MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes

Koen J T Venken[1], Karen L Schulze[1,2], Nele A Haelterman[1], Hongling Pan[1,2], Yuchun He[1,2], Martha Evans-Holm[3], Joseph W Carlson[3], Robert W Levis[4], Allan C Spradling[4], Roger A Hoskins[3] & Hugo J Bellen[1,2,5,6]

We demonstrate the versatility of a collection of insertions of the transposon *Minos*-mediated integration cassette (MiMIC), in *Drosophila melanogaster*. MiMIC contains a gene-trap cassette and the yellow⁺ marker flanked by two inverted bacteriophage ΦC31 integrase *attP* sites. MiMIC integrates almost at random in the genome to create sites for DNA manipulation. The *attP* sites allow the replacement of the intervening sequence of the transposon with any other sequence through recombinase-mediated cassette exchange (RMCE). We can revert insertions that function as gene traps and cause mutant phenotypes to revert to wild type by RMCE and modify insertions to control GAL4 or QF overexpression systems or perform lineage analysis using the Flp recombinase system. Insertions in coding introns can be exchanged with protein-tag cassettes to create fusion proteins to follow protein expression and perform biochemical experiments. The applications of MiMIC vastly extend the *D. melanogaster* toolkit.

Different types of transposons have been used to manipulate the *Drosophila* genome and to assess the function of genes, but each is designed for a specific purpose, and none are truly multifaceted. The most commonly used transposons are the *P* element, *piggyBac* and *Minos*[1–3]. *P* elements mobilize efficiently and often excise imprecisely, but they exhibit a strong insertional bias for the 5′ ends of genes[4,5]. *piggyBac* elements have much less bias[5] but mobilize less efficiently than *P* elements and only excise precisely[6]. *Minos* elements have very little insertional bias[5,7,8], transpose stably and efficiently in many organisms[9], and excise imprecisely at a useful frequency[6,8].

The most popular application of transposons is to create mutations directly by insertion or by imprecise excision[10]. Transposons have been engineered to allow controlled misexpression of genes via upstream activating sequence (UAS) sites in the transposon vector[4,11] or to promote activation of reporters such as GAL4 or β-galactosidase via nearby enhancers[12,13]. Transposons can also function as gene traps if they carry a splice acceptor site followed by stop codons in all three reading frames and a polyadenylation site so that intronic insertions can interrupt transcription and translation[14]. Transposons containing a protein trap harbor a splice acceptor site followed by a coding sequence tag and a splice donor site. When the protein trap is inserted in a coding intron in the appropriate orientation and reading frame, it reveals the protein's expression pattern[15,16]. Unfortunately, the frequency of *P*-element insertions in introns is low[4,5,17], and only one-sixth of insertions in introns have the appropriate orientation and reading frame to function as protein traps. Hence, only about 2.5% of *Drosophila* genes have been tagged with a protein trap, even when a *piggyBac* having a lesser insertional bias had been used[18,19]. Each different application of transposons requires the generation and maintenance of thousands of single-insertion stocks. The burden of stock keeping has limited the availability of these different tools: less than 5% of the transposon stocks that have been generated in the past 25 years are still available[5]. For the vast majority of *Drosophila* genes, only one type of transposon insertion is still available.

Transposons can be engineered to include target sequences recognized by recombinases or integrases[3,20,21] such as Flp recombinase[22] and ΦC31 integrase[23,24], respectively. These enzymes can replace sequences in transposons via recombinase-mediated cassette exchange (RMCE)[25,26]. RMCE has been demonstrated in *Drosophila* with both Flp recombinase and ΦC31 integrase[27,28]. ΦC31 integrase is the preferred enzyme because of its higher efficiency in transgenesis and unidirectional integration[23,24].

Here we describe a new mutagenesis and genome-manipulation system called *Minos*–mediated integration cassette (MiMIC). MiMIC is a *Minos* transposon that carries a dominant marker and a gene-trap cassette flanked by two inverted ΦC31 integrase *attP* sites. This transposon combines unbiased insertional mutagenesis with the ability to replace the gene-trap cassette by RMCE. Using MiMIC insertions, virtually limitless gene modification and genome engineering can be performed. We illustrate the utility of this system in gene- and protein-trap experiments, and reversion of lethal phenotypes.

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA. [2]Howard Hughes Medical Institute, Baylor College of Medicine, Houston, Texas, USA. [3]Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. [4]Department of Embryology, Howard Hughes Medical Institute, Carnegie Institution for Science, Baltimore, Maryland, USA. [5]Department of Neuroscience, Baylor College of Medicine, Houston, Texas, USA. [6]Program in Developmental Biology, Baylor College of Medicine, Houston, Texas, USA. Correspondence should be addressed to K.J.T.V. (kv134369@bcm.edu) or H.J.B. (hbellen@bcm.tmc.edu).

RECEIVED 26 APRIL; ACCEPTED 6 JULY; PUBLISHED ONLINE 31 JULY 2011; DOI:10.1038/NMETH.1662

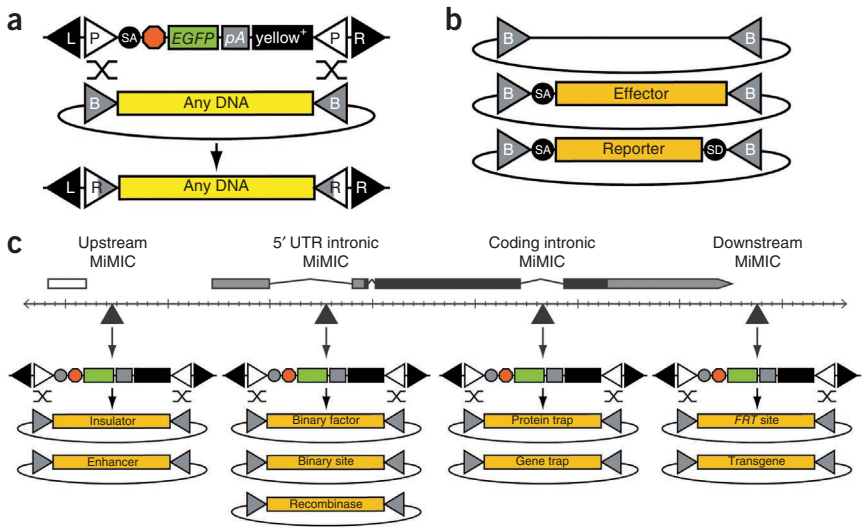

Figure 1 | The MiMIC transposon system. (**a**) MiMIC consists of two *Minos* inverted repeats (L and R), two inverted ΦC31 integrase *attP* sites (P), a gene-trap cassette consisting of a splice acceptor site (SA) followed by stop codons in all three reading frames and the *EGFP* coding sequence with a polyadenylation signal (*pA*), and the yellow+ marker. The sequence between the *attP* sites can be replaced via RMCE with a plasmid containing two inverted *attB* sites (B), resulting in two *attR* sites (R). (**b**) Three *attB* plasmids for RMCE: a correction plasmid consisting of a multiple cloning site, a gene-trap plasmid consisting of an SA fused to a downstream effector, and a protein-trap plasmid consisting of a reporter flanked by SA and splice donor site (SD). (**c**) Various MiMIC insertions in a hypothetical gene with a regulatory element (white), 5′ and 3′ untranslated regions (gray), and coding regions (black) that can be used for several applications as indicated.

RESULTS

The MiMIC transposon

We engineered a transposon vector, MiMIC. Between the *Minos* 255 nucleotide (nt) inverted repeats (**Fig. 1a**), we included the yellow+ dominant body-color marker for identifying insertions, flanked by two inverted ΦC31 integrase *attP* sites. We also included a mutagenic gene-trap cassette consisting of a splice acceptor site followed by stop codons in all three reading frames, the coding sequence of enhanced GFP (EGFP), and an SV40 polyadenylation signal sequence. Sequences between *attP* sites are replaceable through RMCE *in vivo* with any DNA cassette flanked by two inverted ΦC31 integrase *attB* sites[28]. This replaces the yellow+ marker, so RMCE events can be identified by loss of body pigmentation. To test MiMIC, we engineered three replacement cassettes: a neutral correction plasmid for removing the mutagenic gene trap in MiMIC, a gene-trap plasmid for introducing protein-coding sequences under control of a host gene promoter and a protein-trap plasmid for incorporating reporter tags into the coding sequence of a host gene (**Fig. 1b** and **Supplementary Table 1**).

Because donor cassettes can contain any DNA fragment, MiMIC provides enormous flexibility (**Fig. 1c**). MiMIC insertions near the 5′ and 3′ ends of genes allow the integration of regulatory elements such as enhancers or insulators to direct or restrict expression, respectively. Such insertions can also be used to integrate an *FRT* site for creating Flp recombinase–based chromosomal rearrangements[3]. Insertions in 5′ UTR introns allow the incorporation of binary expression components, such as *GAL4*-UAS[29] and *QF*–QF upstream activating sequence (QUAS)[30] and genes encoding recombinases such as Flp[22]. Insertions in coding introns allow integration of protein tags, including an ever-expanding repertoire of fluorescent markers, conditional protein destruction tags and other gene-trap mutator cassettes. Finally, any insertion can be used as a generic docking site for integrating transgenes.

A MiMIC insertion screen

We created 4,464 single-insertion MiMIC lines, determined unique insertion sites for 3,633 insertions (81.3%) by inverse PCR and associated the mapped insertions with features of annotated genes (Online Methods). Of all mapped insertions, 2,293 (63%) mapped within 1,541 annotated genes, and 72% of these intragenic insertions mapped within introns, including 5′ UTR introns and coding introns, both of which are valuable targets for RMCE-based gene manipulation (**Supplementary Table 2**). We deposited 1,269 selected insertion lines in the Bloomington *Drosophila* Stock Center (BDSC) as part of the *Drosophila* Gene Disruption Project (GDP) collection[5]. We will regularly select additional MiMIC lines for the GDP collection and aim to deposit over 6,000 lines during the next 4 years. The GDP maintains an online database (http://flypush.imgen.bcm.tmc.edu/pscreen/) of lines that are available from the BDSC as well as lines that we are still balancing, which may be obtained directly from the GDP.

MiMIC insertion mutants can be reverted by RMCE

The MiMIC transposon contains a gene-trap cassette. Hence, insertions in coding introns should truncate transcripts if MiMIC is inserted in the proper orientation. We selected four MiMIC transposons inserted in the proper orientation to be mutator gene traps (cassettes were named as *Mi{MIC}* followed by the gene name and insertion strain number): *Mi{MIC}Rfx*^{MI00053}, *Mi{MIC}tutl*^{MI00290}, *Mi{MIC}comm*^{MI00380} and *Mi{MIC}wnd*^{MI00494} inserted in *Rfx*, *tutl*, *comm* and *wnd*, respectively. All four alleles were associated with a lethal phenotype. In three cases, the insertion did not complement previously reported mutations of these genes (**Supplementary Table 3**), indicating that the lethality was indeed associated with the insertion. *Mi{MIC}wnd*^{MI00494} was the exception, but complementation data indicate that *Mi{MIC}wnd*^{MI00494} and the previously reported alleles[31] all contained second-site mutations responsible for the lethality, and that none of the transheterozygous *wnd* allelic combinations caused lethality. For *Mi{MIC}Rfx*^{MI00053}, complementation data indicate the presence of uncoordinated escapers for all allelic combinations, a phenotype that has been previously described for *Rfx* loss-of-function mutations[32].

We then removed the gene-trap cassette from *Mi{MIC}Rfx*^{MI00053}, *Mi{MIC}tutl*^{MI00290} and *Mi{MIC}comm*^{MI00380} by RMCE with a correction plasmid (**Fig. 1b**), screened for loss of yellow+ (**Supplementary Fig. 1**) and established that the lethality reverted. Hence, intronic MiMIC insertions are mutagenic, and the mutation can be reverted through microinjection of a plasmid. This demonstrates

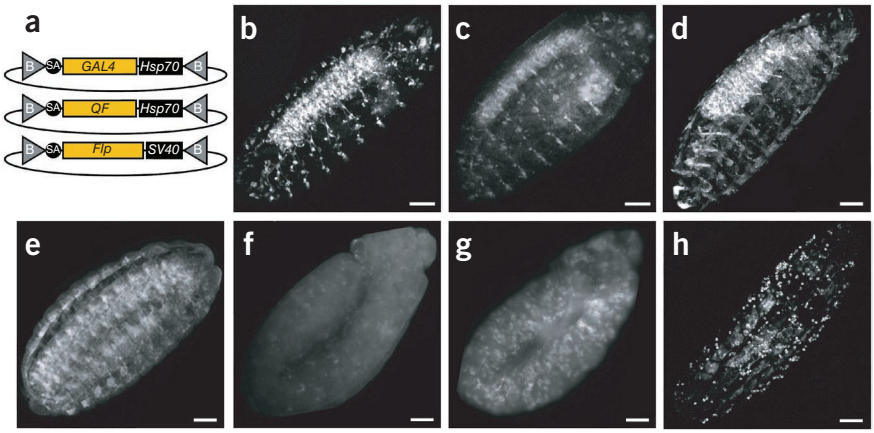

Figure 2 | Binary expression and lineage analysis with MiMIC insertions. (**a**) Gene-trap cassettes that incorporate genes encoding GAL4 or QF trans-activators for binary activation, and the gene encoding Flp recombinase for fate mapping. (**b**–**e**) Live imaging (**b**,**d**) and confocal microscopy analysis using an antibody to mCherry (**c**,**e**) of the expression domain revealed by *GAL4* inserted in *gogo* (**b**,**c**) or *caps* (**d**,**e**) loci. (**f**,**g**) The expression of *MYPT-75D* revealed by *GAL4* (**f**) or *QF* (**g**) integrated in *MYPT-75D*. (**h**) Live imaging of the GAL4 expression pattern revealed by *GAL4* inserted in *BM-40-SPARC*. Scale bars, 50 μm.

that MiMIC is the cause of the lethal phenotypes in these insertion alleles.

Binary expression and lineage analysis with MiMIC

A substantial portion (20.4%) of intragenic MiMICs were localized to 5′ UTR introns (**Supplementary Table 2**). Introduction of exogenous protein-coding sequences into these insertions allows expression under control of the endogenous gene regulatory elements. Hence, we constructed three gene-trap plasmids that encode GAL4, QF and Flp (**Fig. 2a**). We selected five 5′ UTR intronic insertions: *Mi{MIC}gogo*^{MI00065}, *Mi{MIC}Tl*^{MI00181}, *Mi{MIC}caps*^{MI00249}, *Mi{MIC}MYPT-75D*^{MI00314} and *Mi{MIC}BM-40-SPARC*^{MI00329}, inserted in *gogo*, *Tl*, *caps*, *MYPT-75D* and *BM-40-SPARC*, respectively. We used RMCE to incorporate each of the three gene-trap cassettes into these insertions.

We tested the *GAL4* insertions using a *10×UAS-mCherry* cytoplasmic reporter, which comprises 10 copies of the UAS fused to *mCherry* (Online Methods), the *QF* insertions with a *5×QUAS-mtdTomato-3×HA* membrane reporter containing three hemagglutinin (HA) tags[30], and the *Flp* insertions with an *act>y+>GAL4;UAS-GFP* (> indicates *FRT* site) cytoplasmic Flp-out detector (Online Methods). *GAL4* incorporated into the *gogo* insertion revealed expression in the embryonic peripheral and central nervous system, in agreement with RNA *in situ* hybridization data (**Fig. 2b**,**c**) (T. Suzuki, personal communication). GAL4 inserted into *caps* recapitulated the known expression pattern[33] (**Fig. 2d**,**e**). The unknown expression pattern of *MYPT-75D* was revealed by *GAL4*, *QF* and *Flp* integrated in *Mi{MIC}MYPT-75D*^{MI00314}. *GAL4* analysis revealed many scattered cells labeled during germ-band extension (**Fig. 2f**), some of which appeared to be muscle precursors. *QF* analysis also revealed this expression pattern (**Fig. 2g**) but resulted in stronger labeling owing to the membrane marker driven by *QUAS* elements instead of the cytoplasmic marker driven by *UAS* elements. Flp-out analysis revealed a much smaller subset of labeled cells, suggesting inefficient Flp-out (data not shown). Finally, *GAL4* integrated in *BM40-SPARC* revealed an expression pattern very similar to that revealed by antibodies to the endogenous protein, including expression in hemocytes and fat body (**Fig. 2h**)[34,35].

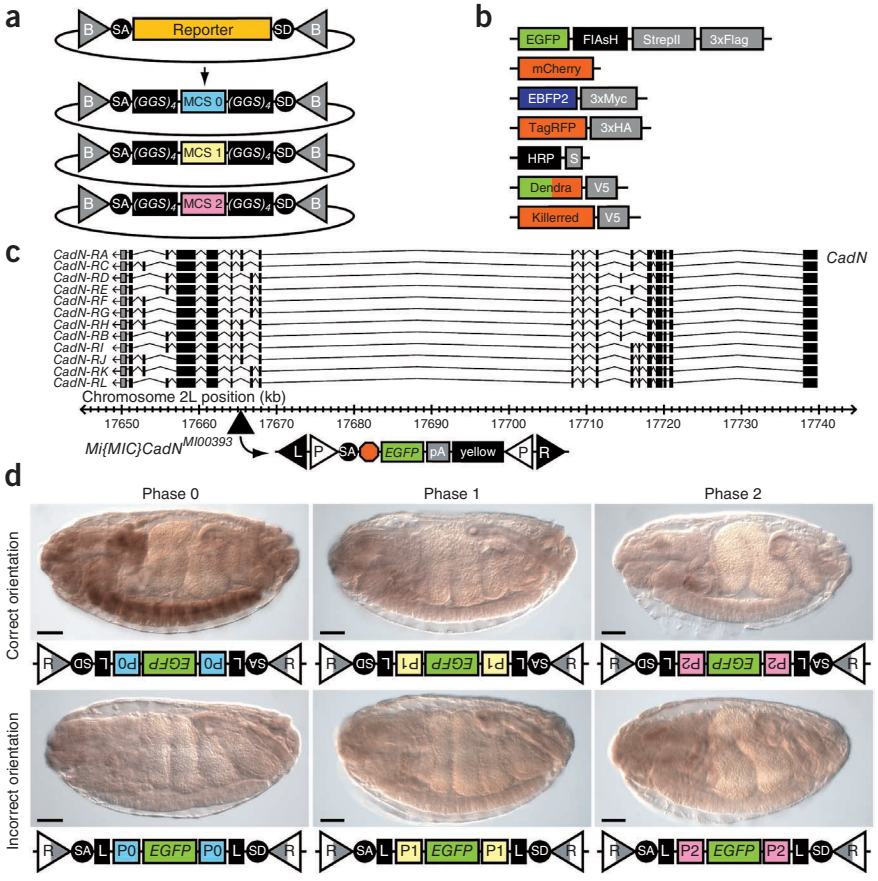

Figure 3 | Protein trapping with MiMIC insertions. (**a**) For each protein-trap cassette, three versions were constructed corresponding to the three intron phases (0, 1 and 2). *(GGS)*$_4$, flexible peptide linker sequence encoding a Gly-Gly-Ser quadruplet tandem repeat. (**b**) Tag and multitag reporters. (**c**) A 100-kb genomic region containing *CadN* with all splice isoforms (*CadN-RA to CadN-RL*) is shown. The location of the *Mi{MIC}CadN*^{MI00393} insertion in a phase 0 coding intron is indicated. (**d**) Integration of a phase-0 *EGFP-FlAsH-StrepII-3xFlag* cassette (*EGFP*) in the indicated orientation and intron phase detected by an antibody to EGFP. L, *(GGS)*$_4$ linker; R, *attR*; P0, splice phase 0; P1, splice phase 1; and P2, splice phase 2. Scale bars, 50 μm.

Table 1 | Protein-trapping experiments

Insertion					RMCE experiment				
Gene	MiMIC line	Gene-trap status	Lethality	Intron phase	Tag[a]	Total lines	Expressing lines	Expressing lines (%)	Lethality[b]
Rfx	MI00053	Yes	Lethal	1	EGFP	1	0	0	NA
					mCherry	2	1	50	L
					EBFP	2	1	50	V
					3xHA	2	1	50	V
					S peptide	5	2	40	2L
					Dendra	6	5	83	5V
					V5	3	3	100	2V/L
tutl	MI00290	Yes	Lethal	1	EGFP	4	3	75	3V
					mCherry	6	2	33	2V
					EBFP	3	2	67	2V
					3xHA	1	0	0	NA
					S peptide	5	3	60	3V
					Dendra	5	2	40	2V
					V5	5	2	40	2L
rhea	MI00296	No	Viable	0	EGFP	5	3	60	V/2L
					mCherry	4	2	50	V/L
					EBFP	6	3	50	2V/L
					3xHA	5	2	40	2V
					S peptide	5	0	0	NA
					Dendra	6	4	67	3V/L
					V5	6	4	67	4V
comm	MI00380	Yes	Lethal	1	EGFP	7	2	29	V/L
					mCherry	3	2	67	V/L
					EBFP	5	2	40	2L
					3xHA	2	1	50	L
					S peptide	5	1	20	L
					Dendra	5	1	20	L
					V5	5	4	80	4L
CadN	MI00393	No	Viable	0	EGFP	5	3	60	3V
					mCherry	2	1	50	V
					EBFP	4	2	50	V/L
					3xHA	4	1	25	V
					S peptide	3	1	33	V
					Dendra	4	2	50	2V
					V5	3	1	33	V
wnd	MI00494	Yes	Lethal	2	EGFP	5	5	100	5L
					mCherry	3	1	33	L
					EBFP	3	1	33	L
					3xHA	3	0	0	NA
					S peptide	4	3	75	3L
					Dendra	3	1	33	V
					V5	1	0	0	NA
Total						166	80	48	

MiMIC insertions in six genes were tagged with different protein-trap cassettes using RMCE.
[a]Tag components of multitags used for expression analysis were EGFP (EGFP-FlAsH-StrepII-3xFlag), mCherry (mCherry), EBFP (EBFP-3xMyc), 3xHA (TagRFP-3xHA), S peptide (HRP-S peptide), Dendra (Dendra-V5) and V5 (Killerred-V5). [b]V, viable (>5% viable homozygotes); L, lethal; and NA, not applicable.

We performed PCR analysis to confirm the molecular nature of RMCE events. In each case, PCR demonstrated that a productive binary activation or recombination activity occured only when the cassette was integrated in the appropriate orientation for expression of the reporter (**Supplementary Fig. 2**). As RMCE can occur in either orientation, we expected a 50% chance of a productive exchange. However, only 25% of the gene-trap RMCE events were in the productive orientation for expression (**Supplementary Table 4**). Although we do not understand the cause, this suggests selection against productive reporter expression.

Protein trapping with MiMIC insertions

To determine the expression pattern of the protein product of a gene, including its subcellular localization, one can analyze it after tagging the protein with an epitope to which antibodies are available or by live imaging. More than 51% of intragenic MiMIC insertions are in coding introns and permit protein trapping (**Supplementary Table 2**). We constructed protein-trap cassette plasmids with splice acceptor and splice donor sites flanking synthetic exons encoding protein tags in three versions (**Fig. 3a**,**b** and **Supplementary Fig. 3**) so as to convert any MiMIC insertion in a coding intron, regardless of its orientation or splicing phase (0, 1 or 2), into a protein trap. We flanked the sequence encoding a protein tag on both sides with a linker sequence encoding a quadruple Gly-Gly-Ser repeat to increase flexibility between the tag and the host protein. We engineered seven multi-tag cassettes for different applications. Most consisted of a fluorescent tag and a peptide tag so that if *in vivo*

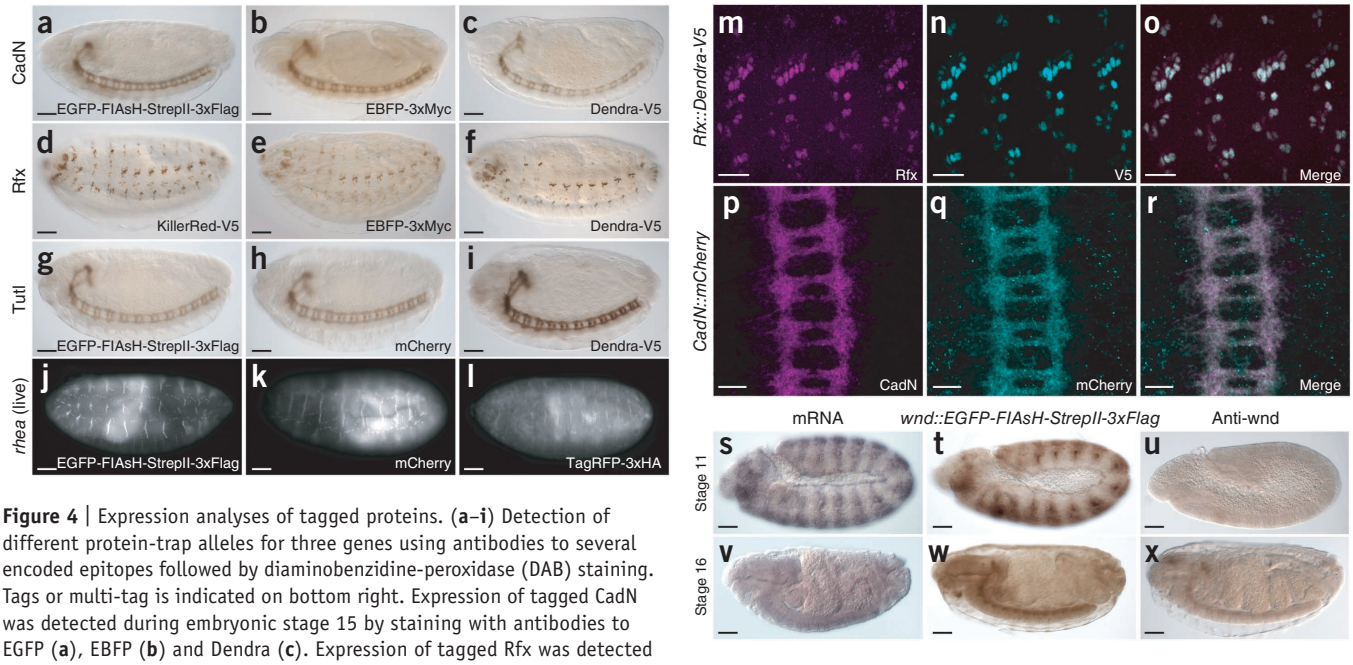

Figure 4 | Expression analyses of tagged proteins. (**a**–**i**) Detection of different protein-trap alleles for three genes using antibodies to several encoded epitopes followed by diaminobenzidine-peroxidase (DAB) staining. Tags or multi-tag is indicated on bottom right. Expression of tagged CadN was detected during embryonic stage 15 by staining with antibodies to EGFP (**a**), EBFP (**b**) and Dendra (**c**). Expression of tagged Rfx was detected in stage-15 embryos with antibodies to V5 (**d**), EBFP (**e**) and Dendra (**f**). Expression of tagged Tutl in the ventral nerve cord of stage-15 embryos was detected with antibodies to EGFP (**g**), mCherry (**h**) and Dendra (**i**). (**j**–**l**) Fluorescence detection of different protein-trap alleles in live, stage-17 embryos expressing tagged Rhea: EGFP (**j**) mCherry (**k**) and TagRFP (**l**) signals are shown. (**m**–**r**) Combined localization of protein traps and endogenous proteins in stage-15 embyos for Rfx (**m**–**o**) and CadN (**p**–**r**). Shown is staining with antibodies to Rfx (**m**), V5 (**n**) and a merged image (**o**) in an *Rfx::Dendra-V5* trap line. Also shown is staining with antibodies to CadN (**p**), mCherry (**q**) and a merged image (**r**) in a *CadN::mCherry* trap line. (**s**–**x**) Expression detected using protein traps: *wnd* mRNA detected using mRNA *in situ* hybridization (**s**,**v**); Wnd detected by staining an *EGFP-FlAsH-StrepII-3xFlag* trap line with an antibody to EGFP (**t**,**w**) and staining with an antibody to Wnd (**u**,**x**) during embryonic stages 11 (**s**–**u**) and 16 (**v**–**x**). Scale bars, 50 μm (**a**–**l**,**s**–**x**) and 20 μm (**m**–**r**).

fluorescence imaging was not possible because of low expression, the tag could still be detected using antibodies.

In a first protein-trapping test, we introduced the EGFP-FlAsH-StrepII-3×Flag multi-tag in all three phases into *Mi{MIC}CadN*^{MI00393}, which we inserted into a phase 0 intron of *CadN* (**Fig. 3c**) and used PCR to identify integration events for each orientation and phase. As expected, only the phase 0 cassette integrated in the correct orientation recapitulated the expected expression pattern, and none of the other five classes of events resulted in detectable expression (**Fig. 3d**).

Protein expression analysis with multi-tag cassettes

Next, we evaluated expression patterns using seven different tags in six different genes in which MiMIC inserted in a coding intron: *Mi{MIC}Rfx*^{MI00053}, *Mi{MIC}tutl*^{MI00290}, *Mi{MIC}rhea*^{MI00296}, *Mi{MIC}comm*^{MI00380}, *Mi{MIC}CadN*^{MI00393} and *MI{MIC}wnd*^{MI00494} inserted in *Rfx* (phase 1), *tutl* (phase 1), *rhea* (also known as *talin*) (phase 0), *comm* (phase 1), *CadN* (phase 0) and *wnd* (phase 2), respectively. We introduced the seven different tag cassettes with the proper intron phase into each of the six insertions (**Fig. 3a**,**b** and **Table 1**). Then we used PCR to determine the orientation of each RMCE event and established that 48% of the integration events were in the desired orientation. This was in agreement with the 50% frequency expected by chance, suggesting that these events are not detrimental to host gene function. The lethality associated with each of the original gene-trap insertions often reverted upon RMCE with protein-trap tags. For *Rfx* and *tutl*, the lethality of the MiMIC gene trap reverted in 69% and 86% of the protein-trap lines, respectively (**Table 1**). This demonstrates that in most cases protein function was at least partially restored when the gene trap was removed and replaced by a protein trap. The reverted lines may be partial loss-of-function mutations or full revertants. Note that reversion of lethality did not occur in most RMCE events for the insertions in *comm* (15%) or *wnd* (9%).

Failure to revert the lethality of a MiMIC insertion by RMCE with a protein-trap cassette may result from the effect of the tags on protein function. Alternatively, some of the MiMIC-bearing chromosomes may contain second-site lethal mutations as previously observed in *P*-element stocks exposed to transposase[13], or mutations could be induced during the RMCE procedure because ΦC31 integrase has been shown to induce DNA damage and chromosome rearrangements[36–38]. Both issues can be obviated by removing the second-site mutations by recombination. To test these possibilities, we crossed protein-trap alleles generated by RMCE that did not revert the lethality of the gene-traps in *comm* (6 lines), *wnd* (4 lines), *rhea* (6 lines) and *CadN* (4 lines) to previously described alleles and deficiencies uncovering the corresponding loci. All of the protein-trap alleles of all the genes tested complemented the established lethal allele or deletion chromosome, providing evidence that the tagged fusion proteins indeed supply sufficient gene function to revert the lethality.

To determine the expression pattern and subcellular localization of the tagged proteins generated by RMCE, we first stained a large sample of tagged proteins using antibodies to GFP, V5 epitope, monomeric (m)Cherry and Dendra. We compared the expression of the same protein fused to different tags. Detection of *CadN* with *EGFP-FlAsH-StrepII-3xFlag*, *EBFP2-3xMyc* or *Dendra-V5* multitags (**Fig. 3b**) shows very similar expression patterns (**Fig. 4a**–**c**).

Similarly, different tag-coding genes integrated into *Rfx* and *tutl* showed highly reproducible expression patterns (**Fig. 4d–i**). In the case of *rhea*, protein trapping allowed live imaging of three different fluorescent tags (**Fig. 4j–l**). The observed expression patterns faithfully recapitulated the previously described expression patterns of *CadN*[39], *Rfx*[40], *tutl*[41] and *rhea*[42]. To determine in more detail whether the fusion protein expression patterns faithfully report the cellular and subcellular localization of the endogenous proteins, we performed simultaneous labeling experiments with antibodies to different tags and the endogenous protein for Rfx and CadN (**Fig. 4m–r**). These experiments showed fully overlapping expression patterns in trans-heterozygotes that expressed both the tagged and the untagged proteins.

In total, we tested 166 independent tagged fusion proteins generated by RMCE for six MiMIC insertions and observed that less than 3% (5/166) exhibited a different expression pattern than we anticipated. The different expression patterns were not associated with any particular gene or tag, suggesting that they were due to a faulty RMCE event, and we confirmed this by an aberrant PCR pattern. These data indicate that RMCE-based protein-trapping results in the precise incorporation of tags and will permit the determination of the expression pattern and subcellular distribution of many uncharacterized proteins.

Detection of new expression patterns

When analyzing the expression patterns of the different tag sequences integrated into *Mi{MIC}wnd*^{MI00494}, we observed an expression pattern that was much broader and more complex than anticipated (**Fig. 4s–x** and **Supplementary Fig. 4**). Antibody staining[31] showed weak expression of Wnd in the embryonic nervous system in stage 16 embryos but not in other tissues (**Fig. 4u,x**). However, the Wnd fusion protein generated by RMCE had a very complex and dynamic expression pattern (**Fig. 4s,v**) and agreed with the pattern revealed by RNA *in situ* hybridization (**Fig. 4t,w**). These and many other immunohistochemical staining experiments on the 166 tagged proteins (data not shown) revealed that well-characterized antibodies to tags that are integrated in fusion proteins were often superior to custom antibodies to the endogenous protein.

DISCUSSION

MiMIC-mediated insertions in 5′ UTR introns allow the expression of transcription factors such as GAL4 and QF, and recombinases such as Flp to generate gene-specific binary expression and recombination systems, respectively. Moreover, these insertions allow any current or future effector to be placed under the control of the endogenous gene's regulatory elements. Genes that are tagged by insertions in coding introns can be manipulated in many ways. One can determine gene expression and subcellular protein distribution using light microscopy and likely immunoelectron microscopy[43]. Tags inserted in transcription factors can be used for chromatin immunoprecipitation[44]. Other applications such as combined immunoprecipitation followed by mass spectrometry and identification of RNA binding partners can also be performed. A complementary *in vivo* swapping approach has been developed recently for enhancer trapping[45].

The widespread adoption of the MiMIC system by the *Drosophila* research community will depend on substantially increasing the number of insertion lines that are available for public distribution. Currently, 1,269 unique MiMIC insertions are available from the BDSC, and we are balancing and validating by resequencing about 900 more. In the GDP, we plan to generate ~6,000 additional insertions during the next four years. As *Minos* integrates almost at random in the genome[5,8] and about 33% of MiMIC insertions are in coding introns (**Supplementary Table 2**), the manipulations documented here will become feasible for many more *Drosophila* genes. The ability to assess gene expression patterns and subcellular protein distributions with high resolution will vastly expand the number and quality of expression patterns of *Drosophila* genes.

Finally, both *attP* sites present in MiMIC can be used as docking sites for integration of gene targeting constructs[46] that can be used to engineer the genome in the vicinity of the transposon insertion. A collection of ~6,000 insertions spaced about 40 kilobases (kb) apart[5] should allow the manipulation of most genes by engineering and integrating large genomic constructs using the P[acman] system[47,48], recombineering methods[3] and RMCE[28].

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Accession codes. GenBank: GU370067 (MiMIC vector), JN222909, (pBS-KS-attB1-2), JN222910 (pBS-KS-attB1-2-GT-SA), JN222911 (pBS-KS-attB1-2-PT-SA-SD-0), JN222912 (pBS-KS-attB1-2-PT-SA-SD-1) and JN222913 (pBS-KS-attB1-2-PT-SA-SD-2).

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank B. Al-Anzi (California Institute of Technology), K. Basler, J. Bischof (University of Zurich), J. Bateman (Bowdoin College), K. Broadie (Vanderbilt University), M. Calos, L. Luo, A. Okada (Stanford University), W. Chia (National University of Singapore), A. DiAntonio (Washington University), B. Durand, A. Laurençon (University of Lyon), F. Karch (University of Geneva), X. Morin (Institute of Developmental Biology of Marseille), A. Nose (University of Tokyo), S. Oehler (University of Crete), A. Pavlopoulos (University of Cambridge), C. Potter (Johns Hopkins University), Y. Rao (McGill University), M. Ringuette, J. Shahab (University of Toronto), C. Savakis (Biomedical Sciences Research Center Alexander Fleming), T. Suzuki (Max Planck Institute of Neurobiology), C. Tan (University of Missouri), G. Tear (King's College London), R. Tsien (University of California San Diego), T. Wu (Harvard University), L. Zipursky (University of California Los Angeles), members of the BDSC and the *Drosophila* Genomics Resource Center (Indiana University), Addgene and members of the Developmental Studies Hybridoma Bank for flies, plasmids, antibodies and communications; S. Park and K. Wan for assistance in mapping MiMIC insertions; D. Bei, Y. Fang, J. Li, Z. Wang, X. Zheng and J. Yue for generating fly stocks; and T. Suzuki for communication of unpublished results. This work was funded by US National Institutes of Health grants 2R01 GM067858 to A.C.S., R.A.H. and H.J.B., and T32 GM07526-33 to K.J.T.V.; A.C.S. and H.J.B. are funded by the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

K.J.T.V. designed the MiMIC technique and vectors, and performed all molecular biology, except for mapping of insertions. R.W.L., A.C.S., R.A.H. and H.J.B. conceived the application of MiMIC to the GDP. H.P. and Y.H. performed microinjections. K.J.T.V., H.P. and Y.H. performed fly genetics. M.E.-H. and R.A.H. mapped insertions. K.J.T.V., Y.H., M.E.-H., J.W.C., R.W.L. and R.A.H. analyzed insertion data, annotated insertions and prepared public database submissions. J.W.C. performed bioinformatic analysis. K.J.T.V., N.A.H. and H.P. verified RMCE events by PCR. K.J.T.V. and K.L.S. did staining of gene-trap events. K.L.S. and N.A.H. did staining of protein trap events. K.J.T.V., K.L.S., N.A.H. and H.J.B. analyzed expression patterns. K.J.T.V. and H.J.B. wrote the paper. R.A.H. and R.W.L. edited the paper.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Ryder, E. & Russell, S. Transposable elements as tools for genomics and genetics in *Drosophila*. *Brief. Funct. Genomics Proteomics* **2**, 57–71 (2003).
2. Venken, K.J. & Bellen, H.J. Emerging technologies for gene manipulation in *Drosophila melanogaster*. *Nat. Rev. Genet.* **6**, 167–178 (2005).
3. Venken, K.J. & Bellen, H.J. Transgenesis upgrades for *Drosophila melanogaster*. *Development* **134**, 3571–3584 (2007).
4. Bellen, H.J. *et al.* The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. *Genetics* **167**, 761–781 (2004).
5. Bellen, H.J. *et al.* The *Drosophila* gene disruption project: progress using transposons with distinctive site specificities. *Genetics* **188**, 731–743 (2011).
6. Witsell, A., Kane, D.P., Rubin, S. & McVey, M. Removal of the bloom syndrome DNA helicase extends the utility of imprecise transposon excision for making null mutations in *Drosophila*. *Genetics* **183**, 1187–1193 (2009).
7. Franz, G. & Savakis, C. Minos, a new transposable element from *Drosophila hydei*, is a member of the Tc1-like family of transposons. *Nucleic Acids Res.* **19**, 6646 (1991).
8. Metaxakis, A., Oehler, S., Klinakis, A. & Savakis, C. Minos as a genetic and genomic tool in *Drosophila melanogaster*. *Genetics* **171**, 571–581 (2005).
9. Pavlopoulos, A., Oehler, S., Kapetanaki, M.G. & Savakis, C. The DNA transposon Minos as a tool for transgenesis and functional genomic analysis in vertebrates and invertebrates. *Genome Biol.* **8** (suppl. 1), S2 (2007).
10. Spradling, A.C. *et al.* The Berkeley *Drosophila* Genome Project gene disruption project: single P-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* **153**, 135–177 (1999).
11. Rorth, P. *et al.* Systematic gain-of-function genetics in *Drosophila*. *Development* **125**, 1049–1057 (1998).
12. Bier, E. *et al.* Searching for pattern and mutation in the *Drosophila* genome with a P-lacZ vector. *Genes Dev.* **3**, 1273–1287 (1989).
13. Bellen, H.J. *et al.* P-element-mediated enhancer detection: a versatile method to study development in *Drosophila*. *Genes Dev.* **3**, 1288–1300 (1989).
14. Lukacsovich, T. *et al.* Dual-tagging gene trap of novel genes in *Drosophila melanogaster*. *Genetics* **157**, 727–742 (2001).
15. Morin, X., Daneman, R., Zavortink, M. & Chia, W. A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **98**, 15050–15055 (2001).
16. Clyne, P.J., Brotman, J.S., Sweeney, S.T. & Davis, G. Green fluorescent protein tagging *Drosophila* proteins at their native genomic loci with small *P* elements. *Genetics* **165**, 1433–1441 (2003).
17. Aleksic, J., Lazic, R., Muller, I., Russell, S.R. & Adryan, B. Biases in *Drosophila melanogaster* protein trap screens. *BMC Genomics* **10**, 249 (2009).
18. Quinones-Coello, A.T. *et al.* Exploring strategies for protein trapping in *Drosophila*. *Genetics* **175**, 1089–1104 (2007).
19. Buszczak, M. *et al.* The carnegie protein trap library: a versatile tool for *Drosophila* developmental studies. *Genetics* **175**, 1505–1531 (2007).
20. Branda, C.S. & Dymecki, S.M. Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev. Cell* **6**, 7–28 (2004).
21. Wirth, D. *et al.* Road to precision: recombinase-based targeting technologies for genome engineering. *Curr. Opin. Biotechnol.* **18**, 411–419 (2007).
22. Golic, K.G. & Lindquist, S. The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. *Cell* **59**, 499–509 (1989).
23. Groth, A.C., Fish, M., Nusse, R. & Calos, M.P. Construction of transgenic *Drosophila* by using the site-specific integrase from phage ϕC31. *Genetics* **166**, 1775–1782 (2004).
24. Bischof, J., Maeda, R.K., Hediger, M., Karch, F. & Basler, K. An optimized transgenesis system for *Drosophila* using germ-line-specific ϕC31 integrases. *Proc. Natl. Acad. Sci. USA* **104**, 3312–3317 (2007).
25. Schlake, T. & Bode, J. Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry* **33**, 12746–12751 (1994).
26. Baer, A. & Bode, J. Coping with kinetic and thermodynamic barriers: RMCE, an efficient strategy for the targeted integration of transgenes. *Curr. Opin. Biotechnol.* **12**, 473–480 (2001).
27. Horn, C. & Handler, A.M. Site-specific genomic targeting in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **102**, 12483–12488 (2005).
28. Bateman, J.R., Lee, A.M. & Wu, C.T. Site-specific transformation of *Drosophila* via ϕC31 integrase-nediated cassette exchange. *Genetics* **173**, 769–777 (2006).
29. Brand, A.H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
30. Potter, C.J., Tasic, B., Russler, E.V., Liang, L. & Luo, L. The Q system: a repressible binary system for transgene expression, lineage tracing, and mosaic analysis. *Cell* **141**, 536–548 (2010).
31. Collins, C.A., Wairkar, Y.P., Johnson, S.L. & Diantonio, A. Highwire restrains synaptic growth by attenuating a MAP kinase signal. *Neuron* **51**, 57–69 (2006).
32. Dubruille, R. *et al.* *Drosophila* regulatory factor X is necessary for ciliated sensory neuron differentiation. *Development* **129**, 5487–5498 (2002).
33. Shishido, E., Takeichi, M. & Nose, A. *Drosophila* synapse formation: regulation by transmembrane protein with Leu-rich repeats, CAPRICIOUS. *Science* **280**, 2118–2121 (1998).
34. Martinek, N., Zou, R., Berg, M., Sodek, J. & Ringuette, M. Evolutionary conservation and association of SPARC with the basal lamina in *Drosophila*. *Dev. Genes Evol.* **212**, 124–133 (2002).
35. Martinek, N., Shahab, J., Saathoff, M. & Ringuette, M. Haemocyte-derived SPARC is required for collagen-IV-dependent stability of basal laminae in *Drosophila* embryos. *J. Cell Sci.* **121**, 1671–1680 (2008).
36. Ehrhardt, A., Engler, J.A., Xu, H., Cherry, A.M. & Kay, M.A. Molecular analysis of chromosomal rearrangements in mammalian cells after ϕC31-mediated integration. *Hum. Gene Ther.* **17**, 1077–1094 (2006).
37. Liu, J., Jeppesen, I., Nielsen, K. & Jensen, T.G. Phic31 integrase induces chromosomal aberrations in primary human fibroblasts. *Gene Ther.* **13**, 1188–1190 (2006).
38. Liu, J., Skjorringe, T., Gjetting, T. & Jensen, T.G. PhiC31 integrase induces a DNA damage response and chromosomal rearrangements in human adult fibroblasts. *BMC Biotechnol.* **9**, 31 (2009).
39. Iwai, Y. *et al.* Axon patterning requires DN-cadherin, a novel neuronal adhesion receptor, in the *Drosophila* embryonic CNS. *Neuron* **19**, 77–89 (1997).
40. Vandaele, C., Coulon-Bublex, M., Couble, P. & Durand, B. *Drosophila* regulatory factor X is an embryonic type I sensory neuron marker also expressed in spermatids and in the brain of *Drosophila*. *Mech. Dev.* **103**, 159–162 (2001).
41. Bodily, K.D., Morrison, C.M., Renden, R.B. & Broadie, K. A novel member of the Ig superfamily, turtle, is a CNS-specific protein required for coordinated motor control. *J. Neurosci.* **21**, 3113–3125 (2001).
42. Brown, N.H. *et al.* Talin is essential for integrin function in *Drosophila*. *Dev. Cell* **3**, 569–579 (2002).
43. Yao, C.K. *et al.* A synaptic vesicle-associated Ca2+ channel promotes endocytosis and couples exocytosis to endocytosis. *Cell* **138**, 947–960 (2009).
44. Negre, N. *et al.* A cis-regulatory map of the *Drosophila* genome. *Nature* **471**, 527–531 (2011).
45. Gohl, D.M. *et al.* A versatile *in vivo* system for directed dissection of gene expression patterns. *Nat. Methods* **8**, 231–237 (2011).
46. Wesolowska, N. & Rong, Y.S. The past, present and future of gene targeting in *Drosophila*. *Fly (Austin)* **4**, 53–59 (2010).
47. Venken, K.J., He, Y., Hoskins, R.A. & Bellen, H.J. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in *D. melanogaster*. *Science* **314**, 1747–1751 (2006).
48. Venken, K.J. *et al.* Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. *Nat. Methods* **6**, 431–434 (2009).

ONLINE METHODS

Resource availability. Genomic sequences flanking MiMIC insertion sites in the 1,269 insertion lines selected for the GDP permanent collection have been submitted to GenBank, and supporting data have been deposited in FlyBase. Plasmids are available through the *Drosophila* Genomics Resource Center (https://dgrc.cgb.indiana.edu/vectors/). Fly strains are available through the BDSC.

Molecular biology. Primers were obtained from Operon or Sigma. PCR for cloning was performed with proofreading enzymes Pfu (Stratagene) or iProof (Biorad). Bacterial colony PCR was done with HotStarTaq DNA polymerase (Qiagen) or Hot MultiTaq DNA polymerase (US DNA). PCR purification and gel extraction were performed with the QIAquick PCR Purification and QIAquick Gel Extraction kits (Qiagen), respectively. Restriction enzymes and T4 DNA ligase were from NEB. The SURE or SURE2 bacterial strains (Stratagene) were used for bacterial transformation experiments. Bacteria were grown in LB broth containing 1% NaCl for plasmid isolation and ampicillin (USB) at a final concentration of 100 μg ml^{-1}. Plasmid purifications were performed using the QIAprep Spin Miniprep kit (Qiagen) or the PureLink HiPure Plasmid kit (Invitrogen) according to the manufacturer's instructions. All cloning experiments were verified by DNA sequencing.

Construction of the MiMIC transposon. pMiLRTetR was used as the *Minos* transposon backbone (gift of S. Oehler and C. Savakis)[49]. pMiLRTetR was digested with HindIII and PacI and ligated with two annealed oligonucleotides, pMiLR-Correction-TOP and pMiLR-Correction-Bottom (see **Supplementary Table 5** for all primer sequences), resulting in the mini *Minos* plasmid, pMiLR-Correction.

Next, two 100-bp *attP* sites were amplified by PCR from pTA-attP (gift of M. Calos)[50]: the first one was amplified with primers attP1-pMiLR-F and attP1-pMiLR-R, and the second one was amplified with primers attP2-pMiLR-F and attP2-pMiLR-R. Both *attP* amplicons, the first cut with HindIII and XhoI, and the second cut with XhoI and SacII, were ligated together into pMiLR-Correction cut with HindIII and SacII, resulting in the mini-Minos-attP plasmid, pMiLR-attP1-2. This plasmid has a stuffer fragment between each *attP* site and *Minos* inverted repeat that allows a specific inverse PCR reaction at either end, as well as a multiple cloning site between the two inverted *attP* sites.

Then the intronless dominant body color marker yellow$^+$ (*y[+mDint2]*) obtained from EPgy2 (ref. 4) was subcloned as a SalI fragment into the MCS of pMiLR-attP1-2, resulting in pMiLR-attP1-2-yellow.

Finally, a gene-trap cassette was constructed, consisting of the *Mhc* intron 18 splice acceptor (SA) site[51] obtained from pP-GC (gift of X. Morin and W. Chia)[15] and amplified with primers MHC-SA-XmaI-F and MHC-SA-EGFP-R, and an *EGFP* with an SV40 polyadenylation signal obtained from pCA-GAP-Mut4-EGFP (gift of A. Okada)[52] and amplified with primers MHC-SA-EGFP-F and EGFP-SpeI-R. The two fragments were fused together using hybrid PCR[53] and subcloned as an XmaI-SpeI fragment into pMiLR-attP1-2-yellow, resulting in the final transposon plasmid pMiLR-attP1-2-yellow-SA-EGFP, abbreviated MiMIC or referred to as Mi{MIC} in construct names.

Generation of single-insertion MiMIC lines. Initial MiMIC transposition experiments were performed by combined injection of both the MiMIC plasmid purified using the PureLink HiPure Plasmid kit and the mRNA encoding the *Minos* transposase generated from pBlueSKMimRNA (gift of A. Pavlopoulos)[54] by *in vitro* transcription after NotI linearization using the mMessage mMachine T7 kit (Ambion) as described previously[47,54]. Combined injection titration experiments were performed with plasmid:transposase concentrations of 300:300 ng μl^{-1}, 300:100 ng μl^{-1}, 100:300 ng μl^{-1} and 100:100 ng μl^{-1}. Injections were performed into a *y* w** strain. Transgenic lines were mapped by genetic crosses using *y* w*; T(2;3)ap^{Xa}/SM5;TM3, Sb* and balanced with *P{w[+mW.Scer\FRT.hs]=RS3}l(1)CB-6411-3^1, w^{1118}/FM7h* (BL6878) for the X chromosome, *y^1 w^{67c23}; In(2LR)Gla, wg^{Gla-1}/SM6a* (BL6600) for chromosome 2, *y* w*; D/TM6b, Hu, Tb* (Bellen lab) for chromosome 3 or *y*; ry^{506}/+; Ci^D/ey^D* for chromosome 4 (Bellen lab).

Subsequently, five MiMIC insertions on the X chromosome (*Mi{MIC}MI00019*, *Mi{MIC}MI00030*, *Mi{MIC}MI00039*, *Mi{MIC}MI00040* and *Mi{MIC}MI00069*), obtained form the co-injection experiments, were remobilized to the autosomes using a transgenic source of transposase under the control of a heat-shock promoter (*P{hsILMiT}*, FlyBase identifier FBtp0021508) inserted on a second chromosome balancer (*P{hsILMiT}2.4*; FlyBase identifier FBti0073645) (gift of C. Savakis)[8]. To allow the identification of new MiMIC remobilization events by screening for *y$^+$*, this transposase source was moved into a *y w* background with *y^1 w*; nub^{2}b^1 noc^{Sco}pr^1 cn^1/CyO* (BL3628) resulting in *y w; nub^{2}b^1 noc^{Sco}pr^1 cn^1/SM6a, P{hsILMiT}2.4*. Heat shocks were performed for 2 h in a 37° water bath on 5 consecutive days. Transposition efficiencies were initially as high as 43% for *Mi{MIC}MI00040*, but eventually dropped to about 7% for all donor insertions tested, for unknown reasons. Insertions in males were balanced with *y* w*; T(2;3)ap^{Xa}/SM5;TM3, Sb*. Insertions on chromosome 4 were balanced with *y*; ry^{506}/+; ci^D/ey^D* as described above.

During a second phase, three MiMIC insertions on the *X* chromosome (*Mi{MIC}MI00019*, *Mi{MIC}MI00030* and *Mi{MIC}MI00040*) were remobilized to the *TM3, Sb* balancer chromosome, resulting in two independent chromosome 3 transposon-donor balancer chromosomes, which we named *MI00000A* and *MI00000B*. These donor chromosomes were used to remobilize MiMIC to the X chromosome and autosomes using the *y w; noc^{Sco}nub^{2}b^1 noc^{Sco}pr^1 cn^1/SM6a, P{hsILMiT}2.4* stock described above. In addition, *Mi{MIC}MI00827* located on the X chromosome was mobilized to the autosomes.

Mapping and annotation of MiMIC insertion sites. We generated 4,464 strains containing insertions of the MiMIC transposon (nearly always single insertions) and mapped 3,633 insertions to a unique site in the reference genome sequence (release 5; http://www.fruitfly.org/). Sequences flanking MiMIC insertions were determined by inverse PCR and DNA sequencing, and mapped by alignment to the genome sequence, as described for MB lines in reference 5 with one modification: genomic DNA was digested with Sau3A I or MboI (which are isoschizomers). A detailed protocol is available on the GDP website (http://flypush.imgen.bcm.tmc.edu/pscreen/). Lines that were selected for the GDP collection were balanced, and their insertion sites were verified by resequencing before delivery to the BDSC.

doi:10.1038/nmeth.1662

We associated MiMIC insertions with annotated genes and gene features (FlyBase r5.32). Annotated features of one gene transcript often overlap those of another transcript, so we adopted a progressive strategy for associating MiMIC insertions with gene features, such that each insertion was assigned to only one feature. We first associated insertions with coding exons, followed by 5′ UTR exons, 3′ UTR exons, coding introns, 5′ UTR introns and 3′ UTR introns. The remaining insertions were classified as intergenic. Note that this is a conservative approach that underestimates the number of insertions associated with lower ranked gene features.

Construction of correction cassettes for RMCE. Two 100-bp fragments containing *attB* sites were obtained by PCR from pTA-attB (gift of M. Calos)[50], the first amplified with primers attB1-pBS-F and attB1-pBS-R, and the second amplified with primers attB2-pBS-F and attB2-pBS-R. The *attB* amplicons, the first cut with SacI and EcoRI, and the second cut with EcoRI and KpnI, were ligated together into pBS-KS and pBS-SK cut with SacI and KpnI, resulting in the mini-attB-RMCE plasmids, pBS-KS-attB1-2 and pBS-SK-attB1-2. These plasmids have two inverted *attB* sites flanking a multiple cloning site (XbaI, SpeI, PstI, EcoRI, XhoI, BamHI and HindIII).

Construction of gene-trap cassettes for RMCE. A gene-trap cassette incorporating the *Mhc* intron 18 SA site[51] was PCR amplified from pP-GC (gift of X. Morin and W. Chia)[15] with primers SA-XbaI-F and SA-PstI-R. The resulting PCR fragment was cut with XbaI and PstI and subcloned into pBS-KS-attB1-2, cut with XbaI and PstI, resulting in the mini gene-trap plasmid, pBS-KS-attB1-2-GT-SA. The SA is followed by a multiple cloning site (PstI, EcoRI, XhoI, BamHI and HindIII).

A mutagenic *GAL4* gene-trap cassette, encompassing the *GAL4* coding sequence and *Hsp70* polyadenylation signal, was obtained from plasmid pChs-GAL4 (*Drosophila* Genomics Resource Center)[55], and PCR amplified with primers GAL4-Hsp70-EcoRI-F and GAL4-Hsp70-BamHI-R. A mutagenic *QF* gene-trap cassette, encompassing the *QF* coding sequence and *Hsp70* polyadenylation signal, was obtained from plasmid pattB-QF-Hsp70 (Addgene)[30] and PCR-amplified with primers QF-SV40-EcoRI-F and QF-SV40-BamHI-R. A mutagenic *Flp* fate mapping gene-trap cassette, encompassing the *FLPo*[56] coding sequence and SV40 polyadenylation signal, was obtained from plasmid pQUAS-DSCP-Flpo (Addgene)[30] and PCR-amplified with primers Flpo-SV40-EcoRI-F and Flpo-SV40-BamHI-R. The resulting PCR fragments were cut with EcoRI and BamHI and subcloned into pBS-KS-attB1-2-GT-SA, cut with EcoRI and BamHI, resulting in the plasmids pBS-KS-attB1-2-GT-SA-GAL4-Hsp70, pBS-KS-attB1-2-GT-SA-Flp-SV40 and pBS-KS-attB1-2-GT-SA-QF-Hsp70, respectively (see **Supplementary Data** for all plasmid insert sequences).

Construction of protein-trap cassettes for RMCE. Protein-trap cassettes were constructed for the three intron phases (0, 1 and 2). The *Mhc* intron 18 SA and intron 17 splice donor (SD) sites[51] were obtained from pP-GC (gift of X. Morin and W. Chia)[15]. The protein trap with splice phase 0 was generated from two PCR fragments, the SA site amplified with primers SA-XbaI-F and SA-SD-Phase-0-R, and the SD site amplified with primers SA-SD-Phase-0-F and SD-HindIII-R. The protein trap with splice phase 1 was generated from two PCR fragments, the SA site amplified with primers SA-XbaI-F and SA-SD-Phase-1-R, and the SD site amplified with primers SA-SD-Phase-1-F and SD-HindIII-R. The protein trap with splice phase 2 was generated from two PCR fragments, the SA site amplified with primers SA-XbaI-F and SA-SD-Phase-2-R, and the SD site amplified with primers SA-SD-Phase-2-F and SD-HindIII-R. For each intron phase construct, the SA PCR fragment was cut with XbaI and BamHI, the SD PCR fragment was cut with BamHI and HindIII, and the digested fragments were subcloned in a three-way ligation into pBS-KS-attB1-2, cut with XbaI and HindIII, resulting in pBS-KS-attB1-2-PT-SA-SD-0, pBS-KS-attB1-2-PT-SA-SD-1 and pBS-KS-attB1-2-PT-SA-SD-2. These plasmids contain the SA site, a phase linker for phase 0, 1 or 2, and the SD site, between two inverted *attB* sites. The phase linker consists of a BamHI site used to sublone protein-trap tags (see below) between two $(GlyGlySer)_4$ peptide-encoding linkers that may provide flexibility between the protein trap tag and the endogenous protein sequences (**Supplementary Fig. 3**).

The fluorescent protein tag mCherry (gift of R.Tsien)[57] was used without codon optimization. The following protein and peptide tags were generated by gene synthesis by GENEART (http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Cloning/gene-synthesis.html?CID=fl-genesynthesis) with codon usage biased to *D. melanogaster*: superfolder GFP[58], enhanced blue fluorescent protein 2 (ref. 59), TagRFP-T[60], HRP[61], Dendra2 (refs. 62,63), KillerRed[64], optimized FlAsH peptide[65], StrepII peptide[66], 3×Flag peptide[67], 3×Myc peptide[68], 3×HA peptide[69] and the TEV protease site[70]. The following peptide tags were generated by PCR amplification and primer addition with codon usage biased to *D. melanogaster* based on the codon use database (http://www.kazusa.or.jp/codon/): S peptide[71] and V5 peptide[72]. The following mult-itags were generated through hybrid PCR[53] or PCR and primer addition, and subcloned in customized vector backbones (K.J.T.V., unpublished data): EGFP-FlAsH-StrepII-TEV-3xFlag, EBFP2-3xMyc, TagRFP-T-3xHA, HRP-S, Dendra-V5 and KillerRed-V5.

The tags were then amplified by PCR as BamHI-insert-GGC-BamHI fragments, to correct for cloning and reconstitution of the $(Gly\text{-}Gly\text{-}Ser)_4$ linker (**Supplementary Fig. 3**) and subcloned into the three-phase protein-trap plasmids described above. *EGFP-FlAsH-StrepII-TEV-3xFlag* was amplified with primers EGFPmultiFINAL-F and EGFPmultiFINAL-R, resulting in pBS-KS-attB1-2-PT-SA-SD-0-EGFP-FlAsH-StrepII-TEV-3xFlag, pBS-KS-attB1-2-PT-SA-SD-1-EGFP-FlAsH-StrepII-TEV-3xFlag and pBS-KS-attB1-2-PT-SA-SD-2-EGFP-FlAsH-StrepII-TEV-3xFlag. PCR amplification of *mCherry* was performed with primers Cherry-F and Cherry-R, resulting in pBS-KS-attB1-2-PT-SA-SD-0-mCherry, pBS-KS-attB1-2-PT-SA-SD-1-mCherry and pBS-KS-attB1-2-PT-SA-SD-2-mCherry. PCR amplification of *EBFP2-3xMyc* was performed with primers EBFP2-Myc-F and EBFP2-Myc-R resulting in pBS-KS-attB1-2-PT-SA-SD-0-EGFP-EBFP2-3xMyc, pBS-KS-attB1-2-PT-SA-SD-1-EBFP2-3xMyc and pBS-KS-attB1-2-PT-SA-SD-2-EBFP2-3xMyc. PCR amplification of *TagRFP-T-3xHA* was performed with primers TagRFP-HA-F and TagRFP-HA-R, resulting in pBS-KS-attB1-2-PT-SA-SD-0-TagRFP-T-3xHA, pBS-KS-attB1-2-PT-SA-SD-1-TagRFP-T-3xHA and pBS-KS-attB1-2-PT-SA-SD-2-TagRFP-T-3xHA. PCR amplification of *HRP-S* was

performed with primers HRP-S-F, HRP-S-R1 and HRP-S-R2, resulting in pBS-KS-attB1-2-PT-SA-SD-0-HRP-S, pBS-KS-attB1-2-PT-SA-SD-1-HRP-S and pBS-KS-attB1-2-PT-SA-SD-2-HRP-S. PCR amplification of *Dendra-V5* was performed with primers Dendra-V5-F, Dendra-V5-R1 and Dendra-V5-R2, resulting in pBS-KS-attB1-2-PT-SA-SD-0-Dendra-V5, pBS-KS-attB1-2-PT-SA-SD-1-Dendra-V5 and pBS-KS-attB1-2-PT-SA-SD-2-Dendra-V5. PCR amplification of *KillerRed-V5* was performed with primers KillerRed-V5-F, KillerRed-V5-R1 and KillerRed-V5-R2 resulting in pBS-KS-attB1-2-PT-SA-SD-0-KillerRed-V5, pBS-KS-attB1-2-PT-SA-SD-1-KillerRed-V5 and pBS-KS-attB1-2-PT-SA-SD-2-KillerRed-V5.

ΦC31 integrase–mediated RMCE. Initial RMCE tests were performed with pBS-KS-attB1-2 to ensure functionality of the plasmid backbone, since this plasmid is the progenitor of all constructs for protein-trap cassettes and other cassettes. pBS-KS-attB1-2 DNA was purified and injected along with mRNA encoding ΦC31 integrase, obtained from pET11ΦC31pA (gift of M. Calos)[23] by *in vitro* transcription after BamHI linearization using the mMessage mMachine T7 kit as described previously[23,47]. Microinjections were performed using the RMCE landing site 25C (gift of J. Bateman and T. Wu)[28] at a plasmid concentration of 123 ng μl^{-1} and an mRNA concentration of 600 ng μl^{-1}. A transgenesis efficiency of 17.5% was obtained.

Subsequent injections were performed with a transgenic ΦC31 integrase gene source driven by *vasa* promoter elements located on the X chromosome (y^1*M{vas-int.B}ZH-2A w**) (gift of J. Bischof, F. Karch and K. Basler)[24]. Plasmid was generally diluted to a concentration between 30 ng μl^{-1} and 100 ng μl^{-1}. Microinjections were performed using the following MiMIC insertion lines: *Mi{MIC}tutl*^{MI00290} and *Mi{MIC}CadN*^{MI00393} on chromosome 2, and *Mi{MIC}Rfx*^{MI00053}, *Mi{MIC}gogo*^{MI00065}, *Mi{MIC}Tl*^{MI00181}, *Mi{MIC}caps*^{MI00249}, *Mi{MIC}rhea*^{MI00296}, *Mi{MIC}MYPT-75D*^{MI00314}, *Mi{MIC}BM-40-SPARC*^{MI00329}, *Mi{MIC}comm*^{MI00380} and *Mi{MIC}wnd*^{MI00494} on chromosome 3. When lines contained a gene-trap insertion, they were injected as heterozygous balanced stocks. Microinjections were performed by crossing males from appropriate MiMIC lines to virgin females containing the ΦC31 integrase source. As RMCE results in a genetically unmarked chromosome owing to the removal of the *y[+mDint2]* marker, fly stocks were generated that contained the ΦC31 integrase source in a balanced background for the second or third chromosome to maintain the MiMIC insertion balanced in G_0 flies: y^1 *M{vas-int.B}ZH-2A w*; noc*^{Sco}*/CyO* and y^1 *M{vas-int.B}ZH-2A w*; Sb/TM6b, Hu, Tb*. Appropriate G_0 flies were crossed to balancer virgins: y^1 w^{67c23}*; In(2LR)Gla, wg*$^{Gla-1}$*/SM6a* (BL6600) for chromosome 2 RMCE experiments, or *y* w*; D/TM6b, Hu, Tb* for chromosome 3 RMCE experiments. Transgenic G_1 flies were scored for the absence of a yellow$^+$ phenotype (loss of the *y[+mDint2]* marker) over a balancer or dominantly marked chromosome appropriate for the chromosome, and crossed to balancer virgins: y^1 w^{67c23}*; In(2LR)Gla, wg*$^{Gla-1}$*/SM6a* (BL6600) (chromosome 2), or *y* w*; D/TM6b, Hu, Tb* (chromosome 3). Balanced transgenic G_2 flies were intercrossed to establish stocks. A list with all RMCE efficiencies is available in **Supplementary Table 6**.

Molecular characterization of integration events. For PCR verification of RMCE integration events, DNA was extracted from 10 to 15 adult flies using the PureLink Genomic DNA Mini kit. PCR was performed with tag sequence–specific primers and MiMIC-specific primers. Tag sequence–specific primers (Tag-F and Tag-R) are mCherry-Seq-F and mCherry-Seq-R for *mCherry*, EGFPdo-Seq-F and EGFPdo-Seq-R for EGFP, EBFP2do-Seq-F and EBFP2do-Seq-R for *EBFP2*, TagRFPdo-Seq-F and TagRFPdo-Seq-R for *TagRFP*, Hrpdo-Seq-F and Hrpdo-Seq-R for *HRP*, Dendrado-Seq-F and Dendrado-Seq-R for *Dendra*, Killerreddo-Seq-F and Killerreddo-Seq-R for *Killerred*, GAL4-1R and GAL4-5F for *GAL4*, FLP0-Seq-R and SV40pA-Long-F for *Flpo*, and QF-Seq-R1 and Hsp70-pA-Alt-F for *QF*. MiMIC specific primers are Orientation-MiL-F and Orientation-MiL-R. PCR reaction conditions were: 1 μl DNA, 1 μl primer 1, 1 μl primer 2, 2 μl 10× buffer, 0.16 μl dNTPs (25 mM each), 0.08 μl Qiagen HotStarTaq DNA Polymerase (Qiagen), 14.76 μl milliQ water. PCR cycling conditions in PTC-225 or DNA Engine (MJ Research) were: denaturation at 94° for 10 min, 40 cycles at 94° for 30 s, 60° for 30 s and 72° for 60 s, and extension at 72° for 10 min.

For each RMCE event, four PCRs were performed: a first PCR with primers Orientation-MiL-F and Tag-R, a second PCR with primers Orientation-MiL-F and Tag-F, a third PCR with primers Orientation-MiL-R and Tag-R, and a fourth PCR with primers Orientation-MiL-R and Tag-F. As the transposon integrates one or two orientations relative to the gene, only one in two RMCE events is productive with respect to creating a gene trap or protein trap, which is reflected in a positive PCR for reactions 1 and 4, or 2 and 3. A '1/4' PCR pattern was always desired for a productive RMCE event (for example, a gene or protein trap), when the gene or transposon configuration is 1/1 or −1/−1. Conversely, a '2/3' PCR pattern is diagnostic of a productive RMCE event, when the gene or transposon configuration is 1/−1 or −1/1. The reverse holds for unproductive RMCE events (**Supplementary Fig. 2**).

Genetic complementation testing. Genetic complementation tests were performed between lethal MiMIC insertion lines or lethal RMCE derivatives of both lethal and viable MiMIC lines, and previously described mutant alleles. *Mi{MIC}Rfx*^{MI00053} was tested with *Rfx*^{49} and *Rfx*^{253} (gifts of A. Laurençon and B. Durand)[32] and *Df(3R)Exel6157* (BL7636)[73]. *Mi{MIC}tutl*^{MI00290} was tested with *P{ry*$^{+t7.2}$*= PZ}tutl*^{01085} (Bloomington stock number BL10979), *tutl*^{4} (gift of K. Broadie)[41], *tutl*^{23} and *tutl*^{GAL4} (gifts of Yong Rao)[74], *tutl*^{ex383} (gift of B. Al-Anzi)[75] and *Df(2L)ed-dp* (BL702). *Mi{MIC}rhea*^{MI00296} was tested with *rhea*^{1} (BL2296) and *Df(3L)W10* (BL2608). *Mi{MIC}comm*^{MI00380} was tested with *comm*^{A490} and *comm*$^{\Delta e39}$ (gifts of G. Tear)[76], and *P{w*$^{+mC}$*y*$^{+mDint2}$*= EPgy2}comm*^{EY10154} (BL17644), *Df(3L)BK10* (BL2992) and *Df(3L)fz-M21* (BL5461). *Mi{MIC}CadN*^{MI00393} was tested with *CadN*^{M19}*FRT40A* and *CadN*$^{\Delta 18A}$*FRT40A* (gifts of L. Zipursky)[39,77], and *Df(2L)Exel7069* (BL7837). *Mi{MIC}wnd*^{MI00494} was tested with *wnd*^{1}, *wnd*^{2} and *wnd*^{3} (gifts of A. DiAntonio)[31], and *Df(3L)XS705* (BL5584) and *Df(3L)Exel9007* (BL7942). Complementation tests for lethal protein-trap events were performed in a similar fashion.

***GAL4*-UAS, *QF-QUAS* and *Flp* experiments.** RMCE experiments using the binary expression factors GAL4 and QF, and the Flp recombinase system were tested as follow: GAL4 swaps were crossed to *y w; VK19::10xUAS-mCherry-SV40* which provides strong mCherry overexpression driven by 10xUAS in a

doi:10.1038/nmeth.1662

customized P[acman] construct (K.J.T.V., unpublished data) and were analyzed as described below. QF swaps were crossed to y^1 w^{1118}; $P\{w^{+mC}$= *QUAS-mtdTomato-3xHA}26* (BL30005)[30] and analyzed as described below. Flp swaps were crossed to a customized actin-GAL4-Flp-out line driving UAS-EGFP (H.J.B., unpublished data) and analyzed as described below.

Expression analysis. The following antibodies were used for expression analysis: mouse antibody to CadN (DN-Ex#8) at 1:200 (ref. 39; Developmental Studies Hybridoma Bank), rabbit antibody to RFX at 1:5,000 (gift from A. Laurençon and B. Durand)[40], rabbit antibody to Wnd at 1:500 (gift from A. DiAntonio)[31], rabbit antibody to GFP at 1:250 (Invitrogen), mouse antibody to DsRed at 1:250 (Clontech), rabbit antibody to TagRFP at 1:500 (Evrogen), rabbit antibody to Dendra2 at 1:5,000 (Evrogen), rabbit antibody to Killerred at 1:1,000 (Evrogen), mouse antibody to Flag at 1:250 (Sigma-Aldrich), mouse antibody to StrepII at 1:200 (Thermo Scientific), mouse antibody to the S epitope at 1:100 (Thermo Scientific), mouse anti-V5 at 1:2,000 (Invitrogen), mouse antibody to c-Myc at 1:250 (Abcam) and mouse antibody to the HA epitope at 1:200 (Covance).

Drosophila embryos (0–24 h) were collected on grape-agar plates and were subsequently fixed for 20 min in a 1:1 mixture of 0.38% formaldehyde in PBS (pH 7.0) and heptane. The fixative was then removed and methanol added. After vigorously shaking, the heptane-methanol mixture was replaced by methanol, and then methanol was replaced by ethanol. Upon rehydration in PBS with 0.2% Triton, embryos were blocked for 1 h in PBS, 10% normal goat serum and incubated overnight with primary antibodies. Fluorescently labeled or HRP-conjugated secondary antibodies were obtained from Jackson ImmunoResearch and were used at a 1:250 dilution.

mRNA *in situ* hybridization. A 1-kb PCR fragment was obtained with primers Wnd-F and Wnd-R from the Wnd cDNA clone LD14856[78] and subcloned into the pGemTeasy vector (Promega). *In vitro* transcription was performed according to standard procedures, using the digoxigenin (DIG) RNA labeling kit (Roche). Fixation and *in situ* hybridization were carried out as previously described[79]. The digoxigenin-labeled RNA probes were detected by alkaline phosphatase reaction as previously described[80].

49. Klinakis, A.G., Loukeris, T.G., Pavlopoulos, A. & Savakis, C. Mobility assays confirm the broad host-range activity of the Minos transposable element and validate new transformation tools. *Insect Mol. Biol.* **9**, 269–275 (2000).
50. Groth, A.C., Olivares, E.C., Thyagarajan, B. & Calos, M.P. A phage integrase directs efficient site-specific integration in human cells. *Proc. Natl. Acad. Sci. USA* **97**, 5995–6000 (2000).
51. Hodges, D. & Bernstein, S.I. Suboptimal 5′ and 3′ splice sites regulate alternative splicing of *Drosophila melanogaster* myosin heavy chain transcripts *in vitro*. *Mech. Dev.* **37**, 127–140 (1992).
52. Okada, A., Lansford, R., Weimann, J.M., Fraser, S.E. & McConnell, S.K. Imaging cells in the developing nervous system with retrovirus expressing modified green fluorescent protein. *Exp. Neurol.* **156**, 394–406 (1999).
53. Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. & Pease, L.R. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61–68 (1989).
54. Pavlopoulos, A., Berghammer, A.J., Averof, M. & Klingler, M. Efficient transformation of the beetle Tribolium castaneum using the Minos transposable element: quantitative and qualitative analysis of genomic integration events. *Genetics* **167**, 737–746 (2004).
55. Apitz, H. *et al.* Identification of regulatory modules mediating specific expression of the roughest gene in *Drosophila melanogaster*. *Dev. Genes Evol.* **214**, 453–459 (2004).
56. Raymond, C.S. & Soriano, P. High-efficiency FLP and ϕC31 site-specific recombination in mammalian cells. *PLoS ONE* **2**, e162 (2007).
57. Shaner, N.C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* **22**, 1567–1572 (2004).
58. Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C. & Waldo, G.S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–88 (2006).
59. Ai, H.W., Shaner, N.C., Cheng, Z., Tsien, R.Y. & Campbell, R.E. Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins. *Biochemistry* **46**, 5904–5910 (2007).
60. Shaner, N.C. *et al.* Improving the photostability of bright monomeric orange and red fluorescent proteins. *Nat. Methods* **5**, 545–551 (2008).
61. Smith, A.T. *et al.* Expression of a synthetic gene for horseradish peroxidase C in *Escherichia coli* and folding and activation of the recombinant enzyme with Ca2+ and heme. *J. Biol. Chem.* **265**, 13335–13343 (1990).
62. Gurskaya, N.G. *et al.* Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nat. Biotechnol.* **24**, 461–465 (2006).
63. Chudakov, D.M., Lukyanov, S. & Lukyanov, K.A. Using photoactivatable fluorescent protein Dendra2 to track protein movement. *Biotechniques* **42**, 553–557 (2007).
64. Bulina, M.E. *et al.* A genetically encoded photosensitizer. *Nat. Biotechnol.* **24**, 95–99 (2006).
65. Martin, B.R., Giepmans, B.N., Adams, S.R. & Tsien, R.Y. Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat. Biotechnol.* **23**, 1308–1314 (2005).
66. Schmidt, T.G., Koepke, J., Frank, R. & Skerra, A. Molecular interaction between the Strep-tag affinity peptide and its cognate target, streptavidin. *J. Mol. Biol.* **255**, 753–766 (1996).
67. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl. Microbiol. Biotechnol.* **60**, 523–533 (2003).
68. Evan, G.I., Lewis, G.K., Ramsay, G. & Bishop, J.M. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol. Cell. Biol.* **5**, 3610–3616 (1985).
69. Wilson, I.A. *et al.* The structure of an antigenic determinant in a protein. *Cell* **37**, 767–778 (1984).
70. Dougherty, W.G., Cary, S.M. & Parks, T.D. Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology* **171**, 356–364 (1989).
71. Hackbarth, J.S. *et al.* S-peptide epitope tagging for protein purification, expression monitoring, and localization in mammalian cells. *Biotechniques* **37**, 835–839 (2004).
72. Southern, J.A., Young, D.F., Heaney, F., Baumgartner, W.K. & Randall, R.E. Identification of an epitope on the P and V proteins of simian virus 5 that distinguishes between two isolates with different biological characteristics. *J. Gen. Virol.* **72**, 1551–1557 (1991).
73. Parks, A.L. *et al.* Systematic generation of high-resolution deletion coverage of the *Drosophila melanogaster* genome. *Nat. Genet.* **36**, 288–292 (2004).
74. Ferguson, K., Long, H., Cameron, S., Chang, W.T. & Rao, Y. The conserved Ig superfamily member Turtle mediates axonal tiling in *Drosophila*. *J. Neurosci.* **29**, 14151–14159 (2009).
75. Al-Anzi, B. & Wyman, R.J. The *Drosophila* immunoglobulin gene turtle encodes guidance molecules involved in axon pathfinding. *Neural Dev.* **4**, 31 (2009).
76. Tear, G. *et al.* commissureless controls growth cone guidance across the CNS midline in *Drosophila* and encodes a novel membrane protein. *Neuron* **16**, 501–514 (1996).
77. Nern, A. *et al.* An isoform-specific allele of *Drosophila* N-cadherin disrupts a late step of R7 targeting. *Proc. Natl. Acad. Sci. USA* **102**, 12944–12949 (2005).
78. Stapleton, M. *et al.* A *Drosophila* full-length cDNA resource. *Genome Biol.* **3**, 0080 (2002).
79. Lecuyer, E., Parthasarathy, N. & Krause, H.M. Fluorescent *in situ* hybridization protocols in *Drosophila* embryos and tissues. *Methods Mol. Biol.* **420**, 289–302 (2008).
80. Lehmann, R. & Tautz, D. *In situ* hybridization to RNA. *Methods Cell Biol.* **44**, 575–598 (1994).