

II. GENERAL AUDIENCE PROJECT SUMMARY

Multicellular animals, such as fruit flies and humans, use DNA sequences known as enhancers to switch “ON” genes in the correct cells (space), and the correct time point during life (time). With enhancers sequences consisting of nothing more than a series of DNA letters (A, C, G, and T), these DNA sequences encode spatial and temporal instructions for switching genes ON. Moreover, mutations in these sequences can cause switch activities to differ, which can result in variation between individuals of the same species, and over evolutionary time an accumulation of such mutations can result in the adaptation of species to changing environments and the evolution of altogether new species. Reporter transgenes are an effective tool to look at the switch like activities of enhancers, where the activity can be seen by detecting the production of a fluorescent protein. However, comparisons of different enhancers in the same animal requires multiple fluorescent proteins whose color emission properties are clearly distinguishable and whose activity at one time point can be distinguished from those occurring at a later time point. My thesis aims to customize tools for the study of enhancer functions by comparing a suite of fluorescent proteins and fluorescent timer proteins in the fruit fly species *Drosophila melanogaster*.

III. PROPOSED THESIS TITLE AND PROPOSED ABSTRACT

Seeing gene expression in space, time, and color: evaluating new fluorescent proteins for the study of gene regulation in fruit flies

The formation of animal bodies occurs through a continuum of changes in the spatial locations of gene expression during the time frame of development. The expression patterns for genes are directed by DNA sequence called enhancers. A major limitation of contemporary genetics is in understanding how these DNA sequences encode information specifying when during development and in what cell types a gene will be expressed. One method to study enhancers is the use of reporter genes, where an enhancer is coupled with the protein coding sequence for a Fluorescent Protein. When inserted into animals, the function of an enhancer can be visualized by observing where and when the Fluorescent Protein is produced. Comparisons of multiple reporter transgenes with distinct fluorescent proteins in the same animal are limited in utility by overlaps in the fluorescence emission spectrum for commonly used Fluorescent Proteins. Compared to the turnover of the endogenous protein regulated by an enhancer, Fluorescent Proteins linger in cells and thereby preclude a visualization of any temporal changes in enhancer activity. The aim of this proposal is to evaluate various Fluorescent Proteins that include diverse and dynamic emission spectra with a *Drosophila melanogaster* enhancer that controls a spatially restricted and temporally changing pattern of expression. The outcomes are likely to identify and adapt an improved set of Fluorescent Proteins tools for the mechanistic study of gene regulation.

IV. Project Description

A. Specific Aims: This project will address two general hypotheses. (1) That blue, red shifted and far-red shifted **Fluorescent Proteins (FPs)** can be coupled with Green Fluorescent Protein for dual labeling experiments. (2) That **Fluorescent Timer (FT)** proteins will enable the visualization of temporally changing enhancer activities.

Aim 1: Compare the color emission spectrum outputs of six Fluorescent Proteins in *Drosophila (D.) melanogaster* pupae. These are EGFP-NLS, Cerulean-NLS, DsRed.T4-NLS, mCherry-NLS, dTomato-NLS, and E2-Crimson-NLS.

Aim 2: Test the capabilities of two different Fluorescent Timer proteins to reveal temporal alterations in enhancer activity during the time course of development.

B. Background and Significance: Embryonic development occurs through the choreographed expression of thousands of genes in an entity that is constantly changing in cell number and cell types, and overall morphology. This choreography is achieved through the process of gene regulation where the expression (making of a gene's functional product) patterns for genes are limited to certain cell types and certain times during development [1]. These expression patterns are encoded in and directed by DNA sequences known as **enhancers**. While much remains to be understood about the encoding of enhancer functions, a general understanding has been reached. In this, the characteristic expression output specified by an enhancer is encoded by the DNA sequence as a collection of short motifs that are bound by certain transcription factor proteins [2]. As these regulatory proteins are limited in expression to certain cell types

and times during development, an enhancer can only direct expression in cells where the right combination of transcription factors are present in the nucleus.

Studies of the human and model organism genomes has revealed that the enhancers outnumber genes in genomes by several fold [3–5]. This stems from genes being regulated by multiple enhancers. It is suspected that mutations in enhancers account for many genetic diseases and disease risk factors, though few causative mutations have been found [6,7]. Several case studies are known where the evolution of animal morphology was shaped by changes in enhancer functions [8,9]. Hence, there is great interest in understanding how an assemblage of “A”, “C”, “G”, and “T” letters – the primary DNA building blocks – can encode enhancer functions and how changes in these letters can result in disease, disease risk, or even evolutionary novelty.

One popular method to study enhancer function is to couple an enhancer to the protein coding sequence for a gene that makes an easily observable product [10]. When these reporter transgenes are put into a transgenic animal, the activity of the enhancer can be visualized by observing where the reporter protein is produced. **Fluorescent Proteins (FPs)** make excellent reporters as their expression can be detected from fluorescent light emissions following excitation by light of a specific wavelength. The most widely used FP is Enhanced Green Fluorescent Protein, or EGFP, that is excited by 488 nanometer light (nm) and emits light between the 480 and 560 nm.

The use of EGFP has two limitations which this proposal seeks to ameliorate. First, the emission spectrum for EGFP overlaps with that for other widely used FPs. This includes DsRed2, which has a spectrum of 540-650 nm and for which newly translated protein is in an immature green fluorescent state that over time matures to a

red state [11]. Due to the spectral overlap, dual-labeling experiments (the use of two transgenes with different enhancers or a normal and mutant version of the same enhancer) are problematic as it remains uncertain whether fluorescence is coming from one, the other, or both FPs. The second limitation has to do with the perdurance of FPs within cells [12]. This lingering presence prevents the observation of temporal changes in the activity of an enhancer, notably the cessation of or a change in expression.

The Williams lab studies how gene expression patterns are encoded in enhancer sequences. The **anterior element** is an enhancer that drives a robust expression of the Bab2 protein in the more anterior A3 and A4 segments of the *D. melanogaster* pupal abdomen (Appendix 1) [13]. Bab1 expression declines dramatically towards the end of pupal development, although a comparable declination in enhancer activity cannot be seen using the EGFP reporter transgene (Appendix 1). I will use the anterior element to drive the expression of various FPs to find those which eliminate or minimize spectral overlap with EGFP (Aim 1). I will also use this enhancer with FPs that change color over their lifespan, and behave like a **Fluorescent Timer (FT)**, to look for changes in anterior element activity between early and late pupal stages (Aim 2).

C. Research Strategy

Aim 1. Research Methods and Strategy: FASTA sequence files were made with the coding sequences of the nuclear localization signal (NLS) of the *D. melanogaster transformer* gene [14] fused in-frame to the 3' end of the coding sequences of *mCherry* and *E2-Crimson* proteins. These sequences were synthesized and cloned into the pUC57 vector (GenScript USA Inc.). Cerulean-NLS, dTomato-NLS, and DsRed.T4-

NLS containing vectors were previously constructed [15,16]. I will separately sub-clone each of these FP sequences into a reporter transgene vector containing the anterior element and for which the *EGFP-NLS* sequence was removed. The resulting reporter transgenes will be integrated into the *D. melanogaster attP40* genome site by site-specific integration methods in order to make transgenic fruit flies [17,18]. Pupa will be removed from their outer pupariums at **40 hours after puparium formation (hAPF)**, a time point of anterior element activity, and mounted on a glass slide. The FP expression patterns will be determined for blue, green, red, and far-red color spectra by confocal microscopy (Appendix 2). Detection of green fluorescent light, FP excitation will be achieved using 488 nm light and emitted light collected between 500-550 nm. Blue, red and far-red light will respectively be detected using 458, 543 and 633 nm light with emitted light collected between 450-490, 560-700 and 610-750 nm (Appendix 2).

Possible Outcomes: I anticipate that EGFP fluorescence will only be detected in the abdomens of pupa when green fluorescence excitation and emission spectrum settings are used. I suspect that DsRed.T4-NLS, mCherry-NLS, and dTomato-NLS will be best detected using the red settings and to a lesser extent the far-red settings, whereas the Cerulean and E2-Crimson-NLS will only be detected using blue and far-red settings respectively. The outcome sought is the identification of FPs suitable for dual-labeling studies. I anticipate this outcome is most likely for E2-Crimson-NLS, as its emission spectrum is most separated from EGFP [19], though it is unclear whether this protein proceeds through an immature green state as seen for DsRed.T4-NLS (Appendix 2).

Progress: The transgenes with the DsRed.T4-NLS, E2-Crimson-NLS and mCherry-NLS sequences were made, and are being incorporated into transgenic fruit flies.

Aim 2. Research Methods and Strategy: The Timer-1 FT proceeds from an immature green fluorescent state to a mature red state [20,21]. The Fast-FT timer proceeds from an immature blue state to a mature red state [22]. I will excise the Timer-1 and Fast-FT coding sequences respectively from the RIP Timer and Fast-FT vectors. These FT sequences will be sub-cloned in the place of the EGFP-NLS sequence in the anterior element containing vector. The resultant reporter transgenes will be incorporated into transgenic fruit flies as described in Aim 1. Transgenic pupa will be collected for confocal analysis at early (40 hAPF) and late time points (85 hAPF). Blue, green, and red fluorescence will be collected using the respective excitation and emission collection wavelengths: 458 nm and 440-490 nm, 488 and 500-550 nm, and 543 and 560-700 nm.

Possible Outcomes: Bab2 expression, but not anterior element-driven EGFP expression, declines between early and late pupal stages, presumably by the anterior element [13]. I suspect that the fluorescence of the FT proteins will recapitulate this temporal change. For Timer-1, green fluorescence will be robust early and new expression will cease at a later pupal stage (Appendix 3). This would be evident as a transition from green (immature) to red (mature) fluorescence. A recent study suggests that the green form Timer-1 does not progress to the mature red form, and instead is a dead end product [23]. This would be evident as a persistent pattern of green fluorescence on top of a later occurring red fluorescence. The blue to red maturation of Fast-FT proceeds by the maturation of an immature blue fluorescent form. I suspect that this FT will better document early and late stage activity of anterior element activity as an absence of blue fluorescence at the later time point (Appendix 3).

Progress – FT Vectors were obtained from the addgene repository (www.addgene.org).

V. Timeline

Spring Term 2014 – Complete construction of anterior element reporter transgenes with DsRed.T4-NLS, mCherry-NLS, and E2-Crimson-NLS (Aim 1). Send the transgenes to Best Gene Inc. to have transgenic *Drosophila melanogaster* lines derived.

Summer 2014 – Complete construction of anterior element reporter transgenes with Cerulean-NLS, dTomato-NLS, Timer-1 and Fast-FT sequences (Aim 2) and have transgenic lines created. Survey fluorescence expression patterns for Aim 1 transgenes via confocal microscopy.

Fall 2014 – Survey fluorescent expression patterns at early and late pupal stages for Aim 2 transgenes via confocal microscopy.

Spring 2015 – During this semester, I will prepare my results for an oral presentation at the University of Dayton's honors symposium (March, 2015), and for a poster presentation at the Stander Symposium (April, 2015).

VI. Bibliography

1. Davidson EH (2006) The Regulatory Genome: Gene Regulatory Networks In Development And Evolution. Burlington, MA: Elsevier Inc. p.
2. Arnone MI, Davidson EH (1997) The hardwiring of development: organization and function of genomic regulatory systems. *Development (Cambridge, England)* 124: 1851–1864. Available:<http://www.ncbi.nlm.nih.gov/pubmed/9169833>.
3. Pennisi E (2012) ENCODE Project Writes Eulogy For Junk DNA. *Science* 337: 1159–1160.

4. Nègre N, Brown CD, Ma L, Bristow CA, Miller SW, et al. (2011) A cis-regulatory map of the *Drosophila* genome. *Nature* 471: 527–531.
Available:<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3179250&tool=pmcentrez&rendertype=abstract>. Accessed 9 March 2012.
5. Shen Y, Yue F, McCleary DF, Ye Z, Edsall L, et al. (2012) A map of the cis-regulatory sequences in the mouse genome. *Nature*: 1–5.
Available:<http://www.nature.com/doi/10.1038/nature11243>. Accessed 2 July 2012.
6. Visel A, Akiyama JA, Shoukry M, Afzal V, Rubin EM, et al. (2009) Functional autonomy of distant-acting human enhancers. *Genomics* 93: 509–513.
Available:<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2683195&tool=pmcentrez&rendertype=abstract>. Accessed 3 November 2011.
7. Sethupathy P, Collins FS (2008) MicroRNA target site polymorphisms and human disease. *Trends in genetics : TIG* 24: 489–497.
Available:<http://www.ncbi.nlm.nih.gov/pubmed/18778868>. Accessed 6 September 2011.
8. Carroll SB (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134: 25–36.
Available:<http://www.ncbi.nlm.nih.gov/pubmed/18614008>. Accessed 24 July 2011.
9. Martin A, Orgogozo V (2013) The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution; international journal of organic evolution* 67: 1235–1250.
Available:<http://www.ncbi.nlm.nih.gov/pubmed/23617905>. Accessed 6 August 2013.
10. Rebeiz M, Williams TM (2011) Experimental Approaches to Evaluate the Contributions of Candidate Cis- regulatory Mutations to Phenotypic Evolution. In: Orgogozo V, Rockman M V., editors. *Methods in Molecular Biology*. Totowa, NJ: Humana Press, Vol. 772. pp. 351–375.
Available:<http://www.springerlink.com/index/10.1007/978-1-61779-228-1>. Accessed 9 November 2011.
11. Baird GS, Zacharias D a, Tsien RY (2000) Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. *Proceedings of the National Academy of Sciences of the United States of America* 97: 11984–11989.
Available:<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=17281&tool=pmcentrez&rendertype=abstract>.

12. Tsien RY (2010) Fluorescence readouts of biochemistry in live cells and organisms. In: Weissleder R, Gambhir SS, Rehemtulla A, editors. *Molecular Imaging: Principles and Practice*. McGraw Hill. pp. 808–828.
13. Williams TM, Selegue JE, Werner T, Gompel N, Kopp A, et al. (2008) The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* 134: 610–623.
Available:<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2597198&tool=pmcentrez&rendertype=abstract>. Accessed 4 August 2011.
14. Hedley ML, Amrein H, Maniatis T (1995) An amino acid sequence motif sufficient for subnuclear localization of an arginine/serine-rich splicing factor. *Proceedings of the National Academy of Sciences of the United States of America* 92: 11524–11528.
Available:<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=40434&tool=pmcentrez&rendertype=abstract>.
15. Barolo S, Castro B, Posakony JW (2004) New *Drosophila* transgenic reporters: insulated P-element vectors expressing fast-maturing RFP. *BioTechniques* 36: 436–440, 442. Available:<http://www.ncbi.nlm.nih.gov/pubmed/15038159>.
16. Livet J, Weissman TA, Kang H, Draft RW, Lu J, et al. (2007) Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* 450: 56–62.
Available:<http://www.ncbi.nlm.nih.gov/pubmed/17972876>. Accessed 6 August 2013.
17. Markstein M, Pitsouli C, Villalta C, Celniker SE, Perrimon N (2008) Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature genetics* 40: 476–483.
Available:<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2330261&tool=pmcentrez&rendertype=abstract>. Accessed 17 July 2011.
18. Groth AC, Fish M, Nusse R, Calos MP (2004) Construction of Transgenic *Drosophila* by Using the Site-Specific Integrase From Phage phiC31. *Genetics* 166: 1775–1782.
19. Strack RL, Hein B, Bhattacharyya D, Hell SW, Keenan RJ, et al. (2009) A rapidly maturing far-red derivative of DsRed-Express2 for whole-cell labeling. *Biochemistry* 48: 8279–8281.
Available:<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2861903&tool=pmcentrez&rendertype=abstract>. Accessed 7 August 2013.

20. Wittkopp PJ, Vaccaro K, Carroll SB (2002) Evolution of yellow gene regulation and pigmentation in *Drosophila*. *Current biology* 12: 1547–1556. Available:<http://www.ncbi.nlm.nih.gov/pubmed/12372246>.
21. Terskikh A, Fradkov A, Ermakova G, Zeraisky A, Tan P, et al. (2000) “Fluorescent Timer”: Protein That Changes Color with Time. *Science* 290: 1585–1588. Available:<http://www.sciencemag.org/cgi/doi/10.1126/science.290.5496.1585>. Accessed 7 August 2013.
22. Subach F V, Subach OM, Gundorov IS, Morozova KS, Kiryl D, et al. (2010) report on cellular trafficking. 5: 118–126. doi:10.1038/nchembio.138.Monomeric.
23. Verkhusha V V, Chudakov DM, Gurskaya NG, Lukyanov S, Lukyanov KA (2004) Common Pathway for the Red Chromophore Formation in Fluorescent Proteins and Chromoproteins. *Chemistry & Biology* 11: 845–854. doi:10.1016/j.

VII. BUDGET

Itemized Budget:

Materials/Supplies:	
Item	Estimated Cost
Gene Synthesis of <i>E2-crimson-NLS</i> and <i>mCherry-NLS</i> coding sequences (GenScript) and purchase of four reporter gene vectors from addgene repository.	\$1,000.00
QIAprep Spin Miniprep kit (Qiagen)	\$250.00
Ascl and SbfI restriction enzymes (New England Biolabs)	\$500.00
T4 DNA ligase (New England Biolabs)	\$250.00
200 DNA sequencing reactions at \$5.00 each (DNA analysis, LTD)	\$1,000.00
4 cases of Drosophila vials and cotton balls (Fisher Scientific)	\$100.00
8 transgene injection services (Best Gene Inc, \$250 each)	\$2,000.00
Travel: (If required for Thesis Project)	
Mileage	X miles X (\$0.55/mile) =
	*The current UD rate for mileage reimbursement will be used.
Other:	
Item	Estimated Cost
Other Source(s) of Funding	
Source	Pending/Secured

Total Amount Requested: \$ 1,500.00

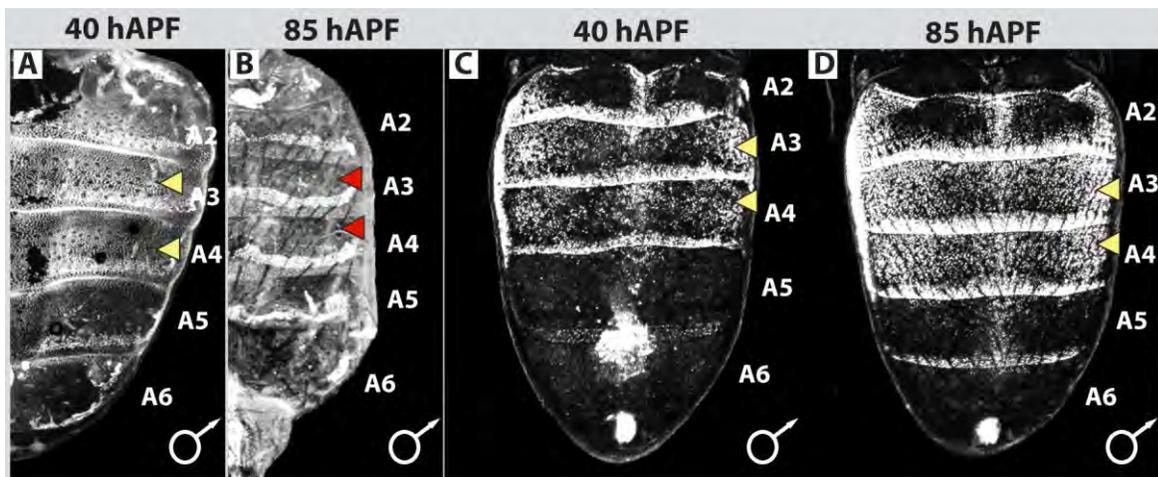
Narrative Budget Justification:

To complete the proposed aims, this project will incur costs that include costs totaling at least \$5,500.00. We are having sequences for custom made genes synthesized by a commercial vendor. This project will require the construction of at least 8 transgene vectors. This requires various enzymes, a DNA plasmid preparation kit, and DNA sequencing reactions. Each vector must be injected into *Drosophila melanogaster* embryos to derive transgenic lines (The Best Gene Inc.). Taking care of transgenic lines will require plastic vials and cotton balls.

VIII. LETTER OF SUPPORT

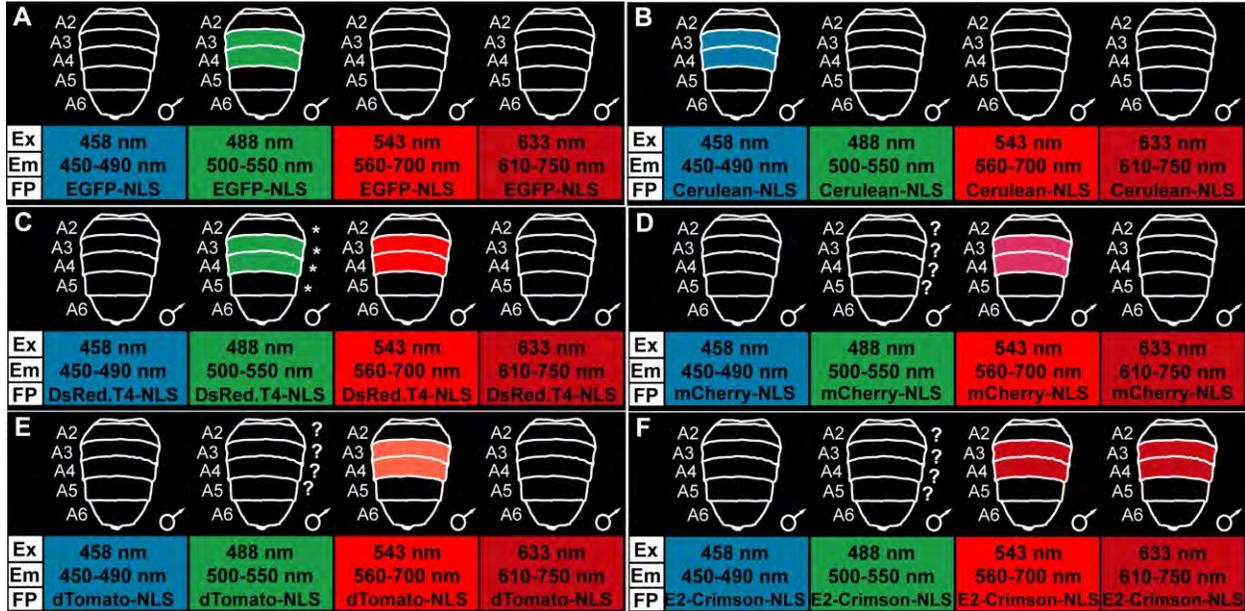
The letter from my advisor, Dr. Thomas Williams an Assistant Professor in the Department of Biology at the University of Dayton, was emailed to **Ramona Speranza** on 3/6/14.

IX. Appendices



Appendix 1. Endogenous Bab2 expression compared to reporter transgene

driven expression. (A) At an early pupal developmental stage, Bab2 is expressed at a high level in the male A3 and A4 abdominal segment epidermis (yellow arrowheads). (B) At a late pupa developmental stage, Bab2 expression is no longer observed in the male A3 and A4 abdominal epidermis (red arrowheads). The bright horizontal stripes on segments A2-A5 are the oenocyte cells which are regulated by another enhancer than the one studied here. EGFP-NLS expression driven by the enhancer known as the anterior element is observed in (yellow arrow heads) the A3 and A4 segments of transgenic pupa at (C) early and (D) late pupal stages.

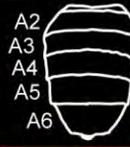


Appendix 2. Conditions and anticipated outcomes for fluorescence properties of

various Fluorescent Proteins. (A-F) Fluorescent proteins will be evaluated at settings optimized for blue, green, red, and far-red fluorescence. Expression will be driven by the *D. melanogaster* anterior element, an enhancer that is active in the A3 and A4 abdominal segments of pupa. “Ex” stands for excitation wavelength, “Em” is the emission spectra to be collected, and “FP” stands for Fluorescent Protein. The “*” indicates a green fluorescence that results from expression of the DsRed.T4-NLS FP. “?”s indicate the uncertainty as to whether other red fluorescent proteins cause a green fluorescence.

Appendix 3. Conditions and anticipated outcomes for fluorescence properties of two Fluorescent Timers. (A-C)

Fluorescent protein expressions driven by the anterior element in transgenic *D. melanogaster* pupa. (A) Green but not red fluorescence is observed for the EGFP-NLS Fluorescent Protein at 40 - early stage - and 85 - late

	Early Pupa		Late Pupa	
A				
Ex	488 nm	543 nm	488 nm	543 nm
Em	500-550 nm	560-700 nm	500-560 nm	560-700 nm
FP	EGFP-NLS	EGFP-NLS	EGFP-NLS	EGFP-NLS
B				
Ex	488 nm	543 nm	488 nm	543 nm
Em	500-550 nm	560-700 nm	500-550 nm	560-700 nm
FP	Timer-1	Timer-1	Timer-1	Timer-1
C				
Ex	458 nm	543 nm	458 nm	543 nm
Em	440-490 nm	560-700 nm	440-490 nm	560-700 nm
FP	Fast-FT	Fast-FT	Fast-FT	Fast-FT

stage - hours after puparium formation. (B) Timer-1 should result in a green fluorescence for newly translated immature Fluorescent Protein, and following maturation the fluorescence will be red. (C) Fast-FT should result in a blue fluorescence for newly translated immature Fluorescent Protein, and following maturation the fluorescence will be red. The expectation is for the anterior element to no longer drive expression at the late time point, which would be supported by an absence of green and blue fluorescence respectively for the Timer-1 and Fast-FT bearing transgenes. “Ex” stands for excitation wavelength, “Em” is the emission spectra to be collected, and “FP” stands for Fluorescent Protein coding sequence included in the reporter transgene. The “?”s indicate the uncertainty as to whether the Timer-1 protein has a green fluorescence resulting from a population of protein product that will not mature to the red form.