UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO
DEPARTAMENTO DE GENÉTICA

**Análise do processo de ativação dos ovários de *Apis mellifera*, aspectos
morfológicos e expressão gênica**

Liliane Maria Fróes de Macedo

Ribeirão Preto

2013

LILIANE MARIA FRÓES DE MACEDO

Análise do processo de ativação dos ovários de *Apis mellifera*, aspectos morfológicos e expressão gênica

Tese apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para obtenção do título de Doutor em Ciências.

Área de concentração:
GENÉTICA

Orientadora: Prof^a Dr^a Zilá Luz Paulino Simões

Ribeirão Preto

2013

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.

Macedo, Liliane Maria Fróes de Macedo

Análise do processo de ativação dos ovários de *Apis mellifera*, aspectos morfológicos e expressão gênica.

Ribeirão Preto, 2013.

181 p. : il. ; 30 cm

Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP. Área de concentração: Genética.

Orientadora: Simões, Zilá Luz Paulino.

1. *Apis mellifera*. 2. Reprodução. 3. Patência. 4. RNA-Seq. 5. microRNAs.

Apoio financeiro:

Este trabalho foi realizado com o apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Proex e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Processo 2010/06336-2 e Projeto Temático 2011/03171-5.

Folha de Aprovação

Nome: Liliane Maria Fróes de Macedo

Título da Tese: Análise do processo de ativação dos ovários de *Apis mellifera*, aspectos morfológicos e expressão gênica.

Tese apresentada ao Programa de Pós-Graduação em Genética da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, como parte dos requisitos para obtenção do título de Doutora em Ciências.

Area de concentração: Genética.

DATA DA DEFESA: ____/____/____

Banca examinadora:

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr. _____ Instituição: _____

Julgamento: _____ Assinatura: _____

*“Por vezes sentimos que aquilo que fazemos não é
senão uma gota de água no mar. Mas o mar seria
menor se lhe faltasse uma gota”.*

Madre Teresa de Calcutá

Dedico este trabalho

Com muito carinho, aos meus pais, Elen e Leda e
às minhas avós: Nena e Edyr

AGRADECIMENTOS

Á Deus por sempre me conceder bênção que me fizeram enfrentar os obstáculos e chegar até aqui;

À minha orientadora Profa. Dra. Zilá Luz Paulino Simões, a quem muito admiro pela competência profissional e genialidade, obrigada pela orientação e pela oportunidade de crescimento pessoal e profissional.

Ao Dr. Daniel Guariz Pinheiro; à Dra. Juliana Ramos Martins pela participação no desenvolvimento deste trabalho, agradeço também pela companhia e amizade durante esses anos.

As Prof. Dr. Klaus Hartfelder por todo auxílio, além de permitir o uso de seu laboratório para execução de experimentos.

À Dra. Márcia M. Gentile Bitondi, por todas as formas de auxílio, pelas sugestões e incentivo durante a realização deste trabalho.

À Dra. Livia Maria Rosatto Moda e ao Dr. Rodrigo Pires Dallacqua; por todas as ajudas, conselhos e participações nos experimentos de microscopia.

À Dra. Érica Donato Tanaka e à Dra. Karina Rosa Guidugli Lazzarini e Dra. Tathyana Rachel Palo Mello, por todos os ensinamentos no laboratório e pelos conselhos e companheirismo.

Ao Prof. Dr. Ademilson Espencer E. Soares, pelos estímulos recebidos, compreensão, flexibilidade e amizade.

À Marcela A. F. Bezerra Laure, pela ajuda essencial nas dissecções e amizade.

À Vera L. C. Figueiredo, pelas conversas alegres, conselhos e ajuda em tudo que tivesse ao seu alcance.

Aos técnicos do apiário experimental do Departamento de Genética, Luiz Aguiar e Rogério Aparecido Pereira pela grande contribuição neste trabalho e pelos ensinamentos da biologia das abelhas.

À todos os colegas e amigos do Laboratório de Biologia do Desenvolvimento de Abelhas, Apilab e Bloco A do Departamento de Genética da FMRP, pelas discussões frutíferas, amizade e conselhos.

À equipe do Laboratório de Microscopia Eletrônica do Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos da Faculdade de Medicina de Ribeirão Preto, pelo tratamento carinhoso e profissional das técnicas Maria Dolores Seabra Ferreira (Tuca), Maria Tereza Piccinoto Maglia, ao técnico José Augusto Maulin

e aos responsáveis pelolaboratório, Prof. Dr. Roy Edward Larson e Prof. Dr. Gutemberg M. Rocha.

Ao Dr. Peter Dearden e Dra. Elizabeth Duncan, por terem me recebido em seu laboratório na Universidade de Otago, Nova Zelândia.

Aos meus familiares pelo carinho e apoio em especial aos meus tios Elizabeth M. de Macedo Viegas, Eliana G. de Macedo Lemos e Manuel Victor Lemos pelo grande estímulo, entusiasmo e paixão contagiante pela Ciência.

Ao João Bortolin, meu grande amor, que em nosso relacionamento sempre prezou pelo respeito, verdade e cumplicidade. Agradeço por acreditar que nosso amor pode ultrapassar qualquer barreira de tempo e distância.

Ao meu irmão Marcelo pelo carinho, amizade, respeito e por não hesitar em me apoiar.

Aos meus pais Elen e Leda que sempre me inspiram, me apoiam, me incentivam e que me deram todo o suporte para chegar até aqui.

À CAPES e à FAPESP pelo suporte financeiro

Enfim, a todos vocês e tantos outros colegas, amigos e familiares, a minha gratidão e que Deus na sua infinita bondade, retribua a vocês em dobro dos gestos de carinho e amizade que dedicaram a mim.

RESUMO

Macedo, LMF. Análise do processo de ativação dos ovários de *Apis mellifera*, aspectos morfológicos e expressão gênica. 2013. 181pp. Tese (Doutorado). Faculdade de Medicina de Ribeirão Preto. Universidade de São Paulo. Brasil.

Inúmeros aspectos da reprodução em *Apis mellifera* já foram extensamente divulgados, no entanto, os mecanismos reguladores da manutenção do estado estéril das operárias, bem como aqueles que permitem a ativação de seus ovários, ainda estão para serem descobertos. Por exemplo, a organização dos folículos ovarianos em crescimento e a arquitetura e papel das células foliculares neste processo. Além disso, para compreender o processo de ativação dos ovários em um contexto mais amplo, também é necessária uma investigação da síntese e maturação de diferentes classes de RNAs as quais modelam redes de interações gênicas extremamente complexas. Portanto, neste doutorado, tivemos como objetivo realizar 1- uma análise morfológica dos ovários ativos de operárias de *A. mellifera* obtidos em condições orfandade, com ênfase nas células foliculares e 2- um estudo aprofundado da regulação da expressão gênica (genes estruturais e reguladores) que é de fundamental importância para ligar os genótipos aos fenótipos. A análise morfológica dos ovários de operárias de *A. mellifera* foi realizada em microscópio de fluorescência ou confocal (priorizou a contagem das células foliculares) e microscópio eletrônico de transmissão, que permitiu a descrição e caracterização, pela primeira vez, da patência em ovários de operárias *A. mellifera*. Paralelamente, por meio da técnica de RNAseq, foi possível analisar o transcriptoma (miRNAs e mRNAs) de amostras específicas de ovários, em diferentes estados fisiológicos, em rainhas e operárias. Os mRNAs e miRNAs que se destacaram em

nossas análises *in silico* foram validados experimentalmente por RT-PCR com alto grau de reprodutibilidade e em harmonia com o estado fisiológico dos ovários. Os transcritos altamente expressos nos ovários ativados foram: *fpps5*, *cad*, *obp7*, *yellow-g* e aqueles representados pelo GB42182 e GB44975. Acreditamos estes genes possam fazer parte da rede que regula o processo de ativação dos ovários em *A. mellifera*. Os miRNAs que se destacaram em nossas análises foram: A) miR-306 e miR-317 - altamente expressos nas amostras de ovários funcionais e B) miR-71 pelo fato de, nas análises *in silico*, ser o mais forte candidato a alvejar a *vitelogenina*, e na análise experimental, apresentarem, microRNAs e mRNAs, perfis de expressão antagônicos. A construção de bibliotecas de microRNAs e mRNAs a partir de ovários funcionais e não funcionais de abelhas operárias e rainhas, a análise de expressão, bem como a predição de uma rede de integração nos deu um retrato do sensível equilíbrio reprodutivo que mantém ambas as castas em aparente harmonia dentro da colônia aonde elas assumem, no momento certo, seus papéis nesta sofisticada sociedade empreendendo ou não a reprodução.

ABSTRACT

Macedo, LMF. Analysis of the activation process of ovaries in *Apis mellifera*, the morphological aspects and gene expression. 2013. 181pp. Ph.D .Thesis – Faculdade de Medicina de Ribeirão Preto. Universidade de São Paulo. Brazil.

Countless aspects of reproduction in *Apis mellifera* have been widely published, however, the regulatory mechanisms for the maintenance of the sterile state of workers as well as those that allow the activation of their ovaries are still to be discovered, as much as the organization of growing ovarian follicles, the architecture and the role of follicular cells during this process. Furthermore, to understand the activation process of the ovaries in a broader context, it is also necessary to investigate the synthesis and maturation of different classes of RNAs which exemplify networks of gene interactions, extremely complex. Therefore, PhD project, we aimed to approach: 1 - A morphological analysis of active ovaries of *A. mellifera* workers obtained in queenless conditions, with emphasis on the follicular cells and 2 - A detailed study of the regulation of gene expression (structural and regulatory genes) that is crucial for linking genotypes to phenotypes. Morphologic analysis of workers' ovaries of *A. mellifera* was performed under a fluorescence microscope or confocal (prioritized follicular cell count) and transmission electron microscope, which allowed, for the first time, a description and characterization of the patency of worker' ovaries in *A. mellifera*. Similarly, by RNAseq technique, it was possible to analyze the transcriptome (miRNAs and mRNAs) of specific samples of ovaries at different physiological states, in queens and workers. mRNAs and miRNAs that stood out in our *in silico* analysis were experimentally validated by RT-PCR with high reproducibility and in harmony with ovaries' physiologic state. Transcripts highly expressed in activated ovaries were *fpps5*, *cad*,

obp7, *yellow-g* and those represented by GB42182 and GB44975. We believe these genes may be part of the network that regulates ovaries' activation process in *A. mellifera*. miRNAs that stood out in our analysis were: - a) - miR-306 and miR-317 - highly expressed in samples of active ovaries and b) -miR-71 by the fact that the *in silico* analysis, was the strongest candidate to target *vitellogenin*, and in experimental analysis, presented antagonistic profile of expression when microRNAs and mRNAs were contrasted. The construction of microRNAs and mRNAs libraries from active and inactive ovaries of worker bees and queens, the analysis' expression, as well as the prediction of a integrative network has given us a portrait of the sensitive reproductive balance that keeps both castes of bees in apparent harmony within the colony, where they take each one, at the right time, their roles in this sophisticated society, undertaking or not the reproduction.

LISTA DE FIGURAS

Figura 1	21
Figura 2	23
Figura 3	25
Figura 4	54
Figura 5	56
Figura 6	58
Figura 7	60
Figura 8	62
Figura 9	64
Figura 10	65
Figura 11	70
Figura 12	71
Figura 13	74
Figura 14	75
Figura 15	77
Figura 16	77
Figura 17	79
Figura 18	80
Figura 19	90
Figura 20	90
Figura 21	94
Figura 22	95
Figura 23	118
Figura 24	119
Figura 25	124
Figura 26	127
Figura 27	128
Figura 28	129
Figura 29	130
Figura 30	136

LISTA DE TABELAS

Tabela I.....	37
Tabela II	55
Tabela III.....	68
Tabela IV.....	83
Tabela V	88
Tabela VI.....	89
Tabela VII	93
Tabela VIII.....	116
Tabela IX.....	117
Tabela X.....	121
Tabela XI.....	122
Tabela XII	159
Tabela XIII.....	160

SUMÁRIO

1.0	Introdução	19
1.1	Aspectos gerais das abelhas <i>Apis mellifera</i>	19
1.2	Aspectos gerais das ovogênese em <i>Drosophila melanogaster</i>	26
1.3	Informações básicas sobre os miRNAs	28
2.0	Objetivo geral.....	32
2.1	Objetivos específicos.....	32
3.0	Metodologia.....	35
3.1	Material Biológico	35
3.2	Análises Morfológicas.....	38
3.2.1	Citoquímica dos ovários de abelhas <i>A. mellifera</i> – marcação com DAPI.....	38
3.2.2	Contagem das células foliculares	39
3.2.3	Microscopia eletrônica de Transmissão	39
3.3	Análises de expressão gênica	40
3.3.1	Extração do RNA total	40
3.3.2	Sequenciamento e construção de bibliotecas de mRNAs e miRNAs.....	41
3.3.3	Análises do sequenciamento das bibliotecas de mRNAs	42
3.3.3.1	Eliminação de Adaptadores.....	43
3.3.3.2	Avaliação de qualidade das leituras	43
3.3.3.3	Mapeamento genômico	44
3.3.3.4	Montagem dos Transcritos e Estimativa da Abundância	44
3.3.3.5	Expressão Gênica Diferencial	46
3.3.3.6	Seleção dos Genes Diferencialmente Expressos	46
3.3.3.7	Caracterização funcional	47
3.3.4	Análises dos dados de sequenciamento das bibliotecas de pequenos RNAs	47
3.3.4.1	Análise de expressão de miRNAs	48
3.3.4.2	Determinação dos alvos dos miRNAs.....	49
3.4	Síntese de cDNA e PCR em Tempo Real para quantificação dos mRNAs e miRNAs	49
4.0	Resultados	53
4.1	Análise da morfologia dos ovários ativados de operárias	53

4.1.1 – Desenvolvimento dos ovócitos e células foliculares	53
4.1.2 – Análise da patência durante a oogênese em ovários de operárias (MET)	60
4.2 Sequenciamento do Transcriptoma de ovários de <i>Apis mellifera</i>	67
4.2.1 Análise do estado funcional <i>versus</i> estado não funcional - ovários de operárias (CRxE1; CRxE2; SRxE1 e SRxE2) e ovários de rainha (RAxRV)	69
4.2.2 Análise comparativa das bibliotecas de ovários inativos (não funcionais) de operárias <i>Apis mellifera</i> (CR, SR e FOR)	78
4.2.3 Categorização funcional dos genes diferencialmente expressos quando comparamos amostras de ovários funcionais com amostras de ovários não funcionais	82
4.2.4 Categorização funcional dos genes diferencialmente expressos em CRxSR	89
4.2.5 Validação experimental por PCR em Tempo Real (RT-PCR)	91
4.2.5.1 GB42182	96
4.2.5.2 FPPS5	97
4.2.5.3 Homeótico caudal	100
4.2.5.4 Obp7	103
4.2.5.5 Yellow-g e Yellow-x2	108
4.2.5.6 GB44975	111
4.2.5.7 GB53748	112
4.2.5.8 Vg	113
4.3 miRNAs expressos em ovários inativos e ativados de <i>Apis mellifera</i>	115
4.3.1 Análise global (RNAseq) dos RNAs curtos	115
4.3.2 - Validação experimental dos miRNAs	122
4.4 Rede integrativa – miRNA-mRNA	134
5.0 Conclusão	138
6.0 Referências Bibliográficas	141
7.0 Apêndice	159

1.0 Introdução

1.1 Aspectos gerais das abelhas *Apis mellifera*

Em abelhas *Apis mellifera*, existem dois fenótipos femininos morfológica e fisiologicamente distintos – as castas - rainhas e operárias. A partir de um ovo fecundado pode-se desenvolver tanto uma rainha como uma operária, o desenvolvimento diferencial das castas é uma resposta à alimentação recebida durante a fase larval. As larvas do sexo feminino, até o segundo estágio larval, recebem apenas geleia real como fonte de alimento (secreção glandular rica em aminoácidos). A partir do terceiro estágio larval, as larvas que são alimentadas apenas com geleia real se desenvolverão em rainhas, e as larvas cuja alimentação consiste em uma mistura de secreções glandulares suplementada com pólen e mel (Rembold, 1987) se desenvolverão em operárias. A alimentação com geleia real é responsável por uma resposta neuroendócrina que estimula o aumento da atividade dos *corpora allata* (CA) - par de glândulas localizadas no complexo retrocerebral que sintetiza o Hormônio Juvenil (HJ) (Nijhout e Wheeler, 1982; Wheeler, 1986; Winston, 1987; Rachinsky e Hartfelder, 1990; Hartfelder e Engels, 1998). Este, considerado o fator endógeno de maior importância para a indução do desenvolvimento de rainhas (Rachinsky, 1994), influencia os processos de proliferação celular e atua sobre o desenvolvimento do ovário estabilizando os elementos do citoesqueleto durante a formação dos polifusomas sinciciais e especificamente, inibe a apoptose nas células germinativas (Schmidt-Capella e Hartfelder, 1998; 2002). Assim, até o quarto estágio larval, os ovários das rainhas e operárias ainda são equivalentes, cada um contendo aproximadamente de 150-180 primórdios de ovaríolos (unidade funcional do ovário). Mas, durante as etapas seguintes

de diferenciação, os ovários das larvas de rainhas continuam crescendo e se desenvolvendo enquanto que nos ovários das larvas de operárias é estabelecido um processo de morte celular, que aparentemente atinge inicialmente as células da linhagem germinativa.

Descobertas recentes demonstraram que uma proteína presente na geleia real, chamada *royalactin*, também é um fator específico e determinante para o desenvolvimento de rainhas (Kamakura, 2011) e confirmam que os altos níveis de açúcares presentes na alimentação das rainhas seriam também fatores limitantes (Leimar *et al*, 2012). As diferenças estabelecidas durante o período larval prevalecem durante toda a vida, sendo as modificações do sistema reprodutor, uma das mais marcantes. Por exemplo, nas operárias, os ovários são menores e menos desenvolvidos que os das rainhas e as modificações morfológicas relativas ao aparelho reprodutor não permitem que elas acasalem (Cruz-Landim, 2009).

Tradicionalmente, em insetos, distinguem-se dois tipos básicos de ovários (em função da organização dos trofócitos ou células nutridoras que os acompanham) os panoísticos e os meroísticos (Figura 1). Os ovários panoísticos são mais comuns em insetos basais, enquanto que os meroísticos são predominantes nos grupos mais derivados. Os ovários das abelhas são do tipo meroístico politrófico, no qual cada folículo no vitelário dos ovaríolos contém um ovócito e várias células nutridoras associadas (Cruz-Landim, 2009).

Figura 1: Reconstrução esquemática de diferentes tipos de ovários dos insetos. (A) ovário de ovário panoístico. B e C – ovário de ovário meroístico; telotrófico ou acrotrofico (B) e politrófico (C). (Modificado de: www.uam.es/departamentos/ciencias/biologia/citologia/Descargas/guion_practica3.pdf).

As rainhas adultas de *A. mellifera* possuem de 180 a 200 ovários em cada um de seus ovários, enquanto que as operárias possuem somente entre 2 e 12 (Oertel, 1930; Snodgrass, 1956; Chaud-Neto e Bueno, 1979; Morini e Bueno, 1993). O número de ovários presente nos ovários das operárias varia entre as linhagens de *A. mellifera* (Kapil 1962; Michener e Brothers, 1974). Há uma indicação de que o número de ovários seja uma propriedade do genótipo bem como da composição genotípica da colônia (Beekman *et al.*, 2000; Allsopp *et al.*, 2003; Linksvayer *et al.*, 2009), esta indicação reforça a importância de efeitos genéticos indiretos. Acredita-se que haja uma arquitetura genética complexa para a determinação do número de ovários nas operárias, esta hipótese tem suporte nas análises de QTL - “*quantitative trait loci*”, realizadas por Linksvayer *et al.*(2009).

Os ovários são estruturas alongadas e afiladas nas extremidades, contêm no seu interior as células germinativas em diferentes fases da ovogênese, além de células somáticas de sustentação. No ovário maduro, cada ovário está longitudinalmente

diferenciado, do ápice para a base, nas seguintes porções: filamento terminal, germário, vitelário e pedicelo, pelo qual se liga ao oviduto lateral. O filamento terminal é formado pelo prolongamento da bainha peritoneal que envolve os ovaríolos contendo células indiferenciadas no interior cuja natureza não está completamente esclarecida. Foi descrito que o filamento terminal está associado à sustentação do ovário dentro da cavidade abdominal além de atuar como um centro de sinalização somática para as células-tronco germinativas (Lin, 1997). Acredita-se também que sua função possa ir além de um simples suspensor ovariano; e é possível que nesta porção do ovaríolo estejam presentes as células-tronco das células pré-foliculares, ou até células germinativas (Cruz-Landim, 2009). No germário encontram-se as células-tronco germinativas e somáticas e é a região onde ocorrem divisões e diferenciações da linhagem germinativa em células nutridoras e ovócito e das células somáticas ou pré-foliculares (King *et al.*, 1982; Margolis e Spradling, 1995; Lin, 1997; Lin e Spradling, 1997). O vitelário, porção do ovário relativamente extensa nas rainhas em postura (ovopositoras), cujo tamanho depende da atividade reprodutiva, caracteriza-se pela presença de folículos contendo ovócitos em processo de maturação (Büning, 1994). Nesta região ocorre a deposição do vitelo no ovócito (Snodgrass, 1956; Louveaux 1976; Chapman, 1998) e, portanto, o seu amadurecimento. Em rainhas recém-emergidas e operárias, todo o ovaríolo apresenta-se fino e com o mesmo diâmetro em toda a sua extensão, com exceção do filamento terminal, e está ocupado pelo germário (Figura 2). Quando o ovário é ativado, ovócitos diferenciam-se na base do ovaríolo e iniciam a vitelogênese, originando um vitelário (Cruz-Landim, 2009).

Figura 2: Ovários não ativados de operárias de *A. mellifera*, neste caso são formados por 4 ovaríolos (setas). Região mais fina e distal caracteriza o filamento terminal as demais porções são caracterizadas pelo germário. Imagem obtida em microscópio confocal (Leica TCS-SP5 ou TCS-SP2 no Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos da FMRP-USP), coloração por DAPI. Aumento, no canto inferior direito das figuras. (Barra =1 µm).

Nas fêmeas das espécies de abelhas solitárias ou com organização colonial sem castas e nas rainhas das espécies eussociais, o acasalamento parece ser crucial para iniciar a ativação dos ovários e consequente ovogênese. Já para as operárias eussociais isto não é verdadeiro, em função dos processos relacionados com a diferenciação das castas nas fases imaturas do desenvolvimento, as operárias possuem um aparelho reprodutor menor, principalmente no que se refere aos dutos genitais, razão pela qual as operárias não se acasalam (Cruz-Landim, 2009). Neste caso, condições fisiológicas individuais associadas a condições ambientais desencadeiam a vitelogênese (Hoover *et al.*, 2006). Em *A. mellifera*, um destes fatores pode ser a orfandade da colônia no qual as abelhas ficam livres dos efeitos dos feromônios produzidos pela rainha (Ruttner e Hesse, 1981).

Um ovário ativado de abelhas *A. mellifera* é caracterizado pela presença dos folículos ovarianos: 1) ovócito, rodeado pelas células foliculares e 2) células nutridorais, rodeadas por células foliculares achatadas. Existe um estreitamento entre o conjunto das

células nutridoras e o ovócito, portanto, o folículo fica dividido em duas câmaras: a câmara ovocítica e a nutridora (Figura 3). Em um primeiro momento, o crescimento do ovócito deve-se exclusivamente ao material recebido pelas células nutridoras. A passagem do conteúdo das células nutridoras para o ovócito faz com que a câmara nutridora comece a decrescer, até se esvair sobrando apenas seus núcleos desintegrados. Ainda não está completamente esclarecido o que determina o movimento do material das células nutridoras para o interior do ovócito. Neste momento os ovócitos são classificados como pré-vitelogênicos, pois a deposição de vitelo ainda não se iniciou. Os ovócitos são classificados como vitelogênicos quando já se consegue verificar, na periferia do ovócito, a deposição do vitelo sob a forma de grânulos acidófilos, a princípio pequenos, de maneira que a parte central do citoplasma continua basófila. O material do espaço peri-ovocítico é absorvido pelo ovócito por pinocitose e é a partir da fusão das vesículas pinocíticas que se formam os grãos de vitelo. A deposição do vitelo encontra-se avançada quando nota-se a presença de grânulos de vários tamanhos ocupando praticamente todo o citoplasma do ovócito. O vitelo das abelhas é predominantemente proteico e basicamente constituído de vitelina que é uma glicofosfolipoproteína.

No período vitelogênico, o crescimento do ovócito se dá, principalmente, por endocitose das proteínas da hemolinfa, estas derivadas, em sua maioria, do corpo gorduroso (vitelogenina - Vg) (para revisão ver Raikhel e Dhadialla, 1992; Izumi *et al.*, 1994; Melo *et al.*, 2000). Porém as células foliculares formam uma camada única ao redor dos folículos, constituindo uma barreira física entre a hemolinfa e o ovócito. A função mais conhecida para as células foliculares, entretanto, é a síntese da membrana vitelina e o cório (envoltórios do ovo) (Büning, 1994). Porém, durante a ovogênese, a função das células foliculares é constituir uma barreira semipermeável entre o ovócito e

a hemolinfa, regulando o fluxo da vitelogenina e outras substâncias desta para o ovócito (King *et al.*, 1982). Já foi verificado em alguns insetos, com exceção da abelha *A. mellifera*, a presença de espaços entre as células foliculares. E foi constatado que estes espaços permitiam a passagem das proteínas derivadas da hemolinfa através do epitélio folicular em direção à superfície e para dentro do ovócito, este fenômeno foi denominado patência (Telfer *et al.*, 1982; Raikhel e Lea, 1991; Davey, 1993; Zimowska *et al.*, 1994). Assim, no presente estudo nos propusemos a investigar este fenômeno com o intuito de verificar se nos ovários ativados de operárias podemos reconhecer os espaços que caracterizam a patência de maneira semelhante ao que já foi encontrado para os demais insetos.

Figura 3: Ovários ativados de operárias de *A. mellifera*, em cada um dos ovaríolos podemos observar os folículos ovarianos (câmara ovocítica – seta amarela e câmara nutridora – seta verde) e células foliculares – seta vermelha. Imagem obtida em microscópio confocal (Leica TCS-SP5 ou TCS-SP2 no Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos da FMRP-USP), coloração por DAPI. Aumento, no canto inferior direito das figuras. (Barra =1 μ m).

Operárias e rainhas sintetizam a Vg, no entanto, as operárias produzem Vg em quantidades bem menores. A Vg acumulada na hemolinfa das operárias possui função de transporte de lipídios (Trenczek *et al.*, 1989; Engels *et al.*, 1990) e não é incorporada nos ovócitos (Bitondi e Simões, 1996), assim, os folículos ovarianos não entram em

vitelogênese. Em condições especiais, como por exemplo: na ausência da rainha ou em operárias com comportamento anarquista (Oldroyd *et al.* 1994) pode ocorrer o aumento da síntese de Vg pelo corpo gorduroso das operárias com subsequente incorporação pelos ovócitos – assim as operárias começam a produzir ovos haplóides. O processo de endocitose de Vg pelo ovócito é mediado por um receptor específico (Amvgr). Em operárias com ovários ativados e rainhas ovopositoras já foi verificado que os níveis de transcritos de *Amvgr* nos ovários são significativamente maiores do que em operárias com ovários inativos e rainhas virgens (Guidugli-Lazzarini *et al.* 2008).

Alguns autores já vêm investindo esforços nos estudos relacionados às análises de expressão gênica em abelhas *A. mellifera* quando estas ativam seus ovários (Grozinger *et al.*, 2007 e Thompson *et al.*, 2008, Cardoen *et al.*, 2011). Estes autores utilizaram da metodologia de microarranjos para identificar genes diferencialmente expressos em operárias com ovários ativados e inativos, para tanto utilizaram como material biológico cérebro, abdômen e corpo inteiro de abelhas. Porém, até o momento, nenhum trabalho realizou uma análise específica do transcriptoma dos ovários. Dentro deste cenário, nos propusemos a realizar uma análise global dos transcritos em diferentes amostras de ovários de ambas as castas, por meio de “sequenciamento de nova geração” (RNAseq), com o intuito de verificar se há um perfil de expressão hierárquica que identifique e/ou defina o estado fisiológico funcional e não funcional dos ovários nas abelhas *A. mellifera*.

1.2 Aspectos gerais das ovogênese em *Drosophila melanogaster*

Entre insetos, os mecanismos moleculares envolvidos na ovogênese que estão mais bem caracterizados são em *Drosophila melanogaster* (para revisão ver King, 1970;

de Cuevas *et al.*, 1997; Gilboa e Lehmann, 2004). As moscas possuem ovários do mesmo tipo que *A. mellifera*, e o desenvolvimento do ovócito depende de uma série de interações entre as células da linhagem germinativa e células foliculares (derivadas de células somáticas). Nas células da linhagem germinativa, os genes *gurken* e *brainiac* são requeridos para individualizar as câmaras ovocíticas, enquanto que nas células somáticas são necessários que os genes *Egfr*, *daughterless*, *Notch*, e *Delta* atuem (Goode *et al.*, 1992, 1996). Inicialmente, o mRNA de *gurken* está localizado na borda posterior do ovócito e sinaliza para as células foliculares de mesma polaridade para se diferenciarem como tal. A assimetria antero-posterior do ovócito é então estabelecida por reorganização de uma rede de microtúbulos que requer um sinal *Notch-Delta*, enviado das células foliculares com polaridade posterior para o ovócito, e também por atividade da proteína quinase A na linhagem germinativa (Ruohola *et al.*, 1991, Lane e Kalderon, 1994). Esta reorganização leva a localização de determinantes maternos antero-posteriores (genes de efeito materno), *bicoid* e *oskar*, para pólos opostos do ovócito e o reposicionamento do núcleo (para revisão ver St Johnston e Nüsslein-Volhard, 1992). No polo anterior do ovócito localiza-se *bicoid*, que será necessário para determinação de estruturas anteriores do embrião, enquanto que no polo posterior estão localizados *oskar* e *nanos* que são necessários para a formação do abdômen (St Johnston, 1993). Para promover a tradução localizada do mRNA *nanos* é necessária a proteína Vasa (Gavis *et al.*, 1996) que também atua na localização de componentes do plasma germinativo, como por exemplo o mRNA *gurken* (Tinker *et al.*, 1998; Ghabrial e Schupbach, 1999). A transcrição de vasa depende diretamente do gene *oskar*, que atua na formação do plasma polar. O mRNA de *gurken*, em combinação com Rhomboid, ativa localmente o receptor EGF e sua cascata *downstream* para direcionar as células adjacentes a se tornarem dorsais (Ruohola-Baker *et al.*, 1993). A polaridade dorso-

ventral do embrião é restrita à ação de três de genes foliculares, *windbeutel*, *nudal* e *pipe* (Hong e Hashimoto, 1995), nas células foliculares ventrais onde levam à ativação localizada de uma cascata serina-protease necessária para produzir o ligante Spätzle para ativar o receptor Toll (Chasan e Anderson, 1993). Posteriormente, a ativação do receptor Torso tirosina quinase é necessária para a padronização final do embrião e isso requer a expressão de *torso-like*. Este último é expresso por um conjunto de células foliculares durante a ovogênese e parece ser o responsável pela expressão de Trunk (Casanova *et al.*, 1995), que é o provável ligante de Torso após a fertilização (Sprenger e Nüsslein-Volhard, 1992). Após esta ligação ocorre a ativação de genes responsáveis pelo término do embrião como *hunchbein* e *tailless* nos pólos do embrião (para revisão ver St Johnston e Nüsslein-Volhard, 1992; Chen *et al.*, 2009).

Na anotação do genoma de *A. mellifera* (The Honeybee Genome Sequencing Consortium, 2006) não foram identificados ortólogos para muitos dos genes citados acima, incluindo *gurken*, *oskar*, *bicoid*, *torso* e *trunk*. Estas ausências mostram que diferentes mecanismos estão ocorrendo na ovogênese de espécies distintas, pois, assim como acontece em *A. mellifera*, os genes *gurken*, *oskar* e *bicoid* também estão ausentes nos genomas de *Bombyx mori* e *Tribolium castaneum* (Schröder, 2003; Dearden *et al.*, 2006). Portanto, está dentro de nossas expectativas identificar mecanismos próprios da abelha *A. mellifera*.

1.3 Informações básicas sobre os miRNAs

Os microRNAs (miRNAs) são pequenos RNAs não codificadores, com tamanho entre 19-24 nucleotídeos. Eles podem se ligar à região 3' não traduzida (3'UTR), sequências codificadoras ou 5'UTR de mRNAs alvos, e desta maneira, regulam a

expressão gênica em nível pós-transcricional (Fernandez-Valverde *et al.*, 2010). Essa ligação, ou pareamento RNA-RNA, leva à inibição da tradução ou à degradação do mRNA (revisado por Almeida *et al.*, 2011 e por Pasquinelli, 2012). Em animais, por não ocorrer um pareamento perfeito dos miRNAs com o alvo, são formados *duplexes* parciais, e geralmente a ligação com o alvo ocorre na região 3'UTR do mRNA alvo (Bartel *et al.*, 2009; Rigoutsos, 2009). Em plantas normalmente ocorre o pareamento perfeito do miRNA com seu alvo, e por certo tempo acreditou-se que somente neste caso o miRNA resultaria na degradação do mRNA, porém, Djuranovic *et al.* (2012) observaram que em células S2 de *D. melanogaster*, os miRNAs primeiramente impedem a tradução do mRNA alvo, e em um segundo momento ocorre a deadenilação e consequente degradação deste mRNA.

A *seed* é região do miRNA (2º ao 8º nucleotídeo na extremidade 5') que comumente forma um pareamento perfeito com o alvo. Quando ocorre do pareamento ser imperfeito na porção da *seed*, em alguns casos, pode haver uma compensação pelo pareamento substancial na região 3' do miRNA (Reinhart *et al.*, 2000, Yekta *et al.*, 2004). Além disso, já foram descritos pareamentos na porção central do miRNA (de 11 a 12 bases contínuas) sem pareamentos substantivos nas extremidades (Shin *et al.*, 2010), e também casos que não seguem nenhum padrão de pareamento específico como os exemplos acima citados (Tay *et al.*, 2008).

Os miRNAs são codificados como transcritos primários longos (pri-miRNAs), com uma estrutura *cap* 7-metil guanosina na extremidade 5' e na extremidade 3' são poliadenilados. No geral, os pri-miRNAs apresentam uma estrutura secundária em forma de grampo (revisado por Bartel, 2004). No núcleo da célula, os pri-miRNAs são processados por um complexo que inclui a RNaseIII Droscha e a proteína Pasha (em insetos e nematóides), resultando no miRNA precursor (pre-miRNA) que possui de 60-

70 nt (Lee *et al.*, 2003). Este é exportado para o citoplasma e, por meio de um complexo que inclui a RNaseIII Dicer, é então clivado produzindo um RNA fita dupla (Förstemann *et al.*, 2005). Uma enzima helicase é responsável por separar as fitas e liberar o miRNA maduro (aproximadamente 20 nt), este é incorporado a um complexo proteico, conhecido como RISC, constituído por membros da família Argonauta. A partir desse momento, o miRNA está funcionalmente ativo para atuar em seu mRNA alvo (Siomi *et al.*, 2010).

Em *D. melanogaster* os miRNAs têm sido amplamente descritos em vários estágios do desenvolvimento e ovogênese, este dado sugere que vários genes relacionados a estes processos devem estar sendo regulados por miRNAs (Aravin *et al.*, 2003; Behura, 2007) inclusive em outras classes de insetos. Em abelhas *A. mellifera*, foram identificados por métodos computacionais 64 sequencias de miRNAs (Weaver *et al.*, 2007) e 267 por meio de “sequenciamento de nova geração” – SOLID system (Chen *et al.*, 2010). No miRBase (mirbase.org, Kozomara *et al.*, 2011), estão disponíveis 221 sequencias de microRNAs (versão 20). E apesar do aumento exponencial no depósito de sequencias no miRBase, ainda não estão esclarecidas as funções de vários miRNA no organismo das abelhas. Dentro deste cenário, pretendemos determinar padrões de expressão do transcriptoma (mRNA e miRNA) tecido específico, sugerir suas prováveis funções na regulação do processo de ativação dos ovários de operárias *A. mellifera* e verificar se há uma alternância na expressão dos miRNAs e mRNAs.

2.0 Objetivo geral

Estudar o processo de ovogênese, com ênfase na vitelogênese, nas operárias e rainhas de abelhas *A. mellifera*, por meio de uma análise global comparativa da expressão de genes (estruturais e reguladores) ligados ao processo de ativação e manutenção do estado infértil. Como suporte, realizar estudos morfológicos evidenciados por marcadores específicos e análises em microscópio eletrônico de transmissão e confocal.

2.1 Objetivos específicos

- Analisar comparativamente a curva de crescimento das câmaras ovocíticas e a proliferação do epitélio folicular correspondente, ao longo de toda a extensão do ovário, em ovários ativados de operárias de *A. mellifera*.
- Investigar o fenômeno da patência em ovários ativados de operárias de *A. mellifera*.
- Realizar, por meio de diferentes amostras de ovários, uma análise global da expressão gênica de abelhas rainhas e operárias de *A. mellifera*, para comparação do estado fisiológico funcional e não funcional.
- Validar experimentalmente os genes diferencialmente expressos, obtidos a partir dos resultados de RNAseq, e com provável participação no processo de ativação dos ovários.
- Realizar uma análise global dos reguladores da expressão gênica em nível pós-transcricional, por meio do sequenciamento dos RNAs curtos não codificadores, em ovários inativos e ativados de operárias *A. mellifera*.

- Validar experimentalmente os miRNAs encontrados como diferencialmente expressos a partir da análise dos dados do RNAseq.
- Construir uma rede regulatória de interação entre mRNAs e miRNAs, relacionada ao processo de ativação dos ovários, por meio de análise *in silico*.

3.0 Metodologia

3.1 Material Biológico

Foram utilizadas abelhas operárias de *A. mellifera* (africanizadas) provenientes do Apiário Experimental do Departamento de Genética da Faculdade de Medicina de Ribeirão Preto – (USP), Ribeirão Preto, São Paulo.

Inicialmente, trabalhamos com quatro colônias diferentes para realizar o estudo com as abelhas operárias. Duas colônias (A e B) forneciam as abelhas operárias adultas enquanto as outras duas (C e D) recebiam as operárias para coletas. Retiramos das colônias A e B quadros de cria contendo pupas prestes a emergir, e estes foram acondicionados em estufa (34°C e 80% de umidade relativa). As abelhas operárias que emergiram num período de 0-16 horas foram coletadas, marcadas (com um ponto de tinta no tórax) e inseridas nas colmeias C e D para coletas futuras. Vale ressaltar que as abelhas provenientes da colônia A eram inseridas na colônia C e as abelhas provenientes da colônia B eram inseridas na colônia D. Após a primeira fase de coleta (descrita abaixo), as colônias C e D foram mantidas órfãs (condição representada pela ausência da rainha). Com sete dias de orfandade, as colônias C e D receberam novamente abelhas recém-emergidas (marcadas com cores distintas no tórax) provenientes das colônias A e B que foram utilizadas para as demais coletas descritas a seguir.

Coletamos ovários de abelhas operárias, em quatro diferentes situações, para compor quatro grupos experimentais distintos. Estas amostras foram utilizadas para a extração de RNA total e para análises morfológicas. O primeiro grupo corresponde às abelhas marcadas e inseridas nas colmeias, sem qualquer outra forma de intervenção.

Estas operárias foram mantidas dentro das colônias até atingirem quatro dias de vida adulta e após este período, as abelhas foram dissecadas para coleta dos ovários. Vale ressaltar que todas as abelhas deste grupo estão inseridas dentro de um contexto normal de colônia (na presença da rainha) e, portanto, possuíam os ovários inativos (não funcionais), desta forma receberam a sigla CR em nosso estudo. Os demais grupos experimentais foram obtidos após a retirada da rainha, ou seja, em colônias órfãs. O segundo grupo experimental, difere do primeiro apenas pelo fato das abelhas estarem em condição de orfandade, isto significa dizer que, as abelhas coletadas possuíam quatro dias de vida adulta e ovários inativos, estas receberam a sigla SR. As demais coletas serviram para compor os grupos amostrais que representam um ovário funcional, ou seja, ativado. Neste caso, não controlamos a idade das abelhas, mas sim o grau de ativação de seus ovários, sendo E2 o grau máximo de ativação e E1 um grau intermediário.

Durante a dissecação, nos preocupamos em analisar as seguintes características para a coleta dos ovários: A) ovários inativos não deveriam apresentar ovariolos contendo folículos em desenvolvimento; B) Os ovários ativados quando apresentavam: 1- folículos em pré- vitelogênese ou vitelogênese receberam a sigla E1 e 2- quando possuíam ovócitos maduros receberam a sigla E2.

Os ovários obtidos a partir da colônia A foram coletados para a extração de RNA total e utilizados para o “sequenciamento de nova geração” dos mRNAs. Os ovários obtidos a partir da colônia B foram coletados para a extração de RNA total e utilizados para a validação por PCR em Tempo Real (RT-PCR) dos mRNAs e para as análises morfológicas. Nossas análises morfológicas utilizaram apenas os ovários que receberam a sigla E2.

Posteriormente, outras 3 amostras também foram coletadas para o “sequenciamento de nova geração” dos mRNAs e validações por RT-PCR que correspondem a: 1- ovários inativos de operárias forrageiras (FOR) proveniente de uma mesma colônia; 2- ovários de rainha virgem (RA - amostras individuais) e 3- ovários de rainha fecundada (RV - amostras individuais).

Para cada uma das condições citadas acima (CR, SR, E1, E2, FOR, RA e RV) uma amostra foi utilizada para o “sequenciamento de nova geração” dos mRNAs (n=1). E para a validação por RT-PCR utilizamos três amostras de cada condição acima citada (n=3). A Tabela I resume todas as características intrínsecas das amostras coletadas para as análises de expressão dos mRNAs.

Tabela I: Amostra utilizadas para os experimentos de “sequenciamento de nova geração” e validação por RT-PCR.

Amostra	Características	Pares de ovários que compõe cada uma das amostras
CR	ovários inativos de abelhas com 4 dias de vida adulta, provenientes de colônias normais	15-18 pares
SR	ovários inativos de abelhas com 4 dias de vida adulta, provenientes de colônias órfãs	15-18 pares
E1	ovários ativados sem ovócitos maduros	7 pares
E2	ovários ativados com ovócitos maduros	5 pares
FOR	ovários de abelhas forrageiras	15-18 pares
RV	ovários de rainhas virgens	1 par
RA	ovários de rainha em postura	1 par

Também utilizamos da técnica de “sequenciamento de nova geração” para obter os níveis de transcritos dos RNAs não codificadores curtos. Para este sequenciamento e posterior validação por RT-PCR foram coletadas 3 amostras que correspondiam a: 1- OA: *pool* de 7 ovários ativos em diferentes estágios de desenvolvimento e 2- NUR: *pool* de 15 ovários inativos de abelhas operárias nutridoras (4 dias de vida adulta). Para as análises de PCR em Tempo Real utilizamos 3 amostras de NUR (n=3, *pool* com 15 ovários) e 3 amostras de ovários OA (n=3, *pool* com 7 ovários). Para o “sequenciamento de nova geração” foi utilizada apenas 1 amostra de OA e uma amostra de NUR.

3.2 Análises Morfológicas

3.2.1 Citoquímica dos ovários de abelhas *A. mellifera* – marcação com DAPI

Após a fixação (24 horas) com solução heptano/formaldeído 1:1, os ovários foram lavados por quatro vezes em PBS 1X [PBS 2M, pH 7,2 (NaCl, KCl, Na₂HPO₄, KH₂HPO₄)] contendo 0,2% TritonX-100, por 15 minutos cada, sob leve agitação, para permitir a emulsificação da bicamada lipídica.

Para a visualização dos núcleos os ovários foram marcados por DAPI (4'6-diamidino-2-phenylindole, Sigma) na proporção de 1:4000 em PBS 1X, durante 4 minutos, sob leve agitação e a temperatura ambiente. Em seguida, foram feitas 3 lavagens de 10 minutos cada com PBS 1X contendo 0,2% TritonX-100 para retirar o excesso de corante do material.

Por fim, os ovários foram montados *in Toto* em glicerol a 75% e as imagens foram adquiridas no microscópio confocal LEICA TCS SP5 (do Departamento de Biologia Celular e Molecular - FMRP).

3.2.2 Contagem das células foliculares

Para estimar o número total de células foliculares dos ovócitos das abelhas de diferentes idades foi utilizada a fórmula matemática descrita por Pascual, 1995. Para cada ovócito analisado, a média do número de células foliculares foi obtida a partir da contagem das mesmas contidas em quatro retângulos delimitados. O comprimento (C) e largura (L) foram medidos para calcular a área da superfície do ovócito: $Pi \times \frac{A}{2} \times \frac{L}{2}$ e por fim foi utilizada a seguinte fórmula: $Total\ Cel.\ Fol. = \frac{2 \times N.de\ Cel. \times \text{Área da Sup.}}{A.do\ Ret.}$.

3.2.3 Microscopia eletrônica de Transmissão

Foram realizadas preparações dos ovaríolos de ovários ativados de abelhas operárias mantidas em colônias órfãs para análises em microscópio de luz e eletrônico de transmissão. Os ovários coletados foram dissecados em solução salina (0,9% de NaCl) esterilizada, individualizando os ovaríolos e retirando-se sua respectiva membrana peritoneal. As amostras foram fixadas em uma mistura de glutaraldeído 2% e paraformaldeído 2% em tampão cacodilato 0,1M, pH 7,2 por 2h a 4°C. Em seguida, foram retiradas da solução fixadora, lavadas em tampão PBS 0,1M (tampão fosfato pH 7,2 em 0,9% de NaCl) e mantidas neste tampão a uma temperatura de 4°C até a realização da etapa seguinte do preparo para a microscopia. Na etapa seguinte, os

ovariolos foram retirados da solução tampão e fixados por 30 min em tetróxido de ósmio 1% em tampão fosfato 0,1M, pH 7,2. Na sequência foram, então, desidratados em acetona e óxido de propileno, pré-embebidos em LX112 (Ladd) em óxido de propileno (1:1) e, finalmente, embebidos em resina pura de LX112 (Leica).

Cortes semifinos (0,5 μ m) foram corados com azul de toluidina e visualizados em microscópio de luz Axioskop II photomicroscope (Zeiss) para seleção dos blocos e amostras a serem utilizadas nos cortes ultrafinos. Após realizar os cortes ultrafinos eles foram corados com acetato de uranila 2% e citrato de chumbo 0,3%. A visualização destes cortes ultrafinos foi realizada em microscópio eletrônico de transmissão JEOL modelo JEM-100CX2 no Laboratório de Microscopia Eletrônica do Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos da Faculdade de Medicina de Ribeirão Preto/USP.

3.3. Análises de expressão gênica

3.3.1. Extração do RNA total

A extração de RNA foi feita utilizando reagente TRIzol® (solução comercial de fenol e isotiocianato de guanidina para isolamento de RNA total, Invitrogen). O protocolo de extração foi aquele recomendado pelo fabricante. O RNA obtido foi ressuspensionado em água tratada com 0,1% dietilpirocarbonato (DEPC). A quantificação do RNA foi realizada em espectrofotômetro (NanoDrop®ND-1000, NanoDrop Technologies) a 260 nm.

3.3.2 Sequenciamento e construção de bibliotecas de mRNAs e miRNAs

Parte do RNA total obtido foi enviado para “sequenciamento de nova geração” por meio da plataforma Illumina. Uma amostra com ovários ativos (OA) em diferentes estágios de ativação (estágio I e II) foi enviada para a *DNAvision* (www.dnavision.com) em colaboração com a Dra. M.D. Piulachs do Instituto de Biologia Evolutiva (CSIC-UPF) - Barcelona, para sequenciamento de RNAs de baixo peso molecular (microRNA, smallRNA, bem como outros da classe *non coding*), por meio da plataforma Illumina (HiSeq 2000, *Life Sciences*).

Amostras distintas compostas pelos grupos experimentais: CR, SR, E1 e E2, foram enviadas para a *High Throughput Sequencing Facility* instalada no Departamento de Genética da Faculdade de Medicina da Universidade da Carolina do Norte, Chapel Hill, Carolina do Norte, EUA para sequenciamento dos mRNAs por meio da plataforma Illumina (Genome AnalyserII, *Life Sciences*). A amostra (NUR) também foi enviada para a *High Throughput Sequencing Facility*, mas para o sequenciamento dos RNAs de baixo peso molecular (microRNA, smallRNA, bem como outros da classe *non coding*).

Amostras de RNA total de ovários de abelhas forrageiras (FOR), rainha virgem (RV) e rainha em atividade de postura (RA) foram encaminhadas à *facility* instalada no Laboratório de Bioquímica de Microrganismos e Plantas do Departamento de Tecnologia da Faculdade de Ciências Agrárias e Veterinárias, Unesp – Jaboticabal, para sequenciamento dos mRNAs por meio do equipamento HiScanTMSQ – Illumina. Todas as etapas do processamento, desde a quantificação no fluorímetro Qubit, análise da qualidade e integridade no Bioanalyzer, preparação da biblioteca, clusterização e sequenciamento puderam ser acompanhados junto à técnica responsável, para aprimoramento do conhecimento sobre esta metodologia.

No caso dos mRNAs, as amostras de RNA total foram preparadas utilizando-se o kit TruSeq RNA™ Sample Preparation (Illumina); e no caso dos RNAs curtos, o kit TruSeq™ Small RNA Sample preparation (Illumina). Os sequenciamentos foram do tipo *single-end*.

3.3.3 – Análises do sequenciamento das bibliotecas de mRNAs

Os resultados do sequenciamento dos mRNAs foram recuperados no servidor de armazenamento local de nosso laboratório sem nenhum ter sido submetido a um processamento prévio. Primeiramente foram definidos dois conjuntos de dados que foram processados de forma independente. Esses dois conjuntos de dados foram definidos conforme o tamanho das leituras nas bibliotecas, o primeiro com todas as bibliotecas (E1, E2, CR, SR, FOR, RV e RA) padronizadas até 36 bases, evitando dessa forma algum tipo de viés no processamento relacionado ao tamanho, e o segundo somente com as bibliotecas de 50 bases (RV, RA e FOR). No primeiro conjunto de dados, nas bibliotecas com leituras de 50 bases as 14 bases finais excedentes foram eliminadas. O pré-processamento dos dados inclui a eliminação de regiões nas leituras derivadas de artefatos do método de sequenciamento (moléculas adaptadoras) e a avaliação da qualidade das leituras.

Após a etapa inicial, as sequências foram submetidas a um alinhamento genômico, este permitiu que fosse realizado o agrupamento das leituras e a montagem dos transcritos gênicos putativos, que foram posteriormente comparados aos transcritos gênicos previamente mapeados e dessa forma identificados. Por fim, esses agrupamentos gênicos foram avaliados, permitindo a estimativa de abundância para cada amostra biológica. Esses valores representativos da expressão gênica em cada

amostra foram então comparados entre si para realizar uma avaliação da expressão gênica diferencial entre as condições biológicas sob investigação.

Esses resultados brutos foram filtrados utilizando critério de razão relativa de expressão procurando selecionar os genes que mais contribuem para as diferenças nos transcritomas de ovários ativos e inativos e entre as castas de rainhas e operárias. Cruzamentos específicos foram realizados a fim de identificar esses genes.

3.3.3.1 Eliminação de Adaptadores

A eliminação de resquícios de sequências das moléculas adaptadoras foram realizadas utilizando os *softwares* Scythe versão 0.981 (<https://github.com/vsbuffalo/scythe>) para a poda na região 3' e o Cutadapt versão 1.0 (Martin, 2011) para a poda na região 5'. O programa Scythe utiliza uma implementação do algoritmo *Naïve* Bayes para a classificação das regiões de alinhamento entre as sequências das leituras e dos adaptadores, considerando a informação de qualidade das bases, já o programa CutAdapt considera apenas a qualidade desse alinhamento. As leituras que não atingiram o tamanho mínimo de 15 bases foram descartadas.

3.3.3.2 Avaliação de qualidade das leituras

O programa Prinseq-lite versão 0.18.2 (Schmieder e Edwards, 2011) foi utilizado para podar as leituras removendo as regiões de baixa qualidade (média de qualidade Phred menor que 25 considerando uma janela de 3 bases, deslizando uma base a partir da região 3' da *read*), as caudas poli(A/T) (mínimo 5 bases A ou T das

extremidades). Após esse processamento as leituras menores que 15 nucleotídeos ou repletas de bases ambíguas (mais do que 80% de Ns) foram descartadas.

3.3.3.3 Mapeamento genômico

As leituras que passaram na etapa de pré-processamento, foram então, alinhadas contra o genoma de *A. mellifera* versão 4.5 utilizando-se o programa TopHat versão 2.0.7 (Kim e Saltzberg, 2011), o qual é apropriado para alinhar leituras RNA-Seq no genoma pois considera também os mapeamentos em regiões de encadeamento de exons. As coordenadas gênicas de referência, que foram utilizadas para auxiliar no processo de alinhamento, foram obtidas do Official Gene Set (OGS) versão 3.2 disponível no repositório *Hymenoptera Genome Database* (Munoz-Torres *et al.*, 2011) (<http://hymenopteragenome.org/beebase/>). Nessa etapa múltiplos alinhamentos foram descartados utilizando um limiar de exclusão (acima de 10 alinhamentos) para evitar mapeamentos ambíguos sem descartar a expressão de alguns genes, de acordo com o recomendado na literatura (Odawara *et al.*, 2011). O programa TopHat foi configurado para realizar um alinhamento priorizando sensibilidade e precisão em detrimento da velocidade. O alinhamento foi realizado por segmentos de 20 nucleotídeos permitindo no máximo 2 substituições (*mismatches*) em todo o alinhamento, com inserção ou deleção de no máximo 3 nucleotídeos.

3.3.3.4 Montagem dos Transcritos e Estimativa da Abundância

Para a montagem dos transcritos gênicos e estimação de seus valores de abundância em cada condição biológica foi utilizado o programa cufflinks do pacote

Cufflinks versão 2.1.1 (Trapnell *et al.*, 2010; Trapnell *et al.*, 2013). O Cufflinks inicialmente precisa definir quais são as estruturas gênicas que estão sendo expressas. Isso é realizado a partir dos fragmentos mapeados. Neste caso foi realizado com o auxílio das coordenadas gênicas previamente definidas (OGS versão 3.2), ou seja, alinhamentos não compatíveis estruturalmente com os modelos gênicos pré-determinados foram ignorados. O procedimento de montagem do Cufflinks é baseado na construção de um grafo de sobreposição que conecta os fragmentos mapeados no genoma que sobrepoem uma determinada região. A caminhada em cada grafo indica a estrutura gênica. A partir daí o Cufflinks estima o valor referente à abundância dos transcritos para cada gene, o FPKM (*Fragments Per Kilobase of exon per Million fragments mapped*), o qual é análogo ao RPKM (*Reads Per Kilobase of exon per Million Mapped reads*) no sequenciamento do tipo *single-end*. A intenção desta medida é obter um valor normalizado para as comparações subsequentes, minimizando os efeitos das diferenças no tamanho dos genes e na quantidade de leituras da biblioteca. Na etapa de estimação da abundância, as leituras mapeadas com TopHat foram agrupadas de acordo com o transcrito ao qual correspondiam, considerando o mapeamento nas coordenadas gênicas obtidas do OGS. O Cufflinks foi configurado para aplicar a estratégia de correção considerando o viés de fragmentos obtidos de forma não uniforme dentro dos transcritos (Roberts *et al.*, 2011). O método de normalização selecionado foi o do quartil superior (*Upper-quartile*) (Bullard *et al.*, 2010). Como foram permitidos alinhamentos múltiplos, outro método de ajuste foi habilitado no Cufflinks, o qual re estima a abundância das leituras mapeadas múltiplas vezes de forma probabilística de acordo com a estimativa inicial.

3.3.3.5 Expressão Gênica Diferencial

Para as análises comparativas de expressão gênica diferencial foi necessário montar uma referência comum com base nas montagens obtidas com Cufflinks para cada biblioteca, para isso o programa cuffmerge integra o pacote de programas Cufflinks. O cuffmerge também associa um identificador do conjunto gênico utilizado como guia a OGS versão 3.2, caso as coordenadas genômicas sejam coincidentes.

Utilizando sempre o resultado do cuffmerge como referência, foram realizadas diversas análises com o programa cuffdiff que também integra o pacote de programas Cufflinks. O cuffdiff testa o *log-fold-change* observado com a hipótese nula de que não há diferença, acessando assim a significância (teste-T) após estimar a variância dos valores de FPKM. Essa estimativa de variância, neste caso em específico em que não há réplicas biológicas, é realizada considerando as bibliotecas sob comparação para a construção de um modelo global.

3.3.3.6 Seleção dos Genes Diferencialmente Expressos

Os resultados brutos foram filtrados utilizando um critério de razão relativa de expressão ($abs\text{-}log_2\text{-}fold\text{-}change > 2$), para evitar divisão ilegal por zero, foi adicionado 1 aos valores de expressão (FPKM). O fato de não haver réplicas biológicas no desenho experimental compromete o poder estatístico do teste de significância, dessa forma nesta etapa observamos apenas a razão relativa de expressão. Para evitar selecionar genes com razão de expressão relativa alta em uma magnitude de expressão muito baixa, evitando também selecionar diferenças relacionadas à presença de artefatos, com maior efeito em baixa expressão, foi aplicado um filtro para valores baixos de FPKM

nas duas bibliotecas (FPKM < 5). Para ilustrar as diferenças e semelhanças entre genes diferencialmente expressos nas comparações em uma mesma análise, utilizamos diagramas de Venn, os quais foram gerados utilizando o pacote R VennDiagram (Chen e Boutros, 2011). Nesses diagramas de Venn foram separados os genes diferencialmente expressos regulados de forma consistente entre análises comparativas consideradas (regulados positivamente ou negativamente) e também os genes regulados de forma inconsistente em ao menos uma das análises, baseado na abordagem do VennPlex (Huan *et al.*, 2013). Diagramas de Venn de cada uma das análises comparativas foram gerados, neste caso destacando os genes regulados positivamente e negativamente com os genes não selecionados na intersecção.

3.3.3.7 Caracterização funcional

A categorização foi realizada durante o curso: “*Anotação, meta-anotação e extração amigável e panorâmica de dados de transcriptomas montados (assemblados) utilizando recursos largamente acessíveis (planilhas Excel)*”, organizado pelo Departamento de Bioquímica e Imunologia da Faculdade de Medicina de Ribeirão Preto, e ministrado pelo Professor Doutor José Marcos Ribeiro (*National Institutes of Health* - NIH). Esta categorização se baseou na metodologia descrita por Ribeiro *et al*, 2007 e utiliza como ferramenta computacional o programa RPS-BLAST.

3.3.4 Análises dos dados de sequenciamento das bibliotecas de pequenos RNAs

Analizamos os dados provenientes do sequenciamento de 2 bibliotecas de ovários de abelhas operárias, funcionais (OA) e não-funcionais (ONUR - nutriz). As

leituras sem processamento possuíam 40 nucleotídeos em OA e 36 nucleotídeos em ONUR. Para identificar as leituras de miRNAs, as sequências foram inicialmente submetidas a um processamento que precedeu o alinhamento contra as sequências dos precursores de miRNAs conhecidos. Essa etapa de processamento inicial incluiu a eliminação de regiões derivadas de artefatos do método de sequenciamento (moléculas adaptadoras) e a avaliação da qualidade e poda das leituras de baixa qualidade. As ferramentas e os parâmetros dessa etapa foram os mesmos aplicados nas bibliotecas de mRNAs e descritos anteriormente.

3.3.4.1 Análise de expressão de miRNAs

As leituras foram colapsadas e contabilizadas para evitar redundância de processamento na etapa de alinhamento. O alinhamento foi realizado com as sequências de precursores de miRNAs de *A. mellifera* obtidos do miRBase (versão 19). Essa etapa foi realizada com o programa miRDeep2. O programa miRDeep (versão 2.0.0.5) (Friedländer *et al.*, 2008) fornece também a quantificação da expressão dos miRNAs nas bibliotecas, através da contagem de leituras que alinham em coordenadas que se sobrepõem à dos miRNAs. O programa original foi alterado para contabilizar apenas o número de fragmentos que estão dentro das coordenadas de mapeamento dos miRNAs maduros independentemente do precursor, ou seja, quando há mais de um precursor com um mesmo maduro, e as leituras alinham consequentemente em todos os precursores, no final ela será contabilizada uma única vez para este miRNA. A normalização é realizada pelo miRDeep considerando as diferenças na quantidade de leituras mapeadas. As razões relativas de expressão para cada par de bibliotecas foram calculadas utilizando o *log2-fold-change*, após somar 1 a cada valor na comparação,

evitando divisão por zero. Para a avaliação da significância estatística das diferenças observadas entre as bibliotecas foi utilizado o teste proposto por Audic e Claverie (1997).

3.3.4.2 Determinação dos alvos dos miRNAs

A busca por alvos de miRNAs foi realizada por dois métodos RNAhybrid (Rehmsmeier *et al.*, 2004) e PITA (Kertesz *et al.*, 2007), considerando a união dos resultados obtidos. Nesta etapa foram utilizadas sequências de 1000 nucleotídeos extraídas após a região codificadora dos transcritos mensageiros e as sequências de miRNAs maduros obtidas do miRBase (versão 19). No caso do RNAhybrid foram considerados somente as predições com energia livre menor ou igual a -20 e p-valor menor ou igual a 0.005 e a região *seed* adotada compreendeu os nucleotídeos das posições 2 a 7. No caso do PITA, somente as predições com *score* menor ou igual a -10 foram consideradas.

Por fim, para melhor compreender o processo de repressão da ativação dos ovários, foi realizada uma rede de interação associando os mRNAs mais expressos no ovários ativados com os miRNAs mais expressos nos ovários inativos por meio do programa Cytoscape (Shannon *et al.*, 2003) o qual desenha as redes de interação.

3.4 Síntese de cDNA e PCR em Tempo Real para quantificação dos mRNAs e miRNAs

Para detecção e quantificação dos mRNAs e miRNAs, moléculas de cDNA de fita simples foram sintetizadas a partir de 2,0 µg de RNA total (item 3.3.1) utilizando-se

o kit NCode™ miRNA First-Strand cDNA Synthesis (Invitrogen), de acordo com as instruções do fabricante. Posteriormente, a primeira fita de cDNA foi utilizada em reações de PCR em tempo real.

Todos os *primers* utilizados foram analisados quanto à sua eficiência, para tanto, foi construída uma curva padrão para cada um dos pares de *primers* 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 para cada gene individualmente a partir do valor obtido de *slope* (obtido para a curva padrão) após as reações de curva padrão e utilizando a fórmula: $E = \frac{(10^{-1})}{\text{slope}}$ (Bulletin # 2, Applied Biosystems, 1997).

Os *primers* específicos utilizados para as quantificações dos mRNAs selecionados (para a validação experimental), encontram-se no Apêndice A item 1. Para assegurarmos que apenas um fragmento havia sido amplificado e que não havia contaminação por DNA genômico, foi feita a verificação através da curva de dissociação do produto da PCR. No caso de contaminação com DNA genômico, dois picos de dissociação seriam visualizados, pois, sempre que possível, os *primers* foram desenhados flanqueando um íntron.

Para cada miRNA foi usado um primer *forward* específico para o miRNA de interesse e um primer *reverse* universal fornecido pelo kit NCode™. O primer *forward* específico para cada miRNA foi desenhado utilizando a sequência completa do miRNA maduro (Apêndice A, item 2), com as bases uracilas das sequências substituídas por timinas. A especificidade dos iniciadores foi avaliada por BLASTN (Altschul *et al.*, 1990) contra o genoma da abelha *A. mellifera*.

Os níveis de expressão dos mRNAs e miRNAs de interesse foram mensurados por PCR quantitativa em tempo real utilizando-se o reagente SYBR® Green Master

Mix 2x (Applied Biosystems) e um aparelho espectrofluorimétrico 7500 Real Time PCR System (Applied Biosystems). As reações foram feitas em placas de 96 poços, em um volume final de 20 µl, sendo 10 µl de SYBR® Green Master Mix 2x (Applied Biosystems), 0,4 µl de cada primer (10 pmol/ µl), 2 µl de cDNA diluído (1:10) e 7,2 µl de água milli-Q autoclavada. Cada amostra foi testada em triplicata, o ciclo da PCR foi: 2 minutos a 50°C, 2 minutos a 95°C seguido por 40 ciclos de 15 segundos a 95°C e 35 segundos a 60°C.

A quantificação dos transcritos dos genes alvo (mRNAs) foi calculada a partir da diferença dos valores de Ct (*threshold PCR cycle*) em relação à média dos transcritos dos genes de controle endógeno, *proteína ribossomal -49 (rp49)* e *elongation factor 1-alpha (ef1-α)*. Já os resultados de miRNA foram normalizados com base na expressão dos miRNAs que apresentaram menores variações de expressão na condição estudada, desta maneira decidimos utilizar a média do miR-2, miR-184 e o gene U5 (componente do spliceossoma). Em nosso laboratório também estamos investigando qual o melhor calibrador para ser utilizado nas reações de PCR em tempo real a partir de moléculas de miRNAs.

Por fim, o cálculo da quantificação relativa dos transcritos foi realizado utilizando o método comparativo Ct (Applied Biosystems, User bulletin # 2).

4.0 Resultados

4.1 – Análise da morfologia dos ovários ativados de operárias

4.1.1 – Desenvolvimento dos ovócitos e células foliculares

A análise da morfologia dos ovários ativados de operárias de *A. mellifera* foi realizada com os seguintes propósitos: 1- analisar comparativamente a curva de crescimento das câmaras ovocíticas em relação a sua posição no vitelário e 2- quantificar o número de células foliculares que circundam cada um dos ovócitos.

Em nossas estimativas utilizamos apenas a região do vitelário uma vez que esta é a porção do ovaríolo que contem os folículos. Isso quer dizer que apenas quando notávamos claramente a presença de uma câmara nutridora e uma câmara ovocítica (folículos), sentido ápice - base, é que começávamos a fotografar as câmaras ovocíticas para posteriormente: A) classificar, B) mensurar e C) contar o número de células foliculares que circundam os ovócitos. A última câmara a ser fotografada e consequentemente analisada (figura 4) era a que precedia o desaparecimento da câmara nutridora (uma vez que após este evento verificamos a presença de apenas um ovócito maduro, praticamente pronto, como um ovo para ser botado). Assim, ao longo de um mesmo ovaríolo, foi possível classificar os ovócitos em uma escala de desenvolvimento de 1-9 ou 1-10 (sendo o menor classificado como número 1 e o maior número classificado ou como 9 ou 10).

No total analisamos onze ovaríolos e cada um deles pertencia a ovários de diferentes abelhas. Considerando-se os parâmetros pré-determinados, verificamos a presença de dez câmaras ovocíticas em seis ovaríolos (dos onze ovaríolos analisados) e

nos cinco demais, foram verificadas apenas nove câmaras ovocíticas em desenvolvimento.

Figura 4: Par de ovários ativados de operária de *A. mellifera*. CO - câmara ovocítica; CN – câmara nutridora; Setas apontam para uma câmara nutridora que já se esvaiu por completo. Microscopia de fluorescência (Axioscop 2, Zeiss) objetiva 5x, coloração por DAPI.

A altura e largura dos ovócitos foram estimadas e classificadas levando-se em consideração a posição dos folículos ao longo do vitelário (1-9 ou 1-10). De posse destes valores, calculamos a média e o desvio padrão juntando os valores obtidos para os onze ovaríolos. Os menores ovócitos, classificados como número 1, em média, apresentaram $38,6 \pm 9,0$ μm de altura e $61,4 \pm 10,3$ μm de largura. Os classificados como número 9 apresentaram $335,6 \pm 180,0$ μm de altura e $184,8 \pm 64,7$ μm de largura. E os classificados como número 10 apresentaram $452,6 \pm 76,2$ μm de altura e $239,4 \pm 26,7$ μm de largura. Na tabela II estão os valores médios e os desvios das câmaras analisadas, bem como o número de células foliculares.

Tabela II: Valores médios de altura; largura; área do folículo e número total de células foliculares calculados para os onze ovários analisados, e seus respectivos desvios.

Câmara ovocítica	Média Altura	Desvio	Média Largura	Desvio	Área do ovócito	Desvio	Número de células foliculares	Desvio
1	38,6	9,0	61,4	10,3	1860,7	491,8	108,5	29,0
2	48,9	10,1	62,4	13,3	2365,6	539,1	181,0	41,3
3	58,5	10,5	70,2	16,2	3329,4	1311,1	260,3	56,5
4	69,0	7,8	73,9	9,9	4016,8	747,4	344,9	95,0
5	77,0	9,0	76,9	11,9	4696,1	1100,6	460,7	94,4
6	93,6	8,7	88,9	13,3	6601,2	1492,8	598,3	138,9
7	113,1	24,0	99,2	16,0	8885,9	2692,0	766,4	173,4
8	176,7	54,0	127,9	27,7	18567,2	9161,5	1156,4	207,8
9	335,6	180,0	184,8	64,7	56604,7	44933,7	1573,9	486,4
10	452,0	76,2	239,4	26,7	86267,2	24240,8	1836,2	371,4

De posse dos valores de altura (A) e largura (L), calculamos a área da superfície dos ovócitos (aplicando a fórmula: $Pi \times \frac{A}{2} \times \frac{L}{2}$). Posteriormente, realizamos os cálculos das respectivas médias e desvios considerando-se os onze ovários em estudo (estes resultados também podem ser visualizados na tabela II). Em relação à área da superfície das câmaras ovocíticas, verificamos que a menor câmara em desenvolvimento (número 1) possui, em média, uma área de $1860,7 \pm 491,8 \mu m^2$ (figura 5). Enquanto que as maiores câmaras ovocíticas em desenvolvimento (classificadas como número 9 e número 10) possuem, em média, a área da superfície com os valores de $56604,7 \pm 44933,7 \mu m^2$ e de $86267,2 \pm 24240,8 \mu m^2$, respectivamente (figura 5). Utilizamos de todos os valores obtidos, relativos à área da superfície dos ovócitos, para realizar o teste estatístico “Kruskal-Wallis One Way Analysis Of Variance On Ranks”, por meio da ferramenta computacional Jandel SigmaStat 3.2. Como resultado obtivemos que os valores médios da área da superfície das câmaras ovocíticas consecutivas são estatisticamente diferentes entre si ($P < 0,001$), explicitamente o que este resultado sugere é que o tamanho médio

da área da superfície da câmara ovocítica de número 1 é significativamente diferente (menor) do tamanho médio da área da superfície da câmara de número 2 que é estatisticamente diferente (menor) do tamanho médio da área da superfície da câmara de número 3 e assim sucessivamente até a última câmara ovocítica (a de número 9 ou 10).

Figura 5: Valores médios da área da superfície das câmaras ovocíticas (N=11). As câmaras foram classificadas de acordo com o grau de desenvolvimento no ovaríolo (1-10 no sentido ápice - base).

Análises quantitativas referentes às variações citológicas durante o desenvolvimento dos ovários já foram realizadas em diferentes espécies de insetos, tais como *Habrobracon juglandis* (Hymenoptera) (Cassidy e King, 1972 e King e Cassidy, 1973), *Lucilia cuprina* (Diptera) (Mohanty, 1981), *Sarcophaga ruficomis* (Diptera) (Mohanty, 1982) e abelhas rainhas de *Apis mellifera* (Bueno e Morini, 1993). Em *A. mellifera*, a curva de crescimento dos folículos ovarianos e de seus componentes (câmara ovocítica e nutridora) em relação à sua posição no vitelário foi determinada em rainhas de abelhas africanizadas, italianas e caucasianas (Bueno e Morini, 1993). A análise dos resultados mostrou que os folículos ovarianos aumentam de tamanho em

uma razão mais ou menos constante até o 12° para a rainha italiana, 8° para a africanizada e 5° para a caucasiana. Daí em diante, a razão do crescimento entre os folículos aumenta. Este dado é concordante com o nosso resultado obtido em operárias órfãs africanizadas (figura 5). Ao analisarmos o crescimento das câmaras ovocíticas ao longo do ovariolo, também podemos observar que a partir da oitava câmara ovocítica há um aumento sensível da área dos ovócitos.

Determinada a curva de crescimento das câmaras ovocíticas (figura 5), por meio do cálculo área da superfície dos ovócitos (Área da Sup.), nosso próximo objetivo foi quantificar o número de células foliculares que circundam cada uma das câmaras ovocíticas (dos onze ovariolos). Para estimar o número total de células foliculares (Total Cel. Fol.), dividimos cada ovócito em quatro quadrantes e contamos o número de células foliculares (N. de Cel.) presentes em um retângulo de área pré-determinada (A. do Ret.), assim para o calculo final (Total Cel. Fol.) utilizamos a seguinte fórmula matemática:
$$Total\ Cel.\ Fol. = \frac{2 \times N.de\ Cel. \times \text{Área da Sup.}}{A.do\ Ret.}$$
 Como resultado, obtivemos que o ovócito de menor tamanho (número 1) possui em média $108,5 \pm 29$ células foliculares enquanto que os ovócitos de maior tamanho (número 9 e número 10) possuem $1573,9 \pm 486,4$ e $1836,2 \pm 371,4$ células foliculares, respectivamente. Os valores das demais câmaras ovocíticas (2-8), juntamente os com os agora citados, encontram-se na tabela II. Utilizamos todos os valores obtidos (número de células foliculares encontradas que rodeiam as câmaras ovocíticas (1-9 ou 1-10)) dos onze ovariolos e realizamos o teste estatístico “Kruskal-Wallis One Way Analysis Of Variance On Ranks”, por meio do programa computacional Jandel SigmaStat 3.2. Este teste revelou que os valores médios do número de células foliculares que circundam as câmaras ovocíticas consecutivas (classificadas de 1-9 ou 1-10) são estatisticamente diferentes

entre si ($P < 0,001$) (figura 6), resultado semelhante ao que verificamos a partir da área da superfície da câmara ovocítica.

Figura 6: Média do número de células foliculares presentes nas câmaras ovocíticas, classificadas de 1 a 10, ao longo da extensão de 11 ovariolos de ovários ativados de operárias de *A. mellifera*.

De posse dos resultados do crescimento das câmaras ovocíticas e da proliferação do epitélio folicular, sentido ápice-base, emergiu uma nova pergunta com relação à existência ou não de uma correlação entre área da superfície da câmara ovocítica e número de células foliculares que as circundam. Para averiguar a veracidade desta correlação utilizamos os seguintes valores: A) Área da Superfície da Câmara Ovocítica e B) Número de células foliculares presentes nesta área ($N.Cel. = \frac{Total\ Cel.Fol.}{2}$) e aplicamos o teste estatístico “Pearson Product Moment Correlation”, por meio da ferramenta computacional Jandel SigmaStat 3.2. O teste evidenciou um coeficiente de correlação positivo, com valor de 0,924, e o valor de p foi de 0,000134 ($P < 0,001$),

corroborando a hipótese de que o aumento no tamanho das câmaras ovocíticas é realmente acompanhado por um incremento no número de células foliculares, resumidamente, concluímos que ambos tendem a crescer juntos e de maneira correlacionada (figura 7).

Resultado semelhante foi o obtido por Ciudad Garrido, 2005 quando investigou o desenvolvimento dos ovócitos em *Blattella germanica*, além disso, descreveu que as três etapas da maturação do ovócito (pré-vitelogênica, vitelogênica e pós-vitelogênica) coincidem com mudanças do epitélio folicular. Na primeira etapa (pre-vitelogênica) o ovócito cresce muito pouco e existe entrada, por pinocitose, de moléculas presentes na hemolinfa; as mitoses das células foliculares acompanham o crescimento do ovócito e encontram-se estreitamente justapostas umas às outras impedindo o fenômeno da patência. Podemos associar as câmaras ovocíticas de 1 a 7 observadas em nosso estudo dentro desta descrição. O período vitelogênico se inicia com a incorporação das proteínas, em especial a vitelogenina, pelo ovócito e encerra-se com o início da síntese das proteínas do cório. A incorporação da vitelogenina foi caracterizada como um processo muito rápido que conduz a um aumento excessivamente grande da câmara ovocítica. As câmaras ovocíticas, de nosso estudo, que se enquadram dentro deste contexto são as de número 8 até a de número 10 (como já mencionado anteriormente). Em *Blattella germanica*, nesta etapa, é observada a formação de espaços entre as células foliculares (processo conhecido como patência), capacitando o ovócito a incorporar a vitelogenina (produzida pelo corpo gorduroso e liberada na hemolinfa) por meio de seus receptores de membrana (receptor de vitelogenina). Os resultados morfológicos mostrados a seguir, nesta tese, representam nossos esforços em caracterizar o fenômeno da patência em ovários de abelhas *A. mellifera*, até então desconhecidos. Na última fase de maturação, pós-vitelogênica, cessa a incorporação da vitelogenina, a patência

começa a desaparecer, as células foliculares secretam a membrana vitelina e as camadas relativas ao cório, uma vez formadas, inicia-se a degeneração progressiva das células foliculares, e segundo Bellés., 1993, todo este processo está sob a regulação da ecdisona. O período pós-vitelogênico não foi incluído em nosso estudo morfológico, mas em nossos estudos de expressão gênica, a presença dos ecdisteróides foi notória e serão posteriormente discutidas no contexto do desenvolvimento dos ovários de abelhas *Apis melliera*.

Figura 7: Média do número de células foliculares em relação ao valor médio da área da superfície da câmara ovocítica (os valores individuais dos onze ovariolos analisados foram compilados para o cálculo da média). Análise estatística (Correlação de Pearson) foi realizada por meio da ferramenta computacional Jandel SigmaStat 3.2, com os seguintes valores: coeficiente de correlação = 0,924 e valor de $p = 0,000134$.

4.1.2 – Análise da patência durante a oogênese em ovários de operárias (MET)

No passado muitos pesquisadores tentaram demonstrar e/ou inferiram que o percurso interfolicular de precursores de proteína do ovo seria universal entre os insetos (Telfer, 1965). Porém, investigações posteriores dedicadas exclusivamente a uma ou no máximo duas espécies demonstraram as particularidades que cada inseto possui em

relação a este processo (Telfer *et al.*, 1982; Lauverjat *et al.*, 1984; Raikhel e Lea, 1991). Em abelhas *A. mellifera*, até então, não havia sido realizado nenhum trabalho dedicado exclusivamente às características do epitélio folicular durante o processo de acumulação das proteínas do ovo. Neste estudo nós propusemos verificar como ocorre a incorporação de proteínas e outros componentes citoplasmáticos advindos da hemolinfa, em ovários ativados de abelhas operárias, mantidas em colônias órfãs. Para tanto, processamos nosso material de estudo para análise em microscópio eletrônico de transmissão.

Uma descrição geral para o processo denominado patência resume-se na contração das células do epitélio folicular que envolve o ovócito (em muitos insetos induzidos pelo hormônio juvenil), e desta maneira, aberturas e/ou espaços intercelulares no epitélio são originados permitindo assim, que a vitelogenina e outras proteínas presentes na hemolinfa, entrem em contato com a membrana celular do ovócito para serem sequestradas por pinocitose (Nation, 2002). Inicialmente analisamos uma câmara ovocítica (provavelmente no estágio 9 de desenvolvimento), em corte transversal, e por meio de nossas imagens, foi possível observar uma “rota” por entre as células foliculares, esta se inicia logo abaixo da lâmina basal e se estende/propaga até alcançar a superfície do ovócito, numa sequência de espaços (vesículas) intercalados por trechos de adesão celular. Há fortes indícios de que as proteínas precursoras do ovo estejam atravessando o interior destes espaços entre as células foliculares, e assim ficaria caracterizada a patência em abelhas *A. mellifera* (figura8).

Figura 8: Corte transversal de uma câmara ovocítica de ovário operária de *A. mellifera*, provavelmente no estágio 9 de desenvolvimento (classificação vide tabelaII). A) Vista panorâmica realizada por meio de microscopia de luz, nota-se no interior ovócito (OV) um número grande de grânulos de vitelo e ao redor do ovócito o epitélio folicular. B) Detalhe para as células foliculares (CF) ao redor do ovócito. C) Imagem realizada por meio de microscopia eletrônica de transmissão dá uma visão geral dos espaços intercelulares presentes entre as células foliculares. D) Detalhe para o espaço formado logo a abaixo da lâmina basal e entre duas células foliculares, materiais granulares e floculados (GR) também podem ser observados no interior deste espaço e dentro de vesículas (VS). E) Padrão de espaço intercelular observado desde a parte superior da câmara ovocítica (logo abaixo da lâmina basal) até as alcançarem o do ovócito. O que se observa são espaços se expandem na forma de vesículas e posteriormente se retraem na forma de finos canais (setas). F) Finos canais (setas) alcançando o espaço entre a o epitélio folicular e o ovócito aonde materiais estão sendo acumulados. Imagens A e B obtidas por meio do microscópio Axioscop 2, Zeiss, objetiva 10x e 40x respectivamente. Imagens C (aumento de 2,7x), D (aumento de 40x), E (aumento de 40x) e F (aumento de 40x) obtidas em microscópio Jeol JEM-100CXII do Laboratório de Microscopia Eletrônica do Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos da Faculdade de Medicina de Ribeirão Preto/USP.

É importante destacar algumas peculiaridades relativas aos espaços observados entre as células foliculares. A figura 8D mostra uma porção do epitélio folicular, aonde pode ser observado o desenvolvimento de um espaço intercelular na região logo abaixo

da lâmina basal, (com a presença, inclusive, de material particulado - GR), não obstante, este espaço se retrai – ganha adesão (flechas figura 8D,E,F) e em seguida se expande – perde adesão (aparentando uma vesícula (VS) – figura 8D,E), do mesmo modo é percebido sucessivas perdas e ganhos de adesão entre as membranas das células foliculares adjacentes que se manifestam até alcançarem a superfície do ovócito (figura 8). Um fato a se destacar é que no interior das vesículas, materiais particulados, também podem ser observados (figura 8D,E). Roth e Porter, 1964, em relação aos espaços que separam as células foliculares, atribuíram este fenômeno à concomitante redução das conexões desmossômicas (junções celulares). Dentro deste contexto, a separação e reaproximação das membranas das células foliculares adjacentes vistas em nossas imagens (figura 8D,E,F), poderiam ser explicadas por um decréscimo/redução desigual das junções desmossômicas ali presentes.

A permeabilidade do epitélio folicular termina quando absorção das proteínas do ovo pelos ovócitos cai drasticamente (Rubenstein, 1979; Lauverjat *et al.*, 1984). Neste momento as junções intercelulares que conferem a coesão, retomam sua atividade, levando a uma vedação dos espaços e assim, impedindo a movimentação de materiais entre as células (Rubenstein 1975, Lauverjat *et al*, 1984). Como podemos observar na figura 8A,B o ovócito já possui seu interior uma grande quantidade de granulos de vitelo (VT) estocados; o espaço entre as células foliculares e o ovócito (EI) está preenchido, aparentemente, por uma grande quantidade de material (provavelmente proteínas precursoras do vitelo) que está sendo gradualmente absorvido pelo ovócito (figura 8F). Um esquema um esquema representativo deste processo foi criado para o melhor entendimento do mecanismo (figura 9).

Figura 9: Esquema do mecanismo de patência observado em ovários de operárias de abelhas *A. mellifera*. A) Divisões das regiões do ovaríolo e destaque para o vitelário aonde se observa a incorporação das proteínas do ovo. Quadrado em vermelho representa a porção apresentada em maior aumento logo abaixo. B) Esquema dos espaços observados entre as células foliculares e a provável maneira de como as proteínas presentes na hemolinfa atingem o ovócito e se acumulam na forma de grânulos de vitelo.

Por não apresentar um padrão clássico de abertura entre as células foliculares, diante das informações acima citadas e analisando nossas imagens, nos questionamos se as células foliculares já estariam iniciando o processo de vedação (o que corresponderia à entrada no período de pós-vitelogênese) e em consequência deste fato não estaríamos verificando, entre as células foliculares, o que de fato corresponderia à patência. Para

conferir esta hipótese selecionamos outra câmara ovocítica, de menor tamanho, provavelmente no estágio 7 de desenvolvimento (figura 10).

Figura 10: Corte transversal de uma câmara ovocítica própria de um ovário de abelha operária *A. mellifera*, provavelmente no estágio 7 de desenvolvimento (classificação

vide tabelaII). A) Vista panorâmica realizada por meio de microscópio de luz, nota-se no interior ovócito (OV) um número grande de vesículas e ao redor do ovócito o epitélio folicular. B) Imagem obtida por Microscopia eletrônica de transmissão, detalhe para 1- o número de vesículas no interior das células foliculares (CF) e 2- o espaço (EI) entre o epitélio folicular e o ovócito (praticamente vazio). C) Materiais granulares estão alcançando a superfície do ovócito e se acumulando no espaço intercelular (EI) D) Entre as células foliculares adjacentes também podem ser observadas as vesículas (VS) com material no seu interior e canais finos (flechas) que conectam as vesículas. E) Finos canais (flechas) alcançando o espaço entre a o epitélio folicular e o ovócito aonde as prováveis proteínas do ovo estão sendo acumulados. F) Outra região do epitélio folicular mostrando a chegada de partículas granulares para o espaço formado entre o epitélio folicular e o ovócito. Imagens A obtidas por meio do microscópio Axioscop 2, Zeiss, objetiva 10x. Imagens B (aumento 5x) C (aumento de 10x), D (aumento de 20x), E (aumento de 20x) e F (aumento de 20x) obtidas em microscópio Jeol JEM-100CXII do Laboratório de Microscopia Eletrônica do Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos da Faculdade de Medicina de Ribeirão Preto/USP.

Ao observar a figura 10 três características se distinguem fortemente em relação à figura 8. A primeira delas é a ausência dos amplos corpos proteicos (grandes granulos de vitelo) no interior do ovócito, acredita-se que estes corpos se formam, no ovócito em um maior grau de desenvolvimento, por meio da fusão das várias unidades protéicas (material particulado) que vão 1- se acumulando no espaço entre o ovócito e as células foliculares e 2- gradualmente vão sendo absorvidos por pinocitose. A segunda delas é a presença marcante de um grande número de vesículas 1- entre as células foliculares; 2- no interior das células foliculares e 3- no interior do ovócito na região próximo a borda. Supõe-se que as vesículas presentes no interior das células foliculares podem estar contribuindo com a saída de moléculas em direção aos ovócitos. As vesículas entre as células foliculares podem realmente representar o mecanismo de patência nas abelhas *A. mellifera*. E é provável que as vesículas presentes no interior dos ovócitos façam parte da formação dos granulos de vitelo contribuindo no processo de aglomeração dos mesmos. Finalmente a última característica é a ausência de um aglomerado de material no interior do espaço entre o ovócito e as células foliculares, mas sim, a presença de pequeno granulos proteicos que certamente estão iniciando o processo de aglomeração

que se mostra intensa nos ovócitos mais desenvolvidos (figura 8A,B). Desta maneira acreditamos que conseguimos acrescentar novas informações a respeito da morfologia dos ovários em desenvolvimento com destaque para o processo de patência, até então não caracterizado para as abelhas *A. mellifera*.

4.2 Sequenciamento do Transcriptoma de ovários de *A. mellifera*

Para avaliar e caracterizar a expressão dos genes de acordo com o estado fisiológico dos ovários de abelhas *A. mellifera*, selecionamos a técnica de “sequenciamento de nova geração” (RNAseq) cuja abordagem nos permite medir os níveis de transcritos (conhecidos e não conhecidos) em larga escala.

Sete diferentes amostras de cDNA foram sequenciadas de acordo com as condições inicialmente estabelecidas: 1- ovários inativos de operária – abelhas com comportamento de forrageira (FOR); 2- ovários inativos de operárias com 4 dias de vida adulta - na presença da rainha (CR); 3- ovários inativos de operárias com 4 dias de vida adulta – órfãs (SR); 4- ovários ativados de operárias – sem ovócito maduro (E1) 5- ovários ativados de operárias – com ovócito maduro (E2); 6- ovário de rainha em atividade de postura (RA) e 7- ovário de rainha virgem (RV). Todas as amostras foram sequenciadas usando o protocolo estabelecido pela *Illumina (Life Sciences)* e resultaram em aproximadamente 200 milhões de *reads*, que variaram em tamanho de acordo com a plataforma utilizada (*Genome Analyser II* - 36pb e *HiScanTMSQ* – 50pb). O tamanho do fragmento sequenciado e a distribuição do número de leituras para cada uma das amostras estão apresentados na tabela III.

Após o tratamento dos fragmentos sequenciados (eliminação dos adaptadores e sequências de má qualidade), foi realizado o mapeamento das sequências obtidas com o genoma de *A. mellifera* (versão 4.5) disponível na base de dados *Beebase* (<http://hymenopteragenome.org/beebase/>). Inicialmente foram removidas as sequências de baixa qualidade para em seguida realizar o alinhamento por meio dos algoritmo *Tophat* (Trapnell *et al.*, 2009). A porcentagem de sequências de baixa qualidade e os resultados do mapeamento no genoma de *A. mellifera* (versão 4.5) também estão apresentados na tabela III.

Tabela III- Análise primária resultante do sequenciamento e alinhamento das sequências com o genoma de *A. mellifera*.

Amostras	Tamanho das leituras	Total de leituras	% Total de leituras mapeadas	% Leituras unicamente mapeadas	% de Leituras de baixa qualidade
FOR	50	30786072	90,1	86,9	5,7
CR	36	34054425	87,1	83,3	6,3
SR	36	28627489	85,4	81,8	7,1
E1	36	39613880	87,6	84,4	4,9
E2	36	29368699	93,8	90,7	4,7
RA	50	34366505	92,5	90,2	4,3
RV	50	28436074	91,2	87,8/	4,9

Após as etapas de normalização, foram calculados os valores de RPKM e de *fold-change* ($\log_2$), utilizando os algoritmos *Cufflinks* e *Cuffdiff* (Trapnell *et al.*, 2010). Os dados gerados para as sete bibliotecas foram utilizados para compor três combinações diferentes de investigação dos dados das bibliotecas de ovários de

diferentes origens: A) ovário em estado funcional *versus* estado não funcional - ovários de operárias (CRxE1; CRxE2; SRxE1 e SRxE2); B) ovários de rainha em postura *versus* ovários de rainha virgem (RAxRV) e C) ovários inativos de operárias de diferentes amostras (FORxCR; FORxSR e CRxSR).

Um limiar de corte de $\log_2$ ratio entre 2,0 e -2,0 foi estabelecido para qualificar os genes como diferencialmente expressos. Além disso, como critério, foi estipulado que para um gene ser considerado diferencialmente expresso deveria apresentar em pelo menos uma das bibliotecas um valor de RPKM ≥ 5 . Vale ressaltar que os valores de RPKM apresentados neste trabalho são, na verdade, RPKM+1; este procedimento foi realizado porque quando um gene, em pelo menos uma das bibliotecas em comparação, possui RPKM=0 não é possível realizar o cálculo de *fold-change* ($\log_2$).

4.2.1 Análise do estado funcional *versus* estado não funcional - ovários de operárias (CRxE1; CRxE2; SRxE1 e SRxE2) e ovários de rainha (RAxRV)

A análise conjunta do estado fisiológico funcional *versus* estado não funcional (ovários de operárias) resultou, de maneira geral, em 3120 genes diferencialmente expressos, dos quais 2616 representam o estado não funcional e 504 representam o estado funcional. Desta associação apenas 45 genes mostraram-se contra regulados, o que em porcentagem, representa uma pequena parcela de 1,44%. Analisando separadamente cada um dos pareamentos observamos que 73 genes são exclusivos do pareamento CRxE1; 425 são exclusivos de CRxE2; 36 são exclusivos de SRxE1 e 91 são exclusivos de SRxE2. Do total de genes diferencialmente expressos (3120 genes), 34,68% (1082 genes) são comuns em todos os pareamentos (CRxE1; CRxE2; SRxE1 e SRxE2). A distribuição completa destes genes pode ser observada na figura 11.

Figura 11: Diagrama de Venn mostrando a distribuição dos 3120 genes diferencialmente expressos nos pareamentos entre as bibliotecas de ovários funcionais *versus* as bibliotecas de ovários não funcionais (E2xCR, E2xSR, E1xCR e E1xSR). A) Distribuição dos 2616 genes mais expressos nos ovários não funcionais. B) Distribuição dos 505 genes mais expressos nos ovários funcionais.

A análise entre rainha em postura *versus* rainha virgem resultou em 2153 genes diferencialmente expressos, dos quais 1612 representam o ovário de uma rainha virgem e 541 representam o ovário de uma rainha em postura (figura 12). Ao comparar este resultado com o resultado obtido em operária (acima descrito), verificamos que 1389 genes são compartilhados nas duas abordagens. Dos 1612 genes que representam o ovário de uma rainha virgem, 1109 também representam um ovário não funcional (ou inativo) de operária, ou seja, 68,8% dos genes altamente expressos em RV estão também altamente expressos em CR e/ou SR. Paralelamente, dos 2616 genes encontrados como mais expressos no ovários inativos de operárias (CR e/ou SR), 1109 estão mais expressos em RV, indicando uma correspondência de 42,4%. Do outro lado, rainhas em postura e operárias com ovários funcionais (ou ativados) compartilham 280 genes o que representa uma correspondência de 51,7% para rainhas em relação às operárias e 55,5% para as operárias em relação às rainhas.

Figura 12: Diagrama de Venn mostrando a distribuição dos genes analisados no pareamento rainha virgem *versus* rainha acasalada. Em rosa estão representados os 1612 genes mais expressos em rainhas virgens, em verde escuro estão representados os genes que não apresentaram diferença de expressão e em verde claro estão representados os 541 genes mais expressos em rainhas em postura.

Reunindo ambos os resultados: A) estado funcional *versus* estado não funcional - ovários de operárias (CRxE1; CRxE2; SRxE1 e SRxE2) e B) rainha em postura *versus* rainha virgem (RAxRV), verificamos também que a quantidade de genes diferencialmente expressos nos ovários inativados supera a quantidade de genes diferencialmente expressos nos ovários ativados e de forma análoga constatamos o mesmo em rainhas (o número de genes diferencialmente expressos em ovário de RV é superior ao número de genes diferencialmente expressos em ovário de RA). Resultado semelhante foi observado por Cardoen *et al.*, 2012, em seu estudo do proteoma de ovários ativados e inativos de abelhas *A. mellifera*. Os autores analisaram 1666 *spots* o que resultou em 613 *spots* com expressão diferencial. O número de *spots* com alta expressão nos ovários inativos foi de 336 e o número de *spots* com expressão aumentada em ovários ativados foi de 227. Ao realizar a identificação destes *spots*, os autores obtiveram 61 proteínas altamente expressas nos ovários ativados e 119 proteínas altamente expressas nos ovários inativos. De acordo com os processos biológicos obtidos, os autores resgataram a hipótese de Tanaka e Hartfelder, 2004, de que a supressão da ativação do ovário deve envolver uma interação constante entre ovogênese

primordial e subsequente degradação. Estes autores aventaram uma hipótese contrária à visão de que a repressão da ativação dos ovários das abelhas operárias seria um processo que ocorreria uma única vez na vida das abelhas. É muito provável que frequentemente esteja ocorrendo repressão da ativação dos ovários. Diante deste cenário inferimos que a manutenção do estado inativo em relação à manutenção do estado ativado e a preparação do órgão para um estado funcional (que produz ovos), possivelmente, demande uma complexidade maior de interações e/ou regulações de genes e de vias metabólicas que por consequência conduz a um gasto energético maior. Seria, portanto, a manutenção do estado inativo um processo mais complexo. Acreditamos que essas circunstâncias possam justificar um maior número de genes diferencialmente expressos nos ovários não funcionais quando os comparamos aos ovários funcionais.

Os resultados acima citados representam todo o universo de genes considerados como diferencialmente expressos nos pareamentos E2xCR; E2xSR; E1xCR; E1xSR e RAxRV. Entretanto, utilizamos um limiar de corte mais rigoroso de $\log_2 \text{ratio}$ entre 5,0 e -5,0, no que tange os níveis de expressão gênica, objetivando estreitar e/ou salientar as diferenças (existentes) entre as bibliotecas em comparação. Dentro deste critério, 245 genes foram considerados como diferencialmente expressos quando comparamos os 4 pareamentos de operárias (ovários funcionais *versus* ovários não funcionais). E quando comparamos rainha virgem *versus* em postura, inserido no contexto desta nova análise, 233 foram considerados como diferencialmente expressos. Ao agrupar os resultados obtidos a partir dos pareamentos de operárias (E2xCR, E2xSR, E1xCR e ExSR) com o pareamento de rainhas (RAxRV), verificamos que A) 75 genes apresentavam $\log_2 \text{fold change} \geq 5$ para ambas as condições, portando unindo os 245 genes de operárias com os 233 genes de rainhas obtivemos em 403 genes diferencialmente expressos; B) 38 genes

se mostravam contra regulados, ou seja, apenas 9,4% dos genes associados possuem um padrão de expressão antagônico quando comparamos operárias (E2xCR, E2xSR, E1xCR e ExSR) com rainhas (RAxRV). Estes genes foram utilizados para realizar agrupamentos hierárquicos supervisionados, com parâmetros de distância Euclidiana e método de ligação média. O agrupamento por distância Euclidiana agrupa os genes de acordo com os níveis de transcrição. Desta forma, para cada um dos genes selecionados, plotamos os valores RPKM obtidos em cada uma das seis bibliotecas (RA; E2; E1; RV; SR e CR), este resultado (figura 13) foi categorizado em um *heatmap*. As cores nestes *heatmaps* indicam a variação na expressão dos genes, sendo que o vermelho indica alto nível expressão, o rosa indica uma expressão intermediária e o branco indica uma baixa expressão (quando comparados os valores referência - RPKM).

O *heatmap* mostra que alguns genes se agrupam por apresentarem perfis de expressão semelhantes nas comparações realizadas, chamadas assinaturas de expressão gênica. Por meio destas assinaturas podemos perceber que rainha em postura e operárias com ovários ativados possuem um padrão de expressão dos genes bastante parecidos. O mesmo se repete em relação às rainhas virgens e operárias com ovários inativos, sugerindo um padrão de expressão gênica similar entre RV, SR e CR (figura 13).

Assim como RA, E2 e E1 formam um grupo distinto, em relação ao perfil de expressão gênica de RV, SR e CR (figura 13) o resultados também nos revela que: A) operárias com ovários ativados (E1 e E2) se relacionam mais entre si do que, por exemplo, operárias com ovócitos maduros (E2) e rainhas em postura (que também possuem ovócitos maduros) e B) operárias com ovários inativos (SR e CR) também formam um grupo mais homogêneo entre si do que, por exemplo, operárias na iminência de ativar seus ovários (SR – não possuem mais um estímulo de repressão) com rainha virgem (estas precisam apenas de um estímulo para ativar seus ovários). Por

este motivo realizamos um novo agrupamento hierárquico (com parâmetros de distância Euclidiana), categorizados em um *heatmap*, utilizando somente os 245 genes diferencialmente expressos nas bibliotecas de operárias (figura 14).

Figura 13: Perfis de expressão (*heatmap*) da associação dos genes de rainhas e operárias com $\log_2 \text{fold change} \geq 5$, quando comparados os estados fisiológicos funcionais e não funcionais dos ovários. Foram plotados, para cada um dos 403 genes, os valores de RPKM calculados em cada uma das 6 bibliotecas (RA, E2, E1, RV, SR, CR).

Figura 14: Perfis de expressão (*heatmap*) dos 245 genes de operárias com $\log_2 \text{fold change} \geq 5$, quando comparados os estados fisiológicos funcionais e não funcionais dos ovários. Foram plotados, para cada um dos genes, os valores de RPKM calculados em cada uma das bibliotecas (E2, E1, SR, CR).

Por meio deste novo *heatmap* (figura 14) foi possível detectar grupos de genes com perfis de transcrição bastante distintos - 8 no total - diferente ao observado na figura 13, onde não conseguimos distinguir, com clareza, os grupos de genes com perfis de transcrição semelhantes. Até o momento, não nos aprofundamos no estudo destes distintos grupos observados na figura 14, porém, a partir dos grupos 2 e 8, que

representam os melhores grupos de genes que discriminam um ovário funcional de um ovário não funcional de operárias, selecionamos alguns para validação experimental por meio de RT-PCR. Outra perspectiva que esta análise gerou foi em relação ao grupo 7, acreditamos que um estudo concentrado neste grupo nos darão subsídios para encontrar as moléculas sinalizadoras que desencadeiam e/ou iniciam o processo de ativação dos ovários, uma vez que a maioria dos genes deste grupo se apresenta expresso apenas na amostra SR.

Para testar a reprodutibilidade dos nossos resultados e verificar se é eficaz a comparação de duas diferentes metodologias de análise global da expressão gênica, utilizamos os resultados de *microarray* (amostras de corpo inteiro de abelhas operárias reprodutivas e não reprodutivas) de Cardoen *et al.*, 2011 e os apresentamos na forma de *heatmaps*. Esta análise foi realizada da seguinte forma: A) listamos os 100 genes mais expressos referentes às abelhas operárias reprodutivas e os 100 genes mais expressos referentes às abelhas operárias não reprodutivas (Cardoen *et al.*, 2011), B) associamos a esta lista de genes – 200 no total (Cardoen *et al.*, 2011) os valores de RPKM obtidos em cada uma de nossas bibliotecas geradas por meio de sequenciamento de “nova geração” (RA, E2, E1, CR, SR, FOR, RV) e B) agrupamos hierarquicamente, com parâmetros de distância Euclidiana, e apresentamos na forma de dois *heatmaps* sendo o primeiro: genes altamente expressos em abelhas operárias reprodutivas (figura 15) e o segundo: genes altamente expressos em abelhas operárias não reprodutivas (figura 16).

Figura 15: Perfis de expressão (*heatmap*) dos 100 genes altamente expressos em corpo inteiro de abelhas operárias reprodutivas obtidas a partir de análises por *microarray* de Cardoen *et al.*, 2011. Foram plotados, para cada um dos genes, os valores de RPKM calculados em cada uma de nossas sete amostras (RA, RV, E2, E1, SR, CR e FOR), cuja análise dos transcritos se baseou no sequenciamento de “nova geração” (RNAseq).

Figura 16: Perfis de expressão (*heatmap*) dos 100 genes altamente expressos em corpo inteiro de abelhas operárias não reprodutivas obtidas a partir do estudo de *microarray* de Cardoen *et al.*, 2011. Foram plotados, para cada um dos genes, os valores de RPKM calculados em cada uma de nossas sete amostras (RA, RV, E2, E1, SR, CR e FOR), cuja análise dos transcritos se baseou no sequenciamento de “nova geração” (RNAseq).

Estas análises nos permitiram verificar uma clara segregação entre abelhas reprodutivas (amostras RA, E2 e E1), em relação ao perfil de expressão gênica, de abelhas não reprodutivas (amostras RV, SR, CR e FOR). Estes resultados foram completamente compatíveis com os resultados obtidos por Cardoen *et al.*, 2011. E assim, podemos demonstrar que: A) as metodologias de *microarray* e RNAseq podem fornecer resultados tão robustos que apesar de utilizarem estratégias experimentais diferentes são perfeitamente comparáveis e com alto grau de confiabilidade, nos permitindo inferir que o sequenciamento do transcriptoma de nossas amostras foi extremamente eficaz mesmo sem dispormos de réplicas biológicas e B) embora nossa biblioteca seja exclusiva de ovários e as amostras de Cardoen *et al.*, 2011 sejam de corpo inteiro, neste caso, o *status* fisiológico do organismo testado prevaleceu, garantindo uma comparação eficaz mesmo partindo de amostras biológicas, não iguais mas, semelhantes.

4.2.2 Análise comparativa das bibliotecas de ovários inativos (não funcionais) de operárias *A. mellifera* (CR, SR e FOR)

Realizamos os três pareamentos possíveis (CRxSR, FORxCR; FORxSR) das bibliotecas de ovários inativos de operárias, e os analisamos inicialmente de maneira isolada. Utilizamos um limiar de corte $\log_2 \text{ratio}$ entre 2,0 e -2,0 e $\text{RPKM} \geq 5$ em pelo menos 1 das bibliotecas, como dito anteriormente. Como resultado obtivemos que quando comparadas as bibliotecas CRxSR, 246 genes encontram-se como diferencialmente expressos, destes, 125 se encontram altamente expressos em CR e 121 estão altamente expressos em SR (figura 17A). Quando comparadas as bibliotecas FORxCR, verificamos a presença de 356 genes diferencialmente expressos, dos quais, 154 possuem uma expressão elevada em FOR e 202 possuem uma expressão elevada

em CR (figura 17B). Por fim, 308 genes foram encontrados como diferencialmente expressos no pareamento FORxSR, dentre eles 162 encontram-se mais expressos em FOR e 146 encontram-se mais expressos em SR (figura 17C).

Figura 17: Diagrama de Venn mostrando a distribuição de genes encontrados como diferencialmente expressos nos ovários ativos e inativos de operárias: A) CRxSR; B) CRxFOR e C) SRxFOR.

Deste universo de genes diferencialmente expressos: A) 43 foram encontrados apenas no pareamento CRxSR sendo 23 altamente expressos em CR e 20 altamente expressos em SR; B) 127 foram encontrados apenas no pareamento CRxFOR, sendo 35 altamente expressos em FOR e 92 altamente expressos em CR e C) 51 genes foram encontrados apenas no pareamento SRxFOR, sendo 27 altamente expresso em FOR e

24 altamente expressos em SR. Estes genes exclusivos de cada pareamento, juntos totalizaram 221 genes. Os demais genes, não exclusivos, ou eram comuns aos três pareamentos (CRxSR + CRxFOR + SRxFOR) ou eram comuns a pelo menos dois 2 pareamentos (CRxSR +CRxFOR ou CRxSR+SRxFOR ou CRxFOR+SRxFOR) e totalizaram 340 genes. Assim geramos uma lista com 561 genes diferencialmente expressos (340 compartilhados e 221 exclusivos), deste listamos 78 genes que apresentavam em pelo menos 1 dos três pareamentos $\log_2 fold change \geq 5$ e realizamos um agrupamento hierárquico supervisionado, com parâmetros de distância Euclidiana e método de ligação média, categorizados em *heatmap* (figura 18).

Figura 18: Perfis de expressão (*heatmap*) dos 78 genes de operárias com $\log_2 fold change \geq 5$, quando comparados entre si os estados fisiológicos não funcionais dos ovários. Foram plotados, para cada um dos genes, os valores de RPKM calculados em cada uma 3 das bibliotecas (FOR, SR, CR).

A análise do *heatmap* nos permitiu verificar que as duas situações de abelhas operárias com ovários inativos na presença da rainha, mesmo estas abelhas apresentando idades diferentes (CR – 4 dias de vida adulta e FOR – mais de 12 dias de vida adulta), são mais próximas entre si do que quando comparadas com a amostra de ovários inativos de operárias (com 4 dias de vida adulta) na ausência da rainha (SR). Esta é mais uma evidência de que a amostra SR deve estar expressando genes próprios da etapa inicial da ativação dos ovários.

Quando realizamos o desenho experimental de nosso estudo decidimos por realizar o sequenciamento de uma amostra de ovário de forrageiras (FOR) por que até aquele momento, experiências pessoais aliados a alguns dados na literatura nos faziam crer que estas abelhas, em função da idade, da alimentação e da atividade que desempenham dentro da colônia, não eram mais capazes de ativar seus ovários. Porém um estudo revelou (Naeger *et al.*, 2013) que forrageiras em colônias órfãs podem ativar seus ovários, além disso, os autores ainda compararam o grau de ativação dos ovários de abelhas forrageiras e não forrageiras de acordo com a idade e concluíram que: A) abelhas forrageiras e não forrageiras com 14 dias de vida adulta não apresentam diferenças no grau de ativação dos ovários, entretanto B) abelhas forrageiras com 21 dias de vida adulta tendem a apresentar um grau de ativação maior do que aquelas que não estão forrageando, e as forrageiras são mais propícias a desenvolver totalmente seus ovários, contendo pelo menos 1 ovócito maduro em cada um de seus ovários (Naeger *et al.*, 2013). Nossas bibliotecas colocam as forrageiras em uma situação intermediária entre as operárias com rainha e sem rainha, o que pode sugerir a existência de potencialidade para assumir tarefas outras que não a de forragear.

4.2.3 Categorização funcional dos genes diferencialmente expressos quando comparamos amostras de ovários funcionais com amostras de ovários não funcionais

Realizamos uma categorização funcional para os 3120 genes encontrados como diferencialmente expressos quando comparamos amostras de ovários funcionais *versus* não funcionais de operárias *A. mellifera* (critério $\text{ratio log}_2 \text{fold change} \geq 2$). E o mesmo realizamos para os 2153 genes encontrados como diferencialmente expressos quando comparamos as amostras de rainha virgem *versus* rainha em postura (critério $\text{ratio log}_2 \text{fold change} \geq 2$). Os nossos resultados referentes a esta análise podem ser visualizados na tabela IV e dentre as diversas informações geradas, dois aspectos, se destacaram e serão explorados em seguida.

Tabela IV – Categorização funcional dos genes encontrados como diferencialmente expressos quando comparamos o estado funcional *versus* o estado não funcional tanto para operárias quanto para rainhas.

Classificação funcional	E1 e E2		RA		SR e CR		RV	
	Número	%	Número	%	Número	%	Número	%
Desconhecido	196	38,9	201	37,2	1257	48,1	652	40,4
Transdução de sinal	49	9,7	48	8,9	315	12,0	186	11,5
Secreção	48	9,5	43	7,9	150	5,7	125	7,8
Regulação nuclear	25	5,0	26	4,8	13	0,5	7	0,4
Maquinaria de transcrição	25	5,0	33	6,1	89	3,4	33	2,0
Citoesqueleto	23	4,6	23	4,3	138	5,3	80	5,0
Transporte/estocagem	21	4,2	21	3,9	112	4,3	86	5,3
Metabolismo de lipídio	14	2,8	16	3,0	26	1,0	30	1,9
Matriz extracelular / Adesão celular	14	2,8	19	3,5	143	5,5	80	5,0
Maquinaria de modificação de proteína	12	2,4	25	4,6	39	1,5	33	2,0
Metabolismo de carboidrato	11	2,2	4	0,7	29	1,1	37	2,3
Fator de transcrição	9	1,8	11	2,0	58	2,2	38	2,4
Maquinaria de transporte de proteína	6	1,2	9	1,7	22	0,8	19	1,2
Metabolismo oxidativo/ Desintoxicação	6	1,2	7	1,3	38	1,5	40	2,5
Sinal de transdução, apoptose	6	1,2	6	1,1	23	0,9	16	1,0
Maquinaria de síntese de proteína	5	1,0	8	1,5	13	0,5	11	0,7
Maquinaria do proteossomo	5	1,0	6	1,1	6	0,2	9	0,6
Viral	5	1,0	4	0,7	21	0,8	16	1,0
Inibidor de proteinase secretada	4	0,8	1	0,2	12	0,5	11	0,7
Metabolismo intermediário	3	0,6	5	0,9	2	0,1	4	0,2
Elemento transponível	3	0,6	3	0,6	26	1,0	9	0,6
Imunidade	3	0,6	4	0,7	10	0,4	10	0,6
Secreção protease	3	0,6	2	0,4	19	0,7	22	1,4
Metabolismo de nucleotídeo	2	0,4	3	0,6	2	0,1	5	0,3
Metabolismo de aminoácido	2	0,4	5	0,9	17	0,6	21	1,3
Metabolismo, energia	1	0,2	4	0,7	12	0,5	12	0,7
Secreção, mucins	1	0,2	1	0,2	9	0,3	2	0,1
Estocagem	1	0,2	1	0,2	3	0,1	3	0,2
Secreção, imunidade	1	0,2	0	0,0	12	0,5	15	0,9
Transporte nuclear	0	0,0	2	0,4	0	0,0	0	0,0
Total	504		541		2616		1612	

O primeiro aspecto a ser discutido se refere ao elevado número de genes desconhecidos e, portanto não classificados funcionalmente. Este ponto em questão pode ser associado a duas diferentes situações: (1) alguns transcritos classificados como desconhecidos são na verdade RNAs não codificadores longos (lncRNA); isto significa que: A) dos 196 transcritos desconhecidos altamente expressos em ovários ativados de operárias 23 são prováveis lncRNA (12%); B) dos 201 transcritos desconhecidos altamente expressos em ovários de rainhas em postura 26 são prováveis lncRNA (13%); C) dos 1257 transcritos desconhecidos altamente expressos em ovários inativos de operárias 98 são prováveis lncRNA (8%) e D) dos 652 transcritos desconhecidos altamente expressos em ovários de rainhas virgens 48 são prováveis lncRNA (7,4%); (2) fizemos uso de uma metodologia nova para estudos moleculares em larga escala (sequenciamento de nova geração), que foi classificada como “sistema aberto”, por Asmann *et al.*, 2008 e Teng e Xiao, 2009, pelo fato de ser possível identificar, em estudos de transcriptoma, além de padrões de expressão, novos mRNA, novos miRNAs e novos tipos de RNAs (como os lncRNA); assim consideramos e verificamos que muitos genes classificados como desconhecidos nem sequer constavam na última versão do genoma de *A. mellifera* (v. 4.0 – até então a versão mais utilizada) e por esse motivo são: 1) pouco explorados e 2) em alguns casos contém erros de anotação - o que dificulta qualquer comparação com outros organismos e consequentemente a busca por genes ortólogos.

O segundo aspecto trata-se das classes que podem representar um ovário funcional e as classes que podem caracterizar um ovário não funcional. A classificação funcional se mostrou, num primeiro momento, muito parecida quando comparamos ovários funcionais e não funcionais. A maioria das diferenças constatadas nos ovários não funcionais é relativa aos processos de: transdução de sinal e matriz extracelular e

adesão. Enquanto que nos ovários funcionais os processos que se destacaram foram: secreção, maquinaria de transcrição, metabolismos de lipídio e regulação nuclear. Dentre todas as classes funcionais citadas, regulação nuclear é a classe que melhor distinguiu, em nossas análises, um ovário funcional de um ovário não funcional. Assim, nos aprofundamos na análise dos genes que constam dentro desta classificação. A tabela V mostra todos os genes, inseridos dentro da classificação - regulação nuclear, presentes nos ovários ativados de operárias e nos ovários de rainhas em postura, enquanto a tabela VI mostra todos os genes, inseridos em regulação nuclear, próprios de ovário inativo de operária e de ovário de rainha virgem. Ao analisar essas tabelas fica evidente a alta concentração de histonas nos ovários funcionais: em operárias dos 25 genes inseridos em regulação nuclear 15 são histonas, e em rainhas dos 26 genes inseridos em regulação nuclear 12 são histonas; quando consideramos os ovários não funcionais, dentro dos critérios aqui impostos, rainhas não possuem nenhuma histona dentre os 7 genes inseridos em regulação nuclear e operárias possuem apenas 1 histona num universo de 13 genes inseridos em regulação nuclear.

Cinco classes principais de histonas são encontradas em todas as células eucarióticas, diferindo no peso molecular e na composição de aminoácidos: H1, H2A, H2B, H3 e H4. As H3 e H4 são praticamente idênticas na sequência de aminoácidos em todos os eucariotos, sugerindo conservação estrita das suas funções, enquanto que as histonas H1, H2A e H2B apresentam um menor grau de homologia de sequência entre as espécies eucarióticas. Nossos resultados mostram que todas estas classes estão altamente expressas nos ovários funcionais de abelhas e 1 histona H4 está altamente expressa em ovários inativos de operárias. Estas proteínas são responsáveis pelo empacotamento e ordenação do DNA em unidades estruturais chamadas nucleossomos, cada nucleossomo contém oito moléculas de histonas, duas cópias de cada uma das

H2A, H2B, H3 e H4, a histona H1, por sua vez, está geralmente associada ao agrupamento dos nucleossomos repetitivos formando então uma estrutura maior denominada cromatina. Assim podemos dizer que metade da massa da cromatina é DNA e a outra metade é histona e quando o DNA se replica antes da divisão celular, grandes quantidades de histonas são também sintetizadas para manter uma proporção 1:1 (Lehninger *et al.*, 2005). Esta pode ser uma primeira explicação para fato de termos encontrado tantas histonas nos ovários funcionais de abelhas *A. mellifera*, visto que estes se encontram em intenso processo de proliferação celular. Além disso, podemos observar na tabela V, a presença de uma ciclina, que são proteínas sintetizadas durante o ciclo celular, já que, durante a ovogênese, há um intenso processo de divisão celular (mitoses) das células foliculares. Outro aspecto extremamente relevante, em relação às histonas, é que elas possuem em sua extremidade uma região amino-terminal, que por se exteriorizar do nucleossomo, pode sofrer modificações pós-transcricionais, como por exemplo: acetilação, fosforilação, metilação e ubiquitinação. A cromatina se encontra normalmente superespiralada, a acetilação de histonas torna a cromatina mais relaxada, permitindo assim que proteínas ligadoras do DNA e/ou regulatórias da transcrição possam se ligar às regiões promotoras dos genes (Ting *et al.*, 2006; Motet e Castronovo, 2008; Esteller, 2008; Valeri *et al.*, 2009). A acetilação pode ocorrer nas histonas H2B, H3 e H4. Podem ser acetilados, na histona H3, os resíduos de lisina (K) 9, 14, 18, 23 e 27 e na histona H4, as lisinas 5,8,12 e 16 (Santos-Rosa e Caldas, 2005). As enzimas acetilases (HAT) e desacetilases (HDAC) são responsáveis pelo estado de acetilação das histonas, estas modificações, assim como a metilação do DNA, são denominadas alterações epigenéticas cujas principais características estão firmadas no fato de serem: estáveis, reversíveis e potencialmente herdáveis na expressão gênica não afetando a sequência de nucleotídeos do DNA (Jiang *et al.*, 2004; Esteller e Almouzni, 2005). A

tabela V nos mostra duas prováveis enzimas histonas acetilases representadas pelo GB46742 e GB42837 e a tabela VI nos mostra uma provável enzima histona desacetilase (GB43234). Além das duas acetilases, também se encontra como altamente expressa, nos ovários funcionais de rainha e operárias, uma provável enzima histona metiltransferase (GB41135). Quando ocorrem modificações das lisinas e/ou argininas da extremidade N-terminal das histonas H3 e H4, cada um destes aminoácidos pode receber um, dois ou três grupos metil através da enzima histona metiltransferase (HMTase), como consequência, há um reflexo na compactação da cromatina. Dependendo do tipo de metilação (mono, di ou tri) e de qual aminoácido sofreu a modificação, a metilação pode tanto compactar a cromatina, impedindo a transcrição, quanto pode descompactá-la, permitindo a atividade gênica. A associação destes resultados suporta que o controle da ativação dos ovários está sobre forte influência de fatores epigenéticos.

Tabela V – Identificação de todos os genes inseridos dentro da classificação regulação nuclear em ovários funcionais. A segunda e terceira coluna informa se o gene consta nas amostras de operárias ou na amostra de rainha. A quarta coluna mostra a descrição conforme consta no NCBI.

Número de acesso (Beebase)	Operária Ativado	RA	Descrição - NCBI
GB47380	sim	não	histone H3-like (LOC725414)
GB47381	sim	não	histone H2A-like (LOC725308)
GB47505	sim	não	histone H2B.3-like (LOC725344)
GB47382	sim	não	histone H4-like (LOC725230)
GB41135	sim	não	histone-lysine N-methyltransferase pr-set7 (pr-set7)
GB53606	sim	não	eggless (egg)
GB42872	sim	não	G1/S-specific cyclin-E (CycE)
GB42920	sim	não	PAP-associated domain-containing protein 5-like (LOC552662)
GB54609	sim	não	CAP-D2 condensin subunit (CAP-D2)
GB47486	sim	sim	histone H3-like (LOC726542)
GB47406	sim	sim	similar to germinal histone H4 (LOC726520)
GB47485	sim	sim	histone H2B.1/H2B.2-like (LOC726487)
GB47483	sim	sim	histone H2A-like (LOC726468)
GB47408	sim	sim	histone H2B.3-like (LOC726447)
GB47507	sim	sim	histone H2A-like (LOC725450)
GB44765	sim	sim	histone H2B.3-like (LOC724633)
GB47407	sim	sim	histone H4 (LOC406132)
GB47482	sim	sim	histone H1-like (LOC100577833)
GB43303	sim	sim	histone H2A variant (His2Av)
GB50658	sim	sim	DNA primase large subunit (DNAPol-alpha60)
GB50678	sim	sim	2-oxoglutarate and iron-dependent oxygenase (LOC725925)
GB49401	sim	sim	protein msta, isoform A-like (LOC725783)
GB50036	sim	sim	Sld5 protein (Sld5)
GB45014	sim	sim	nuclear pore complex protein Nup93-like (LOC410343)
GB47506	sim	sim	uncharacterized LOC725238 (LOC725238)
GB44339	não	sim	histone H2B.3-like (LOC724869)
GB46742	não	sim	histone acetyltransferase type B catalytic subunit-like (LOC552748)
GB46889	não	sim	SWI/SNF- regulator of chromatin subfamily A-like protein 1-like (LOC551616)
GB40362	não	sim	flap endonuclease 1 (LOC412308)
GB42837	não	sim	N-acetyltransferase eco (eco)
GB42106	não	sim	chromatin assembly factor 1 subunit (Caf1)
GB50035	não	sim	DNA replication complex GINS protein SLD5-like (LOC100577899)
GB47822	não	sim	DNA polymerase epsilon subunit 4-like (LOC725835)
GB45490	não	sim	regulator of chromosome condensation 1 (RCC1)
GB50994	não	sim	mitotic spindle assembly checkpoint protein Mad2 (mad2)

Tabela VI – Identificação de todos os genes inseridos dentro da classificação regulação nuclear em ovários não funcionais. A segunda e terceira coluna informa se o gene consta nas amostras de operárias ou na amostra de rainha. A quarta coluna mostra a descrição conforme consta no NCBI.

Número de acesso (Beebase)	Operária Inativo	RV	Descrição - NCBI
GB43234	sim	não	histone deacetylase 4 (HDAC4)
GB48162	sim	não	anoctamin-8-like (LOC408908)
GB53220	sim	nao	ubiquitin carboxyl-terminal hydrolase 3-like (LOC408619)
GB45732	sim	não	uncharacterized LOC726879 (LOC726879)
GB48386	sim	não	suppressor of variegation 3-3 (Su(var)3-3)
GB54998	sim	não	bromodomain adjacent to zinc finger domain protein 2B-like (LOC551071)
GB55527	sim	não	uncharacterized LOC408649 (LOC408649)
GB41520	sim	sim	transducin-like enhancer protein 4-like (LOC411175)
GB43079	sim	sim	mutS protein homolog 4-like (LOC726455)
GB44950	sim	sim	uncharacterized LOC100578910 (LOC100578910)
GB50247	sim	sim	uncharacterized LOC100577551 (LOC100577551)
GB50303	sim	sim	nc-112-related protein-like (LOC408836)
GB51423	sim	sim	WD repeat-containing protein 63-like (LOC409463)
GB52626	não	sim	DNA mismatch repair protein - MLH3 family

4.2.4 Categorização funcional dos genes diferencialmente expressos em CRxSR

Realizamos uma categorização funcional para os 246 genes encontrados como diferencialmente expressos quando comparamos amostras de ovários não funcionais (CRxSR) de operárias *A. mellifera* (critério $\text{ratio } \log_2 \text{ fold change} \geq 2$). A figura 19 mostra a classificação dos genes altamente expressos na amostra SR e a figura 20 apresenta os genes com maior expressão na amostra CR.

Figura 19: Categorização funcional dos 121 genes encontrados como altamente expressos na amostra SR quando comparada com a amostra CR.

Figura 20: Categorização funcional dos 125 genes encontrados como altamente expressos na amostra CR quando comparada com a amostra SR.

Essa análise foi realizada com o intuito de verificar se a amostra SR nos traria subsídios para avaliar os processos que pudessem estar relacionados com a etapa inicial da ativação dos ovários, porém, como podemos observar na figura 19 e na figura 20 o número de genes desconhecidos é tão elevado que qualquer conclusão a respeito das classes funcionais encontradas podem não refletir verdadeiramente quais processos seriam mais representativos para cada uma das amostras.

4.2.5 Validação experimental por PCR em Tempo Real (RT-PCR)

Outra maneira de verificar a eficiência e qualidade dos dados gerados nos experimentos de RNAseq é por meio de validações experimentais. Portanto, procedemos à validação por meio da técnica de PCR em Tempo Real (RT-PCR). Selecionamos 9 genes de interesse que diferenciam um ovário funcional de um ovário não funcional. Os genes selecionados podem ser visualizados na tabela VII, esta tabela também informa os valores calculados, nas análises de dados do RNAseq, de RPKM para amostras RA, E2, E1, RV, SR e CR, além dos valores de $\log_2 fold change$ para os pareamentos RAxRV, E2xCR, E2xSR, E1xCR e E1xSR. Vale ressaltar que os RNAs extraídos para a síntese do cDNAs utilizados nestes experimentos de RT-PCR são totalmente independentes daqueles utilizados para o sequenciamento (RNAseq).

A figura 21 mostra os níveis de transcritos de cada um dos 6 genes selecionados que apresentam uma maior expressão no ovários funcionais. E a figura 22 mostra os níveis de transcritos de cada um dos 3 genes selecionados que apresentam uma alta expressão nos ovários não funcionais. As colunas de ovários inativos de operárias correspondem à média de nove amostras biológicas: três de FOR, três de CR e três de SR. As colunas de ovários ativados de operárias correspondem à média de seis amostras

biológicas: três de E1 e três de E2. Em rainhas, cada uma das colunas corresponde à média de três amostras individuais de ovários RV ou RA. Para proceder estas análises utilizamos o método de $2^{-\Delta\Delta CT}$.

As análises por RT-PCR mostraram que a flutuação dos níveis de transcritos entre as duas condições analisadas (ovários funcionais e não funcionais) foi extremamente marcante. Os transcritos dos genes: GB42182, *ffps5*, *homeótico caudal*, *Obp7*, *yellow-g* e GB44975 encontram-se significativamente mais expressos nos ovários ativados de operárias e nos ovários de rainhas em postura e os transcritos dos genes: GB53748, *yellow-x2* e *vg* apresentam uma expressão estatisticamente maior nos ovários inativos de operárias e nos ovários de rainhas virgens. Como podemos observar, todos os genes mostrados nas figuras 12y e 13y apresentaram diferenças significativas (One Way ANOVA $p \leq 0,001$) entre as duas circunstâncias de estudo (ovários funcionais *versus* ovários não funcionais) e confirmaram os resultados obtidos por RNAseq, ou seja, a reprodutibilidade e eficiência dos nossos resultados de RNAseq foi mais uma vez confirmada. Nos itens a seguir discutiremos individualmente cada um dos genes aqui analisados.

Tabela VII – Genes selecionados para validação experimental por RT-PCR. As colunas RA, E2, E1, RV, SR e CR mostram os valores de RPKM calculados a partir dos experimentos de RNAseq. E as colunas RAxRV, E2xCR, E2xSR, E1xCR e E1xSR mostram os valores de $\log_2$ fold change.

Genes selecionados	RA	E2	E1	RV	SR	CR	RAxRV	E2xCR	E2xSR	E1xCR	E1xSR
<i>fpps5</i>	5469	2602	2388	15	10	2	8,5	10,7	8,0	10,6	7,9
GB42182	3498	776	389	8	2	2	8,7	8,5	8,8	7,6	7,8
GB44975	906	1606	1471	4	10	5	7,7	8,3	7,3	8,1	7,2
<i>Homeótico caudal</i>	223	908	639	3	7	3	6,5	8,1	7,0	7,6	6,5
Obp7	116	295	178	1	1	1	6,4	8,0	7,8	7,3	7,1
Yellow-g	614	578	525	6	4	6	6,6	6,5	7,1	6,4	6,9
<i>vitelogenina</i>	2	2	9	670	39	2	-8,5	-0,4	-4,7	2,3	-2,1
<i>yellow-x2</i>	1	1	7	281	19	24	-7,8	-4,0	-3,7	-1,7	-1,4
GB53748	2	2	3	54	73	118	-5,0	-6,1	-5,4	-5,4	-4,7

Figura 21: Quantificação relativa dos níveis de transcritos dos 6 genes altamente expressos em ovários funcionais quando comparados aos ovários não funcionais. A

coluna de ovário ativo representa a média de 6 amostras biológicas: 3 *pools* com 7 ovários de E1 e 3 *pools* com 5 ovários de E2. A coluna de ovário inativo representa a média 9 amostras biológicas: 3 *pools* com 15 ovários de FOR e o mesmo em relação à CR e SR. No gráficos de rainhas, cada uma das colunas corresponde à média de três amostras individuais de ovários RV ou RA. As análises estatísticas foram realizadas por meio “One Way Anova” utilizando a ferramenta Jandel SigmaStat 3.2.

Figura 22: Quantificação relativa dos níveis de transcritos dos 3 genes altamente expressos em ovários não funcionais quando comparados aos ovários funcionais.. A coluna de ovário ativo representa a média de 6 amostras biológicas: 3 *pools* com 7 ovários de E1 e 3 *pools* com 5 ovários de E2. A coluna de ovário inativo representa a média 9 amostras biológicas: 3 *pools* com 15 ovários de FOR e o mesmo em relação à CR e SR. No gráficos de rainhas, cada uma das colunas corresponde à média de três amostras individuais de ovários RV ou RA. As análises estatísticas foram realizadas por meio “One Way Anova” utilizando a ferramenta Jandel SigmaStat 3.2.

4.2.5.1 GB42182

O gene representado pelo GB42182 sintetiza uma proteína não caracterizada, LOC100578571. Consta nas análises de Cardoen *et al.*, 2011 que este gene, representado pelo GB16065 - na versão 4.0 do genoma - seria um fator de transcrição *mbk-1-like* muito expresso nos corpos cogumelares de abelhas *A. mellifera* e na mosca *Drosophila melanogaster*. Porém, por meio de nossos resultados de RNAseq, verificamos que o gene ao qual o autor se refere não condiz com a sequência de aminoácidos preditos para GB42182/GB16065. Investigamos qual seria o número de acesso para o gene *mbk-1-like* ou *ecdysteroid-regulated gene E93/mbk-1*, como também é conhecido, e constatamos que GB com número de acesso GB50048 (versão 4.5 do genoma) ou GB17328 (versão 4.0 do genoma) é o que representa este gene em abelhas *A. mellifera*, cuja expressão, inclusive, é maior nos ovários inativos com $\log_2 \text{fold change} \geq 2$. Como já foi dito, o GB42182 sintetiza uma proteína não caracterizada e está altamente expressa nos ovários ativados de operárias e nos ovários de rainhas em postura. Busca por ortólogos no site do *National Center for Biotechnology Information* (NCBI), utilizando a ferramenta BlastP, nos trouxeram como resultados proteínas hipotéticas presentes em apenas dois Hymenoptera: a abelha *Bombus impatiens* (com 38% de identidade) e a formiga *Acromyrmex echinator* (com 28% de identidade), sugerindo que seja uma proteína exclusiva da ordem dos himenópteros. Para reforçar nossas análises em relação à busca por ortólogos, utilizamos a ferramenta computacional BlastP disponível no banco de dados *Universal Protein Resource* (UniProt), encontramos uma identidade de 22% para com o a proteína de *Drosophila mojavensis* com número de acesso GI13678. Por meio desta análise fomos capazes de categorizar funcionalmente nossa proteína de acordo com o *Gene Ontology* (<http://www.geneontology.org/>). No que diz respeito à função molecular como resultado

obtivemos que esta proteína possui função molecular associada com microtúbulo/interação de cromatina (acesso GO:0003677), o termo mais próximo, relacionado à este, que se destacou foi: *DNAbinding* (GO:0003674). Em relação ao componente celular, averiguamos que esta proteína provavelmente se encontra dentro do núcleo, para confirmar este resultado utilizamos a ferramenta computacional *Predicting nuclear localization of proteins* – NucPred (Brameier *et al.*, 2007) e verificamos que a proteína representada pelo GB42182 apresentou 77% de probabilidade de estar dentro do núcleo. No que tange os processos biológicos obtivemos que é desconhecido o papel que a proteína codificada por este mRNA desempenha (GO:0008150), os termos mais próximos, relacionado à este, que se destacaram foram: A) processos celulares (GO:0009987) e dentro desta categoria: 1- processos metabólicos (GO:0008152), 2- processos relacionados ao desenvolvimento (GO:0032502) e regulação biológica (GO:0065007); B) regulação de processos biológicos (GO:0050789), e dentro desta classe: 1- organização de componentes celulares ou biogênese (GO:0071840), 2- localização (GO:0051179), 3 - reprodução (GO:0000003) e 4- sinalização (GO:0023052).

4.2.5.2 FPPS5

A enzima *farnesil difosfato sintase* (FPPs) é um membro da família das preniltransferases; as enzimas que pertencem a esta família estão relacionadas com a biosíntese de terpenos (Onhuma *et al.*, 1996). A enzima FPPS tem estrutura homodimérica e cataliza o acoplamento sequencial de dimetilalil difosfato (DMAPP) com duas moléculas de IPP, produzindo o intermediário, farnesil difosfato (FPP) (Szkopinska e Plochocka, 2005). O FPP é utilizado como precursor de uma série de

compostos como dolicois, ubiquinonas, colesterol, fitoesteróides e sesquiterpenos. Em insetos, o sesquiterpeno mais conhecido é hormônio juvenil (Riddiford, 1996; de Kort e Granger, 1996; Gilbert *et al.*, 2000). A FPPS, devido a sua importância biológica na via de síntese do HJ, foi estudada em *A. mellifera* (Bomtorin, *et al.*, 2013, in press) *Bombyx mori* (Koyama *et al.*, 1985, Kikuchi *et al.*, 2001), *Agrotis ipsilon* (Castillo-Gracia e Couillaud, 1999) e *Manduca sexta* (Sen e Sperry, 2002; Kinjoh *et al.*, 2007). Em *B. mori* três diferentes FPPSs foram identificadas, sendo que duas delas têm expressão quase exclusiva nos *corpora allata* (Kinjoh *et al.* 2007).

Em *A. mellifera* são conhecidas 7 isoformas de FPPSs, os genes parecem ter sofrido vários processos de mutação em suas seqüências. A isoforma em estudo (GB48899 – *fpfs5*) perdeu parte de dois éxons e um íntron completo. A isoforma representada pelo GB46095 (*fpfs3*) já foi apontada como sendo a maior responsável pela conversão de geranyl difosfato em farnesil pirofosfato durante a síntese do HJ, no entanto, as demais isoformas não foram descartadas como participantes da síntese do HJ em abelhas *A. mellifera*, ainda que em menor escala (Santos, 2008). Curiosamente, em abelhas adultas, em contraste com as mamangavas e muitos outros insetos (Bloch *et al.*, 2000), a oogenese e a ativação dos ovários não é influenciada pelo HJ (Robinson *et al.*, 1991; Robinson *et al.*, 1992;. Robinson & Vargo 1997). Em vez disso, o HJ em conjunto com vitelogenina, regula a divisão de trabalho relacionado à idade (polietismo estário), incluindo a transição de tarefa, de nutriz para forrageira (Robinson & Vargo 1997; Robinson & Huang, 1998; Guidugli *et al.* 2005). Por outro lado, os ecdisteróides são conhecidos por regular a síntese de vitelogenina em larvas das abelhas (Barchuk *et al.* 2002) e os títulos de 20-hidroxiecdisona (20E) parecem ser ligeiramente maiores em operárias poedeiras, em colônias órfãs (Robinson *et al.*, 1991; Hartfelder *et al.* 2002).

Nós verificamos que a *fpps7* (GB43671) é a única *fpps* que não se expressa em ovários, a *fpps2* (GB52653) não apresentou diferença de expressão entre as amostras e com exceção da *fpps5*, de maneira geral, as demais *fpps* possuem valores de RPKM consideravelmente baixos. As *fpps1* (GB46055) *fpps6* (GB48897) estão preferencialmente expressas nas amostras de ovários inativos e na amostra de rainha virgem, com valores semelhantes de $\log_2 \text{fold-change}$ (*fpps1* $\log_2 \geq 2$ e *fpps6* $\log_2 \geq 1,4$). As enzimas *fpps3*, *fpps4* (GB48898) e *fpps5* apresentaram-se preferencialmente expressas em ovários funcionais de abelhas *A. mellifera* e de maneira geral, todas apresentaram $\log_2 \text{fold-change}$ maior que 1,5 quando comparadas aos estados não funcionais. Sem dúvida a *fpps5* foi a *fpps* que se destacou em nossas análises de bioinformática e, portanto, selecionada para validação experimental. As análises dos transcritos, por qPCR, corroboraram as análises *in silico*. Podemos sugerir, portanto, com base na expressão de *fpps5* que durante a oogenese a sinalização pelos ecdisteróides é um componente importante, senão essencial, para o início do processo de ativação dos ovarios. Até então apenas Cardoen *et al.*, 2012, em sua investigação proteômica e transcriptômica (Cardoen *et al.*, 2011) havia sugerido essa relação. Estes autores apontaram potenciais ligantes de hormônios esteróides (como o GB48124), proteínas contendo repetições tetratricopeptídeos (FKBP59, GB40746) e *heat shock proteins* (*hsp-70* e *hsp-90*, que fazem parte do complexo receptor de esteróides) expressos preferencialmente em ovários ativados. Confrontando estas informações com os resultados de nosso sequenciamento, verificamos que os transcritos que sintetizam estas moléculas também se encontram preferencialmente expressos em ovários ativados de operárias e em rainhas em postura. Os resultados de Cardoen, *et al* 2012, corroboraram anteriores relativo as análises transcriptômicas (*microarray*) (Cardoen; *et al.*, 2011) no qual os autores relataram que os genes *ecdysone-induced gene 74*

(GB49105), *broad complex* (GB48271), ambos com função associada à ovogênese, e o próprio *fpps5* (GB48899) estavam significativamente mais expressos em corpo inteiro de abelhas operárias reprodutivas, quando comparados com as abelhas estéreis. Em nosso sequenciamento os valores de RPKM para os GB49105 e GB48271 são substancialmente baixos em todas as amostras e em relação à expressão diferencial ($\log_2 fold\text{-}change$) não foram observadas diferenças entre as amostras, a conjunção destes resultados nos permite inferir que estes dois genes provavelmente estão se expressando em outro órgão ou tecido que contribuem para o estado fisiológico de uma abelha reprodutiva. Estes resultados lançam uma nova luz sobre o papel dos ecdisteroides na reprodução de abelhas *A. mellifera*, já que no passado suas influências no processo reprodutivo mostraram controvérsias e por um bom tempo foram muito contestadas (Hartfelder *et al.*, 2002).

4.2.5.3 Homeótico caudal

O gene *caudal* (*cad*) é considerado um gene de efeito materno, umas das suas principais características está na expressão dos mRNAs pelas células nutridoras dos ovários de insetos com posterior transferência para dentro ovócitos em desenvolvimento e, por fim, participação nos estágios iniciais do desenvolvimento embrionário (Olesnický *et al.*, 2006). De uma maneira geral, o papel desempenhado pelos genes considerados de efeito materno, nos embriões, está na padronização e/ou especificação do eixo anteroposterior durante as fases iniciais do desenvolvimento. A proteína homeótica caudal (Cad) é fundamental na modulação do polo posterior de muitos embriões, este papel é conservado ao longo da evolução desde *C. elegans* até mamíferos (Olesnický *et al.*, 2006). A localização e expressão do gene que sintetiza a proteína Cad (GB48966) foi analisada em ovários de rainhas em postura de *A. mellifera*, por Wilson

et al., 2010. Os autores observaram que *cad* encontra-se altamente expresso na região do vitelário e que a transcrição ocorre em todas as células nutridoras. As células nutridoras, por sua vez, estão conectadas ao ovócito no seu polo anterior e por meio de pontes citoplasmáticas seus transcritos são transportados para dentro dos ovócitos. Por meio de hibridação *in situ* e microscopia confocal, os autores localizaram os transcritos de *cad* também no córtex anterior dos ovócitos. Esta localização atípica de *cad* no polo anterior se mantém no embriões de 0 a 4 horas após a postura, nos estágios seguintes, os mRNAs do gene *cad* começam a migrar para o polo posterior, durante a blastoderme sincicial inicia-se a transcrição zigótica de *cad* e somente durante a gastrulação a expressão é observada em uma faixa limitada na parte posterior do embrião. A injeção de RNA dupla fita (dsRNA) nos embriões, visando a diminuição de transcritos de *cad*, levaram a uma variedade de fenótipos, independente da localização dos mRNAs de *cad*, todos os embriões apresentaram algum grau de perda nas estruturas e segmentos posteriores e abdominais. Os autores sugeriram que a expressão atípica, no polo anterior, observada nos ovócito e nos embriões de até 4 horas, está diretamente relacionada à função que *cad* desempenha nos ovócitos, visto que em embriões, os papéis desempenhados por *cad* se restringem ao polo posterior. Porém, não só nas abelhas *A. mellifera*, mas na maioria dos insetos, nenhuma função foi atribuída diretamente às proteínas caudais presentes nos ovócitos em desenvolvimento. Em insetos, os autores que mais próximo chegaram a uma atribuição de função para Cad nos ovários foram Olesnicky *et al.*, 2006. Estes autores se propuseram a investigar a função de *cad* nas fêmeas e nos embriões das vespas *Nasonia vitripennis*, para tanto, eles injetaram dsRNA em pupas e observaram o desenvolvimento até a fase adulta. Curiosamente, as fêmeas tratadas com altas concentrações de dsRNA tiveram sua postura afetada (poucos embriões) e os embriões derivados dessas mães deixaram de se

desenvolver não alcançando a fase pupal. As fêmeas injetadas com controle GFP (dsRNA) não apresentam defeitos cuticulares e nenhuma diferença na postura de ovos foi observada. Olesnicky *et al.*, 2006 também chegaram a conclusão de que a função desempenhada por Cad durante a ovogênese seja diferente do seu papel no início do desenvolvimento embrionário. Diante deste resultado os autores postularam que Cad expressa nos ovários pode estar associada à fertilidade e a viabilidade dos ovos de uma maneira geral.

Drosophila melanogaster possui os exemplos mais bem estudados de mRNAs de efeito materno. A maioria deles está localizada nos polos anterior ou posterior do ovócito e conduzem a uma tradução localizada de produtos proteicos no embrião. Alguns exemplos são 1- o gene *bicoid* (ausente em *A. mellifera*) está localizado no polo anterior (Frohnhofer e Nusslein-Volhard, 1986; Berleth *et al.*, 1988) cuja função está associada a repressão da tradução de *cad* nesta região (Rivera-Pomar *et al.*, 1996); 2- o gene *nanos* (GB52596/GB52597), traduzido no polo posterior (Gavis e Lehmann, 1994; Wang *et al.*, 1994) previne que neste polo do embrião seja traduzido *hunchback* (GB50986) (Irish *et al.*, 1989; Wreden *et al.*, 1997). Apesar de *cad* ter sido o mRNA de efeito materno que mais se destacou nas análises *in silico* de nossas bibliotecas, nos levando a validação experimental por RT-PCR, nos propusemos a analisar a expressão digital destes e outros mRNAs de efeito materno em nossas bibliotecas. Como resultado, verificamos que *nanos* possui baixos valores de RPKM e quando comparadas as bibliotecas, os valores obtidos em relação à expressão diferencial ($\log_2$ fold change), foram menores que 1,5. O mesmo se repetiu em relação ao gene *hunchback*. O mRNA *orthodenticle* (GB45500) possui uma localização similar a encontrada para o mRNA *cad* nos ovários de rainhas em postura de abelhas *A. mellifera* (Wilson e Dearden, 2011), mas nos embriões, sua expressão está restrita ao polo anterior, ao analisar seus valores

de RPKM e de $\log_2$ *fold change*, novamente o resultado encontrado para *nanos* e *hunchback* se repetiu. E de maneira intrigante obtivemos o mesmo resultado para o gene *tailless*, que é um fator de transcrição importante para a especificação da extremidade posterior do embrião, é maternalmente expresso no lado dorsal dos ovócitos dos ovários de rainhas (Wilson e Dearden, 2009), e não se expressa maternalmente em outros insetos (Garcia-Solacheetal *et al.*, 2010; Lemke *et al.*, 2010; Liaw e Lengyel, 1993; Lynch *et al.*, 2006; Schroder *et al.*, 2000). Outros mRNAs, em *Drosophila melanogaster*, estão envolvidos na localização espacial dos mRNAs tanto durante a ovogênese como no início do desenvolvimento embrionário como por exemplo: *oskar* (ausente em abelhas), *staufer* (GB54264), *cappuccino* (GB44119), *spire* (GB45685), *swallow* (ausente em abelhas) e *exuperantia* (GB42261) (Lantz *et al.*, 1992). A expressão digital dos transcritos destes genes foi analisada em nossas bibliotecas de ovários e nenhum deles apresentou diferenças de expressão quando comparamos ovários os funcionais com os não funcionais e com exceção de *staufer* e *exuperantia*, todos apresentaram valores de RPKM baixos. Os genes de efeito materno estão discretamente expressos em nossas bibliotecas, com exceção de *caudal*, que deve ter um papel relevante não apenas na embriogênese, mas também no processo de ovogênese. No entanto, estes dados não são conclusivos e exigem estudos funcionais.

4.2.5.4 Obp7

As principais proteínas receptoras olfatórias envolvidas na percepção de odores em insetos, as OBPs, *odorant binding proteins* (OBPs), *odorant-degrading enzymes* (EDOs), *odorant receptors* (ORs), *ionotropic receptors* (IRS) e *sensory neuron membrane proteins* (SNMPs) (para revisão ver Leal, 2013) fazem parte de uma classe

especializada de proteínas, presentes em insetos e vertebrados, cujas principais funções estão no 1- reconhecimento das moléculas hidrofóbicas propagadas pelo ar e na 2- facilitação do transporte destas moléculas até os receptores olfatórios. Apesar de compartilharem o mesmo nome e desempenharem funções similares, filogeneticamente, as OBPs de insetos e vertebrados parecem não se relacionar (Pelosi 1996; Krieger e Breer 1999; Deyu e Leal 2002; Forêt e Maleszka, 2006).

A maioria dos genes codificadores de OBPs, em *A. mellifera*, estão organizados em *clusters* no genoma, refletindo uma expansão relativamente recente dos genes desta família, também observado em *D. melanogaster* e *Anopheles gambiae*. Confrontando com *D. melanogaster* (51 OBPs identificadas), *Anopheles gambiae* (70 OBPs identificadas) e *Tribolium castaneum* (46 OBPs identificadas); *A. mellifera* são os insetos que possuem o menor repertório de OBPs - 21 OBPs identificadas (Forêt e Maleszka, 2006; Robertson e Wanner, 2006). As OBPs de insetos são moléculas pequenas (15 to 18 kDa), solúveis em água e com evidências experimentais de expressão gênica nas sensilas olfativas e gustativas, bem como em outros tecidos especializados (Pelosi *et al.* 2005).

Forêt e Maleszka, 2006, analisaram a expressão dos 21 genes presentes no genoma de *A. mellifera* e verificaram que a maioria se expressava nos tecidos olfatórios. Os autores encontraram três padrões distintos de expressão. O primeiro trata-se daqueles expressos exclusivamente nas antenas de abelhas adultas: *obp1*, *obp2*, *obp4*, *obp5*, *obp6*, *obp8*, *obp11*, e *obp12*. Vale ressaltar que *obp6* e *obp8* são único caso documentado de *splicing* alternativo para as *obps* em abelhas (Forêt e Maleszka, 2006). A segunda categoria inclui *obps* que possuem expressão ubíqua sendo encontradas em quase todas as partes do corpo adulto: *obp3*, *obp16*, *obp17*, *obp18*, *obp19/obp20* e *obp21*. Um caso particularmente interessante dentro deste grupo é *obp3*, que foi

encontrado em todas as partes do corpo com exceção das antenas. Finalmente, a última categoria contém *obps* expressas durante estágios de desenvolvimento relativamente curtos. As *obp14* e *obp15* são encontradas em larvas e após pupação desaparecem. A *obp13* é altamente expressa nas larvas tardias e ao longo das fases de pupa. A *obp10* aparece em pupas e atinge o nível mais alto no cérebro de abelhas recém-emergidas. Os transcritos da *obp9* foram detectados nos ovários de rainhas e nas primeiras horas de do desenvolvimento dos embriões, consistente com padrão de expressão materno, os transcritos de *obp7* foram encontrados exclusivamente nos ovários rainha. Sugere-se que as *obps* encontrados em tecidos não olfatórios, possuem funções restritas ao transporte geral com ampla especificidade para compostos lipofílicos (para revisão ver Leal, 2012).

Analisando os dados de nosso transcriptoma verificamos a expressão das 21 *obps*. Como resultado, observamos que não há transcritos referentes às: *obp1*, *obp 2*, *obp 5*, *obp 6*, *obp 8*, *obp 15*, *obp 16*, *obp 18*, *obp 19*, *obp 20* e *obp 21*. Em suas investigações, Forêt e Maleszka, 2006 verificaram, dentre outras, a expressão de *obp4*, *obp11* e *obp12* como sendo exclusivas de antenas, porém, por meio de nosso sequenciamento, verificamos a expressão das *obps* acima citadas em nossas bibliotecas de ovários. No caso da *obp11* (GB50937) não foram verificadas diferenças de expressão entre as bibliotecas. As análises dos dados do sequenciamento mostraram que a *obp12* (GB40616) e a *obp4* (GB53372), encontram-se mais expressas em ovários inativos e em ovários de rainhas virgens, o mesmo foi verificado em relação à *obp10* (GB50936). Dentre elas a que mais nos chamou a atenção foi a *obp4*, pois, foi verificado um $\log_2$ *fold change* $\geq 5,2$ quando comparadas as bibliotecas de ovários de rainhas virgens e a de ovários de rainhas em postura e, em especial, nas operárias, foi verificada uma maior expressão nos ovários inativos de operárias na ausência da rainha (pareamento SRxE2

com $\log_2 fold\ change \geq 2,3$; pareamento SRxCR $\log_2 fold\ change \geq 1,3$). Os ovários das rainhas recém-emergidas encontram-se na fase inicial da ovogênese e necessitam apenas do acasalamento ou de um estímulo artificial (narcose com CO₂) para se desenvolver por completo e produzir ovos, é provável que alguns mRNAs estejam sendo transcritos nas rainhas recém-emergidas mas as proteínas codificadas por esses genes serão observadas tardiamente (já nos ovários desenvolvidos), o mesmo pode acontecer no caso das operárias, a amostra SR foi classificada como um ovário inativo, mas, pela condição ambiental (ausência da rainha), esta abelha não está mais reprimindo o processo de ativação dos ovários, pelo contrário, esta abelha deve estar se preparando e/ou sintetizando mRNAs cuja tradução será requerida na fase funcional (ovário ativado); sugere-se que este seja motivo pelo qual moléculas transportadoras de lipídios, como o caso da *obp4*, possam estar apresentando uma alta expressão nestas amostras. Em rainhas virgens também observamos uma maior expressão das *obp3* (GB53371), *obp13* (GB46222) e *obp17* (GB46226), nas bibliotecas de operárias elas estão preferencialmente expressas nos ovários ativados. No caso das *obps*: *obp7* (GB53370), *obp9* (GB50151) e *obp14* (GB46223), foi observada uma maior expressão nas rainhas em postura e nas operárias com ovários ativados. As *obp7* e *obp9* já foram observadas como preferencialmente expressas em ovários de rainhas em postura (Forêt e Maleszka, 2006). Em operárias, Cardoen *et al.*, 2011 observou uma maior expressão destes transcritos em corpo inteiro de operárias reprodutivas quando comparado às estéreis. Em nossa validação experimental, por meio de RT-PCR, nós confirmamos uma maior expressão da *obp7* em ovários ativados de operárias e em ovários de rainhas em postura (figura 21).

Os resultados aqui descritos não representam um caso isolado, há outra família genica que codifica proteínas transportadores de odor, cuja expressão não é exclusiva de

tecidos olfativos, esta classe proteínas é denominada *Small Chemosensory Proteins* (CSPs). Diferentemente ao observado para as OBPs, o número de CSPs observados em *D. melanogaster* e *A. gambiae* é semelhante ao identificado para *A. mellifera* (6 CSPs), embora em *T. castaneum*, tenham sido identificadas em torno de 20 CSPs (Forêt *et al.*, 2007). Em relação ao padrão de expressão os autores observaram que apenas a *csp4* (GB52326) possui expressão restrita aos tecidos olfativos e em nossas bibliotecas os valores de RPKM encontrados eram sempre inferiores a 1. Quanto aos *csp2* (GB55547) e *csp3* (GB53324) Forêt *et al.*, 2007, observaram uma expressão ampla nos indivíduos adultos, pupas e larvas, já em nossas bibliotecas observamos uma maior expressão nos ovários inativos de operárias e nos ovários de rainhas virgens. A *csp1* (GB43823), foi observada no corpo todo das abelhas adultas (Forêt *et al.*, 2006) e em nossas bibliotecas as encontramos com RPKM sempre inferior a 4. Já as *cps5* (GB54969) e *cps6* (GB54970) foram encontradas como altamente expressas em ovários de rainhas em postura e embriões, com um diferencial para a *cps5*, enquanto esta apresenta uma expressão restrita aos ovários de rainhas e embriões, a *cps6* persiste durante os estágios larvais e pupais (Forêt *et al.*, 2007). Ao confrontar este resultado aos dados de nosso sequenciamento observamos também uma alta expressão nos ovários de rainhas em postura e em sintonia também verificamos uma expressão preferencial nos ovários ativados de operárias, os valores de $\log_2 \text{fold change}$ calculados para o pareamento CRxE2 foram: *cps5* ≥ 5 em operárias e ≥ 6 em rainhas; *cps6* ≥ 3 em operárias e $\geq 2,9$ em rainhas.

4.2.5.5 Yellow-g e Yellow-x2

Dez genes *yellow* já foram identificados em *A. mellifera*. Em insetos, foram identificados sete genes *yellow* para o bicho-da-seda *Bombyx mori* (Xia *et al.* 2006), 14 para o besouro *T. castaneum* (Arakane *et al.* 2010), 25 para as vespas do gênero *Nasonia* (Werren *et al.* 2010), 13 para *D. melanogaster* (Maleszka e Kucharski 2000; Drapeau, 2001,). Em abelhas a família das proteínas YELLOW compartilha homologia com a família das MRJP (*major royal jelly proteins*) - exclusiva dos Hymenoptera, a MRJP 1 se destaca, nas fases iniciais do desenvolvimento das castas (Schmitzova *et al.* 1998; Albert e Klaidiny 2004; Drapeau, *et al.* 2006, Kamakura, 2011). Há fortes indícios de que os genes que codificam as MRJPs tenham evoluído recentemente via duplicação múltipla a partir da proteína yellow-e3 (Drapeau *et al.*, 2006, Buttstedt *et al.* 2013).

A primeira função sugerida a uma proteína YELLOW referia-se à síntese do pigmento da cutícula dos adultos e das peças bucais larvais de *D. melanogaster* (Brehme 1941; Biessmann 1985). Porém hoje já se sabe que estas proteínas possuem diferentes padrões de expressão sendo consideradas proteínas multifuncionais (Martin e Ollo 1996; De Gregorio *et al.* 2001; Diagana *et al.* 2002; Stanyon *et al.* 2004). As funções já descritas estão relacionadas ao comportamento (Dow 1976; Wilson *et al.* 1976; Burnet and Wilson 1980; Drapeau *et al.* 2003; Drapeau *et al.* 2006; Prud'homme *et al.* 2006), pigmentação e melanização da cutícula (Han *et al.*, 2002; Wittkopp *et al.*, 2002a; Wittkopp *et al.* 2002b; Prud'homme *et al.* 2006), sexo-específicos e maturação reprodutiva (Drapeau, Albert, *et al.* 2006; Xia *et al.* 2006), e podem também desempenhar um papel na diferenciação de castas e sinalização neuronal em abelhas *A. mellifera* (Kucharski *et al.* 1998; Drapeau *et al.* 2006; Kamakura, 2011)

Em especial para a proteína YELLOW-G (GB55171), foi descrito que possui expressão específica em ovários de rainhas e em embriões muito jovens, fato este

consistente com o padrão de expressão materno (Drapeau *et al.*, 2006). Mais tarde foi demonstrado que o gene *yellow-g* também se mostrava altamente expresso em corpo inteiro de operárias reprodutivas (Cardoen *et al.*, 2011). Nossa pergunta em relação a este gene era se os ovários ativados de operárias seriam responsáveis pela alta expressão observada em preparações usando o corpo todo de operárias reprodutivas. Nossos dados permitem concluir que a alta expressão deste gene em ovários ativados de operárias responde pelo seu papel no processo reprodutivo e na fertilidade de abelhas.

Em *D. melanogaster*, a proteína *YELLOW-G* foi associada a funções relacionadas à ovogênese e formação do ovo (Claycomb *et al.* 2004). Os autores demonstraram, por meio de hibridação *in situ*, uma expressão preferencial de *yellow-g* nas células foliculares das câmaras ovocíticas e analisaram o efeito da ausência da proteína codificada por este gene em fêmeas mutantes. Como resultado, a oogênese prosseguiu normalmente nas fêmeas mutantes, porém, os ovos produzidos eram inviáveis em consequência de problemas na integridade da membrana vitelínica. Foi sugerido, que *yellow-g* é necessário para a formação de uma membrana vitelínica apropriada e essencial para a formação adequada (rígida) do envoltório do ovo (Claycomb *et al.* 2004). Quando comparamos *YELLOW-G* em *A. mellifera* e *D. melanogaster* verificamos uma grande homologia na estrutura da proteína, em função desta característica, é provável que em abelhas desempenhe o mesmo papel observado em *D. Melanogaster*. Devido à impossibilidade de produzir fêmeas mutantes em abelhas, o silenciamento gênico seria a melhor metodologia para tentar reproduzir o efeito da ausência desta proteína no desenvolvimento do ovário e do envoltório do embrião.

Por meio de análises filogenéticas foi observado que para a maioria das proteínas *YELLOW* de *D. melanogaster* existe ortólogos em abelhas (Drapeau *et al.*,

2006). Este é um dos motivos pelo qual os ortólogos representantes das abelhas receberam os mesmos nomes que seus homólogos em *D. melanogaster* (ver Drapeau, 2001). No entanto, duas proteínas *YELLOW*, em *A. mellifera*, foram caracterizadas como “orfãs” por não possuírem homólogos óbvios em *D. melanogaster* (Drapeau *et al.*, 2006), elas foram denominados *YELLOW-X1* E *YELLOW-X2* (GB55729). Para o gênero *Tribolium*, foram identificados um clado composto por 5 genes *yellow* que também não possui identidade com os genes *yellow* identificados em *D. melanogaster* (Arakane *et al.*, 2010). Estes cinco genes de *Tribolium* foram relacionados aos *yellow-x* do gênero *Apis* e segundo Ferguson *et al.*, 2011, os gêneros, *Bombyx* e o *Heliconius* também apresentam em seu genoma os genes denominados *yellow-x*. Arakane *et al.*, 2010 analisou a expressão dos 14 genes *yellow* de *T. castaneum* em oito estágios de desenvolvimento (desde pupa farato até adulto recém emergido), a maioria dos genes *yellow* apresentaram padrões de expressão únicos, sugerindo funções fisiológicas distintas. Os resultados obtidos por Arakane *et al.*, 2010 foram associados com a pigmentação e melanização da cutícula.

Em nosso estudo foi analisada a expressão do gene *yellow-x2* nos ovários funcionais e não funcionais rainhas e operárias. Este gene foi selecionado para validação experimental pela intrigante característica de ter uma expressão antagônica a aquela observada para o gene *yellow-g*. Como as análises dos dados do sequenciamento indicavam, foi verificado experimentalmente (RT-PCR) que este gene encontra-se mais expresso nas condições não funcionais dos ovários quando comparados com as condições funcionais dos ovários. É ainda prematuro inferir uma função ou relacionar este gene a alguma via metabólica, porém este resultado acrescenta um novo fator, que deverá ser considerado em investigações futuras sobre a manutenção do estado não funcional dos ovários de *A. mellifera*.

4.2.5.6 GB44975

O GB44975 sintetiza uma proteína não caracterizada LOC100577041. Nós selecionamos para a validação experimental por RT-PCR por ter sido encontrado em nossas análises computacionais como altamente expresso nos ovários de operárias com ovários ativados e em rainhas em postura (tabela VII). Utilizamos da ferramenta BlastP, disponível no site do *National Center for Biotechnology Information* (NCBI), para verificar a existência de ortólogos para proteína hipotética LOC100577041 (GB44975), porém, como resultados obtivemos que esta proteína seria exclusiva de *A. mellifera* (nenhum outro organismo apresentou uma sequência com alguma identidade a esta proteína). Para reforçar nossas análises em relação à busca por ortólogos, utilizamos a ferramenta computacional BlastP disponível no banco de dados *Universal Protein Resource* (UniProt), encontramos uma identidade de 36% com a proteína de *D. erecta* com número de acesso GG17530. Por meio desta análise fomos capazes de categorizar funcionalmente nossa proteína de acordo com o Gene Ontology (<http://www.geneontology.org/>). No que diz respeito à função molecular como resultado verificamos esta proteína possui função molecular desconhecida (acesso GO:0003674), os termos próximos, relacionadas à este, que se destacaram foram: ligação (GO:0005488) e atividade catalítica (GO:0003824). Em relação ao componente celular, averiguamos que esta proteína provavelmente se encontra dentro do núcleo, para confirmar este resultado utilizamos a ferramenta computacional *Predicting nuclear localization of proteins* (NucPred) e corroboramos o resultado, a proteína representada pelo GB44975 apresentou 94% de chance de estar dentro do núcleo. No que tange aos processos biológicos obtivemos que esta proteína deve estar envolvida com a atividade de transporte (GO:0006913) mais especificamente relacionada com o movimento dirigido de moléculas entre o núcleo e o citoplasma.

5.2.5.7 GB53748

Nós selecionamos o gene representado pelo GB53748, que não possui uma proteína caracterizada, por apresentar uma alta expressão nos ovários não funcionais quando comparados aos ovários funcionais. Outro aspecto que chama muito a atenção é o fato deste gene se mostrar, aparentemente, exclusivo da ordem himenóptera. Ao realizar buscas por ortólogos utilizando a ferramenta BlastP disponível no *National Center for Biotechnology Information* (NCBI), os organismos que apresentaram identidade com sequência predita, referente ao GB53748, foram: as abelhas *Apis florea* (80% de identidade) *Bombus terrestris* (65% de identidade), *Bombus impatiens* (66% de identidade) e *Megachile rotundata* (52% de identidade) e a formiga *Harpegnathos saltator* (46% de identidade), para todos estes organismos a sequência se referia ou a proteínas hipotéticas ou a proteínas não caracterizadas.

Apenas um estudo científico foi encontrado na literatura que se refere a este gene (GB53748). Trata-se de uma investigação que consiste na identificação de proteínas presentes na espermateca de rainhas acasaladas e de rainhas virgens, para tanto, os autores coletaram o líquido presente no interior desta estrutura e realizaram uma análise de proteômica, por meio de métodos convencionais, (LC)-MS/MS (Baer *et al.*, 2009). Como resultado, dentre as 122 proteínas identificadas como diferencialmente expressas, a proteína referente ao GB53743 foi encontrada apenas nas amostras de rainhas virgens. Os autores ainda comparam seus achados a duas outras análises semelhantes do fluido seminal e de esperma, mas o GB53748 não foi encontrado em nenhum deles. Seria possível relacionar este estudo ao nosso, uma vez que nossas amostras de rainhas virgens o GB53748 apresenta-se com uma expressão muito superior à expressão encontrada nas rainhas em postura ($\log_2 \text{fold change} \geq 5$). Porém, durante a dissecação para coleta de nossas amostras, tivemos a preocupação de em retirar

espermateca, tanto dos ovários de abelhas virgens, como dos ovários das abelhas em postura, para que expressão da espermateca e seu conteúdo não interferissem em nossos resultados. Além disso, as operárias com ovários não funcionais também apresentaram uma alta expressão deste mRNA, ressaltando que a espermateca nesta casta é apenas vestigial (citar o Snodgrass, 1956). Diante destes fatos podemos concluir que este gene de alguma forma contribui com a esterilidade das abelhas operárias e com o estado não funcional dos ovários das rainhas.

4.2.5.8 Vg

A vitelogenina (Vg) é umas das proteínas mais estudadas em insetos pelo fato de ser a precursora da vitelina, principal proteína do vitelo, suprimento necessário para o desenvolvimento do embrião até a eclosão do ovo (ver Valle, 1993). Nas abelhas, além de desempenhar esta função em rainhas e operárias reprodutoras, a Vg também é encontrada em operárias não reprodutoras (Engels, 1974; Engels *et al.*, 1990) e em zangões adultos (Trenczek e Engels, 1986). Estudos mostraram que esta proteína pode estar envolvida no transporte de açúcares, lipídeos, fosfatos, vitaminas e hormônios (Chen *et al.*, 1997; Sappington e Raikhel, 1998), além de aumentar a longevidade e ter participação na reação imunológica (Seehuus *et al.*, 2006), sendo assim caracterizada com uma proteína multifuncional. O transcrito do gene codificador da Vg, em abelhas *A. mellifera*, possui 5440 nucleotídeos, a proteína contém 1770 aminoácidos com massa molecular de 201 kDa (Piulachs *et al.*, 2003). A Vg, por meio de ligação covalente, se liga a carboidratos (especialmente manose), fosfolipídios e diacilglicerol (Engels *et al.*, 1990; Wheeler e Kawooya, 1990), sendo assim denominada fosfolipoglicoproteína. A principal ocorrência de Vg, nos insetos, é sua síntese no corpo gorduroso, posterior

liberação na hemolinfa para, finalmente, ser incorporada pelos ovócitos em desenvolvimento (Wyatt, 1999; Raikel *et al.*, 2004). Foi verificado por Engels *et al.*, 1990 e Hartfelder e Engels, 1998 que logo após a emergência das rainhas, a Vg começa a acumular na hemolinfa, ao passo que com 3 dias de vida adulta a Vg já representa 70% das proteínas presentes e é mantida elevada por toda a vida da rainha (representando 50% das proteínas presentes na hemolinfa). Todavia, experimentos comprovaram que em algumas espécies de insetos, dentre eles as abelhas, outros órgãos também podem expressar *vg*, como por exemplo, nos ovários de rainhas virgens (Guidugli *et al.*, 2005) e no cérebro de operárias (Seehuus *et al.* 2007, Münch e Amdam 2010 e Nunes *et al.* 2013). Apesar de Guidugli *et al.*, 2005 demonstrarem por meio de *Northern Blot* transcritos presentes nos ovários de rainhas virgens Corona *et al.*, 2007 rebateu o achado, dizendo que a expressão de *vg* só era possível nas células do corpo gorduroso e que os achados de Guidugli *et al.*, 2005 eram consequência de uma contaminação (provavelmente as células do corpo gorduroso que envolviam os ovários). Esta contestação não durou muito tempo uma vez que logo no ano seguinte Seehuus *et al.* 2007 verificou a expressão de *vg* no cérebro e não apenas nas células do corpo gorduroso que circundavam o cérebro. Até o momento, em relação à expressão de *vg* nos ovários de abelhas nada mais foi relatado, porém, nossos resultados de RNAseq mostraram que a *vg* é expressa nos ovários e possui uma maior expressão nas rainhas virgens e nos ovários inativos de operárias (principalmente na amostra SR). Nossa validação experimental nos permitiu corroborar o achado de Guidugli *et al.*, 2005 além de trazer uma prova irrefutável a respeito da veracidade da expressão de *vg* nos ovários que se baseia no fato de rainhas em postura possuírem um nível baixíssimo de expressão (quase nulo) de *vg* e rainhas virgens um nível altíssimo de expressão de *vg*, sendo que ambos ovários são circundados de maneira semelhante pelas células do corpo

gorduroso. Nossas análises de RT-PCR não mostram como está expressa a *vg* nas diferentes amostras de ovários inativos e de ovários ativados, por hora confirmamos somente que operárias com ovários inativos possuem uma maior expressão de *vg*

4.3 miRNAs expressos em ovários inativos e ativados de *A. mellifera*

4.3.1 Análise global (RNAseq) dos RNAs curtos

Duas amostras de mRNAs foram utilizadas para o sequenciamento dos RNAs curtos que correspondem: A) ovários ativados (OA) sem classificação do estágio de ativação e B) ovários inativos de operárias nutrízes (ONUR) com 4 dias de vida adulta. Os sequenciamentos foram realizados em momentos distintos e em diferentes *facilities*. A amostra OA foi a primeira a ser sequenciada (*facility DNAvision*). E a amostra ONUR foi posteriormente sequenciada pela *facility High Throughput Sequencing Facility*.

Foram obtidas 48.584.539 leituras para OA e 35243253 para ONUR. O comprimento dos *contigs* variou no intervalo de 10 a 40 nucleotídeos (representando as diversas classes de pequenos RNAs). A tabela VIII, mostra: A) o número total de sequências geradas pelo RNAseq; B) o número de leituras resultantes após as etapas de *trimagem* dos adaptadores e *trimagem* de qualidade (processamento); C) o número de sequências mapeadas contra os miRNA precursores (pre-miRNA) presentes no genoma de *A. mellifera* e D) o número de sequências unicamente mapeadas também contra os pre-miRNAs no miRBase (19° versão). No final destas análises, observou-se que a maioria das sequências possui entre 19 e 24 nucleotídeos, tamanho esperado para os miRNAs (Chu e Rana, 2007).

Tabela VIII -. Número de sequências obtidas, por meio de experimentos de RNAseq em ovários inativos (ONUR) e ativados (OA) de operárias *A. mellifera*, e após etapas de análise dos dados.

Análise primária das sequencias	ONUR	OA
Nº total de sequências geradas	35243253	88942714
Nº de sequências obtidas após processamento	34183414	73890492
Nº de sequências mapeadas contra os pre-miRNAs	3580391	1412959
Nº mapeamentos não ambíguos contra os pre-miRNAs	3230076	1412959

Além dos miRNAs, outras classes de pequenos RNAs também podem ter sido sequenciadas (tabela IX). Os pequenos RNAs que possuem tamanhos semelhantes aos dos miRNAs são: A) siRNAs pequenos RNAs (19-21 nucleotídeos) cuja maquinaria utilizada é a mesma dos miRNAs e a atuação é o silenciamento do gene alvo (Chun e Rana, 2007), porém, sua origem, na grande maioria dos casos, é exógena o que não é o caso das sequencias obtidas (que são de origem endógena); B) tasiRNAs (21-22 nucleotídeos) são de origem endógena, seus transcritos alvos não são originados dos mesmos genes dos quais os tasiRNAs derivam (Peragine *et al.*, 2004; Vazquez *et al.*, 2004), todavia, não existem tasiRNAs descritos em *A. mellifera* e C) rasiRNAs (24-27 nucleotídeos) suas origens ainda não estão claras, é um tipo de piRNA envolvido no silenciamento de retrotransposons e foram encontrados nas células germinativas de *D. melanogaster* (Shpiz et al., 2009), em *A. mellifera* também ainda não foram descritos.

Tabela IX - Classes de pequenos RNAs, respectivos tamanhos e funções.

Classe	Tamanho (nt)	Função
miRNAs	19-25	Repressão da tradução
siRNAs	19-21	Clivagem do mRNA alvo
tasiRNAs	21-22	Clivagem do mRNA
scnRNAs	~28	Eliminação de DNA
rasiRNAs	24-27	Controle do transposon- silenciamento transcricional
piRNAs	27-31	Controle do transposon em células germinativas

Abreviaturas: miRNAs, *microRNA*; siRNA, *small interfering RNA*; tasiRNA, *trans-acting siRNA*; scn RNA, *small-scan RNA*; rasiRNA, *repeat-associated siRNA*; piRNA, *piwi-interacting RNA*. Dados Adaptados de (Chu e Rana, 2007).

Dos os 221 miRNAs maduros de *A. mellifera* depositados no miRBase (19ª versão), 158 se expressam nos ovários. Destes, dez são exclusivos de ovários ativados: miR-3735, miR- 2765, miR-3792, miR3796, miR-3716b, miR3778, miR-3801, miR3764, miR-3741 e miR-6050. E dezenove são exclusivos de ovários inativos de abelha nutriz, no entanto, onze apresentam níveis de expressão muito baixos e possivelmente não estão presentes em quantidades suficientes para provocar efeitos biológicos significativos. Os demais oito são: miR-193, miR-6005, miR-981, miR-1000, miR-3730, miR-6058, miR-980 e miR-3782. Já foi descrito que todos os miRNAs com nomenclatura acima de 2000, parecem ser exclusivos de abelhas *A. mellifera* (Chen *et al.*, 2010) e por coincidência , a maioria dos miRNAs aqui citados se enquadram nesta categoria.

Na figura 23 são mostrados os 15 miRNAs mais expressos na biblioteca de ovários inativos. E na figura 24 são mostrados 15 miRNAs mais expressos na biblioteca de ovários ativados. Por meio da figura 23 podemos constatar que os três miRNAs mais

expressos nos ONUR são os miR-276, miR-182 e miR-306. E de maneira interessante, por meio da figura 24 podemos constatar que o miR-306 e o miR-184 também estão entre os três miRNAs mais expressos em OA, sendo que o miR-276 é o quinto miRNA mais expresso na amostra OA. Além disso, podemos perceber que nove miRNAs (miR-276, miR-184, miR-306, miR-2, miR-14, miR-263a, miR-317 e miR-279a) são comuns as duas análises (figura 23 e figura 24). Este resultado nos leva a crer que estes miRNAs desempenham funções nos ovários que não restritas ao estado fisiológico dos mesmos.

Figura 23: Distribuição dos 15 miRNAs mais frequentes na biblioteca de ovários inativos (RNAseq). O eixo Y do gráfico indica os valores obtidos para os miRNAs (plotados) de acordo com a contagem normalizada das sequências.

Figura 24: Distribuição dos 15 miRNAs mais frequentes na biblioteca de ovários ativados (RNAseq). O eixo Y do gráfico indica os valores obtidos para os miRNAs (plotados) de acordo com a contagem normalizada das sequências.

Um limiar de corte de $\log_2$ ratio entre 1,5 e -1,5 foi estabelecido para qualificar os miRNAs como diferencialmente expressos no pareamento OAxONUR. Dentro deste critério verificamos que 63 miRNAs não apresentam diferença de expressão e 95 são diferencialmente modulados quando comparamos OA com ONUR. Destes, 68 encontram-se mais expressos em ONUR (tabela X) e 27 em OA (tabela XI).

Em outros organismos, como: *B. germanica* (Cristino *et al.*, 2011) e *D. melanogaster* (Berezikov *et al.*, 2011), a técnica de RNAseq já foi utilizada para identificar miRNAs em ovários. Porém, o contexto de estudo foi outro que não a comparação entre o estado fisiológico funcional *versus* não funcional. Poucos dados sobre miRNA estão disponíveis em *A. mellifera*. A maior contribuição foi a de Chen *et al.* (2010). Estes autores sequenciaram as 3 castas de *A. mellifera* em todas as fases do desenvolvimento e obtiveram um total de 36.796.459 sequências das quais 5,24%

apresentaram identidade com os miRNAs conhecidos até aquele momento (16^o versão do miRBase). Isto permitiu a predição de 267 novos miRNAs dos quais 108 foram depositados no miRBase. Além deste, até o momento, apenas outros dois estudos foram realizados utilizando desta metodologia (RNAseq) para a investigação de miRNAs em *A. mellifera* que são: Liu *et al.*, 2012 (em amostras de nutridoras e forrageiras) e Zondag *et al.*, 2012 (em amostras de embriões).

Para validação experimental levamos em conta os miRNAs identificados e analisados em nossa biblioteca e miRNAs que apresentasse, na literatura, funções que pudéssemos associar ao processo reprodutivo em insetos. Portanto selecionamos os miRNAs: miR-9a, miR-92b, miR-100, miR-184, miR-306, miR-317, miR-1, miR-276, 263a, miR-31a, miR-12, miR-14, miR-8, miR-let-7, miR-125, miR-2, miR-71, miR-13b, miR-13a, miR-3720 e miR-11, para avaliar a expressão em amostras de ovários ativados e ovários inativos e em ovários de rainhas em postura e virgens, por RT-PCR.

Tabela X – miRNAs altamente expressos em ovários inativos de operárias nutriz (ONUR) quando comparados com ovários ativados (OA)

miRNAs com $\log_2 \text{fold change} \leq -2,81$		miRNAs com $\log_2 \text{fold change} \geq -2,81$ e $\leq -1,5$	
miRNAs	$\log_2 \text{fold change}$	miRNAs	$\log_2 \text{fold change}$
ame-bantam-3p	-7,01	ame-miR-2788-3p	-2,74
ame-miR-3785-3p	-6,54	ame-miR-276-3p	-2,74
ame-miR-31a-5p	-6,29	ame-miR-252b-5p	-2,68
ame-miR-277-3p	-6,22	ame-miR-278-3p	-2,65
ame-miR-10-5p	-5,85	ame-miR-87-3p	-2,6
ame-miR-3743-3p	-5,16	ame-miR-375-3p	-2,59
ame-miR-29b-3p	-4,99	ame-miR-34-5p	-2,57
ame-miR-133-3p	-4,79	ame-miR-6001-3p	-2,54
ame-miR-3715-5p	-4,74	ame-miR-13b-3p	-2,48
ame-miR-6043-3p	-4,69	ame-miR-137-3p	-2,45
ame-miR-12-5p	-4,48	ame-miR-279b-3p	-2,43
ame-miR-193-3p	-4,33	ame-miR-13a-3p	-2,41
ame-miR-2765-5p	-4,26	ame-miR-263a-5p	-2,38
ame-miR-3719-3p	-4,1	ame-miR-3730-5p	-2,27
ame-miR-33-5p	-3,89	ame-miR-iab-4-5p	-2,26
ame-miR-3049-3p	-3,81	ame-miR-6051-3p	-2,26
ame-let-7-5p	-3,76	ame-miR-750-3p	-2,21
ame-miR-6005-3p	-3,74	ame-miR-8-3p	-2,05
ame-miR-6000b-3p	-3,71	ame-miR-125-5p	-1,99
ame-miR-281-3p	-3,6	ame-miR-6058-5p	-1,97
ame-miR-3727-3p	-3,59	ame-miR-6047a-3p	-1,95
ame-miR-316-5p	-3,58	ame-miR-3761-5p	-1,83
ame-miR-100-5p	-3,32	ame-miR-993-3p	-1,79
ame-miR-92b-3p	-3,22	ame-miR-9a-5p	-1,75
ame-miR-275-3p	-3,12	ame-miR-92a-3p	-1,72
ame-miR-279c-3p	-3,08	ame-miR-6038-5p	-1,68
ame-miR-283-5p	-3,02	ame-miR-989-3p	-1,66
ame-miR-279a-3p	-2,93	ame-miR-980-3p	-1,61
ame-miR-3718a-3p	-2,92	ame-miR-3782-5p	-1,61
ame-miR-981-3p	-2,9	ame-miR-305-5p	-1,6
ame-miR-6037-3p	-2,84	ame-miR-190-5p	-1,59
ame-miR-1000-5p	-2,84	ame-miR-6041-3p	-1,54
ame-miR-3791-3p	-2,83	ame-miR-315-5p	-1,54
ame-miR-996-3p	-2,81		
ame-miR-927a-5p	-2,81		

Tabela XI – miRNAs altamente expressos em ovários ativados de operárias (OA) quando comparados com ovários inativos (ONUR)

#miRNA	$\log_2 \text{fold change} \geq 1,5$
ame-miR-71-5p	5,16
ame-miR-1-3p	4,2
ame-miR-317-3p	3,87
ame-miR-92c-5p	3,76
ame-miR-6045-5p	3,6
ame-miR-124-3p	3,18
ame-miR-3756-5p	3,13
ame-miR-6040-3p	2,98
ame-miR-6044-5p	2,93
ame-miR-3747a-5p	2,89
ame-miR-6052-5p	2,74
ame-miR-6064-5p	2,64
ame-miR-6049-5p	2,64
ame-miR-3720-5p	2,55
ame-miR-3759-3p	2,51
ame-miR-6063-3p	2,47
ame-miR-306-5p	2,13
ame-miR-932-5p	2,07
ame-miR-318-3p	1,94
ame-miR-6000a-3p	1,93
ame-miR-6050-5p	1,9
ame-miR-3728-5p	1,74
ame-miR-2944-3p	1,69
ame-miR-965-5p	1,66
ame-miR-3741-3p	1,61
ame-miR-3786-5p	1,54
ame-miR-6039-5p	1,5

4.3.2 - Validação experimental dos miRNAs

A expressão dos miRNAs selecionados (miR-9a, miR-92b, miR-100, miR-184, miR-306, miR-317, miR-1, miR-276, 263a, miR-31a, miR-12, miR-14, miR-8, miR-let-7, miR-125, miR-2, miR-71, miR-13b, miR-13^a, miR-3720 e miR-11) foi avaliada por

meio da técnica de PCR em tempo real (RT-PCR). As mostras biológicas utilizadas foram: A) ovários inativos de operárias (NUR), B) ovários ativados de operárias (OA), C) ovários de rainhas em postura (RA) e B) ovários de rainhas virgens (RV). Vale ressaltar que: A) as amostras de rainhas foram as mesmas utilizadas na validação dos mRNAs e B) os RNAs extraídos para a síntese do cDNAs de operárias, utilizados nestes experimentos de RT-PCR, apresentam as mesmas características, mas são totalmente independentes daqueles utilizados para o sequenciamento (RNAseq).

Decidimos analisar e comparar a expressão dos ovários de rainhas em postura e virgens com o intuito de reforçar nossas validações, uma vez que verificamos, nas validações e nas análises de RNAseq (mRNAs), que há um perfil de expressão muito similar entre as duas castas conforme o estado fisiológico dos ovários. Embora, os ovários de rainhas e operárias possuam tamanhos extremamente diferentes e algumas características morfológicas do aparelho reprodutor sejam completamente distintos, os resultados obtidos até o momento indicam que os mecanismos que levam à ovogênese e o processo de vitelogênese, podem ser bastante semelhantes entre estas castas.

A figura 25 representa os 4 miRNAs encontrados em nossas validações como significativamente mais expressos em ovários ativados e rainhas em postura. Dois deles, o miR-306 e miR-317 validaram nossos resultados obtidos a partir do RNAseq, confirmando uma maior expressão nos ovários funcionais quando comparados aos ovários não funcionais. Os outros dois, miR-9a e miR-92b, não validaram os resultados obtidos no RNAseq, pois apresentaram uma maior expressão nos ovários funcionais enquanto a análise dos dados de RNAseq indicavam uma maior expressão nos ovários não funcionais.

Figura 25: Quantificação relativa dos níveis de transcritos dos microRNAs. A) ame-miR-306, B) ame-miR-317, C) ame-miR-9a e D) ame-miR-92b. A validação foi realizada em amostras de operárias com ovários ativados e inativos e em amostras de rainhas com ovários em postura e virgens. As análises foram realizadas por PCR em tempo real, utilizando para normalização dos resultados a média dos valores de Ct dos miR-184, miR-2 e U5. Cada coluna representa a média de três amostras biológicas. O teste estatístico utilizado para comparar os diferentes grupos foi o “One Way Anova” por meio da ferramenta Jandel SigmaStat 3.2.

A figura 26 mostra 7 miRNAs encontrados em nossas validações como significativamente mais expressos em ovários inativos e rainhas virgens. A expressão dos miRNAs: miR-276, miR-263a, miR-31a, miR-12, miR-14 e miR-8, é concordante com os resultados obtidos por meio de RNAseq, pois, apresentaram uma expressão preferencial nos ovários não funcionais. O miR-14 não está incluído na tabela X por apresentar um $\log_2 fold change = -1,2$. Ele foi incluído em nossas validações por fazer parte dos 15 miRNAs mais expressos em ovários de abelhas em ambos os estados

fisiológicos (figuras 23 e 24). Na figura 26 o único miRNA que não corroborou os resultados de RNAseq, por meio da validação experimental, foi o miR-1. O que mais nos impressiona é o fato deste miRNA, inclusive, se apresentar como um dos miRNAs com maior expressão gênica digital quando comparada à biblioteca de OA com NUR ($\log_2 \text{fold change} = 4,2$).

A biogêneses de alguns miRNAs ocorre na mesma região genômica, quando isso ocorre, os miRNAs são agrupados em *clusters*. Nós analisamos a expressão dos miRNAs que compõem dois diferentes *clusters*: A) *cluster let-7*, composto pelos miRNAs: miR-let-7, miR-100 e miR-125 (figura 27) e B) *cluster miR-2*, composto pelos miRNAs: miR-2, miR-13a, miR-13b e miR-71 (figura 28). Os três miRNAs que compõem o *cluster let-7* estão presentes na tabela X, ou seja, nas análises de RNAseq foram verificados como mais expressos nos ovários não funcionais e a validação experimental validou este resultado tanto nas amostras de rainhas como nas amostras de operárias (figura 27). Os quatro miRNAs que compõem o *cluster 2* não apresentam um perfil de expressão similar (figura 28): A) o miR-2 nas análises de RNAseq e na validação experimental não apresentou diferenças quando comparadas as amostras de ovários funcionais e não funcionais, e inclusive serviu como gene de referência, junto com o miR-184 e U5, para normalizar os resultados de expressão dos demais miRNAs; B) o miR-13a é um dos miRNAs apresentados na tabela X, portanto, nas análises de RNAseq, possui uma maior expressão nos ovários inativos, porém, na validação experimental, este miRNA, não apresentou sequer uma tendência preferencial de expressão; C) o miR-13b também está presente na tabela X e sua expressão foi validada experimentalmente, ou seja, ovários inativos de operárias e ovários de rainhas virgens possuem uma expressão estatisticamente maior quando comparados aos ovários: ativados de operárias e em postura de rainhas; e D) o miR-71 faz parte dos miRNAs

como maior expressão gênica digital na amostras de ovários ativados (tabela XI) e este resultados foi validado experimentalmente quando comparadas as amostras de operárias com ovários ativados *versus* operárias com ovários inativo.

Os demais miRNAs, miR-3720, miE-182 e niR-11 não apresentam diferenças estatísticas em nossas validações experimentais (figura 29). O miR-3720, nas análises de RNAseq está mais expresso em ovários ativados (tabela XI). E os miR-184 e miR-11 foram validados pois também não apresentaram diferenças de expressão nas análises de RNAseq. O miR-11 foi selecionado para a validação experimental por estar entre os 15 genes mais expressos em ovários ativados (figura 24). E o miR-184 foi incluído em nossas validações por fazer parte dos 15 miRNAs mais expressos em ovários de abelhas em ambos os estados fisiológicos (figuras 23 e 24), mas devido sua expressão ser similar em todas as amostras testadas, serviu como gene de referência, junto com o miR-2 e U5, para normalizar os resultados de expressão dos demais miRNAs

Figura 26: Quantificação relativa dos níveis de transcritos, por meio de PCR em tempo real, dos microRNAs: ame-miR-276, ame-miR1, ame-miR-263a, ame-miR-31a, ame-miR-12 e ame-miR-14 e ame-miR-8. As amostras utilizadas foram: 1) operárias com ovários ativados e inativos e 2) rainhas com ovários em postura e virgens. Cada coluna representa a média de três amostras biológicas. O teste estatístico utilizado para

comparar os diferentes grupos foi o “One Way Anova” por meio da ferramenta Jandel SigmaStat 3.2.

Figura 27: Quantificação relativa dos níveis de transcritos, por meio de PCR em tempo real, dos microRNAs: A) ame-miR-let-7; B) ame-miR-100 e C) ame-miR-125. As amostras utilizadas foram: 1) operárias com ovários ativados e inativos e 2) rainhas com ovários em postura e virgens. Cada coluna representa a média de três amostras biológicas. O teste estatístico utilizado foi: “One Way Anova” por meio da ferramenta Jandel SigmaStat 3.2.

Figura 28: Quantificação relativa dos níveis de transcritos, por meio de PCR em tempo real, dos microRNAs: A) ame-miR-let-71; B) ame-miR-2; C) ame-miR-13b e D) ame-miR-13a. As amostras utilizadas foram: 1) operárias com ovários ativados e inativos e 2) rainhas com ovários em postura e virgens. Cada coluna representa a média de três amostras biológicas. O teste estatístico utilizado foi: “One Way Anova” por meio da ferramenta Jandel SigmaStat 3.2.

Figura 29: Quantificação relativa dos níveis de transcritos, por meio de PCR em tempo real, dos microRNAs: A) ame-miR- 3720; B) ame-miR-11 e C) ame-miR-184. As amostras utilizadas foram: 1) operárias com ovários ativados e inativos e 2) rainhas com ovários em postura e virgens. Cada coluna representa a média de três amostras biológicas. O teste estatístico utilizado foi: “One Way Anova” por meio da ferramenta Jandel SigmaStat 3.2.

Pela primeira vez, microRNAs específicos foram testados com o intuito de caracterizar os diferentes estados fisiológicos dos ovários de *A. mellifera*. Assim como os mRNAs, os miRNAs também demonstraram que a condição reprodutiva é modulada de maneira similar em ambas as castas. Estudos já demonstraram que os microRNAs são importantes reguladores pós transcricionais envolvidos no desenvolvimento, proliferação celular e desenvolvimento. Nossos resultados nos permitem afirmar que o perfil de expressão destes RNAs retratam a condição funcional e não funcional dos ovários de rainhas e operárias.

O miR-306 e miR-317 foram observados estatisticamente mais expressos nas amostras de ovários funcionais (figura 25), e na biblioteca de ovários ativados (RNAseq), sendo que o miRNA-306 é o miRNA que possui o maior número de sequencias, seguido pelo miRNA-317 (figura 24). Suas funções ainda não estão muito bem caracterizadas, em *B. mori* verificou-se que miR-306 juntamente com miR-317 estão altamente expressos em adultos (Jagadeeswaran *et al.*, 2010). O miR-317, encontra-se conservado em *T. castaneum*, *D. melanogaster* e também em mamíferos. A única função já descrita foi relacionada à resposta imune em *Glyptapanteles flavicoxi* (Gundersen-Rindal e Pedroni 2010). O miR-306 foi encontrado nas células germinativas dos ovários de *D. melanogaster*, além de regular genes específicos na diferenciação dos espermatócitos (Eun *et al.* 2013). Análises *in silico* indicaram apenas um alvo para o miR-306, trata-se do *alpha-tubulin- chain 1* (GB52107), cuja proteína (XP_3919363) é um componente dos microtúbulos. Avaliamos os níveis de transcritos deste gene em nossa biblioteca (RNAseq de mRNAs) e observamos uma maior expressão no ovários inativos de operárias e nos ovários de rainhas virgens (embora o $\log_2 \text{fold change}$ tenha sido inferior a 1,5), esta expressão antagônica aliada à predição *in silico*, não garante, mas é um forte indicativo de que esta relação miRNA-alvo deva ocorrer nos ovários de abelhas *A. mellifera*.

Os miRNAs que não validaram nossos experimentos de RNAseq foram os: miR-9a e o miRNA-92b. Os resultados de RT-PCR apontam uma maior expressão nos ovários funcionais quando comparado aos não funcionais. O miR-9a é altamente conservado e encontrado em animais, de insetos a humanos. Alguns processos biológicos já foram associado a este miRNA como por exemplo: proliferação e morte celular no desenvolvimento, resistência ao estresse e metabolismo de lipídios (Ambros 2003). Weaver *et al.*, (2007) observaram níveis maiores de expressão em abdômen e

tórax de operárias adultas quando comparadas às rainhas, entretanto, em corpo inteiro de pupas, o oposto foi verificado. O autor sugere que este miRNA possivelmente desempenhe papéis relacionados à diferenciação de castas. A principal função descrita para o miRNA-92b está no processo de proliferação celular (Liu *et al.*, 2010; Sengupta *et al.*, 2009). Em mamíferos, foi demonstrado nas células tronco, seu envolvimento no controle do ciclo celular durante a passagem da fase G1 para a fase S (Sengupta *et al.*, 2009).

O miR-71 foi encontrados em ambas análise (RNAseq e RT-PCR), como mais expresso nos ovários ativados de operárias. A figura 24 mostra que este miRNA é sexto mais expresso na biblioteca de ovários ativados. Ele possui a *vitelogenina* como alvo predito, curiosamente, validamos a *vitelogenina* neste trabalho e vimos que a expressão de seu transcrito está elevada nos ovários inativos de operárias. Este resultado sugere que o miR-71 deve estar alvejando os transcritos da *vitelogenina* nos ovários ativados, motivo pela qual a expressão se mostra bastante inferior quando comparada aos níveis de transcritos dos ovários inativos.

O mir-184 é bastante conservado desde insetos até vertebrados, e em nossas bibliotecas, ele está entre os três miRNA com maior expressão digital (figura 24 e 25), Iovino *et al.*, 2009 verificaram que a perda deste miRNA traz severos danos durante a ovogenese e embriogênese. Este gene, como sugere nossos resultados, deve desempenhar um papel importante na manutenção dos ovários, mas não é fator limitante no processo de ativação.

Os miRNAs altamente expressos nos ovários inativos, podem estar envolvidos com a manutenção do estado infértil, possivelmente alvejando importantes mRNAs que desencadeiem o processo de ativação. Para alguns, as funções já fora descritas, como

por exemplo: A) os membros do *cluster* let-7 (figura 27), estão relacionados ao desenvolvimento, à resposta a ecdisona e à regulação de *broad complex* durante a metamorfose de *B. germanica* (Rubio *et al.* 2012), em *A. mellifera* estão relacionados ao comportamento, apresentando uma maior expressão nas operárias nutrízes quando comparadas às forrageiras (Behura e Whitfield, 2010) e em *B. mori* foram observados transcritos nos ovários (Liu *et al.* 2007); B) o miR-12 possui uma importante função nos ovários de *D. melanogaster* (Osei-Amo *et al.*, 2012), ele se liga ao transcrito do gene que regula a formação do córion MCM6, sua função é crucial para uma síntese rápida e uma deposição eficiente deste envoltório durante a formação da membrana externa do ovo; C) os membros da família do miR-263 foram associados aos processos de regulação de apoptose (Nguyen e Frasch 2006; Liu *et al.* 2010); D) miR-14 foi encontrado como sendo um inibidor da morte celular programada (Xu *et al.*, 2003); D) os membros do *cluster* do miR-2 foram visto como essenciais para a supressão da apoptose durante a embriogênese (Leaman *et al* 2005); E) o miR-1 é muito conservado em metazoários, e regularmente expresso durante a embriogênese de *B. mori* e de *D. melanogaster* (Nguyen e Frasch, 2006; Liu *et al.*, 2010) e F) miR-8 está envolvido no crescimento corporal em *D. melanogaster* (Nguyen e Frasch, 2006) e *B. mori* (Liu *et al.*, 2010).

Em humanos, mais de 700 miRNAs já foram identificados (Griffiths *et al.*, 2008), cada um possui o potencial de regular a expressão de milhares de mRNAs (Friedman *et al.*, 2009). A maioria dos miRNAs apresentam padrões de expressão de acordo com o contexto celular, tecidos e/ou órgãos, fases do desenvolvimento (Mattick *et al.*, 2005), o que torna este estudo além de extremamente instigante, uma das mais importantes abordagens para descrição de importantes patologias. A análise de um órgão, em condições fisiológicas distintas, e de certo modo antagônicas, possibilitou

visualizar as correlações entre miRNAs e mRNAs na forma de redes integradas e inferir sobre a regulação deste processo. Este resultado, por certo, traz novas perspectivas para a compreensão da regulação do estado reprodutivo e não reprodutivo em nosso modelo, *A. mellifera*.

4.4 Rede integrativa – miRNA-mRNA

Para a construção da rede integrativa utilizamos a lista que continham os 245 mRNAs encontrados como diferencialmente expressos ($\log_2 \text{fold change} \geq 5$) nas análises entre ovários funcionais *versus* não funcionais de operárias (E2xCR, E2xSR, E1xCE e E1xSR). Inicialmente selecionamos os que se encontravam mais expressos nos ovários ativados (75 no total), destes, fizemos uma nova filtragem e consideramos apenas aqueles que apresentavam $\log_2 \text{fold change} \geq 5$ nos pareamentos E2xCR e E1xCR, o que resultou em 57 genes. A lista dos miRNAs selecionados para compor a rede foram os 68 mais expressos nos ovários inativos. De posse das listas, utilizamos dois métodos: RNAhybrid (Rehmsmeier *et al.*, 2004) e PITA (Kertesz *et al.*, 2007) para realizar as predições das relações miRNAs-alvo e o programa Cytoscape (Shannon *et al.*, 2003) para desenhar as redes de interação. Nosso objetivo ao construir esta rede foi a obtenção de uma imagem do processo de repressão exercido pelos miRNAs que mantém os ovários das operárias inativos.

Ao analisar a rede (figura 30), percebemos que dois mRNAs validados: *fpfs5* (GB48899) e *obp7* (GB53370) não estavam inserido na rede, isto aconteceu devido ao fato destes mRNAs não apresentarem sítios de ligação para nenhum dos 68 miRNAs mais expressos nos ovários inativos. O mRNA representado pelo GB42182 foi o que apresentou o maior número de sítios de ligação para os miRNAs podendo ser alvejado

pelo miR-let-7, miR-3049, miR-981 e miR-375. Como os transcritos do GB42182 e do miR-let-7 foram validados experimentalmente e por possuírem expressões antagônicas em relação ao estado de ativação dos ovários, podemos dizer que há uma possibilidade do miRNA-let-7 estar impedindo a tradução do mRNA representado pelo GB42182 e assim contribuindo de alguma forma no processo de repressão da ativação dos ovários de operárias. O transcrito do gene *yellow-g* (GB55171), assim como o transcrito do gene *cad* (GB48966), de acordo com nossa rede, podem ser regulados por três miRNAs, o primeiro pelos: miR-3719, miR-980 e miR-3785 e o segundo pelos: miR-315, miR-3715 e miR-190. Por fim, para o transcrito do gene representado pelo GB44975 foi verificado que apenas um miRNA, o miR-316, de acordo com nossa predição, poderia alveja-lo.

A rede integrada de expressão gênica construída a partir dos dados aqui apresentados e os resultados das validações experimentais acrescentam elementos importantes para compreensão da regulação do processo reprodutivo em *A. mellifera*, além de abrir novos caminhos nesta área do conhecimento.

Figura 30: Rede de interação entre miRNAs e mRNAs. Os miRNAs que compõem esta rede encontram-se mais expressos nos ovários inativos obtios a partir de resultados de RNAseq e os mRNAs que compõem esta rede encontram-se mais expressos em ovários ativados. As predições miRNAs-alvos foi realizada com base na ferramentas computacionais PITA e RNAhybrid.

CONCLUSÃO

5.0 Conclusão

Em relação ao estudo morfológico dos ovários ativados de operárias *A. mellifera*, concluímos que o tamanho das câmaras ovocíticas consecutivas, em relação a sua posição no vitelário, são estatisticamente diferentes entre si e diretamente correlacionados a um incremento do número de células foliculares. Rainhas e operárias africanizadas a partir da oitava câmara ovocítica aumentam sensivelmente a área dos ovócitos sugerindo uma maior incorporação da vitelogenina, por meio do fenômeno denominado patência. Nas operárias, a patência foi visualizada como um espaçamento não contínuo entre as células foliculares formado por vesículas intercaladas por trechos de adesão celular, que ao se aproximar do espaço que separa o ovócito das células foliculares forma um complexo de vesículas e membranas. Os achados referentes à patência e desenvolvimento dos folículos ovarianos são inéditos e registram importantes aspectos ainda não descritos em outros insetos.

A análise dos dados gerados por meio da técnica de RNAseq nos revelou que as nossas bibliotecas de mRNAs e miRNAs apresentaram qualidades como uniformidade, reprodutibilidade e robustez, essenciais para suportar as conclusões baseadas em nossas análises *in silico* e experimentais. Além disso, os resultados obtidos foram consistentes até mesmo ao serem comparados com outros bancos de dados gerados por protocolos e amostras diferentes. Todos os mRNAs selecionados para a validação experimental foram validados por RT-PCR e 81% dos miRNAs selecionados para a validação corroboraram os dados de expressão obtidos pelo seqüenciamento de última geração.

A utilização da técnica de RNAseq para comparação do estado fisiológico funcional e não funcional, a partir de diferentes amostras de ovários de abelhas rainhas e operárias de *A. mellifera*, nos revelou que existe um perfil de expressão (genes

estruturais e reguladores) que difere o estado reprodutivo do não reprodutivo e que, inclusive, se mostra muito semelhante entre as castas.

Somos motivados a concluir que a diferente modulação dos microRNAs e mRNAs de acordo com o estado fisiológico dos ovários, criam uma rede integrativa onde os genes são reprimidos ou têm a expressão liberada permitindo o estado reprodutivo dos ovários ou mantendo-os como não-reprodutivo.

Finalmente, apontamos uma lista de potenciais (ativadores/inibidores) reguladores do processo de ativação e desenvolvimento dos ovários em operárias e rainhas, dentre eles destacam-se: A) *fpps5*, cuja expressão durante a ovogênese, nos revelou que a sinalização pelos ecdisteróides é um componente importante, senão essencial, para o início do processo de ativação dos ovários; B) gene homeótico *caudal*, altamente expresso nas abelhas reprodutivas, sugerindo um papel relevante não apenas na embriogênese, mas também no processo de ovogênese, como um organizador do citoplasma do ovo; B) *obp7* que quando encontrado em tecido não olfatório, possivelmente assuma funções relacionadas ao transporte dos compostos lipofílicos, que são fundamentais para o processo de ativação e desenvolvimento dos ovários; C) *yellow-g* que por ser muito semelhante ao seu ortólogo em *D. melanogaster* deva desempenhar em *A. mellifera* as mesmas funções, ou seja, participação no processo de formação de uma membrana vitelínica apropriada, além de conferir rigidez ao envoltório do ovo.

REFERÊNCIAS BIBLIOGRÁFICAS

6.0 Referências Bibliográficas

- Albert S.; Klaudiny J. (2004) The MRJP/YELLOW protein family of *Apis mellifera*: identification of new members in the EST library. *J Insect Physiol.*
- Allsopp, M. H.; Calis J. N. M.; Boot W. J. (2003) Differential feeding of worker larvae affects caste characters in the Cape honeybee, *Apis mellifera capensis*. *Behavioral Ecology and Sociobiology*
- Almeida M.I.; Reis R.M.; Calin G.A. (2011) MicroRNA history: discovery, recent applications, and next frontiers. *Mutat Res.*
- Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W. (1990). Lipman DJ. Basic local alignment search tool. *J Mol Biol.*
- Ambros V. (2003) MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell,*
- Applied Biosystems (1997) User Bulletin #2: ABI Prism 7700 Sequence Detection System. Applied Biosystems Inc., Foster City, CA. Disponível online: docs.appliedbiosystems.com/pebiiodocs/04303859.pdf
- Arakane Y.; Dittmer N. T.; Tomoyasu Y.; Kramer K. J.; Muthukrishnan S.; Beeman R. W.; Kanost M. R. (2010) Identification, mRNA expression and functional analysis of several yellow family genes in *Tribolium castaneum*. *Insect Biochem Mol Biol.*
- Aravin, A.A.; Lagos-Quintana, M.; Yalcin, A.; Zavolan, M.; Marks, D.; Snyder, B.; Gaasterland, T.; Meyer, J.; Tuschl, T. (2003). The small RNA profile during *Drosophila melanogaster* development. *Dev Cell.*
- Asmann, Y.W.; Wallace, M.B.; Thompson, E.A. (2008). Transcriptome profiling using next-generation sequencing. *Gastroenterology.*
- Audic S.; Claverie J.M.; (1997) The significance of digital gene expression profiles. *Genome Res.*
- Baer B, Eubel H, Taylor NL, O'Toole N, Millar AH. (2009) Insights into female sperm storage from the spermathecal fluid proteome of the honeybee *Apis mellifera*. *Genome Biol.*
- Barchuk, A.R.; Bitondi, M.M.G.; Simões, Z.L.P. (2002) Effects of juvenile hormone and ecdysone on the timing of vitellogenin appearance in hemolymph of queen and worker pupae of *Apis mellifera*. *J. Insect of Science*, 2.1:8 pages, available online <http://www.insectscience.org/2.1>.
- Bartel D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism and function. *Cell.*
- Bartel D. P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell ,*
- Beekman, M.; Calis, J. N.; Boot, W. J. (2000) Parasitic honeybees get royal treatment. *Nature.*

- Behura, S.K. (2007). Analysis of nuclear copies of mitochondrial sequences in honeybee (*Apis mellifera*) genome. *Mol Biol Evol.*
- Bellés X.; Cassier P.; Cerdá X.; Pascual N.; André M.; Rósso Y.; Piulachs MD. (1993) Induction of choriogenesis by 20-hydroxyecdysone in the German cockroach. *Tissue Cell.*
- Berezikov, E.; Robine, N.; Samsonova, A.; Westholm, J.O.; Naqvi, A.; Hung, J.H.; Okamura, K.; Dai, Q.; Bortolamiol-Becet, D.; Martin, R.; Zhao, Y.; Zamore, P.D.; Hannon, G.J.; Marra, M.A.; Weng, Z.; Perrimon, N.; Lai, E.C. (2001) Deep annotation of *Drosophila melanogaster* microRNAs yields insights into their processing, modification, and emergence. *Genome Res.*
- Berleth T.; Bnrri M., Thoma G.; Bopp D.; Richsteln S.; Frigerlo G.; Noll M.; Nusslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo *EMBO.*
- Behura, S.K.; Whitfield, C.W. (2010). Correlated expression patterns of microRNA genes with age-dependent behavioural changes in honeybee. *Insect Mol Biol.*
- Biessmann H. (1985) Molecular analysis of the yellow gene (y) region of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A*
- Bitondi, M. M. G.; Simões, Z. L. P. (1996) The relationship between level of pollen in the diet, vitellogenin and juvenile hormone titres in Africanized *Apis mellifera* workers. *J. Apic. Res.*
- Bloch G.; Borst D. W, Huang Z.; Robinson G. E.; Cnaani J.; Hefetz A. (2000) Juvenile hormone titers, juvenile hormone biosynthesis, ovarian development and social environment in *Bombus terrestris*. *J Insect Physiol.*
- Brameier M.; Krings A.; MacCallum R.M. (2007) NucPred--predicting nuclear localization of proteins. *Bioinformatics.*
- Brehme K. S. (1941) The effect of adult body color mutations upon the larva of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A.*
- Bueno, O.C.; Morini, M.S.C. (1993). Crescimento das câmaras ovocíticas e células nutridoras durante a ovogênese de rainhas de *Apis mellifera* L. (Hym., Apidae). *Naturalia.*
- Bullard J.H.; Purdom E.; Hansen K.D.; Dudoit S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics.*

- Büning, J. (1994) The insect ovary: ultrastructure, pre-vitellogenic growth and evolution. London. Chapman & Hall. pp. 400.
- Burnet B; Wilson R. (1980) Pattern mosaicism for behavior controlled by the yellow locus in *Drosophila melanogaster*. Genet Res..
- Buttstedt A.; Moritz R. F.; Erler S. (2013) Origin and function of the major royal jelly proteins of the honeybee (*Apis mellifera*) as members of the yellow gene family. Biol Rev Camb Philos Soc.
- Cai H.; Chen H.; Yi T.; Daimon C.M.; Boyle J.P.; Peers C.; Maudsley S.; Martin B. (2013) VennPlex--a novel Venn diagram program for comparing and visualizing datasets with differentially regulated datapoints. PLoS One.
- Cardoen D, Ernst UR, Boerjan B, Bogaerts A, Formesyn E, de Graaf DC, Wenseleers T, Schoofs L, Verleyen P. (2012). Worker honeybee sterility: a proteomic analysis of suppressed ovary activation. J Proteome Res.
- Cardoen D, Wenseleers T, Ernst UR, Danneels EL, Laget D, DE Graaf DC, Schoofs L, Verleyen P. (2011). Genome-wide analysis of alternative reproductive phenotypes in honeybee workers. Mol Ecol.
- Casanova, J.; Furriols, M.; McCormick, C. A.; Struhl, G. (1995) Similarities between trunk and spätzle, putative extracellular ligands specifying body pattern in *Drosophila*. *Genes Dev*.
- Cassid, J.D.; King, R. C. (1972). Ovarian in development in *Habrobracon juglandis* (Asheamed) (Hym.: Braconidae). I – The origin and differentiation of the oocyte – nurse cells complex. Biol. Bull.
- Castillo-Gracia, M.; Couillaud, F. (1999) Molecular cloning and tissue expression of an insect farnesyl diphosphate synthase. Eur. J. Biochem.,
- Ciudad Garrido (2005) Mecanismos reguladores de la maduración ovárica en *B. germanica* (L)(Dictyoptera, Blattellidae). Trabajo realizado en el departamento de Fisiología y Biodiversidad Molecular del Instituto de Biología Molecular de Barcelona (C.S.I.C)
- Chapman, R.F. (1998) The insects: structure and function. 4. Ed. New York: Elsevier. 770p.
- Chasan, R.; Anderson, K.V. (1993) Maternal control of dorsal-ventral polarity and pattern in the embryo. In *The development of Drosophila melanogaster* (ed. M. Bate and A. Martinez Arias), pp. 387–424. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Chaud-Neto, J.; Bueno, O. C. (1979) Number of ovarioles in workers of *Apis mellifera adansonii* and of *Apis mellifera ligustica*: a comparative study. J. Apic. Res.
- Chen H.; Boutros P.C. (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics.
- Chen JS, Sappington W, Raikhel AS. (1997) Extensive sequence conservation among insect, nematode, and vertebrate vitellogenins reveals common ancestry. J Mol Evol.

- Chen, X.; Yu, X.; Cai, Y.; Zheng, H.; Yu, D.; Liu, G.; Zhou, Q.; Hu, S.; Hu, F. (2010) Next-generation small RNA sequencing for microRNAs profiling in the honey bee *Apis mellifera*. *Insect Mol Biol.* Dec.
- Chen, Y. C.; Lin, S. I.; Chen, Y. K.; Chiang, C. S.; Liaw, G. J. (2009) The Torso signaling pathway modulates a dual transcriptional switch to regulate *tailless* expression. *Nucleic Acids Res.*
- Chu, C.Y., Rana, T.M. (2007). Small RNAs: regulators and guardians of the genome. *J. Cell Physiol.*
- Claycomb J. M., Benasutti M., Bosco G., Fenger D. D., Orr-Weaver T. L. (2004) Gene amplification as a developmental strategy: Isolation of two developmental amplicons in *Drosophila*. *Dev. Cell.*
- Corona M, Velarde RA, Remolina S, Moran-Lauter A, Wang Y, Hughes KA, Robinson GE. (2007) Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. *Proc Natl Acad Sci U S A.*
- Cristino A.S., Tanaka E.D., Rubio M., Piulachs M.D. & Belles X. (2011) Deep sequencing of organ- and stage-specific microRNAs in the evolutionarily basal insect *Blattella germanica* (L.) (Dictyoptera, Blattellidae). *PLoS One.*
- Cruz-Landim, C. (2009) *Abelhas: morfologia e função de sistemas*. São Paulo: Editora Unesp, p.407.
- Davey, K. G. (1993) Hormonal integration of egg production in *Rhodnius prolixus*. *Amer. Zool.*
- de Cuevas, M.; Lilly, M. A.; Spradling A. C. (1997) Germline cyst formation in *Drosophila*. *Annual Reviews of Genetics.*
- De Gregorio E.; Spellman P.T.; Rubin, G.M.; Lemaitre, B. (2001) Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. *Proc. Natl. Acad. Sci.*
- de Kort, C.D.A.; Granger, N.A. (1996) Regulation of JH titers: The relevance of degradative enzymes and binding proteins. *Arch. Insect Biochem. Physiol.*
- 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. (2006) Patterns of conservation and change in honey bee developmental genes. *Genome Res.*
- Deyu Z.; Leal W. S. (2002) Conformational isomers of insect odorant-binding proteins. *Arch Biochem Biophys.*
- Diagana T. T.; Thomas U.; Prokopenko S. N.; Xiao B.; Worley P. F.; Thomas J. B. (2002) Mutation of *Drosophila* *homer* disrupts control of

- Djuranovic S.; Nahvi A.; Green R. (2012) miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science*.
- Dow M. A. (1976) The genetic basis of receptivity of yellow mutant *Drosophila melanogaster* females. *Behav Genet*. 1976
- Drapeau M. D. (2001) The family of yellow-related *Drosophila melanogaster* proteins. *Biochem Biophys Res Commun*.
- Drapeau M. D.; Albert S.; Kucharski R.; Prusko C.; Maleszka R. (2006) Evolution of the Yellow/Major Royal Jelly Protein family and the emergence of social behavior in honey bees. *Genome Res*.
- Drapeau M. D, Cyran S. A, Viering M. M, Geyer P. K, Long A. D. (2006) A cis-regulatory sequence within the yellow locus of *Drosophila melanogaster* required for normal male mating success. *Genetics*
- Drapeau M. D, Radovic A, Wittkopp P. J, Long A. D. (2003) A gene necessary for normal male courtship, yellow, acts downstream of fruitless in the *Drosophila melanogaster* larval brain. *J Neurobiol*.
- Engels W (1974) Occurrence and significance of vitellogenins in female castes of social Hymenoptera. *Am Zool*
- Engels, W.; Kaatz, H.; Zillikens, A.; Simões, Z. L. P.; Trube, A.; Braun, R.; Dittrich, F. (1990) Honey bee reproduction: vitellogenin and caste specific regulation of fertility. *Advances in Invertebrate Reproduction*.
- Esteller M. (2008) Epigenetics in cancer. *N Engl J Med*.
- Esteller M.; Almouzni G. (2005) How epigenetics integrates nuclear functions. Workshop on epigenetics and chromatin: transcriptional regulation and beyond. *EMBO Rep*.
- Eun S.H., Stoiber P.M., Wright H.J., McMurdie K.E., Choi C.H., Gan Q., Lim C. & Chen X. (2013) MicroRNAs downregulate Bag of marbles to ensure proper terminal differentiation in the *Drosophila* male germline. *Development*.
- Ferguson LC, Green J, Surridge A, Jiggins CD. (2011) Evolution of the insect yellow gene family. *Mol Biol Evol*.
- Fernandez-Valverde, .SL.; Taft, R.J.; Mattick, J.S. (2010) Dynamic isomiR regulation in *Drosophila* development. *RNA*.
- Friedman JM, Jones PA. (2009) MicroRNAs: critical mediators of differentiation, development and disease. *Swiss Med Wkly*.
- Friedländer M.R.; Chen W.; Adamidi C.; Maaskola J.; Einspanier R.; Knespel S.; Rajewsky N. (2008) Discovering microRNAs from deep sequencing data using miRDeep. *Nat Biotechnol*.

- Forêt S.; Wanner K. W.; Maleszka R. (2007) Chemosensory proteins in the honey bee: Insights from the annotated genome, comparative analyses and expressional profiling. *Insect Biochem Mol Biol*.
- Forêt, S.; Maleszka, R. (2006) Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res*.
- Frohnhofer H. G., Nusslein-Volhard C. (1986) Organization of anterior pattern in the *Drosophila* embryo by the maternal gene bicoid. *Nature*.
- Förstemann K., Tomari Y., Du T., Vagin V.V., Denli A.M., Bratu D.P., Klattenhoff C., Theurkauf W.E., Zamore P.D. (2005). Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biol*.
- Garcia-Solache M.; Jaeger J.; Akam M. (2010) A systematic analysis of the gap gene system in the moth midge *Clogmia albipunctata*. *Dev. Biol*
- Gavis E. R.; Lehmann R. (1994) Translational regulation of nanos by RNA localization. *Nature*.
- Gavis, E. R.; Lunsford, L.; Bergsten, S. E.; Lehmann, R. (1996) A conserved 90 nucleotide element mediates translational repression of nanos RNA. *Development*.
- Ghabrial, A.; Schüpbach, T. (1999) Activation of a meiotic checkpoint regulates translation of Gurken during *Drosophila* oogenesis. *Nat Cell Biol*.
- Gilbert, L.I.; Granger, N.A.; Roe, R.M. (2000) The juvenile hormones: historical facts and speculations on future research directions. *Insect Biochem. Mol. Biol*.
- Gilboa, L.; Lehmann, R. (2004) How different is Venus from Mars? The genetics of germ-line stem cells in *Drosophila* females and males. *Development*. Oct;131(20):4895-905. *Review*.
- Goode, S.; Morgan, M.; Liang, Y. P.; Mahowald, A. P. (1996) *brainiac* encodes a novel, putative secreted protein that cooperates with Grk TGF alpha in the genesis of the follicular epithelium. *Dev Biol*.
- Goode, S.; Wright, D.; Mahowald, A. P. (1992) The neurogenic locus brainiac cooperates with the *Drosophila* EGF receptor to establish the ovarian follicle and to determine its dorsal-ventral polarity. *Development*.
- Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. (2008) miRBase: tools for microRNA genomics. *Nucleic Acids Res*.
- Grozinger C.M.; Fan Y.; Hoover S.E.; Winston M. L. (2007) Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol Ecol*.
- Gundersen-Rindal D.E.; Pedroni M.J. (2010) Larval stage *Lymantria dispar* microRNAs differentially expressed in response to parasitization by *Glyptapanteles flavicoxis* parasitoid. *Arch Virol*.

- Guidugli K. R.; Nascimento A. M.; Amdam G. V.; Barchuk A. R.; Omholt S.; Simões Z. L.; Hartfelder K. (2005) Vitellogenin regulates hormonal dynamics in the worker caste of a eusocial insect. *FEBS Lett.*
- Guidugli K. R.; Piulachs M. D.; Bellés X.; Lourenço A. P.; Simões Z. L. (2005) Vitellogenin expression in queen ovaries and in larvae of both sexes of *Apis mellifera*. *Arch Insect Biochem Physiol.*
- Guidugli-Lazzarini KR, Nascimento AM, Tanaka ED, Piulachs MD, Hartfelder K, Bitondi MG, Simões ZLP. (2008) Expression analysis of putative vitellogenin and lipophorin receptors in honey bee (*Apis mellifera* L.) queens and workers. *Journal of Insect Physiology.*
- Han, Q., Fang, J., Ding, H., Johnson, J.K., Christensen, B.M., and Li, J. (2002) Identification of *Drosophila melanogaster* yellow-f and yellow-f2 proteins as dopachrome-conversion enzymes. *Biochem.*
- Hartfelder, K.; Bitondi, M. M. G.; Santana, W. C.; Simoes, Z. L. P. (2002) Ecdysteroid titer and reproduction in queens and workers of the honey bee and of a stingless bee: loss of ecdysteroid function at increasing levels of sociality? *Insect Biochem. Mol. Biol.*
- Hartfelder, K.; Engels, W. (1998) Social insect polymorphism: hormonal regulation of plasticity in development and reproduction in the honeybee. *Curr Top Dev Biol.*
- Hong, C. C.; Hashimoto, C. (1995) An unusual mosaic protein with a protease domain, encoded by the nudel gene, is involved in defining embryonic dorsoventral polarity in *Drosophila*. *Cell. Sep*
- Hoover, S.E.; Higo, H. A.; Winston, M. L. (2006) Worker honey bee ovary development: seasonal variation and the influence of larval and adult nutrition. *J Comp Physiol B.*
- Iovino, N.; Pane, A.; Gaul, U. (2009). miR-184 has multiple roles in *Drosophila* female germline development. *Dev Cell.*
- Irish V.; Lehmann R.; Akam M. (1989) The *Drosophila* posterior-group gene nanos functions by repressing hunchback activity. *Nature.*
- Izumi, S.; Yano, k.; Yamamoto, Y.; Takahashi, S.Y. (1994) Yolk proteins form insect eggs: Structure, biosynthesis and programmed degradation during embryogenesis. *J. Insect Physiol.*
- Jagadeeswaran G., Zheng Y., Sumathipala N., Jiang H., Arrese E.L., Soulages J.L., Zhang W., Sunkar R. (2010) Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development. *BMC Genomics.*

- Jiang Y. H.; Bressler J.; Beaudet A. L. (2004) Epigenetics and human disease. *Annu Rev Genomics Hum Genet*.
- Kamakura, M (2011). Royalactin induces queen differentiation in honeybees. *Nature.*,
- Kapil, R. P. (1962) Anatomy and histology of the female reproductive system of *Apis indica* F. (Hymenoptera, Apidae). *Insectes Souciaux*.
- Kertesz M.; Iovino N.; Unnerstall U.; Gaul U.; Segal E. (2007) The role of site accessibility in microRNA target recognition. *Nat Genet*.
- Kikuchi, K.; Hirai, M.; Shiotsuki, T.; (2001) Molecular cloning and tissue distribution of farnesyl pyrophosphate synthase from the silkworm *Bombyx mori*. *J. Insect Biotech. Ser*.
- King, R. C. (1970) Ovarian Development in *Drosophila melanogaster*. New York: Academic Press.
- King, R. C.; Cassidy, J. D.; Rousset, A. (1982) The formation of clones of interconnected cells during gametogenesis in insects. In R.C. King & H. Akai (eds). "Insect Ultrastructure" Plenum Press. New York.
- King, R. C.; Cassidy, J. D.; Rousset, A. (1982) The formation of clones of interconnected cells during gametogenesis in insects. In R.C. King & H. Akai (eds). "Insect Ultrastructure" Plenum Press. New York.
- Kinjoh, T.; Kaneko, Y.; Itoyama, K.; Mita, K.; Hiruma, K.; Shinoda, T. (2007) Control of juvenile hormone biosynthesis in *Bombyx mori*: Cloning of the enzymes in the mevalonate pathway and assessment of their developmental expression in the corpora allata. *Insect Biochem. Mol. Biol*.
- Koyama, T.; Matsubara, M.; Ogura, K. (1985) Isoprenoid enzyme systems of silkworm I. Partial purification of isopentenyl pyrophosphate isomerase, farnesyl pyrophosphate synthetase, and geranylgeranyl pyrophosphate synthetase. *J. Biochem*.
- Kozomara A., Griffiths-Jones S. (2011) miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.*, 39(Database issue): D152-157
- Krieger J.; Breer H. (1999) Olfactory reception in invertebrates. *Science*. Oct 22;286(5440):720-3. Review.
- Kucharski R.; Maleszka R.; Hayward D. C.; Ball E. E. (1998) A royal jelly protein is expressed in a subset of Kenyon cells in the mushroom bodies of the honey bee brain. *Naturwissenschaften*.

- Lane, M. E.; Kalderon, D. (1994) RNA localization along the anteroposterior axis of the *Drosophila* oocyte requires PKA-mediated signal transduction to direct normal microtubule organization. *Genes Dev.* Dec.
- Lantz V.; Ambrosio L.; Schedl P. (1992) The *Drosophila* orb gene is predicted to encode sex-specific germline RNA-binding proteins and has localized transcripts in ovaries and early embryos. *Development*.
- Lauverjat S.; Szollosi A.; Marcaillou C. (1984). Permeability of the ovarian follicle during oogenesis in *Locusta migratoria* L. (Insecta, Orthoptera). *J. Ultrastruct.*
- Leal W. S. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol*.
- Leaman, D.; Chen, P.Y.; Fak, J.; Yalcin, A.; Pearce, M.; Unnerstall, U.; Marks, D.S.; Sander, C.; Tuschl, T.; Gaul, U. (2005). Antisense-mediated depletion reveals essential and specific functions of microRNAs in *Drosophila* development. *Cell*. Jul1.
- Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P., Ra°dmark O., Kim S., Kim V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*,
- Leimar, O.; Hartfelder, K.; Laubichler, MD.; Page, RE Jr. (2-12) Development and evolution of caste dimorphism in honeybees – a modeling approach. *Ecol. Evol.*
- Lehninger, A.L.; Nelson, D.L.; Cox, M.M. (1995) *Princípios de bioquímica*. Savier: São Paulo.
- Lemke S.; Busch S. E.; Antonopoulos D. A.; Meyer F.; Domanus M. H.; Schmidt-Ott, U. (2010) Maternal activation of gap genes in the hover fly *Episyrphus*. *Development*
- Liaw G. J.; Lengyel J. A. (1993) Control of tailless expression by bicoid, dorsal and synergistically interacting terminal system regulatory elements. *Mech.Dev.* 40, 47–61.
- Lin, H. (1997) The tao of stem cells in the germline. *Annual Review of Genetics*
- Lin, H.; Spradling, A. C. (1997) A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. *Development*.
- Linksvayer, T. A.; Rueppell, O.; Siegel, A.; Kaftanoglu, O.; Page Jr, R. E.; Amdam, G. V. (2009) The genetic basis of transgressive ovary size in honeybee workers. *Genetics*.
- Liu F, Li W, Li Z, Zhang S, Chen S, Su S. (2011). High-abundance mRNAs in *Apis mellifera*: comparison between nurses and foragers. *J Insect Physiol*.
- Liu F, Peng W, Li Z, Li W, Li L, Pan J, Zhang S, Miao Y, Chen S, Su S. (2012) Next-generation small RNA sequencing for microRNAs profiling in *Apis mellifera*: comparison between nurses and foragers. *Insect Mol Biol*.
- Liu S., Xia Q., Zhao P., Cheng T., Hong K. & Xiang Z. (2007) Characterization and expression patterns of let-7 microRNA in the silkworm (*Bombyx mori*). *Bmc Developmental Biology*.

- Louveaux, J. (1976) Anatomie de l'abeille. IX. L'appareil reproducteur de la reine. France: Fiche Technique, Institute Technique de l'Apiculture, France. p.41-44.
- Lynch J. A.; Olesnick E. C.; Desplan C. (2006). Regulation and function of *tailless* in the long germ wasp *Nasonia vitripennis*. *Dev. GenesEvol.*
- Maleszka R.; Kucharski R. (2000) Analysis of *Drosophila* yellow-B cDNA reveals a new family of proteins related to the royal jelly proteins in the honeybee and to an orphan protein in an unusual bacterium *Deinococcus radiodurans*. *Biochem Biophys Res Commun.*
- Margolis, J.; Spradling, A. (1995) Identification and behavior of epithelial stem cells in the *Drosophila* ovary. *Development.*
- Martin M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *Bioinformatics in Action.*
- Martin, J. R.; Ollo, R. (1996) A new *Drosophila* Ca^{2+} /calmodulin-independent protein kinase (Caki) is localized in the central nervous system and implicated in walking speed. *EMBO J.*
- Mattick JS, Makunin IV. (2005) Small regulatory RNAs in mammals. *Hum Mol Genet.*
- Melo, A. C.; Valle, D.; Machado, E. A.; Salerno, A. P.; Paiva-Silva, G. O.; Cunha, E.; Silva, N. L.; de Souza, W.; Masuda, H.; (2000) Synthesis of vitellogenin by the follicle cells of *Rhodnius prolixus*. *Insect Biochem Mol Biol.*
- Michener, C. D.; Brothers D. J. (1974) Were workers of eusocial Hymenoptera initially altruistic or oppressed? *Proceedings of the National Academy of Sciences of the United States of America* 71:671-674.
- Mohanty, S. C. (1981). Growth of the vitellarium during oogenesis in the fly *Lucilia cuprina* (Diptera) oocyte-nurse cells interrelation . *Growth.*
- Mohanty, S. C. (1982). Relation growth of oocyte, nurse cells and follicle epithelium during oogenesis in the fly *Sarcophaga ruficornis* (Diptera). *Growth.*
- Morini, M. S. C.; Bueno, O. C. (1993) Crescimento das câmaras ovocíticas e células nutridoras durante a ovogênese de rainhas de *Apis mellifera* L. (Hym, Apidae). *Naturalia*, São Paulo.
- Mottet D.; Castronovo V. (2008) Histone deacetylases: target enzymes for cancer therapy. *Clin Exp Metastasis.*
- Münch D, Amdam GV. (2010) The curious case of aging plasticity in honey bees. *FEBS Lett.*
- Munoz-Torres M.C.; Reese J.T.; Childers C.P.; Bennett A.K.; Sundaram J.P.; Childs K.L.; Anzola J.M.; Milshina N.; Elisk C.G. (2011) Hymenoptera Genome Database: integrated community resources for insect species of the order Hymenoptera. *Nucleic Acids Res.*

- Naeger N.L.; Peso M.; Even N.; Barron A.B.; Robinson G.E. (2013) Altruistic behavior by egg-laying worker honeybees. *Curr Biol*.
- Nation, J.L., (2002). *Insect Physiology and Biochemistry*. CRC Press, New York.
- Nijhout, H.F.; Wheeler, D.E. (1982) Juvenile hormone and physiological basis of insect polymorphism. *Quarterly Review Biol*.
- Nguyen H.T.; Frasch M. (2006) MicroRNAs in muscle differentiation: lessons from *Drosophila* and beyond. *Curr Opin Genet Dev*.
- Nunes FM, Ihle KE, Mutti NS, Simões ZL, Amdam GV. (2013) The gene vitellogenin affects microRNA regulation in honey bee (*Apis mellifera*) fat body and brain. *J Exp Biol*.
- Odawara J.; Harada A.; Yoshimi T.; Maehara K.; Tachibana T.; Okada S.; Akashi K.; Ohkawa Y. (2011) The classification of mRNA expression levels by the phosphorylation state of RNAPII CTD based on a combined genome-wide approach. *BMC Genomics*.
- Oertel, E. (1930) Metamorphosis in the honey bee. *J. Morphol*.
- Oldroyd BP, Smolenski AJ, Cornuet JM, Crozier RH. (1994). Anarchy in the beehive. *Nature*
- Olesnicki E. C.; Brent A. E.; Tonnes L.; Walker M.; Pultz M.; Leaf D.; Desplan C. (2006) A caudal mRNA gradient controls posterior development in the wasp *Nasonia*. *Development*.
- Onhuma, S-I.; Nakazawa, T.; Hemmi, H.; Halberg, A.M.; Koyama, T.; Ogura, K.; Nishimo, T. (1996) Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. *J. Biol. Chem*.
- Osei-Amo S., Hussain M., O'Neill S.L. & Asgari S. (2012) Wolbachia-induced aae-miR-12 miRNA negatively regulates the expression of MCT1 and MCM6 genes in Wolbachia-infected mosquito cell line. *PLoS One*.
- Pascual, N. (1995) Fisiologia dels ecdisteroides ovarics de *Blatella germanica* (L.) (Dictyoptera, Blattellidae). *Tesis doctoral Universidad de Barcelona (UB)*. – Departamento de Fisiología, Barcelona.
- Pasquinelli A.E. (2012) MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nat. Rev. Genet*.
- Pelosi P. (1996). Perireceptor events in olfaction. *J. Neurobiol*.
- Pelosi P.; Calvello M.; Ban L. (2005) Diversity of odorant-binding proteins and chemosensory proteins in insects. *Chem Senses*.
- Piulachs MD, Guidugli KR, Barchuk AR, Cruz J, Simões ZLP, Bellés X. (2003) The vitellogenin of the honeybee, *Apis mellifera*: structural analysis of the cDNA and expression studies. *Insect Biochem Mol Biol*.

- Prud'homme B, Gompel N, Rokas A, Kassner V.A, Williams T.M, Yeh S.D, True J.R, Carroll S.B. (2006) Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature*.
- Rachinsky, A. (1994) Octopamine and serotonin influence on corpora allata activity in honey bee (*Apis mellifera*) larvae. *J. Insect Physiol.*, 40: 549-554.
- Rachinsky, A.; Hartfelder, K. (1990) Corpora allata activity, a prime regulating element for caste-specific juvenile hormone titre in honey bee larvae (*Apis mellifera carnica*). *J. Insect Physiol.*
- Raikhel, A. S.; Dhadialla T. S. (1992) Accumulation of yolk proteins in insect oocytes. *Annu Rev Entomol.* 1992;37:217-51. Review.
- Raikhel, A.S.; Lea, A. O. (1991) Control of follicular epithelium development and vitelline envelope formation in the mosquito; role of juvenile hormone and 20-hydroxyecdysone. *Tissue Cell*.
- Raikhel AS, Brown MR, Bellés X. (2004) Hormonal control of reproductive processes. In: Gilbert LI, Iatrou K, Gill SS, editors. *Comprehensive molecular insect science*, Vol. 3. Oxford: Elsevier. p 433–491.
- Rehmsmeier M.; Steffen P.; Hochsmann M.; Giegerich R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA*.
- Reinhart B.J., Slack F.J., Basson M., Pasquinelli A.E., Bettinger J.C., Rougvie A.E., Horvitz H.R., Ruvkun G. (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*.
- Rembold, H. (1987) Caste-specific modulation of juvenile hormone titers in *Apis mellifera*. *Insect. Biochem.*
- Ribeiro, J.M., Arca, B., Lombardo, F., Calvo, E., Phan, V.M., Chandra, P.K., Wikel, S.K., (2007). An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. *BMC Genomics* 8, 6
- Riddiford, L.M. (1996) Juvenile hormone: the status of its status quo action. *Arch. Insect Biochem. Physiol.*, 32: 271-286.
- Rigoutsos I. (2009) New tricks for animal microRNAs: targeting of amino acid coding regions at conserved and nonconserved sites. *Cancer Res*.
- Rivera-Pomar R.; Niessing D.; Schmidt-Ott U.; Gehring WJ.; Jäckle H. (1996) RNA binding and translational suppression by bicoid. *Nature*.

- Roberts A.; Trapnell C.; Donaghey J.; Rinn J.L.; Pachter L. (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol.*
- Robertson H. M.; Wanner K. W. (2006) The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res.*
- Robinson G.E.; Huang Z. Y. (1998). Colony integration in honey bees: genetic, endocrine, and social control of division of labor. *Apidologie.*
- Robinson, G.E.; Strambi, C.; Strambi, A.; Feldlaufer, M.F. (1991) Comparison of juvenile hormone and ecdysteroid titers in adult worker and queen honey bees (*Apis mellifera*). *J. Insect Physiol.*
- Robinson, G.E.; Strambi, C.; Strambi, A.; Huang, Z.Y. (1992) Reproduction in worker honey bees is associated with low juvenile hormone titers and rates of biosynthesis. *General Comparat. Endocrinol.*
- Robinson, G. E.; Vargo, E. L. (1997) Juvenile hormone in adult eusocial Hymenoptera: gonadotropin and behavioral pacemaker. *Arch. Insect Biochem. Physiol.*
- Rubenstein E. C. (1979). The role of an epithelial occlusion zone in the termination of vitellogenesis in *Hyalophora cecropia* ovarian follicles. *Dev. Biol.*
- Rubio M., de Horna A. & Belles X. (2012) MicroRNAs in metamorphic and non-metamorphic transitions in hemimetabolous insect metamorphosis. *BMC Genomics.*
- Ruohola, H.; Bremer, K. A.; Baker, D.; Swedlow, J. R.; Jan, L. Y.; Jan, Y. N. (1991) Role of neurogenic genes in establishment of follicle cell fate and oocyte polarity during oogenesis in *Drosophila*. *Cell.*
- Ruohola-Baker, H.; Grell, E.; Chou, T. B.; Baker, D.; Jan, L. Y.; Jan, Y. N. (1993) Spatially localized rhomboid is required for establishment of the dorsal-ventral axis in *Drosophila* oogenesis. *Cell.*
- Ruttner, F.; Hesse, B. (1981) Rassenspezifische Unterschiede in Ovaentwicklung und Eiablage von weisellosen Arbeiterinnen der Honigbiene (*Apis mellifera* L.), *Apidologie* **12**:159-183
- Santos A. M.; (2008) Biologia molecular de genes envolvidos no metabolismo do hormônio juvenil em *Apis mellifera*. Tese de Doutorado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP – Departamento de Genética
- Santos-Rosa H.; Caldas C. (2005) Chromatin modifier enzymes, the histone code and cancer. *Eur J Cancer.* 2005
- Sappington TW, Raikhel AS. (1998) Molecular characteristics of insect vitellogenins and vitellogenin receptors. *Insect Biochem Mol Biol.*
- Schmidt-Capella, I.C.; Hartfelder, K. (1998) Juvenile hormone effect on DNA synthesis and apoptosis in caste-specific differentiation of the larval honey bee (*Apis mellifera* L.) ovary. *J. Insect Physiol.*

- Schmidt Capella, I. C.; Hartfelder, K. (2002) Juvenile-hormone-dependent interaction of actin and spectrin is crucial for polymorphic differentiation of the larval honey bee ovary. *Cell Tissue Res.*
- Schmieder R.; Edwards R. (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics.*
- Schmitzova J.; Klaudiny J.; Albert S.; Schroder W.; Schreckengost W.; Hanes J.; Judova J.; Simuth J. (1998) A family of major royal jelly proteins of the honeybee *Apis mellifera* L.. *Cell Mol Life Sci.*
- Schroder R.; Eckert C.; Wolff C.; Tautz D. (2000) Conserved and divergent aspects of terminal patterning in the beetle *Tribolium castaneum*. *Proc.Nat.Acad.Sci. U. S.A.* 97,6591–6596.
- Schröder, R. (2003) The genes *orthodenticle* and *hunchback* substitute for *bicoid* in the beetle *Tribolium*. *Nature.*
- Seehuus SC, Norberg K, Gimsa U, Krekling T, Amdam GV. (2006) Reproductive protein protects functionally sterile honey bee workers from oxidative stress. *Proc Natl Acad Sci U S A.*
- Seehuus SC, Norberg K, Krekling T, Fondrk MK Amdam GV (2007) Immunogold localization of vitellogenin in the ovaries, hypopharyngeal glands and head fat bodies of honeybee workers, *Apis mellifera*. *J Insect Sci.*
- Sen, S.E.; Sperry, A.E. (2002) Partial purification of a FPP synthase from whole-body *Manduca sexta*. *Insect Biochem. Mol. Biol.*
- Sengupta, S.; Nie, J.; Wagner, R.J.; Yang, C.; Stewart, R.; Thomson, J.A. (2009). MicroRNA 92b controls the G1/S checkpoint gene p57 in human embryonic stem cells. *Stem Cells.*
- Shannon P.; Markiel A.; Ozier O.; Baliga N.S.; Wang J.T.; Ramage D.; Amin N.; Schwikowski B.; Ideker T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*
- Shin C., Nam J.W., Farh K.K., Chiang H.R., Shkumatava A., Bartel D.P. (2010) Expanding the microRNA targeting code: functional sites with centered pairing. *Mol. Cell.*
- Shpiz S, Kwon D, Rozovsky Y, Kalmykova A. (2009) rasiRNA pathway controls antisense expression of *Drosophila* telomeric retrotransposons in the nucleus. *Nucleic Acids Res.*
- Siomi H., Siomi M.C. (2010) Posttranscriptional regulation of microRNA biogenesis in animals, *Mol. Cell.*
- Snodgrass, R. E. (1956) *Anatomy of the honey bee*. Comstock Publ. Assoc., Ithaca and London.
- Sprenger, F.; Nüsslein-Volhard, C. (1992) Torso receptor activity is regulated by a diffusible ligand produced at the extracellular terminal regions of the *Drosophila* egg. *Cell.*
- Stanyon, C.A., Liu, G., Mangiola, B.A., Patel, N., Giot, L., Kuang, B., Zhang, H., Zhong, J., and Finley Jr., R.L. (2004) A *Drosophila* protein-interaction map centered on cell-cycle regulators. *Genome Biol.* 5: R96.

- St Johnston, D.; Nüsslein-Volhard, C. (1992) The origin of pattern and polarity in the *Drosophila* embryo. *Cell*. Jan 24;68(2):201-19. Review.
- St. Johnston, D. (1993). Pole Plasm and the Posterior Group Genes. In: Bate, M. and Martinez Arias, A. Editors, 1993. The Development of *Drosophila melanogaster* Cold Spring Harbor Press, New York, pp. 325–364.
- Szkopinska, A.; Plochocka, D. (2005) Farnesyl diphosphate synthase of product specificity. *Acta Biochimica Polonica*, 52: 45-55.
- Tanaka, E. D.; Hartfelder K. (2004) The initial stages of oogenesis and their relation to differential fertility in the honey bee (*Apis mellifera*) castes. *Arthropod Struct Dev*.
- Tay Y., Zhang J., Thomson A.M., Lim B., Rigoutsos I. (2008) MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. *Nature*.
- Telfer, W.H. (1965). The mechanism and control of yolk formation. *Annu. Rev. Entomol*.
- Telfer, W.H.; Huebner, E.; Smith, D.S. (1982) The cell biology of vitellogenic follicles in *Hyalophora* and *Rhodnius*. In: R.C. King and H. Akai, Editors, *Insect Ultrastructure*, Plenum, New York, pp. 118–149
- Teng, X.; Xiao, H. (2009) Perspectives of DNA microarray and next-generation DNA sequencing technologies. *Sci China C Life Sci.*
- The Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*.
- Thompson, G. J.; Kucharski, R.; Maleszka, R.; Oldroyd, B. P. (2008) Genome-wide analysis of genes related to ovary activation in worker honey bees. *Insect Mol Biol*.
- Ting A. H.; McGarvey K. M.; Baylin S. B. (2006) The cancer epigenome--components and functional correlates. *Genes Dev*.
- Tinker, R.; Silver, D.; Montell, D. J. (1998) Requirement for the vasa RNA helicase in gurken mRNA localization. *Dev Biol*.
- Trapnell C.; Hendrickson D.G.; Sauvageau M.; Goff L.; Rinn J.L.; Pachter L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol*.
- Trapnell C.; Pachter L.; Salzberg S.L. (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*.
- Trapnell, C.; Williams B.A.; Pertea G.; Mortazavi A.; Kwan G.; van Baren M.J.; Salzberg S.L.; Wold B.J.; Pachter L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*.

- Trenczek T.; Engels W. (1986) Occurrence of vitellogenin in drone honeybees (*Apis mellifica*). *Int.J.Inv.Repr.*
- Trenczek, T.; Zillikens, A.; Engels, W. (1989) Developmental patterns of vitellogenin haemolymph titre and rate of synthesis in adult drone honey bees (*Apis mellifera*). *Journal of Insect Physiology*.
- Valeri N.; Vannini I.; Fanini F.; Calore F.; Adair B.; Fabbri M. (2009) Epigenetics, miRNAs, and human cancer: a new chapter in human gene regulation. *Mamm Genome*.
- Valle D. (1993) Vitellogenesis in insects and other groups: a review. *Mem I Oswaldo Cruz* 88:1–26.
- Wang C.; Dickinson L. K.; Lehmann R. (1994) Genetics of nanos localization in *Drosophila*. *Dev.Dyn.:Off.Publ.Am.Assoc.Anatom.* 199,103–115.
- Weaver, D.B.; Anzola, J.M.; Evans, J.D.; Reid, J.G.; Reese, J.T.; Childs, K.L.; Zdobnov, E.M.; Samanta, M.P.; Miller, J.; Elsik, C.G. (2007). Computational and transcriptional evidence for microRNAs in the honey bee genome. *Genome Biol.* 8(6):R97.
- Werren J. H.; Richards S.; Desjardins C. A.; Niehuis O.; Gadau J.; Colbourne J.; The Nasonia Working Group. (2010) Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science*.
- Wheeler, D.E. (1986) Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. *American Naturalist*.
- Wheeler DE, Kawooya JK (1990) *Arch Insect Biochem Physiol* 14:253–267.
- Wilson R; Burnet B; Eastwood L; Connolly K. (1976) Behavioural pleiotropy of the yellow gene in *Drosophila melanogaster*. *Genet Res*.
- Wilson M. J.; Dearden P. K. (2009) Tailless patterning functions are conserved in the honeybee even in the absence of Torso signaling. *Dev.Biol.*
- Wilson M. J.; Dearden P. K. (2011) Diversity in insect axis formation: two orthodenticle genes and hunchback act in anterior patterning and influence dorso-ventral organization in the honeybee (*Apis mellifera*). *Development*.
- Wilson M. J.; Havler M.; Dearden P. K. (2010) Giant, Krüppel, and caudal act as gap genes with extensive roles in patterning the honeybee embryo. *Dev Biol.* 2010
- Winston, M.L. (1987) *The Biology of the Honey Bee*. Harvard University Press, Cambridge.
- Wittkopp, P.J., True, J.R., and Carroll, S.B. (2002a) Reciprocal functions of the *Drosophila* Yellow and Ebony proteins in the development and evolution of pigment patterns. *Development*.

- Wittkopp, P.J., Vaccaro, K., and Carroll, S.B. (2002b) Evolution of yellow gene regulation and pigmentation in *Drosophila*. *Curr. Biol.*
- Wyatt GR. (1999) Yolk proteins, invertebrates. *Encyclopedia of reproduction*. vol. 4. Amsterdam: Elsevier. p 1029–1035.
- Yekta S., Shih I.H., Bartel D.P. (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science*.
- Xia A. H.; Zhou Q. X.; Yu L. L.; Li W. G.; Yi Y. Z.; Zhang Y. Z.; Zhang Z. F. (2006) Identification and analysis of YELLOW protein family genes in the silkworm, *Bombyx mori*. *BMC Genomics*.
- Xu, P.; Vernooy, S.Y.; Guo, M.; Hay, B.A. (2003). The *Drosophila* microRNA Mir-14 suppresses cell death and is required for normal fat metabolism. *Curr Biol*.
- Zimowska, G.; Shirk, P.D.; Silhacek, D.L.; Shaaya, E. (1994) Yolk sphere formation is initiated in oocytes before development of patency in follicles of the moth *Plodia interpunctella*. *Wilhelm Roux's Archives of Developmental Biology*.
- Zondag L, Dearden PK, Wilson MJ. (2012) Deep sequencing and expression of microRNAs from early honeybee (*Apis mellifera*) embryos reveals a role in regulating early embryonic patterning. *BMC Evol Biol*.

7.0 Apêndice

Tabela XII – Sequência dos nucleotídeos dos mRNAs utilizados para validação experimental por PCR em tempo real.

<i>Gene</i>	<i>Seqüências dos primers</i>
GB42182	Sense → 5' – TTGGGCGAGATTACCGAAGT – 3' Antisense → 5' – TTATTCCCCTGTCCTGCACC – 3'
GB44975	Sense → 5' – TCCTCCACCAAATCGACGAA – 3' Antisense → 5' – TCTCCATCGCTTGTGTACGA – 3'
<i>fpps5</i>	Sense → 5' – TGCCAAGAATCAACAATAGGAT – 3' Antisense → 5' – CCATACGGTCATTAATTTGTTTC – 3'
<i>cad</i>	Sense → 5' – CACAATCCGTCGTAAAGCCG – 3' Antisense → 5' – CTGTTTGCGTTCTTTTGCCC – 3'
<i>obp7</i>	Sense → 5' – ACTGCATAATACACATGGGCT – 3' Antisense → 5' – TCCAAAGCGCACGTAATGAA – 3'
<i>yellow-g</i>	Sense → 5' – AGGCTCTTATGCGTCATCCT – 3' Antisense → 5' – CCAATCCAGGCTGAGTCCA – 3'
<i>yellow-x2</i>	Sense → 5' – TCCATCTGACGGGGAGAAAC – 3' Antisense → 5' – AATCGGAACACCTTGCAACC – 3'
GB53748	Sense → 5' – AGACCGTGGTGCAAGCAGTT – 3' Antisense → 5' – CTTCGCCTCCTCTTTCCCAG – 3'
<i>vitelogenina</i> - <i>vg</i>	Sense → 5' – ACG ACT CGA CCA ACG ACT T – 3' Antisense → 5' – ACG AAA GGA ACG GTC AAT TCC – 3'
<i>EF1-α-F</i>	Sense → 5' – GGT CGT TCA TGT ATA TCA TC – 3' Antisense → 5' – CGG ACA AGC ACA ACT AAG AA – 3'
RP49	Sense → 5' – CGTCATATGTTGCCAACTGGT – 3' Antisense → 5' – TTGAGCACGTTCAACAATGG – 3'

* GB referem-se ao código dos genes preditos depositados na base de dados Official Gene Set – Elsik *et al.*, 2006

Tabela XIII - Sequência dos nucleotídeos dos miRNAs utilizados para validação experimental por PCR em tempo real.

Identificação do miRNA	Sequência de nucleotídeos do miRNA maduro
ame-miR-306	TCAGGTACTGAGTGACTCTGAG
ame-miR-1	TGGAATGTAAAGAAGTATGGAG
ame-miR-276	TAGGAACTTCATACCGTGCTCT
ame-miR-317	TGAACACAGCTGGTGGTATCTCAGT
ame-miR-184	TGGACGGAGAACTGATAAGGGC
ame-miR-100	AACCCGTAGATCCGAACCTGTG
ame-miR-92b	AATTGCACCCGTCCCGGCCTGA
ame-miR-9a	TCTTTGGTTATCTAGCTGTATGA
ame-miR-8	TAATACTGTCAGGTAAAGATGTC
ame-miR-11	CATCACAGGCAGAGTTCTAGTT
ame-miR-12	TGAGTATTACATCAGGTACTGGT
ame-miR-263a	GTAAATGGCACTGGAAGAATTCAC
ame-miR-2	TATCACAGCCAGCTTTGATGAGC
ame-miR-let-7	TGAGGTAGTAGGTTGTATAGT
ame-miR-13a	TATCACAGCCATTTTGATGAG
ame-miR-31a	GGCAAGATGTCGGCATAGCTGA
ame-miR-14	TCAGTCTTTTTCTCTCTCCTA
ame-miR-71	TGAAAGACATGGGTAGTGA
ame-miR-13b	TATCACAGCCATTTTGACGAT
ame-miR-125	CCCCTGAGACCCTAACTTGTGA
ame-miR-3720	ATACGGTGATGAGTTTAAACAG

Title:**MicroRNA signatures characterizing ovarian activity in honey bee (*Apis mellifera* L.) queens and workers****Authors**

Liliane M. F. Macedo¹, Francis M. F. Nunes², Flávia C. P. Freitas¹, Camilla V. Pires¹, Erica D. Tanaka⁵, Juliana R. Martins¹, Maria -D. Piulachs³, Alexandre S. Cristino⁴, Zilé L. P. Simões⁵

¹ Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo

² Departamento de Genética e Evolução – Universidade Federal de São Carlos

³ Institute of Evolutionary Biology, CSIC-Universitat Pompeu Fabra, Barcelona, Spain.

⁴ Queensland Brain Institute, University of Queensland, Brisbane, Australia.

⁵ Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto – Universidade de São Paulo

Correspondence to

ZLPS – zlpssimoe@usp.br

Email addresses:

Keywords: microRNA, transcriptome, plasticity, development, honeybee.

Abstract

In *Apis mellifera* female castes the morphological differences are extended to behavior and to main aspects of the reproductive physiology. During larval development worker ovaries have the ovarioles reduced to 2 to 20 from 200 originals that remain in adult queens. Although, worker ovaries could be activated under special conditions and produce viable haploid eggs. Here, we integrate data from our deep sequenced small RNA library from worker active ovaries with expression analysis of several microRNAs, and target predictions to get information about ovary activation. In the library were detected 165 microRNAs from the 218 registered for honeybee. The majority of the miRNA functions of the top 15 are related to ovary physiology or directly to oogenesis. For instance, the more expressed microRNAs, miR-306 that targets a tubulin, essential for polarizing the oocyte cytoplasm. It is differentially expressed in active ovaries either in workers or queens. Interestingly, less expressed microRNAs are differentially expressed in inactive ovaries, *e.g.* miR- 1, miR-31a and miR-12. qPCR analysis of members of the cluster let-7 and miR-2 demonstrated that the first is co-regulated, with the same number of reads in the library and the same

expression in ovaries, characterizing the inactive state, but the second cluster behave differently, miR-71 is the more expressed of this cluster, prediction tools indicate *vitellogenin* gene as its target, an essential gene for oogenesis. miR-71 is more expressed in active worker ovary. We conclude, the integrative network that coordinates ovary activation is a very fine tuning crosstalk among high and low expression of specific microRNAs whose profiles depict the reproductive condition of the members of the honeybee colony.

Introduction

In honeybees (*Apis mellifera*) female ontogenies are governed by identical genomes however, distinct quantities of proteins and carbohydrates offered to the young larvae trigger specific developmental pathways producing highly specialized female phenotypes, queens and workers. This resultant dimorphism achieves the organism as a whole, and the reproductive structures are markedly affected. The queens are equipped with huge ovaries and a well-developed spermatheca, an organ capable of storing millions of spermatozoa from their multiple mating that ensure, under ideal circumstance, 2,000 eggs per day, for many years. On the other hand, the workers develop profoundly modified reproductive organs, presenting small ovaries and vestigial spermatheca. Also, workers do not mate, and are known as the sterile members of this sophisticated society (Free 1987). The non-reproductive status of thousands of workers is maintained by chemical cues synthesized by the sole queen (Hoover *et al.* 2003) and by the larvae (Oldroyd *et al.* 2001; Maisonnasse *et al.* 2010) in a wild-type queenright colony. However, worker sterility is a reversible trait, under queenless condition some workers can activate their ovaries producing oocytes that are laid as unfertilized producing male eggs that develop into drones (for review see Free 1987).

The genetic mechanisms underlying that reproductive plasticity observed in adult workers is not completely understood. Most of studies focused on gene expression profiles to identify factors that control worker sterility. Comparisons were performed in different biological contexts (whole body, brain and abdominal tissues from wild-type workers, anarchist workers, narcotized workers and queens served as target for experiments) from gene-by-gene analysis (Thompson *et al.* 2007; Koywiwattrakul & Sittipraneed 2009; Vergoz *et al.* 2012) to large-scale genetic studies (Thompson *et al.* 2006; Grozinger *et al.* 2007; Oxley *et al.* 2008; Thompson *et al.* 2008; Cardoen *et al.* 2011). Thus, the control of the reproductive status of female honeybees is partially known in a coding gene context. In addition, the aforementioned studies collectively

point to the existence of an environmentally responsive regulatory network by which workers switch “on” or “off” their ovaries. In turn, it suggests that other genetic factors, such as microRNAs, should play roles in a potential complex network which regulates both ovary physiology and reproductive status.

MicroRNAs are small (19-24 nt) non-coding transcripts acting as inhibitors of the expression of targeted eukaryotic coding genes (reviewed by Bartel 2007). They emerged as key elements in the regulation of essentially all biological processes in animals and plants (Marco *et al.* 2013). For example, microRNAs are related to development (Zhang *et al.* 2009), cell proliferation and tissue size (Nolo *et al.* 2006), and alternative phenotypes (Legeai *et al.* 2010). The functional importance of microRNAs in insect ovaries was determined for holometabolous (John *et al.* 2009; Poulton *et al.* 2011) and hemimetabolous (Cristino *et al.* 2011) species. Since the already studied insects live in a non-social context, the investigation of *A. mellifera* undergoing ovary activation would provide a fascinating opportunity to understand the molecular pathways linked to reproduction in socially regulated environment.

Here we show the microRNA expression profile extracted from a small RNA-Seq library of honeybee worker active ovaries. We selected 18 of the most expressed miRNA to perform a comparative analysis using honeybee ovaries from queens and workers at different conditions of activation. As result we found that some microRNAs characterize the ovary status from virgin and mated queens and worker ovaries, both in the active or inactive state. Interestingly the most expressed microRNAs, miR-306 and miR-317, when tested against the four types of ovary, characterize active ovaries. Beside this, the less expressed in the library characterize the inactive ovaries, miR-1, miR-31a, and miR-12, are some examples. Moreover, it was considered for discussion two miRNA clusters, family of miRNA-2 and family let-7, where, in honeybees have 4 and 3 members respectively. We observed that all members of the family let-7, miR-let7, miR-100 and miR-125 are equally low expressed in the library and more expressed in inactive ovaries. The members of the family 2, miR-13a, miR-13b, miR-71 and miR-2, are observed working independently, having different number of reads in the library and characterizing active or inactive ovaries, depending on the number of reads in the library. The miRNA target prediction permitted us the construction of a coherent integrative miRNA:mRNA network reflecting the ovary state and gave very interesting results. *Vitellogenin* gene is targeted by miR-71, the sole microRNA among the more expressed that have this predicted function. miR-306, at the top of the more expressed

microRNA is suggested having only one target, a tubulin, essential for polarizing the oocyte at the beginning of development.

Material and Methods

Honeybee biological samples and RNA-seq library preparation

All samples were obtained from regular honey bee colonies maintained in the apiary of the Department of Genetics of Ribeirão Preto School of Medicine of the University of São Paulo. To liberate the workers from queen pheromones and initiate ovary activation they are left orphan by removing the queen from a well-developed colony; subsequently all queen-rearing cells were annihilated. The emerged workers 24 h after queen removal were paint marked and returned to the original orphaned colony. After ten days, a total of 70 marked workers were collected and have their ovaries dissected in ice cold saline (0.9%), examined and classified as activated or inactive. We considered as activated those ovaries containing developed oocytes at different stages. Workers presenting inactive ovaries were discarded. A total of 6 pairs of active ovaries were pooled to constitute one sample. The ovaries were transferred to a microtube containing 50 ul TRIzol® (Invitrogen) for total RNA isolation, following the protocol suggested by the manufacturer. RNA sample was quantified using Nanodrop (Nanodrop Technologies) and the quality was verified in agarose gel. Eight micrograms of RNA were used for small RNA-seq library preparation (fragments <200 nt) in accordance to Illumina protocol. Cluster generation and 40 bp single-end sequencing were performed (DNA Vision, Charleroi, Belgian, <http://www.dnavision.com>) on an Illumina Genome Analyzer (HiSeq 2000, Life Sciences).

Computational analysis of sequenced small RNA library

Base calling and adaptor sequences trimming were performed by Bioinformatics facility at DNA Vision. The reads were mapped to *A. mellifera* genome (version 4.0) (Honeybee Genome Sequencing Consortium 2006), using BWA (Li & Durbin 2009) with default parameters (0 to 2 mismatches were accepted). Only reads 19 to 24 nt long and uniquely mapped were analyzed further. Expressed miRNAs were identified by

custom-generated Python script using gff3 file from miRBase database (version 20). MicroRNAs digital expression (DE) was obtained by counting the number of reads mapped to each *A. mellifera* miRNAs.

Quantitative real-time PCR of selected miRNAs

We validate and characterized 19 miRNAs, from a list of the expressed miRNAs detected in our library and indicated by the current literature. The relative expression levels were evaluated in ovaries of virgin and egg-laying mated queens. The worker ovaries were separated in 2 classes: inactive ovaries (regular colony condition), and active (ovaries with oocytes in different stages of development). To evaluate the miRNA expression we performed real-time quantitative PCR using 7500Real Time PCR System (Applied Biosystems), miRNA specific primers and the NCode Universal qPCR reverse primer (Invitrogen). Amplification was carried out in a 20 μ L reaction volume containing 10 μ L SYBR[®] Green Master Mix 2x (Applied Biosystems), 1 μ L cDNA, 7,4 μ L water and 0.8 μ L (10 pmol/ μ L) of each primer (forward/reverse). The PCR conditions were 2 min at 50°C, 2 min at 96°C, followed by 40 cycles of 15 sec at 95°C and 33 sec at 60°C. Each sample was tested in triplicate and three biological samples were prepared from all tested ovary condition. Relative expression was calculated using the comparative C_T method and normalized to the expression of miR-11, miRNA-184 and miRNA-2. The stability these microRNAs used as reference genes were evaluated by two independent statistical applications: geNorm, NormFinder, since there are no universal reference genes described for each condition or species (Rev. Deloffre *et al* 2012, Liman *et al* 2013). The obtained quantitative data were analyzed using one-way ANOVA and t-test. All pair wise comparisons were considered significant when $p < 0.05$. The validated miRNAs and primers are described in the Table S1.

Prediction of a gene regulatory network

In order to investigate potential interactions between microRNAs validated in this study and gene targets, we used a set of proteomic data published by Cardoen *et al.* (2012). These authors identified proteins differentially expressed between active versus inactive ovaries of workers. Using the protein accession number, we recovered the

3'UTR sequences of the mRNA from GenBank database (NCBI). Mature sequences of the 18 studied microRNAs were recovered from miRBase release 20. Both 3'UTR and microRNA sequences were used as input to run target predictions by means of RNAHybrid tool (Rehmsmeier *et al.* 2004). To this end, the same criteria used by Nunes *et al.* (2013) were used, *ie.*, we accepted only the duplexes miRNA:mRNA presenting perfect seed matches (Bartel, 2004) without G:U base pairs. We restricted our results to interactions presenting free energy < -15 kcal/mol. An integrative network was generated using Cytoscape, version 2.7.0 (Shannon *et al.* 2003)(Supplementary).

Results

Description of the library and identification of known miRNA

Here we show the characterization and analysis of small RNA library obtained from activated ovaries of *A. mellifera* honeybee workers deep sequenced in Illumina platform (HiSeq 2000, Life Sciences). The sequencing procedure generated about 49 mi reads (raw data), 63,6% (about 31 mi) mapped in unique regions in *A. mellifera* genome. Reads 19 to 24 nt long represent 47,7% of uniquely mapped reads (Fig. 1). The most recently MiRBase version (version 20) was used for identify microRNA expressed in the worker active ovary library (Supplementary Table S2 and S3). MiRBase is currently composed by 218 honeybee microRNAs. Honeybee worker ovaries express 112 known microRNAs. We elected 19 microRNAs to validate the expression levels by qPCR. The selected microRNAs were the ones highly or poorly expressed in in the library (Fig. 2) and or already known as important to oogenesis and related processes. The microRNAs expression results are presented in the following section.

qPCR analysis of miRNA expression in activate and inactive ovaries

Samples of active and inactive ovaries from queens and workers were used for evaluating microRNA expression. Worker ovaries showing visible signals of activation, presence of oocytes at different stage of development, were classified as active and ovaries whose ovarioles were recognized just as thin filaments were classified as inactive. Fertilized egg-laying queens furnished queen active ovaries and queens inactive ovaries were obtained from virgin queens. All qPCR analyzed microRNA was tested against those samples. We separate the results in 5 groups according to the

number of reads the miRNAs presented in the library. The more expressed microRNAs, mir-306 and mir-317, revealed higher expression in active ovaries, independently of the origin of samples, worker or queen. The microRNAs found in the literature, Mir-9a and 92b, described as involved in honey bee caste differentiation and in the regulation of stem cell proliferation, respectively, gave also the same result, they were more expressed in active ovaries (Fig. 3A). The selected less expressed microRNAs, belonging or not to the top 15 list, identified the inactive ovaries (Fig. 3B). Exception to miR-276 that has very high number of reads in the library (46,500 reads), but more expressed in inactive ovaries of workers and virgin queens. miR-263a is one of the tested microRNA that behave differently among queens and workers. Coincidentally they do not have their function completely determined; miR-276 is highly conserved RNA through the animal kingdom (Campbell *et al.* 2011) and to miR-263 family is suggest role in patterning and in apoptotic process (Nguyen & Frasch 2006; Liu *et al.* 2010). MicroRNAs members of well-known clusters were also tested, the let-7 family, that includes miR-100 and miR-125 and miR-2 family member miR-71, miR-13a and miR-13b. The let-7 family seems to be co-regulated; all members have the same low number of reads in the library (approximately 4,000 reads) and identified inactive ovaries in workers and virgin queens (Fig. 4A). Members of miR-2 family, have different number of reads in the library, 4,000, 54 and 334 reads for miR-2, miR-13a and miR-13b, respectively, and identify the inactive ovaries. But, miR-71 that was predicted targeting *vitellogenin* gene and with 12.000 reads in the library identified the active ovaries in both castes, however significantly only in workers (Fig 4B). miR-2 demonstrated very low variation in all samples tested been because of that used as normalizer in all qPCR reactions (Fig.4B). Other two genes, miR-184 and miR-11 even belonging to the top 15 list have the expression unchanged in all samples tested (Fig.5) and were also used as normalizer together with miR-2.

Integrative network

We recovered 159 predicted or validated 3'UTR sequences from which 67 presented at least one miRNA binding site following the conservative criteria used (see methods). The most microRNA-connected genes were XP_392899.2 (60 kDa heat shock protein [Hsp60], mitochondrial-like), XP_393976.2 (guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1-like) and XP_623199.2 (heat shock protein 70Cb

ortholog), while miR-263a, miR-71, miR-1 presented the greater number of targets (Fig. 6).

Discussion

Although the majority of the worker larval ovarioles have been destroyed during early development in response to the decreasing levels of juvenile hormone (Capella & Hartfelder 1998) due the high expression of juvenile hormone esterase at this developmental stage (Mackert *et al.* 2008), the remaining few units in the ovary (2 to 12 ovarioles, approximately) are completely well formed and capable of produce viable eggs. It has at least two possibilities for workers activate ovaries; they become egg-layer spontaneously as anarchist bees, apparently without special stimuli, due to their genetic background (Chaline *et al.* 2002; Thompson *et al.* 2006). However, anarchist bees are very rare in normal colonies, no more than 1% of the adult workers (Cardoen *et al.* 2011). Other possibility is that they activate ovaries in response to any kind of circumstance meaning absence of the influence of the queen, due overcrowding, queen aging or death. Those few ovarioles contain all elements present in an ovary, a long filament with oocyte precursors, the germ cell line, at the apical end; these cells are dormant under colony regular condition, but assume cell division forming clusters that will follow the oogenesis process producing viable eggs. Many authors reported that in the absence of the queen some behavioral changes occur (Naeger *et al.* 2013). The nurse bees, the more probable member of the colony to be egg layer do not follow the expected age polyethism and will not become foragers. They remain in the colony, their protein profile changes in the hemolymph reflecting the massive protein synthesis in the fat body, which dedicates to the production of huge quantity of vitellogenin, the main protein of the egg and the source of amino acids that sustain embryo development. The ovarioles also change; the first morphological recognizable signal is the swelling of the filaments. This swelling means they are at the beginning of the activation process (Velthuis 1970). Afterwards it could be seen small developing oocytes surrounded by follicle cells with nurse cells on top, the future anterior pole of the embryo. The colony homeostasis was broken. The microRNAs, the regulators of development, cell proliferation and development, depict those physiological changes making recognizable activated and inactive ovaries, either in workers or in virgin or mated queen, throughout their expression profile. For the first time, specific microRNAs was tested intending

characterizing the reproductive state of the ovaries in both, workers and queens, based on a library constructed using worker activate ovaries .

The quantification of specific microRNAs expression in inactive or active worker ovaries and in virgin or fertilized queen demonstrated that they are equally modulated in both castes under the same reproductive conditions. The more expressed among the top 15 microRNAs characterize the activate ovary. The less expressed in the same group characterize the inactive ones. Another observation, members of microRNA clusters tended to have the same expression when they presented the same number of reads in the library. Moreover, the current literature and target prediction suggest strongly that the more expressed microRNAs in the library are bound directly to ovary function and oogenesis. MiR-306 is one case; it is on the top of our library and characterizes the active state of the ovaries. It is described as abundantly expressed in adults of *Bombyx mori* (Liu *et al.* 2009), was found in ovary stem cell line of *Drosophila melanogaster* and also detected regulating specific genes in differentiated spermatocytes in the same species (Eun *et al.* 2013). This microRNA up-regulated in honeybee active ovaries responds rapidly to the absence of the queen (result not shown) indicating that could be one of the signals for ovary activation. *In silico* prediction indicates only one target for miR-306, an alpha-tubulin- chain 1 (XP_3919363) component of microtubules. The microtubules act in a polarized manner in the transportation macromolecules synthesized by the nurse cells to the developing oocyte, been the organization of the cytoskeleton and polarization of this transport crucial for oogenesis (Patrício *et al.* 2011). In bees the bridges of germline cysts, the first differentiate structure in the activate ovariole filament, are reinforced by microtubules and microfilaments of actin (Bilinski & Jaglarz 1999). However, this is not the sole case. The second in our microRNA list, miR-317, even with no function determined in insects, but conserved in *B. mori*, *T. castaneum*, *D. melanogaster* and also in mammals showed the same up-regulated profile, identifying the reproductive state of queens and workers. The only function found was in *Glyptapanteles flavicoxi* where it was described related to immune response (Gundersen-Rindal & Pedroni 2010). Two microRNAs suggested by the literature (Weaver *et al.* 2007) as highly expressed in queen pupae, and also detected in adult workers, miR-9a have also the expression tested been up-regulated in active ovaries in both, queens and workers. Another microRNA that gave the same result and also suggested by the same authors was miR-92b, even though without determined function in insects, it was demonstrated in mammals

controlling G1/S check point in stem cells. We selected and analyzed some less expressed miRNAs from our top 15 list, miR-2, miR- 71, miR-13a, and miR-13b from the cluster 2 and let-7, miR-100, and miR125 from the cluster let-7. Surprisingly those less expressed are up-regulated in worker inactive ovary and ovary of virgin queen. The members of the cluster let-7 have almost the same number of reads in our library and the same behavior, suggesting that they are co-regulated. These microRNAs are related to developmental timing, to ecdysone response and broad complex regulation in *Blattella germanica* metamorphosis, (Rubio *et al.* 2012) nurse bee behavior and *Bombyx mori* ovary (Liu *et al.* 2007) .The tested members of the cluster 2 do not have the same number of reads in the library, miR-71 have the highest number and is more expressed in worker active ovaries. It has an important target, the *vitellogenin* gene that codifies vitellogenin the main protein of the egg and in worker is supposed to be inhibited. Interestingly, it is the unique miRNA of the selected group from the top 15 that target this important gene. The *vitellogenin* gene is crucial for the vitellogenic phase of the oogenesis, been usually more expressed in queens than in workers. This result could reproduce this condition. This microRNA has in *C. elegans* very interesting function interferes in the signaling among germ cell line and nervous system coordinating the organism life span (Boulas & Horvitz 2012). The others members of the cluster either characterize the active state or no difference was detected in the expression, the last corresponds to miR-2. Because of this it was used as normalizer in qPCR reaction, together with miR-184 and miR-11. Although the important function attributed to miR-184, whose loss of function induced severe defects during oogenesis and embryogenesis, having as main effect the complete absence of eggs, its expression is the same in all reproductive conditions. Others microRNAs with interesting function in ovaries but less expressed in the library were miR- 12 that is bound to chorion gene amplification regulating MCM6 in *D. melanogaster* (Osei-Amo *et al.* 2012) crucial function for the fast synthesis and deposition of the external membrane of the egg. The less expressed in the library, miR-263a with no determined function, mir-31a and miR-1, the last targeting mesodermal and muscle related genes (Nguyen & Frasch 2006; Liu *et al.* 2010) all characterize inactive ovaries. Lastly an ancient microRNA was localized with relative high expression in our library, miR-276, present from Tardigrada, sister group of the Arthropods, to mammals (Campbell *et al.* 2011), been the only with high number of reads in the library whose expression characterize inactive ovaries in both castes.

The construction of microRNA library from active ovaries of worker bees, the expression analysis, and the prediction of targets gave us a portrait of the sensitive reproductive equilibrium that maintains both castes in apparent harmony in the colony where they assume at the right time their role in this sophisticated society undertaking or not the reproduction. We are motivated to conclude that up-regulated microRNAs are more expressed in active ovaries and less expressed in inactive, together they create an integrative network where genes are repressed or have the expression liberate allowing the ovaries to a reproductive state or retaining them as non-reproductive.

Acknowledgements

This project was supported by FAPESP 2011/03171-5 to ZLPS. Thanks are also to Daniel Guariz Pinheiro for helping with the informatics analysis and to Luiz R. Aguiar for the work in the apiary providing the bees.

Appendix: Supplementary materials

Legends

Figure 1- Description of the small RNA deep sequenced library of worker bee active ovaries. It was considered reads from 19 nt to 24 nt characterizing the microRNA and reads mapped to honeybee genome version 4.5.

Figure 2 – Categorization of the 15 more expressed microRNA in the honeybee worker active ovaries, mapped to *A. mellifera* genome.

Figure 3 – qPCR analysis of microRNAs in samples obtained from active and inactive ovaries of workers and queens of *A. mellifera*. A- The more abundantly expressed microRNAs, miR-306 and miR-317, are differentially expressed in active ovaries from workers and egg-layer queen. miR-9a and miR-92b indicated in the current literature as involved in caste development are also significantly expressed in active ovaries . B- Less expressed among the top 15 microRNAs and also in the library as a whole. In general these microRNAs are differentially expressed in inactive ovaries, independent of the caste, exception to miR-263a that is equally expressed in virgin and egg-layer

queen. The error bars indicate standard deviation from three technical triplicates and three biological replicates, asterisks indicate statistical significance ($p \leq 0,05$).

Figure 4 - qPCR analysis of members of two clusters of microRNAs in samples from active and inactive ovaries of workers and queens of *A. mellifera*. A- Members of the cluster let-7, miR-100, miR-125 with very number of reads in the library (516, 356 , 159, respectively) are more expressed in inactive ovaries. B- Member of the miR-2 family, miR-71, miR-13a and miR-13b (4000, 12000, 334 and 54 reads respectively) seems to be independently regulated, presenting different number of reads in the library and characterizing differently active and inactive ovaries. The more expressed in the group, miR-71, characterizes active ovaries, probably targeting *vitellogenin* gene, the other members of the cluster either characterize the active ovary or are equally expressed in queens and workers, independently of the reproductive state. The error bars indicate standard deviation from three technical triplicates and three biological replicates, asterisks indicate statistical significance ($p \leq 0,05$).

Figure 5 - qPCR analysis of microRNAs in samples obtained from active and inactive ovaries of workers and queens of *A. mellifera*. These microRNAs, miR- 184 (111600 reads), miR-11 (4900 reads) and miR-3720 (566 reads) are equally expressed in both castes independently of ovary state, because of that they was used as normalizer in all validation testes, exception to miR-3720. Instead this on miR-2 was preferentially used as normalizer. The error bars indicate standard deviation from three technical triplicates and three biological replicates, asterisks indicate statistical significance ($p \leq 0,05$).

Figure 6 - Regulatory network model reconstructed using the validated miRNAs (our data) and proteins differentially expressed in active versus inactive ovaries of worker bees, published by Cardoen et al., 2012. Proteins are marked in grey and microRNAs are marked in white.

Figures

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Supplementary material

Table S1

Table S2

Table S3

miRNA-mRNA predicted network

References

- Bartel D.P. (2004) MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116, 281-97.
- Bartel D.P. (2007) MicroRNAs: Genomics, biogenesis, mechanism, and function (Reprinted from *Cell*, vol 116, pg 281-297, 2004). *Cell* 131, 11-29.
- Bilinski S.M. & Jaglarz M.K. (1999) Organization and possible functions of microtubule cytoskeleton in hymenopteran nurse cells. *Cell Motil Cytoskeleton* 43, 213-20.
- Boulias K. & Horvitz H.R. (2012) The *C. elegans* microRNA mir-71 acts in neurons to promote germline-mediated longevity through regulation of DAF-16/FOXO. *Cell Metab* 15, 439-50.
- Campbell L.I., Rota-Stabelli O., Edgecombe G.D., Marchioro T., Longhorn S.J., Telford M.J., Philippe H., Rebecchi L., Peterson K.J. & Pisani D. (2011) MicroRNAs and phylogenomics resolve the relationships of Tardigrada and suggest that velvet worms are the sister group of Arthropoda. *Proc Natl Acad Sci U S A* 108, 15920-4.
- Capella I.C. & Hartfelder K. (1998) Juvenile hormone effect on DNA synthesis and apoptosis in caste-specific differentiation of the larval honey bee (*Apis mellifera* L.) ovary. *J Insect Physiol* 44, 385-91
- Cardoen D., Wenseleers T., Ernst U.R., Danneels E.L., Laget D., De Graaf D.C., Schoofs L. & Verleyen P. (2011) Genome-wide analysis of alternative reproductive phenotypes in honeybee workers. *Molecular Ecology* 20, 4070-84.
- Chaline N., Ratnieks F.L. & Burke T. (2002) Anarchy in the UK: Detailed genetic analysis of worker reproduction in a naturally occurring British anarchistic honeybee, *Apis mellifera*, colony using DNA microsatellites. *Molecular Ecology* 11, 1795-803.
- Cristino A.S., Tanaka E.D., Rubio M., Piulachs M.D. & Belles X. (2011) Deep sequencing of organ- and stage-specific microRNAs in the evolutionarily basal insect *Blattella germanica* (L.) (Dictyoptera, Blattellidae). *PLoS One* 6, e19350.
- Deloffre L.A., Andrade A., Filipe A.I. & Canario A.V. (2012) Reference genes to quantify gene expression during oogenesis in a teleost fish. *Gene* 506, 69-75.
- Eun S.H., Stoiber P.M., Wright H.J., McMurdie K.E., Choi C.H., Gan Q., Lim C. & Chen X. (2013) MicroRNAs downregulate Bag of marbles to ensure proper terminal differentiation in the *Drosophila* male germline. *Development* 140, 23-30.
- Free J.B. (1987) The Biology of the Honey-Bee - Winston, M.L. *Science* 238, 1591-2.
- Grozinger C.M., Fan Y.L., Hoover S.E.R. & Winston M.L. (2007) Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Molecular Ecology* 16, 4837-48.
- Gundersen-Rindal D.E. & Pedroni M.J. (2010) Larval stage *Lymantria dispar* microRNAs differentially expressed in response to parasitization by *Glyptapanteles flavicoxis* parasitoid. *Arch Virol* 155, 783-7.

Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443, 931-49.

Hoover S.E., Keeling C.I., Winston M.L. & Slessor K.N. (2003) The effect of queen pheromones on worker honey bee ovary development. *Naturwissenschaften* 90, 477-80.

John R., Mark J.S., Paul M.M. & Andreas B. (2009) miRNA-Dependent Translational Repression in the *Drosophila* Ovary. *PLoS One* 4.

Koywiwattrakul P. & Sittipraneed S. (2009) Expression of Vitellogenin and Transferrin in Activated Ovaries of Worker Honey Bees, *Apis mellifera*. *Biochemical Genetics* 47, 19-26.

Legeai F., Rizk G., Walsh T., Edwards O., Gordon K., Lavenier D., Leterme N., Méreau A., Nicolas J., Tagu D. & Jaubert-Possamai S. (2010) Bioinformatic prediction, deep sequencing of microRNAs and expression analysis during phenotypic plasticity in the pea aphid, *Acyrtosiphon pisum*. *BMC Genomics* 11, 281.

Li H. & Durbin R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-60.

Liman M., Wenji W., Conghui L., Haiyang Y., Zhigang W., Xubo W., Jie Q. & Quanqi Z. (2013) Selection of reference genes for reverse transcription quantitative real-time PCR normalization in black rockfish (*Sebastes schlegeli*). *Mar Genomics* 11, 67-73.

Liu S., Gao S., Zhang D., Yin J., Xiang Z. & Xia Q. (2010) MicroRNAs show diverse and dynamic expression patterns in multiple tissues of *Bombyx mori*. *BMC Genomics* 11, 85.

Liu S., Xia Q., Zhao P., Cheng T., Hong K. & Xiang Z. (2007) Characterization and expression patterns of let-7 microRNA in the silkworm (*Bombyx mori*). *Bmc Developmental Biology* 7, 88.

Mackert A., do Nascimento A.M., Bitondi M.M., Hartfelder K. & Simoes Z.L. (2008) Identification of a juvenile hormone esterase-like gene in the honey bee, *Apis mellifera* L.--expression analysis and functional assays. *Comp Biochem Physiol B Biochem Mol Biol* 150, 33-44.

Maisonnasse A., Lenoir J.C., Beslay D., Crauser D. & Le Conte Y. (2010) E-beta-ocimene, a volatile brood pheromone involved in social regulation in the honey bee colony (*Apis mellifera*). *PLoS One* 5, e13531.

Marco A., Ninova M. & Griffiths-Jones S. (2013) Multiple products from microRNA transcripts. *Biochem Soc Trans* 41, 850-4.

Naeger N.L., Peso M., Even N., Barron A.B. & Robinson G.E. (2013) Altruistic behavior by egg-laying worker honeybees. *Curr Biol* 23, 1574-8.

Nguyen H.T. & Frasch M. (2006) MicroRNAs in muscle differentiation: lessons from *Drosophila* and beyond. *Curr Opin Genet Dev* 16, 533-9.

Nolo R., Morrison C.M., Tao C., Zhang X. & Halder G. (2006) The bantam microRNA is a target of the hippo tumor-suppressor pathway. *Curr Biol* 16, 1895-904.

Nunes F.M., Ihle K.E., Mutti N.S., Simoes Z.L. & Amdam G.V. (2013) The gene vitellogenin affects microRNA regulation in honey bee (*Apis mellifera*) fat body and brain. *J Exp Biol* 216, 3724-32.

Oldroyd B.P., Wossler T.C. & Ratnieks F.L.W. (2001) Regulation of ovary activation in worker honey-bees (*Apis mellifera*): larval signal production and adult response thresholds differ between anarchistic and wild-type bees. *Behavioral Ecology and Sociobiology* 50, 366-70.

Osei-Amo S., Hussain M., O'Neill S.L. & Asgari S. (2012) Wolbachia-induced aae-miR-12 miRNA negatively regulates the expression of MCT1 and MCM6 genes in Wolbachia-infected mosquito cell line. *PLoS One* 7, e50049.

Oxley P.R., Thompson G.J. & Oldroyd B.P. (2008) Four quantitative trait loci that influence worker sterility in the honeybee (*Apis mellifera*). *Genetics* 179, 1337-43.

Patricio K., da Cruz-Landim C. & Machado-Santelli G.M. (2011) Cytoskeletal organization of bee ovarian follicles during oogenesis. *Micron* 42, 55-9.

Poulton J.S., Huang Y.C., Smith L., Sun J., Leake N., Schleede J., Stevens L.M. & Deng W.M. (2011) The microRNA pathway regulates the temporal pattern of Notch signaling in *Drosophila* follicle cells. *Development* 138, 1737-45.

Rehmsmeier M., Steffen P., Hochsmann M. & Giegerich R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA* 10, 1507-17.

Rubio M., de Horna A. & Belles X. (2012) MicroRNAs in metamorphic and non-metamorphic transitions in hemimetabolous insect metamorphosis. *BMC Genomics* 13, 386.

Shannon P., Markiel A., Ozier O., Baliga N.S., Wang J.T., Ramage D., Amin N., Schwikowski B. & Ideker T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498-504.

Thompson G.J., Kucharski R., Maleszka R. & Oldroyd B.P. (2006) Towards a molecular definition of worker sterility: differential gene expression and reproductive plasticity in honey bees. *Insect Molecular Biology* 15, 637-44.

Thompson G.J., Kucharski R., Maleszka R. & Oldroyd B.P. (2008) Genome-wide analysis of genes related to ovary activation in worker honey bees. *Insect Molecular Biology* 17, 657-65.

Thompson G.J., Yockey H., Lim J. & Oldroyd B.P. (2007) Experimental manipulation of ovary activation and gene expression in honey bee (*Apis mellifera*) queens and workers: Testing hypotheses of reproductive regulation. *Journal of Experimental Zoology Part a-Ecological Genetics and Physiology* 307A, 600-10.

Velthuis H.H. (1970) Ovarian Development in *Apis-Mellifera* Worker Bees. *Entomologia Experimentalis Et Applicata* 13, 377-&.

Vergoz V., Lim J. & Oldroyd B.P. (2012) Biogenic amine receptor gene expression in the ovarian tissue of the honey bee *Apis mellifera*. *Insect Molecular Biology* 21, 21-9.

Weaver D.B., Anzola J.M., Evans J.D., Reid J.G., Reese J.T., Childs K.L., Zdobnov E.M., Samanta M.P., Miller J. & Elisk C.G. (2007) Computational and transcriptional evidence for microRNAs in the honey bee genome. *Genome Biol* 8, R97.

Zhang Y., Zhou X., Ge X., Jiang J., Li M., Jia S., Yang X., Kan Y., Miao X., Zhao G., Li F. & Huang Y. (2009) Insect-Specific microRNA Involved in the Development of the Silkworm *Bombyx mori*. *PLoS One* 4, e4677.