UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA ANIMAL

Molecular Physiology and Evolution of a New Developmental Stability Pathway

Mestrado em Biologia Evolutiva e do Desenvolvimento

Catarina Sofia Duarte Nunes

Dissertação orientada por:
Dr. Élio Sucena, FCUL, IGC
Dr. Alisson Gontijo, CEDOC

2015

Agradecimentos

Quando chega o momento de agradecer, parece sempre mais fácil fazê-lo por escrito. Especialmente para mim, que sempre tive alguma dificuldade em expressar-me através da palavra falada, mas nunca através da escrita. Serve então este espaço para agradecer, sem rodeios, a todos aqueles que me acompanharam no ultimo ano e que, de alguma forma, contribuíram para que esta fase da minha vida parecesse mais fácil de superar e suportar.

Em primeiro lugar, nunca poderia deixar de agradecer ao Alisson por me ter recebido de uma forma tão calorosa e por sempre me ter ajudado ao longo deste tempo, exigindo de mim sempre mais e melhor. Tudo aquilo que me ensinou ficará para sempre comigo e certamente que me será útil em muitos outros desafios. A aprendizagem não tem preço e por isso nunca lhe poderei agradecer o suficiente. À Fabiana tenho que agradecer por me ter ensinado tudo aquilo que sei acerca de genética de *Drosophila* e por nunca ter desistido de me explicar as coisas que, em certos momentos, me pareciam ser difíceis de perceber. Aos dois, em conjunto, tenho a agradecer pelo laboratório com bom ambiente que criaram e por receberem cada um de nós como se fôssemos parte da vossa família.

Aos meus colegas de laboratório tenho, acima de tudo, que agradecer o companheirismo e a amizade que conto levar comigo. À Andreia, um obrigado sem tamanho por sempre me ter ajudado e acompanhado, por ser minha parceira de experiências e por sempre ter estado disponível para me orientar e ouvir. Lembro-me de ter escrito a meu respeito: “Espero poder retribuir um dia”; e retribuiu, em dobro. Um obrigado ao André por ter estado sempre disponível para me ajudar em tudo o que envolvia informática ou análise de resultados que estava para lá dos meus conhecimentos na altura. O facto de abdicar das suas próprias coisas para ajudar um colega diz muito acerca de si. Um obrigado à Ângela por todo o companheirismo e apoio que me prestou durante os últimos meses. À Maria João, um obrigado pela lufada de ar fresco que trouxe consigo e por ter sido uma fonte de apoio e, ao mesmo tempo, por sempre me ajudar a ver o copo meio cheio. Finalmente, à Márcia, agradeço ter feito esta jornada comigo e por termos partilhado tantos momentos de alegria e de desespero que, de alguma forma, nos ajudaram a criar laços. A todos, agradeço ainda a algazarra que me impede de trabalhar concentrada! Sem vós não seria o mesmo!

À minha colega de mestrado e amiga Margarida, um obrigado por ter partilhado tantas horas de pausa e de trabalho comigo e por sempre me ter dado a confiança suficiente para achar que nada é impossível. A amizade, essa, nem se agradece, retribui-se. Agradeço à Ana Sofia pela companhia em tantas viagens e por ter ouvido tantas vezes os meus desabafos, tendo sempre uma palavra de apoio para me dar.

Um obrigado aos meus pais, por sempre me terem ensinado a lutar pelos meus objetivos e por me terem artilhado para as lutas, fazendo tudo o que estava ao seu

alcance para que a conquista ficasse um passo mais perto. Ao meu Paulo, obrigada por tantas vezes ter suportado que descarregasse nele tantas frustrações e por sempre ter estado disposto a ouvir-me e a aconselhar-me, mesmo sem entender as razões das minhas inquietações. À Nana e à minha irmã, obrigada por sempre terem acreditado em mim, mesmo em momentos em que eu própria duvidei das minhas capacidades.

Aos meus amigos e companheiros Joana, Sara, Marta, Artur, Zé, Ana e Carolina obrigada por estarem sempre presentes e dispostos a proibir conversas sobre trabalho nos nossos encontros.

Um agradecimento especial à Dra. Tatiana Torres e à suas estudantes Gisele Cardoso e Raquel Monfardini por toda a ajuda prestada nas análises de RNAseq e por sempre se terem disponibilizado para me ajudar na interpretação de resultados e no desenho de novas abordagens experimentais.

Last, but definitely not least, I would like to thank Dr. Takashi Koyama, however I cannot. It is not that I don't have anything to thank for, but rather that there are not enough words yet to thank all that he have done for me during the past year. Thank you for encouraging me to always be better and for pushing me. Thank you for never giving up on me and for always finding time to help me. Thanks for all the effort you put in this thesis and for the fast feedback that you always gave on my writing. Thank you for all you taught me. Thank you for being my advisor, my **mentor** but most of all, my friend. I will always keep you in my heart.

Um agradecimento final também à Faculdade de Ciências da Universidade de Lisboa, por ter contribuído para a minha formação académica e por nos proporcionar um ensino de excelência.

Sumário

Em *Drosophila melanogaster*, após a eclosão do ovo, a larva desenvolve-se ao longo de três estádios larvares, que são seguidos pela formação da pupa, terminando assim a fase de crescimento. O tempo das várias transições do desenvolvimento do insecto é influenciado por factores ambientais e é regulado pela hormona esteroide, ecdisona. Os padrões de crescimento e morfogénese a que o animal é sujeito durante o período de desenvolvimento são determinados e, portanto, previsíveis. No entanto, para que seja produzido um organismo com dimensões corporais apropriadas e em sintonia com o ambiente, estes processos podem ser afectados por condições extrínsecas (como a nutrição) ou por perturbações intrínsecas que ocorram, por exemplo, durante a regeneração de um órgão e/ou tecido. Esta capacidade que um organismo tem de lidar com estas perturbações, refletindo proporções corporais corretas é designada por estabilidade de desenvolvimento. Esta é uma propriedade muito relevante para o desenvolvimento, pois a alteração das dimensões corporais e/ou de órgãos de um animal pode afectar a sua *fitness* e, consequentemente, o seu *output* reprodutivo.

Na larva da grande maioria dos insectos holometabólicos (insectos que sofrem uma metamorfose completa entre os estádios de ovo, larva, pupa e adulto), os discos imaginais representam os percursores da maioria dos apêndices do indivíduo adulto, tais como as asas ou as patas. Os discos imaginais têm a capacidade de se regenerarem durante a fase de crescimento larvar, mesmo após dano químico (por agente alquilante, por exemplo) ou físico (por irradiação raio-x). No entanto, quando estas estruturas são danificadas, a transição de larva para pupa é atrasada o tempo suficiente para que as células ajustem as suas taxas de proliferação, permitindo a reparação dos danos. Sendo que os discos imaginais estão remotamente localizados em relação ao cérebro, onde são produzidas as hormonas responsáveis pelo controlo do tempo de desenvolvimento, o mecanismo através do qual as deficiências de crescimento influenciam o tempo de metamorfose tem que ser dependente de um ou mais sinais difusíveis. Ao sentirem que não atingiram o tamanho apropriado, os discos comunicam o seu estado anormal ao resto do corpo, através da secreção de um sinal que impede a pupariação até que seja atingida a dimensão ideal.

Recentemente, dois grupos de investigação independentes descobriram o principal elemento envolvido neste mecanismo de regulação: Dilp8. Dilp8 é um péptido pertencente à superfamília das insulinas/IGFs/relaxinas, que é secretado pelos discos imaginais em situações de crescimento anormal. Todos os péptidos pertencentes a esta família apresentam uma estrutura semelhante àquela das preproinsulinas, consistindo na associação contígua entre um péptido sinal, uma cadeia B, um péptido-C e uma cadeia A. Nas insulinas e relaxinas (mas não nos IGFs), o péptido-C é clivado pela ação de uma convertase, que reconhece sequências de clivagem específicas, dando origem a um péptido maturo ativo. Apesar da importância de Dilp8 na regulação

da estabilidade do desenvolvimento já ter sido demonstrada, até à data continua por determinar qual o mecanismo de ação molecular que permite a coordenação do crescimento anormal e do tempo desenvolvimento.

Este projeto pretendia então explorar os mecanismos moleculares de ação e a evolução de Dilp8. Para tal, foram utilizadas moscas transgénicas e o sistema de sobrexpressão UAS-Gal4. Inicialmente, foi desenvolvido um estudo de evolução molecular que revelou a existência de homólogos de Dilp8 em várias espécies de dípteros, pertencentes ao ramo Brachycera. O alinhamento das sequências proteicas destes homólogos permitiu a identificação de resíduos de aminoácidos absolutamente conservados entre as diferentes espécies. Para tentar perceber quais destes resíduos estavam intimamente associados à função desempenhada por Dilp8 no controlo do tempo de desenvolvimento, foram criadas linhas de moscas que continham uma versão alterada de Dilp8. Nestas linhas, cada um destes aminoácidos foi, individualmente, substituído por outro ou completamente eliminado. O tempo do início da metamorfose foi analisado nestas linhas de moscas sobrexpressando de forma ubíqua cada um dos péptidos alterados e verificou-se que, de facto, a alteração de alguns destes aminoácidos levou à eliminação do atraso no desenvolvimento, provocado pela sobrexpressão da versão *wild-type* de Dilp8. Os dados obtidos parecem indicar que Dilp8 precisa de ser processado para que o péptido maturo promova a estabilidade do desenvolvimento. No entanto, embora a produção da versão não processada do péptido tenha sido confirmada em todas as linhas transgénicas em estudo, em nenhum dos casos foi possível a detecção do péptido maturo por Western Blot.. Tal dificultou a confirmação da hipótese sugerida pelos resultados obtidos. Para obter mais conhecimento acerca do papel do péptido-C na regulação do tempo de desenvolvimento, este foi sobrexpresso sozinho (sem qualquer outro elemento de Dilp8) e verificou-se que este não apresentou a capacidade de atrasar a metamorfose.

Em *D. melanogaster*, a sobrexpressão de Dilp8 tem outro efeito para além do atraso no tempo de desenvolvimento: as moscas tornam-se mais pesadas sem, no entanto, aumentarem as suas dimensões corporais. De modo a explorar o envolvimento do péptido-C neste fenótipo, as pupas de diferentes genótipos de sobrexpressão de péptidos alterados foram pesadas. Não foram obtidas diferenças de peso diretamente relacionadas com a alteração do péptido-C, indicando que esta região de Dilp8, embora possa estar envolvida noutra funções biológicas, não é diretamente responsável pelo aumento de peso verificado.

O estudo das sequências homólogas de Dilp8 permitiu ainda explorar a história evolutiva deste péptido. Para tal, foi utilizada uma espécie mais ancestral, *Hermetia illucens*, para a qual um RNAseq revelou a existência de, pelo menos, duas sequências homólogas de Dilp8. Para verificar se esta versão ancestral do péptido (Hilp8) mantém a capacidade de induzir atrasos no desenvolvimento quando sobrexpresso em *D. melanogaster*, um dos homólogos de Dilp8 foi clonado nesta espécie. Os resultados desta abordagem revelaram que Hilp8 não tem a capacidade de regular o tempo de

desenvolvimento de *Drosophila*. Porém, tal como discutido ao longo do trabalho, tal não indica diretamente que a forma ancestral de Dilp8 apresentava uma função diferente em *H. Illucens*. Este trabalho permitiu ainda confirmar a existência de, pelo menos, duas regiões do genoma diferentes responsáveis pela codificação de duas sequências homólogas de Dilp8. A análise estrutural destas duas sequências permitiu, ainda, a distinção de, pelo menos, três exões para ambas, tal como é característico da maioria dos Ilps.

De modo a entender se Hilp8 apresenta outro tipo de homologia funcional em relação a Dilp8, uma nova análise de RNAseq foi desenvolvida, com a utilização de larvas de *H. Illucens* sujeitas à injeção do agente alquilante EMS. Em *Drosophila*, a exposição a esta droga leva à ativação da via de Dilp8 e, consequentemente, a atrasos na pupariação. Em *H. Illucens* não foi possível a determinação de um aumento da transcrição de *hilp8*, após a exposição a EMS. Porém, foi detectada uma elevada variação de expressão entre amostras biológicas. Esta observação dificulta a determinação do envolvimento de Hilp8 na resposta ao dano. Para estudar a homologia funcional de Dilp8 e das duas formas de Hilp8, a expressão dos genes no tecido reprodutivo de fêmeas foi analisada *H. Illucens* adultas e, contrariamente ao que é verificado em *Drosophila*, estes não se encontram enriquecidos em ovários. Este resultado pode indicar um funcionamento divergente dos dois homólogos.

Em suma, este estudo, com a utilização de abordagens fisiológicas, moleculares e de evo-devo permitiu aumentar o nosso conhecimento acerca do funcionamento e da história evolutiva de Dilp8.

Palavras-chave: Dilp8; *D. melanogaster; H. illucens;* tempo de desenvolvimento; msturação peptídica

Abstract

In insects, development follows predictable growth and morphogenesis patterns that involve both tight control and flexibility in the regulation of cell size and number, in order to produce an animal with proper body and organ size. However, as environmental and intrinsic perturbations can affect these processes, organisms have evolved the capability to buffer these perturbations through developmental stability processes. Abnormally growing imaginal discs of *Drosophila* delay pupariation by secreting Dilp8, an insulin/IGF/relaxin-like peptide. Dilp8 promotes developmental stability by coordinating the growth of the discs with the onset of metamorphosis. Although the importance of Dilp8 in the regulation of metamorphosis has already been demonstrated, its structure, sequence conservation and posttranslational processing have never been studied. My objective was to learn more about the molecular mechanisms of action and evolution of Dilp8. By using physiological and molecular genetics approaches, it was possible to determine conserved amino acids crucial for Dilp8 function and to propose a requirement for Dilp8 C-peptide processing for the coordination of tissue damage and developmental timing. These studies also shed light into the evolutionary history of Dilp8. An Evo-Devo approach was used to study the expression and function of Dilp8 homologues from a brachyceran fly that shared a common ancestor with *Drosophila* about 180 million years ago, *Hermetia illucens*. RNA-Seq analyses suggest *H. illucens* encodes at least two Dilp8 homologues Hilp8a and Hilp8b, none of which responds to ethylmethanesulfonate-induced tissue damage. Furthermore, ubiquitous expression of Hilp8a in *Drosophila* larvae did not delay development when expressed in *Drosophila*. These results suggest that the involvement of Dilp8 in the developmental stability pathway in Drosophila might have evolved after the last common ancestor of all Brachycera lived. However, more studies need to be developed for this to be confirmed.

Key-words: Dilp8; *D. melanogaster; H. illucens;* developmental timing; peptide maturation

Table of Contents

1. Introduction

1.1. *Drosophila melanogaster* development

One of the *Drosophilidae* (dipteran; two-winged insect) family members, *Drosophila melanogaster*, is a widely used model organism in studies in genetics and development [1,2]. After hatching, *Drosophila* larvae develop through three larval instars, followed by the puparium formation. Since pupae do not feed further, most of the growth is restricted to the larval stage (**Figure 1-1**) [3].

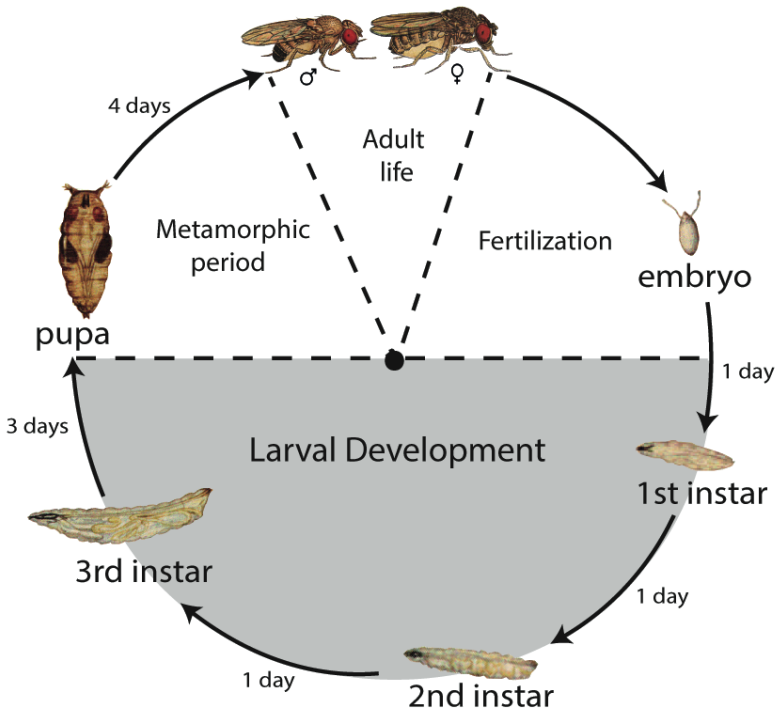


Figure 1-1 – *D.melanogaster* life cycle. Figure based on http://biology.kenyon.edu/courses/biol114/Chap13/Chapter_13A.html

Several environmental cues strongly influence developmental transitions such as molting and metamorphosis in *D. melanogaster*, as seen in other insects. These transitions are primarily regulated by peaks of the steroid molting hormone ecdysone [4]. In *D. melanogaster,* a pair of neurons from each brain hemisphere produces the prothoracicotropic hormone (PTTH) [5-8]. These PTTH-producing neurons innervate directly the prothoracic gland (PG), the primary endocrine organ to produce ecdysone during the larval stage [9]. In the PG, PTTH binds to the tyrosine-kinase receptor *Torso*, leading to the activation of the mitogen-activated protein kinase (MAPK) pathway [10]. Increased MAPK signaling upregulates expression of a series of ecdysone biosynthesis enzymes to convert dietary cholesterol to ecdysone progressively in the PG [11]. Once ecdysone is released from the PG, peripheral organs like the fat body convert ecdysone

to its biological active form, 20-hydroxyecdysone (20E). Since 20E is lipid soluble, it seems to enter most cells directly through cell membranes, from the hemolymph [7, 12].

In other insects, like the moth *Manduca sexta*, Juvenile Hormone (JH) regulates the nature of the molt [13]. In the presence of JH, an ecdysone peak induces molting to the next larval stage. However, once larvae reach a specific size at the last larval instar, JH titers drop, leading to higher peaks of PTTH and ecdysone, inducing metamorphosis (**Figure 1-2)**. The same is assumed in *D. melanogaster* [13, 14].

During the feeding period of the third instar in *D. melanogaster*, there are several small ecdysone peaks that trigger specific developmental events. The first peak induces critical weight [15], an experimentally defined developmental checkpoint. Once larvae reach critical weight, the timing of metamorphosis no longer relies on further nutrition. Therefore, critical weight seems to be a checkpoint to ensure the accumulation of sufficient amount of nutrients to undergo metamorphosis [5]. However, how larvae assess they have reached critical size is still a matter of intense study. Possibly, once a nutrition-sensing organ (presumably, fat body) senses high intracellular dietary macronutrient concentrations, it secretes a signal molecule to induce secretion of *Drosophila* insulin-like peptides (Dilps). Increased insulin/insulin-like growth factor (IGF) signaling (IIS) in the PG induces the first ecdysone peak that ultimately results in critical weight attainment [15, 16].

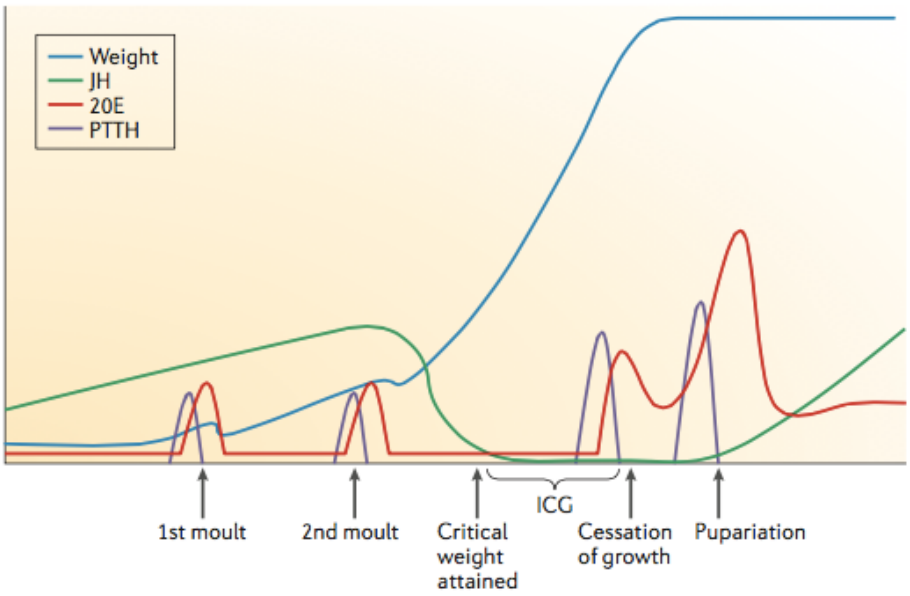


Figure 1-2 – In *Drosophila*, small PTTH and ecdysone peaks accompany every developmental transition. In the final of the third larval instar, a higher peak of 20E (the bioactive form of ecdysone) after critical weight triggers the larva-pupal molting. In *Manduca sexta*, reaching critical weight clears JH through activation of JH degradation pathway, leading to a higher expression of PTTH and ecdysone. In *Drosophila*, the same process is estimated. ICG, Interval to cessation of growth; 20E, 20-hydroxyecdysone; PTTH, prothoracicotropic hormone; JH, Juvenile Hormone. Figure from Edagr, BA (2006) [14].

1.2. Insulin-like peptides: processing and signaling pathway

Both ligand, Ilps, and signal reception pathway, IIS, are well-conserved in most metazoans. IIS integrates extrinsic signals, especially nutrition, to control growth. In *Drosophila*, this pathway regulates many developmental traits, such as growth control, metabolism, fertility and longevity [16-19].

In *D. melanogaster*, there are 8 Dilps produced in different cells at different developmental stages. Three of them (Dilp2-3 and 5) are expressed in the bilateral insulin-producing cells (IPCs), in the brain. The levels of Dilp3 and Dilp5 are dependent on the nutritional state of the flies. Expression of Dilp2, the most similar Dilp to the mammalian insulin and IGFs, is sufficient to rescue the smaller size and the developmental delay observed in the homozygous deficient *dilp1-5* mutant (*Df[dilp1-5-]*) [20, 22, 23]. The peptides produced in the IPCs further regulate the endocrine system and growth in peripheral tissues, because the IPCs innervate the ring gland (RG) and the heart (dorsal vessel) [20].

Of the other Dilps, Dilp4 is expressed in the anterior mesoderm and Dilp7 is expressed in specific cells in the ventral nerve cord. Dilp6, on the other hand, is found both in the larval gut and fat body [18, 24, 25]. Although attributing specific function to individual Dilps is difficult, due to the functional redundancy and compensation, some general functions were described for some Dilps. Overall, Dilp2, 3 and 5 regulate longevity, fecundity, lipid and carbohydrate metabolism and stress resistance [26]. Dilp6 seems to affect fecundity the most, since *dilp6* mutants present a greater reduction in lifetime fecundity [19]. Dilp6 is also involved in growth control, being *dilp6* mutants the smallest [19]. Functional data are still missing for Dilp1, 4 and 7. Dilp8 will be discussed in detail below.

Two types of membrane-bound receptors, which are activated by insulin and insulin-related peptides in vertebrates, are found in *D. melanogaster*: *Drosophila* insulin-like receptor (DInR) and Type C1-Leucine-rich repeat-containing G-protein coupled receptors (Lgrs) [21, 27-29]. The InR is a receptor Tyrosine kinase (RTK), highly similar to human insulin receptor, such that human insulin can bind to and activate DInR with high affinity [21]. As for the Lgrs, it has been hypothesized that an orphan Type C1 Lgr could be capable of binding to a Dilp with basic amino acids in signature positions in the B-chain, just as the vertebrate Type C1 Lgrs bind to Ilps of the relaxin family [21, 29]. More recently, the involvement of an orphan Lgr in a developmental pathway controlled by a specific Dilp was demonstrated in *Drosophila* (see below) [30].

All Dilps are predicted to have a similar structure to that of mammalian insulins, with the contiguous assembly of a signal peptide, B-chain, C-peptide and A-chain. Consensus cleavage sequences between B- and A-chains suggest that the mature peptides consist of two different polypeptide chains [24]. However, this is only true for insulins and relaxins. In these peptides, the C-peptide is excised by the action of a convertase enzyme by digesting cleavage sites to produce active peptides. Mature peptides consist on the B- and A-chains linked covalently by two interchain disulfide bonds and one intrachain disulfide bond (**Figure 1-3**). In contrast, IGFs contain uncleavable shortened C-peptide, resulting in single chain hormones [21, 31].

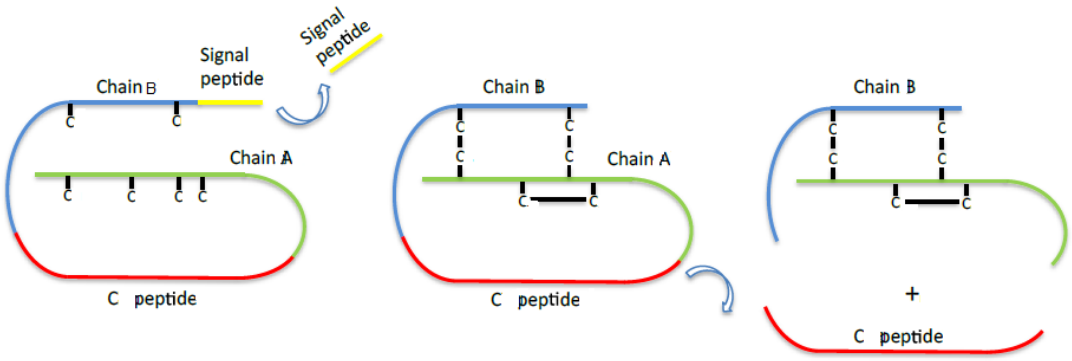


Figure 1-3 – Preproinsulin processing steps. Insulins and relaxins are encoded as preprohormones and are processed as described in the text. IGFs, on the other hand, do not cleave the C peptide and produce a single chain hormone. "C" in the figure stands for the amino acid cystein. Figure from Casimiro, AP Master Thesis (2014) [32].

Reducing IIS phenocopies starved phenotypes, such as delayed developmental timing and reduced adult size [9,17]. However, reducing functions in different components of the IIS sometimes show slightly different effects on developmental timing and body size, presumably because different components functions at different developmental stages. Total developmental time is affected only when changes in IIS occur before larvae reach critical size. After critical weight, changes in IIS only affect body and organ size, because all other physiological processes are irreversibly committed to undergo metamorphosis [16]. On the other hand, activating IIS by Dilp overexpression results in precocious pupariation and bigger flies [20, 21]. Increasing IIS in the PG is sufficient to phenocopy these phenomena [33]. However, it is still unclear whether other factors are involved in the coupling of growth and developmental timing.

1.3 Coupling growth with developmental timing – Dilp8

Proper body and organ size regulation is fundamental for animals, because altered body and/or organ sizes can interfere with an organism's fitness and function [33]. In *Drosophila*, as well as other insects, development follows predictable growth and morphogenesis that are achieved by tight and flexible controls in regulation of cell size and number. By this way, animals can achieve proper body and organ size by sensing environmental fluctuations, like nutrition availability [16]. Besides these extrinsic stimuli, intrinsic perturbations also strongly affect body and organ size, for example, abnormal organ and/or tissue growth. The capacity that an organism has to buffer these perturbations (whether through behavioral or physiological alterations) is named developmental stability [8].

To maintain the organism's integrity, rate and duration of growth must be coordinated among organs. For instance, when damage is inflicted upon an organ, the organ has to communicate with the rest of the body to coordinate proper size regulation among all organs [34]. Several examples of these phenomena are known in the structures called imaginal discs.

In *Drosophila* larvae, the imaginal discs represent the precursors of most adult appendages. During larval development, the imaginal discs remain diploid and are the main site of mitotically active cells. These imaginal tissues have a great capacity to

regenerate during the larval growth phase, following chemical (induced by alkylating agents like ethylmethanosulfonate, EMS) or physical (induced by irradiation, for example) damages [11, 35].

Delays in metamorphosis by x-ray irradiation on growing larvae were reported as early as 1927 [36]. Later, similar growth retardation was also verified in temperature induced cell-lethal mutations and in insects undergoing regeneration of differentiated appendages or imaginal discs [37, 38]. This growth retardation allows additional time to repair the damages as to permit cells to adjust their proliferation rates [39]. Although the intact discs are subjected to extra growth time during this extended developmental time, these do not develop into larger organs or appendages, suggesting two possibilities: 1) either the growth continues at a slower rate (or equal, but accompanied by cell death); or 2) the growth stops and these normal tissues arrest their development until the damaged ones complete their growth [39]. Surprisingly, all these studies suggest that the imaginal discs, aside from the PG (see above), also act as regulators of the developmental timing [34].

Since the imaginal discs are remotely localized relative to the endocrine centers controlling developmental timing (most importantly, the PG), the mechanism through which growth abnormalities influence metamorphosis is presumably associated with a communication system dependent on a humoral signal(s) [8, 39]. Furthermore, since the signal is released from the damaged tissues that need to catch up with the rest of the body parts, this is, most likely, an inhibitory signal. By being secreted, this signal prevents metamorphosis until the damaged disc completes its growth program and reaches specific size [4, 11, 34]. Since larvae lacking imaginal discs can undergo metamorphosis without any developmental delay, growth coordination by the inhibitory signal is probably through the derepression of signals from other tissues or through the direct perception of the damage by inhibition of critical weight attainment [33, 40].

In 2012 two groups independently identified a peptide that is involved in the delay mechanism – Dilp8, a member of the Ilp superfamily [42, 43]. Dilp8 is about 150 amino acids, with a signal peptide followed by a cleavage site at the N-terminus, suggesting it can be secreted to the extracellular compartment. Dilp8 signals to the neuroendocrine centers and ultimately, delays pupariation [8, 42, 43]. Colombabi *et al*. identified this peptide by RNAi screen where reducing the expression of this factor eliminated the developmental delay associated with either tumoral or slow-growing imaginal discs. They speculate that a reduction in Dilp8 levels is a pre-requisite to initiate metamorphosis and that the JNK (c-Jun N-terminal Kinase) signaling pathway may be involved in the regulation of Dilp8 levels, since this pathway is activated under tissue stress in a Dilp8 expression-dependent manner [43]. Garelli and colleagues, on the other hand, identified Dilp8 through microarray analysis, by exploring the property of the factor being regulated in a developmental-delay context caused by tumoral eye discs. This approach showed that Dilp8 is necessary for the delay in metamorphosis. To prove its sufficiency, developmental timing was assessed in larvae overexpressing either wild type or mutated Dilp8 ubiquitously. This study revealed a greater delay in the first scenario [42]. Together, these studies showed that Dilp8 is necessary to delay an ecdysone peak during events of abnormal tissue growth. Dilp8 achieves this by down regulating the expression peaks of the genes *disembodied (dib)* and *phantom* (*phm*), which are necessary for the biosynthesis of ecdysone (**Figure 1-4**) [4, 8, 42, 43, 44].

Dilp8 functional and molecular evolution studies allowed further discovery of some particular characteristics of this peptide, such as distinguishing it from other Dilps. Dilp8 sequences and regulatory properties show characteristics of the relaxin subfamily Ilps, such as similarly positioned basic amino acids between the two B-chain cysteins, abrupt end on cystein C6 and a long C-peptide [31]. At the regulatory level, *dilp8* is enriched in adult reproductive tissue, similar to vertebrate relaxins [42].

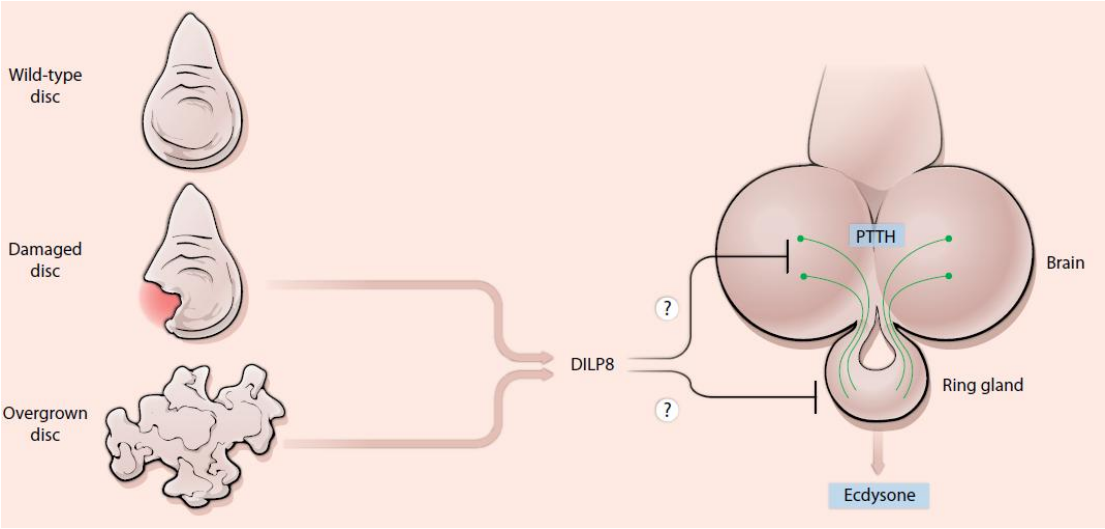


Figure 1-4 – Schematic representation of the mechanism of action of Dilp8. The damaged or overgrown discs secrete Dilp8 that ultimately inhibits ecdysone production, culminating in a developmental delay. Figure from Hariharan, 2012 [8]

Recent work showed that nitric oxide synthase (NOS) activity in the PG contributed to the ability of Dilp8 to inhibit imaginal disc growth, but not to its ability to delay the onset of metamorphosis. NOS activity regulates imaginal discs growth through down regulation of *dib* and *spookier* (*spok*) expression, which leads to decreased ecdysone titers. However, NOS is not required for the Dilp8-dependent delay in development, so it is presently unclear how the reduced ecdysone titers caused by the Dilp8-dependent increase in NOS activity in the PG exclusively affects disc growth and not developmental timing. This is especially intriguing as the length of the third instar larval period is ultimately determined by ecdysone. Despite these possibly contradictory results, this interesting study suggests that Dilp8 could control the timing of the onset of metamorphosis and imaginal disc growth by two separate pathways [45].

Our understanding of how Dilp8 functions and how it evolved would benefic from determining its receptor and tissue(s) directly responding to Dilp8. Recently, Garelli and colleagues have placed an Lgr orphan receptor, Lgr3, in the Dilp8 developmental stability pathway [30]. Lgr3 mutants showed an increased fluctuating asymmetry, a phenotype that is indicative of uncoordinated imaginal disc growth, as seen in Dilp8 deficient animals [30, 42]. Animals lacking Lgr3 did not delay development in response to tissue damage or direct Dilp8-expression, confirming that both genes are acting in the same pathway to control developmental timing in the presence of growth aberrations. Further analysis revealed that that Lgr3 is expressed and required in 2 bilateral Dillp8-responsive central nervous system interneurons called PIL neurons, to transmit the Dilp8 signal from the periphery to the PG. This work revealed a new Dilp8-Lgr3 pathway that is critical to ensure developmental stability. However, it is currently unclear whether Dilp8 acts on Lgr3 directly, because Dilp8 would need to cross the blood-brain-barrier.

Another possibility would be that the Dilp8 signal is transduced through one or more steps before it reaches the PIL neurons and then relayed again to the PG. These results reveal unprecedented complexity in Dilp8 signaling [30]. Consistent with this, this study also reports the biochemical determination of other candidate direct receptors/co-receptors for Dilp8, such as the RTKs, InR and Derailed [30]. The role, if any, of these interacting proteins in the Dilp8-dependent developmental stability pathway remains to be determined.

Besides being involved in growth coordination and developmental timing, Dilp8 could also have other metabolic implications. For instance, larvae overexpressing Dilp8 yield heavier adults which are nevertheless of the same size as wild type flies [42]. While size is determined mostly by the growth of imaginal discs during larval stages, the weight in this case is mostly influenced by growth of larval tissues such as the fat body. These results suggest that Dilp8 might have selective effects depending on whether the tissue is imaginal or not. There are several possible explanations for this. One is that Dilp8 could be posttranslationally modified leading to two different activities. In vertebrates, the cleaved insulin C-peptide, which is secreted equimolarly with insulin, was found to have specific vascular, neural and renal functions, distinct from the mature insulin [46-48]. In invertebrates, including *Drosophila*, the biological roles of any Ilp C-peptides remain unexplored. Dilp8 has a rather long C-peptide, which is likely to be processed and hence secreted equimolarly with it [42]. Importantly, it remains to be established whether all of the previously-described Dilp8-dependent activities, such as the developmental delay, the imaginal disc growth inhibition and the larval weight gain, can be attributed to the mature peptide itself, or whether one or more of these activities can actually be mediated by the long Dilp8 C-peptide. In this thesis, I directly address this problem by testing the ability of the Dilp8 C-peptide to delay development and/or modulate growth.

1.4 *dilp8* evolution

Despite the relevance of the *dilp8* gene in promoting developmental stability, little is known about its evolutionary history. Previous analyses have found *dilp8* homologues in Schizophora flies, but not in Nematocera (mosquitoes) or any other insect or arthropods [42]. These data already indicated that *dilp8* evolved in Diptera after the Schizophora diverged from the Nematocera. A larger survey of Diptera finds clear *dilp8* homologues in many other Muscomorpha, for instance, the Tephritidae fly *Ceratitis capitata* and the Syrphidae *Eristalis pertinax* and *Episyrphus balteatus* (Dr. Darren Obbard, personal communication). *ilp8* is also present in the Asilomorpha clade (Dolichopodidae, Dr. Claude Desplan, personal communication), the Stratiomyomorpha clade (in the black soldier fly, *Hermetia illucens;* Gontijo *et al.* unpublished data) and in the Tabanomorpha clade (*Tabanus bromius,* Dr. D. Obbard, and *T. lineola*, Dr. C. Desplan, personal communications). That *ilp8* is present in the major orders of Brachycera: Muscomorpha Asilomorpha, Tabanomorpha and Stratiomyomorpha, strongly suggests that *dilp8* was present in the last common ancestor to all Brachycera, which should have lived about 180 Million years ago (Figure 1-6). In contrast, clear homologues of *dilp8* are not found in earlier branching Diptera (in all Culicomorpha, mosquitoes), the Bibionomorpha *Mayetiola destructor*, the Psychodidae (*Clogmia albipunctata, Phlebotomus papatasi,* and *Lutzomyia longipalpis*) or any other insect [42] (**Figure 1-6**). Therefore, it is very likely that *ilp8* is a Brachycera innovation. After having arisen, *ilp8* has had an

astonishing fast evolution rate, which is nevertheless notoriously common for Ilps [31]. For instance, the *H. illucens* and *D. melanogaster* Ilp8 sequences share only 24% identity between themselves (Gontijo, unpublished). Nevertheless, within *Drosophila*, Ilp8 is relatively well conserved for an Ilp: there is 57% identify shared between the *D. melanogaster* and *D. grimshawi* Ilp8 proteins. This makes Ilp8 the second most-highly conserved Dilp in *Drosophila* after Ilp7, which has 76% identify, and way above Ilp3, which has 35% identity [Gontijo *et* al, unpublished, 31]. This could suggest that after having quickly diverged following its origination, Dilp8 might have been under more selective constraints in *Drosophila*.

At least three important questions arise from these observations: When and how did *ilp8* become responsive to uncoordinated tissue growth - did the ovarian expression come first? When and how did Ilp8 start delaying development, inhibiting imaginal disc growth and inducing larval weight? Finally, if the other insects do not encode an Ilp8-like peptide in their genomes, how then do they coordinate tissue growth with developmental time?

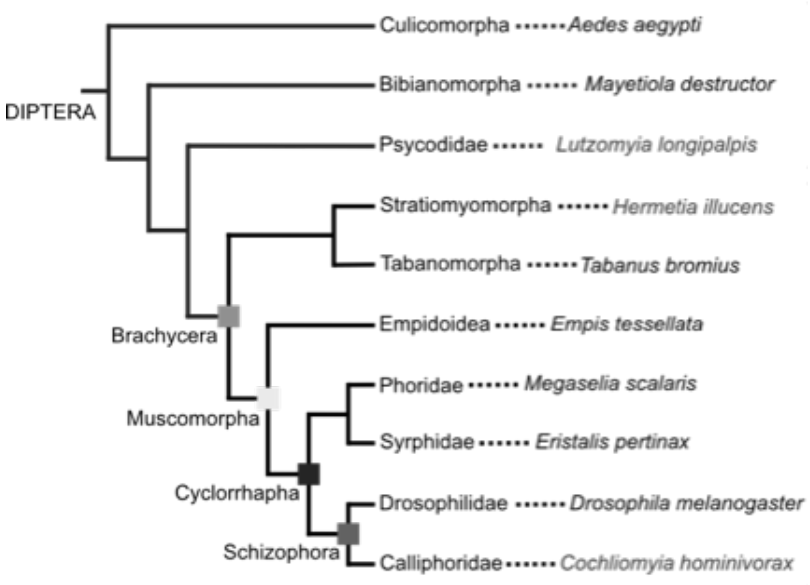


Figure 1-6 – Phylogeny of some dipteran taxa. *H. illucens,* from the Stratiomorpha order is a basal species to *D. melanogaster*.

1.5 Objectives

Dilp8 is a key player in the communication system between abnormally growing peripheral tissues and the neuroendocrine centers that control developmental timing. Understanding the molecular basis of Dilp8 can facilitate the comprehension of processes involved in growth and puberty delays in humans, associated with tissue damages, such as chronic inflammation or infections [11].

Therefore, the main objective of this project is to explore more details of the molecular mechanisms of action and evolution of Dilp8.

At the molecular mechanism level, I aim to use the UAS-Gal4 system to express synthetically-engineered *dilp8* mutant transgenes to clarify if posttranslational processing of Dilp8 into a mature peptide and the C-peptide has any role in its ability to delay development or induce larval weight gain.

In parallel, I aim to further our understanding about the evolution of *dilp8*. Specifically, I aimed at both defining absolutely conserved amino-acids in the Ilp8 protein and test their requirement for the developmental delay and larval weight gain activities; and to gain insight about when the role of Ilp8 as a developmental-stability promoting factor originated. Was it concomitant with the evolution of *ilp8* in the last common ancestor to all Brachycera, or did this role evolve later on in the Brachycera that originated *Drosophila*? To achieve these goals I used an Evo-Devo approach and studied the expression and function of Ilp8 homologues from a brachyceran fly that shared a common ancestor with *Drosophila* about 180 million years ago, *H. illucens*.

Ultimately, this master's project aims to answer both physiological and evolutionary questions associated with this new developmental stability pathway using a combination of molecular and Evo-Devo approaches.

2. Materials and Methods

2.1. Ilp8 homologue sequence alignment

For the molecular evolution study, *Drosophila* species sequences used were obtained from the *Drosophila* 12 Genomes Consortium [55]. *Rhagoletis pomonella* sequences were obtained through the work of Schwarz *et al*, 2009 [56]. The Medfly project (https://www.hgsc.bcm.edu/mediterranean-fruit-fly-genome-project) was used for the *Ceratitis capitata* sequences. *Glossina morsitans* sequences were obtained from a work published by the International *Glossina* Genome Initiative [57]. *Dolichopodidae* and *Tanabus lineola* sequences were generously shared by Claude Desplan, from the University of New York. Darren Obbard from the University of Edinburgh contributed to this work by generously providing *Tabanus bromius* and *Eristallis pertinax* sequences. *H. illucens dilp8* homologues were obtained in the lab, through RNAseq of adult female flies. *H. illucens* genome scaffolds were generously provided by Doris Bachtrog, from the University of California, Berkeley.

All the sequences were aligned by Muscle and curated by hand based on the cysteines, dibasic cleavage sites and the GGY conserved region inside the C-peptide, using MacVector™ software.

2.2. *Drosophila melanogaster* lines and breeding

All fly stocks used in this study were maintained in vials with standard cornmeal-agar medium at 25ºC for experiments or at 18ºC for stock maintenance with controlled humidity conditions. All the stocks used were either generated in the lab or obtained as generous gifts. **Appendix 6-1** represents all the fly stocks used in this project.

2.3. Genomic DNA extraction

Genomic DNA was isolated from five independent head parts of *Hermetia illucens* adults, using the High Pure PCR Template Preparation Kit™ (Roche), according to the manufacturer's instructions.

2.4. RNA extraction

Tissue samples (up to 50 mg) were mixed with 500 µL of TRI Reagent (Zymo Research), macerated using pellet pestles and homogenized and centrifuged at 12000 *g* for 1 min. Then, total RNA was extracted using the Direct-zol™ RNA MiniPrep kit (Zymo Research) according to the manufacturer's instructions. RNA concentration and quality were evaluated using a Thermo Scientific NanoDrop™ 2000 Spectrophotometer. Purified RNA was further treated with Turbo DNase (Ambion), following the manufacturer's instructions. RNA was re-purified and concentrated, using the RNA Clean & Concentrator™-25 kit (Zymo Research), according to the manufacturer's indications. The RNA was re-quantified, as described above and RNA samples were stored at -80ºC until use.

2.5. cDNA synthesis

cDNA was synthesized using the Maxima First Strand cDNA Synthesis Kit for qRT-PCR (ThermoScientific). The 5x Reaction Mix (2 µL), Maxima Enzyme Mix (1µL), total RNA (to be either 500 ng ot 1 µg) and nuclease-free water (to be 10µL of final volume) were gently mixed in an RNase-free tube, centrifuged and incubated for 10 min at room temperature, followed by at least for 30 min at 50ºC and 5 minutes at 85ºC. The samples were diluted to 1-5 ng/µL for further analyses. A reverse-transcriptase negative (RT -) control reaction was performed in parallel for all cDNA reactions to confirm any genomic DNA contamination.

2.6. Primer design

All primers used were designed using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primer specificity was also confirmed on the same program. All primers were synthesized by Sigma, Portugal. Primers used for this study are represented in **Appendix 6-2** and **3.**

2.7. Polymerase Chain Reaction (PCR)

DNA fragments were amplified using a T100 Thermal Cycler™ (BioRad). A commercial Enzyme Mix (NZYTech) was used for PCR reactions (**Appendix 6-4).** The PCR conditions were optimized individually for each aim and the size of the product (extension time – 1kb/min – and the annealing temperature) – **Figure 2-1**.

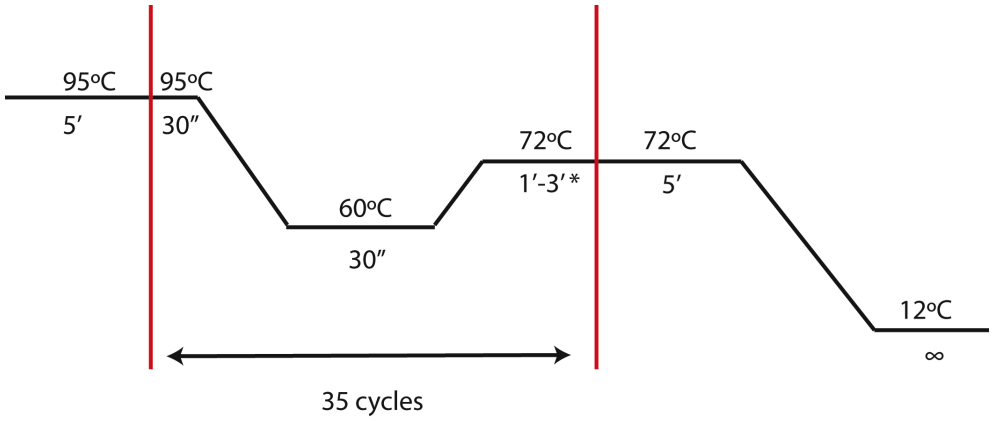


Figure 2-1 – PCR conditions for amplification of interest genes. *1 min per 1kb

2.8. Electrophoresis on agarose gel

Agarose gel was prepared at the concentration of 1.2% by dissolving UltraPure™ Agarose (Life Technologies) in 1x TAE buffer with 2 µL of Green Safe Premium (NZYTech; 2 µL/50 mL). NZYDNA Ladder III (NZYTech) molecular marker was used to estimate DNA size. Electrophoresis was performed in 1x TAE buffer, using a PowerPac 300 system (BioRad). The gel was visualized in Molecular Image Chemidoc™XRS+ (BioRad).

2.9. DNA purification and sequencing

PCR products were confirmed by gel electrophoresis. These fragments were purified using the NZYGel Pure Kit™ (NZYTech), according to manufacturers instructions. Sanger sequencing was performed by StabVida, Portugal. DNA sequences were analyzed using various software packages, such as MacVector, UCSC Genome Browser, BLAST, SnapGene, among others.

2.10. *Hermetia illucens* qPCR analysis

qPCR was performed using the FastStart Essential DNA Green Master dye and polymerase (Roche) and Lightcycler® 96 (Roche). The final volume used was set to be 10µL, consisting of 5 µL of Master Mix, 2 µL of cDNA sample and 3 µL of the specific primer pairs (1µM/µL; see **Figure 2-2)**.

All the qPCR results were normalized to *rp49* (also known as *rpl32*), a housekeeping gene that is expressed throughout development or *ef2*, an elongation factor that is expressed in every tissue. For negative controls, both H_2O and RT- solutions were used as controls. The results were analyzed using the Lightcycler® 96

manufacturer's software and Microsoft Office Excel® 2011. All the Cq values obtained for each primer pair were normalized to those of *rp49* and/or *ef2*. In order to calculate the copy number of any target product relative to the housekeeping gene in percentage, the following formula was applied: **Percentage relative to *rp49*= $2^{-(\Delta Cq)}$ x100**, where ΔCq represents the difference between the Cq of the target and the arithmetic mean of Cq obtained from the three technical replicates of the positive control. The geometric mean ± standard deviations of the experimental replicates (usually n=3) for each condition were presented in the graphs. Prism 6 (GraphPad Prism) was used for the graphical and statistical analyses of the data. A Kruskal-Wallis statistics with a Dunn's multiple comparisons test was used to compare the results.

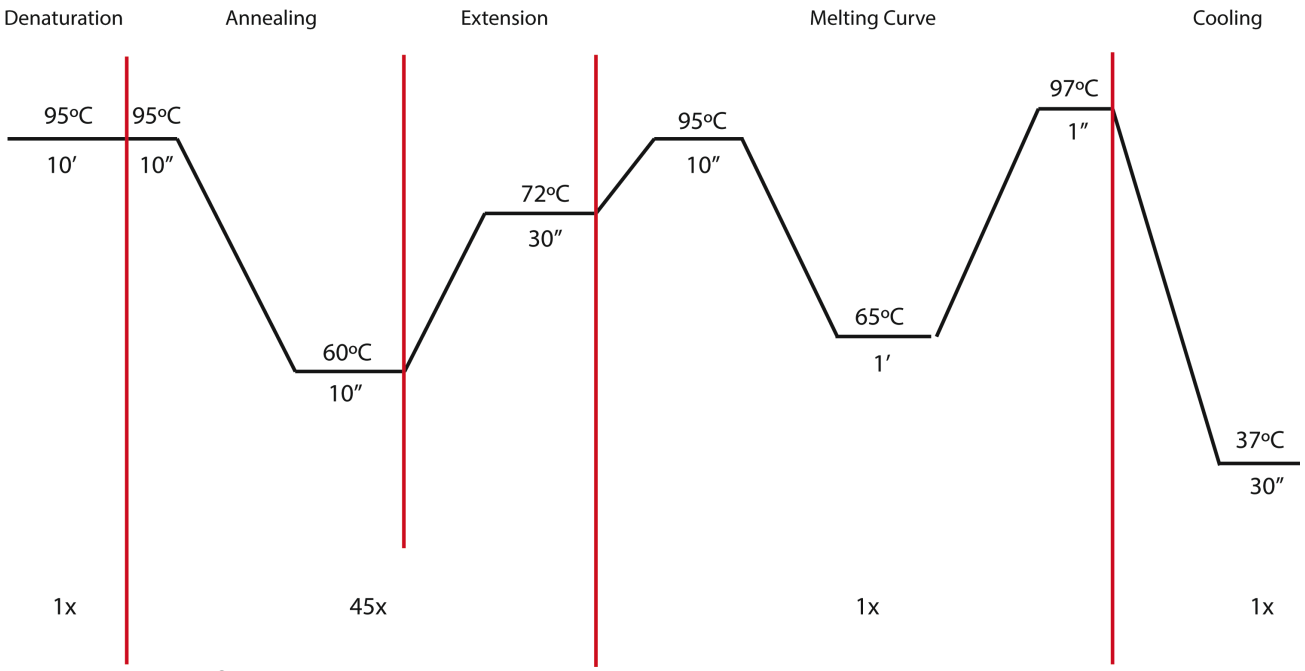


Figure 2-2 - qPCR conditions.

2.11. Plasmid construct and generation of transgenic flies

To generate transgenic lines, two different strategies were chosen: 1) P-element mediated insertion and 2) *attB-attP* site and phiC31 integrase mediated insertion. Using P-element mediated transgenesis, two different mutant overexpression constructs for the C-peptide region of *prodilp8* gene were created, using a P-element-mediated overexpression plasmid, pUASp. A synthesized plasmid, proDilp8/pUC57 was used as a template and two different mutations were induced: 1) a point mutation on L102 to alanine (L102A), and 2) deleting two amino acids G89 and G90 (Δ89, 90). These two mutations were induced by a standard site-directed mutagenesis method with minor modifications [41], using Pfu Polymerase (Fermentas) and two primer pairs (**Appendix 6-5**). The PCR conditions used in this protocol are represented in **Figure 2-3**.

Then, the PCR products were treated with the 4-base cutter restriction enzyme, *DpnI*, which recognizes and digests only methylated target DNA sequence in order to remove all wild type pro-dilp8/pUC57 plasmid used for PCR templates. After PCR products were confirmed on agarose gel, these PCR products were used for transformation into DH5α competent cells. After sequencing to confirm the presence of mutations, the plasmids were digested by *XbaI* and *KpnI* and inserted into pUASp plasmid. Then, a cocktail consisting of two different plasmids: a P-element containing the mutant form of the peptide and a helper plasmid, Δ2-3 (a gift from Dr. Ribeiro, CCU, Portugal) that produces transposase but cannot be inserted into the host genome was injected into w[1118] embryos (at most, 30 minutes after oviposition).

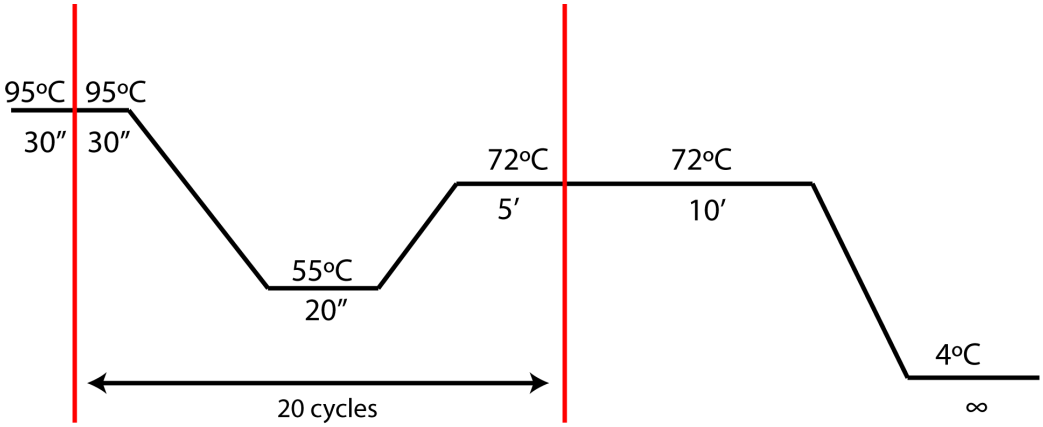


Figure 2-3 – PCR conditions used in the site-directed mutagenesis protocol.

Using the second method, *attB-attP* site and phiC31 integrase mediated insertion, *dilp8* C-peptide constructs with or without OLLAS-tag were created, using a phiC31-*attP-attB*-mediated overexpression plasmid, pUAST *attB* (a gift from Dr. Koyama, IGC, Portugal). For construct making, two different synthesized plasmids were used: Dilp8 C-peptide/pUC57 and OLLAS-Dilp8 C-peptide/pUC57 (Genscript, USA). The plasmids were digested by *XbaI* and *EcoRI*, and inserted into pUAST *attB* plasmid. These plasmids were further injected into the embryos (30 min old, maximum) of an *attP* line, y w p{Y+.nos-int.NLS}; p{CaryP y+}attP2 (a gift from Dr. Manoel, IGC, Portugal).

For both methods, the embryos were previously aligned on a coverslip (approximately 100 per coverslip) and covered with halocarbon oil mixture (87.5% HC-700, 12.5% series 27, Sigma) to diminish the damage caused by the process. The injection was carried out using a micro-injector. The coverslips with the injected embryos were transferred into standard food vials and kept between 22-25ºC.

For P-element insertion, all the newly eclosed flies, after injection, were individually crossed with w[1118] flies. The F1 generation (considering the eclosed embryos as the F0 generation) was observed daily to isolate the adults with colored eyes. These were

then crossed with a double balancer line (w;*If/Cyo;MKRS/TM6B Tb*) and the inserted chromosome was identified.

For *attP-attB* mediated insertion, the newly eclosed adults were collected and were individually crossed with w[1118] flies. The colored eye flies that resulted from this cross were then crossed with the III chromosome balancer line (w;*MRKS/TM6B Tb*). The flies not carrying the integrase gene on the X chromosome were selected for further analyses.

2.12. Functional studies - developmental timing

For the developmental delay assays, the appropriate flies were set in laying pots with apple juice agar plates containing a small amount of yeast-sucrose past for pre-embryo collection about 24 h before embryo collections. The plates were exchanged for new plates every 3 h. Embryos were collected three times a day. Groups of ten second instar larvae (48 hours after egg laying) were transferred into at least three vials (and up to seven) containing 5 mL of standard *Drosophila* food. The number of newly pupariated animals was counted every 6 or 12 h (usually at 9 am, 3 pm and 9 pm). Three independent transgenic lines were analyzed for each genotype to evaluate the effect of insertion sites.

Since PTTH secretion is photoperiod sensitive [5, 42, 43], this assay was carried out either in a normal photoperiod condition or a constant light condition to avoid any photoperiod-sensitive effects.

Microsoft Office Excel® 2011 was used to generate the database, which was then uploaded into Python™ (https://www.python.org/) for the data conversion. After converted, R (http://www.r-project.org/) was used for graphics and statistical analyses. The duration from egg collections to pupariation across different genotypes were compared using a Game-Howell post-hoc statistics.

2.13. Weight study

Sets of 5 *arm-Gal4* females and 3 males of the different pUASp-proDilp8 lines were raised together in a vial containing standard *Drosophila* food and were flipped into new vials daily to control larval density. After approximately 8 d, the pharate adults (between 2-10 h before eclosion) were collected. Male and female pupae were separated (based on the presence or absence of sex combs) and cleaned individually with distilled water and a paintbrush. The pupae were left to dry for at least 15 min before being weighed individually using an ultra-microbalance (Sartorius, SE2). Three independent lines were weighed for each genotype, except for the integrase-mediated insertion line. The

presented results consist on the average of these three independent weighing. The statistical and graphical analyses were carried out using Prism 6 (GraphPad Prism). A Kruskal-Wallis statistics with a Dunn's multiple comparisons test was used to compare the results.

2.14. Western blot

Second instar larvae were used for protein extraction. The whole larvae were homogenized in NB Buffer (150 mM NaCl, 50 mM Tris-HCl pH 7.5, 2 mM EDTA, 0.1% NP-40, 1 mM DDT, 10 mM NaF) with complete Protease inhibitor (Roche) and PhosSTOP phosphatase inhibitor (Roche). After equal amount of 2x Laemmli Buffer was added, the samples were boiled for 5 min followed by centrifugation at full speed for 5 min. Thirty µL of each sample were loaded on 10–20% Mini-PROTEAN® Tris-Tricine Gel (BioRad) and run for approximately 4 h in Running Buffer (25 mM Tris base, 190 mM glycine, 0.1% SDS). Proteins were transferred onto a nitrocellulose 0.2 µm pore membrane (BioRad) using a wet transference method with Blotting Buffer (25 mM Tris base, 190 mM Glycine, and 20% Methanol) at 20mA for 25 min. After blocking with 5% skim milk/PBS with 0.2% Tween 20 the membrane was incubated with a monoclonal ANTI-FLAG® M2, Clone M2 antibody (1:1000, Sigma-Aldrich) for overnight at 4ºC. Then the membrane was incubated with an HRP conjugated anti-Mouse IgG (1:2000, Jackson ImmunoResearch) solution for 1 h at room temperature. Protein detection was performed using Pierce™ ECL Plus Western Blotting (Life Technologies) Substrate, according to the manufacturer's instructions. The signal was exposed to Amersham Hyperfilm™ ECL™ (GE).

2.15. EMS injection into *Hermetia illucens* larvae

In order to determine the effect of the genotoxic agent ethylmethanosulfonate (EMS) in *H. illucens*, 4th instar larvae were obtained from Bioflytech (http://bioflytech.com/en). These were injected in the fourth segment after the head, with either EMS (Sigma-Aldrich) or PBS, using an insulin syringe. According to the National Institute of Environmental Health Sciences (http://tools.niehs.nih.gov/cebs3/ntpviews/index.cfm?action=testarticle.toxicity&cas_number=62-50-0), a dosage of 50-400 mg of EMS/kg of body weight in mice was used as a reference. As an *H. illucens* 4th instar larva weighs approximately 100mg, the stock solution (120 g/mL) was diluted to 1:1000 in PBS and 40 µL were injected (480 mg/larva). The control larvae were treated similarly with PBS. After a 24 h recovery period in *Drosophila* standard food, the larvae that were moving were collected for RNA extraction.

2.16. *Hermetia illucens* RNA extraction for RNAseq

Exclusively for RNAseq analyses, a different RNA was extracted differently, in order to increase the RNA yield and quality. A total of 18 injected larvae (9 for EMS and 9 for PBS) were collected. By combining three larvae from the same treatment, three biologically independent RNA samples per treatment were extracted. Each larva was cut into small pieces for homogenization and the corresponding tissue sample was divided into three different micro-tubes with 1 mL of TRI Reagent (Zymo Research). The samples were then macerated using pellet pestles and homogenized and centrifuged at 12000 *g* for 1 min. The supernatant corresponding to the three larvae was then mixed in a 15 mL tube. From this, the volume corresponding to the maximum capacity of the column (100 mg) was mixed with 0.5 volumes of absolute ethanol and the mixture was loaded into a High Pure Filter Tube combined with a Collection Tube and the RNA was extracted using the High Pure Tissue Kit (Roche), according to the manufacturer's instructions. The remaining samples in TRI Reagent were stored at -80ºC.

After the extraction, evaluation of RNA concentration and quality, second round of DNase treatment, RNA re-purification, re-evaluation of RNA concentrations were performed as described above. Then, the samples were stored at -80ºC until use.

RNAseq analyses were performed by Dr. Tatiana Torres' group. To dry the samples for shipping, 35 µL of each sample (corresponding to at least 4 µg of RNA) were dried, using the RNAStable/RNAStable LD® kit (Biomatrica), according to the manufacturer's instructions.

3. Results

3.1. Identification of conserved Dilp8 amino acid residues

To understand the relevance of Dilp8 posttranslational processing and details about its evolutionary history better, it is important to look for Dilp8 homologues in other species. Dilp8 homologues are found in the 12 *Drosophila* species, as well as in the major order of Brachycera: Muscomorpha, Asilomorpha, Tabanomorpha and Stratiomorpha [42, Gontijo *et al*, unpublished]. In contrast, Dilp8 homologues were not found in the nematocerans. By aligning these Dilp8 homologues, 16 highly conserved amino acid residues were found, 12 of which were absolutely conserved and 4 residues that tolerated other very similar amino acid types (basic, aromatic, etc). These results indicate that some regions of this peptide have been subjected to a severe functional constraint during the evolution of Brachycera (**Figure 3-1**, shaded areas). If we exclude the 6 cysteines that determine the appurtenance of Dilp8 to the Ilp family, one of which has been shown to be critical for Dilp8 activity in *D. melanogaster* [42], there are 6 strictly conserved amino acids: M34, A38, A41, F46, G90 and L102. The first four of these are in the B-chain, and the latter two, G90 and L102, are surprisingly in the C-peptide. Two additional highly conserved positions are evident: a basic residue at position K82 and an aromatic residue at position Y91 in the C-peptide. The presence of some conserved residues in the C-peptide suggested these sites have been subjected to selective pressures during Brachycera evolution. The importance of some of these conserved amino acids for Dilp8 function as a developmental stability factor was explored next.

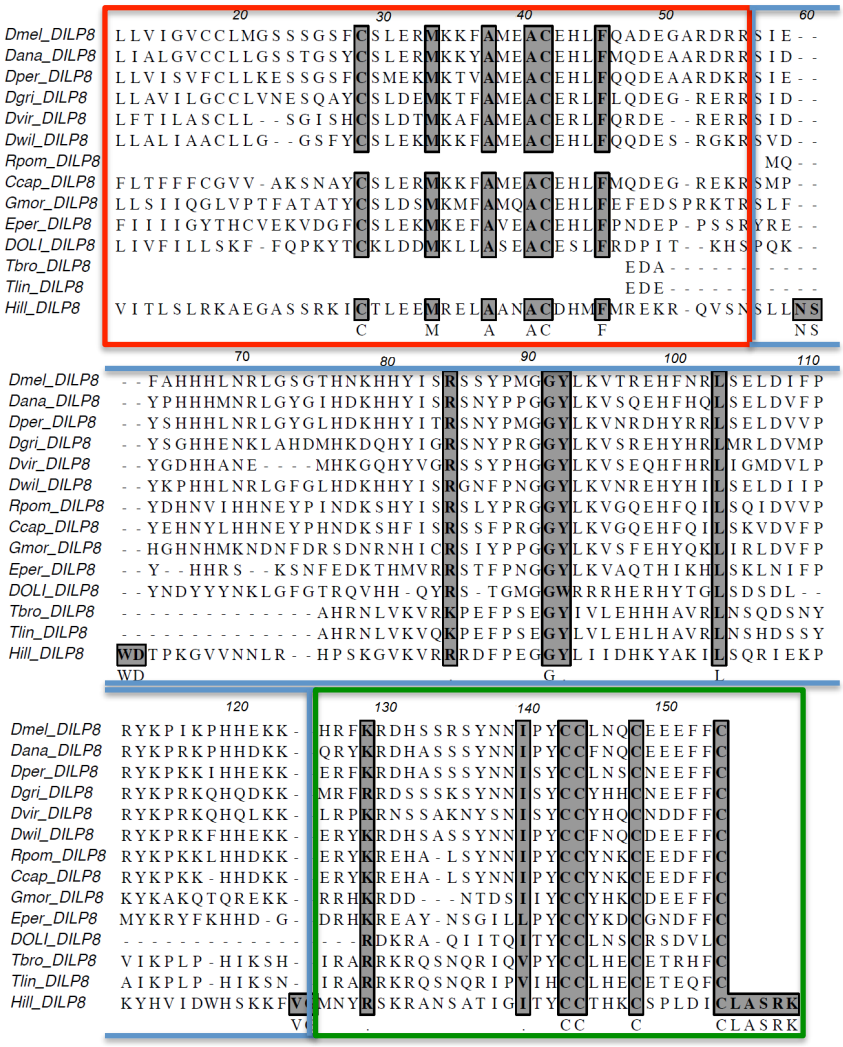


Figure 3-1 – Alignment of Dilp8 homologues revealed a limited amount of absolutely conserved amino acids. Dilp8 homologues were found in several *Drosophila* species, as well as other Brachycera members. The alignment of the homologue sequences revealed the existence of a handful of amino acids that are highly conserved throughout evolution (shaded areas). The red, blue and green boxes indicate the amino acid sequences that correspond to the B-chain, C-peptide and A-chain, respectively. *Dmel, Drosophila melanogaster; Dana, D. ananassae; Dper, D. persimilis; Dgri, D. grimshawi; Dvir, D. virilis; D. wil, D. willistoni*; Rpom, *Rhagoletis pomonella; Ccap, Ceratitis capitata; Gmor, Glossina morsitans; Eper, Eristalis pertinax; DOLI, Dolichopodidae; Tbro, Tabanus bromius; Tlin, T. lineola; Hill, Hermetia illucens*. See the Materials and Methods section for more details.

3.2. Ubiquitous Dilp8 expression causes a developmental delay

To start testing which of the conserved Dilp8 amino acid residues described above are required to delay development, we made a series of *dilp8* mutant constructs using a new *pUASp-preprodilp8::3xFLAG* template and made transgenic flies. To confirm the transgenes are working as expected, the unmodified *pUASp-preprodilp8::3xFLAG* (hereafter, *UAS-prodilp8*) was ubiquitously expressed under the control of the ubiquitous Gal4 driver, *armadillo*, as a control. Consistent with previous results [30, 42], larvae overexpressing the new *dilp8* transgene pupariated ~22 h later than *arm-Gal4* controls (**Figure 3-2**). This delay was observed in all three independent insertion lines.

The data depicted in Figures 3-2, 3-3, 3-4, 3-7, 3-8 and 3-11 were all performed in the same chronological series of experiments under constant darkness conditions. As the same *arm-Gal4/+* controls were always included and were the same for the different conditions, we pooled the *arm-Gal4* into one group for statistical analyses and represent it in these graphs. The same was done for the next large set of experiments depicted in Figures 3-5, 3-6 and 3-9, which all share the same *arm-Gal4/+* and *arm-Gal4/UAS-prodilp8* controls in the statistical analyses and box plot graphics. This last set of experiments was performed under constant light conditions.

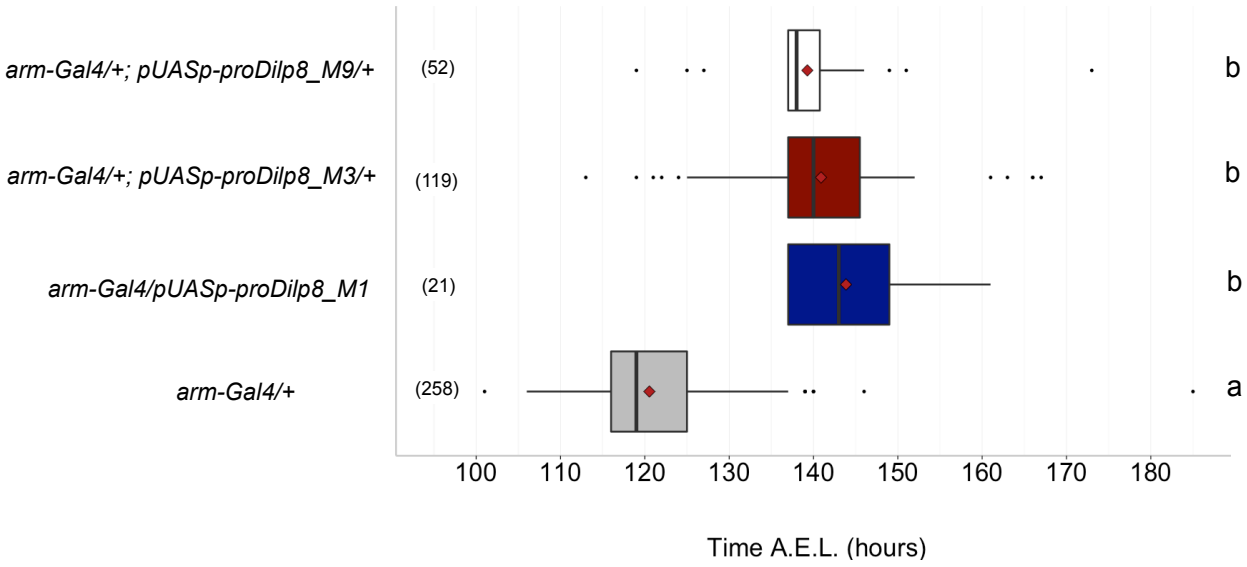


Figure 3-2 – Ubiquitous expression of proDilp8 induces a developmental delay. Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *UAS-prodilp8* under the control of *arm-Gal4.* M1, M3 and M9 correspond to three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* served as control. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

3.3. Dilp8 requires a conserved methionine (M34) in its B-chain to delay development

Here, a point mutation was introduced in the wild-type *pUASp-proDilp8* template to induce an M34A change in the B-chain (white area in **Figure 3-3 (A)**). Ubiquitous expression of two independent *UAS-prodilp8M34A* transgene insertions under the control of *arm-Gal4* showed no developmental delay (**Figure 3-3 (B)**). One of the insertions (*pUASp-prodilp8M34A_M10/arm-Gal4*) showed a small, yet statistically significant delay. However, this particular sample had a very reduced sample size (n=6), so this result should be considered with caution. Overall, these results suggest that the conserved methionine in the B-chain is important for the action of Dilp8 in regulating of developmental timing.

A

B

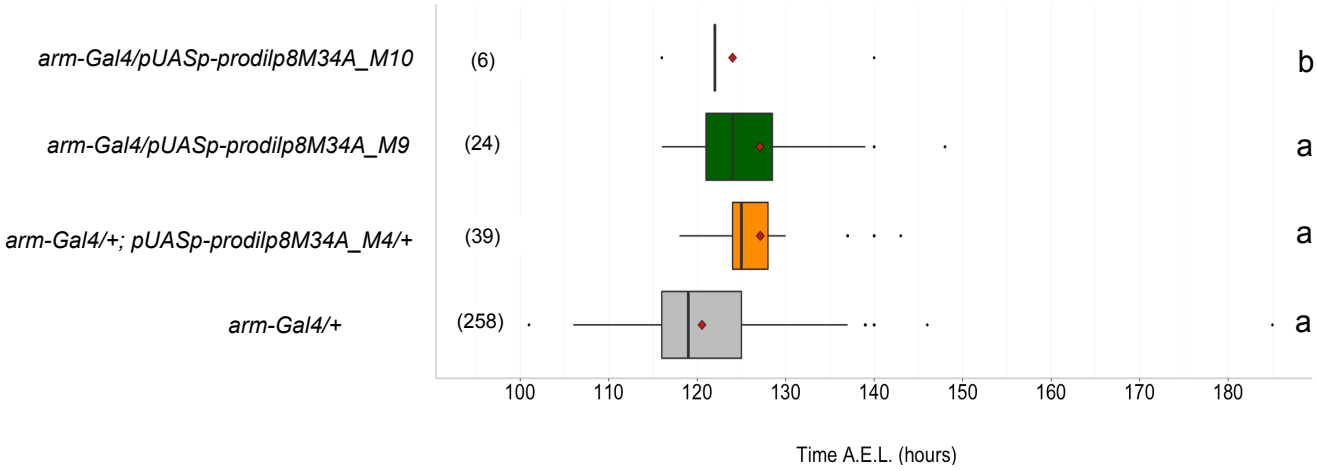
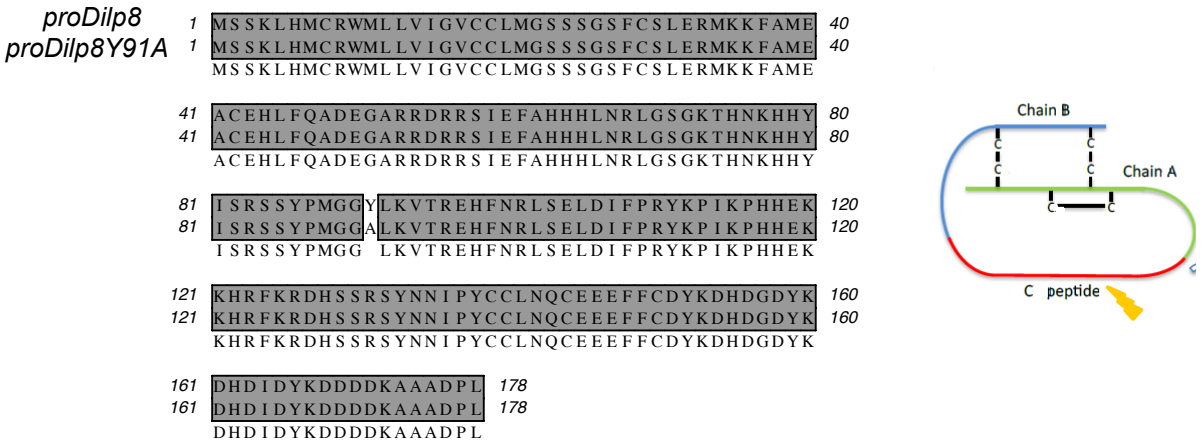


Figure 3-3 – The conserved M34 is required for the Dilp8-dependent developmental delay activity. **(A)** Dilp8 protein alignment between wild-type proDilp8 sequence (upper line) and proDilp8M34A (bottom line). **(B)** Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASp-prodilp8M34A* under the control of *arm-Gal4.* M4, M9 and M10 correspond to three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* served as control. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

3.4. The conserved tyrosine (Y91) or leucine (L102A) residues in the Dilp8 C-peptide are dispensable for the developmental delay function

Here, we introduced a point mutation in the wild-type *UAS-prodilp8* template to induce either a Y91A (white area in **Figure 3-4 (A)**) or L102A change (white area in **Figure 3-5 (A)**) in the Dilp8 C-peptide. Ubiquitous expression of each of the three independent transgene insertions encoding proDilp8Y90A or proDilp8L102A tested under the control of *arm-Gal4* induced a developmental delay (**Figure 3-4 (B)** and **Figure 3-5 (B)**, respectively), similar to that of animals expressing wild-type proDilp8 (**Figure 3-2 (B)**).

A



B

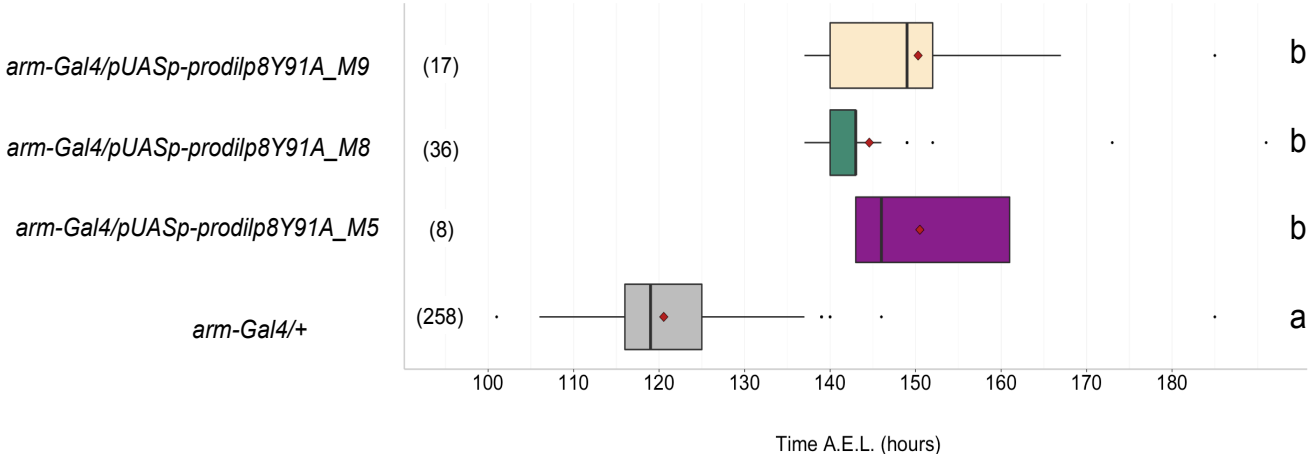


Figure 3-4 – A conserved tyrosine (Y91) in the Dilp8 C-peptide is dispensable for the developmental delay function. (A) Dilp8 amino acid alignment of wild-type proDilp8 sequence (upper line) and proDilp8Y91A (bottom line). **(B) B**ox plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASp-prodilp8Y91A* under the control of *arm-Gal4.* M5, M8 and M9 represent three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* served as control. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

These results demonstrate that two highly conserved amino acids in the Dilp8 C-peptide, namely the aromatic Y residue at position 91 and the absolutely conserved L102 are dispensable for the developmental delay function. However, this does not exclude a possibility that these amino acids are critical for the Dilp8 functions in organ growth control. This possibility was further analyzed in a different assay.

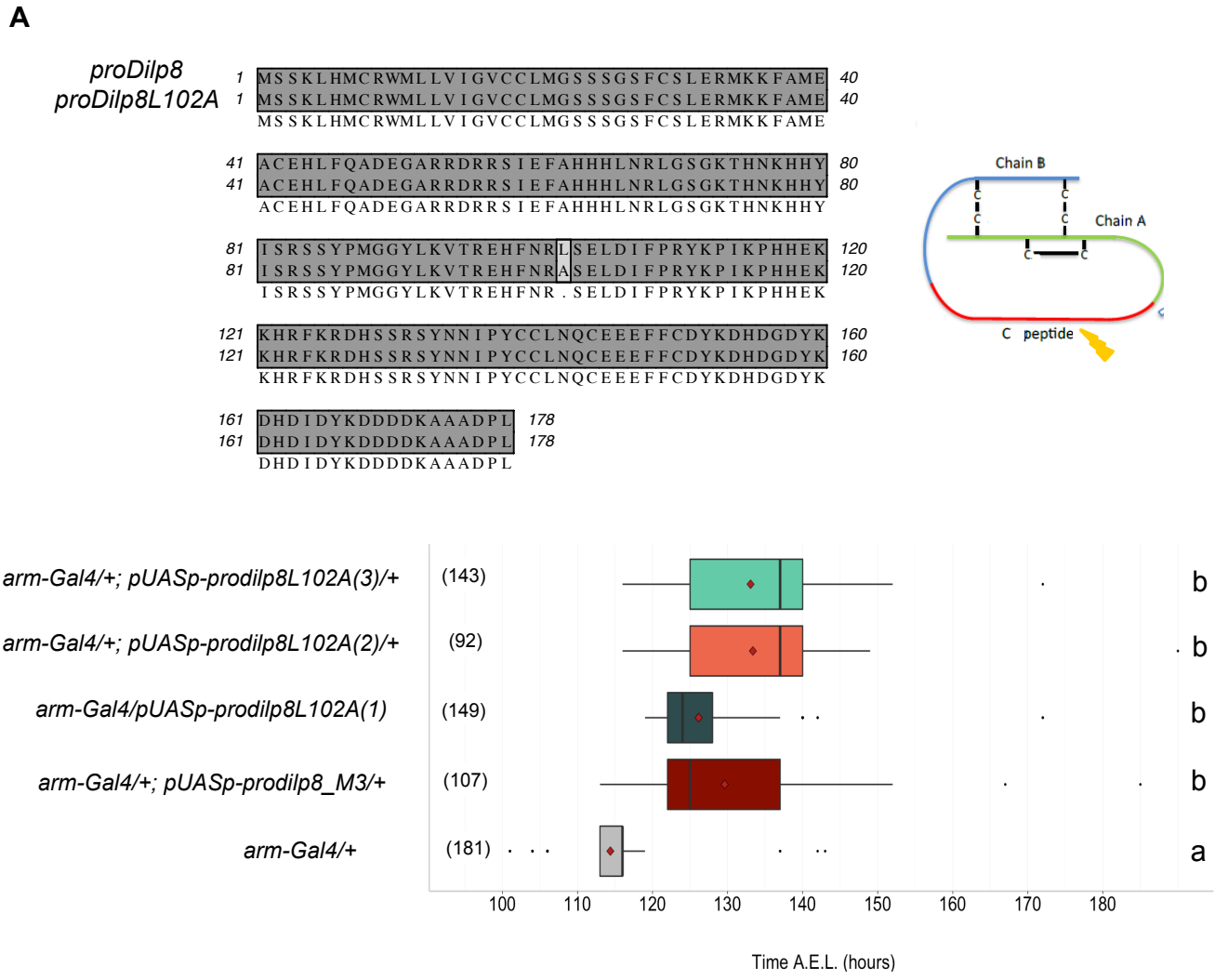


Figure 3-5 – A conserved leucine (L102) in the Dilp8 C-peptide is dispensable for the developmental delay function. (A) Dilp8 amino acid alignment of wild-type proDilp8 sequence (upper line) proDilp8L102A (lower line). **(B)** Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASp-prodilp8L102A* under the control of *arm-Gal4*. The numbers (1), (2) and (3) represent three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* or *UAS-prodilp8* (*pUASp-prodilp8_M3*) served as negative and positive controls, respectively. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

3.5. A diglycine site in the Dilp8 C-peptide (G89G90) is required for the Dilp8-dependent developmental delay activity

The Dilp8 C-peptide has a well-conserved sequence “GGY” (ilp8 motif G/E-G-Y/W) approximately in the middle of the C-peptide (**Figure 3-1**). As substituting the Y for A does not affect developmental timing (**Figure 3-5 (B)**), we also expected that the precise deletion of the diglycine residues (**Figure 3-6 (A)**) would also not affect the ability of Dilp8 to delay development. Surprisingly, ubiquitous expression of three independent *UAS-prodilp8G89G90* transgene insertions under the control of *arm-Gal4* showed no developmental delay (**Figure 3-6 (B)**). On the contrary, the larvae pupariated significantly earlier than controls, even though the difference was very small, <5 h. These results demonstrate that the contrary to the Y91 and the L102 residues, the diglycine residues are critical for the Dilp8-dependent developmental delay activity. The shorter length of the C-peptide per se in prodilp8G89G90 could somehow negatively-affect Dilp8 activity, but as there is considerable variation of the Ilp8 C-peptide length between different Brachycera species, Ilp8 functions in those species need to be unraveled before this can be confirmed. Alternatively, the diglycine motif could be recognized and/or modified to promote either proper folding of proDilp8 and/or the proper processing of the C-peptide.

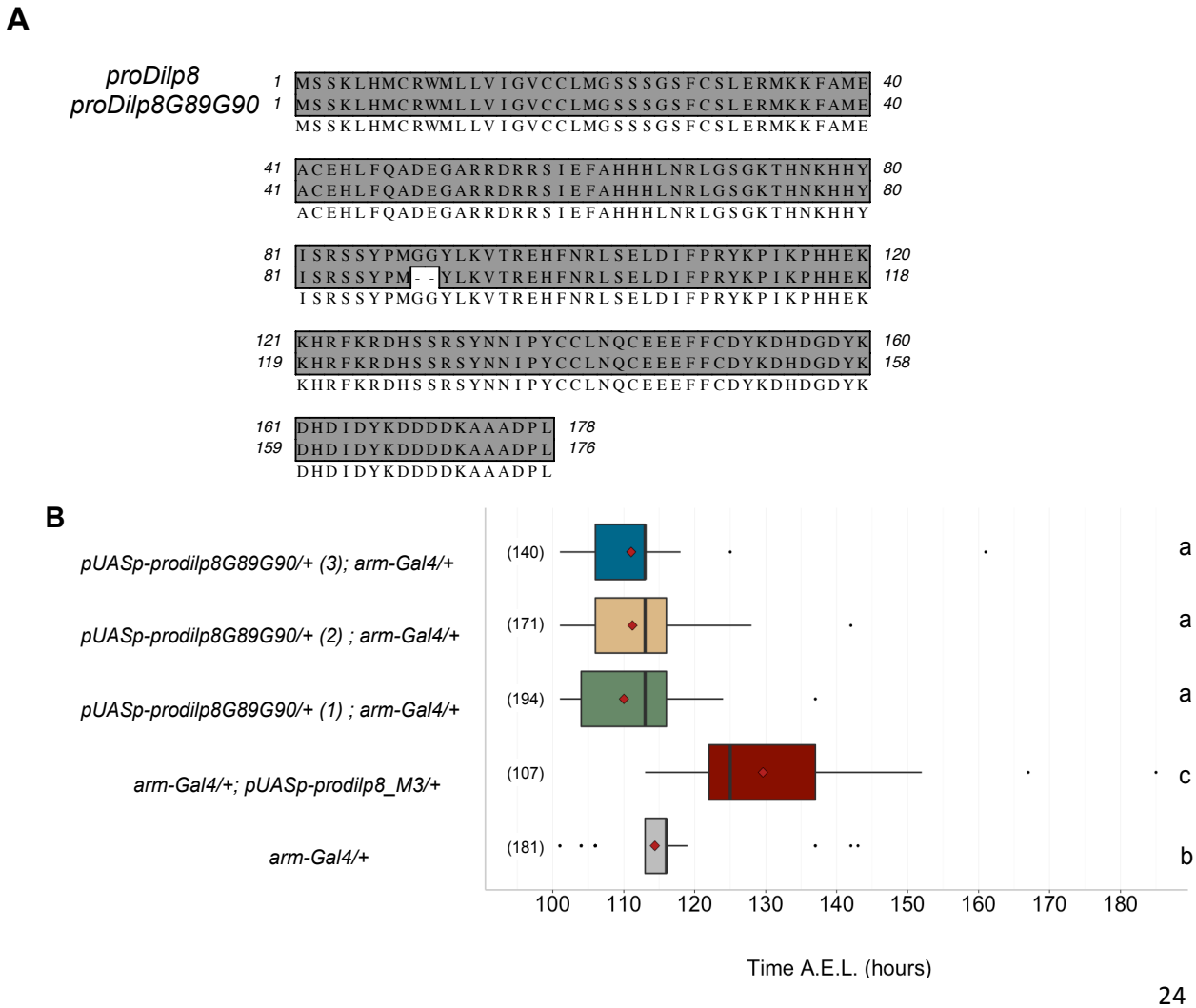


Figure 3-6 – A diglycine site in the Dilp8 C-peptide (G89G90) is required for the Dilp8-dependent developmental delay activity. (A) Dilp8 amino acid alignment of wild-type proDilp8 sequence (upper line) and proDilp8G89G90 (lower line). **(B)** Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASp-proDilp8G89G90* under the control of *arm-Gal4.* The numbers (1), (2) and (3) represent three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* or *UAS-prodip8* (*pUASp-prodilp8_M3*) served as negative and positive controls, respectively. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

3.6. Elimination of putative Dilp8 C-peptide processing sites abrogates the Dilp8-dependent developmental delay activity

In vertebrates, the proinsulin C-peptide is excided by a convertase enzyme that recognizes and cleaves the proinsulin at two specific dibasic sites. To test whether proDilp8 processing is required for the developmental delay activity, we produced a *prodilp8* mutant, *pro*dilp8*, which encodes a proDilp8 version where all the putative dibasic sites are disrupted by substituting the basic amino acids for alanines, as depicted in **Figure 3-6 (A)**. When compared to the control, all the three *pUASp-pro*Dilp8* lines tested did not show any developmental delay when crossed to *arm-Gal4*. These results strongly indicate that Dilp8 processing is critical for the Dilp8-dependent developmental delay activity.

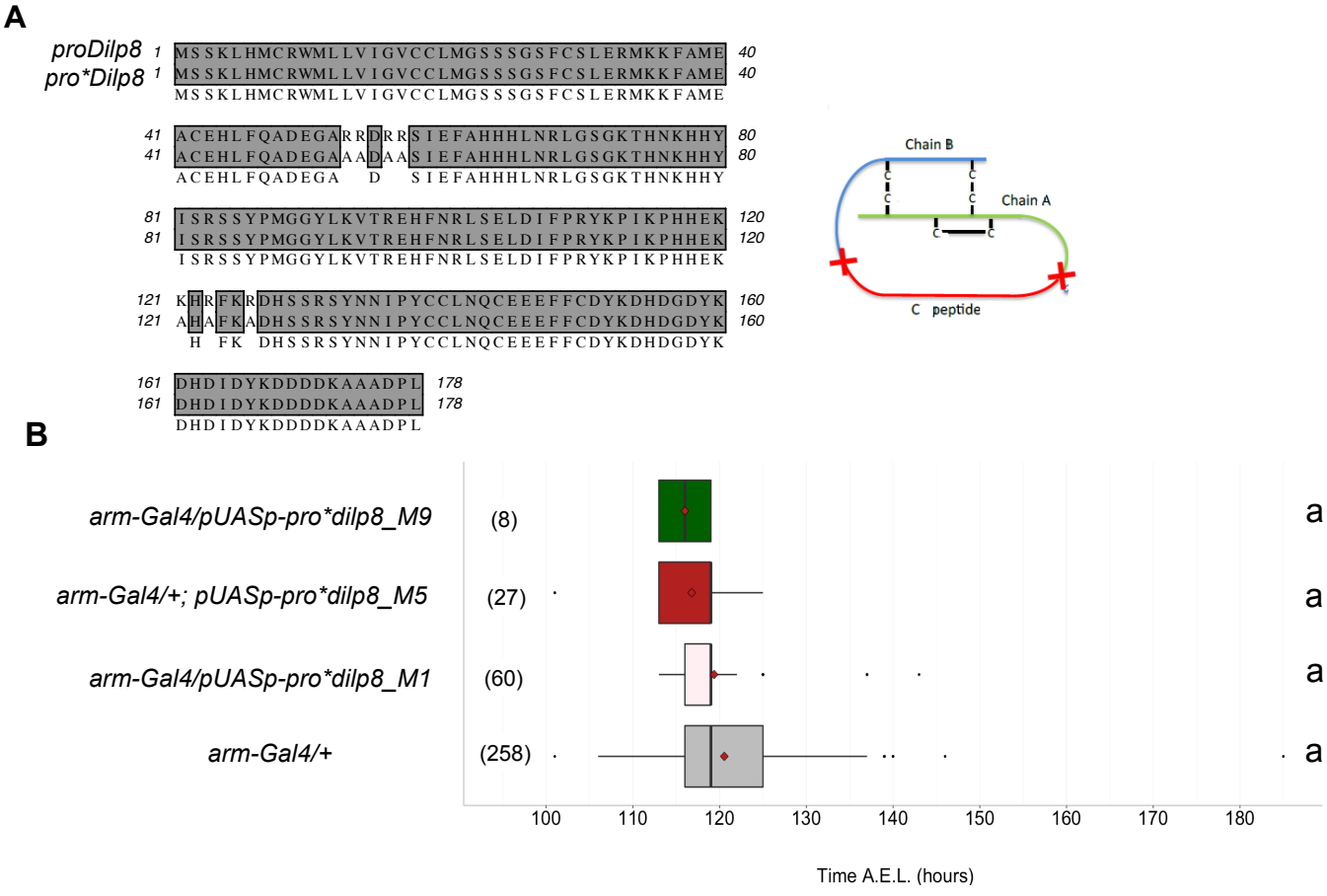
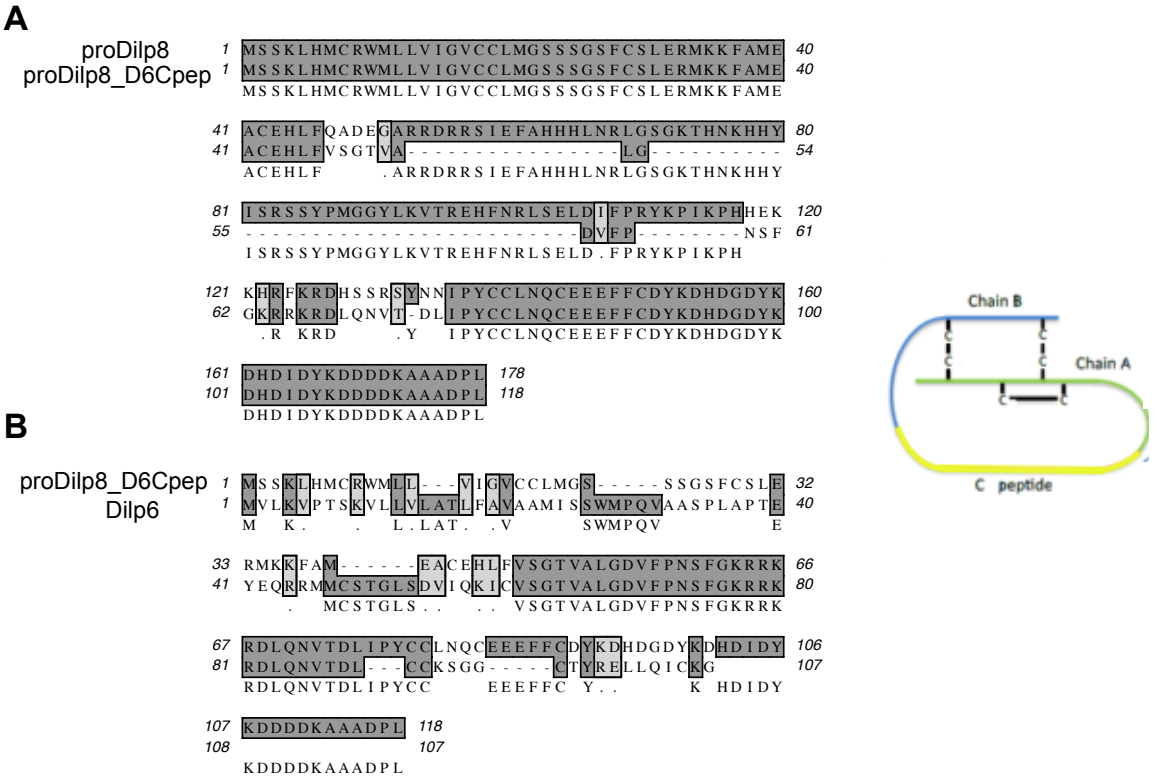


Figure 3-7– Elimination of putative Dilp8 C-peptide processing sites abrogates the Dilp8-dependent developmental delay activity. (A) Dilp8 amino acid alignment of wild-type proDilp8 sequence (upper line) and pro*Dilp8 (lower line). **(B)** Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASp-pro*Dilp8* under the control of *arm-Gal4.* M1, M5 and M9 represent three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* served as control. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

While the C-peptide from vertebrate insulin and relaxin pre-hormones are excised during their maturation process, the equivalent of the C-peptide fragment in the IGF-like "pre"-hormones are retained in the mature form. These peptides either are not cleaved at all or possess only one cleavage site, which does not lead to excision of the C-peptide [19, 21]. Therefore, for an IGF-like protein, the presence of its C-peptide does not interfere with its ability to signal to its cognate receptor. With this in mind, we swapped the Dilp8 C-peptide for the C-peptide of the *Drosophila* IGF-like Ilp, Dilp6, making the chimeric construct *UAS-proDilp8_D6Cpep* (**Figure 3-7 (A)** and **(B)**). If overexpression of this construct would delay development, we could conclude that the original Dilp8 C-peptide is dispensable for Dilp8 activity. If there was no delay, this would indicate either that the Dilp8 C-peptide signals independently from Dilp8 or that Dilp8 cannot tolerate an attached C-peptide, even a small one like that of Dilp6. Results showed that all of the three *UAS-proDilp8_D6Cpep* lines tested did not significantly delay the timing of pupariation when crossed to *arm-Gal4* (**Figure 3-7 (C)**). These findings suggest that proper posttranslational processing of Dilp8 into a mature Dilp8 peptide is necessary for the Dilp8-dependent developmental delay activity.



C

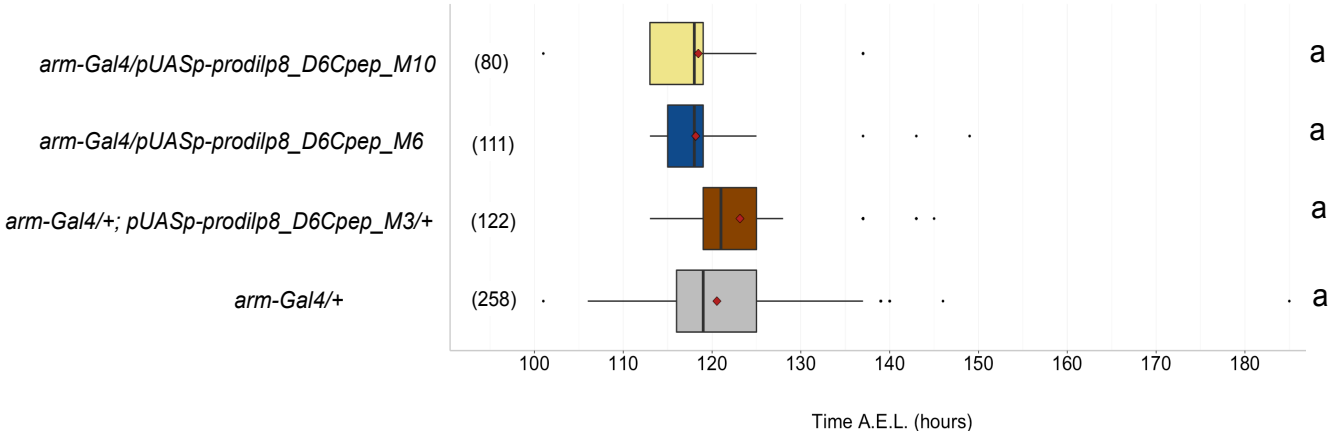


Figure 3-8 – Ubiquitous expression of a chimeric Dilp8 carrying the small Dilp6 C-peptide does not induce a developmental delay. (A) Protein alignment of wild-type proDilp8 (upper line) and proDilp8_D6Cpep (lower line). **(B)** Protein alignment of Dilp6 (upper line) and proDilp8_D6Cpep shows (lower line). **(C)** Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASp-pro*Dilp8* under the control of *arm-Gal4.* M3, M6 and M10 correspond to three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* served as control. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically significantly indistinguishable at alpha = 0.05.

3.7. Ubiquitous expression of the Dilp8 C-peptide is not sufficient to induce a developmental delay

Both *pro*Dilp8* and *proDilp8_D6Cpep* are expected to retain their respective C-peptides. As all of our constructs are C-terminal-tagged with a 3xFLAG, we could theoretically detect the unprocessed form of proDilp8 and the processed A-chain. Thus, to confirm our expectations, extracts of whole larvae ubiquitously expressing these transgenes were prepared and Western Blot was performed using an anti-FLAG antibody. For all the overexpressing proDilp8 peptides, the immature Dilp8 was detected at approximately 17kDa, except for *proDilp8_D6Cpep* that is slightly smaller, due to shorter Dilp6 C-peptide (**Figure 3-8**). Yet, cleaved Dilp8 (without the C-peptide) was not detected in any conditions, presumably due to their size (~5kDa). The detection of the smaller processed peptides could require enrichment by immunoprecipitation prior to Western Blot analyses, something that should be done in the future.

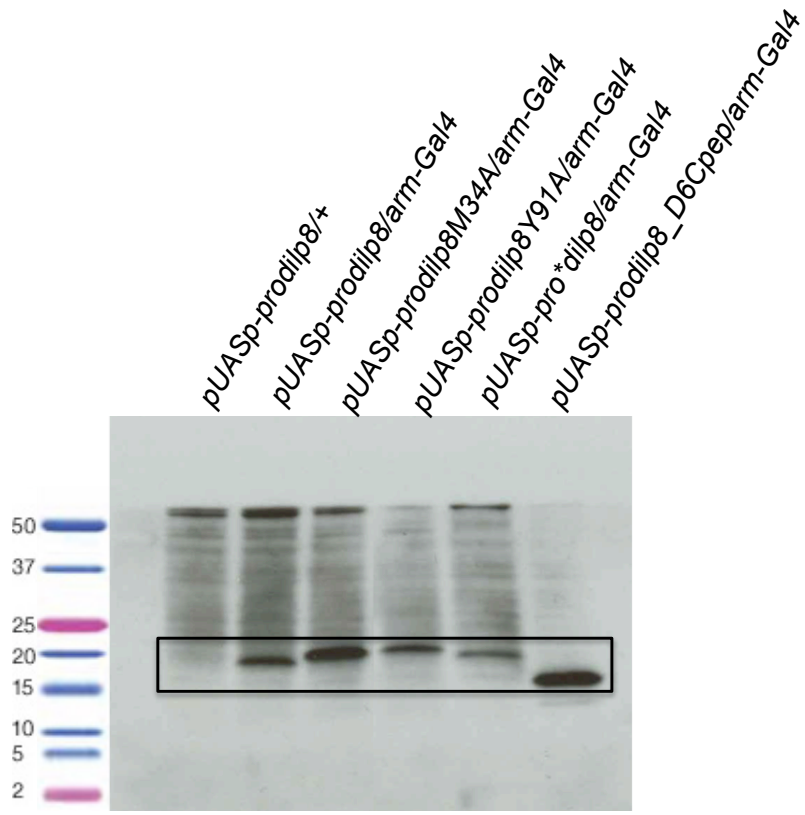


Figure 3-8 – Western blot of different proDilp8 transgene constructs. Western Blot using anti-FLAG antibody showing the presence of proDilp8 in extracts of whole larvae of the depicted genotypes.

As we were unable to detect the processed A-chain to demonstrate the lack of processing in the pro*Dilp8 construct, we tried a parallel genetic strategy to demonstrate or rule-out a possible function of the C-peptide on its own on developmental timing control. We engineered a UAS expression construct that should encode a secreted peptide consisting of a signal peptide followed by the Dilp8 C-peptide tagged with the OLLAS epitope in the C-terminus (*pUASt-Cpeptide::OLLAS*). Overexpressing this construct using *arm-Gal4* did not significantly affect developmental timing (**Figure 3-9**). In summary, this suggests that the C-peptide on its own does not control developmental timing, being this function exclusive of the mature Dilp8.

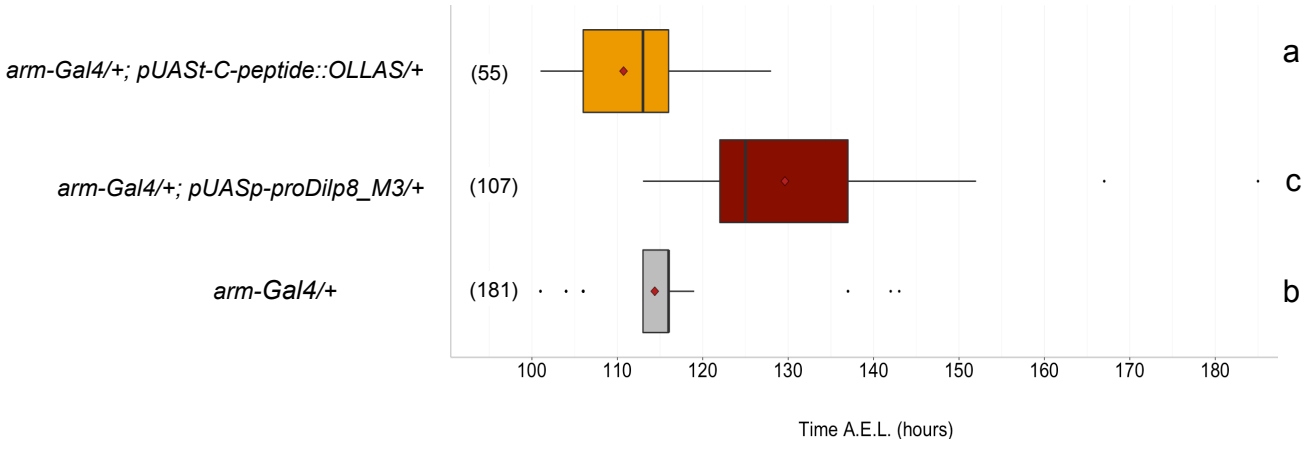


Figure 3-9 – Overexpression of the Dilp8 C-peptide is not sufficient to induce a developmental delay. Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASt-C-peptide::OLLAS/+* under the control of *arm-Gal4. arm-Gal4* crossed to *w[1118]* or *UAS-prodip8* (*pUASp-prodilp8_M3*) served as negative and positive controls, respectively. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

3.8. C-peptide does not influence weight

While the Dilp8 C-peptide does not seem to be directly involved in delaying the onset of metamorphosis, it could nevertheless still be directly responsible for other Dilp8-related phenotypes, such as imaginal disc growth inhibition and/or weight increase. To test the latter hypothesis pharate adults ubiquitously expressing different *proDilp8* mutant lines during development were weighed. **Figure 3-10** represents the average weight of male **(A)** and female **(B)** pupae of three different insertions per genotype, except for the C-peptide::OLLAS construct, for which we only three repeats for the same line. Weight data per insertion is depicted in **Appendix 6-6**. Even though there was a tendency for the genotypes that delayed development to be heavier than the others, none of the differences reached statistical significance ($p > 0.05$, Kruskal-Wallis with Dunn's multiple comparisons test). This could be due to the weaker delay (and hence less time to gain weight) caused by our *pUASp-proDilp8* constructs compared to the *pUASt* constructs that have been used previously [42]. Importantly, these experiments clearly show that ubiquitous expression of the Dilp8 C-peptide alone does not increase weight, suggesting that it is not involved in the weight gain observed in Dilp8 overexpression conditions.

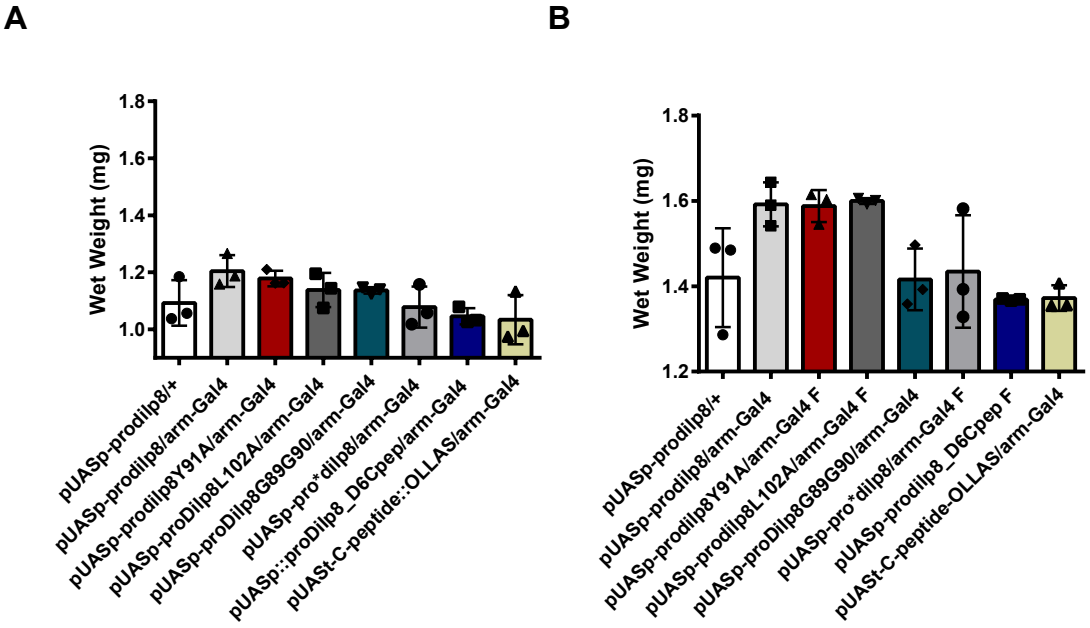


Figure 3-10 – Pupal wet weight in mg for males (**A**) **and females** (**B**). Each symbol represents the average weight of three independent P-element insertions, except for *pUASt-C-peptide::OLLAS*. For this genotype, each symbol represents the average of three independent weight measures of the same line. The bars represent the average of these averages. The genotypes analyzed are statistically indistinguishable.

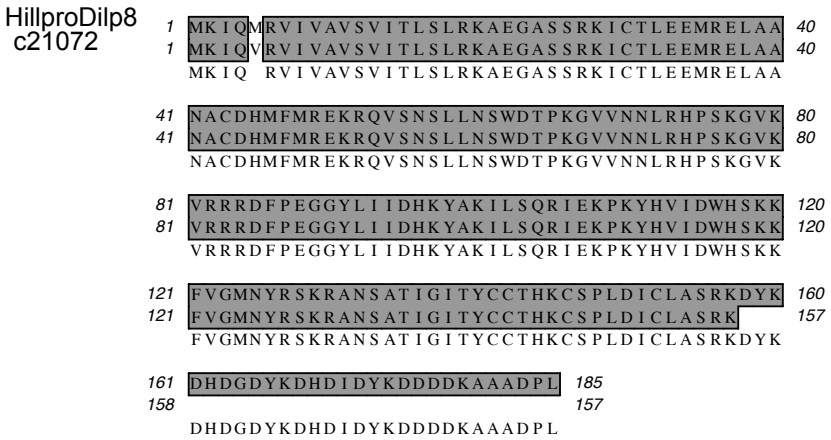
3.9. *Hermetia illucens* Hilp8

H. illucens is a large wasp-like fly, from the major Brachycera order Stratiomorpha. *H. illucens* is thought to have shared a last common ancestor with *D. melanogaster* about ~180 million years ago, which coincides with the time the last common ancestor to all Brachycera should have lived. Hence, by studying the biology of the *H. illucens dilp8* homologue, *hilp8*, we can make inferences about how the *ilp8-like* gene of the last common ancestor of all Brachycera functioned. Moreover, culture methods of *H. illucens* are established and the larvae are easily obtained commercially [49-51]. For these reasons, *H. illucens* was considered an optimal species to learn about the evolution and function of *dilp8*.

Nevertheless, in non-model organisms such as *Hermetia*, the available information and experimental techniques are often restricted. One of the recent techniques available for these organisms is RNAseq. RNAseq analysis using adult *H.* illucens females revealed the presence of at least two *dilp8* homologue sequences, hereafter referred to as *hilp8a* and *hilp8b*. These two sequences are represented in **Figure 3-11 (A)** and **(B)**.

One of the homologues (contig 21072, which encodes a preprohilp8, hereafter referred to as *prohilp8a*) was cloned into pUASp in frame with a C-terminal 3x::FLAG tag and transformed into *D. melanogaster*, making the line *UAS-prohilp8a*. Ubiquitous expression of *prohilp8a* under the control of *arm-Gal4* did not significantly delay development (**Figure 3-12).** On the contrary, two out of three P-element insertions pupariated ~5 h earlier than *arm-Gal4/+* controls. These results demonstrate that ubiquitous *prohilp8a* expression in *D. melanogaster* does not lead to a developmental delay. There are several interpretations to this result, however the simplest one is that the Ilp8-like molecule of the last common ancestor to all Brachycera did not have the ability to delay development, which in turn suggests that this activity arose within Brachycera, in the clade that would then originate *D. melanogaster*.

A



B

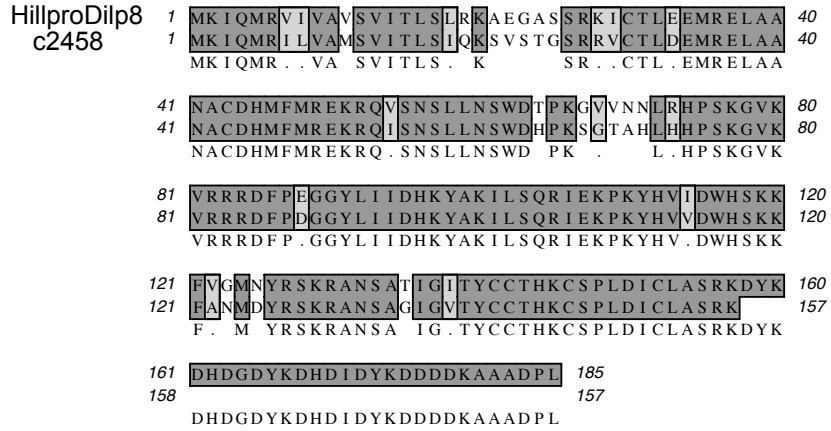


Figure 3-11 – *H. illucens* Ilp8 homologues and proHilp8 protein sequence alignment. (A) Protein alignment of proHilp8a (from the construct *pUASp-proHilp8::3xFLAG*, upper line) and the homologue obtained from contig 21072 (bottom line). **(B)** Protein alignment of proHilp8 (upper line) and the homologue obtained from contig 2458 (lower line; hereafter Hilp8b).

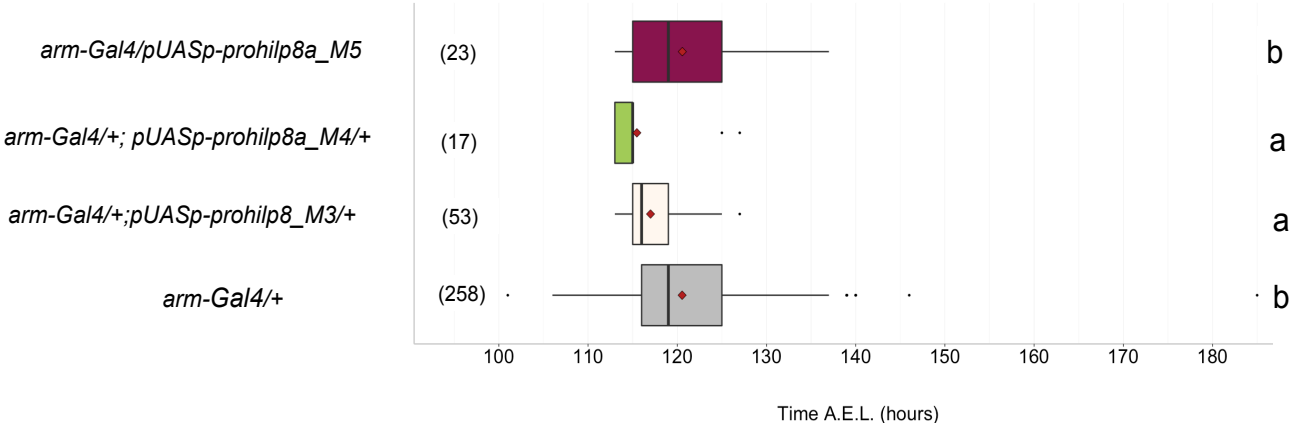


Figure 3-12 – Ubiquitous expression of *prohilp8a* (see Figure 3-9 (A)) in *D. melanogaster* does not delay the onset of metamorphosis. Box plot showing pupariation time (h after egg laying) of *(N)* larvae expressing *pUASp-prohilp8* under the control of *arm-Gal4.* M3, M4 and M5 correspond to three different P-element insertion sites. *arm-Gal4* crossed to *w[1118]* served as control. The black vertical bars represent the median; the firebricks represent the mean. Genotypes sharing the same letter are statistically indistinguishable at alpha = 0.05.

3.10. *Hermetia illucens* RNAseq

To test whether the *dilp8* transcriptional upregulation response to abnormal-tissue growth is also found in *H. illucens,* imaginal tissue damage was induced by injecting EMS (or PBS as controls) into *H. illucens* 4th instar larvae. Then we assessed global transcriptional changes by RNAseq, in collaboration n with Dr. Tatiana Torres (Genomics and Evolution of Arthropods Lab, Instituto de Biociências of the São Paulo University). EMS treatment leads to apoptotic cell death in imaginal discs, leading to *dilp8* upregulation, and consequently a delay in the onset of metamorphosis in *D. melanogaster* [42].

An overall comparison of the two *H. illucens* RNAseq assemblies (the previous one, using adult females and the current one, using larvae) is provided in **Appendix 6-7**. The higher number of contigs >1000 bp in the larval assembly is indicative of superior quality.

The EMS- and PBS-treated samples were compared using a differential expression (DE) analysis, in which the *reads* were aligned with the assembled transcriptome. The number of *reads* is proportional to the expression level of the genes. The DE evaluation revealed the existence of 600 *contigs* differentially expressed in the EMS-treated samples. From these, 467 corresponded to transcripts with induced expression.

We verified if genes known to be involved in the regulating metamorphosis (homologues of ecdysone biosynthesis and ecdysone-response genes in *Drosophila*) or in tissue damage regeneration (*e.g., ilp8*) were amongst these upregulated transcripts. The normalized *read* counts for these selected genes are depicted in **Table 3-2.** A great variation between biological replicates of the same treatment is clear (this topic will be discussed later). Furthermore, considering the differential expression of these same genes (including *ilp8*) none showed a significant difference between EMS and PBS treatments (**Table 3-3**). In **Figure 3-13** the counts for both *hilp8* genes are represented. In summary, it seems that the RNAseq analysis did not reveal a consistent enrichment in *hilp8* levels following EMS challenge

Table 3-2 – RNAseq results from *H. illucens* larvae treated with EMS or PBS (controls).

	Sample			
Candidate Gene	PBS 1	PBS 2	EMS 1	EMS 2
shadow	0	1,42	2,37	0,06
shade	3,27	1,49	2,64	1,02
disembodied	2,31	0,56	1,27	0,13
ecdysone-induced protein	1,72	25,76	1,69	0,70
ultraspiracle	1,02	2,72	1,76	1,15
insulin-like peptide 8 A	5,90	1,92	33,90	1,34
Insulin-like peptide 8 B	3,16	0,81	43,51	0,70

Table 3-3 – Statistical analyses of RNAseq results from *H. illucens* larvae treated with EMS or PBS (controls) for selected genes.

Candidate gene	logFC	logCPM	P-Value	FDR
shadow	0,772	0,108	0,629	1
shade	-0,382	1,157	0,659	1
disembodied	-1,023	0,261	0,345	1
ecdysone-induced protein	-3,513	2,924	0,002	0,068
ultraspiracle	-0,359	0,835	0,665	1
insulin-like peptide 8 A	2,168	3,443	0,064	0,718
insulin-like peptide 8 B	3,469	3,603	0.011	0.222

logFC, log2 Fold Change; logCPM, log2 Counts Per Million; FDR stands for False Discovery Rate and gives a corrected significance value.

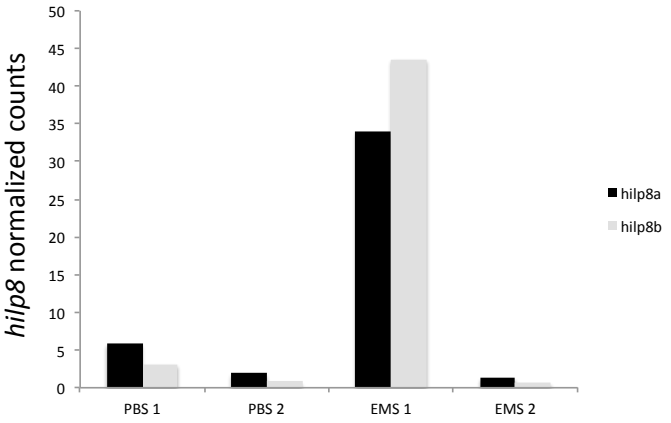


Figure 3-13 – RNASeq *read* counts normalized to the size of the transcript library. Bar graphs representing *hilp8a* and *hilp8b* levels after EMS challenge. There is no statistically significant difference for either *hilp8* gene between treatments.

3.11. qPCR analysis of hilp8 expression levels

To verify our RNAseq results using an independent technique, we quantified *hilp8* mRNA levels in aliquots of the same mRNA samples using qRT-PCR (notice that the analyses of the third PBS and EMS repeats by RNAseq has not been finalized in time for this thesis report). The utilized primer pairs recognize the common region for *hilp8a* and *hilp8b*, meaning that the expression levels detected correspond to both genes. **Figure 3-14** shows the relative *hilp8* mRNA levels normalized the housekeeping genes *H. illucens rp49*, where we find no consistent effect of EMS treatment on *hilp8* levels.

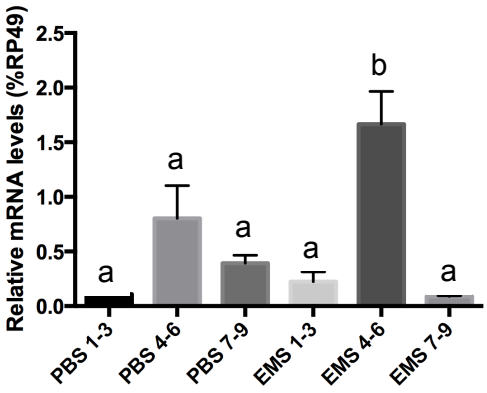


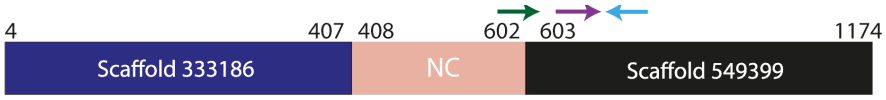
Figure 3-14- EMS treatment does not influence *hilp8* mRNA levels in *H. illucens*. Bar graphs depicting *hilp8* mRNA levels normalized to *H. illucens rp49.* Three independent samples of EMS-injected or PBS (control)-injected samples were analyzed by qRT-PCR. Bars sharing the same letter are statistically indistinguishable at alpha=0.05.

3.12. Hilp8a and Hilp8b are encoded by different genomic sequences

It was unclear whether *hilp8a* and *hilp8b* were isoforms of the same gene or whether they represented paralagous genes or even alleles of the same gene. To elucidate this question, we used BLAST to compare the sequences of both *hilp8a* and *hilp8b* transcripts to *H. illucens* genomic scaffolds (generously provided by Doris Barchtrog) (**Figure 3-15**). We found that *hilp8a* is similar to genomic scaffolds 333186 (between nucleotide 4 and 407) and 549399 (nucleotides 603-11174) – **Figure 3-15 (A)**. On the other hand, parts of *hilp8b* were found in genomic scaffolds 175853 (between nucleotides 1 and 178) and 283090 (between nucleotides 406 and 603). These results show that *hilp8a* and *hilp8b* originate from different genomic regions, suggesting that they are either paralagous genes or highly derived alleles of the same gene. We then performed a series of PCR assays followed by Sanger sequencing with genomic DNA (gDNA) or cDNA using specific primer pairs for both predicted *hilp8a* and *hilp8b* sequences. We could confirm the presence of both *hilp8a* and *hilp8b* by PCR in all five *H. illucens* samples tested (**Figure 3-15 (A) and (B)**, respectively), suggesting that it is unlikely that *hilp8a* and *hilp8b* are segregating as alleles in our population (and in the individuals that had their genomes sequenced by the D Bachthrog group). Taken together, most likely *hilp8a* and *hilp8b* represent two different genes, which are expressed in *H. illucens* larval and adult stages, as shown by RNAseq analysis.

A

hilp8a



B

hilp8b

Figure 3-15 – The *hilp8a* and *hilp8b* are transcribed from distinct genomic regions. (A) The *hilp8a* transcript aligns with the genomic scaffold 333186 between nucleotides 4 and 407. For the region between nucleotide 408 and 602 no corresponding scaffold was found (NC). The region between nucleotide 603 and 1174 aligns to scaffold 549399. **(B) Nucleotides 1 to 178 of the** *hilp8b* transcript aligns to the scaffold 175853. From position 406 to 603 this sequence is similar to the scaffold 283090. No corresponding scaffold was found for both 179-405 and 604-681 nucleotide gaps. Arrows correspond to sets of primers designed to test and/or confirm the existence of *hilp8a* and *hilp8b*. Dark green arrow is ≠ Scaffold 283090; purple and light blue arrows are the set = Scaffold 549399; yellow arrow corresponds to = Scaffold 283090; light green and red arrows correspond to the set ≠ Scaffold 549399 (see Materials and Methods section).

3.13. *hilp8a* and *hilp8b* gene structure

To unravel the structures of *hilp8a* and *hilp8b*, the two RNAseq derived sequences were aligned with the *H. illucens* genome SRA (Sequence Read Archive) trace reads available at the NCBI database. Since the available SRA reads were genomic DNA derived and the *hilp8a* and *hilp8b* sequences were obtained through RNAseq analysis, the alignment of the paired reads could reveal the presence of exon-intron boundaries. Indeed, for *hilp8a* the results show a distribution in three discrete parts (1~400; 401~600; 601~1200, **Figure 3-16 (A)**), in which the block boundaries may represent exon-intron junctions and each block may cover for one exon. Three different blocks are also distinguishable for *hilp8b* (1~180; 181~380; 381~650; **Figure 3-16 (B)**). To confirm if each of these parts correlates with three exons, *hilp8a* and *hilp8b* sequences were aligned with all the respective genomic reads around the exon-intron frontier, to find exon donor and exon acceptor sequences, that delimitate the exon and that are not present in the transcript sequence. Indeed, *hilp8a* alignments revealed the presence of an exon donor sequence (GT) in the 408th position, that is present in the genomic reads but not in *hilp8a* transcript sequence. A second exon donor sequence was identified in position 604. Considering these results, *hilp8a* consists of at least 3 exons (1~407;

408~603; 604~1174 – considering the transcript sequence), just as has been predicted by the alignment of *hilp8a* sequence and the genome SRA trace reads. Similarly, two exon donor sites were found for *hilp8b* (one in the 179th position and other after 350bp), suggesting that *hilp8b* also consists of at least 3 exons (1~178; 179~350; 351~681 – considering the transcript sequence), as suggested before by the BLAST of the transcript sequence with the *H. illucens* genomic reads.

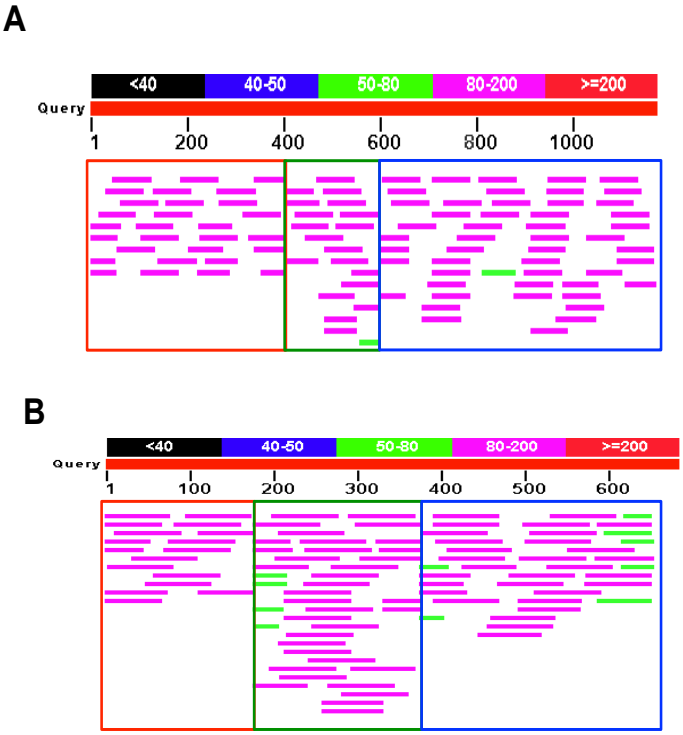


Figure 3-16 – BLAST results for *hilp8a* and *hilp8* transcript sequences with the *H. illucens* genome SRA trace reads. The block-like distribution is due to the alignment of transcript sequences with the genome. (A) *hilp8a* alignment with *H. illucens* genomic reads revealed the existence of three different blocks (1~400 in red; 401~600 in green; 601~1200 in blue) that may correspond to exons. **(B)** *hilp8b* sequence BLAST also distinguishes three blocks (1~180 in red; 181~380 in green; 381;650 in blue), which are putative exons.

3.14. *H. illucens hilp8a* and *hilp8b* ovarian expression

In adult female *D. melanogaster* flies, *dilp8* is highly enriched in the ovary [42]. To examine if *H. illucens* females show similar expression pattern, RNA was extracted from the heads and ovaries of adult females and expression of two *hilp8* genes were quantified by qRT-PCR (five biological replicates with three technical replicates). Results show that *hilp8a* is expressed at a similar level in the head and ovary (**Figure 3-17 (A)**). On the other hand, *hilp8b* expression was significantly reduced in the ovary (**Figure 3-17 (B)**). Although this is not in exact accordance with *dilp8* enrichment in *Drosophila* adult ovaries, it reveals different spatial expression patterns of the two *Hilp8* paralogues.

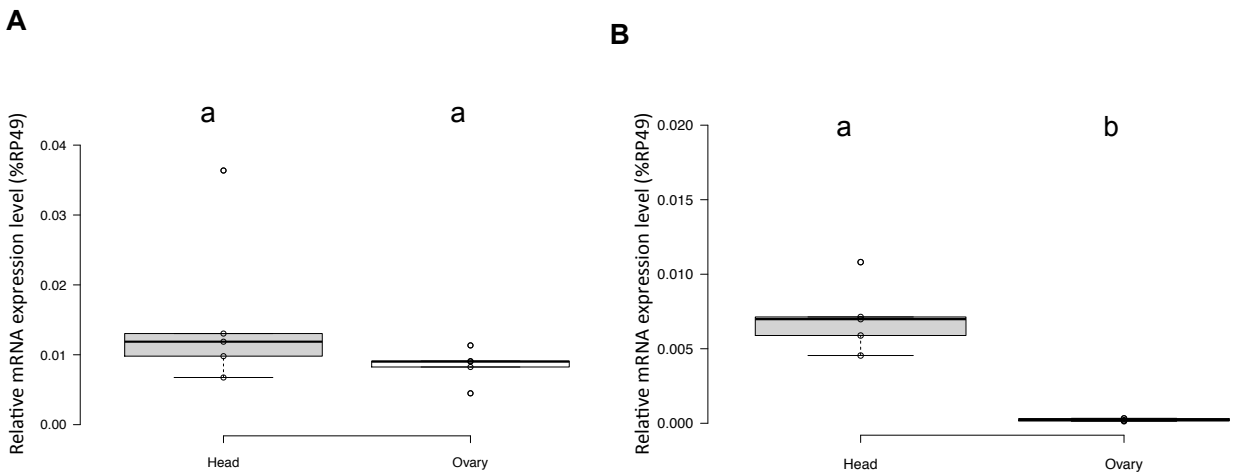


Figure 3-17 – Expression of *hilp8* genes in head and ovary. Box plots representing *hilp8a* and *hilp8b* mRNA expression levels in head and ovary samples estimated by qRT-PCR. **(A)** *hilp8a* expression in the head and in the ovary is not statistically indistinguishable (n=5, $p > 0.05$). **(B)** Significant differential expression is detected for *hilp8b* (n=5; $p = 0.0079$).

4. Discussion

In *Drosophila*, the insulin/relaxin-like peptide Dilp8 is produced and released in response to damages inflicted upon the imaginal discs, acting as a major regulator of developmental stability [42, 43]. Despite of the importance of this peptide in growth coordination, its mechanisms of action and evolutionary history were still unclear. Having this in mind, this study aimed at revealing the molecular mechanisms of action and evolution of Dilp8.

The molecular evolution study revealed the presence of Dilp8 homologues in various dipteran species from the Brachycera branch, but not in the nematocerans (mosquitoes). This observation suggests that Ilp8 might have been a Brachycera innovation, having suffered a very fast evolution rate. Nevertheless, Ilp8 is the second most conserved Ilp in *Drosophila*, indicating that after having quickly diverged following its origin, Ilp8 might have been under high selective constraints in *Drosophila*.

Based on the Ilp8 alignment, several highly conserved amino acids were found, including residues in the C-peptide, which is not predicted to be a part of the mature Dilp8 peptide. The conserved residues in the C-peptide could be required for proper Dilp8 processing or even for a specific activity of the C-peptide as a separate hormonal entity from mature Dilp8, considering that the C-peptide is expected to be secreted together with Dilp8, in equimolar concentrations [60].

Therefore, it was formally possible that one or more of the activities that have been attributed to Dilp8 (i.e., an ability to delay the onset of metamorphosis, to inhibit imaginal disc growth and to increase larval weight), could actually be dependent on the Dilp8 C-peptide, rather than the mature Dilp8. I was further encouraged to test this possibility because the insulin C-peptide of vertebrates retains biological functions, such as improvement of vascular, neural and renal dysfunction in diabetic individuals [46-48]. Also, a study performed in human diabetic patients revealed that plasma C-peptide concentrations are higher in obese subjects [59]. The results in this study are however more consistent with a role for the C-peptide in promoting proper posttranslational processing of the Dilp8 peptide, rather than an independent role for the C-peptide as a separate hormonal entity, even though we cannot exclude the latter. First, I show that the expression of the C-peptide alone does not delay development or increase larval weight. This is strong evidence that these activities require the presence of the mature Dilp8. This study also suggests that the C-peptide needs to be properly processed out of the mature Dilp8 in order for it to delay development and likely increase larval weight. This comes from experiments using proDilp8 constructs in which the C-peptide is expected to be uncleavable (*pro*Dilp8* and *proDilp8_D6Cpep*), all of which render Dilp8 functionless.

This is also consistent with the findings that substituting either of two conserved amino acids in the C-peptide (Y91 and L102) for alanines did not affect Dilp8 functions. However, removal of a conserved diglycine motif (G89G90) in the C-peptide rendered

Dilp8 functionless. There are at least two possibilities that could explain this finding. First, the vertebrate proinsulin C-peptide participates in proinsulin folding by promoting proper alignment of the B- and A-chains [58]. It is therefore possible that the Dilp8 C-peptide lacking the diglycine motif has an improperly aligned A-chain and B-chain, which results in no Dilp8 activity. Yet, truly confirming that the elimination of Dilp8 activity is due to improper alignment of A- and B-chains is complicated. One possibility is to explore if this effect is triggered with the deletion of any two amino acids in the C-peptide. If not, the C-peptide might have another function besides conferring enough flexibility to allow mature Dilp8 folding. One evidence against this hypothesis is that the Ilp8 C-peptide does not seem to be heavily constrained in Brachycera as regards length. An alternative possibility one needs to consider is that the diglycine motif per se is required, not the exact length of the C-peptide. The diglycine motif could be necessary for any aspect of proDilp8 maturation, including its secretion or proper C-peptide processing. One way to distinguish between a length and amino acid type requirement is by changing the diglycine motif into a dialanine motif. If ubiquitous expression of this construct is also unable to delay development, then it would indicate that the diglycine motif itself is critical.

We were unfortunately unable to visualize Dilp8 C-peptide processing by Western blot analyses. If we could devise an assay to monitor Dilp8 maturation, we could determine whether the digylcine motif is critical for C-peptide release. One possibility is to immunoprecipitate the FLAG-tagged proteins from the hemolymph to enrich for the putatively processed fragments, similarly to what was done in [43]. Another alternative would be to try to detect the processed Dilp8 C-peptide by using a specific antibody against it. This could be possible, as one of the anti-Dilp8 antibodies described by Colombani et al. maps to the C-peptide region [43].

Apart from bringing insight into C-peptide processing and biology we have also identified a critical amino acid in the Dilp8 B-chain, M34. It would be interesting to test whether or not the other three absolutely conserved amino acids in the B-chain are also critical for Dilp8 activity. The fact that the B-chain is more conserved than the A-chain could have interesting implications, especially when the direct receptor for Dilp8 is identified.

To understand how *dilp8* evolved, we studied distant *dilp8* homologues from a species that shared a common ancestor with *Drosophila* about 180 million years ago (*H. illucens*) were studied. RNAseq of adult *H. illucens* females revealed the presence of at least two *dilp8* homologues, which were named *hilp8a* and *hilp8b*. Based on PCR and Sanger sequencing, *hilp8a* and *hilp8b* are likely transcribed from two distinct genomic sequences with three exons each, similarly to most *ilps*, including the human insulin [55]. Mapping of the *hilp8a* and *hilp8b* genes of two distinct chromosomal loci will provide the definitive proof that they are two independent genes and not two alleles of the same gene.

Functional experiments showed that ubiquitous expression of one of these genes, *hilp8a,* in *Drosophila* did not delay the timing of pupariation. The most parsimonious explanation for this result is that *ilp8* did not function as a developmental delay-inducing factor early in the evolution of Brachycera, having evolved this function only later during the divergence of the clade that would originate *Drosophila*.

RNAseq was also performed to compare *H. illucens* larvae treated or not with EMS. Our results showed that neither *hilp8a* nor *hilp8b* transcription was significantly increased after EMS challenge. Overall, the high variability observed between biological replicates of the same treatment makes it difficult to understand how *hilp8a* and *hilp8b* are in fact responding to the damage. To overcome this, a new RNAseq will be performed in which the larvae will be resynchronized and single challenged larva will be used as a sample (instead of pools of 3 larvae/sample). This should reduce the variability associated with developmental stage and can elucidate if the variation seen in both RNAseq and qPCR studies is solely related to the effect of EMS on *hilp8* expression. Still it would be of most interest to prove that EMS injection was efficiently producing tissue damage in *H. illucens*. To do that, wing discs of *H. illucens* larvae injected with EMS should be immunostained with an anti-Caspase3 antibody, to show that EMS triggers apoptosis, just as expected in this experimental design.

Finally, the qRT-PCR experiments did not reveal any clear enrichment of the two *hilp8* trasncripts in the female reproductive tissue, showing that even their expression pattern in adults is different from that of *D. melanogaster* [42]. In this scenario, it would be important to look more carefully into the expression pattern of *hilp8a* and *hilp8b*. To do that, specific *in situ* hybridization probes need to be designed and the expression patterns of both *hilp8a* and *hilp8b* should be explored in both adults and larvae.

This work shed new light into the mechanism of action and evolution of Dilp8 and opens up several questions that can be already addressed with new experiments. The exploration of these mechanisms is important, especially because there is clinical evidence that tissue damages (*e.g.*, chronic inflammations, infections or tissue regeneration) can generate a growth and reproductive delay, even in humans [11]. Understanding the mechanisms of action of the peptide and the pathway responsible for controlling developmental timing in insects can further help the study of these phenomena in a wide variety of organisms, including vertebrates.

5. References

[1] Adams, M.D *et al*. (2000). The genome sequence of drosophila melanogaster. *Science* **287:**2185-2195

[2] Reiter L.T. *et al* (2001). A systematic analysis of human disease associated gene sequences in Drosophila melanogaster. *Genome research* **11**:1114-1125

[3] Robertson C.W. (1939) The Metamorphosis of Drosophila melanogaster, Including an Accurately Timed Account of the Principal Morphological Changes. *Insect Morphology* **59**:351-399

[4] Hackney J.F., Zolali-Meybodi O., Cherbas P. (2012). Tissue Damage Disrupts Developmental Progression and Ecdysteroid Byosynthesis in Drosophila. *PLoS ONE* **7**:1–12

[5] Yamanaka N., Rewitz K.F., O'Connor M.B. (2013). Ecdysone Control of Developmental Transitions: Lessons from Drosophila Research. *Annual Review of Entomology* **58**: 497-516

[6] McBrayer Z. *et al* (2007). Prothoracicotropic Hormone Regulates Developmental Timing and Body Size in *Drosophila*. *Developmental Cell* **13**(6):857:871

[7] Smith W., Rybczynski R. (2012). Prothoracicotropic Hormone. In *Insect Endocrinology*, ed. LI Gilbert, pp. 1 – 62. Elsevier

[8] Hariharan I.K. (2012). How Growth Abnormalities Delay "Puberty" in Drosophila. *Science Signaling* **5**:pe27-pe27

[9] Siegmund T., Korge G. (2001). Innervation of the Ring Gland in *Drosophila melanogaster*. *Journal of Comparative Neurology* **431**:481-91

[10] Rewitz K.F., Yamanaka N., Gilbert L.I., O'Connor M.B. (2009). The Insect Neuropeptide PTTH Activates Receptor Tyrosine Kinase Torso to Initiate Metamorphosis. *Science* **326**:1403-1405

[11] Andersen D.S., Colombani J., Léopold P. (2013). Coordination of Organ Growth: Principles and Outstanding Questions From The World of Insects. *Trends Cell Biology* **23**:336-344

[12] Niwa Y.S., Niwa R. (2014). Neural Control of Steroid Hormone Biosynthesis During Development in the fruit fly Drosophila melanogaster. *Genes & Genetic Systems* **89**:27-34

[13] Riddiford L.M., Hiruma, K., Zhou, X., Nelson, C.A. (2003). Insights into the

molecular basis of the hormonal control of molting and metamorphosis from Manduca sexta and Drosophila melanogaster. *Insect biochemistry and molecular biology* ***33***(12):1327-1338

[14] Edgar B.A. (2006). How flies get their size: genetics meets physiology. *Nature Reviews Genetics* ***7***(12):907-916

[15] Koyama T. *et al* (2014). Nutritional Control of Body Size Through FoxO-Ultraspiracle Mediated Ecdysone Biosynthesis. *eLife* **3:**1-20

[16] Shingleton A.W., Das J., Vinicius L., Stern D.L. (2005). The Temporal Requirements for Insulin Signaling During Development in Drosophila. *PLoS Biology* **3**:1607-1617

[17] Colombani J. *et al* (2005). Antagonistic Actions of Ecdysone and Insulins Determine Final Size in Drosophila. *Science* **310**: 667-670

[18] Okamoto N. *et al* (2009). A Fat-Body-Derived IGF-like Peptide Regulates Postfeeding Growth in Drosophila. *Developmental Cell* **17**:885-891

[19] Antonova Y., Arik A.J., Moore W., Riehle M.A., Brown M.R. (2012). Insulin-Like Peptides: Structure, Signaling, and Function. In *Insect Endocrinology*, ed. LI Gilbert, pp. 1 – 62. Elsevier

[20] Orme M.H., Leevers S.J. (2005). Flies on Steroids: The Interplay Between Ecdysone and Insulin Signaling. *Cell Metabolism* **2**:277-278

[21] Wu Q., Brown M.R. (2006). Signaling and Function of Insulin-Like Peptides in Insects. *Annual Review of Entomology* **51**:1-24

[22] Goberdhan D.C.I., Wilson C. (2003). The Functions of Insulin Signaling : Size Isn't Everything, Even in *Drosophila*. *Differentiation* **71**:375-397

[23] Zhang H.*et al* (2009). Deletion of *Drosophila* Insulin-Like Peptides Causes Growth Defects and Metabolic Abnormalities. *Proceedings of the National Academy of Sciences* **106**:19617-19622

[24] Brogiolo W. *et al* (2001). An Evolutionarily Conserved Function of the Drosophila Insulin Receptor and Insulin-like Peptides in Growth Control. *Current Biology* **11**:213-221

[25] Slaidina M. *et al* (2009). A Drosophila Insulin-like Peptide Promotes Growth During Nonfeeding States. *Developmental Cell* **17**:874-884

[26] Nässel D.R., Kubrak O.I., Liu Y., Lushchak O.V. (2013). Factors that regulate insulin producing cells and their output in *Drosophila*. *Frontiers in physiology*, **4**:1-12

[27] Van Loy T. *et al* (2008). Comparative genomics of leucin-rich repeats containing G-protein-coupled receptors and their ligands. *General and Comparative Endocrinology*, **155**:14-21

[28] Van Hiel M.B., Vandersmissen H.P., Van Loy T., Broeck J.V. (2012). An evolutionary comparison of leucin-rich repeat containing G-protein-couples receptors reveals a novel LGR subtype. *Peptides* **34**:193-200

[29] Van Hiel M.B., Vandersmissen H.P., Proost P., Broeck J.V. (2014). Cloning, constitutive activity and expression profiling of two receptors related to relaxin receptors in *Drosophila melanogaster*. *Peptides* **68**:83-90

[30] Garreli A. *et al* (2015). Dilp8 Requires the Neuronal Relaxin Receptor Lgr3 to Couple Growth to Developmental Timing. *Nature Communications*

[31] Grönke S. *et al* (2010). Molecular Evolution and Functional Characterization of Drosophila Insulin-Like Peptides. *PLoS Genetics* **6**, e1000857

[32] Casimiro, AP (2014). Characterizing the function of two *Drosophila* Leucine-rich repeat-containing g-protein coupled receptors. Master's Thesis in Molecular Biology for Health – Escola Superior de Saúde Egas Moniz, Lisboa, 105 pp.

[33] Mirth C., Truman J.W., Riddiford L.M. (2005). The Role of the Prothotacic Gland in Determining Critical Weight for Metamorphosis in *Drosophila melanogaster*. *Current Biology* **15**:1796-1807

[34] Stieper B.C., Kupershtok M., Driscoll M.V., Shingleton A.W. (2008). Imaginal Discs Regulate Developmental Timing in *Drosophila melanogaster*. *Developmental Biology* **321**:18-25

[35] Worley M.I., Setiawan L., Hariharan I.K. (2012). Regeneration and Transdetermination in Drosophila Imaginal Discs. *Annual Review of Genetics* **46**:289-310

[36] Hussey R.G., Thompson W.R., Calhoun E.T. (1927). The Influence of X-rays on the Development of Drosophila larvae. *Science* **66**:65-66

[37] Russel M. (1974). Pattern Formation in Imaginal Discs of a Temperature Senitive Cell Lethal Mutant of Drosophila melanogaster. *Developmental Biology* **40**:24-39

[38] Simpson P., Schneiderman H.A. (1975). Isolation of Temperature Sensitive Mutations Blocking Clone Development in Drosophila melanogaster, and the Effects of a Temperature Sensitive Cell Lethal Mutation on Pattern Formation in Imaginal Discs. *Wilhem Roux's Archives of Developmental Biology* **178**:247-275

[39] Simpson P., Berreur P., Berreur-Bonnenfant J. (1980). The Initiation of Pupariation

in Drosophila: Depedence on Growth of the Imaginal Discs. *Journal of embryology and experimental morphology* **57**:155-165

[40] Shingleton A.W. (2010). The Regulation of Organ Size in Drosophila: Physiology, Plasticity, Patterning and Physical Force. *Organogenesis* **6**(2):76-87

[41] Parker N.F., Shingleton A.W. (2011). The Coordination of Growth Among Drosophila Organs in Response to Localized Growth-Perturbations. *Developmental Biology* **357**(2): 318-325

[42] Garelli A. *et al* (2012). Imaginal Discs Secrete Insulin-Like Peptide 8 to Mediate Plasticity of Growth and Maturation. *Science* **336**:579-582

[43] Colombani J., Andersen D.S., Léopold P. (2013). Secreted Peptide Dilp8 Coordinates Drosophila Tissue Growth with Developmental Timing. *Science* **336**:582-585

[44] Koyama T., Mendes C.C., Mirth C.K. (2013). Mechanisms Regulating Nutrition-Dependent Developmental Plasticity Through Organ-Specific Effects in Insects. *Frontiers in Physiology* **4**:263

[45] Jaszczak J.S., Wolpe J.B., Dao A.Q., Halme A. (2015). Nitric Oxide Regulated Growth Coordination During *Drosophila melanogaster* Imaginal Disc Regeneration. *Genetics* **115**:178053.

[46] Steiner, D.F. (2004). The proinsulin C-peptide—a multirole model. *Experimental Diabetes Research **5***(1):7-14.

[47] Steiner D.F. (1978). On the role of the proinsulin C-peptide*. Diabetes,* **27**(1):145-148.

[48] Wahren J. *et al* (2000). The role of C-peptide in human physiology. *American journal of physiology, endocrinology and metabolism* **278**:759-768

[49] Saurin H.E.M. (2011). Maggot – Bioconversion Research Program in Indonesia. Final Report, Bioconversion Indonesia 2005-2011

[50] Tomberlin J.K., Sheppard D.C., Joyce J.A. (2002). Selected Life-History Traits of the Black Soldier Flies (Diptera:Stratiomyidae) Reared on Three Artificial Diets. *Arthropod Biology*, **95**:379-386.

[51] Tomberlin J.K., Sheppard D.C. (2002). Factors Influencing Mating and Oviposition of Black Soldier Flies (Diptera:Stratiomyidae) in a Colony. *Journal of Entomology Science* **37(**4):345.352

[52] Nagalakshmi U., Waern K., Snyder M. (2001). RNA-Seq: A method for

comprehensive transcriptome analysis. *Current Protocols in Molecular Biology.* **4:11**(89):1-13

[53] Wang Z., Gerstein M., Snyder M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**(1):57-63

[54] Marioni J.C. *et al* (2014). RNA-Seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome research* **18:**1509-1517

[55] *Drosophila* 12 Genomes Consortium (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **450**:203-218

[56] Schwarz D. *et al* (2009). Sympatric ecological speciation meets pyrosequencing: sampling the transcriptome of the apple maggot *Rhagoletis pomonella. BMC Genomics* **10**(1):633

[57] International *Glossina* Genome iniciative (2014). Genome Sequence of the Tsetse Fly (*Glossina morsitans*): Vector of African Trypanosomiasis. *Science* **344**:380-386

[58] Steiner D.F., Chan S.J., Welsh J.M., Kwok S.C.M. (1985). Structure and Evolution of the Insulin Gene. *Annual Reviews in Genetics* **19**:463-484

[59] Reaven G.M. *et al* (1993). Plasma Insulin, C-peptide and Proinsulin Concentrations in Obese and Nonobese Individuals with Varying Degrees of Glucose Tolerance. *Journal of Clinical Endocrinology* **76**(1):44-48

[60] Cauter E.V. *et al* (1992). Estimation of Insulin Secretion Rates from C-peptide Levels. *Peptides* **41**:368-377

6. Appendix

Appedix 6-1 – Fly stocks used in the project

Name	Genotype	Origin
w[1118]	*w[1118]; +/+ ; +/+*	A gift from Maria Dominguez
If/CyO; MKRS/TM6B	*w;If/CyO;MKRS/TM6B Tb*	A gift from António Jacinto
MKRS/TM6B	*w;MRKS/TM6B Tb*	Generated in the Lab
Arm-Gal4	*iso;armGal4/armGal4;iso*	A gift from Pedro Domingos
proDilp8 _M1	*w; pUASp-prodilp8::3xF_M1/ pUASp-prodilp8::3xF_M1*	Generated in the Lab
proDilp8_M3	*w; +/+; pUASp-prodilp8::3xF_M3/ pUASp-prodilp8::3xF_M3*	Generated in the Lab
proDilp8_M9	*w; +/+; pUASp-prodilp8::3xF_M9/ pUASp-prodilp8::3xF_M9*	Generated in the Lab
proDilp8M34A_M4	*w; +/+; pUASp-prodilp8M34A::3xF_M4/ pUASp-prodilp8M34A::3xF_M4*	Generated in the Lab
proDilp8M34A_M9	*w; pUASp-prodilp8M34A::3xF_M9/ pUASp-prodilp8M34A::3xF_M9*	Generated in the Lab
proDilp8M34A_M10	*w; pUASp-prodilp8M34A::3xF_M10/ pUASp-prodilp8M34A::3xF_M10*	Generated in the Lab
proDilp8Y91A_M5	*w; pUASp::prodilp8Y91A::3xF_M5/ pUASp::prodilp8Y91A::3xF_M5*	Generated in the Lab
proDilp8Y91A_M8	*w; pUASp-prodilp8Y91A::3xF_M8/ pUASp-prodilp8Y91A::3xF_M8*	Generated in the Lab
proDilp8Y91A_M9	*w; pUASp-prodilp8Y91A::3xF_M9/ pUASp-prodilp8Y91A::3xF_M9*	Generated in the Lab
pro*Dilp8_M1	*w; pUASp-pro*dilp8::3xF_M1/ pUASp-pro*dilp8::3xF_M1*	Generated in the Lab

Name	Genotype	Origin
pro*Dilp8_M5	*w; +/+; pUASp-pro*dilp8::3xF_M5/ pUASp-pro*dilp8::3xF_M5*	Generated in the Lab
pro*Dilp8_M9	*w; pUASp-pro*dilp8::3xF_M9/ pUASp-pro*dilp8::3xF_M9*	Generated in the Lab
proDilp8D6Cpep_M3	*w; +/+; pUASp-prodilp8D6Cpep::3xF_M3/ pUASp::prodilp8D6Cpep::3xF_M3*	Generated in the Lab
proDilp8D6Cpep_M6	*w; pUASp-prodilp8D6Cpep::3xF_M6/ pUASp-prodilp8D6Cpep::3xF_M6*	Generated in the Lab
proDilp8D6Cpep_M10	*w; pUASp-prodilp8D6Cpep::3xF_M10/ pUASp-prodilp8D6Cpep::3xF_M10*	Generated in the Lab
HillproDilp8_M3	*w; +/+; pUASp-hillprodilp8::3xF_M3/ pUASp-hillprodilp8::3xF_M3*	Generated in the Lab
HillproDilp8_M4	*w; +/+; pUASp-hillprodilp8::3xF_M4/ pUASp-hillprodilp8::3xF_M4*	Generated in the Lab
HillproDilp8_M5	*w; pUASp-hillprodilp8::3xF_M5/ pUASp-hillprodilp8::3xF_M5*	Generated in the Lab
attP-attB	*y w p{Y+.nos-int.NLS}; p{CaryP y+}attP2*	Gift from Dr. Diogo Manoel
proDilp8L102 (1)	*w; pUASp-prodilp8L102A(1)/ pUASp-prodilp8L102A(1)*	Generated in the Lab
proDilp8L102 (2)	*w; +/+ pUASp-prodilp8L102A(2)/ pUASp-prodilp8L102A(2)*	Generated in the Lab
proDilp8L102 (3)	*w; +/+ pUASp-prodilp8L102A(3)/ pUASp-prodilp8L102A(3)*	Generated in the Lab
proDilp8G89G90 (1)	*w pUASp-prodilp8G89G90 (1)/ pUASp-prodilp8G89G90 (1)*	Generated in the Lab
proDilp8G89G90 (2)	*w pUASp-prodilp8G89G90 (2)/ pUASp-prodilp8G89G90 (2)*	Generated in the Lab
proDilp8G89G90 (3)	*w pUASp-prodilp8G89G90 (3)/ pUASp-prodilp8G89G90 (3)*	Generated in the Lab
C-peptide	*w; +/+ ; pUASt – C-peptide::OLLAS/ pUASt – C-peptide::OLLAS*	Generated in the Lab

Appendix 6-2 – *H. illucens* primers used for PCR analysis. The scaffolds indicated are referent to *H. illucens* genomic DNA sequences available on NCBI (http://www.ncbi.nlm.nih.gov)

Primer Name	Gene	Primer sequence	Primer Size (bp)	Fragment Size (bp)
≠ Scaffold 283090	*hilp8*	F- GCAGTGAGCGTAATCACGTTAAGTCTGCG	29	≈ 350 bp (with = Scaffold 549399)
≠ Scaffold 549399	*hilp8*	F - CCACCCCAGCAAAGGGGTGAAAGTGCGG R - GCGTTTACTGCGATAGTCCATGTTGG	28 26	≈ 200 bp
= Scaffold 283090	*hilp8*	F - CAATGAGTGTCATCACACTAAGCATACA	28	≈ 350 bp (with ≠ Scaffold 549399)
= Scaffold 549399	*hilp8*	F- CCATCCAAGCAAGGGAGTGAAAGTAAGA R- GCGTTTGCTTCGGTAGTTCATGCCGA	28 26	≈ 200 bp

Appendix 6-3 - *Hermetia illucens* primers used for qPCR.

Primer name	Gene	Primer sequence	Primer Size (bp)	Fragment Size (bp)
Hill_dilp8	*hilp8*	F – TGCCTGTGATCATATGTTTATGC R- TTTGGCATATTTGTGGTCGATA	23 22	178
Hill_rp49	*rp49*	F – GCGAATTGGAGGTCTTGATG R - TACAATTTCCTTGCGCTTCT	20 20	≈150
Hill_EF2	*ef2*	F- CGCTCTGCGTGTAACTGATGGA R- GCAATTCCAACAAGGCACGGTC	22 22	152
Hilp8A	*hilp8*	F – AGAAGGCGCAAGTTCGAGG R - GGCGCAAGTTATTCACAACAC	19 21	158
Hilp8B	*hilp8*	F – GTGAGCACCGGTTCAAGAAGAG R - GGTGGAGGTGTGCAGTACCAC	22 21	157

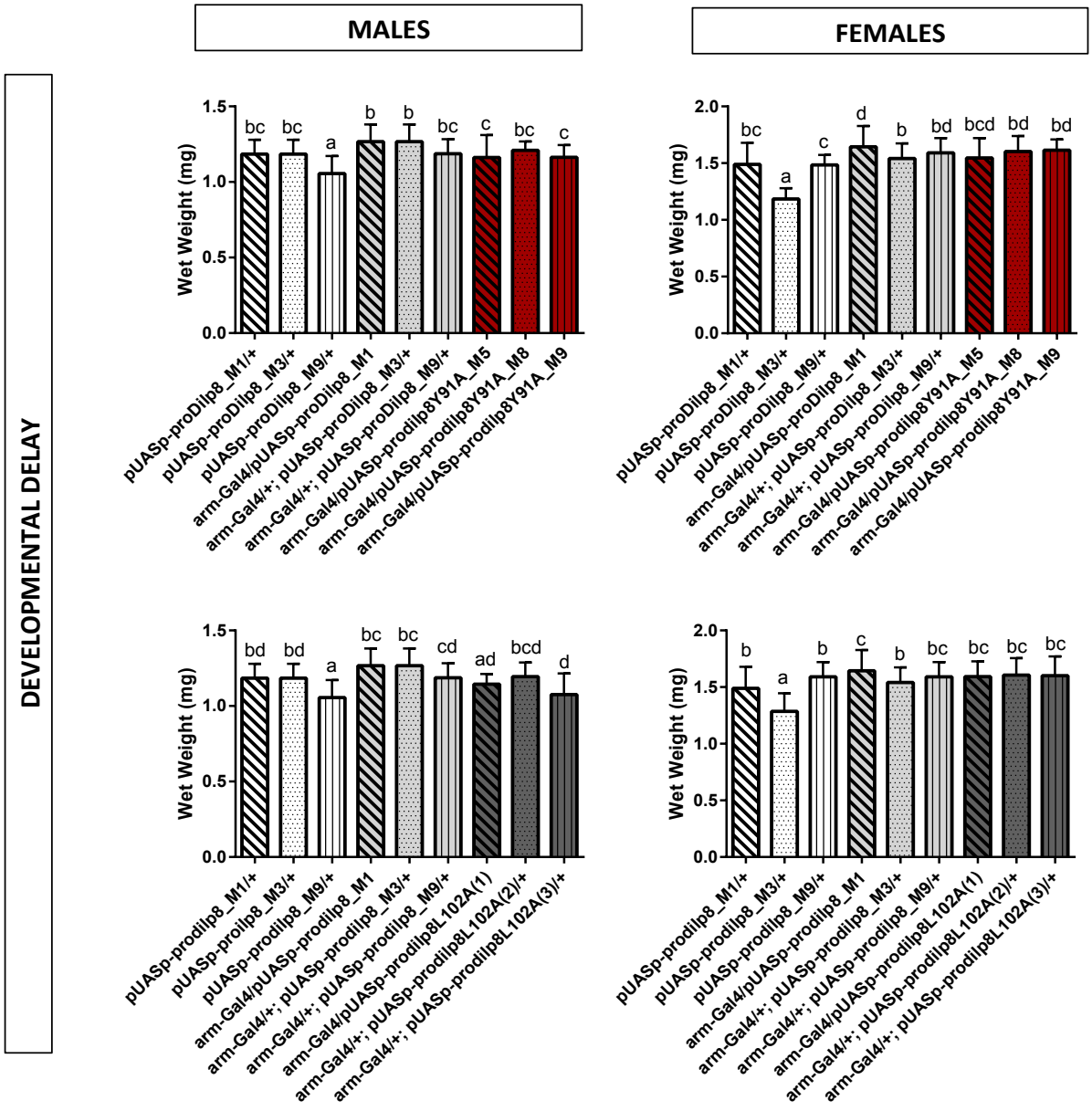
Appendix 6-4 - Standard PCR reaction reagents

Reagent	Concentration	Volume
Green Master Mix (NZYTech)	Enzyme Concentration: 0.2U/μL	
Forward Primer	1 μM	1 μL
Reverse Primer	1 μM	1 μL
RNase-Free Water	Not applicable	2 μL
gDNA	Not applicable	1 μL

Appendix 6-5 – Primer pairs used for the standard site-directed mutagenesis protocol followed for the generation of transgenic flies through P element mediated insertion

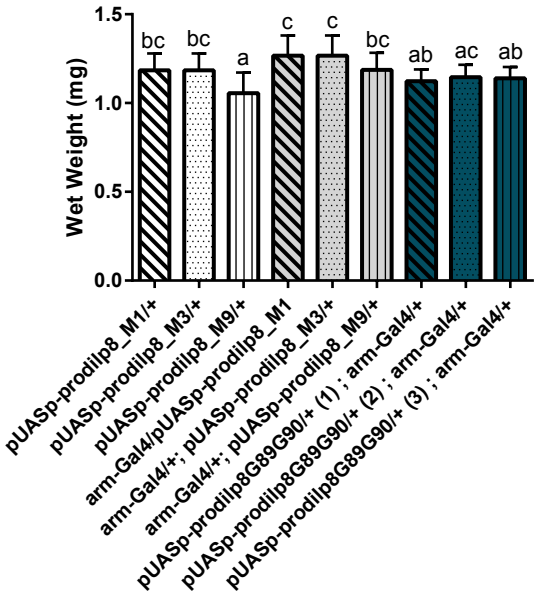
Primer	Primer sequence F	Primer sequence R
L102 point mutation	5'-CACTTCAACCGC*GC*GAGCGAGCTGG-3'	5'-CCAGCTCGCTC*GC*GCGGTTGAAGTG-3'
G89 G90 Deletion	5'-CAGCAGCTACCCCATGTACCTGAAGGT CACCC3'	5'-GGGTGACCTTCAGGTACATGGGGTAG CTGCTG-3'

Appendix 6-6 - Pupal weight (mg) per insertion. Data sharing the same letter are not different at alpha=0.05 (Kruskal-Wallis with Dunn's multiple comparison test).

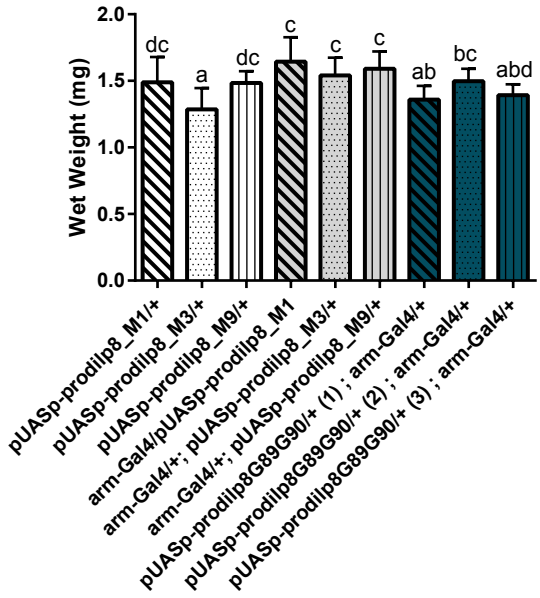


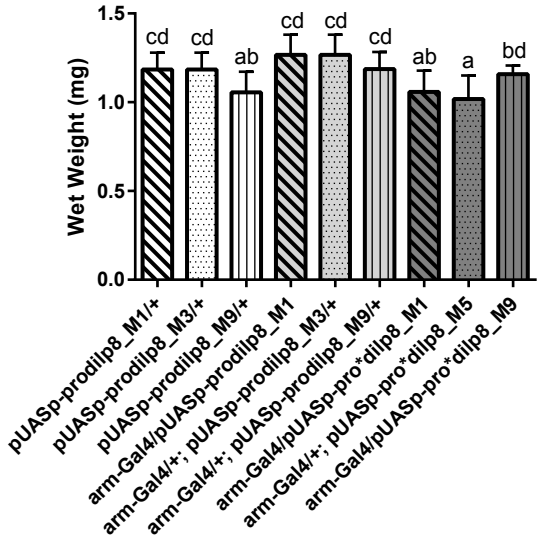
NO DEVELOPMENTAL DELAY

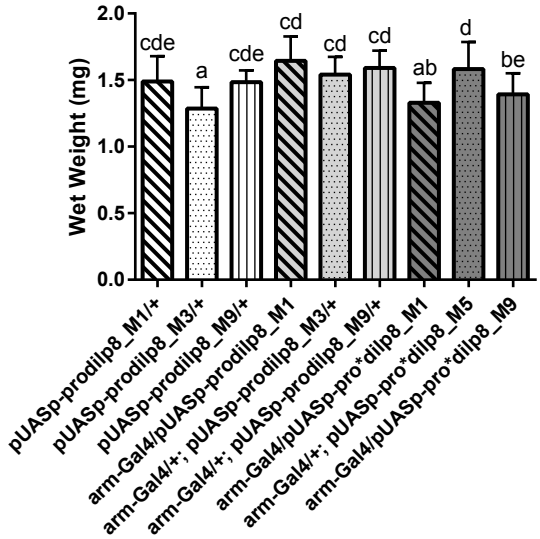
MALES

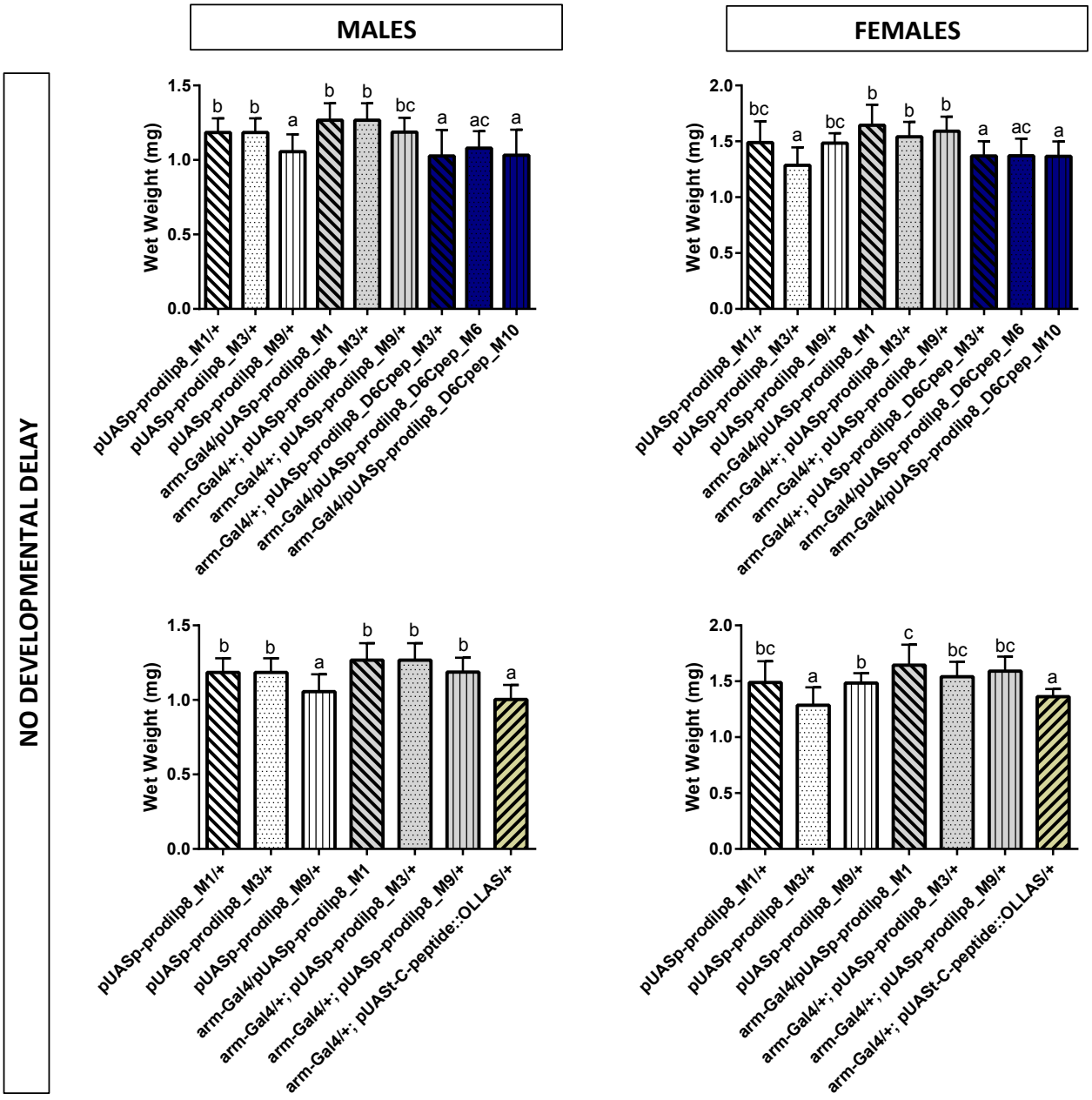


FEMALES









Appenndix 6-7 – RNAseq assembly comparison for Larvae and Adult Female.

Parameter	Larvae	Adult Female
Minimum size	200	200
Maximum size	28,093	12,442
Average	1,078.859	784.276
Total number of *contigs*	58,648	52,820
N *contig* <1000	39,796	39,599
N *contig* ≥ 1000	18,852	13,221