NOTE

A Novel Approach for Directing Transgene Expression in *Drosophila*: T2A-Gal4 In-Frame Fusion

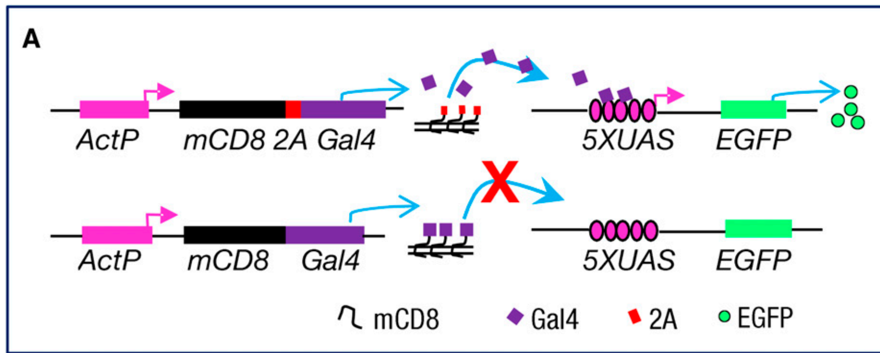
Fengqiu Diao and Benjamin H. White[1]
Laboratory of Molecular Biology, National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland 20892

ABSTRACT In *Drosophila*, the Gal4-UAS system permits a transgene to be expressed in the same pattern as a gene of interest by placing the Gal4 transcription factor under control of the gene's DNA regulatory elements. If these regulatory elements are not known, however, expression of Gal4 in the desired pattern may be difficult or impossible. To solve this problem, we have developed a method for co-expressing Gal4 with the endogenous gene by exploiting the "ribosomal skipping" mechanism of the viral T2A peptide. This method requires explicit knowledge only of the endogenous gene's open reading frame and not its regulatory elements.

ELUCIDATING the functional role of a particular gene often requires manipulating biological processes in the cells that express it. This goal can be accomplished genetically by inducing these cells to also express transgenes that encode products that alter normal cell functions. The most versatile implementation of this strategy is found in binary systems, such as the Gal4-UAS system of *Drosophila*, in which the yeast transcription factor Gal4 is used to drive the expression of a broad array of effectors encoded by other transgenes (Brand and Perrimon 1993; Duffy 2002). To express Gal4 solely in cells that express an endogenous gene of interest, the Gal4-coding sequence is typically placed under the control of the endogenous gene's regulatory elements (*i.e.*, enhancers). When the endogenous gene's enhancers have not been explicitly identified, transgene expression can be directly coupled to expression of the endogenous gene using either enhancer-trap techniques (Brand and Perrimon 1993) or homologous recombination (Rong and Golic 2000; Demir and Dickson 2005; Manoli *et al.* 2005;). In the first instance, transgene constructs randomly inserted into the genome are screened to identify insertions into regulatory regions of the gene of interest and then tested for fidelity of expression. In the second instance, the endogenous gene is simply replaced by inserting the transgene into the endogenous gene's translation start site. Both methods are labor-intensive and prone to failure: the first often results in patterns of transgene expression that fail to fully mimic those of the endogenous gene, and the second requires positive identification of the endogenous gene's start codon, which is often unknown and may differ between splice isoforms. An alternative strategy often used in the mouse yokes expression of the transgene to that of an endogenous gene by interposing a viral *i*nternal *r*ibosomal *e*ntry *s*ite (*i.e.*, IRES) between them (Douin *et al.* 2004). However, no IRES sequences that function robustly in *Drosophila* have been identified (Ye *et al.* 1997). A potentially promising development has been the demonstration that short viral peptides known collectively as "2A-like peptides" can be used, like IRES sequences, to couple transgene expression to the expression of an endogenous gene in mice (Madisen *et al.* 2010; Taniguchi *et al.* 2011).

Viral 2A-like peptides share an Asp-Val/Ile-Glu-X-Asn-Pro-Gly-Pro consensus motif (Donnelly *et al.* 2001a), which, during translation, forces the ribosome to skip from the underlined Gly to the underlined Pro codon without forming a peptide bond (Donnelly *et al.* 2001b). Consequently, the nascent translation product is released after the addition of the glycine residue and a new, independent protein chain is begun with the proline residue. Here, we show that, by inserting a construct consisting of the T2A- and Gal4-coding sequences in-frame into an exon of an endogenous gene, this property of 2A-like peptides can be used to co-express the Gal4 gene and the endogenous gene in *Drosophila*. This method is versatile and allows one to express the Gal4 gene in all cells that express a given gene or only in those that

Copyright © 2012 by the Genetics Society of America
doi: 10.1534/genetics.111.136291
Manuscript received October 31, 2011; accepted for publication December 20, 2011
Supporting information is available online at http://www.genetics.org/content/suppl/2011/12/30/genetics.111.136291.DC1.
[1]Correspondence author: National Institute of Mental Health, NIH, 9000 Rockville Pike, Bethesda, MD 20892. E-mail: benjaminwhite@mail.nih.gov



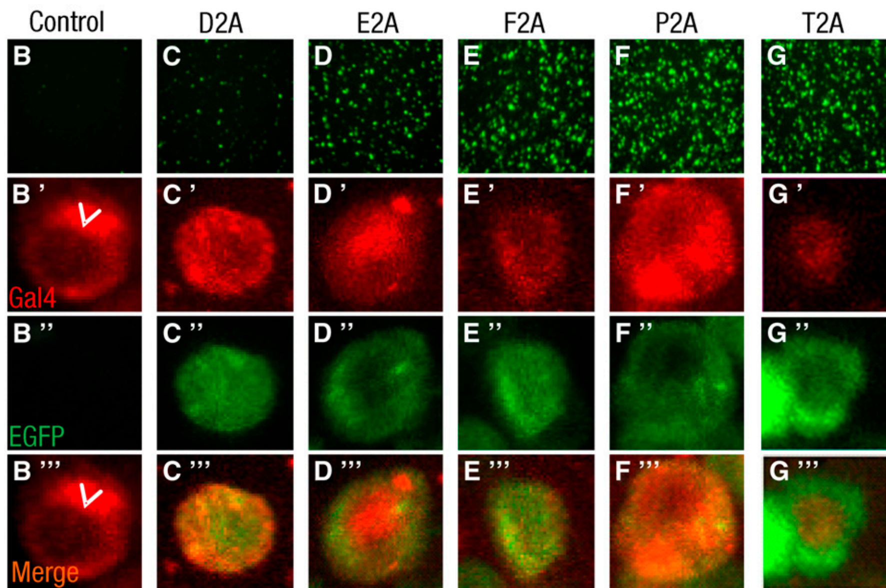


Figure 1 Gal4 separated from the membrane protein mCD8 by T2A-mediated ribosomal skipping drives UAS-EGFP expression in SL2 cells. (A) Schematic of the mCD8-2A-Gal4 construct used to test the capacity of 2A peptides to support ribosomal skipping (top) and the control mCD8-Gal4 construct (bottom). Gal4 released by 2A-mediated ribosomal skipping from the mCD8 translation product of the first type of construct, but not the control construct, should enter the nucleus and drive transcription of UAS-EGFP on a cotransfected reporter plasmid. (B–G) Fluorescence images of EGFP expression in cultured SL2 cells transfected with the control and 2A fusion constructs (indicated in A). Only background EGFP fluorescence is observed in cells transfected with the control construct (B), while robust EGFP expression is observed in cells transfected with the T2A construct (G). (B′–G′′′) Fluorescence photomicrographs showing the following: (B′–G′) anti-Gal4 immunostaining, (B′′–G′′) EGFP fluorescence, and (B′′′–G′′′) merged immunostaining and fluorescence images from representative cells transfected with the control and 2A fusion constructs. In cells transfected with the control construct (B′–B′′′), little Gal4 immunoreactivity is associated with the nucleus (arrowheads) and no EGFP is expressed, while in cells transfected with the T2A construct (G′–G′′′), Gal4 immunoreactivity is strongly localized to the nucleus and accompanied by robust EGFP expression. Consistent with the small percentage of D2A-transfected cells that expressed EGFP (C), many EGFP-negative cells transfected with this construct had Gal4 immunoreactivity outside of the nucleus (data not shown).

express a particular splice variant of that gene. Depending on the site of insertion of the T2A-Gal4 sequence, the translated product of the endogenous gene may be truncated or left functionally intact. This T2A-GIFF (*i.e.*, *T2A-G*al4 *i*n-*f*rame *f*usion) technique thus represents a readily adaptable technique for transgene expression in cells expressing a gene of interest.

Apart from a recent study showing that a 2A peptide from the insect virus *Thosea asigna* known as T2A can support the translation of multiple products from a single transcript in *Drosophila* cell lines (Gonzalez *et al.* 2011), the characterization of 2A peptide activity in *Drosophila* has been limited. To develop the T2A-GIFF approach, we therefore first examined the efficacy of the T2A peptide in promoting ribosomal skipping compared with other candidate 2A peptides when expressed in *Drosophila* SL2 cells. We investigated the efficacy of five different 2A peptides (Supporting Information, File S1, Table S1) using constructs in which the coding sequence of the enhanced green fluorescent protein (*i.e.*, EGFP) or the transcription factor, Gal4, was fused to that of a transmembrane protein, mCD8, separated by an intervening, in-frame 2A peptide sequence (Figure 1A and Figure S1A). Compared to control constructs that lacked the intervening 2A sequences, all five 2A peptides demonstrated at least some promotion of ribosomal skipping in both the EGFP (Figure S1, B–G) and the Gal4 (Figure 1, B–G′′′) assays. The T2A peptide sequence showed the highest efficiency: In SL2 cells transfected with the mCD8-T2A-EGFP construct, EGFP was dispersed throughout the cytosol and nucleus, rather than being localized to the membrane with mCD8 (Figure S1G); in mCD8-T2A-Gal4-expressing cells, Gal4 protein was quantitatively transported to the nucleus (Figure 1, G′ and G′′′) and robustly drove the expression of UAS-EGFP (Figure 1, G–G′′′).

To confirm T2A's ability to support ribosomal skipping *in vivo*, we generated transgenic flies that expressed a membrane-bound mCD8-EGFP-T2A-Gal4 construct only in neurons that express the hormone bursicon (Figure 2A). To detect Gal4 transcription activity, we used flies bearing a UAS-RedStinger reporter transgene. In nervous system whole mounts from these animals, confocal microscopy showed that Gal4 was not tethered to the membrane in the bursicon-expressing neurons and efficiently drove the expression of the UAS-RedStinger reporter (Figure 2, B–B′′′). The lack of tethering of Gal4 was clearly T2A mediated because animals expressing a control construct without the T2A sequence showed only faint mCD8 immunoreactivity from the membrane-associated mCD8-EGFP-Gal4 construct and no UAS-RedStinger expression (Figure 2, C–C′′′).

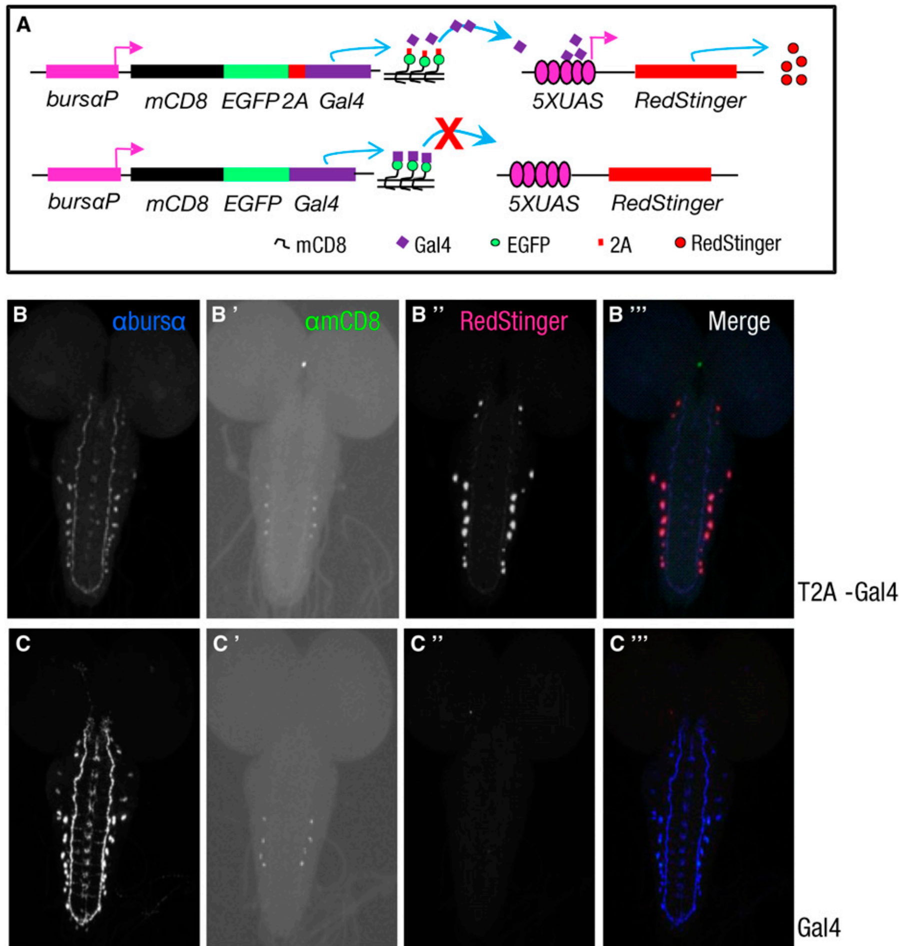


Figure 2 Bicistronic constructs containing the Gal4 transgene can be expressed in targeted cells using the T2A peptide. (A) Schematics of the two constructs used to test the efficacy of the T2A peptide in transgenic flies. As indicated, T2A-mediated ribosomal skipping is expected to cause the Gal4-coding sequence in the mCD8-EGFP-T2A-Gal4 construct (top) to be translated independently of mCD8-EGFP and thus produce transcriptionally active Gal4, whereas the mCD8-EGFP-Gal4 control construct (bottom) is expected to be translated as a single, membrane-bound fusion protein that includes (transcriptionally inactive) Gal4. Both constructs are under control of a promoter that selectively drives their expression in neurons expressing the α-subunit of the hormone bursicon (bursα), and Gal4 activity is monitored by expression of a UAS-RedStinger reporter. (B–B′′′and C–C′′′) Representative confocal images of two CNS whole mounts from third instar larvae expressing the mCD8-EGFP-T2A-Gal4 construct (B–B′′′) or the mCD8-EGFP-Gal4 control (C–C′′′). (B and C) Bursα immunoreactivity in each whole mount shows the bursα-expressing neurons and their processes. (B′ and C′) mCD8-EGFP labeling, as detected by anti-mCD8 immunostaining. (B′′ and C′′) UAS-RedStinger-associated fluorescence. (B′′′ and C′′′) Merged images. Although mCD8-EGFP labeling is observed in preparations expressing both the control and the T2A-containing constructs, only the latter shows RedStinger immunofluorescence, indicating Gal4 transcriptional activity. All images were collected under identical conditions, but contrast and brightness were adjusted in B′ and C′ to enhance mCD8-GFP expression, the low level of which reflects the weakness of the *burs*α promoter. Even with enhancement, expression in some cells remained below the detection threshold of the anti-mCD8 antibody.

The results of both our *in vitro* and *in vivo* studies thus confirm the effectiveness of the T2A peptide in promoting the simultaneous expression of two distinct protein products from a single transcript in *Drosophila* and indicate that the efficacy of T2A in fly neurons is similar to what has been previously reported for mammalian neurons (Tang *et al.* 2009). Importantly for the development of the T2A-GIFF technique, our results also demonstrate that T2A promotes ribosomal skipping during membrane protein translation, when the nascent protein strand is likely to be threaded into the endoplasmic reticulum. Furthermore, they show that the residual proline left at the N terminus of Gal4 upon its separation from the remainder of the T2A peptide does not impair Gal4 transcriptional activity.

To directly validate the T2A-GIFF approach, we implemented it in cells that express the *Drosophila rickets* (*rk*) gene. *rickets* is the fruit-fly ortholog of mammalian stem-cell markers encoded by the *LGR5* and *LGR6* genes (Barker and Clevers 2010). The role of the Rickets protein (RK) in fly stem-cell biology is unknown, but it does have a well-characterized role in promoting adult wing expansion (Baker and Truman 2002). Characterizing RK and the cells that express it has been impeded in the fly by the complexity of the *rk* gene locus. Two splice variants, *rk*-RA and *rk*-RF, with different 5′ start sites, are predicted from the *rk* genomic sequence, and other transcripts have been identified (Eriksen *et al.* 2000; Nishi *et al.* 2000). The longest variant, *rk*-RA, has highly conserved, putative regulatory elements throughout its first, large intron and multiple possible translation start sites, the first of which is preceded by six out-of-frame ATG triplets in the putative 5′ UTR (Eriksen *et al.* 2000). Direct knock-in of the Gal4 transgene into the predicted start site by homologous recombination resulted in no detectable expression of Gal4 (F. Diao and B. H. White, unpublished observations). Gaining genetic access to RK-expressing cells has thus proved challenging, making this an excellent candidate for the T2A-GIFF approach described here.

To target Gal4 expression to *rk*-expressing cells, we inserted the T2A-Gal4 sequence in-frame into the P[acman] genomic clone CH322-119A8 (Venken *et al.* 2006, 2009), which contains the entire predicted coding sequence of the *rk* gene and all the necessary enhancer elements required for *rk* expression (Figure S2). To ensure Gal4 expression in all *rk*-expressing cells, we inserted the T2A-Gal4 sequence into

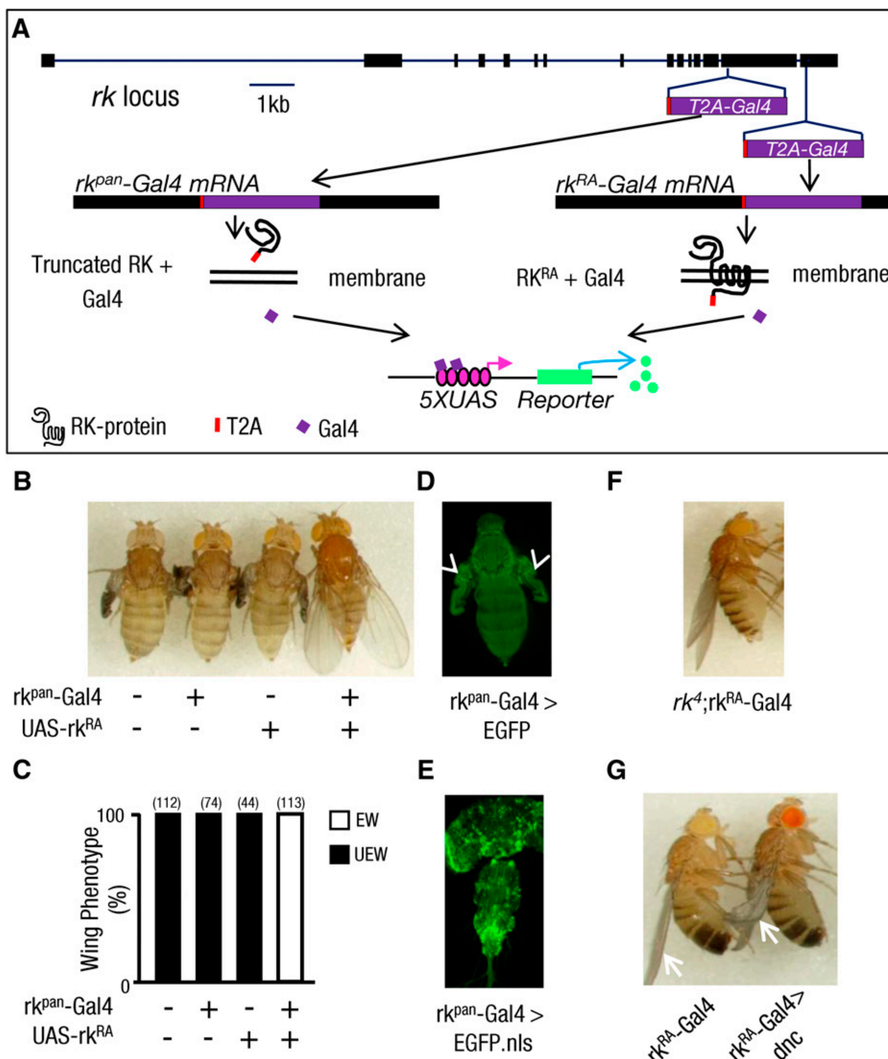


Figure 3 In-frame fusions of T2A-Gal4 can couple Gal4 expression to that of an endogenous gene. (A) Schematic of the intron–exon structure of the *rickets* gene locus. The in-frame insertion sites of the T2A-Gal4 sequence are indicated. Two insertions were made in the *rk* sequence of the genomic P[acman] clone CH322-119A8 (see Figure S2), one within exon 14, which is common to all *rk* transcripts, and one just prior to the stop codon of the *rk*-RA transcript located in exon 15. Transgenic flies bearing the first of these constructs, rk^{pan}-Gal4, produce a truncated, presumably secreted RK protein without transmembrane domains, while flies bearing the second construct produce a full-length RK protein with a C-terminal T2A peptide fusion. (B) Pictures of rk^4 mutant flies (−/−) carrying either the rk^{pan}-Gal4 transgene, a UAS-rk^{RA} rescue transgene, or both. The wing-expansion deficits of rk^4 mutants were rescued only when the UAS-rk^{RA} expression was driven by rk^{pan}-Gal4. (C) Bar graph summarizing the frequency of flies of the various genotypes with unexpanded wings (UEW) *vs.* expanded wings (EW). Total numbers of flies scored is in parentheses above each bar. (D) Fluorescence image of epidermal UAS-EGFP expression driven by rk^{pan}-Gal4 in a newly eclosed fly. Arrowheads indicate prominent labeling of the wing epidermis prior to expansion. (E) Confocal micrograph of UAS-EGFP.nls labeling in the central nervous system driven by rk^{pan}-Gal4 in the pharate adult. (F) Wing-expansion phenotype of an rk^4 mutant fly expressing the rk^{RA}-Gal4 transgene. Rescue of the wing-expansion deficits shows that the RK protein produced by the rk^{RA}-Gal4 construct is functional. (G) A representative fly (right) in which the rk^{RA}-Gal4 transgene drives expression of the *dunce* gene, which encodes a cAMP-phosphodiesterase, to disrupt the second messenger pathway activated by RK protein. Wings (arrows) in such flies remained incompletely expanded, compared with flies not expressing UAS-*dunce* (left), indicating that cells within the rk^{RA}-Gal4 expression pattern contribute to wing expansion.

exon 14 of the *rk* gene, which is shared by all predicted *rk* splice variants, at a highly conserved site just prior to the first transmembrane-spanning region (Figure 3A, rk^{pan}-Gal4). The parent CH322-119A8 clone rescues the wing-expansion deficits of *rk* mutants, but because the T2A-Gal4 insertion should produce a nonfunctional, truncated RK product, we expected the rk^{pan}-Gal4 construct to lack this ability. Consistent with this prediction, wing-expansion deficits were not rescued in flies bearing the rk^{pan}-Gal4 transgene (Figure 3, B and C). However, the rk^{pan}-Gal4 transgene fully rescued wing-expansion deficits when combined with a UAS-*rk*-RA construct encoding the full-length RA isoform of the *rk* gene. This result confirms both the successful release of the Gal4 activity from the upstream RK fragment and the faithful expression of the rk^{pan}-Gal4 driver in the pattern of the endogenous *rk* gene. We further confirmed the fidelity of the rk^{pan}-Gal4 expression pattern by examining the expression of a UAS-EGFP reporter. EGFP was readily detected in tissues known to have RK activity, such as the epidermis and the unfolded wings of newly emerged flies (Davis *et al.* 2007) (Figure 3D), as well as in diverse cells of the central nervous system (Peabody *et al.* 2008) (Figure 3E). The observed tissue distribution of EGFP also closely matched the distribution of *rk* mRNA reported at FlyAtlas (Chintapalli *et al.* 2007), with all tissues having high message levels and showing robust EGFP signals, except the adult fat body (Table 1).

To further demonstrate the utility of the T2A-GIFF approach, we selectively targeted Gal4 expression to cells that express only the *rk*-RA splice variant, by inserting T2A-Gal4 just prior to its stop codon in exon 15 (Figure 3A).

Table 1 Comparison of expression patterns between rk^{pan}-Gal4 and FlyAtlas in adults

Tissue	rk^{pan}-Gal4	FlyAtlas
Brain	++++	+++
Ventral nerve cord	++++	+++
Midgut	++	—
Tubule	+	—
Fat body	—	+
Virgin spermatheca	++++	++++
Mated spermatheca	++++	++++

The number of "+" signs indicate the message levels or relative levels of EGFP fluorescence in different tissues. The expression level of *rk* in FlyAtlas is based on the present call and enrichment of *rk* expression in adults reported at http://flyatlas.org/ (Chintapalli *et al.* 2007).

Unlike rk^{pan}-Gal4, this 3′ terminal insertion construct (*i.e.*, rk^{RA}-Gal4) should yield full-length RK-RF and RK-RA proteins in addition to Gal4. Consistent with this, we found that the rk^{RA}-Gal4 transgene restored wing expansion in *rk* mutants even without driving expression of the UAS-*rk*-RA rescue construct (Figure 3F). However, partial wing-expansion deficits could be induced, along with a variety of other developmental impairments, by rk^{RA}-Gal4-directed expression of a UAS-dunce transgene (Figure 3G, right fly). *dunce* (*i.e. dnc*) encodes a cAMP-phosphodiesterase that is expected to block second-messenger signaling induced by RK activation. The partial effects of *dnc* overexpression on wing expansion may reflect incomplete blockage of the cAMP pathway. Alternatively, the rk^{RA}-Gal4 expression pattern may include many, but not all, cells required for complete wing expansion. Comparison of the rk^{RA}-Gal4 and rk^{pan}-Gal4 expression patterns showed overlap in the wing epidermis and many other tissues (Table 1), but differences at the cellular level may exist and would require more refined analysis to detect.

Taken together, our results show that the T2A-GIFF technique can be used to gain genetic access to otherwise inaccessible cells in *Drosophila* using only the coding sequence of a gene of interest. T2A-GIFF readily allows the full range of Gal4-mediated manipulations of cellular function and can be implemented not only by using recombineered genomic clones as illustrated here, but also by introducing a T2A transgene into the endogenous gene's genomic sequence by homologous recombination (Rong and Golic 2000) or by recombinase-mediated cassette exchange, using, for example, the recently described MiMIC insertions (Venken *et al.* 2011). Finally, we have demonstrated how T2A-GIFF can be used to selectively mark or manipulate cells expressing individual splice variants of a gene of interest and, if desired, to mutate the gene to investigate its role in cellular function. Our approach complements similar 2A-based methods recently described for the cre-lox system in transgenic mice (Madisen *et al.* 2010; Taniguchi *et al.* 2011), and it is readily adaptable for use with other binary expression systems. We thus anticipate that T2A-GIFF and related methods based on in-frame fusions of 2A transgenes will be broadly applicable to the investigation of cellular functions in *Drosophila* as well as other genetic model organisms.

Acknowledgments

We thank Bruce Paterson and Chi-hon Lee for the pPacPL and pC-attB plasmids, respectively; Randall Hewes for UAS-*dnc* flies; and the National Institute of Neurological Disorders and Stroke DNA sequencing facility for sequencing services. We also thank Grace Gray for comments on the manuscript. This research was supported by the Intramural Research Program of the National Institute of Mental Health (Project 1ZIA-MH-002800-07).

Literature Cited

Baker, J. D., and J. W. Truman, 2002 Mutations in the Drosophila glycoprotein hormone receptor, rickets, eliminate neuropeptide-induced tanning and selectively block a stereotyped behavioral program. J. Exp. Biol. 205: 2555–2565.

Barker, N., and H. Clevers, 2010 Leucine-rich repeat-containing G-protein-coupled receptors as markers of adult stem cells. Gastroenterology 138: 1681–1696.

Brand, A. H., and N. Perrimon, 1993 Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118: 401–415.

Chintapalli, V. R., J. Wang, and J. A. T. Dow, 2007 Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat. Genet. 39: 715–720.

Davis, M. M., S. L. O'Keefe, D. A. Primrose, and R. B. Hodgetts, 2007 A neuropeptide hormone cascade controls the precise onset of post-eclosion cuticular tanning in Drosophila melanogaster. Development 134: 4395–4404.

Demir, E., and B. J. Dickson, 2005 fruitless splicing specifies male courtship behavior in Drosophila. Cell 121: 785–794.

Donnelly, M. L., L. E. Hughes, G. Luke, H. Mendoza, E. ten Dam *et al.*, 2001a The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. J. Gen. Virol. 82: 1027–1041.

Donnelly, M. L., G. Luke, A. Mehrotra, X. Li, L. E. Hughes *et al.*, 2001b Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J. Gen. Virol. 82: 1013–1025.

Douin, V., S. Bornes, L. Creancier, P. Rochaix, G. Favre *et al.*, 2004 Use and comparison of different internal ribosomal entry sites (IRES) in tricistronic retroviral vectors. BMC Biotechnol. 4: 16.

Duffy, J. B., 2002 GAL4 system in Drosophila: a fly geneticist's Swiss army knife. Genesis 34: 1–15.

Eriksen, K. K., F. Hauser, M. Schiott, K. M. Pedersen, L. Sondergaard *et al.*, 2000 Molecular cloning, genomic organization, developmental regulation, and a knock-out mutant of a novel Leu-rich repeats-containing G protein-coupled receptor (DLGR-2) from Drosophila melanogaster. Genome Res. 10: 924–938.

Gonzalez, M., I. Martin-Ruiz, S. Jimenez, L. Pirone, R. Barrio *et al.*, 2011 Generation of stable Drosophila cell lines using multicistronic vectors. Sci. Rep. 1: 75.

Madisen, L., T. A. Zwingman, S. M. Sunkin, S. W. Oh, H. A. Zariwala *et al.*, 2010 A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat. Neurosci. 13: 133–140.

Manoli, D. S., M. Foss, A. Villella, B. J. Taylor, J. C. Hall *et al.*, 2005 Male-specific fruitless specifies the neural substrates of Drosophila courtship behaviour. Nature 436: 395–400.

Nishi, S., S. Y. Hsu, K. Zell, and A. J. W. Hsueh, 2000 Characterization of two fly LGR (leucine-rich repeat-containing, G

protein-coupled receptor) proteins homologous to vertebrate glycoprotein hormone receptors: constitutive activation of wild-type fly LGR1 but not LGR2 in transfected mammalian cells. Endocrinology 141: 4081–4090.

Peabody, N. C., F. Diao, H. Luan, H. Wang, E. M. Dewey *et al.*, 2008 Bursicon functions within the Drosophila CNS to modulate wing expansion behavior, hormone secretion, and cell death. J. Neurosci. 28: 14379–14391.

Rong, Y. S., and K. G. Golic, 2000 Gene targeting by homologous recombination in Drosophila. Science 288: 2013–2018.

Tang, W., I. Ehrlich, S. B. Wolff, A. M. Michalski, S. Wolfl *et al.*, 2009 Faithful expression of multiple proteins via 2A-peptide self-processing: a versatile and reliable method for manipulating brain circuits. J. Neurosci. 29: 8621–8629.

Taniguchi, H., M. He, P. Wu, S. Kim, R. Paik *et al.*, 2011 A resource of cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron 71: 995–1013.

Venken, K. J. T., Y. C. He, R. A. Hoskins, and H. J. Bellen, 2006 P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science 314: 1747–1751.

Venken, K. J., J. W. Carlson, K. L. Schulze, H. Pan, Y. He *et al.*, 2009 Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nat. Methods 6: 431–434.

Venken, K. J. T., K. L. Schulze, N. A. Haelterman, H. L. Pan, Y. C. He *et al.*, 2011 MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nat. Methods 8: 737–743.

Ye, X., P. Fong, N. Iizuka, D. Choate, and D. R. Cavener, 1997 Ultrabithorax and Antennapedia 5′ untranslated regions promote developmentally regulated internal translation initiation. Mol. Cell. Biol. 17: 1714–1721.

Communicating editor: P. K. Geyer

GENETICS

Supporting Information
http://www.genetics.org/content/suppl/2011/12/30/genetics.111.136291.DC1

A Novel Approach for Directing Transgene Expression in *Drosophila*: T2A-Gal4 In-Frame Fusion

Fengqiu Diao and Benjamin H. White

Copyright © 2012 by the Genetics Society of America
DOI: 10.1534/genetics.111.136291

File S1

Supporting Materials and Methods

Molecular Biology. The five 2A peptides selected for investigation in this study were the D2A peptide of *Drosophila C* virus, E2A of *Equine rhinitis* A virus, F2A of foot-and-mouth disease, P2A of *Porcine teschovirus1* and T2A of *Thosea asigna* virus (DONNELLY *et al.* 2001). The amino acid sequences of these five 2A peptides are shown in Table S1. Five DNA sequences that encoded the mCD8 transmembrane protein (LEE and LUO 1999) fused to a 2A peptide in a single open reading frame, and incorporating a KpnI restriction between the mCD8 and 2A sequences were synthesized (Epoch Biolabs, Inc. Sugar Land, TX). Following digestion with XbaI and SacI, these constructs were subcloned into the pPacPL vector (a kind gift of Dr. Bruce Paterson; digested with SpeI and SacI) to create the pPacPL-mCD8-2A constructs. The EGFP coding sequence from pEGFP-N2 (BD Biosciences) was subcloned into the pPacPL-mCD8-2A plasmids using the ApaI and NotI restriction sites to generate the five pPacPL-mCD8-2A-EGFP constructs. The pPacPL-mCD8-EGFP control construct was made by digesting both pPacPL-mCD8-2A and pEGFP-N3 (BD Biosciences) with KpnI and NotI and ligating the EGFP fragment into the resulting pPacPL-mCD8-EGFP plasmid.

To generate pPacPL-mCD8-2A-Gal4 constructs, the coding sequence of Gal4 was amplified by PCR using primers that added ApaI (or KpnI) and NotI restriction sites at the 5' and 3' ends. The Gal4 fragment was subcloned into pPacPL-mCD8-2A digested with ApaI or KpnI and NotI, to make the pPacPL-mCD8-2A-Gal4 or pPacPL-mCD8-Gal4 constructs, respectively.

To generate the Bursα-mCD8-EGFP-Gal4 and Bursα-mCD8-EGFP-T2A-Gal4 constructs, a synthetic construct (i.e. Bursα-mCD8-EGFP-T2A-Gal4-Hsp70 pA) including, sequentially, the 252 bp promoter of the bursicon α-subunit gene (PEABODY *et al.* 2008), the mCD8-EGFP-T2A-Gal4 coding sequence, and the Hsp70 polyadenylation signal was generated in pBlueScript with KpnI and SpeI restriction sites at 5' and 3' ends, respectively (Epoch Biolabs, Inc. Sugar Land, TX). Using the KpnI and SpeI restriction sites, the bursα-mCD8-EGFP-T2A-Gal4-Hsp70 construct was subcloned into the pC-attB vector (the kind gift of Dr. Chi-Hon Lee) to obtain pC-attB-Bursα-mCD8-EGFP-T2A-Gal4. The pC-attB-bursα-mCD8-EGFP-Gal4 was obtained by digesting pC-attB-bursα-mCD8-EGFP-T2A-Gal4 with ApaI to excise the T2A sequence.

To generate the UAS-rk^{RA} construct, the coding sequence of the rk-RA cDNA (ERIKSEN *et al.* 2000) was amplified by PCR using primers that introduced an EcoRI restriction site and an optimized translation initiation motif with sequence CAAA immediately before the ATG start codon and a NotI restriction site just after the stop codon. The amplified fragment was then subcloned into the pUAST plasmid for P-element transformation, after EcoRI and NotI restriction digestion.

P[acman] clone CH322-119A8 (VENKEN *et al.* 2009; VENKEN *et al.* 2006) (from BAC/PAC Resources Center of Children's Hospital Oakland Research Institute, Oakland, CA) was used to generate the constructs for rk-Gal4 by recombineering using GalK selection (WARMING *et al.* 2005). The following synthetic primers (IDT, Coralville, Iowa) were used to amplify the GalK targeting

cassette. Sequences representing homologous arms in the *rk* gene are underlined; T2A-Gal4 sequence introduced for higher efficiency recombineering is in bold face:

For rk^{pan}-GalK, rk^{pan}-Galk F:

ACTTCGAGGAGCACGATGTGAGTGGTCCTGCCACGGGATACGGCTTTGGT**GAGGGCCGCGGCAGCCTGCTGACCTGCGGCGATGTGGAGG**CCTGTTGACAATTAATCATCGGCA

and rk^{pan}-GalK R:

ACCGAACCCGGTTGAAAGTCCTCTGTGGACATACCAGAGAATAGTCCAGT**TCATTACTCCTTCTTGGGGTTGGGGGGGGTATCCTCATCA**TCAGCACTGTCCTGCTCCTT

For rk^{RA}-GalK, rk^{RA}-GalK F:

TGGTCCTGCAGCCGGCAAAGCCTCCTCCCGATCCCAACGATGCACCACTATCAGCACTGTCCTGCTCCTT

and rk^{RA}-GalK R: GTCTCCTGGTTGGCCCCACCCAATCTTCGCCAGCGCCAGTCCCACCATCACCTGTTGACAATTAATCATCGGCA.

After obtaining GalK-positive P[acman] clones, the GalK sequence was substituted by the *Drosophilia* codon-optimized T2A-Gal4 sequence. T2A-Gal4 fragments were amplified by PCR with the following primers:

For rk^{pan}-Gal4, rk^{pan}-Gal4 F: TACTTCGAGGAGCACGATGTGAGTGGTCCTGCCACGGGATACGGCTTTGGTGAGGGCCGCGGCAGCCTGCT

and rk^{pan}-Gal4 R: CACCGAACCCGGTTGAAAGTCCTCTGTGGACATACCAGAGAATAGTCCAGTTTACTCCTTCTTGGGGTTGG

For rk^{RA} -Gal4, rk^{RA}-Gal4 F:

TGGTCCTGCAGCCGGCAAAGCCTCCTCCCGATCCCAACGATGCACCACTAGAGGGCCGCGGCAGCCTGCTGACC

and rk^{RA}-Gal4 R: GTCTCCTGGTTGGCCCCACCCAATCTTCGCCAGCGCCAGTCCCACCATCATTACTCCTTCTTGGGGTTGGG.

Plasmid DNAs for all constructs were isolated for fly transformation by S.N.A.P. MidiPrep Kit (Invitrogen, Carlsbad, CA).

Fly Genetics. Transgenic flies were generated using standard injection protocols by either Duke University Model System Genomics Group (Durham, NC) or Rainbow Transgenic Flies, Inc. (Camarillo, CA). UAS-rk^{RA} was generated by P-element transformation. All other constructs were generated by ΦC31-mediated transgenesis into the attP2 site on the 3rd chromosome (GROTH *et al.* 2004). All flies were grown on corn meal-molasses medium and maintained at 25°C in a constant 12 h light–dark cycle. UAS-EGFP, UAS-EGFP.nls, UAS-RedStinger and *$rickets^4$* (*rk^4*) mutants were from the Bloomington Stock Center (Indiana University). UAS-dnc flies were the kind gift of Dr. Randall Hewes.

SL2 Cell Culture. SL2 cells were grown in serum-free HyQ-CCM3 medium (Hyclone, Logan, UT) to a density of 10^6 cells ml^{-1} in 6-well plates (10^6 cells per well). Cells in each well were transfected with 1.0 µg of each DNA construct purified with S.N.A.P.

MidiPrep kit (Invitrogen, Carlsbad, CA) using the Roche FuGENE6 Transfection Reagent (Roche Diagnostics, Indianapolis, IN). pPacPL-mCD8-2A-Gal4 (or PacPL-mCD8-Gal4) and pUAST-EGFP plasmid DNAs were co-transfected in experiments in which Gal4 activity was monitored by UAS-EGFP expression. Cells were analyzed by confocal microscopy after 2 d incubation at 25 °C.

Immunostaining and Image Acquisition. Excised nervous system whole mounts from wandering third-instar larvae or pharate adults were dissected in PBS, and fixed in 4% paraformaldehyde in PBS for 20–30 min, followed by post fixation in 4% paraformaldehyde/PBS plus 0.5% Triton X-100 for 15 min. Primary antibodies were used at the following concentrations: rabbit anti-bursicon α-subunit (PEABODY *et al.* 2008), 1:5000 dilution; mouse anti-mCD8, 1:100 dilution (Invitrogen, Carlsbad, CA); rabbit anti-Gal4, 1:100 dilution (Santa Cruz Biotech, Inc., Santa Cruz, CA). Secondary antibodies used were: AlexaFluor 488 goat anti-mouse, AlexaFluor 568 goat anti-rabbit, and AlexaFluor 680 goat anti-rabbit (all from Invitrogen) were used at 1:500 dilution. Immunolabeled samples were mounted in Vectashield (Vector Laboratories) prior to confocal imaging on a Nikon C-1 confocal microscope. *Z*-series were acquired in 1 µm increments using a 20× objective and 488 nm, and 543 nm, and 633nm laser emission lines for fluorophore excitation. The images shown are maximal projections of volume rendered *z*-stacks of confocal sections collected by incrementally stepping through the entire nervous system.

References

DONNELLY, M. L., L. E. HUGHES, G. LUKE, H. MENDOZA, E. TEN DAM *et al.*, 2001 The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. J Gen Virol **82:** 1027-1041.

ERIKSEN, K. K., F. HAUSER, M. SCHIOTT, K. M. PEDERSEN, L. SONDERGAARD *et al.*, 2000 Molecular cloning, genomic organization, developmental regulation, and a knock-out mutant of a novel Leu-rich repeats-containing G protein-coupled receptor (DLGR-2) from Drosophila melanogaster. Genome Research **10:** 924-938.

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:** 1775-1782.

LEE, T., and L. LUO, 1999 Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. Neuron **22:** 451-461.

PEABODY, N. C., F. DIAO, H. LUAN, H. WANG, E. M. DEWEY *et al.*, 2008 Bursicon functions within the Drosophila CNS to modulate wing expansion behavior, hormone secretion, and cell death. J Neurosci **28:** 14379-14391.

VENKEN, K. J., J. W. CARLSON, K. L. SCHULZE, H. PAN, Y. HE *et al.*, 2009 Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nat Methods **6:** 431-434.

VENKEN, K. J. T., Y. C. HE, R. A. HOSKINS and H. J. BELLEN, 2006 P[acman]: A BAC transgenic platform for targeted insertion of large DNA fragments in D-melanogaster. Science **314:** 1747-1751.

WARMING, S., N. COSTANTINO, D. L. COURT, N. A. JENKINS and N. G. COPELAND, 2005 Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Res **33:** e36.

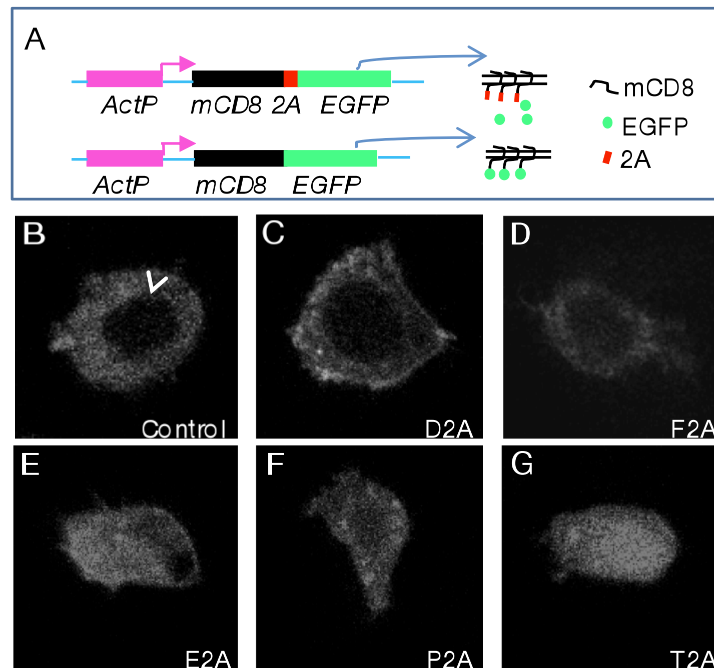


Figure S1 T2A and other 2A peptides facilitate ribosomal skipping in *Drosophila* SL2 cells. (A) Schematic of the mCD8-2A-EGFP construct used to test the faciliation of ribosomal skipping by five 2A peptides (top) and the control mCD8-EGFP construct (bottom). Both constructs are expressed under the control of the *actin* promoter. (B-G) The expression and localization of EGFP in cells transfected with the constructs shown in (A). The images are confocal micrographs of the EGFP distribution in representative individual cells. (B) Tethering EGFP to the membrane by fusing it to mCD8 restricts access of the fluorophore to the nucleus (arrowhead), as is evident in cells made with the control construct lacking a 2A peptide. Constructs made with the D2A (C) or F2A (D) peptides had non-nuclear labeling, similar to controls indicating poor facilitation of ribosomal skipping. In contrast, transfection with the E2A (E), P2A (F), and T2A (G) constructs showed nuclear EGFP localization, with the T2A construct yielding the most uniform labeling patterns suggesting high levels of ribosomal skipping, and the production of soluble, rather than membrane-bound, EGFP. The 2A peptides tested are as indicated in the Materials and Methods.

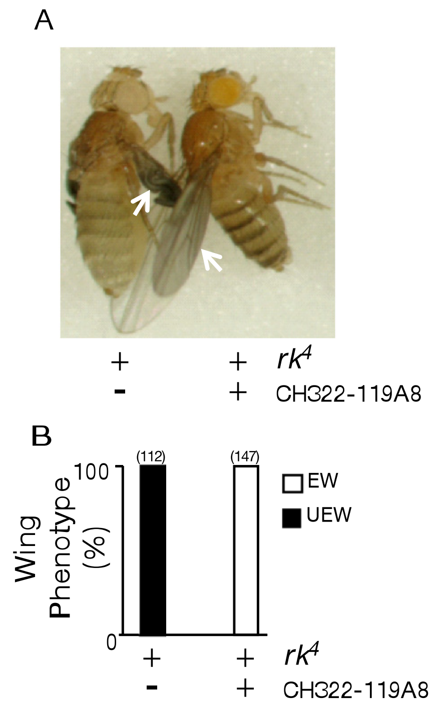


Figure S2 P[acman] clone CH322-119A8, which includes the entire *rk* gene locus as well as 5' and 3' flanking regions, rescues the wing expansion deficits of rk^4 mutants. (A) Unlike flies carrying only the rk^4 mutation (left), flies also bearing a genomic copy of CH322-119A8 (right) expand their wings normally (arrows). (B) Bargraph summarizing the success of rescue with CH322-119A8. Numbers in parentheses indicate the number of flies scored of each genotype.

Table S1 Sequences of 2A peptides tested in this study

2A peptide	Peptide sequence
D2A	AARQMLLLLSG**D**V**E**TN**PG** **P**
E2A	QCTNVALLKLAG**D**V**E**SN**PG** **P**
F2A	VKQTLNFDLLKLAG**D**V**E**SN**PG** **P**
P2A	ATNFSLLKQAG**D**V**E**EN**PG** **P**
T2A	EGRGSLLTCG**D**V**E**EN**PG** **P**

The one-letter amino acid code is used and the conserved –DxExNPG P– motif, which promotes ribosomal skipping, is shown in bold. Sequences are from Donnelly et al. 2001a, as cited in the main text.