A multiplexable TALE-based binary expression system for in vivo cellular interaction studies

Markus Toegel1, Ghows Azzam1,3, Eunice Y. Lee1,4, David J.H.F. Knapp1, Ying Tan2, Ming Fa2 & Tudor A. Fulga1

Binary expression systems have revolutionised genetic research by enabling delivery of loss-of-function and gain-of-function transgenes with precise spatial-temporal resolution in vivo. However, at present, each existing platform relies on a defined exogenous transcription activator capable of binding a unique recognition sequence. Consequently, none of these technologies alone can be used to simultaneously target different tissues or cell types in the same organism. Here, we report a modular system based on programmable transcription activator-like effector (TALE) proteins, which enables parallel expression of multiple transgenes in spatially distinct tissues in vivo. Using endogenous enhancers coupled to TALE drivers, we demonstrate multiplexed orthogonal activation of several transgenes carrying cognate variable activating sequences (VAS) in distinct neighbouring cell types of the Drosophila central nervous system. Since the number of combinatorial TALE–VAS pairs is virtually unlimited, this platform provides an experimental framework for highly complex genetic manipulation studies in vivo.
Binary modular expression systems have revolutionised our ability to manipulate and dissect regulatory networks in vivo. Foremost, the yeast-derived GAL4–UAS system has become an invaluable tool for disentangling intricate genetic interactions in *Drosophila*¹, mice³, and zebrafish⁵. However, the extraordinary complexity of many biological contexts, such as the nervous system, requires the design of more sophisticated tools allowing for parallel and controlled manipulation of multiple tissues and cell types. For example, the field of connectomics, dedicated to deciphering neuronal circuits, would greatly benefit from the ability to simultaneously label different populations of neurons and to control their excitatory state (activation or inhibition) independent of each other. To enable such studies, two other binary expression systems have been developed in *Drosophila*, based on either the viral DNA binding protein LexA, which recognises the LexA operator⁴, or the fungal DNA binding protein QF acting on the QUAS sequence⁵,⁶. Although valuable, each of these systems relies on a single exogenous DNA binding protein (driver) that acts on a unique recognition motif (responder), and thus cannot be used to independently express multiple transgenes in separate tissues. Consequently, the only present modality to carry out complex cellular interaction studies

Fig. 1 Design and optimisation of TALE–VAS driver-responder pairs in *Drosophila* S2 cells. **a** Schematic representation of the TALE–VAS system. TALEs consist of a N-terminal domain (TALE N) containing a nuclear localisation sequence (NLS), a custom TALE-array, and the C-terminal domain (TALE C′) fused to a transcription activation domain (VP64). Each repeat within the TALE-array recognises a specific nucleotide depending on the amino acids 12 and 13 of that repeat (NG = T, NI = A, NN = G, HD = C). Transcription of the gene of interest is initiated upon binding of the TALE to its cognate 20 nt long VAS. The invariable 5′ T at position 0 is shown in blue. **b, c** Comparative analysis of reporter expression between four engineered TALE–VAS driver-responder pairs. Diagrams indicate the experimental conditions tested, reflecting the drivers used (circles) and their respective responders (shapes with matching indentations)—colours and numbers indicate the four different TALE–VAS pairs (b). **c** Flow cytometry analysis of reporter expression in *Drosophila* S2 pMT-GAL4 cells. The original backbone TALE plasmid without a TALE-array (TALECTR) was used to establish background levels of GFP expression. **d–g** Specificity of transgene activation by TALE drivers. Each VAS responder was co-transfected with TALE1, 3, 4 drivers and the control TALE in

![Flow cytometry analysis of reporter expression in *Drosophila* S2 pMT-GAL4 cells.](image-url)
is to combine different expression systems. Such experiments, however, have so far been restricted to parallel analyses involving maximum two platforms, such as GAL4 and LexA or GAL4 and QF. Furthermore, to increase the number of consecutive manipulations, more such orthogonal driver–responder pairs would have to be adopted from different organisms and engineered to render them compatible with in vivo applications and to avoid cross-platform interactions.

To address this limitation and simplify the range of tools, we have developed a customisable binary expression system based on bacterial transcription activator-like effector (TALE) fusion proteins. We reasoned that since TALEs can be programmed to bind any DNA sequence, this platform could enable the generation of virtually unlimited driver–responder pairs suitable for independent parallel transgene expression in vivo. By analogy to the GAL4 upstream activating sequence (UAS), we have termed the programmable TALE responder binding sites variable activating sequences (VAS). Using this logic, we demonstrate that orthogonal TALE drivers and VAS responders can be combined in vivo to enable simultaneous expression of multiple transgenes in interacting tissues, without any apparent crosstalk between individual pairs. Furthermore, we show that the TALE–VAS system can be used in conjunction with other binary expression platforms, further expanding the promise of this technology for complex cellular interactions studies.

Results
Design and optimisation of TALE–VAS pairs in vitro. TALEs were discovered in the bacterial plant pathogen Xanthomonas where they act as secreted effector proteins that induce host gene expression. DNA binding is achieved through arrays of 34 amino acid long repeats each containing a repeat-variable dis-residue (RVD) at position 12 and 13, which provide the code for nucleotide (nt) recognition in the major groove of the DNA (Fig. 1a). In order to assess the general feasibility of our approach, we first generated four different TALE drivers based on the pJC-TALE-VP64 vector, which is optimised for GAL4-mediated transgene expression in Drosophila cells (Supplementary Fig. 1). The four corresponding VAS responders (VAS1–EGFP, VAS2–EGFP, VAS3–EGFP, VAS4–EGFP) were created by cloning TALE target sequences upstream of an EGFP reporter construct in the Drosophila pFRC81 vector scaffold (Supplementary Fig. 2). To ensure high specificity and strong affinity, TALEs were designed to recognise 20 nt long VAS sequences.

Although all current RVDS have high affinity for specific nucleotides, some can tolerate less favourable interactions. Consequently, it has been reported that under certain circumstances both natural and synthetic TALEs could bind to off-target sequences containing up to three base pair mismatches. This is an important consideration for the use of TALEs as programmable nucleases in genome editing applications, where introduction of DNA double-strand breaks at off-target sites can result in unintended mutagenesis events. In the TALE–VAS system confounding off-target effects could manifest in the ectopic activation of endogenous genes. However, the risk of such events is mitigated to some degree by the requirement of TALEs to bind in close proximity to a genomic promoter or enhancer. To identify putative VAS off-target sites, we carried out a comprehensive bioinformatics search for sequences containing up to four nt mismatches across a 4 kb window centred on the transcription start sites (TSS) of all annotated Drosophila genes. This analysis revealed that none of the four TALE drivers had perfect matching endogenous targets or off-target sites containing single-nucleotide mismatches (Supplementary Fig. 3a, b and Supplementary Data 1). Furthermore, only TALE2 had a small number of possible off-targets containing two and three nt mismatches. TALE1, TALE3, and TALE4, however, displayed no relevant off targets with up to three nt mismatches, rendering them well suited for subsequent in vivo applications.

GAL4–UAS responder lines usually carry 5, 10, or 20 UAS copies translating into increasing strength of transgene activation. To assess the requirements of the TALE–VAS system, we first used TALE1 to drive EGFP expression downstream of 1, 5, or 10 VAS repeats. These constructs were tested in Drosophila S2 cells harbouring a stably integrated Cu²⁺-inducible metallothionein promoter GAL4 (pMT-GAL4), which was used to drive TALE expression. In contrast to the GAL4–UAS system, fluorescence levels quantified by flow cytometry already saturated at 5 repeats suggesting that strong transgene expression could be achieved even with a relatively low number of VAS copies (Supplementary Fig. 4). Since further analysis revealed that reducing the number of VAS repeats to 3 only marginally impacted reporter expression, all final responders were designed with 3 VAS copies to reduce the number of repetitive sequences.

To establish the efficacy of transgene expression across different combinations of drivers and responders, we then tested each TALE–VAS pair in S2 pMT-GAL4 cells (Fig. 1b, c and Supplementary Fig. 5). In addition to lacking relevant off-targets (Supplementary Fig. 3a, b), TALE1, TALE3, and TALE4 showed the strongest EGFP reporter activation and thus were used in all subsequent experiments. To test the orthogonality of TALE–VAS pairs we co-transfected each TALE plasmid with either cognate or non-cognate VAS responder plasmids. In all cases, robust EGFP activation above background levels was only observed when the cognate responder was present (Fig. 1d, e and Supplementary Fig. 6). To benchmark the TALE–VAS constructs against the established GAL4–UAS system, we directly compared the performance of these platforms in the same assay. Since the respective DNA binding proteins have different rates of synthesis, folding, maturation, and degradation, this assay could only provide a qualitative rather than quantitative readout. To minimise experimental interference, we used the same setup in pMT-GAL4 S2 cells by swapping the TALE–VAS responder pair for a GAL4 plasmid and corresponding UAS responder. Under these experimental conditions, the TALE–VAS system rendered comparable transgene activation with the GAL4–UAS system (Fig. 1f, g). Finally, to establish basal responder activity in pMT-GAL4 S2 cells, we measured VAS–EGFP and UAS-EGFP expression in the absence of any driver and CuSO₄. In contrast to the strong activation observed in the presence of the corresponding TALE or GAL4 drivers, both responders alone displayed low background signal above the untransfected control cells (Supplementary Fig. 7).

TALE-mediated activation of VAS-responders in vivo. Having established the viability of driving orthogonal transgene expression by the TALE–VAS system in vitro, we next sought to implement this technology in vivo by creating transgenic driver–responder lines in Drosophila. To achieve tissue specificity for targeted transgene expression, we cloned different enhancer elements upstream of the hsp70 minimal promoter in the TALE1, TALE3, and TALE4 vectors. To demonstrate the versatility of this system, we focused on enhancers that have distinct but also broad expression patterns in neighbouring tissues and can be analysed in third larval instars, which are readily amenable to dissections and imaging. To this end, we generated TALE drivers for mature neurons (elavΔ¹.8kb–TALE1), neuroblasts (as8Δ0.8kb–TALE3), muscle cells (mhc2.4kb–TALE3), and glial cells (repo1.9kb–TALE4) (Fig. 2a, Supplementary Fig. 8, and Methods). Additionally, we generated a new set of responder constructs expressing three tagged
flourophores under the control of corresponding VAS elements (VAS1-Citrine-HA, VAS3-V5-mCherry, and VAS4-FLAG-Cerulean) (Fig. 2a). To reduce the number of responder constructs and facilitate subsequent in vivo applications of this technology, we combined the VAS1 and VAS4 responders in a single vector. All TALE drivers and VAS responders were integrated into the Drosophila genome by PhiC31 integrase-mediated germline transformation using five different landing sites (Supplementary Fig. 9). For each construct, at least two independent transgenic lines were obtained that did not show any obvious differences in their expression pattern and thus were used interchangeably.

Whole mount analysis of the two responder lines (VAS3-V5-mCherry and VAS1-Citrine-HA, VAS4-FLAG-Cerulean) revealed weak background expression in a small number of tissues.
(Supplementary Fig. 10a, b). Since this pattern is highly restricted and reproducible, it may reflect a positional integration effect or the presence of putative transcription factor binding sites within the VAS sequence. Ectopic expression was also observed in the epidermis when VAS transgenes were driven by the mhc<sup>c.2kb</sup>, ase<sup>0.8kb</sup> or elav<sup>1.8kb</sup> TALE driver (Supplementary Fig. 10c). Because the mhc<sup>c.2kb</sup>-TALE<sub>3</sub> and the ase<sup>0.8kb</sup>-TALE<sub>3</sub> are both integrated at the same landing site (VK37), it is conceivable that this additional expression also results from TALEs activation by endogenous regulatory elements associated with the integration site. Although the ectopic expression observed using these driver–responder pairs is unlikely to interfere with functional studies, it could under certain circumstances compromise the interpretation of experiments employing effector transgenes. The weak VAS background may also in some instances impact the ability to establish and maintain responder lines encoding lethal transgenes. If necessary, these issues could be addressed by screening the large VAS sequence space for variants devoid of any baseline activity and choosing alternative or random genomic integrations sites for establishing new TALE drivers and VAS responder lines. These considerations will be particularly important if attempting to generate TALE–VAS libraries, in which case the specificity of all system components should be rigorously assessed and empirically validated across multiple tissues and developmental stages.

To assess the levels of transgene activation and their pattern of expression, each TALE driver was crossed individually to its cognate responder line and third larval instar brains and body wall muscles were analysed by immunofluorescence and antibody staining. In all cases, we observed strong tissue-specific activation in a pattern predominantly overlapping with the expression of the mhc, elav, ase, and repo genes, respectively (Fig. 2b–i). It should be noted that although these enhancers were selected to best reflect the endogenous expression patterns of their associated genes, it is unlikely that the chosen sequence boundaries captured the entire set of native regulatory elements. Consequently, only a partial overlap should be expected between the TALE-drivers expression domains and those of the corresponding endogenous genes (antibody staining).

To confirm that the TALE drivers do not cross-activate non-cognate responders in flies, we generated a triple VAS responder line (VAS<sub>1</sub>-Citrine-HA, VAS<sub>2</sub>-V5-mCherry, VAS<sub>3</sub>-FLAG-Cerulean) by classic recombination, and crossed it to each TALE driver line. Similar to the results in S2 cells, expression was only observed between matching TALE–VAS pairs and no unspecific activation of non-cognate responders was detected in the target tissues, demonstrating the specificity of this system in vivo (Supplementary Fig. 11).

The GAL4–UAS system has been frequently used at different temperatures to either enhance or attenuate transgene expression levels<sup>19</sup>. To assess whether the TALE–VAS system also displays temperature sensitivity, we analysed reporter levels by anti-HA antibody staining in brain and eye disc tissues of elav<sup>1.8kb</sup>-TALE<sub>3</sub> > VAS<sub>1</sub>-Citrine-HA third larval instars from crosses reared at 18 or 25 °C. Although expression was robust at both temperatures, Citrine-HA protein levels appeared higher at 18 °C relative to 25 °C, which is opposite to the GAL4–UAS system behaviour (Fig. 3a, b). The amplitude of this effect appeared to vary between tissues, and only reached statistical significance in the eye discs. To determine if this tendency is attributed to a temperature-induced change in TALE production or in the potency of TALE-VP64 to drive VAS transgene expression, we measured TALE and Citrine transcript levels by RT-qPCR. This analysis revealed that while the Citrine mRNA levels displayed an increase trend at 18 °C compared to 25 °C, TALE expression remained largely unaffected by temperature (Fig. 3c). Together, these results suggest that TALE-VP64 activity is slightly stabilised or enhanced at lower temperatures.

**Multiplexed TALE–VAS transgene expression in the nervous system.** Although the TALE–VAS system is a self-sufficient platform for multiplex transgene expression, the ability to combine it with other binary expression systems could be beneficial under certain circumstances. For example, using TALE–VAS driver–responder pairs in parallel with the extensive repertoire of GAL4–UAS constructs could enable genetic and cellular interaction studies of unprecedented complexity. To test this possibility, we first crossed the ase<sup>0.8kb</sup>-TALE<sub>3</sub>; VAS<sub>1</sub>-V5-mCherry line to the UAS-mCD8-GFP; ase-GAL4 line to directly compare the two systems in the same tissue. Analysis of larval central nervous system (CNS) from progeny animals revealed simultaneous and comparable expression of both transgenes. Although the two drivers harbour different ase enhancer elements (see Methods) and the two transgenes are localised to distinct cellular compartments (cytoplasm and membrane), their pattern of expression in neuroblast cells was predominantly overlapping (Fig. 4a and Supplementary Fig. 12). To assess the orthogonality of these systems, we then generated a triple TALE driver line (elav<sup>1.8kb</sup>-
TALE1, asel0.8kb-TALE2, repo1.9kb-TALE4) and crossed it to the mixed VAS, V5-mCherry; UAS-mCD8-GFP responder line. While asel0.8kb-TALE2 drove robust mCherry expression in neuroblasts, none of the TALE drivers activated the UAS transgene to any detectable levels (Fig. 4b). Conversely, crossing the asel-GAL4 driver to the same mixed responder line resulted in specific activation of only the UAS-mCD8-GFP responder in neuroblasts, indicating that GAL4 drivers are also unable to interact with VAS responders (Fig. 4c). To demonstrate the utility of this approach in driving independent transgenes in interacting tissues, we then expressed one marker in neurons using the TALE–VAS system (elas1.8kb-TALE1 > VAS; Citrine-HA) and a second transgene in neuroblasts with GAL4–UAS (asel-GAL4 > UAS-mCD8-mCherry). In this instance, each driver specifically activated its cognate responder in the corresponding tissues with no crosstalk between the two platforms (Fig. 4d, e).

Next, we sought to establish the potential of this platform for driving parallel activation of multiple transgenes in several neighbouring tissues in vivo. To this end, we first crossed the triple TALE driver line (elas1.8kb-TALE1, asel0.8kb-TALE2, repo1.9kb-TALE4) to the triple VAS responder line (VAS; Citrine-HA, VAS; V5-mCherry, VAS; FLAG-Cerulean), thereby generating a viable animal expressing three distinct transgenes in non-overlapping tissues (Fig. 5a). Confocal microscopy analysis revealed a clear separation between the three target tissues, enabling simultaneous imaging of the neuronal, neuroblast, and surface glia compartments in the same third-instar larval CNS (Fig. 5b, c). Notably, in all cases the level of expression appeared to be qualitatively similar to single TALE–VAS experiments, indicating that parallel activation of multiple transgenes has no influence on the performance of each individual TALE–VAS pair (compare Figs. 2 and 5 panels). The ability to drive specific and robust parallel transgene expression is instrumental for understanding complex interactions within the nervous system, such as the interplay between different classes of neuron, glial cells and neurons, or muscles and neurons. To illustrate the potential of employing TALE–VAS binary expression systems for such studies, we sought to image at high-magnification a single segmental nerve and the larval neuromuscular junction (NMJ). To this end, we crossed the triple VAS responder line either with the elastr1.8kb-TALE1, asel0.8kb-TALE2, repo1.9kb-TALE4 triple TALE, or with a new elastr1.8kb-TALE1, mh2.4kb-TALE2, repo1.9kb-TALE4 line, respectively. This approach enabled specific differential labelling of an axon bundle ensheathed by a glial cell in the segmental nerve, and of the neuronal presynaptic terminal and the postsynaptic muscle field at the NMJ (Fig. 5d, e).
Finally, to validate our off-target bioinformatics analysis in vivo, we compared the activity of TALE$_3$ and TALE$_4$ drivers at on-target VAS sites relative to putative off-target sites located upstream and in close proximity of endogenous TSS (Supplementary Fig. 13a, Supplementary Data 1). We focused on two candidate genes, CG16890 and CG5613, which are expressed at low to moderate endogenous levels in third larval instars (modENCODE Tissue Expression Data; FlyBase). Transcript levels of each corresponding transgene (Cerulean and mCherry) and endogenous gene (CG16890 and CG5613) were measured by RT-qPCR in third larval instars of wild-type, homozygous triple VAS-driver, triple VAS-responder, and triple TALE-driver > triple VAS-responder lines (Supplementary Fig. 13b). As expected, this analysis revealed very potent and specific activation of each responder transgene by the corresponding TALE drivers.

In contrast, no apparent changes in expression were detected between control and experimental conditions at the two endogenous off-target loci (Supplementary Fig. 13b).

**Discussion**

Manipulating several tissues in parallel by consecutive expression of distinct transgenes has been challenging so far. We reasoned that this limitation could be addressed by integrating programmable DNA binding proteins in the design of binary expression systems. To this end, we have developed a highly versatile and multiplexable TALE-based driver–responder technology in *Drosophila*, which could facilitate future studies entailing complex tissue-specific manipulations. Using this platform, we demonstrate that three different transgenes can be expressed in various neighbouring tissues in a highly specific and robust manner. Furthermore, we show that TALE–VAS lines can be combined with orthogonal binary expression systems (e.g., GAL4–UAS) without any apparent crosstalk between the platforms, paving the way for future cell–cell interaction studies.

A notable feature of the TALE–VAS system is its apparent generic compatibility with fly development and adult survival. In contrast to the first generation of QF lines, which were lethal when expressed in a broad pattern (e.g., pan-neuronal under an elav enhancer), all crosses ranging from a single TALE–VAS pair to triple TALE–VAS combinations generated viable offspring. TALE–VAS flies could also be reared as stable homozygous stocks (e.g., elav$^{1.8kb}$–TALE$_3$; VAS$_1$–Citrine-HA, VAS$_5$–FLAG–Cerulean). Furthermore, pan-neuronal expression using elav$^{1.8kb}$–TALE$_1$ or pan-muscular expression with mhc$^{2.4kb}$–TALE$_4$ resulted in healthy offspring. Although we have not directly quantified lifespan, these animals did not display any obvious reduction in viability.

Since the TALE–VAS pairs reported in this study displayed low level of background activity in certain tissues, they may not be directly suitable for the expression of lethal transgenes. However, the inherent unrestricted multiplexing capacity of the TALE–VAS system provides a unique opportunity to generate background-free driver–responder pairs, thus extending the range of future implementations to any transgene of interest. Furthermore, the effector domains fused to the TALE array could also be easily replaced in subsequent iterations of the technology to enable an even broader spectrum of applications. For example, TALEs could be fused to GAL4 or QF activation domains, which can be silenced by GAL80 or QS, respectively. This would render the
TAL-E-VAS system suitable for clonal studies employing the mosaic analysis with a repressible cell marker strategy. 34 Finally, we envision that in future implementations of this technology, repositories of TALE drivers could be generated based on available GAL4 driver libraries with well characterised expression patterns in development or adult tissues (i.e., the nervous system). 25,26 In principle, this could be readily achievable by directly replacing the GAL4 sequence in existing drivers with specific TALE sequences using the InSITe system, Trojan-MIMICS, or the CRISPR/Cas9 toolbox 24,26. Considering the successful use of TALE transcription factors in vertebrate systems, it is plausible that the TALE-VAS system will also be applicable to other model organisms, where the use of GAL4–UAS has been limited so far. 27,28

Methods

TAL-E assembly. All TALE sequences were designed de novo to ensure high specificity and optimal TAL dimers. Random sequences were generated using the online Fast sequence toolbox Fador 1.41 (http://users.birc.an.dk/biopv/php/fabox/random_sequence_generator.php) 29 taking into account the rules for optimal TAL design 30; following an invariable T at position 0 that is not bound by the TALE array, avoid T at position 1, avoid A at position 2, and the last 3' nucleotide should be T. Since A at position 1 is frequent in naturally occurring TALEs, all TALE sequences started with a 5' A and ended with a 3' T nucleotide (Supplementary Data 2). To avoid the possibility of genomic on-target binding, each candidate sequence was then queried against the Drosophila genome using the National Center for Biotechnology Information's BLAST similarity search algorithm. 31 Corresponding 19 monomer long TALE arrays (Supplementary Data 2) were assembled using the Golden Gate TALEN and TAL Effect Kit 2.0 (gift from Daniel Voytas and Adam Bogdanove, Addgene kit #1000000024) as previously reported. 32 Briefly, for each position in the TALE array a module plasmid encoding the corresponding RVD was selected from the kit library (Supplementary Data 2). All RVDs corresponding to positions 1–10 were assembled in the array plasmid pFUS_A, while those corresponding to positions 11–18 were assembled in pFUS_B. The two array plasmids and the last RVD (position 19) were subsequently combined in the final backbone plasmid.

Generation of TALE driver constructs. Assembled TALEs were subcloned into either pJCF-VP64-SV4 (gift from David Stern, Addgene plasmid #46147) 33 or pJCF-VP64-SV4-P10 (Supplementary Fig. 1). pJCF-VP64-SV4-P10 was generated from pJCF-VP64-SV4 as follows: The TALE-VP64-SV4 sequence was removed from the pJCF-VP64-SV4 plasmid by XhoI and EcoRI digest. The P10 terminator sequence was then amplified from plasmid pJCRF81 using primers P10_InFusion_Fwd and P10_InFusion_Rev (Supplementary Data 3), and the TALE-VP64 construct was amplified from plasmid pJCF-VP64-SV4 using TAIL programmer TAILv2.1, InFusion HD Fusion cloning and TALE. Finally, the three components were re-assembled into the new pJCF-VP64-P10 vector using the In-Fusion HD Cloning Kit (Clontech, 639648). To generate plasmid pJCF-2xUAS-GAL4-SV4, the GAL4 sequence was amplified from plasmid pC12-GAL4 (gift from A. Bassett) with primers IF-pC12-GAL4-Fwd and IF-pC12-GAL4-Rev. The amplicon was then digested with Hoxl and Xbal and inserted into pJCF-VP64-SV4 replacing the TALE-VP64 sequence. pJCF-2xUAS-GAL4-SV4 was obtained through excision of the EGFP sequence from plasmid pJCR81 with XhoI and Xbal, and insertion of the resulting fragment into pJCF-VP64-SV4 replacing the TALE-VP64 sequence. pJCF-2xUAS-GAL4-SV4 was generated by removing 10xUAS from the plasmid by restriction enzyme cloning (HindIII, Nhel) and insertion annealed oligonucleotides HindIII-260bp-NheI-Rev and HindIII-280bp-NheI-Rev.

Drosophila enhancers. To direct the activity of the TALE–VAS system to specific tissues in flies, four different enhancer elements were cloned upstream of the hsp70 minimal promoter in the TALE1, TALE2, and TALE3 vectors (Supplementary Fig. 8): (1) The ~2.4 kb myosin heavy chain (mhc) enhancer element, representing the region between position +1604 bp and the corresponding RVD was selected from the kit library (Supplementary Data 2). To avoid the possibility of genomic on-target binding, each candidate sequence was then queried against the Drosophila genome using the National Center for Biotechnology Information's BLAST similarity search algorithm. 31 Corresponding 19 monomer long TALE arrays (Supplementary Data 2) were assembled using the Golden Gate TALEN and TAL Effect Kit 2.0 (gift from Daniel Voytas and Adam Bogdanove, Addgene kit #1000000024) as previously reported. 32 Briefly, for each position in the TALE array a module plasmid encoding the corresponding RVD was selected from the kit library (Supplementary Data 2). All RVDs corresponding to positions 1–10 were assembled in the array plasmid pFUS_A, while those corresponding to positions 11–18 were assembled in pFUS_B. The two array plasmids and the last RVD (position 19) were subsequently combined in the final backbone plasmid.

Drosophila cell culture experiments. The Drosophila S2 cell line containing at least one stably integrated pMT-GAL4 construct was a gift from A. Bassett. Cells were cultured in Drosophila Schneider’s media (Gibco, 21730-024) supplemented with 10% fetal calf serum (Gibco, 10276-029) and 0.5% (v/v) penicillin streptomycin (Gibco, 1540-122) at 26 °C and passaged every 3–4 days.

For all transfection experiments, cells were seeded at a density of 1 x 10^6 cells per ml in 6, 12, or 24 well cell culture plates in full media 24 h prior to transfection on the day of transfection. Cells were transfected using Effectene (Qiagen, 301425) at a ratio of 1:3 DNA to Effectene (1 μg DNA:10 μl Effectene, 10 min incubation) in full media; FuGENE (Promega, E2211) at a ratio of 1:3 DNA to FuGENE (1 μg DNA:3 μl FuGENE, 40 min incubation) in serum-free media. GAL4 expression was induced 24 h after transfection by addition of 100 μM CuSO4 stock solution (150 μM) to a final concentration of 0.1–0.7 mM CuSO4 per well. EGFP expression was assessed 12–48 h after induction either by flow cytometry or confocal microscopy.

For flow cytometry measurements, 2 μM GAL4 cells were harvested by gently pipetting a stream of media over the well surface, and the resulting suspension was transferred to a 3 ml tube and centrifuged at 1000 g for 3 min. Media was removed and the cells washed in 1× PBS (phosphate buffered saline) and kept on ice. EGFP levels were measured either with a CyAn ADP flow
cytometer (DakoCytomation) or a BD LSR Fortessa cell analyser (BD Biosciences).

Data analysis was performed with the Kaluza Analysis 1.3 software (Beckman Coulter), using the gating strategy described in Supplementary Fig. 14.

For confocal imaging, cells were either seeded for transfection in a 12-well culture plate on sterile, round cover slips or transfected after transfection and allowed to settle on the cover slips. Cells were then fixed with 4% methanol and 0.5% PFA (Electron Microscopy Sciences, 15710) for 10 min at room temperature (RT) and washed twice with 1× PBS for 5 min each. Following a single wash with 1× PBS-Triton (0.1%) for 10 min, cells were blocked in 5% normal goat serum (NGS) in 1× PBS-Triton (0.1%) for 1 h at RT. Primary antibodies were applied either at RT for 2–4 h or at 4 °C overnight. After three washes with 1× PBS-Triton (0.1%) for 10 min each, the secondary antibodies together with DAPI were applied for 2–4 h at RT or 4 °C overnight. Cells were then washed three times with 1× PBS-Triton (0.1%) for 10 min each and subsequently mounted on a microscope slide by gently lowering the cover slip with the cells facing down onto a drop of SlowFade Diamond Antifade Mountant (Life Technologies, S3972). The mouse anti-β-tubulin antibody (05-119, 1:200) was used at 1:200 anti-rabbit, and Alexa488 at 1:1000 mouse-α-A488 secondary antibody (Invitrogen, A11001) was used at 1:1000 dilution. Nuclear staining was performed with DAPI (Invitrogen, D1306) at 1:2000 dilution.

**Generation of transgenic animals.** Transgenic fly lines were established by phiC31 integrase-mediated germline transformation (GenetiVision, USA). Five different landing sites were chosen to allow for the generation of triple drivers and triple responders by classic recombination: attP1 and attP225, V00020, V00031, and V00033. One third of locations of attP1 (PcaryPattP2) and attP2 (PcaryPattP2) were initially identified as chromosome 2R (cytological band 55C-D, between GM04742 and jockey) and 3L (cytological band 68A1-B2, between genes CG6310 and Mocsi), respectively25,26. These positions were later refined as follows: attP1 is located in plus orientation in band 56C1, 62 bp upstreem of the gene jdb [2R:18,357,321-18,357,321, 55C4]; attP2 is in plus orientation in band 68A4, 44 bp upstream of the gene Mocsi [3L:11,076,538-11,076,538]25,26. The three VK landing sites are in plus orientation for the following genomic loci: VK00020 (Pbac[y+]-attP-9A VK00020) on chromosome 3R (band 99F9; [3R:30,553,313-30,553,313] in the intron of the gene tmd); VK00031 (Pbac[y+]-attP-3B VK00031) on chromosome 3L (band 62E1 [3L:2,395,487-2,395,487] in the intron of the gene CG54186); and VK00037 (Pbac[y+]-attP-3B VK00037) on chromosome 2L (band 22A3 [2L:1,582,820-1,582,820] in the intron of the gene huF) (Supplementary Fig. 9)25,26. All chromosomal locations in square brackets are based on FlyBase (FB2017_02, released 18 April 2017). For presenting an independent integration event.

In addition to the TALE–VAS lines, the following fly stocks were used in this study: UAS-mCD8-mCherry (BL27391), UAS-mCD8-GFP/Cyo. Orgeon R or w1118 were used as wild-type. The ase-GAL4 was a gift from Jf-Long Liu. However, the exact origin of this line could not be retraced. The most commonly used ase-GAL4 lines8,9 are based on the ~2 kb upstream enhancer element described in Jarman et al.40 or were generated by cloning ~2 kb fragments of the ase gene (Janelia lines)25,26. All these lines contain substantially larger genomic fragments compared to the 0.8 kb sequence used for the ase-TALEs. Flies were reared on standard cornmeal media at 18 or 25 °C. All genotypes were confirmed by PCR unless indicated otherwise. Compound stocks, triple driver, and triple responder lines were generated by classic recombination and standard genetic techniques (Supplementary Data 4).

**Larval dissections and immunohistochemistry.** Larval dissections were performed with DAPI (Invitrogen, D1306) for 10 min, cells were blocked in 5% normal goat serum (NGS) in 1× PBS-Triton (0.1%) for 10 min at room temperature (RT) and washed twice with 1× PBS for 5 min each. Following a single wash with 1× PBS-Triton (0.1%) for 10 min, cells were blocked in 5% normal goat serum (NGS) in 1× PBS-Triton (0.1%) for 1 h at RT. Primary antibodies were applied either at RT for 2–4 h or at 4 °C overnight. After three washes with 1× PBS-Triton (0.1%) for 10 min each, the secondary antibodies together with DAPI were applied for 2–4 h at RT or 4 °C overnight. Cells were then washed three times with 1× PBS-Triton (0.1%) for 10 min each and subsequently mounted on a microscope slide by gently lowering the cover slip with the cells facing down onto a drop of SlowFade Diamond Antifade Mountant (Life Technologies, S3972). The mouse anti-β-tubulin antibody (05-119, 1:200) was used at 1:200 anti-rabbit, and Alexa488 at 1:1000 mouse-α-A488 secondary antibody (Invitrogen, A11001) was used at 1:1000 dilution. Nuclear staining was performed with DAPI (Invitrogen, D1306) at 1:2000 dilution.

**Confocal image processing.** 3D confocal stacks were converted into 2D images using the maximum intensity projection tool provided by the Zen software (Zeiss). Images were exported from the Zen software (Zeiss) and exported as TIFs in RGB mode colour and processed in Adobe Photoshop CS5 (Adobe). For publication purposes, brightness and contrast was adjusted, where necessary, using the levels function. In case images were acquired using the same microscope settings, processing was performed using exactly the same parameters. Croping was carried out to render the rectangular marginal area to without modifying the data resolution. The cyan channel was pseudo-coloured by adding the information from the blue channel to the green channel. All figures were assembled in Adobe Illustrator CS6 (Adobe). Bar graphs and statistical analyses were performed with GraphPad Prism 7 (GraphPad).

**Temperature-dependent TALE–VAS experiments.** Multiple crosses between 10 males of the elav39,80-TALE, driver and 10 virgin females of the triple responder (VAS-V, mCherry, VAS-Citrine-HA, VAS-FLAG-Cerulean) were set up at 18 and 25 °C. Wandering third larval instars were picked from vials and dissected imaginal discs and brains were analysed either by confocal microscopy (immunohistochemistry) or RT-qPCR. Maximum intensity projections of brains or eye discs were generated with the Zen software (Zeiss) and exported as TIFs. For the analysis of Citrine-HA transgene expression in eye discs, an area within the retina was cropped using Adobe Photoshop CS5 (Adobe). The mean pixel intensity of each channel within this area was then calculated using the measure tool in the ImageJ package (http://ij.net.sc/Fiji). To account for differences in the staining efficiency, signal intensity (anti-HA staining) was normalised to Elav levels.

**TALE off-target analysis.** To identify possible genomic off-targets, a window spanning 2 kb upstream and 2 kb downstream from each Drosophila TSS annotated in the UCSC transcriptome database was analysed. The R package “TxtDb.Dmelanogaster.UCSC.dm6.encGene” was employed to map TSS using the “BGenome. Dmelanogaster.UCSC.dm6.encDn T. Drosophila genome release. Possible off-target sites were identified by scanning both strands within each 4 kb window for sequences similar to VAS-4 using the R function “matchPattern” allowing up to 4 nt mismatches. To assess the effect of potential off-target binding on endogenous gene regulation, homologous genes were reared at 25 °C, third larval instars were picked from the wall of the culture vessels, and processed for RT-qPCR. The wild-type (w1118) and the triple VAS-responder line (VAS3-Citrine-HA, VAS5-V, mCherry, VAS5-FLAG-Cerulean) served as non-TALE expressing controls, while the triple TALE driver line (elav39,80-TALE, ase80-TALE, repo80-TALE) and the triple TALE-driver > triple VAS-responder line were used to analyse on-target and off-target activation.

**RT-qPCR analysis.** For the analysis of temperature impact on TALE activity, third larval instar brains were dissected one at a time as described above in DEPC (Sigma, D5758) treated 1× PBS and immediately transferred to ice cold RT lys buffer (RNeasy Mini Kit, QiaGen, 74106) containing 1% (v/v) β-Mercaptoethanol (Sigma, M6250). Samples were then vortexed for 20 s at 2000 rpm to facilitate the lys reaction, and stored at −80 °C until RNA extraction. Three biological replicates each containing ten brains were processed for every temperature data point. For the off-target analysis, total third larval instars were homogenised in ice cold RT lys buffer containing 1% (v/v) β-Mercaptoethanol with a motor pestle. After vortexing for 20 s at 2000 rpm, samples were immediately transferred to the −80 °C for RNA extraction. Three biological replicates each containing ten brains were processed for every temperature data point. For the off-target analysis, total third larval instars were homogenised in ice cold RT lys buffer containing 1% (v/v) β-Mercaptoethanol with a motor pestle. After vortexing for 20 s at 2000 rpm, samples were immediately
McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K. & Davis, R. L. Spatiotemporal rescue of memory dysfunction in Drosophila. Science 302, 1765–1768 (2003).

Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker (MARC) for Drosophila neural development. Trends Neurosci. 24, 251–254 (2001).

Jenett, A. et al. A GAL4-driver line resource for Drosophila neurobiology. Cell Rep. 2, 991–1001 (2012).

Kvon, E. Z. et al. Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature 512, 91–95 (2014).

Gohl, D. M. et al. A versatile in vivo system for directed dissection of gene expression patterns. Nat. Methods 8, 231–237 (2011).

Diao, F. et al. Plug-and-play genetic access to drosophila cell types using exchangeable excone cassettes. Cell Rep. 10, 1410–1421 (2015).

Housden, B. E. & Perrimon, N. Cas9-mediated genome engineering in Drosophila melanogaster. Cold Spring Harb. Protoc. 2016, pdbtop086843 (2016).

Zhang, F. et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat. Biotechnol. 29, 149–153 (2011).

Perez-Pinera, P. et al. Synergistic and tunable human gene activation by combinations of synthetic transcription factors. Nat. Methods 10, 239–242 (2013).

Villeisen, P. Falbox: an online toolbox for fasta sequences. Mol. Ecol. Resour. 6, 965–968 (2007).

Cermak, T. et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. Nucleic Acids Res. 39, e82 (2011).

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. J. Mol. Biol. 215, 403–410 (1990).

Klein, P. et al. Ret rescues mitochondrial morphology and muscle degeneration of Drosophila Pink1 mutants. EMBO J. 33, 341–355 (2014).

Yao, K. M. & White, K. Neural specificity of cay expression: defining a Drosophila promoter for directing expression to the nervous system. J. Neurochem. 63, 41–51 (1994).

Lee, B. P. & Jones, B. W. Transcriptional regulation of the Drosophila glial gene repo. Mech. Dev. 122, 849–862 (2005).

Groth, A. C., Fish, M., Nusse, R. & Calos, M. P. Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics 167, 1775–1782 (2004).

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

Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat. Genet. 40, 476–483 (2008).

Zur Lage, P. & Jarman, A. P. Antagonism of EGFR and notch signalling in the reiterative recruitment of Drosophila adult chordosensory sense organ precursors. Development 126, 3149–3157 (1999).

Zhu, S. et al. Gradients of the Drosophila Chrimo BTB-zinc finger protein govern neuronal temporal identity. Cell 127, 409–422 (2006).

Jarman, A. P., Brand, M., Jan, L. Y. & Jan, Y. N. The regulation and function of the helix-loop-helix gene, asense, in Drosophila neural precursors. Development 119, 19–29 (1993).

Pfeifer, B. D. et al. Tools for neuroanatomy and neurogenetics in Drosophila. Proc. Natl. Acad. Sci. USA 105, 9715–9720 (2008).

Acknowledgements

We are grateful to Andrew Bassett for providing the S2 pMT-GAL4 cell line, Tatjana Sauka-Spengler for the Citrine, mCherry and Cerulean constructs, Frank Schnorrer for the mhc enhancer plasmid, and Ilan Davis, Ill-long Liu, and the Bloomington Drosophila Stock Center (NIH P40OD018537) for providing fly stocks. We would like to thank Kevin Clark and Paul Sopp (WIMM FACS Core Facility) for assistance with the mhc enhancer plasmid, and Ilan Davis, Ji-Long Liu, and the Bloomington Drosophila Stock Center (NIH P40OD018537) for providing fly stocks. We would like to thank Kevin Clark and Paul Sopp (WIMM FACS Core Facility) for assistance with flow cytometry measurements and technical expertise, and Christofler Lagerhom (Wolfson Imaging Centre Oxford) for help with confocal imaging. We are grateful to Ilan Davis, Aron Szabo, Quentin Ferry, Yale Michaels, and Toni Baemmler for providing critical comments on the manuscript. M.T. was supported by MRC (#G0902418) to T.A.F. and D.J.H.F.K. is supported by a CIHR Postdoctoral Fellowship (#DT405100). G.A. was supported by a fellowship from the Malaysian Ministry of Education. E.L. was supported by a fellowship from Princeton University and MRC (#G0902418) to T.A.F. D.J.H.F.K. is supported by a CHIR Postdoctoral Fellowship (#201511MFR6833-204362) and BBBSC (#BB/N06550/1). T.A.F. is supported by MRC (#G0902418) and BBBSC (#BB/L010275/1).
Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-017-01592-3.

Competing interests: Y.T. and M.F. are employees of GenetiVision Corporation, a company that provides Drosophila transgenic services. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.