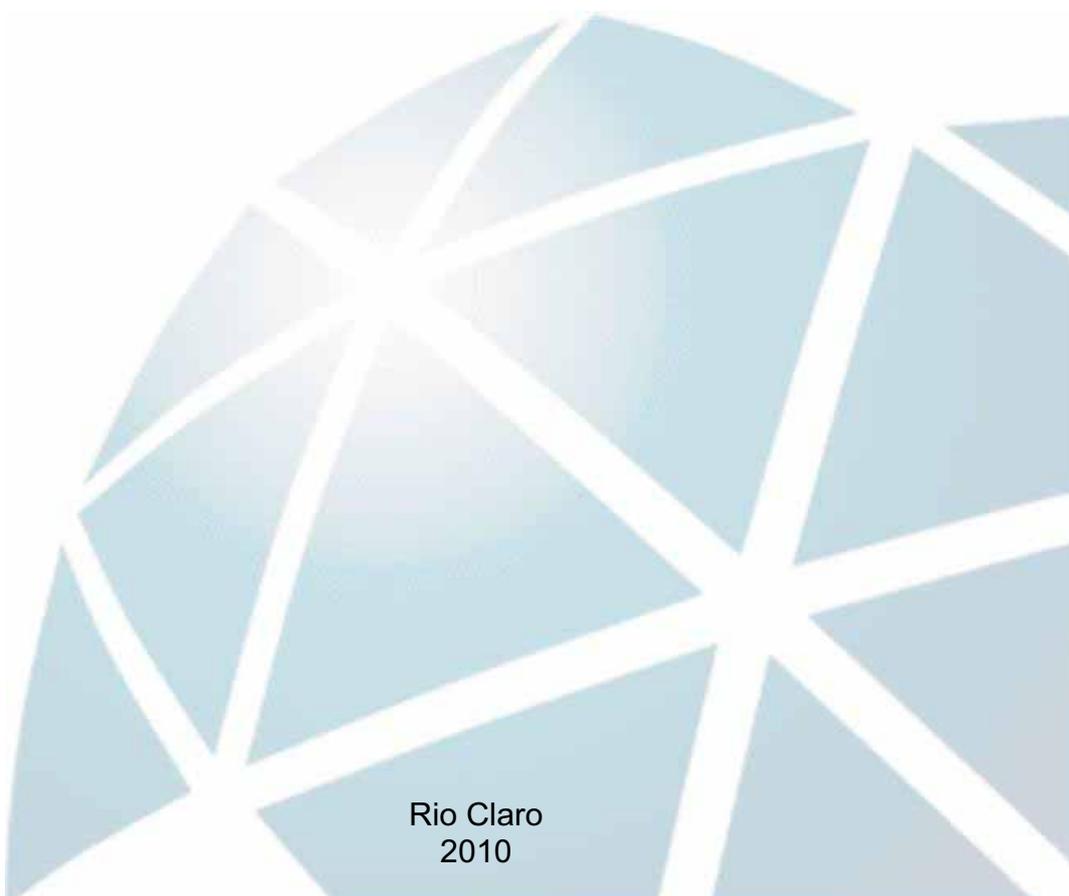

CIÊNCIAS BIOLÓGICAS INTEGRAL

FERNANDO KAMIMURA COCCHI

**ANOTAÇÃO DE 500 ESTs DA SAÚVA *Atta
laevigata***



Rio Claro
2010

FERNANDO KAMIMURA COCCHI

ANOTAÇÃO DE 500 ESTs DA SAÚVA *Atta laevigata*

Orientador: MAURÍCIO BACCI JÚNIOR

Co-orientadora: Cynara de Melo Rodvalho

Trabalho de Conclusão de Curso apresentado ao Instituto de Biociências da Universidade Estadual Paulista “Júlio de Mesquita Filho” - Campus de Rio Claro, para obtenção do grau de Bacharel em Ciências Biológicas.

Rio Claro
2010

595.796 Cocchi, Fernando Kamimura
C659a Anotação de 500 ESTs da saúva *Atta laevigata* / Fernando Kamimura
Cocchi. - Rio Claro : [s.n.], 2010
22 f. : il., figs., gráfs., tabs.

Trabalho de conclusão de curso (bacharelado - Ciências Biológicas) -
Universidade Estadual Paulista, Instituto de Biociências de Rio Claro
Orientador: Maurício Bacci Júnior
Co-Orientador: Cynara de Melo Rodovalho

1. Formiga. 2. Anotação de ESTs. 3. Saúva. 4. *Atta laevigata*. I.
Título.

SUMÁRIO

Resumo.....	3
Abstract.....	4
1. INTRODUÇÃO.....	5
2. OBJETIVOS.....	7
3. MATERIAL E MÉTODOS.....	8
3.1 Material.....	8
3.2 Extração do RNAm.....	8
3.3 Extração de DNA Plasmidial.....	8
3.4 Análises de Pré-Processamento de Sequências.....	9
4. RESULTADOS.....	10
5. DISCUSSÃO.....	17
6. CONCLUSÃO.....	18
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	19

RESUMO

As formigas cortadeiras pertencentes ao gênero *Atta* ocorrem desde as regiões tropicais às subtropicais das Américas. São considerados insetos pragas por causarem grandes prejuízos em áreas agrícolas. Dentre estas, destaca-se a saúva *Atta laevigata*, cuja colônia demanda uma enorme quantia de folhas para o cultivo de seu fungo simbiote, o qual é a principal fonte de alimento da colônia. Dessa forma, o estudo do transcriptoma desta espécie torna-se uma ferramenta útil, pois torna possível a identificação de proteínas potencialmente envolvidas com as suas habilidades como insetos pragas e também, com aquelas relacionadas às diferenças entre as castas presentes nos ninhos. No presente estudo, foi descrito o resultado da análise de parte do transcriptoma da formiga cortadeira praga *A. laevigata*, a partir de sequências de cDNA previamente geradas no Laboratório de Evolução e Molecular (LEM). Os resultados obtidos também poderão ser utilizados para estudos moleculares, ecológicos, metabólicos e evolutivos sobre formigas, além da expressão heteróloga de proteínas importantes como alvos moleculares para o controle de algumas cortadeiras pragas agrícolas.

ABSTRACT

Leaf-cutting ants belonging to the genus *Atta* occur from the tropical to subtropical regions of the Americas. These insects are considered pests because they cause serious damage in agricultural areas. Among these, stands out *Atta laevigata*, species which the colony requires a huge amount of leaves to grow its symbiotic fungus which is the main food source of the nest. Thus, the study of the transcriptome of these ants becomes a useful tool, because it is possible to identify proteins potentially involved with their skills as insect pests and also those related to differences between the varieties present in the nests. In the present study we described results of the partial analysis of the transcriptome of the leaf-cutting ant pest *A. laevigata*, from cDNA sequences previously generated in the Laboratory of Evolution and Molecular (LEM). The results may also be used for molecular, ecological, metabolic and evolutionary studies about ants, and heterologous expression of important proteins as molecular targets for the control of some leafcutting agricultural pests.

1 INTRODUÇÃO

O mutualismo entre insetos e fungos são relações interespecíficas com grande sucesso evolucionário. Uma destas interações ocorre entre formigas (Hymenoptera: Formicidae: tribo Attini) e seus fungos que são cultivados em uma simbiose muito antiga que provavelmente originou-se entre 50 e 60 milhões de anos atrás (MUELLER et al., 2001; SHULTZ; BRADY, 2008).

Esses animais possuem, como principais características, o altruísmo expresso pelos membros da colônia, a complexa divisão do trabalho e uma grande plasticidade. O comportamento social dos insetos é um resultado de interações complexas, em diferentes níveis de organização biológica. Este comportamento é definido principalmente pelas proteínas e peptídeos produzidos pelos genes. No entanto, este comportamento social resulta da interação complexa de indivíduos e são afetados por diversos fatores além daqueles relacionados à expressão dos genes (PAGE JR; ERBER, 2002).

A história evolutiva das formigas demonstra, ainda, que elas apresentam longevidade filética mais acentuada do que outros grupos de animais (baixa taxa de extinção), indicando sua adaptabilidade a mudanças de clima do planeta (WILSON, 1987).

As formigas Attini, conhecidas pelo hábito de cultivar fungos para sua alimentação (HERVEY; NAIR, 1979; VASCONCELOS, 1988; HEBLING et al., 2000), compreendem um grupo monofilético com mais de 230 espécies descritas. Dentre estas, as espécies de *Atta* possuem posição mais apical em relação a grupos taxonômicos próximos dentro desta tribo (WETTERER et al., 1998). Não só este gênero, como o *Acromyrmex* cortam e utilizam principalmente folhas como substrato para o fungo simbiote sendo, por isso, conhecidas como formigas cortadeiras (CHAPELA et al., 1994).

As formigas cortadeiras (*Atta ssp.*) são herbívoras encontradas em regiões neotrópicas e que apresentam destacado sucesso ecológico, uma vez que são amplamente distribuídas no mundo. Elas são capazes de ocupar nichos bastante diferenciados e são dominantes em relação à biomassa animal (HÖLLDOBLER; WILSON, 1990).

A formiga saúva, “cabeça-de-vidro”, é classificada no estudo dos insetos como pertencente à espécie *Atta laevigata* (SMITH, 1858), apresenta soldados com cabeça brilhante, sem pêlos e que quando esmagada não exala odor de folhas de limão, como as cabeças dos soldados de *Atta sexdens*. Sua ocorrência dá-se por praticamente toda a América, desde o Sul dos Estados Unidos até o Centro da Argentina. No Brasil, vive nos estados do Pará, Amazonas, Alagoas, Goiás, Minas Gerais, Rio de Janeiro, São Paulo e Paraná

(BORGMEIER, 1950). Nidifica em terrenos arenosos no litoral, na beira de estradas, em lugares com vegetação rasteira ou capoeiras e mesmo na periferia de áreas de mata, ocasionando danos a numerosas espécies de plantas cultivadas nas fazendas das regiões de ocorrência (DELABIE et al., 1997).

Além de importantes ferramentas ecológicas, também são pragas agrícolas que causam milhões de dólares anuais em prejuízo à agricultura, por onde passam (SOLOMON, 2007), sendo que um saueiro, forrageia cerca de uma tonelada de folhas verdes, por ano (AMANTE, 1967).

Insetos sociais da família Himenoptera (formigas, abelhas, vespas e cupins) representam um dos exemplos mais importantes quando se trata de polimorfismo. Rainhas, operárias e machos deste grupo possuem praticamente o mesmo genótipo (embora, as fêmeas são diplóides e os machos haplóides), diferentemente de quaisquer outros animais, onde quem manda são os cromossomos sexuais. Portanto, as diferenças entre as castas em insetos sociais e o sexo são derivadas das variações entre a expressão gênica (HOFFMAN; GOODISMAN, 2007).

Biblioteca de cDNA não representa o genoma inteiro e sim a população de RNAm. Este é convertido em cDNA (DNA complementar de dupla hélice), com a ajuda da enzima transcriptase reversa, é inserido em um vetor de clonagem e transferido para células (SALEMA, 1999).

Desta forma, a criação de uma biblioteca de cDNA, construída a partir do RNAm através da técnica de recombinação de DNA, permitirá a análise dos genes expressos, não somente em um organismo inteiro, como também em diferentes tecidos e células. Conseqüentemente, é possível desvendar através de comparações de bibliotecas as diferenças entre os “*pools*” gênicos dentro de um mesmo organismo, castas (soldados, rainhas, operárias e machos) estágios de desenvolvimento, entre outras (WOODS et al., 1980).

2 OBJETIVOS

O presente trabalho teve como objetivo geral caracterizar parte do transcriptoma da saúva cortadeira *Atta laevigata*.

Para tanto, os objetivos específicos foram:

- Utilizar minipreparações em placas de 96 poços para extrair o DNA plasmidial contendo fragmentos de interesse obtidos na biblioteca de cDNA;
- Sequenciar as amostras, gerando 500 ESTs (*expressed sequence tags*);
- Identificar e anotar as sequências obtidas utilizando-se das ferramentas de bioinformática disponíveis.

3 MATERIAL E MÉTODOS

3.1 Material

O presente trabalho utilizou uma biblioteca de cDNA já construída no LEM, a partir de operárias maiores e soldados de *A. laevigata*, provenientes de um ninho encontrado no campus da UNESP em Rio Claro, São Paulo, Brasil.

3.2 Extração do RNAm

Foi utilizado o kit PolyAtract, da Promega, para a obtenção de RNAm íntegro e de qualidade e a biblioteca de cDNA foi construída utilizando-se o kit CloneMiner cDNA Library Construction Kit (Invitrogen), segundo recomendações do fabricante para construção de biblioteca de cDNA não radioativa.

3.3 Extração de DNA Plasmidial

O DNA plasmidial foi extraído seguindo o protocolo de minipreparação em placa de 96 poços.

As colônias foram precipitadas ao fundo da placa por centrifugação e o sobrenadante descartado. Foi adicionado GTE (glucose 20%, EDTA 0,5 M, Tris – HCl 1 M) para a remoção dos resíduos do meio de cultivo.

Em uma placa de fundo “U” foi adicionado 5 µL de RNase (10 mg/mL) em cada pocinho, juntamente com 60 µL das células ressuspensas e 60 µL de NaOH 4 M/ SDS 10 %. As placas foram homogeneizadas e deixadas para descansar por 10 minutos. Foram adicionados 60 µL de KOAc 3 M gelado a fim de parar a ação do NaOH, novamente sendo deixado para descansar durante 10 minutos.

Após choque térmico a 90 °C e uma centrifugação de 15 minutos, o sobrenadante foi transferido para a placa filtro e centrifugado novamente. Foram adicionados 110 µL de isopropanol gelado ao filtrado, a mistura foi homogeneizada e colocada para centrifugar por 45 minutos. O sobrenadante foi descartado e ao precipitado foram adicionados 200 µL de álcool 70 %, sendo centrifugado por 10 minutos e o sobrenadante novamente descartado.

Em seguida, as placas foram colocadas em uma estufa a 50° C por 10 minutos para secar. Para eluir o DNA, foram adicionados 50 µL de água MiliQ autoclavada. A mistura foi

então agitada com um pulso no vórtex e deixada por 16 h a 4° C. A qualidade do procedimento foi analisada em Gel de agarose 1 %, corado com Brometo de Etídeo.

O produto de clonagem foi submetido à reação de sequenciamento utilizando *DYEnamic ET Terminator Cycle Sequencing Kit* (GE Healthcare), 300 ng DNA e 5 µM do *primer* M13F (5'-GTAAAACGACGGCCAG-3'), em volume final de 10 µL, o qual foi incubado em termociclador programado para 25 ciclos de 95° C por 20 seg., 50° C por 15 seg. e 60° C por 1 min. O produto de sequenciamento foi purificado segundo orientação do fabricante e analisado em sequenciador automático MegaBACE 1000 DNA Analysis System.

3.4 Análises de Pré-Processamento de Sequências

As sequências ESTs (expressed sequence tags) foram inicialmente mascaradas contra as sequências do primer e do vetor, filtradas por tamanho e qualidade e, as sequências de qualidade phred ≥ 20 foram utilizadas para agrupamento e montagem dos “*contigs*” pelo programa CAP3 (HUANG; MADAN, 1999). Essa etapa, denominada de pré-processamento, foi realizada através do *software* EGene (DURHAM et al., 2005).

Após as etapas de limpeza e montagem das sequências, estas foram analisadas utilizando-se o programa Blast2GO (CONESA; GÖTZ, 1997). Este programa realiza uma busca de similaridade contra um banco de dados de proteínas não redundante, BLAST (ALTSCHUL, et al, 1990), fez também uma busca de termos GO (THE GENE ONTOLOGY CONSORTIUM, 2000), buscou dados de domínio protéico (IPS) com auxílio do InterProScan (QUEVILLON et al, 2005) e por fim executou o EC (*Enzyme Commission*) que classifica numericamente as enzimas com base nas reações químicas que catalizam.

Foram consideradas significantes todas as sequências que apresentarem E-value de BLAST menor ou igual a 10^{-3} . A partir destes resultados, categorias foram estabelecidas para que as proteínas obtidas pudessem ser classificadas de acordo com sua provável função.

4 Resultados

Para a visualização da qualidade das minipreparações, foram preparados géis de agarose 1%. Em cada “pocinho” do gel foram pipetados 2 μ L de amostra, misturado com 2 μ L de *load* e após corrida na cuba eletrolítica, o gel foi corado com Brometo de Etídeo. Cada amostra possui três bandas pelo motivo do plasmídio estar super-enovelado ou “*supercoil*” (banda inferior), enovelado (banda intermediária) e relaxado (banda superior, mais fraca em algumas das preparações).

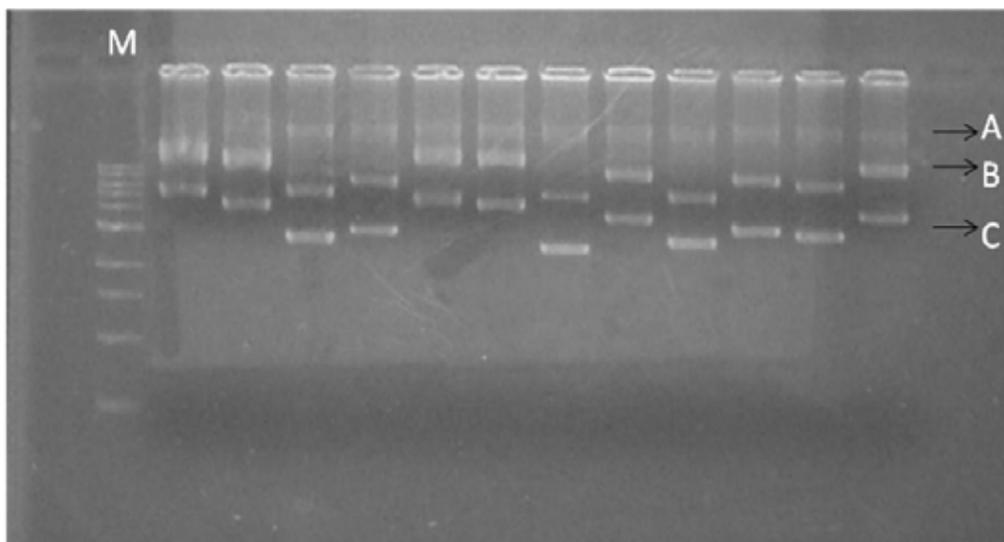


Figura 1. Gel de agarose 1% para análise da qualidade da minipreparação da placa 026. A) Plasmídio relaxado; B) Plasmídio enovelado; C) Plasmídio *supercoil*; M) Marcador molecular de 1 Kb.

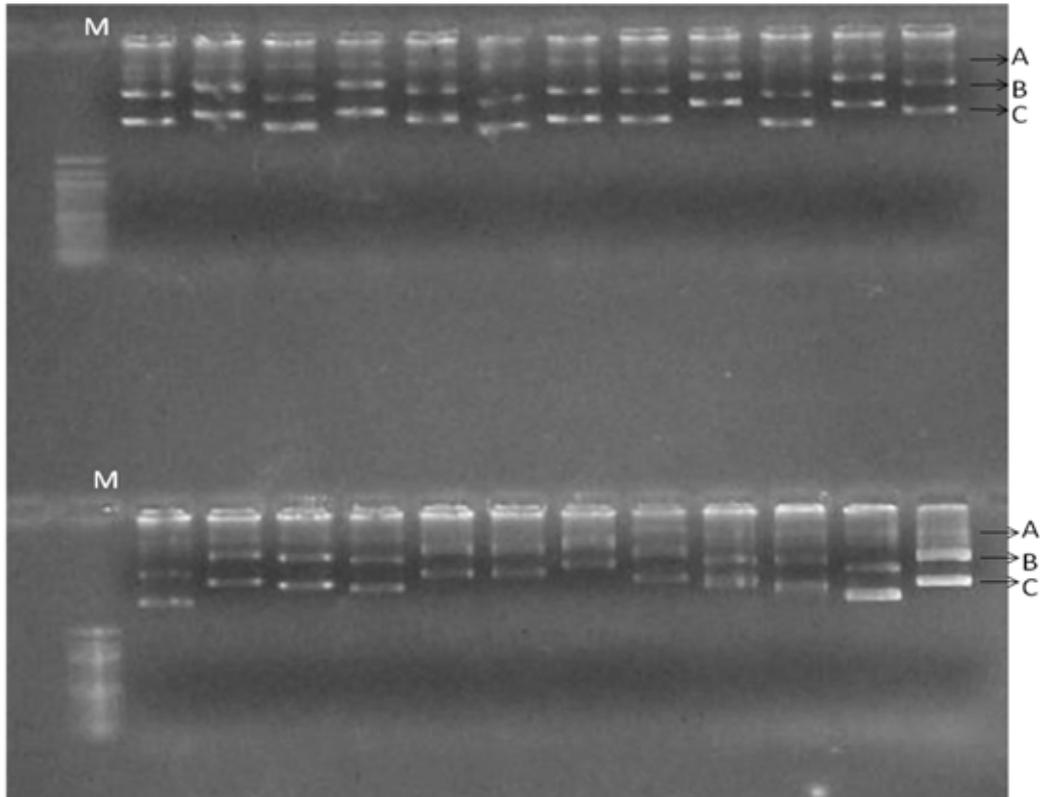


Figura 2. Gel de agarose 1% para análise da qualidade da minipreparação das placas 027 e 028. A) Plasmídeo relaxado; B) Plasmídeo enovelado; C) Plasmídeo *supercoil*; M) Marcador molecular de 1 Kb.

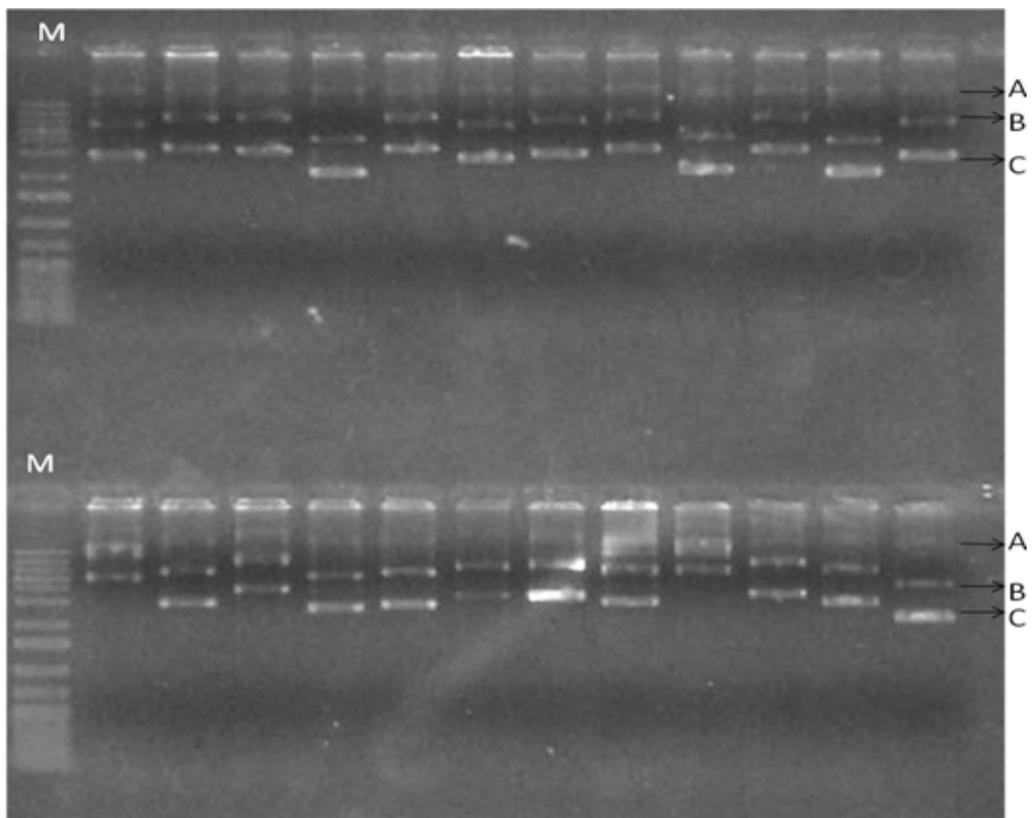


Figura 3. Gel de agarose 1% para análise da qualidade da minipreparação das placas 029 e 030. A) Plasmídeo relaxado; B) Plasmídeo enovelado; C) Plasmídeo *supercoil*; M) Marcador molecular de 1 Kb.

Após sequenciamento, os *reads* foram filtrados por tamanho e qualidade gerando a tabela 1, e depois mascarados contra *primer* e vetor (pré-processamento, tabela 2), restando 392 *reads*, dentre eles, 294 *singlets* e 33 *contigs* foram formados.

Tabela 1. Tabela gerada pelo programa Egene, após o procedimento de limpeza.

SEQUÊNCIAS BOAS			
PROGRAMA	DISCRIMINAÇÃO	SEQUÊNCIAS	PORCENTAGEM
filter_size.pl	Filtro de tamanho	3	0,65%
filter_quality.pl	Filtro de qualidade	70	15,05%
SEQUÊNCIAS BOAS		392	84,30%
TOTAL		465	100%

Tabela 2. Após executar o programa CAP3.

SEQUÊNCIAS	CONTIGS	SINGLETs	REDUNDÂNCIA DOS READS	REDUNDÂNCIA DAS BASES
392	33	294	16.58%	13.55%

O arquivo de saída do Egene, com os *reads* montados, filtrados e mascarados, foram comparados com o banco de dados do Blast. Este resultou em 104 proteínas (31,8% do total das proteínas obtidas) que não obtiveram similaridade (*no hits*) com nenhuma outra depositada (figura 4).

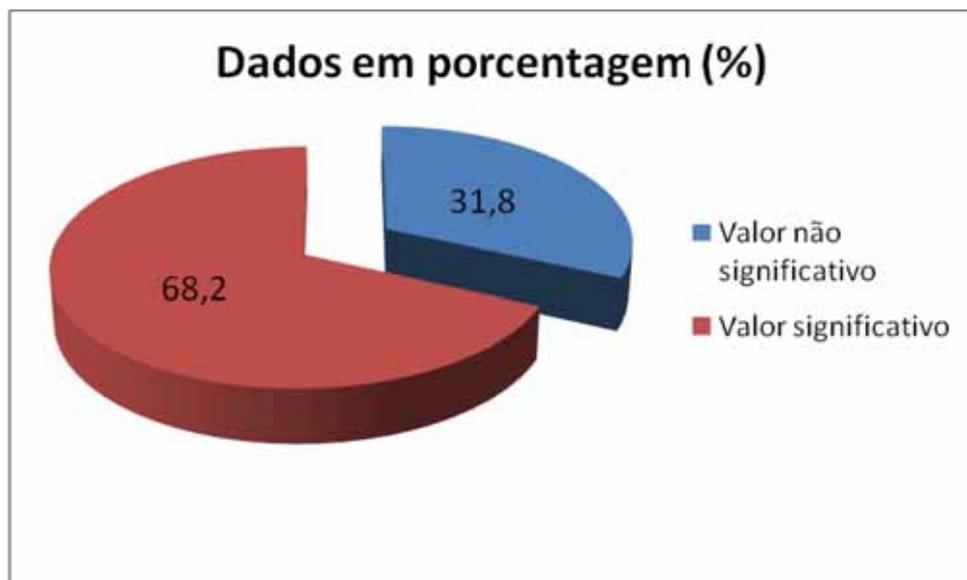


Figura 4. Gráfico em porcentagem das proteínas com “*no hits*”.

Foi possível criar, também, a figura 5 com os organismos que tiveram os menores E-values, sendo o de maior número foi descrita como *Camponotus floridanus*, com 125 das 327 proteínas encontradas no trabalho.

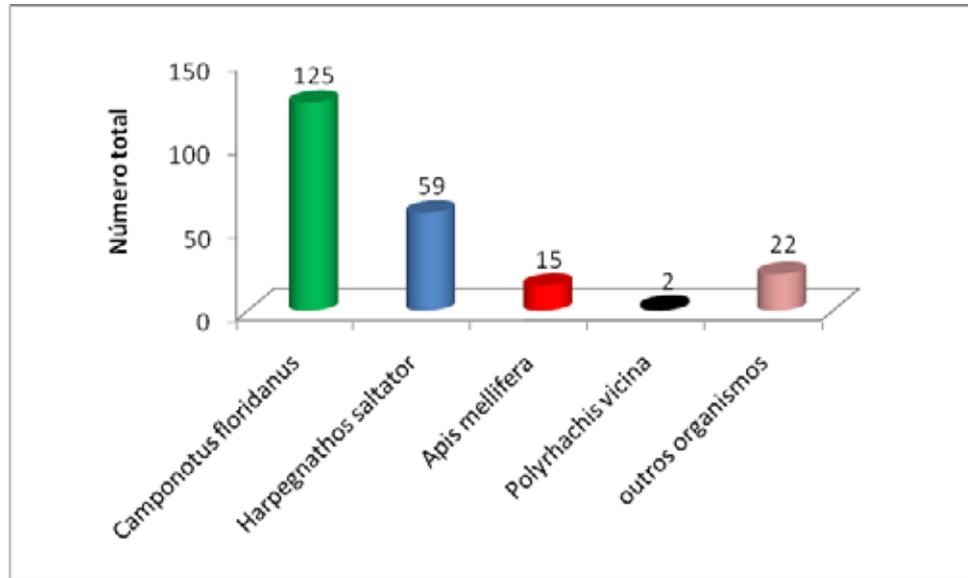


Figura 5. Organismos com menores índices no Blast.

Obtive algumas proteínas hipotéticas (figura 6). Estas somaram 18 (5,5%) de um conjunto de 327 proteínas totais.



Figura 6. Porcentagem das proteínas hipotéticas em relação às demais.

Ao buscar os termos GO, foi possível a criação da figura 7, que é subdividida em três grandes grupos: função celular (verde), processo biológico (azul) e componente celular (vermelho).

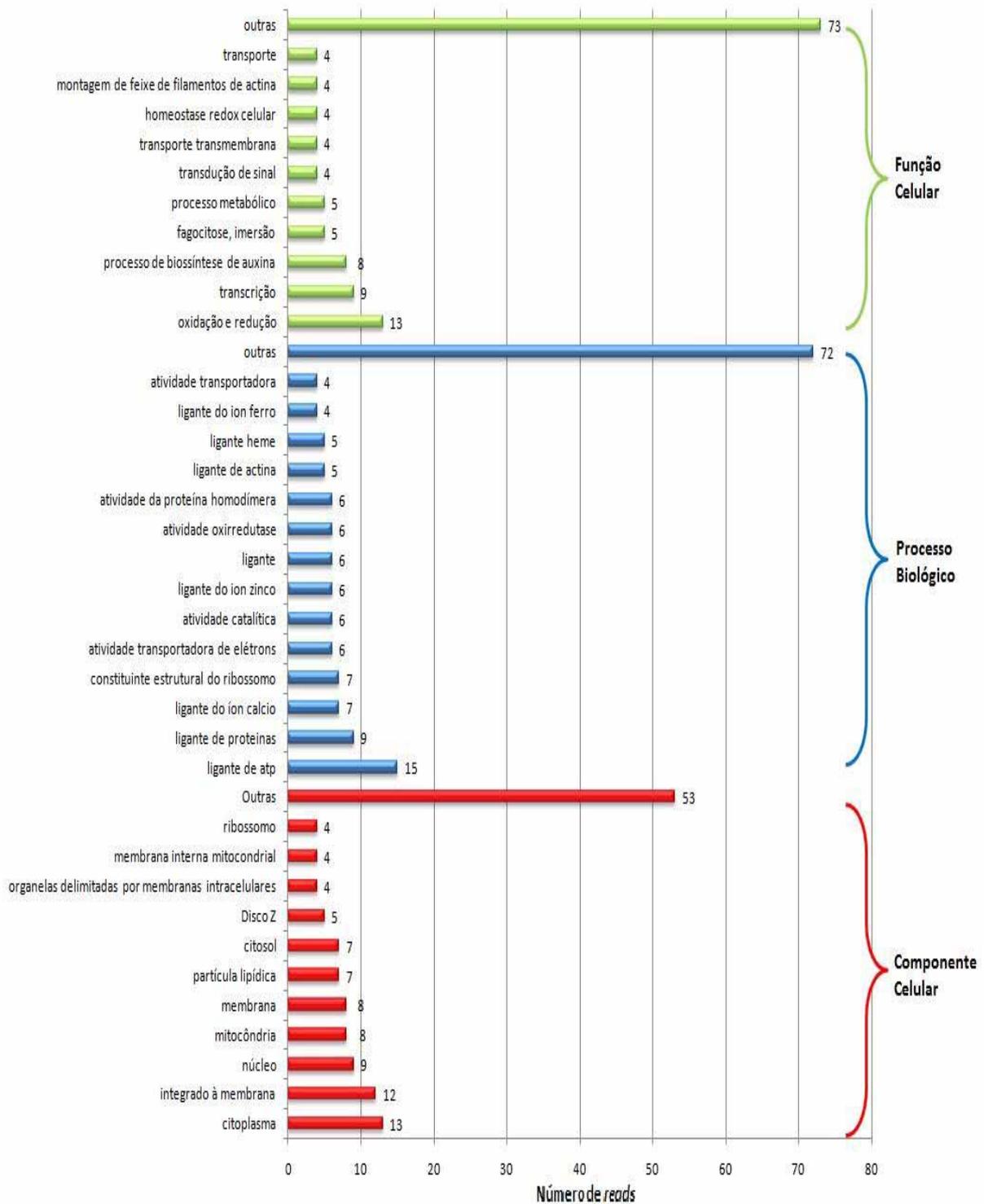


Figura 7. Gráfico com os atributos mais decorrentes do GO, nos três grandes grupos.

Das 327 proteínas, o programa InterProScan encontrou 210 com valores do IPS, ou seja, que possuem domínio protéico e dentre elas, 105 apresentavam GO (figura 8).

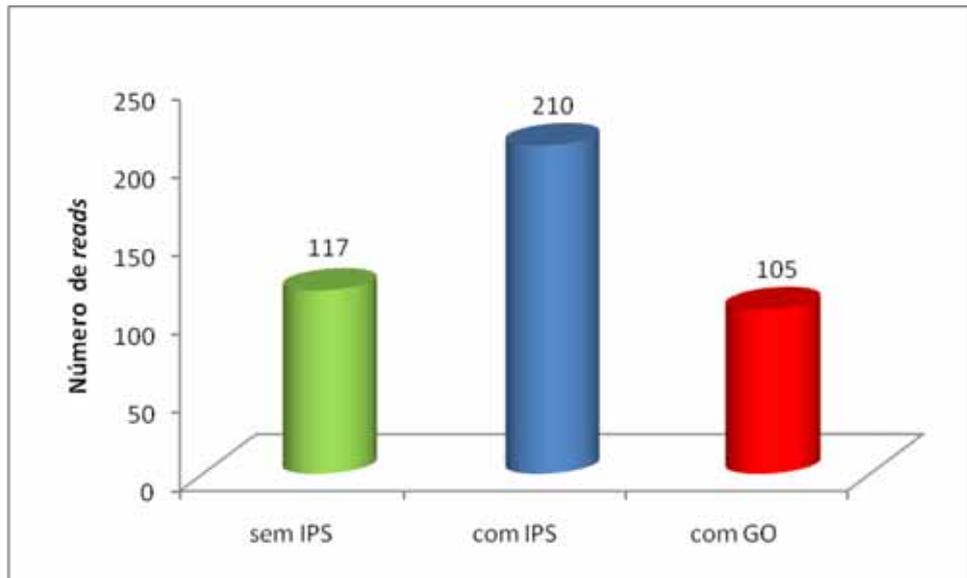


Figura 8. Estatística feita pelo programa InterProScan.

O programa Enzyme Commission localizou apenas 49 proteínas (14,987%) com valores EC (figura 9).

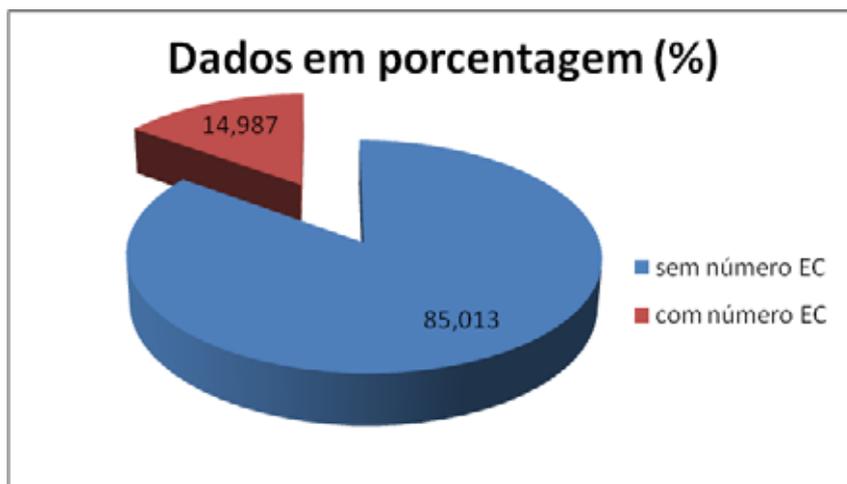


Figura 9. Gráfico em porcentagem das proteínas que possuem e não possuem número EC.

Por fim, realizei uma anotação manual de todas as proteínas, criando grupos de afinidade, resultando no gráfico da figura 10.

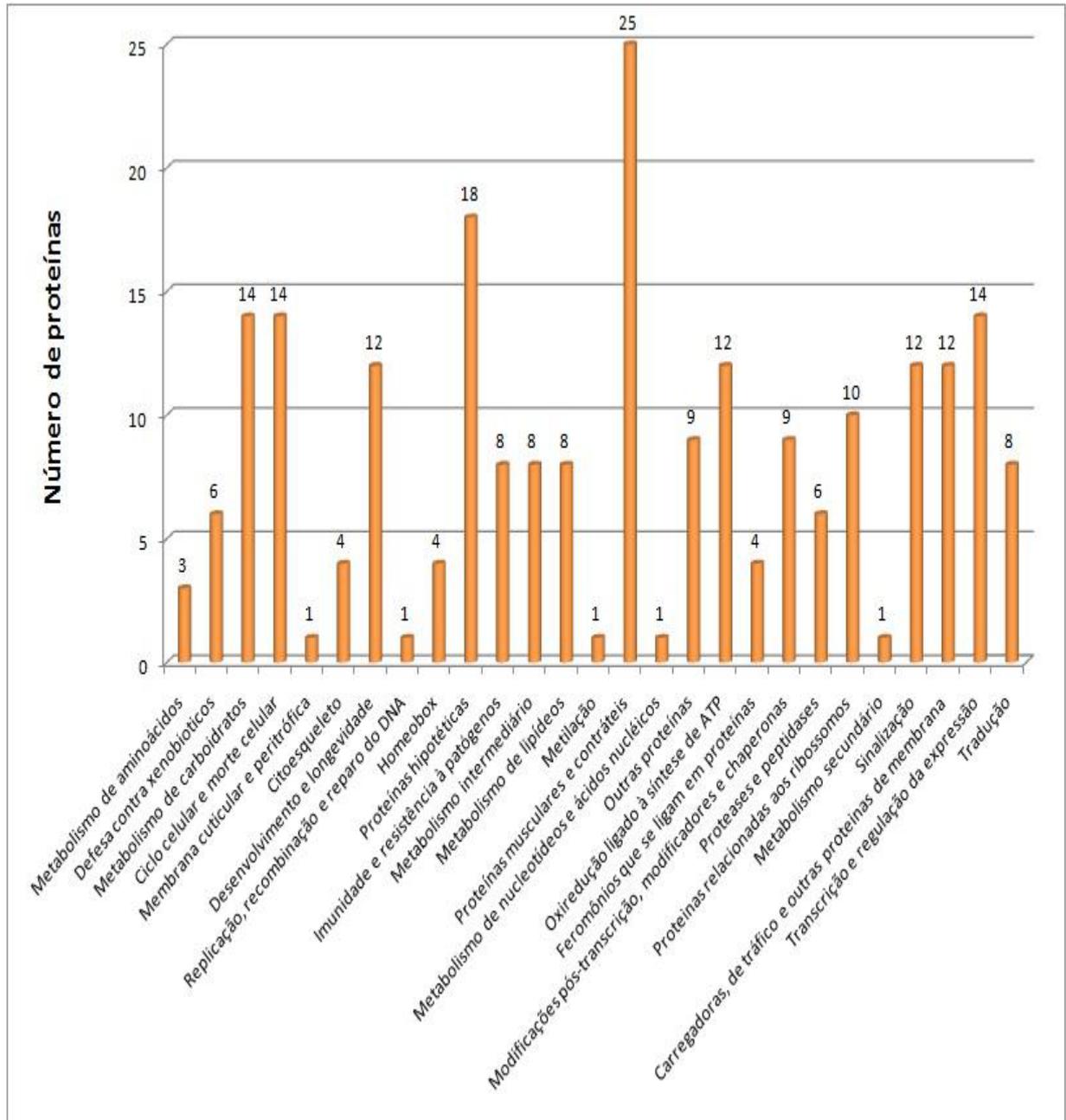


Figura 10. Gráfico com as proteínas agrupadas por afinidade resultante da análise manual.

5 DISCUSSÃO

De acordo com os géis, as minipreparações foram bem feitas, evidenciadas pelo pouco arraste (degradação), pouquíssima proteína no “pocinho” e sem nuvem de RNA no final da corrida.

As proteínas obtidas foram em sua maioria similares às de *Camponotus floridanus* e de *Harpegnathos saltator*, duas espécies de formigas, mas ainda assim, algumas foram similares às de *Apis mellifera*.

O processamento computacional selecionou as sequências de alta qualidade, eliminando erros analíticos. Desta forma, a grande quantidade de “no hits” provavelmente indica um grande número de novos genes prospectados em *A. laevigata*.

6 CONCLUSÃO

Um grande número de genes de *A. laevigata* foi clonado e anotado, gerando material importante para estudos aplicados com estas formigas pragas agrícolas.

7 – REFERÊNCIAS BIBLIOGRÁFICAS

ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** Nucleic Acids Research, v. 25(17), p. 3389-3402, 1997.

AMANTE, E. **Prejuízos causados pela formiga saúva em plantações de Eucalyptus e Pinus no Estado de São Paulo.** Silvicultura em S. Paulo, São Paulo, v. 6, n. único, p. 355-363, 1967.

BATRA, L. R. **First authenticated North American record of *Monilinia fructigena*, with notes on related species.** Mycotaxon, v. 8, p. 476-484, 1979.

BORGMEIER, T. **Estudos sobre *Atta* (Hym. Formicidae).** Memórias do Instituto Oswaldo Cruz, v. 48, p. 237-271, 1950.

CHAPELA, I. H.; REHNER, S. A.; SCHULTZ, T. S.; MUELLER, U. G. **Evolutionary history of the symbiosis between fungus-growing ants and their fungi.** Science, v.266, p.1691-1695, 1994.

CONESA, A.; GÖTZ, S. **Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics.** International Journal of Plant Genomics, v. 2008, 12 pgs., 2008

DELABIE, J. H. C.; LACAU, S.; NASCIMENTO, I. C.; CASIMIRO, A. B.; CAZORLA, I. M. **Communauté des fourmis des souches d'arbres morts dans trois réserves de la**

forêt Atlantique brésilienne (Hymenoptera, Formicidae). Ecol. v. 7, p. 95-103, 1997.

DURHAM, A. M.; KASHIWABARA, A. Y.; MATSUNAGA, F. T.; AHAGON, P. H.; RAINONE, F.; VARUZZA, L.; GRUBER, A. **EGene: a configurable pipeline generation system for automated sequence analysis.** Bioinformatics, v. 21(12), p. 2812-2813, 2005.

HEBLING, M. J. A.; BUENO, O. C.; MAROTI, P. S.; PAGNOCCA, F. C.; da SILVA, O. **A. Effects of leaves of Ipomoea batatas (Convolvulaceae) on nest development and on respiratory metabolism of leaf-cutting ants Atta sexdens L. (Hym., Formicidae).** Journal of Applied Entomology, v. 124, p. 249-252, 2000.

HERVEY, A.; NAIR, M. S. R. **Antibiotic Metabolite of a Fungus Cultivated by Gardening Ants.** Mycologia, Mycological Society of America, v. 71, n. 5, p. 1064-1066, 1979.

HOFFMAN, E. A.; GOODISMAN, M. A. D. **Gene expression and the evolution of phenotypic diversity in social wasps.** BMC Biology, v.5, n. 23, p. 1-9, 2007.

HÖLLDOBLER, B.; WILSON, E. O. **The Ants.** The Belknap Press of Harvard University, Cambridge, MA, p. 732, 1990.

HUANG, X.; MADAN, A. **CAP3: A DNA sequence assembly program.** Genome Research, v. 9(9), p. 868-877, 1999.

MUELLER, U. G.; SCHULTZ, T. R.; CURRIE, C. R.; ADAMS, R. M. M.; MALLOCH,

D. **The origin of the attine ant-fungus mutualism.** Quarterly Review of Biology, v. 76, p. 169-197, 2001.

PAGE JR, R. E.; ERBER, J. **Levels of behavioral organization and the evolution of division of labor.** Naturwissenschaften, n. 89, p. 91-106, 2002.

QUEVILLON, E.; SILVENTOINEN, V.; PILLAI, S.; HARTE, N.; MULDER, N.; APWEILER, R.; LOPEZ, R. **InterProScan: protein domains identifier.** Nucleic Acids Research, v. 33, w. 116-120, 2005.

SALEMA, R. **Biotecnologia vegetal: algumas técnicas e aplicações.** Boletim de Biotecnologia, Bio-opinião, n. 64, p 30-39, 1999.

SCHULTZ, T. R.; BRADY, S. G. **Major evolutionary transitions in ant agriculture.** PNAS, v.105, n. 14, p. 5435-5440, 2008.

SMITH, F. **Catalogue of the hymenopterous insects in the collection of the British Museum.** London, part VI Formicidae, p. 216, 1858.

SOLOMON, S. E. **Biogeography and Evolution of Widespread Leafcutting Ants, Atta Spp. (Formicidae, Attini).** University of Texas at Austin, Dissertation for the Degree of Doctor of Philosophy, august, 2007.

THE GENE ONTOLOGY CONSORTIUM. **Gene ontology: tool for the unification of biology.** Nat. Genet. v. 25, n. 1, p. 25-29, 2000.

TROPPEMAIR, R. H. **Estudo zoogeográfico e ecológico do gênero *Atta* (Hymenoptera), com ênfase sobre *Atta laevigata* Smith, 1858, no Estado de São Paulo.** (Tese - Livre Docência). Rio Claro: FFCL, p. 186, 1973.

VASCONCELOS, H. L. **Distribution of *Atta* (Hymenoptera - Formicidae) in “terra-firme” rain forest of central Amazonia: density, species composition and preliminary results on effects of forest fragmentation.** *Acta Amazonica*, v. 18, p. 309-315, 1988.

ROSS, K. G.; KELLER, L. **Genetic control of social organization in an ant.** *Proc Natl Acad Sci USA*, v. 95, p. 14232-14237, 1998.

WETTERER, J. K.; SCHULTZ, T. R.; MEIER, R. **Phylogeny of fungus-growing ants (tribe Attini) based on mtDNA sequence and morphology.** *Mol Phylogenet Evol*, v. 9, p. 42-47, 1998.

WILSON, E. O. **The arboreal ant fauna of Peruvian Amazon forest: a first assessment.** *Biotropica*, v. 19, p. 245-251, 1987.

WOODS, D.; CRAMPTON, J.; CLARKE, B.; WILLIAMSON, R. **The construction of a recombinant cDNA library representative of the poly(A)⁺ mRNA population from normal human lymphocytes.** *Nucleic Acids Research*, v. 8, n. 22, p. 5157-5168, 1980.