Engineering and application of cGAL, a GAL4 bipartite expression system for *Caenorhabditis elegans*

Thesis by

Jonathan C. Liu

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy in Biology

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2019

(Defended 29th of May, 2019)

Jonathan C. Liu
ORCID: 0000-0002-5761-8704

ACKNOWLEDGEMENTS

To my family, without which this would've been impossible. I thank them for their love and support. This journey is primarily mine, but I did much of it to make them proud.

To Paul, and the random stroke of luck which led to someone suggesting that I consider you as an advisor because of your passion, excellence, and collegiality. When I first came to you, I had no knowledge of *C. elegans* or its excellent community, which I have no doubt you had a hand in creating. It has been a long few years, and I doubt I would have made it with any other advisor. You will be my role model for years to come.

To Han, and to an excellent partnership over the past few years. Your mentorship and support are second to none, and I am grateful to have been able to be your colleague. I wish you the best in the future.

To Allison and James, for being some of my very best friends. Your thanks will come in person.

To the Sternberg lab, past and present alike, for making the lab a home instead of a workplace.

To my committee, and their excellent guidance and role in showing me what excellent scientists look like.

To Liz and Raina, for being patient with me and my awful propensity not to be on time, and for making sure I graduate.

Engineering and application of cGAL, a GAL4 bipartite expression system for *Caenorhabditis elegans*

ABSTRACT

The core objectives of genetics are to dissect and understand the function of genes, the consequence of their perturbation on an organism, and how their collective action influences an organism's biology. For genetic model organisms, transgenesis is a tool that allows researchers to introduce synthetic genetic constructs to determine where a gene acts, when it is required, and infer its function. *Caenorhabditis elegans* is a powerful genetic model organism, with a variety of transgenesis methods available to researchers. Each has its own advantages in speed, efficiency, control of copy number, and control of integration site. However, all methods suffer from issues of reproducibility, reusability, and labor cost. Bipartite systems offer solutions to these issues- they separate the promoter element from the gene product producing strains in which one sex contains the promoter ('driver' strain) and the other contains the gene ('effector' strain). Crossing driver and effector strains reunites promoter and gene in the progeny, which are assayed and analyzed for gene function. This separation of drivers from effectors allows for a variety of benefits. Driver and effector strains can be combinatorially reused, meaning less time-consuming strain construction. Reusing strains allows for more reproducibility and consistency between experiments and between laboratories. Additionally, novel genes and promoters can be crossed to existing strains for novel transgenic patterns requiring minimal effort. Thus, bipartite systems greatly increase the rigor and pace of genetic analysis. This thesis details the engineering of cGAL, a GAL4-based bipartite system for *C. elegans*. It uses a novel GAL4 gene from *Saccharomyces cerevisiae*, a yeast whose optimal growth temperature is similar to that of *C. elegans*. This thesis also describes an intein-based split bipartite system that offers more refined spatiotemporal control, by allowing two promoters to dictate gene expression instead of one. This split method is used to analyze rhythmic feeding in *C. elegans*. Finally, engineering of cGAL using single copy methodology is detailed, with a discussion of future improvements to, and usage of, single copy cGAL. This development of a new bipartite system will greatly accelerate genetic analysis for the *C. elegans*, improve reproducibility for the field, and generate a valuable resource for the community.

Thesis Advisor: Paul W. Sternberg, PhD

Committee Chair: David J. Anderson, PhD

Committee Member: Viviana Gradinaru, PhD

Committee Member: Henry A. Lester, PhD

PUBLISHED CONTENT AND CONTRIBUTIONS

Wang H, **Liu J**, Gharib S, Chai CM, Schwarz EM, Pokala N, Sternberg PW. cGAL, a temperature-robust GAL4-UAS system for Caenorhabditis elegans. Nat Methods. 2017 Feb;14(2):145-148. doi: 10.1038/nmeth.4109.

Co-first author | Experimental design, execution, review | Data analysis

Wang H, **Liu J**, Yuet KP, Hill AJ, Sternberg PW. Split cGAL, an intersectional strategy using a split intein for refined spatiotemporal transgene control in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2018 Apr 10;115(15):3900-3905. doi: 10.1073/pnas.1720063115.

Co-first author | Experimental design, execution, review | Data analysis

Wang H, Park H, **Liu J**, Sternberg PW. An Efficient Genome Editing Strategy To Generate Putative Null Mutants in Caenorhabditis elegans Using CRISPR/Cas9. G3(Bethesda). 2018 Nov 6;8(11):3607-3616. doi: 10.1534/g3.118.200662.

Author | Experimental execution, review | Data analysis

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

FIGURES

TABLES

Chapter 1

Transgenesis in *Caenorhabditis elegans* and Bipartite Gene Expression Systems

1.1 INTRODUCTION

A central goal of biology is to understand how the assortment of genes present in each organism dictates its development, growth, cellular and tissue functions, and behavior, and to examine how genes evolve over time to give rise to diverse forms of life. Each gene may act in different cells, and at different times, for various biological processes. To dissect the function of genes, scientists need to study the functional consequences of perturbing genes, and thus require tools capable of controlling gene expression at will. This can be the introduction of wild-type genes back into certain cells of a null mutant to discover site-of-action, conditional expression of a gene to discover critical time windows, or the introduction of heterologous genes that perform a unique function useful to the researcher. This practice of introducing tailored genetic constructs into organisms, transgenesis, is a workhorse of genetic research.

Caenorhabditis elegans arose as a genetic model organism with Sydney Brenner's seminal work in 1974[1], detailing its genetics, molecular and developmental biology, and neurobiology. This small soil roundworm is found in temperate climates and raised in laboratory conditions on small petri plates containing a lawn of *Escherichia coli* as its food source. It has numerous excellent qualities as a model organism: short generation time, large brood size, and optical clarity. The majority of animals are hermaphroditic, allowing Mendelian segregation of genotypes without a need for mating. Males are found naturally at a small fraction (0.1%), but this fraction can be substantially enlarged under laboratory conditions to allow mating. Generated strains can be thawed and recovered later, allowing for long term storage. It also has a rich scientific history: it was the first multicellular

organism to have its genome sequenced[2], to have a full developmental cell lineage map[3,4], and remains the only organism to have a full connectome of cellular connections mapped[5,6]. Concerning transgenesis, it was one of the first organisms to demonstrate expression of green fluorescent protein, sharing that title with *E. coli*[7].

1.2 Transgenesis in *C. elegans*

There are a variety of transformation methods available in *C. elegans* research. The oldest and perhaps most popular is the method described by Mello *et al.* in 1991[8]. Here, DNA constructs of interest are injected with carrier DNA into the gonad of the hermaphrodite. Transgenic DNA can be in the form of plasmids, cosmids, YACs, or PCR products. In the distal arm of the gonad, the cytoplasm is syncytial, so a single injection of DNA content in this region will distribute itself amongst a large number of oocytes and thus a single worm can produce several transformed progeny, some of which will become stable transformants in following generations (**Figure 1.1**).

Commonly, plasmids containing genetic markers with visible phenotypes are co-injected with the construct of interest, to facilitate the identification of transformants. One strategy is to inject mutant *C. elegans* with plasmids containing rescue genes (e.g. *unc-119*, *pha-1*, *dpy-20*, *lin-15*). Transformants are then identified by wild-type morphology or behavior. Alternatively, gene fragments for dominant visible alleles (*lin-3*, *rol-6*) have also been used as co-injection markers and for injection into wild type hermaphrodites. More novel advancements include the advent of fluorescent proteins as well as selectable antibiotic resistances[9]. Each strategy has its own advantages and disadvantages. Genetic markers can

often be maintained under a standard dissecting microscope and in some cases do not require picking, but the genetic background can be a potential confound. Fluorescent/antibiotic markers require more equipment and reagents, but their genetic background is closer to that of the reference strain N2.

Once injected into the animal, the DNA forms a semi-stable extrachromosomal array. Despite its longstanding use over decades, the exact structure of the array and nature of its inheritance have remained somewhat of a mystery. Using restriction enzyme and Southern blot analysis, Mello *et al.* observed that arrays in the stable transformants contained multiple copies of the injected fragments, and that co-injected DNA fragments always segregated together even when one was present at very low copy number, suggesting that the array was a single heritable structure.

Injection of a P_0 parent commonly gives several F_1 transformed progeny, a fraction of which produce F_2 transformants and a stable line (**Figure 1.1**). This is due to arrays being mitotically unstable; every cell division risks loss of the array, producing mosaic animals. Animals failing to distribute the array to germline progenitors will fail to transmit to future generations, and thus transformants appear in a non-Mendelian fashion. This mosaicism is also a concern when conducting assays- if the array cannot be visibly tracked, an experimenter cannot be sure which cells have received the array and thus which cells are responsible for a phenotype, if any. This can introduce a high degree of variability into a study.

Extrachromosomal array mosaicism can be addressed via chromosomal integration of the array[10]. Double stranded breaks of the genome, via X-ray or ultraviolet irradiation, stimulate incorporation of a portion of the array into the site of the break. This alleviates the mitotic instability, ensuring faithful inheritance from cell to cell. Thus the variability introduced by animal mosaicism can be alleviated.

This workflow is a staple of *C. elegans* research, and a backbone of genetic analysis in the genetic model organism. At the time of writing, there are 11,634 extrachromosomal arrays and 11,217 integrated array alleles documented in WormBase (**Figure 1.2)**. However, the unknown structure of extrachromosomal and integrated arrays still raises several concerns.

In their study, Mello *et al.* performed Southern blot analysis of F_1 transformed progeny to reveal that even siblings of an injected P_0 could produce array structures as different as non-sibling F_1 transformants, but that clonal F_2 progeny and their descendants displayed essentially identical structure. Because of this, researchers often need to verify the behavior of an array across several independent lines to ensure that their results are not an artifact of the structure of a particular array, but a genuine finding. Another concern of multiple-copy arrays and integrations is whether the behavior of the allele evolves over time; it may be subject to rearrangement, silencing, or shortening.

The most we can say about the nature of the integrated array is that it contains multiple copies of the original DNA constructs; what relation it has to the array it was generated from, whether it undergoes structural rearrangement, and how the local genomic context of

the integration site influences expression, are questions the field is unable to answer well. Additionally, using extrachromosomal or integrated arrays raises questions about gene dosage. Commonly, such constructs are used for genetic rescue experiments, but their multi-copy nature may in fact result in overexpression and non-native levels of their gene products.

These issues pervade transgenesis in *C. elegans*. It appears in strains generated by the same researchers, and between researchers or laboratories when they attempt to replicate each other's strains and data.

There exist transgenesis methods which alleviate these issues, namely methods for single copy transgene insertion (SCI). Several methods have been developed in *C. elegans*, namely Mos transposon-based insertion[11–14] and newer CRISPR/Cas9 methods[15–18]. Both work on the principle of creating a single double-stranded break at a defined chromosomal location, and then using homology directed DNA repair (HDR) to repair the break with a supplied template containing the transgene flanked by neighboring homologous sequences. To generate double stranded breaks, the Mos SCI system has developed strains containing the Mos transposon present at various chromosomal locations. An inducible transposase causes excision of the transposon and a double stranded break at that location. CRISPR utilizes the *Streptococcus pyogenes* Cas9 nuclease and a guide RNA, which targets Cas9 to genomic sequence complementary to the guide RNA for cutting. These methods result in a single copy insertion at a defined locus, and control for copy number and genomic position. They also address issues born from the unknown structures and copy numbers in

extrachromosomal array and integrated lines. Furthermore, the single copy transgene can be inserted in specific and well-defined genetic loci, making comparisons between transgenic strains more reproducible and scientifically rigorous. Generally speaking however, single copy transgenes display weaker expression than their extrachromosomal or integrated array counterparts, which are multi-copy.

Each set of methods has its appeal to researchers, with its own set of advantages and disadvantages. Currently, extrachromosomal and integrated arrays are the most popular way of constructing transgenics, likely because extrachromosomal array lines can be obtained in under a week, informing the researcher whether an avenue of investigation is worthwhile. Single copy methods (Mos and CRISPR/Cas9), however, are gaining popularity, considering their more recent development (**Figure 1.2**).

But even with all these methods, most researchers will run into a series of barriers when trying to construct a strain from other researcher's data. If determining expression pattern, they must decide which portion of a promoter to clone in front of their gene. Well-documented papers will list the exact primers, while other, older papers may be more vague*. When reconstructing gene products, genes may have several isoforms and it may not be clear which (if any) introns were kept. Other variables on the researcher's mind might include the backbone vector, 3' UTR, choice of carrier DNA, and injection concentrations; variables that are only sometimes reported. If the experimenter is

* Often due to the state of molecular biology at the time, and not under the researcher's control.

constructing extrachromosomal or integrated arrays, such considerations may be pointless as the researcher cannot control the structure of their array anyways. These issues can make it difficult for laboratories to replicate one another's results, or to compare results across groups. They are compounded when considering the ultimate goal of making such transgenic animals: systematic analysis of the >20,000 *C. elegans* genes in its 959 somatic cells.

The number of promoter::gene combinations that need to be investigated to discover scientific results is potentially staggering. Even more so considering the ever-growing repertoire of useful heterologous genes (e.g. fluorescent reporters[19,20], calcium indicators[21,22], channelrhodopsins[23,24], etc.) For a set of *m* promoters and *n* gene products, the number of transgenes needed is the product *m x n*. Not only is this a large number of constructs, building such a set by conventional extrachromosomal/integrated array methods raises concerns. The expectation is that the same promoter or gene sequences will behave identically between different combinations, between different arrays, and between different integrated lines. But that expectation is a spurious one; we already know that a transformed P_0 progenitor can generate different arrays, and that integration of arrays can place insertions of differing copy number in any location in the genome. If the same promoter is used to drive two different gene products, and the two constructs are integrated in different genomic locations, it would be difficult to dissect whether differences seen between the two constructs is due to the gene products instead of their different positional effects, or perhaps the different structures of the integrated arrays.

1.3 Bipartite Gene Expression Systems

The issues of large numbers of promoter::gene combinations and variability can be mitigated using bipartite systems. Bipartite systems are gene expression systems that decouple the promoter from the gene product. Rather than having the promoter directly drive the gene product, they are separated into two components. The promoter drives a protein intermediate- this pairing is termed a 'driver'. The gene product is modified by control sequences that are only active when the protein intermediate is present, to generate an 'effector'. Each component is incorporated into separate strains, and they are mated to produce progeny with both components. Transgene expression is dependent upon having both components active in a cell (**Figure 1.3)**.

At first glance, the use of bipartite systems to control gene expression might appear paradoxical: why separate a promoter::gene fusion into two parts, thereby doubling the number of constructs and strains necessary to get the same expression pattern? This paradoxical view can be resolved, however, taking a systematic view of transgenesis in *C. elegans.* Using the example above, a set of *m* promoters and *n* gene products requires building only the sum $m + n$ set of driver and effectors strains, which can then be mated together to achieve the product $m \times n$ set of expression patterns. This is considerably more attractive if many of the desired *m* or *n* strains already exist for the community.

Perhaps the key benefit of using bipartite systems however, lies in the reusability and standardization of driver and effector strains. No matter how the driver or effector strains were constructed, whether by extrachromosomal array, integrated array, MosSCI, or

CRISPR, the fact that driver and effector constructs can be shared and reused means that expression is as faithful and reproducible as possible between different laboratories. Researchers are using the exact same alleles between them. As strains become used more and more, any idiosyncrasies or unique features about them are continually compiled and described, and can be addressed. An analogy can be made to computer packages and code. Such packages have many users using and testing them- any bugs found can then be documented and addressed. Similar scrutiny of bipartite expression strains would provide the same type of revision for researchers, which can be used to answer biological questions. Additionally, newly engineered heterologous protein effectors can be incorporated with existing drivers with ease, without the need to reconstruct all pairwise driver::effector strain combinations from scratch.

1.3.1 GAL4/UAS

The original bipartite system was introduced by Brand and Perrimon in 1993[25,26]. They introduced the *Saccharomyces cerevisiae* transcription factor GAL4 into *Drosophila melanogaster*, under the control of a *Drosophila* tissue specific promoter to form a driver. GAL4 recognizes and binds a 17-nucleotide DNA recognition sequence, termed an upstream activating sequences (UAS)[27]. In a separate strain, they placed UAS sites upstream of a gene product to construct the effector. When these two driver and effector parental strains are mated, their progeny drive expression of the effector gene in cells where the promoter is active. The GAL4-UAS system has drastically changed the landscape of *Drosophila* research, becoming an integral part of *Drosophila* transgenesis. It

is a standardized platform with which researchers conduct screens, perform tests of necessity and sufficiency, and discover site- and time-of-action for *Drosophila* genes.

Aside from the benefits mentioned above, bipartite systems also have the advantage of offering more refined spatiotemporal control through a variety of methods. Bipartite systems can incorporate repressive or intersectional strategies.

Often times, a single promoter dictating driver expression may be expressed too widely to make conclusive statements about the site of action of a gene. In this case, subsets of candidate cells can be eliminated by expressing GAL80, a repressor of GAL4. GAL80 binds GAL4, and blocks expression of the effector gene product[28]. Thus cellular subsets in which GAL80 repression of the effector does not affect phenotype can be considered dispensable for that process. GAL80 also affords a layer of temporal control: a temperature-sensitive allele of GAL80 exists whose activity can be controlled by shifting between permissive and restrictive temperatures.[29]

Another strategy to limit expression of the effector is to use split intersectional strategies. Studies of GAL4 had shown that the transcription factor could be split into two modular domains: one domain is responsible for UAS binding and dimerization, the other recruits transcriptional machinery to drive gene expression[30–32]. These two domains could be expressed separately, and each would retain its native activity. This provides the opportunity to dictate effector expression using two promoters instead of one. One promoter dictates expression of the DNA-binding domain; a second dictates expression of

the transcriptional activation domain. If one desires very specific spatial expression, two spatial promoters can direct expression of each module. Alternatively, one promoter can provide spatial information, which the other can afford a layer of temporal control. A functional GAL4 transcriptional activator is formed only in the intersection of time and space where both promoters are active and both components are expressed, assuming each domain has a method of re-associating with the other. The original description of split GAL4 fused leucine zipper domains to each module of the GAL4 domain, which could then reconstitute via non-covalent binding of the zippers[33].

1.3.2 Other Bipartite Systems

A number of other bipartite systems have been engineered for genetic model organisms: LexA, Q, Tet On/Off, Cre/lox, and FLP/FRT. These systems can largely be classified into two classes: transcriptional activator-based systems (encompassing GAL4, LexA, Q, and Tet systems) and recombinase-based systems (encompassing Cre/lox and FLP/FRT).

The LexA[34,35] and Q[36–38] systems resemble that of GAL4. A promoter directs expression of LexA or QF transcriptional activators, respectively. These proteins recognize their cognate DNA motifs (LexAop and QUAS sites, respectively) and drive transcription of the downstream gene product. The Q system has an added layer of control in that QF is repressible, via the QS repressor, much like GAL80 represses GAL4. QS represses the QF activator, and this inhibition is removed in the presence of quinic acid. The Tet system has two versions, Tet-On[39,40] and Tet-Off[41]. Each has a unique transcriptional regulator, both binding the TetO transcriptional operator sequence in different situations. In Tet-On, the presence of doxycycline allows the activating transcriptional regulator rtTA to bind TetO

sequences and drive transcription of effector genes; removal of doxycycline unbinds rtTA from TetO to halt transcription. The functionally opposite Tet-Off system uses the tTA activator, which drives transcription at TetO sites in the absence of doxycycline; addition of doxycycline dissociates rTA from TetO sequences.

In the case of Cre/lox and FLP/FRT promoters drive expression of the Cre or FLP DNA recombinases. These recombinases recognize their cognate target sites (lox and FRT, respectively) to generate site-specific DNA excision or inversion between these sites, depending on their orientation[42,43]. The gene product effector is modified with lox or FRT sites to ensure Cre- or FLP-dependent activation of the gene product. This can be achieved by a variety of methods: insertion of a transcriptional stop cassette flanked by target sites upstream of the gene for excision upon recombinase expression, or by creating an inverted gene flanked by target sites that the recombinase can then restore to its proper orientation. Note that in the case of recombinases, the effector gene must also have its own promoter. This second promoter is often a ubiquitously expressed promoter, such that expression is then limited solely by the driver. If the expression pattern of the second promoter is more limited, the final expression pattern is the intersection of both promoters.

Combining bipartite systems can grant more advanced control of gene expression, much like intersectional and repressive strategies can. A common strategy is to combine a recombinase-based method with a transcriptional activator-based method. One system is driven by a temporal promoter, while the other is controlled by a spatial promoter. The transcriptional activator drives expression of the recombinase, and the effector is flanked

by the DNA motifs recognized by the recombinase. The final spatiotemporal patterning is determined by the intersection of both promoters.

The two differing mechanisms of gene activation between transcriptional activator-based and recombinase-based systems have several experimental ramifications, when considering *C. elegans* transgenesis.

Generally, recombinase-based systems must function at the single copy level. Having multiple copies of recombinase target sites in extrachromosomal or integrated arrays complicates the recombination process. Since such sites work in pairs, it is not possible to control or determine which pairs of sites were recombined, or how many final, in-frame copies of the effector were produced. As might be expected of single copy transgenes, expression is generally weaker, but often is sufficient for assays using genetic rescue or bimodal states (i.e. cell death).

Transcriptional activator-based systems are not limited to single copy methodology and can support construction via extrachromosomal/integrated arrays. This means their expression is often much stronger, which can be desirable when expressing heterologous proteins non-native to the model organism (i.e. fluorescent reporters, sensors, membrane channel proteins, etc.). One major drawback associated with increased expression is increased background, however.

Understanding copy number, expression strength, and background level help to set the sensitivity and specificity of an experimenter's assay. A variety of options of bipartite systems would be of great use to the *C. elegans* community, each suited for a particular design.

1.4 Conclusion

Bipartite systems have the power to greatly improve and accelerate genetic and cellular dissection of biology phenomena in genetic model organisms. Several bipartite systems have shown to function in *C. elegans*, but no GAL4-based system previously existed for *C. elegans*. The following chapters of this thesis describe engineering and application of a GAL4-based bipartite gene expression system for *C. elegans*. Chapter 2 describes the construction of cGAL, the base GAL4 system using a novel GAL4 gene from *Saccharomyces kudriavzevii*, demonstrating using its functionality across various tissues and gene products. Chapter 3 describes a split intersectional strategy using split intein proteins, allowing control of our cGAL system to be dictated by two promoters, providing more refined spatiotemporal control. Chapter 4 briefly presents preliminary data on single copy cGAL design, as well as future directions for the bipartite system. The development of this GAL4 bipartite system will prove to be a valuable resource to the *C. elegans* community, accelerating research for the community and improving reproducibility.

1.5 FIGURES

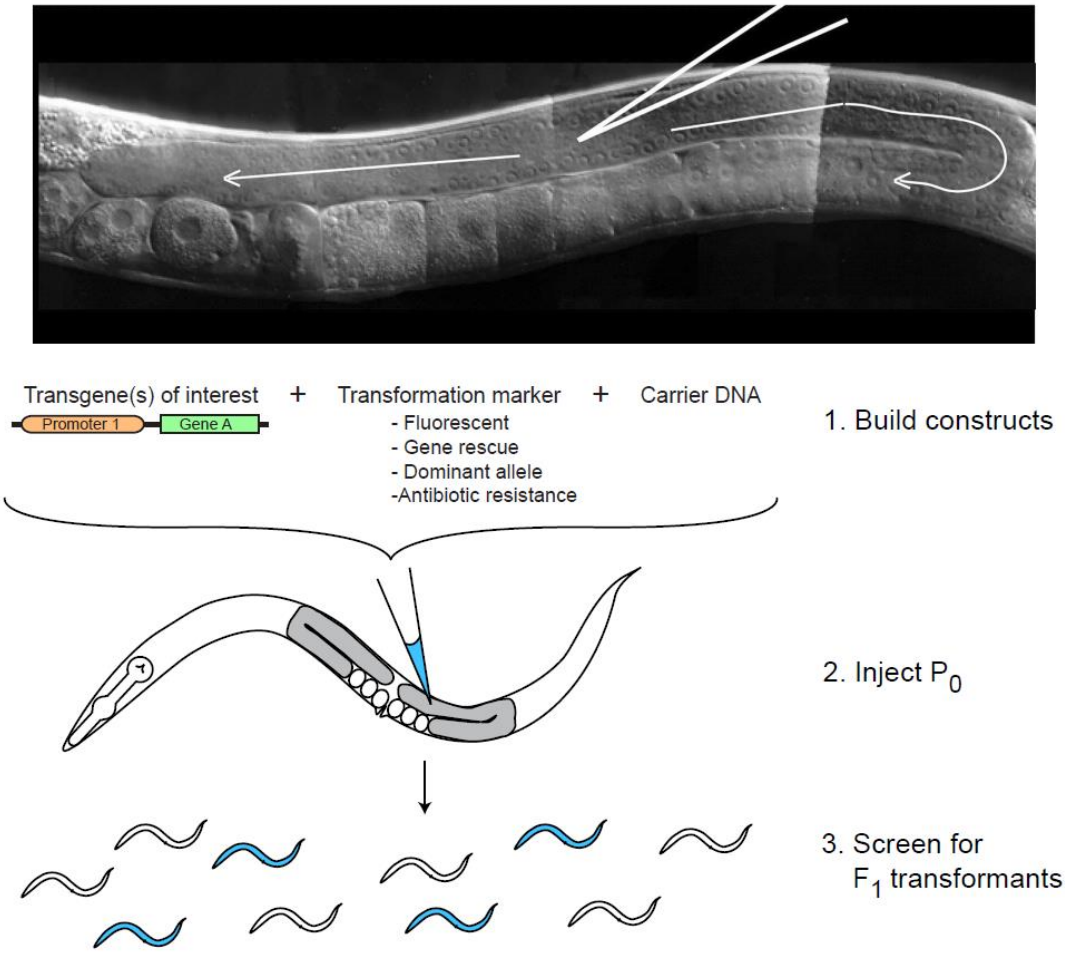


Figure 1.1 | *C. elegans* transformation

Upper: One arm of the *C. elegans* gonad. Transgenic constructs are injected into the distal arm, which consists of syncytial nuclei which will eventually partition into individual oocytes. Injected genetic material is concatenated into a heritable extrachromosomal array. Image taken from WormBase, under the terms of the Creative Commons Attribution License:

http://www.wormbook.org/chapters/www_transformationmicroinjection/transformationmicroinjection.html

(continued from **Figure 1.1***)*

Lower: Standard *C. elegans* transgenesis protocols. Transgenic constructs of interest are combined with visible markers and carrier DNA. Once injected into a P_0 animal, the resulting generations are screened for transformed animals.

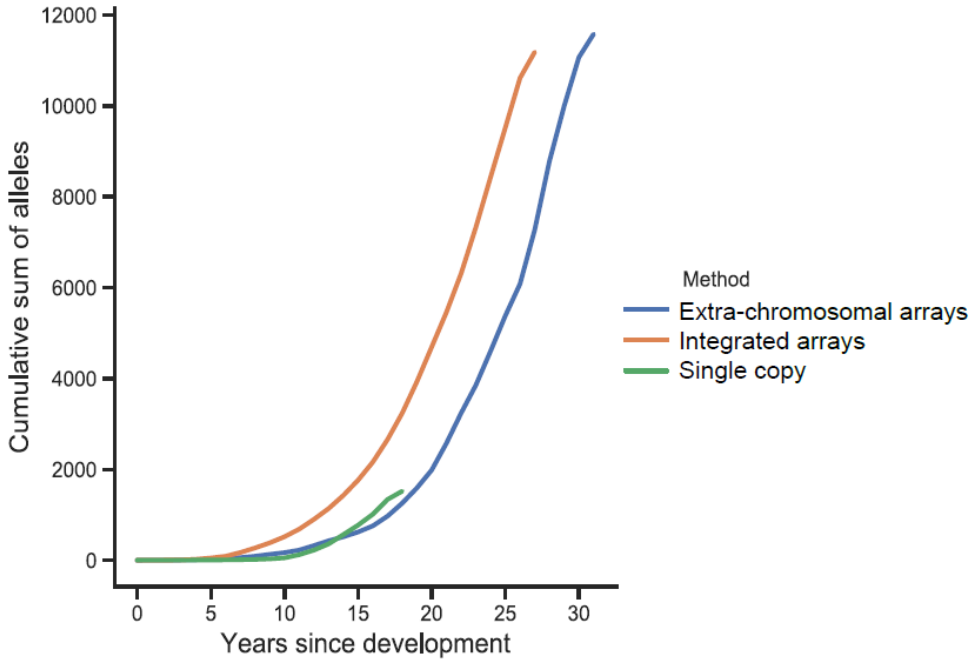


Figure 1.2 | Cumulative sum of various transgene types in *C. elegans*

Documented transgenes generated by various methods, computationally mined from *C. elegans* literature via Textpresso. Transgenes are grouped by *Ex* (extrachromosomal array), *Is* (integrated array), *Ti* and *Si* (single copy) transgenes. The *x*-axis is set to the number of years since the first publication of the transgenesis method for easier comparison.

Special thanks to:

Karen Yook, Juancarlos Chan, Hans-Michael Muller

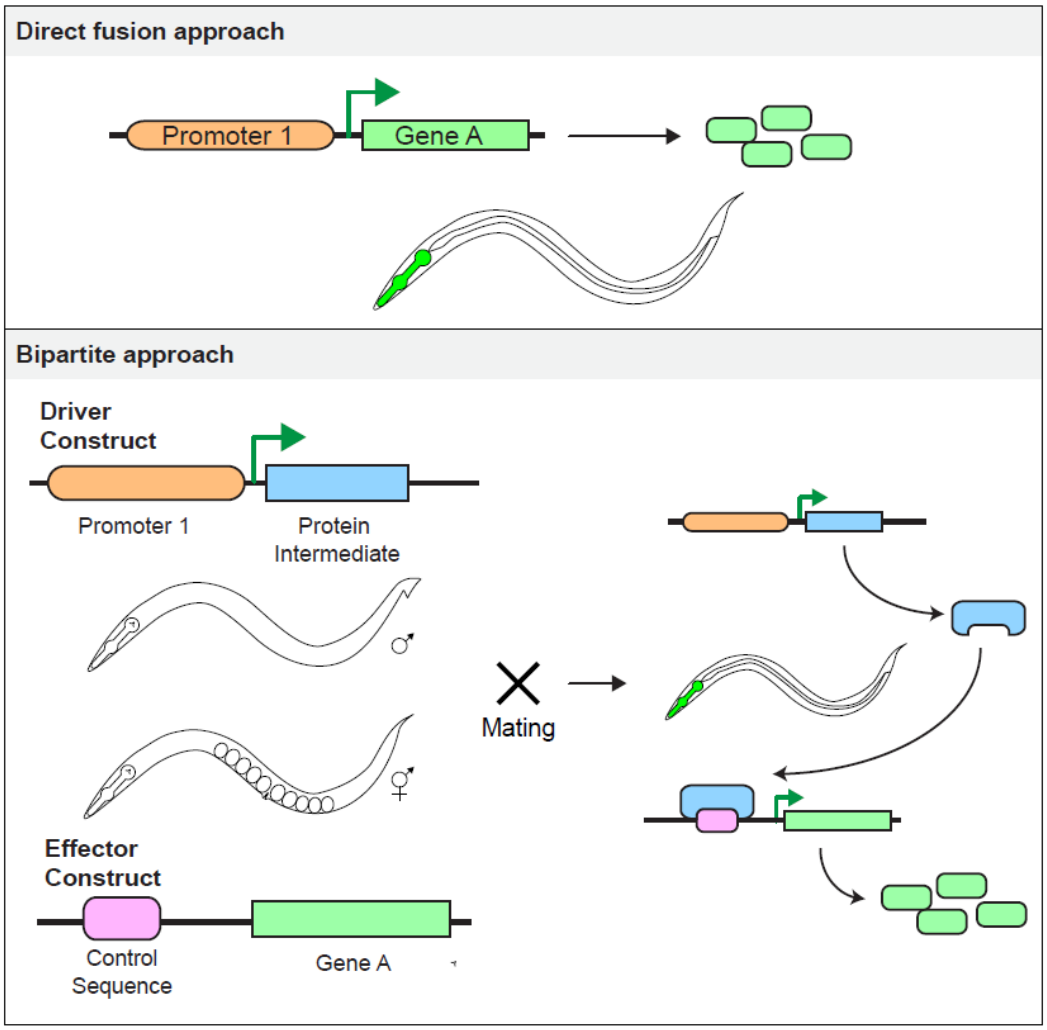


Figure 1.3 | Direct fusion and bipartite approaches to transgenesis

Upper panel: Traditional transgenesis directly couples a promoter element with the gene to be produced. Cells where the promoter element is active produce the gene product.

Lower panel: In the bipartite approach, the promoter and gene are decoupled. A driver strain contains a transgene with a promoter driving a protein intermediate. A separate effector strain contains a transgene with the gene product downstream of control sequences. Mating parents produces progeny containing both driver and effector elements, and the gene product is produced where the promoter is active, much as in the direct fusion

(continued from **Figure 1.3***)*

approach. However, drivers and effector strains can be mated combinatorially to produce vast arrays of promoter::gene combinations and are more efficient and reproducible.

Chapter 2

Engineering cGAL, a GAL4-based Bipartite Gene Expression System for *C. elegans*

This chapter can be found published as an article in:

Wang H, **Liu J**, Gharib S, Chai CM, Schwarz EM, Pokala N, Sternberg PW. cGAL, a temperature-robust GAL4-UAS system for Caenorhabditis elegans. Nat Methods. 2017 Feb;14(2):145-148. doi: 10.1038/nmeth.4109.

2.1 ABSTRACT

Control of gene expression in desired spatiotemporal patterns is a powerful tool for studying gene function. We have established cGAL, a bipartite GAL4-UAS system that effectively controls gene expression in *C. elegans*, by systematically optimizing the three major components: activation domain, UAS copy number, and DNA-binding domain (DBD). Most importantly, the canonical Gal4p DBD from *Saccharomyces cerevisiae* (grown best at 30-34°C) performs poorly across *C. elegans* cultivation temperatures (15-25°C). To overcome this, we use evolutionary analysis to identify a *GAL4* homolog from *Saccharomyces kudriavzevii,* whose optimal growth temperature is 23-24°C, and show that *S. kudriavzevii* Gal4p DBD displays temperature robustness across 15-25°C. We demonstrate the utility of the cGAL system in enabling reporter expression in multiple tissues, site-of-action experiments, and gain-of-function channelrhodopsin experiments. We expect that cGAL will not only significantly aid *C. elegans* research, but also facilitate the application of GAL4-UAS systems in other organisms with low growth temperatures.

2.2 INTRODUCTION

An understanding of genes and their functions is essential for discovering insights about an organism's physiology, development, genetics, and behavior. Spatial and temporal control of gene expression, where possible, is a powerful tool for dissection of gene function. In genetically tractable organisms, this is often done using direct promoter::gene fusion constructs. More sophisticated control has also been achieved with bipartite expression systems such as the GAL4-UAS system.

The GAL4-UAS system was originally described in the yeast species *Saccharomyces cerevisiae*[44]. The protein Gal4p consists of two functional domains: a DNA-binding domain (DBD), and two transcription activation domains (AD). As a homodimer, Gal4p recognizes and binds a 17-nucleotide upstream activating sequence (UAS) in the promoter regions of *GAL* genes in yeast, in order to utilize galactose as a carbon source[45]. The GAL4-UAS system was repurposed as a bipartite system to control transgene expression in flies[25]. Unlike conventional direct promoter::gene fusion constructs, the GAL4-UAS system decouples the promoter from the effector gene into two separate constructs (**Figure 2.1**). Promoters are placed upstream of Gal4p to generate a 'driver' construct, while one or more copies of UAS are placed upstream of the gene product to create an 'effector' construct. Incorporation of these constructs into separate transgenic animals creates standardized driver and effector lines, which can then be crossed together to achieve desired gene expression patterns in the progeny.

This system offers three major advantages over the promoter-gene fusion approach. First, large numbers of strains can be easily generated by crossing different drivers and effectors. With ten driver lines and ten effector lines, a hundred gene expression patterns could be obtained in animals that would normally require a hundred promoter::gene fusions. Second, it takes minimal effort to incorporate novel components into available driver and effector libraries. For instance, the development of optogenetic tools and genetically encoded calcium indicators moves at a rapid pace[46,47]. When a new generation of a tool is developed, only one effector line with such a tool needs to be generated and it can be used by crossing with existing driver lines. Third, once drivers and effectors are tested and

validated, they become community reagents that foster experimental consistency and reproducibility, and allow comparison of findings across studies.

Indeed, with a large library for *GAL4* driver lines and effector lines available, the GAL4-UAS system has become the method of choice for controlling gene expression in *Drosophila*[48]. Furthermore, several improvements of the GAL4-UAS system and its combination with other genetic tools, such as the incorporation of Gal80p, the Gal4p negative regulator, have allowed researchers to perform sophisticated genetic manipulations for the study of many different biological questions in *Drosophila*[48–50]. Because of the power of this bipartite expression system, the GAL4-UAS system has also been adopted in other organisms, such as mice, zebrafish, plants, beetles and mosquitoes[51–55].

However, this system has not previously been successfully adapted for use in *C. elegans*. We therefore modified the components of the GAL4-UAS system for *C. elegans.* Previous efforts to improve the GAL4-UAS system in other organisms have included using more powerful activation domains (ADs), and increasing UAS copy number[56,57]. Furthermore, temperature has been implicated to have an impact on the performance of GAL4-UAS *in vivo*. For example, in *Drosophila*, temporarily shifting animals to 29°C improves GAL4-UAS effectiveness, compared to 25°C or 18°C[26], suggesting that the normal cultivation temperature range of 15-25°C for *C. elegans* might reduce the system's performance. We hypothesized that using a DBD of the Gal4 protein from a colder-growing yeast species might mitigate the effects of low temperature. Taking these

considerations together, we systematically compared the effects of altering AD, the UAS copy number and DBD to develop an optimized GAL4-UAS system for *C. elegans*. We name the new system cGAL, in reference to its initial development in *C. elegans*, and to its engineered performance at cooler temperatures. We demonstrate its robustness for reporter gene expression in multiple tissues, site-specific genetic rescue, and channelrhodopsin experiments in *C. elegans*.

2.3 RESULTS

2.3.1 The more potent activation domain VP64 significantly improves driver activity

Previous unpublished attempts at engineering GAL4-UAS in *C. elegans* suggested that transcriptional activity, while present, was relatively low (our unpublished data and personal communication from H. Korswagen). These attempts used the *S. cerevisiae* Gal4p DBD (residues 1-147, henceforth termed $Gal4_{SC}$) fused to the VP16 viral protein activation domain (VP16) from human herpes virus[58] as the driver components. The effector component consisted of five copies of UAS upstream of *gfp (5xUAS::gfp)*. Because previous attempts in other organisms suggested that stronger activation domains improve the performance of the GAL4-UAS system[56,57], we hypothesized that a stronger transcriptional activation domain might boost the performance of this system in *C. elegans*. A synthetic transcriptional activation domain with four tandem copies of VP16, called VP64, has been shown to be more effective than VP16[59]. We fused the $Gal4_{SC}$ to VP16 and VP64 and placed them under the same promoter (the promoter of *myo-2* gene, *Pmyo-2*, a regulatory promoter specific for pharyngeal muscles), designating them as

Pmyo-2::GAL4$_{SC}$::VP16 and *Pmyo-2::GAL4$_{SC}$::VP64*, respectively. To compare the performance of these two drivers, we first generated a transgene strain containing a chromosomally integrated array of a *15xUAS::gfp* effector (*syIs300)*. We then injected each driver construct at equal concentrations into the effector strain.

At room temperatures (22-23°C), we found that the *Pmyo-2::GAL4$_{SC}$::VP64* driver caused a seven-fold increase of GFP fluorescence in pharyngeal muscles over that seen with the *Pmyo-2::GAL4$_{SC}$::VP16* driver (**Figure 2.2**, $p<0.0001$), demonstrating that the transcription activation domain VP64 greatly outperforms VP16 in *C. elegans*. The GFP fluorescence observed in pharynx is dependent on the presence of both the driver and the effector: neither the parental *Pmyo-2::GAL4$_{SC}$::VP64* transgenic strain alone nor the parental *15xUAS::gfp* effector transgenic strain alone showed GFP fluorescence; only cross progeny of these two strains displayed bright GFP fluorescence in pharynx (**Figure 2.3**). We therefore adopted VP64 as our activation domain of choice for further experiments.

2.3.2 Increasing UAS copy number enhances reporter expression

To further improve the efficacy of the GAL4-UAS system for *C. elegans*, we compared the effects of different UAS copy numbers on the expression of the effector gene. As the Gal4p DBD binds to UAS to recruit transcriptional machinery through the transcriptional activation domain, we reasoned that increasing UAS copy number upstream of the effector gene might enable more Gal4p binding to the promoter region, leading to better expression of the downstream effector gene. To test this, we injected effector constructs with 5x, 10x,

15x, and 20x copies of the UAS sites upstream of *gfp* at equal concentrations, into a transgenic strain with an integrated *Pmyo-2::GAL4$_{SC}$::VP64* driver (*syIs301*). Quantitative fluorescence microscopy revealed a successive increase in GFP fluorescence up until 15 copies of UAS (~2.3 fold vs. 5x, $p<0.0001$; ~1.3 fold vs 10 copies, $p<0.01$), beyond which it appeared to saturate (~1.1 fold vs. 20x, $p=0.51$, not significant; **Figure 2.4**). Thus, increasing UAS copy number generally improves the expression of the downstream effector gene but does eventually saturate; therefore, we adopted 15x copies of UAS for all effector lines used in the remainder of our experiments.

2.3.3 The GAL4-UAS system efficacy is heavily dependent upon temperature

Experimental temperatures for *C. elegans* growth usually range from 15°C to 25°C. We assayed our *Pmyo-2::GAL4$_{SC}$::VP64* driver and *15xUAS::gfp* effector combination at 15°C, 20°C, and 25°C to determine the robustness of the GAL4-UAS system under these conditions. We found that the transcriptional efficacy of the GAL4-UAS system was heavily dependent on temperature. Our driver/effector combination performed well at 25°C, but fell precipitously at lower temperatures (**Figure 2.5**, ~67% drop at 20°C, ~80% drop at 15°C, $p<0.0001$ for both, adjusted). These results were consistent with findings in *Drosophila:* temporarily shifting flies to 29°C instead of the usual 25°C or 18°C increases Gal4p-mediated expression of the effector[26]. Thus, the temperature dependence of the GAL4-UAS system may also contribute to its previously weak and unreliable performance in *C. elegans*.

2.3.4 The Gal4p DBD from *S. kudriavzevii* provides robust and increased performance at low temperatures

The GAL4-UAS system from *S. cerevisiae* displayed improved performance at temperatures closer to the optimal growth temperature of *S. cerevisiae* (around 30-34°C)[26,60]. This observation suggests that Gal4p from *S. cerevisiae* may have evolved to be maximally active around this optimal temperature, and may explain its poor performance across the 15-20°C range. We reasoned that a Gal4p from more cryophilic *Saccharomyces* yeast species with optimal growth temperature ranges closer to that of *C. elegans* (around 20°C) might provide excellent building blocks for a more robust GAL4-UAS system in *C. elegans*. We chose to test the Gal4p DBD from the Portuguese reference strain ZP591 of *Saccharomyces kudriavzevii* [61] (residues 1-147, henceforth termed $Gal4_{SK}$) for two main reasons. First, *S. kudriavzevii* has an optimal growth temperature (23-24°C) closest to that of *C. elegans* amongst the *Saccharomyces* species[60]. Second, the $Gal4_{SK}$ sequence is highly conserved with that of the $Gal4_{SC}$. In particular, the six cysteine residues in the Zn_2Cys_6 binuclear cluster that are essential for DNA binding and two key lysine residues that directly contact the "CGG" nucleotides in the UAS site [45,62,63] are identical between the two yeast species (**Figure 2.6**). Thus we hypothesized that $Gal4_{SK}$ would still bind to the same UAS site, while the remaining subtle changes in protein sequence may confer improved performance at lower temperatures.

To test this, we first generated a new P*myo-2* driver construct (for pharyngeal muscles) by replacing the original $Gal4_{SC}$ with $Gal4_{SK}$ and retaining VP64 as the activation domain. Then, we compared its performance across a temperature series by injecting this new driver

into the same transgenic strain with the integrated *15xUAS::gfp* effector (*syIs300),* at the same concentration as our previous *Pmyo-2::GAL4$_{SC}$::VP64* driver. We noticed a pattern of increased GFP fluorescence intensity with the *Pmyo-2::GAL4$_{SK}$::VP64* driver lines over the *Pmyo-2::GAL4$_{SC}$::VP64* driver lines; blinded researchers could consistently and with perfect accuracy sort each driver by eye through a standard fluorescent dissecting microscope. We chose the brightest line from each for quantitation at 15°C, 20°C and 25°C. Across this temperature series, the *Pmyo-2::GAL4$_{SK}$::VP64* transgenic line performed more robustly than *Pmyo-2::GAL4$_{SC}$::VP64*. Both have comparable fluorescence intensities at 25°C (**Figure 2.5**). However, Gal4$_{SK}$ exhibited more robustness to temperature, experiencing only a ~20% drop at 20°C ($p<0.01$, adjusted; vs ~67% drop for Gal4$_{SC}$), and a ~40% drop at 15°C ($p<0.001$, adjusted; vs ~80% drop for Gal4$_{SC}$). At room temperature, we observed ~30% improvement in GFP fluorescence with the new driver over Gal4$_{SC}$ (2 lines each, **Figure 2.7**). We also analyzed strains that were injected with a direct *Pmyo-2::gfp* fusion at the same concentration for comparison, noting that GFP fluorescence levels were comparable (**Figure 2.5** and **Figure 2.7**). This led us to adopt the Gal4$_{SK}$ domain as our DBD of choice, in conjunction with the VP64 activation domain and the 15xUAS effector, to comprise our fully optimized GAL4-UAS system for *C. elegans.* We designate it the cGAL system to denote its original implementation in *C. elegans* and its potential use in other organisms at cooler temperatures.

2.3.5 The cGAL system performs well across multiple tissues

We further tested whether cGAL would perform in other major tissues, beyond pharyngeal muscles. We generated new driver constructs by replacing the pharyngeal muscle specific

promoter (P*myo-2*) with other tissue-specific promoters (P*nlp-40* for the intestine; P*myo-3* for the body wall muscles). When injected into a strain with the *15xUAS::gfp* effector (*syIs300*), these new drivers produced robust and specific GFP expression in the expected tissues (**Figure 2.8**, *top row*).

However, we encountered an issue when we attempted to drive GFP expression pan-neuronally (P*rab-3*) and in GABAergic neurons (P*unc-47*) in *C. elegans*; we found poor and highly mosaic expression in the nervous system along with intense ectopic GFP fluorescence in the posterior gut (data not shown). We speculated this issue might have been due to the vector (a derivative of the Fire vector pPD49.26) that we used for the driver and effector constructs. This vector contains the *unc-54* 3'UTR, a common 3'UTR used for transgene expression in *C. elegans*. To address this issue, we switched to the pPD117.01 vector which contains a 5' decoy (see Methods) and the *let-858* 3'UTR, which were introduced to reduce ectopic expression in the posterior gut and to improve the transgene expression in a broad range of tissues (A. Fire, personal communication). We generated new P*rab-3* and P*unc-47* driver constructs as well as a new *15xUAS::gfp::let-858 3'UTR* effector construct in the pPD117.01 backbone. We found that these two drivers, when injected into a transgenic strain with a new *15xUAS::gfp::let-858 3'UTR* effector (*syIs343)*, not only displayed decreased ectopic GFP fluorescence in posterior gut (data not shown), but also dramatically increased GFP expression in the entire nervous system and GABAergic neurons, respectively (**Figure 2.8**, *middle row* and data not shown). To further validate pPD117.01 as the vector of choice for neuronal drivers, we generated two additional drivers by cloning in regulatory elements for cholinergic (P*unc-17*) and

glutamatergic (P*eat-4*) neurons. When the P*unc-17* and P*eat-4* drivers were injected into the transgenic strain carrying the integrated new *gfp* effector (*syIs343*), we also observed specific and robust expression in the corresponding neurons (**Figure 2.8,** *bottom row*). Thus, the cGAL system is robust across a variety of tissues in *C. elegans* and we recommend using the pPD117.01 backbone to construct new drivers and effectors.

2.3.6 Using cGAL for tissue-specific rescue

One of the frequent uses of bipartite expression systems is to facilitate site-of-action experiments by rescuing mutant animals via tissue-specific gene expression. We next tested whether the cGAL system could be applied for such functional studies in *C. elegans* by examining the defecation motor program (**Figure 2.9**). In *C. elegans,* defecation occurs approximately every minute and consists of three sequential muscle contractions in the following order: a posterior body wall muscle contraction (pBoc), an anterior body wall muscle contraction (aBoc), and lastly an enteric muscle contraction causing an expulsion event [64,65]. For the expulsion step, the intestine releases the mature neuropeptide from NLP-40, which binds to its receptor AEX-2 on two GABAergic neurons (AVL and DVB) to activate them. Activation of these two neurons causes the release of the neurotransmitter GABA, which triggers enteric muscle contraction and expulsion[66–68].

In *aex-2(sa3)* mutant animals, expulsion is nearly eliminated, while pBoc is unaffected [67]. We applied the cGAL system to test the site-of-action of *aex-2* in the GABAergic neurons for expulsion. Demonstrating the efficiency of cGAL as a bipartite expression system, we re-used drivers that drove *15xUAS::gfp* expression in each of the three tissues involved

with the expulsion circuit: the *Pnlp-40* intestine driver, *Punc-47* GABAergic driver, and *Pmyo-3* body wall muscle driver. We first generated a new transgenic effector line with *15xUAS::aex-2(+)* cDNA (*syEx1444*) in the *aex-2(sa3)* background. We then crossed each of these three driver lines with the new effector line to generate heterozygous animals with *aex-2(+)* cDNA specifically expressed in the intestine, GABAergic neurons, or muscles in the *aex-2(sa3)* background and assayed if any of them had expulsion rescued.

As expected, wild-type animals displayed expulsion events in nearly every defecation cycle, whereas *aex-2(sa3)* animals displayed almost none (**Figure 2.10**). In the *aex-2(sa3)* background, neither the GABAergic driver alone, nor the *15xUAS::aex-2(+)* cDNA effector alone was capable of rescuing expulsion events, demonstrating a lack of leaky expression of *aex-2* in the effector line alone. We found that only *aex-2(sa3)* mutants with both the GABAergic driver and the *15xUAS::aex-2(+)* cDNA effector line displayed rescue (**Figure 2.10**). The incomplete rescue may be due to the mosaic effect of extrachromosomal array of the *15xUAS::aex-2(+)* cDNA effector transgene. Furthermore, ectopic expression of *aex-2(+)* cDNA in either body wall muscle or intestine, the other two tissues in the expulsion circuit, did not rescue, demonstrating a high functional specificity using the cGAL system. These results are highly similar to results from a previous study that tested the site-of-action of *aex-2* for expulsion in the same three tissues using conventional promoter-cDNA fusion transgenes[67]. Thus, we conclude that our cGAL system can be used for tissue-specific rescue experiments in *C. elegans*.

2.3.7 Using cGAL for heterologous channelrhodopsin activation

Next, we tested whether our system could be used to perform gain-of-function experiments in *C. elegans*. Channelrhodopsin is a light-sensitive cation channel; in the presence of all-*trans* retinal and blue light, channelrhodopsin will open and depolarize cells it is expressed in[69]. Activation of GABAergic neurons in *C. elegans* relaxes the body wall muscles and causes the worms to adopt a flaccid, paralyzed state[70]. We injected the GABAergic cGAL driver (*Punc-47::GAL4$_{SK}$::VP64::let-858 3'UTR*) into a transgenic strain carrying an integrated *15xUAS::hChR2(H134R)::eyfp::let-858 3'UTR* effector (*syIs341*) to express channelrhodopsin specifically in GABAergic neurons and tested whether we could use blue light to selectively activate these neurons. We found that in the presence of the co-factor all-*trans* retinal, 475 nm blue light excitation caused an immediate, limp paralysis only in animals possessing both the driver and the effector, but not in animals with just either component alone (**Figure 2.11** and data not shown). Paralysis phenotypes were reversed immediately upon blue light removal. We exposed animals to a three-pulse train, spaced 20 seconds apart, and scored for responses. Most animals with both driver and effector constructs showed full and robust responses to blue light (20 animals, 60 total pulses, ~83% response) in contrast to a complete lack of response in control animals with carrying just the effector (10 animals, 30 total pulses, 0% response, $p<0.0001$; **Figure 2.12**). These results demonstrate that cGAL confers the ability to control the expression of novel exogenous transgenes to dissect neural circuits in *C. elegans*.

2.3.8 Construction of an initial, basic cGAL toolkit

Lastly, we built a basic cGAL driver and effector toolkit (**Tables 2.1 & 2.2**). For drivers, we constructed strains and constructs for major tissues, major neurotransmitter cell types, and some individual sensory neurons. For effectors, we have integrated strains for cell labeling (GFP, mKate2, GFP-H2B and mCherry-H2B), cell ablation (ICE), calcium indicator (GCaMP6s), neuronal activation (ChR2), neuronal inhibition (HisCl1) and synaptic inhibition (TeTx). All effectors are integrated and at least one line for each integrated effector was confirmed functional (**Figure 2.13**).

2.3.9 Discussion

The power of the GAL4-UAS system has been demonstrated in organisms such as *Drosophila*[48], but this tool has not been successfully implemented in *C. elegans*, because of unreliable and weak expression. We solved this problem by establishing cGAL, an optimized GAL4-UAS system for robust control of transgene expression at its preferred growth temperatures (15-25°C) and demonstrating that this system can be used for functional studies in *C. elegans*. Transcriptional efficacy of the system was greatly enhanced by introducing a more powerful transcriptional activation domain (VP64), in conjunction with additional UAS copy numbers (15x). Most importantly, cGAL confers robust performance across the entire temperature range relevant for *C. elegans*, by incorporating a Gal4p DBD from a yeast species adapted to grow at a lower preferred growth temperature.

While the Gal4p DBDs are highly conserved between *S. kudriavzevii* and *S. cerevisiae* (**Figure 2.6**), small differences in the $Gal4_{SK}$ DBD sequence afford better transcriptional activation of effector expression at lower temperatures. This enhancement might be attributed to higher affinity of $Gal4_{SK}$ DBD to the UAS site, more efficient dimerization, greater stability of $Gal4_{SK}$, more favorable folding kinetics, or any combination thereof at lower temperatures. This finding suggests that natural selection may have shaped yeast Gal4 proteins to maximize their performance at species-specific optimal growth temperatures. We predict that this optimized cGAL system and the same engineering principles could be useful in other genetically tractable organisms with optimal growth temperatures are at or below 25°C; in particular for those that have previously lacked tools for genetic analysis but are emerging as new models for developmental biology[71].

Our work has established a fully functional GAL4-UAS system for *C. elegans* research. First, the expression level of the GFP reporter with our cGAL4 bipartite system was robust and comparable with the direct transcriptional reporter approach. Second, further optimization of 3'UTR used in driver and effector constructs enabled robust application of cGAL in various somatic tissues and cell types. Third, we have also demonstrated that the cGAL system works efficiently for site-of-action experiments and for introducing optogenetic tools in *C. elegans.* Fourth, we have built a basic toolkit with cGAL drivers and effectors primarily for the neuroscience field (**Table 2.2**), but these tools can used for studies in major tissues, major neuronal classes, and specific neurons.

Another bipartite expression system, the Q system, has been recently described for *C. elegans*, including the demonstration of temporal control with an inhibitory protein and small molecule[38]. However, the Q system has yet to be widely adopted by the *C. elegans* community; this may be due to a lack of sufficient drivers and effectors.

As is the case in *Drosophila*[48–50], the cGAL system could be combined with other binary expression systems, including the Q system, Cre/loxP system, and FLP/FRT system[38,42,72,73], to enable tighter control of transgene expression using intersectional strategies for more refined spatial control. Furthermore, having multiple independent genetic control systems then enables studies using multiple effectors, each expressed in separate and distinct cellular patterns. This is particularly attractive for neuronal studies where pre-and post-synaptic neurons often need to be stimulated and then recorded in tandem, or for genetic and developmental networks to determine whether genes function cell-autonomously or cell-non-autonomously for each of their specific functions.

In its current form, cGAL already provides several new opportunities for *C. elegans* research. Genetic site-of-action experiments will be greatly facilitated by a collection of cDNA effectors and cell-specific drivers. The vast majority of site-of-action experiments in *C. elegans* are undercontrolled because of the tedium of constructing transgenes and strains. Genome-wide overexpression screens can be performed in a tissue-specific manner with the cGAL system, which will potentially reveal novel functions for many genes; in particular, for genes with redundant paralogs and genes that have lethal consequences when globally over-expressed. These will allow functional dissection of genetic and

developmental networks. Neural circuits can be efficiently probed with a growing collection of light-sensitive and ligand-activated channels and calcium indicator effectors[46,47,74], and importantly, as new tools are developed, a single new effector construct will enable researchers to incorporate such a new tool with existing drivers. The cGAL system will give rise to a continually expanding library of communal resources for the *C. elegans* field, and we expect that the cGAL system will greatly increase the rate and rigor of study in *C. elegans* and potentially in other organisms.

2.4 FIGURES

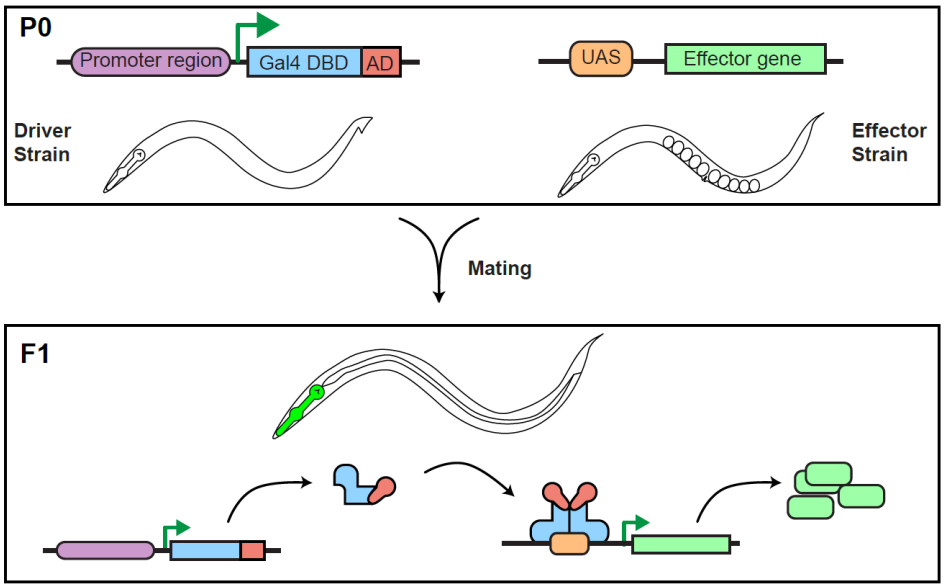


Figure 2.1 | The GAL4-UAS bipartite system

Schematic of the GAL4-UAS bipartite expression system. In one parent strain, a promoter drives the expression of a fusion protein of the DNA binding domain (DBD) of Gal4p and a transcriptional activation domain (AD) in the driver construct; in the other parent, an upstream activation sequence (UAS) is placed 5' to a gene effector. Mating the two parental stains generates offspring containing both components, triggering expression of the effector gene in a pattern dictated by the promoter. Decoupling promoters from effectors with the GAL4-UAS system allows efficient combinatorial control of gene expression. DBD refers to the first 147 amino acids of Gal4p; a synthetic 17-mer was used as the UAS site in all effector constructs (see Methods).

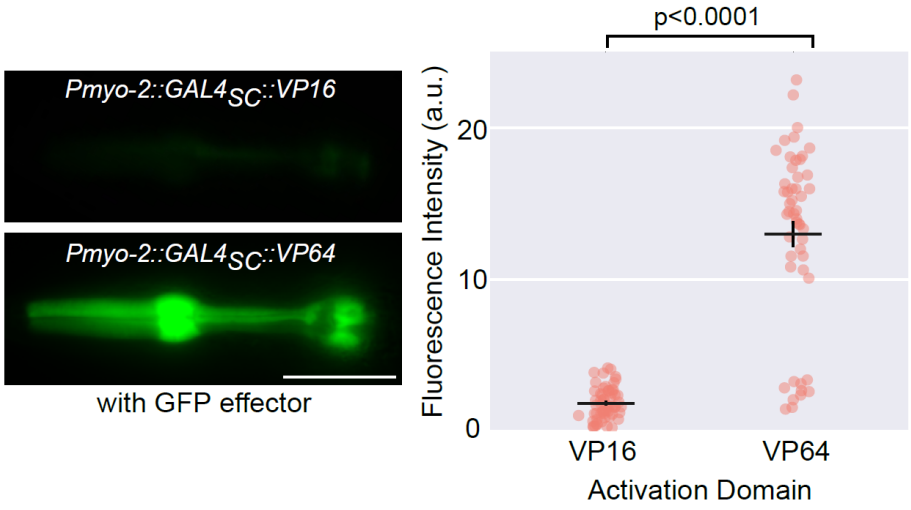


Figure 2.2 | VP64 is superior to VP16 as an activation domain

Left: Representative GFP fluorescence images in the pharynx from an integrated *15xUAS::gfp* effector (*syIs300*), injected with *Pmyo-2*::*GAL4*$_{SC}$*::VP16* driver or *Pmyo-2::GAL4*$_{SC}$*::VP64* driver at the same concentration. *Pmyo-2*, the *myo-2* promoter, is specific for expression in pharyngeal muscle. Scale bar is 20 μm.

Right: Quantitative analysis of GFP fluorescence in the pharynx at room temperature, using pharyngeal muscle drivers with VP16 (n = 58) or VP64 (n = 50), coupled with the same *15xUAS::gfp* effector transgene (*syIs300*). Two independent lines for each driver were used for quantification, with n = ~25 for each line. Bars are mean ± SEM. Two-tailed *t*-test with Welch's correction. a.u., artificial units.

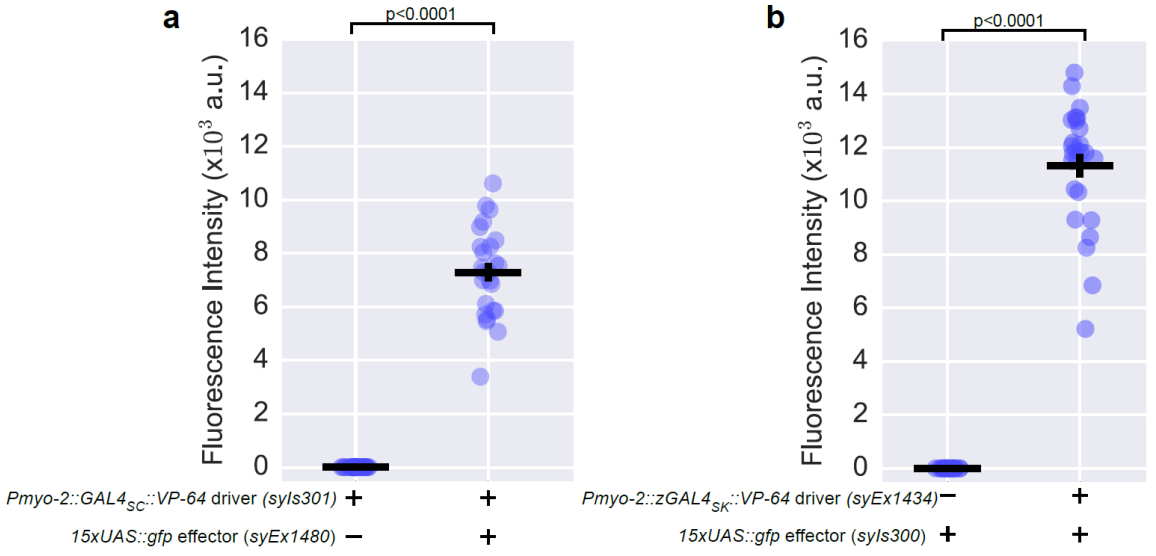


Figure 2.3 | Both components of the cGAL system are required for the expression of an effector gene

Neither the parental *Pmyo-2::GAL4$_{SC}$::VP64* driver alone (*syIs301*, *left*) nor the parental *15xUAS::gfp* effector alone (*syIs302*, *right*) produces expression of GFP in the pharynx. Expression of the effector (GFP) is only seen with both components combined. Bars are mean ± SEM. Two-tailed *t*-test with Welch's correction. a.u., artificial units.

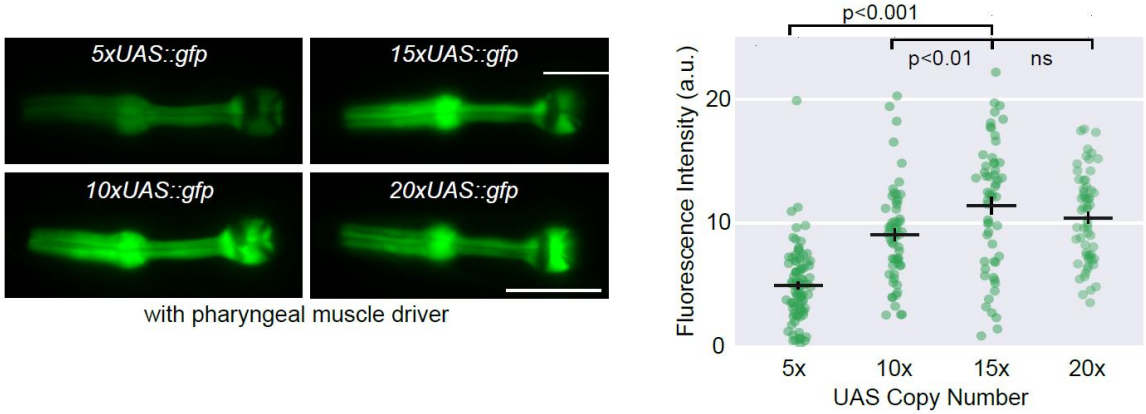


Figure 2.4 | Optimization of UAS copy number

Left: Representative images showing GFP fluorescence in the pharynx of transgenic animals with an integrated *Pmyo-2::GAL4$_{SC}$::VP64* driver (*syIs301*), injected with 5x, 10x, 15x, and 20x copies of the UAS site upstream of *gfp* effector constructs at same concentration. Scale bar is 20 μm.

Right: Quantitative analysis of GFP fluorescence the effector lines with different UAS copy numbers, coupled with the same pharyngeal driver (*syIs301*). n = 89, 58, 56, 60, from left to right. Two to three independent lines were used to quantified, with n = ~30 for each line. Bars are mean ± SEM. ns, not significant. One-way ANOVA with Tukey's post-test. a.u., artificial units.

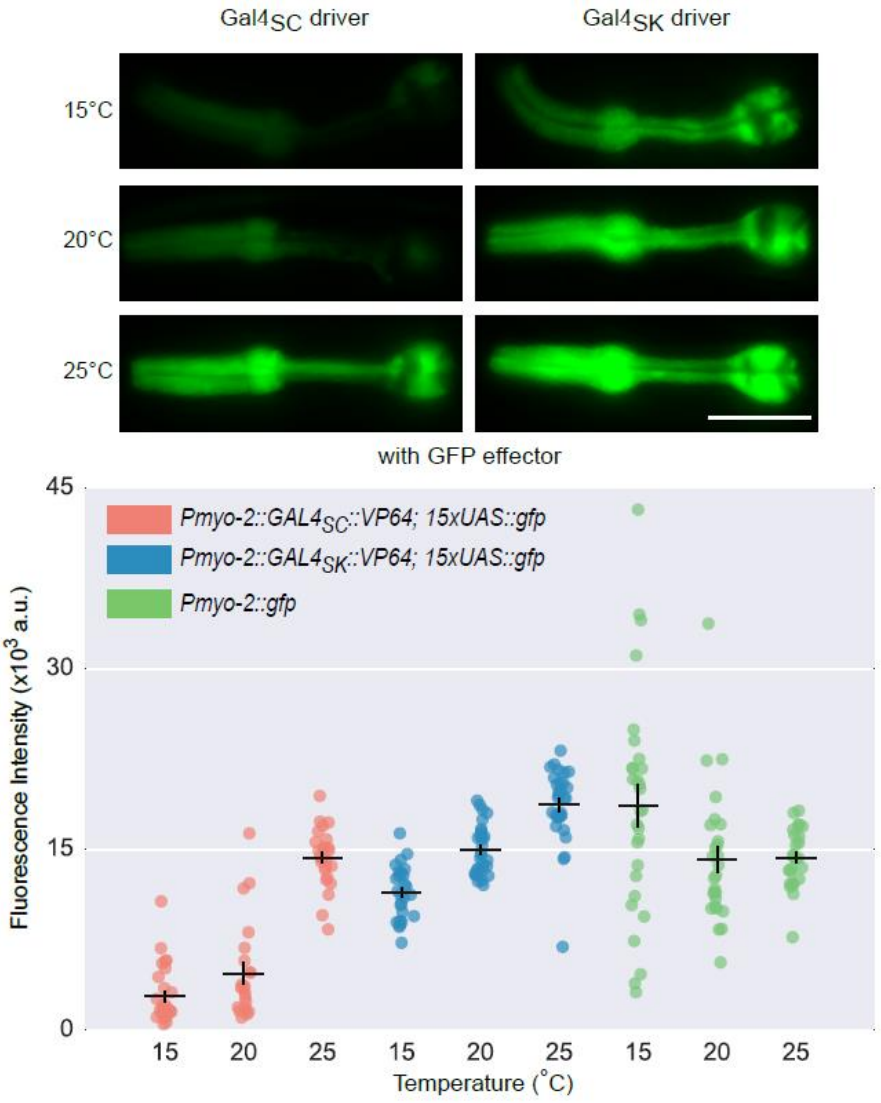


Figure 2.5 | Designing a temperature-robust GAL4 driver via evolutionary analysis

Upper: Representative images showing GFP fluorescence in the pharynx from an integrated *15xUAS::gfp* effector (*syIs300*), injected with same concentration of *Pmyo-2::GAL4$_{SC}$::VP64* (*syEx1434*) or *Pmyo-2::GAL4$_{SK}$::VP64::unc-54 3'UTR* (*syEx1436*), respectively, at 15°C, 20°C, and 25°C. Scale bar is 20 μm.

(continued from **Figure 2.5***)*

Lower: Quantitative analysis of GFP expression of $Gal4_{SC}$ and $Gal4_{SK}$ pharyngeal muscle drivers, along with a *Pmyo-2::gfp* transcriptional fusion (*syEx1437)* across three temperatures. We refer to drivers using $Gal4_{SK}$ and VP64 as cGAL drivers for the rest of this paper. All three arrays were generated by injecting at the same concentration (10 ng/µL). n = ~25 for each condition. Bars are mean ± SEM. All pairwise comparisons within each genotype are significant (** $p<0.01$ or lower), except $Gal4_{SC}$ 15°C vs 20°C, and *Pmyo-2::gfp* 20°C vs 25°C. Two-way ANOVA, Tukey's multiple comparison test. a.u., artificial units.

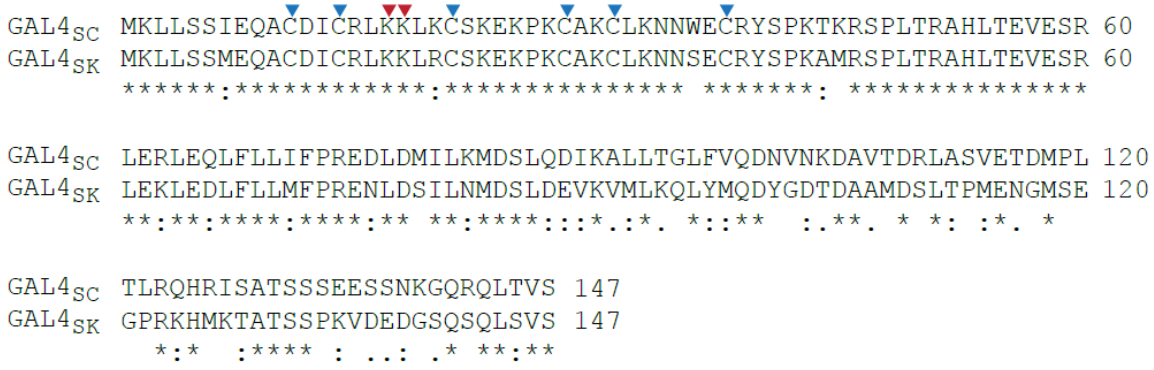


Figure 2.6 | Alignment of *S. cerevisiae* and *S. kudriavzevii* GAL4 DNA-binding domains

Alignment of the DNA binding domains (residues 1-147) of Gal4p sequences from *S. cerevisiae* ($GAL4_{SC}$) and *S. kudriavzevii* ($GAL4_{SK}$). Two critical, conserved DNA-interacting lysine residues are marked with red arrowheads. Six cysteines forming the conserved Zn_2Cys_6 binuclear cluster are marked with blue arrowheads. Asterisk (*), colon (:), and period (.) indicate identical residues, strongly conserved residues and weakly conserved residues, respectively. The alignment was done with the software Clustal Omega[75].

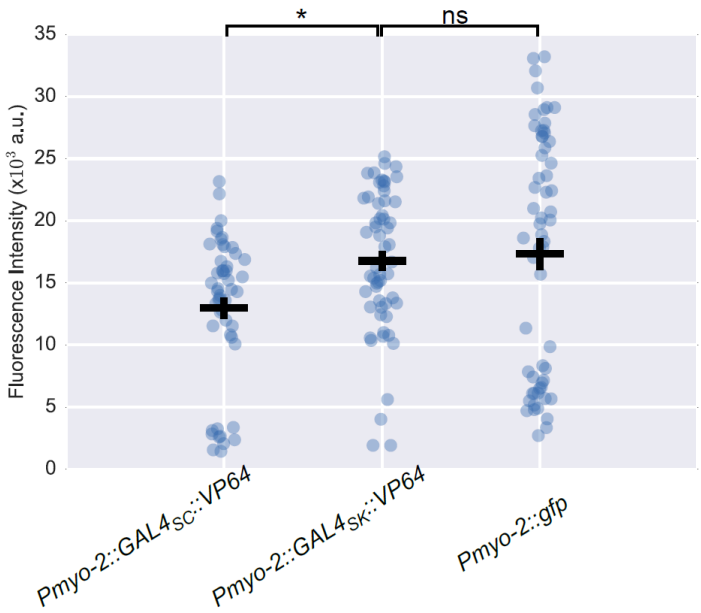


Figure 2.7 | Performance of different DBDs from Gal4 proteins at room temperature

Quantification of GFP fluorescence in the pharynx of transgenic worms with either *Pmyo-2::GAL4$_{SC}$::VP64* or *Pmyo-2::GAL4$_{SK}$::VP64* drivers injected into a strain carrying an integrated *15xUAS::gfp* transgene (*syIs300*) at room temperature (22-23°C). The drivers were both injected at 10 ng/µL. Strains with a direct *Pmyo-2::gfp* fusion array at 10 ng/µL was measured for comparison. Two independent lines were imaged for each genotype. n = 20 - 30 for each line. Bars are mean ± SEM. * $p < 0.05$. ns, not significant. One-way ANOVA with Tukey's post-test. a.u., artificial units. DBD, DNA-binding domain.

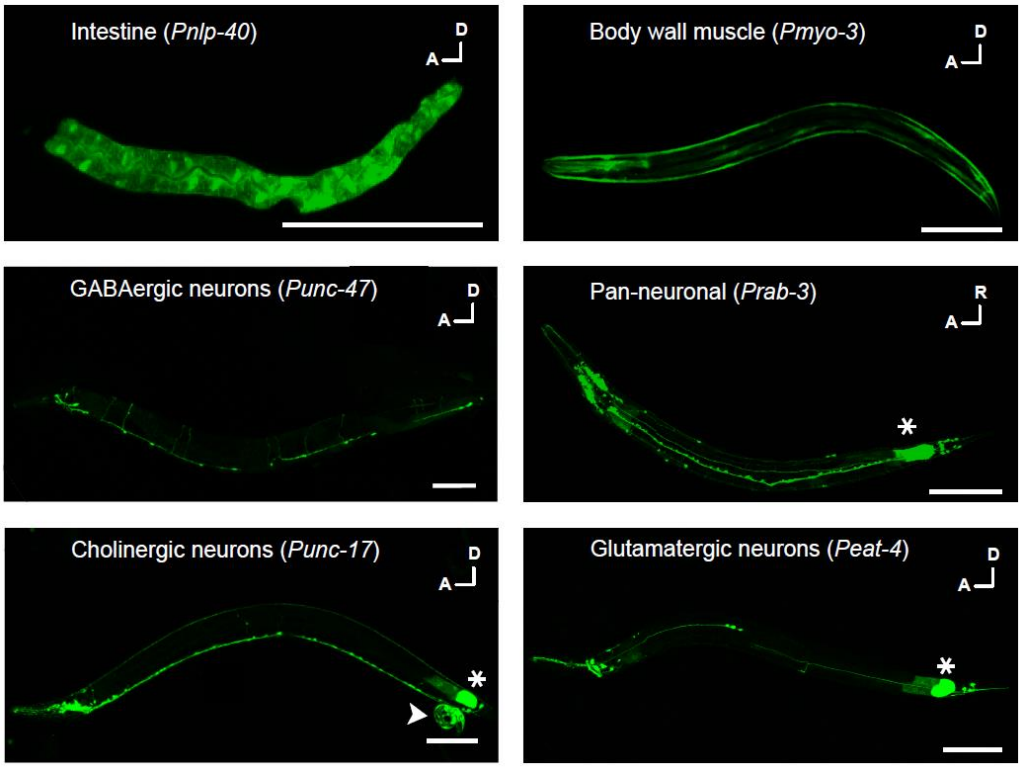


Figure 2.8 | Robust activity of the cGAL system across multiple tissues

Strong GFP fluorescence was observed in corresponding tissues dictated by the promoters used in cGAL drivers, coupled with integrated *15xUAS::gfp* lines. Representative images are shown for the intestine and body wall muscle (*top row*), pan-neuronal and GABAergic neurons (*middle row*), cholinergic and glutamatergic neurons (*bottom row*). Note that in intestinal and body way muscle images, both drivers and the effector (*15xUAS::gfp::unc-54 3'UTR*) are built in the pSM vector. In the middle and bottom rows, both drivers and the effector (*15xUAS::gfp::let-858 3'UTR*) are built in the pPD117.01 vector (see Methods and Supplementary Table 1). Asterisk (*) indicates the canonic ectopic GFP fluorescence in the posterior gut. All scale bars are 100 μm. A, anterior; D, dorsal; R, right.

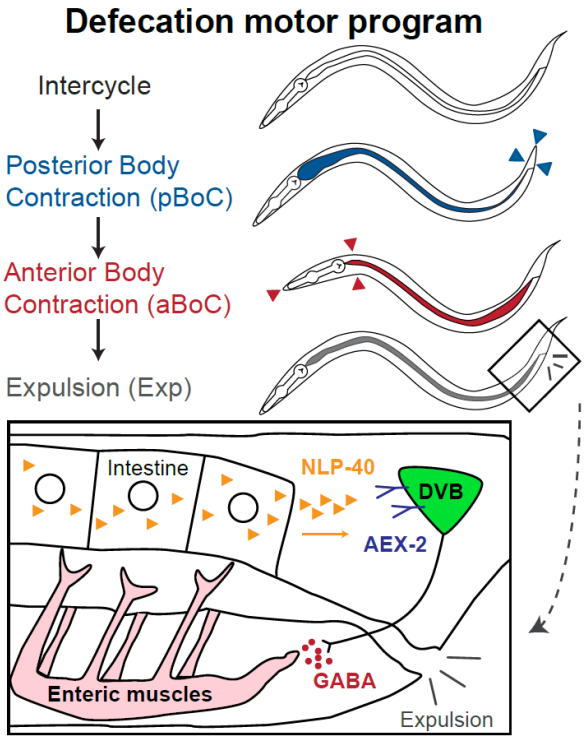


Figure 2.9 | The *C. elegans* defecation motor program

The defecation motor program (DMP) repeats about once per minute and consists of a posterior body wall contraction, followed by an anterior body wall contraction and then a final expulsion step. Inset: the intestine releases the neuropeptide NLP-40, which acts on its receptor AEX-2 in two GABAergic neurons DVB and AVL (AVL not shown), which in turn release GABA to activate enteric muscles to trigger expulsion.

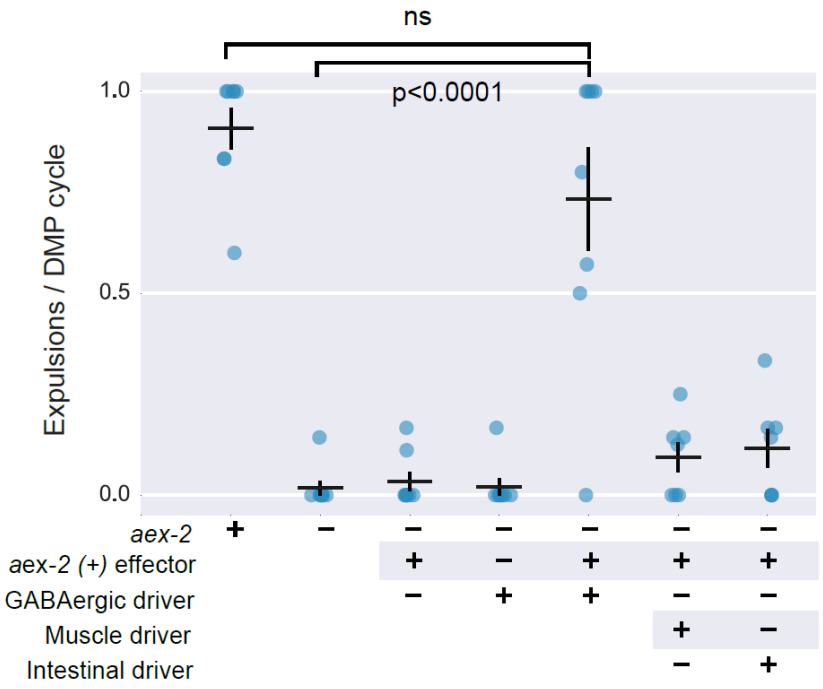


Figure 2.10 | Tissue-specific rescue of *aex-2* for DMP expulsion events

Quantification of expulsion events per defecation cycle in animals with indicated genotypes. Each of the integrated drivers (*Pnlp-40::GAL4$_{SK}$::VP64* for intestine, *Punc-47::GAL4$_{SK}$::VP64* for GABAergic neurons and *Pmyo-3::GAL4$_{SK}$::VP64* for muscles) was crossed with the same *15xUAS::aex-2(+)* cDNA effector line (*syEx1444*) in the *aex-2(sa3)* mutant background. All constructs contained the *unc-54 3'UTR*. Bars are mean ± SEM. n = 8-10 for each genotype. ns, not significant. One-way ANOVA with Tukey's post-test.

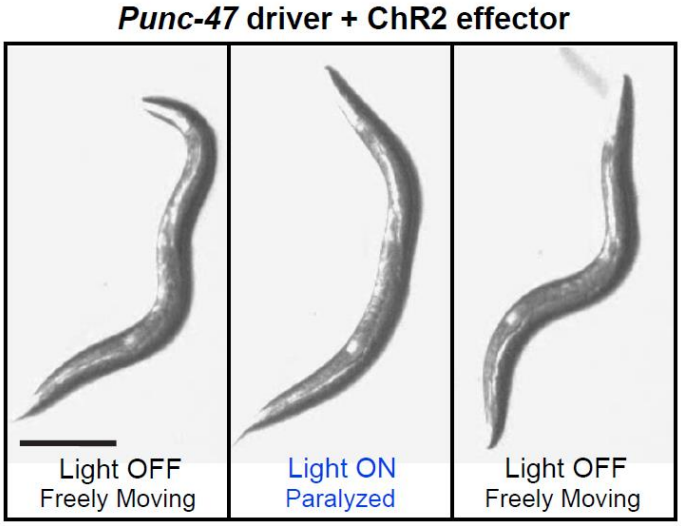


Figure 2.11 | Channelrhodopsin activation in GABAergic neurons results in paralysis

Gain-of-function channelrhodopsin experiment in GABAergic neurons using cGAL. Shown in the figure are three images of the same transgenic worm with both a GABAergic (P*unc-47*) driver and a channelrhodopsin (ChR2) effector from a video recording taken before, during, and after exposure to blue light. Activation of GABAergic neurons produces a limp, paralyzed body posture. Scale bar is 200 μm.

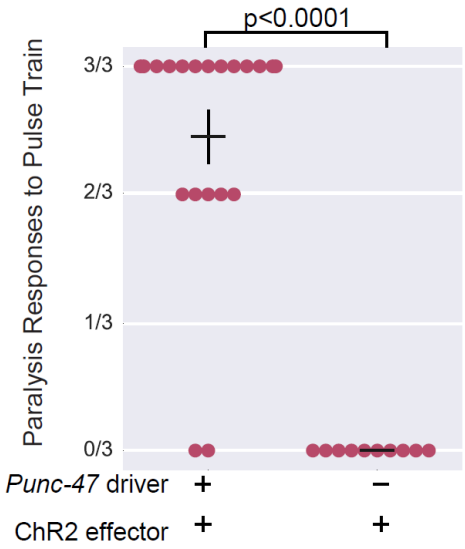


Figure 2.12 | Quantification of GABAergic activation-mediated paralysis

Quantification of light-induced paralysis in worms with the indicated drivers/effectors. Each dot represents an individual animal and its mean response to 3 blue light exposures. Bars are mean ± SEM. n = 20 and 10 for first column and second column, respectively. Mann-Whitney test.

Driver/Promoter	Strain	Genotype	Linkage Group
Pharyngeal Muscle driver, Pmyo-2	PS6844	*syIs301*	LGV
Pharyngeal Muscle driver, Pmyo-2	PS7154	*syIs391*	LGIV
Intestine driver, Pnlp-40	PS6916	*syIs317*	LGIII
Intestine driver, Pnlp-40	PS6933	*syIs318 syIs302*	LGIII
Intestine driver, Pnlp-40	PS6934	*syIs319*	Not on LGIII
Intestine driver, Pnlp-40	PS6935	*syIs320*	Not on LGIII
Body muscle driver, Pmyo-3	PS6936	*syIs321*	Not on LGIII
Pan-neuronal driver, Prab-3	PS6961	*syIs334*	LGX
Pan-neuronal driver, Prab-3	PS6962	*syIs335*	LGX
Pan-neuronal driver, Prab-3	PS6963	*syIs336*	LGX
GABAergic neuron driver, Punc-47	PS7160	*syIs393*	LGIV
GABAergic neuron driver, Punc-47	PS7166	*syIs395 syIs337*	LGIII
GABAergic neuron driver, Punc-47	PS7167	*syIs396 syIs337*	LGIII
Heat shock driver, Phsp16.41	PS7169	*syIs398; syIs337*	
Heat shock driver, Phsp16.41	PS7170	*syIs399; syIs337*	
Heat shock driver, Phsp16.41	PS7171	*syIs400; syIs337*	
Heat shock driver, Phsp16.41	PS7172	*syIs401; syIs337*	
Heat shock driver, Phsp16.41	PS7173	*syIs402; syIs337*	

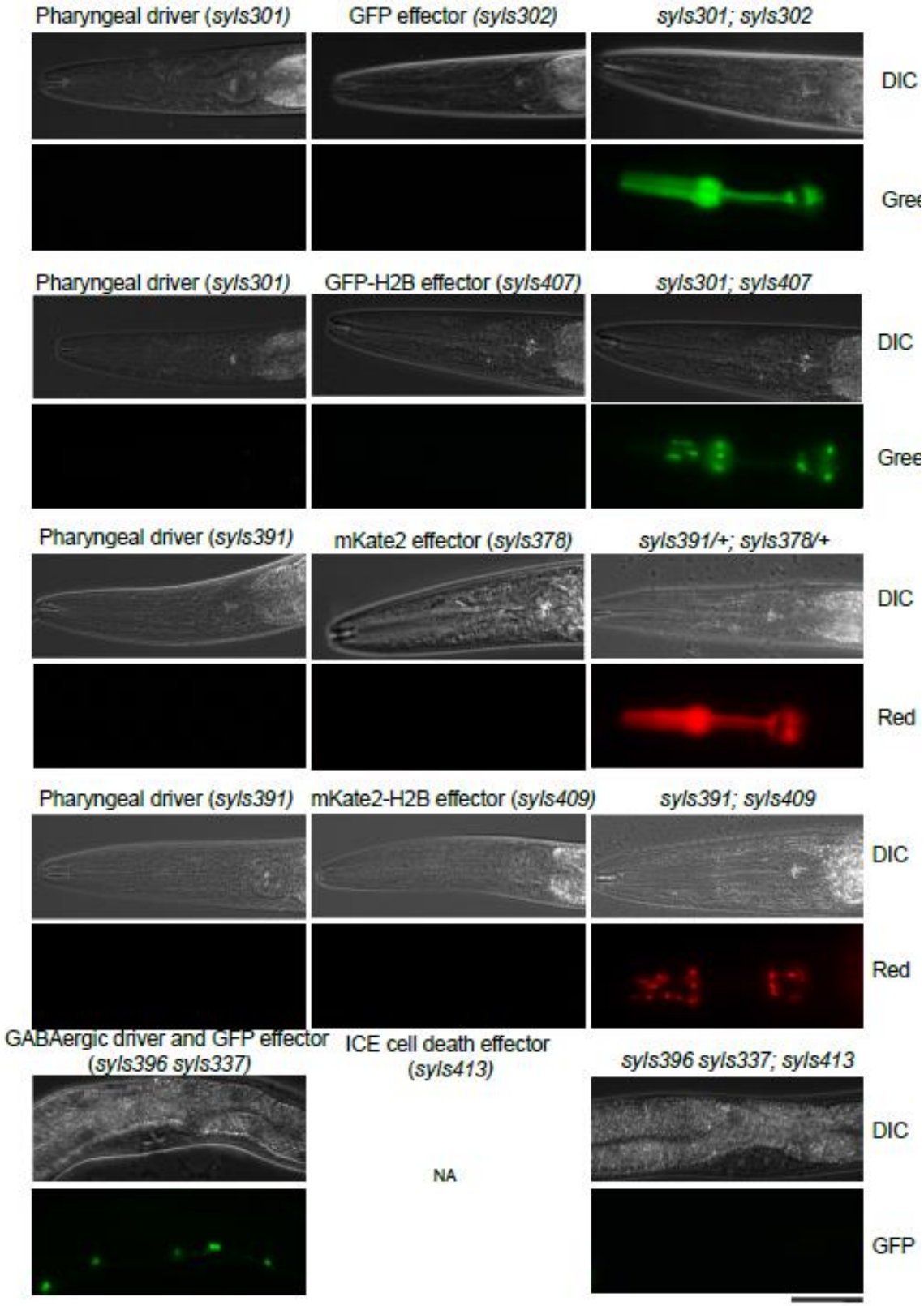
Table 2.1 | Table of integrated cGAL drivers

Abbreviated table of integrated cGAL drivers. A full list can be found by searching elegans.caltech.edu

Effectors	Use	Strain	Genotype	Linkage Group
15xUAS::GFP	cell labeling	PS6843	*syIs300*	LGV
15xUAS::GFP	cell labeling	PS6872	*syIs302*	LGIII
15xUAS::GFP	cell labeling	PS6974	*syIs337*	LGIII
15xUAS::GFP	cell labeling	PS7149	*syIs390*	
15xUAS::GFP	cell labeling	PS7198	*syIs419*	LGIV
15xUAS::mKate2	cell labeling	PS7110	*syIs376*	
15xUAS::mKate2	cell labeling	PS7111	*syIs377*	
15xUAS::mKate2	cell labeling	PS7136	*syIs378*	
15xUAS::mKate2	cell labeling	PS7137	*syIs379*	
15xUAS::mKate2	cell labeling	PS7138	*syIs380*	
15xUAS::GFP::H2B	cell labeling	PS7185	*syIs406*	LGIV
15xUAS::GFP::H2B	cell labeling	PS7186	*syIs407*	
15xUAS::GFP::H2B	cell labeling	PS7187	*syIs408*	LGIII
15xUAS::mCherry::H2B	cell labeling	PS7190	*syIs409*	LGX
15xUAS::GCaMP6s::SL2::mKate2	calcium indicator	PS7203	*syIs423*	LGV
15xUAS::GCaMP6s::SL2::mKate2	calcium indicator	PS7205	*syIs424*	LGIII
15xUAS::GCaMP6s::SL2::mKate2	calcium indicator	PS7206	*syIs425*	LGV
15xUAS::GCaMP6s::SL2::mKate2	calcium indicator	PS7207	*syIs426*	
15xUAS::GCaMP6s::SL2::mKate2	calcium indicator	PS7208	*syIs427*	LG III
15xUAS::hChR2(H134R)::yfp	neuronal activation	PS7043	*syIs340*	
15xUAS::hChR2(H134R)::yfp	neuronal activation	PS7044	*syIs341*	
15xUAS::hChR2(H134R)::yfp	neuronal activation	PS7045	*syIs342*	
15xUAS::HisCl1::SL2::GFP	neuronal inhibition	PS7199	*syIs371*	
15xUAS::HisCl1::SL2::GFP	neuronal inhibition	PS7107	*syIs373*	
15xUAS::HisCl1::SL2::GFP	neuronal inhibition	PS7108	*syIs374*	
15xUAS::TeTx	blocking synaptic transmission	PS7200	*syIs420*	LGIV
15xUAS::TeTx	blocking synaptic transmission	PS7201	*syIs421*	
15xUAS::TeTx	blocking synaptic transmission	PS7202	*syIs422*	LGV
15xUAS::ICE	cell ablation	PS7192	*syIs413*	LGIV
15xUAS::ICE	cell ablation	PS7193	*syIs414*	LGI
15xUAS::ICE	cell ablation	PS7194	*syIs415*	LGI
15xUAS::ICE	cell ablation	PS7195	*syIs416*	LGII
15xUAS::ICE	cell ablation	PS7196	*syIs417*	LGX

Table 2.2 | Table of integrated cGAL effectors

Abbreviated table of integrated cGAL effectors for cell labeling, recording, activation, inhibition, and ablation. A full list can be found by searching elegans.caltech.edu


Pharyngeal driver (syIs301)
GFP effector (syIs302)
syIs301; syIs302
DIC
Gree
Pharyngeal driver (syIs301)
GFP-H2B effector (syIs407)
syIs301; syIs407
DIC
Gree
Pharyngeal driver (syIs391)
mKate2 effector (syIs378)
syIs391/+; syIs378/+
DIC
Red
Pharyngeal driver (syIs391)
mKate2-H2B effector (syIs409)
syIs391; syIs409
DIC
Red
GABAergic driver and GFP effector (syIs396 syIs337)
ICE cell death effector (syIs413)
NA
syIs396 syIs337; syIs413
DIC
GFP

(continued from previous page)

Figure 2.13 | Functional verification of integrated effectors

Expression of integrated drivers (*left column*) or integrated effectors alone (*middle column*) shows no basal expression Only the combination (*right column*) show expression of cytoplasmic or nuclear-localized reports, or death of appropriate cells. DIC, Differential interference contrast. Green, 530 nm green channel. Red, 630 nm red channel. Scale bar is 20 μm.

2.5 METHODS

Maintenance of *C. elegans* Strains

Strains were maintained on NGM plates with *E. coli* OP50 as the food source at room temperature as originally described[1], unless noted otherwise. Bristol strain N2 is the wild-type reference strain. The full list of strains used in this study is detailed in the supplementary information.

Molecular Biology

Plasmids were constructed by standard molecular cloning techniques with either restriction enzyme cleavage and DNA ligation or Gibson assembly using enzymes from New England Biolabs (Beverly, MA). The coding region of $Gal4_{SK}$, residues 1-147 of Gal4p from *Saccharomyces kudriavzevii* (the Portuguese reference strain ZP591, a gift from C. T. Hittinger), was PCR amplified from genomic DNA using the primers:

Forward: 5'-ggaGCTAGCatgaagctgttgtcttcaatgg-3'

Reverse: 5'-cggGAATTCcggcgatacactcaactgactttggc-3'

The synthetic ScaI-17mer sequence (CGGAGTACTGTCCTCCG)[76] was used for the UAS site and was placed upstream of the *pes-10* basal promoter in all effector constructs. All constructs were built in either the pSM vector, a derivative of pPD49.26, which contains the *unc-54* 3'UTR, or the vector pPD117.01 (a gift from A. Fire), which contains the *let-858* 3'UTR and a 5' decoy minigene upstream of the MCS for promoter insertion. Details on plasmids and oligos used in the study are documented in Supplementary Table 1 and Supplementary Table 2.

Transformation

Transgenic animals were generated using standard microinjection techniques[8]. Unless noted otherwise, 100 ng/µl total DNA injection samples were prepared, with either the pBluescript II KS+ plasmid or 1 kb DNA ladder, from New England Biolabs (Beverly, MA), as carrier. Extrachromosomal arrays were integrated into the genome via X-ray irradiation. Most of the integrants were outcrossed at least three times with the wild type strain N2. Full details about transgenic *C. elegans* strains in this study are listed in the Supplementary Information.

Fluorescence imaging

Approximately 25 animals were imaged and quantified for the optimization process of the cGAL system, using the *myo-2* promoter (*Pmyo-2*). Briefly, L4 or young adults animals grown at corresponding temperatures (15°C, 20°C, 25°C or room temperature) were selected and imaged with Leica DMI600 inverted microscope equipped with 40x oil objective and an Andor iXon Ultra 897 EMCCD camera, using Metamorph software (Molecular Devices). Images were captured with the same exposure time (20ms) and the average fluorescence in the pharynx for each animal was analyzed. The representative fluorescent images in Fig. 4 showing the application of cGAL in different tissues were collected with a Zeiss LSM710 confocal microscope with a 20x objective.

Defecation motor program assay

L4 animals raised at room temperature were picked one day before the assay. During the assay, which was performed at 20°C, each individual worm was picked to a new NGM

plate seeded with OP50 and a 18x18mm coverslip was placed over the animal for better optics. After a two-minute acclimation period, each animal was videotaped for five minutes and the number of pBoc and expulsion events were scored. Each pBoc indicated the initiation of each defecation cycle. The ratio of expulsions over pBocs was used to quantify the expulsion phenotype for each animal (n = 8-10 for each genotype).

Optogenetics

One day before the assay, L4 animals raised at room temperature from each strain were picked individually onto NGM plates, seeded with 100 μL OP50 containing 500 μM all-*trans* retinal (Sigma). During the assay, which was performed at 20°C, animals were recorded using a Zeiss Stemi SV11 coupled to a Unibrain Fire-i 501b camera. Channelrhodopsin was activated using blue light generated from a Lumen Dynamics X-Cite series 120 lamp and a standard GFP filter set. Blue 475 nm light intensity was measured to be 0.2 mW/mm^2. After an initial 10-second acclimation period, three light pulses, each 2 seconds in duration, were delivered to each worm at intervals of 20 seconds. The researcher doing the assay was blinded to the genotype of the animals.

Chapter 3

An Intersectional Split Strategy using Split Inteins for Single Cell-type Genetic Access

This chapter can be found published as an article in:

Wang H, **Liu J**, Yuet KP, Hill AJ, Sternberg PW. Split cGAL, an intersectional strategy using a split intein for refined spatiotemporal transgene control in Caenorhabditis elegans. Proc Natl Acad Sci U S A. 2018 Apr 10;115(15):3900-3905. doi: 10.1073/pnas.1720063115.

3.1 ABSTRACT

Bipartite expression systems, such as the GAL4-UAS system, allow fine manipulation of gene expression and are powerful tools for interrogating gene function. We established cGAL, a GAL4-based bipartite expression system for transgene control in *C. elegans*, where a single promoter dictates the expression pattern of a cGAL driver, which then binds target UAS sequences to drive expression of a downstream effector gene. In this chapter, we report a split strategy for cGAL using the split intein gp41-1 for intersectional control of transgene expression. Split inteins are protein domains that associate, self-excise, and covalently ligate their flanking peptides together. We split the DNA-binding domain (DBD) and transcriptional activation domain (AD) of cGAL and fuse them to the N-terminal of gp41-1-N-intein and the C-terminal of gp41-1-C-intein, respectively. In cells where both halves of cGAL are expressed, a functional cGAL driver is reconstituted via intein-mediated protein splicing. This reconstitution allows expression of the driver to be dictated by two promoters for refined spatial control or spatiotemporal control of transgene expression. We apply the split cGAL system to genetically access the single pair of MC neurons (previously inaccessible with a single promoter), and reveal an important role of protein kinase A (PKA) in rhythmic pharyngeal pumping in *C. elegans*. Thus, the split cGAL system gives researchers a greater degree of spatiotemporal control over transgene expression and will be a valuable genetic tool in *C. elegans* for dissecting gene function with finer cell-specific resolution.

3.2 INTRODUCTION

A fundamental goal of biology is to understand how an organism uses its full complement of genes to determine its development, morphology, cellular and tissue functions, and behaviors. Each gene may act in different cells and at different times for various biological processes. Thus, genetic tools that enable precise control of gene expression both spatially and temporally are extremely valuable for dissecting gene function. With its powerful genetics and small size, *C. elegans* is an important genetic model for studying various biological processes and has contributed to the understanding of fundamental mechanisms underlying biology [77]. While a variety of tissue- and cell-specific promoters have long been available to the *C. elegans* community, genetic access for each individual cell type, especially each anatomical neuron type, has not yet been achieved. Providing this type of access would allow much finer resolution of genetic analysis, accelerating full dissection of gene function and understanding of the biology of the worm.

The previous chapter described cGAL, a GAL4-based bipartite expression system, for controlling transgene expression in *C. elegans*[78]. As with other bipartite expression systems, in which a driver specifies the expression pattern of the transgene and an effector dictates the nature of the transgenic perturbation, the cGAL system uses the DNA binding domain (DBD) from a cryophilic yeast strain *Saccharomyces kudriavzevii* and the synthetic VP64 activation domain (AD) for the driver. The cGAL driver triggers expression of the effector gene by binding upstream activation sequence (UAS) sites only in cells in which the promoter used in the driver construct is active (**Figure 3.1**).

However, the extent of transgene control with cGAL in *C. elegans* is limited by available promoters, because the expression pattern is dictated by the single promoter used in each driver. In particular, the majority of neurons in *C. elegans* are not genetically accessible with single promoters[79], which hinders our understanding of the functional importance of different genes and neurons for different behaviors. Furthermore, it is generally impossible to achieve spatial and temporal regulation of transgene expression at the same time using a single promoter in cGAL drivers. Here, we addressed these limitations of the original cGAL system by designing an intersectional split cGAL strategy that provides a logical 'AND' gate for refined transgene control using two distinct promoters (**Figure 3.1**).

Deletion mutant studies in yeast analysis showed that the original Gal4p protein from *Saccharomyces cerevisiae* has two functional modules, the DNA binding domain (DBD) and transcription activation domain (AD)[32,80]. Independently, neither is sufficient to drive the expression of the effector gene downstream of UAS sites. Luan *et al.*[33] took advantage of this modular independence and designed a split GAL4 system for the *Drosophila* community by fusing the DBD and AD to one half of an antiparallel leucine zipper adapter pair, and putting the fusions under the control of two different promoters. With this design, only in cells where both promoters are active would both components be expressed, and the two antiparallel leucine zipper adapters allow the DBD and AD to associate via non-covalent interactions to reconstitute a functional GAL4 driver[33] (**Figure 3.2**). This 'split' system gives spatially restricted expression of GAL4 in cells at the intersection of two promoters, and has made split GAL4 a powerful tool for *Drosophila* researchers to precisely control transgene expression, particularly in the nervous system[33,81].

Since the introduction of the split GAL4 system in *Drosophila*, new adapter protein domains that mediate covalent interactions have been discovered. One example is the SpyTag/SpyCatcher system, an adapter system engineered from *Streptococcus pyogenes*[82]. When proteins are tagged with this adapter pair, SpyTag/SpyCatcher will associate and form a covalent isopeptide bond, uniting their protein partners together[82].

Another example is a class of protein domains called inteins. The first intein was characterized in the yeast *VMA1* gene as an internal portion of the protein (the 'intein') that was capable of simultaneously self-excising and mediating intra-molecular ligation of the two flanking sequences (termed 'exteins') in *cis*[83–85]. Later, sequence analysis in cyanobacteria revealed the presence of split inteins, which could mediate protein splicing in *trans*[86,87]. Here, two separate genes encode two peptide products, each having one extein and one half of the split intein. The two peptides associate via their split intein domains and undergo inter-molecular protein splicing, excising the two split intein domains and fusing the two exteins via a peptide bond[83]. The split inteins gp41-1 and *Npu* DnaE are amongst the most robust and fastest described in the literature[88,89]. However, neither SpyTag/SpyCatcher nor these split inteins have been tested in assembling a functional GAL4 driver from the DBD and AD domains.

With the sole exception of split intein gp41-1, the other three protein adapters described above have been reported to successfully associate proteins in *C. elegans*[79,90–93]. To establish a robust split cGAL system, we systematically compared the efficiency of all four adapters in re-associating the DBD and AD domains to reconstitute a functional cGAL

driver (**Figure 3.2**). We determine that the gp41-1 split intein is the best adapter for our split cGAL system. We also show that split cGAL allows simultaneous spatiotemporal control or refined spatial control of transgene expression in *C. elegans* with two different promoters. Finally, we apply our split cGAL system to reveal a critical role of protein kinase A (PKA) in the single pair of cholinergic MC pharyngeal neurons in the feeding behavior of *C. elegans*.

3.3 RESULTS

3.3.1 Comparing protein adapter domains for reconstitution of the split cGAL driver

To construct our split system, we first wished to test the ability of different adapters to reconstitute a functional cGAL driver from its two modular halves (DBD and AD). We separated our cGAL driver into its modular components and appended one of four different adapters (anti-parallel leucine zipper, SpyTag/SpyCatcher, *Npu* DnaE split intein and gp41-1 split intein, **Figure 3.2**). One half, the cGAL(DBD)-adapter, contained the DNA binding domain from *S. kudriavzevii* and an adapter domain; the other half, adapter-cGAL(AD), contained the cognate adapter domain and the VP64 activation domain (**Figure 3.1**). After placing each gene fusion under the control of a pharyngeal muscle-specific promoter (*myo-2* promoter), we then injected each pair of split cGAL-adapter constructs together at equal concentrations into a transgenic strain with an integrated *15xUAS::gfp* effector (*unc-119 (ed3); syIs300*) and performed quantitative fluorescence imaging to assess GFP levels in pharyngeal muscles.

To our surprise, we found that neither the anti-parallel leucine zipper nor the SpyTag/SpyCatcher could reconstitute the split cGAL(DBD) and cGAL(AD) to drive expression of GFP in pharyngeal muscles (**Figure 3.3**, not statistically significant compared to effector only, one-way ANOVA with Tukey's correction, $p > 0.9999$), although both adapters have been shown to bring together other proteins successfully in *C. elegans* [38,79,90,91]. By contrast, both intein adapters restored the transcriptional activity of split cGAL (**Figure 3.3**, statistically significant compared to effector only, one-way ANOVA with Tukey's correction, $p < 0.0001$). Split cGAL with the DnaE and gp41-1 adapters achieved 37% and 72% of transcriptional activator activity of the intact cGAL, respectively. The gp41-1 intein in particular showed the brightest and most robust expression.

To rule out the possibility that the high level of GFP expression observed with gp41-1-mediated split cGAL is due to recombination of the injected DNA constructs that might have generated an intact cGAL driver fragment in the extrachromosomal array, we injected and integrated each split cGAL driver transgene separately. When individually crossed to the GFP effector (*syIs300*), neither half produced fluorescence. We only observed GFP fluorescence in the cross progeny containing both cGAL halves and the GFP effector transgene demonstrating that split drivers are essential for driving the expression of the effector (**Figure 3.4**).

To test whether successful reconstitution of gp41-1-mediated split cGAL is dependent on protein splicing, we mutated the first cysteine of gp41-1 N-intein to alanine (referred to as

cGAL(DBD)-gp41-1-N-intein (C1A)). The first amino acid (cysteine or serine) of the intein is essential for the first step of intein-mediated protein trans-splicing [94]. We predicted that cGAL(DBD)-gp41-1-N-intein(C1A) would not be able to join cGAL(DBD) and cGAL(AD) together when combined with gp41-1-C-intein-cGAL(AD). Indeed, we found that unlike wild type *Pmyo-2::cGAL(DBD)-gp41-1-N-intein*, the mutated version of *Pmyo-2::cGAL(DBD)-gp41-1-N-intein(C1A)* could not drive GFP expression in pharyngeal muscles when injected into a strain with both *Pmyo-2::gp41-1-C-intein-cGAL(AD)* and *15xUAS::GFP* (**Figure 3.5**). Based on these results, we concluded that the split cGAL system with the intein gp41-1 is most effective for reconstituting a functional cGAL in *C. elegans*. From this point onwards, we will refer to cGAL(DBD)-gp41-1-N-intein as cGAL-N, and to gp41-1-C-intein-cGAL(AD) as cGAL-C, unless stated otherwise.

3.3.2 Spatial and temporal control with split cGAL

During development, genes are turned on at different times to perform their functions. The determination of critical time windows for such genes requires genetic tools that provide temporal control of transgene expression. The use of heat shock promoters is a common way to impart temporal control but it sacrifices spatial control; conversely, tissue-specific promoters in transgenes generally cannot provide temporal control at the same time. We explored the possibility that the split cGAL system could simultaneously achieve spatial and temporal control of transgene expression. As we reported before, cGAL driver constructs in the Fire vector pPD117.01 containing the *let-858* 3'UTR are more robustly expressed than in the Fire vector pPD49.26 with the *unc-54* 3'UTR[78]. Thus, we built new split cGAL drivers in the backbone with the *let-858* 3'UTR. We used a heat shock

promoter (*hsp-16.41* promoter) to drive the expression of cGAL-N, and the constitutive pharyngeal muscle promoter (*myo-2* promoter) to drive the expression of cGAL-C. In cross progeny that were triple heterozygotes for *Phsp-16.41::cGAL-N, Pmyo-2::cGAL-C* , and *15xUAS::gfp*, no GFP expression in pharyngeal muscles was observed without heat shock (**Figure 3.6**). Starting 4 hours after heat shock treatment (33°C for 1 h), we observed a steady increase of GFP expression in pharyngeal muscles all the way up to 16 h after heat shock, the last time point that we assayed (**Figure 3.6**). This induction of GFP effector seemed to be relatively slow, comparing to that from a direct heat shock promoter::GFP fusion (see discussion). Furthermore, we also showed that the conditional expression of the GFP effector in pharyngeal muscles after heat shock required both P*hsp-16.41* and P*myo-2* split cGAL drivers (**Figure 3.7**).

After heat shock, we observed noticeable background GFP expression in the excretory cell in animals containing *Phsp-16.41::cGAL-N* and *15xUAS::gfp*, but not in those containing *Pmyo-2::cGAL-C* and *15xUAS::gfp* (**Figure 3.8,** upper). In the presence of the *15xUAS::gfp* effector, this ectopic expression of GFP in the excretory canal cell was also observed in worms carrying the cGAL-N split driver under control of the ubiquitous *eft-3* promoter (**Figure 3.8,** lower) but not those with *myo-2* promoter (pharyngeal muscle promoter), *rab-3* promoter (pan-neuronal promoter), or *unc-17* promoter (cholinergic neurons), suggesting that the cGAL-N split driver may interact with an unknown transcriptional activator that is specifically expressed in the excretory cell and thus can non-specifically drive the effector gene in this cell. We did not observe *Peft-3::cGAL-C* alone drove ectopic GFP expression in the excretory cell (5 independent lines). Thus, if the

ectopic expression in excretory cell of the promoter in the cGAL-N driver is a concern, the promoter can be swapped to drive cGAL-C instead.

3.3.3 Refined spatial control with split cGAL

C. elegans has 302 neurons in the adult hermaphrodite and 385 neurons in the adult male[95,96]. Despite the relative simplicity of these nervous systems however, many anatomical neuron types cannot be genetically approached using single promoters. As the expression pattern of many *C. elegans* genes are well-characterized, it has been suggested that most anatomical neuron types in *C. elegans* can be genetically accessed with the intersection of two different promoters[79].

We wanted to determine if gp41-1-mediated split cGAL could be used as an 'AND' gate to spatially restrict transgene expression with two overlapping promoters (**Figure 3.1**). We were interested in the regulation of pharyngeal pumping in *C. elegans* by the single pair of MC neurons. Previous work with laser ablation showed that MC neurons are the major excitatory neurons for fast pumping[97,98]. However, there were no previously described single promoters that gave specific access to this neuron type. Thus, we chose to design split cGAL drivers to access the MC neurons. We could validate these split cGAL drivers by crossing them with our existing neuronal effector strain kit [78] to manipulate the activity of MC neurons and examining pumping rate and growth.

The *unc-17* and *ceh-19*b promoters are proposed to specifically overlap in the MC neurons (http://www.wormweb.org/neuralnet). We made two split cGAL constructs *Punc-*

17::cGAL-N and *Pceh-19b::cGAL-C*, and injected both together in the integrated HisCl1 effector line (*syIs371, 15xUAS::HisCl1::SL2::gfp*). HisCl1 encodes a histamine-gated chloride channel, capable of silencing neurons when histamine is applied[74]. As expected, we observed bright green fluorescence in the pair of MC neurons (**Figure 3.9**), suggesting the two split cGAL components successfully reconstituted and drove the expression of both HisCl1 and GFP in MC neurons. However, we also found an additional pair of neurons with weaker GFP fluorescence, likely to be the sensory ADF neurons, as both *unc-17* and *ceh-19b* were also reported to be expressed in ADF [99,100] (**Figure 3.10**).

To functionally validate the MC split cGAL driver, we silenced MC neurons expressing the HisCl1 channel by exposing animals to 10 mM histamine and quantified pharyngeal pumping rate. Animals with both the MC split cGAL driver and the HisCl1 effector raised on 10 mM histamine from hatching grew up to be thinner and less pigmented than counterparts raised in the absence of histamine (**Figure 3.11,** *upper*). Those animals treated with histamine also pumped much slower (62.3 ± 2.8 pumps/min, mean ± SEM, n = 11), compared to the worms of the same genotype but not treated with histamine (217.2 ± 5.3 pumps/min, mean ± SEM, n = 10, **Figure 3.11,** *lower*). We also find that either half split cGAL driver for MC was not sufficient to driver the HisCl1 effector to inhibit pumping in the presence of histamine (**Figure 3.12**). This result is consistent with previous observations in worms with laser-ablated MC neurons[97,98].

3.3.4 Regulation of pharyngeal pumping by protein kinase A in *C. elegans*

Protein kinase A (PKA) is one of the major targets of the second messenger cyclic adenosine monophosphate (cAMP)[101]. The PKA holoenzyme is a tetramer, consisting of two catalytic subunits and two regulatory subunits. In absence of cAMP, the kinase activity of the catalytic subunits is inhibited by the regulatory subunits. When cAMP levels increase, cAMP binds to the regulatory subunit, leading to its dissociation from the catalytic subunit and subsequent disinhibition of PKA[101] (**Figure 3.13**). In *C. elegans*, the catalytic and regulatory subunits of PKA are encoded by *kin-1* and *kin-2*, respectively. Null mutants for both *kin-1* and *kin-2* are lethal[102,103], preventing detailed genetic analysis of PKA signaling in *C. elegans*.

Genetic studies using partial loss-of-function *kin-2* mutants revealed that PKA signaling in the nervous system is involved in the regulation of pharyngeal pumping in *C. elegans*[103,104]. As the MC neurons are the major excitatory motor neurons for pharyngeal pumping[97,98], we hypothesized that normal PKA activity in the MC neurons is necessary for rapid pumping. To test this hypothesis, we used the split cGAL system to block PKA activity specifically in MC neurons. We first created an effector strain with an extrachromosomal array of *15xUAS::kin-2a(G310D)::SL2:::gfp.* The G310D mutation in isoform a of the regulatory subunit KIN-2 prevents its binding with cAMP, thereby maintaining its inhibitory interaction with the catalytic subunit KIN-1 even when cAMP is elevated, and produces a dominant negative form of PKA[105,106] (**Figure 3.13**). Neither this effector strain nor the split cGAL driver strain for MC neurons showed any defect in pharyngeal pumping rate. However, cross progeny from these two parent strains displayed a 33% decrease in pharyngeal pumping rate (162.3 ± 9.8 pumps/min vs. 243.6 ± 5.9 and 250.3 ± 2.5

pumps/min for driver and effector alone, respectively. Mean ± SEM, **Figure 3.14**). These results support the conclusion that PKA signaling in MC neurons is essential for normal fast pharyngeal pumping in *C. elegans*.

3.3.5 Discussion

In this study, we describe our development of a novel split cGAL system using the intein gp41-1 to mediate protein splicing and produce a transcriptionally competent cGAL driver from its split components (DBD and AD). We demonstrate that split cGAL can achieve refined spatial and spatiotemporal control of transgene expression in *C. elegans* using two separate promoters. We also build a cell-type specific split cGAL driver to specifically manipulate PKA activity in the MC pharyngeal neurons, and discover that inhibiting the PKA pathway in MC neurons results in a decrease in pumping rate.

To engineer the split cGAL system, we experimented with four methods of reconstituting cGAL DBD and AD and determine that among the four adapters tested, gp41-1 is the most effective, recapitulating over 70 percent of the intact cGAL driver's performance (**Figure 3.3**). Several reasons may explain this. First, gp41-1 brings the DBD and AD together with a canonical peptide bond. Second, the kinetics of gp41-1-mediated protein splicing are fast, about 10 times faster than *Npu* DnaE[88]. This may explain why gp41-1 outperformed DnaE. Third, similar to DnaE, gp41-1 excises itself and leaves a minimal peptide sequence between DBD and AD of the reconstituted cGAL after protein splicing. Reconstitution of cGAL using the leucine zipper or SpyTag/SpyCatcher results in larger extraneous protein domains between the DBD and AD, likely leading to spatial and steric constraints with

negative functional consequences. Compared to the intact cGAL driver, our split constructs with the DnaE and gp41-1 inteins have additional 13 and 14 amino acids between the cGAL DBD and AD, in contrast to the 80 and 126 amino acids for the leucine zipper and SpyTag/SpyCatcher. This, together with other factors, such as weak association and/or poor kinetics, may account for the failure of leucine zipper and SpyTag/SpyCatcher to successfully reconstitute split cGAL. In support of this explanation, Luan *et al.*[33] observed a 48% reduction in function of the *S. cerevisiae* GAL4 transcriptional activator when split with the leucine zipper in *Drosophila*[33]. Our results suggest that the gp41-1 intein would be an excellent tool for other aspects of protein engineering for *C. elegans* requiring protein splicing. For example, a similar split Q bipartite system using the same leucine zipper pair as we tested was also described in *C. elegans*[38]. It would be interesting to determine if the gp41-1 intein can boost the performance of the split Q system.

Although gp41-1-mediated protein splicing is fast[88], expressing a gene using split cGAL is likely to introduce a temporal delay, when compared to a direct heat shock promoter::gene fusion. The heat shock promoter must first drive expression of one half of cGAL driver, which then must undergo protein splicing with the other cGAL driver half before driving the expression of the effector gene. However, this delay comes at the benefit of adding spatial control of the heat shock promoter, should the experimenter need conditional expression in a restricted subset of cells where the *hsp-16.41* promoter is active. The gp41-1-mediated protein splicing rate *in vivo* is likely to be dependent on the concentration and stoichiometry of both split cGAL components within the cell, as well as temperature. This complexity has unique implications for using the split cGAL system to achieve

spatiotemporal control of gene expression. For instance, the temperature and the duration of the heat shock protocol to induce transgene expression will influence the timing and the expression level of the split cGAL component under the control of the heat shock promoter. The cellular environment of different cell types may also influence the time scale of the intein-mediated protein splicing. Thus, in studies where timing of gene expression is a critical factor, we recommend characterizing the temporal dynamics of a split driver combination with GFP or any other fluorescently tagged effector.

Our study on PKA in MC neurons highlights the importance of genetic tools that allow highly refined spatial control of gene activity, which ultimately will help in understanding the cell-specific roles of genes. PKA has not been reported in forward genetic screens for mutants that are defective in pumping, likely due to the fact that the null alleles of *kin-1* and *kin-2* are lethal. When studying lethal or toxic alleles, expression often must be limited to a small subset of cells. Precise and systemic expression of these alleles is most efficiently achieved with bipartite systems. We used a split cGAL driver and a dominant negative PKA effector and showed that PKA activity in MC neurons is necessary for normal pumping. This finding is in line with a previous observation that serotonin potentiates pumping rate by activating the G protein-coupled receptor SER-7 and $G_{\alpha}s$ signaling to increase cholinergic transmission from MC neurons[103]. PKA may regulate pumping rate by modulating the firing rate of MC neurons or controlling the release of acetylcholine from MC neurons.

With its precision of transgene control, we expect that split cGAL will be particularly useful in providing genetic access to cell types that could not be accessed before, such as many of the individual anatomical neuron types in *C. elegans*. Besides providing highly specific genetic targeting, split cGAL can be used to perturb gene activity and cellular processes with great efficiency, since split cGAL drivers can be reused in combination with UAS effectors by crossing. For example, with our recently published UAS effector toolkit that contains effector strains to manipulate and record neuronal activity[78], any new split cGAL drivers for a neuron type can be crossed with these effector strains to interrogate the function of the neuron in a relevant behavior. Similarly, if new UAS effector strains are generated (e.g. overexpression/interesting alleles of native *C. elegans* genes), they can be crossed to currently available split cGAL drivers to test gene function in cells of interest. As more strains are built, documented, and described, they will contribute to a growing repository of tried and true reagents available to the community at large for extensive, rigorous, and rapid analysis of neural circuits and gene function in *C. elegans*. Furthermore, split cGAL can also be combined with several other systems that have been developed for spatiotemporal control of transgene expression in *C. elegans*[38,42,72,73,107–110] to further improve the precision of transgene expression or achieve orthogonal control of different transgenes to interrogate gene function.

3.4 FIGURES

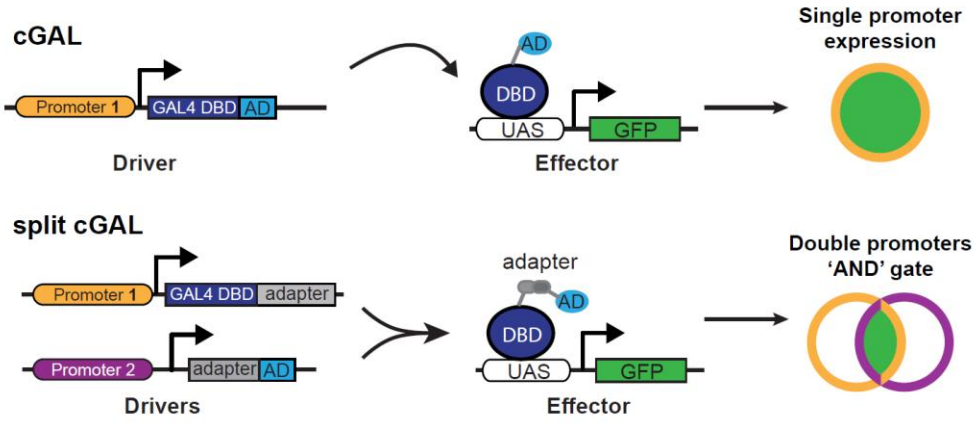


Figure 3.1 | Schematic of cGAL and split cGAL strategies

Upper: The original cGAL bipartite system, in which a single promoter governs expression of the cGAL driver. The driver is composed of the DNA binding domain (DBD) from *S. kudriavzevii* Gal4p, which recognizes upstream activating sequences (UAS), and a transcriptional activation domain (AD) which recruits transcriptional machinery. The cGAL driver then specifies expression of the effector gene (i.e., GFP), under the control of UAS, in cells where the promoter is active.

Lower: Using split strategies, two promoters can be used, providing an "AND" gate to achieve intersectional control of transgene expression. DBD, DNA-binding domain. AD, activation domain.

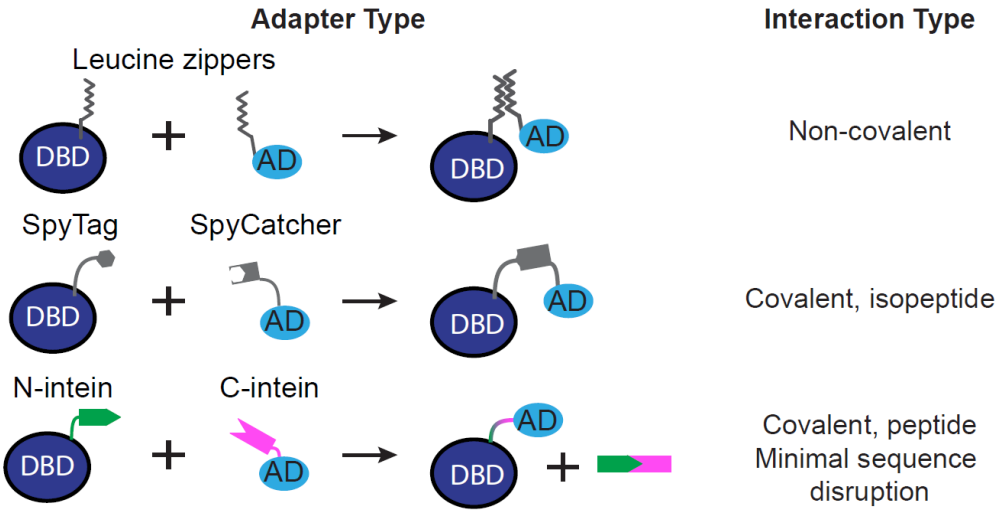


Figure 3.2 | Protein domains that can reconstitute split cGAL components

Splitting the driver components requires a way to reconstitute the split components in cells expressing both. Leucine zippers allow for non-covalent reconstitution of the DBD and AD. The SpyTag/SpyCatcher domains reconstitute via covalent formation of an isopeptide bond. Split intein domains recognize one another and associate, after which they covalently ligate the flanking sequences and self-excise. DBD, DNA-binding domain. AD, activation domain.

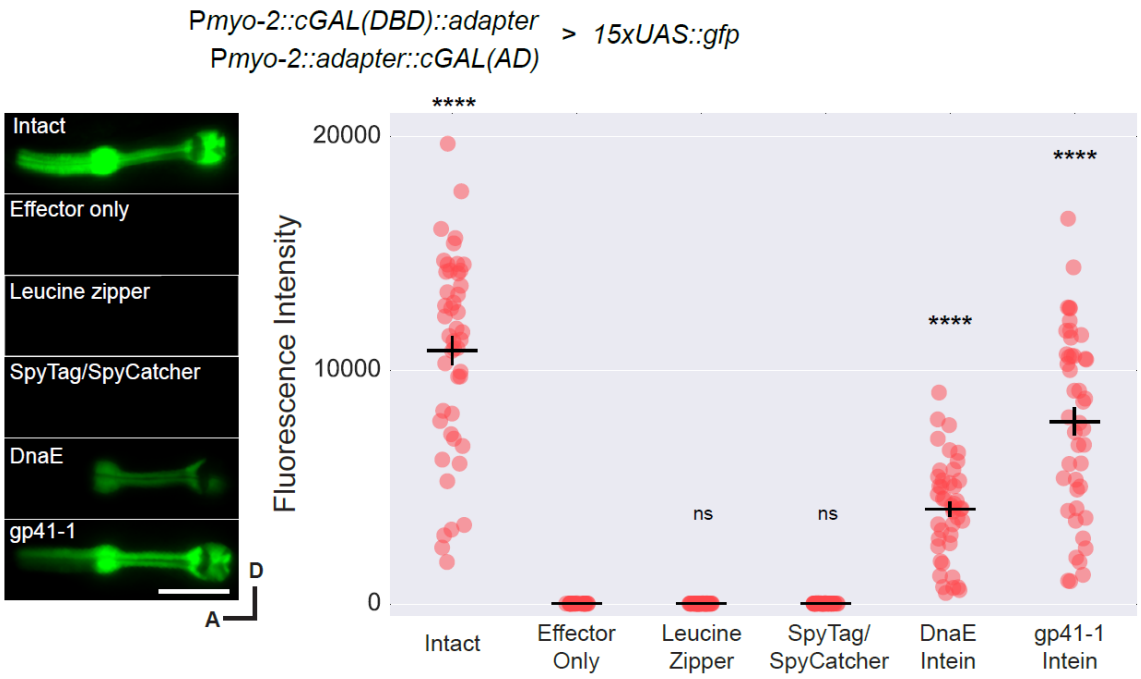


Figure 3.3 | The gp41-1 intein is most efficient in reconstituting split cGAL components

Representative images and quantification of animals with intact cGAL driver, GFP effector only, or the indicated split cGAL driver pairs. The intact cGAL and effector only serve as positive and negative controls. Two independent extrachromosomal transgenic lines were assayed for all groups except the effector alone control, which had only one. Bars are mean ± SEM. From left to right, n = 47, 23, 41, 42, 44, 45. **** $p < 0.0001$, one-way ANOVA and Dunnett's multiple comparisons test to compare the means to the mean of the effector alone. DBD, DNA binding domain. AD, activation domain. D, dorsal. A, anterior. Scale bar is 20 μm.

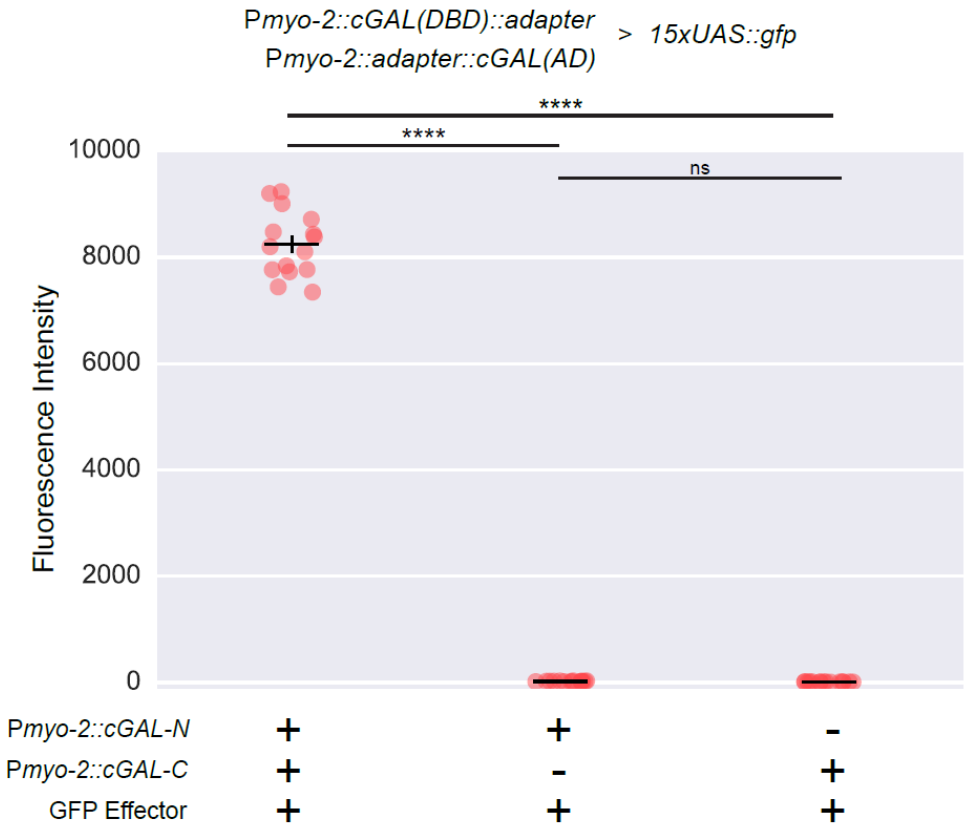


Figure 3.4 | Activation of the GFP effector is dependent on both components of the split cGAL drivers

Quantification of fluorescence in the pharynx of animals with indicated genotypes. All transgenes are integrated into the genome (*syIs431* for *Pmyo-2::cGAL-N*; *syIs433* for *Pmyo-2::cGAL-C*; *syIs300* for GFP effector). +, heterozygote for indicated transgene; -, no indicated transgene. Bars are mean ± SEM. n = 15 for all three genotypes. **** $p < 0.0001$. ns, not significant. One-way ANOVA with Tukey's correction for multiple comparisons. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver.

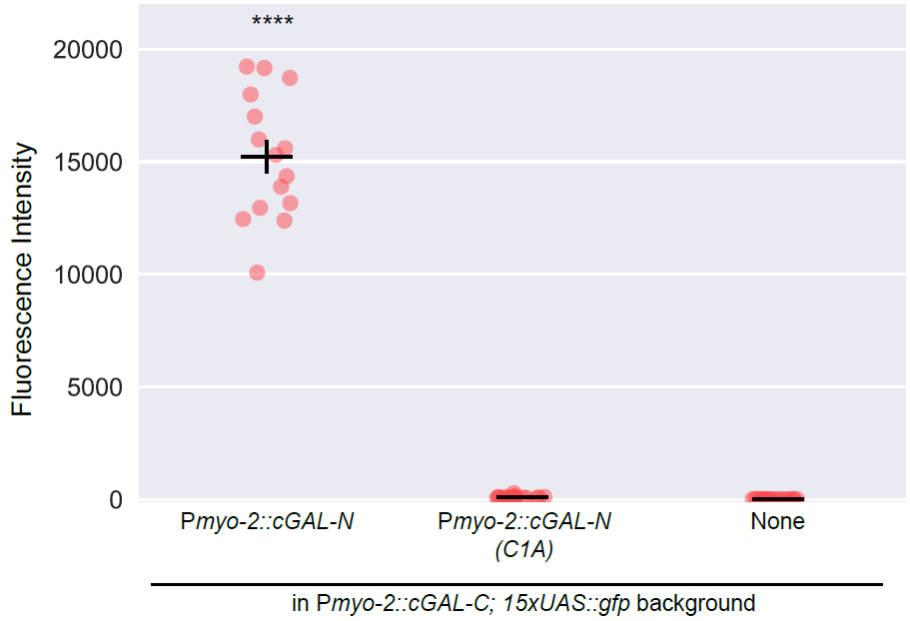


Figure 3.5 | Successful reconstitution of cGAL requires gp41-1-mediated protein splicing

Quantification of fluorescence in the pharynx of animals with indicated genotypes. Mutating the first cysteine of the gp41-1 N-intein to alanine (referred as C1A) disrupts gp41-1-mediated protein splicing. The cGAL-C driver and GFP effector are integrated into the genome (*syIs433* for *Pmyo-2::cGAL-C*; *syIs300* for GFP effector), where P*myo-2::cGAL-N (syEx1589)* and P*myo-2::cGAL-N(C1A) (syEx1590)* are extrachromosomal arrays. Bars are mean ± SEM. For columns from left to right, n = 15, 16, 17. **** $p < 0.0001$. One-way ANOVA with Tukey's correction for multiple comparisons. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver.

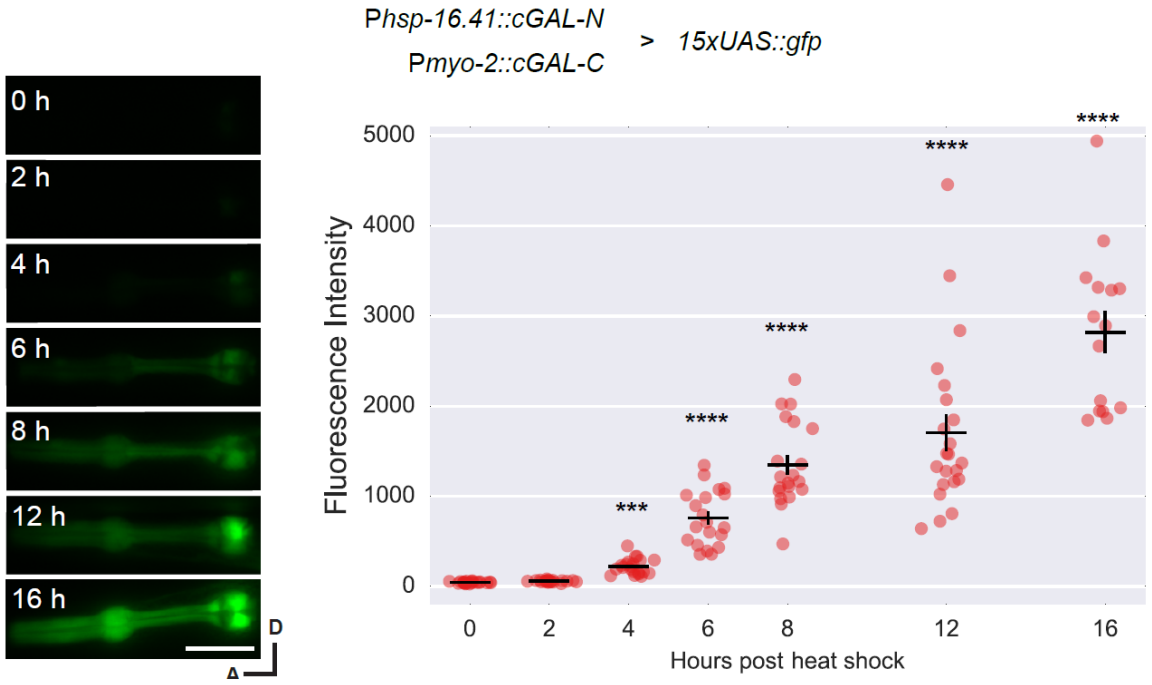


Figure 3.6 | Using split cGAL for spatiotemporal control of gene expression

Representative fluorescence images and quantification of fluorescence in the pharynx of animals that were triple heterozygotes for a conditional cGAL-N driver with the heat shock promoter (*syIs435*), a tissue-specific cGAL-C with the *myo-2* promoter (*syIs433*) and a *15xUAS::GFP* effector (*syIs300*). Bars are mean ± SEM. Each column is a separate group of animals that were imaged at the indicated time point after heat shock, n = 21, 20, 21, 20, 20, 22, and 15 from left to right. *** $p < 0.001$ and **** $p < 0.0001$, One-way ANOVA and Dunnett's multiple comparisons test to compare the means to the mean of no heat-shock control. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver. D, dorsal. A, anterior. Scale bar is 20 μm.

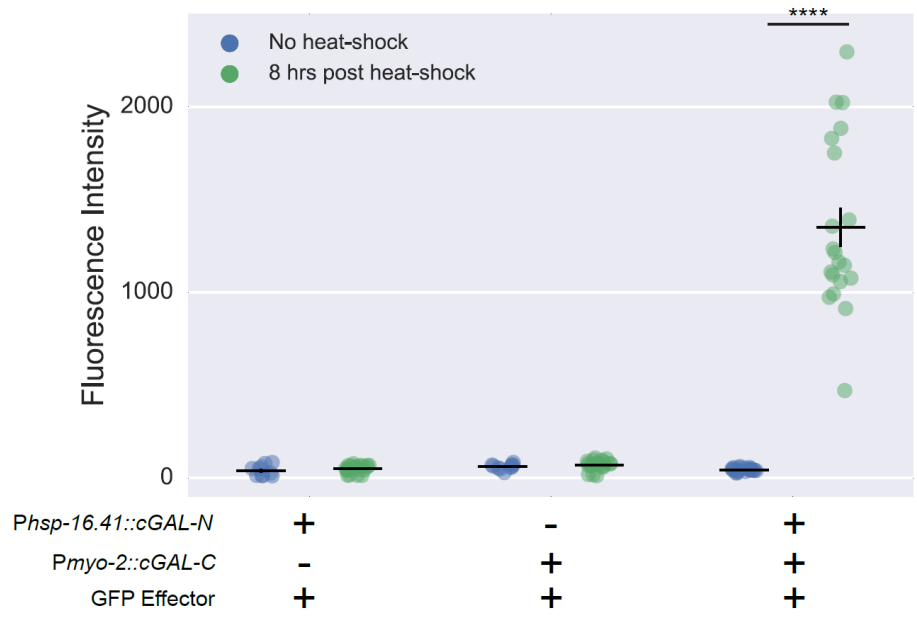


Figure 3.7 | The conditional expression of GFP in pharyngeal muscles required both *hsp-16.41* and *myo-2* split cGAL drivers, in addition to heat shock

Quantification of fluorescence in the pharynx of animals with indicated genotypes, both with and without heat shock. All transgenes are integrated into the genome (*syIs435* for *Phsp16.41::cGAL-N*; *syIs433* for *Pmyo-2::cGAL-C*; *syIs300* for GFP effector). +, heterozygote for indicated transgene; -, no indicated transgene. Bars are mean ± SEM. n = 10, 20, 10, 19, 21, 20 from left to right. **** $p < 0.0001$. Two-way ANOVA with Sidak's correction for multiple comparisons. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver.

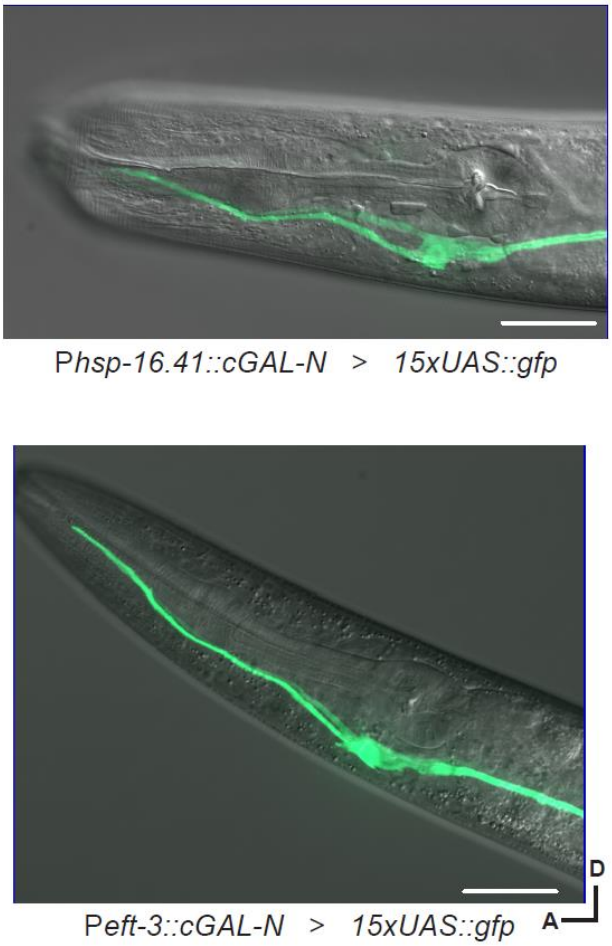


Figure 3.8 | Non-specific expression of GFP in the excretory canal cell

Upper: Merged DIC/GFP image of transgenic worms with *Phsp-16.41::cGAL-N; 15xUAS::gfp* (*syIs435; syIs300*), showing GFP expression in the excretory cell 24 hours after heat shock treatment.

Lower: Merged DIC/GFP image of transgenic worms with *Peft3::cGAL-N; 15xUAS::gfp (syEx1581; syIs300)*, showing GFP expression in the excretory cell.

Scale bar is 20 μm. cGAL-N represents the split cGAL half cGAL(DBD)-gp41-1-N-intein. DIC, differential interference contrast.

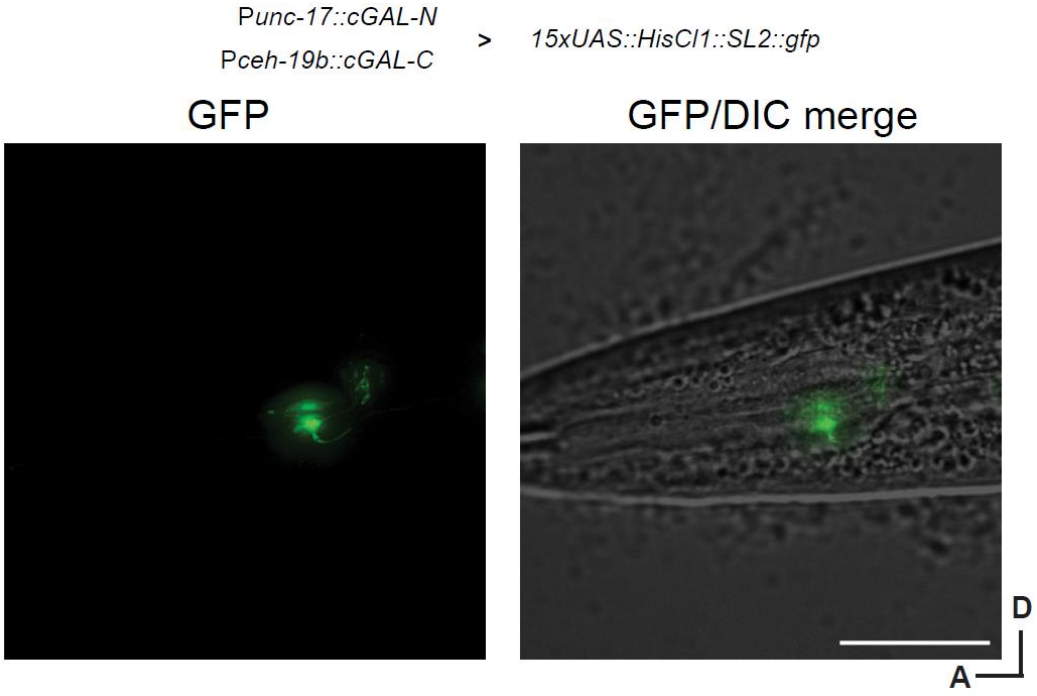


Fig. 3.9 | Using split cGAL for cell-specific expression in MC pharyngeal neurons

Representative images showing specific GFP labeling of bilateral MC motor neurons with the combination of two split cGAL driver constructs using *unc-17* and *ceh-19b* promoters. As indicated, each promoter drives one of the split cGAL components. Co-injection and integration of the components (*syIs483*) is capable of specifically driving a *15xUAS::HisCl1::SL2::gfp* effector (*syIs371*) in MC neurons. D, dorsal. A, anterior**.** DIC, differential interference contrast. Scale bar is 20 μm.

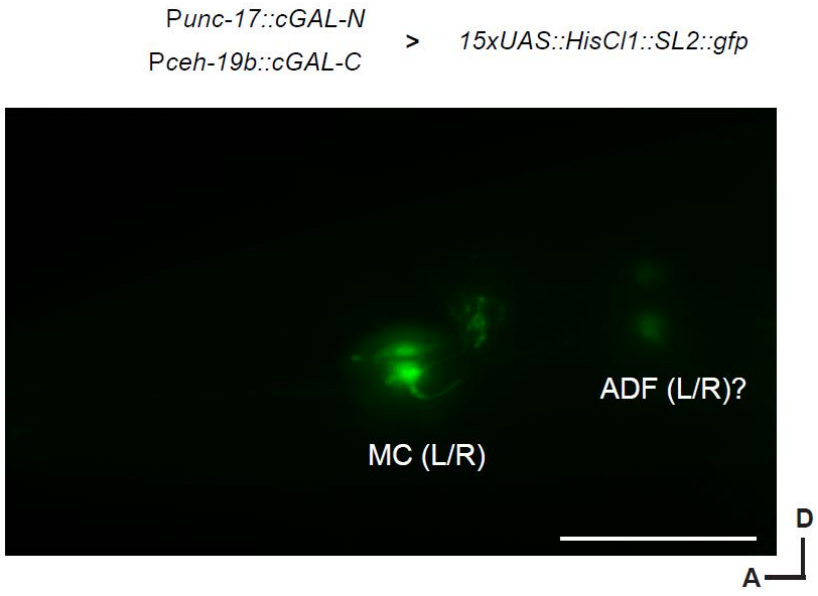


Figure 3.10 | The split cGAL drivers for MC neurons weakly drive expression in ADF

Fluorescence imaging showing transgenic worms with *Punc-17::cGAL-N, Pceh-19b::cGAL-C; 15xUAS::HisCl1::SL2::gfp (syIs483; syIs371),* had strong GFP expression in the MC neurons and weak GFP expression in suspected ADF neurons. Scale bar is 20 μm. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver.

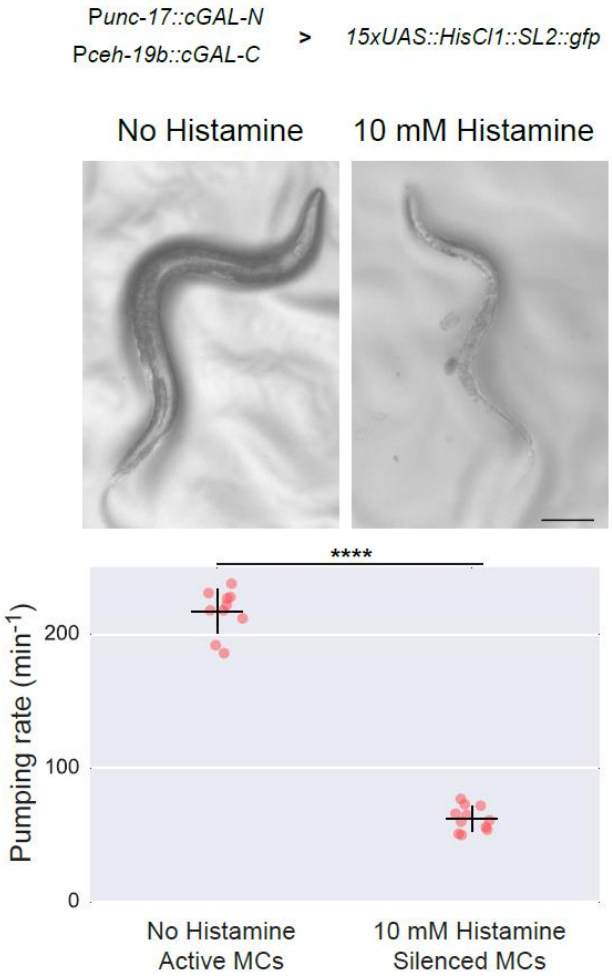


Figure 3.11 | Silencing MCs reduces pumping rate and produces thin, underfed animals

Upper: Light microscopy of *syIs483; syIs371* animals. In the absence of histamine, MC neurons retain their activity and produce pigmented, healthy adults. Raising animals on 10 mM histamine activates the *syIs371* effector to chronically silence the MC neurons, reduces pumping, and produces unhealthy animals with decreased size and pigmentation. Scale bar is 100 μm.

Lower: Quantification of pumping rate of *syIs483; syIs371* animals with or without histamine. Each column represents a separate group of animals with indicated treatments.

(continued from **Figure 3.11***)*

Bars are mean ± SEM. n = 10, 11 from left to right. **** $p < 0.0001$, unpaired Student's *t*-test. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver.

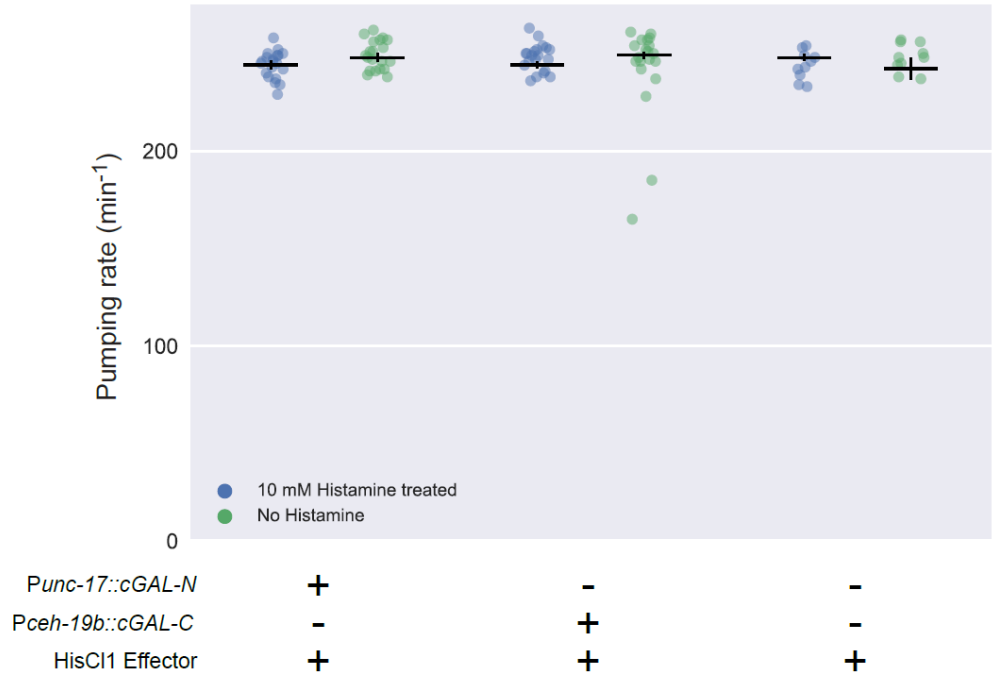


Figure 3.12 | Neither of the MC split cGAL drivers alone is sufficient to reduce pumping rate

Quantification of pumping rate of animals with indicated genotype, treated with or without 10 mM histamine. *Punc-17::cGAL-N* (*syEx1601* and *syEx1602*) and *Pceh-19b::cGAL-C* (*syEx1603* and *syEx1604*) are extrachromosomal arrays, and HisCl1 effector is integrated line (*syIs371*). +, presence of indicated transgene; -, absence of indicated transgene. Bars are mean ± SEM. n = 20, 20, 20, 20, 10 and 10 for columns from left to right. Results are not significant by Two-way ANOVA with Bonferroni's correction. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver.

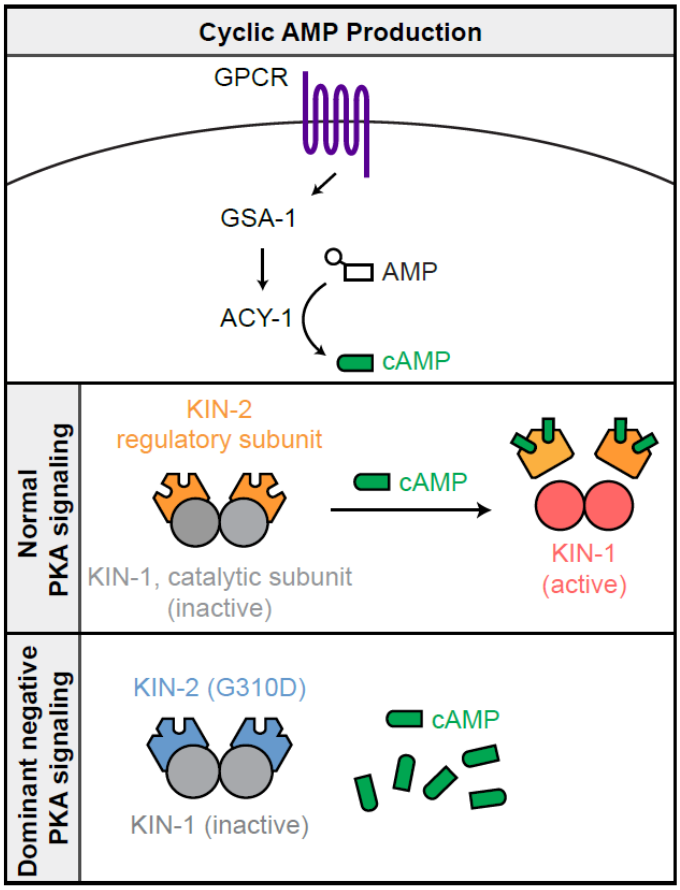


Figure 3.13 | The *C. elegans* Protein Kinase A pathway

Top panel: Diagram of protein kinase A signaling. Ligand binding to a G protein-coupled receptor (GPCR) activates the $G_{\alpha}s$ subunit GSA-1. GSA-1 goes on to activate adenylyl cyclases (i.e. ACY-1), causing conversion of adenosine monophosphate (AMP) to cyclic AMP (cAMP)

Middle panel: In wild-type signaling, cAMP dissociates the inhibitory KIN-2 subunits from the catalytic KIN-1 subunits, leading to the activation of PKA.

Bottom panel: The G310D dominant negative allele of KIN-2 is essentially insensitive to cAMP causing KIN-1 to remain inactive.

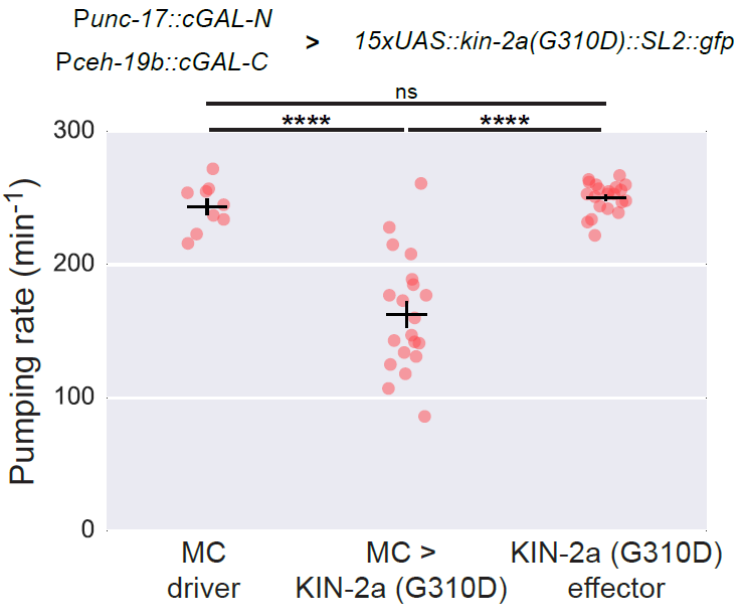


Figure 3.14 | Dominant negative inhibition of protein kinase A signaling in MC neurons reduces pharyngeal pumping rate

Quantification of pumping rate in animals expressing the dominant negative KIN-2a(G310D) in the MC neurons. Here, the *15xUAS::kin-2a(G310D)::SL2::gfp* effector is an extrachromosomal array, and two independent lines were used (*syEx1596* and *syEx1597*). When driven by the *syIs483* driver, both lines showed a significant decrease in pumping rate. Neither driver alone nor effector alone strains displayed aberrant pumping rate. Bars are mean ± SEM. n = 9, 20, 21 from left to right. **** $p < 0.0001$, One-way ANOVA and Tukey's multiple comparisons test to compare all three means. cGAL-N and cGAL-C represent the two halves of the gp-41-1-mediated split cGAL driver.

3.5 METHODS

Strains

The *Caenorhabditis elegans* strains were maintained at 20 °C, as previously described [1] All the strains used in this study are described in detail in Supporting Information.

Molecular biology

All plasmids were constructed in the worm expression vectors pPD49.26 or pPD117.01 from the Fire kit (Addgene). All constructs were generated by standard molecular cloning procedures with restriction digest, PCR, and Gibson assembly or T4 ligation. The coding sequences in the constructs were verified by Sanger sequencing. The complete list of the plasmids and oligos used in this study are listed in Supporting Information (Tables S1-S3).

Transgenic animals

The standard microinjection procedure for *C. elegans* was used to generate transgenic worms with extrachromsomal arrays, some of which were then integrated into the genome using X-ray treatment [111]. The concentrations and compositions of DNA constructs in the injection mixtures of the transgenic worms are described in Supporting Information.

Fluorescence imaging

Worms were paralyzed in M9 buffer supplemented with 30 mg/mL of 2, 3-Butanedione monoxime (Sigma). All fluorescent images for quantification of GFP fluorescence in the pharynx were taken with a Leica DMI6000 inverted microscope equipped with a 40x oil objective and an Andor iXon Ultra 897 EMCCD camera, using Metamorph software

(Molecular Devices). An ROI outlining the entire pharynx, as well as a background ROI, was selected for each image. The background-subtracted mean fluorescence intensity was used to quantify the GFP fluorescence in the pharyngeal muscles of each worm.

Pumping analysis

For the histamine experiments, animals were raised from eggs on regular NGM plates or NGM plates with10 mM histamine dihydrochloride (Sigma), seeded with 150 uL of OP50 bacteria. Gravid animals were bleached, and their eggs were transferred to corresponding plates. Animals were assayed 72 hours later.

For the dominant negative *kin-2* experiments, L4 animals were picked on regular NGM plates seeded with OP50 bacteria. The next day, each adult was transferred to a new NGM plate seeded with OP50 and allowed to acclimate for 10 mins before assaying.

For both experiments, each worm was recorded for 1 min under a Wild Makroskop M420 dissecting microscope equipped with a Unibrain 501b camera. The pumping rate for each worm was determined as total pumping events over the 1-min recording.

Heat shock treatment

L4 cross progeny were picked one day before onto new NGM plates. The next day, plates were sealed with Parafilm and put in a 33 °C water bath for 1 hour with the agar side down. After heat shock, worms were recovered at room temperature and imaged at different times after heat shock.

Statistical Analysis

All the quantification plots were made using custom written Python scripts in Jupyter Notebook [112]. Unpaired Student's *t*-test and one-way ANOVA with Tukey's or Dunnett's tests (GraphPad Prism) were applied when appropriate, as indicated in the figure legends.

Chapter 4

Single copy cGAL and Future Directions for the cGAL Bipartite System

4.1 INTRODUCTION

The prior chapters of this thesis detail engineering of a bipartite cGAL system, functional across multiple tissues using multiple protein effectors. It includes construction of a split intersectional system which allows for more complex spatiotemporal control. We expect this system to greatly further genetic and cellular analysis in *C. elegans*. All data shown previously, however, has been with transgenes consisting of multi-copy extrachromosomal and integrated arrays. While the system in its current form has great utility, multi-copy transgenesis methods have restrictions that limit the full potential of a GAL4 bipartite expression system. The inability to control copy number and (in the case of integrated arrays) inability to control integration site pose challenges to reproducibility during the re-use of strains.

As an example, a common practice is to determine driver expression patterns using a cellular reporter as an effector, e.g. GFP. However, the final expression of GFP is not only dependent upon the activity of the driver, but also the local genomic context of where the effector is embedded. In integrated arrays, it is not possible to control where the effector transgene is integrated, and therefore the local genomic context might (and likely does) differ between different integrated transgenes. If a researcher then crosses that driver to a different multi-copy effector strain, they might falsely assume that the expression pattern of that driver combined with the second integrated effector would be the same as with the GFP effector. If this second effector is not labeled (e.g. with a fluorescent protein) and expression pattern differs, the researcher is blind to this new expression pattern and might falsely conclude that the first expression pattern is responsible for an effect. Therefore,

having cGAL function in single copy form, where genomic context can be controlled, would be desirable.

Single copy methods also have benefits with regards to gene dosage. One of the chief uses of bipartite systems is for site-of-action genetic analysis. In *C. elegans*, rescue is commonly performed with multi-copy methods, but this raises issues of whether the rescued gene is expressed at native levels or over-expressed. Single copy methods would provide expression of rescued genes at amounts resembling native levels.

Improvements to single copy cGAL could also help discover novel drivers via development of enhancer trap methods[116–118]. Enhancer traps insert proteins such as GAL4 randomly into the genome. When crossed to a cellular reporter effector, they can reveal novel expression patterns not known before, and simultaneously serve as driver strains to be analyzed with other effectors. Enhancer traps present a strategy to find new and useful expression patterns in an agnostic and unbiased manner, and do not rely on previous research or genome databases.

Single copy cGAL systems would also allow a method for maximum reusability of split drivers. Maintaining pairs of split drivers in a strain presents several challenges. If kept at separate loci, generating split strains with an effector would require manipulation of three total separate loci. Homozygosing three loci (while not impossible) is generally undesirable and presents an obstacle to efficient strain generation using split systems. In addition to managing multiple loci, the transformation markers for these loci may interfere with

studies. If the marker is fluorescent, a sizable number of cells will have fluorescence that might interfere with studies. If the marker uses genetic rescue, variety of genetic backgrounds could be a potential confound. Antibiotic resistance markers could be of potential use, but are not as convenient to utilize, especially with so many loci.

Split drivers could be all generated at the same locus and maintained as trans-heterozygotes, but without convenient balancing trans-heterozygosity is unstable. A possible strategy would be put two split drivers into a single construct and knock in the pair of drivers as a single genetic locus. But if every time a new combination of split drivers needed to be inserted into a strain, that would defeat the principle of reusability for bipartite systems.

A schema does exist in which only the advantages of all scenarios mentioned above are preserved (**Figure 4.1**). Here, two parental strains contain split drivers that are targeted to the same chromosome, but at different positions. A loxP site resides at a chromosomal position between the two driver loci, and fluorescent markers are present on the side of the chromosome away from its respective driver. Mixing and matching pairs of split drivers occurs by mating two of these parental split driver strains together to obtain a heterozygote animal containing both split drivers. Induction of Cre recombinase (which can be present in either parental strain or both) forces recombination at the loxP sites, simultaneously linking the two split drivers and the two fluorescent markers. Progeny resulting from this recombination event will give fluorescent double positive animals, and fluorescent double negative animals, which are the desired linked pair of split driver species. If these split

drivers are also marked by antibiotic resistance, this linked pair can then be crossed to an effector strain, containing all three components but distributed amongst two functional loci which greatly reduces labor. Recently, a newly discovered antibiotic nourseothricin (NTC) was demonstrated to be functionally orthogonal to hygromycin[114], and therefore these two antibiotic markers could serve as non-fluorescent and non-genetic transformation markers for this split driver linkage scheme. Additionally, several suitable knock-in loci have been documented which could perform as sites for split driver knock-in[115]. With all these considerations in mind, we transferred our cGAL system to single copy methodology.

4.2 RESULTS

4.2.1 Single copy drivers drive robust expression with a multi-copy effector

To determine whether cGAL could function at the single copy level, we cloned our cGAL driver under control of the *myo-2* promoter into a backbone vector containing homology arms for homology directed repair (HDR) onto LG I. For comparison, we also cloned a driver using the original *S. cerevisiae* DNA-binding domain to see if the increased strength of the cGAL driver using the *S. kudriavzevii* DNA-binding domain could be replicated at the single copy level. We injected these constructs into the *syIs337* multi-copy integrated GFP effector. Several lines were obtained, selected for transformation by hygromycin selection, and then screened for transgene integration via PCR.

All lines showed substantial expression of the GFP effector, in some cases comparable to expression levels in multi-copy driver/effector experiments (**Figure 4.2**). However, our *S. kudriavzevii* single copy driver strains displayed multiple distributions of expression

strength. Due to the linearity of expression levels, we surmise that (despite our best efforts) some strains were not truly single copy integrants. Both lines using the *S. cerevisiae* driver appeared to be well behaved, and one line of the cGAL drivers showed expression approximately 40% higher than the *S. cerevisiae* drivers, which is consistent with our initial multi-copy driver/effector experiments (**Figure 2.5**), leading us to believe that these were genuine single copy lines. Expression strength of the single copy drivers with multi-copy drivers are quite robust, nearly on-par with multi-copy drivers/effectors. Thus, single copy drivers might be sufficient to drive expression of the majority of effectors necessary for functional studies.

4.2.2 Single copy drivers and single copy effectors

Next we wanted to assay expression levels of single copy drivers with single copy effectors. We cloned the *15xUAS::gfp::let-858 3'UTR* effector onto LG IV, using MosSCI. As a benchmark, we also inserted a single copy transgene of P*myo-2::gfp* on LG I, in the same genomic location as our assayed drivers. Overall expression of the GFP effector was much lower (**Figure 4.3**) by about 10-fold. The single copy P*myo-2::gfp* transgene expressed the highest levels, and our single copy cGAL driver/ effector combination at approximately 0.41-fold.

4.3 DISCUSSION AND FUTURE DIRECTIONS

This thesis describes the engineering of cGAL: a complete, robust GAL4-based bipartite gene expression system for *C. elegans*. The system makes use of a GAL4 DNA-binding domain from a novel species of yeast, *S. kudriavzevii*, whose optimal growth temperature is

much closer to laboratory conditions of *C. elegans*. The system demonstrates robust activity:

1) across the experimental range of C. elegans (15-25°C),

2) across a variety of tissues,

3) across a variety of functional effector transgenes.

For more precise intersectional control of effector expression, we have engineered a 'split' system whereby driver expression is dictated by two promoters instead of one. This strategy provides refined spatial cellular expression, as well as simultaneous spatiotemporal control. Finally, we have demonstrated the feasibility of constructing cGAL strains using single copy transgenesis for more precise control of transgene copy number and local genomic context.

From here, many directions can be taken to improve upon and expand the capabilities of the cGAL system. The limiting factor appears to be on the effector side, given our experiments with single and multi-copy driver. One possible improvement would be replace the *Δpes-10* basal promoter in our system with a stronger basal promoter, such as the super core promoter[119].

Recently, a hybrid bipartite system consisting of Tet and Q system components reported that the activation domain of QF (QFAD) performed much better than VP16, VP64, and even VPR[120]. Perhaps the combination of *S. kudriavzevii* GAL4 DNA-binding domain with the QFAD might prove to be an even stronger driver component for single copy cGAL.

This combination would have the added benefit of temporal control- using QFAD would render the transcriptional activator sensitive to QS, which can be temporally controlled by supplementation of quinic acid. Other protein domains that could confer temporal control include degrons and temperature-sensitive inteins. Degrons are protein domains that control the half-life of a protein, and can work through ubiquitin-dependent or ubiquitin-independent mechanisms. Several degron systems are available, including light-activated[121] and small molecule-activated[122] systems. Inteins are protein analogs of DNA introns, capable of self-excision from a polypeptide chain without exogenous cofactors or energy sources. For protein engineering, the strategy is that when the intein is retained, it disrupts protein function; splicing and removal of the intein restores protein function. Recently, a series of temperature sensitive inteins have been characterized[123], generated via PCR mutagenesis of the S. cerevisiae *VMA1* intein. This seems a particularly promising option for temporal control of our cGAL system for two reasons. First, the mutants were assayed for their ability to disrupt and subsequently restore *S. cerevisiae GAL4* activity. Second, the mutants possess a variety of temperature thresholds for temperature-sensitive splicing, many of which encompass 18-25°C, ideal for *C. elegans* laboratory settings.

4.4 FIGURES

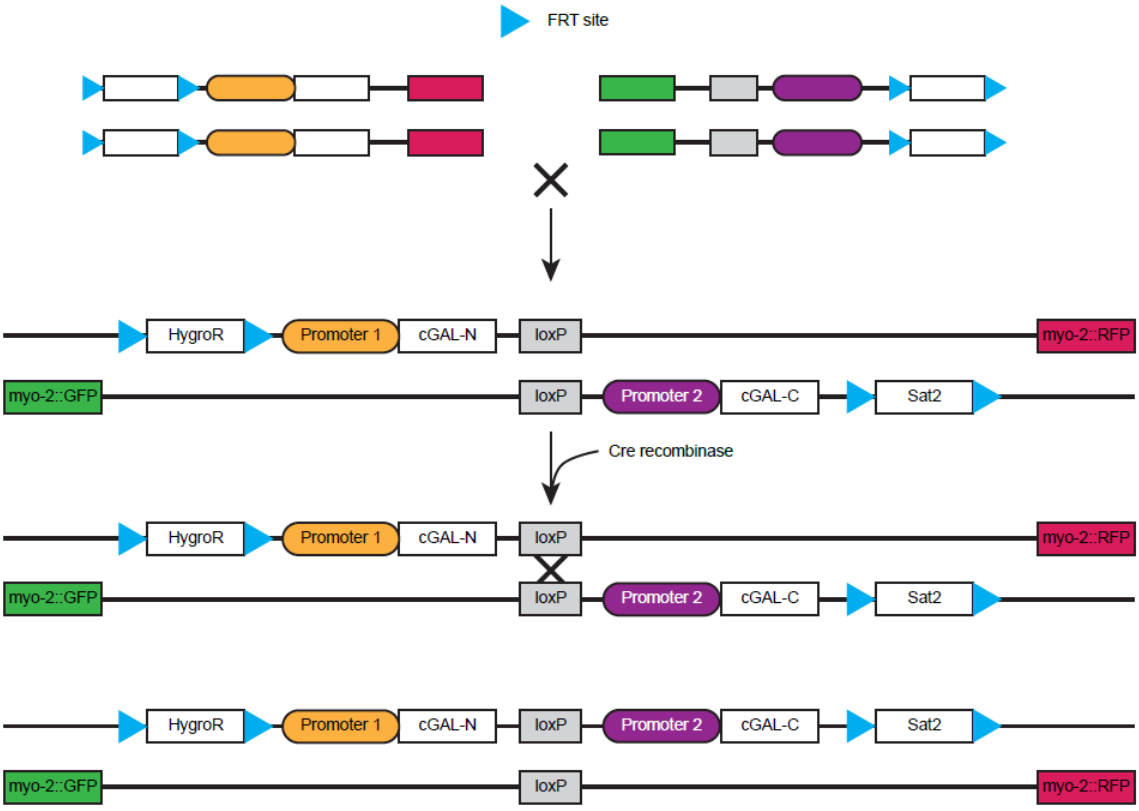


Figure 4.1 | Linkage schema for single copy split drivers

Single copy split drivers have maximum utility when generated as separate strains, but once they are crossed together, maintaining them as two loci is cumbersome. This schema allows for linkage of two different split drivers (orange and purple). Two split drivers are targeted to the same chromosome but at different locations. Each has a loxP site at the same location, and a fluorescent marker (red, green) on the side of the chromosome opposite the loxP site. Mating of these two strains produces the double split trans-heterozygote. If Cre recombinase is then expressed, some fraction of animals will recombine the two strands, producing linked split drivers, and linked fluorescent reporters. Linked split driver animals can then be selected for by lack of fluorescent markers.

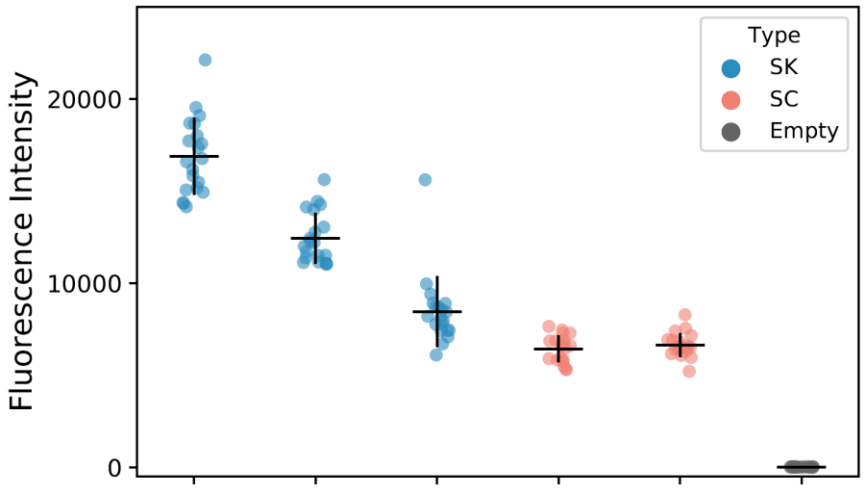


Figure 4.2 | Single copy drivers with *syIs337* multi-copy GFP effector

Single copy drivers are capable of robust expression when combined with a multi-copy effector. Pharyngeal muscle fluorescence was quantitated from single copy cGAL drivers using the *S. kudriavzevii* DNA-binding domain (DBD) in blue, and single copy drivers using the original *S. cerevisiae* DBD are in red for comparison. The last column is the effector alone. The first two columns are suspected not to be true single copy strains. Single lines with n = 20 for all columns.

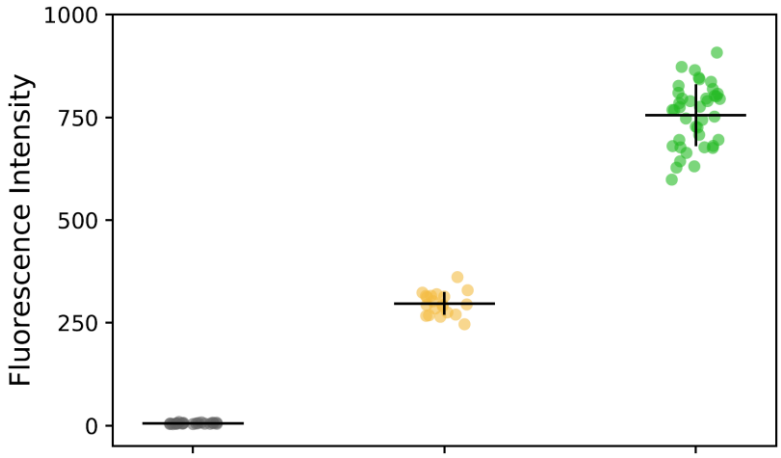


Figure 4.3 | Single copy cGAL driver and effectors

Fluorescence quantitation of *15xUAS::gfp* effector alone, double homozygotes for the *myo-2* cGAL driver and GFP effector in single copy form, and homozygous P*myo-2::gfp* animals. The cGAL animals express GFP at about 40% of the direct fusion. Single lines with n = 20 for all.

4.5 METHODS

Strains

The *Caenorhabditis elegans* strains were maintained at 20 °C, as previously described [1] All the strains used in this study are described in detail in Supporting Information.

Transgenic animals

The standard microinjection procedure for *C. elegans* was used to generate transgenic worms with single copy insertions[13]. The concentrations and compositions of DNA constructs in the injection mixtures of the transgenic worms are described in Supporting Information.

Fluorescence imaging

Worms were paralyzed in M9 buffer supplemented with 30 mg/mL of 2, 3-Butanedione monoxime (Sigma). All fluorescent images for quantification of GFP fluorescence in the pharynx were taken with a Leica DMI6000 inverted microscope equipped with a 40x oil objective and an Andor iXon Ultra 897 EMCCD camera, using Metamorph software (Molecular Devices). An ROI outlining the entire pharynx, as well as a background ROI, was selected for each image. The background-subtracted mean fluorescence intensity was used to quantify the GFP fluorescence in the pharyngeal muscles of each worm.

BIBLIOGRAPHY

1. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* (1974). doi:10.1002/cbic.200300625

2. Genome sequence of the nematode C. elegans: A platform for investigating biology. *Science* (1998). doi:10.1126/science.282.5396.2012

3. Sulston, J. E. & Horvitz, H. R. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev. Biol.* (1977). doi:10.1016/0012-1606(77)90158-0

4. Sulston, J., Schierenberg, E. & White, J. The embryonic cell lineage of the nematode Caenorhabditis elegans. *Development* (1983).

5. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* (1986). doi:10.1098/rstb.1986.0056

6. Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H. & Chklovskii, D. B. Structural properties of the Caenorhabditis elegans neuronal network. *PLoS Comput. Biol.* (2011). doi:10.1371/journal.pcbi.1001066

7. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science (80-.).* (1994). doi:10.1126/science.8303295

8. Mello, C. C., Kramer, J. M., Stinchcomb, D. & Ambros, V. Efficient gene transfer in C.elegans: extrachromosomal maintenance and integration of transforming sequences. *EMBO J.* (1991).

9. Radman, I., Greiss, S. & Chin, J. W. Efficient and Rapid C. elegans Transgenesis by Bombardment and Hygromycin B Selection. *PLoS One* (2013). doi:10.1371/journal.pone.0076019

10. Mello, C. & Fire, A. DNA Transformation. *Methods Cell Biol.* (1995). doi:10.1016/S0091-679X(08)61399-0

11. Bessereau, J.-L. Transposons in C. elegans. *WormBook* (2006). doi:10.1895/wormbook.1.70.1

12. Frøkjær-Jensen, C., Davis, M. W., Ailion, M. & Jorgensen, E. M. Improved Mos1-mediated transgenesis in C. elegans. *Nat. Methods* (2012). doi:10.1038/nmeth.1865

13. Frøkjær-Jensen, C. *et al.* Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat. Genet.* (2008). doi:10.1038/ng.248

14. Frokjaer-Jensen, C. *et al.* Random and targeted transgene insertion in C. elegans using a modified Mosl transposon. *Nat. Methods* (2014). doi:10.1038/nmeth.2889.Random

15. Chiu, H., Schwartz, H. T., Antoshechkin, I. & Sternberg, P. W. Transgene-free genome editing in Caenorhabditis elegans using CRISPR-Cas. *Genetics* (2013). doi:10.1534/genetics.113.155879

16. Dickinson, D. J., Ward, J. D., Reiner, D. J. & Goldstein, B. Engineering the Caenorhabditis elegans genome using Cas9-triggered homologous recombination. *Nat. Methods* (2013). doi:10.1038/nmeth.2641

17. Arribere, J. A. *et al.* Efficient Marker-Free Recovery of Custom Genetic Modifications with CRISPR/Cas9 in Caenorhabditis elegans. *Genetics* **198**, 837–846 (2014).

18. Dickinson, D. J. & Goldstein, B. CRISPR-based methods for caenorhabditis elegans genome engineering. *Genetics* (2016). doi:10.1534/genetics.115.182162

19. Giepmans, B. N. G., Adams, S. R., Ellisman, M. H. & Tsien, R. Y. The fluorescent toolbox for assessing protein location and function. *Science* (2006). doi:10.1126/science.1124618

20. Rodriguez, E. A. *et al.* The Growing and Glowing Toolbox of Fluorescent and Photoactive Proteins. *Trends in Biochemical Sciences* (2017). doi:10.1016/j.tibs.2016.09.010

21. Dana, H. *et al.* Sensitive red protein calcium indicators for imaging neural activity. *Elife* (2016). doi:10.7554/elife.12727

22. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* (2013). doi:10.1038/nature12354

23. McIsaac, R. S., Bedbrook, C. N. & Arnold, F. H. Recent advances in engineering microbial rhodopsins for optogenetics. *Current Opinion in Structural Biology* (2015). doi:10.1016/j.sbi.2015.05.001

24. Zhang, F. *et al.* The microbial opsin family of optogenetic tools. *Cell* (2011). doi:10.1016/j.cell.2011.12.004

25. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* (1993). doi:10.1101/lm.1331809

26. Brand, A. H., Manoukian, A. S. & Perrimon, N. Ectopic expression in Drosophila. *Methods Cell Biol.* **44**, 635–54 (1994).

27. Giniger, E., Varnum, S. M. & Ptashne, M. Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell* (1985). doi:10.1016/0092-8674(85)90336-8

28. Ma, J. & Ptashne, M. The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80. *Cell* (1987). doi:10.1016/0092-8674(87)90670-2

29. McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K. & Davis, R. L. Spatiotemporal Rescue of Memory Dysfunction in Drosophila. *Science (80-.).* (2003). doi:10.1126/science.1089035

30. Ma, J. & Ptashne, M. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* (1987). doi:10.1016/0092-8674(87)90081-X

31. Johnston, S. A., Zavortink, M. J., Debouck, C. & Hopper, J. E. Functional domains of the yeast regulatory protein GAL4. *Proc. Natl. Acad. Sci.* **83**, 6553–6557 (1986).

32. Keegan, L., Gill, G. & Ptashne, M. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. *Science* **231**, 699–704 (1986).

33. Luan, H., Peabody, N. C., Vinson, C. R. & White, B. H. Refined spatial manipulation of neuronal function by combinatorial restriction of transgene expression. *Neuron* **52**, 425–36 (2006).

34. Szüts, D. & Bienz, M. LexA chimeras reveal the function of Drosophila Fos as a context-dependent transcriptional activator. *Proc. Natl. Acad. Sci. U. S. A.* (2000).

35. Lai, S. L. & Lee, T. Genetic mosaic with dual binary transcriptional systems in Drosophila. *Nat. Neurosci.* (2006). doi:10.1038/nn1681

36. Potter, C. J., Tasic, B., Russler, E. V., Liang, L. & Luo, L. The Q system: A repressible binary system for transgene expression, lineage tracing, and mosaic analysis. *Cell* (2010). doi:10.1016/j.cell.2010.02.025

37. Potter, C. J. & Luo, L. Using the Q system in Drosophila melanogaster. *Nat. Protoc.* (2011). doi:10.1038/nprot.2011.347

38. Wei, X., Potter, C. J., Luo, L. & Shen, K. Controlling gene expression with the Q repressible binary expression system in Caenorhabditis elegans. *Nat. Methods* (2012). doi:10.1038/nmeth.1929

39. Bieschke, E. T., Wheeler, J. C. & Tower, J. Doxycycline-induced transgene expression during Drosophila development and aging. *Mol. Gen. Genet.* (1998). doi:10.1007/s004380050770

40. Stebbins, M. J. & Yin, J. C. P. Adaptable doxycycline-regulated gene expression

systems for Drosophila. *Gene* (2001). doi:10.1016/S0378-1119(01)00447-4

41. Bello, B., Resendez-Perez, D. & Gehring, W. J. Spatial and temporal targeting of gene expression in Drosophila by means of a tetracycline-dependent transactivator system. *Development* (1998).

42. Voutev, R. & Hubbard, E. J. A. A 'FLP-out' system for controlled gene expression in Caenorhabditis elegans. *Genetics* (2008). doi:10.1534/genetics.108.090274

43. Hubbard, E. J. A. FLP/FRT and Cre/lox recombination technology in C. elegans. *Methods* (2014). doi:10.1016/j.ymeth.2014.05.007

44. Traven, A., Jelicic, B. & Sopta, M. Yeast Gal4: a transcriptional paradigm revisited. *EMBO Rep.* **7**, 496–9 (2006).

45. Campbell, R. N., Leverentz, M. K., Ryan, L. A. & Reece, R. J. Metabolic control of transcription: paradigms and lessons from Saccharomyces cerevisiae. *Biochem. J.* **414**, 177–87 (2008).

46. Wietek, J. & Prigge, M. Enhancing Channelrhodopsins: An Overview. *Methods Mol. Biol.* **1408**, 141–65 (2016).

47. Rose, T., Goltstein, P. M., Portugues, R. & Griesbeck, O. Putting a finishing touch on GECIs. *Front. Mol. Neurosci.* **7**, 88 (2014).

48. Duffy, J. B. GAL4 system in Drosophila: a fly geneticist's Swiss army knife. *Genesis* **34**, 1–15 (2002).

49. del Valle Rodríguez, A., Didiano, D. & Desplan, C. Power tools for gene expression and clonal analysis in Drosophila. *Nat. Methods* **9**, 47–55 (2012).

50. Venken, K. J. T., Simpson, J. H. & Bellen, H. J. Genetic manipulation of genes and cells in the nervous system of the fruit fly. *Neuron* **72**, 202–230 (2011).

51. Ornitz, D. M., Moreadith, R. W. & Leder, P. Binary system for regulating transgene expression in mice: targeting int-2 gene expression with yeast GAL4/UAS control elements. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 698–702 (1991).

52. Scheer, N. & Campos-Ortega, J. A. Use of the Gal4-UAS technique for targeted gene expression in the zebrafish. *Mech. Dev.* **80**, 153–158 (1999).

53. Laplaze, L. *et al.* GAL4-GFP enhancer trap lines for genetic manipulation of lateral root development in Arabidopsis thaliana. *J. Exp. Bot.* **56**, 2433–2442 (2005).

54. Schinko, J. B. *et al.* Functionality of the GAL4/UAS system in Tribolium requires the use of endogenous core promoters. *BMC Dev. Biol.* **10**, 53 (2010).

55. Lynd, A. & Lycett, G. J. Development of the bi-partite Gal4-UAS system in the African malaria mosquito, Anopheles gambiae. *PLoS One* **7**, e31552 (2012).

56. Pfeiffer, B. D. *et al.* Refinement of tools for targeted gene expression in Drosophila. *Genetics* **186**, 735–55 (2010).

57. Distel, M., Wullimann, M. F. & Köster, R. W. Optimized Gal4 genetics for permanent gene expression mapping in zebrafish. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13365–70 (2009).

58. Triezenberg, S. J., Kingsbury, R. C. & McKnight, S. L. Functional dissection of VP16, the trans-activator of herpes simplex virus immediate early gene expression. *Genes Dev.* **2**, 718–29 (1988).

59. Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14628–33 (1998).

60. Salvadó, Z. *et al.* Temperature adaptation Markedly Determines evolution within the genus Saccharomyces. *Appl. Environ. Microbiol.* **77**, 2292–2302 (2011).

61. Hittinger, C. T. *et al.* Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature* **464**, 54–58 (2010).

62. Liang, S. D., Marmorstein, R., Harrison, S. C. & Ptashne, M. DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA. *Mol. Cell. Biol.* **16**, 3773–3780 (1996).

63. Marmorstein, R., Carey, M., Ptashne, M. & Harrison, S. C. DNA recognition by GAL4: structure of a protein-DNA complex. *Nature* **356**, 408–14 (1992).

64. Thomas, J. H. Genetic analysis of defecation in Caenorhabditis elegans. *Genetics* **124**, 855–72 (1990).

65. Zhao, B. & Schafer, W. R. Neuropeptide signaling: from the gut. *Curr. Biol.* **23**, R481-3 (2013).

66. Wang, H. *et al.* Neuropeptide secreted from a pacemaker activates neurons to control a rhythmic behavior. *Curr. Biol.* **23**, 746–754 (2013).

67. Mahoney, T. R. *et al.* Intestinal signaling to GABAergic neurons regulates a rhythmic behavior in Caenorhabditis elegans. *Proc. Natl. Acad. Sci.* **105**, 16350–16355 (2008).

68. Beg, A. A. & Jorgensen, E. M. EXP-1 is an excitatory GABA-gated cation channel.

Nat. Neurosci. **6**, 1145–52 (2003).

69. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat. Neurosci.* **8**, 1263–8 (2005).

70. Jorgensen, E. M. GABA. *WormBook* 1–13 (2005). doi:10.1895/wormbook.1.14.1

71. Rowan, B. A., Weigel, D. & Koenig, D. Developmental genetics and new sequencing technologies: the rise of nonmodel organisms. *Dev. Cell* **21**, 65–76 (2011).

72. Hoier, E. F., Mohler, W. A., Kim, S. K. & Hajnal, A. The Caenorhabditis elegans APC-related gene apr-1 is required for epithelial cell migration and Hox gene expression. *Genes Dev.* **14**, 874–86 (2000).

73. Davis, M. W., Morton, J. J., Carroll, D. & Jorgensen, E. M. Gene activation using FLP recombinase in C. elegans. *PLoS Genet.* (2008). doi:10.1371/journal.pgen.1000028

74. Pokala, N., Liu, Q., Gordus, A. & Bargmann, C. I. Inducible and titratable silencing of Caenorhabditis elegans neurons in vivo with histamine-gated chloride channels. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 2770–5 (2014).

75. McWilliam, H. *et al.* Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **41**, W597-600 (2013).

76. Webster, N., Jin, J. R., Green, S., Hollis, M. & Chambon, P. The yeast UASG is a transcriptional enhancer in human HeLa cells in the presence of the GAL4 trans-activator. *Cell* **52**, 169–178 (1988).

77. Corsi, A. K., Wightman, B. & Chalfie, M. A Transparent Window into Biology: A Primer on Caenorhabditis elegans. *Genetics* **200**, 387–407 (2015).

78. Wang, H. *et al.* cGAL, a temperature-robust GAL4-UAS system for Caenorhabditis elegans. *Nat. Methods* **14**, 145–148 (2017).

79. Chelur, D. S. & Chalfie, M. Targeted cell killing by reconstituted caspases. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 2283–8 (2007).

80. Ma, J. & Ptashne, M. Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48**, 847–53 (1987).

81. Venken, K. J. T., Simpson, J. H. & Bellen, H. J. Genetic manipulation of genes and cells in the nervous system of the fruit fly. *Neuron* **72**, 202–30 (2011).

82. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E690-7 (2012).

83. Shah, N. H. & Muir, T. W. Inteins: Nature's Gift to Protein Chemists. *Chem. Sci.* **5**, 446–461 (2014).

84. Kane, P. M. *et al.* Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase. *Science* **250**, 651–7 (1990).

85. Hirata, R. *et al.* Molecular structure of a gene, VMA1, encoding the catalytic subunit of H(+)-translocating adenosine triphosphatase from vacuolar membranes of Saccharomyces cerevisiae. *J. Biol. Chem.* **265**, 6726–33 (1990).

86. Wu, H., Hu, Z. & Liu, X. Q. Protein trans-splicing by a split intein encoded in a split DnaE gene of Synechocystis sp. PCC6803. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9226–31 (1998).

87. Iwai, H., Züger, S., Jin, J. & Tam, P.-H. Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. *FEBS Lett.* **580**, 1853–8 (2006).

88. Carvajal-Vallejos, P., Pallissé, R., Mootz, H. D. & Schmidt, S. R. Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *J. Biol. Chem.* **287**, 28686–96 (2012).

89. Zettler, J., Schütz, V. & Mootz, H. D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett.* **583**, 909–14 (2009).

90. Zhang, S., Ma, C. & Chalfie, M. Combinatorial marking of cells and organelles with reconstituted fluorescent proteins. *Cell* **119**, 137–44 (2004).

91. Bedbrook, C. N. *et al.* Genetically Encoded Spy Peptide Fusion System to Detect Plasma Membrane-Localized Proteins In Vivo. *Chem. Biol.* **22**, 1108–21 (2015).

92. Wong, S. S. C., Kotera, I., Mills, E., Suzuki, H. & Truong, K. Split-intein mediated re-assembly of genetically encoded Ca(2+) indicators. *Cell Calcium* **51**, 57–64 (2012).

93. Dong, Y., Gou, Y., Li, Y., Liu, Y. & Bai, J. Synaptojanin cooperates in vivo with endophilin through an unexpected mechanism. *Elife* **4**, 1–50 (2015).

94. Mills, K. V., Johnson, M. A. & Perler, F. B. Protein splicing: How Inteins escape from precursor proteins. *J. Biol. Chem.* **289**, 14498–14505 (2014).

95. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The Structure of the Nervous System of the Nematode Caenorhabditis elegans. *Philos. Trans. R. Soc. B Biol. Sci.* **314**, 1–340 (1986).

96. Lints, R. and Hall, D. H. Male neuronal support cells, overview. *WormAtlas* (2009). doi:10.3908/wormatlas.2.9

97. Avery, L. & Horvitz, H. R. Pharyngeal Pumping Continues after Laser Killing of the Pharyngeal Nervous System of C. elegans.

98. Raizen, D. M., Lee, R. Y. & Avery, L. Interacting genes required for pharyngeal excitation by motor neuron MC in Caenorhabditis elegans. *Genetics* **141**, 1365–82 (1995).

99. Feng, H. & Hope, I. A. The Caenorhabditis elegans homeobox gene ceh-19 is required for MC motorneuron function. *Genesis* **51**, 163–78 (2013).

100. Pereira, L. *et al.* A cellular and regulatory map of the cholinergic nervous system of C. elegans. *Elife* **4**, 1–46 (2015).

101. Sassone-Corsi, P. The cyclic AMP pathway. *Cold Spring Harb. Perspect. Biol.* **4**, (2012).

102. Kim, S., Govindan, J. A., Tu, Z. J. & Greenstein, D. SACY-1 DEAD-Box helicase links the somatic control of oocyte meiotic maturation to the sperm-to-oocyte switch and gamete maintenance in Caenorhabditis elegans. *Genetics* **192**, 905–28 (2012).

103. Song, B.-M. & Avery, L. Serotonin activates overall feeding by activating two separate neural pathways in Caenorhabditis elegans. *J. Neurosci.* **32**, 1920–1931 (2012).

104. Trojanowski, N. F., Nelson, M. D., Flavell, S. W., Fang-Yen, C. & Raizen, D. M. Distinct Mechanisms Underlie Quiescence during Two Caenorhabditis elegans Sleep-Like States. *J. Neurosci.* **35**, 14571–84 (2015).

105. Wang, H. & Sieburth, D. PKA controls calcium influx into motor neurons during a rhythmic behavior. *PLoS Genet.* **9**, e1003831 (2013).

106. Correll, L. A., Woodford, T. A., Corbin, J. D., Mellon, P. L. & McKnight, G. S. Functional characterization of cAMP-binding mutations in type I protein kinase. *J. Biol. Chem.* **264**, 16672–8 (1989).

107. Bacaj, T. & Shaham, S. Temporal control of cell-specific transgene expression in Caenorhabditis elegans. *Genetics* **176**, 2651–5 (2007).

108. Churgin, M. A., He, L., Murray, J. I. & Fang-Yen, C. Efficient single-cell transgene

induction in Caenorhabditis elegans using a pulsed infrared laser. *G3 (Bethesda).* **3**, 1827–32 (2013).

109. Singhal, A. & Shaham, S. Infrared laser-induced gene expression for tracking development and function of single C. elegans embryonic neurons. *Nat. Commun.* **8**, 14100 (2017).

110. Kage-Nakadai, E. *et al.* A conditional knockout toolkit for Caenorhabditis elegans based on the Cre/loxP recombination. *PLoS One* **9**, e114680 (2014).

111. Boulin, T., Etchberger, J. F. & Hobert, O. Reporter gene fusions. *WormBook* 1–23 (2006). doi:10.1895/wormbook.1.106.1

112. Perez, F. & Granger, B. E. IPython: A System for Interactive Scientific Computing. *Comput. Sci. Eng.* **9**, 21–29 (2007).

113. Rubin, G. M. & Spradling, A. C. Genetic transformation of Drosophila with transposable element vectors. *Science* (1982). doi:10.1126/science.6289436

114. Obinata, H., Sugimoto, A. & Niwa, S. Streptothricin acetyl transferase 2 (Sat2): A dominant selection marker for Caenorhabditis elegans genome editing. *PLoS One* (2018). doi:10.1371/journal.pone.0197128

115. Silva-Garcia, C. G. *et al.* Single-copy Knock-In Loci for Defined Gene Expression in C. elegans. *bioRxiv* (2018). doi:10.1101/500892

116. Manseau, L. *et al.* GAL4 enhancer traps expressed in the embryo, larval brain, imaginal discs, and ovary of Drosophila. *Dev. Dyn.* (1997). doi:10.1002/(SICI)1097-0177(199707)209:3<310::AID-AJA6>3.0.CO;2-L

117. Sepp, K. J. & Auld, V. J. Conversion of lacZ enhancer trap lines to GAL4 lines using targeted transposition in Drosophila melanogaster. *Genetics* (1999).

118. Ward, E. J. *et al.* GAL4 enhancer trap patterns duringDrosophila development. *genesis* (2002). doi:10.1002/gene.10138

119. Juven-Gershon, T., Cheng, S. & Kadonaga, J. T. Rational design of a super core promoter that enhances gene expression. *Nat. Methods* (2006). doi:10.1038/nmeth937

120. Mao, S. *et al.* A Tet/Q Hybrid System for Robust and Versatile Control of Transgene Expression in C. elegans. *iScience* (2018). doi:10.1016/j.isci.2018.12.023

121. Hermann, A., Liewald, J. F. & Gottschalk, A. A photosensitive degron enables acute light-induced protein degradation in the nervous system. *Current Biology* (2015). doi:10.1016/j.cub.2015.07.040

122. Zhang, L., Ward, J. D., Cheng, Z. & Dernburg, A. F. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in C. elegans. *Development* (2015). doi:10.1242/dev.129635

123. Tan, G., Chen, M., Foote, C. & Tan, C. Temperature-sensitive mutations made easy: Generating conditional mutations by using temperature-sensitive inteins that function within different temperature ranges. *Genetics* (2009). doi:10.1534/genetics.109.104794

Appendix A

Allele and Strain Information

Chapter 2

The co-injection markers used include *KP708 (Pttx-3::rfp), KP1369(Pmyo-2::nls::mCherry), KP1106(Pmyo-2::nls::gfp), unc-119(+)* rescue plasmid*, Pofm-1::rfp and Punc-122::gfp*.

All initial descriptions of extrachromosomal arrays (***syEx####)*** and integrants ***(syIs####)*** are bolded for convenience. All integrants were generated by X-ray irradiation.

syEx1452 *[15xUAS::Δpes-10::gfp::unc-54 3'UTR, 25ng/µL; Pttx-3::rfp, 40ng/µL; pBlueScript, 35 ng/µL]*, injected into N2, used to generate ***syIs300*** and ***syIs302***.

syEx1431 and ***syEx1432*** *[Pmyo-2::GAL4$_{SC}$::VP16::unc-54 3'UTR, 10ng/µL; unc-119(+), 50ng/µL; pBlueScript, 40 ng/µL]*, injected into the strain *unc-119(ed3); syIs300.*

syEx1433 *and* **syEx1434** *[Pmyo-2::GAL4$_{SC}$::VP64::unc-54 3'UTR, 10ng/µL; unc-119(+), 50ng/µL; pBlueScript, 40 ng/µL]*, injected into the strain *unc-119(ed3); syIs300.*

syEx1435 and ***syEx1436*** *[Pmyo-2::GAL4$_{SK}$::VP64::unc-54 3'UTR; 10ng/µL, unc-119(+), 50ng/µL; pBlueScript, 40 ng/µL]*, injected into the strain *unc-119(ed3); syIs300.*

syEx1437 and ***syEx1438*** *[Pmyo-2::gfp::unc-54 3'UTR, 10ng/µL; unc-119(+), 50ng/µL; pBlueScript, 40 ng/µL]*, injected into the strain *unc-119(ed3).*

syEx1448 and ***syEx1449*** *[Pnlp-40::GAL4$_{SK}$::VP64::unc-54 3'UTR, 10ng/µL; Pmyo-2::nls::mCherry, 10ng/µL; pBlueScript, 80ng/µL]*, injected into *syIs302. syEx1449* was used to generate ***syIs318***, ***syIs319*** and ***syIs320***, as intestine drivers.

syEx1450 and ***syEx1451*** *[Pmyo-3::GAL4$_{SK}$::VP64::unc-54 3'UTR, 10ng/µL; Pmyo-2::nls::mCherry, 10ng/µL; pBlueScript, 80ng/µL]*, injected into the strain carrying *syIs302. syEx1451* was used to generate ***syIs321***, as the body wall muscle driver.

syEx1471 *[Punc-47::GAL4$_{SK}$::VP64::unc-54 3'UTR, 60ng/µL; Pofm-1::rfp, 40ng/µL]*, *syEx1451* was used to generate ***syIs322***, ***syIs323***, ***syIs324*** and ***syIs325***, as GABAergic neuron drivers (These GABAergic drivers were weak, we suggest using drivers built in the pPD117.01 backbone with the *let-858 3'UTR*).

syEx1475, ***syEx1476***, and ***syEx1477*** *[5xUAS::Δpes-10::gfp::unc-54 3'UTR, 25ng/µL; unc-119(+), 50ng/µL; pBlueScript, 25 ng/µL]*, injected into the strain *unc-119(ed3); syIs301*.

syEx1478 and ***syEx1479*** *[10xUAS::Δpes-10::gfp::unc-54 3'UTR, 25ng/µL; unc-119(+), 50ng/µL; pBlueScript, 25 ng/µL]*, injected in to the strain *unc-119(ed3); syIs301*.

syEx1480 and ***syEx1481*** *[15xUAS::Δpes-10::gfp::unc-54 3'UTR, 25ng/µL; unc-119(+), 50ng/µL; pBlueScript, 25 ng/µL],* injected in to the strain *unc-119(ed3); syIs301.*
syEx1482 and ***syEx1483*** *[20xUAS::Δpes-10::gfp::unc-54 3'UTR, 25ng/µL; unc-119(+), 50ng/µL; pBlueScript, 25 ng/µL],* injected in to the strain *unc-119(ed3); syIs301.*

syEx1443 and ***syEx1444*** *[15xUAS::Δpes-10::aex-2(+) cDNA::unc-54 3'UTR, 25ng/µL; Pmyo-2::nls::gfp, 10ng/µL; pBlueScript, 65 ng/µL],* injected into the strain *aex-2(sa3).*

syEx1433 *and* ***syEx1447*** *[Prab-3::GAL4$_{SK}$::VP64::let-858 3'UTR, 10 ng/µL; Pofm-1::rfp, 40ng/µL; pBlueScript, 50 ng/µL]*, injected into N2. *syEx1447* was used to generate ***syIs334***, ***syIs335*** *and* ***syIs336*** as pan-neuronal driver lines.

syEx1430*[Pmyo-2::GAL4$_{SC}$::VP64::unc-54 3'UTR; 10ng/µL, Pofm-1::rfp 40ng/µL; 1kb DNA ladder(NEB), 150 ng/µL]*, also used to generate the ***syIs301*** as the pharyngeal muscle driver.

syEx1488 *[15xUAS::Δpes-10::gfp::let-858 3'UTR, 25 ng/µL; Pttx-3::rfp, 50 ng/µL; 1 kb ladder (NEB), 125 ng/µL]*, injected into N2, used to generate ***syIs337*** and ***syIs343*** for *15xUAS::gfp::let-858 3'UTR* effector lines.

syEx1484 *[Punc-17:: GAL4$_{SK}$::VP64::let-858 3'UTR, 25 ng/µL; Punc-17::mCherry, 25 ng/µL; unc-119(+), 50 ng/µL]*, injected into the strain *syIs343; unc-119(ed3).*

syEx1485 *[Punc-47::GAL4$_{SK}$::VP64::let-858 3'UTR, 25 ng/µL;Punc-47:: mCherry, 25 ng/µL; unc-119(+), 50 ng/µL]*, injected into the strain *syIs343; unc-119(ed3).*

syEx1486 *[Peat-4::GAL4$_{SK}$::VP64::let-858 3'UTR, 25 ng/µL; Peat-4:: mCherry, 25 ng/µL; unc-119(+), 50 ng/µL]*, injected into the strain *syIs343; unc-119(ed3).*

syEx1460 *[15xUAS::Δpes-10::hChR2(H134R)::eyfp::let-858 3'UTR, 25ng/µL; Pttx-3::rfp, 40ng/µL; pBlueScript, 35 ng/µL]*, injected into N2, used to generate ***syIs340***, ***syIs341*** and ***syIs342*** for *15xUAS::hChR2(H134R)::eyfp::let-858 3'UTR* effector lines.

syEx1487 *[Punc-47::GAL4$_{SK}$::VP64::let-858 3'UTR, 25 ng/µL; Pofm-1::rfp, 40 ng/µL; 1 kb ladder (NEB), 35 ng/µL]*, injected into the strain *syIs341*.

Wild type *N2*
PS6041 *unc-119(ed3)* III

Figure 2.1, 2.2
PS6843 *syIs300* V
PS6932 *unc-119(ed3); syIs300*
PS6900 *syEx1431; unc-119; syIs300*
PS6901 *syEx1432; unc-119(ed3); syIs300*

PS6902 *syEx1433; unc-119(ed3); syIs300*
PS6903 *syEx1434; unc-119(ed3); syIs300*

Figure 2.3
PS6872 *syIs302* III
PS6844 *syIs301* V
PS6965 *syIs301; syIs302*

Figure 2.4
PS6844 *syIs301* V
PS6964 *unc-119(ed3); syIs301*
PS7007 *syEx1475; unc-119(ed3); syIs301*
PS7008 *syEx1476; unc-119(ed3); syIs301*
PS7009 *syEx1477; unc-119(ed3); syIs301*
PS7010 *syEx1478; unc-119(ed3); syIs301*
PS7012 *syEx1480; unc-119(ed3); syIs301*
PS7013 *syEx1481; unc-119(ed3); syIs301*
PS7014 *syEx1482; unc-119(ed3); syIs301*
PS7015 *syEx1483; unc-119(ed3); syIs301*

Figure 2.5, 2.7
PS6902 syEx1433; *unc-119(ed3); syIs300*
PS6903 *syEx1434; unc-119(ed3); syIs300*
PS6904 *syEx1435; unc-119(ed3); syIs300*
PS6905 *syEx1436; unc-119(ed3); syIs300*
PS6906 *syEx1437; unc-119(ed3); syIs300*
PS6907 *syEx1438; unc-119(ed3)*

Figure 2.8
PS6933 *syIs318syIs302 III*
PS7067 *syIs321; syIs300*
PS6987 *syIs337; syIs334*
PS7026 *syIs343*
PS7017 *syIs343; unc-119(ed3)*
PS7018 *syEx1484; syIs343; unc-119(ed3)*
PS7019 *syEx1485; syIs343; unc-119(ed3)*
PS7020 *syEx1486; syIs343; unc-119(ed3)*

Figure 2.10
JT3 *aex-2(sa3)* X
PS6975 *syEx1443; aex-2(sa3)*
PS6976 *syEx1444; aex-2(sa3)*
PS6936 *syIs321*
PS6935 *syIs320*
PS6938 *syIs323*

Figure 2.12
PS7021 *syEx1487; syIs341*
PS7044 *syIs341*

Chapter 3

Figures 3.3-3.5
GFP effector (syIs300): pG4US7(15xUAS::Δpes-10::gfp::unc-54 3'UTR), 25 ng/µL; Pttx3::rfp, 40 ng/µL; pBlueScript, 35 ng/µL.

PS6843 *syIs300 V* *outcrossed x7*
PS6932 *unc-119(ed3) III; syIs300 V*

Pmyo-2 intact cGAL driver (syEx1435 and syEx1436): pG4U19, 10 ng/µL; unc-119(+) rescue plasmid, 50 ng/µL; pBlueScript, 40 ng/µL.

PS6904 *syEx1435; unc-119(ed3); syIs300*
PS6905 *syEx1436; unc-119(ed3); syIs300*

Pmyo-2 split cGAL driver with DnaE intein (syEx1463 and syEx1464): pHW438, 10 ng/µL; pHW439, 10 ng/ul; unc-119(+) rescue plasmid, 50 ng/µL; pBlueScript, 30 ng/µL.

PS7034 *syEx1463; unc-119(ed3); syIs300*
PS7035 *syEx1464; unc-119(ed3); syIs300*

Pmyo-2 split cGAL driver with SpyTag/SpyCatcher (syEx1511, syEx1512 and syEx1571): pHW375, 10 ng/µL; pHW378, 10 ng/µL; unc-119(+) rescue plasmid, 50 ng/µL; pBlueScript, 30 ng/µL.

PS7250 *syEx1511; unc-119(ed3); syIs300*
PS7251 *syEx1512; unc-119(ed3); syIs300*
PS7252 *syEx1571; unc-119(ed3); syIs300*

Pmyo-2 split cGAL driver with leucine zipper (syEx1572, syEx1573 and syEx1574): pHW508, 10 ng/µL; pHW509, 10 ng/µL; unc-119(+) rescue plasmid, 50 ng/µL; pBlueScript, 30 ng/µL.

PS7348 *syEx1572; unc-119(ed3); syIs300*
PS7349 *syEx1573; unc-119(ed3); syIs300*
PS7350 *syEx1574; unc-119(ed3); syIs300*

Pmyo-2 split cGAL driver with gp41-1 intein (syEx1575, syEx1576 and syEx1577): pHW510, 10 ng/µL; pHW511, 10 ng/µL; unc-119(+) rescue plasmid, 50 ng/µL; pBlueScript, 30 ng/µL.

PS7351 *syEx1575; unc-119(ed3); syIs300*
PS7352 *syEx1576; unc-119(ed3); syIs300*

PS7353 *syEx1577; unc-119(ed3); syIs300*

Pmyo-2 split cGAL driver with gp41-1 N-intein wild-type. (syEx1589): pAH35, 5 ng/µL; KP1368, 10 ng/µL; pBlueScript, 85 ng/µL.

PS7686 *syEx1589; syIs433 IV; syIs300 V*

Pmyo-2 split cGAL driver with gp41-1 N-intein C1A mutant. (syEx1590): pHW564, 5 ng/µL; KP1368, 10 ng/µL; pBlueScript, 85 ng/µL.

PS7738 *syEx1590; syIs433 IV; syIs300 V*

Figures 3.6-3.8
Phsp-16.41 split cGAL(DBD)-gp41-1-N-intein driver (syEx1579, syIs435): pAH34, 10 ng/µL; Pmyo-2::NLS::mCherry, 10 ng/µL; pBlueScript, 80 ng/µL.

PS7422 *syEx1579*
PS7406 *syIs435 IV* *outcrossed x3*
PS7409 *syIs435 IV; syIs300 V* *outcrossed x5*

Pmyo-2 split cGAL(DBD)-gp41-1-N-intein driver (syIs430, syIs431, and syIs432): pAH35, 10 ng/µL; Pmyo-2::NLS::mCherry, 10 ng/µL; pBlueScript, 80 ng/µL.

PS7400 *syIs431 III* *outcrossed x5*
PS7402 *syIs431 III; syIs300 V* *outcrossed x3*
PS7403 *syIs430 IV* *outcrossed x3*
PS7408 *syIs432 II; syIs300 V outcrossed x3*

Pmyo-2 split gp41-1-C-intein-cGAL(AD): (syEx1580, syIs433, and syIs434): pAH36, 10 ng/µL; Punc-122::rfp, 10 ng/µL; pBlueScript, 80 ng/µL.

PS7423 *syEx1580*
PS7401 *syIs433 IV; syIs300 V* *outcrossed x3*
PS7404 *syIs433 IV* *outcrossed x5*
PS7405 *syIs434 II* *outcrossed x0*

Prab-3 split cGAL(DBD)-gp41-1-N-intein driver (syEx1578): pHW530, 10 ng/µL; Pmyo-2::NLS::mCherry, 10 ng/µL; pBlueScript, 80 ng/µL.

PS7410 *syEx1578; syIs300 V*

Peft-3 split cGAL(DBD)-gp41-1N-intein driver (syEx1581 and syEx1582): pHW533, 10 ng/µL; Pmyo-2::NLS::mCherry, 10 ng/µL; pBlueScript, 80 ng/µL.

PS7424 *syEx1581; syIs300 V*

PS7425 *syEx1582; syIs300 V*

Peft-3 split gp41-1-C-intein-cGAL(AD) driver (syEx1586, syEx1587 and syEx1588): pHW531, 10 ng/µL; unc-119(+) rescue plasmid, 50 ng/µL; pBlueScript, 40 ng/µL.

PS7683 *syEx1586; unc-119(ed3); syIs300 V*
PS7684 *syEx1587; unc-119(ed3); syIs300 V*
PS7685 *syEx1588; unc-119(ed3); syIs300 V*

Figures 3.9-3.12
Split cGAL drivers for MC neurons (syIs483, syIs484 and syIs485): pJL080, 25 ng/µL; pJL081, 25 ng/µL; Punc-122::rfp, 30 ng/µL; 1 kb ladder (NEB), 20 ng/µL.

PS7521 *syIs483 X*	*outcrossed x3*
PS7522 *syIs484*	*outcrossed x0*
PS7523 *syIs485*	*outcrossed x0*

Split MC driver (syIs483) > HisCl1 effector (syIs371, 15xUAS::HisCl1::SL2::gfp::let-858 3'UTR)
PS7524 *syIs371 III; syIs483 X*
PS7199 *syIs371 III*

Split cGAL-N driver alone for MC neurons (syEx1601 and syEx1602): pJL080, 25 ng/µL; Pttx3::rfp, 40 ng/µL; 1kb ladder (NEB), 35 ng/µL.

PS7739 *syEx1601; syIs371 III*
PS7740 *syEx1602; syIs371 III*

Split cGAL-C driver alone for MC neurons (syEx1603 and syEx1604): pJL081, 25 ng/µL; Pttx3::rfp, 40 ng/µL; 1kb ladder (NEB), 35 ng/µL.

PS7741 *syEx1603; syIs371 III*
PS7742 *syEx1604; syIs371 III*

Figures 3.13, 3.14
Dominant PKA effector (syEx1596 and syEx1597): pHW539, 25 ng/µL; Pttx-3::rfp, 40 ng/µL; 1 kb ladder (NEB) 35 ng/µL.

PS7525 *syEx1596*
PS7526 *syEx1597*
PS7527 *syEx1596; syIs483 X*
PS7528 *syEx1597; syIs483 X*