UNIVERSITY OF SÃO PAULO

RIBEIRÃO PRETO MEDICAL SCHOOL

CELLULAR AND MOLECULAR BIOLOGY AND PATHOGENIC BIOAGENTS DEPARMENT

Molecular characterization of the *Drosophila melanogaster* TT (Trachea-Thorax) *cis*-regulatory module

JORGE VICTOR WILFREDO CACHAY WESTER

Ribeirão Preto

2016

UNIVERSITY OF SÃO PAULO

RIBEIRÃO PRETO MEDICAL SCHOOL

CELLULAR AND MOLECULAR BIOLOGY AND PATHOGENIC BIOAGENTS DEPARMENT

Molecular characterization of the *Drosophila melanogaster* TT (Trachea-Thorax) *cis*-regulatory module

Dissertation submitted to the Ribeirão Preto Medical School of São Paulo University to obtain the academic degree of *Master in Science*,

Concentration area: Cellular and Molecular Biology

Student: Jorge Victor Wilfredo Cachay Wester
Advisor: Prof. Dr. Nadia Monesi

Ribeirão Preto

2016

Cachay Wester, Jorge Victor Wilfredo

Molecular characterization of the *Drosophila melanogaster* TT (Trachea-Thorax) *cis*-regulatory module
p.74: il.; 30 cm

Master degree dissertation, presented to the Ribeirão Preto Medical School, University of São Paulo. Concentration area: Cellular and Molecular Biology.

Advisor: Monesi, Nadia.

1. *cis*-regulatory module. 2. DNA Puff. 3. *D. melanogaster*. 4. Tracheal System, 5. Dorsal median muscles

APPROVAL SHEET

Name: Cachay Wester, Jorge Victor Wilfredo

Tittle: Molecular characterization of the *Drosophila melanogaster* TT (Trachea-Thorax) *cis*-regulatory module

Dissertation presented to the Ribeirão Preto Medical School of the São Paulo University to obtain the degree of Master of Science.

Concentration Area: Cellular and molecular Biology

Advisor: Prof. Dr. Nadia Monesi.

Approved at: //2016

EXAMINATION COMMITTEE

Prof. Dr. Nadia Monesi

Institution: Faculty of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo

Evaluation: _____ Signature: _____

Prof. Dr. Ricardo Guelerman Pinheiro Ramos

Institution: Ribeirão Preto Medical School, University of São Paulo

Evaluation: _____ Signature: _____

Prof. Dr. Luiz Paulo Moura Andrioli

Institution: School of Arts, Science and Humanities, University of São Paulo

Evaluation: _____ Signature: _____

Prof. Dr. Claudio Roberto Simon

Institution: Biology department, Federal University of Triângulo Mineiro

Evaluation: _____ Signature: _____

DEDICATION

To Irma Yolanda, for the affection and for believing in me during this journey

To Irma Beatriz, for teaching me the love and giving me the reason to live

To Jose Wilfredo, for trusting in me and supporting me

To María Teresa, for the love, support and for being the special person every day

ACKNOWLEDGMENTS

I am grateful to my advisor Prof. Dr Nadia Monesi, for the opportunity to come to Brazil and for the excellent orientation and support contribution for my scientific development during the realization of the Master. I am also grateful to Dr. Maiaro Cabral and to MSc. Carlos Couto, for the excellent teaching, dedication on the experiments, good friendship and excellent jokes in the lab. I also thanks Marilia, for the technical support, friendship and the great help in take care of my health.

I want to thanks Gabriela from Cellular and Molecular Biology program, for the help and support. I sincerely thank the Ribeirão Preto Medical School and the Cellular and molecular and pathogenic bioagents department for the support during the master studies and the Faculty of Pharmaceutical Sciences, for the infrastructure.

I thank CAPES for the Scholarship and financial support.

I would also like to thank all my friends and colleagues from the lab, Henrique, Vitor, João, Luana, Luiz Miguel, Alexandre, Luciana and Mara, for the good suggestions, language corrections and good times in the lab.

I would like to thank all family in Perú, especially to Carlos Wester, for being a great example of overcoming in life and moral supporting during this journey.

Finally, I really thank *Drosophila melanogaster* for allowing me broaden awareness of Developmental Biology.

.

ABSTRACT

WESTER, J. V. W. C. **Molecular characterization of the *Drosophila melanogaster* TT (Trachea-Torax) *cis*-regulatory module.** 2016. Masters dissertation – Ribeirão Preto Medical School, University of São Paulo.

Previous functional studies identified in the DNA puff *BhC4-1* promoter region a 67 bp (-*253/-187*) *cis*-regulatory module (CRM) that drives reporter gene expression in the ring gland of *D. melanogaster*. A bioinformatics analysis identified 67 *Drosophila melanogaster* sequences that are similar to sequences contained in the ring gland CRM. One of the identified sequences resides in a 657 bp genomic fragment located about 2500 bp upstream *CG13711*, about 400 bp upstream *CG12493*, in a genomic region that constitutes one of the introns of *CG32239* (*Gef64C*). The preliminary characterization of three transgenic lines transformed with a *657 bp-lacZ* construct revealed reporter gene expression in the larval/prepupal tracheal system and in adult thorax. Based on the pattern of expression driven by this CRM we named it Trachea-Thorax (TT). The main goal of this work was to extend the molecular characterization of the lines of the *TT-lacZ* series. Initially β-galactosidase histochemical assays were performed in embryos, first, second and third instar larvae, 0h, 1h and 2h prepupae, 24 h pupae and 1, 3 and 5 days old adults. Reporter gene expression is initially detected during the third larval instar in the tracheal system and continues to be detected in this tissue at 0 h, 1h and 2 h prepupa and, 24 h pupa. During the adult stage, reporter gene expression is verified in the dorsal longitudinal muscles of 3 and 5 days old adults. Since the *TT* CRM lies in an intergenic region and the available information about the nearby *CGs* is still scarce it was not possible to infer which of the *CGs* is regulated by the *TT* CRM. In this context, the mRNA pattern of expression of the *lacZ* reporter gene and of *CG13711*, *CG12493* and *CG32239* was investigated in the tracheal system of both larvae and prepupae and in adult thoraxes of one of the transgenic lines of the *TT-lacZ* series using RT-qPCR. The *lacZ* mRNA expression levels increase about 3 times in 0 h prepupae when compared to the *lacZ* mRNA expression levels present in the tracheal system of third instar larvae. A similar pattern of expression was observed for both *CG32239* and *CG13711*. In three and five days old adult thoraxes *lacZ* mRNA expression levels increase about 37 times and 11 times, respectively, when compared to *lacZ* mRNA expression levels present in one day old thoraxes. In the adult thorax, the only *CG* that presents a similar pattern of expression constitutes *CG12493*. Overall, we conclude that the *TT* CRM drives a dynamic pattern of

expression throughout development. Additionally, based on RT-qPCR results, we suggest that the *TT* CRM regulates the expression of *CG32239* mRNA in the tracheal system during the larvae to prepupae transition, as well as the expression of *CG12493* mRNA in the thorax of 3 and 5 days old adults. Besides extending the functional characterization of a novel CRM our results also contribute new information about the developmental patterns of expression of three *Drosophila melanogaster CGs*.

Key words: *cis*-regulatory module, DNA puff, *D. melanogaster*, tracheal system, dorsal longitudinal muscles.

RESUMO

WESTER, J. V. W. C. **Caracterização molecular do módulo *cis*-regulador TT (Traqueia-Tórax) de *Drosophila melanogaster*.** 2016. Dissertação de Mestrado – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo.

Estudos funcionais anteriores identificaram um módulo *cis*-regulador (MCR) de 67 pb (-253/-187) na região promotora do gene de pufe de DNA *BhC4-1* que dirige a expressão do gene repórter na glândula anelar de *Drosophila melanogaster*. Uma análise bioinformática identificou 67 sequências de *D. melanogaster* que são similares a sequências contidas no MCR de glândula anelar. Uma das sequências identificadas reside em um fragmento genômico de 657 pb localizado aproximadamente 2500 pb à montante do *CG13711*, 400 pb à montante do *CG12493*, em uma região genômica que constitui um dos íntrons do *CG32239* (*Gef64C*). A caracterização preliminar de três linhagens transformadas com a construção *657 pb-lacZ* mostrou expressão do gene repórter no sistema traqueal de larvas e prépupas e no tórax de adultos. Baseado padrão de expressão promovido por este MCR, o mesmo foi denominado *Traqueia-Tórax* (*TT*). O principal objetivo do presente trabalho constituiu estender a caracterização molecular das linhagens da série *TT-lacZ*. Inicialmente embriões, larvas de primeiro, segundo e terceiro estádio, prépupas 0 h, 1 h e 2 h, pupas 24 h e adultos com 1, 3 e 5 dias foram investigados quanto ao padrão de expressão do repórter utilizando ensaio histoquímico que detecta atividade de β-galactosidase. A expressão do gene repórter é inicialmente detectada no sistema traqueal durante o terceiro estádio larval e continua a ser detectada neste tecido em prépupas 0 h, 1 h e 2 h e pupas 24 h. Em adultos, a expressão do gene repórter é verificada nos músculos longitudinais dorsais em adultos de 3 e 5 dias. Uma vez que o MCR *TT* reside em uma região intergênica e a informação disponível sobre os *CGs* próximos ainda é escassa, não foi possível inferir qual dos *CGs* é regulado pelo MCR *TT*. Neste contexto, o padrão de expressão do RNAm do gene repórter *lacZ* e do *CG13711*, *CG12493* e *CG32239* foi investigado no sistema traqueal de larvas e prépupas e no tórax de adultos de uma das linhagens da série TT-lacZ utilizando RT-qPCR. Os níveis de expressão do RNAm *lacZ* aumentam cerca de 3 vezes em prépupas 0 horas, quando comparados com os níveis de expressão do RNAm *lacZ* presentes no sistema traqueal de larvas de terceiro estádio. Um padrão de expressão similar foi observado no caso do *CG32239* e do *CG13711*. Nos tóraxes de adultos de 3 e 5 dias de idade os níveis de expressão do RNAm *lacZ* aumentam cerca de 37 vezes e 11 vezes, respectivamente, quando comparados aos níveis de expressão

do RNAm *lacZ* presentes nos tóraxes de adultos de 1 dia. No tórax de adultos, o único *CG* que apresenta um padrão de expressão similar ao padrão de expressão de *lacZ* constitui o *CG12493*. Em conjunto, nós concluímos que o MCR *TT* promove um padrão dinâmico de expressão durante o desenvolvimento. Além disso, com base nos resultados de RT-qPCR, nós sugerimos que o MCR *TT* regula a expressão do RNAm do *CG32239* no sistema traqueal durante a transição larva-prépupa e também a expressão do RNAm do *CG12493* no tórax de adultos de 3 e 5 dias de idade. Além de estender a caracterização funcional de um novo MCR, nossos resultados também contribuem com novas informações acerca dos padrões de expressão no desenvolvimento de três *CGs* de *D. melanogaster*.

Palavras chave: módulo *cis*-regulador, Pufe de DNA, *D. melanogaster*, sistema traqueal, músculos longitudinais dorsais.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

bp	Base pair
°C	Celsius degrees
DNA	Deoxyribonucleic acid
M	molarity
mM	Mili molarity
pH	Power of hydrogen
PIPES	Piperazine-N,N´-bis (2-ethanesulfonic acid)
EGTA	Ethyleneglycol-bis (2-aminoethylether)-N,N,N´,N´-tetraacetic acid
EDTA	Ethylenediaminetetraacetic acid
µg/mL	Microgram/milliliter
µL	Microliter
RNA	Ribonucleic acid
× *g*	Times gravity
CTAB	Cetyl trimethylammonium bromide
w/v	Weigth/volume
ng/mL	Nanogram/mililiter
mL/cm^2	Milliliter/centimeter square
cDNA	Complementary deoxyribonucleic acid
U	Units
dNTP	Deoxynucleotide phosphate
NCBI	National Center for Biotechnology Information
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
ΔΔCq	Delta delta Cq

TABLE OF CONTENTS

1 INTRODUCTION

1.1. Eukaryotic *cis*-regulatory modules in eukaryotic genomes.

Gene regulatory sequences control the spatial and temporal patterns of gene expression. Regulatory sequences are arranged in units that are named *cis*-regulatory modules (CRMs) (de Leon S. and Davidson E., 2007; Hardison and Taylor, 2012). Every regulatory module receives multiple inputs and processes them in a way that can be mathematically represented as combinations of logic functions like "switch" functions (Davidson E., 2005). A *cis*-regulatory module contains transcription factor binding sites, which enable specific transcription factors to bind the CRMs in order to generate an input. The received input is the regulatory state of the cell and the output is either activation or repression of the gene that is controlled by a given *cis*-regulatory module (de Leon S. and Davidson E., 2007).

The analysis of *cis*-regulatory modules revealed that in bilaterians a gene may present between 5 to 20 *cis*-regulatory modules which can act combinatorially to activate a given gene in a particular time and domain in the organism (Davidson E., 2006). The position of CRMs in relation to the target genes may vary and they can be located from a few kilobases to a hundred kilobases upstream the core promoter. Additionally, CRMs may also reside in the coding sequence, within introns or downstream 3´ UTR sequences (Jeziorska et al., 2009). CRMs are components of gene regulatory networks (Figure 01), which contain the instructions to carry out specific developmental decisions. Many CRMs can act together to contribute to the final transcription rate of a single gene and the interactions between these CRMs modulate the expression of a gene through time and space and also provide a unique code for when, where and at which level a given gene is transcribed (Jeziorska et al., 2009).

Three models have been proposed for the functioning of CRMs based on the way the CRMs interact with the respective target genes. In the first model, named DNA scanning model, transcription factors (TFs) bind to CRMs and move continuously along the DNA sequence until they encounter the promoter region of their target gene (Blackwood et al., 1998). The second model proposes a looping model where CRMs communicate with their target genes through direct protein-protein interactions between the transcription factors bound to the CRM and components of the general transcription machinery assembled at the core promoter, with the intervening DNA looping out (Ptashne et al., 1986; Wang et al., 1988). Finally, the third model integrates elements of the two models above. In this model, activators (ex: transcription factors) and coactivators bind to the CRM sequence and then small loops are formed in the chromatin through which activator-coactivator complexes move

until they encounter their cognate promoter. This model proposes an important role for the coactivator due to its ability to recognize and modify the chromatin structure, which facilitates the movement of the activator-coactivator complex along the chromatin template (Blackwood et al., 1998).

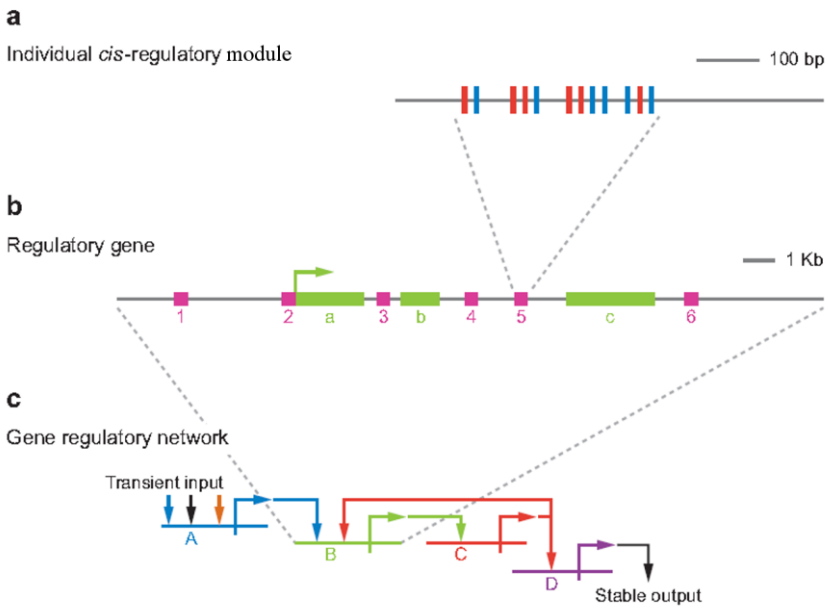

Figure 01. *Cis*-regulatory modules in the gene regulatory hierarchy. (a) Schematic drawing of a *cis*-regulatory module. The red and blue bars represent transcription factor binding sites. **(b)** Schematic drawing of a hypothetical regulatory gene. Exons are represented in green and *cis*-regulatory modules in purple. **(c)** Schematic drawing of a gene regulatory network. The letters A, B, C and D represent regulatory genes. Colored arrows represent the inputs that are processed by the CRMs and the black arrow represents the final stable output that results from the combinatorial activity of the CRMs (modified from Ben-Tabou-de-Leon & Davidson, 2007).

1.2. DNA puffs.

Puffs are the result of unwinding of chromatin and are formed at specific sites in the polytene chromosomes of Diptera. There are two classes of puffs, DNA puffs and RNA puffs. RNA puffs are of general occurrence in Diptera and their formation is related to the transcriptional activity at these *loci* (Ashburner et al., 1974; Daneholt, 1982; Rydlander and Edstrom, 1980; Thummel, 1996). At DNA puff forming sites, which are characteristic of the Sciaridae family, two processes, gene amplification and the abundant transcription of the amplified genes occur in a developmentally regulated manner (Lara, 1991; Stocker, 1996; Monesi, 2009 Simon et al., 2016). Even though the precise mechanism of action of ecdysone

in the regulation of gene amplification and transcription has not been completely elucidated, several studies place this hormone as a regulator of both processes at DNA puff forming sites (Stocker, 1997, Candido-Silva, 2008, Liew et al., 2013, Monesi., 2009, Simon., 2016).

DNA amplification coupled to high levels of transcription of DNA puff genes result in the production of large amounts of DNA puff proteins during a short period of time. The function of DNA puff proteins is partially known and based on their temporal and tissue expression patterns, which are temporally correlated to DNA puff formation, it has been suggested that these proteins are employed to build the cocoon in which the larvae pupate (Winter et al., 1977a,b; Laicine et al., 1984;de-Almeida., 1997, Monesi et al., 2004, Fontes et al., 1999).

1.3. Molecular characterization of the DNA puff C4 *BhC4-1* gene.

In *B. hygida*, the *BhC4-1* gene is amplified 21 times and abundantly expressed in the salivary glands at the end of the fourth larval instar when DNA puff C4 is formed (Paçó-Larson et al., 1992; Monesi et al., 1995). *BhC4-1* expression, *BhC4-1* amplification and DNA puff C4 expansion, are all induced as a late response to the increase in ecdysone titers that trigger metamorphosis (Basso et al., 2002).

The analysis of the deduced amino acid sequence of the *BhC4-1* gene revealed that it encodes a 47,168 Da protein (Monesi et al., 1995). In immunoblots the BhC4-1 protein was detected both in salivary gland extracts and in the saliva of larvae at the age DNA puff C4 is formed. In addition, through immunoelectronmicroscopy it was shown that the BhC4-1 protein is localized in filamentous structures that are part of the saliva (Monesi et al., 2004), which corroborates the idea that DNA puff proteins are employed to build the cocoon in which the larvae pupate.

Recent studies have revealed that at the time DNA puff C4 is formed in the salivary gland, the *BhC4-1* mRNA and the BhC4-1 protein are both expressed in the *B. hygida* prothoracic glands, which is the main source of ecdysone in the larva. In addition, at the time the gene is amplified 21 times in the salivary glands, the *BhC4-1* gene is also amplified at lower levels, 4.8 times, in the prothoracic glands (Candido-Silva et al., 2015). These results confirmed previous studies that have shown that the expression of DNA puff genes is not restricted to the salivary glands (Monesi et al., 2001) and corroborates the suggestion that DNA puff genes might perform additional functions in sciarids, besides being employed to

build the cocoon. Furthermore, the demonstration that developmentally regulated gene amplification also occurs in a second tissue of *B. hygida* suggest that gene amplification is a mechanism that might be more widely employed in insects than previously described (Candido Silva et al., 2015).

1.4. Characterization of the *BhC4-1* gene in transgenic *Drosophila.*

The functional characterization of the *BhC4-1* gene was performed in transgenic *Drosophila*. Initial studies revealed that the *BhC4-1* gene is expressed in a developmentally regulated manner in the prepupal salivary glands of transgenic *Drosophila*, suggesting that *cis*-regulatory elements of *B. hygida* are recognized by *Drosophila trans*-activating factors (Monesi et al 1998). Additionally, *in vitro* studies have shown that in transgenic *Drosophila* salivary glands the *BhC4-1* expression is also induced as a late response to ecdysone, similar to what occurs in *B. hygida*. (Basso et al., 2002). Together these results revealed that the mechanisms that regulate *BhC4-1* expression in the salivary gland are conserved in the heterologous system.

The characterization of the mechanisms that regulate *BhC4-1* expression were extended through functional studies in transgenic *Drosophila*, and led to the identification of *cis*-regulatory modules in the *BhC4-1* promoter region (Figure 02) (Monesi et al., 2003). One of the identified *cis*-regulatory modules is a 97 bp *(-57/+40)* fragment that constitutes the *BhC4-1* core promoter. In the absence of upstream regulatory sequences the *(-57/+40)* fragment drives low levels of *BhC4-1-lacZ* mRNA expression, which is consistent with the activity of a core promoter. On the other hand, when the *(-57/+40)* fragment was placed downstream distinct *cis*-regulatory modules, regulated expression driven by the upstream sequences was verified, confirming that the *BhC4-1* core promoter does not contain other *cis*-regulatory modules beside those characteristic of core promoters (Garcia et al., 2011).

The second identified *cis*-regulatory module is located in a 129 bp *(-186/-58)* fragment that drives *BhC4-1-lacZ* expression in prepupal salivary glands (Monesi et al., 2003; Lecci et al., 2008). Results from bioinformatics analyses revealed that the *Drosophila melanogaster* genome contains 162 sequences similar to sequences contained in the *BhC4-1* proximal promoter *(-253/-58)*. Ninety-five out of the 162 identified sequences align to the 129 bp *(-186/-58)* salivary gland *cis*-regulatory module (Lecci et al., 2008). Interestingly, four of the 95 identified genomic regions contain sequences similar to those located in the *(-79/-58)* region

of the 129 bp *cis*-regulatory module and are located in regulatory regions of genes that are expressed in the salivary gland in response to the increase in ecdysone titers that trigger metamorphosis (Lecci et al., 2008).

Functional assays revealed that the product of the early genes *BR-C* (Z3 isoform), *E74* (A and B isoforms) and *E75* (A isoform) participate in the induction of *BhC4-1-lacZ* in the salivary gland during the larvae to prepupae transition. While the BR-C Z3 isoform plays an essential role in the *BhC4-1-lacZ* expression in the prepupal salivary glands, the isoforms E74A, E74B and E75A are all necessary to promote high levels of *BhC4-1-lacZ* expression in this tissue (Basso et al., 2006). Genetic interaction experiments revealed that the overexpression of the A isoform of *E74* is sufficient to induce *BhC4-1-lacZ* expression in lines transformed with the 129 bp (*-186/-58*) *cis*-regulatory module during the third larval instar, at a developmental time when *BhC4-1-lacZ* is not expressed in the salivary gland (Sanchez and Monesi, unpublished data). More recently we showed that the 129 bp *(-186/-58) cis*-regulatory module directly binds the ETS domain of E74A in the context of the one-hybrid system (Frank H., 2014). Together, these results further confirm that the expression of the *BhC4-1* DNA puff gene in the salivary gland is regulated by the products of the early genes and that at least in the case of E74A the heterologous transcription factor directly binds to the DNA puff promoter regulatory sequences.

The third region identified in the *BhC4-1* proximal promoter is the 67 bp *(-253/-187) cis*-regulatory module, which drives *BhC4-1-lacZ* expression in the ring gland. In this case, transgene expression is initially detected in the ring gland of late embryos and continues throughout the larval and prepupal stages (Monesi et al., 2003; Lecci et al., 2008). In order to extend the characterization of the 67 bp *(-253/-187) cis*-regulatory module, five constructs (*MUT1-MUT5*) containing linker scanning mutations were cloned upstream of the reporter gene *lacZ* and the resulting constructs were employed to transform *Drosophila* (Malta et al., 2006; Furtado I., 2013). The analysis of these transgene lines further refined the characterization of the 67 bp fragment and resulted in the identification of different regions in this *cis*-regulatory module. Specifically, mutations introduced in both *MUT1* and *MUT2* constructs (*-246/-225*) did not alter the pattern of expression driven by the 67 bp *cis*-regulatory module. On the other hand, mutations introduced in the *MUT3*, *MUT4* and *MUT5* constructs strongly affected the activity of the ring gland *cis*-regulatory module indicating that the (*-222/-189*) region contains transcription factor binding sites that are essential for *BhC4-1* expression in the ring gland (Furtado I., 2013).

Two additional *cis*-regulatory modules, which repress *BhC4-1-lacZ* expression in the ring gland, were also identified in the *BhC4-1* promoter region (Monesi et al., 2003). The first one is located in the *(-1293/-254)* region and restricts the onset of *BhC4-1-lacZ* expression in the ring gland to the second larval instar. The second one *(-3314/-1294)*, restricts the onset of *BhC4-1-lacZ* expression in the ring gland to 0 h prepupae (Monesi et al., 2003). The characterization of the proximal repressor module was extended through the analysis of lines transformed with two different constructs, *(-824/-187)* and *(-488/-187)*. In these lines, *BhC4-1-lacZ* ring gland expression is initially detected in second instar larvae, indicating that the elements that repress *BhC4-1-lacZ* expression at developmental times prior do the second larval instar are contained in the *(-488/-254)* fragment. In addition, the results obtained with the *(-488/-187)* lines also reveal that the presence of both the repressor module *(-488/-254)* and the activator module *(-253/-187)* is sufficient to drive gene expression in the ring gland of second instar larvae (Campos Neves, 2009).

The *Drosophila* ring gland is the result of the fusion of the prothoracic glands (the source of ecdysone), the *corpus allatum* (the source of juvenile hormone), and the *corpus cardiacum* (the source of peptide hormones that participate in the control of energy metabolism) (King et al., 1966). In *Bradysia hygida*, the neuroendocrine system comprises a pair of prothoracic glands and a single *corpus allatum* fused to a single *corpus cardiacum* that are located under the retrocerebral aorta (Candido-Silva et al., 2015). Both *BhC4-1* mRNA and BhC4-1 protein are expressed in the prothoracic glands of *B. hygida* at the time DNA puff C4 is formed (Candido-Silva et al., 2015). These results suggest that *BhC4-1-lacZ* expression in the ring gland of lines that contain the ring gland *cis*-regulatory module *(-253/-187)* is also due to the conservation of *cis*-regulatory sequences that are being recognized by *Drosophila* ring gland *trans*-activating factors and reveal that the degree of conservation of the mechanisms that regulate the *BhC4-1* DNA puff gene is higher than previously envisaged (Candido-Silva et al., 2015).

As previously mentioned, a bioinformatics analysis resulted in the identification of 162 *D. melanogaster* sequences that present similarity to the *BhC4-1* proximal promoter *(-253/-58)*. Out of these, 67 sequences align to the 67 bp *(-253/-187) cis*-regulatory module and 91% of these sequences are identical to sequences contained in this *cis*-regulatory module. The length of the identified sequences varies between 15 and 24 nucleotides. Apparently, the identified *D. melanogaster* sequences are not associated to regulatory regions of genes that are expressed in the ring gland (Lecci et al., 2008).

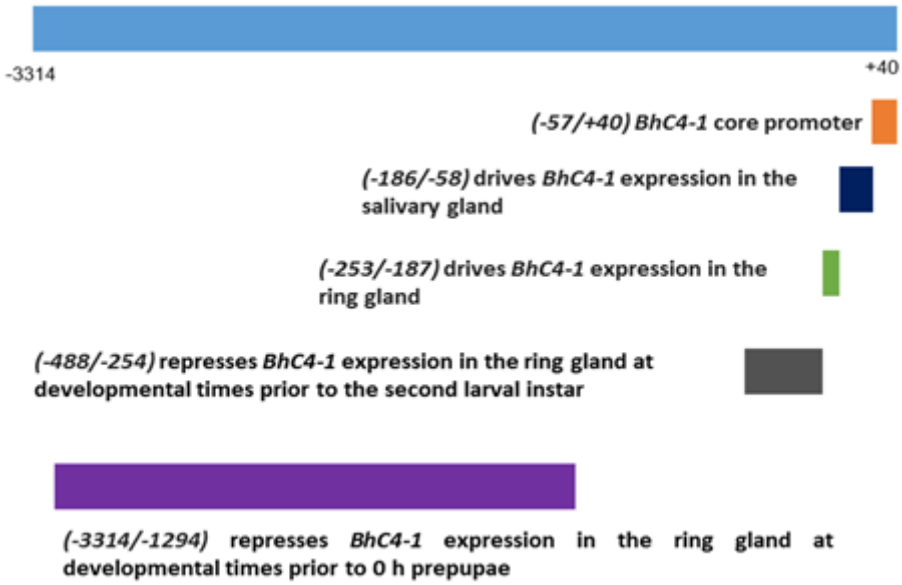

Figure 02. ***Cis*-regulatory modules identified in the *BhC4-1* promoter region *(-3314/+40)*.** In orange the *(-57/+40) BhC4-1* core promoter. In blue, the 129 bp *(-186/-58)* salivary gland *cis*-regulatory module. In green, the 67 *(-253/-187)* ring gland *cis*-regulatory module. Two fragments *(-488/-254)* and *(-3314/-1294)* in grey and purple respectively, represent regions that act as repressors of *BhC4-1* expression in the ring gland at different developmental times.

In order to identify putative ring gland *cis*-regulatory modules in the *D. melanogaster* genome, the 67 *D. melanogaster* identified sequences were further analyzed employing the following criteria: a) the degree of conservation of the identified sequences amongst the 12 *Drosophila* species that had been sequenced, sequences that were conserved in at least 7 *Drosophila* species were selected; b) their position in relation to coding sequences, sequences located upstream coding regions were selected; c) investigation of the data available, sequences associated to genes which could be expressed in the ring gland of embryos were selected. This screening resulted in the selection of 10 sequences (Zampar et al., 2009). Five of these sequences were selected for functional assays and genomic fragments of about 600 bp containing the selected sequences were cloned upstream the *lacZ* reporter gene in the pBP vector (Antoniewski et al., 1996).

Out of the five tested sequences, only one drove detectable levels of reporter gene expression in transgenic animals. This sequence is a 657 bp fragment, located in an intergenic region. This 657 bp fragment is located about 2500 bp upstream *CG13711*, about 400 bp upstream *CG12493*, in a genomic region that constitutes one of the introns of *CG32239* (*Gef64C*) (Figure 3). The analysis of transgenic lines revealed that the 657 bp fragment

contains *cis*-regulatory elements that drive a dynamic pattern of gene expression throughout development. In both third instar larvae and prepupae of lines transformed with this transgene, both *lacZ* mRNA expression and β-galactosidase activity were detected in branches of the tracheal system. Additionally, reporter gene expression was detected in the thoracic muscles of adults. Together these results indicated that the 657 bp fragment contains *cis*-regulatory elements that drive gene expression both in the larval/prepupal tracheal system and in the thoracic muscles of adults (Zampar et al 2009). Due to the observed pattern of expression we named the 657 bp *cis*-regulatory module *Trachea-Thorax* (*TT*).

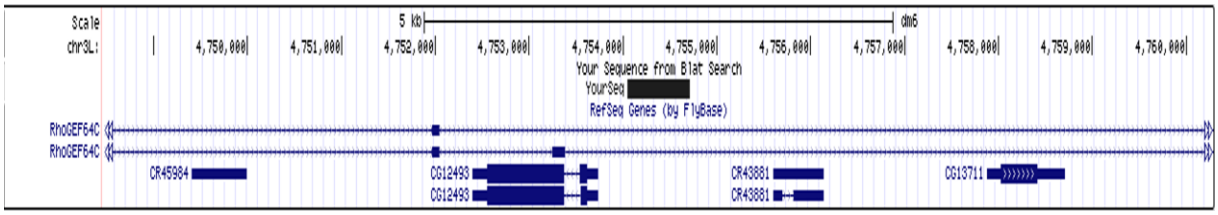

Figure 03. **Location of the *TT cis*-regulatory module in the *D. melanogaster* genome.** On top of the figure are shown the chromosome 3L coordinates (version of August 2014; BDGP R6/dm6). The black rectangle (*your seq*) shows the genomic location of the *TT cis*-regulatory module in the *D. melanogaster* genome. The TT *cis*-regulatory module is located in one of the introns of *CG32239* and upstream of *CG13711* and *CG12493*, which are divergently transcribed with respect to the *TT cis*-regulatory module. Figure obtained from the website http://genome.ucsc.edu.

1.5. Molecular characterization of the *CGs* located near the TT *cis*-regulatory module.

1.5.1. *CG32239*

The *CG32239*, *RhoGEF64C*, is located in chromosome 3L of *Drosophila* in the 4,692,783-4,796,253 region. This gene encodes a guanine nucleotide exchange factor (GEF), which act as GTPase activator by releasing GDP and allowing its replacement by GTP in small GTPases of the Rho subfamily (Simoes et al., 2006). *CG32239* encodes two isoforms, Gef64C_A isoform and Gef64C_B isoform, which are 1985 aminoacids long (Bashaw et al., 2001).

Analysis of RNA-seq data revealed that *RhoGEF64C* presents moderately high levels of expression in the carcass of L3 larvae. *RhoGEF64C* moderate expression levels also occur in 12 to 20 hours embryos, third instar larvae, 12-24 hours prepupae, 3 days pupae, and moderately high expression in 2 days pupae (Figure 04) (Graveley et al., 2010; 2011). *In situ* hybridization experiments revealed that in embryos stage 11-12 *RhoGEF64C* is expressed in

the digestive system, spiracles, Malpighian tubules and tip cell, whereas in embryos stage 13-16 *RhoGEF64C* is expressed in the digestive system, epidermis, Malpighian tubules, tip cell, anal pad and atrium (available online at BDGP, *http://www.fruitfly.org/*).

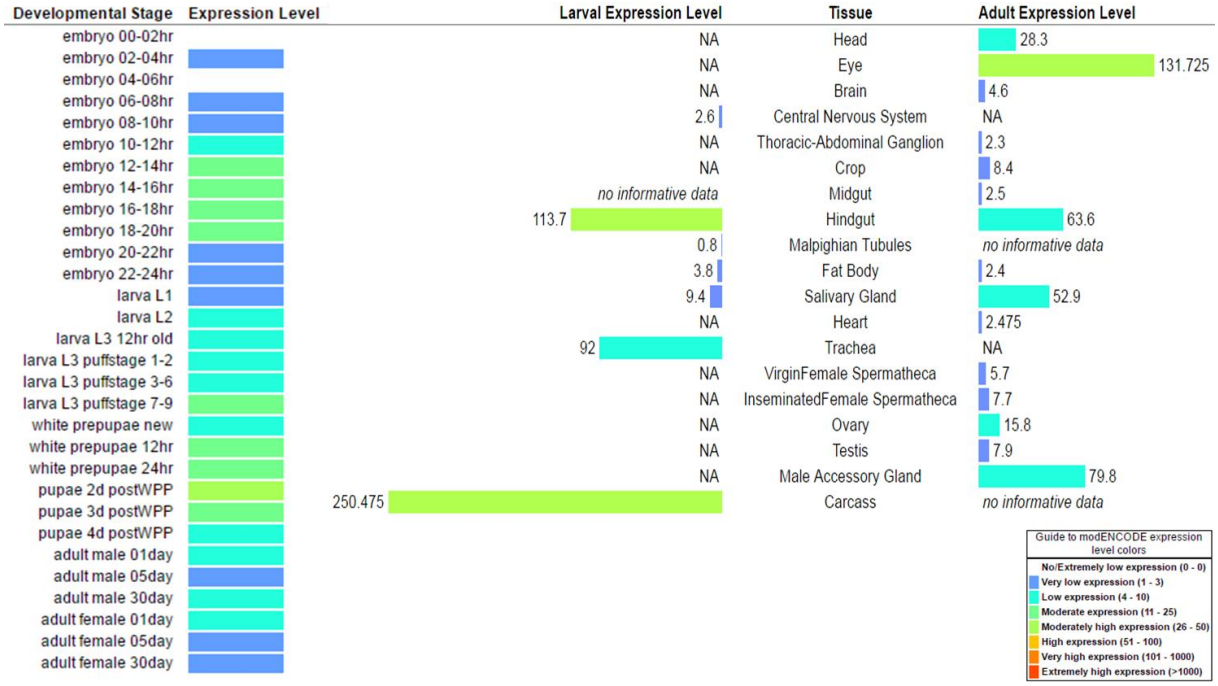

Figure 04. Temporal expression pattern of *CG32239* (*Gef64C*). The figure shows a heatmap of the pattern of *CG32239* expression obtained by RNA-seq analysis at different developmental stages and tissues of *Drosophila melanogaster* (Graveley et al., 2010; 2011; available online at *http://www.flybase.org*).

CG32239 (*Gef64C*) was identified in a genetic screen designed to identify molecules involved in controlling axon guidance decisions. Loss of function experiments revealed that the absence of *CG32239* (*Gef64C*) did not cause strong axon guidance defects. On the other hand, overexpression of *CG32239* (*Gef64C*) results in a phenotype in which many axons abnormally project across the midline indicating that *CG32239* (*Gef64C*) promotes axon attraction to the CNS midline (Bashaw et al., 2001).

Rho1 activity has been shown to be essential for epithelial invagination and tubulogensis and to control the assembly of apical actin during spiracle formation (Simoes et al., 2006, Lovegrove, 2006). Further investigations revealed that Rho1 activity is apically restricted because its regulators, *RhoGEF2* and *CG32239* (*Gef64C*) are enriched in the apical region of the invaginating spiracle cells whereas its inhibitor *RhoGAP Cv-c* occupies the basolateral domain (Simoes et al., 2006). *RhoGEF2*, *CG32239* (*Gef64C*), and Rho1 also participate in the mechanism that restricts the actin nucleation induced by Dia (*Diaphanous)*

to the apical cells of apical region in several tubular structures, including the tracheal network, salivary glands, hindgut and Malpighian tubules (Massarwa et al., 2009). The apical actin cables formed by Dia also seem to have a role in the trafficking of secretory vesicles to the apical surface of tubular organs, a process that employs the motor protein MyoV (Massarwa et al., 2009).

More recently, *CG32239* (*Gef64C*) expression was shown to be restricted to all presumptive leg joints during leg development, indicating that *CG32239* (*Gef64C*) also participates in the regulation of joint morphogenesis during leg development (Greenberg and Hatini 2010). Overall, these studies reveal that *CG32239* (*Gef64C*) participate in the regulation of different cellular mechanisms, in distinct tissues and at different developmental stages.

1.5.2. *CG13711*

CG13711 is located in the 4757883-4758708 region of the 3L chromosome and is entirely located in one of the introns of *CG32239*. This gene presents a single isoform *CG13711*, which encodes a 127 amino acids deduced protein with a molecular mass of 15,049 Da (Bonneaud et al., 2003). *CG13711* is expressed only in the anterior domain in cellularized blastoderm embryos, and accordingly, RNAseq analysis revealed moderately high expression levels of *CG1377* in 2-4 hours old embryos (Figure 05) (Graveley et al., 2010; 2011).

CG13711 is one of the seven members of a protein family that comprises 7 basic proteins that are zigotically expressed in distinct domains during the blastoderm stages and includes *CG13716*, *CG13713*, *CG15876*, *CG13712*, *CG14112* (*SoxNeuro Co-Factor*, *SNCF*) and *CG7428* (*halo*) (Bonneaud et al., 2003; Gross et al 2003). Two of these CGs, *CG14112* (*SoxNeuro Co-Factor*, *SNCF*) and *CG7428* (*halo*) have been further characterized.

SNFC is a 146 amino acids protein encoded by *CG14112* and *SNCF* mRNA is detected in stage 4 embryos (syncytial blastoderms) and stage 5 embryos (cellularized blastoderm). SNCF has been shown to interact with the *Drosophila* SoxN transcription factor (Consortium 2003; Bonneaud et al., 2003). This transcription factor is a member of the Sox protein family that plays multiple important roles in the development of the central nervous system (Bonneaud et al., 2003). SoxN and SNCF are co-expressed only during a short period of time in the neuroectoderm before and shortly after gastrulation (Bonneaud et al., 2003). Since bioinformatics analyses indicated that *CG13711* present similarity to *SNCF* and both

are expressed at the beginning of embryogenesis, it has been speculated that *CG13711* might also act in the modulation of early expressed transcription factors such as SoxN (Bonneaud et al., 2003).

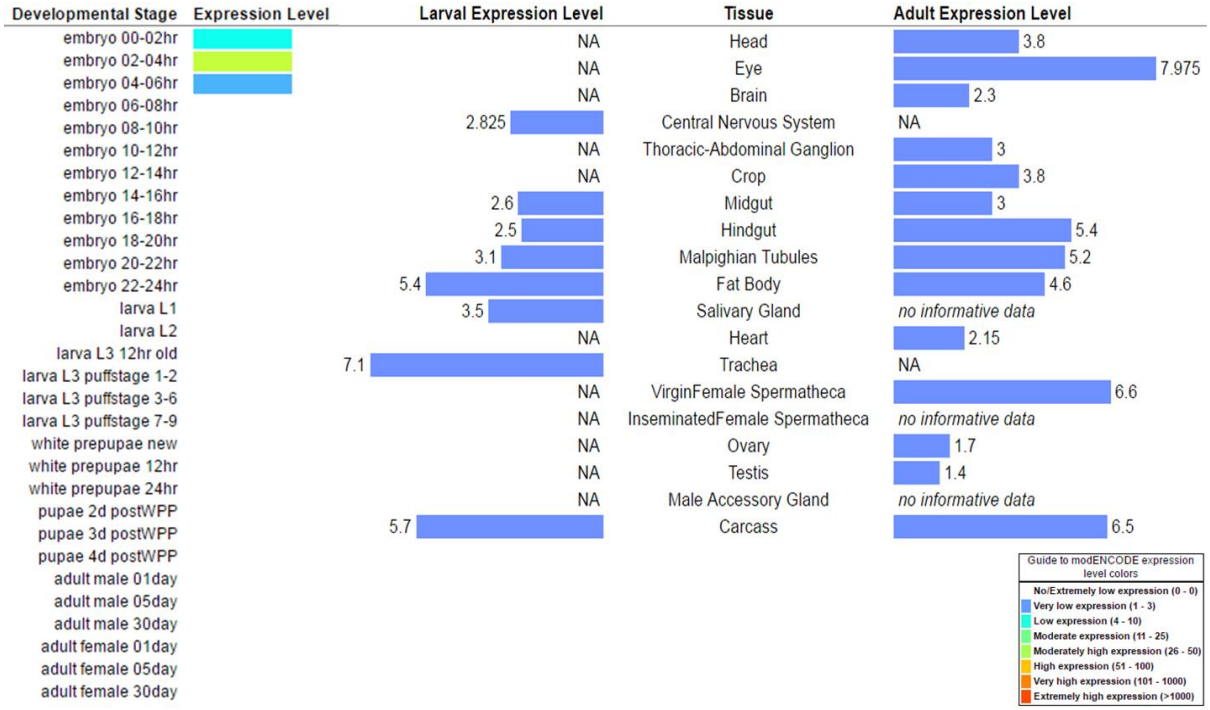

Figure 05. Temporal expression pattern of *CG13711*. The figure shows a heatmap of the pattern of *CG13711* expression obtained by RNA-seq analysis at different developmental stages and tissues of *Drosophila melanogaster* (Graveley et al., 2010; 2011; available online at *http://www.flybase.org*).

Halo is a small 150 amino acids protein encoded by *CG7428 (halo)* that present a dynamic pattern of expression during the syncytial blastoderm and blastoderm cellularization (Gross et al., 2003). Halo is a transporter of lipid droplets. Embryos carrying mutations in *Halo* (Δ*halo)* failed to clear during the blastoderm cellularization stage. Nevertheless, Δ*halo* yields viable and fertile adults (Gross et al., 2003). Even though the exact molecular function of *CG13711* is still unknown, two independent research groups have identified the CG13711 protein as a member of a protein family that is apparently unique of *Drosophila* and that comprises proteins that present diverse functions such as *SNCF* and *Halo* (Bonneaud et al., 2003; Gross et al., 2003).

1.5.3. *CG12493*

CG12493 is located in chromosome 3L in the 4752393-4753738 region. This gene encodes two isoforms: CG12493_A is a 296 amino acids protein of 32,536 Da and CG12493_B is a 293 amino acids protein of 32,261 Da. The available molecular data suggests that *CG12493* encodes a double-stranded RNA binding nuclear protein (Lakso 2000; Saunders and Barber 2003; Lundgren et al., 2005; Gerbasi et al., 2010).

RNA-seq analyses revealed moderate to moderately high levels of *CG12493* expression from the end of the third larval instar until 30 days old adults. The highest levels of expression during the adult stage are found in 5 days males. Moderate levels of expression are found in the fat body of white prepupae and in the digestive system of 1-4 day old adults. High levels of expression were observed in the testis of mated four days old males and in the Malpighian tubules, whereas very high levels of expression were detected in third instar larvae imaginal discs (Figure 06 and data not shown) (Graveley et al., 2010; 2011; available online at *http://www.flybase.org*). *In situ* hybridization experiments revealed *CG12493* expression in the clypeal primordium-labral in 11-12 stage embryos and in the embryonic gonads of embryos stage 13-16 (available online at BDGP, *http://www.fruitfly.org/*).

The *CG12493* deduced amino acid sequence contains a double-stranded RNA binding domain (DRBD). The recognition of dsRNAs by DRBDs involves the recognition of structural motifs which allows them to bind nonspecifically to dsRNAs and ssRNAs with extensive secondary structures (Saunders and Barber 2003). *CG12493* has been described as a paralogue of *blanks* (*CG10630*) (Gerbasi et al., 2010). *blanks* encodes a nuclear protein of 324 amino acids with a molecular mass of 36,500 Da. CG12493 and Blanks share 55% amino acid identity and 67% similarity (Gerbasi et al., 2010). *blanks* was identified in a screen for siRNA duplex binding proteins and was shown to be part of a nuclear dsRNA binding complex that is distinct from previously described RNA induced silencing complexes (RISCs). In *Drosophila*, Blanks participate in sperm differentiation and loss of Blanks causes male sterility, but not female sterility, coupled to up-regulation of genes related to innate immunity pathways (Gerbasi et al., 2010).

The localization and pattern of expression of *CG12493* in male testes is similar to the one observed for Blanks. Interestingly, in Blanks mutants *CG12493* mRNA levels of expression do not change, but the levels of CG12493 protein expression are decreased (Gerbasi et al., 2010). In contrast to Blanks mutant, *CG12493* homozygous mutants are viable

and fertile in spite of presenting a smaller than normal seminal vesicle and reduced mature sperm numbers. In *CG12493* mutants both Blanks mRNA and protein expression are unaltered. Together, these data suggest that CG12493 and Blanks proteins might directly interact in the adult male testes, that this association is necessary for CG12493 protein stability and that the CG12493 protein could perform a supplementary role in sperm differentiation (Gerbasi et al., 2010).

Figure 06. Temporal expression pattern of *CG12493*. The figure shows a heatmap of the pattern of *CG12493* expression obtained by RNA-seq analysis at different developmental stages and tissued of *Drosophila melanogaster* (Graveley et al., 2010; 2011; available online at *http://www.flybase.org*.

1.6. The Tracheal system of *Drosophila melanogaster.*

The tracheal system is a complex network of hollow epithelial tubes, which delivers oxygen directly to the tissues. Oxygen enters into the tracheal system through the spiracles and diffuses along the major branches until it reaches the fine terminal branches, named tracheoles, that end blindly on the surfaces of all tissues (Samakovlis et al., 1996). The tracheal epithelium is a monolayer which does not present a supporting tissue but is prevented from collapsing due to the presence of a stiff extracellular cuticle that lines its apical (luminal) surface (Samakovlis et al., 1996).

The tracheal system arises from 10 clusters of ectodermal cells located on each side of the embryo, one in each hemisegment, from T2 through A8 (Samakovlis et al., 1996). These

clusters of cells invaginate and form elongated sacs of cells that remain connected to the surface by the spiracular branch, a thin stalk. The primary branches appear at stage 12 and consist of six major buds that form along the sac and grow to characteristic lengths and in different directions. Cells that are not part of the primary branches remain and will give rise to the transverse connective. Secondary branches are originated from primary branches at stage 15 and the first terminal branches form at stage 16 as continuations of secondary branches. The terminal branches ramify into extensive arrays of terminal branches (tracheoles) that cover the target tissues (Smakovlis, 1996).

Not all primary and secondary branches complete the branching sequence. The branches that stop branching grow towards tracheal branches from neighboring hemisegments and interconnect the tracheal network. The dorsal trunk is formed at stage 14 and secondary branches fuse with their partners to connect the lateral trunk and to form a dorsal anastomosis that connects the right and left sides of the tracheal network. Overall the branching process takes about 10 hours and terminal branching continues during the larval period. The tracheal system becomes functional, filled with air, about two hours before hatching.

Groups of genes that act at different time points during the branching process were identified from an enhancer trap screening. The *breathless* (*Tracheal-2*) gene is required for primary branch formation and to stimulate expression of *pantip* markers. *pointed* (*Pantip-1*) is required for secondary branch formation and activates expression of terminal markers. *Terminal-1* is required for terminal branch formation. The results derived from the global analyses of the patterns of expression of the genes identified in this screening led the authors to propose that many tracheal genes are organized into a regulatory hierarchy. In this hierarchy the genes expressed at each level of branching serve two functions: a) they control their level of branching and b) they regulate the expression of genes involved in the next level of branching. Through this regulation the different branching events are coordinated and the different steps of branching occur in a proper sequence (Samakovlis et al., 1996).

During the larval stage the tracheal system topology is essentially the same in all three instars. The modifications that occur during the larval stage are related to an increase in the number of trachea and tracheoles in the consecutive stages. The first instar system is metapneustic, whereas the second and third instar systems are amphipneustic and the spiracular tracheae of the non-functional spiracles two to nine are collapsed to form solid spiracular cords. The larvae present a pair of dorsal trunks connected by ten dorsal

anastomoses distributed from the anterior to the posterior region. Additionally, the larval tracheal system presents a pair of lateral longitudinal trunks, which are connected by eight transverse connectives to the dorsal trunks. A dorsal cervical trachea arises from each dorsal trunk and at the ventral side, each ventral cervical trachea gives rise to a medially directed trachea that branch to the central nervous system (Whitten 1957) (Figure 07).

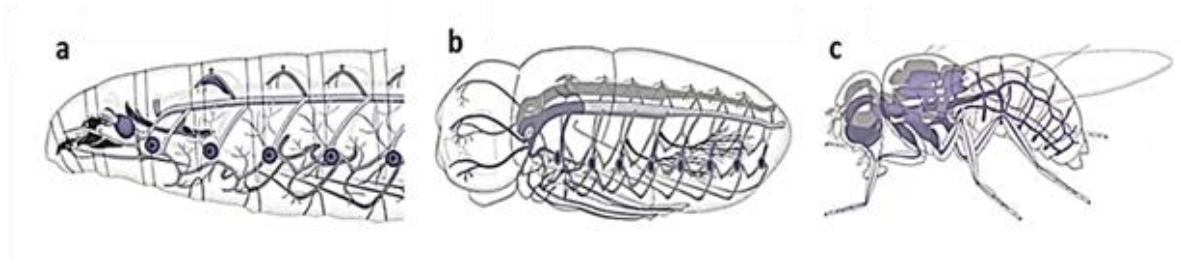

Figure 07. Postembryonic tracheal system in *Drosophila melanogaster*. (a) Represents the anterior side of the larval tracheal system that consists of two spiracles which are continuous to the dorsal trunks. Anastomosis arise dorsally from the dorsal trunks and the transverse connectives and internal branches arise ventrally to the dorsal trunks. **(b)** Represents the pupal tracheal system. Two pairs of trachea arise from the prothoracic spiracle and are located in the head region. Dorsal trunks, anastomosis, transverse connectives and internal branches are located both in the thoracic and abdominal regions. **(c)** Represents the de adult tracheal system. The tracheal system comprises a complex tubular network and air sacs distributed in the head and the thorax, while the abdominal region presents tracheal ramifications that arise from the anterior abdominal region. (Atlas of *Drosophila melanogaster* development).

The pupal tracheal system is distinct from the larval in spite of the fact that it also presents a metameric pattern in the thoracic and anterior abdominal regions. Only the two anterior spiracles are functional. From these prothoracic spiracles originates two pairs of trachea that tracheate the head ventrally and dorsally. The dorsal pair is connected by a transverse anastomosis and sends branches to the developing eyes and to the brain, whereas the ventral branches tracheate the head ventral region. The ventral region is tracheated by numerous ramifications of two pairs of trachea that arise from the anterior abdominal region. The spiracular trachea continues posteriorly as a longitudinal dorsal tracheal trunk from which branches another tracheal longitudinal trunk. Both trunks reach the anterior abdominal region. The two pairs of longitudinal main trunks are connected dorsally by five dorsal anastomoses.

The second pair of anastomosis gives rise to trachea and tracheoles that supply the developing thoracic flight muscles. Laterally, each pair of longitudinal trunks is connected by four trachea. From the lateral trunks arises the ventral tracheas that tracheates the first pair of legs, the nervous system and finally unites with its pairs in the mid-ventral line. The wing and haltere are tracheated by tracheas that originate from the lateral longitudinal trunks. The

posterior abdominal region presents two pairs of branched tracheas that arise from the third and fourth transverse connectives (Whitten, 1957) (Figure 07).

The adult fly presents nine pairs of functional spiracles distributed to each side of the thorax and the anterior abdominal region. In the head there is a pair of dorsal and a pair of ventral air sacs. The dorsal pair of air sacs supplies the brain and the eyes. The paired ventral air sacs pass into a trachea that extends the length of the proboscis and supplies branches to various structures. In the abdomen there is a definite pattern of trachea that includes two pairs of longitudinal trunks which are laterally connected by transverse connectives. The dorsal trunks are connected by nine dorsal anastomoses. The dorsal trunk initiates in the first thoracic spiracular trachea and its diameter is very broad until it reaches the region of the fifth pair of spiracles from where it becomes narrower. Trachea and tracheoles that arise from this trunk pass to the dorsal longitudinal muscles. The second pair of longitudinal trunks joins the dorsal trunk immediately behind the first thoracic spiracle and in the thoracic region it expands into a series of air sacs. In the region of the third spiracle it becomes tracheated and runs in this form to the region of the ninth pair of spiracles (Whitten, 1957) (Figure 07).

1.7. Development of indirect flight muscles of *Drosophila.*

The adult thorax of *D. melanogaster* contains several muscles that are attached to the thoracic exoskeleton and play an important role in the adult life. The adult thoracic musculature can be divided into two types: tubular muscles and fibrillar muscles. The tubular musculature includes the leg, jump and direct flight muscles (DFMs), whereas the indirect flight muscles (IFMs) are fibrillar muscles (Berstein et al., 1993; Allen and Hodenshwege, 2006). The tubular musculature presents a characteristic organization, which consists of rectangular myofibrils that are radially oriented surrounding a centrally located nucleus. Ultrastructural observations revealed that the tubular muscles of the thoracic region present few mitochondria and abundant sarcoplasmic reticulum. (Allen and Godenshwege, 2006).

The fibrillar IFMs include a set of twelve fibers that are oriented dorsal longitudinally, named dorsal longitudinal muscles (DLMs), and a second set of twelve large and two small fibers that are oriented dorsal ventrally, named dorsal ventral muscles (DVMs) (Figure 07). The fibrillar muscles are circular and contain high numbers of mitochondria. Wing myoblast specification as a dorsal lineage occurs as early as the first larval instar (Lawrence, 1982) and these myoblasts give rise to the IFMs (DLMs, DVMs) and the DFMs (direct flight muscles)

during metamorphosis. The adult DLMs are originated from the larval oblique muscles (LOMs) (Fernandes et al., 1991; Fernandes and VijayRaghavan, 1993; Fernandes et al., 1996; Roy et al., 1997; Roy and VijayRaghavan, 1997). Prior to the onset of adult myogenesis, a wave of histolysis destroys most of the larval musculature. On each hemisegment, three mesothoracic muscles escape histolysis and serve as templates for the formation of one group of IFMs, the DLM fibers, through a process that involves myoblast fusion and splitting of the larval muscles (Fernandes et al., 1991). The DVMs are derived from progenitors on the wing imaginal disc and their formation does not employ a larval template, taking place instead by *de novo* fusion of myoblasts at appropriate epidermal sites. The DFMs are also derived from wing imaginal disc progenitors. Notwithstanding the fact that IFMs and DFMs are clonally related and share progenitors at least until the late third larval instar (Lawrence, 1982), it is important to note that they differentiate in different muscle types in terms of molecular markers, anatomy and physiology (Bate, 1993).

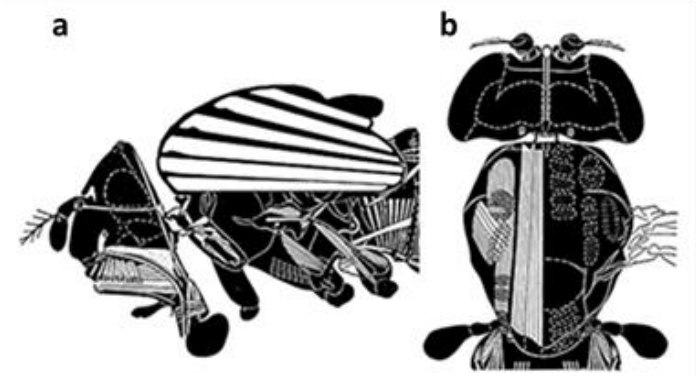

Figure 08. Localization of the indirect flight muscles in the thorax of *Drosophila melanogaster* adults. **(a)** The drawing depicts a sagittal view of the head and thorax of an adult fly. Six fibers of the dorsal longitudinal muscles (DLMs) are shown in white. **(b)** Dorsal plane of the head and thorax of the adult fly. The dorsal longitudinal muscles fibers flank the mid-sagittal plane of the thorax. Dorsal ventral muscles (DVMs) are also shown to the left side of the dorsal longitudinal muscles (DLMs). (Drawing modified from Biology of *Drosophila;* Miller, 1950).

2 OBJECTIVES

General objective

This research has as main objective to contribute to the understanding of the mechanisms that regulate transcription in higher eukaryotes.

Specific objectives

To extend the characterization of the pattern of the reporter gene expression driven by the TT CRM in *D. melanogaster* transgenic lines of the *TT-lacZ* series.

To establish correlations between the pattern of expression of the *TT-lacZ* mRNA and the pattern of expression of *CG13711*, *CG12493* and *CG32239* (*Gef64C*) mRNAs in the trachea of larva and prepupa and in the adult thorax of *D. melanogaster*.

3 MATERIALS AND METHODS

3.1 Materials

3.1.1 Transgenic lines of the 657 bp series

7.1 IIa / Cy, O

18.3 IIIa / TM3, Sb

45.3 IIIb / TM3, Sb

3.2 Methods

3.2.1 *Drosophila melanogaster* maintainance and life cycle

Drosophila lineages were maintained at 25 °C in 500 mL culture bottles, containing standard medium (10.5% corn meal; 1.84% extract yeast, 1.16% agar, 3.03% sucrose, 6.15% dextrose, 0.24% propionic acid, 0.02% phosphoric acid, 0.06% nipagin;). Flies were transferred to fresh medium twice per week. At 25 °C, the *D. melanogaster* life cycle lasts 15 days and at 18°C, 30 days, respectively.

The *Drosophila melanogaster* life cycle comprises an embryonic stage, three larval instars followed by the prepupal, pupal and adult stages. Embryonic development lasts 24 hours. The first, second and third instars last about 24 h, 24 h and 48 h, respectively. Larvae can be staged either based on the morphology of the mouth hooks or by the presence or absence and the appearance of the spiracular opening at the fore and hind end of the larvae. The prepupal stage is initiated by the formation of the puparium (pupariation), which occurs about 120 hours after egg laying (AEL) and is characterized by the shortening of the body and the eversion of the anterior spiracles. Staging of prepupae is performed by selecting 0 hour prepupae, which present a white cuticle during the first 30 minutes after pupariation, and maintaining the prepupae at 25 °C in a humid chamber. The apolysis, a process that involves the retraction of the epidermis from the cuticle of the previous instar occurs between 4 to 6 hours after pupariation. The pupal stage begins 12 hours after pupariation and is characterized by the eversion of the head. The pupal stage lasts 4.5 days and the term pupa is used to describe the animal until the emergence of the adult.

3.2.2 Embryo, L1 and L2 collection

Initially, all the adults obtained from at least two culture bottles were transferred into a fresh bottle containing standard medium sprinkled with yeast extract. Every day, during three days, the flies were transferred to a fresh bottle in order to enable the full development of the ovaries. After this period, the adults were transferred to a plastic beaker, on top of which was placed a petri dish containing oviposition medium (45.5% grape juice; 8.7% sugar; 2.2% agar; 0.0125N NaOH; 0.46% propionic acid and 0.045% phosphoric acid) slightly covered with yeast extract paste. Under these conditions, the females lay the majority of the eggs on top of the oviposition medium. Embryo collections of different time intervals were immediately used or alternatively the collection dishes were incubated at 25 °C for 24 hours or 48 hours, for L1 and L2 collections, respectively.

3.2.3 β-galactosidase histochemical assays in embryos and L1

Initially embryos and first instar larvae were collected from the grape juice agar plates and transferred to a sieve, which was placed on top of a 50 mL tube. After abundant washing with washing solution (0.3M NaCl; 0.01M Tris, pH 8.0; 0.5% Triton X-100) the sieve containing embryos and larvae was placed in a solution of commercial bleach (50%) until chorion removal (about 2 minutes). The sieve was placed again in a 50 mL tube and abundantly washed with washing solution, followed by a final wash with sterile water. The embryos and larvae were transferred to an eppendorf tube. The supernatant was removed and 350 µL of fixation solution (0.1M PIPES, pH 6.9; 0.002M EGTA, pH 7.2; 0.001M $MgSO_4$ $7H_2O$; 9.25% formaldehyde) and 700 µl of n-heptane were added. After 20 minutes incubation, at room temperature, under gentle shaking, the fixation solution and n-heptane were totally removed and the samples were washed twice with sterile water and finally with PBST [1X PBS (137mM NaCl; 2.7mM KCl; 10mM Na_2HPO_4 12 H_2O; 2mM KH_2PO_4, pH 7.4) and 0.3% Triton-X]. After supernatant removal, 500 µl of staining solution (0.01M Phosphate buffer, pH 7.2; 0.0031M $K_4Fe(CN)_6$; 0.0031M $K_3Fe(CN)_6$; 0.15M NaCl; 0.001M $MgCl_2$; 0.3% Triton X-100; 0.2% X gal) were added, followed by incubation at 37 °C, under gently shaking, until β-galactosidase activity was detected or up to a maximum of two hours. The staining solution was removed and the embryos were washed twice with staining solution without X-gal. The samples were kept at 4 °C until image acquisition and analysis.

3.2.4 β-galactosidase histochemical assays in L2, L3, prepupae, pupae and adults

Larvae, prepupae, and adults were dissected prior to assaying for β-galactosidase activity. Larvae were dissected using forceps under a stereoscope, on a glass slide in a drop of fixation solution (0.01M Phosphate buffer, pH 7.0; 0.1M NaCl; 0.001M $MgCl_2$; 1% glutaraldehyde). Dissection consisted in exposing the internal tissues, especially the trachea, salivary glands and ring gland. In the case of prepupae, 0 h prepupae were collected and incubated in a humid chamber at 25 °C during defined times followed by dissection using the same technique employed for larvae dissection.

To obtain staged adult thoraxes, initially the culture bottles were completely emptied of adult flies. On the next day, newly emerged adults were collected (1-day adults) and either dissected or transferred to a fresh vial and incubated at 25 °C during 48 hours (3 day adults) or 96 hours (5 day adults). The flies were anesthetized and the head, abdomen, wings and legs were removed from the adult body. The thoraxes were transferred to an eppendorf tube completely filled with fixation solution in order not to leave any air bubbles in the fixative. The samples were vigorously agitated by hand during 5 minutes until all thoraxes sank to the bottom of the tube. After fixation, the thoraxes were transferred to a glass slide and either sagittally or transversally cut in order to expose the thoraxic muscles. The cut thoraxes were transferred to 1X PBS and were submitted to the staining protocol.

The fixed samples were transferred to a twenty-four well cell culture plate, containing 1 ml of 1X PBS. After two washes with 250 µL of FeNaP buffer (0.01M phosphate buffer, pH 7.2; 0.0031M $K_4Fe(CN)_6$ $3H_2O$; 0.0031M $K_3Fe(CN)_6$; 0.15M NaCl; 0.001M $MgCl_2$), 250 µL of FeNaP solution containing X-gal (final concentration of 0,2%) were added and the samples were incubated during 30 minutes, at 37 °C. After staining, the samples were washed once with FeNaP solution without X-gal. Dehydration was performed by incubating the samples sequentially in 70% ethanol and absolute ethanol, 5 minutes each incubation. After removal of the absolute ethanol, samples were kept at 4 °C in 90% glycerol (1X PBS; 90% glycerol) until image acquisition and analysis.

3.2.5 Image acquisition and analysis

The results from β-galactosidase histochemical assays were analyzed in a $MZ12_5$ stereomicroscope (Leica) coupled to a digital camera (DCF295, Leica), followed by

acquisition and processing using the MultiFocus and Montage MultiFocus modules of the LAS 4.0.0 software (Leica).

3.2.6 *Drosophila melanogaster* genomic DNA extraction

Genomic DNA was extracted from *D. melanogaster* adults. Approximately 150 adults were placed in a 1.5 mL eppendorf and flash frozen in liquid nitrogen. The flies were homogenized in 500 µL of homogenization buffer (0.06M NaCl; 0.01M EDTA, pH 8.0; 5% sucrose; 0.15mM spermine; 0.15mM spermidine; 0.01M Tris, pH 7.5), followed by the addition of 500 µL of T.E. buffer (0.2M Tris, pH 9.0; 0.03M EDTA pH 8.0; 2% SDS; 5% Sucrose) and proteinase K (final concentration 0.2 µg/ml). The homogenates were incubated at 37 °C during 1.5 hours under gentle agitation. After the addition of fresh proteinase K (final concentration 0.4 µg/ml), the homogenates were further incubated for 1.5 hours, at 37 °C, under gentle agitation.

The homogenates were extracted twice with one volume of phenol:chloroform:isoamyl alcohol (25:24:1) and once with one volume of chloroform:isoamyl alcohol (24:1). The DNA was precipitated by the addition of 1/10 of volume of sodium acetate (3M $NaCOOH^+$, pH 5.2) and 2 volumes of ethanol. The DNA fiber was collected and washed twice in 500 µL of 70% ethanol. Finally, DNA samples were dried at room temperature during 10 minutes and resuspended in 80 µL of TE (100mM Tris, pH 8.0; 10mM EDTA, pH 8.0).

3.2.7 Plasmid DNA isolation

The bacterial stock containing the plasmid of interest was streaked in an LB plate containing ampicillin (final concentration 100 µg/ml) and grown at 37 °C during 18 hours. On the next day, one colony was cultured in 3 mL of LB liquid medium containing ampicillin (final concentration 100 µg/ml) and grown at 37 °C, during 18 hours. The bacterial culture was centrifuged at 15000 × *g*, during 1 minute, at room temperature, the supernatant was discarded and the pellet was resuspended in 300 µL of STET buffer (8% sucrose; 50 mM EDTA, pH 8.0; 0.1% Triton^{TM} X-100; 50mM Tris, pH 8.0) containing lysozyme (final concentration 1 µg/mL), followed by incubation during 5 minutes, at room temperature. The bacterial lysates were boiled during 1.5 minutes, followed by centrifugation at 15000 × *g*, during 10 minutes. The pellet was discarded and 12 µL of 5% CTAB were added to the supernatant followed by centrifugation at 15000 × *g*, during 5 minutes. The supernatant was

discarded and the pellet was resuspended in 300 µL of 1.2M NaCl. Plasmid DNA was precipitated with 750 µL of ethanol followed by centrifugation at 15000 × *g*, during 10 minutes, at room temperature. The supernatant was removed and the pellet was rinsed with 300 µL of 70% ethanol followed by centrifugation at 15000 × *g*, during 10 minutes at room temperature. The supernatant was removed, the DNA was air dried and resuspended in 25 µL of TE.

3.2.8 Probe preparation

Plasmid DNA was digested according to the manufacturer's instructions. After separation in an agarose gel, the fragment of interest was excised from the gel and purified using the QUIAquick® Gel Extraction Kit (Quiagen) following the manufacturer's instructions. About 100 ng of the fragment of interest (probe) was labeled using the Amersham Gene Images AlkPhos Direct Labelling and Detection System, following the manufacturer's instructions. In this system the DNA probe is initially denatured and directly labeled with a thermostable alkaline phosphatase enzyme. After probe hybridization to the target DNA immobilized on a membrane and washing steps, the hybrids are detected with the CDP-Star™ detection reagent. This reagent contains a stabilized dioxetane substrate that is decomposed by the alkaline phosphatase that is bound to the probe. The emitted light is detected by exposing the membrane to a film optimized for chemiluminescence detection.

3.2.9 Southern Blot

After electrophoresis of the digested DNA and gel documentation, the gel was trimmed and incubated in 0,125M HCl, during 10 minutes, under gentle shaking at room temperature. The gel was rinsed with sterile water followed by incubation in denaturation solution (1.5M NaCl; 0.5N NaOH), during 30 minutes, with gentle shaking, at room temperature. The gel was rinsed again with sterile water, followed by incubation in neutralization solution (1.5M NaCl; 0.5M Tris, pH 7.5), during 30 minutes, under gentle shaking, at room temperature. DNA transfer to the membrane was performed by capillarity. A glass plate was placed on a stable base and two trays containing 10 x SSC, pH 7.0 (0.75M NaCl; 75mM $Na_3C_6H_5O_7$ $2H_2O$) solution were placed on both sides of the base. A paper strip of 3 MM paper 1 cm larger than the gel, previously soaked in 10 x SSC solution, was placed on top of the glass plate. Another 3MM paper strip the same length of the gel, the salt bridge, was placed on the top of the first one in a way that both ends were immersed in the 10 x SSC

solution. Another two 3MM strips 1 cm larger that the gel and previously soaked in 10X SSC solution were placed on top of the salt bridge. The agarose gel was placed upside down on top of the last 3 MM paper and the nylon membrane (Hybond-N+, GE healthcare) was placed on top of the gel. Plastic film was placed around the gel pile in order to ensure that the 10 x SSC flow passed through the gel. Two 3MM paper strips, the same size of the gel, were placed on top of the nylon membrane and a pile of paper towel was placed on top of the pile, followed by a glass plate. A weight was placed on top of the glass plate and DNA was transferred to the membrane by capillarity overnight. At the end of the transference, the pile was disassembled, the membrane was oriented, air dried and incubated at 80 °C, during 2 hours, for DNA fixation.

4.2.9.1 Hybridization, signal generation and detection

The membrane was placed in a heat-sealable bag and for each square centimeter of membrane, 0.125 mL of AlkaPhos Direct hybridization buffer (GE Healthcare), containing NaCl (final concentration 0.5M) and blocking reagent (GE healthcare, final concentration 4% w/v) were added. The membrane was pre-hybridized at 55°C during 1 hour with shaking in a hybridization oven. After pre-hybridization, the denatured probe (final concentration 10 ng /mL) was added to the hybridization bag and the membrane was hybridized at 55°C overnight in a hybridization oven.

After hybridization, the membrane was transferred to a roller bottle containing an adequate volume (5 mL/cm^2 of membrane) of primary wash buffer (2M Urea; 0.1% SDS; 50mM $NaH_2PO_4.H_2O$, pH 7.0; 150mM NaCl; 1mM $MgCl_2$; 0.2% blocking reagent). The membrane was washed twice in primary wash buffer, during 10 minutes each wash, at 55°C with agitation. The membrane was transferred to a clean tray containing an adequate volume (5 mL/cm^2 of membrane) of secondary wash buffer (0.05M Tris; 0.1M NaCl; 2mM $MgCl_2$), and was washed twice in secondary wash buffer, during 5 minutes each wash, at room temperature with agitation. After the final wash, the blot was placed on a piece of PVC film and an adequate volume (40 µL/cm^2 of membrane) of CDP-*Star* detection reagent was applied on top of the blot. After 5 minutes the excess of detection reagent was drained, the blot was wrapped in PVC film and exposed to Hyperfilm^{TM} ECL (GE Healthcare).

3.2.10 Total RNA isolation from *Drosophila melanogaster* samples using Trizol® (Invitrogen Corporation).

Total RNA was extracted from embryo collections, whole larvae, dissected trachea of late third instar larvae and dissected adult thoraxes. The biological samples were directly placed in 300 µL of Trizol® reagent, manually homogenized, followed by vortexing during 30 seconds at room temperature. The samples were incubated during 5 minutes at room temperature, followed by the addition of 60 µL of chloroform. The preparations were inverted 10 to 15 times, followed by incubation during 3 minutes, at room temperature. After centrifuging at 15000 × *g*, during 18 minutes, at 4 °C, the aqueous phase was transferred to a clean tube, followed by the addition of 180 µL of isopropanol and glycogen (Thermo Scientific, final concentration 1µg/µL). The tubes were mixed and total RNA was precipitated overnight at – 20 °C. The next day, the samples were centrifuged at 15000 × *g*, during 15 minutes, at 4 °C, and the supernatant was discarded. The pellets were washed twice with 500 µL of 75% ethanol, followed by centrifugation at 9000 × *g*, during 6 minutes, at 4 °C. The supernatant was discarded and after air drying the RNA was resuspended in 15 µL of nuclease free water. RNA concentrations were determined by absorbance measurement at 260 nm, assuming that 40 µg/mL of RNA has an absorbance of 1.0. Samples with a 260/280 ratio greater than 1.8 were used for DNase treatment and cDNA synthesis.

3.2.11 DNase treatment

Total RNA samples were treated with DNase to ensure that genomic DNA was totally removed before the cDNA synthesis. An amount of 1.5 µg of RNA was mixed with RQ1 RNase-Free DNase Reaction Buffer (Promega) (final concentration 1x), 1.5 U of RQ1 RNase-Free DNase (Promega) and water to a final volume of 10 µL, followed by incubation at 37 °C, during 30 minutes. DNase treatment was interrupted with the addition of 1 µL of DNase Stop solution (Promega), followed by incubation at 65 °C, during 10 minutes.

3.2.12 cDNA synthesis

cDNA synthesis was immediately performed after DNase treatment. Initially, Random Primers (Promega) (final concentration 0.01µg/µl) and Oligo dT (Promega) (final concentration 0.01µg/µl) were added to the DNase treated RNA, followed by incubation at 70 °C, during 5 minutes. After denaturing, the samples were kept on ice until the reverse transcription reaction. The reverse transcription reactions contained ImProm-II™ reaction

buffer (Promega), $MgCl_2$ (final concentration 3 mM), dNTPs mix (Invitrogen, final concentration 5 mM), 20 U of RNasin Inhibitor (Promega) and ImProm-II™ Reverse Transcriptase (Promega). The negative control reaction was performed in parallel under identical conditions without the addition of reverse transcriptase. Reverse transcriptase reactions were performed in a thermocycler with the following temperature conditions: an initial annealing step, at 25 °C, during 5 minutes, followed by a synthesis step, at 42 °C, during 60 minutes, and a reverse transcriptase inactivation step, at 70 °C, during 15 minutes. The obtained cDNA was stored at -20 °C until use.

3.2.13 Primer design and validation

The coding sequences of *Gef64C (CG32239)*, *CG12493*, *CG13711* and *lacZ* were obtained from public databases (FlyBase and NCBI). Primer sequences were designed using Oligo Explorer *v* 1.1.2. Whenever the gene of interest presented intronic sequences, primers were designed to anneal to the neighboring exons. In the case of *Gef64C* the primers amplify both *Gef64C* isoforms and anneal at the 3´ end of exon 6 and at the 5´ end of exon 7. In the case of *CG12493* the primers anneal at the 3´end of exon 1 and at the 5´ end from the exon 2. Both *CG13711* and the bacterial *lacZ* gene do not present intronic regions (Table 01).

Table 01. List of primers employed in the qPCR reactions, respective sequences, length and size of the amplified fragments.

Primer (Foward/Reverse)	Sequence 5´→3´	Length (bp)	Amplicon length (bp)	Genomic amplicon length (bp)
CG32239-F	ACAAAGGACAGAGTGCATCG	20	159	412
CG32239-R	AGAGAAGCGTTCTTGTCCAC	20		
CG12493-F	GTGCAACCGACAACCAATTG	20	153	326
CG12493-R	GAAGATTCACCATCGGGCATA	21		
CG13711-F	GTACATGAGGTTGTCCAAGG	20	148	-----
CG13711-R	CTGAATCTGGCGACGAAGTT	20		
lacZ-F	GCAGGATATCCTGCTGATGA	20	159	-----
lacZ-R	GATCATCGGTCAGACGATTC	20		

qPCR reactions were performed in a final volume of 12 µL and contained 1X FastEva green qPCR mix (Biotium), 1 pmol of each primer and 1 µL of cDNA. Reactions were carried

out in a 96-well qPCR plate (Applied Biosystems) and performed in the Gene Amp® 7500 Sequence Detection System (Applied Biosystems) equipment. The following cycles were routinely employed: a first step of enzyme activation at 95 °C for 2 minutes; a second step comprising 40 cycles of denaturation at 95 °C for 5 seconds, annealing at 60 °C for 30 seconds and extension at 72 °C for 25 seconds. Reactions without template (cDNA) and reactions containing the product of a cDNA synthesis run in the absence of reverse transcriptase were run in parallel.

The amplification efficiency (E) of each primer was assessed using serial dilutions. In the case of *Gef64C* and *CG12493* genes, a mixture of cDNAs (0-4 hours embryos and late L3 instar larva) were employed to perform linear regression curves (standard curves) for each pair of primers. In the case of *CG13711*, cDNA from 0-4 hours embryos was employed and in the case of the *lacZ* gene cDNA from 3 days adult thoraxes was employed. The amplification efficiency was calculated by using the Slope of the standard curve employing the following formula: $E = [10^{(-1/slope)} - 1] \times 100$. The calculated efficiency for *Gef64C*, *CG12493*, *CG13711* and *lacZ* were 94.5%, 102.1%, 91.6% and 109.5% respectively (Figure 09).

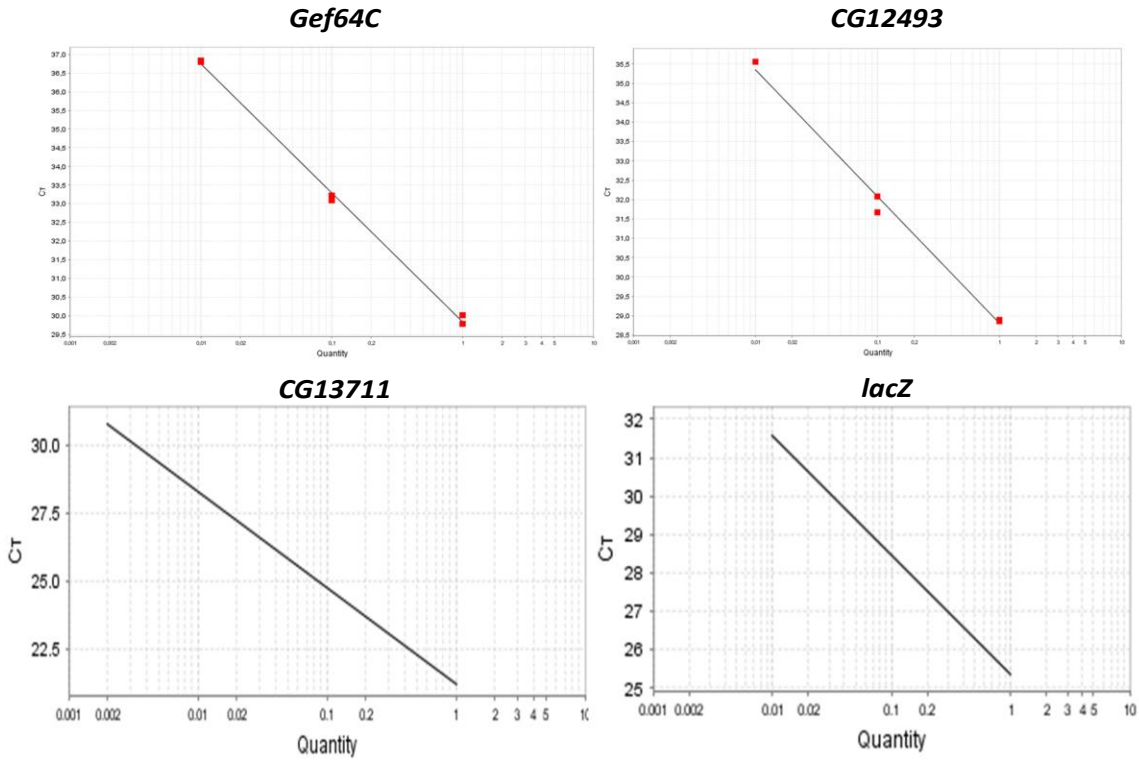

Figure 09. Standard curves obtained for the *Gef64C*, *CG12493*, *CG13711* and *lacZ* primers. Standard curves were obtained for each primer set with 10-fold serial dilutions of the cDNA samples (10-, 100-, and 1000-fold). Slope values of the standard curves, were determined for each primer set and were used to calculate the efficiency (E).

A melting curve analysis was performed for each sample at the end of each run. Melting curves were converted to distinct melting peaks by plotting the first negative derivative of the fluorescence as a function of temperature (-dF/dT). This analysis was employed to determine if a single amplicon was generated in each sample, if any amplicon was generated in the negative controls and enabled us to investigate the occurrence of primer dimers (Figure 10). As a further quality control, the reaction products were routinely analyzed in 2% agarose gels (Figure 11).

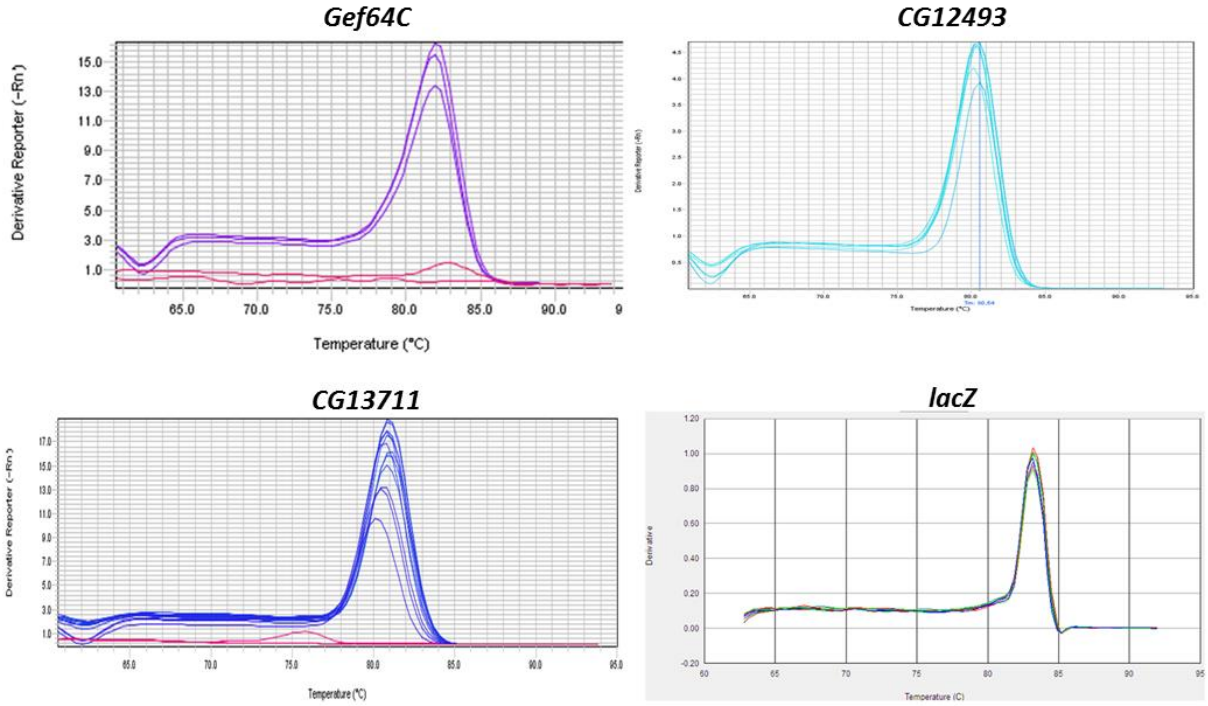

Figure 10. Melting curves obtained for the *Gef64C, CG12493, CG13711* and *lacZ* primers. Melting curves were obtained after each qPCR calibration experiment for each set of primers. On the top of each curve is indicated the set of primers. In the X-axis is indicated the temperature and in the Y-axis is indicated the negative derivative (-Rn).

The relative transcript levels of target genes and reference genes in trachea and dissected thoraxes were estimated using the ΔΔCq method (Livak and Schmittgen, 2001). Four previously described reference genes were employed in RT-qPCR experiments (Ponton et al, 2011). *RpL32* and *RpS20* reference genes were used for dissected trachea and *tubulin* and *Mnf* were used for dissected thoraxes. The relative mRNA levels of expression were calculated from at least two independent biological samples and the analysis of the data was performed using the Excel software (Microsoft Corporation).

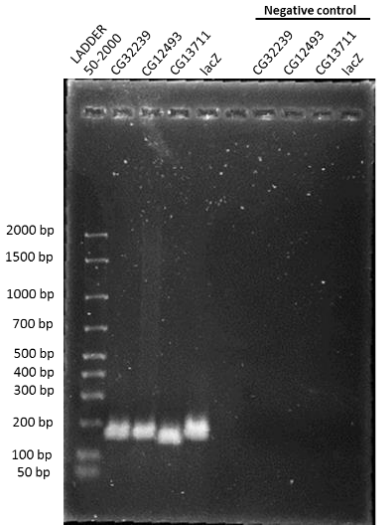

Figure 11. Amplification products of *Gef64C* (*CG32239*), *CG12493*, *CG13711* and *lacZ*. Image of a 2% agarose gel stained with ethidium bromide. On the top of the figure are indicated the sets of primers that were employed. Negative control reactions were run in parallel in the absent of template. On the left is indicated the molecular weight of each fragment of the 50-2000 bp ladder.

4 RESULTS

4.1 Molecular characterization of the *TT-lacZ* transgenic lines

Previous work from our group resulted in the isolation of three transgenic lines transformed with the *TT* construct, lines *TT7.1*, *TT18.3* and *TT45.3*. The *TT* construct contains a 657 bp *Drosophila* genomic fragment, located about 2500 bp upstream *CG13711*, about 400 bp upstream *CG12493*, in a genomic region that constitutes one of the introns of *CG32239* (*Gef64C*), cloned upstream de *lacZ* reporter gene. After integration into the *Drosophila* genome, the transposon presents the following structure: *P* element, / *mini white* gene (marker gene) / 657 bp genomic fragment / *Fbp1* core promoter / *lacZ* (reporter gene) / *P* element (Figure 12).

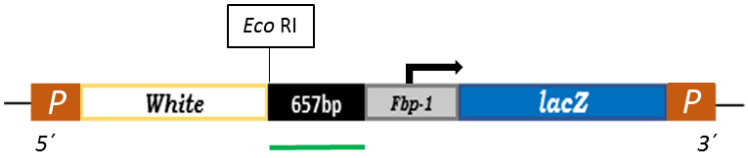

Figure 12. Structure of the transposon containing the *TT cis*-regulatory module. The *orange* boxes represent the *P* elements; the *white* box represents the *mini white* marker gene; the 657 pb *black* box represents the *TT cis*-regulatory module, the *grey* box represents the *Fbp-1* core promoter and the blue box represents the *lacZ* reporter gene. An *Eco* RI restriction site was inserted at the 5' end of the 657 bp fragment during the cloning process and constitutes the only *Eco* RI site in the transposon. The black arrow indicates the transcription start site, located in the *Fbp-1* core promoter and the direction of transcription. The green bar under the diagram indicates the region of the transposon to which the 657 bp probe hybridizes (Zampar et al., 2009).

In order to confirm that the three lines from the *TT* series were the result of independent integration events, a Southern blot containing *Eco* RI digested genomic DNA extracted from each transgenic line was hybridized to the 657 pb fragment. As shown in figure 14, two different genomic *Eco* RI fragments were detected in each transgenic line. One of the fragments is about 29 kb and is detected in all three transgenic lines. The second fragment presents a variable molecular weight in the analyzed transgenic lines (Figure 13).

The *TT cis*-regulatory module is located in *D. melanogaster* chr 3L: 4754052-4754708. A search for *Eco* RI restriction sites in the proximity of the endogenous *TT* CRM in the *D. melanogaster* genome (Aug. 2014 BDGP Release 6 + ISO1MT/dm6) was performed using the *Restriction Enzymes Track* tool of UCSC (Genome Browse University of Santa Cruz). This search resulted in the identification of six *Eco* RI sites in the chromosome 3L 4730000-4785000 region (Figure 14). The endogenous *TT cis*-regulatory module, which does not present *Eco* RI sites, is contained in a 29 kb *Eco* RI genomic fragment which is the higher molecular weight *Eco* RI fragment detected in the three analyzed transgenic lines (Figure 14).

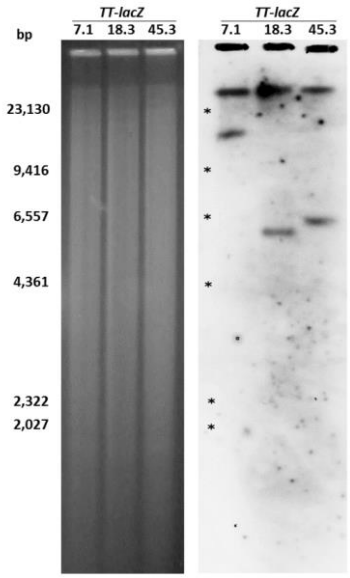

Figure 13. Southern blot analysis of the three independent lines of the TT series. On the left hand side, an image of a 0.8 % agarose gel stained with ethidium bromide, containing 10 µg of *Eco* RI digested genomic DNA extracted from adults of the three transgenic lines of the *TT* series, as indicated on top of the figure. The autoradiogram on the right hand side of the figure is a Southern blot of the gel shown on the left, after hybridization with the 657 bp probe. The numbers on the left indicate the position of migration of the fragments of the λ/*Hind* III molecular weight marker (in base pairs).

The lower molecular weight *Eco* RI fragments of different sizes that were detected in the transgenic lines contain the 657 pb fragment that is part of the transposon that was integrated into the genome. The transposon is about 10 kb long and contains at the 5´ end of the 657 bp fragment a single *Eco* RI restriction site which is located about 4 kb upstream of the 3´ end of the transposon (Figure 13). Since the three genomic fragments that were detected by the probe are longer than 4 kb and are of variable size, we suggest that the transposon is intact in the lines of the *TT* series, that each line contains a singlte copy of the transposon and concluded that the three lines of the *TT* series are the result of independent integration events.

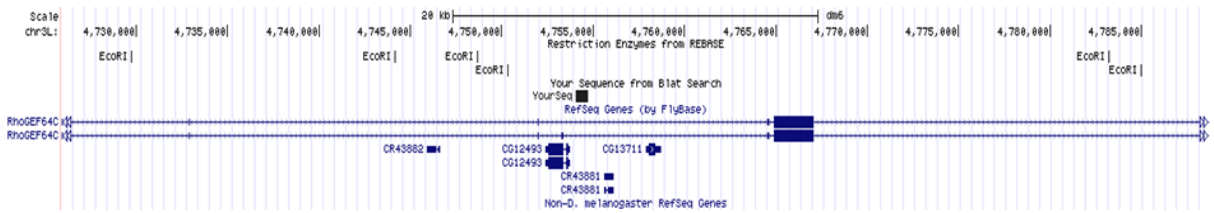

Figure 14. Localization of *Eco* RI restriction sites near the *TT cis*-regulatory module in the *D. melanogaster* genome. On top of the figure the 3L chromosome coordinates are shown (version Aug. 2014 BDGP Release 6 + ISO1MT/dm6). *Eco* RI restriction sites identified using the *Restriction Enzymes Track* tool of UCSC are indicated below the coordinates. The black rectangle (*your seq*) shows the genomic location of the *TT cis*-regulatory module in the *D. melanogaster* genome. (Modified from http://genome.ucsc.edu).

4.2 Characterization of the β-galactosidase patterns of expression driven by the TT CRM

Previous results have shown that the TT CRM drives gene expression in the tracheal system of both third instar larvae and early prepupae and in the adult thoracic muscles (Zampar et al., 2009). In order to refine the characterization of the pattern of expression driven by the *TT* CRM, the three lines from the *TT* series were investigated using a β-galactosidase histochemical assay (Lis et al., 1983). In this analysis, we assayed whole embryos and whole first instar larvae, dissected second and third instar larvae, 0h, 1 h and 2 h prepupae, 24 h pupae and adults 1, 3 and 5 days after emergence.

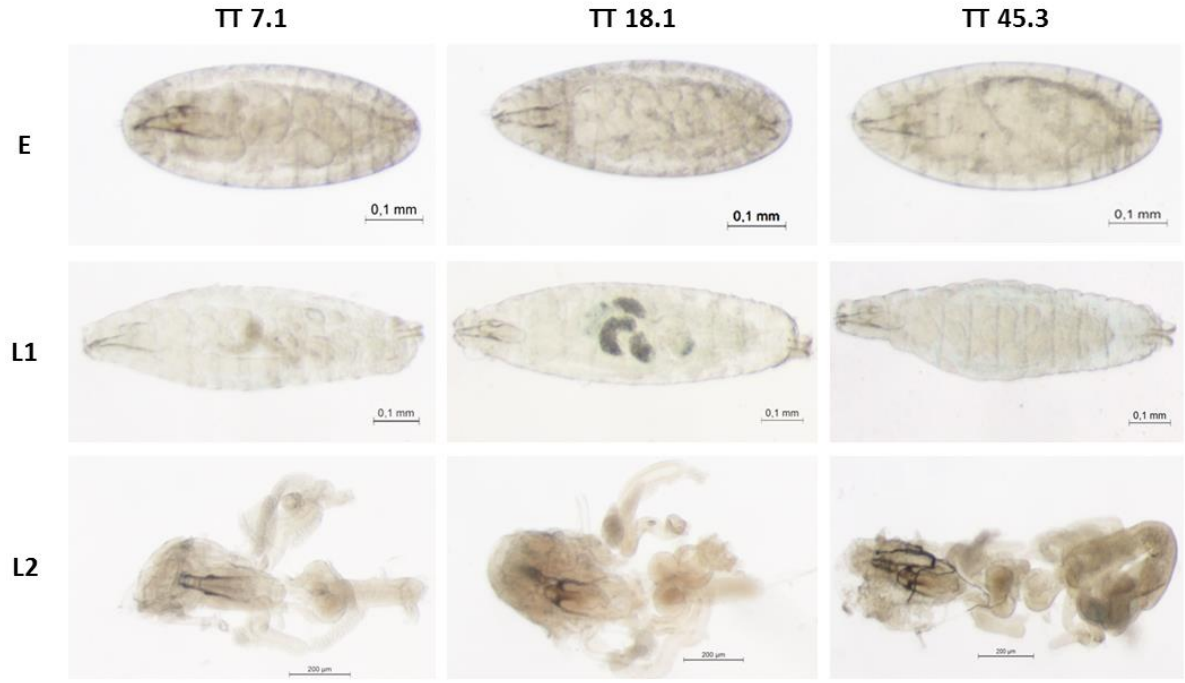

Figure 15: β-galactosidase pattern of expression in late embryos and in first and second instar larvae. (E) late embryos, (L1) first instar larvae, (L2) second instar larvae were submitted to the β-galactosidase histochemical assay (Lis et al., 1983). The analyzed transgenic lines are indicated on top of the figure.

In the lines of the TT series, β-galactosidase activity is not detected in any tissue during the initial developmental stages (late embryos, first and second instar larvae) (Figure 15). During the third larval instar, reporter gene expression is initially detected in the tracheal system and continues to be detected in this tissue during the prepupal stage (0h, 1h and 2h prepupae) and in 24 hours pupae (Figures 16 and 17). β-galactosidase activity in the tracheal system is observed both in the two dorsal longitudinal trunks and in internal branches in third instar larvae, prepupae and pupae. This pattern of expression was consistently observed in the

three independent transgenic lines. Together, our results extend the characterization of the pattern of expression driven by the TT CRM in the larval and prepupal tracheal system and reveal that the transgene is expressed in the tracheal system of pupae. These results confirm that during these developmental stages reporter gene expression is restricted to the tracheal system.

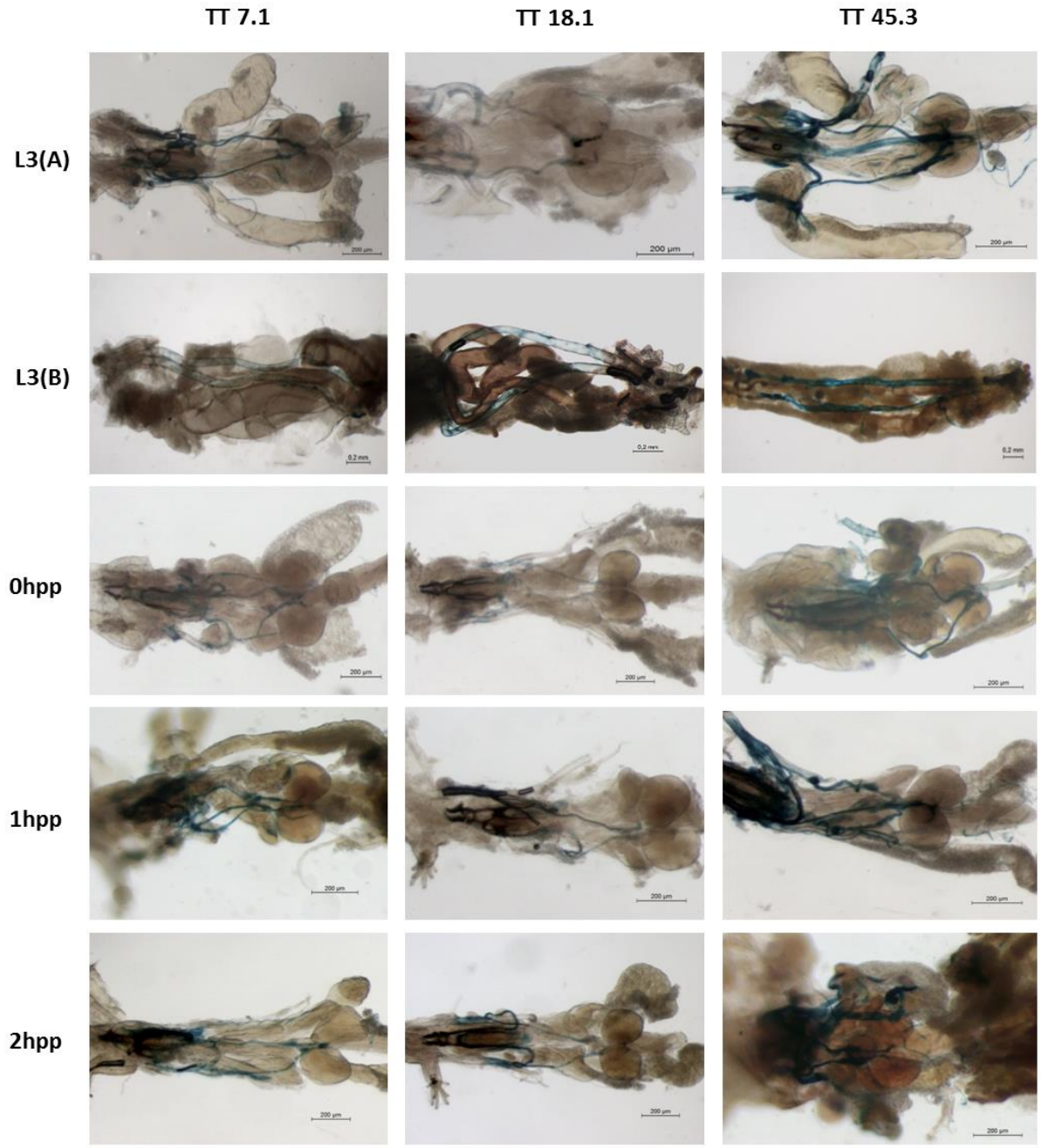

Figure 16: β-galactosidase pattern of expression in third instar larvae and 0h, 1h and 2h prepupae. (L3A) anterior and (L3B) posterior region, respectively, of dissected third instar larvae, (0hpp) 0h prepupae, (1hpp) 1h prepupae and (2hpp) prepupae were submitted to the β-galactosidase histochemical assay (Lis et al., 1983). β-galactosidase activity is observed by

the formation of a blue precipitate. The analyzed transgenic lines are indicated on top of the figure.

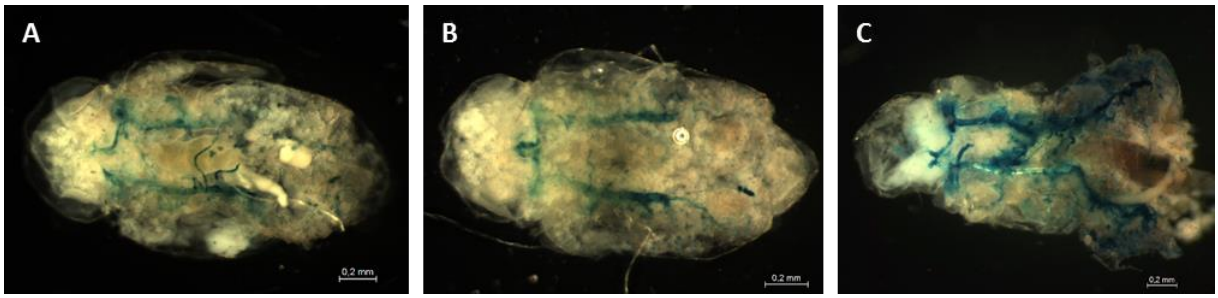

Figure 17: β-galactosidase pattern of expression in 24h pupae. 24h pupae were submitted to the β-galactosidase histochemical assay (Lis et al., 1983). β-galactosidase activity is observed by the formation of a blue precipitate. (A) line *TT 7.1*; (B) line *TT 18.3* and (C) line *TT 45.3*.

In order to further characterize the pattern of expression driven by the *TT* CRM in the adult thorax, the thoraxes of 1, 3 and 5 days old adults were dissected, fixed, cut either at the sagittal plane or transversally, and assayed for gene reporter expression using the β-galactosidase histochemical assay (Lis et al., 1983). As can be observed in figures 19 and 20, reporter gene expression is detected in the thoracic muscles of both male and female adults, 3 and 5 days after emergence, whereas gene reporter expression was not detected in any of the thoracic muscles of 1 day old adults in all lines of the *TT* series. By analyzing both sagittal and transverse sections, we concluded that in these lines reporter gene expression occurs in a group of twelve muscle fibers, which are named the dorsal longitudinal muscles. Together this analysis confirm that the *TT* CRM drives gene expression in the thoracic muscles and extend the previous characterization of the lines of the TT series showing that thoracic gene expression is initially detected in the third day after emergence, is not sex specific and seems to be restricted to a specific group of muscles, the dorsal median muscles.

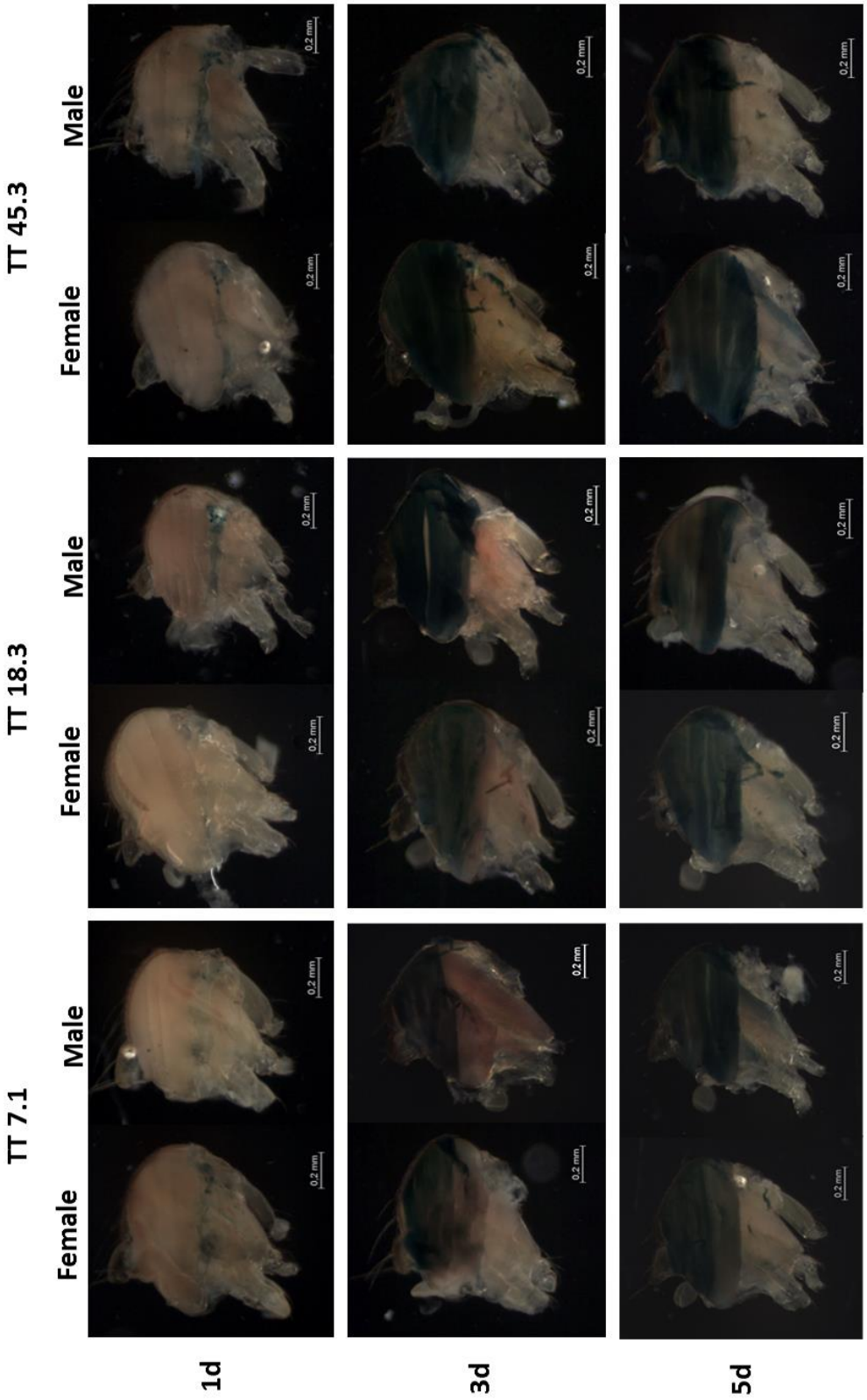

Figure 18: β-galactosidase pattern of expression in the adult thoracic muscles, sagittal plane. Dissected thoraxes of one day (1d), three days (3d) and five days (5d) old adults were submitted to the β-galactosidase histochemical assay (Lis et al., 1983). β-galactosidase activity is observed by the formation of a blue precipitate. The analyzed transgenic lines are indicated on top of the figure.

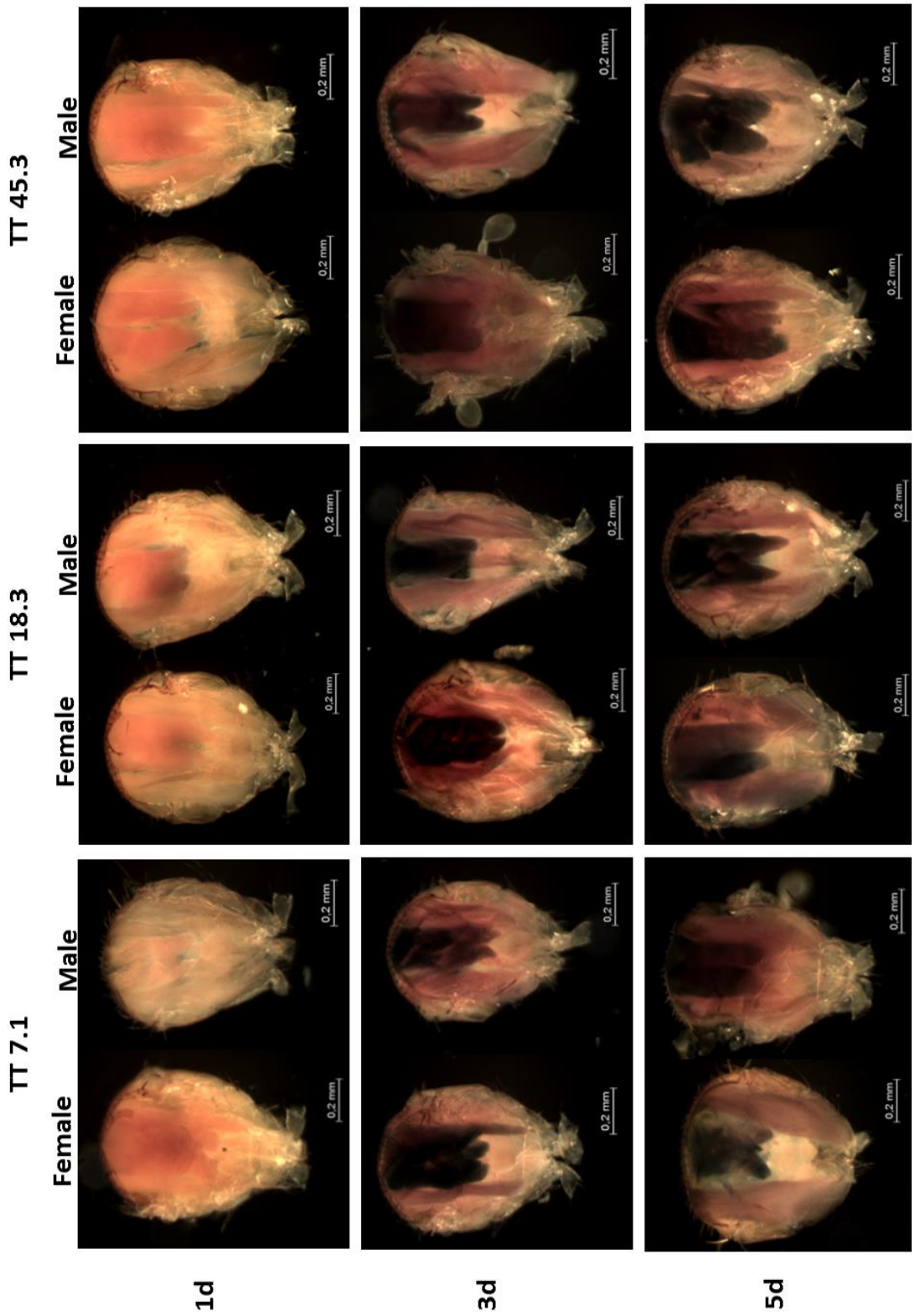

Figure 19: β-galactosidase pattern of expression in the adult thoracic muscles, transversal plane. Dissected thoraxes of one day (1d), three days (3d) and five days (5d) old adults were submitted to the β-galactosidase histochemical assay (Lis et al., 1983). β-galactosidase activity is observed by the formation of a blue precipitate. The analyzed transgenic lines are indicated on top of the figure.

4.3 Relative levels of *lacZ, Gef64C, CG12493, CG13711* expression in the tracheal system of larvae and prepupae and in the adult thorax.

The *TT* CRM drives a dynamic pattern of expression at different developmental stages, prior to metamorphosis and during the adult stage, in two different tissues, the tracheal system and the thoracic muscles. Since the *TT* CRM lies in an intergenic region that is located upstream of *CG13711* and *CG12493* and is also located in one of the introns of *Gef64C* (*CG32239*) and the available information about this set of *CGs* is still scarce, it is not possible to infer which of the *CGs* is regulated by the *TT cis*-regulatory module. Therefore, we investigated the pattern of expression of the *lacZ* reporter gene and of these three *CGs* in the tracheal system of both larvae and prepupae and in the adult thorax using real time quantitative PCR (RT-qPCR) experiments to determine which of the three *CGs* present(s) an expression pattern reminiscent of the *lacZ* expression pattern.

Since the endogenous *TT* CRM resides in chromosome three and we did not determine the cytological location of the transgenes, we employed the transgenic line *TT 7.1*in the RT-qPCR experiments because it was the only line in which the integration of the transposon occurred in chromosome 2. In these experiments, we sought to investigate the change-fold in the relative mRNA expression levels of the *lacZ* reporter gene and of *CG13711, CG12493* and *Gef64C*. Specific primers for each gene were used to quantify the relative mRNA expression levels, which were normalized against *RpS20/CG15693* in the case of the tracheal system of larvae and prepupae and against the *tubulin* gene in the case of the adult thoraxes.

Initially, the relative expression levels of the *lacZ mRNA* and of *CG13711, CG12493* and *Gef64C* mRNAs were determined in dissected trachea of late third instar larvae and 0h prepupa. In this analysis, we employed as a calibrator the normalized expression levels of each of the analyzed genes present in the trachea of third instar larvae, which enabled us to assess if the levels of expression of a given gene were either increasing or decreasing in 0h prepupae. RT-qPCR results reveal that the relative expression levels of the *lacZ* mRNA are about 3 times higher in the trachea of prepupae, when compared to the levels of expression of this gene in third instar larvae (Figure 21).

When the mRNA levels of expression of *CG13711* and *Gef64C* in 0h prepupae were compared to the mRNA expression levels present in the trachea of third instar larvae, an increase was verified for both *CG13711* (1.5 fold) and *CG32239* (2.3 fold). A different

situation was observed in the case of *CG12493*. The relative mRNA expression levels of *CG12493* in 0h prepupae trachea are lower (about 20% reduction) when compared to the mRNA expression levels of *CG12493* in third instar larvae (Figure 21). Together, these results reveal that the expression of the transgene is reminiscent of two of the genes that lie in the vicinity of the *TT* CRM, *CG13711* and *Gef64C*.

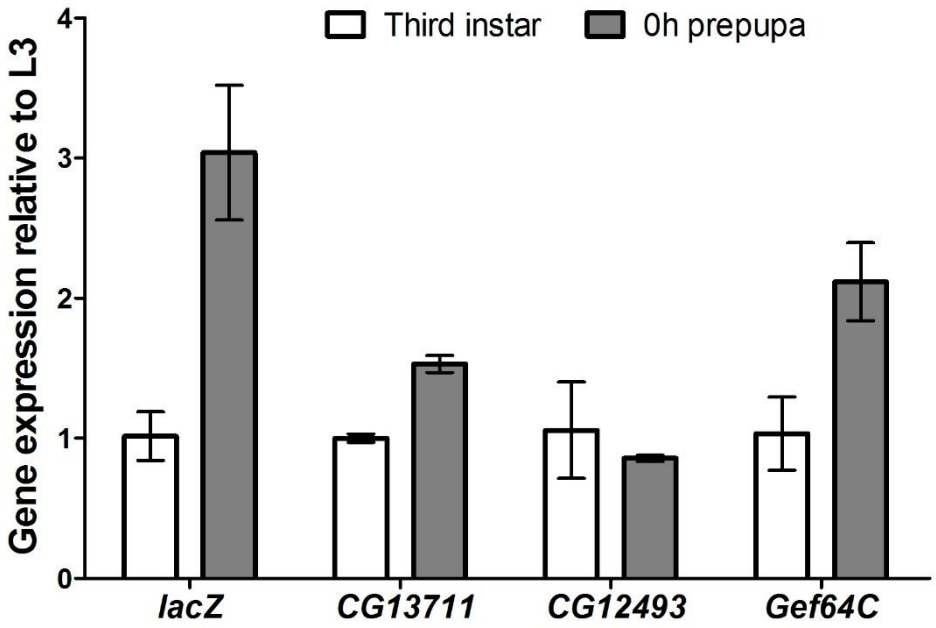

Figure 20: Relative levels of mRNA expression of *lacZ*, *CG13711*, *CG12493* and *Gef64C* in dissected tracheas of third instar larvae and 0h prepupae of line *TT 7.1*. Total RNA was extracted from dissected tracheas at the indicated developmental stages and employed in RT-qPCR experiments. White and grey bars represent the relative mRNA expression levels in third instar larvae and 0h prepupae dissected tracheas, respectively. For each gene, the mRNA expression levels in 0h prepupae were determined in relation to the expression levels in L3, which were arbitrarily set as 1.

Next, a similar approach was undertaken to investigate the mRNA relative expression levels of these four genes in dissected thoraxes of adults 1, 3 and 5 days after emergence. In this case, we have investigated if the levels of expression of a given gene were either increasing or decreasing in 3 or 5 days old adult thoraxes in relation to the mRNA expression levels found in 1 day old adult thoraxes. The RT-qPCR results show that the relative levels of *lacZ* mRNA increase in 3 days old adults (about 38 times in relation to 1 day old adults) and decrease in 5 days old adults (about 12 times in relation to 1 day old adults).

The only gene that presents an increase in the levels of mRNA expression in 3 days old adults constitutes *CG12493*, which increases about 2 times in 3 days old adults. In five

days old adult thoraxes the levels of expression decrease and are about 40 % of those found in 1 day old adults. The two other investigated *CGs*, *CG13711* and *Gef64C* show a reduction in the relative expression levels in the thoraxes of 3 days old and 5 days old adults, where the levels found in five days old adults are lower than those found in three days old thoraxes (Figure 22). Together, these results reveal that the pattern of expression of the transgene is reminiscent of the pattern of expression of *CG12493*.

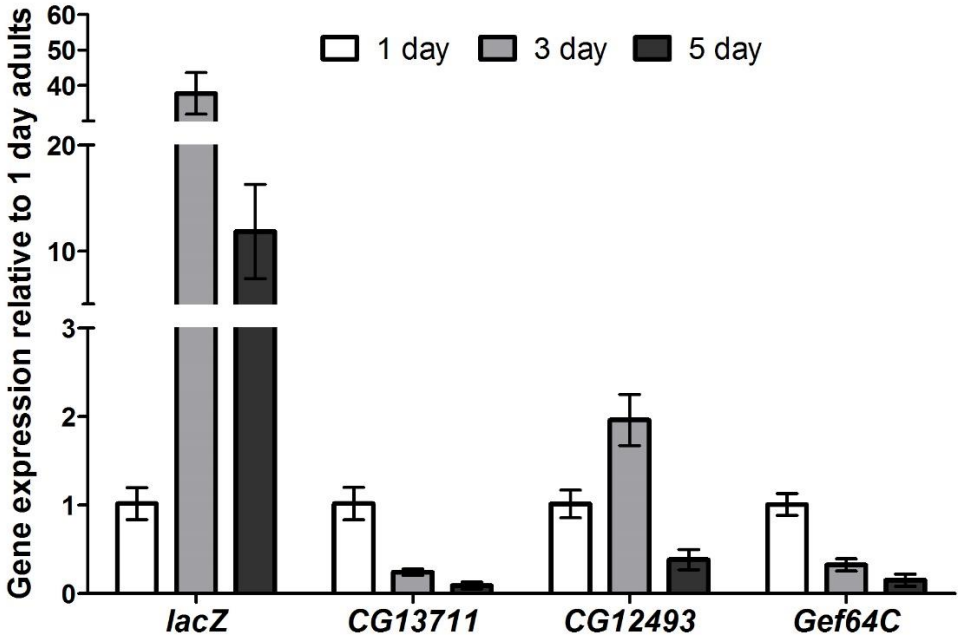

Figure 21: Relative levels of mRNA expression of *lacZ*, *CG13711*, *CG12493* and *Gef64C* in thoraxes of 1, 3 and 5 days old adults of line *TT 7.1*. Total RNA was extracted from thoraxes of 1, 3 and 5 days old adults and employed in RT-qPCR experiments. Grey and black bars represent relative mRNA expression levels in 3 days old and 5 days old adult thoraxes, respectively. For each gene, the mRNA expression levels in 3 days old and 5 days old adult thoraxes were determined in relation to the expression levels in 1 day old thoraxes, which were arbitrarily set as 1.

5 DISCUSSION

5.1. The TT *cis*-regulatory module drives a dynamic pattern of gene expression through development.

CRMs contain transcription factor binding sites that are recognized by transcription factors in distinct cell types at defined developmental times. The interactions between CRMs in a given gene regulatory region modulate gene expression in time and space and provide a unique code for when, where and at which level a particular gene is transcribed (Jeziorska et al., 2009). In this work, we have characterized a novel *D. melanogaster* CRM that we named *TT*. This CRM resides upstream the divergently transcribed *CG13711* and *CG12493*, in a genomic region that corresponds to one of the introns of *CG32239*.

The *TT* CRM drives reporter gene expression in the tracheal system of third instar larvae, prepupae and pupae and in the thoracic DLMs of 3 and 5 days old adults. The present work not only confirm preliminary results obtained from the initial characterization of these lines (Zampar, 2009), but also, reveal that the tracheal *TT-lacZ* expression also occurs during the pupal stage and that the adult thoracic expression is restricted to a specific set of 12 muscle fibers, the DLMs. The observation that the pattern of reporter gene expression is consistent in the three independent transgenic lines of the *TT-lacZ* series indicates that it constitutes the pattern of expression driven by the *TT* CRM.

The 657 bp *TT* CRM contains a 20 bp sequence that is identical to a sequence contained in the *BhC4-1* 67 bp *(-253/-187*) ring gland CRM (Lecci et al., 2008). Besides containing this 20 bp sequence that is also highly conserved in the genome of 22 *Drosophila* species, the 657 bp *TT* CRM contains additional sequences that are conserved in the genome of the other *Drosophila* species that have been sequenced (Figure 22). The 657 bp fragment was originally selected as a candidate CRM that potentially could drive developmentally regulated gene expression in the ring gland (Zampar et al., 2009). In spite of the fact the *TT* CRM does not drive ring gland expression, it is worthy of note that recent studies demonstrated that the trachea and two glands that are part of the ring gland, the *corpora allata* and prothoracic glands, have a homologous origin and that the early endocrine and trachea gene regulatory networks are similar (Sánchez-Higueras et al., 2014). The *corpora allata* and the prothoracic glands develop from ectodermal placodes in the maxilla and labium, respectively, whereas the tracheal system develops from ten trunk placodes. Both the endocrine and trachea primordia invaginate in a similar way but in the endocrine primordia the activation of Snail leads to an epithelial-mesenchymal transition, after which the two

primordia unite and migrate dorsally and join the *corpora cardiaca* to form the ring gland (Sánchez-Higueras et al, 2014). In this sense, it is possible that either the 20 bp sciarid sequence, or other sequences that are highly conserved in the 657 bp CRM, are also present in the CRMs of other tracheal and/or endocrine genes, which are part of gene regulatory networks that act during early tracheal and endocrine development in insects.

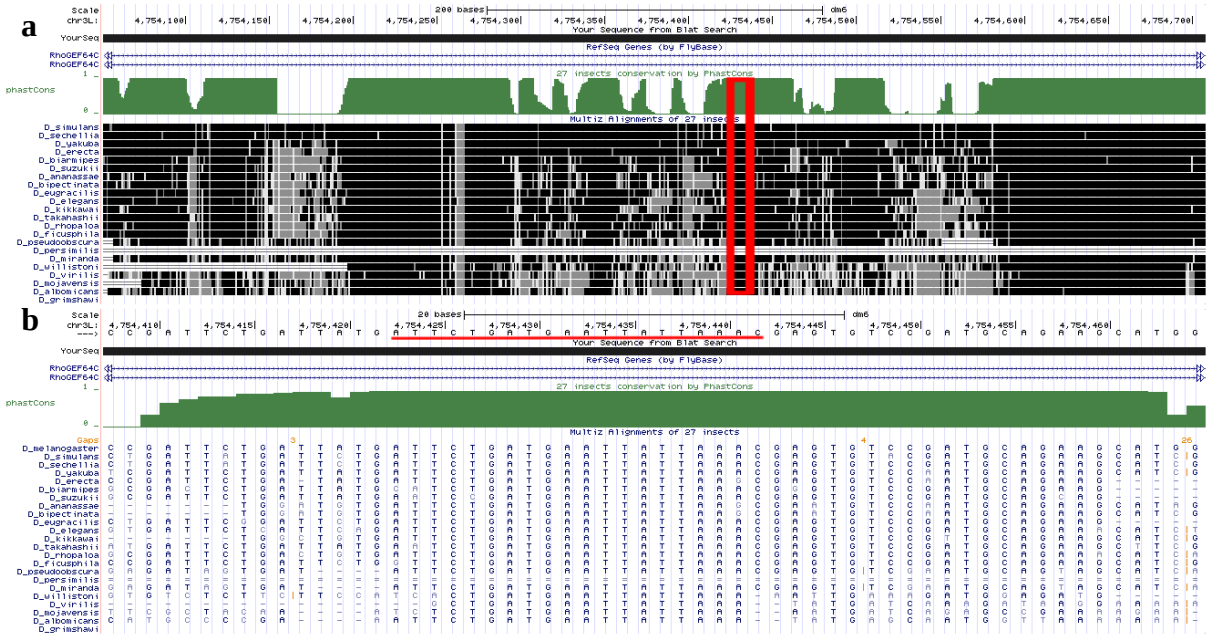

Figure 22. Location of the sequence similar to the ring gland CRM in the *Drosophila melanogaster* genome. (A) On the top is shown the location of the *TT* CRM (*your seq)* in chromosome 3L: 4754052-4754708. The red rectangle indicates the approximate position of the sequence similar to a sequence contained in the *BhC4-1* ring gland CRM. *phastCons* indicates the degree of conservation of *TT* CRM sequences in the genome of the 22 *Drosophila* species that have been sequenced. **(B)** Zoom in of the region of the *D. melanogaster* genome that contains the sequence that is identical to a sequence contained in the *BhC4-1* ring gland CRM (red underlined). This sequence is also highly conserved in the other *Drosophila* genomes that have been sequenced.

5.2. The *TT-lacZ* expression in the thoracic DLMs.

Three *CGs* reside in the vicinity of the *TT* CRM: *CG13711*, *CG12493* and *CG32239* (*Gef64C*). In order to determine if any of these three *CGs* presented a pattern of expression reminiscent to the pattern of expression driven by the *TT* CRM, we searched the available literature. In essence, *CG13711* is expressed in the cellularized blastoderm and its function is still unkown (Bonneaud et al., 2003; Gross et al., 2003). *CG12493* has been characterized as a non-essential gene expressed in male testes and encodes a protein that contains a double-stranded RNA binding domain (DRBD) (Saunders and Barber, 2003; Gerbasi et al., 2010).

The more extensively characterized of these three *CGs* is *CG32239* (*Gef64C*), a guanidine exchange factor that regulates the small GTPase Rho1. Rho1 restricts actin nucleation to the cell apical region during the process of formation of tubular structures such as the tracheal system, salivary glands, hindgut and Malpighian tubules (Lovegrove et al., 2006; Simões et al., 2006). Therefore based on the available information, we could not determine which of these *CGs*, if any, presented an expression pattern reminiscent of the pattern of expression driven by the TT CRM.

In this context, we performed RT-qPCR experiments to determine the relative levels of *CG13711*, *CG12493*, *CG32239* (*Gef64C*) and *lacZ* mRNA expression at different developmental stages in the tissues that presented β-galactosidase activity, namely the tracheal system of larvae and prepupae and in the adult thorax. Interestingly, RT-qPCR experiments revealed that the relative levels of expression of all analyzed genes varied both in the larval to prepupal transition and at different days after adult emergence. To establish a correspondence between the pattern of expression of the three *CGs* and the pattern of expression of the reporter gene we compared the profiles of expression of each *CG* with the profiles of expression of the reporter gene.

In adults, both the *lacZ* reporter gene and *CG12493* show an increase in the relative levels of mRNA expression in the thorax of 3 days old adults when compared to 1 day old adults. In 5 days old adults the relative levels of expression of both genes decrease. On the other hand, the relative levels of expression of both *CG13711* and *CG32239* (*Gef64C*) progressively decrease after emergence. In this context, it is possible that in the adult thorax the *TT* CRM regulates the expression of *CG12493*.

CG12493 has been characterized as a male specific gene expressed at high levels in the adult testes (Gerbasi et al., 2010). This agrees with the available RNAseq data showing that the highest levels of *CG12493* expression are found in the adult testes (REF). In adults of the lines of the *TT-lacZ* series, β-galactosidase expression occurs in the DLM of the thorax of both females and males and is not verified in the male testes. Since in the histochemical assays we did not detect a sex specific pattern of expression. In our RT-qPCR experiments we employed a mixture of dissected thoraxes of both males and females. The available RNAseq data show that the levels of expression of *CG12493* in 1 and 5 old day males are higher than those found in 1 and 5 old day females. In this case, whole adults were employed and therefore the higher levels of *CG12493* expression in males could be due to the sum of the

expression levels of *CG12493* in the testes and in the adult thoracic muscles. On the other hand, lower levels of expression of *CG12493* detected through RNAseq in adult females could be due to the expression of this *CG* in the adult thoracic muscles. Therefore, our results showing that the reporter gene is expressed in the DLM of both males and females might be revealing a new pattern of expression of C*G12493* that is not sex specific. Additionally, since we did not detect reporter gene expression in the male testes, we infer that the elements that regulate the expression of *CG12493* in this tissue are not contained in the TT CRM. *In situ* hybridization experiments should be performed in the near future in order to determine if *CG12493* is actually expressed in the thoracic DLMs, which would confirm that our functional analysis did reveal a new pattern of expression of *CG12493*.

Intriguingly, our results revealed that the expression of the reporter gene occurs in a specific set of thoracic muscles, the DLMs. The adult thoracic musculature comprises two types of muscles, the tubular and fibrilar muscles (Berstein et al., 1993; Allen and Hodenshwege, 2006). The indirect flight muscles (IFMs) are fibrillar muscles and include two main groups of muscle fibers, the dorsal longitudinal, DLMs, and the dorsal ventral, DVMs (Fernandes et al., 1991; Fernandes and VijayRaghavan, 1993; Fernandes et al., 1996; Roy et al., 1997; Roy and VijayRaghavan, 1997). During metamorphosis, the DLMs and the DVMs are both originated from myoblasts that segregate from the wing imaginal disk. In the case of the DLMs, the myoblast fuse to larval muscle scaffolds, whereas the DVMs are formed *de novo* (Costello and Wyman, 1986; Fernandes et al., 1991). It has been suggested that these two distinct ways of muscle formation result in the development of two distinct types of fibrillar muscles. In this context, it is possible that sequences contained in the *TT* CRM are being recognized by DLM specific transcription factors that are not expressed in the DVMs.

5.3. The *TT-lacZ* expression in the tracheal system of larvae, prepupae and pupae..

In the transgenic lines of the *TT-lacZ* series, β-galactosidase is also verified in the tracheal system of third instar larvae, prepupae and pupae. Even though the employed histochemical assay is not suitable for quantification analyses, the observation of both larvae and early prepupae suggested that β-galactosidase expression levels were higher in the tracheal system of early prepupae when compared to third instar larvae (data not sown). Accordingly, RT-qPCR experiments revealed that the relative levels of *lacZ* mRNA expression in the trachea of early prepupae are higher than those present in the trachea of third instar larvae. The relative expression levels of *lacZ* mRNA in the tracheal system at later

developmental times was not investigated. In this sense, the detection of β-galactosidase activity in the trachea of 1h and 2 h prepupae and 24 hours pupae could be either due to the continuous transcription of the reporter gene, or alternatively, consequence of the high stability of the reporter protein. The investigation of the *lacZ* mRNA relative levels of expression in the trachea of older prepupae and pupae should reveal if the *TT-lacZ* transgene drives tracheal gene expression at later developmental stages.

The relative levels of *CG13711*, *CG12493* and *CG32239* (*Gef64C*) mRNA expression were investigated in third instar larvae and prepupae. In the tracheal system the relative levels of expression of *CG12493* decrease in prepupae when compared to the levels detected in third instar larvae. On the other hand, the relative levels of expression of both *CG13711* and *CG32239* (*Gef64C*) increase in early prepupae, similar to what is observed for the *lacZ* reporter gene. In this context, in the larval and prepupal tracheal system the *TT* CRM might be regulating the expression of either *CG13711* or *CG32239* (*Gef64C*) or both.

Even though it is possible that the *TT* CRM drives *CG13711* expression in the tracheal system, all the data available about this *CG* indicates that this gene is only expressed in the blastoderm during the early stages of development, (Bonneaud et al., 2003; Gross et al., 2003). Moreover, the expression of *CG13711* in the larval tracheal system has been investigated through RNAseq and the expression of this *CG* has not been detected in the larval tracheal system (Bonneaud et al., 2003; Gross et al., 2003; Graveley et al., 2011). Since the increase in the relative expression levels of *CG13711* is very modest, only about 1.5 fold, and RTqPCR experiments are very sensitive, we might be detecting very low levels of transcription of this *CG* due to the fact that *CG13711* is located in one of the introns of *CG32239* (*Gef64C*) that is also expressed in the prepupal tracheal system. This suggestion is further supported by the observation that both *CG13711* and *CG32239* (*Gef64C*) are transcribed in the same direction and therefore when *CG32239* (*Gef64C*) is transcribed there will also be *CG13711* transcripts, which will be included in the RNA molecules that are spliced out during the post-transcriptional processing of *CG32239* (*Gef64C*) mRNA.

We suggest that in the larval, prepupal and pupal tracheal system the *TT* CRM regulates *CG32239* (*Gef64C*) mRNA expression. The increase in the relative expression levels of *lacZ* mRNA (about 3 fold) are accompanied by an increase in the relative expression levels of *CG32239* (*Gef64C*) mRNA (about 2.3 fold). Moreover, differently from *CG13711*,

low levels of expression of *CG32239* (*Gef64C*) mRNA expression have also been detected in the tracheal system through RNAseq experiments (Graveley et al., 2011).

We have previously proposed that the *TT* CRM contains elements that drive *CG12493* expression in the DLMs in adult thoraxes. The *TT* CRM is located about 400 bp upstream *CG12493*, in the proximal promoter region of this *CG*. We suggest that the *TT* CRM also contains *cis*-regulatory elements that regulate the expression of *CG32239* (*Gef64C*) in the larval and prepupal tracheal system. The *TT* CRM is localized in one of the introns of *CG32239* (*Gef64C*). CRMs are not always located upstream gene sequences and intronic CRMs have been previously described (Zimmerman et al., 1994; Jeziorska et al., 2009). Furthermore, even though the function of *CG32239* (*Gef64C*) has been extensively investigated in various developmental processes (Bashaw et al., 2001; Simões et al., 2006; Lovegrove, 2006; Massarwa et al., 2009; Greenberg and Hatini 2010), CRMs that regulate the expression of this gene had not been previously characterized.

The role of *CG32239 (Gef64C)* in the formation of the tracheal system and other tubular structures in the embryo has been comprehensively investigated (Bashaw et al., 2001; Simões et al., 2006; Lovegrove, 2006; Massarwa et al., 2009). On the other hand, the role of *CG32239 (Gef64C)* during the larval, prepupal and pupal stages has not been intensively characterized. During the three larval instars the general topology of the tracheal system is not extensively altered. The modifications that take place constitute an increase in the number of trachea and tracheoles in the consecutive stages. During metamorphosis the tracheal system is extensively remodeled leading to the formation of the adult tracheal system that is markedly distinct from the larval tracheal system (Whitten 1957). In this context it is reasonable to propose that the pattern of expression of the *TT-lacZ* transgene in the larvae to prepupae transition might be revealing that *CG32239 (Gef64C)* also participates in the remodeling of the tracheal system during metamorphosis. *In situ* hybridization experiments employing an antisense probe complementary to *CG32239 (Gef64C)* in whole larvae and prepupae should reveal if this *CG* is also expressed in the prepupal tracheal system.

6 REFERENCES

B ALBERTS, A JOHNSON, J LEWIS, M. RAFF, K. ROBERTS, AND P. W. Molecular biology of the cell, 6th edition by B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. In: Biochemistry and Molecular Biology Education. [s.l: s.n.]. v. 36p. 171.

ALLEN, M. J. et al. Making an escape: Development and function of the Drosophila giant fibre systemSeminars in Cell and Developmental Biology, 2006.

ANANT, S.; ROY, S.; VIJAYRAGHAVAN, K. Twist and Notch negatively regulate adult muscle differentiation in Drosophila. Development (Cambridge, England), v. 125, p. 1361–1369, 1998.

ANTONIEWSKI, C. et al. Direct repeats bind the EcR/USP receptor and mediate ecdysteroid responses in *Drosophila melanogaster*. *Mol. Cell. Biol*. 16(6): 2977-2986, 1996.

ALLEN M. J.; GODENSCHWEGE T. A. Making an escape: development and function of the Drosophila giant fibre system. Semin Cell Dev Biol. 17:31–41, 2006.

de ALMEIDA, J. C. A 28-fold increase in secretory protein synthesis is associated with DNA puff activity in the salivary gland of *Bradysia hygida* (Diptera, Sciaridae). *Braz. J. Med. Biol. Res.* 30, 605-614, 1997.

ASHBURNER M. et al. Temporal control of puffing activity in polytene chromosomes. Cold Spring Harb Symp Quant Biol 38:655–662, 1974

BASHAW, G. F.; H. NOBES C.D. and GOODMAN, C.S. A novel Dbl family RhoGEF promotes Rho-dependent axon attraction to the central nervous system midline in *Drosophila* and overcomes Robo repulsion. *The Journal of Cell Biology*, 155, n7. 1117- 1122, 2001.

BASSO JÚNIOR, L. R et al. The induction of DNA puff *BhC4–1* gene is a late response to the increase in 20- hydroxyecdysone titers in last instar dipteran larvae. *Mech Dev*, v. 110, p.15–26, 2002.

BASSO JÚNIOR, L. R. et al. *Broad-Complex*, *E74*, and *E75* early genes control DNA puff *BhC4-1* expression in prepupal salivary glands. *Genesis*, v. 44, p. 505–514, 2006.

BATE, M.; RUSHTON, E. Myogenesis and muscle patterning in Drosophila. C R Acad Sci III, v. 316, n. 9, p. 1047–1061, 1993.

BEN-TABOU DE-LEON, S.; DAVIDSON, E. H. Gene regulation: gene control network in development. Annual review of biophysics and biomolecular structure, v. 36, p. 191, 2007.

BERNSTEIN S. I.; O'DONNELL P. T.; CRIPPS R. M. Molecular genetic analysis of muscle development, structure and function in Drosophila. Int Rev Cytol 143:63-152, 1993.

BLACKWOOD, E. M. Going the Distance: A Current View of Enhancer Action. Science, v. 281, n. 5373, p. 60–63, 1998.

BONNEAUD, N. et al. SNCF, a SoxNeuro interacting protein, defines a novel protein family in Drosophila melanogaster. Gene, v. 319, n. 1-2, p. 33–41, 2003.

BUSTIN, S. et al. The MIQUE guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry* 55:4, 2009.

BUSTIN S. and NOLAN T. Part II: Chapter 7. Primers and Probes. pag 279-319 in A-Z of Quantitative PCR. Edited by Stephen Bustin, 2009.

CAMPOS NEVES, M. Caracterização funcional da região repressora proximal do promotor do gene de pufe de DNA *BhC4-1*. Masters, University of São Paulo, 2009.

CANDIDO-SILVA J. A.; PALERMO D.; RODRIGUES G.; de ALMEIDA J. C. Indirect immune detection of ecdysone receptor (EcR) during the formation of DNA puffs in *Bradysia hygida* (Diptera, Sciaridae). *Chromosome Research* 16: 609–622, 2008.

CANDIDO-SILVA J. A. et al. Amplification and expression of a salivary gland DNA puff gene in the prothoracic gland of *Bradysia hygida* (Diptera: Sciaridae). Journal of Insect Physiology 74: 30-37, 2015.

CONSORTIUM, T.F. The FlyBase database of the *Drosophila* genome projects and community literature. Nucleic Acids Res. 31, 172– 175, 2003.

COSTELLO, W. J.; WYMAN, R. J. Development of an indirect flight muscle in a muscle-specific mutant of Drosophila melanogaster. Developmental Biology, v. 118, n. 1, p. 247–258, 1986.

DANEHOLT B. Structural and functional analysis of Balbiani ring genes in the salivary glands of Chironomus tentans. In: King R, Akai H, editors. Insect ultra-structure. New York: Springer Science1Business Media. pp 382–401, 1982.

DAVIDSON E. H. The regulatory genome: gene regulatory networks in development and evolution. *Academic press*, 2005.

EDWARD, A. C. et al. C. Mutation in many genes affects aggressive behavior in *Drosophila melanogaster*. *BMC Biology* 7:29: 1-13, 2009.

FERNANDES, J.; BATE, M.; VIJAYRAGHAVAN, K. Development of the indirect flight muscles of Drosophila. Development (Cambridge, England), v. 113, p. 67–77, 1991.FERNANDES, J. et al. Muscle development in the four-winged Drosophila and the role of the Ultrabithorax gene. Current Biology, v. 4, n. 11, p. 957–964, 1994.

FERNANDES, J.; VIJAYRAGHAVAN, K. The development of indirect flight muscle innervation in Drosophila melanogaster. Molecular Biology, v. 227, p. 215–227, 1993.

FONTES, A. M. et al. The DNA puff BhB10-1 gene encodes a glycine-rich protein secreted by the late stage larval salivary glands of Bradysia hygida. Gene, v. 231, n. 1-2, p. 67–75, 1999.

FRANK H. Análise da interação entre o fator de transcrição E74A de *Drosophila melanogaster* com o módulo *cis*-regulador (*2186/-58*) de *BhC4-1* no sistema mono-híbrido de levedura. Masters, FMRP, University of São Paulo, 2014.

FURTADO, I. Caracterização do padrão de expressão de BhC4-1-lacZ em linhagens de Drosophila melanogaster transformadas com diferentes versões do ativador de glândula anelar do gene *BhC4-1*. Masters, University of São Paulo, 2013.

GARCIA, A. C. et al. Functional characterization of the sciarid *Bhc4-1* core promoter in transgenic *Drosophila*. *BMC Mol. Biol.* 12:32, 2011.

GERBASI, V. R. et al. Blanks, a nuclear siRNA/dsRNA-binding complex component, is required for Drosophila spermiogenesis. Proceedings of the National Academy of Sciences of the United States of America, v. 108, n. 8, p. 3204–3209, 2011.

GRAVELEY, B. R. et al. The developmental transcriptome of Drosophila melanogaster. Nature, v. 471, n. 7339, p. 473–9, 2011.

GROSS, S. P. et al. A Determinant for Directionality of Organelle Transport in Drosophila Embryos. Current Biology, v. 13, n. 19, p. 1660–1668, 2003.

HARDISON, R. C.; TAYLOR, J. Genomic approaches towards finding cis-regulatory modules in animals. Nature Reviews Genetics, v. 13, n. 7, p. 469–483, 2012.

JAFFE, A. B. AND HALL, A. Rho GTPases: biochemistry and biology. Annu. Rev. Cell Dev. Biol. 21, 247-269, 2005.

JEZIORSKA, D. M.; JORDAN, K. W.; VANCE, K. W. A systems biology approach to understanding cis-regulatory module functionSeminars in Cell and Developmental Biology, 2009.

KING, R. C.; AGGARWAL, S. K.; and BODENSTEIN, D. The comparative submicroscopic morphology of the ring gland of *Drosophila melanogaster* during the second and third larval instars. *Z Zellforsch Mikrosk Anat* 73(2): 272-285, 1996.

LACINE, E. M. et al. Development of DNA puffs and pattern of polypeptide synthesis in the salivary glands of *Bradysia hygida*. *Chromosoma* 89, 280-284, 1984.

LASKO, P. The Drosophila melanogaster genome: Translation factors and RNA binding proteins. Journal of Cell Biology, v. 150, n. 2, 2000.

LARA, F.J.; STOCKER, A.J.; AMABIS, J.M. DNA sequence amplification in sciarid flies: results and perspectives. *Braz J Med Biol Res* 24(3): 233-248, 1991.

LAWRENCE, P. A. Cell lineage of the thoracic muscles of drosophila. Cell, v. 29, n. 2, p. 493–503, 1982.

LECCI, M. S. et al. Functional and bioinformatics analyses reveal conservation of *cis*-regulatory elements between sciaridae and drosophilidae. *Genesis* 46: 43-51, 2008.

LIEW G. M.; FOULK M. S.; GERBI S. A. The ecdysone receptor (ScEcR-A) binds DNA puffs at the start of DNA amplification in *Sciara coprophila. Chromosome Res* 21:345–360, 2013.

LIS, J. T.; SIMON, J. A.; SUTTON, C. A. New heat shock puffs and β-galactosidase activity resulting from transformation of *Drosophila* with an hsp70-lacZ hybrid gene. *Cell* 35: 403-410, 1983.

LIVAK K. J. and SHMITTGEN T. D. Analysis of relative gene expression data using Real-Time Quantitative PCR and the $2^{-\Delta\Delta Ct}$ Method. *METHODS* 25, 402–408, 2001.

LOVEGROVE, B. et al. Coordinated Control of Cell Adhesion, Polarity, and Cytoskeleton Underlies Hox-Induced Organogenesis in Drosophila. Current Biology, v. 16, n. 22, p. 2206–2216, 2006.

LUNDGREN, J. et al. Identification and characterization of a Drosophila proteasome regulatory network. Molecular and cellular biology, v. 25, n. 11, p. 4662–75, 2005.

MALTA, T.M., RUIZ, J.C., MONESI, N. Functional characterization and bioinformatics analysis of the *BhC4-1* ring gland enhancer in transgenic Drosophila. Anais do 16 th International Ecdysone Workshop, página 25, 2006.

MONESI, N. et al. Molecular characterization of an 18 Kb segment of DNA puff C4 of *B. hygida* (Diptera: Sciaridae). *Chromosoma* 103, 715-724, 1995.

MONESI, N.; JACOBS-LORENA, M.; PAÇÓ-LARSON, M. L. The DNA puff gene BhC4–1 of *Bradysia hygida* is specifically transcribed in early prepupal salivary glands of *Drosophila melanogaster*. *Chromosoma*, v. 107, p. 559–569, 1998.

MONESI, N.; SOUSA, J. F. and PAÇÓ-LARSON, M.L. The DNA puff *BhB10-1* gene is differentially expressed in various tissues of *Bradysia hygida* late larvae and constitutively transcribed in transgenic *Drosophila*. *Brazilian Journal of Medical and Biological Research* vol 34: 851-859, 2001.

MONESI, N.; BASSO-JR L. R. and PAÇÓ-LARSON, M. L. Identification of regulatory regions in the DNA puff *BhC4-1* promoter. *Insect Mol Biol* 12: 247-54, 2003.

MONESI, N. et al. Regulation of sciarid DNA puffs by ecdysone: mechanisms and perspectives: in Ecdysone: structures and functions (ed. E. Smagghe), pp. 165-183. Springer Science + Business Media B. V., 2009.

MASSARWA, R.; SCHEJTER, E. D.; SHILO, B. Z. Apical Secretion in Epithelial Tubes of the Drosophila Embryo Is Directed by the Formin-Family Protein Diaphanous. Developmental Cell, v. 16, n. 6, p. 877–888, 2009.

PAÇÓ-LARSON, M. L. et al. Cloning of a developmentally amplified gene sequence in the DNA puff C4 of *Bradysia hygida* (Diptera: Sciaridae) salivary glands. *Insect. Biochem. Molec. Biol.* 22(5), 439-446, 1992.

PHILLIPS, J. M. Part II: Chapter 2. Real Time RT-PCR: What Lies Beneath the Surface pag 47-78 in A-Z of Quantitative PCR/edited by Stephen A. Bustin, 2009.

PONTON, F. et al. Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*. *Journal of Insect Physiology* 57, 840-850, 2011.

PTASHNE, M. Gene regulation by proteins acting nearby and at a distance. Nature, v. 322, n. 6081, p. 697–701, 1986.

ROY, S.; SHASHIDHARA, L. S.; VIJAYRAGHAVAN, K. Muscles in the Drosophila second thoracic segment are patterned independently of autonomous homeotic gene function. Current biology : CB, v. 7, p. 222–227, 1997.

ROY, S.; VIJAYRAGHAVAN, K. Homeotic genes and the regulation of myoblast migration, fusion, and fibre-specific gene expression during adult myogenesis in Drosophila. Development (Cambridge, England), v. 124, n. 17, p. 3333–3341, 1997.

ROY, S.; VIJAYRAGHAVAN, K. Muscle pattern diversification in Drosophila: The story of imaginal myogenesisBioEssays, 1999.

RYDLANDER L. AND EDSTROM JE. Large sized nascent protein as dominating component during protein synthesis in *Chironomus* salivary glands. Chromosoma 81:85–99, 1980.

SAMAKOVLIS, C. et al. Development of the Drosophila tracheal system occurs by a series of morphologically distinct but genetically coupled branching events. Development (Cambridge, England), v. 122, n. 5, p. 1395–1407, 1996a.

SAMAKOVLIS, C. et al. Genetic control of epithelial tube fusion during Drosophila tracheal development. Development (Cambridge, England), v. 122, n. 11, p. 3531–3536, 1996b.

SANCHEZ-HIGUERAS, C.; SOTILLOS, S.; CASTELLI-GAIR HOMBRIA, J. Common origin of insect trachea and endocrine organs from a segmentally repeated precursor. Current Biology, v. 24, n. 1, p. 76–81, 2014.

SATO, M.; KORNBERG, T. B. FGF is an essential mitogen and chemoattractant for the air sacs of the Drosophila tracheal system. Developmental Cell, v. 3, n. 2, p. 195–207, 2002.

SAUNDERS, L. R.; BARBER, G. N. The dsRNA binding protein family: critical roles, diverse cellular functions. FASEB journal : official publication of the Federation of American Societies for Experimental Biology, v. 17, n. 9, p. 961–83, 2003.

SIMON C. R.; SIVIERO F.; MONESI N. Beyond DNA puffs: What can we learn from studying Sciarids?. *Genesis* 00:00-00, 2016.

SIMÕES, S. et al. Compartmentalization of Rho regulators directs cell invagination during tissue morphogenesis. *Development*, 133: 4257-4267, 2006.

STOCKER, A. J. et al. Antibodies against the D-domain of a *Chironomus* ecdysone receptor protein react with DNA puff sites in *Trichosia pubescens Chromosoma*, 106, pp. 456–464, 1997.

SYMONS, M. AND SETTLEMAN, J. Rho family GTPases: more than simple switches. Trends Cell Biol. 10, 415-419, 2000

THUMMEL CS. Flies on steroids-Drosophila metamorphosis and the mechanisms of steroid hormone action. Trends Genet 12:306–310, 1996.

WANG, J. C.; GIAEVER, G. N. Action at a distance along a DNA. Science (New York, N.Y.), v. 240, p. 300–304, 1988.

WHITTEN, J. The Post-embryonic Development of the Tracheal System in Drosophila melanogaster. J Microscopical Sci, v. s3-98, p. 123–150, 1957.

WINTER, C. E. et al. The gigant DNA puffs of *Rhynchosciara Americana* code for polypeptides of the salivary gland secretion. *Journal of Insect Physiology* 23: 1455-1459, 1977a.

WINTER, C. et al. Relationships between newly synthetized proteins and DNA puff patterns in salivary glands of *Rhynchosciara Americana*. *Chromosoma* 61(3): 193-206, 1977b.

ZAMPAR, P. V.; TAVARES, S. S.; and MONESI, N. Functional validation of potential ring gland enhancers identified in the *Drosophila melanogaster* genome. In *IV Simpósio Internacional de Biologia do Desenvolvimento* Taubaté – SP, 2009.

ZIMMERMAN, L. et al. Independent regulatory elements in the nestin gene direct transgene expression to neural stem cells or muscle precursors. **Neuron**, v. 12, n. 1, p. 11–24, 1994.