

# Robust $\Phi$ C31-Mediated Genome Engineering in *Drosophila melanogaster* Using Minimal attP/attB Phage Sites

Roumen Voutev<sup>1</sup> and Richard S. Mann<sup>1</sup>

Departments of Biochemistry and Molecular Biophysics and Systems Biology, Jerome L. Greene Science Center, Columbia University, New York, NY 10027

ORCID IDs: 0000-0003-0066-4731 (R.V.); 0000-0002-4749-2765 (R.S.M.)

**ABSTRACT** Effective genome engineering should lead to a desired locus change with minimal adverse impact to the genome itself. However, flanking loci with site-directed recombinase recognition sites, such as those of the phage  $\Phi$ C31 integrase, allows for creation of platforms for cassette exchange and manipulation of genomic regions in an iterative manner, once specific loci have been targeted. Here we show that a genomic locus engineered with inverted minimal phage  $\Phi$ C31 attP/attB sites can undergo efficient recombinase-mediated cassette exchange (RMCE) in the fruit fly *Drosophila melanogaster*.

## KEYWORDS

$\Phi$ C31  
recombinase  
cassette exchange  
CRISPR/Cas9  
genome editing  
*D. melanogaster*  
attB/attP sites  
Genome Report

The introduction of CRISPR/Cas9 genome editing technique as an everyday molecular biology tool has opened enormous future opportunities for both biological research and gene therapy (reviewed in DELKER AND MANN 2017). As a supplement to this tool, it could be very advantageous to be able to reiteratively modify a locus of interest once it has already been targeted with the CRISPR/Cas9 system. One way to achieve such versatility is by flanking the targeted locus with phage attP or attB sites of one of the already extensively researched site-directed recombinases such as  $\Phi$ C31 (GROTH *et al.* 2000) or Bxb1 (GHOSH *et al.* 2003; KIM *et al.* 2003). Subsequently, the resulting attP(attB)-flanked allele could be edited with admirable precision through recombinase-mediated cassette exchange (RMCE) without adverse effects to the genome, as long

as the attP/attB scars do not cause significant DNA/chromatin changes.

## METHODS & MATERIALS

*Drosophila melanogaster* strain *M[vas-int.Dm]ZH-2A* (#40161, Bloomington *Drosophila* Stock Center, Bloomington, IN) was used as a source of germline integrase.

39bp  $\Phi$ C31 attP site (CCCCAACTGGGGTAACCTTTGAGTTC-TCTCAGTTGGGGG) was introduced in vector pRVV598 (#87629, [www.addgene.org](http://www.addgene.org); (VOUTEV AND MANN 2017)) in forward and reverse orientation (Figure 1A), respectively, flanking a *hs-neo* cassette and replacing the Bxb1 attP sites in vector pRVV598. A loxP site was introduced ahead of this cassette and the resulting vector was used for injection and creation of the allele  $\Phi$ *Phi*<sup>hs-neo</sup>. The  $\Phi$ C31 *ubi-GFP* RMCE vector (Figure 1A) was created by replacing the Bxb1 attP sites in vector pRVV651 (#87631, [www.addgene.org](http://www.addgene.org); (VOUTEV AND MANN 2017)) with 36bp  $\Phi$ C31 attB sites (GGGTGCCAGGGCGTGCCCTTGGGCTCCCGGGCGCG) in forward and reverse orientation, respectively, thus flanking a *Ubi-GFP* cassette (Figure 1A). Plasmid DNA, maps, and complete vector sequences are made available at Addgene (Cambridge, MA, USA; [www.addgene.org](http://www.addgene.org)); Addgene vector IDs: 108279, 108280, 108281, 108282, and 108283.

## Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article.

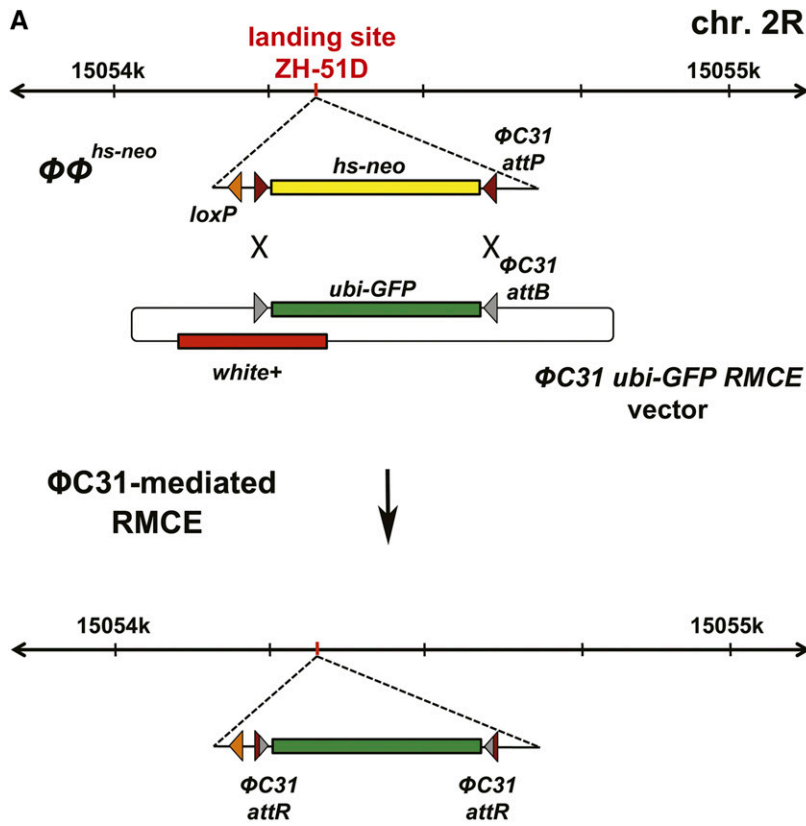
Copyright © 2018 Voutev, Mann

doi: <https://doi.org/10.1534/g3.118.200051>

Manuscript received January 17, 2018; accepted for publication March 7, 2018; published Early Online March 9, 2018.

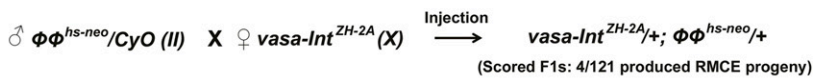
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

<sup>1</sup>Both RV (roumen.voutev@gmail.com) and RSM (rsm10@columbia.edu) are corresponding authors. Departments of Biochemistry and Molecular Biophysics and Systems Biology, Jerome L. Greene Science Center, Columbia University, 3227 Broadway, Quad. 9A, New York, NY 10027.

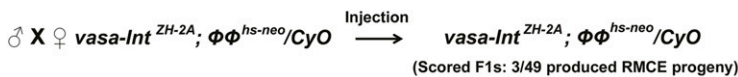


**Figure 1** Genome engineering by using minimal  $\Phi C31$  attP/attB sites. A) Schematic representation of the ZH-51D landing site locus on chromosome 2R and the genome engineering of the  $\Phi\Phi^{hs-neo}$  allele using  $\Phi C31$ -catalyzed recombinase-mediated cassette exchange (RMCE). Brown represents a 39bp  $\Phi C31$  attP site; gray represents a 36bp  $\Phi C31$  attB site. B) Injection schemes employed in this study. Homozygous  $\Phi\Phi^{hs-neo}$  F1 fertile animals segregating in test schemes II and III were included in the scoring as well.

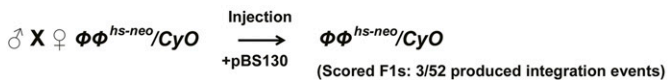
**B Test scheme I:**



**Test scheme II:**



**Test scheme III:**



**RESULTS AND DISCUSSION**

The  $\Phi C31$  site-directed recombinase has already become a common tool in fly genetics for both genome plasmid integration (GROTH *et al.* 2004) and RMCE (VENKEN *et al.* 2011). In addition, 54 bp attB and 50 bp attP  $\Phi C31$  sites have been demonstrated to mediate efficient plasmid integration events in *D. melanogaster* (HUANG *et al.* 2009) but these sites have not been tested for RMCE. Moreover, 40 bp attB/220 bp attP pairs of  $\Phi C31$  sites are capable of performing efficient RMCE in the fruit fly (BATEMAN AND WU 2008). However, even shorter 34 bp attB and 39 bp attP  $\Phi C31$  sites have been shown to function in *E. coli* with close to 100% efficiency, while lowering further the number of base pairs of these sites reduces the efficiency of recombination dramatically (GROTH *et al.* 2000).

Here we test if shorter  $\Phi C31$  att sites would function in *D. melanogaster* for RMCE, because such sites would in principle diminish the effects of the exogenous DNA to a locus of interest during genome engineering.

The features of each locus should carefully be considered before introducing any exogenous attP/attB sites. For example, an enhancer element of interest that is controlled by Mad/Smad (Mothers against Dpp) proteins would not be feasible for flanking and further study with the 54 bp  $\Phi C31$  attB site (HUANG *et al.* 2009) because it contains a consensus GCCGCGGT Mad binding site (KIM *et al.* 1997). In addition, this attB site contains a putative splice donor (agccgcgGTGCGGGT, in vector pGE-attB (HUANG *et al.* 2009)) with a 0.29 score (score ranges from 0 to 1, (REESE *et al.* 1997)), which might interfere with splicing if

introduced as a flank within introns/exons or eRNAs. Using longer attB and attP sites exponentially increases the number of putative transcription factor (TF) binding sites or other regulatory sites, which prevents them from being a viable option for flanking certain loci. For example, the commonly used in RMCE 101 bp attB sites (VENKEN *et al.* 2011) contain additional putative splice donor sites in both the forward and reverse DNA strand (tagcgatGTAGGTCA (0.56 score) and cacatggGTGAGGTG (0.70 score (REESE *et al.* 1997)) respectively, in vector pBS-KS-attB1-2 (VENKEN *et al.* 2011)) and many more putative transcription factor sites from diverse TF families (FlyFactorSurvey, (ZHU *et al.* 2011)). Thus, we decided to test minimal attB/attP sites for RMCE and creation of platforms for cassette exchange in order to strongly reduce the number of transcription factors and other DNA/RNA-binding regulatory proteins that could potentially bind to these exogenous sequences.

To test minimal  $\Phi$ C31 sites in RMCE, we used as a starting point the  $BB^{hs-neo}$  allele that we previously created (VOUDEV AND MANN 2017) in landing site ZH-51D (BISCHOF *et al.* 2007). We introduced through Bxb1-mediated integration in the distal (right) Bxb1 attP site (VOUDEV AND MANN 2017) a plasmid containing the selectable marker  $hs-neo$  (STELLER AND PIRROTTA 1985) flanked by inverted 39bp  $\Phi$ C31 attP sites (Figure 1A). We also positioned a loxP site ahead of this cassette (Figure 1A) that allowed us to excise all intervening plasmid DNA (and the leftover cassette from  $BB^{hs-neo}$ ) through Cre/loxP-mediated excision, which is characteristic for landing site ZH-51D (BISCHOF *et al.* 2007). Thus, we converted the  $BB^{hs-neo}$  into a clean allele of  $hs-neo$  flanked by minimal inverted  $\Phi$ C31 attP sites (Figure 1A), which we called  $\Phi\Phi^{hs-neo}$ .

We also created a compatible  $\Phi$ C31 RMCE vector that contains *ubiquitin-GFP* (*ubi-GFP*) cassette flanked by inverted minimal 36 bp  $\Phi$ C31 attB sites (Figure 1A). In addition, this  $\Phi$ C31 *ubi-GFP* RMCE vector contains *white* (*w+*) selectable marker (Figure 1A) that allows for visually differentiating between vector integration events and RMCE events.

Next, we tested the RMCE efficiency between the  $\Phi$ C31 *ubi-GFP* RMCE vector (injected at 250 ng/ $\mu$ l) and the  $\Phi\Phi^{hs-neo}$  allele in fruit fly embryos by providing germline expression of the  $\Phi$ C31 recombinase in three different ways (Figure 1B). First, we crossed  $\Phi\Phi^{hs-neo}/CyO$  males to  $M[vas-int.Dm]ZH-2A$  (*X*) females (BISCHOF *et al.* 2007) and injected 400 of the resulting embryos from this cross. Second, we established a  $M[vas-int.Dm]ZH-2A; \Phi\Phi^{hs-neo}/CyO$  strain and injected 200 embryos laid by these flies. Third, we co-injected the  $\Phi$ C31 *ubi-GFP* RMCE vector together with the pBS130 plasmid (a source of germline  $\Phi$ C31 integrase (GOHL *et al.* 2011)) at 250:100 ng/ $\mu$ l ratio into 200 embryos laid by the  $\Phi\Phi^{hs-neo}/CyO$  strain. We raised the larvae resulting from each injection at 25° and crossed each hatched individual to *yw* flies (we crossed only the non-*CyO* flies hatching from the first injection scenario).

We scored the progeny of each injected fertile individual for successful RMCE events by the ubiquitous expression of GFP from the *ubi-GFP* cassette. Simultaneously, we could detect any integration *vs.* RMCE events through the presence of the *w+* marker in the fly eyes. In the first case, where each individual was a result of the cross between  $\Phi\Phi^{hs-neo}/CyO$  males and  $M[vas-int.Dm]ZH-2A$  females (Figure 1B), we detected 3.3% RMCE events (4/121 individuals) and each RMCE positive parent was segregating equally complete RMCE and integration events. We sequence-verified four RMCE fly lines and the *ubi-GFP* cassette was exchanged in both forward and reverse orientation, as expected.

In the second case, where we injected  $\Phi$ C31 *ubi-GFP* RMCE vector into the  $M[vas-int.Dm]ZH-2A; \Phi\Phi^{hs-neo}/CyO$  established strain (Figure 1B) we detected higher percentage of RMCE events:

6.1% (3/49 individuals). In addition, only one individual was segregating both RMCE and integration events while the other two individuals were segregating only RMCE events.

Interestingly, in the case where the source of integrase was provided through a co-injected vector (pBS130) rather than an established stock (Figure 1B), we detected only integration events, 5.8% (3/52 individuals), and no full RMCE events. However, we found that each integration allele could be lead to a complete RMCE event through intra-molecular recombination between the intact  $\Phi$ C31 attP/attB sites left at the locus. This can occur by introducing/maintaining the integrated allele in the background of the  $M[vas-int.Dm]ZH-2A$  source of integrase. Surprisingly, such events occurred at much lower rate for  $\Phi$ C31 (2/100 progeny) than in the case of Bxb1 recombinase (67/100 progeny (VOUDEV AND MANN 2017)), which might be due to differences in the recombination mechanism between the two recombinase systems (THORPE AND SMITH 1998; GHOSH *et al.* 2003). This property of the  $\Phi$ C31 recombinase might be useful in experiments where a low-rate switch between an integration allele and an RMCE allele is desired.

Taken together, our results show that using minimal attP/attB  $\Phi$ C31 for RMCE is feasible in *D. melanogaster*. Although the rate of RMCE decreases around ten-fold in comparison with the RMCE rates when using longer  $\Phi$ C31 sites (Venken *et al.* 2011), injecting only 200 embryos is sufficient to generate multiple RMCE fly lines and has the advantage of not introducing unnecessary sequences that might interfere with gene/locus function of the engineered allele. Furthermore, in genome editing it is always better to introduce minimal amount of exogenous DNA since other unforeseeable chromatin disruptions may occur. The orientation of the introduced attB/attP sites should also be taken into account in genome engineering: for example, the core of the attP site contains a consensus Trithorax-like (Trl) binding site, GTTCTCTCAG (ZHU *et al.* 2011), which could potentially lead to binding of Trx group proteins and consequent chromatin remodeling of a locus of interest. However, if the attB/attP  $\Phi$ C31 sites are oriented in the manner shown in Figure 1A, this sequence would be eliminated during the recombination reaction and conversion to an attR site (Figure 1A).

Our findings are applicable to many other organisms as the  $\Phi$ C31 recombinase is being widely used and similar considerations over flanking of loci with attB/attP sites are highly relevant in other biological contexts. Analogous analysis of other recombinase systems and sites is recommended in each particular genomic locus engineering case when exogenous sites are being used.

## ACKNOWLEDGMENTS

Thanks to Rainbow Transgenic Flies Inc. for fly injections. R.V. is a Leukemia and Lymphoma Society Fellow. This work was supported by NIH grant R35GM118336 awarded to R.S.M. R.V. conceived this work, performed the experiments and wrote the article. R.S.M. edited the article and supported the study.

## LITERATURE CITED

- Bateman, J. R., and C. T. Wu, 2008 A simple polymerase chain reaction-based method for the construction of recombinase-mediated cassette exchange donor vectors. *Genetics* 180(3): 1763–1766. <https://doi.org/10.1534/genetics.108.094508>
- 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(9): 3312–3317. <https://doi.org/10.1073/pnas.0611511104>
- Delker, R. K., and R. S. Mann, 2017 From Reductionism to Holism: Toward a More Complete View of Development Through Genome Engineering.

- Adv. Exp. Med. Biol. 1016: 45–74. [https://doi.org/10.1007/978-3-319-63904-8\\_3](https://doi.org/10.1007/978-3-319-63904-8_3)
- Ghosh, P., A. I. Kim, and G. F. Hatfull, 2003 The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of attP and attB. *Mol. Cell* 12(5): 1101–1111. [https://doi.org/10.1016/S1097-2765\(03\)00444-1](https://doi.org/10.1016/S1097-2765(03)00444-1)
- 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(3): 231–237. <https://doi.org/10.1038/nmeth.1561>
- Groth, A. C., M. Fish, R. Nusse, and M. P. Calos, 2004 Construction of transgenic *Drosophila* by using the site-specific integrase from phage phiC31. *Genetics* 166(4): 1775–1782. <https://doi.org/10.1534/genetics.166.4.1775>
- Groth, A. C., E. C. Olivares, B. Thyagarajan, and M. P. Calos, 2000 A phage integrase directs efficient site-specific integration in human cells. *Proc. Natl. Acad. Sci. USA* 97(11): 5995–6000. <https://doi.org/10.1073/pnas.090527097>
- Huang, J., W. Zhou, W. Dong, A. M. Watson, and Y. Hong, 2009 From the Cover: Directed, efficient, and versatile modifications of the *Drosophila* genome by genomic engineering. *Proc. Natl. Acad. Sci. USA* 106(20): 8284–8289. <https://doi.org/10.1073/pnas.0900641106>
- Kim, A. I., P. Ghosh, M. A. Aaron, L. A. Bibb, S. Jain *et al.*, 2003 Mycobacteriophage Bxb1 integrates into the *Mycobacterium smegmatis* groEL1 gene. *Mol. Microbiol.* 50(2): 463–473. <https://doi.org/10.1046/j.1365-2958.2003.03723.x>
- Kim, J., K. Johnson, H. J. Chen, S. Carroll, and A. Laughon, 1997 *Drosophila* Mad binds to DNA and directly mediates activation of vestigial by Decapentaplegic. *Nature* 388(6639): 304–308. <https://doi.org/10.1038/40906>
- Reese, M. G., F. H. Eeckman, D. Kulp, and D. Haussler, 1997 Improved splice site detection in Genie. *J. Comput. Biol.* 4(3): 311–323. <https://doi.org/10.1089/cmb.1997.4.311>
- Steller, H., and V. Pirrotta, 1985 A transposable P vector that confers selectable G418 resistance to *Drosophila* larvae. *EMBO J.* 4: 167–171.
- Thorpe, H. M., and M. C. Smith, 1998 In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc. Natl. Acad. Sci. USA* 95(10): 5505–5510. <https://doi.org/10.1073/pnas.95.10.5505>
- Venken, K. J., K. L. Schulze, N. A. Haelterman, H. Pan, Y. He *et al.*, 2011 MiMIC: a highly versatile transposon insertion resource for engineering *Drosophila melanogaster* genes. *Nat. Methods* 8(9): 737–743. <https://doi.org/10.1038/nmeth.1662>
- Voutev, R., and R. S. Mann, 2017 Bxb1 phage recombinase assists genome engineering in *Drosophila melanogaster*. *Biotechniques* 62(1): 37–38. <https://doi.org/10.2144/000114494>
- Zhu, L. J., R. G. Christensen, M. Kazemian, C. J. Hull, M. S. Enuameh *et al.*, 2011 FlyFactorSurvey: a database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res.* 39(suppl\_1): D111–D117. <https://doi.org/10.1093/nar/gkq858>

Communicating editor: B. Reed