Directed, efficient, and versatile modifications of the *Drosophila* genome by genomic engineering

Juan Huang[1], Wenke Zhou[1], Wei Dong, Annie M. Watson, and Yang Hong[2]

Department of Cell Biology and Physiology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261

Edited by Yuh Nung Jan, University of California School of Medicine, San Francisco, CA, and approved April 2, 2009 (received for review January 21, 2009)

With the completion of genome sequences of major model organisms, increasingly sophisticated genetic tools are necessary for investigating the complex and coordinated functions of genes. Here we describe a genetic manipulation system termed "genomic engineering" in *Drosophila*. Genomic engineering is a 2-step process that combines the ends-out (replacement) gene targeting with phage integrase ϕC31-mediated DNA integration. First, through an improved and modified gene targeting method, a founder knock-out line is generated by deleting the target gene and replacing it with an integration site of ϕC31. Second, DNA integration by ϕC31 is used to reintroduce modified target-gene DNA into the native locus in the founder knock-out line. Genomic engineering permits directed and highly efficient modifications of a chosen genomic locus into virtually any desired mutant allele. We have successfully applied the genomic engineering scheme on 6 different genes and have generated at their loci more than 70 unique alleles.

cell polarity | ends-out targeting | homologous recombination | phiC31 integrase

The development of homologous recombination (HR)-based gene targeting was a major breakthrough in *Drosophila* genetics (1, 2). At present, in *Drosophila* as well as in mice, a HR-based approach is virtually the only way to make directed modifications of a target gene (3, 4). However, because the entire targeting process must be repeated for making each allele, the amount of time and labor may become impractical to make more than just a few targeted alleles. In addition, because of the requirement of HR, it can be very difficult to introduce appreciably complicated DNA sequence modifications by gene targeting. The current lack of adequate genetic tools for directed and efficient modifications of the genome presents a major hurdle in *Drosophila* genetics today. For example, many of the protein pathways that are highly conserved between *Drosophila* and vertebrates, such as the cell polarity pathway (5), appear to be exceedingly complex. Rigorous genetic dissections of such intricate protein networks can be highly challenging, because in most cases the functions of mutated or modified individual genes of such pathways can only be assayed by artificial over-expression methods, which often lack the requisite controllability and fidelity of gene expression. One ideal solution would be for each protein gene of interest to generate, at the gene's native genomic locus, a set of defined mutant alleles that are strategically designed to test hypotheses about the protein's in vivo functions and interactions. Furthermore, being able to generate any conceivable alleles of a target gene, such as functional fusion alleles of fluorescent proteins/purification tags or alleles with conditional activities, would also offer us unprecedented freedom and opportunities to explore unique experiments of imaging, proteomics, and disease models.

To achieve the goal of such directed, efficient, and versatile modifications of the *Drosophila* genome, we have developed an approach we have termed "genomic engineering" (Fig. 1) that combines ends-out (replacement) gene targeting with phage integrase ϕC31-mediated DNA integration. ϕC31 catalyzes unidirectional DNA recombination between the so-called attB and attP sites (6) and works very efficiently in *Drosophila* for transgenesis (6, 7). As illustrated in Fig. 1, our genomic engineering scheme offers several significant benefits. First, regardless of how many distinct mutant alleles will be generated, only 1 ends-out targeting experiment is needed. Second, the efficiency of ϕC31 integration should make the second step of allele generation a rather high-throughput process. Third, because ϕC31 integrase does not appear to discriminate against different DNA substrates (8, 9), DNA constructs for generating mutant alleles are not constrained by the limitations of HR. As a consequence, virtually any conceivable modification of the target-gene sequence can be accommodated.

Results

Minimal attP and attB Sites of ϕC31 Can Mediate Efficient DNA Integration in *Drosophila*. So far, most ϕC31-mediated DNA integration experiments in *Drosophila* used full attB and attP of 200- to 300-bp length or even longer (6, 7, 10), while a few used full-length attP and minimal attB (40 bp) (7, 11). Recombination between attP and attB generates so-called attL and attR sites that are roughly the average size of attB and attP (8), raising the concern that using full-length attP or attB in genomic engineering will result in a long exogenous attR sequence that may interfere with host-gene expression in the final allele (see Fig. 1*F*). To address this concern, we tested and confirmed that minimal attP-50 and attB-53 sites (9) that are 50-bp and 53-bp long, respectively, can mediate efficient integration in *Drosophila* [supporting information (SI) Table S1]. We then constructed new ends-out targeting vectors, such as pGX-attP that carries an attP-50 site, and integration vectors, such as pGE-attB that carries an attB-53 site (Fig. S1), for genomic engineering. By using attP-50 and attB-53 in these genomic engineering vectors, we drastically reduced the attR sequence length in the final engineered allele and minimized the risk of its interference with the expression of the allele.

Generation of Founder Knock-Out Lines of 6 Different Target Genes by Ends-Out Targeting. To apply the genomic engineering scheme on a target locus, a founder knock-out line has to be generated by ends-out targeting. Gene targeting, although successfully developed years ago, has often been considered risky and resource-intensive in *Drosophila*. We have recently developed new reagents and fly stocks that significantly improved the efficiency and throughput of current ends-out targeting (12). Nonetheless, it remains to be demonstrated how our targeting reagents and protocols may perform at different genomic loci. Here we

Author contributions: J.H., W.Z., and Y.H. designed research; J.H., W.Z., W.D., A.M.W., and Y.H. performed research; J.H., W.Z., W.D., and Y.H. analyzed data; and Y.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

The sequences reported in this paper have been deposited in the GenBank database (accession numbers FJ791035, FJ791036, and FJ791037).

[1]J.H. and W.Z. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: yhong@pitt.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0900641106/DCSupplemental.

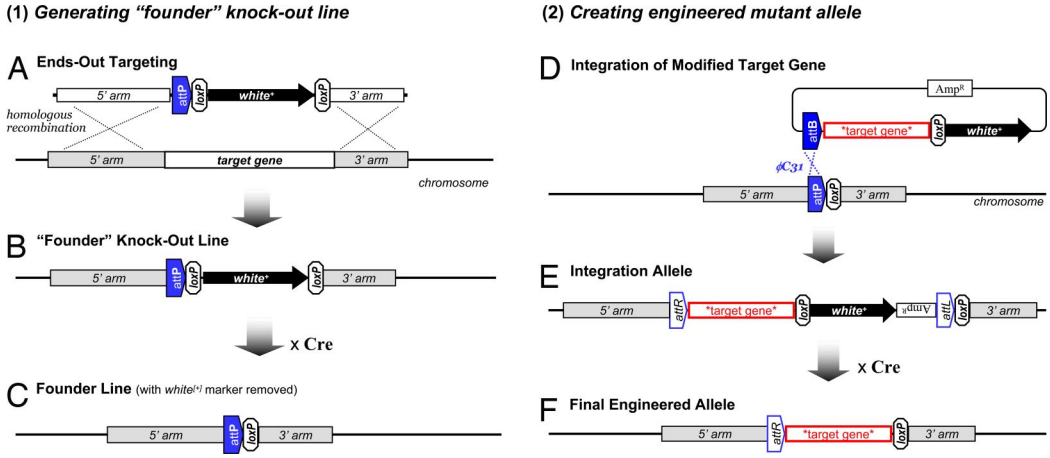

Fig. 1. Genomic engineering by targeted site-specific DNA integration. (*A*) A modified ends-out gene-targeting approach is used to delete the target gene first. The targeting donor DNA fragment contains 5′ and 3′ homologous arms (''arm'') flanking the target-gene locus, a loxP-flanked *white*+ (*w*+) transgenic marker juxtaposed by an attP site of ϕC31. (*B*) In the knock-out mutant (''founder knock-out line'' or founder line), the target gene is effectively replaced by the loxP-flanked *w*+ marker juxtaposed by a single ϕC31 attP site. (*C*) The *w*+ marker is removed by Cre recombinase in the founder line, leaving only the attP and loxP at the deletion locus. (*D*) The deleted genomic DNA of the target gene is engineered in vitro to incorporate desired modifications (''*target gene*'') on an integration vector (pGE-attB) that carries an attB site together with a *w*+ marker. It will then be integrated into the deletion locus of the founder line through ϕC31-mediated DNA integration. (*E*) The resulted ''integration mutant allele'' has the engineered target gene restored (with modifications) at its original genomic locus together with *w*+ and vector sequences (''*Amp*^{R}''). (*F*) Extra vector sequences, together with *w*+ can be removed by Cre recombinase, to generate a final engineered-mutant allele containing solely the engineered target gene flanked by attR and loxP sites.

selected 6 different genes that are distributed over all 3 major chromosomes (X, second, and third), and require deletions from as small as 2.2 kb to as large as 12 kb for genomic engineering (Table 1). Four of these genes, *stardust* (*sdt*), *lethal giant larvae* (*lgl*), *DE-Cadherin* (*DE-Cad*, also known as *shotgun* or *shg*), and *crumbs* (*crb*) are previously characterized polarity protein genes that play highly conserved and essential roles in regulating cell polarity in both *Drosophila* and vertebrates (5, 13, 14), but the detailed molecular and cellular mechanisms by which they control the cell polarity remain to be elucidated. With the help of genomic engineering, we hope that we will be able to generate at each of their native genomic loci a set of defined mutant alleles tailored for our genetic and cell biology assays (see below). *dArf6* (*Arf51F*) and its potential GTPase exchange factor *CG31158* were implicated in controlling cell polarity by our preliminary RNAi screens, and had no published mutant alleles at the time we started their gene-targeting experiments.

Targeting designs for generating founder knock-out lines for the 6 genes are detailed in Table 1 and Figs. S2 to S4. Because our goal is to modify the coding sequence of these 6 genes to test the functions of mutant proteins in vivo, in each of the targeting constructs the 5′ and 3′ homologous arms were designed to delete all or major coding exons in their loci. To minimize the risk that the attR and loxP may interfere with the expression of engineered alleles, in all targeting constructs the attP-50 site was placed into the least-conserved noncoding region, such as introns or upstream sequences of target loci (see *Materials and Methods*), while the loxP site was placed after the 3′UTR of engineered alleles by including the whole 3′UTR into the targeted deletion (see Fig. S2–S4; the only exception is the *CG31158* targeting in Fig. S3*E*). Targeting experiments were carried out based on our improved methods, such as using the new *hs-hid* stocks to streamline the genetic crosses; in the cases of *dArf6* and *sdt* a negative selection marker *UAS-Rpr* was also used to eliminate the majority of false positives (12). Details of *dArf6* targeting were published in Huang et al. (12). For each of the target genes, at least 1 founder knock-out line was obtained and verified by molecular and genetic tests (Table 2 and see Figs. S2–S4). Deletion of *dArf6* or *CG31158* did not cause any lethality or polarity defects, but rather a recessive male sterility, consistent with recent reports that *dArf6* is only essential for male germline development (15). Our targeting results demonstrated that our improved ends-out targeting approach is efficient for generating founder knock-out lines of $>10^{-6}$ HR frequency, especially with the help of *UAS-Rpr* (see Table 2).

Validation of Genomic Engineering Founder Lines by ϕC31-Mediated DNA Integration. To test how efficiently and reliably the target loci in founder knock-out lines could be modified by ϕC31-mediated DNA integration, we selected 1 or 2 founder lines for each of the

Table 1. Design of gene targeting for generating founder knock-out lines for selected polarity genes

Target gene	Target chromosome	Exons/mRNA isoforms	5′ + 3′ arms[a] (kb)	Targeted gDNA deletion[b]	Genomic deletion size (kb)	Protein deletion/full length (aa)
stardust	X	26/7	4.5 + 3.2	*X: 8,129,169–8,134,146*	4.977	741/2,020
lgl	second	9/6	5.2 + 3.1	*2L: 21,725–9,743*	11.982	1,161/1,161
DE-Cad (shotgun)	second	2/1	5.2 + 3.2	*2R: 16,942,814–16,937,965*	4.849	1,298/1,507
dArf6	second	3/5	4.5 + 3.1	*2R: 11,210,875–11,213,032*	2.157	175/175
CG31158	third	14/2	5.3 + 2.8	*3R: 18,424,078–18,431,499*	7.421	1,474/1,480
crumbs	third	13/2	5.2 + 2.9	*3R: 20,130,302–20,140,245*	9.943	2,109/2,189

[a]5′ + 3′ arms: the lengths of 5′ and 3′ homologous arms in targeting construct.
[b]According to *Drosophila* genome release 5.1 at www.flybase.org.

Table 2. Generation of founder knock-out lines by ends-out targeting

Target gene	Screening cross progeny[a]	Preliminary candidates	On-target chromosome	Genetically verified	PCR verified	HR frequency[b]
DE-Cad[c]	≈1.6×10^5	≈1,700	96/≈1,700	22/72[d]	22/22	≈1.9×10^{-4}
lgl[c]	≈2.4×10^5	1,127	95/1,127	22/95[e]	22/22[j]	≈9×10^{-5}
crumbs[c]	≈1.8×10^5	≈400	26/≈400	1/26[f]	1/1	≈6×10^{-6}
CG31158[c]	≈1.5×10^5	1,140	8/1,140	1/8[g]	1/1	≈7×10^{-6}
dArf6	≈7×10^5	315	30/315	5/30[h]	5/5	≈7×10^{-6}
sdt	≈1×10^6	116	4/116	4/4[i]	4/4	≈4×10^{-6}

[a]The total estimated number of screening cross progeny (12) that were screened in each targeting experiment. Because all progeny were mixed and screened together, we did not register the clonality of the preliminary candidates. We assumed that each targeting mutant obtained was because of a distinct targeting event.
[b]Because all female candidates (or male candidates as in the case of *sdt* targeting) were discarded in targeting experiments, the adjusted HR frequency should be twice as high as listed here.
[c]The targeting constructs for these genes were made on an older version of pGX-attP that does not contain the *UAS-Rpr* (see *SI Materials and Methods*), hence the large numbers of false-positives among preliminary candidates.
[d]Only 72 of 96 candidates were tested for noncomplementing the null allele shg^2.
[e]All candidates were tested for noncomplementing the null allele lgl^4.
[f]All candidates were tested for noncomplementing the null allele crb^{11A22}.
[g]A previously generated knock-out allele, $CG31158^{KO\#1}$ (see *Materials and Methods*) was used for complementation assays.
[h]A $dArf6^{\Delta KG\#1}$-deletion allele generated by P-excision (see *Materials and Methods*) was used for complementation assays.
[i]All candidates were tested for noncomplementing the null allele sdt^{XP96} (24).
[j]See Fig. S3 for details.

6 target genes and removed their *w+* marker by Cre recombinase (6) (see Fig. 1 *B* and *C*, and *Materials and Methods*). The resulting *w[−]* founder lines can be readily integrated with DNA constructs, such as pGE-attB, which bears the same w+ marker (see Fig. 1 *D–F*). For each target gene, we first integrated into the founder line a pGE-target (rescue) construct that contains the deleted genomic DNA (gDNA). This generates a so-called $target^{(rescue)}$ allele (Table 3, see Figs. S2–S4 and Table S2) that should fully restore the target locus both molecularly and functionally. Indeed, $sdt^{(rescue)}$, $lgl^{(rescue)\text{-}FRT}$ (see Fig. S3 *C* and *D*), $DE\text{-}Cad^{(rescue)}$, and $crb^{(rescue)}$ alleles fully complemented the lethality of their founder lines and known mutations, while $dArf6^{(rescue)}$ and $CG31158^{(rescue)}$ alleles fully rescued the recessive male sterile phenotype in their founder knock-out lines and in mutants we generated previously (see Table S2, and *SI Materials and Methods*). Furthermore, homozygotes of each of these $target^{(rescue)}$ alleles were viable, healthy and fertile (see Figs. S2–S4), confirming that each of these $target^{(rescue)}$ alleles fully substitutes the original allele throughout development. Finally, our quantitative Western blot analyses showed that there is no significant change of DE-Cad expression levels in $DE\text{-}Cad^{(rescue)}$ homozygous embryos compared to the wild type (Fig. 2*A* and *B*), even though $DE\text{-}Cad^{(rescue)}$ contains an attR in the nonconserved region of the first intron and a loxP after the 3′ UTR (see Fig. S2*C*). Thus, these strategically placed minimal recombination sites in the target loci do not interfere with target gene's function and expression.

Maximizing the Efficiency of ϕC31-Mediated DNA Integration in Founder Lines. The efficiency of ϕC31-mediated integration in founder lines is essential for facile modifications of target loci, as promised by the genomic engineering approach. In our early practice of generating rescue and modified alleles of *lgl*, *DE-Cad*, *crb*, and *CG31158*, we found that the integration efficiency of the attP sites in their founder lines averaged around 1.4% based on the standard ϕC31-integration protocol using ϕC31 mRNA and DNA mixtures (6) (see Table 3). Such efficiency is sufficient for genomic engineering practice (see Table S2), but it is much lower

Table 3. Efficiency of ϕC31-mediated DNA integration in founder lines

Founder line	Number of constructs injected[a]	Survival rate of microinjected embryos	Integration efficiency
$DE\text{-}Cad^{GX23w[-]}/CyO$	6	16.8% (± 3.2%)	1.4% (± 0.6%)
$DE\text{-}Cad^{GX6w[-]}/CyO$	3	23.7% (± 4.9%)	1.3% (± 1.2%)
$crb^{GX24w[-]}/CyO$	8	28.1% (± 3.0%)	0.7% (± 0.6%)
$lgl^{GX7w[-]}/CyO$	3	34.5% (± 3.1%)	0.9% (± 1.2%)
$CG31158^{GX6w[-]}/TM3\ Sb$	3	28.1% (± 1.5%)	2.9% (± 2.5%)
$DE\text{-}Cad^{GX23w[-]}/CyO;\ vasa\text{-}\phi C31^{ZH\text{-}102D}/+$	30	19.4% (± 5.4%)	7.1% (± 3.4%)
$crb^{GX24w[-]}/CyO;\ vasa\text{-}\phi C31^{ZH\text{-}102D}/+$	21	28.8% (± 7.5%)	6.3% (± 2.8%)
$lgl^{GX7w[-]}/CyO;\ vasa\text{-}\phi C31^{ZH\text{-}102D}/+$	7	27.5% (± 8.5%)	1.1% (± 0.6%)
$vasa\text{-}\phi C31^{ZH\text{-}2A}/vasa\text{-}\phi C3^{ZH\text{-}2A};\ dArf6^{GX16w[-]}/CyO;$	3	7.4% (± 2.5%)[b]	9.7% (± 5.4%)
$sdt^{GX73w[-]}/FM7;\ vasa\text{-}\phi C31^{ZH\text{-}102D}/+$	1	21.7% (± n/a)	5.6%[c] (± n/a)

[a]These constructs were based on pGE-attB or pGE-attB^{GMR} (see Table S2).
[b]The low survival rate in $vasa\text{-}\phi C31^{ZH\text{-}2A}/vasa\text{-}\phi C3^{ZH\text{-}2A};\ dArf6^{GX16w[-]}/CyO$ could be related to the homozygous copies of $vasa\text{-}\phi C31^{ZH\text{-}2A}$ on the X chromosome. We have found that having homozygous copies of *vasa-ϕC31* may adversely affect the healthiness and survival rate of micro-injected embryos of several founder lines, such as $DE\text{-}Cad^{GX23w[-]}$ and $crb^{GX24w[-]}$, possibly because of certain unique interactions between *vasa-ϕC31* and the attP-50 landing sites or balancers in *dArf6, DE-Cad* and *crb* founder lines.
[c]Only $sdt^{GX73w[-]}/FM7$ females were used to set up crosses and to calculate the integration efficiency. *FM7/FM7* females and *FM7/Y* males were discarded since they did not carry the $sdt^{GX73w[-]}$ chromosome.

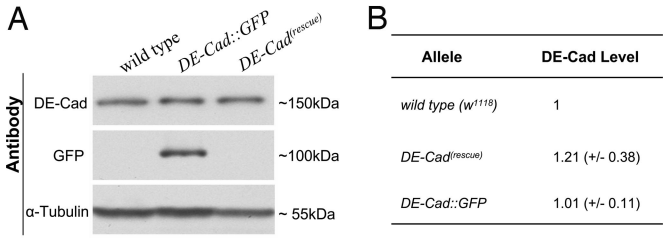

Fig. 2. Quantification of the DE-Cad expression levels in wild type, *DE-Cad*$^{(rescue)}$, and *DE-Cad*::*GFP*. (*A*) A sample Western blot showing the DE-Cad expression levels in wild type, *DE-Cad*::*GFP* (see Fig. 2 *A* and *B*), and *DE-Cad*$^{(rescue)}$ homozygous embryos. Embryos were in mixed stages (24-h collection under 25 °C). (*Top*) Because the majority of DE-Cad proteins undergoes internal cleavage (25), the rat anti-DE-Cad monoclonal antibody (DCAD2) (25) recognizes a single 150kDa band in all 3 samples. (*Middle*) To confirm the identity of DE-Cad::GFP, we also blotted the same samples with anti-GFP antibody. Only in *DE-Cad*::*GFP* sample the antibody recognized a single band around 100 kDa, which corresponds to the carboxyl-terminal half of cleaved DE-Cad::GFP (26). (*Bottom*) α-tublulin was used as loading controls. (*B*) Quantitative measurements of DE-Cad protein levels in wild type, *DE-Cad*$^{(rescue)}$, and *DE-Cad*::*GFP* that were based on multiple Western blot results (*DE-Cad*$^{(rescue)}$: $n = 3$; *DE-Cad*::*GFP*: $n = 4$).

than the average 14% obtained in our attP-50 host lines generated via random P-element insertions (see Table S1). Thus, compared to those attP sites that tend to associate with transposon hot-spots, attP sites inserted into an arbitrary chromosomal location by homologous recombination may indeed be less efficient. Nonetheless, by introducing into these founder lines the *vasa-φC31* transgene, which provides germ-line-specific expression of φC31 integrase (7), the integration efficiency can be drastically increased as much as 9 times as in the case of *crb* founder line (see Table 3). The only exception was *lgl*$^{GX7w[-]}$ (see Table 3), and we speculate that it is either because of the extreme chromosomal location of *lgl*, which is at the very left tip of the second chromosome, or because of the *lgl* genomic locus itself, as it was well documented that the efficiency of attP docking site can suffer strong position effects in the *Drosophila* genome (7, 16, 17).

It is also noteworthy that in the *CG31158*$^{GX6[w-]}$ founder line we initially failed to recover any integration events of the pGE-CG31158(rescue) construct after screening nearly 400 injected adults, while the pGE-attB vector showed decent integration efficiency (see Table S2). These results led us to suspect that the deleted gDNA of *CG31158* may contain strong yet encrypted transcription repressors that inhibit the expression of transgenic marker *w*+. Indeed, by using a new pGE-attB^{GMR} vector in which the *w*+ expression is enhanced by a strong eye-specific enhancer GMR (12, 18), we were able to recover integration events of pGE-attB^{GMR}-CG31158(rescue) at a comparable efficiency to pGE-attB (see Table S2). We later routinely used *vasa-φC31* and pGE-attB^{GMR} to maximize the efficiency and recovery of DNA integration in founder lines. Consistent with previous reports (6, 7, 16), pseudo integration events by φC31 were very rare: in nearly 300 integration events characterized, only 3 of them were found to be nonspecific (see Table S2).

Generation of Unique Genetic Alleles of *lgl*, *DE-Cad*, *crb*, *dArf6*, and *CG31158* by Genomic Engineering. As a proof of the exceptional experimental efficiency and versatility of genomic engineering approach, we have generated an extensive array of nearly 80 unique genetic alleles of *lgl*, *DE-Cad*, *crb*, *dArf6*, and *CG31158* that were tailored to our specific experimental needs. Genomic engineering opens the avenue for many previously difficult or impossible genetic assays. For example, the efficiency of genomic engineering made it an easy practice for us to generate multiple fluorescent protein knock-in alleles of *DE-Cad*, which encodes a core component of adherens junction complex, and *crb*, which encodes a large transmembrane protein of 30 EGF repeats (Fig. 3*A*), to screen for ideal fluorescent markers for their live imaging assays. As shown in Fig. 3, approximately half of these alleles have been validated by genetic and cell biological analyses (Fig. 3 *B–H*), and we are currently using the *DE-Cad*::*GFP* and *DE-Cad*::*PAGFP* to investigate the dynamics and trafficking of AJ in live epithelial cells. GFP knock-in alleles of *lgl*, *CG31158*, and *dArf6* generated by genomic engineering are also fully functional, homozygous viable, and fertile (Fig. 4). Because antibodies against CG31158 and dArf6 (15) were not successful in immunofluorescence studies, the GFP knock-in alleles made it possible to directly visualize their endogenous developmental

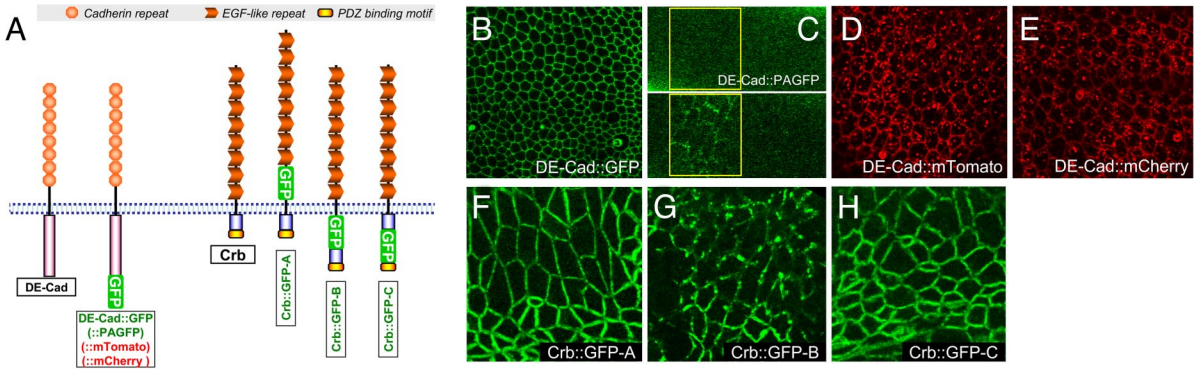

Fig. 3. Fluorescent knock-in alleles of *DE-Cadherin* and *crumbs*. (*A*) Protein domain structures of DE-Cad, Crb, and their fluorescent knock-in alleles. In all *DE-Cad* knock-in alleles, the fluorescent proteins were fused to the C terminus. In 3 Crb::GFP alleles, the GFP was inserted at 2,121 aa (Crb:GFP-A), 2,156 aa (Crb::GFP-B), and 2,189 aa (Crb::GFP-C), respectively. Note that not all of the 30 EGF repeats of Crb are drawn. (*B–E*) Subcellular localization patterns of DE-Cad::GFP, DE-Cad::PAGFP (photoactivatable GFP), DE-Cad::mTomato, DE-Cad::mCherry in live pupal (*B*, *D*, and *E*) or late embryonic (*C*) epithelia. All alleles rescued *DE-Cad* founder lines and were homozygous-viable, but only DE-Cad::GFP and DE-Cad::PAGFP showed clean localization at the adherens junction (*B* and *C*). Note the intracellular aggregates of DE-Cad::mTomato and DE-Cad::mCherry in (*D*) and (*E*). In (*C*) the yellow boxes highlight the region before (*Top*) and after (*Bottom*) the UV laser irradiation in the same sample. DE-Cad::PAGFP is only fluorescent after UV irradiation. *DE-Cad*::*GFP* knock-in homozygotes provide a clean and homogenous background of DE-Cad::GFP, whose expression level is virtually identical to the DE-Cad in wild type (see Fig. 2). (*F–H*) Subcellular localization of Crb::GFP-A, Crb::GFP-B, and Crb::GFP-C in live embryonic epithelia (stage 11). *crb*::*GFP-A* and *crb*::*GFP-C* complemented *crb*^{GX24} and were homozygous viable. They show normal localization along the apical-lateral boundary of the epithelial cells (*F* and *H*). In contrast, Crb::GFP-B shows a disrupted localization pattern (*G*). *crb*::*GFP-B* failed to complement *crb*^{GX24} and *crb*^{11A22}, and is homozygous lethal. All images were taken as the tangential view of the epithelia.

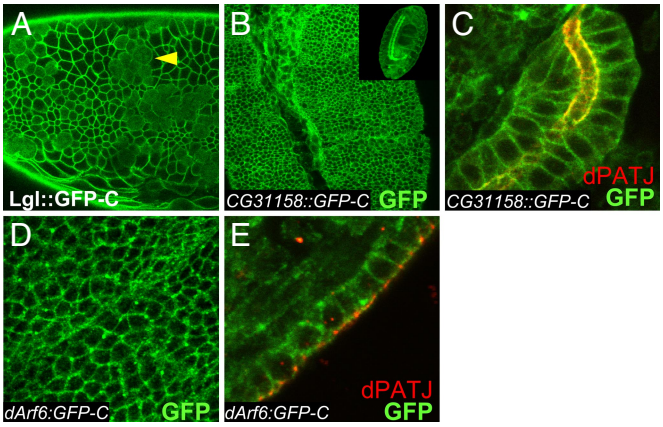

Fig. 4. Tissue and subcellular localization patterns of Lgl::GFP, CG31158::GFP, and dArf6::GFP knock-in mutants. (*A*) Subcellular localization of Lgl::GFP-C in live stage 11 embryonic epithelial cell. Lgl::GPF-C shows localization along the basolateral cortex in postmitotic cells, but is diffused in mitotic cells (one of them highlighted by the *yellow arrowhead*), consistent with previous reports based on Lgl antibodies (27). (*B–E*) Because both CG31158::GFP and dArf6::GFP-C are too weak to be directly detectable in live embryos by confocal microscopy, embryos are immunostained with anti-GFP antibody. (*B*) In this tangential-section view of embryonic epithelial cells, CG31158::GFP-C is cytoplasmic but predominantly cortical. It also shows strong expression in CNS in late stage embryos (*Inset*). (*C*) The subcellular localization of CG31158. In this cross-section view of embryonic gut epithelial cells, the apical polarity protein dPATJ (*red*) is seen exclusively at the apical side facing the lumen, while CG31158::GFP-C (*green*) localizes all around cell cortex. (*D*) Immunofluorescence by anti-GFP antibody visualizes dArf6::GFP-C has a punctuated pattern along the cell cortex or membrane in this tangential-section view of embryonic epithelial cells. dArf6::GFP-C does not show strong CNS expression in later embryos. (*E*) Unlike CG31158::GFP-C, dArf6::GFP-C (*green*) does not overlap extensively with dPATJ (*red*) in this cross-section view of embryonic epithelial cells.

and subcellular expression patterns (see Fig. 4 *B–E*). Besides fluorescent protein knock-in alleles for imaging and immunofluorescence assays, as listed in Table S2, we have generated high-affinity epitope fusion alleles for identifying in vivo protein interactions by proteomics, and alleles carrying specific point mutations and deletions for investigating their specific functions in vivo.

Discussion

We developed a highly efficient genomic engineering approach that allows for directed and versatile modifications of genomic loci in *Drosophila*. Although genomic engineering requires the generation of founder lines by ends-out targeting at first, we would like to emphasize that it is evident in Table 2 that, with our improved ends-out targeting system, targeting experiments of $>10^{-6}$ HR frequency can be accomplished with significantly reduced labor and time requirement (12). Once the founder lines are obtained, generation of engineered alleles by ϕC31-mediated integration is highly efficient, flexible, and straightforward. In most cases the phenotypes of engineered alleles can be examined immediately after the recovery of their integration events. When necessary, the optional removal of *w+* and vector sequences from engineered alleles is virtually 100% efficient and precise by the nature of loxP recombination (*SI Materials and Methods*). As an alternative, we also considered applying recombinase-mediated cassette exchange (RMCE)–based DNA integration (10, 11) in genomic engineering to eliminate the need to remove the *w+* in founder knock-out lines (see Fig. 1 *B* and *C*) and in integration alleles (see Fig. 1 *E* and *F*). For example, by flanking the *w+* marker with a pair of attP-50 in the founder knock-out lines, RMCE-mediated DNA integration can be used to directly replace the *w+* with target allele DNA flanked by a pair of attB-53. Unfortunately, in contrast to the decent efficiency reported by RMCE using full-length attP and attB (10, 11), our initial tests using transgenic *attP-53-[w+]-attP-53* host lines showed that RMCE based on attP-50 and attB-53 was very inefficient (<1%). Thus, in this article we focused on the single attP/attB integration approach in our genomic engineering scheme. Nonetheless, additional dominant transgenic markers like *yellow+* (*y+*) can be used for ϕC31-mediated allele integration (see Fig. 1 *D* and *E*), eliminating the need of removing *w+* marker in the founder knock-out line (see Fig. 1 *B* and *C*). This modification can be easily implemented by replacing the *w+* marker in pGE-attB (or pGE-attB^{GMR}) with *y+*, and could make the whole genomic engineering process even faster. In addition, genomic engineering can be readily modified to accommodate special experimental needs. For example, for target loci that require very large deletions (>20 kb), the recombineering-based P[acman] vectors (16) can be readily adapted for cloning and modifying large DNA fragments in vitro. In rare cases, such as removing the attR or introducing modifications beyond the deleted region in the founder line, double strand break-induced recombination can be adapted into the genomic engineering scheme by adding an *I-Cre*I endonulcease site (1, 19) at the appropriate position in the engineered allele.

While this article was in preparation, Gao et al. reported an approach called "SIRT" that combines the ϕC31-mediated DNA integration with ends-in targeting (20). In contrast to genomic engineering, SIRT results a tandem duplication of target genes after the integration of the allele DNA construct, so an extra step of double strand break-induced recombination is required to produce every desired allele. Because the reduction is based on random recombination events and is not 100% efficient, for each allele multiple recombination events have to be screened and characterized to confirm they contain the expected reduction events (20). The reduction also requires significant overlapping homology between the duplicates, making it potentially difficult in the SIRT approach to introduce certain modifications, such as multiple mutations along a large stretch of the genomic locus.

Recently, Beumer et al. reported that injecting custom-designed target gene-specific zinc-finger nuclease (ZFN) with donor DNA into embryos can directly induce homologous recombination events at increased efficiency (21), bypassing the time-consuming steps of generating transgenic donor lines required in regular targeting experiments. However, because ZFN-induced gene targeting is based on homologous recombination, it will unlikely achieve the highly efficient and virtually unlimited modifications of the target gene through ϕC31-mediated DNA integration in genomic engineering. Generating target-specific ZFN may also require significant efforts, such as screening specialized ZFN libraries (22). Nonetheless, once it becomes truly universal and efficient, ZFN-induced gene targeting can be adopted into the genomic engineering scheme for generating founder knock-out lines, making genomic engineering even more efficient.

The power of integrase-based genetic manipulations offered by genomic engineering, SIRT, and recombineering-based P[acman] transgenesis (16) provide highly useful tools for *Drosophila* researchers to devise unique in vivo and in vitro assays on their biological questions, and will likely make *Drosophila* a more accessible and attractive system for non-fly researchers to consider as their biological model. We envision that a coordinated effort within the *Drosophila* community may eventually make key conserved genes in *Drosophila* available for genomic engineering by systematically generating the required founder lines. To this end, we are working to further improve the gene targeting in *Drosophila* by developing a dual-positive selection system based on a *w+* and neomycin-resistance gene that will

enrich the targeting candidates by 10 to 100 times, so even targeting experiments of very low HR frequency ($<10^{-6}$) may be reliably accomplished. Finally, given the fact that ϕC31 works well in mammalian cells (6), we imagine that similar strategy like genomic engineering can be readily implanted in mammalian systems.

Materials and Methods

Fly Stocks. The following stocks were obtained from the Bloomington stock center: BL#766 (y^1 w^{67c23} *P{Crey}1b*; noc^{Sco}/*CyO*), BL#851 (y^1 w^{67c23} *P{Crey}1b*; *D*/TM3, Sb*), BL#1092 (y^1 w^{67c23}; noc^{Sco}/*CyO, P{w[+mC] = Crew}DH1*), BL#13763 (y^1 w^{67c23}; $Arf51F^{KG02753}$), and BL#3085 (cn^1 shg^2 bw^1 sp^1/*CyO*). lgl^4 was a gift from F. Roegiers (Fox Chase Cancer Center); crb^{11A22} was a gift from K.-W. Choi (Baylor College) and U. Tepass (University of Toronto); sdt^{XP96} was a gift from E. Wieschause (Princeton University). *vasa-*$\phi C31^{ZH\text{-}102D}$ and *vasa-*$\phi C31^{ZH\text{-}2A}$ were provided by K. Basler (7). $dArf6^{\Delta KG\#1}$, used in Table 2, contains a 1.4-kb deletion induced by imprecise excision of P-element KG02753 that deletes all of the coding exons and the 3′UTR of *dArf6*. $CG31158^{KO\#1}$, used in Table 2, was an ends-out targeting mutant that contains a small 74-bp deletion in exon 8.

Ends-Out Gene Targeting. P-element based transgenesis was used to make transgenic donor lines for all of the targeting experiments. w^{1118} was used as the host strain. Gene-targeting experiments and PCR-verifications of targeting candidates were carried out as described in Huang et al. (12). Primers used for making targeting constructs are listed in Table S3. Primers used for PCR verifications as shown in Figs. S2 to S4 are available upon request.

ϕC31-Mediated DNA Integration in Founder Lines. To remove *w+* transgenic marker in founder knock-out lines, *hs-Cre* on X or second chromosome (i.e., BL#766, BL#851 and BL#1092) that constitutively expresses Cre recombinase was crossed into the founder knock-out lines. Single *w[–]* male progeny was then used to establish balanced stocks of $w^{[-]}$ founder lines that were also free of *hs-Cre*. We observed that the efficiency of loxP-recombination by Cre recombinase was virtually 100% in such crosses. To improve the integration efficiency, *w[–]* founder lines were also crossed with *vasa-*$\phi C31^{ZH\text{-}102D}$, which is on the fourth chromosome or *vasa-*$\phi C31^{ZH\text{-}2A}$, which is on the X chromosome (7). Only a single copy of *vasa-*ϕ*C31* was maintained in founder lines, as we found that having homozygous copies of *vasa-*ϕ*C31* could adversely affects the healthiness of founder lines and the survival rate of their embryos after microinjection. Because *vasa-*ϕ*C31* transgene is marked by *3xP3-eGFP* and *3xP3-RFP* (i.e., strong GFP and RFP expressions in the eyes) but not *w+*, it is fully compatible with the *w+* transgenic marker in pGE-attB or pGE-attB^{GMR}.

ϕC31-mediated integration in founder lines was carried out according to the published protocol (6, 16, 23). Mixtures of ϕC31 mRNA and plasmid DNA were always used, regardless whether the particular founder line carried *vasa-*ϕ*C31*. We used plasmid DNA purified from midi-prep (Qiagen), and it is likely that better integration efficiency may be achieved by using higher quality DNA from CsCl-purificaion (16). Before injection, embryos were dechorionated either manually or by 2 min treatment of 50% bleach. We found that the survival rate of microinjected embryos of a particular founder line often favored 1 of these 2 dechorionation methods. Integration events were recovered based on the *w+* marker and were genetically mapped and balanced to confirm that the *w+* is on the target chromosome.

ACKNOWLEDGMENTS. We thank Dr. Michele Calos (Stanford University School of Medicine, Stanford, CA) for ϕC31 cDNA constructs, Dr. Jeff Sekelsky (University of North Carolina, Chapel Hill, NC) for original pEndsOut targeting vectors, Dr. C.T. Wu (Harvard Medical School, Boston) and Dr. Konrad Basler (University of Zurich, Zurich) for ϕC31 reagents and fly stocks, Dr. Roger Tsien (Howard Hughes Medical Institute, University of California San Diego, La Jolla, CA) for mCherry and mTomato constructs, Dr. Jennifer Lippincott-Schwartz (National Institutes of Health, Bethesda, MD) for photoactivatable GFP construct, Chris Rosson and Pu Shi for technical assistances, and the Developmental Studies Hybridoma Bank for antibodies. We also thank Drs. Victor Ambros, Yuh-Nung Jan, Bing Ye, Sige Zou, Peter Soba, Gerard Campbell, Rebecca Yang, Yi Rao, Fen-Biao Gao, Huashun Li, and Simon Watkins for suggestions and comments on developing the genomic engineering. This work is supported by the start-up fund from University of Pittsburgh Medical School (Y.H.) and Grant 1R21RR024869 from National Center for Research Resources, National Institutes of Health (to Y.H.). Vectors for genomic engineering will be donated to the Drosophila Genomic Resource Center.

1. Rong YS, Golic KG (2000) Gene targeting by homologous recombination in Drosophila. *Science* 288:2013–2018.
2. Gong WJ, Golic KG (2003) Ends-out, or replacement, gene targeting in *Drosophila*. *Proc Natl Acad Sci USA* 100:2556–2561.
3. Venken KJ, Bellen HJ (2005) Emerging technologies for gene manipulation in *Drosophila melanogaster*. *Nat Rev Genet* 6:167–178.
4. Adams MD, Sekelsky JJ (2002) From sequence to phenotype: reverse genetics in *Drosophila melanogaster*. *Nat Rev Genet* 3:189–198.
5. Margolis B, Borg JP (2005) Apicobasal polarity complexes. *J Cell Sci* 118:5157–5159.
6. Groth AC, Fish M, Nusse R, Calos MP (2004) Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* 166:1775–1782.
7. Bischof J, Maeda RK, Hediger M, Karch F, Basler K (2007) An optimized transgenesis system for *Drosophila* using germ-line-specific {varphi}C31 integrases. *Proc Natl Acad Sci USA* 104:3312–3317.
8. Groth AC, Calos MP (2004) Phage integrases: biology and applications. *J Mol Biol* 335:667–678.
9. Groth AC, Olivares EC, Thyagarajan B, Calos MP (2000) A phage integrase directs efficient site-specific integration in human cells. *Proc Natl Acad Sci USA* 97:5995–6000.
10. Bateman JR, Lee AM, Wu CT (2006) Site-specific transformation of *Drosophila* via phiC31 integrase-mediated cassette exchange. *Genetics* 173:769–777.
11. Bateman JR, Wu CT (2008) A simple polymerase chain reaction-based method for the construction of recombinase-mediated cassette exchange donor vectors. *Genetics* 180:1763–1766.
12. Huang J, Zhou W, Watson AM, Jan Y-N, Hong Y (2008) Efficient ends-out gene targeting in *Drosophila*. *Genetics* 180:703–707.
13. Nelson WJ (2003) Adaptation of core mechanisms to generate cell polarity. *Nature* 422:766–774.
14. Suzuki A, Ohno S (2006) The PAR-aPKC system: lessons in polarity. *J Cell Sci* 119:979–987.
15. Dyer N, et al. (2007) Spermatocyte cytokinesis requires rapid membrane addition mediated by ARF6 on central spindle recycling endosomes. *Development* 134:4437–4447.
16. Venken KJT, He Y, Hoskins RA, Bellen HJ (2006) P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* 314:1747–1751.
17. Markstein M, Pitsouli C, Villalta C, Celniker SE, Perrimon N (2008) Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nat Genet* 40:476–483.
18. Hay BA, Wolff T, Rubin GM (1994) Expression of baculovirus P35 prevents cell death in *Drosophila*. *Development* 120:2121–2129.
19. Rong YS, et al. (2002) Targeted mutagenesis by homologous recombination in *D. melanogaster*. *Genes Dev* 16:1568–1581.
20. Gao G, McMahon C, Chen J, Rong YS (2008) A powerful method combining homologous recombination and site-specific recombination for targeted mutagenesis in *Drosophila*. *Proc Natl Acad Sci USA* 105:13999–14004.
21. Beumer KJ, et al. (2008) Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc Natl Acad Sci USA* 105:19821–19826.
22. Ramirez CL, et al. (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Meth* 5:374–375.
23. Fish MP, Groth AC, Calos MP, Nusse R (2007) Creating transgenic *Drosophila* by microinjecting the site-specific ϕC31 integrase mRNA and a transgene-containing donor plasmid. *Nat Protocols* 2:2325–2331.
24. Hong Y, Stronach B, Perrimon N, Jan LY, Jan YN (2001) *Drosophila Stardust* interacts with *Crumbs* to control polarity of epithelia but not neuroblasts. *Nature* 414:634–638.
25. Oda H, Tsukita S (1999) Nonchordate classic cadherins have a structurally and functionally unique domain that is absent from chordate classic cadherins. *Dev Biol* 216:406–422.
26. Oda H, Tsukita S (2001) Real-time imaging of cell-cell adherens junctions reveals that *Drosophila* mesoderm invagination begins with two phases of apical constriction of cells. *J Cell Sci* 114:493–501.
27. Betschinger J, Eisenhaber F, Knoblich JA (2005) Phosphorylation-induced autoinhibition regulates the cytoskeletal protein lethal (2) giant larvae. *Curr Biol* 15:276–282.

Supporting Information

Huang et al. 10.1073/pnas.0900641106

SI Materials and Methods

Testing the DNA Integration by Minimal attP-50 and attB-53 in *Drosophila*. We first modified pKIKO (1) to generate a P-element-based construct pAttPC that carries a 50 bp minimal attP-50 site. pAttPC was then used to generate multiple independent transgenic host lines through P-element based transgenesis. The *w+* marker in these attP-50 host lines was then removed by loxP-recombination by crossing with *hs-Cre* fly stocks [detail cross set-up is similar to Maggert et al. (2)]. Following the protocol described by Groth et al. (3), embryos from 4 of these *w[−]* minimal attP host lines were collected and microinjected with a mixture of ϕC31 mRNA and pGE-attB bearing the minimal attB-53. Integration of pGE-attB was first verified by the recovery and genetic mapping of *w+* marker. We further PCR-amplified the integration locus in several integration lines and confirmed the presence of minimal attR site by sequencing.

Generation of Genomic Engineering Vectors and Targeting Constructs. pGX-attP was generated by inserting a minimal attP-50 site into the targeting vector pRK2 (1). pGE-attB was generated by inserting a minimal attB-53 site to a modified targeting vector pKIKO (1). pGE-attB^{GMR} was generated by inserting a GMR enhancer (1) into the pGE-attB. Details about the constructions of these vectors are available upon request. Molecular cloning of targeting constructs of *sdt*, *lgl*, *DE-Cad*, *crb*, *CG31158*, and *dArf6* was carried out according to the protocols described in Huang et al. (1). In brief, 5′ and 3′ homologous arms of each targeting construct were amplified from the corresponding BAC clones (available from Children's Hospital, Oakland, California) by long-range PCR using PfuUltra DNA polymerase (Stratagene). Because of historic reasons, there are several versions of pGX-attP vectors in our laboratory that carry different features for targeting. For *lgl*, *DE-Cad*, *crb*, and *CG31158*, PCR products were cloned into an early version of pGE-attP, which was based on targeting vector pKIKO (1) and did not contain *UAS-Rpr* and GMR-enhanced *w+*. PCR products for *dArf6* targeting were cloned into pRK2 (1) with an added attP-50 (1). PCR products for *sdt* targeting were cloned into the pGX-attP that is shown in supporting information (SI) Fig. S1. All homologous arms were sequenced to verify that there were no PCR errors in the coding sequences. In our practice, PfuUltra polymerase yielded an exceptionally low rate of PCR errors, as we discovered no PCR errors within total ≈106-kb PCR products that were sequenced. Primers used for making targeting constructs are listed in Table S3. We used www.fruitfly.org to compare genomic sequences between *Drosophila melanoganster* and *Drosophila pseudoobscura*, to identify apparently nonconserved noncoding regions for positioning the attP-50 and loxP sites in the target locus.

All constructs for generating genomic engineering alleles in Table S2 were carried out according to standard molecular cloning protocol. Because of the very large number of constructs involved, details, such as primers or sequences and so forth are available upon request. Primers used for making *target$^{(rescue)}$* alleles are listed in Table S3.

Molecular and Genetic Characterizations of Engineered Alleles. To molecularly verify the engineered alleles, at least 2 short PCR reactions as shown in Figs. S2 to S4 were carried out to detect the presence of attR and attL+loxP sequences in the alleles. To remove the *w+* marker and extra vector sequences, alleles were crossed into appropriate *hs-Cre* lines, as described above. The removal of *w+* by Cre was 100% efficient and precise in all of the alleles we generated so far. A single *w[−]* male progeny was used to establish balanced stocks of *w[−]* alleles that are free of *hs-Cre* and *vasa-ϕC31*. Chromosomes carrying *vasa-ϕC31* can be easily excluded during genetic crosses by selecting against its strong GFP and RFP expression in the eyes under a fluorescent dissecting scope. All alleles listed in Table S2 were free of *vasa-ϕC31*.

Complementation ("rescue") assays were carried out by first testing whether the allele, with or without *w+* maker removed, could complement the corresponding founder line or known null alleles (see Table 2). Alleles succeeded in such complementation tests were further tested for being homozygous-viable and fertile. In the cases of *CG31158* and *dArf6*, alleles were only tested for being homozygous-fertile at this step, as *CG31158^{GX6}* and *dArf6^{GX16}* were not lethal.

Live Imaging and Immunohistology. Live imaging in embryos were done by placing the embryos in air-permeable chambers filled with halocarbon oil (#95) on specially made slides, to ensure that they will continue normal development throughout the imaging session (4). Live imaging on pupal epithelial cells was carried out as described (5). Immunostaining experiments on embryos were carried out according to standard protocol (6). The following primary antibodies were used: rabbit anti-GFP which was made against purified EGFP with His-tag, 1:1,500; mouse anti-PATJ, which was made against purified GST::dPATJ fusion protein, 1:500. Secondary antibodies were Cy2- or Cy3-conjugated goat anti-Rabbit IgG or Goat anti-mouse IgG (The Jackson Lab). Images were collected on Leica TCS-NT and Olympus FV1000 confocal microscopes (Center for Biologic Imaging, University of Pittsburgh Medical School).

Quantitative Western Blot Analyses on DE-Cad Protein Expression Levels. Western blot was used to compare the DE-Cad protein level in mixed-staged embryos of *w^{1118}* (wild-type control), *DE-Cad$^{(rescue)}$*, and *DE-Cad::GFP*. Embryos were collected for 24 h under 25 °C, dechorionated by bleach, and were homogenized in lysis buffer [25 mM Tris pH8.0, 27.5 mM NaCl, 20 mM KCl, 25 mM sucrose, 10 mM EDTA, 10 mM EGTA, 10% glycerol, 0.5% Nonidet P-40, 1% Triton X-100, with 1 mM DL-DTT(DTT), 1× CPIM protease inhibitor (Roche) and 1 mM PMSF added immediately before homogenization]. After letting crude lysates sit on ice for 30 min, equal volume of 2× loading buffer [100 mM Tris-HCl(pH6.8), 4% SDS, 0.2% bromophenol blue, 20% glycerol, 200 mM β-mercaptoethanol] was added to the lysates. Lysates were then centrifuged at 20,000 × *g* for 15 min and supernatants were boiled and loaded on 8% SDS/PAGE. Standard Western blot protocol was followed (7). Multiple exposures were made on Kodak BioMax MR films and only properly exposed bands were scanned and measured in Image J. DE-Cad protein levels were measured by bands recognized by DCAD2 (1:10; Developmental Studies Hybridoma Bank), and normalized against the intensity of α-tubulin bands recognized by mouse anti-α-tubulin (AA4.3, Developmental Studies Hybridoma Bank, 1:5000). Rabbit anti-GFP in Fig. 2*A* was used at 1:10,000 dilution.

1. Huang J, Zhou W, Watson AM, Jan Y-N, Hong Y (2008) Efficient ends-out gene targeting in *Drosophila*. *Genetics* 180:703–707.
2. Maggert KA, Gong WJ, Golic KG (2008) Methods for homologous recombination in *Drosophila*. *Methods in Molecular Biology*, ed Dahmann C (Clifton, NJ) 420:155–174.
3. Groth AC, Fish M, Nusse R, Calos MP (2004) Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* 166:1775–1782.
4. Lu B, Roegiers F, Jan LY, Jan YN (2001) Adherens junctions inhibit asymmetric division in the *Drosophila* epithelium. *Nature* 409:522–525.
5. Roegiers F, Younger-Shepherd S, Jan LY, Jan YN (2001) Two types of asymmetric divisions in the *Drosophila* sensory organ precursor cell lineage. *Nat Cell Biol* 3:58–67.
6. Hong Y, Stronach B, Perrimon N, Jan LY, Jan YN (2001) *Drosophila* Stardust interacts with Crumbs to control polarity of epithelia but not neuroblasts. *Nature* 414:634–638.
7. Wodarz A (2008) Extraction and immunoblotting of proteins from embryos. *Methods in Molecular Biology*, ed Dahmann C (Clifton, NJ) 420:335–345.

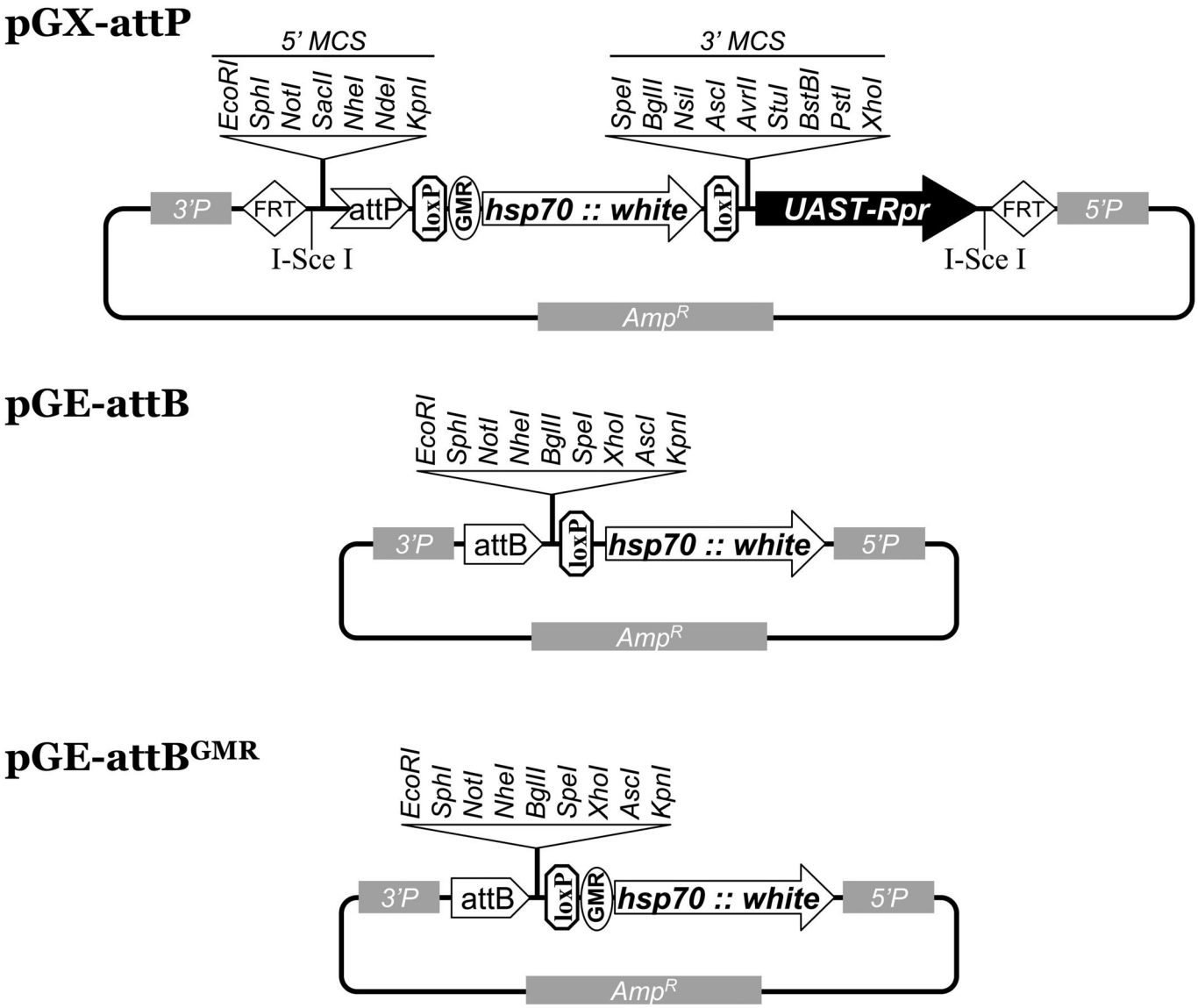

Fig. S1. Vectors for genomic engineering: pGX-attP, pGE-attB, and pGE-attB^{GMR}. MCS: multiple cloning sites. *Amp^R*: ampicillin-resistant gene. *3′P* and *5′P*: 3′ and 5′ P-element sequences for transgenic insertion. Although 3′P and 5′P are not necessary in pGE-attB and pGE-attB^{GMR} for the genomic engineering purpose, they are kept so pGE-attB and pGE-attB^{GMR} can also be used for P-elements-based transgenesis. The only difference between pGE-attB and pGE-attB^{GMR} is the GMR enhancer in the latter. Note that there is also a GMR enhancer in pGX-attP. GenBank accession numbers for the DNA sequences of pGE-attP, pGE-attB, and pGE-attB^{GMR}: FJ791035, FJ791036, and FJ791037.

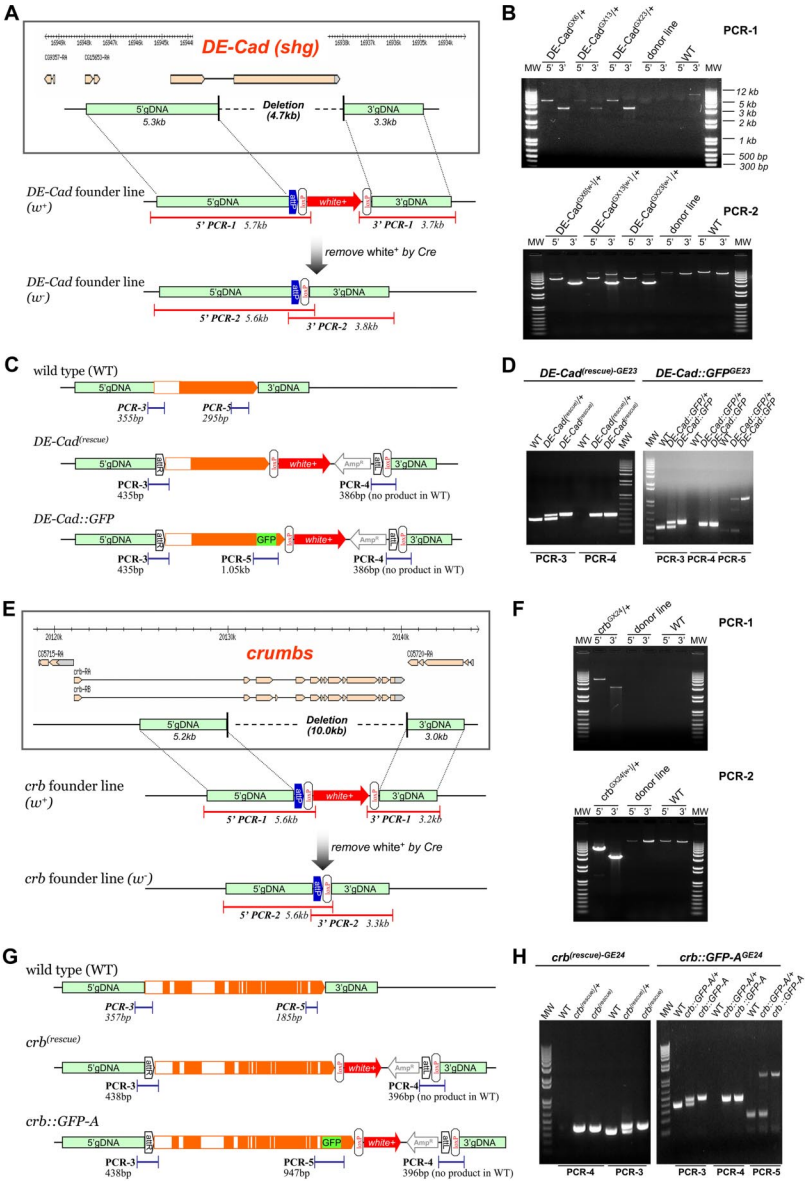

Fig. S2. Genomic engineering of *DE-Cad* and *crb*. (*A–D*) Generation and verification of *DE-Cad* founder lines, *DE-Cad*$^{(rescue)}$ and *DE-Cad*::*GFP* alleles. (*A*) Targeting design and PCR verification scheme of the *DE-Cad* founder line. Boxed is the targeting design of *DE-Cad* founder lines. Green boxes are the genomic DNA (gDNA) regions used for 5′ and 3′ gDNA arms in the targeting construct. In the *DE-Cad* founder knock-out line, a 4.8-kb genomic DNA of *DE-Cad* was deleted, which includes the 3′ UTR and all of the coding sequences for DE-Cad protein except for the signal peptide and the first cadherin repeat (1–209aa), which was coded by the first exon. The original *DE-Cad* founder knock-out lines carrying *w*+ marker are verified by 5′ and 3′ PCR-1. PCR-1 is designed with one primer annealing within the *w*+, while another primer anneals outside the genomic DNA (gDNA) region used for homologous arms (''5′ gDNA'' or ''3′ gDNA'') in targeting construct. Thus, only the expected targeting events will yield PCR products of expected size. *DE-Cad* founder lines with *w*+ removed are further verified by 5′ and 3′ PCR-2. The 5′ or 3′ PCR-2 has one primer annealing at the far side of deletion, while another primer anneals outside the corresponding 5′ or 3′ gDNA region. (*B*) (*Top*) PCR-1 results of founder lines *DE-Cad*^{GX6}, *DE-Cad*^{GX13}, and *DE-Cad*^{GX23} carrying *w*+. There is a faint band in the 3′PCR-1 of wild type (WT), but its size is much bigger than 3.6 kb, so it appears to be nonspecific. (*Bottom*) PCR-2 results of founder lines *DE-Cad*$^{GX6w[-]}$, *DE-Cad*$^{GX13w[-]}$, and *DE-Cad*$^{GX23[-]}$ (with *w*+ marker removed). Donor line: the original transgenic line that carries the donor DNA used for *DE-Cad* targeting. Donor line and WT yielded PCR-2 products that are about 4.7 kb bigger than those produced by founder lines. All long range PCR reactions in this figure and in Figs. S3 and S4 were carried out with Roche 20-kb-plus PCR kit. (*C*) PCR verification scheme of *DE-Cad*$^{(rescue)GE23}$ and *DE-Cad*::*GFP*^{GE23} alleles. PCR-3 and PCR-4 are designed to confirm the 5′ and 3′ attP/attB recombination events, respectively. PCR-5 is designed to confirm the GFP insertion in *DE-Cad*::*GFP* alleles. (*D*) (*Left*) PCR-3 and PCR-4 results from WT, heterozygotes, and homozygotes of *DE-Cad(resuce)*^{GE23}. (*Right*) PCR-3,-4,-5 results from WT, heterozygotes, and homozygotes of *DE-Cad*::*GFP*^{GE23}. The ''*GE23*'' suffix indicates that these alleles were generated in founder line *DE-Cad*$^{GX23w[-]}$. (*E–H*) Generation and verification of *crb* founder lines, *crb*$^{(rescue)}$ and *crb*::*GFP-A* alleles. (*E*) Targeting design and PCR verification scheme of the *crb* founder line. A 10.0-kb genomic DNA of *crb* was deleted in the founder knock-out line. The deletion includes the 3′ UTR and all of the coding sequences for Crb except for the first 80aa. *crb*-specific primers are used in PCR-1 and PCR-2. (*F*) (*Top*) PCR-1 results of founder line *crb*^{GX24} carrying *w*+. (*Bottom*) PCR-2 results of founder line *crb*$^{GX24w[-]}$. Donor line and WT yielded PCR products that are about 10-kb bigger than those produced by founder lines. (*G*) PCR verification scheme of *crb*$^{(rescue)GE24}$ and *crb*::*GFP-A*^{GE24} alleles by PCR-3, PCR-4, and PCR-5 using *crb*-specific primers. (*H*) PCR-3, PCR-4, and PCR-5 results from WT, heterozygotes, and homozygotes of *crb*$^{(rescue)-GE24}$ or *crb*::*GFP-A*^{GE24}. MW: Invitrogen 1kp-plus DNA marker (same in all gel images in Figs. S2–S4). *Amp*^{R}: vector sequence

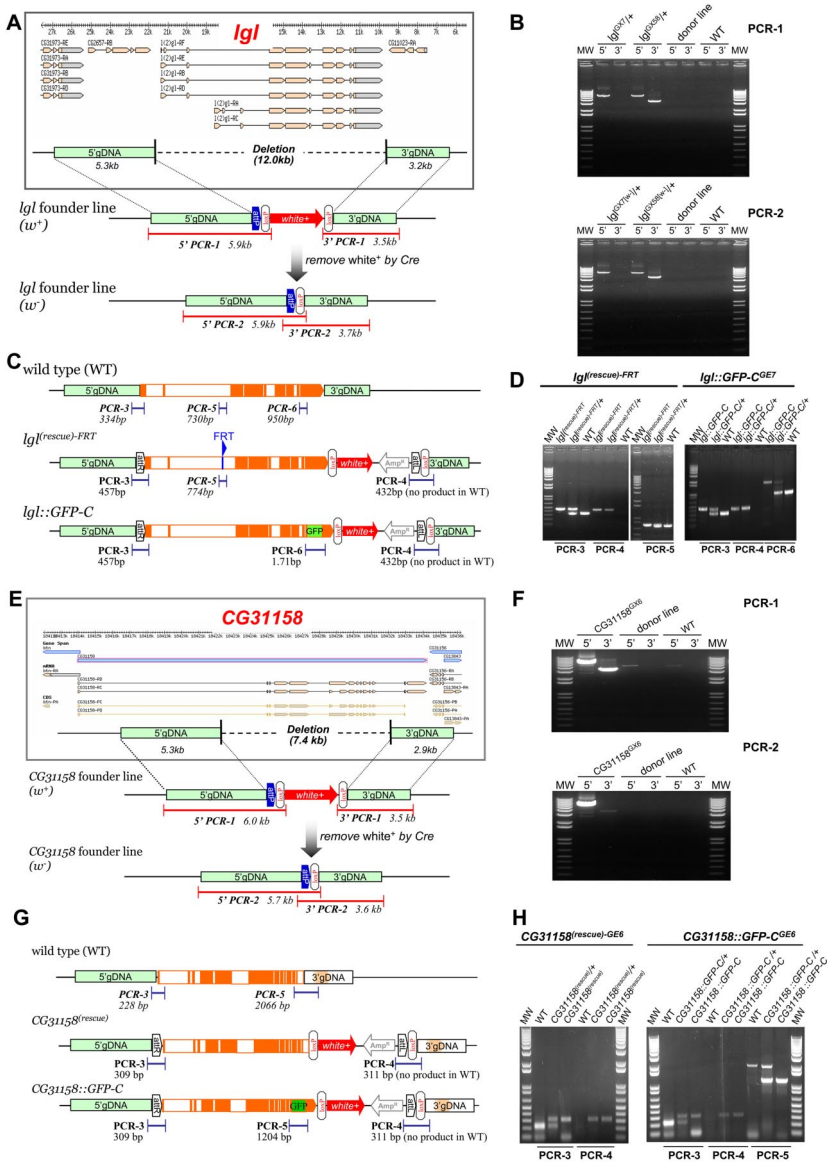

Fig. S3. Genomic engineering of *lgl* and *CG31158*. (*A–D*) Generation and verification of *lgl* founder lines, *lgl*$^{(rescue)\text{-}FRT}$, and *lgl*::*GFP-C* alleles. (*A*) Targeting design and PCR verification scheme of the *lgl* founder line. In the *lgl* founder line, 12-kb gDNA that contains the whole-coding sequence of *lgl* plus 3′UTR and ≈5 kb of *lgl* upstream sequence was deleted. *lgl*-specific primers are used in PCR-1 and PCR-2. (*B*) (*Top*) PCR-1 results of founder lines *lgl*^{GX7} *and lgl*^{GX58} carrying *w*+. Both are positive for 5′PCR-1, but only *lgl*^{GX58} is positive for 3′PCR-1. (*Bottom*) PCR-2 results of *lgl*$^{GX7[w-]}$ and *lgl*$^{GX58w[-]}$ founder lines. Both are positive for 5′PCR-2, but only *lgl*^{GX58} is positive for 3′PCR-2. Among 22 *lgl*^{GX} founder lines, only *lgl*^{GX58} was verified positive in both 5′ and 3′ PCR-1 and PCR-2; the remaining founder lines were only positive in 5′ PCR-1 and 5′ PCR-2. However, all 22 *lgl*^{GX} lines failed to complement null *lgl*^{4} allele and all showed the distinct ''lethal giant larvae'' phonotype in their homozygous third instar larvae. (*C*) PCR verification scheme of *lgl*$^{(rescue)\text{-}FRT\text{-}GE7}$ and *lgl*::*GFP-C*^{GE7} alleles by PCR-3, PCR-4, PCR-5 (for verifying FRT insertion), and PCR-6 (for verifying GFP fusion) using *lgl*-specific primers. Because *lgl*^{GX58} turned out to be very unhealthy (likely because of some unrelated background mutations), we used *lgl*^{GX7} as the founder line for subsequent genomic engineering experiments. In addition, instead of a typical *lgl*$^{(rescue)}$ allele, we made a *lgl*$^{(rescue)\text{-}FRT}$ which contains a FRT site inserted 418 bp away from the ATG (of transcription isoforms D, E, and F). The rescue activity of this allele not only validated the *lgl*$^{GX7[w-]}$ founder line for future generation of engineered alleles, but also confirmed that the insertion of FRT does not disrupt the function of *lgl* locus. This information is critical for generating FRT/FLPase-based conditional alleles of *lgl* in the future. (*D*) PCR-3, PCR-4, PCR-5, and PCR-6 results from WT, heterozygotes, and homozygotes of *lgl*$^{(rescue)\text{-}FRT\text{-}GE7}$ and *lgl*::*GFP-C*^{GE7}. (*E–H*) Generation and verification of *CG31158* founder lines, *CG31185*$^{(rescue)}$ and *CG31158*::*GFP-C* alleles. (*E*) Targeting design and PCR verification scheme of the *CG31158* founder line. The targeted deletion of *CG31158* is 7.4-kb long that removes all of the 1,480 aa of CG31158 protein except for the last 6 amino acids in exon14. In CG31158 founder lines the loxP is placed in the nonconserved region of intron13, other than after the 3′UTR. *CG31158*-specific primers are used in PCR-1 and PCR-2. (*F*) (*Top*) PCR-1 results of founder line *CG31158*^{GX6} carrying *w*+. There is a faint nonspecific 5′ PCR-1 product (≈4 kb) in the donor line and WT. (*Bottom*) PCR-2 results of founder line *CG31158*$^{GX6w[-]}$. (*G*) PCR verification scheme of *CG31158*$^{(rescue)\text{-}GE6}$ and *CG31158*::*GFP-C*^{GE6} alleles by PCR-3, PCR-4, and PCR-5 using *CG31158*-specific primers. Note that in *CG31158*::*GFP-C* allele the exon14 and exon15 (mostly 3′UTR) were duplicated and joined directly to the exon13. The GFP was inserted after the last amino acid in exon14. (*H*) (*Left*) PCR-3 and PCR-4 results from WT, heterozygotes, and homozygotes of *CG31158*$^{(rescue)\text{-}GE6}$. (*Right*) PCR-3, PCR-4, and PCR-5 results from WT, heterozygotes, and homozygotes of *CG31158*::*GFP-C*^{GE6}. Because of the duplication of exon14 and exon15 in *CG31158*::*GFP-C*^{GE6}, the PCR-5 primer, which flanks exon13 and the exon14, yields a 2.066-kb product in WT, but a smaller 1.204-kb product in *CG31158*::*GFP-C*^{GE6} homozygotes.

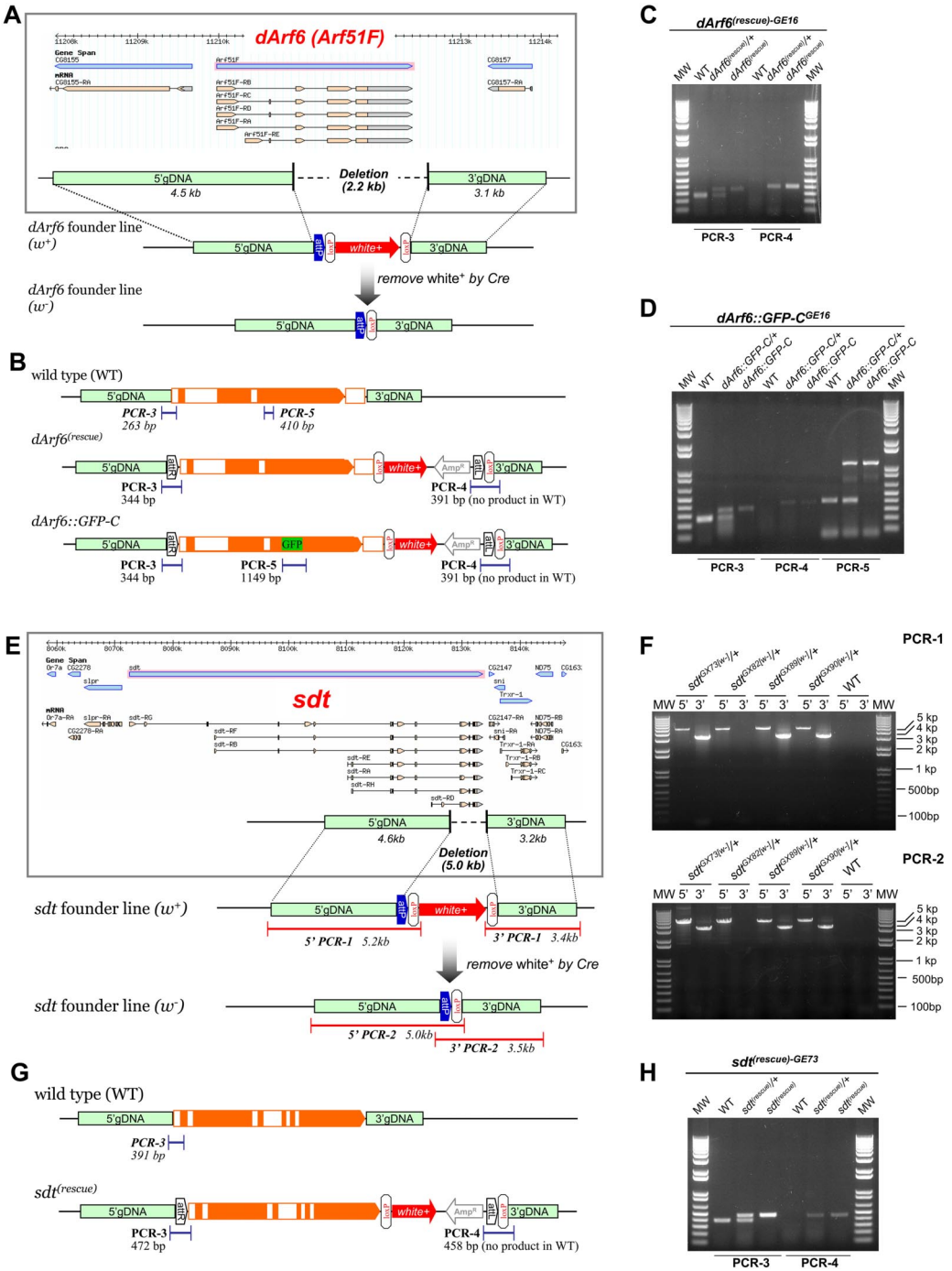

Fig. S4. Genomic engineering of *dArf6* and *sdt*. (*A–D*) Generation and verification of *dArf6* founder lines, *dArf6*$^{(rescue)}$ and *dArf6*::*GFP-C* alleles. (*A*) Targeting design of the *dArf6* founder line. The targeted deletion of *dArf6* is of 2.2-kb long that removes all of the coding exons of *dArf6* plus its 3′UTR. PCR verifications of *dArf6* founder lines were published in Huang et al. [Huang J, Zhou W, Watson AM, Jan Y-N, Hong Y (2008) Efficient ends-out gene targeting in *Drosophila*. *Genetics* 180:703–707.] (*B*) PCR verification scheme of *dAf6*$^{(rescue)\text{-}GE16}$ and *dArf6*::*GFP-C*^{GE16} alleles by PCR-3, PCR-4, PCR-5 using *dArf6*-specific primers. (*C*) PCR-3 and PCR-4 results from WT, heterozygotes, and homozygotes of *dAf6*$^{(rescue)\text{-}GE16}$. (*D*) PCR-3, PCR-4, and PCR-5 results from WT, heterozygotes, and homozygotes of *dAf6*::*GFP-C*^{GE16}. (*E–H*) Generation and verification of *sdt* founder knock-out lines and *sdt*$^{(rescue)}$ allele. (*E*) Targeting design and PCR verification scheme of the *sdt* founder line. The targeted deletion of *sdt* is 5.0-kb long that removes exons coding all of the important protein domains, including PDZ, SH3, and guanylate kinase in the C-terminal half of *Stardust*. *sdt*-specific primers are used in PCR-1 and PCR-2. (*F*) (*Top*) PCR-1 results of founder lines *sdt*^{GX73}, *sdt*^{GX82}, *sdt*^{GX89}, and *sdt*^{GX90} carrying *w*+. (*Bottom*) PCR-2 results of the same founder lines with *w*+ marker removed. (*G*) PCR verification scheme of *sdt*$^{(rescue)\text{-}GE73}$ allele by PCR-3, and PCR-4 using *sdt*-specific primers. (*H*) PCR-3 and PCR-4 results from WT, heterozygotes, and homozygotes of *sdt*$^{(rescue)\text{-}GE73}$.

Table S1. Efficiency of ϕC31-mediated DNA integration through minimal attP-50 and attB-53

Host line	Chromosomal location of attP-50	Embryos injected	Larvae survived	Adults survived	Integration efficiency[a]
attP-50$^{\#1}$	second	560	170 (30%)	38 (22%)	32% (12/38)
attP-50$^{\#1A}$	second	150	77 (51%)	49 (64%)	12% (6/49)
attP-50$^{\#1A}$	second	300	210 (70%)	140 (67%)	11% (15/140)
attP-50$^{\#4}$	third	400	175 (44%)	104 (59%)	5% (5/104)
attP-50$^{\#10}$	third	350	128 (37%)	63 (49%)	11% (7/63)

[a]Integration efficiency was calculated according to Groth et al [Groth AC, Fish M, Nusse R, Calos MP (2004) Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* 166:1775–1782.]

Table S2. Genetic alleles generated by genomic engineering

Founder line	Construct	Vector	Allele description	Integration efficiency	*vasa-φC31*	Rescue
	φC31-attB-53	pGE-attB	test	1.2% (2/162)	N	N
	DE-Cad(rescue)	pGE-attB	rescue	2.0% (5/249)	N	Y
	DE-Cad::GFP	pGE-attB	*GFP* knock-in	1.0% (1/99)*	N	Y
	DE-Cad(rescue)	pGE-attB^{GMR}	rescue	1.7% (4/237)	N	Y
	DE-Cad::PAGFP	pGE-attB	*PAGFP* knock-in	0.6% (2/350)*	N	Y
	DE-Cad::Tomato	pGE-attB	*mTomato* knock-in	1.2% (1/81)	N	Y
	DE-Cad::mCherry	pGE-attB	*mCherry* knock-in	2.4% (2/84)	N	Y
	DE-Cad::GST	pGE-attB	*GST* knock-in	7.0% (3/43)	Y	N
	DE-Cad::His	pGE-attB	*His(× 6) tag* knock-in	16.7% (4/24)	Y	Y
	DE-Cad::PAGFPX2	pGE-attB	*PAGFP(× 2)* knock-in	10.2% (6/59)	Y	Y
	DE-Cad::PAGFPX3	pGE-attB	*PAGFP(× 3)* knock-in	8.6% (5/58)	Y	Y
	DE-Cad::PAGFPX4	pGE-attB	*PAGFP(× 4)* knock-in	3.3% (2/60)	Y	Y
	DE-Cad::FLAG	pGE-attB	*FLAG tag* knock-in	5.1% (4/78)	Y	Y
	DE-Cad::matFLAG	pGE-attB	*Flag tag* knock-in + nanos 3′UTR	8.1% (5/62)	Y	Y
	DE-Cad-ΔC::GST-HA-His	pGE-attB	intracellular domain replaced by *GST-HA-His*	1.4% (1/73)	Y	N
	DE-Cad::HA	pGE-attB	*HA tag* knock-in	0.5% (2/370)	Y	Y
	DE-Cad$^{S1457–60AAA}$::GFP	pGE-attB	point mutation, *GFP-tagged*	3.5% (2/57)	Y**	N
	DE-Cad$^{S1457–63AAAA}$::GFP	pGE-attB	point mutation, *GFP-tagged*	2.4% (2/84)	Y**	N
DE-Cad$^{GX23w[-]}$	DE-Cad^{GGG1377AAA}::GFP	pGE-attB	point mutation, *GFP-tagged*	2.3% (2/86)	Y**	Y
	DE-Cad^{ERD1380AAA}::GFP	pGE-attB	point mutation, *GFP-tagged*	1.9% (1/52)	Y**	Y
	DE-Cad^{TIINY1369FTNPVY}::GFP	pGE-attB	point mutation, *GFP-tagged*	2.5% (2/80)*	Y**	Y
	DE-Cad^{YKDP1391YKDL}::GFP	pGE-attB	point mutation, *GFP-tagged*	2.2% (2/92)	Y**	Y
	DE-Cad::GFP::ubi	pGE-attB	*GFP-ubiquitin* knock-in	2.3% (2/87)	Y**	N
	DE-Cad::GFP::ubiX2	pGE-attB	*GFP-ubiquitin(× 2)* knock-in	3.3% (3/90)	Y**	N
	DE-Cad::GFP::ubiX4	pGE-attB	*GFP-ubiquitin(× 4)* knock-in	1.3% (1/78)	Y**	N
	DE-Cad::αCat::GFP	pGE-attB	*GFP-tagged Ecad-αCat* fusion	1.0% (1/100)	Y**	Y
	DE-CadΔCyt::αCat::GFP	pGE-attB	*GFP-tagged αCat* fusion, minus *DE-Cad* intracellular domain	1.8% (2/109)	Y	ND
	DE-CadΔβ::αCat::GFP	pGE-attB	*GFP-tagged αCat* fusion, minus *β-Catenin* binding domain	2.0% (2/99)	Y	N
	DE-Cad^{YDLN1385YDLL}::GFP	pGE-attB	point mutation, *GFP-tagged*	3.3% (3/90)	Y	Y
	DE-Cad^{S1457A}::GFP	pGE-attB	point mutation, *GFP-tagged*	6.6% (4/61)	Y	Y
	DE-Cad^{S1457D}::GFP	pGE-attB	point mutation, *GFP-tagged*	5.7% (3/53)	Y	Y
	DE-Cad^{S1459A}::GFP	pGE-attB	point mutation, *GFP-tagged*	9.3% (7/75)	Y	Y
	DE-Cad^{S1459D}::GFP	pGE-attB	point mutation, *GFP-tagged*	5.3% (4/76)	Y	Y
	DE-Cad^{S1460AS1463A}::GFP	pGE-attB	point mutation, *GFP-tagged*	8.8% (8/91)	Y	Y
	DE-Cad^{S1460DS1463D}::GFP	pGE-attB	point mutation, *GFP-tagged*	6.7% (6/90)	Y	Y
	DE-Cad^{S1459AS1460AS1463A}::GFP	pGE-attB	point mutation, *GFP-tagged*	6.3% (6/96)	Y	N
	DE-Cad^{S1459DS1460DS1463D}::GFP	pGE-attB	point mutation, *GFP-tagged*	2.2% (2/90)	Y	Y
DE-Cad$^{GX6w[-]}$	φC31-attB-53	pGE-attB	test	2.4% (2/82)	N	N
	DE-Cad::GFP	pGE-attB	*GFP knock-in*	1.6% (1/64)	N	Y
	crb(rescue)	pGE-attB	rescue	0.2% (1/470)	N	Y
	crb^{FRT}	pGE-attB	*FRT* embeded in the last intron	1.2% (3/260)	N	Y
	crb(rescue)	pGE-attB^{GMR}	rescue	0.5% (1/196)	N	Y
	crb::GFP-A	pGE-attB^{GMR}	*GFP* knock-in	0.5% (1/190)	N	Y
	crb::GFP-B	pGE-attB^{GMR}	*GFP* knock-in	1.4% (5/357)	N	N
	crb::GFP-C	pGE-attB^{GMR}	*GFP* knock-in	1.3% (3/238)	N	Y
	crb-attB	pGE-attB^{GMR}	An extra attB site in *crb*	0.2% (1/487)	N	N
	crb::GST-A	pGE-attB^{GMR}	*GST* knock-in	6.3% (9/144)	Y	N
	crb::GST-C	pGE-attB^{GMR}	*GST* knock-in	5.1% (7/137)	Y	N
	crb::HA-A	pGE-attB^{GMR}	*HA tag* knock-in	4.2% (6/143)	Y	Y
	crb::HA-C	pGE-attB^{GMR}	*HA tag* knock-in	4.0% (8/198)	Y	Y
	crb::His-A	pGE-attB^{GMR}	*His(× 6) tag* knock-in	9.4% (9/96)	Y	Y
	crb::His-C	pGE-attB^{GMR}	*His(× 6) tag* knock-in	8.5% (11/129)	Y	Y
crb$^{GX24w[-]}$	crb::mCherry^{FRT}:: crb(intra)	pGE-attB^{GMR}	conditional allele	5.4% (5/92)	Y	Y
	hsFLP	pGE-attB^{GMR}	*hs-FLP* at the *crb* deletion locus	10.5% (10/95)	Y	N
	hsFLP-3′UTR(crb)	pGE-attB^{GMR}	*hs-FLP* at the *crb* deletion locus	11.5% (10/87)	Y	N
	crb-ΔC::GST-HA-His	pGE-attB^{GMR}	intracellular domain replaced by *GST-HA-His*	8.9% (8/90)	Y	N
	crb(NheI)	pGE-attB^{GMR}	Carrying one *NheI* site after the stop codon	3.1% (2/64)	Y	Y
	crb-SV40	pGE-attB^{GMR}	with *SV40* 3′UTR	8.5% (8/94)	Y	Y

Founder line	Construct	Vector	Allele description	Integration efficiency	*vasa-φC31*	Rescue
	crb^{delERLI}	pGE-attB^{GMR}	small deletion, *HA-tagged*	7.0% (7/100)	Y	N
	crb^{T6AT9A}	pGE-attB^{GMR}	point mutation, *HA-tagged*	5.9% (7/118)	Y	Y
	crb^{T6AT9AS11AS13A}	pGE-attB^{GMR}	point mutation, *HA-tagged*	5.8% (4/69)	Y	Y
	crb^{Y10A}	pGE-attB^{GMR}	point mutation, *HA-tagged*	3.5% (4/113)	Y	N
	crb^{E16A}	pGE-attB^{GMR}	point mutation, *HA-tagged*	5.2% (3/58)	Y	Y
	crb^{Y10AP12AE16A}	pGE-attB^{GMR}	point mutation, *HA-tagged*	3.6% (4/111)	Y	N
	crb::HA-SV40	pGE-attB^{GMR}	*HA-tagged*, with SV40 3′UTR	3.7% (4/108)	Y	Y
	*φ*C31-attB-53	pGE-attB	test	2.3% (5/218)	N	N
	lgl(rescue)-FRT	pGE-attB^{GMR}	rescue allele, also contains a FRT site at the 5′ region near *ATG*.	2.3% (2/87)	Y	Y
	lgl::GFP-C	pGE-attB^{GMR}	*GFP* knock-in	0.5% (1/196)	N	Y
	lgl::HA-C	pGE-attB^{GMR}	*HA tag* knock-in	0.9% (2/227)	Y	Y
lgl$^{GX7w[-]}$	lgl::His-C	pGE-attB^{GMR}	*His(× 6) tag* knock-in	0.6% (1/175)	Y	Y
	lgl::GST-N	pGE-attB^{GMR}	*GST* knock-in	0.7% (3/409)	Y	N
	lgl-del5′	pGE-attB^{GMR}	5′ repeat sequences deleted in *lgl* promoter	1.2% (3/244)	Y	Y
	lgl-delC::GST-HA-His	pGE-attB^{GMR}	C-terminal domain replaced by *GST-HA-His*	0.7% (2/293)	Y	N
	*φ*C31-attB-53	pGE-attB	test	5.8%(5/86)	N	N
	CG31158(rescue)	pGE-attB	rescue	0% (0/386)	N	n/a
CG31158$^{GX6w[-]}$	CG31158(rescue)	pGE-attB^{GMR}	rescue	2.1% (3/140)	N	Y
	CG31158::GFP	pGE-attB	*GFP* fusion at the C-terminus	0% (0/146)	N	n/a
	CG31158::GFP	pGE-attB^{GMR}	*GFP* fusion at the C-terminus	0.93% (2/216)	N	Y
	dArf6(rescue)	pGE-attB^{GMR}	rescue	5.6% (2/36)	Y	Y
dArf6$^{GX16w[-]}$	dArf6-C::GFP	pGE-attB^{GMR}	*GFP* fusion at the C-terminus	15.8% (3/19)	Y	Y
	dArf6-N::GFP	pGE-attB^{GMR}	*GFP* fusion at the N-terminus	7.7% (2/26)	Y	N
sdt$^{GX73w[-]}$	sdt(rescue)	pGE-attB^{GMR}	rescue	5.6% (2/36)	Y	Y

*In each of these integration experiments, we discovered a single nonspecific integration event (excluded from the table) based on the chromosomal mapping and PCR verifications.

**These injections were done in *DE-Cad$^{GX23w[-]}$/CyO* flies that had severely reduced presence of *vasa-φC31* in the population, hence the uniformly low integration efficiency.

Rescue: whether the allele complements the knock-out allele or previously characterized null allele.

αCat: α-Catenin. ND: not done. n/a: not applicable.

Table S3. Primers used for generation of founder knock-lines and rescue alleles

Gene	Primers	PCR product
DE-Cad		
5′ homologous arm PCR	HJ138: ATCTGA GCGGCCGC TAATTCGACGCCGGCTATAC	5,316 bp
	HJ139: TCAGTG GGTACC GCGTAAAAGAAAAGGCGTTG	
3′ homologous arm PCR	HJ140: ATCTGA GGCGCGCC CAACTGCGAACAAACCACAA	3,297 bp
	HJ141: TCAGTG AGGCCT GCGATGCCTCCATCAATAAT	
Rescue gDNA PCR	HJ246: TCAAGTGCGGCCGCATTTCCATGTTTGCGACCAG	4,783 bp
	HJ249: TGTCATGGTACCCTTGGCTACATTAGTGTTTGC	
crb		
5′ homologous arm PCR	HJ126: ATCTGA GCGGCCGC GAGGTGGCTCCAAAACAAAA	5,240 bp
	HJ127: TCAGTG GGTACC CCAAAGCGCAGAATCAATTT	
3′ homologous arm PCR	HJ128: TCAGTG ACTAGT GTATCTAAGCGTAAACTTAAGAGACTGTAC	3,025 bp
	HJ129: TCAGTG CCTAGG ACTACTATCGGTGTTAACCGGCAAAAAGAC	
Rescue gDNA PCR	HJ253: TCAAGTGCGGCCGCTGATCAGCGCAAACGAAAG	9,978 bp
	HJ257: TCAAGTACTAGTGGCGCGCCTCTAGAGCAAAATATGTTTTTTATTTGATTTAC	
lgl		
5′ homologous arm PCR	HJ130: ATCTGA GCGGCCGC GAGCACCATTTGGCTTGTTT	5,266 bp
	HJ131: TCAGTG GCTAGC TGTTGTGATTTGTGGGCAGT	
3′ homologous arm PCR	HJ132: TCAGTG CCTAGG CCTCGGTTTTGAGCCTTAGA	3,201 bp
	HJ133: TCAGTG CTCGAG AGAACGAATTTCACCGCAAC	
Rescue gDNA PCR	HJ308: GCATTGGAATTCCTTGAAAGCGGATTGGACAT	1,2057 bp
	HJ311: TAGAGCTCGGCGCGCCTATGATCTTTAAATAAGTCAAAATTAGGAGTTTTCAG	
CG31158		
5′ homologous arm PCR	WK71: CGAGATGCGGCCGCGGTGGAAATGGAGTGGATACTTGAT	5,296 bp
	WK72: CGAGATGGTACCAGAGAGTGAGTGGGTGCCACATATT	
3′ homologous arm PCR	WK25: CGAGATGGCGCGCCTTCCCAATATTTCATTACTGTTGTGT	2,866 bp
	WK26: CGAGATCTCGAGGGCAAGTTGTCTTAAATGAATTGTTA	
Rescue gDNA PCR	WK127: CGAGATGCGGCCGCTTCTCTCTGTCCCCTCTCTCTGAA	7,422 bp
	WK130: CGAGATGGTACCACGATAGGCGCGCCAGCAGCTCAACAAAAAATTAACAAC	
*dArf6**		
Rescue gDNA PCR	WK88: CGAGATGCGGCCGCGCCCTGAATCTCGCCCAGCTATTC	2,158 bp
	WK93: CGAGATGGTACCGTGACTCTACTAATTATTATATATTTTTATTATTATAATAC	
sdt		
5′ homologous arm PCR	WK161: CGAGATGCGGCCGCTTATGATTTTCAGTTGGCGTTTTAG	4,569 bp
	WK162: CGAGATGGTACCAAAAAAACCAATTAACAACACACGT	
3′ homologous arm PCR	WK165: CGAGATACTAGTTGATGGACTAATAATGGATTCTTGG	3,186 bp
	WK166: CGAGATCTCGAGCTAACATTTCACTATTTTCACGCTTG	
Rescue gDNA PCR	WK309: CGAGATGCGGCCGCATTTAATTAGCTCAAAGACTTTTGCATA	4,978 bp
	WK310: CGAGATACTAGTGATCGAAATATGATCTCGAGGACTC	

*PCR primers for the 5′ and 3′ homologous arms of *dArf6* targeting construct were in Huang et al [Huang J, Zhou W, Watson AM, Jan Y-N, Hong Y (2008) Efficient Ends-Out Gene Targeting In Drosophila. *Genetics* 180:703–707].

Rescue gDNA PCR: PCR for amplifying the genomic DNA fragment for making the *target$^{(rescue)}$ allele*.