King's Research Portal

DOI:
10.1073/pnas.0607652103

Document Version
Publisher's PDF, also known as Version of record

Link to publication record in King's Research Portal

Citation for published version (APA):

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 20. Oct. 2018
An efficient promoter trap for detection of patterned gene expression and subsequent functional analysis in *Drosophila*

Camilla Larsen*, Xavier Franch-Marro†, Volker Hartenstein*, Cyrille Alexandre*, and Jean-Paul Vincent‡

*University of California, Life Sciences Building 4214, 621 Charles E. Young Drive South, Los Angeles, CA 90025; and †National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom

Communicated by Matthew P. Scott, Stanford University School of Medicine, Stanford, CA, September 11, 2006 (received for review January 12, 2006)

Transposable elements have been used in *Drosophila* to detect gene expression, inactivate gene function, and induce ectopic expression or overexpression. We have combined all of these features in a single construct. A promoterless GAL4 cDNA is expressed when the construct inserts within a transcriptional unit, and GAL4 activates a GFP-encoding gene present in the same transposon. In a primary screen, patterned gene expression is detected as GFP fluorescence in the live progeny of dysgenic males. Many animals expressing GFP in distinct patterns can be recovered with relatively little effort. As expected, many insertions cause loss of function. After insertion at a genomic location, specific parts of the transposon can be excised by FLP recombinase, thus allowing it to induce conditional misexpression of the tagged gene. Therefore, both gain- and loss-of-function studies can be carried out with a single insertion in a gene identified by virtue of its expression pattern. Using this promoter trap approach, we have identified a group of cells that innervate the calyx of the mushroom body and could thus define a previously unrecognized memory circuit.

Results and Discussion

**Design and Features of the Promoter Trap.** A transposon carrying a promoterless cDNA accurately reflects endogenous gene expression when integrated downstream of a genomic transcription start site (11). However, flies carrying this construct have to be crossed to a GFP expressing reporter line to reveal the expression pattern in live animals. To allow the screening of new patterns in the first generation, we included UAS-GFP within an analogous GAL4-based construct (Fig. 1). Because the original construct by Lukacsovich et al. (11) was shown to trap promoters, the sequences upstream of GAL4 were kept the same, including a splice acceptor site (SA) and a so-called stop-start site (one small variation was added; see below). Because GFP and GAL4 are both present in our construct, the activity of endogenous promoters should be detectable in the first generation progeny of dysgenic animals. Moreover, because insertion of the transposon introduces three transcription termination sites (one downstream of GAL4, one in *white*, and one following GFP), it is expected that transcription of the endogenous gene would be prematurely terminated, thus leading to loss of function. To enable gain-of-function experiments, we introduced sequences that allow easy conversion to an inducible misexpression construct (after insertion at a specific genomic location). Conversion was achieved by introducing FLP recombination target (FRT) variants at suitable positions such that both GAL4 and GFP could be excised. A pair of mutated FRTs (called FRT2 here), which are incompatible with the wild-type FRT but pair with each other in *vitro* (12), were introduced on both sides of the

Author contributions: C.A. and J.-P.V. contributed equally to this work; C.L., C.A., and J.-P.V. designed research; C.L., X.F.-M., and C.A. performed research; C.A. contributed new reagents/analytic tools; C.L. and V.H. analyzed data; and C.L., X.F.-M., and J.-P.V. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviation: FRT, FLP recombination target.

*To whom correspondence should be addressed. E-mail: jp.vincent@nimr.mrc.ac.uk.

© 2006 by The National Academy of Sciences of the USA

www.pnas.org/cgi/doi/10.1073/pnas.0607652103

www.pnas.org

PNAS | November 21, 2006 | vol. 103 | no. 47 | 17813–17817
coding sequence. Another pair of FRTs (FRT1), also incompatible with the wild-type FRT as well as with FRT2, was placed on both sides of the GFP coding sequence (including the polyadenylation site). In theory, FLP expression should excise both GAL4 and GFP while leaving in place the intervening sequence, which includes a miniwhite gene (as a marker) and the UAS-promoter cassette. The latter, which drives GFP expression before excision, should now point downstream into nearby genomic sequences. Excision should allow expression of the downstream gene in the presence of exogenous GAL4 (which would be brought in by a genetic cross). Overall, we expect the transposon to reveal active promoters by triggering transcription and hence GFP expression. By design, therefore, only insertions downstream of active endogenous promoters will be detected, and sequencing of flanking sequences after inverse PCR should reveal active promoters by triggering transcription and the subsequent production of GFP. A splice acceptor site (SA; AATTCTTAT CCTT-TCCTTTAGCTAACGCCGAGGCCAGAAA) and a stop/start (TGATTGAATAAACATG) precede GAL4 as in the construct of Lukacsovich et al. (11). Both GAL4 and GFP are individually flanked by modified 35-bp FRTs (FRT2 and FRT1, respectively). The central core sequence (shown in the figure for FRT2 and FRT1), which determines specificity, was modified from the wild type (TCTAGAAA) to prevent cross-reactivity while still allowing self-pairing. After FLP expression, both GAL4 and GFP are expected to be excised, leaving all other sequences intact, including the miniwhite gene.

Testing the Conversion from Promoter Trap to Misexpression Construct. Because the specificity of the mutated FRT pairs present in our construct had not been tested previously in a heterologous system, we assessed the effect of expressing FLP in three lines (as a marker) and the UAS-promoter cassette. The latter, which drives GFP expression before excision, should now point downstream into nearby genomic sequences. Excision should allow expression of the downstream gene in the presence of exogenous GAL4 (which would be brought in by a genetic cross). Overall, we expect the transposon to reveal active promoters by triggering transcription and hence GFP expression. By design, therefore, only insertions downstream of active endogenous promoters will be detected, and sequencing of flanking sequences after inverse PCR should unambiguously identify the tagged gene. In the cases in which the transposon inserts upstream of the translation start, it should be readily convertible into a misexpression construct after induction of GAL4 expression.

A Pilot Screen. To assess the activity and efficiency of our promoter trap, we mobilized it and screened for GFP expression in embryos and larvae. A silent insertion (no GFP expression) located on the third chromosome was used as a transposon source. To achieve good gamete representation, one dysgenic male was mated to 10 wild-type females. Depending on capacity, embryos from ~3,000 females were screened each day, and a new GFP expression pattern is seen in 1 male out of 10. Thus, ~30 new expression patterns per 3,000 females were identified. Approximately 100 lines were established from single fluorescent animals isolated during a pilot screen. Various expression patterns were selected (examples are shown in Fig. 3). In some

Fig. 1. Schematic representation of the promoter trap after it has inserted into an individual gene. Flanking genomic regions are shown in red with an arrow marking the endogenous start of transcription. The ends of the P element are indicated by black triangles. After insertion of the transposon, the endogenous promoter (UAS, upstream activating sequence) triggers transcription of and the subsequent production of GFP. A splice acceptor site (SA; AATTCTTAT CCTT-TCCTTTAGCTAACGCCGAGGCCAGAAA) and a stop/start (TGATTGAATAAACATG) precede GAL4 as in the construct of Lukacsovich et al. (11). Both GAL4 and GFP are individually flanked by modified 35-bp FRTs (FRT2 and FRT1, respectively). The central core sequence (shown in the figure for FRT2 and FRT1), which determines specificity, was modified from the wild type (TCTAGAAA) to prevent cross-reactivity while still allowing self-pairing. After FLP expression, both GAL4 and GFP are expected to be excised, leaving all other sequences intact, including the miniwhite gene.
Flp-mediated conversion would allow inducible expression of the transposon. For 20 lines, the exact insertion site was determined by inverse PCR (Fig. 5 and Table 3, which are published as supporting information on the PNAS web site). Of particular significance are the patterns of expression detected in the brain (see below).

More than one-half of the insertions were located on the third chromosome (Table 1), perhaps a consequence of the fact that our jumpstarter transposon was on the third chromosome and that local jumping is usually favored (14, 15). After our pilot screen, two silent insertions have been introduced on a marked second chromosome. Such a strain could be used as an alternative jumpstarter line that might favor insertions on the second chromosome. Such a strain could be used as an alternative jumpstarter line that might favor insertions on the second chromosome. Twenty-five percent of our insertions were homozygous lethal (Table 1), indicating that integration of the transposon made without the splice acceptor site, an analogous construct was made without the splice acceptor. It is likely that this could lead to a higher proportion of GFP-expressing animals because out of 21 independent transformants, 6 expressed GFP in specific patterns (no GFP-producing lines were obtained from the original transformants carrying the original jumpstarter trap). A large-scale mobilization experiment will be needed to confirm that efficiency is increased in the absence of a splice acceptor.

lines, GFP expression was restricted, e.g., to the ring gland as in line 71. Other lines have a broad expression profile such as in line 50, which expresses GFP in many tissues such as the gut, trachea, and epidermis. Although the screen was carried out with embryos and early larvae, many insertions produce fluorescence throughout the life of the fly. GFP expression was recorded and is summarily described in Table 2, which is published as supporting information on the PNAS web site. Overall, it appears that the jump starter trap provides a good reporter of endogenous gene activity. The jump starter trap vector was found to insert equally into introns and exons (Table 1). Interestingly, for 19 of the 20 lines, the jump starter trap inserted upstream of the ATG (within the 5′ UTR). In all of these cases, conversion to a misexpression insertion should therefore be possible. Furthermore, some of these lines are homozygous lethal, indicating that insertion of the promoter trap into the 5′ UTR can cause disruption to gene expression. Because there is no obvious benefit from the splice acceptor site, an analogous construct was made without the splice acceptor. It is likely that this could lead to a higher proportion of GFP-expressing animals because out of 21 independent transformants, 6 expressed GFP in specific patterns (no GFP-producing lines were obtained from the original transformants carrying the original jump starter trap). A large-scale mobilization experiment will be needed to confirm that efficiency is increased in the absence of a splice acceptor.
Embryonic expression in scattered cells throughout the embryo. (G) Line 50 shows expression in the tracheal system. (D) Line 4 is expressed in the CNS and the oenocytes. (E) Line 56 shows expression in the tracheal system. (F) Line 86 shows GFP expression in the fat bodies and somatic musculature. (G) In embryos from line 98, GFP is seen in narrow stripes of segmentally repeated epidermal cells. There is also expression in scattered cells throughout the embryo. (H) Line 35 exhibits exclusive expression in the ring gland. (I) Line 90 has GFP expression in segmentally repeated cells of the embryonic epidermis.

Identification of a Novel Group of Cells Innervating the Calyx. As shown above, GFP expression from the promoter trap faithfully reports on the normal endogenous pattern. We expected faithfulness to exceed that of enhancer traps because, with the promoter trap, GAL4 is driven from a fusion transcript. GFP expression was also expected to be strong, thanks to the GAL4-driven amplification step. The strength of the signal turned out to be particularly useful in cases in which expression is confined to a small number of cells. Strong expression can be seen, for example, in larvae carrying the promoter trap in odd-skipped, which is only expressed in ≈5–6 neurons in each brain lobe at the third instar (Fig. 4E). The ease of visualizing these groups of cells led us to postulate that genes expressed in restricted patterns in the larval brain could be screened for at the outset. The feasibility of this approach was tested by small-scale mobilization of the promoter trap and screening for GFP expression in the brain of larval progeny. Some examples are shown in Fig. 4 A–E. In these lines, neuronal processes can later be visualized by crossing the lines to flies carrying UAS-CD8-GFP. Fig. 4A shows the pattern from a promoter trap in vha16, a gene encoding a proton-transporting ATPase (http://flybase.bio.indiana.edu). As can be seen, the trap is expressed within the brain lobes in a restricted number of cells that project predominantly into the nerve cord. Another line (line 42) was found to be expressed exclusively in surface glia (Fig. 4B). To our knowledge, no such expression pattern has been reported previously. Fig. 4C shows the expression pattern from line 2 (B4), which is expressed almost exclusively in a group of cells located in an anterior-medial region of the brain lobe. Broader expression patterns were also identified in Fig. 4D.

The insertion into the odd-skipped gene is expressed at the second instar in a cluster of 5–6 neurons in each brain lobe. With CD8-GFP, these cells are seen to project into the calyx of the mushroom body (white arrow in Fig. 4E) (for 3D visualization, see Movie 1, which is published as supporting information on the PNAS web site). This finding is significant because all of the cells known so far to innervate the calyx are Kenyon cells, yet the odd-skipped cells (white arrow in Fig. 4F) are clearly distinct from the Kenyon cluster: As can be seen in Fig. 4F, odd-skipped-expressing cells (white arrow) do not colocalize with the Dachshund-positive Kenyon cell (light blue arrow). Therefore, the odd-skipped cells may define a previously unrecognized memory circuit. Not only was the promoter trap instrumental in identifying these cells, it also provides tools for future characterization. For example, the ontology of these cells could be uncovered by

Fig. 3. Some examples of GFP patterns. The number at the bottom of each panel refers to individual insertion lines. B, D, and G show photographs of live embryos, and the remaining panels depict live larvae. (A) Line 34 shows expression in the CNS and peripheral nervous system at larval stages. (B) Embryonic expression in segmentally arranged clusters of cells of the epidermis. (C) Line 50 shows expression in the tracheal system, intestinal tract, and the epidermis. (D) Line 4 is expressed in the CNS and the oenocytes. (E) Line 56 shows expression in the tracheal system. (F) Line 86 shows GFP expression in the fat bodies and somatic musculature. (G) In embryos from line 98, GFP is seen in narrow stripes of segmentally repeated epidermal cells. There is also expression in scattered cells throughout the embryo. (H) Line 35 exhibits exclusive expression in the ring gland. (I) Line 90 has GFP expression in segmentally repeated cells of the embryonic epidermis.

Table 1. Insertion sites for some of the lines isolated in a pilot screen

<table>
<thead>
<tr>
<th>P element insertions on</th>
<th>3rd</th>
<th>2nd</th>
<th>X</th>
<th>Homozygotes lethal insertion</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 110</td>
<td>62</td>
<td>27</td>
<td>21</td>
<td>27</td>
</tr>
</tbody>
</table>

A majority of insertions were found on the 3rd chromosome. Twenty-seven percent were homozygous lethal.
tracking them in live and fixed preparations. Our preliminary analysis suggests that these cells originate from a posterior cluster within the embryonic brain. In addition, GAL4 produced by the promoter trap could be used to drive the expression of additional markers such as CD2-HRP (7) to facilitate connectivity studies at the EM level. A toxin could also be expressed for cell ablation (17). Ablation could be done at a defined time if GAL80 [15] is introduced to allow the control of GAL4 activity with temperature (18).

Conclusion
Our construct allows the efficient identification of genes that are expressed in specific patterns within a tissue. Morin et al. (19) reported a promoter trap vector that generates fusion proteins between endogenous gene products and GFP. GFP fusions can report on the subcellular localization of the endogenous product. Seeing the subcellular localization is a distinct advantage, but it comes at the cost of the need for three constructs (one for each frame). Because of its simplicity, our promoter trap allows the identification of the disrupted gene because the insertions occur within the transcription unit. All in all, the gene and promoter traps have increased by the use of a larva/embryo sorter. Efficiency and the ability to screen in the first generation after dysgenesis is such that one can afford to select only the desired patterns of expression for further analysis. Another benefit of our promoter trap is that, as a result of GAL4-mediated amplification, the GFP signal is readily detectable even if expression is restricted to a small number of cells. A further advantage of our construct is that it easily lends itself to gain- and loss-of-function analysis. Both gene and promoter traps allow the identification of the disrupted gene because the insertions occur within the transcription unit. All in all, the gene and promoter traps have distinct benefits (chiefly the creation of GFP fusions for the promoter trap and the bright signal and simplicity for the promoter trap) and hence should complement each other.

As we have shown, our promoter trap can be used to identify genes expressed in a particular tissue at any developmental stage of interest even if the cell population is very small. A proof-of-concept is provided by genes expressed in a subset of cells within the brain. Systematic screening for such expression patterns could provide a palette of tools to probe the development and function of various parts of the brain. Mutations in head genes such as orthodenticle, empty spiracles, tailless, and buttonhead (20, 21) have been known to cause large-scale deletions in the brain. However, relatively few mutations are known to affect restricted neuronal circuits in either embryos or larvae. The features of our promoter trap should help characterizing such circuits. As an assessment of this paradigm, it would be useful to test the role of the odd-skipped cluster and probing the role of odd-skipped in the development and function of mushroom bodies.

Materials and Methods
Construct of the Transposon. Standard techniques of molecular biology were used. The full sequence of the transposon is available upon request.

Fly Stocks and Mating Crosses. The jumpstarter stock used for our pilot screen carried a silent insertion on the third chromosomes. Mobilization was carried out by crossing flies containing the P element to flies carrying Δ2–3 as a source of transposase (20). Male progeny from this cross were mated to females from a white virginizer stock containing heat shock-hid on the Y chromosome (a gift from Ruth Lehman, Skirball Institute, New York, NY). Progeny were screened by using a standard dissecting microscope with a fluorescence source or a COPAS embryo sorter (Union Biometrica, Holliston, Massachusetts). The following fly stocks were used for further experiments: engrailed-GAL4 (a gift from Andrea Brand, Cambridge University, Cambridge, U.K.) and a strain carrying FLP under the control of a testis-specific tubulin promoter (13).

Primers to Test FLP-Mediated Recombination. The region was amplified with CCGACATTGACGCTAGGTAA and GGATT-TGCCATTGATCTTTCG, whereas GFP was amplified with CCGTTCGAGTGATTAGGT and CACGTGCGAGGTGTGATT.

Inverse PCR. Inverse PCR was performed following the Berkeley Drosophila Genome Project standard protocol (as described in FlyBase) except that alternative primers were used in most cases: GGAGGCCGACTACAAGCGAGT and CACCCAAAGC-TCTGCCCCCACAT. In some cases, reliable PCR fragments were generated with the primers suggested by the Berkeley Drosophila Genome Project.


Larsen et al.