TECHNISCHE UNIVERSITÄT DRESDEN
Fakultät Mathematik und Naturwissenschaften

DISSERTATION

A toolkit for visualization of patterns
of gene expression in live *Drosophila* embryos

vorgelegt von:

Radosław Kamil Ejsmont
geboren am 28. Januar 1983 in Szczecin, Polen

zum
Erlangen des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

Gutachter:

1st reviewer: prof. A. Francis Stewart

2nd reviewer: prof. Hugo J. Bellen

Eingereicht am: _____

Verteidigt am: _____

Declaration

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from 14$^{\text{th}}$ September 2006 to 14$^{\text{th}}$ July 2010 under the supervision of Dr. Pavel Tomancak at the Max Planck Institute of Molecular Cell Biology and Genetics.

I declare that I have not undertaken any previous unsuccessful doctorate proceedings.

I declare that I recognize the doctorate regulations of the Fakultät für Mathematik und Naturwissenschaften of the Technische Universität Dresden.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Die Dissertation wurde von Dr. Pavel Tomancak, Max-Planck-Institut für Molekulare Zellbiologie und Genetik betreut und im Zeitraum vom 14. September 2006 bis 14. Juli 2010 verfasst.

Meine Person betreffend erkläre ich hiermit, dass keine früheren erfolglosen Promotionsverfahren stattgefunden haben.

Ich erkenne die Promotionsordnung der Fakultät für Mathematik und Naturwissenschaften, Technische Universität Dresden an.

Date and signature/Datum und Unterschrift

Acknowledgements

I would like to extend an enormous amount of gratitude to my supervisor and mentor Dr. Pavel Tomancak. His dedicated guidance, encouragement and inspiration were crucial for this thesis. Thank you Pavle for giving me the opportunity to do my Ph.D. thesis in an interesting field and friendly group providing me with an excellent research environment.

I would like to thank all current and former members of the Tomancak Lab: Peter, Mano, Maria, Helena, Karolina, Alex, Asli, Pavel, Stephan, Stephan, Vineeth and Michael. Maria and Kamil, summer student, helped me in making the FlyFos libraries, Pavel helped in maintaining them. Without them this project still would be nowhere. Huge thanks for your time and help! Two great summer students, Nicole, Emilia and Michael, diploma student, decided to spend their time imaging flies with SPIM; another extremely skilled diploma student, Peter showed great devotion in implementing some of my insane ideas – thank you all. Both Stephans provided a strong computer-science support for this project. Thank you guys for teaching be how to speak Java and even bigger thanks for SPIM image processing software! Thanks to Alex and Pavel for critical reading of this manuscript.

This work would not be possible without great help of others. I would like to thank Sylke Winkler and the DNA sequencing facility members: Dorit and Liane, for end-sequencing FlyFos libraries. Great thanks to Misho Sarov and the TransGeneOmics unit members, especially Michi, Kristin and Karo for performing high-throughput recombineering experiments and sharing their equipment. I would like to thank Frank Schnorrer and members of his lab at the Max Planck Institute of Biochemistry in Martinsried for great collaboration on the RNAi rescue with *D. pseudoobscura* fosmids. I thank Ivana Viktorinova, Sven Klose, Prof. Eli Knust and Prof. Gunter Reuter for sharing their data on FlyFos rescue in *D. melanogaster*. I thank Carl Zeiss Microimaging for providing and improving the SPIM microscope. Finally, big thank you to Jan Peychl, Dan White and whole Light Microscopy Facility for keeping the SPIM working.

I thank my TAC members Prof. Petra Schwille and Andy Oates for their constant guidance and inspiration.

I would like to thank all my friends for their support and lot's of fun we had in Dresden. I thank my flatmates, Kuba and Andreas for the great time we had living together. I thank Wolfgang, Zoltan, Aliona and Andrei for inspiration and great discussions we had in front of the Instutute.

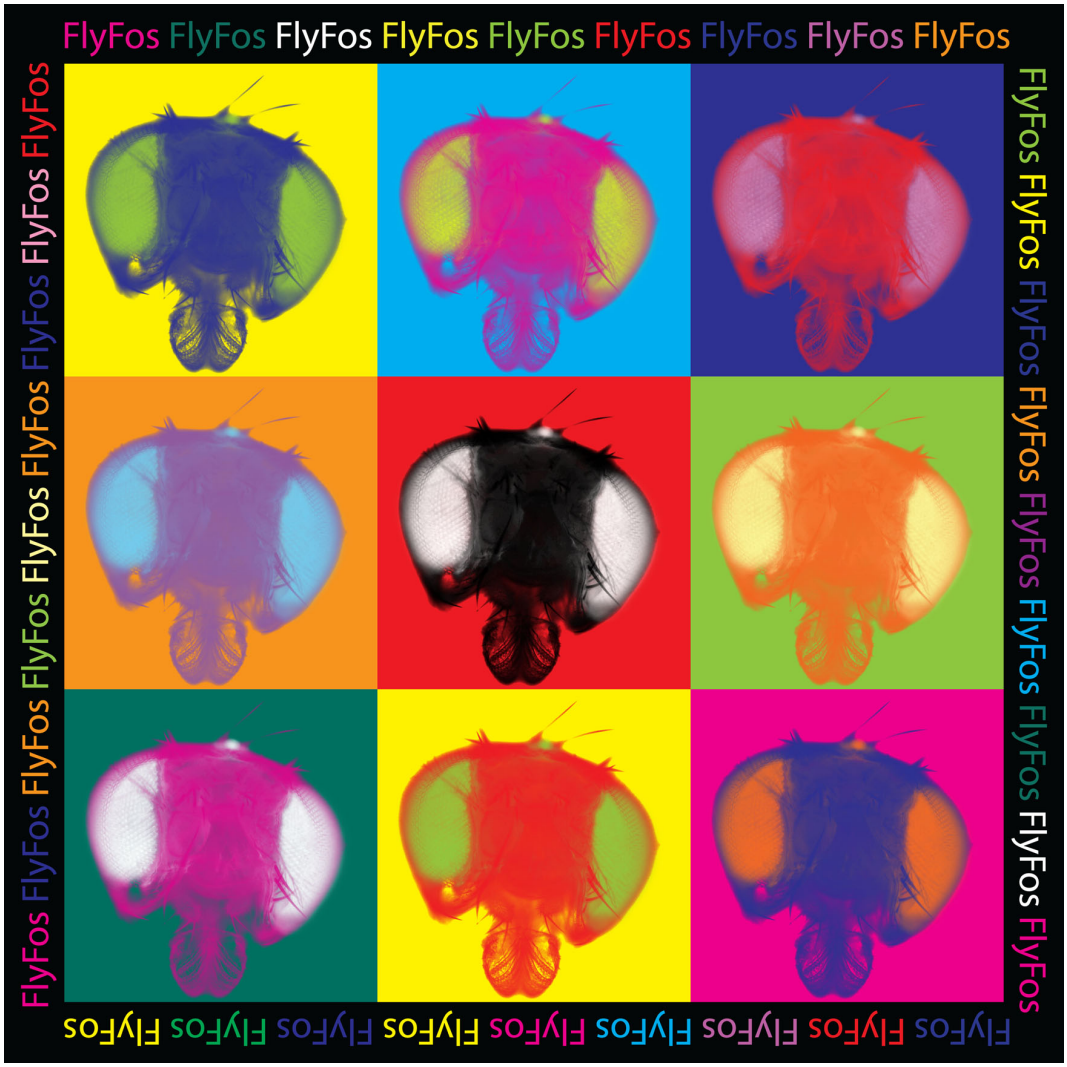
Chciałbym serdecznie podziękować moim Rodzicom za ogromny wkład w moją edukacje, ich cierpliwość, wsparcie i miłość. Dziękuję Wam za to, że zawsze byliście gdy Was potrzebowałem. Dziękuję mojemu bratu za jego uśmiech i energię, które zawsze motywowały mnie do działania.

Finally, I want to express my gratitude and love to someone that plays the leading part in my life. Paulina, without your patience, support and love my life would be less bright and fulfilled.

This work was supported by Human Frontier Science Program grant RGY0084/2008-C. I was supported by Dresden International Graduate School for Biomedicine and Bioengineering PhD stipend.

Abstract

Developing biological systems can be approximately described as complex, three dimensional cellular assemblies that change dramatically across time as a consequence of cell proliferation, differentiation and movements. The presented project aims to overcome problems of limited resolution in both space and time of classical analysis by *in situ* hybridization on fixed tissue. The employment of the newly developed Single Plane Illumination Microscopy (SPIM) combined with new approaches for *in vivo* data acquisition and processing promise to yield high-resolution four-dimensional data of the complete *Drosophila* embryogenesis. We developed a toolkit for high-throughput gene engineering in flies, that provides means for creating faithful *in vivo* reporters of gene expression during *Drosophila melanogaster* development. The cornerstone of the toolkit is a fosmid genomic library enabling high-throughput recombineering and φC31 mediated site-specific transgenesis. The dominant, *3xP3-dsRed* fly selectable marker on the fosmid backbone allows, in principle, transgenesis of the fosmid clones into any non-melanogaster species. In order to extend the capabilities of the gene engineering toolkit to include "evo-devo" studies, we generated genomic fosmid libraries for other sequenced Drosophilidae: *D. virilis*, *D.simulans* and *D. pseudoobscura*. The libraries for these species were constructed in the pFlyFos vector allowing for recombineering modification and φC31 transgenesis of non-melanogaster genomic *loci* into *D. melanogaster*. We have developed a PCR pooling strategy to identify clones for a specific gene from the libraries without extensive clone sequencing and mapping. The clones from these libraries will be primarily used for cross-species gene expression studies. As another application, transgenes originating from closely related species can be used to rescue *D. melanogaster* RNAi phenotypes and establish their specificity. Together with SPIM microscopy, the toolkit will allow to visualize gene expression patterns throughout *Drosophila* development.



"Art and science
have their meeting point in method."

Edward G. Bulwer-Lytton

Table of Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

3D ..... Three-Dimensional

BAC ..... Bacterial Artificial Chromosome

BLAST ..... Basic Local Alignment Search Tool

BLRP ..... Biotin Ligase Recognition Peptide

CCD ..... Charge-Coupled Device

cDNA ..... Complementary DNA

ChIP ..... Chromatin Immunoprecipitation

DNA ..... Deoxyribonucleic Acid

DLSM ..... Digital Scanned Light Sheet Microscopy

dsDNA ..... Double-Stranded DNA

dsRNA ..... Double-Stranded RNA

EGFP ..... Enhanced Green Fluorescent Protein

ECFP ..... Enhanced Cyan Fluorescent Protein

EYFP ..... Enhanced Yellow Fluorescent Protein

FISH ..... Fluorescent *In Situ* Hybridization

FRAP ..... Fluorescence Recovery After Photobleaching

GFP ..... Green Fluorescent Protein

HT ..... High-Throughput

ISH ..... *In Situ* Hybridization

mRNA ..... Messenger RNA

NLS ..... Nuclear Localization Signal

PAC ..... P1-derived Artificial Chromosome

PCR ..... Polymerase Chain Reaction

PMT ..... Photomultiplier Tube

RISC ..... RNA-Induced Silencing Complex

RFLP ..... Restriction Fragment Length Polymorphism

RMCE ..... Recombinase-Mediated Cassette Exchange

RNA ..... Ribonucleic Acid

RNAi ..... RNA Interference

SGFP ..... Superfolder Green Fluorescent Protein

siRNA ..... Small Interfering RNA

SPIM ..... Selective Plane Illumination Microscope/Microscopy

ssDNA ..... Single-Stranded DNA

tRNA ..... Transfer RNA

UAS ..... Upstream Activation Sequence

YAC ..... Yeast Artificial Chromosome

Chapter 1.

1 Introduction

Contents

1.1. Motivation

Animal development can be described as a complex, three-dimensional cellular system that changes dramatically across time as a consequence of cell proliferation, differentiation and movements. These developmental processes are governed by information stored in genomes and differential gene expression is the major mechanism that mediates the realization of genomic information in development. Changes in gene expression are a result of complex interactions on many levels, both inside the cells and among different cells in the organism. Cell autonomous regulatory events on transcriptional, translational and post-translational levels are integrated with inputs resulting from cell-cell communications, forming a huge network that drives cells to their developmental fate. It is clear that to fully understand how genomic information transforms into animal development we need to study the system as a whole (*in toto*). Of course, we currently do not possess the ability to follow all components of the gene regulatory network simultaneously at the molecular level and in the context of the complete developing system. However, we think that emerging reverse genetic and microscopic approaches will allow us to follow developmental events and gene expression regulation in the entire organism at the cellular level.

The *Drosophila* embryo was chosen as the model organism for this project, because it is one of the best characterised developmental systems. During nearly a century of genetic, morphological and molecular studies of the fruit fly's development, both precise embryo anatomy and many gene regulatory networks have been revealed. Availability of full genomic sequence (Adams et al., 2000; Smith et al., 2007) and relatively easy reverse genetic manipulation makes *Drosophila* a perfect organism for genomics studies. The short life cycle of the fruit fly enables results to be gathered in a reasonable time.

Drosophila melanogaster is one of twelve sequenced species in the Drosophilidae group. Since the completition pf the sequencing of twelve Drosophilidae genomes, the comparative genomics era in fruitfly research has begun (Drosophila 12 Genomes Consortium et al., 2007). Expansion of genomic data beyond single species within a phylogenetic group enhanced the quality of *Drosophila melanogaster* genome annotation, enabled discovery of novel regulatory elements (Berman et al., 2004; Kheradpour et al., 2007; Arunachalam et al., 2010) and allowed for the verification of gene regulation evolution theories (Kalinka et al., 2010). Finally, exchange of coding or non-coding genomic elements between closely or distantly related species shed light on evolutionary mechanisms leading to developmental and physiological innovations (Prud'homme et al., 2006). Various applications of cross-genomic data encourage development of tools that can be implemented beyond *Drosophila melanogaster*.

1.2. Capturing the pattern

Developmental processes, and the role of gene expression regulation in them, have been studied for decades using sophisticated microscopy techniques. The gene expression patterns are visible manifestations of constant changes in protein levels in every single cell of the developing embryo. It is standard in developmental biology to draw conclusions about gene expression patterns from subjective assessments of two-dimensional images of fixed and stained biological specimens. The resolution of these images is usually insufficient to distinguish individual cells in the entire three-dimensional specimen. In order to achieve cellular resolution, the microscopy focuses on only a small part of the developing system. The temporal dynamics of changes in gene expression patterns are captured by a sparse sampling of different developmental times in different specimens and the relationship between patterns of gene expression and cellular behavior is often neglected. Levels of gene expression are usually quantified by accompanying *in situ* hybridization (ISH) with microarray analysis. See table 1.1 below for brief summary of some techniques available for description and quantification of gene expression.

Technique	Morphology	Gene expression	Spatial resolution	Temporal resolution	Quantification
Microarray	—	+++	—	+	+++
ISH	++	+++	2D	+	+
FISH	++	+++	3D	+	++
Live imaging	+++	+++	3D	+++	++

Table 1.1.: Comparison of techniques used to describe gene expression

Overview of techniques used to describe gene expression. ISH – *in situ* hybridization, FISH – fluorescent *in situ* hybridization. Description of live imaging patterns assumes use of fluorescent markers for monitoring both morphology and gene expression. (—) – not available, (+) – marginal, (++) – good, (+++) – very good.

1.2.1. Fixed sample approaches

Microarray technology has enabled a quantitative description of gene expression changes in time (Tomancak et al., 2002; Arbeitman et al., 2002; Stolc et al., 2004). While information on gene expression levels in certain stages of development is important for understanding underlying gene function, complete lack of spatial resolution of this technique is usually complemented by classical RNA *in situ* hybridization. A systematic acquisition and annotation of *in situ* expression patterns for over 6,000 *Drosophila* genes was performed in Berkeley, resulting in a comprehensive atlas of

gene expression patterns in embryogenesis (Tomancak et al., 2002, 2007). Expression patterns were described by expert annotators using a controlled vocabulary for embryo anatomy. This standardized, systematic approach allowed relatively easy comparison of expression patterns for different genes and thus, search for those that are co-regulated. The project introduced, to some extent, automated image processing based quantification of gene expression data by applying a triangular mesh to the acquired images and averaging signal over the mesh cell surface (Frise et al., 2010). Numerous approaches have been applied to automate *in situ* data annotation (Peng et al., 2007; Ji et al., 2008; Mace et al., 2010), however wide-field image acquisition limited the spatial resolution of the data and the lack of standards in embryo orientation and depth of focus lead to numerous artifacts during the computational analysis.

Another project aiming to quantify gene expression levels in the *Drosophila* embryo introduced 3D confocal imaging and FISH to overcome spatial resolution problems (Keranen et al., 2006; Luengo Hendriks et al., 2006). Deep sample penetration was achieved by two-photon excitation allowing imaging of nearly the whole embryo. Advanced segmentation algorithms enabled assignment of gene expression levels to single cells (Fowlkes et al., 2008). Thus, this approach resulted in quantitative description of gene expression patterns at cellular resolution at one specific stage of development – cellular blastoderm.

Since both approaches relied on *in situ* hybridization in fixed samples, the tracing of expression patterns in single embryo over time was not achievable. Although some efforts were made to interpolate changes in embryos over time, based on statistical analysis of many samples of different age, such methods did not provide real-time data with high temporal resolution. Recently developed microscopy techniques address this issue.

1.2.2. Selective Plane Illumination Microscopy

Selective Plane Illumination Microscopy (SPIM, Huisken et al., 2004) offers a number of key advantages over other imaging technologies that are also capable of optical sectioning such as confocal microscopes (Carlsson et al., 1985). In SPIM, optical sectioning is achieved by focusing the excitation laser beam into a thin light sheet, which penetrates the living sample embedded in an agarose gel and suspended by gravity in a water-filled chamber. The objective lens is arranged perpendicular to the axis of illumination and the thus laser illuminates only the imaged plane of the specimen (figure 1.1). This minimizes photo-bleaching and laser damage to the living samples and allows very long time-lapse recordings. The images are captured by a

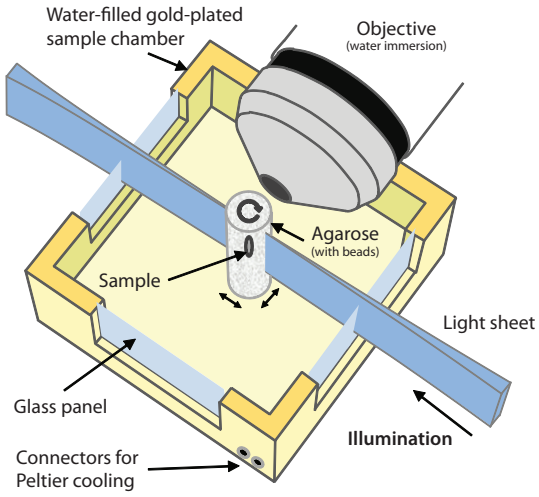


Figure 1.1.: The SPIM microscope

The specimen embedded in agarose is illuminated by a laser light sheet, the CCD camera behind the objective is focused on the center of the light sheet. Optical sectioning is achieved by moving the sample through the light sheet. Multiple acquisitions of the sample can be taken by rotating the agarose column. The water in the sample chamber can be cooled by a Peltier device to slow down developmental processes. Figure courtesy of S. Preibisch (Preibisch et al., 2010).

CCD camera enabling a very fast acquisition rate important for capturing dynamic developmental events in living embryos.

Serial optical sections are recorded by moving the sample through the light sheet. In order to achieve an isotropic resolution uniformly across the sample's volume, it is necessary to rotate the sample and record image stacks for the same specimen from different angles (movie 1). Sample rotation is a unique feature of the SPIM set-up and allows unprecedented flexibility in positioning of large biological specimens for imaging.

The lateral resolution of SPIM-acquired images resembles confocal images, however the axial resolution of raw images is inferior to other optical sectioning techniques, mainly due to physical limitations of light-sheet formation optics. Yet, the ability to acquire image stacks from multiple angles (views) and recently developed image processing techniques allow to reconstruct three dimensional images with isotropic axial and lateral resolution. Preibisch et al. (2008, 2010) described an image processing framework for multiview registration of SPIM-acquired images. In this approach, the samples are embedded in agarose containing subpixel-sized fluorescent beads, which are used as fiduciary markers for sample-independent and fully unguided stack registration (figure 1.2). Together with a content-based image fusion algorithm (Preibisch et al., 2007) this technique provides a comprehensive framework for 3D reconstruction of multiview images acquired with the SPIM.

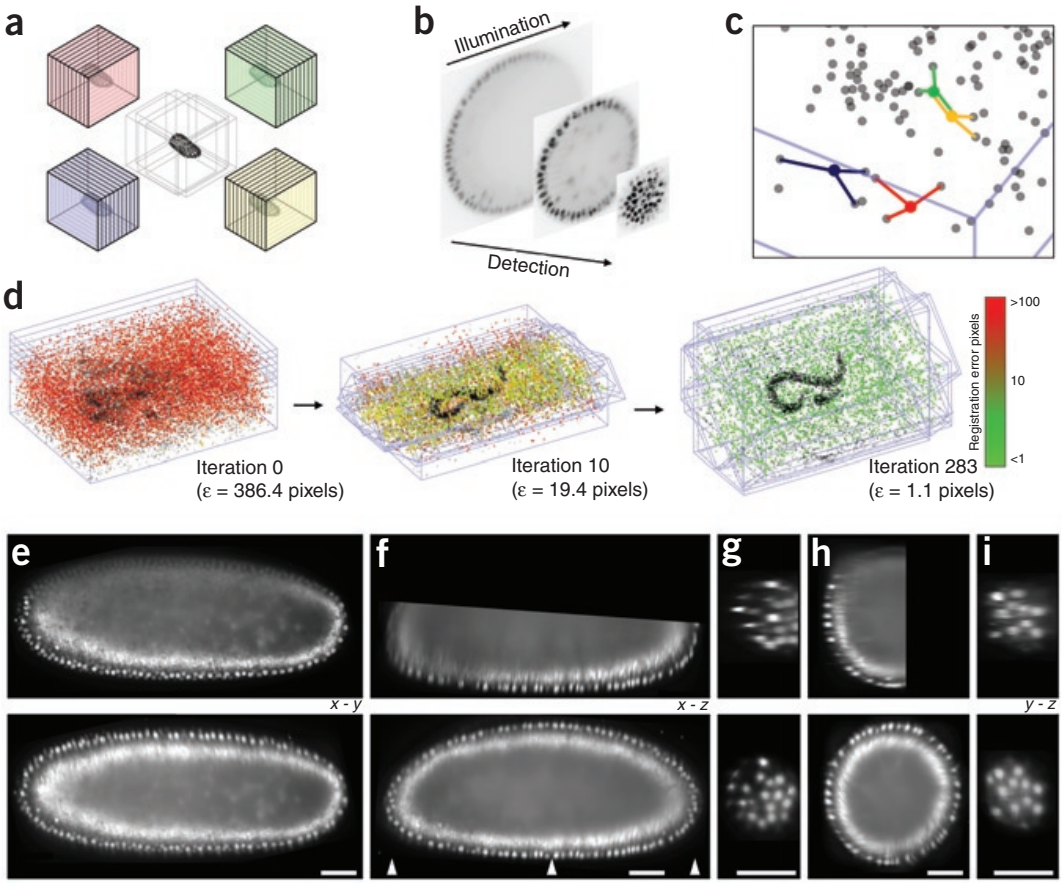


Figure 1.2.: Bead-based registration framework

(a) Several stacks of two-dimensional images of the same specimen acquired from different views have to be registered to obtain a single 3D image. (b) Three SPIM sections of *Drosophila* embryo stained with nuclear marker show the deterioration of the fluorescence signal along the illumination and detection axes. (c) Four color-coded examples of 3D constellations of four beads (central bead and its three nearest neighbors forming a bead descriptor) used to identify corresponding beads in different views (blue lines show view boundary in three dimensions, and gray circles represent the beads). (d) A 3D visualization of the global optimization progress on eight SPIM views of fixed *Caenorhabditis elegans* worm. Displacement of corresponding bead descriptors is color-coded from red (maximum displacement) to green (minimal displacement). The global optimization is initialized with all views on top of each other. Three iterations (0, 10 and 283) are shown along with average displacement across all descriptors. (e–i) Sections through living *Drosophila* embryo expressing *His-EYFP* in all cells; imaged and reconstructed from seven SPIM views (bottom) compared to single SPIM view (top). Single-view acquisitions were stopped approximately in the middle of the embryo to avoid optical aberrations resulting from light scattering and to speed up the acquisition. The lateral resolution in the reconstructed multiview image (e) is comparable with the axial resolution (f) and is superior to the resolution of the single view. y-z sections (g–i) at positions marked by white arrowheads in f. Scale bars, 50 μm. Figure courtesy of S. Preibisch (Preibisch et al., 2010).

1.2.3. Live imaging markers

While imaging fixed samples allows the use of classical staining techniques, live imaging of whole-mount embryos requires fluorescent markers, best originating from the sample itself. Nuclear markers are the well-established standard for highlighting cells inside various tissues, and thus providing information about the embryo's cellular morphology. Fluorescent proteins are suitable for long term *in vivo* imaging during development, because they are continually supplied into all cells when placed under the control of a general promoter. There are two major kinds of fluorescent proteins that specifically mark cell nuclei – those with nuclear localization signal (*NLS*) and fusion proteins between fluorescent marker and components of general nuclear machinery. Davis et al. (1995) used a polyubiquitin promoter to drive expression of green fluorescent protein (*GFP*) fused to the nuclear localization signal. They were able to visualize the cell nuclei throughout development using a wide-field fluorescent microscopy. Clarkson and Saint (1999) fused *GFP* to *Histone 2AvD*, a variant histone of the *H2A* family, which has a general distribution within the chromatin. The *His2AvD-GFP* fusion was expressed under a native promoter and could rescue *His2AvD* mutants. The *GFP* fluorescence was strong, making it possible to observe all phases of mitosis readily using epifluorescence microscopy with living dechorionated embryos. The histone-based nuclear marker, as opposed to *GFP-NLS* fusion is visible in cells throughout the whole cell cycle, enabling cell tracking.

Expressing a membrane-specific marker in imaged embryos may aid in monitoring cellular behaviour in development. Oda and Tsukita (2001) constructed a *ubi-DE-cad-GFP* transgene to visualize cell-cell adherens junctions during mesoderm invagination in living embryos. They cloned the *shotgun* gene under control of the *ubiquitin* promoter. The transgene was able to clearly visualize cell-cell adherens junctions, which formed three-dimensional networks linking apical edges of epithelial cells, in blastoderm cells and later epithelial cells. The construct was also capable of rescuing a *shg* null mutation.

Live imaging reporters for various cell classes and subcellular components, like the aforementioned markers, are broadly available to the research community. They provide information about cellular behavior during development, and can usually be modified to fit specific imaging needs. With a wide range of fluorescent proteins created within the past several years (Nagai et al., 2002; Rizzo et al., 2004; Shaner et al., 2004; Pédelacq et al., 2006; Merzlyak et al., 2007) simultaneous acquisition of multiple markers became possible. Despite these achievements, the large scale visualization of gene expression using live fluorescent reporters still remains a challenge. Techniques and resources that emerged in the past few years have the potential to dramatically expand the possibilities for reverse genetic manipulation of the

Drosophila genome and provide means for generation of fluorescent reporters for every fruitfly gene.

1.3. Gene activity reporters

Classical forward genetic mutagenesis screens pioneered the understanding of animal development in particular by using *Drosophila* as a model system (Nüsslein-Volhard and Wieschaus, 1980). Reverse genetic approaches increasingly complement these traditional ways of studying gene function in development. Transposable elements, like the P-element, became the first tools employed to modify the *Drosophila* genome (Rubin and Spradling, 1982). Traditionally, P-element transposition was used in insertional mutagenesis (Cooley et al., 1988). With the development of transformation vectors and cloning techniques, P-elements were used to reintroduce modified genes into flies. The power of P-element transposition was used to generate expanding collections of gene disruptions (Spradling et al., 1995) and chromosomal deficiencies (Ryder et al., 2007). Coupling P-element transgenesis with the *GAL4-UAS* system (Brand and Perrimon, 1993) unleashed tissue specific gene expression. Large collections of *GAL4* drivers (Manseau et al., 1997) and cDNAs (Stapleton et al., 2002) enable description, as well as loss and gain of function analysis of gene activity at various stages of the *Drosophila* life cycle. Finally, expression of fluorescent proteins under the transcriptional control of the yeast upstream activating sequence that is recognized by *GAL4* provided a fluorescent reporter for monitoring gene expression in living embryos in a tissue specific or temporarly triggered manner (Yeh et al., 1995).

Expressing fluorescent proteins using the *GAL4-UAS* system can provide information about gene expression patterns, however using this technique leads to overexpression of the target construct, and therefore renders quantification of gene expression levels impossible. Moreover, the *GAL4* enhancer traps rely on transcriptional gene expression regulation only. Posttranscriptional control of gene expression plays an important role in the expression pattern formation (Gaul et al., 1987). The only way to enable posttranscriptional control over reporters is to express them as fusions with the target gene in its genomic context. Such a result can be achieved using P-element transposition to generate protein traps with an artificial exon encoding a *GFP* reporter (Morin et al., 2001). Protein traps and enchancer traps allow the expression of reporters natively, however introduction of a transposable element into the genome is random, and therefore genome-wide applications of these techniques is limited. Fluorescent embryo sorting used for selection of the successful protein traps provides strong selection, but due to its limited sensitivity, results in similar insertions

being rediscovered over and over. Moreover, without precise control of insertion location, such approaches often result in disruption of native protein function or even lead to its degradation.

The P-element transposition over years of its use in *Drosophila* reseach proved to be powerful tool, yet integration of P-elements into the fly genome was shown to be biased (Liao et al., 2000; Aleksic et al., 2009), thereby limiting its genome-wide applications. Using other transposons, such as Minos (Loukeris et al., 1995a) or piggyBac (Handler and Harrell, 1999) aided in expanding the coverage of transposon insertions (Bellen et al., 2004; Quiñones-Coello et al., 2007). The piggyBac and Minos transposons, unlike the P-element, can be used for transposition in a variety of insect species (Lobo et al., 1999; Loukeris et al., 1995b). Despite the fact that transposon-mediated gene targeting is a mighty technique, it is random. Targeting a specific gene using transposon traps is a matter of luck and needs plenty of time and labor.

1.3.1. Genomic DNA libraries

Genomic DNA libraries have been widely used for gene cloning, physical mapping and more recently also for whole genome sequencing (Adams et al., 2000; Drosophila 12 Genomes Consortium et al., 2007). A whole new range of application for genomic libraries emerged recently in the fields of cell, developmental and evolutionary biology. In these research areas, it is often desirable to monitor the behavior of modified transgenes re-introduced into the genome to assay tissue specific gene expression, sub-cellular protein localization or affinity purification of protein-protein or protein-DNA complexes. Traditional methods use tagged cDNA clones under the control of various tissue specific or inducible promoters, however these reporters typically do not recapitulate the wild-type gene expression specificity of the gene under study nor its expression levels. Unlike in cDNA constructs, large genomic clones can be selected in such a way that they likely include all the regulatory elements required to recapitulate the native gene expression, both qualitatively and quantitatively.

The genomic libraries can be constructed in a variety of vectors, including cosmids (Collins and Hohn, 1978), fosmids (Kim et al., 1992), P1-phage vectors (Sternberg, 1990), BACs (Shizuya et al., 1992) and YACs (Burke et al., 1987). Virtually all enumerated vectors were used for construction of high-coverage genomic libraries of *Drosophila melanogaster* (Sidén-Kiamos et al., 1990; Ajioka et al., 1991; Kimmerly et al., 1996; Osoegawa et al., 2007). See table 1.2 for comparison of the vectors.

Cosmids are bacterial plasmid vectors that harbor λ phage cohesive-end site (*cos*). They can be packaged into phage particles and used for infecting *E. coli*, resulting

Vector	Origin	Host	Delivery	Insert size	Copy number	Stability
cosmid	λ phage	*E. coli*	phage infection	40 – 45 kb	moderate	+
P1	P1 phage	*E. coli*	phage infection	80 – 100 kb	moderate	++
YAC	yeast chromosome	*S. cerevisiae*	chemical transformation	0.1 – 1 Mb	double*	++
BAC	*E. coli* F factor	*E. coli*	electroporation	20 – 300 kb	single	+++
fosmid	*E. coli* F factor / λ phage	*E. coli*	phage infection	35 – 45 kb	single	+++

Table 1.2.: Overview of frequently used library vectors

* – YACs are maintained similarly to yeast chromosomes, therefore are double-copy in diploid cells (this is how yeast strains are maintained) and single-copy in haploid cells.

in bacterial transformation. The insert DNA size varies depending on the vector size, but usually lies around 40–45 kb. The phage packaging system restricts size of the insert DNA, providing a strict size-selection method in the library production. The phage infection-based transformation of cosmid vectors is extremely efficient considering the size of inserts (Collins and Hohn, 1978). The phage P1 vectors or PACs - similarly to cosmids - rely on phage packaging and infection to deliver constructs into bacteria. Larger capacity of phage P1 particles allows to include inserts of up to 100 kb in size, this however comes at a price of more elaborate packaging system. In the first stage, the packaging site (*pac*) in the vector DNA is cleaved by the pacase extract. In the second stage, that DNA is packaged into phage particles. Phage P1 also requires *in vivo Cre*-mediated recombination between *loxP* sites for circularization of the packaged DNA in bacteria. The P1 vector introduced a copy-control system, where high-copy lytic replicon of phage P1 was put under the control of the inducible *lac* promoter (Sternberg, 1990).

Yeast artificial chromosomes are DNA vectors allowing for very large insert sizes, up to a megabase long. A typical YAC vector is a plasmid that contains a yeast centromere, two sequences that seed telomere formation *in vivo* and yeast-selectable markers. The insert DNA is cloned into one of the chromosome arms, splitting a phenotypically visible marker (such as *ade2-ochre* suppressor, *SUP4*). Upon transformation the linear YAC is maintained in yeast similarly to the native chromosomes. The yeast artificial chromosome libraries, despite their successful applications, usually contain a large number of chimeric clones that are the result of *in vivo* recombination in yeast (Green et al., 1991). Moreover, as YACs require yeast for propagation, the

transformant DNA recovery is more complicated and time consuming than in *E. coli* systems.

BACs and fosmids are library vectors based on *E. coli* fertility factor. Bacterial artificial chromosomes are maintained in low copy number (one to two copies per cell), thus reducing the potential for recombination between DNA fragments carried by the vector. They allow to clone and stably maintain DNA fragments of up to 300 kb in size. The BAC vectors include several genes required for maintenance in *E. coli*. The *oriS* and *repE* genes mediate the unidirectional replication of the F factor, while *parA*, *parB* and *parC* maintain copy number at a level of one or two per cell (Shizuya et al., 1992). Fosmids are hybrid vectors based on BACs that include λ phage cohesive-end site (*cos*) derived from cosmids. While *cos* site provides efficient delivery of the construct into bacteria and strict size selection mechanism, the F factor replication ensures stable maintenance in bacterial cells (Kim et al., 1992).

1.3.2. φC31 transgenesis

Large genomic DNA fragments included in BAC or fosmid libraries cannot be integrated into the fly genome by means of transposon-based transformation techniques. The φC31-based, site-specific recombination addresses this issue. The φC31 integrase belongs to the actinophage resolvase/invertase family. It mediates integration of the viral genome into *Streptomyces* bacteria. The integrase recognizes the chromosomal attachment site (*attB*), and the phage-encoded attachment site (*attP*). Recombination results in hybrid sites called *attL* and *attR* (figure 1.3). The recombination catalyzed by φC31 integrase occurs both *in vivo* and *in vitro*, does not depend on DNA supercoiling and is irreversible (Thorpe and Smith, 1998).

The φC31 integrase was successfully used to integrate exogenous DNA into human tissue culture cells (Groth et al., 2000), mice (Olivares et al., 2002) and *Drosophila* (Groth et al., 2004). In the fly site-specific transgenesis system, the *attP* site recognized by the integrase is introduced into the genome by transposition creating the so called landing site. The *attB* site containing plasmid is usually co-injected together with φC31 integrase mRNA into the landing site strain. To circumvent the need of integrase mRNA co-injection, landing lines harboring genome-encoded source of integrase were created (Bischof et al., 2007). The expression of integrase is localized to the posterior pole, where the injection happens. The φC31-mediated transformation is effective even with BAC-sized constructs (Venken et al., 2006) and, in addition to simple exogenous DNA integration, can be used in recombinase-mediated cassette exchange (RMCE, Bateman et al., 2006). The irreversibility of φC31 integrase-catalysed reaction and activity in a variety of distantly-related species make it the current system of choice for fly transformation.

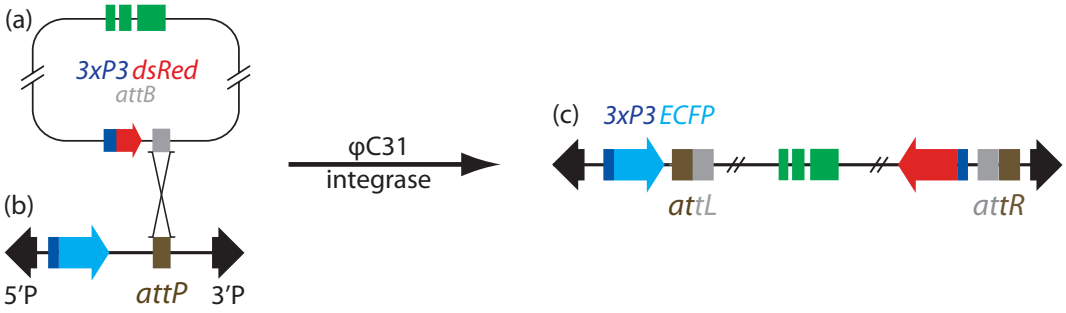


Figure 1.3.: φC31-mediated transgenesis

In this example a vector (a) carying genomic DNA insert and harboring *3xP3-dsRed* as a fly selectable marker is integrated into landing site (b) that was introduced into flies using P-element transgenesis. In presented schema the landing line flies are selected by cyan eye fluorescence comming from *3xP3-ECFP* marker. The φC31 integrase mediates recombination between *attB* site in the vector with *attP* site in the landing line. As a result of irreversible reaction, the vector is inserted into the landing lines and two post-recombination sites, *attL* and *attR*, are formed.

1.3.3. Universal marker for transgenic animal selection

The φC31 integrase system enables species-independent site specific transgenesis with large constructs. The most conventional selectable markers used for selecting transformants in flies are *Drosophila melanogaster*-specific, and therefore restrict possible cross-species applications of the φC31 system. The *3xP3-EGFP* is a fly selectable marker intended to replace *white* in fly transgenesis. The expression of enhanced green fluorescent protein is driven in compound eyes and ocelli by the artificial *3xP3* promoter (Berghammer et al., 1999). The *3xP3* promoter element is bound by three Pax-6 homodimers. The *GFP* fluorescence can be already detected in late (stage 16) embryos, larvae and pupae (Horn et al., 2000). The *3xP3-EGFP* was successfully used for screening piggyBac, Minos and Hermes transposable element insertion events in various insect species (Kokoza et al., 2001; Thomas et al., 2002; Mandrioli and Wimmer, 2003; Marcus et al., 2004; Pavlopoulos et al., 2004). Use of fluorescent proteins other than *EGFP* enables selection in multi-component genetic systems (Horn et al., 2002; figure 1.4).

1.4. Recombineering

Recombineering (recombination-mediated genetic engineering, also known as Red/ET cloning) is a novel technique for DNA engineering using recombination in *Escherichia coli*. Homologous recombination between a linear DNA construct and superhelical target DNA (plasmid, BAC, fosmid, or bacterial genome) is mediated by RecE and

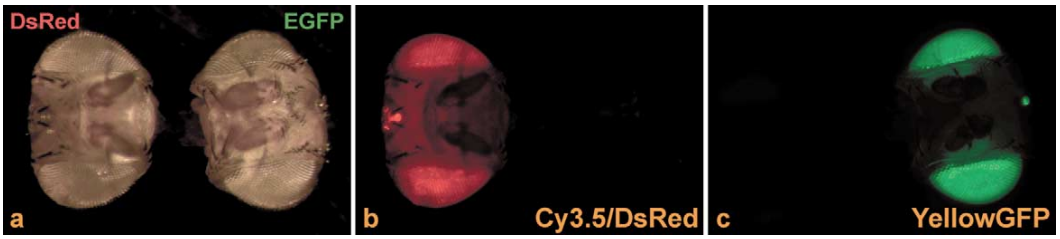


Figure 1.4.: The *3xP3-FP* family of selectable markers

EGFP and *DsRed* serve as distinguishable transformation markers. Comparison of *DsRed* and *EGFP* fluorescence detection using different filter sets. All three panels show the same white-eyed Dm[Bac{3xP3–DsRed}] (left) and Dm[Mos{3xP3–EGFP}] (right) transgenic fly heads. Observations by illumination with a cold light source (a) or with the filter sets Cy3.5/DsRed (b), YellowGFP (c). Figure courtesy of Horn et al. (2002).

RecT in *E. coli sbcA* strain (Zhang et al., 1998) or Redα and Redβ originating from phage λ (Murphy, 1998). In typical recombineering experiment, the linear cassette containing a selectable marker (antibiotic resistance gene) is introduced to bacteria to modify a circular DNA molecule. The recombineering cassette is flanked by ∼50 bp sequence homologous to the target sequence. The selectable marker itself can be flanked by *FRT* or *loxP* sites for its removal by site-specific recombination. Recombinant bacteria are selected on agar medium containing antibiotics that select for both the target plasmid and the recombineering cassette.

The RecE/RecT based recombineering uses a 5'→3' exonuclease encoded by *recE* gene and a single stranded DNA binding protein that also promotes annealing, encoded by *recT* gene. The homologous recombination occurs *via* a double strand break repair mechanism. The ET cloning can be achieved in *recBC+* strains by introduction of a plasmid vector including C-terminal part of *recE* under inducible promoter, constitutively expressed *recT* and *redγ* that inhibits degradation of linear DNA fragments by RecBC complex. Homologous recombination mediated by RecET was shown to work on both plasmids and large (P1-sized) constructs, introducing both insertions and substitutions of the target region with recombineering cassette (Zhang et al., 1998).

The Redα/Redβ recombineering works in a similar manner to RecE/RecT system. The phage λ *reda* encodes a 5'→3' exonuclease, while *redβ* encodes a single stranded DNA binding protein (Murphy, 1998). Homologous recombination with Red*α*β is efficient in *recBC+* strains only when Redγ is introduced (Muyrers et al., 1999). The RecA increases the efficiency of Red/ET recombination several fold, by facilitating cell survival after transformation (Murphy, 1998; Wang et al., 2006). The mechanism underlaying Redαβ recombination with double stranded DNA has been recently described in detail (Maresca et al., 2010; figure 1.5). The recombination requires the

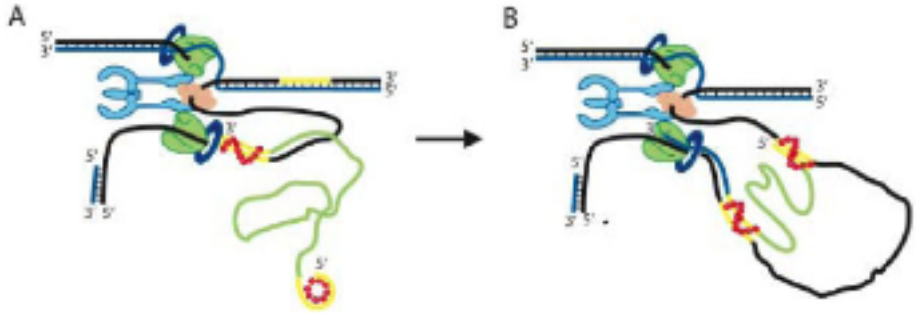


Figure 1.5.: Mechanism of Red/ET-mediated homologous recombination

Model for recombination at the replication fork. Annealing of an ssDNA molecule to complementary regions on the lagging strand at the replication fork is depicted. The ssDNA molecule comprises two flanking homology arms (~50 nt; yellow), interspaced by a heterologous sequence (light green). The Redβ annealing intermediate is shown as a curved line of red dots. The leading strand is shown in blue, lagging strand in black, DnaB helicase in light orange, the two Pol III holoenzymes are green, which are tethered to the γ/τ clamp loader (light blue), and the β sliding clamps are dark blue rings. (A) The Redβ-ssDNA protein complex anneals the 3' end first, which then primes DNA synthesis for an Okazaki fragment. (B) After replication fork progression, the second homology region becomes exposed and annealing of the 5' homology arm creates the ssDNA heteroduplex intermediate. Figure courtesy of M. Maresca and F. Stewart (Maresca et al., 2010).

target circular DNA molecule to initiate its replication. The recombineering cassette is processed by Redα so that one strand is removed completely whilst the other strand remains unresected and contains both homology arms. The single-stranded DNA fragment produced by Redα serves as a primer for lagging strand synthesis during the target DNA replication. Annealing of the fragment is mediated by Redβ. This mechanism, called beta recombination is the main pathway when inserts of up to 3 kb are processed. Recombineering of longer cassettes seems to occur by alternative pathways that do not involve strand preference.

The Red/ET system described above was successfully used for modifications of BACs containing large genomic inserts in the host strain, by introduction of helper plasmids introducing all components of the recombineering machinery (Muyrers et al., 1999; Testa et al., 2003; Wang et al., 2006). Although recombineering is a very powerful method, its application to genome-wide projects was limited by low throughput caused by a need of plating bacteria and screening for recombinants. In liquid culture recombineering the selection for recombineering events is done in liquid medium, without a need for plating bacteria. Such an approach simplifies the recombineering protocol and allows to easily upscale the experiments. Sarov et al. (2006) described a liquid culture recombineering pipeline to tag *Caenorhabditis elegans* genes with green fluorescent protein (figure 1.6). The BAC clones containing the gene of interest were

modified in the host strain. In the first step of the pipeline, bacteria were transformed with pRedFlp – a plasmid containing the Red operon under L-rhamnose promoter and flipase (Buchholz et al., 1998) under anhydrotetracycline promoter. In the second step, expression of the Red operon was induced and bacteria were transformed with recombineering cassette. The third step involved removal of the selectable marker by site-specific recombination between *FRT* sites flanking the selectable marker. In the last step, the BAC vector was retrofitted with a worm-selectable marker (*unc-119*). The recombineering protocol, thanks to all reactions being done in liquid culture, was soon expanded to the 96-well format (Poser et al., 2008).

1.5. RNAi rescue

RNA interference (RNAi) is an RNA-dependent gene-silencing process that is controlled by the RNA-induced silencing complex, called RISC (see figure 1.7). The process is initiated by short double-stranded RNA molecules that when introduced, cause the RNAi machinery to knock down the RNA targets of dsRNA in a sequence-specific manner (Fire et al., 1998). RNAi allows to analyse the loss-of-function phenotype of the genes where mutant alleles are not available or hard to handle. The relative simplicity in achieving knock-down of a specific gene encouraged development of genome-wide RNAi libraries in *Drosophila* (Dietzl et al., 2007; Matsumoto et al., 2007; Ni et al., 2009). Since in these libraries the gene encoding interfering RNAs were put under control of the *Gal4-UAS* binary system, it is possible to achieve knock-down of gene targets in a specific tissue (Roignant et al., 2003). This enabled genome-wide RNAi screens to study organ development (Mummery-Widmer et al., 2009; Schnorrer et al., 2010) and neuronal function (Yapici et al., 2008) in an intact fly.

A major pitfall of any RNAi approach are potential false positives resulting from unspecific knock-down of other genes than the anticipated target, the so called "off-target" effect. In case of randomly inserted hairpin transgenes, false positives may arise from missexpression of neighbouring genes. Despite the relatively low false positive rate in the systematic screens performed thus far (5–7%) (Mummery-Widmer et al., 2009; Schnorrer et al., 2010), its presence necessitates the confirmation of the association of an RNAi phenotype with a particular gene by an independent method. The best proof is the recapitulation of the RNAi phenotype by a classical mutant, however such an approach is not universal as mutants are either not available or may display uninterpretable, pleiotropic phenotypes. Alternatively, the RNAi phenotype can be confirmed by a second hairpin construct targeting a different region of the target gene that should show no or a different off-target effect. However,

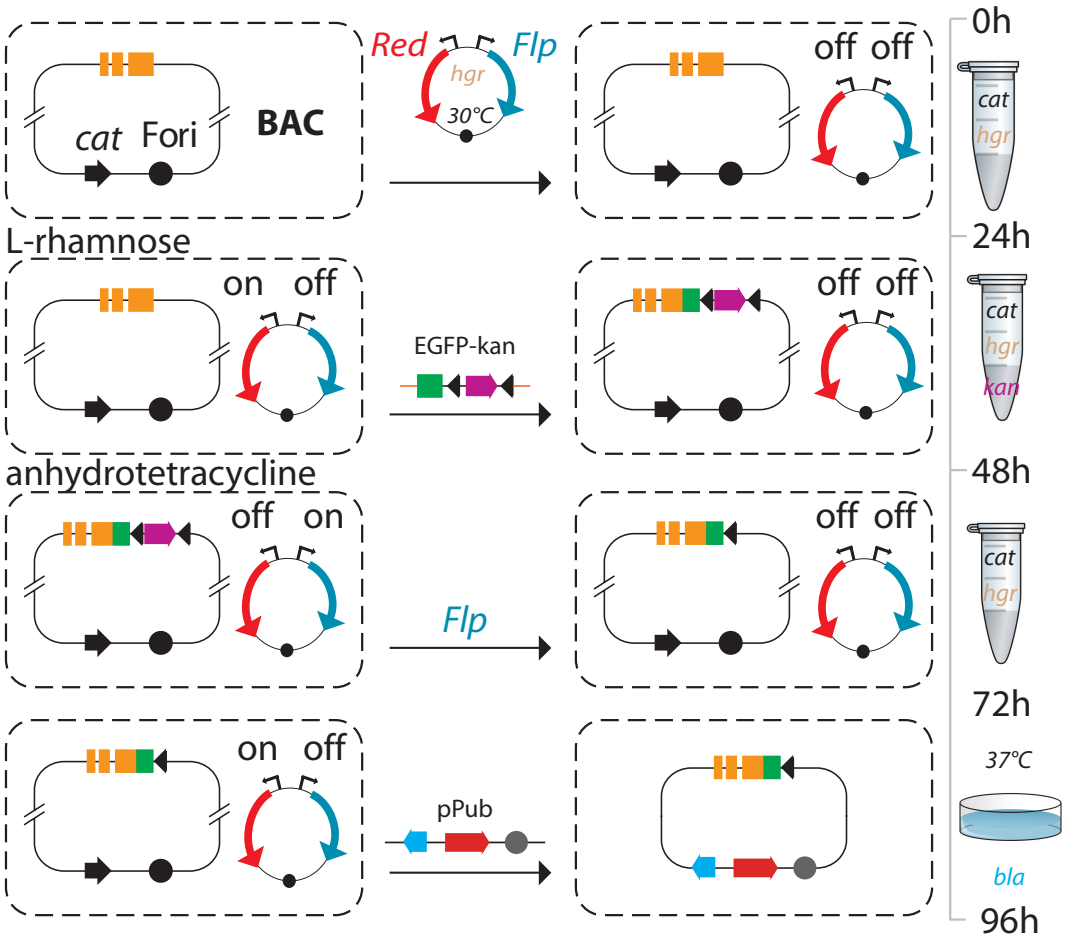


Figure 1.6.: Liquid culture recombineering pipeline for gene tagging in *C. elegans*

A suitable BAC clone for the gene of interest (orange) is chosen and all recombineering steps are done in the original BAC host cells (broken line). (0h–24h) The host is transformed with the dual expression plasmid pRedFlp. Fori, BAC replication origin; *cat*, chloramphenicol resistance gene. (24h–48h) Expression of the Red operon (red) is induced with rhamnose and the cells are then electroporated with the *EGFP-kan* cassette. (48h–72h) Expression of Flp recombinase is induced with anhydrotetracycline. Flp binds to the *FRT* sites (blue triangles) and excises the *kan* gene. (72h–96h) Expression of the Red operon is again induced with rhamnose; the cells are then electroporated with the pPUB subcloning vector, followed by selection for blasticidin, temperature shift to 37°C and omission of all other antibiotics. Figure courtesy of M. Sarov (Sarov et al., 2006).

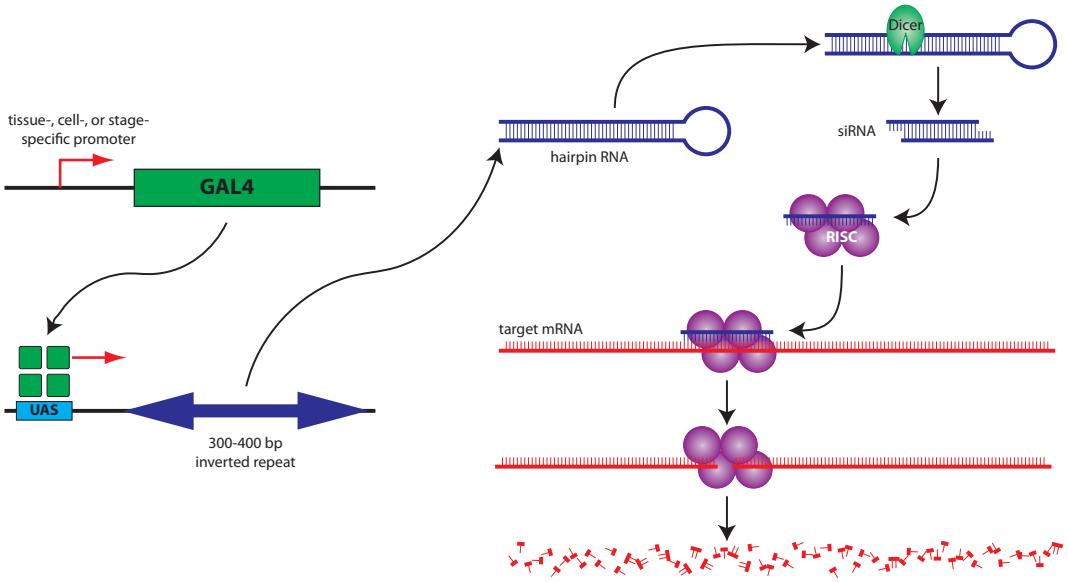


Figure 1.7.: RNAi-induced gene knockdown in *Drosophila*

The *GAL4/UAS* system is used to drive the expression of a hairpin RNA. These double-stranded hairpin RNAs are processed by Dicer into siRNAs which direct RISC-mediated degradation of the target mRNA. Figure modified from `http://www.vdrc.at/typo3temp/pics/52ad173258.jpg`.

not all hairpins work to the same efficiency of knock-down and hence the observed phenotypes may differ despite the fact that only the correct on-target is knocked-down. Furthermore, not all genes are suited to generate several optimal 300 bp long hairpin sequences without overlap.

A conclusive proof of RNAi specificity is a rescue with a transgene that is immune to the RNAi and complements the loss of function of the target gene (Sarov and Stewart, 2005). A convenient source of an RNAi-immune transgene is an orthologous gene from another closely related species that is divergent enough on the nucleotide sequence level to diminish RNAi efficiency while still functionally complementing the knock-down of the endogenous gene activity. Such an approach was successfully applied in human tissue culture RNAi using BAC transgenes from mouse (Kittler et al., 2005) and in *C. elegans* with subcloned genomic BAC from *C. briggsae* (Sarov et al., 2006).

1.6. Aims of the work

This project aims to develop a set of reverse genetic and imaging techniques to capture the *Drosophila* embryos *in toto* and *in vivo* across their development with at least cellular resolution. We would like to trace all major cellular level changes occurring

during development, namely cell divisions, cell motility and cell death. We also want to record spatial and temporal information on expression of selected patterning genes expressed during embryogenesis and couple them with morphological and anatomical changes within the embryo. As changes in gene expression for many genes are connected with cell differentiation, this kind of events can also be registered in the project. By tracing all cells in the embryo though their divisions and migrations, and assigning gene expression levels for patterning genes to these traced cells, we will generate an unprecedented systemic description of animal development. It will serve as a foundation for addressing questions about global cellular behaviour during morphogenesis and the role of dynamics of gene expression regulation in developmental events.

In the first step towards achieving this long-term vision, we have created a toolkit allowing manipulation of Drosophilidae genomes in a high-throughput manner, providing faithful reporters for visualization of gene expression patterns. By combining genomic DNA libraries, φC31 transgenesis and selectable markers for isolation of transgenic animals in a broad range of insect species we provide a reliable source of transgenes, enabling functional analysis of any gene in the context of its intact *cis*-regulatory neighborhood. With high-throughput, liquid culture recombineering we furnish efficient means to modify genomic constructs with fluorescent markers, allowing to visualize gene expression patterns *in vivo*. Finally, with selective plane illumination microscopy, we bring *in vivo* and *in toto* imaging with cellular resolution to the *Drosophila* research.

Chapter 2.

2 Materials and methods

Contents

2.1. General remarks

Many methods described below rely on the protocols developed during the thesis work. For convenience, these protocols are listed in appendix B. Bacterial strains, fly stocks, enzymes, kits and other reagents required in described experiments are listed in appendix A. If not stated otherwise in the text, the following standard procedures were employed:

- Restriction digests were done in a 50 µl reaction using enough substrate to give at least 1 µg of the product. The manufacturer's supplied buffer was used in the digest. Double digests were done using a buffer suggested by the NEB double digest finder (`http://www.neb.com/nebecomm/DoubleDigestCalculator.asp`). Whenever possible, enzymes with reduced star activity (HF enzymes) were used. Digests were done overnight at the optimal temperature (usually at 37°C).

- Digestion products were gel-purified using QIAquick Gel Extraction Kit. Fragments shorter than 70 bp or longer than 10 kb were purified using QIAEX II Gel Extraction Kit.

- PCR reactions were done using Phusion® High-Fidelity DNA Polymerase (for cloning the PCR product) or Taq DNA Polymerase (for screening). Reactions were assembled according to the manufacturer's recommendations. PCR products were purified using QIAquick PCR Purification Kit.

- Ligation reactions were done in 10 µl volume using T4 DNA Ligase overnight at 16°C. If ligations were used in non-directional cloning, the vector has been dephosphorylated using Antarctic Phosphatase.

- Ligation products were electroporated into SmartCells *E. coli* (or other strain) as described in protocol 1.

- Plasmids were transformed into SmartCells *E. coli* (or other strain) as described in protocol 2.

- Flies were handled in standard conditions at 25°C. Viable, stable fly stocks were stored at 18°C and flipped once a month.

2.2. Generation of nuclear and membrane markers for *in vivo* embryo imaging

2.2.1. Marker for visualization of cell nuclei (HisEYFP)

Construction of the marker The *HisEYFP* marker was generated from *Histone 2AvD-GFP* construct created by Clarkson and Saint (1999). The cassette containing *His2AvD-GFP* fusion gene was amplified from pONIXAvDGFP using primers R4MCSinAfwd and R4MCSinBrev. The resulting PCR fragment was TOPO-cloned into pCR-XL-TOPO vector (Invitrogen) and sequenced by primer-walking using M13uni(-21) and M13rev(-29) as starting primers. A complete sequence of the resulting pHis2AvD-GFP plasmid has been assembled. After assembly, the recombineering primers HisBglIIcatF and HisNheIcatR were designed to replace *EGFP* with *Bgl*II- and *Nhe*I-flanked chloramphenicol acetylotransferase (*cat*) gene in reverse orientation. The *cat* gene has been amplified using the recombineering primers from pBAD33 (Guzman et al., 1995). The recombineering cassette has been digested with *Dpn*I to remove the PCR template, and electroporated (1 µg) together with pHis2AvD-GFP plasmid (1.2 µg) into recombineering-competent *E. coli* (protocol 3). Recombinant clones were selected on chloramphenicol plates and verified by RFLP using *Bgl*II/*Nhe*I. In the final step, *EYFP* has been amplified from p{SL-FRT-EYFP-linotte-FRT3} (Horn and Handler, 2005) using EYFP_cEX_F and EYFP_cEX_R primers, and cloned into *Nhe*I/*Bgl*II digested vector. Resulting clones were selected on kanamycin plates and replicated onto chloramphenicol plates. Clones sensitive to chloramphenicol were analyzed by *Nhe*I/*Bgl*II RFLP and sequencing-verified using HisEYFP_F and HisEYFP_R primers. The resulting construct, pHisEYFP was used to subclone *His2AvD-EYFP* reporter into *Not*I site of P-element vector pCaSpeR4 (Thummel and Pirrotta, 1992). Both forward (pCaSpeR4::HisEYFPfwd) and reverse (pCaSpeR4::HisEYFPrev) orientation variants were obtained, however only the forward variant was used in further experiments.

Fly transgenesis Transformation of *Drosophila* was achieved using P-element transposition (Rubin and Spradling, 1982). The pCaSpeR4::HisEYFPfwd was purified using QIAGEN Plasmid Maxi Kit and mixed with pTurbo (Tomlinson et al., 1988) as source of P-transposase at final concentrations of 500 ng/µl (construct DNA) and 100 ng/µl (pTurbo) in water. The mixture was injected into w^- embryos. The w^+ G_1 flies were crossed to *Sp/CyO* and *TM6B/MKRS* virgins and males for genetic mapping. Since the insertion was found on second chromosome, *HisEYFP/CyO* stock was established. *HisEYFP/CyO* virgins and males were used to establish stable homozygous *HisEYFP* stock.

2.2.2. Marker for visualization of cell membranes (CadECFP)

The *CadECFP* marker was generated from *DE-cad-GFP* construct created by Oda and Tsukita (2001). Since obtaining the pCaSpeR-ubi-DE-cad-GFP plasmid was not possible, genomic DNA was isolated from *ubi-DE-cad-GFP* fly line using QIAGEN DNeasy Blood & Tissue Kit. The cassette containing *DE-cad-GFP* fusion gene was amplified from *ubi-DE-cad-GFP* genomic DNA using primers R4MCSexFwd and R4MCSexRev. The resulting PCR fragment was TOPO-cloned into pCR-XL-TOPO vector (Invitrogen) and sequenced by primer-walking using M13uni(-21) and M13rev(-29) as starting primers. A complete sequence of the resulting pCad-GFP plasmid has been assembled. After assembly, the recombineering primers CadAscIcatF and CadNheIcatR were designed to replace *EGFP* with *Asc*I- and *Nhe*I-flanked chloramphenicol acetylotransferase (*cat*) gene in reverse orientation. The *cat* gene has been amplified using the recombineering primers from pBAD33 (Guzman et al., 1995). The recombineering cassette has been digested with *Dpn*I to remove the PCR template, and electroporated (1 μg) together with pHis2AvD-GFP plasmid (1.2 μg) into recombineering-competent *E. coli* (protocol 3). Recombinant clones were selected on chloramphenicol plates and verified by RFLP using *Asc*I/*Nhe*I. In the final step, *ECFP* has been amplified from pBac{3xP3-FRT-ECFP-linotte-FRT3} (Horn and Handler, 2005) using EYFP_cEX_F and ECFP_cEX_R primers, and cloned into *Nhe*I/*Asc*I digested vector. Resulting clones were selected on kanamycin plates and replicated onto chloramphenicol plates. Clones sensitive to chloramphenicol were analyzed by *Nhe*I/*Asc*I RFLP and sequencing-verified using CadECFP_F and CadECFP_R primers. The resulting construct, pCadECFP was used together with pHisEYFP to create universal cell imaging marker as described in section 2.2.3.

2.2.3. Universal cell imaging marker (HisEYFP/CadECFP)

Construction of the marker The universal cell imaging marker contains both *HisEYFP* and *CadECFP* constructs and allows to image both cell nuclei and cell membranes simultaneously. To ease transformation of flies, the pFlyFos vector (Ejsmont et al., 2009; section 2.3.1) and ΦC31-mediated transgenesis (Groth et al., 2004; Bischof et al., 2007) were used. *CadECFP* was excised from pCadECFP with *Eco*RI and cloned into *Eco*RI site of pFlyFos. Transformants were selected on chloramphenicol plates and analyzed by *Eco*RI RFLP. Only the forward orientation variant (pFlyFos::CadECFP) was used as a vector for cloning HisEYFP. *HisEYFP* was excised from pHisEYFP with *Not*I and cloned into *Not*I site of pFlyFos::CadECFP. Transformants were selected on chloramphenicol plates and analyzed by *Eco*RI RFLP. Resulting pFlyFos::HisEYFP-CadECFP plasmid was used for fly transgenesis.

Fly transgenesis The pFlyFos::HisEYFP-CadECFP has been purified with QIAGEN Plasmid Maxi Kit and injected into *attP40* landing line (Markstein et al., 2008) on the second chromosome. The injections were performed by Genetic Services. Received G_1 flies were crossed to *Sp/CyO* to establish *HisEYFPCadECFP/CyO* line. Males and virgins from heterozygous line were crossed to establish homozygous *HisEYFPCadECFP* stock.

2.3. Fosmid libraries production

2.3.1. pFlyFos fosmid vector

Construction of the vector The fosmid library vector, pFlyFos, was based on pCC2fos, a part of EPICENTRE CopyControl™ HTP Fosmid Library Production Kit. The 606 bp *Apa*LI/*Sfi*I fragment of pCC2fos has been resynthesized to include *3xP3-dsRed-SV40* cassette from pSL{FRT-EYFP-linotte-FRT3-3xP3-DsRed} (Horn and Handler, 2005), *LacZ* region from pCC2fos and *attB* integration site from attB-P[acman]-Cm^R(Venken et al., 2006). The 2,055 bp construct synthesized by Sloning BioTechnology was cloned into *Apa*LI/*Sfi*I sites of pCC2fos. Positive clones were selected on 15 µg/ml chloramphenicol plates and verified by *Sal*I RFLP. The resulting pFlyFos vector was used for fosmid library production and as a general vector for fly transgenesis.

Fly transgenesis The pFlyFos vector has been injected into *attP40* landing line (Markstein et al., 2008) to test the φC31-mediated transgenesis efficiency. Plasmid DNA was purified as described in protocol 5 and injected (500 ng/µl in water) into *attP40* embryos. The $dsRed^+$ G_1 flies were counted to assess the transformation efficiency.

2.3.2. Simulation of the fosmid library production

Before production of the fosmid genomic libraries for has started, the amount of clones required to include a certain number of genes has been estimated. The simulation program has been implemented in C programming language. The program requires six arguments: (1) mean clone length, (2) clone length standard deviation, (3) minimal distance between clone and gene boundaries, (4) file containing chromosomes and their lengths, (5) file containing gene annotations, and (6) number of clones to generate. The algorithm 2.1 shows pseudocode for the simulation. The source code is available on the attached DVD. In short, the whole genome is treated as one long

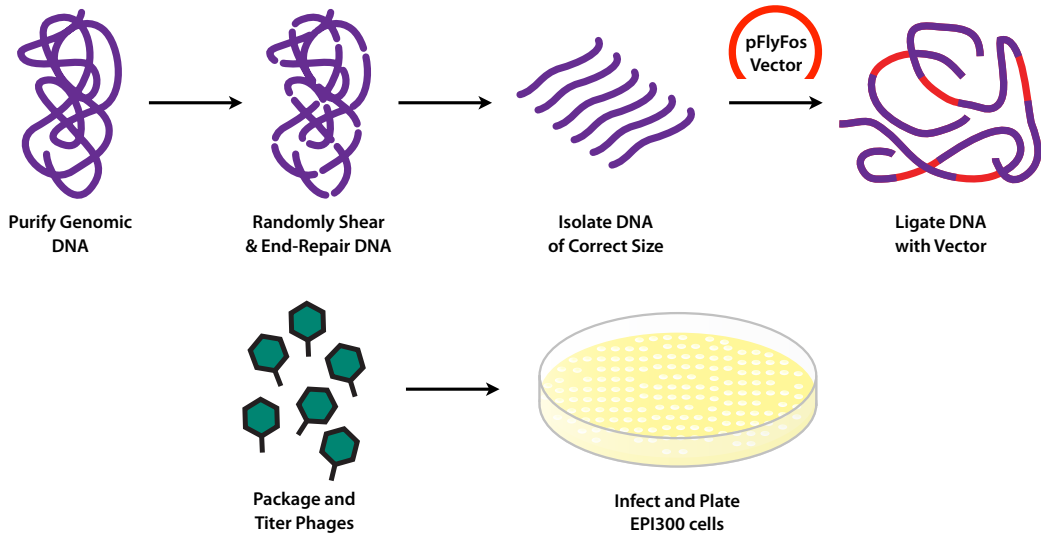


Figure 2.1.: Fosmid library production

sequence with marked beginnings and ends of the chromosomes. A random number is generated that sets beginning of a new clone. Another random number sets the clone length as a Gaussian variable generated using polar Box-Muller transformation (Devroye, 1986) with preset mean (μ) and standard deviation (σ). The end of the clone is set at the beginning plus acquired clone length. If the clone spans chromosome boundary, it is rejected and a new clone is generated. For each generated clone, the number of cloned genes in incremented by the number of genes that did not appear in previously generated clones and meet the minimum required distance from the clone boundaries. The whole procedure is repeated until a certain number of clones is generated.

2.3.3. Construction of the fosmid libraries

The fosmid genomic libraries for *D. melanogaster*, *D. pseudoobscura*, *D. simulans* and *D. virilis* were constructed from sequences strains as described in protocol 7. High molecular weight genomic DNA was isolated from about 1 ml of embryos as described in protocol 4. DNA (250 ng/µl) was sheared using HydroShear device with 4–40 kb (large) shearing assembly. The following parameters were used for DNA shearing: speedcode 17, retraction speed 40, 25 shearing cycles, 200 µl sample volume. Sheared DNA was end-repaired, purified and used directly (without size-selection) for ligation with *Pml*I digested, dephosphorylated pFlyFos vector. Ligated fosmids were packaged into phage particles and used to infect EPI300 cells. Libraries were plated on 15 µg/ml chloramphenicol plates. The library production process in summarized in figure 2.1.

1: **for** $n = 1$ to $N_{chromosomes}$ **do**
2: $chromosome_n.start \Leftarrow genome.length + 1$
3: $chromosome_n.end \Leftarrow chromosome_n.start + chromosome_n.length - 1$
4: $genome.length \Leftarrow chromosome_n.end$
5: **for** $m = 1$ to $chromosome_n.N_{genes}$ **do**
6: $chromosome_n.gene_m.start$
$\Leftarrow chromosome_n.start + chromosome_n.gene_m.start - 1$
7: $chromosome_n.gene_m.end$
$\Leftarrow chromosome_n.start + chromosome_n.gene_m.end - 1$
8: $chromosome_n.gene_m.new \Leftarrow$ **true**
9: **end for**
10: **end for**
11: $i \Leftarrow 0$
12: $N_{cloned} \Leftarrow 0$
13: **while** $i \leq N_{clones}$ **do**
14: **repeat**
15: $n \Leftarrow 0$
16: $clone_i.start \Leftarrow$ RANDOM $k \in \mathbb{N} : 1 \leq k \leq genome.length$
17: $clone_i.length \Leftarrow$ RANDOM $l \sim \mathbb{G}(\mu, \sigma)$
18: $clone_i.end \Leftarrow clone_i.start + clone_i.length - 1$
19: **repeat**
20: $n \Leftarrow n + 1$
21: **until** $clone_i.start \geq chromosome_n.start$
22: **until** $clone_i.end \leq chromosome_n.end$
23: **for** $m = 1$ to $chromosome_n.N_{genes}$ **do**
24: **if** $chromosome_n.gene_m.start \geq clone_i.start + flank$
and $chromosome_n.gene_m.end \leq clone_i.end - flank$
and $chromosome_n.gene_m.new =$ **true then**
25: $chromosome_n.gene_m.new \Leftarrow$ **false**
26: $N_{cloned} \Leftarrow N_{cloned} + 1$
27: **end if**
28: **end for**
29: $i \Leftarrow i + 1$
30: **end while**
31: **return** N_{cloned}

Algorithm 2.1: Simulation of random DNA shearing

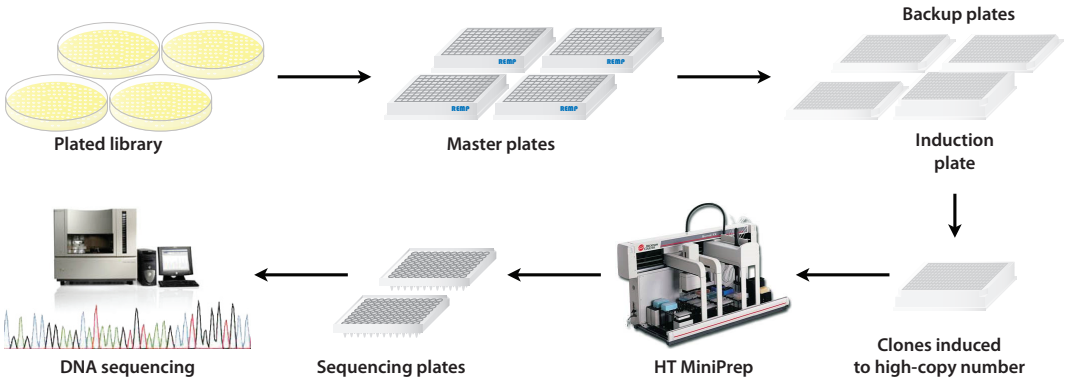


Figure 2.2.: Library re-arraying and sequencing schema

Clones were manually picked into 200 µl LB + 25 µg/ml chloramphenicol in 96-well REMP plates and cultured overnight at 37°C. Clones were induced to high-copy number by inoculating 100 µl LB + 25 µg/ml chloramphenicol + 0.1% arabinose in 384-well deepwell plates with 5 µl of the overnight culture. Remaining primary overnight cultures were supplemented with glycerol to 10% final concentration and distributed (40 µl per well) into three 384-well backup plates. Cultures remaining in primary plates and backup plates were frozen at -80°C for long-term storage. Fosmid DNA was isolated from induced cultures using high-throughput, 384-well format MiniPrep (protocol 6). Liquid handling was done on Beckman Coulter Biomek FX Laboratory Automation Workstation. Isolated DNA was used for fosmid end-sequencing with pCC2FOSfwd and pCC2FOSrev primers. See figure 2.2 for plate processing schema.

2.3.4. Clone mapping

Clone mapping was performed using console script written in PHP. Mapping data was stored in MySQL database (figure 2.3). Pseudocode for the mapping software is presented in algorithm 2.2. The end sequences for each clone were BLASTed (word size 11, expect threshold 10) against appropriate genomic sequence (*D. melanogaster*, *D. pseudoobscura*). The forward and reverse BLAST results were analyzed for each clone, starting from the best matches. Only alignments longer than 100 bp were taken into account. Clone was considered as successfully mapped if (1) both forward and reverse BLAST hits were on the same sequence (chromosome, scaffold or contig), (2) the forward sequencing read had orientation opposite to the reverse read, and (3) the clone size based on the BLAST results was between 10 kb and 80 kb. If these conditions were not met, the clone was considered as chimeric.

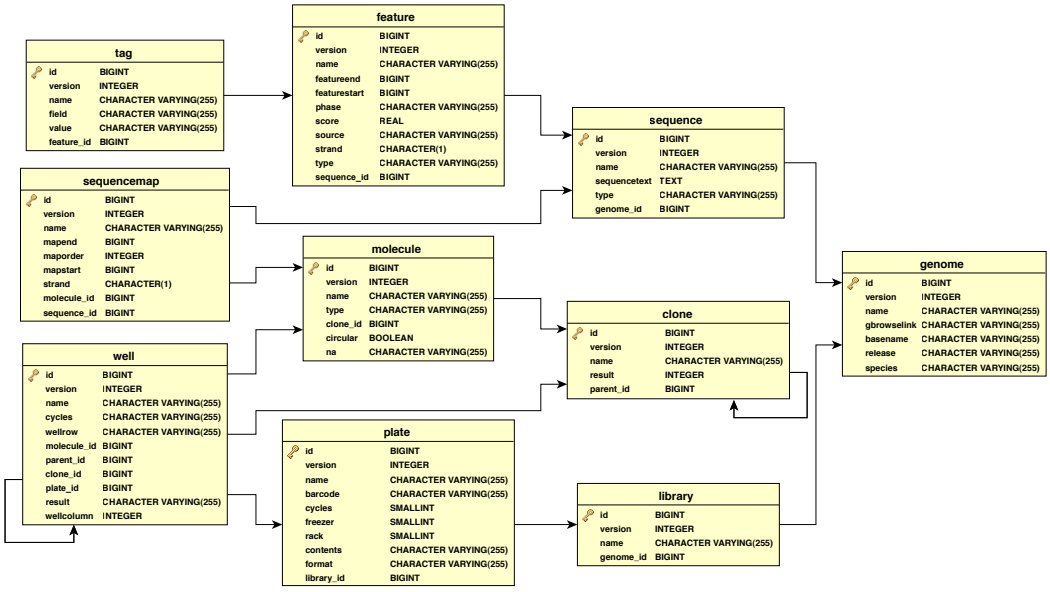


Figure 2.3.: Simplified schema of the FlyFos database

The sequence table stores reference sequences (genomic sequence, vector sequence). Features annotated on the reference sequences are stored in table feature together with all the feature information that is stored in table tag. If a given sequence is a reference genomic sequence, a proper reference is assigned to the table genome, where information about the genomes is stored. Reference sequences build up molecules (stored in the molecule table) via mapping in the sequencemap table. This way, a DNA construct can be stored as an ordered set of references to the reference sequences, and thus reducing redundancy. Clones (table clone) are carriers of the one or more DNA molecules (plasmids, fosmids). They are stored in wells (table well) of a 96- or 384-well plates (table plate). A set of plates containing clones with a certain content build up a library for a given species (genome).

```
1: result ⇐ FAILED
2: while i ≤ N_fwdBLASTs do
3:     while j ≤ N_revBLASTs do
4:         if fwdBLAST_i.subject = revBLAST_j.subject
           and fwdBLAST_i.start < fwdBLAST_i.end
           and revBLAST_j.start > revBLAST_j.end
           and revBLAST_j.start − fwdBLAST_i.start > 10000
           and revBLAST_j.start − fwdBLAST_i.start < 80000 then
5:             clone_n.start ⇐ fwdBLAST_i.start
6:             clone_n.end ⇐ revBLAST_j.start
7:             clone_n.strand ⇐ forward
8:             result ⇐ MAPPED
9:             return result
10:        else if fwdBLAST_i.subject = revBLAST_j.subject
           and fwdBLAST_i.start > fwdBLAST_i.end
           and revBLAST_j.start < revBLAST_j.end
           and fwdBLAST_i.start − revBLAST_j.start > 10000
           and fwdBLAST_i.start − revBLAST_j.start < 80000 then
11:            clone_n.start ⇐ revBLAST_j.start
12:            clone_n.end ⇐ fwdBLAST_i.start
13:            clone_n.strand ⇐ reverse
14:            result ⇐ MAPPED
15:            return result
16:        else
17:            result ⇐ CHIMERA
18:            return result
19:        end if
20:        j ⇐ j + 1
21:    end while
22:    i ⇐ i + 1
23: end while
24: return result
```

Algorithm 2.2: Clone mapping

2.4. Liquid culture recombineering

2.4.1. Construction of pTag[ubi-mCherry-NLS-T2A]

The *mCherry* (Shaner et al., 2004) tagging vector, pTag[ubi-mCherry-NLS-T2A], was based on the pTag3 plasmid by M. Sarov. The *ubiquitin-mCherry* fusion gene has been amplified from pCS2+UbCherry plasmid by D. Soroldoni using CherryNoXhoIFwd and CherryBspEIrev primers. The *NLS-T2A* fusion was generated by PCR. Nuclear localization signal (*NLS*) has been amplified from pStinger (Barolo et al., 2000) using nlsT2Afwd and nlsT2Arev primers. The 54 bp sequence encoding *T2A* (Osborn et al., 2005) was included on the reverse primer. The *Nhe*I/*Xho*I digested *ubi-mCherry* and *Xho*I/*Bsp*EI digested *NLS-T2A* constructs were ligated together with 3286 bp *Nhe*I/*Bsp*EI fragment from pTag3. Ligation products were electroporated into *pir-116* cells. Transformants were selected on plates with 25 μg/ml kanamycin and 50 μg/ml ampicillin. Clones were verified by *Pst*I RFLP and sequencing with pTag4chkFwd and pTag4chkRev primers. The amplified pTag[ubi-mCherry-NLS-T2A] plasmid was re-sequenced using pTag4seq primers.

2.4.2. *CG4702* gene tagging

Recombineering The *CG4702* gene was used to test the recombineering in FlyFos clones. Tagging was performed as described in protocol 8. The whole procedure is summarized in figure 2.4. The FlyFos014971 clone was selected as a source of *CG4702* gene. Tagging with *ubi-mCherry-NLS-T2A* was done N-terminally, between the start codon and the second codon. Tagging with *2xTY1-EGFP-3xFLAG* was done C-terminally, between the last codon and the stop codon. The tagging cassettes were amplified from pTag[ubi-mCherry-NLS-T2A] (section 2.4.1) and pTag[2xTY1-EGFP-3xFLAG] (Sarov et al., 2006). The primers used for the tagging cassette amplification contained a 25 bp priming region and a 50 bp homology arm, complement to the sequence flanking tagging site. See section A.6.3 for primer sequences.

An aliquot of the frozen glycerol stock containing selected clone was plated on 15 μg/ml chloramphenicol plates. A single colony was used to inoculate 1 ml of LB + 25 μg/ml chloramphenicol and cultured overnight at 37°C. Bacteria were transformed with pRedFlp4 recombineering helper (Sarov et al., 2006) and grown overnight at 30°C in LB + 25 μg/ml chloramphenicol + 50 μg/ml hygromycin. A fresh culture was grown until OD600 of 0.2, induced with 0.5% L-rhamnose for 1 hour at 37°C and transformed with the tagging cassette. Recombinants were selected in liquid culture on LB + 25 μg/ml chloramphenicol + 50 μg/ml hygromycin + 25 μg/ml kanamycin. The *FRT*-flanked selectable marker was removed by growing the bacteria

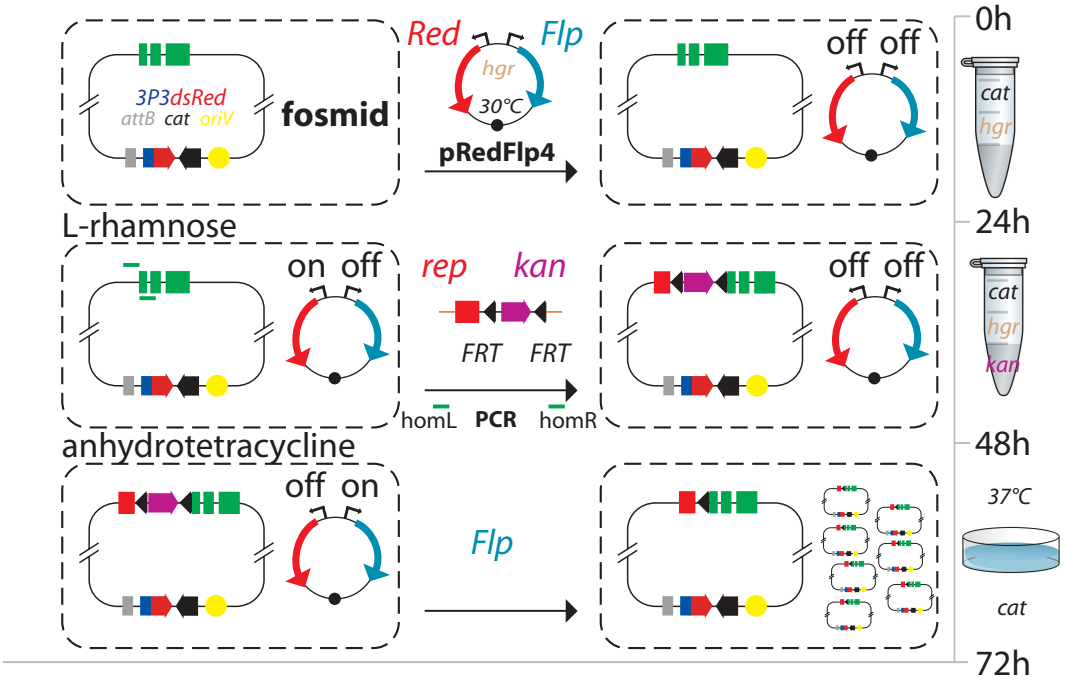


Figure 2.4.: Recombineering pipeline used for tagging *CG4702*.

(0h–24h) *E. coli* cultures containing a fosmid clone of interest are transformed with pRedFlp4 plasmid carrying an inducible Red operon rendering them competent for homologous recombination. (24h–48h) A PCR product carrying 50 bp homology arms surrounding the tagging cassette and *FRT* flanked kanamycin resistance gene (*kan*), is electroporated into the cells. Only recombinant fosmids are able to grow efficiently in the presence of kanamycin. (48h–72h) The *kan* gene is removed by inducing a flippase on the pRedFlp4 plasmid leaving the tagged transgene with a residual *FRT* sequence on the gene-tag boundary.

overnight in LB + 25 μg/ml chloramphenicol + 50 μg/ml hygromycin + 200 mM anhydrotetracycline at 30°C. The pRedFlp4 helper was removed during an overnight culture in LB + 25 μg/ml chloramphenicol at 37°C. Fosmid DNA from recombinant clones was isolated as described in protocol 5 and analyzed by sequencing with ubi-2NR and T2A-1CF primers for *CG4702-mCherry*, EGFP-N and EGFP-C primers for *CG4702-EGFP.*

Fly transgenesis Both *CG4702-mCherry* and *CG4702-EGFP* fosmids were injected into *attP40* landing line (Markstein et al., 2008). Purified fosmid DNA was injected at 1 μg/μl in water. G_1 flies were selected for *dsRed* expression in the eyes and crossed to *Sp/CyO* to establish heterozygous lines. Males and virgins from heterozygous lines were crossed to establish *CG4702-mCherry* and *CG4702-EGFP* homozygous stocks.

Verifying expression pattern of *CG4702* transgenes The *CG4702-mCherry* and *CG4702-EGFP* stocks were amplified and used to set-up embryo collection cages. Embryos were collected for 24 hours with 3-hour intervals. Collected embryos were dechorionated, devitellinized and frozen in -80°C as described by Weiszmann et al. (2009). *In situ* probes were prepared for *mCherry*, *EGFP* and *CG4702* from PCR fragments amplified from pTag[ubi-mCherry-NLS-T2A], pTag[2xTY1-EGFP-3xFLAG] and pFlyFos014971 respectively. The following primer pairs were used for amplification: ubi_mCherry_fwd and ubi_mCherry_rev_T7, EGFP_fwd and EGFP_rev_T7, CG4702_fwd and CG4702_rev_T7. RNA probe synthesis, *in situ* hybridization and staining was performed following the aforementioned protocol. Stained embryos were mounted on microscopy slides in 70% glycerol and imaged on Zeiss Axioplan2 with EC Plan-Neofluar 40x/0.75 objective.

In addition to *in situ* hybridization, the *CG4702-EGFP* embryos were stained by immunofluorescence. The fixed embryos were washed for 10 minutes twice with PBT and once with PBT + 5% normal goat serum. After washing, embryos were incubated overnight at 4°C with anti-GFP rabbit polyclonal antibody (Invitrogen A11122) in PBT + 5% normal goat serum (1:250). After overnight incubation, embryos were washed three times with PBT for 10 minutes. Washed embryos were incubated with donkey anti-rabbit IgG Cy2-antibody conjugate (dianova 711-225-152) in PBT + 5% normal goat serum (1:250). After incubation with the secondary antibody, embryos were washed three times with PBT for 10 minutes. Stained embryos mounted on microscopy slides in 70% glycerol and were imaged on Zeiss Axioplan with EC Plan-Neofluar 40x/0.75 objective. Living and GFP antibody-stained *CG4702-EGFP* embryos were also imaged on Zeiss SPIM prototype microscope as described in section 2.8.

2.4.3. High-throughput gene tagging

Gene and clone selection Genes for tagging experiments were selected based on an annotated embryonic gene expression pattern in the Atlas of Patterns of Gene Expression (Tomancak et al., 2007) and the availability of a suitable fosmid clone. The annotated patterns were extracted from the APOGEE database, collapsed to the level of organ systems and temporally grouped into three ranges covering the early (blastoderm), mid (stages 4–10) and late embryogenesis (stages 11–16). The patterns were organized by hierarchical clustering and cross-referenced with the MySQL database of FlyFos clones. Forty eight genes that are included in fosmid clones and together cover most of the organ systems in the annotation hierarchy were selected for tagging.

Fosmid clones suitable for recombineering were selected using a console script written in PHP. First, clones that contain the complete gene model of interest including at least 2.5 kb of upstream and downstream noncoding region were selected. Next, the clones containing (in sorting order) either both, upstream only or downstream only neighboring genes were prioritized. Finally, the clones have been ordered by a score s calculated using the formula listed below:

$$s = 10 \times \left(\log_2\left(\frac{u}{d}\right)\right)^2 + \frac{(u+d)^2}{500 \times f^2} + \frac{50 \times f^2}{(u+d)^2}$$

where u is the length of the upstream sequence, d is the length of the downstream sequence, and f the length of the gene.

This scoring formula assigns the highest score to clones where the gene has more upstream than downstream sequence and where the clone size is proportional to the size of a given gene. Pseudocode for clone selection script is presented in algorithm 2.3. The start codon (N-terminal tagging) or the stop codon (C-terminal tagging) that contributes to the most protein isoforms was selected as a tagging site. Recombineering primers were designed automatically using a console script written in PHP. Pseudocode for primer design and tagging site selection scripts in presented in algorithm 2.5 and algorithm 2.4 respectively.

Recombineering The selected 48 genes were tagged in 96-well format as described in protocol 9. Tagging was performed by the MPI-CBG TransGeneOmics Unit using four tags: (1) *ubi-mCherry-NLS-T2A* for N-terminal tagging, and (2) *2xTY1-EGFP-3xFLAG,* (3) *2xTY1-T2A-EGFP-3xFLAG,* (4) *2xTY1-tdTomato-3xFLAG* for C-terminal tagging. Tags were amplified for each gene using gene-specific recombineering primers. Primers contained a 25 bp priming region and a 50 bp homology arm,

1: **for all** *clones* **where** $clone.start < gene.start - 2500$
and $clone.end > gene.end + 2500$ **do**
2: $u \Leftarrow gene.start - clone.start$
3: $d \Leftarrow clone.end - gene.end$
4: $f \Leftarrow gene.end - gene.start + 1$
5: $clone.score \Leftarrow 10 \times \left(\log_2\left(\frac{u}{d}\right)\right)^2 + \frac{(u+d)^2}{500 \times f^2} + \frac{50 \times f^2}{(u+d)^2}$
6: **for all** *clone.genes* **do**
7: **if** $clone.gene.end < gene.start$ **then**
8: $clone.uneighbor \Leftarrow$ **true**
9: **else if** $clone.gene.start > gene.end$ **then**
10: $clone.dneighbor \Leftarrow$ **true**
11: **end if**
12: **end for**
13: **end for**
14: **order** *clones* **by** (*clone.uneighbor*, *clone.dneighbor*), *clone.score*
15: **return** *clones*

Algorithm 2.3: Fosmid clone selection

1: $primerFwd \Leftarrow tag.subSequence(1, primerLength, forward)$
2: $primerRev \Leftarrow tag.subSequence($
$tagLength - primerLength + 1, tagLength, reverse)$
3: **if** $gene.strand = forward$ **then**
4: $homologyFwd \Leftarrow clone.subSequence($
$tagPoint - homologyLength + 1, tagPoint, forward)$
5: $homologyRev \Leftarrow clone.subSequence($
$tagPoint + 1, tagPoint + homologyLength, reverse)$
6: **else**
7: $homologyFwd \Leftarrow clone.subSequence($
$tagPoint + 1, tagPoint + homologyLength, reverse)$
8: $homologyRev \Leftarrow clone.subSequence($
$tagPoint - homologyLength + 1, tagPoint, forward)$
9: **end if**
10: $primers.forward \Leftarrow homologyFwd + primerFwd$
11: $primers.reverse \Leftarrow homologyRev + primerRev$
12: **return** *primers*

Algorithm 2.4: Recombineering primer design

```
1: for all gene.mRNAs do
2:     if taggingTerminus = N then
3:         if gene.strand = forward then
4:             tagPoint.position ⇐ mRNA.CDS.start + 2
5:         else
6:             tagPoint.position ⇐ mRNA.CDS.end − 3
7:         end if
8:     else
9:         if gene.strand = forward then
10:            tagPoint.position ⇐ mRNA.CDS.end − 3
11:        else
12:            tagPoint.position ⇐ mRNA.CDS.start + 2
13:        end if
14:    end if
15:    if gene.tagPoints(tagPoint).count > 0 then
16:        gene.tagPoints(tagPoint).count ⇐ gene.tagPoints(tagPoint).count + 1
17:    else
18:        gene.tagPoints(tagPoint).count ⇐ 1
19:        gene.tagPoints(tagPoint).position ⇐ tagPoint.position
20:    end if
21: end for
22: finalTagPoint.count ⇐ 0
23: finalTagPoint.position ⇐ 0
24: for all gene.tagPoints do
25:    if tagPoint.count > finalTagPoint.count then
26:        finalTagPoint.count ⇐ tagPoint.count
27:        finalTagPoint.position ⇐ tagPoint.position
28:    end if
29: end for
30: return finalTagPoint.position
```

Algorithm 2.5: Selection of the tagging site

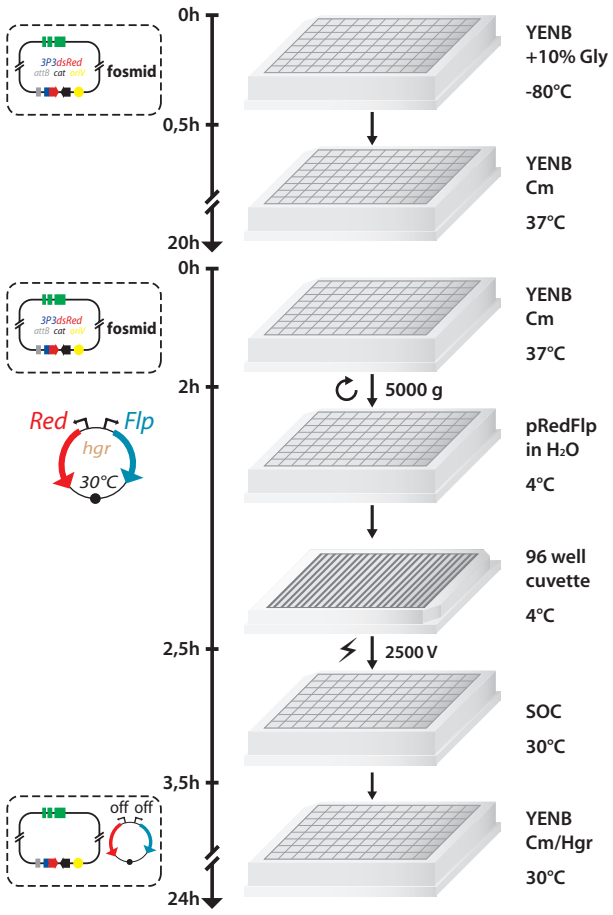
complement to the sequence flanking tagging site. See section A.6.3 for primer sequences. All C-terminal tagging cassettes were amplified using the same primers for each gene. The recombineering pipeline is summarized in figure 2.5. Fosmid DNA from recombinant clones was isolated as described in protocol 6. Clones were analyzed by *Xba*I and *Pac*I fingerprinting and sequencing with ubi-2NR and T2A-1CF primers for *mCherry*, EGFP-N and EGFP-C primers for *EGFP* and *T2A-EGFP.* Since designing effective primers for sequencing of *tdTomato*-tagged clones has failed, they were analyzed fingerprinting only.

Fly transgenesis and expression pattern validation Fosmid DNA from selected 12 *mCherry*-tagged clones, four *EGFP*-tagged clones and one *T2A-EGFP*-tagged clone has been purified with QIAGEN Plasmid Maxi Kit and injected into *attP40* landing line (Markstein et al., 2008). The injections were performed by Genetic Services. Received G_1 flies were crossed to *Sp/CyO* to establish heterozygous line. Males and virgins from heterozygous line were crossed to establish homozygous stocks. Homozygous lines were amplified and used to set-up embryo collection. Collected embryos were fixed and subjected to *in situ* hybridization, as described previously in section 2.4.2.

2.5. RNAi rescue in *D. melanogaster* with *D. pseudoobscura* FlyFos clones

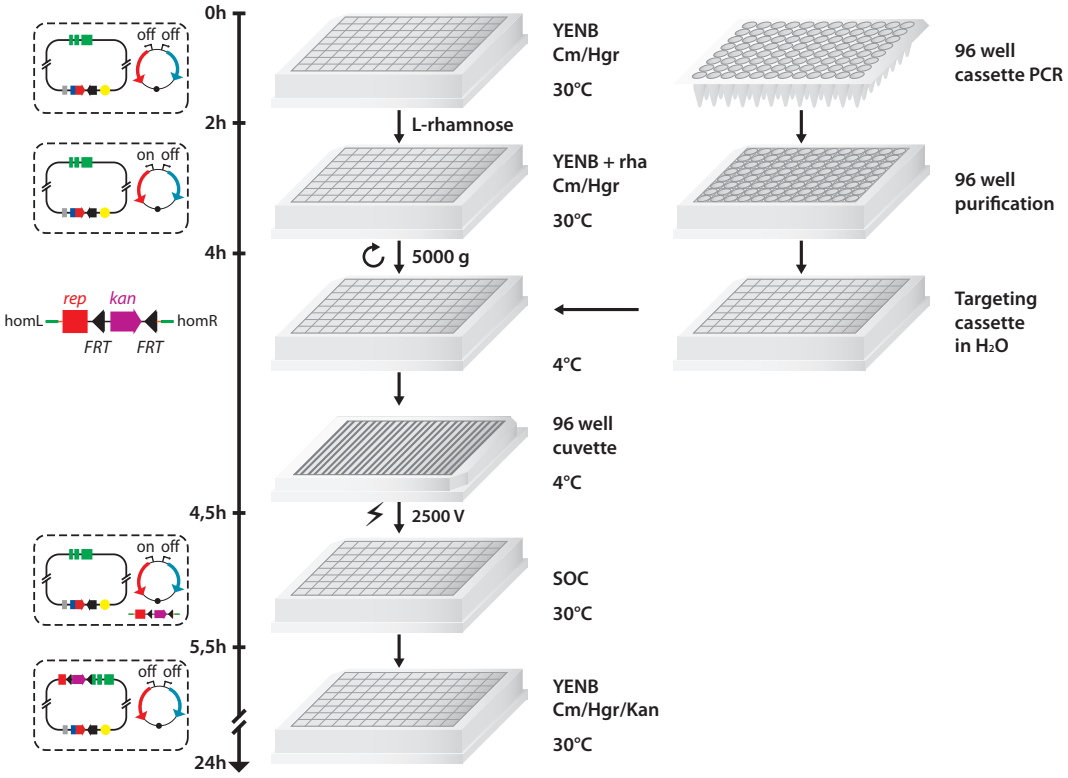
2.5.1. Analysis of hairpin sequence divergence

The bioinformatics analysis of hairpin sequence divergence was done described in Langer et al. (2010). The pairwise alignments of genomic sequences between *Drosophila melanogaster* and five non-melanogaster Drosophilidae were downloaded from UCSC database (`http://hgdownload.cse.ucsc.edu/downloads.html`). The following reference sequences were used: dm3 (*D. melanogaster*), droSim1 (*D. simulans*), droAna3 (*D. ananassae*), dp4 (*D. pseudoobscura*), droPer1 (*D. persimilis*) and droVir3 (*D. virilis*). The portions of pairwise alignments corresponding to the transcripts annotated in FlyBase release 5 of *D. melanogaster* genome were extracted using a console script written in PERL. In case a transcript had multiple isoforms, the longest transcript was used as a reference. Extracted sequence regions were grouped using *D. melanogaster* sequence as a reference to create a multiple alignment file for each gene. The files were searched with RNAi hairpin sequences from genome wide transgenic RNAi library (Dietzl et al., 2007). To simplify the search, only one hairpin per gene was used in the search. Sequences for genes not covered by

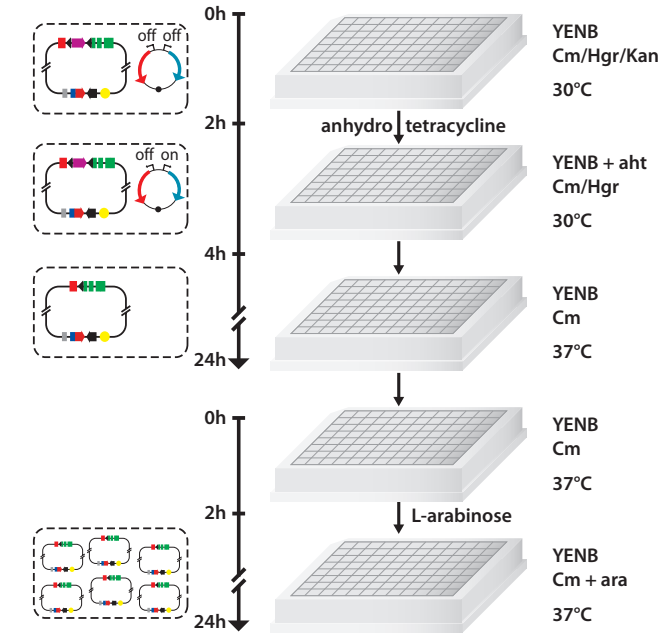


(a) Transformation of recombineering helper

Figure 2.5.: High-throughput recombineering pipeline



(b) Tagging by Red/ET recombination



(c) Removal of selectable marker and recombineering helper

Figure 2.5.: High-throughput recombineering pipeline (continued)

UCSC genome wide alignments and the ones that did not map completely were excluded from the analysis. The number of conserved nucleotides and the length of longest uninterrupted stretch of the identical sequence were counted in the alignment sequences corresponding to the mapped RNAi hairpins.

2.5.2. FlyFos clone selection

PCR screening proof of principle Since exact mapping of *D. pseudoobscura* clones was not complete at the beginning of the RNAi rescue project, a strategy for identifying a clone containing the gene of interest using a series of PCR reactions was developed. Saturated 1 ml cultures for *D. pseudoobscura* clones were split into two 500 µl aliquots. The first set of aliquots was pooled platewise (i.e. cultures from each plate were pooled together) and placed (2 ml) in a 96-well plate. Fosmid DNA was isolated from the library plates (500 µl culture) and the pool plate (2 ml culture) using an automated MiniPrep as described in protocol 6. The isolated fosmid DNA was used as a template in two consecutive PCR reactions with gene-specific primers (MICAL gene was used as a target for search with dpse/MICAL_fwd and dpse/MICAL_rev primers). The first 96-well PCR reaction was run using fosmid DNA from the pool plate as a template. That reaction was performed to identify plate containing the clone of interest. The second PCR reaction was run using fosmid DNA from the plate identified in the previous step, to identify well that contains the desired clone. The FlyFos clone from the identified well was amplified as described in protocol 5 and mapped to the *D. pseudoobscura* genome using data from sequencing with pCC2fos_fwd and pCC2fos_rev primers.

Clone selection The genes identified in the genome-wide screen for muscle phenotypes with *Mef2-Gal4* driver (Schnorrer et al., 2010) were used to search the FlyFos database of *D. pseudoobscura* fosmid clones. Identified fosmids were manually inspected and five fosmids were arbitrarily selected, based on the phenotype given by RNAi knockdown and the position of a given gene within the fosmid. FlyMine (Lyne et al., 2007) was used as a tool to integrate the data for fosmid clone selection. The fosmid DNA for selected clones was isolated as described in protocol 5 and injected into *attP2* landing line (Markstein et al., 2008) on the third chromosome. The injections were performed by Genetic Services.

2.5.3. RNAi rescue

The described rescue experiment was performed in F. Schnorrer's lab (MPIB Martinsried) by CCH. Langer and C. Schönbauer, as described in Langer et al., 2010.

The flies received from injection were used to recombine the fosmid insertions with *Mef2-Gal4* driver located on the third chromosome. Recombinants were identified by *dsRed* fluorescence in the *ocelli*. The recombinant flies were crossed with RNAi lines for selected genes. In case the hairpin construct was present on the third chromosome, it was recombined with the *Mef2-Gal4/FlyFos-pse* lines. The *Mef2-Gal4/FlyFos-pse* flies were also crossed to the available mutant and deficiency lines for the selected genes. The phenotypes of the RNAi-induced knockdown, mutants and results of rescue with *D. pseudoobscura* fosmids were determined by inspection of the muscle morphology in larvae and by measurement of the larval length at 48 or 72 hours, depending on the strength of the phenotype.

2.6. Improved recombineering tags

2.6.1. pTagNG vector

The pTagNG is a common vector for the next generation tags. It is based on the pR6K backbone (Sarov et al., 2006). The main part of the vector has been synthesized as a single fragment introducing *Not*I and *Eco*RV sites at the 5' and the 3' end respectively. The fragment includes (in 5'–3' order): *2xTY1* epitope, *Eco*RI site, *T2A* ribosomal cleavage site, *Nhe*I, and *Xho*I sites separated by a spacer sequence, nuclear localization signal (*NLS*), *Bam*HI- and *Sal*I-flanked selection-counter selection cassette and *3xFLAG* epitope. The selection-counter selection cassette harbors *FRT*-flanked *rpsL* (streptomycin sensitivity) and *kanR* (kanamycin resistance) genes. All aforementioned restriction sites are unique and allow for easy insertion of various tags. The pR6K backbone has been amplified by PCR with pR6K_core_fwd and pR6K_core_rev primers. Primer sequences included *Eco*RV and *Not*I sites on the 5' and the 3' end respectively. The 2,457 bp TagNG fragment synthesized by GeneArt and the 1.787 bp PCR product were digested with *Not*I and *Eco*RV. Both restriction fragments were gel-purified and ligated in an overnight reaction. Ligation products were electroporated into *pir-116* cells. Transformants were selected on plates with 25 μg/ml kanamycin and 50 μg/ml ampicillin. Clones were verified by *Pst*I RFLP and sequencing with pTag_chk primers.

2.6.2. Fluorescent protein tags

Universal protein fusion tags Five codon-optimized fluorescent protein genes: *Cerulean* (Rizzo et al., 2004), *Venus* (Nagai et al., 2002), *EGFP* (Heim and Tsien, 1996), *mCherry* (Shaner et al., 2004) and *tagRFP* (Merzlyak et al., 2007) were synthesized by GeneArt for cloning into pTagNG core. Codon optimization for *Drosophila*

melanogaster was done using the OPTIMIZER webtool (Puigbò et al., 2007). The *Drosophila melanogaster* entry in the Codon Usage Database (Nakamura et al., 2000) was used as a reference for the codon optimization. Coding sequences for all constructs were free from commonly used restriction sites. Fluorescent protein genes were amplified by PCR using gene-specific primers introducing *Eco*RI site on the 5' end and *Bam*HI site on the 3' end (see table 2.1). In addition to the synthesized fluorescent protein genes, the superfolder *GFP* (*SGFP* - Pédelacq et al., 2006) gene has been amplified in a similar manner from pEGFPmultiFINAL provided by K. Venken. PCR products were digested with *Eco*RI/*Bam*HI and cloned into the pTagNG vector. Ligation products were electroporated into *pir-116* cells. Transformants were selected on plates with 25 μg/ml kanamycin and 50 μg/ml ampicillin. Clones were verified by *Pst*I RFLP and sequencing with pTag_chk_2 and pTag_chk_5 primers.

T2A-NLS C-terminal tags The T2A-NLS C-terminal tags were based on the universal tags described above. Fluorescent protein genes were amplified by PCR using gene-specific primers introducing *Nhe*I site on the 5' end and *Xho*I site on the 3' end (see table 2.2). PCR products were digested with *Nhe*I/*Xho*I and cloned into the pTagNG vector. Ligation products were electroporated into *pir-116* cells. Transformants were selected on plates with 25 μg/ml kanamycin and 50 μg/ml ampicillin. Clones were verified by *Pst*I RFLP and sequencing with pTag_chk_2 and pTag_chk_5 primers.

2.6.3. Biotin, V5 and birA tags

The biotin tags contain the sequence encoding biotin ligase recognition peptide (*BLRP*) in conjunction with *V5* epitope, *SGFP* or both. The *BLRP* is separated from the rest of the tag by two protease sites: *PreScission* and *TEV*. Since upon protein purification biotinylated *BLRP* remains bound to the streptavidin column while the purified protein is cleaved-off with the protease, the position on the *BLRP* determines whether the tag can be used N- or C-terminally. Therefore, two version for each of the biotin tags have been produced: N-terminal (*BLRP-preTEV-V5*, *BLRP-preTEV-SGFP*, *BLRP-preTEV-V5-SGFP*) and C-terminal (*V5-preTEV-BLRP*, *SGFP-preTEV-BLRP* and *SGFP-V5-preTEV-BLRP*). In addition to the BLRP tags, a tag containing *V5* epitope only and a *T2A-birA* (biotin ligase) tag have been constructed.

For the *V5* tags construction, the *BLRP-preTEV-V5* has been amplified from pRK2-Neo-N-term, the *V5-preTEV-BLRP* has been amplified from pRK2-Neo-C-term. The *V5* sequence been amplified from pRK2-Neo-N-term. The PCR primers for these constructs introduced *Eco*RI site on the 5' end and *Bam*HI site on the 3' end

(see table 2.3). All PCR products were digested with *Eco*RI/*Bam*HI and cloned into *Eco*RI/*Bam*HI-digested pTagNG vector. For N-terminal *BLRP-SGFP* tags, the *BLRP-preTEV-V5* and *BLRP-preTEV* have been amplified from pRK2-Neo-N-term with primers introducing *Eco*RI and *Nhe*I site on the 5' and 3' end respectively. *SGFP* has been amplified from pEGFPmultiFINAL with NheI_SGFP_fwd and BamHI_SGFP_rev primers. The *Eco*RI/*Nhe*I-digested *BLRP* PCR products and *Nhe*I/*Bam*HI-digested *SGFP* were combined in a ligation reaction with *Eco*RI/*Bam*HI-digested pTagNG vector. For C-terminal *SGFP-BLRP* tags, the *V5-preTEV-BLRP* and *preTEV-BLRP* have been amplified from pRK2-Neo-C-term with primers introducing *Xho*I and *Bam*HI site on the 5' and 3' end respectively. *SGFP* has been amplified as described previously with EcoRI_SGFP_fwd and XhoI_SGFP_rev primers. The *Xho*I/*Bam*HI-digested *BLRP* PCR products and *Eco*RI/*Xho*I-digested *SGFP* were combined in a ligation reaction with *Eco*RI/*Bam*HI-digested pTagNG vector. The biotin ligase gene (*birA*) has been amplified from pUASTattB-3xHABirA using NheI_birA_fwd and BamHI_birA_rev primers. The *Nhe*I/*Bam*HI-digested PCR product was cloned into *Nhe*I/*Bam*HI sites of pTagNG vector. All ligation products were electroporated into *pir-116* cells. Transformants were selected on plates with 25 μg/ml kanamycin and 50 μg/ml ampicillin. Clones were verified by *Pst*I RFLP and sequencing with pTag_chk_2 and pTag_chk_5 primers.

2.7. The "Ultimate" system

The Ultimate is an *in vivo* recombinase mediated cassette exchange (RMCE) based tagging system. In consists of two parts: pTagUltimate – a recombineering tag that when incorporated at either end of a target gene acts as an RMCE acceptor, and pUltimate – an RMCE donor. Both components of the system are integrated into the fly genome at the same location, allowing for an exchange of the selectable marker included in TagUltimate for a fluorescent tag included in the Ultimate construct.

2.7.1. pTagUltimate

The pTagUltimate RMCE acceptor tag has been constructed from three fragments synthesized by GeneArt. The TagUltimate_part_1 fragment was used as a scaffold for construction of the tag. It contained the *Eco*RI site for cloning the whole construct into pTagNG vector, the *FRT3* recombination site followed by *Gal4*Δ gene under *DSCP* minimal promoter. The gene is followed by *Pml*I and *Bam*HI sites used for insertion of the further parts of the tag. The TagUltimate_part_2 contained the *Venus* ORF together with an *Adh* terminator sequence, flanked with *Bam*HI and *Spe*I

Fluorescent protein	Forward primer	Reverse primer
Cerulean	EcoRI_Cerulean_dmel_fwd	BamHI_Cerulean_dmel_rev
Venus	EcoRI_Venus_dmel_fwd	BamHI_Venus_dmel_rev
EGFP	EcoRI_eGFP_dmel_fwd	BamHI_eGFP_dmel_rev
mCherry	EcoRI_mCherry_dmel_fwd	BamHI_mCherry_dmel_rev
tagRFP	EcoRI_tagRFP_dmel_fwd	BamHI_tagRFP_dmel_rev
SGFP	EcoRI_SGFP_fwd	BamHI_SGFP_rev

Table 2.1.: Primers used to amplify fluorescent proteins for cloning in pTagNG

Fluorescent protein	Forward primer	Reverse primer
Cerulean	NheI_Cerulean_dmel_fwd	XhoI_Cerulean_dmel_rev
Venus	NheI_Venus_dmel_fwd	XhoI_Venus_dmel_rev
EGFP	NheI_eGFP_dmel_fwd	XhoI_eGFP_dmel_rev
mCherry	NheI_mCherry_dmel_fwd	XhoI_mCherry_dmel_rev
tagRFP	NheI_tagRFP_dmel_fwd	XhoI_tagRFP_dmel_rev
SGFP	NheI_SGFP_fwd	XhoI_SGFP_rev

Table 2.2.: Primers used to amplify fluorescent proteins for cloning in pTagNG as *T2A-NLS* fusion constructs

Tag	Forward primer	Reverse primer
V5	EcoRI_V5_fwd	BamHI_V5_rev
BLRP-preTEV-V5	EcoRI_BLRP_fwd	BamHI_V5_rev
BLRP-preTEV-SGFP	EcoRI_BLRP_fwd	NheI_preTEV_rev
BLRP-preTEV-V5-SGFP	EcoRI_BLRP_fwd	NheI_V5_rev
V5-preTEV-BLRP	EcoRI_V5_fwd	BamHI_BLRP_rev
SGFP-preTEV-BLRP	XhoI_preTEV_fwd	BamHI_BLRP_rev
SGFP-V5-preTEV-BLRP	XhoI_V5_fwd	BamHI_BLRP_rev

Table 2.3.: Primers used to amplify biotin tags for cloning in pTagNG

Fluorescent protein	Forward primer	Reverse primer
Cerulean	XhoI_Cerulean_dmel_fwd	EcoRI_Cerulean_dmel_rev
Venus	XhoI_Venus_dmel_fwd	EcoRI_Venus_dmel_rev
EGFP	XhoI_eGFP_dmel_fwd	EcoRI_eGFP_dmel_rev
mCherry	XhoI_mCherry_dmel_fwd	EcoRI_mCherry_dmel_rev
tagRFP	XhoI_tagRFP_dmel_fwd	EcoRI_tagRFP_dmel_rev
SGFP	XhoI_SGFP_fwd	EcoRI_SGFP_rev

Table 2.4.: Primers used to amplify fluorescent proteins for cloning in pUltimate

sites on the 5' end and *Pml*I site on the 3' end. The TagUltimate_part_3 contained a *Bam*HI- and *Spe*I-flanked *GMR-SCP1* eye promoter. The TagUltimate_part_2 was excised with *Pml*I and *Bam*HI from the supplied pMA vector and cloned into *Pml*I/*Bam*HI-digested pMK::TagUltimate_part_1. Transformants were selected on kanamycin plates and verified by *Pml*I/*Bam*HI digest. In the next step, the TagUltimate_part_3 fragment was excised from the supplied vector using *Spe*I and *Bam*HI and cloned into *Spe*I/*Bam*HI-digested pMK:TagUltimate_part_1_2. Transformants were selected on kanamycin plates and verified by *Spe*I/*Bam*HI RFLP. Finally, the whole TagUltimate was excised from pMK with *Eco*RI and *Bam*HI and ligated with *Eco*RI/*Bam*HI-digested pTagNG vector. Ligation products were electroporated into *pir-116* cells. Transformants were selected on plates with 25 μg/ml kanamycin and 50 μg/ml ampicillin. Clones were verified by *Pst*I RFLP and sequencing with pTag_chk_2 and pTag_chk_5 primers.

2.7.2. pUltimate

The pUltimate RMCE donor vector has been constructed from two fragments synthesized by GeneArt and cloned into the pFlyFos vector. The Ultimate_part_1 fragment contained the *3xP3-Cerulean-SV40* gene followed by the *FRT3* recombination site. The construct was flanked by *Not*I site on the 5' end and *Xho*I/*Pml*I sites separated with a linker sequence on the 3' end. The Ultimate_part_2 fragment contained the *tagRFP* fluorescent protein followed by the *FRT* recombination site. The construct was flanked by *Xho*I site on the 5' end and the *Pml*I site on the 3' end. The Ultimate_part_2 fragment was excised from the supplied pMA vector using *Xho*I and *Pml*I and cloned into *Xho*I/*Pml*I-digested pMK::Ultimate_part_1. Transformants were selected on kanamycin plates and verified by *Xho*I/*Pml*I restriction digest. The complete Ultimate construct was excised from pMK with *Not*I and *Pml*I and cloned into *Not*I and *Pml*I sites of pFlyFos vector. Transformants were selected on 15 μg/ml chloramphenicol plates and verified by *Sal*I RFLP. The pUltimate plasmid has been purified as described in protocol 5 and sequenced using pUltimate_seq primers.

Codon optimized fluorescent proteins described in section 2.6.2 were amplified using gene-specific primers introducing *Xho*I site on the 5' end and *Eco*RI site on the 3' end (see table 2.4). Products of the PCR reaction were digested with *Xho*I/*Eco*RI and gel-purified. Digested fragments were cloned into *Xho*I/*Eco*RI sites of pUltimate to create a collection of RMCE donors. Transformants were selected on 15 μg/ml chloramphenicol plates and verified by sequencing using pUltimate_seq_2 and pUltimate_seq_5 primers.

2.8. Live Imaging

2.8.1. Embryo collection

The flies were held in collection cages overnight at 25°C. The next day, embryos were precollected twice for two hours. The precollected embryos were discarded, and the sample embryos were collected for two hours. Embryos were removed from the agar plates with a smooth brush and washed twice in PBT. Washed embryos were dechorionated in 50% bleach for two minutes, followed by two washes in PBT. Dechorionated embryos were inspected under the binocular and mounted for live imaging.

2.8.2. Laser Scanning Microscope

Dechorionated embryos were embedded in air-permable halocarbon oil, mounted on a standard microscopy slide and covered with a cover slip placed on 170 µm support. Samples were imaged on Zeiss LSM 405/594 confocal microscope. Samples were identified in bright field using Zeiss Plan-Apochromat 10x/0.45 objective. Imaging of the *EGFP* fluorescence was done with 975 nm infrared laser excitation using Zeiss Plan-Apochromat 63x/1.4 Oil objective. Acquired images were processed in ImageJ (Rasband, 1997-2010).

2.8.3. Two-photon Laser Scanning Microscope

Dechorionated embryos were embedded in air-permable halocarbon oil, mounted on a standard microscopy slide and covered with a cover slip placed on 170 µm support. Samples were imaged on Bio-Rad two-photon laser scanning microscope. Samples were identified in bright field using Nikon PlanApo 20x/0.75 DIC objective. Imaging of the *EGFP* fluorescence was done with 975 nm infrared laser excitation using Nikon Sfluor 40x/1.3 Oil DIC objective. Acquired images were processed in ImageJ (Rasband, 1997-2010).

2.8.4. Spinning Disk Confocal Microscope

Dechorionated embryos were embedded in 1% low gelling temperature agarose (Sigma type-VII) and mounted on cell culture dish. After the agarose has solidified, the dish was filled with PBS. Samples were imaged on Andor Revolution XD spinning disk confocal microscope. Samples were identified in bright field using Olympus UPlanSApo 10x/0.4 objective. Imaging of the *EYFP* fluorescence was done with

488 nm laser excitation using Olympus UPlanSApo 60x/1.20W objective. Acquired images were processed in ImageJ (Rasband, 1997-2010).

2.8.5. Selective Plane Imaging Microscope

Dechorionated embryos were immersed in PBT with 1:1000 dilution of 0.5 µm yellow fluorescent beads (Estapor F-Y050). Warm 2% low gelling temperature agarose (Sigma type-VII) was added to the tube and sample was mixed by vortexing for 1 second. Agarose containing embryos was drawn from the tube into 20 µm capillaries with a piston. Filled capillaries were immersed in PBS and incubated at room temperature for 10 minutes. Finally, the capillaries were mounted on the Zeiss Selective Plane Imaging Microscope. Imaging chamber was filled with PBS. The imaging temperature was maintained at 18°C. Samples were identified in bright field using Zeiss ACHROPLAN 20x/0.5W objective. Imaging of the *EYFP* fluorescence was done with 488 nm laser sheet excitation using Zeiss ACHROPLAN 20x/0.5W objective. Acquired images were processed in FIJI (Schindelin et al., 2008-2010).

Chapter 3.
3 Results

Contents

3.1. Microscopy techniques

The microscopy technique used for imaging a whole organism live must ensure penetration throughout the sample, high spatial and temporal resolution, environmental control and low phototoxicity. To find the most suitable microscopy technique, *GFP-NLS* and *HisEYFP* embryos were imaged using single-photon confocal microscope, two-photon laser scanning microscope, spinning-disk microscope and selective plane illumination microscope. Advantages and drawbacks of these techniques are summarized in table 3.1.

Technique	Acquisition speed	Resolution		Penetration	Bleaching	Image processing
		lateral	axial			
Confocal	Low	High	Low	Medium	High	Low
Two-photon	Very low	High	Medium	High	Medium	Low
Spinning disk	High	High	Very low	Low	Low	Low
SPIM	High	High	Medium	Medium	Low	Low
SPIM multiview	High	High	High	High	Low	High

Table 3.1.: Comparison of confocal, spinning-disk, two-photon and SPIM microscopy

Confocal microscopy (figure 3.1a) provided high lateral resolution data, however both speed of the stack acquisition and sample penetration were relatively low. It took over 5 minutes to acquire a stack reaching up to half of the embryo depth. Moreover, during time-lapse acquisition significant fluorophore bleaching was observed. Increasing excitation laser power reduced stack acquisition time but introduced even more photo-bleaching. The sample penetration was limited by scattering of both excitation and emission light in the embryonic tissue.

The two-photon microscopy (figure 3.1b), where near-infrared light is used for excitation allowed to increase the sample penetration. Lower energy of the exciting photons resulted in lower bleaching (Helmchen and Denk, 2005). The images acquired with the two-photon microscope exhibited lateral resolution comparable to the single-photon confocal, higher axial resolution and better signal to noise ratio. Although the sample penetration was significantly increased, it was only possible to reach up to two-thirds of the embryo depth, mostly due to emitted fluorescent light being scattered by the tissue. The acquisition time of a single stack was even higher than in the confocal microscope, reaching up to 15 minutes.

The Nipkow disk (spinning disk) microscope (figure 3.2, movie 2) enabled high-speed imaging. Thanks to CCD-based image acquisition, frame-rates up to 30 images per second were possible, which is over ten times higher than in PMT-based

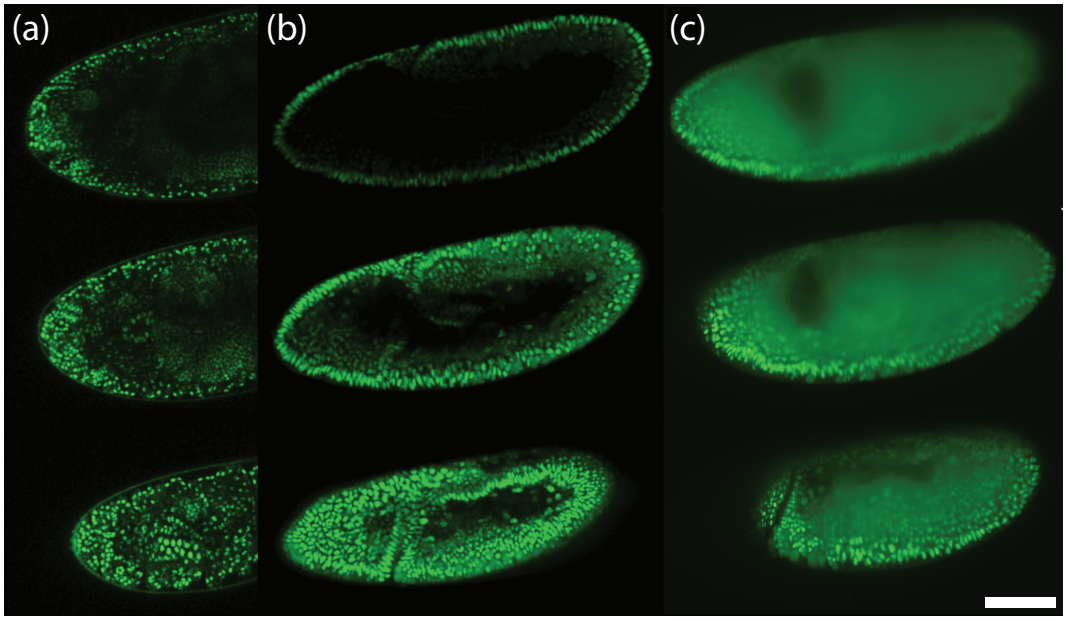


Figure 3.1.: Comparison of raw images acquired with single-photon, two-photon and SPIM microscopy

The *GFP-NLS* flies imaged with (a) Zeiss LSM 405/594 confocal, (b) Bio-Rad two-photon laser scanning microscope, and (c) Zeiss Selective Plane Imaging Microscope. Imaging was done as described in section 2.8. The image shows three optical sections through the embryo for each technique. Scale bar is 100 μm.

laser scanning microscopes. Despite the gain in acquisition speed and good lateral resolution, spinning-disk microscope had the lowest penetration and poor axial resolution.

The Single Plane Illumination Microscope, similarly to the spinning-disk microscope features high frame rates (up to 5 frames per second in the set-up used) thanks to CCD-based image acquisition. Sample penetration and lateral resolution were comparable to the confocal microscope, however large amount of artifacts affected the overall image quality (figure 3.1c).

The moderate z-resolution, penetration and SPIM artifacts were complemented by the ability to image the sample from multiple angles. Multiview imaging combined with recently developed image processing solutions for bead-based registration (Preibisch et al., 2010) and content-based fusion (Preibisch et al., 2007) enabled visualizing the embryo *in toto* with isotropic resolution.

A series of tests using both *GFP-NLS* and *HisEYFP Drosophila* strains have shown, that the speed of embryo development in fruitflies exceeds temporal resolution of the SPIM microscope. Temporal resolution could be increased either by faster imaging, or by slowing down embryo development. The first approach, although more logical, was

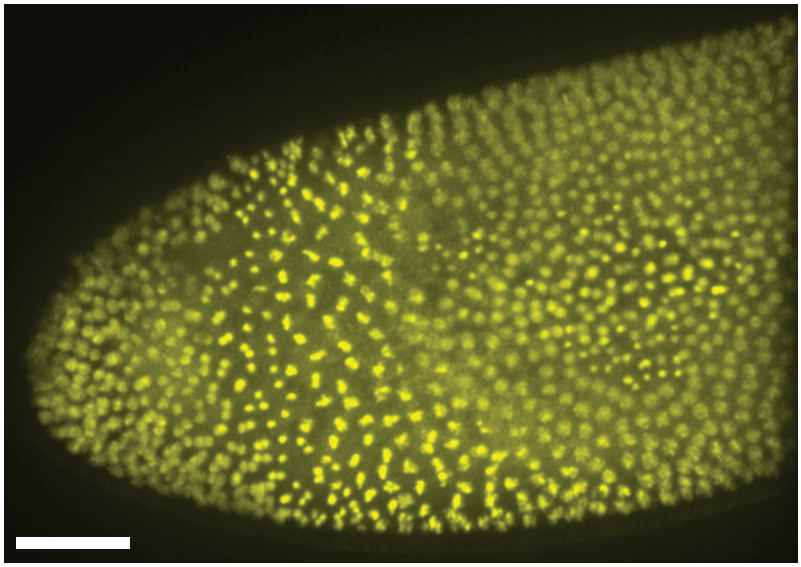


Figure 3.2.: Early *HisEYFP* embryo imaged with a spinning-disk microscope

The *HisEYFP* embryos were imaged with Andor Revolution XD system. Imaging was done as described in section 2.8. Scale bar is 50 µm.

impossible to achieve due to limitations of the SPIM set-up. Out SPIM microscope is equiped with an old generation laser module that limits effective acquisition frame rate. Since the set-up is a loan from Zeiss the laser module cannot be replaced with a new one.

Drosophila development speed is a temperature dependent process. It can be slowed down by decreasing the system temperature within a reasonable range. The commonly used temperature of 18°C allows to slow down fly development and increase lifespan about twofold (Ashburner et al., 2005). The SPIM set-up was equipped with a temperature control unit based on a brass chamber and a Peltier element. Although the system was performing well in maintaining temperatures higher or slightly lower than ambient temperature, it failed in cooling the sampling chamber to temperatures below 20°C. Introducing a water bath heat-sink to the system allowed to cool the sampling chamber even down to 10°C.

The brass used for the chamber construction is an ideal heat-conductor, however copper, which is the main component of brass, tends to corrode when salt-based media (like PBS) are used for imaging. To overcome the corrosion, the entire chamber has been plated with gold. This solution preserved the good heat-conducting properties of the chamber and made it chemically inert. Such modified SPIM set-up was able to produce long time-lapse movies of the whole *Drosophila* embryonic development (figure 3.3, movies 3–5), making it the system of choice for *in toto* live imaging.

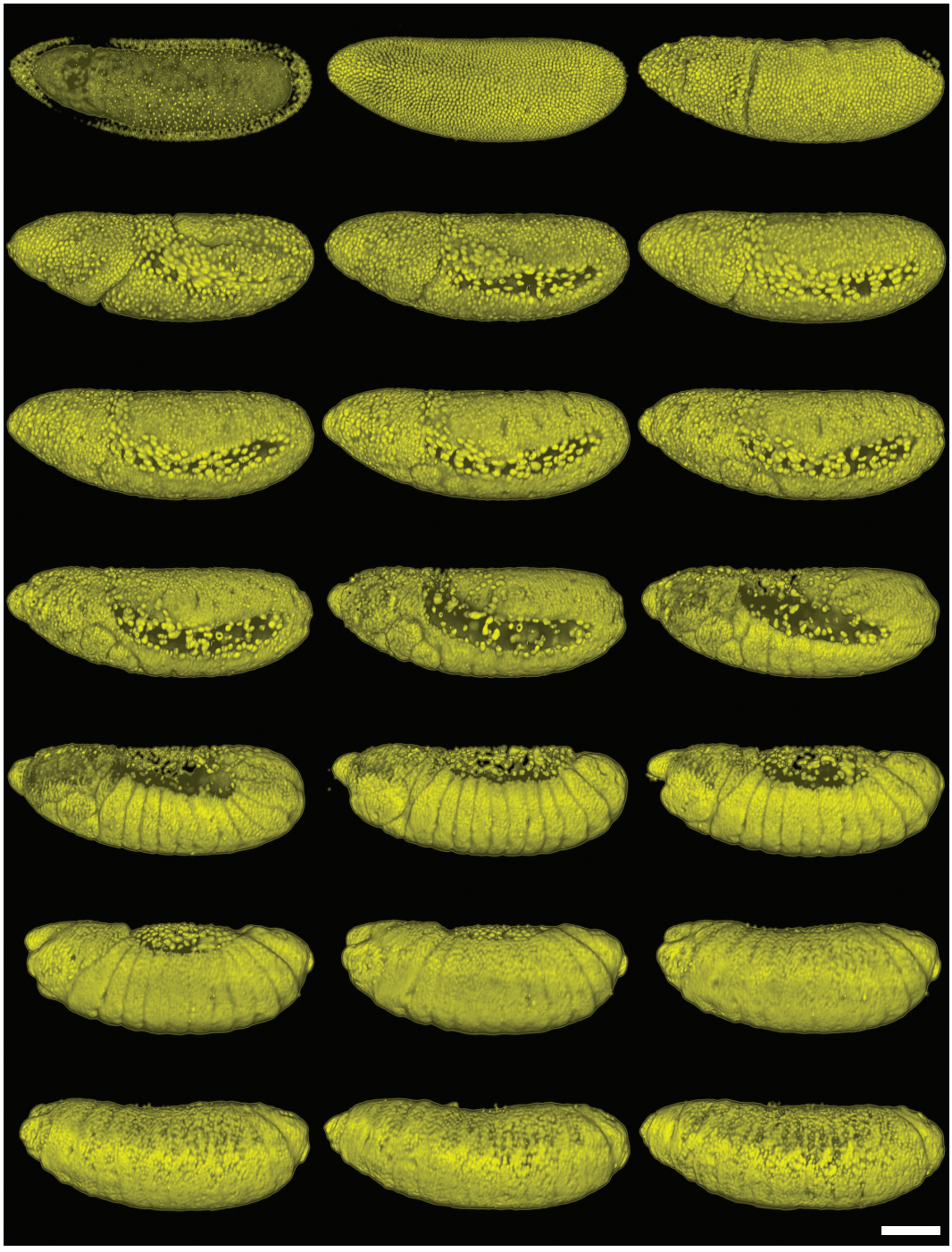


Figure 3.3.: *Drosophila* embryonic development captured with SPIM

The whole development of the *HisEYFP* embryo captured with SPIM. The image shows 21 out of 210 time-points acquired with 5-minutes interval. The interval between presented time-points is 50 minutes. The embryo was imaged from six angles. Image stacks acquired from different angles and between time-points were registered and fused as described by Preibisch et al. (2010). Images were acquired by M. Weber as described in section 2.8. Scale bar is 100 µm.

3.2. Markers for *in vivo* embryo imaging

Tracking cellular behavior during development requires a nuclear marker that is visible during the whole cell cycle. The *Histone 2AvD-EYFP* fusion protein has been constructed for this purpose. The construct has been cloned in a pCaSpeR4 vector (figure 3.4) and used for fly transformation. The obtained *HisEYFP* stock was homozygous viable. Flies from established *HisEYFP* stock were used for imaging with SPIM. The *His2AvD-EYFP* marker exhibits approximately twice more fluorescence than *GFP-NLS* construct used before. In SPIM image acquisition, 150 ms exposure with 488 nm laser was enough to saturate *HisEYFP* images, compared to 300 ms required for *GFP-NLS* marker, even though suboptimal excitation wavelength for EYFP (optimal is 514 nm) was used. The brightness of *EYFP* is nearly twice as high as that of *GFP*, as reported by Shaner et al. (2005). Moreover, since histone is strongly bound to DNA, it produces intensive signal in small volumes occupied by the nucleic acid, whereas *GFP-NLS* signal is dispersed in the whole nuclear volume. While *Histone-EYFP* is detectable during the whole cell cycle (figure 3.5), the *GFP-NLS* gets released into the cytoplasm during the cell division, which makes dividing cells untraceable.

In order to aid in tracking cellular behavior during embryo development, the second marker was introduced. *E-cadherin* is a member of a complex forming intercellular adherens junctions in epithelial cells (Steinberg and McNutt, 1999). Since during early development the embryo is mainly composed of epithelial tissue and adherens junctions are abundant (Oda and Tsukita, 2001), the *E-cadherin* based reporter marks apical membrane in the majority of cells. The *DE-cadherin-ECFP* was cloned together with *HisEYFP* in pFlyFos vector (figure 3.6) and used for φC31-mediated transgenesis of the *attP40* fly line. Flies expressing *CadECFP* were imaged with SPIM (figure 3.7, movie 6). The marker performed well in imaging, although exposure times significantly longer than for *HisEYFP* were required to saturate images (300 ms compared to 100 ms). The combined markers are suitable for dual-channel imaging. Since the pFlyFos vector utilizes φC31 integrase for transgenesis, the construct can be integrated into virtually any *Drosophila* species.

3.3. Fosmid genomic libraries

Tracking gene expression during development requires a reliable source of genes expressed in native genomic constructs. Since the genes have to be modified to facilitate expression pattern visualization, use of their chromosomal copies for that purpose is impractical. Therefore, genomic libraries were constructed in a fosmid

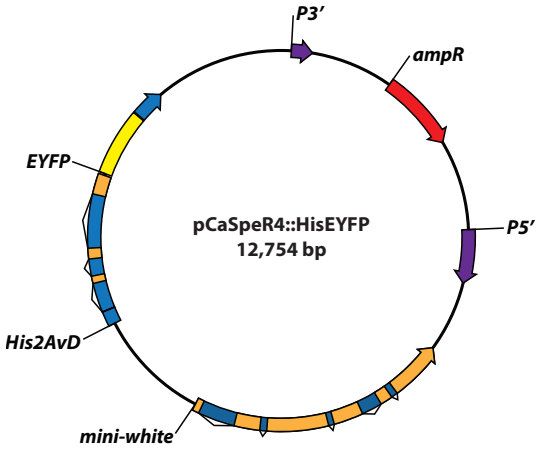


Figure 3.4.: pCaSpeR4::HisEYFP construct map

The His2AvD-EYFP construct was cloned in pCaSpeR4 vector as described in section 2.2.1. The construct was used for P-element transgenesis of w^- flies.

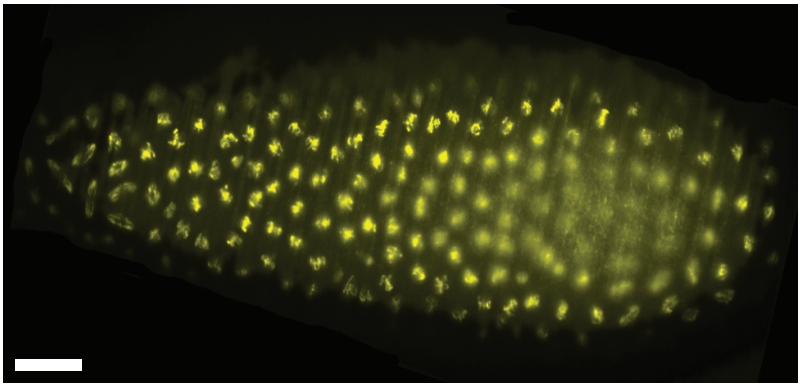


Figure 3.5.: *HisEYFP* flies imaged with SPIM

The blastoderm stage embryo expressing *HisEYFP* was imaged with Zeiss SPIM. Images were acquired from six angles. Image stacks were registered and fused as described by Preibisch et al. (2010). Scale bar is 50 µm.

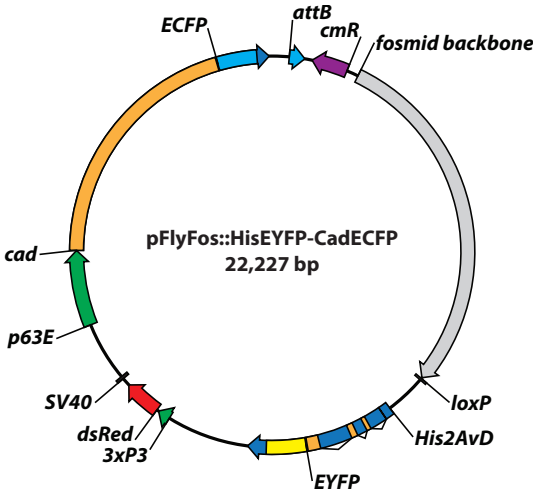


Figure 3.6.: pFlyFos::HisEYFP-CadECFP construct map

The DE-Cadherin-ECFP construct was constructed as described in section 2.2.2 and cloned together with His2AvD-EYFP in pFlyFos vector (section 2.2.3). The construct was used for φC31-mediated transgenesis of *attP40* flies.

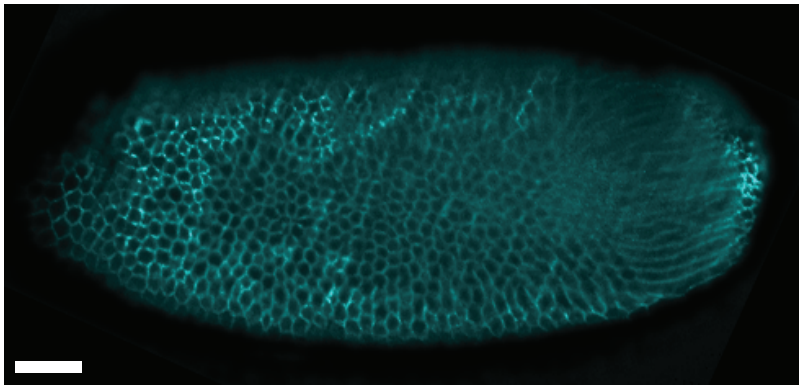


Figure 3.7.: Cadherin-ECFP membrane marker imaged with SPIM

The cellular blastoderm stage embryo expressing *CadECFP* was imaged with Zeiss SPIM. Images were acquired from six angles. Image stacks were registered and fused as described by Preibisch et al. (2010). Scale bar is 50 µm.

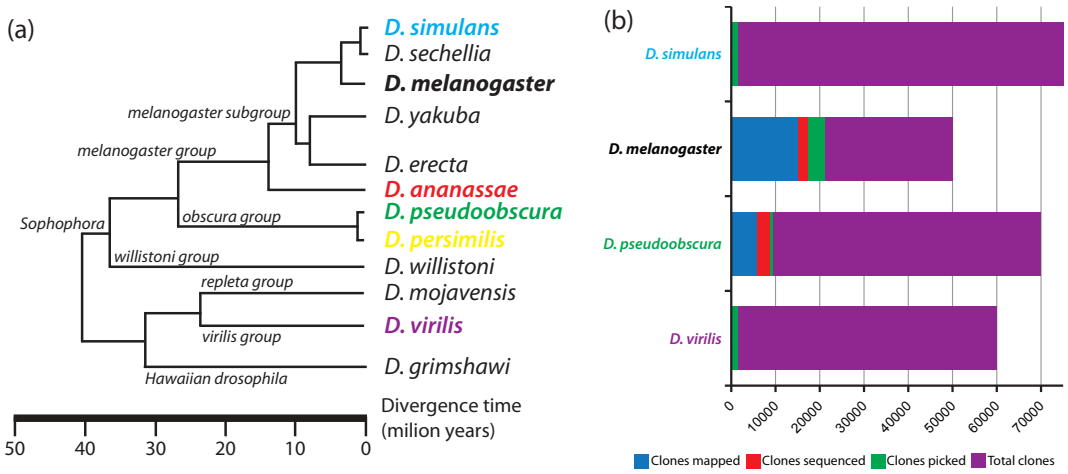


Figure 3.8.: Status of the fosmid library production

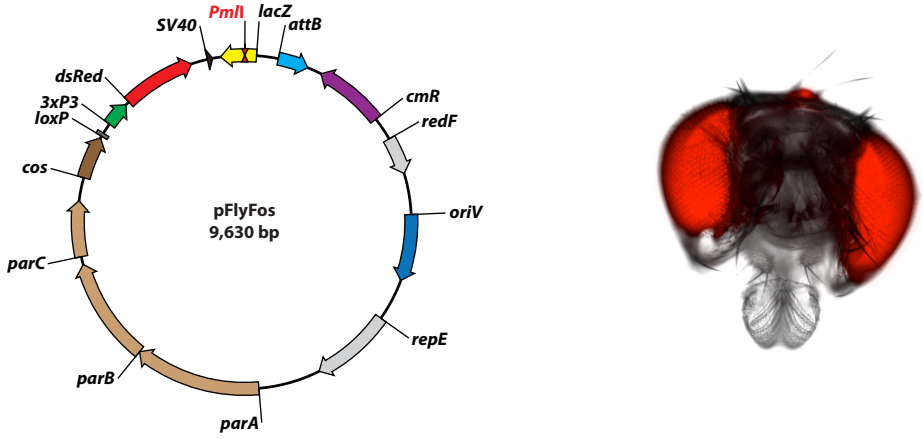
(a) The phylogenetic tree of Drosophilidae with highlighted species that were used or are planned for genomic library construction. (b) Status of the library production. The libraries for *D. melanogaster* and *D. pseudoobscura* have been characterized. The libraries for *D. simulans* and *D. virilis* still remain uncharacterized.

vector to provide constructs containing genes in their native genomic context. Since the focus of the project is to provide a cross-species toolkit, the libraries were constructed for four species of a widespread divergence: *Drosophila melanogaster*, *Drosophila simulans*, *Drosophila pseudoobscura*, and *Drosophila virilis*. Further two libraries for *D. ananassae* and *D. persimilis* are planned (figure 3.8).

The libraries were constructed in a novel fosmid vector (pFlyFos – figure 3.9) containing the *attB* sequence recognized by φC31 integrase and the eye promoter-driven dominant selectable marker (*3xP3-dsRed-SV40*) for isolation of transformants. Thanks to the arabinose-inducible origin of replication (*oriV*) the fosmid can be induced from single copy to moderate copy number for efficient DNA preparation. Since both φC31 integrase (Groth et al., 2004) and *3xP3-dsRed* selectable marker (Horn et al., 2002) are active in a variety of species, the constructs cloned in pFlyFos vector can be utilized in a wide range of model organisms. The vector has been tested in fly transgenesis experiment (section 2.3.1). The *dsRed* fluorescence in the eyes and ocelli was observed in 10.3% of fertile G_1 crosses when empty vector was injected.

3.3.1. *Drosophila melanogaster* library

The *Drosophila melanogaster* library was first out of two that were characterized. Out of approximately 50,000 clones in the genomic library 21,120 clones were picked and 17,280 clones were analyzed as described in section 2.3.4. Nearly 88% of the analyzed clones (15,204) were successfully mapped to the *Drosophila melanogaster*



(a) (b)

Figure 3.9.: The pFlyFos vector

(a) The pFlyFos vector used for library production. The fosmid contains *3xP3-dsRed-SV40* selectable marker and *attB* site for integration into the host genome. The vector backbone harbours chloramphenicol resistance gene (*cmR*), resolvase (*redF*), *oriV* for high-copy induction, replication initiation gene (*repE*), genes required for fosmid partitioning during cell division (*parA*, *parB*, *parC*) and *cos* site for phage packaging. (b) dsRed fluorescence in the eyes used for transformant screening.

genome, for 8.25% clones (1,426) one of the sequencing reactions failed, 3.65% of clones (630) are likely chimeric (figure 3.10a). The average size of the genomic insert was 36 kb, with 95% of clone sizes between 20 kb and 50 kb (figure 3.10c). The library covers all chromosomes, including heterochromatin regions (figure 3.10e) with average coverage of 3.3-fold. The mapped clones contain sequence of 89.27% of annotated *D. melanogaster* genes with at least 10 kb of upstream an 5 kb of downstream sequence. The actual number of genes included in the library was compared to the computer prediction described in section 2.3.2 (figure 3.10g). Seven percent of the *D. melanogaster* genes were too long to be included in the fosmid library. Therefore, the library covers 96% of the genes that can in principle be included in the fosmid clones.

3.3.2. *Drosophila pseudoobscura* library

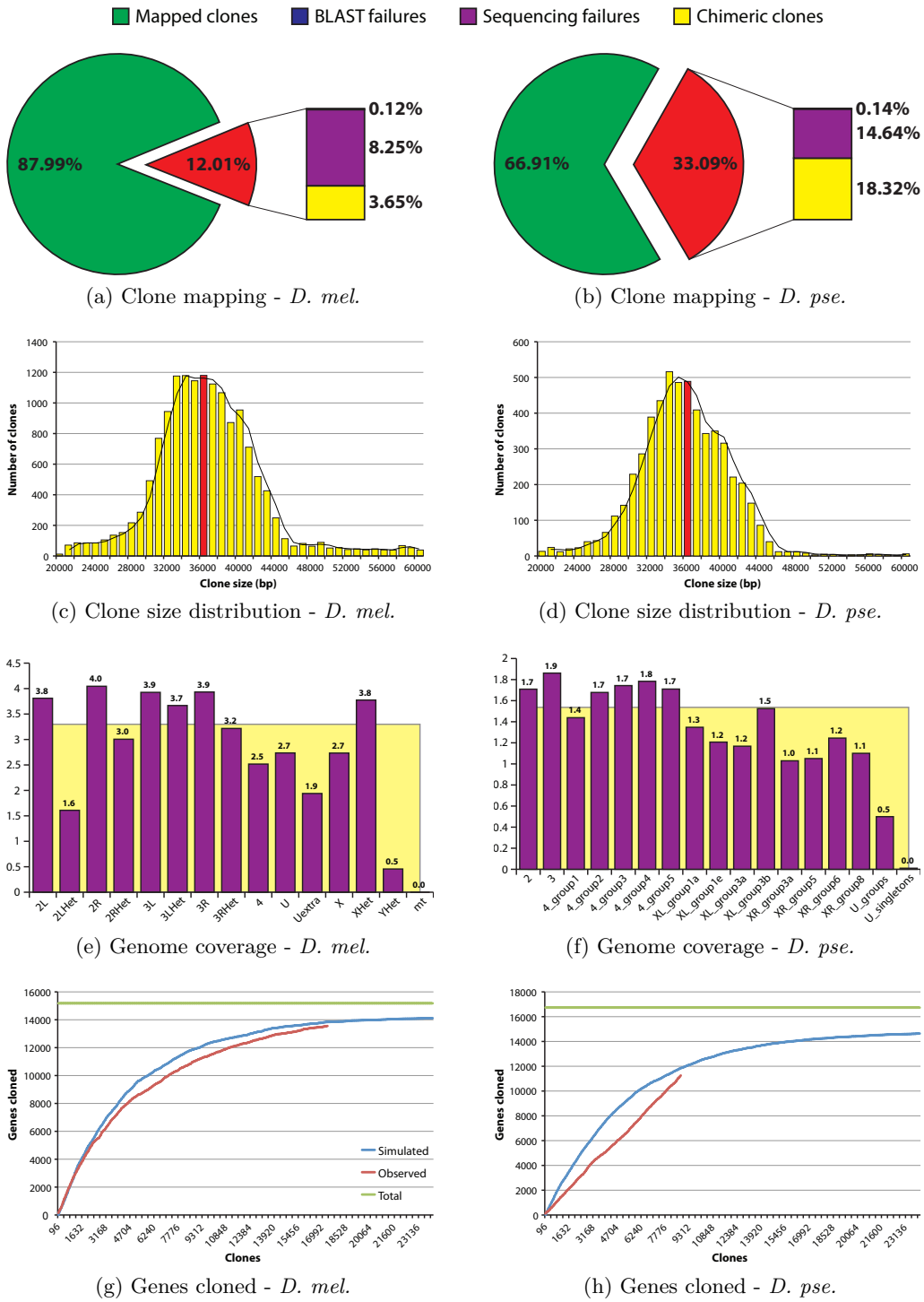
The second characterized library was the one for *Drosophila pseudoobscura*. From a total number of 75,000 clones in the library 9,504 clones were picked and 8,751 clones were analyzed similarly to clones from *D. melanogaster* library. Mapping was successful for 66.91% of clones (5845), 14.64% of clones (1,279) failed in the sequencing reaction, 18.32% of clones (1600) were reported as chimeric, however

since the assembly of the *D. pseudoobscura* genome is incomplete, most of these simply span the scaffold boundaries (figure 3.10b). The clone size distribution was very similar to the *D. melanogaster* library (figure 3.10d). The library covers all large scaffolds and partially covers unmapped groups (figure 3.10e) with an average coverage of 1.5-fold. The mapped clones cover sequence of 67.28% of annotated *D. pseudoobscura* genes with at least 10 kb of upstream and 5 kb of downstream sequence. Surprisingly, the actual gene coverage data does not follow the computer simulation (figure 3.10h). Instead of the expected logarithmic rise of the number of genes cloned in a certain number of clones, the number of genes was increasing linearly. This might be caused by the incomplete assembly of the *D. pseudoobscura* genome and as a consequence large mapping failure rate. According to the computer simulation data, 10% of *D. pseudoobscura* genes could not be included in fosmids due to their size. Therefore, the library covers 75% of the genes that can be cloned in fosmids.

3.4. Recombineering

3.4.1. *CG4702* tagging

Clones from *D. melanogaster* library were used for gene tagging with fluorescent markers. As a proof of principle, the first tagging experiment was performed on FlyFos014971 clone containing *CG4702* gene (figure 3.12a). The *CG4702* was chosen because of its dynamic expression pattern. It is highly expressed in the early stages of development with a striped ectodermal pattern. During germ band elongation the gene expression diminishes (except from ventral ectoderm primodium) to reappear, as a typical epidermal marker, after the germ band retraction is finished (Tomancak et al., 2007). The gene was tagged C-terminally with Tag[2xTY1-EGFP-3xFLAG] (figure 3.11c and C.2) and N-terminally with Tag[ubi-mCherry-NLS-T2A] (figure 3.11a and C.1). Both tagging experiments were successful, as verified by junction sequencing. FlyFos014971::CG4702-EGFP and FlyFos014971::CG4702-mCherry fosmids were used for transgenesis of the *attP40* landing line strain. Transformation efficiency was close to 1%. Activity of the fluorescent proteins was verified by simple fluorescent imaging. While *mCherry* expression was not detectable in fluorescent imaging, the *EGFP* fluorescence was observed in the late embryo. The expression pattern of *EGFP*-tagged *CG4702* gene was compared to native expression pattern by *EGFP*-specific *in situ* staining (figure 3.12b). Late *CG4702-EGFP* embryos were fixed and stained with DRAQ5 and *EGFP* antibody. Both fixed and live embryos were imaged SPIM (figure 3.12c, 3.13 and movie 7). The *EGFP* expression in the tracheal system, anal pads, salivary gland ducts and epidermis was consistent with the *in situ* data.



(a) Clone mapping - *D. mel.*

(b) Clone mapping - *D. pse.*

(c) Clone size distribution - *D. mel.*

(d) Clone size distribution - *D. pse.*

(e) Genome coverage - *D. mel.*

(f) Genome coverage - *D. pse.*

(g) Genes cloned - *D. mel.*

(h) Genes cloned - *D. pse.*

Figure 3.10.: Fosmid library statistics

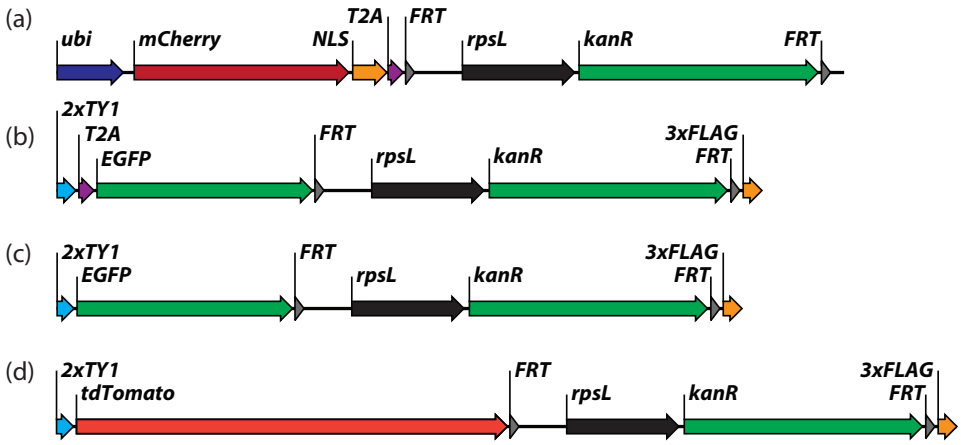


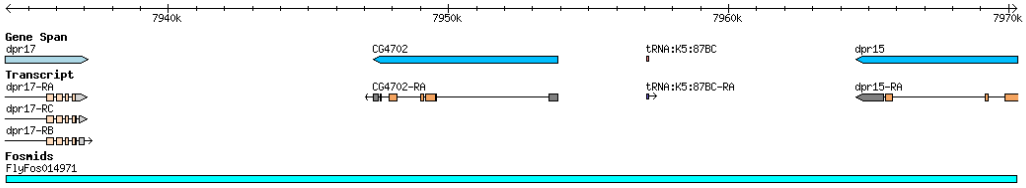
Figure 3.11.: Recombineering tags used in tagging experiments

3.4.2. High-throughput recombineering

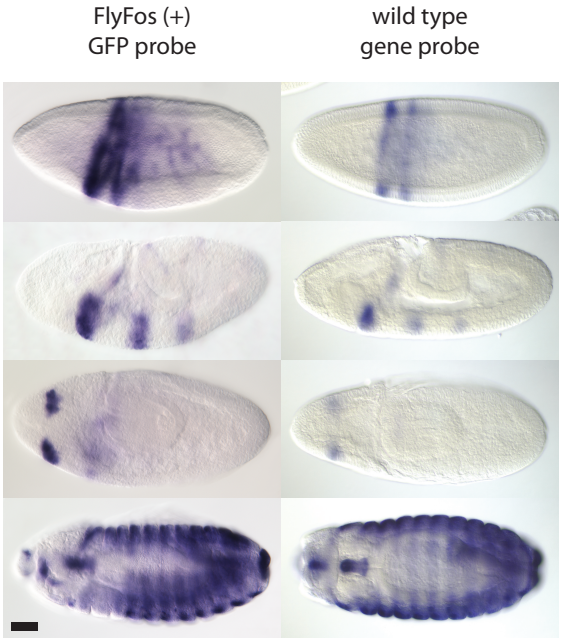
Subsequently, to bring recombineering to the high-throughput level the FlyFos clones containing 48 genes listed in table 3.2 (see section D.1 for fosmid maps) were used for C-terminal tagging with Tag[2xTY1-EGFP-3xFLAG] (figure 3.11c and C.2), Tag[2xTY1-T2A-EGFP-3xFLAG] (figure 3.11b and C.2), Tag[2xTY1-tdTomato-3xFLAG] (figure 3.11d and C.2) and N-terminal tagging with Tag[ubi-mCherry-NLS-T2A] (figure 3.11a and C.1). The genes selected for tagging were chosen so that they cover most organ systems in *Drosophila* embryogenesis (figure 3.14). Selected tissue-specific marker genes were tagged as described in section 2.4.3. The recombineering success rate improved from 75% to 100% through four tagging experiments (figure 3.15), suggesting that the method is easily scalable for genome-wide application.

3.4.3. Expression pattern analysis

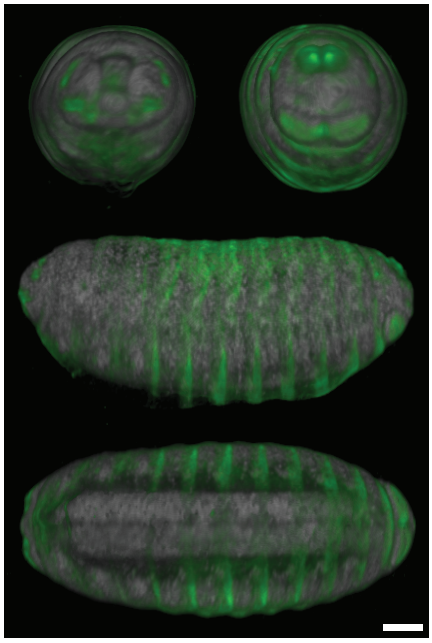
Twelve *mCherry*-tagged fosmids were injected into *attP40* landing line embryos to test whether the recombinant transgenes recapitulate wild-type gene expression patterns. Generated transgenic lines are listed in table 3.3. The expression patterns were visualized by RNA *in situ* staining using *mCherry*-specific probe (figure 3.16). Eight out of twelve transgenes (66%) recapitulated the wild-type gene expression patterns suggesting that *cis*-regulatory elements required for control of gene expression were included in the fosmids. One tagged transgene showed no expression, and three were detected in the wrong tissues possibly because of the absence of surrounding insulator elements (table 3.3). The transgene expression could be visualized by different techniques, such as RNA *in situ*, immunofluorescence and fluorescent protein imaging (figure 3.17).



(a)



(b)



(c)

Figure 3.12.: *CG4702* gene tagged with *EGFP*

(a) FlyBase map of FlyFos014971 clone used for *CG4702* tagging. (b) In situ staining of *CG4702-EGFP* embryos using *EGFP*-specific probe compared to the native expression pattern from the Atlas of Patterns of Gene Expression (Tomancak et al., 2007). (c) A three-dimensional rendering of a late-stage embryo *CG4702-EGFP* embryo stained with DRAQ5 nuclear marker imaged with a single plane illumination microscope. Shown are the frontal (top left), caudal (top right), lateral (middle) and ventral (bottom) views of the same embryo. Scale bars are 50 µm. Figure reproduced from Ejsmont et al. (2009).

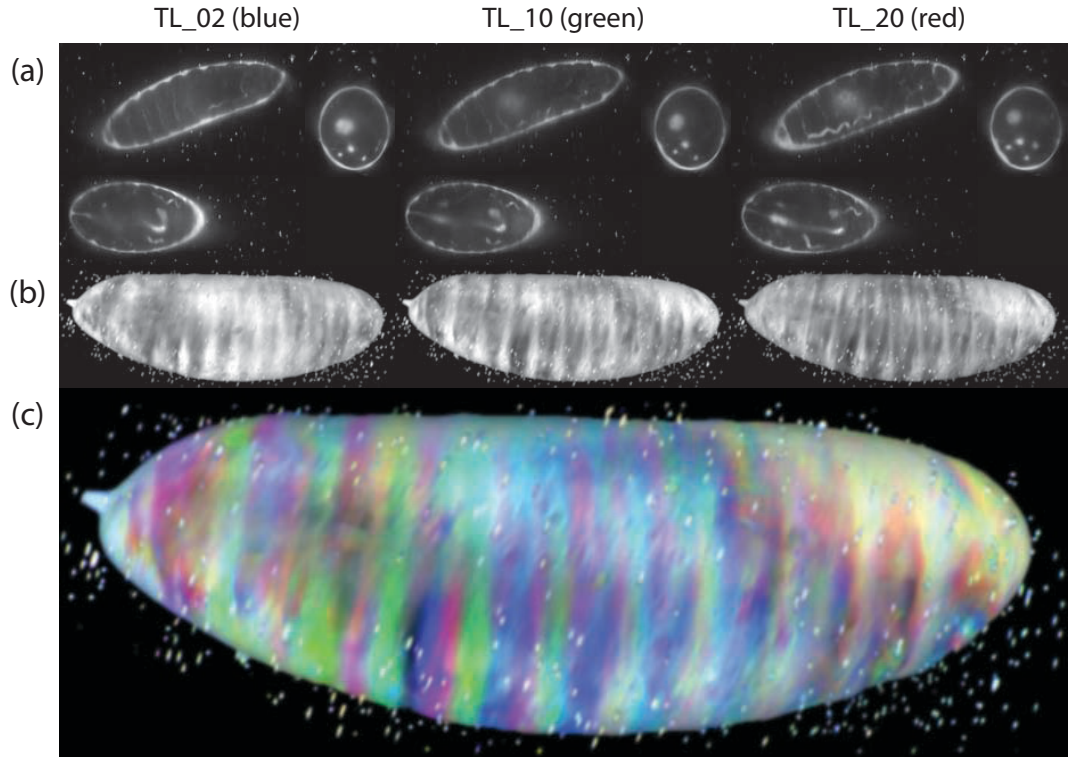


Figure 3.13.: *CG4702-EGFP* embryos imaged with SPIM

(a) Orthogonal sections (xy, xz, yz) through reconstructed SPIM embryo acquisitions for three time-points approximately 50 minutes apart. (b) 3-D rendering of the SPIM multi-view reconstruction of the embryo at the three time-points. (c) The three time- points color-coded (red, green and blue) and overlaid in a single 3-D rendering, showing morphological changes that occurred during imaging. Only the epidermal expression is visible in 3-D renderings (b,c) due to transparency settings. Image stacks were registered and fused as described by Preibisch et al. (2010). Fluorescent beads that were used as fiduciary markers for SPIM multi-view reconstruction surround the embryos. Figure reproduced from Ejsmont et al. (2009).

Well	Gene	Clone
A01	*odd*	FlyFos016895
A02	*numb*	FlyFos015836
A03	*CG6113*	FlyFos015925
A04	*sna*	FlyFos015520
A05	*Ugt36Bc*	FlyFos015822
A06	*CG1962*	FlyFos016667
A07	*CG9336*	FlyFos015601
A08	*Spn43Aa*	FlyFos016654
A09	*Optix*	FlyFos016694
A10	*CG8193*	FlyFos016218
A11	*CG8850*	FlyFos016487
A12	*Lac*	FlyFos016233
B01	*Mp20*	FlyFos016401
B02	*CG17041*	FlyFos016401
B03	*CG5493*	FlyFos016960
B04	*CG9416*	FlyFos016005
B05	*Obp56a*	FlyFos016413
B06	*CG13506*	FlyFos017141
B07	*ken*	FlyFos015857
B08	*gsb-n*	FlyFos015278
B09	*CG13920*	FlyFos016980
B10	*CG12011*	FlyFos015257
B11	*CG10591*	FlyFos015827
B12	*CG32354*	FlyFos015127

Well	Gene	Clone
C01	*Doc3*	FlyFos016847
C02	*CG14110*	FlyFos016260
C03	*mex1*	FlyFos015648
C04	*comm2*	FlyFos016035
C05	*disp*	FlyFos016541
C06	*pyd3*	FlyFos016094
C07	*CG8147*	FlyFos016563
C08	*CG14687*	FlyFos016339
C09	*mfas*	FlyFos015057
C10	*Ect3*	FlyFos015057
C11	*CG15887*	FlyFos015939
C12	*Fsh*	FlyFos015812
D01	*CG7678*	FlyFos016718
D02	*CG18594*	FlyFos014991
D03	*CG13653*	FlyFos016224
D04	*HLHm5*	FlyFos015754
D05	*CG14253*	FlyFos015266
D06	*Obp99a*	FlyFos015387
D07	*PH4alphaSG2*	FlyFos016922
D08	*l(1)sc*	FlyFos015631
D09	*CG4194*	FlyFos016428
D10	*CG15209*	FlyFos015174
D11	*wgn*	FlyFos016415
D12	*run*	FlyFos015034

Table 3.2.: Fosmid clones used for high-throughput tagging

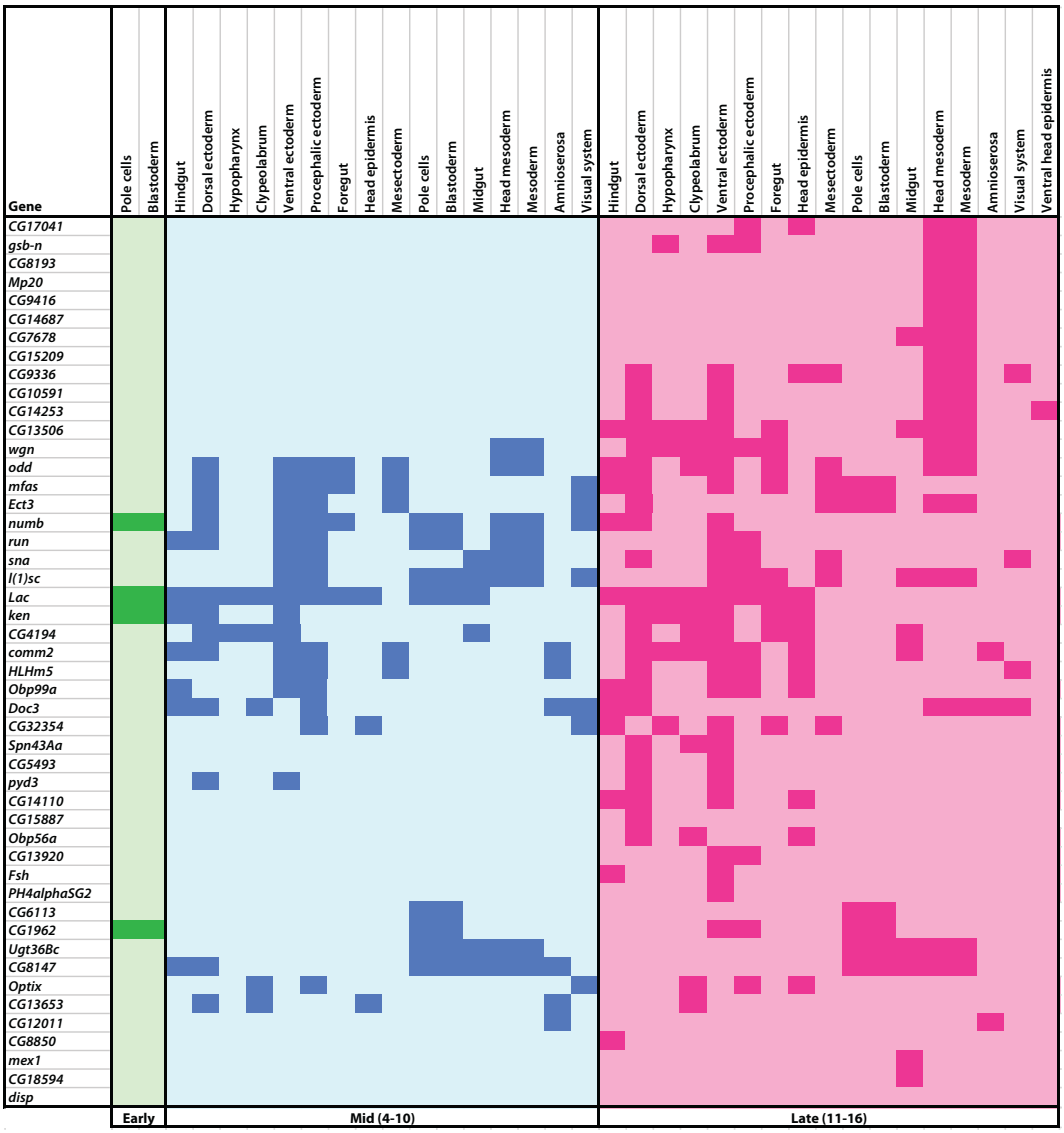


Figure 3.14.: Genes selected for high-throughput tagging and their expression patterns

Hierarchical clustering of gene expression pattern annotations for 48 genes selected for tagging experiments. This set maximizes the coverage of organ systems at all stages of embryogenesis. Figure reproduced from Ejsmont et al. (2009).

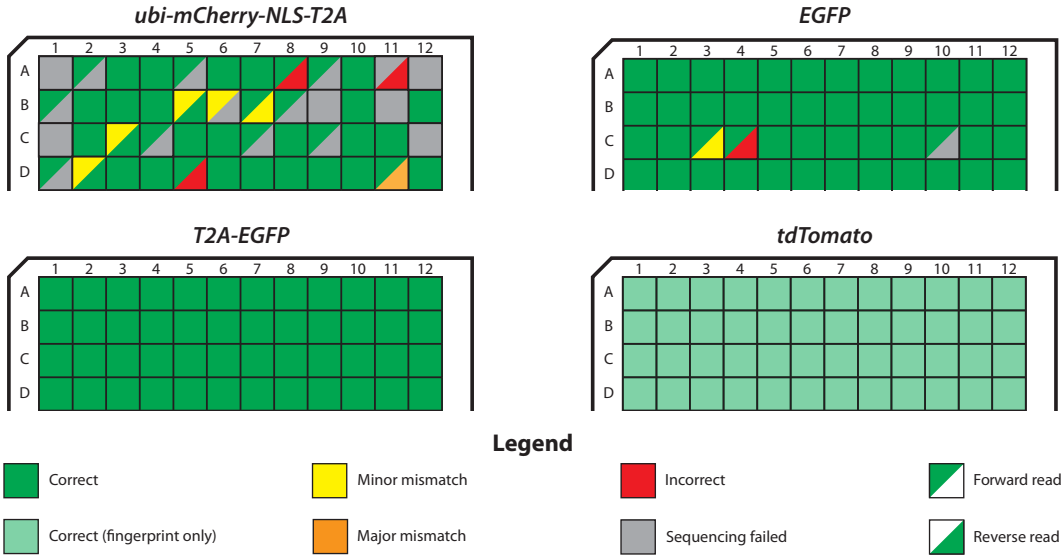


Figure 3.15.: High throughput tagging results

The structure of the modified fosmids was verified by two sequencing reactions with primers extending from within the tag toward the tag-fosmid junction. See section 2.4.3 and appendix E for details.

Gene	*mCherry-T2A-NLS*	*EGFP*	*T2A-EGFP*	*in situ* result
CG1962	+/+	+/+		ectopic expression
CG9336	+/+			recapitulates native pattern
Spn43Aa	+/+	+/+		ectopic expression
Optix	+/+	+/+		no expression
CG8193	+/+			recapitulates native pattern
CG5493	+/+			recapitulates native pattern
CG12011	+/+			recapitulates native pattern
CG32354	+/+			recapitulates native pattern
pyd3	+/+			recapitulates native pattern
CG15887	+/+			ectopic expression
HLHm5			+/+	not assayed
wgn	+/+			recapitulates native pattern
CG4702	+/+	+/+		recapitulates native pattern

Table 3.3.: Transgenic lines established using modified fosmids

All generated lines were homozygous viable, and therefore marked as +/+ in the above table. Results of transgene expression pattern validation are summarized in the last column. All genes except *CG4702* were assayed with *mCherry*-specific *in situ* probes. The *CG4702* was assayed with *EGFP*-specific probe.

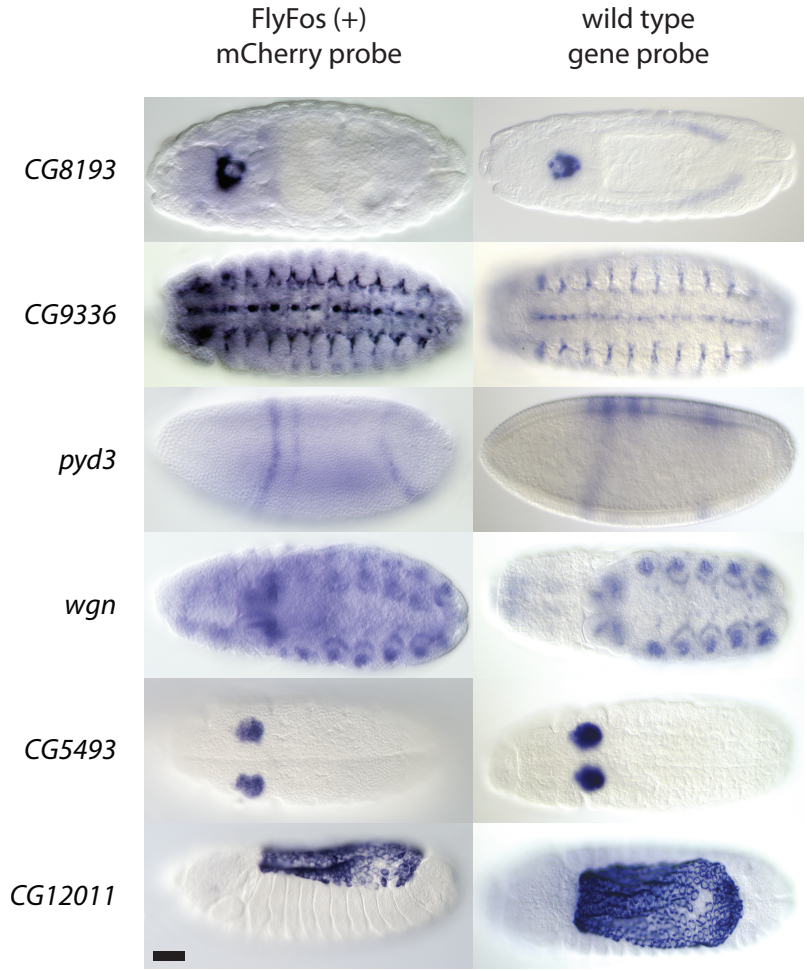


Figure 3.16.: RNA *in situ* for mCherry-tagged genes

Examples of RNA *in situ* patterns recapitulated by the fosmid transgenes tagged with *mCherry* compared to the native expression patterns. Scale bar is 50 μm. Figure reproduced from Ejsmont et al. (2009).

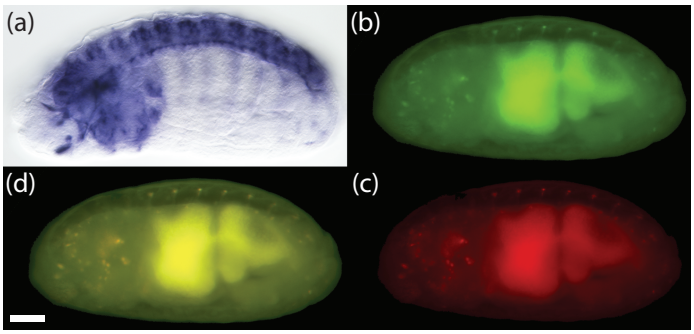


Figure 3.17.: *mCherry* expression visualized by different techniques

The *ubi-mCherry-NLS* transgene expression visualized by RNA *in situ* (a), immunofluorescence (b), mCherry fluorescence (c), merge image (d). Scale bar is 50 μm.

3.5. RNAi rescue

3.5.1. Bioinformatics analysis

Clones from *D. pseudoobscura* library were used in cross-species RNAi rescue experiment. *Drosophila pseudoobscura* has been chosen as a source of rescue constructs since it provides a good balance between sequence divergence and conservation of the gene function in respect to *Drosophila melanogaster*. As described in section 2.5.1, sequences from non-melanogaster species homologous to hairpin sequences from *D. melanogaster* RNAi library (Dietzl et al., 2007). The hairpin sequence similarity followed the phylogeny (3.18a and 3.18b). The largest sequence conservation was observed in *D. simulans*. The sequences are not only nearly identical (94.75% are more than 90% conserved) but also the vast majority of uninterrupted stretches of identity (as illustrated in figure 3.18c) were longer than 18 nucleotides (figure 3.18d), which is known to be enough for RNAi induction (Kulkarni et al., 2006; Ma et al., 2006; Perrimon and Mathey-Prevot, 2007). The sequence homology quickly drops outside of the *melanogaster* subgroup. The number of genes that are conserved in 90% decreases from 2,98% for *D. ananassae*, through 1,78% for *D. pseudoobscura*, 1.63% for *D. persimilis* down to 0,41% for *D. virilis*. Basing on the number of identical sequence stretches shorter than 19 nucleotides, a fraction of genes that is likely to be RNAi-refractory was estimated as shown in figure 3.18e. Since 94% of the RNAi-refractory orthologs could be found either *D. pseudoobscura* or *D. virilis*, which are established model systems, they are both well suited to serve as a donor for RNAi rescue experiment from the sequence divergence point of view. Besides sequence divergence, the second important criterion for successful RNAi rescue is the ability of the transgene to complement the RNAi phenotype. Comparative microarray time-course analysis of embryogenesis revealed that 24.7% of *D. virilis* genes exhibits differential gene expression profiles relative to *D. melanogaster* compared to 18.8% for *D. pseudoobscura* (Kalinka et al., submitted manuscript). Based on this, *D. pseudoobscura* genomic transgenes are more likely to complement *D. melanogaster* loss-of-function phenotypes and are thus best suited for RNAi rescue.

3.5.2. Clone selection

Since at the time of the experiment the *D. pseudoobscura* library was not fully characterized, a PCR-based strategy allowing to identify clones containing genes of interest was established (section 2.5.2). To validate the pooling approach, the pool plates were screened with a primer pair targeting *D. pseudoobscura* ortholog of *D. melanogaster* gene *Mical* to identify the library plates harboring a potential *Mical*

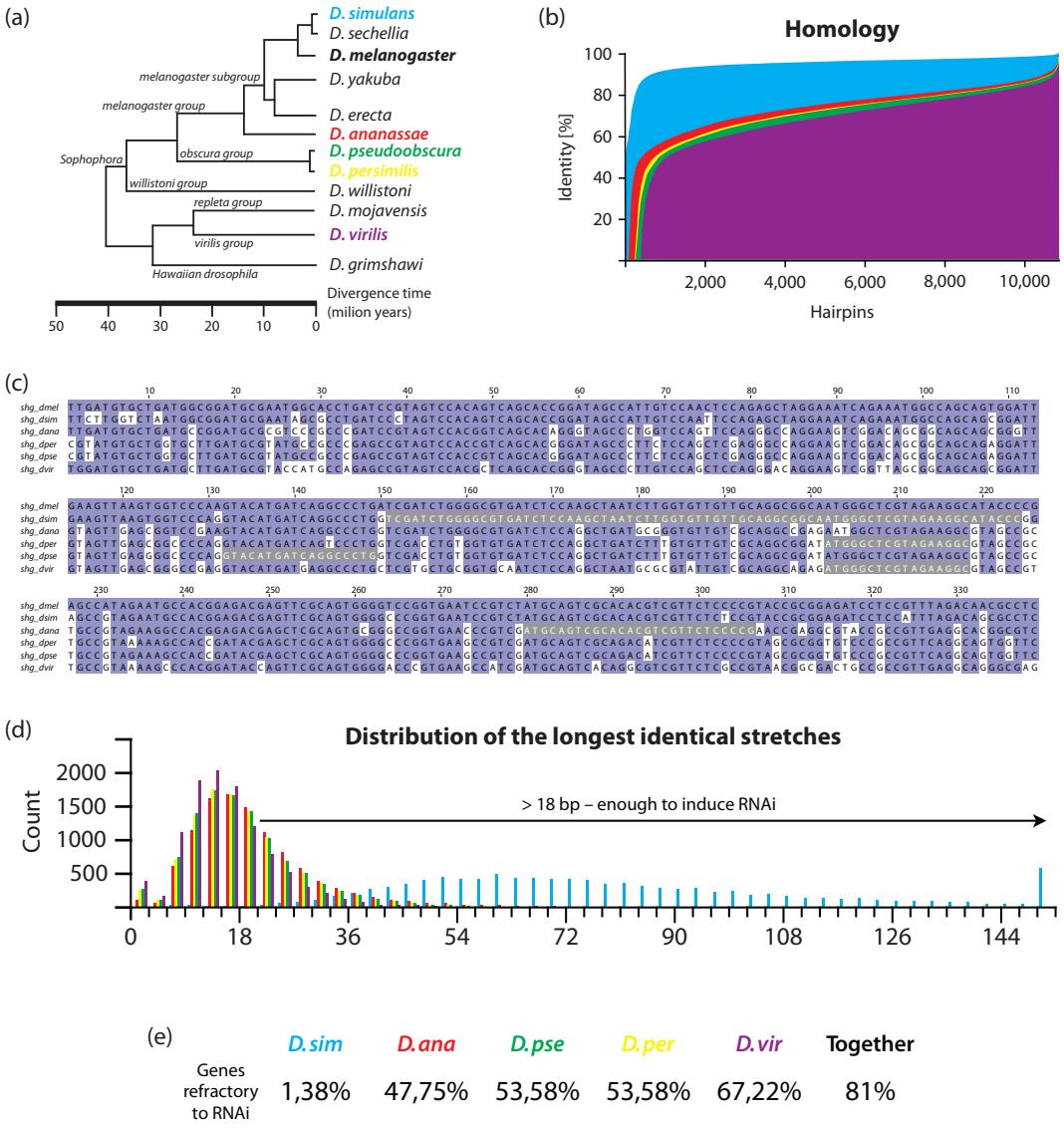


Figure 3.18.: Bioinformatics analysis of hairpin sequence divergence in Drosophilidae

(a) Phylogenetic tree of twelve sequenced *Drosophila* species. (b) Summary of the conservation of RNAi hairpins in pair-wise genome alignments with *D. melanogaster* as common reference. The percentage of nucleotides identical across the hairpin alignment (y-axis) is plotted for all hairpins ordered by increasing conservation (x-axis). Species are colour-coded according to (a and e). (c) An example of 6 species multiple sequence alignment for a hairpin targeting *shotgun* (*shg*). Nucleotides identical to *D. melanogaster* are shaded in blue. The longest uninterrupted stretch of identical nucleotides is shaded grey for each species. (d) Histogram of longest uninterrupted stretches for all hairpins binned in size groups of 3. (d) Number of genes likely to be RNAi-refractory in each of the analyzed species. Figure reproduced from Langer et al. (2010).

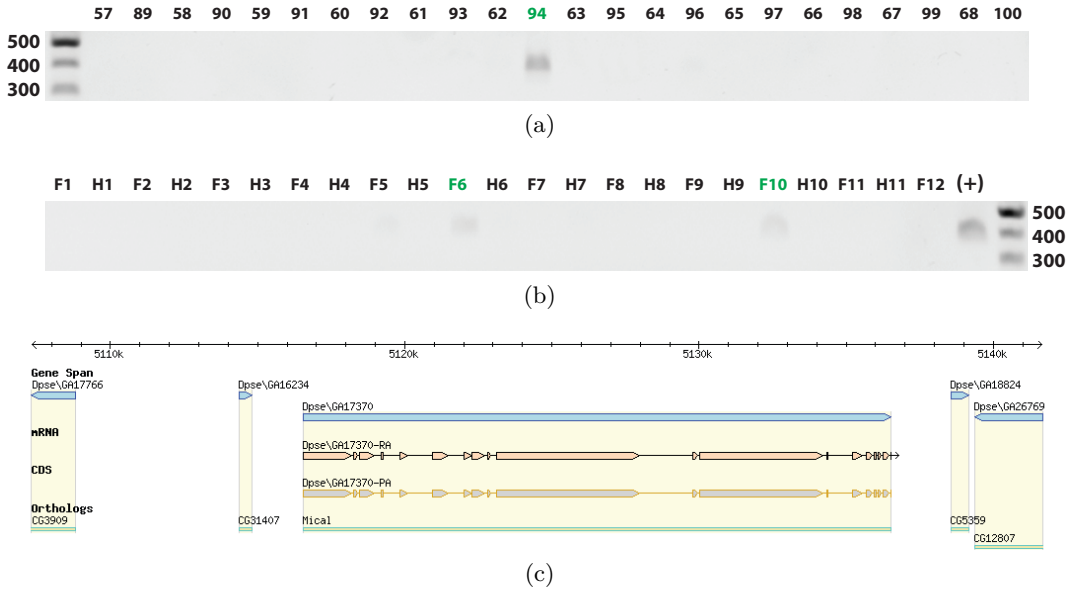


Figure 3.19.: PCR pooling strategy for fosmid clone identification

(a) First round PCR agarose gel showing a single band that identifies library plate (94) containing a *Mical*-positive fosmid clone. (b) Second round PCR agarose gel showing two bands of which one is a false positive (F6) and the other one (F10) identified a fosmid clone (c) containing *Mical* gene.

fosmid clone (figure 3.19a). With a second round of PCR on the single, *Mical*-positive library plate, the candidate clones were identified (figure 3.19b). Identified clones were end-sequenced and mapped to the *D. pseudoobscura* genome to reveal that one of them, indeed, includes the *Mical* locus.

Five genes that lead either to larval lethality or a flightless phenotype were selected for the RNAi rescue experiment (table 3.4, see section D.2 for fosmid maps). All selected fosmids span at least to the next gene 5' and 3' from the gene assayed. The sequence similarity between *D. melanogaster* and *D. pseudoobscura* for the gene regions targeted by the used hairpins ranges from 73–94%. The largest stretch of exact match varies from 17–104 nucleotides. In order to estimate the ability of the siRNAs derived from the hairpins to function in RNAi, the sequences were analyzed using DEQOR (Henschel et al., 2004) – see figure 3.20. DEQOR evaluates all possible 19-mers from the hairpin sequence for a number of criteria (GC content, GC balance across the length of the siRNA and polynucleotide stretches) resulting in a score that reflects the efficiency of each 19-mer in RNAi. DEQOR scores were used to ask whether the long identical stretches between *D. melanogaster* and *D. pseudoobscura* sequences are efficient in RNAi and thus likely to cross-silence the rescue transgene.

Interestingly, most of the long identical stretch sequences were predicted to perform poorly in RNAi suggesting that used hairpins will not significantly affect the *D. pseudoobscura* transgenes.

3.5.3. *In vivo* RNAi fosmid rescue

Transgenic *D. melanogaster* lines were obtained for all five fosmids. In case of the *Mical* fosmid (FlyFos045847), the *dsRed* fluorescence used for transgenic selection was observed in the thorax instead of the eye. As this fosmid was not able to rescue a *Mical* mutant allelic combination that recapitulates observed RNAi phenotype, this transgenic line, as non-functional, was not investigated further. To test cross-species functionality of the D. *pseudoobscura* fosmid in *D. melanogaster* classical mutants of *shg* and *sar1* were rescued to viability and flight ability with FlyFos045685 and FlyFos045459 fosmids, respectively (table 3.4) demonstrating that the *D. pseudoobscura* genes are fully functional in *D. melanogaster*. For *shg* RNAi in muscle, a flightless phenotype caused by missing indirect flight muscles in the thorax (Schnorrer et al., 2010) was observed. The *shg* fosmid does not rescue this phenotype, indicating that the RNAi phenotype is either unspecific or the *D. pseudoobscura* gene is targeted by the hairpin.

Collagen IV (Cg25C) is strongly expressed in embryonic hemocytes and supposedly has an important role in basement membrane function (figure 3.21b). When *Cg25C* is knocked-down in muscle with *Mef2-GAL4*, the *collagen IV* diminishes (figure 3.21c) and larvae die at early stage (Schnorrer et al., 2010). The *D. pseudoobscura Cg25C* fosmid (FlyFos045318) rescues larval growth significantly but not completely compared to knock-down (figure 3.21a) and wild type demonstrating the specificity of the RNAi knock-down. This incomplete rescue suggests that the *Cg25C* expressed from FlyFos045318 fosmid is either not fully functional or not entirely immune to the *Cg25C* hairpin. Antibody staining against *collagen IV* argue for the latter as its localisation around the muscles is still markedly reduced in the rescued larvae (figure 3.21d).

Muscles require the integrin complex for stable attachment to tendons (Bökel et al., 2005). Upon knock-down of *parvin*, a putative member of the *Drosophila* integrin complex, early larval lethality with body muscles displaying a myospheroid phenotype is observed (figure 3.22c and Schnorrer et al., 2010). The myospheroid phenotype is entirely rescued by the *D. pseudoobscura parvin* fosmid (figure 3.22d). Similarly, the growth defect in *parvin* knock-down larva is rescued; interestingly two copies of the fosmid increase the level of rescue (figure 3.22a).

D. mel gene	FlyFos clone	RNAi phenotype	RNAi fosmid rescue result	Mutant allelic combination	Mutant phenotype	Mutant fosmid rescue result
Cg25C (collagen IV)	FlyFos 045318	larval lethal	larval growth rescued; few pupa and adults	$Cg25C^{k00405}$/ Df(2L)Exel7022	embryo or larval lethal	n. a.
CG32528 (parvin)	FlyFos 044975	myospheroid phenotype; early larval lethal	myospheroid phenotype rescued; 2x fosmid survive until early pupae	—	—	—
sar1	FlyFos 045459	sarcomere defect; larval lethal	larval growth and sarcomere phenotype rescued; survive until early pupae	sar^{105712}/ Df(3R)ED6085	embryo or larval lethal	few adult survivors (small size, can fly)
shg	FlyFos 045685	missing flight muscles	no rescue	shg^{E17D}/shg^{2}	embryo or larval lethal	viable adults that fly
Mical	FlyFos 045847	irregular flight muscle myofibrils	no rescue	$Mical^{k1496}$/ Dr(3R)Exel6155	irregular flight muscle myofibrils	no rescue

Table 3.4.: Overview of genes and fosmids used in RNAi rescue

Overview of all genes, RNAi constructs and fosmids used. The degree of homology between the genes in the targeted region is indicated. The RNAi and mutant phenotypes and their rescue by the fosmids is summarized. Table reproduced from Langer et al. (2010).

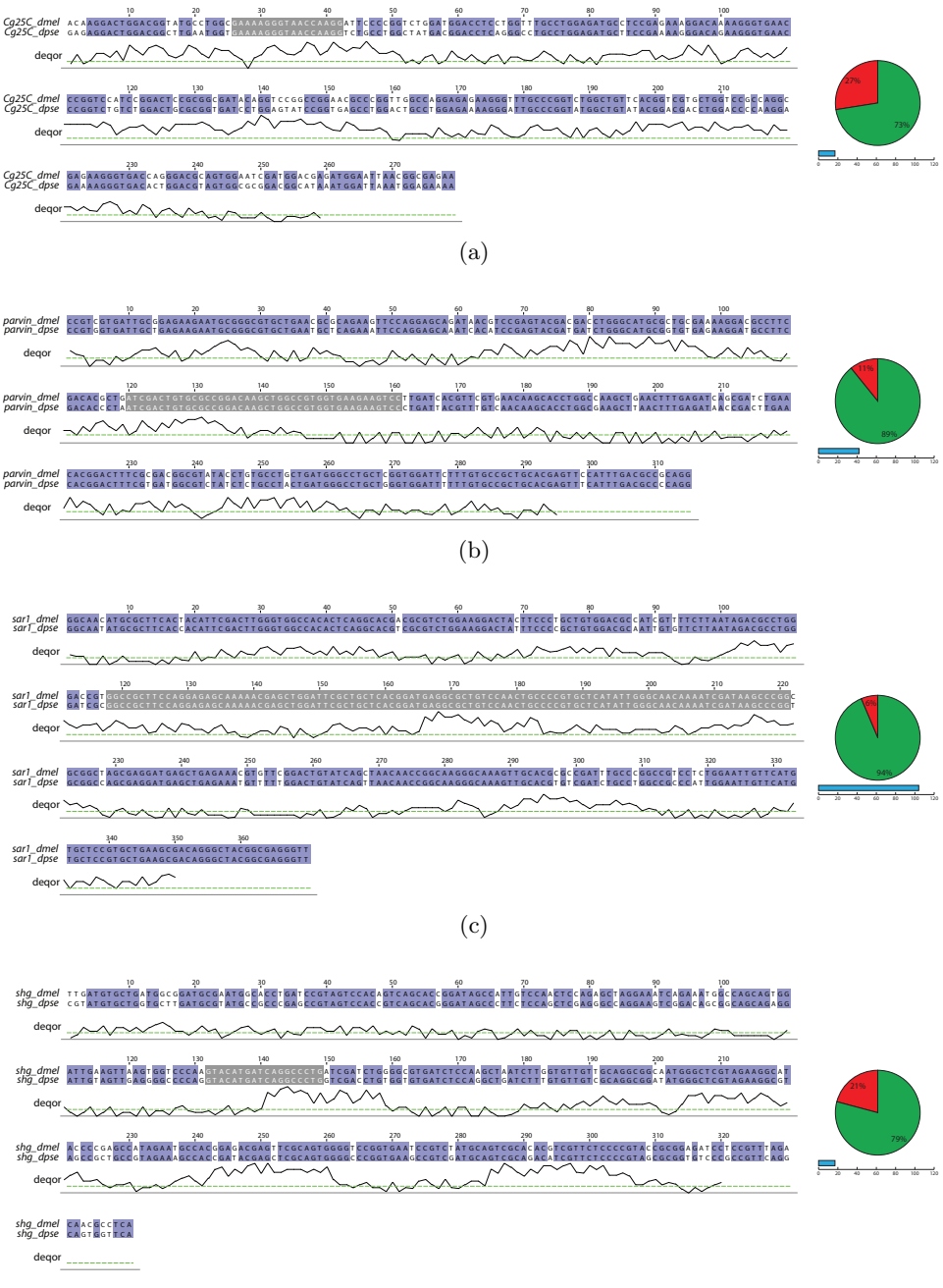


Figure 3.20.: Pairwise sequence alignment of hairpins used in rescue experiments

Alignments between *D. melanogaster* and *D. pseudoobscura* for hairpins targeting (a) *Cg25c (collagen IV)*, (b) *CG32528 (parvin)*, (c) *sar1*, (d) *shg*. The extent of homology and the longest identical nucleotide stretch are graphically depicted next to each alignment. Matching nucleotides are shaded purple, mismatches white and the longest identical stretches are shaded grey within the alignments. The DEQOR scores are plotted below the alignments (a–d) and the score 5 cut-off above which the siRNA at that position is considered RNAi inefficient is depicted by a green line. Figure reproduced from Langer et al. (2010).

Finally, the small GTPase *sar1* implicated in vesicle transport (Aridor et al., 2001) and heart formation in the embryo (Olson, 2006) was investigated. Knock-down of *sar1* in muscle causes a muscle sarcomere phenotype. Both the myosin thick filaments and the Z-line anchoring the actin filaments show a "fading-Z" phenotype or, in extreme cases, a partial loss of sarcomeres (figure 3.23c–d). The FlyFos045459 completely rescues sarcomere phenotype (figure 3.23e) demonstrating a specific role of *sar1* for sarcomere formation and in turn larval growth (figure 3.23a).

3.6. Next-generation recombineering tags

The next generation recombineering tags were designed to provide flexibility and ease of modification. Tags were cloned in a modified pTag-based vector, called pTagNG. The vector features unique restriction sites flanking every feature, and thus, allows for easy modifications. Similarly to the pTag vector family (figure C.2), the FRT-flanked selection cassette contains *rpsL* (streptomycin sensitivity gene) and *kanR* (kanamycin resistance gene). All pTagNG-based tags contain *2xTY1* and *3xFLAG* epitopes on 5' and 3' end respectively. These enable amplification of different tagging cassettes using the same recombineering primers.

The first set of recombineering tags contains *Drosophila* codon-optimized fluorescent proteins (figure C.3). Fluorescent proteins that were chosen for recombineering tags were rated as best choices for respective spectrum ranges (Shaner et al., 2005; Merzlyak et al., 2007; Pédelacq et al., 2006). Codon optimization of a coding sequence is believed to improve expression level of a given gene. Expression rate of codon-optimized genes is not limited by availability of rare tRNAs and therefore can reach maximal levels (Gustafsson et al., 2004). Most commercially available fluorescent protein vectors are optimized for expression in human or *Escherichia coli.* Since codon usage between *Drosophila melanogaster* and these species differs, especially when comparing fruitfly and *E. coli* (figure 3.25), fluorescent protein expression levels and brightness should benefit from codon optimization.

The second set of recombineering tags contains the same fluorescent proteins as the previous set, however this time they are preceded by a *T2A* sequence and followed by a nuclear localization signal (figure C.4). The picoviral *T2A* sequence is a short (18 aminoacids) signal that, due to its sterical properties, causes ribosomal cleavage of a nascent peptide in mechanism of ribosomal skipping (Osborn et al., 2005). This unique feature of *2A*-like sequences allows creation of polycistronic constructs, where a single mRNA encodes for multiple peptides. Such strategy was used in C-terminal recombineering tags to create nuclear fluorescent reporters that are expressed from fosmid in one-to-one molar ratio with the target gene.

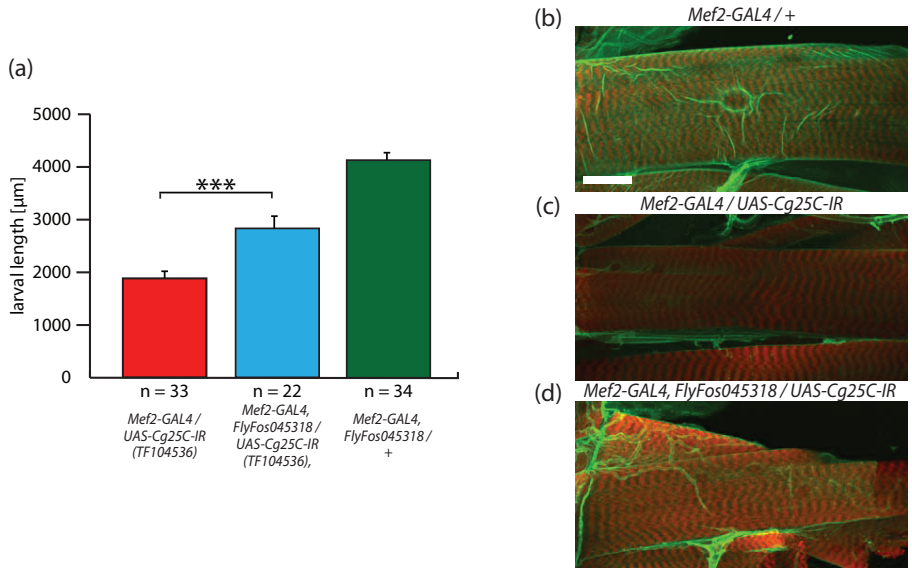


Figure 3.21.: Phenotypic rescue of *Cg25C* by *D. pseudoobscura* fosmid

(a) Quantification of larval size in *Mef2-GAL4 / UAS-Cg25C-IR* larvae (red) rescued by FlyFos045318 (blue) and wild-type (green). ***p,0.0001 (unpaired two-tailed t-test). Larvae 72–96 h after egg laying were assayed. Error bars indicate standard error of the mean (SEM). *Collagen IV* (green) wraps the larval muscles in wild-type (b) and is strongly reduced in *Mef2-GAL4 / UAS-Cg25C-IR* (c) but rescued by FlyFos045318 (d). Actin was visualised with phalloidin. Scale bar corresponds to 25 µm. Figure courtesy of F. Schnorrer, modified from Langer et al. (2010).

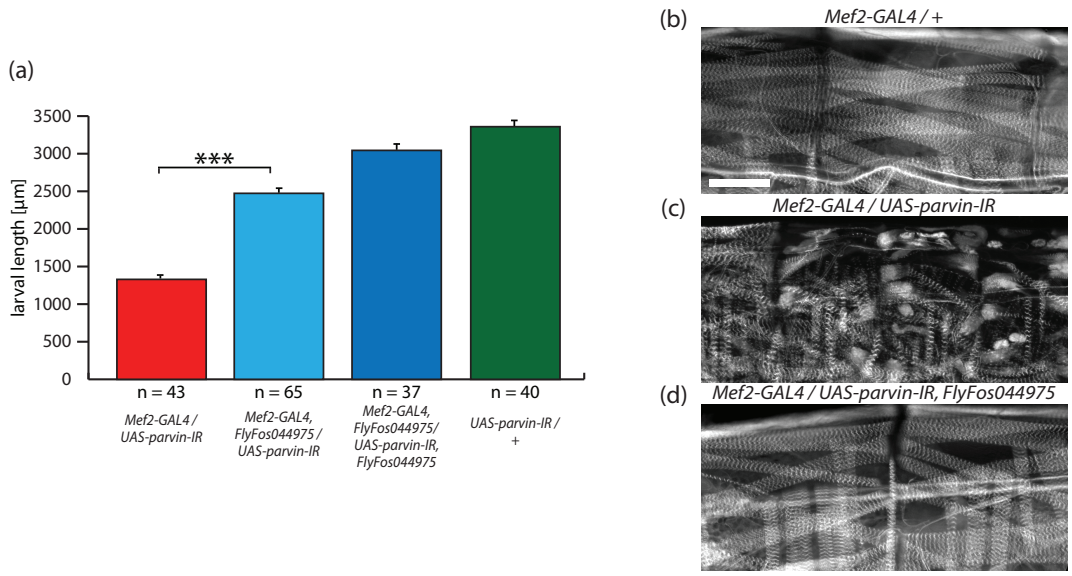


Figure 3.22.: Phenotypic rescue of *parvin* by *D. pseudoobscura* fosmid

(a) Quantification of larval size in *Mef2-GAL4 / UAS-parvin-IR* larva (red), rescued by one (light blue) or two copies of FlyFos044975 (dark blue), compared to wild-type (green). Larvae 48–72 h after egg laying were assayed. Error bars indicate standard error of the mean (SEM), ***p,0.0001 (unpaired two-tailed t-test) compared to rescued larvae. Rounded/myospheroid muscle phenotype in *Mef2-GAL4 / UAS-parvin-IR* (b) is rescued by FlyFos044975 (d) to wild-type (b). Scale bar corresponds to 100 µm. Figure courtesy of F. Schnorrer, modified from Langer et al. (2010).

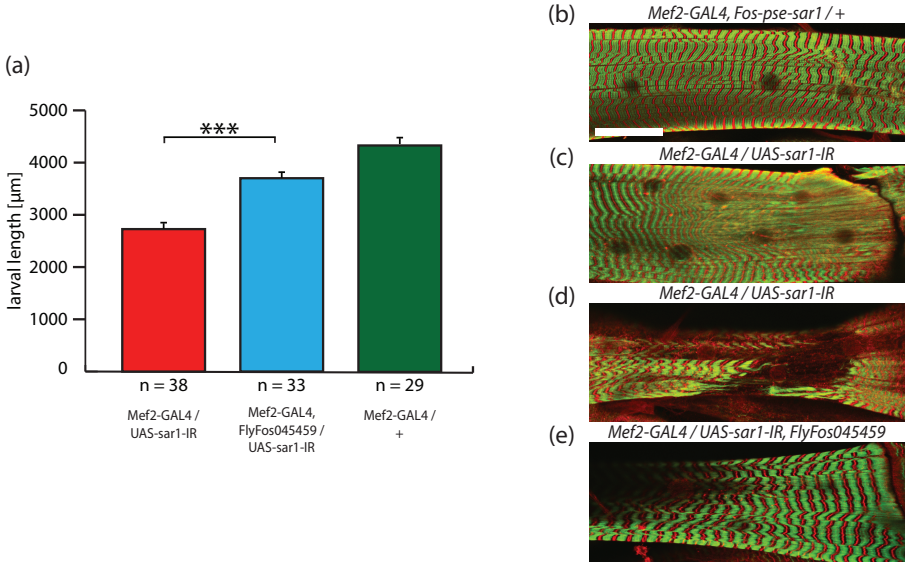


Figure 3.23.: Phenotypic rescue of *sar1* by *D. pseudoobscura* fosmid

(a) Quantification of larval length in *Mef2-GAL4 / UAS-sar1-IR* larvae (red), compared to FlyFos045459 rescued (blue) and wild type (green). Larvae 72–96 h after egg laying were assayed. Error bars indicate standard error of the mean (SEM), ***p,0.0001 (unpaired two-tailed t-test) compared to rescued larvae. Fading Z- and M-line or loss of sarcomeres in *Mef2-GAL4 / UAS-sar1-IR* (c–d) is rescued by FlyFos045459 (e) to wild type (b). Z-lines are visualised with anti-Kettin (red), M-lines with anti-Mhc antibody (green). Scale bar corresponds to 50 µm. Figure courtesy of F. Schnorrer, modified from Langer et al. (2010).

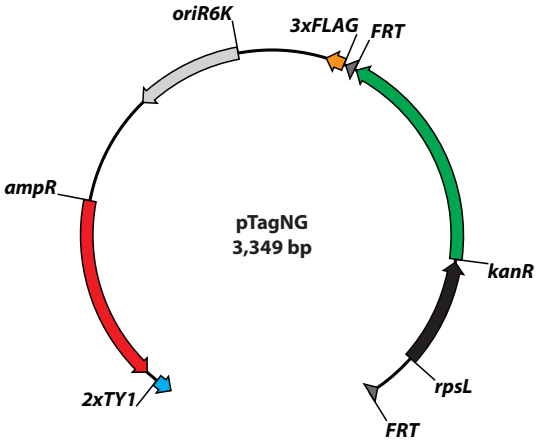


Figure 3.24.: The pTagNG vector backbone

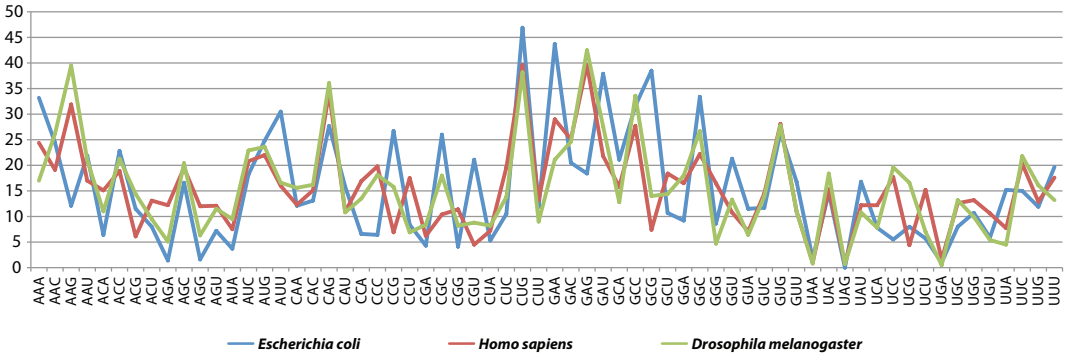


Figure 3.25.: Codon usage in *E. coli*, *H. sapiens* and *D. melanogaster*

The last set of tags utilizes biotin ligase recognition peptide (*BLRP*) and *T2A*-biotin ligase (*birA*) fusion tag. *BLRP*-tagged proteins can undergo *in vivo* biotinylation when co-expressed with *birA* (Tirat et al., 2006). Biotin-tagged proteins can be purified on *avidin* columns or used in immunoprecipitation with anti-biotin antibodies. Biotin ligase tag utilizes previously described *T2A* sequence for bicistronic expression with the tagged gene. One can combine two genes having partially overlapping expression patterns, of which one is tagged with *BLRP* and the other with *T2A-birA*, to specifically biotinylate *BLRP*-tagged protein in a tissue where expression patterns overlap. This creates a sort of binary system for tissue specific protein purification. The *BLRP* tag has been constructed in conjunction with the *V5* epitope and/or *SGFP*, expanding its applications. All *BLRP* tags were cloned in both N-terminal (figure C.5) and C-terminal (figure C.6) variants. The *birA* tag has been designed to produce C-terminal bicistronic fusion gene (figure C.6).

3.7. The "Ultimate" system

The "Ultimate" is a two component *in vivo* recombinase mediated cassette exchange (RMCE) system, consisting of two components: a recombineering tag (pTagNG[Ultimate] - figure C.7) that acts as an RMCE acceptor and pUltimate set of fly-transformable vectors containing RMCE donor cassettes (figure C.8). The Tag[Ultimate] is cloned in the pTagNG backbone and includes the same priming sequences as all other tags developed for *Red/ET* recombineering (using *2xTY1* for forward and *3xFLAG* for reverse primer). Core of the tag is flanked by *FRT3* and *FRT* to enable recombinase mediated cassette exchange. The tag itself consists of *Gal4Δ* under minimal *DSCP* promoter and *Venus* under eye promoter (*GMR-SCP1*). The *FRT*-flanked *rpsL-kanR* cassette is used as a selection-counter-selection marker. The *Gal4Δ* gene can be used as an enhancer trap to drive expression of marker genes under *UAS* promoter. Eye-expressed *Venus* is used in negative fly selection for RMCE.

The pUltimate is a FlyFos based RMCE donor, that contains a fluorescent marker (multiple variants were prepared, including *Venus*, *Cerulean*, *tagRFP* and *EGFP*) flanked by *FRT3* and *FRT* recombination sites. The pUltimate construct harbors a *3xP3-Cerulean* negative selection marker upstream of the donor cassette. Since pFlyFos backbone contains *attP*, the construct is directly fly-transformable. The system will allow to tag genes of interest with pTagUltimate and transform them into flies, resulting in a collection of RMCE acceptor fly lines. The pUltimate will be used to generate a collection of RMCE donor stocks capable of expressing flipase under inducible heat-shock promoter.

For successful RMCE, both acceptor and donor constructs must be integrated into same *locus*. Upon crossing TagUltimate and Ultimate flies and induction of flipase two site-specific, *flp*-mediated recombination events cause an exchange of *FRT3-FRT*-flanked *DSCP-Gal4Δ-GMR-SCP1-Venus* cassette to *FRT3-FRT*-flanked fluorescent marker originating from pUltimate (figure 3.26a). These leads to loss of yellow and blue eye fluorescence in the recombinant progeny. Single recombination events leading to rearrangements in the target sequence can be easily selected for by screening for remnant yellow and/or blue eye fluorescence (figure 3.26b).

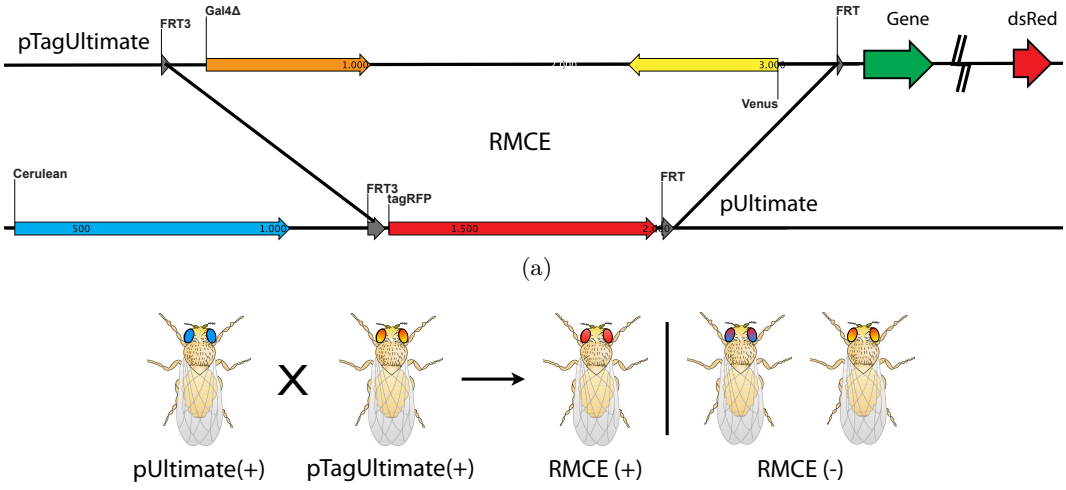


Figure 3.26.: The “Ultimate” system RMCE

(a) The RMCE reaction that occurs when flipase is induced in a cross of TagUltimate and Ultimate flies. (b) Eye-fluorescence phenotypes selected for in the “Ultimate” system. The fluorescent markers on both components of the “Ultimate” system are placed in such way, that only two recombination events required for RMCE result in the correct eye-color phenotype. Upon RMCE, the red and yellow eye-colored acceptor line flies (pTagUltimate(+)) lose the yellow eye fluorescence comming from *Venus* in the RMCE target region. Since the cyan fluorescent marker (*Cerulean*) in the donor strain is outside of the RMCE donor cassete, the recombinant flies do not acquire cyan fluorescence. The red fluorescence comming from pFlyFos vector harborig genomic insert is the only remaining fluorescent marker, and therefore the target strain eye color is red. Upon single recombination event that leads to chromosomal rearrangements, the *Cerulean* marker is introduced and allows for easy rejection of RMCE-negative flies.

Chapter 4.
4 Discussion

Contents

4.1. Overview

We presented a set of tools suitable for live imaging of gene expression patterns in *Drosophila*. We combined state-of-the-art microscopy, accurate nuclear and membrane markers, reliable source of genomic constructs, an efficient way of tagging genes with fluorescent markers and transformation technique allowing to integrate large genomic constructs into the genome. Although accomplishing the final goal, which is creating a comprehensive atlas of gene expression in various Drosophilidae, has not yet begun, we believe that the resulting toolkit can and will be used to achieve it.

4.2. Imaging development *in vivo*

Selective plane illumination microscopy was chosen for imaging because it allows to image *Drosophila* embryos *in toto*. While the two-dimensional resolution of such a microscope is comparable to classical confocal microscopy, the ability to image deep into the embryo from multiple angles (known as multiview imaging) results in three dimensional images of isotropic resolution. The design of the sample chamber allows to keep imaged specimens alive long enough to capture complete developmental time-courses. Fast, CCD-based image acquisition contributes to high temporal resolution of SPIM. The combination of different laser lines, emission filters and fluorescent markers enables recording of multiple channels, and therefore visualization of both cellular behavior and gene expression.

Light-sheet based imaging with digital scanned laser light sheet fluorescence microscopy (DLSM) has already allowed to digitize early development of zebrafish (*Danio rerio*) embryo. Keller et al. (2008) used DLSM to image zebrafish embryo from the 64-cell stage onwards. Similar to our imaging approach, *Histone-GFP* fusion protein was used as a nuclear marker. The whole 24-hour recording was performed in a controlled environment with the tempetarure maintained at 26.5°C, which is optimal for zebrafish development. A two-angle dataset consisting of about 400,000 images was processed with a parallelized image segmentation pipeline to detect positions of nuclei during each timepoint. The authors of that manuscript were able to determine the positions of 92% of the nuclei in the entire embryo. This approach was used to compare the development of wild-type embryo with *one-eyed pinhead* mutants to reveal the mechanism of mesendoderm formation in zebrafish.

We believe, that with light-sheet based microscopy we will be able to provide a similar dataset for *Drosophila melanogaster* embryogenesis. Although attempts were already taken to digitize the development of the fruit fly (Keller et al., 2010), the quality of acquired images is disputable (Tomancak P., manuscript in preparation). The imaged

embryo seems to be dying past the gastrulation. Moreover an imprecise algorithm used for registration may lead to errors in segmentation the of nuclei. We have found that the survival rate of embryos embedded in agarose is limited and therefore multiple imaging sessions are usually required to obtain satisfactory data including whole embryonic development. The multiview reconstructions of *Drosophila* embryos expressing *His2AvD-EYFP* nuclear marker that we have completed so far provide data set suitable for segmentation of individual nuclei early in development. Expanding accurate segentation beyond the early developmental stages will require improvement of both spacial and temporal resolution of the SPIM set-up. Since precise algorithms suitable for reliable segmentation and tracking of nuclei from multiview imaging data are already under development, the fully digital reconstruction of *Drosophila* embryonic development is only a matter of time.

4.3. Genomic DNA libraries

We described here an efficient method to produce genomic fosmid libraries that enable cross-species transgenesis. We have identified steps that are crucial for successful library production. First, we designed a fast and efficient protocol for the isolation of high molecular weight genomic DNA. Second we used mechanical shearing that allows production of the unbiased, sequence-independent DNA fragments for library production. We have found that exposure of genomic DNA to UV light results in irreversible damage, rendering exposed DNA inappropriate for library production. Therefore, for cases where direct use of sheared DNA fragments is impossible, we developed a safe and accurate protocol for gel purification of genomic fragments.

Using the designed protocols, we generated four genomic DNA libraries for *D. melanogaster*, *D. pseudoobscura*, *D. simulans* and *D. virilis*. Two of the constructed libraries, for *D. melanogaster* and *D. pseudoobscura*, were characterized. The libraries were cloned in a unique fosmid vector that features a dominant selectable marker with wide species specificity and φC31-mediated transgenesis. The clones in the libraries exhibit a tight distribution of clone sizes due to the phage packaging step. Comparison of shearing simulations and actual clone mapping revealed that for sequenced and annotated genomes the amount of clones required for a whole genome coverage can be predicted. We developed a simple, yet powerful robotic miniprep protocol that can produce up to 12,000 sequencing grade DNA templates in 2 days. The success rate of sequencing from the isolated template is significantly higher then bacterial culture sequencing allowing for complete mapping of more then 90% of the clones, which is nice. We proposed a hybrid strategy for characterizing the libraries. During the initial sequencing phase virtually every clone is different and the number of cloned

genes increases linearly. When clone coverage approaches single genome complement, we switched to a pooling strategy that allows identification of clones containing the gene of interest by two rounds of PCR in less than one day. This hybrid approach will make it possible to generate libraries not only for all sequenced species of flies, but also for individual strains of a single fly species. Moreover, the method is obviously applicable to any species and particularly among insects it will enable assaying of the activity of divergent genomic regions in the context of *Drosophila melanogaster* genome. The clones from characterized libraries are available to the community. The TransGeneOmics project website (`http://transgeneome.mpi-cbg.de`) contains tools for identifying fosmid clones containing genes of interest.

Fosmid clones containing fragments of fly genome can be integrated into fly genome with satisfactory efficiencies. As we have shown, majority of transgenes introduced to flies recapitulated native expression patterns, however we have observed cases, where the expression pattern of a transgene was different. Ectopic expression of transgenes might be caused by lack of insulator elements in the fosmid constructs. The modENCODE ChIP-Chip- and ChIP-Seq-based map of *D. melanogaster* insulator elements (Nègre et al., 2010) may serve as an aid in choosing fosmid clones that are likely to provide natively expressed transgenes.

Transgenes originating from *D. melanogaster* library were shown to be able to rescue mutant phenotypes. Klose et al. (unpublished data) used FlyFos019790 clone containing *crumbs* locus to rescue *crb* null phenotype in *Drosophila* eyes (figure 4.1). The fosmid clone allowed for full rescue of *crb* mutation lethality, which was not achievable using available cDNA constructs. Viktorinová et al. (2009) used FlyFos021145 clone containing *fat2/kugelei* locus in studies on establishment of planar cell polarity in the *Drosophila* ovary. The *fat2* gene was tagged with *EGFP* using our recombineering pipeline. The tagged transgene was able to rescue *fat2* mutant allelic combination (figure 4.2). The fact that *EGFP*-tagged construct is functional *in vivo* eases functional studies by providing a construct that can readily be used for both rescue experiments and imaging. The FlyFos clones were also used in physical mapping of various DNA-methylation-related mutations by a group of G. Reuter from Martin-Luther-Universität Halle-Wittenberg (personal communication).

The *D. melanogaster* FlyFos library complements recently developed p[ACMAN] libraries (Venken et al., 2009b). The p[ACMAN] libraries were constructed in a BAC vector, functionally similar to pFlyFos, which uses *attB* site for φC31-mediated transgenesis, mini-*white* as a fly-selectable marker and *oriV* for copy-control. The average clone size of the libraries named CHORI-321 and CHORI-322 is 83.3 kb (± 21.5 kb) and 21.0 kb (± 4.0 kb) respectively. The average clone size of FlyFos library (36 kb ± 16 kb) fills a gap between the CHORI libraries (figure 4.3). Majority from

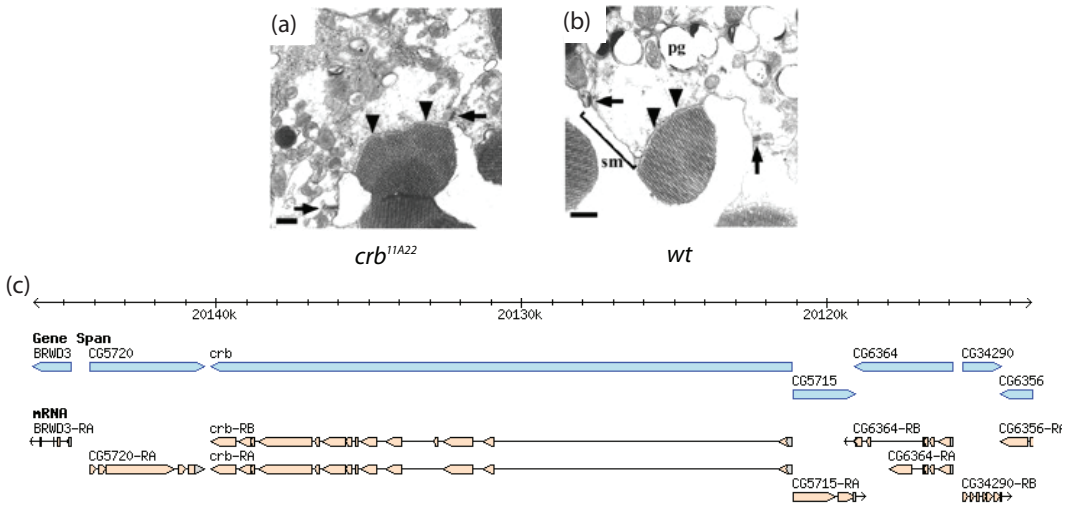


Figure 4.1.: FlyFos019790 rescues *crb* mutants.

The *crb* mutation results in defects in rhabdomere formation (a). FlyFos019790 fosmid clone (c) rescues the mutant phenotype to wild-type (b). Figure courtesy of Klose et al.

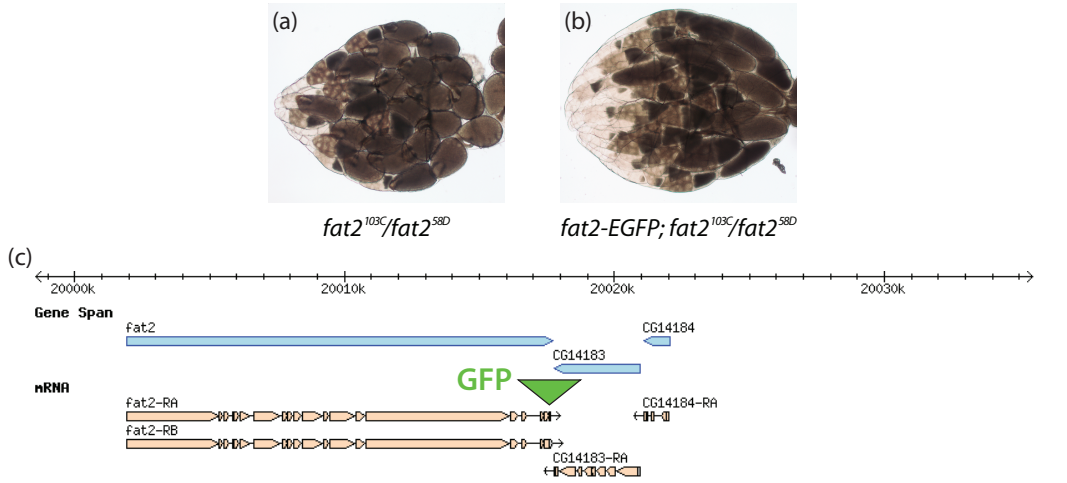


Figure 4.2.: FlyFos021145 rescues *fat2* mutants.

The *fat2* mutant allelic combination results in rounded, disfunctional eggs (a). FlyFos021145 fosmid clone (c) rescues the mutant phenotype to wild-type (b). Figure courtesy of Viktorinová et al.

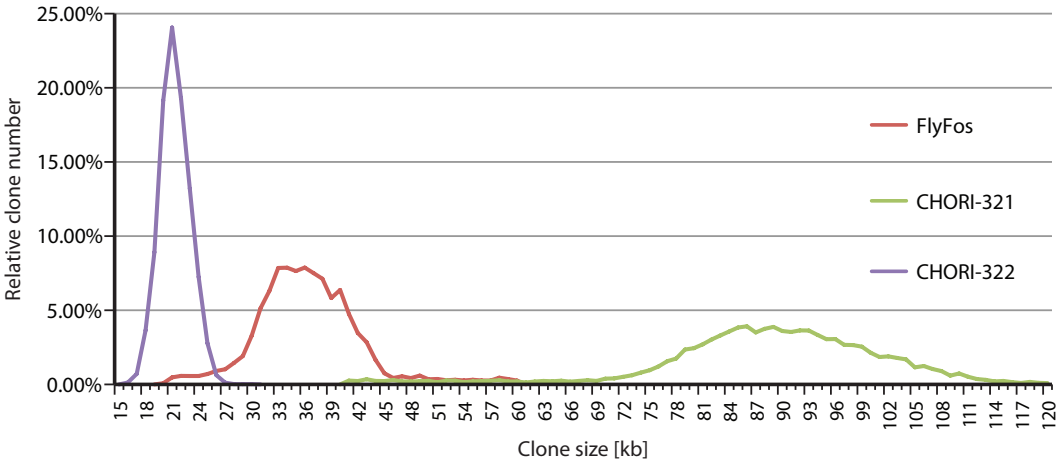


Figure 4.3.: Comparison of clone sizes in p[ACMAN] and FlyFos libraries

10% of genes that due to fosmid packaging size restrictions could not be cloned in our library, are included in the CHORI-321 library. While p[ACMAN] libraries provide larger gene coverage, the selectable marker of pFlyFos vector enables transgenesis in a variety of insect species.

4.4. Recombineering

We showed that FlyFos clones can be highly efficiently modified by liquid culture recombineering. The recombineering pipeline consists of a single homologous recombination step, is routinely performed in 96 well plate format in three days, does not require laborious plating and screening of the recombinant clones at any stage and yet achieved efficiency of successful tagging of up to 100%. The F-factor origin of replication that is included in the fosmid vector keeps fosmid clones single copy in bacteria. If fosmids were maintained at higher copy number (as in cosmid clones) rare recombineering events would be unlikely to occur in all copies, resulting in positively selectable bacteria where only one copy of target sequence is modified. The resulting large background would render liquid culture recombineering inefficient. On the other hand, isolation of low copy fosmid DNA requires large culture volumes and time-consuming handling. By including the inducible high-copy origin of replication (*oriV*) we enable efficient preparation of fosmid DNA in amounts required for fly transformation using a midi-scale protocol.

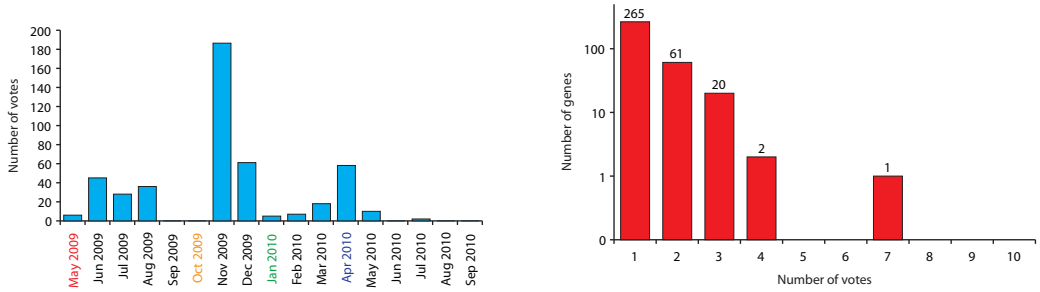
The expression pattern of tagged genes can be verified by *in situ* hybridization. We have shown that expression pattern of genes tagged with *EGFP* can be visualized *in vivo* using fluorescent microscopy and SPIM. The *ubi-mCherry-NLS-T2A* N-terminal tag that we have used in the high-throughput experiment, was visible in

fluorescent imaging only in highly-expressed genes. We deduce that fast turnover of ubiquitinized *mCherry* prevented fluorophone maturation in levels sufficient for imaging, and therefore rendered this tag hard to image.

High efficiencies of liquid culture recombineering that we were able to achieve suggest that the presented method can be upscaled to a genome-wide application. Availability of liquid handling stations and development in automation enables massively parallel sample processing, where dozens of 96-well plates are processed in a single run. Such a high-throughput approach was recently applied in the tagging of over 10,000 *C. elegans* genes with green fluorescent protein (Sarov M., manuscript in preparation). Since we have observed large demand for fluorescently tagged genes in the *Drosophila* community (figure 4.4), we recently began a similar project, where nearly thousand *D. melanogaster* genes are tagged with *GFP*. The tagged constructs can be used for visualization and of gene expression patterns in whole-mount embryo imaging. Because tagged genes are expressed under their native regulation, quantification of expression levels is also possible. Other imaging-related applications may include determining subcellular localization of proteins or membrane dynamics studies using FRAP. But fluorescent imaging is not a limit of tagged protein applications. Various existing antibodies against *GFP* or *TY1* and *FLAG* epitopes allow for purification of tagged protein which can be applied in protein complex studies or chromatin immunoprecipitation (ChIP).

Despite many possible applications of *GFP*-tagged constructs users of the growing FlyFos community have requested different types of tags for various applications. Since the long primers (~75 bp) that include homology arms targeting the tagging cassette to the correct *locus* have the largest contribution to the recombineering costs, we have decided to construct a comon backbone for our tags that includes priming sites used for PCR amplification of the tag. Therefore, primers that allow amplification of a given tagging cassette for one gene, will also be suitable for amplification of the other tags. We have created a collection of tags including broad spectrum of fluorescent proteins that are considered as best choices for a given wavelength range. Since the fluorescent proteins we have chosen include cyan, yellow and red, which are easily separable with commonly used filters, the combination of proteins tagged with these readily allows for three-channel imaging.

In some cases tagging protein on either of its termini disrupts its function. In these cases, use of a large fluorescent marker for internal tagging might cause protein misfolding. Therefore, we have developed a small cassette for internal tagging that includes three commonly used epitopes: *2xTY1*, *V5* and *3xFLAG*. With a total length of 75 amino acids the tag is more likely to work than the 300 residues long *EGFP* tag. Availability of many commercial antibodies against



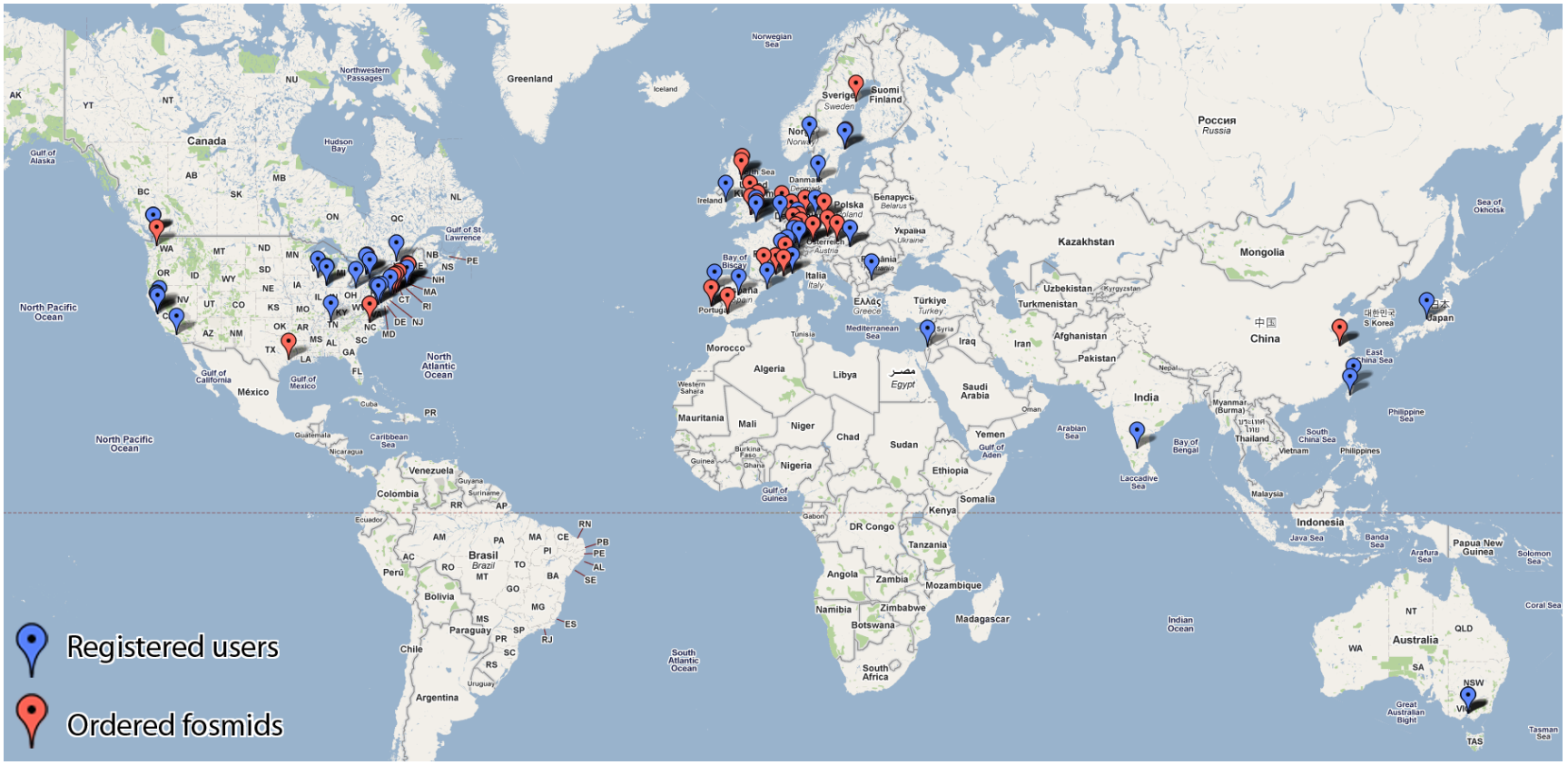
(a) (b)

Figure 4.4.: Community interest in the FlyFos system

The TransGeneOmics website includes an online voting system where users are can submit genes they would like to have tagged. So far nearly 500 genes were voted for by over 60 users. The plot (a) shows number of votes submitted since publication of the Nature Methods paper (Ejsmont et al., 2009). This date is highlighted on the plot in red. The date of Janelia Conference "Improving the Toolkit for Drosophila Neurogenetics" is highlighted in orange, the publication date of RNAi rescue paper (Langer et al., 2010) is highlighted in green, the date of 51st Annual Drosophila Research Conference is highlighted in blue. The plot (b) shows the number of genes that received certain number of votes. Most genes were voted only once. The voting system is available at `http://transgeneome.mpi-cbg.de/transgenomics/user/vote.html`.

2xTY1, *V5* and *3xFLAG* makes this small tag useful in various applications. The proteins tagged with *2xTY1-V5-3xFLAG* can be visualized in fixed samples using fluorescently labeled antibodies, purified by affinity chromatography or used in chromatin immunoprecipitation experiments.

Increasing need for inexpensive protein purification methods encouraged us to develop a recombineering based *in vivo* biotinylation system. Biotinylated proteins and complexes can be purified on streptavidin columns or immunoprecipitated with anti-biotin antibody. We have created a recombineering tag that contains biotin ligase recognition peptide (*BLRP*), which can be biotinylated *in vivo* by biotin ligase (*birA*). To provide a reliable source of biotin ligase, we have constructed a second tag, that harbors *birA* separated from the target protein with *T2A* ribosomal cleavage site. Such an approach allows to either express biotin ligase under control of a strong constitutive promoter of any *Drosophila* housekeeping gene, or express it in a defined subset of cells using tissue-specific drivers. The latter method establishes biotin binary system, where both biotin ligase and *BLRP*-tagged target meet only in overlapping section of two distinct expression patterns. Using this technique, we hope to provide a simple system for tissue-specific protein purification.



(c)

Figure 4.4.: Community interest in the FlyFos system (continued)

The map (c) shows places from which users registered at the TransGeneOmics database come (blue placemarks) and to which fosmid clones were shipped (red placemarks). Map was generated using login data from the TransGeneOmics database and individual fosmid requests.

4.5. The “Ultimate” system

An important, and yet unresolved, question is what tag should be used in genome-wide tagging to cover the broadest range of possible applications. We see several alternative strategies. Firstly, the stunning efficiency of the liquid culture recombineering can be leveraged to generate fosmids tagged with various tags optimized for different purposes. Secondly, composite tags containing sequences for various applications (fluorescent proteins for live imaging, *FLAG, TY1* or *V5* for immunoprecipitation, *HRP* for electron microscopy analysis) could be developed and tested. Since these tags would be relatively large, the efficiency of HT recombineering may be compromised and the resulting fusion proteins may not be functional.

We believe that the best strategy for genome-wide tagging is to tag first with a universal single tag that has immediate versatile use, such as *GAL4*, and subsequently develop recombination strategies to exchange the tag *in vivo*. Recombinase Mediated Cassette Exchange allows efficient replacement of transgenes flanked by recombinase target site (*FLP*, *Cre* or φC31) (Horn and Handler, 2005; Oberstein et al., 2005; Bateman et al., 2006). The existing RMCE systems usually rely on cDNA constructs that can be modified when incorporated into flies. An interesting approach, Minos-mediated integration casette (MIMIC), was presented by Venken et al. (2009a). In this system an artificial exon containing an RMCE acceptor site is integrated into the fly genome using Minos-mediated transposition. If the transposon insertion happens to land in an intron, the artificial exon is incorporated into the gene and allows the creation of protein fusions with reporters introduced by φC31-mediated RMCE. Yet another technique, IMAGO (Choi et al., 2009), enables RMCE in native *loci* by using ends-out Gong and Golic (2003) gene targetting.

With *in vivo* RMCE as a goal we started work on a two-component “Ultimate” RMCE system. Instead of using cDNA constructs or randomly targeting the genome, we will provide a genome-wide resource of RMCE acceptor constructs that include full genomic context of the targeted gene and precisely engineered recombination sites at C- and N-terminus of each fly gene. We believe that *in vivo* RMCE will revolutionize fly transgenesis by eliminating the need for multiple injections required to produce alternatively tagged constructs.

4.6. RNAi specificity assessment

As an example use for non-melanogaster genomic libraries we have established a system for RNAi specificity verification. We have shown that RNAi-induced phenotype that is rescued using a transgene from a related species can be considered specific.

We identified *D. pseudoobscura* and *D. virilis* as suitable species for transgenic RNAi rescue and chose *D. pseudoobscura* FlyFos fosmid library to test the rescue performance. Despite the sequence similarity, which in some cases goes well beyond the well recognized 19 nt threshold (*sar1* 104 nt stretch), we were able to demonstrate rescue of the RNAi phenotype for three of the five genes tested. Similarly we showed rescue of classical mutants for *shg* and *sar1*.

The idea of using orthologous genes in rescuing RNAi phenotypes is not new, and have already been shown to work in *C. elegans (Sarov et al., 2006), D. melanogaster* cell culture and flies (Kondo et al., 2009). Unlike in approach presented by Kondo et al. (2009) where *D. pseudoobscura* fosmids had to be retrofitted in a fly transformable vector, our transgenic libraries allow for direct transgenesis, reducing time and cost of rescue experiments. Another RNAi rescue approach presented recently by Schulz et al. (2009) uses engineered silent mutations in the part of coding sequence targeted by siRNA to generate RNAi-immune alleles. While this strategy may be very useful in cases when rescue by orthologous genes is impossible, we believe that wide application of this technique would be very costly and laborious. Our approach is simple and does not require engineering or any processing of rescue constructs prior to fly transformation. After transgenesis that can be efficiently performed in-house or by a company, the fosmids marked with *dsRed* in eyes and ocelli can be easily recombined with most existing *GAL4* lines or hairpin constructs.

We did not obtain a full rescue of the RNAi phenotypes. Since we observed full rescue of classical mutant phenotypes in two out of three cases and Kondo et al. (2009) reported successful rescue in four out of four cases, we believe that in most cases the *D. pseudoobscura* gene products are able to functionally replace the *D. melanogaster* gene. We hypothesize that the incompleteness of the RNAi rescue is mainly caused by the sequence similarity of the genes between *D. melanogaster* and *D. pseudoobscura* which still results in partial knock-down of the *D. pseudoobscura* gene. In case of *parvin* we have strong evidence supporting this notion as two copies of the fosmid rescue better than a single copy. Kondo et al. (2009) reports full rescue of a *rough-eye* phenotype induced by over-expressing dsRNA directed against apoptotic gene *diap1* with an eye specific driver (*GMR-GAL4*) raising the possibility that the efficiency of the cross-species RNAi rescue will depend on the strength of the *GAL4* driver, the tissue and the gene tested.

Interestingly, the extent of the rescue does not necessarily correlate with the similarity of the hairpin-targeted sequences as measured by longest identity stretches. Hence assessing the efficiency of theoretical siRNAs generated from the hairpin by the DEQOR protocol may represent a more realistic measure of cross-silencing potential.

Analysis of larger sets of cross-species rescue experiments will be required to evaluate the predictive power of the DEQOR analysis.

We observed a broad range of outcomes in our cross-species RNAi rescue experiments that allow us to define simple rules for their interpretation. We propose that if a phenotypic rescue, albeit incomplete, is observed, the specificity of the RNAi knock-down need not be questioned any longer. If, however, no rescue is observed, it is necessary to determine whether the rescuing construct is active. This can be done by rescuing a classical mutant allele if available, or by showing, using antibody staining or RNA *in situ*, that the expression of the hetero-specific transgene mimics the expression of the wild-type ortholog and is unperturbed in the RNAi genetic background. For the purpose of visualizing the rescue construct in a straightforward manner, it may be useful to tag the construct with a reporter such as *GFP*. When these controls establish that the rescue construct is functional, the absence of RNAi rescue indicates that the observed phenotype is caused by an off-target knock-down.

4.7. Outlook

We believe that the non-melanogaster libraries will become essential tools for "evo-devo" studies. We showed that unmodified clones from *D. pseudoobscura* are capable of rescuing RNAi phenotypes when transformed into the *D. melanogaster* genome carrying hairpin transgenes targeting the orthologous locus. With a library for *D. virilis* characterized, we hope to provide a source of transgenes that represents the complementary resource to RNAi libraries in controlling the "off target" effects. The clones from non-melanogaster species can be modified by HT recombineering to easily distinguish them from the endogenous *D. melanogaster* orthologs. This approach will enable quantitative comparisons of gene expression patterns of non-melanogaster transgenes in the cellular enviroment of *D. melanogaster* assessing the contribution of *cis*-regulatory sequences and *trans*-acting factors to the pattern divergence. The reciprocal experiment of assessing *D. melanogaster* transgenes in non-melanogaster specifes will become feasible when landing sites for other Drosophilid genomes become available. The dominant selectable marker with broad species specificity employed in the FlyFos vector facilitates the routine production of hybrid Drosophilid genomes.

The recombineering pipeline that we described is easily realizable in any laboratory at least in its low-throughput version. We plan to generate a genome-wide resource of tagged fosmid clones. By exploring the capabilities of recently developed tags, we will expand our set of tools for protein localization and function studies. We plan to extensively test red-fluorescent tags in multichannel imaging with *HisEYFP* and *CadECFP* and if successful, use this strategy to systematically generate a

genome-wide resource for expression pattern imaging. With a growing number of light-sheet illumination microscope set-ups we hope to establish a comunity-driven effort in generating an atlas of gene expression in *Drosophila*. With libraries for non-melanogaster species characterized we could expand this effort for other Drosophilidae, thus generating valuable resource for comparative genomics. With the Ultimate system ready, we will be able to generate a genome-wide resource of transgenes ready to pop-in any tag by simply crossing flies, and therefore circumventing costly and time-consuming injections.

Since DNA modifications introduced with recombineering are not limited to tagging, we plan to explore high-throughput approaches to generate deletions and substitutions in both coding and non-coding DNA. To achieve high-efficiency, scarless deletions, we will develop efficient liquid culture selection-counter-selection strategy. While selection-counter-selection approaches were already used in recombineering (Zhang et al., 1998), the counter-selection gene used (*rpsL*) was prone to giving a large background. In our approach, we plan to combine a lethal *ccdB* counter-selection cassette included in the tag with the *ccdB* resistanc gene (*ccdA*) under an inducible promoter. Quick turnover of *ccdA* when its expression is ceased will kill bacteria where counter selectable marker was not removed (either by flipout, or by second recombineering event). We believe that this technique will work efficiently in liquid culture, expanding applications of high-throughput recombineering in studying gene function and regulation.

The major drawback of presented fosmid transgenesis system, is that the transgenes are introduced as additional, third-copy alleles. While for most cases such approach should work, in extreme examples of genes that are highly dosage-specific, introduction of additional alleles may result in an overexpression phenotype. We think that it would be worth to combine the fosmid recombineering approach with *in vivo* homologous recombination in *Drosophila*, thus providing a framework for modification of genes in their genomic *loci*. Two approaches that enable homologous gene targetting in flies, ends-out Gong and Golic (2003) and ends-in Xie and Golic (2004) seem to be integratable with fosmid technology. We believe that combining gene targetting with fosmid recombineering and recombinase-mediated cassette exchange will provide a completely new quality in generating complex reporters in *Drosophila*.

Bibliography

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, et al. (2000) The genome sequence of Drosophila melanogaster. Science 287(5461): 2185–95.

Ajioka JW, Smoller DA, Jones RW, Carulli JP, Vellek AE, Garza D, Link AJ, Duncan IW and Hartl DL (1991) Drosophila genome project: one-hit coverage in yeast artificial chromosomes. Chromosoma 100(8): 495–509.

Aleksic J, Lazic R, Müller I, Russell SR and Adryan B (2009) Biases in Drosophila melanogaster protein trap screens. BMC Genomics 10: 249.

Arbeitman MN, Furlong EEM, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW and White KP (2002) Gene expression during the life cycle of Drosophila melanogaster. Science 297(5590): 2270–5.

Aridor M, Fish KN, Bannykh S, Weissman J, Roberts TH, Lippincott-Schwartz J and Balch WE (2001) The Sar1 GTPase coordinates biosynthetic cargo selection with endoplasmic reticulum export site assembly. J Cell Biol 152(1): 213–29.

Arunachalam M, Jayasurya K, Tomancak P and Ohler U (2010) An alignment-free method to identify candidate orthologous enhancers in multiple Drosophila genomes. Bioinformatics .

Ashburner M, Golic KG and Hawley RS (2005) Drosophila: a laboratory handbook. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2nd edn.

Barolo S, Carver LA and Posakony JW (2000) GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in Drosophila. Biotechniques 29(4): 726, 728, 730, 732.

Bateman JR, Lee AM and Wu Ct (2006) Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. Genetics 173(2): 769–77.

Bellen HJ, Levis RW, Liao G, He Y, Carlson JW, Tsang G, Evans-Holm M, Hiesinger PR, Schulze KL, Rubin GM, et al. (2004) The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes. Genetics 167(2): 761–81.

Berghammer AJ, Klingler M and Wimmer EA (1999) A universal marker for transgenic insects. Nature 402(6760): 370–371.

Berman BP, Pfeiffer BD, Laverty TR, Salzberg SL, Rubin GM, Eisen MB and Celniker SE (2004) Computational identification of developmental enhancers: conservation and function of transcription factor binding-site clusters in Drosophila melanogaster and Drosophila pseudoobscura. Genome Biol 5(9): R61.

Bischof J, Maeda RK, Hediger M, Karch F and Basler K (2007) An optimized transgenesis system for Drosophila using germ-line-specific phiC31 integrases. Proc Natl Acad Sci U S A 104(9): 3312–3317.

Bökel C, Prokop A and Brown NH (2005) Papillote and Piopio: Drosophila ZP-domain proteins required for cell adhesion to the apical extracellular matrix and microtubule organization. J Cell Sci 118(Pt 3): 633–42.

Brand AH and Perrimon N (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118(2): 401–15.

Buchholz F, Angrand PO and Stewart AF (1998) Improved properties of FLP recombinase evolved by cycling mutagenesis. Nat Biotechnol 16(7): 657–62.

Burke DT, Carle GF and Olson MV (1987) Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236(4803): 806–12.

Carlsson K, Danielsson PE, Lenz R, Liljeborg A, Majlöf L and Aslund N (1985) Three-dimensional microscopy using a confocal laser scanning microscope. Opt Lett 10(2): 53–5.

Choi CM, Vilain S, Langen M, Van Kelst S, De Geest N, Yan J, Verstreken P and Hassan BA (2009) Conditional mutagenesis in Drosophila. Science 324(5923): 54.

Clarkson M and Saint R (1999) A His2AvDGFP fusion gene complements a lethal His2AvD mutant allele and provides an in vivo marker for Drosophila chromosome behavior. DNA Cell Biol 18(6): 457–462.

Collins J and Hohn B (1978) Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. Proc Natl Acad Sci U S A 75(9): 4242–6.

Cooley L, Kelley R and Spradling A (1988) Insertional mutagenesis of the Drosophila genome with single P elements. Science 239(4844): 1121–8.

Davis I, Girdham CH and O'Farrell PH (1995) A nuclear GFP that marks nuclei in living Drosophila embryos; maternal supply overcomes a delay in the appearance of zygotic fluorescence. Dev Biol 170(2): 726–729.

Devroye L (1986) Non-Uniform Random Variate Generation. Springer-Verlag, New York.

Dietzl G, Chen D, Schnorrer F, Su KC, 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(7150): 151–6.

Drosophila 12 Genomes Consortium, Clark AG, Eisen MB, Smith DR, Bergman CM, Oliver B, Markow TA, Kaufman TC, Kellis M, Gelbart W, et al. (2007) Evolution of genes and genomes on the Drosophila phylogeny. Nature 450(7167): 203–18.

Ejsmont RK, Sarov M, Winkler S, Lipinski KA and Tomancak P (2009) A toolkit for high-throughput, cross-species gene engineering in Drosophila. Nat Methods 6(6): 435–7.

EPICENTRE (2010) CopyControl Fosmid Library Production Kit. EPICENTRE Biotechnologies.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE and Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391(6669): 806–11.

Fowlkes CC, Hendriks CLL, Keränen SVE, Weber GH, Rübel O, Huang MY, Chatoor S, DePace AH, Simirenko L, Henriquez C, et al. (2008) A quantitative spatiotemporal atlas of gene expression in the Drosophila blastoderm. Cell 133(2): 364–74.

Frise E, Hammonds AS and Celniker SE (2010) Systematic image-driven analysis of the spatial Drosophila embryonic expression landscape. Mol Syst Biol 6: 345.

Gaul U, Seifert E, Schuh R and Jäckle H (1987) Analysis of Krüppel protein distribution during early Drosophila development reveals posttranscriptional regulation. Cell 50(4): 639–47.

GeneMachines (2010) HydroShear DNA Shearing Device User Manual. GeneMachines.

Gong WJ and Golic KG (2003) Ends-out, or replacement, gene targeting in Drosophila. Proc Natl Acad Sci U S A 100(5): 2556–61.

Green ED, Riethman HC, Dutchik JE and Olson MV (1991) Detection and characterization of chimeric yeast artificial-chromosome clones. Genomics 11(3): 658–69.

Groth AC, Olivares EC, Thyagarajan B and Calos MP (2000) A phage integrase directs efficient site-specific integration in human cells. Proc Natl Acad Sci U S A 97(11): 5995–6000.

Groth AC, Fish M, Nusse R and Calos MP (2004) Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics 166(4): 1775–82.

Gustafsson C, Govindarajan S and Minshull J (2004) Codon bias and heterologous protein expression. Trends Biotechnol 22(7): 346–53.

Guzman LM, Belin D, Carson MJ and Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177(14): 4121–30.

Handler AM and Harrell RAn (1999) Germline transformation of Drosophila melanogaster with the piggyBac transposon vector. Insect Mol Biol 8(4): 449–457.

Heim R and Tsien RY (1996) Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. Curr Biol 6(2): 178–82.

Helmchen F and Denk W (2005) Deep tissue two-photon microscopy. Nat Methods 2(12): 932–40.

Henschel A, Buchholz F and Habermann B (2004) DEQOR: a web-based tool for the design and quality control of siRNAs. Nucleic Acids Res 32(Web Server issue): W113–20.

Horn C and Handler AM (2005) Site-specific genomic targeting in Drosophila. Proc Natl Acad Sci U S A 102(35): 12483–12488.

Horn C, Jaunich B and Wimmer EA (2000) Highly sensitive, fluorescent transformation marker for Drosophila transgenesis. Dev Genes Evol 210(12): 623–629.

Horn C, Schmid BGM, Pogoda FS and Wimmer EA (2002) Fluorescent transformation markers for insect transgenesis. Insect Biochem Mol Biol 32(10): 1221–1235.

Huisken J, Swoger J, Del Bene F, Wittbrodt J and Stelzer EHK (2004) Optical sectioning deep inside live embryos by selective plane illumination microscopy. Science 305(5686): 1007–1009.

Ji S, Sun L, Jin R, Kumar S and Ye J (2008) Automated annotation of Drosophila gene expression patterns using a controlled vocabulary. Bioinformatics 24(17): 1881–8.

Kalinka AT, Varga KM, Gerrard DT, Preibisch S, Jarrells J, Corcoran D, Ohler U, Bergman CM and Tomancak P (2010) Gene expression divergence recapitulates the developmental hourglass. submitted manuscript .

Keller PJ, Schmidt AD, Wittbrodt J and Stelzer EHK (2008) Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. Science 322(5904): 1065–9.

Keller PJ, Schmidt AD, Santella A, Khairy K, Bao Z, Wittbrodt J and Stelzer EHK (2010) Fast, high-contrast imaging of animal development with scanned light sheet-based structured-illumination microscopy. Nat Methods 7(8): 637–42.

Keranen SVE, Fowlkes CC, Luengo Hendriks CL, Sudar D, Knowles DW, Malik J and Biggin MD (2006) Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution II: dynamics. Genome Biol 7(12): R124.

Kheradpour P, Stark A, Roy S and Kellis M (2007) Reliable prediction of regulator targets using 12 Drosophila genomes. Genome Res 17(12): 1919–31.

Kim UJ, Shizuya H, de Jong PJ, Birren B and Simon MI (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Res 20(5): 1083–5.

Kimmerly W, Stultz K, Lewis S, Lewis K, Lustre V, Romero R, Benke J, Sun D, Shirley G, Martin C, et al. (1996) A P1-based physical map of the Drosophila euchromatic genome. Genome Res 6(5): 414–30.

Kittler R, Pelletier L, Ma C, Poser I, Fischer S, Hyman AA and Buchholz F (2005) RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. Proc Natl Acad Sci U S A 102(7): 2396–401.

Kokoza V, Ahmed A, Wimmer EA and Raikhel AS (2001) Efficient transformation of the yellow fever mosquito Aedes aegypti using the piggyBac transposable element vector pBac[3xP3-EGFP afm]. Insect Biochem Mol Biol 31(12): 1137–43.

Kondo S, Booker M and Perrimon N (2009) Cross-species RNAi rescue platform in Drosophila melanogaster. Genetics 183(3): 1165–73.

Kulkarni MM, Booker M, Silver SJ, Friedman A, Hong P, Perrimon N and Mathey-Prevot B (2006) Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. Nat Methods 3(10): 833–8.

Langer CCH, Ejsmont RK, Schönbauer C, Schnorrer F and Tomancak P (2010) In vivo RNAi rescue in Drosophila melanogaster with genomic transgenes from Drosophila pseudoobscura. PLoS One 5(1): e8928.

Liao GC, Rehm EJ and Rubin GM (2000) Insertion site preferences of the P transposable element in Drosophila melanogaster. Proc Natl Acad Sci U S A 97(7): 3347–51.

Lobo N, Li X and Fraser MJ Jr (1999) Transposition of the piggyBac element in embryos of Drosophila melanogaster, Aedes aegypti and Trichoplusia ni. Mol Gen Genet 261(4-5): 803–10.

Loukeris TG, Arcà B, Livadaras I, Dialektaki G and Savakis C (1995a) Introduction of the transposable element Minos into the germ line of Drosophila melanogaster. Proc Natl Acad Sci U S A 92(21): 9485–9.

Loukeris TG, Livadaras I, Arcà B, Zabalou S and Savakis C (1995b) Gene transfer into the medfly, Ceratitis capitata, with a Drosophila hydei transposable element. Science 270(5244): 2002–5.

Luengo Hendriks CL, Keranen SVE, Fowlkes CC, Simirenko L, Weber GH, DePace AH, Henriquez C, Kaszuba DW, Hamann B, Eisen MB, et al. (2006) Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution I: data acquisition pipeline. Genome Biol 7(12): R123.

Lyne R, Smith R, Rutherford K, Wakeling M, Varley A, Guillier F, Janssens H, Ji W, Mclaren P, North P, et al. (2007) FlyMine: an integrated database for Drosophila and Anopheles genomics. Genome Biol 8(7): R129.

Ma Y, Creanga A, Lum L and Beachy PA (2006) Prevalence of off-target effects in Drosophila RNA interference screens. Nature 443(7109): 359–63.

Mace DL, Varnado N, Zhang W, Frise E and Ohler U (2010) Extraction and comparison of gene expression patterns from 2D RNA in situ hybridization images. Bioinformatics 26(6): 761–9.

Mandrioli M and Wimmer EA (2003) Stable transformation of a Mamestra brassicae (lepidoptera) cell line with the lepidopteran-derived transposon piggyBac. Insect Biochem Mol Biol 33(1): 1–5.

Manseau L, Baradaran A, Brower D, Budhu A, Elefant F, Phan H, Philp AV, Yang M, Glover D, Kaiser K, et al. (1997) GAL4 enhancer traps expressed in the embryo, larval brain, imaginal discs, and ovary of Drosophila. Dev Dyn 209(3): 310–22.

Marcus JM, Ramos DM and Monteiro A (2004) Germline transformation of the butterfly Bicyclus anynana. Proc Biol Sci 271 Suppl 5: S263–5.

Maresca M, Erler A, Fu J, Friedrich A, Zhang Y and Stewart AF (2010) Single-stranded heteroduplex intermediates in lambda Red homologous recombination. BMC Mol Biol 11: 54.

Markstein M, Pitsouli C, Villalta C, Celniker SE and Perrimon N (2008) Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat Genet 40(4): 476–83.

Matsumoto A, Ukai-Tadenuma M, Yamada RG, Houl J, Uno KD, Kasukawa T, Dauwalder B, Itoh TQ, Takahashi K, Ueda R, et al. (2007) A functional genomics strategy reveals clockwork orange as a transcriptional regulator in the Drosophila circadian clock. Genes Dev 21(13): 1687–700.

Merzlyak EM, Goedhart J, Shcherbo D, Bulina ME, Shcheglov AS, Fradkov AF, Gaintzeva A, Lukyanov KA, Lukyanov S, Gadella TWJ, et al. (2007) Bright

monomeric red fluorescent protein with an extended fluorescence lifetime. Nat Methods 4(7): 555–7.

Milne AA (1926) Winnie-the-Pooh. Methuen & Co. Ltd., London, 1st edn.

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. Proc Natl Acad Sci U S A 98(26): 15050–5.

Mummery-Widmer JL, Yamazaki M, Stoeger T, Novatchkova M, Bhalerao S, Chen D, Dietzl G, Dickson BJ and Knoblich JA (2009) Genome-wide analysis of Notch signalling in Drosophila by transgenic RNAi. Nature 458(7241): 987–92.

Murphy KC (1998) Use of bacteriophage lambda recombination functions to promote gene replacement in Escherichia coli. J Bacteriol 180(8): 2063–71.

Muyrers JP, Zhang Y, Testa G and Stewart AF (1999) Rapid modification of bacterial artificial chromosomes by ET-recombination. Nucleic Acids Res 27(6): 1555–7.

Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K and Miyawaki A (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat Biotechnol 20(1): 87–90.

Nakamura Y, Gojobori T and Ikemura T (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic Acids Res 28(1): 292.

Nègre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, Feng X, Ahmad K, Russell S, White RAH, et al. (2010) A comprehensive map of insulator elements for the Drosophila genome. PLoS Genet 6(1): e1000814.

Ni JQ, Liu LP, Binari R, Hardy R, Shim HS, Cavallaro A, Booker M, Pfeiffer BD, Markstein M, Wang H, et al. (2009) A Drosophila resource of transgenic RNAi lines for neurogenetics. Genetics 182(4): 1089–100.

Nüsslein-Volhard C and Wieschaus E (1980) Mutations affecting segment number and polarity in Drosophila. Nature 287(5785): 795–801.

Oberstein A, Pare A, Kaplan L and Small S (2005) Site-specific transgenesis by Cre-mediated recombination in Drosophila. Nat Methods 2(8): 583–5.

Oda H and Tsukita S (2001) Real-time imaging of cell-cell adherens junctions reveals that Drosophila mesoderm invagination begins with two phases of apical constriction of cells. J Cell Sci 114(Pt 3): 493–501.

Olivares EC, Hollis RP, Chalberg TW, Meuse L, Kay MA and Calos MP (2002) Site-specific genomic integration produces therapeutic Factor IX levels in mice. Nat Biotechnol 20(11): 1124–8.

Olson EN (2006) Gene regulatory networks in the evolution and development of the heart. Science 313(5795): 1922–7.

Osborn MJ, Panoskaltsis-Mortari A, McElmurry RT, Bell SK, Vignali DAA, Ryan MD, Wilber AC, McIvor RS, Tolar J and Blazar BR (2005) A picornaviral 2A-like sequence-based tricistronic vector allowing for high-level therapeutic gene expression coupled to a dual-reporter system. Mol Ther 12(3): 569–574.

Osoegawa K, Vessere GM, Li Shu C, Hoskins RA, Abad JP, de Pablos B, Villasante A and de Jong PJ (2007) BAC clones generated from sheared DNA. Genomics 89(2): 291–299.

Pavlopoulos A, Berghammer AJ, Averof M and Klingler M (2004) Efficient transformation of the beetle Tribolium castaneum using the Minos transposable element: quantitative and qualitative analysis of genomic integration events. Genetics 167(2): 737–46.

Pédelacq JD, Cabantous S, Tran T, Terwilliger TC and Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. Nat Biotechnol 24(1): 79–88.

Peng H, Long F, Zhou J, Leung G, Eisen MB and Myers EW (2007) Automatic image analysis for gene expression patterns of fly embryos. BMC Cell Biol 8 Suppl 1: S7.

Perrimon N and Mathey-Prevot B (2007) Matter arising: off-targets and genome-scale RNAi screens in Drosophila. Fly (Austin) 1(1): 1–5.

Poser I, Sarov M, Hutchins JRA, Hériché JK, Toyoda Y, Pozniakovsky A, Weigl D, Nitzsche A, Hegemann B, Bird AW, et al. (2008) BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals. Nat Methods 5(5): 409–15.

Preibisch S, Rohlfing T, Hasak M and Tomancak P (2007) Mosaicing of Single Plane Illumination Miscroscopy Images Using Groupwise Registration and Fast Content-Based Image Fusion .

Preibisch S, Saalfeld S, Rohlfing T and Tomancak P (2008) Bead-based Mosaicing of Single Plane Illumination Microscopy Images using Geometric Local Descriptor Matching. In SPIE.

Preibisch S, Saalfeld S, Schindelin J and Tomancak P (2010) Software for bead-based registration of selective plane illumination microscopy data. Nat Methods 7(6): 418–9.

Prud'homme B, Gompel N, Rokas A, Kassner VA, Williams TM, Yeh SD, True JR and Carroll SB (2006) Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. Nature 440(7087): 1050–3.

Puigbò P, Guzmán E, Romeu A and Garcia-Vallvé S (2007) OPTIMIZER: a web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res 35(Web Server issue): W126–31.

QIAGEN (2010) QIAGEN Plasmid Purification Handbook. QIAGEN.

Quiñones-Coello AT, Petrella LN, Ayers K, Melillo A, Mazzalupo S, Hudson AM, Wang S, Castiblanco C, Buszczak M, Hoskins RA, et al. (2007) Exploring strategies for protein trapping in Drosophila. Genetics 175(3): 1089–104.

Rasband WF (1997-2010) ImageJ.
URL `http://rsb.info.nih.gov/ij/`

Rizzo MA, Springer GH, Granada B and Piston DW (2004) An improved cyan fluorescent protein variant useful for FRET. Nat Biotechnol 22(4): 445–449.

Roignant JY, Carré C, Mugat B, Szymczak D, Lepesant JA and Antoniewski C (2003) Absence of transitive and systemic pathways allows cell-specific and isoform-specific RNAi in Drosophila. RNA 9(3): 299–308.

Rubin GM and Spradling AC (1982) Genetic transformation of Drosophila with transposable element vectors. Science 218(4570): 348–53.

Ryder E, Ashburner M, Bautista-Llacer R, Drummond J, Webster J, Johnson G, Morley T, Chan YS, Blows F, Coulson D, et al. (2007) The DrosDel deletion collection: a Drosophila genomewide chromosomal deficiency resource. Genetics 177(1): 615–29.

Sarov M and Stewart AF (2005) The best control for the specificity of RNAi. Trends Biotechnol 23(9): 446–8.

Sarov M, Schneider S, Pozniakovski A, Roguev A, Ernst S, Zhang Y, Hyman AA and Stewart AF (2006) A recombineering pipeline for functional genomics applied to Caenorhabditis elegans. Nat Methods 3(10): 839–44.

Schindelin J, Saalfeld S, Cardona A, Tinevez JY, Longair M, Schmid B, Arganda-Carreras I, Preibisch S, Elliott C, Weiler N, et al. (2008-2010) FIJI is just ImageJ.
URL `http://pacific.mpi-cbg.de/`

Schnorrer F, Schönbauer C, Langer CCH, Dietzl G, Novatchkova M, Schernhuber K, Fellner M, Azaryan A, Radolf M, Stark A, et al. (2010) Systematic genetic analysis of muscle morphogenesis and function in Drosophila. Nature 464(7286): 287–91.

Schulz JG, David G and Hassan BA (2009) A novel method for tissue-specific RNAi rescue in Drosophila. Nucleic Acids Res 37(13): e93.

Shaner NC, Campbell RE, Steinbach PA, Giepmans BNG, Palmer AE and Tsien RY (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol 22(12): 1567–1572.

Shaner NC, Steinbach PA and Tsien RY (2005) A guide to choosing fluorescent proteins. Nat Methods 2(12): 905–909.

Shizuya H, Birren B, Kim UJ, Mancino V, Slepak T, Tachiiri Y and Simon M (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc Natl Acad Sci U S A 89(18): 8794–7.

Sidén-Kiamos I, Saunders RD, Spanos L, Majerus T, Treanear J, Savakis C, Louis C, Glover DM, Ashburner M and Kafatos FC (1990) Towards a physical map of the Drosophila melanogaster genome: mapping of cosmid clones within defined genomic divisions. Nucleic Acids Res 18(21): 6261–70.

Smith CD, Shu S, Mungall CJ and Karpen GH (2007) The Release 5.1 annotation of Drosophila melanogaster heterochromatin. Science 316(5831): 1586–1591.

Spradling AC, Stern DM, Kiss I, Roote J, Laverty T and Rubin GM (1995) Gene disruptions using P transposable elements: an integral component of the Drosophila genome project. Proc Natl Acad Sci U S A 92(24): 10824–30.

Stapleton M, Carlson J, Brokstein P, Yu C, Champe M, George R, Guarin H, Kronmiller B, Pacleb J, Park S, et al. (2002) A Drosophila full-length cDNA resource. Genome Biol 3(12): RESEARCH0080.

Steinberg MS and McNutt PM (1999) Cadherins and their connections: adhesion junctions have broader functions. Curr Opin Cell Biol 11(5): 554–60.

Sternberg N (1990) Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. Proc Natl Acad Sci U S A 87(1): 103–7.

Stolc V, Gauhar Z, Mason C, Halasz G, van Batenburg MF, Rifkin SA, Hua S, Herreman T, Tongprasit W, Barbano PE, et al. (2004) A gene expression map for the euchromatic genome of Drosophila melanogaster. Science 306(5696): 655–60.

Testa G, Zhang Y, Vintersten K, Benes V, Pijnappel WWMP, Chambers I, Smith AJH, Smith AG and Stewart AF (2003) Engineering the mouse genome with bacterial artificial chromosomes to create multipurpose alleles. Nat Biotechnol 21(4): 443–7.

Thomas JL, Da Rocha M, Besse A, Mauchamp B and Chavancy G (2002) 3xP3-EGFP marker facilitates screening for transgenic silkworm Bombyx mori L. from the embryonic stage onwards. Insect Biochem Mol Biol 32(3): 247–53.

Thorpe HM and Smith MC (1998) In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proc Natl Acad Sci U S A 95(10): 5505–10.

Thummel CS and Pirrotta V (1992) New pCaSpeR P element vectors. Drosoph Inf Serv 71: 150.

Tirat A, Freuler F, Stettler T, Mayr LM and Leder L (2006) Evaluation of two novel tag-based labelling technologies for site-specific modification of proteins. Int J Biol Macromol 39(1-3): 66–76.

Tomancak P, Beaton A, Weiszmann R, Kwan E, Shu S, Lewis SE, Richards S, Ashburner M, Hartenstein V, Celniker SE, et al. (2002) Systematic determination of patterns of gene expression during Drosophila embryogenesis. Genome Biol 3(12): RESEARCH0088.

Tomancak P, Berman B, Beaton A, Weiszmann R, Kwan E, Hartenstein V, Celniker S and Rubin G (2007) Global analysis of patterns of gene expression during Drosophila embryogenesis. Genome Biol 8(7): R145.

Tomlinson A, Kimmel BE and Rubin GM (1988) rough, a Drosophila homeobox gene required in photoreceptors R2 and R5 for inductive interactions in the developing eye. Cell 55(5): 771–84.

Venken KJT, He Y, Hoskins RA and Bellen HJ (2006) P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. Science 314(5806): 1747–51.

Venken KJT, Bellen HJ and Hoskins RA (2009a) MIMIC transposable element ("Minos Mediated Integration Casette"). FlyBase communication .

Venken KJT, Carlson JW, Schulze KL, Pan H, He Y, Spokony R, Wan KH, Koriabine M, de Jong PJ, White KP, et al. (2009b) Versatile P[acman] BAC libraries for transgenesis studies in Drosophila melanogaster. Nat Methods 6(6): 431–4.

Viktorinová I, König T, Schlichting K and Dahmann C (2009) The cadherin Fat2 is required for planar cell polarity in the Drosophila ovary. Development 136(24): 4123–32.

Wang J, Sarov M, Rientjes J, Fu J, Hollak H, Kranz H, Xie W, Stewart AF and Zhang Y (2006) An improved recombineering approach by adding RecA to lambda Red recombination. Mol Biotechnol 32(1): 43–53.

Weiszmann R, Hammonds AS and Celniker SE (2009) Determination of gene expression patterns using high-throughput RNA in situ hybridization to whole-mount Drosophila embryos. Nat Protoc 4(5): 605–18.

Xie HB and Golic KG (2004) Gene deletions by ends-in targeting in Drosophila melanogaster. Genetics 168(3): 1477–89.

Yapici N, Kim YJ, Ribeiro C and Dickson BJ (2008) A receptor that mediates the post-mating switch in Drosophila reproductive behaviour. Nature 451(7174): 33–7.

Yeh E, Gustafson K and Boulianne GL (1995) Green fluorescent protein as a vital marker and reporter of gene expression in Drosophila. Proc Natl Acad Sci U S A 92(15): 7036–40.

Zhang Y, Buchholz F, Muyrers JP and Stewart AF (1998) A new logic for DNA engineering using recombination in Escherichia coli. Nat Genet 20(2): 123–8.

A Appendix A.

Reagents used

Contents

A.1. Kits

Name	Supplier	Catalog №
QIAEX II Gel Extraction Kit	QIAGEN	20021
QIAquick PCR Purification Kit	QIAGEN	28104
QIAquick Gel Extraction Kit	QIAGEN	28704
QIAprep Spin Miniprep Kit	QIAGEN	27104
QIAGEN Plasmid Midi Kit	QIAGEN	12143
QIAGEN Plasmid Maxi Kit	QIAGEN	12163
DNeasy Blood & Tissue Kit	QIAGEN	69504
RNeasy Mini Kit	QIAGEN	74104
AllPrep DNA/RNA/Protein Mini Kit	QIAGEN	80004
CopyControl™ HTP Fosmid Library Production Kit	EPICENTRE Biotechnologies	CCFOS059
TOPO® XL PCR Cloning Kit	Invitrogen	K4700-10
Zero Blunt® TOPO® PCR Cloning Kit	Invitrogen	K2860-20
SuperScript® III One-Step RT-PCR System	Invitrogen	12574-018

A.2. Antibodies

Name	Source	Clonality	Supplier	Catalog №
anti-GFP	rabbit	polyclonal	Invitrogen	A11122
anti-mCherry	rabbit	polyclonal	Clontech	632496
anti-rabbit IgG Cy2-conjugate	donkey	polyclonal	dianova	711-225-152

A.3. Enzymes

A.3.1. Polymerases and other DNA/RNA modifying enzymes

Name	Supplier	Catalog №
Taq DNA Polymerase	MPI-CBG	–
Phusion® High-Fidelity DNA Polymerase	NEB	F-530S
T4 DNA Ligase	NEB	M0202S
Antarctic Phosphatase	NEB	M0289S
T4 Polynucleotide Kinase	NEB	M0201S
DNA Polymerase I, Klenow Fragment	NEB	M0210S

A.3.2. Restriction Enzymes

Name	Recognition site	Supplier	Catalog №
*Afe*I	AGC^GCT	NEB	R0652S
*Alu*I	AG^CT	NEB	R0137S
*Apa*I	GGGCC^C	NEB	R0114S
*Apa*LI	G^TGCAC	NEB	R0507S
*Asc*I	GG^CGCGCC	NEB	R0558S
*Bam*HI	G^GATCC	NEB	R0136S
*Bam*HI-HF	G^GATCC	NEB	R3136S
*Bgl*II	A^GATCT	NEB	R0144S
*Bsp*EI	T^CCGGA	NEB	R0540S
*Cla*I	AT^CGAT	NEB	R0197S
*Dpn*I	GÃ^TC	NEB	R0176S
*Eco*RI	G^AATTC	NEB	R0101S
*Eco*RI-HF	G^AATTC	NEB	R3101S
*Eco*RV	GAT^ATC	NEB	R0195S
*Eco*RV-HF	GAT^ATC	NEB	R3195S
*Fse*I	GGCCGG^CC	NEB	R0588S
*Hind*III	A^AGCTT	NEB	R0104S
*Kpn*I	GGTAC^C	NEB	R0142S
*Kpn*I-HF	GGTAC^C	NEB	R3142S
*Nco*I	C^CATGG	NEB	R0193S
*Nco*I-HF	C^CATGG	NEB	R3193S
*Nhe*I	G^CTAGC	NEB	R0131S
*Nhe*I-HF	G^CTAGC	NEB	R3131S
*Not*I	GC^GGCCGC	NEB	R0189S
*Not*I-HF	GC^GGCCGC	NEB	R3189S
*Pac*I	TTAAT^TAA	NEB	R0547S
*Pml*I	CAC^GTG	NEB	R0532S
*Psi*I	TTA^TAA	NEB	R0657S
*Psp*XI	VC^TCGAGB	NEB	R0656S
*Pst*I	CTGCA^G	NEB	R0140S
*Pst*I-HF	CTGCA^G	NEB	R3140S
*Sfi*I	GGCCNNNN^NGGCC	NEB	R0123S
*Spe*I	A^CTAGT	NEB	R0133S
*Xba*I	T^CTAGA	NEB	R0145S
*Xho*I	C^TCGAG	NEB	R0146S

A.4. Bacterial strains

Name	Genotype	Source
SmartCells™	*F⁻ recA1 endA1 hsdR17 supE44 thi-1 gyrA96 relA1 Φ80lacZΔM15 Δ(lacZYA-argF)U169*	Genlantis
EPI300™-T1^R	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 recA1 endA1 araΔ139 Δ(ara, leu)7697 galU galK λ⁻ rpsL (Str^R) nupG trfA tonA*	EPICENTRE
EC100D™ *pir-116*	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZΔM15 ΔlacX74 recA1 endA1 araΔ139 Δ(ara, leu)7697 galU galK λ⁻ rpsL (Str^R) nupG pir-116(DHFR)*	EPICENTRE
ccdB Survival™-T1^R	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araΔ139 Δ(ara-leu)7697 galU galK λ⁻ rpsL (Str^R) nupG tonA::Ptrc -ccdA*	Invitrogen
TOP10	*F⁻ mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araΔ139 Δ(ara-leu)7697 galU galK λ⁻ rpsL (Str^R) nupG*	Invitrogen
dam⁻/dcm⁻	*ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbΔ1 R(zgb210::Tn10) Tet^S endA1 rspL136 (Str^R) dam13::Tn9 (Cm^R) xylA-5 mtl-1 thi-1 mcrB1 hsdR2*	NEB

A.5. Fly stocks

Species	Name	Genotype	Stock ID	Source
D. mel	*w⁻*	*y^1 w^{1118}*	n/a	MPI-CBG
D. mel	*Sp/CyO*	*y^1 w^{1118}; wg^{Sp-1}/ CyO*	n/a	MPI-CBG
D. mel	*MKRS/TM6B*	*y^1 w^{1118}; MKRS / TM6B, Tb^1*	n/a	MPI-CBG
D. mel	*GFP-NLS*	*y^1 w^{67c23}*; P{Ubi-GFP.nls}ID-2; P{Ubi-GFP.nls}ID-3	1691	Bloomington
D. mel	*attP2*	*y^1 sc^1 v^1 P{nos-phiC31\int.NLS}X; P{CaryP}attP2*	25710	Bloomington
D. mel	*attP40*	*y^1 v^1 P{nos-phiC31\int.NLS}X; P{CaryP}attP40*	25709	Bloomington
D. mel	*seq*	*y^1; Gr22b^1 Gr22d^1 cn^1 CG33964$^{R4.2}$ bw^1 sp^1; LysC^1 MstProx^1 GstD5^1 Rh6^1*	2057	Bloomington
D. pse	*seq*	*wt*	14011-0121.94	San Diego
D. sim	*seq*	*w^{501}*	14021-0251.195	San Diego
D. vir	*seq*	*b^1; tb^1 gp-L2^1; cd^1; pe^1*	15010-1051.87	San Diego

A.6. Oligonucleotides

A.6.1. PCR primers

Name	Sequence	Bases	GC [%]	Tm [°C]
R4MCSinAfwd	CAG AGA AGG AGG CAA ACA	18	50	48,0
R4MCSinBrev	CGT GGG GTT TGA ATT AAC T	19	42	46,8
R4MCSexFwd	AAA AAA CAA ACA AAA ATA AGA AGC GAG AGG A	31	29	55,1
R4MCSexRev	AAT AAG TGC GAG TGA AAG GAA TAG TAT TCT	30	33	56,2
HisBglIIcatF	TTC GAA GGC GTC CGA TTG CCG ACT GGC TTA GTA GGC TAG AAG ATC TGC GCT GAT GTC CG	59	56	76,4
HisNheIcatR	GAA GGG CAA CGT CAT TCT GTC GCA GGT CGA GGG TAC CAA AGC TAG CGG GTC GAA TTT GCT TT	62	53	75,9
CadAscIcatF	CGT CAT CAT ACC GCG GCT GGC GAA GAT TCC TAC TCC ATG ATG GCG CGC CGC GCT GAT GTC CG	62	63	79,8
CadNheIcatR	CAA CGT GGA CGA TGA CCA GGG CTG GCG CAT CGC CAT GAG CGC TAG CGG GTC GAA TTT GCT TT	62	60	78,5
EYFP_cEX_F	GGG TAC CAA AGC TAG CAT GGT GAG CAA GGG C	31	58	67,0
EYFP_cEX_R	GTA GGC TAG AAG ATC TTT ACT TGT ACA GCT CG	32	44	61,8
ECFP_cEX_R	GTA GGC TAG AGG CGC GCC TTA CTT GTA CAG CTC G	34	59	69,2
CherryNoXhoIFwd	TAC GAT GTT CGC TAG CAT GCA GAT CTT CGT GAA GAC TCT GAC TGG TAA GAC CAT CAC CCT GGA GGT TGA GCC CAG TGA	78	51	77,3
CherryBspEIrev	GAA CAT CGT ATC CGG ACA GGC TCG AGT TAC T	31	52	64,4
nlsT2Afwd	AGT AAC TCG AGC CTG AAG AGC AGG CAC AGA	30	53	64,4
nlsT2Arev	GAA CAT CGT ATC CGG ATG GGC CAG GAT TCT CCT CGA CGT CAC CGC ATG TTA GCA GAC TTC CTC TGC CCT CGC GTC TTC GTT CAC T	85	56	80,1
ubi_mCherry_fwd	TCA GAG GTG TGG TGG GGA AG	20	60	55,9
ubi_mCherry_rev_T7	TAA TAC GAG TCA CTA TAG GGA CTC GTT GTG GGA GGT GAT GT	41	46	67,5
EGFP_fwd	TGG AGA GGG TGA GGG TGA TG	20	60	55,9
EGFP_rev_T7	TAA TAC GAG TCA CTA TAG GGA CGA AAG GGC AGA TTG TGT GGA	42	45	67,4
CG4702_fwd	TAC CGG GAT GAG AGT GTG GC	20	60	55,9
CG4702_rev_T7	TAA TAC GAG TCA CTA TAG GGA CTC CGA GGT GCG ATA GGT TC	41	49	68,5

Continued on the next page

Name	Sequence	Bases	GC [%]	Tm [°C]
pR6K_core_fwd	GTA TGC ACT TCG CGG CCG CTT CAA AAA AAA GCC CGC TCA TTA GGC GGG CTG GGT TAC	57	56	76,1
pR6K_core_rev	CGA CGA CAA GGA TAT CTT GCC CTC ATC TGT TAC GCC G	37	54	68,9
dpse/MICAL_fwd	TAT TTC CTA AAG CAA AAA TCC CAC A	25	32	51,1
dpse/MICAL_rev	TAC CCG GAC AGT GCA TTA TAC TTT T	25	40	54,4
NheI_mCherry_dmel_fwd	ACT GAC TGG CTA GCA TGG TGT CCA AGG GCG AAG AGG ATA AT	41	51	69,5
XhoI_mCherry_dmsl_rev	ACT GAC TGC TCG AGC AGA CTG GAG TTC GAG GTA CAC CGC TCA TCC ATG C	49	57	74,6
EcoRI_mCherry_dmel_fwd	ACT GAC TGG AAT TCA TGG TGT CCA AGG GCG AAG AGG ATA AT	41	46	67,5
BamHI_mCherry_dmel_rev	ACT GAC TGG GAT CCC AGA CTG GAG TTC GAG GTA CAC CGC TCA TCC ATG C	49	57	74,6
XhoI_mCherry_dmel_fwd	ACT GAC TGC TCG AGA TGG TGT CCA AGG GCG AAG AGG ATA AT	41	51	69,5
EcoRI_mCherry_dmel_rev	ACT GAC TGG AAT TCC AGA CTG GAG TTC GAG GTA CAC CGC TCA TCC ATG C	49	53	72,9
NheI_eGFP_dmel_fwd	ACT GAC TGG CTA GCA TGG TCA GTA AGG GAG AAG AAT TGT TCA CGG GCG TAG T	52	50	72,5
XhoI_eGFP dmel_rev	ACT GAC TGC TCG AGC TTG TAC AAC TCA TCC ATG CCC AGC GTG ATT CCT	48	52	72,2
EcoRI_eGFP_dmel_fwd	ACT GAC TGG AAT TCA TGG TCA GTA AGG GAG AAG AAT TGT TCA CGG GCG TAG T	52	46	70,9
BamHI_eGFP_dmel_rev	ACT GAC TGG GAT CCC TTG TAC AAC TCA TCC ATG CCC AGC GTG ATT CCT	48	52	72,2
XhoI_eGFP_dmel_fwd	ACT GAC TGC TCG AGA TGG TCA GTA AGG GAG AAG AAT TGT TCA CGG GCG TAG T	52	50	72,5
EcoRI_eGFP_dmel_rev	ACT GAC TGG AAT TCC TTG TAC AAC TCA TCC ATG CCC AGC GTG ATT CCT	48	48	70,5
NheI_Cerulean_dmel_fwd	ACT GAC TGG CTA GCA TGG TGT CCA AGG GCG AGG AAC TGT TTA	42	52	70,4
XhoI_Cerulean_dmel_rev	ACT GAC TGC TCG AGC TAC TTA TAC AGT TCG TCC ATG CCG AGG GTG AT	47	51	71,5
EcoRI_Cerulean_dmel_fwd	ACT GAC TGG AAT TCA TGG TGT CCA AGG GCG AGG AAC TGT TTA	42	48	68,4
BamHI_Cerulean_dmel_rev	ACT GAC TGG GAT CCC TAC TTA TAC AGT TCG TCC ATG CCG AGG GTG AT	47	51	71,5
XhoI_Cerulean_dmel_fwd	ACT GAC TGC TCG AGA TGG TGT CCA AGG GCG AGG AAC TGT TTA	42	52	70,4
EcoRI_Cerulean_dmel_rev	ACT GAC TGG AAT TCC TAC TTA TAC AGT TCG TCC ATG CCG AGG GTG AT	47	47	69,8
NheI_tagRFP_dmel_fwd	ACT GAC TGG CTA GCA TGT CCG AAG AAT TGA TCA AGG AGA ACA TGC ACA	48	46	69,7
XhoI_tagRFP_dmel_rev	ACT GAC TGC TCG AGA TTC AGC TTG TGC CCC AGC TTA CT	38	53	68,8

Continued on the next page

Name	Sequence	Bases	GC [%]	Tm [°C]
EcoRI_tagRFP_dmel_fwd	ACT GAC TGG AAT TCA TGT CCG AAG AAT TGA TCA AGG AGA ACA TGC ACA	48	42	68,0
BamHI_tagRFP_dmel_rev	ACT GAC TGG GAT CCA TTC AGC TTG TGC CCC AGC TTA CT	38	53	68,8
XhoI_tagRFP_dmel_fwd	ACT GAC TGC TCG AGA TGT CCG AAG AAT TGA TCA AGG AGA ACA TGC ACA	48	46	69,7
EcoRI_tagRFP_dmel_rev	ACT GAC TGG AAT TCA TTC AGC TTG TGC CCC AGC TTA CT	38	47	66,6
NheI_Venus_dmel_fwd	ACT GAC TGG CTA GCA TGA GTA AGG GAG AGG AGC TAT TCA CGG GTG T	46	52	71,7
XhoI_Venus_dmel_rev	ACT GAC TGC TCG AGC TTA TAC AGT TCG TCC ATG CCA TGC GTA ATT C	46	48	69,9
EcoRI_Venus_dmel_fwd	ACT GAC TGG AAT TCA TGA GTA AGG GAG AGG AGC TAT TCA CGG GTG T	46	48	69,9
BamHI_Venus_dmel_rev	ACT GAC TGG GAT CCC TTA TAC AGT TCG TCC ATG CCA TGC GTA ATT C	46	48	69,9
XhoI_Venus_dmel_fwd	ACT GAC TGC TCG AGA TGA GTA AGG GAG AGG AGC TAT TCA CGG GTG T	46	52	71,7
EcoRI_Venus_dmel_rev	ACT GAC TGG AAT TCC TTA TAC AGT TCG TCC ATG CCA TGC GTA ATT C	46	43	68,1
EcoRV_pR6K_fwd	ACT GAC TGG ATA TCT TGC CCT CAT CTG TTA CGC	33	48	64,4
NotI_pR6K_rev	ACT GAC TGG CGG CCG CTT CAA AAA AAA GCC CGC TCA TTA GG	41	54	70,5
NheI_SGFP_fwd	ACT GAC TGG CTA GCA TGG TGT CCA AGG GCG AGG	33	61	69,4
XhoI_SGFP_fwd	ACT GAC TGC TCG AGA TGG TGT CCA AGG GCG AGG	33	61	69,4
XhoI_SGFP_rev	ACT GAC TGC TCG AGC TTG TAC AGC TCA TCC ATG CCC	36	56	69,0
EcoRI_SGFP_fwd	ACT GAC TGG AAT TCA TGG TGT CCA AGG GCG AGG	33	55	66,9
EcoRI_SGFP_rev	ACT GAC TGG AAT TCC TTG TAC AGC TCA TCC ATG CCC	36	50	66,7
BamHI_SGFP_rev	ACT GAC TGG GAT CCC TTG TAC AGC TCA TCC ATG CCC	36	56	69,0
EcoRI_V5_fwd	ACT GAC TGG AAT TCG GCA AGC CCA TCC CCA ACC CCC	36	61	71,3
BamHI_BLRP_rev	ACT GAC TGG GAT CCG CTG CCG CCG GCG TTG CTG C	34	71	74,1
XhoI_V5_fwd	ACT GAC TGC TCG AGG GCA AGC CCA TCC CCA ACC CCC	36	67	73,6
EcoRI_BLRP_fwd	ACT GAC TGG AAT TCA TGG CCA GCA GCC TGC GCC AG	35	60	70,3

Continued on the next page

Name	Sequence	Bases	GC [%]	Tm [°C]
BamHI_V5_rev	ACT GAC TGG GAT CCG GTG CTA TCC AGG CCC AGC AGG	36	64	72,4
XhoI_preTEV_fwd	ACT GAC TGC TCG AGC TGG AGG TGC TGT TCC AGG GC	35	63	71,5
NheI_preTEV_rev	ACT GAC TGG CTA GCG GGG CCC TGG AAC AGC ACC TCC	36	67	73,6
NheI_V5_rev	ACT GAC TGG CTA GCG GTG CTA TCC AGG CCC AGC AGG	36	64	72,4

A.6.2. Sequencing primers

Name	Sequence	Bases	GC [%]	Tm [°C]
M13uni(-21)	TGT AAA ACG ACG GCC AGT	18	50	48,0
M13rev(-29)	CAG GAA ACA GCT ATG ACC	18	50	48,0
HIS-rev-107R	TTC CAC TCA AAG TCA GC	17	47	44,6
HIS-fwd-189R	ATC GGA GTT GGA GGA TTC G	19	53	51,1
HISF-C1-199R	TAT GGA CAG CAA GCG AAC	18	50	48,0
HISR-C1-3434R	CGT GTC TTG TAG TTC CCG TC	20	55	53,8
HIS-rev-817R	CGG ACT GCT CTG TGT ATC AG	20	55	53,8
HIS-fwd-853F	CAG TTA GAA TCA CCG AGT GC	20	50	51,8
HISF-C1-2502F	AAG TGT GCT TCC GCC	15	60	44,7
HISR-C1-124R	GCC ACT GGT AAC AGG ATT AGC	21	52	54,4
CadA1-fwd-157R	TTT GTC AGC GGT TTC GTG	18	50	48,0
CadA1-fwd-886F	TAC GGC TTG CTG TTC TTC G	19	53	51,1
CadA1-C1R-225R	CGC TGT AGG TAT CTC AGT TCG	21	52	54,4
CadA1-C1R-2611F	CAT TGA ACA CCA TAG CAC AG	20	45	49,7
CadA1-rev-822F	GCT CCC ATT CAT CAG TTC C	19	53	51,1
CadA1-rev-118R	CAC TCA GAC TCA ATA CGA CAC TC	23	48	55,3
CadA1-C1F-860R	CCG CTC GAG CAT GCA TC	17	65	51,9
CadA1-C1F-2645F	CGA ACG GTC GCC TTC	15	67	47,4
CadA1-C2R-202R	GGA ACG GCA CTG GTC AAC	18	61	52,6
CadA1-C2R-4430F	GGT GCG TTC ACA TTG AGG	18	56	50,3
CadA1-C2F-231R	GAG AGG CTA TTC GGC TAT G	19	53	51,1
CadA1-C2F-3822F	TGA GCA CCA GTG TCC AGC G	19	63	55,4
CadA1-C3-111rev	CCT CGC ACT TGA ACT TCT C	19	53	51,1
CadA1-C3-10273fwd	AAC GAC CGA TTG ACA AGA G	19	47	48,9
CadA1-C4-155R	ACT GAT GCG AAT GGG G	16	56	45,9
CadA1-C4-11904F	GCC AAT ACG AAT ACC GAG G	19	53	51,1
hisEYFP_F	TCA TTC TGT CGC AGG	15	53	41,9
hisEYFP_R	CCG ACT GGC TTA GTA	15	53	41,9
cadECFP_F	ACG ATG ACC AGG GCT	15	60	44,7
cadECFP_R	CTG GCG AAG ATT CCT	15	53	41,9
pCC2FOSfwd	GTA CAA CGA CAC CTA GAC	18	50	48,0
pCC2FOSrev	CAG GAA ACA GCC TAG GAA	18	50	48,0
pTag4chkFwd	AGC GCT TCA TAC CCA TAC	18	50	48,0
pTag4chkRev	CAC CGG TTC ACG AAG TTC	18	56	50,3

Continued on the next page

Name	Sequence	Bases	GC [%]	Tm [°C]
pTag4seq_fwdA	GCT GGC TGG TTT ATT GCT	18	50	48,0
pTag4seq_revA	GAT AAC TAC GAT ACG GGA	18	44	45,8
pTag4seq_fwdB	CGA CAT CCC CGA CTA CTT	18	56	50,3
pTag4seq_revB	TCT TCT TCT GCA TTA CGG	18	44	45,8
pTag4seq_fwdC	GCT TAA GGA GGA CAA TCA	18	44	45,8
pTag4seq_revC	TGA CAA AAA GAA CCG GGC	18	50	48,0
pTag4seq_fwdD	CCT CAT GGC TAA CGT ACT	18	50	48,0
pTag4seq_revD	CAA ACA TGA GAG CTT AGT ACG	21	43	50,5
ubi-2NR	GGA TGC CTT CCT TAT CTT GG	20	50	51,8
T2A-1CF	GCT AAC ATG CGG TGA CGT CG	20	60	55,9
EGFP-N	CGC CGT CCA GCT CGA CCA	18	72	57,2
EGFP-C	ATG GTC CTG CTG GAG TTC GT	20	55	53,8
pTag_chk_1	GGG CGC AAG GGC TGC TAA	18	67	54,9
pTag_chk_2	CAA ACG ACG AGC GTG ACA CC	20	60	55,9
pTag_chk_3	CCG AAG GAG AGG GCA AGC	18	67	54,9
pTag_chk_4	ACC GAA GCA AGA AGC CCG CA	20	60	55,9
pTag_chk_5	AGG GCG ATG CCG AAA AGG T	19	58	53,2
pTag_chk_6	CGC TTG GGT GGA GAG GCT ATT	21	57	56,3
pTag_chk_7	AGT CCC TTC CCG CTT CAG T	19	58	53,2
pTag_chk_8	TCA CTG TCC CTT ATT CGC ACC T	22	50	54,8
pTag_chk_9	AGC GGA AAA CGG CCA CGA	18	61	52,6
pTag_chk_10	CGG CTG GGT GTG GCG GAC	18	78	59,4

A.6.3. Recombineering primers

Name	Sequence
CG4702_mCherry_fwd	TTT CAA ATA GGT TTA ACC CAT TCT CGT CTC GGT CTC TTC CAG TAG TCA TG gtg agc aag ggc gag gag gat aac a
CG4702_mCherry_rev	CTG GCT ACT ATG GAG AGC AGC AGG TAG GCG TGT ACT TTC CGT TGA TTC CA atc cat atg ttg tct ttc gaa ttt g
CG4702_uni_fwd	GTC GCT CCG GAT TCG ACT GTC GCC AGA TCT ATC TGG AGT GCA ACG AGG TC gaa gtg cat acc aat cag gac ccg c
CG4702_uni_rev	AAT TTA AAT TAT TTT AGG GAG AGT CCA GTC GGT GGC CAC TGG GAA GTC TA ctt gtc gtc gtc atc ctt gta gtc a
odd_mCherry_fwd	CAA TTT TAA GCC AAA TAA AAC TAC ACA AGG CCA ACA AAG ACA GTA TAA TG gtg agc aag ggc gag gag gat aac a
odd_mCherry_rev	AGC TCG TCA TCC ACG GTT ATG TTG CTG ATG GGT GAG GCC GAT GTG GAA GA atc cat atg ttg tct ttc gaa ttt g
odd_uni_fwd	AGA AGC CCA AGC GGA TGC TGG GCT TCA CCA TCG ATG AGA TCA TGA GCA GA gaa gtg cat acc aat cag gac ccg c
odd_uni_rev	GAA TGT CTC AAA AGA AAC CGG ATA CCG GGT TTC GCT GGT CCT TCA ATC TA ctt gtc gtc gtc atc ctt gta gtc a
numb_mCherry_fwd	AGT GCA GCG AAA CCA GCG AGT GCG AGC GAG AGG GCT AGC GAA CAG GCA TG gtg agc aag ggc gag gag gat aac a
numb_mCherry_rev	CGT GTG AAG CCG CGC TCG AGT GGT TCG TGC GTG TGT GAC GAG GAG TTT CC atc cat atg ttg tct ttc gaa ttt g
numb_uni_fwd	CGT TCA TCT CAC CGC CCA AGG CGC CGG CGC AGT CAT TCC AGG TGC AGC TC gaa gtg cat acc aat cag gac ccg c

Continued on the next page

Name	Sequence
numb_uni_rev	TTC AAG CGA ACC ACA CTT TGT CCC CCT GGT GAC CCC GCA ATC GCT GCC TA ctt gtc gtc gtc atc ctt gta gtc a
CG6113_mCherry_fwd	GTA GTA CGA GTG TAA CCG CTG AGA TTA GTC GTA AAA TCG GTG AAA TAA TG gtg agc aag ggc gag gag gat aac a
CG6113_mCherry_rev	CCC GAC TGC AGG AGG CAC AAA CTA AGC GCT ACA ATT AGC AAT TTC ACC GA atc cat atg ttg tct ttc gaa ttt g
CG6113_uni_fwd	TGT GGG ATC GAA TGC TGG AAA TAA TGC GAA ATC ATG AGA ATT CAA TTA TC gaa gtg cat acc aat cag gac ccg c
CG6113_uni_rev	ATA ATA TAC TGT GAA TAA CTA AAT TAT AAC CAT AGT TCT ATT CCC CTT TA ctt gtc gtc gtc atc ctt gta gtc a
sna_mCherry_fwd	TCT CGA TCA GTA CCG GAA ACT AAA ACT TAA TCA CAC ACA CAT CAA AAA TG gtg agc aag ggc gag gag gat aac a
sna_mCherry_rev	ACG AAG ACA ATG GGG CGC TTC TTT AGC GGG CAG CTT TTG TAG TTG GCG GC atc cat atg ttg tct ttc gaa ttt g
sna_uni_fwd	TGT CGC TCC TGA ACA AGC ACT CCA GCT CCA ACT GCA CCA TCA CTA TTG CG gaa gtg cat acc aat cag gac ccg c
sna_uni_rev	TTG CTT AGG TAA TTG TGT CCT GCT AAG GGA TTC ATA TGT CGA GAA TCC TA ctt gtc gtc gtc atc ctt gta gtc a
Ugt36Bc_mCherry_fwd	GAA CCA TAT CAG TTT CCA TTC GTA CTT GGA CTT GAA CGG AGC GAG TCA TG gtg agc aag ggc gag gag gat aac a
Ugt36Bc_mCherry_rev	ACC AAC AGG CCG CCC AAG CTG CAT CCA ATC CAT GTG CTC CGG TTT TGT GT atc cat atg ttg tct ttc gaa ttt g
Ugt36Bc_uni_fwd	TAT TTG GCA AAT CGA ACA AAG TTT CCA AGG GAA AGA AGG TGA AGA AGC AG gaa gtg cat acc aat cag gac ccg c
Ugt36Bc_uni_rev	TTT TAA TCT TTT AAG TAT AAA TTA GTT AAA ATT TCT ATA TTT CCA ACC TA ctt gtc gtc gtc atc ctt gta gtc a
CG1962_mCherry_fwd	CGC CAG AGT TCG CAT CAA GCC CAA GCT TAG AAA GGT CCA AGT CCA AGA TG gtg agc aag ggc gag gag gat aac a
CG1962_mCherry_rev	ATG AAC TGC GAT ACG TTT TCA CAG CCA GCC GAA CCG TGA TTG GAT TCC TC atc cat atg ttg tct ttc gaa ttt g
CG1962_uni_fwd	ACA ACG GAG CCA ACA TTC GCA AGA GTC ATC ATC ATC AGT TTC GTC AAA AG gaa gtg cat acc aat cag gac ccg c
CG1962_uni_rev	GGC TAG AAT GCA CTG TTT AAT TGC TTA TTT ACA TTC TCT AAA CAA GTT TA ctt gtc gtc gtc atc ctt gta gtc a
CG9336_mCherry_fwd	AAA ATC GTT TTC GAA AAG CAA TTC CCA CAC TCG AAG TAT TCG CGA AAA TG gtg agc aag ggc gag gag gat aac a
CG9336_mCherry_rev	GCC AGA CTG ATC ATA ACG GCC ACG GCC AAA CTG CAT TTC AGA GCG GAC AC atc cat atg ttg tct ttc gaa ttt g
CG9336_uni_fwd	CCA TCG CCG GAG CCA TCC TGC TCT TCT TCG GCG TGG CTC GTC TGC TGG CC gaa gtg cat acc aat cag gac ccg c
CG9336_uni_rev	AAG ATC GTT AAA TAC TAC GCA CAG GTA ATT TAC TAG CTA GTT AAG ATC TA ctt gtc gtc gtc atc ctt gta gtc a
Spn43Aa_mCherry_fwd	TGG CTG GGC CAT TTC ACT TTT AGT CTC GAG GTG TCG ACG CAG GCG CAA TG gtg agc aag ggc gag gag gat aac a
Spn43Aa_mCherry_rev	CTG GGC ATT AGG AAT CTT CGG TTA GGG TGC TCA CTA CTT AGC CAG TGG TT atc cat atg ttg tct ttc gaa ttt g
Spn43Aa_uni_fwd	TCA TTC GCG ACA AGC ACG CTG TCT ATT TCA CCG GAC ACA TTG TCA AGT TT gaa gtg cat acc aat cag gac ccg c
Spn43Aa_uni_rev	TAA TGA TTA CAA TTT AAA GTG CTT AAT CTG AGG GAA ATG TGT GAC GAT TA ctt gtc gtc gtc atc ctt gta gtc a

Continued on the next page

Name	Sequence
Optix_mCherry_fwd	GTG GAG AAT AGT ATG CCA ATT TGT TCA CAG TGG ATT CAA CGA ATA AAA TG gtg agc aag ggc gag gag gat aac a
Optix_mCherry_rev	GGC GAG AAG CTC TCT GAG GGC GGC TGT TTG CCC TCC GTC GGT CCA ACG GC atc cat atg ttg tct ttc gaa ttt g
Optix_uni_fwd	GGC CCT TCT CCA CGT CGC CGG AGC TGA AGC ACA GTG CTC CCG AGA TCA CA gaa gtg cat acc aat cag gac ccg c
Optix_uni_rev	ATT TAG TCT GGG TGG GGT TAC AGG TGG ACA CCT CAG ACC CGC TGA TAT CA ctt gtc gtc gtc atc ctt gta gtc a
CG8193_mCherry_fwd	TAG TTT GCT CCG CGA TCC AGC AGG TCC TCC CTG ACA TCC CAT TGA AAA TG gtg agc aag ggc gag gag gat aac a
CG8193_mCherry_rev	ACT GGC TCG GTG GGA TGG TCG AAA AGC AGG AGG AGA TTC TTC TTG TCG GC atc cat atg ttg tct ttc gaa ttt g
CG8193_uni_fwd	TCG TGG ACG TGA ACA TCC GCC ACG AGA ACC GCA CCG TGC AGC GCC CAA AC gaa gtg cat acc aat cag gac ccg c
CG8193_uni_rev	GGA ATC GGG TCA GGG CGA ATC GGG AAA GCG AAT CAG GTG CGG ATC GTC TA ctt gtc gtc gtc atc ctt gta gtc a
CG8850_mCherry_fwd	TCT AAA AAA TGG ATG ATG ATG CAG AAT ACC AGA AGC TCC GGC GAA ACA TG gtg agc aag ggc gag gag gat aac a
CG8850_mCherry_rev	ATG CCA TCA TTT GAT CCT GGA TGT CCT TCT CTC GAT CCT TGA GCT CCC TG atc cat atg ttg tct ttc gaa ttt g
CG8850_uni_fwd	CGG ACT TTA AGA GCC CAC GTG GCG GAT ACT TGT TCG ACA ATA TCT TTG GC gaa gtg cat acc aat cag gac ccg c
CG8850_uni_rev	TTG CAT TAA TAA TAA AGC GTA ACA CTA AAA TAC ACT AGA GTA GTC CTT TA ctt gtc gtc gtc atc ctt gta gtc a
Lac_mCherry_fwd	TTA AAT CGC GCG CTT GCA GGG TGT GGT GCT AAA AGT CAA TTT CTA AGA TG gtg agc aag ggc gag gag gat aac a
Lac_mCherry_rev	ATG GCC AGG AGC AGG GTG CTC CAC ACG CAA TTC GAG ATA CTC GGC CGC CA atc cat atg ttg tct ttc gaa ttt g
Lac_uni_fwd	CCA CTT CGT TCG CTC TTG TGG GCA TCC TGG CGG CGT TGC TCT TCG CCA GA gaa gtg cat acc aat cag gac ccg c
Lac_uni_rev	GGA TTG GAC CTG AAG CGC TTG GAG TCG ACG GCC GTG GGC CCA TTG GCT TA ctt gtc gtc gtc atc ctt gta gtc a
Mp20_mCherry_fwd	TTA GTG AAG ATC CCG CAG GAC CCG AAA CCA AAA ACC AAG AAT CAA ACA TG gtg agc aag ggc gag gag gat aac a
Mp20_mCherry_rev	TTC TTC ACT GAT TGA GTA TTC ACC TTG GCA CGA ACG GCA CGC TCA AGA GA atc cat atg ttg tct ttc gaa ttt g
Mp20_uni_fwd	CCC AGG CTG GCC AGA ACC TCG GCG CTG GCC GCA AGA TCC TGC TCG GCA AG gaa gtg cat acc aat cag gac ccg c
Mp20_uni_rev	GCA TAA GTG TAG AAA AGG GTG TGG ACA TCC TGG CCA TCC TTT GGC GCT TA ctt gtc gtc gtc atc ctt gta gtc a
CG17041_mCherry_fwd	AGG AAG AGA TCC GGT ATC TCT TCG CCA GAG AAC GGG TAA ACA AAG CGA TG gtg agc aag ggc gag gag gat aac a
CG17041_mCherry_rev	GCC TTT GAC TTG GAC TCC GGC ACG GAC CAT AGC AGA CCC AGT TGT TCC GG atc cat atg ttg tct ttc gaa ttt g
CG17041_uni_fwd	ACA CCA TTG AGA ACA TGC TGA TGG CCC TGC CCA GCG CCT CCA AGG CCA AG gaa gtg cat acc aat cag gac ccg c
CG17041_uni_rev	CAG CTT CTG CTG CCA GGG ATA AGC CGG TTA GCT GGG TGA ACG GCT CCT CA ctt gtc gtc gtc atc ctt gta gtc a
CG5493_mCherry_fwd	CTT CCG CTT CCC TTT TCG CAA CCT AGG TCA ATC AGA GCA AGC CCA AAA TG gtg agc aag ggc gag gag gat aac a

Continued on the next page

Name	Sequence
CG5493_mCherry_rev	TAC TTC TCC TGG TCA ACC TGC TCG ATC TCG GTG TCG ATC TTG GAC AGG GC atc cat atg ttg tct ttc gaa ttt g
CG5493_uni_fwd	AGG TCA CCT TCT GGA GCA AAT ACG GCG TGA GGA CGA AGC AGA ACG AGC AG gaa gtg cat acc aat cag gac ccg c
CG5493_uni_rev	TAT TAT GCA ACA ATA CGA GAG GCT ATA TTT TTA GGA TCT TGA TGG ATC TA ctt gtc gtc gtc atc ctt gta gtc a
CG9416_mCherry_fwd	CGG TGA CCA ATG AAG TGA ATA TGT TTA TCT CCC CTT TTT AGG CCA CAA TG gtg agc aag ggc gag gag gat aac a
CG9416_mCherry_rev	CCG ATC TTG CTG CGA TTG TAG ATG ATC TTC ATG TTT TCG TAC TTT GAT TT atc cat atg ttg tct ttc gaa ttt g
CG9416_uni_fwd	CGT GGG CGC ATG TTA GCT CCT GGC TGG GAT CCT ACA ATA GTT GGC AGC TT gaa gtg cat acc aat cag gac ccg c
CG9416_uni_rev	CAA AAA TGG GAT TGA GTT GAG TTC GTG AAG GAA CTT TGA GTC TAC ACT TA ctt gtc gtc gtc atc ctt gta gtc a
Obp56a_mCherry_fwd	GCA TCA GAA CTT CCC CAA CGT TCT AAC AAG TCA AAG TAT TTC TCA ACA TG gtg agc aag ggc gag gag gat aac a
Obp56a_mCherry_rev	ACA GCC AGA GTC ACA AAA AGA GCA CTC AAA GCG ATC ACG AAG TAG GAG TT atc cat atg ttg tct ttc gaa ttt g
Obp56a_uni_fwd	AGT TGT ACG ATT GCT TCG AGA GCT TCA AGC CCG CCC CCG AGG CTA AGG CC gaa gtg cat acc aat cag gac ccg c
Obp56a_uni_rev	ATT TTT TCC CGA ATC ACA ATT TGC CAA GCA TTA AAT CCC TAA CTT CTT TA ctt gtc gtc gtc atc ctt gta gtc a
CG13506_mCherry_fwd	AAA ACG AAC TCA CCA AAA CCC AGT CAG GCA AAC AAA CAC CAA TCA TCA TG gtg agc aag ggc gag gag gat aac a
CG13506_mCherry_rev	CCA ATT AGC AGA CTA ATG AGC AGC AGC CTC GTC GAA TCT CTC GCC TTG AT atc cat atg ttg tct ttc gaa ttt g
CG13506_uni_fwd	TGA ACG TGG GAG TGA TCC TGC TGG CAG CGC TCC TGC TGC GAG TCC GCC TC gaa gtg cat acc aat cag gac ccg c
CG13506_uni_rev	TCC GGA AAT ACG TAT GTA CAC ATC TCG GTC TGG ATG GGT GGC ACC CCT TA ctt gtc gtc gtc atc ctt gta gtc a
ken_mCherry_fwd	TGA CTA TAT TCA TCC TGG GAT TAA CCA ACT GCT GAA CAT CCA ACT TAA TG gtg agc aag ggc gag gag gat aac a
ken_mCherry_rev	ATG AAA ATC TCG GGG AAA AGC TTG CTG CCG AAC GGC AGA CTT ACC TCT TT atc cat atg ttg tct ttc gaa ttt g
ken_uni_fwd	ACC TGT CCG GCC ACC ACA ATA ACC TGC TGC TGA CAA AGA ATC TGC GCG AA gaa gtg cat acc aat cag gac ccg c
ken_uni_rev	GAC AGA TGG GAT ACT CTC GGA TTA TGC GGT TGC CGA TGC CCG GGA CAC TA ctt gtc gtc gtc atc ctt gta gtc a
gsb-n_mCherry_fwd	AGC GTC GCT TAG ATT CTC GAT TGC TAT AAG CTC TTT GCA TTC GGA CCA TG gtg agc aag ggc gag gag gat aac a
gsb-n_mCherry_rev	AAG GGA TAC CCT GCG AAA AGG GGC CGC AAC GAG TTC GCG CTG GAC ATA TC atc cat atg ttg tct ttc gaa ttt g
gsb-n_uni_fwd	CCG CCT ACA GCC ACC CCC TGC CGA CGC AGG GTC AGG CCA AGT ACT GGT CA gaa gtg cat acc aat cag gac ccg c
gsb-n_uni_rev	AAA TGT AAT GAT TCT CGG TCG TAC AGC GAC GTC GGC GAT TCA TTA AAT CA ctt gtc gtc gtc atc ctt gta gtc a
CG13920_mCherry_fwd	CCC AGT GAA CTC CGC CCC TCC GAG TAT TTA CCC ATA ACC GGG CCA AGA TG gtg agc aag ggc gag gag gat aac a
CG13920_mCherry_rev	CCC AGC AGC ACG GAG AGG CTC TTC AGC ACG ATC GTA TTG GAT GCA GGA GG atc cat atg ttg tct ttc gaa ttt g

Continued on the next page

Name	Sequence
CG13920_uni_fwd	ACG AGG CCA CGA CGG CAG CGC AGC CTG CTG CGA ACG GCG TAA AGC AGG AC gaa gtg cat acc aat cag gac ccg c
CG13920_uni_rev	TGG ATC AGT GAA TGA CTA TGA GTA CGT GAT GCA CCG GTA GCA GGC ACC TA ctt gtc gtc gtc atc ctt gta gtc a
CG12011_mCherry_fwd	AGC CCG AGA AAA TTC TAA ATT GGC ACA GTT CAA CTG AAA CCC TCA TCA TG gtg agc aag ggc gag gag gat aac a
CG12011_mCherry_rev	AAT ACT TAA CAA TTG TAC TAC CCA AGG ACT CGC TCT CTC GTT TGA CTT AC atc cat atg ttg tct ttc gaa ttt g
CG12011_uni_fwd	CGA AGC CCA TCT ATC GCT TCT TCA AGG GCA TCT TTG GCG GTT TCT CCA AC gaa gtg cat acc aat cag gac ccg c
CG12011_uni_rev	CTC GTT TCT TTG TTG TTC TTT CGG TAT TGA TCA CTT TAG TTA GCC TCT CA ctt gtc gtc gtc atc ctt gta gtc a
CG10591_mCherry_fwd	CGA ATT TTG GAT CTC AGT CCG ATC TGA AGA GAA ATC CGA AGT ACA TCA TG gtg agc aag ggc gag gag gat aac a
CG10591_mCherry_rev	TGT TTT ATA AAC TTA CCA GCA ACC AGA AAA GCC AAG AGT CCC AAA AAA CT atc cat atg ttg tct ttc gaa ttt g
CG10591_uni_fwd	ATG CCA CGC TCA TCC AAC CGC GTA ACT CAA ACC AAT ATG CAG TCA TCA TT gaa gtg cat acc aat cag gac ccg c
CG10591_uni_rev	CAG TTT TTA TTA AAT TAT TAA ACT AAA TTT AAT CTT AAG TTT TCC CTT CA ctt gtc gtc gtc atc ctt gta gtc a
CG32354_mCherry_fwd	AAT CAT CAA GCG TCT AAT AGG AAA AGT GCA GCA GAC AGC CAG CGA AAA TG gtg agc aag ggc gag gag gat aac a
CG32354_mCherry_rev	GCG TTT GGG GCG GGG GGC TGC AGA TGC AGA TGC AGA TGC TGT TTG TGG TG atc cat atg ttg tct ttc gaa ttt g
CG32354_uni_fwd	ACT ACG GCG CCT GTG GCC GCC CCG AAG CAC CAT CCA CTA ACT TCC TTT AC gaa gtg cat acc aat cag gac ccg c
CG32354_uni_rev	AAA TAT TAA ATG CCA AGT GAA ATG AAG ACG CCA CGC ATA CAT ACG TCC TA ctt gtc gtc gtc atc ctt gta gtc a
Doc3_mCherry_fwd	TCA CAA AAA TAA TAG ATT ACG CAC ATA GCT CCA CGA AGA CCC CAA ATA TG gtg agc aag ggc gag gag gat aac a
Doc3_mCherry_rev	TGG GCG ATC TGC TGC TGC AAT CGC AGA TCG GCG ATG TTG GGC AAG GTC AA atc cat atg ttg tct ttc gaa ttt g
Doc3_uni_fwd	AGC GCA GCA GCT TCA GCA TCT CGG ACA TAT TAG GAA CCA GCT CGT CCA TT gaa gtg cat acc aat cag gac ccg c
Doc3_uni_rev	ATT CCA AGC CAA ACG GGA GCA CAT GCA CCC GTC TGC ACT TGA TAG TTT TA ctt gtc gtc gtc atc ctt gta gtc a
CG14110_mCherry_fwd	AAA GCA GCG GGA TTT GTG TCA CTT GTC ACA GAA GTT GAC CAA CTG CAA TG gtg agc aag ggc gag gag gat aac a
CG14110_mCherry_rev	AGA ATC AGT AGC TGA ATC TTC CAG ACT GGA ACT CCG ACA CTA GTG GCT CC atc cat atg ttg tct ttc gaa ttt g
CG14110_uni_fwd	GAA ATT TTG AAC TTC AAA AAC TAA GAA GAG CCA ATA AAG TGC AAA AAT AT gaa gtg cat acc aat cag gac ccg c
CG14110_uni_rev	GAA TCC AAT TGA AAA CCC GTG GTT TTG TTT TCG TTA ATA ATG TAT ATC TA ctt gtc gtc gtc atc ctt gta gtc a
mex1_mCherry_fwd	GTA CAG TAT ATC TAT CTA TAA TAG AAT AAC CCA AAA AAG TCA TCA CCA TG gtg agc aag ggc gag gag gat aac a
mex1_mCherry_rev	CAG CAA ACC ACT TTG CCG GGA CAT TTG AGG CAT TCA CAG AGA GCG TTG CA atc cat atg ttg tct ttc gaa ttt g
mex1_uni_fwd	TGA CGC CCA TTG TGA AGC GCA GCA TAC GCG ACT ACT TCA ACA AGG AGT AC gaa gtg cat acc aat cag gac ccg c

Continued on the next page

Name	Sequence
mex1_uni_rev	CGG AGT TAT TAT TAA TAT GGC CTA TAG TAT AAA TTC ATG TCC TCA GAT CA ctt gtc gtc gtc atc ctt gta gtc a
comm2_mCherry_fwd	TAC AAA TAG AGA TTA CAC CCG CCA GTC GAC CGA TAT AAA AGT AAA CCA TG gtg agc aag ggc gag gag gat aac a
comm2_mCherry_rev	AAA TGC AAA TCG TGC GAG AGT TCG TAG TTT AAT GCG CGC GGC AAT TCC TC atc cat atg ttg tct ttc gaa ttt g
comm2_uni_fwd	AAA AGA AAT CGA AGT CAA AAG ACA GCC AGT CGA AAG ACG ATA TCA AGC GG gaa gtg cat acc aat cag gac ccg c
comm2_uni_rev	AGT TGC TTG GCT AGG TAT TCG TAT TCG TAT TCT CGT CAC TTT TAC CTT TA ctt gtc gtc gtc atc ctt gta gtc a
disp_mCherry_fwd	GCA ATG CGA AGA GGG TAA AGA GGA TTC GGG CAT CAC ATT CTA CTG ACA TG gtg agc aag ggc gag gag gat aac a
disp_mCherry_rev	CTG GCC AGG ACG TGG TAG TAC CAG TTC ATC CTC TCC GAG TCG AAG CAC AA atc cat atg ttg tct ttc gaa ttt g
disp_uni_fwd	TCC AGA CGA TGC ACG AGT GCA AAT ATC AAA CGT ATC CGT CTA CAT CCA AT gaa gtg cat acc aat cag gac ccg c
disp_uni_rev	GAT TCT TCT TGG CAA CAA CAA CGC GTT GTC CGG CAA TAG TAA CTA ACT CA ctt gtc gtc gtc atc ctt gta gtc a
pyd3_mCherry_fwd	TGA GTC CGA TAA TTG ATG AGA TAT TTT GTT GCT GTA AAA TTG GAA AAA TG gtg agc aag ggc gag gag gat aac a
pyd3_mCherry_rev	GGT AAA TGC TTT TCC AAG CAA TCA TTT AAA TTT TTC AGT TCA AAT GCT GA atc cat atg ttg tct ttc gaa ttt g
pyd3_uni_fwd	TCA AAA AGG CAT CCG AAC ATG GCT TCA AGC CGC AGA TCA TCA AGG AAA CA gaa gtg cat acc aat cag gac ccg c
pyd3_uni_rev	ATC AAT TCC GCT AAT CAA CAA AGT CAA TCG TAA AAT ACT CTT TTC TCT TA ctt gtc gtc gtc atc ctt gta gtc a
CG8147_mCherry_fwd	AGC TCG TAA ACT GAG AAA CTC TAA AAC TCA GAA GAA AGT ATA GAA AAA TG gtg agc aag ggc gag gag gat aac a
CG8147_mCherry_rev	CCT CCA CTG ACT AGA ACA CTC AGG CCG AGA AAA AAG AAG AGC TGC AGC CT atc cat atg ttg tct ttc gaa ttt g
CG8147_uni_fwd	TCA GTG ATC GTA ACA TGT GCG TGG ATG GGG GCG TGG CAC GGA GAC CAC GC gaa gtg cat acc aat cag gac ccg c
CG8147_uni_rev	CGC TTA GAC TCA TTG TGC TTC ACT TAA CCT TTG ATC CCC GCC AAG TTT CA ctt gtc gtc gtc atc ctt gta gtc a
CG14687_mCherry_fwd	AAA CGC GGA CTC AGA TTG CCA TTT TTG TTG CAG TGC ACC AGA GGA TCA TG gtg agc aag ggc gag gag gat aac a
CG14687_mCherry_rev	TCT TCG GAC TGC AAG GAC GAT GCT GAT TCC GGT GAG CTG GGG TAA ATG TA atc cat atg ttg tct ttc gaa ttt g
CG14687_uni_fwd	TCC GGG GAT TCA AAA CAC GCA AAG AAT TGA AAC AAT GCG AGC CCA TTG TG gaa gtg cat acc aat cag gac ccg c
CG14687_uni_rev	ATG TAC AGT AGA AAT CAG AGT TGT GGA ACG ACC AGC GCG CAA AGT CAT TA ctt gtc gtc gtc atc ctt gta gtc a
mfas_mCherry_fwd	TAC TTA GCT CCC AAC CGA GGC TCC AGA TTA AAA TTG TGA TAC CAA ACA TG gtg agc aag ggc gag gag gat aac a
mfas_mCherry_rev	GCC TGG ATC TGG ATT GAT CCC AGG AGG AGC AGG CAG GCC CAC AGC CGT AG atc cat atg ttg tct ttc gaa ttt g
mfas_uni_fwd	TCC CGC CCG GAG CTG GCT ATC AGC CAC AGG GCG ATT TCG ATG TCT TCT TC gaa gtg cat acc aat cag gac ccg c
mfas_uni_rev	ATA TCT GCT TTA TAT ATC AAT ATG CAT GAG ATG GGG ACG AGC ACC GCT CA ctt gtc gtc gtc atc ctt gta gtc a

Continued on the next page

Name	Sequence
Ect3_mCherry_fwd	CCA GTA TCT TAA TTG ATA TAT TAT CTT TCC TAC TGC AAT CCT TTA GAA TG gtg agc aag ggc gag gag gat aac a
Ect3_mCherry_rev	ACA GCT CCT AGC AGC GGT AGT AGG GCC ACC AGG ACC ACC ACA CTG AAC TT atc cat atg ttg tct ttc gaa ttt g
Ect3_uni_fwd	CAC AGG AGC TGC ACT TCC GAG ATA CCC CCA TTC TGA ACG CGA GGA CCG TT gaa gtg cat acc aat cag gac ccg c
Ect3_uni_rev	ATG GCT ACA GCG ATG GAT GCA AGT CGC AAT GGT CCT TAG GCC GAC GAC TA ctt gtc gtc gtc atc ctt gta gtc a
CG15887_mCherry_fwd	TAC TCA TCC TCA AGG AAT CAA ATC ACC AAC AGT CAA ATC AAA TCG AAA TG gtg agc aag ggc gag gag gat aac a
CG15887_mCherry_rev	CAT GCC AAA AAG AGG CAG ACG AGG GCG AAC ACA ATC TTC TGG TTG GCG GC atc cat atg ttg tct ttc gaa ttt g
CG15887_uni_fwd	TCT ACT CCC ACT CGC ACA CCC AGC AGC CCG TCT GGT TGG AGA AGG AGT GG gaa gtg cat acc aat cag gac ccg c
CG15887_uni_rev	CTG GTC TTG CAT TGT CAT CCT GGT CAG CCA AAA CCG AAA CAA ATC ATC TA ctt gtc gtc gtc atc ctt gta gtc a
Fsh_mCherry_fwd	GAA CAT TAT GTA CCT AGG ATC GCT GGA CGG AAA AGA CAG TGA GAG CCA TG gtg agc aag ggc gag gag gat aac a
Fsh_mCherry_rev	CTC GGA CGG TAA GTG GTA CCC ATC CGC TGG GAC AGA CTC GGG TGC TTT TC atc cat atg ttg tct ttc gaa ttt g
Fsh_uni_fwd	TGA TGG GCG CTG AGA CGC AGA AAA TGC TGA AGA ACA GCG AGG ATT ATG TT gaa gtg cat acc aat cag gac ccg c
Fsh_uni_rev	ATC ACA ACG GGG TGG GAA GAG GTG AGT TCA CAA TAC CCG GGT TCG GTT TA ctt gtc gtc gtc atc ctt gta gtc a
CG7678_mCherry_fwd	CGG AAT TGT TTT AAC CCA AGG AGC AAG GAT CAA CAG GAT CAG CTA TCA TG gtg agc aag ggc gag gag gat aac a
CG7678_mCherry_rev	CGG AAG ATG CTG TTG CTC TCC TGA TTG GAG CCG CAG CTC CAC CAC TTG GA atc cat atg ttg tct ttc gaa ttt g
CG7678_uni_fwd	CGT TCA CGC CAT TCA GCT TTA AGG ATA TTT TGA TCG TCG TCG AAG ATG AT gaa gtg cat acc aat cag gac ccg c
CG7678_uni_rev	ATT AAA CCT TCT TCA AGG AGC GAA TCA CGA TGG CTA TCC TGA TGG TTC TA ctt gtc gtc gtc atc ctt gta gtc a
CG18594_mCherry_fwd	ACT CGC TAA CGC TAC ACC GAG CAG AAG CAA CAG AAC TAC CAG CTA ACA TG gtg agc aag ggc gag gag gat aac a
CG18594_mCherry_rev	TTG GAG GCG GGC TTG ACG TCG ATG ATG TCG GGA ATA ATG CCG GCG GTG TC atc cat atg ttg tct ttc gaa ttt g
CG18594_uni_fwd	TCC AGG CCC AAT ACG ATG ACT ACG TGA AGA CCC TCA TCG AGA CGG TCC AG gaa gtg cat acc aat cag gac ccg c
CG18594_uni_rev	ATT TAA TAT TTA TTA TTT CAC AGA GAG CTG ATC AGT TGG TGG CCA GAT TA ctt gtc gtc gtc atc ctt gta gtc a
CG13653_mCherry_fwd	GAC ATC ACA TTC GCC ACA ACC ACC GAA CGA AGC ACA TCG ATC TGA AGA TG gtg agc aag ggc gag gag gat aac a
CG13653_mCherry_rev	CTT TGC TGG ATG GCC GCC AGC AAA ATC AGA ATG ATA GCT TTT CCC AAC TG atc cat atg ttg tct ttc gaa ttt g
CG13653_uni_fwd	GGG AGA GCG AAG AAA CCA AAT TGC ACG GCC CCG ACA ATG ATG ACT ACA TC gaa gtg cat acc aat cag gac ccg c
CG13653_uni_rev	ATA TGA ACC TCT ATT GCA CAG CCC TCC ATC ATC GAT AGT CTA GGG ATT TA ctt gtc gtc gtc atc ctt gta gtc a
HLHm5_mCherry_fwd	CTC AGC ACA TTT CTA CAA ATC TTC CAA AAC AAA AAA CAC ATT ACA AAA TG gtg agc aag ggc gag gag gat aac a

Continued on the next page

Name	Sequence
HLHm5_mCherry_rev	TAG TGC TGG GTC TTG GAG ACG AAT GTG GTG CTG TTG TTG CTC TGT GGT GC atc cat atg ttg tct ttc gaa ttt g
HLHm5_uni_fwd	AAT CCG CCG CCA GCC CCA AGC CAG TCG AAG AAA CCA TGT GGC GCC CTT GG gaa gtg cat acc aat cag gac ccg c
HLHm5_uni_rev	TCA GCT GGA AGA CTG GAT TCG ATG TCG ATG ATG ATG GTG ATG GTG ATT TA ctt gtc gtc gtc atc ctt gta gtc a
CG14253_mCherry_fwd	AAA GAC GTA TGT AAT TAG ATG CGG CTG CCA AGT GCC GCG GAT CAG AGA TG gtg agc aag ggc gag gag gat aac a
CG14253_mCherry_rev	AGC ATC TGC AGA CTA TCA CCA ATC CGG CGT ATC CTC TGC GGC CGG CTA CC atc cat atg ttg tct ttc gaa ttt g
CG14253_uni_fwd	CAC TTC TGG AGG TGC AAA GCC AGT CCC AGA TTC CGC CAA CTA GCT TGG CC gaa gtg cat acc aat cag gac ccg c
CG14253_uni_rev	GTA TAA AAA TAA ACC TGA GAT TGC TGA TCT GTT GGC TAC CCG CAG GAT TA ctt gtc gtc gtc atc ctt gta gtc a
Obp99a_mCherry_fwd	CGC TCG ATC GCT GGA GGA ATA CAT ACA TAG GTG GAA AGA AAG TGA AAA TG gtg agc aag ggc gag gag gat aac a
Obp99a_mCherry_rev	TAT CGA GCA CTC ACC AGT CCA ATC AGC ACG CAG ATG GCA ACG AAA ACC TT atc cat atg ttg tct ttc gaa ttt g
Obp99a_uni_fwd	TGC TGA AGG AGA ACC TGG CCC AGA TCC AGA AGA GCC TGG CCC CGA AGG CC gaa gtg cat acc aat cag gac ccg c
Obp99a_uni_rev	CTA AAC TAA TGC TTA TCG TTA CAT CCG TCC AAC TAG GAC CTA AGC CAC TA ctt gtc gtc gtc atc ctt gta gtc a
PH4alphaSG2_mCherry_fwd	GAA CCA GTC TAC ATC AGT AAC TCG TGG TTC ACA GTG CTC TGG TCA TAA TG gtg agc aag ggc gag gag gat aac a
PH4alphaSG2_mCherry_rev	ACC CAA ATT ATA AGC TGG AAA ATC CCA ATA TAA AGA CAG TGC CGA TCC AA atc cat atg ttg tct ttc gaa ttt g
PH4alphaSG2_uni_fwd	GTC CCT GCA ACC TCA CCT CAG ATA GCT ACA AGT CGC TAG CCT ATC GAG AT gaa gtg cat acc aat cag gac ccg c
PH4alphaSG2_uni_rev	TAA AAT AAG ATA TTA ATT TAT TAA ATG TTT ATT ATA TTA AAT TAC ACC TA ctt gtc gtc gtc atc ctt gta gtc a
l(1)sc_mCherry_fwd	GTC AAC ATC TGT AAA CTA AAT CTT AGA AAA CTC TCG CAA GGA TTA CCA TG gtg agc aag ggc gag gag gat aac a
l(1)sc_mCherry_rev	TTG GTC AGC TGG TAA TGC TGC TGC TGG AAT TTG CTG CTG CAA ATG CTC GT atc cat atg ttg tct ttc gaa ttt g
l(1)sc_uni_fwd	AGC CAG ATG ACG AGG AGC TAC TCG ATT ATA TTT CAT CTT GGC AAG AGC AG gaa gtg cat acc aat cag gac ccg c
l(1)sc_uni_rev	ACA GTT TGT ACA ATA TTT GTT TGT TTG GGA CTT TTA GTA AGA CCC CTT CA ctt gtc gtc gtc atc ctt gta gtc a
CG4194_mCherry_fwd	CTT GCT GCT CAG GTA GAA ACA ACA AAA AAC GAA TAT CAG TCG AGA AAA TG gtg agc aag ggc gag gag gat aac a
CG4194_mCherry_rev	GAT CCC AGG ATC ATA AGC AAA CAC ACA AAG ACT TCT GAA CGC GAG CGA AG atc cat atg ttg tct ttc gaa ttt g
CG4194_uni_fwd	CTA TTA CGC GAC TGG AGC TCT TGG CCG TCA AGA AGG GCT CGA ACA AAA AC gaa gtg cat acc aat cag gac ccg c
CG4194_uni_rev	GGA TAG ATG GAT TTA AGT GTT TGA TTT ACT TAG ATT TTC CAC ATA TTT TA ctt gtc gtc gtc atc ctt gta gtc a
CG15209_mCherry_fwd	GGC GGG CAG CAG TGC CAC AGC AAA GCT ACT AGC AGT CGG ACG TAA ACA TG gtg agc aag ggc gag gag gat aac a
CG15209_mCherry_rev	CCG AGG ATC ACC AGG AAG ATG ACC GGA ATA TTG TTG CAC CCA ATG GGT TT atc cat atg ttg tct ttc gaa ttt g

Continued on the next page

Name	Sequence
CG15209_uni_fwd	ACA ATG CCA AAG GAC CGG AGG AGC AGC CCA ATC AGG CCA TCG ATG AGC GT gaa gtg cat acc aat cag gac ccg c
CG15209_uni_rev	GTG AAC TTT GGC TTT AAA TAA ACG AGC TTA CTC ATT AAG CAA CTA AAT TA ctt gtc gtc gtc atc ctt gta gtc a
wgn_mCherry_fwd	TCT TAC AAA ATC CAT AAA AGT ATC GTT CTC TCG CTT CTC TGC TGC AGA TG gtg agc aag ggc gag gag gat aac a
wgn_mCherry_rev	CTC CGA CTA CGC ATG GCT CCT CCA TGG CCG CCT GGC AGT CTT GGC GGC AT atc cat atg ttg tct ttc gaa ttt g
wgn_uni_fwd	GCG GCA TGG GCG TGG GCC TGG GCG TCC GCG GCT GTT CCG GCC TGA AGG GC gaa gtg cat acc aat cag gac ccg c
wgn_uni_rev	GCC TGC ATA GCC ACT CCT TAA GTA TGT CCC TTG ACC ACA TCT ACG GCT CA ctt gtc gtc gtc atc ctt gta gtc a
run_mCherry_fwd	TCG ACA GCC CCA GGA TTA CGG CTA CGA TTT CCA CAT TCG GAT ACG AGA TG gtg agc aag ggc gag gag gat aac a
run_mCherry_rev	GCC AGG ACC TGT GTG TTG TTG GCC ACC ATC GTT GGA CCC GCT GGC AGA TG atc cat atg ttg tct ttc gaa ttt g
run_uni_fwd	CCA AGA TCA AGA GCG CCG CCG TGC AGC AGA AGA CCG TGT GGC GGC CCT AC gaa gtg cat acc aat cag gac ccg c
run_uni_rev	ATC ACT TTG TTT TCT TCA TTC CTC CAG ATT TTT GGG GAT CAG ATG CCC TA ctt gtc gtc gtc atc ctt gta gtc a

B Appendix B.

Protocols

Contents

Protocol 1.

Preparation and transformation of electrocompetent *E. coli*

This protocol describes a large scale (~100 aliquots) preparation of electrocompetent *E. coli* cells for DNA transformation. For highest efficiency, all the operations should be performed in a coldroom (4°C).

Reagents needed:

- LB medium
- Ice cold water
- Ice cold 10% glycerol
- Liquid nitrogen
- SOC medium

Preparation protocol:

1. Inoculate **50 ml** of LB with a single colony. Culture cells **overnight** at **37°C** with vigorous shaking.
2. Inoculate **1l** of LB with **5 ml** of an overnight culture.
3. Culture cells at **37°C** with vigorous shaking until OD(600) reaches **0.5**–**0.8**.
4. Chill cells **on ice** for **30 min**.
5. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
6. Resuspend cells in **1l** of **ice cold** water.
7. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
8. Resuspend cells in **500 ml** of **ice cold** water.
9. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
10. Resuspend cells in **50 ml** of **ice cold** 10% glycerol.
11. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
12. Resuspend cells in **5 ml** of **ice cold** 10% glycerol.
13. Aliquot **50 µl** of cell suspension into **2 ml** tubes.
14. Freeze aliquots in liquid nitrogen and store at **-80°C**.

Transformation protocol:

1. Prepare DNA for transformation.
   *If the DNA solution contains salt, perform a microdialysis on the Millipore 0.025 µm VSWP filter for **1–2** hours.*
2. Thaw an aliquot (**50 µl**) of electrocompetent cells on ice.
3. Add DNA solution (**up to 10 µl**) to bacteria and mix by pipetting. Incubate on ice for **5 min**.
4. Transfer bacteria to 1 or 2 mm electroporation cuvette and electroporate at **1.5–3 kV**. Expect time-constant in the range of **4.5–5.3 ms**.
5. **Immediately** add **1 ml** of SOC and move bacteria to a 2 ml tube.
6. Culture for **1h** at **37°C**.
7. Plate on LA with respective antibiotics.

Protocol 2.

Preparation and transformation of ultracompetent *E. coli*

This protocol describes a large scale (~100 aliquots) preparation of ultracompetent *E. coli* cells for DNA transformation. For highest efficiency, all the operations should be performed in a coldroom (4°C).

Reagents needed:

- SOB medium
- TB, pH 6.7 (10 mM PIPES, 15 mM $CaCl_2$, 250 mM KCl, 55 mM $MnCl_2$)
- DMSO
- Liquid nitrogen
- SOC medium

Preparation protocol:

1. Inoculate **5 ml** of SOB with a single colony. Culture cells **overnight** at **37°C** with vigorous shaking.
2. Inoculate **100 ml** of SOB with **500 µl** of an overnight culture.
3. Culture cells at **19°C** with vigorous shaking until OD(600) reaches **0.5**.
4. Chill cells **on ice** for **30 min**.
5. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
6. Resuspend cells in **50 ml** of **ice cold** TB.
7. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
8. Resuspend cells in **5 ml** of **ice cold** TB + **350 µl** DMSO.
9. Aliquot **50 µl** of cell suspension into **2 ml** tubes.
10. Freeze aliquots in liquid nitrogen and store at **-80°C**.

Transformation protocol:

1. Prepare DNA for transformation.
2. Thaw an aliquot (**50 µl**) of chemocompetent cells on ice.
3. Add DNA solution (**up to 10 µl**) to bacteria and mix by pipetting. Incubate on ice for **5 min**.
4. Heat shock bacteria at **42°C** for **45 seconds**.
5. **Immediately** add **1 ml** of SOC and move bacteria to a 2 ml tube.
6. Culture for **1h** at **37°C**.
7. Plate on LA with respective antibiotics.

Protocol 3.

Preparation and transformation of electrocompetent recombineering-ready *E. coli*

This protocol describes a large scale (~100 aliquots) preparation of electrocompetent recombineering-ready *E. coli* cells for Red/ET recombination. The cells can be co-transformed with the target plasmid and the recombineering cassette. For highest efficiency, all the operations should be performed in a coldroom (4°C).

Reagents needed:

- LB medium
- SOC medium
- LB + Tetracycline (10 μg/ml)
- LB + Tetracycline (10 μg/ml) + L-arabinose (0.1%)
- Ice cold water
- Ice cold 10% glycerol
- Liquid nitrogen

Preparation protocol:

1. Inoculate **1 ml** of LB with a single colony. Culture cells **overnight** at **37°C** with vigorous shaking.
2. Use **20 μl** of the overnight culture to inoculate **1 ml** of LB. Culture cells for **2h** at **37°C**.
3. Centrifuge cells at **10,000g** for **30 sec** at **2°C**. Remove supernatant entirely.
4. Resuspend cells in **1 ml** of **ice cold** water.
5. Centrifuge cells at **10,000g** for **30 sec** at **2°C**. Remove supernatant entirely.
6. Resuspend cells in **500 ml** of **ice cold** water.
7. Centrifuge cells at **10,000g** for **30 sec** at **4°C**. Remove supernatant entirely.
8. Resuspend cells in **50 μl** of **ice cold** 10% glycerol.
9. Add **1 μl** of pSC101-BAD-gbaA (**100 ng/μl**) to the cells. Mix briefly by pipetting. Transfer the cell suspension into a chilled **2 mm** electroporation cuvette.
10. Electroporate at **3000V**, **25 μF**, **200 Ω**.
11. **Immediately** add **1 ml** of SOC and transfer bacteria into a new 2 ml tube.

12. Culture cells for **1h** at **30°C**.
13. Inoculate **50 ml** of LB+Tet^{10} with **1 ml** of transformed bacteria. Wrap the flask with an aluminium foil and culture **overnight** at **30°C** with vigorous shaking.
14. Inoculate **1l** of LB+Tet^{10}+Ara$^{0.1\%}$ with **25 ml** of an overnight culture.
15. Wrap the flask with an aluminium foil and culture cells at **30°C** with vigorous shaking until OD(600) reaches **0.5**–**0.8**.
16. Chill cells **on ice** for **30 min**.
17. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
18. Resuspend cells in **1l** of **ice cold** water.
19. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
20. Resuspend cells in **500 ml** of **ice cold** water.
21. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
22. Resuspend cells in **50 ml** of **ice cold** 10% glycerol.
23. Centrifuge cells at **6,000g** for **15 min** at **4°C**. Remove supernatant entirely.
24. Resuspend cells in **5 ml** of **ice cold** 10% glycerol.
25. Aliquot **50 µl** of cell suspension into **2 ml** tubes.
26. Freeze aliquots in liquid nitrogen and store at **-80°C**.

Transformation protocol:

1. Prepare plasmid DNA and the PCR-amplified recombineering cassette for transformation.
   *If the DNA solutions contain salt, perform a microdialysis on the Millipore 0.025 µm VSWP filter for **1–2** hours.*
2. Thaw an aliquot (**50 µl**) of electrocompetent cells on ice.
3. Add mixed plasmid and recombineering cassette solution (**up to 10 µl**) to bacteria and mix by pipetting. Incubate on ice for **5 min**.
4. Transfer bacteria to 1 or 2 mm electroporation cuvette and electroporate at **1.5**–**3 kV**. Expect time-constant in the range of **4.5**–**5.3 ms**.
5. **Immediately** add **1 ml** of SOC and move bacteria to a 2 ml tube.
6. Culture for **1h** at **37°C**.
7. Plate on LA with respective antibiotics.

Protocol 4.

HMW genomic DNA isolation from *Drosophila* embryos

This protocol describes a quick way of isolating high quality and molecular weight genomic DNA from *Drosophila* embryos. DNA isolated in the described way is suitable for genomic library production.

Reagents needed:

- 1x PBS
- 1x PBT
  (0.1% Tween 20 in 1x PBS)
- 100% n-Heptane
- 100% Methanol
- Lysis buffer
  (50 mM Tris-HCl pH 8.0, 100 mM EDTA, 100 mM NaCl, 0.5% SDS, 50 µg/ml Proteinase K, 100 µg/ml RNAse A)
- Phenol:Chloroform:Isoamyl alcohol (25:24:1) pH 7.5
- Chloroform:Isoamyl alcohol (24:1)
- 3M Potassium acetate, pH 5.2
- 100% Isopropanol
- 70% Ethanol
- 1x TE

Protocol:

1. Collect embryos for **24h**.
   *Optional: Let them age for up to **12h** at **room temperature**.*
2. Decorionate embryos for **2 minutes** in **100%** bleach fluid.
3. Wash embryos with 1x PBS
4. Wash embryos with 1x PBT
5. Transfer embryos into a bottle containing **1 volume** of PBS and **1 volume** of n-Heptane. Use **20 ml** of PBS per **1 ml** of embryos. Mix by briefly shaking the bottle.
6. Remove PBS (lower phase). Leave the interphase intact.
7. Add **1 volume** of methanol and shake vigorously by hand for **1 minute**.
8. Remove n-heptane and interphase.
9. Transfer embryos into the Falcon tube and wash **twice** with **1 volume** of methanol.
10. Remove methanol completely.

11. Add **1 volume** of lysis buffer. Lyse for **2–3 hours** at **55°C**. Gently mix by inverting the tube every 15 minutes.
12. Centrifuge at **4,000g** for **30 minutes**. Transfer supernatant to a new Falcon tube. *Optional: Remove **200 µl** for quality analysis.*
13. Add **1 volume** of Phenol:Chloroform:Isoamyl alcohol. Incubate on a rotating wheel or a nutator for **1 hour** at **4°C**.
14. Centrifuge at **4,000g** for **10 minutes**. Transfer aqueous (upper) phase to a new Falcon tube.
15. Repeat steps 13–14.
16. Add **1 volume** of Chloroform:Isoamyl alcohol. Incubate on a rotating wheel or a nutator for **1 hour** at **4°C**.
17. Centrifuge at **4,000g** for **10 minutes**. Transfer aqueous (upper) phase to a new Falcon tube.
18. Add **0.05 volume** of **3M** KAc. Mix by gently inverting the tube.
19. Add **0.7 volume** of isopropanol. Incubate on a rotating wheel or a nutator for **30 minutes** at **4°C**.
20. Centrifuge at **6,000g** for **15 minutes**. Remove supernatant.
21. Wash the pellet **twice** with **1 volume** of 70% ethanol.
22. Air-dry the pellet for **10 minutes** at **room temperature**.
23. Dissolve the pellet in 1x TE prewarmed to **55°C**. Store DNA at **4°C**.

Protocol 5.

Isolation of FlyFos fosmid DNA from *E. coli*

This protocol describes an efficient way of isolating injection-quality fosmid DNA from FlyFos clones. The ▲ MidiPrep and ● MaxiPrep protocols are included. The protocol is based on QIAGEN Plasmid Purification Handbook.

Reagents needed:

- LB + Chloramphenicol (25 μg/ml)
- LB + Chloramphenicol (25 μg/ml) + L-Arabinose (0.1%)
- Buffer P1 (QIAGEN)
- Buffer P2 (QIAGEN)
- Buffer P3 (QIAGEN)
- Buffer QBT (QIAGEN)
- Buffer QC (QIAGEN)
- Buffer QF (QIAGEN)
- 100% Isopropanol
- 70% Ethanol
- 1x TE

Protocol:

1. Inoculate **50 ml** of LB+Cm^{25} with a single colony of FlyFos strain. Culture overnight at **37°C** with vigorous shaking.
2. Use ▲ 2 x 1 ml or ● 2 x 5 ml to inoculate ▲ 2 x 100 ml or ● 2 x 500 ml LB+Cm^{25}+$Ara^{0.1\%}$ in ▲ 500 ml or ● 2500 ml flasks. Culture overnight at 37°C. Shake cultures vigorously – **250 rpm** in a bare minimum.
3. Harvest the bacterial cells by centrifugation at **6,000g** for **15 min** at **4°C**.
4. Resuspend the bacterial pellet from both flasks combined in ▲ 8 ml or ● 50 ml of Buffer P1.
5. Add ▲ 8 ml or ● 50 ml of Buffer P2, mix thoroughly by vigorously inverting 4–6 times, and incubate at **room temperature** for **5 min**.
6. Add ▲ 8 ml or ● 50 ml of chilled Buffer P3, mix immediately and thoroughly by vigorously inverting 4–6 times, and incubate **on ice** for **30 min**.
7. Centrifuge at ≥**20,000g** for **30 min** at **4°C**. Remove supernatant containing fosmid DNA promptly.
8. Place folded Whatmann filter in a 50 ml syringe. Prewet and compress filter by passing water through the syringe. Use such prepared syringe for filtering supernatant.

9. Precipitate the DNA by adding ▲ 17 ml or ● 105 ml (**0.7 volumes**) of room temperature isopropanol to the lysate. Centrifuge at ≥**15,000g** for **30 min** at **4°C**, and carefully decant the supernatant.
10. Redissolve the DNA pellet in **500 µl** warm (**60°C**) TE buffer, pH 8.0, and add Buffer QBT to obtain a final volume of ▲ 5 ml or ● 12 ml for selected ▲ QIAGEN-tip 100 or ● QIAGEN-tip 500, respectively.
11. Equilibrate a ▲ QIAGEN-tip 100 or ● QIAGEN-tip 500 by applying ▲ 4 ml or ● 10 ml Buffer QBT, and allow the column to empty by gravity flow.
12. Apply the DNA solution from step 10 to the QIAGEN-tip and allow it to enter the resin by gravity flow.
13. Wash the QIAGEN-tip with ▲ 2 x 10 ml or ● 2 x 30 ml Buffer QC.
14. Elute DNA with ▲ 5 ml or ● 15 ml Buffer QF.
15. Precipitate DNA by adding ▲ 3.5 ml or ● 10.5 ml (**0.7 volumes**) of room temperature isopropanol to the eluted DNA. Mix and centrifuge immediately at ≥**15,000g** for **30 min** at **4°C**. Carefully decant the supernatant.
16. Wash DNA pellet with ▲ 2 ml or ● 5 ml room-temperature 70% ethanol, and centrifuge at ≥**15,000g** for **10 min**. Carefully decant the supernatant without disturbing the pellet.
17. Wash DNA pellet again with ▲ 2 ml or ● 5 ml room-temperature 70% ethanol, and centrifuge at ≥**15,000g** for **10 min**. Carefully decant the supernatant without disturbing the pellet.
18. Air-dry the pellet for 5–10 min, and redissolve the DNA in a suitable volume (▲ 50 µl or ● 250 µl) of warm (**60°C**) nuclease-free water.
19. You should obtain in total ▲ 100 µg or ● 500 µg of pure injection-quality fosmid DNA.

Protocol 6.

Manual 96-well and automated 384-well MiniPrep

This high-throughput protocol describes an efficient way of isolating sequencing-quality fosmid DNA from FlyFos clone cultures in multiwell format. The ▲ manual 96-well and ● automated 384-well protocols are included.

Reagents needed:

- LB + Chloramphenicol (25 µg/ml)
- LB + Chloramphenicol (25 µg/ml) + L-Arabinose (0.1%)
- Buffer P1 (QIAGEN)
- Buffer P2 (QIAGEN)
- Buffer P3 (QIAGEN)
- 100% Isopropanol
- 70% Ethanol

Protocol:

1. Use ▲ **50 µl** or ● **5 µl** of the primary culture to inoculate ▲ **1000 µl** or ● **100 µl** of LB+Cm^{25}+$Ara^{0.1\%}$. Seal plates with air-permable seal and culture overnight at 37°C with vigorous shaking.
2. Harvest the bacterial cells by centrifugation at **6,000g** for **15 min** at **4°C**. ▲ Discard supernatant by inverting plates over the sink and placing them on a stack of paper towels. ● Remove supernatant by aspirating 1 mm from the well bottom at speed of 10 µl/s, move within a well at 50% of speed. Discard supernatant to the waste container. Wash tips in ethanol and the wash station after pipetting is finished.
3. Transfer ▲ **350 µl** or ● **15 µl** of Buffer P1 to each well. ● Wash tips in the wash station after pipetting is finished.
4. Vortex plates vigorously to resuspend bacteria.
5. Transfer ▲ **350 µl** or ● **15 µl** of Buffer P2. ▲ Mix by inverting sealed plate 4–6 times. ● Wash tips in the wash station after pipetting is finished.
6. Incubate plates at **room temperature** for **5 min**.
7. Transfer ▲ **350 µl** or ● **15 µl** of Buffer P3. ▲ Mix by vigorously inverting sealed plate 4–6 times. ● Wash tips in the wash station after pipetting is finished.
8. Centrifuge plates at ≥**6,000g** for **45 min** at **4°C**.

9. Transfer ▲ **900 µl** or ● **40 µl** of supernatant into the new plates. ▲ Be careful to avoid touching the precipitate. If transferred supernatant contains precipitate, repeat centrifugation (step 8) and transfer supernatant into the new plates. ● Wash tips in the wash station between each pipetting step. Aspirate 2 mm from the well bottom at speed of 10 µl/s, move within a well at 50% of speed.
10. Precipitate DNA by adding ▲ **600 µl** or ● **25 µl** (~ 0.7 volume) of isopropanol into each well.
11. Mix by vortexing and centrifuge plates at ≥**6,000g** for **45 min** at **4°C**. ▲ Discard supernatant by inverting plates over the sink and placing them on a stack of paper towels. ● Remove supernatant by aspirating 2 mm from the well bottom at speed of 10 µl/s, move within a well at 50% of speed. Discard supernatant to the waste container. Wash tips in the wash station after pipetting is finished.
12. Wash DNA pellet with ▲ **1000 µl** or ● **75 µl** of 70% ethanol, and centrifuge at ≥**6,000g** for **15 min**. ▲ Discard supernatant by inverting plates over the sink and placing them on a stack of paper towels. ● Remove supernatant by aspirating 2 mm from the well bottom at speed of 10 µl/s, move within a well at 50% of speed. Discard supernatant to the waste container. Wash tips in the wash station after pipetting is finished.
13. Wash DNA pellet again with ▲ **1000 µl** or ● **75 µl** of 70% ethanol, and centrifuge at ≥**6,000g** for **15 min**. ▲ Discard supernatant by inverting plates over the sink and placing them on a stack of paper towels. ● Remove supernatant by aspirating 2 mm from the well bottom at speed of 10 µl/s, move within a well at 50% of speed. Discard supernatant to the waste container. Wash tips in the wash station after pipetting is finished.
14. Place inverted plates on a stack of paper towels. Allow the remaining ethanol to be completely absorbed through capillary forces. Replace towels when they become wet.
15. Air-dry the plates for **15–30 min**.
16. Redissolve the DNA in ▲ **200 µl** or ● **20 µl** of nuclease-free water.

Protocol 7.

Fosmid library production

This protocol describes the production of fosmid genomic libraries for Drosophilidae. The protocol is based on the manual for EPICENTRE CopyControl™ Fosmid Library Production Kit and HydroShear device user manual.

Reagents needed:

- 0.5x TBE
- 10 mg/ml Ethidium bromide (EtBr)
- Bio-Rad Pulse Field Agarose
- SeaPlaque LMP Agarose (LONZA)
- MidRange II PFG Marker (NEB)
- Fosmid Control DNA (EPICENTRE)
- 10x End-Repair Buffer (EPICENTRE)
- 2.5 mM dNTP Mix (EPICENTRE)
- 10 mM ATP (EPICENTRE)
- End-Repair Enzyme Mix (EPICENTRE)
- GELase 50x Buffer (EPICENTRE)
- GELase Enzyme (EPICENTRE)
- 3M Potassium acetate (KAc), pH 7.0
- 5M Lithium chloride (LiCl)
- 100% Isopropanol
- 100% Ethanol
- 70% Ethanol
- NEBuffer 1 (NEB)
- *Pml*I (*Eco*72I) restriction enzyme (NEB)
- Bovine Serum Albumin (BSA) 10 mg/ml (NEB)
- Antarctic phosphatase (NEB)
- Antarctic phosphatase buffer (NEB)
- T4 DNA Ligase (NEB)
- 10x Ligase Buffer (NEB)
- QIAquick Gel Extraction Kit (QIAGEN)
- LB
- LB + $MgSO_4$(10 mM) + Maltose (0.2%)
- MaxPlax Lambda Packaging Extract (EPICENTRE)
- Phage Dilution Buffer (10 mM Tris-HCl pH 8.3, 100 mM NaCl, 10 mM $MgCl_2$)
- LA plates + Chloramphenicol (15 µg/ml)

Protocol:

Shearing and end-repair of the genomic DNA

1. Isolate high molecular weight genomic DNA as described in protocol 4.
2. Dilute the genomic DNA to final concentration of **250 ng/µl** with water.

3. Shear the DNA using the HydroShear device (DigiLab). Use **4–40 kb** (large) **shearing assembly**. Since every shearing assembly has slightly different shearing properties, test different speedcodes by shearing about **5 µg** of the DNA (minimal shearing volume is 50 µl – dilute the DNA accordingly). The following parameters are suggested: speedcode 17, retraction speed 40, 25 shearing cycles, 200 µl sample volume.

4. Verify the shearing results by running a pulse-field gel electrophoresis (PFGE) with **1 µg** of the sheared DNA. Include Fosmid Control DNA (**100 ng**) and MidRange II PFG Marker (**500 ng**) on the gel for reference. The following parameters are suggested for the Bio-Rad CHEF Mapper XA system. Use **0,8%** Bio-Rad Pulse-Field Agarose in **0.5x TBE**. Setup a two-state program at **6.0 V/cm**, initial switch at **1.5s**, final switch at **7.0s**, **120°** angle and **linear** ramping factor. Run the gel at **14°C** for **19h 45 mins**.

5. Stain the gel for **30 minutes** with **0.5 µg/ml** ethidium bromide in **0.5x** TBE.

6. Destain the gel for **1 hour** in **0.5x** TBE.

7. Visualize the sheared DNA in UV and determine the best shearing conditions. Choose the speedcode that produces maximal amount of DNA in the range of **30–60 kb** and nearly **no DNA below 20 kb**. Including fragments smaller than 20 kb in the library production process may result in large number of chimeric clones. If you cannot find the shearing conditions that yield DNA that is directly suitable for library production, you will need to size-select the DNA.

8. Shear **100 µg** of the genomic DNA (2 x 200 µl) using the determined conditions. Use the newly sheared DNA for further processing.

9. Setup an end-repair reaction. If you intend to size-select the DNA by PFGE, use **80 µg** of sheared DNA in a **240 µl** reaction. Otherwise set up an **80 µl** reaction using **20 µg** sheared DNA, **8 µl** 10x End-Repair Buffer, **8 µl** 2.5 mM dNTP Mix, **8 µl** 10 mM ATP, **4 µl** End-Repair Enzyme Mix and water up to **80 µl**.

10. Incubate the reaction at **room temperature** for **45 minutes**.

11. Heat-inactivate the End-Repair Enzyme Mix at **55°C** for **10 minutes**. If you do not need to size select the DNA by PFGE, proceed directly to final purification of the genomic DNA.

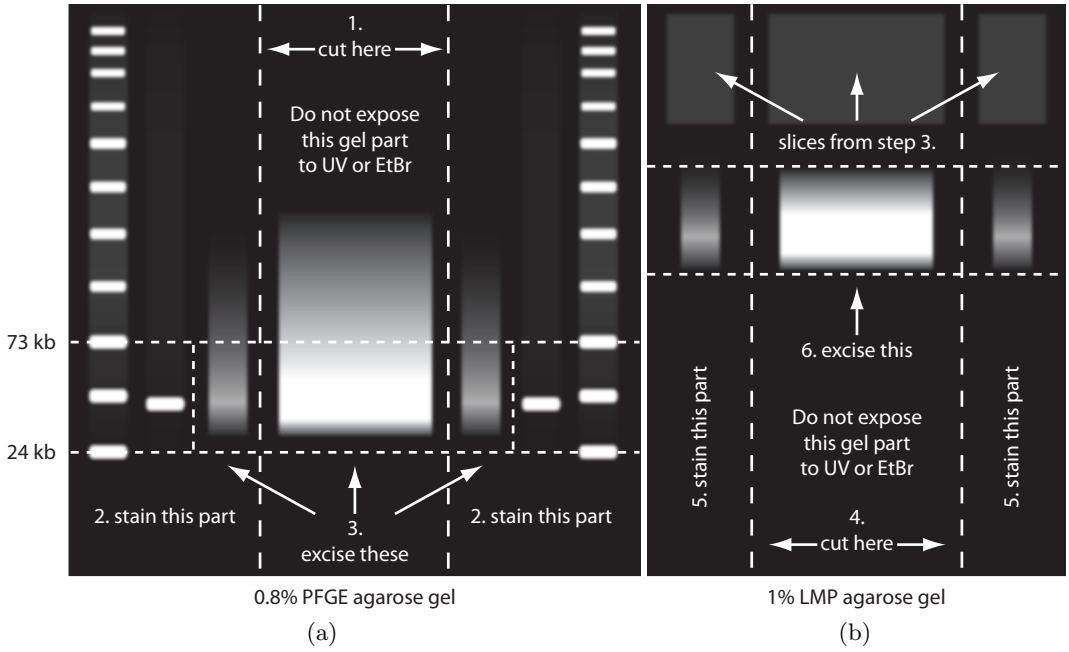


Figure 1.: Running and cutting the PFGE gel and LMP gel.

The sheared DNA is run on the PFGE gel (a), together with markers (see text for details). After electrophoresis, the marker lanes are cut (1) and stained with EtBr (2). The identified range is excised from not stained part of the gel containing sample DNA, together with reference lanes (3) and run on the LMP gel (b). Again, after electrophoresis, the marker lanes are cut (4), stained and visualized (5). The gel slice containing size-selected DNA in the LMP agarose is finally excised (6).

Size-selection of the genomic DNA (optional)

1. Load the end-repair reaction onto the **0,8%** PFGE gel. Run the gel as described previously. Run both markers (Fosmid Control DNA and MidRange II PFG Marker) on both sides of the gel. In addition, include aliquots (**1 µg**) of the end-repaired DNA on both sides of the sample for reference. See figure 1a for reference.

2. Cut off the marker lanes from the gel, and stain them as described previously. Mark the position between **24 kb** and **73 kb** bands of the MidRange II PFG Marker with a razor blade.

3. Reassemble the gel and excise a gel slice containing the sheared DNA between the marked positions. Excise the reference bands containing the sheared DNA as well. **Do not expose sample DNA to the UV light**.

4. Embed the sample DNA slice flanked by reference slices in **1%** SeaPlaque LMP Agarose in **0.5x TBE** buffer. See figure 1b for reference.

5. Run the gel at **5 V/cm** in the coldroom for **1.5–2 hours** to transfer DNA into the LMP agarose.
6. Cut off the reference bands and stain them as described previously. Mark the position of the DNA smear with a razor blade.
7. Reassemble the gel and excise a gel slice containing the sheared DNA between the marked positions. **Do not expose sample DNA to the UV light**.
8. Weight the sample DNA slice in a tared tube.
9. Warm the GELase 50x Buffer to **45°C**. Melt the LMP agarose by incubating the tube at **70°C** for **10–15 minutes**. Quickly transfer the tube to **45°C**.
10. Add the appropriate volume of warmed GELase 50x Buffer to **1x final concentration**. Carefully add **2U** (2 µl) of GELase Enzyme Preparation to the tube for each **100 µl** of melted agarose. Keep the melted agarose solution at **45°C** and gently mix the solution. Incubate the solution at **45°C overnight**.
11. Transfer the reaction to **70°C** for **10 minutes** to inactivate the GELase enzyme.
12. Remove **500 µl aliquots** of the solution into sterile 1.5 ml microfuge tube(s).
13. Chill the tubes on ice for **5 minutes**. Centrifuge the tubes in a microcentrifuge at ≥**20,000g** for **20 minutes** to pellet any insoluble oligosaccharides. Carefully remove the upper 90%-95% of the supernatant, which contains the DNA, to a sterile 1.5-ml tube. Be careful to avoid the gelatinous pellet.

Final purification of the genomic DNA

1. Add **0.1 volume** of 3M KAc (pH 7.0) to the end-repaired DNA or the DNA that you have purified from the LMP agarose gel. Mix gently by inverting the tube.
2. Add **2.5 volumes** of ethanol. Mix gently by inverting the tube.
3. Incubate sample at room temperature for **10 minutes** and centrifuge at ≥**20,000g** for **15 min**. Remove the supernatant.
4. Wash DNA pellet with **1 ml** room-temperature 70% ethanol, and centrifuge at ≥**20,000g** for **5 min**. Carefully decant the supernatant without disturbing the pellet.
5. Wash DNA pellet again with **1 ml** room-temperature 70% ethanol, and centrifuge at ≥**20,000g** for **5 min**. Carefully decant the supernatant without disturbing the pellet. Use a pipet to completely remove the remaining ethanol.

6. Air-dry the pellet for **5–10 min**, and redissolve the DNA in **10 µl** of warm (**55°C**) nuclease-free water.
7. Dialyse the DNA solution against water on the Millipore 0.025 µm VSWP membrane for **1h**.
8. Use **1 µl** of the solution to determine the DNA concentration by running it on a gel and using **100 ng** of the Fosmid Control DNA as a reference. Store the prepared DNA at **-20°C** or use it directly for ligation (recommended).

Preparation of the fosmid vector

1. Purify the pFlyFos vector as described in the MaxiPrep version of the protocol 5.
2. Set up a **100 µl** restriction digest of the pFlyFos DNA. Use **10 µl** NEBuffer 1; **1 µl** BSA; **30 µg** pFlyFos DNA; **5 µl** *Pml*I; water to **100 µl**. Incubate at **37°C overnight**.
3. Run all of the digested vector on the **0.8%** agarose gel. Include undigested vector (**500 ng**) and an aliquot of digested vector (**500 ng**) as a reference.
4. Cut out the agarose slice containing digested DNA (the linear vector migrates slower than superhelical reference plasmid). **Avoid UV exposure**. Use undigested and digested vector reference samples to determine where agarose should be cut. As an alternative, crystal violet in-gel staining (1 µg/ml of the gel) can be used.
5. Weight the agarose slice and isolate DNA using QIAquick Gel Extraction Kit. Use two columns (each per 50 µg of restriction digest). Elute vector DNA from each column with **50 µl** water. Combine the eluates.
6. Add **12 µl** antarctic phosphatase buffer and **5 µl** antarctic phosphatase to the eluate. Adjust the volume to **120 µl** with water and incubate at **37°C** for **3 hours**. Heat inactivate enzyme at **65°C** for **15 minutes**.
7. Precipitate DNA by adding **6 µl** 5M LiCl and **90 µl** isopropanol. Mix by vortexing and centrifuge at ≥**20,000g** for **15 min** at **4°C**. Remove the supernatant.
8. Wash DNA pellet with **1 ml** room-temperature 70% ethanol, and centrifuge at ≥**20,000g** for **5 min**. Carefully decant the supernatant without disturbing the pellet.
9. Wash DNA pellet again with **1 ml** room-temperature 70% ethanol, and centrifuge at ≥**20,000g** for **5 min**. Carefully decant the supernatant without disturbing the pellet. Use a pipet to completely remove the remaining ethanol.

10. Air-dry the pellet for **5–10 min**, and redissolve the DNA in **10 µl** of warm (**55°C**) nuclease-free water.
11. Dialyse the DNA solution against water on the Millipore 0.025 µm VSWP membrane for **1h**.
12. Use **1 µl** of the solution to measure the DNA concentration and adjust it to **500 ng/µl** with nuclease-free water. Store the prepared vector at **-20°C** or use it directly for ligation (recommended).

Ligation, phage packaging and infection

1. Inoculate **50 ml** of LB with a single colony of EPI300-T1^{R} cells. Culture **overnight** at **37°C**. Store culture at **4°C** for up to 48 hours.
2. Set up a **10 µl** ligation reaction with **500 ng** of the cut pFlyFos, **0.25–5 µg** sheared genomic DNA, **1 µl** 10x Ligase Buffer and **1 µl** T4 DNA Ligase. The optimal amount of genomic DNA can differ depending on DNA quality. For our ligations, it was 2 µg. Incubate ligation reaction **overnight** at **16°C**.
3. Inoculate **50 ml** of LB + $MgSO_4^{10mM}$ + Maltose$^{0.2\%}$ with **0.5 ml** of the EPI300-T1^{R} overnight culture. Culture cells at **37°C** with vigorous shaking until OD(600) reaches **0.8–1.0**.
4. Thaw on ice one tube of the MaxPlax Lambda Packaging Extract. When thawed, immediately transfer **25 µl** of the packaging extract to a new tube. Keep the tube on ice. Return the remaining **25 µl** of the packaging extract to a **-80°C** freezer. Avoid exposing MaxPlax Lambda Packaging Extracts to any source of CO_2.
5. Add **10 µl** of the ligation reaction to **25 µl** of the packaging extract. Mix by pipetting, avoid introduction of the air bubbles. Incubate at **30°C** for **2 hours**.
6. Add the remaining **25 µl** of the packaging extract to the reaction tube. Incubate at **30°C** for **2 hours**.
7. Add **950 µl** of the Phage Dilution Buffer (PDB) to the packaging reaction. Mix gently by inverting the tube.
8. Add **25 µl** of chloroform to precipitate unassembled phage proteins. Mix gently by inverting the tube.
9. Prepare **1:10**, **1:100** and **1:1000** serial dilutions of the phage particles in Phage Dilution Buffer.

10. Use **10 µl** of each dilution and the undiluted phage individually to infect **100 µl** of the EPI300-T1^{R} cells. Incubate each tube for **1 hour** at **37°C.** Store remaining phage dilutions and undiluted phage at **4°C** for up to 48h.
11. Plate cells on LA + Cm^{15}. Incubate plates **overnight** at **37°C**. Sometimes longer incubation times (up to 36h) are necessary to obtain large colonies.
12. Count colonies on the plates and determine the phage titer using the following formula:

$$\frac{(\#\,of\,colonies)\cdot(dilution\,factor)\cdot(1000\,\mu g/ml)}{(volume\,of\,phage\,extract\,[\mu l])} = x\,[cfu/ml]$$

13. Inoculate **50 ml** of LB + $MgSO_4{}^{10mM}$ + Maltose$^{0.2\%}$ with **0.5 ml** of the EPI300-T1^{R} overnight culture. Culture cells at **37°C** with vigorous shaking until OD(600) reaches **0.8**–**1.0**.
14. Dilute phages accordingly to obtain **100 colonies** from **100 µl** of cells infected with **10 µl** of phage particles. Infect EPI300-T1^{R} cells for **one hour** at **37°C**.
15. Plate the library on LA + Cm^{15}. During plating, keep the infected cells **on ice** to prevent formation of duplicate clones. Incubate plates overnight at **37°C**. Sometimes longer incubation times (up to 36h) are necessary to obtain large colonies.

Protocol 8.

Liquid culture recombineering

This protocol describes tagging of genes in FlyFos clones with a variety of markers by Red/ET recombination. It is based on the recombineering pipeline for generation of tagged transgenes from *C. elegans* genomic fosmid clones by M. Sarov.

Reagents needed:

- Phusion® High-Fidelity DNA Polymerase (NEB)
- 5x Phusion™ HF Buffer
- 10mM dNTPs
- 5M LiCl
- 100% Isopropanol
- 70% Ethanol
- LA plates + Chloramphenicol (15 μg/ml)
- 10% Glycerol
- 25% L-Rhamnose
- SOC medium
- LB + Chloramphenicol (25 μg/ml)
- LB + Chloramphenicol (25 μg/ml) + Hygromycin (50 μg/ml)
- LB + Chloramphenicol (25 μg/ml) + Hygromycin (50 μg/ml) + Kanamycin (25 μg/ml)
- LB + Chloramphenicol (25 μg/ml) + Hygromycin (50 μg/ml) + Anhydrotetracycline (200 nM)

Protocol:

Amplification of the tagging cassette

1. Design recombineering primers. Primers include priming part (25 bp) that is complementary to the ends of the tag sequence (forward and reverse) on 5' end and 50 bp homology arms complementary to the target sequence. Verify the orientation of both primers.
2. Set up a **100 μl** PCR reaction to amplify the tagging cassette. Use **50–100 ng** of the tagging vector as a template. Use HPLC-purified recombineering primers at **10 nmol/μl** final concentration. Run the PCR reaction for **20–25** cycles.
3. Verify the PCR by running **5 μl** of the reaction on an agarose gel.
4. Precipitate DNA by adding **5 μl** 5M LiCl and **70 μl** isopropanol. Mix by vortexing and centrifuge at ≥**20,000g** for **15 min** at **4°C**. Remove the supernatant.

5. Wash DNA pellet with **1 ml** room-temperature 70% ethanol, and centrifuge at ≥**20,000g** for **5 min**. Carefully decant the supernatant without disturbing the pellet.
6. Wash DNA pellet again with **1 ml** room-temperature 70% ethanol, and centrifuge at ≥**20,000g** for **5 min**. Carefully decant the supernatant without disturbing the pellet. Use a pipet to completely remove the remaining ethanol.
7. Air-dry the pellet for **5–10 min**, and redissolve the DNA in **10 µl** of nuclease-free water.
8. Store the amplified tagging cassette at **-20°C**.

Transformation of pRedFlp4 recombineering helper

1. Plate the fosmid clone on LA+Cm^{15}. Incubate plates **overnight** at **37°C**.
2. Use a single colony to inoculate **1 ml** of LB+Cm^{25}. Culture **overnight** at **37°C** with vigorous shaking.
3. Use **20 µl** of the overnight culture to inoculate **1 ml** of LB+Cm^{25}. Culture cells for **2h** at **37°C**.
4. Centrifuge cells at **10,000g** for **30 sec** at **2°C**. Remove supernatant entirely.
5. Resuspend cells in **1 ml** of **ice cold** water.
6. Centrifuge cells at **10,000g** for **30 sec** at **2°C**. Remove supernatant entirely.
7. Resuspend cells in **500 ml** of **ice cold** water.
8. Centrifuge cells at **10,000g** for **30 sec** at **4°C**. Remove supernatant entirely.
9. Resuspend cells in **50 µl** of **ice cold** 10% glycerol.
10. Add **1 µl** of pRedFlp4 (**100 ng/µl**) to the cells. Mix briefly by pipetting. Transfer the cell suspension into a chilled **2 mm** electroporation cuvette.
11. Electroporate at **3000V.**
12. **Immediately** add **1 ml** of SOC and transfer bacteria into 2 ml tube.
13. Culture cells for **1h** at **30°C**.
14. Inoculate **1 ml** of LB+Cm^{25}+Hgr^{50} with **100 µl** of transformed bacteria. Culture **overnight** at **30°C** with vigorous shaking.

Tagging by Red/ET recombination

1. Inoculate **1 ml** of LB+Cm^{25}+Hgr^{50} with **30 µl** of overnight culture. Incubate for **2 hours** at **30°C** with vigorous shaking.

2. Induce Red operon expression by supplementing medium with **10 µl** of 25% L-rhamnose. Incubate for **2 hours** at **30°C** with vigorous shaking.
3. Centrifuge cells at **10,000g** for **30 sec** at **2°C**. Remove supernatant entirely.
4. Resuspend cells in **1 ml** of **ice cold** water.
5. Centrifuge cells at **10,000g** for **30 sec** at **2°C**. Remove supernatant entirely.
6. Resuspend cells in **500 ml** of **ice cold** water.
7. Centrifuge cells at **10,000g** for **30 sec** at **4°C**. Remove supernatant entirely.
8. Resuspend cells in **50 µl** of **ice cold** 10% glycerol.
9. Add **1 µl** of the tagging cassette (**500 ng/µl**) to the cells. Mix briefly by pipetting. Transfer the cell suspension into a chilled **2 mm** electroporation cuvette.
10. Electroporate at **3000V**.
11. **Immediately** add **1 ml** of SOC and transfer bacteria into 2 ml tube.
12. Culture cells for **1h** at **30°C**.
13. Inoculate **1 ml** of LB+Cm^{25}+Hgr^{50}+Kan^{25} with **100 µl** of transformed bacteria. Incubate at **30°C** with vigorous shaking **until the culture is saturated** (30 hours).

Removal of the selectable marker and pRedFlp helper

1. Inoculate **1 ml** of LB+Cm^{25}+Hgr^{50}+AHT^{200} with **10 µl** of the saturated culture. Incubate **overnight** at **30°C** with vigorous shaking.
2. Inoculate **1 ml** of LB+Cm^{25} with **10 µl** of the saturated culture. Incubate **overnight** at **37°C** with vigorous shaking.

Protocol 9.

High-throughput liquid culture recombineering

This is a high-throughput and improved version of protocol 8. The whole process is performed in 96-well format.

Reagents needed:

- Phusion® High-Fidelity DNA Polymerase (NEB)
- 5x Phusion™ HF Buffer
- 10mM dNTPs
- 96-well PCR purification kit
- 10% Glycerol
- 25% L-Rhamnose
- SOC medium
- YENB + Chloramphenicol (25 μg/ml)
- YENB + Chloramphenicol (25 μg/ml) + Hygromycin (50 μg/ml)
- YENB + Chloramphenicol (25 μg/ml) + Hygromycin (50 μg/ml) + Kanamycin (25 μg/ml)
- YENB + Chloramphenicol (25 μg/ml) + Hygromycin (50 μg/ml) + Anhydrotetracycline (200 nM)

Protocol:

Amplification of the tagging cassette

1. Design recombineering primers for each sample. Primers include priming part (25 bp) that is complementary to the ends of the tag sequence (forward and reverse) on 5' end and 50 bp homology arms complementary to the target sequence. Verify the orientation of both primers.
2. Set up **50 μl** 96-well PCR reactions to amplify the tagging cassettes. Use **25–50 ng** of the tagging vector as a template. Use HPLC-purified recombineering primers at **10 nmol/μl** final concentration. Run the PCR reaction for **20–25** cycles.
3. Verify the PCR by running **5 μl** of the reaction on an agarose gel.
4. Purify the DNA with 96-well PCR purification kit following the manufacturer's instructions. Elute DNA with **500 μl** of nuclease-free water.
5. Store the amplified tagging cassettes at **-20°C**.

Transformation of pRedFlp4 recombineering helper

1. Use a glycerol stocks to inoculate **1 ml** of YENB+Cm^{25} in a 96-well deep well plate. Seal the plate with an air-permable seal and culture **overnight** at **37°C** with vigorous shaking.
2. Use **40 µl** of the overnight cultures to inoculate **1 ml** of YENB+Cm^{25} per well. Seal the plate with an air-permable seal and culture cells for **2h** at **37°C** with vigorous shaking.
3. Centrifuge the plate at **5,000g** for **10 min** at **2°C**. Discard supernatant by inverting the plate over the sink and placing it on a stack of paper towels.
4. Add **1 ml** of **ice cold** 10% glycerol into each well. Seal the plate with an aluminium or plastic seal.
5. Resuspend bacteria by shaking the plate at **1400 rpm** for **1 min** at **2°C**.
6. Centrifuge the plate at **5,000g** for **10 min** at **2°C**. Discard supernatant by inverting the plate over the sink and placing it on a stack of paper towels.
7. Add **100 µl** of pRedFlp4 (**0.1 ng/µl** in **ice-cold** water) into each well. Resuspend cells by pipetting.
8. Transfer the cell suspension into a chilled 96-well electroporation cuvette and electroporate at **2500V**.
9. **Immediately** transfer the cell suspension into a new plate with **1 ml** of SOC per well.
10. Seal the plate with an air-permable seal and culture for **1h** at **30°C** with vigorous shaking.
11. Use **100 µl** of the transformed bacteria to inoculate **1 ml** of YENB+Cm^{25}+Hgr^{50} per well. Seal the plate with an air-permable seal and culture **overnight** at **30°C** with vigorous shaking.

Tagging by Red/ET recombination

1. Use **40 µl** of the overnight cultures to inoculate **1 ml** of YENB+Cm^{25}+Hgr^{50} per well.
2. Seal the plate with an air-permable seal and culture cells for **2h** at **30°C** with vigorous shaking.
3. Induce Red operon expression by adding **20 µl** of 25% L-rhamnose into each well.

4. Seal the plate with an air-permable seal and incubate plate for **2 hours** at **30°C** with vigorous shaking.
5. Centrifuge the plate at **5,000g** for **10 min** at **2°C**. Discard supernatant by inverting the plate over the sink and placing it on a stack of paper towels.
6. Add **1 ml** of **ice cold** 10% glycerol into each well. Seal the plate with an aluminium or plastic seal.
7. Resuspend bacteria by vigorously shaking the plate for **1 min** at **2°C**.
8. Centrifuge the plate at **5,000g** for **10 min** at **2°C**. Discard supernatant by inverting the plate over the sink and placing it on a stack of paper towels.
9. Add **100 µl** of the tagging cassette (**5 ng/µl** in **ice-cold** water) into each well. Resuspend cells by pipetting.
10. Transfer the cell suspension into a chilled 96-well electroporation cuvette and electroporate at **2500V**.
11. **Immediately** transfer the cell suspension into a new plate with **1 ml** of SOC per well.
12. Seal the plate with an air-permable seal and culture for **1h** at **30°C** with vigorous shaking.
13. Use **100 µl** of the transformed bacteria to inoculate **1 ml** of YENB+Cm^{25}+Hgr^{50} +Kan^{25} per well. Seal the plate with an air-permable seal and culture **overnight** at **30°C** with vigorous shaking.

Removal of the selectable marker and pRedFlp helper

1. Use **10 µl** of the overnight cultures to inoculate **1 ml** of YENB+Cm^{25}+Hgr^{50} +AHT^{200} per well.
2. Seal the plate with an air-permable seal and culture for **2 hours** at **30°C** with vigorous shaking.
3. Use **100 µl** of the cultures to inoculate **1 ml** of YENB+Cm^{25} per well.
4. Seal the plate with an air-permable seal and culture **overnight** at **37°C** with vigorous shaking.

Appendix C.

Plasmid maps of recombineering vectors

Contents

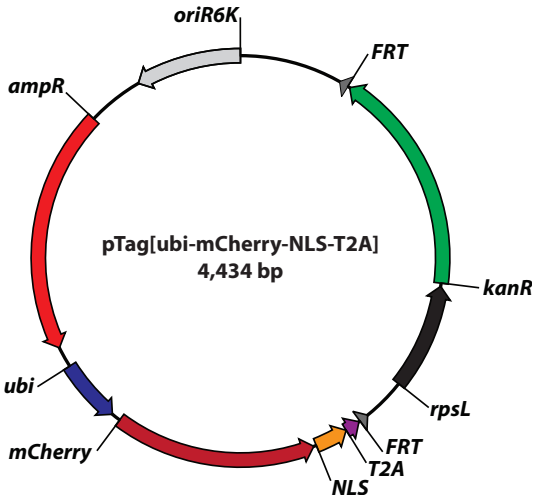


Figure C.1.: pTag[ubi-mCherry-NLS-T2A] N-terminal tagging vector map

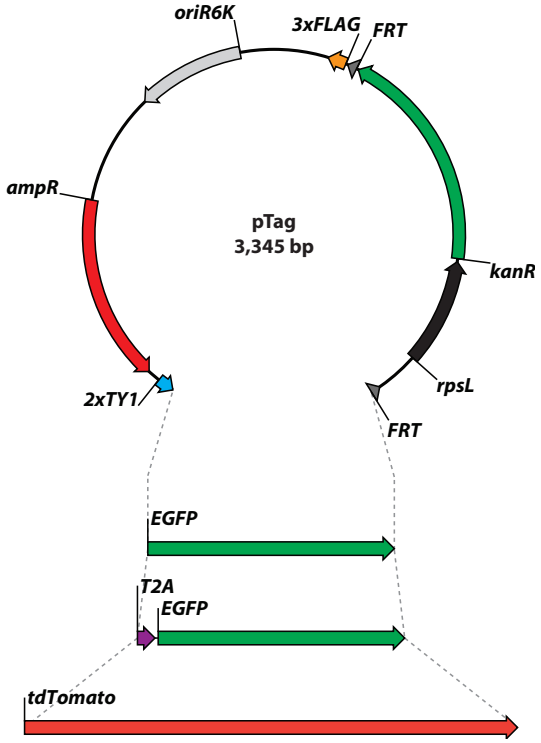


Figure C.2.: Maps of EGFP, T2A-EGFP and tdTomato tagging constructs

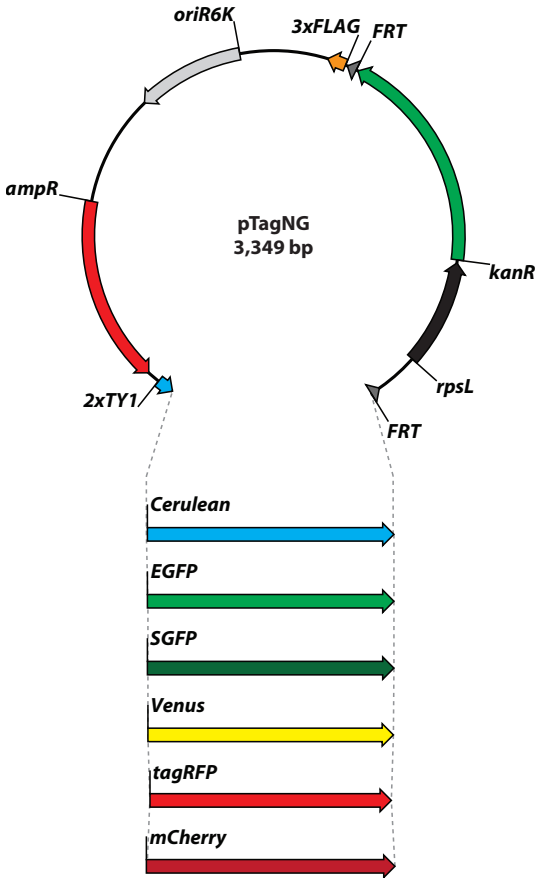


Figure C.3.: Maps of pTagNG fluorescent protein tags

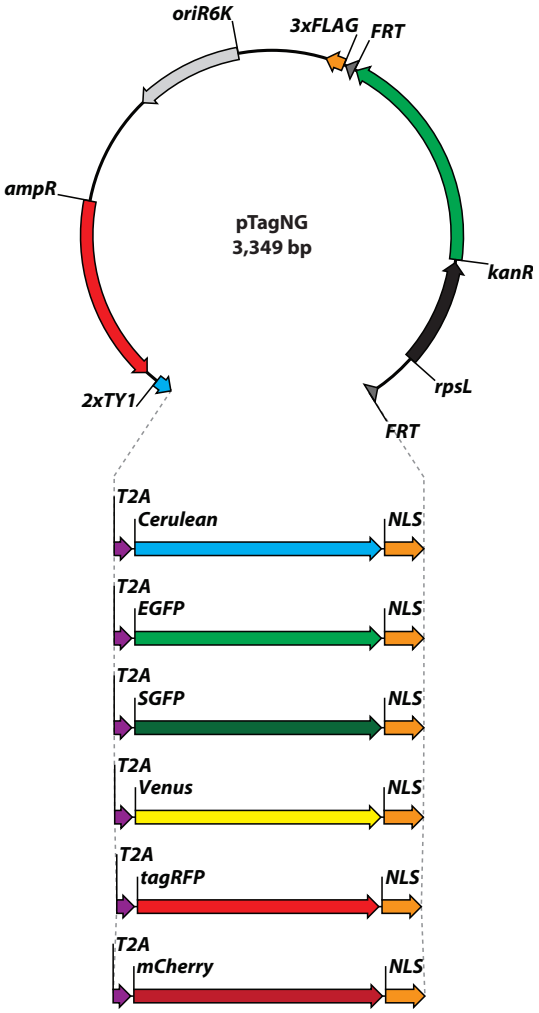


Figure C.4.: Maps of pTagNG T2A-NLS fluorescent protein tags

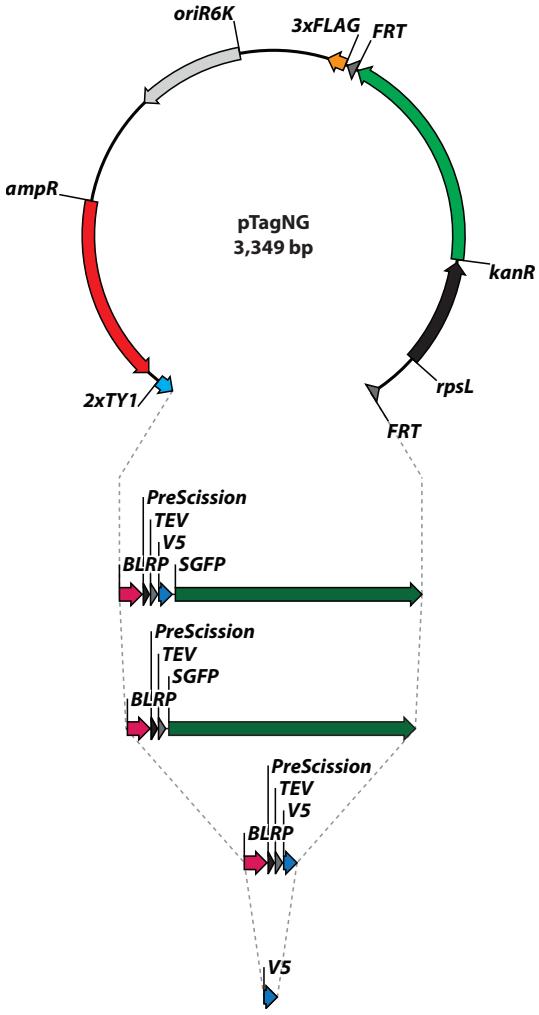


Figure C.5.: Maps of pTagNG N-terminal biotin tags and V5 tag

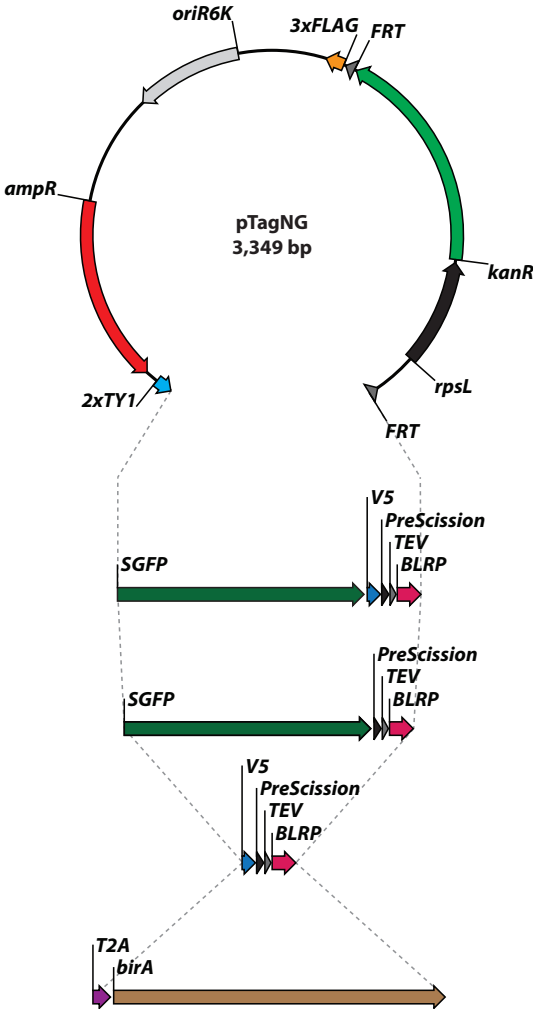


Figure C.6.: Maps of pTagNG C-terminal biotin tags and T2A-birA tag

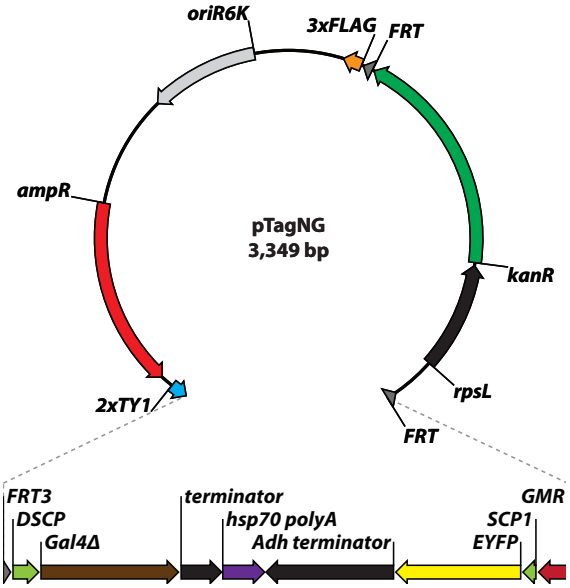


Figure C.7.: pTagNG[Ultimate] RMCE acceptor tag map

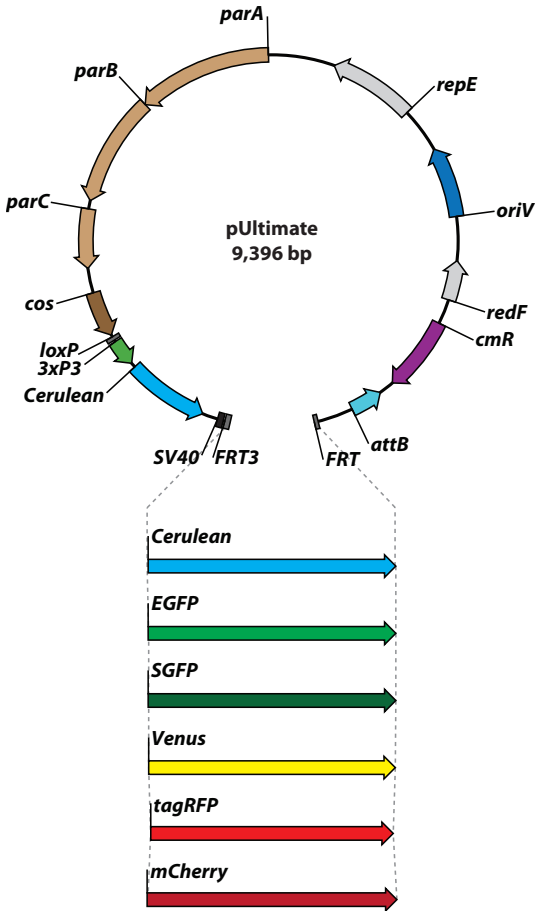


Figure C.8.: Maps of pUltimate fluorescent protein RMCE donors

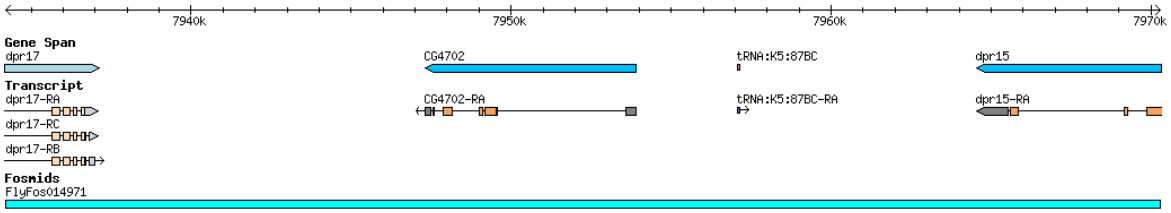
D Appendix D.

FlyBase maps of used fosmid clones

Contents

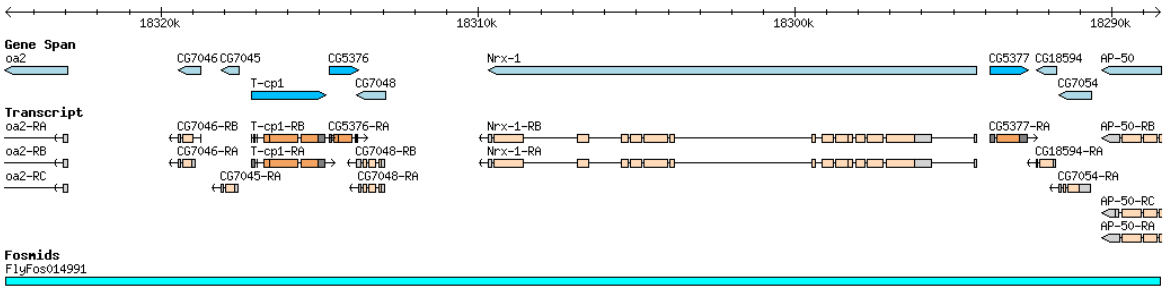
D.1. *Drosophila melanogaster* fosmids

FlyFos014971



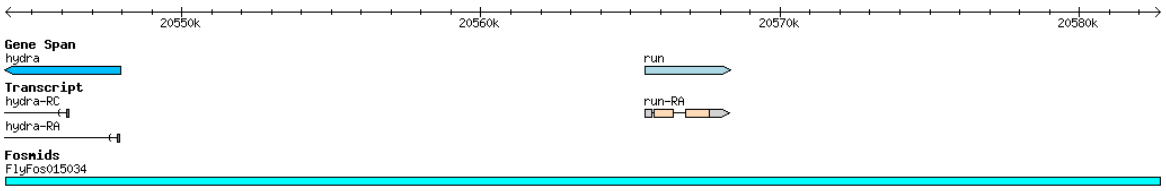
3R [7934206..7970287] (+)

FlyFos014991



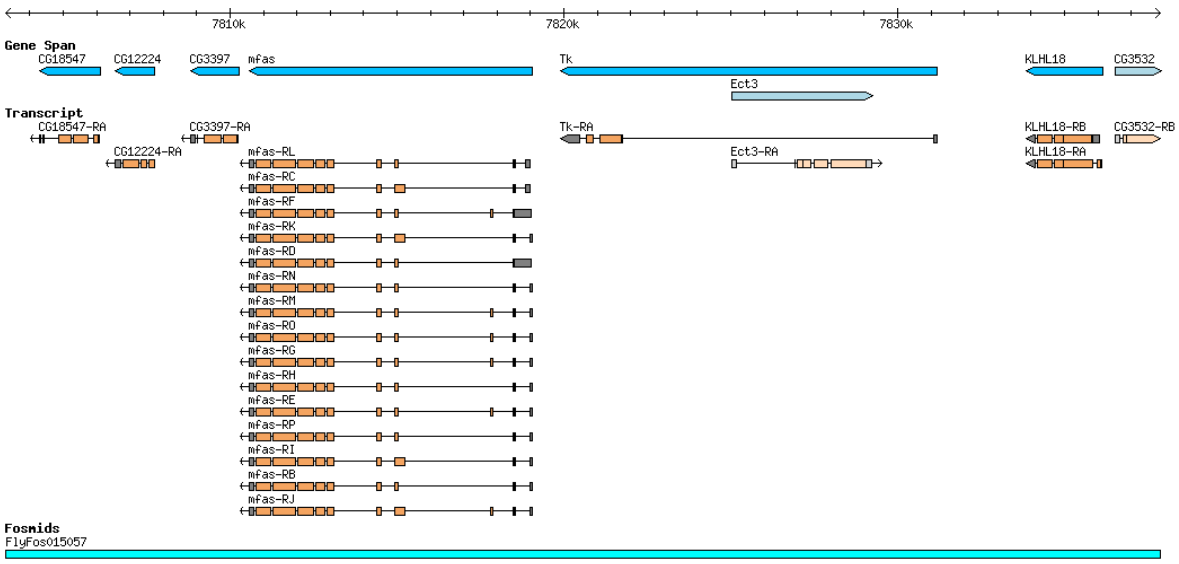
3R [18288474..18324899] (-)

FlyFos015034



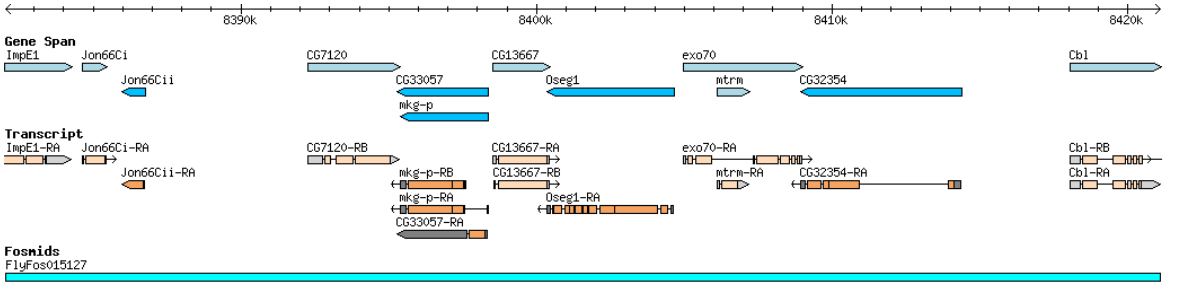
X [20544115..20582706] (+)

FlyFos015057



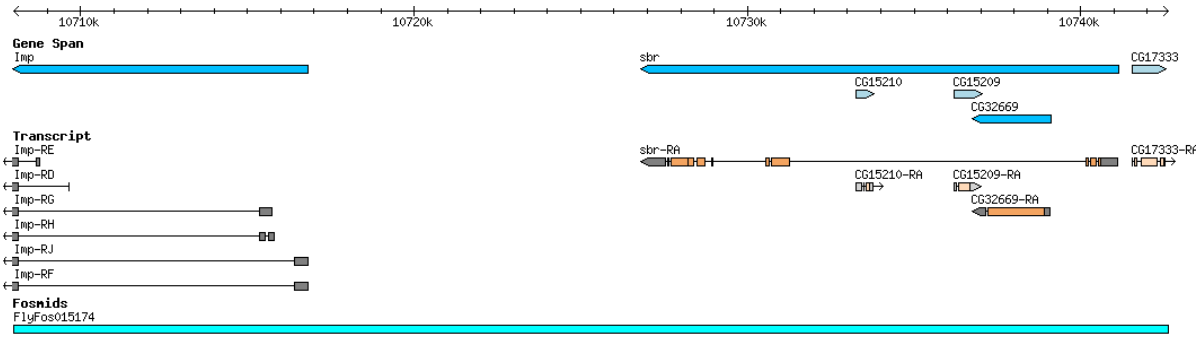
3R [7803280..7837867] (+)

FlyFos015127



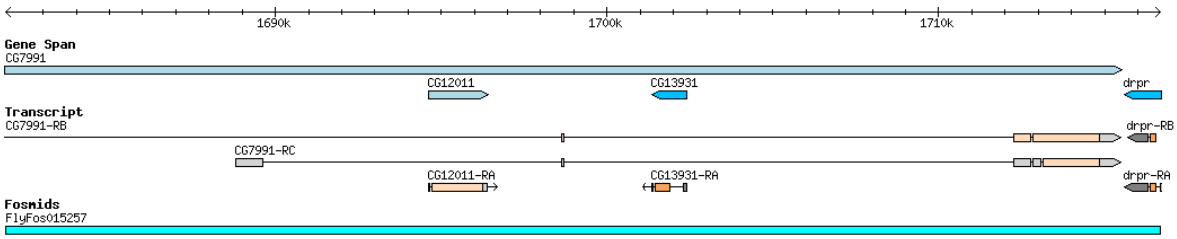
3L [8382038..8421105] (+)

FlyFos015174



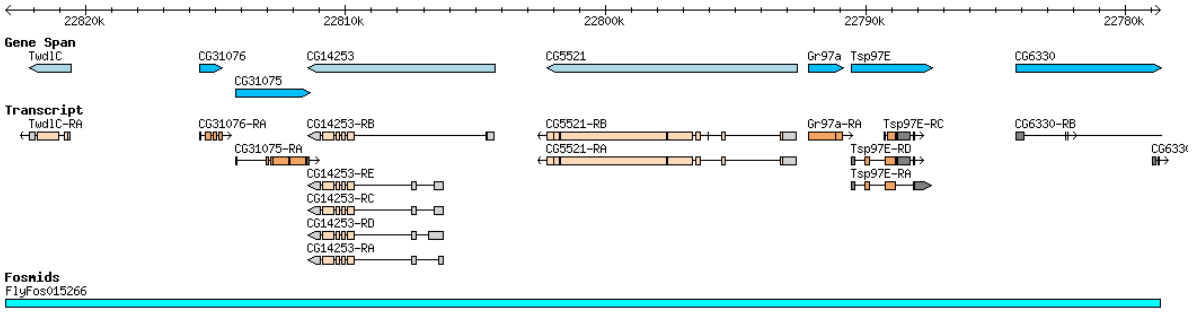
X [10707988..10742625] (+)

FlyFos015257



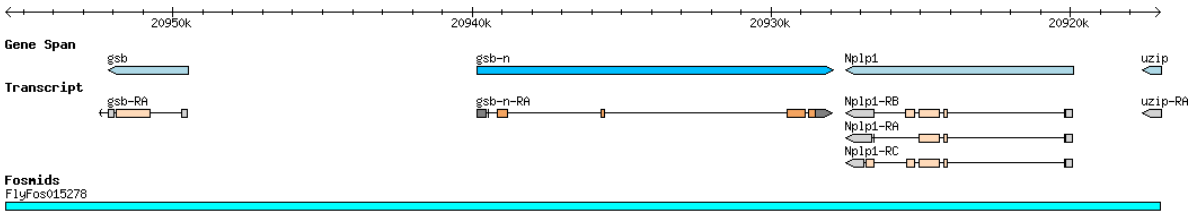
3L [1681872..1716680] (+)

FlyFos015266



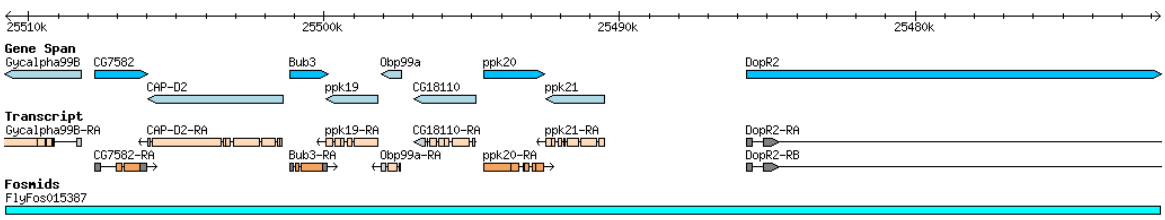
3R [22778671..22823063] (-)

FlyFos015278



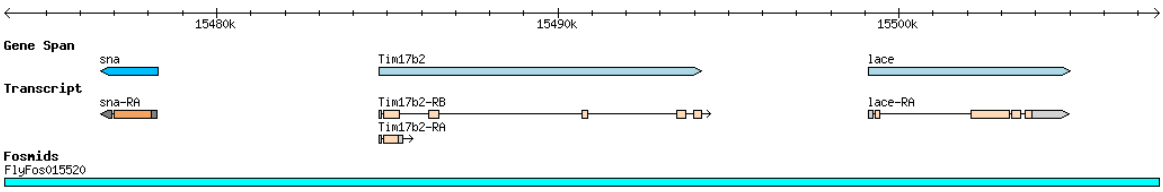
2R [20916988..20955595] (-)

FlyFos015387



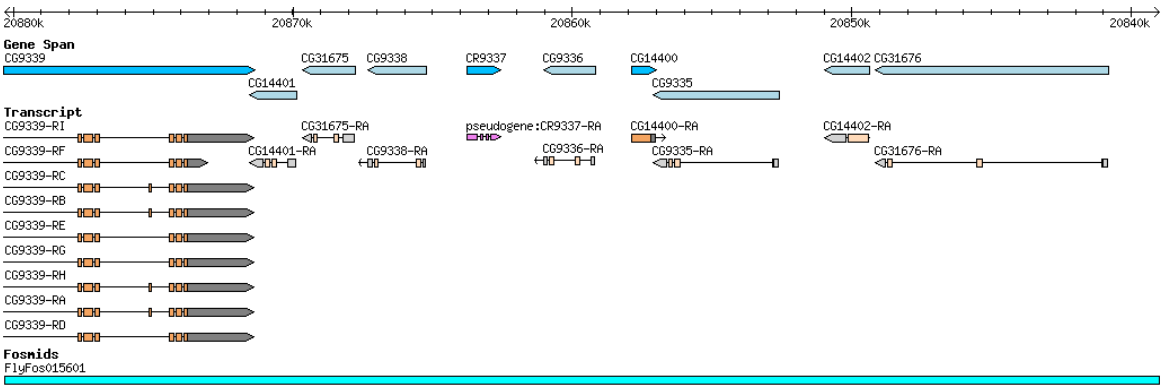
3R [25471702..25510775] (-)

FlyFos015520



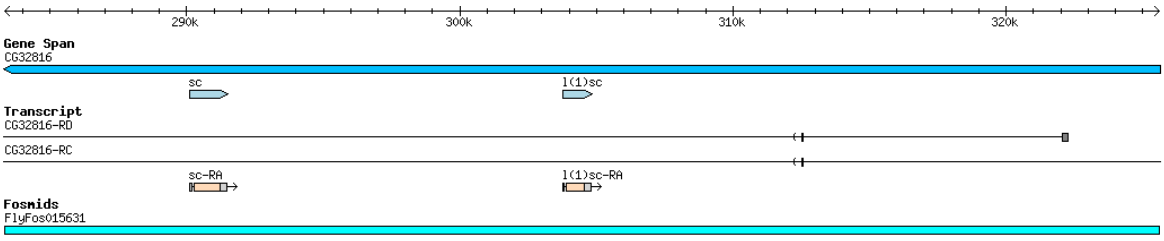
2L [15473788..15507625] (+)

FlyFos015601



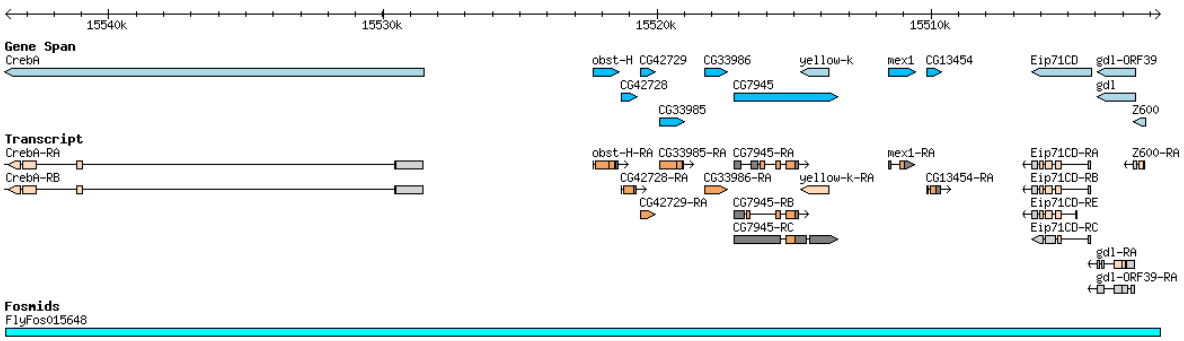
2L [20839002..20880305] (-)

FlyFos015631



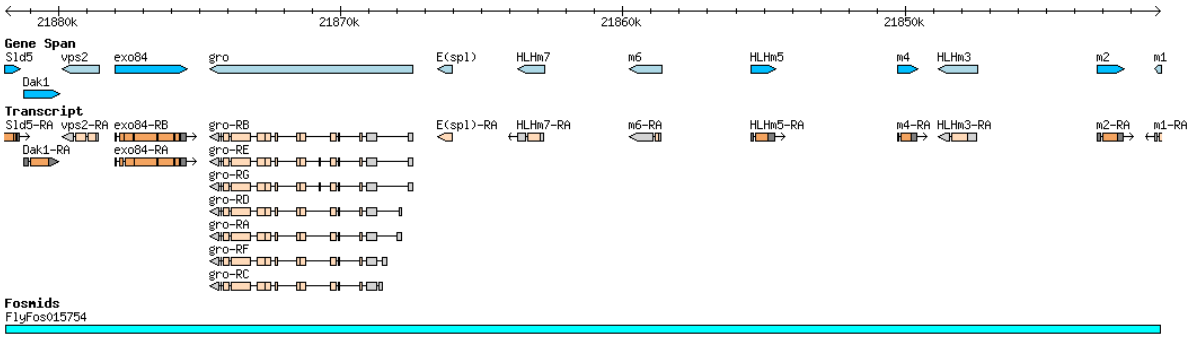
X [283337..325611] (+)

FlyFos015648



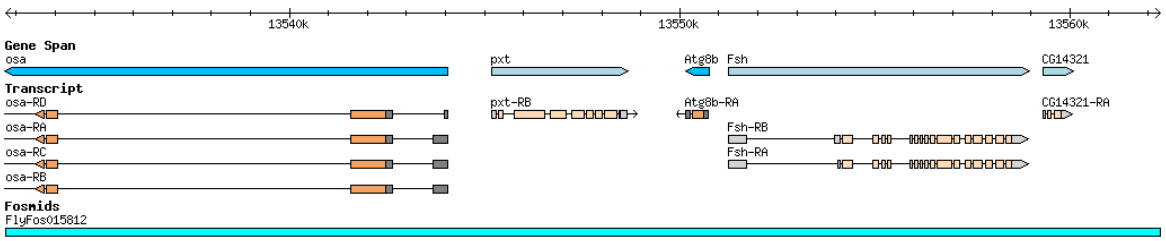
3L [15501633..15543762] (-)

FlyFos015754



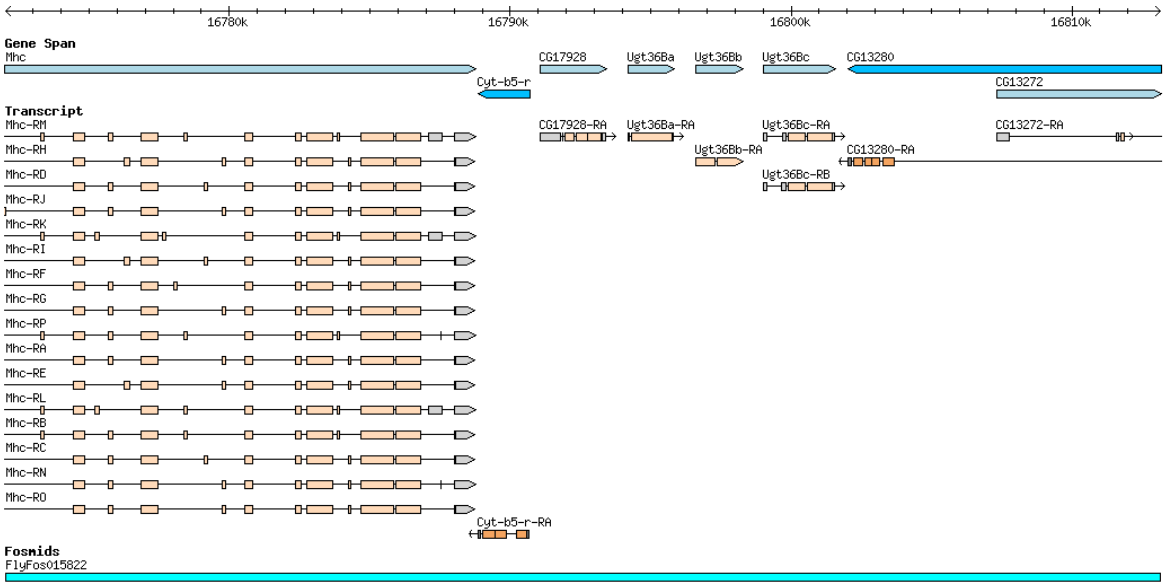
3R [21840961..21881867] (-)

FlyFos015812



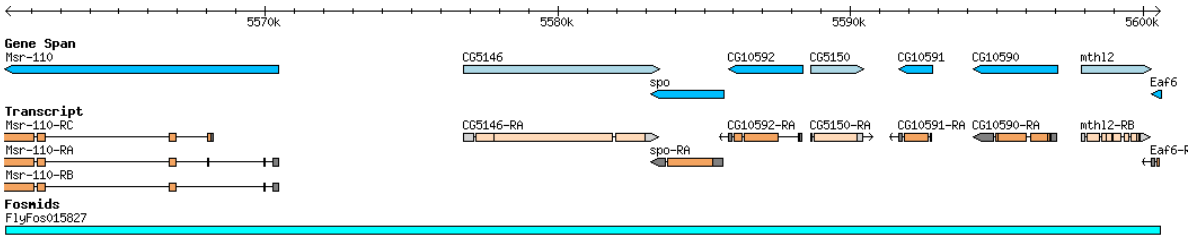
3R [13532729..13562316] (+)

FlyFos015822



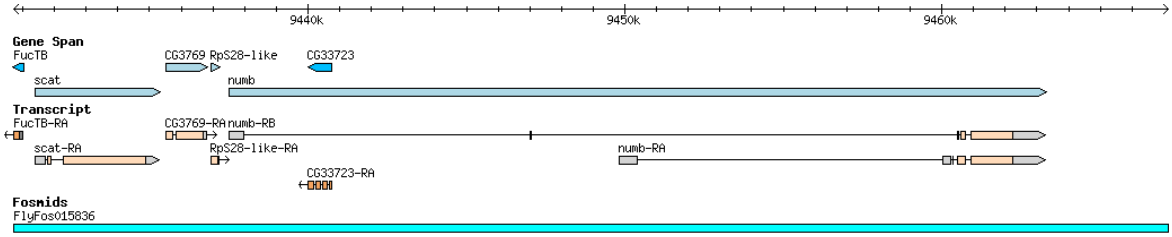
2L [16772035..16813148] (+)

FlyFos015827



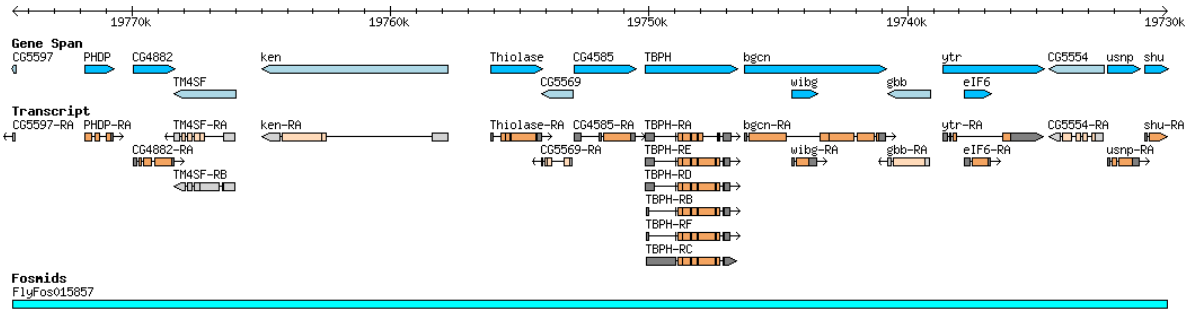
3L [5561128..5600585] (+)

FlyFos015836



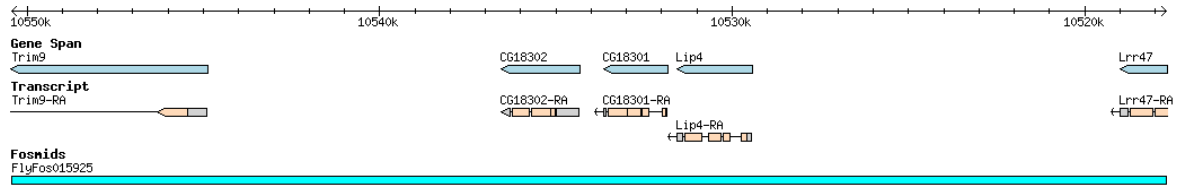
2L [9430713..9467147] (+)

FlyFos015857



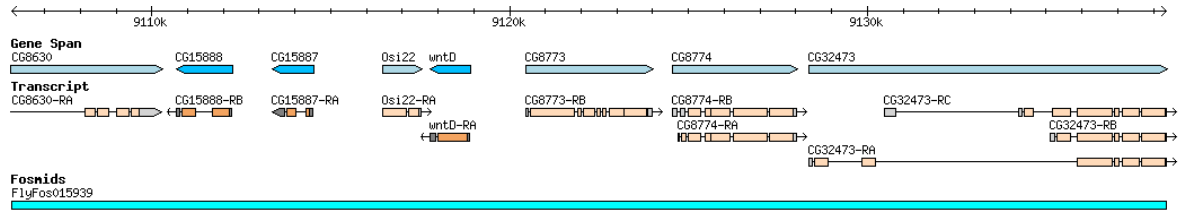
2R [19729951..19774627] (-)

FlyFos015925



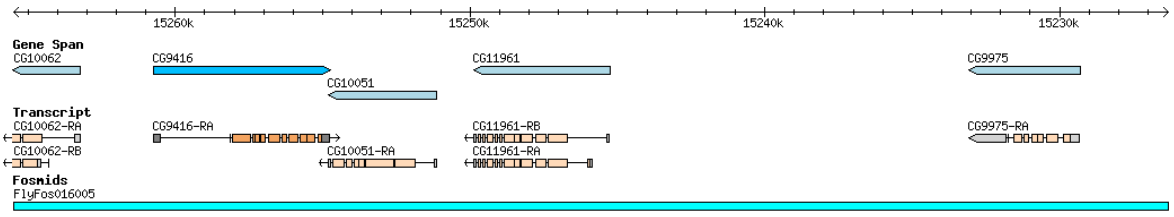
2L [10517686..10550437] (-)

FlyFos015939



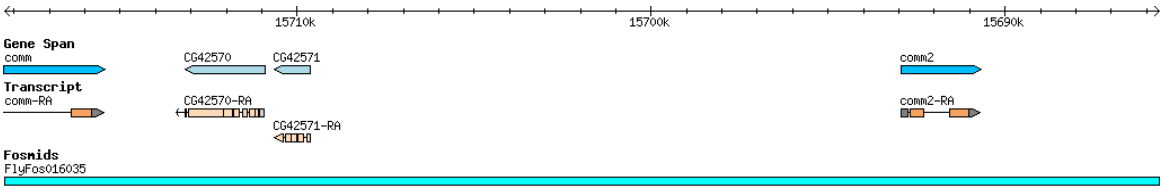
3R [9106084..9138358] (+)

FlyFos016005



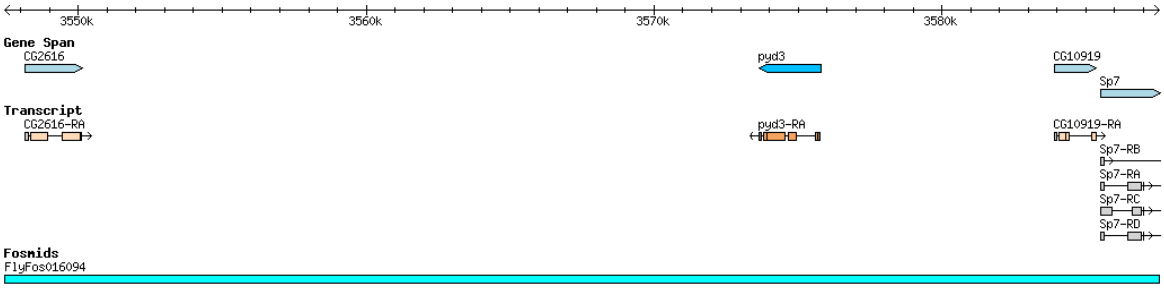
2R [15226312..15265481] (-)

FlyFos016035



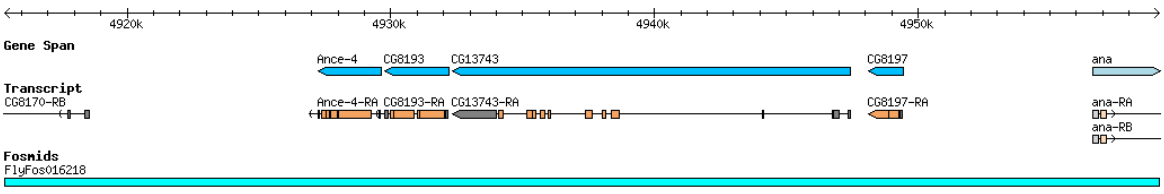
3L [15685645..15718262] (-)

FlyFos016094



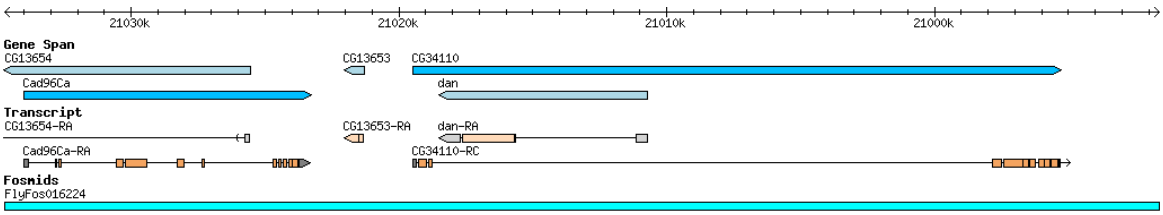
3R [3547444..3587548] (+)

FlyFos016218



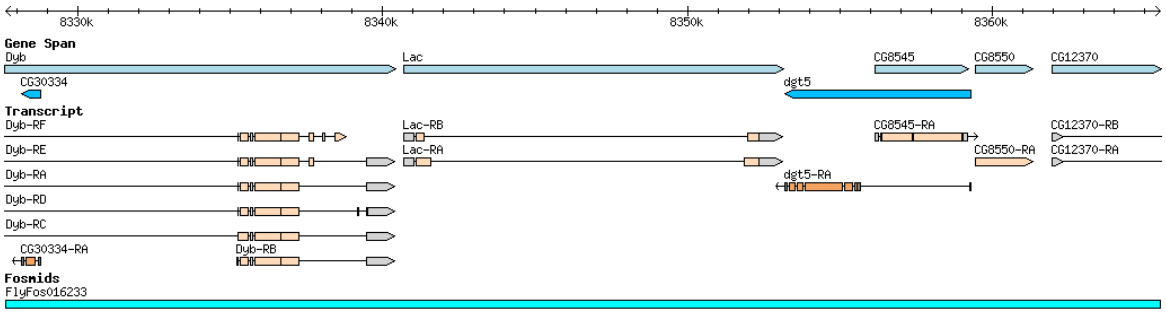
2R [4915358..4959144] (+)

FlyFos016224



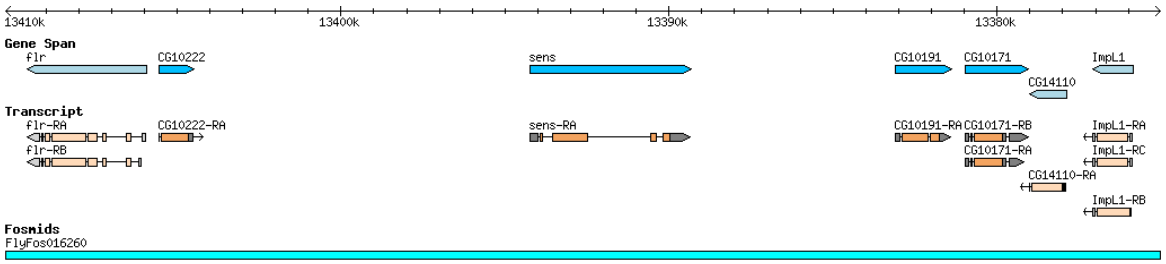
3R [20991603..21034727] (-)

FlyFos016233



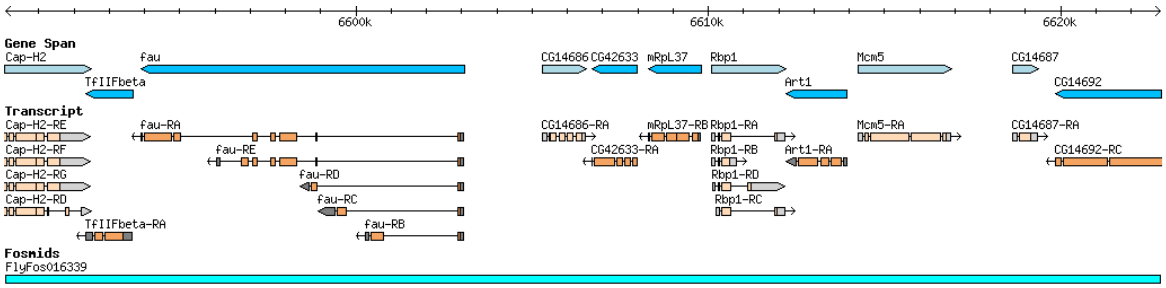
2R [8327639..8365500] (+)

FlyFos016260



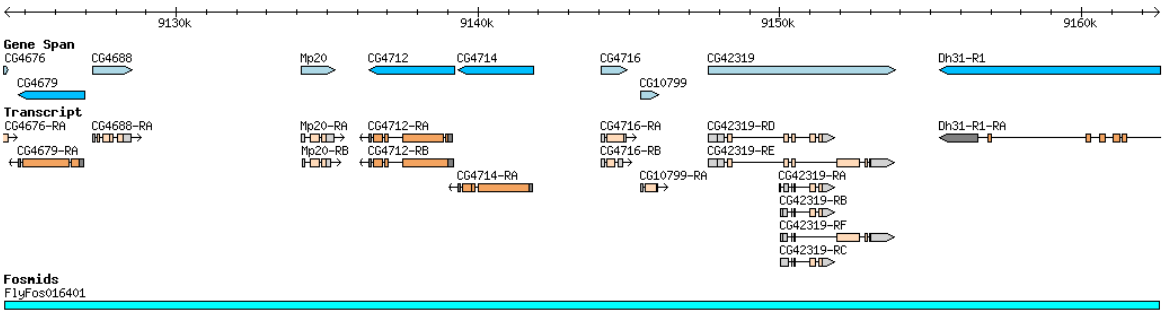
3L [13375025..13410221] (-)

FlyFos016339



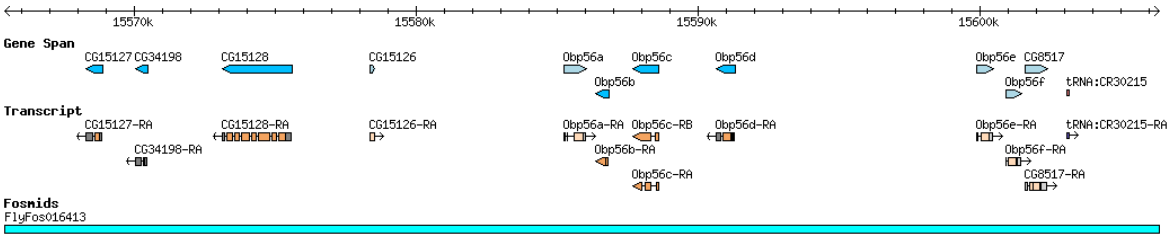
3R [6590063..6622825] (+)

FlyFos016401



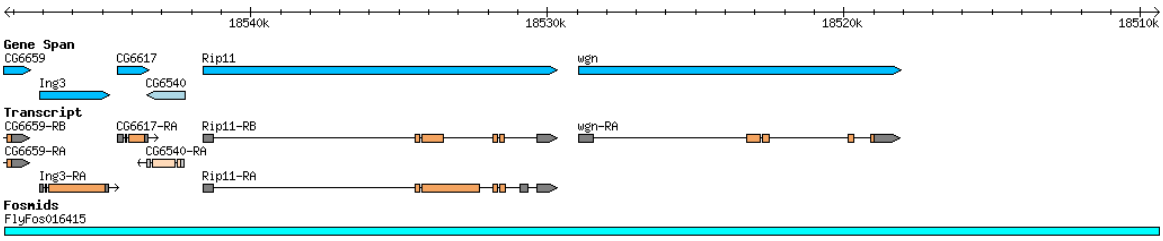
2R [9124314..9162577] (+)

FlyFos016413



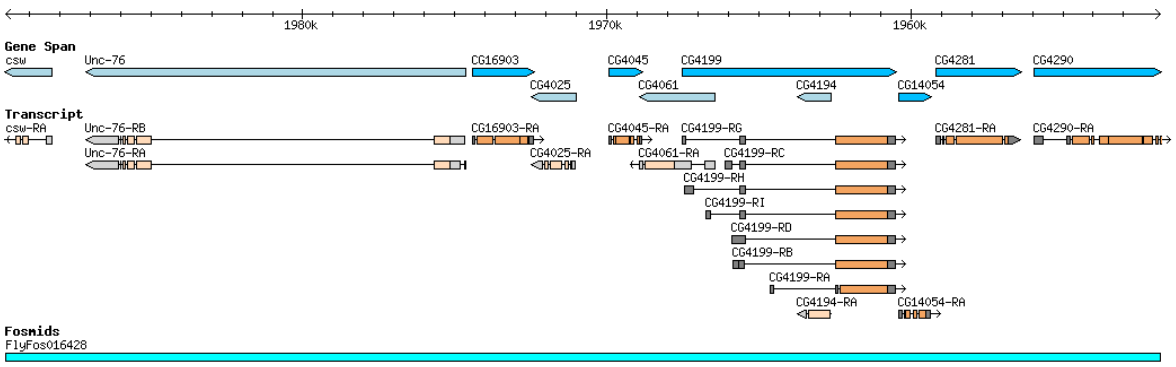
2R [15565402..15606342] (+)

FlyFos016415



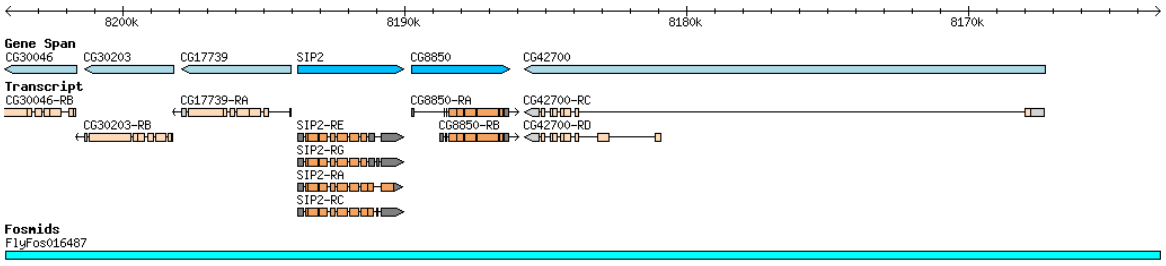
X [18509363..18548300] (-)

FlyFos016428



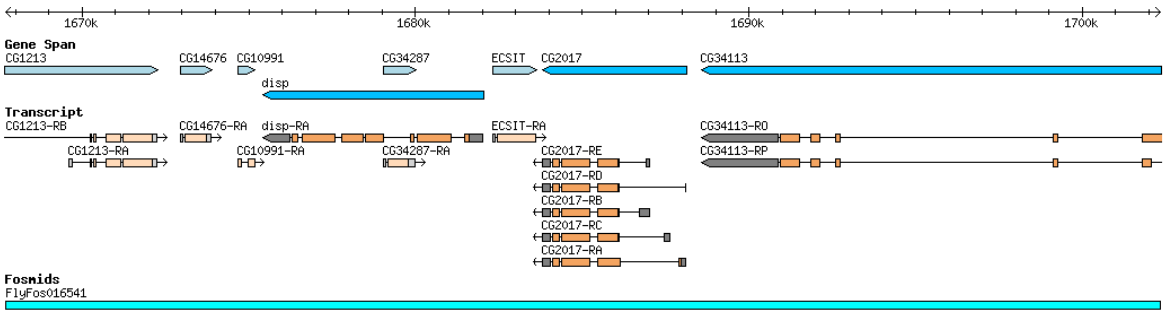
X [1951793..1989772] (-)

FlyFos016487



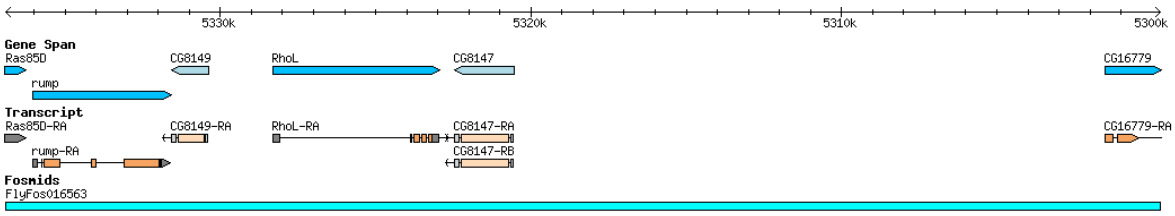
2R [8163217..8204141] (-)

FlyFos016541



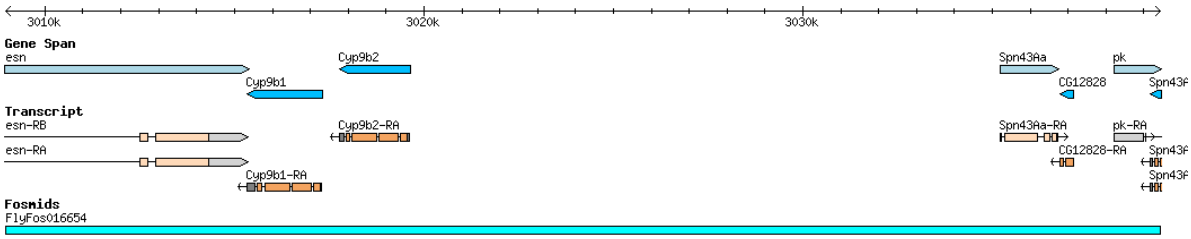
3R [1667710..1702337] (+)

FlyFos016563



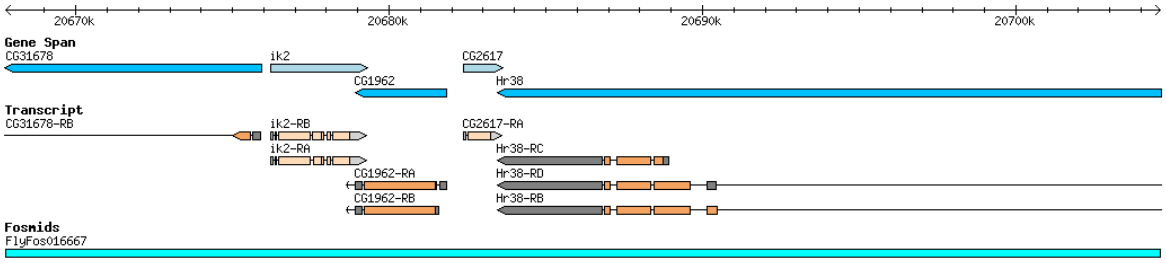
3R [5299736..5336911] (-)

FlyFos016654



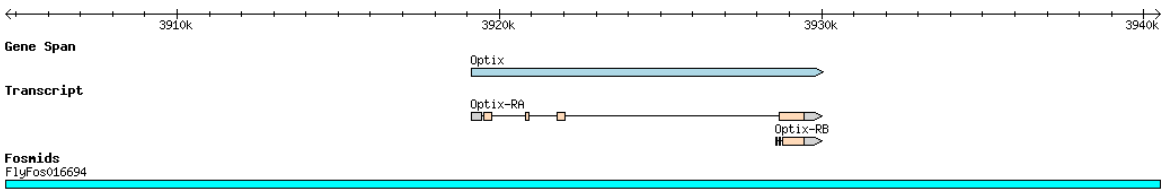
2R [3008962..3039431] (+)

FlyFos016667



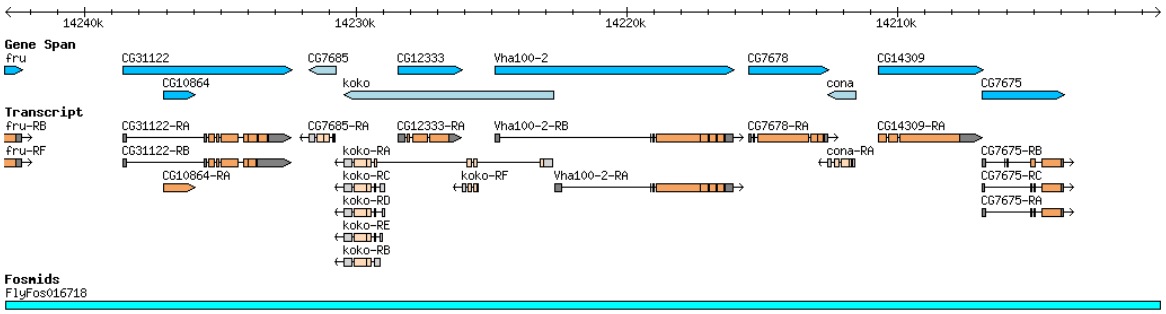
2L [20667764..20704608] (+)

FlyFos016694



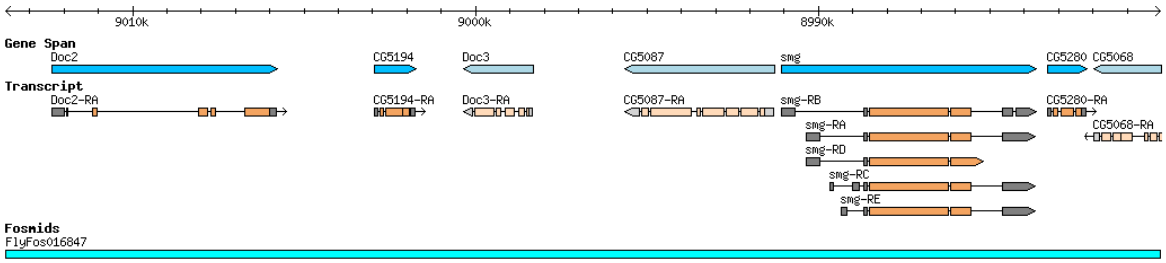
2R [3904681..3940513] (+)

FlyFos016718



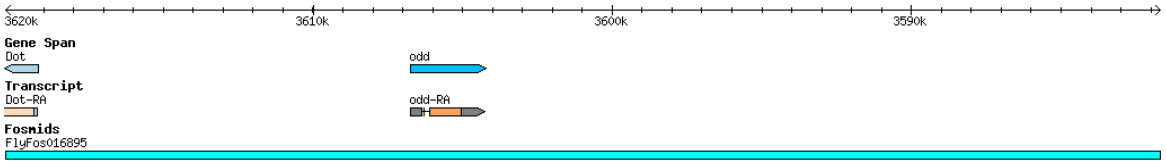
3R [14200322..14242930] (-)

FlyFos016847



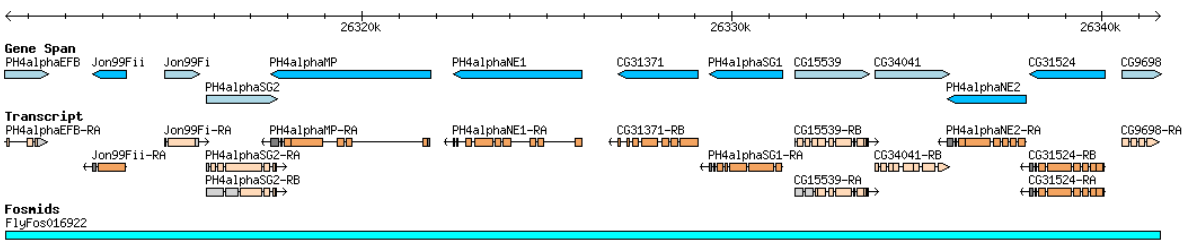
3L [8980033..9013701] (-)

FlyFos016895



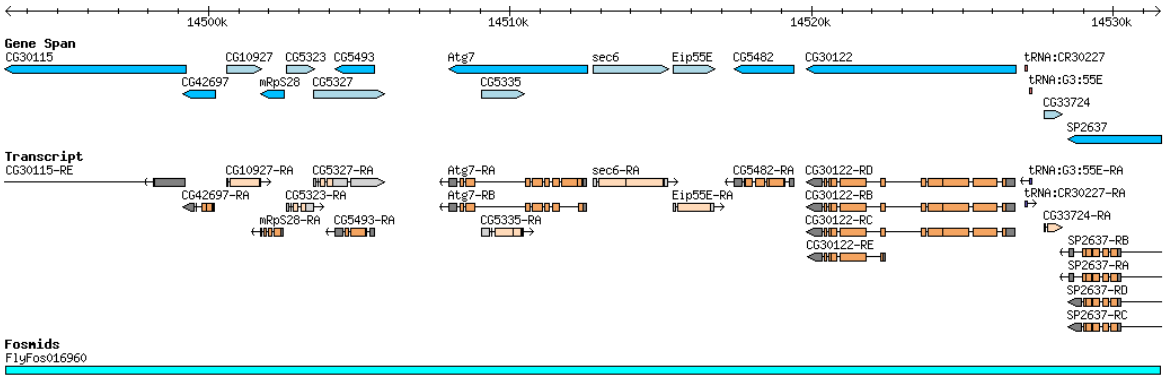
2L [3581684..3620267] (-)

FlyFos016922



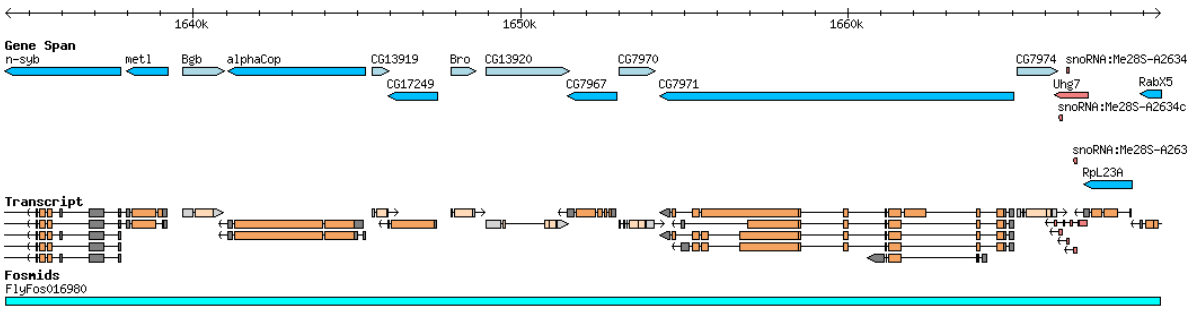
3R [26310377..26341570] (+)

FlyFos016960



2R [14493279..14531555] (+)

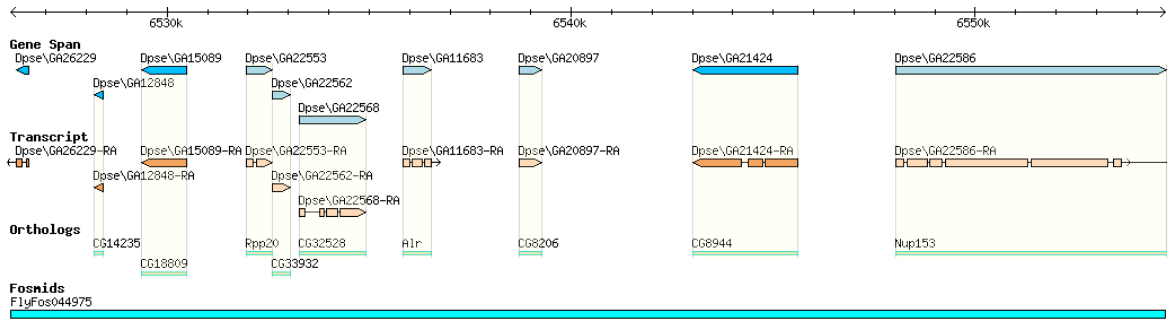
FlyFos016980



3L [1634267..1669527] (+)

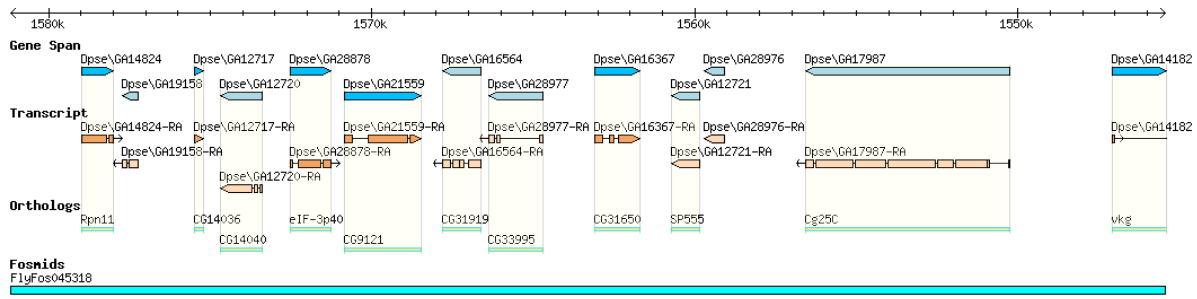
D.2. *Drosophila pseudoobscura* fosmids

FlyFos044975



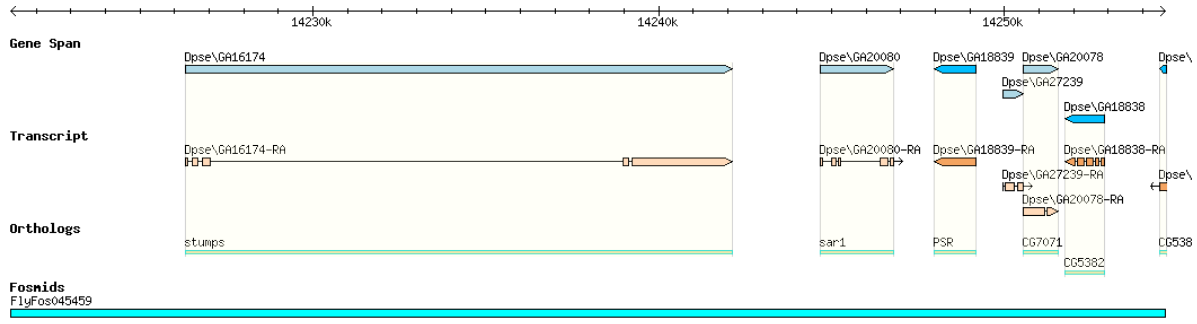
XL_group1e [6526126..6554702] (+)

FlyFos045318



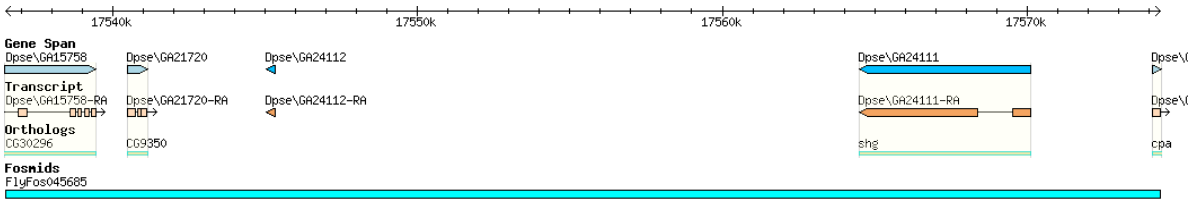
4_group4 [1545444..1581170] (-)

FlyFos045459



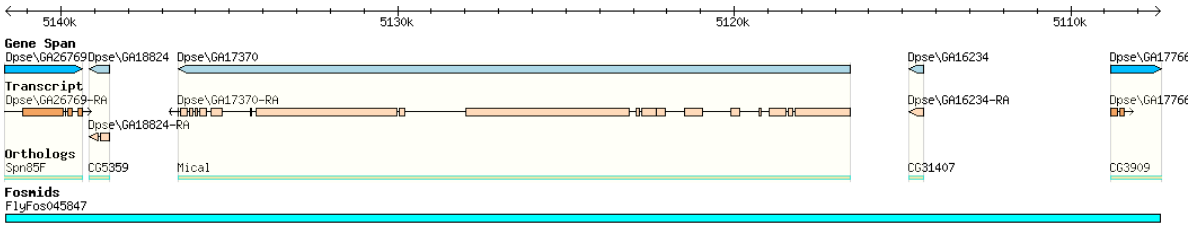
2 [14221258..14254652] (+)

FlyFos045685



3 [17536500..17574365] (+)

FlyFos045847



2 [5107339..5141645] (-)

Appendix E.

Tagging verification alignments

Contents

E.1. Tagging verification principle

The recombineering was validated by two sequencing reactions with primers complementary to the tag sequence, extending towards the tag-fosmid junction (see figure E.1 below). The sequencing results were cropped to 100 bp and evaluated by an automated computer algorithm that predicts the hypothetical, ideal recombineered construct and aligns it with the sequence reads to evaluate the alignment particularly at the tag-fosmid junction. The full results and alignments for all tagging reactions discussed in this work are presented on the following pages.

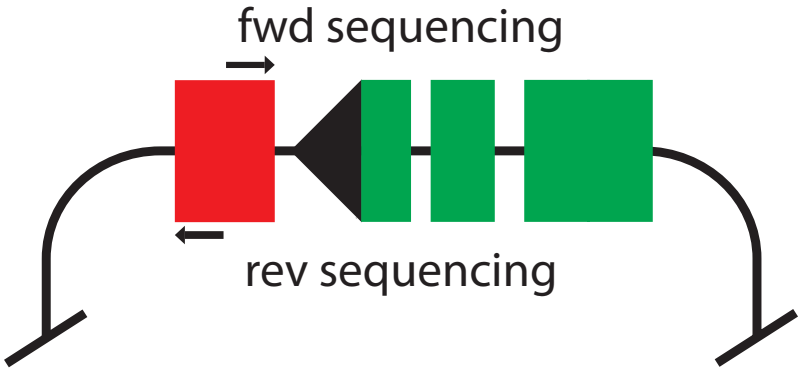


Figure E.1.: Tagging verification principle

In the tables on the following pages, the sequencing reads are shown in the upper line of alignments. Sequence of the tag is in uppercase. Adjacent fosmid sequence is in lowercase. The reference (predicted) sequences are shown in the bottom line of alignments and are in uppercase. Alignment matches are marked with vertical lines. Mismatches are marked with dots. Gaps in the alignment sequences are marked with tildes. The tagging result is color-coded as in figure 3.15: green – correct, yellow – minor mismatch, orange – major mismatch, red – incorrect, gray – sequencing failure. The colors were assinged based on the amount and severity of mismatches in the sequence.

E.2. Tagging with *ubi-mCherry-NLS-T2A*

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>A01</td><td>FWD</td><td></td></tr>
<tr><td>A01</td><td>REV</td><td></td></tr>
<tr><td>A02</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATggaaactcctcgtcacacacgcacgaaccactcgagcgcggcttcacacgcggaaaattcggtgatgttaaaat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGGAAACTCCTCGTCACACACGCACGAACCACTCGAGCGCGGCTTCACACGCGGAAAATTCGGTGATGTTAAAAT</td></tr>
<tr><td>A02</td><td>REV</td><td></td></tr>
<tr><td>A03</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATtcggtgaaattgctaattgtagcgcttagtttgtgcctcctgcagtcgggcatcgtcgagggtatatcattatca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTCGGTGAAATTGCTAATTGTAGCGCTTAGTTTGTGCCTCCTGCAGTCGGGCATCGTCGAGGGTATATCATTATCA</td></tr>
<tr><td>A03</td><td>REV</td><td>cccatatatctattgtattataatagtagtacgagtgtaaccgctgagattagtcgtaaaatcggtgaaataatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCATATATCTATTGTATTATAATAGTAGTACGAGTGTAACCGCTGAGATTAGTCGTAAAATCGGTGAAATAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>A04</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgccgccaactacaaaagctgcccgctaaagaagcgccccattgtcttcgtggaggagcgtctgccacaaacggag
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGCCGCCAACTACAAAAGCTGCCCGCTAAAGAAGCGCCCCATTGTCTTCGTGGAGGAGCGTCTGCCACAAACGGAG</td></tr>
<tr><td>A04</td><td>REV</td><td>gatcgccgatctcccgatttacccatctcgatcagtaccggaaactaaaacttaatcacacacacatcaaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GATCGCCGATCTCCCGATTTACCCATCTCGATCAGTACCGGAAACTAAAACTTAATCACACACACATCAAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>A05</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATacacaaaaccggagcacatggattggatgcagcttgggcggcctgttggtcgctctattggccctgcaaacgatg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATACACAAAACCGGAGCACATGGATTGGATGCAGCTTGGGCGGCCTGTTGGTCGCTCTATTGGCCCTGCAAACGATG</td></tr>
<tr><td>A05</td><td>REV</td><td></td></tr>
<tr><td>A06</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgaggaatccaatcacggttcggctggctgtgaaaacgtatcgcagttcatgctcgatgacctacaattggcagca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGAGGAATCCAATCACGGTTCGGCTGGCTGTGAAAACGTATCGCAGTTCATGCTCGATGACCTACAATTGGCAGCA</td></tr>
<tr><td>A06</td><td>REV</td><td>aaaatcgattaaatctttaaattttcgccagagttcgcatcaagcccaagcttagaaaggtccaagtccaagatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAAATCGATTAAATCTTTAAATTTTCGCCAGAGTTCGCATCAAGCCCAAGCTTAGAAAGGTCCAAGTCCAAGATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>A07</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgtgtccgctctgaaatgcagtttggccgtggccgttatgatcagtctggcttgttcgggtgcgttggttgcgaaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGTGTCCGCTCTGAAATGCAGTTTGGCCGTGGCCGTTATGATCAGTCTGGCTTGTTCGGGTGCGTTGGTTGCGAAA</td></tr>
<tr><td>A07</td><td>REV</td><td>ataatcgtgcctgactttaaaaaaaaaaaatcgttttcgaaaagcaattcccacactcgaagtattcgcgaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATAATCGTGCCTGACTTTAAAAAAAAAAAATCGTTTTCGAAAAGCAATTCCCACACTCGAAGTATTCGCGAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>A08</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaaccactggctaagtagtgagcaccctaaccgaagattcctaatgcccagctgagaactaatccttttcaattct
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAACCACTGGCTAAGTAGTGAGCACCCTAACCGAAGATTCCTAATGCCCAGCTGAGAACTAATCCTTTTCAATTCT</td></tr>
<tr><td>A08</td><td>REV</td><td>tataaatgacaggtggctgggccatttcacttttagtctcgaggtgtc~~~~~gacgcaggcgca~~~~~~~atgCAGA~TCT~~~TCGTGAAGACTCTG
|||||||||||||||||||||||||||||||||||||||||||.||||.....|.||||...|||........||.|||.|||...|.||.|||||..|.
TATAAATGACAGGTGGCTGGGCCATTTCACTTTTAGTCTCGAGATGTCACGCAGGCGCA~ATGCAGATCTTCGTGAAGACTCTGACTGGT~AAGACCATC</td></tr>
<tr><td>A09</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgccgttggaccgacggagggcaaacagccgccctcagagagcttctcgcccacgcaccaccagattatagcaccc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGCCGTTGGACCGACGGAGGGCAAACAGCCGCCCTCAGAGAGCTTCTCGCCCACGCACCACCAGATTATAGCACCC</td></tr>
<tr><td>A09</td><td>REV</td><td></td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>A10</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgccgacaagaagaatctcctcctgcttttcgaccatcccaccgagccagtgttcatggacaagggcaagagggtg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGCCGACAAGAAGAATCTCCTCCTGCTTTTCGACCATCCCACCGAGCCAGTGTTCATGGACAAGGGCAAGAGGGTG</td></tr>
<tr><td>A10</td><td>REV</td><td>cagacttccagtcacattccccatttagtttgctccgcgatccagcaggtcctccctgacatcccattgaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGACTTCCAGTCACATTCCCCATTTAGTTTGCTCCGCGATCCAGCAGGTCCTCCCTGACATCCCATTGAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>A11</td><td>FWD</td><td>cctacgtgattgtg
||||||||||||||
CCTACGTGATTGTG</td></tr>
<tr><td>A11</td><td>REV</td><td>caaat~~~~tagctgtaaatctaaaaaatggatgatgatgcag~~~~~~aataccagaagctccggcgaaacatgCAGAT
.||||.....|..||.|.||||......|.|.|||.||..|.|......||.||||.....||......|.|.||.||.|
~AAATCTAAAAAATGCAGATCT~~~~~~TCG~TGAAGACTCTGACTGGTAAGACCA~~~~~TC~~~~~~ACCCTGGAGGT</td></tr>
<tr><td>A12</td><td>FWD</td><td>c
|
C</td></tr>
<tr><td>A12</td><td>REV</td><td></td></tr>
<tr><td>B01</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATtctcttgagcgtgccgttcgtgccaaggtgaatactcaatcagtgaagaaaaaagatccttaagaaaacatagat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTCTCTTGAGCGTGCCGTTCGTGCCAAGGTGAATACTCAATCAGTGAAGAAAAAAGATCCTTAAGAAAACATAGAT</td></tr>
<tr><td>B01</td><td>REV</td><td></td></tr>
<tr><td>B02</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATccggaacaactgggtctgctatggtccgtgccggagtccaagtcaaaggcgcccatcatcaaggtgtcctgcggc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATCCGGAACAACTGGGTCTGCTATGGTCCGTGCCGGAGTCCAAGTCAAAGGCGCCCATCATCAAGGTGTCCTGCGGC</td></tr>
<tr><td>B02</td><td>REV</td><td>tttataaataaaaccagcggatagcaggaagagatccggtatctcttcgccagagaacgggtaaacaaagcgatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTATAAATAAAACCAGCGGATAGCAGGAAGAGATCCGGTATCTCTTCGCCAGAGAACGGGTAAACAAAGCGATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>B03</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgccctgtccaagatcgacaccgagatcgagcaggttgaccaggagaagtacctgcgtcaggccactagctactat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGCCCTGTCCAAGATCGACACCGAGATCGAGCAGGTTGACCAGGAGAAGTACCTGCGTCAGGCCACTAGCTACTAT</td></tr>
<tr><td>B03</td><td>REV</td><td>ctcaactgatgcaccaccactctaacttccgcttccctttttcgcaacctaggtcaatcagagcaagcccaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCAACTGATGCACCACCACTCTAACTTCCGCTTCCCTTTTCGCAACCTAGGTCAATCAGAGCAAGCCCAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>B04</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaaatcaaagtacgaaaacatgaagatcatctacaatcgcagcaagatcggctggtactgggctccgttgttcgta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAAATCAAAGTACGAAAACATGAAGATCATCTACAATCGCAGCAAGATCGGCTGGTACTGGGCTCCGTTGTTCGTA</td></tr>
<tr><td>B04</td><td>REV</td><td>ctccgaacactcttaatatttatttcggtgaccaatgaagtgaatatgtttatctccccttttaggccacaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTCCGAACACTCTTAATATTTATTTCGGTGACCAATGAAGTGAATATGTTTATCTCCCCTTTTTAGGCCACAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>B05</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaactcctacttcgtgatcgctttgagtgctctttttgtgactctggctgttggatcggtgagtttgagaaaacta
|||||||||||||||||||||||||||||.||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAACT~CTACTTCGTGATCGCTTTGAGTGCTCTTTTTGTGACTCTGGCTGTTGGATCGGTGAGTTTGAGAAAACTA</td></tr>
<tr><td>B05</td><td>REV</td><td>aaagccgcgatcccatcccatgtcggcatcagaacttccccaacgttctaacaagtcaaagtatttctcaacatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAAGCCGCGATCCCATCCCATGTCGGCATCAGAACTTCCCCAACGTTCTAACAAGTCAAAGTATTTCTCAACATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>B06</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATatcaaggcgagagattcgacgaggctgctgctcattagtctgctaattggacaactatacggtaagtcaaggacc
|||.||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAGTTCGAAAGACAACATATGGATATCAAGGCGAGAGATTCGACGAGGcTGCTGCTCATTAGTCTGCTAATTGGACAACTATACGGTAAGTCAAGGACC</td></tr>
<tr><td>B06</td><td>REV</td><td></td></tr>
<tr><td>B07</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaaagaggtaagtctgccgttcggcagcaagcttttccccgagattttcatcatctttgggcattgcaacatcgct
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAAAGAGGTAAGTCTGCCGTTCGGCAGCAAGCTTTTCCCCGAGATTTTCATCATCTTTGGGCATTGCAACATCGCT</td></tr>
<tr><td>B07</td><td>REV</td><td>cagaaagtgcgcaagtgaatagcagtgactatattcatcctgggattaaccaactgctgaacatccaacttaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||.|||||||||||||||||||||||||||||||||||||||||||||||||||
CAGAAAGTGCGCAAGTGAATAGCAGTGACTATATTCATCCTGGGATTATCCAACTGCTGAACATCCAACTTAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>B08</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgatatgtccagcgcgaactcgttgcggccccttttcgcagggtatccctttcaaggtaagtaatttcaacatata
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGATATGTCCAGCGCGAACTCGTTGCGGCCCCTTTTCGCAGGGTATCCCTTTCAAGGTAAGTAATTTCAACATATA</td></tr>
<tr><td>B08</td><td>REV</td><td></td></tr>
<tr><td>B09</td><td>FWD</td><td></td></tr>
<tr><td>B09</td><td>REV</td><td></td></tr>
<tr><td>B10</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgtaagtcaaacgagagagcgagtccttgggtagtacaattgttaagtattggccttccaatttctccaccagcat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGTAAGTCAAACGAGAGAGCGAGTCCTTGGGTAGTACAATTGTTAAGTATTGGCCTTCCAATTTCTCCACCAGCAT</td></tr>
<tr><td>B10</td><td>REV</td><td>ttgtccagggctagctagtataaatagcccgagaaaattctaaattggcacagttcaactgaaaccctcatcatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTGTCCAGGGCTAGCTAGTATAAATAGCCCGAGAAAATTCTAAATTGGCACAGTTCAACTGAAACCCTCATCATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>B11</td><td>FWD</td><td></td></tr>
<tr><td>B11</td><td>REV</td><td></td></tr>
<tr><td>B12</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATcaccacaaacagcatctgcatctgcatctgcagccccccgccccaaacgccacccaaacacaggcccacggactt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATCACCACAAACAGCATCTGCATCTGCATCTGCAGCCCCCCGCCCCAAACGCCACCCAAACACAGGCCCACGGACTT</td></tr>
<tr><td>B12</td><td>REV</td><td>gcatccgatccgagcggcaacaacaaatcatcaagcgtctaataggaaaagtgcagcagacagccagcgaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCATCCGATCCGAGCGGCAACAACAAATCATCAAGCGTCTAATAGGAAAAGTGCAGCAGACAGCCAGCGAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>C01</td><td>FWD</td><td>cccccatcagc
|||||||||||
CCCCCATCAGC</td></tr>
<tr><td>C01</td><td>REV</td><td></td></tr>
<tr><td>C02</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATggagccactagtgtcggagttccagtctggaagattcagctactgattctgctgagtgcaggtgagaatattcta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGGAGCCACTAGTGTCGGAGTTCCAGTCTGGAAGATTCAGCTACTGATTCTGCTGAGTGCAGGTGAGAATATTCTA</td></tr>
<tr><td>C02</td><td>REV</td><td>acaggttgcggctgggttgcctataaaagcagcgggatttgtgtcacttgtcacagaagttgaccaactgcaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACAGGTTGCGGCTGGGTTGCCTATAAAAGCAGCGGGATTTGTGTCACTTGTCACAGAAGTTGACCAACTGCAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>C03</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATtgcaacgctctctgtgaatgcctcaaatgtcccggcaaagtggtttgctggtaagtttccatcggtttctggctc
|||||||||||||||||||||||||||||||||||||||||||||.||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTGCAACGCTCTCTGTGAATGTCTCAAATGTCCCGGCAAAGTGGTTTGCTGGTAAGTTTCCATCGGTTTCTGGCTC</td></tr>
<tr><td>C03</td><td>REV</td><td>acgcatcgggaatcggaaatatactgtacagtatatctatctataatagaataacccaaaaaagtcatcaccatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACGCATCGGGAATCGGAAATATACTGTACAGTATATCTATCTATAATAGAATAACCCAAAAAAGTCATCACCATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>C04</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgaggaattgccgcgcgcattaaactacgaactctcgcacgatttgcatttcgatcattacgccggtgccgcagct
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGAGGAATTGCCGCGCGCATTAAACTACGAACTCTCGCACGATTTGCATTTCGATCATTACGCCGGTGCCGCAGCT</td></tr>
<tr><td>C04</td><td>REV</td><td></td></tr>
<tr><td>C05</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATttgtgcttcgactcggagaggatgaactggtactaccacgtcctggccaggcgtccctacctggtggtcgtctcc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTTGTGCTTCGACTCGGAGAGGATGAACTGGTACTACCACGTCCTGGCCAGGCGTCCCTACCTGGTGGTCGTCTCC</td></tr>
<tr><td>C05</td><td>REV</td><td>aacccaaggagcatttgatgttcccgcaatgcgaagagggtaaagaggattcgggcatcacattctactgacatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AACCCAAGGAGCATTTGATGTTCCCGCAATGCGAAGAGGGTAAAGAGGATTCGGGCATCACATTCTACTGACATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>C06</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATtcagcatttgaactgaaaaatttaaatgattgcttggaaaagcatttaccacccgatgaactaaaggaggttaag
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTCAGCATTTGAACTGAAAAATTTAAATGATTGCTTGGAAAAGCATTTACCACCCGATGAACTAAAGGAGGTTAAG</td></tr>
<tr><td>C06</td><td>REV</td><td>ggtcggagaaggttgttctatcaattgagtccgataattgatgagatattttgttgctgtaaaattggaaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGTCGGAGAAGGTTGTTCTATCAATTGAGTCCGATAATTGATGAGATATTTTGTTGCTGTAAAATTGGAAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>C07</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaggctgcagctcttcttttttctcggcctgagtgttctagtcagtggaggaggcaagttcttatataggtagatc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAGGCTGCAGCTCTTCTTTTTTCTCGGCCTGAGTGTTCTAGTCAGTGGAGGAGGCAAGTTCTTATATAGGTAGATC</td></tr>
<tr><td>C07</td><td>REV</td><td></td></tr>
<tr><td>C08</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATtacatttaccccagctcaccggaatcagcatcgtccttgcagtccgaagagagtgcggccattaagattcaggcc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTACATTTACCCCAGCTCACCGGAATCAGCATCGTCCTTGCAGTCCGAAGAGAGTGCGGCCATTAAGATTCAGGCC</td></tr>
<tr><td>C08</td><td>REV</td><td>gatgaaagccgtatttgttggatagaaacgcggactcagattgccattttgttgcagtgcaccagaggatcatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GATGAAAGCCGTATTTGTTGGATAGAAACGCGGACTCAGATTGCCATTTTGTTGCAGTGCACCAGAGGATCATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>C09</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATctacggctgtgggcctgcctgctcctcctgggatcaatccagatccaggcggttccattctacggcgagtgagta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATCTACGGCTGTGGGCCTGCCTGCTCCTCCTGGGATCAATCCAGATCCAGGCGGTTCCATTCTACGGCGAGTGAGTA</td></tr>
<tr><td>C09</td><td>REV</td><td></td></tr>
<tr><td>C10</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaagttcagtgtggtggtcctggtggccctactaccgctgctaggagctgtttcagcgaatcgcacttttgtggtg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAAGTTCAGTGTGGTGGTCCTGGTGGCCCTACTACCGCTGCTAGGAGCTGTTTCAGCGAATCGCACTTTTGTGGTG</td></tr>
<tr><td>C10</td><td>REV</td><td>tgataaatataatgaaatttttttcccagtatcttaattgatatattatctttcctactgcaatcctttagaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGATAAATATAATGAAATTTTTTTCCCAGTATCTTAATTGATATATTATCTTTCCTACTGCAATCCTTTAGAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>C11</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgccgccaaccagaagattgtgttcgccctcgtctgcctctttttggcatgtgatttggtgctgggtcagcagcag
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGCCGCCAACCAGAAGATTGTGTTCGCCCTCGTCTGCCTCTTTTTGGCATGTGATTTGGTGCTGGGTCAGCAGCAG</td></tr>
<tr><td>C11</td><td>REV</td><td>gtgcagtgtcgaacggattaccagatactcatcctcaaggaatcaaatcaccaacagtcaaatcaaatcgaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGCAGTGTCGAACGGATTACCAGATACTCATCCTCAAGGAATCAAATCACCAACAGTCAAATCAAATCGAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>C12</td><td>FWD</td><td></td></tr>
<tr><td>C12</td><td>REV</td><td></td></tr>
<tr><td>D01</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATtccaagtggtggagctgcggctccaatcaggagagcaacagcatcttccgcagcgaggtgatgtccttggtgcaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTCCAAGTGGTGGAGCTGCGGCTCCAATCAGGAGAGCAACAGCATCTTCCGCAGCGAGGTGATGTCCTTGGTGCAA</td></tr>
<tr><td>D01</td><td>REV</td><td></td></tr>
<tr><td>D02</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgacaccgccggcattattcccgacatcatcgacgtcaagcccgcctccaaggccaccatcacctatccttccggc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||.|.|||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGACACCGCCGGCATTATTCCCGACATCATCGACGTCACGGCCGCCTCCAAGGCCACCATCACCTATCCTTCCGGC</td></tr>
<tr><td>D02</td><td>REV</td><td>acatcagggatccgggccagagtcaactcgctaacgctacaccgagcagaagcaacagaactaccagctaacatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ACATCAGGGATCCGGGCCAGAGTCAACTCGCTAACGCTACACCGAGCAGAAGCAACAGAACTACCAGCTAACATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>D03</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATcagttgggaaaagctatcattctgattttgctggcggccatccagcaaagctgcctggctctctacatcaagagc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATCAGTTGGGAAAAGCTATCATTCTGATTTTGCTGGCGGCCATCCAGCAAAGCTGCCTGGCTCTCTACATCAAGAGC</td></tr>
<tr><td>D03</td><td>REV</td><td>taagggtagctacgcaggagcacaggacatcacattcgccacaaccaccgaacgaagcacatcgatctgaagatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TAAGGGTAGCTACGCAGGAGCACAGGACATCACATTCGCCACAACCACCGAACGAAGCACATCGATCTGAAGATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th></tr>
<tr><td>D04</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATgcaccacagagcaacaacagcaccacattcgtctccaagacccagcactatttgaaggtgaagaagccccttttg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGCACCACAGAGCAACAACAGCACCACATTCGTCTCCAAGACCCAGCACTATTTGAAGGTGAAGAAGCCCCTTTTG</td></tr>
<tr><td>D04</td><td>REV</td><td>gcatcctgcaagcagttcagatcagctcagcacatttctacaaatcttccaaaacaaaaaacacattacaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCATCCTGCAAGCAGTTCAGATCAGCTCAGCACATTTCTACAAATCTTCCAAAACAAAAAACACATTACAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>D05</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATggtagccggccgcagaggatacgccggattggtgatagtctgcagatgctaacccgggaagccagaagcgaagtg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATGGTAGCCGGCCGCAGAGGATACGCCGGATTGGTGATAGTCTGCAGATGCTAACCCGGGAAGCCAGAAGCGAAGTG</td></tr>
<tr><td>D05</td><td>REV</td><td>taaacgacgaaaaagacgtatgta~~~~attagatgcggc~~~~tgcc~~~~~~aagtgccgcggatcagagatgCAGA~TCT~~~TCGTGAAGACTCTG
|||||||||||||||||||||..|....||..|..||.||....||||......||.|..||.||||...|.||.|||.|||...|.||.|||||..|.
TAAACGACGAAAAAGACGTATTAATTAGAT~~GCGGCTGCCAAGTGCCGCGGATCAGAGATGCAGATCTTCG~TGAAGACTCTGACTGGT~AAGACCATC</td></tr>
<tr><td>D06</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaaggttttcgttgccatctgcgtgctgattggactggtgagtgctcgatacagataaacggcgaccaggaccaga
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAAGGTTTTCGTTGCCATCTGCGTGCTGATTGGACTGGTGAGTGCTCGATACAGATAAACGGCGACCAGGACCAGA</td></tr>
<tr><td>D06</td><td>REV</td><td>cctggccctatttcaaaacagtcttcgctcgatcgctggaggaatacatacataggtggaaagaaagtgaaaatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTGGCCCTATTTCAAAACAGTCTTCGCTCGATCGCTGGAGGAATACATACATAGGTGGAAAGAAAGTGAAAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>D07</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATttggatcggcactgtctttatattgggattttccagcttataatttgggttggagtagcgaacggtgagttttac
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATTTGGATCGGCACTGTCTTTATATTGGGATTTTCCAGCTTATAATTTGGGTTGGAGTAGCGAACGGTGAGTTTTAC</td></tr>
<tr><td>D07</td><td>REV</td><td>tttggctcagttcgtacctgaacaagaaccagtctacatcagtaactcgtggttcacagtgctctggtcataatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TTTGGCTCAGTTCGTACCTGAACAAGAACCAGTCTACATCAGTAACTCGTGGTTCACAGTGCTCTGGTCATAATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
<tr><td>D08</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATacgagcatttgcagcagcaaattccagcagcagcattaccagctgaccaacagtaacattttcttgctgcaacat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATACGAGCATTTGCAGCAGCAAATTCCAGCAGCAGCATTACCAGCTGACCAACAGTAACATTTTCTTGCTGCAACAT</td></tr>
<tr><td>D08</td><td>REV</td><td>tcattgcggaatctgattccacacagtcaacatctgtaaactaaatcttagaaaactctcgcaaggattaccatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCATTGCGGAATCTGATTCCACACAGTCAACATCTGTAAACTAAATCTTAGAAAACTCTCGCAAGGATTACCATGCAGATCTTCGTGAAGACTCTGACTG</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment</th><th></th></tr>
<tr><td>D09</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATcttcgctcgcgttcagaagtctttgtgtgtttgcttatgatcctgggatcagttctggtccgttcgaatcttggg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATCTTCGCTCGCGTTCAGAAGTCTTTGTGTGTTTGCTTATGATCCTGGGATCAGTTCTGGTCCGTTCGAATCTTGGG</td><td></td></tr>
<tr><td>D09</td><td>REV</td><td>gggtccaagggaaattctggcaccgcttgctgctcaggtagaaacaacaaaaaacgaatatcagtcgagaaaatgCAGATCTTCGTGAAGACTCTGACTG
|||||||||||||||||||||||||||||||||||||?|??|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGGTCCAAGGGAAATTCTGGCACCGCTTGCTGCTCAGNTNNAAACAACAAAAAACGAATATCAGTCGAGAAAATGCAGATCTTCGTGAAGACTCTGACTG</td><td></td></tr>
<tr><td>D10</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATaaacccattgggtgcaacaatattccggtcatcttcctggtgatcctcggcatggtcagcctggccaattcgctg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATAAACCCATTGGGTGCAACAATATTCCGGTCATCTTCCTGGTGATCCTCGGCATGGTCAGCCTGGCCAATTCGCTG</td><td></td></tr>
<tr><td>D10</td><td>REV</td><td>ggccacgatcgatcgacatgcgacaggcgggcagcagtgccacagcaaagctactagcagtcggacgtaaacatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCCACGATCGATCGACATGCGACAGGCGGGCAGCAGTGCCACAGCAAAGCTACTAGCAGTCGGACGTAAACATGCAGATCTTCGTGAAGACTCTGACTG</td><td></td></tr>
<tr><td>D11</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATatgccgccaagactgccaggcggccatggaggagccatgcgtagtcggagcagcagcagtggccaccacttaaac
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATATGCCGCCAAGACTGCCAGGCGGCCATGGAGGAGCCATGCGTAGTCGGAGCAGCAGCAGTGGCCACCACTTAAAC</td><td></td></tr>
<tr><td>D11</td><td>REV</td><td>caacgagaccattaaaccattaaatcttac~aaaatccataaaagtatcgttctctcgcttctctgctgcagatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||.|||||||||||||||||||.|||||||||||||||||||||||||||||||||||||||||||||||||
CAACGAGACCATTAAACCATTAAATCTTACAAAAATCCATAAAAGTATCGCTCTCTCGCTTCTCTGCTGCAGATGCAGATCTTCGTGAAGACTCTGACTG</td><td></td></tr>
<tr><td>D12</td><td>FWD</td><td>CAAATTCGAAAGACAACATATGGATcatctgccagcgggtccaacgatggtggccaacaacacacaggtcctggccgctgccgccgccgcagcagccgcc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAAATTCGAAAGACAACATATGGATCATCTGCCAGCGGGTCCAACGATGGTGGCCAACAACACACAGGTCCTGGCCGCTGCCGCCGCCGCAGCAGCCGCC</td><td></td></tr>
<tr><td>D12</td><td>REV</td><td>atccgcaatagaaaaccggcaattgtcgacagccccaggattacggctacgatttccacattcggatacgagatgCAGATCTTCGTGAAGACTCTGACTG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCCGCAATAGAAAACCGGCAATTGTCGACAGCCCCAGGATTACGGCTACGATTTCCACATTCGGATACGAGATGCAGATCTTCGTGAAGACTCTGACTG</td><td></td></tr>
</table>

E.3. Tagging with *EGFP*

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>A01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagattgaaggaccagcgaaacccggtatccggtttcttttgagacattctaaaagagattcagagagctggacg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATTGAAGGACCAGCGAAACCCGGTATCCGGTTTCTTTTGAGACATTCTAAAAGAGATTCAGAGAGCTGGACG</td></tr>
<tr><td>A01</td><td>REV</td><td>ctatcctcgtcatcctcgagctcagagaagcccaagcggatgctgggcttcaccatcgatgagatcatgagcagaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTATCCTCGTCATCCTCGAGCTCAGAGAAGCCCAAGCGGATGCTGGGCTTCACCATCGATGAGATCATGAGCAGAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggcagcgattgcggggtcaccagggggacaaagtgtggttcgcttgaaaaaaacgtaacttagtttaaagttc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGCAGCGATTGCGGGGTCACCAGGGGGACAAAGTGTGGTTCGCTTGAAAAAAACGTAACTTAGTTTAAAGTTC</td></tr>
<tr><td>A02</td><td>REV</td><td>cctctggccaggcacagcaccaatccgttcatctcaccgcccaaggcgccggcgcagtcattccaggtgcagctcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTCTGGCCAGGCACAGCACCAATCCGTTCATCTCACCGCCCAAGGCGCCGGCGCAGTCATTCCAGGTGCAGCTCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaggggaatagaactatggttataatttagttattcacagtatattattattagttaggtatttattgtttag
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGGGGAATAGAACTATGGTTATAATTTAGTTATTCACAGTATATTATTATTAGTTAGGTATTTATTGTTTAG</td></tr>
<tr><td>A03</td><td>REV</td><td>tggggcatcgatgcccgggaacttttgtgggatcgaatgctggaaataatgcgaaatcatgagaattcaattatcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGGGCATCGATGCCCGGGAACTTTTGTGGGATCGAATGCTGGAAATAATGCGAAATCATGAGAATTCAATTATCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggattctcgacatatgaatcccttagcaggacacaattacctaagcaaactcacattccgctggcgaagacca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGATTCTCGACATATGAATCCCTTAGCAGGACACAATTACCTAAGCAAACTCACATTCCGCTGGCGAAGACCA</td></tr>
<tr><td>A04</td><td>REV</td><td>gtgtgccacaaatctttctccaggatgtcgctcctgaacaagcactccagctccaactgcaccatcactattgcgGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGTGCCACAAATCTTTCTCCAGGATGTCGCTCCTGAACAAGCACTCCAGCTCCAACTGCACCATCACTATTGCGGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>A05</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggttggaaatatagaaattttaactaatttatacttaaaagattaaaaaaaaaaaaatagtaaaaccacaaaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGTTGGAAATATAGAAATTTTAACTAATTTATACTTAAAAGATTAAAAAAAAAAAAATAGTAAAACCACAAAA</td></tr>
<tr><td>A05</td><td>REV</td><td>gtgtggaaatttgtgtggcgcaagttatttggcaaatcgaacaaagtttccaagggaaagaaggtgaagaagcagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGTGGAAATTTGTGTGGCGCAAGTTATTTGGCAAATCGAACAAAGTTTCCAAGGGAAAGAAGGTGAAGAAGCAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A06</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaacttgtttagagaatgtaaataagcaattaaacagtgcattctagccatagggcattctaccattttaaat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAACTTGTTTAGAGAATGTAAATAAGCAATTAAACAGTGCATTCTAGCCATAGGGCATTCTACCATTTTAAAT</td></tr>
<tr><td>A06</td><td>REV</td><td>agatcacattcccagcgacagcaacacaacggagccaacattcgcaagagtcatcatcatcagtttcgtcaaaagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGATCACATTCCCAGCGACAGCAACACAACGGAGCCAACATTCGCAAGAGTCATCATCATCAGTTTCGTCAAAAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A07</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagatcttaactagctagtaaattacctgtgcgtagtatttaacgatcttgttctctggaaatttcttctaaatt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATCTTAACTAGCTAGTAAATTACCTGTGCGTAGTATTTAACGATCTTGTTCTCTGGAAATTTCTTCTAAATT</td></tr>
<tr><td>A07</td><td>REV</td><td>tgcaacggatcctcctccctggcccccatcgccggagccatcctgctcttcttcggcgtggctcgtctgctggccGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGCAACGGATCCTCCTCCCTGGCCCCCATCGCCGGAGCCATCCTGCTCTTCTTCGGCGTGGCTCGTCTGCTGGCCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A08</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatcgtcacacatttccctcagattaagcactttaaattgtaatcattacatcaataaataaatgcggagaacc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATCGTCACACATTTCCCTCAGATTAAGCACTTTAAATTGTAATCATTACATCAATAAATAAATGCGGAGAACC</td></tr>
<tr><td>A08</td><td>REV</td><td>gtggccgatcatccatttgcgttcatcattcgcgacaagcacgctgtctatttcaccggacacattgtcaagtttGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGGCCGATCATCCATTTGCGTTCATCATTCGCGACAAGCACGCTGTCTATTTCACCGGACACATTGTCAAGTTTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A09</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgatatcagcgggtctgaggtgtccacctgtaaccccacccagactaaatcaacaaccccaaaccgaaatcccaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGATATCAGCGGGTCTGAGGTGTCCACCTGTAACCCCACCCAGACTAAATCAACAACCCCAAACCGAAATCCCAA</td></tr>
<tr><td>A09</td><td>REV</td><td>cacagtggccaactgatgctgcatcggcccttctccacgtcgccggagctgaagcacagtgctcccgagatcacaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACAGTGGCCAACTGATGCTGCATCGGCCCTTCTCCACGTCGCCGGAGCTGAAGCACAGTGCTCCCGAGATCACAGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>A10</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagacgatccgcacctgattcgctttcccgattcgccctgacccgattccatgacctcgcccacttctagatata
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGACGATCCGCACCTGATTCGCTTTCCCGATTCGCCCTGACCCGATTCCATGACCTCGCCCACTTCTAGATATA</td></tr>
<tr><td>A10</td><td>REV</td><td>aacttcctcacgcccaacatgagcatcgtggacgtgaacatccgccacgagaaccgcaccgtgcagcgcccaaacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AACTTCCTCACGCCCAACATGAGCATCGTGGACGTGAACATCCGCCACGAGAACCGCACCGTGCAGCGCCCAAACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaggactactctagtgtattttagtgttacgctttattattaatgcaattggtaattaatatattcttaggct
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGGACTACTCTAGTGTATTTTAGTGTTACGCTTTATTATTAATGCAATTGGTAATTAATATATTCTTAGGCT</td></tr>
<tr><td>A11</td><td>REV</td><td>tacataatccatcgaaaacgcgaggcggactttaagagcccacgtggcggatacttgttcgacaatatctttggcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACATAATCCATCGAAAACGCGAGGCGGACTTTAAGAGCCCACGTGGCGGATACTTGTTCGACAATATCTTTGGCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagccaatgggcccacggccgtcgactccaagcgcttcaggtccaatccatcaaccagccctcgaatgcataaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGCCAATGGGCCCACGGCCGTCGACTCCAAGCGCTTCAGGTCCAATCCATCAACCAGCCCTCGAATGCATAAA</td></tr>
<tr><td>A12</td><td>REV</td><td>atcgccggagccgaggatgtgtccgccacttcgttcgctcttgtgggcatcctggcggcgttgctcttcgccagaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCGCCGGAGCCGAGGATGTGTCCGCCACTTCGTTCGCTCTTGTGGGCATCCTGGCGGCGTTGCTCTTCGCCAGAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagcgccaaaggatggccaggatgtccacaccctttttctacacttatgctaagtgaacacacccatatatattt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGCGCCAAAGGATGGCCAGGATGTCCACACCCTTTTCTACACTTATGCTAAGTGAACACACCCATATATATTT</td></tr>
<tr><td>B01</td><td>REV</td><td>caggccggttccaacaagggagccacccaggctggccagaacctcggcgctggccgcaagatcctgctcggcaagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGCCGGTTCCAACAAGGGAGCCACCCAGGCTGGCCAGAACCTCGGCGCTGGCCGCAAGATCCTGCTCGGCAAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaggagccgttcacccagctaaccggcttatccctggcagcagaagctgtcgcatccctacacctgcactgaga
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGGAGCCGTTCACCCAGCTAACCGGCTTATCCCTGGCAGCAGAAGCTGTCGCATCCCTACACCTGCACTGAGA</td></tr>
<tr><td>B02</td><td>REV</td><td>ggcgacaaccaggagctgcgcacgaacaccattgagaacatgctgatggccctgcccagcgcctccaaggccaagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCGACAACCAGGAGCTGCGCACGAACACCATTGAGAACATGCTGATGGCCCTGCCCAGCGCCTCCAAGGCCAAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>B03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagatccatcaagatcctaaaaatatagcctctcgtattgttgcataatactaagaatctttattactagatatt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATCCATCAAGATCCTAAAAATATAGCCTCTCGTATTGTTGCATAATACTAAGAATCTTTATTACTAGATATT</td></tr>
<tr><td>B03</td><td>REV</td><td>gacaacctgaagaagaccaccgccaaggtcaccttctggagcaaatacggcgtgaggacgaagcagaacgagcagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACAACCTGAAGAAGACCACCGCCAAGGTCACCTTCTGGAGCAAATACGGCGTGAGGACGAAGCAGAACGAGCAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagtgtagactcaaagttccttcacgaactcaactcaatcccatttttgccatgacacctcagctactcttaat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGTGTAGACTCAAAGTTCCTTCACGAACTCAACTCAATCCCATTTTTGCCATGACACCTCAGCTACTCTTAAT</td></tr>
<tr><td>B04</td><td>REV</td><td>cgtgagttcctggccacgtttccgccgtgggcgcatgttagctcctggctgggatcctacaatagttggcagcttGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTGAGTTCCTGGCCACGTTTCCGCCGTGGGCGCATGTTAGCTCCTGGCTGGGATCCTACAATAGTTGGCAGCTTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B05</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaagaagttagggatttaatgcttggcaaattgtgattcgggaaaaaatgtaacaaaatttaaataaattcttt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGAAGTTAGGGATTTAATGCTTGGCAAATTGTGATTCGGGAAAAAATGTAACAAAATTTAAATAAATTCTTT</td></tr>
<tr><td>B05</td><td>REV</td><td>gagaacaagtgtgataccgcctccaagttgtacgattgcttcgagagcttcaagcccgcccccgaggctaaggccGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGAACAAGTGTGATACCGCCTCCAAGTTGTACGATTGCTTCGAGAGCTTCAAGCCCGCCCCCGAGGCTAAGGCCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B06</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaggggtgccacccatccagaccgagatgtgtacatacgtatttccggactactcagctatcgaggctatcgga
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGGGGTGCCACCCATCCAGACCGAGATGTGTACATACGTATTTCCGGACTACTCAGCTATCGAGGCTATCGGA</td></tr>
<tr><td>B06</td><td>REV</td><td>ggagctgcgtcatcgatgcagcgcctgaacgtgggagtgatcctgctggcagcgctcctgctgcgagtccgcctcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGAGCTGCGTCATCGATGCAGCGCCTGAACGTGGGAGTGATCCTGCTGGCAGCGCTCCTGCTGCGAGTCCGCCTCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B07</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagtgtcccgggcatcggcaaccgcataatccgagagtatcccatctgtccgatccgatccaagtcgatccgagg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGTGTCCCGGGCATCGGCAACCGCATAATCCGAGAGTATCCCATCTGTCCGATCCGATCCAAGTCGATCCGAGG</td></tr>
<tr><td>B07</td><td>REV</td><td>gagtcggtacaggagctcgtccgtcacctgtccggccaccacaataacctgctgctgacaaagaatctgcgcgaaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGTCGGTACAGGAGCTCGTCCGTCACCTGTCCGGCCACCACAATAACCTGCTGCTGACAAAGAATCTGCGCGAAGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>B08</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgatttaatgaatcgccgacgtcgctgtacgaccgagaatcattacattttcgcgttagttttatgcatttcaat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGATTTAATGAATCGCCGACGTCGCTGTACGACCGAGAATCATTACATTTTCGCGTTAGTTTTATGCATTTCAAT</td></tr>
<tr><td>B08</td><td>REV</td><td>gcagcggctcacgggaacccggcctccgcctacagccaccccctgccgacgcagggtcaggccaagtactggtcaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCAGCGGCTCACGGGAACCCGGCCTCCGCCTACAGCCACCCCCTGCCGACGCAGGGTCAGGCCAAGTACTGGTCAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B09</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggtgcctgctaccggtgcatcacgtactcatagtcattcactgatccagctttgttttagcaccttaagttgg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGTGCCTGCTACCGGTGCATCACGTACTCATAGTCATTCACTGATCCAGCTTTGTTTTAGCACCTTAAGTTGG</td></tr>
<tr><td>B09</td><td>REV</td><td>tggtaccagaccgcccgcctccaggacgaggccacgacggcagcgcagcctgctgcgaacggcgtaaagcaggacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGTACCAGACCGCCCGCCTCCAGGACGAGGCCACGACGGCAGCGCAGCCTGCTGCGAACGGCGTAAAGCAGGACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B10</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagaggctaactaaagtgatcaataccgaaagaacaacaaagaaacgaggtggaaactaaggcatatccttgta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGAGGCTAACTAAAGTGATCAATACCGAAAGAACAACAAAGAAACGAGGTGGAAACTAAGGCATATCCTTGTA</td></tr>
<tr><td>B10</td><td>REV</td><td>tccgttctgtttgccgtggagattccgaagcccatctatcgcttcttcaagggcatctttggcggtttctccaacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCGTTCTGTTTGCCGTGGAGATTCCGAAGCCCATCTATCGCTTCTTCAAGGGCATCTTTGGCGGTTTCTCCAACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaagggaaaacttaagattaaatttagtttaataatttaataaaaactgtactgataatgtctaaaaagaatat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAAGGGAAAACTTAAGATTAAATTTAGTTTAATAATTTAATAAAAACTGTACTGATAATGTCTAAAAAGAATAT</td></tr>
<tr><td>B11</td><td>REV</td><td>gctttggtcgaactgaaggagaagtatgccacgctcatccaaccgcgtaactcaaaccaatatgcagtcatcattGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTTTGGTCGAACTGAAGGAGAAGTATGCCACGCTCATCCAACCGCGTAACTCAAACCAATATGCAGTCATCATTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggacgtatgtatgcgtggcgtcttcatttcacttggcatttaatatttgtaggctatagtcttgtattgtact
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGACGTATGTATGCGTGGCGTCTTCATTTCACTTGGCATTTAATATTTGTAGGCTATAGTCTTGTATTGTACT</td></tr>
<tr><td>B12</td><td>REV</td><td>gcgaaccagacggtgaacgtcaactactacggcgcctgtggccgccccgaagcaccatccactaacttcctttacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCGAACCAGACGGTGAACGTCAACTACTACGGCGCCTGTGGCCGCCCCGAAGCACCATCCACTAACTTCCTTTACGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th><th></th></tr>
<tr><td>C01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaactatcaagtgcagacgggtgcatgtgctcccgtttggcttggaatcggtgccttgtacatttaattagcg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAACTATCAAGTGCAGACGGGTGCATGTGCTCCCGTTTGGCTTGGAATCGGTGCCTTGTACATTTAATTAGCG</td><td></td></tr>
<tr><td>C01</td><td>REV</td><td>agcgatccgacgatggccaagccgaagcgcagcagcttcagcatctcggacatattaggaaccagctcgtccattGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCGATCCGACGATGGCCAAGCCGAAGCGCAGCAGCTTCAGCATCTCGGACATATTAGGAACCAGCTCGTCCATTGAAGTGCATACCAATCAGGACCCGC</td><td></td></tr>
<tr><td>C02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagatatacattattaacgaaaacaaaaccacgggttttcaattggattcatgttttaatgtactaacaacaatg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATATACATTATTAACGAAAACAAAACCACGGGTTTTCAATTGGATTCATGTTTTAATGTACTAACAACAATG</td><td></td></tr>
<tr><td>C02</td><td>REV</td><td>cgtgctctttcgcaaaacttaattcgaaattttgaacttcaaaaactaagaagagccaataaagtgcaaaaatatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTGCTCTTTCGCAAAACTTAATTCGAAATTTTGAACTTCAAAAACTAAGAAGAGCCAATAAAGTGCAAAAATATGAAGTGCATACCAATCAGGACCCGC</td><td></td></tr>
<tr><td>C03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgatctgaggacatgaatttatactataggccatattaataataactccgtcgaaatcgaaattgaaacgaacta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGATCTGAGGACATGAATTTATACTATAGGCCATATTAATAATAACTCCGTCGAAATCGAAATTGAAACGAACTA</td><td></td></tr>
<tr><td>C03</td><td>REV</td><td>gaggtgcagaagcaggtcgcccaactgacgcccattgtgaagcgcagcatacgcgactacttcaacaaggagtacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||.|||||||||||||||||||||||||||||||||||||||||||||||||
GAGGTGCAGAAGCAGGTCGCCCAACTGACGCCCATTGTGAAGCGCAGCATTCGCGACTACTTCAACAAGGAGTACGAAGTGCATACCAATCAGGACCCGC</td><td></td></tr>
<tr><td>C04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaggtaaaagtgacgagaatacgaatacgaatacctagccaagcaactgagctctgtgatattttcatgttca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGGTAAAAGTGACGAGAATACGAATACGAATACCTAGCCAAGCAACTGAGCTCTGTGATATTTTCATGTTCA</td><td></td></tr>
<tr><td>C04</td><td>REV</td><td>gtcgacgagaaaaagaaatcgaagtcaaaagacagccagtcga~~~~~~aagacgatatcaa~~~~~~~~~gcggGAAGTGCATACCAATCAGGACCCGC
|||||||||||||||||||||||||||||||||||||||||||......|||...|||.|||.........||.|||..||.|..|||..||..|.||..
GTCGACGAGAAAAAGAAATCGAAGTCAAAAGACAGCCAGTCGAAGCGGGAAGTGCATACCAATCAGGACCCGCTGGA~~TGAAGTCCA~~CACAAACC~~</td><td></td></tr>
<tr><td>C05</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagttagttactattgccggacaacgcgttgttgttgccaagaagaatcaggcaactgcatttttatacagggt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGTTAGTTACTATTGCCGGACAACGCGTTGTTGTTGCCAAGAAGAATCAGGCAACTGCATTTTTATACAGGGT</td><td></td></tr>
<tr><td>C05</td><td>REV</td><td>ctgccgagggacttcgagcactcattccagacgatgcacgagtgcaaatatcaaacgtatccgtctacatccaatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGCCGAGGGACTTCGAGCACTCATTCCAGACGATGCACGAGTGCAAATATCAAACGTATCCGTCTACATCCAATGAAGTGCATACCAATCAGGACCCGC</td><td></td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>C06</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagagaaaagagtattttacgattgactttgttgattagcggaattgattttgaagaaaattgcattttgattt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGAGAAAAGAGTATTTTACGATTGACTTTGTTGATTAGCGGAATTGATTTTGAAGAAAATTGCATTTTGATTT</td></tr>
<tr><td>C06</td><td>REV</td><td>agagttccgctgtatgcggagtccttcaaaaaggcatccgaacatggcttcaagccgcagatcatcaaggaaacaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGAGTTCCGCTGTATGCGGAGTCCTTCAAAAAGGCATCCGAACATGGCTTCAAGCCGCAGATCATCAAGGAAACAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C07</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaaacttggcggggatcaaaggttaagtgaagcacaatgagtctaagcgacaaacgtattattctcgtttaaga
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAAACTTGGCGGGGATCAAAGGTTAAGTGAAGCACAATGAGTCTAAGCGACAAACGTATTATTCTCGTTTAAGA</td></tr>
<tr><td>C07</td><td>REV</td><td>cacgcccttggttacgcctcttgcctcagtgatcgtaacatgtgcgtggatgggggcgtggcacggagaccacgcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACGCCCTTGGTTACGCCTCTTGCCTCAGTGATCGTAACATGTGCGTGGATGGGGGCGTGGCACGGAGACCACGCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C08</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatgactttgcgcgctggtcgttccacaactctgatttctactgtacatacaaatatttgtattcaaatcctac
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATGACTTTGCGCGCTGGTCGTTCCACAACTCTGATTTCTACTGTACATACAAATATTTGTATTCAAATCCTAC</td></tr>
<tr><td>C08</td><td>REV</td><td>gccgctgttaaaattcaggctggcttccggggattcaaaacacgcaaagaattgaaacaatgcgagcccattgtgGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCGCTGTTAAAATTCAGGCTGGCTTCCGGGGATTCAAAACACGCAAAGAATTGAAACAATGCGAGCCCATTGTGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C09</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagcggtgctcgtccccatctcatgcatattgatatataaagcagatatttatatttactcttaacgatttgtc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGCGGTGCTCGTCCCCATCTCATGCATATTGATATATAAAGCAGATATTTATATTTACTCTTAACGATTTGTC</td></tr>
<tr><td>C09</td><td>REV</td><td>cagcaacagcagcgaccccagctgatcccgcccggagctggctatcagccacagggcgatttcgatgtcttcttcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGCAACAGCAGCGACCCCAGCTGATCCCGCCCGGAGCTGGCTATCAGCCACAGGGCGATTTCGATGTCTTCTTCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C10</td><td>FWD</td><td>gtggagtatcagcagacacccgcctcacaggagctgcacttccgagataccccccattctgaacgcgaggaccgttGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGGAGTATCAGCAGACACCCGCCTCACAGGAGCTGCACTTCCGAGATACCCCCATTCTGAACGCGAGGACCGTTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C10</td><td>REV</td><td></td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>C11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagatgatttgtttcggttttggctgaccaggatgacaatgcaagaccagggataacggcgagctggtagcgagt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATGATTTGTTTCGGTTTTGGCTGACCAGGATGACAATGCAAGACCAGGGATAACGGCGAGCTGGTAGCGAGT</td></tr>
<tr><td>C11</td><td>REV</td><td>cccgtccaggcacccgtgccggtggtctactcccactcgcacacccagcagcccgtctggttggagaaggagtggGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCGTCCAGGCACCCGTGCCGGTGGTCTACTCCCACTCGCACACCCAGCAGCCCGTCTGGTTGGAGAAGGAGTGGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaccgaacccgggtattgtgaactcacctcttcccaccccgttgtgatatatgatacatatatgtaatacata
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAACCGAACCCGGGTATTGTGAACTCACCTCTTCCCACCCCGTTGTGATATATGATACATATATGTAATACATA</td></tr>
<tr><td>C12</td><td>REV</td><td>tcactcacgtgcaaaatgcagactgtgatgggcgctgagacgcagaaaatgctgaagaacagcgaggattatgttGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCACTCACGTGCAAAATGCAGACTGTGATGGGCGCTGAGACGCAGAAAATGCTGAAGAACAGCGAGGATTATGTTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagaaccatcaggatagccatcgtgattcgctccttgaagaaggtttaattaaaaaagcattacagaataaaaag
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGAACCATCAGGATAGCCATCGTGATTCGCTCCTTGAAGAAGGTTTAATTAAAAAAGCATTACAGAATAAAAAG</td></tr>
<tr><td>D01</td><td>REV</td><td>aagttctatgtgggaaacggatatccgttcacgccattcagctttaaggatattttgatcgtcgtcgaagatgatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAGTTCTATGTGGGAAACGGATATCCGTTCACGCCATTCAGCTTTAAGGATATTTTGATCGTCGTCGAAGATGATGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatctggccaccaactgatcagctctctgtgaaataataaatattaaatatgtactagttctcataaaagttat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATCTGGCCACCAACTGATCAGCTCTCTGTGAAATAATAAATATTAAATATGTACTAGTTCTCATAAAAGTTAT</td></tr>
<tr><td>D02</td><td>REV</td><td>ggcggtcccgtggccggcaacttcttccaggcccaatacgatgactacgtgaagaccctcatcgagacggtccagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCGGTCCCGTGGCCGGCAACTTCTTCCAGGCCCAATACGATGACTACGTGAAGACCCTCATCGAGACGGTCCAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaatccctagactatcgatgatggagggctgtgcaatagaggttcatatgctggcattggacttgtctttaggc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAATCCCTAGACTATCGATGATGGAGGGCTGTGCAATAGAGGTTCATATGCTGGCATTGGACTTGTCTTTAGGC</td></tr>
<tr><td>D03</td><td>REV</td><td>caggccaagatgaacgagtgggagcgggagagcgaagaaaccaaattgcacggccccgacaatgatgactacatcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGCCAAGATGAACGAGTGGGAGCGGGAGAGCGAAGAAACCAAATTGCACGGCCCCGACAATGATGACTACATCGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>D04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaatcaccatcaccatcatcatcgacatcgaatccagtcttccagctgaagaatcttcttcagcatcgacattg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAATCACCATCACCATCATCATCGACATCGAATCCAGTCTTCCAGCTGAAGAATCTTCTTCAGCATCGACATTG</td></tr>
<tr><td>D04</td><td>REV</td><td>taccacagcgataacgaggactctcaatccgccgccagccccaagccagtcgaagaaaccatgtggcgcccttggGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACCACAGCGATAACGAGGACTCTCAATCCGCCGCCAGCCCCAAGCCAGTCGAAGAAACCATGTGGCGCCCTTGGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D05</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatcctgcgggtagccaacagatcagcaatctcaggtttatttttatacgttgtgttagtgtttagtatatcta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATCCTGCGGGTAGCCAACAGATCAGCAATCTCAGGTTTATTTTTATACGTTGTGTTAGTGTTTAGTATATCTA</td></tr>
<tr><td>D05</td><td>REV</td><td>tcgcacttgagcgatcttgcagttccacttctggaggtgcaaagccagtcccagattccgccaactagcttggccGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGCACTTGAGCGATCTTGCAGTTCCACTTCTGGAGGTGCAAAGCCAGTCCCAGATTCCGCCAACTAGCTTGGCCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D06</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagtggcttaggtcctagttggacggatgtaacgataagcattagtttagttaataaagtaattgatttcccata
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGTGGCTTAGGTCCTAGTTGGACGGATGTAACGATAAGCATTAGTTTAGTTAATAAAGTAATTGATTTCCCATA</td></tr>
<tr><td>D06</td><td>REV</td><td>tgggcctaccgcggagccacctgtctgctgaaggagaacctggcccagatccagaagagcctggccccgaaggccGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGGCCTACCGCGGAGCCACCTGTCTGCTGAAGGAGAACCTGGCCCAGATCCAGAAGAGCCTGGCCCCGAAGGCCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D07</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggtgtaatttaatataataaacatttaataaattaatatcttattttaattggcgttgagtgaattttctagt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGTGTAATTTAATATAATAAACATTTAATAAATTAATATCTTATTTTAATTGGCGTTGAGTGAATTTTCTAGT</td></tr>
<tr><td>D07</td><td>REV</td><td>cactcgggataccaggagttccgtcgtccctgcaacctcacctcagatagctacaagtcgctagcctatcgagatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTCGGGATACCAGGAGTTCCGTCGTCCCTGCAACCTCACCTCAGATAGCTACAAGTCGCTAGCCTATCGAGATGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D08</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaaggggtcttactaaaagtcccaaacaaacaaatattgtacaaactgtaaataccctaaattgttgccttagt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAAGGGGTCTTACTAAAAGTCCCAAACAAACAAATATTGTACAAACTGTAAATACCCTAAATTGTTGCCTTAGT</td></tr>
<tr><td>D08</td><td>REV</td><td>tcctttgatagcttcagtgacgagcagccagatgacgaggagctactcgattatatttcatcttggcaagagcagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCTTTGATAGCTTCAGTGACGAGCAGCCAGATGACGAGGAGCTACTCGATTATATTTCATCTTGGCAAGAGCAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (EGFP tagging)</th></tr>
<tr><td>D09</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaatatgtggaaaatctaagtaaatcaaacacttaaatccatctatccaaaagttgagctttgagattaaaca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAATATGTGGAAAATCTAAGTAAATCAAACACTTAAATCCATCTATCCAAAAGTTGAGCTTTGAGATTAAACA</td></tr>
<tr><td>D09</td><td>REV</td><td>gagcagcgcaggtttctggacgatgctattacgcgactggagctcttggccgtcaagaagggctcgaacaaaaacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGCAGCGCAGGTTTCTGGACGATGCTATTACGCGACTGGAGCTCTTGGCCGTCAAGAAGGGCTCGAACAAAAACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D10</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatttagttgcttaatgagtaagctcgtttatttaaagccaaagttcacttaatatatatatacatatatatat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATTTAGTTGCTTAATGAGTAAGCTCGTTTATTTAAAGCCAAAGTTCACTTAATATATATATACATATATATAT</td></tr>
<tr><td>D10</td><td>REV</td><td>catctgtcccagaatcgcaatgtttacaatgccaaaggaccggaggagcagcccaatcaggccatcgatgagcgtGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CATCTGTCCCAGAATCGCAATGTTTACAATGCCAAAGGACCGGAGGAGCAGCCCAATCAGGCCATCGATGAGCGTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagccgtagatgtggtcaagggacatacttaaggagtggctatgcaggcgcacggacgcaggacgcgggacgaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGCCGTAGATGTGGTCAAGGGACATACTTAAGGAGTGGCTATGCAGGCGCACGGACGCAGGACGCGGGACGAA</td></tr>
<tr><td>D11</td><td>REV</td><td>attgaggagagcaacgctggattgggcggcatgggcgtgggcctgggcgtccgcggctgttccggcctgaagggcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATTGAGGAGAGCAACGCTGGATTGGGCGGCATGGGCGTGGGCCTGGGCGTCCGCGGCTGTTCCGGCCTGAAGGGCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagggcatctgatccccaaaaatctggaggaatgaagaaaacaaagtgatataacagcggcgacgcagagcggca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGGCATCTGATCCCCAAAAATCTGGAGGAATGAAGAAAACAAAGTGATATAACAGCGGCGACGCAGAGCGGCA</td></tr>
<tr><td>D12</td><td>REV</td><td>cgccagccgtcgccggaaacgaccaccaagatcaagagcgccgccgtgcagcagaagaccgtgtggcggccctacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGCCAGCCGTCGCCGGAAACGACCACCAAGATCAAGAGCGCCGCCGTGCAGCAGAAGACCGTGTGGCGGCCCTACGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

E.4. Tagging with *T2A-EGFP*

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>A01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagattgaaggaccagcgaaacccggtatccggtttcttttgagacattctaaaagagattcagagagctggacg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATTGAAGGACCAGCGAAACCCGGTATCCGGTTTCTTTTGAGACATTCTAAAAGAGATTCAGAGAGCTGGACG</td></tr>
<tr><td>A01</td><td>REV</td><td>ctatcctcgtcatcctcgagctcagagaagcccaagcggatgctgggcttcaccatcgatgagatcatgagcagaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTATCCTCGTCATCCTCGAGCTCAGAGAAGCCCAAGCGGATGCTGGGCTTCACCATCGATGAGATCATGAGCAGAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggcagcgattgcggggtcaccaggggggacaaagtgtggttcgcttgaaaaaaacgtaacttagtttaaagttc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGCAGCGATTGCGGGGTCACCAGGGGGGACAAAGTGTGGTTCGCTTGAAAAAAACGTAACTTAGTTTAAAGTTC</td></tr>
<tr><td>A02</td><td>REV</td><td>cctctggccaggcacagcaccaatccgttcatctcaccgcccaaggcgccggcgcagtcattccaggtgcagctcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCTCTGGCCAGGCACAGCACCAATCCGTTCATCTCACCGCCCAAGGCGCCGGCGCAGTCATTCCAGGTGCAGCTCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaggggaatagaactatggttataatttagttattcacagtatattattattagttaggtatttattgtttag
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGGGGAATAGAACTATGGTTATAATTTAGTTATTCACAGTATATTATTATTAGTTAGGTATTTATTGTTTAG</td></tr>
<tr><td>A03</td><td>REV</td><td>tggggcatcgatgcccgggaacttttgtgggatcgaatgctggaaataatgcgaaatcatgagaattcaattatcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGGGCATCGATGCCCGGGAACTTTTGTGGGATCGAATGCTGGAAATAATGCGAAATCATGAGAATTCAATTATCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggattctcgacatatgaatcccttagcaggacacaattacctaagcaaactcacattccgctggcgaagacca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGATTCTCGACATATGAATCCCTTAGCAGGACACAATTACCTAAGCAAACTCACATTCCGCTGGCGAAGACCA</td></tr>
<tr><td>A04</td><td>REV</td><td>gtgtgccacaaatctttctccaggatgtcgctcctgaacaagcactccagctccaactgcaccatcactattgcgGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGTGCCACAAATCTTTCTCCAGGATGTCGCTCCTGAACAAGCACTCCAGCTCCAACTGCACCATCACTATTGCGGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

Well	Read	Alignment (*T2A-EGFP* tagging)
A05	FWD	TGACTACAAGGATGACGACGACAAGtaggttggaaatatagaaattttaactaatttatacttaaaagattaaaaaaaaaaaaaatagtaaaaccacaaaa |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| TGACTACAAGGATGACGACGACAAGTAGGTTGGAAATATAGAAATTTTAACTAATTTATACTTAAAAGATTAAAAAAAAAAAAAATAGTAAAACCACAAAA
A05	REV	gtgtggaaatttgtgtggcgcaagttatttggcaaatcgaacaaagtttccaagggaaagaaggtgaagaagcagGAAGTGCATACCAATCAGGACCCGC |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| GTGTGGAAATTTGTGTGGCGCAAGTTATTTGGCAAATCGAACAAAGTTTCCAAGGGAAAGAAGGTGAAGAAGCAGGAAGTGCATACCAATCAGGACCCGC
A06	FWD	TGACTACAAGGATGACGACGACAAGtaaacttgtttagagaatgtaaataagcaattaaacagtgcattctagccatagggcattctaccatttttaaat |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| TGACTACAAGGATGACGACGACAAGTAAACTTGTTTAGAGAATGTAAATAAGCAATTAAACAGTGCATTCTAGCCATAGGGCATTCTACCATTTTTAAAT
A06	REV	agatcacattcccagcgacagcaacacaacggagccaacattcgcaagagtcatcatcatcagtttcgtcaaaagGAAGTGCATACCAATCAGGACCCGC |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| AGATCACATTCCCAGCGACAGCAACACAACGGAGCCAACATTCGCAAGAGTCATCATCATCAGTTTCGTCAAAAGGAAGTGCATACCAATCAGGACCCGC
A07	FWD	TGACTACAAGGATGACGACGACAAGtagatcttaactagctagtaaattacctgtgcgtagtatttaacgatcttgttctctggaaatttcttctaaatt |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| TGACTACAAGGATGACGACGACAAGTAGATCTTAACTAGCTAGTAAATTACCTGTGCGTAGTATTTAACGATCTTGTTCTCTGGAAATTTCTTCTAAATT
A07	REV	tgcaacggatcctcctccctggcccccatcgccggagccatcctgctcttcttcggcgtggctcgtctgctggccGAAGTGCATACCAATCAGGACCCGC |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| TGCAACGGATCCTCCTCCCTGGCCCCCATCGCCGGAGCCATCCTGCTCTTCTTCGGCGTGGCTCGTCTGCTGGCCGAAGTGCATACCAATCAGGACCCGC
A08	FWD	TGACTACAAGGATGACGACGACAAGtaatcgtcacacatttccctcagattaagcactttaaattgtaatcattacatcaataaataaatgcggagaacc |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| TGACTACAAGGATGACGACGACAAGTAATCGTCACACATTTCCCTCAGATTAAGCACTTTAAATTGTAATCATTACATCAATAAATAAATGCGGAGAACC
A08	REV	gtggccgatcatccatttgcgttcatcattcgcgacaagcacgctgtctatttcaccggacacattgtcaagtttGAAGTGCATACCAATCAGGACCCGC |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| GTGGCCGATCATCCATTTGCGTTCATCATTCGCGACAAGCACGCTGTCTATTTCACCGGACACATTGTCAAGTTTGAAGTGCATACCAATCAGGACCCGC
A09	FWD	TGACTACAAGGATGACGACGACAAGtgatatcagcgggtctgaggtgtccacctgtaaccccacccagactaaatcaacaaccccaaaccgaaatcccaa |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| TGACTACAAGGATGACGACGACAAGTGATATCAGCGGGTCTGAGGTGTCCACCTGTAACCCCACCCAGACTAAATCAACAACCCCAAACCGAAATCCCAA
A09	REV	cacagtggccaactgatgctgcatcggcccttctccacgtcgccggagctgaagcacagtgctcccgagatcacaGAAGTGCATACCAATCAGGACCCGC |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| CACAGTGGCCAACTGATGCTGCATCGGCCCTTCTCCACGTCGCCGGAGCTGAAGCACAGTGCTCCCGAGATCACAGAAGTGCATACCAATCAGGACCCGC

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>A10</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagacgatccgcacctgattcgctttcccgattcgccctgacccgattccatgacctcgcccacttctagatata
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGACGATCCGCACCTGATTCGCTTTCCCGATTCGCCCTGACCCGATTCCATGACCTCGCCCACTTCTAGATATA</td></tr>
<tr><td>A10</td><td>REV</td><td>aacttcctcacgcccaacatgagcatcgtggacgtgaacatccgccacgagaaccgcaccgtgcagcgcccaaacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AACTTCCTCACGCCCAACATGAGCATCGTGGACGTGAACATCCGCCACGAGAACCGCACCGTGCAGCGCCCAAACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaggactactctagtgtattttagtgttacgctttattattaatgcaattggtaattaatatattcttaggct
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGGACTACTCTAGTGTATTTTAGTGTTACGCTTTATTATTAATGCAATTGGTAATTAATATATTCTTAGGCT</td></tr>
<tr><td>A11</td><td>REV</td><td>tacataatccatcgaaaacgcgaggcggactttaagagcccacgtggcggatacttgttcgacaatatctttggcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACATAATCCATCGAAAACGCGAGGCGGACTTTAAGAGCCCACGTGGCGGATACTTGTTCGACAATATCTTTGGCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>A12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagccaatgggcccacggccgtcgactccaagcgcttcaggtccaatccatcaaccagccctcgaatgcataaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGCCAATGGGCCCACGGCCGTCGACTCCAAGCGCTTCAGGTCCAATCCATCAACCAGCCCTCGAATGCATAAA</td></tr>
<tr><td>A12</td><td>REV</td><td>atcgccggagccgaggatgtgtccgccacttcgttcgctcttgtgggcatcctggcggcgttgctcttcgccagaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATCGCCGGAGCCGAGGATGTGTCCGCCACTTCGTTCGCTCTTGTGGGCATCCTGGCGGCGTTGCTCTTCGCCAGAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagcgccaaaggatggccaggatgtccacacccttttctacacttatgctaagtgaacacacccatatatattt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGCGCCAAAGGATGGCCAGGATGTCCACACCCTTTTCTACACTTATGCTAAGTGAACACACCCATATATATTT</td></tr>
<tr><td>B01</td><td>REV</td><td>caggccggttccaacaagggagccacccaggctggccagaacctcggcgctggccgcaagatcctgctcggcaagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGCCGGTTCCAACAAGGGAGCCACCCAGGCTGGCCAGAACCTCGGCGCTGGCCGCAAGATCCTGCTCGGCAAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaggagccgttcacccagctaaccggcttatccctggcagcagaagctgtcgcatccctacacctgcactgaga
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGGAGCCGTTCACCCAGCTAACCGGCTTATCCCTGGCAGCAGAAGCTGTCGCATCCCTACACCTGCACTGAGA</td></tr>
<tr><td>B02</td><td>REV</td><td>ggcgacaaccaggagctgcgcacgaacaccattgagaacatgctgatggccctgcccagcgcctccaaggccaagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCGACAACCAGGAGCTGCGCACGAACACCATTGAGAACATGCTGATGGCCCTGCCCAGCGCCTCCAAGGCCAAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>B03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagatccatcaagatcctaaaaatatagcctctcgtattgttgcataatactaagaatctttattactagatatt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATCCATCAAGATCCTAAAAATATAGCCTCTCGTATTGTTGCATAATACTAAGAATCTTTATTACTAGATATT</td></tr>
<tr><td>B03</td><td>REV</td><td>gacaacctgaagaagaccaccgccaaggtcaccttctggagcaaatacggcgtgaggacgaagcagaacgagcagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GACAACCTGAAGAAGACCACCGCCAAGGTCACCTTCTGGAGCAAATACGGCGTGAGGACGAAGCAGAACGAGCAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagtgtagactcaaagttccttcacgaactcaactcaatcccatttttgccatgacacctcagctactcttaat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGTGTAGACTCAAAGTTCCTTCACGAACTCAACTCAATCCCATTTTTGCCATGACACCTCAGCTACTCTTAAT</td></tr>
<tr><td>B04</td><td>REV</td><td>cgtgagttcctggccacgtttccgccgtgggcgcatgttagctcctggctgggatcctacaatagttggcagcttGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTGAGTTCCTGGCCACGTTTCCGCCGTGGGCGCATGTTAGCTCCTGGCTGGGATCCTACAATAGTTGGCAGCTTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B05</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaagaagttagggatttaatgcttggcaaattgtgattcgggaaaaaatgtaacaaaatttaaataaattcttt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGAAGTTAGGGATTTAATGCTTGGCAAATTGTGATTCGGGAAAAAATGTAACAAAATTTAAATAAATTCTTT</td></tr>
<tr><td>B05</td><td>REV</td><td>gagaacaagtgtgataccgcctccaagttgtacgattgcttcgagagcttcaagcccgcccccgaggctaaggccGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGAACAAGTGTGATACCGCCTCCAAGTTGTACGATTGCTTCGAGAGCTTCAAGCCCGCCCCCGAGGCTAAGGCCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B06</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaggggtgccacccatccagaccgagatgtgtacatacgtatttccggactactcagctatcgaggctatcgga
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGGGGTGCCACCCATCCAGACCGAGATGTGTACATACGTATTTCCGGACTACTCAGCTATCGAGGCTATCGGA</td></tr>
<tr><td>B06</td><td>REV</td><td>ggagctgcgtcatcgatgcagcgcctgaacgtgggagtgatcctgctggcagcgctcctgctgcgagtccgcctcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGAGCTGCGTCATCGATGCAGCGCCTGAACGTGGGAGTGATCCTGCTGGCAGCGCTCCTGCTGCGAGTCCGCCTCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B07</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagtgtcccgggcatcggcaaccgcataatccgagagtatcccatctgtccgatccgatccaagtcgatccgagg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGTGTCCCGGGCATCGGCAACCGCATAATCCGAGAGTATCCCATCTGTCCGATCCGATCCAAGTCGATCCGAGG</td></tr>
<tr><td>B07</td><td>REV</td><td>gagtcggtacaggagctcgtccgtcacctgtccggccaccacaataacctgctgctgacaaagaatctgcgcgaaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGTCGGTACAGGAGCTCGTCCGTCACCTGTCCGGCCACCACAATAACCTGCTGCTGACAAAGAATCTGCGCGAAGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>B08</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgatttaatgaatcgccgacgtcgctgtacgaccgagaatcattacattttcgcgttagttttatgcatttcaat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGATTTAATGAATCGCCGACGTCGCTGTACGACCGAGAATCATTACATTTTCGCGTTAGTTTTATGCATTTCAAT</td></tr>
<tr><td>B08</td><td>REV</td><td>gcagcggctcacgggaacccggcctccgcctacagccaccccctgccgacgcagggtcaggccaagtactggtcaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCAGCGGCTCACGGGAACCCGGCCTCCGCCTACAGCCACCCCCTGCCGACGCAGGGTCAGGCCAAGTACTGGTCAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B09</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggtgcctgctaccggtgcatcacgtactcatagtcattcactgatccagctttgttttagcaccttaagttgg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGTGCCTGCTACCGGTGCATCACGTACTCATAGTCATTCACTGATCCAGCTTTGTTTTAGCACCTTAAGTTGG</td></tr>
<tr><td>B09</td><td>REV</td><td>tggtaccagaccgcccgcctccaggacgaggccacgacggcagcgcagcctgctgcgaacggcgtaaagcaggacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGTACCAGACCGCCCGCCTCCAGGACGAGGCCACGACGGCAGCGCAGCCTGCTGCGAACGGCGTAAAGCAGGACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B10</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagaggctaactaaagtgatcaataccgaaagaacaacaaagaaacgaggtggaaactaaggcatatccttgta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGAGGCTAACTAAAGTGATCAATACCGAAAGAACAACAAAGAAACGAGGTGGAAACTAAGGCATATCCTTGTA</td></tr>
<tr><td>B10</td><td>REV</td><td>tccgttctgtttgccgtggagattccgaagcccatctatcgcttcttcaagggcatctttggcggtttctccaacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCGTTCTGTTTGCCGTGGAGATTCCGAAGCCCATCTATCGCTTCTTCAAGGGCATCTTTGGCGGTTTCTCCAACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaagggaaaacttaagattaaatttagtttaataatttaataaaaactgtactgataatgtctaaaaagaatat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAAGGGAAAACTTAAGATTAAATTTAGTTTAATAATTTAATAAAAACTGTACTGATAATGTCTAAAAAGAATAT</td></tr>
<tr><td>B11</td><td>REV</td><td>gctttggtcgaactgaaggagaagtatgccacgctcatccaaccgcgtaactcaaaccaatatgcagtcatcattGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCTTTGGTCGAACTGAAGGAGAAGTATGCCACGCTCATCCAACCGCGTAACTCAAACCAATATGCAGTCATCATTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>B12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggacgtatgtatgcgtggcgtcttcatttcacttggcatttaatatttgtaggctatagtcttgtattgtact
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGACGTATGTATGCGTGGCGTCTTCATTTCACTTGGCATTTAATATTTGTAGGCTATAGTCTTGTATTGTACT</td></tr>
<tr><td>B12</td><td>REV</td><td>gcgaaccagacggtgaacgtcaactactacggcgcctgtggccgccccgaagcaccatccactaacttcctttacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCGAACCAGACGGTGAACGTCAACTACTACGGCGCCTGTGGCCGCCCCGAAGCACCATCCACTAACTTCCTTTACGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>C01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaactatcaagtgcagacgggtgcatgtgctcccgtttggcttggaatcggtgccttgtacatttaattagcg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAACTATCAAGTGCAGACGGGTGCATGTGCTCCCGTTTGGCTTGGAATCGGTGCCTTGTACATTTAATTAGCG</td></tr>
<tr><td>C01</td><td>REV</td><td>agcgatccgacgatggccaagccgaagcgcagcagcttcagcatctcggacatattaggaaccagctcgtccattGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCGATCCGACGATGGCCAAGCCGAAGCGCAGCAGCTTCAGCATCTCGGACATATTAGGAACCAGCTCGTCCATTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagatatacattattaacgaaaacaaaaccacgggttttcaattggattcatgttttaatgtactaacaacaatg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATATACATTATTAACGAAAACAAAACCACGGGTTTTCAATTGGATTCATGTTTTAATGTACTAACAACAATG</td></tr>
<tr><td>C02</td><td>REV</td><td>cgtgctctttcgcaaaacttaattcgaaattttgaacttcaaaaactaagaagagccaataaagtgcaaaaatatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGTGCTCTTTCGCAAAACTTAATTCGAAATTTTGAACTTCAAAAACTAAGAAGAGCCAATAAAGTGCAAAAATATGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgatctgaggacatgaatttatactataggccatattaataataactccgtcgaaatcgaaattgaaacgaacta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGATCTGAGGACATGAATTTATACTATAGGCCATATTAATAATAACTCCGTCGAAATCGAAATTGAAACGAACTA</td></tr>
<tr><td>C03</td><td>REV</td><td>gaggtgcagaagcaggtcgcccaactgacgcccattgtgaagcgcagcatacgcgactacttcaacaaggagtacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGGTGCAGAAGCAGGTCGCCCAACTGACGCCCATTGTGAAGCGCAGCATACGCGACTACTTCAACAAGGAGTACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaggtaaaagtgacgagaatacgaatacgaatacctagccaagcaactgagctctgtgatattttcatgttca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAGGTAAAAGTGACGAGAATACGAATACGAATACCTAGCCAAGCAACTGAGCTCTGTGATATTTTCATGTTCA</td></tr>
<tr><td>C04</td><td>REV</td><td>agctcgtacaaaaccgtcgacgagaaaaagaaatcgaagtcaaaagacagccagtcgaaagacgatatcaagcggGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGCTCGTACAAAACCGTCGACGAGAAAAAGAAATCGAAGTCAAAAGACAGCCAGTCGAAAGACGATATCAAGCGGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C05</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagttagttactattgccggacaacgcgttgttgttgccaagaagaatcaggcaactgcatttttatacagggt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGTTAGTTACTATTGCCGGACAACGCGTTGTTGTTGCCAAGAAGAATCAGGCAACTGCATTTTTATACAGGGT</td></tr>
<tr><td>C05</td><td>REV</td><td>ctgccgagggacttcgagcactcattccagacgatgcacgagtgcaaatatcaaacgtatccgtctacatccaatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CTGCCGAGGGACTTCGAGCACTCATTCCAGACGATGCACGAGTGCAAATATCAAACGTATCCGTCTACATCCAATGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>C06</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaagagaaaagagtattttacgattgactttgttgattagcggaattgattttgaagaaaattgcattttgattt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAGAGAAAAGAGTATTTTACGATTGACTTTGTTGATTAGCGGAATTGATTTTGAAGAAAATTGCATTTTGATTT</td></tr>
<tr><td>C06</td><td>REV</td><td>agagttccgctgtatgcggagtccttcaaaaaggcatccgaacatggcttcaagccgcagatcatcaaggaaacaGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AGAGTTCCGCTGTATGCGGAGTCCTTCAAAAAGGCATCCGAACATGGCTTCAAGCCGCAGATCATCAAGGAAACAGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C07</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaaacttggcggggatcaaaggttaagtgaagcacaatgagtctaagcgacaaacgtattattctcgtttaaga
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAAACTTGGCGGGGATCAAAGGTTAAGTGAAGCACAATGAGTCTAAGCGACAAACGTATTATTCTCGTTTAAGA</td></tr>
<tr><td>C07</td><td>REV</td><td>cacgcccttggttacgcctcttgcctcagtgatcgtaacatgtgcgtggatgggggcgtggcacggagaccacgcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACGCCCTTGGTTACGCCTCTTGCCTCAGTGATCGTAACATGTGCGTGGATGGGGGCGTGGCACGGAGACCACGCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C08</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatgactttgcgcgctggtcgttccacaactctgatttctactgtacatacaaatatttgtattcaaatcctac
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATGACTTTGCGCGCTGGTCGTTCCACAACTCTGATTTCTACTGTACATACAAATATTTGTATTCAAATCCTAC</td></tr>
<tr><td>C08</td><td>REV</td><td>gccgctgttaaaattcaggctggcttccggggattcaaaacacgcaaagaattgaaacaatgcgagcccattgtgGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GCCGCTGTTAAAATTCAGGCTGGCTTCCGGGGATTCAAAACACGCAAAGAATTGAAACAATGCGAGCCCATTGTGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C09</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagcggtgctcgtccccatctcatgcatattgatatataaagcagatatttatatttactcttaacgatttgtc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGCGGTGCTCGTCCCCATCTCATGCATATTGATATATAAAGCAGATATTTATATTTACTCTTAACGATTTGTC</td></tr>
<tr><td>C09</td><td>REV</td><td>cagcaacagcagcgaccccagctgatcccgcccggagctggctatcagccacagggcgatttcgatgtcttcttcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGCAACAGCAGCGACCCCAGCTGATCCCGCCCGGAGCTGGCTATCAGCCACAGGGCGATTTCGATGTCTTCTTCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C10</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagtcgtcggcctaaggaccattgcgacttgcatccatcgctgtagccataaatcatgcatcatcatccgtatat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGTCGTCGGCCTAAGGACCATTGCGACTTGCATCCATCGCTGTAGCCATAAATCATGCATCATCATCCGTATAT</td></tr>
<tr><td>C10</td><td>REV</td><td>gtggagtatcagcagacacccgcctcacaggagctgcacttccgagataccccattctgaacgcgaggaccgttGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GTGGAGTATCAGCAGACACCCGCCTCACAGGAGCTGCACTTCCGAGATACCCCATTCTGAACGCGAGGACCGTTGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>C11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagatgatttgtttcggttttggctgaccaggatgacaatgcaagaccagggataacggcgagctggtagcgagt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGATGATTTGTTTCGGTTTTGGCTGACCAGGATGACAATGCAAGACCAGGGATAACGGCGAGCTGGTAGCGAGT</td></tr>
<tr><td>C11</td><td>REV</td><td>cccgtccaggcacccgtgccggtggtctactcccactcgcacacccagcagcccgtctggttggagaaggagtggGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CCCGTCCAGGCACCCGTGCCGGTGGTCTACTCCCACTCGCACACCCAGCAGCCCGTCTGGTTGGAGAAGGAGTGGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>C12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaccgaacccgggtattgtgaactcacctcttcccaccccgttgtgatatatgatacatatatgtaatacata
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAACCGAACCCGGGTATTGTGAACTCACCTCTTCCCACCCCGTTGTGATATATGATACATATATGTAATACATA</td></tr>
<tr><td>C12</td><td>REV</td><td>tcactcacgtgcaaaatgcagactgtgatgggcgctgagacgcagaaaatgctgaagaacagcgaggattatgttGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCACTCACGTGCAAAATGCAGACTGTGATGGGCGCTGAGACGCAGAAAATGCTGAAGAACAGCGAGGATTATGTTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D01</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagaaccatcaggatagccatcgtgattcgctccttgaagaaggtttaattaaaaaagcattacagaataaaaag
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGAACCATCAGGATAGCCATCGTGATTCGCTCCTTGAAGAAGGTTTAATTAAAAAAGCATTACAGAATAAAAAG</td></tr>
<tr><td>D01</td><td>REV</td><td>aagttctatgtgggaaacggatatccgttcacgccattcagctttaaggatattttgatcgtcgtcgaagatgatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AAGTTCTATGTGGGAAACGGATATCCGTTCACGCCATTCAGCTTTAAGGATATTTTGATCGTCGTCGAAGATGATGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D02</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatctggccaccaactgatcagctctctgtgaaataataaatattaaatatgtactagttctcataaaagttat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATCTGGCCACCAACTGATCAGCTCTCTGTGAAATAATAAATATTAAATATGTACTAGTTCTCATAAAAGTTAT</td></tr>
<tr><td>D02</td><td>REV</td><td>ggcggtcccgtggccggcaacttcttccaggcccaatacgatgactacgtgaagaccctcatcgagacggtccagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GGCGGTCCCGTGGCCGGCAACTTCTTCCAGGCCCAATACGATGACTACGTGAAGACCCTCATCGAGACGGTCCAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D03</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaatccctagactatcgatgatggagggctgtgcaatagaggttcatatgctggcattggacttgtctttaggc
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAATCCCTAGACTATCGATGATGGAGGGCTGTGCAATAGAGGTTCATATGCTGGCATTGGACTTGTCTTTAGGC</td></tr>
<tr><td>D03</td><td>REV</td><td>caggccaagatgaacgagtgggagcgggagagcgaagaaaccaaattgcacggccccgacaatgatgactacatcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CAGGCCAAGATGAACGAGTGGGAGCGGGAGAGCGAAGAAACCAAATTGCACGGCCCCGACAATGATGACTACATCGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>D04</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaatcaccatcaccatcatcatcgacatcgaatccagtcttccagctgaagaatcttcttcagcatcgacattg
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAATCACCATCACCATCATCATCGACATCGAATCCAGTCTTCCAGCTGAAGAATCTTCTTCAGCATCGACATTG</td></tr>
<tr><td>D04</td><td>REV</td><td>taccacagcgataacgaggactctcaatccgccgccagccccaagccagtcgaagaaaccatgtggcgcccttggGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TACCACAGCGATAACGAGGACTCTCAATCCGCCGCCAGCCCCAAGCCAGTCGAAGAAACCATGTGGCGCCCTTGGGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D05</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatcctgcgggtagccaacagatcagcaatctcaggtttatttttatacgttgtgttagtgtttagtatatcta
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATCCTGCGGGTAGCCAACAGATCAGCAATCTCAGGTTTATTTTTATACGTTGTGTTAGTGTTTAGTATATCTA</td></tr>
<tr><td>D05</td><td>REV</td><td>tcgcacttgagcgatcttgcagttccacttctggaggtgcaaagccagtcccagattccgccaactagcttggccGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCGCACTTGAGCGATCTTGCAGTTCCACTTCTGGAGGTGCAAAGCCAGTCCCAGATTCCGCCAACTAGCTTGGCCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D06</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagtggcttaggtcctagttggacggatgtaacgataagcattagtttagttaataaagtaattgatttcccata
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGTGGCTTAGGTCCTAGTTGGACGGATGTAACGATAAGCATTAGTTTAGTTAATAAAGTAATTGATTTCCCATA</td></tr>
<tr><td>D06</td><td>REV</td><td>tgggcctaccgcggagccacctgtctgctgaaggagaacctggcccagatccagaagagcctggccccgaaggccGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGGGCCTACCGCGGAGCCACCTGTCTGCTGAAGGAGAACCTGGCCCAGATCCAGAAGAGCCTGGCCCCGAAGGCCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D07</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaggtgtaatttaatataataaacatttaataaattaatatcttattttaattggcgttgagtgaattttctagt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGTGTAATTTAATATAATAAACATTTAATAAATTAATATCTTATTTTAATTGGCGTTGAGTGAATTTTCTAGT</td></tr>
<tr><td>D07</td><td>REV</td><td>cactcgggataccaggagttccgtcgtccctgcaacctcacctcagatagctacaagtcgctagcctatcgagatGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CACTCGGGATACCAGGAGTTCCGTCGTCCCTGCAACCTCACCTCAGATAGCTACAAGTCGCTAGCCTATCGAGATGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D08</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgaaggggtcttactaaaagtcccaaacaaacaaatattgtacaaactgtaaataccctaaattgttgccttagt
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAAGGGGTCTTACTAAAAGTCCCAAACAAACAAATATTGTACAAACTGTAAATACCCTAAATTGTTGCCTTAGT</td></tr>
<tr><td>D08</td><td>REV</td><td>tcctttgatagcttcagtgacgagcagccagatgacgaggagctactcgattatatttcatcttggcaagagcagGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TCCTTTGATAGCTTCAGTGACGAGCAGCCAGATGACGAGGAGCTACTCGATTATATTTCATCTTGGCAAGAGCAGGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Continued on the next page

<table>
<tr><th>Well</th><th>Read</th><th>Alignment (T2A-EGFP tagging)</th></tr>
<tr><td>D09</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaaaatatgtggaaaatctaagtaaatcaaacacttaaatccatctatccaaaagttgagctttgagattaaaca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAAAATATGTGGAAAATCTAAGTAAATCAAACACTTAAATCCATCTATCCAAAAGTTGAGCTTTGAGATTAAACA</td></tr>
<tr><td>D09</td><td>REV</td><td>gagcagcgcaggtttctggacgatgctattacgcgactggagctcttggccgtcaagaagggctcgaacaaaaacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
GAGCAGCGCAGGTTTCTGGACGATGCTATTACGCGACTGGAGCTCTTGGCCGTCAAGAAGGGCTCGAACAAAAACGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D10</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtaatttagttgcttaatgagtaagctcgtttatttaaagccaaagttcacttaatatatatatacatatatatat
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAATTTAGTTGCTTAATGAGTAAGCTCGTTTATTTAAAGCCAAAGTTCACTTAATATATATATACATATATATAT</td></tr>
<tr><td>D10</td><td>REV</td><td>catctgtcccagaatcgcaatgtttacaatgccaaaggaccggaggagcagcccaatcaggccatcgatgagcgtGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CATCTGTCCCAGAATCGCAATGTTTACAATGCCAAAGGACCGGAGGAGCAGCCCAATCAGGCCATCGATGAGCGTGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D11</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtgagccgtagatgtggtcaagggacatacttaaggagtggctatgcaggcgcacggacgcaggacgcgggacgaa
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTGAGCCGTAGATGTGGTCAAGGGACATACTTAAGGAGTGGCTATGCAGGCGCACGGACGCAGGACGCGGGACGAA</td></tr>
<tr><td>D11</td><td>REV</td><td>attgaggagagcaacgctggattgggcggcatgggcgtgggcctgggcgtccgcggctgttccggcctgaagggcGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
ATTGAGGAGAGCAACGCTGGATTGGGCGGCATGGGCGTGGGCCTGGGCGTCCGCGGCTGTTCCGGCCTGAAGGGCGAAGTGCATACCAATCAGGACCCGC</td></tr>
<tr><td>D12</td><td>FWD</td><td>TGACTACAAGGATGACGACGACAAGtagggcatctgatccccaaaaatctggaggaatgaagaaaacaaagtgatataacagcggcgacgcagagcggca
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TGACTACAAGGATGACGACGACAAGTAGGGCATCTGATCCCCAAAAATCTGGAGGAATGAAGAAAACAAAGTGATATAACAGCGGCGACGCAGAGCGGCA</td></tr>
<tr><td>D12</td><td>REV</td><td>cgccagccgtcgccggaaacgaccaccaagatcaagagcgccgccgtgcagcagaagaccgtgtggcggccctacGAAGTGCATACCAATCAGGACCCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
CGCCAGCCGTCGCCGGAAACGACCACCAAGATCAAGAGCGCCGCCGTGCAGCAGAAGACCGTGTGGCGGCCCTACGAAGTGCATACCAATCAGGACCCGC</td></tr>
</table>

Appendix F.

Contents of the attached DVD

Ejsmont RK Thesis.pdf

PDF version of this document.

Movies

Movie 1 – SPIM in Action.mov

Agarose bar with two Drosophila embryos mounted in the SPIM imaging chamber and imaged in bright field. Movie courtesy of S. Preibisch.

Movie 2 – Histone EYFP Spinning Disk.mov

Early Histone-EYFP embryo imaged with spinning disk microscope, showing waves of cell divisions.

Movie 3 – Histone EYFP SPIM multiview.mov

3D reconstruction of Histone-EYFP embryo whole-development SPIM time-lapse. Lateral and dorsal views are shown. Movie courtesy of S. Preibisch.

Movie 4 – Histone EYFP SPIM lateral.mov

3D volume rendering of Histone-EYFP embryo whole-development SPIM time-lapse. Lateral view. Movie courtesy of M. Weber.

Movie 5 – Histone EYFP SPIM dorsal.mov

3D volume rendering of Histone-EYFP embryo whole-development SPIM time-lapse. Dorsal view. Movie courtesy of M. Weber.

Movie 6 – Cadherin ECFP.mov

Cadherin-ECFP embryo imaged with SPIM – z-stack movie.

Movie 7 – CG4702 3D SPIM.mov

3D rendering of CG4702 expression in late *Drosophila* embryo. CG4702 immunostaining is in green, nuclei are in gray. 3D rendering performed by D. White.

FlyFos

dmel-5.4-FlyFos.gff

Mapping data for *Drosophila melanogaster* FlyFos library in GFF format.

dpse-2.3-FlyFos.gff

Mapping data for *Drosophila pseudoobscura* FlyFos library in GFF format.

Sequencing

End-sequencing results for *D. mel.* and *D. pse.* clones in FastA format.

Sequences

pFlyFos.gb

Annotated sequence of pFlyFos vector in GenBank format.

pTag

Annotated sequences of first generation tags in GenBank format.

pTagNG

Annotated sequences of second generation tags in GenBank format.

Ultimate

Annotated sequences of pTagNG[Ultimate] and pUltimate constructs in GenBank format.

Software

Shear

Source code and input data for DNA shearing simulation program. GNU C compiler and GNU Scintific Library are required to compile this program.