

**Universidade Federal de Pernambuco**  
**Centro de Ciências Biológicas**  
**Programa de Pós-Graduação em Genética**

**Marco Antonio de Oliveira dos Santos**

**Reconstrução Filogenética do Filo Arthropoda Baseada no  
Genoma Mitocondrial**

**Recife**

**2012**

**Marco Antonio de Oliveira dos Santos**

**Reconstrução Filogenética do Filo Arthropoda Baseada  
no Genoma Mitocondrial**

Tese apresentada ao Programa de Pós-Graduação  
em Genética da Universidade Federal de  
Pernambuco como parte dos requisitos exigidos  
para obtenção do título de Mestre em Genética.

**Orientador:** Valdir de Queiroz Balbino

**Recife**

**2012**

Catálogo na Fonte:  
Bibliotecário Bruno Márcio Gouveia, CRB-4/1788

S237r Santos, Marco Antonio de Oliveira dos  
Reconstrução filogenética do filo Arthropoda baseada no genoma mitocondrial /  
Marco Antonio de Oliveira Santos. – Recife: O Autor, 2012.

90 f. : il., fig., tab.

Orientador: Valdir de Queiroz Balbino

Tese (doutorado) – Universidade Federal de Pernambuco. Centro de  
Ciências Biológicas. Pós-graduação em Genética, 2012.

Inclui bibliografia e anexos

1. Artrópode 2. Filogenia 3. Genoma I. Balbino, Valdir de Queiroz  
(orientador) II. Título.

595

CDD (22.ed.)

UFPE/CCB-2013-225

**Marco Antonio de Oliveira dos Santos**

**Reconstrução Filogenética do Filo Arthropoda Baseada no  
Genoma Mitocondrial**

**Aprovado em 06/08/2012**

**Banca Examinadora:**

---

**Dr. Valdir de Queiroz Balbino  
Universidade Federal de Pernambuco**

---

**Dr. Marcos Antônio de Moraes Junior  
Universidade Federal de Pernambuco**

---

**Dr. Martin Alejandro Montes  
Universidade Federal Rural de Pernambuco**

**Recife  
2012**

A todos que acreditam em mim.

## **Agradecimentos**

À minha família que me apoiou e incentivou a sempre buscar o melhor de mim.

À minha noiva, Clarice, por sempre ficar ao meu lado, mesmo nas horas mais difíceis, e ter se mostrado uma eterna companheira e sempre compreensiva.

Ao professor e orientador Valdir Balbino, pela oportunidade, enorme paciência e vontade de indicar os melhores caminhos.

Ao pessoal do laboratório pelo companheirismo, troca de conhecimento e bom humor. Principalmente a Carlos, Marcus, Moisés e Tiago que tornaram esse trabalho mais fácil de realizar.

Ao CNPQ pela ajuda financeira e ao LABBE pela disponibilidade da infraestrutura.

## Resumo

Os Arthropoda formam o grupo animal mais abundante e diverso na terra, com mais de um milhão de espécies descritas, apresentando extrema importância econômica, ambiental e médica. Devido a grande diversidade e antiguidade do grupo, inúmeras hipóteses para explicar a história evolutiva dos Arthropoda foram sugeridas ao longo dos anos. No entanto, as relações entre e dentro dos quatro principais grupos do filo (Crustacea, Hexapoda, Myriapoda e Chelicerata) permanecem como uma das grandes questões em aberto na sistemática. Neste contexto, o genoma mitocondrial representa uma importante fonte de informação para a reconstrução filogenética de um grupo taxonômico tão diverso. Assim, esse trabalho teve como objetivo elucidar as relações filogenéticas do filo Arthropoda a partir dos genomas mitocondriais, completamente sequenciados e disponíveis em bancos de dados públicos, utilizando diferentes abordagens computacionais. As análises dos genomas mitocondriais sugerem, com altos valores de suporte, o monofilétismo dos grupos Pancrustacea (Hexapoda + Crustacea) e Mandibulata (Pancrustacea + Myriapoda) indicando que o grupo irmão dos demais Arthropoda é Chelicerata. Além disso, a utilização da entropia de Shannon se mostrou altamente eficiente na seleção de regiões filogeneticamente informativas, sendo mais efetiva que outros métodos utilizados tradicionalmente na análise de genomas mitocondriais. A entropia também se mostrou como ótima fonte de informações para seleção de regiões conservadas destinadas à determinação de iniciadores degenerados para a amplificação de regiões homólogas. Esse trabalho representa um importante passo para a elucidação da história filogenética do filo Arthropoda.

**Palavras-chave:** Arthropoda; Filogenia Molecular; Genoma Mitocondrial; Entropia.

## Abstract

Arthropoda are the most abundant and diverse animal group on earth, with over a million described species. They have very important economic and medical, environmental. Given the great diversity and seniority of the group, several hypotheses to explain the evolutionary history of Arthropoda have been suggested over the years. The relationships between and within these four major lineages of the phylum (Crustacea, Hexapoda, Myriapoda and Chelicerata) lineages remain one of the most contentious issues in systematics. The mitochondrial genome represents an important source of information for phylogenetic reconstructions of a taxonomic group as diverse. Thus, this study aimed to elucidate the phylogenetics relationships of the phylum Arthropoda by use the completely sequenced and available mitochondrial genomes in public databases, using different computational approaches. Analyses of mitochondrial genomes suggest with high values support the monophyly of groups Pancrustacea (Hexapoda + Crustacea) and Mandibulata (Pancrustacea + Myriapoda) indicating that Chelicerata is sister group of other Arthropoda. Furthermore, the use of Shannon's entropy proved highly effective in the selection of phylogenetically informative regions. Shannon's entropy was more effective than other methods traditionally used in the analysis of mitochondrial genomes. The entropy also showed like great source of information for selection of conserved regions to determination of degenerate primers. This work represents an important step to elucidate the phylogenetics history of the phylum Arthropoda.

**Key words:** Arthropoda; Molecular phylogeny; Mitochondrial Genome; Entropy.



## Lista de Ilustrações

<b>Figura 1.</b> Diversidade do filo Arthropoda.	3
<b>Figura 2.</b> Esquema genoma mitocondrial.	10
<b>Figura 3.</b> Árvore de UPGMA baseada na ordem gênica do genoma mitocondrial dos Arthropoda.	33
<b>Figura 4.</b> Gráficos mostrando os valores de entropia para os genes ATP6, ATP8, COX1, COX2, COX3, CYTB, ND1 e ND2.	40
<b>Figura 4.</b> Gráficos mostrando os valores de entropia para os genes ND3, ND4, ND4L, ND5 e ND6.	41
<b>Figura 5.</b> Gráficos mostrando a saturação das sequências de DNA iniciais e para os dois métodos de seleção empregados nos genes ATP6, COX1 e COX2.	43
<b>Figura 6.</b> Gráficos mostrando a saturação das sequências de DNA iniciais e para os dois métodos de seleção empregados nos genes COX3, CYTB e ND1.	44
<b>Figura 8.</b> Gráficos mostrando a saturação das sequências de DNA iniciais e para os dois métodos de seleção empregados nos genes ND3, ND4 e ND5.	45
<b>Figura 7.</b> Gráfico mostrando os valores de entropia para os genomas mitocondriais da ordem Diptera.	47
<b>Figura 8.</b> Árvore filogenética da ordem díptera baseada nas regiões amplificadas com iniciadores degenerados.	50
<b>Figura 9.</b> Árvore filogenética do filo Arthropoda baseada nas regiões selecionadas por entropia.	53
<b>Figura 10.</b> Continuação da árvore filogenética do filo Arthropoda baseada nas regiões selecionadas por entropia.	54
<b>Figura 11.</b> Árvore filogenética do filo Arthropoda baseada nos blocos conservados selecionadas.	55

**Figura 12.** Continuação da árvore filogenética do filo Arthropoda baseada nos blocos conservados selecionadas.

56

## Lista de Tabelas

<b>Tabela 1.</b> Variação no viés GC% dos genes mitocondriais	34
<b>Tabela 2.</b> Variabilidade dos genes mitocondriais.	36
<b>Tabela 3.</b> Comparação entre número de sítios identificados como de baixa entropia e pertencentes a blocos conservados.	38
<b>Tabela 4.</b> Iniciadores degenerados definidos para a amplificação dos genomas mitocondriais de dípteros.	48
<b>Tabela 5.</b> Modelos evolutivos aplicados na reconstrução filogenética dos Arthropoda.	51

## Lista de Abreviaturas

ATP	Adenosina Trifosfato
AIC	<i>Akaike Information Criterion</i>
ATP6	ATP F0 Sintase Subunidade 6
ATP8	ATP F0 Sintase Subunidade 8
COX1	Subunidade 1 da citocromo c oxidase
COX2	Subunidade 2 da citocromo c oxidase
COX3	Subunidade 3 da citocromo c oxidase
CYTB	Citocromo B
DNA	Ácido Desoxirribonucléico
EST	Etiquetas de Genes Expressos
MCMC	Cadeia de Markov de Monte Carlo
MitDNA	DNA Mitocondrial
NCBI	Centro Nacional para Informação Biotecnológica
ND1	Subunidade 1 da NADH Desidrogenase
ND2	Subunidade 2 da NADH Desidrogenase
ND3	Subunidade 3 da NADH Desidrogenase
ND4	Subunidade 4 da NADH Desidrogenase
ND4L	Subunidade 4L da NADH Desidrogenase
ND5	Subunidade 5 da NADH Desidrogenase
ND6	Subunidade 6 da NADH Desidrogenase
NNI	<i>Nearest Neighbor Interchange</i>

pb	Pares de Bases
PCR	Reação da Polímera em Cadeia
RNA	Ácido Ribonucléico
RNA <sub>t</sub>	Ácido Ribonucléico Transportador
UPGMA	Método de Agrupamento Pareado Não Ponderado com Média Aritmética

## Sumário

Resumo	vi
Abstract	vii
Lista de ilustrações	viii
Lista de tabelas	ix
Lista de abreviaturas	xi
1. Introdução	1
2. Revisão da Literatura	3
2.1. Filo Arthropoda	3
2.1.1 Evolução dos Arthropoda	4
2.1.2. Ordem Diptera	7
2.1.2.1.Gênero <i>Chrysomya</i> (Diptera: Calliphoridae)	9
2.2. Genoma Mitocondrial	9
2.2.1. Genoma Mitocondrial em Estudos Evolutivos	11
2.2.2. Ordem Genética em Mitocôndrias	12
2.3. Filogenia Molecular	13
2.3.1. Métodos para Seleção de Regiões	16

<b>Filogeneticamente Informativas</b>	
<b>2.4. Genômica Comparativa</b>	<b>19</b>
<b>3. Objetivos</b>	<b>21</b>
<b>4. Material e Métodos</b>	<b>22</b>
<b>4.1. Criação do Banco de Dados</b>	<b>22</b>
<b>4.1.1. Construção do Banco de Dados Local</b>	<b>22</b>
<b>4.2. Seleção de Genomas</b>	<b>23</b>
<b>4.3. Análise de Sintenia</b>	<b>24</b>
<b>4.4. Alinhamento de Sequências</b>	<b>24</b>
<b>4.5. Análises de Variabilidade</b>	<b>25</b>
<b>4.6. Identificação de Regiões Problemáticas</b>	<b>26</b>
<b>4.7. Definição de Iniciadores Para Amplificação do Genoma         Mitochondrial de Diptera</b>	<b>27</b>
<b>4.8. Experimentos de Biologia Molecular</b>	<b>28</b>
<b>4.8.1. Tratamento e Análise das Sequências             Amplificadas</b>	<b>29</b>
<b>4.9. Análises Filogenéticas</b>	<b>29</b>
<b>5. Resultados</b>	<b>31</b>
<b>5.1. Banco de Dados Local</b>	<b>31</b>
<b>5.2. Análise de Sintenia</b>	<b>31</b>

<b>5.3. Análise do Viés GC%</b>	<b>34</b>
<b>5.4. Análise de Variabilidade</b>	<b>36</b>
<b>5.5. Análise Comparativa Entre Métodos de Seleção de Regiões Filogeneticamente Informativa</b>	<b>37</b>
<b>5.6. Definição de Iniciadores Degenerados para Amplificação do Genoma Mitocondrial da Ordem Diptera</b>	<b>46</b>
<b>5.6.1. Teste dos Iniciadores Definidos</b>	<b>49</b>
<b>5.7. Reconstrução Filogenética do Filo Arthropoda</b>	<b>51</b>
<b>6. Discussão</b>	<b>57</b>
<b>7. Conclusões</b>	<b>64</b>
<b>8. Referências Bibliográficas</b>	<b>65</b>
<b>9. Anexo 1. Viés GC% dos genes mitocondriais</b>	<b>79</b>
<b>10. Currículo Lattes</b>	<b>89</b>



## 1. Introdução

Os Arthropoda formam o grupo animal mais abundante e diverso na terra, com mais de um milhão de espécies descritas. Apresentam extrema importância econômica, ambiental e médica, sendo os únicos organismos multicelulares conhecidos que comprovadamente afetaram a evolução humana.

Os Arthropoda também são importantes na reciclagem de nutrientes e na polinização de espécies vegetais. Várias espécies do grupo são pragas de espécies vegetais cultivadas e de animais de criação. Além disso, os Arthropoda também são reconhecidos como vetores dos agentes etiológicos de inúmeras doenças humanas, como leishmanioses, arboviroses, Doença de Chagas, entre outras.

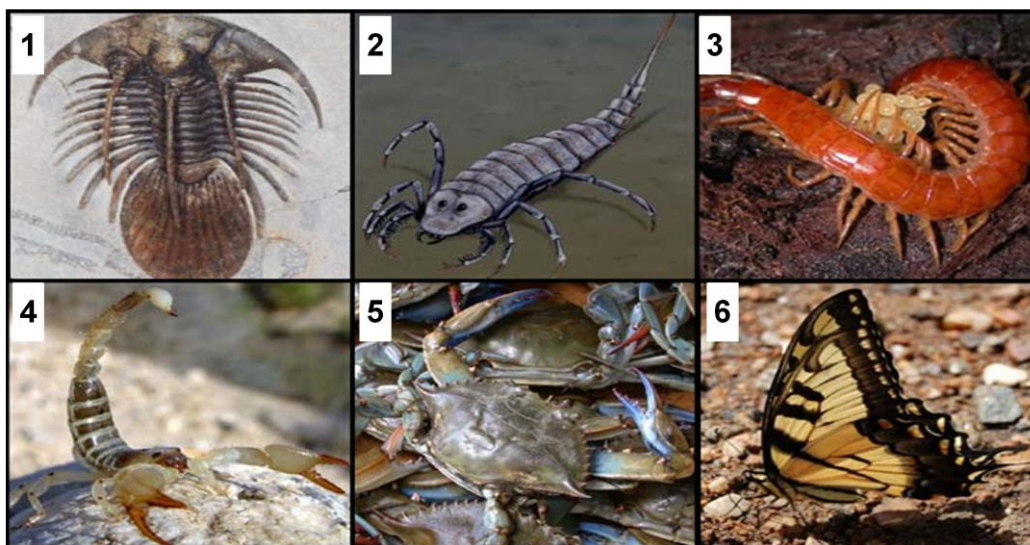
Os primeiros membros do filo surgiram provavelmente nos oceanos do Precambriano, há cerca de 600 milhões de anos. Devido a sua grande diversidade e antiguidade, inúmeras hipóteses para explicar a história evolutiva dos Arthropoda foram sugeridas ao longo dos anos. No entanto, as relações entre e dentro dos quatro principais grupos do filo (Crustacea, Hexapoda, Myriapoda e Chelicerata) permanecem como uma das grandes questões em aberto na sistemática. Neste contexto, o genoma mitocondrial representa uma importante fonte de informação para a reconstrução filogenética de um grupo taxonômico tão diverso. O DNA mitocondrial oferece algumas vantagens em relação aos outros marcadores moleculares utilizados em análises filogenéticas: ausência de parálogos; padrão de herança uniparental; e taxas

de substituição diferentes da maioria dos genes nucleares. Adicionalmente, existe um grande número de genomas mitocondriais nos bancos de dados públicos de sequências biológicas. Assim, novas abordagens no uso da informação derivada da análise do genoma mitocondrial representam uma forma de contribuir para a compreensão da história evolutiva do filo Arthropoda.

## 2. Revisão de Literatura

### 2.1. Filo Arthropoda

Os Arthropoda (Figura 1) representam o grupo animal mais abundante e diverso na Terra, com mais de um milhão de espécies descritas. Tradicionalmente, os Arthropoda são divididos em grandes subfilos: Crustacea (e.g. caranguejos, camarões, etc.); Hexapoda (e.g. insetos, colembolas, etc.); Myriapoda (e.g. centopeias e diplópodes); e Chelicerata (e.g. caranguejos-ferradura e aracnídeos). As primeiras espécies do grupo surgiram provavelmente no período Pré-cambriano. O grupo sofreu uma forte radiação adaptativa, e hoje, seus representantes ocorrem em praticamente todos os ambientes da terra, explorando um grande número de estilos de vida e habitats. Os Arthropoda modernos variam de pequenos ácaros, com menos de um milímetro, a crustáceos, como o caranguejo aranha japonês, que pode alcançar mais de três metros (Brusca e Brusca, 2003).



**Figura 13.** Diversidade do filo Arthropoda. 1 Trilobitomorpha (subfilo extinto). 2 Chelicerata. 3 Myriapoda. 4 Chelicerata. 5 Crustacea. 6 Hexapoda. (Fonte: <http://www.brighthub.com>).

Grimaldi e Engel (2005) destacaram algumas das características morfológicas compartilhadas pelos Arthropoda, a saber: segmentação interna e externa do corpo, com especialização regional (ou tagmosis); exoesqueleto composto por uma cutícula que é endurecida por meio de calcificação ou por esclerotização; exoesqueleto formado por placas articuladas; apêndices articulados; celoma reduzido a partes do trato reprodutivo e sistema excretor (a cavidade do corpo principal é uma hemocele aberta); sistema circulatório aberto com coração dorsal; trato digestivo completo; cordão nervoso ventral; crescimento por etapas através de mudas.

### **2.1.1. Evolução dos Arthropoda**

Há muito tempo é reconhecido que dois grupos vivos, os Onychophora (vermes aveludados de corpo mole) e os microscópicos Tardigrada (ursos d'água) são evolutivamente relacionados aos Arthropoda. Os membros destes grupos são formados por animais segmentados, com presença de apêndices e são coletivamente referidos como Panarthropoda (Brusca e Brusca, 2003). Todas as evidências morfológicas e moleculares suportam a existência de um grupo monofilético formado pelos Arthropoda e Onychophora, sendo que a posição taxonômica dos Tardigrada ainda é discutível. Apesar de serem tradicionalmente considerados como os parentes vivos mais próximos dos Euarthropoda, alguns estudos de filogenia molecular colocam os Tardigrada como os Panarthropoda mais basais, ou mesmo como grupo irmão de Nematoda. Entretanto, Budd e Telford (2009) consideram que a dificuldade de determinar a posição filogenética de Tardigrada está relacionada com artefatos causados pela rápida evolução das regiões genômicas utilizadas nos estudos já realizados.

Devido à grande diversidade dos Arthropoda, por muito tempo houve um considerável debate sobre a condição monofilética ou polifilética do grupo. No entanto, como revisado por Budd e Telford (2009), no início da década de 1990 a aplicação de ferramentas de sistemática molecular encerrou a discussão, favorecendo a hipótese acerca da condição monofilética do grupo. Em outras palavras, muitas das características morfológicas compartilhadas pelos Arthropoda tiveram uma única origem e, a partir desse ponto, diversificaram-se ao longo do tempo evolutivo.

As relações entre e dentro dos quatro grupos de Arthropoda permanecem como uma das grandes questões em aberto da sistemática, e diferentes hipóteses têm sido propostas na literatura (Hassanin, 2006). A monofilia dos Chelicerata, por exemplo, é tida como um dos assuntos menos controversos na sistemática dos Arthropoda, sendo suportada por uma série de estudos morfológicos e moleculares. Por outro lado, o estado parafilético de Myriapoda (em relação aos hexápodes) é comumente defendido em favor da hipótese do grupo Atelocerata. No entanto, a união entre Myriapoda e Hexapoda é fortemente rejeitada em vários estudos moleculares que também suportam o reconhecimento do estado monofilético de Myriapoda (Edgecombe, 2010).

A monofilia dos Crustacea é normalmente suportada por análises cladísticas de dados morfológicos. Porém, com o passar do tempo, a discussão do estado parafilético do grupo, em relação aos hexápodes, passou a ser recorrente, com o grupo irmão dos Hexapoda variando entre diferentes grupos de Crustacea, a depender da natureza do estudo realizado. Ainda com relação aos hexápodes, os estudos morfológicos e de genes nucleares apontam para a monofilia do grupo, enquanto que os dados de genes mitocondriais suportam um estado parafilético ou inconclusivo para o subfilo (Edgecombe, 2010).

Uma união entre Chelicerata e Myriapoda, como grupos irmãos, foi sugerida inicialmente na década de 1990, a partir da análise de genes do RNAs ribossomais

nucleares (Friedrich e Tautz, 1995; Giribet *et al.*, 1996). Vários estudos posteriores usando diferentes dados moleculares tais como sequências do gene HOX, de hemocianinas, de genomas mitocondriais e dados de expressão gênica suportaram a formação de um grupo que une os dois subfilos, chamado Paradoxopoda ou Myriochelata. Os dados morfológicos que suportam o monofiletismo de Paradoxopoda são provenientes de estudos focados na embriogênese e na formação do sistema nervoso dos Arthropoda e Onychophora (Edgecombe, 2010).

Um grupo dos Arthropoda em que o apêndice pós-tritocerebral assume a forma de um maxilar ou mandíbula tem proporcionado a tradicional base para unir Myriapoda, Crustacea e Hexapoda em um clado chamado Mandibulata (Edgecombe, 2010). A homologia dos Mandibulata é suportada tanto por dados morfológicos (Bitsch, 2001; Edgecombe *et al.*, 2003) quanto por dados de expressão gênica (Prpic e Tautz, 2003). A hipótese dos Mandibula exclui a formação do grupo Paradoxopoda, já que nessa hipótese Chelicerata seria o grupo irmão de Mandibulata e não de Myriapoda.

A relação entre Crustacea e Hexapoda (excluindo Myriapoda), formando o grupo Pancrustacea ou Tetraconata, é um tema frequente na sistemática dos Arthropoda e foi primeiramente suportada por estudos moleculares realizados por Wheeler *et al.* (1993). Do ponto de vista morfológico, a maioria dos argumentos a favor da monofilia do grupo vem de estudos do sistema nervoso. As evidências moleculares suportam as semelhanças neuroanatômicas entre os dois subfilos, não apenas por causa dos altos valores de suporte encontrados nas análises com dados de sequências de DNA, mas também porque o grupo é independentemente recuperado por vários tipos de dados moleculares. Entre estes estão dados de genes ribossomais, genes nucleares codificantes de proteínas, genomas mitocondriais, dados de expressão e a combinação de diferentes informações moleculares (Edgecombe, 2010).

A tradicional hipótese do grupo Atelocerata ou Tracheata, formado por Myriapoda e Hexapoda, rivaliza com a ideia dos Pancrustacea na tentativa de determinar o grupo mais próximo aos hexápodes. O clado Atelocerata apresenta um grande suporte morfológico, como membros sem segmentos intercalares, endoesqueleto tensorial, órgãos pós-antenas, túbulos de Malpigi e um único músculo pré-tarsal (Edgecombe, 2010). No entanto, Atelocerata é uma hipótese baseada apenas em dados morfológicos. Nenhum estudo molecular, até agora, suporta a união de Myriapoda e Hexapoda, excluindo Crustacea e Chelicerata, em um mesmo clado [com exceção de um experimento realizado por Wägele e Stanjek (1995), com um pequeno fragmento de rRNA 12S]. A falta de suporte das análises de marcadores moleculares diminui a robustez da hipótese dos Atelocerata, uma vez que a hipótese que reconhece os Pancrustacea é suportada por informações de diversas fontes, bem como por numerosas características do sistema nervoso (Edgecombe, 2010).

### **2.1.2. Ordem Diptera**

A ordem Diptera, que compreende moscas, mosquitos e espécies afins, com cerca de 120.000 espécies, é um dos grupos de insetos mais diverso, tanto ecologicamente quanto em termos de riqueza de espécies. Os dípteros estão distribuídos por todos os continentes e têm colonizado com sucesso praticamente qualquer tipo de habitat (Courtney e Merritt, 2008). O fóssil mais antigo de um díptero data do Triássico Superior, há cerca de 225 milhões de anos atrás. Entre os dípteros, muitas espécies apresentam importância econômica e médica, como pragas de plantas e vetores de doenças perigosas para pessoas e animais (Evenhuis, 1994).

Certamente, uma das razões mais relevantes para o estudo dos dípteros diz respeito a sua importância como vetores de doenças humanas. De fato, as doenças transmitidas por mosquitos, flebotomíneos, borrachudos, tsé-tsé e outros, afetam atualmente mais pessoas do que todas as doenças transmitidas por outros Arthropoda combinadas. Doenças como malária, febre amarela, leishmanioses e doença do sono, por exemplo, têm tido um impacto profundo na história humana, alterando o curso de guerras e até mesmo a demografia humana. Na verdade, os dípteros, juntamente com as pulgas, são os únicos organismos multicelulares conhecidos por terem afetado a evolução humana (Grimaldi e Engel, 2005).

Os dípteros também apresentam grande importância para a ciência, já que a *Drosophila melanogaster* foi o primeiro eucarioto complexo a ter o seu genoma completamente sequenciado e foi usada nos primórdios da genética para aumentar o entendimento sobre evolução (Adams *et al.*, 2000). O estudo do genoma desta espécie ampliou a compreensão do papel de diferentes genes durante o desenvolvimento animal (Lindsley e Zimm, 1992).

A família Calliphoridae é formada por moscas robustas com coloração em geral brilhante e metálica. O nome mais comum usado para designar os membros desta família é varejeira. No mundo todo, existem 1100 espécies de moscas-varejeiras, com 228 espécies na região Neotropical. Os indivíduos adultos põem ovos em qualquer carne fresca ou cozida, geralmente de animais de grande porte, bem como excremento animal. Além de também causarem ferimentos graves em animais domésticos, devido ao processo de alimentação das larvas. Os adultos se alimentam e vivem em flores, plantas e animais decompostos (Grimaldi e Engel, 2005).



### 2.1.2.1. Gênero *Chrysomya* (Diptera: Calliphoridae)

Espécies do gênero *Chrysomya* (Diptera: Calliphoridae) apresentam considerável importância médica e sanitária, pois ocorrem em lixo, fezes e carcaças, podendo ser vetores mecânicos de patógenos, além de poderem atuar como causadores de miíases no homem e em animais (Baumgartner e Greenberg, 1984). Como as larvas das espécies do gênero *Chrysomya* alimentam-se de matéria orgânica em decomposição, o gênero é considerado de grande importância para estudos de biologia forense, uma vez que suas larvas podem ser utilizadas como indicadores do intervalo *post-mortem* (Hall e Wall 1995; Guimarães e Papavero, 1999).

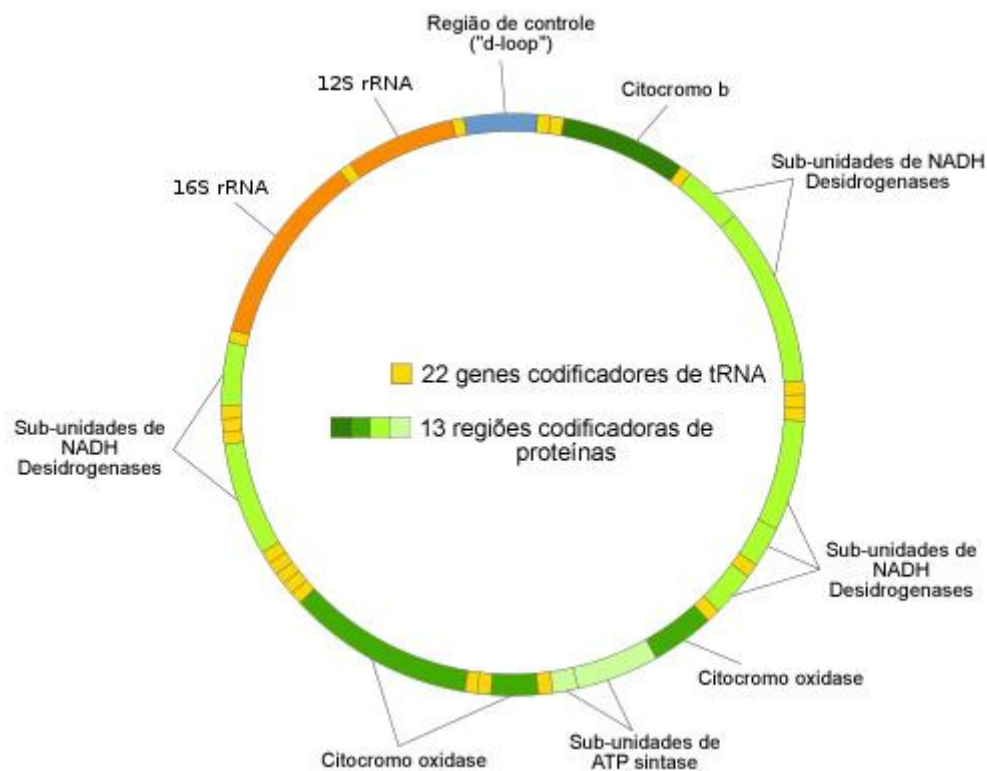
O gênero *Chrysomya* é constituído por cerca de 30 espécies nativas do Velho Mundo e Australásia; no entanto, na década de 1970, quatro espécies deste grupo foram introduzidas acidentalmente no Continente Americano: *C. megacephala*; *C. albiceps*; *C. putoria*; e *C. rufifacies*. As três primeiras dispersaram-se rapidamente na América do Sul, causando um declínio significativo em populações locais de *Cochliomyia macellaria* e de *Lucilia eximia*, espécies de califorídeos nativas do continente (Guimarães *et al.*, 1978; Baumgartner e Greenberg, 1984; Dear, 1985; Gião e Godoy, 2006).

## 2.2. Genoma Mitocondrial

As mitocôndrias são as organelas onde ocorre a fosforilação oxidativa, reação essencial para a produção de ATP, bem como uma variedade de outras funções bioquímicas. Elas desempenham um papel central na apoptose, no metabolismo, e no envelhecimento celular. Dentro destas organelas há um genoma exclusivo, separado da

cromatina nuclear, referido como DNA mitocondrial (mtDNA), comumente utilizado em estudos de filogenia molecular (Alberts *et al.*, 2008).

O genoma mitocondrial (Figura 2) varia amplamente em tamanho e conteúdo gênico em diversos grupos eucarióticos, mas a sua estrutura é notadamente uniforme entre os metazoários, apresentando uma organização gênica muito compacta, sem *introns*, geralmente com pequenas regiões intergênicas, em alguns casos, há curtas sobreposições de genes. O genoma mitocondrial típico de um metazoário é circular, com 14kb a 18kb de extensão, e codifica 13 subunidades proteicas de algumas das enzimas responsáveis pela fosforilação oxidativa, além das duas subunidades (12S e 16S) do ribossomo mitocondrial e dos 22 tRNAs necessários para a tradução das proteínas codificadas pelo mtDNA. Todos os 37 genes tipicamente encontrados no mtDNA de animais têm homólogos no mtDNA de plantas, fungos e protistas. Há também, de modo geral, uma grande região não codificante rica em AT, que, para alguns animais, é conhecida por conter elementos controladores de replicação e transcrição (Boore, 1999).



**Figura 14.** Esquema genoma mitocondrial. (Fonte: <http://en.wikipedia.org>)

O mtDNA codifica apenas 13 subunidades de proteínas envolvidas na fosforilação oxidativa (três subunidades da enzima citocromo c oxidase: COXI, COXII e COXIII; sete subunidades da NADH desidrogenase: ND1-ND6 e ND4L; citocromo b [CYTB]; e duas subunidades da ATP F0 sintase: ATP6 e ATP8). As outras subunidades envolvidas na fosforilação oxidativa, assim como as enzimas responsáveis pela replicação e transcrição do mtDNA, são codificadas por genes nucleares, sintetizadas no citoplasma celular e importados para a organela por meio de sinais específicos (Alberts *et al.*, 2008).

### **2.2.1. Genoma Mitocondrial em Estudos Evolutivos**

A comparação de genomas mitocondriais pode resultar em novas hipóteses sobre a evolução dos organismos e dos próprios genomas (Boore, 1999). O DNA mitocondrial animal oferece vantagens distintas dos outros marcadores moleculares para a análise filogenéticas. Por estarem presentes em cópias únicas, a comparação entre genes parálogos geralmente não é uma preocupação; o padrão de herança maternal, característico do mtDNA da maioria dos animais, permite a reconstrução direta de uma topologia de árvore bifurcada; a herança uniparental também reduz o tamanho efetivo da população de genes mitocondriais, implicando que as mutações são fixadas mais rapidamente; o DNA mitocondrial também tem uma taxa de substituição de bases muito maior que a maioria dos genes nucleares. Estas características fizeram do mtDNA um marcador genético bastante frequente em estudos evolutivos (Cuore e Kocher, 1999).

Muitos aspectos da evolução do genoma (*e.g.* evolução dos rearranjos gênicos; regulação gênica; sistemas de processamento de mensagens; e mecanismos de replicação) podem ser estudados em sistemas mitocondriais em níveis de detalhamento e

de compreensão superiores aos observados nos sistemas nucleares. Questões que podem ser abordadas pela genômica mitocondrial comparativa incluem: 1) quais são os padrões na evolução dos mecanismos de expressão gênica?; 2) que tipos de genes se reorganizam com mais frequência?; 3) que aspectos da variação genômica são correlacionados com os aspectos da fisiologia, de mecanismos moleculares ou da história de vida?; 4) como fatores moleculares, que interagem, co-evoluem?; 5) a taxa de evolução das sequências é correlacionada com a taxa de rearranjos gênicos dentro de uma mesma linhagem? (Boore, 1999).

No entanto, o DNA mitocondrial apresenta três principais dificuldades para a utilização em reconstruções filogenéticas de Arthropoda: 1) a saturação de mutações, em consequência da antiguidade do grupo a vários eventos de mutação em um mesmo sítio é um grande problema na reconstrução das árvores filogenéticas. E, no caso do mtDNA, a ocorrência de saturações é mais importante, pelo fato do genoma mitocondrial evoluir mais rapidamente do que o genoma nuclear (revisado por Burger *et al*, 2003); 2) As taxas de mutação são diferentes entre as linhagens e táxons que evoluem mais rápido podem se agrupar artificialmente por causa do fenômeno da atração de ramos longos (Felsenstein, 1978); 3) Outro problema que pode induzir a erro nas inferências filogenéticas, é o fato que os genes mitocondriais de alguns táxons foram afetados por mudanças importantes no padrão de substituição (Hassanin, 2005).

### **2.2.2. Ordem Genética em Mitocôndrias**

Embora as sequências mitocondriais animais sejam conhecidas por evoluir rapidamente, seus arranjos gênicos frequentemente permanecem inalterados por longos períodos evolutivos. Com poucas exceções, os arranjos gênicos são relativamente

estáveis dentro de grandes grupos, mas apresentam variação quando esses mesmos grupos são comparados entre si, dessa forma, esses arranjos têm grande potencial para elucidar as relações filogenéticas de muitos ramos dos metazoários (Boore, 1999).

O grande número de arranjos gênicos possíveis faz com que seja muito improvável que diferentes táxons adotem, de forma independente, um estado idêntico. Assim, é pouco provável que o compartilhamento de rearranjos tão complexos seja fruto de convergência evolutiva. Comparações de arranjos gênicos mitocondriais têm fornecido convincentes reconstruções filogenéticas em casos em que todos os outros dados foram ambíguos, incluindo as relações entre grupos de Echinodermata e Arthropoda (Boore, 1999).

### **2.3. Filogenia Molecular**

A análise filogenética de sequências de DNA ou de proteínas é uma importante ferramenta no estudo da história evolutiva dos organismos, e estudos de filogenia molecular podem ser realizados para elucidar o processo evolutivo em todos os níveis das classificações de organismos. As análises filogenéticas são importantes também para elucidar os padrões evolutivos de uma família gênica e para entender os processos da evolução adaptativa em um nível molecular (Nei e Kumar, 2000).

A reconstrução de árvores filogenéticas mediante o uso de métodos estatísticos teve início, de forma independente, na taxonomia numérica com dados morfológicos e na genética de populações com dados de frequência gênica. Alguns métodos desenvolvidos para esses propósitos ainda são usados para análises filogenéticas com dados moleculares, porém, mais recentemente, muitos outros métodos foram desenvolvidos (Nei e Kumar, 2000). Atualmente, os métodos mais utilizados na reconstrução filogenética com

dados moleculares são: 1) os métodos baseados em distância; 2) os métodos baseados na parcimônia; 3) os métodos baseados na verossimilhança (Schmidt e Haeseler, 2009).

Nos métodos baseados em distância, as distâncias evolutivas são computadas para todos os pares de táxons, e uma árvore filogenética é construída comparando-se os valores das distâncias entre esses pares. Existem muitos métodos baseados em distância, como o método de agrupamento pareado não ponderado com média aritmética (*Unweighted Pair Group Method with Arithmetic Mean* ou UPGMA), que presume que todas as linhagens evoluíram com uma taxa de substituição gênica constante, e a Evolução Mínima, que computa todas as topologias possíveis, e a topologia com o menor valor de comprimento de ramo é escolhida como a melhor árvore (Nei e Kumar, 2000).

O método de máxima parcimônia foi desenvolvido inicialmente para análise de dados morfológicos, e mais tarde foi aplicado em dados moleculares. Nesse método, os nucleotídeos (ou aminoácidos) do táxon ancestral são avaliados separadamente para cada sítio para uma dada topologia. O menor número de substituições capaz de explicar a topologia de cada uma das árvores possíveis é computado, e a topologia que requer o menor número de substituições é escolhida como a melhor árvore (Nei e Kumar, 2000).

O conceito de verossimilhança se refere a situações que normalmente surgem nas ciências naturais na qual a partir de algum dado, uma decisão pode ser tomada para construir uma explicação adequada para os dados. Assim, um modelo específico e uma hipótese são formulados geralmente o modelo não estando em questão. Em análises filogenéticas, uma parte do modelo é que sequências evoluem em um padrão de árvore e isso não muda entre as análises. As possíveis hipóteses incluem todas as estruturas de árvores diferentes, além dos comprimentos de ramo, dos parâmetros do modelo evolutivo, entre outros. Para valores atribuídos a esses elementos, é possível calcular a probabilidade dos dados sob estes parâmetros e fazer declarações sobre a sua

plausibilidade. Ao testar diferentes hipóteses, algumas delas produzirão os dados com maior probabilidade do que outros (Schmidt e Haeseler, 2009).

Na evolução, mutações pontuais são consideradas eventos probabilísticos. Portanto, pelo menos em princípio, a probabilidade de encontrar uma mutação ao longo de um ramo de uma árvore filogenética pode ser calculada usando o conceito de máxima verossimilhança. A principal ideia por trás da inferência filogenética com a máxima verossimilhança é determinar a topologia da árvore, comprimentos de ramo e os parâmetros do modelo evolutivo (por exemplo, razão transição/transversão, frequências de bases e variação na taxa de substituição entre os sítios) que maximizam a probabilidade de observar as sequências em questão (Schmidt e Haeseler, 2009).

Na máxima verossimilhança, a verossimilhança de um conjunto de sequências para um determinado modelo de substituição específica é maximizado para cada topologia, e a topologia que tem o maior valor de verossimilhança é escolhida como árvore final. Os parâmetros a serem considerados não são as topologias, mas os comprimentos de ramos da cada topologia, e a verossimilhança é otimizada para estimar os comprimentos dos ramos (Nei e Kumar, 2000).

A inferência bayesiana é apenas a formalização matemática de um processo de decisão que nada mais é do que uma análise de probabilidade. Nesse sentido, a inferência bayesiana é mais simples do que métodos estatísticos clássicos, que dependem da teoria da amostragem, comportamento assintótico, significância estatística, e outros conceitos mais complexos (Ronquist *et al.*, 2009). Na análise filogenética, a inferência bayesiana parte de uma teoria sobre as relações entre as espécies, que deve ser especificada anteriormente. Na ausência de uma teoria *a priori*, pode-se atribuir probabilidades iguais a todas as árvores possíveis (Ronquist *et al.*, 2009).

Para determinar qual árvore tem uma maior probabilidade *a posteriori* são utilizados os dados, geralmente na forma de um alinhamento de sequências moleculares, e um modelo estocástico no processo de geração dos dados de cada árvore. Em princípio, a regra de Bayes é então utilizada para obter a distribuição da probabilidade *a posteriori*, que é o resultado da análise. A probabilidade *a posteriori* especifica a probabilidade de cada uma das árvores dado o modelo, a teoria *a priori*, e os dados. Quando os dados são filogeneticamente informativos, a maior probabilidade *a posteriori* é tipicamente concentrada em uma única árvore (ou um pequeno subconjunto de árvores) (Ronquist *et al.*, 2009).

Para calcular a probabilidade *a posteriori* podemos utilizar uma técnica de simulação chamada Cadeia de Markov de Monte Carlo (MCMC). As Cadeias de Markov têm a propriedade de convergirem para um estado de equilíbrio, independentemente do ponto de partida. Se a Cadeia de Markov é iniciada a partir de uma árvore aleatória e com comprimentos de ramos arbitrários, as chances são de que a probabilidade inicial seja baixa. À medida que a cadeia se move para outras árvores, a probabilidade normalmente aumenta muito rapidamente. Esta fase inicial da corrida é conhecida como *burn-in*, e geralmente é descartada por ser fortemente influenciada pela o ponto de partida (Ronquist *et al.*, 2009).

### **2.3.1. Métodos para Seleção de Regiões Filogeneticamente Informativas**

A análise filogenética de sequências biológicas requer como primeira etapa o alinhamento de sequências nucleotídicas ou de aminoácidos, de tal forma que os caracteres sejam idênticos em todas as posições. Dessa forma, o alinhamento de sequências biológicas pretende determinar as substituições que ocorreram durante o



processo evolutivo a partir de um ancestral comum às sequências. Assim, a qualidade de um alinhamento múltiplo de sequências pode ter um forte impacto sobre a precisão da árvore filogenética inferida, seja qual for o método utilizado na inferência. Apesar da melhoria constante dos algoritmos usados para o alinhamento de sequências, um alinhamento pode conter regiões onde a homologia dos caracteres é ambígua. Além disso, regiões evolutivamente muito distantes (mesmo quando corretamente alinhadas) podem apresentar um alto grau de saturação mutacional, que vem a ser uma importante fonte de viés nas reconstruções filogenéticas (Criscuolo e Gribaldo, 2010).

A saturação de substituições diminui a informação filogenética contida nas sequências, e tem prejudicado as análises filogenéticas que envolvem grupos antigos, tais como os Arthropoda. Em casos extremos, quando as sequências apresentam um alto grau de saturação, a similaridade entre as sequências dependerá inteiramente da semelhança nas frequências dos nucleotídeos, que muitas vezes não reflete as verdadeiras relações filogenéticas (Xia, 2009).

A fim de minimizar o viés introduzido pelas saturações e por sítios não homólogos, uma abordagem frequentemente utilizada é a detecção e remoção dessas regiões dos alinhamentos múltiplos, antes da análise filogenética. De fato, tem-se observado que a remoção dessas regiões permite a inferência de árvores mais precisas. Nesse sentido, diferentes métodos de corte foram desenvolvidos para a retirada de regiões problemáticas e seleção dos sítios que são adequados para inferência filogenética em um alinhamento múltiplo de sequências. Tais métodos também possibilitam a diminuição do custo computacional para a realização das inferências filogenéticas (Criscuolo e Gribaldo, 2010).

O método desenvolvido por Castresana (2000) e aplicado no *software* GBLOCKS, define um conjunto de blocos conservados de um alinhamento múltiplo de acordo com uma combinação de requisitos simples. Esses blocos são selecionados a partir de

parâmetros relacionados com o número de posições contíguas conservadas, com a quantidade de *gaps*, e com a variabilidade dos sítios que flanqueiam as posições conservadas.

O método desenvolvido por Shannon (1948) é baseado no conceito de entropia da informação. Que para um número  $N$  de sequências alinhadas de comprimento  $L$ , a entropia no sítio  $i$  é dada por:

$$H_i = - \left( \sum_{j=1}^4 p_j \log_2 p_j \right)$$

onde  $j = 1, 2, 3, 4$ , correspondendo aos nucleotídeos A, C, G e T, e  $p_j$  é a proporção de nucleotídeo  $j$  no site  $i$ . O valor de  $H_i$  é dois, quando as frequências dos quatro nucleotídeos em uma posição são iguais, e zero quando o sítio é completamente conservado, com a ocorrência de uma única base.

Batista *et al.* (2011) destacam que em conjuntos de dados altamente diversos, é importante analisar regiões conservadas dos genomas, já que elas têm maior probabilidade de estarem associadas a domínios funcional em proteínas, melhorando a busca por sítios homólogos entre os genomas. Assim, a entropia é uma medida numérica que indica quais regiões tem maior probabilidade de serem homólogas.

Criscuolo e Gribaldalo (2010), utilizando sequências de Archaeplastida e Unikonts, aplicaram a análise de entropia para selecionar regiões filogeneticamente informativas em alinhamentos múltiplos de nucleotídeos e aminoácidos. Estes autores concluíram que o uso da análise de entropia pode melhorar a capacidade de inferência dos métodos usados para a reconstrução filogenética, além de aumentar os valores de suporte das árvores. Batista *et al.* (2011), trabalhando com genomas de Pappilomavírus, também

construíram árvores filogenéticas com valores mais altos de confiança quando utilizaram a entropia para a seleção de regiões filogeneticamente informativas.

## **2.4. Genômica Comparativa**

A sequência completa do genoma de um organismo ou organela pode ser considerada como o mapa gênico final, no sentido de que as características hereditárias são codificadas dentro do DNA e que a ordem de todos os nucleotídeos ao longo de cada cromossomo é conhecida na sequência (Hardison, 2003).

A genômica comparativa se baseia em uma hipótese simples: características comuns a dois organismos, muitas vezes, são codificadas dentro de segmentos de DNA que são conservados entre as espécies. Ou seja, as regiões de DNA que codificam as proteínas e RNAs responsáveis por funções que são conservadas desde o ancestral comum de duas espécies provavelmente são preservadas nas sequências genômicas contemporâneas. Dessa forma, diferentes questões podem ser abordadas por meio da comparação de genomas (Hardison, 2003).

Os genomas mitocondriais completos fornecem bons modelos para estudos de evolução molecular e abundantes marcadores moleculares para estudos filogenéticos e populacionais. Dessa forma, os dados oriundos das análises de genomas mitocondriais completos têm um potencial óbvio para os estudos das relações filogenéticas em diferentes níveis taxonômicos (Wei *et al.*, 2010).

Desde o sequenciamento dos primeiros genomas mitocondriais, uma série de trabalhos foram realizados com o intuito de utilizar esta informação para estudos evolutivos e filogenéticos. Com o grupo dos Arthropoda não foi diferente; tentativas de elucidar as relações filogenéticas do filo baseadas em rearranjo dos genes mitocondriais

foram desenvolvidas por Boore *et al.* (1995; 1998), Lavrov *et al.* (2002; 2004) e Cook *et al.* (2005), entre outros. Além disso, Cook *et al.* (2005) e Hassanin (2006) utilizaram as sequências de todos os genes mitocondriais com o mesmo objetivo.

Dessa forma, o genoma mitocondrial se mostra como uma importante fonte de dados para estudos evolutivos e filogenéticos que envolvam o filo Arthropoda, tendo em vista que centenas de genomas mitocondriais deste grupo já foram completamente sequenciados e se encontram disponíveis nos bancos de dados biológicos públicos. Apesar da sua importância, como fonte primária de informação, os genomas mitocondriais de Arthropoda atualmente disponíveis têm sido pouco utilizados na elucidação das relações filogenéticas do grupo.

Portanto, esse trabalho se justifica por apresentar e discutir a utilização de um extenso conjunto de dados, composto de genomas mitocondriais completos, na reconstrução da história evolutiva dos Arthropoda, ao mesmo tempo em que compara métodos para identificação de regiões filogeneticamente informativas desses genomas.

### **3. Objetivos**

#### **3.1. Objetivo Geral**

Elucidação a história evolutiva do filo Arthropoda utilizando a informação contida em genomas mitocondriais completamente sequenciados.

#### **3.2. Objetivos Específicos**

- Estabelecer a ordem gênica mitocondrial padrão para o filo Arthropoda.
- Comparar a eficiência de diferentes métodos de identificação de regiões filogeneticamente informativas e identificar as regiões que possam melhorar as reconstruções filogenéticas.
- Reconstruir as relações filogenéticas dos Arthropoda a partir da informação contida nos genomas mitocondriais.
- Desenvolver um conjunto de iniciadores degenerados capaz de amplificar todo o genoma mitocondrial da ordem díptera e testar parte dos iniciadores definidos no sequenciamento de fragmentos dos genomas mitocondrial de *Chrysomya megacephala*.

## 4. Material e Métodos

### 4.1. Criação do Banco de Dados

Foram recuperadas do banco de dados do Centro Nacional para Informação Biotecnológica (NCBI, do inglês *National Center for Biotechnology Information* - <http://www.ncbi.nlm.nih.gov>) as sequências nucleotídicas dos genomas mitocondriais completos de 360 Arthropoda. As sequências foram recuperadas no formato *GenBank* que, além da sequência propriamente dita, também disponibiliza outras informações importantes, como regiões codificadas e ordem dos genes.

#### 4.1.1. Construção do Banco de Dados Local

Um *script* na linguagem de programação Perl foi escrito para realizar a construção e alimentação de um banco de dados local de forma automática. Esta linguagem, desenvolvida por Wall (1986), tem sido extensivamente utilizada em áreas como a bioinformática e programação *web*. Os algoritmos em Perl se tornaram populares entre os biólogos por exibirem uma grande eficiência na análise de dados em bioinformática, principalmente dados em formato de texto (Tisdall, 2001).

A partir dos genomas no formato *GenBank*, o *script* foi utilizado para extrair as informações pertinentes para a construção do banco de dados local. As informações, de cada um dos genomas, utilizadas para a construção do banco de dados foram: sequência nucleotídica do genoma inteiro; sequência nucleotídica de cada um dos genes; informações taxonômicas do organismo; ordem gênica; conteúdo de guanina e citosina (GC%); e comprimento do genoma e dos genes mitocondriais.

O banco de dados local foi baseado no modelo de banco de dados relacional. O modelo de banco de dados relacional é uma representação lógica dos dados que permite que os relacionamentos entre os dados sejam considerados sem se preocupar com a estrutura física do banco de dados (Deitel *et al.*, 2001). Para o gerenciamento do banco de dados foi utilizado o sistema de gerenciamento de banco de dados MySQL.

## 4.2. Seleção de Genomas

Inicialmente cada um dos genes codificantes de proteínas teve seu viés de GC% calculado. O viés de GC% é definido como o excesso de G sobre C normalizado para uma dada sequência, podendo ser calculado como:

—

O valor de viés de GC% pode variar de 1 a -1. Um valor igual à 0 indica que a sequência tem a mesmas proporções de G e C, enquanto que um valor próximo de 1 indica que a sequência é enriquecida em G. Variações no viés GC% têm sido relatadas em vários genomas mitocondriais de metazoários (Saccone *et al.*, 1999) e representam uma fonte clara de perturbação do sinal filogenético nos Arthropoda (Jones *et al.*, 2007; Hassanin *et al.*, 2005; Hassanin, 2006). Assim, este índice é uma importante fonte de informação para a escolha dos genomas mitocondriais a serem utilizados nas reconstruções filogenéticas dos Arthropoda.

Os genomas que apresentaram algum gene com desvio do viés GC% foram retirados das análises seguintes. Além disso, para evitar que as análises filogenéticas fossem afetadas por um viés na representação dos táxons, alguns genomas também

foram excluídos das análises, mantendo-se os genomas com viés GC% mais próximos da média dos Arthropoda.

### **4.3. Análise de Sintenia**

A partir das informações de ordem gênica presente no banco de dados local uma matriz binária foi construída. Tal matriz foi construída levando-se em conta a posição relativa de cada um dos genes codificantes de proteínas, genes ribossomais e genes de RNAs transportadores dentro do genoma mitocondrial. Para cada uma das ordens das quatro classes dos Arthropoda, foi considerada apenas uma sintenia. Em grupos com mais de uma sintenia, levou-se em consideração apenas àquela mais representativa dentro do grupo.

Em seguida, a matriz foi utilizada para a construção de um dendograma, utilizando-se o método UPGMA (*Unweighted Pair Group Method with Arithmetic Mean*), com o auxílio do programa MEGA5 (*Molecular Evolutionary Genetics Analysis v. 5.0*) (Tamura *et al.*, 2011) e 1000 de *bootstrap*. A análise de sintenia teve como objetivo determinar quais genomas mitocondriais apresentavam ordens gênicas similares e quais ordens gênicas são mais ou menos derivadas ao longo da evolução do genoma mitocondrial dos Arthropoda.

### **4.4. Alinhamento de Sequências**

As sequências de nucleotídeos de cada um dos genes codificantes de proteínas foram inicialmente traduzidas para sequências de aminoácidos com o programa MEGA5



(Tamura *et al.*, 2011), utilizando-se o código genético mitocondrial de invertebrados. Em seguida, as sequências protéicas foram alinhadas com o programa MUSCLE (Edgar, 2004) incorporado no MEGA5, regiões com alinhamento de baixa qualidade foram ajustadas manualmente. Por fim, os alinhamentos de aminoácidos foram traduzidos de volta para os códons para obtenção das sequências de nucleotídeos alinhadas. Esta abordagem foi utilizada para evitar que *gaps* fossem inseridos em regiões inadequadas, mudando a matriz de leitura dos códons.

Além do alinhamento gene a gene dos Arthropoda, também foi realizado o alinhamento de todo o genoma mitocondrial de 34 espécie de Diptera (*Aedes aegypti*; *A. albopictus*; *Anopheles darlingi*; *A. gambiae*; *A. quadrimaculatu*; *Bactrocera carambolae*; *B. dorsalis*; *B. minax*; *B. olea*; *B. papayae*; *B. philippinensis*; *B. tryoni*; *Ceratitidis capitata*; *Chrysomya putoria*; *Cochliomyia hominivorax*; *Culex quinquefasciatus*; *C. pipiens*; *Cydistomyia duplonotata*; *Dermatobia hominis*; *Drosophila littoralis*; *D. mauritiana*; *D. sechellia*; *D. simulans*; *D. yakuba*; *Exorista sorbillans*; *Haematobia irritans*; *Hypoderma lineatum*; *Liriomyza sativae*; *L. trifolii*; *Lucilia sericata*; *Mayetiola destructor*; *Rhopalomyia pomum*; *Simosyrphus grandicornis*; *Trichophthalma punctata*) com o intuito de prover um conjunto de dados viável para definição de iniciadores degenerados para a ordem. O alinhamento dos genomas dos dípteros também foi realizado com a versão do MUSCLE incorporada ao MEGA5.

#### **4.5. Análises de Variabilidade**

Para cada gene codificante e sua respectiva sequência protéica, foi calculado o número total de sítios alinhados, sítios variáveis e invariáveis, sítios variáveis *singletons*,

sítios informativos para parcimônia, número de aminoácidos conservados e variáveis, através dos programas MEGA5.

Para verificar o possível decaimento do sinal filogenético causado pela saturação mutacional, para cada um dos genes codificantes, o programa DAMBE (Xia e Xie, 2001) foi utilizado para a construção de gráficos que representam o comportamento das taxas de transição e transversão à medida que se aumenta a distância genética entre as sequências.

#### **4.6. Identificação de Regiões Problemáticas**

Com o intuito de excluir regiões dos alinhamentos que pudessem vir a afetar a precisão das inferências filogenéticas, as sequências nucleotídicas foram avaliadas por dois métodos de identificação e exclusão dessas regiões de forma independente.

O primeiro método utilizado foi a entropia de Shannon. Para cada um dos genes, gráficos com os valores de entropia para cada sítio foram construídos com o auxílio do programa DAMBE (Xia e Xie, 2001), onde foi utilizada uma janela de 100 nucleotídeos com o objetivo de reduzir o ruído nos gráficos. Em seguida, todos os sítios com baixa complexidade, ou seja, com valores de entropia menor ou igual a um, foram selecionados para a construção da filogenia.

Para aplicação do segundo método foi utilizado o programa GBLOCKS v. 0.91 (Castresana, 2000). O método utilizado nessa ferramenta consiste em identificar blocos conservados em um alinhamento múltiplo, de acordo com uma combinação de alguns requisitos simples. Os parâmetros utilizados para identificar os blocos conservados nos genes mitocondriais foram: número mínimo de sequências com uma posição conservada igual a 50%; número mínimo de sequências em uma posição flaqueadora igual a 85%;

número máximo de posições não conservadas contínuas igual a oito; tamanho mínimo de um bloco igual a nove.

Tanto nas análises de entropia como na identificação dos blocos conservados as sequências nucleotídicas foram tratadas como códons, de forma a manter a correspondência com as sequências proteicas. Nos casos em que o códon ficou incompleto, após a retirada das regiões pouco conservadas, ele foi desconsiderado. Em seguida, as sequências proteicas foram determinadas a partir dos códons.

Dessa forma, após a seleção das regiões menos divergentes de cada gene mitocondrial codificante dos Arthropoda, as sequências foram concatenadas, formando assim, quatro grandes conjuntos de dados: dois resultantes da concatenação das regiões de baixa entropia em nucleotídeos e aminoácidos; e outros dois resultantes da concatenação dos blocos conservados também em nucleotídeos e aminoácidos.

O alinhamento dos genomas dos dípteros também foi avaliado segundo o método de entropia, sendo que nenhuma região foi excluída desse alinhamento já que o objetivo desta análise era identificar regiões candidatas para a determinação de iniciadores degenerados para a amplificação do genoma mitocondrial dos dípteros.

#### **4.7. Definição de Iniciadores Para Amplificação do Genoma Mitocondrial de Diptera**

O gráfico resultante da análise de entropia dos genomas de Diptera foi utilizado para determinar as regiões mais conservadas e, por isso, possíveis regiões para a definição de iniciadores degenerados para a amplificação dos genomas por PCR.

Possíveis iniciadores foram manualmente definidos a partir do alinhamento das regiões com menor valor de entropia. Cada par de iniciadores foi avaliado, com o auxílio

(<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/>), de forma a identificar a formação de possíveis grampos e de outras características indesejáveis em um par de iniciadores funcional. Os iniciadores foram definidos de forma a não produzir *amplicons* maiores que 1200 pares de bases (pb) e apresentando uma sobreposição de pelo menos 100 pares de bases com dois outros *amplicons*.

#### 4.8. Experimentos de Biologia Molecular

Para a confirmação da funcionalidade dos iniciadores definidos, dois pares foram escolhidos aleatoriamente com o intuito de amplificar e sequenciar parte do genoma mitocondrial de *Chrysomya megacephala*.

A extração do DNA foi realizada a partir do tórax de um exemplar previamente estocado e preservado em etanol 96%, utilizando-se um protocolo baseado em Chellex 5% e proteinase K (Sepp *et al.*, 1994). O DNA extraído foi quantificado por espectrofotometria de massa e estocado em freezer a -20°C. As amostras de DNA foram amplificadas a partir de parte do conjunto de iniciadores definidos, os segmentos amplificados corresponderam a parte dos genes mitocondriais COX1, NAD2 e RNA transportador de metionina). Os *amplicons* foram purificados e sequenciados para o posterior tratamento das sequências. O sequenciamento das amostras foi realizado na Plataforma Tecnológica de Genômica e Expressão Gênica do laboratório central do centro de ciências biológicas da Universidade Federal de Pernambuco.

#### 4.8.1. Tratamento e Análise das Sequências Amplificadas

Os cromatogramas resultantes do sequenciamento dos *amplicons* foram tratados com o pacote de programas Staden (Staden *et al.*, 2003). Nesse tratamento, as extremidades de baixa qualidade foram retiradas e em seguida as sequências foram clusterizadas. Para a retirada das regiões de baixa qualidade foi utilizado um valor de PHRED (Ewing *et al.*, 1998) igual a 40 como ponto de corte.

Após o tratamento das sequências, foi utilizado o programa BLASTn (*Basic Local Alignment Search Tool*) (Altschul *et al.*, 1990) para determinar se os iniciadores realmente amplificaram as regiões previstas. Em seguida, foi realizado um alinhamento múltiplo entre as sequências de *Chrysomya megacephala* e as regiões correspondentes dos genomas de outros dípteros. Esse alinhamento foi realizado com o programa MUSCLE incorporado ao MEGA5. Por fim, os dois alinhamentos foram concatenados.

#### 4.9. Análises Filogenéticas

Para as reconstruções filogenéticas dos Arthropoda baseadas nas sequências dos genes mitocondriais, três grupos externos foram utilizados de forma independente: Tardigrada, Onychophora e Priapulida. Tanto para as regiões de baixa entropia quanto para os blocos conservados foram utilizadas as sequências nucleotídicas e de aminoácidos. Em todas as análises os *gaps* foram tratados como dados perdidos.

O programa ModelTest versão 3.7 (Posada e Crandall, 1998) foi utilizado para definir a escolha do modelo de substituição de nucleotídeos que melhor se adéqua-se aos dados. Para as sequências de aminoácidos, o programa ProtTest versão 2.4 (Abascall *et*

al., 2005) teve essa função. Em ambos os casos foi utilizada a estrutura de seleção de modelos *Akaike Information Criterion* (AIC).

A análise de máxima verossimilhança foi realizada com a ajuda do programa PhyML versão 3.0 (Guindon e Gascuel, 2003). Os modelos evolutivos e os parâmetros que mais se adequaram ao conjunto de dados foram utilizados, assim como o método NNI (*Nearest Neighbor Interchange*) para estimar a topologia da árvore. Uma árvore de BioNJ (algoritmo desenvolvido por Gascuel [1997] a partir do *neighbor-joining*) foi utilizada como árvore inicial e os táxons foram adicionados aleatoriamente.

A análise bayesiana foi realizada com o programa MrBayes v. 3.1 (Ronquist e Huelsenbeck, 2003). Mais uma vez, os parâmetros e modelos mais adequados foram utilizados. A topologia final foi estimada com 1.000.000 de interações usando 2.500 ciclos descartados e amostrando a cada 100 interações.

O alinhamento com as sequências de *Chrysomya megacephala* e outros Diptera também foi utilizados para gerar uma filogenia do grupo. Nesse momento, a metodologia empregada foi a mesma dos outros conjuntos de dados, com a diferença que *Bittacus pilicornis*, *Boreus elegans* e *Neopanorpa pulchra*, pertencentes à ordem Mecoptera, formaram o grupo externo.

## **5. Resultados**

### **5.1. Banco de Dados Local**

Após as buscas no NCBI por genomas mitocondriais completos de Arthropoda um total de 360 genomas foi recuperado. Nesse momento, os genomas com duplicação ou ausência de um dos genes codificantes de proteínas não foram considerados. Dos genomas recuperados, 49 foram do subfilo Cheliceratha, 57 do subfilo Crustacea, 247 do subfilo Hexapoda e sete do subfilo Myriapoda. O Anexo 1 apresenta todos os organismos que tiveram seus genomas recuperados e utilizados nas análises, além de apresentar a classificação taxonômica desses organismos.

### **5.2. Análise de Sintenia**

Um total de 113 ordens gênicas diferentes foi identificado nos genomas mitocondriais dos Arthropoda. Deste total, 94 estavam presentes em uma única espécie e representavam, assim, uma autapomorfia ou um efeito da diferença na representação dos táxons, já que alguns grupos estão representados por apenas um genoma mitocondrial.

A Figura 3 apresenta o agrupamento de 44 ordens dos Arthropoda que apresentam ao menos uma espécie com genoma mitocondrial completamente sequenciado. Nela podemos observar que 23 táxons apresentam a mesma ordem gênica, formando um grande clado com valor de suporte igual 66. Dessas ordens, 19 são Hexapoda e quatro são Crustacea. Esse compartilhamento da ordem gênica entre tantos grupos de Crustacea e Hexapoda suporta com grande robustez a hipótese do grupo Pancrustacea

indicando que um grande número de espécies desse grupo ainda apresenta a ordem gênica ancestral.

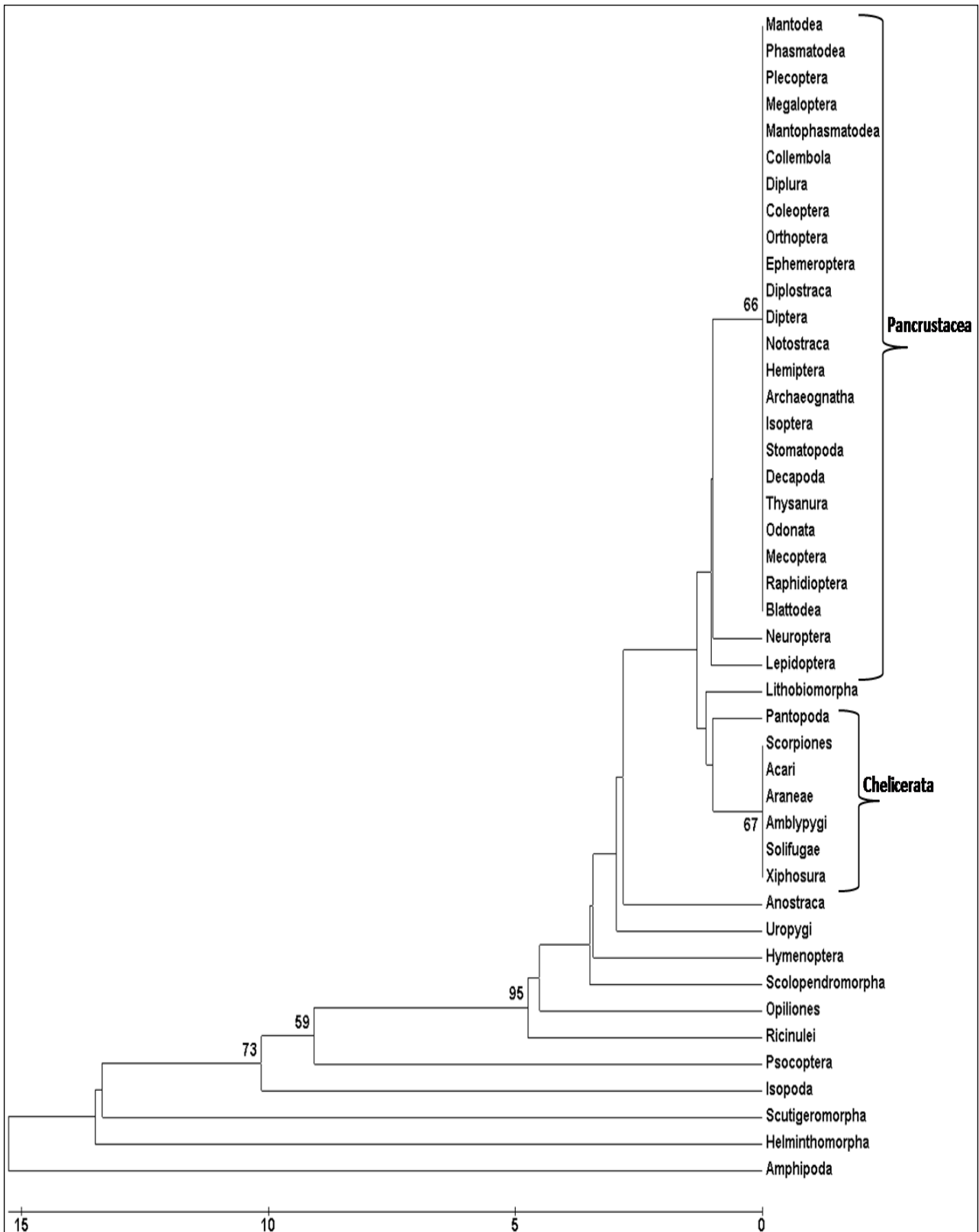
Os demais grupos de Hexapoda (Neuroptera, Lepidoptera, Hymenoptera e Pscoptera) e Crustacea (Anostraca, Isopoda e Amphipoda) se distribuíram ao longo da árvore por apresentarem ordens gênicas derivadas e distintas da sintenia ancestral de Pancrustacea, indicando a ocorrência de mutações que formaram as sintenias específicas desses grupos. Os grupos Lepidoptera e Neuroptera, por exemplo, apresentaram sintenias bem próximas dos outros Pancrustacea. Em Neuroptera houve uma troca de lugares entre dois RNA transportadores vizinhos (Triptofano e Cisteína) que ficam entre os genes ND2 e COX1. Já em Lepidoptera, a mudança foi a alteração relativa do RNA transportados de metionona. Corroborando assim a hipótese dos Pancrustacea e a definição de uma ordem gênica ancestral para o grupo.

Nas análises de ordem gênica o grupo irmão de Pancrustacea foi um clado formado por oito ordens (Acari, Amblypygi, Aranae, Lithobiomorpha, Pantopoda, Scorpiones, Solifugae e Xiphosura). Dessas, apenas Lithobiomorpha (Myriapoda) não pertence ao subfilo Chelicerata. A ordem gênica mais abundante nos Chelicerata apresenta apenas uma diferença em relação à sintenia do grupo Pancrustacea: a alteração da posição relativa do RNA transportador de leucina, que em Pancrustacea está entre os genes COX1 e COX2 e em Chelicerata se localiza imediatamente antes do gene da maior subunidade ribossomal. Assim a análise de sintenia suporta a formação de um clado composto pelos grupos Pancrustacea (Crustacea + Hexapodas) e Chelicerata.

A posição filogenética do subfilo Myriapoda, baseada na análise de ordem gênica é difícil de determinar. Pois, esse grupo apresenta um pequeno número de genomas completamente sequenciados e ainda não tem uma sintenia conservada representada nos bancos de dados públicos. Dessa forma, a análise de ordem gênica não gerou



nenhum clado robusto para a determinação da posição filogenética do grupo ou determinação de uma sintenia ancestral.



**Figura 15.** Árvore de UPGMA baseada na ordem gênica do genoma mitocondrial dos Arthropoda. Valores de *bootstrap* menores que 50% não são mostrados.

### 5.3. Análise do Viés GC%

Todos os genes mitocondriais dos Arthropoda apresentaram grande variação no viés GC%. O gene que apresentou maior variação foi o ATP8, que variou de -0,889 a 0,692. O segundo gene mais variável, ND6, variou de -0,820 a 0,683. Os genes que menos variaram em relação ao viés GC% o CYTB que variou de -0,677 a 0,326; o COX2 que variou de -0,573 a 0,392 e o gene COX1 que variou de -0,446 a 0,365 (Tabela 1).

**Tabela 1.** Variação no viés GC% dos genes mitocondriais.

Gene	Menor Viés GC	Maior Viés GC
ATP8	-0,889	0,692
ND6	-0,82	0,682
ND3	-0,726	0,705
ND2	-0,743	0,614
ND4L	-0,434	0,855
ATP6	-0,828	0,488
ND4	-0,478	0,748
ND5	-0,488	0,703
COX3	-0,624	0,488
ND1	-0,302	0,747
CYTB	-0,677	0,326
COX2	-0,573	0,392
COX1	-0,446	0,367

O Anexo 1 apresenta os valores de viés GC% calculados para todos os genes de cada um dos 360 genomas recuperados do NCBI. Além de apresentar quais genes se distanciaram do padrão para genes mitocondriais de Arthropoda, marcados em cinza escuro. As espécies marcadas com essa mesma cor foram espécies retiradas por apresentar um ou mais dos genes com desvio do viés GC%.

Ao todo 78 genomas apresentaram ao menos um dos genes, além do ATP8, com desvio do viés GC%. Desses 28 foram do subfiló Chelicerata, 12 foram Crustacea, 36

Hexapoda e dois Myriapoda. Apesar do maior número de Hexapoda com algum gene com desvio, o subfilo que teve a maior proporção de genomas com genes enviesados, 57,14%, foi o Chelicerata por apresentar 49 genomas, enquanto os hexapodas apresentaram 14,5% das espécies com genes fora do padrão.

Entre todos os genes, o ATP8 foi o que apresentou o maior número de espécies com grande desvio em seu viés GC%, 66 genomas. No entanto as informações de viés GC% do gene ATP8 não foram utilizadas para selecionar os genomas para as análises filogenéticas, pois tal gene apresenta um pequeno tamanho, variando de 97 a 228 nucleotídeos, e uma grande variabilidade. Características essas, que causaram a retirada de grande parte ou de todo o gene, a depender do método, após a retirada de regiões problemáticas para filogenia.

Com exceção dos genes ND1, ND4 e ND5, os genes com desvio do viés GC% apresentaram valores positivos, e próximos de um, com maior frequência que números menores que zero. Indicando que esses genes em geral apresentam um enriquecimento de guanina e uma diminuição de citosina em suas sequências.

O Anexo 1 também apresenta os genomas (marcados em cinza claro) que foram retirados das análises para evitar um viés causado pela representação excessiva de um determinado táxon. Desses, quatro pertencem ao subfilo Chelicerata, treze são Crustacea, 125 Hexapoda e nenhum Myriapoda. O maior número de Hexapoda retirados nessa fase se deve ao grande número de espécies desse grupo com genoma mitocondrial completamente sequenciado.

#### 5.4. Análise de Variabilidade

Os alinhamentos para cada gene mitocondrial codificante de proteína dos Arthropoda foram realizados com um total de 139 genomas; cinco Myriapoda, 17 Chelicerata, 32 Crustacea s e 85 Hexapoda.

O gene que apresentou o maior número de sítios alinhados foi o ND5, com 1895 sítios, e o menor gene foi o ATP8 com apenas 234 sítios. O gene mais variável foi o ND3 com 93,5% de sítios variáveis. Os demais genes também apresentaram altos valores de variação: ND2, 92,2%; ATP8, 91,4%; ND5 88,1%; ND6, 87,8%; ND4, 86,8%; ND1, 86,5%; CYTB, 85,6%; COX3, 84,9%; ATP6, 84,8%; COX2, 84,2%; ND4L, 78,5%; COX1, 72,4%. Com relação às proteínas, o ND3 foi a proteína com maior número de aminoácidos variáveis, com 100% dos sítios variando. Enquanto o COX1 foi o menos variável, com 72% (Tabela 2).

**Tabela 2.** Variabilidade dos genes mitocondriais.

<b>Gene</b>	<b>SAL</b>	<b>SV</b>	<b>SI</b>	<b>SING</b>	<b>SP</b>	<b>aaC</b>	<b>aaV</b>
ATP6	756	641	60	36	630	17	219
ATP8	234	214	17	29	167	3	73
COX1	1643	1190	381	142	1048	128	394
COX2	722	608	96	46	562	29	204
COX3	857	728	108	71	645	35	243
CYTB	1188	1017	159	96	911	42	349
ND1	1085	939	70	71	852	9	326
ND2	1133	1045	43	42	989	43	377
ND3	371	347	15	32	315	0	123
ND4	1565	1359	137	101	1240	38	459
ND4L	377	296	20	13	282	2	102
ND5	1895	1670	171	114	1532	42	570
ND6	647	568	26	23	532	2	215

SAL. Sítios alinhados; SV Sítios variáveis; SI. Sítios invariáveis; SING. *Singletons*; SP. Sítios informativos para parcimônia; aaC. Aminoácidos conservador; aaV aminoácidos Variáveis.

## **5.5. Análise Comparativa Entre Métodos de Seleção de Regiões Filogeneticamente Informativa**

Todos os genes alinhados foram avaliados segundo o método de entropia (Shannon, 1948) e também pelo método de seleção de blocos (Castresana, 2000). Foi possível identificar um total de 5773 sítios em regiões de baixa entropia nos genes mitocondriais e 7224 em blocos conservados. 4915 sítios foram selecionados por ambos os métodos (Tabela 3). Indicando que o método de entropia, menor ou igual a 1, é mais sensível à variação do que o método de busca por blocos conservados.

O gene COX1 foi o gene com maior número de sítios selecionados, pelos dois métodos. Esse resultado é explicado pelo grande tamanho do gene e por ser o gene mitocondrial mais conservado nos Arthropoda. O gene ND4L não apresentou nenhuma região com entropia menor ou igual a 1, e apenas dois blocos conservados, com 63 nucleotídeos no total, foram identificados. O gene ATP8 também teve um pequeno número de sítios em regiões de baixa entropia ou pertencentes a blocos conservados, 6 e 21 respectivamente. Esse resultado também foi esperado já que os dois genes apresentam um pequeno tamanho e grande diversidade.

**Tabela 3.** Comparação entre número de sítios identificados como de baixa entropia e pertencentes a blocos conservados.

Gene	Entropia	Nº de Blocos	GBLOCKS	Ambos*
ATP6	327	9	465	324
ATP8	6	1	21	0
COX1	1320	12	1407	932
COX2	519	4	582	487
COX3	615	4	744	600
CYTB	903	10	894	737
ND1	564	12	636	492
ND2	33	15	438	33
ND3	138	7	180	100
ND4	564	11	810	541
ND4L	0	2	63	0
ND5	753	15	879	649
ND6	30	3	105	20
Total	5772	105	7224	4915

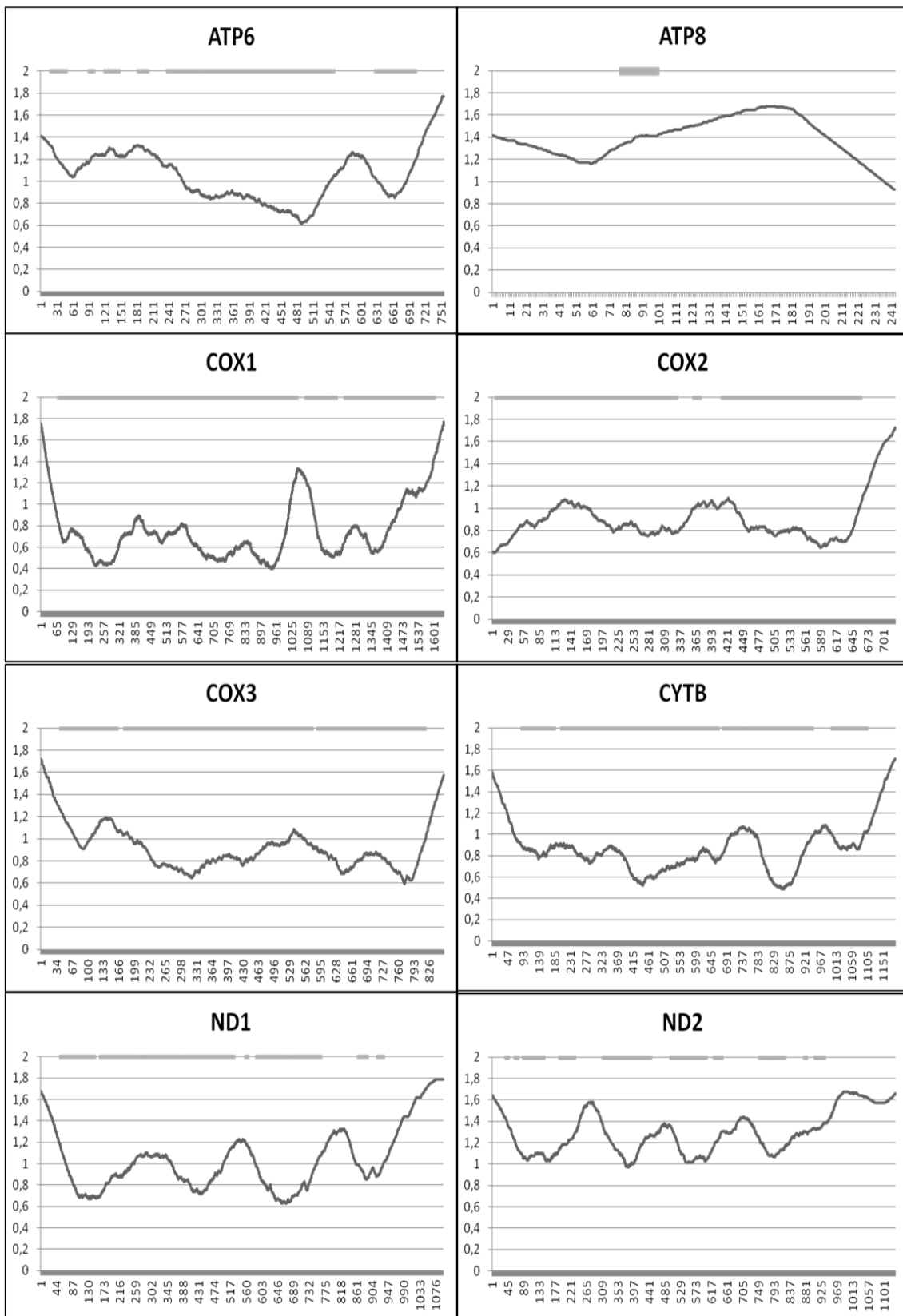
\*Número de sítios selecionados por ambos os métodos

Os genes com maior número de blocos conservados detectados foram ND2 e ND5. O gene ND5 apresentou um número próximo de sítios com baixa entropia e em blocos conservados, indicando que o gene é composto por várias regiões conservadas intercaladas com regiões com grande diversidade. Porém, o gene ND2 apresentou apenas 33 nucleotídeos em uma única região de baixa entropia e 438 sítios distribuídos em 15 blocos sendo, dessa forma, o gene com maior divergência entre as duas análises. Indicando, também, que o método de entropia é mais sensível à variação.

As figuras 4 e 5 apresentam graficamente a comparação entre os dois métodos, gene a gene. O gene COX1 apresenta baixo valor de entropia em quase todo seu comprimento, com um pico de alta entropia entre os sítios 960 e 1100. Região essa, que em parte, pertence a um dos doze blocos conservados identificados. O gene ATP6 apresentou seis pequenos blocos, totalizando 93 sítios, em uma região inicial de alta entropia. O único bloco, sítios 79-99, identificado no gene ATP8 se localiza em uma região de alta entropia, enquanto apenas os seis sítios finais apresentam entropia menor

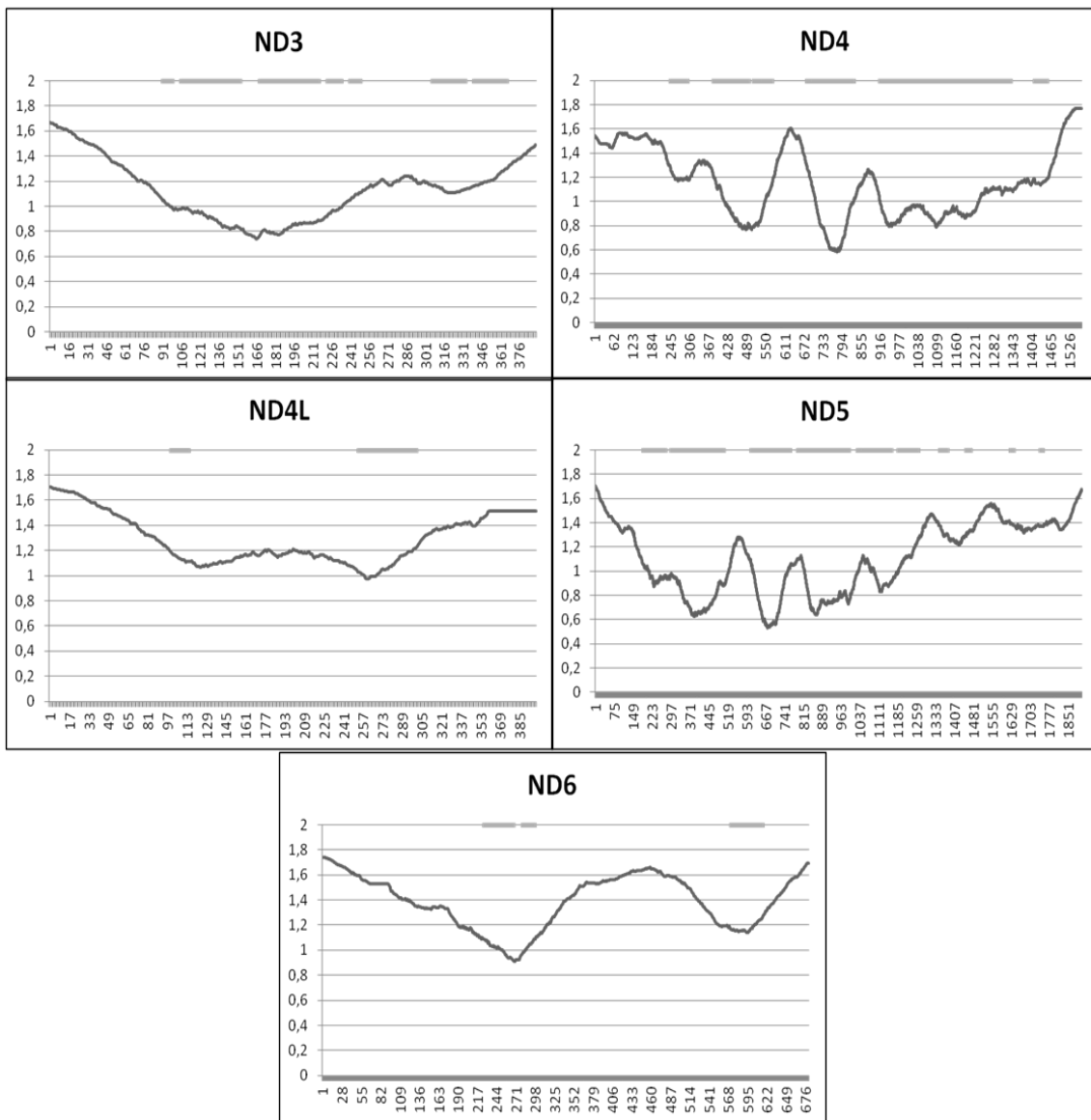
que 1. O gene ND2 apresenta apenas 33 sítios com baixa entropia. No entanto, 438 sítios estão distribuídos em 15 blocos conservados, assim como os genes ND4L e ND6, que apresentam poucos sítios com baixa entropia e alguns blocos conservados. Os demais genes possuem uma boa correlação entre os blocos identificados e as regiões de baixa entropia. Indicando que os métodos são capazes de identificar, em certo grau, as regiões com menor divergência, o que possibilita o galho de precisão nas análises filogenéticas.

Em níveis de divergência mais antigos, sítios variáveis estão sujeitos ao acúmulo de substituições múltiplas (especialmente de mutações do tipo transição nas terceiras posições dos códons), levando, portanto à saturação. A saturação reduz o sinal filogenético e é uma das maiores causas de homoplasia em conjuntos de dados formados por sequências nucleotídicas.



**Figura 16.** Gráficos mostrando os valores de entropia para os genes ATP6, ATP8, COX1, COX2, COX3, CYTB, ND1 e ND2. Eixo horizontal, numero de bases; eixo vertical, valor de entropia. Blocos selecionados pelo GBLOCKS são evidenciados com linhas a cima de cada gráfico (em cinza claro). Alguns blocos podem não ser distinguíveis devido ao grande tamanho do gene e proximidade entre os blocos.



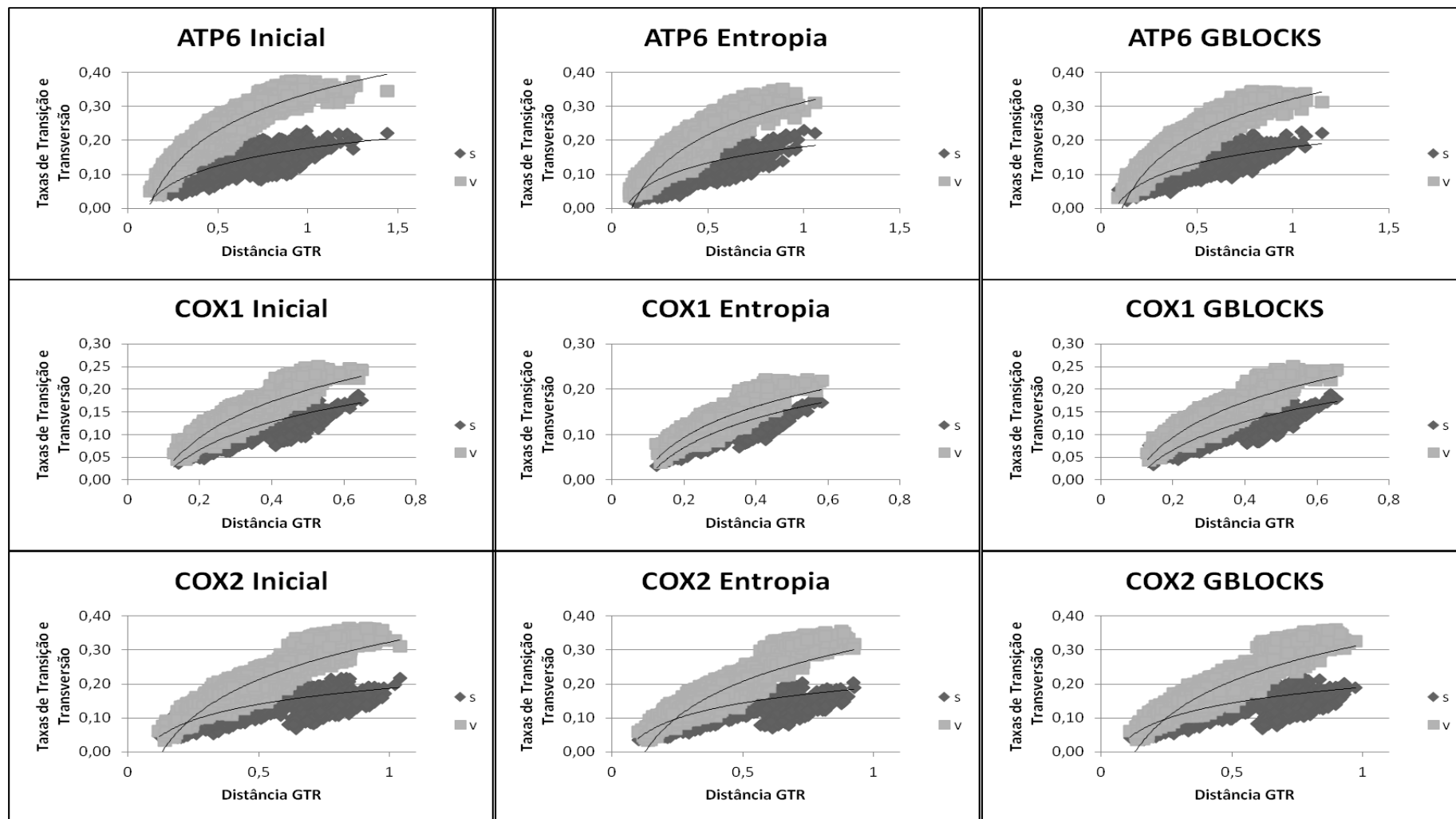


**Figura 17.** Gráficos mostrando os valores de entropia para os genes ND3, ND4, ND4L, ND5 e ND6. Eixo horizontal, numero de bases; eixo vertical, valor de entropia. Blocos seleccionados pelo GBLOCKS são evidenciados com linhas a cima de cada gráfico (em cinza claro). Alguns blocos podem não ser distinguíveis devido ao grande tamanho do gene e proximidade entre os blocos.

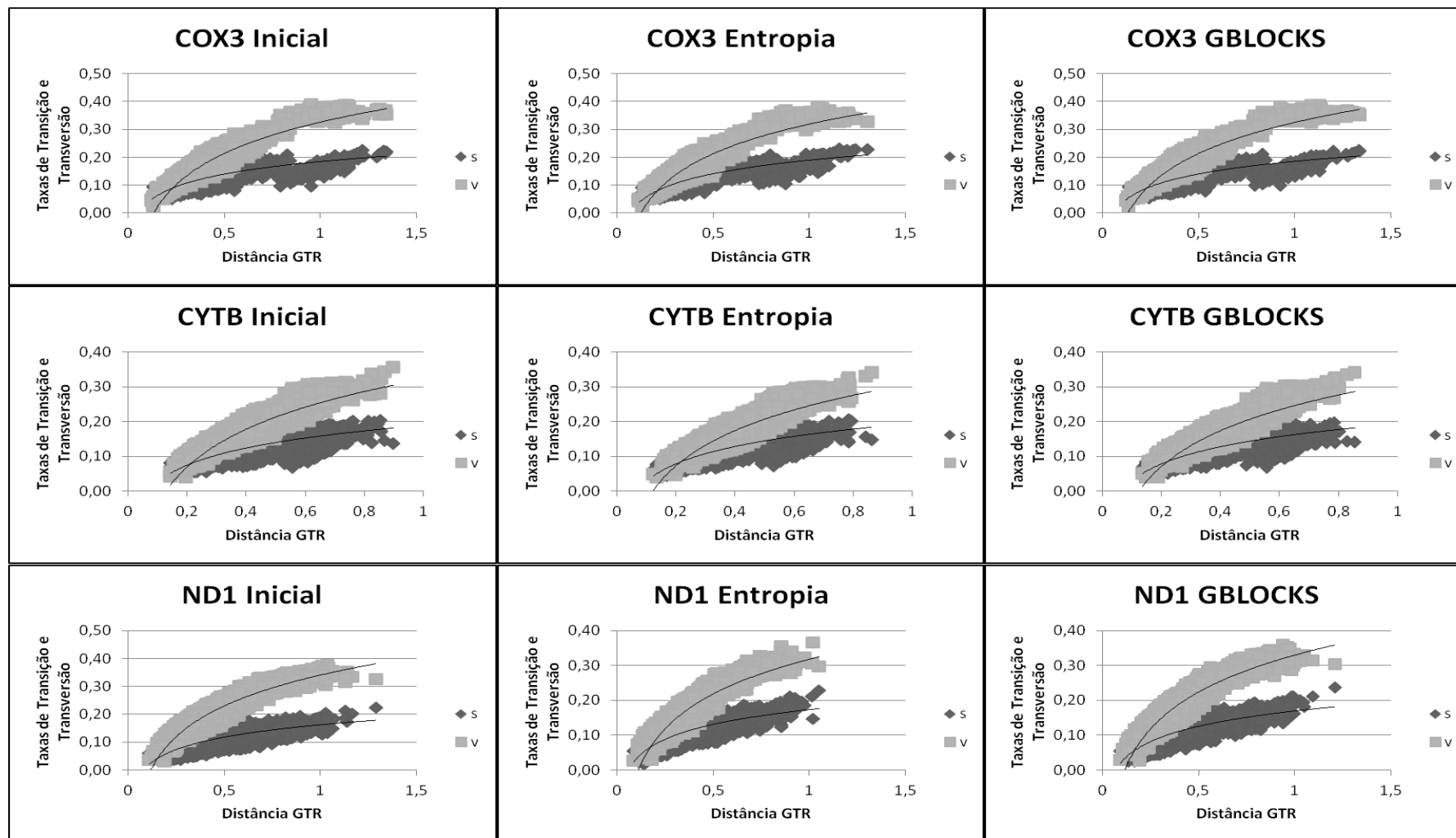
Para determinação da eficiência na redução das saturações, os dois métodos foram testados e comparados graficamente com a condição inicial de cada gene (Figura 6 e 7). Nos gráficos, à medida que se aumenta a distância genética, as taxas de transição e transversão também deveriam aumentar. Quando as sequências estão saturadas, isso não acontece e é possível observar um platô no gráfico. Não foi possível construir os

gráficos de saturação para os genes ATP8, ND2, ND4L e ND6 devido ao pequeno número de sítios selecionados por um ou ambos os métodos.

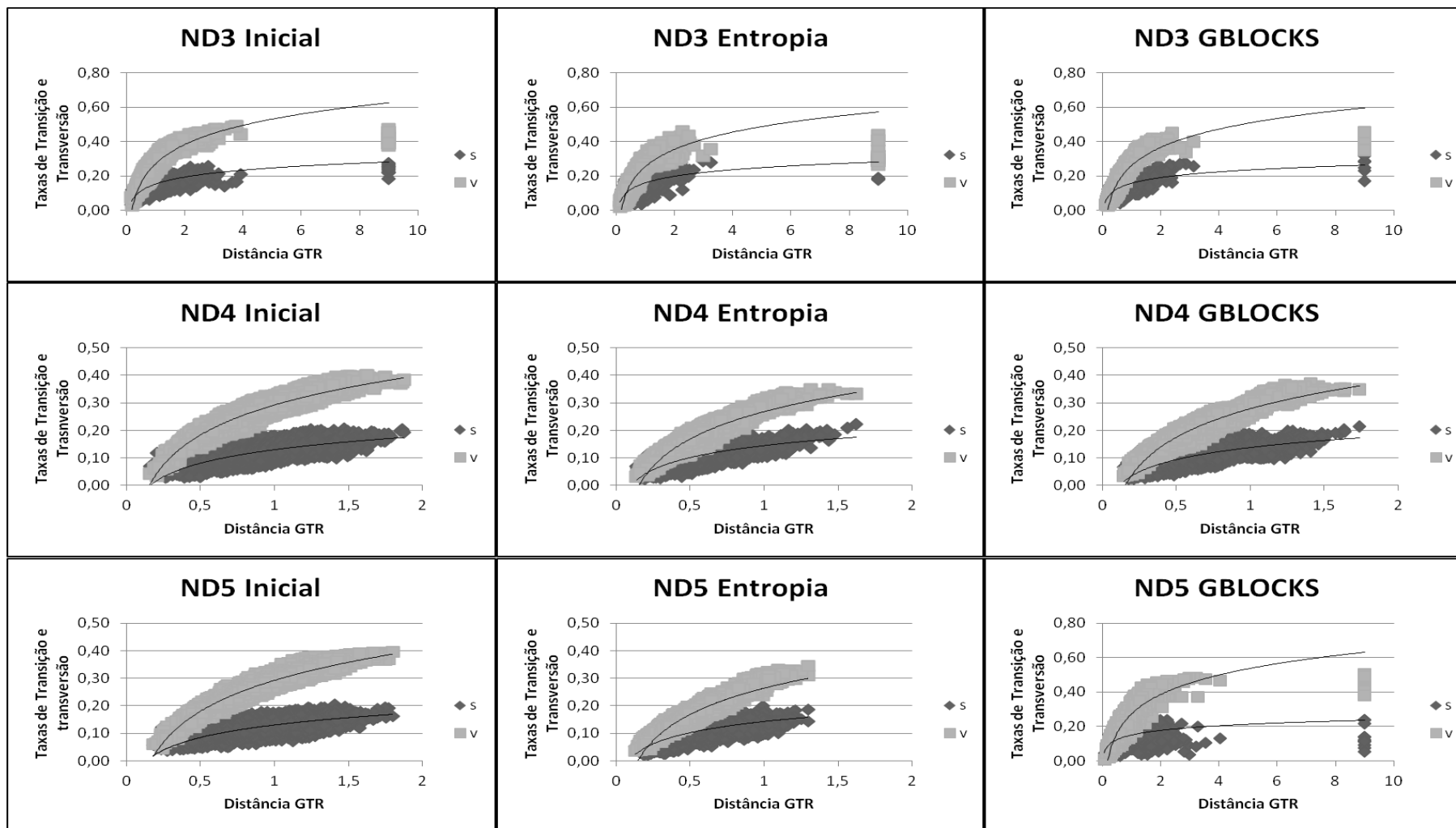
Os testes de saturação mostraram que todos os genes mitocondriais estavam saturados. Os dois métodos apresentaram redução do platô característico das sequências com saturação de mutações. Ambos os métodos também apresentaram diminuição da distância genética entre as espécies e aumento da taxa de transição em relação à taxa de transversão, indicativos da diminuição da saturação de mutações. Tanto o método de entropia quanto a seleção de blocos conservados se mostraram capazes de diminuir o grau de saturação dos genes mitocondriais. No entanto, o método de entropia apresentou uma maior eficiência, principalmente devido a maior redução das distâncias entre os genomas e aumento das taxas de transição quando comparado com o método de seleção de blocos.



**Figura 18.** Gráficos mostrando a saturação das sequências de DNA iniciais e para os dois métodos de seleção empregados nos genes ATP6, COX1 e COX2. S = taxa de transição e V = taxa de transversão.



**Figura 19.** Gráficos mostrando a saturação das sequências de DNA iniciais e para os dois métodos de seleção empregados nos genes COX3, CYTB e ND1. S = taxa de transição e V = taxa de transversão.



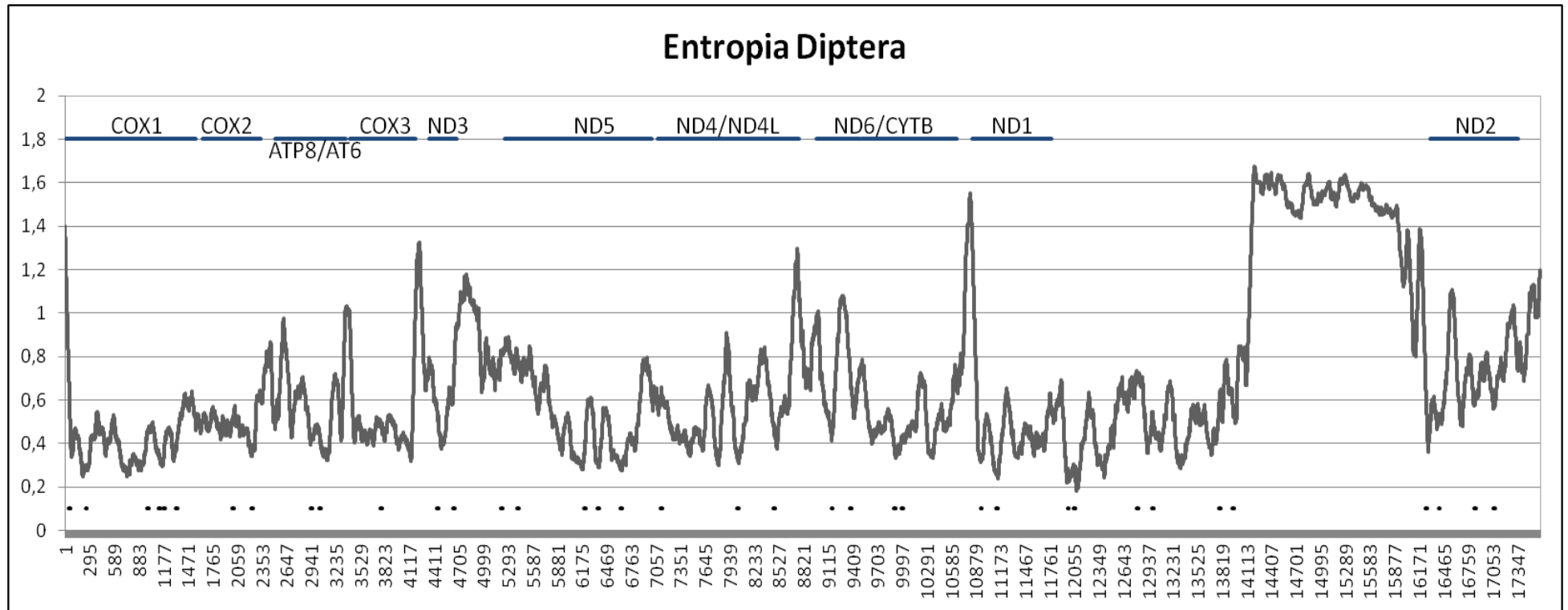
**Figura 20.** Gráficos mostrando a saturação das sequências de DNA iniciais e para os dois métodos de seleção empregados nos genes ND3, ND4 e ND5. S = taxa de transição e V = taxa de transversão.

## 5.6. Definição de Iniciadores Degenerados para Amplificação do Genoma Mitocondrial da Ordem Diptera

Os valores de entropia calculados para o alinhamento formado pelos genomas mitocondriais de trinta e quatro Diptera são apresentados na figura 9. As regiões de baixa entropia foram utilizadas para definir um conjunto de vinte e dois pares de iniciadores degenerados capaz de amplificar todo o genoma mitocondrial através de PCR (Tabela 4).

A maior parte das regiões de baixa entropia nos genomas dos dípteros foi identificada nos genes codificantes de proteínas, principalmente o gene COX1 que, por ser o gene mitocondrial mais conservado, possibilitou a determinação de seis iniciadores em sua extensão. O gene ATP8, por ser curto e muito variável, não mostrou nenhuma região conservada o suficiente para a definição de um *iniciador*.

Pôde-se observar entre o gene ND1 e ND2 uma região de alta entropia. Essa região é a região de controle, caracterizada por uma baixa conservação entre os genomas. Impossibilitando assim, a utilização dessa região para definição de iniciadores eficientes em todos os dípteros.



**Figura 21.** Gráfico mostrando os valores de entropia para os genomas mitocondriais da ordem Diptera. Pontos na parte inferior indicam as regiões, 18-26 pares de bases, utilizadas para a definição de iniciadores degenerados. Eixo horizontal, numero de bases; eixo vertical, valor de entropia.

**Tabela 4.** Iniciadores degenerados definidos para a amplificação dos genomas mitocondriais de Diptera.

	<b>FW (5' - 3')</b>	<b>Tm(°c)</b>	<b>RW (5' - 3')</b>	<b>Tm(°c)</b>	<b>FRAGMENTO*</b>
<b>LABBE-Dip1</b>	TCAACTAAYCATAARGATATTGG	47	CAACWACATAATAWGTATCGTG	47,4	1099
<b>LABBE-Dip 2</b>	AGTTGATTAGCTACWTTWCAYGG	52,3	TCTGAGTAWCGTCGAGGTATTCC	55,3	362
<b>LABBE-Dip 3</b>	GTMTTGGCAATTATAGCAGG	46	CCRCARATTTCTGAACATTG	47	1061
<b>LABBE-Dip 4</b>	GAATTTGAYTCWTATATAGTTCCAAC	51	ACTAAATGAGCAAATATATGTTGWGT	50,1	1035
<b>LABBE-Dip 5</b>	TTYCCTTATATTTTTACAAGAAC	45	GGAATTTGRAAGGGATTAAA	45	805
<b>LABBE-Dip 6</b>	CTATATAYCAATGATGACGAGA	48	TGRTTTCATTCATGRTATAATCC	48	1023
<b>LABBE-Dip 7</b>	CATTTGAATGCGGRTTTGA	45	GATCARGGTTGAASAGAATWT	47	800
<b>LABBE-Dip 8</b>	ACATCTTCARTGTCATRCTCT	50,5	CAATAGCAGCCCCTACTCC	53,2	1155
<b>LABBE-Dip 9</b>	AATTTGCYCCTAATCCTGCYATA	52	CATTGATTWCCTCTCAATTTAYT	54,5	923
<b>LABBE-Dip 10</b>	CCTATATGAGCWACAGAWGAA	48,5	CGAACWCATGGWAATGATTAT	46,5	897
<b>LABBE-Dip 11</b>	CTARTCCTAAACCATCTCAHCC	52,5	TTGCTTATTCWTCWGTAGCTCA	49,2	955
<b>LABBE-Dip 12</b>	ARTAAATTGAGAGGWAATCAATG	47	GWTGAGGGTATCAACCTGA	48,9	922
<b>LABBE-Dip 13</b>	TCAGGTTTATACCCTCAWC	46,8	ATMAATAAAACTAGTATWCCTC	44,5	1041
<b>LABBE-Dip 14</b>	AAATAATCATTACCATGWGT	41,5	TAGGAAAATRTATAGTTT	38	843
<b>LABBE-Dip 15</b>	GAGGWATACTAGTTTTATTKAT	44,5	CCTCAWACYCATTGAACTA	45	846
<b>LABBE-Dip 16</b>	TAGGWTATGYTTACCWTGAGG	50	GTTTCTGGRTTTAATGTWGARTA	48	1139
<b>LABBE-Dip 17</b>	TTATCATAWCGAAATCGAGG	45,6	GACCACGATGTTGGATTAA	46,8	1122
<b>LABBE-Dip 18</b>	TATGAACTCAGATCATGTAA	43,6	TCTCGAWTTTGAAAGAKTTA	42	986
<b>LABBE-Dip 19</b>	AGCTCATCCCTTAACATATT	45,6	CTGCTRCTGCGTTATAC	48	1118
<b>LABBE-Dip 20</b>	AGGGTATCTAATCCTAGTTT	45,6	ATAGATGMTAACAAGTTAATGTC	47	1200
<b>LABBE-Dip 21</b>	AGCTAYTGAGYTCATACCCC	51,5	AATCCTATAAAWGGAGGTARTCC	50,5	818
<b>LABBE-Dip 22</b>	TAGCTTTTTCWTCWATCAAYCA	46,5	ARTCAATTTCCAAAYCCWCC	47,5	937

\*Fragmento: tamanho do fragmento esperado, baseando-se no genoma de *Anopheles gambiae*.



### 5.6.1. Teste dos Iniciadores Definidos

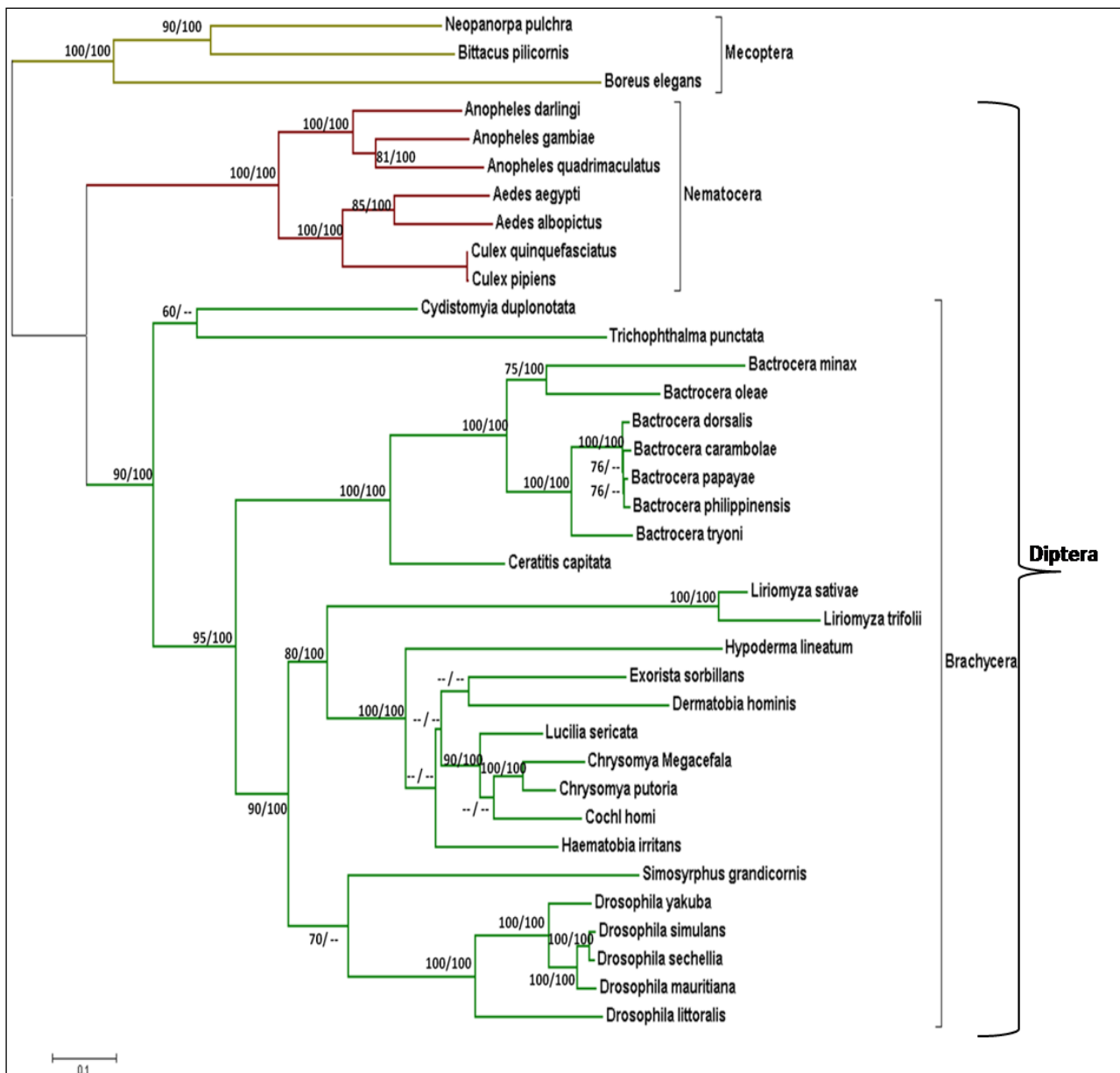
Com o intuito de determinar a eficiência dos iniciadores definidos. Dois pares de iniciadores (LABBE-Dip 2 e 21) foram utilizados para a amplificação de parte do genoma mitocondrial de *Chrysomya megacephala*.

Cada um dos pares de iniciadores amplificou um único fragmento. Após o tratamento das sequências, o par LABBE-Dip 2 amplificou um fragmento de 362 pares de bases correspondentes à parte do gene COX1, confirmado com a ferramenta BLASTn. O par LABBE-Dip 21 amplificou uma região de 821 pares de bases correspondentes às 55 pares de bases finais do RNA transportador de metionina e 766 iniciais do gene ND2, o que corresponde com aproximadamente 75% do gene. Já que, em média, o gene ND2 dos dípteros apresenta 1020 pares de bases. As regiões amplificadas pelo par LABBE-Dip 21 também foram confirmadas com a ferramenta BLASTn. A correspondência do tamanho do fragmento e regiões esperadas, com as amplificadas pelos dois pares de iniciadores, demonstra que o método de entropia é extremamente eficiente na identificação de regiões viáveis para a determinação de iniciadores degenerados capazes de amplificar os genomas de mais de uma espécie.

As sequências determinadas foram alinhadas com as regiões correspondentes dos genomas mitocôndrias das outras espécies pertencentes à ordem Diptera. Como grupo externo, foram utilizados e três espécies da ordem Mecoptera. O modelo evolutivo mais adequado para a reconstrução filogenética com esse conjunto de dados foi o TVM+I+G. A árvore filogenética resultante dessa análise é apresentada na Figura 10.

Como esperado, a espécie irmã de *Chrysomya megacephala* foi *C. putoria* com 100 de *bootstrap* e probabilidade bayesiana. A árvore apresenta dois grandes clados que correspondem com as subordens Nematocera e Brachycera. Todas as espécies que

pertencem a um mesmo gênero formaram clados monofiléticos correspondentes com aos gêneros: *Anopheles*, *Aedes*, *Bactrocera*, *Chrysomya*, *Culex*, *Drosophila* e *Liriomyza*. Em geral, os valores de suporte foram altos e as topologias das árvores de máxima verossimilhança e inferência bayesiana foram congruentes. Demonstrando que os fragmentos utilizados são eficientes na reconstrução filogenética da ordem díptera.



**Figura 22.** Árvore de máxima da ordem díptera baseada nas regiões amplificadas com iniciadores degenerados concatenadas (1183 pb). Topologia determinada por Máxima Verossimilhança, os valores de suporte dos ramos representam: *bootstrap* para Máxima Verossimilhança/Probabilidade Bayesiana. Incongruências nas topologias e valores de *bootstrap* abaixo de 50% não estão representadas.

## 5.7. Reconstrução Filogenética do Filo Arthropoda

O modelos evolutivos que melhor se adequaram a maior parte dos *data-sets* utilizados foram o GTR+I+G, para nucleotídeos, e MtArt+I+G, para aminoácidos, o modelo TVM+I+G se mostrou a melhor opção nas análises de nucleotídeos com baixa entropia nas análises em que os grupos externos foram Priapulida e Tardigrada (Tabela 5).

**Tabela 5.** Modelos evolutivos aplicados na reconstrução filogenética dos Arthropoda.

Grupo externo/Método	Nucleotídeo	Aminoácido
Onychophora /Entropia	GTR+I+G	MtArt+I+G
Priapulida/Entropia	TVM+I+G	MtArt+I+G
Tardigrada/Entropia	TVM+I+G	MtArt+I+G
Onychophora /Blocos	GTR+I+G	MtArt+I+G
Priapulida/Blocos	GTR+I+G	MtArt+I+G
Tardigrada/Blocos	GTR+I+G	MtArt+I+G

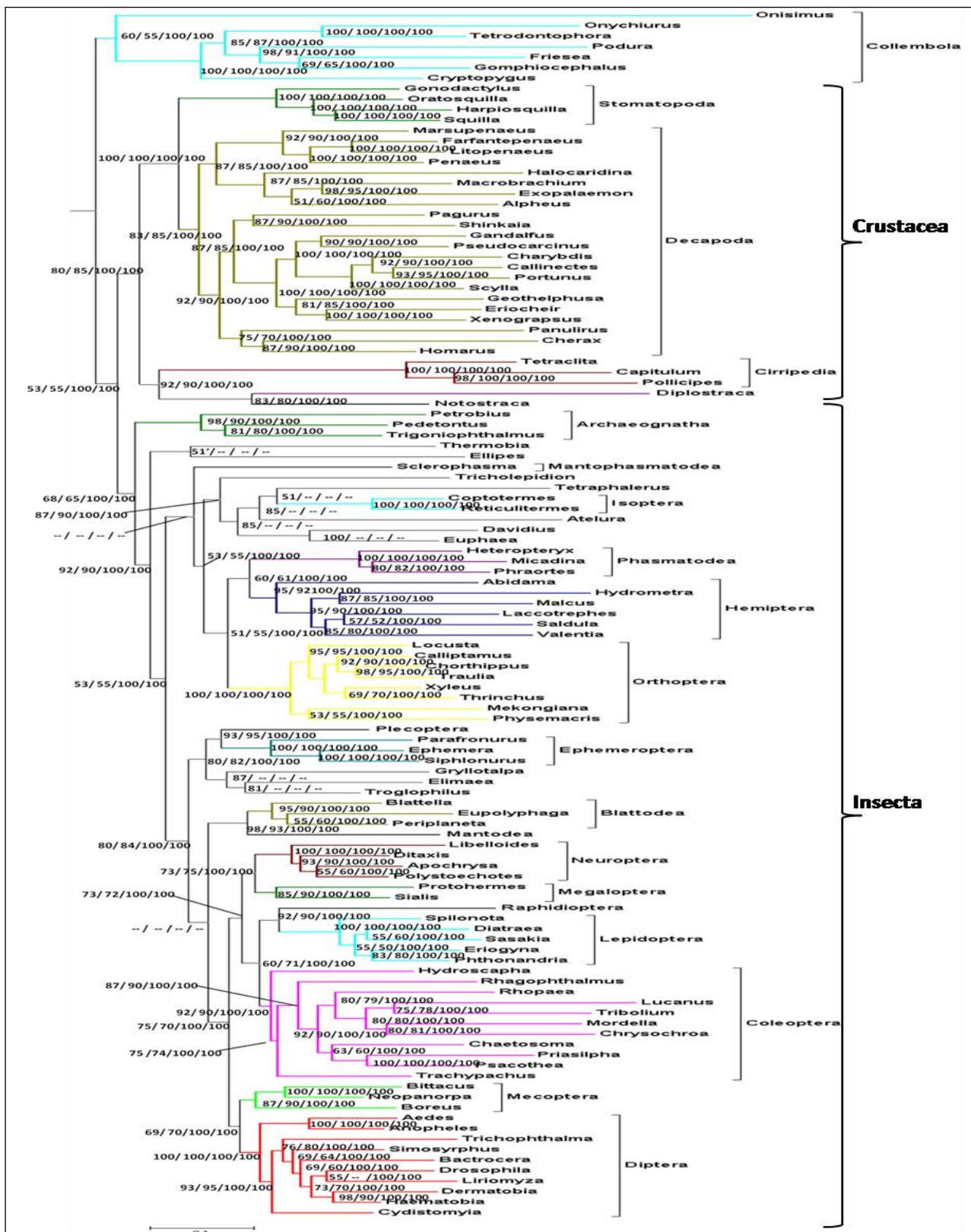
Entre os três grupos externos utilizados para a reconstrução filogenética do filo Arthropoda, Tardigrada foi o que apresentou o menor número de inconsistências entre as análises de máxima verossimilhança e inferência bayesiana tanto para nucleotídeos quanto para aminoácidos. Além disso, as análises realizadas com o grupo Tardigrada geraram um maior número de clados monofiléticos quando comparadas às análises usando Priapulida ou Onychophora como grupos externos.

As figuras 11 e 12 apresentam a reconstrução filogenética dos Arthropoda com as regiões de baixa entropia e as figuras 13 e 14 apresentam a reconstrução filogenética dos Arthropoda com os blocos conservados. As filogenias reconstruídas a partir de regiões de baixa entropia e blocos conservados apresentaram grande correspondência entre si, além de também apresentarem altos valores de suporte na maioria dos ramos.

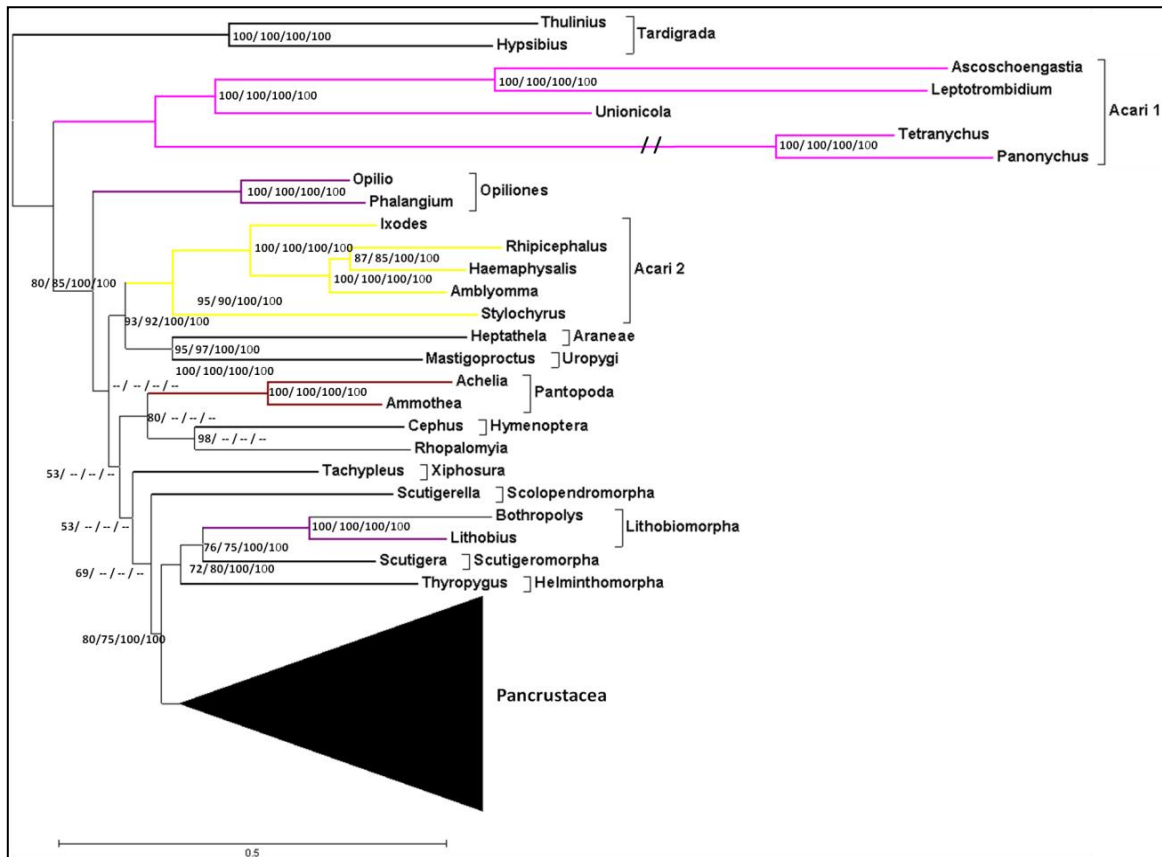
Foi possível observar a formação de um grande grupo composto pela maior parte das ordens dos Hexapoda. Esse clado corresponde à classe Insecta e apresenta altos valores de suporte e grande congruência entre os métodos utilizados. No entanto, nas análises que foram utilizados os sítios de baixa entropia, a ordem irmã das demais desse clado foi Archaeognatha. Enquanto que nas análises com os sítios de blocos conservados o grupo irmão foi formado pelas ordens Phasmatodea, Hymenoptera e Hemiptera.

Ambos os conjuntos de dados suportaram a combinação de Crustacea e Hexapoda em um único clado monofilético. No entanto, a posição do subfilo Crustacea variou, em relação ao grupo Insecta, a depender do método de seleção de sítios empregado. Nas análises com sítios de baixa entropia, o grupo Crustacea se posicionou como grupo irmão de Insecta. Nas análises com sítios em blocos conservados, o grupo Crustacea ficou inserido no clado Insecta. Ou seja, ambos os métodos suportam a hipótese de Pancrustacea, no entanto, não suportam a mesma posição filogenética dos Crustacea. A ordem Collembola foi o grupo irmão das demais ordens de Pancrustacea nos dois métodos empregados.

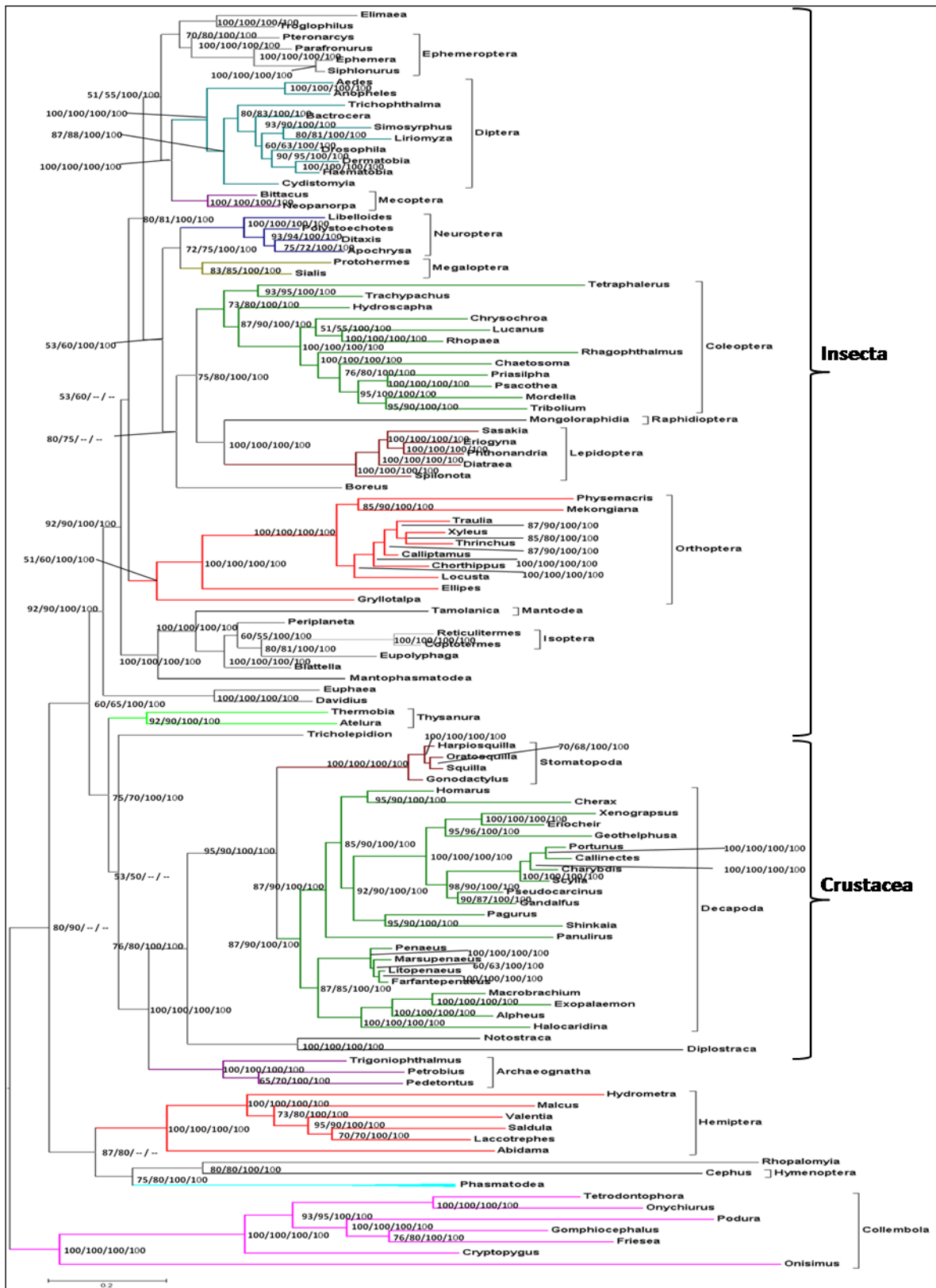
O subfilo Myriapoda não se mostrou monofilético em nenhuma das análises filogenéticas realizadas. No entanto, em todas as análises, o grupo formou um clado com Pancrustacea suportando a hipótese do grupo Mandibulata. O subfilo Chelicerata ficou na base da árvore do filo Artropoda, indicando que esse subfilo representa a origem do filo sendo o grupo mais antigo dos Arthropoda. Dentro do subfilo, a ordem Acari se dividiu em dois grupos que foram nomeados como Acari 1 e Acari 2 para melhor entendimento das árvores filogenéticas aqui apresentadas.



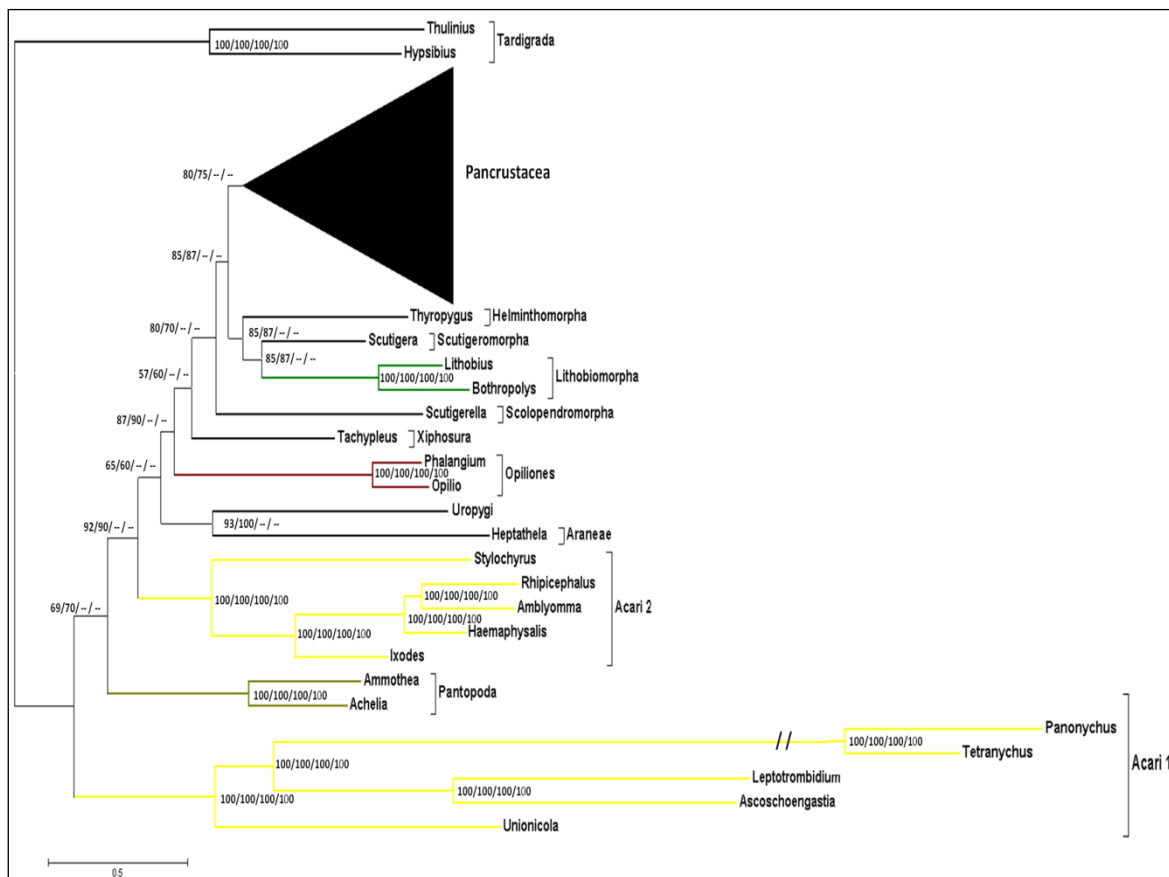
**Figura 23.** Árvore filogenética do filo Arthropoda baseada nas regiões selecionadas por entropia (5772 pb). Topologia determinada por ML, os valores de suporte dos ramos representam: *bootstrap* da ML com nucleotídeos/*bootstrap* da ML com aminoácidos/Probabilidade Bayesiana com nucleotídeos/Probabilidade Bayesiana com aminoácidos. Incongruências nas topologias e valores de *bootstrap* abaixo de 50% não estão representadas.



**Figura 24.** Continuação da árvore filogenética do filo Arthropoda baseada nas regiões selecionadas por entropia (5772 pb). Topologia determinada por ML, os valores de suporte dos ramos representam: *bootstrap* da ML com nucleotídeos/*bootstrap* da ML com aminoácidos/Probabilidade Bayesiana com nucleotídeos/Probabilidade Bayesiana com aminoácidos. Incongruências nas topologias e valores de *bootstrap* abaixo de 50% não estão representadas.



**Figura 25.** Árvore filogenética do filo Arthropoda baseada nos blocos conservados selecionados (7224 pb). Topologia determinada por ML, os valores de suporte dos ramos representam: *bootstrap* da ML com nucleotídeos/*bootstrap* da ML com aminoácidos/Probabilidade Bayesiana com nucleotídeos/Probabilidade Bayesiana com aminoácidos. Incongruências nas topologias e valores de *bootstrap* abaixo de 50% não estão representadas.



**Figura 26.** Continuação da árvore filogenética do filo Arthropoda baseada nos blocos conservados selecionadas (7224 pb). Topologia determinada por ML, os valores de suporte dos ramos representam: *bootstrap* da ML com nucleotídeos/*bootstrap* da ML com aminoácidos/Probabilidade Bayesiana com nucleotídeos/Probabilidade Bayesiana com aminoácidos. Incongruências nas topologias e valores de *bootstrap* abaixo de 50% não estão representadas.



## 6. Discussão

Uma série de trabalhos tem sido desenvolvida com o intuito de elucidar a história filogenética do filo Arthropoda, utilizando parte ou toda a informação do genoma mitocondrial (Hassanin, 2006; Carapelli *et al.*, 2007 e Koenemann *et al.*, 2010). No entanto, este é o primeiro trabalho a utilizar um número tão grande de genomas, assim como também é o primeiro a comparar o efeito da utilização de regiões de baixa entropia na reconstrução filogenética do grupo e na definição de iniciadores para a amplificação de genomas mitocondriais.

Já havia sido sugerido por Boore *et al.* (2002) e por Lavrov *et al.* (2004) que rearranjos na ordem gênica mitocondrial são fontes úteis de informação no estudo das relações evolutivas entre os grandes grupos dos Arthropoda. O arranjo gênico identificado no presente trabalho, presente na maioria das ordens de Pancrustacea, já havia sido apontado por Lavrov *et al.* (2004) como sendo àquele que provavelmente reflete a sintenia ancestral do grupo. Dessa forma, ficou claro que o arranjo gênico ancestral dos Pancrustacea é bem representado nos bancos de dados públicos, sugerindo que a reorganização do genoma mitocondrial não afetou a maior parte dos grupos representados.

A alta diversidade de ordens gênicas em Myriapoda, assim como o seu baixo nível de representação nos bancos de dados, dificultou a determinação da posição filogenética do grupo em relação aos outros grupos de Arthropoda. Gai *et al.* (2008) e Podsiadlowski *et al.* (2008) também ressaltaram a dificuldade de se indicar a posição filogenética deste grupo através da ordem gênica mitocondrial.

A organização sintênica mais representada em Cheliceratha foi a mesma que Boore *et al.* (1995) indicaram como sendo a ancestral dos Arthropoda; tal afirmação, no entanto,

somente poderia ser feita ao se assumir que este subfilo é o grupo irmão dos demais Arthropoda, o que é suportado pelas análises filogenéticas realizadas com os conjuntos de dados de nucleotídeos e de aminoácidos analisados neste estudo.

O viés de GC% se mostrou uma ferramenta importante na seleção de genomas mitocondriais para análises filogenéticas. Desvios neste parâmetro têm sido relatados em vários genomas mitocondriais de metazoários (Saccone *et al.*, 1999) e foram relacionados como fonte de perturbação do sinal filogenético nos Arthropoda (Jones *et al.*, 2007; Hassanin *et al.*, 2005; Hassanin, 2006). Rota-Stabelli *et al.* (2010) também demonstraram a eficiência dessa variável na identificação de um conjunto de genomas propícios para a reconstrução filogenética baseada em genomas mitocondriais. Assim, a utilização do viés GC% na seleção dos genomas a serem analisados representa um importante passo metodológico para a melhoria das inferências filogenéticas do filo Arthropoda.

A análise das sequências dos genes mitocondriais demonstrou um alto grau de variabilidade, sendo que até mesmo o gene mais conservado (COX1) apresentou uma considerável proporção de sítios variáveis (72%). A alta taxa de variação dos genes mitocondriais de Arthropoda reflete a antiguidade do grupo e a alta taxa mutacional característica desse marcador. Ressalta-se, deste modo, a necessidade de se adotar passos anteriores às reconstruções filogenéticas, visando à exclusão de regiões e de genomas que possam enviesar as análises.

O uso do método de seleção de blocos conservados para a seleção de regiões filogeneticamente informativas, utilizado na ferramenta GBLOCKS (Castresana, 2000), foi empregado em diferentes estudos com genomas mitocondriais de Arthropoda (Cook *et al.*, 2005; Podsiadlowski *et al.*, 2008; Rota-Stabelli *et al.*, 2010) e com outros marcadores moleculares e organismos (eg. Ibarra-Laclette *et al.*, 2011). Talavera e Castresana (2007), trabalhando com sequências reais e simuladas com alto grau de divergência, demonstraram a eficiência deste método na melhoria das reconstruções filogenéticas.

Neste trabalho, este método de análise foi comparado com um baseado na entropia de Shannon. O método de entropia mostrou-se mais adequado na identificação de regiões filogeneticamente informativas dos genomas mitocondriais e na conseqüente diminuição da saturação mutacional. Rota-Stabelli e Telford (2008) já consideravam que se o sinal filogenético de um conjunto de dados é fraco, a inclusão de pequenas regiões sem informação filogenética pode até mesmo eliminar o verdadeiro sinal. Nestes casos, a utilização de métodos que ajudem a discernir e remover as regiões pouco informativas é extremamente recomendada.

O desempenho superior da entropia de Shannon na recuperação de regiões filogeneticamente informativas justifica a sua aplicação em análises filogenéticas, principalmente para estudos de grupos antigos como os Arthropoda. Batista *et al.* (2011), trabalhando com genomas de Pappilomavírus, também obtiveram árvores filogenéticas com níveis mais altos de confiança quando utilizaram a entropia para a seleção de regiões filogeneticamente informativas.

Linhart e Shamir (2002) sugeriram a aplicação da entropia de Shannon na seleção de regiões para a definição de iniciadores degenerados capazes de amplificar um conjunto de sequências homólogas. O método, entretanto, tem sido pouco utilizado deste então, sendo este o primeiro trabalho a fazer uso da entropia de Shannon na indicação de iniciadores degenerados para a amplificação do genoma mitocondrial de Arthropoda.

Gibson *et al.* (2011) e Simon *et al.* (2006) revisaram e compilaram uma lista com iniciadores utilizados em estudos filogenéticos da ordem Diptera. Boa parte dessa lista corresponde a iniciadores voltados para a amplificação de regiões do genoma mitocondrial (principalmente dos genes CYTB, COX1 e ribossomal curto), que são amplamente utilizados em estudos filogenéticos. No entanto, os fragmentos produzidos por esses iniciadores cobrem apenas uma pequena região do genoma mitocondrial desta ordem.

Barau *et al.* (2005), baseados nas análises do genoma de *Haematobia irritans*, sugeriram um conjunto de seis iniciadores capazes de amplificar todo o genoma mitocondrial dos dípteros e de algumas outras ordens de Arthropoda. Porém, esse conjunto de iniciadores se baseia na técnica de PCR longa, exigindo assim, etapas de clonagem dos *amplicons* e a posterior utilização do método *shotgun* para a realização do sequenciamento do genoma mitocondrial.

O conjunto de 22 pares de iniciadores sugerido no presente estudo foi estabelecido com o intuito de se produzir *amplicons* com tamanho máximo de 1200 pares de bases e com sobreposição mínima de 100 pares de bases entre os *amplicons*, resultando na cobertura total do genoma mitocondrial. Assim, os iniciadores podem ser utilizados no sequenciamento de todo o genoma mitocondrial dos dípteros, sem a necessidade das etapas de clonagem e do uso do método *shotgun*. Dessa forma, os iniciadores aqui sugeridos representam o primeiro conjunto de iniciadores para a amplificação de todo o genoma mitocondrial dos dípteros, sem a necessidade de métodos de clonagem e *shotgun*.

A reconstrução filogenética da ordem Diptera aqui descrita coincide com a topologia atualmente aceita para o grupo (Grimaldi e Engel, 2005), demonstrando que os dois fragmentos (correspondendo a porções dos genes COX1, ND2 e RNAt de metionina) utilizados nas análises são eficientes na reconstrução da história evolutiva da ordem. Dessa forma, considera-se que esses iniciadores se apresentam como boas alternativas para o estudo evolutivo das espécies de Diptera, demonstrando ainda a viabilidade do uso do conjunto de 22 iniciadores voltados para a amplificação de todo o genoma mitocondrial da ordem.

Na reconstrução filogenética do filo Arthropoda, a melhor adequação dos Tardigrada como grupo externo não coincide com observações de trabalhos anteriores, que também realizaram comparações entre vários grupos, visando a determinação de qual táxon se

mostraria como o grupo externo mais adequado para inferências filogenéticas entre os Arthropoda. Rota-Stabelli e Telford (2008) concluíram que o filo Priapulida representava a melhor opção como grupo externo em análises filogenéticas do filo Arthropoda. No entanto, estes autores não puderam avaliar a eficiência do grupo Tardigrada como tal, uma vez que nenhum genoma mitocondrial deste grupo estava disponível para análise. Em Rota-Stabelli *et al.* (2010), Priapulida também foi tido como o grupo externo mais indicado, apesar da maior proximidade filogenética entre Tardigrada, Onychophora e Arthropoda. No entanto, os autores destacaram a melhora do desempenho de Tardigrada após a retirada de algumas espécies da ordem Araneae das análises filogenéticas, indicando que estas espécies estariam enviesando as análises. No presente estudo, a melhor adequação de Tardigrada como grupo externo provavelmente resultou da retirada prévia de cinco das seis espécies da ordem Araneae, decorrente da detecção de alteração do viés GC%, e de um número maior de genomas analisados, já que Rota-Stabelli *et al.* (2010) descrevem a análise de apenas 66 genomas.

A monofilia de Chelicerata é geralmente considerada como uma das questões menos controversa na sistemática dos Arthropoda [por exemplo, Giribet *et al.* (2005) demonstraram a monofilia do grupo ao combinar dados morfológicos com os de nove genes codificantes de proteínas]. No entanto, no presente estudo Chelicerata se mostrou parafilético, tal resultado pode ter sido causado pela grande diversidade e número de genomas utilizados, já que outros estudos que utilizaram um número menor de genomas mitocondriais suportaram o estado monofilético de Chelicerata (Cook *et al.*, 2005; Hassanin *et al.*, 2005; Rota-Stabelli *et al.*, 2010).

Com relação ao clado Acari, que também se mostrou parafilético, autores de trabalhos com dados moleculares vêm sugerindo a elevação taxonômica do grupo para superclasse, e integração do grupo Opiliones à superordem Parasitiformes (Murrell *et al.*, 2005). Os dados mitocondriais aqui analisados não suportam a integração dos Opiliones

ao grupo Parasitiformes, já que o clado Opiliones se mostrou monofilético e não apresentou relação de grupo irmão com Parasitiformes (Acari 2 nas figuras 12 e 14). No entanto, Klompen *et al.*, 2007, resalta que as relações filogenéticas do grupo Acari ainda apresentam muitas divergências, deixando em aberto a questão.

As análises filogenéticas suportaram o monofiletismo do subfilo Crustacea, tanto com uso de entropia quanto com os blocos conservados. Análises de genes nucleares têm tipicamente suportado Crustacea como grupo não monofilético (Edgecombe, 2010). Trabalhos baseados em dados morfológicos e genômica mitocondrial defendem o monofiletismo do grupo (Giribet *et al.*, 2005).

O estado parafilético de Crustacea em relação a Hexapoda é um tema recorrente, principalmente no que diz respeito à definição de qual grupo de Crustacea é irmão de Hexapoda. Os candidatos apontados na literatura como o grupo irmão de Hexapoda são os Branchiopoda (Fanenbruck e Harzsch, 2005; Harzsch, 2006) ou Copepoda mais Branchiopoda (Mallatt *et al.*, 2004, Mallatt e Giribet, 2006, von Reumont *et al.*, 2009), e um clado que combina Remipedia e Cephalocarida (Regier *et al.*, 2008). Apesar da monofilia de Crustacea não ser suportada pela maioria dos dados moleculares, no presente trabalho os dados de genoma mitocondrial indicaram, com valores de suporte elevados, a monofilia do grupo, resultado também alcançado por Hassanin *et al.* (2005). Assim, as análises do genoma mitocondrial suportam o monofiletismo do grupo e corroboram com os dados morfológicos.

A reconstrução filogenética baseada nos genomas mitocondriais indicou Hexapoda como grupo parafilético, com Collembola em posição basal para a união entre Hexapoda e Crustacea. Esse posicionamento de Collembola em relação à Hexapoda geralmente é identificado com dados de genomas mitocondriais (Nardi *et al.*, 2003; Hassanin, *et al.*, 2005; Hassanin, *et al.*, 2006). Outras análises também suportaram esse padrão parafilético de Hexapoda, com o grupo Diplura colocado em posição ainda mais basal que

Collembola (Carapelli *et al.*, 2007). Reconstruções filogenéticas geradas a partir de genes nucleares contrastam com àquelas da mitogenômica, que suportam o monofiletismo de hexapoda. Assim como os dados morfológicos que suportam Collembola como um grupo de Hexapoda primitivamente sem asas (Regier *et al.*, 2004, 2008;. Mallatt e Giribet, 2006;. Timmermans *et al.*,2008; von Reumont *et al.*, 2009).

Os genomas mitocondriais analisados no presente trabalho suportaram a formação do clado Pancrustacea composto por Crustacea e Hexapoda. A formação de tal grupo tem sido recorrente em análises moleculares de Arthropoda, sendo primeiro defendida por Wheeler *et al.* (1993). Do ponto de vista morfológico, a maioria dos argumentos em favor da monofilia deste grupo vêm de estudos do sistema nervoso (Richter, 2002; Harzsch, 2006, e Strausfeld *et al.* 2006).

As evidências moleculares corroboram as semelhanças neuroanatômicas entre Hexapoda e Crustacea, não apenas por causa dos altos valores de suporte encontrados em análises de sequências biológicas, mas também porque Pancrustacea é independentemente recuperada por vários tipos de informações moleculares, conforme revisto por Edgecombe (2010).

A monofilia de Mandibulata (Pancrustacea + Myriapoda) tem sido suportada por vários dados morfológicos (Bitsch, 2001; Edgecombe *et al.*, 2003; Harzsch, 2005). Além disso, estudos moleculares também apresentaram o clado Mandibulata como monofilético. Esses estudos utilizaram a análise de genes ribossomais nucleares (Giribet e Ribera, 1998), sequências de hemocianina (Kusche *et al.*, 2003), genes nucleares codificantes de proteínas (Regier *et al.*, 2008), combinação de genes nucleares e genes que codificam os ribossomos mitocondriais (Bourlat *et al.*, 2008), e expressão gênica (Prpic e Tautz, 2003). Assim, a mitogenômica se une a análises morfológicas e outras análises moleculares no suporte ao monifiletismo de Mandibulata.

## 7. Conclusões

Os genomas mitocondriais do filo Arthropoda apresentam um alto grau de variabilidade genética e viés GC%.

O uso da ordem gênica do genoma mitocondrial suporta a formação dos grupos Pancrustacea (Hexapoda + Crustacea).

O uso da ordem gênica do genoma mitocondrial não é capaz de indicar uma posição filogenética mais robusta para Myriapoda devido à alta diversidade da ordem gênica mitocondrial e ao pequeno número de genomas sequenciados do grupo.

A utilização da entropia de Shannon se mostra altamente eficiente na seleção de regiões filogeneticamente informativas, sendo mais efetiva que outros métodos utilizados tradicionalmente na análise de genomas mitocondriais.

A entropia de Shannon também se mostra como ótima fonte de informações para seleção de regiões conservadas destinadas à definição de iniciadores degenerados para a amplificação de regiões homólogas.

As análises dos genomas mitocondriais sugerem com altos valores de suporte a formação e monofiletismo do grupo Crustacea.

O grupo Tardigrada se mostra melhor grupo externo para análises filogenéticas do grupo Arthropoda com dados de genomas mitocondriais.

As análises dos genomas mitocondriais sugerem com altos valores de suporte a formação e monofiletismo dos grupos Pancrustacea (Hexapoda + Crustacea), e Mandibulata (Pancrustacea + Myriapoda). As informações aqui apresentadas indicam que o grupo irmão dos demais Arthropoda é Chelicerata.



## 8. Referências Bibliográficas

Abascal F, Zardoya R, Posada D (2005) ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics*: 21:2104-2105.

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science* 24:2185-2195.

Alberts B, Johnson A, Lewis J, Roberts MRK, Walter P (2008) *Molecular Biology of Cell*. 5ª edição. Garland science, New York. 1725pp.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403-410.

Barau JG., Azeredo-Espin AML, Lessinger AC (2005), Conservation and versatility of a new set of iniciadores for long-PCR amplification of complete insect mitochondrial genomes based on *Haematobia irritans* mtDNA sequences. *Mol Ecol Notes* 5:885–887.

Batista MV, Ferreira TA, Freitas AC, Balbino VQ (2011) An entropy-based approach for the identification of phylogenetically informative genomic regions of Papillomavirus. *Infect Genet Evol.* 11:2026-2033.

Baumgartner OL. e Greenberg B (1984). The genus *Chrysomya* (Diptera: Calliphoridae) in the New World. *J Med Entomol* 21:105-113.

Bitsch J (2001). The arthropod mandible: morphology and evolution. Phylogenetic implications. *Ann Soc Entomol* 37:305-321.

Boore JL, Collins TM, Stanton DD, Brown LL, Wesley MTI (1995) Deducing the pattern of arthropod phylogeny from mitochondrial DNA Rearrangements *Nature*. *Nature* 376:163-165.

Boore JL, Lavrov DV, Brown WM (1998). Gene translocation links insects and crustaceans. *Nature* 392:667-668.

Boore JL (1999). Animal mitochondrial genomes. *Nucleic Acids Res* 27:1767–1780.

Bourlat SJ, Nielsen C, Economou AD, Telford MA (2008). Testing the new animal phylogeny: a phylum level molecular analysis of the animal kingdom. *Mol Phylogenet Evol* 49, 23–31.

Brusca RC e Brusca GJ (2003) *Invertebrates*. 2nd edition. Sinauer, Sunderland, Massachusetts. 966pp.

Budd GE, Telford MJ (2009). The origin and evolution of arthropods. *Nature* 457:812–817.

Carapelli A, Liò P, Nardi F, van der Wath E, Frati F (2007) Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evol Biol* 7:Suppl 2:S8.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540-552.

Cook CE, Yue Q, Akam M (2005). Mitochondrial genomes suggest that hexapods and crustaceans are mutually paraphyletic. *Proc Biol Sci* 272:1295–1304.

Courtney GW e Merritt RW (2008). Capítulo 22. Aquatic Diptera. Part one. Larvae of aquatic Diptera. In Merritt RW, Cummins KW, Berg MB (eds.). *An Introduction to the Aquatic Insects of North America*. 4a edição. Kendall/Hunt Publishing Co. Dubuque, Iowa, pp. 687-722.

Criscuolo A, Gribaldo S (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol* 13;10:210.

Curole, JP e Kocher TD (1999) Mitogenomics: digging deeper with complete mitochondrial genomes. *Trends Ecol Evol* 14:394-398.

Dear, J.P. 1985. A revision of the New World Chrysomini (Diptera: Calliphoridae). *Rev Bras Zool* 3: 109-169.

Deitel HM, Deitel TR, Nieto DC, McPhie PJ (2001) Perl: How to Programming. 1st Edition, Prentice Hall, 325pp.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792-1797.

Edgecombe GD, Richter S, Wilson GDF (2003). The mandibular gnathal edges: homologous structures across Mandibulata? In: Hamer, M. (ed.), Proceedings of the Twelfth International Congress of Myriapodology. *African Invertebrates* 44, 115–135.

Edgecombe GD (2010) Arthropod phylogeny: An overview from the perspectives of morphology, molecular data and the fossil record. *Arthropod Struct Dev* 39:74–87.

Evenhuis NL (1994). Catalogue of the Fossil Flies of the World (Insecta:Diptera). Backhuys Publishers, Leiden, Netherlands, 600 pp.

Ewing B, Hillier L, Wendl M, Green P: (1998) Basecalling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8:175-185.

Fanenbruck M e Harzsch S (2005) A brain atlas of *Godzillionomus frondosus* Yager, 1989 (Remipedia, Godzilliidae) and comparison with the brain of *Speleonectes tulumensis* Yager, 1987 (Remipedia, Speleonectidae): implications for arthropod relationships. *Arthropod Struct Dev* 34:343–378.

Friedrich M e Tautz D (1995) Ribosomal DNA phylogeny of the major extant arthropod classes and the evolution of myriapods. *Nature* 376:165–167.

Gai Y, Song D, Sun H, Yang Q, Zhou K (2008). The complete mitochondrial genome of *Symphylella* sp. (Myriapoda: Symphyla): Extensive gene order rearrangement and evidence in favour of Progoneata. *Mol Phylogenet Evol* 48:103–111.

Gião JZ e Godoy WAC (2006) Seasonal population dynamics in *Lucilia eximia* (Wiedemann) (Diptera, Calliphoridae). *Neotrop Entomol* 35:753–756.

Joel F, Gibson, SK, Jackson MD, Kits JH, Miranda GFG, SKEVINGTON JH (2011) Diptera-Specific Polymerase Chain Reaction Amplification Inicialdores of Use in Molecular Phylogenetic Research. *Ann Entomol Soc Am* 104:976-997

Gascuel O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol* 14:685-695.

Giribet G, Carranza S, Bagunà J, Riutort M, Ribera C (1996). First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Mol Biol Evol* 13:76–84.

Giribet G e Ribera C (1998) The position of arthropods in the animal kingdom: a search for a reliable outgroup for internal arthropod phylogeny. *Mol Phylogenet Evol* 9:481–488.

Giribet G, Richter S, Edgecombe GD, Wheeler WC (2005) The position of crustaceans within Arthropoda – evidence from nine molecular loci and morphology. In: Koenemann S, Jenner R (Eds.) Crustacea and Arthropod Relationships. Crustacean Issues, 16. CRC Press, Taylor e Francis, pp. 307–352.

Grimaldi D e Engel MS (2005). Evolution of the Insects. 1st Edition, Cambridge University Press, Cambridge, 770pp.

Guimarães JH e Papavero N, (1999). Myiasis in man and animals in the Neotropical Region. Bibliographic database. Editora Plêiade/Fapesp. 308 pp.

Guimarães JH, Prado AP, Linhares AX (1978) Tree newly introduced blowfly species in southern Brazil (Diptera, Calliphoridae). Revista Brasileira de Entomologia 22:53–60.

Guindon S e Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696-704.

Hall M e Wall R (1995) Myiasis os humans and domestic animals. Adv. Parasitol 35:257-334.

Hardison RC, Roskin KM, Yang S, Diekhans M, Kent WJ (2003) Covariation in frequencies of substitution, deletion, transposition and recombination during eutherian evolution. *Genome Res* 13:13–26.

Harzsch S, Muöller CHG, Wolf H (2005). From variable to constant cell numbers: cellular characteristics of the arthropod nervous system argue against a sistergroup relationship of Chelicerata and “Myriapoda” but favour the Mandibulata concept. *Dev Genes Evol* 215:53–68.

Harzsch S (2006) Neurophylogeny: architecture of the nervous system and a fresh view on arthropod phylogeny. *Integr Comp Biol* 46:162–194.

Hassanin A, Léger N, Deutsch J (2005). Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of metazoa, and consequences for phylogenetic inference. *Syst Biol* 54:277–298.

Hassanin A (2006) Phylogeny of Arthropoda inferred from mitochondrial sequences: strategies for limiting the misleading effects of multiple changes in pattern and rates of substitution. *Mol Phylogenet Evol* 38:100–116.

Ibarra-Laclette E, Albert VA, Pérez-Torres CA, Zamudio-Hernández F, Ortega-Estrada MJ, Herrera-Estrella A, Herrera-Estrella L (2011) Transcriptomics and molecular evolutionary rate analysis of the bladderwort (*Utricularia*), a carnivorous plant with a minimal genome. *BMC Plant Biol* 11:101.

Jones M, Gantenbein B, Fet V, Blaxter M (2007) The effect of model choice on phylogenetic inference using mitochondrial sequence data: Lessons from the scorpions. *Mol Phylogenet Evol* 43:583–595.

Klompen H, Lekveishvili M, Black IV WC. Phylogeny of parasitiform mites (Acari) based on rRNA. *Mol Phylogenet Evol* 43:936–951.

Koenemann S, Jenner RA, Hoenemann M, Stemme T, von Reumont BM (2010) Arthropod phylogeny revisited, with a focus on crustacean relationships. *Arthropod Struct Dev* 39:88–110.

Kusche K, Hembach A, Hagner-Holler S, Gebauer W, Burmester T (2003). Complete subunit sequences, structure and evolution of the 6x6-mer hemocyanin from the common house centipede, *Scutigera coleoptrata*. *Eur J Biochem* 270:2860–2868.

Lavrov DV, Boore JL, Brown WM (2000) The Complete Mitochondrial DNA Sequence of the Horseshoe Crab *Limulus polyphemus*. *Mol Biol Evol* 17:813–824.

Lavrov DV, Brown WM, Boore JL (2004) Phylogenetic position of the Pentastomida and (pan)crustacean relationships. *Proc R Soc Lond B Biol Sci* 271, 537–544.



Lindsley DL e Zimm GG (1992). *The Genome of Drosophila melanogaster* (San Diego, California: Academic Press). San Diego, CA.133 pp.

Linhart C e Shamir R. The degenerate iniciador design problem (2002) *Bioinformatics*18 Suppl 1:S172-81.

Mallatt JM, Garey JR, Shultz JW (2004) Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin. *Mol Phylogenet Evol* 31:178–191.

Mallatt JM, Giribet G (2006) Further use of nearly complete 28S and 18S rRNA genes to classify Ecdysozoa: 37 more arthropods and a kinorhynch. *Mol Phylogenet Evol* 40:772–794.

Murrell A, Dobson SJ, Walter DE, Campbell NJH, Shao RF, Barker SC (2005) Relationships among the three major lineages of the Acari (Arthropoda: Arachnida) inferred from small subunit rRNA: paraphyly of the parasitiformes with respect to the opilioacariformes and relative rates of nucleotide substitution. *Invert Syst* 19:383-389.

Nardi F, Spinsanti G, Boore JL, Carapelli A, Dallai R, Frati F (2003) Hexapod origins: monophyletic or paraphyletic? *Science* 299:1887–1889.

Nei M e Kumar S (2000) *Molecular Evolution and Phylogenetics*. 1ª edição, Oxford University Press, New York, 333pp.

Podsiadlowski L, Braband A, Mayer G (2008) The Complete Mitochondrial Genome of the Onychophoran *Epiperipatus biolleyi* Reveals a Unique Transfer RNA Set and Provides Further Support for the Ecdysozoa Hypothesis. *Mol Biol Evol* 25(1):42–51.

Posada D e Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817-8.

Prpic NM e Tautz D (2003) The expression of the proximodistal axis patterning genes *Distal-less* and *dachshund* in the appendages of *Glomeris marginata* (Myriapoda: Diplopoda) suggests a special role of these genes in patterning the head appendages. *Dev Biol* 260:97–112.

Regier JC, Shultz JW, Ganley ARD, Hussey A, Shi D, Ball B, Zwick A, Stajich JE, Cummings MP, Martin JW, Cunningham CW (2008) Resolving arthropod phylogeny: exploring phylogenetic signal within 41 kb of protein-coding nuclear gene sequences. *Syst Biol* 57:920–938.

Regier JC, Shultz JW, Kambic RE (2004) Phylogeny of basal hexapod lineages and estimates of divergence times. *Ann Entomol Soc Am* 97:411-419.

Richter S (2002) The Tetraconata concept: hexapod-crustacean relationships and the phylogeny of Crustacea. *Organ Div Evol* 2:217–237.

Rota-Stabelli O e Telford MJ (2008). A multi criterion approach for the selection of optimal outgroups in phylogeny: Recovering some support for Mandibulata over Myriochelata using mitogenomics. *Mol Phylogenet Evol* 48:103–111.

Rota-Stabelli O, Kayal E, Gleeson D, Daub J, Boore JL, Telford MJ, Pisani D, Blaxter M, Lavrov DV (2010) Ecdysozoan Mitogenomics: Evidence for a Common Origin of the Legged Invertebrates, the Panarthropoda. *Genome Biol Evol* 2:425–440.

Saccone C, De Giorgi C, Gissi C, Pesole G, Reyes A. (1999) Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene* 238:195–209.

Sepp S, Szabó I, Uda H, Sakamoto H. (1994) Rapid techniques for DNA extraction from routinely processed archival tissue for use in PCR. *J Clin Pathol* 47:318-323.

Shannon CE e Weaver W. *The Mathematical Theory of Communication*. (1949) Urbana, Illinois: University of Illinois Press.

Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT (2006) Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction iniciadores for animal mitochondrial DNA. *Annu Rev Ecol Evol Syst* 37:545-579.

Staden R, Judge DP, Bonfield JK (2003) Analysing Sequences Using the Staden Package and EMBOSS. In Krawetz SA e Womble DD (eds) Introduction to Bioinformatics. A Theoretical and Practical Approach. Human Press Inc., Totawa. pp393-410.

Ronquist F e Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.

Ronquist F, van der Mark P, Huelsenbeck JP (2009) Bayesian phylogenetic analysis using MRBAYES. In Lemey P, Salemi M, Vandamme AM (eds) *The Phylogenetic Handbook* Second Edition. pp210-236.

Schmidt HA e von Haeseler (2009) A Phylogenetic inference using maximum likelihood methods. In Lemey P, Salemi M, Vandamme AM (eds) *The Phylogenetic Handbook* Second Edition. pp181-207.

Strausfeld NJ, Strausfeld CM, Loesel R, Rowell D, Stowe S (2006) Arthropod phylogeny: onychophoran brain organization suggests an archaic relationship with a chelicerate stemlineage. *Proc Biol Sci* 273:1857–1866.

Talavera G e Castresana J (2007) Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst Biol* 56:564-577.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 28:2731-2739.

Timmermans MJTN, Roelofs D, Mariën J, van Straalen NM (2008) Revealing pancrustacean relationships: phylogenetic analysis of ribosomal protein genes places Collembola (springtails) in a monophyletic Hexapoda and reinforces the discrepancy between mitochondrial and nuclear DNA markers. *BMC Evol Biol* 8:83.

Tisdall J (2001) *Beginning Perl for Bioinformatics*. 1st edition, O'REILLY, Cambridge. 369pp.

von Reumont BM, Meusemann K, Szucsich NU, Dell'Amplio E, Gowri-Shakar V, Bartel D, Simon S, Letsch HO, Stocsits RR, Luan YX et al., (2009). Can comprehensive background knowledge be incorporated into substitution models to improve phylogenetic analyses? A case study on major arthropod relationships. *BMC Evol Biol* 9:119.

Wägele JW e Stanjek G (1995) Arthropod phylogeny inferred from 12S rRNA revisited: monophyly of the Tracheata depends on sequence alignment. *J Zoo Syst Evol Res* 33:75–80.

Wei S, Shi M, Sharkey MJ, van Achterberg C, Chen X (2010) Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to Holometabolous insects. *BMC Genomics* 11:371.

Wheeler WC, Cartwright P, Hayashi CY (1993) Arthropod phylogeny: a combined approach. *Cladistics* 9:1–39.

Xia X e Xie Z (2001) DAMBE: data analysis and molecular biology and evolution. *J Hered* 92:371-373.

Xia X (2009) Assessing substitution saturation with Dambe. In Lemey P, Salemi M, Vandamme AM (eds) *The Phylogenetic Handbook Second Edition*. pp615-630.

## 9. Anexo 1. Viés GC% dos genes mitocondriais.

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso			
Chelicerata	Arachnida	Acari	Dermatophagoides farinae	0,198	-0,077	0,185	0,162	0,313	-0,224	0,031	-0,356	0,313	0,300	0,465	0,246	-0,293	NC_013184		
			Dermatophagoides pteronyssinus	0,204	0,171	0,120	0,157	0,254	-0,215	0,004	-0,284	0,273	0,277	0,323	0,175	-0,340	NC_012218		
			Steganacarus magnus	-0,252	0,077	0,104	0,033	0,075	-0,108	0,163	-0,106	0,289	-0,032	-0,097	0,005	-0,140	NC_011574		
			Panonychus citri	-0,175	0,000	0,009	-0,046	-0,008	0,005	0,000	-0,114	-0,059	0,124	-0,037	0,064	-0,179	NC_014347		
			Panonychus ulmi	-0,135	-0,333	0,003	0,019	0,041	-0,030	0,000	-0,132	-0,100	0,118	0,000	-0,033	-0,448	NC_012571		
			Tetranychus cinnabarinus	-0,034	-0,333	-0,011	0,066	0,095	-0,086	0,167	-0,120	0,056	-0,018	0,111	-0,037	0,000	NC_014399		
			Tetranychus urticae	-0,159	-0,167	-0,006	0,008	0,141	-0,141	0,036	-0,061	-0,070	0,000	0,120	-0,013	-0,106	NC_010526		
			Unionicola foili	-0,522	-0,421	-0,182	-0,268	-0,255	-0,268	0,300	-0,386	-0,181	0,099	0,406	0,297	-0,531	NC_011036		
			Unionicola parkeri	-0,445	-0,600	-0,205	-0,346	-0,214	-0,281	0,269	-0,390	-0,143	0,208	0,481	0,307	-0,413	NC_014683		
			Ascoschoengastia sp. TATW1	-0,047	0,000	-0,138	-0,014	0,016	-0,077	0,086	-0,092	-0,200	-0,044	0,081	-0,079	-0,018	NC_010596		
			Leptotrombidium pallidum	-0,159	-0,355	-0,057	-0,186	-0,084	-0,078	-0,120	-0,113	0,026	0,024	0,094	-0,033	-0,057	NC_007177		
			Walchia hayashii	-0,467	-0,600	-0,205	-0,426	-0,325	-0,269	0,408	-0,383	-0,341	0,250	0,521	0,285	-0,468	NC_010595		
			Carios capensis	-0,467	-0,722	-0,200	-0,282	-0,301	-0,409	0,508	-0,393	-0,405	0,398	0,865	0,442	-0,663	NC_005291		
			Ornithodoros moubata	-0,519	-0,515	-0,151	-0,323	-0,231	-0,382	0,555	-0,463	-0,380	0,527	0,821	0,433	-0,607	NC_004357		
			Ornithodoros porcinus	-0,404	-0,297	-0,133	-0,318	-0,252	-0,413	0,531	-0,309	-0,333	0,509	0,842	0,410	-0,554	NC_005820		
			Amblyomma triguttatum	-0,182	-0,391	-0,102	-0,288	-0,016	-0,213	0,069	-0,461	0,212	0,172	0,364	0,058	-0,220	NC_005963		
			Haemaphysalis flava	-0,165	-0,500	-0,064	-0,191	0,033	-0,165	0,108	-0,298	0,125	0,132	0,280	0,045	-0,324	NC_005292		
			Ixodes hexagonus	-0,463	-0,737	-0,179	-0,390	-0,314	-0,322	0,490	-0,579	-0,460	0,463	0,618	0,318	-0,657	NC_002010		
			Ixodes holocyclus	-0,315	-0,600	-0,083	-0,256	-0,173	-0,184	0,419	-0,422	-0,179	0,380	0,462	0,389	-0,242	NC_005293		
			Ixodes persulcatus	-0,401	-0,545	-0,098	-0,295	-0,170	-0,247	0,368	-0,360	-0,250	0,294	0,529	0,318	-0,507	NC_004370		
			Ixodes uriae	-0,464	-0,742	-0,157	-0,237	-0,280	-0,302	0,484	-0,469	-0,378	0,455	0,423	0,468	-0,476	NC_006078		
			Rhipicephalus sanguineus	-0,248	-0,625	-0,063	-0,219	-0,049	-0,084	-0,035	-0,279	0,093	0,092	0,375	0,006	-0,265	NC_002074		
			Stylochyrus rarior	-0,406	-0,813	-0,160	-0,254	-0,205	-0,295	0,398	-0,330	-0,299	0,407	0,536	0,453	-0,443	NC_013474		
			Phytoseiulus persimilis	-0,394	-0,360	0,176	0,283	-0,065	-0,183	0,354	-0,282	-0,246	0,380	0,632	0,268	0,348	NC_014049		
			Varroa destructor	0,215	-0,077	0,202	0,232	0,255	0,166	-0,044	0,451	0,347	-0,153	0,077	-0,264	0,143	NC_004454		
			Amblypygi		Damon diadema	-0,472	-0,765	-0,233	-0,379	-0,336	-0,379	0,619	-0,545	-0,486	0,645	0,696	0,619	-0,635	NC_011293
					Phrynus sp. 1 SEM2008	-0,475	-0,628	-0,192	-0,362	-0,338	-0,471	0,647	-0,487	-0,527	0,559	0,786	0,569	-0,592	NC_010775
			Araneae		Nephila clavata	0,267	0,304	0,232	0,178	0,339	0,184	-0,134	0,476	0,343	-0,395	-0,208	-0,321	0,683	NC_008063
					Habronattus oregonensis	0,264	0,515	0,244	0,347	0,220	0,309	-0,266	0,514	0,364	-0,421	-0,283	-0,376	0,677	NC_005942
					Hypochilus thorelli	0,266	0,257	0,229	0,297	0,253	0,273	-0,302	0,484	0,287	-0,366	-0,277	-0,343	0,381	NC_010777
					Heptathela hangzhouensis	-0,284	-0,590	-0,072	-0,196	-0,144	-0,214	0,403	-0,396	-0,270	0,282	0,507	0,253	-0,435	NC_005924
					Calisoga longitarsis	0,448	0,536	0,365	0,343	0,431	0,310	-0,247	0,614	0,411	-0,411	-0,434	-0,426	0,529	NC_010780
			Haplopelma schmidti	0,295	0,487	0,271	0,384	0,360	0,326	-0,257	0,586	0,526	-0,478	-0,325	-0,444	0,645	NC_005925		
Opiliones		Opilio parietinus	-0,385	-0,561	-0,121	-0,215	-0,173	-0,196	0,225	-0,235	-0,263	0,166	0,600	0,332	-0,316	NC_014700			
		Phalangium opilio	-0,340	-0,556	-0,109	-0,247	-0,172	-0,233	0,331	-0,219	-0,242	0,246	0,634	0,366	-0,293	NC_010766			
Ricnulei		Pseudocellus pearsei	-0,482	-0,745	-0,277	-0,414	-0,341	-0,504	0,588	-0,556	-0,543	0,588	0,872	0,650	-0,587	NC_009985			
Scorpiones		Buthus occitanus	0,427	0,640	0,316	0,392	0,414	0,078	0,095	0,516	0,545	-0,443	-0,132	-0,459	0,160	NC_010765			
		Centruroides limpidus	0,409	0,544	0,288	0,391	0,409	0,092	0,057	0,444	0,609	-0,470	-0,297	-0,488	0,268	NC_006896			

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		Mesobuthus martensii	0,447	0,692	0,354	0,333	0,448	0,020	0,043	0,537	0,615	-0,477	-0,273	-0,486	0,352	NC_009738
		Uroctonus mordax	0,243	0,484	0,167	0,284	0,284	0,099	0,132	0,174	0,383	-0,132	-0,016	-0,255	0,236	NC_010782
	Solifugae	Nothopuga sp. 1 LP2008	-0,442	-0,512	-0,145	-0,238	-0,216	-0,323	0,516	-0,377	-0,405	0,531	0,698	0,456	-0,568	NC_009984
		Eremobates cf. palpietulosus SEM2008	-0,422	-0,674	-0,163	-0,274	-0,261	-0,350	0,531	-0,417	-0,471	0,546	0,659	0,516	-0,614	NC_010779
	Uropygi	Mastigoproctus giganteus	-0,350	-0,467	-0,121	-0,336	-0,167	-0,298	0,408	-0,372	-0,327	0,332	0,562	0,379	-0,431	NC_010430
Merostomata	Xiphosura	Limulus polyphemus	-0,526	-0,721	-0,210	-0,331	-0,333	-0,427	0,511	-0,570	-0,524	0,485	0,636	0,396	-0,629	NC_003057
		Tachypleus tridentatus	-0,389	-0,533	-0,146	-0,284	-0,183	-0,310	0,415	-0,430	-0,341	0,393	0,662	0,345	-0,612	NC_012574
Pycnogonida	Pantopoda	Achelia bituberculata	-0,015	-0,167	0,094	0,029	0,146	-0,136	0,230	0,108	0,205	0,055	0,294	0,015	-0,087	NC_009724
		Ammothera carolinensis	-0,257	-0,478	-0,028	-0,273	-0,106	-0,188	0,410	-0,458	-0,370	0,221	0,463	0,243	-0,472	NC_014671
		Nymphon gracile	-0,328	-0,889	0,158	0,077	-0,108	-0,149	0,348	0,242	-0,352	0,288	0,545	0,249	-0,523	NC_008572
		Tanystylum orbiculare	0,000	-0,185	0,122	0,107	0,127	0,083	0,144	0,080	0,123	-0,068	0,240	-0,062	0,059	NC_014505
Crustacea	Branchiopoda	Artemia franciscana	-0,127	-0,200	0,000	-0,054	0,096	-0,075	-0,038	-0,078	-0,009	-0,172	-0,051	0,000	-0,065	NC_001620
	Diplostraca	Daphnia pulex	-0,165	-0,241	-0,006	-0,180	-0,125	-0,151	0,188	-0,043	-0,102	0,087	0,146	0,113	-0,196	NC_000844
	Notostraca	Triops cancrivormis	-0,307	-0,438	-0,040	-0,143	-0,120	-0,149	0,234	-0,265	-0,019	0,126	0,250	0,112	-0,156	NC_004465
		Triops longicaudatus	-0,181	-0,297	0,004	-0,100	-0,103	-0,127	0,288	-0,273	-0,020	0,123	0,295	0,134	-0,168	NC_006079
	Malacostraca	Farfantepenaeus californiensis	-0,233	-0,404	-0,068	-0,064	-0,140	-0,227	0,205	-0,263	-0,097	0,249	0,356	0,181	-0,481	NC_012738
		Fenneropenaeus chinensis	-0,255	-0,385	-0,057	-0,136	-0,148	-0,187	0,234	-0,356	-0,113	0,280	0,357	0,214	-0,453	NC_009679
		Litopenaeus stylirostris	-0,175	-0,350	-0,084	-0,104	-0,111	-0,191	0,227	-0,342	-0,124	0,275	0,376	0,226	-0,410	NC_012060
		Litopenaeus vannamei	-0,212	-0,526	-0,071	-0,081	-0,098	-0,187	0,236	-0,322	-0,155	0,227	0,356	0,222	-0,408	NC_009626
		Marsupenaeus japonicus	-0,249	-0,455	-0,039	-0,073	-0,166	-0,239	0,225	-0,341	-0,242	0,314	0,304	0,264	-0,453	NC_007010
		Penaeus monodon	-0,163	-0,353	-0,042	-0,005	-0,123	-0,107	0,047	-0,264	-0,115	0,221	0,235	0,139	-0,374	NC_002184
		Panulirus japonicus	-0,289	-0,122	-0,060	-0,124	-0,111	-0,148	0,240	-0,233	-0,152	0,223	0,316	0,131	-0,407	NC_004251
		Panulirus ornatus	-0,292	-0,277	-0,100	-0,126	-0,070	-0,184	0,266	-0,296	-0,250	0,276	0,402	0,196	-0,252	NC_014854
		Panulirus stimpsoni	-0,355	-0,364	-0,124	-0,149	-0,135	-0,193	0,331	-0,266	-0,210	0,246	0,479	0,208	-0,375	NC_014339
		Shinkaia crosnieri	-0,310	-0,613	-0,155	-0,210	-0,187	-0,268	0,379	-0,377	-0,340	0,447	0,543	0,383	-0,475	NC_011013
		Pagurus longicarpus	-0,299	-0,231	-0,065	-0,126	-0,103	-0,204	0,279	-0,260	-0,067	0,314	0,444	0,250	-0,449	NC_003058
		Homarus americanus	-0,163	-0,487	-0,102	-0,117	-0,076	-0,134	0,258	-0,351	-0,132	0,312	0,195	0,233	-0,338	NC_015607
		Cherax destructor	-0,223	-0,520	-0,131	-0,234	-0,205	-0,286	0,321	-0,326	-0,189	0,374	0,396	0,322	-0,379	NC_011243
		Gandalfus yunohana	-0,264	-0,568	-0,156	-0,239	-0,138	-0,193	0,347	-0,455	-0,224	0,347	0,483	0,367	-0,373	NC_013713
		Eriocheir hepuensis	-0,221	-0,438	-0,101	-0,145	-0,075	-0,178	0,279	-0,394	-0,250	0,389	0,443	0,360	-0,358	NC_011598
		Eriocheir japonica	-0,223	-0,455	-0,096	-0,148	-0,110	-0,197	0,311	-0,394	-0,228	0,374	0,514	0,337	-0,379	NC_011597
		Eriocheir sinensis	-0,236	-0,471	-0,109	-0,162	-0,086	-0,198	0,307	-0,432	-0,189	0,374	0,487	0,387	-0,376	NC_006992
		Xenograpsus testudinatus	-0,307	-0,385	-0,115	-0,171	-0,144	-0,173	0,336	-0,422	-0,281	0,424	0,616	0,363	-0,466	NC_013480
		Callinectes sapidus	-0,252	-0,478	-0,170	-0,210	-0,204	-0,217	0,338	-0,437	-0,200	0,389	0,422	0,306	-0,485	NC_006281
		Charybdis japonica	-0,266	-0,400	-0,102	-0,119	-0,124	-0,188	0,260	-0,421	-0,103	0,286	0,370	0,301	-0,373	NC_013246
		Portunus trituberculatus	-0,267	-0,444	-0,133	-0,127	-0,154	-0,152	0,317	-0,369	-0,252	0,313	0,413	0,303	-0,437	NC_005037
		Scylla olivacea	-0,264	-0,535	-0,155	-0,203	-0,170	-0,243	0,310	-0,396	-0,211	0,347	0,576	0,320	-0,464	NC_012569
		Scylla paramamosain	-0,249	-0,529	-0,122	-0,160	-0,107	-0,193	0,344	-0,415	-0,167	0,363	0,463	0,312	-0,398	NC_012572
		Scylla serrata	-0,222	-0,619	-0,095	-0,115	-0,110	-0,157	0,377	-0,442	-0,248	0,315	0,532	0,322	-0,366	NC_012565
		Scylla tranquebarica	-0,250	-0,625	-0,071	-0,168	-0,115	-0,177	0,360	-0,439	-0,222	0,384	0,507	0,341	-0,423	NC_012567



Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso	
		<i>Geothelphusa dehaani</i>	-0,270	-0,737	-0,128	-0,170	-0,149	-0,207	0,375	-0,485	-0,313	0,426	0,570	0,372	-0,421	NC_007379	
		<i>Pseudocarcinus gigas</i>	-0,257	-0,600	-0,088	-0,231	-0,114	-0,193	0,353	-0,427	-0,176	0,354	0,535	0,321	-0,507	NC_006891	
		<i>Alpheus distinguendus</i>	-0,376	-0,443	-0,159	-0,195	-0,248	-0,306	0,275	-0,366	-0,318	0,285	0,415	0,274	-0,455	NC_014883	
		<i>Halocaridina rubra</i>	-0,260	-0,517	-0,228	-0,251	-0,138	-0,324	0,353	-0,429	-0,191	0,369	0,366	0,348	-0,435	NC_008413	
		<i>Exopalaemon carinicauda</i>	-0,336	-0,544	-0,160	-0,253	-0,193	-0,253	0,273	-0,385	-0,265	0,319	0,481	0,316	-0,444	NC_012566	
		<i>Macrobrachium lanchesteri</i>	-0,273	-0,489	-0,155	-0,230	-0,239	-0,206	0,293	-0,371	-0,285	0,315	0,407	0,273	-0,463	NC_012217	
		<i>Macrobrachium nipponense</i>	-0,278	-0,469	-0,145	-0,214	-0,232	-0,264	0,296	-0,341	-0,302	0,365	0,367	0,256	-0,497	NC_015073	
		<i>Macrobrachium rosenbergii</i>	-0,318	-0,551	-0,208	-0,262	-0,260	-0,272	0,304	-0,368	-0,293	0,360	0,316	0,269	-0,533	NC_006880	
	Stomatopoda	<i>Gonodactylus chiragra</i>	-0,294	-0,415	-0,057	-0,104	-0,138	-0,190	0,342	-0,348	-0,140	0,377	0,260	0,304	-0,338	NC_007442	
		<i>Harpiosquilla harpax</i>	-0,158	-0,436	0,022	-0,009	-0,025	-0,038	0,151	-0,178	-0,074	0,246	0,277	0,200	-0,207	NC_006916	
		<i>Oratosquilla oratoria</i>	-0,238	-0,389	-0,004	-0,048	-0,072	-0,065	0,224	-0,255	0,009	0,318	0,238	0,231	-0,229	NC_014342	
		<i>Squilla empusa</i>	-0,229	-0,463	-0,010	-0,035	-0,107	-0,148	0,252	-0,266	-0,103	0,316	0,154	0,268	-0,307	NC_007444	
		<i>Squilla mantis</i>	-0,183	-0,515	0,025	0,014	-0,026	0,002	0,250	-0,220	-0,067	0,332	0,191	0,219	-0,120	NC_006081	
	Amphipoda	<i>Caprella mutica</i>	-0,299	-0,455	-0,066	-0,157	-0,088	-0,232	0,176	-0,251	-0,113	0,226	0,253	-0,166	-0,269	NC_014492	
		<i>Caprella scaura</i>	-0,162	-0,280	-0,047	-0,146	-0,168	-0,165	0,137	-0,108	-0,063	0,242	0,326	-0,089	-0,114	NC_014687	
		<i>Metacrangonyx longipes</i>	0,019	-0,217	0,078	0,044	0,076	-0,192	0,216	0,206	0,224	0,211	0,373	0,115	-0,200	NC_013032	
		<i>Onisimus nanseni</i>	-0,240	-0,429	-0,051	-0,119	0,011	-0,135	0,297	-0,225	0,098	0,313	0,605	0,293	-0,203	NC_013819	
	Isopoda	<i>Ligia oceanica</i>	0,038	-0,021	0,083	0,110	0,048	-0,235	0,020	0,226	0,160	-0,235	-0,118	0,334	0,303	NC_008412	
		<i>Eophreatoicus sp. 14 FK2009</i>	0,281	0,238	-0,091	0,183	0,283	-0,295	-0,176	-0,178	0,440	-0,246	-0,205	0,387	0,594	NC_013976	
Maxillopoda	Cirripedia	<i>Pollicipes polymerus</i>	-0,196	-0,319	-0,048	-0,155	-0,144	-0,101	0,183	-0,111	-0,172	0,089	0,242	0,164	-0,275	NC_005936	
		<i>Capitulum mitella</i>	-0,385	-0,520	-0,156	-0,206	-0,262	-0,184	0,333	-0,321	-0,266	0,358	0,302	0,350	-0,302	NC_008742	
		<i>Megabalanus volcano</i>	-0,253	-0,409	-0,051	-0,169	-0,141	-0,177	0,221	-0,137	-0,185	-0,012	-0,022	-0,038	-0,510	NC_006293	
		<i>Tetraclita japonica</i>	-0,370	-0,560	-0,133	-0,202	-0,237	-0,257	0,325	-0,269	-0,292	0,290	0,277	0,301	-0,458	NC_008974	
	Copepoda	<i>Tigriopus californicus</i>	0,367	0,300	0,157	0,222	0,229	0,152	0,306	0,297	0,493	0,320	0,540	0,368	0,396	NC_008831	
		<i>Tigriopus japonicus</i>	0,286	0,000	0,201	0,288	0,228	0,222	0,330	0,394	0,475	0,342	0,284	0,242	0,311	NC_003979	
		<i>Armillifer armillatus</i>	-0,828	-0,800	-0,446	-0,573	-0,624	-0,677	0,747	-0,743	-0,726	0,748	0,863	0,703	-0,820	NC_005934	
Ostracoda	Myodocopa	<i>Vargula hilgendorffii</i>	-0,313	-0,447	-0,206	-0,282	-0,224	-0,305	0,529	-0,513	0,667	0,551	0,638	0,522	-0,459	NC_005306	
Remipedia	Nectiopoda	<i>Speleonectes tulumensis</i>	-0,405	-0,073	-0,220	-0,356	-0,035	-0,358	0,329	-0,384	-0,258	0,349	0,429	0,308	-0,315	NC_005938	
Hexapoda	Entognatha	Collembola	<i>Orchesella villosa</i>	-0,238	-0,282	-0,015	-0,198	-0,055	-0,106	0,052	0,004	-0,133	0,273	0,127	0,154	-0,149	NC_010534
		<i>Cryptopygus antarcticus</i>	-0,182	-0,488	-0,054	-0,120	-0,098	-0,070	0,267	-0,201	-0,216	0,197	0,155	0,145	-0,107	NC_010533	
		<i>Gomphiocephalus hodgsoni</i>	-0,200	-0,316	-0,076	-0,182	-0,004	-0,056	0,276	-0,339	-0,106	0,245	0,377	0,198	-0,154	NC_005438	
		<i>Friesea grisea</i>	-0,362	-0,722	-0,143	-0,295	-0,180	-0,179	0,224	-0,331	-0,232	0,165	0,324	0,183	-0,194	NC_010535	
		<i>Bilobella aurantiaca</i>	-0,313	-0,556	-0,167	-0,340	-0,156	-0,224	0,477	-0,323	-0,242	0,398	0,652	0,293	-0,333	NC_011195	
		<i>Onychiurus orientalis</i>	-0,292	-0,489	-0,104	-0,173	-0,168	-0,160	0,270	-0,318	-0,298	0,245	0,275	0,195	-0,333	NC_006074	
		<i>Tetradontophora bielensis</i>	-0,293	-0,600	-0,148	-0,285	-0,216	-0,215	0,322	-0,422	-0,341	0,371	0,569	0,330	-0,436	NC_002735	
		<i>Podura aquatica</i>	-0,231	-0,660	-0,151	-0,232	-0,143	-0,212	0,369	-0,208	-0,308	0,175	0,364	0,206	-0,119	NC_006075	
		<i>Sminthurus viridis</i>	-0,157	-0,077	-0,021	-0,071	0,033	-0,073	0,034	-0,029	-0,129	-0,023	0,116	-0,086	-0,051	NC_010536	
	Diplura	<i>Campodea lubbocki</i>	-0,312	-0,514	-0,121	-0,234	-0,176	-0,236	0,353	-0,360	-0,365	0,425	0,848	0,415	-0,558	NC_008234	
Insecta	Archaeognatha	<i>Pedetontus silvestrii</i>	-0,310	-0,680	-0,059	-0,135	-0,155	-0,214	0,276	-0,275	-0,172	0,233	0,500	0,240	-0,223	NC_011717	
		<i>Petrobius brevistylis</i>	-0,267	-0,400	-0,042	-0,087	-0,201	-0,212	0,203	-0,285	-0,169	0,286	0,461	0,240	-0,357	NC_007688	

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		<i>Trigoniophthalmus alternatus</i>	-0,284	-0,568	-0,082	-0,243	-0,162	-0,301	0,324	-0,313	-0,299	0,388	0,563	0,360	-0,484	NC_010532
	Coleoptera	<i>Trachypachus holmbergi</i>	-0,195	-0,375	0,052	-0,012	-0,014	-0,061	0,256	-0,219	-0,042	0,217	0,200	0,216	-0,250	NC_011329
		<i>Aspidytes niobe</i>	-0,218	-0,250	0,042	-0,031	-0,059	-0,059	0,252	-0,187	-0,135	0,254	0,381	0,234	-0,286	NC_012139
		<i>Macrogyrus oblongus</i>	-0,260	-0,545	-0,063	-0,150	-0,064	-0,168	0,294	-0,313	-0,125	0,264	0,378	0,241	-0,372	NC_013249
		<i>Tetraphalerus bruchi</i>	-0,417	-0,463	-0,174	-0,286	-0,175	-0,275	0,263	-0,396	-0,418	0,166	0,341	0,221	-0,244	NC_011328
		<i>Hydroscapha granulum</i>	-0,115	-0,429	0,002	-0,022	-0,042	-0,128	0,312	-0,129	-0,035	0,297	0,172	0,236	-0,229	NC_012144
		<i>Sphaerius</i> sp. BT0074	-0,050	-0,579	0,046	-0,090	-0,052	-0,040	0,250	-0,216	-0,099	0,278	0,318	0,236	-0,161	NC_011322
		<i>Apatides fortis</i>	-0,397	-0,478	-0,218	-0,314	-0,216	-0,311	0,340	-0,449	-0,276	0,258	0,211	0,250	-0,458	NC_013582
		<i>Psacotheta hilaris</i>	-0,212	-0,333	-0,031	-0,141	-0,044	-0,148	0,339	-0,250	-0,184	0,248	0,200	0,270	-0,184	NC_013070
		<i>Anoplophora glabripennis</i>	-0,179	-0,600	-0,002	-0,167	-0,040	-0,151	0,336	-0,288	-0,156	0,244	0,217	0,238	-0,307	NC_008221
		<i>Crioceris duodecimpunctata</i>	-0,283	-0,333	-0,020	-0,128	-0,014	-0,122	0,288	-0,119	-0,268	0,243	0,429	0,169	-0,303	NC_003372
		<i>Chaetosoma scaritides</i>	-0,148	-0,500	-0,018	-0,138	-0,073	-0,150	0,218	-0,239	-0,158	0,228	0,424	0,297	-0,366	NC_011324
		<i>Mordella atrata</i>	-0,320	-0,333	-0,099	-0,159	-0,115	-0,236	0,320	-0,361	-0,341	0,265	0,410	0,259	-0,370	NC_013254
		<i>Priasilpha obscura</i>	-0,251	-0,364	-0,075	-0,169	-0,098	-0,216	0,306	-0,364	-0,195	0,311	0,356	0,260	-0,451	NC_011326
		<i>Adelium</i> sp. NCS2009	-0,214	-0,487	-0,106	-0,198	-0,127	-0,218	0,278	-0,274	-0,308	0,203	0,364	0,238	-0,344	NC_013554
		<i>Tribolium castaneum</i>	-0,355	-0,379	-0,191	-0,328	-0,129	-0,310	0,278	-0,371	-0,390	0,355	0,397	0,372	-0,474	NC_003081
		<i>Chrysochroa fulgidissima</i>	-0,262	-0,297	-0,132	-0,198	-0,200	-0,269	0,301	-0,267	-0,296	0,283	0,400	0,234	-0,453	NC_012765
		<i>Acmaeodera</i> sp. NCS2009	-0,330	-0,476	-0,201	-0,311	-0,243	-0,258	0,249	-0,303	-0,317	0,177	0,280	0,165	-0,356	NC_013580
		<i>Chauliognathus opacus</i>	-0,166	-0,571	-0,041	-0,165	-0,055	-0,126	0,195	-0,238	-0,157	0,101	0,347	0,158	-0,207	NC_013576
		<i>Pyrophorus divergens</i>	-0,339	-0,350	-0,184	-0,253	-0,204	-0,319	0,345	-0,344	-0,386	0,325	0,385	0,311	-0,437	NC_009964
		<i>Pyrocoelia rufa</i>	-0,304	-0,400	-0,035	-0,253	-0,028	-0,214	0,293	-0,243	-0,218	0,197	0,519	0,179	-0,224	NC_003970
		<i>Rhagophthalmus lufengensis</i>	-0,206	-0,368	-0,019	-0,189	-0,064	-0,121	0,269	-0,224	-0,088	0,256	0,490	0,240	-0,231	NC_010969
		<i>Rhagophthalmus ohbai</i>	-0,252	-0,500	-0,068	-0,164	-0,078	-0,190	0,288	-0,266	-0,125	0,213	0,434	0,275	-0,325	NC_010964
		<i>Cyphon</i> sp. BT0012	-0,188	-0,543	-0,068	-0,238	-0,136	-0,191	0,328	-0,322	-0,130	0,239	0,517	0,288	-0,461	NC_011320
		<i>Lucanus mazama</i>	-0,322	-0,395	-0,122	-0,243	-0,226	-0,248	0,347	-0,368	-0,345	0,272	0,486	0,235	-0,417	NC_013578
		<i>Rhopaea magnicornis</i>	-0,228	-0,394	-0,052	-0,228	-0,069	-0,124	0,373	-0,409	-0,165	0,313	0,464	0,217	-0,451	NC_013252
	Diptera	<i>Drosophila littoralis</i>	-0,247	-0,419	-0,014	-0,052	-0,076	-0,129	0,265	-0,156	-0,184	0,271	0,286	0,285	-0,419	NC_011596
		<i>Drosophila mauritiana</i>	-0,139	-0,417	0,046	-0,032	-0,013	-0,047	0,248	-0,097	-0,194	0,264	0,277	0,242	-0,333	NC_005779
		<i>Drosophila melanogaster</i>	-0,152	-0,481	0,034	0,039	-0,039	-0,058	0,282	-0,091	-0,176	0,250	0,269	0,220	-0,351	NC_001709
		<i>Drosophila sechellia</i>	-0,190	-0,429	0,015	-0,011	0,022	-0,047	0,251	-0,158	-0,171	0,221	0,234	0,199	-0,279	NC_005780
		<i>Drosophila simulans</i>	-0,160	-0,417	0,027	-0,027	0,036	-0,031	0,244	-0,160	-0,188	0,250	0,234	0,219	-0,325	NC_005781
		<i>Drosophila yakuba</i>	-0,166	-0,429	0,028	0,050	0,013	-0,064	0,259	-0,116	-0,151	0,265	0,234	0,224	-0,300	NC_001322
		<i>Haematobia irritans irritans</i>	-0,138	-0,481	0,046	-0,006	-0,040	0,003	0,246	-0,078	-0,211	0,222	0,280	0,172	-0,244	NC_007102
		<i>Trichophthalma punctata</i>	-0,296	-0,455	-0,114	-0,175	-0,169	-0,221	0,344	-0,318	-0,205	0,337	0,284	0,307	-0,383	NC_008755
		<i>Chrysomya putoria</i>	-0,164	-0,353	-0,017	-0,059	-0,078	-0,036	0,309	-0,180	-0,291	0,293	0,241	0,269	-0,340	NC_002697
		<i>Cochliomyia hominivorax</i>	-0,181	-0,444	-0,060	-0,090	-0,115	-0,098	0,304	-0,230	-0,318	0,348	0,263	0,277	-0,374	NC_002660
		<i>Lucilia sericata</i>	-0,172	-0,375	-0,011	-0,043	-0,063	-0,049	0,300	-0,155	-0,243	0,283	0,259	0,271	-0,302	NC_009733
		<i>Dermatobia hominis</i>	-0,243	-0,500	-0,053	-0,137	-0,118	-0,132	0,336	-0,214	-0,316	0,320	0,410	0,321	-0,429	NC_006378
		<i>Hypoderma lineatum</i>	-0,234	-0,520	-0,068	-0,173	-0,155	-0,204	0,321	-0,192	-0,378	0,331	0,467	0,296	-0,370	NC_013932
		<i>Exorista sorbillans</i>	-0,227	-0,538	-0,031	-0,055	-0,039	-0,116	0,269	-0,198	-0,270	0,239	0,143	0,225	-0,247	NC_014704

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		<i>Liriomyza sativae</i>	-0,198	-0,515	-0,013	-0,082	-0,068	-0,096	0,302	-0,102	-0,139	0,282	0,333	0,291	-0,363	NC_015926
		<i>Liriomyza trifolii</i>	-0,193	-0,533	-0,018	-0,147	-0,106	-0,117	0,278	-0,164	-0,108	0,241	0,347	0,316	-0,317	NC_014283
		<i>Simosyrphus grandicornis</i>	-0,118	-0,462	0,036	-0,061	0,000	-0,047	0,204	-0,241	-0,176	0,202	0,360	0,214	-0,216	NC_008754
		<i>Bactrocera carambolae</i>	-0,268	-0,268	-0,085	-0,125	-0,116	-0,198	0,294	-0,293	-0,313	0,302	0,333	0,297	-0,450	NC_009772
		<i>Bactrocera dorsalis</i>	-0,266	-0,250	-0,087	-0,128	-0,125	-0,209	0,309	-0,308	-0,297	0,310	0,353	0,308	-0,455	NC_008748
		<i>Bactrocera papayae</i>	-0,286	-0,250	-0,086	-0,129	-0,109	-0,218	0,317	-0,299	-0,292	0,297	0,353	0,300	-0,431	NC_009770
		<i>Bactrocera philippinensis</i>	-0,276	-0,250	-0,088	-0,129	-0,124	-0,202	0,309	-0,285	-0,305	0,287	0,324	0,289	-0,455	NC_009771
		<i>Bactrocera tryoni</i>	-0,289	-0,289	-0,081	-0,157	-0,142	-0,243	0,317	-0,290	-0,212	0,278	0,344	0,310	-0,406	NC_014611
		<i>Bactrocera oleae</i>	-0,342	-0,488	-0,128	-0,241	-0,179	-0,203	0,370	-0,350	-0,292	0,362	0,391	0,394	-0,492	NC_005333
		<i>Bactrocera minax</i>	-0,341	-0,643	-0,214	-0,287	-0,164	-0,295	0,318	-0,300	-0,382	0,365	0,429	0,391	-0,557	NC_014402
		<i>Ceratitidis capitata</i>	-0,157	-0,500	-0,023	-0,073	-0,098	-0,109	0,270	-0,229	-0,139	0,319	0,321	0,287	-0,326	NC_000857
		<i>Cydistomyia duplonotata</i>	-0,151	-0,636	-0,004	-0,065	-0,088	-0,068	0,258	-0,232	-0,151	0,354	0,311	0,200	-0,326	NC_008756
		<i>Culicoides arakawae</i>	-0,269	-0,500	-0,087	-0,154	-0,172	-0,178	0,274	-0,245	-0,176	0,288	0,390	0,180	-0,360	NC_009809
		<i>Anopheles darlingi</i>	-0,159	-0,357	0,011	-0,029	-0,044	-0,024	0,286	-0,109	-0,111	0,222	0,259	0,211	-0,200	NC_014275
		<i>Anopheles gambiae</i>	-0,149	-0,379	-0,017	-0,087	-0,047	-0,054	0,285	-0,126	-0,123	0,185	0,308	0,256	-0,190	NC_002084
		<i>Anopheles quadrimaculatus A</i>	-0,200	-0,571	-0,026	-0,096	-0,074	-0,048	0,322	-0,161	-0,143	0,262	0,308	0,275	-0,291	NC_000875
		<i>Aedes aegypti</i>	-0,196	-0,394	-0,052	-0,193	-0,093	-0,098	0,342	-0,189	-0,183	0,299	0,333	0,306	-0,287	NC_010241
		<i>Aedes albopictus</i>	-0,227	-0,250	-0,058	-0,146	-0,073	-0,021	0,282	-0,086	-0,138	0,274	0,385	0,242	-0,224	NC_006817
		<i>Culex pipiens pipiens</i>	-0,179	-0,200	0,007	-0,080	-0,035	-0,032	0,286	-0,122	0,030	0,299	0,346	0,263	-0,250	NC_015079
		<i>Culex quinquefasciatus</i>	-0,148	-0,200	0,002	-0,080	-0,085	-0,079	0,348	-0,118	-0,061	0,299	0,346	0,279	-0,250	NC_014574
		<i>Mayetiola destructor</i>	-0,197	-0,333	0,020	-0,119	0,037	-0,047	0,104	-0,091	-0,061	0,155	0,517	0,163	-0,116	NC_013066
		<i>Rhopalomyia pomum</i>	-0,304	-0,273	-0,003	-0,048	0,039	-0,009	0,152	-0,176	0,000	0,158	0,200	0,236	-0,081	NC_013063
	Hymenoptera	<i>Apis mellifera ligustica</i>	-0,269	-0,412	-0,093	-0,167	-0,188	-0,148	0,359	-0,304	-0,375	0,337	0,676	0,297	-0,545	NC_001566
		<i>Bombus hypocrita sapporoensis</i>	-0,353	0,091	-0,117	-0,145	-0,309	-0,253	0,404	-0,541	-0,476	0,457	0,622	0,359	-0,571	NC_011923
		<i>Melipona bicolor</i>	-0,370	-0,600	-0,009	-0,130	-0,176	-0,180	0,397	-0,398	-0,333	0,357	0,692	0,255	-0,579	NC_004529
		<i>Pristomyrmex punctatus</i>	-0,439	-0,600	-0,181	-0,374	-0,225	-0,372	0,423	-0,625	-0,500	0,504	0,674	0,420	-0,511	NC_015075
		<i>Solenopsis geminata</i>	-0,494	-0,667	-0,233	-0,370	-0,346	-0,400	0,451	-0,700	-0,467	0,492	0,750	0,408	-0,709	NC_014669
		<i>Solenopsis invicta</i>	-0,478	-0,652	-0,230	-0,398	-0,348	-0,391	0,483	-0,705	-0,443	0,503	0,885	0,417	-0,698	NC_014672
		<i>Solenopsis richteri</i>	-0,464	-0,644	-0,209	-0,400	-0,338	-0,390	0,462	-0,697	-0,506	0,473	0,821	0,430	-0,686	NC_014677
		<i>Radoszkowskii oculata</i>	-0,513	-0,579	-0,134	-0,186	-0,232	-0,195	0,386	-0,267	-0,434	0,464	0,778	0,436	-0,325	NC_014485
		<i>Abispa ephippium</i>	-0,496	-0,652	-0,179	-0,302	-0,202	-0,279	0,524	-0,447	-0,412	0,481	0,708	0,432	-0,651	NC_011520
		<i>Evania appendigaster</i>	-0,523	-0,440	-0,129	-0,284	-0,195	-0,299	0,477	-0,444	-0,319	0,448	0,719	0,427	-0,565	NC_013238
		<i>Spathius agrili</i>	0,105	0,250	0,212	0,248	0,263	0,200	0,000	0,339	0,345	-0,182	0,111	-0,188	0,375	NC_014278
		<i>Cotesia vestalis</i>	0,033	-0,167	0,199	0,186	0,101	-0,025	-0,065	0,091	0,200	0,170	0,091	0,080	-0,048	NC_014272
		<i>Vanhornia eucnemidarum</i>	-0,457	-0,571	-0,225	-0,303	-0,366	-0,292	0,337	-0,473	-0,259	0,402	0,500	0,417	-0,448	NC_008323
		<i>Cephus cinctus</i>	-0,255	-0,613	-0,084	-0,244	-0,143	-0,240	0,302	-0,490	-0,169	0,331	0,409	0,386	-0,310	NC_012688
		<i>Orussus occidentalis</i>	-0,437	-0,643	-0,133	-0,268	-0,161	-0,280	0,347	-0,467	-0,281	0,395	0,654	0,358	-0,536	NC_012689
	Lepidoptera	<i>Bombyx mandarina</i>	-0,257	-0,375	-0,045	-0,110	-0,095	-0,117	0,317	-0,223	-0,267	0,293	0,532	0,340	-0,183	NC_003395
		<i>Bombyx mori</i>	-0,255	-0,500	-0,036	-0,148	-0,055	-0,168	0,306	-0,254	-0,220	0,289	0,583	0,327	-0,077	NC_002355
		<i>Antheraea pernyi</i>	-0,241	-0,429	-0,077	-0,130	-0,031	-0,152	0,330	-0,260	-0,156	0,294	0,442	0,296	-0,377	NC_004622

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		Antheraea yamamai	-0,257	-0,429	-0,062	-0,143	-0,049	-0,163	0,373	-0,260	-0,172	0,328	0,429	0,273	-0,351	NC_012739
		Eriogyna pyretorum	-0,232	-0,400	-0,047	-0,139	-0,019	-0,147	0,307	-0,249	-0,241	0,291	0,455	0,260	-0,200	NC_012727
		Saturnia boisduvalii	-0,244	-0,600	-0,036	-0,092	-0,057	-0,199	0,320	-0,239	-0,207	0,333	0,476	0,310	-0,171	NC_010613
		Manduca sexta	-0,163	-0,647	-0,010	-0,067	0,000	-0,120	0,282	-0,219	-0,179	0,303	0,535	0,308	-0,061	NC_010266
		Phthonandria atrilineata	-0,220	-0,538	0,017	-0,138	-0,033	-0,154	0,295	-0,226	-0,254	0,285	0,422	0,315	-0,121	NC_010522
		Hyphantria cunea	-0,262	-0,625	-0,043	-0,108	-0,069	-0,178	0,360	-0,250	-0,353	0,328	0,467	0,345	-0,260	NC_014058
		Lymantria dispar	-0,305	-0,455	-0,076	-0,158	-0,090	-0,228	0,382	-0,278	-0,188	0,315	0,519	0,359	-0,173	NC_012893
		Sesamia inferens	-0,247	-0,429	-0,028	-0,110	-0,108	-0,179	0,379	-0,183	-0,238	0,336	0,585	0,318	-0,373	NC_015835
		Helicoverpa armigera	-0,211	-0,538	-0,026	-0,130	-0,037	-0,119	0,343	-0,203	-0,262	0,292	0,511	0,270	-0,175	NC_014668
		Ochrogaster lunifer	-0,345	-0,667	-0,108	-0,243	-0,145	-0,287	0,395	-0,382	-0,280	0,401	0,656	0,405	-0,356	NC_011128
		Coreana raphaelis	-0,184	-0,467	0,015	-0,057	-0,049	-0,073	0,322	-0,143	-0,123	0,237	0,297	0,270	-0,186	NC_007976
		Apatura metis	-0,274	-0,385	-0,031	-0,134	-0,057	-0,170	0,316	-0,250	-0,235	0,326	0,565	0,318	-0,114	NC_015537
		Sasakia charonda	-0,225	-0,500	-0,041	-0,195	-0,063	-0,167	0,337	-0,247	-0,239	0,301	0,422	0,343	-0,273	NC_014224
		Sasakia charonda kuriyamaensis	-0,225	-0,500	-0,041	-0,195	-0,063	-0,096	0,333	-0,234	-0,239	0,285	0,422	0,337	-0,333	NC_014223
		Calinaga davidis	-0,197	-0,600	-0,068	-0,103	-0,043	-0,191	0,306	-0,280	-0,324	0,197	0,458	0,280	-0,159	NC_015480
		Acraea issoria	-0,293	-0,167	-0,091	-0,143	-0,158	-0,189	0,362	-0,262	-0,360	0,305	0,362	0,279	-0,205	NC_013604
		Hipparchia autonoe	-0,301	-0,636	-0,055	-0,186	-0,091	-0,210	0,387	-0,291	-0,375	0,299	0,481	0,324	-0,263	NC_014587
		Papilio maraho	-0,351	-0,500	-0,108	-0,205	-0,159	-0,234	0,355	-0,193	-0,324	0,299	0,551	0,325	-0,310	NC_014055
		Teinopalpus aureus	-0,246	-0,467	-0,082	-0,083	-0,094	-0,191	0,363	-0,290	-0,389	0,281	0,418	0,321	-0,309	NC_014398
		Parnassius bremeri	-0,212	-0,600	-0,014	-0,108	-0,045	-0,120	0,306	-0,184	-0,254	0,267	0,476	0,256	-0,227	NC_014053
		Pieris melete	-0,246	-0,529	-0,034	-0,086	-0,100	-0,229	0,333	-0,221	-0,189	0,309	0,455	0,314	-0,294	NC_010568
		Pieris rapae	-0,241	-0,538	-0,039	-0,130	-0,126	-0,197	0,396	-0,208	-0,246	0,336	0,511	0,337	-0,244	NC_015895
		Chilo suppressalis	-0,283	-0,571	-0,046	-0,138	-0,117	-0,160	0,361	-0,239	-0,333	0,344	0,480	0,339	-0,314	NC_015612
		Diatraea saccharalis	-0,273	-0,467	-0,087	-0,237	-0,132	-0,221	0,299	-0,232	-0,395	0,371	0,593	0,358	-0,325	NC_013274
		Ostrinia furnacalis	-0,169	-0,455	-0,041	-0,090	-0,034	-0,139	0,307	-0,208	-0,267	0,313	0,488	0,284	-0,200	NC_003368
		Ostrinia nubilalis	-0,139	-0,333	-0,034	-0,121	-0,039	-0,140	0,317	-0,183	-0,258	0,310	0,478	0,269	-0,205	NC_003367
		Spilonota lechriaspis	-0,132	-0,200	-0,037	-0,069	-0,074	-0,083	0,310	-0,228	-0,292	0,331	0,409	0,274	-0,237	NC_014294
		Grapholita molesta	-0,101	-0,444	-0,039	-0,069	-0,062	-0,130	0,224	-0,160	-0,242	0,298	0,345	0,282	0,000	NC_014806
		Adoxophyes honmai	-0,253	-0,412	-0,052	-0,106	-0,111	-0,089	0,327	-0,168	-0,343	0,311	0,302	0,263	-0,120	NC_008141
	Mecoptera	Bittacus pilicornis	-0,232	-0,659	-0,054	-0,100	-0,052	-0,134	0,333	-0,367	-0,109	0,341	0,322	0,343	-0,170	NC_015118
		Boreus elegans	-0,205	-0,583	-0,028	-0,174	-0,110	-0,155	0,239	-0,311	-0,242	0,281	0,400	0,240	-0,299	NC_015119
		Neopanorpa pulchra	-0,171	-0,407	-0,022	-0,011	-0,042	-0,107	0,206	-0,169	-0,157	0,284	0,345	0,298	-0,205	NC_013180
	Megaloptera	Corydalus cornutus	-0,323	-0,586	-0,086	-0,166	-0,116	-0,194	0,366	-0,236	-0,146	0,391	0,649	0,341	-0,468	NC_011276
		Protohermes concolorus	-0,274	-0,647	-0,037	-0,168	-0,090	-0,213	0,368	-0,215	-0,213	0,331	0,491	0,360	-0,481	NC_011524
		Sialis hamata	-0,190	-0,478	0,006	-0,138	-0,080	-0,157	0,271	-0,122	-0,143	0,261	0,472	0,251	-0,256	NC_013256
	Neuroptera	Ascaloptynx appendiculatus	-0,278	-0,481	-0,045	-0,208	-0,102	-0,210	0,314	-0,250	-0,231	0,274	0,377	0,223	-0,293	NC_011277
		Libelloides macaronius	-0,236	-0,212	-0,067	-0,162	-0,117	-0,144	0,251	-0,259	-0,169	0,218	0,356	0,210	-0,276	NC_015609
		Apochrysa matsumurae	-0,089	-0,440	0,009	-0,101	-0,056	-0,072	0,301	-0,084	-0,167	0,294	0,333	0,224	-0,184	NC_015095
		Chrysoperla nipponensis	-0,058	-0,241	0,025	-0,043	-0,045	-0,038	0,248	-0,134	0,015	0,261	0,388	0,223	-0,028	NC_015093
		Ditaxis biseriata	-0,195	-0,440	0,009	-0,131	-0,074	-0,069	0,282	-0,179	-0,280	0,227	0,388	0,236	-0,294	NC_013257

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		<i>Polystoechotes punctatus</i>	-0,097	-0,407	0,013	-0,083	-0,036	-0,081	0,241	-0,143	-0,227	0,262	0,265	0,253	-0,195	NC_011278
	Raphidioptera	<i>Mongoloraphidia harmandi</i>	-0,230	-0,520	-0,094	-0,110	-0,068	-0,136	0,367	-0,351	-0,314	0,339	0,532	0,308	-0,306	NC_013251
	Blattodea	<i>Blattella germanica</i>	-0,202	-0,588	0,010	-0,067	0,004	-0,096	0,236	-0,185	-0,209	0,263	0,273	0,282	-0,333	NC_012901
		<i>Periplaneta fuliginosa</i>	-0,188	-0,333	0,008	-0,054	-0,043	-0,109	0,274	-0,138	-0,172	0,214	0,346	0,265	-0,188	NC_006076
		<i>Eupolyphaga sinensis</i>	-0,229	-0,576	-0,037	-0,155	-0,120	-0,201	0,343	-0,273	-0,250	0,328	0,367	0,310	-0,250	NC_014274
	Isoptera	<i>Coptotermes formosanus</i>	-0,339	-0,660	-0,192	-0,293	-0,268	-0,333	0,349	-0,267	-0,344	0,365	0,379	0,401	-0,413	NC_015800
		<i>Reticulitermes flavipes</i>	-0,356	-0,702	-0,185	-0,260	-0,240	-0,302	0,373	-0,220	-0,333	0,303	0,318	0,380	-0,393	NC_009498
		<i>Reticulitermes hageni</i>	-0,336	-0,636	-0,167	-0,237	-0,228	-0,313	0,335	-0,231	-0,306	0,319	0,326	0,348	-0,411	NC_009501
		<i>Reticulitermes santonensis</i>	-0,356	-0,708	-0,181	-0,276	-0,234	-0,298	0,379	-0,212	-0,298	0,310	0,310	0,381	-0,389	NC_009499
		<i>Reticulitermes virginicus</i>	-0,333	-0,617	-0,191	-0,257	-0,258	-0,300	0,320	-0,219	-0,305	0,339	0,357	0,367	-0,418	NC_009500
	Mantodea	<i>Tamolanica tamolana</i>	-0,283	-0,750	-0,084	-0,098	-0,121	-0,099	0,364	-0,336	-0,234	0,299	0,308	0,299	-0,205	NC_007702
	Mantophasmatodea	<i>Sclerophasma pairesense</i>	-0,191	-0,500	-0,006	-0,063	-0,081	-0,144	0,246	-0,120	-0,182	0,200	0,188	0,154	-0,228	NC_007701
	Orthoptera	<i>Acrida cinerea</i>	-0,230	-0,429	-0,060	-0,088	-0,086	-0,064	0,163	-0,137	-0,200	0,232	0,164	0,216	-0,175	NC_014887
		<i>Acrida willemsei</i>	-0,241	-0,400	-0,070	-0,083	-0,092	-0,068	0,167	-0,127	-0,256	0,260	0,246	0,215	-0,133	NC_011303
		<i>Calliptamus italicus</i>	-0,241	-0,368	-0,061	-0,041	-0,110	-0,102	0,172	-0,249	-0,097	0,273	0,284	0,193	-0,245	NC_011305
		<i>Ognevia longipennis</i>	-0,203	-0,462	-0,013	-0,091	-0,072	-0,069	0,202	-0,216	-0,098	0,222	0,158	0,238	-0,111	NC_013701
		<i>Prumna arctica</i>	-0,259	-0,481	-0,044	-0,088	-0,061	-0,081	0,142	-0,197	-0,111	0,216	0,036	0,155	-0,109	NC_013835
		<i>Traulia szetschuanensis</i>	-0,229	-0,588	-0,034	-0,108	-0,057	-0,075	0,144	-0,207	-0,109	0,264	0,270	0,221	-0,200	NC_013826
		<i>Schistocerca gregaria gregaria</i>	-0,287	-0,450	-0,059	-0,070	-0,101	-0,118	0,176	-0,235	-0,157	0,260	0,231	0,213	-0,223	NC_013240
		<i>Arcyptera coreana</i>	-0,184	-0,448	-0,011	-0,027	-0,083	-0,085	0,126	-0,183	-0,077	0,200	0,100	0,190	-0,181	NC_013805
		<i>Chorthippus chinensis</i>	-0,190	-0,529	-0,029	-0,058	-0,054	-0,044	0,170	-0,221	-0,098	0,196	0,231	0,219	-0,234	NC_011095
		<i>Euchorthippus fusigeniculatus</i>	-0,276	-0,484	-0,038	-0,031	-0,058	-0,082	0,139	-0,172	-0,034	0,228	0,139	0,239	-0,152	NC_014449
		<i>Gomphocerippus rufus</i>	-0,224	-0,706	-0,045	-0,110	-0,080	-0,084	0,176	-0,188	-0,136	0,226	0,242	0,232	-0,161	NC_014349
		<i>Gomphocerus licenti</i>	-0,213	-0,563	-0,010	-0,026	-0,071	-0,073	0,156	-0,169	-0,075	0,224	0,268	0,205	-0,159	NC_013847
		<i>Gomphocerus sibiricus tibetanus</i>	-0,278	-0,704	-0,008	-0,056	-0,030	-0,078	0,143	-0,140	-0,068	0,220	0,257	0,220	-0,149	NC_015478
		<i>Phlaeoba albonema</i>	-0,256	-0,371	-0,048	-0,019	-0,099	-0,097	0,202	-0,194	-0,126	0,255	0,265	0,231	-0,231	NC_011827
		<i>Gastrimargus marmoratus</i>	-0,282	-0,714	-0,089	-0,188	-0,167	-0,227	0,246	-0,304	-0,205	0,283	0,292	0,232	-0,375	NC_011114
		<i>Locusta migratoria</i>	-0,244	-0,615	-0,055	-0,152	-0,098	-0,185	0,175	-0,259	-0,154	0,227	0,343	0,165	-0,354	NC_001712
		<i>Locusta migratoria manilensis</i>	-0,234	-0,615	-0,061	-0,150	-0,103	-0,175	0,188	-0,268	-0,136	0,234	0,314	0,151	-0,354	NC_014891
		<i>Locusta migratoria migratoria</i>	-0,200	-0,630	-0,078	-0,131	-0,094	-0,184	0,179	-0,268	-0,133	0,229	0,303	0,137	-0,373	NC_011119
		<i>Locusta migratoria tibetensis</i>	-0,222	-0,643	-0,057	-0,137	-0,112	-0,206	0,181	-0,259	-0,119	0,242	0,303	0,141	-0,347	NC_015624
		<i>Oedaleus decorus asiaticus</i>	-0,217	-0,643	-0,063	-0,147	-0,145	-0,186	0,198	-0,228	-0,233	0,239	0,188	0,196	-0,380	NC_011115
		<i>Oxya chinensis</i>	-0,161	-0,655	-0,021	-0,061	-0,061	-0,067	0,164	-0,127	-0,060	0,216	0,123	0,215	-0,133	NC_010219
		<i>Xyleus modestus</i>	-0,344	-0,429	-0,073	-0,091	-0,107	-0,176	0,294	-0,215	-0,180	0,280	0,284	0,261	-0,185	NC_014490
		<i>Thrinchus schrenkii</i>	-0,310	-0,463	-0,094	-0,196	-0,162	-0,161	0,242	-0,233	-0,204	0,296	0,159	0,262	-0,228	NC_014610
		<i>Physemacris variolosa</i>	-0,326	-0,590	-0,169	-0,260	-0,224	-0,231	0,264	-0,313	-0,248	0,201	0,344	0,188	-0,406	NC_014491
		<i>Atractomorpha sinensis</i>	-0,217	-0,688	-0,047	-0,078	-0,145	-0,136	0,233	-0,136	-0,217	0,297	0,206	0,218	-0,165	NC_011824
		<i>Mekongiana xiangchengensis</i>	-0,293	-0,538	-0,062	-0,170	-0,181	-0,098	0,230	-0,128	-0,186	0,274	0,284	0,244	-0,255	NC_014450
		<i>Mekongiella xizangensis</i>	-0,263	-0,600	-0,098	-0,198	-0,192	-0,225	0,203	-0,137	-0,229	0,267	0,343	0,258	-0,392	NC_014451
		<i>Ellipes minuta</i>	-0,328	-0,492	-0,170	-0,260	-0,248	-0,319	0,346	-0,412	-0,339	0,247	0,425	0,264	-0,456	NC_014488

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		<i>Teleogryllus emma</i>	-0,297	-0,556	-0,080	-0,183	-0,098	-0,227	0,344	-0,360	0,023	0,380	0,532	0,317	-0,636	NC_011823
		<i>Myrmecophilus manni</i>	-0,363	-0,636	-0,181	-0,303	-0,193	-0,279	0,319	-0,390	-0,339	0,333	0,429	0,320	-0,542	NC_011301
		<i>Gryllotalpa orientalis</i>	-0,350	-0,222	-0,158	-0,227	-0,216	-0,312	0,394	-0,350	-0,277	0,396	0,380	0,397	-0,583	NC_006678
		<i>Gryllotalpa pluvialis</i>	-0,353	-0,529	-0,112	-0,254	-0,206	-0,281	0,410	-0,445	-0,287	0,446	0,463	0,395	-0,482	NC_011302
		<i>Troglophilus neglectus</i>	-0,259	-0,487	-0,085	-0,181	-0,155	-0,227	0,248	-0,306	-0,143	0,284	0,406	0,290	-0,470	NC_011306
		<i>Deracantha onos</i>	-0,361	-0,610	-0,139	-0,204	-0,226	-0,255	0,305	-0,346	-0,322	0,354	0,342	0,297	-0,486	NC_011813
		<i>Ruspolia dubia</i>	-0,251	-0,633	-0,111	-0,202	-0,153	-0,146	0,274	-0,297	-0,176	0,310	0,408	0,283	-0,433	NC_009876
		<i>Elimaea cheni</i>	-0,240	-0,590	-0,080	-0,176	-0,175	-0,215	0,366	-0,298	-0,286	0,354	0,371	0,338	-0,460	NC_014289
		<i>Anabrus simplex</i>	-0,403	-0,677	-0,137	-0,186	-0,197	-0,239	0,307	-0,325	-0,263	0,348	0,467	0,278	-0,458	NC_009967
		<i>Gampsocleis gratiosa</i>	-0,367	-0,621	-0,158	-0,286	-0,201	-0,322	0,333	-0,352	-0,270	0,341	0,461	0,311	-0,505	NC_011200
	Phasmatodea	<i>Micadina phluctainoides</i>	-0,175	-0,412	0,009	-0,030	-0,065	-0,083	0,214	-0,176	-0,190	0,142	0,321	0,106	-0,151	NC_014673
		<i>Phraortes illepidus</i>	-0,182	-0,217	-0,039	-0,106	-0,097	-0,159	0,270	-0,200	-0,317	0,200	0,373	0,201	-0,378	NC_014695
		<i>Phraortes sp. Iriomote Island</i>	-0,229	-0,444	-0,057	-0,146	-0,108	-0,107	0,238	-0,251	-0,315	0,204	0,345	0,235	-0,297	NC_014705
		<i>Entoria okinawaensis</i>	-0,138	-0,333	-0,078	-0,163	-0,074	-0,147	0,226	-0,231	-0,266	0,239	0,279	0,235	-0,400	NC_014694
		<i>Ramulus hainanense</i>	-0,299	-0,448	-0,087	-0,165	-0,084	-0,217	0,339	-0,231	-0,269	0,240	0,313	0,270	-0,443	NC_013185
		<i>Ramulus irregulariterdentatus</i>	-0,178	-0,273	-0,029	-0,038	0,009	-0,126	0,189	-0,198	-0,126	0,152	0,133	0,197	-0,395	NC_014702
		<i>Phobaeticus serratipes</i>	-0,242	-0,455	-0,085	-0,200	-0,104	-0,164	0,298	-0,259	-0,358	0,307	0,333	0,297	-0,448	NC_014678
		<i>Megacrania alpheus adan</i>	-0,258	-0,364	-0,085	-0,167	-0,155	-0,208	0,265	-0,240	-0,213	0,313	0,475	0,290	-0,446	NC_014688
		<i>Heteropteryx dilatata</i>	-0,210	-0,481	-0,060	-0,173	-0,129	-0,187	0,246	-0,228	-0,341	0,284	0,345	0,273	-0,457	NC_014680
	Hemiptera	<i>Philaenus spumarius</i>	-0,135	-0,524	0,070	0,000	0,000	-0,036	0,221	-0,046	0,054	0,164	0,296	0,090	-0,231	NC_005944
		<i>Abidama producta</i>	-0,189	-0,556	-0,088	-0,243	-0,191	-0,191	0,354	-0,158	-0,250	0,275	0,429	0,330	-0,422	NC_015799
		<i>Laodelphax striatellus</i>	-0,443	-0,538	-0,192	-0,202	-0,141	-0,230	0,165	-0,315	-0,155	0,254	0,349	0,211	-0,195	NC_013706
		<i>Geisha distinctissima</i>	-0,353	-0,357	-0,157	-0,227	-0,181	-0,269	0,381	-0,430	-0,184	0,315	0,467	0,280	-0,276	NC_012617
		<i>Lycorma delicatula</i>	-0,333	-0,630	-0,182	-0,322	-0,288	-0,265	0,366	-0,539	-0,341	0,299	0,483	0,300	-0,327	NC_012835
		<i>Gerris sp. NKMT033</i>	-0,228	-0,556	-0,057	-0,141	-0,065	-0,160	0,245	-0,161	-0,244	0,138	0,313	0,112	-0,287	NC_012841
		<i>Hydrometra sp. NKMT020</i>	-0,283	-0,500	-0,034	-0,188	-0,094	-0,078	0,194	-0,182	-0,015	0,197	0,236	0,091	-0,351	NC_012842
		<i>Orius niger</i>	-0,106	-0,185	-0,023	-0,121	-0,017	-0,067	0,200	-0,126	-0,100	0,175	0,226	0,027	-0,113	NC_012429
		<i>Agriosphodrus dohrni</i>	-0,092	-0,200	-0,055	-0,047	-0,039	-0,090	0,187	-0,068	-0,068	0,087	0,382	0,079	-0,162	NC_015842
		<i>Valentia hoffmanni</i>	-0,215	-0,333	-0,066	-0,159	-0,042	-0,081	0,281	-0,298	-0,238	0,338	0,417	0,237	-0,379	NC_012823
		<i>Triatoma dimidiata</i>	-0,358	-0,486	-0,119	-0,181	-0,184	-0,238	0,328	-0,255	-0,295	0,358	0,385	0,347	-0,276	NC_002609
		<i>Saldula arsenjevi</i>	-0,194	-0,231	-0,059	-0,094	-0,081	-0,135	0,200	-0,137	-0,150	0,128	0,000	0,202	-0,102	NC_012463
		<i>Nerthra sp. NKMT022</i>	-0,264	-0,486	-0,004	-0,103	-0,034	-0,206	0,291	-0,103	-0,231	0,276	0,449	0,261	-0,313	NC_012838
		<i>Ilyocoris cimicoides</i>	-0,168	0,255	-0,111	-0,169	-0,089	-0,189	0,209	-0,135	-0,071	0,112	0,183	0,228	-0,113	NC_012845
		<i>Laccotrephes robustus</i>	-0,202	-0,395	-0,088	-0,179	-0,138	-0,251	0,295	-0,187	-0,140	0,307	0,375	0,225	-0,321	NC_012817
		<i>Enithares tibialis</i>	-0,188	-0,217	-0,067	-0,164	-0,040	-0,190	0,229	-0,232	-0,200	0,232	0,134	0,220	-0,235	NC_012819
		<i>Ochterus marginatus</i>	-0,280	-0,520	-0,115	-0,209	-0,206	-0,300	0,350	-0,244	-0,233	0,298	0,378	0,295	-0,382	NC_012820
		<i>Paraplea frontalis</i>	-0,258	-0,037	-0,043	-0,120	-0,094	-0,139	0,219	-0,236	-0,096	0,190	0,303	0,227	-0,152	NC_012822
		<i>Neuroctenus parus</i>	-0,257	-0,366	-0,152	-0,283	-0,148	-0,215	0,345	-0,224	-0,083	0,271	0,265	0,240	-0,289	NC_012459
		<i>Riptortus pedestris</i>	-0,257	-0,200	-0,041	-0,081	-0,040	-0,120	0,241	-0,157	-0,117	0,234	0,400	0,211	-0,122	NC_012462
		<i>Hydaropsis longirostris</i>	-0,279	-0,333	-0,063	-0,104	-0,065	-0,226	0,369	-0,228	-0,181	0,291	0,429	0,299	-0,244	NC_012456

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		Aeschyntelus notatus	-0,183	-0,333	-0,023	-0,124	-0,022	-0,126	0,313	-0,055	-0,200	0,302	0,355	0,235	-0,208	NC_012446
		Stictopleurus subviridis	-0,202	-0,448	-0,051	-0,077	-0,038	-0,069	0,264	-0,132	-0,209	0,233	0,294	0,259	-0,184	NC_012888
		Yemmalysus parallelus	-0,096	-0,091	0,036	-0,112	0,086	-0,027	0,227	-0,022	-0,026	0,079	0,276	0,146	-0,060	NC_012464
		Phaenacantha marcida	-0,253	-0,355	-0,103	-0,144	-0,139	-0,175	0,208	-0,174	-0,306	0,165	0,304	0,235	-0,370	NC_012460
		Geocoris pallidipennis	-0,129	-0,259	-0,019	-0,075	-0,036	-0,177	0,250	-0,105	-0,221	0,151	0,311	0,198	-0,157	NC_012424
		Malcus inconspicuus	-0,274	-0,429	-0,026	-0,174	-0,096	-0,167	0,215	-0,283	-0,151	0,173	0,527	0,285	-0,167	NC_012458
		Macroscytus subaeneus	-0,097	-0,171	-0,063	-0,136	-0,033	-0,118	0,205	-0,031	-0,082	0,225	0,362	0,249	-0,240	NC_012457
		Halyomorpha halys	-0,172	-0,357	-0,040	-0,094	-0,087	-0,111	0,211	-0,098	-0,084	0,115	0,281	0,206	-0,224	NC_013272
		Nezara viridula	-0,146	-0,217	-0,028	-0,093	-0,023	-0,068	0,280	-0,069	-0,116	0,173	0,377	0,151	-0,191	NC_011755
		Coptosoma bifaria	-0,135	-0,182	-0,070	-0,104	-0,081	-0,132	0,256	-0,075	-0,070	0,025	0,086	0,104	-0,125	NC_012449
		Megacopta cribraria	-0,193	-0,444	-0,121	-0,176	-0,036	-0,227	0,205	-0,221	-0,184	0,176	0,342	0,157	-0,304	NC_015342
		Physopelta gutta	-0,291	-0,241	-0,108	-0,198	-0,098	-0,113	0,291	-0,333	-0,242	0,335	0,531	0,239	-0,169	NC_012432
		Dysdercus cingulatus	-0,236	-0,172	-0,057	-0,131	-0,119	-0,149	0,295	-0,240	-0,282	0,392	0,533	0,218	-0,098	NC_012421
		Homalodisca vitripennis	-0,117	-0,133	-0,018	-0,024	0,005	-0,151	0,253	0,010	-0,194	0,219	0,292	0,125	-0,205	NC_006899
		Aleurodicus dugesii	0,260	0,125	0,275	0,237	0,308	0,232	-0,122	0,077	0,550	-0,212	-0,310	-0,279	0,273	NC_005939
		Aleurochiton aceris	0,169	0,360	0,216	0,230	-0,019	0,172	-0,121	0,127	-0,143	-0,119	-0,077	-0,198	0,289	NC_006160
		Bemisia tabaci	-0,059	0,020	0,176	0,200	-0,099	0,081	-0,119	0,235	0,282	-0,117	0,017	-0,127	0,221	NC_006279
		Neomaskellia andropogonis	0,218	0,647	0,282	0,247	0,041	0,206	-0,240	0,286	-0,176	-0,328	-0,021	-0,279	0,214	NC_006159
		Tetraleurodes acaciae	0,184	0,200	0,201	0,271	0,220	0,141	-0,098	0,250	0,171	-0,169	-0,221	-0,190	0,184	NC_006292
		Trialeurodes vaporariorum	0,191	0,348	0,208	0,222	0,246	0,155	-0,298	0,344	0,390	-0,220	-0,227	-0,298	0,415	NC_006280
		Schizaphis graminum	-0,393	-0,826	-0,106	-0,227	-0,293	-0,134	0,269	-0,297	-0,319	0,347	0,353	0,322	-0,172	NC_006158
		Acyrtosiphon pisum	-0,436	-0,667	-0,125	-0,203	-0,286	-0,172	0,271	-0,333	-0,273	0,337	0,500	0,279	-0,224	NC_011594
	Phthiraptera	Heterodoxus macropus	0,108	0,048	0,063	0,090	0,234	0,063	0,111	-0,034	0,273	0,050	0,321	0,106	0,000	NC_002651
		Bothriometopus macronemesis	0,056	0,083	0,070	0,085	0,082	-0,031	0,149	0,085	0,041	0,000	0,014	0,004	0,033	NC_009983
		Campanulotes bidentatus compar	0,385	0,314	0,328	0,336	0,318	0,297	0,390	0,489	0,705	0,256	0,472	0,388	0,613	NC_007884
	Psocoptera	Thrips imaginis	-0,178	-0,538	-0,124	-0,176	-0,175	-0,117	-0,174	-0,347	-0,250	0,189	0,045	0,091	-0,170	NC_004371
	Plecoptera	Pteronarcys princeps	-0,214	-0,400	-0,046	-0,122	-0,153	-0,164	0,328	-0,277	-0,219	0,293	0,324	0,271	-0,328	NC_006133
	Ephemeroptera	Siphonurus immanis	-0,227	-0,522	-0,023	-0,143	-0,082	-0,174	0,356	-0,231	-0,082	0,290	0,183	0,274	-0,403	NC_013822
		Ephemeria orientalis	-0,300	-0,522	-0,072	-0,170	-0,151	-0,186	0,356	-0,240	-0,245	0,347	0,324	0,322	-0,391	NC_012645
		Parafironurus youi	-0,239	-0,417	-0,084	-0,145	-0,173	-0,196	0,381	-0,249	-0,294	0,304	0,341	0,268	-0,386	NC_011359
	Odonata	Davidius lunatus	-0,227	-0,378	-0,059	-0,089	-0,113	-0,147	0,242	-0,277	-0,292	0,251	0,307	0,260	-0,315	NC_012644
		Euphaea formosa	-0,257	-0,297	-0,063	-0,048	-0,115	-0,144	0,262	-0,190	-0,200	0,271	0,395	0,254	-0,152	NC_014493
	Thysanura	Tricholepidion gertschi	-0,227	-0,610	-0,083	-0,071	-0,185	-0,237	0,226	-0,365	-0,238	0,300	0,477	0,313	-0,443	NC_005437
		Thermobia domestica	-0,400	-0,571	-0,106	-0,247	-0,259	-0,272	0,233	-0,390	-0,366	0,322	0,351	0,196	-0,452	NC_006080
		Atelura formicaria	-0,280	-0,520	-0,186	-0,255	-0,224	-0,277	0,345	-0,411	-0,343	0,365	0,415	0,322	-0,533	NC_011197
Myriapoda	Chilopoda	Scutigera coleoptrata	-0,352	-0,459	-0,174	-0,270	-0,258	-0,307	0,298	-0,386	-0,333	0,373	0,615	0,392	-0,486	NC_005870
		Lithobiomorpha														
		Bothropolys sp. SP2004	-0,377	-0,771	-0,184	-0,291	-0,233	-0,233	0,375	-0,466	-0,364	0,291	0,457	0,377	-0,504	NC_009458
		Lithobius forficatus	-0,257	-0,644	-0,140	-0,236	-0,257	-0,217	0,259	-0,331	-0,300	0,274	0,465	0,367	-0,463	NC_002629
	Diplopoda	Helminthomorpha														
		Antrokoreana gracilipes	-0,142	-0,063	-0,035	-0,047	0,042	-0,116	0,147	0,349	0,014	0,089	0,019	-0,020	-0,090	NC_010221
		Narceus annularus	-0,520	-0,661	-0,240	-0,360	-0,300	-0,394	0,511	-0,432	-0,493	0,531	0,671	0,433	-0,549	NC_003343

Classe	Ordem	Espécie	ATP6	ATP8	COX1	COX2	COX3	CYTB	ND1	ND2	ND3	ND4	ND4L	ND5	ND6	Nº de Acesso
		Thyropygus sp. DVL2001	-0,416	-0,545	-0,181	-0,264	-0,257	-0,306	0,365	-0,376	-0,484	0,335	0,386	0,247	-0,534	NC_003344
Symphyla	Scolopendromorpha	Scutigereilla causeyae	-0,278	-0,579	-0,127	-0,167	-0,146	-0,233	0,380	-0,406	-0,372	0,378	0,636	0,439	-0,411	NC_008453

Campos em cinza escuro indicam genes com viés GC% fora do padrão ou espécies com genes fora do padrão. Campos em cinza claro indicam espécies retiradas das análises para evitar excesso de um táxon.



## 10. Currículo Lattes.

### Resumos publicados em anais de congressos:

1. SANTOS, M. A. O. ; M. V. A. Batista ; Balbino, V. Q. . METHODS FOR SELECTION OF PHYLOGENETIC INFORMATIVE REGIONS IN MITOCHONDRIAL GENES. In: 7th International Conference of th Brazilian Association for Bioinformatics and Computational Biology - X-Meeting, 2011, Florianópolis. Anais X-meeting 2011, 2011.
2. G. Z. G. Silva ; S. M. O. Medeiros ; R. C. A. Gomes ; SANTOS, M. A. O. ; M. V. A. Batista ; W. Cunha ; L. Gomes ; T. L. D. Lima ; T. A. E. Ferreira ; V. Q. Balbino . SANDFLYDATABASE: DEVELOPMENT OF AN INTEGRATED PLATFORM OF BIOLOGICAL AND MOLECULAR DATA OF SANDFLIES (DIPTERA: PSYCHODIDAE) OF MEDICAL AND VETERINARY IMPORTANCE. In: IV Workshop de Genética e Biologia Molecular de Insetos Vetores de Doenças Tropicais, 2010, Recife. Anais do IV Workshop de Genética e Biologia Molecular de Insetos Vetores de Doenças Tropicais, 2010.
3. Silva, G. G. Z. ; Medeiros S M O ; Gomes, R. C. A. ; SANTOS, M. A. O. ; Batista, M. V. A. ; Gomes, L. ; Lima, T. L. D. ; Ferreira, T. A. E. ; Balbino, V. Q. . SANDFLY DATABASE: IN SILICO EXPLOITATION OF EST DATA TO EXPLORE EXPRESSION PATTERNS IN SANDFLY SPECIES. In: 6th International Conference of the Brazilian Association for Bioinformatics and Computational Biology, 2010, Ouro Preto. X-meeting Abstracts Book, 2010.

### Curso de curta duração ministrado:

1. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . III CURSO DE BIOINFORMÁTICA: ANÁLISE DE DADOS MOLECULARES (Turma I). 2012. (Curso de curta duração ministrado/Extensão).
2. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . III CURSO DE BIOINFORMÁTICA: ANÁLISE DE DADOS MOLECULARES (Turma II). 2012. (Curso de curta duração ministrado/Extensão).
3. SANTOS, M. A. O. . UTILIZAÇÃO DE FERRAMENTAS DE BIOINFORMÁTICA PARA ANÁLISE DE SEQUÊNCIAS ALVO PARA DIAGNOSE VEGETAL. 2012. .
4. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. ; Lima, T. L. D. . II CURSO DE BIOINFORMÁTICA: ANÁLISE DE DADOS MOLECULARES. 2011. (Curso de curta duração ministrado/Extensão).
5. SANTOS, M. A. O. . Bioinformática. 2011. (Curso de curta duração ministrado/Extensão).
6. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . I CURSO DE BIOINFORMÁTICA: FERRAMENTAS E APLICAÇÕES (Turma I). 2011. (Curso de curta duração ministrado/Extensão).
7. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . I CURSO DE BIOINFORMÁTICA: FERRAMENTAS E APLICAÇÕES (Turma II). 2011. (Curso de curta duração ministrado/Extensão).
8. SANTOS, M. A. O. . INTRODUÇÃO A BIOINFORMÁTICA. 2011. (Curso de curta

duração ministrado/Extensão).

9. SANTOS, M. A. O. . Bioinformática: métodos computacionais em Biologia Molecular. 2010. (Curso de curta duração ministrado/Extensão).

**Organização de eventos:**

1. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . III CURSO DE BIOINFORMÁTICA:ANÁLISE DE DADOS MOLECULARES (Turma I). 2012.
2. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . III CURSO DE BIOINFORMÁTICA:ANÁLISE DE DADOS MOLECULARES (Turma II). 2012.
3. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. ; Lima, T. L. D. . II CURSO DE BIOINFORMÁTICA: ANÁLISE DE DADOS MOLECULARES. 2011.
4. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . I CURSO DE BIOINFORMÁTICA: FERRAMENTAS E APLICAÇÕES (Turma I). 2011.
5. Balbino, V. Q. ; SANTOS, M. A. O. ; Batista, M. V. A. . I CURSO DE BIOINFORMÁTICA: FERRAMENTAS E APLICAÇÕES (Turma II). 2011.