

# Genetic and Transgenic Reagents for *Drosophila simulans*, *D. mauritiana*, *D. yakuba*, *D. santomea*, and *D. virilis*

David L. Stern,<sup>\*,1</sup> Justin Crocker,<sup>\*</sup> Yun Ding,<sup>\*</sup> Nicolas Frankel,<sup>†</sup> Gretchen Kappes,<sup>\*\*</sup> Elizabeth Kim,<sup>\*</sup> Ryan Kuzmickas,<sup>‡</sup> Andrew Lemire,<sup>\*</sup> Joshua D. Mast,<sup>§</sup> and Serge Picard<sup>††</sup>

<sup>\*</sup>Janelia Research Campus, Ashburn, Virginia 20147, <sup>†</sup>Departamento de Ecología, Genética y Evolución, Instituto de Ecología, Genética y Evolución de Buenos Aires-Consejo Nacional de Investigaciones Científicas y Técnicas, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, C1428EHA Argentina, <sup>‡</sup>Genetics Division, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02115, <sup>§</sup>Department of Biological Sciences, California State University East Bay, Hayward, California 94542, <sup>\*\*</sup>Department of Ecology and Evolutionary Biology, and <sup>††</sup>Lewis-Sigler Institute for Integrative Genomics, Carl Icahn Laboratory, Princeton University, New Jersey 08544

ORCID ID: 0000-0002-1847-6483 (D.L.S.)

**ABSTRACT** Species of the *Drosophila melanogaster* species subgroup, including the species *D. simulans*, *D. mauritiana*, *D. yakuba*, and *D. santomea*, have long served as model systems for studying evolution. However, studies in these species have been limited by a paucity of genetic and transgenic reagents. Here, we describe a collection of transgenic and genetic strains generated to facilitate genetic studies within and between these species. We have generated many strains of each species containing mapped *piggyBac* transposons including an *enhanced yellow fluorescent protein (EYFP)* gene expressed in the eyes and a  $\phi$ C31 *attP* site-specific integration site. We have tested a subset of these lines for integration efficiency and reporter gene expression levels. We have also generated a smaller collection of other lines expressing other genetically encoded fluorescent molecules in the eyes and a number of other transgenic reagents that will be useful for functional studies in these species. In addition, we have mapped the insertion locations of 58 transposable elements in *D. virilis* that will be useful for genetic mapping studies.

## KEYWORDS

*Drosophila*  
genetics  
transgenics  
 $\phi$ C31 integrase  
speciation  
evolution

Ever since A. Sturtevant discovered *Drosophila simulans*, the sister species to *D. melanogaster*, in 1919, species of the *D. melanogaster* species subgroup have played a central role in studies of evolution and speciation (Powell 1997; Barbash 2010). Most species of the subgroup display superficially similar anatomy, although all species can be distinguished by both qualitative and quantitative anatomical differences (Orgogozo and Stern 2009). In addition, the species display

enormous variation in ecology and behavior, with some having evolved into ecological specialists on unusual food sources (R'Kha *et al.* 1991; Yassin *et al.* 2016).

One of the major advantages of this subgroup for evolutionary studies is that many of the species can be crossed to *D. melanogaster* to generate sterile hybrids and some can be crossed to each other to generate fertile hybrid females (Powell 1997). An unusual and important feature of these fertile pairs is that strains of each species can be found that share synteny across all chromosomes (Lemeunier and Ashburner 1976; Moehring *et al.* 2006a). This allows comprehensive genetic interrogation of the entire genome through recombination mapping. This is an uncommon feature for fertile pairs of *Drosophila* species; most species that have been examined exhibit major chromosomal inversions that are fixed between species (Powell 1997).

The combination of relatively straightforward genetics with diversity in anatomy, physiology, and behavior has encouraged many groups to perform genetic analyses of these species (*e.g.*, Liu *et al.* 1996; True *et al.* 1997; Macdonald and Goldstein 1999; Gleason and Ritchie 2004; Moehring *et al.* 2004, 2006a,b; Carbone *et al.* 2005; Gleason *et al.*

Copyright © 2017 Stern *et al.*

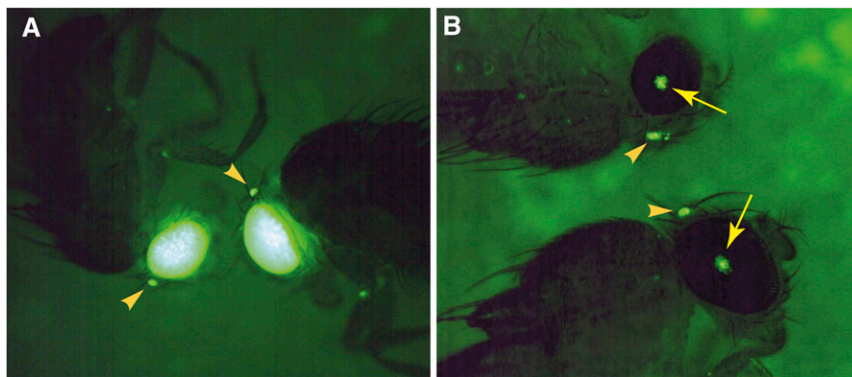
doi: <https://doi.org/10.1534/g3.116.038885>

Manuscript received December 28, 2016; accepted for publication February 21, 2017; published Early Online March 7, 2017.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental material is available online at [www.g3journal.org/lookup/suppl/doi:10.1534/g3.116.038885/-/DC1](http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.116.038885/-/DC1).

<sup>1</sup>Corresponding author: Janelia Research Campus, 19700 Helix Drive, Ashburn, VA 20147. E-mail: [sternd@janelia.hhmi.org](mailto:sternd@janelia.hhmi.org)



**Figure 1** Appearance of enhanced yellow fluorescent protein (EYFP) fluorescence in fly eyes. (A) In flies carrying a  $w^-$  mutation, fluorescence is often intense and observable throughout the compound eye and in the ocelli (arrowheads). (B) In flies carrying wild-type eye coloration, fluorescence is observed in the compound eye as small dots including  $\sim 10$  ommatidia (arrows) and in the ocelli (arrowheads).

2005; Orgogozo *et al.* 2006; Cande *et al.* 2012; Arif *et al.* 2013; Peluffo *et al.* 2015). However, in the vast majority of cases, these studies have stopped after quantitative trait locus mapping of traits of interest. One factor that has limited further genetic study of these traits is a limited set of genetic markers, which can facilitate fine-scale mapping. J. True and C. Laurie established a large collection of strains carrying  $P$ -element transposons marked with a  $w^+$  mini-gene in a  $w^-$  background of *D. mauritiana* (True *et al.* 1996a,b). These have been used for introgression studies (True *et al.* 1996b; Coyne and Charlesworth 1997; Tao *et al.* 2003a,b; Masly and Presgraves 2007; Masly *et al.* 2011; Arif *et al.* 2013; Tanaka *et al.* 2015; Tang and Presgraves 2015) and for high-resolution mapping studies (McGregor *et al.* 2007; Araripe *et al.* 2010), demonstrating the utility of dominant genetic markers for evolutionary studies. One limitation of these strains is that the  $w^+$  marker is known to induce behavioral artifacts (Zhang and Odenwald 1995; Campbell and Nash 2001; Xiao and Robertson 2016). We have also observed that mutations in the *white* gene and some  $w^+$  rescue constructs cause males to generate abnormal courtship song (Y. Ding and D. Stern, unpublished data). Other pigmentation genes that are commonly used in *D. melanogaster* are also known to disrupt normal behavior (Bastock 1956; Kyriacou *et al.* 1978; Drapeau *et al.* 2006; Suh and Jackson 2007); therefore, it would be preferable to employ dominant genetic markers that do not interfere with normal eye color or pigmentation.

We were motivated by the phenotypic variability and genetic accessibility of these species to establish a set of reagents that would allow, simultaneously, a platform for site-specific transgenesis (Groth *et al.* 2004) and reagents useful for genetic mapping studies. Therefore, we set out to establish a collection of strains carrying transposable elements marked with innocuous dominant markers for four of the most commonly studied species of the *D. melanogaster* species subgroup: *D. simulans*, *D. mauritiana*, *D. yakuba*, and *D. santomea*. We chose the *piggyBac* transposable element to minimize bias of insertion sites relative to gene start sites (Thibault *et al.* 2004) and integrated transposable elements carrying *EYFP* and *DsRed* driven by a *3XP3* enhancer, which is designed to drive expression in the eyes (Horn *et al.* 2003). A large subset of the lines described here also include a  $\phi$ C31 *attP* landing site to facilitate site-specific transgene integration. Here, we describe the establishment and mapping of many lines of each species carrying *pBac{3XP3::EYFP,attP}* and *pBac{3XP3::DsRed}* (Horn *et al.* 2003). We have characterized a subset of the *pBac{3XP3::EYFP, attP}* lines from each species for  $\phi$ C31 integration efficiency of plasmids containing an *attB* sequence. In addition, we have integrated transgenes carrying the *even-skipped* stripe 2 enhancer to characterize embryonic expression generated by a subset of *attP* landing sites. We have employed CRISPR/Cas9 to knock out the *3XP3::EYFP* gene in a subset of lines to facilitate integration of reagents for neurogenetics. We

also describe several other genetic and transgenic reagents that may be useful to the community, including the map positions for *pBac* transposons integrated in the *D. virilis* genome.

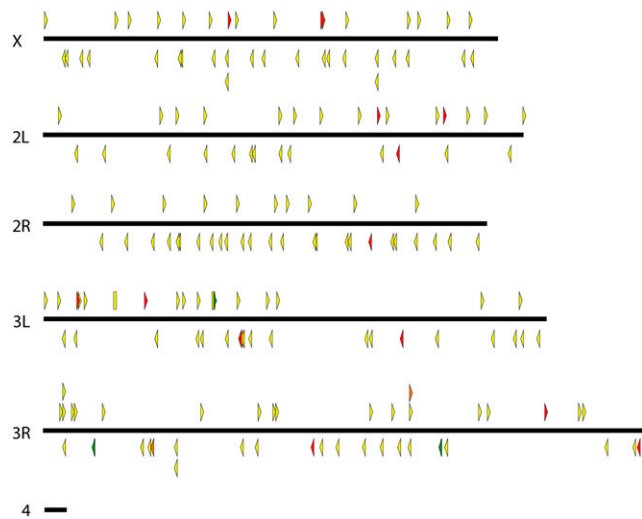
## MATERIALS AND METHODS

### Transposable elements employed

We used *piggyBac* transposable elements (Horn *et al.* 2003) to mobilize markers to random locations within the genomes of *D. simulans white* [501] (San Diego Species Stock Center stock number 14021-0251.011), *D. simulans yellow*[1] *white*[1] (San Diego Species Stock Center stock number 14021-0251.013), *D. mauritiana white*<sup>-</sup> (San Diego Species Stock Center stock number 14021-0241.60), *D. yakuba white*<sup>-</sup> (San Diego Species Stock Center stock number 14021-0261.02), *D. santomea* STO CAGO 1482 (provided by P. Andolfatto), and *D. virilis w*[50112] (San Diego Species Stock Center number 15010-1051.53). We constructed *pBac{3XP3::EYFP-attP}* by cloning a *Bgl*II fragment containing the *attP* site from *pM{3XP3-RFPattP}* (Bischof *et al.* 2007) into the single *Bgl*II site of *pBac{3XP3::EYFPafm}* (Horn and Wimmer 2000).

We constructed *pBac* plasmids carrying a source of  $P$ -element transposase marked with *3XP3::EYFP* or *3XP3::DsRed* as follows. We digested the plasmid pACNNTNPII-S129A (Beall *et al.* 2002) with *Eco*RI and *Not*I and cloned the  $\sim 5$  kb fragment resulting from digestion into pSLFa1180fa (Horn and Wimmer 2000). This plasmid was digested with *Asc*I or *Fse*I and the  $\sim 5$  kb fragment was cloned into the *Asc*I or *Fse*I restriction sites of *pBac{3XP3::DsRed}* or *pBac{3XP3::EGFP,attP}* (Horn and Wimmer 2000) to generate *pBac{Pactin::Ptrsps, 3XP3::DsRed}* and *pBac{Pactin::Ptrsps 3XP3::EGFP,attP}*, respectively. These plasmids were injected into strains of *D. simulans* and *D. mauritiana*. We also injected *pBac{3XP3::DsRed}* (Horn *et al.* 2003) into strains of *D. simulans*, *D. mauritiana*, *D. yakuba*, and *D. santomea*. The complete sequences of *pBac{3XP3::EYFP-attP}*, *pBac{3XP3::DsRed}*, and *phsp-pBac* are provided as Supplemental Material, File S2. These plasmids were co-injected with 250 ng/ $\mu$ l *phsp-pBac* (Handler and Harrell 1999), a heat shock-inducible source of *piggyBac* transposase, and 1 hr after injection embryos were heat shocked at 37° for 1 hr. All embryo injections were performed by Rainbow Transgenic Flies Inc. G0 flies were backcrossed to uninjected flies of the same strain and G1 flies were screened for fluorescence in their eyes.

Fluorescence could be detected easily in the compound eyes and ommatidia in all of the *white*<sup>-</sup> strains (*D. simulans*, *D. mauritiana*, *D. yakuba*, and *D. virilis*) using any dissecting microscope we tried with epi-fluorescence capability (Figure 1A). In flies with wild-type eye coloration, fluorescence in the compound eye is limited to a small spot of  $\sim 10$  ommatidia (Figure 1B). However, we found that fluorescence was very weak, and usually unobservable, in the eyes of flies with wild-type eye coloration using a Leica 165 FC stereomicroscope. This microscope

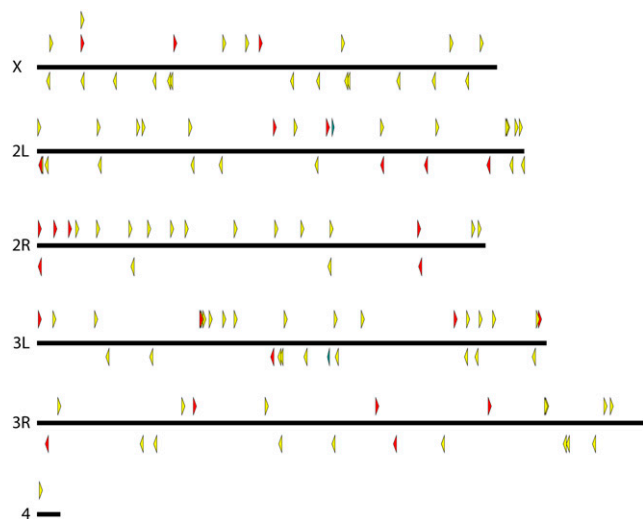


**Figure 2** Genomic insertion sites of *pBac* transposable elements in *D. simulans*. Each triangle represents a unique *pBac* element insertion. Some strains carry multiple insertion events. Some insertion sites are present in multiple strains, at least one of which contains multiple insertions. These strains were maintained to maximize the diversity of insertion sites in the collection. *pBac* insertions oriented forward are indicated above each chromosome and point to the right while reverse insertions are indicated below each chromosome and point to the left. Rectangles represent inserted elements whose orientation could not be determined. Yellow, green, and red indicate elements carrying *3XP3::EYFP*, *3XP3::EGFP*, and *3XP3::DsRed*, respectively.

uses “TripleBeam Technology” to deliver excitation light along a separate light path from the emission light. Unfortunately, the excitation light in this system appears to illuminate ommatidia adjacent to the ommatidia that are viewed for the emission light. Fluorescence can still be detected in the ocelli of these flies with this microscope, although this requires a bit more patience than when using a standard epi-fluorescence microscope to screen for fluorescence in the compound eyes.

### Mapping of transposable element insertion sites

We mapped the genomic insertion sites of all *pBac* elements using both inverse PCR (iPCR) (Ochman *et al.* 1988) and TagMap (Stern 2016). iPCR was not ideal for our project for several reasons. First, many isolated strains appeared to contain multiple insertion events, even though they were isolated from single G0 animals. These multiple events could sometimes be detected by segregation of offspring with multiple strengths of fluorescence in the eyes. In these cases, iPCR sometimes produced uninterpretable sequences and occasionally only a single insertion event was amplified. Second, many iPCR sequences were too short to allow unambiguous mapping to the genome. Third, sometimes iPCR reactions failed for no obvious reason. For all of these reasons, it was difficult to unambiguously map all of the *pBac* insertions with iPCR. Therefore, we developed and applied TagMap (Stern 2016) to map the insertion positions of all *pBac* elements. TagMap combines genome fragmentation and tagging using Tn5 transposase with a selective PCR to amplify sequences flanking a region of interest. This method provides high-throughput, accurate mapping of transposon insertions. Tagmap provided transposon insertion positions for all but a few strains. Transposable element insertion sites in the *D. simulans* and *D. mauritiana* strains were mapped to *D. simulans*



**Figure 3** Genomic insertion sites of *pBac* transposable elements in *D. mauritiana*. *pBac* insertions oriented forward are indicated above each chromosome and point to the right while reverse insertions are indicated below each chromosome and point to the left. Rectangles represent inserted elements whose orientation could not be determined. Yellow and red indicate elements carrying *3XP3::EYFP* and *3XP3::DsRed*, respectively.

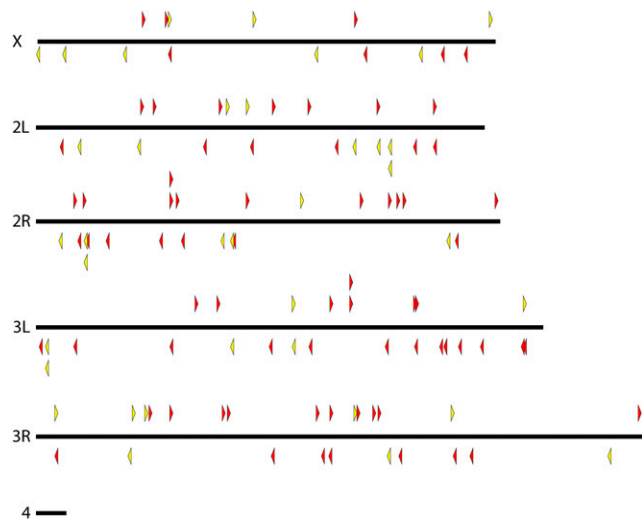
genome release 2 (Hu *et al.* 2013), available from [ftp://ftp.flybase.net/genomes/Drosophila\\_simulans/dsim\\_r2.01\\_FB2015\\_01/](ftp://ftp.flybase.net/genomes/Drosophila_simulans/dsim_r2.01_FB2015_01/). Insertion sites in *D. yakuba* and *D. santomea* were mapped to *D. yakuba* genome release 1.3 (*Drosophila* 12 Genomes Consortium *et al.* 2007), available from [ftp://ftp.flybase.net/genomes/Drosophila\\_yakuba/dyak\\_r1.3\\_FB2014\\_03/](ftp://ftp.flybase.net/genomes/Drosophila_yakuba/dyak_r1.3_FB2014_03/). The actual genomes used for mapping and the mapped positions of the transposable elements are provided in the Geneious files supplied as File S1.

### Mapping *pBac* transposon insertion sites in *D. virilis*

We previously generated multiple *pBac*(*enhancer-lacZ*) insertions into *D. virilis* to study the *svb* gene (Frankel *et al.* 2012). However, none of these *pBac* (*enhancer-lacZ*) insertions have been mapped previously. These reagents may be useful for genetic mapping studies. Therefore, we have mapped positions of these inserts using TagMap. The larger scaffolds from the *D. virilis* CAF1 assembly project (<http://insects.eugenes.org/species/data/dvir/>) (*Drosophila* 12 Genomes Consortium *et al.* 2007) have been mapped to Muller elements (Schaeffer *et al.* 2008). We combined this information with genetic linkage data to assemble ~159 Mbp of the *D. virilis* genome into the six Muller arms (N. Frankel and D. Stern, unpublished data). We mapped insertion sites to this unpublished version of the *D. virilis* genome.

### Generation of a *D. santomea white*<sup>-</sup> allele

We began to generate this collection of reagents prior to the availability of a *white*<sup>-</sup> strain of *D. santomea*. However, soon after CRISPR/Cas9-mediated genome editing became available, we generated a *white*<sup>-</sup> strain derived from *D. santomea* STO-CAGO 1482 as follows. *In vitro*-transcribed Cas9 mRNA, generated with an *EcoRI*-digested T7-*Cas9* template plasmid and the mMACHINE T7 Transcription Kit (Thermo Fisher Scientific), together with two gRNAs targeting the third exon of the *white* gene were injected into preblastoderm embryos by Rainbow Transgenics. The sequence for the T7-*Cas9*

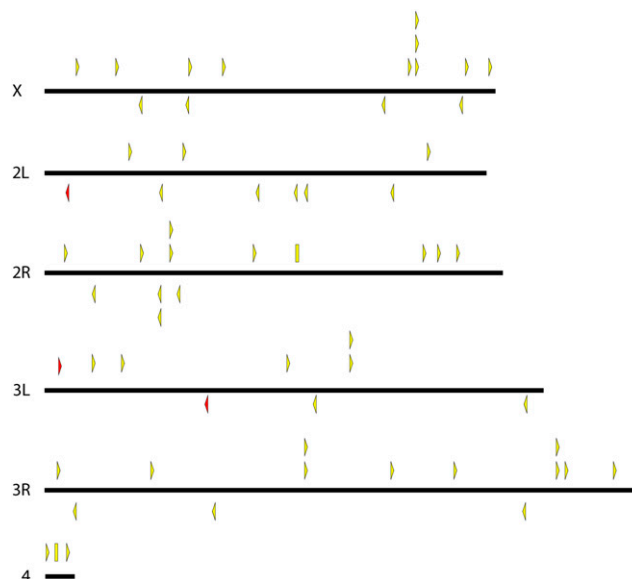


**Figure 4** Genomic insertion sites of *pBac* transposable elements in *D. yakuba*. *pBac* insertions oriented forward are indicated above each chromosome and point to the right while reverse insertions are indicated below each chromosome and point to the left. Rectangles represent inserted elements whose orientation could not be determined. Yellow and red indicate elements carrying *3XP3::EYFP* and *3XP3::DsRed*, respectively.

plasmid is provided in File S2. The gRNAs were generated by separate *in vitro* transcription reactions, using the MEGAscript T7 Transcription Kit (Thermo Fisher Scientific), of PCR-amplified products of the following forward and reverse primers: Forward primer CRISPRF-san-w12, 5'-GAA ATT AAT ACG ACT CAC TAT AGG CAA CCT GTA GAC GCC AGT TTT AGA GCT AGA AAT AGC-3'; Forward primer CRISPRF-san-w17, 5'-GAA ATT AAT ACG ACT CAC TAT AGG GCC ACG CGC TGC CGA TGT TTT AGA GCT AGA AAT AGC-3'; Reverse primer gRNA-scaffold, 5'-AAA AGC ACC GAC TCG GTG CCA CTT TTT CAA GTT GAT AAC GGA CTA GCC TTA TTT TAA CTT GCT ATT TCT AGC TCT AAA AC-3'. All PCR reactions described in this paper were performed using Phusion High Fidelity DNA Polymerase (New England Biolabs) using standard conditions. Injected G0 flies were brother-sister mated and G1 flies were screened for white eyes. Once we identified a *white*<sup>-</sup> strain, we backcrossed the *pBac*{*3XP3::EYFP-attP*} markers generated previously in *D. santomea* STO-CAGO 1482 to the *white*<sup>-</sup> strain. The *pBac* insertion sites in these new *white*<sup>-</sup> strains were then remapped with TagMap.

### Testing $\phi$ C31-mediated integration efficiency

Different *attP* landing sites provide different efficiencies of integration of *attB*-containing plasmids (Bischof *et al.* 2007). We performed a preliminary screen of integration efficiency on a subset of the *attP* landing sites that we generated. Preblastoderm embryos were co-injected with 250 ng/ $\mu$ l of plasmids containing *attB* sites and 250 ng/ $\mu$ l pBS130 (Gohl *et al.* 2011), a heat shock-inducible source of  $\phi$ C31 integrase, and 1 hr after injection were incubated at 37° for 1 hr. G0 offspring were backcrossed to the parental line and G1 offspring were screened for the relevant integration marker. We performed this screen using a heterogeneous collection of plasmids that we were integrating for other purposes. Therefore, the integration efficiencies we report are not strictly comparable between sites. Nonetheless, we were able to identify a subset of sites that provide reasonable integration efficiency and which can be made homozygous after integration of transgenes. We report these statistics for all sites that we have tested (File S3).



**Figure 5** Genomic insertion sites of *pBac* transposable elements in *D. santomea*. *pBac* insertions oriented forward are indicated above each chromosome and point to the right while reverse insertions are indicated below each chromosome and point to the left. Rectangles represent inserted elements whose orientation could not be determined. Yellow and red indicate elements carrying *3XP3::EYFP* and *3XP3::DsRed*, respectively.

### Testing expression patterns and levels of transgenes integrated in different *attP* sites

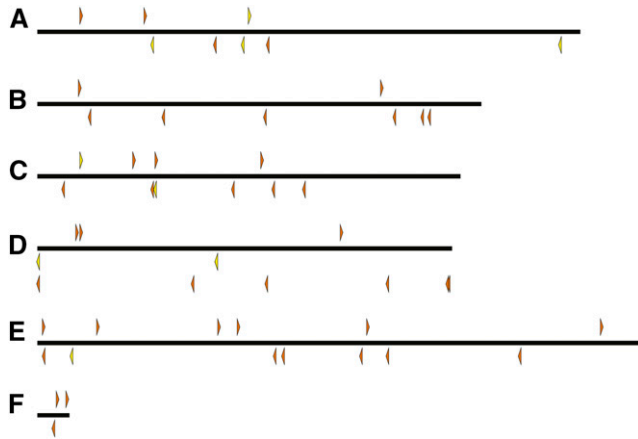
Different *attP* landing sites drive different levels and patterns of transgene expression (Pfeiffer *et al.* 2010). We have tested a subset of the *attP* sites in our collection for embryonic expression of an integrated *D. melanogaster even-skipped stripe 2* enhancer (Small *et al.* 1992). A plasmid containing the *D. melanogaster eveS2-placZ* was co-injected with 250 ng/ $\mu$ l pBS130 into ~10 *pBac*{*3XP3::EYFP-attP*} strains of each species. We isolated transgenic lines for seven *D. simulans*, four *D. mauritiana*, two *D. yakuba* strains, and four *D. santomea* strains. We performed mRNA fluorescent *in situ* hybridization (FISH) and imaged midstage 5 embryos on a Leica TCS SPE confocal microscope (antibody staining is less sensitive at these stages than FISH due to slow production of reporter gene protein products.) Embryos of all samples were scanned with equal laser power to allow quantitative comparisons of expression patterns between strains.

We performed staining experiments for all sites from each species in parallel; embryo collection, fixation, hybridization, image acquisition, and processing were performed side-by-side under identical conditions. Confocal exposures were identical for each series. Image series were acquired in a single day, to minimize signal loss. Sum projections of confocal stacks were assembled, embryos were scaled to match sizes, background was subtracted using a 50-pixel rolling-ball radius, and fluorescence intensity was analyzed using ImageJ software (<http://rsb.info.nih.gov/ij/>).

### Killing EYFP expression from *attP* landing sites

Expression of the *EYFP* genes associated with the *attP* sites may conflict with some potential uses of the *attP* landing sites, for example for integration of transgenes driving *GFP* derivatives, such as *GCaMP*, in the brain. Therefore, we have started generating *pBac*{*3XP3::EYFP-attP*} strains where we have killed the *EYFP* activity using CRISPR/Cas9-mediated targeted mutagenesis. We first built a derivative of the





**Figure 6** Genomic insertion sites of *pBac* transposable elements on each chromosome (Muller element) of *D. virilis*. Each triangle represents a unique *pBac* element insertion. Some strains carry multiple insertion events. *pBac* insertions oriented forward are indicated above each chromosome and point to the right and reverse insertions are indicated below each chromosome and point to the left. Yellow and orange indicate elements carrying *3XP3::EYFP* and *w+*, respectively.

*pCFD4-U61-U63* tandem gRNAs plasmid (Port *et al.* 2014) where we replaced the *vermillion* marker with a *3XP3::DsRed* dominant marker. The *vermillion* marker was removed by *HindIII* digestion of *pCFD4-U61-U63* and isolation of the 5253 bp band. The *3XP3::DsRed* cassette was amplified from a *pUC57{3xP3::DsRed}* plasmid using the following primers: 5'-TAC GAC TCA CTA TAG GGC GAA TTG GGT ACA CCA GTG AAT TCG AGC TCG GT-3' and 5'-TTG GAT GCA GCC TCG AGA TCG ATG ATA TCA ATT ACG CCA AGC TTG CAT GC-3'. The PCR product and vector backbone were assembled with Gibson assembly (Gibson *et al.* 2009) following [http://openwetware.org/wiki/Gibson\\_Assembly](http://openwetware.org/wiki/Gibson_Assembly) to generate *p{CFD4-3xP3::DsRed-BbsI}*. To remove the *BbsI* restriction site from *DsRed*, which conflicts with the *BbsI* restriction site used for cloning gRNA sequences, we digested this plasmid with *NcoI* and isolated the ~6 kb fragment, PCR-amplified this region with primers that eliminated the *BbsI* restriction site (forward primer: 5'-CGG GCC CGG GAT CCA CCG GTC GCC ACC ATG GTG CGC TCC TCC AAG AAC GTC A-3' and reverse primer: 5'-CGC TCG GTG GAG GCC TCC CAG CCC ATG GTT TTC TTC TGC ATT ACG GGG CC-3'), and Gibson cloned the PCR product into the plasmid backbone. This yielded plasmid *p{CFD4-3xP3::DsRed}*.

To make a plasmid for mutating *EYFP* in fly lines, we digested *p{CFD4-3xP3::DsRed}* with *BbsI* and gel purified the 5913 bp fragment. A gBlocks Gene Fragment (IDT) (5'-CAA GTA CAT ATT CTG CAA GAG TAC AGT ATA TAT AGG AAA GAT ATC CGG GTG AAC TTC GGG TGG TGC AGA TGA ACT TCA GTT TTA GAG CTA GAA ATA GCA AGT TAA AAT AAG GCT AGT CCG TTA TCA ACT TG-3'), which contained a gRNA sequence targeting *EYFP* that was previously validated by direct injection of gRNA, was synthesized and Gibson assembled with the *BbsI*-digested fragment of *p{CFD4-3xP3::DsRed}* to make *p{CFD4-EYFP-3xP3::DsRed}*.

This plasmid contains *attB* and can be integrated into *attP* sites. We tested this by integrating this plasmid into the *attP* site of *D. simulans* line 930. This plasmid is a potent source of gRNA targeting *EYFP*, which we confirmed by crossing this line to a transgenic strain carrying *nos-Cas9*. We have generated transgenic strains of *D. simulans*, *D. mauritiana*, and

**Table 1** Number of *attP* strains of each of five species that did not or did support integration of *attB* plasmids

Species	Number of Strains with Zero Integrants	Number of Strains with at Least One Integrant
<i>D. mauritiana</i>	14	21
<i>D. simulans</i>	13	29
<i>D. santomea</i>	1	8
<i>D. yakuba</i>	1	19
<i>D. virilis</i>	9	0

Details are available in File S3.

*D. yakuba* carrying *nos-Cas9* [Addgene plasmid 62208, described in Port *et al.* (2014)] and details of these lines are provided as File S3.

To knock out *EYFP* in specific strains carrying *pBac{3XP3::EYFP-attP}*, we co-injected 500 ng/μl *in vitro*-transcribed Cas9 mRNA and 250 ng/μl *p{CFD4-EYFP-3xP3::DsRed}*. G0 individuals were brother-sister mated and we screened for reduction or loss of *EYFP* expression in G1 progeny. Individuals displaying reduced or no *EYFP* expression were crossed to generate strains homozygous for *EYFP*<sup>-</sup>.

#### Data availability

Plasmid *pBac{3XP3::EYFP-attP}* is available from D. Stern upon request. The *p{CFD4}* derivative plasmids have been deposited with Addgene (plasmid IDs 86863 and 86864). All fly stocks are maintained in the Stern lab at Janelia Research Campus and all requests for fly stocks should be directed to D. Stern. The raw iPCR and TagMap data are available upon request from D. Stern. We continue to produce new fly strains based on the reagents described in this paper. An Excel sheet containing information about all strains in this paper and any new lines is available at [http://research.janelia.org/sternlab/Strains\\_and\\_Integration\\_Efficiencies.xlsx](http://research.janelia.org/sternlab/Strains_and_Integration_Efficiencies.xlsx). Geneious files containing genomic insertion sites for all transgenes will be updated with new strains and are available at the following sites: [http://research.janelia.org/sternlab/D.simulans\\_mauritiana\\_insertions.geneious](http://research.janelia.org/sternlab/D.simulans_mauritiana_insertions.geneious); [http://research.janelia.org/sternlab/D.yakuba\\_santomea\\_insertions.geneious](http://research.janelia.org/sternlab/D.yakuba_santomea_insertions.geneious); and [http://research.janelia.org/sternlab/D.virilis\\_insertions.geneious](http://research.janelia.org/sternlab/D.virilis_insertions.geneious). All of these files can be accessed via our lab web page at <https://www.janelia.org/lab/stern-lab/tools-reagents-data>.

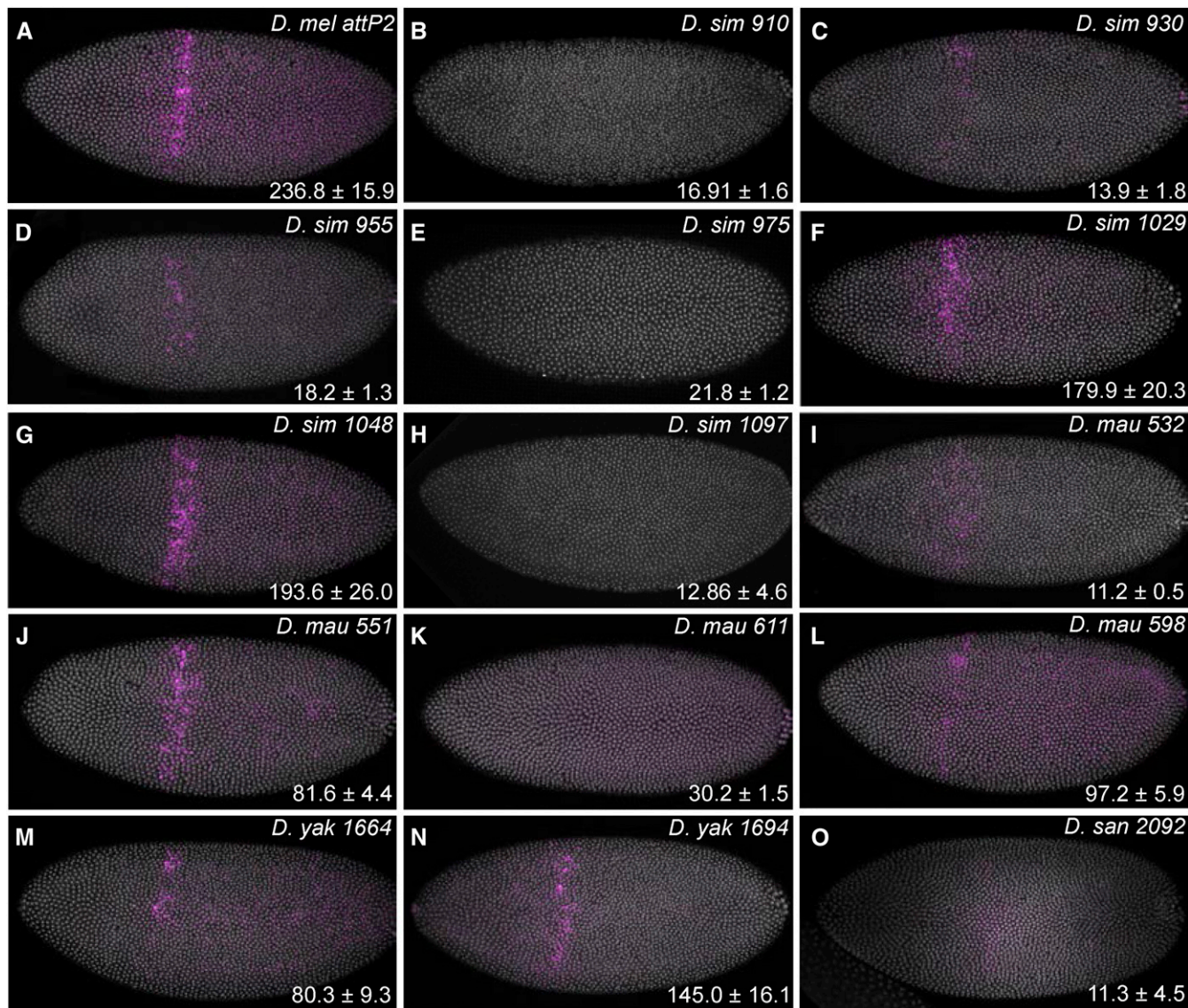
## RESULTS

### Generation and mapping of *pBac{3XP3::EYFP-attP}* strains

We generated many strains carrying *pBac{3XP3::EYFP-attP}* and *pBac{3XP3::DsRed}* insertions, mapped these, and culled the collection to unique lines that could be maintained as homozygotes. The final collection includes 184 *D. simulans* lines, 122 *D. mauritiana* lines, 104 *D. yakuba* lines, 64 *D. santomea* lines, and 9 *D. virilis* lines. Maps indicating the insertion site locations are shown in Figure 2, Figure 3, Figure 4, Figure 5, and Figure 6, and are provided as searchable Geneious files (<http://www.geneious.com/>) in File S1. Details of the transgenic strains are provided in File S3.

### Mapping *pBac* transposon insertion sites in *D. virilis*

To assist with genetic experiments in *D. virilis*, we mapped the insertion locations for all *pBac* lines generated in our lab for a previously published study (Frankel *et al.* 2012). We mapped 58 transposon insertions



**Figure 7** Variation in transgene expression supported by different *attP* landing sites in four species. An *eveS2* transgene driving expression in the *even-skipped stripe 2* domain of early embryos was inserted into multiple *attP* sites of each of four species: *D. simulans* (*D. sim*), *D. mauritiana* (*D. mau*), *D. yakuba* (*D. yak*), and *D. santomea* (*D. san*). *eveS2* expression is shown in purple and DNA was counterstained with 4',6-diamidino-2-phenylindole (DAPI) and shown in white. Expression levels in the stripe 2 domain were quantified in 10 embryos of each strain and the mean ± SD are reported in the bottom right corner of each panel in arbitrary units of fluorescence intensity. (A) As a control, we stained a line containing the same plasmid inserted into the *attP2* site of *D. melanogaster*. (B–N) Seven *attP* strains of *D. simulans* (B–H), four *attP* strains of *D. mauritiana* (I–L), and two *attP* strains of *D. yakuba* (M and N) support different levels of *eveS2* expression. (N) Strain 1694 contains two *attP* landing sites, and we have not determined which landing site contains the *eveS2* transgene or whether both do. (O) None of the four *D. santomea* *attP* strains we tested supported high levels of spatio-temporally correct *eveS2* expression. The strain displaying the strongest expression (2092) is shown here.

from 39 *pBac{enhancer-lacZ}* strains plus nine new *pBac{3XP3::EYFP-attP}* strains. Some strains contained multiple insertions and some insertions mapped to contigs that are not currently associated with Muller arm chromosomes. These results are shown in Figure 6 and are available in a Geneious file and File S3.

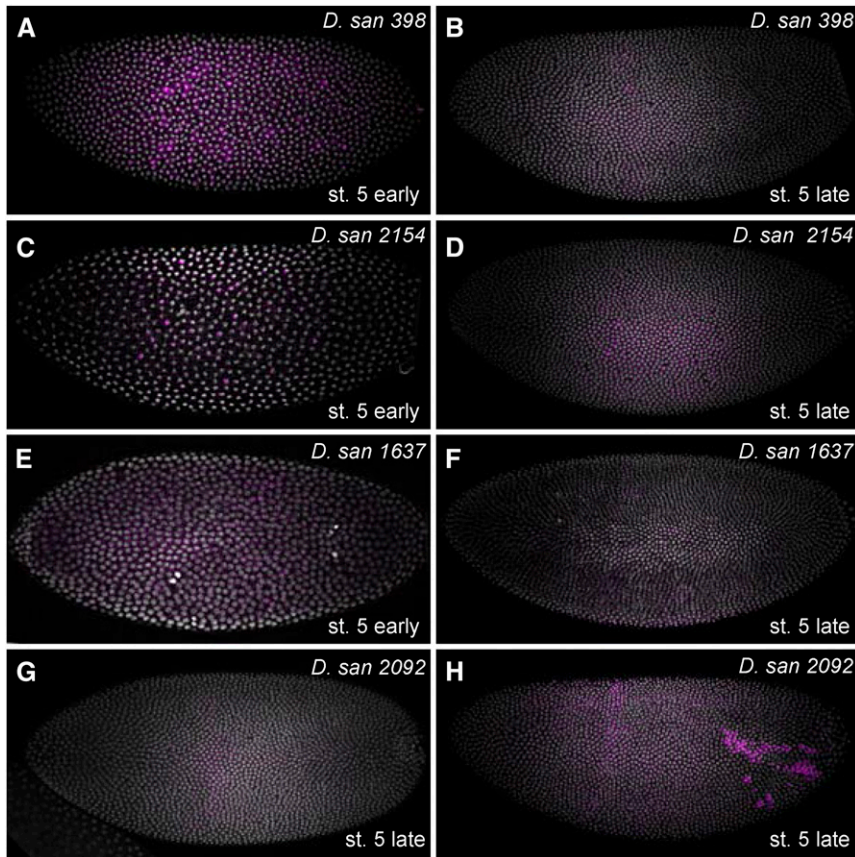
#### Testing $\phi$ C31-mediated integration efficiency

We tested efficiency of integration of *attB* plasmids into *attP* landing sites of multiple strains of each species. There are strong differences in integration efficiencies between landing sites. Some landing sites in *D. simulans*, *D. mauritiana*, *D. santomea*, and

*D. yakuba* supported integration of *attB* plasmids, although many landing sites did not support integration at reasonable frequency (Table 1). Details of integration efficiencies for each line are provided in File S3.

In addition, we tested nine *D. virilis* strains carrying *pBac{3XP3::EYFP-attP}* for integration of the *eveS2-placZ* plasmid, which contains an *attB* site. We screened ~100 fertile G0 offspring for each of these nine strains and did not recover any integrants. This is a surprising result, and we do not yet know whether this failure of *attB* integration is specific to these lines or reflects a general low efficiency of *attP-attB* integration in *D. virilis*.





**Figure 8** Four *D. santomea* *attP* landing sites do not support spatio-temporally correct *eveS2* transgene expression. (A–G) In early stage 5 embryonic stages, the lines displayed variable levels of diffuse expression, as is often observed with *eveS2* transgenes (A, C, and E). However, at late stage 5, none of the lines drove strong expression in the stripe 2 region (B, D, F, and G). (H) Strain 2092 sometimes displayed strong ectopic expression outside of the stripe 2 domain.

### Testing expression patterns of transgenes integrated in different *attP* sites

We integrated a *D. melanogaster* *eveS2-placZ* plasmid into multiple *attP* landing site strains of each species to examine variability in expression at different landing sites. Levels of reporter gene expression varied between strains (Figure 7). In *D. simulans*, *D. mauritiana*, and *D. yakuba*, we identified at least one strain that drove strong and temporal-spatially accurate levels of *eveS2* expression. However, of the four landing sites we tested in *D. santomea*, none provided strong expression of *eveS2* (Figure 7 and Figure 8). *eveS2* transgenes often drive weak, spatially diffuse expression prior to stage 5, and all of the *D. santomea* strains displayed similar diffuse, weak expression at early stages. We also observed ectopic expression of the *eveS2* transgene in *D. santomea* 2092 (Figure 8H). It is not clear if the poor expression of *eveS2* in these *D. santomea* landing sites reflects differential regulation of the *D. melanogaster* *eveS2* enhancer in *D. santomea* or suppression of expression caused by position effects of these specific landing sites.

### Unmarked *attP* landing sites

To facilitate integration of plasmids expressing fluorescent proteins that overlap with the excitation and emission spectra of EYFP, we have generated a subset of strains in which we induced null mutations in the *EYFP* gene marking the *attP* landing sites. These strains were generated by CRISPR/Cas9-induced mutagenesis. All strains were sequenced to ensure that the mutations did not disrupt the *attP* landing site. We have so far generated two strains in *D. mauritiana*, and three strains in each of *D. santomea*, *D. simulans*, and *D. yakuba* (File S3).

### DISCUSSION

We have generated a collection of transgenic strains that will be useful for multiple kinds of experiments. First, the *3XP3::EYFP-attP* strains provide a collection of *attP* landing sites for each species that will facilitate transgenic assays in these species. Integration efficiencies vary widely between strains and our experiments provide some guidance to identify landing sites with the highest efficiency of integration. Second, these transgenes carry markers that will be useful for genetic mapping experiments. Several published studies have already used these reagents and illustrate the power of these strains for genetic studies (Andolfatto *et al.* 2011; Erezyilmaz and Stern 2013; Ding *et al.* 2016).

We have generated transgenic strains using these *attP* landing sites and found that they show variation in embryonic expression patterns (Figure 7 and Figure 8). These results provide a rough guide to which strains may be useful for experiments that require low or high levels of embryonic expression. However, these results may not be predictive of transgene expression patterns at other developmental stages and in other tissues, and we strongly encourage colleagues to test a variety of landing sites for their experiments and report their experiences to us. We plan to continue to maintain a database reporting on integration efficiencies and expression patterns, and we will periodically update the Excel file associated with this manuscript.

This collection of reagents complements the existing resources available for studying species of the genus *Drosophila*, including the availability of multiple genome sequences (*Drosophila* 12 Genomes Consortium *et al.* 2007) and BAC resources (Song *et al.* 2011). This resource will accelerate research on gene function in diverse *Drosophila* species and the study of evolution in the genus *Drosophila*.

## ACKNOWLEDGMENTS

We thank Peter Andolfatto and an anonymous referee for helpful comments that improved this paper. Most fly stocks were provided by the San Diego Species Stock Center and *D. santomea* STO CAGO 1482 was kindly provided by Peter Andolfatto. We thank Ernst Wimmer for providing the original *piggyBac* reporter plasmids. The authors declare no competing financial interests.

Author Contributions: D.L.S. conceived of the project. N.F. made *pBac* {3XP3::*EYFP-attP*}. D.L.S. made all other plasmids. D.L.S., Y.D., G.K., and J.D.M. screened injected flies for integration events. Y.D., R.K., A.L., E.K., and S.P. performed iPCR experiments. E.K. prepared DNA samples for TagMap. D.L.S. performed TagMap. A.L. and S.P. sequenced the TagMap libraries. D.L.S., Y.D., G.K., and E.K. performed the genetics. J.C. performed the embryo *in situ* hybridization experiments. D.L.S. wrote the paper.

## LITERATURE CITED

- Andolfatto, P., D. Davison, D. Erezyilmaz, T. T. Hu, J. Mast *et al.*, 2011 Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.* 21: 610–617.
- Ararape, L. O., H. Montenegro, B. Lemos, and D. L. Hartl, 2010 Fine-scale genetic mapping of a hybrid sterility factor between *Drosophila simulans* and *D. mauritiana*: the varied and elusive functions of “speciation genes”. *BMC Evol. Biol.* 10: 385.
- Arif, S., M. Hilbrant, C. Hopfen, I. Almudi, M. D. S. Nunes *et al.*, 2013 Genetic and developmental analysis of differences in eye and face morphology between *Drosophila simulans* and *Drosophila mauritiana*. *Evol. Dev.* 15: 257–267.
- Barbash, D. A., 2010 Ninety years of *Drosophila melanogaster* hybrids. *Genetics* 186: 1–8.
- Bastock, M., 1956 A gene mutation which changes a behavior pattern. *Evolution* (N. Y.) 10: 421–439.
- Beall, E. L., M. B. Mahoney, and D. C. Rio, 2002 Identification and analysis of a hyperactive mutant form of *Drosophila P*-element transposase. *Genetics* 162: 217–227.
- Bischof, J., R. K. Maeda, M. Hediger, F. Karch, and K. Basler, 2007 An optimized transgenesis system for *Drosophila* using germ-line-specific  $\phi$ C31 integrases. *Proc. Natl. Acad. Sci. USA* 104: 3312–3317.
- Campbell, J. L., and H. A. Nash, 2001 Volatile general anesthetics reveal a neurobiological role for the *white* and *brown* genes of *Drosophila melanogaster*. *J. Neurobiol.* 49: 339–349.
- Cande, J., P. Andolfatto, B. Prud'homme, D. L. Stern, and N. Gompel, 2012 Evolution of multiple additive loci caused divergence between *Drosophila yakuba* and *D. santomea* in wing rowing during male courtship. *PLoS One* 7: e43888.
- Carbone, M. A., A. Llopart, M. deAngelis, J. A. Coyne, and T. F. Mackay, 2005 Quantitative trait loci affecting the difference in pigmentation between *Drosophila yakuba* and *D. santomea*. *Genetics* 171: 211–225.
- Coyne, J. A., and B. Charlesworth, 1997 Genetics of a pheromonal difference affecting sexual isolation between *Drosophila mauritiana* and *D. sechellia*. *Genetics* 145: 1015–1030.
- Ding, Y., A. Berrocal, T. Morita, K. D. Longden, and D. L. Stern, 2016 Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 536: 329–332.
- Drapeau, M. D., S. A. Cyran, M. M. Viering, P. K. Geyer, and A. D. Long, 2006 A *cis*-regulatory sequence within the *yellow* locus of *Drosophila melanogaster* required for normal male mating success. *Genetics* 172: 1009–1030.
- Drosophila 12 Genomes Consortium, Clark, A. G., M. B. Eisen, D. R. Smith, C. M. Bergman, B. Oliver *et al.*, 2007 Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450: 203–218.
- Erezyilmaz, D. F., and D. L. Stern, 2013 Pupariation site preference within and between *Drosophila* sibling species. *Evolution* (N. Y.) 67: 2714–2727.
- Frankel, N., S. Wang, and D. L. Stern, 2012 Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proc. Natl. Acad. Sci. USA* 109: 20975–20979.
- Gibson, D. G., L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, III *et al.*, 2009 Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* 6: 343–345.
- Gleason, J. M., and M. G. Ritchie, 2004 Do quantitative trait loci (QTL) for a courtship song difference between *Drosophila simulans* and *D. sechellia* coincide with candidate genes and intraspecific QTL? *Genetics* 166: 1303–1311.
- Gleason, J. M., J. M. Jallon, J. D. Rouault, and M. G. Ritchie, 2005 Quantitative trait loci for cuticular hydrocarbons associated with sexual isolation between *Drosophila simulans* and *D. sechellia*. *Genetics* 171: 1789–1798.
- Gohl, D. M., M. A. Silies, X. J. Gao, S. Bhalerao, F. J. Luongo *et al.*, 2011 A versatile *in vivo* system for directed dissection of gene expression patterns. *Nat. Methods* 8: 231–237.
- Groth, A. C., M. Fish, R. Nusse, and M. P. Calos, 2004 Construction of transgenic *Drosophila* by using the site-specific integrase from phage  $\phi$ C31. *Genetics* 166: 1775–1782.
- Handler, A. M., and R. A. Harrell, II, 1999 Germline transformation of *Drosophila melanogaster* with the *piggyBac* transposon vector. *Insect Mol. Biol.* 8: 449–457.
- Horn, C., and E. A. Wimmer, 2000 A versatile vector set for animal transgenesis. *Dev. Genes Evol.* 210: 630–637.
- Horn, C., N. Offen, S. Nystedt, U. Hacker, and E. A. Wimmer, 2003 *piggyBac*-based insertional mutagenesis and enhancer detection as a tool for functional insect genomics. *Genetics* 163: 647–661.
- Hu, T. T., M. B. Eisen, K. R. Thornton, and P. Andolfatto, 2013 A second-generation assembly of the *Drosophila simulans* genome provides new insights into patterns of lineage-specific divergence. *Genome Res.* 23: 89–98.
- Kyriacou, C. P. P., B. Burnet, and K. Connolly, 1978 The behavioural basis of overdominance in competitive mating success at the *ebony* locus of *Drosophila melanogaster*. *Anim. Behav.* 26: 1195–1206.
- Lemeunier, F., and M. A. Ashburner, 1976 Relationships within the *Drosophila* species subgroup of the genus *Drosophila* (Sophophora). II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. *Proc. R. Soc. Lond. B Biol. Sci.* 193: 275–294.
- Liu, J., J. M. Nercer, L. F. Stam, G. C. Gibson, Z.-B. B. Zeng *et al.*, 1996 Genetic analysis of a morphological shape difference in the male genitalia of *Drosophila simulans* and *D. mauritiana*. *Genetics* 142: 1129–1145.
- Macdonald, S. J., and D. B. Goldstein, 1999 A quantitative genetic analysis of male sexual traits distinguishing the sibling species *Drosophila simulans* and *D. sechellia*. *Genetics* 153: 1683–1699.
- Masly, J. P., and D. C. Presgraves, 2007 High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. *PLoS Biol.* 5: e243.
- Masly, J. P., J. E. Dalton, S. Srivastava, L. Chen, and M. N. Arbeitman, 2011 The genetic basis of rapidly evolving male genital morphology in *Drosophila*. *Genetics* 189: 357–374.
- McGregor, A. P., V. Orgogozo, I. Delon, J. Zanet, D. G. Srinivasan *et al.*, 2007 Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* 448: 587–590.
- Moehring, A. J., J. Li, M. D. Schug, S. G. Smith, M. deAngelis *et al.*, 2004 Quantitative trait loci for sexual isolation between *Drosophila simulans* and *D. mauritiana*. *Genetics* 167: 1265–1274.
- Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. Mackay, 2006a The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference. *Genetics* 173: 215–223.
- Moehring, A. J., A. Llopart, S. Elwyn, J. A. Coyne, and T. F. Mackay, 2006b The genetic basis of postzygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to hybrid male sterility. *Genetics* 173: 225–233.
- Ochman, H., A. S. Gerber, and D. L. Hartl, 1988 Genetic applications of an inverse polymerase chain reaction. *Genetics* 120: 621–623.
- Orgogozo, V., and D. L. Stern, 2009 How different are recently diverged species? *Fly* (Austin) 3: 117.



- Orgogozo, V., K. W. Broman, and D. L. Stern, 2006 High-resolution quantitative trait locus mapping reveals sign epistasis controlling ovariole number between two *Drosophila* species. *Genetics* 173: 197–205.
- Peluffo, A. E., I. Nuez, V. Debat, R. Savisaar, D. L. Stern *et al.*, 2015 A major locus controls a genital shape difference involved in reproductive isolation between *Drosophila yakuba* and *Drosophila santomea*. *G3 (Bethesda)* 5: 2893–2901.
- Pfeiffer, B. D., T. T. Ngo, K. L. Hibbard, C. Murphy, A. Jenett *et al.*, 2010 Refinement of tools for targeted gene expression in *Drosophila*. *Genetics* 186: 735–755.
- Port, F., H. Chen, T. Lee, and S. L. Bullock, 2014 Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 111: E2967–E2976.
- Powell, J. R., 1997 Progress and Prospects in Evolutionary Biology: The *Drosophila* Model. Oxford University Press, Oxford, UK.
- R’Kha, S., P. Capy, and J. R. David, 1991 Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis. *Proc. Natl. Acad. Sci. USA* 88: 1835–1839.
- Schaeffer, S. W., A. Bhutkar, B. F. McAllister, M. Matsuda, L. M. Matzkin *et al.*, 2008 Polytene chromosomal maps of 11 *Drosophila* species: the order of genomic scaffolds inferred from genetic and physical maps. *Genetics* 179: 1601–1655.
- Small, S., A. Blair, and M. Levine, 1992 Regulation of even-skipped stripe 2 in the *Drosophila* embryo. *EMBO J.* 11: 4047–4057.
- Song, X., J. L. Goicoechea, J. S. S. Ammiraju, M. Luo, R. He *et al.*, 2011 The 19 genomes of *Drosophila*: a BAC library resource for genus-wide and genome-scale comparative evolutionary research. *Genetics* 187: 1023–1030.
- Stern, D. L., 2016 Tagmentation-based mapping (TagMap) of mobile DNA genomic insertion sites. *bioRxiv*. Available at: <https://doi.org/10.1101/037762>.
- Suh, J., and F. R. Jackson, 2007 *Drosophila* ebony activity is required in glia for the circadian regulation of locomotor activity. *Neuron* 55: 435–447.
- Tanaka, K. M., C. Hopfen, M. R. Herbert, C. Schlotterer, D. L. Stern *et al.*, 2015 Genetic architecture and functional characterization of genes underlying the rapid diversification of male external genitalia between *Drosophila simulans* and *Drosophila mauritiana*. *Genetics* 200: 357–369.
- Tang, S., and D. C. Presgraves, 2015 Lineage-specific evolution of the complex *Nup160* hybrid incompatibility between *Drosophila melanogaster* and its sister species. *Genetics* 200: 1245–1254.
- Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie, 2003a Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. *Genetics* 164: 1383–1397.
- Tao, Y., Z. B. Zeng, J. Li, D. L. Hartl, and C. C. Laurie, 2003b Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. II. Mapping hybrid male sterility loci on the third chromosome. *Genetics* 164: 1399–1418.
- Thibault, S. T., M. A. Singer, W. Y. Miyazaki, B. Milash, N. A. Dompe *et al.*, 2004 A complementary transposon tool kit for *Drosophila melanogaster* using *P* and *piggyBac*. *Nat. Genet.* 36: 283–287.
- True, J. R., J. M. Mercer, and C. C. Laurie, 1996a Differences in crossover frequency and distribution among three sibling species of *Drosophila*. *Genetics* 142: 507–523.
- True, J. R., B. S. Weir, and C. C. Laurie, 1996b A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics* 142: 819–837.
- True, J. R., J. Liu, L. F. Stam, Z.-B. Zeng, and C. C. Laurie, 1997 Quantitative genetic analysis of divergence in male secondary sexual traits between *Drosophila simulans* and *Drosophila mauritiana*. *Evolution (N. Y.)* 51: 816–832.
- Xiao, C., and R. M. Robertson, 2016 Timing of locomotor recovery from anoxia modulated by the *white* gene in *Drosophila*. *Genetics* 203: 787–797.
- Yassin, A., V. Debat, H. Bastide, N. Gidaszewski, J. R. David *et al.*, 2016 Recurrent specialization on a toxic fruit in an island *Drosophila* population. *Proc. Natl. Acad. Sci. USA* 113: 4771–4776.
- Zhang, S. D., and W. F. Odenwald, 1995 Misexpression of the *white (w)* gene triggers male-male courtship in *Drosophila*. *Proc. Natl. Acad. Sci. USA* 92: 5525–5529.

Communicating editor: J. A. Birchler