria and the 3Xp3-RFP-polyA marker.

pCis-{gRNAs_Chinmo-Exon1/Exon2/Exon3}: following the same design described above, a DNA fragment was synthetized (Genscript), which contained 4 gRNAs (6 for Exon2) either targeting the corresponding exon or the exon-intron junction. This fragment was then inserted into pCis-{4gRNAs_mCD8}, thus removing the previous gRNAs cassette.

The *Chinmo* UTRs were amplified from Drosophila genomic DNA. DarkGFP fused to V5, cmyc or ollas were amplified from previously existing plasmids. Standard molecular biology techniques were used to clone the DarkGFP fusions containing one or two *Chinmo* UTRs into 13XLexAop2 (pJFRC15, (30)). If the construct ended with *Chinmo* 3'UTR, the SV40 signal was removed from the vector backbone. *13XLexAop2-5' UTR-GFP-3' UTR* was further modified to create d, bD, and D23 reporters containing various deletions in 5' UTR.

Drosophila genetics.

***ebony* and UAS-shibire mutagenesis by *bamP(898)-Cas9*.** Female or male founders (*bamP(898)-Cas9/U6:3-gRNA-e*) were mated to *Df(3R)ED10838/TM6B*, and chromosomes over *Df(3R)ED10838* were scored for ebony loss-of-function phenotype. In total, 709 progeny from 5 female frounders and 1574 progeny from 9 male founders were collected and phenotyped. Male *UAS-shibire* mutagenesis founders were crossed with *GMR-Gal4*, and the wildtype-eyed 919 progeny were sacrificed for next generation sequencing (NGS).

CAMIO on UAS-mCD8::GFP. Male founders (*UAS-mCD8::GFP*, *bamP(898)-Cas9*, *nos-phiC31*, and one or two copies of *pCis-{4gRNAs_mCD8}* were mated with *act5C-Gal4* females for scoring of loss of green fluorescence in the progeny. For one copy of *pCis-{4gRNAs_mCD8}*, we screened 807 progeny from 20 founder males. 30 loss-of-GFP progeny from two founders were further subject to sequence analysis. For two copies of *pCis-{4gRNAs_mCD8}*, we screened 570 progeny from 16 founders. 110 loss-of-GFP progeny were analyzed and grouped into three deletion categories.

CAMIO on *chinmo* 5' UTR. After mating with females carrying second chromosome balancer for stock keeping, 1200 male progeny from 4 CAMIO on *chinmo* gRNA array combinations were sacrificed for genomic study. We designed primer sets that produce amplicons covering exon 1, 2, 3, or exons

1+2. Males from each combination were intentionally numbered 1-300, and their genomic amplicons were carefully matched and mixed with counterparts from other 2 combinations. Finally, 300 DNA mixtures plus 84 amplicons from combination 4 were tagmented (Nextera XT DNA Library Prep Kit, illumina) and barcoded (Nextera XT Index Kit v2) for NGS.

Immunohistochemistry and confocal imaging.
Fly brains at indicated larval, pupal, and adult stages were dissected, fixed, and immunostained as described previously (28,31). The following primary antibodies were used in this study: chicken GFP polyclonal antibody (1:500, Invitrogen, A10262); rat anti-Deadpan (1:100, abcam, ab195173); mouse 1D4 anti-Fasciclin II (1:50, DSHB); rabbit polyclonal anti-Trio (1:1000) (34); rat anti-Chinmo (1:500, a gift from the Sokol lab) (35). Secondary antibodies from Molecular Probes were all used in a 1:200 dilution. The immunofluorescent signals were collected using Zeiss LSM 880 confocal microscope and processed using Fiji and Adobe Illustrator.

Bioinformatics.
Sequence reads (FASTQ data) are first processed with cutadatp [https://github.com/marcelm/cutadapt] to remove adapter sequences with options: --overlap=7 --minimum-length=30 -a "CTGTCTCTTATACACATCTCTGAGCGGGCTGGCAAGGCAGACCG". Then they were mapped to genomic sequence corresponding to *chinmo* region using Bowtie2 (36) with the following options: --local --score-min G,20,0 -D 20 -R 3 -L 20 -i S,1,0.50 --no-unal. Resulting SAM files are parsed with Pysam [https://github.com/pysam-developers/pysam] to extract deletion/insertion information from the Cigar strings. When the cigar string contains ‘D’ or ‘N’, we extracted mapped sequence as deletion and designated as type D. When the cigar string contains ‘I’, we extracted it as insertion and designated as type I. While parsing the SAM file, soft clipped reads (reads partially mapped to the genome) are detected and clipped (unmapped) portions are set aside in a FASTA file. Then these sequences in this FASTA file are mapped using bowtie2 again to the genomic sequence encompassing *chinmo* gene. Then, mapped portions of the first mapping and that of the second mapping (when the second one exists) are merged to form a deletion event which we denote as type L (large gap). This type of gap often contained inserted sequences in the middle. We discarded events with less than 10 reads.

RESULTS

Independent targeted mutagenesis in individual male germ cells

We had shown in the original Golic+ paper that *bam* promoter could effectively drive flippase induction in female germ cells while sparing germline stem cells (GSCs) (27). The restriction to germ cells suggests independent gene targeting

in each offspring with Golic+. To rigorously examine this proposition, we exploited Cas9-mediated mutagenesis which should have sufficient diversity to confer independently mutated offspring with different mutations.

To this end, we made *bamP(898)-Cas9* and tested its ability to mutate gRNA-targeted sites in female as well as male germline. We chose the *ebony* gene for targeted mutagenesis because of easy detection of loss-of-*ebony* mutations with *ebony* deficiency and the availability of transgenic gRNA targeting *ebony* (Fig. 1A) (19). First, we determined the mutagenesis efficiency in male versus female founders (Fig. 1B). We recovered loss-of-function (LoF) *ebony* mutations in over 25% of male gametes as opposed to only ~3% of female gametes. This result demonstrates that the male germline is particularly susceptible to *bamP-Cas9*-mediated genome editing.

To examine if the Cas9-mediated editing events did occur independently, we next sequenced a part of the *ebony* locus in individual mutants carrying *ebony* deficiency. From a single male founder, we recovered ten mutants that show eight distinct indels around the Cas9 cut site (Fig. 1C). The three siblings with an identical *ebony* mutation could result from differential deletions of GTC repeats, or from microhomology-mediated repair (37).

The recovery of many distinct indels from a single founder argues for independent Cas9 actions in individual germ cells. This result encouraged us to establish *bamP* induced CRISPR in male germline as a targeted mutagenesis pipeline. Indels of various sizes uncovering the same Cas9 cut site within protein-coding sequences likely cripple the targeted protein comparably (e.g. creating truncated proteins of similar sizes). Nevertheless, we explored the possibility of making dominant-negative proteins via targeted CRISPR. *shibire* encodes *Drosophila* Dynamin, a motor protein crucial for synaptic vesicle endocytosis (38), and shi^{ts1} with a missense mutation is widely used in fly behavior assays to shut off neuronal activity with raised temperature (39). We wondered if targeted mutagenesis around the shi^{ts1} mutation site could permit recovery of additional useful *shi* alleles.

We designed a gRNA against a *UAS-shibire* transgene at the region where the temperature-sensitive shi^{ts1} point mutation is located. Male founders (*bamP(898)-Cas9*, *dU6-3-gRNA-shi*, *UAS-shibire*) were mated with *GMR-Gal4*, and the progeny were raised at 29°C for us to screen for *UAS-shibire* indel alleles causing rough eye phenotypes. We observed a wide spectrum of compound eye roughness even among the progeny carrying the same UAS-shibire allele. The evident variability in the strength of *GMR-Gal4* especially at a higher temperature severely compromised our ability to identify interesting alleles. Therefore, we resorted to surveying normal-looking offspring instead, which should be selectively devoid of dominant mutations. We pooled around 1000 phenotypically normal progeny collected from 20 male founders for amplicon analysis with next generation sequencing (NGS). We obtained a big collection of diverse indels,

with the majority of deletions smaller than 30 bps (Fig. 1D). Notably, there is a clear under-representation of in-frame indels (Fig. 1E). This observation strongly supports the notion that lacking some amino acids around the *shi*^{ts1} point mutation site makes the otherwise full-length Shibire toxic.

Taken together, our data demonstrate that *bamP(898)* effectively restricts Cas9-mediated mutagenesis to germ cells. There is no evidence for clonal expansion contributing to the exceptionally high mutation efficiency in male founders. Therefore, transgenic CRISPR induced by bamP can effectively serves as a pipeline for mass production of targeted mutations.

CAMIO: Cas9-mediated arrayed mutagenesis of individual offspring

Despite independent mutagenesis in each germ cell, the derived offspring diversity with only one gRNA is rather limited as all indels were anchored around the same Cas9 cut site. To expand the diversity of deletions that one can recover from a single population cross, we next explored the possibility of multiplexing single-gRNA mutagenesis. We envision in male germ lines expression of various combinatorial gRNAs stochastically selected from a gRNA repertoire rather than all gRNAs of the repertoire simultaneously. Supplying random gRNA subsets prevents contamination of rare deletions by a much more frequent second-site deletion. This way, discrete clusters of simple deletions can be recovered from a single population cross. Various clusters would consist of deletions induced by different gRNAs. Therefore, we can achieve tiling a sizable DNA region of interest with diverse small deletions with a repertoire of evenly spaced gRNAs.

We targeted a *UAS-mCD8::GFP* transgene with four independent gRNAs to examine the feasibility of our multiplexing design. For stochastically activating only one out of the four gRNAs, we made a conditional U6-gRNAs that can be restored by phiC31-mediated recombination (Fig. 2A). A single U6 promoter followed by an attB site is separated from multiple attP-preceding gRNAs by an eye reporter gene (Fig. 2A). Upon induction of the phiC31 recombinase in GSCs using the *nos* promoter, the lone attB site can irreversibly recombine with one of the following attP sites. The intra-chromosomal recombination excises the intervening eye reporter gene and reconstitutes a functional U6-gRNA transgene. Given the rather small size of each gRNA as compared to the large eye reporter gene, the differences in the length of intervening DNA sequences for the activation of the tandemly arranged gRNAs are relatively trivial. We are therefore hopeful that each gRNA will be expressed at comparable frequencies, based on a previous construct for multicolor imaging (40). For brevity, we name this conditional U6-gRNA transgene as, pCis-{4gRNAs_targeted DNA}, to indicate a set of four mutually exclusive gRNAs that are further numbered in sequence from 5' to 3'.

For the multiplex targeted mutagenesis of mCD8::GFP, we established male founders carrying *UAS-mCD8::GFP*, *bamP(898)-Cas9*, *nos-phiC31*, and *pCis-*

{4gRNAs_mCD8}, and crossed them to *act5C-Gal4* females for easy scoring of GFP fluorescence in the progeny. Overall, approximately 35% of the progeny lost GFP expression. We collected 30 GFP-negative offspring from two founder males. Sequencing the mCD8-coding region revealed that each GFP-negative offspring carried an indel corresponding to a chosen gRNA (Supplementary Fig. 1). Encouragingly, we recovered various deletions resulting from activation of each of the four gRNAs. However, the manifested frequency of mutations resulting from activation of distinct gRNAs varied. Both founders yielded many more deletions around the gRNA#1/#4 targets than the gRNA#2/#3 targets, possibly reflecting their different on-target potencies.

As to the extent of deletions, we found that the majority (83.3%) of deletions uncovered 20 or fewer bps and that the largest one lost 85 bps. To tile a sizable DNA region with such small deletions would require many gRNAs bombarding the target at a density of around one gRNA per 100 bps. Alternatively, we should be able to create larger deletions spanning two Cas9 cuts elicited by two gRNAs acting in distance. We realized this possibility by providing two copies of *pCis-{4gRNAs_mCD8}* for multiplexed dual mutagenesis of *UAS-mCD8::GFP* (Fig. 2B). We obtained a comparable loss-of-GFP mutation rate at ~35% despite co-expressing two identical or distinct U6-gRNAs. This phenomenon implicates Cas9 activity (either level or duration) limiting the efficiency of gRNA-directed mutagenesis in germ cells. Nonetheless, we could recover from the dual gRNA-derived GFP-negative progeny further diverse mutations, including many single-site deletions (76.4%) and quite a few double-site deletions (15.4%) as well as some big deletions spanning two gRNA target sites (8.2%) (Fig. 2C). Notably, the single-site deletions greatly outnumbered those involving two sites. This outcome is favored, as it would increase the chance of recovering some sizable deletions without a second-site contamination.

With dual U6-gRNA sets, we can further create various large deletions defined by pairs of distant gRNAs. Such defined deletions occurred in about 3% (8.2% of 35%) of the total progeny. Random assortment from two identical copies of four mutually exclusive gRNAs should yield six types of defined deletions. Encouragingly, we recovered five of the six anticipated kinds of defined deletions from a collection of only nine defined deletions. In contrast with the variable efficiencies in mutating GFP by distinct single gRNAs, all four gRNAs appeared equally potent in acting together to delete the intervening sequences despite the size ranging from ~100 to ~250 bps. Such size-independent performance permits flexible selection of gRNA targets in systematic DNA deletion analysis with dual gRNAs. However, physical hindrance may prevent two Cas9 complexes from acting simultaneously on very close gRNA targets. This might explain why we failed to recover the smallest defined deletion of 37 bps between the Cas9 cut sites of gRNA#2/#3 targets.

In sum, we established an effective strategy for expressing various combinations of two gRNAs in GSCs. In combination with restricting Cas9 to germ cells, we

built a germline pipeline for multiplex targeted mutagenesis. We dub this genetic system as CAMIO, which can derive from a single population cross a matrix of variable deletions spanning a sizable DNA region at both coarse and fine levels. Next, we exemplified the power of CAMIO in structure-functional analysis of a 2.2kb-long genomic fragment.

Structure-functional analysis of *chinmo* 5'UTR

Drosophila mushroom body (MB) is composed of γ, α'β', and αβ neurons, which arise sequentially from common MB neuroblasts during embryonic, larval and pupal development (41). A descending gradient of Chinmo proteins expressed among serially derived MB neurons guides the neuronal temporal fate diversification (28). Multiple lines of evidence, obtained from in situ hybridization as well as reporter transgenes with or without endogenous untranslated region (UTR), led us to believe there exists a strong post-transcriptional control through *chinmo* 5' UTR for its protein expression (28,42). However, a much longer *chinmo* 3'UTR (~8.5 kb) was later annotated and further identified as the site of action for miRNAs let-7 and miR-125 to down-regulate Chinmo expression (35). To clarify which UTR plays a more significant role, we generated transgenic 13XLexAop2-GFP reporters flanked by different combinations of *chinmo* 5' and 3' UTRs, and examined their expression throughout protracted MB neurogenesis (Supplementary Fig. 2). Only *13XLexAop2-5'UTR-GFP* displayed a prominent declining expression profile, ascribing the major down-regulating activity to *chinmo*'s 5' UTR. To narrow down the involved region, we conducted in-situ deletion analysis of *chinmo* 5' UTR with CAMIO.

chinmo's 5'UTR is dispersed in 3 adjacent exons, with a distant 3rd exon separated from the first two near-by exons by 36kb (Fig. 3A). Three gRNA sets were designed to target these 3 exons for CRISPR mutagenesis (Fig. 3B). We provided two copies of the same set for induction of both indels and inter-target deletions within an exon. Additionally, we paired gRNA sets for exon 1 and 2 to create larger deletions that extend across the exon1/2 junction. Thus, we could delete various parts of *chinmo* 5'UTR in its endogenous locus by simple fly pushing.

Based on the previously observed deletion rate of around 35%, we collected 300 male progeny from each CAMIO genetic cross with the hope of saturating each 5'UTR exon with ~100 various deletions. In total, 1200 CAMIO males were harvested from 4 different gRNA array combinations. We mapped potential indels through sequencing indexed PCR products in a high-throughput manner.

We detected numerous indels around a single gRNA target site (Supplementary Fig. 3), and also recovered quite a few large inter-target deletions that together allow efficient coverage of the entire 5'UTR (Fig. 3C). Screening these large inter-target deletions in homozygous organisms revealed aberrant MB morphology in $chinmo^{Ex2L56}$ and $chinmo^{Ex2L130}$ that uncover a common ~300 bps

fragment in exon 2 (Fig. 4A). The penetrance of this phenotype is sensitive to Chinmo dosage (Fig. 4B). The defective MBs with shorten or absent αβ lobes bear striking resemblances to the abnormal MBs observed in *fussel* mutants. *fussel encodes* a Smad co-factor required for activation of Ecdysone Receptor-B1 (EcR-B1) expression in late larval MBs (43). Marchetti and Tavosanis recently proposed a negative feedback loop in the suppression of Chinmo expression by EcR-B1 to promote α'β' to αβ fate transition at the prepupal stage (44). Indeed, we observed aberrantly elevated Chinmo in young MB neurons around pupa formation in both *chinmo*^{Ex2L56} and *chinmo*^{Ex2L130} homozygous mutants (Fig. 4C). Taken together, just one round of CAMIO has allowed we to identify a 300 bps (from 2.2kb) untranslated region critical for proper MB development. We should be able to narrow down further the candidate region if we had kept relevant small-deletion lines for follow-up phenotypic analysis.

Meanwhile, notably absent were exon 2 indels or large deletions involving the target sites of g2_2 and g2_3. This area might carry essential sequences for Chinmo's dosage-sensitive activities critical for organism viability. Alternative explanations for the failure in recovering those indels include: an overall shallow sequencing depth of the whole exon 2 region, our small sampling size, unexpectedly low gRNA on-target strength of these 2 gRNAs, or flawed designs in the exon 2 gRNA set. To address the last concern that bothered us most, we further assessed the usage of specific gRNAs for the exon 2 set in GSCs (Supplementary Fig. 4). While g2_2 and g2_3 were not recruited as frequently as others, they each still emerged in 6-7% of the time. This rate should be sufficient for us to recover some indels assuming comparable on-target efficiencies.

To overcome the possible heterozygous viability issue, we exploited mosaic approaches by creating brain-specific somatic mutations. A transgene, *dU6_g2+3*, was hence assembled to express g2_2 and g2_3 ubiquitously. We began with MB-specific CRISPR mutagenesis by inducing *UAS-Cas9* specifically in the MB lineage using a MB specific Gal4 (*OK107-Gal4*). We saw no temporal fate changes in the MB, the classic phenotype of *chinmo* misregulation. We next elicited CRISPR mutagenesis in all neural stem cells (neuroblast: NB) with NB-restricted Cas9 (*dpn-Cas9*) and *dU6_g2+3*. These animals were viable and showed no abnormal tumor-like NBs in both larval and adult brains (typical of Chinmo overexpression) (42). These data failed to provide evidence in support of presence of critical brain regulatory elements in the region targeted by g2_2 and g2_3. We strengthened this negative conclusion by directly removing various small-to-large fragments around g2_2 and g2_3 targets from the above *chinmo* 5' UTR GFP transgene. In developing MBs, we observed indistinguishable GFP expression profiles between wild-type and modified 5' UTRs (Supplementary Fig. 5). By contrast, inducing early somatic mutations nonspecifically with *act5C-Cas9* and *dU6_g2+3* triggered severe embryonic or early larval lethality. This dominant lethality provides a direct explanation for our failure in recovering viable organisms carrying g2_2 or g2_3 induced indels.

In conclusion, a single round of CAMIO has successfully led us to uncover two critical regions of *chinmo* 5' UTR. The first critical region lies around g2_2 and g2_3 targets and carries essential sequences for dosage-sensitive organism viability apparently related with non-brain Chinmo functions. However, we could not rule out the possibility that the two gRNAs with adjacent *chinmo* targets accidently hit another essential gene(s) carrying the same sequence motifs, which appears very unlikely. As to the second critical region uncovered by the overlapping $chinmo^{Ex2L56}$ and $chinmo^{Ex2L130}$ deletions, we confirmed its importance in down-regulating Chinmo expression to ensure proper MB development. This fruitful case study well exemplifies the power of CAMIO in high-throughput unbiased deletion analysis of the genome.

DISCUSSION

Two innovations synergistically enable CAMIO, a germline pipeline for arrayed CRISPR mutagenesis. First, bamP-Cas9 can specifically and efficiently elicit independent Cas9 actions in male germ cells. We obtained an indel mutation rate of about 50% when employing a set of mutually exclusive gRNAs (estimated based on the observed 35% for those two thirds with out-of-frame indels). Second, supplying an array with multiple sets of mutually exclusive gRNAs permits combinatorial expressions of same or distinct gRNAs (one gRNA recruited from one set). Co-existing gRNAs can then elicit small on-target indels as well as large inter-target deletions. This way, we can conduct targeted deletion analysis of the genome in vivo with complementary coverages at coarse and fine levels.

We were happily surprised to discover a much higher CRISPR mutagenesis rate in male than female germ cells using *bamP*. This sex difference was also observed when improving *bamP*-dependent CRISPR-induced HDR in Enhanced-Golic+ (manuscript in preparation). Currently, we do not have a proven explanation for this phenomenon. *bamP* has a striking similar expression pattern in both female and male germline: absence in GSCs and an early onset of expression during the four amplificatory mitoses with incomplete cytokinesis to produce 16 germline cysts (45). We haven't investigated if *bamP* activity itself is higher in male than female germline. Also, it is known that meiotic recombination does not occur in male *Drosophila*. Perhaps in the male cellular context under *bamP*, indels were favored outcomes due to a reduced access to homologous chromosomes as a template for homology-mediated repair. In any case, this feature allows us to utilize *bamP* to build a high-efficiency pipeline for targeted CRISPR mutagenesis in male germ cells.

Conventional gRNA multiplexing provides all gRNAs at once as a cocktail, which expands indel diversity but inevitably creates complex and often biased deletion patterns. The off-target effects of a gRNA cocktail also accumulate in an additive manner. By contrast, CAMIO pre-selects gRNAs from multiple sets with one

gRNA from one set. Thus, CAMIO confers every gRNA with some autonomy while achieving multiplexed mutagenesis as a whole. Off-target concerns in CAMIO can be adequately addressed by examining multiple independent mutations of similar kinds. Also, arrays of targeted mutations can be introduced into specific genetic backgrounds with CAMIO. For example, we purposefully targeted a 2L chromosome arm that also carries transgenes for twin-spot MARCM in the CAMIO on *chinmo* 5' UTR. Hence, all the CAMIO *chinmo* indels were immediately ready for mosaic analysis. Such independent indels could be directly screened for visible phenotypes in mosaic organisms. However, we favor mapping the indels first by NGS, which can be conducted in a high-throughput manner via combinatorial sample indexing. We have further reduced the costs by pooling distinct amplicons for co-indexing.

In general, we have recovered similar indel spectrums as what's been commonly described. For gRNAs that are inherently more potent, like g1_4 for *chinmo* 5' UTR, we obtained more indels around the cut site. Despite recovering numerous single-site deletions, we rarely see those deletions exceeding 30 bps in length. Therefore, the observed ease in creating diverse inter-target deletions by CAMIO is particularly valuable for systematic dissection of sizable DNAs. In the case of CAMIO on *chinmo* 5'UTR, we have recovered most of the predicted inter-target deletions if not considering the 'toxic' g2_2 and g2_3. Notably, the largest inter-targeting deletion we have identified exceeds 2 kb in length. These observations suggest that we can be more aggressive in choosing farther adjacent gRNA targets to cover bigger genomic regions of interest.

After all, the capacity of CAMIO is mainly determined by how many gRNAs one can pack into a single set. Given the small size of gRNAs, we expect no problem in packing many gRNAs without compromising the system. We validated this intuition by checking the actual recruitment frequency for each of the six tandem gRNAs in the *chinmo* exon 2 set. Further, if needed, we can possibly improve the fairness in recruitment by increasing the distance between the U6 promoter and the gRNA set. Taken together with the consistence in recovering almost all predicted inter-target deletions so far, we have apparently not explored the potential of CAMIO to its maximum in terms of the size of DNA one CAMIO experiment can afford to analyze.

With CAMIO's fine dissection on *chinmo*'s 5' UTR, we successfully identified a region responsible for Chinmo downregulation around papa formation. Elevated Chinmo expression affected MB morphogenesis, possibly due to abnormal MB neuronal fate transition. There evidently exist multiple mechanisms governing *chinmo* expression throughout development. Notably, the mechanisms that contribute to a declining Chinmo protein profile primarily act on *chinmo* transcripts. Resolving such delicate post-transcriptional regulations requires additional rounds of CAMIO experiments. Given the multiplex and combinatorial power, CAMIO should also aid greatly the dissection of promoters, enhancers, long non-coding RNAs, and DNA repeats etc..

In theory, CAMIO should work for any organisms where a *bamP*-like promoter exists. Particularly, if a founder parent (possibly a father) can produce a large number of targeted mutants, CAMIO may become a desirable genetic screening platform for interrogating the genome. A mouse gene, *Gm114*, was identified as a putative ortholog of *Drosphila bam (46)*. Encouragingly, it is greatly enriched in undifferentiated spermatocytes and spermatids but absent or extremely low in undifferentiated spermatogonia, a pattern strikingly similar to *bam*'s. Orthologs of *bam* and *Gm114* were also found in zebrafish, chicken, macaque, human, and others. The human ortholog (*HT013*) exists abundantly in testis as well. Taken together, we have pioneered CAMIO as a germline pipeline for arrayed CRISPR mutagenesis in *Drosophila*. As exemplified by CAMIO on *chinmo* 5'UTR, we expect CAMIO to expedite systematic structure-functional analysis of the genome across diverse model organisms.

ACKNOWLEDGEMENT

We thank Janelia Fly Core and Quantitative Genomics for technical support. We thank Dr. Ling-Yu Liu for assistance on analysis of Chinmo expression in the MB. We thank Crystal Di Pietro and Kathryn Miller for administrative support. This work was supported by Howard Hughes Medical Institute.

REFERENCES

1. Chalfie, M. and Kain, S.R. (2005) *Green fluorescent protein: properties, applications and protocols*. John Wiley & Sons, Inc. .
2. Boulin, T., Etchberger, J.F. and Hobert, O. (2006) *Reporter gene fusions*. WormBook, ed. The C. elegans Research Community, WormBook.
3. Kvon, E.Z., Stampfel, G., Yanez-Cuna, J.O., Dickson, B.J. and Stark, A. (2012) HOT regions function as patterned developmental enhancers and have a distinct cis-regulatory signature. *Genes Dev*, **26**, 908-913.
4. Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D. *et al.* (2006) In vivo enhancer analysis of human conserved non-coding sequences. *Nature*, **444**, 499-502.
5. Dupuy, D., Li, Q.R., Deplancke, B., Boxem, M., Hao, T., Lamesch, P., Sequerra, R., Bosak, S., Doucette-Stamm, L., Hope, I.A. *et al.* (2004) A first version of the Caenorhabditis elegans Promoterome. *Genome research*, **14**, 2169-2175.
6. Kim, Y.-G., Cha, J. and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences*, **93**, 1156-1160.
7. Bibikova, M., Golic, M., Golic, K.G. and Carroll, D. (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics*, **161**, 1169-1175.
8. Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.-L., Rutter, J., Gall, J.G. and Carroll, D. (2008) Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. *Proceedings of the National Academy of Sciences*, **105**, 19821-19826.
9. Kim, H. and Kim, J.-S. (2014) A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics*.
10. Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E. and Hinkley, S.J. (2011) A TALE nuclease architecture for efficient genome editing. *Nature biotechnology*, **29**, 143-148.
11. Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, **326**, 1509-1512.
12. Moscou, M.J. and Bogdanove, A.J. (2009) A simple cipher governs DNA recognition by TAL effectors. *Science*, **326**, 1501-1501.
13. Mussolino, C., Alzubi, J., Fine, E.J., Morbitzer, R., Cradick, T.J., Lahaye, T., Bao, G. and Cathomen, T. (2014) TALENs facilitate targeted genome editing in human cells with high specificity and low cytotoxicity. *Nucleic acids research*, **42**, 6762-6773.
14. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816-821.

15. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W. and Marraffini, L.A. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819-823.
16. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.J. and Joung, J.K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature biotechnology*, **31**, 227-229.
17. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823-826.
18. Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J. and O'Connor-Giles, K.M. (2013) Genome engineering of Drosophila with the CRISPR RNA-guided Cas9 nuclease. *Genetics*, **194**, 1029-1035.
19. Port, F., Chen, H.M., Lee, T. and Bullock, S.L. (2014) Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, E2967-2976.
20. Komor, A.C., Badran, A.H. and Liu, D.R. (2017) CRISPR-Based Technologies for the Manipulation of Eukaryotic Genomes. *Cell*, **169**, 559.
21. Lieber, M.R. (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual review of biochemistry*, **79**, 181-211.
22. San Filippo, J., Sung, P. and Klein, H. (2008) Mechanism of eukaryotic homologous recombination. *Annual review of biochemistry*, **77**, 229-257.
23. Fellmann, C., Gowen, B.G., Lin, P.C., Doudna, J.A. and Corn, J.E. (2017) Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nature reviews. Drug discovery*, **16**, 89-100.
24. Hales, K.G., Korey, C.A., Larracuente, A.M. and Roberts, D.M. (2015) Genetics on the Fly: A Primer on the Drosophila Model System. *Genetics*, **201**, 815-842.
25. Bier, E., Harrison, M.M., O'Connor-Giles, K.M. and Wildonger, J. (2018) Advances in Engineering the Fly Genome with the CRISPR-Cas System. *Genetics*, **208**, 1-18.
26. Chen, D. and McKearin, D.M. (2003) A discrete transcriptional silencer in the bam gene determines asymmetric division of the Drosophila germline stem cell. *Development*, **130**, 1159-1170.
27. Chen, H.M., Huang, Y., Pfeiffer, B.D., Yao, X. and Lee, T. (2015) An enhanced gene targeting toolkit for Drosophila: Golic+. *Genetics*, **199**, 683-694.
28. Zhu, S., Lin, S., Kao, C.F., Awasaki, T., Chiang, A.S. and Lee, T. (2006) Gradients of the Drosophila Chinmo BTB-zinc finger protein govern neuronal temporal identity. *Cell*, **127**, 409-422.
29. Miyares, R.L. and Lee, T. (2018) Temporal control of Drosophila central nervous system development. *Curr Opin Neurobiol*, **56**, 24-32.
30. Pfeiffer, B.D., Ngo, T.-T.B., Hibbard, K.L., Murphy, C., Jenett, A., Truman, J.W. and Rubin, G.M. (2010) Refinement of tools for targeted gene expression in Drosophila. *Genetics*, **186**, 735-755.

31. Yu, H.-H., Chen, C.-H., Shi, L., Huang, Y. and Lee, T. (2009) Twin-spot MARCM to reveal the developmental origin and identity of neurons. *Nature neuroscience*, **12**, 947-953.
32. Pfeiffer, B.D., Truman, J.W. and Rubin, G.M. (2012) Using translational enhancers to increase transgene expression in Drosophila. *Proceedings of the National Academy of Sciences*, **109**, 6626-6631.
33. He, Y., Zhang, T., Yang, N., Xu, M., Yan, L., Wang, L., Wang, R. and Zhao, Y. (2017) Self-cleaving ribozymes enable the production of guide RNAs from unlimited choices of promoters for CRISPR/Cas9 mediated genome editing. *Journal of genetics and genomics = Yi chuan xue bao*, **44**, 469-472.
34. Awasaki, T., Saito, M., Sone, M., Suzuki, E., Sakai, R., Ito, K. and Hama, C. (2000) The Drosophila trio plays an essential role in patterning of axons by regulating their directional extension. *Neuron*, **26**, 119-131.
35. Wu, Y.C., Chen, C.H., Mercer, A. and Sokol, N.S. (2012) Let-7-complex microRNAs regulate the temporal identity of Drosophila mushroom body neurons via chinmo. *Dev Cell*, **23**, 202-209.
36. Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods*, **9**, 357-359.
37. Sfeir, A. and Symington, L.S. (2015) Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends in biochemical sciences*, **40**, 701-714.
38. Ferguson, S.M. and De Camilli, P. (2012) Dynamin, a membrane-remodelling GTPase. *Nature reviews. Molecular cell biology*, **13**, 75-88.
39. Kitamoto, T. (2001) Conditional modification of behavior in Drosophila by targeted expression of a temperature-sensitive shibire allele in defined neurons. *Journal of neurobiology*, **47**, 81-92.
40. Kanca, O., Caussinus, E., Denes, A.S., Percival-Smith, A. and Affolter, M. (2014) Raeppli: a whole-tissue labeling tool for live imaging of Drosophila development. *Development*, **141**, 472-480.
41. Lee, T., Lee, A. and Luo, L. (1999) Development of the Drosophila mushroom bodies: sequential generation of three distinct types of neurons from a neuroblast. *Development*, **126**, 4065-4076.
42. Dillard, C., Narbonne-Reveau, K., Foppolo, S., Lanet, E. and Maurange, C. (2018) Two distinct mechanisms silence chinmo in Drosophila neuroblasts and neuroepithelial cells to limit their self-renewal. *Development*, **145**.
43. Takaesu, N.T., Stinchfield, M.J., Shimizu, K., Arase, M., Quijano, J.C., Watabe, T., Miyazono, K. and Newfeld, S.J. (2012) Drosophila CORL is required for Smad2-mediated activation of Ecdysone Receptor expression in the mushroom body. *Development*, **139**, 3392-3401.
44. Marchetti, G. and Tavosanis, G. (2017) Steroid Hormone Ecdysone Signaling Specifies Mushroom Body Neuron Sequential Fate via Chinmo. *Current biology : CB*, **27**, 3017-3024.e3014.
45. Fuller, M.T. and Spradling, A.C. (2007) Male and female Drosophila germline stem cells: two versions of immortality. *Science*, **316**, 402-404.

46. Tang, H., Ross, A. and Capel, B. (2008) Expression and functional analysis of Gm114, a putative mammalian ortholog of Drosophila bam. *Dev Biol*, **318**, 73-81.
47. Park, J., Lim, K., Kim, J.S. and Bae, S. (2017) Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics (Oxford, England)*, **33**, 286-288.
48. Allen, F., Crepaldi, L., Alsinet, C., Strong, A.J., Kleshchevnikov, V., De Angeli, P., Palenikova, P., Khodak, A., Kiselev, V., Kosicki, M. *et al.* (2018) Predicting the mutations generated by repair of Cas9-induced double-strand breaks. *Nat Biotechnol*.

FIGURES LEGENDS

Figure 1. *bamP-Cas9* induces efficient CRISPR targeted mutagenesis in male germ cells. (A) *U6:3-gRNA-e* was designed to target *ebony* close to the 5' end of its CDS. (B) 3.3% of the progeny of female founders displayed LoF *ebony* phenotype, while 26.1% of male founders' progeny were Lof *ebony*. Error bars represent SE. (C) Sequences of 10 randomly selected Lof ebony progeny from a common male founder. (D) Deletion profile of 919 phenotypically wildtype progeny. NGS data was analyzed and presented by Cas-analyzer, www.rgenome.net (47). (E) We use FORECasT (48) to predict indels from *gRNA-shi* and *UAS-shibire* sequences, and group them into 3 reading frames, with frame 0 represents the in-frame indels. Actual deletion sizes calculated from mapping results (obtained with CAS-analyzer), are converted into frame shift and the read counts are aggregated accordingly. Chi-square test assuming equal distribution is used to assess the significance. The result is below machine precision and thus set to zero.

Figure 2. CAMIO produces diverse arrayed mutations around selected gRNA target sites. (A) The design of our conditional U6-gRNAs, pCis-(mutagenesis). The U6 promoter is separated from the gRNAs by a large fragment containing a 3xP3-RFP marker. In each GSC, *nos-phiC31* can bring only one gRNA within the gRNA set under the U6 control and complete the gRNA selection and activation. Shown here for example, gRNA#2 is chosen. (B) 4 gRNA target sites along mCD8 CDS were selected for pCis-{4gRNAs_mCD8} to disrupt mCD8::GFP expression. Simple or combined multiplexing mutagenesis can be achieved by incorporating one or multiple *pCis-{mutagenesis}* transgenes with *nos-phiC31* in male *bamP(898)-Cas9* founders. (C) Three categories of deletions arose from applying two copies of pCis-{4gRNAs_mCD8}: single-site deletion (76.4%), defined deletion (8.2%), and double-site deletion (15.4%). Deletion size and count are depicted for single-site deletions; gRNA choices deduced from sequencing the GFP negative progeny are summarized for defined deletions and double-site deletions.

Figure 3. Applying CAMIO on *chinmo* 5' UTR. (A) An illustration of the *chinmo* gene. (B) We selected 4 target sites on *chinmo* exon 1 (455 bps), 6 target sites on exon 2 (1061 bps), and 4 target sites on exon 3 (639 bps). Each exon was dissected by CAMIO with a pair of pCis-{mutagenesis} gRNA sets. Additionally, we had combination 4, *pCis-{4gRNAs_chinmo Exon1}* + *pCis-{6gRNAs_chinmo Exon2}*, to see if larger deletions could be produced across target sites much farther apart. (C) In total, we sequenced 984 offspring (300 each for combination 1-3, and another 84 for combination 4) from CAMIO on *chinmo* 5' UTR. Shown here are the larger deletions detected by our NGS analysis. Predicted deletions over 100 bps were of great interest and were subject to second round Sanger sequencing. Confirmed >100 bps deletions are marked in red and named by L (large) followed by their stock ID. Homozygous viable L deletions that cover

larger proportions of each exons (marked in green) were selected for MB development studies.

Figure 4. Recovery of two *chinmo* 5'UTR deletions, *chinmo*^{Ex2L56} and *chinmo*^{Ex2L130}, which show defective adult MB morphology and elevated Chinmo expression. (A) Stacked confocal images of the adult MBs stained for Fas2 (green, α, β and γ lobes) and Trio (red, α', β' and γ lobes). About half of *chinmo*^{Ex2L56} and *chinmo*^{Ex2L130} MBs have missing or misshapen αβ lobes. Scale bars: 50 μm. (B) Homozygous *chinmo*^{Ex2L56} (39.2%) and *chinmo*^{Ex2L130} (50%) and transheterozygous *chinmo*^{Ex2L56}/*chinmo*^{Ex2L130} (27.9%) have strong MB αβ lobe defects, while such phenotypes are much less penetrant in hemizygous *chinmo*^{Ex2L56}*/Df* (1.6%) and *chinmo*^{Ex2L130}*/Df* (5%). (C) Single optical sections of multiple MB lineages (labeled by *OK107-Gal4*, *UAS-mCD8::GFP*) at white pupa stage for wild type, homozygous *chinmo*^{Ex2L56} and *chinmo*^{Ex2L130}. Abnormally elevated Chinmo is observed in newly derived MB neurons (weaker GFP signals, outlined by white dashed lines). Green: GFP; Magenta: Chinmo. Scale bars: 20 μm.

Supplementary Figure 1. Targeted mutagenesis of mCD8::GFP with pCis-{4gRNAs_mCD8}. Deletions located around all four gRNA target sites were recovered after sequencing 30 progeny from two male founders. However, much less deletions originated from gRNA#2 and gRNA#3.

Supplementary Figure 2. Three *13XLexAop2-GFP* reporters for investigating *chinmo* 5' and 3' UTRs. We generated three 13XLexAop2-GFP reporters flanked by both *chinmo* 5' and 3' UTRs, only 5' UTR, or only 3' UTR. We induced and restricted their expression in MB lineages by immortalizing a transient MB NB expression (*41A10-KD*) into a sustained MB NB production of these three GFPs, with this genetic setup: *DpnEE-KO-LexAp65; 13XLexAop2-GFP-UTRs; 41A10-KD*. By examining their expression profiles at two time points (48 and 72 hours after larval hatching, ALH), we observed *13XLexAop2-GFP-3' UTR* is expressed at a much higher level than *13XLexAop2-5' UTR-GFP-3' UTR*, while expression of *13XLexAop2-5' UTR-GFP* decreases over time. Green: GFP; Red: Deadpan. Scale bars: 20 μm.

Supplementary Figure 3. The small indels created by CAMIO in *chinmo* Exon 1, 2, and 3. CAMIO generated arrayed indels around designed gRNA target sites in all four array combinations. However, we did not recover any indels around g2_2 and g2_3.

Supplementary Figure 4. Examining gRNA array selection of *pCis-(6gRNAs_chinmo Exon2)*. We studied how frequently each gRNA gets selected after nos-phiC31 mediated recombination by scoring progeny of male *nos-phiC31; pCis-(6gRNAs_chinmo Exon2*). (A) 92 progeny's gRNA choices from 9 male founders are presented in this heat map. (B) Counts for each gRNA from all the founders are aggregated and whether there is bias in selection is tested by

Chi-square test assuming equal distribution. g2_2 (6.5%, 6/92) and g2_3 (7.6%, 7/92) are underrepresented.

Supplementary Figure 5. Three 5'UTR-GFP-3'UTR reporters carrying small to large *chinmo* Exon 2 deletions for studying the region around g2_2 and g2_3 target sites. We generated three additional chinmo 5' UTR-GFP-3' UTR reporters: d (50 bps deletion around g2_3), bD (larger 158 bps deletion), and D23 (removing all the Exon 2 sequence upstream of g2_3). GFPs were induced in the same *41A10-KD* immortalization strategy. We did not observe difference in expression among the four GFP UTR reporters at three development time points: 48h AEL (after egg laying), 96h AEL and, white pupa. Scale bars: 30 μm.

FIGURES

Figure 1. *bamP-Cas9* induces efficient CRISPR targeted mutagenesis in male germ cells.

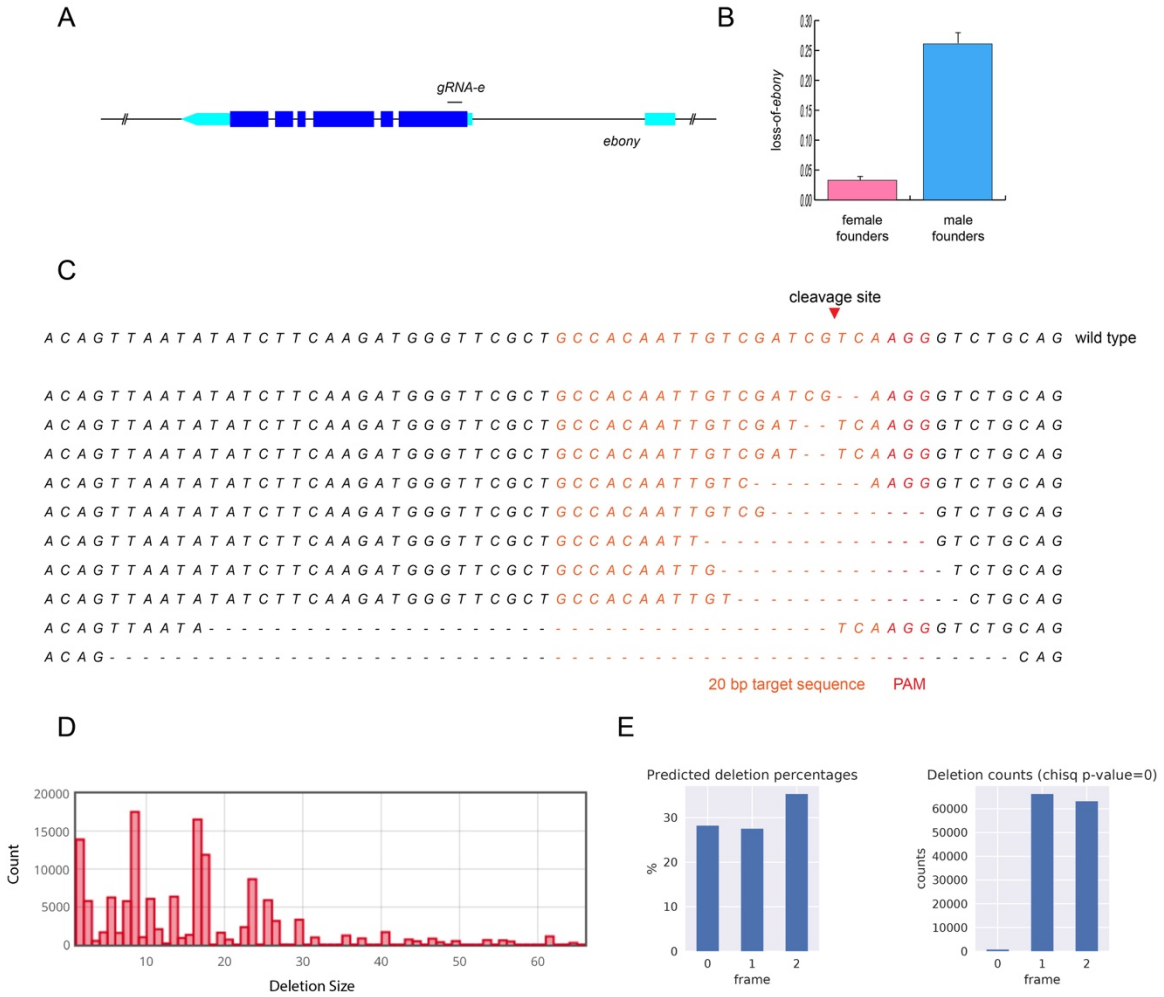


Figure 2. CAMIO produces diverse arrayed mutations around selected gRNA target sites.

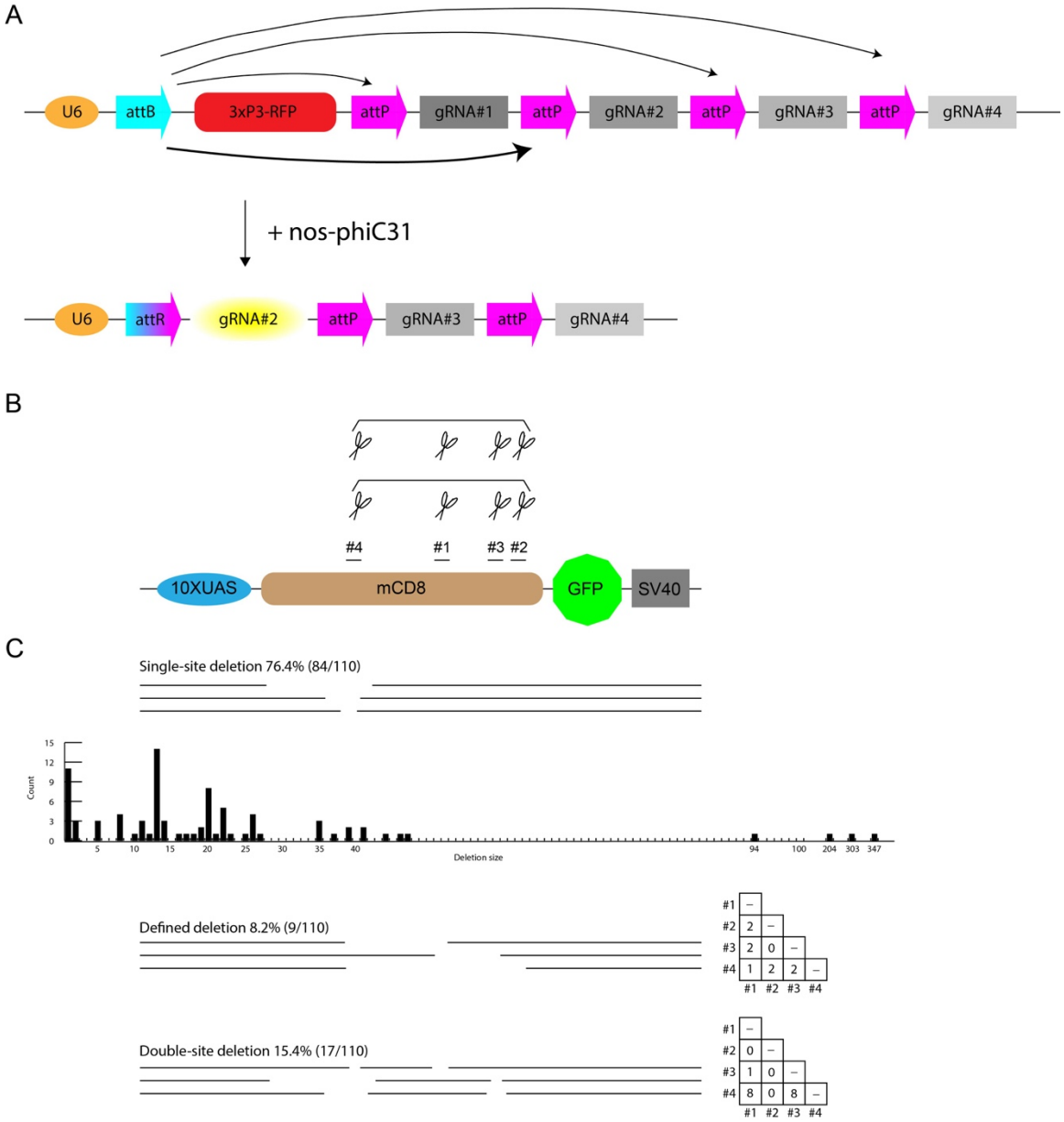


Figure 3. Applying CAMIO on *chinmo* 5' UTR.

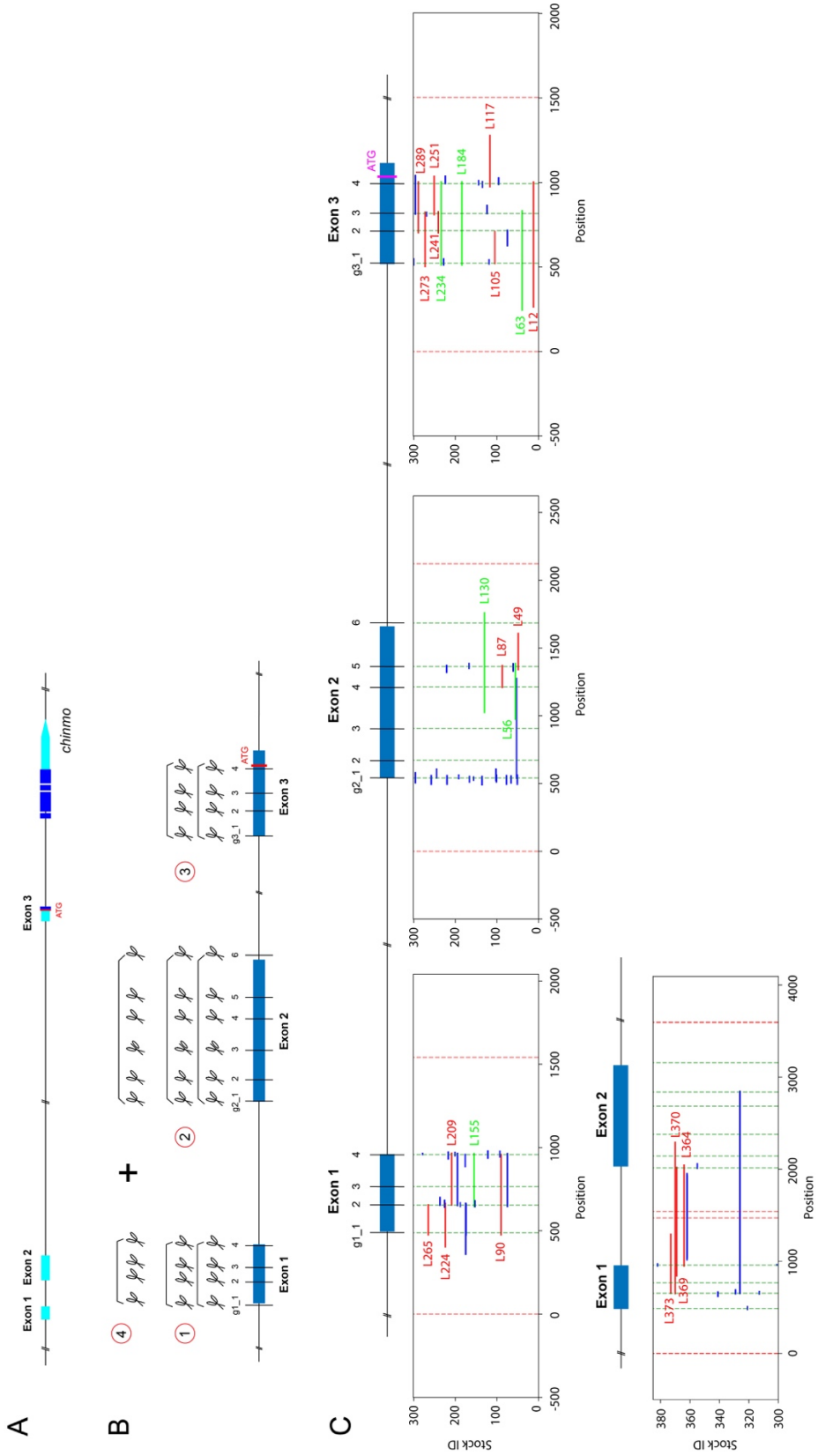
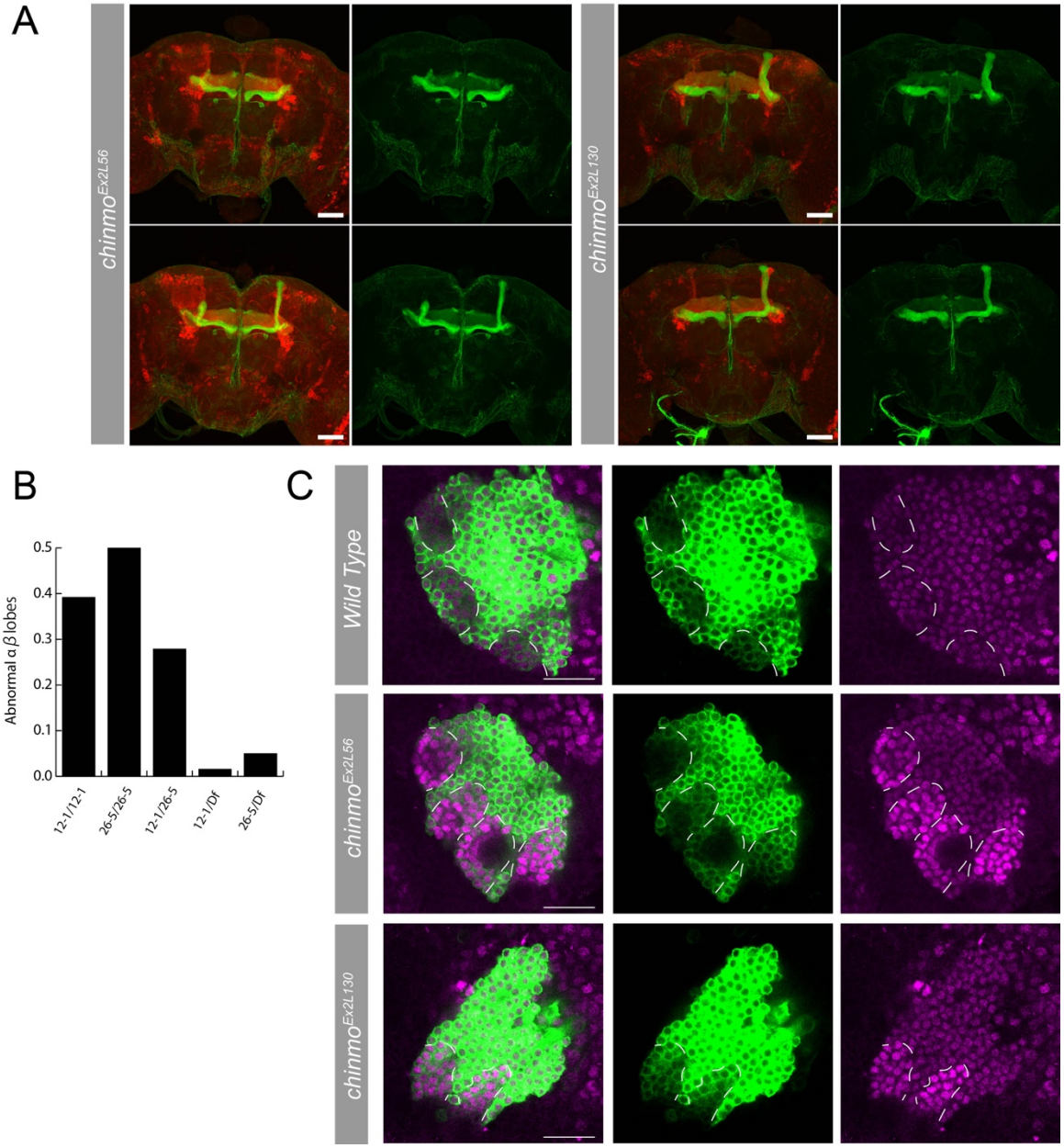
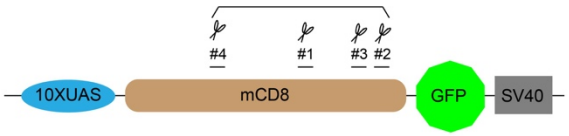


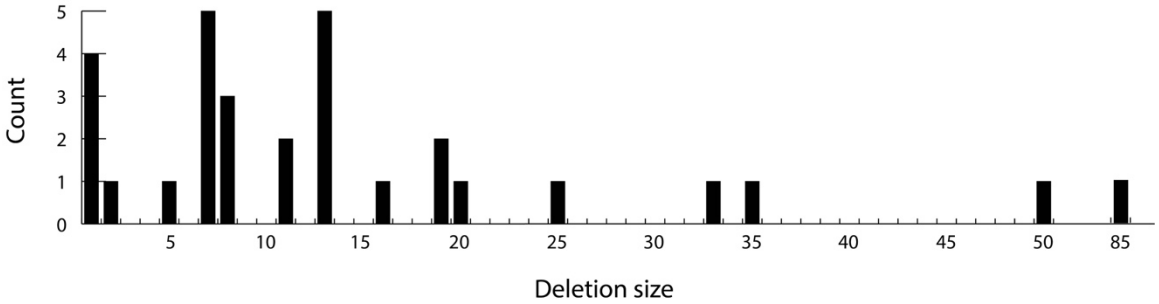
Figure 4. Recovery of two *chinmo* 5'UTR deletions, *chinmo*^{Ex2L56} and *chinmo*^{Ex2L130}, which show defective adult MB morphology and elevated Chinmo expression.



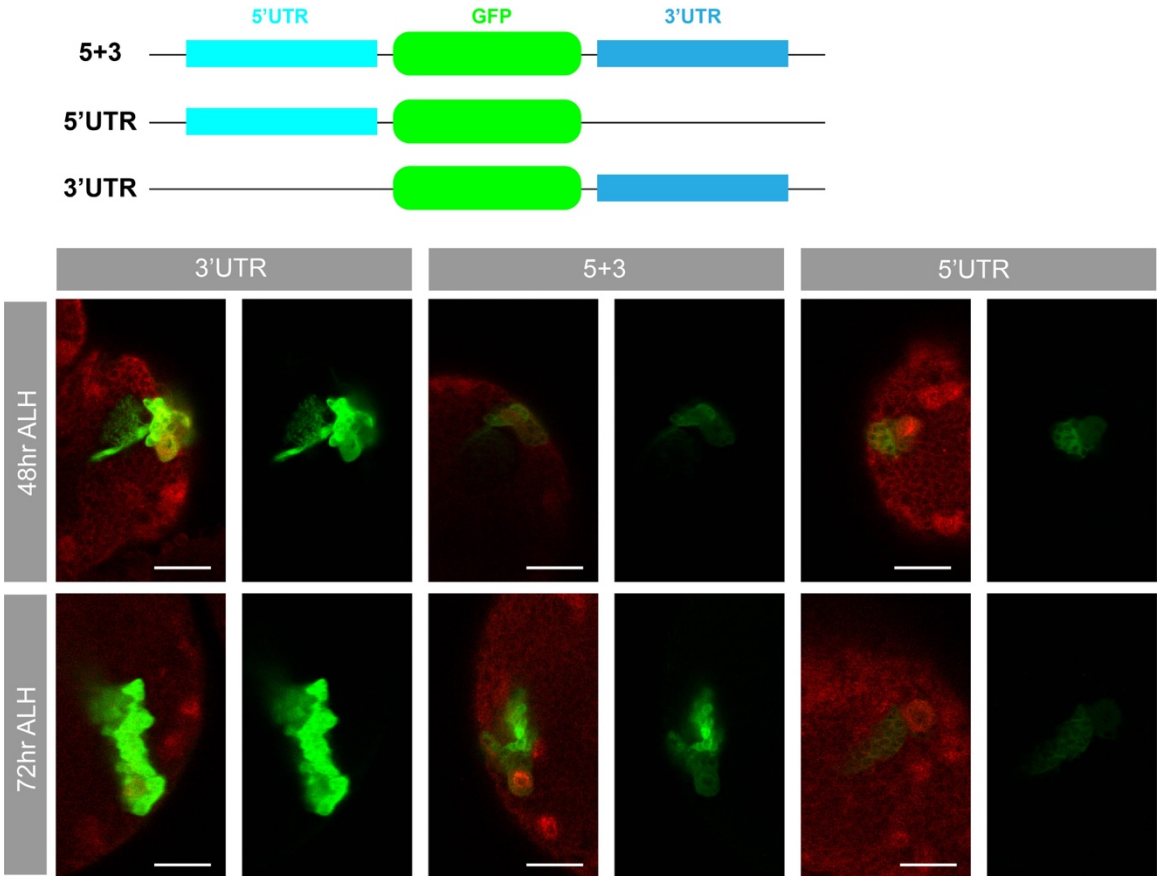
Supplementary Figure 1. Targeted mutagenesis of mCD8::GFP with pCis-{4gRNAs_mCD8}.



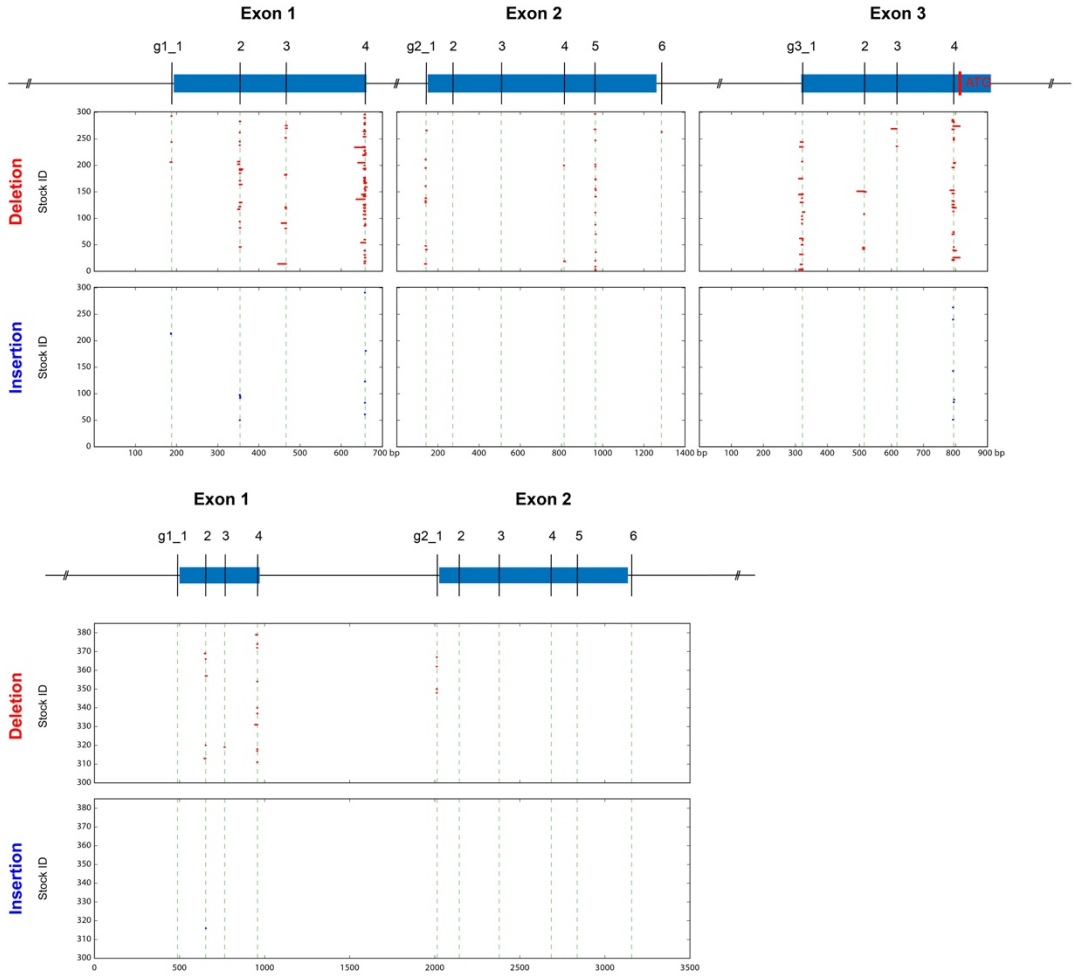
	#1	#2	#3	#4
Founder A	10	0	1	4
Founder B	9	2	1	3
Overall	62%	7%	7%	24%



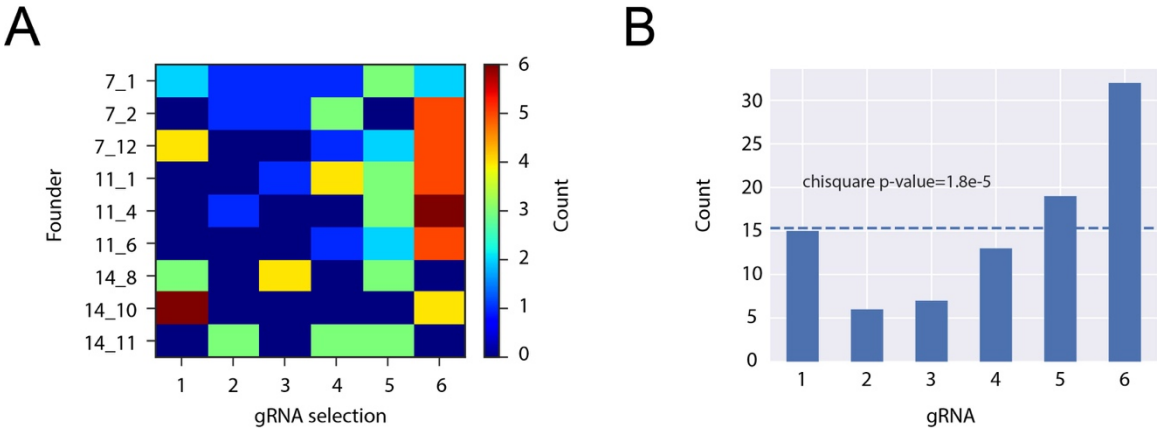
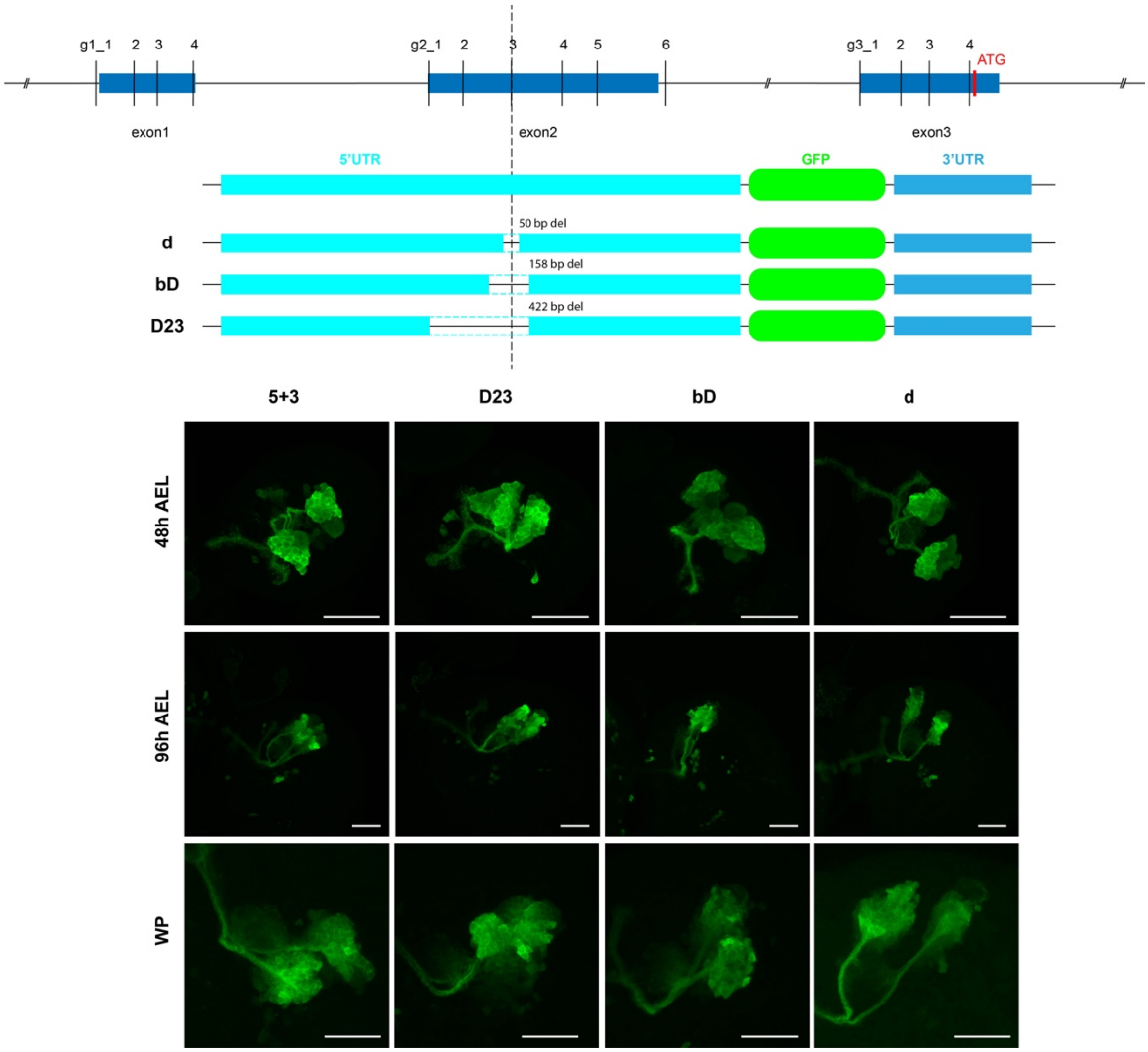
Supplementary Figure 2. Three *13XLexAop2-GFP* reporters for investigating *chinmo* 5' and 3' UTRs.



Supplementary Figure 3. The small indels created by CAMIO in *chinmo* Exon 1, 2, and 3.



Supplementary Figure 4. Examining gRNA array selection of *pCis-(6gRNAs_chinmo Exon2)*.



Supplementary Figure 5. Three 5'UTR-GFP-3'UTR reporters carrying small to large *chinmo* Exon 2 deletions for studying the region around g2_2 and g2_3 target sites.