

# Inducible DamID systems for genomic mapping of chromatin proteins in *Drosophila*

Alexey V. Pindyurin<sup>1,2,\*</sup>, Ludo Pagie<sup>1</sup>, Elena N. Kozhevnikova<sup>2</sup>, Joris van Arensbergen<sup>1</sup> and Bas van Steensel<sup>1,\*</sup>

<sup>1</sup>Division of Gene Regulation, Netherlands Cancer Institute, 1066 CX Amsterdam, the Netherlands and <sup>2</sup>Institute of Molecular and Cellular Biology SB RAS, Novosibirsk 630090, Russia

Received January 04, 2016; Revised March 02, 2016; Accepted March 08, 2016

## ABSTRACT

**Dam identification (DamID) is a powerful technique to generate genome-wide maps of chromatin protein binding. Due to its high sensitivity, it is particularly suited to study the genome interactions of chromatin proteins in small tissue samples in model organisms such as *Drosophila*. Here, we report an intein-based approach to tune the expression level of Dam and Dam-fusion proteins in *Drosophila* by addition of a ligand to fly food. This helps to suppress possible toxic effects of Dam. In addition, we describe a strategy for genetically controlled expression of Dam in a specific cell type in complex tissues. We demonstrate the utility of the latter by generating a gliaspecific map of Polycomb in small samples of brain tissue. These new DamID tools will be valuable for the mapping of binding patterns of chromatin proteins in *Drosophila* tissues and especially in cell lineages.**

## INTRODUCTION

Structural and regulatory components of chromatin are responsible for control of gene expression, DNA replication, and genome maintenance. Genome-wide mapping of chromatin proteins contributes to the understanding of the mechanisms of these processes. One of the techniques for this purpose is Dam identification (DamID) (1,2), which is based on *in vivo* expression of a chromatin protein of interest fused to DNA adenine methyltransferase (Dam). This way, Dam is targeted to the native binding sites of its fusion partner and methylates adenines in GATC sequences in their immediate vicinity. The methylated regions are subsequently selectively amplified from genomic DNA and identified by microarray hybridization or high-throughput sequencing. DamID has been successfully applied in a broad range of species (3–9). Particularly in *Drosophila* the method has been used to study a large di-

versity of DNA-binding factors and chromatin proteins (10,11).

In general, a requirement for DamID experiments is that the expression level of the Dam fusion proteins is kept very low in order to avoid saturating methylation of genomic DNA due to the high enzymatic activity of Dam, and to prevent mis-targeting of the Dam fusion protein (2,12). In *Drosophila* cultured cells, larvae and flies, this low expression is typically achieved by means of the low leaky activity of the *hsp70* heat-shock promoter in the absence of heat shock.

However, it is often desirable that the expression of the Dam fusion protein is controllable. For example, one may need to express a Dam fusion protein in a defined time window, for example to study a specific developmental stage or after application of a certain stimulus. In addition, it is useful to be able to activate Dam expression only in a particular cell type of interest. Because the amplification of Dam-methylated DNA is highly specific (13,14) and extremely sensitive (15), such selective expression would make it possible to generate cell type specific protein binding maps from small heterogeneous tissue samples without the need to purify the cells of interest. A requirement for this strategy is that the Dam activity is strictly limited to the cell type of interest, otherwise signals will be picked up from irrelevant cell types.

Another reason why controllable expression of Dam fusion proteins is desirable lies in possible toxic effects. Although it was originally reported that expression of Dam in *Drosophila* has no detectable effects on fly development and viability (16), subsequent experiments indicated that transgenesis with Dam-expressing vectors can be inefficient (14,17). For example, in an attempt to generate Dam-only, SUUR-Dam and Dam-SUUR transgenic flies by random integration we found that the transgenes were inactivated in 14 out of 14 independent lines due to sequence rearrangements (A.V.P., unpublished data), suggesting strong negative selection. Furthermore, we observed very low efficiency of phiC31 integrase-mediated site-specific insertion

\*To whom correspondence should be addressed. Tel: +31 20 5122040; Email: b.v.steensel@nki.nl  
Correspondence may also be addressed to Alexey V. Pindyurin. Tel: +7 383 3639042; Fax: +7 383 3639078; Email: a.pindyurin@mcb.nsc.ru

of Dam-histone H1 and Dam-BRM transgenes: 0 and 1 line from 200 and 400 embryos injected, respectively (J. van Bommel and B.v.S., unpublished data). The underlying cause of the inefficient transgenesis of Dam constructs has never been satisfactorily explained. However, the suspected toxicity problem may be overcome by reducing the activity of the Dam fusion protein, or by restricting its expression to the cells or tissue of interest.

Here, we report two approaches to achieve controllable activity of Dam fusion proteins. First, we designed a ligand-dependent version of Dam that enables the tuning of Dam activity in *Drosophila* tissues. In addition, we present a genetically controllable binary switch to express Dam fusion proteins selectively in a *Drosophila* cell type or tissue of choice, and we demonstrate its utility for generating cell type specific binding maps of the Polycomb protein. We discuss the merits of these and other recently reported systems for controllable Dam expression in *Drosophila*.

## MATERIALS AND METHODS

### Dam-containing plasmid constructs

Plasmid constructs are described in the Extended Experimental Procedures.

### Fly stocks and handling

Transgenic flies were generated using phiC31 integrase-mediated site-specific transgenesis system (18) by the BestGene company (<http://www.thebestgene.com/>). All Dam-containing plasmid constructs were injected into embryos containing attP site at cytogenetic location 51C (*y[1] M{vas-int.Dm}ZH-2A w[\*]; M{3xP3-RFP.attP}ZH-51C*; Bloomington stock no. 24482). Up to five independent transformant lines were established for each construct and only homozygous-viable lines were chosen for the experiments. All transgenic lines were verified by genotyping PCR (data not shown).

The *Oregon-R* strain was used as wild type. The repo-FLP stock (19) was provided by Christian Klämbt (Institut für Neurobiologie, Universität Münster, Münster, Germany). The  $\beta$ Tub85D-FLP stock was obtained from the Bloomington *Drosophila* Stock Center (no. 7196). The *y[\*] w[\*]* strain was provided by François Karch (University of Geneva, Geneva, Switzerland). Flies were raised at 25°C on standard cornmeal/molasses/agar medium. To induce 4-HT-intein splicing, fly food was supplemented with 4-hydroxytamoxifen (Sigma-Aldrich, H7904) at a final concentration of 25  $\mu$ M. Mated female flies were then allowed to lay eggs on this food, so the resulting larvae were exposed to 4-HT from hatching until they were collected in the third instar stage.

### RNA-seq

For each gene expression sample, larval central brain or fat bodies were dissected into TRIzol reagent (Invitrogen, 15596-018). Tissues were dissected from male wandering third instar larvae. For each tissue/condition, two independent samples were processed. RNA-seq was performed using an Illumina HiSeq 2000 instrument and standard Il-

lumina protocol. Differentially expressed genes were identified by using the R package DESeq (20,21; <http://www-huber.embl.de/users/anders/DESeq/>).

### DamID-seq

Isolation of total DNA from whole adults, larval central brain and fat bodies is described in the Extended Experimental Procedures. For each stage/tissue/protein/condition, two independent samples were processed. For the larval central brain and fat body samples, the appropriate tissues were dissected from male wandering third instar larvae. For the repo-positive glial cell samples, no gender selection was employed; a mix of male and female wandering third instar larvae was used for dissection of larval central brains. Methylated GATC DNA sequences were amplified by DpnI digestion followed by ligation-mediated PCR as described previously (22) with a minor modification: 500 ng instead of 2.5  $\mu$ g of DNA was used as an input. For the experiments with an individual *Drosophila* larval central brain, total amount of isolated DNA was used as an input. The amplified methylated DNA fragments were purified using QIAquick PCR Purification kit (QIAGEN, 28104) according to the manufacturer's instructions, eluted in nuclease-free water (Ambion, AM9938) and quantified on a NanoDrop spectrophotometer. Preparation of the purified DNA for Illumina sequencing is described in the Extended Experimental Procedures. 50 bp single-end read cycles of sequencing were performed on an Illumina HiSeq 2000 machine.

### Processing of DamID-seq fastq files

Because reads with GATC sites contain the DamID adapter sequence, this sequence was first clipped from such reads using the software package 'cutadapt' (version 1.2.1; <https://cutadapt.readthedocs.org/en/stable/>). All reads were then aligned to the *Drosophila melanogaster* genome (release 5) using the short read alignment software 'bowtie2' (version 2.0.0-beta2; <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>). Next, the reads were mapped onto GATC fragments using the software 'HTSeq-count' (version 0.5.3p3; <http://www-huber.embl.de/users/anders/HTSeq/doc/index.html>). Next, counts were normalized to the total number of reads. Finally, Dam<sup>4-HT-intein@L127C</sup>-PC and STOP#1-Dam-PC data were normalized to the Dam<sup>4-HT-intein@L127C</sup> and STOP#1-Dam alone controls, respectively, and the replicates were averaged. A detailed description of data transformation and the code are available upon request.

### Data release

DamID-seq and RNA-seq data are available from the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>), accession number GSE75835.

## RESULTS

### Construction of a ligand-controlled Dam vector for use in *Drosophila* tissues

In order to establish a ligand-inducible version of Dam we tested an intein-based molecular switch (Figure 1A). Inteins are internal protein elements catalyzing their own excision and joining of the flanking polypeptide sequences. Modified inteins can promote conditional protein splicing in a ligand-dependent manner (23,24). We reasoned that insertion of such a modified intein within Dam might block its enzymatic activity, which can be restored by the activation of intein splicing. Based on the crystal structure of Dam we chose eight different positions (S57, C87, L127, R128, C162, A170, C180 and H220) for the insertion of the 4-hydroxytamoxifen-inducible intein (23) (hereafter, 4-HT-intein). Because intein splicing requires the presence of a cysteine residue in the host protein immediately downstream of the inserted intein (23–25), we first tested whether S57C, L127C, R128C, A170C and H220C mutations affect the functioning of the Dam enzyme. We introduced these mutations in the pNDamMyc plasmid encoding Dam under the control of the *Drosophila* heat-shock promoter *hsp70* (1).

To assay the activity of the Dam mutants, we took advantage of a fortuitous observation that the *hsp70* promoter can drive Dam expression in *E. coli*. We amplified the mutant-encoding plasmids in a *dam<sup>-</sup>* *E. coli* strain and then isolated and digested the plasmids with DpnI and DpnII restriction enzymes that cut only G<sup>m6</sup>ATC and unmethylated GATC motifs, respectively. This revealed that all tested Dam mutants retained DNA adenine methyltransferase activity (Figure 1B).

Next, we cloned the 4-HT-intein in the pNDamMyc plasmid at the selected (and, where necessary, mutated) positions in Dam. We did not detect any adenine methylation in *dam<sup>-</sup>* bacteria transformed by the constructed plasmids and grown in the absence of 4-hydroxytamoxifen (4-HT). However, in the presence of the ligand (10  $\mu$ M 4-HT), we observed partial methylation of some plasmid constructs (Figure 1C). We chose one of these promising chimeric constructs, with the 4-HT-intein inserted just before L127C (hereafter, Dam<sup>4-HT-intein@L127C</sup>), for the subsequent experiments in *Drosophila* flies.

For expression of Dam<sup>4-HT-intein@L127C</sup> proteins in *Drosophila* we decided to rely on the basal activity of the full-length *hsp70* promoter. We inserted all Dam transgenes in the same genomic location (51C) by phiC31 integrase-mediated site-directed transgenesis (18). With this experimental setup, stable germline transformation of *Drosophila* with the Dam<sup>4-HT-intein@L127C</sup>-containing transgenes occurred with normal frequencies (Figure 2A). However, we detected Dam methylation in the transgenic animals grown in the absence of 4-HT, meaning that Dam<sup>4-HT-intein@L127C</sup> is only partially inactivated by the intein (Figure 2B). In some tissues in the absence of 4-HT we also observed a clear banded pattern in the amplified methylated DNA (Figure 2B). Sequencing of this DNA revealed that these discrete bands represent mitochondrial DNA. This suggests that the Dam<sup>4-HT-intein@L127C</sup> is par-

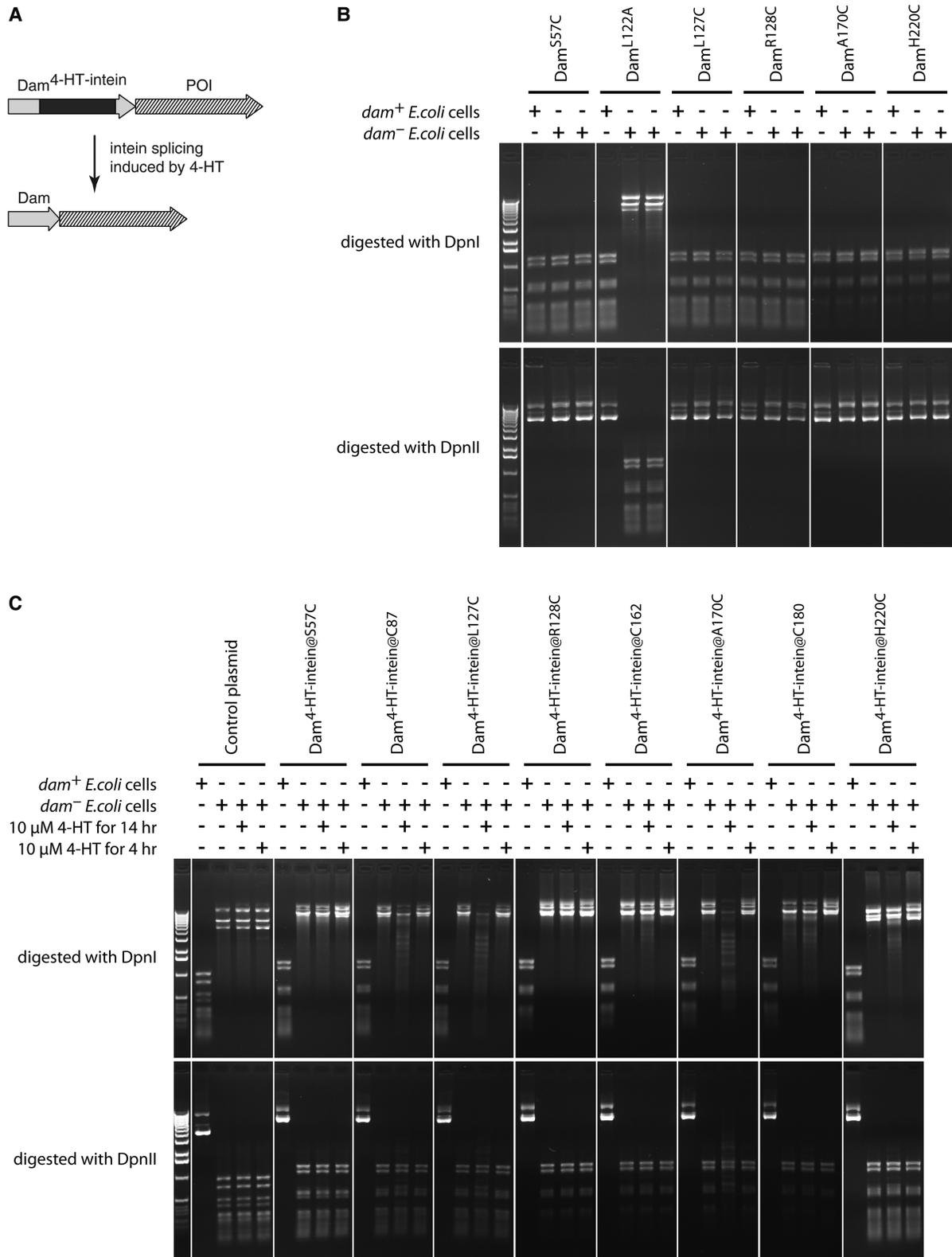
tially targeted to mitochondria, for which we do not have an explanation. Because mitochondrial DNA sequences are easily discerned from nuclear DNA sequences in DamID-seq, this does not substantially compromise the mapping. Induction of 4-HT-intein splicing by growing animals on fly food supplemented with 4-HT increases the activity of Dam and, as a result, the methylation of nuclear DNA, which in its turn reduces the proportion of mitochondrial DNA sequences in DamID samples (Figure 2B). Notably, mRNA sequencing of fat body tissues (two replicate experiments) suggested that the effects of 4-HT on gene expression are mild (Figure 2C and D).

To test whether Dam<sup>4-HT-intein@L127C</sup> can be used to generate a genome-wide DamID profile, we generated DamID-seq maps for Dam<sup>4-HT-intein@L127C</sup>-Polycomb (Dam<sup>4-HT-intein@L127C</sup>-PC) as well as unfused Dam<sup>4-HT-intein@L127C</sup> from dissected larval central brain and fat body tissue. These maps show the typical domain-like pattern of PC binding, similar to what we have previously observed in Kc cells (Figure 3A). Remarkably, the Dam<sup>4-HT-intein@L127C</sup>-PC DamID profile in larval central brain reproduced very well the profile generated previously by the original DamID approach (26) (Figure 3B, C) verifying the usability of the Dam<sup>4-HT-intein@L127C</sup> system. Thus, we constructed a ligand-inducible system for DamID profiling that appears to overcome toxicity problems, enables the tuning of Dam activity in *Drosophila* tissues, and can be used to generate DamID profiles from dissected tissues.

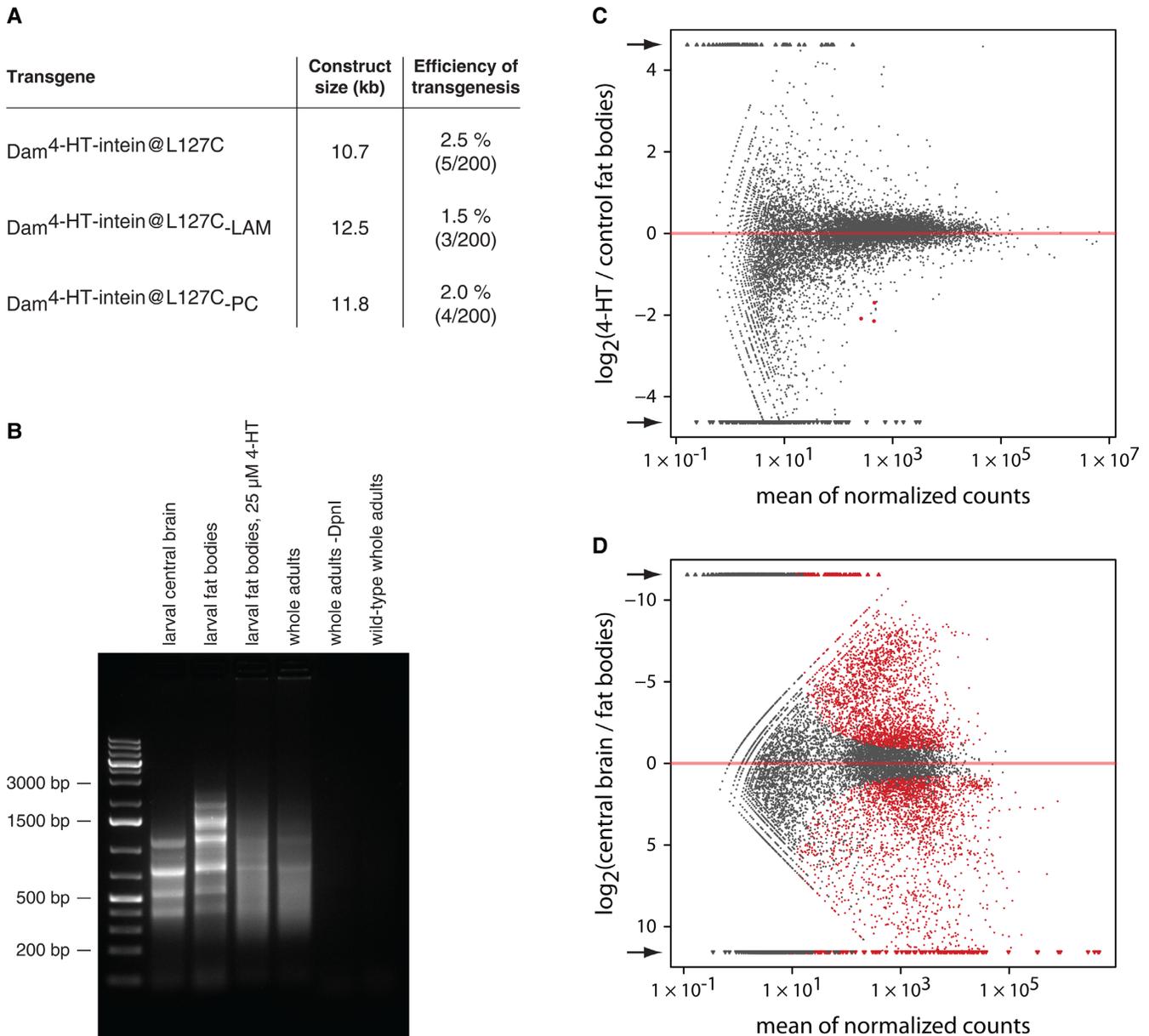
### FLP-inducible Dam vectors for tissue-specific DamID profiling in *Drosophila*

Next, we constructed a system for cell type specific expression of Dam proteins. We placed a cassette consisting of a transcription terminator flanked by two directly repeated DNA recombinase recognition sites, between the promoter and the Dam protein coding region. This should prevent expression of Dam unless the cassette is excised by expression of the recombinase (Figure 4A). Thus, cell type-specific expression of the recombinase enables the generation of the corresponding DamID profiles using whole tissue as input.

We chose the FLP/FRT recombination system, which has been extensively used in *Drosophila* (27,28). Since even trace amounts of Dam can lead to substantial background methylation of the genome, the terminator should prevent transcription of the transgene coding region extremely effectively. The sequences previously successfully used as terminators in *Drosophila* transgenic constructs were mostly used in combination with genes encoding proteins without enzymatic activities (for example, GFP). Therefore, we first tested five candidate cassettes named STOP#1 through #5 (Table 1) for their ability to effectively prevent transcription of the Dam-containing transgene. We inserted each of the five cassettes, flanked by FRT sites, between the *hsp70* promoter and an open reading frame encoding either Dam alone or a fusion protein of Dam and Lamin Dm0 (Dam-LAM). We chose LAM as Dam fusion partner because it is particularly suited to study genome—nuclear lamina interactions (29,30), which can undergo profound changes during differentiation (31).



**Figure 1.** Ligand-controlled approach to regulate Dam activity. (A) The principle of the approach. Insertion of the 4-HT-intein in the middle of the Dam disrupts the enzyme activity, which can be restored in the presence of 4-HT. POI, protein of interest. (B) Effect of cysteine substitutions on Dam activity in *E. coli* cells. *dam*<sup>+</sup> and *dam*<sup>-</sup> (in duplicate) *E. coli* cells were transformed with plasmids expressing the indicated Dam point mutants. After overnight culturing, the plasmids were purified and digested with DpnI and DpnII restriction enzymes that cut only G<sup>m</sup>6ATC and unmethylated GATC motifs, respectively. L122A is a mutation known to reduce Dam activity (50) and serves as a control. (C) Some Dam-4-HT-intein fusions exhibit inducible enzyme activity in bacterial cells. *dam*<sup>+</sup> and *dam*<sup>-</sup> *E. coli* cells were transformed with plasmids encoding the indicated Dam-4-HT-intein fusions. To activate the intein splicing, 4-HT was added to bacterial cultures for 4 or 14 hr. A plasmid without Dam coding sequence was used as a negative control.

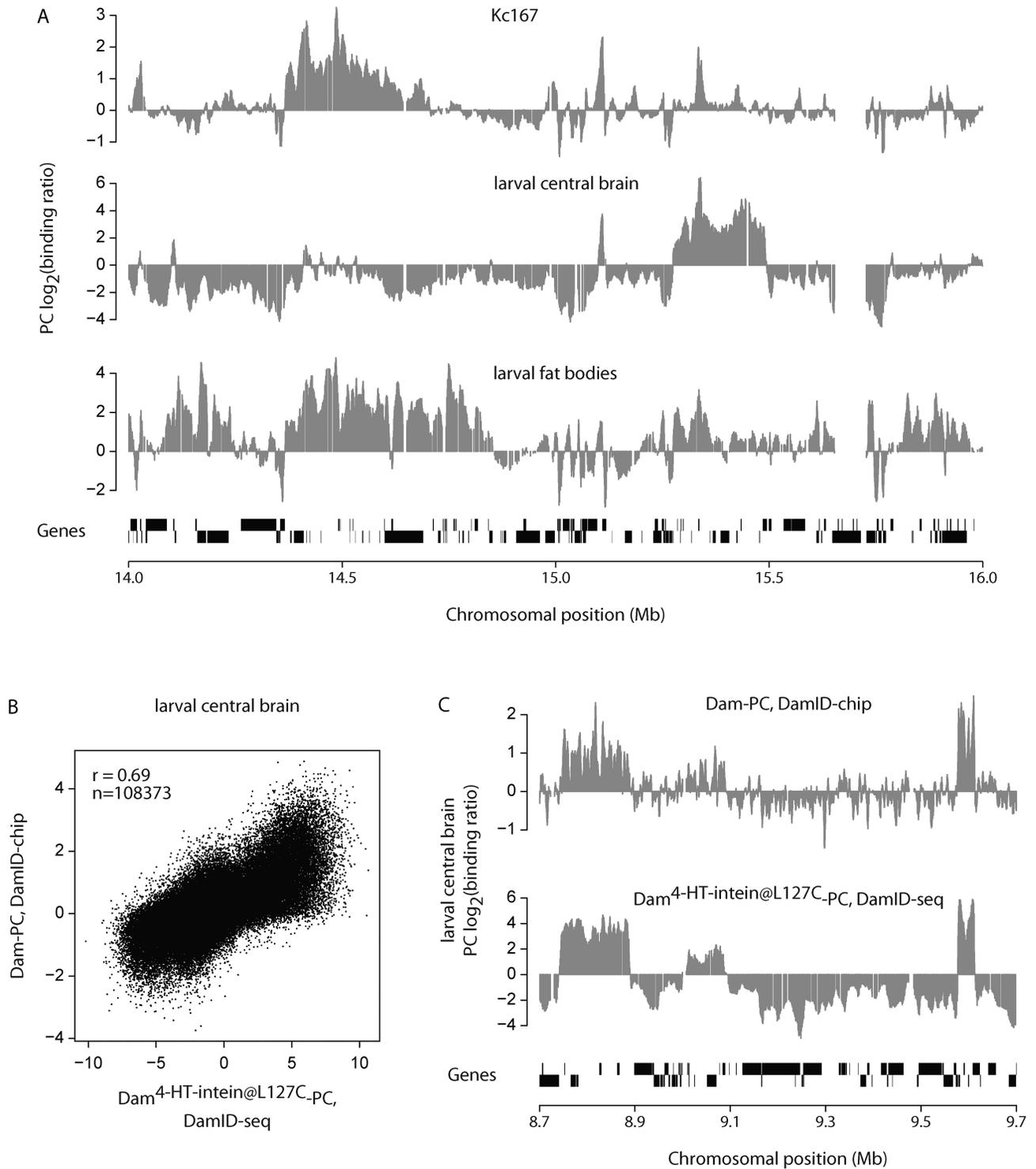


**Figure 2.** 4-HT-regulated *Drosophila* DamID system. (A) Transgenesis efficiency with Dam<sup>4</sup>-HT-intein@L127C-containing constructs. All transgenes were integrated at the same locus. For each construct, 200 embryos were injected, but no more than five independent transformants were collected, meaning that the efficiency of transgenesis could be even higher than 2.5% in the case of Dam<sup>4</sup>-HT-intein@L127C. LAM, Lamin Dm0; PC, Polycomb. (B) Bulk methylation by Dam<sup>4</sup>-HT-intein@L127C-LAM in dissected larval central brain, larval fat bodies and whole animals, detected by digestion with the G<sup>m6</sup>ATC-specific restriction endonuclease DpnI followed by ligation-mediated PCR (22). The presence of methylation in the absence of 4-HT indicates that 4-HT-intein has some background activity and that Dam<sup>4</sup>-HT-intein@L127C behaves as a hypomorphic Dam mutant. No signal is observed when DpnI is omitted (-DpnI) or in non-transgenic (wild-type) control flies. (C) Comparison of gene expression between fat bodies of larvae grown on regular fly food (control) and food supplemented with 25  $\mu$ M 4-HT. Only three genes are differentially expressed ( $P$  value adjusted for multiple testing with the Benjamini–Hochberg procedure  $< 0.05$ ; highlighted in red) according to the analysis with the R package DESeq (20). Arrows indicate data points outside the graph area, which were clipped. (D) Comparison of gene expression between larval fat bodies and larval central brain. 5328 genes are differentially expressed (adjusted  $p$  value  $< 0.05$ ).

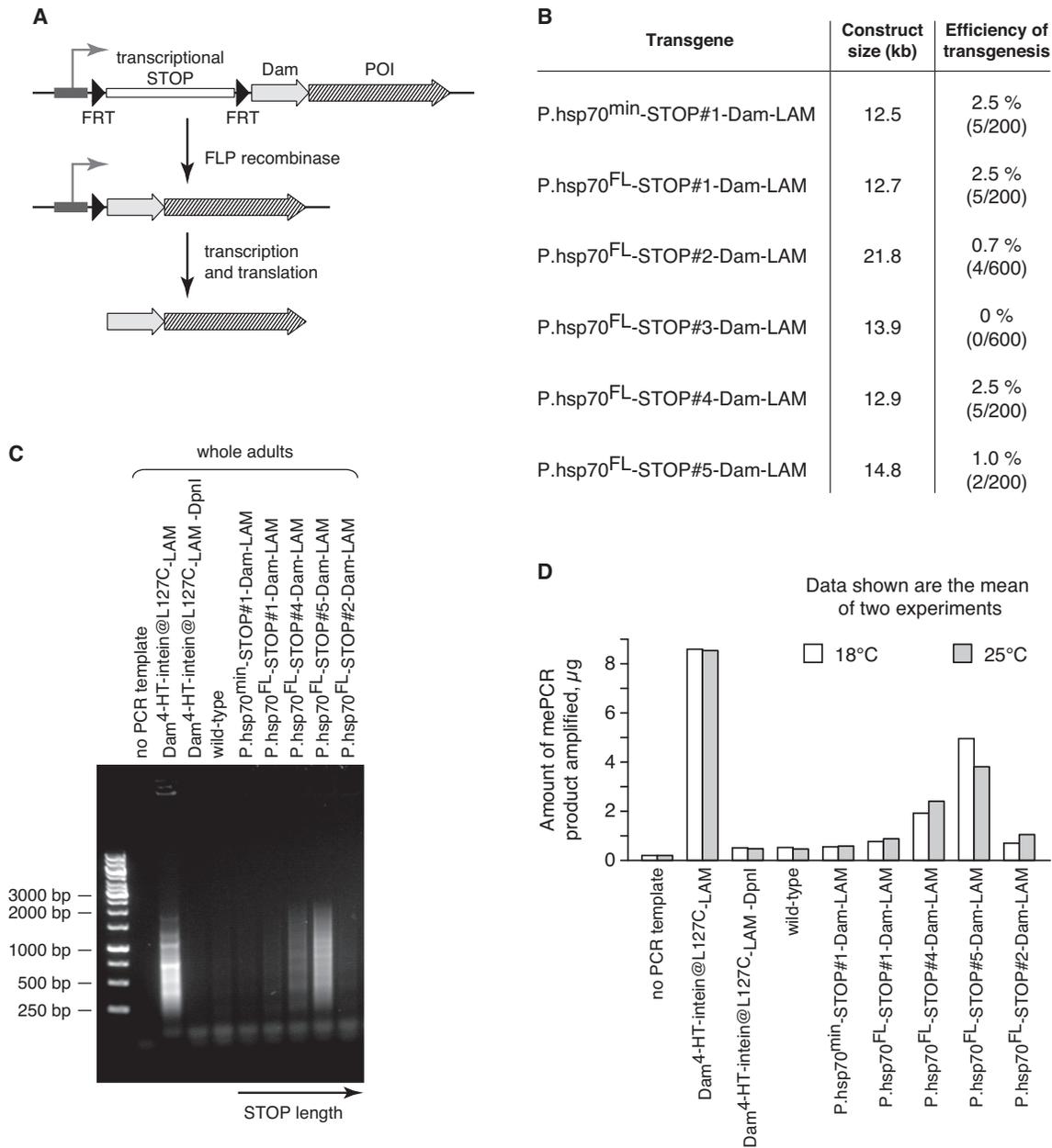
Already at the step of *Drosophila* transgenesis we observed significant differences between the terminator cassettes (Figure 4B). Efficient transgenesis was obtained only for STOP#1- and STOP#4-containing Dam-LAM constructs, indicating that other terminators might allow transcriptional read-through to some extent, causing toxicity. Indeed, the single fly line that we obtained with the

STOP#3-Dam transgene showed high background levels of adenine methylation in the genomic DNA (data not shown), whereas not a single fly line was established with the STOP#3-Dam-LAM transgene.

To directly test whether a STOP cassette can prevent toxicity we devised a viability assay with one of the STOP#2-Dam-LAM lines (Figure 5). Although the FLP-out effi-



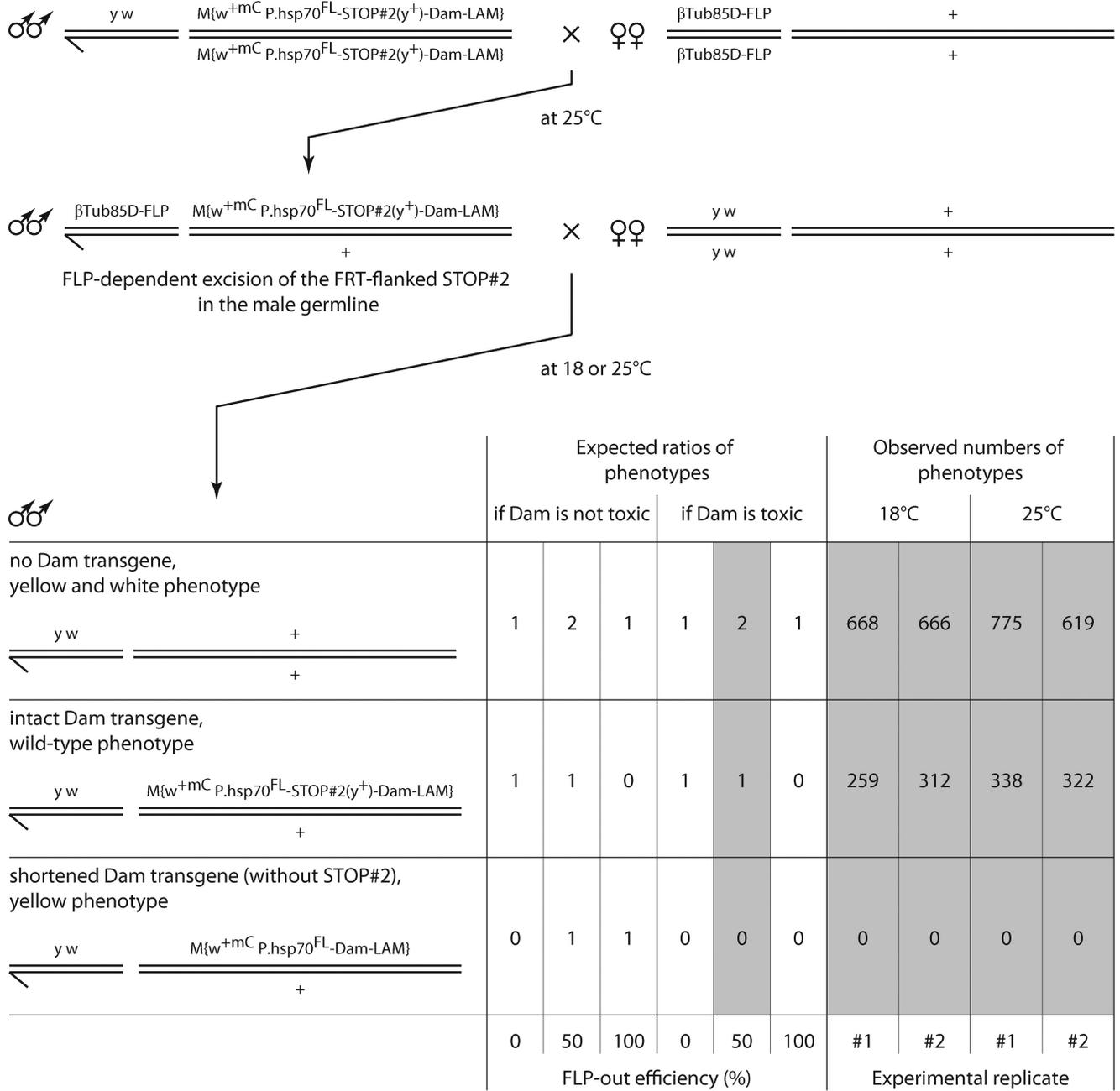
**Figure 3.** Dam<sup>4-HT-intein@L127C</sup>-PC DamID profiles in *Drosophila* tissues. (A) Comparison of PC DamID profiles in Kc cells, *Drosophila* larval central brain and fat bodies. A 2.0 Mb fragment of chromosomal arm 2L is shown. Data for Kc cells are from (10). (B and C) Comparison of conventional Dam-PC profile (26) and Dam<sup>4-HT-intein@L127C</sup>-PC DamID profile, both in larval central brain. (B) Genome-wide correlation between the data sets.  $r$ , Pearson's correlation coefficient;  $n$ , number of GATC fragments compared. (C) A 1.0 Mb fragment of chromosomal arm 2L is shown. The lower noise level and wider dynamic range of the Dam<sup>4-HT-intein@L127C</sup>-PC DamID profile might be due to different detection platform used (high-throughput sequencing instead of microarrays). In (A) and (C) a running mean algorithm (a sliding window of 10 GATC fragments, one GATC fragment per step) was applied to the PC binding data; genes are indicated by black rectangles.



**Figure 4.** FLP-inducible *Drosophila* DamID system. (A) Principle of the approach. Insertion of a transcriptional terminator (STOP) between the promoter and Dam protein coding sequence blocks the expression of the latter at the level of mRNA synthesis. FRT sites flanking the STOP allow its removal in the presence of FLP recombinase. POI, protein of interest. (B) Dam transgenes with different STOP sequences demonstrate diverse efficiency of *Drosophila* transgenesis. P.hsp70<sup>min</sup>, minimal *hsp70* promoter; P.hsp70<sup>FL</sup>, full-length *hsp70* promoter. All transgenes were integrated at the 51C locus by phiC31 integrase-mediated recombination. For each construct, at least 200 embryos were injected, but no more than 5 independent transformants were collected, meaning that the efficiency of transgenesis could be even higher than 2.5% in some cases. LAM, Lamin Dm0. (C) Methylation in whole adults of STOP-Dam-LAM transgenic flies grown at 18°C. Dam<sup>4</sup>-HT-intein@L127C-LAM flies were used as a positive control. (D) Quantification of amplified methylated GATC fragments from transgenic flies grown at 18 and 25°C.

**Table 1.** Tested terminator cassettes

Terminator cassette	Length (bp)	Elements	Reference
STOP#1	1317	C-terminal sequence of yeast <i>His3</i> gene, the SV40 polyadenylation signal region, a false translation initiation signal and a 5' splice donor site	(47)
STOP#2	10 361	inverted sequence of the <i>yellow</i> ( <i>y</i> <sup>+</sup> ) gene placed in the middle of the <i>hsp70</i> 3'UTR	(45)
STOP#3	2436	3'UTR of the <i>hsp70</i> gene alone	(45)
STOP#4	1522	SV40 and <i>Tub84B</i> polyadenylation signal regions	(48)
STOP#5	3425	<i>HcRed</i> gene and the <i>Glutamine synthetase 1</i> ( <i>gs1</i> ) 3'UTR	(49)



**Figure 5.** Test for the toxicity of the Dam-Lamin Dm0 (Dam-LAM) protein for *Drosophila*. First, flies bearing two transgenes,  $\beta$ Tub85D-FLP and P.hsp70<sup>FL</sup>-STOP#2-Dam-LAM, were obtained. The  $\beta$ Tub85D-FLP transgene is expressed during spermatogenesis (44). The STOP#2 cassette contains a functional *yellow* ( $y^+$ ) gene (45). In the progeny of males carrying the two transgenes, the following three classes were expected: [1] flies without Dam-LAM transgene, [2] flies with the intact (non-recombined) transgene and [3] flies with shortened (FRT-recombined) transgene. Male adults were counted on days 1–3 after eclosion. Experiments were performed at two different temperatures, 18°C and 25°C.  $w^{+mC}$ , mini-white gene. P.hsp70<sup>FL</sup>, full-length *hsp70* promoter.

ciency in the male germline was about two-fold lower than reported earlier (32), the genotype frequencies obtained were only compatible with a Dam toxicity scenario and prevention of this toxicity by the STOP#2 cassette. The relatively low level of transgenesis by the STOP#2-containing construct (Figure 4B) may have been caused by its relatively big size.

Next, we measured the background levels of DNA adenine methylation in adult transgenic flies. In animals with the STOP#1 and STOP#2 transgenes, leakage of Dam-LAM expression was minimal, since the amounts of amplified methylated fragments were only slightly higher compared to wild-type *Oregon* flies used as a negative control (Figure 4C and D). We detected substantially higher

DNA methylation levels in the STOP#4 and STOP#5 animals. Because all the transgenes were inserted in the same genomic location (51C) to avoid position effect, we conclude that the STOP#1 and STOP#2 sequences are the best transcriptional terminators in *Drosophila*. We chose the STOP#1 sequence for subsequent experiments because of its smaller size and high transgenesis efficiency.

Comparison of the STOP#1-Dam-LAM transgenes driven by the full-length and the minimal *hsp70* promoters revealed a negligible difference (Figure 4B and D). The minimal promoter does not respond to heat shock and therefore can be used in the combination with the *hsp70*-FLP transgene (33,34) to ubiquitously express Dam proteins at the low levels as an alternative to the inducible system based on the Dam<sup>4-HT-intein@L127C</sup>. Therefore, we finally selected the minimal *hsp70* promoter along with the STOP#1 sequence for FLP-inducible expression of Dam proteins in *Drosophila*.

### Proof-of-principle of cell type specific DamID profiling using STOP#1-Dam system

To test the utility of STOP#1-Dam system, we applied it to study the genome-wide binding profile of the Polycomb (PC) protein in glial cells, which constitute no more than 10% of all *Drosophila* larval brain cells (35). No Dam-dependent DNA methylation was detected in central brains dissected from STOP#1-Dam larvae, confirming that the STOP#1 sequence effectively terminates transcription initiated at the *hsp70* promoter. To activate expression of STOP#1-Dam constructs, we used the repo-FLP transgene that provides FLP activity exclusively in glial cells, but not in other cell types (19). Indeed, a substantial level of DNA methylation at GATC sites was observed in central brains of repo-FLP/STOP#1-Dam trans-heterozygous larvae (Figure 6A and B). Comparison of the PC DamID profile for repo<sup>+</sup> glial cells with that in whole larval central brain revealed many tissue-specific differences in addition to the overall similarity (Figure 6C). For example, we observed that the *gcm* gene—a glial lineage determination factor that is only expressed in glial cells and not in neurons (36)—is bound by PC in whole larval central brain, but not in repo<sup>+</sup> glial cells (Figure 6D).

For this experiment and other ones described above, we combined tissue material from about 20 larvae. With the aim to identify the minimum amount of input material required for the STOP#1-Dam system, we checked whether a tissue dissected from a single larva is enough to detect Dam-specific methylation occurring only in a small proportion of its cells. Analysis of several individual repo-FLP/STOP#1-Dam and repo-FLP/STOP#1-Dam-LAM trans-heterozygous larval central brains showed that methylated GATC fragments can be specifically amplified in each case (Figure 6E). This high sensitivity suggests that the STOP#1-Dam-containing transgene may be used to generate DamID profiles in small subsets of cells within a complex tissue.

## DISCUSSION

Here, we report two different systems to control the expression or activity of Dam-fusion proteins for DamID exper-

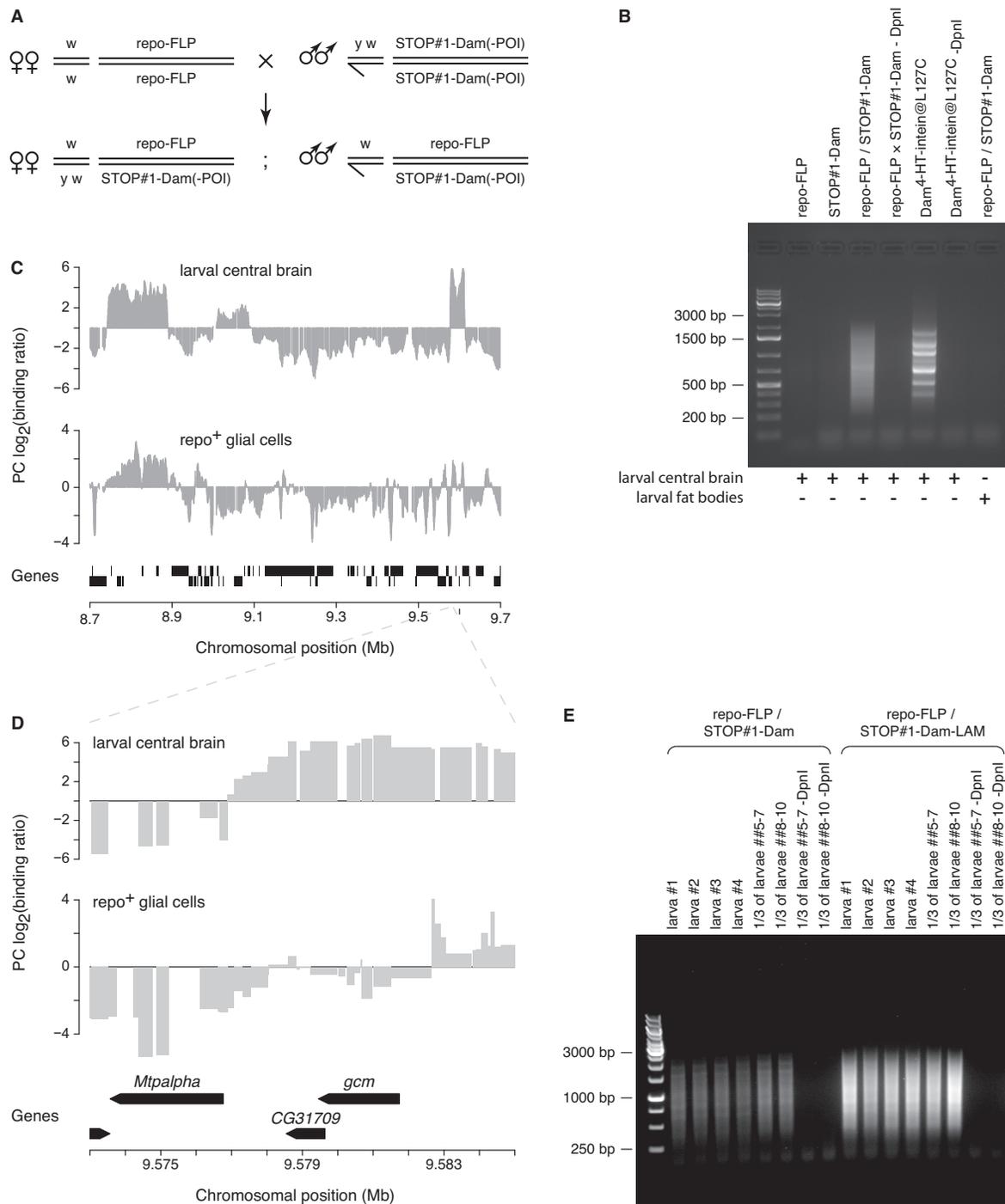
iments in *Drosophila*. The Dam<sup>4-HT-intein@L127C</sup> system may be used to fine-tune Dam activity by addition of 4-HT to fly food. It appears relatively leaky in the absence of 4-HT, which may be due to spontaneous excision of the intein. Improvements in inducible intein technology (37) may reduce this background further. Nevertheless, Dam<sup>4-HT-intein@L127C</sup> construct is useful because it substantially reduces toxicity problems and it allows for tuning of the Dam activity when needed.

We find that the STOP#1-Dam system is much more tightly controllable than the intein-based system; in our experiments, STOP#1-Dam exhibited essentially no genome methylation in the absence of FLP recombinase. This low leakiness combined with the extreme sensitivity and specificity of the protocol for amplification of Dam-methylated DNA (14,15; this study) creates opportunities to map chromatin protein binding in specific cell types in complex tissues. For this purpose many GAL4 driver lines are available to express FLP selectively in a tissue or cell type of choice. For example, hundreds of fly lines have been established that express FLP in specific subsets of brain cells (38). Unlike cell type specific chromatin immunoprecipitation methods (39,40) the STOP#1-Dam strategy does not require purification of the cell type of interest. We illustrated this here for glial cells, which constitute only ~10% of the fly brain.

Earlier reports also described FLP-inducible (41) and Cre-inducible (42) DamID systems employing respectively the terminator cassettes STOP#4 and STOP#3, but their utility for cell type specific DamID was not tested. Considering that STOP#3 and STOP#4 are relatively leaky in our hands, these termination cassettes may not be suited for this purpose. Southall and colleagues developed another system for inducible expression of Dam fusion proteins (14). Here, the expression vector consists of a GAL4-inducible promoter followed by open reading frame that encodes green fluorescent protein (GFP), and a second open reading frame that encodes the Dam fusion protein; expression of the latter is very low because it relies on inefficient ribosome reinitiation. As a proof-of-principle for this method, lineage-specific binding maps of RNA polymerase II in fly brain were produced (14). An additional opportunity for control of DamID activity is the use of a split-protein complementation version of DamID that was recently reported for mammalian cells (43). How these systems and ours compare in terms of cell type specificity of the DamID data (which is mostly determined by the degree of leaky expression in the non-target cells) may be tested in the future. Possibly these strategies may be combined to further reduce crosstalk from non-target cell types. Regardless, our data provide new evidence for the feasibility of cell type specific DamID mapping in small tissue samples. Our new vectors expand the suite of available tools for cell type specific studies of chromatin in complex tissues, which is a largely unexplored area of biology.

## AVAILABILITY

Vectors for transgenesis with Dam<sup>4-HT-intein@L127C</sup> and STOP#1-Dam proteins are available through Addgene (accession numbers 71805–71812). Fly lines and other plasmids are available upon request.



**Figure 6.** Lineage-specific activation of STOP#1-Dam-constructs. (A) Genetic crosses used to activate the STOP#1-Dam-containing transgenes in repo-positive glial cells. POI, protein of interest. (B) Methylation detected in genomic DNA isolated from dissected larval tissues. Specificity of the induction of the STOP#1-Dam transgene expression was confirmed by the presence of methylated DNA fragments only in central brain of repo-FLP/STOP#1-Dam, but not in repo-FLP and STOP#1-Dam larvae, nor in fat bodies of repo-FLP/STOP#1-Dam larvae, where the *repo* promoter is not active (46). Dam<sup>4-HT-intein@L127C</sup> larvae served as a positive control; the banded pattern is derived from mitochondrial DNA. (C and D) Polycomb (PC) DamID profiles in *Drosophila* whole larval central brain and repo-positive glial cells in the brain obtained using the Dam<sup>4-HT-intein@L127C</sup>- and STOP#1-Dam-containing transgenes, respectively. (C) A 1.0 Mb fragment of chromosomal arm 2L is shown. A running mean algorithm (sliding window of 10 GATC fragments, one GATC fragment per step) was applied to the PC binding data. Genes are indicated by black rectangles. (D) A region of the chromosomal arm 2L spanning the *glial cells missing* (*gcm*) gene is shown. Bar widths correspond to lengths of GATC fragments. Genes are indicated by black arrows. (E) DamID is sensitive enough for application to individual *Drosophila* larval central brains in which a STOP#1-Dam-containing transgene is expressed only in a subset of cells. Methylated GATC fragments amplified from repo-FLP/STOP#1-Dam and repo-FLP/STOP#1-Dam-LAM larval brains. In each case, either central brains from 4 individual larvae or from two sets of 3 larvae were dissected and used for the isolation of genomic DNA. Specificity of amplification of the methylated GATC fragments was confirmed by the ‘-Dpnl’ control reactions performed on one-third of DNA isolated from 3 larval central brain samples.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank T. Sixma for help with the selection of positions within Dam to insert the intein; V.V. Shloma, J. van Bommel and G. Filion for help with experiments and data analyses; S. de Vries and the NKI Genomics Core Facility for technical support; J. Kind and M. Amendola for helpful suggestions; B. Tolhuis and J. van Bommel for sharing unpublished results; D. Liu, G. Struhl, E. Marois, C. Klämbt and F. Karch for flies and plasmids.

## FUNDING

EURYI, NWO/ALW-VICI, the Netherlands Consortium for Systems Biology and ERC Advanced Grant 293662—CHROMATINPRINCIPLES to B.v.S.; Russian fundamental scientific research project 0310-2014-0002 to A.V.P. Funding for open access charge: ERC Advanced Grant [293662].

*Conflict of interest statement.* None declared.

## REFERENCES

- van Steensel, B. and Henikoff, S. (2000) Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat. Biotechnol.*, **18**, 424–428.
- van Steensel, B., Delrow, J. and Henikoff, S. (2001) Chromatin profiling using targeted DNA adenine methyltransferase. *Nat. Genet.*, **27**, 304–308.
- Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W. *et al.* (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, **453**, 948–951.
- Germann, S., Juul-Jensen, T., Letarnc, B. and Gaudin, V. (2006) DamID, a new tool for studying plant chromatin profiling in vivo, and its use to identify putative LHP1 target loci. *Plant J.*, **48**, 153–163.
- Venkatasubrahmanyam, S., Hwang, W.W., Meneghini, M.D., Tong, A.H. and Madhani, H.D. (2007) Genome-wide, as opposed to local, antisilencing is mediated redundantly by the euchromatic factors Set1 and H2AZ. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 16609–16614.
- Woolcock, K.J., Gaidatzis, D., Punga, T. and Buhler, M. (2011) Dicer associates with chromatin to repress genome activity in *Schizosaccharomyces pombe*. *Nat. Struct. Mol. Biol.*, **18**, 94–99.
- Towbin, B.D., Gonzalez-Aguilera, C., Sack, R., Gaidatzis, D., Kalck, V., Meister, P., Askjaer, P. and Gasser, S.M. (2012) Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell*, **150**, 934–947.
- Gonzalez-Aguilera, C., Ikegami, K., Ayuso, C., de Luis, A., Iniguez, M., Cabello, J., Lieb, J.D. and Askjaer, P. (2014) Genome-wide analysis links emerin to neuromuscular junction activity in *Caenorhabditis elegans*. *Genome Biol.*, **15**, R21.
- Schuster, E., McElwee, J.J., Tullet, J.M., Doonan, R., Matthijssens, F., Reece-Hoyes, J.S., Hope, I.A., Vanfleteren, J.R., Thornton, J.M. and Gems, D. (2010) DamID in *C. elegans* reveals longevity-associated targets of DAF-16/FoxO. *Mol. Syst. Biol.*, **6**, 399.
- Filion, G.J., van Bommel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J. *et al.* (2010) Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell*, **143**, 212–224.
- van Bommel, J.G., Filion, G.J., Rosado, A., Talhout, W., de Haas, M., van Welsem, T., van Leeuwen, F. and van Steensel, B. (2013) A network model of the molecular organization of chromatin in *Drosophila*. *Mol. Cell*, **49**, 759–771.
- Greil, F., Moorman, C. and van Steensel, B. (2006) DamID: mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase. *Methods Enzymol.*, **410**, 342–359.
- Vogel, M.J., Pagie, L., Talhout, W., Nieuwland, M., Kerkhoven, R.M. and van Steensel, B. (2009) High-resolution mapping of heterochromatin redistribution in a *Drosophila* position-effect variegation model. *Epigenet. Chromatin*, **2**, 1.
- Southall, T.D., Gold, K.S., Egger, B., Davidson, C.M., Caygill, E.E., Marshall, O.J. and Brand, A.H. (2013) Cell-type-specific profiling of gene expression and chromatin binding without cell isolation: assaying RNA Pol II occupancy in neural stem cells. *Dev. Cell*, **26**, 101–112.
- Kind, J., Pagie, L., de Vries, S.S., Nahidiazar, L., Dey, S.S., Bienko, M., Zhan, Y., Lajoie, B., de Graaf, C.A., Amendola, M. *et al.* (2015) Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. *Cell*, **163**, 134–147.
- Wines, D.R., Talbert, P.B., Clark, D.V. and Henikoff, S. (1996) Introduction of a DNA methyltransferase into *Drosophila* to probe chromatin structure in vivo. *Chromosoma*, **104**, 332–340.
- Choksi, S.P., Southall, T.D., Bossing, T., Edoff, K., de Wit, E., Fischer, B.E., van Steensel, B., Micklem, G. and Brand, A.H. (2006) Prospero acts as a binary switch between self-renewal and differentiation in *Drosophila* neural stem cells. *Dev. Cell*, **11**, 775–789.
- Bischof, J., Maeda, R.K., Hediger, M., Karch, F. and Basler, K. (2007) An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 3312–3317.
- Silies, M., Yuva, Y., Engelen, D., Aho, A., Stork, T. and Klämbt, C. (2007) Glial cell migration in the eye disc. *J. Neurosci.*, **27**, 13130–13139.
- Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
- Anders, S., McCarthy, D.J., Chen, Y., Okoniewski, M., Smyth, G.K., Huber, W. and Robinson, M.D. (2013) Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat. Protoc.*, **8**, 1765–1786.
- Vogel, M.J., Peric-Hupkes, D. and van Steensel, B. (2007) Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nat. Protoc.*, **2**, 1467–1478.
- Buskirk, A.R., Ong, Y.C., Gartner, Z.J. and Liu, D.R. (2004) Directed evolution of ligand dependence: small-molecule-activated protein splicing. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 10505–10510.
- Skretas, G. and Wood, D.W. (2005) Regulation of protein activity with small-molecule-controlled inteins. *Protein Sci.*, **14**, 523–532.
- Noren, C.J., Wang, J. and Perler, F.B. (2000) Dissecting the chemistry of protein splicing and its applications. *Angew. Chem. Int. Ed. Engl.*, **39**, 450–466.
- Tolhuis, B., Blom, M., Kerkhoven, R.M., Pagie, L., Teunissen, H., Nieuwland, M., Simonis, M., de Laat, W., van Lohuizen, M. and van Steensel, B. (2011) Interactions among Polycomb domains are guided by chromosome architecture. *PLoS Genet.*, **7**, e1001343.
- Theodosiou, N.A. and Xu, T. (1998) Use of FLP/FRT system to study *Drosophila* development. *Methods*, **14**, 355–365.
- Kao, C.F. and Lee, T. (2012) In vivo single cell labeling techniques. *The Making and Un-making of Neuronal Circuits in Drosophila*. doi:10.1007/978-1-61779-830-6.4.
- Pickersgill, H., Kalverda, B., de Wit, E., Talhout, W., Fornerod, M. and van Steensel, B. (2006) Characterization of the *Drosophila melanogaster* genome at the nuclear lamina. *Nat. Genet.*, **38**, 1005–1014.
- van Bommel, J.G., Pagie, L., Braunschweig, U., Brugman, W., Meuleman, W., Kerkhoven, R.M. and van Steensel, B. (2010) The insulator protein SU(HW) fine-tunes nuclear lamina interactions of the *Drosophila* genome. *PLoS One*, **5**, e15013.
- Peric-Hupkes, D., Meuleman, W., Pagie, L., Bruggeman, S.W., Solovei, I., Brugman, W., Graf, S., Flicek, P., Kerkhoven, R.M., van Lohuizen, M. *et al.* (2010) Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell*, **38**, 603–613.
- Golic, M.M., Rong, Y.S., Petersen, R.B., Lindquist, S.L. and Golic, K.G. (1997) FLP-mediated DNA mobilization to specific target sites in *Drosophila* chromosomes. *Nucleic Acids Res.*, **25**, 3665–3671.

33. Golic, K.G. and Lindquist, S. (1989) The FLP recombinase of yeast catalyzes site-specific recombination in the *Drosophila* genome. *Cell*, **59**, 499–509.
34. Golic, K.G. and Golic, M.M. (1996) Engineering the *Drosophila* genome: chromosome rearrangements by design. *Genetics*, **144**, 1693–1711.
35. Stork, T., Engelen, D., Krudewig, A., Silies, M., Bainton, R.J. and Klämbt, C. (2008) Organization and function of the blood-brain barrier in *Drosophila*. *J. Neurosci.*, **28**, 587–597.
36. Jones, B.W. (2005) Transcriptional control of glial cell development in *Drosophila*. *Dev. Biol.*, **278**, 265–273.
37. Peck, S.H., Chen, I. and Liu, D.R. (2011) Directed evolution of a small-molecule-triggered intein with improved splicing properties in mammalian cells. *Chem. Biol.*, **18**, 619–630.
38. Bohm, R.A., Welch, W.P., Goodnight, L.K., Cox, L.W., Henry, L.G., Gunter, T.C., Bao, H. and Zhang, B. (2010) A genetic mosaic approach for neural circuit mapping in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 16378–16383.
39. Deal, R.B. and Henikoff, S. (2010) A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev. Cell*, **18**, 1030–1040.
40. Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczynski, B., Riddell, A. and Furlong, E.E. (2012) Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.*, **44**, 148–156.
41. Luo, S.D., Shi, G.W. and Baker, B.S. (2011) Direct targets of the *D. melanogaster* DSXF protein and the evolution of sexual development. *Development*, **138**, 2761–2771.
42. Maksimov, D.A., Koryakov, D.E. and Belyakin, S.N. (2014) Developmental variation of the SUUR protein binding correlates with gene regulation and specific chromatin types in *D. melanogaster*. *Chromosoma*, **123**, 253–264.
43. Hass, M.R., Liow, H.H., Chen, X., Sharma, A., Inoue, Y.U., Inoue, T., Reeb, A., Martens, A., Fulbright, M., Raju, S. *et al.* (2015) SpDamID: marking DNA bound by protein complexes identifies notch-dimer responsive enhancers. *Mol. Cell*, **59**, 685–697.
44. Struhl, G., Fitzgerald, K. and Greenwald, I. (1993) Intrinsic activity of the Lin-12 and Notch intracellular domains in vivo. *Cell*, **74**, 331–345.
45. Struhl, G. and Basler, K. (1993) Organizing activity of wingless protein in *Drosophila*. *Cell*, **72**, 527–540.
46. Chintapalli, V.R., Wang, J. and Dow, J.A. (2007) Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.*, **39**, 715–720.
47. Lakso, M., Sauer, B., Mosinger, B. Jr, Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L. and Westphal, H. (1992) Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 6232–6236.
48. Stockinger, P., Kvitsiani, D., Rotkopf, S., Tirian, L. and Dickson, B.J. (2005) Neural circuitry that governs *Drosophila* male courtship behavior. *Cell*, **121**, 795–807.
49. Marois, E. and Eaton, S. (2007) RNAi in the Hedgehog signaling pathway: pFRiPE, a vector for temporally and spatially controlled RNAi in *Drosophila*. *Methods Mol. Biol.*, **397**, 115–128.
50. Horton, J.R., Liebert, K., Bekes, M., Jeltsch, A. and Cheng, X. (2006) Structure and substrate recognition of the *Escherichia coli* DNA adenine methyltransferase. *J. Mol. Biol.*, **358**, 559–570.