

# FUNÇÕES DO FATOR DE TRANSCRIÇÃO *ZELDA* EM INSETOS

LUPIS RIBEIRO GOMES NETO

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY  
RIBEIRO - UENF

CAMPOS DOS GOYTACAZES- RJ

JUNHO – 2017

# FUNÇÕES DO FATOR DE TRANSCRIÇÃO ZELDA EM INSETOS

LUPIS RIBEIRO GOMES NETO

Tese apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Biociências e Biotecnologia.

Orientador: Prof. Dr. Thiago Motta Venancio

Co-orientador: Prof. Dr. Rodrigo Nunes da Fonseca

UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE DARCY  
RIBEIRO - UENF

CAMPOS DOS GOYTACAZES- RJ

JUNHO – 2017



# FUNÇÕES DO FATOR DE TRANSCRIÇÃO ZELDA EM INSETOS

LUPIS RIBEIRO GOMES NETO

Tese apresentada ao Centro de Biociências e Biotecnologia da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Biociências e Biotecnologia.

Orientador: Prof. Dr. Thiago Motta Venancio

Co-orientador: Prof. Dr. Rodrigo Nunes da Fonseca

**Comissão Examinadora:**

---

Prof. Doutor Thiago Motta Venancio - UENF

---

Profa. Doutora Antônia Elenir Amâncio Oliveira - UENF

---

Prof. Doutor Carlos Jorge Logullo de Oliveira - UENF

---

Prof. Doutor Pedro Lagerblad de Oliveira – UFRJ

CAMPOS DOS GOYTACAZES – RJ

JUNHO – 2017

## **AGRADECIMENTOS**

Ao meu filho Pedro, que após seu nascimento mudou todo o mote da minha vida. Pelos momentos de felicidade proporcionados em meio ao estresse.

À minha família, pela motivação das mais variadas maneiras possíveis. Minha mãe Bernadete, pelo apoio e esperança em seu filho. Meu pai Lucio pelo apoio financeiro e paciência por todo esse período de aprendizado. Meu irmão Rian pelo companheirismo e por ter desempenhando tão bem a função de tio por esse tempo de doutoramento.

Aos professores Thiago Venâncio e Rodrigo Nunes da Fonseca por terem aceitado a tarefa de me orientarem durante o doutorado. Aproveito este parágrafo para citar outros professores que foram importantíssimos nessa caminhada: Jackson Menezes, Natalia Feitosa, Fábio Olivares, Gustavo Rezende, José Nepomuceno, José Roberto, Carlos Logullo e o pós-doutorando Evenilton Pessoa.

Aos colegas de laboratório da UENF: Daniel, Cris, Ana Laura e Newton, os mais recentes Raj, Kahnu, Gustavo e Filipe pelas discussões, contribuições e troca de experiências. Agradeço também aos colegas de outros grupos: Eduardo, Kayan, Gustavo, João e Helena por terem compartilhado momentos especiais durante a minha readaptação a UENF.

Aos colegas de laboratório de NUPEM: Vitória, Marcelle, Rebeca, Victor, Dani, Ariadne, Paula, Luiz, Bruno, Felipe, João Paulo, Renata, Geórgia pelos “galhos quebrados”, ajuda nos experimentos e oportunidades de ensinar e me aperfeiçoar ainda mais nas técnicas de laboratório.

Dedico a todos os funcionários do NUPEM/UFRJ e da UENF, que também passaram por momentos difíceis devido a atrasos no pagamento durante o período.

À professora Olga Machado e as secretárias do Programa de Pós Graduação em Biociências e Biotecnologia.

Às agências de fomento FAPERJ, CAPES e CNPq pelo suporte no projeto.

## SUMÁRIO

1. INTRODUÇÃO .....	1
1.1. Considerações Gerais .....	1
1.2. Fatores de Transcrição .....	1
1.3. A Transição Materno Zigótica .....	2
1.4. <i>Zelda</i> .....	4
1.5. O desenvolvimento inicial de <i>Tribolium castaneum</i> .....	5
1.6. Padronização do eixo anteroposterior: <i>D. melanogaster</i> e <i>T. castaneum</i> .....	7
1.7. Silenciamento gênico por RNAi em <i>T. castaneum</i> .....	11
2. JUSTIFICATIVA .....	11
3. OBJETIVO GERAL .....	12
3.1. Objetivos específicos .....	12
4. METODOLOGIA .....	12
4.1. Identificação de genes ortólogos e análises filogenéticas .....	12
4.2. Cultivo dos besouros e obtenção larvas .....	13
4.3. Fixação de embriões para imunohistoquímica e hibridização <i>in situ</i> .....	13
4.4. Imunohistoquímica de embriões de <i>T. castaneum</i> .....	14
4.5. Western Blotting .....	14
4.6. Extração de RNA total .....	15
4.7. Síntese de DNA Complementar .....	15
4.8. Desenho de iniciadores ( <i>primers</i> ) para PCR .....	16
4.9. Amplificação do <i>zelda</i> em <i>T. castaneum</i> .....	16
4.10. Síntese dos Templates para dsRNA .....	16
4.11. Injeções embrionárias de dsRNA .....	18
4.12. Silenciamento por RNAi parental .....	18
4.13. Sequenciamento de Nova Geração .....	19
4.14. Confeção de sondas de RNA anti-senso e hibridização <i>in situ</i> .....	21
4.15. Experimentos com larvas de <i>T. castaneum</i> .....	23
4.16. Análises morfológicas .....	24
4.17. Cálculo do Índice próximo-distal .....	25
4.18. Estatística .....	26
5. RESULTADOS E DISCUSSÃO .....	26
5.1. Identificação de domínios conservados e a origem de <i>zld</i> .....	26
5.2. <i>Zelda</i> é expresso na região da zona de crescimento em <i>T. castaneum</i> .....	30
5.3. A ativação zigótica em <i>T. castaneum</i> .....	32

5.4.	A predição da interação de <i>zelda</i> no genoma de <i>T. castaneum</i> .....	34
5.5.	Expressão de alvos preditos de Tc-Zld no embrião de <i>T. castaneum</i> .....	43
5.6.	<i>Tc-zld</i> desempenha papéis específicos na região posterior de crescimento	46
5.7.	Análise do Transcriptoma de embriões silenciados de <i>Tc-zld</i> .....	50
5.8.	Identificação de genes diferencialmente expressos por análise do transcriptoma .....	52
5.9.	Zelda possui um papel conservado no hemíptero <i>Rhodnius prolixus</i> .....	54
5.10.	A função pós embrionária de <i>Tc-zld</i> .....	57
6.	Considerações finais .....	67
7.	CONCLUSÕES .....	69
8.	REFERÊNCIAS BIBLIOGRÁFICAS .....	70
9.	ANEXOS .....	78

## ÍNDICE DE FIGURAS E TABELAS

<b>Figura 1: A transição materno zigótica em diferentes animais.</b> Todos os animais sofrem a transição materno-zigótica que consiste na degradação dos mRNAs depositados pela mãe e a síntese de novos mRNAs pelo zigoto, existindo variação no tempo em que a degradação e síntese irão ocorrer. As bandas vermelhas indicam degradação dos produtos maternos, em azul claro a primeira onda de transcrição zigótica, em azul escuro o estabelecimento dos transcritos zigóticos. (Tadros e Lipshitz, 2009).....	4
<b>Figura 2: Eventos do desenvolvimento embrionário de <i>T. castaneum</i>.</b> Quadro ilustrando da formação do blastoderma ao início dos movimentos musculares de embriões de <i>T. castaneum</i> e seus tempos de desenvolvimento a 35 °C e temperatura ambiente ( <i>room temperature</i> , RT) Adaptado de (Strobl <i>et al.</i> , 2015). .....	7
<b>Figura 3: Ovos de insetos de diferentes tipos de banda germinal.</b> Os ovos estão apresentados pela vista lateral com a parte anterior para cima. A área sombreada representa a banda germinal. Orthoptera: a, <i>Oecanthus pellucens</i> ; b, <i>Acheta domesticus</i> . Odonata: c, <i>Platycnemis pennipes</i> . Hemiptera: d, <i>Euscelis plebejus</i> . Coleoptera: e, <i>Atrachya menetriesi</i> ; f, <i>Leptinotarsa decemlineata</i> ; h, <i>Bruchidius obtectus</i> . Lepidoptera: g, <i>Bombyx mori</i> . Diptera: i, <i>Smittia sp.</i> ; j, <i>Drosophila melanogaster</i> ; k, <i>Calliphora erythrocephala</i> . Hymenoptera: l, <i>Apis melífera</i> . (Davis e Patel, 2002). .....	10
<b>Figura 4: Cálculo do índice PD para asas e élitros.</b> (A) e (B) as medidas de largura “B” e “C”, e comprimento “A” de élitro e asa utilizados para o cálculo do índice PD. (C) A fórmula utilizada para o cálculo do índice PD e a demonstração entre as relações dos valores do índice e a forma das estruturas. Adaptado de (Ravisankar <i>et al.</i> , 2016). .....	25
<b>Figura 5: Alinhamento múltiplo de sequências de proteínas <i>Zelda</i> representando os principais grupos dos artrópodes.</b> Alinhamento realizado com o programa <i>muscle</i> (Edgar, 2004) e visualização feita a partir do programa JalView (Waterhouse <i>et al.</i> , 2009). .....	28
<b>Figura 6: Características conservadas das proteínas <i>Zelda</i>.</b> Asteriscos e “x” indicam presença e ausência, respectivamente. O ponto de interrogação denota que a característica da sequência está parcialmente preservada ou ausente devido a erros na montagem do genoma. O cladograma foi organizado de acordo com o estudo filogenético realizado por (Misof <i>et al.</i> , 2014). Dois grupos externos com indivíduos ausentes de ortólogos de <i>zld</i> foram incluídos. ....	29
<b>Figura 7: A expressão de <i>Tc-zld</i> é de origem materno e posteriormente confinada na zona de crescimento durante o desenvolvimento embrionário.</b> (A) A expressão relativa de <i>Tc-zld</i> no ovário e carcaça. (B) Expressão relativa de <i>Tc-zld</i> 0-3, 3-6 e 6-9 horas após a oviposição. A expressão foi normalizada em relação ao gene constitutivo <i>rps3</i> em ambos os experimentos. Hibridização <i>in situ</i> para o mRNA de <i>Tc-zld</i> ao longo do desenvolvimento de <i>T. castaneum</i> . (C) Transcritos de <i>Tc-zld</i> são uniformemente distribuídos no embrião no estágio de pré-blastoderma (0-3 horas);(D) DAPI do respectivo estágio. (E) No estágio de blastoderma celular (3-6 horas), transcritos começam a ocupar o primórdio de banda germinal;(F) marcação nuclear por DAPI. (G) Pouco tempo após a invaginação, a expressão de <i>zld</i> começa a se concentrar na região posterior de onde a zona de crescimento irá se formar (6-9	



horas). (H) pouco após o início do alongamento da banda germinal, a expressão de *Tc-zld* é confinada à zona de crescimento. ....31

**Figura 8: A expressão de *Tc-zelda* durante os estágios mais tardios do desenvolvimento.** (A) No início da extensão da banda germinal, ou seja, aproximadamente 13 horas após a oviposição, *Tc-zld* é expresso no região posterior de crescimento (Seta). (B) No período de 18-21 horas após oviposição, *Tc-zld* mantém sua expressão na região mais posterior do embrião (seta), além de ser expresso nos lobos cefálicos (pontas de setas) e segmento da gnata (asterisco). (C) Aproximadamente 39-48 horas após a oviposição, durante o fechamento dorsal, a expressão pode ser observada ao longo sistema nervoso (seta). ....32

**Figura 9. A transição materno zigótica (MZT) em *Tribolium castaneum*.** (A-C) Marcação nuclear com DAPI em embriões de *Tribolium* entre 0-1 horas (A), 3-6 horas (B) e 6-9 horas após a oviposição (C). (D,H) Western blots de extratos de embriões de 0-1, 0-3, 3-6, 6-9 horas após oviposição utilizando anticorpos para a forma transcricionalmente ativa de RNA pol II como descrito por (Nestorov *et al.*, 2013). (D) ou anticorpo contra a proteína  $\alpha$ -tubulina utilizado como controle (H). ....33

**Figura 10: Predição computacional e identificação de genes alvos de *Tc-Zld*:** (A) Comparação feita pelo software TOMTOM do motivo obtido com o programa DREME no genoma de *Tribolium castaneum* e o motivo de ligação de *Zelda* em *D. melanogaster*. (B) Diagrama de Venn dos genes regulados durante a MZT descritos por (Harrison *et al.*, 2011) e os genes preditos pelo programa FIMO utilizando o motivo predito de *D. melanogaster* (amarelo). (C) Diagrama de Venn com os alvos de *Zld* em *D.melanogaster* (Harrison *et al.*, 2011) em verde e os ortólogos do tipo 'one-to-one' de genes alvos preditos para *Tc-Zld*. ....35

**Figura 11: Padrão de expressão do gene TC008343.** (A,A') Expressão do gene no estágio inicial da embriogênese de *T. castaneum*, durante as primeiras divisões mitóticas (A'). (B-B') O gene TC008343 é expresso na banda germinal tendo maior presença na região posterior no início da gastrulação.(C-C') Em momentos pré-fechamento da janela da serosa, o gene TC008343 mantém sua expressão na banda germinal e na região da zona de crescimento, não sendo expresso em células extra-embriônicas. ....44

**Figura 12: Padrão de expressão do gene TC014638.** (A-A') O gene TC014638 é expresso na banda germinal momentos antes da gastrulação (A). Vista lateral do embrião de *T. castaneum* no estágio de blastoderma diferenciado evidenciado pelos núcleos marcados por DAPI (A') (B-B') Visão ventral do embrião de *T. castaneum* evidenciando a marcação por hibridização in situ para o gene TC014638 na banda germinal durante o movimento de fechamento da janela da serosa. ....45

**Figura 13: Dinâmica de expressão do gene *Tc-Toll7* (TC004474).** (A,A') Embrião de *T. castaneum* no estágio pré-blastoderma celular (A') não apresenta expressão de *Tc-Toll7*. (B,B') *Tc-Toll7* exibe expressão semelhante em grandes faixas no estágio de blastoderma celular uniforme. (C,C') Na sequência, durante o blastoderma diferenciado, *Tc-Toll7* tem sua expressão na região posterior reduzida. (D,D') No momento da invaginação, *Tc-Toll7* é expresso no embrião na cabeça e na região que dará origem ao mesoderma. (E,E') A expressão de *Tc-Toll7* começa a se localizar nos segmentos à medida em que são formados. (F,F') Durante o processo de segmentação, *Tc-Toll7* é expresso na cabeça e nos parasegmentos do embrião. ....46

**Figura 14: O silenciamento embrionário de *Tc-zld* interrompe a segmentação.** A injeção de solução de dsRNA de *Tc-zld* nos embriões expressando GFP afetou a segmentação a partir da zona posterior de crescimento (B), quando comparado ao embrião controle em pleno estágio de banda germinal alongada (A). Escala: 250  $\mu$ m.

.....47

**Figura 15: O RNAi parental de *Tc-zld* afeta genes importantes para a padronização do embrião de *T. castaneum*.** (A,B) A expressão de *Tc-zen* na região da serosa é reduzida após o pRNAi de *Tc-zld* (B, seta) quando comparado ao controle injetado com ds*neo*. (C) A expressão do gene de padronização dorsoventral *Tc-cactus* é reduzida após o pRNAi de *Tc-zld* (D, seta). (E,F) A expressão do gene *Tc-toll7* sobrepõe-se a zona de crescimento na região posterior no embrião controle (E), enquanto sua expressão na região posterior do embrião se encontra dispersa em pequenos grupamentos no embrião silenciado para *Tc-zld* (F). (G,H) O gene de padronização dorsoventral *Tc-sog* é expresso na região ventral do embrião controle (G), porém no embrião ds*zld*, um grande domínio perde parte de sua expressão como indicado pela seta (H). (I,J) *Tc-eve* é expresso em faixas na região posterior de crescimento, porém no embrião silenciado ds*zld* sua expressão não obedece os limites das faixas. (K,L) A segmentação proveniente da zona de crescimento sofre interrupção de após silenciamento por ds*zld*, o padrão de expressão do gene de parasegmentos *Tc-gsb* é presente em todos os segmentos do embrião controle (I) enquanto no silenciado, somente a expressão até os segmentos torácicos é detectada, tendo o embrião falhado em se segmentar (K).....49

**Figura 16: Densidade de expressão gênica em FPKM para as amostras controle (EC) e ds*zld* (ZLD).** O gráfico demonstra a distribuição da expressão gênica entre as diferentes amostras. Onde  $\log_{10} \text{fkpm} 0 = 1 \text{ FPKM}$ .....51

**Figura 17: Dendograma da distância JS demonstrando as relações entre as replicatas de cada tratamento.** Esta análise calcula a divergência no nível de expressão de genes (em FPKM) entre as amostras. Amostras similares tendem a se agrupar, neste caso replicatas silenciadas (ZLD\_1, ZLD\_0 e ZLD\_2) e amostras controle (EC\_1, EC\_0 e EC\_2). .....52

**Figura 18: Mapa de calor com os níveis transcricionais dos genes classificados como *Segmentation* (GO:0035282) pela análise de DAVID (Huang *et al.*, 2009)** ..As cores azul e vermelha indicam os valores de  $\log_{10}$ FPKM mais aproximados de 2,5 e zero, respectivamente. Nomes dos genes acompanhados de "XLOC\_" ,identificadores adotados pelo programa cufflinks para regiões delimitadas no genoma. ....54

**Figura 19: O gene *zld* é requerido para a geração dos segmentos posteriores de *Rhodnius prolixus*.** (A,C,E) Embriões controle (ds*neo*) de *R. prolixus*. (B,D,F,F') Embrião proveniente de fêmea injetada com dsRNA de *Rp-zld*. (F,F') Embrião dentro do córion, vista ventral e dorsal. (C,E) Embrião removido do ovo. (G,H) Representação esquemática dos embriões controle e *zld* pRNAi. (D) O asterisco indica o olho, observado na região ventral do embrião (F). (D) Após a dissecação, o olho pode ser identificado por sua pigmentação vermelha. Escala: 0,5 mm. ....55

**Figura 20: Expressão de *Tc-zld* e potenciais genes alvo relacionados a metamorfose:** (A-D) Morfologia da larva de *T. castaneum* no terceiro (A-L3), quinto (B-L5), sétimo instares larvares (C-L7) e primeiro estágio de pupa (D-P1). (E) Expressão relativa de *zld*, *ubx*, *wingless* (*wg*), *distalless* (*dll*), *Iroquois* (*iro*), *lim-1* e

*kruppel-homolog-1* (*krh-1*) nos estágios L3, L5, L7 e P1. O RT-PCR em tempo real foi normalizado pela expressão do gene *Tc-rps3* como descrito por (Lord *et al.*, 2010). Asteriscos representam diferenças significativas entre os estágios ( $P < 0,05$ ). ..... 59

**Figura 21: Eficiência do RNAi de *Tc-zld* no estágio pós embrionário:** Larvas de *T. castaneum* foram injetadas no quinto instar larval como descrito na literatura (Tomoyasu e Denell, 2004). (A) Expressão relativa do gene *Tc-zld* após injeção de dsRNAs de *Tc-zld* e *neo*. O RT-PCR em tempo real foi normalizado pela expressão do gene *Tc-rps3* como descrito por (Lord *et al.*, 2010). (B) Número de besouros adultos que emergiram onze dias após silenciamento larval de *Tc-zld*. ..... 60

**Figura 22: O silenciamento de *Tc-zld* afeta diferentes estágios da vida do besouro:** Larvas de *T. castaneum* foram injetadas com soluções de dsRNA nos estágios mais jovem (3<sup>o</sup>) e mais tardio (6<sup>o</sup>) da larva. (A,B) Morfologia da pupa obtida pela injeção de *dsneo* e *dszld*. Pupa de *dszld* apresentando pigmentação diferencial na região da cabeça (seta) e asas reduzidas, enquanto a pupa *dsneo* apresenta características do tipo selvagem. (C,D) Adultos obtidos por injeção no estágio larval por injeções de *dsneo*(C) e *dszld*(D), onde as asas não se dobram corretamente sob os élitros. .... 60

**Figura 23: Élitros e asas de besouros injetados com dsRNA de *Tc-zld* apresentam menor resistência a fixação:** Os apêndices dorsais de *T. castaneum* provenientes de larvas injetadas com *dsneo* e *dszld* foram dissecados e fixados em etanol 95% por 24 h. (A,B) O élitro dissecado de besouros *dszld* (B) apresenta rugas (seta) e menor resistência física que a estrutura do besouro controle (A). (C) A asa membranosa obtida do besouro *dsneo* apresenta um padrão de venação e pigmentação característica. (D) Asas destacadas de besouros *dszld* e submetidas à fixação em etanol apresentam maior fragilidade e ressecamento drástico, assim como maior melanização na região distal (seta). Escala: 1 mm. .... 62

**Figura 24: RNAi de *Tc-zld* nos estágios larvares afeta élitros e asas no estágio adulto:** (A) Élitros extraídos do besouro controle *dsneo* apresentam o padrão de venação paralelo e uma estrutura rígida. (B) O élitro do besouro silenciado *dszld* apresenta uma perturbação no padrão de venação e menor rigidez. (C,D) Asas dissecadas não apresentaram diferenças observáveis em sua morfologia. (E) Índices PD de comparação entre élitros e asas *dsneo* e *dszld*. O índice PD exprime a forma da asa baseada em razões de suas dimensões (Ravisankar *et al.*, 2016), indicando que o silenciamento de *Tc-zld* afeta o formato das asas membranosas. (F) O comprimento das asas e élitros de *dszld* não apresenta diferenças significativamente estatísticas quando comparado às respectivas estruturas em *dsneo*. As análises estatísticas foram realizadas utilizando teste-t sobre a variância ( $* = P < 0,0001$ ). .... 64

**Figura 25: O knockdown (silenciamento) de *Tc-zld* induz fenótipos alterados nas antenas e patas de adultos de *T. castaneum*:** (A) Expressão relativa do gene *Tc-Dll* após injeção de *dsneo* e *dszld*. (B-D) Antenas de indivíduos controle apresentam o número normal de segmentos: dois proximais (escapo e pedicelo), seis intermediários (flagelo) e três segmentos distais diferenciados formando a clava. (E,F) A injeção de *dszld* causou a fusão dos segmentos da clava enquanto em fenótipos mais “fortes “ a antena não se segmenta, desenvolvendo-se como um rudimento de antena. As patas de *T. castaneum* possuem seis podômeros: Coxa (Cx), trocânter (Tr), fêmur (Fe), Tibia (Ti), Tarso (T) e pós-tarso (ou pré-tarso (pT)). (D) As patas mesotorácicas de indivíduos injetados com *dsneo* apresentam cinco

segmentos tarsais, ou tarsômeros. (E) Em indivíduos silenciados para *Tc-zld*, alguns segmentos do tarso são perdidos ou fundidos, resultando em redução do tarso, enquanto outras estruturas da pata aparentemente não são afetadas. ....66

## TABELAS

Tabela 1: Componentes da reação de síntese de cDNA. ....	15
Tabela 2: Concentração e volume dos componentes da reação padrão de PCR. ....	17
Tabela 3: Parâmetros utilizados no termociclador para a realização do PCR. ....	17
Tabela 4: Oligonucleotídeos iniciadores utilizados nos ensaios de qRT-PCR. ....	24
Tabela 5: Lista de 141 alvos de Zld durante a MZT de <i>D.melanogaster</i> que possuem ortólogos <i>one-to-one</i> entre os alvos preditos de Tc-Zld. ....	35
Tabela 6: Análise funcional com termos de ontologia dos genes listados na Tabela 4 realizada pela ferramenta DAVID. ....	39
Tabela 7: Sumário do alinhamento das leituras no genoma de <i>T. castaneum</i> . Dados de alinhamento das replicatas utilizadas no experimento controle EC e silenciado ZLD. ....	50

**LISTA DE ABREVIATURAS**

AP – anteroposterior

cDNA – DNA complementar

CRM – cis regulatory modules (módulos regulatórios em cis)

DAPI – 4',6 diamino-2-phenylindole

Dm - *Drosophila melanogaster*

DV – dorsoventral

eRNAi – RNAi embrionário

FPKM – Fragmentos por quilobase por milhões de leituras mapeadas

GRN – gene regulatory networks

miRNA – microRNA

mRNA – RNA mensageiro

MZT - Transição materno zigótica

nGFP – nuclear GFP

Pb – pares de bases

PBS – tampão fosfato salino

PD – próximo-distal

pRNAi – RNAi parental

PVDF – *polyvinylidene difluoride* (fluoreto de polivinilideno)

RNA – ácido ribonucleico

RNA pol II – RNA polimerase II

RNAi – RNA de interferência

RT- *room temperature*

Tc- *Tribolium castaneum*

TF – fator de transcrição

TSS – sítio de início de transcrição

*zld* – *zelda* (gene)

Zld – Zelda (proteína)

## RESUMO

Redes regulatórias gênicas (Gene regulatory networks, GRNs) evoluem como um resultado de processos que agem sobre fatores de transcrição (TFs) e os módulos cis-regulatórios (CRM) aos quais se ligam. Um TF do tipo dedo de zinco (*zinc finger*) denominado *zelda* (*zld*) é essencial para a transição materno-zigótica (MZT) em *Drosophila melanogaster*, onde se liga diretamente a milhares de CRMs para regular a acessibilidade da cromatina. *D. melanogaster* possui o tipo de desenvolvimento característico de insetos de banda germinal longa, onde todos os segmentos são gerados simultaneamente ao longo do ovo. Entretanto, ainda não se encontra descrito o envolvimento de *zld* na MZT de insetos de banda germinal curta, assim como sua participação em outros processos biológicos. Neste trabalho, demonstramos que *zld* é uma inovação que surgiu nos Pancrustacea, sendo ausente em artrópodes mais distantes e em outros organismos. Para entender melhor a função ancestral de *zld*, investigamos suas funções em um inseto de banda curta, o besouro *Tribolium castaneum*, utilizando técnicas de biologia molecular, sequenciamento de segunda geração e abordagens computacionais. Nossos resultados demonstram que o papel de *zld* não está restrito apenas à MZT, participando também da segmentação posterior do embrião e da padronização de estruturas derivadas de discos imaginiais no inseto adulto. Além disso, demonstramos que *zld* é necessário para a segmentação posterior no hemíptero *Rhodnius prolixus*, indicando que o surgimento de parte de suas funções precede da origem dos insetos holometábolos e foi perdida em insetos de banda germinal longa. Nossos resultados demonstram novas funções de *zld* na manutenção do estado pluripotente das células progenitoras na região posterior e sugerem que mudanças na expressão de *zld* e provavelmente outros TFs pioneiros foram críticas para a evolução das GRNs de insetos.

## ABSTRACT

Gene regulatory networks (GRN) result from the evolution of transcription factors (TF) and the cis-regulatory modules (CRMs) which they bind to. The zinc-finger TF *zelda* (*zld*) is essential for maternal zygotic transition (MZT) in *Drosophila melanogaster*, when it directly binds over thousand CRMs and regulates chromatin accessibility. *D. melanogaster* displays a long germ type of embryonic development, where all segments are simultaneously generated along the whole egg length. It remains unclear if *zld* is also involved in MZT of short-germ insects (including those from basal lineages), as well as in other biological processes. Here we show that *zld* is an innovation of the Pancrustacea lineage, being absent in more distant arthropods (e.g. chelicerates) and other organisms. To better understand *zld*'s ancestral function, we thoroughly investigated its roles in a short-germ insect, the beetle *Tribolium castaneum*, using molecular biology and computational approaches. Our results demonstrate roles for *zld* not only during the MZT, but also in posterior segmentation and patterning of imaginal disc derived structures. Further, we also demonstrate that *zld* is critical for posterior segmentation in the hemipteran *Rhodnius prolixus*, indicating that this function predates the origin of holometabolous insects and was subsequently lost in long-germ insects. Our results unveil new roles of *zld* in maintaining pluripotent state of progenitor cells at the posterior region and suggest that changes in expression of *zld* (and probably other pioneer TFs) are critical in the evolution of insect GRNs.

# 1. INTRODUÇÃO

## 1.1. Considerações Gerais

O desenvolvimento de um organismo multicelular a partir de uma única célula, o ovo fertilizado, é um triunfo da evolução. Durante o desenvolvimento embrionário, milhões de células são geradas a partir de um único ovo. Ao longo desse processo, essas células vão formar estruturas variadas e complexas. A partir desses eventos, uma série de questões fundamentais são levantadas: Como as células provenientes de uma mesma origem se diferenciam tanto umas das outras? Qual o princípio de organização incluído no desenvolvimento do organismo no que diz respeito a material genético? Muitas dessas perguntas têm sido respondidas a partir do estudo da Biologia do Desenvolvimento.

Com o desenvolvimento das técnicas de biologia molecular nas últimas décadas, atualmente já se compreende muito sobre como a informação contida no DNA direciona os principais processos da embriogênese, incluindo como atuam os programas de regulação da expressão gênica, que constituem extensas redes regulatórias formadoras do plano corporal do indivíduo. No presente trabalho demonstraremos como um fator de transcrição fortemente atuante no início da embriogênese pode desempenhar funções em outros processos biológicos e em estágios distintos do desenvolvimento.

## 1.2. Fatores de Transcrição

A expressão gênica é regulada pela ação de fatores de transcrição (TFs) que são proteínas que possuem domínios de ligação ao DNA. TFs geralmente se ligam a regiões específicas do genoma, ativando ou reprimindo a transcrição de genes adjacentes. Os TFs exercem sua função a partir da interação com regiões regulatórias em cis, que incluem promotores e elementos proximais (Grossman *et al.*, 2017), assim como módulos localizados a maiores distâncias do sítio de início de transcrição (TSS), como por exemplo, os *enhancers*.

Além de regular a expressão gênica, alguns TFs operam na facilitação do acesso de outros TFs ao DNA, sendo chamados portanto de TFs pioneiros. Em eucariotos, os genes se encontram contidos e organizados em um complexo de DNA e proteína denominado cromatina. O nucleossomo, a unidade básica da



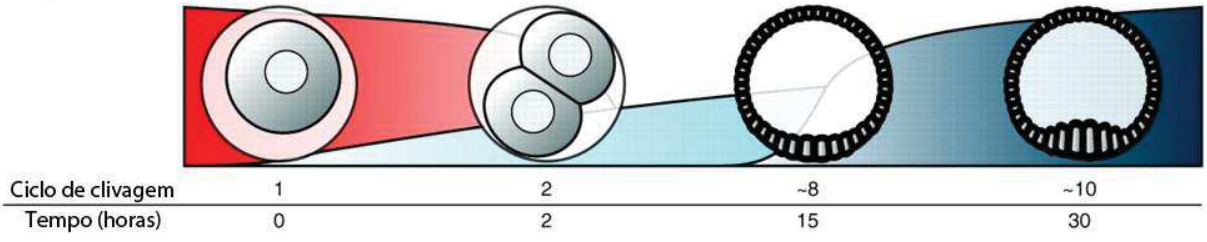
cromatina, é composto por um octâmero de proteínas histonas envolto duas vezes por alças contendo aproximadamente 200 pares de bases (pb) de DNA, ocasionando uma restrição a qualquer interação entre esse DNA e outras proteínas (Lorch e Kornberg, 2015). Os TFs pioneiros tem a propriedade de superar tais restrições, permitindo que outros TF alcancem a cromatina outrora inacessível. Um momento importante onde os TFs pioneiros desempenham sua função é na transcrição dos primeiros genes no início da embriogênese, durante o processo de ativação zigótica. A ativação zigótica é o segundo ato do fenômeno presente em todos os metazoários sexuais, onde os produtos do genoma maternal dão lugar a RNAs e proteínas de origem zigótica, no evento denominado transição materno zigótica (Drouin, 2014).

### **1.3. A Transição Materno Zigótica**

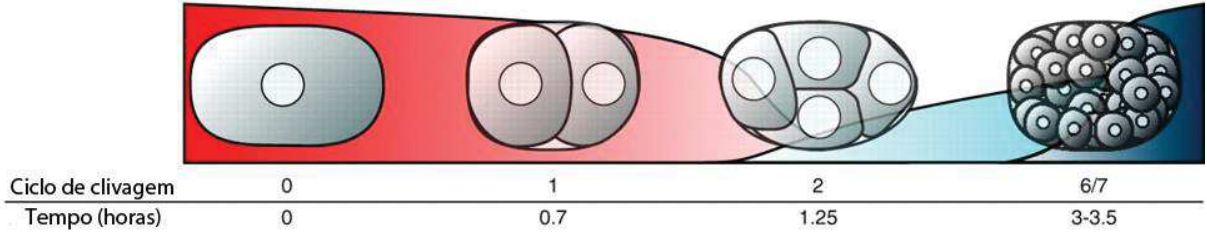
Fatores derivados do genoma maternal controlam praticamente todos os acontecimentos dos primeiros momentos do desenvolvimento. Transcritos e proteínas estocados no citoplasma do ovócito são de importância fundamental para assegurar o desenvolvimento embrionário inicial até o momento em que o genoma zigótico é ativado (De Renzis *et al.*, 2007). Este evento é denominado transição materno-zigótica (MZT, do inglês *maternal zygotic transition*). Dois processos principais caracterizam a MZT: a de gradação de um conjunto de RNAs mensageiros (mRNAs) maternos e a ativação do genoma zigótico. Neste momento, a transcrição a partir do genoma zigótico leva à produção de microRNAs (miRNAs) que aumentam ainda mais a degradação de transcritos maternos, garantindo maior eficiência do processo. Além disso, os primeiros mRNAs transcritos no embrião são provenientes de genes que codificam ativadores transcricionais, que por sua vez aumentam a eficiência da transcrição zigótica (Walser e Lipshitz, 2011).

A MZT é um evento que ocorre em diversos organismos e já foi amplamente estudada em equinodermos, nematóides, insetos, peixes, anfíbios e mamíferos. Em *Drosophila melanogaster*, a MZT é concluída por volta de 2,5 horas após a oviposição, enquanto no camundongo, por exemplo, ela ocorre no estágio de duas células, por volta de 22 horas após a fecundação (Figura 1) (Tadros e Lipshitz, 2009).

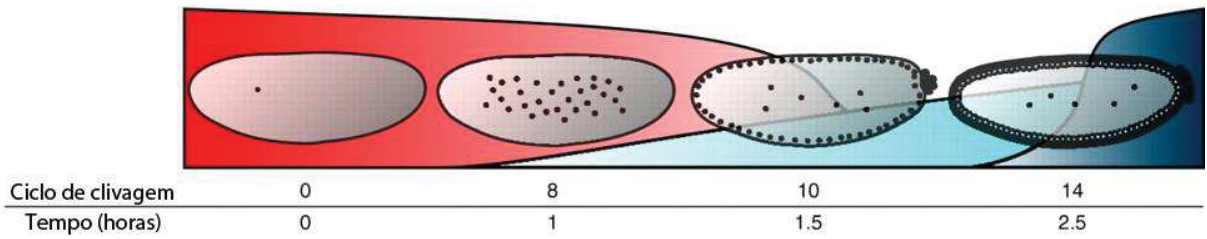
Ouriço-do-mar (*S. purpuratus*)



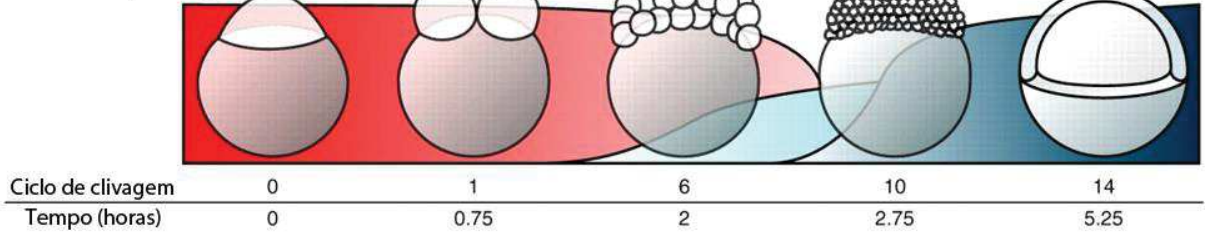
Nematoide (*C. elegans*)



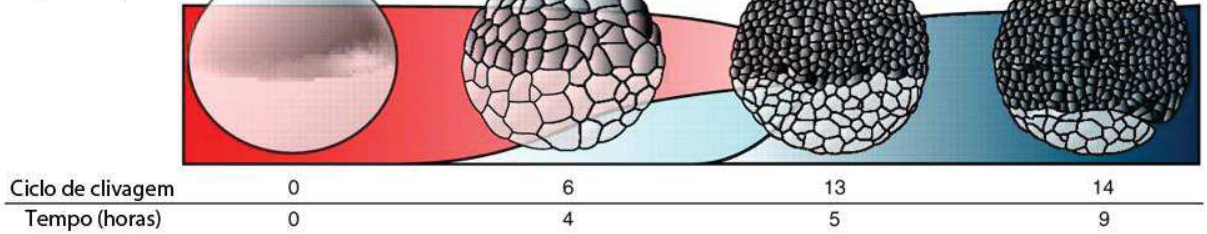
Mosca da fruta (*D. melanogaster*)



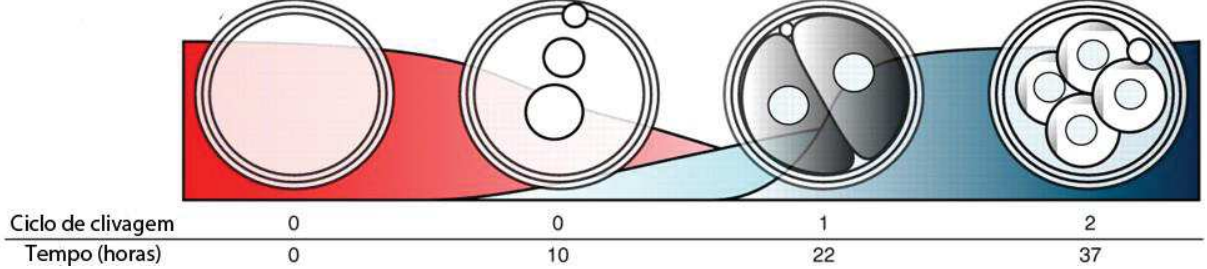
Peixe zebra (*D. rerio*)



Rã (*X. laevis*)



Camundongo (*M. musculus*)



**Figura 1: A transição materno zigótica em diferentes animais.** Todos os animais sofrem a transição materno-zigótica que consiste na degradação dos mRNAs depositados pela mãe e a síntese de novos mRNAs pelo zigoto, existindo variação no tempo em que a degradação e síntese irão ocorrer. As bandas vermelhas indicam degradação dos produtos maternos, em azul claro a primeira onda de transcrição zigótica, em azul escuro o estabelecimento dos transcritos zigóticos. (Tadros e Lipshitz, 2009).

Nas etapas iniciais da embriogênese de *D. melanogaster*, as proteínas e mRNAs depositados no ovócito durante a ovogênese são extremamente necessárias para as primeiras divisões nucleares. Entre os ciclos mitóticos 10 e 14, concomitantemente com o evento da celularização, ocorre a MZT (Tadros e Lipshitz, 2009). Os primeiros genes zigóticos ativados são determinantes para o desenvolvimento do embrião particularmente na definição dos destinos celulares e padronização de eixos dorsoventral e anteroposterior. Para que estes genes sejam expressos, o TF *zelda* atua como um dos principais ativadores zigóticos já descritos. A ação dessa proteína nas primeiras horas do desenvolvimento permite a transcrição eficiente destes genes.

#### 1.4. *Zelda*

O TF *Zelda* (*Zld*) foi identificado pela primeira vez no embrião de *D. melanogaster* onde atua como ativador global da expressão zigótica (Liang *et al.*, 2008). O mRNA materno de *zld* é necessário para a ativação da transcrição dos primeiros genes zigóticos e *Zld* é exigido para o início do desenvolvimento embrionário (Liang *et al.*, 2008). Embriões deficientes de *zld* materno morrem durante a MZT e apresentam dificuldade de celularização (Liang *et al.*, 2008). *Zld* se liga a sequências denominadas elementos TAGteam (Liang *et al.*, 2008; Harrison *et al.*, 2010; Nien *et al.*, 2011). Esses elementos se encontram enriquecidos em promotores e *enhancers* de genes expressos durante a MZT (Ten Bosch *et al.*, 2006; De Renzis *et al.*, 2007). A ocupação de *Zld in vivo* é regulada pela sequência de nucleotídeos TAGteam, sendo 64% desses a sequência canônica CAGGTAG, ocupada por *Zld* no embrião em início de desenvolvimento (Harrison *et al.*, 2011; Nien *et al.*, 2011).

Em *D. melanogaster*, a maior isoforma de *zld* codifica uma proteína de 1596 aminoácidos, compreendendo seis domínios do tipo dedo de zinco C2H2 e regiões de sequências altamente repetitivas (Staudt *et al.*, 2006; Liang *et al.*, 2008; Hamm *et al.*, 2015). Sua região carboxi-terminal possui quatro dos seis domínios

dedo de zinco, os quais realizam a atividade de ligação com o domínio específico de DNA. Estes domínios são bastante conservados em outros insetos e construções baseadas em ortólogos de Zld de *Nasonia vitripennis* e *Anopheles gambiae* são capazes até mesmo de ativar a transcrição por interação com TAGteam em células S2 de *D. melanogaster* (Hamm *et al.*, 2015). Além de sua função no início da embriogênese, Zld também está relacionado à regulação da expressão de genes nos estágios tardios da embriogênese (por exemplo, em processos como formação do tecido nervoso) e provavelmente pós-embrionários devido a sua presença na fase larvar e a sua expressão nos discos imaginais, tecidos que darão origem aos apêndices na fase adulta (Pearson *et al.*, 2012; Giannios e Tsitilou, 2013). Recentemente foi demonstrada a expressão de *zld* também em embriões no estágio inicial de *Apis mellifera*, reforçando a conservação de sua expressão durante a MZT em insetos (Pires *et al.*, 2016).

O besouro *Tribolium castaneum* emergiu nas últimas décadas como modelo de genética de insetos alternativo a *D. melanogaster* (Brown *et al.*, 2009). Os estudos com *Tribolium* alcançaram o sucesso devido à facilidade de cultivo, tempo de desenvolvimento relativamente curto e seu *status* de praga de estocagem de grãos e outros produtos. Os primeiros trabalhos sobre *Tribolium* abordaram principalmente aspectos sobre ecologia, genética de populações e genética quantitativa. (Park, 1934; Levene *et al.*, 1965; Englert e Bell, 1970). Quanto à nutrição, comportamento e fisiologia, muitas informações foram publicadas em estudos condensados por Alexander Sokoloff (Sokoloff, 1972).

De longe o grupo mais bem sucedido entre os insetos, os besouros possuem diferentes adaptações, como por exemplo, luminescência, protrusão de líquidos defensivos e mimetismo (Engelhardt *et al.*, 1965; Hackett *et al.*, 1992; Hafernik e Saul-Gershenz, 2000). No ambiente de laboratório, o *Tribolium* requer cuidados considerados mínimos se comparados a outros insetos. Os besouros não necessitam de suprimento de água, vivem em ambientes com grande amplitude de umidade e podem sobreviver a temperaturas que variam de 20 a 37 °C. Normalmente os experimentos são realizados com besouros cultivados a 30 °C, condições sob as quais apresentam maior fecundidade e maior taxa de sobrevivência (Sokoloff, 1972).

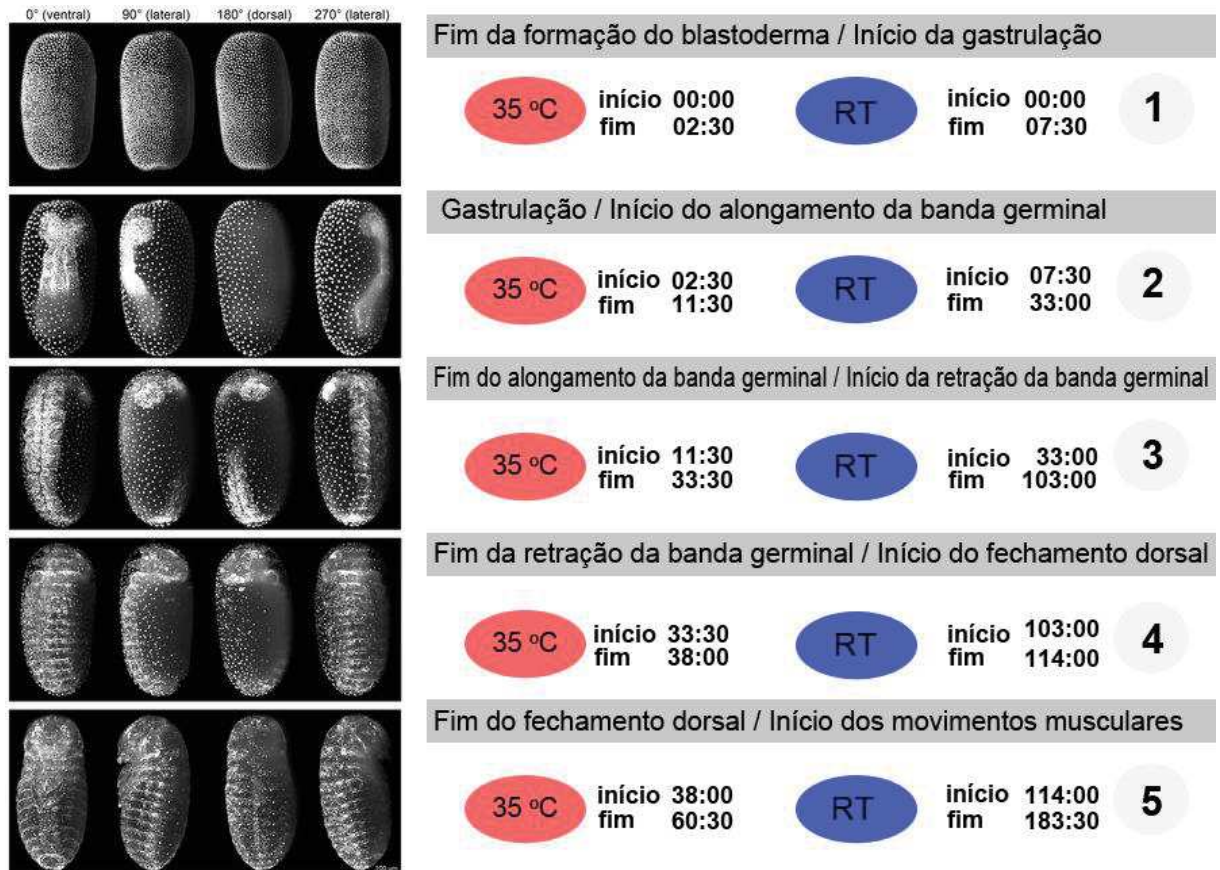
### **1.5. O desenvolvimento inicial de *Tribolium castaneum***

Os ovos de *T. castaneum* tem aproximadamente meio milímetro e possuem um leve estreitamento na região posterior. De maneira similar à *Drosophila*, *T. castaneum* apresenta as divisões nucleares mitóticas iniciais no interior do ovo sem o evento da citocinese. Após as divisões, os núcleos migram para a periferia (Figura 2, quadro 1), e ocorre a celularização e a formação do blastoderma celular. Em seguida, o embrião de *T. castaneum* passa por reorganizações celulares que resultam na formação do rudimento embrionário (Figura 2, quadro 1). Uma invaginação na região posterior, a linha primitiva, é formada à medida que células se agregam na região posterior ventral do ovo. Este agregado de células forma o embrião, enquanto as células periféricas remanescentes formam a cobertura extraembrionária. O lado ventral da linha primitiva se torna a parte caudal do embrião enquanto o lado dorsal contribui com a amnioserosa que irá revestir o embrião ventralmente. O primórdio embrionário desliza ao longo da região ventral do ovo e a linha primitiva se move anteroventralmente para longe do polo posterior. Conforme a região caudal do primórdio embrionário adentra o vitelo, o sulco digestivo é formado em uma progressão postero-anterior para gerar o mesoderma. Concomitantemente a amnioserosa se fecha na região ventral do embrião, também na direção postero-anterior. Finalmente, os lobos cefálicos se formam conforme a camada amnioserosa se dobra a partir dos limites laterais anteriores (Figura 2, quadro 2).

O rudimento embrionário resultante deste processo ocupa 80% do ovo e se encontra na região ventral, com os lobos da cabeça localizados na região anterior. Durante a fase subsequente da embriogênese, este rudimento se alonga e se diferencia na chamada banda germinal segmentada (Figura 2, quadro 2). Enquanto isso a região cefálica se desenvolve e se estende pelo polo anterior do ovo. Simultaneamente a banda germinal se alonga e alcança o polo posterior.

Durante o alongamento da banda germinal, o embrião já demonstra alguns sinais de segmentação. Conforme a banda germinal é formada, os segmentos mandibulares e maxilares já são perceptíveis próximos ao protoencéfalo já são perceptíveis, apesar de ainda existir uma região não segmentada denominada zona de crescimento. A segmentação ocorre de maneira rápida no sentido anteroposterior a partir da região torácica. A banda germinal continua a se estender e segmentar até que o plano corporal do embrião esteja completo (Figura 2, quadro 2). O próximo

passo envolve a retração da banda germinal até que o embrião atinja a mesmo comprimento do ovo. Durante a extensão e retração do embrião brotos de apêndices se formam em alguns segmentos (Figura 2, quadro 3). Por fim, o embrião cresce dorsalmente encobrindo o vitelo remanescente.



**Figura 2: Eventos do desenvolvimento embrionário de *T. castaneum*.** Quadro ilustrando da formação do blastoderma ao início dos movimentos musculares de embriões de *T. castaneum* e seus tempos de desenvolvimento a 35 °C e temperatura ambiente (*room temperature*, RT) Adaptado de (Strobl *et al.*, 2015).

Durante os estágios da embriogênese supracitados, alguns genes desempenham um papel fundamental na regulação do crescimento e desenvolvimento do organismo. Um exemplo de evento cujos fatores atuantes são bem estudados é a padronização de eixos embrionários.

### 1.6. Padronização do eixo anteroposterior: *D. melanogaster* e *T. castaneum*

Em *D. melanogaster*, o inseto mais bem estudado, três sistemas derivados maternalmente dispõem informação espacial que progressivamente subdividem o embrião em seu eixo anteroposterior (AP): os sistemas anterior, posterior e terminal. Dois desses sistemas, o anterior e posterior, dependem de mRNAs responsáveis

pela polaridade interna do ovócito (*bicoid* e *nanos*, respectivamente), enquanto o sistema terminal se basta de modificações na casca do ovo durante a ovogênese para exercer sua função na padronização (Porcher *et al.*, 2010).

O eixo AP de *D. melanogaster* é padronizado pela expressão de uma hierarquia de genes que progressivamente aumentam a resolução e a precisão da informação posicional. Esse processo se inicia com genes maternos (*bicoid* e *nanos*) que provém informação ampla e graduada em ambos os pólos do embrião. Esses gradientes são interpretados pelos genes do tipo gap (os primeiros genes de segmentação zigoticamente ativados). Os domínios de expressão desses genes subdividem o embrião em blocos e disponibilizam informação de curto alcance. Genes do tipo gap (*hunchback* e *Krüppel*) codificam TFs e são denominados gap devido ao seu fenótipo de perda de função em *D. melanogaster*, onde regiões do embrião contendo vários segmentos são deletadas (Nusslein-Volhard e Wieschaus, 1980). O padrão espacial dos genes maternos e do tipo gap é responsável pela definição dos padrões de expressão dos genes da regra de pares ("*pair-rule*").

Os genes da regra de pares interpretam esses gradientes de curto alcance e são expressos em uma periodicidade de segmentos duplos para dar as primeiras formas ao plano corporal da mosca-da-fruta. Os genes da regra dos pares codificam TFs e foram caracterizados originalmente por seus fenótipos de perda de função em *D. melanogaster*, neste caso, deleções na cutícula do embrião ocorrendo periodicamente de dois em dois segmentos. A expressão dos genes da regra dos pares é consistente com seus fenótipos, sendo tipicamente expressos em faixas em uma periodicidade de dois segmentos durante o desenvolvimento dos blastodermas celular e sincicial. Dentre os genes *pair-rule* conhecidos em *D. melanogaster* estão *even-skipped (eve)*, *hairy*, *runt*, *fushi-tarazu (ftz)* e *paired (prd)* (Nüsslein-Volhard *et al.*, 1987).

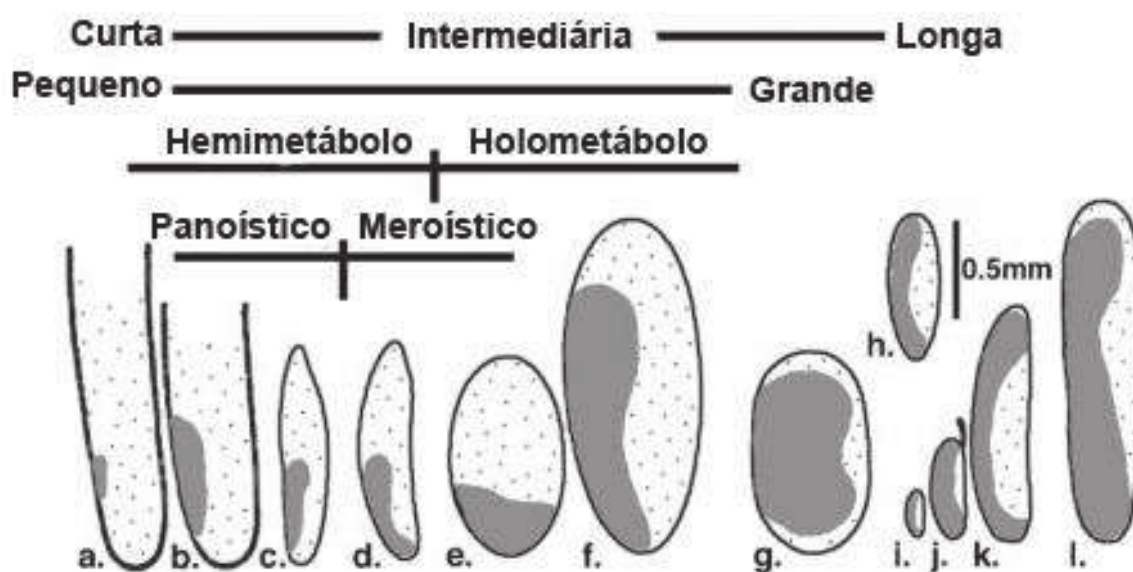
O padrão de expressão dos genes da regra dos pares é interpretado pelos genes de polaridade de segmento, que em sua maioria são expressos nos segmentos durante o início da gastrulação e através do estágio de banda germinal segmentada. Ao contrário dos genes do tipo gap e *pair-rule*, nem todos os genes de polaridade de segmento são TFs, por exemplo, alguns componentes de vias de sinalização. Dois dos genes desta classe mais estudados em insetos são *engrailed (en)* e *wingless (wg)*, que codificam um TF e uma molécula secretada,

respectivamente. Os genes de polaridade de segmentos também são denominados devido aos fenótipos de perda de função em *D. melanogaster*, onde suas funções envolvem estabelecer os destinos dos segmentos e limites entre os mesmos (Akam, 1994). Os genes de polaridade de segmento são dispostos ao longo de todos os segmentos do blastoderma de forma simultânea, refletindo o modo de segmentação de banda germinal longa

Em *T. castaneum* este sistema é substancialmente diferente daquele observado em *D. melanogaster*, o que se deve em parte à diferenças no mapa de destino do blastoderma de embriões de banda germinal curta. Por exemplo, ao invés de haver vários segmentos da cabeça que requerem altos níveis de informação de padronização para estabelecer suas fronteiras, a maior parte da porção anterior do embrião de *T. castaneum* é destinada a células extraembrionárias, destino que seria restrito a células na parte dorsal em *D. melanogaster* (Schmidt-Ott, 2005).

Embora o conhecimento do processo de formação do eixo AP em *D. melanogaster* seja bastante importante, este conhecimento não pode ser diretamente estendido a outros artrópodes basais como quelicerados ou mesmo a outros insetos holometábolos, uma vez que em *D. melanogaster* e outros dípteros a região do primórdio do embrião, a banda ou disco germinal, ocupa toda a extensão do ovo. Esta banda germinal, chamada de banda germinal longa é encontrada em Diptera (moscas, mosquitos, etc), Lepidoptera (mariposas, borboletas) Hymenoptera (vespãs, abelhas, formigas) e em alguns Coleoptera (besouros) (Davis e Patel, 2002).





**Figura 3: Ovos de insetos de diferentes tipos de banda germinal.** Os ovos estão apresentados pela vista lateral com a parte anterior para cima. A área sombreada representa a banda germinal. Orthoptera: a, *Oecanthus pellucens*; b, *Acheta domesticus*. Odonata: c, *Platycnemis pennipes*. Hemiptera: d, *Euscelis plebejus*. Coleoptera: e, *Atrachya menetriesi*; f, *Leptinotarsa decemlineata*; h, *Bruchidius obtectus*. Lepidoptera: g, *Bombyx mori*. Diptera: i, *Smittia sp.*; j, *Drosophila melanogaster*; k, *Calliphora erythrocephala*. Hymenoptera: l, *Apis mellifera*. (Davis e Patel, 2002).

Algumas diferenças importantes podem ser destacadas quando comparamos embriões de banda germinal curta e longa. Nos embriões de banda longa a região mais anterior (futura região da cabeça), e a mais posterior (futuro abdômen), se desenvolvem concomitantemente (Figura 3 g-i). Em contrapartida, embriões banda curta (Figura 3 a-f) dividem seu processo de padronização em uma fase inicial, onde os segmentos anteriores da cabeça são padronizados, e uma fase secundária, em que a zona posterior é responsável pela geração de novas células que darão origem a região abdominal (Fonseca *et al.*, 2009). Nos estágios iniciais de desenvolvimento existe uma diferença extrema no modo e na velocidade de formação das membranas plasmáticas entre os ovos com embriogênese mais lenta, correspondendo aos embriões de banda curta, e os de embriogênese rápida, como *D. melanogaster*, que apresenta banda longa (Figura 3). Enquanto nas espécies que apresentam banda curta as membranas celulares existem e limitam a difusão de TFs livremente pelo ovo, nos banda longa, TFs podem estabelecer os diversos territórios AP, que no final, corresponderão aos domínios de expressão dos genes de polaridade de segmentos. Ser um inseto de banda curta pode ser listado como um

dos fatores que contribuíram para a emergência do besouro *T. castaneum* como um modelo de estudos comparativos de biologia do desenvolvimento e evolução. Outro fator importante para o sucesso do *T. castaneum* como modelo experimental é a eficiência do RNA de interferência (RNAi), que auxilia no estudo de função gênica em todas as fases de vida do besouro (Tomoyasu e Denell, 2004; Brown *et al.*, 2009).

### **1.7. Silenciamento gênico por RNAi em *T. castaneum***

Ao contrário de *D. melanogaster*, *T. castaneum* possui uma resposta robusta ao RNAi sistêmico, tornando-o um excelente modelo para estudos de genômica funcional. Em combinação com as ferramentas de manipulação gênica, essas metodologias estabeleceram o *T. castaneum* como um dos principais modelos invertebrados para estudos genéticos e moleculares (Bucher *et al.*, 2002).

A ferramenta de RNA de interferência pode ser utilizada em qualquer estágio da vida do *T. castaneum*. Quando fêmeas (pupas ou adultas) são tratadas, a supressão dos transcritos é observada não somente das fêmeas, mas também na progênie, um fenômeno denominado RNAi parental (pRNAi). Em *T. castaneum*, a resposta típica ao RNAi é potente e sistêmica, ou seja, se espalha pelo animal após a injeção local produzindo fenótipos de silenciamento diferentes daqueles conhecidos por mutantes nulos (Brown *et al.*, 2009).

A injeção duplas fitas de RNA (dsRNAs) diretamente em ovos no estágio de pré-blastoderma celular (Brown *et al.*, 1999) pode produzir fenótipos fortes, embora a variação de embrião para embrião possa ser alta se comparada aos obtidos por RNAi parental. Embora seja útil para a produção de fenótipos de silenciamento no estágio embrionário, o processo de injeção de ovos é bastante laborioso e a taxa de sobrevivência dos embriões é de 50 a 70 %. Felizmente, a abrangência sistêmica da resposta ao RNAi torna possível o estudo de funções gênicas pela injeção de dsRNA em fêmeas adultas ou pupas e posterior obtenção de sua prole contendo os efeitos fenotípicos por até 2 semanas após a sessão de injeções (Bucher *et al.*, 2002).

## **2. JUSTIFICATIVA**

Embora existam evidências da conservação de *zelda* em outros artrópodes, sua função fora de insetos drosofilídeos permanece elusiva. No presente trabalho primeiramente analisaremos a história evolutiva de *zelda*. Em seguida investigaremos as funções distintas de *zld* nos insetos de banda curta *T. castaneum* e *Rhodnius prolixus*, empregando de ferramentas de biologia molecular. Abordaremos também as recém-descobertas funções pós-embrionárias relacionadas a este importante TF.

### 3. OBJETIVO GERAL

Investigar a função do gene *zelda* no besouro *T. castaneum*.

#### 3.1. Objetivos específicos

- Analisar a possível origem de *zld* na linhagem dos artrópodes;
- Prospectar o possível sítio de ligação de *zld* em *T. castaneum* e identificar computacionalmente os potenciais alvos deste TF;
- Investigar o papel zigótico de *Tc-zld* na extensão da banda germinal de *T. castaneum*;
- Analisar os genes diferencialmente expressos durante a gastrulação do embrião de *T. castaneum* face ao silenciamento de *Tc-zld*;
- Realizar o estudo de função em estágios pós-embrionários de *T. castaneum*;

### 4. METODOLOGIA

#### 4.1. Identificação de genes ortólogos e análises filogenéticas

As sequências da proteína Zld de *D. melanogaster* ([www.flybase.org](http://www.flybase.org)) e *T. castaneum* (<http://beetlebase.org/>) foram usadas para as buscas por BLAST contra outros genomas de artrópodes. Genes preditos foram utilizados após análise de domínios do CDD (Marchler-Bauer *et al.*, 2017) e BLAST recíproco. As sequências peptídicas previstas de Zelda foram alinhadas com a ferramenta de alinhamento múltiplo *Muscle* (Edgar, 2004) e os domínios foram preditos através do banco de dados *Pfam* (Finn *et al.*, 2016). O alinhamento foi visualizado com  *JalView* (Waterhouse *et al.*, 2009) e processado para melhor visualização com *Adobe*

*Photoshop CS4* (versão 11.0, Adobe). A filogenia foi baseada no estudo publicado por (Misof *et al.*, 2014), realizada a partir de 1478 genes codificantes entre as classes de insetos.

A parte de experimentação funcional foi realizada sob a orientação do Prof. Rodrigo Nunes da Fonseca, no Laboratório de Ciências Morfofuncionais, localizado na Universidade Federal do Rio de Janeiro - Pólo Macaé.

#### **4.2. Cultivo dos besouros e obtenção larvas**

Para os experimentos foram utilizados besouros da espécie *Tribolium castaneum* mantidos estufa tipo BOD a 30 °C (Sokoloff, 1972). Para a alimentação dos besouros *T. castaneum*, utilizou-se farinha de trigo estéril. Para a esterilização, a farinha é mantida no freezer por 24h a -20 °C e em seguida colocada em estufa por mais 24h a 50 °C. A farinha foi peneirada em malhas de 710 e 300 µM, para posterior separação de besouros e ovos, respectivamente.

Os besouros foram mantidos em caixas plásticas de aproximadamente 15 × 15 cm, com abertura de ar na tampa. A troca de farinha foi realizada de 3 em 3 dias juntamente com a troca dos potes plásticos.

Para obtermos larvas nos estágios desejados sincronizamos a colônia de acordo com o tempo de oviposição. Para isso, a farinha contida no pote dos besouros foi trocada de modo que apenas besouros adultos estivessem presentes. Os potes reservados para oviposição foram acompanhados na estufa tipo BOD por 24 horas à 30 °C. Os ovos obtidos após 24 horas foram separados em outro pote e identificados como “0-24h” e armazenados em estufa tipo BOD à 30 °C.

#### **4.3. Fixação de embriões para imunohistoquímica e hibridização *in situ***

Os embriões foram coletados e decorionados em solução de hipoclorito de sódio comercial em proporção 1:1 em água. Em seguida foram lavados e adicionados em uma mistura bifásica de solução de paraformaldeído 4% diluído em tampão fosfato salino (PBS: NaCl 8 g/L, KCl 0,2 g/L , Na<sub>2</sub>HPO<sub>4</sub> 1,42 g/L; KH<sub>2</sub>PO<sub>4</sub> 0,24 g/L ) e heptano 1:1 por 30 minutos. Após a fixação, os embriões passaram por agitação em meio metanol/heptano 1:1 para perda de membranas extraembrionárias e em seguida foram armazenados em metanol ou PBST (PBS 0,5% Tween® 20 ).

#### 4.4. Imunohistoquímica de embriões de *T. castaneum*

Os embriões foram fixados como descrito anteriormente no item 4.3, exceto pela não utilização de metanol durante o processo. A marcação foi realizada conforme descrito previamente por (Patel *et al.*, 1989). Os embriões foram incubados com anticorpo monoclonal para RNA polimerase II CTD *repeat* YSPTSPS (phospho S2) (Abcam, ab24758) à diluição de 1:100, durante 16h à 4 °C. O anticorpo *AlexaFluor 546 anti-mouse IgG* (Invitrogen) na diluição de 1:1000 foi utilizado como secundário. Os embriões foram marcados com 4',6-diamidino-2-phenilindole (DAPI) para terem seus núcleos visualizados. As imagens fluorescentes foram capturadas em microscópio invertido Leica DMI4000, utilizando os filtros FITC 404/20 para o azul e BP 576/20 para o vermelho, as imagens coloridas no software LAS AF 2.6.0 (Leica).

#### 4.5. Western Blotting

Os extratos proteicos para a análise de imunoblot foram isolados de estágios iniciais do desenvolvimento. Embriões foram homogeneizados em 200 µL de “*deyolking buffer*” (55mM NaCl, 1,8mM NaHCO<sub>3</sub>) utilizando um pistilo seguido agitação em vortex. As amostras foram centrifugadas a 300 g - por 5 minutos a 4 °C. O sobrenadante foi descartado e os precipitados foram lavados com 200 µL de “*wash buffer*” (110mM NaCl, 3,5 mM KCl, 2,7 mM CaCl<sub>2</sub>). Após outros 5 minutos de centrifugação a 300 g, os precipitados foram lavados com 300 µL de “*wash buffer*”. Após a centrifugação final a 300 g por 5 minutos, o sobrenadante foi removido e os precipitados ressuspensos em 10 µL de água deionizada. A concentração de proteínas foi medida em espectrofotômetro (absorbância 280 nm) e amostras foram diluídas apropriadamente segundo suas respectivas concentrações de forma que a quantidade de proteínas fosse igual em todas as amostras.

Foram adicionados 8 µL de solução aquosa de proteína ao 4× *SDS loading buffer* (Fermentas) e em seguida as amostras foram incubadas em gelo por 15 minutos. As soluções foram fervidas a 95 °C por 5 minutos e submetidas à eletroforese desnaturante em gel de poliacrilamida (SDS-PAGE) 8% a 200V até que a banda equivalente a 200 kDa do padrão de peso atingisse a metade do gel. A transferência para a membrana de fluoreto de polivinidileno (PVDF) foi realizada a 200 mA a 4 °C. A membrana foi bloqueada com albumina sérica bovina (BSA) 5%

em PBST por 60 minutos. Em seguida a membrana foi incubada durante aproximadamente 1h com anticorpo primário anti-pRNAPII (ab24758) à diluição de 1:250 a 4 °C. O anticorpo secundário peroxidase anti-mouse IgG foi utilizado a 1:2000 e incubado por uma hora a temperatura ambiente. A revelação foi feita com o reagente *BM Blue POD Substrate* (Roche), seguindo o procedimento fornecido pelo fabricante.

#### 4.6. Extração de RNA total

O isolamento de RNA total foi realizado utilizando o reagente *Trizol* (Life Technologies), seguindo recomendações do fabricante. O RNA total de ovos de *T.castaneum* foi extraído de 10 mg de ovos decorionados e macerados com pistilo em microtubos de fundo cônico de 1,5 mL, livres de RNAses. A amostra foi centrifugada a 12000 g por 10 minutos, o material insolúvel foi descartado e ao sobrenadante foram adicionados 500 µL de clorofórmio. Em seguida, o material foi homogeneizado por inversão, incubado a temperatura ambiente por 3 minutos, submetido à centrifugação 12000 g por 15 minutos. O sobrenadante foi recolhido e adicionado 500 µL de isopropanol. Posteriormente, a amostra foi homogeneizada por inversão e incubada a -20 °C por 1 hora. Em seguida, a amostra foi centrifugada a 12000 g por 15 minutos e seu sobrenadante foi descartado. A amostra foi lavada com 500 µL de etanol 75 %, que foi então evaporado. O precipitado foi ressuspensão em 20 µL de água estéril tratada com dietil pirocarbonato (DEPC). Um microlitro de RNA foi utilizado para dosagem e pureza do RNA total em espectrofotômetro (260nm/280nm), de acordo com (Chomczynski, 1993).

#### 4.7. Síntese de DNA Complementar

Para a síntese de DNA complementar (cDNA) foi utilizado o kit comercial *High Capacity cDNA Reverse Transcription* (Applied Biosystem), utilizando 2 µg de RNA e seguindo as recomendações do fabricante listadas abaixo.

Tabela 1: Componentes da reação de síntese de cDNA.

Componente	Volume
10X RT Buffer	2.0 µL
25X dNTP Mix (100 mM)	0.8 µL

---

<b>10X RT Random Primers</b>	2.0 $\mu$ L
<b>MultiScribe™ Reverse Transcriptase</b>	1.0 $\mu$ L
<b>Inibidor de RNase</b>	1.0 $\mu$ L
<b>H2O livre de nuclease</b>	3.2 $\mu$ L
<b>Total por reação</b>	10.0 $\mu$ L

---

#### **4.8. Desenho de iniciadores (*primers*) para PCR**

Para o desenho de oligonucleotídeos iniciadores (*primers*) foi utilizado o programa online *Primer3* (<http://frodo.wi.mit.edu/>) (Rozen e Skaletsky, 2000). Os parâmetros utilizados para a síntese de moldes para dsRNA e ribossondas foram os seguintes: O tamanho do produto deveria ter o tamanho de 700-850 pares de bases e o tamanho ótimo dos *primers* deveria ser de 20-22 pares de bases.

Foram escolhidos os pares de *primers* com as melhores características e adicionados na extremidade 5' para cada *primer* a seu sequência complementar para *primers* universais., Ou seja, para o *primer* senso foi adicionada a sequência ggccgcg à extremidade 5' e para o *primer* antisenso foi adicionada a sequência cccgggc, também na extremidade 5'.

#### **4.9. Amplificação do *zelda* em *T. castaneum***

A sequência do gene *zelda* foi obtida no banco de dados *BeetleBase* (<http://beetlebase.org>), utilizando-se o nome do gene ou seu número de identificação TC014798 (Wang *et al.*, 2007). Os *primers* foram dissolvidos a um estoque de 0.1 nanomol/microlitro (100  $\mu$ M) em água deionizada. Para a concentração de trabalho, foi feita a diluição do estoque à concentração de 2  $\mu$ M, diluindo o estoque a 1:50.

#### **4.10. Síntese dos Templates para dsRNA**

A primeira reação de PCR, realizada para a confecção dos moldes para ampliações posteriores, seguiu o seguinte protocolo:

Tabela 2: Concentração e volume dos componentes da reação padrão de PCR.

<b>Componente</b>	<b>Volume Final</b>	<b>Concentração Final</b>
<b>Tampão (5×)</b>	10 µL	1×
<b>MgCl<sub>2</sub> (25 mM)</b>	4 µL	2,0 mM
<b>dNTP (10 mM)</b>	1 µL	0,2 mM cada dNTP
<b>Primer senso</b>	5 µL	0,2 µM
<b>Primer antisenso</b>	5 µL	0,2 µM
<b>Taq DNA polimerase (5u/ µL)</b>	0,25 µL	1,25 unidades
<b>DNA</b>	1 µL	2 ng
<b>H<sub>2</sub>O</b>	Até 50 µL	

As reações serão feitas em microtubos de fundo cônico de de 0,2 mL devidamente estéreis e livres de contaminação por DNAses e RNAses. Os reagentes foram descongelados e homogeneizados, com toda a reação sendo preparada no gelo para evitar degradação dos compostos e amplificação precoce.

A reação pronta foi levada a um aparelho termociclador *Applied Biosystems 2720 Thermal Cycler* configurado com os seguintes parâmetros:

Tabela 3: Parâmetros utilizados no termociclador para a realização do PCR.

<b>Etapa</b>	<b>Temperatura</b>	<b>Tempo</b>	<b>Número de ciclos</b>
<b>Desnaturação inicial</b>	95 °C	2 minutos	1 ciclo
<b>Desnaturação</b>	95 °C	1 minuto	35 ciclos
<b>Anelamento</b>	55 °C	45 segundos	



<b>Extensão</b>	72 °C	50 segundos	
<b>Extensão final</b>	72 °C	5 minutos	1 ciclo

#### 4.11. Injeções embrionárias de dsRNA

As injeções feitas diretamente no ovo para silenciamento zigótico do gene *Tc-zld* foram realizadas de acordo com a metodologia descrita previamente (Brown *et al.*, 1999; Nunes Da Fonseca *et al.*, 2008). Embriões transgênicos expressando o gene da proteína fluorescente verde (GFP) foram coletados por uma hora e deixados a 30 °C por mais 3 horas para se desenvolverem (El-Sherif *et al.*, 2012; Sarrazin *et al.*, 2012). Os embriões foram então, decorionados em solução de hipoclorito de sódio a 2%, alinhados em uma lâmina de microscopia e cobertos com óleo mineral 700 (Sigma). Os embriões foram injetados na sua região anterior com soluções contendo dsRNA para os genes *neo* e *zld* a concentração de 1 µg/µL com ajuda do instrumento *Nanoinject II* (Drummond Scientific Company). Após a injeção os embriões foram fotografados durante 16 horas (25 °C) em intervalos de 5 em 5 minutos no microscópio invertido Leica DMI4000 utilizando filtro de fluorescência para GFP (espectro de excitação 470nm/ emissão 525 nm). Ao fim da captura, um vídeo utilizando as sequências de imagens foi criado com auxílio do software *Leica Application Suite AF* (versão 3.0.2/Leica) e posteriormente tratado no software *Windows Movie Maker* (versão 10.0/Microsoft).

#### 4.12. Silenciamento por RNAi parental

A sequência molde para transcrição in vitro do gene *Tc-zld* foi amplificada a partir do cDNA contendo adaptadores para *primers* com sequência promotoras para T7 RNA polimerase. O dsRNA foi produzido in vitro utilizando o kit comercial *Megascript T7* (Ambion). O RNAi parental foi realizado injetando-se fêmeas adultas (Bucher *et al.*, 2002). Fêmeas sexualmente maduras do besouro foram colocadas em placas de Petri sobre gelo e resfriadas por 10 minutos para que adormecessem. Usando-se uma pinça, o tórax do animal foi pressionado até que o abdômen e genitália se expusessem, possibilitando que o dsRNA do gene *Tc-zld* fosse injetado na lateral do abdômen, a uma concentração de 1 µg/µl. Como controle, foi feita a

injeção de dsRNA do gene de resistência à neomicina (*dsneo*), um gene ausente no genoma de *T. castaneum*, gerando fenótipos idênticos ao do tipo selvagem.

Após a injeção, as fêmeas do besouro foram colocadas em potes com farinha estéril e deixadas por 24h em estufa tipo BOD. Seguida a recuperação, foram adicionados machos do besouro para acasalamento e obtenção de ovos fecundados. Após o primeiro dia, os embriões foram coletados para sua devida finalidade. Os embriões utilizados para observação de padrão de expressão seguiram para a fixação enquanto os embriões utilizados para análise do transcriptoma tiveram seu RNA extraído e armazenado.

#### 4.13. Sequenciamento de Nova Geração

Para realizar o sequenciamento de RNA, amostras contendo 2 µg de RNA total de embriões controle (*dsneo/EC*) e silenciados (*dszld/ZLD*), após 7,5-12,5 horas de injeção, foram tratadas com DNase I (Ambion, Life Technologies) por 30 minutos a 37 °C seguido de inativação da enzima na temperatura de 65 °C. As amostras foram armazenadas em etanol 100% e enviadas para a unidade de sequenciamento do Laboratório Central de Tecnologias de Alto Desempenho em Ciências da Vida (LaCTAD/Unicamp), onde foram preparadas as bibliotecas de mRNA. As bibliotecas foram preparadas com o kit “TruSeq Sample Prep Kit” da Illumina. Em seguida, as bibliotecas foram quantificadas e sequenciadas na plataforma *HiSeq 2500* (Illumina) no modo *paired-end*, com leituras de 100 pb. Ao todo, foram sequenciadas 8 bibliotecas, constituindo triplicatas biológicas de RNAs extraídos de embriões controle e de embriões silenciados para *Tc-zld* e *Tc-eve*, respectivamente. Mapeamento das leituras e análise de expressão gênica

A análise do transcriptoma foi realizada de acordo com um protocolo padrão, utilizando a metodologia *Tophat/Cufflinks* (Trapnell *et al.*, 2012). Antes do mapeamento, algumas etapas de filtragem foram realizadas. Os segmentos de sequências gênicas obtidos no sequenciamento (também denominados leituras) tiveram as sequências de adaptadores removidas utilizando o software *Trimmomatic* (Bolger *et al.*, 2014). Em seguida, análises de qualidade do sequenciamento utilizando o software *FastQC* foram realizadas (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). Após as análises, foram

detectadas bases com baixa qualidade nas primeiras 12 posições dos fragmentos, que foram removidas.

A partir das leituras obtidas pelo sequenciamento, realizamos o mapeamento das sequências no genoma de referência de *T. castaneum* utilizando o programa *TopHat* v2.0.11 (Kim *et al.*, 2013). Foi utilizada a versão 3.0 e seu respectivo arquivo GTF contendo as coordenadas dos genes, ambos adquiridos no banco de dados *Beetlebase* (<http://beetlebase.org/>) (Wang *et al.*, 2007; Tribolium Genome Sequencing *et al.*, 2008).

Em seguida, o programa *Cufflinks* v2.2.1 (Trapnell *et al.*, 2010) foi utilizado (com parâmetros default) para quantificar a abundância dos transcritos contra os seus respectivos genes no genoma de referência. A abundância foi normalizada usando fragmentos por quilobase por milhões de leituras mapeadas (*Fragments Per Kilobase of transcript/gene per Million mapped reads*: FPKM). Os genes diferencialmente expressos foram identificados usando as ferramentas *cuffmerge* e *cuffdiff*, ambas do pacote *cufflinks*. Os dados foram explorados por meio do ambiente R (<http://www.r-project.org/>), com auxílio da biblioteca *cummeRbund* (Trapnell *et al.*, 2012). Para quantificar a expressão diferencial dos genes entre as amostras foi adotado o valor de corte  $\alpha = 0,05$  no nível de significância para filtrar genes com diferença significativa de expressão através do comando *getSig*. Genes com FPKM de valor maior que 1 foram considerados expressos. Ainda baseando-se no FPKM, genes com valor de *fold change* (razão de mudança) em  $\log_2$  maior que 1 entre as amostras controle e Zld foram classificados como tendo sua expressão aumentada (*up regulated*) e genes com valor de *fold change* menor que -1 foram classificados como tendo sua expressão reduzida (*down regulated*). Dessa forma, tais genes tiveram sua expressão no tratamento silenciado para *zelda* maior ou menor se comparados ao controle, respectivamente.

Os genes diferencialmente expressos foram analisados utilizando a plataforma DAVID (<http://david.abcc.ncifcrf.gov/tools.jsp>) (Huang *et al.*, 2009), visando identificar categorias funcionais enriquecidas que possam direcionar novas análises e experimentos de validação. Seguindo estas primeiras análises será possível identificar novos transcritos e possíveis candidatos para análises individuais (Roberts *et al.*, 2011). Estes genes candidatos serão utilizados para um rastreamento posterior de genes envolvidos nos eventos presentes no início do

desenvolvimento como a padronização da zona de crescimento (*growth-zone*), segmentação e morfogênese envolvendo experimentos de perda de função (RNAi) e padrão de expressão (hibridização *in situ*).

#### 4.14. Confeção de sondas de RNA anti-senso e hibridização *in situ*

Com a finalidade de descrever o padrão de expressão do gene *zelda* em *T. castaneum* e investigar se o silenciamento do mesmo afetaria o padrão de expressão de genes importantes para o desenvolvimento, foram realizados experimentos de hibridização *in situ*. Para a síntese dos moldes para sonda de RNA anti-senso, foram confeccionados os oligonucleotídeos iniciadores descritos abaixo, sendo que os mesmos apresentavam em sua região 5' uma sequência de reconhecimento para iniciadores universais que contém o sítio de ligação para a enzima RNA polimerase T7 (em minúsculo):

<b>Oligonucleotídeo/gene</b>	<b>Sequência (5' – 3')</b>
<b>Tc-zldF</b>	ggccgcggACGACGAGTACCGCTTGACT
<b>Tc-zldR</b>	cccggggcCTTACCACAGGTGTCGCAGA
<b>Tc-cactusF</b>	ggccgcggTGTCCGGGAAACAGGATAAG
<b>Tc-cactusR</b>	cccggggcCGAGTAGTGGAGGGCGGTGTA
<b>Tc-KrüppelF</b>	ggccgcggGAACGAAAAACGTGGAAGGTG
<b>Tc-KrüppelR</b>	cccggggcCAATAGTCACCGGCTTGCATC
<b>Tc-sogF</b>	ggccgcggTGGGTTTGACTCGTGTACGG
<b>Tc-sogR</b>	cccggggcACTGCGCCATGTTCAAACC
<b>Tc-Toll7F</b>	ggccgcggCTCACTTCGAAATTGCGTTTG
<b>Tc-Toll7R</b>	cccggggcCTTTTGCTGTCAATGCGAGTC
<b>Tc-twi</b>	Plasmídeo de (Handel <i>et al.</i> , 2005)
<b>Tc-gsb</b>	Plasmídeo (Davis <i>et al.</i> , 2001)
<b>Tc-zen</b>	Plasmídeo de (Van Der Zee <i>et al.</i> , 2005)

Após a primeira amplificação, para a confeção da sonda adicionamos a 1 µg produtos de PCR purificado: 1 µL tampão de transcrição 10× (Ambion T7 Maxiscript), 1 µL *DIG RNA labelling mix* (Roche), água deionizada tratada com DEPC (qsp 10 uL) e por último 1 µL (15U/ µL) da enzima RNA polimerase T7 (Ambion T7 Maxiscript). Todos estes reagentes foram adicionados em microtubos de

fundo cônico livres de contaminação por nucleases (DNase/RNase). Após rápida centrifugação, procedeu-se com a síntese por 2 horas à 37 °C. Após este tempo adicionou-se sequencialmente: 40 uL de água deionizada tratada com DEPC e 45 uL de acetato de amônio 5M (Ambion T7 Maxiscript), 150 µL de etanol absoluto e 2 µL de RNA transportador de levedura (20 mg/mL). Esta mistura foi incubada por 2 horas à -20 °C e então centrifugada à 10000 g e à 4 °C por 30 minutos. O precipitado foi lavado em etanol 70% e novamente centrifugado à 10000 g e à 4 °C por 15 minutos. O precipitado final foi então resuspendido em 50 µL de solução de hibridização. Uma vez confeccionadas as sondas de RNA-antisense, deu-se início ao processo de hibridização *in situ*.

Com a técnica “*whole mount*” de hibridização *in situ* usando sondas não radioativas é possível detectar moléculas de mRNA expressos em regiões do embrião sem a necessidade de efetuar cortes histológicos. Para isso, utilizam-se sondas de RNA antisense para o mRNA alvo contendo um nucleotídeo modificado (no caso o trifosfato de uridina/UTP marcado com Digoxigenina). Esta sonda será reconhecida por um anticorpo acoplado a uma enzima que é capaz de transformar um dado substrato em um produto cromogênico. Este procedimento, por fim, precipita uma substância colorida no local onde a sonda hibridizou (Tautz e Pfeifle, 1989).

Para realizar a hibridização *in situ*, os embriões foram coletados, decorionados e fixados. Os embriões foram então passados gradativamente de 100% metanol, em seguida 1:1 metanol/PBST, à solução 100% PBST. Em seguida os embriões foram incubados em solução de hibridização (em anexo) suplementada com DNA de esperma de salmão (Roche) por 1 hora a 55 °C, no procedimento de pré-hibridização, para prevenir a formação de ligações não específicas. Após a pré-hibridização, foram adicionados 2 µL das respectivas sondas, seguido de incubação por cerca de 16 h a 55 °C. Após esse tempo, os embriões foram submetidos a quatro lavagens com solução de hibridização seguidas de 3 lavagens de 30 minutos cada com a mesma solução. Então os embriões foram retornados gradativamente de 100% solução de hibridização para 100% PBST e tratados com solução de bloqueio (*Western Blocking Solution*, Roche). Neste momento, foi preparada a solução de bloqueio contendo anticorpo anti-digoxigenina acoplado à fosfatase alcalina (Anti-Digoxigenin-AP Cat. No. 1093274 Roche Diagnosis GmbH Fab

Fragments) na diluição de 1:2000. Os embriões foram incubados por aproximadamente 16 h à 4 °C com a solução contendo o anticorpo. Em seguida os embriões foram lavados com PBST por no mínimo 5 vezes, durante 5 minutos em cada lavagem. Os embriões foram então lavados com solução tampão AP (em anexo), para fosfatase alcalina, por 3 vezes durante 15 minutos em cada lavagem, e em seguida adicionou-se a solução de revelação (*BM purple AP substrate*, Roche). Sob um período de observação que durou de 30 minutos a 24 horas, observou-se o surgimento da coloração azul. A visualização dos embriões foi realizada com auxílio de uma lupa e placa de vidro de relojoeiro. Quando os embriões alcançaram a coloração adequada, foram lavados com PBST por três vezes para a interrupção da reação. Após as sucessivas lavagens com PBST, foi realizada a marcação com DAPI para identificação do estágio de desenvolvimento do embrião, por microscopia óptica de fluorescência. Os embriões foram então passados para glicerol e colocados em lâminas de microscopia contendo os devidos espaçadores para observação ao longo dos seus eixos anteroposterior e dorsoventral. A captura das imagens foi realizada na Lupa Leica DFC 550 de fluorescência utilizando-se de aumento de 161×.

#### **4.15. Experimentos com larvas de *T. castaneum***

O besouro *T. castaneum* possui de seis a sete estágios larvares, também conhecidos como *instars*. Para nossos experimentos, utilizamos dois estágios distintos: o estágio quatro e o estágio seis. Realizamos as injeções de dsRNA com auxílio do microinjetor *Nanoject II* (Drummond Scientific Company). A aplicação das soluções era feita da seguinte maneira: as larvas eram alocadas em placa de Petri de forma que sua lateral e seus intersegmentos se tornassem expostos. A injeção era feita entre os segmentos abdominais, de modo que o mínimo de lesões mecânicas fosse causado. As larvas foram injetadas com o volume de 0,21 µL ou 3 vezes o volume máximo de 0,69 nL por injeção do microinjetor. Em seguida as larvas, com aparente aumento no volume, eram colocadas em potes com farinha esteril e levadas à estufa tipo BOD a 30 °C. As larvas foram observadas a cada 2 dias e acompanhadas até seu desenvolvimento em adulto.

Com o objetivo de confirmar o silenciamento, realizamos coletas das larvas do grupo controle (*dsneo*) e *dszld* um dia após a injeção. Para isso foram seguidos

os protocolos de extração de RNA (ver item 4.6), síntese de cDNA (ver item 4.7) e PCR quantitativo em tempo real.

As reações de PCR em tempo real (qRT-PCR) foram realizadas em triplicatas biológicas em equipamento *StepOne Plus* (Applied Biosystems), utilizando o reagente *SYBR® Green PCR Master Mix* (Applied Biosystems). Uma reação controle na ausência do cDNA molde foi realizada para cada triplicata. Para investigar a expressão dos genes possivelmente regulados por *Tc-zld* durante a metamorfose, decidimos quantificar os genes *Tc-zld*, *Tc-Ubx*, *Tc-Wg*, *Tc-Dll*, *Tc-iro*, *Tc-lim1* e *Tc-krh1*. Os oligonucleotídeos iniciadores utilizados para cada gene estão listados na Tabela 3. Para quantificação comparativa da expressão, seguimos o método  $\Delta\Delta C_t$  com os valores sendo transformados para escala linear através da fórmula  $2^{-\Delta\Delta C_t}$ , de acordo com (Livak e Schmittgen, 2001).

Tabela 4: Oligonucleotídeos iniciadores utilizados nos ensaios de qRT-PCR.

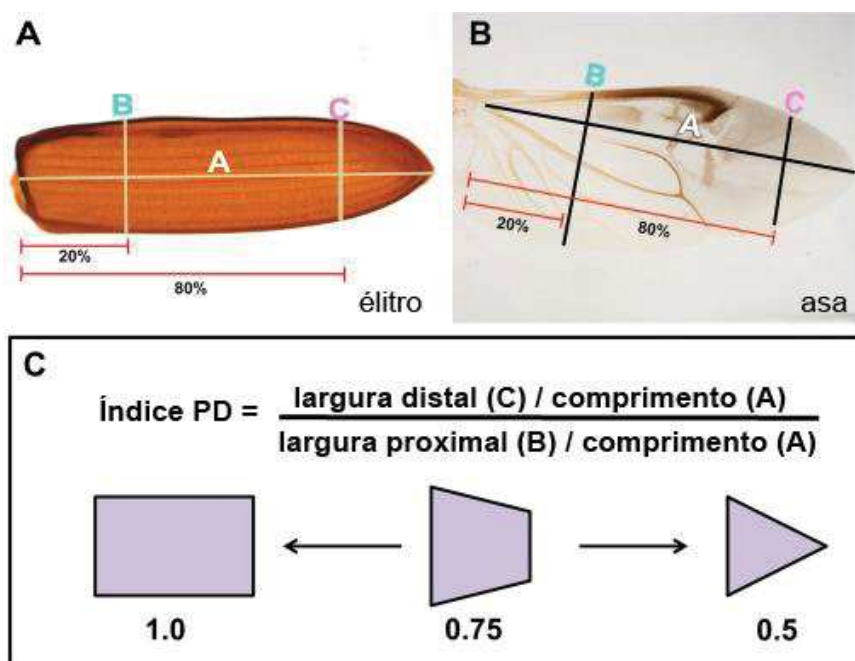
Oligonucleotídeo/Gene	Sequência (5' – 3')
<b>Tc-dll-QF1</b>	GCACCAGAACTCCTACACGGGATAC
<b>Tc-dll-QR1</b>	CGGCCTCTCCAACGATAAACACTTT
<b>Tc-iro-QF1</b>	AAGATGAAGAGGACAAACGGGACGA
<b>Tc-iro-QR1</b>	CCGAGGATGTGGTGCATGTTGTTG
<b>Tc-lim1-QF1</b>	ATTGTTTCACCTGTCTCGTGTGCAG
<b>Tc-lim1-QR1</b>	GCCGTCTTGTTTCCGTTGAGGTAAT
<b>Tc-ubx-QF1</b>	CCACACAAACCATTACCTGACACGG
<b>Tc-ubx-QR1</b>	TGCTCGTTGAGTTCTTTGATCGCTT
<b>TcKr-h1-QF1</b>	GTTTGCTCCAAGGGGTTTCACG
<b>TcKr-h1-QR1</b>	GGTTGTAGCCGAAGGATTTGCC
<b>Tc-zld-QF</b>	CAACGGCCCCTTACTGCTACT
<b>Tc-zld-QR</b>	GTGAGCACAGTCAAGCGGTACT
<b>Tc-rps3 QF</b>	CGCCTTGTAAGCGTCTCTATC
<b>Tc-rps3 QR</b>	TGGATGGAAAGGCCAATTC

#### 4.16. Análises morfológicas

Os indivíduos foram observados periodicamente e seus fenótipos iniciais anotados. Apêndices como antenas, asas e élitros e patas foram dissecados e fixados em etanol 95% para posterior análise. As estruturas foram levadas à lupa Leica, modelo M205, onde foram capturadas imagens em alta definição dos fenótipos observados. As imagens foram organizadas utilizando-se o software Adobe Photoshop CS4 (versão 11.0 /Adobe).

#### 4.17. Cálculo do Índice próximo-distal

As medidas de asa e élitros para cálculo do índice PD foram realizadas de acordo com a metodologia descrita por (Ravisankar *et al.*, 2016). O índice é obtido medindo-se o comprimento e a largura de duas posições das asas (Figura 4 A,B). A largura proximal (medida B) encontra-se a 20% do comprimento da asa (medida A) a partir da região proximal, e a largura distal (medida C) é medida a 80% deste comprimento. Os valores das larguras distal e proximal são normalizados contra o valor do comprimento e a razão entre as os valores normalizados gera o índice PD (Figura 4 C). Esta análise morfométrica exprime a redução distal na morfologia de asas e élitros em valores que representam da forma retangular (próxima de 1) a triangular (próxima de 0,5) dos apêndices (Figura 4)



**Figura 4: Cálculo do índice PD para asas e élitros.** (A) e (B) as medidas de largura “B” e “C”, e comprimento “A” de élitro e asa utilizados para o cálculo do índice PD. (C) A fórmula utilizada para o cálculo do índice PD e a demonstração entre as relações dos valores do índice e a forma das estruturas. Adaptado de (Ravisankar *et al.*, 2016).



#### 4.18. Estatística

A análise dos dados foi realizada com auxílio do programa estatístico GraphPad Prism versão 5.0. Para os dados de PCR em tempo real em larvas foi utilizado two-way ANOVA seguido de pós-teste de Bonferroni. Os dados de índice PD foram analisados com o teste de normalidade Shapiro-Wilk seguido do teste t não pareado.

### 5. RESULTADOS E DISCUSSÃO

Um dos conjuntos de resultados desta tese envolveu a identificação e análise de homólogos do gene *zelda* entre os Pancrustacea, algo não esperado, tendo em vista que uma hipótese inicial considerava *zelda* como uma novidade da classe Insecta. Essas observações foram realizadas em conjunto com o pesquisador Dr. Aravind Lakshminarayanan Iyer do *National Center for Biotechnology Information* (USA) e contribuíram para entender a história deste importante fator de transcrição.

#### 5.1. Identificação de domínios conservados e a origem de *zld*

Enquanto estudos anteriores reportaram o envolvimento de *zld* na MZT em *D. melanogaster* (Liang *et al.*, 2008), sua história evolutiva permanecia em questão. Nós investigamos a distribuição filética de *zld* e encontramos um único ortólogo em todos os genomas de insetos avaliados, incluindo o besouro *T. castaneum* (Figura 5), indicando que insetos são provavelmente sensíveis ao aumento de cópias desse regulador, o que é interessante considerando que diferentes famílias de TFs são propensas a expansão em linhagens de insetos (Vidal *et al.*, 2016). Além disso, homólogos de *zld* também são encontrados em genomas de collembola e crustáceos (Figura 5). A arquitetura canônica de Zld é descrita tendo a presença de um domínio em dedo de zinco JAZ (Pfam:zf-C2H2\_JAZ) e um agrupamento C-terminal com quatro domínios em dedos de zinco C2H2 (zf-C2H2) (Staudt *et al.*, 2006; Liang *et al.*, 2008; Tribolium Genome Sequencing *et al.*, 2008). Realizamos uma análise detalhada das proteínas Zld de múltiplas espécies e encontramos outros domínios conservados e características estruturais dignas de observação. Primeiramente, encontramos dois domínios zf-C2H2 N-terminais ao domínio JAZ (Figura 5 e Figura 6). O domínio zf-C2H2 mais N-terminal encontra-se parcialmente completo (com ausência de cisteínas e histidinas conservadas) em algumas espécies, incluindo *D. melanogaster* (Figura 5 e Figura 6). Por outro lado, o segundo

domínio N-terminal zf-C2H2 é conservado em virtualmente todos os insetos existentes, mas ausentes ou degenerados em outros Pancrustacea (por exemplo, em *Daphnia magna*, ordem Cladocera, encontra-se parcialmente conservado, exceto pela perda de uma cisteína) (Figura 5). Além disso, entre este segundo domínio zf-C2H2 e o domínio JAZ existe uma região de aminoácidos de natureza acídica que podem estar envolvidos no recrutamento de proteínas remodeladoras de cromatina baseados na conservação com regiões acídicas da proteína remodeladora de cromatina CECR2 (Aravind e Iyer, 2012) (Figura 5 e Figura 6). Esse trecho é caracterizado por um motivo conservado, de consenso [DE]I[LW]DLD, que é susceptível a adotar uma conformação helicoidal entre regiões desordenadas. Estes trechos acídicos também são encontrados nos domínios HUN de chaperonas de histonas (Balaji *et al.*, 2009). Uma grande região desordenada, localizada entre o domínio JAZ e o agrupamento conservado de quatro domínios zf-C2H2, foi demonstrada de importância na transativação de *Dm-zld in vitro* (Hamm *et al.*, 2015). Nossa análise das restrições evolutivas na sequência proteica revelaram um motivo de forma hP[IVM]SxHHHPxRD que aparenta estar sob seleção por retenção apesar da grande divergência nesta região. Conseqüentemente, é possível que esta região desempenhe um papel específico na função de transativação. Além disso, entre os dois domínios N-terminais zf-C2H2, existe o domínio RYPHY, que é bastante conservado em proteínas *Zld* e pode estar envolvido na determinação da localização nuclear (Figura 5), como predito para alguns repressores transcricionais (Pereira *et al.*, 2016). Considerada a conservação destes domínios adicionais, surge a hipótese que estes também desempenhem importantes papéis nas funções de *Zld*, previamente atribuídas exclusivamente aos domínios C-terminais.

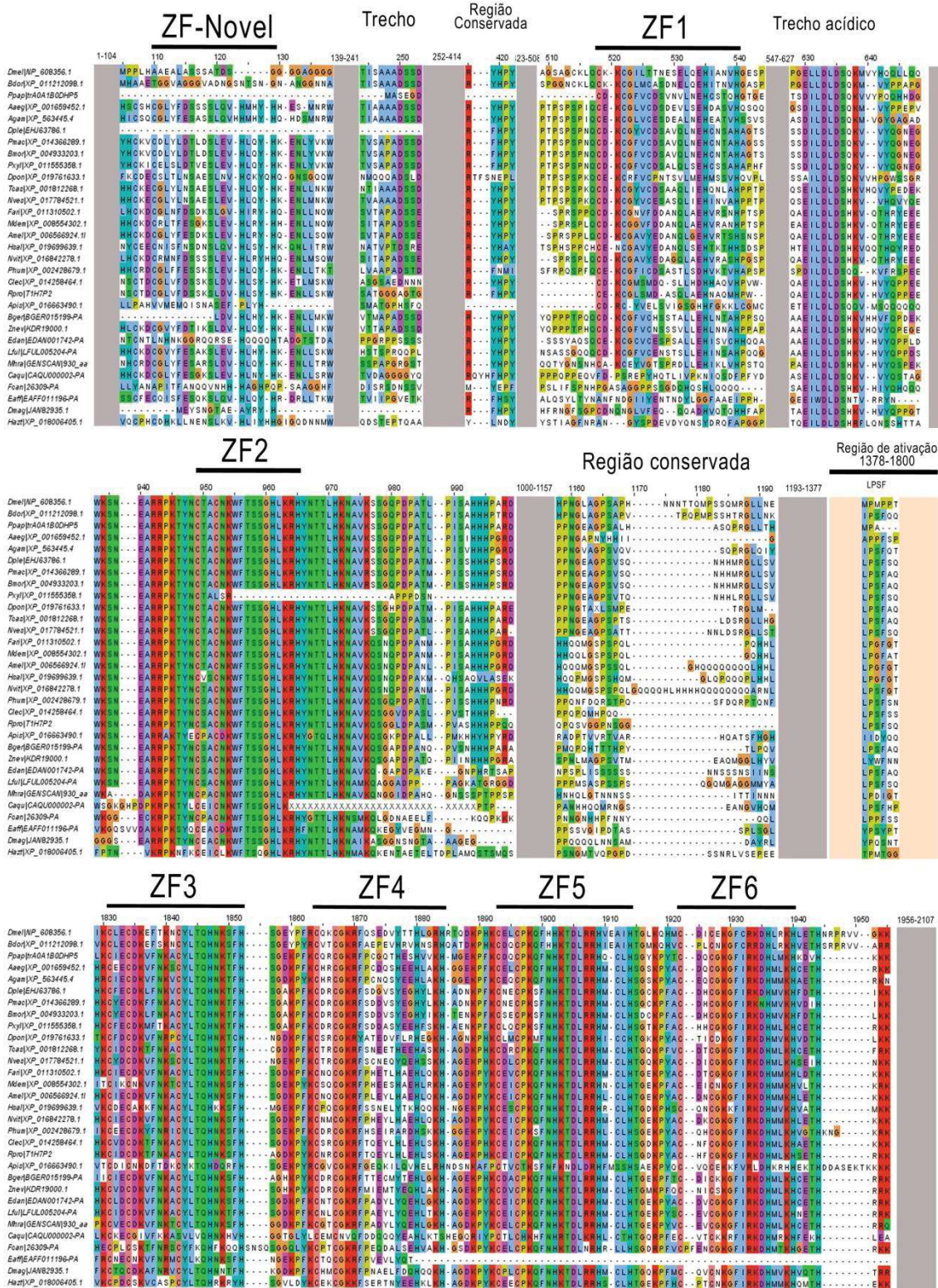
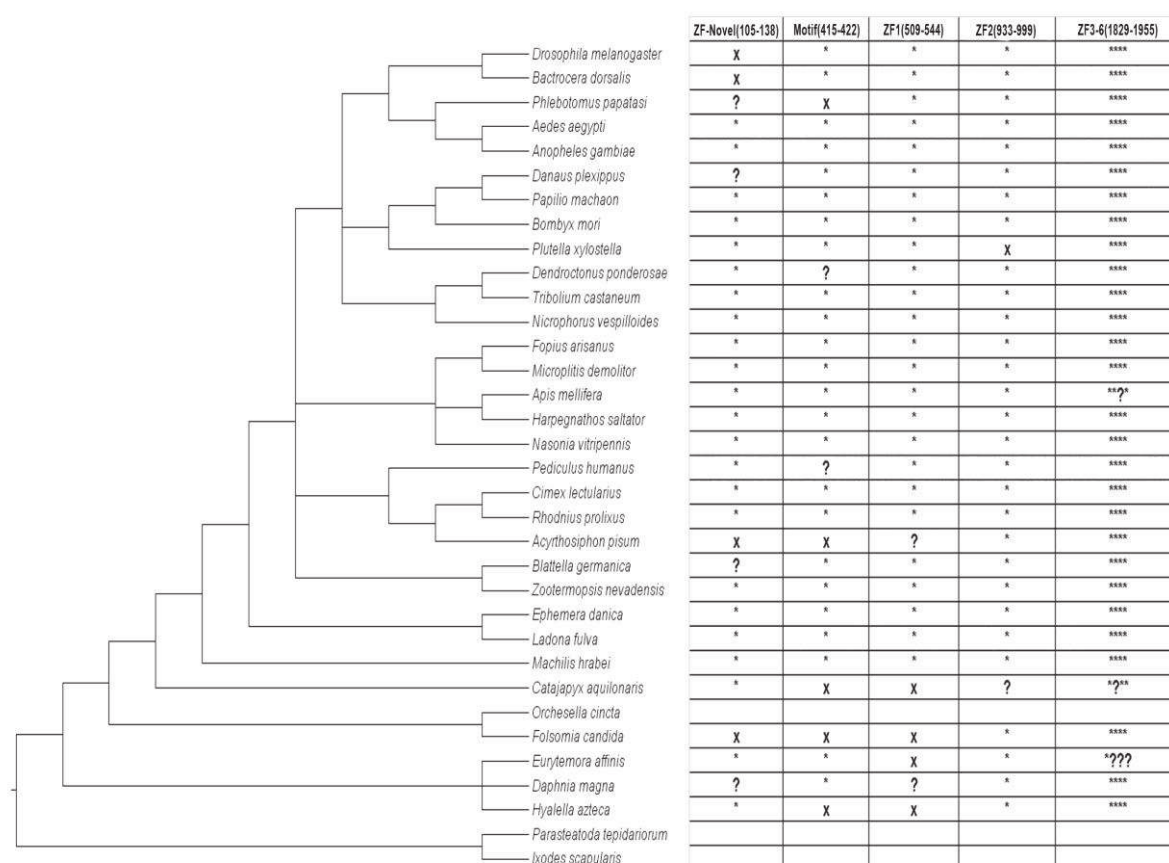


Figura 5: Alinhamento múltiplo de seqüências de proteínas Zeldin representando os principais grupos dos artrópodes. Alinhamento realizado com o programa muscle (Edgar, 2004) e visualização feita a partir do programa JalView (Waterhouse et al., 2009).



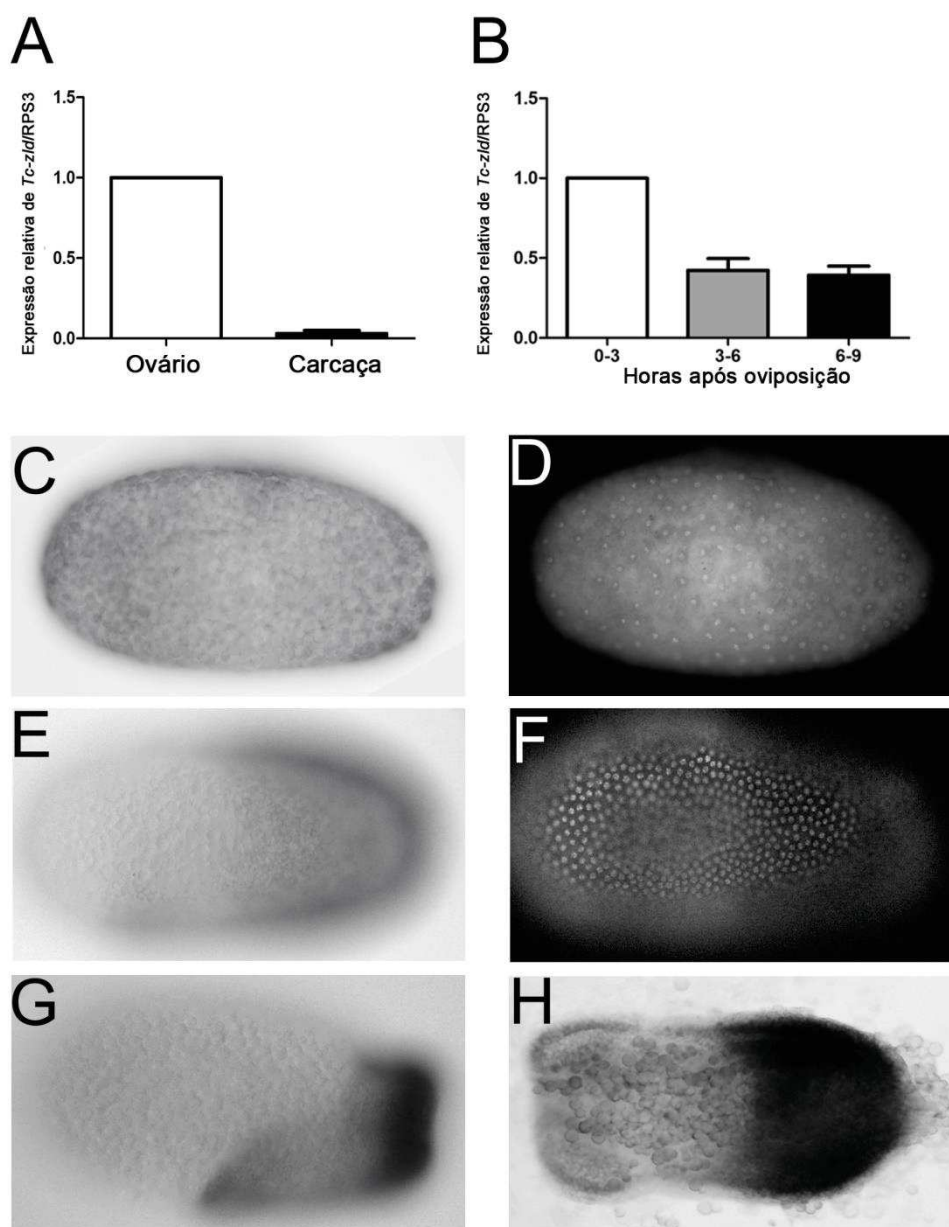
**Figura 6: Características conservadas das proteínas *Zelda*.** Asteriscos e “x” indicam presença e ausência, respectivamente. O ponto de interrogação denota que a característica da sequência está parcialmente preservada ou ausente devido a erros na montagem do genoma. O cladograma foi organizado de acordo com o estudo filogenético realizado por (Misof *et al.*, 2014). Dois grupos externos com indivíduos ausentes de ortólogos de *zld* foram incluídos.

Buscando esclarecer as origens de *zld*, nós realizamos uma extensa busca por sequências homólogas e fomos incapazes de encontrar genes ortólogos com arquitetura de domínios conservados fora de Pancrustacea, indicando que *zld* é uma inovação deste clado. Nós detectamos homólogos de *zld* entre os insetos, incluindo o cupim *Zootermopsis nevadensis* (Ordem Isoptera), na libélula *Ladona fulva* (Ordem Odonata), em efêmera *Ephemera danica* (Ordem Ephemeroptera) e em *Machilis hrabei* (Ordem Archaeognatha). Homólogos de *Zld* também são encontrados em crustáceos pertencentes a diferentes ordens: *Daphnia magna* (Diplostraca), *Hyalella azteca* (Amphipoda) e *Eurytemora affinis* (Calanoida), assim como no colembola *Folsomia cândida*. Curiosamente não foram encontrados homólogos de *zld* nos genomas de *Orchesella cincta* (collembola) e *Daphnia pulex* (crustáceo), sugerindo que *zld* não está absolutamente conservado fora dos insetos ou a sequência se perdeu devido a genomas depositados de forma incompleta nos

bancos de dados. Adicionalmente realizamos buscas minuciosas em genomas de outros artrópodes não-insetos, incluindo quelicerados (no carrapato *Ixodes scapularis* e na aranha *Parasteatoda tepidariorum*) e não foi possível encontrar proteínas com a organização de domínios semelhante a Zld (Figura 6). Buscas gerais no banco de dados do NCBI contra proteínas de artrópodes não-Pancrustacea também não retornaram ortólogos de Zld. Embora buscas realizadas com o algoritmo BLAST utilizando sequências peptídicas de Zld contra os bancos de dados *nr* (sequências não redundantes) e *refseq* (sequências referência) tenham retornado um número significativo de proteínas em eucariotos distantes, a similaridade estava quase sempre restrita ao agrupamento C-terminal de quatro domínios Zf-C2H2, que são bastante comuns entre diversas famílias de TFs (por exemplo, *glass*, *earmuff/fez*, *senseless/gfi-1* e *jim*). Quando considerados em conjunto, nossos resultados suportam a emergência recente de *zld* no clado Pancrustacea, com possíveis perdas subsequentes em espécies em particular. Cabe ressaltar que todos os genomas de insetos que analisamos possuem exatamente uma cópia do gene *zld* por genoma haploide.

## **5.2. *Zelda* é expresso na região da zona de crescimento em *T. castaneum***

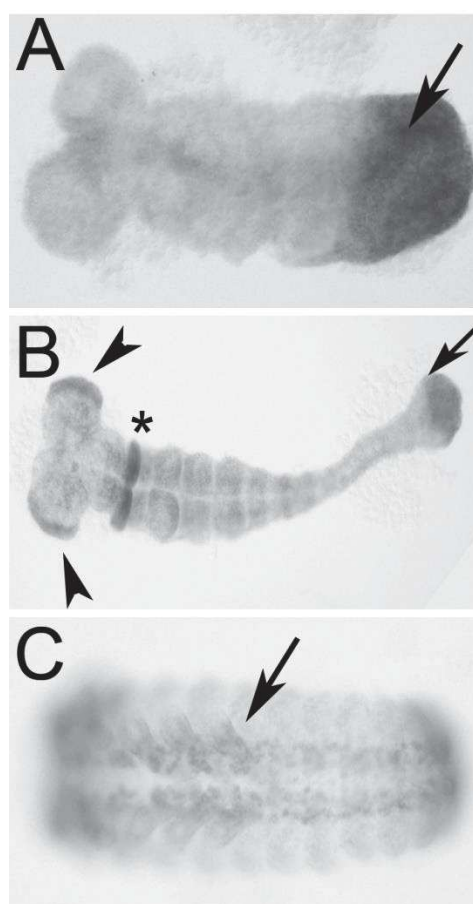
A partir do conhecimento de que *zld* é maternalmente expresso nas células germinativas de *D. melanogaster* (Liang *et al.*, 2008), comparamos os níveis de mRNA de ovários extraídos de fêmeas de *T. castaneum* e suas carcaças por qRT-PCR. Transcritos de *Tc-zld* foram altamente expressos em ovários (Figura 7 A), sugerindo que *Tc-zld* é expresso células germinativas. O desenvolvimento de *T. castaneum* se inicia com divisões mitóticas sincronizadas e com ausência de celularização durante as primeiras três horas de embriogênese (à 30 °C), seguido da migração dos núcleos em direção ao córtex do ovo e da segregação desses em células em 7 à 8 horas após a oviposição (Benton *et al.*, 2013). Os níveis do mRNA de *Tc-zld* são maiores nas primeiras três horas de desenvolvimento (Figura 7 B) quando comparados aos períodos posteriores (3-6 e 6-9 horas), sugerindo que o mRNA de *Tc-zld* é degradado após as primeiras 3 horas de desenvolvimento.



**Figura 7: A expressão de *Tc-zld* é de origem materna e posteriormente confinada na zona de crescimento durante o desenvolvimento embrionário.** (A) A expressão relativa de *Tc-zld* no ovário e carçaça. (B) Expressão relativa de *Tc-zld* 0-3, 3-6 e 6-9 horas após a oviposição. A expressão foi normalizada em relação ao gene constitutivo *rps3* em ambos os experimentos. Hibridização *in situ* para o mRNA de *Tc-zld* ao longo do desenvolvimento de *T. castaneum*. (C) Transcritos de *Tc-zld* são uniformemente distribuídos no embrião no estágio de pré-blastoderma (0-3 horas);(D) DAPI do respectivo estágio. (E) No estágio de blastoderma celular (3-6 horas), transcritos começam a ocupar o primórdio de banda germinal;(F) marcação nuclear por DAPI. (G) Pouco tempo após a invaginação, a expressão de *zld* começa a se concentrar na região posterior de onde a zona de crescimento irá se formar (6-9 horas). (H) pouco após o início do alongamento da banda germinal, a expressão de *Tc-zld* é confinada à zona de crescimento.

A hibridização *in situ* utilizando sondas para o gene *Tc-zld* confirmou a expressão materna nas primeiras 3 horas de desenvolvimento (Figura 7 CD) e

demonstrou um confinamento progressivo do mRNA de *Tc-zld* na região posterior do ovo, onde o embrião começa a ser formado (Figura 7 EF). Entre 6 e 9 horas de desenvolvimento, a expressão de *Tc-zld* permanece concentrada na região posterior (Figura 7 GH), de onde surgirão os segmentos abdominais a partir da zona de crescimento (El-Sherif *et al.*, 2012; Sarrazin *et al.*, 2012). Mais tarde no desenvolvimento, a expressão de *Tc-zld* ainda é detectada na zona de crescimento (Figura 8 AB) e no sistema nervoso (Figura 8 C), evidenciando a conservação de *zelda* nas células do tecido nervoso, como relatado para *D. melanogaster* (Staudt *et al.*, 2006; Liang *et al.*, 2008; Pearson *et al.*, 2012).

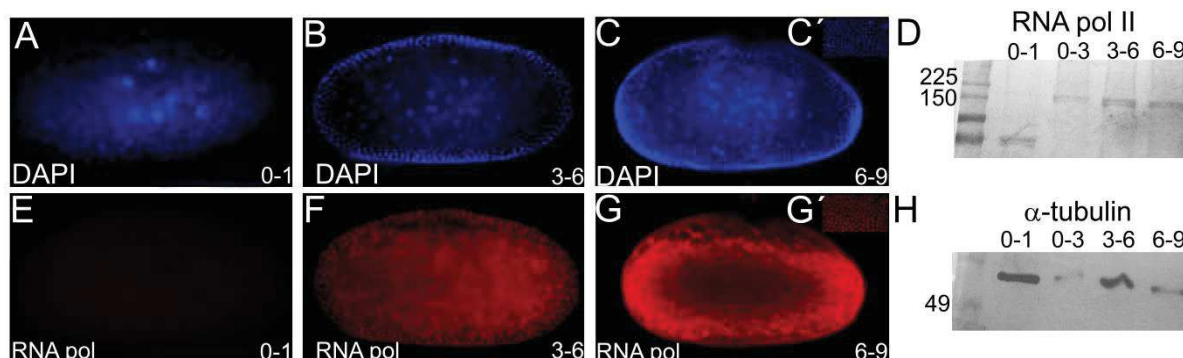


**Figura 8: A expressão de *Tc-zelda* durante os estágios mais tardios do desenvolvimento.** (A) No início da extensão da banda germinal, ou seja, aproximadamente 13 horas após a oviposição, *Tc-zld* é expresso no região posterior de crescimento (Seta). (B) No período de 18-21 horas após oviposição, *Tc-zld* mantém sua expressão na região mais posterior do embrião (seta), além de ser expresso nos lobos cefálicos (pontas de setas) e segmento da gnata (asterisco). (C) Aproximadamente 39-48 horas após a oviposição, durante o fechamento dorsal, a expressão pode ser observada ao longo sistema nervoso (seta).

### 5.3. A ativação zigótica em *T. castaneum*

O gene *zld* é descrito como principal ativador zigótico em *D. melanogaster*, atuando principalmente durante a MZT do inseto ao ativar a primeira onda de expressão de genes zigóticos e favorecer a expressão dos genes nos estágios subsequentes ao oitavo ciclo mitótico (Liang *et al.*, 2008; Harrison *et al.*, 2011). O seu ortólogo de *T. castaneum*, *Tc-zld*, também desempenha um papel importante no início da embriogênese e concentra sua expressão na zona posterior de crescimento em momentos anteriores à extensão da banda germinal (Figura 7 CH). Com relação a MZT, não existem dados disponíveis na literatura sobre o momento da ativação zigótica no besouro de banda curta.

Utilizando um anticorpo contra a forma ativa da enzima RNA polimerase II (RNA pol II), em método previamente descrito por (Nestorov *et al.*, 2013), foi possível demonstrar que a transcrição dos genes zigóticos em *T. castaneum* se inicia entre três e seis horas de desenvolvimento (Figura 9 BF). Nesse período, os núcleos (energídeos) formados pelas primeiras divisões mitóticas migram em direção à periferia do ovo, onde serão celularizados, dando origem ao blastoderma celular. Com auxílio das alunas Vitória Tobias Santos, Geórgia Feltran e Daniele Santos, foi realizado um westernblot (Figura 9 DH) com extratos de embriões nos mesmos estágios utilizados na imunohistoquímica, confirmando a presença da forma fosforilada da RNA pol II observada nas imagens de microscopia de fluorescência.



**Figura 9. A transição materno zigótica (MZT) em *Tribolium castaneum*.** (A-C) Marcação nuclear com DAPI em embriões de *Tribolium* entre 0-1 horas (A), 3-6 horas (B) e 6-9 horas após a oviposição (C). (D,H) Western blots de extratos de embriões de 0-1, 0-3, 3-6, 6-9 horas após oviposição utilizando anticorpos para a forma transcricionalmente ativa de RNA pol II como descrito por (Nestorov *et al.*, 2013). (D) ou anticorpo contra a proteína  $\alpha$ -tubulina utilizado como controle (H).

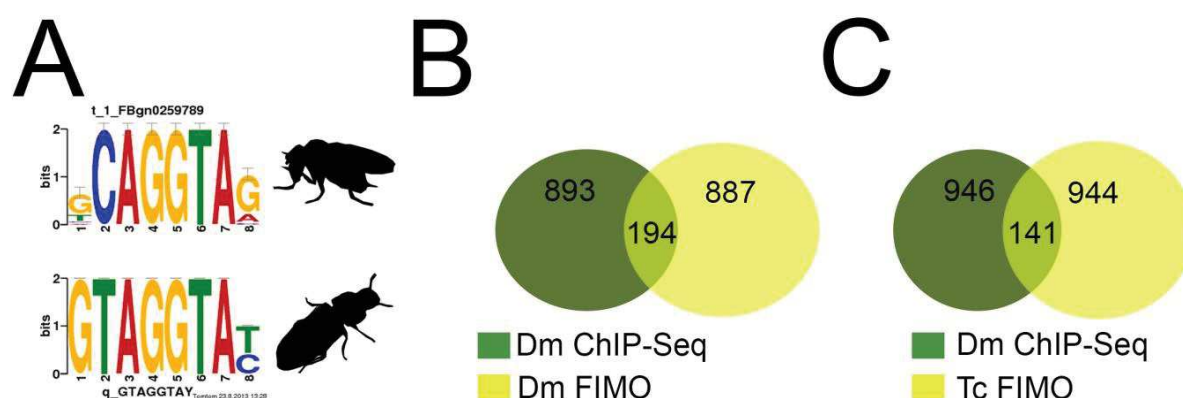
Conforme observado através da imunohistoquímica em *T. castaneum*, a ativação do genoma zigótico começa momentos antes e se estabelece momentos após a celularização do embrião, condizente com o observado através da expressão



de miRNAs no próprio besouro (Ninova *et al.*, 2016). Os resultados indicam que este acontecimento da ativação zigótica aparenta ser conservado entre os insetos descritos até o momento, em *N. vitripennis*, um ortólogo de *zelda* também é depositado de forma maternal e alcança níveis altos de expressão após a celularização do embrião (Arsala e Lynch, 2017). Em *A. melífera* a MZT ocorre por volta das 24 horas de desenvolvimento, também durante a celularização, mas em tempo diferente da *Drosophila* e *Tribolium* devido ao seu processo de desenvolvimento mais lento (Pires *et al.*, 2016). Nos organismos recém-citados o TF *zelda* parece ser o principal componente da ativação zigótica. Devido a sua conservação, *zelda* despertou o interesse para uma abordagem que envolve a predição da interação da proteína com regiões do genoma conhecidas como sítios de interação de TFs (TF *binding sites*, TFBS). Baseado na hipótese de que famílias de TFs de estruturas diferentes possuem especificidades divergentes e a similaridade de sequências de aminoácidos pode auxiliar na predição de interação entre os TFs e as sequências de DNA (Nitta *et al.*, 2015), foi realizada a busca por regiões de ligação de *Tc-zld* no genoma de *T. castaneum* e identificação de genes possivelmente regulados.

#### **5.4. A predição da interação de *zelda* no genoma de *T. castaneum***

Primeiramente utilizamos o programa DREME (Bailey, 2011) para analisar as regiões 2000 pb à 5' (ou "a montante") de todos os genes codificantes de proteínas da versão 3.0 do genoma de *T. castaneum*. Esta análise encontrou uma sequência motivo (GTAGGTAY, onde "Y" representa uma pirimidina) bastante similar ao motivo TAGteam previamente descrito (Ten Bosch *et al.*, 2006) (Figura 10 A). Em seguida, nós utilizamos o genoma de *D. melanogaster* e dados experimentais (Liang *et al.*, 2008; Harrison *et al.*, 2011; Nien *et al.*, 2011) para validar a metodologia e encontrar uma sobreposição significativa entre os alvos preditos e experimentais de Zld em *D. melanogaster* (Figura 10 B). A sequência do motivo encontrado em *T. castaneum* foi então utilizada para a busca no genoma, resultando na identificação de 3250 alvos putativos de Tc-Zld, representando aproximadamente 19% do genoma de *T. castaneum* (Figura 10 C).



**Figura 10: Predição computacional e identificação de genes alvos de Tc-Zld:** (A) Comparação feita pelo software TOMTOM do motivo obtido com o programa DREME no genoma de *Tribolium castaneum* e o motivo de ligação de *Zelda* em *D. melanogaster*. (B) Diagrama de Venn dos genes regulados durante a MZT descritos por (Harrison *et al.*, 2011) e os genes preditos pelo programa FIMO utilizando o motivo predito de *D. melanogaster* (amarelo). (C) Diagrama de Venn com os alvos de Zld em *D. melanogaster* (Harrison *et al.*, 2011) em verde e os ortólogos do tipo 'one-to-one' de genes alvos preditos para Tc-Zld.

A comparação entre os alvos putativos de Tc-Zld com os 1087 genes regulados por Zld durante a MZT de *D. melanogaster* (Harrison *et al.*, 2011; Nien *et al.*, 2011) permitiu a identificação de 141 genes de *D. melanogaster* com homologia *one-to-one* (ou seja, um único identificador correspondente da proteína no organismo alvo) com os alvos putativos de Tc-Zld em *T. castaneum* (Figura 10 C, Tabela 5). Análises funcionais deste grupo de genes, utilizando a ferramenta DAVID (Huang *et al.*, 2009), revelaram o enriquecimento funcional de importantes categorias, incluindo um grupo de 26 genes do tipo *homeobox* (Tabela 6) além de outros grupos de genes envolvidos em especificação de segmentos e regionalização, formação de disco imaginal e metamorfose (Tabela 6). Este grupo de genes inclui diversos reguladores do desenvolvimento, como genes de padronização anteroposterior (AP), dorsoventral (DV), do tipo gap, *regra* dos pares e genes homeóticos.

Tabela 5: Lista de 141 alvos de Zld durante a MZT de *D. melanogaster* que possuem ortólogos *one-to-one* entre os alvos preditos de Tc-Zld.

Identificador do gene	Identificador do gene de <i>T. castaneum</i>	Nome do gene de <i>T. castaneum</i>
FBgn0000014	TC000894	ABD-A
FBgn0000015	TC000889	ABD-B
FBgn0041171	TC006451	AGO
FBgn0262739	TC005857	AGO1
FBgn0013751	TC003238	AWH
FBgn0024250	TC000748	BRK

FBgn0000250	TC002003	CACT
FBgn0000251	TC007576	CAUDAL-1
FBgn0000320	TC008985	CLIFT
FBgn0000439	TC000920	DFD
FBgn0000157	TC009351	DLL
FBgn0000497	TC007180	DS
FBgn0000504	TC010597	DSX
FBgn0000659	TC013245	FKH
FBgn0001150	TC007492	GIANT
FBgn0033252	TC007382	GLEAN_07382
FBgn0005677	TC007637	GLEAN_07637
FBgn0261954	TC007655	GLEAN_07655
FBgn0035432	TC007872	GLEAN_07872
FBgn0003028	TC008099	GLEAN_08099
FBgn0033607	TC008107	GLEAN_08107
FBgn0037797	TC008170	GLEAN_08170
FBgn0038057	TC008189	GLEAN_08189
FBgn0020762	TC008293	GLEAN_08293
FBgn0035767	TC010995	GLEAN_10995
FBgn0053977	TC014898	GLEAN_14898
FBgn0004878	TC015058	GLEAN_15058
FBgn0001987	TC010824	GLI
FBgn0001168	TC012851	HAIRY
FBgn0008636	TC008926	HBN
FBgn0261434	TC010992	HKB
FBgn0001319	TC001270	KN
FBgn0001325	TC011460	KR
FBgn0002522	TC000926	LAB
FBgn0026411	TC014939	LIM1
FBgn0014343	TC003634	MIRR
FBgn0005771	TC000693	NOC
FBgn0015524	TC008928	OTP
FBgn0051481	TC000925	PB
FBgn0003129	TC008838	POXM
FBgn0003130	TC008672	POXN
FBgn0003339	TC000917	SCR
FBgn0002573	TC013474	SENS
FBgn0003396	TC009542	SHN
FBgn0027364	TC003852	SIX4
FBgn0042630	TC013165	SOX21B
FBgn0029123	TC014065	SOXN
FBgn0000606	TC009469	TC-EVE
FBgn0029114	TC004898	TOLLO
FBgn0003944	TC000903	UBX
FBgn0004009	TC014084	WNT1
FBgn0000180	TC010832	
FBgn0000227	TC005343	
FBgn0000492	TC011744	
FBgn0001942	TC014113	
FBgn0002778	TC013186	
FBgn0002931	TC005579	
FBgn0002941	TC012332	
FBgn0003002	TC010234	
FBgn0003118	TC014512	

FBgn0003499	TC004846
FBgn0004583	TC012545
FBgn0010246	TC011553
FBgn0010313	TC014556
FBgn0010660	TC014339
FBgn0011278	TC011748
FBgn0013303	TC010475
FBgn0013750	TC004744
FBgn0014002	TC003523
FBgn0014092	TC011681
FBgn0014143	TC002813
FBgn0019890	TC001202
FBgn0020307	TC001741
FBgn0023528	TC030776
FBgn0024244	TC006347
FBgn0024291	TC002584
FBgn0025455	TC009709
FBgn0026160	TC000880
FBgn0026239	TC001065
FBgn0026533	TC008666
FBgn0026718	TC011057
FBgn0027660	TC003241
FBgn0029092	TC010167
FBgn0029095	TC001466
FBgn0030096	TC008817
FBgn0031645	TC006413
FBgn0031770	TC010818
FBgn0031850	TC008608
FBgn0031950	TC010160
FBgn0031961	TC003018
FBgn0032723	TC009337
FBgn0032922	TC004041
FBgn0033351	TC011565
FBgn0033507	TC003187
FBgn0034046	TC009488
FBgn0034282	TC030635
FBgn0034644	TC001676
FBgn0034795	TC003334
FBgn0034876	TC006736
FBgn0034972	TC013623
FBgn0035164	TC000276
FBgn0035204	TC010872
FBgn0035643	TC013073
FBgn0036032	TC002269
FBgn0037203	TC001934
FBgn0037235	TC012171
FBgn0037427	TC012819
FBgn0037648	TC015973
FBgn0037788	TC008569
FBgn0037892	TC015959
FBgn0038126	TC000982
FBgn0038531	TC014646
FBgn0039056	TC012753
FBgn0039683	TC010058

<b>FBgn0040071</b>	TC008695
<b>FBgn0041156</b>	TC009461
<b>FBgn0043364</b>	TC014124
<b>FBgn0043900</b>	TC004669
<b>FBgn0051248</b>	TC014389
<b>FBgn0051522</b>	TC010977
<b>FBgn0052369</b>	TC000081
<b>FBgn0052423</b>	TC005341
<b>FBgn0052451</b>	TC011526
<b>FBgn0053169</b>	TC014457
<b>FBgn0053196</b>	TC011986
<b>FBgn0053207</b>	TC009138
<b>FBgn0083978</b>	TC014638
<b>FBgn0086712</b>	TC001370
<b>FBgn0086898</b>	TC006777
<b>FBgn0259986</b>	TC004602
<b>FBgn0261552</b>	TC008998
<b>FBgn0261930</b>	TC011745
<b>FBgn0262699</b>	TC001036
<b>FBgn0262733</b>	TC002522
<b>FBgn0262737</b>	TC000237
<b>FBgn0263397</b>	TC010149
<b>FBgn0265274</b>	TC011709
<b>FBgn0266084</b>	TC003980
<b>FBgn0266756</b>	TC001061
<b>FBgn0267002</b>	TC006253
<b>FBgn0267791</b>	TC009519

---

Tabela 6: Análise funcional com termos de ontologia dos genes listados na Tabela 4 realizada pela ferramenta DAVID.

Annotation Cluster 1		Enrichment Score: 5.02					
Category	Term	Count	%	List Total	Fold Enrichment	PValue	
SP_PIR_KEYWORDS	dna-binding	46	33,58	133	2,12	1,38E-08	
INTERPRO	IPR001356:Homeobox	26	18,98	127	2,97	3,09E-08	
SP_PIR_KEYWORDS	Homeobox	26	18,98	133	2,96	3,39E-08	
SMART	SM00389:HOX	26	18,98	80	2,77	4,81E-08	
INTERPRO	IPR012287:Homeodomain-related	25	18,25	127	2,97	6,05E-08	
INTERPRO	IPR017970:Homeobox, conserved site	25	18,25	127	2,91	1,03E-07	
SP_PIR_KEYWORDS	nucleus	51	37,23	133	1,85	3,35E-07	
GOTERM_BP_FAT	GO:0006355~regulation of transcription, DNA-dependent	48	35,04	110	1,83	5,11E-07	
GOTERM_BP_FAT	GO:0051252~regulation of RNA metabolic process	48	35,04	110	1,78	1,43E-06	
GOTERM_MF_FAT	GO:0003700~transcription factor activity	40	29,20	110	1,92	2,26E-06	
GOTERM_BP_FAT	GO:0045449~regulation of transcription	52	37,96	110	1,66	4,26E-06	
UP_SEQ_FEATURE	DNA-binding region:Homeobox	18	13,14	63	2,74	9,74E-06	
GOTERM_MF_FAT	GO:0043565~sequence-specific DNA binding	30	21,90	110	2,07	1,76E-05	
GOTERM_MF_FAT	GO:0030528~transcription regulator activity	49	35,77	110	1,61	2,96E-05	
GOTERM_MF_FAT	GO:0003677~DNA binding	48	35,04	110	1,62	3,35E-05	
GOTERM_MF_FAT	GO:0003704~specific RNA polymerase II transcription factor activity	17	12,41	110	2,17	1,47E-03	
SP_PIR_KEYWORDS	developmental protein	39	28,47	133	1,55	1,64E-03	
GOTERM_MF_FAT	GO:0003702~RNA polymerase II transcription factor activity	25	18,25	110	1,79	1,64E-03	
SP_PIR_KEYWORDS	transcription regulation	26	18,98	133	1,53	1,70E-02	

SP_PIR_KEYWORDS	Transcription	23	16,79	133	1,44	5,00E-02
GOTERM_BP_FAT	GO:0006350~transcription	25	18,25	110	1,36	6,72E-02

---

Annotation Cluster 2 **Enrichment Score: 2.41**

---

Category	Term	Count	%	List Total	Fold Enrichment	PValue
GOTERM_BP_FAT	GO:0003002~regionalization	34	24,82	110	1,70	4,32E-04
GOTERM_BP_FAT	GO:0007379~segment specification	10	7,30	110	3,14	1,24E-03
GOTERM_BP_FAT	GO:0007389~pattern specification process	34	24,82	110	1,61	1,37E-03
GOTERM_MF_FAT	GO:0003704~specific RNA polymerase II transcription factor activity	17	12,41	110	2,17	1,47E-03
SP_PIR_KEYWORDS	developmental protein	39	28,47	133	1,55	1,64E-03
GOTERM_MF_FAT	GO:0003702~RNA polymerase II transcription factor activity	25	18,25	110	1,79	1,64E-03
GOTERM_BP_FAT	GO:0035287~head segmentation	8	5,84	110	3,05	6,88E-03
GOTERM_BP_FAT	GO:0007350~blastoderm segmentation	18	13,14	110	1,85	7,51E-03
GOTERM_BP_FAT	GO:0009880~embryonic pattern specification	18	13,14	110	1,82	9,32E-03
GOTERM_BP_FAT	GO:0035289~posterior head segmentation	7	5,11	110	3,12	1,25E-02
GOTERM_BP_FAT	GO:0007380~specification of segmental identity, head	5	3,65	110	3,82	2,49E-02
GOTERM_BP_FAT	GO:0035282~segmentation	19	13,87	110	1,61	2,67E-02

---

Annotation Cluster 3 **Enrichment Score: 1.54**

---

Category	Term	Count	%	List Total	Fold Enrichment	PValue
GOTERM_BP_FAT	GO:0007447~imaginal disc pattern formation	16	11,68	110	2,44	5,13E-04
GOTERM_BP_FAT	GO:0007450~dorsal/ventral pattern formation, imaginal disc	8	5,84	110	2,38	3,31E-02
GOTERM_BP_FAT	GO:0035222~wing disc pattern formation	8	5,84	110	2,14	5,83E-02

GOTERM_BP_FAT	GO:0009953~dorsal/ventral pattern formation	10	7,30	110	1,72	9,94E-02
GOTERM_BP_FAT	GO:0048190~wing disc dorsal/ventral pattern formation	5	3,65	110	2,06	2,07E-01

## Annotation Cluster 4

Enrichment Score: 1.41

Category	Term	Count	%	List Total	Fold Enrichment	PValue
GOTERM_BP_FAT	GO:0007480~imaginal disc-derived leg morphogenesis	7	5,11	110	3,40	7,38E-03
GOTERM_BP_FAT	GO:0035109~imaginal disc-derived limb morphogenesis	7	5,11	110	3,40	7,38E-03
GOTERM_BP_FAT	GO:0035127~post-embryonic limb morphogenesis	7	5,11	110	3,40	7,38E-03
GOTERM_BP_FAT	GO:0007444~imaginal disc development	29	21,17	110	1,53	8,40E-03
GOTERM_BP_FAT	GO:0007478~leg disc morphogenesis	7	5,11	110	3,12	1,25E-02
GOTERM_BP_FAT	GO:0035110~leg morphogenesis	8	5,84	110	2,67	1,65E-02
GOTERM_BP_FAT	GO:0060173~limb development	8	5,84	110	2,67	1,65E-02
GOTERM_BP_FAT	GO:0035108~limb morphogenesis	8	5,84	110	2,67	1,65E-02
GOTERM_BP_FAT	GO:0035218~leg disc development	9	6,57	110	2,41	1,91E-02
GOTERM_BP_FAT	GO:0035120~post-embryonic appendage morphogenesis	18	13,14	110	1,69	2,04E-02
GOTERM_BP_FAT	GO:0048737~imaginal disc-derived appendage development	19	13,87	110	1,61	2,67E-02
GOTERM_BP_FAT	GO:0035114~imaginal disc-derived appendage morphogenesis	19	13,87	110	1,61	2,67E-02
GOTERM_BP_FAT	GO:0048736~appendage development	19	13,87	110	1,59	3,14E-02
GOTERM_BP_FAT	GO:0035107~appendage morphogenesis	19	13,87	110	1,59	3,14E-02
GOTERM_BP_FAT	GO:0048569~post-embryonic organ development	21	15,33	110	1,50	4,12E-02
GOTERM_BP_FAT	GO:0035220~wing disc development	18	13,14	110	1,46	7,92E-02
GOTERM_BP_FAT	GO:0009791~post-embryonic development	23	16,79	110	1,37	8,05E-02
GOTERM_BP_FAT	GO:0048563~post-embryonic organ morphogenesis	19	13,87	110	1,43	8,20E-02



GOTERM_BP_FAT	GO:0007560~imaginal disc morphogenesis	19	13,87	110	1,43	8,20E-02
GOTERM_BP_FAT	GO:0048707~instar larval or pupal morphogenesis	20	14,60	110	1,39	9,44E-02
GOTERM_BP_FAT	GO:0009886~post-embryonic morphogenesis	20	14,60	110	1,37	1,05E-01
GOTERM_BP_FAT	GO:0007552~metamorphosis	20	14,60	110	1,34	1,29E-01
GOTERM_BP_FAT	GO:0007472~wing disc morphogenesis	14	10,22	110	1,41	1,60E-01
GOTERM_BP_FAT	GO:0007476~imaginal disc-derived wing morphogenesis	14	10,22	110	1,41	1,60E-01
GOTERM_BP_FAT	GO:0002165~instar larval or pupal development	20	14,60	110	1,26	2,02E-01
GOTERM_BP_FAT	GO:0007469~antennal development	3	2,19	110	2,00	4,54E-01

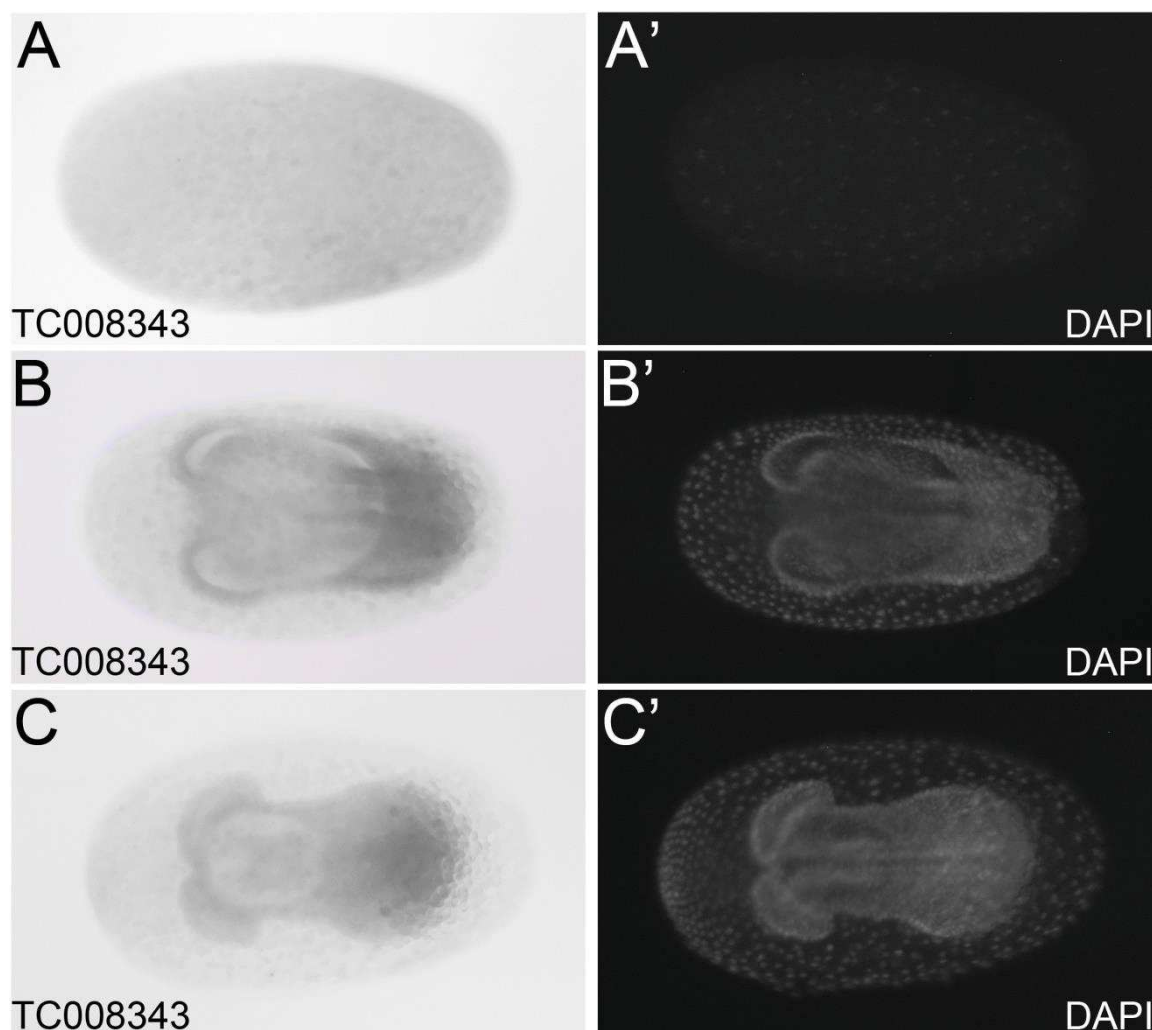
### 5.5. Expressão de alvos preditos de Tc-Zld no embrião de *T. castaneum*

Os genes conservados e possivelmente regulados por *zelda* em *T. castaneum*, encontrados pelo algoritmo de predição descrito acima, revelaram funções relacionadas a TFs e ativadores iniciais da embriogênese (Tabela 6), genes envolvidos com a segmentação, e até mesmo eventos pós-embrionários, que serão discutidos no próximo item. Contudo, alguns genes identificados não possuem função descrita na literatura e despertaram a curiosidade quanto à sua expressão espacial no embrião de *T. castaneum*. Para isso, os genes: TC008343, TC014638, TC004474, TC007382 e TC004898 foram amplificados e a partir de seus produtos foram sintetizadas sondas de RNA para hibridização *in situ*. Conjuntos de embriões em estágios mistos foram utilizados para a hibridização *in situ* dos genes alvos de Tc-Zld, para fim de observação de uma possível dinâmica de expressão ao longo da embriogênese. As sondas para os genes TC007382 e TC004898 não apresentaram marcação de seus transcritos após a hibridização, e dessa forma prosseguimos com a análise dos outros três candidatos.

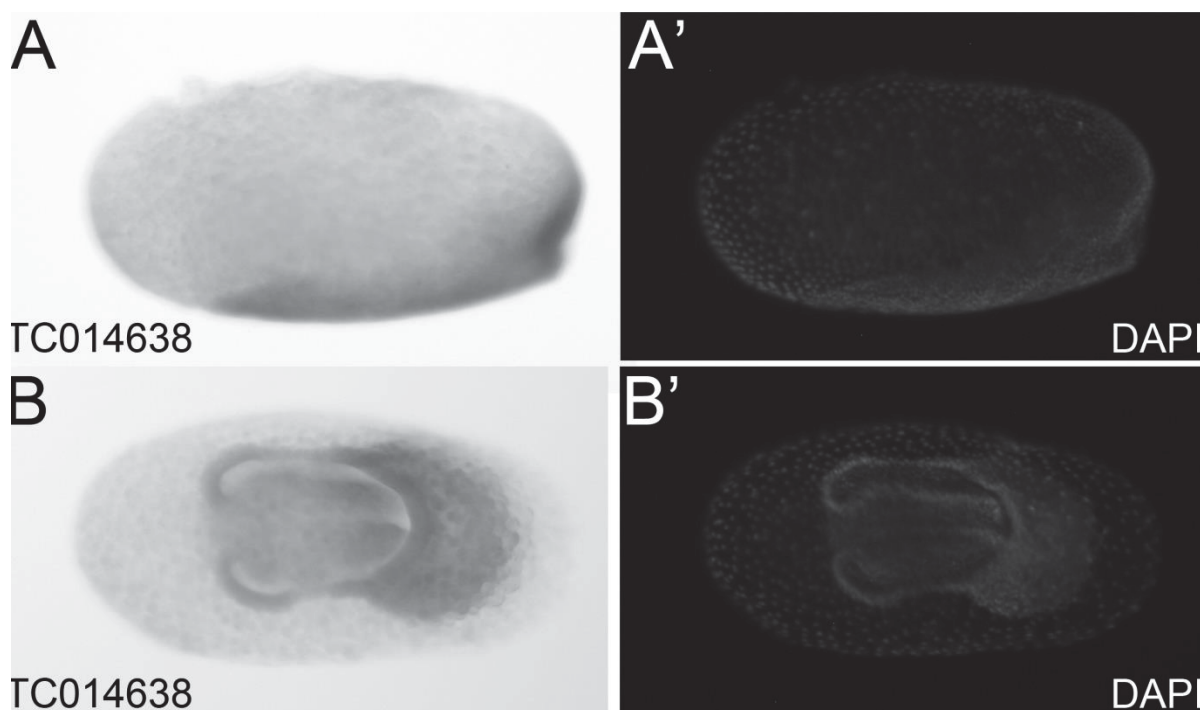
O gene TC008343 é expresso de maneira branda poucas horas após oviposição (Figura 11 AA'). Durante a invaginação, a presença dos mRNAs passa a se revelar na banda germinal, nas regiões das células do ectoderma (Figura 11 BB'). No instante do fechamento da janela da serosa, a expressão de TC008343 permanece restrita às células do embrião, com presença mais forte na região de crescimento (Figura 11 CC'). O homólogo de TC008343 em *D. melanogaster* é o gene codificante de uma proteína semelhante à Calumenina. A Calumenina é descrita como uma proteína ligadora de cálcio, expressa em células precursoras do sistema nervoso central no estágio embrionário de camundongos (Vasiljevic *et al.*, 2012).

Os transcritos do gene TC014638 são detectados nas células da banda germinal nos momentos iniciais da invaginação no embrião de *T. castaneum* (Figura 12 AA'). Durante a invaginação, a expressão de TC014638 pode ser verificada nas células da banda germinal, incluindo a região posterior que dará origem à zona de crescimento. O seu ortólogo em *D. melanogaster* codifica a proteína Slit, cuja

clivagem resulta em dois polipeptídeos que são encontrados na região da glia, ao longo da linha média do embrião (Kidd *et al.*, 1999). Recentemente, subunidades de originadas pela clivagem de Slit foram relacionadas à diferenciação do tecido muscular, sendo localizadas em células formadoras de tendões (Ordan e Volk, 2015).



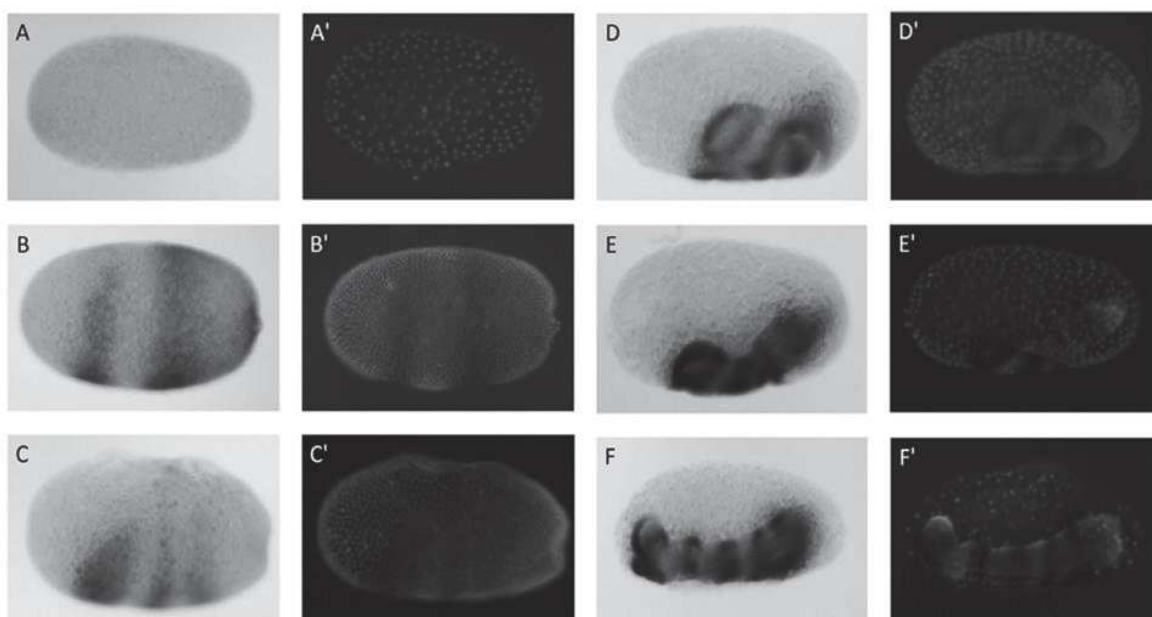
**Figura 11: Padrão de expressão do gene TC008343.** (A,A') Expressão do gene no estágio inicial da embriogênese de *T. castaneum*, durante as primeiras divisões mitóticas (A'). (B-B') O gene TC008343 é expresso na banda germinal tendo maior presença na região posterior no início da gastrulação.(C-C') Em momentos pré-fechamento da janela da serosa, o gene TC008343 mantém sua expressão na banda germinal e na região da zona de crescimento, não sendo expresso em células extra-embrionárias.



**Figura 12: Padrão de expressão do gene TC014638.** (A-A') O gene TC014638 é expresso na banda germinal momentos antes da gastrulação (A). Vista lateral do embrião de *T. castaneum* no estágio de blastoderma diferenciado evidenciado pelos núcleos marcados por DAPI (A') (B-B') Visão ventral do embrião de *T. castaneum* evidenciando a marcação por hibridização in situ para o gene TC014638 na banda germinal durante o movimento de fechamento da janela da serosa.

Estudos recentes demonstram que a padronização da zona de crescimento em *T. castaneum* depende de movimentação e intercalação de células e não da proliferação celular para a geração dos segmentos (Sarrazin *et al.*, 2012; Nakamoto *et al.*, 2015). Dentre os alvos conservados preditos (Tabela 5), um gene em especial: TC004474, identificado como *Tc-Toll7* pertence à família de genes envolvidos neste processo, assim como os já descritos *Tc-Toll2*, *Toll6* e *Toll8* (Paré *et al.*, 2014). Torna-se, portanto, interessante investigar a expressão e função de *Tc-Toll7*. Após a marcação, verificou-se que *Tc-Toll7* apresenta uma expressão dinâmica que se inicia na formação do blastoderma e se estende até a segmentação do embrião (Figura 13), indicando que seu produto gênico possa desempenhar alguma função neste processo. Pouco tempo depois de realizados esses experimentos, Benton e colaboradores (2016) demonstraram que *Tc-Toll7* em conjunto com *Tc-Toll10* são necessários para a intercalação celular durante a extensão da banda germinal de *T. castaneum*. Além disso, esta família gênica tem o padrão de expressão semelhante em outras seis espécies de artrópodes e poderia desempenhar esta função de alongamento do embrião em artrópodes basais (Benton *et al.*, 2016). Desta forma,

nos restringimos a avaliar a possibilidade de *Tc-zld* regular a expressão de *Tc-Toll7* na região posterior do embrião, conforme descrito no próximo item.

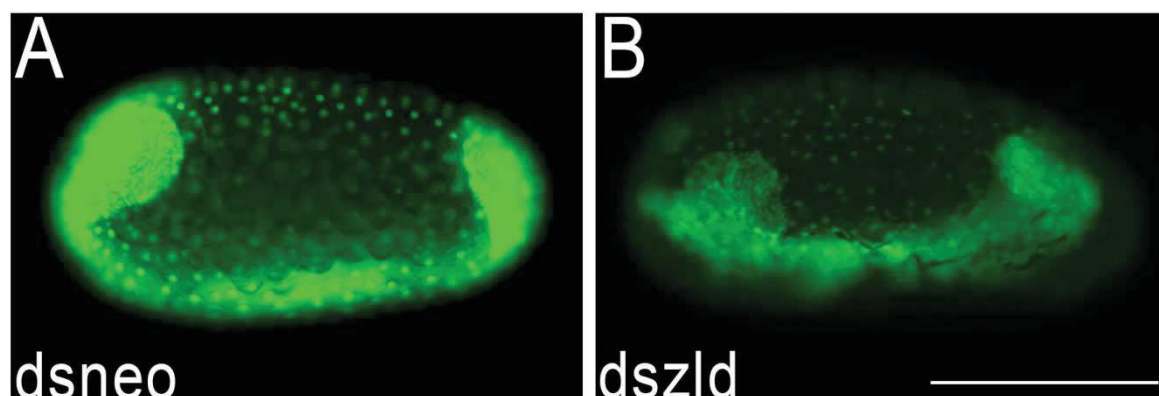


**Figura 13: Dinâmica de expressão do gene *Tc-Toll7* (TC004474).** (A,A') Embrião de *T. castaneum* no estágio pré-blastoderma celular (A') não apresenta expressão de *Tc-Toll7*. (B,B') *Tc-Toll7* exibe expressão semelhante em grandes faixas no estágio de blastoderma celular uniforme. (C,C') Na sequência, durante o blastoderma diferenciado, *Tc-Toll7* tem sua expressão na região posterior reduzida. (D,D') No momento da invaginação, *Tc-Toll7* é expresso no embrião na cabeça e na região que dará origem ao mesoderma. (E,E') A expressão de *Tc-Toll7* começa a se localizar nos segmentos à medida em que são formados. (F,F') Durante o processo de segmentação, *Tc-Toll7* é expresso na cabeça e nos parasegmentos do embrião.

### 5.6. *Tc-zld* desempenha papéis específicos na região posterior de crescimento

Enquanto o RNAi parental reduz a expressão maternal e zigótica de genes em *T. castaneum* (Bucher *et al.*, 2002), injeções feitas diretamente no embrião (eRNAi) afetam somente a expressão zigótica (Brown *et al.*, 1999; Nunes Da Fonseca *et al.*, 2008). Para investigar se *zld* seria requerido para a padronização posterior do embrião, foi injetada solução de dsRNA do gene *Tc-zld* na região posterior de embriões que expressam GFP em seus núcleos (nGFP) (Sarrazin *et al.*, 2012; Benton *et al.*, 2013). Injeções embrionárias de ds*zld* após a MZT impediram a geração de segmentos a partir da zona de crescimento, enquanto embriões injetados com ds*neo* se desenvolveram como embriões do tipo selvagem (Figura 14 AB). Além disso, a expressão do alvo predito *Tc-eve*, um TF importante para a padronização da zona de crescimento (El-Sherif *et al.*, 2012), teve sua expressão

diminuída após o silenciamento de *zld*. Portanto, nossos resultados fornecem evidências de que *Tc-zld* não está envolvido somente na MZT, padronização inicial e formação do sistema nervoso como descrito para *D. melanogaster* (Liang *et al.*, 2008), mas também desempenha um papel importante para a geração de segmentos a partir da zona de crescimento, estrutura exclusiva de insetos do tipo de banda germinal curta, como o *T. castaneum*.

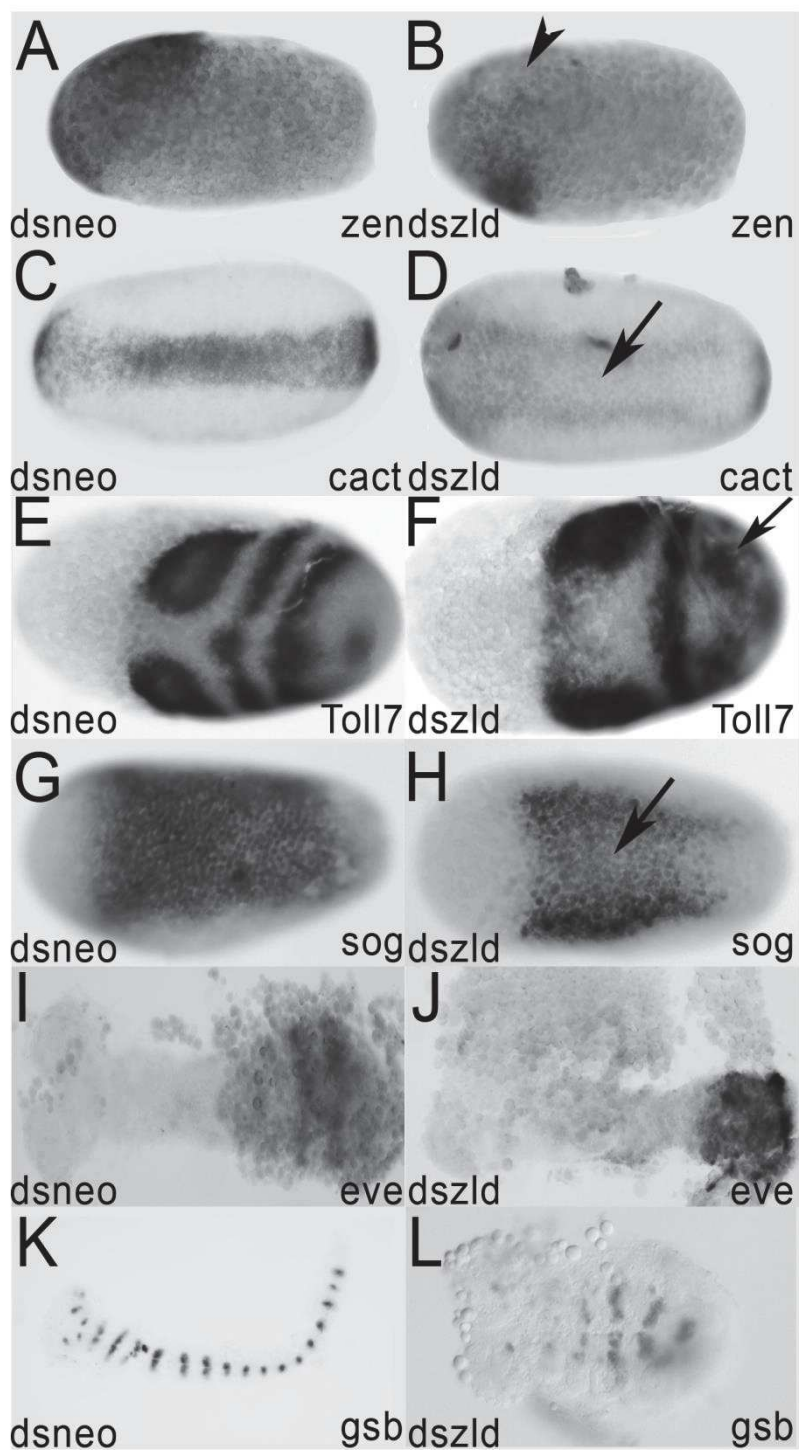


**Figura 14: O silenciamento embrionário de *Tc-zld* interrompe a segmentação.** A injeção de solução de dsRNA de *Tc-zld* nos embriões expressando GFP afetou a segmentação a partir da zona posterior de crescimento (B), quando comparado ao embrião controle em pleno estágio de banda germinal alongada (A). Escala: 250  $\mu$ m.

Visto que o silenciamento zigótico de *Tc-zld* tem influência direta na segmentação do embrião, buscamos definir quais genes poderiam sofrer alterações na sua expressão devido ao silenciamento parental de *Tc-zld*. Após o pRNAi de *Tc-zld*, encontramos genes alvo conservados que sofreram diminuição na expressão no embrião, como genes zigóticos envolvidos na padronização anteroposterior, como por exemplo o gene *Tc-zerknüllt* (*Tc-zen*), expresso na região da serosa (Figura 15 AB) (Van Der Zee *et al.*, 2005).

Mudanças na distribuição espacial de transcritos preditos para os genes alvo dorsoventrais também foram observadas após o pRNAi de *Tc-zld*. Nos embriões de *T. castaneum*, o TF Dorsal forma um gradiente dinâmico e transiente, que ativa *Tc-cactus* (*Tc-cact*) e *Tc-short-gastrulation* (*Tc-sog*) em sua região ventral (Chen *et al.*, 2000; Nunes Da Fonseca *et al.*, 2008). Após o RNAi de *Tc-zld*, a expressão de *Tc-cact* e *Tc-sog* foi observada em um padrão de dois domínios laterais, em contraste à expressão em um único domínio ventral encontrada nos embriões *dsneo*. Estes resultados sugerem que *Tc-zld* é requerido para a atividade adequada de Tc-Dorsal na região ventral do embrião.

Conforme demonstrado, *Tc-zld* é expresso na região posterior do embrião (Figura 7 H) e está relacionado à segmentação e regionalização (Tabela 6). Além disso, certos alvos prováveis de *Tc-zld* estão envolvidos em segmentação posterior, como *caudal* (*Tc-Cdx*), *even-skipped* (*Tc-Eve*) e alguns genes Hox, como *Ultrabithorax*, *Abdominal-A* e *Abdominal-B*. O gene *Tc-eve*, por exemplo, é essencial para o estabelecimento da rede regulatória requerida para a segmentação posterior (Choe *et al.*, 2006). Neste caso, o silenciamento de *Tc-zld* demonstrou um padrão de expressão de *Tc-eve* como um domínio contínuo na zona de crescimento ao invés de um padrão em faixas que é requerido para a segmentação do tipo selvagem (Figura 15 IJ). O gene *Tc-Toll7* (TC004474) não demonstrou ter sua expressão afetada na região anterior do embrião, enquanto na região posterior sua expressão foi reduzida (Figura 15 EF). Em suma, o produto de *Tc-zld* regula a expressão de diversos genes que são necessários para a padronização anteroposterior (*Tc-zen*, *Tc-mlpt*) e dorsoventral (*Tc-sog*, *Tc-cact* e *Tc-twist*) e, em um segundo momento, de genes requeridos para o alongamento posterior (*Toll7* e *Tc-eve*) e polarização de segmentos (*Tc-gsb*).



**Figura 15: O RNAi parental de *Tc-zld* afeta genes importantes para a padronização do embrião de *T. castaneum*.** (A,B) A expressão de *Tc-zen* na região da serosa é reduzida após o pRNAi de *Tc-zld* (B, seta) quando comparado ao controle injetado com dsneo. (C) A expressão do gene de padronização dorsoventral *Tc-cactus* é reduzida após o pRNAi de *Tc-zld* (D, seta). (E,F) A expressão do gene *Tc-toll7* sobrepõe-se a zona de crescimento na região posterior no embrião controle (E), enquanto sua expressão na região posterior do embrião se encontra dispersa em pequenos grupamentos no embrião silenciado para *Tc-zld* (F). (G,H) O gene de padronização dorsoventral *Tc-sog* é expresso na região ventral do embrião controle (G), porém no embrião dszld, um grande domínio perde parte de sua expressão como indicado pela seta (H). (I,J) *Tc-eve* é expresso em faixas na região



posterior de crescimento, porém no embrião silenciado *dsz/d* sua expressão não obedece os limites das faixas. (K,L) A segmentação proveniente da zona de crescimento sofre interrupção de após silenciamento por *dsz/d*, o padrão de expressão do gene de parasegmentos *Tc-gsb* é presente em todos os segmentos do embrião controle (I) enquanto no silenciado, somente a expressão até os segmentos torácicos é detectada, tendo o embrião falhado em se segmentar (K).

### 5.7. Análise do Transcriptoma de embriões silenciados de *Tc-zld*

As ferramentas de sequenciamento de segunda geração têm sido utilizadas para na compreensão em nível molecular de processos biológicos complexos em diversos organismos. Em insetos, vários grupos utilizaram o sequenciamento do transcriptoma (RNAseq) para avaliar a expressão diferencial de genes sob determinadas condições. No caso do *T. castaneum*, existem estudos publicados de transcriptoma para identificação de transcritos maternalmente carregados no ovócito do besouro (Preuss *et al.*, 2012), trabalhos abordando o sistema imune (Altincicek *et al.*, 2013; Behrens *et al.*, 2014) e de biologia do desenvolvimento (Stappert *et al.*, 2016). Para avaliar a influência de *Tc-zld* na expressão global durante a embriogênese, realizamos sequenciamento de RNA na plataforma Illumina HiSeq 2500 de embriões de 7,5h a 11,5 h de desenvolvimento, correspondendo ao fechamento da janela da serosa e início da extensão da banda germinal.

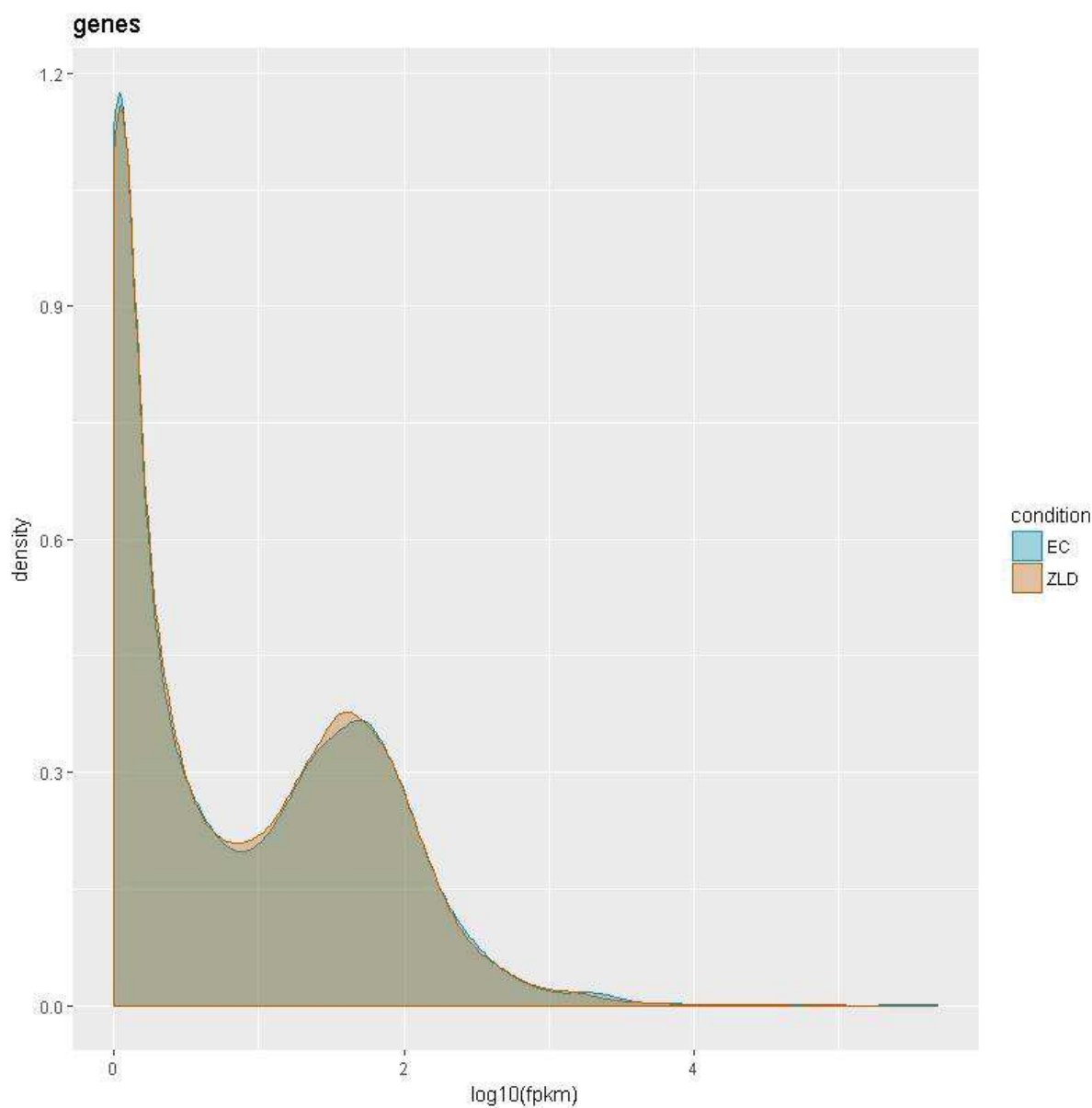
Foram gerados nesta análise um total 95 milhões de leituras *paired-end* por tratamento (Tabela 7). A taxa de mapeamento das leituras ao genoma de referência foi de 79%, sendo que 6% alinharam à múltiplas regiões do genoma de referência. Foram encontrados 2583 genes expressos (com FPKM > 1) no estágio em que os embriões foram coletados, na janela correspondente a 7,5 - 11,5 h após a oviposição, o que corresponde a aproximadamente 15% dos 16590 genes preditos contidos no genoma de referência de *T. castaneum* (Tribolium Genome Sequencing *et al.*, 2008).

Tabela 7: Sumário do alinhamento das leituras no genoma de *T. castaneum*. Dados de alinhamento das replicatas utilizadas no experimento controle EC e silenciado *ZLD*.

Amostra	Total de leituras		Total de pares alinhados	Múltiplos alinhamentos	Leituras não mapeadas	
	Left	Right			Left	Right
EC_0	7742683	7742683	6150233	464452	1357346	1331470
EC_1	13934033	13934033	9884792	2135843	3155090	3428026
EC_2	12489888	12489888	10002039	717626	2101983	2065747
ZLD_0	42603539	42603539	35443090	1300513	5970321	5940441

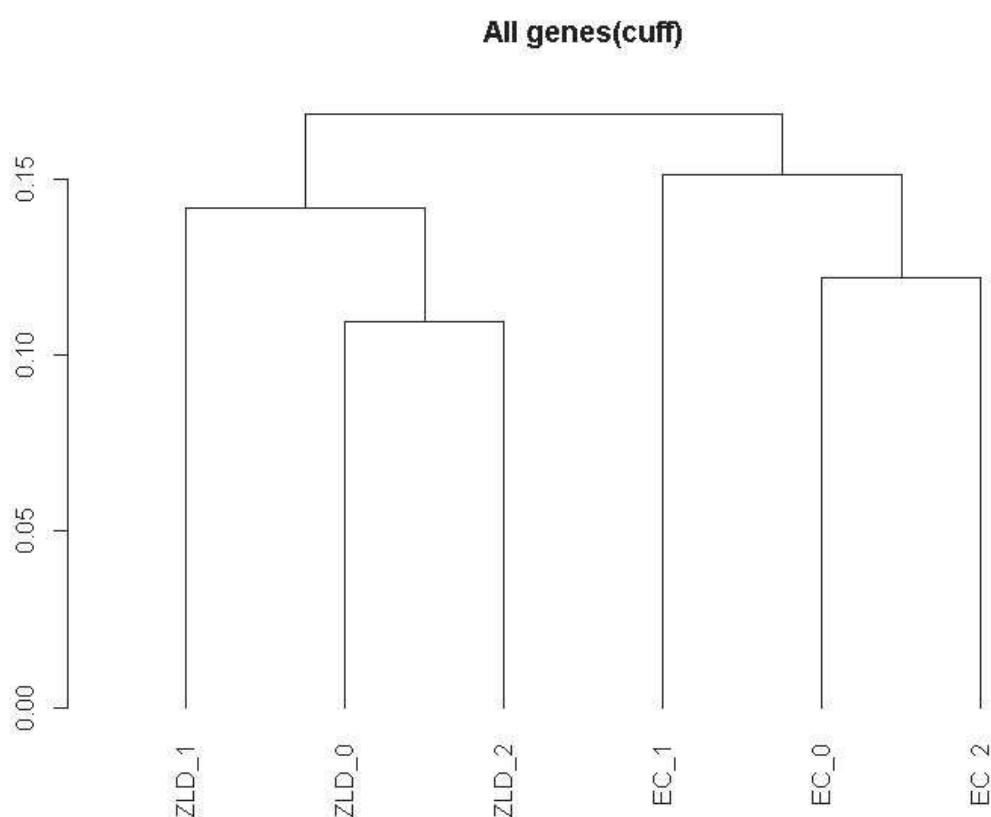
<b>ZLD_1</b>	4118237	4118237	3200132	297363	750007	768747
<b>ZLD_2</b>	14997442	14997442	11739199	1055295	2631448	2714424

A densidade de expressão gênica encontrada em ambos os tratamentos é representada pelo gráfico na Figura 16, onde temos a distribuição dos valores de FPKM entre as amostras controle (EC) e *dszld* (ZLD). O resultado demonstra uma maior expressão gênica global no tratamento controle (EC), tendo a amostra ZLD maior densidade de transcritos com baixos níveis de expressão.



**Figura 16: Densidade de expressão gênica em FPKM para as amostras controle (EC) e *dszld* (ZLD).** O gráfico demonstra a distribuição da expressão gênica entre as diferentes amostras. Onde  $\log_{10} \text{fpkm } 0 = 1 \text{ FPKM}$ .

Para demonstrar a qualidade do sequenciamento, analisamos a similaridade e divergência da expressão entre as replicatas. Baseado no método estatístico Jensen-Shannon (*JS distance*) tornou-se possível estimar a divergência entre as amostras dos transcriptomas. Cada replicata foi comparada entre si e a partir disso foi construído um dendograma para demonstrar a relação entre as amostras. Segundo o dendograma, as amostras relativas a cada tratamento co-segregaram, constituindo dois grupos distintos e evidenciando a divergência entre o grupo controle e silenciado para *Tc-zld* (Figura 17).



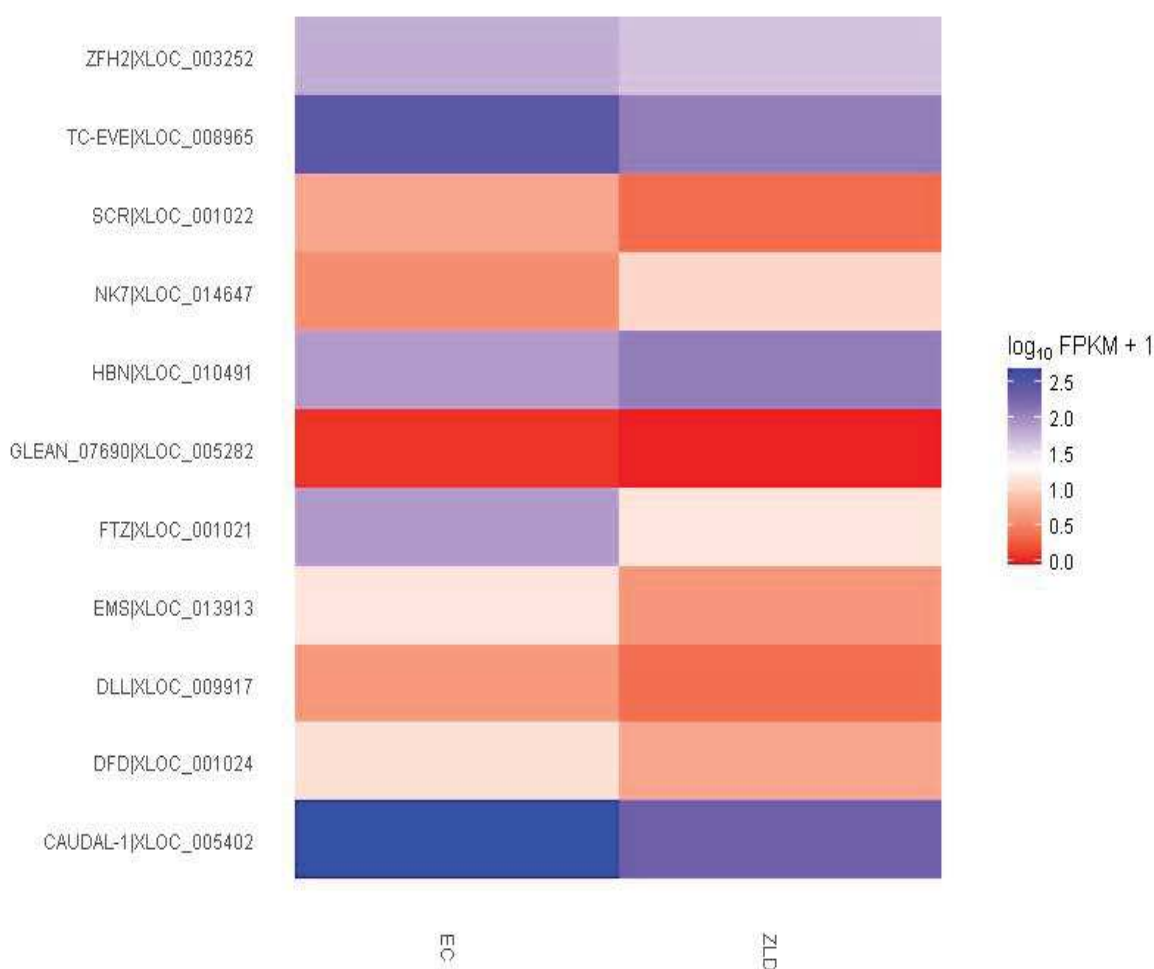
**Figura 17: Dendograma da distância JS demonstrando as relações entre as replicatas de cada tratamento.** Esta análise calcula a divergência no nível de expressão de genes (em FPKM) entre as amostras. Amostras similares tendem a se agrupar, neste caso replicatas silenciadas (ZLD\_1, ZLD\_0 e ZLD\_2) e amostras controle (EC\_1, EC\_0 e EC\_2).

### 5.8. Identificação de genes diferencialmente expressos por análise do transcriptoma

Ao compararmos o transcriptoma dos embriões provenientes de fêmeas injetadas com dsRNA controle e *Tc-zld* buscamos identificar genes que poderiam ser regulados por *Tc-zld*. Como controle, um dsRNA portando a sequência do gene de

resistência à neomicina (*dsneo*) foi usado para distinguir os efeitos observados de possíveis mudanças de expressão gênica causadas pelo procedimento de injeção propriamente dita. Após a filtragem dos genes expressos e realização de análises estatísticas, identificamos genes que sofreram alteração estatisticamente significativa ao compararmos os tratamentos EC e ZLD, controle e *zld* respectivamente. Ao todo 2583 genes apresentaram expressão diferencial após o silenciamento de *Tc-zld*. Deste conjunto, 397 genes foram considerados regulados positivamente ( $\log_2\text{fold} > 1$ ) e 396 genes regulados negativamente ( $\log_2\text{fold} < 1$ ) quando comparados ao controle ( $p < 0.05$ ). Notavelmente, alguns genes encontrados em nossa predição *in silico* figuraram entre os genes que sofrem influência de zelda pela análise do transcriptoma. Para melhor analisarmos o perfil de expressão, agrupamos alguns genes de acordo com sua função. Como análise preliminar, foram selecionados genes classificados pelo termo *Segmentation* (GO:0035282), segundo a tabela de predição de função criada a partir de DAVID (Tabela 6).

Os genes *Tc-eve* e *Tc-Dll* foram reprimidos nos embriões silenciados *dszld*, demonstrando uma forte influência de *Tc-zld* sobre a expressão desses genes de segmentação (Figura 18). Esses dados demonstram o potencial de *Tc-zld* na regulação de outros TFs importantes para o desenvolvimento em *T. castaneum*. Além disso, os genes *sex combs reduced* (*scr*) e *fushi-tarazu* (*ftz*), genes homeótico e da regra dos pares, respectivamente, também tiveram sua expressão reduzida após o silenciamento de *Tc-zld*, indicando o seu potencial de regulação nos eventos da embriogênese. Nota-se também a diminuição da expressão de *Tc-caudal*, um gene importante para a regulação de *Tc-eve* nas células da região posterior do embrião (El-Sherif *et al.*, 2014).

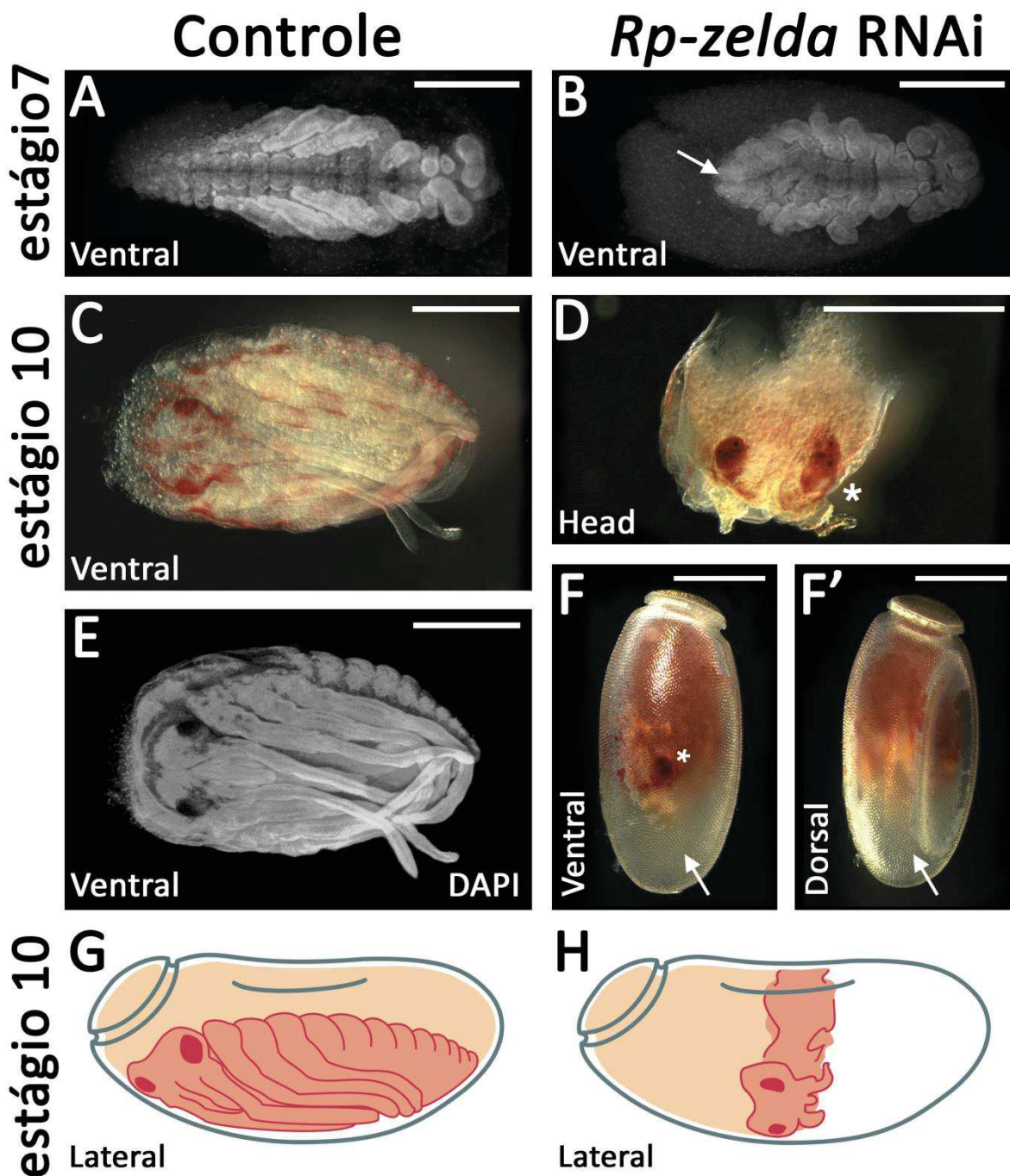


**Figura 18: Mapa de calor com os níveis transcricionais dos genes classificados como *Segmentation* (GO:0035282) pela análise de DAVID (Huang *et al.*, 2009) ..As cores azul e vermelha indicam os valores de  $\log_{10}$ FPKM mais aproximados de 2,5 e zero, respectivamente. Nomes dos genes acompanhados de “XLOC\_”, identificadores adotados pelo programa cufflinks para regiões delimitadas no genoma.**

### 5.9. ZeldA possui um papel conservado no hemíptero *Rhodnius prolixus*

Embora tenha sido demonstrado o envolvimento de *zld* na segmentação de *T. castaneum*, é importante também conhecer a conservação dessa função em outras espécies. Com esse objetivo, analisamos a função do ortólogo de *zld* no hemíptero *Rhodnius prolixus* (*Rp-zld*), um representante dos insetos hemimetábolos, ou seja, de insetos que não apresentam metamorfose completa (Lavore *et al.*, 2012; Mesquita *et al.*, 2015). O silenciamento de *Rp-zld* por pRNAi resultou em um fenótipo no qual os embriões desenvolvem somente as regiões anteriores, que compreendem os segmentos da cabeça, maxila e tórax (Figura 19). Estes resultados suportam que *zld* está envolvido na segmentação não somente no *T. castaneum*,

mas também em outro inseto de banda germinal curta mais basal, como o *R. prolixus*. Com isso, sugerimos que o papel de *zld* na segmentação posterior estaria presente no ancestral comum dos insetos hemimetábolos e holometábolos.



**Figura 19: O gene *zld* é requerido para a geração dos segmentos posteriores de *Rhodnius prolixus*.** (A,C,E) Embriões controle (*dsneo*) de *R. prolixus*. (B,D,F,F') Embrião proveniente de fêmea injetada com dsRNA de *Rp-zld*. (F,F') Embrião dentro do córion, vista ventral e dorsal. (C,E) Embrião removido do ovo. (G,H) Representação esquemática dos embriões controle e *zld* pRNAi. (D) O asterisco indica o olho, observado na região ventral do embrião (F). (D) Após a dissecação, o olho pode ser identificado por sua pigmentação vermelha. Escala: 500  $\mu$ m.

Durante a MZT em *D. melanogaster*, Zld age como um TF pioneiro favorecendo a abertura da cromatina para os genes expressos no início da embriogênese (Harrison *et al.*, 2011; Li *et al.*, 2014). Em especial, Zld aumenta a acessibilidade da cromatina de TFs importantes envolvidos na padronização DV(*dorsal*) e AP (*bicoid*) (Foo *et al.*, 2014; Xu *et al.*, 2014). A adição ou remoção de sítios de ligação de *Zld* influencia no tempo de ativação dos primeiros genes zigóticos em *D. melanogaster* (Nien *et al.*, 2011; Foo *et al.*, 2014), sugerindo que *Zld* atua como um controlador do tempo de ativação durante o desenvolvimento. Nossas análises na expressão do mRNA de *Tc-zld* demonstraram uma contribuição maternal importante, seguida de expressão zigótica no sentido anteroposterior. Pouco depois da gastrulação, a expressão de *Tc-zld* é confinada na região posterior do embrião. Desse modo, propõe-se que *zld* contribua progressivamente para a abertura da cromatina no eixo anteroposterior em *T. castaneum* e que as células da zona de crescimento retêm um estado desdiferenciado. Em *Danio rerio* (o peixe zebra), a proteína Pou5f1, um homólogo do TF Oct4 de células pluripotentes de mamíferos, ocupa os sítios de ligação SOX-POU antes do estabelecimento da transcrição zigótica e ativa os primeiros genes zigóticos (Leichsenring *et al.*, 2013). Além disso, os sítios de ligação de Pou5f1 e Sox2, seus estados da cromatina, e a ligação de RNA pol II são similares em embriões de *D. rerio* e células tronco embrionárias de mamíferos, sugerindo uma função ancestral do fator de pluripotência de vertebrados na ativação de genes zigóticos e no controle temporal da expressão no início da embriogênese de vertebrados. O presente trabalho propõe que células expressando *zld* localizadas na região embrionária posterior são capazes de manter a pluripotência durante os estágios iniciais do alongamento da banda germinal, enquanto passam por movimentos de extensão convergente (Sarrazin *et al.*, 2012; Nakamoto *et al.*, 2015).

A perda da expressão de *zld* levaria à diferenciação prematura dos tecidos e à perda da extensão convergente, conduzindo conseqüentemente à falha na segmentação. Sendo *zld* também expresso e importante na região posterior do inseto hemimetábolo *R. prolixus*, o papel de *zld* no desenvolvimento da região posterior pode datar desde o ancestral comum de Paraneoptera. Além do papel conhecido de *zelda* na MZT em *D. melanogaster* (Liang *et al.*, 2008), nossos resultados revelaram dois novos papéis biológicos de *zld* no besouro *T. castaneum*:

a regulação da geração de segmentos a partir da região posterior de crescimento (*growth zone*) durante a embriogênese e a padronização das estruturas derivadas de discos imaginais. A padronização da zona de crescimento e os discos imaginais possuem em comum o fato destes processos envolverem mudanças morfológicas extensas e modificações do *status* de diferenciação das células precursoras.

#### **5.10. A função pós embrionária de *Tc-zld***

Alguns insetos possuem versões menores de suas estruturas corporais adultas que sofrem crescimento em tamanho ao longo do desenvolvimento, como patas em insetos ametábolos. Entretanto, em insetos como a *D. melanogaster* e *T. castaneum*, que passam por mudanças incisivas durante seu ciclo de vida, as estruturas precursoras não sofrem somente crescimento, mas passam por mudanças substanciais durante a metamorfose. As estruturas que darão origem aos apêndices externos no adulto permanecem protegidas internamente na larva até que ocorra a metamorfose. Essas estruturas precursoras são referidas como imaginais, levando em conta que irão dar origem a estruturas do corpo adulto, conhecido como “imago”, e se desenvolvem a partir da epiderme (Truman e Riddiford, 1999).

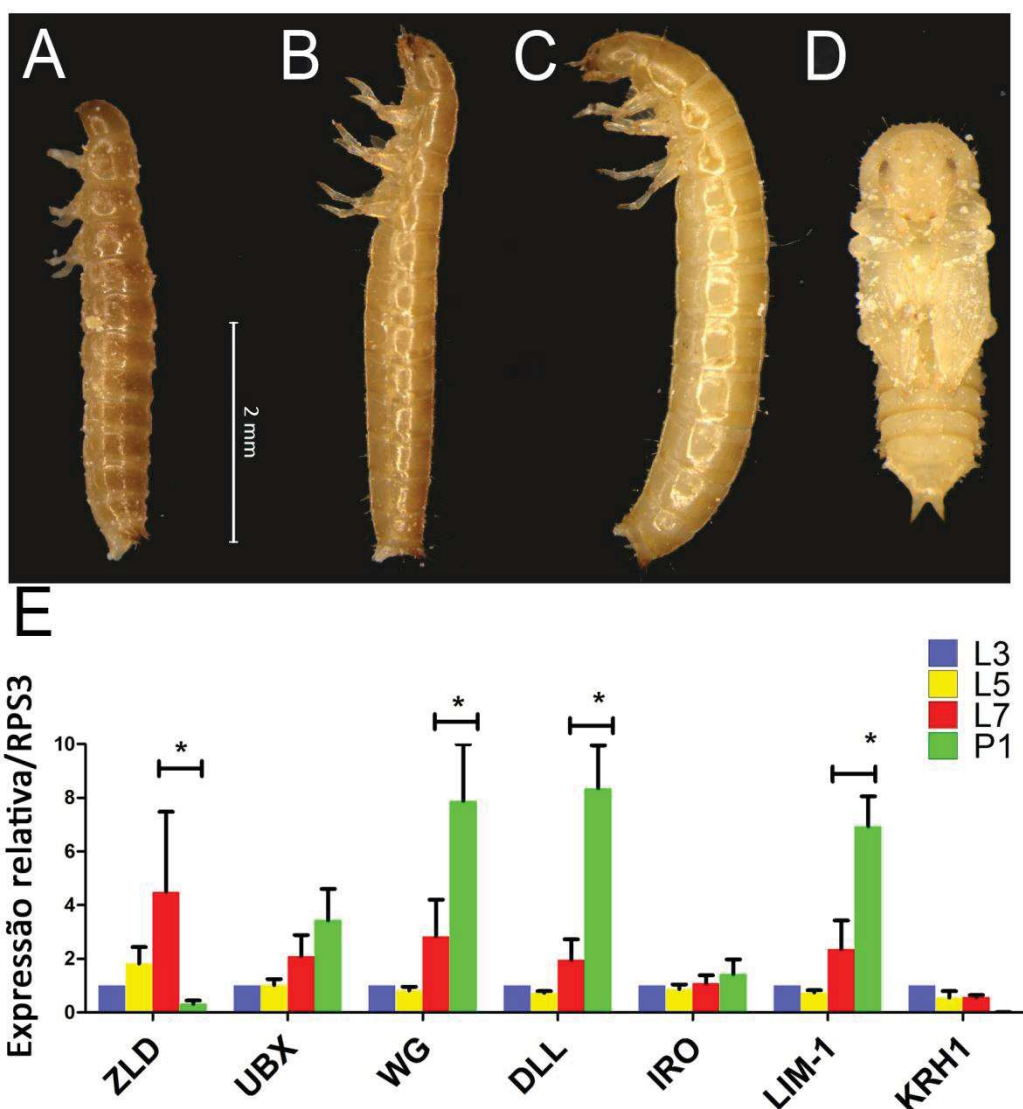
Morfologicamente, discos imaginais são tecidos invaginados que persistem e são padronizados por múltiplos instares larvais para dar origem a estruturas distintas em adultos após a metamorfose. Em *D. melanogaster*, células que formam os discos imaginais são destacadas durante a embriogênese, não contribuem para o desenvolvimento larvar e formam não só os apêndices, mas também a parede corporal da cabeça e tórax do adulto (Cohen *et al.*, 1993). Na metamorfose, a epiderme larvar se degenera, sendo substituída por tecidos derivados dos discos imaginais. Acredita-se que os discos imaginais sofreram evolução independente em conjunto com perda dos apêndices larvares em Holometabola (Truman e Riddiford, 1999). Por exemplo, larvas de mosca não possuem apêndice algum na cabeça e no tórax. Os discos imaginais servem como uma forma de iniciar o desenvolvimento de asas antes do desenvolvimento do último instar larvar e não evoluíram somente em Diptera, mas também em Coleoptera, Lepidoptera e alguns Himenoptera (Svácha, 1992).

As análises computacionais da interação de *Tc-zld* com o genoma de *T. castaneum* revelaram um conjunto de genes possivelmente regulados em estágios



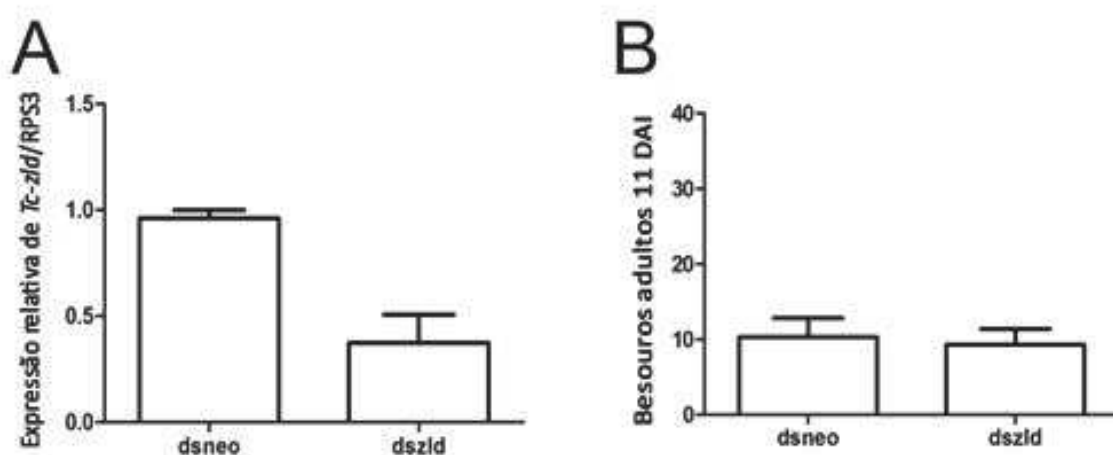
pós-embrionários. Observando os resultados das análises de função expostos pela ferramenta DAVID (Tabela 6), os 29 alvos putativos de *Tc-zld* estão, de alguma forma, relacionados com desenvolvimento de discos imaginais (Tabela 6, GO:0007444). Entre os genes listados estão alguns TFs contendo homodomínios como: *distalless (Dll)*, *abdominal A (Abd-A)*, *abdominal B (Abd-B)*, *zen*, *engrailed (En)*, *caudal (cad)*, *defective proventriculus (dve)*, *mirror (mirr)*, *araucan (ara-iroquois)*, *drop (dr)* e outros fatores como *dachsund (dac)*, *taranis (tara)* e *lim1*, *poxN*, *kn*, *sob*, *drm*, *awh* e *dp*.

Para iniciar a caracterização pós-embrionária da função de *zld*, foi analisada a expressão de diversos genes, incluindo o próprio *Tc-zld*, por qRT-PCR em larvas no terceiro (L3), quinto (L5) e sétimo (L7) estágios, além do primeiro estágio de pupa (P1) (Figura 20). Segundo a análise, a expressão de *Tc-zld* aumenta ao longo dos estágios larvares, decaindo somente após a metamorfose para pupa (Figura 20 D). Curiosamente, um processo que acontece durante os últimos estágios larvares é o desenvolvimento dos discos imaginais, sugerindo que *Tc-zld* possa ser requerido para a padronização de estruturas derivadas dos discos, como antena, patas e asas (Tomoyasu e Denell, 2004; Tomoyasu *et al.*, 2005). Além disso, constatamos que três dos cinco genes alvo preditos de *Tc-zld* analisados apresentam aumento na expressão durante os últimos estágios larvares e desenvolvimento em pupa (Figura 20 E). A função de *Tc-wg* vai além da padronização de eixos durante a embriogênese e tem sua função de formação dos ligamentos entre segmentos nos membros de insetos (Ober e Jockusch, 2006). *Tc-lim1* é expresso nas patas, antena e envolvidos na padronização e segmentação de apêndices do adulto (Angelini *et al.*, 2009). *Tc-Dll*, um conhecido marcador molecular de discos imaginais, possui expressão característica nas regiões mais distais de membros, e sua função conservada em *T.castaneum* com fenótipos resultantes defeitos de segmentação nas patas (Panganiban *et al.*, 1994; Suzuki *et al.*, 2009).



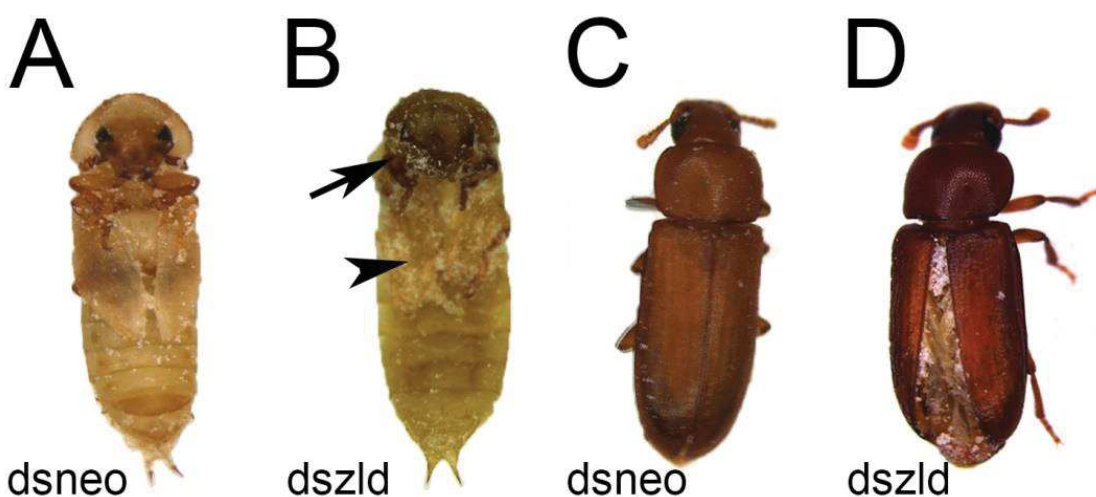
**Figura 20: Expressão de *Tc-zld* e potenciais genes alvo relacionados a metamorfose:** (A-D) Morfologia da larva de *T. castaneum* no terceiro (A-L3), quinto (B-L5), sétimo instares larvares (C-L7) e primeiro estágio de pupa (D-P1). (E) Expressão relativa de *zld*, *ubx*, *wingless* (*wg*), *distalless* (*dll*), *Iroquois* (*iro*), *lim-1* e *kruppel-homolog-1* (*krh-1*) nos estágios L3, L5, L7 e P1. O RT-PCR em tempo real foi normalizado pela expressão do gene *Tc-rps3* como descrito por (Lord *et al.*, 2010). Asteriscos representam diferenças significativas entre os estágios ( $P < 0,05$ ).

Para investigar os possíveis papéis de *Tc-zld* no estágio pós-embriônico, injetamos duas construções de dsRNA de *Tc-zld* nos estágios larvares L3 e L6, como previamente descrito (Tomoyasu e Denell, 2004) e em seguida monitoramos a expressão relativa de *Tc-zld* para confirmação do silenciamento. De acordo com os dados de qRT-PCR, *Tc-zld* apresentou diminuição de expressão após a injeção de dsRNA (Figura 21 A). Embora a expressão de *Tc-zld* tenha sido reduzida, as injeções de dsRNA em larvas não afetaram significativamente a emergência de adultos, se comparado ao controle dsneo (Figura 21 B).



**Figura 21: Eficiência do RNAi de *Tc-zld* no estágio pós embrionário:** Larvas de *T. castaneum* foram injetadas no quinto instar larval como descrito na literatura (Tomoyasu e Denell, 2004). (A) Expressão relativa do gene *Tc-zld* após injeção de dsRNAs de *Tc-zld* e *neo*. O RT-PCR em tempo real foi normalizado pela expressão do gene *Tc-rps3* como descrito por (Lord *et al.*, 2010). (B) Número de besouros adultos que emergiram onze dias após silenciamento larval de *Tc-zld*.

Injeções de dsRNA em estágios larvares diferentes produziram fenótipos em diferentes estágios ao longo da metamorfose. Ao injetarmos *dszld* em larvas mais jovens (L3), pudemos observar fenótipos marcantes ainda no estágio de pupa. A pupa *dszld* apresentou pigmentação semelhante a de besouros adultos na região da cabeça e do tórax, além de defeitos nas asas em formação (setas, Figura 22 B).



**Figura 22: O silenciamento de *Tc-zld* afeta diferentes estágios da vida do besouro:** Larvas de *T. castaneum* foram injetadas com soluções de dsRNA nos estágios mais jovem (3°) e mais tardio (6°) da larva. (A,B) Morfologia da pupa obtida pela injeção de *dsneo* e *dszld*. Pupa de *dszld* apresentando pigmentação diferencial na região da cabeça (seta) e asas reduzidas, enquanto a pupa *dsneo* apresenta características do tipo selvagem. (C,D) Adultos obtidos por injeção no estágio larval por injeções de *dsneo*(C) e *dszld*(D), onde as asas não se dobram corretamente sob os élitros.

Por outro lado, indivíduos injetados no estágio L6 não apresentaram defeitos durante a fase de pupa, vindo a demonstrar fenótipo somente no estágio adulto. Um defeito mais visível em besouros *dsz/d* foi a aparência de “carapaça aberta” causada pelo não fechamento do élitro na região dorsal do abdômen e a dificuldade das asas permanecerem dobradas sob os élitros. Os besouros (50% dos indivíduos injetados) que apresentavam este fenótipo morriam após algum tempo devido a exposição direta da região dorsal ao meio externo, o que provavelmente causava o ressecamento dos indivíduos.

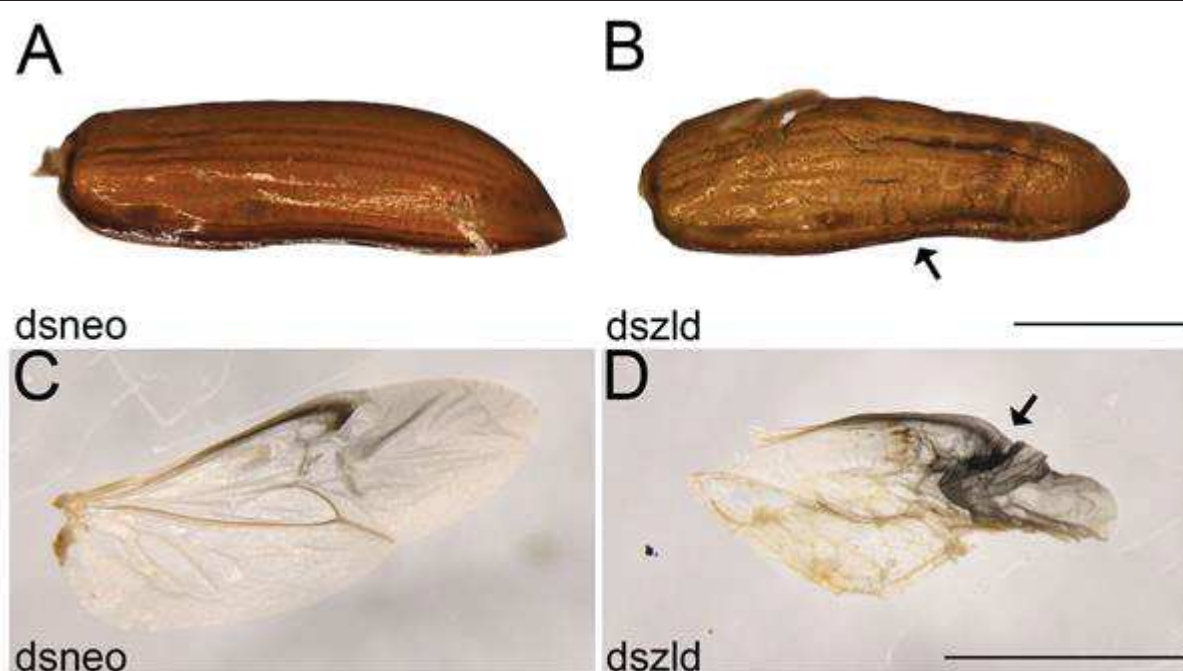
Ao observarmos inicialmente o fenótipo nos apêndices dorsais, dissecamos as asas e élitros e fixamos essas estruturas em etanol 95% por 24 h para posterior análise morfológica. Curiosamente, as asas e élitros do besouro *dsz/d* sofreram um processo de ressecamento, causando diminuição em seu volume, além de apresentarem élitro mais flexível. Estruturas retiradas de indivíduos injetados com *dsneo* apresentavam morfologia idêntica ao tipo selvagem (Figura 23 A,C). Os élitros, asas modificadas conhecidas por sua rigidez que confere proteção contra choques mecânicos, predação e desidratação (Tomoyasu *et al.*, 2009), apresentaram rugas em sua superfície e maior fragilidade perante a manipulação em besouros *dsz/d* (seta, Figura 23 B). Já as asas membranosas, que se localizam dobradas sob os élitros e tem como principal função o voo, sofreram diminuição drástica em seu volume e apresentavam maior melanização na região distal (seta, Figura 22 D).

O élitro rígido quando fechado sela as asas e segmentos abdominais sob o espaço subelital. Esta adaptação permite os besouros penetrarem em uma diversidade de nichos como solo, madeira, cascas de árvores e água, facilitando sua irradiação adaptativa. O selamento do élitro ocorre por meio de diversos encaixes entre os perímetros dos élitros e do corpo, entre o élitro e as asas e entre os dois élitros (Frantsevich *et al.*, 2005).

Nossos resultados demonstraram que após o silenciamento de *Tc-z/d* no estágio larvar, indivíduos adultos que apresentam fenótipo desenvolvem élitros com aparência fragilizada e mais flexíveis (Figura 23 B). Juntamente a essa característica, os élitros fixados em etanol exibem baixíssima resistência à dissecação, qualidade imprescindível às estruturas do seu par tipo selvagem. Essas propriedades de barreira contra a perda de água e defesa contra outros estresses

físicos tornam os élitros essenciais para a proteção das asas membranosas e da própria região dorsal do abdômen (Linz *et al.*, 2016). Mesmo o *T. castaneum* possuindo uma grande tolerância à perda de água, diante do silenciamento de *Tc-zld*, os besouros que apresentaram fenótipo morreram depois de duas semanas em exposição ao meio composto por farinha de trigo, provavelmente devido à exposição de seu fino tegumento abdominal.

Logo após a eclosão, os élitros se apresentam como estruturas de baixa pigmentação e rigidez, contendo uma cutícula frágil e branca. Os élitros vão se tornando rígidos e mais escuros durante a maturação da cutícula. Estruturalmente, os élitros do besouro demonstram um padrão de venação distinto das asas. O élitro é composto de veias paralelas que se estendem ao longo de toda sua estrutura e contribuem para sua resistência física como forma de distribuição de forças mecânicas exercidas sobre ele (Lomakin *et al.*, 2010). Os élitros de besouros silenciados para *Tc-zld*, além de apresentarem menor rigidez durante a maturação também sofreram a perda de seu padrão de venação.

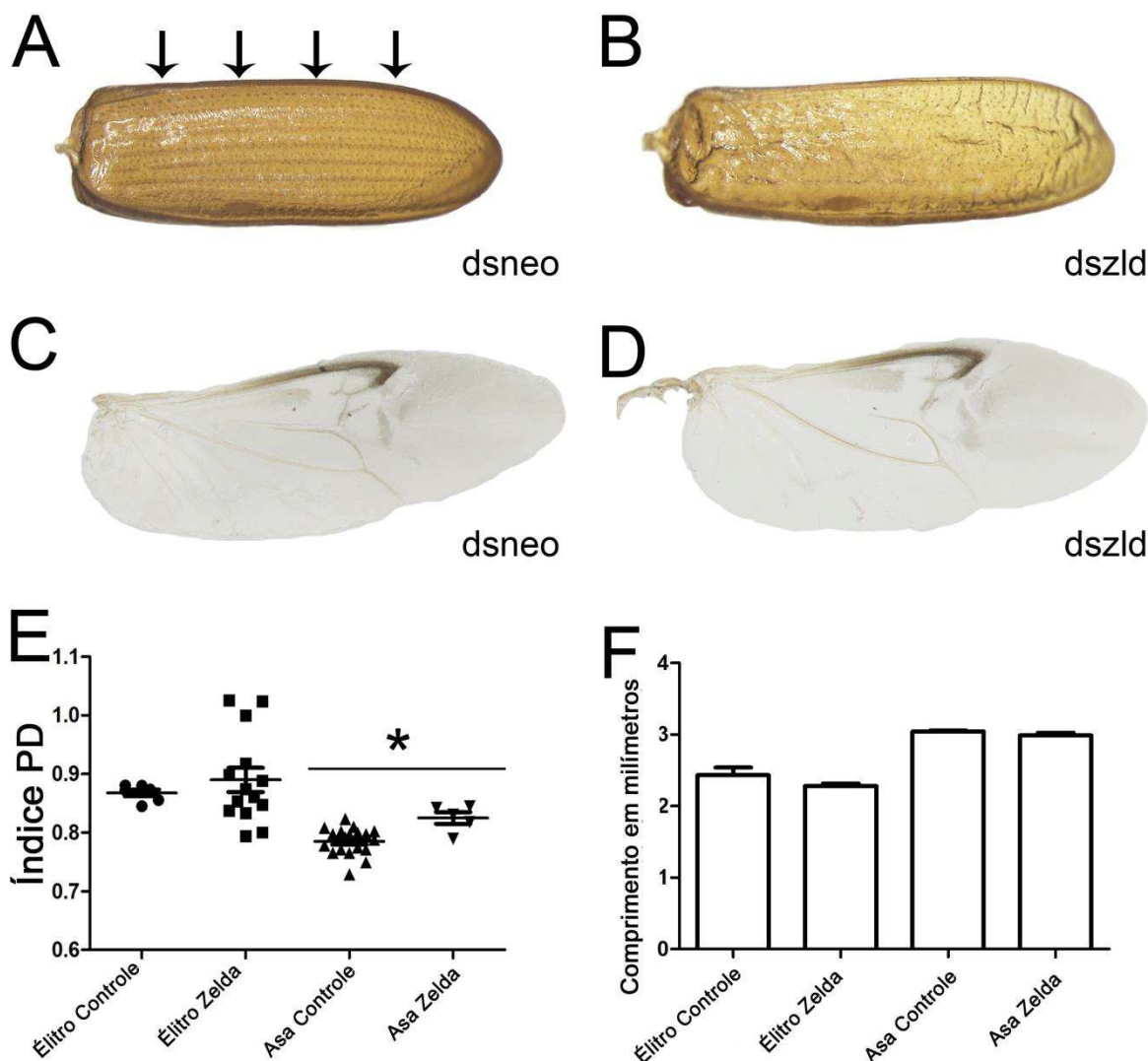


**Figura 23: Élitros e asas de besouros injetados com dsRNA de *Tc-zld* apresentam menor resistência a fixação:** Os apêndices dorsais de *T. castaneum* provenientes de larvas injetadas com *dsneo* e *dszld* foram dissecados e fixados em etanol 95% por 24 h. (A,B) O élitro dissecado de besouros *dszld* (B) apresenta rugas (seta) e menor resistência física que a estrutura do besouro controle (A). (C) A asa membranosas obtida do besouro *dsneo* apresenta um padrão de venação e pigmentação característica. (D) Asas destacadas de besouros *dszld* e submetidas à fixação em etanol apresentam maior fragilidade e ressecamento drástico, assim como maior melanização na região distal (seta). Escala: 1 mm.

Devido a influência do processo de fixação causar dano a constituição de asas e élitros derivados de embriões silenciados, analisou-se o material recém-extraído e mantido em PBST a fim de observar as características morfológicas originais dessas estruturas. Em seguida, realizamos a análise morfológica para investigar se os defeitos relacionados ao silenciamento de *Tc-zld* ocorreriam *in natura* nos élitros e asas de *T. castaneum* (Figura 24). Observa-se que os élitros controle possuem um padrão de venação formado por veias paralelas (setas na Figura 24 A) e que esse padrão é interrompido em élitros *dszld* (Figura 24 B). Já as asas não apresentaram diferenças morfológicas visíveis e nem padrão de venação anormais, quando comparamos os dois tratamentos (Figura 24 C,D).

Buscando observar mais diferenças entre as estruturas após o silenciamento de *Tc-zld* utilizamos uma metodologia descrita recentemente, o índice próximo-distal (PD) (Ravisankar *et al.*, 2016). Enquanto os valores do índice PD para élitros silenciados para *Tc-zld* e controle são similares, um pequeno aumento no índice PD de asas foi observado (Figura 24 E). Já entre as medidas do comprimento de asas e élitros não foram observadas diferenças significativas causadas pelo silenciamento de *Tc-zld*. Os dados indicam que *Tc-zld* é necessário para a padronização correta da venação dos élitros, mas não para o estabelecimento de sua forma geral.

Análises de *zld* realizadas previamente em *D. melanogaster* demonstraram sua expressão em discos imaginais de asas, particularmente onde as células com atividade mitótica estão localizadas (Staudt *et al.*, 2006). Além disso, a expressão ectópica de *zld* nos discos imaginais de *D. melanogaster* ocasiona má formação nas asas e perda do tecido (Staudt *et al.*, 2006; Giannios e Tsitilou, 2013).



**Figura 24: RNAi de *Tc-zld* nos estágios larvares afeta élitros e asas no estágio adulto:** (A) Élitros extraídos do besouro controle *dsneo* apresentam o padrão de venação paralelo e uma estrutura rígida. (B) O élitro do besouro silenciado *dszld* apresenta uma perturbação no padrão de venação e menor rigidez. (C,D) Asas dissecadas não apresentaram diferenças observáveis em sua morfologia. (E) Índices PD de comparação entre élitros e asas *dsneo* e *dszld*. O índice PD exprime a forma da asa baseada em razões de suas dimensões (Ravisankar *et al.*, 2016), indicando que o silenciamento de *Tc-zld* afeta o formato das asas membranosas. (F) O comprimento das asas e élitros de *dszld* não apresenta diferenças significativamente estatísticas quando comparado às respectivas estruturas em *dsneo*. As análises estatísticas foram realizadas utilizando teste-t sobre a variância (\* =  $P < 0,0001$ )

Com relação aos besouros que apresentavam o fenótipo de “carapaça aberta” (Figura 22 D), a análise morfológica mais detalhada revelou que esses mesmos indivíduos também exibiam fenótipos distintos nas antenas e patas. Antenas e patas tem seu desenvolvimento governado por redes regulatórias gênicas similares e por isso são conhecidos como homólogos seriais. O gene *distalless* (*dll*), um alvo predito de *Tc-zld*, é essencial para a segmentação de apêndices em insetos (Panganiban *et al.*, 1994). *Tc-dll* também é expresso durante o final da

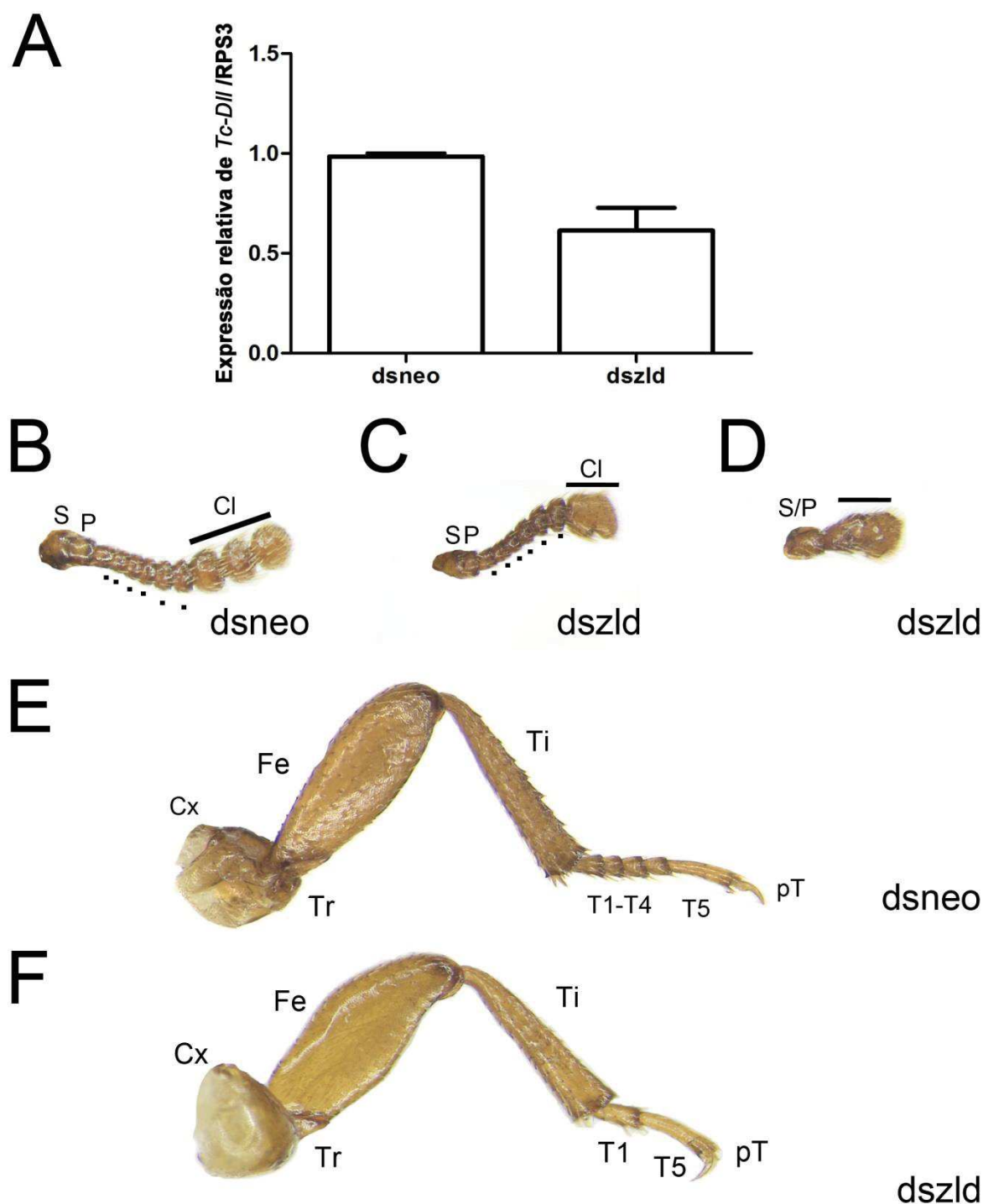
embriogênese na região distal da pata e, como seu próprio nome sugere, a ausência de sua função leva a perda da segmentação das regiões distais de pata e antenas (Beermann *et al.*, 2001). Para confirmar a influência de *Tc-zld* na expressão de *Tc-dll*, investigamos a sua expressão por qRT-PCR. Os níveis de mRNA de *Tc-dll* nas larvas injetadas com *dszld* apresentaram uma queda em comparação com o controle injetado com *dsneo* (Figura 25A), indicando que, durante esses processos, o gene *Tc-dll* tem sua expressão regulada por *Tc-zld* e sugerindo também uma via de sinalização que pode estar envolvida no surgimento desse fenótipo.

O silenciamento de *Tc-zld* resultou em uma variedade de fenótipos afetando as antenas (Figura 25 B-D). Besouros injetados com *dsneo* em sua fase larvar apresentaram antenas com características de indivíduos do tipo selvagem, com os onze segmentos íntegros e dispostos em: escapo (S), pedicelo (P), seis segmentos intermediários do flagelo e os três segmentos distais modificados, formando a clava (Cl) (Figura 25 B). Em fenótipos mais fortes, *dszld* resultou em má formação dos segmentos distais do flagelo, causando o aspecto de clava fusionada em um único segmento (Figura 25 C). Por outro lado, fenótipos fortes de *Tc-zld* resultaram na fusão do escapo e pedicelo, na perda de segmentos e na má formação do flagelo, gerando um único segmento rudimentar (Figura 25 D).

As patas de *T. castaneum* são originadas durante a embriogênese tardia e podem ser reconhecidas como pequenos brotamentos da parede corporal, os brotos de membros. No estágio adulto, elas se apresentam em três pares de patas segmentadas constituídas por seis segmentos: coxa, fêmur, trocanter, tibia, tarso e pós-tarso (Angelini e Kaufman, 2005). Em *T. castaneum*, o tarso é subdividido em segmentos menores denominados tarsômeros, cinco deles nas patas protorácicas e mesotorácicas e quatro deles nas patas metatorácicas. A segmentação do tarso ocorre durante a metamorfose do besouro e acredita-se que sua subdivisão ocorreu no ancestral comum dos insetos, visto que o tarso não se encontra subdividido em outros hexápodes além dos insetos (Boxshall, 2004). Após o silenciamento de *Tc-zld* os segmentos do tarso estiveram ausentes ou foram observados fusionados, resultando no encurtamento das patas, enquanto em besouros *dsneo* as patas têm características idênticas às de animais do tipo selvagem (Figura 25 EF). Estes resultados indicam que *Tc-zld* está envolvido no desenvolvimento de segmentos e na formação de articulações. Sabe-se que um grande domínio de expressão do



gene *Tc-Dll*, alvo de *Tc-zld*, pode ser encontrado no tarso do besouro (Angelini *et al.*, 2012) e o silenciamento de *Tc-Dll* em besouros também gera patas com deleção de tarsômeros (Suzuki *et al.*, 2009; Angelini *et al.*, 2012). Portanto, o silenciamento de *Tc-zld* gera fenótipos de perda de segmentos tanto nas antenas quanto nas patas do besouro, tratando-se de um fenótipo pleiotrópico relacionado à homologia serial destes apêndices (Angelini *et al.*, 2012).



**Figura 25:** O *knockdown* (silenciamento) de *Tc-zld* induz fenótipos alterados nas antenas e patas de adultos de *T. castaneum*: (A) Expressão relativa do gene *Tc-Dll* após

injeção de *dsneo* e *dszld*. (B-D) Antenas de indivíduos controle apresentam o número normal de segmentos: dois proximais (escapo e pedicelo), seis intermediários (flagelo) e três segmentos distais diferenciados formando a clava. (E,F) A injeção de *dszld* causou a fusão dos segmentos da clava enquanto em fenótipos mais “fortes “ a antena não se segmenta, desenvolvendo-se como um rudimento de antena. As patas de *T. castaneum* possuem seis podômeros: Coxa (Cx), trocânter (Tr), fêmur (Fe), Tibia (Ti), Tarso (T) e pós-tarso (ou pré-tarso (pT)). (D) As patas mesotorácicas de indivíduos injetados com *dsneo* apresentam cinco segmentos tarsais, ou tarsômeros. (E) Em indivíduos silenciados para *Tc-zld*, alguns segmentos do tarso são perdidos ou fundidos, resultando em redução do tarso, enquanto outras estruturas da pata aparentemente não são afetadas.

## 6. Considerações finais

Nossos resultados evidenciam que além de necessário para a ativação zigótica e segmentação do embrião, o produto do gene *Tc-zld* também participa de eventos no desenvolvimento pós-embrionário. Evidenciamos que a expressão de *Tc-Zld* é importante na segmentação das patas e antenas (homólogos seriais), onde desempenha um papel semelhante ao que exerce na segmentação que ocorre no estágio embrionário, e também no desenvolvimento de asas e élitros. Conforme as funções de *Tc-zld* se revelam, pode-se sugerir que sua ação esteja relacionada com populações de células precursoras nas quais essa proteína provavelmente adquire a função de TF “pioneiro”, facilitando a ligação de outros TFs nas regiões vizinhas do DNA. Dessa forma, *Tc-zld* pode representar um primeiro passo para o funcionamento das redes regulatórias e ao mesmo tempo participar na regulação de genes a jusante da cadeia de sinalização.

Recentemente, foi descrito que a competição entre TFs e histonas parece regular o início da transcrição zigótica em embriões de *D. rerio* (Joseph *et al.*, 2017), onde se verifica que TFs como *Pou5f3* e *Sox19b* competem com histonas pela ligação no DNA. Os TFs ditos pioneiros podem interagir com DNA ligado a histona e, sob essa nova perspectiva, seria de interesse investigar se *Zelda* exerceria essa função de competição em insetos.

O transcriptoma embrionário comparativo revelou uma grande quantidade de genes com expressão diferencial após o silenciamento de *Tc-zld*. Foram identificados alguns genes de função já conhecida como TFs e proteínas de sinalização pertencentes a vias de padronização de eixos e segmentação, além de genes *Homeobox*. Os dados obtidos por essa análise transcricional possuem ainda um grande potencial a ser explorado, com a investigação funcional também de

genes únicos em *T. castaneum* e de diversos genes de função desconhecida, que podem ajudar a desvendar a rede de interações reguladas por Zelda.

## 7. CONCLUSÕES

1. Zelda está presente no clado Pancrustacea e possui função de ativador zigótico conservada nos insetos estudados;
2. Zelda possui domínios protéicos conservados que não haviam sido previamente identificados;
3. Os eventos de ativação zigótica e segmentação da banda germinal são dependentes de Zelda;
4. O silenciamento de Zelda resulta na modulação de um grande número de genes, dentre eles vários TFs envolvidos em etapas cruciais do desenvolvimento;
5. Zelda é atuante em eventos pós-embrionários como metamorfose e desenvolvimento de apêndices nos adultos.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

- AKAM, M. (1994). Insect development. Is pairing the rule? Nature, v. 367, n. 6462, p. 410-411.
- ALTINCICEK, B. et al. (2013). Next generation sequencing based transcriptome analysis of septic-injury responsive genes in the beetle *Tribolium castaneum*. PLoS One, v. 8, n. 1, p. e52004.
- ANGELINI, D. R.; KAUFMAN, T. C. (2005). Insect appendages and comparative ontogenetics. Developmental Biology, v. 286, n. 1, p. 57-77.
- ANGELINI, D. R.; KIKUCHI, M.; JOCKUSCH, E. L. (2009). Genetic patterning in the adult capitata antenna of the beetle *Tribolium castaneum*. Developmental Biology, v. 327, n. 1, p. 240-251.
- ANGELINI, D. R.; SMITH, F. W.; JOCKUSCH, E. L. (2012). Extent With Modification: Leg Patterning in the Beetle *Tribolium castaneum* and the Evolution of Serial Homologs. G3 (Bethesda, Md.), v. 2, n. 2, p. 235-248.
- ARAVIND, L.; IYER, L. M. (2012). The HARE-HTH and associated domains: novel modules in the coordination of epigenetic DNA and protein modifications. Cell Cycle, v. 11, n. 1, p. 119-131.
- ARSALA, D.; LYNCH, J. A. (2017). Ploidy has little effect on timing early embryonic events in the haplo-diploid wasp *Nasonia*. genesis, p. n/a-n/a.
- BAILEY, T. L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics, v. 27, n. 12, p. 1653-1659.
- BALAJI, S.; IYER, L. M.; ARAVIND, L. (2009). HPC2 and ubinuclein define a novel family of histone chaperones conserved throughout eukaryotes. Molecular Biosystems, v. 5, n. 3, p. 269-275.
- BEERMANN, A. et al. (2001). The Short antennae gene of *Tribolium* is required for limb development and encodes the orthologue of the *Drosophila* Distal-less protein. Development, v. 128, n. 2, p. 287-297.
- BEHRENS, S. et al. (2014). Infection routes matter in population-specific responses of the red flour beetle to the entomopathogen *Bacillus thuringiensis*. BMC Genomics, v. 15, p. 445.
- BENTON, M. A.; AKAM, M.; PAVLOPOULOS, A. (2013). Cell and tissue dynamics during *Tribolium* embryogenesis revealed by versatile fluorescence labeling approaches. Development, v. 140, n. 15, p. 3210-3220.
- BENTON, M. A. et al. (2016). Toll genes have an ancestral role in axis elongation. Current Biology, v. 26, n. 12, p. 1609-1615.

- BOLGER, A. M.; LOHSE, M.; USADEL, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, v. 30, n. 15, p. 2114-20.
- BOXSHALL, G. A. (2004). The evolution of arthropod limbs. Biological Reviews of the Cambridge Philosophical Society, v. 79, n. 2, p. 253-300.
- BROWN, S. J. et al. (1999). Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. Evolution & Development, v. 1, n. 1, p. 11-15.
- BROWN, S. J. et al. (2009). The red flour beetle, *Tribolium castaneum* (Coleoptera): a model for studies of development and pest biology. Cold Spring Harb Protoc, v. 2009, n. 8, p. pdb emo126.
- BUCHER, G.; SCHOLTEN, J.; KLINGLER, M. (2002). Parental rnaï in tribolium (coleoptera). Current Biology, v. 12, n. 3, p. R85-6.
- CHEN, G.; HANDEL, K.; ROTH, S. (2000). The maternal NF-kappaB/dorsal gradient of *Tribolium castaneum*: dynamics of early dorsoventral patterning in a short-germ beetle. Development, v. 127, n. 23, p. 5145-5156.
- CHOE, C. P.; MILLER, S. C.; BROWN, S. J. (2006). A pair-rule gene circuit defines segments sequentially in the short-germ insect *Tribolium castaneum*. Proceedings of the National Academy of Sciences of the United States of America, v. 103, n. 17, p. 6560-6564.
- CHOMCZYNSKI, P. (1993). A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. Biotechniques, v. 15, n. 3, p. 532-4, 536-7.
- COHEN, B.; SIMCOX, A. A.; COHEN, S. M. (1993). Allocation of the thoracic imaginal primordia in the *Drosophila* embryo. Development, v. 117, n. 2, p. 597-608.
- DAVIS, G. K.; JARAMILLO, C. A.; PATEL, N. H. (2001). Pax group III genes and the evolution of insect pair-rule patterning. Development, v. 128, n. 18, p. 3445-58.
- DAVIS, G. K.; PATEL, N. H. (2002). Short, long, and beyond: molecular and embryological approaches to insect segmentation. Annual Review of Entomology, v. 47, p. 669-699.
- DE RENZIS, S. et al. (2007). Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo. PLoS Biology, v. 5, n. 5, p. e117.
- DROUIN, J. (2014). Minireview: pioneer transcription factors in cell fate specification. Molecular Endocrinology, v. 28, n. 7, p. 989-998.
- EDGAR, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics, v. 5, p. 113.

- EL-SHERIF, E.; AVEROF, M.; BROWN, S. J. (2012). A segmentation clock operating in blastoderm and germband stages of *Tribolium* development. Development, v. 139, n. 23, p. 4341-4346.
- EL-SHERIF, E. et al. (2014). Caudal regulates the spatiotemporal dynamics of pair-rule waves in *Tribolium*. PLoS Genet, v. 10, n. 10, p. e1004677.
- ENGELHARDT, M.; RAPOPORT, H.; SOKOLOFF, A. (1965). Odorous secretion of normal and mutant *Tribolium confusum*. Science, v. 150, n. 3696, p. 632-3.
- ENGLERT, D. C.; BELL, A. E. (1970). Selection for Time of Pupation in *TRIBOLIUM CASTANEUM*. Genetics, v. 64, n. 3-4, p. 541-52.
- FINN, R. D. et al. (2016). The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research, v. 44, n. D1, p. D279-85.
- FONSECA, R. N.; LYNCH, J. A.; ROTH, S. (2009). Evolution of axis formation: mRNA localization, regulatory circuits and posterior specification in non-model arthropods. Curr Opin Genet Dev, v. 19, n. 4, p. 404-11.
- FOO, S. M. et al. (2014). Zelda potentiates morphogen activity by increasing chromatin accessibility. Current Biology, v. 24, n. 12, p. 1341-1346.
- FRANTSEVICH, L. et al. (2005). Geometry of elytra opening and closing in some beetles (Coleoptera, Polyphaga). The Journal of Experimental Biology, v. 208, n. Pt 16, p. 3145-3158.
- GIANNIOS, P.; TSITILLOU, S. G. (2013). The embryonic transcription factor Zelda of *Drosophila melanogaster* is also expressed in larvae and may regulate developmentally important genes. Biochemical and Biophysical Research Communications, v. 438, n. 2, p. 329-333.
- GROSSMAN, S. R. et al. (2017). Systematic dissection of genomic features determining transcription factor binding and enhancer function. Proceedings of the National Academy of Sciences of the United States of America, v. 114, n. 7, p. E1291-E1300.
- HACKETT, K. J. et al. (1992). Lampyridae (Coleoptera): A plethora of mollicute associations. Microb Ecol, v. 23, n. 2, p. 181-93.
- HAFERNIK, J.; SAUL-GERSHENZ, L. (2000). Beetle larvae cooperate to mimic bees. Nature, v. 405, n. 6782, p. 35-6.
- HAMM, D. C.; BONDRA, E. R.; HARRISON, M. M. (2015). Transcriptional activation is a conserved feature of the early embryonic factor Zelda that requires a cluster of four zinc fingers for DNA binding and a low-complexity activation domain. The Journal of Biological Chemistry, v. 290, n. 6, p. 3508-3518.

- HANDEL, K. et al. (2005). *Tribolium castaneum* twist: gastrulation and mesoderm formation in a short-germ beetle. Development Genes and Evolution, v. 215, n. 1, p. 13-31.
- HARRISON, M. M.; BOTCHAN, M. R.; CLINE, T. W. (2010). Grainyhead and Zelda compete for binding to the promoters of the earliest-expressed *Drosophila* genes. Developmental Biology, v. 345, n. 2, p. 248-255.
- HARRISON, M. M. et al. (2011). Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. PLoS Genetics, v. 7, n. 10, p. e1002266.
- HUANG, D. W.; SHERMAN, B. T.; LEMPICKI, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols, v. 4, n. 1, p. 44-57.
- JOSEPH, S. R. et al. (2017). Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos. Elife, v. 6,
- KIDD, T.; BLAND, K. S.; GOODMAN, C. S. (1999). Slit is the midline repellent for the robo receptor in *Drosophila*. Cell, v. 96, n. 6, p. 785-94.
- KIM, D. et al. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol, v. 14, n. 4, p. R36.
- LAVORE, A. et al. (2012). The gap gene giant of *Rhodnius prolixus* is maternally expressed and required for proper head and abdomen formation. Dev Biol, v. 361, n. 1, p. 147-55.
- LEICHSENDRING, M. et al. (2013). Pou5f1 transcription factor controls zygotic gene activation in vertebrates. Science, v. 341, n. 6149, p. 1005-1009.
- LEVENE, H. et al. (1965). Genetic load in *tribolium*. Proc Natl Acad Sci U S A, v. 53, n. 5, p. 1042-50.
- LI, X.-Y. et al. (2014). Establishment of regions of genomic activity during the *Drosophila* maternal to zygotic transition. eLife, v. 3,
- LIANG, H.-L. et al. (2008). The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. Nature, v. 456, n. 7220, p. 400-403.
- LINZ, D. M. et al. (2016). Functional value of elytra under various stresses in the red flour beetle, *Tribolium castaneum*. Sci Rep, v. 6, p. 34813.
- LIVAK, K. J.; SCHMITTGEN, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods, v. 25, n. 4, p. 402-8.



- LOMAKIN, J. et al. (2010). Mechanical properties of elytra from *Tribolium castaneum* wild-type and body color mutant strains. J Insect Physiol, v. 56, n. 12, p. 1901-6.
- LORCH, Y.; KORNBERG, R. D. (2015). Chromatin-remodeling and the initiation of transcription. Quarterly Reviews of Biophysics, v. 48, n. 4, p. 465-470.
- LORD, J. C. et al. (2010). Evaluation of quantitative PCR reference genes for gene expression studies in *Tribolium castaneum* after fungal challenge. Journal of Microbiological Methods, v. 80, n. 2, p. 219-221.
- MARCHLER-BAUER, A. et al. (2017). CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Research, v. 45, n. D1, p. D200-D203.
- MESQUITA, R. D. et al. (2015). Genome of *Rhodnius prolixus*, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. Proc Natl Acad Sci U S A, v. 112, n. 48, p. 14936-41.
- MISOF, B. et al. (2014). Phylogenomics resolves the timing and pattern of insect evolution. Science, v. 346, n. 6210, p. 763-767.
- NAKAMOTO, A. et al. (2015). Changing cell behaviours during beetle embryogenesis correlates with slowing of segmentation. Nature Communications, v. 6, p. 6635.
- NESTOROV, P. et al. (2013). The maternal transcriptome of the crustacean *Parhyale hawaiiensis* is inherited asymmetrically to invariant cell lineages of the ectoderm and mesoderm. Plos One, v. 8, n. 2, p. e56049.
- NIEN, C.-Y. et al. (2011). Temporal coordination of gene networks by Zelda in the early *Drosophila* embryo. PLoS Genetics, v. 7, n. 10, p. e1002339.
- NINOVA, M.; RONSHAUGEN, M.; GRIFFITHS-JONES, S. (2016). MicroRNA evolution, expression, and function during short germband development in *Tribolium castaneum*. Genome Res, v. 26, n. 1, p. 85-96.
- NITTA, K. R. et al. (2015). Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. Elife, v. 4,
- NUNES DA FONSECA, R. et al. (2008). Self-regulatory circuits in dorsoventral axis formation of the short-germ beetle *Tribolium castaneum*. Developmental Cell, v. 14, n. 4, p. 605-615.
- NÜSSLEIN-VOLHARD, C.; FROHNHÖFER, H. G.; LEHMANN, R. (1987). Determination of anteroposterior polarity in *Drosophila*. Science, v. 238, n. 4834, p. 1675-1681.
- NUSSLEIN-VOLHARD, C.; WIESCHAUS, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. Nature, v. 287, n. 5785, p. 795-801.

- OBBER, K. A.; JOCKUSCH, E. L. (2006). The roles of wingless and decapentaplegic in axis and appendage development in the red flour beetle, *Tribolium castaneum*. Dev Biol, v. 294, n. 2, p. 391-405.
- ORDAN, E.; VOLK, T. (2015). Cleaved Slit directs embryonic muscles. Fly (Austin), v. 9, n. 2, p. 82-5.
- PANGANIBAN, G.; NAGY, L.; CARROLL, S. B. (1994). The role of the Distal-less gene in the development and evolution of insect limbs. Current Biology, v. 4, n. 8, p. 671-675.
- PARÉ, A. C. et al. (2014). A positional Toll receptor code directs convergent extension in *Drosophila*. Nature, v. 515, n. 7528, p. 523-527.
- PARK, T. (1934). Observations on the General Biology of the Flour Beetle, *Tribolium Confusum*. The Quarterly Review of Biology, v. 9, n. 1, p. 36-54.
- PATEL, N. H. et al. (1989). Expression of engrailed proteins in arthropods, annelids, and chordates. Cell, v. 58, n. 5, p. 955-968.
- PEARSON, J. C.; WATSON, J. D.; CREWS, S. T. (2012). *Drosophila melanogaster* Zelda and Single-minded collaborate to regulate an evolutionarily dynamic CNS midline cell enhancer. Developmental Biology, v. 366, n. 2, p. 420-432.
- PEREIRA, F. et al. (2016). Evolution of the NET (NocA, Nlz, Elbow, TLP-1) protein family in metazoans: insights from expression data and phylogenetic analysis. Scientific reports, v. 6, p. 38383.
- PIRES, C. V. et al. (2016). Transcriptome Analysis of Honeybee (*Apis Mellifera*) Haploid and Diploid Embryos Reveals Early Zygotic Transcription during Cleavage. Plos One, v. 11, n. 1, p. e0146447.
- PORCHER, A. et al. (2010). The time to measure positional information: maternal hunchback is required for the synchrony of the Bicoid transcriptional response at the onset of zygotic transcription. Development, v. 137, n. 16, p. 2795-2804.
- PREUSS, K. M. et al. (2012). Identification of maternally-loaded RNA transcripts in unfertilized eggs of *Tribolium castaneum*. BMC Genomics, v. 13, p. 671.
- RAVISANKAR, P. et al. (2016). Comparative developmental analysis of *Drosophila* and *Tribolium* reveals conserved and diverged roles of abrupt in insect wing evolution. Developmental Biology, v. 409, n. 2, p. 518-529.
- ROBERTS, A. et al. (2011). Identification of novel transcripts in annotated genomes using RNA-Seq. Bioinformatics, v. 27, n. 17, p. 2325-9.
- ROZEN, S.; SKALETSKY, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology, v. 132, p. 365-386.

- SARRAZIN, A. F.; PEEL, A. D.; AVEROF, M. (2012). A segmentation clock with two-segment periodicity in insects. Science, v. 336, n. 6079, p. 338-341.
- SCHMIDT-OTT, U. (2005). Insect serosa: a head line in comparative developmental genetics. Current Biology, v. 15, n. 7, p. R245-7.
- SOKOLOFF, A. **The biology of Tribolium: with special emphasis on genetic aspects**. Oxford,: Clarendon Press, 1972. ISBN 0198573537 (v. 1).
- STAPPERT, D. et al. (2016). Genome-wide identification of Tribolium dorsoventral patterning genes. Development, v. 143, n. 13, p. 2443-54.
- STAUDT, N. et al. (2006). Mutations of the Drosophila zinc finger-encoding gene vielfältig impair mitotic cell divisions and cause improper chromosome segregation. Molecular Biology of the Cell, v. 17, n. 5, p. 2356-2365.
- STROBL, F.; SCHMITZ, A.; STELZER, E. H. (2015). Live imaging of Tribolium castaneum embryonic development using light-sheet-based fluorescence microscopy. Nat Protoc, v. 10, n. 10, p. 1486-507.
- SUZUKI, Y.; SQUIRES, D. C.; RIDDIFORD, L. M. (2009). Larval leg integrity is maintained by Distal-less and is required for proper timing of metamorphosis in the flour beetle, Tribolium castaneum. Developmental Biology, v. 326, n. 1, p. 60-67.
- SVÁCHA, P. (1992). What are and what are not imaginal discs: reevaluation of some basic concepts (Insecta, Holometabola). Developmental Biology, v. 154, n. 1, p. 101-117.
- TADROS, W.; LIPSHITZ, H. D. (2009). The maternal-to-zygotic transition: a play in two acts. Development, v. 136, n. 18, p. 3033-3042.
- TAUTZ, D.; PFEIFLE, C. (1989). A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback. Chromosoma, v. 98, n. 2, p. 81-5.
- TEN BOSCH, J. R.; BENAVIDES, J. A.; CLINE, T. W. (2006). The TAGteam DNA motif controls the timing of Drosophila pre-blastoderm transcription. Development, v. 133, n. 10, p. 1967-1977.
- TOMOYASU, Y. et al. (2009). Repeated co-options of exoskeleton formation during wing-to-elytron evolution in beetles. Current Biology, v. 19, n. 24, p. 2057-2065.
- TOMOYASU, Y.; DENELL, R. E. (2004). Larval RNAi in Tribolium (Coleoptera) for analyzing adult development. Development Genes and Evolution, v. 214, n. 11, p. 575-578.

- TOMOYASU, Y.; WHEELER, S. R.; DENELL, R. E. (2005). Ultrabithorax is required for membranous wing identity in the beetle *Tribolium castaneum*. Nature, v. 433, n. 7026, p. 643-647.
- TRAPNELL, C. et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc, v. 7, n. 3, p. 562-78.
- TRAPNELL, C. et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol, v. 28, n. 5, p. 511-5.
- TRIBOLIUM GENOME SEQUENCING, C. et al. (2008). The genome of the model beetle and pest *Tribolium castaneum*. Nature, v. 452, n. 7190, p. 949-955.
- TRUMAN, J. W.; RIDDIFORD, L. M. (1999). The origins of insect metamorphosis. Nature, v. 401, n. 6752, p. 447-52.
- VAN DER ZEE, M.; BERNS, N.; ROTH, S. (2005). Distinct functions of the *Tribolium* *zerknüllt* genes in serosa specification and dorsal closure. Current Biology, v. 15, n. 7, p. 624-636.
- VASILJEVIC, M. et al. (2012). Spatio-temporal expression analysis of the calcium-binding protein calumenin in the rodent brain. Neuroscience, v. 202, p. 29-41.
- VIDAL, N. M. et al. (2016). Transcription factors, chromatin proteins and the diversification of Hemiptera. Insect Biochemistry and Molecular Biology, v. 69, p. 1-13.
- WALSER, C. B.; LIPSHITZ, H. D. (2011). Transcript clearance during the maternal-to-zygotic transition. Current Opinion in Genetics & Development, v. 21, n. 4, p. 431-443.
- WANG, L. et al. (2007). BeetleBase: the model organism database for *Tribolium castaneum*. Nucleic Acids Research, v. 35, n. Database issue, p. D476-9.
- WATERHOUSE, A. M. et al. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics, v. 25, n. 9, p. 1189-1191.
- XU, Z. et al. (2014). Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in *Drosophila*. Genes & Development, v. 28, n. 6, p. 608-621.

## 9. ANEXOS

### **Solução de hibridização**

Formamida 15mL

SSC 20× 7.5mL

DNA de esperma de salmão (10 mg/mL) 300 µL

Heparina (5 mg/mL) 30 µL

Tween 20 30 µL

Ajustar com H<sub>2</sub>O para 30 mL

### **SSC 20x (pH 5,5)**

NaCl 175,3 g

Citrato de Sódio 2H<sub>2</sub>O 88,2 g

Ajustar com H<sub>2</sub>O para 1000 mL

### **Tampão fosfatase alcalina**

Tris-HCl 100 mM (pH 9.5)

MgCl<sub>2</sub> 5 mM

NaCl 100 mM

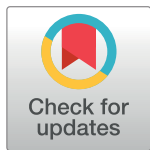
RESEARCH ARTICLE

# Evolution and multiple roles of the Pancrustacea specific transcription factor *zelda* in insects

Lupis Ribeiro<sup>1,2</sup>, Vitória Tobias-Santos<sup>1</sup>, Daniele Santos<sup>1</sup>, Felipe Antunes<sup>1</sup>, Geórgia Feltran<sup>1</sup>, Jackson de Souza Menezes<sup>1</sup>, L. Aravind<sup>3</sup>, Thiago M. Venancio<sup>2\*</sup>, Rodrigo Nunes da Fonseca<sup>1\*</sup>

**1** Laboratório Integrado de Bioquímica Hatisaburo Masuda, Núcleo em Ecologia e Desenvolvimento SócioAmbientaI de Macaé (NUPEM), Campus UFRJ Macaé, Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular - INCT-EM, Macaé, Brazil, **2** Laboratório de Química e Função de Proteínas e Peptídeos, Centro de Biociências e Biotecnologia, Universidade Estadual do Norte Fluminense Darcy Ribeiro, Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular - INCT-EM, Rio de Janeiro, Brazil, **3** National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland, United States of America

\* [thiago.venancio@gmail.com](mailto:thiago.venancio@gmail.com) (TMV); [rodrigo.nunes.da.fonseca@gmail.com](mailto:rodrigo.nunes.da.fonseca@gmail.com) (RNdF)



 OPEN ACCESS

**Citation:** Ribeiro L, Tobias-Santos V, Santos D, Antunes F, Feltran G, de Souza Menezes J, et al. (2017) Evolution and multiple roles of the Pancrustacea specific transcription factor *zelda* in insects. PLoS Genet 13(7): e1006868. <https://doi.org/10.1371/journal.pgen.1006868>

**Editor:** Claude Desplan, New York University, UNITED STATES

**Received:** April 17, 2017

**Accepted:** June 14, 2017

**Published:** July 3, 2017

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by: NIH Intramural Funding to LA; FAPERJ/RJ/BRAZIL to TMV and RNdF [www.faperj.br](http://www.faperj.br); and CNPq/BRAZIL to TMV and RNdF [www.cnpq.br](http://www.cnpq.br). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Gene regulatory networks (GRNs) evolve as a result of the coevolutionary processes acting on transcription factors (TFs) and the cis-regulatory modules they bind. The zinc-finger TF *zelda* (*zld*) is essential for the maternal-to-zygotic transition (MZT) in *Drosophila melanogaster*, where it directly binds over thousand cis-regulatory modules to regulate chromatin accessibility. *D. melanogaster* displays a long germ type of embryonic development, where all segments are simultaneously generated along the whole egg. However, it remains unclear if *zld* is also involved in the MZT of short-germ insects (including those from basal lineages) or in other biological processes. Here we show that *zld* is an innovation of the Pancrustacea lineage, being absent in more distant arthropods (e.g. chelicerates) and other organisms. To better understand *zld*'s ancestral function, we thoroughly investigated its roles in a short-germ beetle, *Tribolium castaneum*, using molecular biology and computational approaches. Our results demonstrate roles for *zld* not only during the MZT, but also in posterior segmentation and patterning of imaginal disc derived structures. Further, we also demonstrate that *zld* is critical for posterior segmentation in the hemipteran *Rhodnius prolixus*, indicating this function predates the origin of holometabolous insects and was subsequently lost in long-germ insects. Our results unveil new roles of *zld* in different biological contexts and suggest that changes in expression of *zld* (and probably other major TFs) are critical in the evolution of insect GRNs.

## Author summary

Pioneer transcription factors (TFs) are considered the first regulators of chromatin accessibility in fruit flies and vertebrates, modulating the expression of a large number of target

**Competing interests:** The authors have declared that no competing interests exist.

genes. In fruit flies, *zelda* resembles a pioneer TF, being essential during early embryogenesis. However, the evolutionary origins and ancestral functions of *zelda* remain largely unknown. Through a number of gene silencing, microscopy and evolutionary analysis, the present work shows that *zelda* is an innovation of the Pancrustacea lineage, governing not only the MZT in the short-germ insect *Tribolium castaneum*, but also posterior segmentation and post-embryonic patterning of imaginal disc derived structures such as wings, legs and antennae. Further, *zelda* regulation of posterior segmentation predates the origin of insects with complete metamorphosis (holometabolous), as supported by gene silencing experiments in the kissing bug *Rhodnius prolixus*. We hypothesize that the emergence of *zelda* contributed to the evolution of gene regulatory networks and new morphological structures of insects.

## Introduction

Gene regulatory networks (GRNs) depend on the coevolution of transcription factors (TFs) and their relationship with the cis-regulatory modules (CRMs) they bind [1, 2]. In insects, the detailed role of a number of TFs and CRMs have been well-described, particularly during the embryogenesis of the fruit fly *Drosophila melanogaster* [3].

In metazoans, the period following fertilization is typically characterized by rapid and near-synchronous mitotic divisions and cleavages that occur under conditions of minimal cellular differentiation. Cleavages typically depend on maternally supplied factors and zygotic genome transcription is constrained during this early period of development [4]. A conserved process of metazoan embryogenesis is the maternal-to-zygotic transition (MZT), which is characterized by two critical steps: 1) the elimination of a maternal set of mRNAs and proteins and; 2) the beginning of zygotic transcription, which leads to the zygotic genomic control of development [5]. In *D. melanogaster*, the majority of the first set of zygotic transcripts are regulated by *zelda* (*zld*) [6], a zinc finger TF with particular affinity for promoter regions containing TAG-team sites—heptamers constituted by CAGGTAG and its variants [7–9]. *zld* (*Dm-zld*, in *D. melanogaster*) binding sites have been identified in *D. melanogaster* embryos (cycles 8 to 14) by chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-Seq) [8]. *Dm-zld* regulates a large set of genes involved in important processes such as cytoskeleton organization, cellularization, germ band development, pattern formation, sex determination and miRNA biogenesis [6]. *Dm-zld* was also suggested to participate in larval wing disc development [10] and its overexpression during wing imaginal disc formation led to wing blisters in adults, an indicative of improper adherence of ventral and dorsal wing epithelia [11]. Nevertheless, while *zld*'s functions have been thoroughly investigated in *D. melanogaster* MZT, its roles in other organisms and biological processes remain elusive.

*D. melanogaster* displays a long germ type of embryonic development, during which all segments are simultaneously generated along the whole egg. In contrast, short germ insects generate anterior (e.g. head) segments early in development, while the remaining segments are patterned from the posterior region, the growth zone (GZ). Since short germ development is considered to be the ancestral mode of development, short-germ insects have been established as developmental model systems [12, 13]. The short-germ red flour beetle *Tribolium castaneum* (*Tc*) was the first beetle species to have its genome completely sequenced [14]. *T. castaneum* displays a short life-cycle and is amenable to gene silencing via RNAi [15], gene overexpression [16], specific tissue expression [17] and fluorescence labeling during early development [18]. Several developmental processes that have been investigated in *T.*

*castaneum* were lost or extensively modified in the *D. melanogaster* lineage, such as GZ formation [19], extensive extra-embryonic morphogenesis [20] and the formation of a morphologically complex head during embryogenesis [21]. Early development of *T. castaneum* is similar to most other insect groups, in which synchronous rounds of division are followed by nuclear migration to the egg cortex and cellularization, forming the so called uniform blastoderm [18, 22]. Taken together, all the genetic and morphological information on *T. castaneum* early development, along with the established techniques mentioned above, make this beetle an ideal model to understand the evolution of *zld*'s function during insect development.

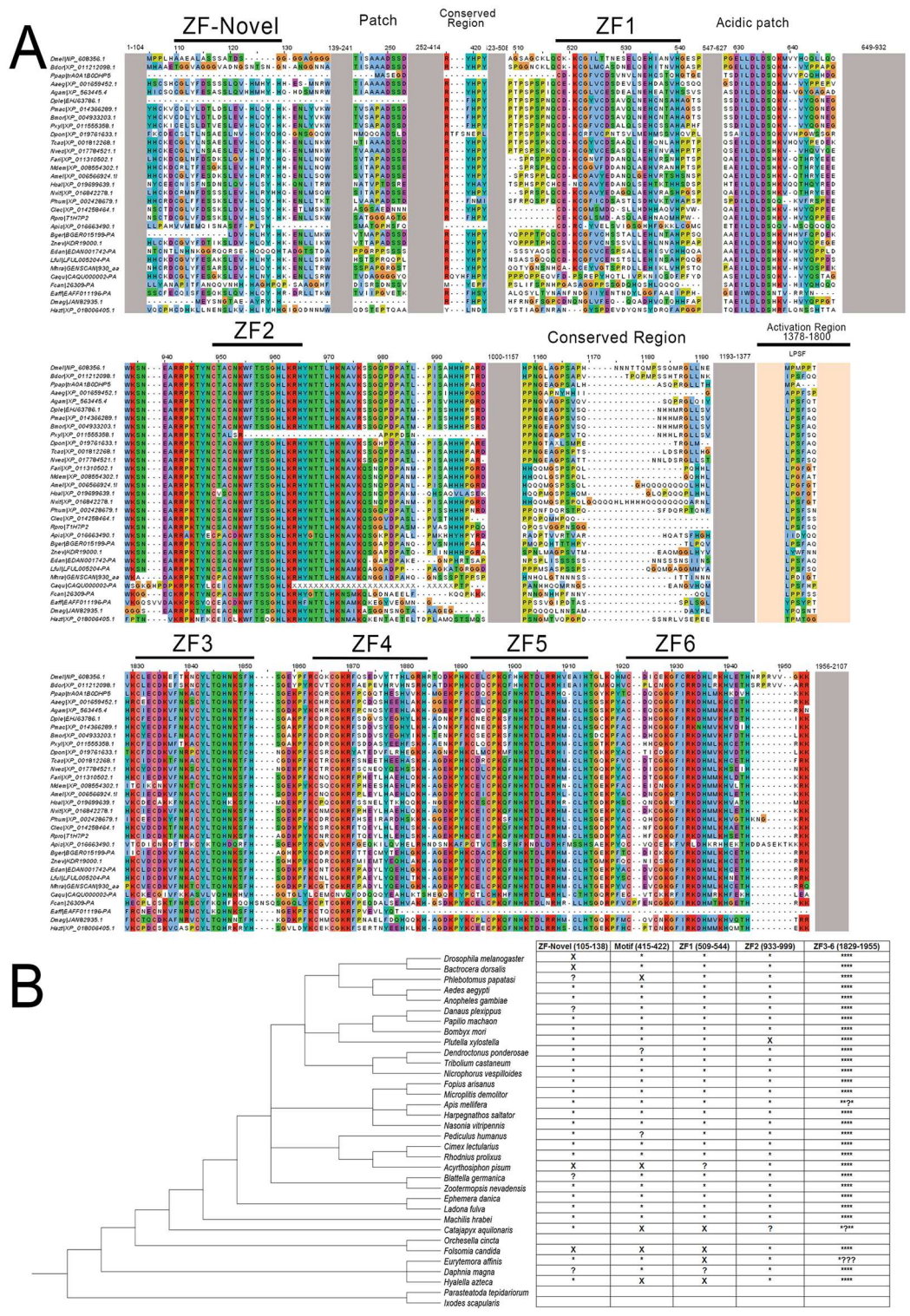
In the present work, we provide the first comprehensive analysis of *zld* orthologs across a wide range of species. Further, we provide functional analysis of a *zld* ortholog in a non-Drosophilid insect, the short germ beetle *T. castaneum* (thenceforth referred as *Tc-zld*). Among our main results are: 1) The identification of some previously overlooked conserved domains in Zld; 2) An inference of the evolutionary origin of Zld, based on phyletic analysis in various hexapods and crustaceans; 3) The identification of a conserved set of 141 putative *Tc-zld* targets (i.e. genes with upstream TAGteam sequences), enriched in TFs, whose homologs that have been demonstrated to be *zld* targets in *D. melanogaster*; 4) Identification of key roles played by *Tc-zld* during the MZT; 5) Identification and experimental validation of two new biological roles of *zld* in *T. castaneum*: segment generation from the posterior GZ during embryogenesis and postembryonic imaginal disc development; 6) Demonstration that *zld* is also involved in GZ patterning in the hemipteran *Rhodnius prolixus*, supporting a conserved role in the GZ of a distant short-germ species. Altogether, our results unveil new roles of *zld* as a pleiotropic TF acting in various developmental processes across distantly-related insects.

## Results/Discussion

### Identification of conserved domains and tracing the origin of *zld*

While previous studies reported that *zld* is involved in the MZT of *D. melanogaster* [6], its evolutionary history remains unclear. We investigated the phyletic distribution of *zld* and found a single ortholog in all the inspected insect genomes, including the beetle *T. castaneum* (Fig 1; S1 Table), indicating that insects are sensitive to increased copy number of this important regulator, which is interesting considering that different TF families are particularly prone to expansions across insect lineages [23]. Further, *zld* homologs were also found in some (but not all) collembolan and crustacean genomes (Fig 1B; S1 Table). The canonical domain architecture of Zld has been reported as comprising a JAZ zinc finger (Pfam: zf-C2H2\_JAZ) domain and a C-terminal cluster of four DNA binding C2H2 zinc finger domains (zf-C2H2) [6, 11, 24]. However, we performed a sensitive and detailed analysis of Zld proteins from multiple species and found other notable conserved protein domains and structural features. Firstly, we found two additional zf-C2H2 domains, N-terminal to the JAZ domain (Fig 1A). The most N-terminal zf-C2H2 is absent or partially eroded (i.e. without the conserved cysteines and histidines) in some species (Fig 1B), including *D. melanogaster*. We define this N-terminal zf as ZF-Novel, since it has not been reported in previous studies. This observation was confirmed by inspecting *Dm-zld* alternative splicing isoforms. On the other hand, the second N-terminal zf-C2H2 domain is conserved in virtually all extant insects, but absent or degenerate in the other Pancrustaceans (e.g. *Daphnia magna*; partially conserved, with one lost cysteine) (Fig 1). Further, between this second zf-C2H2 and the JAZ domain, there is a strikingly conserved acidic patch that is characterized by an absolutely conserved motif of the form [DE]I[LW]DLD (Fig 1), which is predicted to adopt an extended conformation (using the JPRED software [25]) amidst surrounding disordered regions. A related conserved acidic motif was also found in the chordate protein CECR2, C-terminal to the DDT and WHIM motifs, which constitute a helical





**Fig 1. Tc-Zld proteins from insects and crustaceans.** (A) Multiple sequence alignment of Zelda proteins, representing major groups of arthropods. (B) Conserved protein architecture features of Zelda proteins. Asterisks and x marks represent presence and absence, respectively. Question marks denote that the feature is either partially preserved or could be flagged as absent due to sequencing or assembly errors (e.g. wrong start codons). The cladogram was organized according to a previously reported phylogenetic study [79]. Two outgroups without *zld* orthologs were also included.

<https://doi.org/10.1371/journal.pgen.1006868.g001>

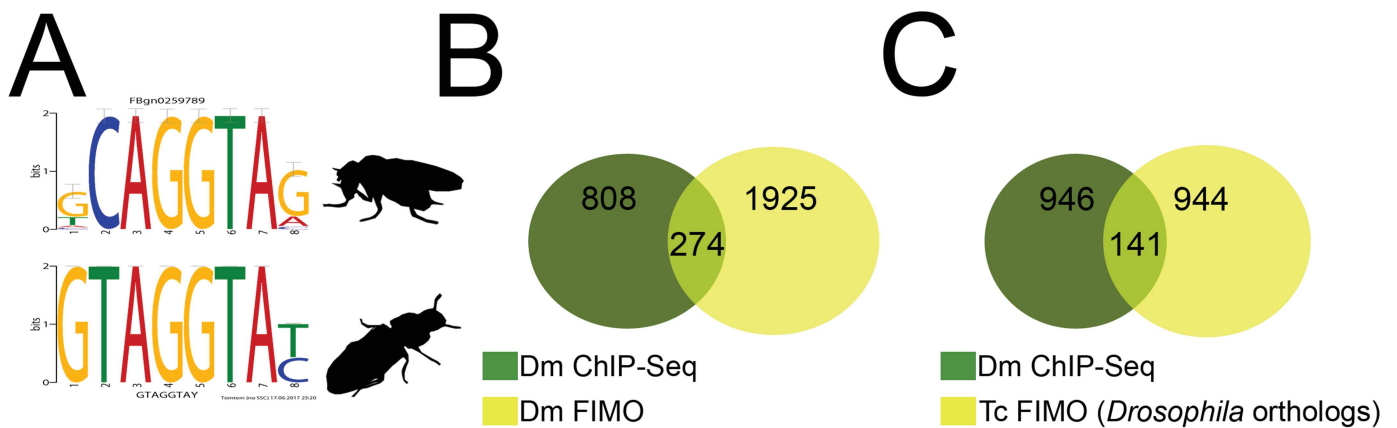
domain involved in setting the nucleosome spacing in conjunction with the ISWI ATPase during chromatin remodeling [26]. An analogous conserved acidic patch is also seen in the HUN domain which functions as a histones chaperone [27]. Taken together, these observations raise the possibility that this Zld acidic region interacts with positively charged chromatin proteins such as histones. A largely disordered region, between the JAZ finger and the cluster of 4 widely-conserved zf-C2H2 domains, has been shown to be important for *Dm-zld* transactivation *in vitro* [24]. Our analysis of evolutionary constraints on the protein sequence also revealed a motif of the form hP[IVM]SxHHHPxRD, which appears to be under selection for retention despite the strong divergence in this region. Hence, it is possible that it specifically plays a role in transactivation. Further, between the two N-terminal zf-C2H2 domains, there is a highly conserved RYHPY motif, which could be involved in nuclear localization (Fig 1A), as predicted for other TFs [28]. Given the conservation of these additional domains, we hypothesize that they also play important roles in Zld functions that were previously attributed exclusively to the C-terminal domains.

Aiming to elucidate the origins of *zld*, we performed extensive sequence searches and were unable to find homologs with the conserved canonical domain architecture outside of Pancrustacea, indicating that *zld* is an innovation of this lineage. We detected clear *zld* homologs across insects, including the termite *Zootermopsis nevadensis* (Order Isoptera), the scarer chaser *Ladona fulva* (Order Odonata), the mayfly *Ephemera danica* (Order Ephemeroptera) and *Machilis hrabei* (Order Archaeognatha). *zld* homologs were also found in crustaceans belonging to different classes (*Daphnia magna*, *Hyalella azteca* and *Eurytemora affinis*), as well as in the collembola *Folsomia candida*. Curiously, we found no *zld* in the genome of *Orchesella cincta* (collembola) and *Daphnia pulex* (crustacean), suggesting either that it is not absolutely conserved outside of insects or missing due to incompleteness of the deposited genomes. Specifically, we carefully searched the genomes of other non-insect arthropods, including chelicerates (e.g. the tick *Ixodes scapularis* and the spider *Parasteatoda tepidariorum*), and found no proteins with the canonical Zld domain organization (Fig 1B). General searches on Genbank against non-Pancrustacea arthropod proteins also returned no Zld orthologs. Although BLAST searches with Zld proteins against the *nr* and *refseq* databases recovered a number of significant hits in several distant eukaryotes, the similarity is almost always restricted to the C-terminal cluster of four Zf-C2H2 domains, which are very common across several TF families (e.g. *glass*, *earmuff/fez*, *senseless/gfi-1* and *jim*). Taken together, our results support the early emergence of *zld* in the Pancrustacea lineage, with subsequent losses in particular species. Importantly, all the insect genomes we inspected have exactly one *zld* gene, indicating that this gene became essential in hexapods.

### *Tc-zld* is a master regulator of signaling genes and other transcription factors

Previous studies in *D. melanogaster* using microarrays and ChIP-Seq revealed that *Dm-zld* regulates the transcription of hundreds of genes during early embryogenesis [6, 9, 29]. In *D. melanogaster*, enhancers bound by Dm-Zld are characterized by a consensus sequence CAGGTAG (i.e. TAGteam sequence), which is overrepresented in early zygotic activated genes, including TFs involved in AP and DV patterning [7, 8]. Since the TAGteam motif identified in *D. melanogaster* is conserved in *A. aegypti* [30], we investigated whether we could predict *Tc-zld* targets by detecting TAGteam motifs in the upstream regions of *T. castaneum* genes.

Firstly, an *ab initio* approach using DREME [31] was employed to analyze 2kb upstream regions of all *T. castaneum* protein-coding genes. This analysis uncovered a motif (i.e. GTAGGTAY) that is nearly identical to the TAGteam motif (Fig 2A). We used the *D.*



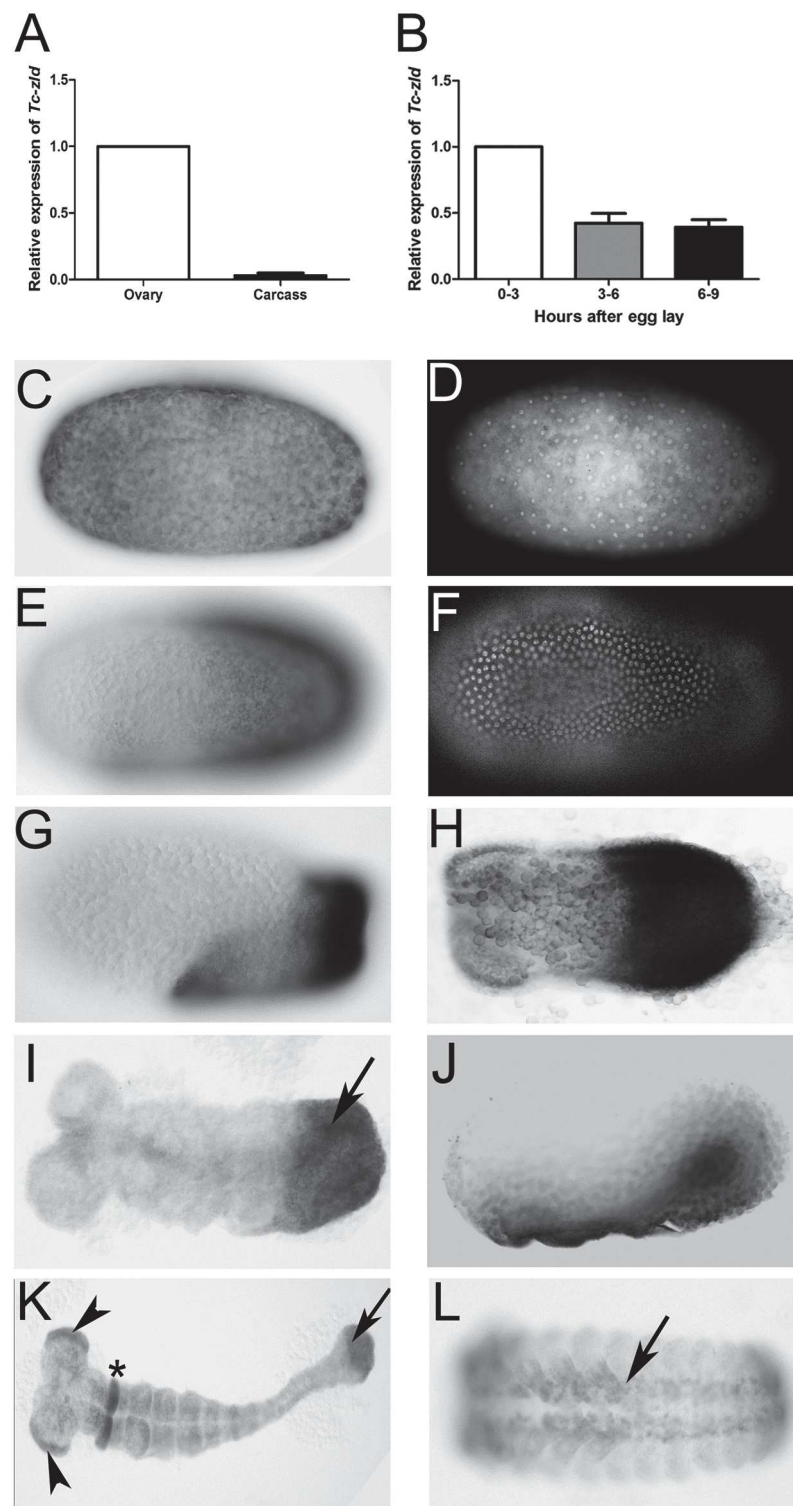
**Fig 2. Computational identification of *Tc-zelda* target genes.** (A) TOMTOM comparison of *D. melanogaster* motif similar to the TAGteam [7] obtained by DREME and putative *T. castaneum zld* DREME motif. (B) Venn diagram of the *D. melanogaster* ChIP-Seq MZT regulated genes (green) from [8] and *D. melanogaster* genes predicted by FIMO analysis with the putative DREME motif (yellow). (C) Venn diagram of the *D. melanogaster* Dm-Zld MZT targets (green) [8] and *D. melanogaster* one-to-one orthologs of putative *Tc-zld* targets.

<https://doi.org/10.1371/journal.pgen.1006868.g002>

*melanogaster* genome and experimental data [6, 9, 29] to validate our approach and found a significant overlap between experimental and predicted *zld* targets in *D. melanogaster* (Fig 2B). The putative *T. castaneum* motif was then used to screen the *T. castaneum* genome, resulting in the identification of 3,250 putative *zld* targets, representing ~19% of the *T. castaneum* genome (Fig 2C, S2 Table). Comparison of the putative *Tc-zld* targets with 1,087 genes regulated by Zld during *D. melanogaster* MZT [8, 29] allowed the identification of 141 *D. melanogaster* genes for which one-to-one orthologs figured among the putative *Tc-zld* targets (hypergeometric distribution,  $P < 4.5 \times 10^{-4}$ ; Fig 2C, S3 Table). Functional analysis of this gene set using DAVID [32] uncovered the enrichment of important categories, including a highly significant cluster of 26 homeobox TFs (S4 Table) and other significant clusters comprising genes involved in regionalization and segment specification, imaginal disc formation and metamorphosis (S4 Table). Interestingly, this gene set included multiple developmental regulators such as anteroposterior (AP), gap, pair-rule, homeotic and dorsoventral (DV) genes (S2 Table). Since several of these genes are involved in early developmental processes, we focused our initial analysis of *Tc-zld*'s function during embryogenesis.

### *Tc-zld* is maternally provided and dynamically expressed during embryogenesis

Since *zld* is maternally expressed in the germ line in *D. melanogaster*, we compared its transcription in *T. castaneum* female ovaries and carcass by qRT-PCR. *T. castaneum* early development starts with synchronous divisions during the first three hours of embryogenesis (at 30°C), followed by nuclear migration to the egg cortex and membrane segregation of nuclei into separate cells, approximately 7–8 hours after egg lay [18, 22]. We found that *Tc-zld* is highly expressed in the ovaries (Fig 3A), supporting the transcription of *Tc-zld* in the germ line. The abundance of *Tc-zld* transcripts is also higher in the first three hours of development than in the next two 3-hour periods (i.e. 3–6 and 6–9 hours) (Fig 3B), suggesting that *Tc-zld* mRNA is maternally provided and degraded after the first 3 hours of development. An antibody against the transcriptionally active form of RNA polymerase II, previously used in other ecdysozoan species [33, 34], showed that zygotic transcription in *T. castaneum* begins between three and six hours of development, shortly after the nuclei have reached the periphery (Fig 4).



**Fig 3. *Tc-zelda* is maternally provided and progressively confined to the posterior growth-zone during embryonic development.** In all embryos, the anterior region points to the left. (A) Relative expression of *Tc-zld* in ovary and carcass. (B) Relative expression of *Tc-zld* 0–3, 3–6 and 6–9 hours after egg lay. Expression values were normalized using the constitutive gene *rps3* in both experiments. (C) Pre-blastoderm stage embryo (0–3 hours) shows *Tc-zld* transcripts uniformly distributed and its respective DAPI staining in D. (E) At uniform blastoderm stage (3–6 hours) transcripts begin to occupy the germ rudiment, respective DAPI

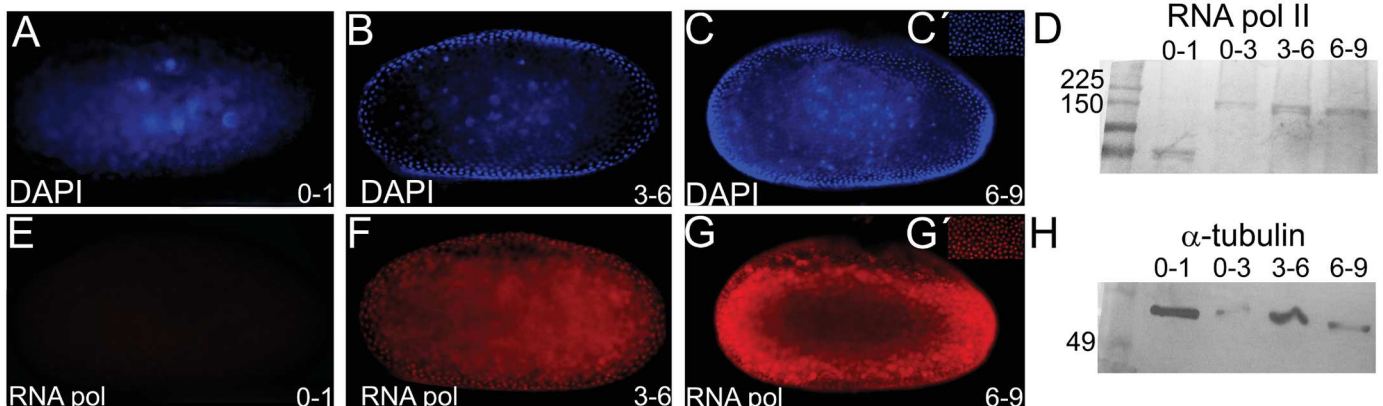
in F. (G) Shortly before posterior invagination, *zld* expression occurs along the whole germ rudiment, with higher levels at the posterior region where growth zone will form (6–9 hours). (H) Shortly before the beginning of germ band elongation *zld* expression is highly expressed at the GZ, although lower levels of mRNA can be observed at the along the whole embryonic region. (I) Approx. 13 hours after egg lay (AEL), during the beginning of germ band extension, *Tc-zld* is expressed at the posterior GZ (arrow). (J) An embryo slightly older than the one in H, highlighting *Tc-zld* expression at ventral serosal cells during serosal window closure. (K) Approx. 18–21 hours AEL *zld* is still expressed at the posterior most region (arrow), head lobes (arrowheads) and a single gnathal segment (asterisk). (L) Approx. 39–48 hours AEL, during early dorsal closure, expression is observed at the nervous system (arrow).

<https://doi.org/10.1371/journal.pgen.1006868.g003>

*In situ* hybridization confirmed maternal ubiquitous expression of *Tc-zld* in the first three hours of development (Fig 3C and 3D) and showed a progressive confinement of *Tc-zld* mRNA to the posterior region of the egg, where the germ rudiment will be formed (Fig 3E and 3F). Between 6 and 9 hours of development, *Tc-zld* expression is observed at the embryonic tissue (Fig 3G), with higher levels at the posterior region (Fig 3H), where the GZ will generate new segments [19, 35]. In addition, *Tc-zld* is also expressed in the ventral serosa, during the serosal window closure (Fig 3I). Later in development, *Tc-zld* expression is still detected at the GZ (Fig 3J), at the head lobes and a single gnathal segment (Fig 3K) and, subsequently, in the nervous system (Fig 3L). Although *zld* is maternally provided, and zygotically expressed at the neural progenitors of both *D. melanogaster* and *T. castaneum* [6, 11], biased posterior expression is a feature so far described only in short-germ insects like the latter.

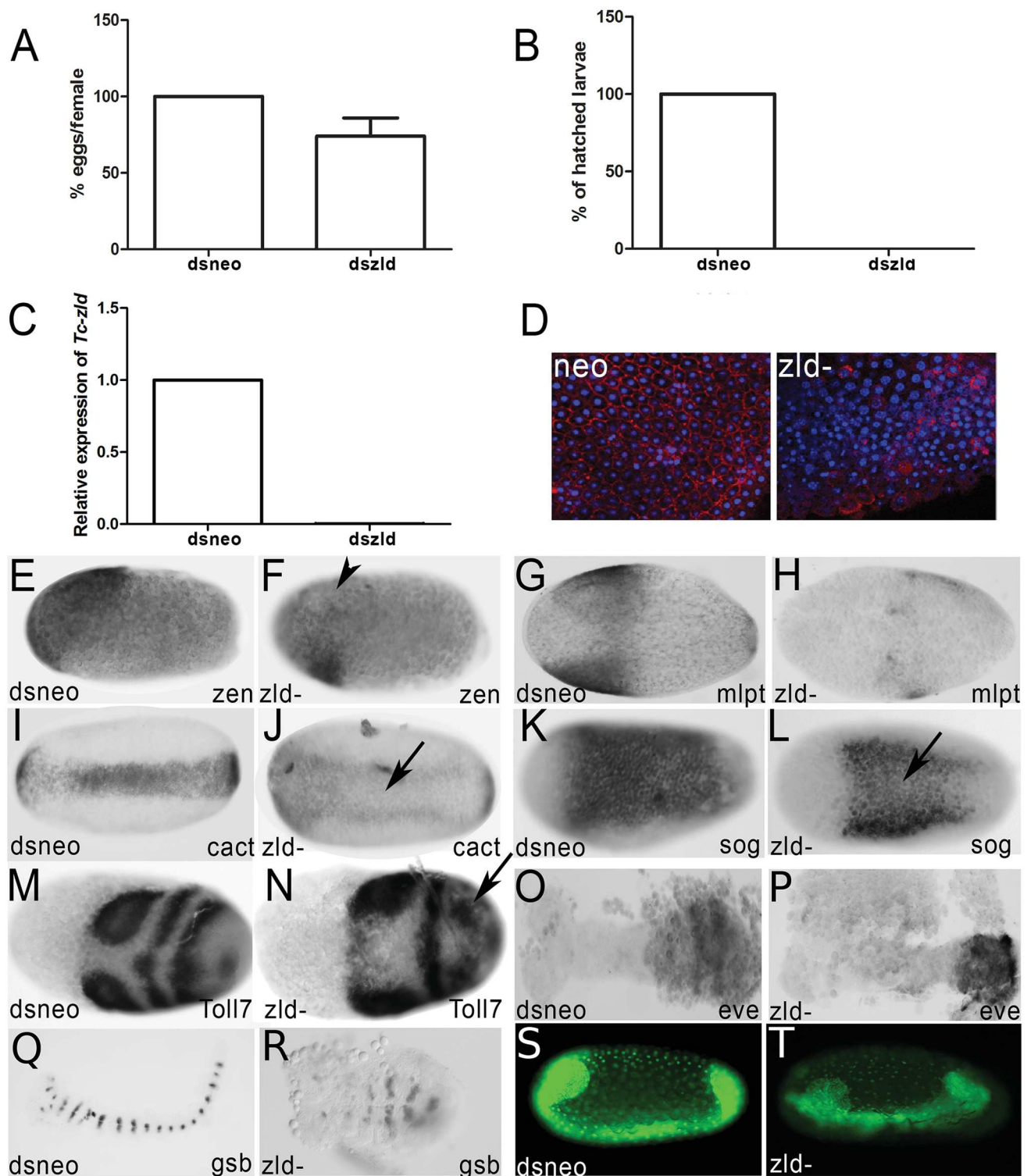
### Parental injection of *Tc-zld* dsRNA reduces number of eggs laid and impairs embryogenesis

It has been shown that injection of dsRNA in *T. castaneum* females leads to reduced expression of a given gene in the females and their offspring, in a phenomenon called parental RNAi (pRNAi) effect [15]. We injected *zld* dsRNA in females and analyzed *Tc-zld* transcriptional levels during embryogenesis by qRT-PCR. After *zld* pRNAi, *Tc-zld* mRNA levels were reduced in the first two weeks of egg laying, severely impairing larval hatching (Fig 5A–5C). Importantly, identical knockdown phenotypes during embryogenesis were obtained by using a second,



**Fig 4. The onset of Maternal Zygotic Transition (MZT) in *Tribolium castaneum*.** (A–C) Nuclear DAPI staining of *T. castaneum* embryos between 0–1 hours (A), 3–6 hours (B) and 6–9 hours after oviposition (C). (D,H) Western-blots of embryonic extracts from 0–1, 0–3, 3–6 and 6–9 hours after oviposition using an antibody against the transcriptionally active form of RNA pol II, as previously described [33]. (D) or an antibody against the  $\alpha$ -tubulin protein as a loading control (H). (E,F,G) Immunostaining showing nuclear RNA pol II at 3–6 hours (F) and 6–9 hours (G), but not at 0–1 hour (E) after oviposition. Coupling RNA pol II staining with nuclear DAPI staining shows that *T. castaneum* zygotic transcription starts between 3–6 hours after egg laying, when the energids reaches the periphery (Fig 2A–C–DAPI and E–G–RNA pol II). Similar results were obtained by western-blots using the same antibody (D).

<https://doi.org/10.1371/journal.pgen.1006868.g004>



**Fig 5. *Tc-zelda* parental RNAi affects oviposition and cellularization, while embryonic RNAi affects embryonic posterior patterning.** (A) Normalized number of *dsneo* and *dszld* RNAi collected eggs. *zld* pRNAi reduces oviposition. (B) Percentage of hatched larvae obtained from eggs after *zld* pRNAi in comparison to the control (*dsneo* RNAi). *zld* pRNAi leads to lethality during embryogenesis. (C) Relative expression of *zld* in *dsneo* RNAi (control) and *dszld* RNAi during the first two egg lays. *zld* pRNAi almost completely abolishes *zld* expression. (D) Phospho-tyrosine (red) and DAPI (blue) staining in *dsneo* and *zld* RNAi eggs during cellularization. *zld* pRNAi impairs cellularization. (E,F) *Tc-zen* expression in the serosa [36] is reduced after *Tc-zld* pRNAi (arrowhead) when compared to the *dsneo* control. (G,H) Expression of the

gap gene *Tc-mlpt* [37] in control (G) and *zld* RNAi (H). *Tc-mlpt* loses its anterior expression domain in *zld* RNAi when compared to control. (I,J) The expression of the dorsoventral gene *Tc-cactus* [38] was reduced after *Tc-zld* pRNAi (J-arrow) when compared to the (I) control. (K,L) The expression of the dorsoventral gene *Tc-sog* was reduced after *Tc-zld* pRNAi (J-arrow) when compared to the (I) control. (M,N) Expression of *Tc-Toll7* was affected in the posterior region of *Tc-zld* RNAi embryos (N) when compared to *dsneo* (M). (O,P) Expression of the segmentation gene *Tc-eve* in control (O) and *zld* RNAi (P). *Tc-eve* expression is essential for segmentation via a pair rule circuit in *T. castaneum* [39]. After *zld* RNAi the characteristic *Tc-eve* stripe expression in the growth zone is lost. (Q,R) Segment generation from the growth-zone (GZ) is disrupted, as judged by the analysis of *Tc-gsb* expression after *Tc-zld* pRNAi (R) and *dsneo* (Q). (S,T) *Tc-zld* dsRNA embryonic injection (eRNAi) into nuclear GFP transgenic line affects segment generation from the posterior GZ (L-asterisk), when compared to nGFP embryo injected with *dsneo* dsRNA.

<https://doi.org/10.1371/journal.pgen.1006868.g005>

non-overlapping, dsRNA construct (S1 Fig). Further, morphological analyses showed that cellularization was severely disrupted in over 50% of the *zld* dsRNA embryos (Fig 5D), similarly to what was previously reported in *D. melanogaster zld* mutants [6, 11]. The remaining *zld* pRNAi embryos were not severely affected during cellularization and developed beyond that stage. Finally, we also found that some putative conserved target genes were down-regulated in embryos after *zld* pRNAi, such as the early zygotic genes involved in AP patterning, the serosal gene *Tc-zerknüllt* [36] (Fig 5E and 5F) and the gap gene *milli-pattes* [37] (Fig 5G and 5H).

Changes in the spatial distribution of transcripts from predicted dorsoventral target genes were also observed after *Tc-zld* RNAi. In wild-type (WT) embryos, the TF Dorsal forms a dynamic transient gradient, which activates *Tc-cactus* (*Tc-cact*) and *Tc-short-gastrulation* (*Tc-sog*) at the ventral region [38, 39]. After *Tc-zld* pRNAi, the expression of *Tc-cact* and *Tc-sog* is observed in two lateral domains, in contrast to the expression at the single ventral domain in *dsneo* RNAi embryos (Fig 5L, 5J, 5K and 5L). These results suggest that *Tc-zld* is required for proper activity of Dorsal at the ventral-most region of the embryo.

As discussed above, *Tc-zld* is highly expressed at the posterior region of the embryo (Fig 3) and likely associated with segmentation and regionalization (S4 Table). Further, several putative *Tc-zld* targets are involved in posterior segmentation, such as *caudal* (*Cdx*), *even-skipped* (*Eve*) and several Hox genes (e.g. *Ultrabithorax*, *Abdominal-A* and *Abdominal-B*). *Tc-eve*, for example, is essential for the establishment of a genetic circuit required for posterior segmentation [40]. Interestingly, *Tc-zld* pRNAi embryos showed a continuous *Tc-eve* expression domain instead of the typical stripe patterning required for WT segmentation (Fig 5O and 5P).

Elegant studies on *T. castaneum* GZ patterning showed that cell proliferation is not essential for segment generation, which rather occurs by coordinated cell movement and intercalation [19, 41]. We then evaluated whether *Tc-zld* regulates genes involved in cell intercalation, such as *Toll2*, *Toll6* and *Toll8* [42, 43]. Interestingly, *Toll7* (TC004474) and *Toll8* (*Tollo*:TC004898) are among the common *zld* targets conserved between *D. melanogaster* and *T. castaneum* (S3 Table). Since *Tc-Toll7* is expressed during early segmentation [43], we compared its expression in *dsneo* and *Tc-zld* RNAi embryos. While anterior expression of *Tc-Toll7* is apparently unaffected, the striped expression at the posterior region is lost when *Tc-zld* expression is reduced (Fig 5M and 5N). Further support for the loss of posterior segmentation after *Tc-zld* RNAi is also provided by the analysis of expression of the segment-polarity gene, *Tc-gooseberry* (*Tc-gsb*) (Fig 5Q and 5R). In summary, *Tc-zld* regulates the expression of several genes that are critical for early AP (*zen*, *mlpt*) and DV (*sog*, *cact*) patterning and, in a second phase, genes required for posterior elongation (e.g. *Toll7*, *gsb*).

### *Tc-zld* plays specific roles in the posterior GZ

While pRNAi diminishes maternal and zygotic expression of *Tc-zld* in *T. castaneum* [15], embryonic dsRNA injections (eRNAi) may affect only the zygotic component, since eggs can

be injected after the MZT [38, 44]. To investigate if *Tc-zld* is specifically required for embryonic posterior patterning, we injected *Tc-zld* dsRNA in transgenic embryos expressing nuclear GFP (nGFP), as previously described [18, 19]. Embryonic injections of *Tc-zld* dsRNA after the MZT (see [Methods](#) for details) impaired segment generation from the GZ, while *dsneo*-injected embryos developed like WT ones ([Fig 5S and 5T, S1 and S2 Movies](#)). In addition, expression of the predicted target *Tc-eve*, a key TF involved in GZ patterning [35], has been largely down-regulated upon *Tc-zld* eRNAi, as previously observed for *zld* pRNAi ([Fig 5O and 5P](#)). In summary, our results imply that *Tc-zld* is involved not only in the MZT, early patterning and nervous system formation, as described for *D. melanogaster* [6], but also play roles in segment generation from the GZ, a structure found only in embryos of short-germ insects like *T. castaneum*.

### Functional analysis in the hemipteran *Rhodnius prolixus* shows a conserved role of *zld* in short-germ insect posterior region

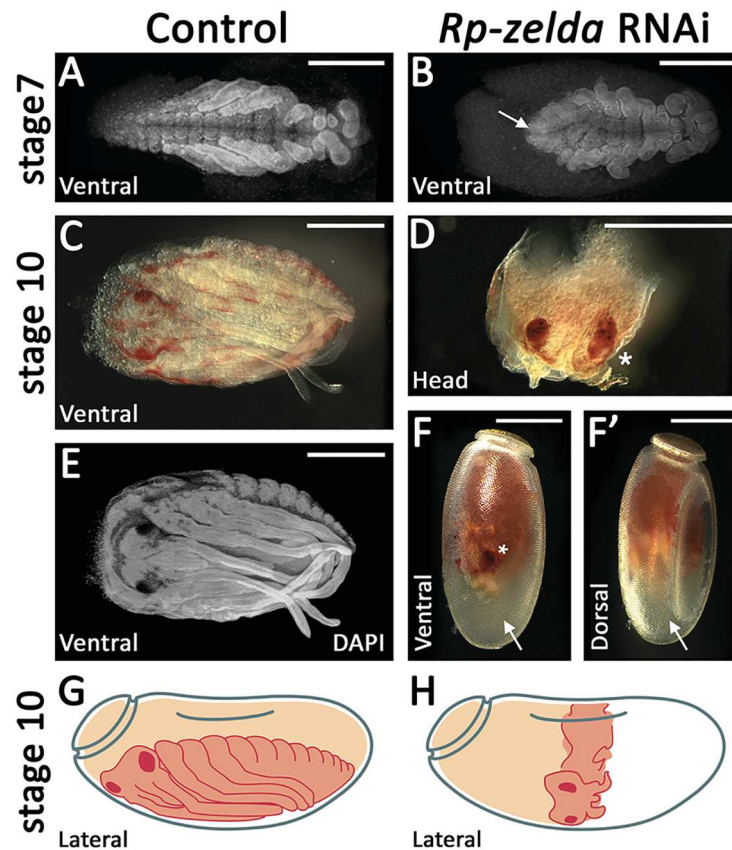
Although we demonstrated the involvement of *zld* in *T. castaneum* GZ, the conservation of this regulatory mechanism in other species remained unclear. Thus, we sought to analyze the functions of the *zld* ortholog in the hemipteran *R. prolixus* (*Rp*; *Rp-zld* gene), which is a hemimetabolous insect and lacks the complete metamorphosis present in holometabolous species such as *T. castaneum* [45, 46]. *Rp-zld* knockdown via pRNAi resulted in two types of embryonic phenotypes: 1) severe defects in gastrulation and lack of any appendage development; 2) embryos that developed only the anterior-most embryonic regions comprising the head, gnathal and thoracic segments ([Fig 6](#)). These results support a role for *zld* in posterior elongation that was likely present in the common ancestor of hemimetabolous true bugs and holometabolous insects, if not earlier. To our knowledge this is the first direct description of *zld* function in insects other than *D. melanogaster*. A recent report showed *zld* as maternally transcribed in the hymenoptera *Apis mellifera* (also a long-germ insect), while zygotic transcripts are concentrated at the central region of the embryo during blastodermal stages [47]. Further, during the preparation of this manuscript, *zld* has been also proposed to be required for MZT and cellularization in the wasp *Nasonia vitripennis* [48].

### *Tc-zld* is essential for patterning of imaginal disc-derived structures: Wing, legs, elytra and antenna

DAVID analysis of *D. melanogaster* orthologs of the putative *Tc-zld* targets uncovered a cluster of 29 genes involved in imaginal disc development ([S4 Table](#); GO:0007444). Among these genes are several homeodomain TFs, such as *distalless* (*Dll*), *Abdominal A* (*Abd-A*), *Abdominal B* (*Abd-B*), *zen*, *Engrailed* (*En*), *caudal* (*cad*), *defective proventriculus* (*dve*), *mirror* (*mirr*), *arauacan* (*ara-iroquois*) and *Drop* (*dr*), as well as other TFs such as *dachsund* (*dac*), *taranis* (*tara*) and *Lim1*, *PoxN*, *kn*, *sob*, *drm*, *Awh*, *dp*.

As the first step towards the characterization of the post-embryonic role of *zld*, we analyzed *Tc-zld* expression by qRT-PCR in larvae of third (L3), fifth (L5) and seventh (L7) stages, and first pupal stage (P1) ([Fig 7A–7D](#)). Interestingly, *Tc-zld* expression increases during successive larval stages and sharply decreases after pupal metamorphosis ([Fig 7E](#)). This suggests that *Tc-zld* might be required for late larval stages, which take place during larvae-pupae metamorphosis, such as growth and patterning of structures derived from imaginal discs in *D. melanogaster* (e.g. antennae, legs, fore- and hindwings) [49, 50]. Further, we found that three out of five predicted *Tc-zld* targets, namely *Dll*, *Wg* and *Lim-1*, also displayed an increase in expression during late larval and pupal development ([Fig 7E](#)).



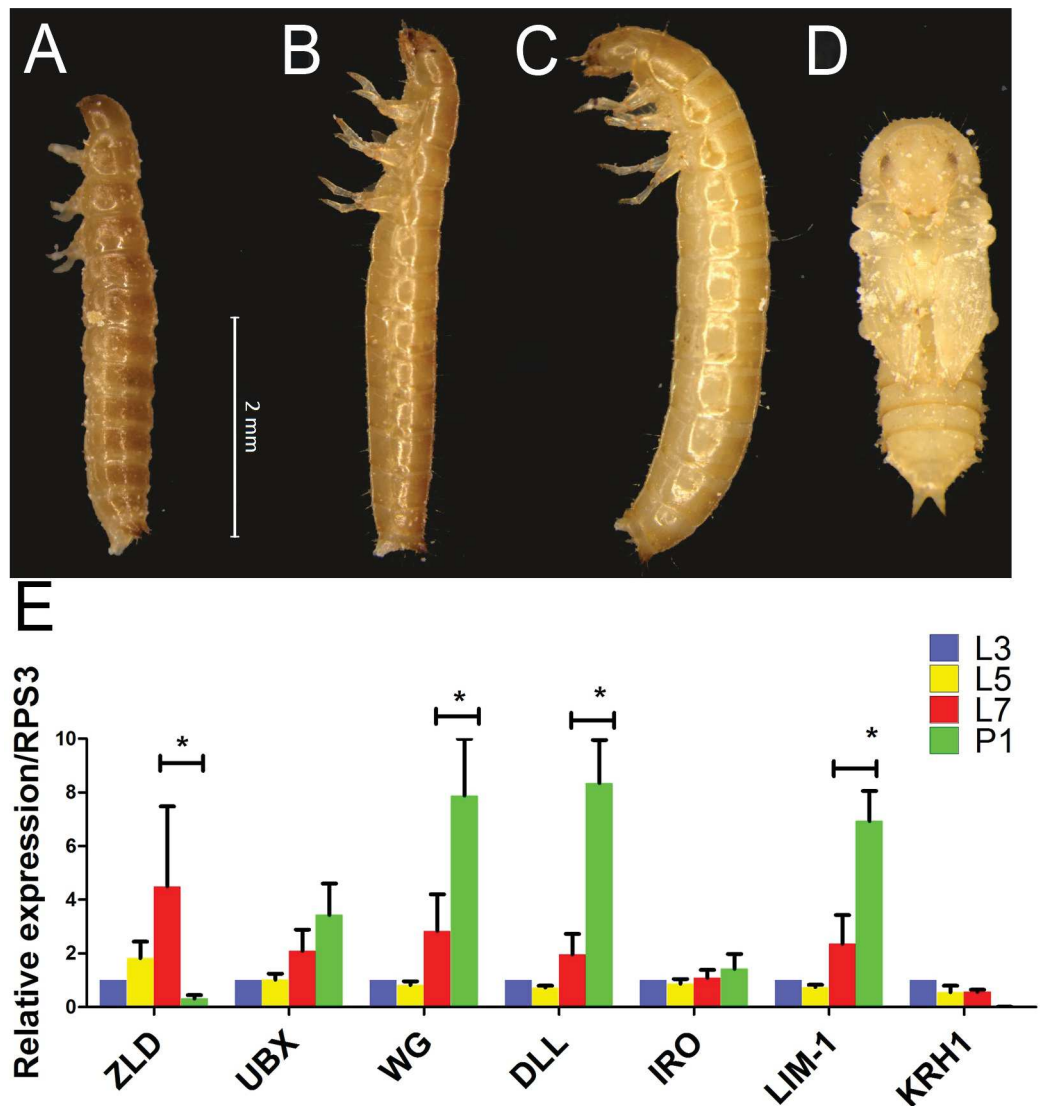


**Fig 6. *zld* is required for the generation of the posterior region in the hemiptera *Rhodnius prolixus*.** (A, C, E) *R. prolixus* control (*dsneo* embryos). (B, D, F, F') A representative embryo collected from *R. prolixus zld* pRNAi. (F, F') Embryo inside the chorion, ventral and dorsal views. (C, E) Embryo removed from the egg shell. (G, H) Schematic drawings of control and *zld* pRNAi embryos. (A) DAPI stainings of control (A, E) and *zld* RNAi (B) embryos. (D) Asterisk denotes an eye which can be observed at the ventral side (F). (D) After dissection, the eye can be identified due to its characteristic red pigmentation and shape. Scale bar corresponds to 500 $\mu$ M.

<https://doi.org/10.1371/journal.pgen.1006868.g006>

To investigate *zld*'s post-embryonic roles, we injected two non-overlapping *Tc-zld* dsRNA constructs into early (L3) and late larval (L6) stages, as previously described [50]. qRT-PCR confirmed that *Tc-zld* was down-regulated after dsRNA injection (Fig 8A). *Tc-zld* dsRNA injections at early larval stages (L3) led to over 50% lethality during pupal stages. Atypical adult pigmentation in the pupal head and reduction in the wing size were observed after early *Tc-zld* dsRNA injection, suggesting that proper *Tc-zld* expression is required for metamorphosis and wing growth (Fig 8B and 8C). Interestingly, *Tc-zld* dsRNA injections at late larval stage (L6) displayed a different phenotype when compared with early larval dsRNA injections (L3). L6 larvae injected with *dszld* and *dsneo* reached adulthood at comparable rates (Fig 8D). Specifically, *Tc-zld* dsRNA adults showed a series of morphological alterations in tissues undergoing extensive morphological changes during metamorphosis, such as fore- and hindwings, antennae and legs (Figs 8E, 8F, 9 and 10).

The most visible effect of *Tc-zld* dsRNA beetles was a failure of the forewings (elytra) to enclose the hindwings, leading to the exposure of the dorsal abdomen (Fig 8E and 8F). Elytra, which are highly modified beetle forewings, have been proposed to be an important beetle innovation, being required for protection against mechanical stresses, dehydration, and

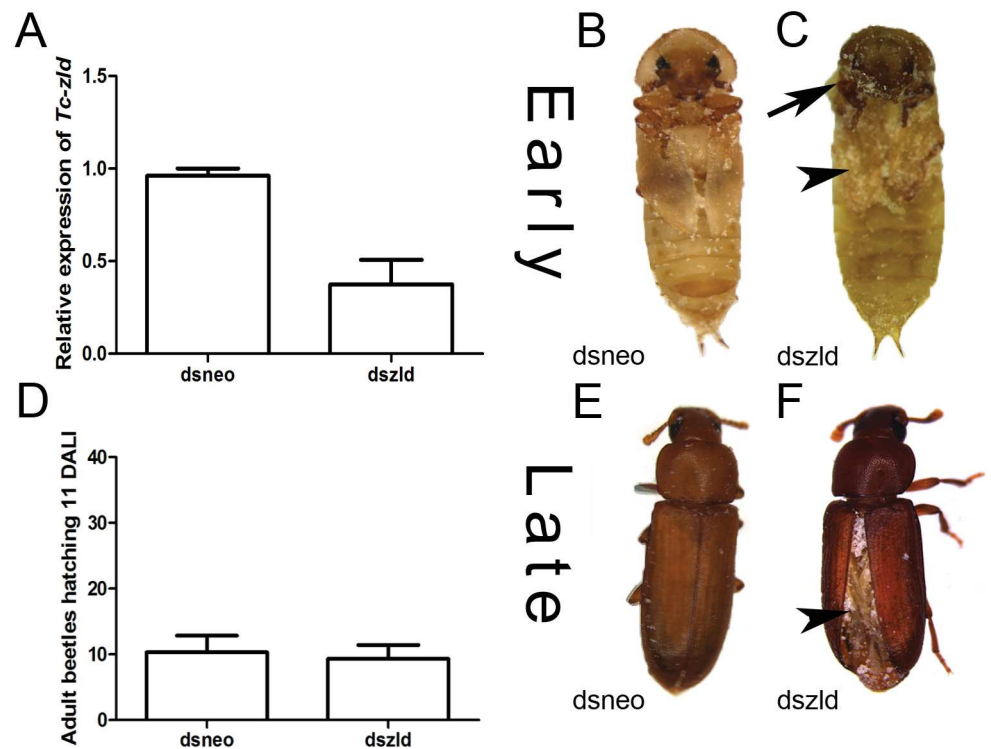


**Fig 7. Larval stages and expression dynamics of *zld* and its putative target genes.** (A-D) Morphology of *Tribolium castaneum* larvae on 3rd (A-L3), 5th (B-L5), 7th (C-L7) and early pupal stages (D-P1). (E) Relative expression of *zld*, *Ubx*, *wingless (wg)*, *distalless (dll)*, *Iroquois (iro)*, *Lim-1* and *Kruppel-homolog-1 (KRH-1)* at L3, L5, L7 and P1. Asterisks represent significant differences between stages (P < 0.05).

<https://doi.org/10.1371/journal.pgen.1006868.g007>

predation [51]. In line with this hypothesis, *Tc-zld* dsRNA adults with exposed abdomens started to die a few days after metamorphosis, probably due to dehydration.

Next, we performed a detailed morphological analysis to investigate if patterning defects resulting from *Tc-zld* dsRNA occurred in the sclerotized elytra (forewing) or in the hindwing (Fig 9A–9D). This analysis showed that the parallel vein pattern of the elytra (Fig 9A) is disrupted after *Tc-zld* dsRNA in comparison to the control (Fig 9A and 9B). Nevertheless, hindwings of *Tc-zld* dsRNA beetles showed no signs of abnormal venation (Fig 9C and 9D). To address if fore or hindwing shapes have changed upon *Tc-zld* dsRNA knockdown, we applied the recently developed Proximo-Distal (PD) index, a morphometric analysis consisting of the measurement of the wing length and width at two positions [52]. While the PD index values of forewings (elytra) of *Tc-zld* and controls were similar, a slight but significant increase in the



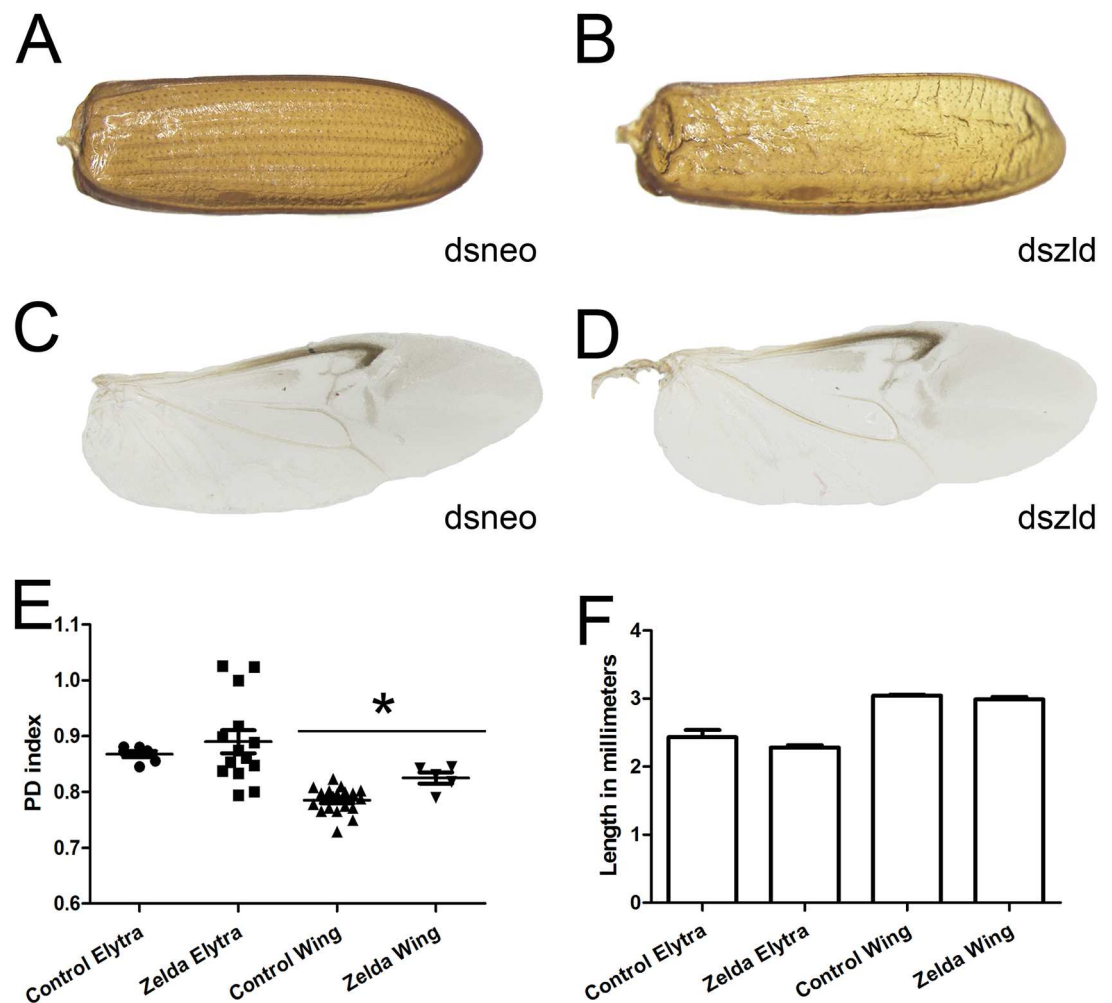
**Fig 8. *Tc-zelda* larval RNAi affects elytra enclosure and molting.** *Tribolium castaneum* larvae were injected during early (3rd) and late (6th) stages as previously described [50, 80]. (A) Relative expression of *Tc-zld* during pupal stages after *Tc-zld* dsRNA or *neo* dsRNA injection. qRT-PCR was normalized using *Tc-rps3* gene, as previously described [78]. (B,C) Morphology of late pupae obtained from early larvae injected with *zld* or *neo* dsRNA. Differential pigmentation in the head (arrow) and reduced wings (arrowhead) were observed in *zld* dsRNA and not in *dsneo* dsRNA pupae. (D) Number of emerging adult beetles eleven days after larval injection (DALI) at late (6th) stage. (E,F). Adults obtained by late larval injections of *neo* (E) or *zld* (F) dsRNA. Hindwings are usually not properly folded underneath the forewing (elytra).

<https://doi.org/10.1371/journal.pgen.1006868.g008>

hindwing PD index was observed (Fig 9E). Neither fore- nor hindwings length were altered upon *Tc-zld* knock down (Fig 9F). In conclusion, *Tc-zld* is required for proper venation pattern, but not shape, of the elytra. Previous analysis of *zld* expression in *D. melanogaster* showed expression in wing imaginal discs, particularly where mitotically active cells are located [11]. Moreover, *Dm-zld* overexpression during wing imaginal disc formation leads to adult wing blisters or tissue loss [10, 11]. Nevertheless, to our knowledge, our study provides the first direct evidence of *zld*'s role in insect wing formation.

Interestingly, all *Tc-zld* dsRNA beetles that showed this 'opened wing' phenotypes (Fig 8F) also displayed defects in the legs and antennae. Antennae and legs share similar developmental GRNs, involving the so called serial homologs. *Distalless (Dll)*, one of the putative *Tc-zld* targets, is essential for appendage segmentation [53]. *Tc-Dll* is also expressed during late embryogenesis on the distal part of the leg and, as its name suggests, disruption of its function leads to the absence of distal leg and antennae segmentation [54]. Interestingly, we found that *Tc-zld* RNAi resulted in a significant decrease of *Tc-Dll* mRNA levels (Fig 10A), indicating that this gene is indeed downstream of *Tc-zld*.

Insect antennae possess three primary segments: scape, pedicel and flagellum. In *T. castaneum*, the adult antennae display eleven segments, out of which nine form the flagellum. The three most distal flagellar segments are enlarged and form the club, while the six intermediate

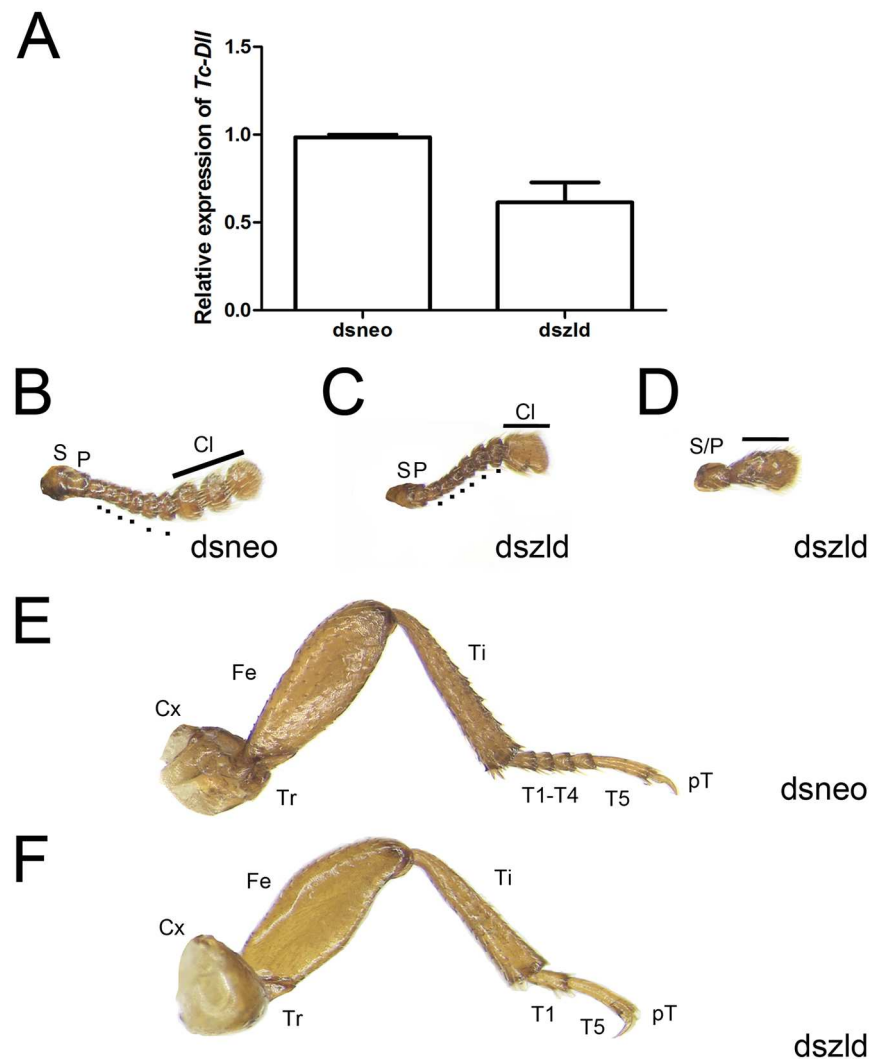


**Fig 9. *Tc-zld* knockdown in larval stages affects elytra and wings in adult stage.** (A) Control elytron extracted from a *dsneo* adult shows a parallel venation pattern and rigid chitinous structure. (B) *Tc-zld* RNAi elytron displays a disrupted vein pattern and a less resistant structure. (C,D) Hindwings were dissected and photographed. Although the overall morphological pattern is not affected in both wings, *zld* dsRNA wings show severe dehydration after ethanol fixation. (E) PD indexes comparison between *dsneo* and *dszld* RNAi hindwings and forewings, a PD index reflects the shape of the wing based on its dimensions ratio [52]. The statistical analysis was carried out by unpaired t-test assuming unequal variances (asterisks refer to  $P < 0.0001$  while ns stands for “no significance”), indicating that *Tc-zelda* RNAi affects hindwing shape. (F) *Tc-zld* RNAi elytra and hindwing do not show statistically significant differences in length when compared to their respective *dsneo* RNAi controls.

<https://doi.org/10.1371/journal.pgen.1006868.g009>

flagella are called the funicle (Fig 10B) [55]. In mild phenotypes, *dszld* caused a joint malformation in distalmost flagellar segments resulting in a fusion of the club (Fig 10C). On the other hand, strong *Tc-zld* dsRNA phenotypes resulted in fusion of the scape and pedicel, leading to severe loss of flagellar joints and formation of a single truncated segment (Fig 10D). In contrast, no differences in scape and pedicel were observed.

*T. castaneum* legs originate during late embryogenesis and can be recognized as a small outgrowth of the body wall, the limb bud. In the adult stage, there are three pairs of segmented legs with six segments: coxa, femur, trochanter, tibia, tarsus and pretarsus [56]. In *T. castaneum* the tarsus is subdivided in smaller segments (i.e. the tarsomeres), five in prothoracic and mesothoracic legs and four in metathoracic legs. Tarsal segmentation occurs during beetle



**Fig 10. *Tc-zld* knockdown induces phenotypes in antennae and legs of adult *T. castaneum*.** (A) Relative expression of *Tc-Dll* at pupal stages after *neo* or *zld* dsRNA injections. (B–D) Antennae extracted from control injected individuals (B) display 11 segments, including two proximal (S = scape, P = pedicel), six intermediate (Flagellum = dots) and three distal modified segments that form the club (Cl). (C,D) Injection of *zld* dsRNA caused fusion of the club segments in mild phenotypes (C), while in stronger phenotypes (D) the antennae fail to segment, developing as an antennal rudiment. (E,F) Legs of *T. castaneum* display six podomeres: Coxa (Cx), Trochanter (Tr), Femur (Fe), Tibia (Ti), Tarsus (T) and pretarsus (pT). (E) Mesothoracic legs of *neo* dsRNA injected individuals show five tarsal segments, or tarsomeres. (F) *zld* dsRNA beetles lacks tarsal segments resulting in a reduced tarsus, while other leg structures do not display large morphological defects.

<https://doi.org/10.1371/journal.pgen.1006868.g010>

metamorphosis and this subdivision of the tarsus evolved in the common ancestor of insects, since the tarsus is not subdivided in non-insect hexapods [57]. After *Tc-zld* RNAi, tarsal segments were absent or fused, resulting in leg shortening; in *dsneo* insects, legs were identical to that of WT animals (Fig 10E and 10F). This indicates that some of the *Tc-zld* targets are involved in segment development or joint formation. Interestingly, a large domain of *Dll* expression is observed in beetle leg tarsus [58], and *Tc-Dll* knockdown in beetles also generated legs with tarsomere deletion [58, 59].

## *zld* is required for diverse biological processes during embryonic and postembryonic development

Besides the conservation of *zld* roles in the MZT, our results uncovered two new biological roles of *zld* in *T. castaneum*: regulation of segment generation from the posterior GZ during embryogenesis and; patterning of imaginal disc derived structures. But what do posterior GZ patterning, MZT and imaginal disc development have in common? All these processes require an accurate temporal and spatial coordination of complex GRNs to properly pattern cell populations.

Recently, *Dm-zld* has been demonstrated to be a key factor in the establishment of the early chromatin architecture, particularly for the formation of topologically associating domain (TAD) boundaries [60]. During *D. melanogaster* MZT, *zld* resembles a pioneer TF marking the chromatin of earliest expressed genes [8, 61]. *Dm-zld* also increases chromatin accessibility of the most important TFs involved in DV and AP patterning (*Dorsal* and *Bicoid*, respectively) [62–65]. Further, the addition or removal of *Dm-zld* binding sites influences the timing of activation of *Dorsal* early zygotic targets in *D. melanogaster* [62, 63], suggesting that *Zld* acts as a developmental timer.

Our results in *T. castaneum* showed an extensive and ubiquitous maternal contribution of *Tc-zld* mRNAs, followed by zygotic *Tc-zld* expression along the embryonic rudiment, particularly at the posterior region (Fig 3). It is possible that *Tc-zld* mediates a progressive anteroposterior opening of the chromatin in *T. castaneum*, shortly before these posterior GZ cells undergo convergent extension movements required for germ band elongation [19, 41]. Loss of *Tc-zld* expression might lead to lack of convergent extension due to loss of *Toll7* and *eve* expression and, ultimately, segmentation failure (Fig 5). Since *zld* is also important for the development of the posterior region of the hemimetabolous insect *R. prolixus* (Fig 6), *zld*'s role in posterior region dates back at least to the last common ancestor of Paraneoptera.

*Tc-zld* expression was also observed at the ventral serosa (Fig 3) during embryonic stages. While we cannot rule out that *Tc-zld* plays a role during normal serosal development, lack of segmentation after *Tc-zld* eRNAi cannot be explained by loss of *Tc-zld* expression in this tissue, since embryos without serosa, e.g. *Tc-zen1* RNAi, do not display phenotypic defects under normal developmental conditions [36, 66]. Future studies are warranted to determine if *Tc-zld* plays a specific role in the beetle serosa, particularly by regulating some of its predicted target genes involved in immune responses (e.g. *Toll*, *cactus*).

*Tc-zld* might also play a similar role in the regulation of leg and antenna segmentation during metamorphosis (Fig 10). The number of tarsomere segments in the leg and intermediate funicle are reduced after *Tc-zld* dsRNA injection, suggesting that the post-embryonic role of *Tc-zld* in appendage segmentation might also involve the regulation of complex GRNs, as observed in the posterior embryonic region.

The evolutionary success of hexapods is attributed to a combination of features: their segmented body plan and jointed appendages, which were inherited from their arthropod ancestor and; wings and holometaboly, two features that arose later in insect evolution [67]. It is interesting to note that the specific Pancrustacea gene *zld* is required for most of these processes, such as embryonic segment formation, wing (elytra) patterning, and appendage (antennae and leg) formation during beetle development. Like several other zinc finger proteins that show tremendous lineage-specific diversity in eukaryotes, *zld* appears to have specifically arisen within Pancrustacea and risen to play an important role as a “master TF”. Hence, future studies focusing on how this TF was integrated to the conserved backdrop of developmental TFs and existing GRNs of arthropods would be of great interest.

## Materials & methods

### Beetle rearing

*T. castaneum* beetles were cultivated in whole-wheat flour. For sterilization, the flour was kept for 24h at -20°C and another 24h at 50°C. The beetles were maintained inside plastic boxes of approximately 15x15cm with humidity between 40–80%.

### Bioinformatic analyses

Protein sequence analyses were performed using BLAST [68] and PFAM searches using HMMER3 [69] against proteins available on Genbank [70] and Vectorbase [71]. Dm-Zld (NP\_608356) was used as initial query for sequence searches. Genomic data from the following genomes were obtained from the Baylor College of Medicine Human Genome Sequencing Center: *Eurytemora affinis*, *Hyalella azteca*, *Blattella germanica*, *Catantopus aquilonaris*, *Machilis hrabei*, *Libellula (Ladona) fulva* and *Ephemera danica*. BLAST results and domain architectures were manually inspected. *In silico* searches in the *T. castaneum* genome for over-represented motifs were performed using the DREME software [72], part of the MEME suite software toolkit [73]. The motif with highest identity with Dm-Zld binding site (CAGGTAY) was compared to previously described motifs in the FlyFactorSurvey [74] database with TOMTOM. Further, this motif was used to scan 2000 bp upstream of all predicted *T. castaneum* genes with FIMO [14, 75, 76]. Upstream regions and orthology information were retrieved from Ensembl Metazoa Biomart (<http://metazoa.ensembl.org/biomart/>) [77].

### dsRNA synthesis, parental RNAi and embryonic phenotypic analysis

Two non-overlapping PCR fragments containing T7 promoter initiation sites at both ends were used as templates for Ambion T7 Megascript Kit (Cat. No. AM1334) following the manufacturer instructions (for details see S1 Fig). The amount and integrity of the dsRNA samples were measured by spectrophotometry and agarose gel electrophoresis, respectively. For parental RNAi (pRNAi) analysis, ~0.5µl of dsRNA were injected from a solution containing 1µg/µl of dsRNA into adult female beetles [36]. Eggs were collected for four egg lays (2 day each) and *zld*'s down regulation was estimated by quantitative Real Time PCR (see below).

### *T. castaneum* embryonic dsRNA injections

Egg injections were performed as previously described [38, 44]. Briefly, for the analysis of *Tc-zld* zygotic role, embryos containing nuclear-localized green fluorescent protein (GFP) were collected for one hour and let to develop for an additional three hours (30°C) [19, 35]. After this period, twenty embryos were dechorionated with bleach (2% solution), aligned onto a glass slide and covered with Halocarbon oil 700 (Sigma). Embryos were immediately microinjected at the anterior region with *zld* or *neo* dsRNA at 1 µg/µL concentration with the help of a Nanoinject II instrument (Drummond Scientific Company). After injection, a single nGFP embryo was photographed every five minutes during the following 16 hours (25°C) in a Leica DMI4000 inverted microscope using a GFP filter. Single photographs were used to generate a movie using Windows Movie Maker (S1 and S2 Movies). Phenotypes of all injected embryos (*neo* or *zld* dsRNA) were scored at the end of the experiment.

### *T. castaneum* larval dsRNA injections

Larvae were injected with *zld* or *neo* dsRNA as previously described [50]. Knockdown phenotypes in pupae and adult beetles were generated by injection of *zld* or *neo* dsRNA solutions at a concentration of 1 µg/µL in the dorsal abdomen of individuals on third and sixth larval instars

(n = 40). Following injection, larvae were reared in flour at 30°C and collected periodically for RNA extraction and phenotype annotation. Adult beetles were then fixed in ethanol 95% overnight for further morphological analysis. Immunostainings have been performed as previously described [38].

### Morphological analysis of imaginal disc derived tissues

Antennae, legs, elytra and wings were dissected using forceps and placed in a petri dish for observation. Phenotypic analyses and documentation were performed under a Leica stereoscope model M205.

### PD index

The methods for wing and elytra measurement and PD index were performed according to described by [52]. Leica AF Lite software was used for the wing measurements. Image properties were adjusted in Adobe Photoshop CS4.

### Quantitative real-time PCR

For experiments using embryos, total RNA was isolated from 100 mg of eggs collected from specific development stages (0–3, 3–6 and 6–9 hours after egg laying), ovary and carcass (whole beetle without ovary) using Trizol (Invitrogen), according to the manufacturer's instructions. Three independent biological replicates were used for each assay. First strand complementary DNA (cDNA) was synthesized from 2 µg of RNA using Superscript III reverse transcriptase (Invitrogen) and oligo(dT). The cDNA was used as template for real time qRT-PCR analysis using SYBR green based detection. qRT-PCR reactions were carried out in triplicate, and melting curves were examined to ensure single products. Results were quantified using the “delta-delta Ct” method and normalized to *rps3* transcript levels, as previously described [78]. Primer sequences used during the study are provided at the supplemental data (S5 Table).

### *zld* pRNAi in the hemiptera *Rhodnius prolixus* (*Rp*)

*zld* cDNA sequence was initially identified by BLAST and included in Fig 1. Parental RNAi against *Rp-zld* was performed as previously described [46].

### Supporting information

**S1 Fig. Schematic drawings of *Tc-zelda* non-overlapping dsRNA constructs.** *Tc-zld* gene corresponds to the Beetlebase ID: TC014798. The following primer pairs were used: C1-Forward ggccgcggAGCGCATCTTCTCCCTATCA and C1-Reverse cccggggcGCCGTTTTGT CGTTTCTCAT. This primer pair amplifies leads to the amplification of a fragment of 528 bp (349–857). A second primer pair C2-Forward ggccgcggACGACGAGTACCGCTTGACT and C2-Reverse—cccggggcCTTACCACAGGTGTCGAGA was also used and lead to similar results. This primer pair amplifies a fragment of 476 bp (1823–2279) covering the four zinc-finger domains. The lowercase letters contain the primer sequence used as a template for a second PCR using universal primers which adds a T7 promoter at both sides of the PCR template as previously described [81].

(TIF)

**S1 Table. *Zelda* homologs in several Pancrustacea species.**

(XLSX)



**S2 Table.** Gene IDs of 3250 genes with the motif GTAGGTAY in the 2kb upstream regions of *T. castaneum* genes protein-coding genes.

(XLS)

**S3 Table.** List of 141 *Dm-zld* targets during the MZT with one-to-one orthologs among *Tc-zld* predicted targets.

(XLS)

**S4 Table.** DAVID functional analysis and enriched ontology terms of the genes listed in [S3 Table](#).

(XLS)

**S5 Table.** Primers used in quantitative qRT-PCR experiments.

(XLSX)

**S1 Movie.** A representative nGFP embryo injected with *dsneo* dsRNA (see [Methods](#) for a detailed description).

(MP4)

**S2 Movie.** A representative nGFP embryo injected with *Tc-zld* dsRNA (see [Methods](#) for a detailed description).

(MP4)

## Acknowledgments

RNdaF would like to thank Prof. Michalis Averof for providing the nuclear GFP line and Prof. Siegfried Roth (Univ. Cologne, Germany) for the support during the beginning of this project.

## Author Contributions

**Conceptualization:** Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Data curation:** Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Formal analysis:** Lupis Ribeiro, Vitória Tobias-Santos, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Funding acquisition:** L. Aravind, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Investigation:** Lupis Ribeiro, Vitória Tobias-Santos, Daniele Santos, Felipe Antunes, Geórgia Feltran, Jackson de Souza Menezes, L. Aravind, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Methodology:** Lupis Ribeiro, Daniele Santos, Felipe Antunes, Geórgia Feltran, Jackson de Souza Menezes, L. Aravind, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Project administration:** Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Resources:** Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Software:** Lupis Ribeiro, L. Aravind, Thiago M. Venancio.

**Supervision:** Jackson de Souza Menezes, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Validation:** Lupis Ribeiro, Vitória Tobias-Santos, Daniele Santos, Felipe Antunes, Geórgia Feltran, Jackson de Souza Menezes, L. Aravind, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Visualization:** Lupis Ribeiro, Vitória Tobias-Santos, Daniele Santos, Geórgia Feltran, Jackson de Souza Menezes, L. Aravind, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Writing – original draft:** Lupis Ribeiro, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

**Writing – review & editing:** Lupis Ribeiro, L. Aravind, Thiago M. Venancio, Rodrigo Nunes da Fonseca.

## References

1. Nowick K, Stubbs L. Lineage-specific transcription factors and the evolution of gene regulatory networks. *Briefings in functional genomics*. 2010; 9(1):65–78. <https://doi.org/10.1093/bfpg/elp056> PMID: 20081217
2. Aravind L, Anantharaman V, Venancio TM. Apprehending multicellularity: regulatory networks, genomics, and evolution. *Birth Defects Research Part C, Embryo Today: Reviews*. 2009; 87(2):143–64. <https://doi.org/10.1002/bdrc.20153> PMID: 19530132
3. Kvon EZ, Kazmar T, Stampfel G, Yáñez-Cuna JO, Pagani M, Schernhuber K, et al. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature*. 2014; 512(7512):91–5. <https://doi.org/10.1038/nature13395> PMID: 24896182
4. Blythe SA, Wieschaus EF. Coordinating Cell Cycle Remodeling with Transcriptional Activation at the *Drosophila* MBT. *Current Topics in Developmental Biology*. 2015; 113:113–48. <https://doi.org/10.1016/bs.ctdb.2015.06.002> PMID: 26358872
5. Tadros W, Lipshitz HD. The maternal-to-zygotic transition: a play in two acts. *Development*. 2009; 136(18):3033–42. <https://doi.org/10.1242/dev.033183> PMID: 19700615
6. Liang H-L, Nien C-Y, Liu H-Y, Metzstein MM, Kirov N, Rushlow C. The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*. *Nature*. 2008; 456(7220):400–3. <https://doi.org/10.1038/nature07388> PMID: 18931655
7. De Renzis S, Elemento O, Tavazoie S, Wieschaus EF. Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo. *PLoS Biology*. 2007; 5(5):e117. <https://doi.org/10.1371/journal.pbio.0050117> PMID: 17456005
8. Harrison MM, Li X-Y, Kaplan T, Botchan MR, Eisen MB. Zelda binding in the early *Drosophila melanogaster* embryo marks regions subsequently activated at the maternal-to-zygotic transition. *PLoS Genetics*. 2011; 7(10):e1002266. <https://doi.org/10.1371/journal.pgen.1002266> PMID: 22028662
9. ten Bosch JR, Benavides JA, Cline TW. The TAGteam DNA motif controls the timing of *Drosophila* preblastoderm transcription. *Development*. 2006; 133(10):1967–77. <https://doi.org/10.1242/dev.02373> PMID: 16624855
10. Giannios P, Tsitilou SG. The embryonic transcription factor Zelda of *Drosophila melanogaster* is also expressed in larvae and may regulate developmentally important genes. *Biochemical and Biophysical Research Communications*. 2013; 438(2):329–33. <https://doi.org/10.1016/j.bbrc.2013.07.071> PMID: 23891688
11. Staudt N, Fellert S, Chung H-R, Jäckle H, Vorbrüggen G. Mutations of the *Drosophila* zinc finger-encoding gene *vielfältig* impair mitotic cell divisions and cause improper chromosome segregation. *Molecular Biology of the Cell*. 2006; 17(5):2356–65. <https://doi.org/10.1091/mbc.E05-11-1056> PMID: 16525017
12. Lynch JA, El-Sherif E, Brown SJ. Comparisons of the embryonic development of *Drosophila*, *Nasonia*, and *Tribolium*. *Wiley interdisciplinary reviews Developmental biology*. 2012; 1(1):16–39. <https://doi.org/10.1002/wdev.3> PMID: 23801665
13. Davis GK, Patel NH. Short, long, and beyond: molecular and embryological approaches to insect segmentation. *Annual Review of Entomology*. 2002; 47:669–99. <https://doi.org/10.1146/annurev.ento.47.091201.145251> PMID: 11729088
14. *Tribolium* Genome Sequencing C, Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, et al. The genome of the model beetle and pest *Tribolium castaneum*. *Nature*. 2008; 452(7190):949–55. <https://doi.org/10.1038/nature06784> PMID: 18362917
15. Bucher G, Scholten J, Klingler M. Parental *rnai* in *tribolium* (coleoptera). *Current Biology*. 2002; 12(3):R85–6. PMID: 11839285
16. Schinko JB, Hillebrand K, Bucher G. Heat shock-mediated misexpression of genes in the beetle *Tribolium castaneum*. *Development Genes and Evolution*. 2012; 222(5):287–98. <https://doi.org/10.1007/s00427-012-0412-x> PMID: 22890852

17. Schinko JB, Weber M, Viktorinova I, Kiupakis A, Averof M, Klingler M, et al. Functionality of the GAL4/UAS system in *Tribolium* requires the use of endogenous core promoters. *BMC Developmental Biology*. 2010; 10:53. <https://doi.org/10.1186/1471-213X-10-53> PMID: 20482875
18. Benton MA, Akam M, Pavlopoulos A. Cell and tissue dynamics during *Tribolium* embryogenesis revealed by versatile fluorescence labeling approaches. *Development*. 2013; 140(15):3210–20. <https://doi.org/10.1242/dev.096271> PMID: 23861059
19. Sarrazin AF, Peel AD, Averof M. A segmentation clock with two-segment periodicity in insects. *Science*. 2012; 336(6079):338–41. <https://doi.org/10.1126/science.1218256> PMID: 22403177
20. Panfilio KA, Oberhofer G, Roth S. High plasticity in epithelial morphogenesis during insect dorsal closure. *Biology open*. 2013; 2(11):1108–18. <https://doi.org/10.1242/bio.20136072> PMID: 24244847
21. Kittelmann S, Ulrich J, Posnien N, Bucher G. Changes in anterior head patterning underlie the evolution of long germ embryogenesis. *Developmental Biology*. 2013; 374(1):174–84. <https://doi.org/10.1016/j.ydbio.2012.11.026> PMID: 23201022
22. Handel K, Grunfelder CG, Roth S, Sander K. *Tribolium* embryogenesis: a SEM study of cell shapes and movements from blastoderm to serosal closure. *Dev Genes Evol*. 2000; 210(4):167–79. <https://doi.org/10.1007/s004270050301> PMID: 11180819.
23. Vidal NM, Graziotin AL, Iyer LM, Aravind L, Venancio TM. Transcription factors, chromatin proteins and the diversification of Hemiptera. *Insect Biochemistry and Molecular Biology*. 2016; 69:1–13. <https://doi.org/10.1016/j.ibmb.2015.07.001> PMID: 26226651
24. Hamm DC, Bondra ER, Harrison MM. Transcriptional activation is a conserved feature of the early embryonic factor Zelda that requires a cluster of four zinc fingers for DNA binding and a low-complexity activation domain. *The Journal of Biological Chemistry*. 2015; 290(6):3508–18. <https://doi.org/10.1074/jbc.M114.602292> PMID: 25538246
25. Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res*. 2015; 43(W1):W389–94. <https://doi.org/10.1093/nar/gkv332> PMID: 25883141;
26. Aravind L, Iyer LM. The HARE-HTH and associated domains: novel modules in the coordination of epigenetic DNA and protein modifications. *Cell Cycle*. 2012; 11(1):119–31. <https://doi.org/10.4161/cc.11.1.18475> PMID: 22186017
27. Balaji S, Iyer LM, Aravind L. HPC2 and ubinuclein define a novel family of histone chaperones conserved throughout eukaryotes. *Molecular Biosystems*. 2009; 5(3):269–75. <https://doi.org/10.1039/b816424j> PMID: 19225618
28. Pereira F, Duarte-Pereira S, Silva RM, da Costa LT, Pereira-Castro I. Evolution of the NET (Noca, Nlz, Elbow, TLP-1) protein family in metazoans: insights from expression data and phylogenetic analysis. *Scientific reports*. 2016; 6:38383. <https://doi.org/10.1038/srep38383> PMID: 27929068
29. Nien C-Y, Liang H-L, Butcher S, Sun Y, Fu S, Gocha T, et al. Temporal coordination of gene networks by Zelda in the early *Drosophila* embryo. *PLoS Genetics*. 2011; 7(10):e1002339. <https://doi.org/10.1371/journal.pgen.1002339> PMID: 22028675
30. Biedler JK, Hu W, Tae H, Tu Z. Identification of early zygotic genes in the yellow fever mosquito *Aedes aegypti* and discovery of a motif involved in early zygotic genome activation. *Plos One*. 2012; 7(3): e33933. <https://doi.org/10.1371/journal.pone.0033933> PMID: 22457801
31. Brown P, Baxter L, Hickman R, Beynon J, Moore JD, Ott S. MEME-LaB: motif analysis in clusters. *Bioinformatics*. 2013; 29(13):1696–7. <https://doi.org/10.1093/bioinformatics/btt248> PMID: 23681125
32. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*. 2009; 4(1):44–57. <https://doi.org/10.1038/nprot.2008.211> PMID: 19131956
33. Nestorov P, Battke F, Levesque MP, Gerberding M. The maternal transcriptome of the crustacean *Parhyale hawaiiensis* is inherited asymmetrically to invariant cell lineages of the ectoderm and mesoderm. *Plos One*. 2013; 8(2):e56049. <https://doi.org/10.1371/journal.pone.0056049> PMID: 23418507
34. Wang J, Garrey J, Davis RE. Transcription in pronuclei and one- to four-cell embryos drives early development in a nematode. *Current Biology*. 2014; 24(2):124–33. <https://doi.org/10.1016/j.cub.2013.11.045> PMID: 24374308
35. El-Sherif E, Averof M, Brown SJ. A segmentation clock operating in blastoderm and germband stages of *Tribolium* development. *Development*. 2012; 139(23):4341–6. <https://doi.org/10.1242/dev.085126> PMID: 23095886
36. van der Zee M, Berns N, Roth S. Distinct functions of the *Tribolium* *zerknüllt* genes in serosa specification and dorsal closure. *Current Biology*. 2005; 15(7):624–36. <https://doi.org/10.1016/j.cub.2005.02.057> PMID: 15823534

37. Savard J, Marques-Souza H, Aranda M, Tautz D. A segmentation gene in tribolium produces a polycistronic mRNA that codes for multiple conserved peptides. *Cell*. 2006; 126(3):559–69. <https://doi.org/10.1016/j.cell.2006.05.053> PMID: 16901788
38. Nunes da Fonseca R, von Levetzow C, Kalscheuer P, Basal A, van der Zee M, Roth S. Self-regulatory circuits in dorsoventral axis formation of the short-germ beetle *Tribolium castaneum*. *Developmental Cell*. 2008; 14(4):605–15. <https://doi.org/10.1016/j.devcel.2008.02.011> PMID: 18410735
39. Chen G, Handel K, Roth S. The maternal NF- $\kappa$ B/dorsal gradient of *Tribolium castaneum*: dynamics of early dorsoventral patterning in a short-germ beetle. *Development*. 2000; 127(23):5145–56. PMID: 11060240
40. Choe CP, Miller SC, Brown SJ. A pair-rule gene circuit defines segments sequentially in the short-germ insect *Tribolium castaneum*. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(17):6560–4. <https://doi.org/10.1073/pnas.0510440103> PMID: 16611732
41. Nakamoto A, Hester SD, Constantinou SJ, Blaine WG, Tewksbury AB, Matei MT, et al. Changing cell behaviours during beetle embryogenesis correlates with slowing of segmentation. *Nature Communications*. 2015; 6:6635. <https://doi.org/10.1038/ncomms7635> PMID: 25858515
42. Paré AC, Vichas A, Fincher CT, Mirman Z, Farrell DL, Mainieri A, et al. A positional Toll receptor code directs convergent extension in *Drosophila*. *Nature*. 2014; 515(7528):523–7. <https://doi.org/10.1038/nature13953> PMID: 25363762
43. Benton MA, Pechmann M, Frey N, Stappert D, Conrads KH, Chen Y-T, et al. Toll genes have an ancestral role in axis elongation. *Current Biology*. 2016; 26(12):1609–15. <https://doi.org/10.1016/j.cub.2016.04.055> PMID: 27212406
44. Brown SJ, Mahaffey JP, Lorenzen MD, Denell RE, Mahaffey JW. Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. *Evolution & Development*. 1999; 1(1):11–5.
45. Lavore A, Pagola L, Esponda-Behrens N, Rivera-Pomar R. The gap gene giant of *Rhodnius prolixus* is maternally expressed and required for proper head and abdomen formation. *Developmental Biology*. 2012; 361(1):147–55. <https://doi.org/10.1016/j.ydbio.2011.06.038> PMID: 21763688
46. Berni M, Fontenele MR, Tobias-Santos V, Caceres-Rodrigues A, Mury FB, Vionette-do-Amaral R, et al. Toll signals regulate dorsal-ventral patterning and anterior-posterior placement of the embryo in the hemipteran *Rhodnius prolixus*. *EvoDevo*. 2014; 5:38. <https://doi.org/10.1186/2041-9139-5-38> PMID: 25908955
47. Pires CV, Freitas FCdP, Cristino AS, Dearden PK, Simões ZLP. Transcriptome Analysis of Honeybee (*Apis Mellifera*) Haploid and Diploid Embryos Reveals Early Zygotic Transcription during Cleavage. *Plos One*. 2016; 11(1):e0146447. <https://doi.org/10.1371/journal.pone.0146447> PMID: 26751956
48. Arsala D, Lynch JA. Ploidy has little effect on timing early embryonic events in the haplo-diploid wasp *Nasonia*. *Genesis*. 2017; 55(5). <https://doi.org/10.1002/dvg.23029> PMID: 28432826;
49. Tomoyasu Y, Wheeler SR, Denell RE. Ultrathorax is required for membranous wing identity in the beetle *Tribolium castaneum*. *Nature*. 2005; 433(7026):643–7. <https://doi.org/10.1038/nature03272> PMID: 15703749
50. Tomoyasu Y, Denell RE. Larval RNAi in *Tribolium* (Coleoptera) for analyzing adult development. *Development Genes and Evolution*. 2004; 214(11):575–8. <https://doi.org/10.1007/s00427-004-0434-0> PMID: 15365833
51. Tomoyasu Y, Arakane Y, Kramer KJ, Denell RE. Repeated co-options of exoskeleton formation during wing-to-elytron evolution in beetles. *Current Biology*. 2009; 19(24):2057–65. <https://doi.org/10.1016/j.cub.2009.11.014> PMID: 20005109
52. Ravisankar P, Lai Y-T, Sambrani N, Tomoyasu Y. Comparative developmental analysis of *Drosophila* and *Tribolium* reveals conserved and diverged roles of *abrupt* in insect wing evolution. *Developmental Biology*. 2016; 409(2):518–29. <https://doi.org/10.1016/j.ydbio.2015.12.006> PMID: 26687509
53. Panganiban G, Nagy L, Carroll SB. The role of the *Distal-less* gene in the development and evolution of insect limbs. *Current Biology*. 1994; 4(8):671–5. PMID: 7953552
54. Beermann A, Jay DG, Beeman RW, Hülskamp M, Tautz D, Jürgens G. The *Short antennae* gene of *Tribolium* is required for limb development and encodes the orthologue of the *Drosophila Distal-less* protein. *Development*. 2001; 128(2):287–97. PMID: 11124123
55. Angelini DR, Kikuchi M, Jockusch EL. Genetic patterning in the adult capitate antenna of the beetle *Tribolium castaneum*. *Developmental Biology*. 2009; 327(1):240–51. <https://doi.org/10.1016/j.ydbio.2008.10.047> PMID: 19059230
56. Angelini DR, Kaufman TC. Insect appendages and comparative ontogenetics. *Developmental Biology*. 2005; 286(1):57–77. <https://doi.org/10.1016/j.ydbio.2005.07.006> PMID: 16112665

57. Boxshall GA. The evolution of arthropod limbs. *Biological Reviews of the Cambridge Philosophical Society*. 2004; 79(2):253–300. PMID: [15191225](https://pubmed.ncbi.nlm.nih.gov/15191225/)
58. Angelini DR, Smith FW, Jockusch EL. Extent With Modification: Leg Patterning in the Beetle *Tribolium castaneum* and the Evolution of Serial Homologs. *G3 (Bethesda, Md)*. 2012; 2(2):235–48. <https://doi.org/10.1534/g3.111.001537> PMID: [22384402](https://pubmed.ncbi.nlm.nih.gov/22384402/)
59. Suzuki Y, Squires DC, Riddiford LM. Larval leg integrity is maintained by Distal-less and is required for proper timing of metamorphosis in the flour beetle, *Tribolium castaneum*. *Developmental Biology*. 2009; 326(1):60–7. <https://doi.org/10.1016/j.ydbio.2008.10.022> PMID: [19022238](https://pubmed.ncbi.nlm.nih.gov/19022238/)
60. Hug CB, Grimaldi AG, Kruse K, Vaquerizas JM. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell*. 2017; 169(2):216–28 e19. <https://doi.org/10.1016/j.cell.2017.03.024> PMID: [28388407](https://pubmed.ncbi.nlm.nih.gov/28388407/)
61. Li X-Y, Harrison MM, Villalta JE, Kaplan T, Eisen MB. Establishment of regions of genomic activity during the *Drosophila* maternal to zygotic transition. *eLife*. 2014; 3. <https://doi.org/10.7554/eLife.03737> PMID: [25313869](https://pubmed.ncbi.nlm.nih.gov/25313869/)
62. Xu Z, Chen H, Ling J, Yu D, Struffi P, Small S. Impacts of the ubiquitous factor Zelda on Bicoid-dependent DNA binding and transcription in *Drosophila*. *Genes & Development*. 2014; 28(6):608–21. <https://doi.org/10.1101/gad.234534.113> PMID: [24637116](https://pubmed.ncbi.nlm.nih.gov/24637116/)
63. Foo SM, Sun Y, Lim B, Ziukaite R, O'Brien K, Nien C-Y, et al. Zelda potentiates morphogen activity by increasing chromatin accessibility. *Current Biology*. 2014; 24(12):1341–6. <https://doi.org/10.1016/j.cub.2014.04.032> PMID: [24909324](https://pubmed.ncbi.nlm.nih.gov/24909324/)
64. Schulz KN, Bondra ER, Moshe A, Villalta JE, Lieb JD, Kaplan T, et al. Zelda is differentially required for chromatin accessibility, transcription factor binding, and gene expression in the early *Drosophila* embryo. *Genome Res*. 2015; 25(11):1715–26. <https://doi.org/10.1101/gr.192682.115> PMID: [26335634](https://pubmed.ncbi.nlm.nih.gov/26335634/);
65. Sun Y, Nien CY, Chen K, Liu HY, Johnston J, Zeitlinger J, et al. Zelda overcomes the high intrinsic nucleosome barrier at enhancers during *Drosophila* zygotic genome activation. *Genome Res*. 2015; 25(11):1703–14. <https://doi.org/10.1101/gr.192542.115> PMID: [26335633](https://pubmed.ncbi.nlm.nih.gov/26335633/);
66. Jacobs CG, Rezende GL, Lamers GE, van der Zee M. The extraembryonic serosa protects the insect egg against desiccation. *Proc Biol Sci*. 2013; 280(1764):20131082. <https://doi.org/10.1098/rspb.2013.1082> PMID: [23782888](https://pubmed.ncbi.nlm.nih.gov/23782888/);
67. Nicholson DB, Ross AJ, Mayhew PJ. Fossil evidence for key innovations in the evolution of insect diversity. *Proc Biol Sci*. 2014; 281(1793). Epub 2014/08/29. <https://doi.org/10.1098/rspb.2014.1823> PMID: [25165766](https://pubmed.ncbi.nlm.nih.gov/25165766/);
68. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997; 25(17):3389–402. PMID: [9254694](https://pubmed.ncbi.nlm.nih.gov/9254694/);
69. Finn RD, Coghill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*. 2016; 44(D1):D279–85. <https://doi.org/10.1093/nar/gkv1344> PMID: [26673716](https://pubmed.ncbi.nlm.nih.gov/26673716/)
70. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res*. 2016; 44(D1):D67–72. <https://doi.org/10.1093/nar/gkv1276> PMID: [26590407](https://pubmed.ncbi.nlm.nih.gov/26590407/);
71. Giraldo-Calderon GI, Emrich SJ, MacCallum RM, Maslen G, Dialynas E, Topalis P, et al. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. *Nucleic Acids Res*. 2015; 43(Database issue):D707–13. <https://doi.org/10.1093/nar/gku1117> PMID: [25510499](https://pubmed.ncbi.nlm.nih.gov/25510499/);
72. Bailey TL. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics*. 2011; 27(12):1653–9. <https://doi.org/10.1093/bioinformatics/btr261> PMID: [21543442](https://pubmed.ncbi.nlm.nih.gov/21543442/)
73. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research*. 2009; 37(Web Server issue):W202–8. <https://doi.org/10.1093/nar/gkp335> PMID: [19458158](https://pubmed.ncbi.nlm.nih.gov/19458158/)
74. Zhu LJ, Christensen RG, Kazemian M, Hull CJ, Enuameh MS, Basciotta MD, et al. FlyFactorSurvey: a database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res*. 2011; 39(Database issue):D111–7. <https://doi.org/10.1093/nar/gkq858> PMID: [21097781](https://pubmed.ncbi.nlm.nih.gov/21097781/);
75. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biology*. 2007; 8(2):R24. <https://doi.org/10.1186/gb-2007-8-2-r24> PMID: [17324271](https://pubmed.ncbi.nlm.nih.gov/17324271/)
76. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics*. 2011; 27(7):1017–8. <https://doi.org/10.1093/bioinformatics/btr064> PMID: [21330290](https://pubmed.ncbi.nlm.nih.gov/21330290/)

77. Kinsella RJ, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database: the Journal of Biological Databases and Curation*. 2011; 2011:bar030. <https://doi.org/10.1093/database/bar030> PMID: [21785142](https://pubmed.ncbi.nlm.nih.gov/21785142/)
78. Lord JC, Hartzler K, Toutges M, Oppert B. Evaluation of quantitative PCR reference genes for gene expression studies in *Tribolium castaneum* after fungal challenge. *Journal of Microbiological Methods*. 2010; 80(2):219–21. <https://doi.org/10.1016/j.mimet.2009.12.007> PMID: [20026205](https://pubmed.ncbi.nlm.nih.gov/20026205/)
79. Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, et al. Phylogenomics resolves the timing and pattern of insect evolution. *Science*. 2014; 346(6210):763–7. <https://doi.org/10.1126/science.1257570> PMID: [25378627](https://pubmed.ncbi.nlm.nih.gov/25378627/)
80. Minakuchi C, Namiki T, Shinoda T. Krüppel homolog 1, an early juvenile hormone-response gene downstream of Methoprene-tolerant, mediates its anti-metamorphic action in the red flour beetle *Tribolium castaneum*. *Developmental Biology*. 2009; 325(2):341–50. <https://doi.org/10.1016/j.ydbio.2008.10.016> PMID: [19013451](https://pubmed.ncbi.nlm.nih.gov/19013451/)
81. Lynch JA, Peel AD, Drechsler A, Averof M, Roth S. EGF signaling and the origin of axial polarity among the insects. *Current Biology*. 2010; 20(11):1042–7. <https://doi.org/10.1016/j.cub.2010.04.023> PMID: [20471269](https://pubmed.ncbi.nlm.nih.gov/20471269/)