



UNIVERSIDADE FEDERAL DO PAMPA
CAMPUS SÃO GABRIEL

CLARISSA DE LIMA BARCELOS

**DETERMINAÇÃO DE POTENCIAIS PROTEÍNAS ALVO DE PROTEASES 3C DE
VÍRUS QUE INFECTAM ABELHAS NO CONJUNTO DE PROTEÍNAS ANOTADAS
DE *APIS MELLIFERA***

SÃO GABRIEL

2014

CLARISSA DE LIMA BARCELOS

**DETERMINAÇÃO DE POTENCIAIS PROTEÍNAS ALVO DE PROTEASES 3C DE
VÍRUS QUE INFECTAM ABELHAS NO CONJUNTO DE PROTEÍNAS ANOTADAS
DE *APIS MELLIFERA***

Trabalho de conclusão de curso III, como pré-requisito parcial para obtenção do grau de Bacharel em Biotecnologia na Universidade Federal do Pampa.

Orientador: Andrés Delgado Cañedo

Co-orientadora: Raíssa Ochôa Golin

SÃO GABRIEL

2014

CLARISSA DE LIMA BARCELOS

**DETERMINAÇÃO DE POTENCIAIS PROTEÍNAS ALVO DE PROTEASES 3C DE
VÍRUS QUE INFECTAM ABELHAS NO CONJUNTO DE PROTEÍNAS ANOTADAS
DE *APIS MELLIFERA***

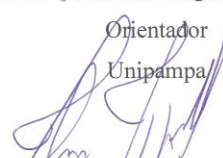
Trabalho de conclusão de curso III, como
pré-requisito parcial para obtenção do
grau de Bacharel em Biotecnologia na
Universidade Federal do Pampa.

Trabalho de conclusão de curso III, defendido e aprovado em: 22 de agosto de 2014.


Banca examinadora:



Prof. Dr. Andrés Delgado Cañedo

Orientador
Unipampa


Prof. Dr. Juliãno Tomazzoni Boldo
Unipampa



Prof. Dr. Paulo Marcos Pinto
Unipampa

Dedico este trabalho aos meus amados pais, Carlos Adão Barcelos e Paula Marta de Lima Barcelos, e às minhas amadas irmãs, Carla de Lima Barcelos de Vasconcellos e Quélen de Lima Barcelos.

AGRADECIMENTOS

Agradeço primeiramente ao meu orientador/chefe/pai, Andrés, por ter me aceitado como sua IC e por acreditar em mim, sempre. Pela enorme paciência que teve comigo durante todos esses anos em que trabalhamos juntos, por ter me ensinado e me incentivado em todos os momentos, por ser o chefe que cobrava sempre o melhor de mim, por ser meu pai/amigo quando eu precisava de atenção e conselhos sobre minha vida pessoal. E principalmente por ter me dado a oportunidade de trabalhar com ele, onde eu pude crescer não só no meio acadêmico mas também como pessoa. Como eu sempre digo: Se eu sair da graduação com no mínimo 1% do que meu chefe foi pra mim, sei que terei sucesso em toda minha vida!

Aos meus pais e irmãs, por me darem muita força desde o início da minha graduação e principalmente nos semestres finais, por estarem do meu lado, me apoiando em todas as minhas decisões e por me ajudarem a enfrentar os momentos ruins. Foram a fé e amor incondicional que vocês tem em mim que me fizeram e me farão seguir em frente, neste e em todos os momentos da minha vida.

Aos meus professores que souberam exercer sua profissão de ensinar e desenvolver o pensamento crítico em seus alunos.

À Gringa, pela enorme amizade, atenção e carinho, por aguentar as minhas crises, por me apoiar e por sempre ter uma palavra doce pra me trazer calma e principalmente, por ser um exemplo pela garra de como conduz e enfrenta sua vida.

Ao Leandro Lemos, por conversar e me fazer rir todos os dias e pelos 07 papers que ele baixou para eu escrever o TCC.

Ao Maikel, ao Tiago, a Vivi e a Perin, pela amizade, carinho, cumplicidade e cuidados que só irmãos podem dar.

E, ao Leo, por ser meu companheiro, pelo incentivo, por me fazer ver e crer todos os dias o quanto eu era e sou capaz de realizar qualquer coisa em minha vida. Pela amizade, cumplicidade, atenção, pelo enorme amor e carinho e pela esperança viva de um futuro da maneira que sempre desejamos.

RESUMO

A principal característica das abelhas da espécie *Apis mellifera* L., é sua efetiva capacidade de polinização e de coleta de néctar floral, assim, desempenham papel de extrema importância na manutenção da biodiversidade e no ecossistema do planeta, tendo em vista o seu trabalho na polinização de inúmeras espécies florais. Com o alto índice de polinização de diferentes cultivares na agricultura global, é de extrema importância a detecção e prevenção dos inúmeros agentes patológicos que acometem as abelhas. Um dos principais focos no estudo da sanidade apícola é o fenômeno chamado Desordem do Colapso das Colônias, caracterizado pela perda rápida e inexplicável da população adulta da colmeia, não sendo encontradas abelhas mortas dentro do ninho nem ao redor das colmeias. Logo, o motivo desta desordem está possivelmente ligado a muitas variáveis como, o uso de pesticidas, carga de patógenos associados, especialmente vírus, bactérias, fungos e ácaros. Dentre os patógenos virais, destacam-se os vírus pertencentes à superfamília Picorna-like que conserva um conjunto de genes que codificam a RNA polimerase dependente de RNA (RdRp), protease quimiotripsina-like (3C-Pro), a superfamília da helicase 3 (S3H) e uma proteína de ligação ao genoma (chamada proteína viral - VPg). A maioria das clivagens na poliproteína viral são realizadas pela protease 3C viral. Além do seu papel essencial no processamento e maturação da poliproteína viral, a protease 3C viral participa da regulação das funções das células hospedeiras através da degradação proteolítica evitando e bloqueando os mecanismos antivirais, possibilitando assim a instalação e replicação do vírus na célula alvo. O objetivo deste trabalho foi identificar prováveis sítios alvo da protease 3C dos vírus DWV, SBPV, e vírus do complexo KIA no conjunto de proteínas de *A. mellifera* e identificar grupos funcionais superrepresentados dentro destas proteínas alvo. Para isto, identificamos na literatura os sítios de clivagem de cada protease estudada, realizamos a busca de sequências alvos da protease 3C no conjunto proteico de *Apis mellifera* e analisamos as sequências potencialmente clivadas pela protease através da ferramenta Gene Ontology. Os resultados obtidos para a protease do vírus DWV mostraram que as funções proteicas afetadas agregaram-se em três grandes grupos: ligação a ácidos nucleicos, proteólise e desenvolvimento neuronal, destacando a presença de genes como: Apidaecina, Dachous, Dicer-1 e Dead-box1. Nas análises para o SBPV, a busca resultou na formação de 33 agrupamentos onde destacaram-se genes como: Roundabout e Shibire. Por fim, as análises para o complexo KIA resultaram em 31 grupamentos onde destacaram-se genes como:

Brahma, Blue Cheese e Aurora Borealis. A correlação entre a função de várias das proteínas potencialmente clivadas com os sintomas da infecção dos vírus estudados sugerem que estas proteases clivam estas proteínas nas abelhas infectadas para facilitar o processo de infecção a partir da clivagem de proteínas envolvidas em vias de sinalização antiviral e, ao mesmo tempo, clivando outras proteínas que terminam por desencadear os sintomas de cada infecção viral. Estes dados poderão auxiliar futuramente no desenvolvimento de mecanismos de prevenção ou tratamento da infecção viral em abelhas, como também na busca de mutações nas sequências que codificam estes sítios de clivagem em abelhas sem sintomas de infecção.

Palavras-chave: *Apis mellifera* - vírus - Protease 3C - proteoma

ABSTRACT

The bees of the specie *Apis mellifera L.* are known/characterized by their/the high capacity for harvesting and storing pollen and nectar from flowers, they produce honey and play a key role on the planet ecosystem as a floral biodiversity maintainer by promoting polonization of numerous floral species. Due to its fundamental function on global agriculture, the detection and prevention of possible pathological agents with threats the bee health is an crucial task to beekeepers researchers. The major subject in bee research is the unidentified suddenly vanishment of bee population from the colony. This phenomenon is known as Colony Collapse Disorder (CCD) and it is described as an unexplained abruptly disappearance of work bees from the colony. The possible reason for this disorder is linked to various factors such as use of chemical pesticide, mites, bacterias, fungis and viruses. Under the pathogenic viruses which affect bees, the Picornavirus-like superfamily has a set of conserved genes which codify: a RNA-dependent RNA polymerase RdRp, a helicase-3 S3H superfamily, a viral genome protein binder VPg and a protease chymotrypsin-like 3C-Pro. The 3C-Pro has a fundamental role not only in the processing and maturation of the viral polyprotein by cleaving it but also performs metabolism regulation on the host cell by proteolytic degradation avoiding and blocking the antiviral mechanisms, consequently, enabling the virus insertion and replication on the host cell. The main objective of this work was to identify possible targets of viral protease 3C possible cleavage sites from DWV, SBPV, and the KIA complex and aim for the functions of those proteins in the proteome of *Apis mellifera*. The potential 3C protease targets sequence sites in *Apis mellifera* proteome were inquired and analyzed through the online database Gene Ontology. After further observation on DWV 3C protease targets, we observed three major group with altered functions: DNA-Binding proteins, proteolysis and neural development. The analysis detected the presence of the genes Apidaecina, Dachsous, Dicer-1 and Dead-box1 which functions play fundamental role in viral infection. The genes Roundabout and Shibire were highlighted in the query for SBPV resulting in a clusterization of 33 groups for affected proteins. Lastly, the results for the KIA complex clustered 31 groups highlighting the genes Brahma, Blue Cheese and Aurora Borealis. The correlation between the function of the potential cleaved proteins with the studied symptoms of the viral infection suggest that those proteases cleave proteins related to the antiviral cell signaling pathways in honey bees and this process would be performed in order to facilitate the infection process and, consequently, initiating the

symptoms of each viral infection. These data would be useful on the development of prevention mechanisms or even treatments for viral infection in honeybees as well as the search for mutations in the sequences which encode the cleavage site in nonsymptomatic honey bees.

Keywords: *Apis mellifera* - Viruses - 3C Protease – Proteome.

SUMÁRIO

1 INTRODUÇÃO.....	10
1.1 Abelhas.....	10
1.2 Sanidade apícola e CCD (<i>Colony Collapse Disorder</i>).....	12
1.3 Vírus das abelhas.....	13
1.4 Protease 3C.....	14
2 OBJETIVOS.....	17
2.1 Objetivo geral.....	17
2.1 Objetivo específico.....	17
3 METODOLOGIA.....	17
3.1 Identificação dos sítios.....	17
3.2 Busca das sequencias alvos da protease 3C no proteoma de <i>Apis mellifera</i>.....	18
3.3 Análise das sequencias potencialmente clivadas pela protease 3C com a ferramenta Gene Ontology.....	18
4 RESULTADOS E DISCUSSÃO.....	20
4.1 DWV.....	20
4.2 SBPV.....	21
4.3 Complexo KIA.....	21
5 CONSIDERAÇÕES FINAIS.....	23
REFERENCIAS BIBLIOGRÁFICAS.....	24
ANEXO.....	28

1. INTRODUÇÃO

1.1 Abelhas

As abelhas, da espécie *Apis mellifera* L., se caracterizam pela sua alta capacidade de coletar e armazenar o néctar das flores e com isso produzir mel para suprir suas necessidades energéticas. Elas desempenham um importante papel na manutenção da biodiversidade floral, bem como no ecossistema do planeta, principalmente por promoverem a polinização de inúmeras espécies florais. Além disso, ajudam na manutenção e sustentabilidade de diversos hectares de plantações, sendo de grande importância para a agricultura global (Francis, 2012). As abelhas originaram-se de vespas que deixaram seus hábitos predatórios e passaram ao aprisionamento em seus ninhos. Os primeiros indivíduos provavelmente apareceram concomitantemente ao surgimento das plantas angiospermas, durante a Era terciária do paleocontinente Gondwana, e a coevolução destes dois grupos está intimamente ligada à diversidade de formas e cores das flores, e também à produção de néctar e pólen, ao passo que as abelhas são atraídas por estas características e oferta de alimento e, em troca, agem como um mecanismo de transferência de pólen entre as plantas. Assim, provavelmente evoluíram seu aparato bucal com a finalidade de propiciar a ingestão de néctar para a sua nutrição e para alimentar suas crias. A característica morfológica que mais difere as abelhas das vespas é o desenvolvimento da corbícula, ou cesta de pólen, em suas pernas traseiras, o que lhes permite a coleta do pólen para ser levado até a colmeia onde será armazenado (Winston, 2003).

A população de cada colmeia é dividida em três castas: rainha, zangão e operárias, e cada colmeia pode ter aproximadamente 60.000 operárias, 400 zangões e 1 rainha. A determinação destas castas deve-se à fertilização ou não fertilização dos ovos. Quando fertilizados, desenvolvem-se abelhas operárias e rainha, sendo que a diferenciação entre as duas se dá, principalmente pelo tipo de alimentação e pela quantidade de alimento que recebem. As operárias recebem alimentação leve e controlada, enquanto a rainha recebe alimentação mais rica em nutrientes e geleia real. Estudos mostram que a composição química do alimento que as operárias e rainha recebem, difere não somente nos percentuais de açúcares como, frutose e glucose, como também na presença de pequenas porcentagens de levedura (Kaftanoglu *et al.*, 2011). Dos ovos não fertilizados, as larvas crescem e tornam-se zangões. O desenvolvimento de todas as abelhas passa pelos estágios de ovo, larva, pupa e adulto. Na fase larval, cada indivíduo fica alojado em um alvéolo aberto onde recebem a alimentação. Posteriormente, quando atingem a fase de pupa, o alvéolo é fechado para a

metamorfose até o estágio adulto. Estes estágios até a fase adulta, podem durar aproximadamente, 21 dias, 16 dias e 24 dias para abelhas operárias, rainha e zangões, respectivamente (Winston, 2003).

Além dos benefícios da polinização onde as abelhas promovem a biodiversidade floral e manutenção do ecossistema, pode-se desenvolver a atividade de apicultura, acarretando em fatores positivos também no âmbito social e econômico do país, pois gera empregos e fluxo de renda na agricultura e auxilia no desenvolvimento do homem nas atividades rurais.

No Brasil, a abelha *Apis mellifera* é um híbrido entre as abelhas europeias (*Apis mellifera mellifera*, *Apis mellifera lingustica*, *Apis mellifera caucasica* e *Apis mellifera carnica*) com a abelha africana, *Apis mellifera scutellata*. Desde a década de 70 a apicultura brasileira vem se adaptando às novas características desse poli-híbrido. Com a criação da Confederação Brasileira de Apicultura, em 1970, foram debatidas as dificuldades de adaptação dos apicultores no manejo dessa nova raça, devido a agressividade dessas abelhas, característica herdada da raça *A. scutellata*, e então, a terminologia “abelha africanizada” passou a ser usada para referir essa espécie. Desde então, muitas metodologias foram desenvolvidas com a finalidade de ajudar os apicultores com o manejo das abelhas africanizadas. Este foi um passo muito importante no cenário apícola brasileiro, pois as características dessa nova raça, como a sua alta adaptação aos diversos climas e regiões do nosso país, resistência e maior tolerância a diferentes doenças, e a alta capacidade de enxamear, resultou no aumento da produtividade da colônia. Assim, as abelhas africanizadas fornecem excelentes condições para o avanço e crescimento do mercado apícola brasileiro (Gonçalves, 1974).

Atualmente, o Brasil está na 11ª posição do ranking mundial de países exportadores de mel, e estima-se que cerca de 350 mil pessoas vivam da renda da apicultura. No momento atual, esta atividade já gerou, até o período de junho de 2014 uma receita de mais de 55 milhões de dólares. Dados importantes que evidenciam que a receita gerada pelo primeiro semestre deste ano já supera o montante anual de 2013 totalizado em 54 milhões de dólares (Abemel, 2014). Essa alta produtividade é reflexo tanto da vasta biodiversidade da flora brasileira que, possibilita a obtenção de méis de diversas floradas, durante todos os meses do ano, com cores, aromas e sabores únicos, como também, da alta difusão da apicultura em todas as regiões do país (Sabbag & Nicodemo, 2011).

Com o grande avanço na apicultura e na produção de mel, também foi preciso aumentar e melhorar os cuidados, não somente na manutenção das colônias, mas também nos processos subseqüentes da produção. Há uma constante busca de novas técnicas de manejo

dos apiários, bem como inúmeras pesquisas sendo desenvolvidas com o objetivo de melhorar a sanidade tanto na criação das abelhas como na produção do mel. Toda a atenção quanto à saúde dos apiários, se deve à vasta lista de agentes patógenos à abelhas como: ácaros, fungos, bactérias, protozoários, entre outros agentes, que causam inúmeras enfermidades às colônias, tendo em vista a grande densidade populacional de cada colônia e conseqüentemente a alta taxa de transmissão e disseminação dessas doenças que terminam conseqüentemente acarretando na perda destas populações. Somado a isto, tem se observado nos últimos anos, um declínio na polinização devido à redução na diversidade de abelhas, causa que está diretamente ligada ao avanço do desmatamento, ao uso intenso de pesticidas em culturas agrícolas, e na introdução de espécies que competem principalmente pelos recursos florais com as abelhas nativas (Rocha & Alencar, 2012).

1.2 Sanidade apícola e CCD (*Colony Collapse Disorder*)

Com o alto índice de polinização de diferentes cultivares na agricultura global, é de extrema importância a detecção e prevenção dos inúmeros agentes patológicos que acometem as abelhas. Os organismos que podem causar danos às abelhas tem a capacidade de afetar os indivíduos da colmeia em qualquer etapa das fases de desenvolvimento e, na maioria das vezes, causando grandes quedas na produtividade pela elevação as taxas de mortalidade.

Um dos principais focos no estudo da sanidade apícola é o fenômeno chamado Desordem do Colapso das Colônias (conhecido pela sigla CCD, do inglês, Colony Collapse Disorder), caracterizado pela perda rápida e inexplicável da população adulta da colmeia, não sendo encontradas abelhas mortas dentro do ninho nem ao redor das colmeias, e por manter-se intacta quanto ao alimento e crias. No inverno de 2006, na costa leste dos Estados Unidos da América, foi observada uma mortandade em larga escala de abelhas da espécie *Apis mellifera* L., este cenário se repetiu nos invernos dos anos de 2007 e 2008 (Gifford, 2011). Fato igualmente descrito nesse período por apicultores em diversos países europeus. Entretanto, perdas em larga escala de colônias não é um fato novo para a apicultura, relatos evidenciam que desde 1869 houve pelo menos dezoito episódios documentados em todo o mundo (Vanengelsdorp *et al.*, 2009).

As causas do CCD ainda não são específicas e são intensamente estudadas pela comunidade científica. Em seus estudos Cox-Foster *et al.* (2007) constataram que agentes patógenos como vírus, bactérias e fungos estavam presentes na grande maioria de amostras de *A. mellifera* onde foram relatados casos de CCD. Logo, o motivo desta desordem está

possivelmente ligado a muitas variáveis como, o uso de pesticidas, carga de patógenos associados, especialmente vírus, bactérias, fungos e ácaros, fisiologia da abelha adulta, deficiência nutricional, mudanças climáticas, apicultura migratória, estresse imunológico, entre outros (Francis, 2012; Vanengelsdorp *et al.*, 2009).

1.3 Virus das abelhas

Entre os possíveis agentes ligados ao CCD, estão os causadores de infecções virais. Os vírus são parasitas celulares obrigatórios, pois utilizam praticamente todo o maquinário da célula hospedeira para se replicar, assim que se replicam conseguem infectar rapidamente toda a colônia, independentemente da fase de desenvolvimento que a abelha se encontra, devido ao contato físico dado pela alta densidade populacional das colmeias (Gifford, 2011; Francis, 2012).

Doenças virais que ameaçam a saúde das abelhas apresentam grande ameaça à apicultura mundial, portanto, cada vez mais são desenvolvidas estratégias para a identificação, monitoramento e estudo desses patógenos visando prevenir e conter epidemias. Cerca de vinte e quatro espécies virais já foram identificadas e caracterizadas em *Apis mellifera* L. (Miranda *et al.*, 2013). Os vírus mais comumente observados e que se tem maior conhecimento são os vírus de RNA de fita simples (Genersch, 2010).

A maioria dos agentes virais causam sintomas visíveis de detecção em seu hospedeiro quando em níveis elevados de infecção. Nas abelhas, a maioria dos vírus persiste naturalmente sem causar sintomas a população afetada, entretando, a presença de sintomas visíveis é fator determinante para o diagnóstico fácil e robusto das doenças nos apiários. As desvantagens destes diagnósticos estão relacionadas ao fato de muitos vírus não apresentarem sintomas visíveis; diferentes vírus podem causar os mesmos sintomas e até mesmo o fato de alguns vírus não apresentarem sintomas durante todo o ciclo de vida da abelha. A transmissão viral apresenta diferentes características biológicas, podendo ocorrer por vias: horizontal (quando passadas via agente externo ou contato individual) e vertical (quando passada dos pais à prole); associados a outros organismos como fungos e ácaros; aos estágios de vida do hospedeiro e também associados as estações do ano (Miranda *et al.*, 2013).

Dentre estes, destacam-se os vírus *Deformed wing virus* (DWV), *Slow bee paralysis virus* (SBPV), pertencentes à família Iflaviridae e os vírus *Kashmir bee virus* (KBV), *Israeli acute paralysis virus* (IAPV) e *Acute bee paralysis virus* (ABPV), pertencentes à família Dicistroviridae (Baker & Schroeder, 2008).

O *Deformed wing virus*, vírus deformador de asas, pode também ser encontrado em abelhas assintomáticas. Estudos revelam que o DWV se comporta de maneira diferente em abelhas sintomáticas e assintomáticas. Nas que apresentam sintomas visíveis o vírus está totalmente presente no corpo do inseto: cabeça, tronco e abdômen. Entretanto, nas abelhas sem sintomas de asas deformadas, o vírus foi identificado nas porções do tórax e abdômen, mas nunca na cabeça (Yue & Genersch, 2005). Somando-se a isto, pesquisas mostram que a replicação desse vírus aumenta nos meses compreendidos entre a primavera e o outono, estabelecendo-se tanto em abelhas adultas quanto em pupas. O vírus DWV também infecta e se replica em ácaros parasitas, como o *Varroa destructor*, implicando ser este um possível vetor de transmissão do DWV as abelhas (Boncristiani *et al.*, 2009; Tentcheva *et al.*, 2004).

O *Slow bee paralysis virus*, descoberto no ano de 1974 na Inglaterra, é largamente encontrado como patógeno de *Apis mellifera*, levando a um quadro de paralisia dos dois pares de patas anteriores das abelhas, por volta de 10 dias após o seu contágio. Por ser também identificado em larvas e pupas, além das abelhas adultas, sua dispersão está intimamente associada ao ácaro *Varroa destructor* tornando uma doença agressiva e elevando as taxas de mortalidade da população da colméia (Miranda *et al.*, 2010a).

Os vírus *Kashmir bee virus*, *Acute paralysis virus* e *Israeli paralysis virus*, são analisados como um grupamento denominado Complexo KIA. Amplamente distribuídos pelo mundo e fortemente associados ao fenômeno do CCD, suas patologias em níveis baixos de contaminação não demonstram sintomas visuais nas colônias afetadas porém, em altas taxas de contaminação por ABPV ou IAPV, as abelhas adultas infectadas morrem poucos dias após a infecção, e em fase larval o patógeno compromete o desenvolvimento da larva acarretando a altos níveis de mortalidade na colônia. Considerando a alta virulência destes vírus, e por estarem igualmente associados ao *Varroa destructor*, esse ácaro se torna um poderoso vetor na transmissão deste agente viral (Miranda *et al.*, 2013; Boncristiani *et al.*, 2009).

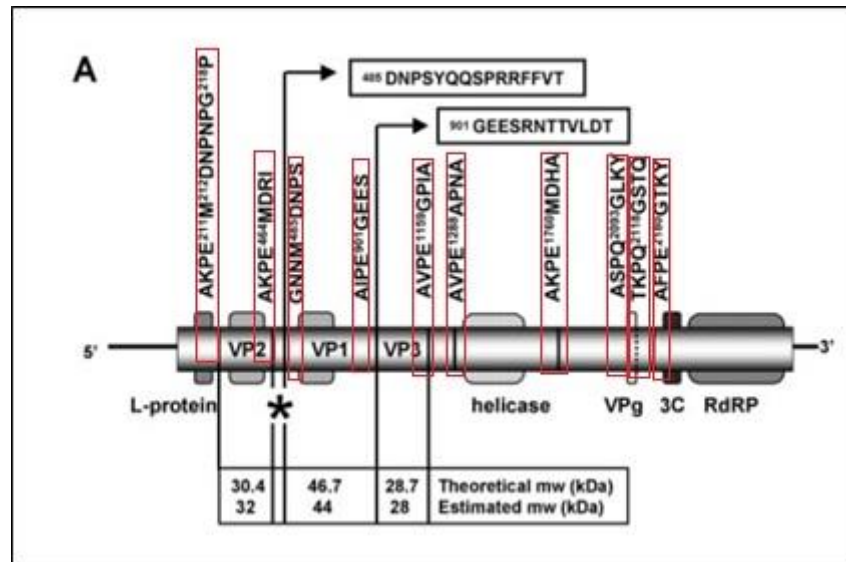
1.4 Protease 3C

A superfamília dos vírus picorna-like é em particular caracterizada pela conservação parcial de um conjunto de genes que codificam a RNA polimerase dependente de RNA (RdRp), protease quimiotripsina-like (3C-Pro), a superfamília da helicase 3 (S3H) e uma proteína de ligação ao genoma (chamada proteína viral - VPg) (Koonin *et al.*, 2008). A maioria das clivagens na poliproteína viral são realizadas pela protease 3C viral, que é uma enzima *cisteína-like* (Arnold *et al.*, 1987). Além do seu papel essencial no processamento e

maturação da poliproteína viral, a protease 3C viral participa da regulação das funções das células hospedeiras através da degradação proteolítica evitando e bloqueando os mecanismos antivirais, possibilitando assim a instalação e replicação do vírus na célula alvo, processo bem descrito para Rhinovirus (Porter, 1993; Amineva *et al.*, 2004; Wang & Chen, 2007).

Dentre os vírus que infectam as abelhas, DWV, VDV, Kakugo, SBPV, IAPV, KBV e ABPV tem seus sítios de clivagem mapeados. Lanzi *et al.* (2006), em suas pesquisas, identificaram os sítios de clivagem da protease 3C no genoma do vírus DWV, conforme Figura 1.

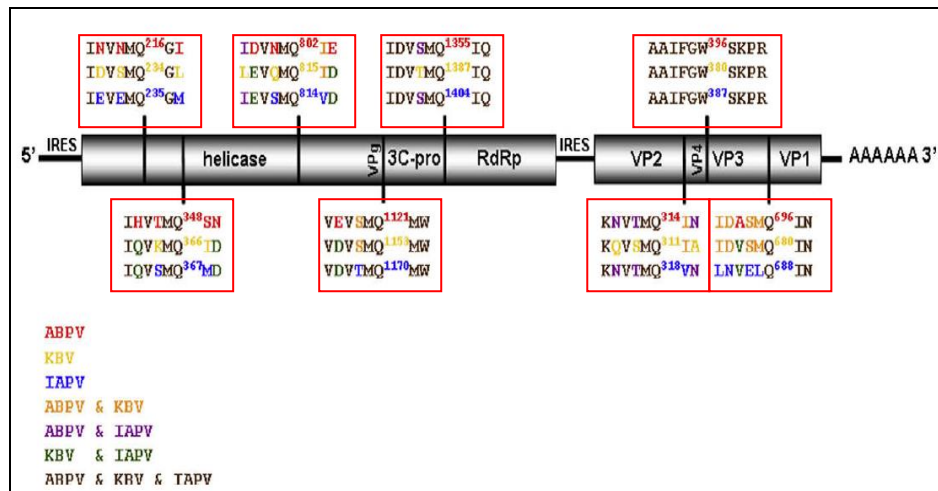
Figura 1 - Mapa demonstrando os sítios de clivagem pela 3C protease no genoma do DWV



Fonte: Adaptado de Lanzi *et al.* (2006).

Em estudos posteriores, Miranda *et al.* (2010a) identificaram os sítios de clivagem da protease 3C no genoma dos vírus do complexo KIA, conforme Figura 2.

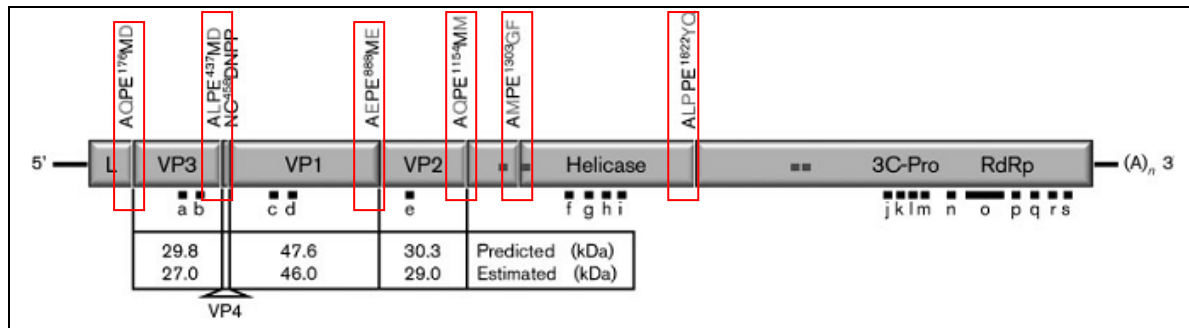
Figura 2 - Esquema demonstrando a organização do genoma do complexo KIA, indicando os sítios de clivagem da protease 3C



Fonte: Adaptado de Miranda *et al.*, (2010a).

E, em pesquisas similares, Miranda *et al.* (2010b) identificaram os sítios de clivagem pela protease 3C no genoma do Slow bee paralysis virus, esquematizado na Figura 3.

Figura 3 - Mapa genômico do vírus SBPV, evidenciando os sítio de clivagem da protease 3C



Fonte: Adaptado de Miranda *et al.*, (2010b).

A identificação dos sítios de clivagem da protease 3C nos diferentes genomas virais, possibilita o entendimento da biologia viral a partir da clivagem da sua poliproteína e, por outro lado, permitiria a identificação e o conhecimento funcional de proteínas do hospedeiro que são alvo de clivagem potenciais pelas proteases destes vírus, influenciando o mecanismo de infecção da célula hospedeira. Neste trabalho, foi analisado o proteoma de *Apis mellifera* na procura de alvos potenciais das proteases 3C dos vírus *Deformed wing virus*, *Slow bee paralysis virus*, e os vírus do complexo KIA (*Kashmir bee virus*, *Israeli acute paralysis virus* e *Acute bee paralysis virus*).

2. OBJETIVOS

2.1 Objetivo geral

Identificar proteínas de *A. mellifera* potencialmente clivadas pelos vírus que as infectam.

2.2 Objetivos específicos

- Identificar as proteínas de *A. mellifera* potencialmente clivadas pela protease 3C do vírus DWV;
- Identificar as proteínas de *A. mellifera* potencialmente clivadas pela protease 3C do vírus SBPV;
- Identificar as proteínas de *A. mellifera* potencialmente clivadas pela protease 3C do complexo viral KIA;
- Identificar grupos funcionais entre as proteínas potencialmente clivadas pelas proteinases 3C dos diferentes vírus.

3. METODOLOGIA

3.1 Identificação dos sítios

Para a identificação dos sítios de clivagem reconhecidos pela protease 3C do *Deformed wing virus* (DWV), foi seguido o trabalho de Lanzi *et al.* (2006). Dos sítios identificados pelo autor, utilizamos os sítios: AKPE, AIPE, AVPE e AFPE.

Para a identificação dos sítios de clivagem reconhecidos pela protease 3C do *Slow bee paralysis virus* (SBPV) e dos sítios de clivagem reconhecidos pela protease 3C dos vírus do complexo KIA, foram usados os trabalho de Miranda *et al.* (2010a;b), respectivamente. Para o SBPV, dos sítios identificados pelo autor, utilizamos os sítios: AQPE, ALPE, AEPE, AMPE e ALPPE. Por fim, para o complexo KIA, dos sítios identificados pelo autor, utilizamos os sítios: VLQM, VVQM, VIQM, VTQM, VKQM, VFQM, VMQM e VNQM.

3.2 Busca de sequências alvos da protease 3C no proteoma de *Apis mellifera*

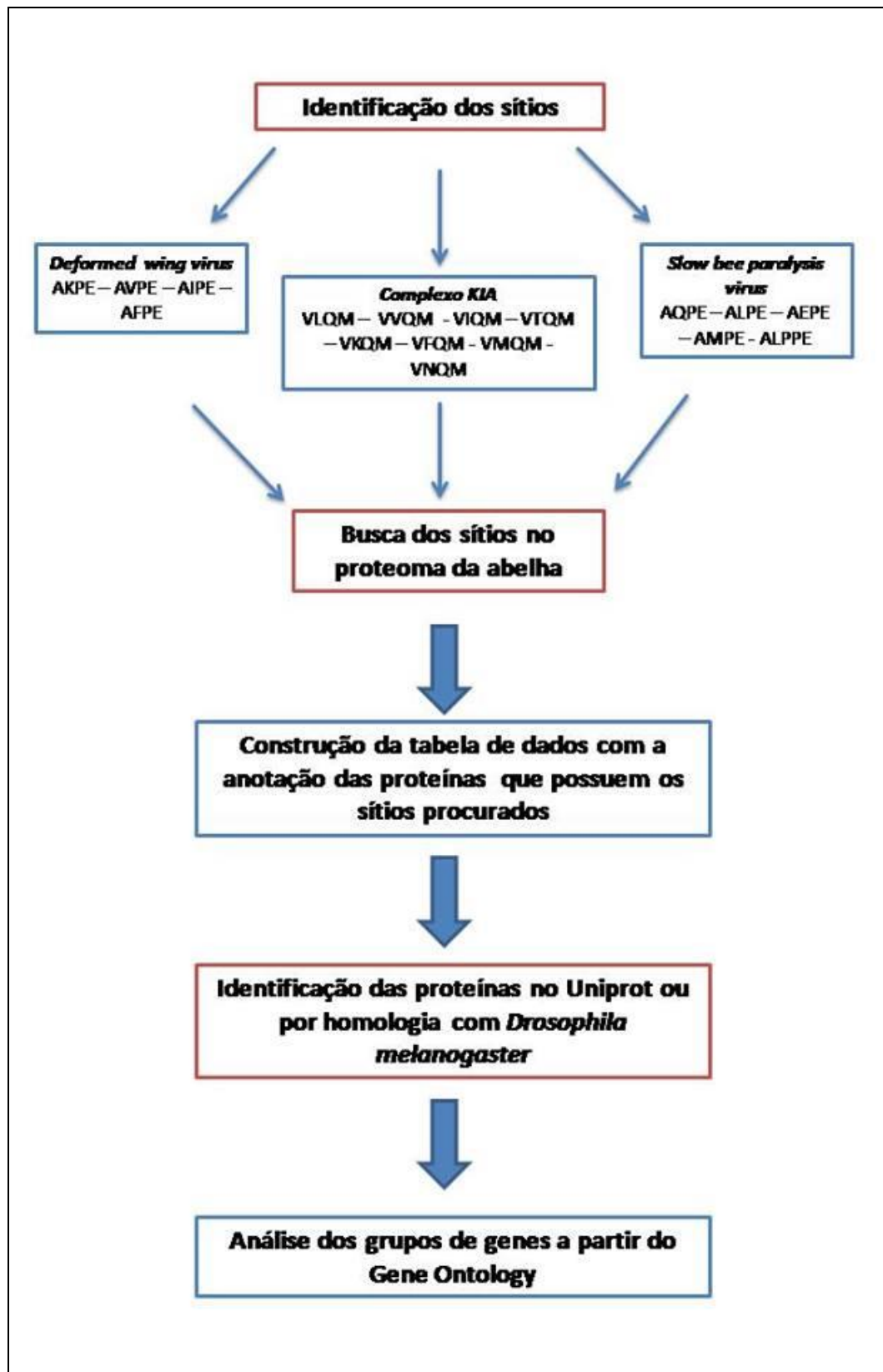
O proteoma da *Apis mellifera* foi obtido a partir do banco de dados online NCBI (National Center of Biotechnology Information) em 2013. O proteoma foi submetido ao software BioEdit e então foi realizada a busca manual de cada sítio. A partir disto, foi anotado o código de cada proteína onde encontrava-se o sítio em questão.

3.3 Análise das sequências potencialmente clivadas pela protease 3C com a ferramenta Gene Ontology.

A partir da obtenção dos códigos, foi realizada a identificação de cada proteína utilizando o banco de dados online Uniprot (2014). Em seguida, os dados foram submetidos à análise no banco de dados Gene Ontology, para o reconhecimento da função de cada proteína. Quando a proteína não era identificada, foi realizado BLASTp com organismo de referência para a busca, a *Drosophila melanogaster*, pois o banco de dados para *Apis mellifera* do Gene Ontology encontra-se defasado.

Sendo assim, as anotações dos códigos das proteínas resultaram na obtenção dos grupos de genes referentes às proteínas identificadas, com suas respectivas anotações funcionais. O fluxograma do processo metodológico pode ser visto a seguir na figura 4.

Figura 4 – Modelo esquemático da metodologia utilizada



Fonte: Clarissa Barcelos

4. RESULTADOS E DISCUSSÃO

4.1 DWV

Nas análises dos sítios destacou-se fortemente a presença de diferentes membros da família de peptídeos Apidaecina, proteína característica do proteoma de *Apis mellifera*. Estes peptídeos encontram-se presente na hemolinfa e desempenham atividade antimicrobiana principalmente em resposta a infecções por bactérias gram-negativas (Li *et al.*, 2006). Ainda não foram relatados estudos evidenciando uma correlação entre o vírus DWV e algum patógeno bacteriano de abelhas, entretanto a clivagem deste sítio provavelmente se dá como uma estratégia viral em moldar o ambiente celular do hospedeiro a fim de torná-lo mais suscetível a doença.

A análise dos dados através da ferramenta Gene Ontology (GO) mostrou que as funções proteicas afetadas congregam-se em 3 grandes grupos: ligação a ácidos nucleicos, proteólise e desenvolvimento neuronal. Neste destacaram-se genes fundamentalmente responsáveis por morfogênese; desenvolvimento e organização celular; desenvolvimento e morfogênese neuronal. Como exemplo, pode-se citar o gene *Dachsous*, responsável pela codificação de uma proteína transmembrana, que atua na polarização celular e também por um segundo tipo de polarização, em células epiteliais, conhecido como polarização planar como foi relatado por Thomas & Strutt (2012) em suas pesquisas. Mutações neste gene resultam em defeitos na morfogênese dos tecidos do tórax, asas e pernas (Rodríguez, 2004).

No grupo caracterizado por cluster de genes que desempenham funções de ligação a ácidos nucleicos aparecem: ligação de ATP, ligação a adenil ribonucleotídeo, ligação a ribonucleotídeo purínico, ligação a nucleotídeo, ligação a ribonucleotídeo; ligação a íon zinco, ligação a íon metálico de transição, ligação a cátion, entre outros. Neste grupo destacamos o gene *Dicer-1*, as proteínas codificadas por esse gene possuem atividade de RNase sendo responsáveis pela clivagem e produção de pequenos RNAs. A proteína *Dicer* é essencial nos processos de RNA interferência que é capaz de degradar moléculas RNA mensageiro, assumindo papel importante como mediadora da resposta antiviral citoplasmática (MacKay *et al.*, 2014). Neste cluster também se destaca a presença do gene *Dead-box-1*, que codifica proteínas responsáveis na participação na maioria dos processos de metabolismo do RNA, agindo como RNases ou RNA disruptases (proteínas que desfazem interações RNA-proteína) atuando tanto em processos enzimáticos de ligação a RNA, ATPase, atividade de anelamento

e desanelamento, como na atuação em processos biológicos de transcrição, splicing, síntese ribossomal, entre outros (Cordin, *et al.*, 2006).

Todos os grupos corroboram a ideia de que o DWV na infecção do patógeno, utiliza das mais variadas ferramentas para replicação na célula hospedeira evitando mecanismos antivirais e também intervém em processos biológicos importantes no desenvolvimento da abelha. Este último dado é de extrema importância, pois além da preocupação na infecção deste vírus pelos problemas no desenvolvimento das abelhas, este também poderia contribuir na infecção com outros patógenos imunossuprimindo os indivíduos infectados. Um exemplo dos efeitos deletérios deste vírus é o caso da entrada de uma nova variante do vírus na Ilha Oahu, do complexo de ilhas de Hawaii, no ano de 2007. Este evento foi devastador para os apicultores da ilha que perderam 65% das suas colmeias (Martin *et al.*, 2012).

As tabelas com os resultados das análises de cada cluster, a partir do software on line DAVID (Huang *et al.*, 2009) estão em anexo.

4.2 SBPV

Para o *Slow bee paralysis virus*, identificaram-se 33 clusters. Destacando-se os agrupamentos: divisão celular, ligação nucleotídica, morfogênese epitelial, regulação da transcrição, desenvolvimento glandular exócrino, desenvolvimento do gameta masculino, atividade cinase, morfogênese neuronal, metamorfose, sinalização celular, morfogênese do olho, morfogênese do instar larval e pupal. Nestes grupamentos destacamos a família de genes Roundabout (ROBO) que desempenham papéis na orientação do axônio. Cada membro dessa família tem um grau de envolvimento diferente que, além da disposição e orientação dos axônios, também desempenha função na regulação do crescimento traqueal e na orientação da migração de mioblastos. As proteínas desta família desenvolvem função de ligação proteica a sítios como da tirosina cinase e também na interação com proteínas de ativação da GTPase havendo algumas evidencias de que estas proteínas podem funcionar, *in vivo*, como moléculas de adesão celular para promover o crescimento axônico (Parsons *et al.*, 2003).

No grupamento de desenvolvimento glandular exócrino, destaca-se o gene Shibire (*shi*) que em *Drosophila* codifica a proteína GTPase dinamina, a qual participa na formação de vesículas tanto nas membranas celulares quanto no complexo de Golgi, estando assim envolvida nos processos de endocitoses e é essencial no processo de reciclagem de vesículas sinápticas (Kitamoto, 2001). Em suas pesquisas, Inada *et al.* (2011), evidenciou que a

modulação na expressão deste gene, por temperatura, resulta em defeitos mutagênicos nas larvas levando a deformidades, paralisia e problemas na atividade locomotora.

Tais grupos proteicos alvos de clivagem da protease 3C colaboram com as evidências da sintomatologia mais comum visualizada em abelhas com esse vírus, a paralisia dos primeiros pares de patas e uma letargia constante.

As tabelas com os resultados das análises de cada cluster, a partir do software on line DAVID (Huang *et al.*, 2009) estão em anexo.

4.3 Complexo KIA

Para o complexo KIA, através do Gene Ontology foram identificados 31 clusters. Dentre os agrupamentos de genes conforme a função, destacam-se: grupo de ligação nucleotídica, oogênese, localização e transporte de proteínas, regulação celular, atividade ATPase, atividade helicase, regulação positiva do metabolismo celular, desenvolvimento celular do ovário, ankirin, ligação com tubulina, regulação e modificação da cromatina, atividade cinase, WD40, morfogênese celular, desenvolvimento epitelial, fibronectina, fagocitose, processamento de RNA mensageiro, divisão celular, segmentação da blastoderme, proteólise, atividade motora, transmissão sináptica, metamorfose, atividade GTPase, atividade peptidase. Foram identificados genes envolvidos em desenvolvimento muscular e neuronal, como o gene Brahma que é responsável pela codificação do complexo proteico Brahma e, em *Drosophila*, é encarregado na organização do nucleossomo desempenhando papel no desenvolvimento biológico das asas, controle do ciclo celular, autorenovação da células tronco do sistema neural e proliferação das células tronco intestinais (Shi *et al.*, 2014). Evidenciou-se também, a presença do gene Blue cheese, gene envolvido em processos biológicos de transporte lisossomal, determinação da vida útil adulta, regulação de ubiquitinação protéica, regulação da organização sináptica e também em componentes celulares como na junção muscular, ativação da zona pré-sináptica entre outras (Finley *et al.*, 2003). Nos processos de divisão celular, podemos citar o gene Aurora Borealis, envolvido na ativação da proteína Aurora A cinase, que está ligada a maturação dos centrômeros e localização de proteínas nos processos mitóticos (Seki *et al.*, 2008).

Com base nesses grupos de genes, podemos compreender algumas lacunas existentes quanto aos sintomas desse complexo de vírus. O vírus Kashmir não apresenta uma descrição geral dos seus sintomas, sendo muitas vezes não identificado visualmente; porém, há relatos

de que muitas abelhas não conseguem chegar à fase adulta, o que pode ser explicado pela clivagem de proteínas que pertencem ao grupo de genes relacionados à morfogênese celular e desenvolvimento da blastoderme. Os vírus israelita e da paralisia aguda tem como sintomatologia a perda do controle dos movimentos da abelha, o que pode ser compreendido pela clivagem de proteínas pertencentes ao grupo de genes com função motora e transmissão sináptica.

As tabelas com os resultados das análises de cada cluster, a partir do software on line DAVID (Huang *et al.*, 2009) estão em anexo.

5. CONSIDERAÇÕES FINAIS

Com este estudo, foi possível realizar uma análise dos sítios alvo de clivagem da protease 3C viral dos vírus que infectam abelhas. A partir dessa indentificação, pode-se inferir que estas proteases clivam proteínas do hospedeiro que possuem as mesmas sequências de reconhecimento, e desta forma, a evolução poderia ter selecionados sítios de clivagens que afetem proteínas envolvidas nos mecanismos antivirais fazendo da protease 3C não apenas uma ferramenta para a clivagem da sua poliproteína, como também uma ferramenta para facilitar a infecção no seu hospedeiro.

Como perspectivas futuras, este estudo possibilita uma gama de novas pesquisas na busca de entender o mecanismo de modulação desta protease na clivagem dos sítios. Visando compreender estas interações, os dados obtidos nestas análises deverão ser comprovados em atividades experimentais, pois esses sítios alvo são proteínas que desempenham funções correspondentes a alguns sintomas dos vírus estudados e poderão ser usadas como alvo para o desenvolvimento de alternativas antivirais para o tratamento das abelhas.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABEMEL, Associação Brasileira dos Exportadores de Mel. **Dados das exportações de mel**, 2014. Disponível em: <http://www.brazilletsbee.com.br/dados-setoriais.aspx>, acessado em: 04 ago 2014.
- AMINEVA, S. P., AMINEV, A. G., PALMENBERG, A. C., GERN J. E. **Rhinovirus 3C protease precursors 3CD and 3CD' localize to the nuclei of infected cells**. Journal of General Virology, 85: 2969-2979, 2004.
- ARNOLD, E., LUO, M., VRIEND, G., ROSSMANN, M. G., PALMENBERG, A. C., PARKS, G. D., NICKLIN, M. J. H., WIMMER, E. **Implications of the picornavirus capsid structure for polyprotein processing**. Proc. Natl. Acad. Sci. USA, 84: 21-25, 1987.
- BAKER, A. C. & SCHROEDER, D. C. **The use of RNA-dependent RNA polymerase for the taxonomic assignment of Picorna-like viruses (order Picornavirales) infecting *Apis mellifera* L. populations**. Virology Journal, 5:10, 2008.
- BONCRISTIANI, H. F., DI PRISCO, G., PETTIS, J. S., HAMILTON, M., & CHEN, Y. P. **Molecular approaches to the analysis of deformed wing virus replication and pathogenesis in the honey bee *Apis mellifera***. Virology journal, 2009.
- CORDIN, Olivier, *et al.* **The Dead-box protein family of RNA helicases**. Gene, 367: 17-37, 2006.
- COX-FOSTER, D. L. *et al.* **A metagenomic survey of microbes in honeybee Colony Collapse Disorder**. Science, 318: 283-286, 2007.
- FINLEY, K. D. *et al.* **Blue cheese Mutations Define a Novel, Conserved Gene Involved in Progressive Neural Degeneration**. The Journal of Neuroscience, 23(4):1254 –1264, 2003.
- FRANCIS, Roy Mathew. **Viral Diseases in Honey bees**. Ph.D. Thesis Science and technology. Aarhus University, 2012. Disponível em: https://pure.au.dk/portal/files/52004373/RoyPhDThesis_PrintVersion.pdf. Acessado em: 22 maio 2013.
- GENERSCH, E. **Honey bee pathology: current threats to honey bees and beekeeping**. Applied Microbiology Biotechnology. p. 87, 2010.

GIFFORD, Chelsea. **Colony collapse disorder - The vanishing honeybee (*Apis mellifera*)**. A thesis submitted to the University of Colorado at Boulder .May, 2011. Disponível em: http://digitool.library.colostate.edu/exlibris/dtl/d3_1/apache_media/L2V4bGlicmlzL2R0bC9kM18xL2FwYWNoZV9tZWRpYS8xMTk1ODc=.pdf, acessado em: 22 maio 2013.

GONÇALVES, L. S..**The Introduction of the African Bees (*Apis mellifera adansonii*) into Brazil and Some Comments on their Spread in South America**.*American Bee Journal*.114 (11), 1974.

HUANG, DW; *et al.* **Systematic and interactive analysis of large gene lists using DAVID Bioinformatic Resources**. *Nature Protoc.* 2009;4(1);44-47

INADA, Kengo *et al.* **Optical Dissection of Neural Circuits Responsible for *Drosophila* Larval Locomotion with Halorhodopsin**. *PLoS ONE* 6 (12), 2011.

KAFTANOGLU, O; LINKSVAYER, T. A. ; PAGE, R. E.. **Rearing honey bees, *Apis mellifera*, in vitro1: Effects of sugar concentrations on survival and development**. *Journal of Insect Science* 11:96, 2011.

KITAMOTO, Toshihiro. **Conditional Modification of Behavior in *Drosophila* by Targeted Expression of a Temperature-Sensitive shibire Allele in Defined Neurons**. *Developmental Neurobiology - Journal of Neurobiology*, 47 (2): 81-92, 2001.

KOONIN, E. V.; WOLF, Y. I., NAGASAKI, K., DOLJA, V. V..**The Big Bang of Picorna-like virus evolution antedates the radiation of eucariotic supergroups**. *Nature Reviews - Microbiology*, vol. 6, p. 925-939, 2008.

LANZI, G. *et al.***Molecular and biological characterization of Deformed Wing Virus of honeybees (*Apis mellifera* L.)**. *Journal of Virology*, 80: 4998-5009, 2006.

LI, Wei-Fei *et al.* **Apidaecin-type peptides: Biodiversity, structure-function relationships and mode of action**. *Peptides*, 27:2350-2359, 2006.

MACKAY, C. R. *et al.***Dicer's role as an antiviral: still an enigma**. *Science Direct*, 26:49-55, 2014.

MARTIN, S. J. *et al.* **Global Honey Bee Viral Landscape Altered by a Parasitic Mite.** *Science* 336, 2012.

MIRANDA, J. R. de. *et al.* **The Acute bee paralysis virus - Kashmir bee virus - Israeli acute paralysis virus complex.** *Journal of invertebrate Pathology*, 103: S30-S47, 2010 (a).

MIRANDA, J. R. de. *et al.* **Genetic characterization of slow bee paralysis virus of the honeybee (*Apis mellifera* L.).** *Journal of General Virology*, 91: 2524-2530, 2010 (b).

MIRANDA, J. R. de. *et al.* **Standard Methods for virus research in *Apis mellifera*.** *Journal of Apicultural Research* 52 (4), 2013.

PARSONS, Linda *et al.* **Roundabout gene family functions during sensory axon guidance in the *Drosophila* embryo are mediated by both Slit-dependent and Slit-independent mechanisms.** *Developmental Biology*, 264: 363-375, 2003.

PORTER, A.G. **Picornavirus nonstructural proteins: emerging roles in virus replication and inhibition of host cell functions.** *Journal of Virology* 67(12): 6917-6921, 1993.

ROCHA, Maria Cecília de Lima & ALENCAR, Sá de. **Efeitos Agrotóxicos sobre as Abelhas Silvestres no Brasil: Proposta metodológica de acompanhamento.** IBAMA. Brasília, 2012. Disponível em: <http://www.ibama.gov.br/publicadas/publicacao-do-ibama-aponta-efeitos-dos-agrotoxicos-sobre-as-abelhas-silvestres-no-brasil.pdf>, acessado em: 22 maio 2013.

RODRÍGUEZ, Isabel. **The dachsous gene, a member of the cadherin family, is required for Wg-dependent pattern formation in the *Drosophila* wing disc.** *Development*, 131:3195-3206, 2004.

SABBAG, O. J., & NICODEMO, D. **Viabilidade Econômica Para Produção De Mel Em Propriedade Familiar.** *Pesquisa Agropecuária Tropical*, 2011. Disponível em: <http://www.revistas.ufg.br/index.php/pat/article/view/10414>, acessado em: 22 maio 2013.

SEKI, Akiko *et al.* **Bora and Aurora A Cooperatively Activate Plk1 and Control the Entry into Mitosis.** *Science*, 320(5883): 1655–1658, 2008.

SHI, Jiejun *et al.* **Drosophila Brahma complex remodels nucleosome organizations in multiple aspects.** Nucleic Acids Research, 2014.

TENTCHEVA, D., GAUTHIER, L., ZAPPULLA, N., DAINAT, B., COUSSERANS, F., EDOUARD, M., BERGOIN, M., *et al.* **Prevalence and Seasonal Variations of Six Bee Viruses in *Apis mellifera* L . and *Varroa destructor* Mite Populations in France.** Applied and Environmental Microbiology, p. 7185-7191, 2004.

THOMAS, Chloe & STRUTT, David. **The roles of Cadherins Fat and Dachous in Planar Polarity Specification in *Drosophila*.** Developmental Dynamics, 241:27–39, 2012.

UNIPROT. **The Uniprot Consortium Activies at the Universal Protein Resource (Uniprot).** Nucleic Acids Res. 42: D191-D198 (2014)

VANENGELSDORP, D., EVANS, J. D., SAEGERMAN, C., MULLIN, C., HAUBRUGE, E., NGUYEN, B. K., FRAZIER, M., *et al.* **Colony collapse disorder: a descriptive study.** PloS one, 2009.

WANG A, Q. M. & CHEN, S. **Human Rhinovirus 3C Protease as a Potential target for the development of antiviral agents.** Current Protein and Peptide Science, 8, 19-27, 2007.

WINSTON, Mark L. **A biologia da abelha.** Tradução de Carlos A. Osowski. Porto Alegre: Magister, 2003.

YUE, Constanze& GENERSCH, Elke. **RT-PCR analysis of *Deformed wing virus* in honeybees (*Apis mellifera*) and mites (*Varroa destructor*).** Journal of General Virology, 86: 3919-3924, 2005.

ANEXOS

1. DWV






















































Figura 1: Clusters formados a partir das análises funcionais no Gene Ontology.

Category	Term	RT	Genes in Group	Count	LF	PH	PI	%	P-Value	Fold Enrichment	Genes not in Group
GOTERM_BP_FAT	neuron projection morphogenesis	RT	4	4	286	7937	100,0	4,6E-5	2.1E1	(3)	
GOTERM_BP_FAT	neuron projection development	RT	4	4	287	7937	100,0	4,7E-5	2.1E1	(3)	
GOTERM_BP_FAT	cell morphogenesis involved in neuron differentiation	RT	4	4	288	7937	100,0	4,7E-5	2.1E1	(3)	
GOTERM_BP_FAT	cell morphogenesis involved in differentiation	RT	4	4	303	7937	100,0	5,5E-5	2.0E1	(5)	
GOTERM_BP_FAT	cell projection morphogenesis	RT	4	4	320	7937	100,0	6,5E-5	1.9E1	(3)	
GOTERM_BP_FAT	cell part morphogenesis	RT	4	4	331	7937	100,0	7,2E-5	1.8E1	(3)	
GOTERM_BP_FAT	neuron development	RT	4	4	347	7937	100,0	8,3E-5	1.7E1	(3)	
GOTERM_BP_FAT	cell projection organization	RT	4	4	365	7937	100,0	9,6E-5	1.6E1	(5)	
GOTERM_BP_FAT	neuron differentiation	RT	4	4	409	7937	100,0	1,4E-4	1.5E1	(5)	
GOTERM_BP_FAT	cell morphogenesis	RT	4	4	442	7937	100,0	1,7E-4	1.3E1	(8)	
GOTERM_BP_FAT	cellular component morphogenesis	RT	4	4	518	7937	100,0	2,8E-4	1.1E1	(9)	
INTERPRO	Protein kinase, core	RT	5	6	216	10196	83,3	9,6E-7	3.1E1		
GOTERM_BP_FAT	protein amino acid phosphorylation	RT	5	6	259	7937	83,3	1,1E-6	2.5E1		
GOTERM_MF_FAT	ATP binding	RT	6	6	748	7918	100,0	7,4E-6	8.8E0	(7)	
GOTERM_MF_FAT	adenyl ribonucleotide binding	RT	6	6	750	7918	100,0	7,5E-6	8.8E0	(7)	
GOTERM_MF_FAT	protein kinase activity	RT	5	6	283	7918	83,3	7,8E-6	1.9E1	(1)	
GOTERM_BP_FAT	phosphorylation	RT	5	6	425	7937	83,3	8,1E-6	1.5E1	(1)	
GOTERM_MF_FAT	adenyl nucleotide binding	RT	6	6	812	7918	100,0	1,1E-5	8.1E0	(9)	
GOTERM_MF_FAT	purine nucleoside binding	RT	6	6	817	7918	100,0	1,2E-5	8.1E0	(9)	
GOTERM_MF_FAT	nucleoside binding	RT	6	6	824	7918	100,0	1,2E-5	8.0E0	(9)	
GOTERM_MF_FAT	purine ribonucleotide binding	RT	6	6	923	7918	100,0	2,1E-5	7.1E0	(9)	
GOTERM_MF_FAT	ribonucleotide binding	RT	6	6	923	7918	100,0	2,1E-5	7.1E0	(9)	
GOTERM_BP_FAT	phosphate metabolic process	RT	5	6	551	7937	83,3	2,3E-5	1.2E1	(4)	
GOTERM_BP_FAT	phosphorus metabolic process	RT	5	6	551	7937	83,3	2,3E-5	1.2E1	(4)	
GOTERM_MF_FAT	purine nucleotide binding	RT	6	6	989	7918	100,0	3,0E-5	6.7E0	(11)	
INTERPRO	Serine/threonine protein kinase, active site	RT	4	6	156	10196	66,7	3,4E-5	3.3E1		
INTERPRO	Serine/threonine protein kinase-related	RT	4	6	168	10196	66,7	4,3E-5	3.0E1		
GOTERM_MF_FAT	nucleotide binding	RT	6	6	1206	7918	100,0	8,1E-5	5.5E0	(13)	
SP_PIR_KEYWORDS	transferase	RT	5	6	873	12980	83,3	9,6E-5	9.9E0	(5)	
GOTERM_MF_FAT	protein serine/threonine kinase activity	RT	4	6	203	7918	66,7	1,6E-4	2.0E1	(1)	
INTERPRO	Serine/threonine protein kinase	RT	3	6	112	10196	50,0	1,2E-3	3.0E1		
SMART	S_TKc	RT	3	6	112	4824	50,0	3,1E-3	1.7E1		
SP_PIR_KEYWORDS	atp-binding	RT	3	6	585	12980	50,0	1,9E-2	7.4E0	(4)	
SP_PIR_KEYWORDS	nucleotide-binding	RT	3	6	743	12980	50,0	2,9E-2	5.8E0	(6)	
SP_PIR_KEYWORDS	serine/threonine-protein kinase	RT	2	6	148	12980	33,3	5,6E-2	1.5E1		
GOTERM_BP_FAT	cell adhesion	RT	2	6	174	7937	33,3	8,5E-2	9.1E0	(4)	
INTERPRO	Protein kinase, ATP binding site	RT	2	6	182	10196	33,3	8,6E-2	9.3E0		
GOTERM_BP_FAT	biological adhesion	RT	2	6	188	7937	33,3	9,1E-2	8.4E0	(4)	
GOTERM_MF_FAT	zinc ion binding	RT	4	4	1026	7918	100,0	2,2E-3	5.8E0	(3)	
GOTERM_MF_FAT	transition metal ion binding	RT	4	4	1354	7918	100,0	5,0E-3	4.4E0	(8)	
GOTERM_MF_FAT	metal ion binding	RT	4	4	1718	7918	100,0	1,0E-2	3.5E0	(12)	
GOTERM_MF_FAT	cation binding	RT	4	4	1764	7918	100,0	1,1E-2	3.4E0	(12)	
GOTERM_MF_FAT	ion binding	RT	4	4	1772	7918	100,0	1,1E-2	3.4E0	(13)	
GOTERM_MF_FAT	exopeptidase activity	RT	2	4	106	7918	50,0	4,0E-2	1.9E1		
GOTERM_MF_FAT	metallopeptidase activity	RT	2	4	168	7918	50,0	6,2E-2	1.2E1		
INTERPRO	Zinc finger, C2H2-type	RT	2	4	320	10196	50,0	9,1E-2	8.0E0		
GOTERM_BP_FAT	proteolysis	RT	2	4	731	7937	50,0	9,2E-2	5.4E0	(4)	

Fonte: Clarissa Barcelos

2. SBPV

Figura 2: Clusters formados a partir das análises funcionais no Gene Ontology.

Annotation Cluster 1		Enrichment Score: 1.95			Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_CC_FAT	microtubule cytoskeleton	RT		8	3.1E-3	3.5E-1
<input type="checkbox"/>	GOTERM_CC_FAT	cytoskeletal part	RT		8	1.3E-2	6.0E-1
<input type="checkbox"/>	GOTERM_CC_FAT	cytoskeleton	RT		8	3.5E-2	7.1E-1
Annotation Cluster 2		Enrichment Score: 1.61			Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	nucleotide-binding	RT		13	2.3E-3	1.1E-1
<input type="checkbox"/>	GOTERM_MF_FAT	purine nucleotide binding	RT		15	2.7E-2	1.0E0
<input type="checkbox"/>	GOTERM_MF_FAT	ribonucleotide binding	RT		14	3.5E-2	9.7E-1
<input type="checkbox"/>	GOTERM_MF_FAT	purine ribonucleotide binding	RT		14	3.5E-2	9.7E-1
Annotation Cluster 3		Enrichment Score: 1.1			Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	cell cycle	RT		4	2.3E-2	2.7E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	mitosis	RT		3	4.1E-2	3.6E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	cell division	RT		3	5.8E-2	3.9E-1
Annotation Cluster 5		Enrichment Score: 0.96			Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	atp-binding	RT		9	3.1E-2	3.2E-1
Annotation Cluster 6		Enrichment Score: 0.95			Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	morphogenesis of an epithelium	RT		5	9.8E-2	9.9E-1
Annotation Cluster 7		Enrichment Score: 0.92			Count	P_Value	Benjamini
<input type="checkbox"/>	INTERPRO	Immunoglobulin-like fold	RT		4	5.0E-2	9.8E-1
<input type="checkbox"/>	INTERPRO	Immunoglobulin I-set	RT		3	6.1E-2	9.7E-1
Annotation Cluster 8		Enrichment Score: 0.85			Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	transcription regulation	RT		6	9.8E-2	4.4E-1
Annotation Cluster 9		Enrichment Score: 0.84			Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	salivary gland development	RT		4	9.4E-2	9.9E-1
<input type="checkbox"/>	GOTERM_BP_FAT	exocrine system development	RT		4	9.4E-2	9.9E-1
Annotation Cluster 10		Enrichment Score: 0.82			Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_MF_FAT	microtubule motor activity	RT		3	7.7E-2	9.3E-1
Annotation Cluster 11		Enrichment Score: 0.8			Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	detection of stimulus involved in sensory perception	RT		3	5.8E-2	1.0E0
Annotation Cluster 12		Enrichment Score: 0.8			Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	gtp-binding	RT		4	7.1E-2	3.8E-1
Annotation Cluster 15		Enrichment Score: 0.61			Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	serine/threonine-protein kinase	RT		4	6.7E-2	4.0E-1
<input type="checkbox"/>	UP_SEQ_FEATURE	binding site:ATP	RT		4	7.7E-2	9.9E-1
Annotation Cluster 16		Enrichment Score: 0.56			Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	dendrite morphogenesis	RT		4	6.3E-2	9.9E-1
<input type="checkbox"/>	GOTERM_BP_FAT	dendrite development	RT		4	6.3E-2	9.9E-1
Annotation Cluster 19		Enrichment Score: 0.52			Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_CC_FAT	nucleoplasm	RT		6	8.7E-2	7.5E-1

Fonte: Clarissa Barcelos













































3. Complexo KIA

Figura 3: Clusters formados a partir das análises funcionais no Gene Ontology.

Annotation Cluster 1		Enrichment Score: 4.08		Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	nucleotide-binding	RT	19	3.8E-6	4.3E-4
<input type="checkbox"/>	SP_PIR_KEYWORDS	atp-binding	RT	16	1.4E-5	8.0E-4
<input type="checkbox"/>	GOTERM_MF_FAT	purine nucleotide binding	RT	22	2.3E-5	4.5E-3
<input type="checkbox"/>	GOTERM_MF_FAT	ribonucleotide binding	RT	21	2.9E-5	2.8E-3
<input type="checkbox"/>	GOTERM_MF_FAT	purine ribonucleotide binding	RT	21	2.9E-5	2.8E-3
<input type="checkbox"/>	GOTERM_MF_FAT	nucleotide binding	RT	24	4.7E-5	3.0E-3
<input type="checkbox"/>	GOTERM_MF_FAT	adenyl nucleotide binding	RT	19	6.2E-5	3.0E-3
<input type="checkbox"/>	GOTERM_MF_FAT	purine nucleoside binding	RT	19	6.7E-5	2.6E-3
<input type="checkbox"/>	GOTERM_MF_FAT	nucleoside binding	RT	19	7.5E-5	2.4E-3
<input type="checkbox"/>	GOTERM_MF_FAT	ATP binding	RT	18	7.7E-5	2.1E-3
<input type="checkbox"/>	GOTERM_MF_FAT	adenyl ribonucleotide binding	RT	18	7.9E-5	1.9E-3
<input type="checkbox"/>	GOTERM_MF_FAT	ATPase activity	RT	9	7.9E-3	1.4E-1
<input type="checkbox"/>	GOTERM_MF_FAT	ATPase activity, coupled	RT	8	1.2E-2	1.9E-1
Annotation Cluster 2		Enrichment Score: 2.45		Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	cytoplasm	RT	11	1.0E-3	2.3E-2
<input type="checkbox"/>	GOTERM_BP_FAT	ovarian follicle cell development	RT	8	1.2E-3	3.7E-1
<input type="checkbox"/>	GOTERM_BP_FAT	female gamete generation	RT	13	1.2E-3	2.7E-1
<input type="checkbox"/>	GOTERM_BP_FAT	gamete generation	RT	15	1.4E-3	2.4E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	phosphoprotein	RT	15	1.7E-3	3.2E-2
<input type="checkbox"/>	GOTERM_BP_FAT	sexual reproduction	RT	15	1.8E-3	2.4E-1
<input type="checkbox"/>	GOTERM_BP_FAT	reproductive cellular process	RT	12	2.0E-3	2.2E-1
<input type="checkbox"/>	GOTERM_BP_FAT	reproductive developmental process	RT	12	2.2E-3	1.9E-1
<input type="checkbox"/>	GOTERM_BP_FAT	multicellular organism reproduction	RT	15	3.1E-3	2.1E-1
<input type="checkbox"/>	GOTERM_BP_FAT	reproductive process in a multicellular organism	RT	15	3.1E-3	2.1E-1
<input type="checkbox"/>	GOTERM_BP_FAT	oogenesis	RT	12	3.5E-3	2.1E-1
<input type="checkbox"/>	GOTERM_BP_FAT	chorion-containing eggshell formation	RT	5	6.8E-3	2.3E-1
<input type="checkbox"/>	GOTERM_BP_FAT	eggshell formation	RT	5	7.1E-3	2.3E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	developmental protein	RT	9	3.7E-2	2.4E-1
<input type="checkbox"/>	GOTERM_BP_FAT	germ cell development	RT	5	1.3E-1	7.8E-1
Annotation Cluster 3		Enrichment Score: 2.16		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	protein localization	RT	12	2.6E-4	1.8E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	protein transport	RT	5	5.7E-3	7.8E-2

Fonte: Clarissa Barcelos

Figura 4: Clusters formados a partir das análises funcionais no Gene Ontology.

Annotation Cluster 5		Enrichment Score: 1.91	G		Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	atp-binding	RT		16	1.4E-5	8.0E-4
<input type="checkbox"/>	GOTERM_MF_FAT	ATPase activity	RT		9	7.9E-3	1.4E-1
Annotation Cluster 6		Enrichment Score: 1.9	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_MF_FAT	helicase activity	RT		6	1.9E-3	4.1E-2
<input type="checkbox"/>	INTERPRO	Helicase, superfamily 1 and 2, ATP-binding	RT		5	2.9E-3	2.7E-1
<input type="checkbox"/>	INTERPRO	DEAD-like helicase, N-terminal	RT		5	3.2E-3	2.1E-1
<input type="checkbox"/>	INTERPRO	DNA/RNA helicase, C-terminal	RT		5	3.4E-3	1.7E-1
<input type="checkbox"/>	SMART	DEXDc	RT		5	5.8E-3	2.7E-1
<input type="checkbox"/>	SMART	HELICc	RT		5	6.1E-3	1.5E-1
<input type="checkbox"/>	GOTERM_MF_FAT	ATPase activity	RT		9	7.9E-3	1.4E-1
<input type="checkbox"/>	GOTERM_MF_FAT	ATPase activity, coupled	RT		8	1.2E-2	1.9E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	helicase	RT		4	1.7E-2	1.8E-1
<input type="checkbox"/>	INTERPRO	DNA/RNA helicase, DEAD/DEAH box type, N-terminal	RT		3	4.8E-2	6.6E-1
<input type="checkbox"/>	GOTERM_MF_FAT	RNA helicase activity	RT		3	6.1E-2	5.2E-1
Annotation Cluster 7		Enrichment Score: 1.81	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	positive regulation of molecular function	RT		5	2.6E-3	2.0E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	kinase	RT		7	3.9E-3	6.2E-2
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of transferase activity	RT		4	7.9E-3	2.3E-1
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of protein kinase activity	RT		4	7.9E-3	2.3E-1
<input type="checkbox"/>	GOTERM_BP_FAT	regulation of kinase activity	RT		4	7.9E-3	2.3E-1
<input type="checkbox"/>	GOTERM_BP_FAT	intracellular signaling cascade	RT		6	9.1E-2	6.6E-1
Annotation Cluster 8		Enrichment Score: 1.72	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	ovarian follicle cell development	RT		8	1.2E-3	3.7E-1
<input type="checkbox"/>	GOTERM_BP_FAT	actin cytoskeleton organization	RT		6	5.8E-3	2.6E-1
<input type="checkbox"/>	GOTERM_BP_FAT	actin filament-based process	RT		6	5.9E-3	2.5E-1
<input type="checkbox"/>	GOTERM_BP_FAT	chorion-containing eggshell formation	RT		5	6.8E-3	2.3E-1
<input type="checkbox"/>	GOTERM_BP_FAT	eggshell formation	RT		5	7.1E-3	2.3E-1
Annotation Cluster 10		Enrichment Score: 1.42	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_MF_FAT	microtubule binding	RT		3	9.4E-2	6.6E-1
Annotation Cluster 11		Enrichment Score: 1.3	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	chromosome organization	RT		8	7.3E-3	2.3E-1
<input type="checkbox"/>	GOTERM_BP_FAT	chromatin organization	RT		5	4.8E-2	4.8E-1
<input type="checkbox"/>	GOTERM_BP_FAT	chromatin remodeling	RT		3	5.3E-2	5.1E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	chromatin regulator	RT		3	7.6E-2	3.8E-1
Annotation Cluster 12		Enrichment Score: 1.24	G		Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	kinase	RT		7	3.9E-3	6.2E-2
<input type="checkbox"/>	GOTERM_BP_FAT	cell growth	RT		4	6.7E-3	2.4E-1
<input type="checkbox"/>	GOTERM_MF_FAT	protein kinase activity	RT		8	8.2E-3	1.4E-1
<input type="checkbox"/>	INTERPRO	Protein kinase, ATP binding site	RT		5	6.1E-2	6.5E-1
<input type="checkbox"/>	INTERPRO	Serine/threonine protein kinase	RT		4	6.3E-2	6.4E-1
<input type="checkbox"/>	SP_PIR_KEYWORDS	serine/threonine-protein kinase	RT		4	8.8E-2	4.1E-1
<input type="checkbox"/>	GOTERM_BP_FAT	intracellular signaling cascade	RT		6	9.1E-2	6.6E-1
<input type="checkbox"/>	SMART	S_TKc	RT		4	9.5E-2	7.5E-1

Fonte: Clarissa Barcelos

Figura 5: Clusters formados a partir das análises funcionais no Gene Ontology.

Annotation Cluster 13		Enrichment Score: 1.22	G		Count	P_Value	Benjamini
<input type="checkbox"/>	INTERPRO	WD40 repeat	RT		5	6.6E-2	6.3E-1
<input type="checkbox"/>	INTERPRO	WD40/YVTN repeat-like	RT		5	7.4E-2	6.5E-1
Annotation Cluster 14		Enrichment Score: 1.1	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	muscle organ development	RT		8	5.3E-3	2.5E-1
<input type="checkbox"/>	GOTERM_BP_FAT	neuron development	RT		9	6.4E-3	2.4E-1
<input type="checkbox"/>	GOTERM_BP_FAT	neuron projection morphogenesis	RT		8	7.9E-3	2.3E-1
<input type="checkbox"/>	GOTERM_BP_FAT	cellular component morphogenesis	RT		11	8.0E-3	2.2E-1
<input type="checkbox"/>	GOTERM_BP_FAT	neuron projection development	RT		8	8.0E-3	2.1E-1
<input type="checkbox"/>	GOTERM_BP_FAT	cell morphogenesis involved in neuron differentiation	RT		8	8.1E-3	2.1E-1
<input type="checkbox"/>	GOTERM_BP_FAT	axon extension	RT		3	9.0E-3	2.2E-1
Annotation Cluster 15		Enrichment Score: 1.07	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	ovarian follicle cell development	RT		8	1.2E-3	3.7E-1
<input type="checkbox"/>	GOTERM_BP_FAT	intracellular signaling cascade	RT		6	9.1E-2	6.6E-1
Annotation Cluster 16		Enrichment Score: 1.05	G		Count	P_Value	Benjamini
<input type="checkbox"/>	INTERPRO	Fibronectin, type III-like fold	RT		3	5.6E-2	6.5E-1
<input type="checkbox"/>	INTERPRO	Fibronectin, type III	RT		3	9.6E-2	7.1E-1
Annotation Cluster 17		Enrichment Score: 1	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	vesicle-mediated transport	RT		10	3.9E-3	2.1E-1
Annotation Cluster 19		Enrichment Score: 0.88	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	chromosome organization	RT		8	7.3E-3	2.3E-1
Annotation Cluster 20		Enrichment Score: 0.72	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	cell division	RT		7	6.0E-3	2.4E-1
Annotation Cluster 21		Enrichment Score: 0.69	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_BP_FAT	regionalization	RT		8	7.2E-2	6.0E-1
<input type="checkbox"/>	GOTERM_BP_FAT	pattern specification process	RT		8	9.0E-2	6.7E-1
Annotation Cluster 22		Enrichment Score: 0.63	G		Count	P_Value	Benjamini
<input type="checkbox"/>	SP_PIR_KEYWORDS	ubiquitin conjugation pathway	RT		4	3.2E-2	2.2E-1
Annotation Cluster 27		Enrichment Score: 0.48	G		Count	P_Value	Benjamini
<input type="checkbox"/>	GOTERM_CC_FAT	endomembrane system	RT		6	7.0E-2	1.0E0

Fonte: Clarissa Barcelos