A genome-wide resource for the analysis of protein localisation in *Drosophila*

Mihail Sarov[1], Christiane Barz[2], Helena Jambor[1], Marco Y. Hein[3], Christopher Schmied[1], Dana Suchold[1], Bettina Stender[2], Stephan Janosch[1], Vinay K.J. Vikas[4], R.T. Krishnan[4], K. Aishwarya[4], Irene R. S. Ferreira[2], Radoslaw K. Ejsmont[1], Katja Finkl[2], Susanne Hasse[1], Philipp Kämpfer[5], Nicole Plewka[2], Elisabeth Vinis[1], Siegfried Schloissnig[5], Elisabeth Knust[1], Volker Hartenstein[6], Matthias Mann[3], Mani Ramaswami[7], K.VijayRaghavan[4], Pavel Tomancak[1] and Frank Schnorrer[2]

[1] Max Planck Institute of Cell Biology and Genetics, Pfotenhauerstr. 108, 01307 Dresden, Germany

[2] Muscle Dynamics Group, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

[3] Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

[4] Centre for Cellular and Molecular Platforms, National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bald Bellary Road, Bangalore 560065, India

[5] Heidelberg Institute of Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany

[6] Department of Molecular Cell and Developmental Biology, University of California, Los Angeles, 610 Charles E. Young Drive, 5009 Terasaki Life Sciences Building, Los Angeles, CA 90095, USA

[7] Institute of Neuroscience, Trinity College Dublin, Dublin 2, Ireland

26 correspondence should be addressed to:

27 sarov@mpi-cbg.de

28 tomancak@mpi-cbg.de

29 schnorrer@biochem.mpg.de

Abstract

The *Drosophila* genome contains >13,000 protein coding genes, the majority of which remain poorly investigated. Important reasons include the lack of antibodies or reporter constructs to visualise these proteins. Here we present a genome-wide fosmid library of ≈10,000 GFP-tagged clones, comprising tagged genes and most of their regulatory information. For 880 tagged proteins we have created transgenic lines and for a total of 207 lines we have assessed protein expression and localisation in ovaries, embryos, pupae or adults by stainings and live imaging approaches. Importantly, we can visualise many proteins at endogenous expression levels and find a large fraction of them localising to subcellular compartments. Using complementation tests we demonstrate that two-thirds of the tagged proteins are fully functional. Moreover, our clones also enable interaction proteomics from developing pupae and adult flies. Taken together, this resource will enable systematic analysis of protein expression and localisation in various cellular and developmental contexts.

Impact statement

We provide a large-scale transgenic resource, which enables live imaging, subcellular localisation and interaction proteomics of selected gene products at all stages of *Drosophila* development.

Introduction

With the complete sequencing of the *Drosophila* genome (Adams et al., 2000) genome-wide approaches have been increasingly complementing the traditional single gene, single mutant studies. This is exemplified by the generation of a genome-wide transgenic RNAi library (Dietzl et al., 2007) to systematically assess gene function in the fly or by the documentation of the entire developmental transcriptome during all stages of the fly's life cycle by mRNA sequencing (Graveley et al., 2010). Furthermore, expression patterns were collected for many genes during *Drosophila* embryogenesis by systematic mRNA *in situ* hybrisation studies in different tissues (Hammonds et al., 2013; Tomancak et al., 2002; 2007). Particularly for transcription factors (TFs) these studies revealed complex and dynamic mRNA expression patterns in multiple primordia and organs during development (Hammonds et al., 2013), supposedly driven by specific, modular enhancer elements (Kvon et al., 2014). Furthermore, many mRNAs are not only dynamically expressed but also subcellularly localised during *Drosophila* oogenesis (Jambor et al., 2015) and early embryogenesis (Lécuyer et al., 2007). Together, these large-scale studies at the RNA level suggest that the activity of many genes is highly regulated in different tissues during development. Since the gene function is mediated by the encoded protein(s), the majority of proteins should display particular expression and subcellular localisation patterns that correlate with their function.

However, a lack of specific antibodies or live visualisation probes thus far hampered the systematic survey of protein expression and localisation patterns in various developmental and physiological contexts in *Drosophila*. Specific antibodies are only available for about 450 *Drosophila* proteins (Nagarkar-Jaiswal et al., 2015), and the versatile epitope-tagged UAS-based overexpression collection that recently

74 became available (Bischof et al., 2013) is not suited to study protein distribution at
75 endogenous expression levels. Collections of knock-in constructs are either limited to
76 specific types of proteins (Dunst et al., 2015) or rely on inherently random genetic
77 approaches, such as the large-scale protein-trap screens or the recently developed
78 MiMIC (Minos Mediated Insertion Cassette) technology (Venken et al., 2011). The
79 classical protein-trap screens are biased for highly expressed genes, and altogether
80 recovered protein traps in 514 genes (Buszczak et al., 2007; Lowe et al., 2014; Morin
81 et al., 2001; Quiñones-Coello et al., 2007). The very large scale MiMIC screen
82 isolated insertions in the coding region of 1,862 genes, 200 of which have been
83 converted into GFP-traps available to the community (Nagarkar-Jaiswal et al., 2015).
84 Both approaches rely on transposons to mediate cassette insertion and require
85 integration into an intron surrounded by coding exons for successful protein tagging.
86 Thus, about 3,000 proteins, whose ORF is encoded within a single exon, cannot be
87 tagged by these approaches. Together, this creates a significant bias towards trapping
88 a particular subset of the more than 13,000 protein coding genes in the fly genome.

89 Hence, the *Drosophila* community would profit from a resource that enables
90 truly systematic protein visualisation at all developmental stages for all protein coding
91 genes, while preserving the endogenous expression pattern of the tagged protein. One
92 strategy to generate a comprehensive resource of tagged proteins is to tag large
93 genomic clones by recombineering approaches in bacteria and transfer the resulting
94 tagged clones into animals as third copy reporter allele as was previously done in *C.*
95 *elegans* (Sarov et al., 2012). In *Drosophila* it is possible to insert this tagged copy of
96 the gene as a transgene at a defined position into the fly genome (Venken et al.,
97 2006). The third copy reporter allele approach was used successfully with large
98 genomic BAC or fosmid clones derived from the fly genome. It has been shown that

these transgenes recapitulate the endogenous expression pattern of the gene in flies and thus likely provide a tagged functional copy of the gene (Ejsmont et al., 2009; Venken et al., 2009).

Here, we introduce a comprehensive genome-wide library of almost 10,000 C-terminally tagged proteins within genomic fosmid constructs. For 880 constructs, covering 826 different genes we generated transgenic lines, 765 of which had not been tagged by previous genetic trapping projects. Rescue experiments showed that two thirds of the tagged proteins are functional. We characterised the localisation pattern for more than 200 tagged proteins at various developmental stages from ovaries to adults by immunohistochemistry and by live imaging. This identified valuable markers for various tissues and subcellular compartments, many of which are detectable *in vivo* by live imaging. Together, this shows the wide range of possible applications and the potential impact this publically available resource will have on *Drosophila* research and beyond.

Results

Our goal was to generate a comprehensive resource that allows the investigation of protein localisation and physical interactions for any fly protein of interest through a robust, generic tagging pipeline in bacteria, which is followed by a large-scale transgenesis approach (**Figure 1**). We based our strategy on a *Drosophila melanogaster* FlyFos library of genomic fosmid clones, with an average size of 36 kb, which covers most *Drosophila* genes (Ejsmont et al., 2009). Our two-step tagging strategy first inserts a generic 'pre-tag' at the C-terminus of the protein, which is then replaced by any tag of choice at the second tagging step, for example with a superfolder-GFP (sGFP) tag to generate the sGFP TransgeneOme clone library. These tagged clones are injected into fly embryos to generate transgenic fly-TransgeneOme (fTRG) lines, which can be used for multiple *in vivo* applications. (**Figure 1**).

sGFP TransgeneOme – a genome-wide tagged FlyFos clone library

We aimed to tag all protein coding genes at the C-terminus of the protein, because a large number of regulatory elements reside within or overlap with the start of genes, including alternative promoters, enhancer elements, nucleosome positioning sequences, etc. These are more likely to be affected by a tag insertion directly after the start codon. Signal sequences would also be compromised by an inserted tag after the start codon. Additionally, the C-termini in the gene models are generally better supported by experimental data than the N-termini due to an historical bias for 3' end sequencing of ESTs. Thus, C-terminal tagging is more likely to result in a functional tagged protein than N-terminal tagging, although we are aware of the fact that some proteins will be likely inactivated by addition of a tag to the C-terminus. Moreover, only about 1,400 protein coding genes contain alternative C-termini, resulting in all

protein isoforms labelled by C-terminal tagging in almost 90 % of all protein coding genes.

In a series of pilot experiments we tested the functionality of several tagging cassettes with specific properties on a number of proteins (**Figure 2 Supplement 1, Table 1**). For the genome-wide resource we applied a two-step tagging strategy, whereby we first inserted a non-functional 'pre-tagging' cassette consisting of a simple bacterial selection marker, which is flanked with linker sequences present in all of our tagging cassettes. This strategy enables a very efficient replacement of the pre-tag by any tag of interest by homologous recombination mediated cassette exchange in bacteria. As fluorescent proteins and affinity tags with improved properties are continuously being developed, specific clones or the entire resource can be easily re-fitted to any new tagging cassette (**Figure 1**). For the genome-scale resource we selected a tagging cassette suitable for protein localisation and complex purification studies, consisting of the 2xTY1 tag as a flexible linker, the superfolder GFP coding sequence (Pédelacq et al., 2005), the V5 tag, followed by specific protease cleavage sites (for the PreScission- and TEV-proteases), the biotin ligase recognition peptide (BLRP) tag allowing for specific *in vivo* or *in vitro* biotinylation (Deal and Henikoff, 2010; Vernes, 2014), and the 3xFLAG tag (**Figure 2 Supplement 1**).

Of the 13,937 protein coding genes in the dmel5.43 genome assembly, 11,787 genes (84.6 %) were covered by a suitable fosmid from the original FlyFos library (Ejsmont et al., 2009), extending at least 2.5 kb upstream and 2.5 kb downstream of the annotated gene model. For picking clones, designing oligonucleotides for recombineering and for tracking all steps of the transgene engineering process, as well

as for providing access to all construct sequences and validation data we used the previously developed TransgeneOme database (Sarov et al., 2012).

For high-throughput tagging of the *Drosophila* FlyFos clones, we developed an improved version or our previously applied high-throughput, 96-well format liquid culture recombineering pipeline (Ejsmont et al., 2011; Sarov et al., 2012). The high efficiency of recombineering in *E. coli* allowed for multi-step DNA engineering in 96-well format liquid cultures with single clone selection only at the last step. The specific pipeline consists of five steps (**Figure 2A**). First, the pRedFlp helper plasmid containing all genes required for homologous recombination as well as the Flippase-recombinase (under L-rhamnose and tetracyclin control, respectively) was introduced into *E. coli* by electroporation. Second, the 'pre-tagging' cassette containing only a bacterial antibiotic resistance gene was inserted by homologous recombination with gene-specific homology arms of 50 base pairs. Third, the sGFP-V5-BLRP tagging cassette, including an FRT-flanked selection and counter-selection cassette, was inserted to replace the 'pre-tagging' cassette. Since the linker sequences in the 'pre-tagging' cassette are identical to the tagging cassette, the tagging cassette was simply excised from a plasmid by restriction digest and no PCR amplification was required. This strongly reduced the risk of PCR induced mutations in the tagging cassette. Fourth, the selection marker was excised by the induction of Flippase expression. Fifth, the helper plasmid was removed by suppression of its temperature sensitive replication at 37 ºC (Meacock and Cohen, 1980) and single clones were isolated from each well, by plating on selective solid agar plates.

All five steps of the engineering pipeline were highly efficient (between 95.8 and 99.7 %), resulting in an overall efficiency of 93.6 % or 10,995 growing cultures (**Figure 2A**). To validate the sequence of the engineered clones we developed a new

next-generation-sequencing (NGS)-based approach (**Figure 2B**). In short, we pooled single clones from all 96-well plates into 8 rows and 12 columns pools and deep sequenced the barcoded and pooled mate pair libraries using the Illumina platform. The mate pair strategy allowed us to map the otherwise common tag coding sequence to a specific clone in the library and thus to verify the integrity of the tagging cassette insertion in the clones with single nucleotide resolution (see Material and Methods for details). When applied to the final sGFP TransgeneOme collection we detected no mutations for 9,580 constructs (87.1 %). 8,005 (72.8 %) of these clones had complete sequence coverage in the tag and thus represent the most reliable subset of the tagged library (**Figure 2B**). For 1,417 of the clones (12.8 %) one or more differences to the expected sequences were detected. The most common differences were point mutations, which cluster almost exclusively to the homology regions in the oligonucleotides used to insert the 'pre-tagging' cassette. This is suggestive of errors in the oligonucleotide synthesis but could also reflect polymorphisms in the genomic sequence of the homology arms. Another subset of point mutations clustered around the junctions between the homology arms and the rest of the tagging cassette, indicating an imprecise resolving of the homology exchange reaction in small subset of clones (**Figure 2C**). Finally, a small group of clones (165) still contained an un-flipped selection cassette. The NGS results were confirmed by Sanger sequencing of the entire tag coding sequence for a subset of constructs (**Supplementary Table 1**).

Taken together, the sGFP TransgeneOme and our pilot tagging experiments resulted in 10,711 validated tagged clones, representing 9,993 different *Drosophila* genes. (**Table 1, Supplementary Table 1**). Moreover, the 'pre-tagged' TransgeneOme library is a versatile resource for generating fosmid clones with arbitrary tags at the C-terminus of the gene models.

Fly TransgeneOme (fTRG) – a collection of flies with tagged fosmids

We next established a pipeline to systematically transform the tagged TransgeneOme clones into flies. To efficiently generate fly transgenic lines we injected the tagged fosmid constructs into a recipient stock carrying the attP landing site VK00033 located at 65B on the third chromosome using a transgenic *nanos*-ΦC31 source (Venken et al., 2006). For some genes positioned on the third chromosome we injected into VK00002 located on the second chromosome at 28E to simplify genetic rescue experiments. In total, we have thus far generated lines for 880 tagged constructs representing 826 different genes (**Table 1**). These genes were partially chosen based on results of a public survey amongst the *Drosophila* community to identify genes for which there is the strongest demand for a tagged genomic transgenic line. 765 (87 %) of the newly tagged genes have not been covered by the previous protein-trap projects (**Supplementary Table 2**), hence, these should be particularly useful for the fly community. From our pilot tagging experiments, we made 51 lines for the 2xTY1-sGFP-3xFLAG tag and 30 lines for the 2xTY1-T2A-sGFPnls-FLAG transcriptional reporter. The majority of the lines (799) were generated with the versatile 2xTY1-sGFP-V5-Pre-TEV-BLRP-3xFLAG tag, used for the genome-wide resource (**Figure 2 Supplement 1, Table 1**). The collection of fly lines is called 'tagged FlyFos TransgeneOme' (fTRG) and all 880 fTRG lines have been deposited at the VDRC stock centre for ordering (http://stockcenter.vdrc.at).

To assess whether the tagged fosmids in our transgenic library are functional, we have chosen a set of 46 well-characterised genes, mutants of which result in strong developmental phenotypes. For most cases, we tested null or strong hypomorphic alleles for rescue of the respective phenotypes (embryonic lethality, female sterility,

237 flightlessness etc.) with the tagged fosmid lines. More than two-thirds of the lines (31
238 of 46), including tagged lines of *babo, dlg1, dl, fat, Ilk, LanB1, numb, osk, rhea, sax,*
239 *smo* and *yki* completely rescued the mutant phenotypes (**Figure 3**, **Supplementary**
240 **Table 3**), demonstrating that the majority of the tagged fosmids are functional. Our
241 rescue test set is biased towards important developmental regulators; 10 of the 15
242 genes that did not show a complete rescue are transcription factors with multiple
243 essential roles during development, such as *esg, eya, odd, sna* and *salm*. Thus, their
244 expression is likely regulated by complex cis-regulatory regions that may not be
245 entirely covered by the available fosmid clone; for example wing-disc enhancers are
246 located more than 80 kb away from the start of the *salm* gene (de Celis, 1999). Hence,
247 we expect that a typical gene, which is embedded within many other genes in the
248 middle of the fosmid clone, is more likely to be functional. Together, these data
249 suggest that both the genome-wide tagged construct library and the transgenic fTRG
250 library provide functional reagents that are able to substitute endogenous protein
251 function.
252
253 **Expression of fTRG lines in the ovary**
254 To demonstrate the broad application spectrum of our TransgeneOme library we
255 analysed tagged protein expression and subcellular localisation in multiple tissues at
256 various developmental stages. Germline expression in flies differs substantially from
257 somatic expression, requiring particular basal promoters and often specialised
258 3'UTRs (Ni et al., 2011; Rørth, 1998). Therefore, we used ovaries to test the fTRG
259 library and probed the expression of 115 randomly selected lines in germline cells
260 versus somatic cells during oogenesis (**Figure 4A**). From the 115 lines 91 (79 %)
261 showed detectable expression during oogenesis, with 45 lines being expressed in both,

germ cells and the somatic epithelial cells (**Figure 4B, C** and **Supplementary Table 4**). 76 (66 %) fTRG lines showed interesting expression patterns restricted to subsets of cells or to a subcellular compartment (**Figure 4B - D**). For example, Tan-GFP is expressed in germline stem cells only, whereas the ECM protein Pericardin (Prc-GFP) is concentrated around the neighbouring cap cells and the transcription factor Delilah (Dei-GFP) is specifically localised to the nuclei of somatic stem cells, which will give rise to the epithelial cells surrounding each egg chamber (**Figure 4A, C**). In early egg chambers Reph-GFP is expressed in germ cells only, whereas the ECM protein Viking (Vkg-GFP) specifically surrounds all the somatic epithelial cells. Interestingly, the transcription factor Auracan (Ara-GFP) is only expressed in posterior follicle cells, whereas the putative retinal transporter CG5958 is only detectable in the squamous epithelial cells surrounding the nurse cells (**Figure 4C**).

We further investigated the subcellular localisation of the tagged proteins, which revealed a localisation for the RNA helicase l(2)35Df to all nuclei, whereas the predicted C_2H_2-Zn-finger transcription factor Crooked legs (Crol-GFP) is restricted to the nuclei of the epithelial cells (**Figure 4D**). Interestingly, Corolla-GFP is exclusively localised to the oocyte nucleus in early egg chambers. This is consistent with the function of Corolla at the synaptonemal complex attaching homologous chromosomes during early meiosis (Collins et al., 2014). In contrast, the uncharacterised homeobox transcription factor E5 (E5-GFP) is largely restricted to the nuclei of anterior and posterior epithelial cells (**Figure 4D**). Apart from nuclear patterns, we found a significant number of cortical localisations, including the well characterised Crumbs (Crb-GFP) (Bulgakova and Knust, 2009) and the PDZ-domain containing Big bang (Bbg-GFP) (Bonnay and Cohen-Berros, 2013) at the apical cortex of the epithelial cells, the Na^+/K^+ transporter subunit Nervana 2 (Nrv2-GFP) at

the lateral epithelial membrane, and the EGF-signalling regulator Star (S-GFP) as well as the TGF-β receptor Saxophone (Sax-GFP) localised to the cortex or membrane of the germ cells **(Figure 4D** and **Supplementary Table 4**). Furthermore, we find a perinuclear enrichment for the uncharacterised predicted NAD binding protein CG8768, and oocyte enrichments for the Tom22 homolog Maggie (Mge-GFP) (Vaskova et al., 2000), glycosyltransferase Wollknäuel (Wol-GFP) (Haecker et al., 2008) and the TGF-α homolog Gurken (Grk-GFP), the latter with its well established concentration around the oocyte nucleus (Neuman-Silberberg and Schüpbach, 1993) (**Figure 4D**).

To test if genes expressed from FlyFos system also undergo normal post-transcriptional regulation we analysed the *osk-GFP* line, which was recently used as a label for germ granules (Trcek et al., 2015). *osk* mRNA is transcribed from early stages of oogenesis onwards in the nurse cell nuclei and specifically transported to the oocyte, where it localises to the posterior pole (Ephrussi et al., 1991; Kim-Ha et al., 1991). Only after the mRNA is localised, it is translated from stage 9 onwards (Kim-Ha et al., 1995). Indeed, fosmid derived *osk-GFP* mRNA localizes normally during all stages of oogenesis and its translation is repressed during mRNA transport, as Osk-GFP can only be detected at the posterior pole from stage 9 onwards (**Figure 4 Supplement 1A, B**). *osk*-GFP also fully rescues all aspects of an *osk* null allele (**Figure 3B, C**). Additionally, we discovered a post-transcriptional regulation for *corolla*. *corolla-GFP* mRNA is localised to the oocyte at stage 6 and Corolla-GFP protein is transported into the oocyte nucleus. However, despite the presence of the *corolla-GFP* mRNA at stage 8, Corolla protein is undetectable, suggesting either a translational block of the RNA or targeted degradation of the protein (**Figure 4 Supplement 1C - F**). Taken together, these expression and protein localisation data

recapitulate known patterns accurately and identify various unknown protein localisations in various cell types during oogenesis, and thus emphasise the value of the fly TransgeneOme resource.

Live in toto imaging of fTRG lines during embryogenesis

For many genes, the expression patterns at the mRNA level are particularly well characterised during *Drosophila* embryogenesis (Hammonds et al., 2013; Tomancak et al., 2002; 2007). However, *in situ* hybridisation techniques on fixed tissues do not visualise dynamics of expression over time and thus do not allow tracking of the expressing cells during development. As our tagging approach enables live imaging at endogenous expression levels we set out to test if *in toto* imaging using the SPIM (Selective Plane Illumination Microscopy) technology (Huisken et al., 2004) can be applied to the fly TransgeneOme lines. We pre-screened a small subset of lines (**Supplementary Table 5**) and selected the Na^+/K^+ transporter subunit Nrv2, as it shows high expression levels, for long-term time-lapse live imaging with a multi-view dual-side SPIM (Huisken and Stainier, 2007). During embryogenesis Nrv2 expression was reported in neurons (Sun et al., 1999) and glial cells (Younossi-Hartenstein et al., 2002). Interestingly, we find that Nrv2-GFP is already expressed from stage 11 onwards in most likely all cell types, where it localises to the plasma membrane, similar to the localisation in ovaries (**Figure 4D**). The expression level increases during stage 15 in all cells, with a particularly strong increase in the developing central nervous system (CNS) labelling the CNS and motor neuron membranes (**Figure 5, Supplementary Movie 1**[1]). These live *in toto* expression data are

[1] Examine raw data in BigDataViewer (Fiji -> Plugins -> BigDataViewer -> Browse BigDataServer and http://bds.mpi-cbg.de:8087

consistent with expression data of a recently isolated GFP trap in *nrv2* (Lowe et al., 2014), thus validating our methodology.

We wanted to extend our approach beyond highly expressed structural genes towards transcription factors that enable to follow cell lineages in the embryo. For this purpose we crossed the fTRG line of the homeobox transcription factor *gooseberry* (Gsb-GFP) to H2A-mRFPruby, which labels all nuclei, and recorded a two-colour multi-view dual-side SPIM movie. We find that Gsb-GFP is expressed in the presumptive neuroectoderm of the head region, labelling segmentally reiterated stripe-like domains at stage 10 (**Figure 5 Supplement 1A, B, Supplementary Movie 2**) as was described from fixed images (Gutjahr et al., 1993). Focusing on the deuterocerebral domain, we could reconstruct the delamination of two neuroblasts, which up-regulate Gsb-GFP while initiating their asymmetric divisions (**Figure 5 Supplement 1C - F**). It was possible to individually follow their neural progeny. Gsb-GFP expression also allowed us to directly follow the gradual down-regulation of Gsb-GFP in ectodermal cells that remained at the head surface after neuroblast delamination. As opposed to the neuroblasts, these cells, which give rise to epidermis, did not divide at all, or underwent only one further division (**Figure 5 Supplement 1G - J, Supplementary Movie 2**).

gsb is in part required for *gooseberry-neuro* (*gsb-n*; also called *gsb-d*) expression (He and Noll, 2013), which marks a defined set of brain neuroblasts in the embryo, with three *gsb-n* positive neuroblasts in the deuterocerebral domain (Urbach and Technau, 2003). We assume that these cells correspond to the three Gsb-GFP expressing neuroblasts visible in our live 4D stack. Notably expression in the Gsb-n-GFP fTRG line becomes detectable only at the end of germ-band extension (end of stage 11) in the developing CNS, where it lasts until stage 17 (**Figure 5 Supplement**

360 **2**, **Supplementary Movie 3**). Again this is consistent with published immuno-
361 histochemistry data (Gutjahr et al., 1993; He and Noll, 2013). We conclude that our
362 fly TransgeneOme library can be used for live *in toto* imaging, even for transcription
363 factors expressed at endogenous levels. This will be of significant importance for on-
364 going efforts linking the transcription factor expression patterns of embryonic
365 neuroblasts to the morphologically defined lineages that structure the larval and adult
366 *Drosophila* brain (Hartenstein et al., 2015; Lovick et al., 2013; Pereanu, 2006).
367
368 **Expression of fTRG lines in the adult thorax**
369 Cells in the embryo are generally rather small and thus not ideally suited to document
370 subcellular protein localisation patterns. Thus, we decided to apply our
371 TransgeneOme library to the adult thorax that contains some very large cells. We
372 scored expression in the large indirect flight muscles (IFMs), in leg muscles, in
373 visceral muscles surrounding the gut, in the gut epithelium, the tendon epithelium, the
374 trachea and the ventral nerve cord including the motor neurons. In total, we found
375 detectable expression in at least one tissue for 101 of 121 (83.5 %) analysed fTRG
376 lines, thus creating a large number of valuable markers for cell types and subcellular
377 structures (**Supplementary Tables 6, 7**).
378 The large IFMs are fibrillar muscles, which have a distinct transcriptional
379 program resulting in their distinct morphology (Schönbauer et al., 2011). This is
380 recapitulated by the expression of Act88F-GFP, which localises to the thin filaments
381 of IFMs only (**Figure 6A - C**), whereas Mlp84B-GFP is not in IFMs but at the
382 peripheral Z-discs of leg and visceral muscles only (**Figure 6D - F**), similar to the
383 published localisation in larval muscle (Clark et al., 2007). We find various dotty
384 patterns indicating localisation to intracellular vesicles; a particularly prominent

example is Tango1-GFP in the midgut epithelium (**Figure 6G, H, Supplementary Table 6**). Tango1 regulates protein secretion in S2 cells, where it localises to the Golgi apparatus upon over-expression (Bard et al., 2006), suggesting that the pattern described here is correct. We find Par6 with the expected apical localisation in the proventriculus epithelium and in trachea (**Figure 6I, J**), whereas we identified a surprising pattern for the TRP channel Painless (Tracey et al., 2003). Pain-GFP is not only highly expressed in motor neurons (**Figure 6K**) but also in particular cells in the gut epithelium and most surprisingly, in the tendon cells to which the IFMs attach (**Figure 6L, M**). At this point we can only speculate that Pain might be involved in mechanical stretch-sensing in these cells. We have also tagged various ECM components, with LamininB1 (LanB1-GFP), LamininA (LanA-GFP) and BM40-SPARC resulting in the most prominent expression patterns. All three ensheath most adult tissues, particularly the muscles (**Figure 6N, O**, **Figure 6 Supplement 1A, B, E, F**). Interestingly, LanA-GFP and LanB1-GFP also surround the fine tracheal branches that penetrate into the IFMs, whereas BM40-SPARC is only detected around the large tracheal stalk and the motor neurons (**Figure 6P**, **Figure 6 Supplement 1C, D, G, H**). Finally, we also detected prominent neuromuscular junction (NMJ) markers; the IκB homolog Cactus shows a distinct pattern on leg muscles, visceral muscles and IFMs, the latter we could confirm by co-staining with the neuronal marker Futsch (**Figure 6 Supplement 1I - M**). Interestingly, such a NMJ pattern for Cactus and its binding partner Dorsal has been shown in larval body muscle by antibody stainings (Bolatto et al., 2003). Together, these results suggest that our fly TransgeneOme library provides a rich resource for tissue-specific markers in the adult fly that can routinely be used to visualise subcellular compartments in various tissues.

To further validate the advantages of our TransgeneOme lines to label subcellular structures we imaged the large IFMs of the same 121 lines at high resolution. We found various markers for the thick filaments (e.g. the myosin associated protein Flightin, Fln-GFP) (Vigoreaux et al., 1993), for the myofibrils (e.g. the protein kinase Fray-GFP), the M-lines (e.g. the titin related protein Unc-89/Obscurin-GFP) (Katzemich et al., 2012), the Z-discs (e.g. CG31772-GFP) and the muscle attachment sites (e.g. Integrin-linked-kinase, Ilk-GFP). Furthermore, we identified markers for the T-tubules (e.g. Dlg1-GFP), for different vesicular compartments (e.g. the TGFβ receptor Baboon-GFP) and for mitochondria (CG12118) within the IFMs (**Figure 7, Supplementary Tables 6, 7**). Additionally, we documented the nuclear localisation in IFMs and leg muscles for a variety of fTRG proteins, including the uncharacterised homeodomain protein CG11617 and the C_2H_2 Zinc-fingers CG12391 and CG17912 (**Figure 7 Supplement 1A - C, E - G**); both of the latter result in flightless animals when knocked-down by muscle-specific RNAi (Schnorrer et al., 2010) suggesting that these genes play an essential role for IFM morphogenesis or function. Interestingly, the well characterised C_2H_2 Zinc-finger protein Hunchback (Hb) is only localised to leg muscle nuclei, but absent from IFMs suggesting a leg muscle-specific function of Hb (**Figure 7 Supplement 1G, H**).

However, differences between muscle types are not only controlled transcriptionally but also by alternative splicing (Oas et al., 2014; Spletter and Schnorrer, 2014; Spletter et al., 2015). To investigate if our tagging approach can be used to generate isoform-specific lines, we have chosen two prominent muscle genes, *mhc* and *rhea* (the fly Talin), both of which have predicted isoforms with different C-termini (**Figure 7 Supplement 2A, H**). Interestingly, we found that Mhc-isoforms K, L, M are expressed in IFMs and all leg muscles, however the predicted Mhc-isoforms

434 A, F, G, B, S, V with the distal STOP codon are selectively expressed in visceral
435 muscle and in a subset of leg muscles, however absent in adult IFMs (**Figure 7**
436 **Supplement 2B - G**). Even more surprisingly, while the long 'conventional' *rhea*
437 (Talin) isoforms B, E, F, G show the expected localisation to muscle attachment sites
438 in IFMs and leg muscles (Weitkunat et al., 2014), the short Talin-isoforms C and D
439 do not localise to muscle attachment sites, but are selectively concentrated at
440 costamers of leg muscles (**Figure 7 Supplement 2I - N**). Hence, our TransgeneOme
441 library is ideally suited to label subcellular compartments and protein complexes, and
442 in some cases can even distinguish between closely related protein isoforms.
443
444 **Expression of the fTRG lines in the living pupal thorax**
445 An attractive application of the fly TransgeneOme library is live *in vivo* imaging. In
446 the past, we had established live imaging of developing flight muscles in the pupal
447 thorax using overexpressed marker proteins (Weitkunat et al., 2014). Here, we wanted
448 to test, if live imaging of proteins at endogenous expression levels is also possible in
449 the thick pupal thorax. We selected six fTRG lines for well established genes and
450 indeed could detect expression and subcellular localisation for all of them using a
451 spinning-disc confocal microscope either at the level of the pupal epidermis or below
452 the epidermis, in the developing flight muscles, or both (**Figure 8**). The adducin-like
453 Hts-GFP labels the cytoplasm of fusing myoblasts from 10 to 20 h APF (after
454 puparium formation) and the developing SOPs (sensory organ precursors) with a
455 particular prominent concentration in developing neurons and their axons (**Figure 8B**
456 **- E**). In contrast, Dlg1-GFP localises to cell-cell-junctions of the pupal epidermis and
457 to a network of internal membranes in the developing IFMs (**Figure 8F - I**) that may
458 resemble developing T-Tubules, for which Dlg1 is a well established marker (Razzaq

et al., 2001). Interestingly, the long isoforms of Talin-GFP (*rhea* isoforms B, E, F, G) are largely in the cytoplasm and at the cortex of the epidermal cells, with a marked enrichment in the developing SOPs at 10 to 20 h APF (**Figure 8J, K**). Further, Talin-GFP is strongly concentrated at muscle attachment sites of developing IFMs from 24 h onwards (**Figure 8L, M**) consistent with antibody stainings of IFMs (Weitkunat et al., 2014).

The dynamics of the extracellular matrix is little described thus far as very few live markers existed. Hence, we tested our LamininB1 fosmid and found that LanB1-GFP is readily detectable within the developing basement-membrane basal to the epidermal cells of the pupal thorax at 10h APF (**Figure 8N**). It also labels the assembling basement-membrane around the developing IFMs from 16 to 30h APF without a particularly obvious concentration at the muscle attachment sites (**Figure 8O - Q**). To specifically visualise the developing IFMs we chose Actin88F, which is specifically expressed in IFMs and a few leg muscles (Nongthomba et al., 2001). We find that the Act88F-GFP fTRG line indeed very strongly labels the IFMs from about 18 h APF but is also expressed in the developing pupal epidermis again with an enrichment in the forming SOPs from 10 to 20 h APF (**Figure 8R - U**). The latter is not surprising as Act88F-lacZ reporter has been shown to also label the developing wing epithelium (Nongthomba et al., 2001), again suggesting that our fTRG line recapitulates the endogenous expression pattern. We finally tested the βTub60D fTRG-line, as βTub60D was reported to label the myoblasts and developing myotubes in embryonic and adult muscles (Fernandes et al., 2005; Leiss et al., 1988; Schnorrer et al., 2007). Indeed, we detect βTub60D-GFP in fusing myoblasts and the developing IFMs, with particularly prominent label of the microtubule bundles at 24h APF (**Figure 8V - Y**). In addition, βTub60D-GFP also strongly marks the developing hairs

484 of the sensory organs of the pupal epidermis (**Figure 8X**, see also **Supplementary**
485 **Movie 6**).

486 In order to test, if the fly TransgeneOme lines and the sGFP-tag are indeed
487 suited for long-term live imaging we chose Act88F-GFP and βTub60D-GFP and
488 imaged the developing IFMs for more than 19 h with a two-photon microscope using
489 an established protocol for over-expressed markers (Weitkunat and Schnorrer, 2014).
490 For both proteins we can detect strongly increasing expression after 18h APF in the
491 developing IFMs, with Act88F-GFP being restricted to the myotubes and the
492 developing myofibrillar bundles (**Supplementary Movie 4**, **Figure 8 Supplement**
493 **1A - F**) whereas βTub60D-GFP also labels the fusing myoblasts and is largely
494 incorporated into prominent microtubule bundles (**Supplementary Movie 6**, **Figure 8**
495 **Supplement 1L - Q**).

496 As photo bleaching was no serious problem in these long movies we also
497 recorded movies at higher time and spatial resolution. We labelled the developing
498 IFMs with Act88F-GFP and the myoblasts with a *him*-GAL4, UAS-palm-Cherry and
499 acquired a 3D stack every two minutes using a spinning disc-confocal. This enabled
500 us to visualise single myoblast fusion events in developing IFMs of an intact pupa
501 (**Supplementary Movie 5**, **Figure 8 Supplement 1G - K**). The six dorsal
502 longitudinally oriented IFMs develop from three larval template muscles to which
503 myoblasts fuse to induce their splitting into six myotubes (Fernandes et al., 1991).
504 Using high resolution imaging of βTub60D-GFP we find that most myoblasts fuse in
505 the middle of the developing myotube during myotube splitting, with prominent
506 microtubules bundles located at the peripheral cortex of the splitting myotube
507 (**Supplementary Movie 7, Figure 8 Supplement 1R**). These prominent microtubules
508 bundles are then relocated throughout the entire developing myotube

509 (**Supplementary Movie 7**, **Figure 8 Supplement 1S - V**). Taken together, these live
510 imaging data suggest that many of the fTRG lines will be well suited for high
511 resolution live imaging of dynamic subcellular protein localisation patterns in
512 developing *Drosophila* organs. This will strongly expand the set of live markers
513 available for research in flies.
514
515 **FlyFos library as bait for proteomics**
516 For the proper composition, localisation and *in vivo* function of most protein
517 complexes the expression levels of the individual components are critical (Rørth et al.,
518 1998; Tseng and Hariharan, 2002). Hence, the TransgeneOme library would be an
519 ideal experimental set-up to purify protein complexes from different developmental
520 stages using endogenous expression levels of the bait protein. In principle, all the
521 small affinity tags (TY1, V5, FLAG) (**Figure 2 Supplement 1**) can be used for
522 complex purifications. The presence of precision and TEV cleavage sites even allow
523 two-step purifications. For proof of principle experiments, we selected four tagged
524 proteins as baits: Ilk, Dlg1, Talin and LanB1, and analysed two different
525 developmental stages. In each case we homogenised hundred 24 to 48 h pupae and
526 hundred adult flies per experiment and mixed the cleared lysate with a GFP antibody
527 matrix to perform single step affinity enrichment and mass-spec analysis modified
528 from the QUBIC protocol (Hein et al., 2015; Hubner et al., 2010; Keilhauer et al.,
529 2014). Each affinity-enrichment was performed in triplicate and intensity profiles of
530 all identified proteins were quantified in a label-free format by running all 30
531 purifications consecutively on the same Orbitrap mass-spectrometer and analysing the
532 data with the MaxQuant software suite (Cox and Mann, 2008; Cox et al., 2014)
533 (**Supplementary Table 8**). Interestingly, enriching Ilk-GFP from both developing

pupae and adult flies recovered the entire Ilk, PINCH, Parvin, RSU-1 complex (**Figure 9**), which had previously been purified *in vitro* from *Drosophila* S2 cells (Kadrmas et al., 2004) and mammalian cells (Dougherty et al., 2005; Tu et al., 2001) giving us confidence in our methodology. We also successfully enriched Talin-GFP from pupae or adults, however did not identify an obvious strong and specific binding partner (**Figure 9**, **Supplementary Table 8**). In contrast, we identified Mesh as a novel interactor of Dlg1 from pupae and adult flies. Mesh colocalises with Dlg1 at septate junctions of the embryonic *Drosophila* midgut, however a molecular interaction of both proteins was not established (Izumi et al., 2012). Finally, we purified the laminin complex by pulling on LanB1, which recovered LanB2 and LanA roughly stoichiometrically, both from pupae and adult flies, as had been found in cell culture experiments (Fessler et al., 1987), showing that extracellular matrix complexes can also be purified from *in vivo* samples with our methodology. In summary, these data demonstrate that interaction proteomics with the fly TransgeneOme library can confirm known interaction partners and discover novel *in vivo* complex members, making the system attractive for a variety of biochemical applications.

Discussion

The TransgeneOme resource presented here adds new powerful component into the arsenal of tools available to the *Drosophila* research community. It complements the genetic resources for gene disruption and localisation (Buszczak et al., 2007; Lowe et al., 2014; Morin et al., 2001; Nagarkar-Jaiswal et al., 2015; Quiñones-Coello et al., 2007; St Johnston, 2012; Venken and Bellen, 2012) with a comprehensive genome-scale library that does not suffer the biases of random mutagenesis. Analogously to the powerful MiMIC system (Nagarkar-Jaiswal et al., 2015) the TransgeneOme resource is versatile and can be adapted to the developments in tag chemistry and to various specialised applications. Although the resource is designed to study behaviour of proteins, it can for example be easily converted into a toolkit for live imaging of mRNAs. By designing a tagging cassette with an array of MS2 binding sites (Forrest and Gavis, 2003) the existing 'pre-tagged' TransgeneOme can be converted into an MS2-tagged TransgeneOme by a simple liquid culture recombineering step in bacteria (**Figure 1**).

However, any new TransgeneOme has to be transformed into flies and this process still represents a significant bottleneck. We present here an optimised protocol for transgenesis of fosmid size clones into *Drosophila melanogaster*. It took three years and four dedicated technicians to generate the 880 fly lines presented in this study. Although, the systematic transgenesis is a continuing process in our laboratories, the value of the TransgeneOme collection is highlighted by the fact that any specific set of genomically tagged gene clones is now available. These can be efficiently transformed by in house transgenesis of *Drosophila* labs around the world using the optimised protocol presented here. For that purpose we will not only deposit the sGFP TransgeneOme but also the 'pre-tagged' TransgeneOme collection at Source

Biosciences from which any subset of clones can be conveniently ordered (http://www.sourcebioscience.com/).

One caveat of designed expression reporters is the necessity to place the tag into a defined position within the gene model. We chose to generate our 'pre-tagged' collection at the most commonly used C-terminus predicted by the gene model, thus labelling most isoforms. In a few cases a tag at the C-terminus will inactivate the protein, however such a reagent can still be useful for visualising the protein. This has been demonstrated for a number of sarcomeric GFP-traps, some of which lead to lethality when homozygous, yet result in interesting localisation patterns when heterozygous (Buszczak et al., 2007; Morin et al., 2001 and F.S., unpublished observation). For particular genes, it will be useful to tag differential protein isoforms, which in some cases can be done by tagging alternative C-termini, as shown here for *Mhc* and *rhea*. However, tagging a particular isoform requires a very informed construct design, which cannot easily be automated at the genome-scale.

Genome engineering is experiencing a tremendous growth with the introduction of CRISPR/Cas technology and it will be only a matter of time before a larger collection of precisely engineered fusion proteins at endogenous loci will become available in flies. However to date, such examples are still limited to a few genes (Baena-Lopez et al., 2013; Gratz et al., 2014; Port et al., 2014; Zhang et al., 2014), which had been carefully picked and were individually manipulated with custom-designed, gene-specific tools. It remains to be tested which proportion of such engineered loci will be fully functional and thus potentially superior to the fTRG collection. Having a transgenic third allele copy, as is the case in our TransgeneOme collection, might even be advantageous, if the tagging interferes with protein function, because the TransgeneOme lines still retain two wild-type endogenous gene

601 copies. In some cases, addition of GFP might destabilise the protein, regardless of N-
602 or C-terminal fusion, as recently shown for the Engrailed protein (Sokolovski et al.,
603 2015). However, our ability to detect the protein product in the vast majority of our
604 tagged lines argues that this could be a relatively rare, gene specific phenomenon.
605 Nevertheless, caution should be taken with respect to protein turnover dynamics of
606 any tagged protein.

607 Together, the FlyFos library, the fly TransgeneOme library and the fTRG
608 collection of strains, enable genome-scale examination of expression and localisation
609 of proteins comparable with the high-throughput mRNA *in situ* screens (Tomancak et
610 al. 2002, Tomancak et al. 2007). Our data for tagged Oscar protein show that fosmid
611 reporters can in principle recapitulate all aspects of gene expression regulation at
612 transcriptional and post-transcriptional levels. It will be particularly interesting to
613 combine the spatial expression data of mRNAs with that of proteins. Since many
614 transcripts show subcellular localisation in various developmental contexts (Jambor et
615 al., 2015; Lécuyer et al., 2007), the question arises whether RNA localisation
616 generally precedes localised protein activity. Systematic examination of protein
617 patterns expressed from localised transcripts in systems such as the ovary will provide
618 a genome-scale overview of the extent and functional role of translational control. At
619 the tissue level, the patterns of mRNA expression may be different from the patterns
620 of protein expression, for example due to translational repression in some cells or
621 tissue specific regulation of protein stability, as shown here for the Corolla protein.
622 The combined mRNA and protein expression patterns may therefore uncover a hidden
623 complexity in overall gene activity regulation and the fTRG lines will help to reveal
624 these combinatorial patterns in a systematic manner.

The fTRG lines faithfully recapitulate gene expression patterns in ovaries, embryos, larvae, pupae and adults suggesting that they can be used to visualise proteins in every tissue during the life cycle of the fly. This includes adult tissues such as the flight or leg muscles, which thus far had not been subjected to systematic protein expression and localisation studies. However, due to their size and the conservation of the contractile apparatus these tissues are particularly attractive to study with this new resource. In general, antibody or FISH stainings with a single standard anti-tag reagent are easier to optimise, compared to antibody stainings or mRNA *in situ* with gene specific antibodies/probes. This simplicity makes it possible to explore the expression of the available genes across multiple tissues, as has been done for the *rab* collection (Dunst et al., 2015). Such an approach is orthogonal to the collections of expression data generated thus far, in which many genes were examined systematically but only at particular stages or in certain tissues, i.e. embryos or ovaries (Jambor et al., 2015; Lécuyer et al., 2007; Tomancak et al., 2007). We are confident that the analysis of regularly studied as well as less explored *Drosophila* tissues will be stimulated by the fTRG collection.

When protein expression levels are sufficiently high, the fusion proteins can be visualised by live imaging approaches in intact animals. It is difficult to estimate the absolute expression levels required for live visualisation, as this depends on the imaging conditions, the accessibility and transparency of the tissue and importantly on the observed protein pattern. A strongly localised protein can result in a very bright local signal, such as Talin or Ilk at the muscle attachment sites, compared to a protein homogenously distributed throughout the entire cell. In particular for tissues, such as the adult legs, antennae or the adult fat body, which are difficult to dissect and stain without losing tissue integrity, these live markers should be enormously beneficial.

One important limitation for examining the pattern of protein expression is the accessibility of the tissue of interest for imaging. We have shown that light sheet microscopy can be used to image the dynamics of tagged protein expression throughout embryogenesis. We further demonstrated that two-photon microscopy can be applied to study protein dynamics during muscle morphogenesis in developing pupae. Other confocal or light sheet based imaging paradigms could be adapted for *in toto* imaging of living or fixed and cleared specimen from other life cycle stages. Establishing standardized protocols for preparation, staining and imaging of *Drosophila* stages, isolated tissues and organs will be necessary to realise the full potential of the fTRG collection.

Protein interaction data in fly are available from a number of studies (Formstecher et al., 2005; Giot et al., 2003; Guruharsha et al., 2011). These results were generated using yeast-two-hybrid, or overexpression in a tissue culture system, followed by affinity-purification and mass spectrometric analysis. Despite high-throughput, these approaches face the problem that the interacting proteins might not be present at the same place within a cell, or not even co-expressed in a developing organism. This is circumvented by affinity purifications of endogenously expressed proteins, which thus far at genome-scale was only reported from yeast (Gavin et al., 2002; Ho et al., 2002; Krogan et al., 2006). In higher organisms, BAC-based systems, which are closely related to our fosmid approach, elegantly solved these issues, as shown by a recent human interactome study (Hein et al., 2015).

The collection of transgenic flies covers currently only about 10 % of the available tagged fosmids. Expanding the collection to include most genes of the genome and importantly characterising the transgenes-encoded proteins by imaging in various biological contexts is best achieved by spreading the clones and transgenic

675 lines amongst the community of researchers using *Drosophila* as a model system.
676 Therefore, all transgenic lines are available from the VDRC stock collection. Despite
677 the expanding CRISPR-based genome engineering technologies, the fTRG collection
678 will continue to be an important resource for the fly community, in particular, if the
679 full functionality of certain fTRG lines has been demonstrated, as we did here for a
680 selection of important developmental regulators. As with many genome-scale
681 resources it is typically easier to produce them than to fully characterise and exploit
682 their potential. Comprehensive generation of thousands of transgenes and their
683 thorough analysis takes time; it took us 4 years to assemble the collection presented
684 here. Development of protocols and techniques to image these collections of tagged
685 lines and assembling open access databases to share the data needs to continue and
686 will eventually become useful also for the characterisation of resources whose
687 production began only recently.

688 Wangler, Yamamoto and Bellen convincingly argued that the *Drosophila*
689 system remains an indispensable model for translational research because many
690 essential fly genes are homologs of Mendelian disease genes in humans (Wangler et
691 al., 2015). Yet, even after decades of research on fruit flies only about 2000 of the
692 estimated 5000 lethal mutations have been investigated. Resources like ours will
693 therefore provide essential functional information about gene expression and
694 localisation in *Drosophila* tissues that can serve as a starting point for the mechanistic
695 understanding of human pathologies and their eventual cures.

Materials and Methods

TransgeneOme clone engineering

Fosmids were engineered as described previously (Ejsmont et al., 2009; 2011), except for the inclusion of the 'pre-tagging' step in the genome-wide TransgeneOme set. All tagging cassettes were generated from synthetic DNA and cloned into R6K carrying plasmids, which require the presence of the *pir* gene product for replication (Metcalf et al., 1996). The *pir* gene is not present in the FlyFos library host strain, thereby ensuring near-complete lack of background resistance in the absence of the correct homologous recombination event.

Details of the recombineering steps are as follows (**Figure 2A**): Step 1. The *E. coli* cells containing a FlyFos clone covering the gene locus of interest are transformed with the pRedFlp plasmid, containing the genes necessary for the homologous recombination and the Flp recombinase under independently inducible promoters. Step 2. Next, a 'pre-tagging' cassette carrying an antibiotic resistance gene (NatR, nourseothricin resistance) surrounded by regions of homology to all specific tagging cassettes (**Figure 2 Supplement 1**) and flanked by gene specific homology arms is electroporated as linear DNA fragment produced by PCR. By combination of induced (L-rhamnose) pRedFlp homologous recombination enzyme action and strong selection with a cocktail of three antibiotics (one to maintain the fosmid (chloramphenicol, Cm), one to maintain the pRedFlp (hygromycin, Hgr) and nourseothricin (Ntc) to select for the inserted fragment) the electroporated linear 'pre-tagging' fragment becomes inserted in front of the STOP codon of the gene of interest. Step 3. The 'pre-tagging' cassette is exchanged for a cassette of the chosen tag coding sequence including an FRT flanked selection / counter selection marker (rpsL-neo). This cassette is now universally targeting the homologous sequences

721 shared by the tagging and pre-tagging cassettes and is produced in bulk by restriction
722 enzyme mediated excision from a plasmid. Note that in this way, no PCR induced
723 mutations can be introduced at this step. Step 4. Upon Flp induction (with
724 anhydrotetracycline) the rpsL-neo cassette is excised, leaving a single FRT site,
725 positioned in frame with the tag coding sequence. Step 5. Finally, the recombineering
726 plasmid is removed from the cells containing the engineered fosmids by inhibition of
727 its temperature sensitive origin of replication and release from Hgr selection. The
728 cells are plated on a selective chloramphenicol agar plate from which a single colony
729 is picked and further validated.
730
731 **NGS-based validation of the TransgeneOme clones**
732 For NGS-based resource validation single colonies for each TransgeneOme clone
733 were picked into 96-well plates and individual wells of all 96-well plates were pooled
734 into 8 rows and 12 columns pools. Fosmid DNA was isolated from these pools, and
735 mate pair fragment libraries were prepared and sequenced on an Illumina platform.
736 First, adapters and low quality sequences were trimmed with Trimmomatic0.32.
737 (parameters: ILLUMINACLIP:NexteraPE-PE.fa:2:30:10 LEADING:3 TRAILING:3
738 SLIDINGWINDOW:4:15 MINLEN:36). Second, in order to detect un-flipped
739 fosmid sequences (where the FLP-mediated excision of selection cassette failed), the
740 read pairs were mapped with Bowtie2 (Langmead and Salzberg, 2012) against the un-
741 flipped tag sequence and the genome. If any read of the mate of the pair mapped to
742 the un-flipped sequence while the second mate mapped to the genome consistent with
743 the estimated mate pair insert size of 3000 bp ± 1000 bp, the fosmid was flagged a un-
744 flipped and was not further analysed. Third, in order to identify mutations in the tag
745 and in the immediate genomic surrounding (± 1000 bp), the NGS reads were mapped

against the fosmid references that included the flipped tag. The Bowtie2 was set to report only hits where both reads of the pairs map concordantly to the insert size in the tag and in the genome (parameters: -I 2200 -X 3700 --rf --no-discordant --no-unal --no-mixed). This mapping was further filtered by deleting PCR duplicated read pairs with samtools1.1 *rmdup* (Li et al., 2009). Mutations were identified by utilising SNP calling implemented in FreeBayes (Garrison and Marth, 2012) using the standard filters and *vcffilter* to eliminate reported SNPs with scores < 20. Finally, in the last step, the information of the row and column pools were compared and summarized using a custom C-program that read the results of the SNP calling and the Bowtie mappings and counted the coverage for each read pair anchored in tag sequence with at least 20 bp. To correct for random PCR or sequencing errors the reported SNPs were compared for the row and column pools of each fosmid and SNPs occurring in both pools with coverage of 3 or more reads were considered as real.

Drosophila stocks and genetic rescue experiments

Fly stocks were maintained using standard culture conditions. All crosses were grown at 25°C unless otherwise noted. Most of the fly mutant or deficiency strains for the rescue experiments were obtained from the Bloomington *Drosophila* Stock Center and if located on X- or 2nd chromosome crossed together with the respective fTRG line. If the mutant gene was located on the 3rd chromosome it was recombined with the fTRG line. Rescue was generally tested in trans-heterozygotes as indicated in **Supplemental Table 3**. The rescue for 6 genes (*bam, fat, mask, rap, RhoGEF2* and *yki*) was done by others, who communicated or published the results (**Supplementary Table 3**). For rescue of flightlessness a standard flight test was used (Schnorrer et al., 2010).

771

772 **Generation of transgenic FlyFos (fTRG) lines**

773 Most TransgeneOme fosmid clones were injected into the y[1], w[*], P{nos-
774 phiC31\int.NLS}X; PBac{y+-attP-3B}VK00033 (BL-32542). This stock has white
775 eyes and no fluorescent eye markers, which would interfere with screening for the red
776 fluorescent eye marker used in the FlyFos clones (Ejsmont et al., 2009). A few fosmid
777 clones were also injected into y[1], w[*], P{nos-phiC31\int.NLS}X; PBac{y+-attP-
778 3B}VK00002, with the attP site located on the 2nd chromosome. The *osk-GFP* fosmid
779 was injected into attP40.

780

781 *Detailed injection protocol*

782 A) Bacterial culture of fosmid clones: 1. Inoculate 2 ml LB-medium plus
783 chloramphenicol (Cm 25 µg / ml) with fosmid clone and grow over night at 37 °C. 2.
784 Dilute to 10 ml (9 ml LB-medium + Cm and 1 ml bacterial culture) and add 10 µl 10
785 % arabinose (final concentration 0.01 %) to induce the fosmid to high copy number.
786 3. Grow at 37 °C for 5 h and collect the pellet by 10 min centrifugation at 6000 rpm.
787 Pellet can be stored at -20 ºC.

788 B) Preparation of fosmid DNA: Use the HiPure Plasmid Miniprep Kit from
789 Invitrogen (order number: K2100-03) according to the supplied protocol
790 (MAN0003643) with following modifications: before starting: pre-warm the elution
791 buffer (E4) to 50 °C; step 4: incubate the lysate for 4 min at room temperature; step 5:
792 incubate 4 min on ice before centrifuging at 4 °C for 10 min; step 8: add 850 µl
793 elution buffer (pre-warmed to 50 °C) to the column; step 9: add 595 µl isopropanol to
794 the elution tube, centrifuge 20 min at 4 °C; wash pellet with 800 µl 70 % ethanol;
795 centrifuge for 2 min; step 12: air dry the pellet for 4 min. Add 20 µl EB-buffer

(Qiagen) to the pellet and leave at 4 °C overnight to dissolve without pipetting to avoid shearing of the DNA. Do not freeze the DNA. Adjust the concentration to 250 ng / µl and centrifuge 5 min at full speed before injections. Do not inject DNA older than one week.

C) Embryo injections: Collect young embryos (0 - 30 min) on an agar plate, bleach away the chorion, wash and collect the embryos on a cellulose filter (Whatman 10409814). Align the embryos, transfer them to a glued slide and dry them with silica gel for 10 - 15 min (Roth T199.2). Cover the embryos with Voltalef 10S oil (Lehmann & Voss) and inject the prepared fosmid DNA using a FemtoJet set-up (Eppendorf 5247). The injected DNA should be visible within the embryo. Incubate the injected embryos for 48 h at 18 ºC in a wet chamber and collect the hatched larvae with a brush. Cross the surviving mosaic adults individually to *y, w* males or virgins.

Immuno-stainings

Ovaries: sGFP-protein detection in egg-chambers was done as previously described (Dunst et al., 2015). Detection of the *oskar-GFP* mRNA was performed with a *gfp*-antisense probe (Jambor et al., 2014) and co-staining of *osk* mRNA and Osk protein was done as previously described (Jambor et al., 2011) using a *gfp*-antisense probe and a rabbit anti-GFP antibody (1:1000, ThermoFisher).

Adult thoraces: Antibody stainings of adult thoraces, including flight, leg and visceral muscles, were done essentially as described for adult IFMs (Weitkunat and Schnorrer, 2014). Briefly, thoraces from young adult males were fixed for 15 min in relaxing solution (20 mM phosphate buffer, pH 7.0; 5 mM $MgCl_2$; 5 mM EGTA, 5 mM ATP, 4 % PFA) + 0.5 % Triton X-100, cut sagittally with a sharp microtome blade and blocked for 1 hour at room temperature with 3 % normal goat serum in PBS-0.5%

Triton X-100. Samples were stained with primary antibodies overnight at 4 ºC (rabbit anti-GFP 1:2000 Amsbio; mouse anti-Futsch 1:100, Hybridoma Bank) washed and incubated with secondary antibodies coupled to Alexa dyes and rhodamine-phalloidin or phalloidin-Alexa-660 (all from Molecular Probes). After washing, samples were mounted in Vectashield containing DAPI. Images were acquired with a Zeiss LSM 780 confocal microscope and processed with Fiji (Schindelin et al., 2012) and Photoshop (Adobe).

Live imaging

SPIM imaging of embryos: De-chorionated embryos of the appropriate age were embedded in 1 % low melting point agarose and mounted into a glass capillary. Fluorescent microspheres (FY050 Estapor microspheres, Merck Millipore; 1:4000) were included in the embedding medium for multi-view registration. The embryos were imaged using the Zeiss Lightsheet Z.1 with a Zeiss 20x/1.0 water-immersion Plan Apochromat objective lens with 0.8x zoom at 25 °C using 488 nm laser set at 4 mW. Five views were imaged using dual-sided illumination with Zeiss 10x/0.2 illumination lenses. A mean fusion was applied to fuse both illumination sides after acquisition using the ZEN software (Zeiss). The views were acquired at 72° angles with a stack size of 130 µm and a step size of 1.5 µm. Exposure time were 30 ms per slice. Each slice consists of 1920 x 1200 pixels with a pixel size of 0.29 µm and a bit depth of 16 bits. The light sheet thickness was 4 µm at the center of the field of view. The embryos were imaged from onset of expression of the fosmids transgenes (determined empirically) until late embryogenesis with a time resolution of 15 min. Multi-view processing of the dataset was carried out using the Fiji plugin for multi-view reconstruction (Preibisch et al., 2009; Schmied et al., 2014), which was executed

on a high performance computing cluster (Schmied et al., 2015). The multi-view reconstruction was followed by multi-view deconvolution (Preibisch et al., 2014), for which the images were down sampled by a factor of two. Movies were extracted via the Fiji plugin BigDataViewer (Pietzsch et al., 2015).

The Gsb-GFP fTRG line was crossed with the H2Av-mRFPruby line (Fischer et al., 2004; Preibisch et al., 2014), the embryos of this cross were imaged using a 40x/1.0 water immersion Plan Apochromat lens from Zeiss with 1x zoom at 25 °C at 17.5 mW of the 488 nm laser and 4 mW of the 561 nm laser. A single angle with dual sided illumination was imaged. The stack size was 82.15 µm with a step size of 0.53 µm. Exposure time was 30 ms per slice. Each slice consisted of 1920 x 1920 pixels with a pixel size of 120 nm and a bit depth of 16-bit. The light sheet thickness was 3.21 µm at the center of the field of view. The embryos were imaged from early blastoderm onwards until late embryogenesis focusing on the head with a time resolution of 7 min.

Imaging of pupae: Staging and live imaging of the pupae were performed at 27 ºC. Live imaging of pupae at the appropriate stage was done as described previously (Weitkunat and Schnorrer, 2014). Briefly, the staged pupa was cleaned with a brush and a small observation window was cut into the pupal case with sharp forceps. The pupa was mounted on a custom-made slide and the opening was covered with a small drop of 50 % glycerol and a cover slip. Z-stacks of either single time points or long-term time lapse movies were acquired using either a spinning disc confocal microscope (Zeiss, Visitron) or a two-photon microscope (LaVision), both equipped with heated stages.

Proteomics

Per sample about hundred pupae or adult flies were snap-frozen in liquid nitrogen and ground to a powder. The powder was re-suspended and further processed as described in the quantitative BAC-GFP interactomics protocol (Hubner et al., 2010). In brief, 800 µl of lysate per sample were cleared by centrifugation. The cleared lysate was mixed with magnetic beads pre-coupled to anti-GFP antibodies and run over magnetic micro-columns (both Miltenyi Biotec). Columns were washed, and samples subjected to in-column tryptic digestion for 30 min. Eluates were collected and digestion continued overnight, followed by desalting and storage on StageTips. Eluted peptides were analysed with an Orbitrap mass spectrometer (Thermo Fisher). Raw data were analysed in MaxQuant version 1.4.3.22 (Cox and Mann, 2008) using the MaxLFQ algorithm for label-free quantification (Cox et al., 2014). Interacting proteins were identified by the similarity of their intensity profiles to the respective baits (Keilhauer et al., 2014). Heat maps were plotted in the Perseus module of the MaxQuant software suite.

Author contributions

D.S., S.H., E.V. and M.S performed the liquid culture recombineering experiments.

M.S., S.J., P.K. and S.S. analysed the NGS data and contributed to the TransgeneOme database.

B. S., N.P., K.F. and F.S performed the transgenesis of about half the fTRG lines.

V.K.J.V., R.T.K., K.A., M.R. and K.V performed the transgenesis of the other half of the fTRG lines.

C.B. performed most of the genetic rescue experiments.

H.J. performed and analysed all experiments in ovaries.

C.S. collected and analysed SPIM in toto images of embryos.

C.B. and F.S. performed and analysed all imaging data in pupae and adults.

M.Y.H and M.M. performed the proteomics analysis and analysed the data.

V.H. contributed analysis of Gsb-GFP cell tracking.

R.K.E. constructed the tagging cassettes.

I.R.S.F. performed initial proof of principle tagging experiments.

P.T., F.S., M.S. and E.K. initiated the collaborative project and obtained dedicated funding.

P.T. and F.S. conceived and coordinated the study and co-wrote the manuscript with input from the other authors.

Acknowledgements

Stocks obtained from the Bloomington *Drosophila* Stock Center (NIH P40OD018537) were used in this study. We thank Anne Ephrussi for sharing of fly lines. We thank Franziska Friedrich for drawing the ovariole scheme used for Figure 4A. We thank Andreas Dahl from the Deep Sequencing Group at CRTD/BIOTECH, Dresden for the NGS library preparation and sequencing. We also thank light microscopy facility and the computer department for assistance with imaging and data processing. We are grateful to Sandra Lemke, Aynur Kaya-Copur and Xu Zhang for help with the genetic rescue experiments and to Caroline Sonsteby for assistance during some of the pupal imaging experiments. We thank the entire Schnorrer lab for helpful comments on this manuscript. We are particularly grateful to Reinhard Fässler and Herbert Jäckle for continuous support of this work. This work was funded by the EMBO Young Investigator Program (F. S.) the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant 310939 (F. S.), the European Research Council under the European Union's Seventh

Framework Programme (FP/2007-2013)/ERC Grant 260746 (P.T.), European Commission GENCODYS (P.T. and H.J.), Human Frontier Science Program (HFSP) RGY0093/2012 (P.T. and C. S.) and the Max Planck Society (P.T., F.S., M.S., M.M. and E.K.).

Competing interests

The authors declare that no competing interests exist.

References

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. (2000). The genome sequence of Drosophila melanogaster. Science *287*, 2185–2195.

Baena-Lopez, L.A., Alexandre, C., Mitchell, A., Pasakarnis, L., and Vincent, J.P. (2013). Accelerated homologous recombination and subsequent genome modification in Drosophila. Development *140*, 4818–4825.

Bard, F., Casano, L., Mallabiabarrena, A., Wallace, E., Saito, K., Kitayama, H., Guizzunti, G., Hu, Y., Wendler, F., DasGupta, R., et al. (2006). Functional genomics reveals genes involved in protein secretion and Golgi organization. Nature *439*, 604–607.

Bischof, J., Bjorklund, M., Furger, E., Schertel, C., Taipale, J., and Basler, K. (2013). A versatile platform for creating a comprehensive UAS-ORFeome library in Drosophila. Development *140*, 2434–2442.

Bolatto, C., Chifflet, S., Megighian, A., and Cantera, R. (2003). Synaptic activity modifies the levels of Dorsal and Cactus at the neuromuscular junction of Drosophila. J. Neurobiol. *54*, 525–536.

Bonnay, F., and Cohen-Berros, E. (2013). big bang gene modulates gut immune tolerance in Drosophila.

Bulgakova, N.A., and Knust, E. (2009). The Crumbs complex: from epithelial-cell polarity to retinal degeneration. Journal of Cell Science *122*, 2587–2596.

Buszczak, M., Paterno, S., Lighthouse, D., Bachman, J., Planck, J., Owen, S., Skora, A.D., Nystul, T.G., Ohlstein, B., Allen, A., et al. (2007). The carnegie protein trap library: a versatile tool for Drosophila developmental studies. Genetics *175*, 1505–1531.

Clark, K.A., Bland, J.M., and Beckerle, M.C. (2007). The Drosophila muscle LIM

protein, Mlp84B, cooperates with D-titin to maintain muscle structural integrity. Journal of Cell Science *120*, 2066–2077.

Collins, K.A., Unruh, J.R., Slaughter, B.D., and Yu, Z. (2014). Corolla is a novel protein that contributes to the architecture of the synaptonemal complex of Drosophila. ….

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology *26*, 1367–1372.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell Proteomics *13*, 2513–2526.

de Celis, J. (1999). Regulation of the spalt/spalt-related gene complex and its function during sensory organ development in the Drosophila thorax. Development *126*, 2653.

Deal, R.B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. Developmental Cell *18*, 1030–1040.

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.

Dougherty, G.W., Chopp, T., Qi, S.-M., and Cutler, M.L. (2005). The Ras suppressor Rsu-1 binds to the LIM 5 domain of the adaptor protein PINCH1 and participates in adhesion-related functions. Experimental Cell Research *306*, 168–179.

Dunst, S., Kazimiers, T., Zadow, von, F., Jambor, H., Sagner, A., Brankatschk, B., Mahmoud, A., Spannl, S., Tomancak, P., Eaton, S., et al. (2015). Endogenously tagged rab proteins: a resource to study membrane trafficking in Drosophila. Developmental Cell *33*, 351–365.

Ejsmont, R.K., Sarov, M., Winkler, S., Lipinski, K.A., and Tomancak, P. (2009). A toolkit for high-throughput, cross-species gene engineering in Drosophila. Nature Methods *6*, 435–437.

Ejsmont, R.K., Ahlfeld, P., Pozniakovsky, A., Stewart, A.F., Tomancak, P., and Sarov, M. (2011). Recombination-mediated genetic engineering of large genomic DNA transgenes. Methods Mol. Biol. *772*, 445–458.

Ephrussi, A., Dickinson, L.K., and Lehmann, R. (1991). Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. Cell *66*, 37–50.

Fernandes, J., Bate, M., and VijayRaghavan, K. (1991). Development of the indirect flight muscles of Drosophila. Development *113*, 67–77.

Fernandes, J.J., Atreya, K.B., Desai, K.M., Hall, R.E., Patel, M.D., Desai, A.A., Benham, A.E., Mable, J.L., and Straessle, J.L. (2005). A dominant negative form of

994 Rac1 affects myogenesis of adult thoracic muscles in Drosophila. Developmental
995 Biology *285*, 11–27.

996 Fessler, L.I., Campbell, A.G., Duncan, K.G., and Fessler, J.H. (1987). Drosophila
997 laminin: characterization and localization. Journal of Cell Biology *105*, 2383–2391.

998 Fischer, M., Haase, I., Simmeth, E., Gerisch, G., and Müller-Taubenberger, A. (2004).
999 A brilliant monomeric red fluorescent protein to visualize cytoskeleton dynamics in
1000 Dictyostelium. FEBS Letters *577*, 227–232.

1001 Formstecher, E., Aresta, S., Collura, V., Hamburger, A., Meil, A., Trehin, A.,
1002 Reverdy, C., Betin, V., Maire, S., Brun, C., et al. (2005). Protein interaction mapping:
1003 a Drosophila case study. Genome Research *15*, 376–384.

1004 Forrest, K.M., and Gavis, E.R. (2003). Live imaging of endogenous RNA reveals a
1005 diffusion and entrapment mechanism for nanos mRNA localization in Drosophila.
1006 Current Biology *13*, 1159–1168.

1007 Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-
1008 read sequencing

1009 . 1–9.

1010 Gavin, A.-C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J.,
1011 Rick, J.M., Michon, A.-M., Cruciat, C.-M., et al. (2002). Functional organization of
1012 the yeast proteome by systematic analysis of protein complexes. Nature *415*, 141–147.

1013 Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi,
1014 C.E., Godwin, B., Vitols, E., et al. (2003). A protein interaction map of Drosophila
1015 melanogaster. Science *302*, 1727–1736.

1016 Gratz, S.J., Ukken, F.P., Rubinstein, C.D., Thiede, G., Donohue, L.K., Cummings,
1017 A.M., and O'Connor-Giles, K.M. (2014). Highly specific and efficient CRISPR/Cas9-
1018 catalyzed homology-directed repair in Drosophila. Genetics *196*, 961–971.

1019 Graveley, B., Brooks, A., Carlson, J., and Duff, M. (2010). The developmental
1020 transcriptome of Drosophila melanogaster. Nature.

1021 Guruharsha, K.G., Rual, J.-F., Zhai, B., Mintseris, J., Vaidya, P., Vaidya, N.,
1022 Beekman, C., Wong, C., Rhee, D.Y., Cenaj, O., et al. (2011). A Protein Complex
1023 Network of Drosophila melanogaster. Cell *147*, 690–703.

1024 Gutjahr, T., Patel, N.H., Li, X., Goodman, C.S., and Noll, M. (1993). Analysis of the
1025 gooseberry locus in Drosophila embryos: gooseberry determines the cuticular pattern
1026 and activates gooseberry neuro. Development *118*, 21–31.

1027 Haecker, A., Bergman, M., Neupert, C., Moussian, B., Luschnig, S., Aebi, M., and
1028 Mannervik, M. (2008). Wollknauel is required for embryo patterning and encodes the
1029 Drosophila ALG5 UDP-glucose:dolichyl-phosphate glucosyltransferase.
1030 Development *135*, 1745–1749.

1031 Hammonds, A.S., Bristow, C.A., Fisher, W.W., Weiszmann, R., Wu, S., Hartenstein,

V., Kellis, M., Yu, B., Frise, E., and Celniker, S.E. (2013). Spatial expression of transcription factors in Drosophila embryonic organ development. Genome Biology *14*, R140.

Hartenstein, V., Younossi-Hartenstein, A., Lovick, J.K., Kong, A., Omoto, J.J., Ngo, K.T., and Viktorin, G. (2015). Lineage-associated tracts defining the anatomy of the Drosophila first instar larval brain. Developmental Biology 1–26.

He, H., and Noll, M. (2013). Differential and redundant functions of gooseberry and gooseberry neuro in the central nervous system and segmentation of the Drosophila embryo. Developmental Biology *382*, 209–223.

Hein, M., Hubner, N., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I., Weisswange, I., Mansfeld, J., Buchholz, F., et al. (2015). A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell *in press*.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., and Moore, L. (2002). Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. Nature *415*, 180–183.

Hubner, N.C., Bird, A.W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010). Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. The Journal of Cell Biology *189*, 739–754.

Huisken, J., and Stainier, D.Y.R. (2007). Even fluorescence excitation by multidirectional selective plane illumination microscopy (mSPIM). Opt Lett *32*, 2608–2610.

Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E.H.K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science *305*, 1007–1009.

Izumi, Y., Yanagihashi, Y., and Furuse, M. (2012). A novel protein complex, Mesh-Ssk, is required for septate junction formation in the Drosophila midgut. Journal of Cell Science *125*, 4923–4933.

Jambor, H., Brunel, C., and Ephrussi, A. (2011). Dimerization of oskar 3' UTRs promotes hitchhiking for RNA localization in the Drosophila oocyte. Rna *17*, 2049–2057.

Jambor, H., Mueller, S., Bullock, S.L., and Ephrussi, A. (2014). A stem-loop structure directs oskar mRNA to microtubule minus ends. Rna *20*, 429–439.

Jambor, H., Surendranath, V., Kalinka, A.T., Mejstrik, P., Saalfeld, S., and Tomancak, P. (2015). Systematic imaging reveals features and changing localization of mRNAs in Drosophila development. eLife *4*, e05003.

Jenny, A., Hachet, O., Závorszky, P., Cyrklaff, A., Weston, M.D.J., Johnston, D.S., Erdélyi, M., and Ephrussi, A. (2006). A translation-independent role of oskar RNA in early Drosophila oogenesis. Development *133*, 2827–2833.

Kadrmas, J.L., Smith, M.A., Clark, K.A., Pronovost, S.M., Muster, N., Yates, J.R., and Beckerle, M.C. (2004). The integrin effector PINCH regulates JNK activity and epithelial migration in concert with Ras suppressor 1. Journal of Cell Biology *167*, 1019–1024.

Katzemich, A., Kreisköther, N., Alexandrovich, A., Elliott, C., Schöck, F., Leonard, K., Sparrow, J., and Bullard, B. (2012). The function of the M-line protein obscurin in controlling the symmetry of the sarcomere in the flight muscle of Drosophila. Journal of Cell Science *125*, 3367–3379.

Keilhauer, E.C., Hein, M.Y., and Mann, M. (2014). Accurate Protein Complex Retrieval by Affinity Enrichment Mass Spectrometry (AE-MS) Rather than Affinity Purification Mass Spectrometry (AP-MS). Molecular & Cellular Proteomics *14*, 120–135.

Kim-Ha, J., Kerr, K., and Macdonald, P.M. (1995). Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. Cell *81*, 403–412.

Kim-Ha, J., Smith, J.L., and Macdonald, P.M. (1991). oskar mRNA is localized to the posterior pole of the Drosophila oocyte. Cell *66*, 23–35.

Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637–643.

Kvon, E.Z., Kazmar, T., Stampfel, G., Yáñez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., and Stark, A. (2014). Genome-scale functional characterization of Drosophila developmental enhancers in vivo. Nature *512*, 91–95.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*, 357–359.

Leiss, D., Hinz, U., Gasch, A., Mertz, R., and Renkawitz-Pohl, R. (1988). Beta 3 tubulin expression characterizes the differentiating mesodermal germ layer during Drosophila embryogenesis. Development *104*, 525–531.

Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell *131*, 174–187.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lovick, J.K., Ngo, K.T., Omoto, J.J., Wong, D.C., Nguyen, J.D., and Hartenstein, V. (2013). Postembryonic lineages of the Drosophila brain_ I. Development of the lineage-associated fiber tracts. Developmental Biology *384*, 228–257.

Lowe, N., Rees, J.S., Roote, J., Ryder, E., Armean, I.M., Johnson, G., Drummond, E., Spriggs, H., Drummond, J., Magbanua, J.P., et al. (2014). Analysis of the expression patterns, subcellular localisations and interaction partners of Drosophila proteins

using a pigP protein trap library. Development *141*, 3994–4005.

Meacock, P.A., and Cohen, S.N. (1980). Partitioning of bacterial plasmids during cell division: a cis-acting locus that accomplishes stable plasmid inheritance. Cell *20*, 529–542.

Metcalf, W.W., Jiang, W., Daniels, L.L., Kim, S.K., Haldimann, A., and Wanner, B.L. (1996). Conditionally replicative and conjugative plasmids carrying lacZ alpha for cloning, mutagenesis, and allele replacement in bacteria. Plasmid *35*, 1–13.

Morin, X., Daneman, R., Zavortink, M., and Chia, W. (2001). A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in Drosophila. Proceedings of the National Academy of Sciences of the United States of America *98*, 15050–15055.

Nagarkar-Jaiswal, S., Lee, P.-T., Campbell, M.E., Chen, K., Anguiano-Zarate, S., Gutierrez, M.C., Busby, T., Lin, W.-W., He, Y., Schulze, K.L., et al. (2015). A library of MiMICs allows tagging of genes and reversible, spatial and temporal knockdown of proteins in Drosophila. eLife *4*.

Neuman-Silberberg, F.S., and Schüpbach, T. (1993). The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGF alpha-like protein. Cell *75*, 165–174.

Ni, J.-Q., Zhou, R., Czech, B., Liu, L.-P., Holderbaum, L., Yang-Zhou, D., Shim, H.-S., Tao, R., Handler, D., Karpowicz, P., et al. (2011). A genome-scale shRNA resource for transgenic RNAi in Drosophila. Nature Methods *8*, 405–407.

Nongthomba, U., Pasalodos-Sanchez, S., Clark, S., Clayton, J., and Sparrow, J. (2001). Expression and function of the Drosophila ACT88F actin isoform is not restricted to the indirect flight muscles. J Muscle Res Cell Motil *22*, 111–119.

Oas, S.T., Bryantsev, A.L., and Cripps, R.M. (2014). Arrest is a regulator of fiber-specific alternative splicing in the indirect flight muscles of Drosophila. The Journal of Cell Biology *206*, 895–908.

Pereanu, W. (2006). Neural Lineages of the Drosophila Brain: A Three-Dimensional Digital Atlas of the Pattern of Lineage Location and Projection at the Late Larval Stage. Journal of Neuroscience *26*, 5534–5553.

Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2005). Engineering and characterization of a superfolder green fluorescent protein. Nature Biotechnology *24*, 79–88.

Pietzsch, T., Saalfeld, S., Preibisch, S., and Tomancak, P. (2015). BigDataViewer: visualization and processing for large image data sets. Nature Methods *12*, 481–483.

Port, F., Chen, H.-M., Lee, T., and Bullock, S.L. (2014). Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. Proceedings of the National Academy of Sciences *111*, E2967–E2976.

Preibisch, S., Amat, F., Stamataki, E., Sarov, M., Singer, R.H., Myers, E., and

1151 Tomancak, P. (2014). Efficient Bayesian-based multiview deconvolution. Nature
1152 Methods *11*, 645–648.

1153 Preibisch, S., Saalfeld, S., and Tomancak, P. (2009). Globally optimal stitching of
1154 tiled 3D microscopic image acquisitions. Bioinformatics *25*, 1463–1465.

1155 Quiñones-Coello, A.T., Petrella, L.N., Ayers, K., Melillo, A., Mazzalupo, S., Hudson,
1156 A.M., Wang, S., Castiblanco, C., Buszczak, M., Hoskins, R.A., et al. (2007).
1157 Exploring strategies for protein trapping in Drosophila. Genetics *175*, 1089–1104.

1158 Razzaq, A., Robinson, I., McMahon, H., Skepper, J., Su, Y., Zelhof, A., Jackson, A.,
1159 Gay, N., and O'Kane, C. (2001). Amphiphysin is necessary for organization of the
1160 excitation–contraction coupling machinery of muscles, but not for synaptic vesicle
1161 endocytosis in Drosophila. Genes & Development *15*, 2967.

1162 Rørth, P. (1998). Gal4 in the Drosophila female germline. Mechanisms of
1163 Development *78*, 113–118.

1164 Rørth, P., Szabo, K., Bailey, A., Laverty, T., Rehm, J., Rubin, G.M., Weigmann, K.,
1165 Milán, M., Benes, V., Ansorge, W., et al. (1998). Systematic gain-of-function
1166 genetics in Drosophila. Development *125*, 1049–1057.

1167 Sarov, M., Murray, J.I., Schanze, K., Pozniakovski, A., Niu, W., Angermann, K.,
1168 Hasse, S., Rupprecht, M., Vinis, E., Tinney, M., et al. (2012). A genome-scale
1169 resource for in vivo tag-based protein function exploration in C. elegans. Cell *150*,
1170 855–866.

1171 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T.,
1172 Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source
1173 platform for biological-image analysis. Nature Methods *9*, 676–682.

1174 Schmied, C., Stamataki, E., and Tomancak, P. (2014). Open-source solutions for
1175 SPIMage processing. Methods Cell Biol. *123*, 505–529.

1176 Schmied, C., Steinbach, P., Pietzsch, T., Preibisch, S., and Tomancak, P. (2015). An
1177 automated workflow for parallel processing of large multiview SPIM recordings.
1178 *arXiv:1507.08575*, 1–13.

1179 Schnorrer, F., Kalchhauser, I., and Dickson, B. (2007). The transmembrane protein
1180 Kon-tiki couples to Dgrip to mediate myotube targeting in Drosophila.
1181 Developmental Cell *12*, 751–766.

1182 Schnorrer, F., Schönbauer, C., Langer, C.C.H., Dietzl, G., Novatchkova, M.,
1183 Schernhuber, K., Fellner, M., Azaryan, A., Radolf, M., Stark, A., et al. (2010).
1184 Systematic genetic analysis of muscle morphogenesis and function in Drosophila.
1185 Nature *464*, 287–291.

1186 Schönbauer, C., Distler, J., Jährling, N., Radolf, M., Dodt, H.-U., Frasch, M., and
1187 Schnorrer, F. (2011). Spalt mediates an evolutionarily conserved switch to fibrillar
1188 muscle fate in insects. Nature *479*, 406–409.

1189 Sokolovski, M., Bhattacherjee, A., Kessler, N., Levy, Y., and Horovitz, A. (2015).

1190 Thermodynamic Protein Destabilization by GFP Tagging: A Case of Interdomain
1191 Allostery. Biophysical Journal *109*, 1157–1162.

1192 Spletter, M.L., and Schnorrer, F. (2014). Transcriptional regulation and alternative
1193 splicing cooperate in muscle fiber-type specification in flies and mammals.
1194 Experimental Cell Research *321*, 90–98.

1195 Spletter, M.L., Barz, C., Yeroslaviz, A., Schönbauer, C., Ferreira, I.R.S., Sarov, M.,
1196 Gerlach, D., Stark, A., Habermann, B.H., and Schnorrer, F. (2015). The RNA-binding
1197 protein Arrest (Bruno) regulates alternative splicing to enable myofibril maturation in
1198 Drosophila flight muscle. EMBO Rep *16*, 178–191.

1199 St Johnston, D. (2012). Using mutants, knockdowns, and transgenesis to investigate
1200 gene function in Drosophila. WIREs Dev Biol *2*, 587–613.

1201 Sun, B., Xu, P., and Salvaterra, P.M. (1999). Dynamic visualization of nervous
1202 system in live Drosophila. Proceedings of the National Academy of Sciences of the
1203 United States of America *96*, 10438–10443.

1204 Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S.E., Richards,
1205 S., Ashburner, M., Hartenstein, V., Celniker, S.E., et al. (2002). Systematic
1206 determination of patterns of gene expression during Drosophila embryogenesis.
1207 Genome Biology *3*, RESEARCH0088.

1208 Tomancak, P., Berman, B.P., Beaton, A., Weiszmann, R., Kwan, E., Hartenstein, V.,
1209 Celniker, S.E., and Rubin, G.M. (2007). Global analysis of patterns of gene
1210 expression during Drosophila embryogenesis. Genome Biology *8*, R145.

1211 Tracey, W.D., Wilson, R.I., Laurent, G., and Benzer, S. (2003). painless, a Drosophila
1212 gene essential for nociception. Cell *113*, 261–273.

1213 Trcek, T., Grosch, M., York, A., Shroff, H., Lionnet, T., and Lehmann, R. (2015).
1214 Drosophila germ granules are structured and contain homotypic mRNA clusters.
1215 Nature Communications *6*, 7962.

1216 Tseng, A.-S.K., and Hariharan, I.K. (2002). An overexpression screen in Drosophila
1217 for genes that restrict growth or cell-cycle progression in the developing eye. Genetics
1218 *162*, 229–243.

1219 Tu, Y., Huang, Y., Zhang, Y., Hua, Y., and Wu, C. (2001). A new focal adhesion
1220 protein that interacts with integrin-linked kinase and regulates cell adhesion and
1221 spreading. Journal of Cell Biology *153*, 585–598.

1222 Urbach, R., and Technau, G.M. (2003). Molecular markers for identified neuroblasts
1223 in the developing brain of Drosophila. Development *130*, 3621–3637.

1224 Vaskova, M., Bentley, A.M., Marshall, S., Reid, P., Thummel, C.S., and Andres, A.J.
1225 (2000). Genetic analysis of the Drosophila 63F early puff. Characterization of
1226 mutations in E63-1 and maggie, a putative Tom22. Genetics *156*, 229–244.

1227 Venken, K.J.T., and Bellen, H.J. (2012). Genome-Wide Manipulations of Drosophila
1228 melanogaster with Transposons, Flp Recombinase, and ΦC31 Integrase. Methods Mol.

Biol. *859*, 203–228.

Venken, K.J.T., Carlson, J.W., Schulze, K.L., Pan, H., He, Y., Spokony, R., Wan, K.H., Koriabine, M., de Jong, P.J., White, K.P., et al. (2009). Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nature Methods *6*, 431–434.

Venken, K.J.T., 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.T., Schulze, K.L., Haelterman, N.A., Pan, H., He, Y., Evans-Holm, M., Carlson, J.W., Levis, R.W., Spradling, A.C., Hoskins, R.A., et al. (2011). MiMIC: a highly versatile transposon insertion resource for engineering Drosophila melanogaster genes. Nature Methods *8*, 737–743.

Vernes, S.C. (2014). Genome wide identification of fruitless targets suggests a role in upregulating genes important for neural circuit formation. Sci Rep *4*, 4412.

Vigoreaux, J., Saide, J., Valgeirsdottir, K., and Pardue, M. (1993). Flightin, a novel myofibrillar protein of Drosophila stretch-activated muscles. Journal of Cell Biology *121*, 587.

Wangler, M.F., Yamamoto, S., and Bellen, H.J. (2015). Fruit flies in biomedical research. Genetics *199*, 639–653.

Weitkunat, M., and Schnorrer, F. (2014). A guide to study Drosophila muscle biology. Methods *68*, 2–14.

Weitkunat, M., Kaya-Copur, A., Grill, S.W., and Schnorrer, F. (2014). Tension and force-resistant attachment are essential for myofibrillogenesis in Drosophila flight muscle. Curr Biol *24*, 705–716.

Younossi-Hartenstein, A., Salvaterra, P.M., and Hartenstein, V. (2002). Early development of theDrosophila brain: IV. Larval neuropile compartments defined by glial septa. The Journal of Comparative Neurology *455*, 435–450.

Zhang, X., Koolhaas, W.H., and Schnorrer, F. (2014). A versatile two-step CRISPR- and RMCE-based strategy for efficient genome engineering in Drosophila. G3 (Bethesda) *4*, 2409–2418.

Figure 1. Overview of the tagging strategy and its applications

Liquid culture recombineering is used to insert a 'pre-tagging' cassette into FlyFos genomic clones in bacteria. This cassette can then be replaced by a simple, universal, recombineering reaction with any tag of choice, here, a superfolder GFP tag (sGFP) to generate the sGFP TransgeneOme clone library. These clones are transformed into flies generating transgenic FlyFos libraries that can be used for multiple *in vivo* applications. In the future, additional libraries with other tags can be generated easily.

Figure 2. Generation of the TransgeneOme library

(**A**) TransgeneOme resource engineering. The steps of the recombineering pipeline are shown on the left with the success rate of each step indicated on the right (red colour denotes bacterial clones that did not grow). The *E. coli* cells are schematically represented with a dotted circle. With the first two steps the 'pre-tagging' cassette is inserted, which is replaced in the next three steps with the sGFP cassette to generate the sGFP TransgeneOme library. See text and methods section for details. (**B**) Next-generation-sequencing (NGS)-based validation of the sGFP TransgeneOme library. Schematic of the bar coding (BC) strategy of row and column pools is shown to the left and sequencing results to the right. Only the magenta and blue mate pairs contribute to the analysis of tag sequence. Fully sequenced tags within fosmids without point mutations are shown in solid green, clones without mutation in tagging cassette but incomplete coverage in light green and clones with mutation(s) or un-flipped cassette are shown in red. (**C**) Statistics of the mutation distributions with deletions indicated by green, substitutions by red and insertions by blue lines. Note that most mutations reside within the primer sequences used as homology arms for recombineering (HA-1 and HA-2).

Figure 2 - Supplement 1: Tagging cassettes

Tags tested and used in this study. Shown is the form of the tagging cassette after insertion into target site and flip-out of the counter selection sequences (rpsl-neo) leaving behind the FRT sequence colored red.

Figure 3: Functionality of the fTRG lines

(**A**) Genetic rescue statistics of null/strong mutant alleles for 46 selected fTRG lines. Note that more than two-thirds of the lines show a full rescue (see **Supplementary Table 3**). (**B**, **C**) *osk*-GFP mRNA (in yellow) expressed from fTRG1394 fully rescues egg-chamber development of an *osk* null allele (Jenny et al., 2006). *osk*-GFP mRNA enriches in the early oocyte (**B**, stage 6) and rescues the oogenesis arrest and the DNA condensation defect of the *osk* mutant (**B'**, yellow arrow). At stage 10 *osk*-GFP RNA enriches at the posterior pole (**C**) and produces sufficient protein to ensure proper embryogenesis. *osk*-GFP mRNA is shown in yellow, DAPI in magenta; scale bars indicate 30 µm.

Figure 4: Expression of tagged transgenes in ovaries

(**A**) Schematic overview of oogenesis stages and cell types. (**B**) Summary of the identified expression patterns; see also **Supplementary Table 4**. (**C**) Selected examples for cell type specific FlyFos expression patterns at germarium, early- and mid-oogenesis stages visualised by anti-GFP antibody staining. (**D**) Selected examples of subcellular localisation patterns, highlighting nuclear, cortical and cytoplasmic patterns at different oogenesis stages. GFP is show in green, DAPI in magenta; scale bars indicate 30 µm.

Figure 4 - Supplement 1: Posttranscriptional regulation of protein expression during oogenesis

(**A**, **B**) *osk-GFP* mRNA visualised by an anti-GFP labelled RNA probe (yellow, DAPI in magenta) at stage 6 and stage 10 of oogenesis. (**A**', **B**') Osk-GFP protein visualised by anti-GFP antibody (green, DAPI in magenta) at stage 6 and stage 10. Note that Osk-GFP protein is not detectable at stage 6. (**C, D**) *corolla-GFP* mRNA (yellow, DAPI in magenta) at stage 6 and stage 8. (**E, F**) Corolla-GFP protein (green, DAPI in magenta) at stage 6 and stage 8. Note that Corolla-GFP protein in only detectable at stage 6 but not stage 8. Scale bars indicate 30 µm.

Figure 5: Live *in toto* imaging during embryogenesis with SPIM

(**A** - **C**) Nrv2-GFP protein is enriched in cell membrane of the epidermis and the CNS of late stage 16 embryos, as shown by a lateral section (**A**) and high magnifications of the posterior epidermis (**B**) and the ventral CNS (**C**). (**D** - **H**) Still image from a Nrv2-GFP time-lapse movie with lateral section views on the left, ventral sections in the middle and transverse sections on the right. Note that Nrv2-GFP is first expressed in the developing epidermal epithelial cells (**D**, **E**) and then becomes enriched in the CNS (**F** - **H**, see **Supplementary Movie 1**). (**I**) Schemes of the lateral, ventral and transverse section views through the embryo. Scale bars indicate 50 µm.

Figure 5 - Supplement 1: Live Gsb-GFP imaging during embryogenesis with SPIM

(**A**, **B**) Lateral view of a live stage 10 embryo expressing Gsb-GFP (green in A) and Histone2A-mRFPruby (red in A); anterior is to the left (see **Supplementary Movie 2**). (**C** - **F**) Reconstruction of *gsb*-positive deuterocerebral proneural domain (from a

similar movie and position as boxed in B). Fate of cells is symbolized by different colours (blue: epidermal precursor undergoing no further mitosis; purple: epidermal precursor undergoing one mitosis; red: neuroblasts). (C) and (D) show optical section through neurectoderm of stage 10 embryo prior to neuroblast delamination; (E) and (F) 30 minutes later, after neuroblasts have delaminated (stage 11), with a superficial optical section of surface ectoderm (E), and deep section of neuroblast layer (F). (**G - I**) Optical cross sections of similar live embryo as in (A) expressing GFP-tagged Gsb (green) and Histone-2A-mRFPruby (red) at stage 10 (G), early 11 (H) and late 11 (I) showing neuroblasts delaminating from Gsb-GFP domain. **J**: Schematic cross section of stage 10 (left) and stage 11 (right) ectoderm illustrating fate of cells forming part of Gsb-positive pro-neural domain. Scale bars indicate 25µm (A, B) and 10 µm (G - I).

Figure 5 - Supplement 2: Live Gsb-n-GFP imaging during embryogenesis with SPIM

(**A, B**). Ventral view of Gsb-n-GFP expression of a stage 12 embryo during germ-band retraction (A) and stage 14 during head involution (B**)**. Note that Gsb-n-GFP remains expressed in neuronal precursors during stage 14 (**Supplementary Movie 3**). Scale bars indicate 50 µm.

Figure 6: FlyFos expression in tissues of the adult thorax.

Antibody stainings of the adult thorax with anti-GFP antibody (green) and phalloidin (red). (**A - F**) Act88-GFP expression is specific to the IFMs, where it labels the thin filaments (**B**), whereas Mlp84B specifically labels the Z-discs of leg muscles (**F**). (**G - J**) Tango1-GFP concentrates in a vesicle-like pattern in the gut epithelium (**H, H'**), whereas Par6-GFP is highly expressed in trachea (**I**) and the gut epithelium, where it

concentrates at the apical membrane, as shown for a cross-section of the proventriculus (**J, J'**), nuclei are labelled with DAPI (blue). (**K - M**) Pain-GFP expression in the flight muscle motor neurons (**K**), cells close to the visceral muscles cells (possibly hemocytes) (**L, L'**) and tendon cells (**M, M'**). (**N - P**) LanB1-GFP labels the extracellular matrix surrounding the IFMs, the motor neurons and the trachea (**N**), as well as the visceral muscles (**O**). Even the finest trachea marked by UV auto-fluorescence (white) (**P**) are surrounded by LanB1-GFP (**P'**). Scale bars indicate 100 µm (A, D, I, K, N), 20 µm (H, J, L, O) and 5 µm (B, C, E, F, M, P).

Figure 6 - Supplement 1: ECM and synaptic markers of the adult thorax.

(**A - H**) Antibody stainings of the adult thorax with anti-GFP antibody (green or white in the single colour images) and phalloidin (red). LanA-GFP and BM-40-SPARC-GFP labels the ECM around motor neurons, visceral muscle and trachea. Note that thin trachea within the IFMs (marked by UV auto-fluorescence in white) are surrounded by LanA-GFP but not BM-40-SPARC-GFP (**D**, **H**). (**I - M**) Cact-GFP (green) shows a distinct localisation in IFMs, leg and visceral muscle reminiscent of a neuromuscular junction pattern. Note the partial co-localisation with the motor neuron marker Futsch (in red, **M, M'**), whereas no co-localisation with trachea in IFMs (in white, **M**). Scale bars indicate 100 µm (A, E, I), 20 µm (B, C, F, G, L), 10 µm (M) and 5 µm (D, H, J, K).

Figure 7: Subcellular patterns in adult flight muscles

Antibody stainings of the adult thorax with anti-GFP antibody (green or white in the single colour images) and phalloidin (red). (**A - D**) Localisation to specific myofibrillar sub-regions; Fln-GFP marks the thick filaments (**A**, **A'**), Fray-GFP

surrounds the myofibrils with an enrichment at M-lines and Z-discs (**B**, **B'**), Unc-89-GFP marks only M-lines (**C**, **C'**) and CG31772-GFP only Z-discs (**D**, **D'**). (**E** - **H**) Ilk-GFP strongly concentrates at the muscle-tendon attachment sites (**E**, **E'**), Dlg1-GFP labels the T-tubular membranes (**F**, **F'**), Bab-GFP shows a dotty, vesicular pattern (**G, G'**) and CG12118 displays a mitochondrial pattern (**H**, **H'**). Scale bars indicate 5 µm.

Figure 7 - Supplement 1: Nuclear localisations in adult flight muscles

Antibody stainings of the adult thorax with anti-GFP antibody (green or white in the single colour images), phalloidin (red) and DAPI (blue). (**A** - **H**) CG11617-GFP (**A**, **E**), CG12391-GFP (**B**, **F**) and CG17912 (**C**, **G**) are localised to the nuclei of IFMs and leg muscles, whereas Hb-GFP is only found in leg muscle nuclei (**H**) and not detectable in IFM nuclei (**D**). Scale bars indicate 5 µm.

Figure 7 - Supplement 2: Alternative splicing into alternative C-termini

Antibody stainings of the adult thorax with anti-GFP antibody (green or white in the single colour images) and phalloidin (red). (**A, H**) Gene models of the 3' end of *Mhc* (**A**) and *rhea* (**H**) listing the predicted isoforms; coding exons are shown in pink, 3'UTRs in yellow boxes. The positions of the GFP tag insertions are marked by green arrows. (**B** - **G**) The shorter Mhc-GFP isoforms (Iso K, L, M) are expressed in IFMs and all leg muscles (**B** - **D**), whereas the slightly longer Mhc-GFP isoforms (Iso A, F, G etc.) are not detectable in IFMs but present in visceral muscles and a subset of leg muscles (**E** - **G**). (**I** - **N**) The long Talin-GFP isoforms (*rhea* Iso B, E, F, G) localise to muscle-tendon attachment sites in IFMs (**J**) and leg muscles (**K**), whereas the shorter Talin-GFP isoforms (*rhea* Iso C, D) are not detectable at muscle-tendon attachment

sites in IFMs (**M**, arrowheads) and leg muscles (**N**, arrowhead), however do localise to costamers of leg muscles (**N**). Scale bars indicate 100 µm (B, E, I, L), 10µm (G) and 5 µm (C, D, F, J, K, M, N).

Figure 8: Live imaging of fTRG expression in living pupal thorax

(**A**) Schematic drawing of a 10 - 12 h (left) and a 30h pupal thorax (right). The developing epidermis is shown in blue, with the SOP precursors in yellow (developing neurons in red), the differentiating tendons are shown in orange, the myoblasts and muscle fibers in green, and the muscle-tendon junction in red. The schematic positions of the optical sections through epithelium and muscles are indicated with blue and green dotted lines, respectively. (**B** - **Y**) Live imaging of pupal thoraces at the indicated stages acquired with a spinning disc confocal (except **S** and **T**, which were acquired with a two-photon microscope). Blue bars above the image indicate epithelial sections and green bars indicate muscle sections (as explained in **A**). Hts-GFP is expressed in fusing myoblasts (**B**, **C**) and strongly in developing SOPs (**D**, **E**). Dlg1-GFP labels the epithelial junctions (**F**), internal muscle structures (green dots, **G**) and an unidentified additional developing epithelium (yellow dots, **H**, **I**). Talin-GFP is higher expressed in developing SOPs (**J**, **K**) and strongly localised to the muscle-tendon junction from 24h APF (red arrowheads, **L**, **M**). LanB1-GFP localises to the basal side of the developing epithelium (**N**) and surrounds the forming muscle fibers (green dots, **O** - **Q**) with a slight concentration at the muscle-tendon junction at 30h APF (red arrowheads, **Q**). Act88F-GFP weakly labels the developing epithelium, with a slight concentration in the SOPs until 20h APF (**R**, **S**) and very strongly marks the IFMs from 24h onwards (**T**, **U**). βTub60D-GFPis expressed in the fusing myoblasts (**V**, **W**) and also labels the microtubule

bundles in the developing muscle fibers (**X**, **Y**) and hair cells of the developing sensory organs (light blue arrow heads in **X**). Scale bars indicate 10 µm.

Figure 8 - Supplement 1: Live imaging during pupal development

(**A - F**) Stills from live two-photon imaging of an intact 14h APF pupa expressing Act88F-GFP strongly labelling the developing IFMs (see **Supplementary Movie 4**). (**G - K**) Stills from a two-colour spinning disc movie expressing Act88F-GFP (green) and *him*-GAL4, UAS-palm-Cherry (red) labelling the myoblasts (see **Supplementary Movie 5**). Note the sudden green label of single myoblasts after fusion had occurred (yellow arrowheads, see **Supplementary Movie 5**). (**L - Q**) Stills from two-photon movie of an intact 14h APF pupa expressing βTub-60D-GFP in fusing myoblasts and developing myofibers (See **Supplementary Movie 6**). (**R - V**) Still from a high resolution two-photon movie of an intact 16h APF pupa expressing βTub-60D-GFP. Single myoblast during fusion can be resolved (See **Supplementary Movie 7**). Strong microtubule bundles (red arrow heads) are visible close to the edges of the splitting myotube (white dashed lines, **R**); splitting is complete in (**V**). Scale bars indicate 50 µm (A - F, L - Q) and 10 µm (G - K) and (R - V).

Figure 9: Proteomics with fTRG bait proteins

GFP pull-downs from Ilk-GFP, Dlg1-GFP, Talin-GFP, LanB1-GFP and control pupal (left) or adult fly (right) protein extracts were analysed by mass spectrometry and MaxQuant. The label-free quantification (LFQ) intensities are colour coded. They allow the comparison of protein amounts across samples and correlate with absolute protein amounts. Baits and specific interactors are characterized by quantitative enrichment over the negative controls and samples with non-interacting baits.

Table 1: TransgeneOme constructs and fTRG lines - overview

Overview of TransgeneOme constructs generated and verified by sequencing for the different pilot sets and the genome-wide set, including the respective numbers of the transgenic fTRG lines generated.

Supplementary Table 1: TransgeneOme constructs

Construct and clone names, tag locations, as well the sequencing validation data are listed for all TransgeneOme constructs generated. Sheet 1 lists the sGFP TransgeneOme (TY1-sGFP-V5-BLRP-FLAG tag, NGS sequenced), sheet 2 the TY1-sGFP-FLAG pilot set clones (junctions Sanger sequenced, only exact matches are counted as verified), sheet 3 the TY1-T2A-sGFPnls-FLAG pilot set clones (entire tag Sanger sequenced) and sheet 4 the TY1-sGFP-V5-BLRP-FLAG pilot set clones (entire tag Sanger sequenced). Sheet 5 summarises all verified tagged genes in these sets.

Supplementary Table 2: Transgenic FlyFos (fTRG) lines

Table listing all 880 transgenic FlyFos (fTRG) lines, with fTRG numbers, construct and clones names, as well as nature of the tag and the used landing site. The second sheet compares the genes tagged by the fTRG lines to the available GFP gene trap lines. 765 genes are only found in the TransgeneOme resource.

Supplementary Table 3: Genetic rescue of the fTRG lines

Table listing fTRG lines and respective genetic alleles as well as rescue assays that were used to assess the functionality of the fTRG lines. Note that about two-thirds of the lines fully rescue the mutant phenotypes (marked green).

Supplementary Table 4: FlyFos (fTRG) expression in ovaries

Table listing the expression patterns for 115 fTRG lines in ovaries. Expression was detected in 94 lines by anti-GFP antibody stainings. Cell type specific expression and subcellular localisations were monitored for these lines.

Supplementary Table 5: *in toto* SPIM imaging of FlyFos (fTRG) lines in the embryo

Table listing the fTRG lines that were imaged in the embryo using Zeiss Lightsheet Z.1 from multiple angles over time. *nrv2, gsb* and *gsb-n* are discussed in the text. For the remaining lines we list broad categorisation of the expression detected by SPIM imaging.

Supplementary Table 6: FlyFos (fTRG) expression in the adult thorax

Table listing the expression pattern for 121 fTRG lines in adult thoraces. Expression was detected in101 lines by anti-GFP antibody stainings. Cell type specific expression and subcellular localisations were monitored for these lines.

Supplementary Table 7: Summary of adult muscle patterns of FlyFos (fTRG) lines

54 detected adult muscle localisation patterns (flight muscle, leg muscle and visceral muscle) from Supplementary Table 6 are summarised.

Supplementary Table 8: Proteomics quantification

Quantitative mass spectrometry values of all detected protein obtained with the MaxQuant software suite for all the GFP-enrichment experiments are listed.

Supplementary Movie 1

Multi-view SPIM movie of a stage 12 Nrv2-GFP expressing embryo. A stack was acquired every 15 minutes, lateral, dorsal, ventral and transverse views of the same time points are displayed. From stage 11 onwards Nrv2-GFP is present ubiquitously in the plasma membrane. Later its expression increases in the CNS, particularly in the neuropil and the motor neurons. Movie plays with 7 frames per second. Time is given in hh:mm. Scale bar indicates 50 µm.

Supplementary Movie 2

Lateral head section from a SPIM movie of a stage 10 Gsb-GFP (green, white in the top movie), Histone-2A-mRFPruby (red) embryo. A stack was acquired every 7 minutes. The segmentally re-iterated stripe-like *gsb* expression domain in the head neuroectoderm is visible. Later, *gsb* is expressed in ganglion mother cells and nerve cells that are the progeny of *gsb* expressing neuroblasts. Movie plays with 7 frames per second. Time is given in hh:mm. Scale bar indicates 50 µm.

Supplementary Movie 3

Ventral view of a SPIM movie of a stage 6 Gsb-n-GFP embryo. A stack was acquired every 15 minutes. Gsb-n-GFP is only detectable at the end of germ-band extension. During germ-band retraction it is expressed in characteristic L-shaped expression

domains in the hemi-segments of the trunk. In the late stage embryo Gsb-n-GFP is present in the neurons of the shortening ventral nerve cord. Movie plays with 7 frames per second. Time is given in hh:mm. Scale bar indicates 50 µm.

Supplementary Movie 4

Z-projection of a two-photon movie of an about 14h APF pupa expressing Act88F-GFP. A stack was acquired every 20 min for 19 h. Expression of Act88F-GFP increases in the indirect flight muscles dramatically, thus contrast was reduced several times in course of the movie to avoid over-exposure. Movie plays with 5 frames per second. Time is given in hh:mm.

Supplementary Movie 5

Single plane of a spinning disc confocal movie of an about 14 h APF old pupa expressing Act88F-GFP (green) in the flight muscle myotubes and *him*-GAL4; UAS-palm-Cherry in the myoblasts. An image stack was acquired every two minutes. Note the newly fused myoblasts acquired the GFP label within a single time interval (highlighted by green arrows). Movie plays with 5 frames per second. Time is given in minutes.

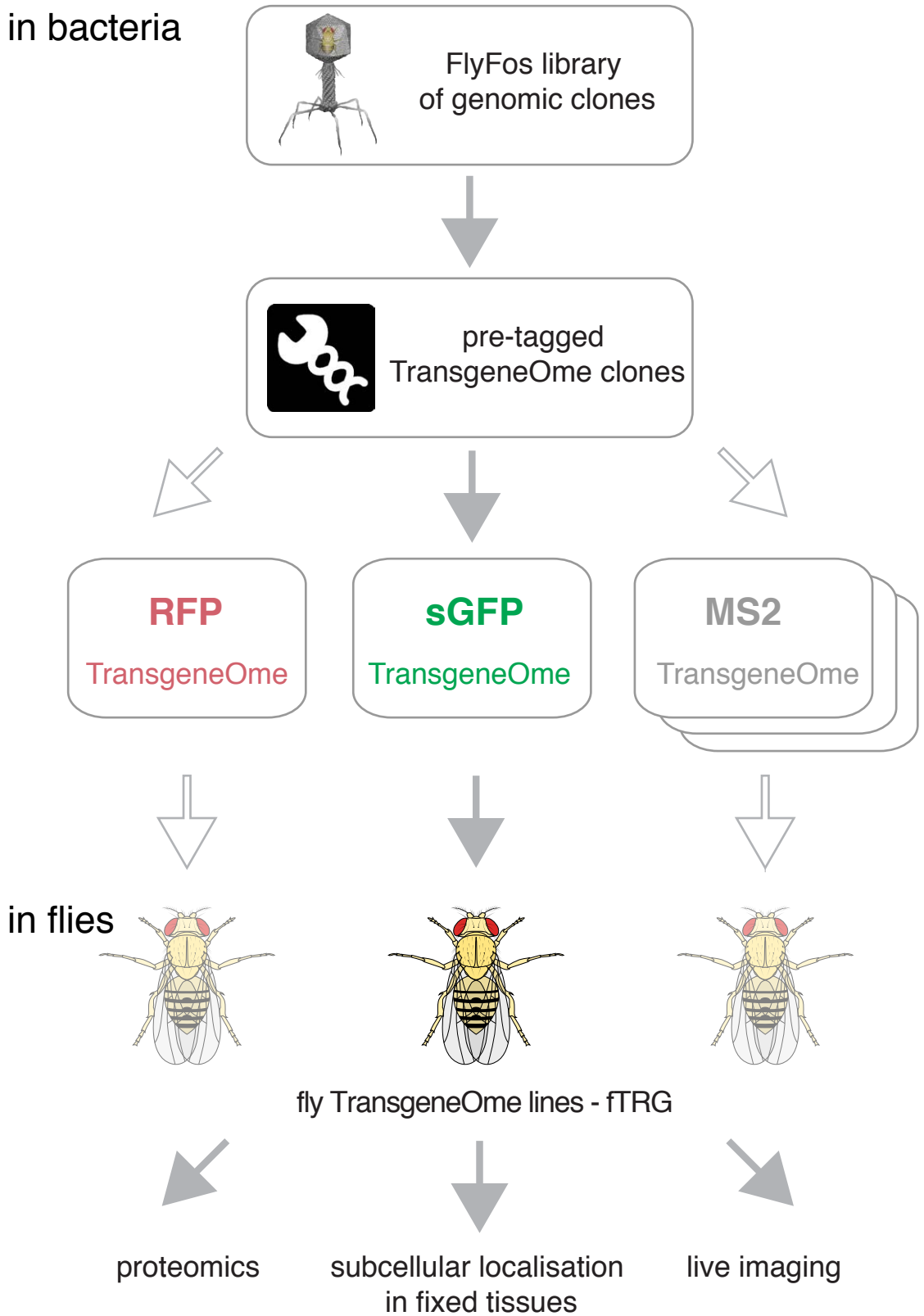
Supplementary Movie 6

Z-projection of a two-photon movie of an about 14h APF pupa expressing βTub60D-GFP. A stack was acquired every 20 min for 25 h. Note the high expression of βTub60D-GFP in fusing myoblasts and the thick microtubules bundles in the developing flight muscles. Hair cells of the developing sensory organs also show

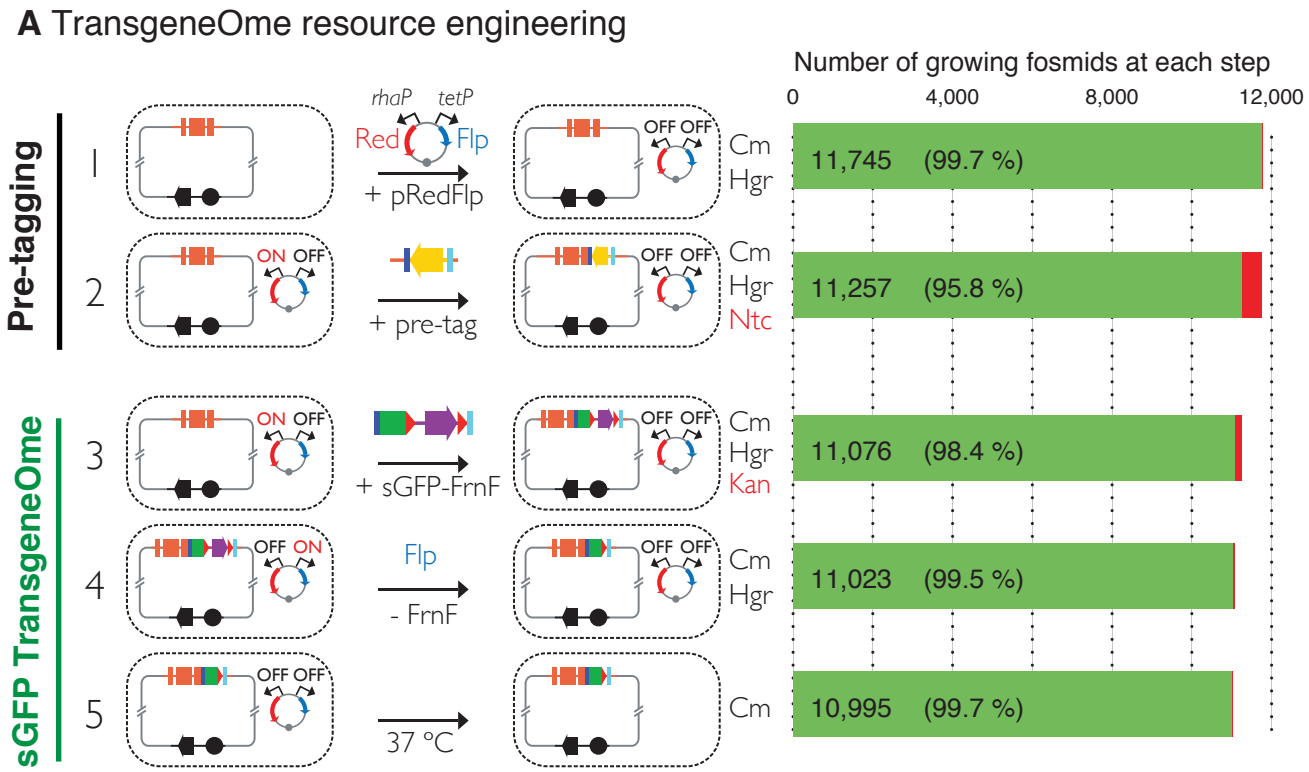
strong expression, however move out of the Z-stack over time. Movie plays with 5 frames per second. Time is given in hh:mm.

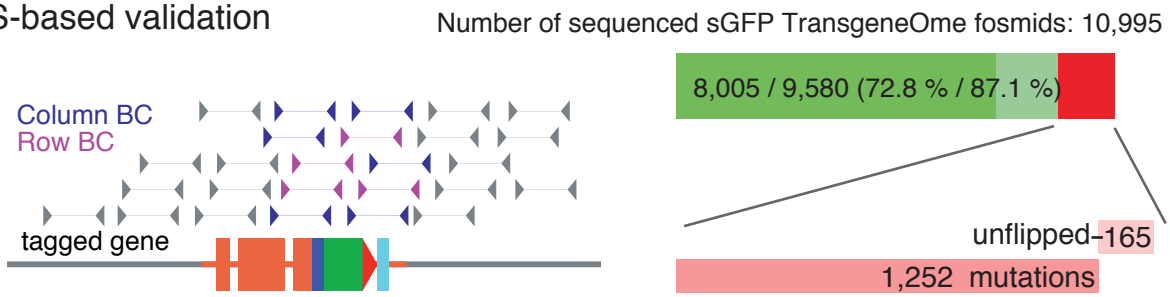
Supplementary Movie 7

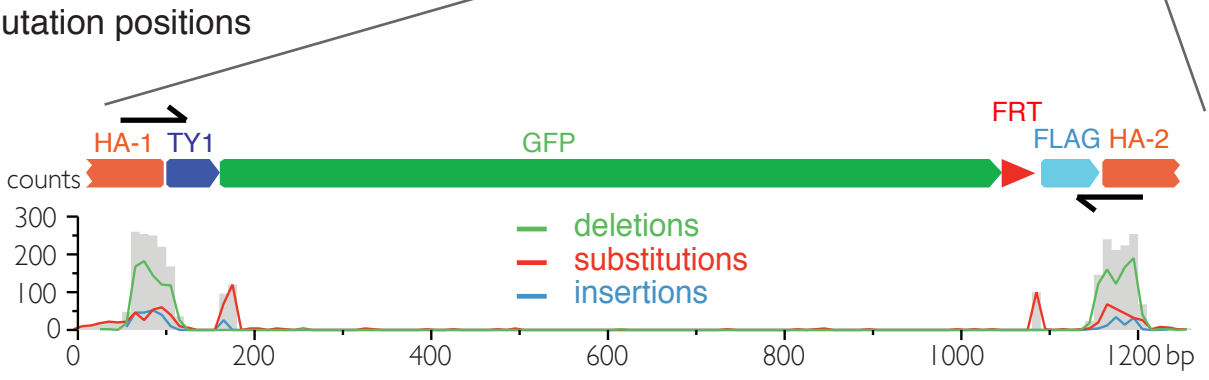
Single plane of a two-photon movie of an about 16 h APF old pupa expressing βTub60D-GFP in myoblasts and the forming flight muscle myotubes. An image stack was acquired every two minutes for more than 3 h. Note that single myoblasts can be followed during fusion. Most myoblasts fuse in the center of the myotube, which gradually splits into two myotubes. Movie plays with 5 frames per second. Time is given in hh:mm.



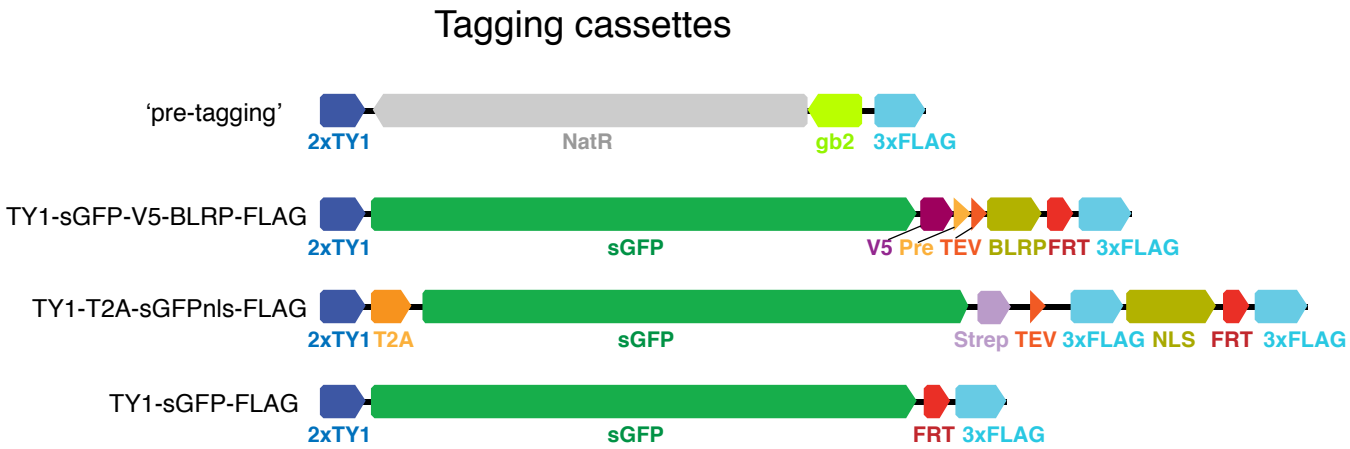
Sarov et al. Figure 1



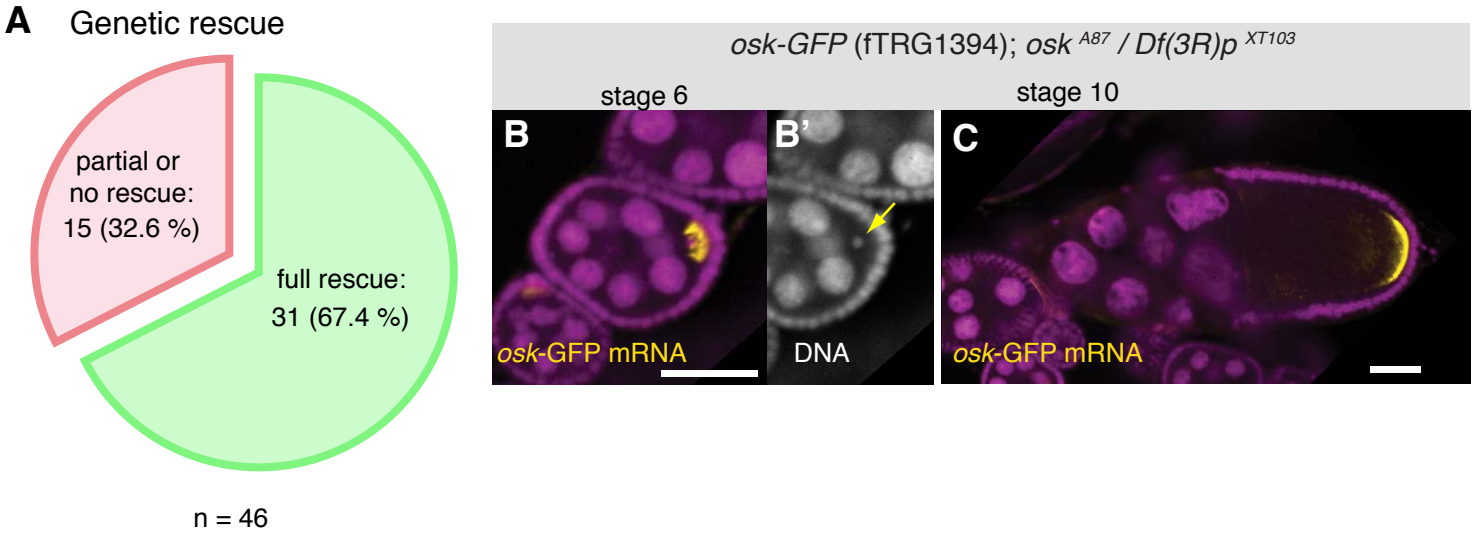




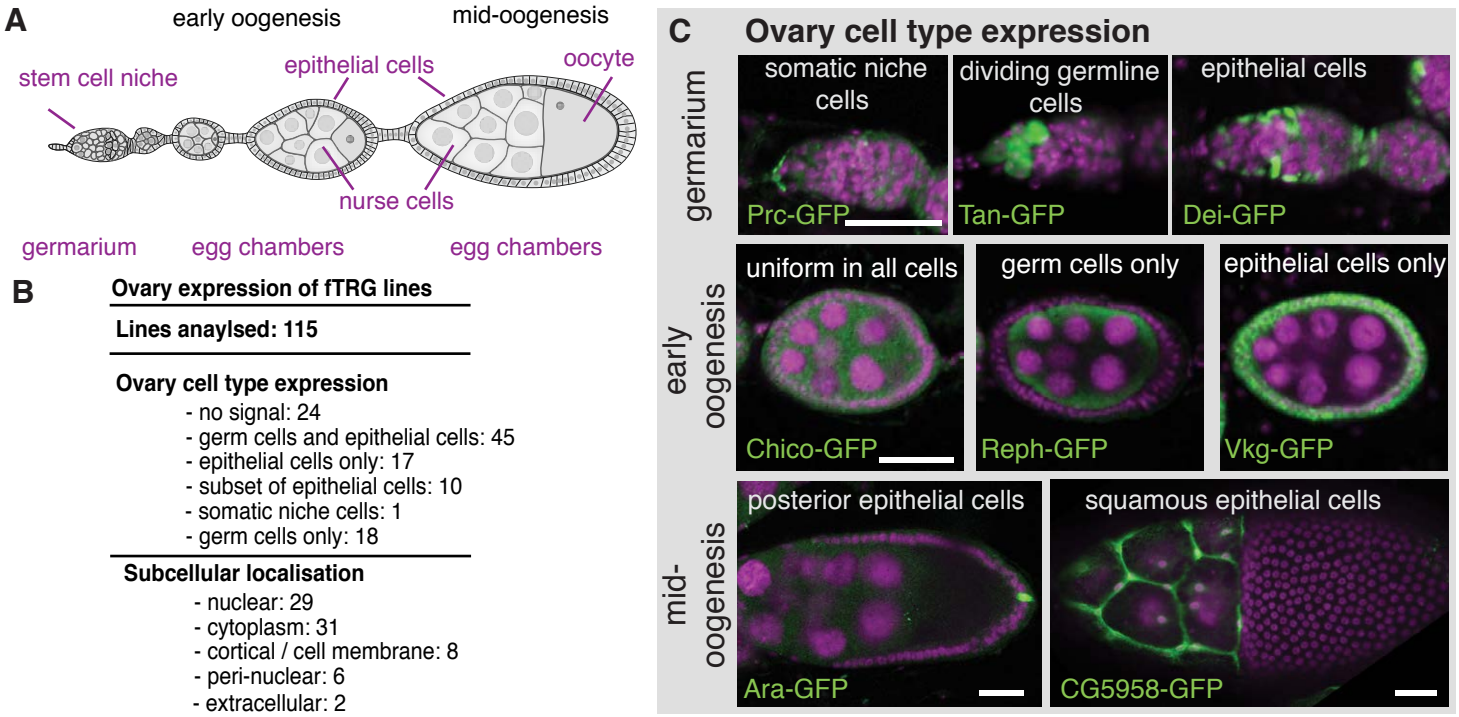
Sarov et al. Figure 2



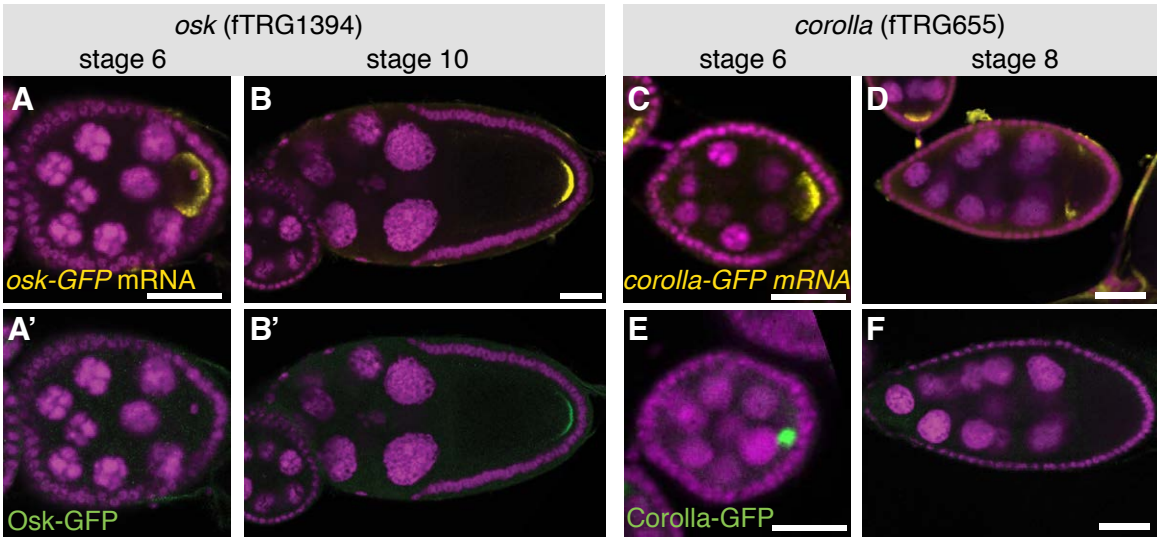
Sarov et al. Figure 2 Supplement 1



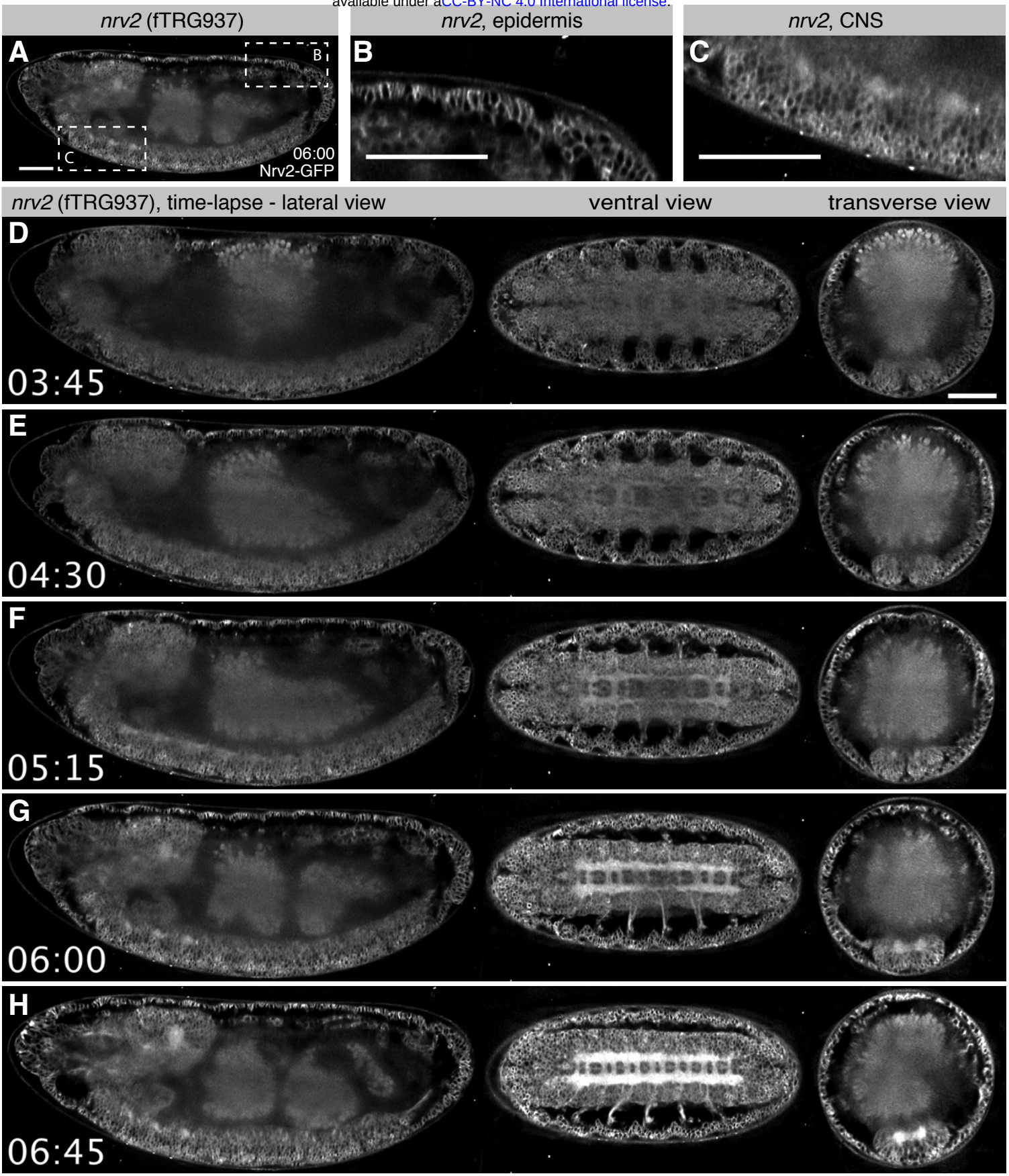
Sarov et al. Figure 3



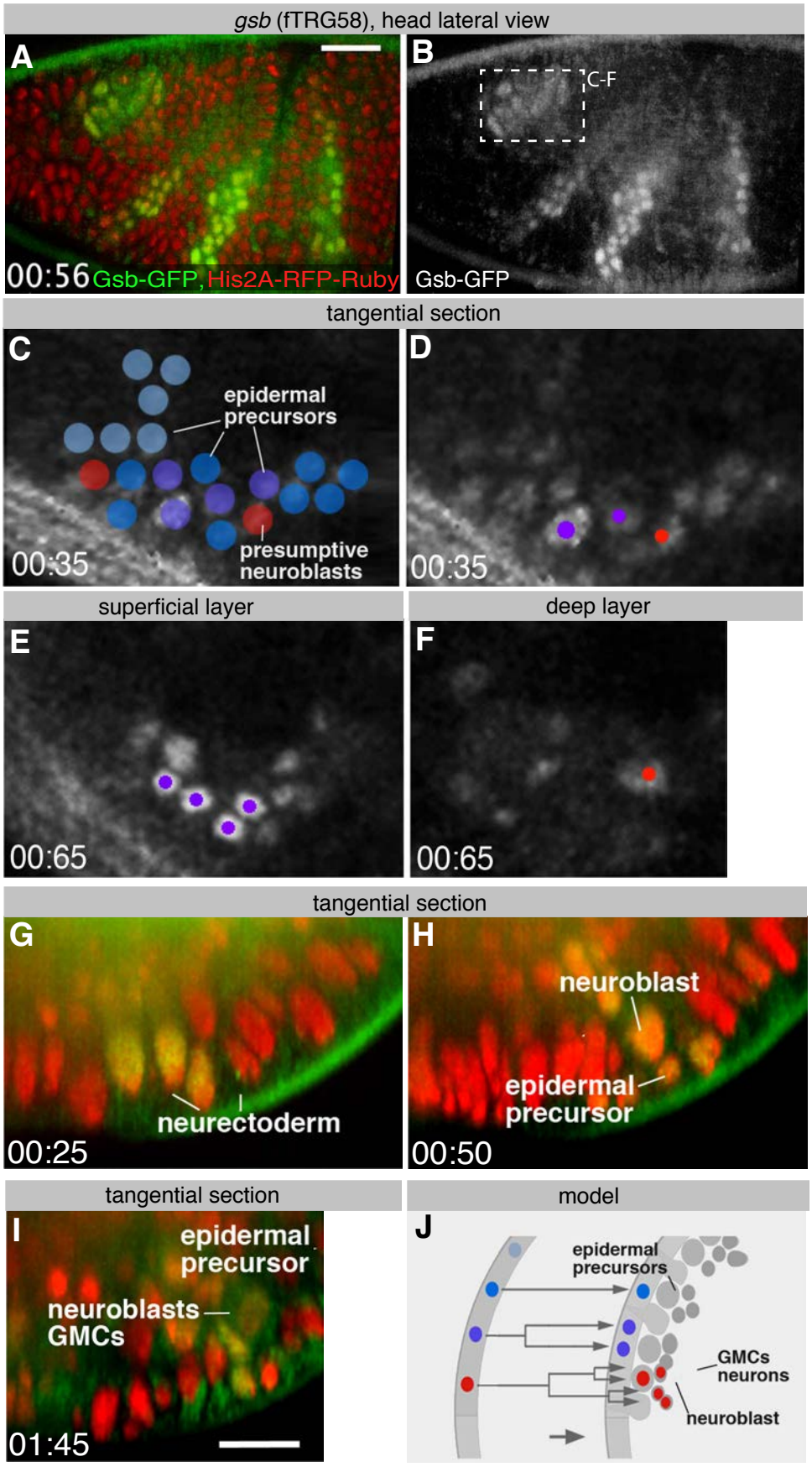
Sarov et al. Figure 4



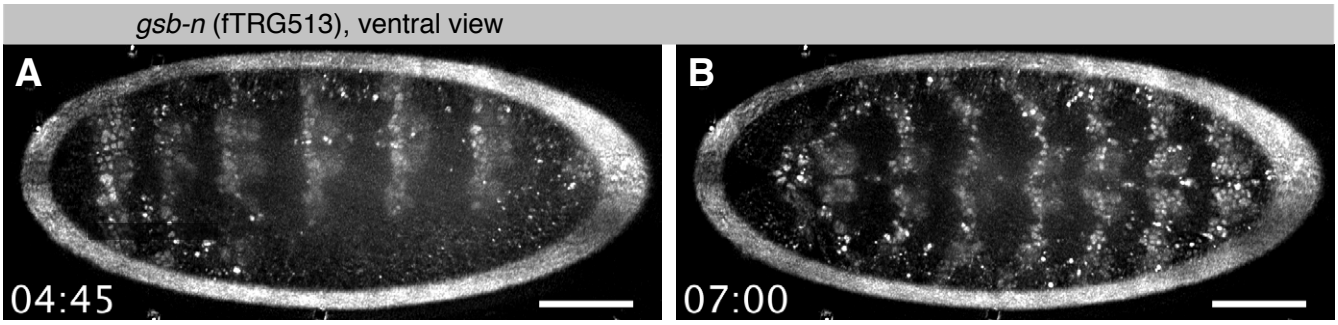
Sarov et al. Figure 4 Supplement 1



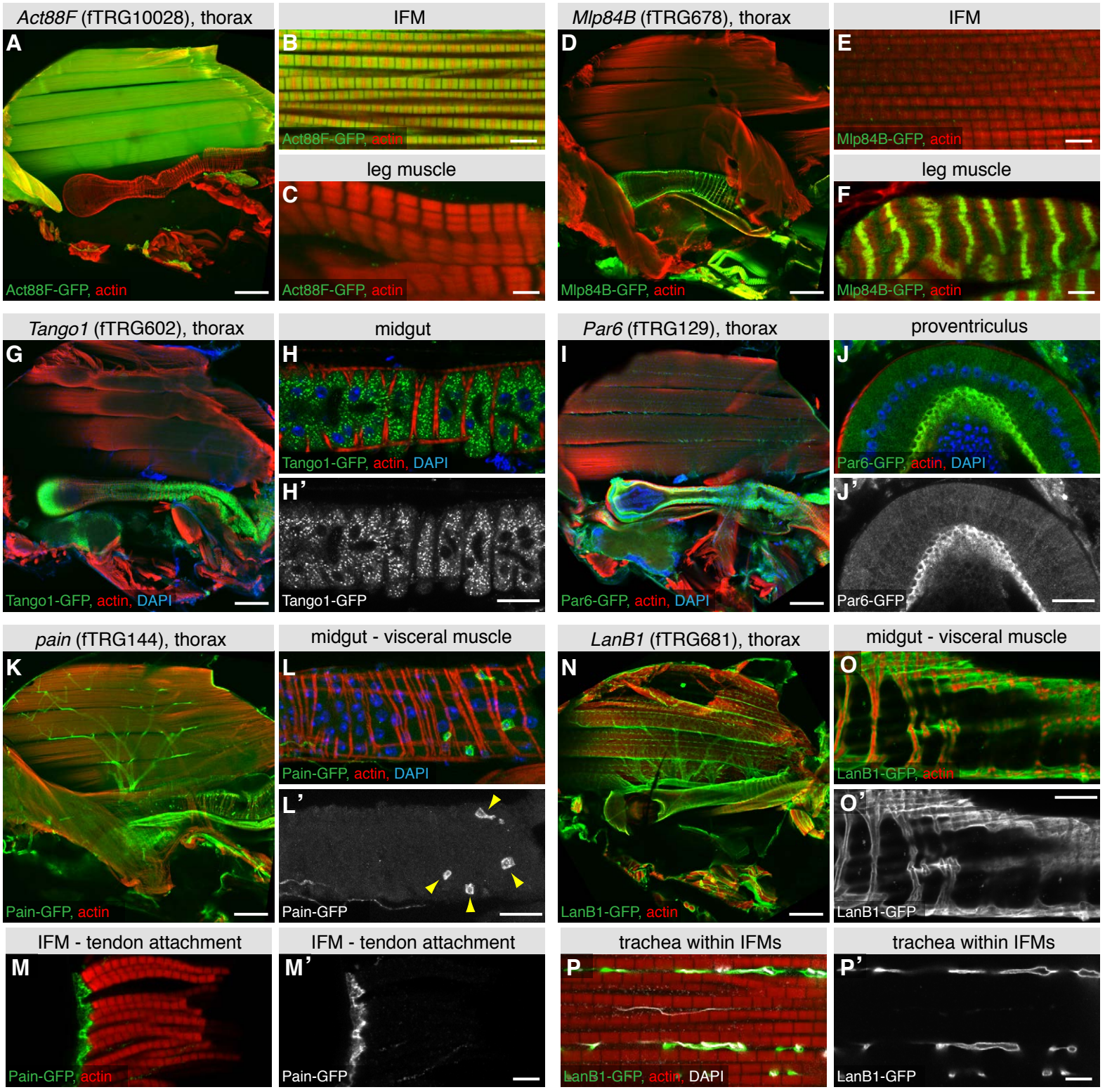
Sarov et al. Figure 5



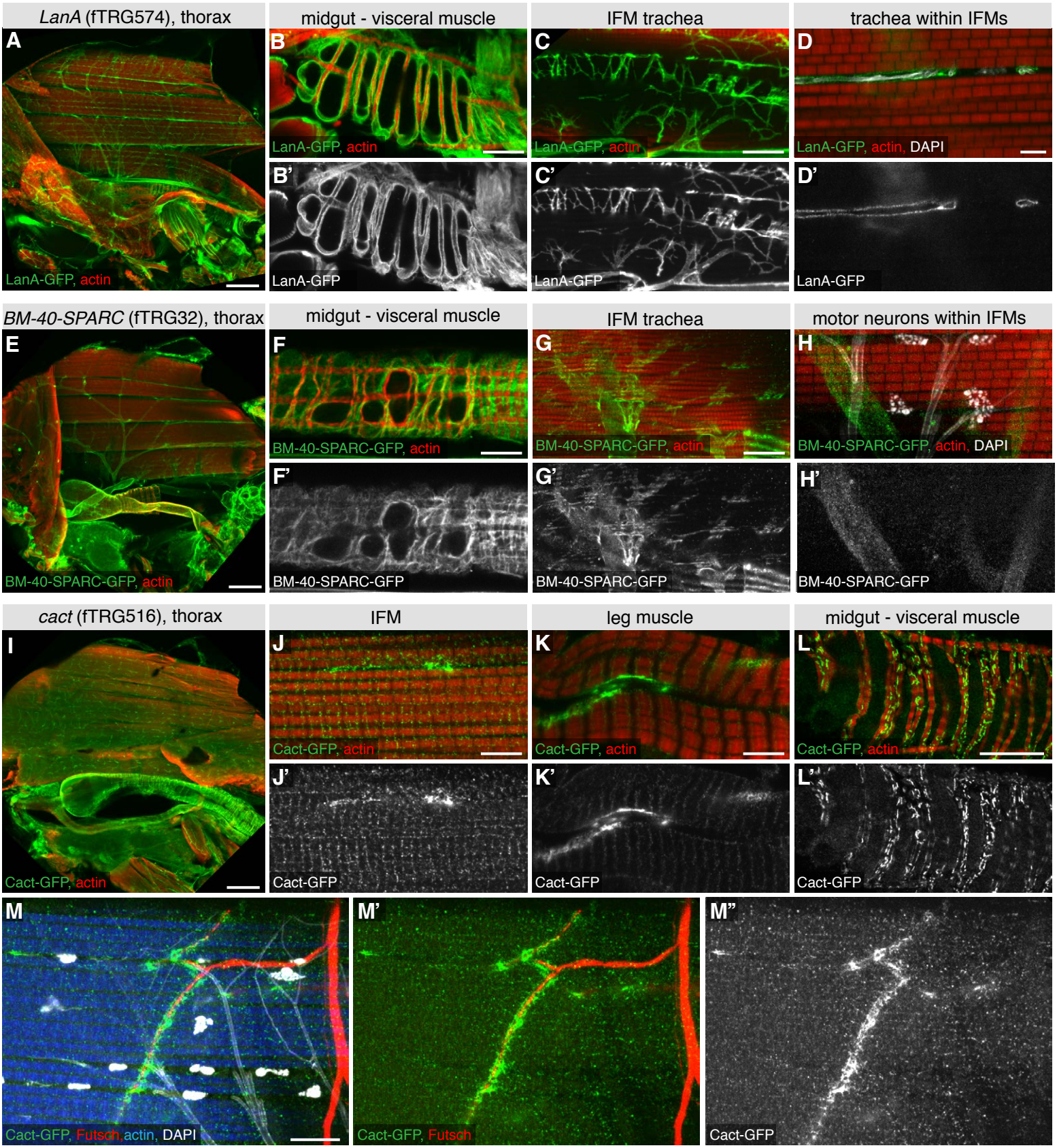
Sarov et al. Figure 5 Supplement 1



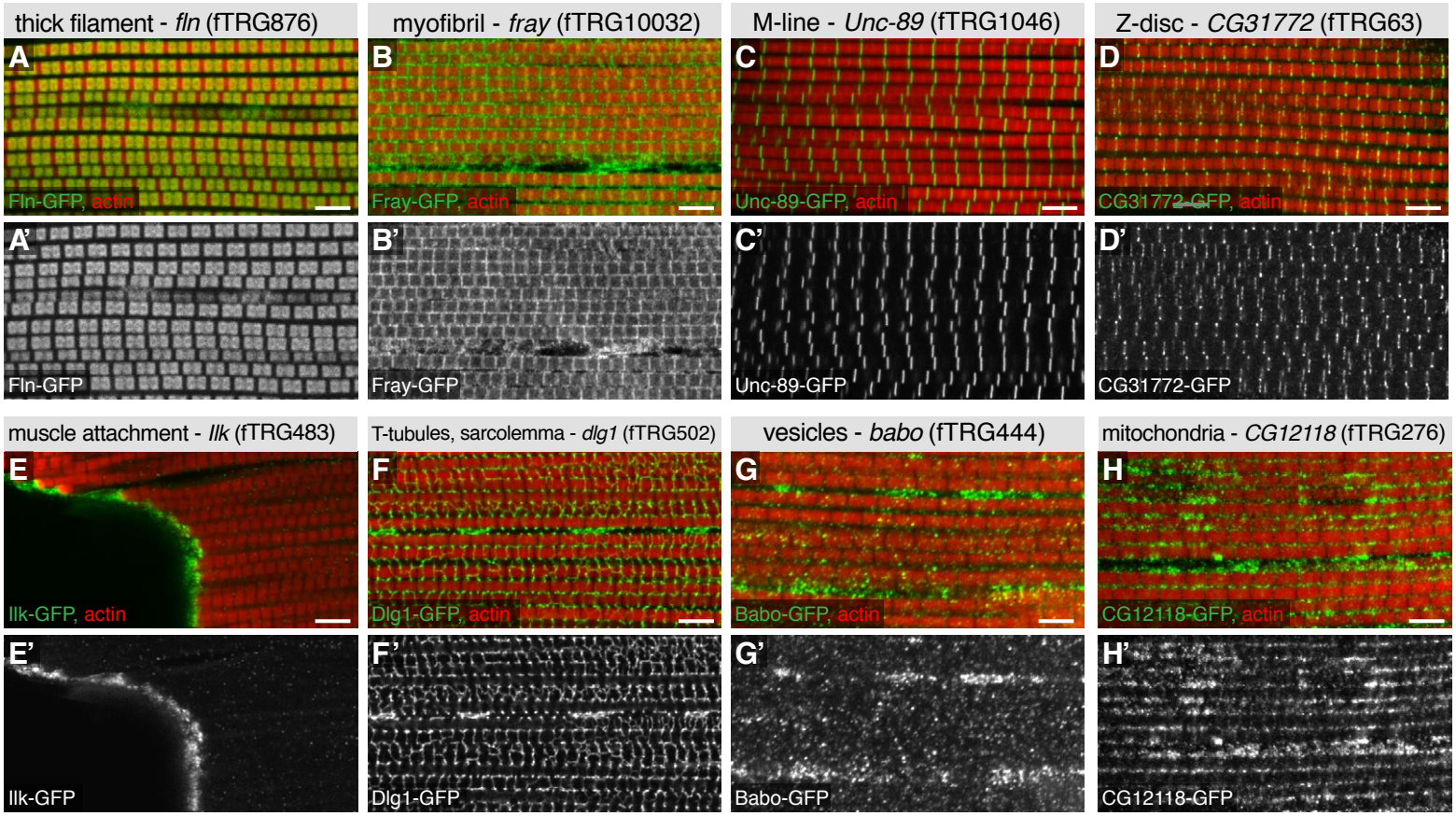
Sarov et al. Figure 5 Supplement 2



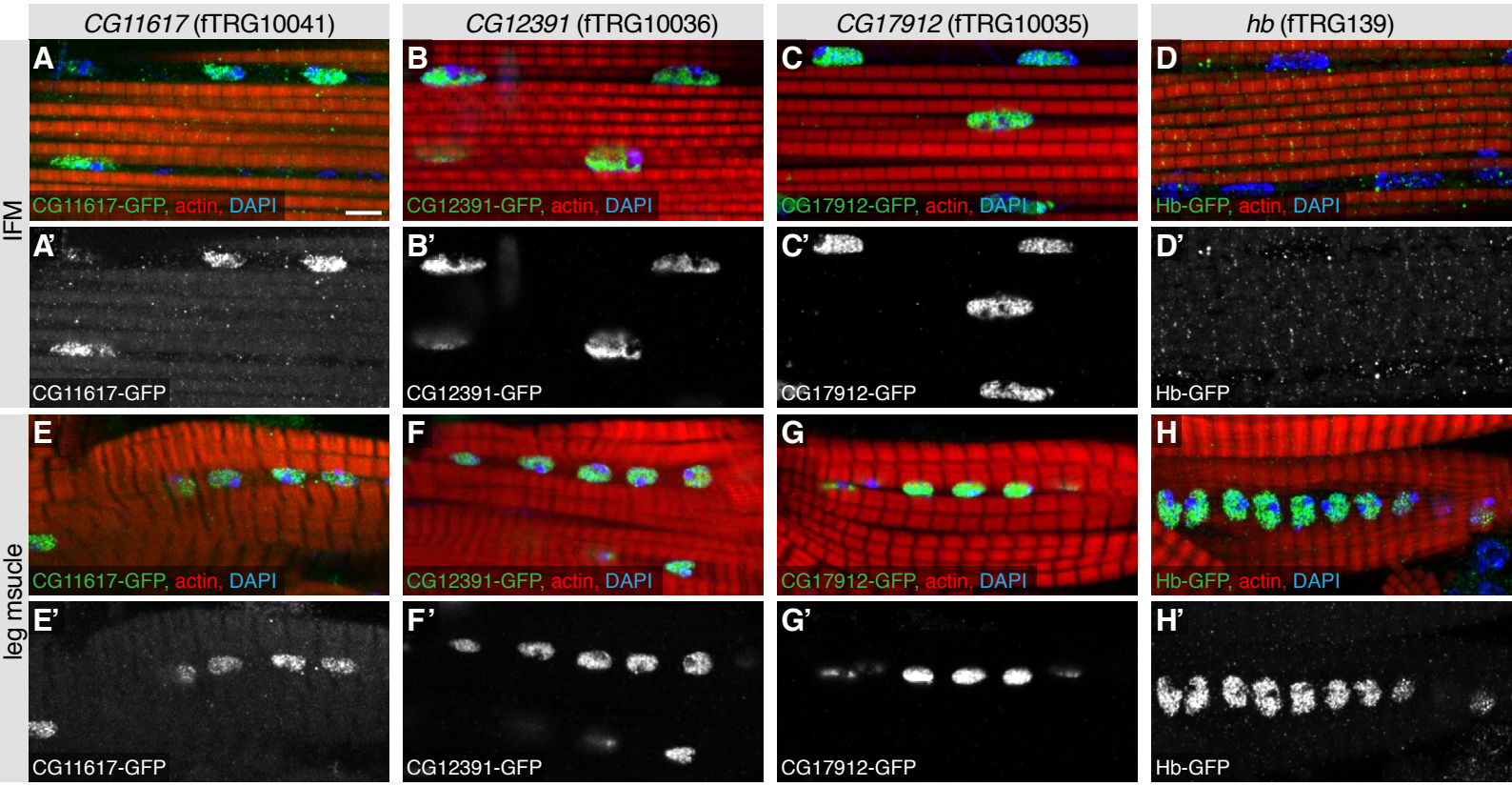
Sarov et al. Figure 6



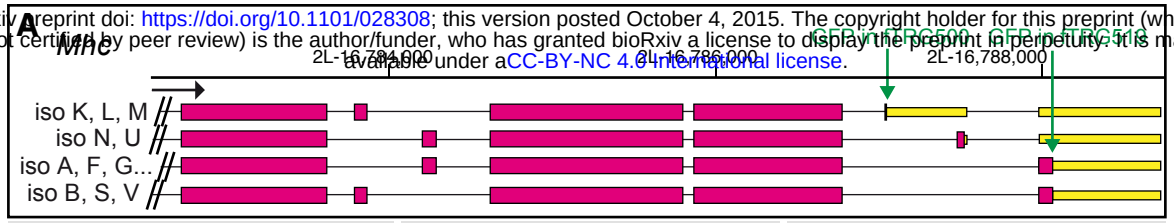
Sarov et al. Figure 6 - Supplement 1

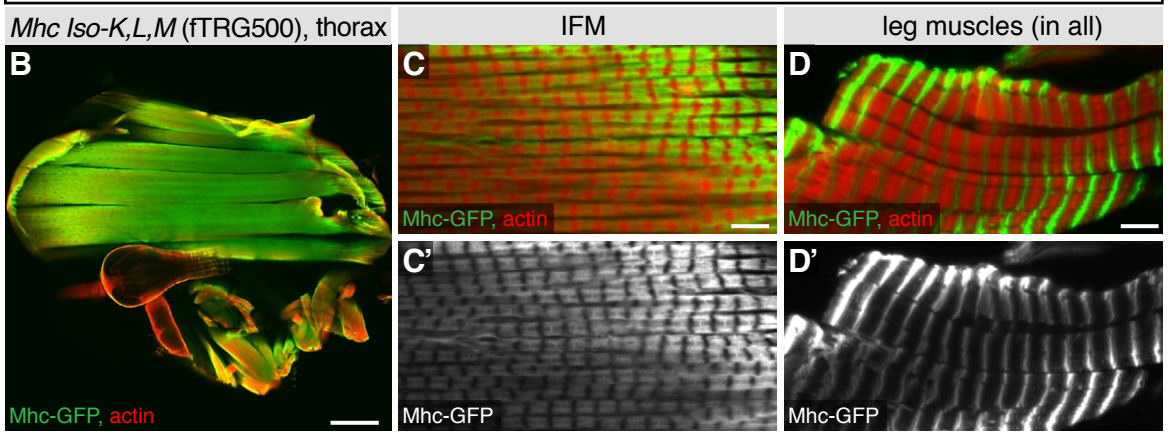


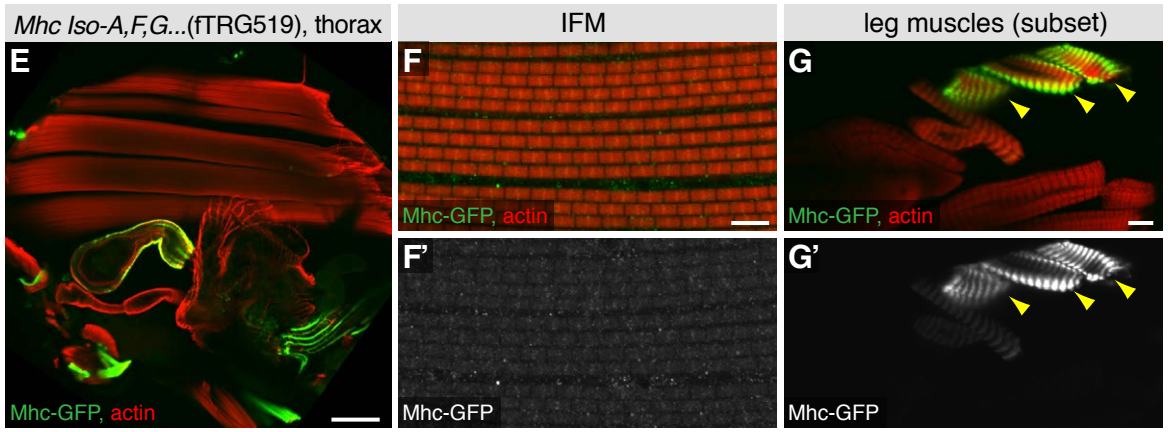
Sarov et al. Figure 7

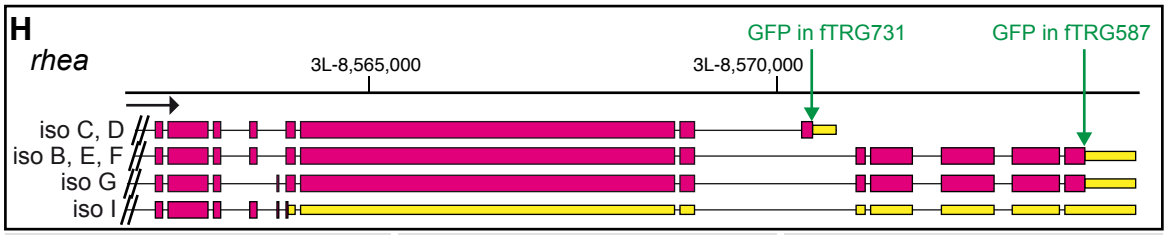


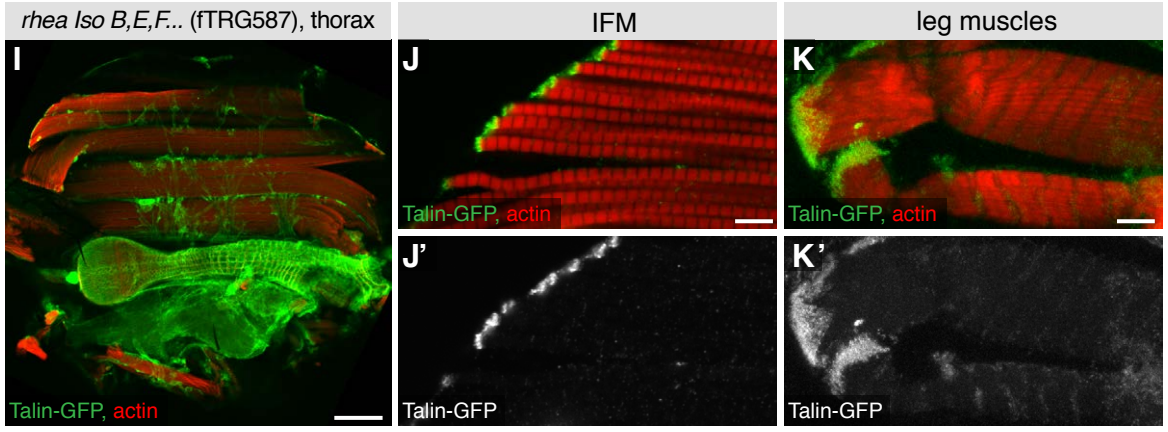
Sarov et al. Figure 7 - Supplement 1

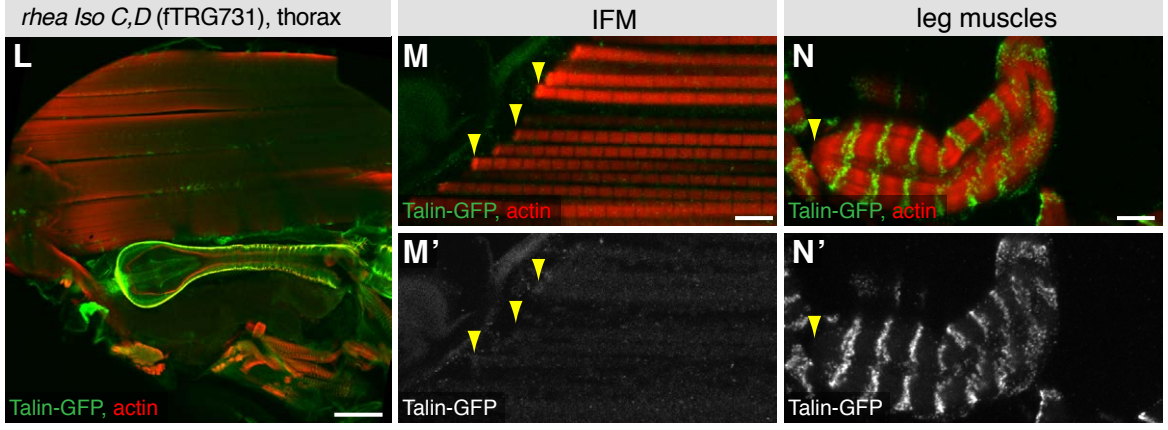




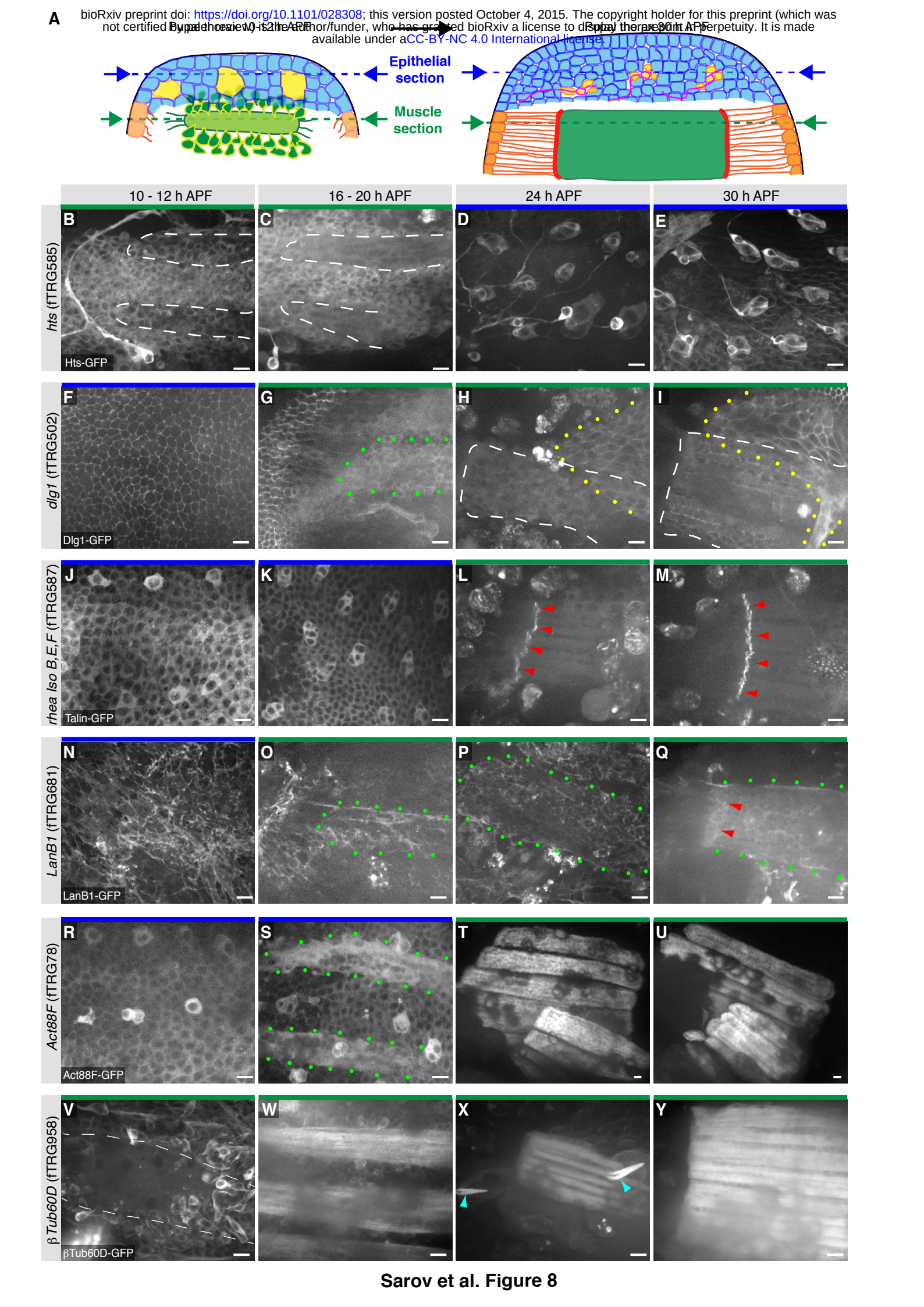




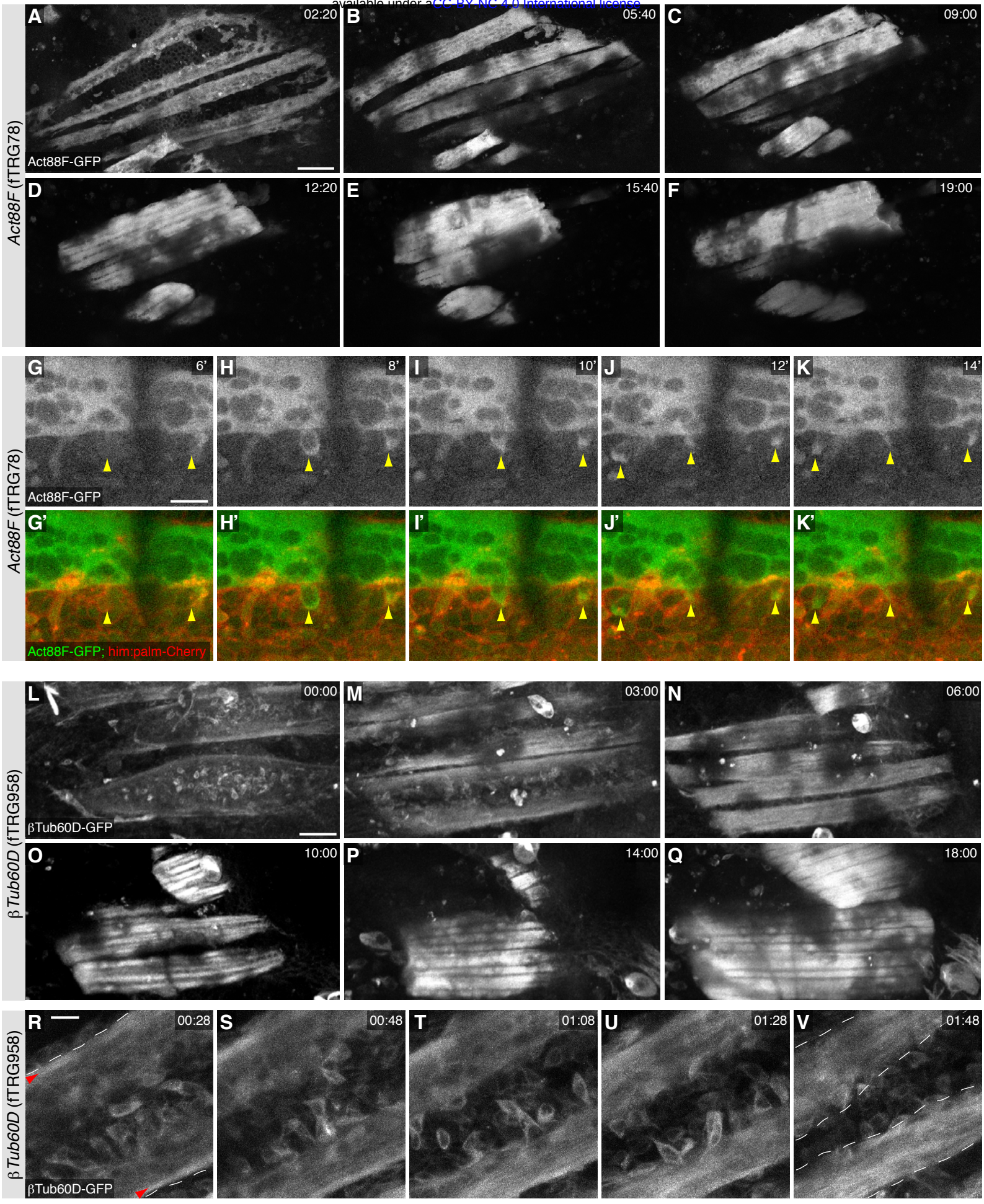




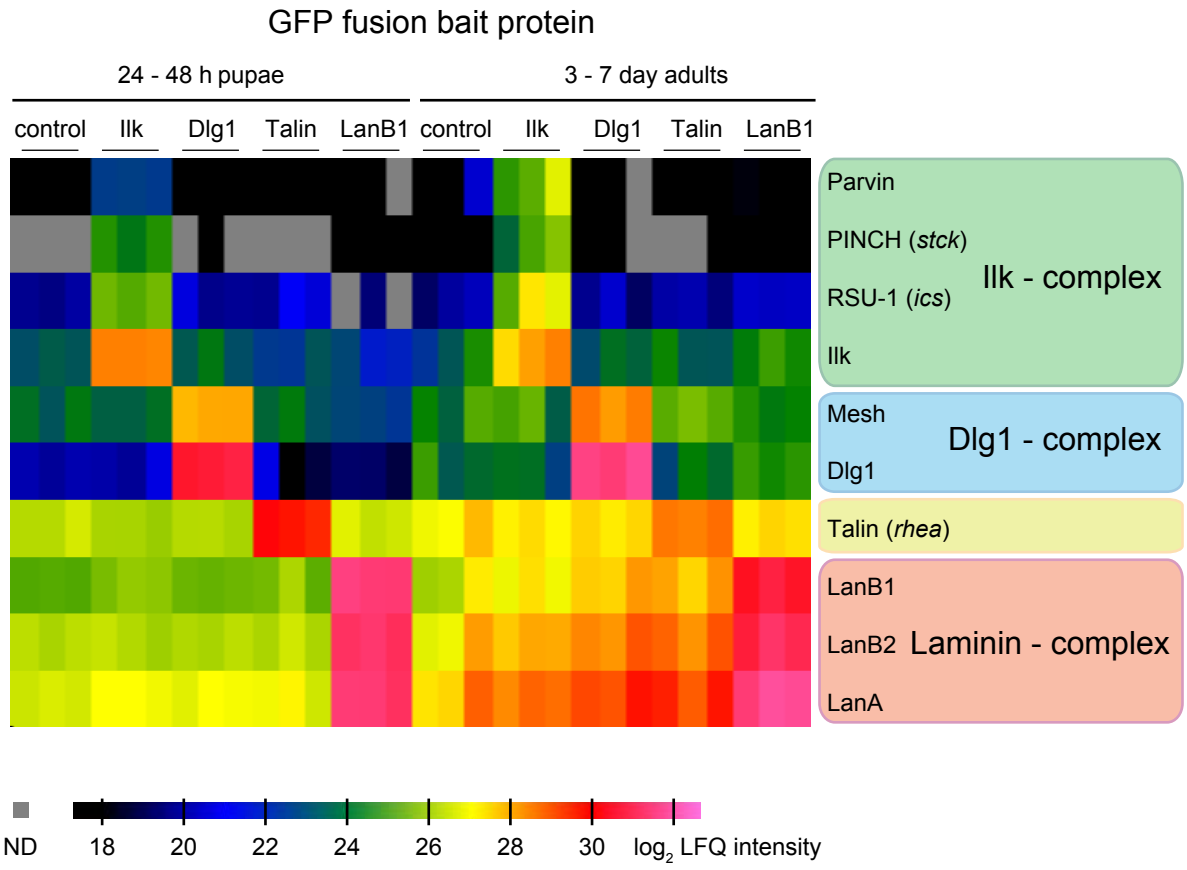
Sarov et al. Figure 7 - Supplement 2



Sarov et al. Figure 8



Sarov et al. Figure 8 - Supplement 1



Sarov et al. Figure 9

Tagged constructs and transgenic lines

	constructs	verified constructs	transgenic lines
'pre-tagging' - TransgeneOme	11,257	--	--
TY1-sGFP-V5-BLRP-FLAG - TransgeneOme - pilot set	10,995 1,328	9,580 1,328	799
TY1-T2A-sGFPnls-FLAG - pilot set	273	273	30
TY1-sGFP-FLAG - pilot set	644	483	51
unique constructs	**23,169**	**10,711**	**880**
unique genes	**11,257**	**9,993**	**826**

Sarov et al. Table 1

Gene	Chromosome	fTRG line	Tag	Rescue?	Rescue assay	Alleles, deficiencies used in trans for rescue assay	Reference
amos	2nd	**fTRG_218**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	antenna size and bristle number rescued to normal	*amos[3]*	
anterior open (aop, Yan)	2nd	**fTRG_142**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	embryonic lethality rescued to viable adults	*aop[1]* (BL-3101); *aop[Yan1]* (BL-8780)	
aubergine (aub)	2nd	**fTRG_581**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	female sterility entirely rescued	*aub[HN2]* (BL-8517); Df(2L)BSC145 (BL-9505)	
baboon (babo)	2nd	**fTRG_444**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	lethality rescued to viable adults	*babo[32]* (BL-5399); *babo[k16912]* (BL-11207)	
bag of marbles (bam)	3rd	**fTRG_3**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	female sterility entirely rescued	*bam[delta86]; Df(3R)exel9020*	Christian Bökel, pers. comm.
cactus (cact)	2nd	**fTRG_516**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	lethality and female sterility rescued	*cact[1]; cact[4]*	
CG32121	3rd	**fTRG_92**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	flightlessness rescued	*CG32121[XZ1]* (X. Zhang and F.S., unpublished); Df(3L)ED4502 (BL-8097)	
CG6509 (dlg5)	2nd	**fTRG_10251**	2xTY1-sGFP-3xFLAG	**yes**	lethality rescued to viable adults	*CG6509[KG006748]* (BL13692); Df(2L)BSC244 (BL-9718)	
discs large 1 (dlg1)	X	**fTRG_502**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	male lethality rescued to viable adults	*Dlg1[5]* (BL-36280)	
dorsal (dl)	2nd	**fTRG_29**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	bristle number rescued to normal	*dl[1]; dl[4]*	
ebi	2nd	**fTRG_10141**	2xTY1-sGFP-3xFLAG	**yes**	lethality rescued to viable adults	*ebi[CCS-8]* (BL-8397); *ebi[E90]* (BL-30720)	
escargot (esg)	2nd	**fTRG_10170**	2xTY1-sGFP-3xFLAG	**no**	lethality not rescued	*esg[35Ce-1]* (BL-3900); *esg[35Ce-3]* (BL-30475)	
eyes absent (eya)	2nd	**fTRG_492**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**no**	lethality not rescued	*eya[C0233]; eya[C0275]*	
fat (ft)	2nd	**fTRG_10233**	2xTY1-T2A-nlsGFP-3xFLAG	**yes**	lethality rescued to viable adults	*ft[G-rv]* (BL-1894); *ft[8]* (BL-5406)	
48 related 2 (Fer2)	3rd	**fTRG_334**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	defective climbing rescued to wild type	*Fer2[e03248]*	Bou Dib *et al.* 2014
fizzy (fzy)	2nd	**fTRG_10250**	2xTY1-T2A-nlsGFP-3xFLAG	**no**	lethality not rescued	*fzy[1]* (BL-2492); *fzy[3]* (BL-25143)	
flightless I (fliI)	X	**fTRG_467**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	lethality or flightlessness rescued	*fliI[14]* (BL-7481); *fliI[3]* (BL-4730)	
Hand	2nd	**fTRG_10163**	2xTY1-sGFP-3xFLAG	**yes**	lethality rescued to viable adults	*Hand[173]*	
hippo (hpo)	2nd	**fTRG_10130**	2xTY1-sGFP-3xFLAG	**yes**	larval lethality rescued to viable adults	*hpo[KS240]* (BL-25085); *hpo[KC202]* (BL-25090)	
HLH54F	2nd	**fTRG_153**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	lethality rescued to viable adults	*bHLH54F[598];* Df(2R)Exel7150 (BL-7891)	
Integrin linked kinase (Ilk)	3rd	**fTRG_483**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	embryonic lethality rescued to viable adults (wing blisters)	*Ilk[1]* (BL-4861) ; *Df(3L)BSC733* (BL-26831)	
Kinesin heavy chain (Khc)	2nd	**fTRG_10243**	2xTY1-T2A-nlsGFP-3xFLAG	**yes**	lethality rescued to viable adults	*Khc[8]* (BL-1607); *Khc[1ts]* (BL-31994)	
LanB1	2nd	**fTRG_681**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	lethality rescued to viable adults	*LanB1[KG03456]* (BL-13957); *Df(2L)Exel7032* (BL-7806)	
multiple ankyrin repeats single KH domain (mask)	3rd	**fTRG_486**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	lethality rescued to viable adults	*mask[10.22]/Df(3R)BSC317*	Barry Thompson, pers. comm
midline (mid)	2nd	**fTRG_490**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**no**	lethality not rescued	*mid[B1295]; mid[C2372]*	
numb	2nd	**fTRG_25**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	lethality rescued to viable adults	*numb[1]* (BL-4096); Df(2L)30A-C(BL-3702)	
odd skipped (odd)	2nd	**fTRG_47**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**no**	lethality not rescued	*odd[5]* (BL-5345); Df(2L)Exel7018 (BL-7789)	
optomotor-blind-related-gene-1 (org-1)	X	**fTRG_485**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**no**	male lethality not rescued	*org-1[OJ487]*	
oskar (osk)	3rd	**fTRG_1394**	2XTY1-SGFP-V5-preTEV-BLRP-3XFLAG	**yes**	female sterility entirely rescued	*osk[A87]/Df(3R)p-XT103*	
Pabp2	2nd	**fTRG_565**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**no**	lethality not rescued	*Pabp2[01]* (BL-9838); *Pabp2[55]* (BL-38390)	
patched (ptc)	2nd	**fTRG_82**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**yes**	lethality rescued to viable adults	*ptc[9]* (BL-3377); *ptc[16]* (BL-35500)	
retina abarrent in pattern (rap)	X	**fTRG_1253**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	lethality rescued to viable adults	*rap[ie28]*	Yuu Kimata, pers. comm.
rhea (Talin)	3rd	**fTRG_587**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	embryonic lethality rescued to viable adults	*rhea[1]* ; *rhea[79]*	
RhoGEF2	2nd	**fTRG_591**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	embryonic lethality rescued to viable adults	*RhoGEF2[04291]*	Jörg Großhans, pers. comm.
roundabout (robo)	2nd	**fTRG_567**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**no**	lethality not rescued	*robo[1]* (BL-8755); *robo[2]* (BL-8756)	
saxophone (sax)	2nd	**fTRG_10070**	2xTY1-sGFP-3xFLAG	**yes**	lethality rescued to viable adults	*sax[4]* (BL-5404); *sax[5]* (BL-8785)	
scribbler (sbb)	2nd	**fTRG_443**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**no**	lethality not rescued	*sbb[04440]* (BL-11376); Df(2R)BSC334 (BL-24358)	
Sin3A	2nd	**fTRG_596**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**no**	lethality not rescued	*Sin3A[08269]* (BL-12350); *Sin3A [B0948]*	
smoothened (smo)	2nd	**fTRG_599**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	lethality rescued to viable adults	*smo[3]* (BL-3277); *smo[119B6]* (BL-24772)	
snail (sna)	2nd	**fTRG_71**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**no**	lethality not rescued	*sna[18]* (BL-2311); *sna[1]* (BL-25127)	
spalt major (salm)	2nd	**fTRG_165**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**no**	lethality not rescued	*salm[1]* (BL-3274); Df(2L)32FP-5 (BL-29717)	
Target of rapamycin (Tor)	2nd	**fTRG_713**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**no**	lethality not rescued	*Tor[deltaP]* (BL-7014); Df(2L)Exel7055 (BL-7823)	
traffic jam (tj)	2nd	**fTRG_163**	2xTY1-sGFP-V5-preTEV-BLRP-3xFLAG	**no**	sterility not rescued	*tj[PL3]* (BL-4987); Df(2L)Exel8041 (BL-7849)	
viking (vkg)	2nd	**fTRG_595**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**no**	lethality not rescued	*vkg[01209]* (BL-11003); Df(2L)Exel7022 (BL-7794)	
Unc-89/Obscurin	2nd	**fTRG_1046**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	flightlessness rescued	*Unc-89[EY15484]*	
yorkie (yki)	2nd	**fTRG_875**	2xTY1-sGFP-V5-preTEV-BLRP-3XFLAG	**yes**	lethality rescued to viable adults	yki[B5]	Barry Thompson, pers. comm

Sarov et al. Supplementary Table 3

Transformant_fTRGnumber	FBgn_id	Gene	ovary cell type expression	sub-cellular localisation	nuclear	cytoplasm	cortical	interesting
4	FBgn0011763	Dp	germ cells and epithelial cells	nuclear	1			1
9	FBgn0037992	CG4702	no signal	n.a.				
13	FBgn0038032	CG10096	no signal	n.a.				
25	FBgn0002973	numb	germ cells and epithelial cells	cortical / cell membrane			1	1
32	FBgn0026562	BM-40-SPARC	no signal	n.a.				
53	FBgn0037513	pyd3	subset of epithelial cells: anterior/posterior cells	cytoplasm		1		1
64	FBgn0032157	Etl1	germ cells and epithelial cells	nuclear	1			1
74	FBgn0035471	sc2	germ cells and epithelial cells	oocyte enriched / nurse cells perinuclear				1
76	FBgn0037443	CG1021	germ cells and epithelial cells	oocyte enriched / nurse cells perinuclear				1
84	FBgn0039044	p53	germ cells	nuclear	1			1
85	FBgn0259685	Dr	germ cells and epithelial cells	nuclear	1			1
87	FBgn0035625	Blimp-1	no signal	n.a.				
88	FBgn0025360	Optix	no signal	n.a.				
91	FBgn0000658	fj	subset of epithelial cells	cytoplasm		1		1
101	FBgn0008646	E5	subset of epithelial cells: anterior/posterior cells	nuclear	1			1
125	FBgn0039269	veli	germ cells and epithelial cells	uniform in all cells				
134	FBgn0015218	eIF-4E	germ cells and epithelial cells	uniform in all cells				
137	FBgn0259685	crb	germ cells and epithelial cells	oocyte enriched / apical in follicle cells			1	1
148	FBgn0015904	ara	subset of epithelial cells: anterior/posterior cells	cytoplasm		1		1
150	FBgn0028980	tan	germ cells	cytoplasm		1		1
166	FBgn0263118	dei	subset of epithelial cells	nuclear	1			1
168	FBgn0024191	sip1	germ cells and epithelial cells	cytoplasm / nuclear	1	1		1
178	FBgn0041188	Atx2	epithelial cells	cytoplasm		1		1
179	FBgn0035856	CG13679	no signal	n.a.				
182	FBgn0016076	vri	epithelial cells	nuclear	1			1
192	FBgn0005642	wdn	germ cells and epithelial cells	nuclear	1			1
207	FBgn0032130	CG3838	no signal	n.a.				
240	FBgn0004101	bs	germ cells and epithelial cells	nuclear	1			1
245	FBgn0020309	crol	epithelial cells	nuclear	1			1
257	FBgn0032858	CG10949	no signal	n.a.				
264	FBgn0261363	Dox-A3	germ cells	cytoplasm		1		1
275	FBgn0027561	CG18659	no signal	n.a.				
286	FBgn0038179	CG9312	no signal	n.a.				
288	FBgn0035878	CG7182	germ cells and epithelial cells	uniform in all cells				
289	FBgn0034084	CG8435	germ cells and epithelial cells	nuclear	1			1
296	FBgn0036483	CG12316	germ cells	cytoplasm / perinuclear		1		1
308	FBgn0040305	MTF-1	epithelial cells	cytoplasm		1		1
323	FBgn0015371	chn	germ cells and epithelial cells	nuclear	1			1
329	FBgn0036004	CG3654	germ cells and epithelial cells	cytoplasm / nuclear	1	1		1
333	FBgn0259785	CG7752	epithelial cells	nuclear	1			1
340	FBgn0003254	rib	no signal	n.a.				
349	FBgn0024294	Spn43Aa	subset of epithelial cells: anterior/posterior cells	cytoplasm		1		1
351	FBgn0020617	Rx	no signal	n.a.				
356	FBgn0022935	D19A	epithelial cells	nuclear	1			1
368	FBgn0010313	Corto	subset of epithelial cells: anterior/posterior cells	cytoplasm		1		1
372	FBgn0032015	Ostgamma	germ cells	cytoplasm / perinuclear		1		1
373	FBgn0034194	CG15611	germ cells	cytoplasm		1		1
376	FBgn0038768	CG4936	germ cells and epithelial cells	nuclear	1			1
377	FBgn0029905	Nf-YC	germ cells and epithelial cells	cytoplasm / nuclear	1	1		1
401	FBgn0040318	HGTX	no signal	n.a.				
408	FBgn0028996	onecut	no signal	n.a.				
409	FBgn0035036	CG4707	germ cells and epithelial cells	nuclear	1			1
425	FBgn0085065	CG5669	germ cells and epithelial cells	nuclear	1			1
436	FBgn0261020	wol	germ cells	oocyte enriched				1
443	FBgn0010575	sbb	germ cells and epithelial cells	nuclear	1			1
448	FBgn0020386	Pdk1	no signal	n.a.				
456	FBgn0024248	chico	germ cells and epithelial cells	uniform in all cells				
458	FBgn0035872	CG7185	germ cells and epithelial cells	nuclear	1			1
467	FBgn0000709	flil	no signal	n.a.				
480	FBgn0011591	fng	epithelial cells	cytoplasm		1		1
482	FBgn0010389	htl	no signal	n.a.				
483	FBgn0028427	Ilk	germ cells and epithelial cells	uniform in all cells				
488	FBgn0004635	rho	subset of epithelial cells: anterior/posterior cells	cytoplasm		1		1
492	FBgn0000320	eya	epithelial cells	cytoplasm		1		1
496	FBgn0004054	zen2	no signal	n.a.				
497	FBgn0028573	prc	somatic niche cells	extracellular				1
511	FBgn0015946	grim	no signal	n.a.				
519	FBgn0264695	Mhc	germ cells and epithelial cells	uniform in all cells				
521	FBgn0000625	eyg	subset of epithelial cells: anterior/posterior cells	cytoplasm		1		1
527	FBgn0087007	bbg	epithelial cells	apical cell membrane			1	1
539	FBgn0003459	stwl	germ cells and epithelial cells	uniform in all cells				
550	FBgn0000541	E(bx)	germ cells and epithelial cells	nuclear	1			1
565	FBgn0005648	Pabp2	germ cells and epithelial cells	nuclear	1			1
568	FBgn0010341	Cdc42	germ cells and epithelial cells	uniform in all cells				
576	FBgn0011655	Med	no signal	n.a.				
577	FBgn0262526	vas	germ cells	cytoplasm		1		1
595	FBgn0016075	vkg	epithelial cells	cytoplasm		1		1
601	FBgn0033402	Myd88	germ cells and epithelial cells	uniform in all cells				
603	FBgn0035473	mge	germ cells and epithelial cells	germ cells: oocyte enriched, nurse cells aggregates				1
608	FBgn0034068	casp	epithelial cells	cytoplasm		1		1
616	FBgn0029664	CG10802	no signal	n.a.				
617	FBgn0028473	Non1	epithelial cells	nuclear	1			1
621	FBgn0021800	Reph	germ cells	cytoplasm		1		1
634	FBgn0033717	CG8839	epithelial cells	cytoplasm		1		1
640	FBgn0028327	l(1)G0320	no signal	n.a.				
651	FBgn0022153	l(2)k05819	no signal	n.a.				
655	FBgn0267967	corolla	germ cells	oocyte nucleus				1
674	FBgn0003310	S	germ cells	cortical / cell membrane			1	1
681	FBgn0261800	LanB1	epithelial cells	extracellular				1
685	FBgn0052412	QC	germ cells and epithelial cells	nuclear	1			1
691	FBgn0033769	CG8768	germ cells	perinuclear				1
693	FBgn0259243	Pka-R1	dividing germ cells	cytoplasm		1		1
699	FBgn0050404	Tango11	dividing germ cells	cytoplasm		1		1
700	FBgn0002962	nos	germ cells	uniform				1
705	FBgn0024273	WASp	germ cells and epithelial cells	uniform in all cells				
713	FBgn0021796	Tor	germ cells and epithelial cells	cytoplasm		1		
722	FBgn0016070	smg	germ cells	oocyte enriched				1
733	FBgn0030719	eIF5	germ cells and epithelial cells	uniform in all cells				
735	FBgn0263108	BtbVII	germ cells and epithelial cells	nuclear	1			1
739	FBgn0035761	RhoGEF4	no signal	n.a.				
921	FBgn0001986	l(2)35Df	germ cells and epithelial cells	nuclear	1			1

927	FBgn0035149	MED30	germ cells and epithelial cells	uniform in all cells				
937	FBgn0015777	nrv2	epithelial cells	lateral cell membrane			1	1
958	FBgn0003888	betaTub60D	epithelial cells	cytoplasm		1		1
960	FBgn0001137	grk	germ cells	oocyte enriched / anterior				1
1394	FBgn0003015	osk	germ cells	oocyte enriched / posterior				1
10013	FBgn0031913	CG5958	subset of epithelial cells: sqamous cells	cytoplasm		1		1
10044	FBgn0005386	ash1	germ cells and epithelial cells	nuclear	1			1
10070	FBgn0003317	sax	germ cells and epithelial cells	cortical / cell membrane			1	1
10082	FBgn0026192	par-6	germ cells and epithelial cells	uniform in all cells				
10143	FBgn0011020	Sas-4	germ cells and epithelial cells	cytoplasm		1		
10152	FBgn0037526	cg10092	germ cells and epithelial cells	in nurse cells: perinuclear				1
10195	FBgn0266671	Sec6	germ cells and epithelial cells	in epithelial cells: cortical / cell membrane			1	1
10202	FBgn0266673	sec10	germ cells and epithelial cells	in epithelial cells: cortical / cell membrane			1	1
10206	FBgn0025700	CG5885	epithelial cells	cytoplasm		1		1

Supplementary Table 4

fTRG number	FBgn_id	Gene symbol	Signal	Embryonic expression	Movie	bds
58	FBgn0001148	gsb	strong	tissue specific expression	Yes	Yes
71	FBgn0003448	snail	weak	tissue specific expression	Yes	Yes
88	FBgn0025360	Optix	medium	tissue specific expression	Yes	Yes
94	FBgn0010433	ato	weak	tissue specific expression	Yes	Yes
137	FBgn0259685	crb	medium	tissue specific expression	Yes	Yes
155	FBgn0029123	SoxN	strong	tissue specific expression	Yes	Yes
349	FBgn0024294	spn43Aa	strong	late expression, deposited in the cuticle	Yes	Yes
513	FBgn0001147	gsb-n	medium	tissue specific expression	Yes	Yes
937	FBgn0015777	nrv2	strong	ubiquitous expression, membrane signal	Yes	Yes

Sarov et al. Supplementary Table 5

Transformant	Gene	Isoforms	Clone_ID	expression levels thorax	IFM	IFM localisation	leg muscle	leg muscle localisation	tendons IFM	tendon localisation	visceral muscle	visceral muscle localisation	gut epidermis	gut epidermis localisation	trachea IFM	trachea localisation	motorneurons IFM	ventral ganglion	ganglion localisation
1	**mfas**		033559757814831 89 A02	low	no	none	no	none	no	none	no	none	no	none	YES	cell membrane	no	YES	n.a.
21	**gro**		033559757814831 89 B08	intermediate	YES	nucleus	YES	nucleus	n.a.	n.a.	YES	nucleus	YES	nucleus	YES	nucleus	n.a.	YES	nucleus
25	**numb**		033559757814831 89 B09	intermediate	no	none	no	none	YES	cell membrane	no	none	YES	cell membrane	YES	cell membrane	no	YES	cell membrane
27	**ro**		033559757814831 89 D05	very low	no	none	no	none	no	none	no	none	YES	nucleus	no	none	no	n.a.	n.a.
29	**dl**		033559757814831 89 C02	low	no	none	YES	nucleus	n.a.	n.a.	no	none	YES	cytoplasm	no	none	no	no	none
31	**Rho1**		033559757814831 89 C06	intermediate	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	n.a.	n.a.	n.a.	n.a.	YES	cytoplasm	YES	dotty structure - vesicles?	no	n.a.	n.a.
32	**BM-40-SPARC**		033559757814831 89 C12	high	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	YES	extracellular matrix
34	**Mnt**		8431894728673938 B08	intermediate	YES	nucleus	YES	nucleus	n.a.	n.a.	YES	nucleus	YES	nucleus	n.a.	n.a.	n.a.	n.a.	n.a.
39	**sens**		033559757814831 89 G09	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
53	**pyd3**		033559757814831 89 D12	intermediate	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	cytoplasm	YES	cytoplasm	YES	cytoplasm	n.a.	n.a.	YES	YES	cytoplasm
57	**CG5625**		8431894728673938 A07	high	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	YES	dotty structure - vesicles?
59	**Tsc1**		8431894728673938 E01	low	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	n.a.	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	n.a.	n.a.	n.a.	YES	n.a.
63	**CG31772**		6290259782156287 A05	intermediate	YES	Z-disc	YES	Z-disc	no	none	YES	Z-disc	YES	n.a.	no	none	no	YES	cytoplasm
74	**Sc2**		6290259782156287 E03	intermediate	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	n.a.	no	none	YES	cytoplasm	YES	cytoplasm	n.a.	YES	cytoplasm
76	**CG1021**		6290259782156287 B04	intermediate	no	none	no	none	n.a.	n.a.	no	none	YES	cell membrane	YES	cell membrane	no	YES	n.a.
78	**Act88F**		6290259782156287 D01	high	YES	thin filament	no	none	no	none	no	none	no	none	no	none	no	no	none
84	**p53**		8431894728673938 D12	low	YES	nucleus	YES	nucleus	n.a.	n.a.	YES	nucleus	YES	nucleus	n.a.	n.a.	n.a.	YES	nucleus
87	**Blimp-1**		033559757814831 89 G04	low	no	none	YES	nucleus	no	none	no	none	YES	nucleus	no	none	no	no	none
91	**fj**		8431894728673938 H07	very low	no	none	no	none	YES	n.a.	no	none	YES	cytoplasm	no	none	no	n.a.	n.a.
92	**CG32121**		8431894728673938 A05	high	YES	nucleus	YES	nucleus	no	none	no	none	no	none	no	none	no	n.a.	n.a.
96	**GluRIIC**		033559757814831 89 C07	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
103	**CG9336**		033559757814831 89 D07	high	no	none	no	none	YES	cell membrane	no	none	YES	cell membrane	YES	cell membrane	YES	n.a.	n.a.
105	**uzip**		033559757814831 89 D01	intermediate	no	none	no	none	no	none	no	none	no	none	no	none	YES	YES	cell membrane
108	**sty**		6290259782156287 C03	intermediate	no	none	no	none	no	none	no	none	YES	dotty structure - vesicles?	no	none	no	no	none
115	**E2f2**		6290259782156287 F02	intermediate	YES	nucleus	YES	nucleus	YES	nucleus	YES	nucleus	YES	nucleus	n.a.	n.a.	n.a.	YES	nucleus
117	**CG2790**		6290259782156287 G01	none	no	none	no	none	no	none	no	none	no	none	no	none	no	no	none
125	**fray**		6290259782156287 H03	high	YES	myofibril	YES	myofibril	no	none	no	none	YES	cytoplasm	no	none	YES	YES	cytoplasm
129	**par-6**		6745694817474591 E03	intermediate	no	none	no	none	YES	cell membrane	no	none	YES	apical membrane	YES	cell membrane	no	n.a.	n.a.
134	**eIF-4E**		033559757814831 89 D09	intermediate	YES	cytoplasm	YES	cytoplasm	YES	cytoplasm	YES	cytoplasm	YES	cytoplasm	YES	cytoplasm	YES	YES	cytoplasm
137	**crb**		6745694817474591 C04	intermediate	no	none	no	none	YES	cell membrane	no	none	YES	apical membrane	YES	cell membrane	no	n.a.	n.a.
139	**hb**		6745694817474591 H05	intermediate	no	none	YES	nucleus	no	none	YES	nucleus	no	none	n.a.	n.a.	n.a.	n.a.	n.a.
142	**aop**		033559757814831 89 F06	low	no	none	no	none	no	none	YES	nucleus	YES	nucleus	no	none	no	n.a.	n.a.
144	**pain**		24648923033663195 G03	intermediate	no	none	no	none	YES	cell membrane	no	none	YES	cell membrane	no	none	YES	YES	n.a.
145	**Mmp1**		24648923033663195 G11	intermediate	no	none	no	none	no	none	no	none	no	none	YES	extracellular matrix	no	n.a.	n.a.
147	**CG15715**		8431894728673938 F05	low	no	none	no	none	no	none	no	none	YES	cytoplasm	no	none	no	n.a.	n.a.
150	**tan**		5827172721875329 A12	very low	no	none	YES	nucleus	no	none	YES	n.a.	YES	cytoplasm	no	none	YES	YES	cytoplasm
151	**orb2**		5827172721875329 A10	low	no	none	no	none	no	none	no	none	YES	cytoplasm	no	none	no	YES	n.a.
156	**CLIP-190**		8431894728673938 B09	intermediate	no	none	YES	dotty structure - vesicles?	no	none	no	none	YES	dotty structure - vesicles?	no	none	no	YES	n.a.
158	**ergic53**		24648923033663195 H11	intermediate	no	none	no	none	no	none	no	none	YES	dotty structure - vesicles?	no	none	no	no	none
161	**myoglianin**		6196214900843875 E05	very low	no	none	no	none	no	none	no	none	YES	nucleus	no	none	no	n.a.	n.a.
165	**salm**		6196214900843875 H05	very low	YES	nucleus	YES	nucleus	no	none	no	none	no	none	n.a.	n.a.	n.a.	no	none
166	**dei**		066568815868833885 H04	very low	no	none	YES	nucleus	no	none	no	none	no	none	no	none	no	n.a.	n.a.
182	**vri**		6290259782156287 C04	intermediate	YES	nucleus	YES	nucleus	n.a.	n.a.	YES	nucleus	YES	nucleus	n.a.	n.a.	n.a.	YES	nucleus
220	**scrt**		014526891357111271 F02	low	no	none	no	none	no	none	no	none	no	none	n.a.	n.a.	n.a.	YES	nucleus
276	**CG12118**		7778023793434855 E04	high	YES	mitochondria	YES	mitochondria	n.a.	n.a.	YES	mitochondria	YES	mitochondria	n.a.	n.a.	n.a.	YES	mitochondria
286	**CG9312**		7778023793434855 D08	very low	no	none	YES	dotty structure - vesicles?	no	none	no	none	YES	nucleus	no	none	no	n.a.	n.a.
288	**CG7182**		7778023793434855 H08	low	no	none	no	none	no	none	no	none	YES	cytoplasm	no	n.a.	YES	YES	cytoplasm
290	**Pp1alpha-96A**		7778023793434855 F11	intermediate	no	none	no	none	YES	cytoplasm	no	none	YES	cytoplasm	no	none	YES	YES	cytoplasm
444	**babo**		014526891357111271 A07	intermediate	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	cell membrane	YES	cell membrane	YES	cell membrane	YES	cell membrane	YES	YES	cell membrane
448	**Pdk1**		7144464337096436 H02	low	no	none	no	none	no	none	YES	nucleus	YES	cytoplasm	no	none	no	n.a.	n.a.
449	**Hml**		7778023793434855 E09	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
453	**l(1)G0230**		7778023793434855 E06	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
455	**Tim8**		7144464337096436 A11	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
456	**chico**		7778023793434855 H07	intermediate	no	none	no	none	no	none	no	none	YES	cell membrane	no	none	no	no	none
463	**CG32667**		8431894728673938 F03	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
467	**flil**		7778023793434855 E07	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
471	**Mf**	Iso L, O, M	6744927827587379 B02	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
475	**Prm**	Iso A, B, C, D	6744927827587379 D01	high	YES	M-line	YES	thick filament	no	none	YES	thick filament	no	none	no	none	no	no	none
478	**Fhos**		6744927827587379 D02	none	no	none	no	none	no	none	no	none	no	none	no	none	no	no	none
481	**Mp20**		17337204868928058 B04	intermediate	no	none	YES	cytoplasm	no	none	YES	cytoplasm	no	none	no	none	no	n.a.	n.a.
483	**Ilk**		8921672467940407 B06	intermediate	YES	muscle attachment site	YES	muscle attachment site	YES	n.a.	n.a.	n.a.	YES	n.a.	no	none	YES	n.a.	n.a.
484	**TpnC4**		7778023793434855 B04	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
485	**org-1**		1582528031925018 E07	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
486	**mask**		23029740539567312 D06	intermediate	YES	Z-disc	YES	dotty structure - vesicles?	no	none	no	none	YES	nucleus	no	none	no	YES	cytoplasm
488	**rho**		014526891357111271 C02	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
492	**eya**		9077719496880051 C06	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
495	**sas**		014526891357111271 E08	none	no	none	no	none	no	none	no	none	no	none	no	none	no	no	none
500	**Mhc**	Iso K, L, M	6744927827587379 G01	high	YES	thick filament	YES	thick filament	no	none	YES	thick filament	no	none	no	none	no	no	none
501	**Mf**	Iso A, G, N	6744927827587379 A02	high	YES	thick filament	YES	thick filament	no	none	YES	thick filament	no	none	no	none	no	no	none
502	**dlg1**		6744927827587379 C01	high	YES	T-tubules, cell membrane	YES	cell membrane	YES	n.a.	YES	cell membrane	YES	cell membrane	no	none	YES	YES	n.a.
514	**Pink1**		033559757814831 89 C05	none	no	none	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
516	**cact**		23029740539567312 E01	intermediate	YES	neuromuscular junction	YES	neuromuscular junction	YES	cytoplasm	YES	neuromuscular junction	YES	cytoplasm	no	none	YES	n.a.	n.a.

519	**Mhc**	Iso A, F, G etc.	17337204868928058 C04	high	no	none	YES	thick filament	no	none	YES	thick filament	no	none	no	none	no	n.a.	n.a.
565	**Pabp2**		17080591458582148 H03	low	no	none	no	none	YES	n.a.	YES	nucleus	YES	nucleus	no	none	no	YES	nucleus
568	**Cdc42**		1582528031925018 H07	intermediate	no	none	no	none	YES	cell membrane	no	none	YES	cell membrane	YES	cell membrane	YES	YES	cell membrane
569	**Kettin / sls**		6744927827587379 D03	high	YES	Z-disc	YES	Z-disc	no	none	YES	Z-disc	no	none	no	none	no	n.a.	n.a.
570	**Adar**		6744927827587379 C03	intermediate	YES	nucleus	YES	nucleus	no	none	YES	nucleus	YES	nucleus	no	none	no	YES	nucleus
574	**LanA**		6196214900843875 D06	high	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	n.a.	n.a.
576	**Med**		06656881586833885 B11	low	no	none	no	none	YES	cytoplasm	no	none	YES	cytoplasm	no	none	no	YES	cytoplasm
579	**Lis-1**		7778023793434855 G01	low	no	none	no	none	no	none	no	none	YES	dotty structure - vesicles?	no	none	YES	YES	n.a.
584	**Lmpt**		6744927827587379 E02	intermediate	no	none	YES	I-band	YES	cytoplasm	YES	I-band	no	none	no	none	no	n.a.	n.a.
585	**hts**		6744927827587379 B03	high	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	cell membrane	no	none	YES	cell membrane	no	none	YES	YES	n.a.
587	**rhea**	Iso B, E, F	6744927827587379 G02	high	YES	muscle attachment site	YES	muscle attachment site	YES	muscle attachment site	YES	Z-disc	YES	cell membrane	YES	cell membrane	YES	YES	cell membrane
595	**vkg**		98545048033824267 E08	low	no	none	no	none	no	none	no	none	YES	extracellular matrix	no	none	no	YES	extracellular matrix
600	**sqh**		6290259782156287 H07	intermediate	no	none	no	none	YES	cell membrane	no	none	YES	cell membrane	YES	cell membrane	YES	YES	n.a.
602	**Tango1**		8921672467940407 E01	intermediate	no	none	no	none	no	none	no	none	YES	dotty structure - vesicles?	no	none	no	YES	dotty structure - vesicles?
678	**Mlp84B**		9077719496880051 B07	intermediate	no	none	YES	Z-disc	no	none	YES	Z-disc	no	none	no	none	no	n.a.	n.a.
679	**mib2**		24648923033663195 A11	none	no	none	no	none	no	none	no	none	no	none	no	none	no	no	none
680	**zormin**		6744927827587379 C02	none	no	none	no	none	no	none	no	none	no	none	no	none	no	no	none
681	**LanB1**		23029740539567312 H12	high	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	extracellular matrix	YES	n.a.	n.a.
699	**Tango11**		98545048033824267 F04	intermediate	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	n.a.	YES	dotty structure - vesicles?
703	**Lmpt**		6744927827587379 F02	intermediate	YES	cytoplasm	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
709	**Mlp60A**	Iso E, F	6744927827587379 F03	low	no	none	YES	Z-disc	no	none	YES	Z-disc	no	none	no	none	no	n.a.	n.a.
731	**rhea**	Iso C, D	6744927827587379 H02	intermediate	no	none	YES	Z-disc	no	none	YES	Z-disc	no	none	YES	cell membrane	no	no	none
876	**fln**		8921672467940407 D08	intermediate	YES	thick filament	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
925	**wupA**		2733665080679012 B04	intermediate	no	none	YES	thin filament	no	none	YES	thin filament	no	none	no	none	no	n.a.	n.a.
932	**mys**		5827172721875329 D11	intermediate	YES	muscle attachment site	YES	muscle attachment site	YES	cell membrane	no	none	YES	cell membrane	no	none	no	n.a.	n.a.
1046	**Unc-89**		6536462692990512 G05	intermediate	YES	M-line	YES	M-line	no	none	YES	M-line	no	none	no	none	no	n.a.	n.a.
1257	**TpnC25D**		8332520144274483 C10	intermediate	no	none	YES	thin filament	no	none	YES	thin filament	no	none	no	none	no	n.a.	n.a.
10010	**Hsp83**		6329535328138883 A04	high	YES	myofibril	YES	myofibril	YES	cytoplasm	YES	cytoplasm	YES	cytoplasm	YES	cytoplasm	YES	n.a.	n.a.
10013	**CG5958**		4562672055690222 B10	none	no	none	no	none	no	none	no	none	no	none	n.a.	n.a.	n.a.	no	none
10017	**CG5885**		9180222713131253 D02	intermediate	no	none	YES	dotty structure - vesicles?	no	none	no	none	YES	cytoplasm	no	none	YES	YES	cytoplasm
10028	**Act88F**		9180222713131253 D01	high	YES	thin filament	no	none	no	none	no	none	no	none	no	none	no	n.a.	n.a.
10032	**fray**		9180222713131253 H03	high	YES	myofibril	YES	myofibril	no	none	no	none	YES	cytoplasm	no	none	YES	YES	cytoplasm
10035	**CG17912**		9180222713131253 G06	high	YES	nucleus	YES	nucleus	YES	nucleus	YES	nucleus	YES	nucleus	n.a.	n.a.	n.a.	YES	nucleus
10036	**CG12391**		3015068842848645 A04	high	YES	nucleus	YES	nucleus	YES	nucleus	YES	nucleus	YES	nucleus	n.a.	n.a.	n.a.	YES	nucleus
10039	**Sc2**		9180222713131253 E03	intermediate	YES	dotty structure - vesicles?	YES	dotty structure - vesicles?	YES	n.a.	no	none	YES	cytoplasm	YES	cytoplasm	n.a.	YES	cytoplasm
10041	**CG11617**		6329535328138883 C08	intermediate	YES	nucleus	YES	nucleus	YES	n.a.	YES	nucleus	YES	nucleus	n.a.	n.a.	n.a.	n.a.	n.a.
10059	**CG32121**		3015068842848645 A05	intermediate	YES	nucleus	YES	nucleus	no	none	no	none	no	none	no	none	no	YES	n.a.
10070	**sax**		6329535328138883 D05	high	YES	T-tubules, cell membrane	YES	cell membrane	n.a.	n.a.	YES	cell membrane	YES	cell membrane	no	none	no	n.a.	n.a.
10075	**sqh**		9180222713131253 H07	intermediate	no	none	no	none	YES	cell membrane	no	none	YES	cell membrane	YES	cell membrane	YES	YES	cell membrane
10085	**crb**		6329535328138883 C04	low	no	none	no	none	YES	cell membrane	no	none	YES	apical membrane	YES	cell membrane	no	n.a.	n.a.
10125	**veli**		6329535328138883 H03	high	YES	neuromuscular junction	YES	neuromuscular junction	no	none	no	none	no	none	no	none	YES	YES	cytoplasm
10141	**ebi**		33294093041722905 A06	very low	no	none	no	none	no	none	no	none	YES	nucleus	no	none	no	n.a.	n.a.
10149	**Blimp-1**		33294093041722905 G04	intermediate	YES	nucleus	YES	nucleus	YES	n.a.	YES	nucleus	YES	nucleus	YES	nucleus	n.a.	YES	nucleus
10163	**Hand**		07323567509660489 C02	low	no	none	no	none	no	none	YES	nucleus	no	none	no	none	no	n.a.	n.a.
10174	**CG5625**		7654110682321876 A07	intermediate	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	n.a.	YES	nucleus (nls-GFP)
10199	**CG2157**		7654110682321876 F01	none	no	none	no	none	no	none	no	none	no	none	no	none	no	no	none
10236	**rhea**		20271910681552O2 F10	low	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	YES	n.a.	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	no	none	no	YES	n.a.
10243	**Khc**		20271910681552O2 E09	low	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	YES	n.a.	YES	nucleus (nls-GFP)	YES	nucleus (nls-GFP)	n.a.	n.a.	n.a.	n.a.	n.a.
10251	**CG6509**		1522781026558664 B11	low	YES	dotty structure - vesicles?	no	none	YES	cell membrane	no	none	YES	cell membrane	YES	cell membrane	YES	n.a.	n.a.

Sarov et al. Supplementary Table 6

Adult muscle localisation patterns

thick filament	thin filament / myofibril	M-line	Z-disc	muscle attachment site	T-tubules / sarcolemma	dotty pattern / vesicles (?)	mitochondria	nucleus	neuromuscular junction
Fln (876, IFM)	Act88F (78, 10028, IFM)	Prm (475, IFM)	CG31772 (63)	Ilk (483)	Dlg1(502)	Babo (444)	CG12118 (276)	Adar (570)	Cact (516)
Mf (Iso-A,G, N, 501)	Fray (125, 10032)	Unc-89 (1046)	Kettin (Sls-Isoform, 569)	Talin (*rhea*, Iso-B, E, F, G, 587)	Sax (10070)	CG5885 (10017, leg m.)		Blimp-1 (10149)	Veli (10125)
Mhc (Iso-K, L, M, 500)	Hsp83 (10010)		Lmpt (584, I-band, leg m.)	β-PS Integrin (*mys*, 932)		CLIP-190 (156)		CG11617 (10041)	
Mhc (Iso-A, F, 519, leg m. subset & visc. m.)	TpnC25D (1257, leg m. & visceral m.)		Mask (486, IFM)			Dlg5 (*CG6509*, 10251, IFM)		CG12391 (10036)	
Prm (475, leg m. & visceral m.)	TpnI (*wupA*, 925, leg m. & visceral m.)		Mlp60A (709, leg m. & visceral m.)			Hts (585)		CG17912 (10035)	
			Mlp84B (678, leg m. & visceral m.)			Mask (486, leg m.)		CG32121 (92)	
			Talin (Iso-C, D, 731, leg m. & visceral m.)			Pyd3 (53)		Dorsal (29, leg m.)	
						Rho1 (31)		E2F2 (115)	
						Sc2 (79, 10039)		Gro (21)	
						Tango11 (699)		Hand (10163, visceral m.)	
						Tsc1 (59)		Hb (139, leg m.)	
						Vps35 (*CG5625*, 57)		Mnt (34)	
								P53 (84)	
								Salm (165)	
								Vri (182)	

Sarov et al. Supplementary Table 7