Development 140, 2434-2442 (2013) doi:10.1242/dev.088757

A versatile platform for creating a comprehensive UAS-ORFeome library in *Drosophila*

Johannes Bischof[1,*], Mikael Björklund[2,3,*], Edy Furger[1], Claus Schertel[1], Jussi Taipale[3,4] and Konrad Basler[1,‡]

SUMMARY

Overexpression screens are used to explore gene functions in *Drosophila*, but this strategy suffers from the lack of comprehensive and systematic fly strain collections and efficient methods for generating such collections. Here, we present a strategy that could be used efficiently to generate large numbers of transgenic *Drosophila* strains, and a collection of 1149 UAS-ORF fly lines that were created with the site-specific ΦC31 integrase method. For this collection, we used a set of 655 genes that were cloned as two variants, either as an open reading frame (ORF) with a native stop codon or with a C-terminal 3xHA tag. To streamline the procedure for transgenic fly generation, we demonstrate the utility of injecting pools of plasmids into embryos, each plasmid containing a randomised sequence (barcode) that serves as a unique identifier for plasmids and, subsequently, fly strains. We also developed a swapping technique that facilitates the rapid exchange of promoters and epitope tags *in vivo*, expanding the versatility of the ORF collection. The work described here serves as the basis of a systematic library of Gal4/UAS-regulated transgenes.

KEY WORDS: *Drosophila*, ΦC31, ORFeome library, Overexpression

INTRODUCTION

Traditionally, genetic analysis in *Drosophila* has relied on loss-of-function approaches to determine the specific function of genes. Such mutagenesis screens have contributed substantially to our understanding of gene functions and biological processes. Over the years, this strategy was extended and recently culminated in the creation of a genome-wide RNAi library for targeted knockdown of genes *in vivo* (Dietzl et al., 2007) (other similar resources are 'TRiP' and 'NIG-Fly', see www.flyrnai.org and www.shigen.nig.ac.jp/fly/nigfly, respectively). Despite their merits, loss-of-function studies have failed to reveal the function of the majority of all *Drosophila* genes, as they show no obvious loss-of-function phenotype (Miklos and Rubin, 1996). Some of this is likely to be due to functional redundancy, which could be addressed by, for example, combinatorial RNAi strategies. A likely reason for the frequent 'absence' of loss-of-function phenotypes is the use of assays that are simply not sensitive enough to detect subtle defects (e.g. Ramani et al., 2012). Thus, the function of the majority of all *Drosophila* genes cannot be identified by conventional loss-of-function or knockdown screens, and alternative approaches, such as the use of sensitised backgrounds, as well as an increase in assay sensitivity are required to mitigate this limitation.

By contrast, extensive mis- and overexpression screens in *Drosophila* are less common, though they can be equally informative for elucidating gene functions (e.g. Wittwer et al., 2005; Prelich, 2012). So far, such screens have almost exclusively relied on random transposon insertions containing an enhancer-promoter (EP) element that can be used to drive the expression of genes flanking the insertions (Rørth, 1996; Bellen et al., 2011). In another case, a collection of transgenic lines generated from a specific set of heterologous open reading frame (ORF) constructs was used for misexpression (Xu et al., 2008; for a review, see Zhong and Yedvobnick, 2009). As with the RNAi library, the binary Gal4/upstream activating site (UAS) system (Brand and Perrimon, 1993) is used in these gain-of-function screens to achieve spatiotemporal control of gene expression; naturally, these screens can also be performed in sensitised backgrounds. Screens utilising such random insertion lines, however, are not saturating, an inherent limitation of using transposons. Furthermore, unambiguous identification of the phenotype-causing 'event', which does not necessarily have to be the activation of a neighbouring gene, often causes considerable complications and efforts.

Thus, a more systematic and controlled method of misexpression is required to overcome these shortcomings. In particular, the new strategy should potentially allow saturation and any resulting phenotype should be precisely linked to a specific gene. We propose the generation of a transgenic ORFeome library originating from specifically designed overexpression constructs. Up to now, however, the *in vivo* use of even small sets of UAS-ORF constructs in *Drosophila* has been very limited (Xu et al., 2008). The considerable efforts required for cloning and sequencing full-length expression constructs and for establishing and maintaining the large number of resulting transgenic lines has hampered implementation of such a strategy. The recent development of the site-specific ΦC31 integrase system in *Drosophila* (Groth et al., 2004; Bateman et al., 2006; Venken et al., 2006; Bischof et al., 2007) overcomes some of these challenges. In particular, it is now feasible to inject large collections of constructs, as this method guarantees efficient integration. Targeted insertion eliminates the need for mapping, and selection of a suitable *attP* (phage attachment) site minimises unpredictable position effects. Consequently, the effects of transgenes can be compared more reliably. A further advantage of the ΦC31 integrase system is that lost transgenic lines can be easily and identically reconstructed.

Here, we present a strategy for establishing a genome-wide UAS-ORF *in vivo* library, exemplified by the generation of a pilot library

[1]Institute of Molecular Life Sciences, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland. [2]College of Life Sciences, Division of Cell and Developmental Biology, University of Dundee, DD1 5EH, Scotland, UK. [3]Genome-Scale Biology Research Program, University of Helsinki, FI-00014, Finland. [4]Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, SE-171 77, Sweden.

*These authors contributed equally to this work
‡Author for correspondence (konrad.basler@imls.uzh.ch)

Accepted 26 February 2013

containing 1149 fly lines. Site-specific integration and the use of molecular barcodes facilitate the efficient generation and identification of large numbers of transgenic flies. Furthermore, we investigated the effects of epitope tags on transgene functionality. We present a system that enables modification of the ORFeome library *in vivo*, including the exchange of promoter regions and epitope tags simply by crossing appropriate fly strains. In summary, this library provides a powerful tool for *Drosophila* genetics and offers a rapid screening method for identifying novel regulators or effectors of genetic pathways.

MATERIALS AND METHODS

Plasmid constructions

pUASg.attB and pUASg-HA.attB

First, pUAST.attB (Bischof et al., 2007) was digested with *Nhe*I-*Hin*dIII, releasing the loxP site fragment, which was replaced by a duplex oligonucleotide containing the *Swa*I site (*Nhe*I and *Hin*dIII sites were destroyed). Next, the UAS-hsp70 promoter fragment was amplified from pUAST.attB and cloned into the plasmid using *Swa*I and *Kpn*I. The SV40 trailer (cleaved with *Kpn*I-*Bam*HI) was replaced with a duplex oligonucleotide containing the following restriction sites: *Kpn*I, *Nhe*I, *Xho*I, *Hin*dIII and a *Bam*HI compatible overhang (destroying the *Bam*HI site after insertion). Next, we digested the plasmid with *Nhe*I-*Xho*I and inserted a *tubulin* 3′UTR, PCR-amplified from pM{3xP3-RFPattP} (Bischof et al., 2007). We digested this clone with *Kpn*I, followed by blunting, and then with *Nhe*I to insert a Gateway cassette, either without tag (pUASg.attB) or with the 3xHA tag (pUASg-HA.attB). The cassette for pUASg.attB was released from the Gateway plasmid pAW by *Sac*I, followed by blunting, and then *Nhe*I digestion. The cassette for the pUASg-HA.attB plasmid was released from the Gateway plasmid pAWH by *Eco*RV and *Nhe*I digest. Details of plasmids pAW and pAWH can be found at the *Drosophila* Genomics Resource Center (DGRC) vector collection (https://dgrc.cgb.indiana.edu/vectors/store/vectors.html).

pTF-HA.attB

We digested pUASg.attB with *Bgl*II-*Age*I and inserted a duplex oligonucleotide containing the shortened and mutated *FRT5* variant (CGAAGTTCCTATTCTTCAAAAGGTATAGGAACTTCA) (Schlake and Bode, 1994).

Then we digested with *Age*I-*Not*I and inserted the 5′ part of the Gateway cassette, amplified from pUASg.attB. Next, the plasmid was digested with *Xba*I-*Nhe*I, and both the 3′ part of the Gateway cassette (amplified from pUASg.attB, pre-digested with *Xba*I-*Kpn*I) and the 3xHA stretch (amplified from pUASg-HA.attB, pre-digested with *Kpn*I-*Nhe*I) were simultaneously inserted.

Next, we cleaved with *Swa*I-*Bgl*II and inserted a loxP-UAS-hsp70 fragment amplified from pUASg.attB (loxP sequence was inserted by the forward primer). Finally, we cleaved the plasmid with *Kpn*I and inserted an *FRT2* oligonucleotide (CGAAGTTCCTATTCTCTACTTAGTATAGGAACTTC) (Schlake and Bode, 1994).

pGW-HA.attB

This is our current destination vector used for the further cloning of the ORFeome library. This plasmid differs from pTF-HA.attB in that the sequence between the *FRT2* and the 3xHA tag was replaced with a 24-aa-long linker region (FL sequence: ILGAPSGGGATAGAGGAGGPAGLI) that is thought to minimise steric interference between the native protein and the epitope tag (Jankovics and Brunner, 2006). We cleaved pTF-HA.attB with *Kpn*I-*Nhe*I and inserted the synthesised fragment *FRT2*-FL-3xHA via these sites, giving rise to pGW-HA.attB.

pattB

We digested pUAST.attB (Bischof et al., 2007) with *Bam*HI and removed the UAS-hsp70-SV40 cassette. This cassette was replaced by a duplex oligonucleotide containing a multiple cloning site.

pPSlexO.attB

The MCS (multiple cloning site) of the pattB plasmid was deleted by *Nhe*I-*Xba*I and replaced with the linker *Nhe*I-*Mlu*I-*Avr*II-*Xba*I (*Nhe*I site being destroyed), followed by *Dra*II digest and insertion of a oligonucleotide containing the sites *Dra*II-*Xho*I-*Nhe*I-*Not*I-*Kpn*I-*Bgl*II. The *Bgl*II site was used to insert the mutant *FRT5* variant. The plasmid was opened with *Nhe*I-*Kpn*I, made blunt, and used to insert a blunted lexA operator (lexO) *Bam*HI-*Eco*RI fragment from pLOT (Lai and Lee, 2006). Finally, a *yellow* marker gene released with *Sal*I from a flp-out cassette (Basler and Struhl, 1994) was inserted into the compatible *Xho*I site, giving rise to pPSlexO.attB.

pTSeGFP.attB

We deleted the *white* gene and the loxP site from pattB by *Dra*II-*Nhe*I digestion, blunted, re-ligated, digested with *Bam*HI-*Xho*I and used the plasmid in a four-fragment ligation together with the following fragments: *Bam*HI-FRT2-*Hin*dIII, *Hin*dIII-eGFP-*Not*I and *Not*I-tubulin 3′UTR-*Xho*I. The resulting plasmid from this ligation was then digested with *Xho*I to insert a *yellow* gene (identical to pPSlexO.attB), generating pTSeGFP.attB.

pTSVNm9.attB and pTSVC155.attB

The VNm9 and VC155 fragments were PCR-amplified from plasmids VNm9 and VC155 in pCS2 (Saka et al., 2007), introducing the flanking restriction sites *Hin*dIII and *Not*I. Additionally, a myc tag was introduced at the 5′ end of VNm9. Plasmid pTSeGFP.attB (without the *yellow* insertion) was digested with *Hin*dIII-*Not*I, followed by insertion of VNm9 and VC155, respectively. Finally, the *yellow* marker was inserted as already described.

p3xP3-eGFP/vas-dΦC31.attB

This plasmid is identical to p3xP3-eGFP/vas-ΦC31(+/−NLS)attB in Bischof et al. (Bischof et al., 2007), except for dΦC31 replacing *Streptomyces* phage ΦC31. dΦC31 is a *Drosophila* codon-optimised ΦC31 integrase, differing in 172 nucleotides from the phage integrase ORF.

Reporter plasmids

Detailed information on the construction of the reporter plasmids placZ.attB, placZ-2.attB, pEGFP.attB and pEGFP-2.attB and on the *FRT* test constructs is available upon request.

Barcoding

Molecular barcoding using randomised duplex oligonucleotides was used to facilitate identifications of plasmids and fly lines. We first isolated a *Xho*I-*Hin*dIII flanked *lacZ* gene as a stuffer fragment and inserted it into the *Xho*I-*Hin*dIII site of the various vectors to be barcoded. Plasmids were then digested and the stuffer replaced with the barcode oligonucleotides. For the pilot library we used the following barcode design: 5′-*Xho*I-ANNTANNNATNNNNTAANNNNTANNNATNNNTANNG-*Hin*dIII-3′, i.e. 36 nucleotides of which 21 are randomised. The 5′-phosphorylated barcode oligonucleotide mixtures were annealed according to standard protocols and the restriction sites were created as ‘sticky-end’ overhangs. The barcodes were ligated into the plasmids in a 250 μl ligation mix overnight at 16°C. This we next used to transform 50 tubes of 50 μl aliquots of DB3.1 cells with 2.5 μl ligation mix each; these reactions were heat-shocked for 45 seconds, then combined and added to 100 ml SOC medium. After 1 hour shaking at 37°C, 100 μl aliquots were plated on LB+Amp plates to calculate the diversity of the barcodes. We scored ~450 colonies, which results into an actual diversity of 450,000 (dilution 1:1000). 2xTY medium (300 ml) containing 75 μg/ml ampicillin and 10 μg/ml chloramphenicol was added to the remaining culture, which was grown for 19 hours at 37°C. Plasmids were purified using Qiagen Maxi Prep Kit.

The barcode used for the latest version of our destination vectors, pGW-HA.attB, has the following composition: 5′-*Xho*I-ANNTGNNNA-CNNNNTGANNNNACNNNATNNNGANNG-*Hin*dIII-3′. The main difference from the previous version is an increase in the GC content.

Cloning and verification of ORFs

Full-length ORFs were cloned from the *Drosophila* gene collections releases 1-3 or from the Schneider cell SD-pOT2 cDNA library (Berkeley *Drosophila* Genome Project) with gene-specific primers using a two-step strategy in which full-length *attB1* and *attB2* sites are introduced into the PCR product as described in the Gateway Cloning manual (Invitrogen). The forward primer had a sequence AAAAAGCAGGCTTCAAC before methionine codon and gene-specific sequence [sequence corresponding to

Drosophila Kozak consensus CAA(A/C) underlined] (Cavener, 1987). Two versions of the reverse oligonucleotide were used, one with native stop codon and one without stop, with AGAAAGCTGGGTC flanking sequence. The inserts were cloned to pDNR221 vector by BP recombination reaction and plated on kanamycin-containing LB agar plates. Single clones were picked into 96-well deep well plates containing 1.3 ml Terrific Broth medium and cultured for 22-24 hours at 37°C. Miniprep DNA was purified using the Promega Wizard SV96 Kit. Correct recombinants were identified by end-sequencing with M13F and M13R primers.

For the entry clones for which full-length ORF sequence could not be obtained by end sequencing, we amplified the inserts by PCR using *attB1* and *attB2* primers. The PCR products were pooled, nebulised and sequenced using a Roche 454 Sequencer at the Institute of Biotechnology, University of Helsinki.

The entry clones were transferred to barcoded pUASg.attB (for stop-containing clones) and pUASg-HA.attB (for ORFs without stop codon) destination vectors by a Gateway LR recombination reaction and selected on ampicillin.

Preparation of ORF pools

Single ORF clones were arrayed in 96-well plates and the plasmid concentration normalised to 100 ng/μl. To maximise the recovery of the transgenes in the pool injections, the ORF clones were arrayed according to their size over several multiwell plates, i.e. the clones within one multiwell plate span the smallest possible size range. Before pooling, the barcodes were sequenced from the expression clones to associate the barcode information with the individual ORF.

The ORF clones were pooled with multi-channel pipettes and the mixes were cleaned using Midiprep Qiagen-tips. For injections, we diluted the pools to a concentration of 100 ng/μl.

Generation of *ΦX-86Fb* and other fly strains

We inserted a *Drosophila* codon-optimised ΦC31 integrase construct, p3xP3-eGFP/vas-dΦC31attB, into the *attP* landing site *ZH-attP-2A* (Bischof et al., 2007). The resulting transgenic line, *vas-dΦ-zh2A*, is doubly fluorescence-marked with *RFP* and *eGFP*. Both markers are under the control of the *3xP3* promoter (Horn et al., 2000). The *RFP* marker originates from the *attP* landing site construct pM{3xP3-RFPattP}, the *eGFP* from the integrase construct. To create the line ΦX-86Fb, we combined the vasa-integrase transgene with the *ZH-attP-86Fb* landing site. The genotype of this line is *y*¯*w*¯ *M{eGFP.vas-int.Dm}ZH-2A; +; M{3xP3-RFP.attP}ZH-86Fb*. The line *ΦX-86Fb* is available at the Bloomington *Drosophila* Stock Center (Indiana University), along with other lines created in a similar fashion. We maintain these ΦC31 integrase containing stocks at 18°C.

PSlexO-86Fb, *TSeGFP-86Fb*, *TSVNm9-86Fb*, *TSVC155-86Fb*: The constructs pPSlexO.attB, pTSeGFP.attB, pTSVNm9.attB and pTSVC155.attB were injected into line *ΦX-86Fb*. Lines were made homozygous for these transgenes and combined with an X chromosome-linked hsp70-flp construct. The fly strains used for bimolecular fluorescence complementation (BiFC) analysis were combined with a *GMR-Gal4* on the second chromosome. Other used driver lines were *MS1096-GAL4*, *GMR-GAL4*, *ey-Gal4*, *dpp-GAL4*, *en-GAL4* and *C765-Gal4*.

ΦC31 integrase-mediated germline transformation

We injected the plasmid DNA pools into *ΦX-86Fb*. As a rule of thumb we injected ~100 embryos per ten different constructs, i.e. for a pool of ~100 ORFs we injected ~1000 embryos (ten slides). The embryos were dechorionated, dried, covered with 10S Voltatef oil and injected with either a simple syringe-microinjection device or with an Eppendorf FemtoJet-TransferMan setup combined with a micromanipulator mounted on a Zeiss Axi inverted microscope. Glass capillaries were pulled on a Narishige PN-30 puller and opened on a grinder.

Injected embryos were kept in a moist chamber at 18°C for 2 days, then larvae were transferred to fly food vials and raised at 25°C (~80-120 per vial). Two G0 males were crossed to three to four *y*¯*w*¯ virgins per vial; G0 females were discarded. Single F1 transgenic progeny males were crossed to a third chromosome balancer line (*y*¯*w*¯; *D gl3 / TM3 Sb Ser*) to subsequently create a balanced stock. We often took two F1 transgenic males from the same vial and crossed them out individually, as we often acquired different transgenic lines from these males. Established balanced stocks were not subsequently monitored in order to create homozygous stocks.

Molecular determination of transgenic progeny

Three days after setting up the F1 crosses, we subjected the males to single fly PCR procedures. PCR was carried out with a primer pair aligning to the *tubulin* 3′UTR of the vector and to the flanking genomic region of the 86Fb landing site (tub-F2 new: ATTTATGTGACTATGGTAGGTCG; 86Fb-Rev: GCTTAGCTTCTGGGTGCATGTGACCG). PCR products were sequenced to identify the barcode composition. This information permits association of the individual outcrosses with specific ORFs, from which we could decide which lines to keep and balance, and which to discard (discarded lines represent recurring transgenes).

Swapping procedure

We used a vial setup and tested several heat-shock regimes, guided by conditions used in Parks et al. (Parks et al., 2004). We crossed four males from the ORF library to approximately ten females carrying the swap construct and a hsp70-flp transgene (all flies homozygous for the constructs). At day 3, we subjected the progeny to a single 30-minute heat-shock. Offspring virgin females were collected and mated to *y*¯*w*¯ males: four females and three males per vial (F1 cross). The offspring of these crosses was screened for the appropriate marker combinations (see Results). We noticed that often only a few flies with the desired marker combinations were found per positive outcross. Note that in other experiments we outcrossed males instead of females in the F1 crosses and obtained similarly efficient swapping rates.

Western blotting

Protein extracts were obtained by boiling (5 minutes at 95°C) heads or third instar wing imaginal discs in 30 μl NuPAGE LDS sample buffer (Invitrogen). Protein extracts were centrifuged at 14,100 ***g*** for 5 minutes, separated on NuPAGE 4-12% Bis-Tris Gel (Invitrogen) and transferred onto Amersham Hybond-C Extra membrane (Amersham Biosciences). HA-tagged proteins were detected with mouse anti-HA antibody (HA.11, Covance, 1:5000); equal loading was estimated with mouse anti-α-tubulin antibody (DM1A, Sigma, 1:5000).

Immunohistochemistry

Antibodies used were mouse anti-HA antibody (HA.11, Covance, 1:500) and secondary goat anti-mouse Alexa594 (1:400, Molecular Probes). DAPI (0.4 μg/ml) was used for staining DNA. Fixation, immunohistochemistry and imaging were conducted via standard protocols.

Other methods

RNA isolation, cDNA synthesis and single fly PCR were carried out according to standard protocols.

Accession numbers

GenBank accession numbers are as follows: pUASg.attB, KC896836; pUASg-HA.attB, KC896837; pGW-HA.attB, KC896838; pattB, KC896839; placZ.attB, KC896840; placZ-2.attB, KC896841; pEGFP.attB, KC896842; pEGFP-2.attB, KC896843.

Fly lines

The transgenic UAS-ORF lines will be available from FlyORF (http://www.flyorf.ch).

RESULTS

Vector design, barcoding and *attP* site selection

We first constructed two Gateway-based destination vectors (Fig. 1B): one for the expression of full-length ORFs with their native stop codon (untagged clones, stop clones), and a second for ORFs that are fused to a C-terminal 3xHA (hemagglutinin) tag after the shuttling event (tagged clones, HA or 3xHA clones). Three tandem copies of the HA epitope tag were used to increase the sensitivity and signal-to-noise ratio in biochemical and histochemical assays (Jarvik and Telmer, 1998). These vectors are equipped with a partially

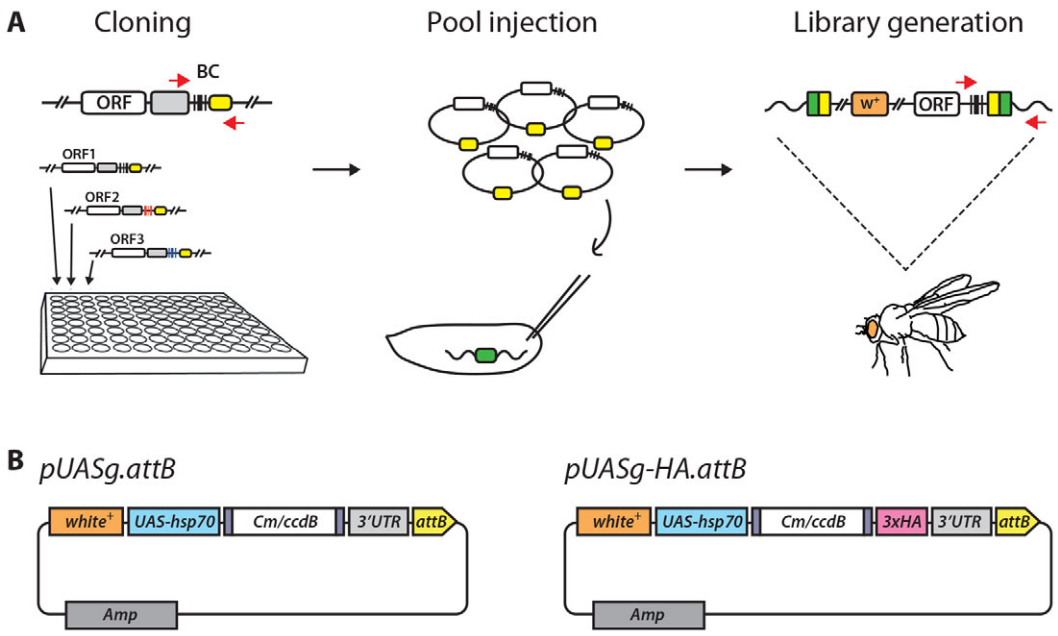

Fig. 1. Library creation strategy and vector design. (**A**) Flowchart of the ORF library generation process. Left: ORFs are cloned into Gateway-based destination/expression vectors that contain randomised 'barcode' sequences. The well position for each ORF clone is known and the associated barcode composition is determined after the ORF cloning is completed by sequencing over the barcode region. Schematic at top: grey box, 3′UTR; yellow box, *attB* recombination site; barcode is indicated by multiple bars (BC); red arrows, primers. Centre: the ORF clones of 96-well plates are pooled and bulk injected into strain *ΦX-86Fb* (green box, *attP* landing site). Right: After outcrossing, the transformants are determined by single fly PCR followed by sequencing the barcode (arrows indicate the primers used for the PCR). The reverse primer binds to the flanking genomic region of the *attP-86Fb* landing site. The green/yellow boxes represent the hybrid sites that are created upon integration of the plasmid into an *attP* site. The wavy lines represent flanking genomic sequences next to the *attP* docking site. Note that there is a *3xP3-RFP* transgene present next to the *attP* site (not indicated). (**B**) Destination vector design. The two constructs used for the pilot library are identical except for the 3xHA epitope tag (the vectors are depicted with the Gateway cassette containing the genes *Cm* and *ccdB*, flanked by indicated Gateway recombination sites). These Gateway-based vectors contain five Gal4-responsive UAS elements, a basal *hsp70* promoter, the *tubulin alpha 1* 3′UTR and an *attB* site for site-specific integration.

randomised oligonucleotide of 36 bp for vector identification. Some positions within the barcode are kept constant to provide a recognisable signature, and the remaining positions are completely randomised. The used barcode has a theoretical diversity of 4^{21} (21 random positions, 15 fixed), allowing any transgene in a genome-wide collection to be unambiguously identified even in the presence of sequencing errors. In practice, the diversity is limited by the transformation efficiency of the barcode library. Our current version of the vector has ~450,000 independent clones (see Materials and methods). In subsequent steps, this molecular barcode serves as a unique identifier for any expression clone and the corresponding fly line. These vectors were used to generate a 'pilot library' of 655 different *Drosophila* genes, and two versions of this gene set were created: an untagged set and a 3xHA-tagged set. These genes were selected based on their loss-of-function phenotypes in *Drosophila* S2 cells (Björklund et al., 2006) or because they belong to pathways that have previously been implicated in growth, cell size regulation and cell cycle progression. All the ORFs were provided with a *Drosophila* Kozak consensus sequence to support efficient translation (see Materials and methods) (Cavener, 1987).

Transgenic constructs were inserted into the *Drosophila* genome at the cytological position 86F, using the previously generated *attP* insertion *ZH-attP-86Fb* (Bischof et al., 2007). This landing site was chosen based on a number of parameters that are crucial for the efficient generation of a large transgenic library and for its subsequent utilisation. We previously evaluated this line, together with others, based on the expression profile in wing discs (Bischof et al., 2007). Additionally, we quantified expression levels in 11 *ZH-attP* lines by measuring β-galactosidase activity upon expression of *UAS-lacZ* reporter transgenes by the *C765-Gal4* driver. In these assays the site *attP-86Fb* provided one of the highest expression levels among the tested lines (supplementary material Fig. S1). In summary, *ZH-attP-86Fb* offers (1) a high integration rate, (2) strong transgene expression, (3) ease of scoring transgenic offspring owing to a moderately strong *white* marker expression, which enables differentiation of hetero- and homozygous transgenics, and (4) good overall fitness and high fertility.

Probably owing to a genomic location (supplementary material Fig. S2A) that favours strong transgene expression, the line also displays some background expression (detected during early embryonic development; not shown). We achieve a high homozygosity rate of ~90% with inserted transgenes at this *attP* site (based on 1107 UAS-ORF lines), further indicating that this site is well suited for library construction. The site *attP-86Fb* was combined with a *Drosophila* codon-optimised ΦC31 integrase (Bischof et al., 2007) located on the X chromosome, giving rise to the line *ΦX-86Fb*, which we used for creating the pilot library.

In addition to the library vectors mentioned above, we constructed multiple transformation vectors suitable for the ΦC31 integration method. These include the vector *pattB* for cloning of, for example, genomic rescue constructs, and the four reporter constructs *placZ.attB*, *placZ-2.attB*, *pEGFP.attB* and *pEGFP-2.attB* (supplementary material Fig. S2B; see also the FlyC31 website at

www.flyc31.org). These vectors are perfectly suited to increase the versatility of the ΦC31 integrase system.

Enhanced transgenesis efficiency by pooled plasmid injections

To avoid time-consuming and tedious single construct injections, we mixed ORF constructs together and injected them as pools into the *ΦX-86Fb* host (Fig. 1A). Advantages of the pooling strategy include: (1) specific transgene plasmid purifications for injections are reduced to one per pool instead of many individual purifications; (2) the exchanges of injection needles are drastically reduced, which saves considerable time; (3) injected embryos do not have to be kept separately, which simplifies the handling; and, most importantly, (4) pool injections reduce the number of embryos injected substantially as one round of injection leads to the recovery of many different transgenes versus only one in the case of single construct injections.

Initially, we tested several pool sizes of up to 400 plasmids without seeing a specific effect of pool sizes on transgene recovery rates (supplementary material Table S1). Based on practical considerations, we continued to use pools of ~100 constructs for injections ('96-multiwell format'). Such a pool size led to a recovery of ~60% of the included ORFs from 100-120 outcrosses per injection round (supplementary material Table S1). The individual plasmid concentrations are normalised in the pool to maximise the recovery of all transgenes. The transgenic progeny is determined by single fly PCR, i.e. amplifying the plasmid region containing the barcode, followed by Sanger sequencing. Standard vector/*86Fb*-specific PCR primer pairs simultaneously identify the barcode and confirm the site-specific integration into the *86Fb* site (Fig. 1A). The previous assignment of an individual barcode to each ORF uniquely identifies the transgenic lines without needing to sequence into the actual ORFs. Each new line that we obtained was balanced, whereas repeatedly occurring lines were discarded. Constructs that were not recovered in the transgenic offspring were pooled again and injected in a subsequent round. Altogether, we created transgenic lines for 547 untagged and 602 HA-tagged ORFs (supplementary material Table S2).

Functional comparison of untagged and 3xHA-tagged transgenes

A tagged library holds major advantages over an untagged library: (1) a single antibody can be used to detect any ORF; (2) cross-reaction with related proteins can be avoided, as an antibody specific to the tag can be used; (3) the tagged protein can be distinguished from the endogenous, untagged protein; and (4) immunochemistry becomes possible for even poorly immunogenic proteins or proteins that lack a specific antibody (Jarvik and Telmer, 1998). However, because epitope tagging can also interfere with protein function, we examined the potential of the C-terminal 3xHA tag to alter protein function in overexpression experiments. We crossed all UAS-ORF lines that we had obtained in both versions to the wing-specific *MS1096-Gal4* driver line and recorded the phenotypes. The crosses were scored blind. We classified the effects on a scale from 0 (no effect) to 4 (lethal). Categories 1-3 represent different phenotypic strengths with 1 indicating mostly mild patterning or growth defects, 2 indicating severe growth defects, and 3 indicating complete or almost complete absence of wings. Overall, we could test 473 genes in both versions. Of these, 230 showed no phenotype in either version and were therefore not informative for this purpose. We also excluded 46 genes that showed a weak effect (category 1) in the untagged version but not the tagged version and 22 with the opposite behaviour, as they could

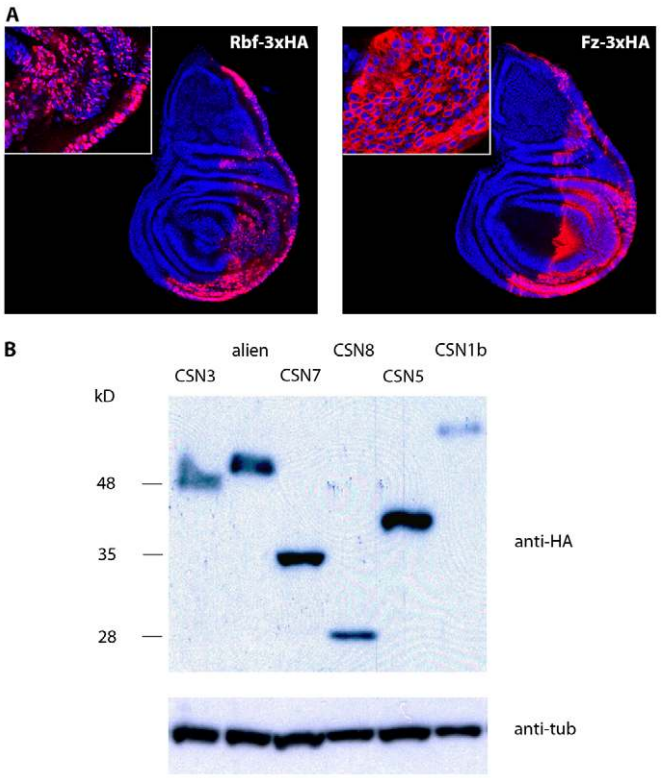

Fig. 2. Analysis of epitope-tagged proteins. (**A**) Anti-HA staining in third instar wing discs of overexpressed Rbf-3xHA (Retinoblastoma-family protein, left) and Fz-3xHA (Frizzled, right). The transgenes were specifically expressed in the posterior compartment by an *en-Gal4* (*engrailed*) driver (pink stained area to the right of each wing disc). The magnified insets highlight the expected nuclear (left panel) and cytoplasmic membrane (right panel) localisation of the respective proteins in wing discs. Nuclei are stained with DAPI (blue). (**B**) Western blot analysis showing that six components of the COP9 signalosome are specifically detected upon *GMR-Gal4* driver expression in the eye tissue. Protein molecular weights (in kDa) are indicated on the left.

have been false negatives due to slight experimental alterations, such as incubation temperature. Of the remaining 175 transgenes, 20 showed a strong phenotype (>2) in the tagged version but no effect in the untagged version (11.4%, 'false positive'). In addition, 38 showed a strong phenotype (>2) in the untagged version but no effect in the 3xHA version (21.7%, 'false negative'). The remaining genes induced similar phenotypes that sometimes varied in phenotypic strength by no more than one category (11 cases showed a divergence by two categories).

We also used the 3xHA epitope tag to monitor the subcellular localisation and the protein level of overexpressed proteins. Transgene expression was induced during larval development with the *en-Gal4* (*engrailed*) driver and their resulting expression was assayed in imaginal wing discs. The epitope-specific antibody detected the protein in the posterior compartment without obvious fluctuations in total abundance (Fig. 2A). Furthermore, protein abundance in the uninduced parts of the disc was below the detection limit of our staining protocol. It was also possible to identify subcellular expression patterns: nuclear (Rbf) and cytoplasmic membrane localisation (Fz) could be distinguished (Fig. 2A).

We also compared protein abundance in adult tissue samples. We chose components of the multiprotein COP9 signalosome (Wei et al., 2008) and tested their individual abundance by western blot

assays after misexpression in eye tissue with *GMR-Gal4*. The proteins are specifically detected without significant signs of degradation products. The total abundance, however, varies between the different proteins, probably owing to differences in translation rates and protein turnover (Fig. 2B).

In summary, these functional tests demonstrate that overexpression of HA epitope-tagged proteins typically, but not always, results in phenotypes similar to the untagged clones. The results also point out the need to evaluate potential effects of epitope tags with additional experiments. A detailed study validating the biological usefulness of this UAS-ORF library is presented elsewhere (Schertel et al., 2013).

An improved Gateway-compatible expression system

Cloning sequence-verified ORFs at a genome-wide scale into a system suitable for *in vivo* use represents a challenge for creating an ORFeome library. Importantly, the ORFs should be in a system that guarantees the flexible re-use of the ORFs without the need to re-amplify and re-verify them.

The Berkeley *Drosophila* Genome Project (BDGP) offers such a resource (Yu et al., 2011). The BDGP has generated thousands of sequence-verified 'movable ORFs', either with a native stop codon or without such a stop for C-terminal tagging. These ORFs are provided in a Cre/*loxP*-based gene transfer system, also known as Creator System (Clontech). We tested this system to make use of this ORF collection, but found it unsatisfactory in our *in vivo* misexpression tests. In our hands, the splicing required to attach the tag to the ORF was unspecific, generating multiple splice variants. Compared with our 3xHA-tagged Gateway clones, this resulted in lower expression of the correct splice version and therefore explains the complete absence of misexpression phenotypes (a detailed assessment of this approach can be found in supplementary material Appendix S1 and Fig. S4).

For further development of the cloning system, we reverted to the Gateway technology to preserve the advantage of re-using ORFs for multiple purposes once they are cloned and sequence verified. However, we made some modifications to the 3xHA-tagged destination vector by introducing two shortened and mutated FRTs (FLP recognition targets, namely *FRT2* and *FRT5*) (Schlake and Bode, 1994), which immediately flank the Gateway cassette (Fig. 3A). Recombination between these mutant FRT sites is reported to be incompatible (Larsen et al., 2006), which is essential for preventing the excision of the intervening sequence in presence of a FLP recombinase (FLP).

Because incompatibility between the selected FRT sites is crucial for our system, we performed multiple tests of the behaviour of the two FRT variants with clonal GFP assays *in vivo* (supplementary material Fig. S3). We examined each of the FRTs for recombination with an identical FRT, between each other, and together with a wild-type FRT. These tests were carried out in an intra-chromosomal fashion (supplementary material Fig. S3), essentially to exclude any recombination between the two mutated FRTs, and in an inter-chromosomal fashion with the purpose of demonstrating that efficient recombination does occur in *trans* between identical FRTs (not shown). We confirmed that the two FRT variants were incompatible, supporting the use of the selected FRTs for our promoter and tag swapping strategy, described below.

Exchanging promoters and tags *in vivo*

The two mutant FRTs allow distinct modification of the UAS-ORF library by FLP/FRT recombination *in vivo*. The C-terminal 3xHA tag can be exchanged with any other tag of choice, e.g. to switch to another epitope/antibody pair. Similarly, the UAS promoter region can be exchanged with any other promoter, including the possibility of adding N-terminal tags. The swapping events can be tracked by specific gain and loss of markers, e.g. the C-terminal exchange events can be recovered by selecting for w^+y^+ recombinants, the N-terminal exchange by selecting for y^+ (see Fig. 3B). Furthermore, N- and C-terminal exchanges can be combined sequentially by adjusting the markers in the 'promoter' and '3′-tagging' lines accordingly. For example, a combined promoter swap and tag swap could be achieved by first initiating the promoter swap and selecting for y^+ and then performing a tag swap with a w^+ marked tag-line and choosing y^+w^+ recombinants. The swapping device makes the ORF lines highly flexible for different applications and assays. Repeated injections to obtain new constructs are avoided after the one-time creation of the ORF and the specific swapping lines.

Currently, we have created two promoter lines (for *lexO* and *actin5C*) and three C-terminal tag lines (for *eGFP* and the two Venus YFP fragments *VNm9* and *VC155*) at cytological position *86F* that are compatible with our library. All five lines carry an *hsp70-FLP* construct on the X chromosome, necessary for the FLP/FRT inter-chromosomal recombination.

To test the efficiency and accuracy of this system, we generated a few UAS-ORF lines with a vector containing all the features depicted in Fig. 3A, named pTF-HA.attB and later upgraded to pGW-HA.attB. One of these lines, containing *bicoid* (*bcd*), was used to test swapping rates: males of *UAS-bcd* were crossed to either *lexA operator* (*lexO*) or to *eGFP* tag females, followed by a single 30-minute heat shock of the progeny at day 3; hatching females were outcrossed (four females per vial; F1 males can be used similarly, see Materials and methods) and the progeny was scored for the appropriate marker combinations, indicating successful swapping events. Under these conditions, the *lexO* promoter swapping resulted in recombinants found in 84% of the F1 crosses (n=39), whereas the exchange of the C-terminal tags occurred in 71% of the F1 crosses (n=24). These high rates are achievable with a convenient vial setup and a single heat-shock treatment. Finally, we confirmed the accuracy of the exchanges by PCR amplification and sequencing of the involved regions undergoing recombination (not shown).

Next, we validated the swapping by measuring protein expression before and after the swap (Fig. 4). Misexpression of *UAS-bcd*^{HA} by *dpp-Gal4* resulted in a distinct expression domain along the anterior/posterior (A/P) axis (Fig. 4A). Following a *UAS-hsp70* for *lexO-hsp70* promoter exchange, we overexpressed the *bcd*^{HA} transgene with the lexA TA (transactivator) driver *dpp-LHG* (Yagi et al., 2010), leading again to the distinct expression along the A/P axis (Fig. 4B). In this LHG-driven misexpression experiment, the wing discs displayed a pronounced fold in the wing pouch domain, different to the Gal4-driven expression for which we observed only a weak morphological effect. Misexpression of *bcd*^{HA} with *MS1096-Gal4* is lethal in males and causes significantly smaller wings in females (not shown). Thus, the observed morphological abnormality in the wing discs (Fig. 4B) is an early sign of the detrimental effects of *bicoid* misexpression. This early morphological alteration only appears weakly in the Gal4-driven *bicoid* expression (Fig. 4A), probably because the transcriptional activation by the Gal4 driver is substantially weaker than that observed for the driver LHG (Yagi et al., 2010). To visualise the 3′-tag exchange we used the transgene *UAS-emc* (*extra macrochaetae*) and monitored the expression of the protein via the 3xHA tag before (Fig. 4C) and after (Fig. 4D) the exchange with the *eGFP* tag. Both proteins, Emc^{HA} and Emc^{eGFP},

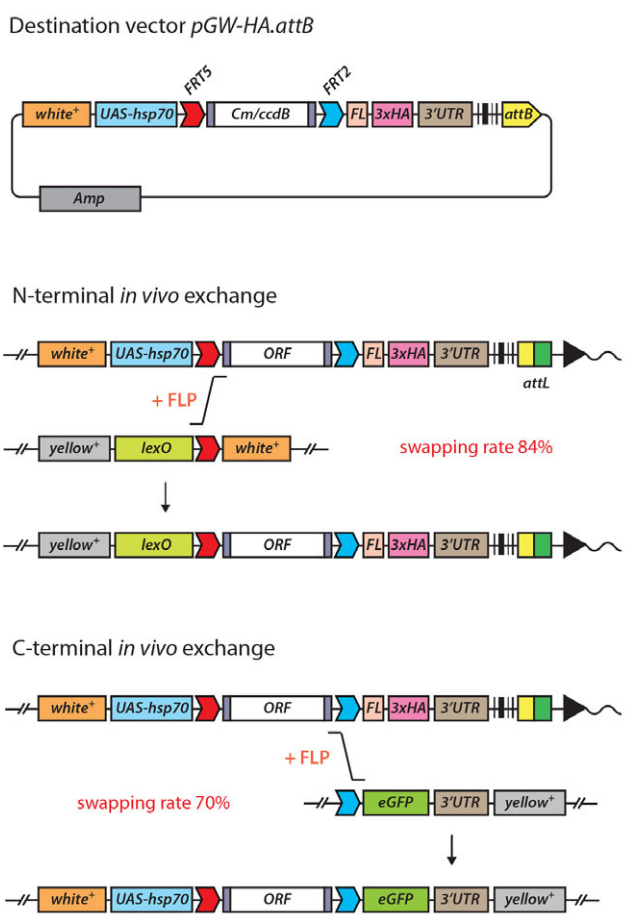

Fig. 3. Schematic of the *in vivo* swapping strategy. (**A**) Design of pGW-HA.attB destination vector. The mutant FRTs immediately flank the Gateway cassette, and the 3xHA tag is separated from the FRT by a 24-aa-long flexible linker (FL). (**B**) Illustration of possible *in vivo* swapping events. For N-terminal promoter swapping, the desired recombination event can be identified in the progeny by the single appearance of the *yellow* marker. For C-terminal tag swapping, the desired recombination can be monitored by the simultaneous occurrence of w^+ and y^+. Swapping rates shown indicate the number of vials containing transgenic flies (originating from female outcrosses, see Materials and methods). Activation of the FLP recombinase (+FLP) originates from a heat-inducible *flp* transgene present in the *lexO* promoter and the *eGFP* tag line (not indicated).

led to identical expression domains when driven by the *en-Gal4* line.

The ability to test for direct physical interactions between selected proteins with bimolecular fluorescence complementation (BiFC) analysis further demonstrates the versatility of the FRT-mediated swapping tool (Hu et al., 2002; Saka et al., 2007). This method visualises direct protein interactions by reconstituting a functional YFP from two non-fluorescent subfragments of YFP, which are fused to interacting proteins. We tested a previously documented (Jaw et al., 2000) (Fig. 4E) and a potential (Fig. 4F) protein-protein interaction by C-terminally tagging one ORF with the Venus YFP fragment VNm9 and the other with fragment VC155, followed by co-expression of the fusion proteins in the eye tissue by *GMR-Gal4*. A strong fluorescence signal was detected upon co-expression of HTH-VNm9 and EXD-VC155 (i.e. Homothorax and Extradenticle; Fig. 4E), confirming *in vivo* a direct protein-protein interaction as previously documented (Jaw et al., 2000). No specific fluorescence was detected when testing for interaction between HLHm5-VNm9 and HLH54F-VC155 (Fig. 4F).

Taking together, the modified FRT sites allow us to manipulate the promoters and tags of the library *in vivo* and validation tests have shown that this works accurately and efficiently. The BiFC analysis is one demonstration of using the swapping device, in this case to gain knowledge of potential *in vivo* protein-protein interactions.

DISCUSSION

Here, we provide an efficient strategy for creating a UAS-ORFeome library in *Drosophila*. Our effort to create such a library is mainly motivated by the experience that the majority of genes do not show 'obvious' loss-of-function phenotypes in conventional genetic screens (e.g. Miklos and Rubin, 1996). We propose that comprehensive and efficient gain-of-function screens using a UAS-ORF library will be a useful approach towards uncovering phenotypes for genes that have eluded loss-of-function screens. We extensively tested our approach with a pilot library of ~1200 transgenic lines and optimised the cloning, embryo injections and *in vivo* expression. The high integration rates and site specificity achieved with the ΦC31 integrase system reduces production time and improves the overall quality of the transgenic library, while avoiding the labour-intensive mapping required with traditional approaches. Molecular barcodes and plasmid pool injections further streamline the whole process.

The strategy presented here is a powerful way to create a genome-wide library for gain-of-function screens. However, some aspects require further consideration. The comparison between native and 3xHA-tagged clones highlights the problem of using tagged lines. Although the vast majority of the C-terminally 3xHA-tagged proteins behave like the untagged counterparts, we might miss up to 20% of candidates in a screen because of this epitope tag (false negative). In cases in which the tagged version causes a phenotype contrary to the wild-type protein (false positive), the tag might alter the protein stability, e.g. by masking a domain that is required for degradation. Potential interference of tags with protein function, for example leading to misrouting, has been demonstrated (Romano et al., 1998; Brothers et al., 2003). Furthermore, the accurate subcellular localisation of a protein might also be altered simply by overexpression. Approaches that tag proteins in their endogenous loci are likely to be more reliable strategies to reflect the accurate subcellular localisation (see Venken et al., 2011). Together, these results and considerations strongly emphasise the importance of including appropriate controls to address the effects of epitope tags. Despite this limitation, we favour the creation of a tagged version of the library, as this facilitates further analysis, such as immunohistochemistry, mass spectrometry or ChIP assays.

Generally, dominant-negative effects can result from overexpression (Herskowitz, 1987). For example, components of multiprotein complexes might be more prone to this effect, as an imbalance in the subunit composition can be fatal. Suspected dominant-negative effects, however, could be confirmed by screening corresponding mutants or RNAi lines.

Despite these limitations, this growing UAS-ORFeome library greatly extends the currently available tools for elucidating gene function in *Drosophila*. The Gal4/UAS system, clearly the most widely used binary expression system in *Drosophila* genetics, provides a rich repertoire of cell- and tissue-specific expression, which is continuously being expanded. Overexpression levels can be tuned in various ways and, therefore, using the UAS-ORF library will allow a broad range of phenotypic read-outs. Furthermore, this

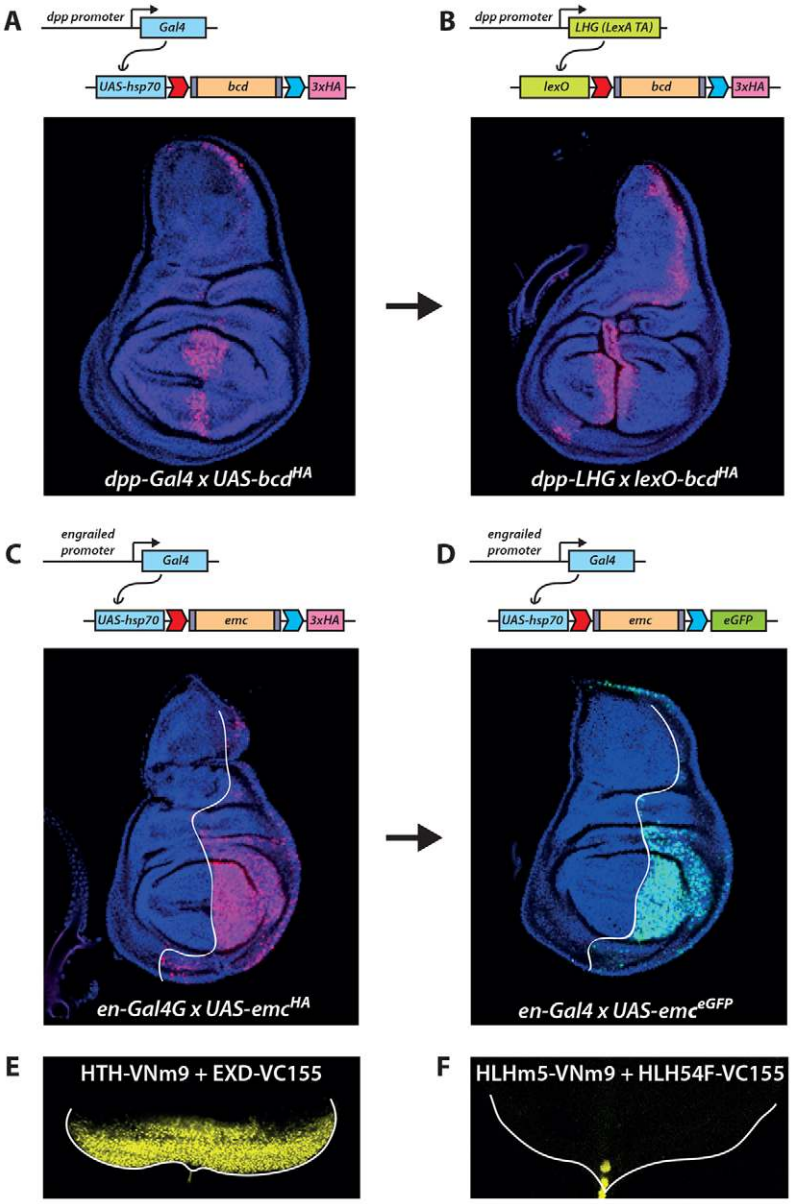

Fig. 4. Validation of promoter and C-terminal tag swapping. (**A-D**) The swapping events are demonstrated with two different transgenes in third instar wing discs: A and B represent a promoter exchange, and C and D represent a C-terminal tag exchange. Images show the situation before (A,C) and after (B,D) the *in vivo* exchange. The genetic status of each wing disc is illustrated by the schematics. (A) *UAS-bcd* overexpressed by *dpp-Gal4* and visualised by anti-HA immunostaining along the anterior/posterior border. (B) Localisation of Bicoid after the promoter swap, now being activated by LexA driver *dpp-LHG*. Here, the wing disc displays a pronounced fold in the wing pouch expression domain. This morphological abnormality is probably due to the considerably stronger expression caused by the specific LHG transactivator. (C) Localisation of 3xHA-tagged Emc (extra macrochaetae) in the posterior compartment (marked by pink staining). (D) Localisation of Emc after swapping, now visualised by the eGFP signal. The Emc^{eGFP} localisation resembles that of Emc^{HA}. Nuclei are stained with DAPI (blue). The (partially presumptive) boundaries between the anterior and posterior compartments are indicated by lines in C and D. (**E**,**F**) BiFC analysis of protein-protein interactions, tested in third instar eye-antennal discs. (E) Co-expression of the HTH-VNm9 and EXD-VC155 with *GMR-Gal4* resulted in a strong Venus YFP fluorescence signal, documenting physical interaction between the two proteins. (F) Absence of a fluorescence signal upon co-expression of HLHm5-VNm9 and HLH54F-VC155, indicating no interaction between these proteins. Boundaries of the eye disc are indicated by lines in E and F.

resource can also be used in combination with other genetic tools, such as the MARCM (mosaic analysis with a repressible cell marker) technique (Lee and Luo, 1999). Doing overexpression screens in modified backgrounds will often be the reasonable strategy. Unlike loss-of-function phenotypes, interpreting overexpression phenotypes will usually be more challenging (Prelich, 2012). For example, whether an observed phenotype results from an activating or inhibiting mechanism has to be determined. Testing the candidates for loss-of-function effects (i.e. mutants, RNAi) will be a likely next step.

The introduced FLP/FRT-mediated swapping technology is a convenient method for customising the library towards specific applications, as the appropriate tagging constructs need to be injected only once. Owing to the high exchange efficiencies, many transgenic lines can be converted in parallel. A very basic application is the creation of fluorescence-tagged transgenes for expression and localisation studies. This swapping device is also optimally suited to test potential protein-protein interactions *in vivo* with the BiFC method, as we have demonstrated. Recently, different protein inactivation methods were developed, such as 'TEV protease-induced protein inactivation' (TIPI) (Taxis et al., 2009) and 'degrade Green Fluorescent Protein' (deGradFP) (Caussinus et al., 2011). These methods require the attachment of a degron unit to the target protein or the creation of a GFP (or close derivative) fusion with the target protein, respectively. With the swapping device implemented in the library, this can be easily achieved. Generally, any sequence-encoded N- or C-terminal modification can be swapped onto a target protein with this exchange system.

The full genome-wide ORFeome library will take many years to complete. Here, we present the first set of UAS-ORF lines together with the detailed description of the technology underlying this project. The current subset comprises lines carrying cell cycle and growth control genes and, thus, is an interesting stand-alone library that can be readily used for various overexpression screens. These UAS-ORF lines will become publicly available (http://www.flyorf.ch) and new sub-collections, such as kinases/phosphatases and transcription factors, will be added continuously. In the longer term, we may consider expanding the

scope of the library to also include isoforms, specifically altered sets of genes (e.g. constitutively active or catalytically inactive mutant forms) or heterologous gene sets. This ORFeome project offers a powerful, new and continuously expanding resource for systematically uncovering and testing the function of genes in any given genetic pathway.

Acknowledgements
We thank Kenneth H. Wan, Mark Stapleton and Susan E. Celniker for a *pDNR-Dual* vector and information; Sini Miettinen, Aileen McLeod, Reto Städeli, Eliane Escher, Nellcia Wang and Cristina Bastos for technical assistance; Robert P. Zinzen, David Luginbuhl and Liqun Luo for expression analysis of our *UAS-lacZ-attP* lines; Jean-Paul Vincent for information on the mutated *FRT* versions; Damian Brunner for alerting us to an alternative linker sequence; Roland Baumgartner for advice and reagents for BiFC; Korneel Hens and Bart Deplancke for providing information on specific protein-protein interactions. We are very grateful to Jeremiah J. Zartman and Ludovic Baillon for critical reading of the manuscript.

Funding
This work was supported by the National Center of Competence in Research 'Frontiers in Genetics'; the Swiss National Science Foundation; the Kanton of Zürich; the European Research Council; the Academy of Finland; and the Scottish Universities Life Sciences Alliance.

Competing interests statement
The authors declare no competing financial interests.

Author contributions
J.T. and K.B. conceived and coordinated the project. J.B., M.B., J.T. and K.B. established the methodology and designed the experiments. J.B., M.B., E.F. and C.S. performed the experiments, compiled and analysed the data. J.B. wrote the manuscript with contributions from the co-authors.

Supplementary material
Supplementary material available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.088757/-/DC1

References

Basler, K. and Struhl, G. (1994). Compartment boundaries and the control of Drosophila limb pattern by hedgehog protein. *Nature* **368**, 208-214.
Bateman, J. R., Lee, A. M. and Wu, C. T. (2006). Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. *Genetics* **173**, 769-777.
Bellen, H. J., Levis, R. W., He, Y., Carlson, J. W., Evans-Holm, M., Bae, E., Kim, J., Metaxakis, A., Savakis, C., Schulze, K. L. et al. (2011). The Drosophila gene disruption project: progress using transposons with distinctive site specificities. *Genetics* **188**, 731-743.
Bischof, J., Maeda, R. K., Hediger, M., Karch, F. and Basler, K. (2007). An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. USA* **104**, 3312-3317.
Björklund, M., Taipale, M., Varjosalo, M., Saharinen, J., Lahdenperä, J. and Taipale, J. (2006). Identification of pathways regulating cell size and cell-cycle progression by RNAi. *Nature* **439**, 1009-1013.
Brand, A. H. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.
Brothers, S. P., Janovick, J. A. and Conn, P. M. (2003). Unexpected effects of epitope and chimeric tags on gonadotropin-releasing hormone receptors: implications for understanding the molecular etiology of hypogonadotropic hypogonadism. *J. Clin. Endocrinol. Metab.* **88**, 6107-6112.
Caussinus, E., Kanca, O. and Affolter, M. (2011). Fluorescent fusion protein knockout mediated by anti-GFP nanobody. *Nat. Struct. Mol. Biol.* **19**, 117-121.
Cavener, D. R. (1987). Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. *Nucleic Acids Res.* **15**, 1353-1361.
Dietzl, G., Chen, D., Schnorrer, F., Su, K. C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oppel, S., Scheiblauer, S. et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* **448**, 151-156.
Groth, A. C., Fish, M., Nusse, R. and Calos, M. P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. *Genetics* **166**, 1775-1782.
Herskowitz, I. (1987). Functional inactivation of genes by dominant negative mutations. *Nature* **329**, 219-222.
Horn, C., Jaunich, B. and Wimmer, E. A. (2000). Highly sensitive, fluorescent transformation marker for Drosophila transgenesis. *Dev. Genes Evol.* **210**, 623-629.
Hu, C. D., Chinenov, Y. and Kerppola, T. K. (2002). Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell* **9**, 789-798.
Jankovics, F. and Brunner, D. (2006). Transiently reorganized microtubules are essential for zippering during dorsal closure in Drosophila melanogaster. *Dev. Cell* **11**, 375-385.
Jarvik, J. W. and Telmer, C. A. (1998). Epitope tagging. *Annu. Rev. Genet.* **32**, 601-618.
Jaw, T. J., You, L. R., Knoepfler, P. S., Yao, L. C., Pai, C. Y., Tang, C. Y., Chang, L. P., Berthelsen, J., Blasi, F., Kamps, M. P. et al. (2000). Direct interaction of two homeoproteins, homothorax and extradenticle, is essential for EXD nuclear localization and function. *Mech. Dev.* **91**, 279-291.
Lai, S. L. and Lee, T. (2006). Genetic mosaic with dual binary transcriptional systems in Drosophila. *Nat. Neurosci.* **9**, 703-709.
Larsen, C., Franch-Marro, X., Hartenstein, V., Alexandre, C. and Vincent, J. P. (2006). An efficient promoter trap for detection of patterned gene expression and subsequent functional analysis in Drosophila. *Proc. Natl. Acad. Sci. USA* **103**, 17813-17817.
Lee, T. and Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451-461.
Miklos, G. L. and Rubin, G. M. (1996). The role of the genome project in determining gene function: insights from model organisms. *Cell* **86**, 521-529.
Parks, A. L., Cook, K. R., Belvin, M., Dompe, N. A., Fawcett, R., Huppert, K., Tan, L. R., Winter, C. G., Bogart, K. P., Deal, J. E. et al. (2004). Systematic generation of high-resolution deletion coverage of the Drosophila melanogaster genome. *Nat. Genet.* **36**, 288-292.
Prelich, G. (2012). Gene overexpression: uses, mechanisms, and interpretation. *Genetics* **190**, 841-854.
Ramani, A. K., Chuluunbaatar, T., Verster, A. J., Na, H., Vu, V., Pelte, N., Wannissorn, N., Jiao, A. and Fraser, A. G. (2012). The majority of animal genes are required for wild-type fitness. *Cell* **148**, 792-802.
Romano, J. D., Schmidt, W. K. and Michaelis, S. (1998). The Saccharomyces cerevisiae prenylcysteine carboxyl methyltransferase Ste14p is in the endoplasmic reticulum membrane. *Mol. Biol. Cell* **9**, 2231-2247.
Rørth, P. (1996). A modular misexpression screen in Drosophila detecting tissue-specific phenotypes. *Proc. Natl. Acad. Sci. USA* **93**, 12418-12422.
Saka, Y., Hagemann, A. I., Piepenburg, O. and Smith, J. C. (2007). Nuclear accumulation of Smad complexes occurs only after the midblastula transition in Xenopus. *Development* **134**, 4209-4218.
Schertel, C., Huang, D., Björklund, M., Bischof, J., Yin, D., Li, R., Wu, Y., Zeng, R., Wu, J., Taipale, J. et al. (2013). Systematic screening of a Drosophila ORF library in vivo uncovers Wnt/Wg pathway components. *Dev. Cell* **25**, 207-219.
Schlake, T. and Bode, J. (1994). Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry* **33**, 12746-12751.
Taxis, C., Stier, G., Spadaccini, R. and Knop, M. (2009). Efficient protein depletion by genetically controlled deprotection of a dormant N-degron. *Mol. Syst. Biol.* **5**, 267.
Venken, K. J., He, Y., Hoskins, R. A. and Bellen, H. J. (2006). P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science* **314**, 1747-1751.
Venken, K. J., Simpson, J. H. and Bellen, H. J. (2011). Genetic manipulation of genes and cells in the nervous system of the fruit fly. *Neuron* **72**, 202-230.
Viktorinová, I. and Wimmer, E. A. (2007). Comparative analysis of binary expression systems for directed gene expression in transgenic insects. *Insect. Biochem. Mol. Biol.* **37**, 246-254.
Wei, N., Serino, G. and Deng, X. W. (2008). The COP9 signalosome: more than a protease. *Trends Biochem. Sci.* **33**, 592-600.
Wittwer, F., Jaquenoud, M., Brogiolo, W., Zarske, M., Wüstemann, P., Fernandez, R., Stocker, H., Wymann, M. P. and Hafen, E. (2005). Susi, a negative regulator of Drosophila PI3-kinase. *Dev. Cell* **8**, 817-827.
Xu, R., Deng, K., Zhu, Y., Wu, Y., Ren, J., Wan, M., Zhao, S., Wu, X., Han, M., Zhuang, Y. et al. (2008). A large-scale functional approach to uncover human genes and pathways in Drosophila. *Cell Res.* **18**, 1114-1127.
Yagi, R., Mayer, F. and Basler, K. (2010). Refined LexA transactivators and their use in combination with the Drosophila Gal4 system. *Proc. Natl. Acad. Sci. USA* **107**, 16166-16171.
Yu, C., Wan, K. H., Hammonds, A. S., Stapleton, M., Carlson, J. W. and Celniker, S. E. (2011). Development of expression-ready constructs for generation of proteomic libraries. *Methods Mol. Biol.* **723**, 257-272.
Zhong, J. and Yedvobnick, B. (2009). Targeted gain-of-function screening in Drosophila using GAL4-UAS and random transposon insertions. *Genet. Res. (Camb.)* **91**, 243-258.