Universidade de São Paulo

Faculdade de Medicina de Ribeirão Preto

Programa de Pós-Graduação em Genética

Franciene Rabiço Oliveira

De ovo a embrião: Análise dos genes expressos nas primeiras fases embrionárias de *Apis mellifera*

Ribeirão Preto - SP

2019

Franciene Rabiço Oliveira

De ovo a embrião: Análise dos genes expressos nas primeiras fases embrionárias de *Apis mellifera*

Versão Original

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto (FMRP) da Universidade de São Paulo (USP) para obtenção do título de Mestre em Ciências.

Área de concentração Genética

Orientadora: Profa. Dra. Zilá Luz Paulino Simões

Ribeirão Preto - SP

2019

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Rabico, Franciene Oliveira

De ovo a embrião: Análise dos genes expressos nas primeiras fases embrionárias de *Apis mellifera*– Ribeirão Preto, 2019. 100p.: il. 30 cm.

Dissertação de Mestrado apresentada à Faculdade de Medicina de Ribeirão Preto/USP.
Área de concentração: Genética.

Orientadora: Simões, Zilá Luz Paulino

1. embriogênese 2. *zelda* 3. *smaug* 4. Desenvolvimento 5. *Apis mellifera*.

Folha de Aprovação

Nome: Franciene Rabiço Oliveira

Título da Dissertação: De ovo a embrião: Análise dos genes expressos nas primeiras fases embrionárias de *Apis mellifera*

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto (FMRP) da Universidade de São Paulo (USP) para obtenção do título de Mestre em Ciências (Área de Concentração: Genética).

Aprovado em:

Banca Examinadora

Prof. Dr._____Instituição: _____

Julgamento:_____Assinatura: _____

Prof. Dr._____Instituição: _____

Julgamento:_____Assinatura: _____

Prof. Dr._____Instituição: _____

Julgamento:_____Assinatura: _____

Ribeirão Preto,____de_____de

Agradecimentos

À Profa**.** Dra. Zilá Luz Paulino Simões por ter me acolhido no laboratório desde a Iniciação Científica. Pela orientação, sempre com muito zelo, atenção, compartilhamento de conhecimento, encorajamento para vencer as etapas do projeto, paciência e pelo exemplo de como ser uma excelente professora e pesquisadora.

À Profa. Dra. Márcia Maria Gentile Bitondi por todo o conhecimento científico compartilhado, sugestões e incentivos durante a realização desse trabalho. Também agradeço por, juntamente com a Profa. Dra. Zilá, ter despertado em mim, ainda durante a graduação, o meu interesse e amor pelas abelhas e pela embriologia.

Ao técnico Luiz Roberto Aguiar pelo auxílio na coleta das amostras e no manuseio das colônias de *Apis mellifera.*

À técnica Vera Lúcia Castelo Figueiredo pela constante transmissão de experiências das práticas laboratoriais e protocolos.

À Profa. Dra. Flávia Cristina de Paula Freitas pela ajuda nas fases iniciais do projeto e também pelo conhecimento compartilhado ao longo dos anos.

Aos colegas de laboratório Dr. Fabiano Carlos Pinto de Abreu, MSc Fábio Oliveira Barbin, MSc. Luana Bataglia, MSc Lucas Trevisoli, MSc. Thiago da Silva Depintor, Patrick Souza, Fernanda Dalarmi, Fernanda Sanches, Leonardo Nascimento de Paula e Rodrigo Leite de Barros Souza, por estarem sempre prontos a ajudar, tanto no dia-a-dia das práticas de laboratório, quanto nas trocas de ideias teóricas.

À MSc. Danielle Luna Lucena pela amizade, companheirismo em vários “projetos paralelos”, pela ajuda nas análises e também pelo encorajamento diante de problemas.

Às minhas amigas de graduação Amanda Rosanna Alves Barboza, Giovana Craveiro e Priscila Cristina Corrêa, por todos os anos de amizade e companheirismo. Que por, também serem pós-graduandas, entendem perfeitamente as inseguranças, os medos e as dúvidas que fazem parte dessa fase. Obrigada por toda a ajuda e pelos momentos de distração.

À funcionária Dona Maria pelo bom humor, alegria e torcida e, também, por manter sempre as salas limpas.

Ao curso de Pós-Graduação em Genética (FMRP/USP) e a todos os seus docentes e

funcionários pela oportunidade e apoio.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela bolsa concedida (Processo nº 2017/1894-7).

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Processos nº 2016/ 06657-0 e 2019/ 02347-1), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

Ao meu pai Francisco José Resende Oliveira por todo apoio e incentivo durante o mestrado.

À minha mãe Fátima Rabiço Oliveira, obrigada por não me deixar desistir, por ser sempre meu porto seguro.

À Deus, por tudo, sempre.

RESUMO

De ovo a embrião: Análise dos genes expressos nas primeiras fases embrionárias de *Apis mellifera*

O início do desenvolvimento embrionário é único e possui características próprias, onde cada componente é programado para atuar em um local e hora precisamente definidos. Em sua fase inicial, o embrião não apresenta transcrição zigótica, sendo mantido por componentes moleculares depositados pela mãe durante a oogênese. A ativação transcricional do genoma zigótico é coordenada com a degradação dos mRNAs de origem materna, resultando em uma mudança no perfil transcricional do embrião. Assim, durante essa transição, denominada materno-zigótica (MZT), o zigoto estabelece o controle de seu desenvolvimento. A degradação dos produtos maternos ocorre principalmente pela ação conjunta de Smaug e Zelda. Em *Drosophila melanogaster*, a transcrição zigótica aumenta gradualmente do ciclo 2 ao ciclo 13 do desenvolvimento embrionário. *Zelda* é um gene com papel essencial no início da embriogênese. A homologia, a sequência deste gene e os domínios funcionais são altamente conservados na ordem Hymenoptera, bem como em outras espécies investigadas. Experimntos piloto detectaram a expressão *zelda* logo nos primeiros ciclos do desenvolvimento embrionário de *Apis mellifera*, o que motivou a investigação de outros genes igualmente importantes nos passos iniciais do desenvolvimento embrionário. A partir de bibliotecas RNAseq selecionamos *RpL32, troponina C tipo IIb, takeout, engrailed even-skipped, hairy, kruppel e odd-skipped* para análise de expressão apartir de ovócito (tempo zero, 0), 2 e 6 horas de desenvolvimento. Para isto foram desenhados primers para as duas formas de transcritos, pré-splicing e pós-splicing, como forma de validar a transcrição pelo embrião. Os resultados apontam para uma atividade de síntese de novos transcritos na primeira hora de desenvolvimento embrionário. Para complementar os achados e verificar os ciclos mitóticos daquele período, os núcleos de embriões diplóides com 1 hora de desenvolvimento após a postura foram marcados com DAPI, a fim de verificar os ciclos mitóticos desse período. Observamos, com 1 hora de desenvolvimento, três núcleos de clivagem, indicando que eles passaram por dois ciclos mitóticos, sugerindo uma divisão assíncrona. Esses achados indicam fortemente que pode haver expressão do gene zigótico nos primeiros ciclos mitóticos do embrião de *A. mellifera*. Por fim, foi feita a reconstrução de redes de interação gene – proteínas em fases iniciais do desenvolvimento embrionário. Os resultados apontam que vários genes e proteínas podem estar envolvidos nas fases iniciais da ativação do zigoto. Acreditamos que esses resultados certamente ajudarão a esclarecer os primeiros passos da embriogênese nesse organismo modelo, abrindo novos e promissores campos para estudos nessa área.

Palavras chave: *Apis melífera*, embriogênese, transição materno-zigótica, *smaug*, zelda.

ABSTRACT

From egg to embryo: Analysis of genes expressed in early embryonic stages of *Apis mellifera*

The onset of embryonic development is unique and has its own characteristics, where each component is programmed to act in a precisely defined place and time. In its initial phase, the embryo has no zygotic transcription, being maintained by molecular components deposited by the mother during oogenesis. The transcriptional activation of the zygotic genome is coordinated with the degradation of maternally derived mRNAs, resulting in a change in the embryo transcriptional profile. Thus, during this transition, called materno-zygote (MZT), the zygote establishes control of its development. The degradation of maternal products occurs mainly by the joint action of Smaug and Zelda. In *Drosophila melanogaster*, zygotic transcription gradually increases from cycle 2 to cycle 13 of embryonic development. Zelda is a gene that plays an essential role in early embryogenesis. The homology, sequence of this gene and functional domains are highly conserved in the order *Hymenoptera*, as well as in other investigated species. Pilot experiments detected *zelda* expression early in the early stages of *Apis mellifera* embryonic development, which prompted the investigation of other equally important genes in the early stages of embryonic development. From RNAseq libraries we selected *RpL32, tropoinin C type IIb, takeout, engrailed even-skipped, hairy, kruppel and odd-skipped* for expression analysis from oocyte (time zero, 0), 2 and 6 hours of development. For this purpose, primers were designed for both transcript forms, pre-splicing and post-splicing, as a way to validate transcription by the embryo. The results point to a synthesis activity of new transcripts in the first hour of embryonic development. To complement the findings and verify the mitotic cycles of that period, the nuclei of diploid embryos with 1 hour of development after laying were marked with DAPI in order to verify the mitotic cycles of this period. We observed at 1 hour of development, three cleavage nuclei, indicating that they went through two mitotic cycles, suggesting an asynchronous division. These findings strongly indicate that there may be zygotic gene expression in the early mitotic cycles of the *A. mellifera* embryo. Finally, the reconstruction of gene - protein interaction networks in early embryonic development was performed. The results indicate that several genes and proteins may be involved in the early stages of zygote activation. We believe that these results will certainly help to clarify the first steps of embryogenesis in this model organism, opening new and promising fields for studies in this area.

Key words: *Apis mellifera*, embryogenesis, genome activation, *smaug*, *zelda.*

SUMÁRIO

1- INTRODUÇÃO.....14
1.1- Caracterização da ativação do ovo em insetos.....15
1.2- Fases iniciais da embriogênese.....17
1.3- Transcrição zigótica no início da embriogênese.....19
1.4- Genes expressos precocemente nos embriões.....23
1.5- Resultados preliminares.....24
2- OBJETIVO.....27
2.1- Objetivo geral.....27
2.2- Objetivos específicos.....27
3- MATERIAL E MÉTODOS.....28
3.1- Descrição geral do gene *zelda*.....28
3.2- Material biológico.....29
3.2.1- Obtenção de ovócitos maduros e de embriões com idades controladas.....29
3.3- Análise de expressão gênica por qPCR.....30
3.3.1- Desenho de primers.....30
3.3.2- Extração do RNA total.....33
3.3.3- Síntese de cDNA.....34
3.3.4- Reação em cadeia da polimerase (PCR).....34
3.3.5- PCR quantitativa (qPCR).....35
3.3.6- Análise estatística dos dados de expressão quantitativa.....36
3.3.7- Sequenciamento das regiões amplificadas.....36
3.4- Marcação de embriões diploides com DAPI.....37
3.5- Construção de Redes Gênicas.....37
3.5.1- Construção da rede inicial.....37
3.5.2- Análise topológica da rede pelo software Cytoscape.....38
4- RESULTADOS.....39
4.1- Descrição geral do gene *zelda*.....39
4.1.1- Comparação da sequência proteica de Zelda de *A. mellifera* com demais espécies.....39
4.1.2- Árvore filogenética de Zelda.....45
4.2- Escolha de genes candidatos para análise por qPCR.....56
4.3- Análise da expressão gênica de *takeout, troponin C e RpL32* em embriões diploides e haploides no intervalo de 0- 6 horas após a postura.....56

4.4- Análise da expressão gênica de *engrailed, even-skipped, hairy, Kruppel e odd-skipped* em embriões diploides e haploides no intervalo de 0- 6 horas após a postura.....63

4.5- Marcação com DAPI dos núcleos de embriões diploides com 1 hora.....73

4.6- Análise Topológica de Rede.....75

4.6.1- Redes das proteínas even skipped (eve) e Kruppel (Kr).....75

4.6.2- Análise das interações diretas de eve e Kr.....79

5- CONCLUSÃO.....89

6- REFERÊNCIAS BIBLIOGRÁFICAS.....90

ANEXO A.....97

LISTA DE FIGURAS

1- Esquema representativo dos primeiros estágios do desenvolvimento de embrião de *D. melanogaster*.....19
2- Os níveis de *zelda* (*zld*) aumentam aproximadamente 1 hora após a fertilização..... 20
3- Ligação de Smaug ao mRNA materno de *nanos,* na região de elementos de reconhecimento de Smaug (SRE), localizado na região 3'UTR.....22
4- O fator de transcrição Zelda modula a transcrição dos RNAs zigóticos.....22
5- Perfil de expressão obtido por qPCR dos transcritos de *zelda* em ovócito maduro e embriões na fase de clivagem de *A. mellifera*.....25
6- Perfil de expressão de *smaug* em ovócitos maduros e embriões haploides e diploides em bibliotecas de mRNA.....25
7- Esquema do gene *smaug,* que é composto por 8 éxons, em um total 1839 pb.....32
8- Esquema do gene *zelda,* que é composto por um único éxon de 1093 pb.....33
9- Alinhamento de 19 sequências de espécies de insetos.....44
10- Árvore filogenética do gene *zelda*.....45
11- Árvore filogenética de 10 espécies de abelhas corbiculadas.....47
12- Árvore filogenética de 10 espécies de abelhas.....47
13- Árvore filogenética das principais ordens de insetos.....48
14- Análise em gel de agarose da amplificação dos genes candidatos, por PCR convencional, usando os *primers* POS (Pós-processamento) e PRS (Pré-processamento) 57
15- Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário diploide de *A. mellifera* (E1/2 a E6 horas).....61
16- Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário haploide de *A. mellifera* (E1/2 a E6 horas).....62

17- Gel de agarose 1% para análise dos fragmentos de RNA amplificados usando-se *primers* POS (Pós-processamento) e PRS (Pré-processamento) 63
18- Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário diploide de *A. mellifera* (E1/2 a E6 horas).....69
19- Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário haploide de *A. mellifera* (E1/2 a E6 horas).....71
20- Marcação com DAPI de embriões diploides de *A. mellifera* com 1 hora de desenvolvimento após a postura.....74
21- Rede topológica da proteína even-skipped (eve).....77
22- Rede topológica da proteína Kruppel (Kr).....78
23- Rede topológica dos vizinhos diretos de eve.....79
24- Rede topológica dos vizinhos diretos de Kr.....84

LISTA DE TABELAS

1- Material biológico utilizado para os experimentos de qPCR e marcação com DAPI.....30
2- *Primers* propostos construídos pelo programa *Primer3*.....31
3- Genes selecionados com base na biblioteca de mRNA, para validação por qPCR.....51
4- Processos biológicos dos genes selecionados para as análises.....54
5- Proteínas homólogas as vizinhas de eve de *D. melanogaster* encontras em *A. mellifera*.....80
6- Processos biológicos da rede de vizinhos diretos de eve.....82
7- Proteínas homólogas as vizinhas de kr de *D. melanogaster* encontras em *A. mellifera*.....85
8- Processos biológicos da rede de vizinhos diretos de kr.....87

1- INTRODUÇÃO

Como já dizia o poeta João Cabral de Melo Neto "...o ovo, e apesar / da pura forma concluída, / não se situa no final: / está no ponto de partida". Mesmo que pareça simples e óbvia essa afirmação, a singularidade que é um ovo, ou embrião, têm despertado o interesse da humanidade desde a antiguidade, quando Hipócrates (cerca de 460-377 a.C.) e Aristóteles (cerca de 384-322 a.C.) faziam observações sobre os gametas masculino e feminino.

Desde da antiguidade, inúmeros progressos foram feitos na área da embriologia e no estudo dos zigotos e embriões de diferentes espécies. Porém, ainda há muito o que se estudar nessa temática e um recorte bem interessante sobre essa fase são os momentos iniciais do desenvolvimento embrionário. O momento da passagem de ovócito para embrião, ou seja, o início do desenvolvimento embrionário, um momento único e dependente da finalização da meiose. Além disto, as divisões celulares durante este período têm regras especiais e são dependentes do material biológico depositado e organizado pela mãe durante o processo de ovogênese. Assim, não é difícil concluir que o ovo é uma célula especial, com características próprias, onde cada componente está programado para atuar em local e momento criteriosamente programados. Em muitos organismos estudados, as primeiras divisões celulares ou nucleares, após a finalização do processo de meiose, se caracterizam por uma velocidade rápida e exclusiva desta fase do desenvolvimento, com pouco ou nenhum aumento no volume do citoplasma. As primeiras mitoses, reguladas pelos fatores maternos, são binárias (estão presentes apenas as fases S e M), as fases de crescimento G1 e G2, serão retomadas paulatinamente e constitui também uma

característica de cada embrião (Newport & Kirschner, 1982a, b). Em alguns organismos, a velocidade do ciclo de divisão é ainda mais acentuada pelo fato de que essas divisões acontecem sem citocinese, por exemplo, em *Drosophila melanogaster*, dentro de 2 horas, os núcleos já sofreram 13 ciclos mitóticos (Foe & Alberts, 1983; Shermoen *et al.,* 2010).

A medida que o desenvolvimento avança, os produtos maternos vão se extinguindo e são substituídos pelos transcritos produzidos pelo novo indivíduo, o zigoto, um processo especial denominado Transição Materno-Zigótica (MZT) (Darbo *et al.,* 2013; Siddiqui *et al.,* 2012; Tadros *et al.,* 2007, Tadros & Lipshitz, 2009). Assim, durante o desenvolvimento inicial, há uma remodelação coordenada do ciclo de divisão e do perfil transcricional do embrião.

Em modelos como *D. melanogaster, Caenorhabiditis elegans* e alguns vertebrados, os eventos iniciais do desenvolvimento são relativamente bem conhecidos. No entanto, até mesmo nestes organismos, aspectos importantes deste processo estão por serem descobertos. As abelhas *Apis mellifera* ainda são pouco estudadas e oferecem, pelo seu modo de desenvolvimento e determinação do sexo, possibilidades para o entendimento deste processo. O sistema de determinação do sexo nos *Hymenoptera*, em que a fêmea é diploide e se desenvolve a partir de um ovo fecundado e os machos de ovos haploides, ou seja, de ovos não fecundados, constitui um experimento da natureza, não havendo, portanto, necessidade de experimentos muito invasivos para estudo destas duas modalidades de desenvolvimento.

1.1- Caracterização da ativação do ovo em insetos

O processo de ativação do ovo é o passo inicial em que uma única célula, subitamente, começa a sofrer mitoses especiais e até certo ponto sincronizadas, dando origem aos diversos tipos celulares presentes no organismo adulto. A compreensão dos mecanismos iniciais da ovogênese fornece dados importantes para o entendimento de como uma única célula indiferenciada originará tecidos e órgãos morfológica e funcionalmente diversos.

Em muitos organismos, ovócitos maduros permanecem paralisados em um estágio da meiose, sendo que o momento em que ocorre esta pausa é característico de cada espécie. A retomada da meiose ocorre pela ativação do ovo (Von Stetina & Orr-Weaver, 2011). Esse processo ativa vias moleculares e eventos celulares no ovócito iniciando o processo de embriogênese. As mudanças moleculares incluem modificações na estrutura da membrana vitelina, ativação de fatores de transcrição, alterações no perfil de proteínas e nas inúmeras famílias de RNAs maternos, por meio da tradução e degradação, ou modificações pós-traducionais das proteínas já existentes (Tadros & Lipshitz, 2009, Robertson & Lin, 2013, Kronja *et al.,* 2014). Esses eventos moleculares dirigem a maioria das mudanças a nível celular, tais como, finalização da meiose, fusão dos pró-núcleos em organismos diploides e preparação para a mitose embrionária.

Nos vertebrados e invertebrados, a ativação do ovo requer um aumento no nível do cálcio intracelular (Townley et *al*., 2006; Whitaker, 2008; Sartain & Wolfner, 2013). Em muitos casos, essa liberação de cálcio é causada pela fertilização do ovo por um espermatozoide. Entretanto, em alguns casos como nos *Diptera* e *Hymenoptera,* a ativação do ovo é independente da fecundação. Além disto, em *Hymenoptera* os ovos não fertilizados se desenvolvem em

machos, ou seja, independente da fecundação. Em *A. mellifera* não é necessário ocorrer a fecundação para a ativação dos ovócitos, sendo que ovos não fecundados originam os zangões. Nesses casos, é possível que a ativação do ovo ocorra devido a mudanças no ambiente iônico, mudanças de pH ou estimulação mecânica durante sua passagem pelo oviduto, independente da fecundação, como descrito para *D. melanogaster* (Tadros & Lipshitz, 2009).

Acredita-se que a ativação do desenvolvimento em *A. mellifera* ocorra em consequência da passagem do ovo pelo trato reprodutivo materno no qual os músculos semi-circulares que envolvem o oviduto estimulariam mecanicamente a ativação do ovo e também por meio do influxo de cálcio do ambiente externo para o interior do ovo (Sasaki & Obara, 2002; Sartain & Wolfner, 2013).

O processo que culmina com a ativação do ovo durante sua passagem pelo oviduto é considerado um período de maturação do ovócito, como foi descrito para *D. melanogaster* (Hara et al., 2017).

1.2- Fases iniciais da embriogênese

Os primeiros passos da embriogênese nos metazoários são rápidos, precisos e possuem um padrão semelhante. Nos primeiros momentos após o término da meiose, ou seja, após a ativação do ovo, ocorrem mitoses rápidas, apenas com as fases S e M e com um período de intérfase curto (Newport & Kirschner, 1982a, b). Em *D. melanogaster*, nas primeiras duas ou três horas, ocorrem treze divisões nucleares sinciciais, que geram aproximadamente 6.000 núcleos ao redor do córtex do embrião. É amplamente aceito que grande parte da maquinaria que governa o início da embriogênese foi depositada no ovócito, pela mãe, durante a ovogênese e são independentes das informações presentes

no genoma do zigoto. Essa afirmação foi aceita e difundida durante um longo período e era apoiada nos seguintes fatores: os primeiros ciclos nucleares ocorrem em menos de dez minutos, fazendo com que a expressão gênica do zigoto seja, aparentemente, pouco expressiva; os estudos iniciais sobre o tema não conseguiam detectar transcrição antes do oitavo ciclo nuclear; foram identificadas muitos produtos maternos que estão presentes nos ovários e nos ovócitos em formação que são necessários, durante a embriogênese, para dar manutenção as primeiras divisões nucleares dos embriões, minimizando as necessidade de novos produtos sintetizados pelo zigoto na fase inicial (Pritchard & Schubiger, 1996, ten Bosh *et al.,* 2006 Edgar & Schubiger, 1986). Porém, com o avanço das técnicas moleculares, novas evidências sobre a atividade precoce do genoma do zigótico estão surgindo.

Os primeiros ciclos mitóticos ocorrem aproximadamente a um terço da região do polo anterior do embrião. Durante os três primeiros ciclos mitóticos, os núcleos possuem um arranjo levemente esférico, à medida que se expandem para a região posterior. Nos ciclos 4 a 6, os núcleos se expandem ao longo do eixo do embrião, formando um elipsoide, próximo à periferia. Nas divisões seguintes, até o ciclo 10 os núcleos se organizam no citoplasma cortical, caracterizando o primeiro estágio de blastoderme sincicial (ciclo 10). A partir do ciclo 14, se inicia o processo de celularização (blastoderma celular) (figura 1). Ao longo destas fases iniciais, os ciclos nucleares estão bem sincronizados, o espaçamento entre os núcleos é regular e os movimentos nucleares são semelhantes em quase todos os embriões de *D. melanogaster* (Ji et al., 2004, Foe & Alberts, 1983). As primeiras clivagens e as fases de desenvolvimento embrionário inicial são muito semelhantes entre *A. mellifera* e *D. melanogaster,*

entretanto, elas ocorrem em períodos de tempo diferentes, sendo que a clivagem de *D. melanogaster* ocorre em, aproximadamente, 1 hora e 20 minutos, enquanto que a de *A. mellifera* em 7 horas (Fleig & Sander 1986, Campos-Ortega & Hartenstein, 1997).

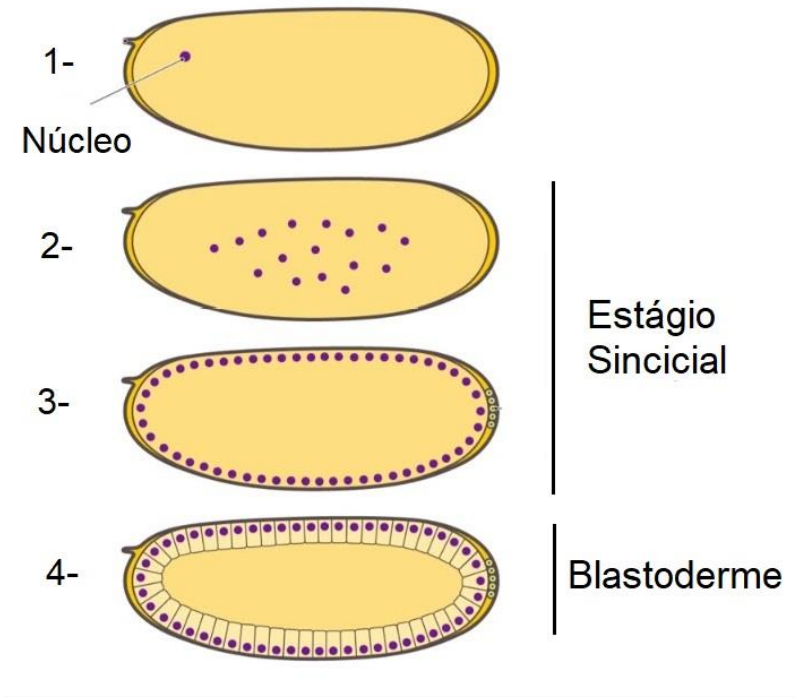


Figura 1: Esquema representativo dos primeiros estágios do desenvolvimento de embrião de *D. melanogaster*. 1- Zigóto com apenas um núcleo localizado próximo ao polo anterior. 2- Embrião entre os ciclos 4 a 6, no qual os núcleos de clivagem formam uma estrutura elipsoide no centro do embrião. 3- Embrião entre os ciclos 10 a 13, quando os núcleos se dirigem para o córtex. 4- Celularização do embrião com o estabelecimento da blastoderme celular.

1.3- Transcrição zigótica no início da embriogênese

O início do desenvolvimento embrionário está sob controle dos componentes moleculares depositados pela mãe. Esses produtos maternos são degradados concomitantemente à ativação do genoma zigótico, também por estes mesmos produtos (Schier, 2007).

Os trabalhos iniciais sobre a MZT, em *D. melanogaster*, demonstraram que a produção de RNA aumenta drasticamente à medida que o embrião se

desenvolve, desde os estágios iniciais após a postura, até a fase de blastoderme (Anderson & Lengyel, 1979, Lamb & Laird, 1976). Estudos mais recentes mostram que a ativação do genoma zigótico é um processo gradual, que resulta na expressão de fatores de transcrição e microRNAs que modelam os processos moleculares do embrião e que genes como *engrailed* (*en*), *smaug* (*smg*) e *zelda* (*zld*) possuem expressão zigótica nos primeiros ciclos embrionários (Ali-Murthy *et al.* 2013; Pires *et al.* 2016). Esses mesmos estudos apontaram para outros genes que provavelmente são expressos antes da formação da blastoderme (Figura 2).

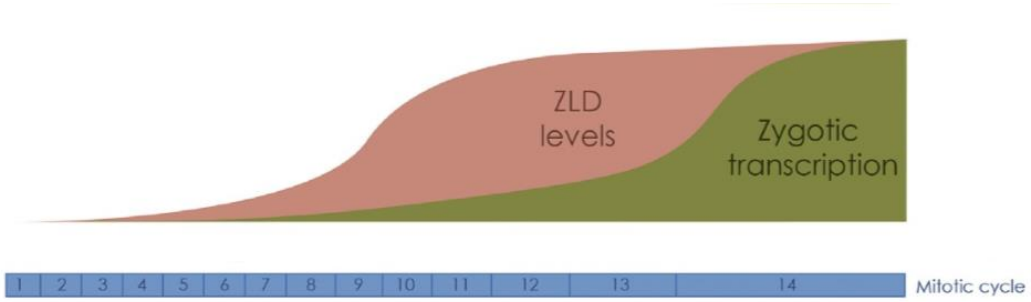


Figura 2: Os níveis de *zelda* (zld) aumentam aproximadamente 1 hora após a fertilização. Transcritos zigóticos aumentam gradualmente a partir do 2º ciclo até o 13º. No 14º ciclo, milhares de genes encontram-se trancricionalmente ativos (modificado de: Harrison & Eisen, 2015).

Geralmente, os genes que são transcritos no início do desenvolvimento são pequenos e não possuem íntrons. Dos 59 primeiros genes expressos identificados por De Renzis *e* cols*.,* 70% não tinham íntrons (De Renzis *et al.,* 2007). O comprimento reduzido destas unidades de transcrição pode ser um requisito para a expressão inicial devido às divisões nucleares rápidas neste momento em desenvolvimento. Dessa forma, como a intérfase nos embriões iniciais é muito curta, somente genes pequenos e de rápida transcrição seriam sintetizados (Shermoen & O'Farrell, 1991). Essas características dos primeiros genes zigóticos provavelmente são compartilhadas entre organismos cujo

desenvolvimento inicial é caracterizado por divisões rápidas, já que os genes zigóticos iniciais em *zebra-fish* também são, na maioria das vezes, curtos e com poucos íntrons (Heyn *et al.,* 2014). Além do tamanho menor dos genes zigóticos iniciais, há também um enriquecimento dos fatores de transcrição que são essenciais nos primeiros estágios da embriogênese, quando o genoma está sendo gradualmente ativado (De Renzis et al., 2007). Além disso, os mRNAs não são os únicos RNAs expressos durante o MZT. MicroRNAs (miRs) também são expressos (Aboobaker *et al.,* 2005; Biemar *et al.,* 2005; Fu *et al.,* 2014), e sua expressão coordena a ativação do genoma zigótico com a degradação dos mRNAs maternos (Bushati *et al.,* 2008; Lund *et al.,* 2009).

Em dípteros, a mudança da razão núcleo/citoplasma (N/C) determina o início da transcrição zigótica (Foe & Alberts, 1983; Newport & Kirschner, 1982 a,b; Pritchard & Schubiger, 1996). Esse processo parece ser conservado em insetos (*D. melanogaster*), anfíbios e *zebra-fish* (Tadros *et al.,* 2007). Entretanto, em himenópteros, a relação N/C não parece ser essencial para ativação do genoma zigótico (Arsala & Lynch, 2017), sendo mais relevante a atuação em conjunto das proteínas Smaug e Zelda. Uma vez que o ovo esteja ativo, Smaug emprega múltiplos mecanismos para desestabilizar quase 2/3 dos RNAs maternos, sendo que a remoção destes é necessária para a ativação do genoma zigótico (Smibert *et al.,* 1999; Tadros *et al.,* 2003; Tadros *et al.,*2007; Benoit *et al.,* 2009) (Figura 3). Zelda é um fator de transcrição que atua em conjunto com Smaug, se ligando a motivos denominados CAGGTAG (De Renzis, *et. al*., 2007) ou TAGteam (ten Bosch *et al.* 2006; Sung *et al.* 2013) e ativando a transcrição de milhares de genes zigóticos (Tadros *et al.,* 2007; Liang *et al.,* 2008; Benoit *et al.,* 2009; Tadros & Lipshitz, 2009; Nien *et al.,* 2011) (Figura 4). Além disso, Zelda

é um remodelador de cromatina, controlando a acessibilidade à cromatina e/ou modificando histonas na região na qual se liga (Harrison *et al.,* 2011). Esses avanços nas buscas pelos sinais das primeiras expressões zigóticas foram fundamentais para que outros genes fossem identificados como essenciais para o desenvolvimento inicial.

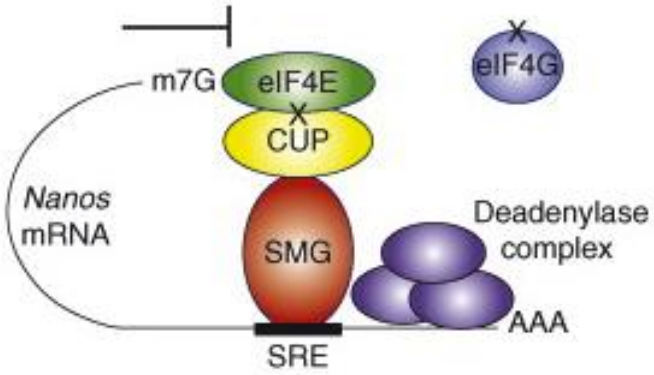


Figura 3: Ligação de Smaug ao mRNA materno *nanos,* na região de elementos de reconhecimento de Smaug (SRE), localizado na região 3'UTR. Além de Smaug, a proteína CUP também é recrutada, e essa se liga à eIF4E impedindo a ligação com eIF4G. Esse mecanismo leva a repressãoda tradução do transcrito (modificado de: Vardy & Orr-Weaver, 2007).

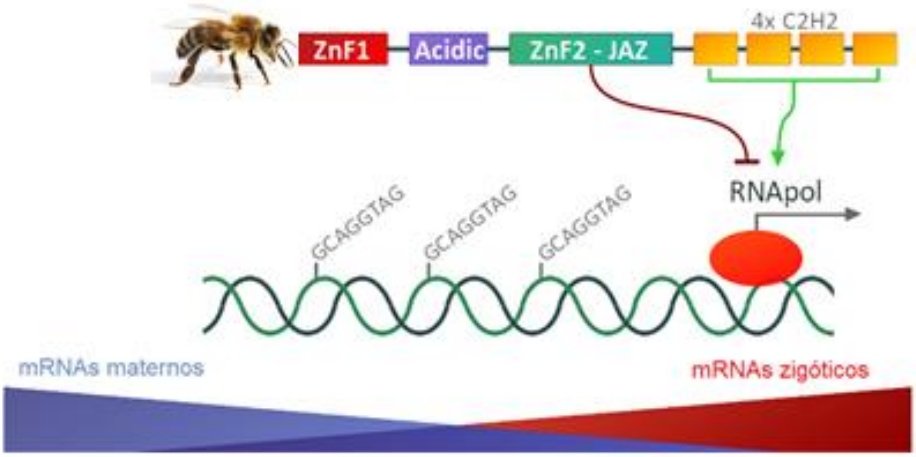


Figura 4: O fator de transcrição Zelda modula a transcrição dos RNAs zigóticos. O motivo ZnF2 – JAZ atua inibindo a transcrição zigótica, enquanto que os motivos C2H2, estimulam a transcrição de novos genes (modificado de: Fonseca & Venâncio, 2018).

1.4- Genes expressos precocemente nos embriões

Os estudos iniciais que apontavam para um genoma zigótico quiescente nos insetos eram resultados de conclusões geradas por técnicas pouco sensíveis à detecção dos transcritos zigóticos. À medida que métodos com maior sensibilidade se tornam disponíveis, tais como RNAseq e microarranjos de DNA (RNA), possibilitou identificar transcritos sintetizados pelo zigoto, sendo possível concluir que já nas primeiras horas do desenvolvimento embrionário dos metazoários a expressão gênica é variável e dinâmica (Arbeitman, 2002).

Cada vez mais, várias evidências têm apontado para a expressão de genes na fase de pré-blastoderme em *D. melanogaster.* No ciclo 7, quando o embrião possui aproximadamente 32 núcleos, foi possível verificar a presença dos genes zigóticos *engrailed* (*en*), *hairy* (*h*), *odd skipped* (*odd*), *Kruppel* (*Kr*) e *even skipped* (*eve*) (Ali-Murthy *et al.,* 2013). Além disso, foi possível verificar que o gene *engrailed* é expresso pela primeira vez no ciclo 2 e que as mutações zigóticas nesse gene dão origem a diferentes fenótipos antes da formação da blastoderme (Ali-Murthy *et al.,* 2013).

Também foi visto em *A. mellifera* que a expressão zigótica pode ter início precocemente, com a expressão de miRNAs e mRNAs, como *zld* e *smg* já nas primeiras horas após a postura, tanto nos ovos haploides quanto nos diploides (Pires *et al.,* 2016). A anotação do genoma de *A. mellifera* (Honeybee Genome Sequencing Consortium 2006) também contribuiu para que comparações entre os genomas de *A. mellifera* e *D. melanogaster* pudessem ser feitas nas fases embrionárias iniciais, porém, não foram identificados ortólogos para muitos dos genes de *D. melanogaster* ativos nesta fase da vida.

Essas variações nos dois genomas poderiam indicar que diferentes mecanismos estariam atuando na embriogênese inicial dessas duas espécies. Porém, é possível encontrar similaridades na expressão de genes e vias que são essenciais para a formação e manutenção do padrão e estrutura corpórea do embrião. Isso ocorre com os genes *Hox* (homeobox), principais responsáveis pelo controle do posicionamento e diferenciação das células nos embriões (Campos- Ortega & Hastenstein 1997), sendo que tanto em *A. mellifera*, quanto em *D. melanogaster*, se mostram bem conservados (Dearden *et al.* 2006). Além desses, podemos citar também genes envolvidos em na formação do eixo ântero-posterior embrionário, como o *eve* e o *paired* (*prd*) (Osborne & Dearden, 2005).

1.5- Resultados preliminares

Com dados prévios obtidos por Pires, 2016 e por Oliveira, 2017, foi possível observar que *zelda* e *smaug* possuem picos de expressão em embriões diploides de *A. mellifera.* Para *zelda*, esse pico foi visto em idades entre zero e uma hora após a postura, quando comparado com outros pontos da embriogênese (Figura 5). Já para *smaug*, foi observado uma maior expressão no intervalo de 0-2 horas após a postura, tanto em embriões diploides quanto em haploides (Pires *et al.,* 2016) (Figura 6). Como em *D. melanogaster* a transcrição zigótica é prematura e foi detectada já na segunda divisão nuclear (Arsala & Linch, 2017) e em *A. mellifera* também há fortes evidências de transcrição nos primeiros ciclos mitóticos (Pires et al., 2016), acredita-se que *zelda* e *smaug*, essenciais para a o início da transcrição, estejam sendo expressos pelo embrião já nos primeiros ciclos de clivagem sincicial.

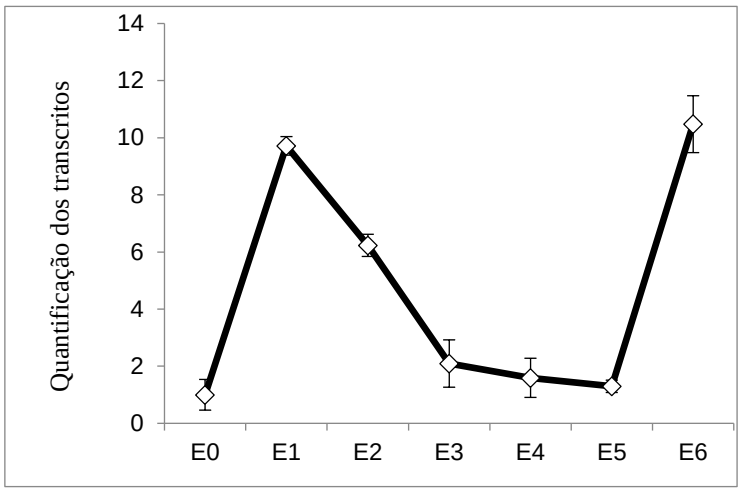


Figura 5: Perfil de expressão obtido por qPCR dos transcritos de *zelda* em ovócito maduro e embriões na fase de clivagem de *A. mellifera*. Há uma marcante variação na fase inicial do desenvolvimento, seguido de uma queda em E2, estabilizando em seguida (E3, E4, E5) e voltando a níveis altos em E6. 0h, Ovócito maduro (E0), 0-1h (E1), de 1- 2h (E2), de 2- 3h (E3), de 3- 4h (E4), de 4- 5h (E5), de 5- 6h (E6), após a postura. Cada ponto é a média de três réplicas biológicas e cada uma, um pool de 35 ovos. Barra corresponde a desvio padrão das réplicas biológicas (retirado de: Oliveira, 2017).

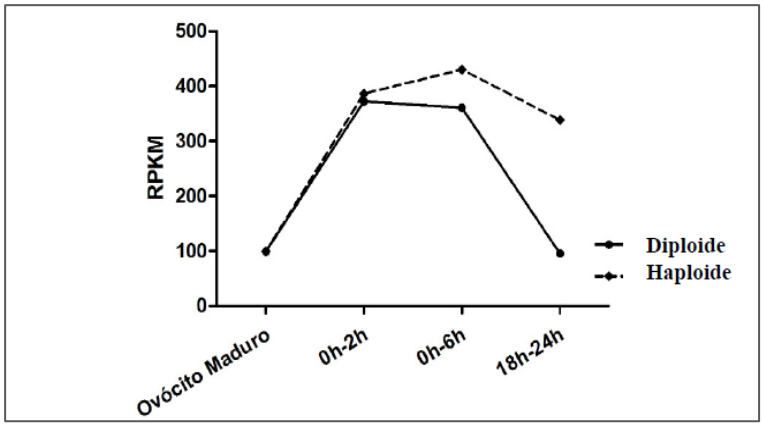


Figura 6: Perfil de expressão de *smaug* em ovócitos maduros e embriões haploides e diploides em bibliotecas de mRNA. Tanto no embrião haploide, quanto no diploide, os níveis de expressão do gene *smaug* aumenta no intervalo de 0-2h, em relação ao ponto anterior de ovócito maduro. Dados retirados da biblioteca de RNA-seq de embriões de *A. mellifera* (*accession numbers*: SRR1538449; SRR1538849; SRR1538934; SRR1538877) (modificado de: Pires, 2014).

Entretanto, há fortes evidências de que não somente esses genes estejam sendo expressos nas primeiras horas de desenvolvimento embrionário de *A. mellifera*. Outros genes como *eve*, *en*, *h*, *Kr* e *odd* podem estar sendo transcritos já nos primeiros ciclos mitóticos. Além disso, outros genes, sem ortólogos

conhecidos para *D. melanogaster* ou com pouca participação nas primeiras fases embrionárias, também foram identificados como sendo expressos na biblioteca RNAseq de embriões haploides e diploides (*accession numbers*: SRR1538449; SRR1538849; SRR1538934; SRR1538877) (Pires *et al.,* 2016). Neste estudo, foram analisados os genes *troponin C type IIb (TpnCIIb), protein takeout (to), smaug (smg)* presentes na biblioteca RNAseq.

Dessa forma, uma abordagem mais aprofundada pode ser útil na elucidação de questões como: estariam os genes zigótos de *D. melanogaster* expressos em fases iniciais também expressos nas fases iniciais de *A. mellifera*? Os genes presentes nas bibliotecas de RNA indicados como expressos em embriões de *A. mellifera* podem ser validados por análises por qPCR, e se sim, qual a possível função deles na embriogênese inicial? A análise da estrutura e das homologias do gene *zelda*, a fim de confirmar sua identidade em *A. melífera,* também são importantes. Além disso, informações sobre as iferenças quanto as expressões zigóticas iniciais nos embriões haploides e diploides em *A. mellifera* para os genes selecionados podem elucidar questões importantes. A reunião de todas estas informações em um único gráfico de interação um estudo exploratório sobre redes topológicas foi iniciado visando abrir novas possibilidades de estudos que envolvam interações entre genes e proteínas. Esses dados em conjunto podem ser úteis para melhorar e aprimorar nosso entendimento sobre os processos de ativação do genoma do zigoto em organismos em que esses processos são menos estudados, como é o caso de *A. mellifera* e, assim, contribuir para abertura de novos caminhos nesta área do conhecimento.

2- OBJETIVOS

2.1- Objetivo Geral

O presente estudo teve por objetivo analisar a expressão de genes reguladores nos embriões haploides e diploides de *A. mellifera*, durante as primeiras divisões mitóticas nucleares, que caracterizam a fase de clivagem.

2.2- Objetivos específicos

- Análise comparativa da estrutura, domínios funcionais, sequência e relações filogenéticas do gene *zelda*, a fim de validar sua identidade em *A. mellifera*.
- Análise do perfil de expressão dos pré-RNAs e mRNAs de *TpnCIIb, to* e *smg* durante fases iniciais (pré-bastodérmicas) de embriões haploides e diploides de *A. mellifera*, no intervalo de 0-6h, usando qPCR, para investigação da transcrição inicial em embriões. Estes genes foram selecionados de bibliotecas RNA-seq de embriões de *A. mellifera,* utilizando como parâmetro seus níveis de expressão.
- Análise do perfil de expressão dos pré-RNAs e mRNAs de *em, eve, h, kr* e *odd* durante fases iniciais (pré-bastodérmicas) de embriões haploides e diploides de *A. mellifera*, no intervalo de 0-6h, usando qPCR, para investigação da transcrição inicial em embriões. Estes genes foram selecionados com base em estudos anteriores em embriões de *D. melanogaster.*
- Contagem de núcleos em embriões diploides de *A. mellifera* no intervalo de 0-1h após a postura, para estimativa das divisões mitóticas neste intervalo.

- Reconstrução de redes de interação gene – proteínas em fases iniciais do desenvolvimento embrionário.

3- MATERIAL E MÉTODOS

3.1- Descrição geral do gene *zelda*

O objetivo inicial deste projeto foi motivado por resultados preliminares obtidos em experimentos piloto e análise da expressão gênica diferencial em bibliotecas de mRNA embrionário construídas por sequenciamento de última geração, que apontaram o gene *zelda* como expresso precocemente no desenvolvimento embrionário de *A. melífera.*

Para conhecer melhor a estrutura do gene *zelda,* foi feita uma descrição geral da estrutura do gene com base nas informações disponibilizadas pelos bancos de dados de *A. mellifera*, bem como, uma busca pelos motivos funcionais usando ferramentas disponíveis no NCBI (National Center for Biotechonology Information) e anotação manual, usando o programa Artemis. Em seguida, foi feita uma busca por ortólogos de *zelda* em uma ampla gama de espécies e a partir da sequência de aminoácidos presentes em *A. mellifera* e foi feito um Blastp (https://blast.ncbi.nlm.nih.gov/Blast.cgi) (Johnson *et al.,* 2008). Foram selecionadas sequências de espécies de insetos que apresentaram uma identidade ≥ 50%, com a sequência de *A. mellifera*. Com essas sequências, foi feito um alinhamento pelo software MEGA® for Windows, v.7.0.26 (https://www.megasoftware.net) (Kumar *et al.,* 2016), utilizando a ferramenta de alinhamento MUSCLE (Edgar, 2004). Em seguida, a história evolutiva foi inferida usando o método de Neighbor-Joining (Saito & Nei, 1987). As distâncias evolutivas foram calculadas usando o método de correção de Poisson (Zuckerkandl & Pauling, 1965). Os domínios funcionais foram identificados através de buscas no Banco de dados Pfam v. 32.0 (http://pfam.sanger.ac.uk/) (Finn et al., 2014) e também em Ribeiro, 2017.

3.2- Material biológico

Para a realização deste experimento foram utilizadas abelhas *A. mellifera* provenientes do apiário experimental do Departamento de Genética da Faculdade de Medicina de Ribeirão Preto (FMRP) da Universidade de São Paulo.

3.2.1 - Obtenção de ovócitos maduros e de embriões com idades controladas

Para a obtenção de ovócitos maduros, equivalentes a embriões de hora 0, uma rainha fecundada foi dissecada e os ovócitos localizados mais próximos ao oviduto foram coletados e armazenados em *TRIzol*® (Invitrogen) para as análises posteriores.

Para a obtenção de embriões com idades controladas, é coletada uma postura na qual uma rainha é confinada, por 30 e 60 minutos, sobre favos específicos para postura de zangões, para coleta de embriões haploides, ou sobre favos específicos para operárias, para coleta de embriões diploides. Após esse período, os embriões de 30 e 60 minutos são coletados e armazenados. Para obtenção de embriões mais velhos, os favos com embriões de 60 minutos foram colocados em estufa a 34ºC, envoltos em pano levemente umedecido, até completar o número de horas desejadas. As amostras com idades que variavam de 30 minutos a 6 horas foram utilizadas no procedimento de extração de RNA para qPCR (Tabela 1). Amostras com idade de 1 hora foram utilizadas no experimento de marcação dos núcleos com *4',6-Diamidine-2'phenylindole dihydrochloride* (DAPI) (Sigma-Aldrich ®).

Tabela 1: Material biológico utilizado para os experimentos de qPCR e marcação com DAPI. Os embriões coletados foram separados conforme o número de horas após a postura, com intervalos variando de 30 minutos a 6h30 minutos. N/A (não aplicável).

Tipo de material	Grupo	Tempo de Postura	Tempo de incubação na estufa	Número de embriões que compõem cada amostra (*pool*)	Número de amostras biológicas utilizadas
Ovócito maduro	E 0	N/A	N/A	10 – 15	5
Embrião com 30 minutos	E ½	30 min.	0	10 – 15	5
Embrião com 1 hora	E 1	1 hora	0	10 – 15	5
Embrião com 2 horas	E 2	1 hora	1 hora	10 – 15	5
Embrião com 3 horas	E 3	1 hora	2 horas	10 – 15	5
Embrião com 4 horas	E4	1 hora	3 horas	10 – 15	5
Embrião com 5 horas	E 5	1 hora	4 horas	10 – 15	5
Embrião com 6 horas	E 6	1 hora	5 horas	10 – 15	5

3.3- Análise de expressão gênica por qPCR

3.3.1- Desenho de primers

Os *primers* para amplificar os transcritos dos genes *smaug (smg), troponin C type IIb (TpnCIIb), protein takeout (to), engrailed* (*en*), *hairy* (*h*), *odd skipped* (*odd*), *Kruppel* (*Kr*) e *even skipped* (*eve*), além do gene normalizador *ribosomal protein L32 (RPL32),* foram desenvolvidos pelo programa Primer3 (http://fokker.wi.mit.edu/primer3/input.htm) (Tabela 2).

Tabela 2: *Primers* propostos construídos pelo programa *Primer3.* Foram desenhados dois tipos de *primers*: os *primers* POS, que amplificam transcritos após o processo de *splicing* e *primers* PRS, que amplificam transcritos que possuem íntron, ou seja, antes do *splicin*g. O *primer reverse* foi mantido o mesmo para ambos os tipos de *primers* (POS e PRS), sendo sempre desenhado em uma região de éxon.

Gene	Sequência dos Primers	Região do *primer Forward*	Tipo de *primers*	Pb amplificados
SMG	F-5' GGTATGAGAGACGTTCCATC 3' R- 5' CCCTTTTGTCACACCCTGTT 3'	Éxon	POS	135
	F- 5' CACGGAAAAGAAGAAAAGTTCG 3' R- 5' CCCTTTTGTCACACCCTGTT 3'	Íntron	PRS	169
To	F- 5' TACCCAATACACATAGACTGGGAT 3' R- 5' CCGCTCCTCTGACAGGTAAA 3'	Éxon	POS	124
	F-5'CAGAAAACGTTCAGTCAATAATAATAA 3' R- 5' CCGCTCCTCTGACAGGTAAA 3'	Íntron	PRS	160
TpnCIIb	F- 5' GAGGAAGGAAGCGGCAAAT 3' R- 5' GTTCCCCTCCTTGTCGTAAA 3'	Éxon	POS	129
	F- 5' TCTTCTTAAATCGTGCCTTTCA 3' R- 5' GTTCCCCTCCTTGTCGTAAA 3'	Íntron	PRS	145
RPL32	F- 5' ACATAATGTCAAGGAACTGGAAGT 3' R- 5' TTGAGCACGTTCAACAATGG 3'	Éxon	POS	112
	F- 5' TGAAATTTATTTTCTTTGATTTACAGG 3' R- 5' TTGAGCACGTTCAACAATGG 3'	Íntron	PRS	125
Eve	F- 5' ACGATAGTGGTGGACCTCGT 3' R- 5' GGTCCATTGAGGACAACGAT 3'	Éxon	POS	130
	F-5' CCGTTTGCATCATCATTTTGTTCA 3' R- 5' GGTCCATTGAGGACAACGAT 3'	Íntron	PRS	167
En	F- 5' TACTCGGACAGGCCATCTTC 3' R- 5' CCCTCTTCAATCTTGCCAGT 3'	Éxon	POS	133
	F-5' CGAGGGATTAATGCGATGAT 3' R- 5' CCCTCTTCAATCTTGCCAGT 3'	Íntron	PRS	137
h	F- 5' GACCACGACCACGAGAAGAT 3' R- 5' GGGTCTTCAGATCGTTTAAGCA 3'	Éxon	POS	107
	F- 5' CCAAAACCTGTGCCAACTCT 3' R- 5' GGGTCTTCAGATCGTTTAAGCA 3'	Íntron	PRS	158
Kr	F- 5' TTCACCTGTGGAGTGTGCTC 3' R- 5' ATGTGGGTCTTGAGGTGGTG 3'	Éxon	POS	143
	F- 5' ACGGATAATGCACAACATCTCCC 3' R- 5' ATGTGGGTCTTGAGGTGGTG 3'	Íntron	PRS	189
odd	F- 5' CCACACGGATGAGAGACCTT 3' R- 5' AGGCTTCTCCTTGCTGTGAA 3'	Éxon	POS	103
	F- 5' CCCTCTCCTCCCAAAAATCTCTGT 3' R- 5' AGGCTTCTCCTTGCTGTGAA 3'	Íntron	PRS	150

Foram desenhados dois tipos de pares de *primers* para cada gene. O primeiro par foi desenhado em dois éxons diferentes e detecta transcritos de mRNA que já passaram pelo processamento de RNA (Pós-processamento- POS), e o segundo par foi desenhado com um dos *primers* no éxon e o outro no íntron, cujo objetivo é amplificar pré-mRNAs que foram transcritos recentemente, antes do processo de *splicing* alternativo, com a retirada dos íntrons (Pré-processamento- PRS) (Figura 7). Dessa forma, esperamos identificar e avaliar os níveis de transcrição dos produtos maternos e os sintetizados pelo zigoto.

Para o gene *zelda*, não foi possível utilizar essa abordagem, pois o gene é composto por um único éxon (Figura 8).

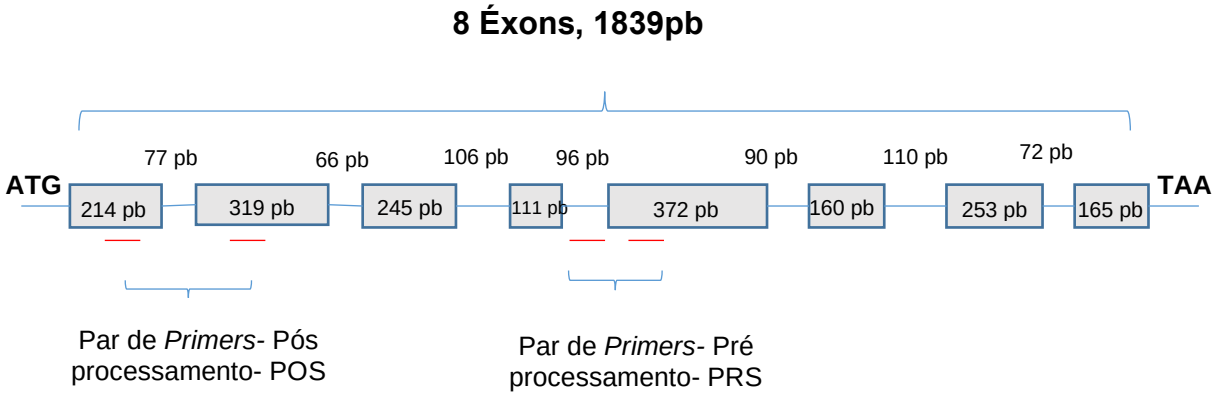


Figura 7: Esquema do gene *smaug,* que é composto por 8 éxons, num total 1839 pb. ATG- códon de início; TAA- códon de parada**.** Modelo dos dois tipos de *primers* (POS e PRS) desenhados para esse experimento estão destacados no esquema acima.

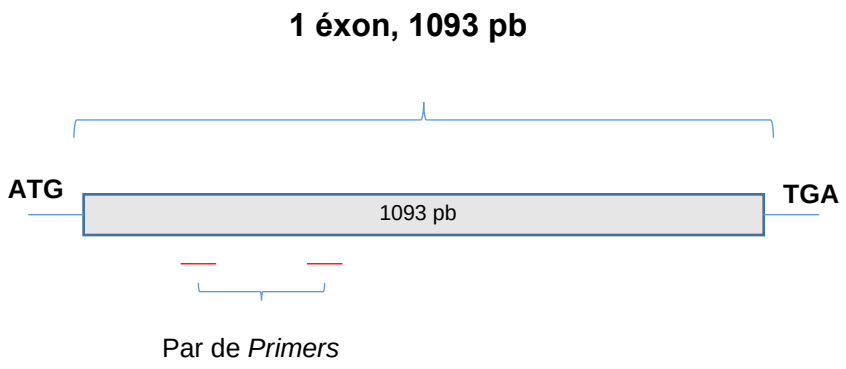


Figura 8: Esquema do gene *zelda,* que é composto por um único éxon de 1093 pb. ATG- códon de início; TGA- códon de parada. Em destaque, a região utilizada para desenhos do part de *primers*.

3.3.2- Extração do RNA total

Depois de coletados, os materiais biológicos foram colocados em 300 µL de *TRIzol*® (Invitrogen) e, em seguida, as amostras foram congeladas a -80°C até a extração do RNA total. As amostras em *TRIzol*® foram descongeladas a temperatura ambiente. A cada amostra, foram adicionados 100 µL de clorofórmio e os tubos foram agitados vigorosamente por 15 segundos e levados a uma centrifugação a 12.000 rpm (centrífuga com raio de 5,5 cm), a 4°C por 20 minutos. Ao final da centrifugação, as amostras apresentaram duas fases: uma fase branca aquosa (onde estava o RNA) e outra vermelha (contendo proteínas e o restante do *TRIzol*®). A fase branca foi transferida para outro tubo onde foram adicionados 250 µL de álcool isopropílico e 1µL de glicogênio (concentração de 10 µg/µL). As amostras foram então, armazenadas no freezer -20°C por 48 horas para que o glicogênio potencializasse a precipitação.

Após essa etapa de precipitação, as amostras foram retiradas do *freezer* e deixadas em temperatura ambiente por 40 minutos. Em seguida o material foi submetido a uma segunda centrifugação a 14.000 g a 4°C por 30 minutos. Após a centrifugação o sobrenadante foi descartado e 500 µL de etanol 75% foram

adicionados. Uma nova centrifugação foi realizada a 10.000 rpm (centrífuga com raio de 5,5 cm), a 4°C por 10 minutos para separação do RNA total. O sobrenadante foi descartado com cuidado e os tubos abertos foram colocados em um termobloco (Digital Dry Bath Incubator) a 55°C por 10 minutos para evaporação do álcool remanescente. O RNA foi, então, ressuspendido em 15-30 µL de água ultrapura tratada com dietilpirocarbonato (DEPC) e autoclavada. Para completa dissolução do RNA precipitado, as amostras foram deixadas em um termobloco a 55 ºC por 5 minutos.

A pureza (estimada por meio da razão 260/280 nm entre os valores da leitura) e concentração (dada em µg/µL) da solução final de cada amostra foram obtidas por absorbância óptica a 280 nm, em espectrofotômetro (NanoDrop® ND-1000, NanoDrop Technologies).

3.3.3- Síntese de cDNA

Para a síntese de cDNA foi utilizado o protocolo *SuperScript™ Reverse Transcriptase II* (200U/µl, Invitrogen). A cada amostra, composta de 1 a 3 µL do RNA total, foram adicionados: 1 µL de oligo (dt)12-18, 1 µL de dNTP Mix e 12 µL de água estéril e destilada. A mistura foi aquecida a 65°C por 5 minutos, seguido de rápido resfriamento em gelo. Após breve centrifugação, foram adicionados 5 µL de tampão First-Strand (5x) (250 mM Tris-HCl, 375 mM KCl; 15 mM MgCl2), e 2 µL de DTT (0,1 M). O conteúdo dos tubos foi misturado suavemente e incubado por 2 minutos a uma temperatura de 42°C. Em seguida, foi adicionado 1 µL de SuperScript® II e a mistura foi incubada por 50 minutos a 42°C. A reação foi interrompida por aquecimento a 70°C durante 15 minutos.

3.3.4- Reação em cadeia da polimerase (PCR)

A amplificação do cDNA pela técnica de PCR foi feita utilizando-se 10 µL de *Master Mix* (Promega), 8 µL de água autoclavada, 0,5 µL de cada *primer* específico (10 pmol/µL) e 1µL de cDNA. Para a separação das fitas de cDNA foi utilizada uma temperatura de 94 ºC por 30 segundos, seguido da temperatura ideal de pareamento específica de cada *primer* por 45 segundos e 72 ºC por 45 segundos. A temperatura de pareamento utilizada nas amplificações por PCR variou conforme as características de cada par de *primers*. A visualização dos fragmentos amplificados ocorreu em géis de agarose 1% corados com *SYBR® Safe* (Invitrogen) e submetidos à luz ultravioleta.

3.3.5- PCR quantitativa (qPCR)

Primers específicos foram usados para amplificar o cDNA das amostras de *A. mellifera,* usando qPCR. As amplificações foram realizadas utilizando a solução *SyBr Green®* (Thermo Fisher Scientific) em termociclador (Step One Plus™ Real Time PCR System- Applied Biosystem). A mistura da reação foi constituída de: 4,8 µL de água destilada, 7,5 µL de *SyBr Green®*, 0,6 µL de *forward primer*, 0,6 µL de *reverse primer* e 1,5 µL de cDNA.

As condições de PCR foram iguais para todos os genes testados e foi adicionado também uma curva de dissociação. Para analisar os dados de expressão gênica foi usado o método $2^{-\Delta\Delta c}T$ (Livak & Schmittgen, 2001).

Todos os *primers* utilizados foram checados quanto à sua eficiência e, para tanto, foi construída uma curva padrão utilizando diluições seriadas de cDNA (sem diluição; 1:10; 1:100; 1:1000 e 1:10000). Os valores de eficiência da PCR (E) foram calculados a partir do valor obtido de *slope* (obtido para a curva

padrão) após as reações de curva padrão e utilizando a fórmula: E= (10-1)/slope (Bulletin # 2, Applied Biosystems, 1997).

3.3.6- Análise estatística dos dados de expressão quantitativa

Para o presente estudo, foi utilizada a análise de variância *one-way* ANOVA*,* seguido pelo teste de Tukey, quando foi comparado apenas um fator. Essas análises foram feitas pelo software BioEstat 5.3 e foi considerada diferença significativa valores de $p < 0,05$ entre cada amostra.

3.3.7- Sequenciamento das regiões amplificadas

Após a extração do RNA total e síntese de cDNA e PCR, foi realizada uma purificação dos produtos de PCR, antes da reação de sequenciamento. Para a purificação da PCR, foi utilizado o *PCR purification Kit* (Cellco). Inicialmente, foram adicionados 20 µl DNA, 150 µl de *binding buffer*, 100 µl de Isopropanol. Essa solução foi transferida para uma coluna de separação dentro de um tubo de 2 ml e este conjunto foi centrifugado a 10.000 g por 30 segundos. O filtrado resultante foi descartado e foi adicionado 700 µl de *washing buffer*, seguido de centrifugação a 10.000 g por 30 segundos. Para a eluição foi utilizado 30 µl de *elution buffer,* com incubação de 1 minuto a temperatura ambiente, seguido de centrifugação a 10.000 g por 1 minuto.

Os fragmentos purificados foram sequenciados nas direções *forward* (F) e *reverse* (R) utilizando o kit *Big Dye Terminator* (Applied Biosystems) em sequenciador automático modelo *ABI 3500xL Genetic Analyzer*, disponível no Laboratório de Genética Molecular e Bioinformática (LGMB), do hemocentro FMRP/USP-RP.

Os eletroferogramas emitidos pelo sequenciador foram interpretados pelo programa Sequencing Analysis versão5.4 (Applied Biosystems) e analisados no programa CodonCode Aligner 8.0.2 for Windows. A sequência consenso resultante foi confrontada com o banco de dados GenBank (National Center for Biotechnology Information - NCBI) utilizando-se BLAST (Altschul *et al.,* 1990).

3.4 - Marcação de embriões diploides com DAPI

Embriões diploides com 1 hora após a postura foram coletados e fixados conforme os procedimentos de fixação propostos por Dearden *et al.* (2009a, 2009b). Após a fixação, o cório foi retirado e os embriões foram encubados em uma solução de diamidino-2-fenilindole (DAPI) e PTw (PBS 1x contendo 0,1% de Tween 20) (1:1) por 50 minutos. Em seguida, os embriões foram lavados 5 vezes por 5 minutos com tampão PBS. Por fim, os embriões foram transferidos para uma solução de glicerol 70% em PTw e montados em lâminas histológicas com Slow Fade® Gold, observados e fotografados em microscópio confocal (TCS SP5, Leica).

3.5- Construção de Redes Gênicas

3.5.1- Construção da rede inicial

Primeiramente, foi construída uma rede que mostra as interações proteína -proteína dos genes *eve* e Kr que estão relacionados com os processos do desenvolvimento inicial do embrião. As redes foram construídas por meio do software STRING 11.0 (https://string-db.org/cgi/input.pl), sendo utilizado *D. melanogaster* como organismo modelo para realizar as buscas, pois para *A. mellifera*, as informações genéticas são mais escassas. Com o objetivo de

produzir uma rede com pelo menos 150 nós, foram utilizadas 200 interações e um valor médio de confiança, correspondente a 0,400.

3.5.2- Análise topológica da rede pelo software Cytoscape

Os resultados obtidos no STRING foram analisados fazendo uso do software Cytoscape 3.7.1 for Windows (Shannon *et al.,* 2003).

A análise de ontologia gênica foi realizada através do banco de dados *Flybase*. Com essa análise, é possível recuperar os processos biológicos que estão relacionados com as proteínas que fazem parte das redes.

4- RESULTADOS E DISCUSSÃO

4.1- Descrição geral do gene *zelda*

4.1.1- Comparação da sequência proteica de Zelda de *A. mellifera* com as demais espécies

Durante o desenvolvimento inicial dos metazoários, os mRNAs maternos são gradualmente degradados e são substituídos por transcritos embrionários, durante a MZT (Tadros & Lipshitz, 2009). Um componente-chave nesse processo é o fator de transcrição Zelda (*zld*). Em *D. melenogaster*, *zld* apresenta 6 domínios *zinc-finger* (ZnF) bem caracterizados (Liang *et al.,* 2008). Estudos recentes mostraram que *zld* é conservado no grupo *Pancrustacea* e possui regiões especificas conservadas, como um *acidic patch* (que está relacionado com o recrutamento de proteínas remodeladoras de cromatina) e dois ZnF N-terminal (ZnF-Novel e ZnF1) (Ribeiro *et al.,* 2017). Porém, o ZnF-Novel aparece erodido em *Diptera* (Ribeiro *et al.,* 2017). Quando comparamos o alinhamento de 19 espécies de insetos (*Apis dorsata, Apis florea, Apis mellifera, Blatella germanica, Bombus impatiens, Bombus terrestris*, *Ceratina calcarata, Cephus cinctus, Dufourea* novaeangliae, *Drosophila melanogaster, Eufriesea mexicana*, *Habropoda laboriosa, Helicoverpa armigera, Megachile rotundata, Melipona quadrifasciata, Nasonia vitripennis, Osmia bicornis bicornis,Orussus abietinus, Tribolium castaneum),* sendo a maioria pertencente ao grupo *Hymenoptera*, é possível perceber que várias regiões são bem conservadas em quase todas as espécies (Figura 9). Contudo, algumas regiões relacionadas a domínios de ZnF não são conservadas em todas as espécies analisadas. Nas regiões do ZnF-Novel e do ZnF-1, esse domínio não está conservado em *H. armigera, B.*

germânica e em *D. melanogaster*. Na região *acidic patch*, essas mesmas 3 espécies e *T. castaneum*, aparentemente, estão em dissonância com as demais. Em relação ao ZnF 2, em *H. laboriosa,* uma parte do domínio não foi detectada. Essa falha provavelmente se deve a algum erro de sequenciamento e montagem do genoma, embora não possamos excluir uma alteração biológica. Os ZnF 3-6, estão presentes em todas as espécies do alinhamento, porém as espécies *H. armigera, B. germânica e D. melanogaster* novamente apresentam as maiores diferenças em relação às demais espécies do alinhamento. Nas regiões conservadas, é possível perceber que entre as espécies pertencentes ao grupo *Hymenoptera*, de forma geral, observa-se um melhor alinhamento das sequências, quando comparado com os demais.

Pelo alinhamento, foi possível verificar regiões e domínios que estão conservados em várias partes do genoma das espécies selecionadas, dando maior confiança sobre a homologia e conservação do gene *zld* em outras espécies de insetos. Para dar ainda mais suporte à homologia do gene *zld* entre essas espécies, foi construído uma árvore filogenética utilizando apenas a sequência desse gene.

ZnF Novel – 45 A 109

1 a 44 | 110 a 221

Species/Abbrv	
1. Apis mellifera	GGNSNPEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWANQA---QQEESNNNNSKAGNHNS
2. Apis cerana	GGSSNPEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWANQA---QQEESNNNNSKAGNHNS
3. Apis florea	GGSSNPEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWASQA---QQEESNNNNSKAGNHNS
4. Eufriesea mexicana	GGSSNPEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWASQA---QQEESNNNNSKAGNHNS
5. Ceratina calcarata	GGTSNPEL--HCKDCGLNFEGEKSLEVHLRY-HQENLLTQWANQA---QQEESNNNNSKAGNHNN
6. Megachile rotundata	GGSSNPEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWASQA---QQEESNNNNSKAGNHNS
7. Dufourea novaeangliae	GGSSNPEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWASQA---QQEESNNNNSKAGNHNS
8. Osmia bicornis bicornis	GGSSNPEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWASQA---QQEESNNNNSKAGNHNS
9. Habropoda laboriosa	GGSSNSEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWASQA---QQEESNNNNSKAGNHNS
10. Bombus impatiens	GGSSNSEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWASQA---QQEESNNNNSKAGNHNS
11. Bombus terrestris	GGSSNSEH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWANQA---QQEESNNNNNKAGNHIS
12. Melipona quadrifasciata	GGSSNLEH--HCKDCDVYFESDKSLEVHLRY-HHENLLSQWASQA---QQEESNNNNSKAGNHNS
13. Helicoverpa armigera	STLPSRDY--HCKICDLYLDTVDSLEVHLQY-HKENLYVKWGTQN---SQNDTDNNNGAKVKIET
14. Blattella germanica	----MEGH--HCKDCGLNFDSVKSLDVHLHY-HKENLLMKWG------QPDEANNNNAGSNNKRV
15. Drosophila melanogaster	-----PPL--HA----------AEALASSSATDS---GGGGAGGGGGGGGGGGSG
16. Cephus cinctus	GGSANAEH--HCKDCGLYFESDKSLEVHLRY-HQDNLLSQWASQA---QQEESNNNNSKAGNHNS
17. Orussus abietinus	GGAASPKH--HCKDCGLYFESDKSLEVHLRY-HQENLLSQWATQA---QQEESNNNNSKAGNHNS
18. Nasonia vitripennis	GGASSDKSSLHCKDCRMNFDSSSSLEVHLRYHHQENLMSQWASQA---SQEESNNNNSKAGNHNS
19. Tribolium castaneum	PQLNTSPH--HCKECGLYLNSAESLEVHLQY-HKENLLNKWATQAATSHSEETNNNNPKAIKREF

Região conservada 222 a 241 | 222 a 322 | Região Conservada 323 a 329 | 320 a 352 | 353 a 358 | 359 a 388 | ZnF 1- 389 a 416 | 417 a 514

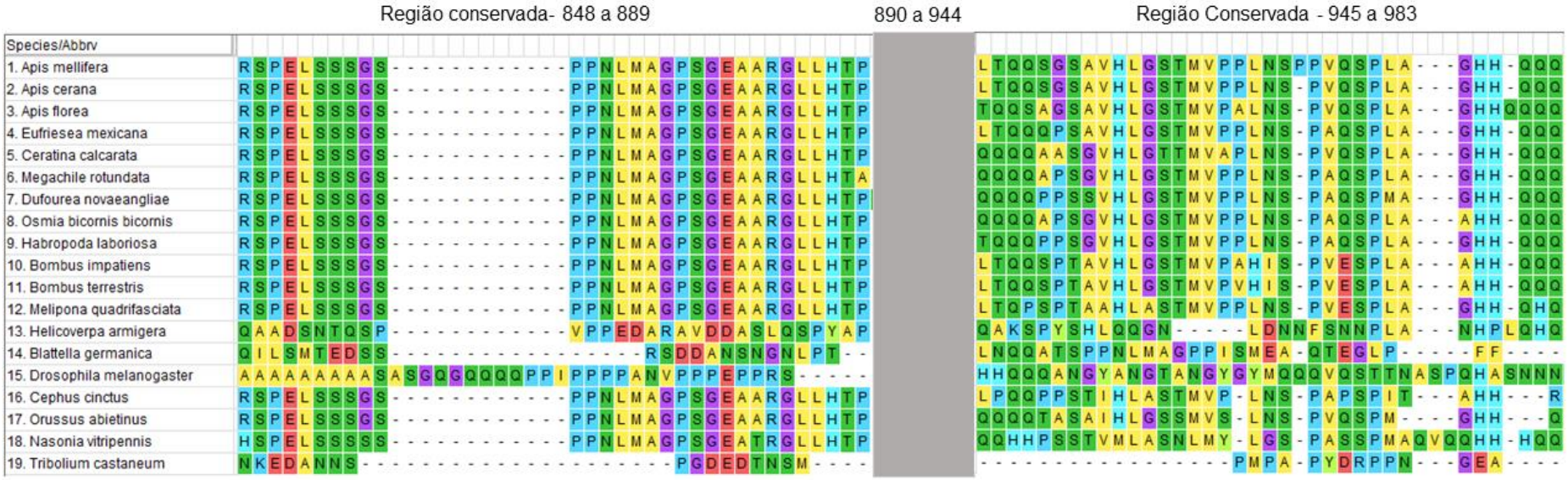
Species/Abbrv				
1. Apis mellifera	T-PMFGET--N-YFMQNDQA	PAERTSS	VSSTSP	LQCDKCGAVYEDANQLGEHVRTSHSNSP
2. Apis cerana	T-PMFGET--N-YFMQNDQA	QPERTSS	VSSTSP	LQCDKCGAVYEDANQLGEHVRTSHSNSP
3. Apis florea	T-PMFGET--N-YFMQNDQA	PPERTSS	VSSTSP	LQCDKCGAVYEDANQLGEHVRTSHSNSP
4. Eufriesea mexicana	A-PMFGET--N-YFMQNDQA	PPERTSS	VSSTSP	LQCDKCGAVYEDANQLGEHVRTNHSNSP
5. Ceratina calcarata	P-PMFSEA--S-YFMQNEPA	PTERTGS	VSSTSP	LQCEKCGAVYEDANQLGEHVRTSHTDSP
6. Megachile rotundata	T-PMFGEA--T-FFMQNDPA	PPERTSS	VSSTSP	IQCDKCGAVYEDANQLGEHVRTNHSNSP
7. Dufourea novaeangliae	T-PMFGET--N-YFMQNEPA	PPERTSS	VSSTSP	LQCDKCGAVYEDANQLGEHVRTNHSNSP
8. Osmia bicornis bicornis	A-SMFGEA--N-YFMQNEPA	PPERTSS	VSSTSP	IQCDKCGAVYEDANQLGEHVRTNHSNSP
9. Habropoda laboriosa	T-PMFGET--N-YFMQNEPA	PPERTSS	VSSTSP	IQCDKCGAVYEDANQLGEHVRTNHSNSP
10. Bombus impatiens	T-SMYGD---N-YFMQNDQA	PPERANS	VSSTSP	LQCEKCGAVYEDANQLGEHVRASHSNSP
11. Bombus terrestris	T-SMYGE---N-YFMQNDQA	PSERANS	VSSTSP	LQCEKCGAVYEDANQLGEHVRTNHSNSP
12. Melipona quadrifasciata	P-SMFGET--S-YFMQNDQA	APDRTSS	VSSTSP	LQCDKCGAVYENANQLGEHVRANHSNSP
13. Helicoverpa armigera	PGPQFGNNYSHSGFPQNQHS	MQERVSQ	VSSSSP	NQCDKCGFVCDSAVQLNEHCNSAHASTA
14. Blattella germanica	N-YLASE---S-FFMQQQQS	VPERNNG	VSPSGP	QQCDKCGFVCHSSTALLEHLNTAHPPSQ
15. Drosophila melanogaster	THHMYSAS--G-GQQQQQQQ	IPTASSA	VSSSSP	LQCKKCGILTTNESELQEHIANVHGESP
16. Cephus cinctus	T-PMFSET--S-YFMQNDQA	PPERANS	VSSTSP	LQCDKCGAVYDDANQLGEHVRSSHPTSP
17. Orussus abietinus	A-PMFSES--G-YFMQNEPP	PPERANS	VSSTSP	LQCDKCGAVYDDANQLTEHVRTSHPSSP
18. Nasonia vitripennis	T-PMFSDT--Q-YFMHNETP	PVERVNS	VSSTSP	LQCDKCGAVYEDAGQLAEHVRASHPGSP
19. Tribolium castaneum	N-SAFSPNFQN-FQLKQERT	RSQSNAQ	VSSSSP	KQCDKCGYVCDSAAQLIEHQNLAHPPTP

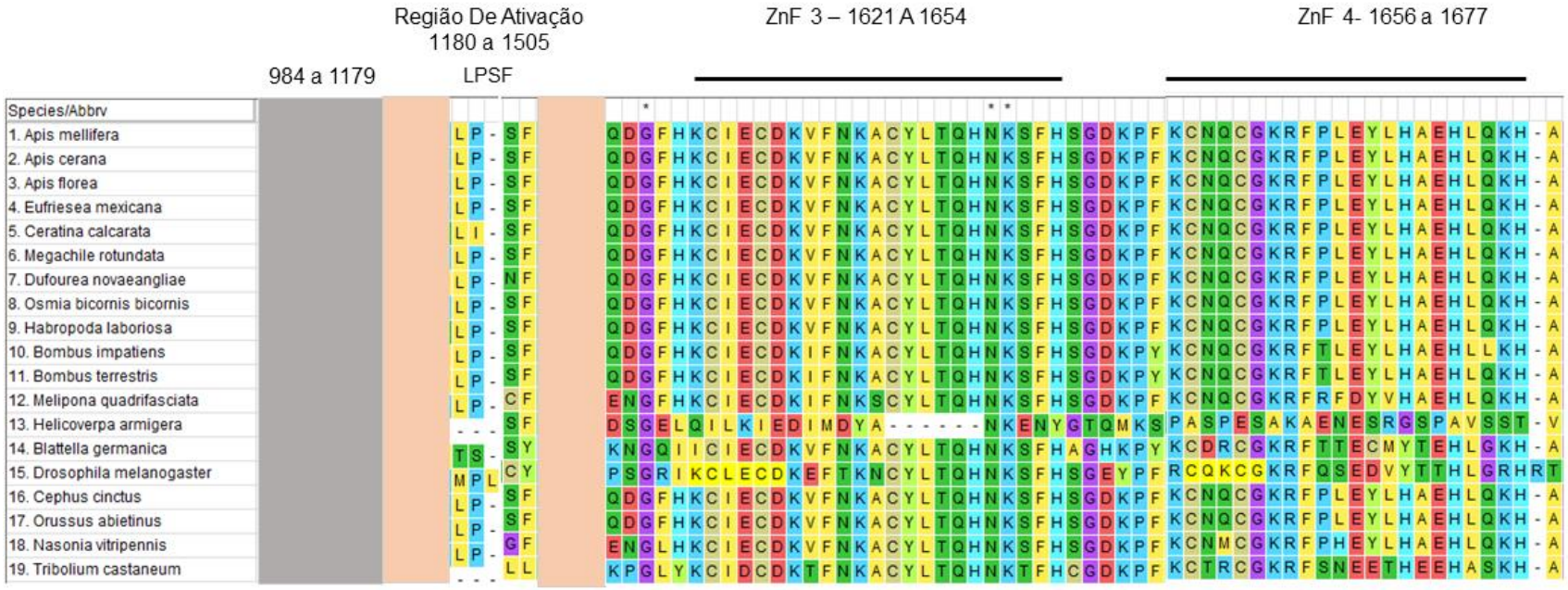
Região Conservada 515 - 537 | 538 a 601 | Região Conservada 602 a 618 | 619 a 631 | Acidic patch- 633 a 658 | 659 a 759

Species/Abbrv	Região Conservada 515 - 537	538 a 601	Região Conservada 602 a 618	619 a 631	Acidic patch- 633 a 658	659 a 759
1. Apis mellifera	VKLDVKQEPDEQAEILDLDSHKV		HQQQQRSGS--HSVSSM		LG-W-PTAA-QSHD--YH--GLSSMGP	
2. Apis cerana	VKLDVKQEPDEQAEILDLDSHKV		HQQQQRSGS--HSVSSM		LG-W-PTAA-QSHD--YH--GLSSMGP	
3. Apis florea	VKLDVKQEPDEQAEILDLDSHKV		HQQQQRSGS--HSVSSM		LG-W-PPAA-QSHD--YH--GLSSMGP	
4. Eufriesea mexicana	VKTEVKQEPEEQAEILDLDSHKV		HQQQQRAGS--HSVSSM		LG-W-PPAA-QPHD--YHP-GLSPMGP	
5. Ceratina calcarata	VKTEVKQEPEEQAEILDLDSHKV		HQQQQRTGA--HSVSSM		LG-W-PPAA-QPHD--YHP-GLSPMGP	
6. Megachile rotundata	TKTELKQEPEEQAEILDLDSHKV		HQQQQRAGS--HSVSSM		LG-W-PPAA-QPHD--YHP-GLSPMGP	
7. Dufourea novaeangliae	VKTEVKQEPEEQAEILDLDSHKV		HQQQQRAGS--HSVSSM		LG-W-PPTA-QPHD--YHP-GLSPMGP	
8. Osmia bicornis bicornis	MKTELKQEPEEQAEILDLDSHKV		HQQQQRAGS--HSVSSM		LG-W-PPAA-QPHD--YHP-GLSPMGP	
9. Habropoda laboriosa	VKTEVKQEPEEQAEILDLDSHKV		HQQQQRAGS--HSVSSM		LG-W-PPAA-QPHD--YHP-GLSPMGP	
10. Bombus impatiens	VKLDITKEPDEQAEILDLDSHKV		HQQQQRAGA--HSVSSM		LA-W-PPAA-QSHD--YHP-GLSPMIP	
11. Bombus terrestris	VKLDMTKEPDEQAEILDLDSHKV		HQQQQRAGT--HSVSSM		LG-W-PPAA-QPHD--YHP-GLSPMVS	
12. Melipona quadrifasciata	VEIEAKQEPDEQAEILDLDSHKV		HQQQQRAGT--HSVSSM		LG-W-PPAA-QQHD--YHP-GLSPMGP	
13. Helicoverpa armigera	ENAKVKEEHEESSDILDLDSQKV		YSNPQLNGD--VSLFKE		QKMFAEQKAYQTEAKMFHP-EQKFAYT	
14. Blattella germanica	--SEIKTEDKPAAEILDLDSHKV		ETNGRGSNP--HSVSGM		L--WGAVPR-QDIA--HSP-EFQTVSS	
15. Drosophila melanogaster	--PGVGSATANPGELLDLDSQKM		WEQQPQQTV--ASVEGL		VGVDKSPYY-SPKQSPYHQ-STGVGGA	
16. Cephus cinctus	IKVEIKQEPEEQAEILDLDSHKV		QQQQQRGGS--HSVSSM		LG-W-PPA--QPHE--YHPAGVSPMGP	
17. Orussus abietinus	VKSEVKQEPEEQAEILDLDSHKL		QQHQQRGGP--HS----		----------	
18. Nasonia vitripennis	IKSEAKQEPEEQAEILDLDSHKV		QQHGPRGAT--HSVSSM		LS-WPPPAS-QPHD--YHS-GLSPMGP	
19. Tribolium castaneum	--PQIKSEEDAQSEILDLDSHKV		RQNGDMNGPNPHSVTAL		LNTWSPPQQ-QQQQKMYQQQQQYMNGP	

ZnF 2 - 760 a 847

Species/Abbrv	
1. Apis mellifera	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDSHS--G-GRAGGPPP
2. Apis cerana	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDTHS--G-GRTGGPPS
3. Apis florea	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDSHS--G-GRAGGPPP
4. Eufriesea mexicana	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHK--XKQSNQPDPANMPISAHHHPGRDNHS--AGGGRGGGPS
5. Ceratina calcarata	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANMPISAHHHPGRDNHT--T-GGRGSGSS
6. Megachile rotundata	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDSHS----GGRGGGPS
7. Dufourea novaeangliae	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDNHS--GSGRGGGGPP
8. Osmia bicornis bicornis	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANMPISAHHHPGRDSHS--G-GGRGGGPS
9. Habropoda laboriosa	PP---LAN-QSWKSNEARRPKTYNCTACNK----------VKQSNQPDPANLPISAHHHPGRDSHS--V-GGRGGGPS
10. Bombus impatiens	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDSHS--S-GRGG-GPS
11. Bombus terrestris	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDSHS--S-GRGG-GAS
12. Melipona quadrifasciata	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDSHS--G-GRGG-GPS
13. Helicoverpa armigera	LPTIGGKG-ANWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVRSSGQPDPATLPISNHHHPSREMSR---------AQQ
14. Blattella germanica	QS---SGN-QSWKSNEARRPKTYNCSACNKWFTSSGHLKRHYNTTLHKNAVKQSGAPDPANQPVSNHHHPARAAEI--SPSPRRQSPA
15. Drosophila melanogaster	PANSLTAKAATWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKSSGQPDPATLPISAHHHPARDPVT--KPSRRGNAAA
16. Cephus cinctus	PP---MAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANMPISAHHHPGRDAASLGG-GGGGRGSA
17. Orussus abietinus	PP---LAN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANLPISAHHHPGRDAVS--S-SGSRAPPV
18. Nasonia vitripennis	PP---LGN-QSWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKQSNQPDPANMPISAHHHPGRDNNS--R-GAGAAASR
19. Tribolium castaneum	PV---TKS-ANWKSNEARRPKTYNCTACNKWFTSSGHLKRHYNTTLHKNAVKSSGQPDPASMPISAHHHPARDSAS----------


Região conservada- 848 a 889
890 a 944
Região Conservada - 945 a 983
Species/Abbrv
1. Apis mellifera
2. Apis cerana
3. Apis florea
4. Eufriesea mexicana
5. Ceratina calcarata
6. Megachile rotundata
7. Dufourea novaeangliae
8. Osmia bicornis bicornis
9. Habropoda laboriosa
10. Bombus impatiens
11. Bombus terrestris
12. Melipona quadrifasciata
13. Helicoverpa armigera
14. Blattella germanica
15. Drosophila melanogaster
16. Cephus cinctus
17. Orussus abietinus
18. Nasonia vitripennis
19. Tribolium castaneum


Região De Ativação
1180 a 1505
ZnF 3 – 1621 A 1654
ZnF 4- 1656 a 1677
984 a 1179
LPSF
Species/Abbrv
1. Apis mellifera
2. Apis cerana
3. Apis florea
4. Eufriesea mexicana
5. Ceratina calcarata
6. Megachile rotundata
7. Dufourea novaeangliae
8. Osmia bicornis bicornis
9. Habropoda laboriosa
10. Bombus impatiens
11. Bombus terrestris
12. Melipona quadrifasciata
13. Helicoverpa armigera
14. Blattella germanica
15. Drosophila melanogaster
16. Cephus cinctus
17. Orussus abietinus
18. Nasonia vitripennis
19. Tribolium castaneum

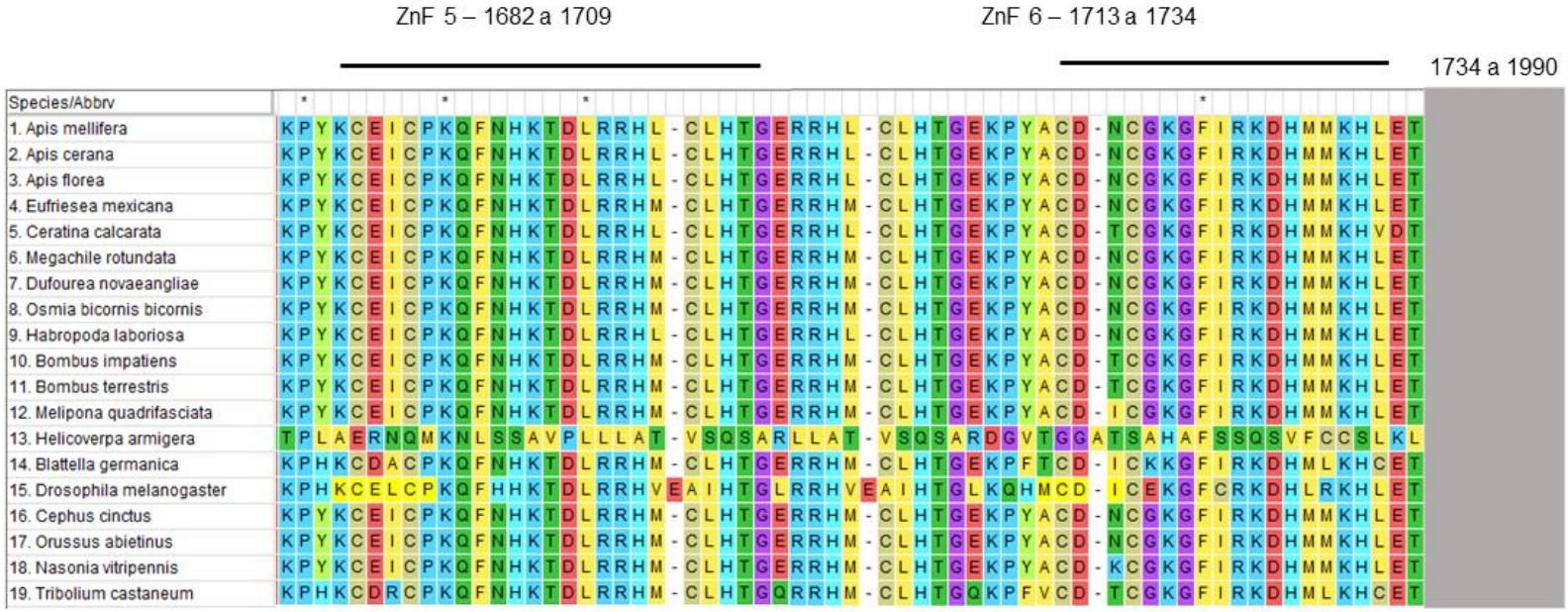
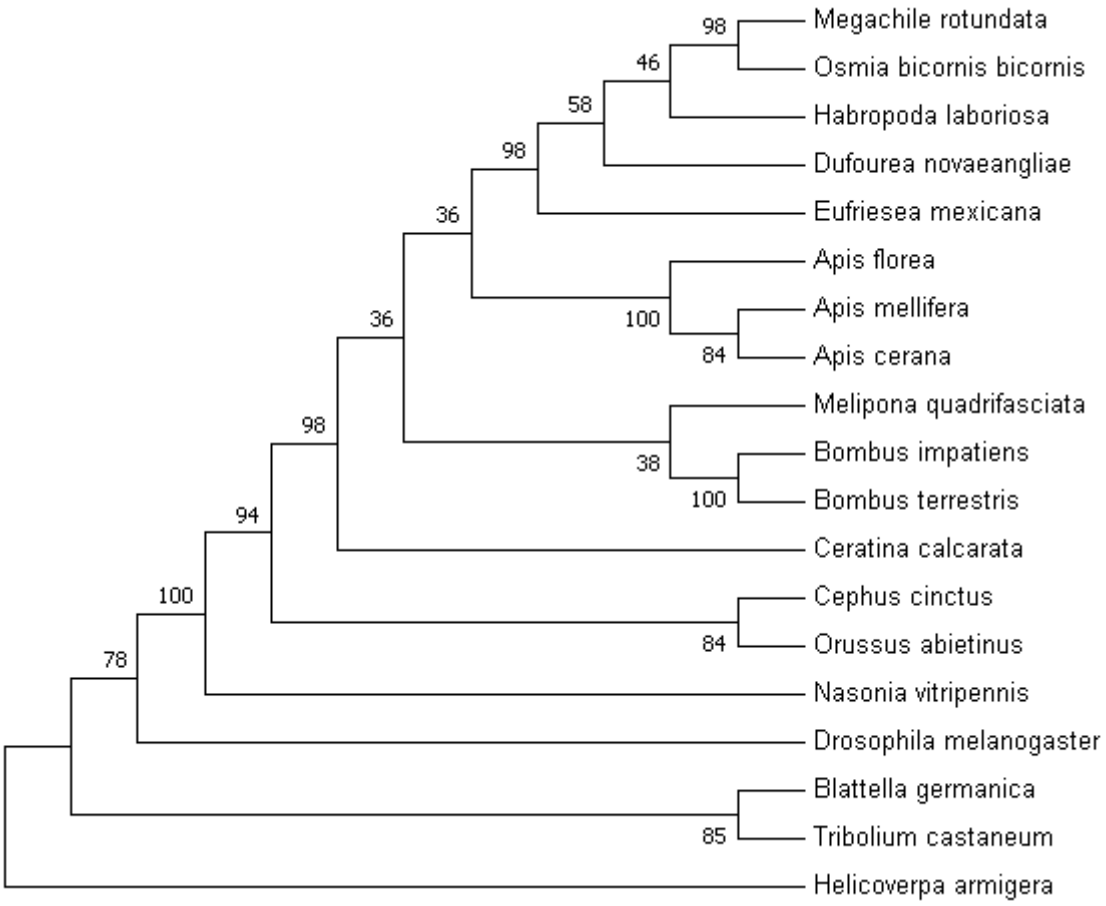


Figura 9: Alinhamento de 19 sequências de espécies de insetos. As sequências foram eleitas de acordo com a porcentagem de identidade no Blast P (NCBI), em relação a sequência de *zld* presente em *A. mellifera.* Todas as sequências analisadas no estudo possuem no mínimo 50% de identidade com a sequência de *A. mellifera.* O alinhamento foi realizado pelo software MEGA®, por meio da ferramenta MUSCLE.

4.1.2- Árvore filogenética de Zelda

A região codificadora do gene *zelda* possui um único éxon de 1093 pares de bases (pb), iniciando no *start codon* ATG (metionina) e terminando no *stop codon* TGA e está localizado no LG 12 - Bee Genome, versão 4.5. A busca pelos ortólogos de *zelda* recuperou 18 insetos com, no mínimo, 50% de identidade com a sequência de *A. mellifera*. Com esses dados, uma árvore filogenética enraizada foi construída pelo *software* MEGA® (Figura 10).

Figura 10: Árvore filogenética do gene *zelda*. Para a construção da árvore filogenética foram



consideradas as seguintes espécies: *Apis dorsata, Apis florea, Apis mellifera, Blatella germanica, Bombus impatiens, Bombus terrestris*, *Ceratina calcarata, Cephus cinctus, Dufourea* novaeangliae, *Drosophila melanogaster, Eufriesea mexicana*, *Habropoda laboriosa, Helicoverpa armigera, Megachile rotundata, Melipona quadrifasciata, Nasonia vitripennis, Osmia bicornis bicornis,Orussus abietinus, Tribolium castaneum.* Os clados formados estão de acordo com o encontrado com outros trabalhos (Kapheim *et al.,* 2015, Romiguier *et al.,* 2016, Rehan *et al.,* 2016, Ribeiro *et al.,* 2017, Wheeler *et al.,* 2001). A análise filogenética foi feita utilizando o software MEGA®, utilizando o método de Neighbor-Joining.

A reconstrução da história filogenética do gene *zelda* nas espécies selecionadas é um grande desafio, pois apenas uma pequena amostra de espécies foi selecionada, sendo a maioria pertencente aos *Hymenoptera*. Contudo, mesmo que possua alguns nós controversos, uma árvore filogenética de um gene nos permite enxergar as relações entre a sequência estudada e compara-la a filogenias presentes em outros estudos.

Alguns representantes do grupo das abelhas *Corbiculata* foram incluídos na análise do gene *zelda* (*M. rotundata, H. laboriosa, E. mexicana, A. florea, A. mellifera, A. cerana, M. quadrifasciata, B. impatiens, B. terrestris, C. calcarata*). Os *Corbiculata* incluem as abelhas que apresentam uma eussocialidade complexa, que é caracterizado por colônias grandes e perenes e castas diferenciadas morfologicamente, com comportamentos especializados, como ocorre nas castas de operária e rainha. Esse grupo é amplamente estudado para entender a evolução da eussocialidade, que é considerada uma das maiores transições na evolução (Kapheim *et al.,* 2015). Com o advento das técnicas de sequenciamento de genomas, foi possível obter uma maior confiabilidade entre as relações filogenéticas das espécies.

As outras espécies de do grupo *Hymenoptera* foram representadas por abelhas solitárias (*O. bicornis, D. novaeangliae, C. cinctus*) ou por vespas (*N.vitriennis O. abietinus*). Estudos anteriores que tinham como objetivo estabelecer as relações filogenéticas de abelhas corbiculadas (Romiguier *et al.,* 2016) ou elucidar a transição do comportamento social (Kapheim *et al.,* 2015) fizeram uso dos genomas das abelhas sequenciadas para estabelecer as relações filogenéticas entre as espécies. Os resultados produzidos por esses

estudos foram utilizados para comparativo com a árvore do gene *zelda* e foram aqui trazidos com o intuito de facilitar as comparações (Figura 11 e 12).

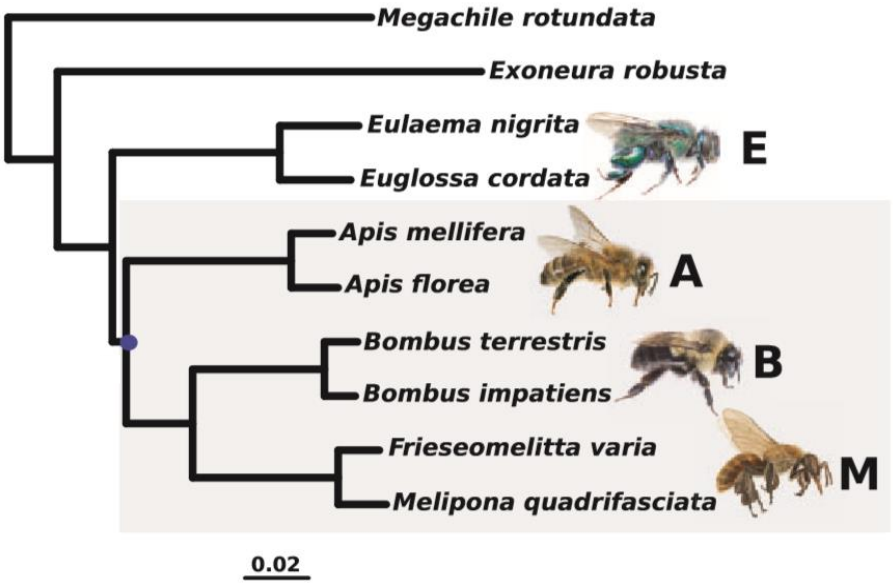


Figura 11: Árvore filogenética de 10 espécies de abelhas corbiculadas. As relações entre os clados *Apini*, *Bombini* e *Meliponini* foram utilizados como comparativo com a filogenia do gene *zelda* (modificado de Romiguier *et al.,* 2016).

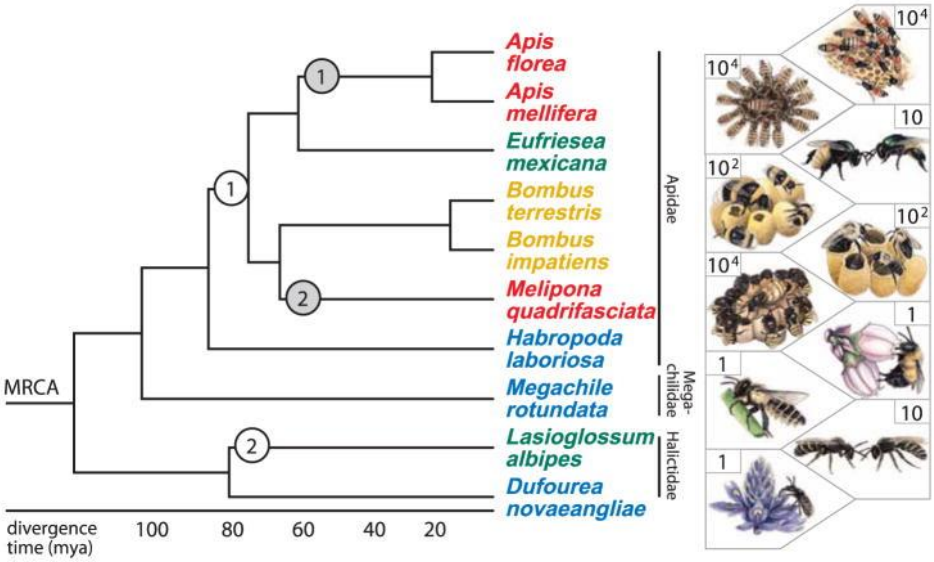


Figura 12: Árvore filogenética de 10 espécies de abelhas. Com exceção da espécie *L. albipes*, todas as demais espécies de abelhas encontram-se na filogenia do gene *zelda* (modificado de Kapheim *et al.,* 2015).

Além dos *Hymenoptera*, foram incluídos representantes dos grupos *Blattaria* (*B. germanica*), *Coleoptera* (*T. castaneum*), *Diptera* (*D. melanogaster*), *Lepdoptera* (*H. armigera*). Para as relações de parentesco entre essas espécies, foram consideradas as ordens das respectivas espécies incluídas na análise (Figura 13).

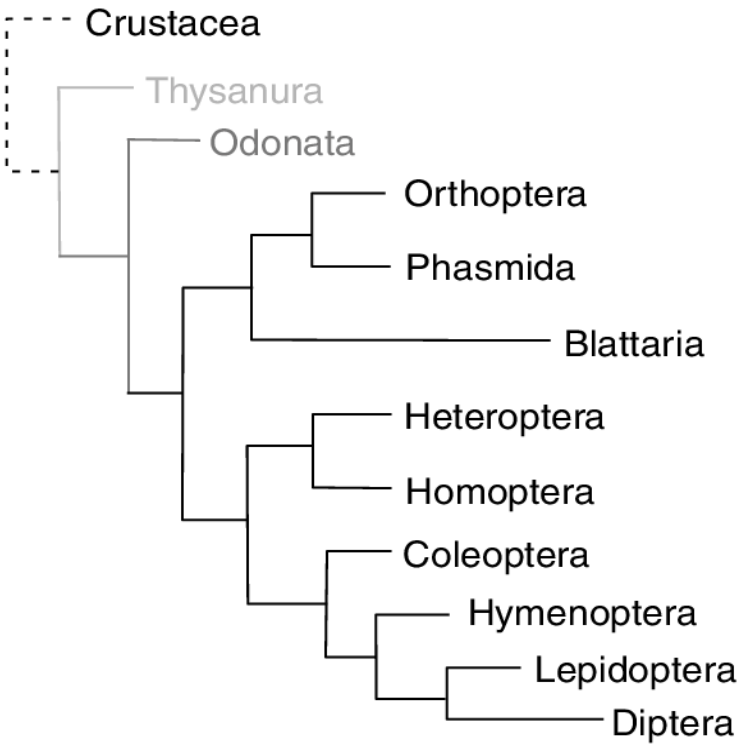


Figura 13: Árvore filogenética das principais ordens de insetos. As espécies *B. germanica* (*Blattaria*), *T. castaneum* (*Coleoptera*), *D. melanogaster* (*Coleoptera*), *H. armigera* (*Lepidoptera*) foram utilizadas na construção da filogenia de *zelda*. Embora seja uma forma generalista e imprecisa, as ordens representativas de cada espécie foram utilizadas para a comparação das árvores aqui apresentadas (modificado de Wheeler *et al.,* 2001).

Na filogenia obtida para o gene *zelda*, é possível ver muitas semelhanças entre as filogenias das figuras 11, 12 e 13. O grupo *Apini, Meliponini* e *Bombini* na filogenia de *zelda* apresenta a mesma disposição dos clados das filogenias construídas por Kapheim e Romiguier. As espécies *M. rotundata, O. bicornis, H. laboriosa e E. mexicana* também apresentam distribuição cladística semelhante

ao estudo de Kapheim (Figura 12). Uma exceção é *D. novaeangliae*, cuja localização é diferente da proposta por Kapheim. A espécie *C. calcarata, C. cinctus e O. abietinus* possuem localização semelhante à filogenia proposta por Rehan e colaboradores (2016).

Quando tentamos analisar os grupos formados na árvore de *zelda,* fazendo uso da ordem de cada espécie, algumas diferenças puderam ser facilmente notadas. Com base na filogenia proposta por Wheeler (Figura 13), *Lepidoptera* e *Diptera* são grupos irmãos. Contudo, na árvore de zelda, as espécies *H. armigera* (*Lepidoptera*) e *D. melanogaster* (*Diptera*) não são espécies irmãs, sendo a sequência do gene de *D. melanogaster* mais semelhante as sequências de *Hymenoptera*. Já a espécie *H. armigera* ficou na condição de grupo externo. Isso provavelmente se deve a pouca representatividade de espécies da mesma ordem, enviesando o resultado da filogenia. As espécies representantes das ordens *Blattaria* (*B. germanica*) e *Coleoptera* (*T. castaneum*) também não seguem a distribuição de grupos proposta por Wheeler.

As relações filogenéticas se encontram próximas ao proposto pelos estudos citados, mostrando uma alta conservação na sequência e estrutura deste gene, ao menos nas espécies estudadas. Os resultados da busca por ortólogos para o gene *zelda*, principalmente em insetos nos quais já está bem caracterizado, podem ser muito úteis para concluir sobre as funções deste gene no início do desenvolvimento de *A. mellifera*.

4.2- Escolha de genes candidatos para análise por qPCR

A análise por qPCR teve por objetivo validar a expressão de genes candidatos durante as primeiras horas da embriogênese, no intervalo de 30 minutos a 6 horas, comparando com dados das bibliotecas de RNA. Os genes propostos inicialmente para essa investigação foram o fator de transcrição *zelda* e o ligante de RNA *smaug,* que além de estarem relacionados a esse período do desenvolvimento, também apresentaram resultados estimulantes em estudos anteriores.

Para verificar se outros genes possuíam o mesmo comportamento de *zelda* e *smaug*, foram analisadas bibliotecas de RNAseq (mRNA) de embriões haploides e diploides de *A. mellifera* presentes no banco de bibliotecas do Laboratório de Biologia do Desenvolvimento de Abelhas (LBDA). Essas bibliotecas revelaram que alguns genes não apresentam expressão em ovócitos maduros, 0 horas, ou seja, antes que seja feita a postura. Porém, 1 hora após a postura já apresentam algum nível de expressão, suficiente para ser detectado pela análise de transcriptoma. Casos inversos, no qual determinados genes que se expressam no tempo 0 e, posteriormente, com 1 hora apresentam expressão nula ou muito baixa também foram selecionados para o estudo. Como exceção, foi mantido o gene *smaug*, que apresenta expressão em ambos os períodos de desenvolvimento (0 e 1 hora) (Tabela 3). Além desses genes, também foi incluído na análise o gene normalizador *ribosomal protein L32 (RPL32)* (Lourenço, 2008).

Assim, seria esperado que os genes que estão expressos 1 hora após a postura e não são detectados no tempo 0 foram transcritos pelo genoma do

embrião e não foram depositados, no ovócito, pela mãe. Para análise dos genes candidatos, ou seja, transcritos em uma fase inicial da embriogênese, foram usados os *primers* PRS, que se ligam na região intrônica dos mRNAs, o que significa a amplificação de transcritos recém-sintetizados. Dessa forma, se houver expressão desses genes acredita-se que eles foram produzidos pelo genoma zigótico e não foram depositados no ovócito durante a ovogênese.

Tabela 3: Níveis de expressão (RPKM) dos genes selecionados de bibliotecas de mRNA, para validação por qPCR. Os genes listados são: *cuticular protein 21 (CPR21), take-out-like carrier protein (JHBP-1B), troponin C type IIb (TpnCIIb), protein takeout (to), UDP-glucuronosyltranferase 1-8 (UDP), matrix-metalloproteinase-14 (Mmp14), calcium-dependent protein Kinase 4 (PDCALCIUM), smaug (smg) e Ribosomal protein L32 (RpL32).* 1H_dipl (embriões diploides 1 hora após a postura), 1H_hapl (embriões haploides 1 hora após a postura).

GB_Amel	Genes	0H	1H_dipl	6H_dipl	1H_hapl	6H_hapl
GB46311	CPR21	0	2.2640421156	0	303.771482608	0.3229665836
GB48492	JHBP-1B	0	0.7405481077	0	158.342521318	0.1584592601
GB48216	TpnCIIb	0	1.5586536186	0	11.3212131498	0
GB42799	to	952.0378315	0	0	0.3520041063	0
GB55864	UDP	9.596366316	0	0.071657904	0.2854995137	0.0538277639
GB55889	MmP14	83.82544897	0	0	6.6084944834	0.1962139012
GB48539	PDCALCIUM	8.244200374	0	0	2.9792666184	0.0802439496
GB48452	smg	199.14251256	307.663912743	360.6026778	320.582141088	430.149164699
GB47227	RpL32	438.3421956	427.177438869	177.8427958	1309.35019973	143.344138737

Além dos genes escolhidos pelo critério da quantidade de transcritos no ovócito maduro e no embrião, selecionados de bibliotecas de RNA-seq, também foram eleitos outros genes para a análise, com base em dados de literatura.

Estudos em *D. melanogaster*, que também abordavam os primeiros ciclos mitóticos dos embriões, apontaram que o gene *engrailed* é sintetizado pelo embrião antes do oitavo ciclo mitótico, sendo essencial para a sincronia das mitoses, funcionamento normal dos ciclos celulares e manutenção da simetria

das células polares com a linha média do embrião (Ali-Murthy *et al.,* 2013). Este mesmo estudo identificou que os genes *even-skipped, hairy, Kruppel e odd-skipped* também são transcritos nos ciclos mitóticos inicias do embrião. Por esse motivo, esses genes também se tornaram alvo dos nossos estudos de qPCR.

Os genes que são transcricionalmente ativados no início do desenvolvimento de *D. melanogaster*, geralmente, são pequenos e não possuem íntrons. De Renzis *et al.* (2007) identificou que dos 59 primeiros genes expressos pelo embrião, 70% não possuíam íntrons. O comprimento reduzido destas unidades de transcrição pode ser um requisito para a expressão precoce devido às divisões nucleares rápidas neste momento em desenvolvimento (Shermoen & O'Farrell, 1991). Essa característica dos primeiros genes zigóticos são provavelmente compartilhadas entre organismos cujo desenvolvimento inicial é caracterizado por divisões rápidas. Um outro exemplo são os genes zigóticos iniciais no peixe *paulistinha (Danio rerio),* que também apresentam poucos ou nenhum íntron (Heyn *et al.,* 2014). Além disso, foi visto que muitos destes genes são fatores de transcrição, enquanto que aqueles que são depositados maternalmente são, preferencialmente, genes de manutenção (De Renzis *et al.,* 2007). Assim como em *D. melanogaster* e *D. rerio,* nosso modelo experimental, aparentemente, também possui essas características, uma vez que, dos 30 genes selecionados incialmente na biblioteca de mRNA de embriões, somente 8 apresentaram as condições necessárias para que o desenho experimental fosse aplicado, ou seja, ao menos um ou mais íntrons separando os éxons. Os genes *en, eve, h, kr e odd* que fazem parte dos ciclos iniciais da embriogênese de *D. melanogaster* também são pequenos e apresentam poucos íntrons.

Para verificar em quais processos biológicos os genes selecionados estariam inseridos e em quais fases do desenvolvimento esses genes apresentam pico de expressão, foi feita uma busca no banco de dados de *D. melanogaster,* FlyBase (Tabela 4).

Tabela 4: Processos biológicos dos genes selecionados para as análises. Os processos biológicos dos genes selecionados foram retirados do banco de dados Flybase. Por essa tabela, é possível ver que todos os genes participam do processo de desenvolvimento, sendo que a maioria possui pico de expressão no estágio embrionário.

Genes	Org. cel./biogênese	Desenvolvi-mento	Resp. ao estímulo	Sinalização	Ciclo cel./ proliferação	Expressão gênica	Transporte/ localização	Reprodução	Proc. Sist. nervoso	Comporta-mento	Metabolismo de proteínas	Estágio do pico de expressão
En	■	■	■			■		■				Embrionário- 0-6 horas
Eve	■	■	■			■						Embrionário- 0- 6 horas
H	■	■	■			■						Embrionário- 0- 6 horas e 18h- 24 horas
kr	■	■	■		■	■						Embrionário- 0- 12 horas
odd		■				■						Embrionário- 0- 6 horas

Genes	Org. cel./biogênese	Desenvolvi-mento	Resp. ao estímulo	Sinalização	Ciclo cel./ proliferação	Expressão gênica	Transporte/ localização	Reprodução	Proc. Sist. nervoso	Comporta-mento	Metabolismo de proteínas	Estágio do pico de expressão
CPR21		■										Larval- final
JHBP-1B		■	■					■		■		Embrionário e larval
TpnCIIb		■	■	■								Embrionário e Larval
to		■	■					■		■		Embrionário e larval
UDP		■	■	■							■	Embrionário e larval
MmP14	■	■	■								■	Pupal
PDCALCIUM		■	■	■							■	Pupal
smg		■				■					■	Embrionário-0-6horas
RpL32		■				■					■	Todas as fases

Uma rápida análise da Tabela 4 mostra que os genes candidatos tem participação durante os primeiros passos do desenvolvimento embrionário, o que não exclui a participação dos mesmos em outras fases do desenvolvimento.

4.3- Análise da expressão de *takeout, troponin C e RpL32* em embriões diploides e haploides no intervalo de 0- 6 horas após a postura

Inicialmente, foi realizado um teste por PCR convencional para amplificar todos os genes selecionados a partir de amostras do *pool* de embriões de 0 e 1 hora, e a presença ou ausência da banda específica foi analisada em eletroforese em gel de agarose (Figura 14). Esse teste inicial foi importante para verificar o tamanho dos transcritos amplificados, eliminando assim a suspeita de contaminação por DNA genômico. Após essa análise inicial, os *primers* foram testados quanto à sua eficiência. Nesse teste, somente os genes *smaug, takeout, troponin C type IIb* e *RpL32* apresentaram curvas de eficiência satisfatórias para ambos os tipos de *primers* e, por conseguinte, foram utilizados nas análises por qPCR. Para garantir que os *primers* desenhados amplificavam os seus respectivos genes, foi realizado o sequenciamento dos segmentos que foram amplificados por cada *primer* (Anexo A). Os demais *primers* não foram utilizados a princípio.

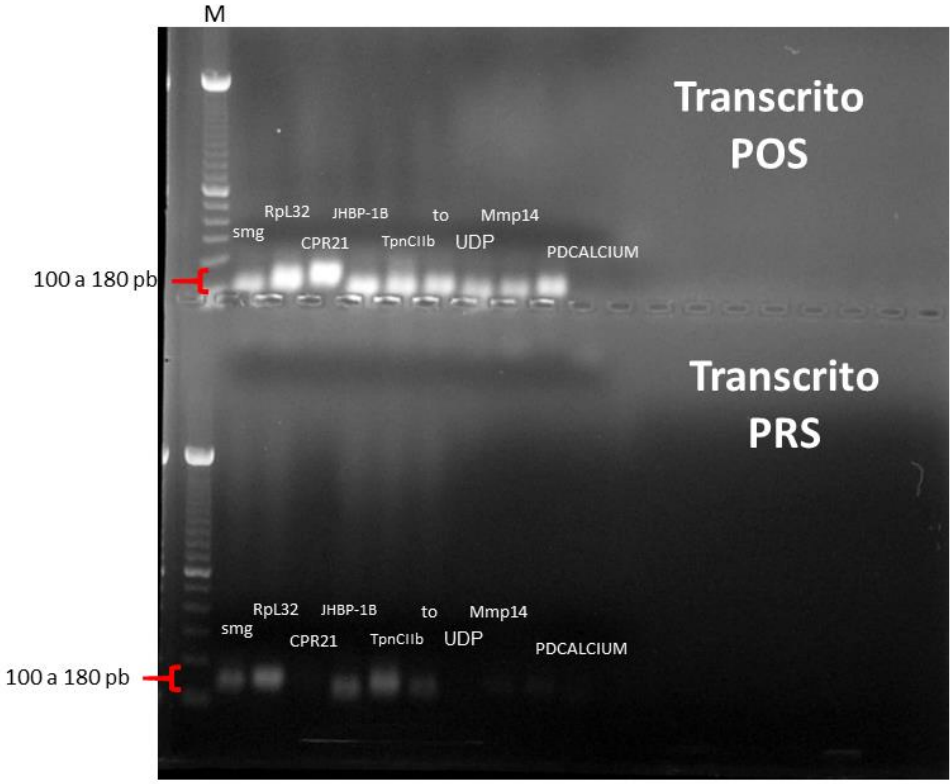


Figura 14: Análise em gel de agarose da amplificação dos genes candidatos, por PCR convencional, usando os *primers* POS (Pós-processamento) e PRS (Pré-processamento). No primeiro *front* (de cima para baixo), estão presentes as bandas geradas pelos *primers* POS e no segundo, PRS. Todos os fragmentos dos genes amplificados possuem um tamanho esperado (entre 100 e 180 pares de base). Da esquerda para a direita, estão presentes os seguintes genes: *smaug (smg), Ribosomal protein L32 (RpL32), cuticular protein 21 (CPR21), take-out-like carrier protein (JHBP-1B), troponin C type IIb (TpnCIIb), protein takeout (to), UDP-glucuronosyltranferase 1-8 (UDP), matrix-metalloproteinase-14 (Mmp14), calcium-dependent protein Kinase 4 (PDCALCIUM).* As amostras biológicas foram compostas de *pools* de embriões antes da postura (0 hora) e 1 hora após a postura (1 hora). Marcador (M) de 100 pb.

Em virtude das caraterísticas únicas dessa fase de desenvolvimento, foi visto, nos resultados de qPCR, que o gene *smaug* é melhor candidato a gene endógeno do que o gene *RpL32*, tanto para o primer POS quanto para o PRS em decorrência da sua expressão estável, pelo menos no período do desenvolvimento testado. Assim, o gene *RpL32* também foi analisado, juntamente com os demais genes, porém não mais como normalizador.

Os resultados prévios obtidos pela análise de bibliotecas de mRNA de *A. mellifera* demonstraram que os genes selecionados apresentam expressão

gênica diferencial entre os períodos de 0 e 1 hora de desenvolvimento (Tabela 3). A Figura 15 mostra os resultados das análises por qPCR, usando primers POS (A) e os primers PRS (B) em embriões diploides com idades variando de 30 minutos a 6 horas após a postura e também em ovócitos maduros. Na Figura 15 A, é possível ver que a expressão do mRNA já processado de *RpL32* se mantém relativamente estável até E4. A partir desse ponto, os valores de expressão diminuem consideravelmente. Quando observamos os resultados relativos ao pré-processamento (Fig. 15 B), podemos ver um pico de expressão em, embrião de 30 minutos (E1/2). Pequenos picos de expressão também podem ser identificados em E5 e E6. Essa expressão em E1/2 pode indicar uma ativação precoce do genoma zigótico, sendo que, nos próximos 30 minutos (E1) esta expressão já está menor, ou seja, ele está sendo pouco expresso. Possivelmente, o transcrito produzido em E1/2 no zigoto (Fig. 15 B) passa pelo processamento de mRNAs e é detectado até 3 horas após a postura (E1/2 a E3) como mostra a Figura 15 A, que é referente aos transcritos de RNA já processados.

Os dados mostrados na Figura 15 C e D se referem à expressão do gene *takeout,* mRNA e pré-RNA respectivamente. O gene *takeout* apresenta um claro pico de expressão em ovócitos maduros (E0), seguido de uma grande redução nas demais idades. Em D, um perfil semelhante é observado, embora a magnitude entre os pontos seja menor. Esse dado sugere que a maior parte dos transcritos desse gene são de origem materna. Entretanto meia hora após a postura(E1/2) foi detectado pré-mRNA, ainda que em quantidades pequenas, indicando que este gene também está sendo transcrito bem cedo pelo genoma zigótico.

Na Figura 15 E e F temos a expressão do gene *troponin C type IIb*, mRNA e pré-RNA respectivamente. Um nível alto de expressão é observado em ovócito maduro (E0), sendo que nas demais idades a expressão é baixa, sendo próximo a 0, principalmente em E5 e E6. Possivelmente esses transcritos foram depositados durante a ovogênese, atuando principalmente no ovócito maduro. Esse transcrito materno, apresenta indicativos de degradação logo no início da embriogênese, ao mesmo tempo em que começa a ser expresso pelo genoma do embrião, em E1/2(Figura 15F), se mantendo relativamente estável.

A Figura 16 mostra os resultados obtidos por qPCR, usando primers POS (A) e os *primers* PRS (B), para os genes apresentados na Figura 15, contudo em embriões haploides com idades variando de 30 minutos a 6 horas após a postura e também em ovócitos maduros.

Na Figura 16 A, é possível ver um pico de expressão de *RpL32* uma hora após a postura (E1), que corresponde aos mRNAs já processados, enquanto que quando observamos (Figura 16 B) os transcritos não processados, a expressão é alta em E1/2 e E1, indicando uma possível expressão zigótica. Assim, é possível assumir que o pico em E1 da Figura 16 A seja de transcritos processados e de origem zigótica.

Na Figura 16 C, relativa ao transcrito do gene *takeout* já processado observamos um pico de expressão já em 30 minutos após a postura (E1/2), bem como em uma hora após a postura (E1), sendo esse mRNA possivelmente de origem materna. Isso pode ser afirmado, pois no gráfico D não foi observado níveis de transcritos pré-processados expressivos após a postura.

Os níveis de expressão de *Troponin C* mostrados em E e F da Figura 16 indicam uma possível atividade do genoma zigótico desde 30 minutos (E1/2), sendo essa síntese inicial do embrião refletida em E1/2 e E1. Seis horas após a postura vemos uma significativa redução dos níveis de expressão, o que pode indicar uma possível degradação dos transcritos já processados.

Quando comparamos, de forma geral, os perfis de expressão mostrados nos gráficos dos embriões haploides e diploides, é possível dizer que há fortes indícios de estar ocorrendo transcrição zigótica nos primeiros ciclos nucleares do embrião, nos dois tipos de desenvolvimento. Assim, pode ser afirmado que tanto os embriões diploides, quanto os haploides de *A. mellifera* possuem processos de ativação do genoma zigótico já nas fases iniciais da embriogênese, como descrito para outros insetos, inclusive *D. melanogaster* (Tadros et al., 2007, Benoit *et al.,* 2009, Ali-Murthy, et al., 2013).

Pelo menos para os genes analisados não foi encontrada uma defasagem de horário, como proposto por Sasaki & Obara 2002. O atraso no desenvolvimentho dos embriões haploides, sugerido pelos autores, seria devido à uma espera pela fertilização, por pelo menos 30 minutos. Contudo, podemos observar uma modulação diferencial na expressão de *Troponin C*, quando comparamos o desenvolvimento haploide e diploide.

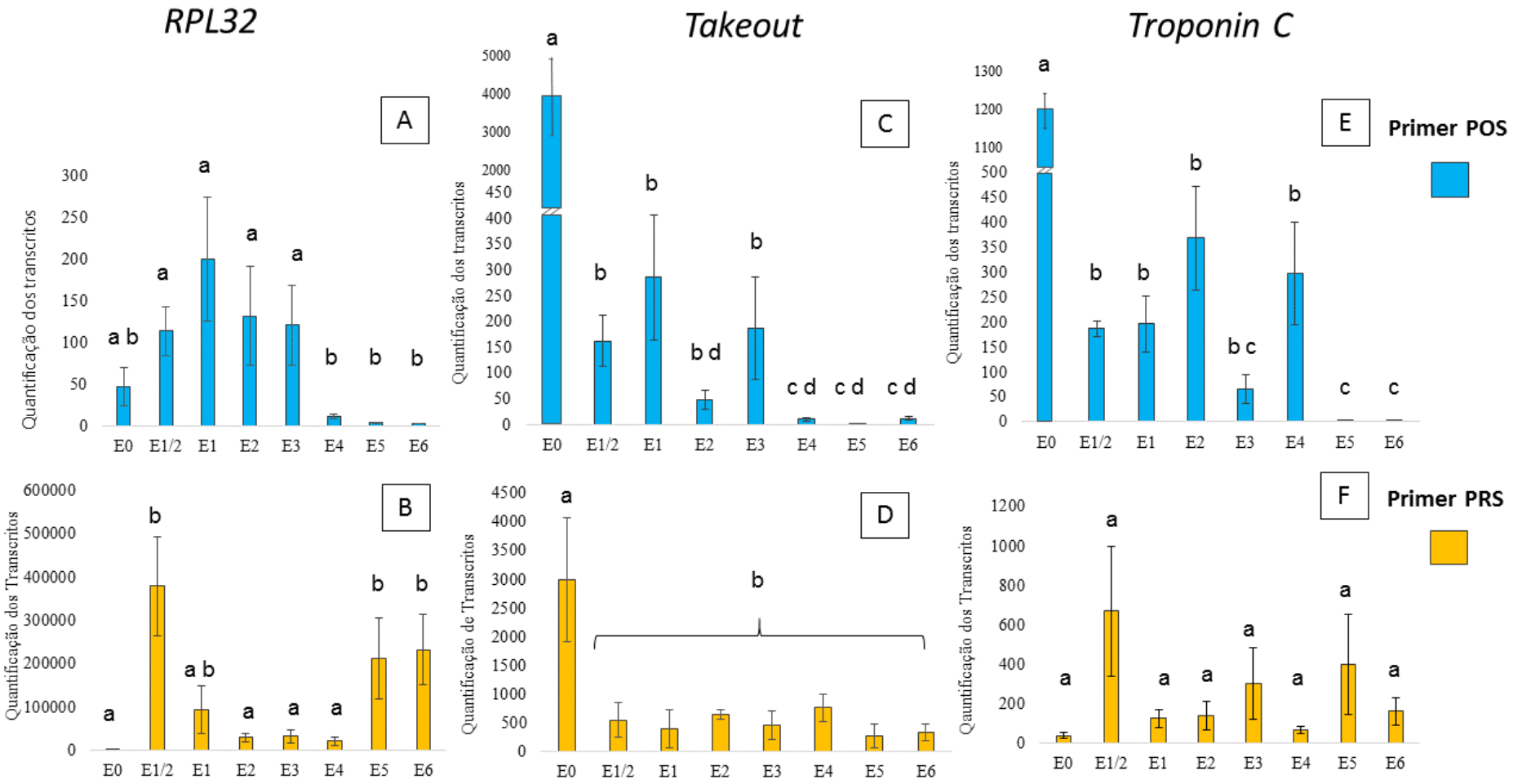


Figura 15: Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário diploide de *A. mellifera* (E1/2 a E6 horas). Expressão relativa de RpL32, Takeout e Troponin C type IIb, usando primers específicos para mRNA e pré-RNA, quantificada por qPCR e analisada usando o método ΔΔCt. Os níveis de expressão foram normalizados pela expressão relativa de *smaug.* Cada barra representa a média e desvio padrão de três réplicas biológicas. Letras diferentes indicam diferenças significativas determinadas pelas análises por ANOVA one-way e Tukey post hoc, quando aplicável ($p \leq 0,05$).

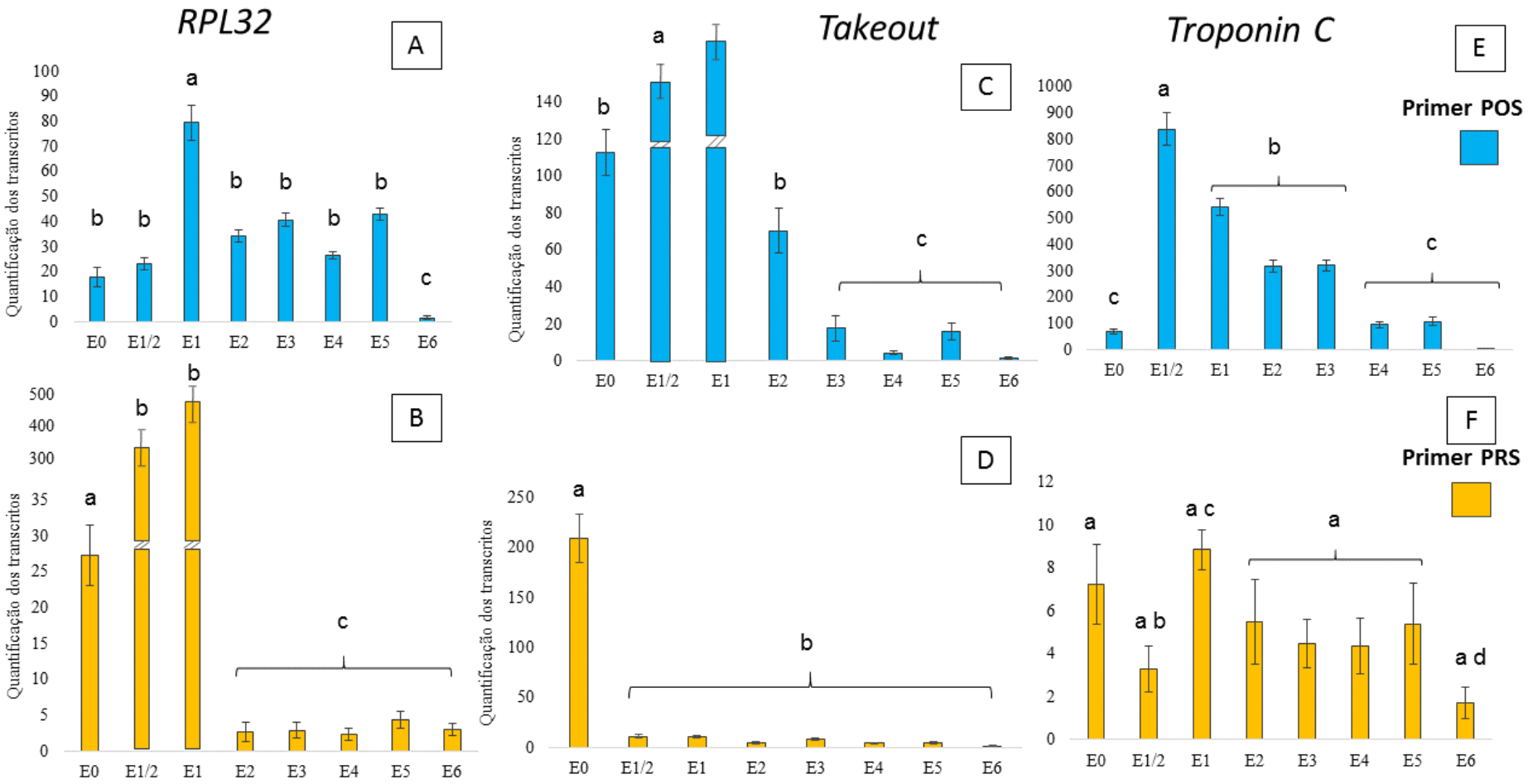


Figura 16: Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário haploide de *A. mellifera* (E1/2 a E6 horas). Expressão relativa de RpL32, Takeout e Troponin C type IIb, usando primers específicos para mRNA e pré-RNA, quantificada por qPCR e analisada usando o método ΔΔCt. Os níveis de expressão foram normalizados pela expressão relativa de *smaug.* Cada barra representa a média e desvio padrão de três réplicas biológicas. Letras diferentes indicam diferenças significativas determinadas pelas análises por ANOVA one-way e Tukey post hoc, quando aplicável ($p \leq 0,05$).

4.4- Análise da expressão gênica de *engrailed, even-skipped, hairy, Kruppel e odd-skipped* em embriões diploides e haploides no intervalo de 0- 6 horas após a postura

Para testar a integridade do material genético, bem como a especificidade dos *primers*, foram feitas PCRs convencionais, usando como amostra um *pool* de ovócitos maduros e embriões de 1 hora tanto haploide, quanto diploide. Em seguida, os fragmentos amplificados foram analisados em gel de agarose a 1% (Figura 17). Pela análise do gel, foi possível observar que todos os transcritos formam uma única banda e que o tamanho dos segmentos amplificados era o esperado (entre 100- 150 pb, variando para cada para de *primer*).

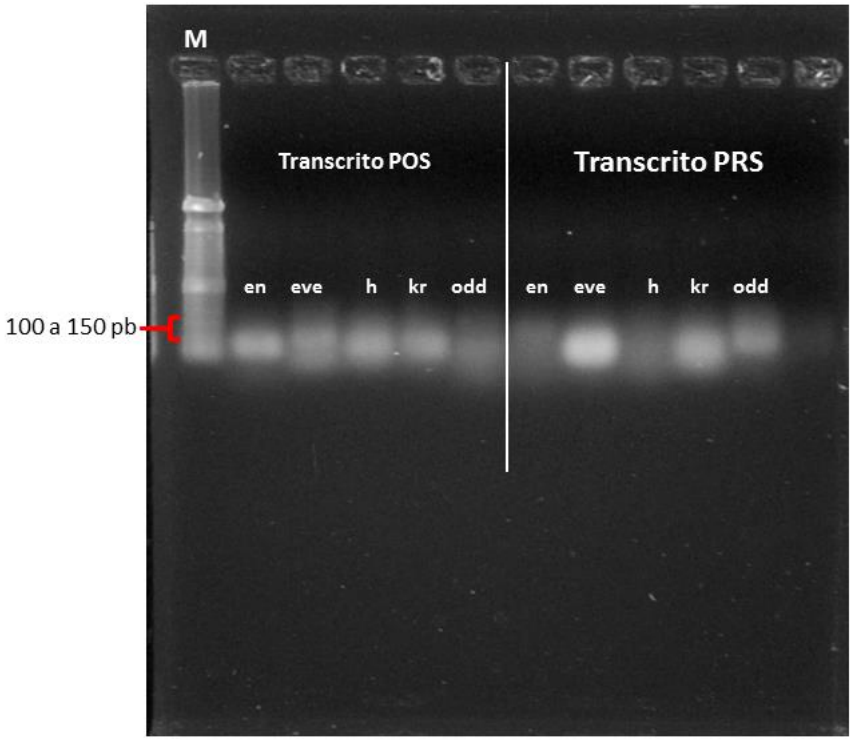


Figura 17: Gel de agarose 1% para análise dos fragmentos de RNA amplificados usando-se *primers* POS (Pós-processamento) e PRS (Pré-processamento). As cinco primeiras bandas, da esquerda para a direita, correspondem aos primers de transcritos pós-processados (POS), e as cinco bandas posteriores correspondem aos primers de transcritos pré-processados (PrS). Da esquerda para a direita estão presente os genes*: engrailed (en), even-skipped (eve), hairy (h), kruppel (kr) e odd-skipped (odd).* Marcador (M) de 100 pb.

Após a verificação dos primers por meio da PCR convencional, foram feitas as curvas de eficiência para cada primer e, em seguida, os experimentos de qPCR. A figura 18 mostra o perfil de expressão dos genes *en, eve, h, kr* e *odd* em embriões diploides. Na figura 18 A, que compreende os transcritos já processados, há uma maior quantidade do gene *en* em embriões 1 hora após a postura (E1). O gráfico 18 B é referente aos transcritos pré-processamento e aponta para uma maior expressão de *en* a partir da quinta hora após a postura. Com isso, é possível sugerir que os níveis de transcritos observados em E1 (18 A) é decorrente de transcritos de origem materna e os níveis observados em E5 e E6 são produtos de origem zigótica.

A análise do gene *eve* pós-processado (18 C) mostra um pico em ovócitos (E0), e pequenos sinais de expressão de E1 a E4. Na Figura 18 D, referente ao transcrito de *eve* pré-processado, observa-se níveis maiores de E1 a E3. Isso é um forte sinal de que os transcritos vistos em E1 a E4, Figura18 C sejam de mRNAs zigóticos, já processados.

Na Figura 18 E, relativa ao gene *h* pós-processado, há maior expressão em ovócitos (E0) e 1 hora após a postura (E1), estes últimos, provavelmente, zigóticos. Em 18 F, ocorre um aumento gradual dos transcritos até o ponto E4, indicando um aumento da síntese zigótica gradualmente.

Nos gráficos mostrados na Figura 18 G, observamos dois picos bem proeminentes em E1/2 e E2 de *kr*. O primeiro, 30 minutos após a postura(E1/2) é provável que os transcritos sejam de origem materna. No entanto, duas horas após a postura (E2) pode ser uma mistura de *kr* materno e produzido pelo

embrião, pois na Figura 18 H é possível identificar maior expressão em E1. Em E5 e E6 já são transcritos produzidos pelo embrião.

Os níveis de transcritos do gene *odd* após o processamento estão mostrados na Figura 18 I. É possível observar um E0, de transcritos oriundos da mãe. Na Figura 18 J, a partir de uma hora (E1) há transcrição de *odd* proveniente do embrião. Contudo, Na Figura 18 I (E1) observamos um pequeno pico, provavelmente resultante do processamento dos transcritos detectados em E1 (Figura 18J), no entanto, não há mais picos de expressão nas demais horas de desenvolvimento embrionário, mas sim uma produção constante, que provavelmente não é processada, como mostra os gráficos na Figura 18 I, cujos níveis são muito baixos

Os gráficos da Figura 19 A-J correspondem ao perfil de expressão de embriões haploides. No gráfico da figura 19 A, há um pico de expressão de *en* em 1 hora após a postura (E1), cujos transcritos podem ser de origem materna ou zigótica, pois no gráfico 19 B há um pico de expressão em E1/2. Também há um outro pico em E5 que provavelmente é advinda do embrião.

A figura 19 C, de transcritos pós-processados do gene *eve,* apresenta somente um pico de expressão em ovócito maduro (E0), de origem materna. Já na figura 19 D, tem dois picos de expressão nas idades E5 e E6 de transcritos pré-processados, indicando uma transcrição zigótica do gene *eve* a partir da quinta hora em embriões haploides.

O transcrito do gene *h* pós-processado possui seu maior pico de expressão meia hora após a postura (E1/2), e é provável que seja depositado no embrião durante a ovogênese, conforme pode ser visto em 19 E. Já os

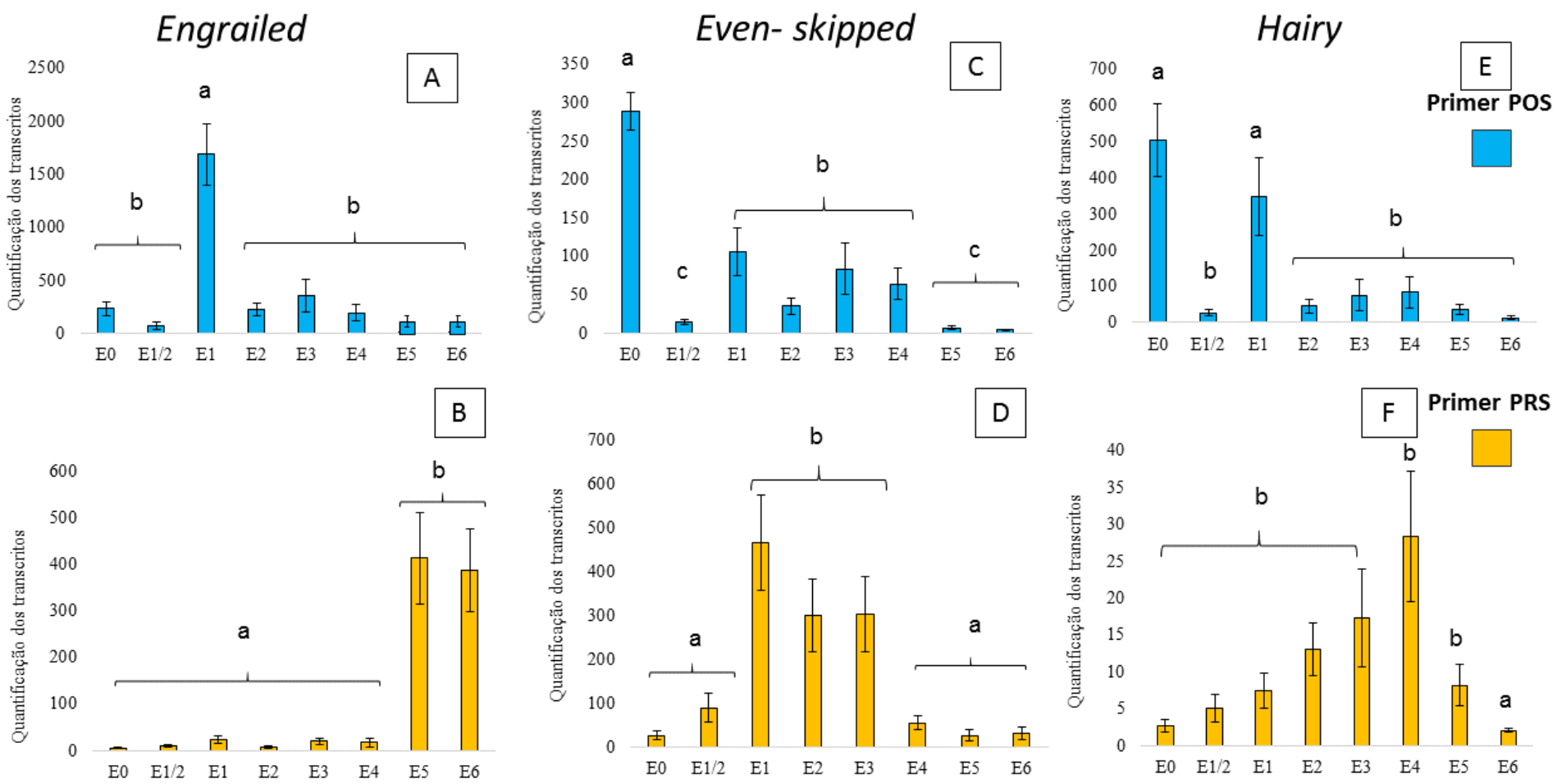
transcritos pré-processados de *h* têm uma sutil expressão em E1/2, E1 e E2, podendo indicar o início da transcrição zigótica deste gene.

A figura 19 G relativo ao gene *Kr* pós-processado possui um pico elevado em E1/2, que possivelmente foi depositado pela mãe. Em embriões haploides, ocorre aumento de expressão de transcritos pré-sintetizados em E5 e E6, como pode ser visto em 19 H.

Na figura 19 I, é possível identificar um pico de expressão do gene *odd* em ovócito maduro E0, que é de origem materna. Em 19 J, há indícios de expressão zigótica de *odd* a partir do ponto E1, sendo detectado transcritos até E6, porém, aparentemente, a quantidade de produção não é constante.

Os dados obtidos para esses genes apontam para pequenas diferenças nos tempos da ativação da transcrição entre os embriões diploides e haploides, porém, todos os genes analisados apresentam algum nível de expressão zigótica em, pelo menos, alguma idade estudada. Estudos anteriores em *D. melanogaster* mostraram que a ativação do genoma zigótico é gradual, com centenas de genes sendo ativados a cada ciclo mitótico do embrião (Li *et al.,* 2014, Lott *et al.,* 2011). Ao mesmo tempo que os transcritos zigóticos começam a ser produzidos, ocorre a degradação dos transcritos maternos (Bushati *et al.,* 2008; Lund *et al.,* 2009). A dinâmica de degradação de transcritos maternos e expressão de transcritos zigóticos é coordenada por mRNAs e miRs (Aboobaker *et al.,* 2005; Biemar *et al.,* 2005; Fu *et al.,* 2014). O fator de transcrição *zelda* pode ativar a expressão de vários genes zigóticos (Kanodia *et al.,* 2012), sendo possível que os genes analisados sejam regulados pelo próprio. Já na degradação dos fatores maternos, o gene *smaug* pode estar interferindo,

ajudando a desestabilizar a estrutura dos mRNAs depositados durante a ovogênese (Gotze & Whale, 2014). Porém, ainda assim é bem claro que a ativação do genoma zigótico pode ocorrer nos ciclos iniciais de clivagem de embriões de *A. mellifera*.


Engrailed
Even- skipped
Hairy
A
B
C
D
E
F
Primer POS
Primer PRS
Quantificação dos transcritos
E0
E1/2
E1
E2
E3
E4
E5
E6

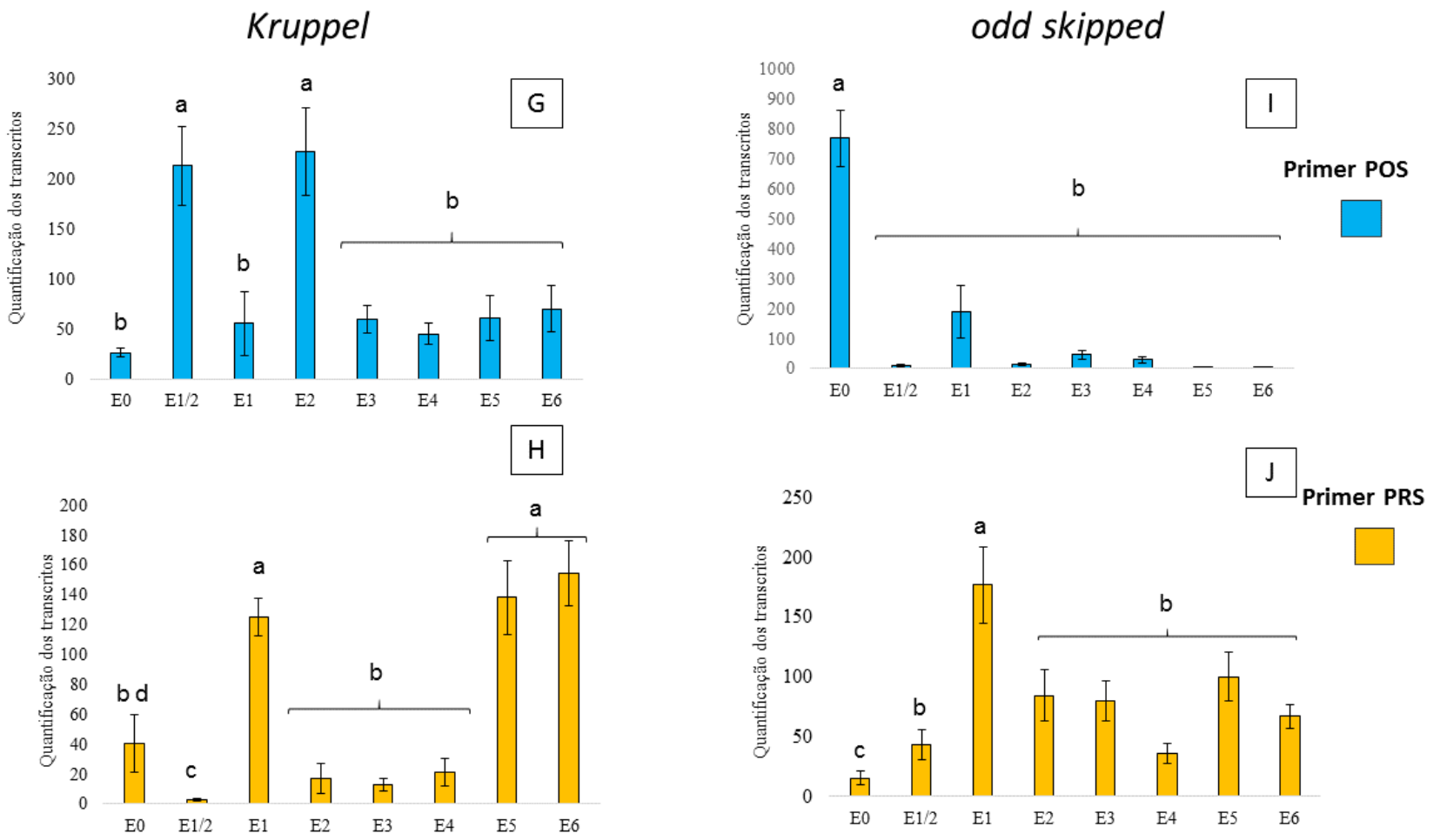
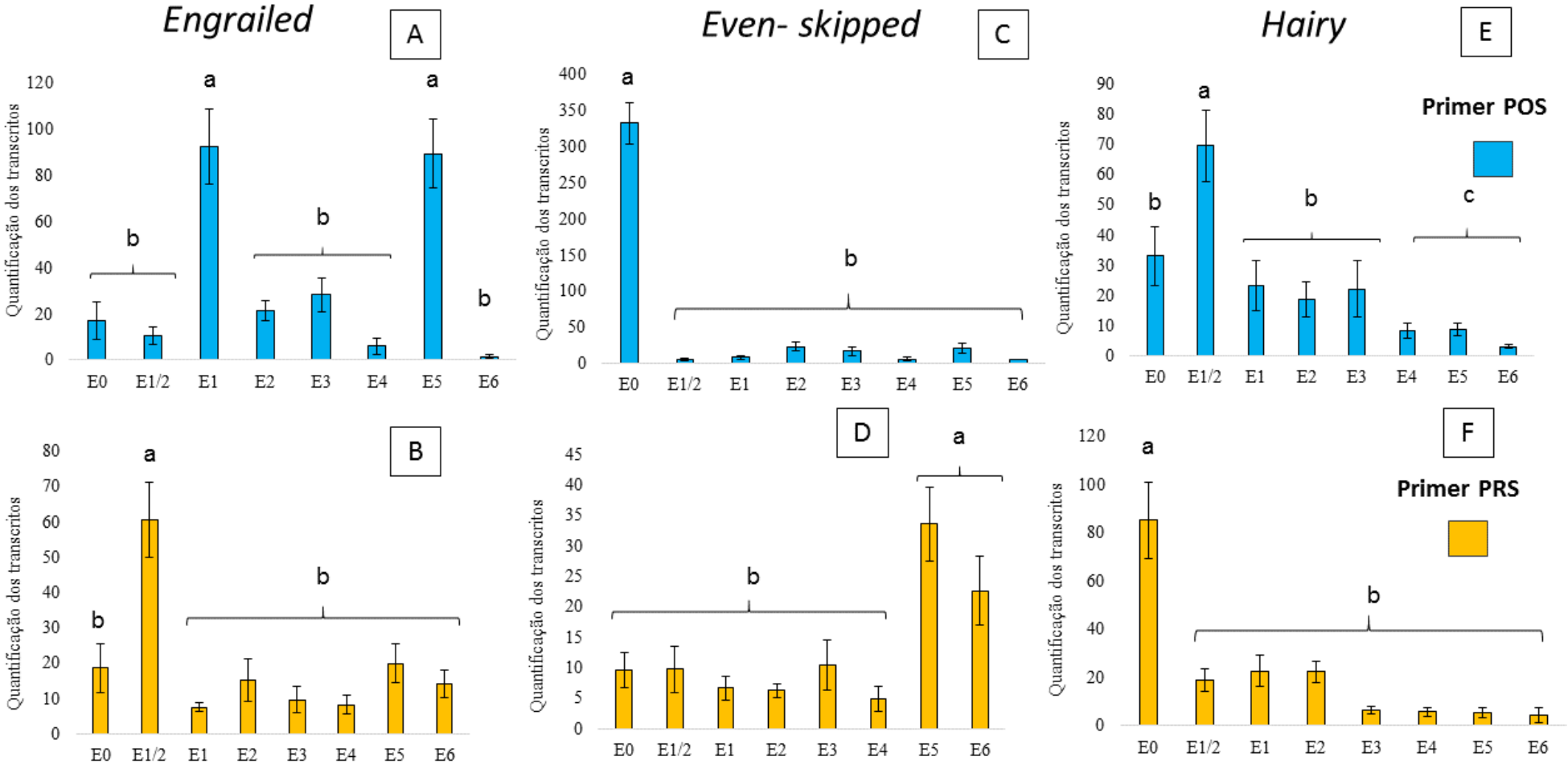


Figura 18: Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário diploide de *A. mellifera* (E1/2 a E6 horas). Expressão relativa de *engrailed, even-skipped, hairy, kruppel, odd-skipped*, usando primers específicos para mRNA e pré-RNA, quantificada por qPCR e analisada usando o método ΔΔCt. Os níveis de expressão foram normalizados pela expressão relativa de *smaug.* Cada barra representa a média e desvio padrão de três réplicas biológicas. Letras diferentes indicam diferenças significativas determinadas pelas análises por ANOVA one-way e Tukey post hoc, quando aplicável ($p \leq 0,05$).


Engrailed
A
Even- skipped
C
Hairy
E
Primer POS
B
D
F
Primer PRS
Quantificação dos transcritos
E0
E1/2
E1
E2
E3
E4
E5
E6

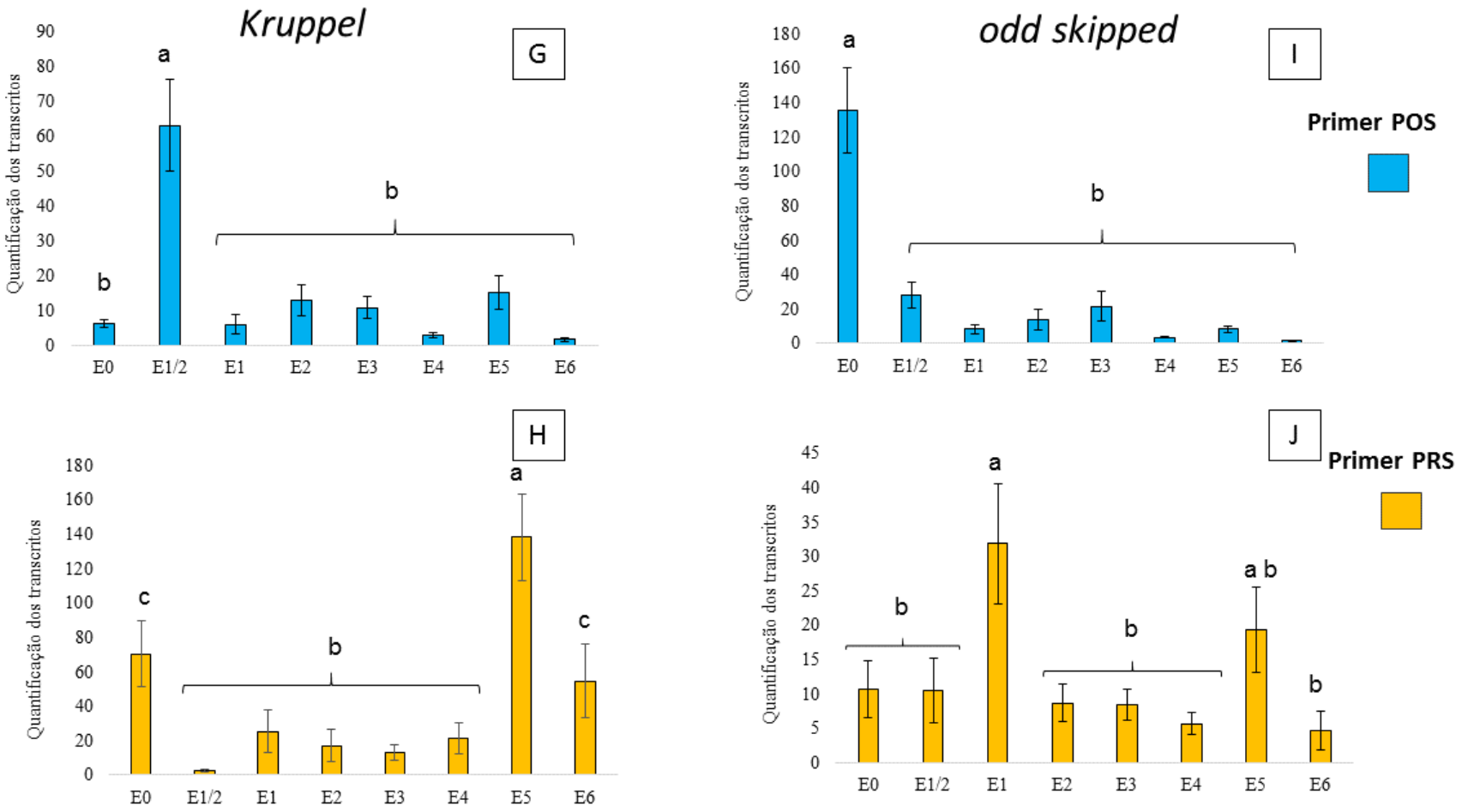


Figura 19: Análise da expressão de mRNAs (POS, pós-processamento) e pré- mRNAs (PRS, pré-processamento) de ovócitos maduros (E0) e durante as primeiras fases do desenvolvimento embrionário haploide de *A. mellifera* (E1/2 a E6 horas). Expressão relativa de *engrailed, even-skipped, hairy, kruppel, odd-skipped*, usando primers específicos para mRNA e pré-RNA, quantificada por qPCR e analisada usando o método ΔΔCt. Os níveis de expressão foram normalizados pela expressão relativa de *smaug.* Cada barra representa a média e desvio padrão de três réplicas biológicas. Letras diferentes indicam diferenças significativas determinadas pelas análises por ANOVA one-way e Tukey post hoc, quando aplicável ($p \leq 0,05$).

Em *D. melanogaster,* experimentos de micoarranjos *de DNA* e embriões com deleções cromossômicas específicas, identificaram no 14º ciclo mitótico,1158 genes expressos pelo zigoto e 6485 transcritos que são maternalmente depositados (De Renzis et al., 2007). Outra estratégia para identificar transcritos zigóticos foi realizada utilizando os polimorfismos em diferentes linhagens de *D. melanogaster,* no qual foi possível identificar 5598 genes exclusivamente maternos, 2210 genes exclusivamente zigóticos e 1195 genes que tinham origem materna, mas que também estavam sendo expressos pelo embrião até o 13º ciclo mitótico (Lott et al., 2011). Além disso, análise de hibridização *in situ* de alto rendimento foi usado para determinar a expressão de aproximadamente 7000 genes em quatro estágios distintos na embriogênese: estágios 1 a 3, estágios 4 a 5, estágios 6 a 7 e estágios 8 a 9 (Lecuyer et al., 2007), sendo que aproximadamente 5000 genes são expressos em embriões nos estágios 1-3, quando o embrião tem até 2,5 horas após a postura. Esses dados, assim como os resultados obtidos por qPCR de ovócitos maduros e embriões de *A. mellifera* indicam que a transcrição zigótica em insetos não é comparável, analogamente, a um degrau, mas sim a uma rampa, na qual aos poucos vão sendo transcritos novos produtos pelo embrião. Porém, o mais importante é que este processo se inicia muito antes do que pensava.

4.5- Marcação com DAPI dos núcleos de embriões diploides com 1 hora

Para a verificação do número de núcleos de clivagem, e em consequência, determinar em qual ciclo mitótico se encontram os embriões com idade de 1 hora após a postura, foram feitas marcações com DAPI. Foram encontrados 3 núcleos de clivagens nos embriões diploides, indicando que ao menos o embrião nessa idade já sofreu duas divisões mitóticas (Figura 20). Além disso, a presença de três núcleos indica que as divisões podem ser assincrônicas.

Esses resultados nos permitem estimar que durante a primeira hora de desenvolvimento, o embrião passou por duas mitoses nucleares. Estudos anteriores que também tentaram estimar o ciclo das primeiras divisões nucleares mostraram que após 15 minutos de postura, ocorre a primeira clivagem em *A. mellifera* (Sasaki & Obara, 2002). Assim, os embriões das fotos provavelmente possuem idade próximas de 30 a 40 minutos após postura. Quando esses dados são combinados com os resultados de qPCR, é possível dizer que no intervalo observado ocorreram pelo menos dois ciclos mitóticos e que genoma de *A. mellifera* vai se tornando gradualmente ativo desde as primeiras clivagens.

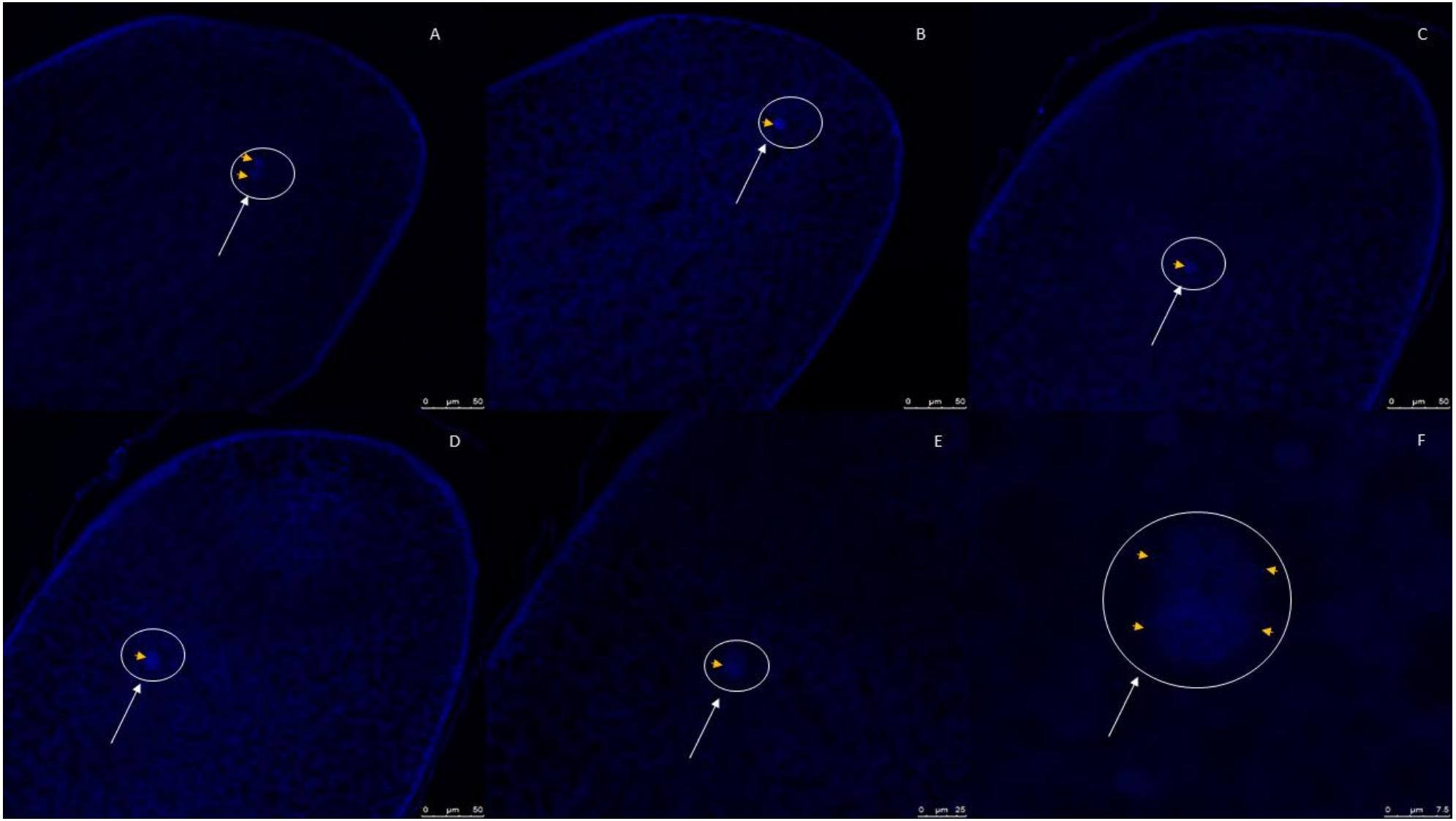


Figura 20: Marcação com DAPI de embriões diploides de *A. mellifera* com 1 hora de desenvolvimento após a postura. As imagens A e B correspondem ao mesmo embrião em planos de imagem diferentes. Em A é possível ver dois núcleos em um plano mais profundo do embrião. Em B é possível ver um núcleo em um plano mais superficial do embrião. As imagens C, D, E e F correspondem ao mesmo embrião em planos diferentes. Na imagem C é possível ver um único núcleo em um plano mais superficial. As imagens D, E e F é de um plano mais profundo do embrião e nele é possível ver 2 núcleos bem próximo de um ao outro. As setas e círculos brancos indicam a região dos núcleos de clivagem de cada embrião. As cabeças de setas laranjas apontam diretamente para os núcleos de clivagem. Imagens obtidas pelo microscópio confocal Leica® Sp5 - FMRP-USP.

4.6- Análise Topológica de Rede

Para a descrição completa da função de uma proteína é essencial o conhecimento sobre seus parceiros de interação específicos. As redes de proteínas são um resumo topológico das interações das proteínas e é de grande auxílio para determinar suas funções, interações com outras proteínas e também pode auxiliar no entendimento dos genes e proteínas que estão sendo, potencialmente, transcritos e traduzidos.

As redes aqui analisadas são de Interação Proteína-Proteína (*Protein-Protein Interaction*- PPIs) que são formadas por contatos físicos de alta especificidade, estabelecidos entre duas ou mais moléculas proteicas como um resultado de eventos bioquímicos, tais como forças eletrostáticas (De Las Rivas & Fontanillo, 2010). Essas redes foram primeiramente obtidas pelo banco de dados STRING (Szklarczyk *et al.,* 2014) e posteriormente, analisadas pelo software Cytoscape (Shannon *et al.,* 2003).

4.6.1- Redes das proteínas even skipped (eve) e Kruppel (Kr)

A segmentação em embriões de *D. melanogaster* é controlada por interações de várias classes de genes (Nüisslein-Volhard and Wieschaus 1980; Nüsslein-Volhard *et al.,* 1985). As informações são transmitidas em uma progressão temporal, que vai dos genes *gap,* para os genes *pair-rule* e, finalmente, para os genes de polaridade de segmento (Ingham, 1988). O gene even-skipped (eve), é tradicionalmente classificado como um gene *pair-rule,* além de ser um importante regulador transcricional. Ele é expresso ao longo de todo o embrião (Frasch *et al.* 1987) e sua distribuição é do polo anterior para o posterior. Além

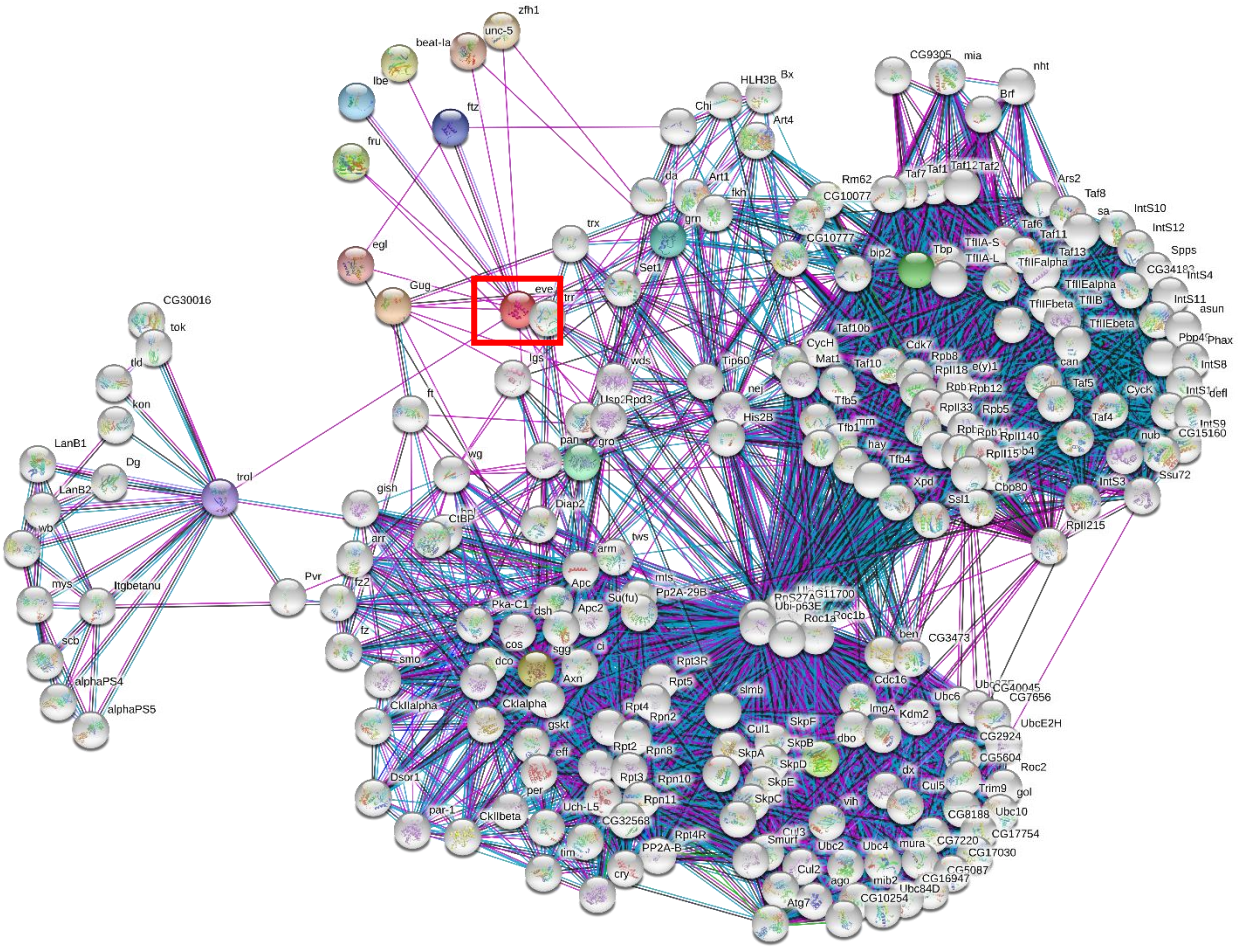
disso, foi visto que *eve* realiza suas funções de 2 a 3 horas após a postura do embrião (Goto *et al.,* 1989).

Estudos mais recentes identificaram *eve* já nos primeiros ciclos de clivagem do embrião (Ali-Murthy et al., 2013). Pelo motivo de ser um importante gene *pair-rule*, desempenhando um papel importante na segmentação dos embriões, e também por estar sendo expresso nos primeiros ciclos mitóticos embrionários, *eve* foi escolhido para análises de redes topológicas.

Kruppel (Kr) é um gene *gap,* que se expressa no centro do embrião. Ele é um dos responsáveis por subdividir o embrião ao longo do eixo antero-posterior, criando uma região na qual futuramente irão se expressar os genes *pair-rule*. Além disso, Kr é responsável pelo destino das células nos Túbulos de Malpighi (Hoch, 1994).

As evidências disponíveis têm mostrado que *Kruppel* pode funcionar adicionalmente como um ativador da expressão gênica. *Kruppel* se liga à sequência de consenso AAAAC / GGGGTTAA (Rosenberg *et al.,* 1986). O domínio de *zinc finger* da proteína é formado por dois domínios trans-repressores evolutivamente conservados (TR1 N-terminal - domínio trans-repressor 1) e C64 (domínio repressor C-terminal) e um único domínio trans-ativador fraco (TA1). Os dois domínios repressores de *Kr* demonstraram controlar a expressão do gene *pair-rule hairy* (La Rosee-Borggreve *et al.,* 1999). Estudos dos ciclos iniciais de clivagem em *D.melanogaster*, reportam que Kr é expresso nos ciclos iniciais de clivagem (Ali-Murthy et al., 2013).

Rede inicial de eve (Figura 21) e Kr (Figura 22) recuperada do banco de dados STRING.


zfh1
unc-5
beat-Ia
lbe
ftz
fru
HLH3B
Bx
Chi
Art4
CG9305
mia
nht
Brf
Rm62
CG10077
Taf7
Taf1
Taf12
Taf2
Ars2
Taf8
sa
IntS10
IntS12
Spps
CG34162
IntS4
IntS11
asun
Pbp4
Phax
IntS8
defl
IntS9
CG15160
nub
Ssu72
IntS3
Rpll215
Rpll140
Rpb5
Rpb12
Rpb8
Rpll18
Cdk7
CycH
Mat1
Taf10b
Taf10
Tfb5
Tfb1
Tfb4
Xpd
Cbp80
Ssl1
Taf5
Taf4
CycK
Tbp
TfIIA-S
TfIIA-L
TfIIEalpha
TfIIEbeta
TfIIFbeta
TfIIB
blp2
Tip60
His2B
Rpd3
Usp2
Set1
trx
Gug
egl
eve
CG30016
tok
tld
kon
LanB1
LanB2
Dg
trol
Pvr
htgbetanu
mys
scb
alphaPS4
alphaPS5
gish
CtBP
Diap2
wg
ft
arr
fz2
fz
smo
Cklalpha
Dsor1
par-1
Ckllbeta
Pka-C1
dsh
Apc
Apc2
Su(fu)
Pp2A-29B
cos
dco
sgg
ci
Axn
Rpt3R
Rpt5
Rpt4
Rpn2
Rpt2
Rpn8
Rpt3
Rpn10
Rpn11
Uch-L5
CG32568
Rpt4R
PP2A-B-55
slmb
Cul1
SkpF
SkpB
SkpA
SkpD
SkpE
SkpC
dbo
ben
CG3473
Cdc16
Ubc6
CG40045
CG7656
UbcE2H
ImgA
Kdm2
CG2924
CG5604
Roc2
Cul5
Trim9
gol
CG8188
Ubc10
CG17754
CG7220
CG17030
CG5087
CG16947
Ubc84D
CG10254
Atg7
Cul2
ago
mib2
Ubc2
Ubc4
mura
Smurf
Cul3
vih
dx
Roc1a
Roc1b
Ubi-p63E
Rbx1
CG11700
Rpn6
Rpn7
Rpn5
Rpn9
Rpn12
Rpn3
Rpn1
Rpn2
Rpn11
Rpn8
Rpn10

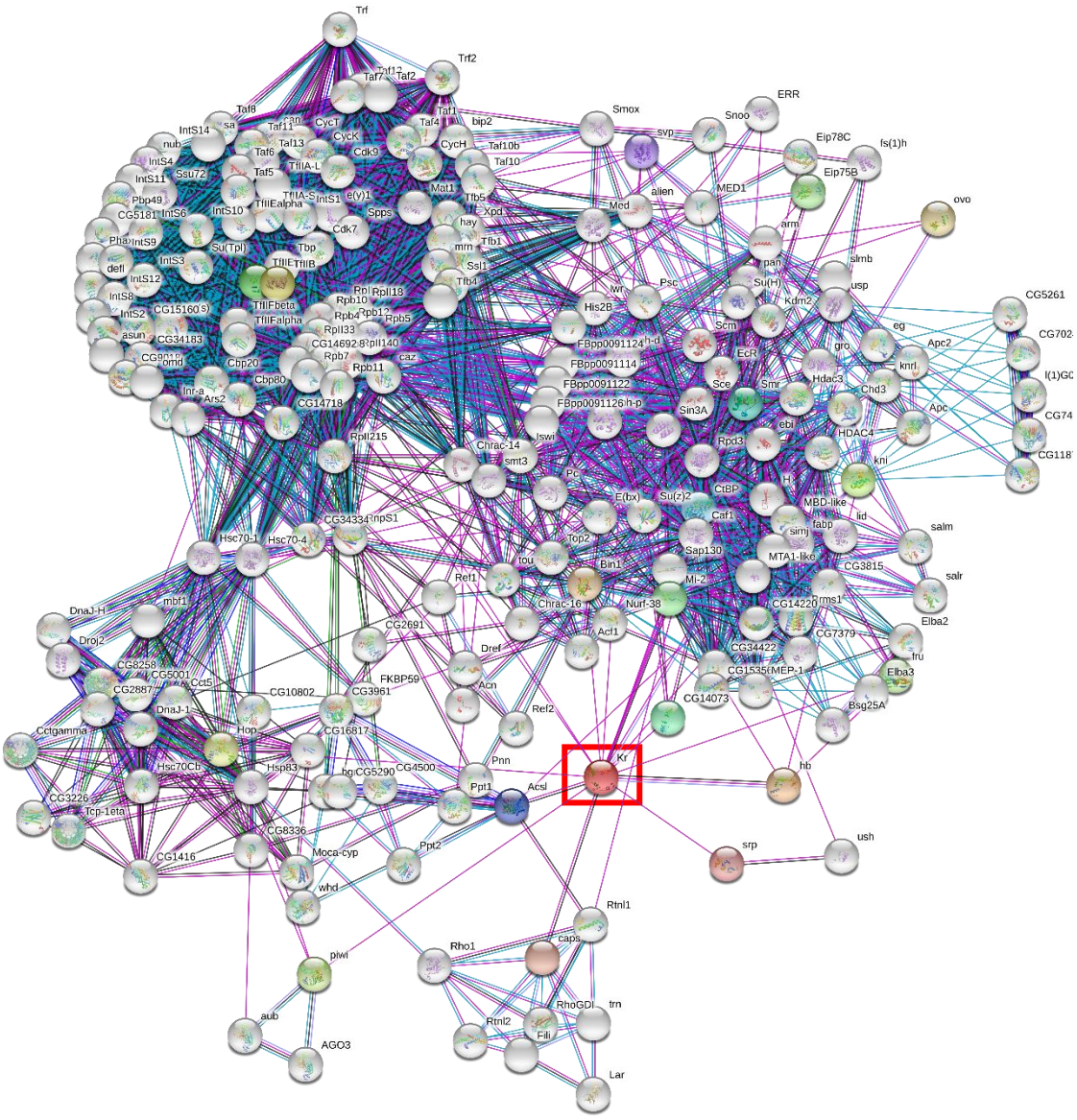


Figura 22: Rede topológica da proteína Kruppel (Kr). Rede criada pelo banco de dados STRING mostra a interação física entre a proteína Kr (retângulo vermelho) e as demais proteínas que constituem a rede.

Nessas redes, é possível identificar centenas de ligações às proteínas eve e Kr, sendo que muitas dessas relações não serão utilizadas em nossas análises. Para filtrar melhor as relações, foi utilizado o software *Cytoscape*.

4.6.2- Análise das interações diretas de eve e Kr.

O gene *eve* está presente em várias fases da embriogênese, e por ser um gene regulador, está envolvido na regulação de vários genes. Na tentativa de entender um pouco da complexidade das interações da proteína eve, uma rede topológica de interações diretas (vizinhos diretos) foi criada pelo *software* Cytoscape (figura 23).

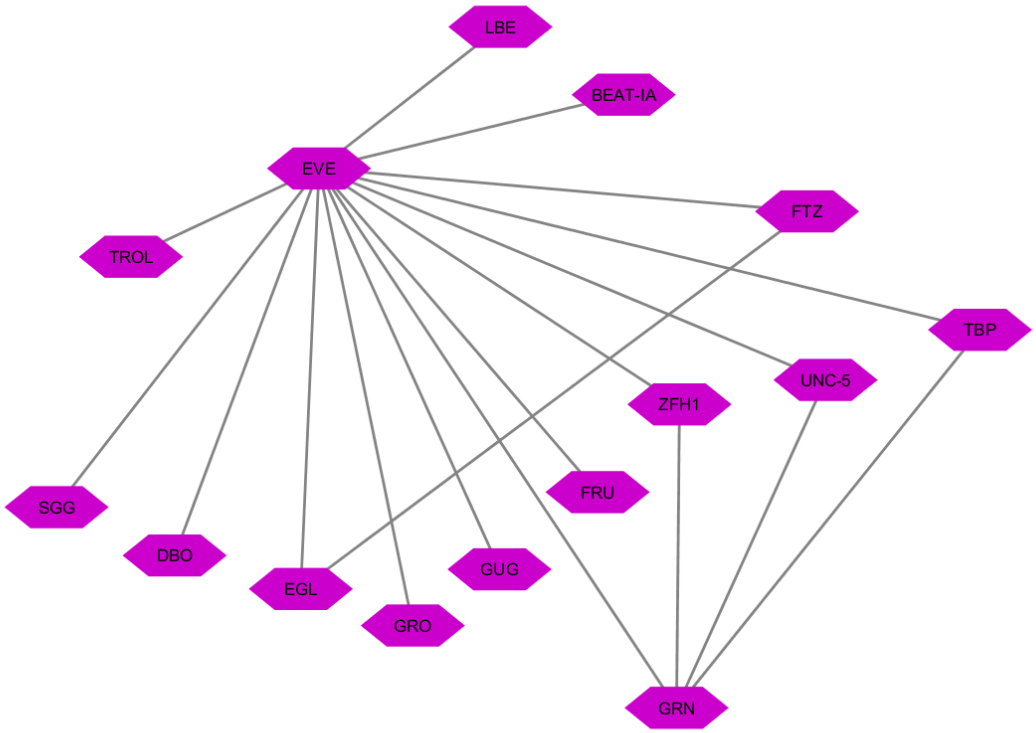


Figura 23: Rede topológica dos vizinhos diretos de eve. As proteínas presentes na figura possuem interação direta com a proteína eve. terribly reduced optic lobes (trol), shaggy (sgg), diablo (dbo), egalitarian (egl), groucho (gro), grunge (gug), fruitless (fru), grain (grn), zinc finger homeodomain 1(zfh1), unc-5 (unc-5), TATA binding protein (Tbp), fushi tarazu (ftz), beaten path Ia (beat-Ia), ladybird early (lbe).

As proteínas que possuem interação direta com eve são: terribly reduced optic lobes (trol), shaggy (sgg), diablo (dbo), egalitarian (egl), groucho (gro), grunge (gug), fruitless (fru), grain (grn), zinc finger homeodomain 1(zfh1), unc-5 (unc-5),

TATA binding protein (Tbp), fushi tarazu (ftz), beaten path Ia (beat-Ia), ladybird early (lbe). Essas proteínas foram identificadas para *D. melanogaster*. Os homólogos dessas proteínas em *A. mellifera* estão presentes na tabela 5.

Tabela 5: Proteínas homólogas as vizinhas de eve de *D. melanogaster* encontras em *A. mellifera*. De todas as proteínas listadas que fazem contato físico direto com eve, apenas zinc finger homeodomain 1 (zfh1) não possui homólogo em *A. mellifera*. As buscas pelos homólogos foi feita pela ferramenta Blastp, disponível no National Center for Biotechnology Information (NCBI).

Organismo / Genes	*D. melanogaster*_XP	*FlyBase ID*	*A. mellifera*_XP	BeeBase ID
trol	NP_001245494.2	FBgn0284408	XP_006569679.1	GB44798
sgg	NP_001259193.1	FBgn0003371	XP_006567426.1	GB53394
dbo	NP_001261892.1	FBgn0040230	XP_006558840.1	GB48469
egl	NP_726360.3	FBgn0000562	XP_026298567.1	GB44828
gro	NP_001247309.1	FBgn0001139	XP_006565473.1	GB41521
gug	NP_523973.2	FBgn0010825	XP_016771698.1	GB54092
fru	NP_524397.2	FBgn0004652	XP_006560812.2	GB17617
grn	NP_731211.1	FBgn0001138	XP_016769155	GB41785
zfh1	NP_733401.3	FBgn0004606	-	-
unc-5	NP_001286450.1	FBgn0034013	XP_006558291.2	GB47918
Tbp	NP_523805.1	FBgn0003687	XP_623088.2	GB47180
ftz	NP_477498.1	FBgn0001077	XP_006557455.1	GB42142

Organismo / Genes	*D. melanogaster*_XP	*FlyBase ID*	*A. mellifera*_XP	BeeBase ID
Beat-Ia	NP_001285975.1	FBgn0013433	XP_026299298.1	GB53593
lbe	NP_524435.2	FBgn0011278	XP_016773403.2	GB47490

Com base nas buscas pelos homólogos, foi possível identificar homólogos de todas as proteínas de *D. melanogaster* em *A. mellifera*, com exceção de zfh1, para a qual não foi encontrado homólogo.

Para saber as funções das proteínas relacionadas com eve foi feita uma análise de GO pelo banco de dados *FlyBase* (Tabela 5). Com exceção do gene TATA binding protein (Tbp), todos os demais possuem funções importantes e essenciais durante o desenvolvimento. Outra função que se destacou por estar presente na maioria dos genes analisados foi a de organização celular e biogênese, cujos processos estão principalmente presentes durante a embriogênese.

Tabela 6: Processos biológicos da rede de vizinhos diretos de eve. Os processos biológicos dos genes que se ligam diretamente a eve foram retirados do banco de dados Flybase. Por essa tabela, é possível ver que todos os genes participam do processo de desenvolvimento, com exceção da proteína tbp. O processo de organização celular e biogênese também está presente em quase todos as proteínas, com exceção de gro e ftz. Retângulos roxos indicam os processos biológicos de genes que possuem homólogos em *A. mellifera*, retângulos cinzas representam aqueles que não possuem homólogos em *A. mellifera*.

Genes	Org. cel./biogênese	Desenvolvimento	Resp. ao estímulo	Sinalização	Ciclo cel./ proliferação	Expressão gênica	Transporte/ localização	Reprodução	Proc. Sist. nervoso	Comportamento	Metabolismo de proteínas
lbe											
trol											
sgg											
dbo											
egl											
gro											
gug											

	Org. cel./biogênese	Desenvolvimento	Resp. ao estímulo	Sinalização	Ciclo cel./ proliferação	Expressão gênica	Transporte/ localização	Reprodução	Proc. Sist. nervoso	Comportamento	Metabolismo de proteínas
fru											
grn											
zfh1											
unc-5											
Tbp											
ftz											
beat-1a											

A proteína Kruppel (Kr) também teve seus vizinhos diretos selecionados e analisados pelo software *Cytoscape* (Figura 24). As proteínas que possuem interação direta com Kr são: P-element induced wimpy testis (piwi), hopscotch (hop), Acyl-CoA synthetase long-chain (Acsl), Transcription factor IIB (TfIIB), Transcription factor IIEbeta (TfIIEbeta), Bicoid interacting protein 1 (Bin1), seven up (svp), CG14073, Smrter (Smr), ovo (ovo), fruitless (fru), hunchback (hb), ecdysone-induced protein 75B (Eip75B), serpent (srp), capricious (caps), Knirps (kni), C-terminal binding protein (CtBP), Mi-2 (Mi-2).

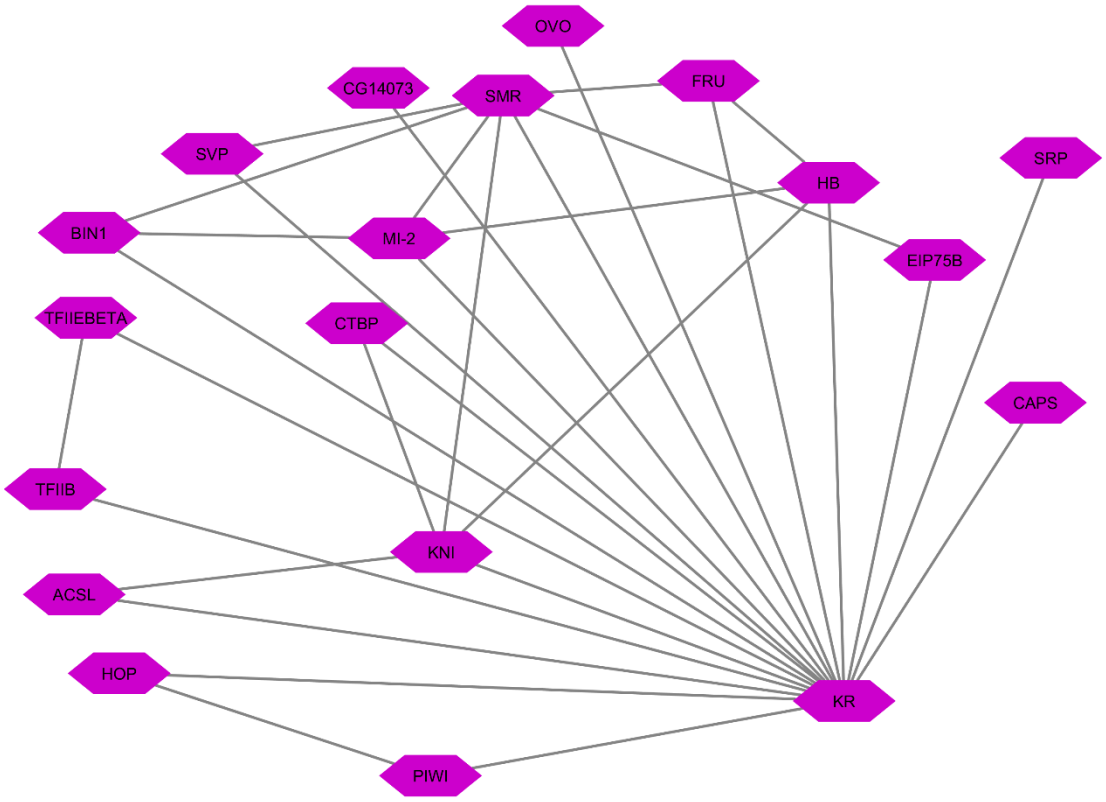


Figura 24: Rede topológica dos vizinhos diretos de Kr. As proteínas presentes na figura possuem interação direta com a proteína Kr.

Com o propósito de encontrar os homólogos em *A. mellifera* das proteínas listadas da rede de Kr, foi feita uma busca no banco de dados NCBI e *BeeBase*

(Tabela 7). Os resultados encontrados mostram que as proteínas Bin1, CG14073, srp e caps não possuem homólogos em *A. mellifera.*

Tabela 7: Proteínas homólogas as vizinhas de Kr de *D. melanogaster* encontras em *A. mellifera*. De todas as proteínas listadas que fazem contato físico direto com Kr, Bicoid interacting protein 1 (Bin1), CG14073, serpente (srp) e capricious (caps) não possuem homólogos em *A. mellifera*. As buscas pelos homólogos foi feita pela ferramenta Blastp, disponível no National Center for Biotechnology Information (NCBI).

Organismo / Genes	*D. melanogaster*_XP	*FlyBase ID*	*A. mellifera*_XP	BeeBase ID
piwi	NP_476875.1	FBgn0004872	XP_026300549.1	GB49909
hop	NP_511119.2	FBgn0004864	XP_006567752.1	GB44594
Acsl	NP_001014511.1	FBgn0263120	XP_006561670.1	GB42732
TfIIB	NP_001260349.1	FBgn0004915	XP_006569641.1	GB44780
TfIIEbeta	NP_001163346.1	FBgn0015829	XP_026297680.1	GB40984
Bin1	NP_001303441.1	FBgn0024491	-	-
svp	NP_731682.1	FBgn0003651	XP_026296916.1	GB17100
CG14073	NP_001262022.1	FBgn0036814	-	-
Smr	NP_727634.2	FBgn0265523	XP_026297043.1	GB53206
ovo	NP_001162673.1	FBgn0003028	XP_026301718.1	GB51354
fru	NP_524397.2	FBgn0004652	XP_006560812.2	GB17617
hb	NP_731267.1	FBgn0001180	NP_001239355.1	GB50986
Eip75B	NP_001246822.1	FBgn0000568	NP_001073579.1	GB47224

Organismo / Genes	*D. melanogaster*_XP	*FlyBase ID*	*A. mellifera*_XP	BeeBase ID
srp	NP_001247128.1	FBgn0003507	-	-
caps	NP_524056.3	FBgn0023095	-	-
kni	NP_524187.1	FBgn0001320	XP_026298877.1	GB51231
CtBP	NP_001014617.1	FBgn0020496	XP_006557700.1	GB43266
mi-2	NP_649154.2	FBgn0262519	XP_016768938.1	GB41803

Para saber as funções das proteínas relacionadas com Kr, foi feita uma análise de GO pelo banco de dados *FlyBase* (Tabela 7). A maioria das proteínas listadas estão envolvidas nos processos de desenvolvimento, expressão gênica e organização celular/biogênese.

Os dados coletados com as redes topológicas das proteínas eve e Kr são de caráter exploratório, visando dar perspectivas para futuros estudos sobre o que ocorre no cenário da embriogênese inicial de *A. mellifera.* Assim, conhecimentos prévios sobre as interações entre proteínas, mRNAs e miRNAs serão muito úteis para dar encaminhamento para as próximas fases. Estes resultados preliminares corroboram nossos achados ao analisar a expressão destes genes durantes as fases iniciais do desenvolvimento embrionário de *A. melífera*. Ou seja, as funções das proteínas associadas são essenciais para a organização e ritmo das divisões celulares nos estágios iniciais do desenvolvimento.

Tabela 8: Processos biológicos da rede de vizinhos diretos de Kr. Os processos biológicos dos genes que se ligam diretamente a Kr foram retirados do banco de dados Flybase. Retângulos roxos indicam os processos biológicos de genes que possuem homólogos em *A. mellifera*, retângulos cinzas representam aqueles que não possuem homólogos em *A. mellifera*.

Genes	Org. cel./biogênese	Desenvolvimento	Resp. ao estímulo	Sinalização	Ciclo cel./ proliferação	Expressão gênica	Transporte/ localização	Reprodução	Proc. Sist. nervoso	Comportamento	Metabolismo de proteínas	Sistema imune
piwi												
hop												
Acsl												
TfIIB												
TfIIEbeta												
Bin1												
svp												

Genes	Org. cel./biogênese	Desenvolvimento	Resp. ao estímulo	Sinalização	Ciclo cel./ proliferação	Expressão gênica	Transporte/ localização	Reprodução	Proc. Sist. nervoso	Comportamento	Metabolismo de proteínas	Sistema imune
CG14073												
Smr												
ovo												
fru												
hb												
Eip75B												
srp												
caps												
kni												
CtBP												
mi-2												

5- CONCLUSÃO

Até o presente momento, os dados obtidos indicam que a transcrição zigótica se inicia nos primeiros estágios do desenvolvimento embrionário de *A. mellifera,* sendo detectado pré-mRNAs dos genes *RPL32, Takeout, Troponin C Type IIb*, *Engrailed, Even-skipped, Hairy, Kruppel* e *Odd-skipped* após 30 minutos após a postura. As imagens obtidas dos núcleos de embriões mostram que, durante a primeira hora após a postura, provavelmente já ocorreram ao menos 2 ciclos mitóticos, sendo possível identificar até 3 núcleos nos embriões. Além disso, as redes topológicas de proteínas cujos mRNAs se expressam muito precocemente, justificam, ainda que sejam experimentos preliminares, os achados deste trabalho.

6- REFERÊNCIA BIBLIOGRÁFICAS

Altschul S. F.; Gish W.; Miller W.; Myers E. W.; Lipman D. J. Basic local alignment search tool. *Journal of Molecular Biology*, v. 215, p. 403-410, 1990.

Aboobaker A. A., Tomancak P., Patel N., Rubin G. M., Lai E. C. Drosophila microRNAs exhibit diverse spatial expression patterns during embryonic development. *Proceedings of the National Academy of Sciences of the United States of America*, v. 102, p. 18017–18022, 2005.

Ali-Murthy Z., Lott S. E., Eisen M. B., Kornberg T. B. An essential role for zygotic expression in the pre-cellular Drosophila embryo. *PLoS Genetics*, doi: 10.1371/journal.pgen.1003428, 2013.

Arbeitman M. N. Gene Expression During the Life Cycle of Drosophila melanogaster. *Science,* v. 297, p. 2270-2275, 2002

Anderson K. V., Lengyel J. A. Rates of synthesis of major classes of RNA in Drosophila embryos. *Developmental Biology*, v. 70, p. 217–231, 1979.

Applied Biosystems. User Bulletin #2: ABI Prism 7700 Sequence Detection System. Applied Biosystems Inc., Foster City, CA. Disponível online: docs.appliedbiosystems.com/pebiodocs/04303859.pdf, 1997.

Arsala D.; Lynch J. A. Ploidy has little effect on timing early embryonic events in the haplo-diploid wasp *Nasonia*. *Genesis.* v. 55:e23029. https://doi.org/10.1002/dvg.23029*,* 2017.

Benoit B.; He C.; Zhang F.; Votruba S.; Tadros W.; Westwood J.; Smibert C.; Lipshitz H.; Theurkauf W. An essential role for The RNA-binding protein Smaug during the D. melanogaster maternal-to-zygotic transition. *Devolopment,* v. 136, p. 923-55, 2009.

Biemar F., Zinzen R., Ronshaugen M., Sementchenko V., Manak J. R., Levine M. S. Spatial regulation of microRNA gene expression in the Drosophila embryo. *Proceedings of the National Academy of Sciences of the United States of America,* v. 102, p. 15907–15911, 2005.

Bushati N., Stark A., Brennecke J., Cohen S. M. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in Drosophila. *Current Biology*, v.18, p. 501–506, 2008.

Campos-Ortega J. A., Hartenstein V. The embryonic development of Drosophila melanogaster. Springer-Verlag, Berlin; New York, 1997.

Darbo E., Hermann C., Lecuit T., Thieffry D., van Helden J. Transcriptional and epigenetic signatures of zygotic genome activition during early D. melanogaster embryogenesis. *BMC Genomics,* v. 14, p. 226-230, 2013.

De Las Rivas J., Fontanillo C. Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Computational Biology*. DOI: 10.1371/journal.pcbi.1000807, 2010.

De Renzis S.; Elemento O.; Travzoie S. Wieschaus E. F. Unmasking activation of the zygotic genome using chromosomal deletions in the D. melanogaster embryos. *PLoS Biol.* https://doi.org/10.1371/journal.pbio.0050117, 2007.

Dearden P. K., Wilson M. J., Sablan L., Osborne P. W., Havler M., McNaughton E., Kimura K., Milshina N. V., Hasselmann M., Gempe T., Schioett M., Brown S. J., Elsik C. G., Holland P. W., Kadowaki T., Beye M. Patterns of conservation and change in honey bee developmental genes. *Genome Res,* v. 16, p. 1376-84, 2006.

Dearden P., Duncan E., Wilson M. Fixation and storage of honeybee (*Apis mellifera*) tissues. *Cold Spring Harbor Protocols*, 2009a.

Dearden P., Duncan E., Wilson M. Whole-Mount In situ Hybridization of honeybee (*Apis mellifera*) tissues*. Cold Spring Harbor Protocols*, 2009b.

Edgar B. A., Schubiger G. Parameters controlling transcriptional activation during early Drosophila development. *Cell*, v. 44, p. 871–877, 1986.

Edgar R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, v. 32, p. 1792-1797, 2004.

Finn R. D., Bateman A., Clements J., Coggill P., Eberhardt R. Y., Eddy S. R., Punta M. Pfam: the protein families database. *Nucleic acids research*, v. *42*(Database issue), p. 222- 230, 2014.

Fleig R., Sander K. Embryogenesis of the honeybee Apis mellifera l. (Hymenoptera: Apidae): An sem study. *International Journal of Insect Morphology and Embryology*, v. 15, p. 449-62, 1986.

Foe V. E., Alberts B. M. Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in Drosophila embryogenesis. *Journa ofl Cell Scence,* v. 61, p. 31–70, 1983.

Fonseca R. N., Venâncio T. M. Maternal or zygotic: Unveiling the secrets of the Pancrustacea transcription factor zelda. PLoS Genet DOI:10.1371/journal.pgen.1007201, 2018.

Frasch M., Hoey T., Rushlow C., Doyle H., Levine M. Characterization and localization of the even-skipped protein of Drosophila. *EMBO Journal,* v. 6, p. 749-759, 1987.

Fu S., Nien C. Y., Liang H. L., Rushlow C. Co-activation of microRNAs by Zelda is essential for early Drosophila development. *Development*, v. 141, p. 2108–2118, 2014.

Goto T., MacDonald P., Maniatis T. Early and late periodic patterns of even skipped expression are controlled by distinct regulatory elements that respond to different spatial cues. *Cell*, v. 57 p. 413 - 422, 1989.

Gotze M., Wahle E. Smaug destroys a huge treasure. *Genome Biology*, v. 15, p. 101-103, 2014.

Hara M., Petrova B., Orr-Weaver T. L. Control of PNG kinase, a key regulator of mRNA translation, is coupled to meiosis completion at egg activation. *eLife.* DOI: 10.7554/eLife.22219, 2017.

Harrison M. M., Li X. Y., Kaplan T., Botchan M. R., Eisen M. B. Zelda binding in the early D. melanogaster melanogaster embryo marks regions subsequently activated atthematernal-to-zygotic transition. *PLoS Genetics*, DOI: 10.1371/journal.pgen.1002266, 2011.

Heyn P., Kircher M., Dahl A., Kelso J., Tomancak P., Kalinka A. T., et al. The earliest transcribed zygotic genes are short, newly evolved, and different across species. *Cell Reports*, v. 6, p. 285–292, 2014.

Hoch M., Broadie K., Jäckle H., Skaer H. Sequential fates in a single cell are established by the neurogenic cascade in the Malpighian tubules of Drosophila. *Development*, v. 120, p. 3439-3450, 1994.

Honeybee Genome Sequencing Consortium. Insights into social insects from the genome of honeybee *Apis melliera*. *Nature*, v. 443, p. 931-949, 2006.

Horner V., Wolfner M. Transitioning from egg to embryo: triggers and mechanisms of egg activation. *Devlopmental Dynamics*, v. 237, p. 527-71, 2008.

Ingham P. W. The molecular genetics of embryonic pattern formation in Drosophila. *Nature,* v. 335, p. 25-34, 1988.

Ji J. Y., Squirrell J. M., Schubiger G. Both cyclin B levels and DNAreplication checkpoint control the early embryonic mitoses in Drosophila. *Development*, v. 131, p. 401–411, 2004.

Johnson M., Zaretskaya I., Raytselis Y., Merezhuk Y., McGinnis S., Madden T. L. NCBI BLAST: a better web interface. Nucleic Acids Research, v. 36:W5-W9, 2008.

Kanodia J. S., Liang H. L., Kim Y., Lim B., Zhan M., Lu H., *et al*. Pattern formation by graded and uniform signals in the early Drosophila embryo. *Biophysical Journal,* v. 102, p. 427–433, 2012.

Kapheim K. M., Pan H., Li C., Wang J., Robinson G. E., Zhang G. Genomic signatures of evolutionary transitions from solitary to group living. *Science*, v. 348, p. 1138-1144, 2015.

Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, v. 33, p. 1870-1874, 2016.

Kronja I., Whitfield Z. J., Yuan B., Dzeyk K., Kirkpatrick J., Krijgsveld J., Orr-Weaver T. L. Quantitative proteomics reveals the dynamics of protein changes during Drosophila oocyte maturation and the oocyte-to-embryo transition. *PNAS* v. 111, p. 16023–16028, 2014.

La Rosée-Borggreve A., Häder T., Wainwright D., Sauer F., Jäckle H. hairy stripe 7 element mediates activation and repression in response to different domains and levels of Krüppel in the Drosophila embryo. *Mechanisms of Development,* v. 89, p. 133-140, 1999.

Lamb M. M., Laird C. D. Increase in nuclear poly(A)-containing RNA at syncytial blastoderm in Drosophila melanogaster embryos. *Developmental Biology*, v. 52, p. 31–42, 1976.

Lecuyer E., Yoshida H., Parthasarathy N., Alm C., Babak T., Cerovina, T., et al. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell,* v.131, p. 174–187, 2007.

Li X. Y., Harrison M. M., Villalta J. E., Kaplan T., Eisen M. B. Establishment of regions of genomic activity during the maternal to zygotic transition. *eLife,* v. 3, 3, 2014.

Liang H. L., Nien C. Y., Liu H. Y., Metzstein M. M., Kirov N., Rushlow C. The zinc-finger protein zelda is a key activator of the early zygotic genome in D. melanogaster. *Nature*, v. 456, p. 400-403, 2008.

Livak K. J., Schmittgen T. D. Analysis of Relative Gene Expression Data Using Real Time Quantitative PCR and the 22DDCT Method. *METHODS*, v. 25, p. 402–408, 2001.

Lott S. E., Villalta J. E., Schroth G. P., Luo S., Tonkin L. A., Eisen M. B. Noncanonical compensation of the zygotic X transcription in early D. melanogaster melanogaster development revealed through single-embryo RNA-seq. *PLoS Biology*, DOI:10.1371/journal.pbio.1000590, 2011.

Lourenço A., P., Mackert A., Cristino A. S., Simões, Z. L. P. Validation of reference genes for gene expression studies in the honey bee, *Apis mellifera*, by quantitative real-time RT-PCR. *Apidologie,* v. 39, p. 372-85, 2008.

Lund E., Liu M., Hartley R. S., Sheets M. D., Dahlberg J. E. Deadenylation of maternal mRNAs mediated by miR-427 in Xenopus laevis embryos. *RNA society*, v. 15, p. 2351–2363, 2009.

Maere S., Heymans K., Kuiper M., BiNGO: a cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. *Bioinformatics*, v. 21, p. 3448-3449, 2005.

Newport J., Kirschner M. A major developmental transition in early Xenopus embryos. I. characterization and timing of cellular changes at the midblastula stage. *Cell,* v. 30, p. 675–686, 1982a.

Newport J., Kirschner M. A major developmental transition in early Xenopus embryos. II. Control of the onset of transcription. *Cell,* v.30, p. 687–696, 1982b.

Nien C.Y., Liang H. L., Butcher S., Sun Y., Fu S., Gocha T., Kirov N., Manak J. R., Rushlow C. Temporal coordenation of gene networks by Zelda in the early D. melanogaster embryo. *PloS Genetics,* DOI: 10.1371/journal.pgen.1002339, 2011.

Nüisslein-Volhard C., Wieschaus E. Mutations affecting segment number and polarity in Drosophila. *Nature*, v. 287, p. 795-801, 1980.

Nüsslein-Volhard C., Kluding H., Jürgens G. Genes affecting the segmental subdivision of the Drosophila embryo. *Cold Spring Harbor Symposia on Quantitative Biology,* v. 50, p .145-154, 1985.

Oliveira F. R. Análise do fator de transcrição Zelda durante o desenvolvimento de *Apis mellifera.* 76. Trabalho de conclusão de curso. Universidade de São Paulo, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto- Departamento de Biologia, 2017.

Osborne P. W., Dearden P. K. Expression of Pax group III genes in the honeybee (Apis mellifera). *Development Genes Evolution,* v. 215, p. 499-508, 2005.

Pires C. V. Regulação gênica dos processos iniciais do desenvolvimento de embriões haploides e diploides de *Apis mellifera.* 152. Tese de doutorado. Universidade de São Paulo, Faculdade de medicina de Ribeirão Preto-Departamento de genética, 2014.

Pires C. V. Freitas F. C. P., Cristino A. S., Dearden P. K., Simões Z. L. P. Transcriptome Analysis of Honeybee (*Apis Mellifera*) Haploid and Diploid Embryos Reveals Early Zygotic Transcription during Cleavage. *PLoS One*, DOI:10.1371/journal.pone.0146447, 2016.

Pritchard D. K., Schubiger G. Activation of transcription in Drosophila embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes Development,* v. 10, p. 1131–1142, 1996.

Rehan S. M., Glastad K. M., Lawson S. P., Hunt B. G. The genome and Methylome of a subsocial small Carpenter bee, Ceratina Calcarata. *Genome Biology and Evolution,* v. 8, p. 1401-1410, 2016.

Ribeiro L., Tobias-Santos V., Santos D., Antunes F., Feltran G., de Souza Menezes J., et al. Evolution and multiple roles of the Pancrustacea specific transcription factor *zelda* in insects. *PLoS Genet*ics DOI:10.1371/journal.pgen.1006868, 2017.

Robertson S., Lin R. The oocyte-to-embryo transition. Advances in Experimental Medicine and Biology 757: 351–372, 2013.

Romiguier J., Cameron S. A., Woodard S. H, Fischman B. J., Keller L., Praz C. J. Phylogenomics controlling for base compositional bias reveals a single origin of eusociality in corbiculate Bees. *Molecular Biology Evolution*, v. 33, p. 670-678, 2016.

Rosenberg U. B., Schroeder C., Preiss A., Kienlin A., Cote S., Riede I., Jäckle H. Structural homology of the product of the Drosophila Kruppel gene with Xenopus transcription factor IIIA. *Nature,* v. 319, p. 336-339, 1986.

Saitou N., Nei M. The neighbor-joining method: A new method for reconstruction phylogenetic trees. *Molecular Biology and Evolution*, v. 4, p. 406-425, 1987.

Sartain C. V., Wolfner M. F. Calcium and Egg Activation in D. melanogaster. *Cell Calcium*, v. 53, p. 10–15, 2013.

Sasaki K., Obara Y. Egg activation and timing of sperm acceptance by an egg in honeybees (*Apis mellifera L*.). *Insectes Sociaux*, v. 49, p. 234-240, 2002.

Schier A. F. The maternal-zygotic transition: death and birth of RNAs. *Science,* v. 316, p. 406–407, 2007.

Shannon P., Markiel A., Ozier O., Baliga N. S., Wang J. T., Ramage D., Amin N., Schwikowski B., Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research,* v. 13, p. 498-504, 2003.

Shermoen A. W., O'Farrell P. H. Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. *Cell,* v. 67, p. 303–310, 1991.

Shermoen A. W., McCleland M. L., O'Farrell P. H. Developmental control of late replication and S phase length. *Current Biology,* v. 20, p. 2067–2077, 2010.

Siddiqui N. U., Li X., Lou H., Karaiskakis A., Hou H., Kislinger T., Westwood J. T., Morris Q., Lipshitz H. D. Genome-wide analysis of the maternal-to-zygotic transition in D. melanogaster primordial germ cells. *Genome Biology*, v. 13, R11, 2012.

Smibert C. A., Lie Y. S., Shillinglaw W., Henzel W. J., Macdonald P. M. Smaug, a novel and conserved protein, contributes to repression of nanos mRNA translation in vitro. *RNA society,* v. 5, p. 1535–1547, 1999.

Sung H. W., Spangenberg S., Vogt N., Grosshans J. Number of nuclear divisions in the D. melanogaster blastoderm controlled by onset of zygotic transcription. *Current Biology,* v. 23, p. 133-138, 2013.

Szklarczyk D., Franceschini A., Wyder S., Forslund K., Heller D., Huerta-Cepas J., Simonovic M., Roth A., Santos A., Tsafou K. P, Kuhn M., Bork P., Jensen L. J., Mering C. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, v. 43, p. 447-452, 2015.

Tadros W., Houston S. A., Bashirullah A., Cooperstock R. L., Semotok J. L., Reed B. H., Lipshitz H. D. Regulation of Maternal Transcript Destabilization During Egg Activation in D. melanogaster. *Genetics,* v. 1001, p. 989–1001, 2003.

Tadros W., Goldman A. L., Babak T., Menzies F., Vardy L., Orr-Weaver T., Hughes T. R., Westwood J. T., Smibert C. A., Lipshitz H. D. SMAUG is a major regulator of maternal mRNA destabilization in D. melanogaster and its translation is activated by the PAN GU kinase. *Developmental Cell,* v. 12, p. 143-55, 2007.

Tadros W. Lipshitz H. D. The maternal- to- zygotic transition: a play in two acts. *Development*, v. 36, p. 3033-3042, 2009.

ten Bosch J. R., Benavides J. A., Cline T. W. The TAGteam DNA motif controls the timing of D. melanogaster pre-blastoderm Transcription. *Development,* v. 133, p. 1967-1977, 2006.

Vardy L., Orr-Weaver T. Regulating translation of maternal messages: Multiple repression mechanisms. *Trends in Cell Biology,* v. 17, p. 547-554, 2007.

Von Stetina J. R., Orr-Weaver T. L. Developmental control of oocyte maturation and egg activation in metazoan models. *Cold Spring Harbor Perspectives in Biology*, DOI: 10.1101/cshperspect.a005553, PMID: 21709181, 2011.

Townley I. K., Roux M. M., Foltz K. R. Signal transduction at fertilization: the Ca2+ release pathwayin echinoderms and other invertebrate deuterostomes. *Seminars in Cell and Developmental Biology,* v. 17, p. 293–302, 2006.

Wheeler W. C., Whiting M., Wheeler Q. D., Carpenter J. M. The phylogeny of the extant hexapods orders. *Cladistics*, v. 17, p. 113-169, 2001.

Whitaker M. Calcium signalling in early embryos. *Philosophical Transactions of the Royal Society B: Biological Sciences,* v. 363, p. 1401–1418, 2008.

Zuckerkandl E., Pauling L. Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H. J. Vogel, 97-166, *Academic Press,* New York, 1965.

ANEXO A

RpL32 - Primer F POS

PREDICTED: Apis mellifera ribosomal protein L32 (RpL32), transcript variant X3, mRNA

Sequence ID: XM_016914656.2 **Length:** 601 **Number of Matches:** 1

Range 1: 397 to 428 GenBank Graphics ▼ Next Match ▲ Previous

Score	Expect	Identities	Gaps	Strand
41.0 bits(44)	0.003	28/32(88%)	0/32(0%)	Plus/Minus

```
Query  44   TGTGTTGCATCATTAATACTTCCAGTGCCGTG  75
            ||| ||||||||||| ||||||||| || ||
Sbjct  428  TGTTTTGCATCATTAAAACTTCCAGTTCCTTG  397
```

RpL32 - Primer F PRS

PREDICTED: Apis mellifera ribosomal protein L32 (RpL32), transcript variant X3, mRNA

Sequence ID: XM_016914656.2 **Length:** 601 **Number of Matches:** 1

Range 1: 413 to 499 GenBank Graphics ▼ Next Match ▲ Previou

Score	Expect	Identities	Gaps	Strand
120 bits(132)	4e-27	81/87(93%)	3/87(3%)	Plus/Plus

```
Query  6    TTTATGATGCAAA-CAGACA-TTTTGTGCTGAAATTGCTCATGGTGTGAGCAGTAAAAAA  63
            || |||||||||| |||| | |||||||||||||||||||||||||||||||||||||||
Sbjct  413  TTAATGATGCAAAACAGAAAATTTTGTGCTGAAATTGCTCATGGTGTGAGCAGTAAAAAA  472

Query  64   CGT-TATCCATTGTTGAACGTGCTCAA  89
            |||  |||||||||||||||||||||
Sbjct  473  CGTAAATCCATTGTTGAACGTGCTCAA  499
```

RpL32 - Primer R

PREDICTED: Apis mellifera ribosomal protein L32 (RpL32), transcript variant X3, mRNA

Sequence ID: XM_016914656.2 **Length:** 601 **Number of Matches:** 1

Range 1: 387 to 457 GenBank Graphics ▼ Next Match ▲ Previous Matc

Score	Expect	Identities	Gaps	Strand
109 bits(120)	7e-24	68/72(94%)	1/72(1%)	Plus/Minus

```
Query  13   ACCATGAGCGCTTGTCAGCACAAAATGTTCTGTTTTGCATCATTAAAACTTCCAGTTCCT  72
            |||||||||   || |||||||||||| |||||||||||||||||||||||||||||||||
Sbjct  457  ACCATGAGCAATT-TCAGCACAAAATTTTCTGTTTTGCATCATTAAAACTTCCAGTTCCT  399

Query  73   TGACATTATGTA  84
            ||||||||||||
Sbjct  398  TGACATTATGTA  387
```

TpnCIIb - Primer F POS

PREDICTED: Apis mellifera troponin C type IIb (TpnCIIb), transcript variant X2, mRNA

Sequence ID: XM_026444560.1 **Length:** 745 **Number of Matches:** 1

Range 1: 256 to 333 GenBank Graphics ▼ Next Match ▲ Previ

Score	Expect	Identities	Gaps	Strand
121 bits(133)	4e-27	74/78(95%)	2/78(2%)	Plus/Plus

```
Query  17   CACTTCCAAGAGGAGGACGACGAGGCTCTGCAGAAGGAGTTGAAAGAGG--TACAGGCTT  74
            |||||||||||||||||||||||||| |||||||||||||||||||||||  | |||||||
Sbjct  256  CACTTCCAAGAGGAGGACGACGAGGCCCTGCAGAAGGAGTTGAAAGAGGCTTTCAGGCTT  315

Query  75   TACGACAAGGAGGGGAAC  92
            ||||||||||||||||||
Sbjct  316  TACGACAAGGAGGGGAAC  333
```

TpnCIIb - Primer F PRS

PREDICTED: Apis mellifera troponin C type IIb (TpnCIIb), transcript variant X2, mRNA

Sequence ID: XM_026444560.1 **Length:** 745 **Number of Matches:** 1

Range 1: 230 to 333 GenBank Graphics ▼ Next Match ▲ Previou

Score	Expect	Identities	Gaps	Strand
159 bits(176)	6e-39	100/104(96%)	3/104(2%)	Plus/Plus

```
Query  9    TCGACTCGTTCTATCGGGTTGCCTGC-ACTTCCAAGAGGAGG-CGACGAGTCCCTG-AGA  65
            |||||||||||||||||||||||||| ||||||||||||||| ||||||| ||||| |||
Sbjct  230  TCGACTCGTTCTATCGGGTTGCCTGCCACTTCCAAGAGGAGGACGACGAGGCCCTGCAGA  289

Query  66   AGGAGTTGAAAGAGGCTTTCAGGCTTTACGACAAGGAGGGGAAC  109
            |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  290  AGGAGTTGAAAGAGGCTTTCAGGCTTTACGACAAGGAGGGGAAC  333
```

TpnCIIb - Primer R

PREDICTED: Apis mellifera troponin C type IIb (TpnCIIb), transcript variant X2, mRNA

Sequence ID: XM_026444560.1 **Length:** 745 **Number of Matches:** 1

Range 1: 205 to 291 GenBank Graphics ▼ Next Match ▲ Previou

Score	Expect	Identities	Gaps	Strand
153 bits(169)	8e-37	86/87(99%)	0/87(0%)	Plus/Minus

```
Query  11   CTTCTGCAGAGCCTCGTCGTCCTCCTCTTGGAAGTGGCAGGCAACCCGATAGAACGAGTC  70
            ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  291  CTTCTGCAGGGCCTCGTCGTCCTCCTCTTGGAAGTGGCAGGCAACCCGATAGAACGAGTC  232

Query  71   GAAATTTAATTTGCCGCTTCCTTCCTC  97
            |||||||||||||||||||||||||||
Sbjct  231  GAAATTTAATTTGCCGCTTCCTTCCTC  205
```

Takeout - Primer F POS

Apis mellifera take-out-like carrier protein (JHBP-1), mRNA

Sequence ID: NM_001011640.1 **Length:** 991 **Number of Matches:** 1

See 1 more title(s)

Range 1: 391 to 481 GenBank Graphics ▼ Next Match ▲ P

Score	Expect	Identities	Gaps	Strand
152 bits(168)	7e-37	89/91(98%)	1/91(1%)	Plus/Plus

```
Query  3    CTGTCATCGGA-TCGTACAATCCTCAAGTAGATTTTGTCGCCGATTACAAAATCGAAGGA  61
            ||||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  391  CTGTCATCGGAATCATACAATCCTCAAGTAGATTTTGTCGCCGATTACAAAATCGAAGGA  450

Query  62   AAAGTGTTGCTTTTACCTGTCAGAGGAGCGG  92
            ||||||||||||||||||||||||||||||||
Sbjct  451  AAAGTGTTGCTTTTACCTGTCAGAGGAGCGG  481
```

Takeout - Primer F PRS

Apis mellifera take-out-like carrier protein (JHBP-1), mRNA

Sequence ID: NM_001011640.1 **Length:** 991 **Number of Matches:** 1

See 1 more title(s)

Range 1: 383 to 483 GenBank Graphics ▼ Next Match ▲ P

Score	Expect	Identities	Gaps	Strand
132 bits(145)	4e-30	91/102(89%)	1/102(0%)	Plus/Plus

```
Query  27   AATCTATACTGTCATCGTGCATCGTACAATCCTCATGTAGATTTTGTCGCCGATGACGAA  86
            ||| |||||||||||||| | ||| ||||||||||| |||||||||||||||||| || ||
Sbjct  383  AATGTATACTGTCATCG-GAATCATACAATCCTCAAGTAGATTTTGTCGCCGATTACAAA  441

Query  87   AACGAAGAAAAAGTGTTGCTTTTGCCTGTAAGAGGAGCGGGA  128
            | ||||| |||||||||||||||| ||||| ||||||||||||
Sbjct  442  ATCGAAGGAAAAGTGTTGCTTTTACCTGTCAGAGGAGCGGGA  483
```

Takeout - Primer R

Apis mellifera take-out-like carrier protein (JHBP-1), mRNA

Sequence ID: NM_001011640.1 **Length:** 991 **Number of Matches:** 1

See 1 more title(s)

Range 1: 368 to 441 GenBank Graphics ▼ Next Match ▲ P

Score	Expect	Identities	Gaps	Strand
113 bits(125)	6e-25	71/74(96%)	2/74(2%)	Plus/Minus

```
Query  8    TTTGTA-TCGGCGACAAAATCTACTTGAGG-TTGTACGATTCCGATGACAGTATACATTT  65
            |||||| |||||||||||||||||||||||| ||||| |||||||||||||||||||||||
Sbjct  441  TTTGTAATCGGCGACAAAATCTACTTGAGGATTGTATGATTCCGATGACAGTATACATTT  382

Query  66   ATCCCAGTCTATGT  79
            ||||||||||||||
Sbjct  381  ATCCCAGTCTATGT  368
```

Smaug - Primer F POS

PREDICTED: Apis mellifera protein Smaug homolog 1 (LOC409472), mRNA

Sequence ID: XM_006562886.3 **Length:** 7433 **Number of Matches:** 1

Range 1: 1202 to 1303 GenBank Graphics ▼ Next Match ▲ P

Score	Expect	Identities	Gaps	Strand
132 bits(145)	3e-30	96/105(91%)	5/105(4%)	Plus/Plus

```
Query  4     TTGCG-CTTCATCAG-ATAGTTATTTATTTGTCACTCTTAGTTATGAAGAAATGCTGCAA  61
             ||||| |||||| |  |||||||||||||||||||||||||||||||||||||||||||
Sbjct  1202  TTGCGTCTTCATAAATATAGTTATTTATTTGTCACTCTTAGTTATGAAGAAATGCTGCAA  1261

Query  62    TTAACTCGAAGAACAATTGCGCCGAACAGGGTGTCGACAAAAGGG  106
             |||||| |||||||||||  || |||||||||| |||||||||
Sbjct  1262  TTAACT-GAAGAACAATT-AGCAGAACAGGGTGT-GACAAAAGGG  1303
```

Smaug - Primer F PRS

PREDICTED: Apis mellifera protein Smaug homolog 1 (LOC409472), mRNA

Sequence ID: XM_006562886.3 **Length:** 7433 **Number of Matches:** 1

Range 1: 1204 to 1284 GenBank Graphics ▼ Next Match ▲ P

Score	Expect	Identities	Gaps	Strand
88.7 bits(97)	4e-17	71/82(87%)	3/82(3%)	Plus/Plus

```
Query  38    GCATCTTCATAAACTAT-GTTATTTATTTGTCACTCTTAGTTATGAAGAAAGGCCGCCTT  96
             || |||||||||| ||| |||||||||||||||||||||||||||||||| || ||   |
Sbjct  1204  GCGTCTTCATAAA-TATAGTTATTTATTTGTCACTCTTAGTTATGAAGAAATGCTGCAAT  1262

Query  97    TAACAGAAGAGC-ATAAGCAGA  117
             |||| ||||| | || ||||||
Sbjct  1263  TAACTGAAGAACAATTAGCAGA  1284
```

Smaug - Primer R

PREDICTED: Apis mellifera protein Smaug homolog 1 (LOC409472), mRNA

Sequence ID: XM_006562886.3 **Length:** 7433 **Number of Matches:** 1

Range 1: 1168 to 1262 GenBank Graphics ▼ Next Match ▲ P

Score	Expect	Identities	Gaps	Strand
172 bits(190)	9e-43	95/95(100%)	0/95(0%)	Plus/Minus

```
Query  9     ATTGCAGCATTTCTTCATAACTAAGAGTGACAAATAAATAACTATATTTATGAAGACGCA  68
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1262  ATTGCAGCATTTCTTCATAACTAAGAGTGACAAATAAATAACTATATTTATGAAGACGCA  1203

Query  69    AAGTTTTAAGCCACGATGGAACGTCTCTCATACCA  103
             |||||||||||||||||||||||||||||||||||
Sbjct  1202  AAGTTTTAAGCCACGATGGAACGTCTCTCATACCA  1168
```

Anexo A: Alinhamento das sequências obtidas pelo sequenciamento dos produtos de PCR dos genes candidatos, com sequências gênicas depositadas no banco de dados do NCBI. As sequências do fragmento obtido por sequenciamento dos genes *RpL32, TpnCIIb, To* e *smg* foram alinhadas com as sequências dos respectivos genes depositadas no banco de dados do NCBI.