

**DINÂMICA EVOLUTIVA DE DNAS REPETITIVOS COM
ÊNFASE EM ESPÉCIES DA TRIBO PHANAEINI**

SARAH GOMES DE OLIVEIRA

Botucatu – SP

2013

UNIVERSIDADE ESTADUAL PAULISTA

“Julio de Mesquita Filho”

INSTITUTO DE BIOCÊNCIAS DE BOTUCATU

**DINÂMICA EVOLUTIVA DE DNAS REPETITIVOS COM
ÊNFASE EM ESPÉCIES DA TRIBO PHANAEINI**

CANDIDATA: SARAH GOMES DE OLIVEIRA

ORIENTADOR: CESAR MARTINS

CO-ORIENTADORA: RITA DE CÁSSIA DE MOURA

Tese apresentada ao Instituto de
Biociências, Câmpus de Botucatu,
UNESP, para obtenção do título de
Doutora no Programa de Pós-Graduação
em Ciências Biológicas (Genética).

Botucatu – SP

2013

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO DE AQUIS. E TRAT. DA INFORMAÇÃO
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CAMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: **ROSEMEIRE APARECIDA VICENTE**

Oliveira, Sarah Gomes de.

Dinâmica evolutiva de DNAs repetitivos com ênfase em espécies da tribo Phanaeini / Sarah Gomes de Oliveira – Botucatu : [s.n], 2013

Tese (doutorado) - Universidade Estadual Paulista, Instituto de
Biociências de Botucatu

Orientador: Cesar Martins

Coorientador: Rita de Cássia de Moura

Capes: 20204000

1. DNA. 2. Cromossomos. 3. Coleóptero. 4. Evolução (Biologia).
5. Mapeamento cromossômico.

Palavras-chave: Coleoptera; DNA repetitivo; Elementos de transposição;
Evolução cromossômica; Evolução em concerto; Famílias multigênicas;
Transferência horizontal.

Dedico aos meus amados pais,
por acreditarem.

AGRADECIMENTOS

A realização desta tese marca o final de uma importante etapa da minha vida. Gostaria de agradecer a todos que contribuíram de forma decisiva para a sua concretização:

À minha amada família, que sempre me estimulou a crescer científica e pessoalmente; apoiando-me nos momentos de ansiedade, de desespero e de empolgação. Acima de tudo aos meus pais, Márcia e Manoel, pelo inestimável apoio familiar, pelo incentivo por toda a minha vida e, principalmente, durante esta trajetória na pós-graduação. À minha vizinha pelo carinho, amor e paciência revelados ao longo destes anos.

Ao meu querido irmão Mauro e sua adorável esposa, pela compreensão e ternura manifestadas apesar da falta de atenção e ausências; e pela excitação e orgulho com que sempre reagiram aos meus resultados acadêmicos ao longo dos anos.

Ao meu orientador, Prof. Dr. Cesar Martins, pela competência científica, acompanhamento do trabalho, disponibilidade e generosidade reveladas ao longo destes anos. Assim como pelas críticas, correções e sugestões relevantes durante a orientação.

À minha co-orientadora, Prof^a Dr^a Rita de Cássia de Moura, por sua colaboração, conhecimentos transmitidos e capacidade de estímulo ao longo de todo o trabalho.

À FAPESP pela bolsa concedida, que permitiu que eu me dedicasse exclusivamente aos estudos, à pesquisa e à elaboração da tese de doutorado.

Ao IBAMA pela concessão das licenças necessárias para a coleta e envio de amostras.

Ao Prof. Dr. Thomas Eickbush, Danna Eickbush e William Burke, da University of Rochester, pela confiança em meu trabalho, e por me ensinarem com prazer e dedicação parte do que sei, além da disponibilidade e amizade demonstradas.

Ao Dr. Jerzy Jurka e a todos de sua equipe do Genetic Information Research Institute, pela maneira amável, aberta e atenciosa como fui recebida, e por me mostrarem uma outra forma de olhar a ciência e as relações pessoais.

Ao Programa de Pós-Graduação em Genética e seus professores, por toda dedicação no processo constante de melhoria do Programa e por me fazer ter certeza de que fiz a escolha certa. Aos funcionários da Pós-Graduação, que com sua simpatia e atenção, sempre estiveram à disposição para qualquer dúvida.

Aos funcionários do Departamento de Morfologia, em especial Luciana, D. Terezinha, e D. Yolanda. Ao Departamento de Morfologia, ao Instituto de Biociências de Botucatu e à Universidade Estadual Paulista, pela estrutura cedida para a realização deste trabalho, pela excelência da formação prestada e conhecimentos transmitidos.

Aos meus colegas e amigos do Laboratório de Genômica Integrativa, por compartilhar o dia-a-dia: Érica Ramos, Guilherme Valente, Rafael Nakajima, Bruno Fantinatti, Juliana Giusti, Marielly Campos, Diego Marques e Marcos Dias. E aos ex-companheiros: Diogo Cabral-de-Mello, Juliana Mazzuchelli, Marcos Geraldo, Pedro Nachtigall e Danillo Pinhal. Obrigada pelas horas agradáveis de estudo, debates e cafezinhos, principalmente pela parceria nesta viagem doida que é a pós-graduação (que a gente adora!).

Aos amigos do LBGI... A nossa parceria continua, ainda que estejamos distantes.

Aos amigos Talita Sarah Mazzoni e Julio Santana, que se tornaram grandes amigos sem se darem conta. Pois entre brincadeiras e sextas-feiras fizeram o tempo passar de maneira mais divertida.

À amiga Érica Ramos pela amizade, pelos momentos de desabafo e descontração. Nos encontraremos e desencontraremos pelo caminho... Nossa despedida continua!

Às amigas Juliana Mazzucheli e Tatiane Mariguela pelos chás e por todas as “galinhas gordas”... A amizade construída nestes anos não tem preço e não terá fim.

Aos amigos que fiz no período em que morei em Rochester e Mountain View, por me ouvirem, auxiliarem, e cederem uma mão amiga nos dias em que a luz não brilhava tanto (literalmente, em muitas ocasiões).

Ao Alejandro, por todos os momentos e sonhos compartilhados.

O doutorado foi uma experiência essencial para a minha construção como pesquisadora, e agradeço a todos que fizeram parte desta jornada. A todos que me apoiaram e acreditaram em mim e no trabalho que eu estava desenvolvendo.

"O correr da vida embrulha tudo, a vida é assim: esquenta e esfria, aperta e daí afrouxa, sossega e depois desinquieta. O que ela quer da gente é coragem."

Fragmento do livro "Grande Sertão Veredas",
autoria de Guimarães Rosa.

RESUMO

O estudo de DNAs repetitivos tem se mostrado uma ferramenta esclarecedora para diversas questões, que incluem desde a organização molecular e o entendimento da estrutura cromossômica, à análises relacionadas à diversificação e evolução cariotípica. Desta forma, o objetivo deste trabalho foi analisar a organização cromossômica e genômica dos DNAs repetitivos com ênfase em representantes da tribo Phanaeini (Coleoptera: Scarabaeidae: Scarabaeinae), visando a compreensão dos mecanismos evolutivos envolvidos na dinâmica dos DNAs repetitivos no genoma e suas implicações. Em representantes da tribo Phanaeini, especialmente membros do gênero *Coprophanaeus*, foi observado que a expansão de DNA repetitivo ocorreu no início da diversificação do grupo; estando estas sequências envolvidas com a diversificação dos mecanismos sexuais de *Coprophanaeus*, e com a origem e evolução do cromossomo B observado em *C. cyanescens*. O isolamento e mapeamento de transposons *Mariner* revelou que estas sequências sofreram uma elevada diversificação durante a história evolutiva de Phanaeini, podendo também exercer alguma função na região pericentromérica dos cromossomos. Comparações entre sequências de famílias relacionadas do transposon *Mariner* mostrou que estas sequências podem estar envolvidas em um processo de transferência horizontal (HT) em diversos grupos animais não-relacionados, especialmente entre insetos e mamíferos; o que contribuiu para a ampla distribuição destes elementos transponíveis. Em relação às famílias multigênicas analisadas, observou-se uma grande variação para o DNAr 18S em contraste com o padrão conservado de DNAr 5S, sugerindo que as regiões genômicas que abrigam as classes de genes ribossomais são governadas por distintas forças evolutivas. Adicionado à isso, as análises das sequências de DNAs ribossomais 28S e retrotransposons não-LTR R2 mostraram que o auto nível de inserções de R2 não afetaram significativamente a evolução em concerto dos genes ribossomais, mas podem estar envolvidas na dispersão dos genes de rRNA para diferentes cromossomos. Desta maneira, em representantes da tribo Phanaeini, o uso do mapeamento de seqüências repetitivas se mostrou uma ferramenta útil no entendimento do cariótipo do grupo, contribuindo para o esclarecimento dos processos que governam a evolução de seus cariótipos e genomas como um todo.

Palavras-chave: Coleoptera, DNA repetitivo, elementos de transposição, evolução cromossômica, evolução em concerto, famílias multigênicas, transferência horizontal

ABSTRACT

The study of repetitive DNAs has been explored as an important tool to answer various biological questions, including the molecular organization and understanding of chromosome structure, and the analysis of karyotype diversification and evolution. Thus, this study focused in the chromosomal and genomic organization of repetitive DNA with emphasis on representatives of the Phanaeini tribe (Coleoptera: Scarabaeidae: Scarabaeinae), with the aim in understanding the mechanisms involved in the evolutionary dynamics of repetitive DNAs in the genome and their implications. In Phanaeini, especially members of *Coprophanaeus* genus, it was observed that the expansion of repetitive DNAs occurred early in the diversification of the group. Also, these sequences have been involved in the diversification of sex chromosomes of *Coprophanaeus*, and the origin and evolution of the B chromosome observed in *C. cyanescens*. Isolation and mapping of *Mariner* transposons revealed that these sequences had a high diversification during the evolutionary history of Phanaeini, and therefore may also play a role in the pericentromeric region of the chromosomes. Comparative analysis between sequences of related families of *Mariner* transposons showed that these sequences may have been involved in a process of horizontal transfer (HT) in several unrelated groups of animals, especially between insects and mammals, which contributed to the widespread distribution of transposable elements (TEs). Regarding to the chromosomal mapping of multigene families, there was a large variation for 18S rDNA in contrast to the conserved pattern of 5S rDNA, suggesting that distinct evolutionary forces govern the genomic regions that harbor both ribosomal. Furthermore, the chromosomal mapping of repetitive sequences is a useful tool in understanding the processes that govern the evolution of karyotypes and genomes, as here seen for Phanaeini. Additionally, the analysis of 28S ribosomal DNA sequences and non-LTR retrotransposons R2 showed that the level of R2 insertions did not affect significantly the concerted evolution of the ribosomal genes, but may have been involved in the dispersion of rRNA genes to different chromosomes.

Keywords: chromosome evolution, Coleoptera, concerted evolution, horizontal transfer, multigene families, repetitive DNA, transposable elements

SUMÁRIO

Resumo

Abstract

1. INTRODUÇÃO	10
1.1. Características gerais dos DNAs repetitivos e organização dos genomas	10
1.1.1. DNAs repetitivos <i>in tandem</i>	11
1.1.2. Famílias multigênicas	13
1.1.3. Elementos de transposição	16
1.2. Citogenética clássica e molecular em besouros da família Scarabaeidae, com ênfase em Phanaeini	22
2. OBJETIVOS	25
2.1. Objetivo geral	25
2.2. Objetivos específicos	25
3. CAPÍTULOS	26
3.1. Capítulo 1: Heterochromatin, sex chromosomes and rRNA gene clusters in <i>Coprophanaeus</i> beetles (Coleoptera, Scarabaeidae)	27
3.2. Capítulo 2: B chromosome in the beetle <i>Coprophanaeus cyanescens</i> (Scarabaeidae): emphasis in the organization of repetitive DNA sequences	44
3.3. Capítulo 3: <i>Mariner</i> transposable elements in Scarabaeinae coleopterans and their relationship to other <i>Mariner</i> families	58
3.4. Capítulo 4: Horizontal transfers of <i>Mariner</i> transposons between mammals and insects	89
3.5. Capítulo 5: High diversity of R2 transposable elements in the Phanaeini coleopterans	112
4. CONCLUSÕES	152
5. REFERÊNCIAS BIBLIOGRÁFICAS	153

ANEXOS	176
Anexo A1: Heterochromatin, sex chromosomes and rRNA gene clusters in <i>Coprophanaeus</i> beetles (Coleoptera, Scarabaeidae)	177
Anexo A2: B chromosome in the beetle <i>Coprophanaeus cyanescens</i> (Scarabaeidae): emphasis in the organization of repetitive DNA sequences	187
Anexo A3: Horizontal transfers of <i>Mariner</i> transposons between mammals and insects	193
Anexo B: Relação de trabalhos paralelos publicados relacionados ao tema da tese	199

1. INTRODUÇÃO

1.1. Características gerais dos DNAs repetitivos e organização dos genomas

O tamanho do genoma de eucariotos pode variar entre espécies sem relação óbvia com a complexidade do organismo, o número de genes ou o nível de ploidia (Gregory, 2005). Segundo dados do Genome Size Database (<http://www.genomesize.com/>) o tamanho dos genomas já descritos para eucariotos está entre cerca de 0.0023 pg no parasita intestinal *Encephalitozoon intestinalis* (Vivares, 1999) e 1.400 pg na ameoba de vida livre *Chaos chaos* (Friz, 1978), diferença maior do que 600 vezes. Esta ampla variação na quantidade de genoma e a fraca correlação entre complexidade do organismo foi denominada de “paradoxo do valor C” (Thomas, 1971), onde o termo “valor C” refere-se a quantidade de DNA por célula haplóide (geralmente expressa em picogramas). A descoberta, a partir da década de 1970, da presença de grande quantidade de DNAs repetitivos na maioria dos genomas sugeria que o paradoxo do valor C estaria resolvido. No entanto, importantes aspectos que norteiam a biologia das espécies, como variação do tamanho do genoma nos diferentes taxa, quantidade, tipo, distribuição e função de DNA não codificante, precisam também ser analisados para inferências mais precisas da relação tamanho do genoma versus complexidade do organismo (Gregory, 2005).

Parte da variação do tamanho do genoma relatada para eucariotos deve-se ao acúmulo de diferentes quantidade de seqüências repetitivas (Petrov, 2001; Kidwell, 2002). Resumidamente, a expansão da porção repetitiva do genoma envolve processos de amplificação destas seqüências por conversão gênica, *crossing-over* desigual, replicação circular e replicação *slippage* (Charlesworth et al., 1994; Hancock, 1999; Li et al., 2002).

Durante muito tempo as seqüências de DNAs repetitivos foram consideradas seqüências de “DNA lixo” (*junk DNAs*) devido ao desconhecimento de suas funções (Nowak, 1994; Makalowski, 2003; Biémont e Vieira, 2006). Contudo, diversos estudos têm sugerido, e muitas vezes demonstrado, que este tipo de DNA está envolvido com a organização estrutural e funcional do genoma, atuando na regulação gênica, replicação, reparo do DNA, rearranjos cromossômicos, diferenciação e segregação cromossômica, podendo, também, estar relacionados com doenças (Biet et al., 1999; Kidwell, 2002; Li et al., 2002; Shapiro e Sternberg, 2005; Biémont e Vieira, 2006; Feschotte e Pritham, 2007).

O DNA repetitivo é composto por seqüências curtas, idênticas ou similares, que se encontram *in tandem* (lado a lado) ou dispersas pelo genoma. Aquelas que se encontram *in*

tandem compõem o DNA satélite (DNAsat), microssatélites, minissatélites e algumas famílias multigênicas, enquanto as que estão dispersas são representadas pelos elementos transponíveis (Charlesworth et al., 1994; Sumner, 2003; Wicker et al., 2007).

1.1.1. DNAs repetitivos *in tandem*

O número de ocorrências de um padrão de repetição é chamado de número de cópias. O número de cópias em uma determinada região repetitiva *in tandem* é denominada região de número de cópias; sendo que a classificação das seqüências repetitivas *in tandem* é baseada no tamanho do fragmento de repetição e no tamanho do arranjo total das repetições (Ugarkovic e Plohl, 2002; Rao et al., 2010; Figura 1).

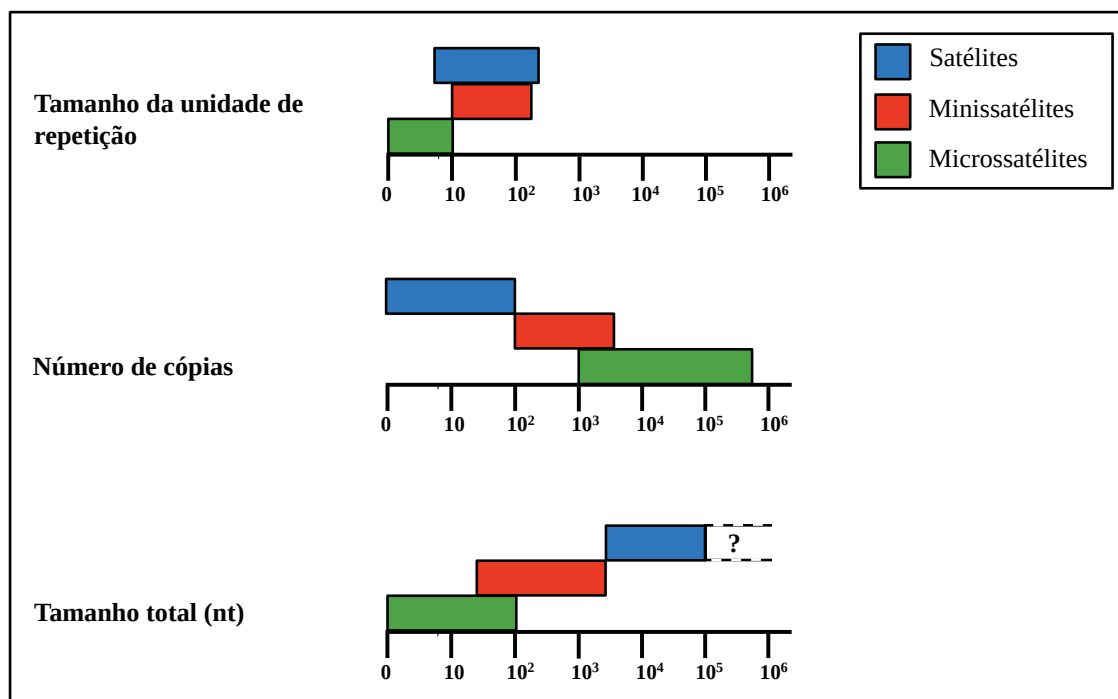


Figura 1. Características gerais dos DNAs repetitivos *in tandem*.

Repetições *in tandem* possuem papéis estruturais e funcionais significativos, ocorrendo em abundância em áreas estruturais, como telômeros, centrômeros, histonas e regiões de ligação. Estas seqüências desempenham um papel regulador perto dos genes e, talvez, até mesmo dentro de genes (Vogt et al., 1992; Gemayel et al., 2010; Rao et al., 2010).

Os DNAs repetitivos *in tandem* podem ser amplamente distribuídos no genoma de uma família ou gênero taxonômico, ou pode ser específico para uma espécie ou cromossomo. Estas repetições podem adquirir variações na seqüência e no número de cópias ao longo da

evolução, podendo ser utilizadas em estudos taxonômicos e filogenéticos (Rao et al., 2010).

As repetições *in tandem* são consecutivas, de acordo com um padrão a partir de um local de duplicação. Os padrões de repetição podem ser classificados como direto, indireto, complemento, complemento reverso ou palíndrome. Uma repetição direta ou forward é a ocorrência de um segmento na mesma fita de DNA e na mesma ordem de nucleotídeos. Uma repetição indireta, inversa ou reversa ocorre na mesma fita, mas a ordem de nucleotídeos é reversa. Repetições complementares são repetições nas quais os nucleotídeos são complementares, de acordo com o pareamento de bases; a repetição ocorre na mesma fita, mas os nucleotídeos são complementares e a ordem dos nucleotídeos é reversa. Palíndromes são repetições combinadas de duas ocorrências em orientações opostas, e lidas da mesma maneira da direita para a esquerda ou vice-versa (Grumbach e Tahi, 1994; Rao et al., 2010).

O primeiro grupo de DNAs repetitivos *in tandem* é representado pelo DNA satélite, sendo este o principal constituinte da heterocromatina (HC), localizando-se preferencialmente nas regiões pericentroméricas e subteloméricas (Yunis e Yasmineh, 1971; John e Miklos, 1979; Juan et al., 1993; Chen et al., 2004; Shapiro e Sternberg, 2005; Eymeri et al., 2009).

As famílias de DNA satélite podem surgir *de novo* devido a mecanismos moleculares como *crossing-over* desigual, replicação circular, replicação *slippage* e mutação. Este tipo de DNA é composto por seqüências altamente repetitivas, agrupadas em um ou alguns locais ao longo de um ou mais cromossomos, e intercaladas com seqüências de cópia única. As unidades repetitivas medem, geralmente, de 100 a 1000 nucleotídeos; enquanto o comprimento total da repetição pode formar matrizes longas ocupando cerca de 100 Mb (Charlesworth et al., 1994; Kubis et al., 1998; Schmidt e Heslop-Harrison, 1998; Vergnaud e Denoeud, 2000; Ugarkovic e Plohl, 2002; Figura 1).

O segundo grupo é composto por um DNA moderadamente repetitivo, com uma unidade de repetição com cerca de 10 a 100 pb, denominado minissatélite ou seqüências com número variável de repetições (VNTR - *Variable Number of Tandem Repeats*) (Jeffreys et al., 1985). Contudo, as unidades de repetição geralmente são inferiores a 50 pb, com um comprimento total variando de 0,5 a 30 kb (Jeffreys et al., 1985; Figura 1). Cada repetição constitui um loco de minissatélite cujos alelos se diferenciam por variações em seu tamanho total, podendo ser usado como marcador genético em uma técnica denominada impressão digital do DNA (DNA *fingerprint*) (Mueller e Wolfenbarger, 1999).

Minissatélites são muitas vezes caracterizados por taxas de mutação elevadas (até 5%), que podem envolver heterogeneidade das repetições ou do número de repetições. As taxas de mutação também apresentam correlação positiva com o tamanho total da matriz de

repetições. Os mecanismos responsáveis pela variabilidade dos minissatélites incluem, possivelmente, replicação *slippage*, transposição, conversão gênica, e eventos de recombinação ou troca desigual entre cromátides irmãs ou entre cromossomos homólogos (Jarman e Wells, 1989; Jeffreys et al., 1990; Richards e Sutherland, 1992; Wolff et al., 1991; Rao et al., 2010).

O terceiro grupo é composto pelos microssatélites, também conhecidos como sequências repetitivas simples (SSR - *Simple Sequence Repeats*). Possuem um comprimento padrão curto, de 1 a 6 pb. São co-dominantes, multialélicos e altamente polimórficos, sendo por isso utilizados como marcadores em estudos de genética de populações, conservação, epidemiologia, testes de parentesco, identificação de indivíduos e mapeamento (Balloux et al., 1998; Röder et al., 1998; Schlötterer, 2000). Microsatélites possuem distribuição ampla no genoma, tendo sido frequentemente encontrados em regiões não-centroméricas, muitos deles sendo localizado perto ou dentro de genes. Entre as funções atribuídas aos microssatélites estão a participação na organização da cromatina, replicação do DNA, recombinação e regulação da atividade gênica (Li et al., 2002).

1.1.2. Famílias multigênicas

O termo “família multigênica” é utilizado para incluir grupos de genes oriundos de um ancestral comum, e que possuem sequências estruturalmente e funcionalmente semelhantes. Exemplos de famílias multigênicas incluem os genes codificadores de hemoglobinas, imunoglobulinas, tubulinas, interferons, histonas e RNAs ribossomais (Nei e Rooney, 2005).

Um dos exemplos mais conhecidos de famílias multigênicas é representado pelos DNAs ribossomais (DNAr), que transcrevem os RNAs ribossomais (RNAr), principais componentes dos ribossomos (Hillis e Dixon, 1991; Eickbush e Eickbush, 2007; Smit et al., 2007). Os ribossomos são um complexo de RNA e proteínas que, conjuntamente, catalisam uma das atividades bioquímicas fundamentais mais conservadas: a síntese protéica (Noller et al., 1992). Algumas das regiões universalmente conservadas do RNAr podem remontar ao “mundo do RNA”, um estágio hipotético de evolução em que RNA realizava todas as reações bioquímicas importantes (Gilbert, 1986). Os genes ribossomais estão presentes em todas as espécies analisadas, sendo observadas múltiplas cópias nos genomas eucarióticos, possibilitando sua utilização para a elaboração das relações evolutivas entre os organismos (Smit et al., 2007).

Na maioria dos eucariotos as sequências de DNAr estão organizadas em dois grupos

arranjados *in tandem*. O menor arranjo é formado pelas seqüências gênicas que transcrevem o RNAr 5S. Estas seqüências gênicas são altamente conservadas e espaçadas por seqüências não transcritas (NTS - *Non Transcribed Spacer*) (Long e Dawid, 1980; Eickbush e Eickbush, 2007; Figura 2b). O arranjo maior, DNAr 45S, é formado pelos genes que transcrevem os RNAs ribossomais 18S, 5.8S e 28S. Cada cluster de DNAr 45S é separado por espaçadores transcritos externos (ETS - *External Transcribed Spacer*) e por espaçadores intergênicos (IGS - *Intergenic Spacers*). Dentro de cada cluster, por sua vez, as regiões que transcrevem os RNAs ribossomais 18S, 5.8S e 28S são espaçadas por espaçadores intergênicos transcritos internos (ITS - *Internal Transcribed Spacer*) (Hillis e Dixon, 1991; Eickbush e Eickbush, 2007; Figura 2a).

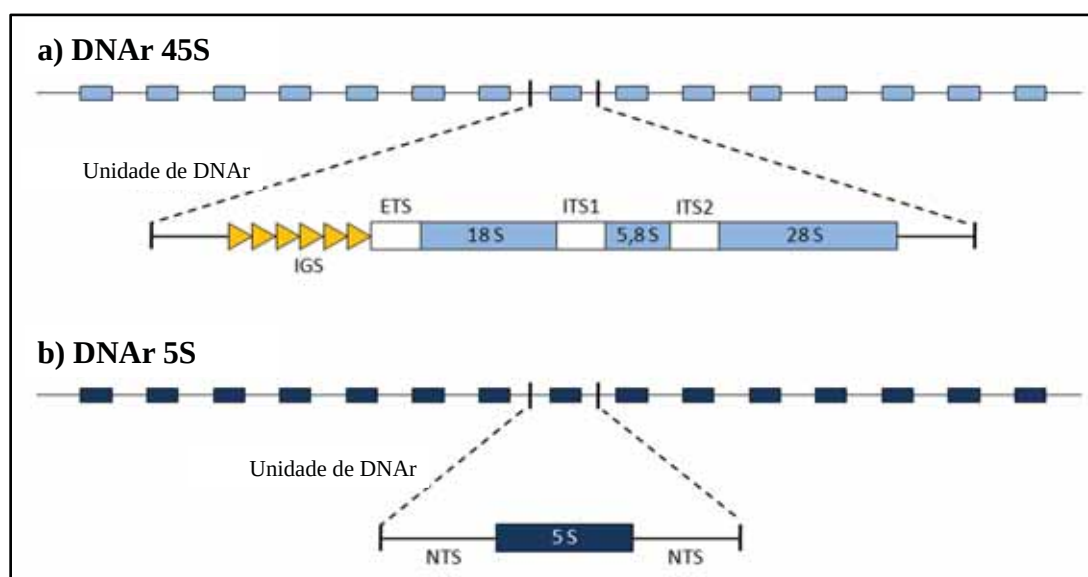


Figura 2. Organização dos genes de RNA ribossomais (RNAr) eucarióticos. Os genes estão organizados em unidades de repetição *in tandem*, como esquematizadas na parte superior. As unidades típicas são mostradas em detalhes na parte inferior. IGS, espaçador intergênico, composto por unidades repetitivas (setas amarelas); ETS, espaçador transcrito externo; ITS, espaçador transcrito interno; NTS, espaçador não-transcrito. Esquema modificado de Martins et al., 2011.

A determinação de qual é o melhor modelo para explicar a evolução das famílias multigênicas tem sido alvo de controvérsias. Os primeiros estudos focando a dinâmica evolutiva das famílias multigênicas começaram utilizando a hemoglobina e a mioglobina como sistemas modelo (Ingram, 1961). A observação de que os genes que codificam estas proteínas são filogeneticamente relacionados e que eles adquiriram novas funções através de

uma divergência gradual levou à proposta do primeiro modelo geral da evolução destas famílias multigênicas, referido como "evolução divergente" (Eirín-López et al., 2012; Figura 3a).

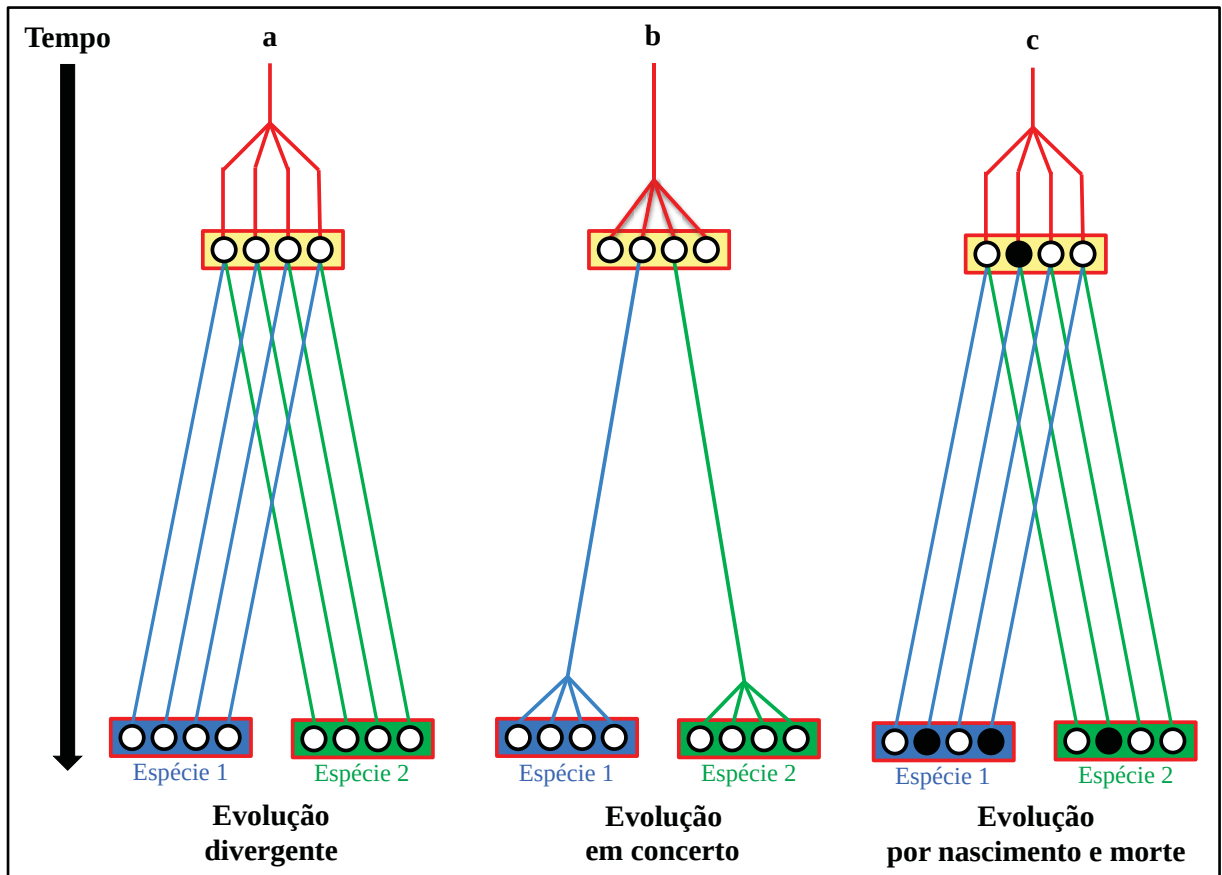


Figura 3. Três diferentes modelos propostos para a evolução de famílias multigênicas. Círculos brancos significam genes funcionais, e círculos pretos pseudogenes. Esquema modificado de Ney e Rooney, 2005.

A validade do modelo de evolução divergente foi contestada pela crescente quantidade de dados obtidos a partir de estudos com outras famílias multigênicas, principalmente estudos enfocando os genes ribossomais. O desenvolvimento de técnicas de sequenciamento de DNA possibilitou a análise de padrões de variação em regiões codificantes e não-codificantes, revelando que as sequências nucleotídicas dos diferentes membros de uma família multigênicas estão mais estreitamente relacionadas dentro de cada espécie do que entre as espécies. Estas observações foram explicadas por um modelo alternativo de evolução das famílias multigênicas denominado "evolução em concerto" (Figura 3b). De acordo com este modelo, depois da separação de uma espécie ancestral em dois descendentes, os membros de uma família gênica poderiam evoluir conjuntamente como um bloco, exibindo um elevado

grau de homogeneidade dentro de uma espécie descendente, considerando que as sequências divergiram gradualmente a partir de espécies estreitamente relacionadas. De acordo com este modelo, uma modificação que ocorre em uma repetição se propaga através dos membros de uma família, e a homogeneização das sequências são um resultado do *crossing-over* desigual e conversão gênica entre os membros de uma família multigênica (Ney e Rooney, 2005; Eirín-López et al., 2012).

A evolução em concerto começou a ser questionada a partir da crescente disponibilidade de dados moleculares provenientes da “era genômica”. Estes dados revelam que a maioria dos membros das famílias multigênicas possuem diversidade genética e funcional inconsistente com os mecanismos de homogeneização. A proposta de um novo modelo de evolução, “nascimento-e-morte” (*birth-and-death*) (Figura 3c), forneceu uma explicação para a geração de novas famílias gênicas. Neste modelo novos genes são criados por duplicação gênica, e alguns genes duplicados são mantidos no genoma por um longo período de tempo, enquanto que outros são eliminados ou tornam-se não-funcionais por meio de mutações deletérias (Nei et al., 1997; Nei e Rooney, 2005; Eirín-López et al., 2012).

Contudo, a controvérsia sobre a evolução das famílias multigênicas continua. As famílias gênicas que produzem produtos gênicos variados, em geral estão sujeitas à evolução por nascimento-e-morte, considerando que este modelo explica a variação genética. Contudo, algumas famílias multigênicas estão sujeitas a evolução em concerto, ou a processos mistos; enquanto outras evoluem seguindo o modelo de nascimento-e-morte submetidas a uma forte seleção purificadora, ao passo que outras evoluem pelo mesmo modelo submetidas à seleção positiva (Nei e Rooney, 2005; Eirín-López et al., 2012).

1.1.3. Elementos de transposição

Os elementos transponíveis (TEs, do Inglês *Transposable Elements*) são um exemplo notável de sucesso evolutivo, exibem grande diversidade e estão presentes em quase todos os reinos, de eubactéria e arqueobactéria até os organismos multicelulares. Apresentam uma grande diversidade genética e funcional, e parecem ter explorado durante o processo evolutivo as mais relevantes maneiras de se duplicar e manterem-se no genoma (Wicker et al., 2007; Rouzic e Capy, 2009).

Os TEs se diferenciam das outras sequências do genoma por possuírem a capacidade de se mover e replicar, gerando plasticidade (Hartl et al., 1992). Uma característica importante destes elementos está nos polimorfismos de sítios de inserção e na variabilidade

no número de cópias que podem surgir dentro e entre as espécies (Shapiro e Sternberg, 2005; Feschotte e Pritham, 2007; O'Donnell e Burns, 2010).

Atualmente, o número e a variedade de TEs descritos está aumentando. A análise dos dados das seqüências dos projetos de sequenciamento de genomas nos permite identificar novos elementos; alguns intimamente relacionado a outros já conhecidos, enquanto novos grupos também estão sendo reportados. Estas descrições têm facilitado os estudos comparativos, e possibilitando muitas análises de evolução de genoma. Contudo, uma anotação consistente e uma classificação geral dos TEs tornam-se ainda necessária (Wicker et al., 2007; Treangen e Salzberg, 2012).

Uma classificação de TEs deve refletir uma filogenia. No entanto, as filogenias são geralmente baseados em uma parte (ou módulo) de um TE. Isso significa que se deve assumir que a filogenia de um módulo de um elemento transponível pode ser equiparada à filogenia das seqüências completas. Portanto, assume-se que as relações filogenéticas de diferentes módulos de TEs são congruentes. Entretanto, frequentemente, é mais apropriado analisar os TEs como conjuntos de módulos com suas próprias histórias (Capy e Maisonhaute, 2002; Capy, 2005).

Wicker et al. (2007) propuseram um sistema de classificação de TE, fornecendo um consenso entre as várias classificações propostas anteriormente. Este sistema unificado de classificação destina-se a facilitar estudos comparativos e evolutivos de TEs, estabelecendo uma nomenclatura: um código de três letras com cada letra, respectivamente, denotando classe, ordem e superfamília; o nome da família (ou subfamília); a seqüência (acesso ao banco de dados) em que o elemento foi encontrado; e o 'running number', que define a inserção individual no acesso (Wicker et al., 2007; Kapitonov e Jurka, 2008).

A classificação dos elementos transponíveis é baseada em suas características estruturais e moleculares, sendo agrupados em classes, subclasses, superfamílias e subfamílias (Wicker et al., 2007). O critério para a classificação de classes se baseia no mecanismo de transposição: a) Classe I: elementos que se transpõem utilizando a transcriptase reversa de uma cópia intermediária de RNA, e b) Classe II: elementos que se transpõem de DNA a DNA utilizando a enzima transposase (Charlesworth et al., 1994; Hua-Van et al., 2005; Kapitonov e Jurka, 2008) (Figura 4).

Os mecanismos de transposição estão relacionados com os meios utilizados pelos elementos transponíveis para se inserir em um novo sítio dentro do genoma. O meio de transposição via DNA pode ser conservativo ou replicativo. No primeiro, o TE é removido de um local e inserido em outro, enquanto no segundo o TE é duplicado antes de ser transportado

para um novo local, aumentando o número de cópias no genoma (Kapitonov e Jurka, 2008). No processo de transposição via RNA, um RNA intermediário é transcrito a partir da sequência de DNA do TE, sendo posteriormente transcrito reversamente em uma nova cópia de DNA, de forma basicamente replicativa (Xiong e Eickbush, 1990; Flavell, 1995). Contudo, muitos TEs mostram vários mecanismos de transposição (Figura 4) e diferentes propriedades dinâmicas, que nem sempre correspondem à sua classificação (Hua-Van et al., 2005).

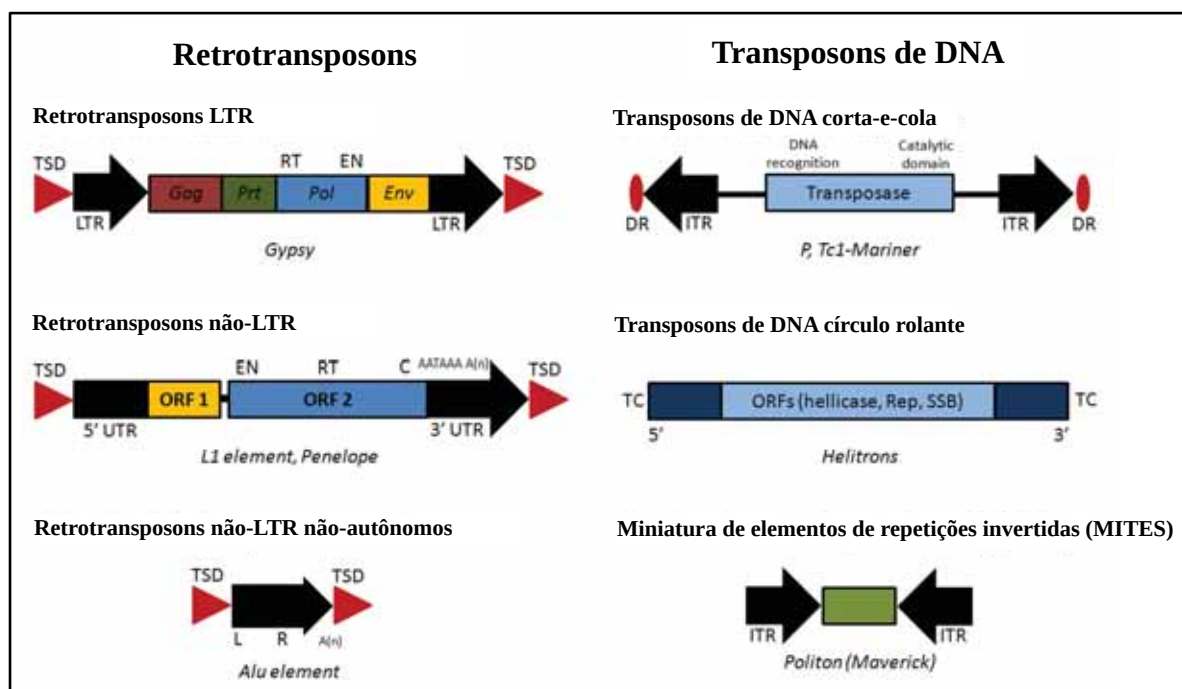


Figura 4. Estrutura generalizada dos principais tipos de elementos transponíveis. Esquema modificado de Martins et al. 2011.

As classes são analisadas utilizando critérios como presença ou ausência de domínios e assinaturas nas proteínas codificadas pelos TEs. A classe I não se divide em subclasses, considerando que os membros deste grupo não clivam ou transferem fitas de DNA para um sítio doador. Os retrotransposons, contudo, são subdivididos em cinco ordens de acordo com seus mecanismos de transposição e filogenia da transcriptase reversa. A classe II, por sua vez, se divide em duas subclasses, que se distinguem pelo número de fitas de DNA que são cortadas durante a transposição. A subclasse I é representada, entre outros, pelos clássicos transposons de DNA do tipo “corta-e-cola” que possuem repetições terminais invertidas (TIRs – *Terminal Inverted Repeats*); enquanto os elementos da subclasse II são representados por TEs que passam por um processo de transposição sem realizar uma clivagem dupla da cadeia de DNA (Capy, 2005; Wicker et al., 2007). Este é o caso, por exemplo, dos elementos

que se transpõem por um mecanismo similar ao de um círculo rolante, clivando somente um filamento de DNA e sem formação de TSDs (Kapitonov e Jurka, 2001).

As superfamílias de uma ordem compartilham uma estratégia de replicação, mas são distintas em relação à algumas características, como a estrutura das proteínas e/ou dos domínios não codificantes. As superfamílias também diferem quanto à presença e tamanho dos sítios de duplicação alvo (TSD – *Target Site Duplication*), uma repetição curta direta, gerada em ambas as pontas de um TE como consequência da inserção. Superfamílias são subdivididas em famílias, que são definidas pela conservação da sequência de DNA. A semelhança entre proteínas é geralmente elevada entre as diferentes famílias que pertencem à mesma superfamília. Embora TEs possam ser classificados em poucas ordens e superfamílias, um único genoma pode conter centenas ou milhares de diversas famílias de TE. Subfamílias são definidas com base em dados filogenéticos e, em casos específicos, podem servir para distinguir internamente elementos autônomos de não-autônomos (Kazazian, 2004; Wicker et al., 2007). O menor nível de classificação dos TEs é o padrão de inserção, que corresponde a um evento específico de transposição e inserção. Em todos os níveis acima de inserção, um elemento pode ser temporariamente classificado como "desconhecido" (Wicker et al., 2007).

As sequências de TEs parecem estar submetidas a diversas forças antagônicas, conduzindo-os a vários padrões evolutivos. Estes padrões dependem de muitos fatores, como a biologia da espécie hospedeira, as características do TE, ou simples acaso. Considerando o nível molecular, quanto mais eficiente o processo de transposição utilizado, melhor será a colonização do genoma hospedeiro. Contudo, se os elementos são deletérios para o hospedeiro, indivíduos carregando muitas cópias podem ser eliminados por seleção natural. A evolução dos genomas também poderia facilmente levar ao aparecimento de sistemas de controle ou regulação da replicação. Mutações genômicas recorrentes podem levar à deleção ou inativação parcial ou completa das cópias de TE, enquanto alguns elementos ou fragmentos podem permanecer integrados ao genoma e participar na função adaptativa do organismo (Charlesworth, 1989; Rouzic e Capy, 2009).

Estudos teóricos sobre a dinâmica dos TEs são geralmente desafiados pela complexidade do processo (Charlesworth et al., 1994). A evolução de cada inserção é similar à dinâmica de um único locus gênico exposto por seleção natural, mutações e deriva genética. Desta maneira, muitos “alelos” podem co-existir em cada inserção, e cada um deles pode possuir uma razão de transposição e diferentes impactos na aptidão do organismo. Cogita-se que o número total dos sítios de inserção varia, e cada evento de transposição leva a um novo locus de inserção (Charlesworth, 1989; McDonald, 1993; Rouzic e Capy, 2009).

Quando um TE se acumula no genoma, torna-se necessário considerar a taxa do número de cópias (Charlesworth e Charlesworth, 1983). Os eventos de transposição conduzem ao aparecimento de cópias em um novo sítio de inserção, enquanto a deleção resulta em perda da cópia de um sítio de inserção original. Considerando estas informações, propõe-se que duas forças evolutivas contrabalaceam a invasão dos TEs: a regulação da transposição e a seleção natural (Figura 5). A transposição (ou duplicação) aumentará a média do número de cópias, enquanto vários tipos de deleções irão eliminar cópias do genoma. Se as inserções são deletérias, os indivíduos que possuem poucas cópias reproduzirão melhor do que outros, e a seleção natural diminuirá a média do número de cópias da população. Em pequenas populações, por sua vez, a deriva genética randômica pode alterar o número de cópias, diminuindo ou aumentando em relação ao valor esperado (Figura 5). No começo do processo de invasão, a razão de transposição é provavelmente alta, e o número de cópias no genoma aumenta. Um estado de equilíbrio pode ser alcançado quando o aumento e a diminuição das forças evolutivas que estão atuando são balanceadas (Rouzik e Capy, 2009).

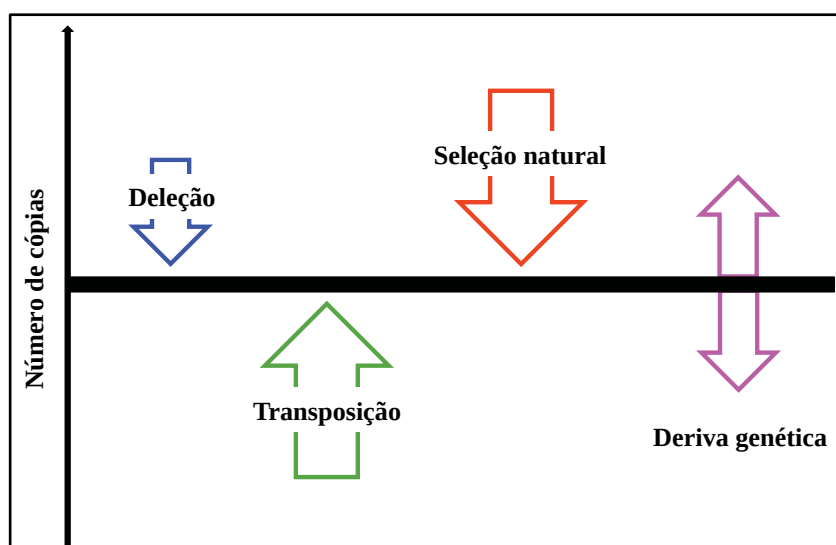


Figura 5. Representação das forças evolutivas atuando sobre o número de cópias de elementos transponíveis no genoma das espécies. Esquema modificado de Lankenau e Volff, 2009.

Todos os modelos descrevem a colonização de uma família de TE como um processo determinístico. O espalhamento de um TE em uma população, e o progressivo aumento do número de cópias, entretanto, não parece ser um mecanismo predito, fornecido por uma grande população; limitando, desta maneira, a influência da deriva genética. Contudo, apesar do tamanho da população, um elemento não pode escapar da aleatoriedade do começo de uma

invasão (Biémont, 1994; Brookfield e Badge, 1997; Rouzik e Capy, 2009).

Cada novo elemento que coloniza o genoma de uma espécie provém de uma sequência de um TE relacionado do mesmo genoma ou do genoma de uma outra espécie. Os genomas estão repletos de cópias de TEs inativos ou incompletos, os quais podem potencialmente recombinar e gerar um TE novo e funcional. Contudo, muitas invasões de TEs parecem estar relacionadas a transferências horizontais (HTs) interespecíficas, evidenciando a habilidade que estas sequências possuem em invadir genomas, não importando a distância filogenética entre as espécies (Kidwell, 1992; Leaver, 2001; Kurlang et al., 2003; Davis e Wurdack, 2004; Schaak et al., 2010).

A capacidade de invasão de uma família de TE depende diretamente de sua razão de transposição inicial, sendo que algumas especificidades da biologia do TE pode alterar a probabilidade de fixação (Rouzik e Capy, 2009). Apesar de poucas exceções, quase todos os modelos de dinâmica de TEs propõem que depois de um estágio de invasão inicial, uma família de TE atinge o equilíbrio (Charlesworth, 1991). Após um rápido estágio de invasão de um elemento autônomo (ativo), mecanismos reversíveis e não-reversíveis limitarão a razão de transposição e o espalhamento do número de cópias. Cópias não autônomas podem, eventualmente, apresentar uma vantagem para multiplicar-se em relação às cópias autônomas. Desta maneira, as únicas sequências de TE que persistem no genoma são elementos inativos, que serão lentamente eliminados e fragmentados (Rouzik e Capy, 2009).

Após a inclusão em uma nova espécie, uma única cópia de TE ativa pode se amplificar, ou será rapidamente eliminada por seleção natural e deriva genética. O número de cópias, então, aumenta, e algumas mutações podem ocorrer nestes elementos funcionais. Alguns destes elementos “mutados” podem conduzir ao aparecimento de cópias não-autônomas, que são capazes de se amplificar quando cópias autônomas estão presentes no mesmo genoma. Algumas outras cópias podem trazer um aspecto adaptativo ao hospedeiro, sendo, então, domesticadas e fixadas, ainda que não possuam capacidade transcricional. A atividade de uma família de TE desaparece quando os seus elementos ativos são progressivamente perdidos devido a deleções e mutações, considerando que o aumento de sequências não-autônomas resulta em diminuição da transposição. Contudo, algumas cópias autônomas podem escapar do declínio, e podem iniciar um novo processo de invasão na mesma espécie ou em outra espécie através de HT (Kidwell e Lisch, 2001; Rouzik e Capy, 2009). Desta maneira, ao longo de milhões de anos de evolução, os TE têm-se equilibrado entre os efeitos prejudiciais e benéficos em uma espécie mediante a modificações no genoma (Kazazian, 2004).

1.2. Citogenética clássica e molecular em besouros da família Scarabaeidae, com ênfase em Phanaeini

A família Scarabaeidae tem destaque com cerca de 400 espécies estudadas citogeneticamente. Essa família tem sido considerada conservada cariotipicamente, em relação ao número diplóide $2n = 20$, mecanismo sexual Xy_p e morfologia meta-submetacêntrica (meiofórmula $9II + Xy_p$). Pode-se observar uma variabilidade de número diplóide de $2n = 8$ a $2n = 30$ e sete diferentes tipos de mecanismos sexuais, XY, Xy, Xy_p , XY_p , Xy_r , neo-XY e X0 (Smith e Virkki, 1978; Yadav e Pillai, 1979; Yadav et al., 1979; Moura et al., 2003; Cabral-de-Mello et al., 2007, 2008). Dentre as subfamílias de Scarabaeidae, Scarabaeinae é a menos conservada cariotipicamente, apresentando variação no número diplóide de $2n = 8$ a $2n = 24$, e seis mecanismos sexuais (XY, Xy_p , XY_p , Xy_r , neo-XY e X0). A variabilidade cromossômica observada em Scarabaeinae foi originada a partir de rearranjos cromossômicos, que modificaram o cariótipo $2n = 20$, Xy_p e a morfologia cromossômica meta-submetacêntrica, considerados modal e ancestral para o grupo (Smith e Virkki, 1978; Yadav et al., 1979; Bione et al., 2005a; Cabral-de-Mello et al., 2008). Nos representantes desta subfamília, os rearranjos são bastante comuns, contudo, os mecanismos de evolução cromossômica nas diversas tribos são diferentes (Cabral-de-Mello et al., 2008).

A tribo Phanaeini é bastante diversa cariotipicamente, devido principalmente a rearranjos do tipo fusão autossômica (Smith e Virkki, 1978; Yadav e Pillai, 1979; Vidal, 1984). Algumas espécies da tribo apresentam fusões entre autossomos, e entre autossomos e cromossomos sexuais, ocasionando a redução do número diplóide destas espécies e a modificação do mecanismo sexual de Xy_p para neo-XY (Vidal, 1984; Cabral-de-Mello et al., 2008). Rearranjo do tipo perda do cromossomo y também foi evidenciado, gerando o cariótipo $2n = 19$, X0 (Cabral-de-Mello et al., 2008).

O padrão de distribuição de HC observado em espécies de Scarabaeidae, em geral, é pericentromérico, coincidindo com o comumente relatado para a ordem Coleoptera. Contudo, também têm sido descritos blocos intersticiais, teloméricos e cromossomos difásicos, como observado em representantes da subfamília Scarabaeinae (Vidal e Giacomozzi, 1978; Colomba et al., 1996; Bione et al., 2005a). Neste subfamília há ocorrência de ampla variabilidade da HC, entretanto, a maioria das espécies analisadas apresentou blocos localizados na região pericentromérica (Vidal e Nocera, 1984; Colomba et al., 1996; Wilson e Angus, 2005; Angus et al., 2007). Em Coleoptera, no geral, o cromossomo X apresenta-se quase completamente heterocromático, embora algumas espécies possuam blocos

paracêntricos, pericentroméricos e subteloméricos. Estudos realizados até o momento quanto à composição de bases da HC têm mostrado uma grande heterogeneidade em riqueza de pares de bases (Colomba et al., 2000; Moura et al., 2003; Vitturi et al., 2003; Bione et al., 2005a, b). Em Scarabaeinae, estudos cromossômicos utilizando fluorocromos base específicos foram realizados em um número reduzido de espécies, entretanto, os resultados são bastante distintos e a HC tem se caracterizado como um material cromossômico com alta variabilidade e heterogeneidade (Colomba et al., 1996; Colomba et al., 2000, 2006; Bione et al., 2005a).

Em representantes da tribo Phanaeini foi observada uma grande quantidade de HC e a ocorrência de cromossomos difásicos (Vidal, 1980 apud Virkki, 1983; Bione et al., 2005a; Oliveira et al., 2010). O cromossomo X destas espécies é quase totalmente heterocromático e o Y eucromático, padrão comumente relatado para Scarabaeidae (Vitturi et al., 2003; Bione et al., 2005b; Angus et al., 2007; Oliveira et al., 2010).

Embora ainda escasso, o mapeamento de DNAs repetitivos em Coleoptera tem se focado na análise do DNAr 45S (De La Rúa et al., 1996; Gómez-Zurita et al., 2004; Almeida et al., 2010; Cabral-de-Mello et al., 2011a, b). Em Scarabaeidae o mapeamento do DNAr 45S revelou nas espécies analisadas uma ampla variabilidade em relação ao número e localização dos clusters. Em representantes desta família o DNAr tem sido observado em pelo menos um par, podendo ser localizado restritamente em cromossomos autossômicos, sexuais ou ambos; apresentando uma variação de dois clusters (comum para várias espécies) a 15 sítios em *C. ensifer* (Colomba et al., 2000, 2006; Moura et al., 2003; Vitturi et al., 2003; Bione et al., 2005a, b; Oliveira et al., 2010; Cabral-de-Mello et al., 2011a). O mapeamento de DNAr 5S e histone H3, por sua vez, tem mostrado uma grande conservação em relação ao número de sítios. Em Coleoptera, este tipo de estudos foi realizado em um reduzido número de espécies, e apresentam-se co-localizados nas espécies de Scarabaeinae, indicando uma ligação entre estas duas famílias multigênicas na organização genômica desta subfamília (Cabral-de-Mello et al., 2010a; Cabral-de-Mello et al., 2011a, b).

Em Coleoptera, estudos moleculares objetivando a análise e caracterização do DNA satélite (DNAsat) foram realizados principalmente em representantes das famílias Cincidelidae e Tenebrionidae, sendo a última mais bem estudada (Ugarkovic et al., 1992, 1995, 1996; Vogler e DeSalle, 1993; Plohl et al., 1993; Pons et al., 1993, 1997, 2002, 2004; Plohl e Ugarkovic, 1994; Bruvo et al., 2003, 2007; Gálian e Vogler, 2003; Mravinac et al., 2004, 2005, 2007; Gibbs et al., 2008; Wang et al., 2008). O mapeamento destes DNAs satélites têm revelado preferencialmente um padrão pericentromérico, compartilhado por todos os cromossomos autossômicos, e com menor frequência também pelos cromossomos

sexuais (Dover, 1982; Juan et al., 1993; Barceló et al., 1998; Lorite et al., 2001; Gálian e Vogler, 2003; Palomeque et al., 2005).

Estudos moleculares visando a análise de elementos transponíveis (TEs) em Coleoptera foram realizados em representantes das famílias Chrysomelidae, Coccinellidae, Meloidae, Scarabaeidae, Staphylinidae e Tenebrionidae (Jakubczak, et al., 1991; Robertson, 1993; Robertson e Lampe, 1995a; Braquart et al., 2001; Robertson e MacLeod, 1993; Lampe et al., 2003). Exceto Tenebrionidae, que possui um alto número de espécies analisadas por diversas metodologias, estes estudos limitam-se ao estudo de uma ou poucas espécies de cada família (Nakakita et al., 1981; Pavlopoulos et al., 2004; Angelini e Jockusch, 2008). Até o momento foram descritos em coleópteros os elementos transponíveis DCE1, DCE2, Het-A, *Mariner*, *Minos*, MITE, *PstI*, R1, R2 e TART (Burke et al., 1993; Robertson, 1993; Robertson e MacLeod, 1993; Robertson e Lampe, 1995a, b; Braquart et al., 2001; Pons, 2003; Lampe et al., 2003; Pavlopoulos et al., 2007). Contudo, nenhum mapeamento cromossômico de TEs foi descrito para a ordem.

O estudo de DNAs repetitivos tem se mostrado uma ferramenta esclarecedora para diversas questões, desde o entendimento da estrutura de diferentes regiões cromossômicas, à análises relacionadas à diversificação cariotípica. Além disso, o mapeamento cromossômico destas sequências tem contribuído para o entendimento da estrutura e evolução dos genomas de eucariotos. Em Phanaeini, o uso do mapeamento de sequências repetitivas pode ser uma ferramenta útil no entendimento da ampla diversidade cariotípica observada no grupo, podendo contribuir para o esclarecimento dos processos que governam a evolução de seus cariótipos e genomas como um todo.

2. OBJETIVOS

2.1. Objetivo geral

- Analisar a organização cromossômica e genômica dos DNAs repetitivos com ênfase em representantes da tribo Phanaeini (Coleoptera: Scarabaeidae: Scarabaeinae), visando a compreensão dos mecanismos evolutivos envolvidos na dinâmica dos DNAs repetitivos no genoma e suas implicações.

2.2. Objetivos específicos

- Caracterizar a evolução macro-cromossômica, cromossomos sexuais e diversificação de DNAs repetitivos em 5 espécies do gênero *Coprophanaeus*;
- Analisar a composição e organização do cariótipo de *Coprophanaeus cyanescens*, enfatizando a investigação da presença de cromossomo B;
- Investigar a fração e organização de DNAs repetitivos em três espécies de Phanaeini que possuem extensos blocos de heterocromatina constitutiva (*Coprophanaeus cyanescens*, *C. ensifer* e *Diabroctis mimas*);
- Analisar a transferência horizontal (HT) de famílias de transposons *Mariner* entre insetos e mamíferos;
- Investigar a variação nucleotídica do DNAr 28S e sua relação com a inserção do retrotransposons não-LTR R2 em Phanaeini.

3. CAPÍTULO

Os resultados obtidos durante o desenvolvimento da presente tese foram divididos em cinco capítulos apresentados na forma de manuscritos:

3.1. Capítulo 1: Heterochromatin, sex chromosomes and rRNA gene clusters in *Coprophanaeus* beetles (Coleoptera, Scarabaeidae)

3.2. Capítulo 2: B chromosome in the beetle *Coprophanaeus cyanescens* (Scarabaeidae): emphasis in the organization of repetitive DNA sequences

3.3. Capítulo 3: *Mariner* transposable elements in Scarabaeinae coleopterans and their relationship to other *Mariner* families

3.4. Capítulo 4: Horizontal transfers of *Mariner* transposons between mammals and insects

3.5. Capítulo 5: High diversity of R2 transposable elements in the Phanaeini coleopterans

3.1. Capítulo 1:

Heterochromatin, sex chromosomes and rRNA gene clusters in *Coprophanaeus* beetles (Coleoptera, Scarabaeidae)

Sárah G Oliveira¹, Diogo C Cabral-de-Mello², Amanda P Arcanjo³, Crislaine Xavier⁴, Maria J Souza³, Cesar Martins^{1*}, Rita C Moura⁴

¹Departamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista/UNESP, Botucatu, São Paulo, Brazil

²Departamento de Biologia, Instituto de Biociências, Universidade Estadual Paulista/UNESP, Rio Claro, São Paulo, Brazil

³Departamento de Genética, Centro de Ciências Biológicas/CCB, Universidade Federal de Pernambuco/UFPE, Recife, Pernambuco, Brazil

⁴Departamento de Biologia, Instituto de Ciências Biológicas/ICB, Universidade de Pernambuco/UPE, Recife, Pernambuco, Brazil

* *Corresponding author*

Abstract

Repetitive DNA sequences constitute a high fraction of eukaryotic genomes and are considered a key component for the chromosome and karyotype evolution. For a better understanding of their evolutionary role in beetles, we examined the chromosomes of 5 species of the genus *Coprophanaeus* by C-banding, fluorochrome staining CMA3/DA/DAPI, and fluorescence *in situ* hybridization (FISH) with probes for 18S and 5S rRNA genes. The *Coprophanaeus* species have identical chromosome numbers and a conserved chromosome morphology. However, they show different sex chromosome forms, XY, Xy, XY_p, and heterochromatin seems to be involved in the origin and diversification of these forms. C-banding showed primarily the presence of di-phasic chromosomes in all species examined. After CMA₃/DA/DAPI staining, 1-9 autosomal pairs showed CMA₃-positive blocks depending on the species, while DAPI-positive blocks were detected only in *Coprophanaeus dardanus*. FISH mapping revealed 5S rDNA signals in one autosomal pair in each species, whereas the number of pairs with 18S rDNA signals varied from 1-8 between the *Coprophanaeus* species. Our results suggest that distinct genetic mechanisms had been involved in the karyotype evolution of *Coprophanaeus* species, i.e. mechanisms maintaining the conserved number of 5S rDNA clusters and those generating variability in the amount of heterochromatin, sex chromosome forms, and distribution of 18S rDNA clusters.

Key Words: Chromosomal evolution, Comparative cytogenetics, Karyotype, Meiosis, Multigene family, Repetitive DNA

Introduction

A large fraction of eukaryotic genomes is composed of repetitive DNA sequences. Although these sequences were long considered ‘junk’ DNA, they are now considered a major player in genomic architecture and function and are thought to play a role in the structural and functional organization of chromosomes [Nowak, 1994; Biémont and Vieira, 2006]. The variation in total DNA content between species is mostly due to variation in the repetitive DNA content which is often present at many chromosomal sites [Sumner, 2003]. Repetitive DNA sequences have been extensively explored as markers for cytogenetic mapping because their reiterated number of copies generates easily visualizable signals on chromosomes. The cytogenetic mapping of repetitive DNA provides useful chromosome markers that can be used in studies of genome organization, species evolution, and applied genetics and in the identification of specific chromosomes, homologous chromosomes, chromosome rearrangements, and sex chromosomes.

The subfamily Scarabaeinae (Scarabaeidae, Polyphaga) comprises a diverse and cosmopolitan group of Coleoptera that play an important role in the conservation of ecosystems as seed dispersers, pollinators, and recyclers of organic matter [Louzada, 2008]. The group encompasses approximately 250 genera and over 6,000 species, with 618 species belonging to 49 genera in Brazil [Hanski and Cambefort, 1991; Vaz-de-Mello, 2000]. Among them, the tribe Phanaeini comprises approximately 150 species distributed in 12 genera, of which *Gromphas* and *Oruscatus* are considered the most basal taxa, while *Coprophanæus*, *Dendropaemon*, *Megatharsis*, and *Tetramereia* are more derived [Philips et al., 2004]. The phanaeines are restricted to the Neotropics [Edmonds, 1967, 1972; Philips et al., 2004], and their representatives have been detected in all Brazilian regions [Edmonds and Zidek, 2010].

The karyotype variation observed for Polyphaga coleopterans is clearly represented in the high diversity of sex chromosomes identified in the group that includes the X0, XY, neo-XY, Xy_p, XY_p, XY_r, XY_c, and multiple systems [White, 1973; Smith and Virkki, 1978; Ferreira et al., 1984; Mesa and Fontanetti, 1985; Yadav et al., 1990; Galián et al., 2002; Rozék et al., 2004; Cabral-de-Mello et al., 2010a]. In the different forms of association between X and Y chromosomes the ‘y’ represents a smaller chromosome than the X, the ‘p’ refers to a ‘parachute’ meiotic conformation between the X and Y(y), ‘r’ refers to a road-shaped configuration, and ‘c’ to a rare central association of X and Y. The Scarabaeinae species follows the high level of

karyotype diversity of the group, with variation in the diploid number from $2n = 8$ in *Eurysternus caribaeus* to $2n = 24$ in *Oniticellus spinipes* (= *Timiocellus spinipes*) and 7 types of sex chromosomes mechanisms (XY_p , Xy_p , XY , Xy , Xy_r , neo- XY , and $X0$) [Smith and Virkki, 1978; Yadav et al., 1979; Cabral-de-Mello et al., 2008]. Although diverse with respect to their chromosomal characteristics, there is a predominance of the $2n = 20$ karyotype, Xy_p sex chromosome mechanism, and meta-submetacentric chromosome morphology among Scarabaeinae species. Among the 12 tribes of Scarabaeinae, Phanaeini predominantly present a chromosome number of $2n = 20$ and a high variability among the sex chromosomes mechanisms (XY , Xy , XY_p , Xy_p , neo- XY , and $X0$) [Cabral-de-Mello et al., 2008, 2010c; Oliveira et al., 2010].

To characterize the macro-chromosomal evolution, sex chromosomes, and diversification of repetitive DNA, including heterochromatin and rDNAs, the chromosomes of 5 species of *Coprophanaeus* belonging to the 3 subgenera of the group (*Coprophanaeus*, *Megaphanaeus*, and *Metallophanaeus*) were studied with classical cytogenetic techniques (conventional staining, C-banding, and fluorochrome CMA₃/DA/DAPI staining) and fluorescence *in situ* hybridization (FISH) using probes for the 18S and 5S rRNA genes. With the exception of previously published information concerning the basic karyotype data of *C. (Coprophanaeus) dardanus* [Cabral-de-Mello et al., 2008], this study is the first to describe the basic karyotypes, the heterochromatin composition and distribution, and the rRNA gene organization of *C. (Coprophanaeus) acrisius*, *C. (Coprophanaeus) dardanus*, *C. (Metallophanaeus) pertyi*, *C. (Metallophanaeus) horus*, and *C. (Megaphanaeus) bellicosus*. The results indicated the conservation of macro-chromosomal structures in the genus, the conservation of the number of 5S rDNA sites in contrast to heterochromatin heterogeneity, and variation in the chromosomal distribution of 18S rRNA gene clusters and sex chromosomes.

Materials and Methods

Samples from adult males of *Coprophanaeus (Coprophanaeus) acrisius* (MacLeay, 1819), *C. (Coprophanaeus) dardanus* (MacLeay, 1819), *C. (Metallophanaeus) pertyi* (Olsoufieff, 1924), *C. (Metallophanaeus) horus* (Waterhouse, 1891), and *C. (Megaphanaeus) bellicosus* (Olivier, 1789) were collected from different regions in Pernambuco and Minas Gerais States, Brazil (table 1). The animals were collected in the wild according to Brazilian laws for

environmental protection (wild collection permit, MMA/ IBAMA/SISBIO no. 2376–1). The experimental research on animals was conducted according to the international guidelines followed by São Paulo State University (Protocol no. 35/08 CEEA/IBB/UNESP). The testes were fixed in Carnoy solution (3:1 ethanol:acetic acid) and stored in the freezer at -20°C.

Chromosome preparations were obtained by the classical testicular follicle squashing technique. The C-banding technique was performed according to Sumner [1972] with modifications to analyze the male meiotic spreads and heterochromatin regions. Fluorochrome staining using the combination Chromomycin A3/Distamycin/4'6-diamino-2'-phenylindol (CMA₃/DA/DAPI) method was performed according to Schweizer [1976] with modifications to quantify the heterochromatin in relation to the AT/ GC base pair content.

DNA probes for the 18S and 5S rRNA genes were obtained from the dung beetle *Dichotomius semisquamosus* [Cabral-de-Mello et al., 2010b]. The 18S rRNA gene probe was labeled by nick translation using biotin-11-dATP (Invitrogen, San Diego, Calif., USA), and the 5S rRNA gene was labeled with digoxigenin-11-dUTP (Roche, Mannheim, Germany) via PCR. The FISH procedures were performed according to the protocol adapted by Cabral-de-Mello et al. [2010b] for Coleoptera. The chromosome preparations were counterstained with DAPI, and the slides were mounted in Vectashield mounting medium (Vector, Burlingame, Calif., USA). Images were captured using an Olympus DP71 digital camera coupled to a BX61 Olympus microscope and were optimized for brightness and contrast using Adobe Photoshop CS2.

Results

Karyotypes and Heterochromatin Characterization

Coprophanaeus species showed similar karyotypes consisting of $2n = 20$ and meta-submetacentric chromosomes with a gradual reduction in size. The species analyzed, however, differed in sex-chromosome mechanisms: XY in *C. (Coprophanaeus) acrisius* and *C. (Megaphanaeus) bellicosus*, with the X and Y chromosomes similar in size; Xy in *C. (Coprophanaeus) dardanus*, with the y being smaller than the X; and XY_p in *C. (Metallophanaeus) horus* and *C. (Metallophanaeus) pertyi* (fig. 1; table 1) where the 'p' refers to a 'parachute' configuration.

The C-banding technique primarily showed the presence of diphasic chromosomes (with the long arm heterochromatic) in the karyotypes of the 5 analyzed species (fig. 1; table 1). In *C.*

(*Megaphanaeus*) *bellicosus*, all the autosomal chromosomes were diphasic (fig. 1e), while in *C. (Coprophanaeus) acrisius*, *C. (Metallophanaeus) pertyi*, and *C. (Metallophanaeus) horus* one autosomal pair had a larger pericentromeric block, and the remaining autosomes had a diphasic pattern (fig. 1a, c, d). In *C. (Coprophanaeus) dardanus*, 2 autosomal pairs had large heterochromatic blocks in the pericentromeric regions and 7 pairs were diphasic (fig. 1b). With respect to the sex chromosomes, the X was almost completely heterochromatic in all species analyzed (fig. 1), but the Y chromosome was almost completely heterochromatic in *C. (Coprophanaeus) acrisius* and *C. (Megaphanaeus) bellicosus* (fig. 1a, e) and had a large pericentromeric block in *C. (Metallophanaeus) pertyi* (fig. 1c).

In the 5 analyzed species, CMA₃/DA/DAPI staining revealed positive signals only in the autosomes (fig. 2; table 1). Chromomycin A₃-positive (CMA₃⁺) signals were observed in 2 chromosomal pairs of *C. (Coprophanaeus) acrisius* and *C. (Metallophanaeus) pertyi*, with one block in the terminal region and the other in the pericentromeric region of another pair (fig. 2a, c). Which chromosomes carried the blocks could not be determined due to the high chromosome condensation level. In *C. (Megaphanaeus) bellicosus*, CMA₃⁺ blocks were observed in the pericentromeric region of one pair and in the terminal region of 2 other pairs (fig. 2e). CMA₃⁺ signals were observed only in the pericentromeric region of one pair in *C. (Metallophanaeus) horus* (fig. 2d). In these 4 species, DAPI-positive (DAPI⁺) signals were not visualized (fig. 2a, c-e). In *C. (Coprophanaeus) dardanus*, the CMA₃/DA/DAPI staining identified DAPI⁺ blocks in 5 autosomal pairs, 3 with the blocks in the terminal region and 2 with the blocks in the pericentromeric region. Two autosomal pairs that have pericentromeric DAPI⁺ blocks also showed adjacent CMA₃⁺ sites. CMA₃⁺ blocks were also observed in a third autosomal bivalent region, but this region was DAPI-negative (DAPI⁻) (fig. 2b).

Cytogenetic Mapping of 5S and 18S rDNA

Chromosomal mapping of 5S rDNA revealed signals in the pericentromeric region of one autosomal bivalent with 2 chiasmata in all 5 species analyzed (fig. 3; table 1). In contrast, the mapping of 18S rDNA revealed variability for this marker among the species studied. These gene clusters were located in the pericentromeric regions of 2 autosomal pairs in *C. (Coprophanaeus) dardanus* (fig. 3b) and in the terminal regions of 3 autosomal pairs in *C. (Coprophanaeus) acrisius* (only one homolog member was labeled for 1 of the 3 chromosomal pairs) (fig. 3a). The

18S rDNA sites were observed in the terminal region of one autosomal pair in *C. (Metallophanaeus) horus* and *C. (Metallophanaeus) pertyi* and in the pericentromeric region of the X chromosome in *C. (Metallophanaeus) horus* (fig. 3c, d). In *C. (Megaphanaeus) bellicosus*, the 18S rDNA sites were identified in the pericentromeric region of 8 autosomal pairs, although 2 pairs were heteromorphic, possessing only one site in one of the homologs (fig. 3e).

Discussion

Heterochromatin Features

The diploid number, $2n = 20$, and the meta-submetacentric morphology observed in *Coprophanaeus* species are in agreement with the karyotype considered modal for Scarabaeidae, for the suborder Polyphaga, and for Coleoptera [Smith and Virkki, 1978; Wilson and Angus, 2005; Cabral-de-Mello et al., 2008]. The presence of $2n = 20$ is also shared among other distinct Phanaeini species, such as *Bolbites ornitoides*, *Diabroctis mimas*, *Gromphas lacordairei*, and *Sulcophanaeus imperator* [Vidal, 1984; Bione et al., 2005a; Cabral-de-Mello et al., 2008], and in species belonging to closely related groups such as Onitini, Eucraniini, and Dichotomini [Smith and Virkki, 1978; Vidal, 1984; Colomba et al., 1996; Angus et al., 2007; Cabral-de-Mello et al., 2008]. These data indicate that $2n = 20$ could be the ancient condition for this group. In contrast, variability for this characteristic was observed in some species of the Phanaeini tribe, such as *Oxysternon silenus* and *Phanaeus daphnis* [Smith and Virkki, 1978; Cabral-de-Mello et al., 2008].

The high amount of heterochromatin observed in *Coprophanaeus* species differs from the most common pattern observed in the family Scarabaeidae which generally presents heterochromatic blocks in the pericentromeric region of the autosomes, although variations in the location of the heterochromatin along the sex chromosomes were observed [Moura et al., 2003; Wilson and Angus, 2004, 2005; Bione et al., 2005b; Wilson and Angus, 2006; Angus et al., 2007; Silva et al., 2009]. For the tribe Phanaeini, diphasic autosomes were also observed in *C. (Coprophanaeus) cyanescens*, *C. (Megaphanaeus) ensifer*, and *Diabroctis mimas* [Bione et al., 2005a; Oliveira et al., 2010], and a large amount of heterochromatin in the genome appears to be common in this group. These large heterochromatic blocks indicate that the chromosomes of the species suffered an expansion of repetitive DNA, involving amplification by different spreading processes as described for vertebrates and other insects [Charlesworth et al., 1994; Hancock,

1999; Landais et al., 2000; Li et al., 2002]. All the Phanaeini species analyzed showed a high amount of heterochromatin [Bione et al., 2005a; Oliveira et al., 2010; Cabral-de-Mello et al., 2011a; present work] which could indicate that the amplification of heterochromatin occurred early during group diversification.

The heterochromatic regions have shown a heterogeneous pattern for base pair richness in representatives of the Scarabaeidae family [Colomba et al., 1996; Vitturi et al., 1999; Colomba et al., 2000; Moura et al., 2003; Vitturi et al., 2003; Bione et al., 2005b; Colomba et al., 2006; Cabral-de-Mello et al., 2010b, 2010c; Oliveira et al., 2010; present work]. The presence of CMA₃⁺ blocks is predominant in the genomes of Phanaeini species. DAPI⁺ blocks, however, were observed in *C. (Coprophanaeus) dardanus* and *Bolbites onitoides* [Virkki, 1983]. Different mechanisms can be involved in the heterogeneous pattern of the base pair richness observed in *Coprophanaeus* species. Bione et al. [2005a] suggested the occurrence of small duplications in tandem, resulting in the additional heterochromatic blocks observed in *Diabroctis mimas*. Indeed, different families of satellite DNA have been previously observed in the genome of some Coleoptera representatives [Juan et al., 1993; Ugarkovic et al., 1995; Zinic et al., 2000]. Lohe and Roberts [2000] proposed that the heterochromatin can undergo different rearrangements, allowing the expansion or reduction of satellite repeats during evolution and may allow the emergence of new satellites. The spread of new satellite DNA can occur through a variety of mechanisms, including gene conversion, unequal crossing-over, slippage replication, transposition, and RNA-mediated exchanges [Dover, 2002; Palomeque and Lorite, 2008]. These events may have occurred during the chromosomal evolution of *Coprophanaeus* species which explains the increase in the quantity and variability of heterochromatin, including the variability observed in the sex chromosomes.

Sex Chromosome Systems

Distinct chromosomal rearrangements involving inversions, autosome fusions and fissions, X-autosome fusions, and y chromosome loss appear to be the most probable events involved in the generation of diploid number and sex chromosome variability observed in Scarabaeinae and in the Phanaeini tribe. The mechanisms of chromosomal evolution differ between different tribes and between different genera in the group. Although high chromosomal variability was observed in Phanaeini, *Coprophanaeus* showed conservation of the diploid

number and chromosome morphology as well as of sex chromosome mechanisms. Although the *Coprophanaeus* species analyzed have X and y chromosomes, these chromosomes showed different forms of association. During meiosis in the Xy and XY systems, the sex chromosomes associate in the classical configuration pattern of sex chromosomes [White, 1973; Smith and Virkki, 1978]. In the Xy_p mechanism, however, the X chromosome represents the dorsal of the parachute, and the y chromosome represents the parachutist, connected by 2 thin wires. Usually, the X is a medium metacentric chromosome and the y is an extremely small metacentric chromosome. In cases where the X and Y chromosomes are relatively large and of similar size, the sex mechanism is called XY_p [White, 1973; Mesa and Fontanetti, 1985].

The XY_p mechanism observed in *C. (Metallophanaeus) horus*, *C. (Metallophanaeus) pertyi*, and the previously analyzed *C. (Coprophanaeus) cyanescens* [Oliveira et al., 2010] was also observed in other Scarabaeidae species, such as *Deltochilum (Calhyboma) verruciferum* and *Malagoniella aff astianax*, and *Phyllophaga (Phyllophaga) aff capilata* [Moura et al., 2003; Cabral-de-Mello et al., 2008, 2010c; Oliveira et al., 2010]. The XY_p mechanism is not a common mechanism for the group; it is considered to be derived and is characterized by sex chromosomes with similar sizes, differing from the Xy_p mechanism often observed in Coleoptera. This mechanism is thus considered primitive. From the observation of the XY_p mechanism in *Itu zeus* (Myxophaga), Mesa and Fontanetti [1985] proposed 2 hypotheses to explain the origin of this mechanism: (i) from an achiasmatic XY system, with X and Y chromosomes of similar sizes and (ii) by the addition of heterochromatin in an Xy_p bivalent system. Recently, Cabral-de-Mello et al. [2010c] proposed a third hypothesis based on the analysis of the XY_p mechanism in *Deltochilum (Calhyboma) verruciferum* which allowed the observation of argyrophilic proteins in the lumen of the sex bivalent region and the presence of a large, diphasic Y chromosome, reinforcing the idea that this mechanism could have been derived from a primitive Xy_p , at least in *D. (Calhyboma) verruciferum*.

In the Phanaeini tribe, the chiasmatic Xy mechanism without the formation of the parachute observed in *C. (Coprophanaeus) dardanus* was described only in *Bolbites onitoides* [Vidal, 1984; Bione et al., 2005a; Cabral-de-Mello et al., 2008]. The XY sex mechanism observed in *C. (Coprophanaeus) acrisius* and *C. (Megaphanaeus) bellicosus*, however, was also observed in 3 other Phanaeini species: *C. (Megaphanaeus) ensifer*, *Phanaeus chalcomelas*, and *P. igneus* [Smith and Virkki, 1978; Martins, 1994; Cabral-de-Mello et al., 2008]. The variability

of the sex chromosomes in the *Coprophanaeus* species is enhanced by the occurrence of different chiasmatic (XY and Xy) and achiasmatic (XY_p and Xy_p) sex mechanisms. The variations observed relative to the amount and distribution of heterochromatin suggest the involvement of heterochromatin in the origin of the sex chromosome diversity in this group which may have involved the amplification and gain of heterochromatin.

Chromosomal Organization of 5S and 18S rRNA Gene Clusters

The presence of only one chromosome bivalent carrying the 18S rDNA clusters is the ancient condition in Scarabaeinae [Cabral-de-Mello et al., 2011a] and in the order Coleoptera, at least for Polyphaga representatives [Schneider et al., 2007; Cabral-de-Mello et al., 2011a]. The presence of more than one chromosome bivalent carrying clusters of 18S rDNA was previously observed in *C. (Coprophanaeus) cyanescens* and *C. (Megaphanaeus) ensifer*, with 5 and 15 sites (the largest number of 18S rDNA sites of the order Coleoptera), respectively [Oliveira et al., 2010; Cabral-de-Mello et al., 2011a]. Among the *Coprophanaeus* species studied in this work, only *C. (Metallophanaeus) pertyi* showed one autosomal pair with 18S rDNA sites, while *C. (Coprophanaeus) bellicosus* had 14 sites of 18S rDNA. *Coprophanaeus (Coprophanaeus) cyanescens*, *C. (Metallophanaeus) horus*, and *C. (Megaphanaeus) ensifer* show intraspecific polymorphisms with respect to both the number and/or position of 18S rDNA sites [Oliveira et al., 2010; Cabral-de-Mello et al., 2011a; present work]. Intraspecific polymorphisms, the repositioning of the rDNA sites, increasing numbers of rDNA sites, and the movement of rDNA sites to different autosomes and sex chromosomes were observed [Oliveira et al., 2010; Cabral-de-Mello et al., 2011a].

Correlating the amount of heterochromatin with the number of 18S rDNA sites was not possible because the heterochromatin amount and distribution is very similar in all *Coprophanaeus* analyzed. Considering the subgenera distribution, *Metallophanaeus* appears to present the ancestral condition, showing conservation in the number of sites, while *Megaphanaeus* appears to have undergone amplification of 18S rDNA sites, with *Coprophanaeus* showing an intermediate position. Despite these observations, addressing a correlation between the phylogenetic position and the variation of 18S rDNA sites was not possible.

The data obtained by FISH with the 18S rDNA probes in *Coprophanaeus* species support

the hypothesis proposed by Sánchez-Gea et al. [2000] for inter- and intra-specific polymorphisms observed in the genus *Zabrus* (Carabidae) and by Oliveira et al. [2010] for the large number of rDNA sites in *C. (Megaphanaeus) ensifer*. This hypothesis postulates that the association of rDNA with heterochromatic regions, including the presence of transposable elements, could affect the occurrence of these polymorphisms.

In contrast to the variability in the number of 18S rDNA sites, the 5 *Coprophanaeus* species analyzed in this work, as well as the previously analyzed *C. (Coprophanaeus) cyanescens* and *C. (Megaphanaeus) ensifer* [Cabral-de-Mello et al., 2011a], showed only one chromosomal pair carrying 5S rRNA genes. Although defining which autosomal pair carries the 5S rDNA clusters is not possible, it appears to be the same pair in all species due to its meiotic behavior. In *C. (Megaphanaeus) ensifer*, the 5S rDNA clusters are located on the sex chromosomes [Cabral-de-Mello et al., 2011a], possibly as a derived condition that resulted from chromosomal rearrangements. In addition to the species analyzed here, 14 other previously analyzed Scarabaeinae species showed only one pair of autosomal chromosomes carrying the 5S rDNA sites [Cabral-de-Mello et al., 2011a]. Some species have shown more than one 5S rDNA site, although the presence of 5S rDNA at only one site with a distinct location from the 18S rDNA sequences is commonly observed in diverse eukaryote groups [Martins and Galetti, 2001; Vitturi et al., 2002; Ribeiro and Fernandez, 2004; Loreto et al., 2008a; Puerma et al., 2008; Cabral-de-Mello et al., 2011a, b]. The physical separation observed between the 5S and 18S rDNA sites could be the result of a functional advantage for these ribosomal sequences, considering that the 18S rRNA gene is transcribed by RNA polymerase I and that the 5S rRNA gene is transcribed by RNA polymerase III [Roussel and Hernandez-Verdun, 1994; Sumner, 2003; Raska et al., 2004]. Despite the diversification of the species, the presence of one 5S rDNA site in different species of *Coprophanaeus* and in other Coleoptera may indicate that this represents the ancestral condition that has been preserved from major changes, possibly due to a purifying selection that prevented the spread of these sequences to other regions of the genome [Rooney and Ward, 2005; Fujiwara et al., 2009]. In contrast, the clusters of 45S rDNA evolution appear to be under the action of ectopic recombination generated from different mechanisms such as unequal crossing-over and gene conversion and transposable elements [Dover, 1986; Montgomery et al., 1991; Charlesworth et al., 1994; Petrov et al., 2003], allowing the spreading of this repetitive family to new locations as here observed for *Coprophanaeus*.

Conclusions

The *Coprophanaeus* species showed karyotypes considered to be the possible ancient condition for Coleoptera. The high amount of heterochromatin appears to be common in *Coprophanaeus* species and indicates that the chromosomes of the species underwent an expansion of repetitive DNA early during group diversification. The different sex mechanisms observed suggest that the amplification and distribution of heterochromatin were involved in the origin of these mechanisms. The heterogeneous pattern for base pair richness also indicates the heterogeneity of the heterochromatin. The variation observed for the 18S rDNA in contrast to that of conserved 5S rDNA patterns suggests that distinct evolutionary forces govern the genomic regions that harbor both rRNA gene classes.

Acknowledgements

The authors are grateful to J. Louzada, C.M.Q. Costa, F. França, and F.A.B. Silva for the viability of the collection in Minas Gerais State and the taxonomic identification of the species studied. This study was supported by the Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE – BPD-003.2.02/08 e APQ-0464-2.02/10), the Programa de Fortalecimento Acadêmico da UPE (PFAUPE), and the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP).

References

As referências citadas neste e nos demais capítulos encontram-se no item 5 (Referências Bibliográficas).

Figure legends

Fig. 1. C-banding of metaphase I chromosomes of *Coprophanaeus* species. **a** *C. (Coprophanaeus) acrisius*. **b** *C. (Coprophanaeus) dardanus*. **c** *C. (Metallophanaeus) pertyi*. **d** *C. (Metallophanaeus) horus*. **e** *C. (Megaphanaeus) bellicosus*. The arrows indicate the sex chromosome bivalents, the asterisks indicate the autosomal pairs with pericentromeric blocks of heterochromatin, and the arrowhead in (b) indicates the autosomal pair with a heterochromatic block in the short arm. In (d) the autosomal pair with pericentromeric blocks of heterochromatin is shown in detail in 1 (mitotic pair) and 2 (meiotic pair). Bars = 5 μ m.

Fig. 2. Fluorochrome staining (CMA₃/DA/DAPI) of metaphases I of *Coprophanaeus* species. **a** *C. (Coprophanaeus) acrisius*. **b** *C. (Coprophanaeus) dardanus*. **c** *C. (Metallophanaeus) pertyi*. **d** *C. (Metallophanaeus) horus*. **e** *C. (Megaphanaeus) bellicosus*. The arrows indicate the sex chromosome bivalents. In (b) the DAPI⁺ blocks adjacent to the CMA₃⁺ blocks are indicated with asterisks (pericentromeric blocks) and arrowheads (terminal blocks). Bars = 5 μ m.

Fig. 3. FISH with 18S (green) and 5S rRNA (red) genes probed in metaphases I of *Coprophanaeus* species. **a** *C. (Coprophanaeus) acrisius*. **b** *C. (Coprophanaeus) dardanus*. **c** *C. (Metallophanaeus) pertyi*. **d** *C. (Metallophanaeus) horus*. **e** *C. (Megaphanaeus) bellicosus*. The arrows indicate the sex chromosome bivalents. The X and Y sex chromosomes of the metaphase plate in (d) are indicated in the box. Bars = 5 μ m.

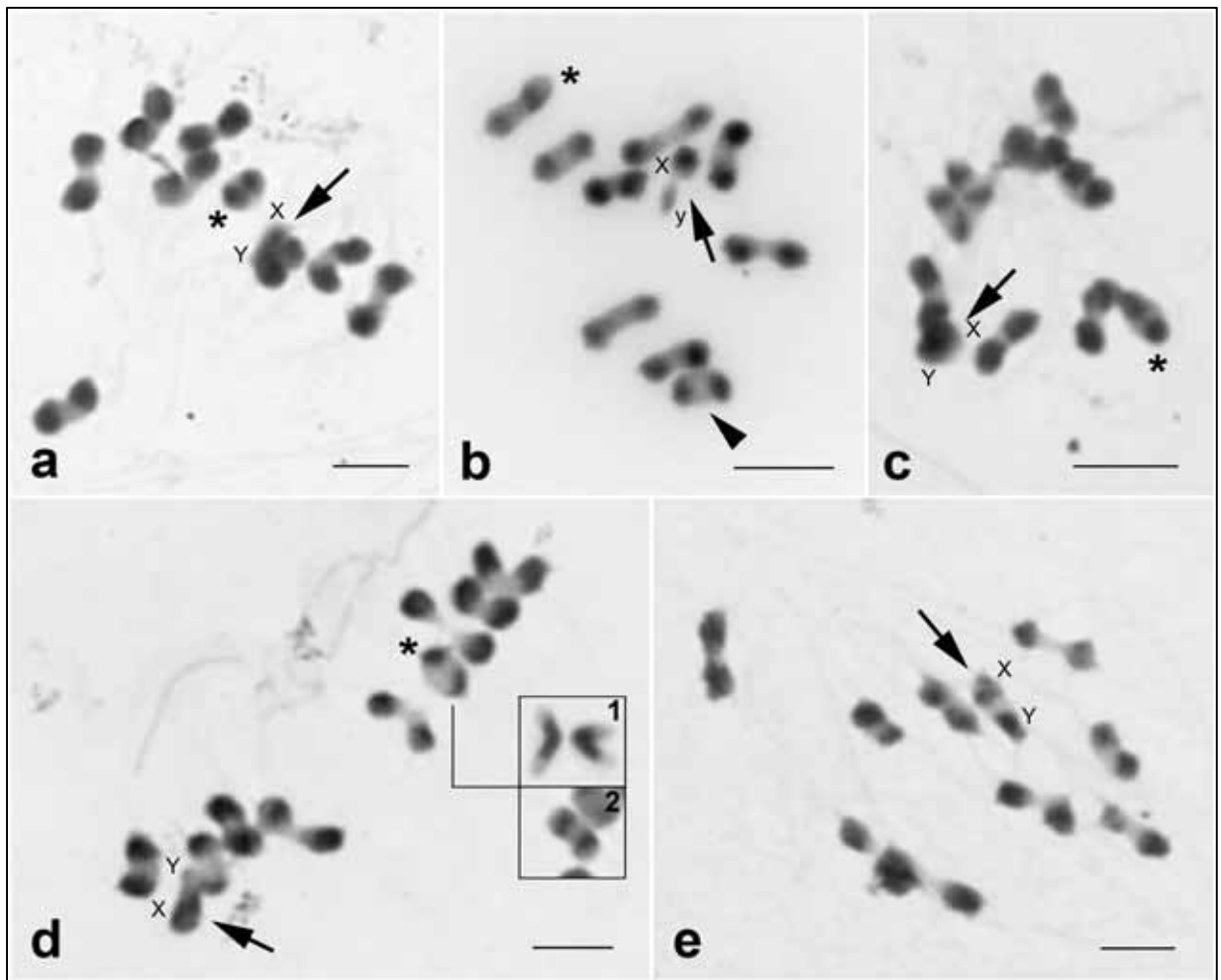


Figure 1

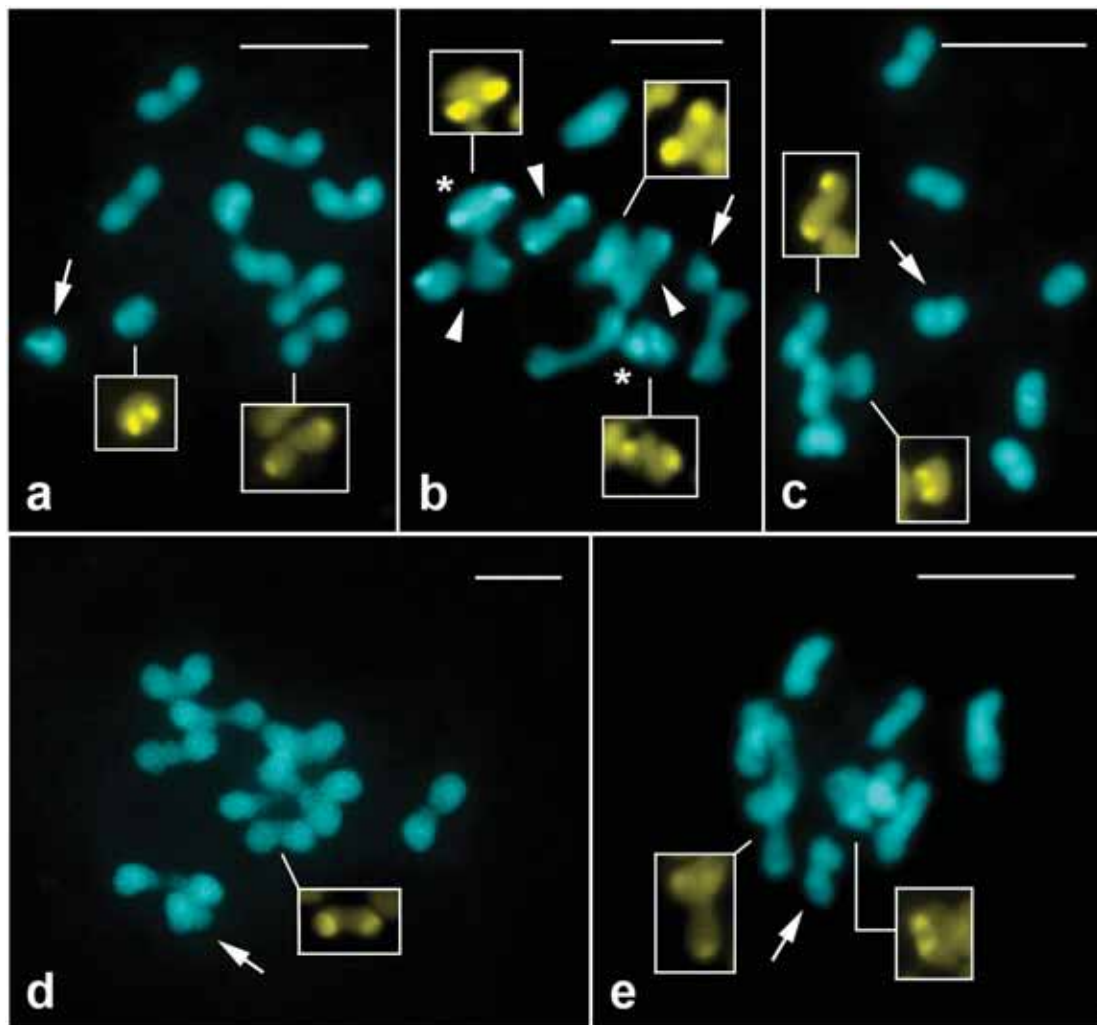


Figure 2

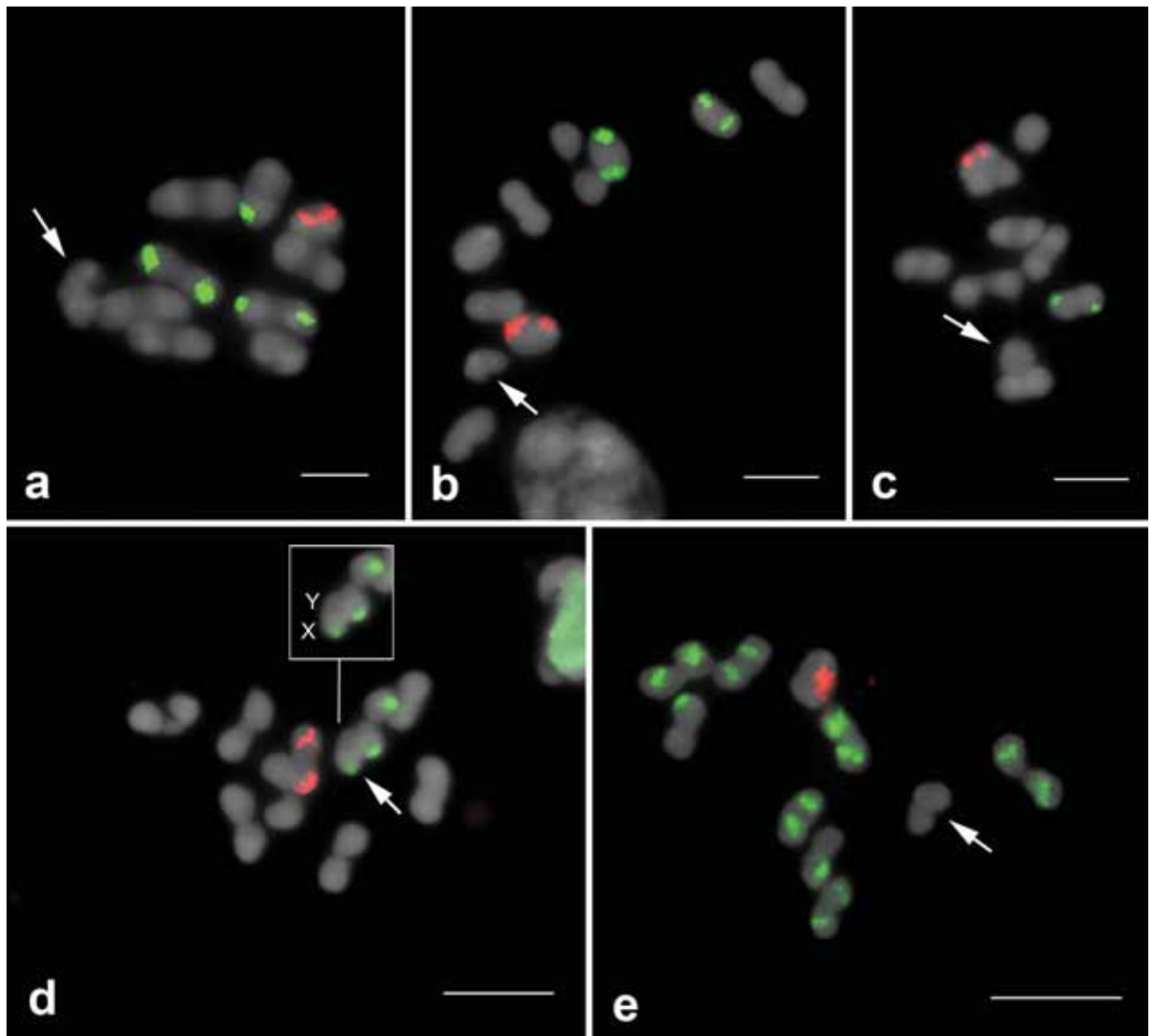


Figure 3

Table 1. Diploid numbers, heterochromatin patterns, and chromosomal location of rRNA gene clusters in *Coprophanaeus* species

Species	Chromosomal formula (males)	Hetero-chromatin distribution	CMA ₃ /DA/DAPI		45S rDNA		5S rDNA		Collection sites in Brazil	References
			auto-some	sex chro-mosome	auto-some	sex chro-mosome	auto-some	sex chro-mosome		
<i>Coprophanaeus (Coprophanaeus) acrisius</i>	20 = 9 + XY	diphasic	4: CMA ₃ ⁺	–	5	–	2	–	Refúgio Ecológico Charles Darwin, Igarassu-PE	this work
<i>C. (Coprophanaeus) cyaneus</i>	20 = 9 + XY _p	diphasic	2: CMA ₃ ^{±a}	–	5/4 ^p	–/X ^p	2	–	Refúgio Ecológico Charles Darwin, Igarassu-PE; Andaraí, Chapada Diamantina-BA	Oliveira et al. [2010]; Cabral-de-Mello et al. [2011a]
<i>C. (Coprophanaeus) dardanus</i>	20 = 9 + Xy	diphasic	10: DAPI ⁺ 6: CMA ₃ ⁺	–	4	–	2	–	Refúgio Ecológico Charles Darwin, Igarassu-PE	Cabral-de-Mello et al. [2008]; this work
<i>C. (Metallophanaeus) horus</i>	20 = 9 + XY _p	diphasic	2: CMA ₃ ⁺	–	2	X	2	–	Carrancas-MG	this work
<i>C. (Metallophanaeus) pertyi</i>	20 = 9 + XY _p	diphasic	4: CMA ₃ ⁺	–	2	–	2	–	Brejo Novo, Caruaru-PE	this work
<i>C. (Megaphanaeus) bellicosus</i>	20 = 9 + XY	diphasic	6: CMA ₃ ⁺	–	14	–	2	–	Parque Ecológico João Vasconcelos Sobrinho, Caruaru-PE	this work
<i>C. (Megaphanaeus) ensifer</i>	20 = 9 + XY	diphasic	18: CMA ₃ ⁺	Y: CMA ₃ ⁺	14/10 ^p	X/– ^p	–	X, Y	Aldeia, Paudalho-PE; Andaraí, Chapada Diamantina-BA	Martins [1994]; Oliveira et al. [2010]; Cabral-de-Mello et al. [2011a]

PE = Pernambuco State; BA = Bahia State; MG = Minas Gerais State.

^a Patterns previously misidentified for the sex chromosomes. ^b Presence of a polymorphism in the same species.

3.2. Capítulo 2

B chromosome in the beetle *Coprophanaeus cyanescens* (Scarabaeidae): emphasis in the organization of repetitive DNA sequences

Sarah Gomes de Oliveira¹, Rita de Cassia de Moura², Cesar Martins^{1,*}

¹UNESP - Sao Paulo State University, Bioscience Institute, Morphology Department, Botucatu, SP, 18618-970, Brazil.

²UPE - Pernambuco State University, Biological Sciences Institute, Department of Biology, Recife, PE, 50100-130, Brazil

*Corresponding author

Email addresses:

SGO: sarahg.oliveira@gmail.com

RCM: ritamoura.upe@gmail.com

CM: cmartins@ibb.unesp.br

Abstract

Background: To contribute to the knowledge of coleopteran cytogenetics, especially with respect to the genomic content of B chromosomes, we analyzed the composition and organization of repetitive DNA sequences in the *Coprophanaeus cyanescens* karyotype. We used conventional staining and the application of fluorescence *in situ* hybridization (FISH) mapping using as probes *Cot*-1 DNA fraction, the 18S and 5S rRNA genes, and the *LOA-like* non-LTR transposable element (TE)

Results: The conventional analysis detected 3 individuals (among 50 analyzed) carrying one small metacentric and mitotically unstable B chromosome. The FISH analysis revealed a pericentromeric block of *Cot*-1 DNA in the B chromosome but no 18S or 5S rDNA clusters in this extra element. Using the *LOA-like* TE probe, the FISH analysis revealed large pericentromeric blocks in eight autosomal bivalents and in the B chromosome and a pericentromeric block extending to the short arm in one autosomal pair. No positive hybridization signal was observed for the *LOA-like* element in the sex chromosomes.

Conclusions: The results indicate that the origin of the B chromosome is associated with the autosomal elements, as demonstrated by the hybridization with *Cot*-1 DNA and the *LOA-like* TE. The present study is the first report on the cytogenetic mapping of a TE in coleopteran chromosomes. These TEs could also have been involved in the origin and evolution of the B chromosome in *C. cyanescens*.

Keywords: Chromosomal rearrangements, Heterochromatin, Multigene families, Supernumerary chromosome, Transposable elements

Background

Eukaryote genomes are composed of classical genes and genetic elements, including transposable elements (TEs), B chromosomes and several cytoplasmic factors that do not follow Mendelian laws of inheritance (Camacho et al., 2000). B chromosomes (also called supernumerary or accessory chromosomes) are not essential for the life of a species and are thus considered “dispensable” additional chromosomes. B chromosomes have been observed in approximately 15% of living species (White, 1973; Jone, 1991; Camacho et al., 2000; Palestis et al., 2004). Most B chromosomes are heterochromatic and composed of repetitive DNA sequences, supporting the idea that these chromosomes are non-coding. However, some B chromosomes show the presence of active genes (Green, 2004; Marschner et al., 2007; Houben et al., 2011). B chromosomes demonstrate an irregular behavior during mitosis and meiosis that allows them to accumulate in the germ line in a non-Mendelian pattern of inheritance (Kean et al., 1982; Palestis et al. 2004). Although B chromosomes have been the focus of intensive work in a diversity of eukaryotic species (Houben et al., 1996; Houben et al., 1997; McAllister e Werren, 1997; Sharbel et al., 1998; Goodwin et al. 2001; Perfectti e Werren, 2001; Bertolotto et al., 2004; Coluccia et al., 2004; Borisov, 2012), several questions concerning their origin, evolutionary mechanism and function remain unanswered.

In Coleoptera, the presence of B chromosomes has been described in approximately 80 species belonging to several families, including Buprestidae (Moura et al., 2008), Cantharidae (James and Angus, 2007), Cicindelidae (Proença et al., 2002) and Scarabaeidae (Angus et al., 2007; Cabral-de-Mello et al., 2010a). In general, the studies in Coleoptera have concentrated on the presence or absence of B chromosomes in species, with few reports covering their frequency in populations and/or their molecular content (Camacho, 2005; Angus et al., 2007; Moura et al., 2008; Cabral-de-Mello et al., 2010a). There are a few reports that document the presence of B chromosomes in the Scarabaeidae family, including species of the Scarabaeinae and Cetoniinae subfamilies (Angus et al., 2007; Cabral-de-Mello et al., 2010a). Among scarabaeines, the *Coprophanaeus* species (Phanaeinae) showed similar karyotypes consisting of $2n = 20$ and meta-submetacentric chromosomes with a gradual reduction in size, three types of sex chromosomes mechanisms (XY, Xy, XY_p), a high amount of constitutive heterochromatin, and there is no description of B chromosomes for this group until now (Palomeque et al., 2005; Cabral-de-Mello et al., 2010c; Oliveira et al., 2010). Besides their karyotype characteristics, the phanaeines are

restricted to the Neotropical region and play an important role in the ecosystems including nutrient recycling (Edmonds, 1972; Philips et al., 2004; Edmonds and Zidek, 2010).

Although the cytogenetic mapping of repetitive DNA sequences has been performed for several species of coleopterans, the data are limited to the analysis of satellite DNA, rRNA and H3 histone genes e.g. (Pons et al., 1997; Gálian and Vogler, 2003; Bione et al., 2005b; Palomeque et al., 2005; Cabral-de-Mello et al., 2010a, c; Oliveira et al., 2010; Cabral-de-Mello et al., 2011a; Oliveira et al., 2012a). Based in the heterochromatin nature of the B chromosomes and that several families of TEs are particularly enriched in heterochromatin, it is particularly interesting the analysis of TE sequences in relation to their organization in B chromosomes. Considering the gap of knowledge on the genomic content of Coleoptera B chromosomes, the present work performed molecular cytogenetic mapping of repetitive DNAs in the beetle *Coprophanaeus cyanescens*, with emphasis in the investigation of the B chromosome.

Results

The standard karyotype observed in *C. cyanescens* was $2n = 20$, XY_p (“p” refers to a “parachute” meiotic conformation between the X and Y), with meta-submetacentric chromosomes that showed a gradual reduction in size (Figure 1a). In addition, three individuals among the 50 analyzed (6%) carried 1 small-sized B meta-submetacentric chromosome. For each individual carrying the B chromosome, at least 30 metaphase I stages were analyzed, and 13.8% of the cells did not present the extra chromosome, indicating mitotic instability. The B chromosome had a condensation pattern similar to that of the autosomal chromosomes and was easily recognized as a small univalent structure in metaphase I (Figure 1).

The FISH analysis using the *Cot*-1 DNA probe revealed positive hybridization in the long arms of all the autosomal chromosomes and the X and Y chromosome and in a pericentromeric block in the B chromosome (Figure 1b). The chromosomal mapping using the 18S and 5S rDNA probes showed clusters on distinct chromosomes (Figure 1c). The 18S rDNA clusters were observed at nine sites (four autosomal pairs plus one single chromosome), and the 5S rDNA clusters were observed at two sites (one autosomal pair) (Figure 1c). None of the rDNA probes hybridized with the B chromosome (Figure 1c).

Analysis of the non-LTR retrotransposon sequence (hereafter named the *LOA-like* non-LTR retrotransposon), which was isolated by polymerase chain reaction (PCR) and subsequently

cloned, revealed a segment of 223 bp that shared ~65% similarity to the *Baggins-1_Nvi* family previously identified in *Nasonia vitripennis* (Jurka, 2009). The alignment of these sequences is shown in Additional file 1. FISH analysis using probes for the *LOA-like* element revealed large pericentromeric blocks in eight autosomal bivalents and the B chromosome and a pericentromeric block extending to the short arm in one autosomal pair; a positive hybridization signal was not observed in the sex chromosomes (Figure 1d).

Discussion

Basic characteristics of the *C. cyanescens* karyotype

The basic karyotype structure for *C. cyanescens* (composed of $2n = 20$, XY_p , with meta-submetacentric chromosomes) is in concordance with previous karyotype data reported for *Coprophanaeus* species (Oliveira et al., 2010, Oliveira et al., 2012a). However, this is the first study to identify a B chromosome in this species as well as in the Phanaeini tribe. In contrast to the small size of the B chromosome observed in *C. cyanescens*, the B chromosomes were medium- or large-sized in the other Scarabaeidae species (Wilson and Angus, 2005; Angus et al., 2007; Cabral-de-Mello et al., 2010a). In *Onthophagus vacca*, the presence of one medium-sized B chromosome was observed in heterochromatin in the centromeric region, whereas *Onthophagus similis* and *O. gazella* showed respectively medium- and small-sized B chromosomes; however, there was no information about the heterochromatic pattern. Large heterochromatic B chromosomes, ranging in number from three to nine, were detected in all the specimens studied for *Bubas bubalus* (Angus et al., 2007). Individuals carrying one heterochromatic B chromosome in two populations of *Dichotomius geminatus*, corresponding to an average prevalence rate of 20.93% and 25.00% in each of the populations, were observed (Cabral-de-Mello et al., 2010a).

The frequency with which B chromosomes are detected in natural populations varies widely between populations. B chromosomes can be present in high frequencies based on the degree to which a species can tolerate the extra chromosome and their power of accumulation (Camacho, 2005). It is difficult to determine the factors that are involved in the low frequency of B chromosomes in the population studied, and several mechanisms may be involved, including selection, random transmission, and historical factors.

Among Coleoptera species, the studies reporting the presence of B chromosomes have

generally focused on the presence or absence of this element and have not considered their frequency in the population or their molecular content (Camacho, 2005; Wilson and Angus, 2005; Angus et al., 2007; Moura et al., 2008). The presence of B chromosomes was reported in representatives of the Cetoniinae and Scarabaeinae, subfamilies of Scarabaeidae (Angus et al., 2007; Cabral-de-Mello et al., 2010a). The evolution of the Scarabaeinae karyotype appears to have occurred under diverse mechanisms of chromosomal rearrangements (Cabral-de-Mello et al., 2008), which could have contributed to the origin of the B chromosome in this group.

Molecular cytogenetic mapping of *C. cyanescens*

The hybridization of the *Cot*-1 DNA to the pericentromeric regions extending up to the long arms of *C. cyanescens* chromosomes is in agreement with the heterochromatin distribution pattern observed in this species (Oliveira et al., 2010). Although heterochromatin analyses were not conducted in the present work, the accumulation of repeated DNAs in the pericentromeric region of the B suggests also the compartmentalization of heterochromatin in the same region. The formation of the heterochromatic chromocenters in the Phanaeini species (Bione et al., 2005a; Colomba et al., 2006) indicates that this mechanism of heterochromatin amplification may be involved in the formation of diphasic chromosomes, including the large pericentromeric block of the B chromosome.

The distribution of *Cot*-1 DNA in the A complement and the B chromosome suggests an intraspecific origin of the extra element and the occurrence of homogenization mechanisms in the heterochromatic regions between the B and A elements. Generally, B chromosomes of more recent origin are enriched in repetitive DNA sequences when compared with the genome from which they originated (Camacho, 2000; Camacho, 2005). This enrichment is indicative of a massive amplification of repetitive sequences over a relatively short time-scale; and, it has also been suggested that repetitive sequences amplification may be a mechanism through which a chromosome fragment (as a neo-B chromosome) may become stabilized and selected (Camacho, 2000; Camacho, 2005). This does not appear to be the case for *C. cyanescens*, indicating that the B chromosome may not have been recently established in this species. Although the data obtained indicates an intraspecific origin of the B chromosome, it was not possible to identify which chromosomal A element was involved in the process. However, the chromosomes carrying the 5S and 18S RNA genes are probably not involved in this process, as the B element does not

contain rRNA gene sequences.

The cytogenetic mapping of the *LOA-like* non-LTR retrotransposon mostly to the pericentromeric regions, including those of the B chromosome, indicates the exchange of genetic material between the A and B chromosomes, implying that the B chromosome has coexisted with the A chromosomes during the period of transposition. However, it is not possible to reject the hypothesis that the B chromosome originated from a segment without *LOA-like* that was received later, by transposition. According to a previous report (Beukeboom, 1994), B chromosomes can accumulate DNA from various sources, including transposable elements, and may affect the structure of the genome by ectopic recombination. A study in *Drosophila melanogaster* identified 25 transposon-mediated rearrangements by ectopic recombination in the region flanking the white locus (Montgomery et al., 1991). The B chromosomes could act as a refuge for TEs, which in turn would generate structural variability in the whole genome. The hybridization that occurred in homologous regions, such as the pericentromeric regions, is another indication of recombination between the A complement and the B chromosome, and this recombination event could be explained by the chromocenter formation during the beginning of meiosis (Cabral-de-Mello et al., 2008).

The present study is the first report on the cytogenetic mapping of a transposable element in coleopteran chromosomes. The *LOA* non-LTR retrotransposon was first isolated from the genome of *Drosophila silvestris*, a species that is endemic to the Hawaiian Islands (Felger, 1992). These elements belong to evolutionarily younger clades of non-LTR retrotransposons (Volff et al., 2004), contain very few known elements, and have mostly been identified in *Drosophila*, *Aedes* and *Ciona* genomes (Rho and Tang, 2009).

The distribution of *LOA-like* elements in the chromosomes reinforces an evolutionary relationship between the A complement and the B chromosome at least in the pericentromeric area. Recent work involving the centromere-enriched retrotransposons indicates that these elements preferentially insert into the centromeric regions (Birchler and Presting, 2012). The *LOA-like* elements may have been maintained in the genome of *C. cyanescens* due to a possible functional role they play in the maintenance of the pericentromeric regions. The absence of *LOA-like* elements in the sex chromosomes suggests that sex differentiation occurs before the distribution of this transposable element into the genome. Subsequently, the suppression of recombination could have produced the differences observed in the distribution of TEs between

the A complement and the sex chromosomes. These results suggest that *LOA-like* element could be involved in the maintenance of the pericentromeric regions and might contribute to the origin of the B chromosome.

Conclusions

The results obtained by the hybridization of *Cot*-1 DNA and the *LOA-like* non-LTR retrotransposon indicate that the origin of the B chromosome is associated with autosomal elements. The present study is the first report on the cytogenetic mapping of a transposable element in coleopteran chromosomes. Our work further suggests that TEs could also have been involved in the origin and evolution of the B chromosome in *C. cyanescens*.

Materials and methods

Animal sampling and cytogenetic analysis

Fifty adult specimens of *Coprophanaeus cyanescens* (Olivier, 1789) (Coleoptera: Scarabaeidae: Scarabaeinae: Phanaeini) were obtained from Parque João Vasconcelos Sobrinho, Caruaru, Pernambuco State, Brazil. The specimens were collected in the wild according to Brazilian laws for environmental protection (wild collection permit, MMA/IBAMA/SISBIO n°. 2376-1). The experimental research on animals was conducted according to the international guidelines followed by São Paulo State University (Protocol no. 35/08 – CEEA/IBB/UNESP).

The testes were fixed in Carnoy solution (3:1 ethanol: acetic acid) and later stored at -20°C. The chromosome preparations were obtained by using the classical testicular follicle squashing technique. Conventional chromosome analysis was performed after staining the slides with 5% Giemsa.

Chromosomal probe isolation

The DNA samples were obtained from frozen tissues collected from specimens. The procedure for extraction of genomic DNA followed the protocol previously described (Sambrook and Russel, 2001) with minor modifications. The quality and quantity of purified DNA was evaluated in 0.8% agarose gel and spectrophotometry.

Three sets of DNA sequences were used as probes for fluorescence *in situ* hybridization (FISH) as follow: (i) sequences for the 18S and 5S rRNA genes were obtained from cloned

sequences of the dung beetle, *Dichotomius semisquamosus* (Cabral-de-Mello et al., 2010a); (ii) sequences of the *LOA-like* non-LTR retrotransposon were obtained from *C. cyanescens* by PCR with the RF-Co (5' CGC CTA CTT CAG GAC CAG AG 3') and RR-Co (5' AGA CTG CAG GCC GTA GAA AA 3') primers (Burke et al., 1993); (iii) *Cot-1* DNA sequences were isolated from *C. cyanescens* based on the DNA re-association kinetics (Zwick et al., 1997) with modifications (Ferreira and Martins, 2008).

PCR products from the non-LTR retrotransposons were inserted into the pGEM-T plasmid (Promega) according to the manufacturer's recommendations, and the recombinant plasmids were used to transform competent *Escherichia coli* cells (Invitrogen, San Diego, CA, USA). The presence of the inserts in the recombinant plasmids was analyzed by PCR, and the recombinant clones were stored at -80°C. The recombinant plasmids were subjected to nucleotide sequencing using an Applied Biosystems sequencer (3500 Genetic Analyzer).

Analysis of transposable elements

The *LOA-like* non-LTR retrotransposon sequences isolated by PCR from *C. cyanescens* were used as queries to detect related TEs in other genomes available from the Repbase (<http://www.girinst.org/rebase/>) and NCBI (National Center for Biotechnology Information - <http://www.ncbi.nlm.nih.gov/>) databases. The search included whole genome shotgun contigs, nucleotide collections, and high throughput genomic sequences. Analysis of the recovered DNA sequences was performed with the LIRMM software (Laboratoire Le d'Informatique, Robotique et de Microélectronique of Montpellier) available online (<http://www.phylogeny.fr/>) (Guindon and Gascuel, 2003; Chevenet et al., 2006; Dereeper et al., 2008).

Fluorescence *in situ* hybridization

The DNA probes were labeled by nick translation with biotin-11-dATP (Invitrogen) or digoxigenin-11-dUTP (Roche, Mannheim, Germany) by PCR. The FISH technique was performed according to a protocol adapted for Coleoptera [22]. The chromosome spreads were counterstained with DAPI (4', 6-diamidino-2-phenylindole), and the slides were mounted in Vectashield mounting medium (Vector, Burlingame, CA, USA). The images were captured using an Olympus DP71 digital camera coupled to a BX61 Olympus microscope and were optimized for brightness and contrast using Adobe Photoshop CS2 and Corel Photo-Paint 13.

List of abbreviations

CEEA: Comissão de Ética em Experimentação Animal

DAPI: 4', 6-diamidino-2-phenylindole

FISH: fluorescence *in situ* hybridization

IBAMA: Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis

IBB: Instituto de Biociências de Botucatu

LIRMM: Laboratoire Le d'Informatique, Robotique et de Microélectronique of Montpellier

LTR: long terminal repeat

MMA: Ministério do Meio Ambiente

NCBI: National Center for Biotechnology Information

PCR: polymerase chain reaction

rDNA: ribosomal DNA

rRNA: ribosomal RNA

SISBIO: Sistema de Autorização e Informação em Biodiversidade

UNESP: Universidade Estadual Paulista “Júlio de Mesquita Filho”

TE(s): transposable element(s)

Competing interests

The authors declare that they have no competing interests.

Author contributions

SGO, RCM and CM contributed to the development of the hypothesis, specimen collection and preparation, and analysis and interpretation of data. SGO and CM drafted the first version of the manuscript. RCM revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors are grateful to CMQ Costa and FAB Silva for the taxonomic identification of the studied species. The study was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), the Conselho Nacional de Desenvolvimento Científico e Tecnológico

(CNPq) and the Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) of Brazil.

References

As referências citadas neste e nos demais capítulos encontram-se no item 5 (Referências Bibliográficas).

Figure legend

Figure 1: Metaphase I stages of *Coprophanaeus cyanescens* carrying 1 B chromosome. Conventional staining (a), FISH mapping of *Cot*-1 DNA (b), 18S (green) and 5S (red) rRNA genes (c) and *LOA-like* non-LTR retrotransposon (d). The B chromosome and the XY sex chromosomes are indicated. Bar = 5 μ m.

Additional file

File name: Additional file 1

File format: PDF

Title of data: Alignment of the *LOA* non-LTR retrotransposon nucleotide sequences from *Nasoniavitripennis*(Baggins-1_NVi) and *Coprophanaeus cyanescens* (Cc-1 to Cc-3). The asterisks (*) indicate similarity in sequence, and the dashes (-) indicate indels.

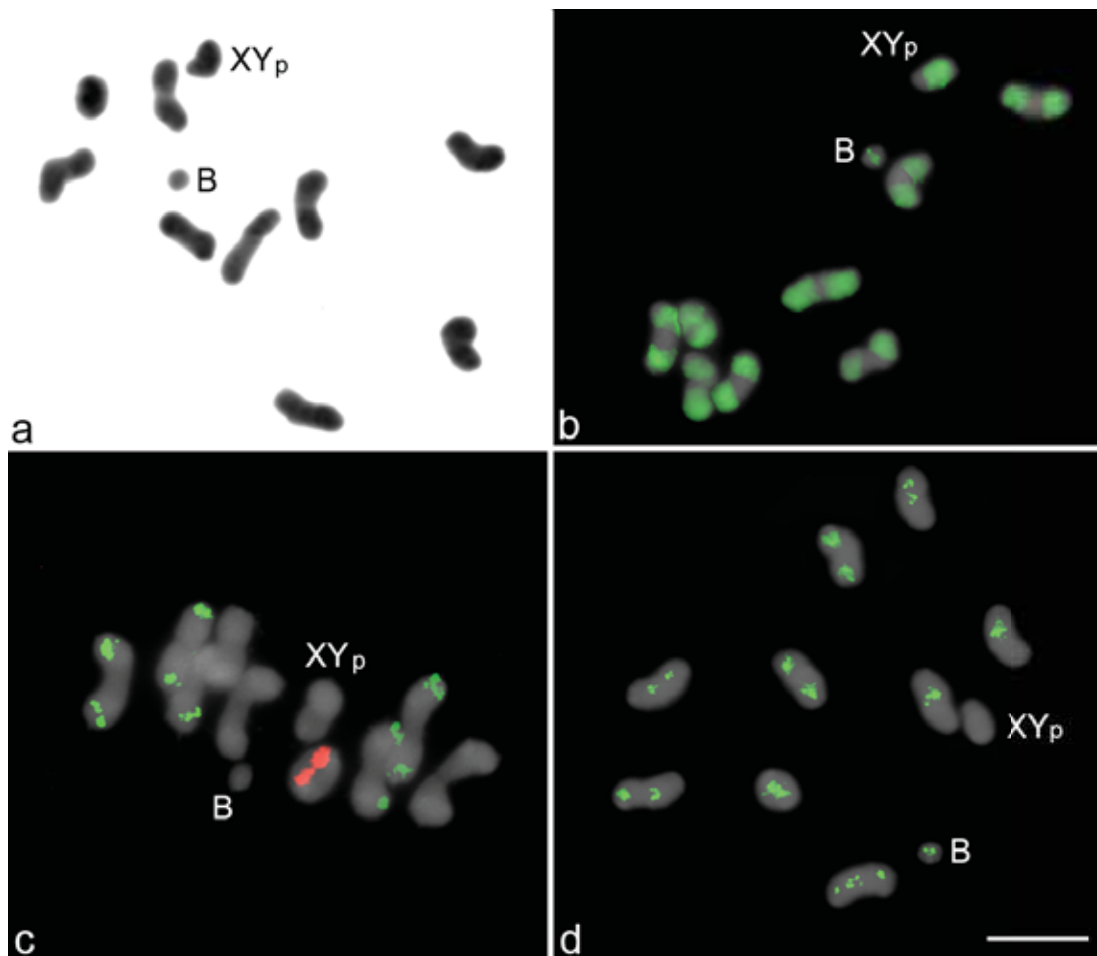


Figure 1

Additional file 1: Alignment of the LOA non-LTR retrotransposon nucleotide sequences from *Nasonia vitripennis* (Baggins-1_NVi) and *Coprophanaeus cyaneescens* (Cc-1 to Cc-3). The asterisks (*) indicate similarity in sequence, and the dashes (-) indicate indels.

3.3. Capítulo 3:

***Mariner* transposable elements in Scarabaeinae coleopterans and their relationship to other *Mariner* families**

Sárah G. Oliveira¹, Diogo C. Cabral-de-Mello², Rita C. Moura³ and Cesar Martins^{1*}

¹UNESP - São Paulo State University, Bioscience Institute, Morphology Department, Botucatu, SP, 18618-970, Brazil.

²UNESP - São Paulo State University, Bioscience Institute, Biology Department, Rio Claro, SP, 13506-900, Brazil.

³UPE - Pernambuco State University, Biological Sciences Institute, Department of Biology, Recife, PE, 50100-130, Brazil.

Short running title: Mariner transposable elements in Scarabaeinae coleopterans

*Corresponding author

Email: cmartins@ibb.unesp.br

Abstract

Mapping repetitive DNA sequences has improved the knowledge of genome organization and chromosomal differentiation during the evolutionary history of species. For Coleoptera genomes, the organization of repetitive DNAs has been poorly investigated, with only one cytogenetic mapping study involving transposable elements. With the aim of increasing the knowledge on the evolution of coleopteran genomes, we investigated *Mariner* transposons in three Scarabaeinae species (*Coprophanaeus cyanescens*, *C. ensifer* and *Diabroctis mimas*) through cytogenetics and nucleotide sequence analysis. The cytogenetic mapping revealed an accumulation of *Mariner* transposons in pericentromeric repetitive-rich regions characterized through heterochromatin and *Cot*-1 DNA analysis. Nucleotide sequence analysis of *Mariner* revealed the presence of two major groups of *Mariner* copies in the three investigated species and other insects, indicating that *Mariner* lineages had an early origin in the base of insect diversification. Comparative analysis detected low levels of similarity in the *Mariner* transposases of the coleopteran species, indicating high diversification of these sequences during the evolutionary history of the group. Furthermore, comparisons of the coleopterans sequences with other insects and mammals suggested that horizontal transfer could have acted in spreading *Mariner* in diverse non-related animal groups.

Key Words: Chromosomal rearrangements, Evolution, Heterochromatin, Horizontal transfer, Repetitive DNA, Transposition

Introduction

Repetitive DNAs represent a significant fraction of eukaryotic genomes and are primarily enriched in heterochromatic regions, although some of them have been observed in euchromatic regions (Dimitri and Junakovic, 1999; Torti et al. 2000; Pikaard and Pontes, 2007; Montiel et al. 2012). Among repeated DNAs, transposable elements (TEs) are DNA sequences capable of changing their location in the genome by moving from one site to another, which seems to benefit only the elements, and for a long time, they were considered "parasitic" and/or "selfish" elements. However, TEs represent an evolutionary force that provides potential conditions for the emergence of new genes, the modification of gene expression, and adaptation to new environmental challenges (Kazazian, 2004; Feschotte, 2008; Pritham, 2009). In this way, TEs play a major role in shaping and influencing the structure and function of genomes (Walisko et al., 2009).

Among several groups of TEs, *Mariner*-like elements (MLEs) are a superfamily of DNA transposons that consist of a single gene without introns flanked by two terminal inverted repeats (TIRs) of approximately 30 bp, with a total length of approximately 1,300 bp. Each terminal repeat is flanked by a TA dinucleotide, which results from target site duplication (TSD) (Robertson, 1995; Wicker et al., 2007). The *Mariner* transposase gene encodes a protein of 330-360 amino acids, which recognizes TIRs and cuts both strands at each end, and it is responsible for transposition through excising, exchanging, and fusing DNAs in a coordinated manner (Lampe et al., 1996; Plasterk, 1996).

The *Mariner* superfamily is most likely the most widespread and diverse group of TEs found in animals, persisting in genomes throughout evolutionary time (Hartl et al., 1997a). The MLE history began with their discovery in insects (observed in several orders), and their distribution has now been reported in multiple invertebrate and vertebrate genomes (e.g., Robertson, 1993; Robertson, 1995; Miskey et al., 2005; Feschotte and Pritham, 2007). The high similarity between sequences from distantly related organisms, the incongruence between TEs and phylogeny, and the unequal distribution of some *Mariner* subfamilies among closely related taxa indicate that horizontal transfer (HT) contributed to this widespread distribution (e.g., Robertson and Lampe, 1995b; Ray et al., 2008; Oliveira et al., 2012c; Sormacheva et al., 2012).

The cytogenetic mapping of repetitive DNAs has improved the knowledge of genome organization and chromosomal differentiation during the evolutionary history of species. In contrast, the genome organization of repetitive DNAs has been poorly investigated in Coleoptera, with only one study involving the cytogenetic mapping of TEs (Oliveira et al., 2012b). With the aim of contributing to the knowledge of coleopteran genome evolution at the molecular and chromosomal level, we investigated the repetitive DNA fraction of three Scarabaeinae species, which are characterized by the presence of extensive blocks of heterochromatin (Oliveira et al. 2010, 2012a; Bione et al. 2005a). The overall chromosomal distribution of repetitive DNAs was investigated through the chromosomal hybridization of the *C₀t*-1 DNA fraction (DNA enriched with highly and moderately repeated sequences), and the genomic features of *Mariner* TEs were addressed through nucleotide sequencing and chromosomal mapping.

Materials and methods

Animals and DNA samples

Adult male samples of *Coprophanaeus cyanescens*, *C. ensifer* and *Diabroctis mimas* were collected in Caruaru, Igarassu, Paudalho and Saloá, Pernambuco State, Brazil. The animals were collected in the wild according to Brazilian laws for environmental protection (wild collection permit, MMA/IBAMA/SISBIO n. 2376-1). The experimental research on animals was conducted according to the international guidelines followed by Sao Paulo State University (Protocol no. 35/08 – CEEA/IBB/UNESP). The testes were fixed in Carnoy solution (3:1 ethanol: acetic acid) and stored in a freezer at -20°C. The DNA samples were obtained from living specimens and immediately frozen in the freezer at -20°C. The procedure for the extraction of genomic DNA followed the protocol described by Sambrook and Russel (2001), with minor modifications. The quality and quantity of purified DNA were analyzed using electrophoresis and spectrophotometry.

Obtaining *C₀t*-1 DNA

C₀t-1 DNA fractions were obtained from *C. cyanescens*, *C. ensifer* and *D. mimas* based on the reassociation kinetics proposed by Zwick et al. (1997) and the modifications described by Ferreira and Martins (2008). The *C₀t*-1 DNA fractions obtained were labeled and used directly as

probes for chromosome hybridization; hybridizations were performed within the same species and between different species.

Isolation and characterization of *Mariner* TEs

The *Mariner* transposable elements were isolated through polymerase chain reaction (PCR) with the set of primers MAR-188F (5' ATC TGR AGC TAT AAA TCA CT) and MAR-251R (5' CAA AGA TGT CCT TGG GTG TG), which were designed based on conserved regions of the amino acid sequence of the putative transposase gene of the *Mariner* element (Lampe et al., 2003). The PCR products were cloned using the pGEM-T kit (Promega, Madison, WI, USA) according to the manufacturer's recommendations. The recombinant plasmids were submitted to nucleotide sequencing using a 3500 Genetic Analyzer sequencer (Applied Biosystems, Foster City, CA, USA).

The DNA sequences obtained were used as an initial query to search against a database of repetitive DNA elements (Repbse database) (<http://www.girinst.org/repbse/>), which contains repetitive DNA sequences of various eukaryotic species (Jurka et al., 2005). Additionally, the obtained sequences were analyzed against the nucleotide collection of The National Center for Biotechnology (NCBI) (<http://www.ncbi.nlm.nih.gov>) using the Blast search tool. Family consensus sequences were constructed whenever possible. DNA sequences were analyzed with the software LIRMM (Laboratoire Le d'Informatique, Robotique et de Microélectronique of Montpellier), which is available online at <http://www.phylogeny.fr/> (Guindon and Gascuel 2003; Chevenet et al., 2006; Dereeper et al., 2008). Phylogenetic trees were built with neighbor joining (NJ) and confirmed as consistent with trees built by PhyML (Guindon et al. 2010). The *Mariner* protein-coding sequences were searched against the Pfam database (<http://www.pfam.sanger.ac.uk>).

Chromosome isolation and fluorescence *in situ* hybridization (FISH)

Meiotic chromosomes for FISH were obtained from testes of *C. cyanescens*, *C. ensifer* and *D. mimas* according to the classic technique of squashing the testicular follicles in a drop of 45% acetic acid, following by dipping in liquid nitrogen to remove the coverslip.

The PCR products, which contained a pool of *Mariner* sequences, and the *Cot*-1 DNA fraction were labeled with biotin-11-dATP by nick translation using the Bionick Labeling System

kit (Invitrogen, San Diego, CA, USA). The FISH protocol was performed according to Pinket et al. (1986) using the adaptations described by Cabral-de-Mello et al. (2010c). The probes were labeled with biotin-14-dATP and detected by avidin-FITC (fluorescein isothiocyanate) (Sigma-Aldrich, St. Louis, MO, USA). The chromosomes were counterstained with 4,6-diamidino-2-phenylindole (DAPI), and the slides were mounted with Vectashield (Vector, Burlingame, CA, USA). The images were captured using an Olympus DP71 digital camera coupled to a BX61 Olympus microscope and DP Control program and processed using Corel Photo-Paint 12 and Adobe Photoshop CS2.

Results

Cytogenetic mapping of *C₀t*-1 DNA and *Mariner*

The *C₀t*-1 DNA fraction was isolated from the genomes of *C. cyanescens*, *C. ensifer* and *D. mimas* and hybridized to their respective chromosomes (Figure 1a-c). This *C₀t*-1 DNA revealed positive hybridization in the long arms of all autosomes, the X chromosome of the two *Coprophanaeus* species and the long arm of the Y chromosome of *C. ensifer* (Figure 1a, b). The *C₀t*-1 DNA fraction hybridization in *D. mimas* showed large pericentromeric blocks in all autosomal pairs and the sex chromosomes, extending to the short arms of some autosomal chromosomes (never to the long arm), including the terminal region, and in the terminal region of one autosomal pair (Figure 1c). Cross-species hybridization of the *C₀t*-1 DNA fraction showed positive hybridization only among species of the same genus, and the same pattern was observed for the hybridization of probes that originated from the same genome (as shown in Supplement S1). These patterns of *C₀t*-1 DNA hybridization were similar to the data that were previously generated by C-banding in the three species (Bione et al., 2005a; Oliveira et al., 2010). However, *C₀t*-1 DNA hybridization was not observed in the C-positive banded centromeric region of the Y chromosome of *C. cyanescens* or the interstitial blocks of the short arms of three autosomal pairs or the telomeric block in a small autosomal pair observed in *C. ensifer* (Oliveira et al., 2010). Additionally, *C₀t*-1 DNA blocks observed in the terminal region of an autosomal pair and the centromeric region of the Y chromosome of *D. mimas* were not observed by C-banding (Bione et al., 2005a).

FISH using probes of *Mariner* sequences in *C. cyanescens* labeled the X and Y sex chromosomes and four autosomal pairs with large pericentromeric blocks and three large

autosomal pairs with pericentromeric labeling that extended to the long arm (Figure 1d). In *C. ensifer*, *Mariner* mapping revealed small pericentromeric blocks in the X and all autosomal chromosomes (Figure 1e). In *D. mimas*, the pattern was similar to the pattern obtained by *Cot*-1 DNA hybridization; however, the blocks observed were smaller (Figure 1f).

Analysis of *Mariner* sequences

Nucleotide transposase (partial DDE domain – region with three conserved amino acids: asparagine-asparagine-glutamine) sequences of approximately 230 bp were obtained for *C. cyanescens* (eight sequences), *C. ensifer* (six sequences) and *D. mimas* (nine sequences) (Supplement S2). The comparative analysis of *Mariner* with several vertebrates and invertebrates showed that the *Mariner* sequences were organized into two major groups (I and II), and group I was subdivided into 3 branches (Figure 2). In the first branch, only insect sequences were observed, represented by flies (*Drosophila ficusphila*), ants (*Harpegnathos saltator*), bees (*Apis florea* and *Apis mellifera*), earwig (*Forficula auricularia*) and beetles (*C. cyanescens*, *C. ensifer*, and *D. mimas*). In the second branch, mammalian (*Erinaceus europaeus* and *Tupaia belangeri*), planaria (*Schmidtea mediterranea*) and insect sequences, represented by ants (*Atta cephalotes*, *Harpegnathos saltator*, *Linepithema humile*, and *Solenopsis invicta*), bees (*Megachile rotundata*) and beetles (*C. ensifer* and *D. mimas*), were observed. In the third branch, only one sequence from a fly was observed (*Chymomyza amoena*). The sequences of group II were organized in one major branch and contained mammalian (*Bos taurus*) and insect sequences, represented by flies (*Drosophila elegans* and *Drosophila erecta*) and ants (*Acromyrmex echinator*, *Atta cephalotes*, *Camponotus floridanus*, *Harpegnathos saltator*, and *Solenopsis invicta*).

In the first branch (group I sequences), the sequences of beetles showed similarity with other insect sequences, although they formed a separate branch with a high branch support value (0.87). In the second branch, the insect sequences were related to the *Mariner1_Tbel* family in mammals. Even including the beetle sequences, a high sequence similarity of the mammals *E. europaeus* and *T. belangeri* with the genome of the ants *Pogonomyrmex barbatus* and *H. saltator* was clear (higher than 92% compared with *E. europaeus* and higher than 98% compared with *T. belangeri*) (Supplement S3). The genetic distance within *Mariner1_Tbel* sequences between mammal and insect (including beetles) species was relatively low (0.017-0.359%) (Supplement

S3). In turn, the genetic distances observed within *Mariner-1_BT* sequences among insect species were 0.258-0.525% (Supplement S3).

The *Mariner* sequences of beetles branched out into two groups (seen in the first and second branches of group I). Within the first branch, sequences from *C. cyanescens*, *C. ensifer* and *D. mimas* were observed. However, within the second branch, only sequences from the species *C. ensifer* and *D. mimas* were observed. The genetic distance between these two beetle groups was relatively high (0.338-0.518).

Discussion

General aspects of heterochromatin and repeated DNA organization

The presence of large blocks of heterochromatin and the C_{0t} -1 DNA fraction in the studied species suggests the amplification of repetitive DNAs and/or heterochromatin transfer between chromosomes during the karyotype differentiation of species, as previously observed in other animals (Schweizer and Loidl, 1987; Sumner, 2003; Loreto et al., 2005; Cabral-de-Mello et al., 2010a). This statement is supported by the common pattern of heterochromatic blocks that are primarily located in pericentromeric areas in related species, including coleopterans (Juan and Petitpierre, 1989; Rozek et al., 2004; Angus et al., 2007). Most information concerning heterochromatin in coleopterans is focused on the description of chromosomal distribution, with few data regarding its molecular content. C_{0t} -1 DNA fraction hybridization showed a general pattern coinciding with the data generated by C-banding (Bione et al., 2005a; Oliveira et al., 2010), indicating that heterochromatin is enriched in highly repetitive DNA. The presence of large blocks of C_{0t} -1 suggests an abundance of repetitive sequences, and cross-species hybridization analysis among Phanaeini species shows high conservation between the fractions of repetitive DNA within genera and divergence between the two different genera. However, the use of C_{0t} -1 DNA fractions as probes in *Dichotomius* species (Coleoptera, Scarabaeidae) allowed for the observation of heterochromatin distribution patterns that were highly conserved in the terminal/sub-terminal region but had extensive variation in relation to pericentromeric heterochromatin (Cabral-de-Mello et al., 2011b), which contrasts with the Phanaeini species studied. These data reinforce the intense evolutionary dynamics of the repeated DNA fraction through mutation, gene conversion, unequal crossing over, circular replication and slippage replication (Charlesworth et al., 1994; Li et al., 2002; Shapiro, 2005), generating high divergence

among taxa above the genus level. The processes of heterochromatin spreading and accumulation could be favored by the presence of chromocenters during prophase, as previously observed in *C. cyanescens* and *D. mimas* (Bione et al., 2005a; Cabral-de-Mello et al., 2011a). In coleopteran meiosis, the association of non-homologous heterochromatin forming chromocenter fusions involving various chromosomes appears to be common (e.g., Smith and Virkki, 1978; Cabral-de-Mello et al., 2011a).

Chromosomal organization of *Mariner* transposable elements

It is a common observation that some transposable elements may be overabundant in specific regions of chromosomes, and the results obtained with mapping *Mariner* showed that these sequences were not randomly distributed but accumulated in heterochromatic areas. However, the accumulation of this element in euchromatic areas was recently reported in *Eyprepocnemis plorans* (Montiel et al. 2012). The accumulation of a large amount of copies in heterochromatic regions can indicate selection against insertions of TEs in euchromatin based on ectopic exchanges. Different major forces can affect TEs in heterochromatic and euchromatic regions of the genome, as accumulation in heterochromatic regions is explained by the absence of selection against insertional mutations in genetically inert regions and the stochastic accumulation of deleterious elements in regions with no recombination (Dimitri and Junakovik, 1999; Torti et al., 2000; Kaminker et al., 2002).

The accumulation of *Mariner* sequences in pericentromeric regions is possibly due to the low rate of recombination that is characteristic of these regions and could indicate that this element is enriched in regions in which the damage of its insertion is reduced (Gray, 2000; Delaurière et al., 2009). Although it is not possible to predict the possible role of these elements in Coleoptera, they may be involved in chromosomal rearrangements, as evidenced by the pericentromeric inversion observed in *D. mimas*. This species presents meta-submetacentric (pairs 1, 2, 3 and 7) and acrocentric (pairs 5, 6, 8 and 9) autosomal chromosomes (Bione et al., 2005a), while *C. cyanescens* and *C. ensifer* have meta-submetacentric morphology for all autosomal chromosomes (Oliveira et al., 2010). In *D. mimas*, the presence of four acrocentric autosomal pairs indicates the occurrence of pericentromeric inversions, unlike the standard meta-submetacentric karyotype described for the family Scarabaeidae (Cabral-de-Mello et al., 2008). Chromosomal rearrangements, such as that observed in *D. mimas*, are possible because

transposable elements have been reported to be involved in various types of rearrangements through transposition and recombination (Deininger et al. 2003; Chen et al. 2005; Slotkin and Martienssen, 2007).

Another approach to explaining the accumulation of transposable elements is that the *Mariner* transposon could have been maintained in the pericentromeric region by presenting a functional role in the maintenance of this region (Dimitri et al., 1997). For example, during the evolution of the genome, heterochromatic transposable elements may lose the ability to transpose and accumulate mutations and structural rearrangements, acquiring new functions (von Stenberg et al., 1992; McDonald, 1993). Feschotte (2008) proposed that the movement and accumulation of TEs and their derived proteins have played an important role in the evolution of the genome. The association of TEs with the structure and/or function of centromeres seems to be a usual occurrence and has been observed in diverse species (e.g., Dawe 2003; Wong and Choo 2004). Proteins from TEs with centromeric function have been observed in yeast and mammals through a process of convergent domestication of the *Pogo*-like transposase, which has adapted to give rise to proteins with centromere-binding activity (Casola et al., 2008).

The mapping of *Mariner* in the sex chromosomes of the three species could be related to the common spreading of the TEs in most heterochromatic areas of the genome or because sex chromosomes can act as a refuge for transposable elements, as previously reported (Steinemann and Steinemann 1992; Mandrioli, 2003). Several genetic processes can cause an accumulation of TEs in genomic regions in which crossing over is reduced or absent (Charlesworth et al, 1994). In some cases, for example, the sex chromosomes show a tendency of non-recombining in genomic regions to accumulate transposable elements (Bachtrog, 2005; Charlesworth et al., 2005). Another possibility is that recombination suppression itself could inhibit recombination in nearby regions of the sex chromosomes (Charlesworth et al., 2005).

The transposition/selection model establishes that the distribution and abundance of TEs are indicative of their evolutionary history (Langley et al., 1988; Charlesworth et al., 1994). The stages of this process are the following: (i) invasion of the host genome, (ii) rapid spread by replicative transposition, and (iii) vertical inactivation and accumulation in heterochromatin. Considering that hypothesis, the *Mariner* present in *C. cyanescens*, *C. ensifer* and *D. mimas* could be considered ancient because active and recently acquired elements are expected to be preferentially located in euchromatin. TEs are expected to be overabundant in heterochromatin,

where recombination is strongly reduced, because the TEs cannot be easily removed from heterochromatin once they have been inserted (Junakovic et al., 1998; Gomulski et al. 2004).

Mariner transposable elements in Scarabaeinae coleopterans

The *Mariner* sequences of Scarabaeinae coleopterans branched out into two groups, which showed a relatively high genetic distance, indicating early divergence from an ancestral element. This finding is consistent with previous studies, which have proposed that members of the *Tc1/Mariner* family are most likely monophyletic in origin and diversified in various groups through the accumulation of modifications and/or horizontal transfer mechanisms (Robertson, 1995; Capy et al., 1996; Plasterk et al., 1999).

Most likely, beetle *Mariner* copies have evolved independently of each other, according to the pattern of molecular evolution related to *Mariner* transposons. When divergent elements do exist, they display, as observed, a low percentage of similarity to the full-length sequences. This finding suggests that TEs are highly active within the genome and that the highly divergent copies reflect relics of ancient mobilizations, as described in *D. melanogaster* (Lerat et al., 2003).

The relatively high genetic distances observed between the two classes of *Mariner* sequences in beetles and their distribution in other animals indicate two possibilities: i) these two classes had an early origin in the base of insect diversification, or ii) considering the highest similarity for each group with other insect sequences (particularly ants) and similarity to mammals in one of the groups, there is evidence that the sequences may have originated by HT.

Mariner horizontal transfer

Mariner transposable elements have been described in many arthropods and were possibly spread by HT (Robertson and Lampe, 1995a, b; Lampe et al., 2003). In general, the phylogenies based on *Mariner* sequences are not always congruent with the phylogenies of the taxa, suggesting HT (Robertson, 1993; Robertson and MacLeod, 1993).

The high sequence similarity between sequences from distantly related organisms, the incongruence between TE distribution and phylogeny, and the unequal distribution of some *Mariner* subfamilies among closely related taxa indicate that HT contributed to this widespread distribution (e.g., Robertson and Zuppano, 1997; Ray et al., 2008; Sormacheva et al., 2012). Several TEs have been introduced into mammalian lineages through HT (e.g., Kordis and

Gubensek, 1998; Piskurek and Okada, 2007; Gilbert et al., 2009; Schaack et al., 2010), including *Mariner* (Pace et al., 2008; Oliveira et al., 2012c). Comparative analyses of mammalian genomes show the presence of high amounts of TEs, but their content could vary among different lineages (Böhne et al., 2008; Ray et al., 2008). The genetic distance within *Mariner1_Tbel* sequences between mammals and insects species was relatively low, which is consistent with their phylogenetic distances and reinforces the role of HT in the spread of these elements to different taxa (Oliveira et al., 2012c).

The *Mariner* tree topology clearly indicates the involvement of HT during the evolutionary history of insects and mammals, although it is not possible to show at which evolutionary moment this transfer occurred. Multiple mechanisms may be related to the spread of TE by horizontal transfer through different types of vectors (external parasites, infectious agents, intracellular parasites and symbionts, DNA viruses, RNA viruses, and retroviruses) (Kidwell, 1993; Hartl et al., 1997a; Silva and Kidwell, 2000). Thus, a model of transfer could be implemented for each described case of a proposed HT. Our results are consistent with the criteria of HT and reveal interesting patterns of patchy distribution among animals, suggesting the repeated invasion of *Mariner* from insects to mammals.

Acknowledgments

The authors are grateful to CMQ Costa and FAB Silva for the taxonomic identification of the species studied. The study was supported by the Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP), and the Fundação de Amparo a Ciência e Tecnologia do Estado de Pernambuco (FACEPE).

References

As referências citadas neste e nos demais capítulos encontram-se no item 5 (Referências Bibliográficas).

Figure captions

Figure 1: Metaphase I of *Coprophanæus cyaneus* (a, d), *C. ensifer* (b, e) and *Diabroctis mimas* (c, f) analyzed with FISH with *Cot*-1 DNA (a-c) and *Mariner* transposable element (d-f) probes. Bar = 5 µm.

Figure 2: Alignment guide tree of *Mariner* families performed with the software LIRMM. The sequences obtained for beetles in this work are indicated in green. Except for individual sequences with accession numbers provided, all other sequences are represented by consensus sequences deposited in Repbase with their Repbase ID. The sequences are indicated in different colors based on whether they are derived from mammals (red), planaria (pink), beetles (green) or other insects (blue). The species are Cc (*Coprophanæus cyaneus*), Ce (*C. ensifer*), Dm (*Diabroctis mimas*), Ac (*Atta cephalotes*), Ae (*Acromyrmex echinator*), Af (*Apis florea*), Am (*Apis mellifera*), Bte (*Bombus terrestris*), Bt (*Bos taurus*), Ca (*Chymomyza amoena*), Cf (*Camponotus floridanus*), Del (*Drosophila elegans*), Der (*Drosophila erecta*), Df (*Drosophila ficusphila*), Ee (*Erinaceus europaeus*), Fa (*Forficula auricularia*), Hs (*Harpegnathos saltator*), Lh (*Linepithema humile*), Mr (*Megachile rotundata*), Pb (*Pogonomyrmex barbatus*), Si (*Solenopsis invicta*), Sm (*Schmidtea mediterranea*) and Tb (*Tupaia belangeri*). The sequences Cc_4 to Cc_7 are smaller than 200 bp and do not have GenBank accession numbers. The *Mariner* sequences are organized into two major groups (I and II, as indicated), and the groups are subdivided into branches (colored blocks). The branch support values are indicated on the nodes. The scale bar indicates the genetic distance.

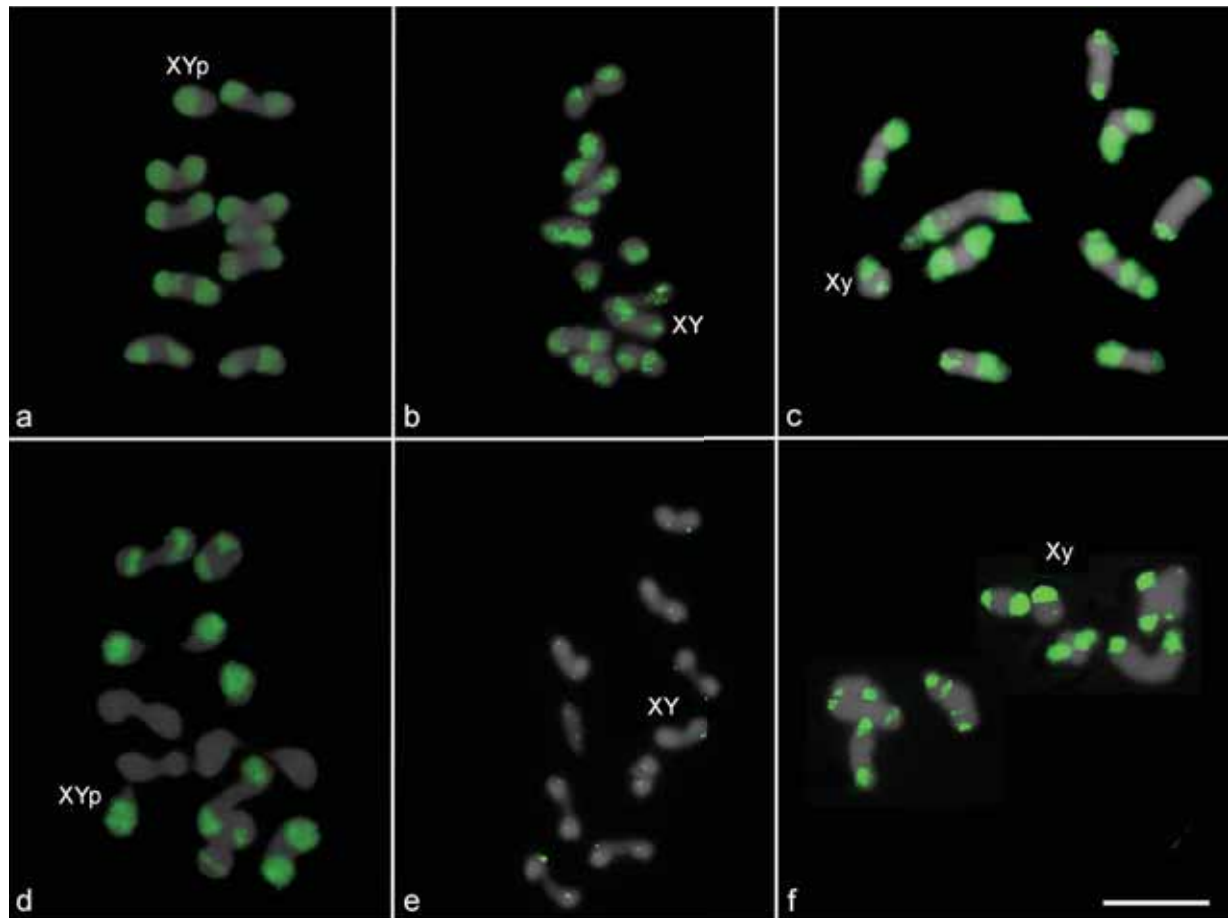


Figure 1

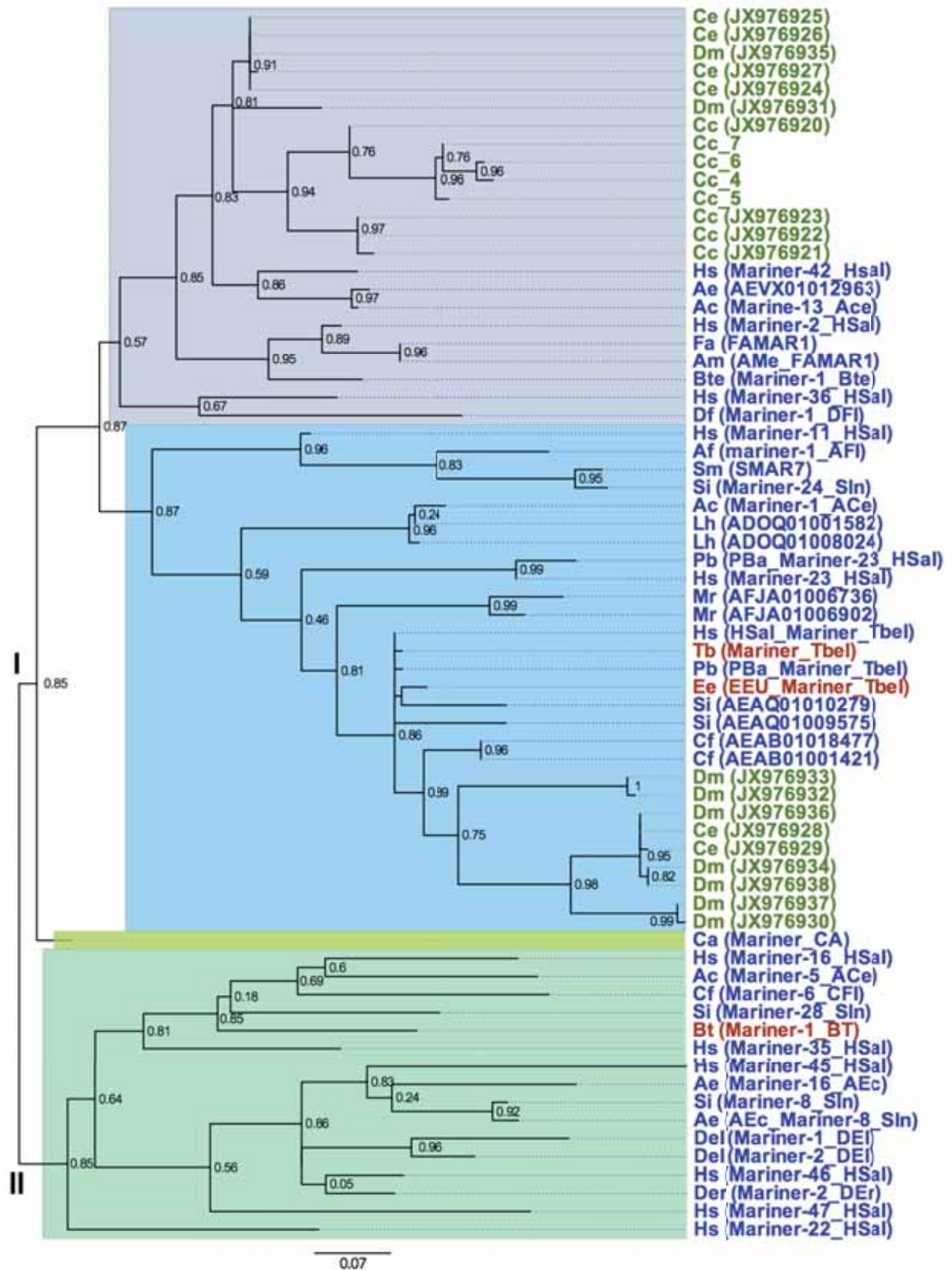


Figure 2

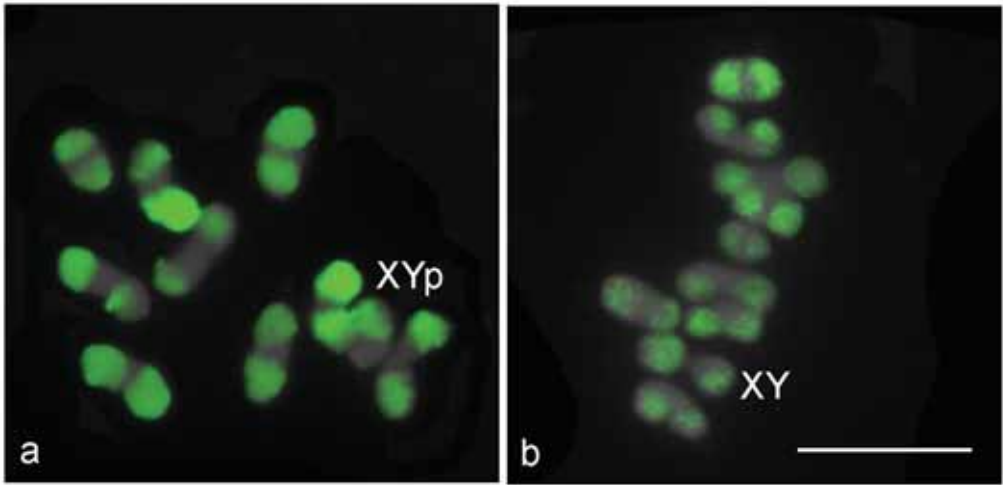
Supplementary Material

Supplementary Material S1: Cross-species hybridization of the C_{ot} -1 DNA fraction in metaphase I of *Coprophanaeus* species. Probe of *C. ensifer* hybridized to *C. cyanescens* (a) and probe of *C. cyanescens* hybridized to *C. ensifer* (b). The sex chromosome bivalents are indicated. Bar = 5 μ m.

Supplementary Material S2: Sequence alignment of related *Mariner* families of diverse organisms retrieved from public databases and sequences obtained in the present work. The abbreviations correspond to species names and IDs as described in the caption of Figure 2. Dashes represent indels.

Supplementary Material S3: Pairwise similarity alignment matrix of related *Mariner* families. The abbreviations correspond to the species names and IDs as described in the caption of Figure 2.

Supplementary Material S1



Supplementary Material S2

Hs (Mariner-24_SIn) AAACAACCTCGAAGGCGGGATTGCATCCGAAGAAAATAAT-----GCTCTCCGTATGGT
 Hs (Mariner-45_HSal) AGTCTGTGGCAAAGGCCACACTGACGCCAGGAAGGTTAT-----GCTGTGTGTTTGGT
 Cf (Mariner-6_CFl) TAACCACTTCGAAGGCAGGCTTACATCCGAAAAAGGTAAT-----GCTCTGCATTTGGT
 Hs (Mariner-35_HSal) AAACCACACCGAAGCAAAACATCCATACGAAAAAAGTCAT-----GCTGTGTATTTGGT
 Df (Mariner-1-DF) AAGCCACATCAAAGGCCGATATCCATCAGAGGAAAGTTAT-----GCTGTCTGTTTGGT
 Ac (Mariner-1_Ace) AAAGGCAAGCAAAGGCCGAGATCCACCAAAAGAAGGTAAT-----GCTGTCAATGTGGT
 Lh (ADOQ01008024) AAAGGCAAGCAAAGGCCGACATCCACCAAAAGAAGGTGAT-----GCTGTCAATCTGGT
 Lh (ADOQ01001582) ---CACCGAAAGGCCGACATCCACCAAAAGAAGGTGAT-----GCTGTCAATCTGGT
 Dm (JX976930) -----TATCAAAAGCTGGTATTCATCATAGGAAGTTTAT-----GCTCTCAGTGTGGT
 Dm (JX976937) -----TATCAAAAGCTGATATTCATCATAGGAAGTTTAT-----GCTCTCAGTGTGGT
 Dm (JX976938) -----AAAGCTGATATTCATCATAGAAAGTTTAA-----GCTCTCACTGTGGT
 Ce (JX976929) -----AAAGCTGGTATTCATCATAGAAAGTTTAT-----GCTCTCACTGTGGT
 Dm (jx976934) -----TATCAAAAGCTGGTATTCATCATAGAAAGTTTAT-----GCTCTCACTGTGGT
 Ce (JX976928) -----TATCAAAAGCTGATATTCATCATAGTAAGTTTAT-----GCTCTCACTGTGGT
 Dm (JX976936) -----AAAGCTGATATTCATCATAGAAAGTTTAT-----GCTCTCACTGTGGT
 Mr (AFJA01006902) AAACGAGATCGAAACCAGAGATCCACCAGAGGTAGATTAT-----GCTCTCTGTGTGGT
 Mr (AFJA01006736) AAACGAGATCGAAACCAGAGATCCACCACAGGAAGGTAAT-----GCTCTCTGTGTGGT
 Dm (JX976932) -----AAAGCTGGTATTCATCAGAAGAAGGTTAT-----GCTCACAGTGTAGT
 Dm (JX976933) -----TATCAAAAGCTGGTATTCATCAGAAGAAGGTTAT-----GCTCACAGTGTAGT
 Si (AEAQ01010279) ---CGACATCAAAAGCAAACATTCATCAGAGAAAAAATTATGTTGTGGCTCTCAGTGTGGT
 Si (AEAQ01009575) ---CGACATCAAAAGCAAACATTCA--GAGGAAAGTTAT-----GCTCTTAATGTGAT
 Cf (AEAB01001421) ---CAACTTCAAAAGCAGAAGTTTAT-----GCTCTCAGTGTGGT
 Cf (AEAB01018477) ---CAACTTCAAAAGCAGAAGTTTAT-----GCTCTCAGTGTGGT
 Ee (EEu_Mariner_Tbel) AAATGACATCGAAAGCAGACATTCATCAGAGGAAGGTTAT-----GCTCTCAGTGTGGT
 Pb (PBa_Mariner_Tbel) AAACGACATCAAAAGCAAACATTCATCAGAGGAAGGTTAT-----GCTCTCAGTGTGGT
 Tb (Mariner_Tbel) AAACGACATCGAAAGCAGACATTCATCAGAGGAAGGTTAT-----GCTCTCAGTGTGGT
 Hs (HSal_Mariner_Tbel) AAACGACATCGAAAGCAGACATTCATCAGAGGAAGGTTAT-----GCTCTCAGTGTGGT
 Hs (Mariner-22_HSal) AAACCACTTCGAAGGCAGGGCTTCACCCGAGCAAGATCAT-----GCTCTCAATTTGGT
 Ae (Mariner-16_AEc) AAGCAATCGCAAAACCTGGATTGACGACGAAGAAGGTGAT-----GTTGTGTGTGTGGT
 Ae (AEC_Mariner-8_SIn) AAACGACTTCAAAGCCTGGATTGACGACAAATAAGGTGAT-----GCTGTGTGTGTGGT
 Si (Mariner-8_SIn) AAACGACTTCAAAGCCTGATTGACGACAAATAAGGTGAT-----GCTGTGTGTGTGGT
 Bte (Mariner-1_Bte) AGACGACACCAAAAGCTGAGATTCACCA-AAAAAGATTAT-----GCTGTCAAGTTTGGT
 Am (AMe_FAMAR1) AAACAACATCAAAAGCTGGTATTCATCAAAAGAAGGTTTT-----GTTATCAGTTTGGT
 Fa (FAMAR1) AAACAACATCAAAAGCTGGTATTCATCAAAAGAAGGTTTT-----GTTATCAGTTTGGT
 Hs (Mariner-2_HSal) AAACAACATCCAAAGCTGATATTCATCAAAAGAAAGTTAT-----GTTATCAGTTTGGT
 Hs (Mariner-42_HSal) AAACCGTATCGAAAGCTGAATTGCATCAAAAAAAGATCAT-----GCTGTCAATTTGGT
 Ac (Mariner-13_Ace) AAACCTCATCGAAAGCCGAAGTGCATCAAAAAAAGATAAT-----GCTGTCAATTTGGT
 Ae (AEVX01012963) AAACCTCATCGAAAGCTGAAGTGCATCAAAAAAAGATAAT-----GCTGTCAATTTGGT
 Cc_4 -----AAAAGCTGGTATTCATC-AAAAAGATCAT-----GTTGTCAATTTGGT
 Cc_6 ---TATCTCAAAAGCTGATATTCATCA-AAAAAGATCAT-----GTTGTCAATTTGGT
 Cc_5 -----AAAAGCTGGTATTCATCA-AAAAAGATCAT-----GTTGTCAATTTGGT
 Cc_7 -----AAAAGCTGGTATTCATCA-AAAAAGATCAT-----GTTGTCAATTTGGT
 Cc (JX976920) ---TATCATCAAAAGCTGATATTCATCA-AAAAGGATCAT-----GCTATCAGTTTGGT
 Cc (JX976921) -----AAAAGCTGGTATTCATCA-AAAAGGATCAT-----GCTATCAGTTTGGT
 Cc (JX976922) ---TATCATCAAAAGCTGGTATTCATCA-AAAAGGATCAT-----GCTATCAGTTTGGT
 Cc (JX976923) -----AAAAGCTGGTATTCATCA-AAAAGGATCAT-----GCTATCAGTTTGGT
 Dm (JX976931) -----AAAAGCTGGTATTCATCATA--AAGGTCAT-----GCTCTCAATTTGGT
 Ce (JX976927) AAA-----GAAAGCTGATATTCATCA-AAAAAGATCAT-----GCTGTCAATTTGGT
 Dm (JX976935) ---TATCAAAAAAAGCTGATATTCATCA-AAAAAGATCAT-----GCTGTCAATTTGGT
 Ce (JX976926) -----AAAAGCTGATATTCATCA-AAAAAGATCAT-----GCTGTCAATTTGGT
 Ce (JX976924) AAA-----AAAAGCTGATATTCATCA-AAAAAGGTCAT-----GCTGTCAATTTGGT
 Ce (JX976925) -----AAAAAGCTGATATTCATCA-AAAAAGGTCAT-----GCTGTCAATTTGGT
 Hs (Mariner-36_HSal) AAAGCACTTCTAAAGCAGACATTCACCAAAAGGAAAGTTTT-----GCTATCTGTTTGGT
 Hs (Mariner-23_HSal) AAAGCACATCAAAAGCAGACATCCACCAGAGAAAGGTGAT-----GCTCTCAATCTGGT
 Pb (PBa_Mariner-23_HSal) AAAGCACATCAAAAGCAAACATCCACCAGAGGAAGGTGAT-----GCTCTCAATCTGAT
 Ca (Mariner_CA) AAACCACTTCAAAGGCTGATATCCACCAAAAGAAGGTTAT-----GCTGTCTGTTTGGT
 Si (Mariner-24_SIn) AAAGCACTTCAAAGGCCGATATTCATCAAAAAAAGGTGAT-----GCTATCTGTTTGGT
 Sm (SMAR7) AATCCACTTCAAAGGCCGATATTCACCAAAAAAAGGTTAT-----GCTATCTGTTTGGT
 Af (Mariner-1_AFl) AAAGCACTTCAAAGGCCAACATTACCAAAAAAAGGTGAT-----GCTCTCTGTTTGGT
 Hs (Mariner-11_HSal) AAAGCACATCGAAAGCTGATATCCATCAAAAGAAGGTGAT-----GCTCTCTGTTTGGT
 Hs (Mariner-16_HSal) AGACGGTCGCAAGGCTGGATTACCCCGAAGAAGGTCAT-----GCTCTGTATCTGGT

Der (Mariner-2_DER) AAACAGTGGCCAAGCCGGGATTGACGGCCAGGAAGGTTTT - - - - - GCTGTGTGTTTGGT
Hs (Mariner-26_HSal) AAACGATCGCCAAGCCCGGATTAACGGCCAAGAAGGTTTT - - - - - ACTGTGTGTTTGGT
Del (Mariner-2_DEL) AGGCGGTGGCTAAGCCAGGATTGACGGCCAGAAAGGTTTT - - - - - GCTGTGTGTTTGGT
Del (Mariner-1_DEL) AGACGGTGGCCAAGCCTGGATTGACGGCCAGGAAGGTTCT - - - - - TCTGTGTGTTTGGT
Bt (Mariner-1_BT) CAACCAACCCAAAGGCCGGTCTTCATCCAAAGAAGGTGAT - - - - - GTTGTGTATATGGT
Ac (Mariner-5_Ace) TAGCCACCCCAAAAGCCGGTCTTCATCCAAAGAAGGTCAT - - - - - GCTCTGTGTCTGGT
Si (Mariner-28_SIn) TGAC CACTCCAAAGCCTGGGCTTCATCCGGAGAAGATCAT - - - - - GCTGTGTATTTGAT
* * * * *

Hs (Mariner-24_SIn) GGGATTTCAAGGGTG - - TCATCTATTACGAACTGTTACCGTCAGGCCAAAACGATTGATTC
Hs (Mariner-45_HSal) GGGATTGGAAGGGCA - - TCGTGCATCACGAGCTGCTGCCGCCAGGCAGAACGATAGACTC
Cf (Mariner-6_CFl) GGAATTGGAAGGGGA - - TAGTATATTATGAGCTACTACCGGACAACGAAACCATCGATTC
Hs (Mariner -35_HSal) GGGATTGGAAGGGAA - - TCGTGTA CTACGAGCTGCTTCCGCAAAACCAAACAATTGATTC
Df (Mariner-1-DF) GGAATTGGAAGGGTG - - TAGTGCACTTTGAGCTGCTGGGACGGAATGAAACCATCAATTC
Ac (Mariner-1_Ace) GGGATTGGAAGGGAC - - CAGTCTTCTATGAGCTACTTCCAAAGAACAAAACGATTAATTC
Lh (ADOQ01008024) GGGATTGGAAGGGAC - - CAGTCTTCTATGAGCTGCTTCCAAAGAATAAAACGATTAATTC
Lh (ADOQ01001582) GGGATTGGAAGGGAC - - CAGTCTTCTATGAGCTGCTTCCAAAGAATAAAACGATTAATTC
Dm (JX976930) AGG - - - - - GTG - - CTATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Dm (JX976937) AGG - - - - - GTG - - CTATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Dm (JX976938) GGA - - - - - GTG - - CCATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Ce (JX976929) GGA - - - - - GTG - - CCATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Dm (JX976934) GGA - - - - - GTG - - CCATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Ce (JX976928) GGA - - - - - GTG - - CCATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Dm (JX976936) GGA - - - - - GTG - - CCATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Mr (AFJA01006902) GGGATTTCAAAGGCA - - CAGTGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Mr (AFJA01006736) GGGATTTGAAAGGCA - - TAGTGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Dm (JX976932) GGGATTTTAAAGGTG - - ACATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Dm (JX976933) GGGATTTTAAAGGTG - - ACATGTCTTTTGAAGTCTCTACTAGGGA ACCAAAGAATCAATTC
Si (AEAQ01010279) GAGATGTT AAAGGTA - - TCGTGTCTTCTGAGCTCCTACCAAGGAACCAAACAATCAATTC
Si (AEAQ01009575) GGAATTTTAAAGGTG - - TCGTGTCTTCTGAGCTCCCACCAAGGAACCAAACAATCAATTC
Cf (AEAB01001421) GGGATTGGAAGGGTG - - TGGTGTCTTTTGAAGCTCCTACCAAGGAACCAAACAATCAATTC
Cf (AEAB01018477) GGGATTGGAAGGGTG - - TGGTGTCTTTTGAAGCTCCTACCAAGGAACCAAACAATCAATTC
Ee (EEu_Mariner_Tbel) GGGATTTTAAAGGCA - - TT - - - - - TTTTGAAGTCTCTACCAAGGAACCAAACAATCAATTC
Pb (PBa_Mariner_Tbel) GGGATTTTAAAGGTG - - TCGTGTCTTTTGAAGTCTCTACCAAGGAACCAAACAATCAATTC
Tb (Mariner_Tbel) GGGATTTTAAAGGTG - - TCGTGTCTTTTGAAGTCTCTACCAAGGAACCAAACAATCAATTC
Hs (HSal_Mariner_Tbel) GGGATTTTAAAGGTG - - TCGTGTCTTTTGAAGTCTCTACCAAGGAACCAAACAATCAATTC
Hs (Mariner-22_HSal) GGGATTGGAAGGGTG - - TTGTGTCTTACGAGCTACTTCCGAAGAACGATCAATTC
Ae (Mariner-16_AEc) GGGATTGGAAGGGAA - - TCGTTCATCATGAGCTGTTATCGCCGATCAAACAATTAATTC
Ae (Aec_Mariner-8_SIn) GGGATTGGAAGGGAG - - TCGTCCATCATGAGGTGTTACCACATGGCCT AACGATTAATTC
Si (Mariner-8_SIn) GGGATTGGAAGGGAG - - TCGTCCATTATGAGGTGTTACCACATGGCCTAACGATTAATTC
Bte (Mariner-1_Bte) GGGATTATAAAGGAA - - TTCTGTACTTTGAAGTCTTACCAAGGAACCAAACAATCAATTC
Am (Ame_FAMAR1) GGGATTACAAAGGAA - - TTGTCTATTTTGAAGTCTTACCAAGGAACCAAACAATCAATTC
Fa (FAMAR1) GGGATTACAAAGGAA - - TTGTCTACTTTGAAGTCTTACCAAGGAACCAAACAATCAATTC
Hs (Mariner-2_HSal) GGGATTACAAAGGAA - - TTGTCTACTTTGAAATTTTACCAAGGAACCAAACAATCAATTC
Hs (Mariner-42_HSal) GGGATTACAAAGGTG - - TTGTGTATTTTGAAGTCTTCAAGCAACCAAACAATCAATTC
Ac (Mariner-13_Ace) GGGATTTCAAAGGTG - - TTGTGTATTTTGAAGTCTTCAAGGAATCAAACAATCAATTC
Ae (AEVX01012963) GGGATTTCAAAGGTG - - TTGTGTATTTTGAAGTCTTCAAGGAATCAAACAATCAATTC
Cc_4 GGGATTACAAAGGTG - - TTGTGTATTTTGAAGTCTTCAAGGAATCAGACC - - - - -
Cc_6 GGGATTACAAAGGTG - - TTGTGTATTTTGAAGTCTTCAAGGAATCAGACC - - - - -
Cc_5 GGGATTACAAAG GTG - - TTGTGTATTTTGAAGTCTTCAAGGAATCAGACC - - - - -
Cc_7 GGGATTACAAAGGGTG - - TTGTGTATTTTGAAGTCTTCAAGGAATCAGACC - - - - -
Cc (JX976920) GGAATTACAAACGTA - - TTGTGTATTTTGAAGTCTTCCAGGAACCAAGACGATCAATTC
Cc (JX976921) GGAATTACAAACGTA - - TTGTGTATTTTGAAGTCTTCCAGGAACCAAGACGATCAATTC
Cc (JX976922) GGAATTACAAACGTA - - TTGTGTATTTTGAAGTCTTCCAGGAACCAAGACGATCAATTC
Cc (JX976923) GGAATTACAAACGTA - - TTGTGTATTTTGAAGTCTTCCAGGAACCAAGACGATCAATTC
Dm (JX976931) GGAA - - - - - TTTTGAAGTCTTCTCAAGGGAGCAGA - - ATAAATTT
Ce (JX976927) GGAATTACAAAGGCT - - GTGTGCATTTTGAAGTCTTCAAGGAACCAAGACATCAATTC
Dm (JX976935) GGAATTACAAAGGCT - - GTGTGCATTTTGAAGTCTTCAAGGAACCAAGACATCAATTC
Ce (JX976926) GGAATTACAAAGGCT - - GTGTGCATTTTGAAGTCTTCAAGGAACCAAGACATCAATTC
Ce (JX976924) GGAATTACAAAGGCT - - GTGTGCATTTTGAAGTCTTCAAGGAACCAAGACATCAATTC
Ce (JX976925) GGAATTACAAAGGCT - - GTGTGCATTTTGAAGTCTTCAAGGAACCAAGACATCAATTC
Hs (Mariner-36_HSal) GGGATTACAAAGGTA - - TTGTGTCTTTTGAAGTCTTAAACGTAATCAAACGATTGATTC
Hs (Mariner-23_HSal) GGGATTGGAAGGGTG - - TGGTGTCTTTTGAAGTCTTACCAAGGAATGTTACGATCAATTC
Pb (PBa_Mariner-23_HSal) GGGATTGGAAGGGTG - - TGGTGTCTTTTGAAGTCTTACCAAGGAATGTTACGATCAATTC
Ca (Mariner_CA) GGGATTGGAAGGGTG - - TGGTGTCTTTTGAAGTCTTCAAGGAACCAAACAATCAATTC

Si (Mariner-24_SIn) GGGATTTCAAAGAAATCTTTTTTTTTTTGAGCTTCTACCGGACAACACAACGATCAATT C
 Sm (SMAR7) GGGATTTCAAAGGGGA--TCGT-TTTTTTGAGCTTCTACCGGACAATACCACGATTAATT C
 Af (Mariner-1_AFl) GGGATTTCAAAGGGAG--TTGTTTTTTTTTGAGCTTCTACCTGACAATTGCACGATTAATT C
 Hs (Mariner-11_HSal) GGGATTGGAAGGAA--TGTTTTTTTTTGAGCTGCTACCAACAACCAACGATTGATT C
 Hs (Mariner-16_HSal) GGGATTGGAAGGAA--TCGTCTATTATGAGCTGCTGCCACCCAACAAAACGATTGATT C
 Der (Mariner-2_DER) GGGATTGGAAGGGAA--TCATCCACTATGAGCTGCTCCCATATGGCCAGACGCTTAATT C
 Hs (Mariner-26_HSal) GGGATTGGAAGGGAA--TCATCCACTATGAGCTGCTCCCATATGGCCAGACGCTTAATT C
 Del (Mariner-2_DEL) GGGATTGGCAAGGAA--TTATCTACTATGAGCTGCTCCCTATGGCCAAACACTTAATT C
 Del (Mariner-1_DEL) GGGATTGGCAGGGAA--TAATCCACTATGAGCTGCTCCCTATGGCCAAACGCTCAATT C
 Bt (Mariner-1_BT) GGGATTGGAAGGGAG--TCCTCTATTATGAGCTCCTTCCGGAACCAACGATTAATT C
 Ac (Mariner-5_Ace) GGGATTGGAAGGGGA--TCCTGTATTATGAGCTTTTACCAACAACGAGACGATAAATT C
 Si (Mariner-28_SIn) GGGACTGGAA--GGTG--TCGTGTA TTATGAGCTCCTTCCAAGAACGTACGGCTTAATT C

Hs (Mariner-24_SIn) AACTGTATACTGTTTCGCAATTGACGAAATT---GAACCAAGCAATCCGCACAAAACG--
 Hs (Mariner-45_HSal) GGACCTGTACTGTGCGACAATTAGCGCGATT---GCACCTAGCAATTCAAAAGAAACG--
 Cf (Mariner-6_CFl) GGACAAGTACTGTTTACAGTTGGATAAATT---GAAGATAGAAATCGCCAAAAAGTG--
 Hs (Mariner-35_HSal) CAACAAGTATTGTTCCCAATTGGAAGTGT---GAAGGCAGCAATCGATGAGAAGCG--
 Df (Mariner-1_DF) AGATGTGTACTGTGCGCAGCTATCCAATTT---GGCGGAAAAAATTAAAGAAAAGGA--
 Ac (Mariner-1_Ace) CGATGTCTACTGTGAGCAGCTGCAGAAATT---AAGTGATGCCATCGCACAGAAACG--
 Lh (ADOQ01008024) CGATGTCTCCTGTGAGCAGCTACAGAAATT---AAGTGATGCCATCGCACAGAAACG--
 Lh (ADOQ01001582) CGATGTCTACTGTGAGCAGCTACAGAAATT---AAGTGATGCCATCGCACAGAAACG--
 Dm (JX976930) AAATCCATACTGTTAACAATTT-----GAACGAATCCGTTACCTAGAAAGG--
 Dm (JX976937) AAATCCATACTGTTAACAATTT-----GAACGAATCCGTTACCTAGAAAGG--
 Dm (JX976938) AAATCTATACTGTTAACAATTT-----GAACGAATCTGTTATCAAGAAAGG--
 Ce (JX976929) AAATCTATACTGTTAACAATTT-----GAACGAATCTGTTATCTAGAAAGG--
 Dm (JX976934) AAATCTATACTGTTAACAATTT-----GAACGAATCTGTTATCTAGAAAGG--
 Ce (JX976928) AAATCTATACTGTTAACAATTT-----GAACGAATCTGTTATCTAGAAAGG--
 Dm (JX976936) AAATCTAT ACTGTTAACAATTT-----GAACGAATCTGTTATCTAGAAAGG--
 Mr (AFJA01006902) TGACGTCTACTGCCGGCAATTAGAAAGTTT---GAAAGAATCAATTGAACAAAAACG--
 Mr (AFJA01006736) TGATATCTACTGCCGGTAATTAGAGAGTTT---GAAAGAATCAATTGAACAAAAACG--
 Dm (JX976932) GAGTGTGTACTATCAGCAATTGGACAGTTTAAAGAAAGAATCCTACGTGCAGAAACA--
 Dm (JX976933) GAGTGTGTACTATCAGCAATTGGACAGTTTAAAGAAAGAATCCTTCGTGCAGAAACA--
 Si (AEAQ01010279) AAACGTGTACTGTTGGCAATTGGACAGTTT---TAACGAATCTATCACCTAGAAATG--
 Si (AEAQ01009575) GCGTATGTACTGTGCGCAATTGGACAATTT---AAACAAATCCATCATCCAGAAACGT
 Cf (AEAB01001421) GGATGTGTACTGTGCGCAATTGGACAGTTT---GAACGAATCTGTCTATCCAGAAACG--
 Cf (AEAB01018477) GGATGTGTACTGTGCGCAATTGGACAGTTT---GAACGAATCTGTCTATCCAGAAACG--
 Ee (EEU_Mariner_Tbel) AAATGTGTACTGTGCGCAATTGGACAGTTT---GAACGAATCCATCATCCAGAAATG--
 Pb (PBa_Mariner_Tbel) GAATGTGTACTGTGCGCAATTGGACAGTTT---GAACGAATCCATCA TCCAGAAACG--
 Tb (Mariner_Tbel) GAATGTGTACTGTGCGCAATTGGACAGTTT---GAACGAATCCATCATCCAGAAACG--
 Hs (HSal_Mariner_Tbel) GAATGTGTACTGTGCGCAATTGGACAGTTT---GAACGAATCCATCATCCAGAAACG--
 Hs (Mariner-22_HSal) AGATGTGTACTGTAGTCAGCTGGATAAATT---GAATGCAGCGATCCACGAGAAGCG--
 Ae (Mariner-16_AEc) CGAACTCTACTGTGAACAACCTGGAGAGATT---ACAACAAGCAATTGAGAGGAAGCG--
 Ae (AEC_Mariner-8_Sln) AGAGCTCTACTGTTTACAACCTGGATAGATT---ACAGGAAGCGATTAAGGAAAAACG--
 Si (Mariner-8_SIn) AGAGCTCTACTGTTTACAACCTGGATAGATT---ACAGGAAGCGATTAAGGAAAAACG--
 Bte (Mariner-1_Bte) AAACGTGTACGTTCAACAACCTGCCAAACT---GAGCGATGCAGTTCAAGAAAAAGCG--
 Am (Ame_FAMAR1) TGTGTCTACATTGAACAACCTAACGAAATT---AAACAATGCAGTTGAAGAAAAAGCG--
 Fa (FAMAR1) TGTGTCTACATTGAACAACCTAACGAAATT---AAACAATGCAGTTGAAGAAAAAGCG--
 Hs (Mariner-2_HSal) AGATGTATACATTCAACAACCTAACCAACT---GAACAATGCAATTCAGAAAAAGCG--
 Hs (Mariner-42_HSal) AGATGTTTACTG CCAACAACCTATGAAATT---GGAGGAAGCAATCAAAGAGAAAAAG--
 Ac (Mariner-13_Ace) AGATGTCTACTGCCAACAACCTAATGAAACT---GGAGGAAGCAATCAAAGAAAAACG--
 Ae (AEVX01012963) AGATGTCTACTGCCAACAACCTAATGAAACT---GGAGGAAGCAATTAAGAAAAACG--
 Cc_4 -----GGATGAAGCAATCAAAGAAAAAGCT--
 Cc_6 -----GGATGAAGCAATCAAAGAAAAAGCT--
 Cc_5 -----GGATGAAGCAATCAAAGAAAAAGCT--
 Cc_7 -----GGATGAAGCAATCAAAGAAAAAGCT--
 Cc (JX976920) AGATCCTTATTGTTAACAACCTTATGAAAAAT---GGATGAAGCAACCAAAGAAAAAGCT--
 Cc (JX976921) AAATCCTTATTGTTAACAACCTTATGAAAAAT---GGATGAAGCAACCAAAGAAAAATA--
 Cc (JX976922) AAATCCTTATTGTTAACAACCTTATGAAAAAT---GGATGAAGCAACCAAAGAAAAATA--
 Cc (JX976923) AAATCCTTATTGTTAACAACCTTATGAAAAAT---GGATGAAGCAACCAAAGAAAAATA--
 Dm (JX976931) ATATGTTTCTGTTAACAACCTAATGAAACT---GGATGAAGTAATCAAAGAAAAACG--
 Ce (JX976927) ACATGTTTACTGTCAACAACCTAATAAAACT---GAATGAAGCAATCAAAGAAAAACG--
 Dm (JX976935) ACATGTTTACTGTCAACAACCTAATAAAACT---AAATGAAGCAATCAAAGAAAAACG--
 Ce (JX976926) ACATGTTTACTGTCAACAACCTAATAAAACT---GAATGAAGCAATCAAAGAAAAACG--
 Ce (JX976924) ACATGTTTACTGTCAACAACCTAATAAAACT---GAATGAAGCAATCAAAGAAAAACG--

Ce (JX976925) ACATGTTTACTGTCAACAACCTAATAAACT - - - GAATGAAGCAATCAAAGAAAAACG - -
 Hs (Mariner-36_HSal) GGAATTGTACTGTCTCAATTGGACAAATT - - - ACACGAAGCAATCAAGCAGAAGCG - -
 Hs (Mariner-23_HSal) GGACGTGTACTGTCAACAGCTCGACAAATT - - - GAACGAAGCAATAGCAGAAAAAGCG - -
 Pb (PBa_Mariner-23_HSal) GGACGTGTACTGTCAACAGCTAGACAAACT - - - GAACGAAGCAATAGCAGAAAAAGCG - -
 Ca (Mariner_CA) GGATGTTTACTGTCAACAACCTGGACAAATT - - - GAATGCAGCCATCAACGAGAAACG - -
 Si (Mariner-24_SIn) TGAAGTGTACTGTCTCAGCTGGACAAATT - - - GAATGATTCACTTAAACAGAAAAAG - -
 Sm (SMAR7) TGAA GTGTACTGTGATCAACTGGACAAATT - - - GAATGATTGCTCAAACAGAAAAAG - -
 Af (Mariner-1_AFL) GGAAGTGTACTGCAATCAATTGGACAAATT - - - AAACGATTCCATCAAACAAAAGAG - -
 Hs (Mariner-11_HSal) GAATGTGTATTGTCTCAATTGGACAAATT - - - GAATGATTCCATCAAGCAGAAGAG - -
 Hs (Mariner-16_HSal) GACCAAGTACTGCTCACAACCTGGCCAAATT - - - AAAGCGAGCAATCGATCAGAAGCG - -
 Der (Mariner-2_DER) TACCATCTACTGCGAACAACCTGGACCGCTT - - - GAAGCAGGCGATCGACCAGAAGCG - -
 Hs (Mariner-26_HSal) GGACCTACACTGTCAACAACCTGAC CAGATT - - - GAAGCAGGCGATCGACGAGAAGCG - -
 Del (Mariner-2_ DEL) GGTTTTGTACTGTCAACAATTGGACCGTCT - - - GAAGGAAGCAATTGCCCAGAAGCG - -
 Del (Mariner-1_ DEL) GGACCTGTACTGCCAACAACCTGGACCGCTT - - - GAAGGCAGCACTCATG CAGAAGAG - -
 Bt (Mariner-1_BT) CAACAAGTACTGCTCCAGTTAGACCAACT - - - GAAAGCAGCACTCGACGAAAAAGCG - -
 Ac (Mariner-5_Ace) AGAGAAGTATTGTTCCCAATTAGACGAATT - - - GAAGACAGCAATTGAACAAAAACG - -
 Si (Mariner-28_SIn) GGACAAGTACTGTGCCCAATTGGACAAACT - - - GAAGGAAGCC ATTGCAGAAAAACG - -

* *

Hs (Mariner-24_SIn) - - - TCCAGAATTGGCAAAT - CGCAAGGGTGTCTCTTCCACCACGACAACGCCAGACCT
 Hs (Mariner-45_HSal) - - - GCCGGAACCTGGTCAAC - AGAAAAGGGTGTCTGTATACCACTGACAACGCCAGACCG
 Cf (Mariner-6_CFL) - - - TCCGGAATTGATCAAT - AGGAAAGGCGTCTGTTTTTCATCATGATAACGCCAGACCT
 Hs (Mariner -35_HSal) - - - TCCAGAATTGAGCAAT - CGTTATGGTGTATATCCCAAGACAACGCTAGACCT
 Df (Mariner-1_DF) - - - GCCGGCACTAGCTAAT - CGCAAGGGTATAGTCTTTCCACATGACAACGCTAGGCC
 Ac (Mariner-1_Ace) - - - TCCCGAGCTAATCAAT - CGTAAGGGTGTGGTGTTCACCACGACAATGCGAGACCA
 Lh (ADOQ01008024) - - - TCCCGAGCTAATCAAT - CGTAAGGGTGTGGTGTTCACCACGACAATGCGAGACCA
 Lh (ADOQ01001582) - - - TCCCGAGCTGATCAAT - CGTAAGGGTGTGGTGTTCACCACGACAATGTGAGACCA
 Dm (JX976930) - - - TGCAAAGCACCTTAAT - AAGTTAGGATTGGCATTTCATCGCGGCAATGCAAGGCCA
 Dm (JX976937) - - - TGCA AAGCACCTTAAT - AAGTTAGGATTGGCATTTCATCGCGGCAATGCAAGGCCA
 Dm (JX976938) - - - TGCAGAGCTTCTTAATAAGTAAAGGAGTGGCGTTTTCTGACGGCAATGCAAGGCCA
 Ce (JX976929) - - - TGTAGAGCTTCTTAATAAGTAAAGGAGTGGCGTTTTCTGACGGCAATGCAAGGCCA
 Dm (JX976934) - - - TGCAGAGCTTCTTAATAAGTAAAGGAGTGGCGTTTTCTGACGGCAATGCAAGGCCA
 Ce (JX976928) - - - TGCAGAGCTTCTTAATAAGTAAAGGAGTGGCGTTTTCTGACGGCAATGCAAGGCCA
 Dm (JX976936) - - - TGCAGAGCTTCTTAATAAGTAAAGGAGTGGCGTTTTCTGACGGCAATGCAAGGCCA
 Mr (AFJA01006902) - - - CCCAGAAGTACGCAAC - CGTAAAGGAGTTGTGTTCCACACGATAATGCGAGACCA
 Mr (AFJA01006736) - - - CCCAGAAGTGGCCAAT - CGTAAAGGATTTATGTTCCACCCGATAATGCGAGACCA
 Dm (JX976932) - - - TCCAGAGCTTGTTAAC - CATAAAGGAGTTGTATTTTCATCATGACAATGCAAGGCCA
 Dm (JX976933) - - - TCCAGAGCTTGTTAAC - CATAAAGGAGTTGTATTTTCATCATGACAATGCAAGGCCA
 Si (AEAQ01010279) - - - TATA - AATTCATTAAT - CGTAAAGGAGTTGTGTTCCATCAGCAACGCGTGACCA
 Si (AEAQ01009575) AATATCCAGAAGCTCGTTAAC - CATAAAGGAGTTGTGTTCCATCAGCAACGCGAGACCA
 Cf (AEAB01001421) - - - TCCGGAGCTCGTCAAT - CGTAAAGGAGTTGTGTTCCATCAGCAACGCGAGACCA
 Cf (AEAB01018477) - - - TCCGGAGCTCGTCAAT - CGTAAAGGAGTTGTGTTCCATCAGCAACGCGAGACCA
 Ee (EEu_Mariner_Tbel) - - - TCCAGAGCTCGTTAAT - CGTAAAGGAGTTGTGTTCCATCATGACAACGCAAGACCA
 Pb (PBa_Mariner_Tbel) - - - TCCAGAGCTCGTTAAT - CGTAAAGGAGTTGTGTTCCATCAGCAACGCGAGACCA
 Tb (Mariner_Tbel) - - - TCCAGAGCTCGTTGAT - CGTAAAGGAGTTGTGTTCCATCAGCAACGCGAGACCA
 Hs (HSaL_Mariner_Tbel) - - - TCCAGAGCTCGTTAAT - CGTAAAGGAGTTGTGTTCCATCAGCAACGCGAGACCA
 Hs (Mariner-22_HSal) - - - TCCAGAATTGGTGAAT - CGTAGAGGCGTCTCTTTCCACATGACAACGCTAGGCCG
 Ae (Mariner-16_AEc) - - - GCCAGAATTAATCAAT - AGGAGGGGTGCTGCTTTCCATCAGCAACGCTCGACCA
 Ae (AEc_Mariner-8_SIn) - - - ACCAGAATTGATTAAC - AGAAAAGGTGTTGTCTTCCATCATGACAACGCCAGACCA
 Si (Mariner-8_SIn) - - - ACCAGAAT TGATCAAC - AGAAAAGGTGTTGTCTTCCATCATGACAACGCCAGACCA
 Bte (Mariner-1_Bte) - - - GCCAGAATTGGCAAAT - CGTAAAGGGTGTGTTTTCCAGCATGATAATGCAAAGCCC
 Am (AMe_FAMAR1) - - - GCCCGAATTGACAAAT - CGAAAAGGTGTTGTATTCCATCATGACAATGCAAGGCCA
 Fa (FAMAR1) - - - GGCCGAATTGACAAAT - CGAAAAGGTGTTGTATTCCATCATGACAATGCAAGGCCA
 Hs (Mariner-2_HSal) - - - ACCAGAATTGGCAAAT - CGAAAAGGTATTGTGTTCCACCATGACAATGCAAGGCCA
 Hs (Mariner-42_HSal) - - - GCCAGAATTGGCAAAT - CGTAAAGGAAT CGTGTTCCACCACGACAATGCGAGGCC
 Ac (Mariner-13_Ace) - - - GCCAGAGTTAGCGAAT - CGCAAAGGAATCGTCTTTTCATCATGACAATGCAAGACCA
 Ae (AEVX01012963) - - - GCCAGAGTTAGCGAAT - TGCAAAGGAATCGTCTTTTCATCATGACGATGCAAGACCA
 Cc_4 - - - GCCATAATCACCACAT - CACATACGAATTGTGTTTCGATCATGATAATGCAAGGCCA
 Cc_6 - - - GCCATAATCAGGCACAT - CACATACGAATTGTGTTTCGATCATGATAATGCAAGGCCA
 Cc_5 - - - GCCATAATTTGGCAAAT - CACATAGGAAGTGTGTTTCGATCATGAT - - - CAAGGCCA
 Cc_7 - - - GCCATAATTTGGCAAAT - CACATAGGAAGTGTGTTTCGATCATGAT - - - CAAGGCCA
 Cc (JX976920) - - - GCCATAATTTGGCAAAT - CACATAGGAAGTGTGTTTCGATCATGAT - - - CAAGGCCA
 Cc (JX976921) - - - GCTAAAATTGGAAAAT - CGCTATAGAATTATGTTCCACCATGTTAAAGCAAGGCCA
 Cc (JX976922) - - - GCTAAAATTGGAAAAT - CGCTATAGAATTGTGTTCCACCATGTTAAAGCAAGGCCA
 Cc (JX9723) - - - GCTAAAATTGGAAAAT - CGCTATAGAATTGTGTTCCACCATGTTAAAGCAAGGCCA

```

Dm (JX976931)      - - - - GCCAAATTTGGCAAAT - CGCAAAGGGAGTTGTGTTCCACCGTGATAATGCAAGGCCA
Ce (JX976927)      - - - - ATCAGAATTGGCAAAT - CGCAAAGGAATTGTGTTCCACCATGATAATGCAAGGCCA
Dm (JX976935)      - - - - ATCAGAATTGGCAAAT - CGCAAAGGAATTGTGTTCCACCATGATAATGCAAGGCCA
Ce (JX976926)      - - - - ATCAGAATTGGCAAAT - CGCAAAGGAATTGTGTTCCACCATGATAATGCAAGGCCA
Ce (JX976924)      - - - - ATCAGAATTGGCAAAT - CGCAAAGGAATTGTGTTCCACCATGATAATGCAAGGCCA
Ce (JX976925)      - - - - ATCAGAATTGGCAAAT - CGCAAAGGAATTGTGTTCCACCATGATAATGCAAGGCCA
Hs (Mariner-36_HSal) - - - - CCCAGAACTGGTGAAT - CGTAAAGATGTTGTCTTTCCCATGACAACGCTAGACCA
Hs (Mariner-23_HSal) - - - - GCCAGAATTGATAAAT - CGCAAAGGAGTGGTCTTCCATCATGACAACGCTCGACCA
Pb (PBa_Mariner-23_HSal) - - - - GCCAGAATTGATAAAT - CGCAAAGGAGTAGTATTCCAACATGACAACGCTCGACCA
Ca (Mariner_CA)     - - - - GCCAGAATTGATCAAT - CGTAAAGGTGTCATATTCCATCAGGACAACGCCAGACCA
Si (Mariner-24_SIn) - - - - GCCAGAATTGATCAAT - AGAAAAGGTGTAGTGTTCACCAAGATAATGCGAGACCT
Sm (SMAR7)          - - - - GCCAGAATTGATCAAT - AGA AAAGGTATAGTGTTCACCAAGATAATGCGAGACCT
Af (Mariner-1_AFl)  - - - - ATCTGAATTAATTAAC - AGGAAAGATGTAGTGTTCATCAGGATAACGCTAGACCT
Hs (Mariner-11_HSal) - - - - ACCAGAATTGGCGAAC - AGAAAAGGTGTTGTGTTTCATCATGACAACGCCAGACCT
Hs (Mariner-16_HSal) - - - - GCCCGAATTGGCGAAC - AGAAAAGGCGTTGTGTTCCACCAAGGACAACGCTAGACCA
Der (Mariner-2_DER) - - - - TCCAGAATTGGCCAAC - AGGAAGGGTGTAGTGTTCACCAAGGACAACGCCAGACCA
Hs (Mariner-26_HSal) - - - - GCCAGAATTGGCCAAT - AGGAAGGGTGTGTTGTGTTCCATCAAGACAATGCCAGGCCG
Del (Mariner-2_DEL) - - - - CCCCGCTTTGGCCAAT - AGGAAAGGAATTGTTTTCCATCAGGACAACGCTAGACCA
Del (Mariner-1_DEL) - - - - GCCATCTTTGATCAAC - AGAGGCCGAATTGTCTTCCATCAGGACAACGCCAGGCCA
Bt (Mariner-1_BT)   - - - - TCCGGAATTAGTCAAC - AGAAAACGCATAATCTTCCATCAGGATAACGCAAGACCG
Ac (Mariner-5_Ace)  - - - - TCCAGAAATAGCGAAT - CGGAAGGGCGTCGTGTTTCATCAGGACAATGCGCGGCCCT
Si (Mariner-28_SIn) - - - - CCCAGAATTGGTGAAT - AGAAAAGGTGTTTTATTCCATCAGGACAACGCCAGGCCCT

```

*

*

*

**

```

Hs (Mariner-24_SIn) CATACGTCATTGACCACTCGGAA -
Hs (Mariner-45_HSal) CATACATCTTTAACGACCCGTCAG
Cf (Mariner-6_CFl)  CATGCAAGTTTGCAGACTCGCCAA
Hs (Mariner-35_HSal) CACGTATCTTTGGCAACCCGGCAA
Df (Mariner-1-DF)   CACACATCTTTGGTCACTCGGCA -
Ac (Mariner-1_Ace)  CACACAAGTTTGGTCACTCGGC - -
Lh (ADOQ01008024)  CACACAAGTTTGGTCACTCGGC - -
Lh (ADOQ01001582)  CACACAAGTTTGGTCACTCGGC - -
Dm (JX976930)      CACACATCTTTGA - - - - -
Dm (JX976937)      CACACATCTTTGA - - - - -
Dm (JX976938)      CACACATCTTTGA - - - - -
Ce (JX976929)      CACACATCTTTGA - - - - -
Dm (JX976934)      CACACATCTTTGA - - - - -
Ce (JX976928)      CACACATCTTTGA - - - - -
Dm (JX976936)      CACACATCTTTGA - - - - -
Mr (AFJA01006902)  CACACAAGTTTGGTCACCCGC - -
Mr (AFJA01006736)  CAGACGAGTGTGGTCACCCGC - -
Dm (JX976932)      CACACATCTTTGA - - - - -
Dm (JX976933)      CACACATCTTTGA - - - - -
Si (AEAQ01010279)  TACATAAGTTTGATCAC - - - - -
Si (AEAQ01009575)  CACACAATTTTGATCAG - - - - -
Cf (AEAB01001421)  CACACATCTTTGGC - - - - -
Cf (AEAB01018477)  CACACATCTTTGGC - - - - -
Ee (EEu_Mariner_Tbel) CACACAAGTTTGATCACCTGT - -
Pb (PBa_Mariner_Tbel) CACACAAGTTTGACCACCCGT - -
Tb (Mariner_Tbel)   CACACAAGTTTGATCACCCGT - -
Hs (HSal_Mariner_Tbel) CACACAAGTTTGATCACCCGT - -
Hs (Mariner-22_HSal) CACACATCTCTACAAACCCGGCAA
Ae (Mariner-16_AEc) CACACATCTTTGATAACTCGGCAA
Ae (AEC_Mariner-8_SIn) CACACATCTTTGATGACGCGCCAA
Si (Mariner-8_SIn)  CACACATCTTTGATGACGCGCCAA
Bte (Mariner-1_Bte) CACACATCTTTGGTCACTCGCCA -
Am (AME_FAMAR1)    CACACATCTTTGGTCACTCGGCA -
Fa (FAMAR1)        CACACATCTTTGGTCACTCGGCA -
Hs (Mariner-2_HSal) CACACATCTTTGGTCACTCGGCA -
Hs (Mariner-42_HSal) CATACATCTTTAGCAACACGTAC -
Ac (Mariner-13_Ace) TACACATCTTTGGCCACGCGGCA -
Ae (AEVX01012963)  TACACATCTTTGGCCACGCGGCA -
Cc_4                CACACATCTTTGTTGA - - - - -
Cc_6                CACACATCTTTGGA - - - - -
Cc_5                CACACATCTTTGTTGA - - - - -

```


Supplementary Material S3

Sequences	Hs (Mariner-47_HSal)	Hs (Mariner-45_HSal)	Cf (Mariner-6_CFI)	Hs (Mariner-35_HSal)	Df (Mariner-1_DF)	Ac (Mariner-1_ACe)	Lh (ADOQ01001582)	Lh (ADOQ01008024)	Dm (JX976930)
Hs (Mariner-47_HSal)	ID								
Hs (Mariner-45_HSal)	0.665	ID							
Cf (Mariner-6_CFI)	0.685	0.632	ID						
Hs (Mariner-35_HSal)	0.669	0.653	0.722	ID					
Df (Mariner-1_DF)	0.639	0.628	0.644	0.653	ID				
Ac (Mariner-1_ACe)	0.647	0.583	0.657	0.653	0.676	ID			
Lh (ADOQ01001582)	0.635	0.583	0.640	0.648	0.684	0.971	ID		
Lh (ADOQ01008024)	0.635	0.579	0.644	0.644	0.676	0.930	0.950	ID	
Dm (JX976930)	0.520	0.461	0.489	0.469	0.528	0.539	0.526	0.527	ID
Dm (JX976937)	0.516	0.461	0.485	0.473	0.532	0.543	0.530	0.531	0.995
Dm (JX976938)	0.514	0.455	0.483	0.467	0.518	0.536	0.524	0.533	0.844
Ce (JX976929)	0.514	0.451	0.483	0.463	0.514	0.532	0.520	0.529	0.853
Dm (JX976934)	0.526	0.467	0.495	0.475	0.534	0.549	0.536	0.537	0.882
Ce (JX976928)	0.522	0.467	0.487	0.475	0.538	0.553	0.540	0.541	0.877
Dm (JX976936)	0.514	0.455	0.483	0.471	0.522	0.540	0.528	0.537	0.853
Mr (AFJA01006902)	0.643	0.604	0.616	0.636	0.663	0.707	0.703	0.679	0.557
Mr (AFJA01006736)	0.635	0.608	0.616	0.640	0.655	0.699	0.691	0.674	0.549
Dm (JX976932)	0.572	0.518	0.534	0.558	0.576	0.607	0.591	0.600	0.659
Dm (JX976933)	0.584	0.534	0.546	0.570	0.596	0.623	0.607	0.609	0.685
Si (AEAQ01010279)	0.608	0.529	0.601	0.629	0.596	0.626	0.614	0.617	0.622
Si (AEAQ01009575)	0.604	0.545	0.593	0.605	0.604	0.630	0.618	0.621	0.605
Cf (AEAB01001421)	0.647	0.612	0.636	0.624	0.668	0.687	0.670	0.683	0.668
Cf (AEAB01018477)	0.647	0.612	0.636	0.624	0.668	0.687	0.670	0.683	0.668
Ee (EEu_Mariner_Tbel)	0.655	0.587	0.653	0.648	0.672	0.703	0.691	0.674	0.648
Pb (PBa_Mariner_Tbel)	0.688	0.624	0.653	0.685	0.684	0.728	0.716	0.699	0.665
Tb (Mariner_Tbel)	0.688	0.608	0.657	0.677	0.684	0.728	0.716	0.699	0.661
Hs (HSal_Mariner_Tbel)	0.692	0.612	0.661	0.681	0.688	0.732	0.720	0.703	0.665
Hs (Mariner-22_HSal)	0.693	0.644	0.734	0.718	0.693	0.710	0.689	0.689	0.530
Ae (Mariner-16_AEc)	0.677	0.706	0.657	0.689	0.644	0.673	0.673	0.665	0.489
Ae (AEc_Mariner-8_Sin)	0.689	0.710	0.685	0.669	0.640	0.653	0.648	0.640	0.497
Si (Mariner-8_Sin)	0.685	0.710	0.689	0.685	0.648	0.665	0.661	0.653	0.497
Bte (Mariner-1_Bte)	0.668	0.604	0.616	0.657	0.700	0.684	0.672	0.663	0.549
Am (AMe_FAMAR1)	0.688	0.608	0.640	0.624	0.692	0.700	0.696	0.688	0.569
Fa (FAMAR1)	0.680	0.600	0.632	0.624	0.692	0.700	0.696	0.688	0.569
Hs (Mariner-2_HSal)	0.704	0.595	0.624	0.644	0.737	0.684	0.680	0.672	0.565
Hs (Mariner-42_HSal)	0.692	0.628	0.632	0.657	0.696	0.680	0.676	0.663	0.565
Ac (Mariner-13_ACe)	0.704	0.608	0.644	0.657	0.745	0.713	0.717	0.700	0.569
Ae (AEVX01012963)	0.692	0.604	0.636	0.644	0.737	0.700	0.704	0.688	0.573
Cc_4	0.479	0.408	0.457	0.473	0.520	0.473	0.481	0.481	0.515
Cc_6	0.495	0.420	0.465	0.489	0.536	0.477	0.485	0.497	0.524
Cc_5	0.479	0.412	0.457	0.473	0.520	0.465	0.473	0.481	0.497
Cc_7	0.487	0.416	0.461	0.469	0.524	0.469	0.477	0.485	0.502
Cc (JX976920)	0.577	0.493	0.567	0.555	0.622	0.576	0.563	0.570	0.590
Cc (JX976921)	0.573	0.493	0.546	0.559	0.602	0.567	0.555	0.569	0.588
Cc (JX976922)	0.598	0.506	0.571	0.575	0.627	0.584	0.572	0.579	0.607
Cc (JX976923)	0.581	0.502	0.555	0.559	0.610	0.576	0.563	0.577	0.594
Dm (JX976931)	0.549	0.489	0.518	0.493	0.596	0.576	0.584	0.589	0.589
Ce (JX976927)	0.598	0.538	0.575	0.587	0.659	0.633	0.633	0.634	0.588
Dm (JX976935)	0.598	0.534	0.575	0.587	0.659	0.637	0.637	0.650	0.590
Ce (JX976926)	0.594	0.538	0.571	0.579	0.655	0.629	0.629	0.644	0.590
Ce (JX976924)	0.602	0.546	0.587	0.600	0.668	0.646	0.646	0.640	0.590
Ce (JX976925)	0.590	0.542	0.575	0.583	0.659	0.633	0.633	0.648	0.590
Hs (Mariner-36_HSal)	0.668	0.640	0.657	0.673	0.713	0.684	0.688	0.684	0.536
Hs (Mariner-23_HSal)	0.655	0.604	0.718	0.657	0.717	0.724	0.736	0.728	0.563
Pb (PBa_Mariner23_HSal)	0.651	0.608	0.693	0.648	0.709	0.699	0.716	0.707	0.543
Ca (Mariner_CA)	0.700	0.677	0.710	0.738	0.737	0.762	0.754	0.754	0.581
Si (Mariner-24_Sin)	0.638	0.607	0.700	0.655	0.662	0.715	0.707	0.703	0.524
Sm (SMAR7)	0.655	0.620	0.706	0.657	0.672	0.729	0.725	0.729	0.516
Al (Mariner-1_AFI)	0.631	0.575	0.693	0.624	0.647	0.688	0.692	0.680	0.512
Hs (Mariner-11_HSal)	0.672	0.628	0.718	0.697	0.684	0.725	0.721	0.713	0.553
Hs (Mariner-16_HSal)	0.697	0.697	0.714	0.734	0.628	0.673	0.669	0.661	0.469
Der (Mariner-2_DEr)	0.685	0.718	0.636	0.673	0.608	0.657	0.648	0.636	0.465
Hs (Mariner-46_HSal)	0.677	0.722	0.657	0.669	0.640	0.663	0.632	0.624	0.493
Del (Mariner-2_DEI)	0.624	0.685	0.644	0.665	0.608	0.608	0.600	0.591	0.489
Del (Mariner-1_DEI)	0.575	0.697	0.620	0.628	0.604	0.595	0.587	0.575	0.477
Bt (Mariner-1_BT)	0.653	0.636	0.730	0.738	0.636	0.636	0.636	0.628	0.497
Ac (Mariner-5_ACe)	0.676	0.624	0.722	0.702	0.663	0.643	0.643	0.647	0.520
Si (Mariner-28_Sin)	0.663	0.648	0.697	0.714	0.631	0.639	0.622	0.631	0.514

Sequences	Dm (JX976937)	Dm (JX976938)	Ce (JX976929)	Dm (JX976934)	Ce (JX976928)	Dm (JX976936)	Mr (AFJA01006902)	Mr (AFJA01006736)	Dm (JX976932)	Dm (JX976933)	Si (AEAQ01010279)
Hs (Mariner-47_HSal)											
Hs (Mariner-45_HSal)											
Cf (Mariner-6_CFI)											
Hs (Mariner-35_HSal)											
Df (Mariner-1_DF)											
Ac (Mariner-1_ACe)											
Lh (ADOQ01001582)											
Lh (ADOQ01008024)											
Dm (JX976930)											
Dm (JX976937)		ID									
Dm (JX976938)	0.849										
Ce (JX976929)		0.980	ID								
Dm (JX976934)	0.877	0.962	0.971	ID							
Ce (JX976928)	0.882	0.962	0.962	0.990	ID						
Dm (JX976936)	0.858	0.990	0.990	0.971	0.971	ID					
Mr (AFJA01006902)	0.561	0.559	0.555	0.572	0.576	0.563	ID				
Mr (AFJA01006736)	0.553	0.547	0.543	0.559	0.563	0.551	0.909	ID			
Dm (JX976932)	0.655	0.679	0.684	0.673	0.669	0.684	0.654	0.634	ID		
Dm (JX976933)	0.681	0.669	0.673	0.699	0.695	0.673	0.670	0.650	0.974	ID	
Si (AEAQ01010279)	0.626	0.619	0.628	0.640	0.640	0.628	0.705	0.673	0.673	0.693	ID
Si (AEAQ01009575)	0.609	0.615	0.611	0.632	0.636	0.619	0.697	0.681	0.722	0.742	0.773
Cf (AEAB01001421)	0.672	0.673	0.669	0.686	0.690	0.678	0.735	0.714	0.754	0.771	0.784
Cf (AEAB01018477)	0.672	0.673	0.669	0.686	0.690	0.678	0.735	0.714	0.754	0.771	0.784
Ee (EEu_Mariner_Tbel)	0.652	0.641	0.637	0.654	0.658	0.646	0.776	0.760	0.739	0.756	0.806
Pb (PBa_Mariner_Tbel)	0.669	0.662	0.658	0.679	0.683	0.666	0.801	0.785	0.768	0.788	0.842
Tb (Mariner_Tbel)	0.665	0.662	0.658	0.674	0.679	0.666	0.809	0.793	0.768	0.784	0.834
Hs (HSal_Mariner_Tbel)	0.669	0.666	0.662	0.679	0.683	0.670	0.814	0.797	0.772	0.788	0.838
Hs (Mariner-22_HSal)	0.526	0.524	0.528	0.540	0.536	0.528	0.648	0.640	0.598	0.610	0.625
Ae (Mariner-16_AEc)	0.485	0.471	0.471	0.487	0.483	0.471	0.632	0.612	0.538	0.554	0.561
Ae (AEc_Mariner-8_Sln)	0.493	0.491	0.495	0.512	0.512	0.495	0.632	0.616	0.554	0.570	0.593
Si (Mariner-8_Sln)	0.502	0.500	0.495	0.512	0.520	0.504	0.644	0.628	0.554	0.570	0.589
Bte (Mariner-1_Bte)	0.553	0.551	0.546	0.563	0.563	0.555	0.713	0.680	0.608	0.625	0.632
Am (AMc_FAMAR1)	0.565	0.538	0.542	0.563	0.558	0.542	0.692	0.676	0.608	0.629	0.604
Fa (FAMAR1)	0.565	0.538	0.542	0.563	0.559	0.542	0.684	0.668	0.600	0.620	0.600
Hs (Mariner-2_HSal)	0.569	0.555	0.551	0.567	0.571	0.559	0.700	0.692	0.604	0.620	0.616
Hs (Mariner-42_HSal)	0.569	0.571	0.567	0.587	0.587	0.575	0.688	0.692	0.604	0.625	0.624
Ac (Mariner-13_ACe)	0.573	0.567	0.563	0.579	0.579	0.571	0.709	0.696	0.616	0.633	0.632
Ae (AEVX01012963)	0.577	0.571	0.567	0.583	0.583	0.575	0.704	0.692	0.608	0.625	0.620
Cc_4	0.511	0.500	0.504	0.504	0.495	0.504	0.524	0.500	0.536	0.536	0.497
Cc_6	0.529	0.495	0.491	0.513	0.513	0.500	0.541	0.524	0.531	0.553	0.510
Cc_5	0.493	0.486	0.490	0.491	0.482	0.490	0.516	0.500	0.536	0.536	0.497
Cc_7	0.497	0.490	0.495	0.495	0.486	0.495	0.516	0.500	0.532	0.531	0.493
Cc (JX976920)	0.594	0.566	0.562	0.583	0.583	0.570	0.603	0.590	0.597	0.614	0.597
Cc (JX976921)	0.584	0.578	0.592	0.581	0.573	0.583	0.578	0.578	0.610	0.604	0.597
Cc (JX976922)	0.603	0.575	0.587	0.600	0.592	0.579	0.599	0.590	0.605	0.622	0.618
Cc (JX976923)	0.590	0.585	0.598	0.587	0.579	0.589	0.586	0.578	0.616	0.610	0.605
Dm (JX976931)	0.584	0.597	0.601	0.600	0.595	0.601	0.557	0.561	0.590	0.588	0.551
Ce (JX976927)	0.593	0.635	0.631	0.629	0.629	0.640	0.652	0.644	0.662	0.659	0.613
Dm (JX976935)	0.594	0.622	0.618	0.630	0.630	0.626	0.648	0.640	0.648	0.661	0.630
Ce (JX976926)	0.594	0.637	0.633	0.630	0.630	0.641	0.648	0.640	0.663	0.661	0.626
Ce (JX976924)	0.594	0.637	0.633	0.630	0.630	0.641	0.657	0.657	0.668	0.665	0.610
Ce (JX976925)	0.594	0.634	0.630	0.630	0.630	0.639	0.644	0.644	0.665	0.665	0.622
Hs (Mariner-36_HSal)	0.540	0.530	0.526	0.538	0.542	0.534	0.725	0.713	0.612	0.625	0.648
Hs (Mariner-23_HSal)	0.567	0.561	0.557	0.577	0.577	0.565	0.736	0.716	0.623	0.643	0.702
Pb (PBa_Mariner23_HSal)	0.547	0.528	0.524	0.545	0.549	0.532	0.724	0.703	0.603	0.623	0.686
Ca (Mariner_CA)	0.586	0.571	0.567	0.583	0.587	0.575	0.692	0.696	0.653	0.669	0.668
Si (Mariner-24_Sln)	0.528	0.522	0.518	0.530	0.530	0.526	0.699	0.695	0.588	0.600	0.634
Sm (SMAR7)	0.520	0.518	0.514	0.526	0.526	0.522	0.704	0.700	0.580	0.592	0.644
Af (Mariner-1_AFI)	0.516	0.502	0.497	0.510	0.510	0.506	0.676	0.655	0.572	0.584	0.644
Hs (Mariner-11_HSal)	0.557	0.546	0.542	0.559	0.563	0.551	0.709	0.700	0.657	0.673	0.664
Hs (Mariner-16_HSal)	0.465	0.459	0.463	0.479	0.475	0.463	0.624	0.620	0.566	0.582	0.577
Der (Mariner-2_DEr)	0.461	0.455	0.459	0.471	0.467	0.459	0.612	0.608	0.538	0.550	0.537
Hs (Mariner-46_HSal)	0.489	0.483	0.487	0.500	0.495	0.487	0.604	0.595	0.554	0.566	0.545
Del (Mariner-2_DEI)	0.495	0.467	0.471	0.483	0.475	0.471	0.604	0.620	0.546	0.558	0.545
Del (Mariner-1_DEI)	0.473	0.451	0.455	0.467	0.463	0.455	0.583	0.571	0.534	0.546	0.521
Bt (Mariner-1_BT)	0.493	0.475	0.479	0.495	0.491	0.479	0.608	0.608	0.558	0.574	0.561
Ac (Mariner-5_ACe)	0.516	0.510	0.514	0.526	0.522	0.514	0.651	0.647	0.584	0.596	0.596
Si (Mariner-28_Sln)	0.510	0.495	0.500	0.512	0.508	0.500	0.610	0.606	0.568	0.580	0.584

Sequences	Si (AEAQ01009575)	Cf (AEAB01001421)	Cf (AEAB01018477)	Ee (EEu_Mariner_Tbel)	Pb (PBa_Mariner_Tbel)	Tb (Mariner_Tbel)	Hs (HSal_Mariner_Tbel)	Hs (Mariner-22_HSal)	Ae (Mariner-16_AEc)
Hs (Mariner-47_HSal)									
Hs (Mariner-45_HSal)									
Cf (Mariner-6_CFI)									
Hs (Mariner-35_HSal)									
Df (Mariner-1_DF)									
Ac (Mariner-1_ACe)									
Lh (ADOQ01001582)									
Lh (ADOQ01008024)									
Dm (JX976930)									
Dm (JX976937)									
Dm (JX976938)									
Ce (JX976929)									
Dm (JX976934)									
Ce (JX976928)									
Dm (JX976936)									
Mr (AFJA01006902)									
Mr (AFJA01006736)									
Dm (JX976932)									
Dm (JX976933)									
Si (AEAQ01010279)									
Si (AEAQ01009575)	ID								
Cf (AEAB01001421)	0.784	ID							
Cf (AEAB01018477)	0.784		1,000 ID						
Ee (EEu_Mariner_Tbel)	0.794	0.822	0.822	ID					
Pb (PBa_Mariner_Tbel)	0.846	0.871	0.871	0.925	ID				
Tb (Mariner_Tbel)	0.838	0.863	0.863	0.933	0.983	ID			
Hs (HSal_Mariner_Tbel)	0.842	0.867	0.867	0.938	0.987	0.995	ID		
Hs (Mariner-22_HSal)	0.621	0.710	0.710	0.689	0.714	0.718	0.722	ID	
Ae (Mariner-16_AEc)	0.557	0.616	0.616	0.608	0.624	0.624	0.628	0.648	ID
Ae (AEc_Mariner-8_Sin)	0.585	0.632	0.632	0.640	0.653	0.653	0.657	0.677	0.787
Si (Mariner-8_Sin)	0.589	0.640	0.640	0.644	0.657	0.657	0.661	0.677	0.795
Bte (Mariner-1_Bte)	0.608	0.655	0.655	0.696	0.692	0.692	0.696	0.677	0.636
Am (AMe_FAMAR1)	0.608	0.663	0.663	0.680	0.680	0.680	0.684	0.661	0.697
Fa (FAMAR1)	0.600	0.655	0.655	0.672	0.672	0.672	0.676	0.661	0.697
Hs (Mariner-2_HSal)	0.620	0.659	0.659	0.704	0.696	0.700	0.704	0.677	0.661
Hs (Mariner-42_HSal)	0.608	0.688	0.688	0.688	0.713	0.713	0.717	0.718	0.628
Ac (Mariner-13_ACe)	0.616	0.696	0.696	0.709	0.721	0.721	0.725	0.722	0.648
Ae (AEVX01012963)	0.604	0.684	0.684	0.696	0.709	0.709	0.713	0.710	0.653
Cc_4	0.495	0.525	0.525	0.508	0.524	0.528	0.528	0.510	0.465
Cc_6	0.506	0.560	0.560	0.524	0.545	0.545	0.545	0.526	0.465
Cc_5	0.493	0.534	0.534	0.512	0.528	0.528	0.533	0.514	0.457
Cc_7	0.489	0.534	0.534	0.508	0.524	0.524	0.528	0.518	0.461
Cc (JX976920)	0.585	0.655	0.655	0.628	0.623	0.619	0.623	0.620	0.546
Cc (JX976921)	0.556	0.615	0.615	0.615	0.603	0.603	0.607	0.595	0.530
Cc (JX976922)	0.576	0.650	0.650	0.640	0.628	0.623	0.628	0.620	0.542
Cc (JX976923)	0.564	0.621	0.621	0.623	0.611	0.611	0.615	0.604	0.538
Dm (JX976931)	0.551	0.595	0.595	0.616	0.596	0.596	0.590	0.571	0.514
Ce (JX976927)	0.621	0.679	0.679	0.673	0.665	0.661	0.665	0.644	0.559
Dm (JX976935)	0.643	0.693	0.693	0.677	0.665	0.661	0.665	0.640	0.563
Ce (JX976926)	0.634	0.676	0.676	0.673	0.661	0.661	0.665	0.636	0.559
Ce (JX976924)	0.627	0.676	0.676	0.685	0.677	0.673	0.677	0.648	0.567
Ce (JX976925)	0.639	0.672	0.672	0.677	0.665	0.665	0.669	0.632	0.563
Hs (Mariner-36_HSal)	0.668	0.721	0.721	0.733	0.745	0.750	0.754	0.734	0.657
Hs (Mariner-23_HSal)	0.682	0.728	0.728	0.740	0.777	0.777	0.781	0.738	0.669
Pb (PBa_Mariner23_HSal)	0.682	0.703	0.703	0.716	0.761	0.753	0.757	0.710	0.640
Ca (Mariner_CA)	0.672	0.754	0.754	0.745	0.774	0.774	0.778	0.771	0.714
Si (Mariner-24_Sin)	0.603	0.670	0.670	0.715	0.711	0.719	0.723	0.692	0.647
Sm (SMAR7)	0.612	0.680	0.680	0.695	0.709	0.717	0.721	0.702	0.661
Af (Mariner-1_AFI)	0.648	0.655	0.655	0.692	0.713	0.713	0.717	0.677	0.640
Hs (Mariner-11_HSal)	0.668	0.704	0.704	0.766	0.766	0.774	0.778	0.710	0.665
Hs (Mariner-16_HSal)	0.601	0.612	0.612	0.628	0.648	0.648	0.653	0.689	0.718
Der (Mariner-2_DEr)	0.553	0.600	0.600	0.620	0.624	0.628	0.632	0.632	0.746
Hs (Mariner-46_HSal)	0.549	0.620	0.620	0.624	0.640	0.636	0.640	0.648	0.755
Del (Mariner-2_DEl)	0.553	0.632	0.632	0.624	0.636	0.632	0.636	0.624	0.697
Del (Mariner-1_DEl)	0.525	0.600	0.600	0.600	0.600	0.604	0.608	0.624	0.702
Bt (Mariner-1_BT)	0.577	0.620	0.620	0.620	0.632	0.632	0.636	0.689	0.657
Ac (Mariner-5_ACe)	0.592	0.639	0.639	0.643	0.659	0.659	0.663	0.677	0.673
Si (Mariner-28_Sin)	0.596	0.651	0.651	0.618	0.663	0.655	0.659	0.706	0.628

Sequences	Ae (Aec_Mariner-8_Sin)	Si (Mariner-8_Sin)	Bte (Mariner-1_BTe)	Am (AMe_FAMAR1)	Fa (FAMAR1)	Hs (Mariner-2_HSal)	Hs (Mariner-42_HSal)	Ac (Mariner-13_ACe)	Ae (AEVX01012963)	Cc_4	Cc_6
Hs (Mariner-47_HSal)											
Hs (Mariner-45_HSal)											
Cf (Mariner-6_Cf)											
Hs (Mariner-35_HSal)											
Df (Mariner-1_Df)											
Ac (Mariner-1_ACe)											
Lh (ADOQ01001582)											
Lh (ADOQ01008024)											
Dm (JX976930)											
Dm (JX976937)											
Dm (JX976938)											
Ce (JX976929)											
Dm (JX976934)											
Ce (JX976928)											
Dm (JX976936)											
Mr (AFJA01006902)											
Mr (AFJA01006736)											
Dm (JX976932)											
Dm (JX976933)											
Si (AEAQ01010279)											
Si (AEAQ01009575)											
Cf (AEAB01001421)											
Cf (AEAB01018477)											
Ee (EEu_Mariner_Tbel)											
Pb (PBa_Mariner_Tbel)											
Tb (Mariner_Tbel)											
Hs (HSal_Mariner_Tbel)											
Hs (Mariner-22_HSal)											
Ae (Mariner-16_AEc)											
Ae (Aec_Mariner-8_Sin)	ID										
Si (Mariner-8_Sin)	0.975	ID									
Bte (Mariner-1_Bte)	0.657	0.665	ID								
Am (AMe_FAMAR1)	0.697	0.702	0.807	ID							
Fa (FAMAR1)	0.689	0.693	0.807	0.991	ID						
Hs (Mariner-2_HSal)	0.681	0.693	0.844	0.901	0.901	ID					
Hs (Mariner-42_HSal)	0.653	0.657	0.733	0.729	0.721	0.770	ID				
Ac (Mariner-13_ACe)	0.685	0.689	0.745	0.750	0.741	0.774	0.856	ID			
Ae (AEVX01012963)	0.689	0.693	0.745	0.750	0.741	0.774	0.848	0.983			
Cc_4	0.473	0.477	0.580	0.577	0.569	0.581	0.598	0.614	ID		
Cc_6	0.473	0.485	0.600	0.590	0.581	0.606	0.627	0.631	0.606	ID	
Cc_5	0.473	0.477	0.580	0.573	0.565	0.581	0.602	0.610	0.606	0.925	0.886
Cc_7	0.477	0.481	0.580	0.573	0.565	0.581	0.598	0.606	0.602	0.936	0.896
Cc (JX976920)	0.571	0.583	0.691	0.696	0.688	0.717	0.733	0.733	0.729	0.673	0.714
Cc (JX976921)	0.559	0.563	0.670	0.659	0.651	0.680	0.713	0.692	0.684	0.663	0.646
Cc (JX976922)	0.575	0.579	0.691	0.688	0.680	0.709	0.737	0.725	0.717	0.639	0.670
Cc (JX976923)	0.567	0.571	0.679	0.668	0.659	0.688	0.721	0.700	0.692	0.665	0.648
Dm (JX976931)	0.538	0.542	0.631	0.618	0.610	0.639	0.655	0.668	0.659	0.587	0.576
Ce (JX976927)	0.604	0.616	0.720	0.696	0.692	0.750	0.778	0.770	0.762	0.666	0.681
Dm (JX976935)	0.608	0.620	0.724	0.709	0.704	0.754	0.774	0.774	0.766	0.652	0.683
Ce (JX976926)	0.604	0.616	0.720	0.692	0.688	0.745	0.770	0.762	0.754	0.682	0.673
Ce (JX976924)	0.616	0.628	0.724	0.709	0.704	0.762	0.782	0.774	0.766	0.663	0.678
Ce (JX976925)	0.608	0.620	0.716	0.696	0.692	0.750	0.766	0.758	0.750	0.675	0.669
Hs (Mariner-36_HSal)	0.681	0.689	0.717	0.717	0.709	0.729	0.704	0.733	0.721	0.520	0.545
Hs (Mariner-23_HSal)	0.689	0.697	0.729	0.721	0.713	0.713	0.704	0.737	0.729	0.559	0.572
Pb (PBa_Mariner23_HSal)	0.657	0.665	0.704	0.704	0.696	0.713	0.688	0.713	0.704	0.530	0.543
Ca (Mariner_CA)	0.763	0.779	0.737	0.750	0.741	0.750	0.766	0.782	0.774	0.553	0.565
Si (Mariner-24_Sin)	0.659	0.676	0.715	0.735	0.727	0.743	0.727	0.699	0.695	0.487	0.508
Sm (SMAR7)	0.677	0.693	0.704	0.717	0.709	0.725	0.717	0.704	0.692	0.491	0.512
Af (Mariner-1_Af)	0.673	0.677	0.663	0.676	0.672	0.680	0.655	0.684	0.672	0.471	0.475
Hs (Mariner-11_HSal)	0.718	0.730	0.721	0.725	0.717	0.750	0.717	0.721	0.717	0.512	0.528
Hs (Mariner-16_HSal)	0.718	0.714	0.665	0.685	0.677	0.665	0.665	0.653	0.648	0.491	0.485
Der (Mariner-2_DEr)	0.751	0.751	0.628	0.640	0.640	0.640	0.624	0.616	0.608	0.436	0.436
Hs (Mariner-46_HSal)	0.755	0.755	0.669	0.677	0.677	0.685	0.657	0.653	0.640	0.461	0.473
Del (Mariner-2_DEl)	0.697	0.697	0.636	0.624	0.624	0.628	0.628	0.636	0.636	0.440	0.448
Del (Mariner-1_DEl)	0.738	0.738	0.608	0.600	0.600	0.608	0.628	0.608	0.604	0.440	0.440
Bt (Mariner-1_BT)	0.693	0.702	0.644	0.648	0.640	0.648	0.616	0.673	0.661	0.506	0.497
Ac (Mariner-5_ACe)	0.665	0.669	0.680	0.684	0.676	0.692	0.663	0.680	0.672	0.504	0.504
Si (Mariner-28_Sin)	0.693	0.693	0.663	0.643	0.635	0.631	0.659	0.672	0.676	0.483	0.491

Sequences	Cc_5	Cc_7	Cc (JX976920)	Cc (JX976921)	Cc (JX976923)	Cc (JX976924)	Dm (JX976931)	Ce (JX976927)	Dm (JX976935)	Ce (JX976926)	Ce (JX976924)	Ce (JX976925)	Hs (Mariner-36_HSal)	
Hs (Mariner-47_HSal)														
Hs (Mariner-45_HSal)														
Cf (Mariner-6_CFI)														
Hs (Mariner-35_HSal)														
Df (Mariner-1_DF)														
Ac (Mariner-1_ACe)														
Lh (ADOQ01001582)														
Lh (ADOQ01008024)														
Dm (JX976930)														
Dm (JX976937)														
Dm (JX976938)														
Ce (JX976929)														
Dm (JX976934)														
Ce (JX976928)														
Dm (JX976936)														
Mr (AFJA01006902)														
Mr (AFJA01006736)														
Dm (JX976932)														
Dm (JX976933)														
Si (AEAQ01010279)														
Si (AEAQ01009575)														
Cf (AEAB01001421)														
Cf (AEAB01018477)														
Ee (EEu_Mariner_Tbel)														
Pb (PBA_Mariner_Tbel)														
Tb (Mariner_Tbel)														
Hs (HSal_Mariner_Tbel)														
Hs (Mariner-22_HSal)														
Ae (Mariner-16_AEc)														
Ae (AEc_Mariner-8_Sln)														
Si (Mariner-8_Sln)														
Bte (Mariner-1_Bte)														
Am (AMe_FAMAR1)														
Fa (FAMAR1)														
Hs (Mariner-2_HSal)														
Hs (Mariner-42_HSal)														
Ac (Mariner-13_ACe)														
Ae (AEVX01012963)														
Cc_4														
Cc_6														
Cc_5		ID												
Cc_7		0.978	ID											
Cc (JX976920)		0.716	0.711	ID										
Cc (JX976921)		0.663	0.663	0.854	ID									
Cc (JX976922)		0.643	0.639	0.909	0.944	ID								
Cc (JX976923)		0.669	0.665	0.858	0.986	0.948	ID							
Dm (JX976931)		0.592	0.592	0.693	0.736	0.714	0.745	ID						
Ce (JX976927)		0.671	0.666	0.793	0.798	0.798	0.803	0.765	ID					
Dm (JX976935)		0.656	0.652	0.822	0.781	0.826	0.786	0.748	0.952	ID				
Ce (JX976926)		0.687	0.682	0.786	0.823	0.790	0.837	0.785	0.965	0.944	ID			
Ce (JX976924)		0.668	0.663	0.790	0.794	0.794	0.800	0.770	0.986	0.948	0.960	ID		
Ce (JX976925)		0.679	0.675	0.782	0.815	0.786	0.828	0.786	0.956	0.944	0.991	0.961	ID	
Hs (Mariner-36_HSal)		0.524	0.532	0.622	0.598	0.618	0.606	0.553	0.643	0.647	0.639	0.647	0.635	ID
Hs (Mariner-23_HSal)		0.563	0.567	0.654	0.621	0.641	0.629	0.604	0.670	0.670	0.670	0.683	0.674	0.733
Pb (PBA_Mariner23_HSal)		0.534	0.539	0.625	0.596	0.617	0.604	0.592	0.654	0.654	0.654	0.666	0.658	0.717
Ca (Mariner_CA)		0.557	0.561	0.668	0.655	0.663	0.655	0.627	0.704	0.700	0.704	0.717	0.709	0.778
Si (Mariner-24_Sln)		0.487	0.483	0.617	0.605	0.626	0.613	0.573	0.650	0.646	0.646	0.662	0.650	0.727
Sm (SMAR7)		0.491	0.495	0.614	0.602	0.622	0.610	0.572	0.643	0.639	0.639	0.651	0.643	0.741
Af (Mariner-1_AFI)		0.467	0.471	0.561	0.536	0.553	0.545	0.516	0.594	0.598	0.594	0.606	0.598	0.741
Hs (Mariner-11_HSal)		0.520	0.516	0.631	0.606	0.627	0.614	0.577	0.655	0.655	0.655	0.668	0.659	0.770
Hs (Mariner-16_HSal)		0.485	0.481	0.559	0.563	0.571	0.571	0.546	0.591	0.595	0.595	0.604	0.600	0.673
Der (Mariner-2_DER)		0.440	0.444	0.510	0.510	0.522	0.518	0.510	0.559	0.555	0.559	0.571	0.563	0.640
Hs (Mariner-46_HSal)		0.461	0.465	0.555	0.538	0.551	0.546	0.538	0.604	0.595	0.600	0.616	0.604	0.628
Del (Mariner-2_DEI)		0.444	0.440	0.530	0.506	0.522	0.514	0.493	0.567	0.563	0.563	0.575	0.567	0.648
Del (Mariner-1_DEI)		0.432	0.436	0.526	0.518	0.530	0.526	0.489	0.538	0.534	0.538	0.551	0.542	0.616
Bt (Mariner-1_BT)		0.493	0.497	0.555	0.555	0.563	0.555	0.522	0.583	0.587	0.587	0.595	0.591	0.628
Ac (Mariner-5_ACe)		0.495	0.491	0.602	0.594	0.606	0.602	0.573	0.610	0.610	0.614	0.618	0.618	0.647
Si (Mariner-28_Sln)		0.487	0.487	0.565	0.561	0.573	0.565	0.543	0.598	0.602	0.602	0.602	0.598	0.655

Sequences	Hs (Mariner-23_HSal)	Pb (Pba_Mariner23_HSal)	Ca (Mariner_CA)	Si (Mariner-24_SIn)	Sm (SMAR7)	Af (Mariner-1_AfI)	Hs (Mariner-11_HSal)	Hs (Mariner-16_HSal)	Der (Mariner-2_DEr)
Hs (Mariner-47_HSal)									
Hs (Mariner-45_HSal)									
Cf (Mariner-6_CfI)									
Hs (Mariner-35_HSal)									
Df (Mariner-1_DF)									
Ac (Mariner-1_ACe)									
Lh (ADOQ01001582)									
Lh (ADOQ01008024)									
Dm (JX976930)									
Dm (JX976937)									
Dm (JX976938)									
Ce (JX976929)									
Dm (JX976934)									
Ce (JX976928)									
Dm (JX976936)									
Mr (AFJA01006902)									
Mr (AFJA01006736)									
Dm (JX976932)									
Dm (JX976933)									
Si (AEAQ01010279)									
Si (AEAQ01009575)									
Cf (AEAB01001421)									
Cf (AEAB01018477)									
Ee (EEu_Mariner_Tbel)									
Pb (Pba_Mariner_Tbel)									
Tb (Mariner_Tbel)									
Hs (HSal_Mariner_Tbel)									
Hs (Mariner-22_HSal)									
Ae (Mariner-16_AEc)									
Ae (AEc_Mariner-8_SIn)									
Si (Mariner-8_SIn)									
Bte (Mariner-1_Bte)									
Am (AMe_FAMAR1)									
Fa (FAMAR1)									
Hs (Mariner-2_HSal)									
Hs (Mariner-42_HSal)									
Ac (Mariner-13_ACe)									
Ae (AEVX01012963)									
Cc_4									
Cc_6									
Cc_5									
Cc_7									
Cc (JX976920)									
Cc (JX976921)									
Cc (JX976922)									
Cc (JX976923)									
Dm (JX976931)									
Ce (JX976927)									
Dm (JX976935)									
Ce (JX976926)									
Ce (JX976924)									
Ce (JX976925)									
Hs (Mariner-36_HSal)									
Hs (Mariner-23_HSal)	ID								
Pb (Pba_Mariner23_HSal)	0.950	ID							
Ca (Mariner_CA)	0.786	0.766	ID						
Si (Mariner-24_SIn)	0.723	0.707	0.747	ID					
Sm (SMAR7)	0.717	0.700	0.774	0.906	ID				
Af (Mariner-1_AfI)	0.737	0.725	0.729	0.804	0.823	ID			
Hs (Mariner-11_HSal)	0.754	0.729	0.811	0.808	0.790	0.823	ID		
Hs (Mariner-16_HSal)	0.661	0.640	0.726	0.635	0.648	0.640	0.722	ID	
Der (Mariner-2_DEr)	0.608	0.591	0.706	0.623	0.653	0.616	0.681	0.755	ID
Hs (Mariner-46_HSal)	0.628	0.612	0.726	0.631	0.644	0.595	0.677	0.726	0.840
DeI (Mariner-2_DEI)	0.612	0.587	0.669	0.603	0.608	0.595	0.657	0.693	0.804
DeI (Mariner-1_DEI)	0.604	0.575	0.689	0.611	0.624	0.583	0.657	0.685	0.804
Bt (Mariner-1_BT)	0.653	0.648	0.730	0.663	0.669	0.657	0.677	0.742	0.677
Ac (Mariner-5_ACe)	0.672	0.659	0.696	0.686	0.684	0.655	0.709	0.726	0.648
Si (Mariner-28_SIn)	0.672	0.684	0.741	0.630	0.643	0.622	0.680	0.722	0.644

Sequences	Hs (Mariner-46_HSal)	Del (Mariner-2_DEI)	Del (Mariner-1_Del)	Bt (Mariner-1_BT)	Ac (Mariner-5_Ac)	Si (Mariner-28_Sin)
Hs (Mariner-47_HSal)						
Hs (Mariner-45_HSal)						
Cf (Mariner-6_CFI)						
Hs (Mariner-35_HSal)						
Df (Mariner-1_DF)						
Ac (Mariner-1_Ac)						
Lh (ADOQ01001582)						
Lh (ADOQ01008024)						
Dm (JX976930)						
Dm (JX976937)						
Dm (JX976938)						
Ce (JX976929)						
Dm (JX976934)						
Ce (JX976928)						
Dm (JX976936)						
Mr (AFJA01006902)						
Mr (AFJA01006736)						
Dm (JX976932)						
Dm (JX976933)						
Si (AEAQ01010279)						
Si (AEAQ01009575)						
Cf (AEAB01001421)						
Cf (AEAB01018477)						
Ee (EEu_Mariner_Tbel)						
Pb (PBa_Mariner_Tbel)						
Tb (Mariner_Tbel)						
Hs (HSa_Mariner_Tbel)						
Hs (Mariner-22_HSal)						
Ae (Mariner-16_AEc)						
Ae (AEC_Mariner-8_Sin)						
Si (Mariner-8_Sin)						
Bte (Mariner-1_Bte)						
Am (AMe_FAMAR1)						
Fa (FAMAR1)						
Hs (Mariner-2_HSal)						
Hs (Mariner-42_HSal)						
Ac (Mariner-13_Ac)						
Ae (AEVX01012963)						
Cc_4						
Cc_6						
Cc_5						
Cc_7						
Cc (JX976920)						
Cc (JX976921)						
Cc (JX976922)						
Cc (JX976923)						
Dm (JX976931)						
Ce (JX976927)						
Dm (JX976935)						
Ce (JX976926)						
Ce (JX976924)						
Ce (JX976925)						
Hs (Mariner-36_HSal)						
Hs (Mariner-23_HSal)						
Pb (PBa_Mariner23_HSal)						
Ca (Mariner_CA)						
Si (Mariner-24_Sin)						
Sm (SMAR7)						
Al (Mariner-1_AFI)						
Hs (Mariner-11_HSal)						
Hs (Mariner-16_HSal)						
Der (Mariner-2_DEI)						
Hs (Mariner-46_HSal)	ID					
Del (Mariner-2_DEI)	0.775	ID				
Del (Mariner-1_DEI)	0.787	0.820	ID			
Bt (Mariner-1_BT)	0.636	0.628	0.644	ID		
Ac (Mariner-5_Ac)	0.644	0.616	0.575	0.742	ID	
Si (Mariner-28_Sin)	0.665	0.644	0.616	0.742	0.745	ID

3.4. Capítulo 4:

Horizontal transfers of *Mariner* transposons between mammals and insects

Sarah G Oliveira^{1,*}, Weidong Bao², Cesar Martins¹, Jerzy Jurka²

¹UNESP –Sao Paulo State University, Bioscience Institute, Morphology Department, Botucatu, SP, 18618-970, Brazil

²Genetic Information Research Institute, 1925 Landings Drive, Mountain View, CA 94043, USA

* *Corresponding author*: UNESP –Sao Paulo State University, Bioscience Institute, Morphology Department, Botucatu, SP, Brazil, 18618-970 Botucatu, SP, Brazil Phone/Fax: 55 14 38800462.
e-mail: sarahg.oliveira@gmail.com

SGO: sarahg.oliveira@gmail.com

WB: weidong@girinst.org

CM: cmartins@ibb.unesp.br

JJ: jjurka@gmail.com

Abstract**Background**

Active transposable elements (TEs) can be passed between genomes of different species by horizontal transfer (HT). This may help them to avoid vertical extinction due to elimination by natural selection or silencing. HT is relatively frequent within eukaryotic taxa, but rare between distant species.

Findings

Closely related Mariner-type DNA transposon families, collectively named as *Mariner-1_Tbel* families, are present in the genomes of two ants and two mammalian genomes. Consensus sequences of the four families show pairwise identities greater than 95%. In addition, mammalian *Mariner 1_BT* family shows a close evolutionary relationship with some insect Mariner families. Mammalian *Mariner 1_BT* type sequences are present only in species from three groups including ruminants, tooth whales (*Odontoceti*), and New World leaf-nosed bats (*Phyllostomidae*).

Conclusions

Horizontal transfer accounts for the presence of *Mariner_Tbel* and *Mariner 1_BT* families in mammals. *Mariner_Tbel* family was introduced into hedgehog and tree shrew genomes approximately 100-69 million years ago (MYA). Most likely, these TE families were transferred from insects to mammals, but details of the transfer remain unknown.

Keywords

DNA transposon, genome evolution, horizontal transfer, *Mariner*

Findings

In contrast to the vertical transmission of the genetic material from parents to offspring, the horizontal transfer (HT) is a process in which new genetic information is transmitted between different, sometimes distant, species (Syvanen, 1994; Brown, 2003). HT is likely to be one of the factors leading to the persistence of TEs in eukaryotes (Hartl et al., 1997b; Sanchez-Gracia et al., 2005; Schaack et al., 2010), and complicating the evolutionary trees.

The detection of HT is mostly inferential, mainly based on the combination of two types of evidence: unusually high similarity between TE sequences from species that have long diverged from each other, and a limited distribution of one particular TE family within a group of species (Silva et al., 2004). To date, numerous HTs have been detected in eukaryotes (Daniels et al., 1990; Maruyama and Hartl, 1991; Robertson and Lampe 1995b; Silva et al., 2004; Loreto et al., 2001), but of particular interest are HTs across distant branches. A recent example of such a rare event is HT of hAT DNA transposon families between vertebrate and invertebrate species (Gilbert et al., 2010).

Here we report two families of *Mariner*-type DNA transposons that have possibly undergone HT from insects to mammals. The first family, called *Mariner_Tbel*, was originally identified in the tree shrew (*Tupaia belangeri*), but families nearly identical to *Mariner_Tbel* were also found in the genome of another mammal, European hedgehog (*Erinaceus europaeus*), and in two ant species: red harvester ant (*Pogonomyrmex barbatus*), and Jerdon's jumping ant (*Harpegnathos saltator*) (Table 1, 2). Although the copy numbers and divergence vary between the families, the family consensus sequences reconstructed in each genome show a high level of identity to each other throughout the entire length (~1.3-kb) (Table 1). The lowest identity is found between the two families in *E. europaeus* and ant *P. barbatus* (95.84%), and the highest identity is found between the two ant families (98.45%). Therefore, unless otherwise stated, the four families in the genomes are referred collectively to as *Mariner_Tbel* families. Given the long divergence time between insects and mammals (~1 billion years) (Hedges et al., 2004; Moreau et al., 2006; Brady et al., 2009), this high identity strongly indicates that HT took place during the evolutionary history of *Mariner_Tbel* families. This notion is consistent with the fact that mammal *Mariner_Tbel* sequences were found only in two distantly related mammalian species, even though over 30 mammalian genomes were sequenced to date.

We then estimated the approximate ages of the four *Mariner_Tbel* families in each genome. In mammals, we compared the sequence divergences of *Mariner_Tbel* to an older *Mariner*-type family (*TIGGER1*), relatively common in the mammalian genomes (Table 2). *TIGGER1* elements are present in multiple copies in eutherian mammals, but only one or two degenerated copies were found in marsupial genomes, including *Macropus eugenii*, *Monodelphis domestica*, and *Sarcophilus harrisii* (Fig 1A). Therefore, mammalian *TIGGER1* families likely expanded after the split of marsupials and placentals (190 MYA), but before the placental radiation (~100 MYA) (Meredith et al., 2011). In the genome of the tree shrew and European hedgehog the divergence of the *TIGGER1* family is $21.2 \pm 2\%$ and $28.0 \pm 3\%$, respectively (Table 2). Therefore, based on the divergence of *Mariner_Tbel* in the two mammal genomes ($15.0 \pm 2\%$ and $19.4 \pm 3\%$, respectively), the ages of *Mariner_Tbel* in the two mammals were calculated to be ~134-70 MYA and ~131-69 MYA, respectively. Because it is unlikely that the mammalian *Mariner_Tbel* expanded in the common ancestor of placental mammals before 100 MY, we adjusted the ages to be 100-70 MYA and 100-69 MYA, respectively (see Fig. 1A and Meredith et al., 2011).

In the ant genomes, no *Mariner* family was yet identified as unambiguously present in the common ancestor of all ant species. Among potential candidates are the oldest known *Mariner* families present in some of the ant genomes (e.g. *Mariner-28_SIn* or *Mariner-94_HSal*; Fig. 1). These small families may have expanded in the common ancestor of all ant species (140 Mya) (Moreau et al., 2006), assuming that they were lost in some ant species. Alternatively, these old families might have expanded in some ant species after they split from their common ancestor. Under either scenario, the outermost ages when the two ant *Mariner_Tbel* families expanded could be still estimated by comparing their diversities with the diversity of *Mariner-28_SIn* (Table 2). Based on that, *Mariner_Tbel* family in the red harvester ant (*P. barbatus*) and Jerdon's jumping ant (*H. saltator*) expanded at most ~43 and ~50 million years ago, respectively.

The above age estimates suggest that the two ant *Mariner_Tbel* families are possibly younger than the mammalian *Mariner_Tbel* families. However, the history of *Mariner_Tbel* can be traced further back in ants and their insect relatives than in mammals. Individual *Mariner_Tbel*-like elements from distinct families, such as AEAQ01009575, AEAB01001421 and AFJA01006902 (Fig. 1B), were also found in the genomes of two other ants (*Solenopsis invicta* and *Camponotus floridanus*) as well as in the alfalfa leafcutting bee (*Megachile*

rotundata). These *Mariner_Tbel*-like sequences and *Mariner_Tbel* sequences form a single lineage in the phylogenetic tree (Fig. 1B), with the bee sequences in a more ancestral position (Fig. 1B). The topology of this particular lineage mirrors the evolutionary history of the ant and bee species (Fig. 1A). Furthermore, figure 1B indicates that the *Mariner_Tbel* family and many other similar *Mariner* families in ants and other insects shared a common ancestral sequence. These observations suggest the ancestor of ants *Mariner_Tbel* may have been present in some ant or other insect species very long time ago, probably as far back as the common ancestor of bees and ants (~150 Mya) (Brady et al., 2009). Thus, the mammalian *Mariner_Tbel* families probably originated by HTs from insects to mammals through some unknown vectors. Given that the two mammals belong to two distinct lineages, *Mariner_Tbel* in tree shrew and hedgehog may represent two independent HTs (Fig. 1A). Notably, we cannot rule out the possibility that the *Mariner_Tbel* families in one of the two ant species, or both, also originated by HTs. This possibility is suggested by the two facts: (a) the relatively young ages (at most ~43-50 Mya) of the two families, (b) the high identity (98.5%) between the two family consensus sequences, even *H. saltator* and *P. barbatus* diverged from each other ~100 million years ago (Moreau et al., 2006). Among insect species, frequent HTs have been documented in flies (Bartolome et al., 2009). Alternatively, *Mariner_Tbel* sequences could have survived for a very long time in either of the two ant genomes before the most recent family expansions.

In addition to mammalian *Mariner_Tbel* families, *Mariner1_BT* DNA families might also have originated by HT from insects (Fig 1A). We were able to obtain high quality consensus of *Mariner1_BT* from bovine species (*Bos taurus*, *Bos indicus*, *Bos grunniens*, *Bubalus bubalis*) and bottlenosed dolphin *Tursiops truncatus* (Table 2). All the derived consensus sequences show similar lengths (~1280-bp), and high pairwise identities throughout the entire length (> 98%). Blast screening against NCBI databases using *Mariner1_BT* consensus sequence as query also detected this family in several other mammalian species, including one bat species, *Carollia perspicillata* (Seba's short-tailed bat), additional ruminants, and whale (Table 3). These BlastN hits show similar score and query coverage (>80% identity to the consensus and >90% coverage). In summary, *Mariner1_BT* type TEs were found only in three taxonomic groups to date: ruminants, tooth whales (*Odontoceti*), and New World leaf-nosed bats (*Phyllostomidae*) (Fig. 1A.). Notably, in *C. perspicillata* (short-tailed fruit bat) and *Desmodus rotundus* (vampire bat),

we also detected a family of nonautonomous DNA transposon, called *Mariner-N1_CPe*, which was likely derived from the bat *Mariner1_BT* family (Fig 1C).

Remarkably, other closest relatives of *Mariner1_BT* are all found in ant species: *Mariner1_BT* co-clusters significantly (bootstrap = 83) with three other ant *Mariner* families (*Mariner-5_ACe*, *Mariner-28_SIn* and *Mariner-35_HSal*) (Fig. 1B). Given the vast diversity of *Mariners* found in insects (Fig. 1B), and the confined distribution of *Mariner1_BT* in mammals, we propose *Mariner1_BT* family could also originate from a horizontally transferred insect-like element. Using a similar method above, i.e. based on the family divergence and mammalian phylogeny (Table 2, and Fig. 1A), we estimated the ages of bovine *Mariner1_BT* to be 90-85 MYA, and 90-63 MYA for dophin *T. truncatus* *Mariner1_BT* family. The age of *Mariner1_BT* in bat could not be estimated due to insufficient data. We also could not determine if HT happened in mammals more than once, because the three taxonomic groups that include *Mariner1_BT* are relatively close.

In summary, this is the first report of two cases of horizontally transferred *Mariner* elements (*Mariner_Tbel* and *Mariner1_BT*) between insects and mammals. Previously, four families of DNA transposons from the *hAT* superfamily were also found to be involved in multiple waves of HT between insects and other vertebrates including mammals (Gilbert et al., 2010). This could partially be attributed to the fact that insects are the largest and the most diverse group of invertebrate animals on earth. While insects are the most likely source of the horizontally transferred transposons, the original source or possible intermediaries, such as parasitic insects (Gilbert et al., 2010) or viruses, remain unclear. This is complicated by the possibility that recurrent HTs of related *Mariner* elements are likely to take place between different insects (Bartolome et al., 2009). The role of viruses in HT proposed a while ago (Kidwell, 1993), still remains to be understood. As more genome sequence data become available, more mechanistic details on HT between mammals and insects are likely to emerge.

Materials and methods

Mariner transposable elements from Repbase (<http://www.girinst.org/repbase/>) were used as an initial query to screen *Mariners* in diverse genomes available at NCBI (National Center for Biotechnology Information - <http://www.ncbi.nlm.nih.gov/>). Family consensus sequences were constructed whenever possible. The copy numbers in each family were determined by BLASTN

using consensus sequences as queries. Sequence divergence within each family was assessed by the average pairwise k-distance (Kimura two-parameter model) between individual insertions and the corresponding consensus sequences. The k-distance was calculated using the software MEGA 4 (Tamura et al, 2007). For a given family, individual sequences used in k-distance calculation were randomly chosen from the family members; in most cases individual sequences matched >70% of the consensus length.

We used *Mariner_Tbel* and *Mariner1_BT* as BlastN queries against Repbase to select top scoring TE entries for phylogeny analysis. Individual sequences selected from GenBank were also used in the tree if Repbase consensus sequences were not available. The sequence alignments are shown in Additional File 1. The alignments were done using the online MAFFT server (<http://mafft.cbrc.jp/alignment/software/>). The phylogeny tree was inferred using MEGA 4 (Tamura et al, 2007), using neighbor joining (NJ) method and k-distances. Branch support was estimated using 1000 bootstrap replicates.

List of abbreviations

Ace: *Atta cephalotes*

AEc: *Acromyrmex echinator*

AFI: *Apis florae*

AMe: *Apis mellifera*

BT: *Bos Taurus*

BTe: *Bombus terrestris*

CA: *Chymomyza amoena*

CFl: *Camponotus floridanus*

Del: *Drosophila elegans*

DEr: *Drosophila erecta*

DF: *Drosophila ficusphila*

EEu: *Erinaceus europeus*

HSal: *Harpegnathos saltator*

HT: horizontal transfer

LHu: *Linepithema humile*

MRO: *Megachile rotundata*

Mya: million years ago

NCBI: National Center for Biotechnology Information

PBa: *Pogonomyrmex barbatus*

SIn: *Solenopsis invicta*

SMAR7: *Schmidtea mediterranea*

Tbel: *Tupaia belangeri*

TTr: *Tursiops truncatus*

TE(s): transposable element(s)

Competing interests

The authors declare that they have no competing interests.

Author's contributions

SGO contributed to development of the hypothesis, collection, preparation, analysis and interpretation of data, wrote the first draft of the manuscript, and revised the text. WB contributed to the analysis and interpretation of data, writing and revising the manuscript. CM contributed to the discussion of data and revisions of the manuscript. JJ contributed to development of the hypothesis, interpretation of data and final revisions. All the authors read and approved the final manuscript.

Acknowledgements

This work was supported by funds from the Sao Paulo Research Foundation (FAPESP), Sao Paulo State University (UNESP) and the National Institutes of Health grant 5 P41 LM006252. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Library of Medicine or the National Institutes of Health.

References

As referências citadas neste e nos demais capítulos encontram-se no item 5 (Referências Bibliográficas).

Tables

Table 1. Pairwise identities (%) between the *Mariner_Tbel* consensus from mammals (*T. belangeri*, *E. europeus*) and insects (*P. barbatus*, *H. saltator*)

	<i>T. belangeri</i>	<i>E. europeus</i>	<i>P. barbatus</i>
<i>E. europeus</i>	96.55		
<i>P. barbatus</i>	98.05	95.84	
<i>H. saltator</i>	97.90	95.69	98.45

Table 2. Divergence of Mariner TE families in mammalian and insect genomes

Family	Length (bp)	Copy No.	Divergence (%)*
<i>Mariner-1_Tbel</i> (TBel)	1279	>400	15.0±2 (183)
<i>Mariner-1_Tbel</i> (EEr)	1266	>70	19.4±3 (34)
<i>Mariner-1_Tbel</i> (PBa)	1285	>90	6.3±1 (53)
<i>Mariner-1_Tbel</i> (HSa)	1285	>30	7.2±2 (16)
<i>TIGGER1</i> (Tbel)	2413	>80	21.2±2 (27)
<i>TIGGER1</i> (EEu)	2410	>47	28.0±3 (25)
<i>TIGGER1</i> (BT)	2408	>500	17.3±2 (102)
<i>TIGGER1</i> (TTr)	2419	>580	12.1±1 (159)
<i>Mariner-28_SIn</i> (SIn)	1226	~14	20.1±3 (12)
<i>Mariner1_BT</i> (BT) [#]	1277	>400	14.7±2 (95)
<i>Mariner1_BT</i> (TTr)	1285	>700	7.6±1 (101)

* The divergence represents average pairwise k-distance between individual copies and the consensus. Numbers of individual sequences used for the k-distance calculation are indicated in parentheses.

[#] *Mariner1_BT* families from other *Bos indicus*, *Bos grunniens*, *Bubalus bubalis* are not shown.

Table 3. *Mariner1_BT* sequences detected in mammals. The table shows only top scores using *Mariner1_BT* as BlastN query.

Groups	Species	Accession	Score	Query coverage	E value	Identity	Gaps
bat	<i>Carollia perspicillata</i>	AC152852.2	1324	98%	0	1068/1291 (83%)	87/1291 (7%)
Ruminants	<i>Odocoileus hemionus</i>	AY330343.1	1278	98%	0	1061/1294 (82%)	73/1294 (6%)
	<i>Ovis aries</i>	AC148039.3	1274	99%	0	1078/1309 (82%)	64/1309 (5%)
	<i>Muntiacus reevesi</i>	AC174385.3	1265	99%	0	1074/1312 (82%)	68/1312 (5%)
	<i>Muntiacus muntjak vaginalis</i>	AC152844.3	1242	99%	0	1071/1313 (82%)	71/1313 (5%)
	<i>Capra hircus</i>	EU870890.1	1142	98%	0	1031/1296 (80%)	109/1296 (8%)
Whale	<i>Pseudorca crassidens</i>	AP011081.1	1833	98%	0	1180/1288 (92%)	22/1288 (2%)

Figure legend

Figure 1. A) Distribution of five Mariner families (*Mariner_Tbel*, *Mariner1_BT*, *TIGGER1*, *Mariner-28_SIn*, and *Mariner-N1_CPe*) in placental mammals, marsupials, and insects. The phylogeny is adopted from published trees of insects and mammals (Hedges et al., 2004; Moreau et al., 2006; Murphy et al., 2007; Brady et al., 2009; Meredith et al., 2011). Marsupials are represented by three species (*Monodelphis domestica*, *Macropus eugenii*, *Sarcophilus harrisii*). The scale bar on the bottom left indicates time of evolutionary branching in the tree. Species with genomic sequences available from public databases are highlighted in orange. B) Phylogenetic position of the two horizontally transferred Mariner (*Mariner_Tbel* and *Mariner1_BT*: red color), relative to other closely related Mariner. Most of the families are from insects with color-coded branches: ants (green), bees (orange), flies (pink) and other insects (blue). The remaining branches (black) are from mammals and planarian (SMAR7). Except for a few individual sequence segments (with accession numbers), all other families are represented by consensus sequences deposited in Repbase (except of highly similar copies). C) Structural relationship between *Mariner1_BT* and *Mariner-N1_CPe* family. The species abbreviations are: ACe (*Atta cephalotes*), AEc (*Acromyrmex echinator*), AFl (*Apis florea*), AMe (*Apis mellifera*), BTe (*Bombus terrestris*), BT (*Bos taurus*), CA (*Chymomyza amoena*), CFl (*Camponotus floridanus*), Del (*Drosophila elegans*), DEr (*Drosophila erecta*), DF (*Drosophila ficusphila*), EEu (*Erinaceus europeus*), HSal (*Harpegnathos saltator*), LHu (*Linepithema humile*), MRo (*Megachile rotundata*), PBa (*Pogonomyrmex barbatus*), SIn (*Solenopsis invicta*), SMAR7 (*Schmidtea mediterranea*), Tbel (*Tupaia belangeri*), CPe (*Carollia perspicillata*).

Figure 1

Additional file

File name: Additional file 1

File format: PDF

Title of data: Sequence alignments of three *Mariner* families (*Mariner_Tbel*, *Mariner1_BT*, *Mariner-28_SIn*) from eutherian mammals and insects.

Description of data: Except for a few individual sequence segments (with accession numbers), all other families are represented by consensus sequences deposited in Repbase (excluding highly similar copies). The species are: ACe (*Atta cephalotes*), AEc (*Acromyrmex echinator*), AFl (*Apis florea*), AMe (*Apis mellifera*), BTe (*Bombus terrestris*), BT (*Bos taurus*), CA (*Chymomyza amoena*), CFl (*Camponotus floridanus*), Del (*Drosophila elegans*), DEr (*Drosophila erecta*), DF (*Drosophila ficusphila*), EEu (*Erinaceus europeus*), HSal (*Harpegnathos saltator*), LHu (*Linepithema humile*), MRo (*Megachile rotundata*), PBa (*Pogonomyrmex barbatus*), SIn (*Solenopsis invicta*), SMAR7 (*Schmidtea mediterranea*), Tbel (*Tupaia belangeri*).

Additional file 1. Sequence alignment of three *Mariner* families (*Mariner_Tbel*, *Mariner1-BT*, *Mariner-28_Sln*) from eutherian mammals and insects.

1
Mariner-24_Sln
SMAR7
Mariner-11_HsAl
Mariner-1_AFl
Mariner-36_HsAl
Mariner-23_HsAl
PBA(Mariner-23_HsAl)
Mariner-42_HsAl
Mariner-13_Ace
AEC_AEVX01012963
Mariner-2_HsAl
AME(FAMAR1)
FAMAR1
Mariner-1_BT
MARINER_CA
Mariner-1_Ace
LHu_ADOQ01001582
LHu_ADOQ01008024
Mariner_Tbel
HsAl(Mariner_Tbel)
PBA(Mariner_Tbel)
EEU(Mariner_Tbel)
CFL_(AEAB01001421)
CFL_(AEAB01018477)
Sln_(AEA01009575)
Sln_(AEA01010279)
MRO_(AFJA01006902)
MRO_(AFJA01006736)
Mariner-25_HsAl
Mariner-35_HsAl
Mariner1_BT
Mariner-5_Ace
Mariner-1_DF
Mariner-6_CFl
Mariner-16_HsAl
Mariner-47_HsAl
Mariner-45_HsAl
AEC(Mariner-8_Sln)
Mariner-8_Sln
Mariner-16_Ace
Mariner-1_DEL
Mariner-46_HsAl
Mariner-2_DEL
Mariner-2_DER
Mariner-28_Sln
2
Mariner-24_Sln
SMAR7
Mariner-11_HsAl
Mariner-1_AFl
Mariner-36_HsAl
Mariner-23_HsAl
PBA(Mariner-23_HsAl)
Mariner-42_HsAl
Mariner-13_Ace
AEC_AEVX01012963
Mariner-2_HsAl
AME(FAMAR1)
FAMAR1
Mariner-1_BT
MARINER_CA
Mariner-1_Ace
LHu_ADOQ01001582
LHu_ADOQ01008024
Mariner_Tbel
HsAl(Mariner_Tbel)
PBA(Mariner_Tbel)
EEU(Mariner_Tbel)
CFL_(AEAB01001421)
CFL_(AEAB01018477)
Sln_(AEA01009575)
Sln_(AEA01010279)
MRO_(AFJA01006902)
MRO_(AFJA01006736)
Mariner-22_HsAl
Mariner-35_HsAl
Mariner1_BT
Mariner-5_Ace
Mariner-1_DF
Mariner-6_CFl
Mariner-16_HsAl
Mariner-47_HsAl
AEC(Mariner-8_Sln)
Mariner-8_Sln
Mariner-1_DEL
Mariner-46_HsAl
Mariner-2_DEL
Mariner-2_DER
Mariner-28_Sln
93
TATATATTTTACAGCTGATATAGCAAGGCTTGCTT--GTGAAAAAGTTCGATTGATTCTACGCAG-TAGTTTTTGTGGTGCCTATC
TTATATATTTTACAGCTGATATAGTAAAGGTTTGT--GTGAAAAATTTTGTGTTGATTCTTTGCAA-TAGTTTTTGGAAAAAT-----
TTTTACATTTTACAACGCACATTCAAGCTTCGTTTG--AAAAAAGTTGTTTGGAACTCTGTCAG-TTGTTTTTTGTGGCGTGCATC
TTATATATTTTACAGCTTGGCATTTCAGGTTTCTTTA--AGAAAAATTCGTTGATTCTGTGCGG-TTGTGTTTTTGTGGTGTGTAAC
TTATATCATTTTACAGCTGACATTTCAGGCTTCTTTA--AAAAAAGTTCGCTTTCGTTGAAAGTAAAG-AATTGTTTGTGGCATAGATT
CATGCGCGTTTACAGCTGACACTTATAGCTTCCATC--AAGGAAAAAAGTTCGCGCAATCGATCAAC-AATTATCTTTCTATACGCTTTT
TATGCGCGTTTATAGCTGACACTTGTAGGTTCCATCA--AAAAAAGTTCGCGCAATCGATCAAC-AATTATCTTTCTATACGCTTTT
TTTTCAGATTGGTATGTTGACACTTTCAGGTTACTATA--GAAACAAATTCGCTGTAATTTTGTGTTAC-AACTGTGGTTTACCGTCACTT
TATGTTTGTGTTGACAGCTGATCTTTAGCTTTCATTC--AGTAAAAAATACATCGATCGAACATT-TAGTTTTTGTTCAGCAATTGTTT
TATGTTTGTGTTGAGCTGTTACTTCAATTGGCATTC--AGTAAAAAATGATCGATCGGTTGTG-TAGTTTTTGTTCGCGTTCGCTTG
TTTTGTTATTTGATAGTTGGCAATTCAGCTGTCAATC--AGTAAAAAAGTTTGTGATCGGTTGCG-TAGTTTTTGTTCGCGTTCGCTTG
TTTTGTTATTTGATAGTTGGCAATTCAGCTGTCAATC--AGTAAAAAAGTTTGTGATCGGTTTCT-TAGTTTTTGTTCGCGTTCGCTTG
TATTGTTTCTTGTGTTGAGCTCCACCTTTAACTTT--AGTAAAAAATGATCGATTAGTTATT-TAGTTTCGTTTGTATCCCAATTT
TTTTGACATTTTATAGCTGTGACTTATAGCTTTCATTA--GAAAAAAGTTCGCGTAATTTTGTGTTAC-ATTTGTTTGTTCGCGCTTTT
TACGCGCTTTTGTGTTGACACTCAACCTAAAAACCA--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGCTTGCATTTGTGGTTT
TATGCGCATTTTGTGTTGACACTTAACTTAACTAAAC--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
TATGCGCATTTTGTGTTGACACTTAACTTAACTAAAC--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
CTATGTCATTTGACAGCTGACAGTTCTAGCTTCATTCT--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
CTATGTCATTTGACAGCTGACAGTTCTAGCTTCATTCT--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
CTATGTCATTTGACAGCTGACAGTTCTAGCTTCATTCT--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
ATATGCCGTTTAAACAGCTGACGTTTAAAGCTTCATT--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
TATGTCATTTGACAGTTTACAGTTCTAGTTTATTTCT--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
ACATGTCATTTGACAGTTTGGCAGTTCTAGTTTCCATTCT--AAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
TGTGCTATGCGATAGTCGTTTGTGCTAGCTTCATTCT--AGAAAAAAGTTCGCGGATTTCGCTGAG-TTTTGTGTTGTAATTTGTGGTTT
TATATTTTCTGAAAGGTTAGGTTTCTGCTGATTTT--AAGCTGTTATTTGTGCAAAATCTATTGAG-TAATTTTGTGTTGCGCAGATTG
GATTGTTATTTGACAGCTGTCATATCAGCGTGCATTTT--TTATCAAAATTTGGTGAAT-TTTTGTGTTGAGCATTATTAATTT
GTACATATCGTAAAGAGCGGTTCTAGCAGCATCTTG--GAAAAAATTTGGTTCGATACGTTAAT-TTTTGTGTTTATACGCTTT
TAGTATCGTTTAAAGCTTATGTCTTAAAGCTATCT--GAGCAAAATTTGGTGAATTTATCGATAAC-TTTTGTGTTTATAGCAAAAT
TTTCTGTTTGTGTTAGTGACATTCAACGCTATCTTA--GTGATATTAATTTATCAAAATTTAAAC-AACTTTTACTTGACATTTAGTTT
GTACATTAAAGTAAAGCTGAAATTTGTGCGGCACTTT--AGAAAGATTTTATCGAGATCCGTCAG-TATACGATTTTAAATTTATTTT
AATACACATTTGAAAGGTTATGTTTCACTGTTTATGTT--GTGCTGTTTGGTTTGGTCAACATCAG-TATTGTTGTTGCTGCGTACGCTG
TAACGTCATTTGACAGCTGACAGTTGAAGCTGATAGAC--GTACAAAAAAGTTCGGAATTTTATATA--TCAGACGTAAGGTTTGTGCGGCT
AACGCGCTCTGAAAGATGACATTTTGTGTTTATTTT--TTGTAACAAAGTTCGGAATTTTATATA--CGTGTCAAAAGTTATATTTGTTT
AACGCGCTCTGAAAGATGACATTTTGTGTTTATTTT--TTGTAACAAAGTTCGGAATTTTATATA--CGTGTCAAAAGTTATATTTGTTT
GACGTGCGTTTAAAGATGACATTTTCAACCTTTGTTT--GAGCGGAAAAAGTGAATTCGCTGAG-TGGATCAAAAGTTATTTTATTTT
-----CCTTTGATTGATTGCTTTT-----TCTTTTAAAG-CGTTTGTGAGTTTATAGCTGCG
ATACTGCGACTTAAAGGTGACATAACGCTTAAAAAA--GTTCTTTGACAGTTTGTGTTTGGTGAAG-CCAGTTGTGTTTAGAGCG-TTC
AGGTGCGCTTGAAGAGTTGAGCTTTGCTGATCAAAAGCTTCTTAGAAAGTTTGTGATCGATTAACA-CAACGCTTTTGTAGAGCTG--AT
ATACATAATTTGAAGAGCTTGTGACATAACCTTAAAC--GAGCTATGCAATGATTAGTTTGGATATT-CGCTTCAACAGTTTATAGAGCTG
CATGCTTTTGTGATAGCTGTAACCTTAAAGCATCAATT-----AGTGTATGTAGAAATTTGAATGCA---GAAGCGTTTGTATTCACTTA


```

Mariner-24_SIn 369 ATGCACCA-----
SMAR7          ACGCACCA-----
Mariner-11_HSal ATGCACCA-----
Mariner-1_AFl  ATGCGCCA-----
Mariner-36_HSal ATGCACCA-----
Mariner-23_HSal ATGCACCT-----
PBa(Mariner-23_HSal) ATGCACCT-----
Mariner-42_HSal ATGACCAA-----
Mariner-13_Ace ATAATCAA-----
AEC_AEVX01012963 ATAACCAA-----
Mariner-2_HSal ATGAACAA-----
AMe(FAMAR1)    ACGAAAAA-----
FAMAR1         ATGAAAAA-----
Mariner-1_BTe  ATGCTCAA-----
MARINER_CA     ATGAGCCA-----
Mariner-1_Ace  ACGCTCAT-----
LHu_ADOQ01001582 ACGCTCAT-----
LHu_ADOQ01008024 ACGCTCAT-----
Mariner_Tbel   ACGCTCCT-----
HSal(Mariner_Tbel) ACGCTCCT-----
PBa(Mariner_Tbel) ACGCTCCT-----
EEu(Mariner_Tbel) ATGCTCCT-----
CFL_(AEAB01001421) ATGCTACT-----
CFL_(AEAB01018477) ATGCTACT-----
SIn_(AEAQ01009575) ATTTCTCT-----
SIn_(AEAQ01010279) ACTCTCCT-----
MRO_(AFJA01006902) ACACTCCT-----
MRO_(AFJA01006736) ACACTCCT-----
Mariner-22_HSal ACAAACCA-----
Mariner-35_HSal ATGGTGAG-----
Mariner1_BT     ATGCTCCA-----
Mariner-5_Ace   ATCAAGAA-----
Mariner-1_DF     ACGCCCCA-----
Mariner-6_CFl    ATGCTCCT-----
Mariner-16_HSal ATGCTCCT-----
Mariner-47_HSal ATGCGCCC-----
Mariner-45_HSal ATGCTCCT-----
AEC(Mariner-8_SIn) ATCGACCT-----
Mariner-8_SIn   ATCGACCT-----
Mariner-16_AEc  ATGAACCT-----
Mariner-1_DEl    ATGCGCCA-----
Mariner-46_HSal ATGAGACA-----
Mariner-2_DEl    ATGAGGCA-----
Mariner-2_DEr    ATAAGCCAGCCGGCGGAAGACCTGTGACGACAAATACCGATCAAATCATGGAATACATCGAGTTAGACCGGCATGTGGCATCTCGTGACATC
Mariner-28_SIn  ATGAGAAG-----

Mariner-24_SIn 461 -----
SMAR7          -----
Mariner-11_HSal -----
Mariner-1_AFl  -----
Mariner-36_HSal -----
Mariner-23_HSal -----
PBa(Mariner-23_HSal) -----
Mariner-42_HSal -----
Mariner-13_Ace -----
AEC_AEVX01012963 -----
Mariner-2_HSal -----
AMe(FAMAR1)    -----
FAMAR1         -----
Mariner-1_BTe  -----
MARINER_CA     -----
Mariner-1_Ace  -----
LHu_ADOQ01001582 -----
LHu_ADOQ01008024 -----
Mariner_Tbel   -----
HSal(Mariner_Tbel) -----
PBa(Mariner_Tbel) -----
EEu(Mariner_Tbel) -----
CFL_(AEAB01001421) -----
CFL_(AEAB01018477) -----
SIn_(AEAQ01009575) -----
SIn_(AEAQ01010279) -----
MRO_(AFJA01006902) -----
MRO_(AFJA01006736) -----
Mariner-22_HSal -----
Mariner-35_HSal -----
Mariner1_BT     -----
Mariner-5_Ace   -----
Mariner-1_DF     -----
Mariner-6_CFl    -----
Mariner-16_HSal -----
Mariner-47_HSal -----
Mariner-45_HSal -----
AEC(Mariner-8_SIn) -----
Mariner-8_SIn   -----
Mariner-16_AEc  -----
Mariner-1_DEl    -----
Mariner-46_HSal -----
Mariner-2_DEl    -----
Mariner-2_DEr    GCCCAGGAGATGGGAGTTAGTCACCAAACCATTTTAAACCATCTGCAGAAAGGCTGGATACAAAAAAGCTTGATGTTTGGGTGCCGCATGA
Mariner-28_SIn  -----

```


3.5. Capítulo 5:

High diversity of R2 transposable elements in the Phanaeini coleopterans

Oliveira, SG¹; Burke, WD²; Moura, RC³; Martins, C¹; Eickbush, TH²

¹UNESP – Sao Paulo State University, Bioscience Institute, Morphology Department, Botucatu, SP, 18618-970, Brazil

²University of Rochester, Department of Biology, Rochester, NY, 14627, USA

³UPE-Pernambuco State University, Biological Sciences Institute, Department of Biology, Recife, PE, 50100-130, Brazil

SGO: sarahg.oliveira@gmail.com

BWD: bill.burke@rochester.edu

RCM: ritamoura.upe@gmail.com

CM: cmartins@ibb.unesp.br

THE: thomas.eickbush@rochester.edu

Abstract

The R non-LTR retrotransposons insert into a highly conserved region of the 28S rRNA gene, and have been described in many animals, including arthropods, fish, tunicates, flatworms and nematodes. This study focused on understanding the pattern of nucleotide variation observed in rDNA loci and the involvement of R2 superfamily elements with the expansion of rDNA loci number (2-15 clusters) in the chromosomes of Phanaeini coleopterans. Polymerase chain reaction (PCR) was used to obtain the following sequences: the complete ITS2 region; ~670 bp of the 28S rRNA gene containing the R2 insertion site region; ~800 bp of R2 elements; and ~840 bp of the 28S rRNA gene downstream of the R2 site. The comparative analysis of 32 R2 sequences showed extremely high levels of intraspecies nucleotide divergence: from 0.1 to 47.1% in *Coprophanaeus bellicosus*, from 13.8 to 81.7% in *C. cyanescens*, and from 0.9 to 80.8% in *C. ensifer*. Interspecies nucleotide divergences between species were from 6.4 to 81.0%. In other insect species nucleotide divergence between R2 is typically <1% for copies from the same R2 family, and >25% between families. The high diversity of R2 elements in each species is likely due to the intrinsic biological characteristics of the phaneines, considering the species are distributed in small and structured populations, which may have allowed the maintenance or loss of multiple R2 lineages. Southern blot data showed that ~48% of the 28S rRNA gene copies have R2 elements inserted in accordance to data observed in other insects (10-50%). The ITS2 region isolated from *C. cyanescens* and *C. ensifer* showed a great variability in size for both species (ranging between 31-513 bp, and 174-445 bp, respectively). Unlike the high diversity of R2 sequences, analyses of the 28S rRNA gene sequences isolated revealed very low levels of nucleotide diversity (0.0008%, 0.0005 and 0.0011%, for *C. bellicosus*, *C. cyanescens*, and *C. ensifer*, respectively). This data suggests that the R2 insertions in phaneines did not affect the concerted evolution of the 28S rRNA gene sequences, but may have been involved in the dispersion of rRNA genes to different chromosomes.

Keywords: R2 non-LTR retrotransposon, Coleoptera, evolution, ribosomal DNA

Background

In eukaryotes, the 45S ribosomal RNAs (rRNA) are represented by hundreds of gene copies organized in tandem arrays (the rDNA loci) (Long and Dawid 1980; Hillis and Dixon 1991). 45S rDNA units contain one copy of three major rRNA genes: 18S, 5.8S and 28S, transcribed from a single promoter by RNA polymerase I. Each contiguous unit contains the transcribing sequence separated by an intergenic spacer (IGS). In addition to the transcribing sequences, rDNA units have an external transcribed spacer (ETS) upstream of the 18S sequence, and two internal transcribed spacers (ITS1 and ITS2) located, respectively, between the 18S and 5.8S sequences, and the 5.8S and 28S sequences (Tautz et al. 1988; Eickbush and Eickbush, 2007). While the sequence of the rRNA genes evolves slowly, the internal transcribed spacers (ITS1 and ITS2), the external transcribed spacer (ETS), and the intergenic spacer (IGS) evolve rapidly (Hillis and Dixon, 1991; Eickbush and Eickbush, 2007; Ganley and Kobayashi, 2007; Stage and Eickbush, 2007).

Early studies with evolution of ribosomal genes suggested that the rRNA gene copies evolve horizontally, indicating that a mutation in one copy spreads to all other copies, homogeneously, in a process called “concerted evolution” (Brown et al., 1972; Zimmer et al., 1980). The concerted evolution theory explains the lack of genetic variability observed among rRNA gene copies in many different species, and it became a fundamental dogma to understand the rRNA multigene family evolution (Coen et al., 1982; Liao, 1999; Liao 2000). Although the majority of studies have shown that rRNA multigene families evolve in a concerted manner, recent studies have proposed that different mechanisms could also be involved (e.g. Carranza et al., 1999; Rooney, 2004; Rooney and Ward, 2005; Mukha et al., 2010; Nei e Rooney, 2005; Eirín-López et al., 2012).

Transposable elements (TEs) can insert randomly or into specific target sites within the host genome. As occurs with other genes and repetitive DNAs, the rRNA genes in many eukaryote lineages are targeted for insertion by different TEs. Non-LTR retrotransposons have evolved site specificity by adapting the endonuclease that initiates the integration reaction to recognize specific DNA sequences (Aksoy *et al.*, 1990; Okazaki *et al.*, 1995; Feng et al., 1998; Christensen et al., 2000; Takahashi and Fugiwara 2002). Among TEs, the R non-LTR retrotransposons (R, in reference to rRNA genes) insert into a highly conserved region of the 28S rRNA gene and have been described in many animals, including arthropods, cnidarians, fishes,

tunicates, platelmyntes, rotifers, mollusks and nematodes (e.g. Burke et al., 1995; Burke et al., 1998; Burke et al., 2003; Kojima and Fujiwara, 2004; Christensen et al., 2006; Kojima *et al.*, 2006; Eickbush and Eickbush, 2007; Gladyshev and Arkhipova, 2009; Mingazzini et al., 2010; Luchetti et al., 2012). The first elements discovered within these “rDNA arrays” were the R1 and R2 superfamilies of *Drosophila melanogaster* (Dawid and Rebbert, 1981; Roiha et al., 1981). These two elements are the most extensively studied non-LTR retrotransposable elements of arthropods (Jakubczak et al., 1991; Burke et al., 1998; Eickbush, 2002) and appear to have been vertically inherited into the 28S genes of most arthropods since the origin of this phylum (Burke et al., 1998; Malik et al., 1999). The R2 elements inserts into the specific sequence 5'-TTAAGG↓TAGCCA-3' of the 28S rRNA gene (Eickbush, 2002).

The association of rDNA with heterochromatic regions, including the presence of transposable elements, could affect the occurrence of inter- and intra-specific polymorphisms for the number of rDNA sites. Despite of only one bivalent chromosome carrying the rDNA clusters represents the ancient condition in the order Coleoptera (Schneider et al., 2007; Cabral-de-Mello et al., 2011a), a high variation of rDNA loci number (2-15 clusters) has been described in the chromosomes of Phanaeini tribe (Oliveira et al., 2010; Cabral-de-Mello et al., 2011a; Oliveira et al., 2012a). This study focused on understanding the pattern of nucleotide variation observed in rDNA loci and the involvement of R2 superfamily elements with the expansion of rDNA loci number in the chromosomes of Phanaeini coleopterans.

Materials and methods

Animals and DNA samples

Samples from adult *Coprophanaeus bellicosus*, *C. cyanescens*, *C. dardanus*, *C. ensifer*, *C. pertyi*, *Diabroctis mimas* and *Phanaeus splendidulus* (Coleoptera: Scarabaeidae: Scarabaeinae: Phanaeini) were collected from different regions of Pernambuco State, Brazil. The animals were collected in the wild according to Brazilian laws for environmental protection (wild collection permit, MMA/IBAMA/SISBIO n. 2376-1). The experimental research on animals was conducted according to the international guidelines followed by São Paulo State University (Protocol no. 35/08 – CEEA/IBB/UNESP). Whole animals were frozen and stored for DNA extraction. The genomic DNA was isolated from the muscle of the paws, following the protocol described by Sambrook and Russel (2001), with minor modifications.

Characterization of rDNA loci

Based on reported data for insects (Xiong and Eickbush 1990, Burke et al., 1993; Burke et al., 1999; Malik et al., 1999), several sets of primers were constructed for polymerase chain reaction (PCR) as follow. The isolation of the insertion site region of R2 in the 28S rRNA gene of seven Phanaeini species (*C. bellicosus*, *C. cyanescens*, *C. dardanus*, *C. ensifer*, *C. pertyi*, *Diabroctis mimas* and *Phanaeus splendidulus*) was performed with the primer set Up20_for (5'-TGC CCA GTG CTC TGA ATG TC-3'), annealing approximately 20 nucleotides upstream to the region of R2 insertion, and Cla-region_rev (5'-GAG CCG ACA TCG AAG GAT C-3'), for the cutting region of the enzyme *Cla*I (a conserved region of the 28S rRNA). The obtained PCR samples were directly sequenced (Macrogen, Rockville, MD USA) and the sequences were aligned using ClustalW allocated in the MacVector software (version 12.7.0). From the alignment obtained it was possible to design the primer R2+50_rev (5'-GTT AAT CCA TTC ATG CGC GTC-3'), positioned in a region approximately 50 base pairs downstream the insertion site of the R2 element, the primer R2-junc_for (5'-CTC TCT TAA GGT AGC CAA ATG-3') that includes the junction region of the R2 element (used to amplify sites without R2 insertions), and the primer R2-HincII_rev (5'-CCG TCG ACT AAG TCT CCC AC-3'), which includes the cutting region of the enzyme *Hinc*II (Figure 1).

A region of approximately 830 bp of the 28S rRNA gene of *C. bellicosus*, *C. cyanescens* and *C. ensifer* was isolated by PCR with the primers Cla-region_rev (5'-GAG CCG ACA TCG AAG GAT C-3'), and R2-up120_for (5'-TCC GAC TGT CTA ATT AAA ACA AAG C-3') positioned in a region approximately 50 base pairs upstream the insertion site of the R2 element. The PCR products obtained were cloned using the TOPO cloning vector (Invitrogen, Grand Island, NY USA) and subsequently sequenced.

The ITS2 region of *C. cyanescens* and *C. ensifer* was obtained with the primers 5.8Srd_for (5'-GAA CAT CGA CAT TTT GAA CGC-3') and 28S_rev (5'-AAC TTT CCC TCA CGG TAC TTG-3'). The PCR products obtained from *C. cyanescens* and *C. ensifer* were cloned and sequenced as described above. Similarity analyses among the sequences were performed using the MacVector, using the multiple sequence alignment (ClustalW).

Characterization of R2 elements

Several forward degenerated primers were designed to the QGDPLS and AFADD conserved regions of the R2 element, based on amino acid sequences of arthropods R2 elements (Burke et al., 1993). The designed primers were as follows: AFADD-1_for (5'-GCN TTY GCN GAY GAY-3'), AFADD-2_for (5'-GCN TTC GCN GAC GAC-3'), AFADD-3_for (5'-GCN TTT GCN GAT GAT-3'), AFADD-4_for (5'-GCN TTT GCN GAT GAC-3'), AFADD-5_for (5'-GCN TTC GCN GAC GAT-3'), QGDPLS-1_for (5'-CAR GGN GAY CCN YTM TC-3'), QGDPLS-2_for (5'-CAR GGN GAY CCN YTK TC-3'), QGDPLS-3_for (5'-CAA GGM GAC CCA CTA TC-3'), QGDPLS-4_for (5'-CAA GGM GAC CCC CTA TC-3'), QGDPLS-5_for (5'-CAG G GK GAT CCG TTC TC-3'), QGDPLS-6_for (5'-CAG G GK GAT CCT TTC TC-3'), QGDPLS-7_for (5'-CAA GGM GAC CCA CTG TC-3'), QGDPLS-8_for (5'-CAA GGM GAC CCC CTG TC-3'), QGDPLS-9_for (5'-CAG G GK GAT CCG TTT TC-3'), QGDPLS-10_for (5'-CAG G GK GAT CCT TTT TC-3'). These degenerated primers for the regions QGDPLS and AFADD were used in combination with the primer R2+50_rev, in order to isolate the R2 non-LTR retrotransposons. The R2 elements isolated were cloned into TOPO cloning vector and sequenced for *C. bellicosus*, *C. cyanescens* and *C. ensifer*. The sequences analysis, translation, multiple alignments (ClustalW) and the construction of a maximum likelihood tree were performed using MacVector software.

For the Southern blot experiments, ~10µg of DNA of *C. cyanescens* was restriction digested with *HincII* (cleave once within the element and once within the rDNA unit, immediately flanking the R2 element) and separated on 1.0% agarose gels. After transfer of the genomic DNA to nylon membrane, it was hybridized in 2X SSC, 5X Denhardt's at 65°C for ~14 hours. Final washing of the filters was in 0.5X SSC at 65°C. The sequence used as probe for the hybridization was amplified via PCR from genomic DNA using the primers R2-junc_for and R2-HincII_rev, which amplified a region of ~300 bp of the 28S rRNA gene without the insertion of R2. Final PCR products to be used as probes were random labeled with ³²P dCTP according to Random Primers DNA Labeling System (Invitrogen). Hybridized and washed filters were exposed to a PhosphorImager screen for 48-72 hours and the images were scanned with a Molecular Dynamics PhosphorImager Scanner. The image was analyzed with Quantity One program to produce signal traces of band intensities and to estimate the copy number of sequences.

Results

rDNA loci characterization

A region of approximately 670 bp of the 28S rRNA gene was isolated from seven Phanaeini species (*Coprophanaeus bellicosus*, *C. cyanescens*, *C. dardanus*, *C. ensifer*, *C. pertyi*, *Diabroctis mimas* and *Phanaeus splendidulus*) based in the application of the primers Up20_for and Cla-region_rev (Figure 1). After the alignment of these sequences, it was possible to identify the region of R2 insertion and the design of the following primers: R2-junc_for, R2+50_rev and R2-HincII_rev (Figure 1), which were necessary for the development of further analyses.

A region of approximately 840 bp of the 28S rRNA gene from the species *C. bellicosus*, *C. cyanescens* and *C. ensifer* was isolated by PCR using the primers R2-up120_for and Cla-region_rev (Additional file 1). The alignments pairwise similarity matrix showed low levels of nucleotide divergence: 0 to 0.4% in *Coprophanaeus bellicosus* (Table 1), from 0 to 3.9% in *C. cyanescens* (Table 2), and from 0 to 0.6% in *C. ensifer* (Table 3). The analyses of the 28S rRNA gene isolated revealed low levels of intraspecific nucleotide divergence (0.0008%, 0.0005%, and 0.0011% for *C. bellicosus*, *C. cyanescens*, and *C. ensifer*, respectively) (Figure 2). The alignment of 28S sequences of the three species showed that these sequences are organized into distinct groups (Additional file 2).

The isolation of the ITS2 showed a high variability in size, ranging between 31 to 513 bp in *C. cyanescens* (seven sequences), and 174 to 445 bp in *C. ensifer* (seven sequences) (Figure 3, Additional file 3). A base mutation was observed in the end of the 5.8S in three sequences of *C. ensifer* (Figure 3, Additional file 3c), as well a high variability in the beginning of the 28S in all sequences of *C. cyanescens* and *C. ensifer* (Figure 3, Additional file 3).

R2 elements

Sequences of 32 R2 sequences of approximately 800 bp were isolated with the primer R2+50_rev and degenerated primers for the regions QGDPLS and AFADD, aligned (Additional file 4) and the maximum likelihood tree showed the sequences grouped in different branches (Figure 4). The similarity of 32 R2 sequences showed extremely high levels of intraspecific nucleotide divergence: from 0.1 to 47.1% in *C. bellicosus*, from 13.8 to 81.7% in *C. cyanescens*, and from 0.9 to 80.8% in *C. ensifer* (Table 4). Interspecies nucleotide divergences between species were from 6.4 to 81.0% (Table 4). The intraspecific aminoacid divergence was also

extremely high: from 0.4 to 49.1% in *C. bellicosus*, 4.9 to 76.9% in *C. cyanescens*, and from 0.0 to 73.2% in *C. ensifer* (Table 5). Interspecific aminoacid divergences were from 4.8 to 76.9 % (Table 5).

Southern blot were also performed to determine what fraction of the 28S rRNA genes had insertion of R2. According to Figure 1, the region of the 28S rRNA gene between the R2 junction and the *HincII* insertion is approximately 300 bp. Considering that information, the Southern blot data showed that ~48% of the 28S rRNA gene copies of *C. cyanescens* have R2 elements inserted into the 28S rRNA genes (Figure 5).

Discussion

Previous studies suggest that unequal crossover and gene conversions between the different sequences of rDNA could be involved with the concerted evolution of rDNA loci resulting in a high intraspecific level of sequence identity between the many copies of the rDNA units in the species (reviewed by Eickbush and Eickbush, 2007; Stage and Eickbush, 2007; Stage and Eickbush, 2010). It was observed a low level of nucleotide divergence in the 28S rRNA genes for the coleopteran investigated indicating that the rRNA genes are also governed by concerted evolution mechanisms in this insect group. On the other hand, even considering the high level of 28S rRNA genes disrupted by R2 insertion (~48%), R elements insertion seem not to affect the concerted evolution of the rRNA genes. According to Mingazzini et al. (2010) the presence of R2 within a 28S rRNA gene can influence the ribosomal sequence homogeneity, and can hide the concerted evolution. However, it was observed here that the presence of R2 does not impact on 28S gene diversity, which can be explained through the rapid elimination of new insertions (Eickbush and Eickbush, 2007). These considerations are consistent with previous analysis, which indicated that R1 and R2 inserted units are rapidly lost from the rDNA locus, and these elements maintain their presence only by active retrotransposition (Jakubczak et al., 1992; Eickbush and Eickbush, 2003; Stage and Eickbush, 2007; Beauregard et al., 2008; Stage and Eickbush, 2010).

Previous works identified that 10 to over 50% of 28S rRNA gene copies could have insertion of R1 and R2 elements (Jakubczak et al., 1991; Lathe et al. 1995; Lathe and Eickbush, 1997; Eickbush and Eickbush, 2003). Those results are consistent to the obtained in the present work that determined ~48% of the 28S rRNA gene copies have R2 elements inserted.

Another question concerning the distribution of R2 elements in the rDNA locus is related to the properties of rRNA transcription. Only a small fraction of the many rDNA units in eukaryotes are actively transcribed at any one time (Conconi et al., 1989; Dammann et al., 1995; Eickbush et al., 2008). To optimize the synthesis of functional rRNA, the species may have evolved mechanisms to active these regions of the rDNA locus that contain the lowest ratio of R2 inserted units (Malik et al., 1999; Eickbush et al., 2008).

In contrast to the slowly evolving rRNA genes, the ITS2 regions changes rapidly in sequence, showing a great intra and interspecies variability in *C. cyanescens* (31-513 bp) and *C. ensifer* (174-445 bp). The data indicate that the intraspecific and interspecific variation in the ITSs were due to the occurrence of rearrangements and recombinations in the rDNA regions within individuals, affecting only the beginning of the 28S rDNA. Dover and Falvell (1984) proposed that a molecular coevolution could explain a DNA divergence and the maintenance of the function. The mechanism responsible for the pattern of variation, together with the gene conversion, could evolve unequal exchange both within a chromosome and between homologous and non-homologous chromosomes. The coevolution is a plausible model to explain the rapid sequence changes and the maintenance of spacer functions and chromatin organization. Using this idea, it is possible to propose that the spread of variant repeats creates the conditions for accumulation of changes on the ITSs.

The intra- and intraspecific comparisons of R2 sequences showed extremely high levels of nucleotide and amino acid divergences. High diversity of R2 elements was also observed in the Japanese beetle *Popillia japonica* (Coleoptera: Scarabaeidae) previously analyzed, that showed a high nucleotide sequence divergence (Burke et al., 1993; Burke et al., 1998), and even the level of amino acid divergence was high, ~55% (Burke et al., 1993). The high level of variation observed in the R2 sequences of coleopteran species suggests the action of another direction of evolution than the classical concerted evolution mechanisms. In other insect species nucleotide divergence between R2 is typically <1% for copies from the same R family, and >25% between families (Burke et al., 1998; Pérez-González and Eickbush, 2011).

The divergence of R2 elements, also present in *P. japonica*, does not represent a single large intraspecific population of elements that has a wide range of nucleotide divergence. These analyses suggest that the R2 elements in these species can be divided into distinct number of families. The presence of multiple families of R2 in the same species clearly differs from the

most common situation observed in previously studied species, where R2 elements from the same species are nearly identical in sequence. These analyses suggest that the R2 elements in these species can be divided into a distinct number of families that are each quite similar in sequence (Roiha et al., 1981; Jakubczak et al., 1990; Burke et al., 1998).

The high divergence of R2 sequences observed in *Coprophanaeus* species is possibly the result of the rapid turnover of the element and thus the preservation of a different stock of copies in each species. In the absence of concerted evolution, multiple families of R2 elements may accumulate sequence changes independently of each other. These sequences can further accumulate sufficient differences becoming distinct families. Thus, a large number of R2 families could be simultaneously maintained in a species, as detected in a few species, for example, in *P. japonica*, *Tenebrio molitor* and *Nasonia vitripennis* (Burke et al., 1993; Burke et al., 1998) and here for *Coprophanaeus* species. Multiple families of R2 are probably regulated independently and each family would be prevented from elimination by expansion and recombination between copies (Burke et al., 1993).

Eickbush et al. (1995) observed that the R1 and R2 of Diptera appear to evolve similarly to nuclear genes, suggesting that these elements must have very low rates of transposition. However, retrotransposons might be expected to evolve more quickly than would genes of their hosts, considering that their transposition will impose a high rate of mutation. These considerations may indicate that these elements are evolving faster than rDNA, and/or transposing more frequently. This higher level of transposition could explain the variability of ribosomal rDNA chromosomal clusters observed in Phanaeini species (Oliveira et al., 2010; Cabral-de-Mello et al., 2011a; Oliveira et al., 2012a).

The high diversity of R2 elements in each species could have been facilitated due to the intrinsic biological characteristics of the phaneines, related to their ecological and populational structure. Phanaeini has wide distribution in the Neotropical region, organized in small and structured populations (Edmonds, 1972; Edmonds, 1994; Edmonds, 2000; Edmonds and Zidek, 2004; Philips et al., 2004; Edmonds and Zidek, 2010), which may have allowed the maintenance or loss of multiple R2 families. Individual sequences of R2 element may be lost in a local population only after being re-introduced by occasional migrants in other population, favoring the maintenance of multiple families of R2.

In summary, R2 insertions in phaneines did not affect the concerted evolution of the 28S

rRNA gene sequences, but may have been involved in the dispersion of rRNA genes to different chromosomes.

Acknowledgements

The authors are grateful to CMQ Costa, FAB Silva and JW Barreto for the taxonomic identification of the studied species. This work was supported by funds from the São Paulo Research Foundation (FAPESP), the National Institutes of Health (USA), the Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE).

References

As referências citadas neste e nos demais capítulos encontram-se no item 5 (Referências Bibliográficas).

Figure legends

Figure 1: Multiple sequence alignment of seven Phanaeini species obtained by PCR using the primers Up20_for and Cla-region_rev. Specific regions are indicated: region of R2 insertion (bold), and the primers R2-junc_for (blue), R2+50_rev (green) and R2-HincII_rev (orange). Read: Cb, *Coprophanaeus bellicosus*; Cc, *C. cyanescens*; Cd, *C. dardanus*; Ce, *C. ensifer*; Cp, *C. pertyi*; Dm, *Diabroctis mimas*; Ps, *Phanaeus splendidulus*.

Figure 2: Maximum likelihood tree based in the analysis of approximately 840 bp of the 28S rRNA gene sequences. The sequences were isolated with the primers 28S-rev and R2-up120_for from *C. bellicosus* (a), *C. cyanescens* (b), and *C. ensifer* (c).

Figure 3: Schematic diagram of the multiple sequences of ITS2 region from *Coprophanaeus cyanescens* and *C. ensifer*. The 5.8S (green), ITS2 (orange) and the 28S (blue) regions are indicated. The end of the 5.8S gene copies are show, evidencing a nucleotide variation in the last base position. The filled region (blue box) of the 28S gene indicates a highly variable region in the beginning of the 28S rRNA gene. The nucleotide length of the ITSs is indicated.

Figure 4: Maximum likelihood tree of approximately 800 bp conserved segment of R2 elements. The sequences in red, blue and green were derived, respectively, from *Coprophanaeus bellicosus*, *C. cyanescens* and *C. ensifer*. The branch support values are indicated on the nodes. The scale bar indicates the genetic distance.

Figure 5: Southern blot of *Coprophanaeus cyanescens* using a 28S rRNA gene probe generate by PCR with the primers R2-junc_for and R2-HincII_rev (~300 bp). The numbers below the asterisk refer to band intensities (in % adj. vol.). The fragment DNA sizes are indicated on the left.

Figure 1

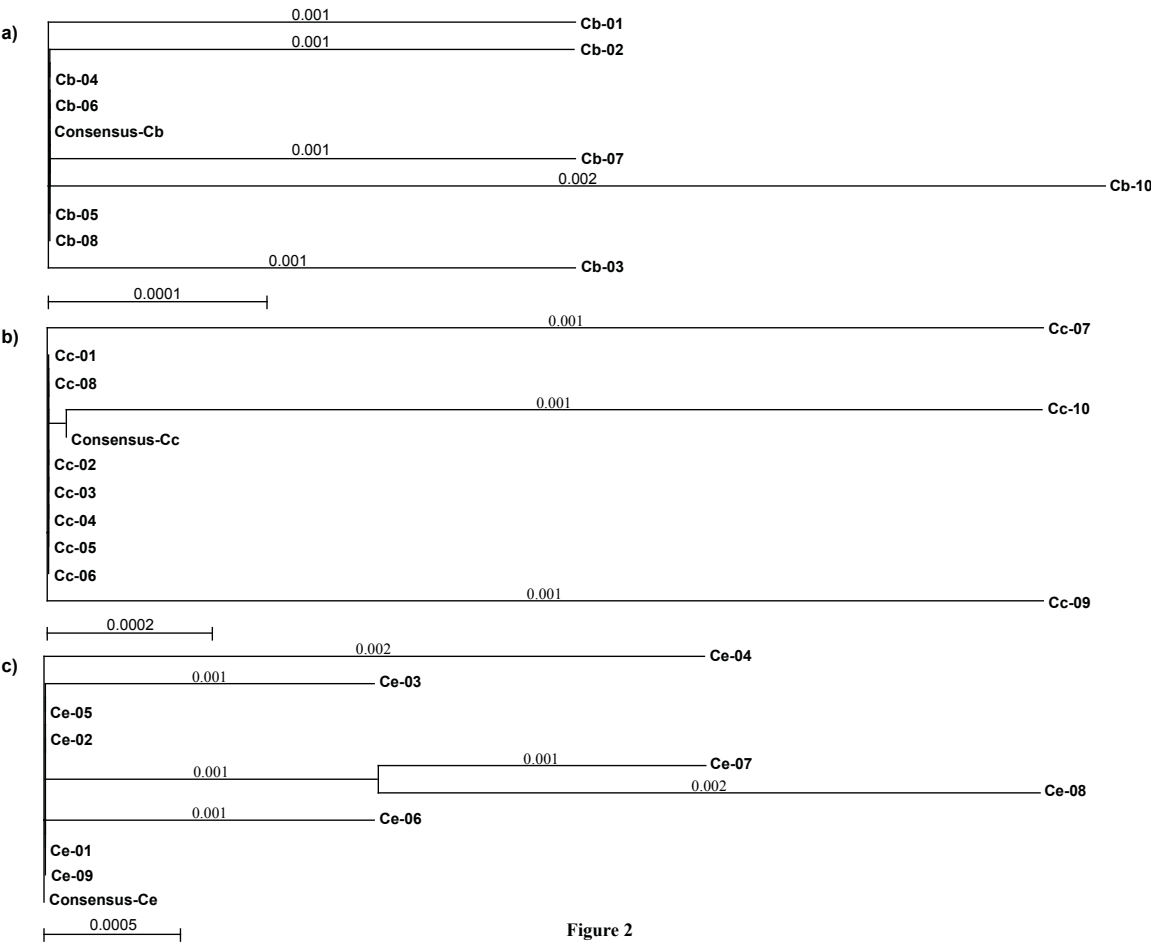


Figure 2

Figure 2

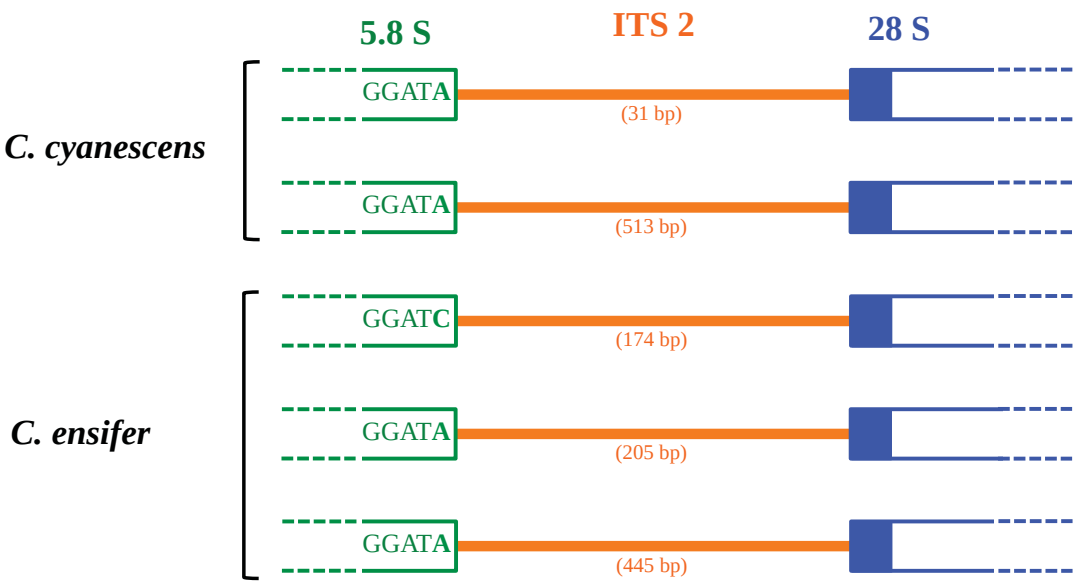


Figure 3

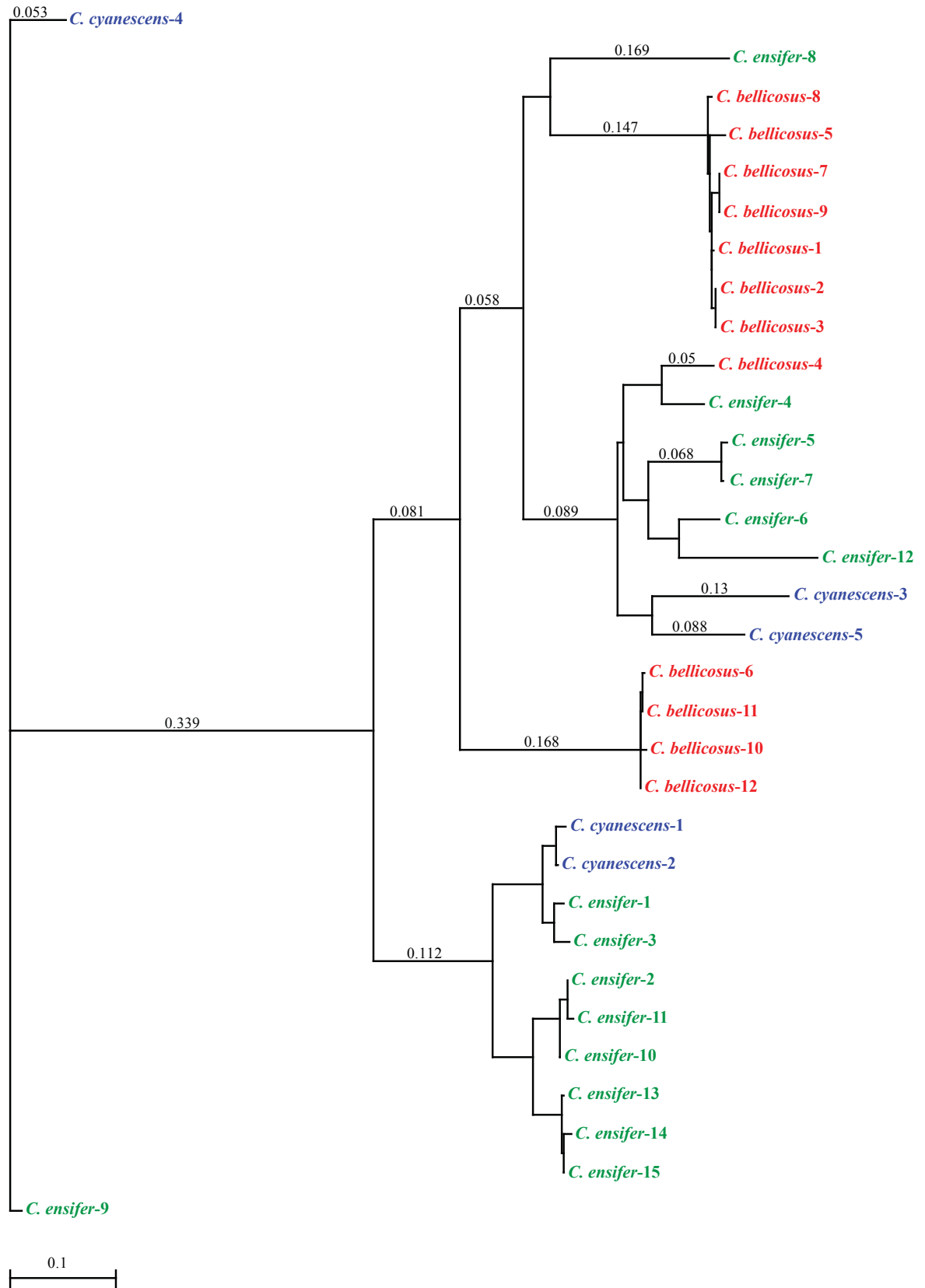


Figure 4

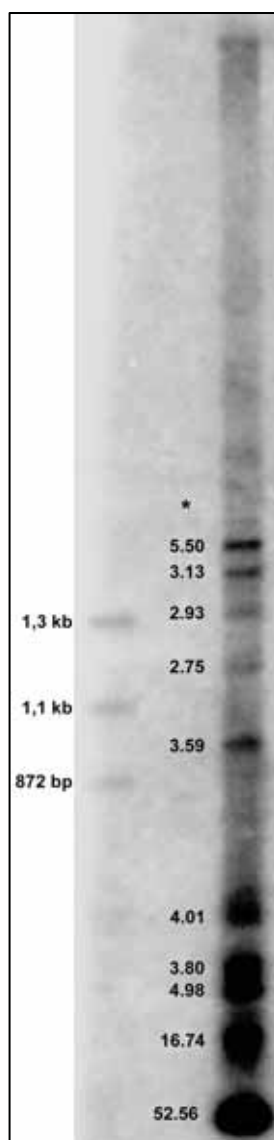


Figure 5

Tables

Table 1: Alignments pairwise similarity matrix of 28 rDNA sequences of approximately 840 bp isolated by PCR with the primer 28S-rev (including the R2 element insertion) the primer R2-up120_for to *Coprophanaeus bellicosus* (Cb).

	Identity scores (%)								
	Cb-1	Cb-2	Cb-3	Cb-4	Cb-5	Cb-6	Cb-7	Cb-8	Cb-9
Cb-1	100.0	99.8	99.8	99.9	99.9	99.9	99.8	99.9	99.6
Cb-2	99.8	100.0	99.8	99.9	99.9	99.9	99.8	99.9	99.6
Cb-3	99.8	99.8	100.0	99.9	99.9	99.9	99.8	99.9	99.6
Cb-4	99.0	99.9	99.9	100.0	100.0	100.0	99.9	100.0	99.8
Cb-5	99.9	99.9	99.9	100.0	100.0	100.0	99.9	100.0	99.8
Cb-6	99.9	99.9	99.9	100.0	100.0	100.0	99.9	100.0	99.8
Cb-7	99.8	99.8	99.8	99.9	99.9	99.9	100.0	99.9	99.6
Cb-8	99.9	99.9	99.9	100.0	100.0	100.0	99.9	100.0	99.8
Cb-9	99.6	99.6	99.6	99.8	99.8	99.8	99.6	99.8	100.0
Consensus-Cb	99.9	99.9	99.9	100.0	100.0	100.0	99.9	100.0	99.8
Similarity scores (%)									

Table 2: Alignments pairwise similarity matrix of 28 rDNA sequences of approximately 850 bp isolated by PCR with the primer 28S-rev (including the R2 element insertion) the primer R2-up120_for to *Coprophanaeus cyanecens* (Cc).

	Identity scores (%)										
	Cc-1	Cc-2	Cc-3	Cc-4	Cc-5	Cc-6	Cc-7	Cc-8	Cc-9	Cc-10	Consensus-Cc
Cc-1	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	99.9	96.2	96.3
Cc-2	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	99.9	96.2	96.3
Cc-3	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	99.9	96.2	96.3
Cc-4	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	99.9	96.2	96.3
Cc-5	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	99.9	96.2	96.3
Cc-6	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	99.9	96.2	96.3
Cc-7	99.9	99.9	99.9	99.9	99.9	99.9	100.0	99.9	99.8	96.1	96.2
Cc-8	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	99.9	96.2	96.3
Cc-9	99.9	99.9	99.9	99.9	99.9	99.9	99.8	99.9	100.0	96.1	96.2
Cc-10	96.2	96.2	96.2	96.2	96.2	96.2	96.1	96.2	96.1	100.0	99.9
Consensus-Cc	96.3	96.3	96.3	96.3	96.3	96.3	96.2	96.3	96.2	99.9	100.0
Similarity scores (%)											

Table 3: Alignments pairwise similarity matrix of 28 rDNA sequences of approximately 850 bp isolated by PCR with the primer 28S-rev (including the R2 element insertion) the primer R2-up120_for to *Coprophanaeus ensifer* (Ce).

	Identity scores (%)								
	Ce-1	Ce-2	Ce-3	Ce-4	Ce-5	Ce-6	Ce-7	Ce-8	Ce-9
Ce-1	100.0	100.0	99.9	99.8	100.0	99.9	99.8	99.6	100.0
Ce-2	100.0	100.0	99.9	99.8	100.0	99.9	99.8	99.6	100.0
Ce-3	99.9	99.9	100.0	99.6	99.9	99.8	99.6	99.5	99.9
Ce-4	99.8	99.8	99.6	100.0	99.8	99.6	99.5	99.4	99.8
Ce-5	100.0	100.0	99.9	99.8	100.0	99.9	99.8	99.6	100.0
Ce-6	99.9	99.9	99.8	99.6	99.9	100.0	99.6	99.5	99.9
Ce-7	99.8	99.8	99.6	99.5	99.8	99.6	100.0	99.6	99.8
Ce-8	99.6	99.6	99.5	99.4	99.9	99.5	99.6	100.0	99.6
Ce-9	100.0	100.0	99.9	99.8	100.0	99.9	99.8	99.6	100.0
Consensus-Ce	100.0	100.0	99.9	99.8	100.0	99.9	99.8	99.6	100.0
Similarity scores (%)									

Table 4: Alignments pairwise similarity matrix of 32 DNA sequences of a conserved region of R2 elements of approximately 800 bp by ClustalW (MacVector program). The sequences were isolated with the primers R2+50_rev and degenerate primers for the regions QGSPLS and AFADD from the species *C. bellicosus* (Cb), *C. cyanescens* (Cc) and *C. ensifer* (Ce). In orange, blue and green are highlighted the comparisons between the sequences, respectively, from the species *C. bellicosus*, *C. cyanescens* and *C. ensifer*.

	Cb-1	Cb-2	Cb-3	Cb-4	Cb-5	Cb-6	Cb-7	Cb-8	Cb-9	Gb-10	Gb-11	Gb-12	Ce-1	Ce-2	Ce-3	Ce-4	Ce-5	Ce-6	Ce-7	Ce-8	Ce-9	Ce-10	Ce-11	Ce-12	Ce-13	Ce-14	Ce-15
Cb-1	100.0	99.4	99.2	61.8	98.2	53.6	99.0	98.9	98.7	98.6	53.4	53.8	25.7	25.5	29.8	35.9	58.8	61.3	60.8	67.2	36.2	45.8	45.4	52.0	44.6	44.8	44.4
Cb-2	99.4	100.0	99.9	61.9	97.9	53.8	98.9	98.7	98.6	54.1	53.7	54.1	25.9	25.6	29.9	35.9	58.4	61.4	60.9	67.0	36.2	46.1	45.7	52.3	44.6	44.8	44.4
Cb-3	99.2	99.9	100.0	61.8	97.7	53.7	98.7	98.6	98.5	53.9	53.6	53.9	25.8	25.7	29.8	35.8	58.4	61.4	60.8	67.1	36.1	45.9	45.6	52.1	44.4	44.7	44.3
Cb-4	61.8	61.9	61.8	100.0	61.3	54.2	61.7	61.8	61.7	54.3	54.1	54.4	23.4	22.8	33.1	34.8	77.7	77.7	77.7	85.5	34.9	45.6	45.3	72.2	45.4	45.6	45.7
Cb-5	98.2	97.9	97.7	61.3	100.0	53.1	97.5	97.6	97.2	53.3	52.9	53.3	25.1	25.0	29.8	35.9	58.9	61.9	60.2	61.2	36.2	45.9	45.6	51.6	44.7	44.9	44.6
Cb-6	53.6	53.8	53.7	54.2	53.1	100.0	53.6	53.6	53.6	99.1	99.6	99.6	28.9	26.5	29.2	35.9	51.9	51.9	52.4	53.6	36.4	44.9	45.0	48.4	45.1	45.6	45.1
Cb-7	99.0	98.9	98.7	61.7	97.5	53.6	100.0	98.4	99.7	53.8	53.4	53.8	25.9	25.6	29.7	35.4	58.9	61.4	60.9	66.8	35.7	45.9	45.6	52.1	44.4	44.6	44.3
Cb-8	98.9	98.7	98.6	61.8	97.6	53.6	98.4	100.0	98.1	53.8	53.4	53.8	25.9	25.6	29.7	35.8	58.4	61.4	60.9	67.3	36.1	46.2	45.8	51.9	44.7	44.7	44.6
Cb-9	98.7	98.6	98.5	61.7	97.2	53.6	98.7	98.1	100.0	53.8	53.4	53.8	25.9	25.6	29.8	36.2	58.9	61.4	60.9	66.5	35.4	45.8	45.4	52.1	44.3	44.9	44.2
Cb-10	53.8	54.1	53.9	54.3	53.3	99.1	53.8	53.8	53.8	100.0	99.0	99.3	26.5	26.1	29.2	36.4	52.3	52.9	52.7	53.6	36.9	45.3	45.4	48.6	45.5	44.6	45.5
Cb-11	53.4	53.7	53.6	54.1	52.9	99.6	53.4	53.4	53.4	99.0	100.0	99.5	26.6	26.3	29.5	35.8	52.1	52.7	52.7	54.8	36.3	45.1	45.3	48.8	45.4	46.0	45.4
Cb-12	53.8	54.1	53.9	54.4	53.3	99.6	53.8	53.8	53.8	99.3	99.5	100.0	26.8	26.3	29.1	36.2	51.9	52.7	53.8	53.1	36.7	45.1	45.3	48.7	45.4	45.9	45.4
Ce-1	25.7	25.9	26.0	23.4	25.1	26.9	25.9	25.9	26.0	26.5	26.6	26.8	100.0	86.2	86.2	86.2	86.2	86.2	86.2	86.2	86.2	86.2	86.2	86.2	86.2	86.2	86.2
Ce-2	25.5	25.6	25.7	22.8	25.0	26.5	25.6	25.6	25.6	26.1	26.2	26.3	86.2	100.0	25.2	19.2	22.8	22.8	22.8	22.8	22.8	22.8	22.8	22.8	22.8	22.8	22.8
Ce-3	29.8	29.9	29.8	33.1	29.8	29.2	29.7	29.7	29.8	29.2	29.5	29.1	25.5	25.2	100.0	21.7	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1	35.1
Ce-4	35.9	35.9	35.8	34.8	35.9	35.9	35.4	35.8	35.2	36.4	35.8	36.2	18.3	19.2	21.7	100.0	34.2	34.2	34.2	34.2	34.2	34.2	34.2	34.2	34.2	34.2	34.2
Ce-5	58.8	58.5	58.4	77.7	58.5	51.9	58.9	58.4	58.9	52.3	52.1	51.9	24.2	22.8	35.1	34.2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Ce-6	26.0	26.1	26.2	23.3	25.7	26.2	26.1	26.1	26.1	25.9	26.0	26.2	85.0	81.2	25.7	19.5	25.3	25.3	25.3	25.3	25.3	25.3	25.3	25.3	25.3	25.3	25.3
Ce-7	45.8	46.1	45.9	45.6	45.8	45.3	45.9	46.2	45.8	45.6	45.5	45.4	27.4	26.9	29.2	33.7	44.9	44.9	44.9	44.9	44.9	44.9	44.9	44.9	44.9	44.9	44.9
Ce-8	26.0	26.1	26.2	23.6	25.6	26.5	25.8	26.1	25.8	26.2	26.2	26.4	89.2	80.7	25.0	19.0	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5
Ce-9	62.5	62.4	62.3	90.9	61.9	55.4	62.0	62.3	62.0	55.7	55.3	55.7	23.4	23.6	33.2	35.4	76.7	76.7	76.7	76.7	76.7	76.7	76.7	76.7	76.7	76.7	76.7
Ce-10	60.7	60.8	60.8	81.5	60.2	52.4	60.8	60.8	60.8	52.7	52.3	52.7	23.6	23.3	33.0	34.6	75.2	75.2	75.2	75.2	75.2	75.2	75.2	75.2	75.2	75.2	75.2
Ce-11	61.3	61.4	61.3	85.5	61.2	53.6	61.4	61.3	61.4	53.6	53.4	53.8	24.6	23.8	32.2	36.1	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8
Ce-12	60.8	60.9	60.8	81.6	60.5	52.8	60.9	60.9	60.9	52.9	52.7	53.1	24.6	23.8	32.2	36.1	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8	76.8
Ce-13	67.2	67.0	66.9	67.5	66.2	54.9	66.6	67.3	66.5	54.9	54.8	55.2	28.5	26.3	29.9	35.1	61.2	61.2	61.2	61.2	61.2	61.2	61.2	61.2	61.2	61.2	61.2
Ce-14	36.2	36.2	36.1	34.9	36.2	36.4	35.7	36.1	35.4	36.9	36.3	36.7	18.5	19.4	21.4	93.6	33.6	33.6	33.6	33.6	33.6	33.6	33.6	33.6	33.6	33.6	33.6
Ce-15	45.8	46.1	45.9	45.6	45.9	44.9	45.9	46.2	45.8	45.3	45.1	45.1	28.1	28.1	29.6	34.1	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3
Ce-16	45.4	45.7	45.6	45.3	45.6	45.0	45.6	45.8	45.4	45.4	45.4	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3	45.3
Ce-17	52.0	52.3	52.1	72.2	51.6	48.4	52.1	51.9	52.1	48.6	48.8	48.7	25.7	24.5	31.6	30.1	70.3	70.3	70.3	70.3	70.3	70.3	70.3	70.3	70.3	70.3	70.3
Ce-18	44.6	44.6	44.4	44.4	44.7	45.1	44.4	44.7	44.3	45.5	45.4	45.4	27.2	27.2	29.6	35.1	44.7	44.7	44.7	44.7	44.7	44.7	44.7	44.7	44.7	44.7	44.7
Ce-19	44.8	44.8	44.7	45.6	44.9	45.6	44.7	44.8	44.6	46.0	45.9	45.9	27.1	27.1	29.7	34.7	43.6	43.6	43.6	43.6	43.6	43.6	43.6	43.6	43.6	43.6	43.6
Ce-20	44.4	44.4	44.3	45.7	44.6	45.1	44.3	44.6	44.2	45.5	45.4	45.4	27.3	27.2	29.5	35.3	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8
Ce-21	44.4	44.4	44.3	45.7	44.6	45.1	44.3	44.6	44.2	45.5	45.4	45.4	27.3	27.2	29.5	35.3	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8	44.8

Table 5: Alignments pairwise similarity matrix of 32 aminoacid sequences of a conserved region of R2 elements of approximately 800 bp by ClustalW (MacVector program). The sequences were isolated with the primers R2+50_rev and degenerate primers for the regions QGSPLS and AFADD from the species *C. bellicosus* (Cb), *C. cyanescens* (Cc) and *C. ensifer* (Ce). In orange, blue and green are highlighted the comparisons between the sequences, respectively, from the species *C. bellicosus*, *C. cyanescens* and *C. ensifer*.

	Cb-1	Cb-2	Cb-3	Cb-4	Cb-5	Cb-6	Cb-7	Cb-8	Cb-9	Gb-10	Gb-11	Gb-12	Ce-1	Ce-2	Ce-3	Ce-4	Ce-5	Ce-6	Ce-7	Ce-8	Ce-9	Ce-10	Ce-11	Ce-12	Ce-13	Ce-14	Ce-15
Cb-1	100.0	99.2	98.9	64.7	99.2	52.4	98.6	98.5	99.2	52.4	52.4	53.2	38.1	35.4	63.5	29.4	64.3	62.4	62.4	63.5	65.4	29.0	39.7	39.3	62.4	40.1	40.1
Cb-2	99.6	100.0	99.6	64.3	96.6	52.4	98.6	98.5	99.2	52.4	52.4	53.2	38.1	35.4	63.5	29.4	64.3	62.4	62.4	63.5	65.4	29.0	39.7	39.3	62.4	40.1	40.1
Cb-3	99.2	99.6	100.0	63.9	96.2	52.1	99.2	98.1	98.9	52.1	52.1	52.8	38.1	35.4	63.2	29.0	63.2	62.4	62.4	63.2	65.0	28.6	39.7	39.3	62.0	40.1	40.1
Cb-4	82.3	82.7	82.3	100.0	62.8	55.1	64.3	63.9	63.9	55.1	55.1	55.8	38.8	38.8	90.2	26.0	89.5	91.4	89.1	89.1	71.4	26.8	41.6	41.2	91.7	40.8	40.8
Cb-5	97.7	97.4	97.0	80.8	100.0	50.9	97.0	96.6	96.6	50.9	50.9	51.7	37.7	35.1	61.3	29.7	62.0	60.2	60.2	61.3	63.9	29.4	39.3	39.0	60.5	39.7	40.1
Cb-6	70.0	69.7	69.3	70.8	68.2	100.0	52.8	52.8	52.4	99.3	98.9	99.3	37.5	34.6	53.6	31.9	53.6	54.7	55.4	55.8	53.9	31.1	39.2	39.9	55.4	39.2	39.2
Cb-7	100.0	99.6	99.2	82.3	97.7	70.0	100.0	98.9	99.6	52.8	52.8	53.6	38.4	35.8	62.8	29.7	63.5	61.7	61.7	62.8	65.0	29.4	40.1	39.7	62.0	40.4	40.4
Cb-8	98.9	98.5	98.1	82.3	97.7	70.0	98.9	100.0	98.9	52.8	52.4	53.2	38.1	35.8	62.8	29.7	63.5	61.7	61.7	62.8	64.7	28.6	39.7	39.3	61.7	40.1	40.1
Cb-9	100.0	99.6	99.2	82.3	97.7	70.0	100.0	98.9	100.0	52.4	52.4	53.2	38.1	35.8	62.8	29.7	63.5	61.7	61.7	62.8	64.7	28.6	39.7	39.3	61.7	40.1	40.1
Cb-10	70.0	69.7	69.3	70.8	68.2	99.3	70.0	70.0	70.0	100.0	98.9	99.3	37.9	34.9	53.6	31.9	53.6	54.7	55.4	55.8	53.9	31.1	39.6	40.3	55.4	39.6	39.6
Cb-11	70.0	69.7	69.3	70.8	68.2	99.3	70.0	70.0	70.0	100.0	98.9	99.3	37.9	34.9	53.6	31.9	53.6	54.7	55.4	55.8	53.9	31.1	39.6	40.3	55.4	39.6	39.6
Cb-12	70.8	70.4	70.0	71.5	68.9	99.3	70.8	70.8	70.8	99.3	99.3	100.0	100.0	98.9	98.9	98.9	98.9	98.9	98.9	98.9	98.9	98.9	98.9	98.9	98.9	98.9	98.9
Cc-1	57.8	58.5	57.8	59.3	57.8	59.9	57.8	57.8	57.8	60.6	60.2	60.2	100.0	86.4	38.1	26.1	37.3	37.3	37.3	37.3	37.3	27.7	27.7	27.7	27.7	27.7	27.7
Cc-2	54.5	54.9	54.5	56.7	54.5	56.1	54.5	54.5	54.5	56.1	56.5	56.5	88.3	100.0	38.1	23.1	37.3	37.3	37.3	37.3	37.3	27.7	27.7	27.7	27.7	27.7	27.7
Cc-3	81.2	81.6	81.2	96.2	79.7	70.8	81.2	81.2	81.2	70.8	70.8	71.5	59.0	56.3	100.0	27.1	95.1	95.1	95.1	95.1	95.1	27.9	27.9	27.9	27.9	27.9	27.9
Cc-4	46.1	45.7	45.4	46.8	46.1	47.8	46.1	46.5	46.1	48.5	48.1	48.1	42.2	37.7	47.2	100.0	27.1	95.1	95.1	95.1	95.1	27.9	27.9	27.9	27.9	27.9	27.9
Cc-5	80.8	80.5	80.1	95.9	79.3	70.4	80.8	80.8	80.8	70.4	70.4	71.2	58.2	55.6	98.1	47.2	100.0	27.1	95.1	95.1	95.1	27.9	27.9	27.9	27.9	27.9	27.9
Ce-1	61.4	61.8	61.4	61.8	61.0	61.8	61.4	61.4	61.4	62.7	62.3	62.3	89.8	82.7	60.7	44.2	59.9	41.6	41.6	41.6	41.6	27.3	27.3	27.3	27.3	27.3	27.3
Ce-2	61.0	61.4	61.0	62.2	60.7	61.6	61.0	61.0	61.0	62.3	61.9	61.9	89.1	82.3	61.0	44.2	60.3	41.6	41.6	41.6	41.6	27.7	27.7	27.7	27.7	27.7	27.7
Ce-3	61.8	62.2	61.8	62.2	61.4	61.6	61.8	61.8	61.8	62.3	61.9	61.9	89.8	82.3	61.0	44.2	60.3	41.6	41.6	41.6	41.6	27.7	27.7	27.7	27.7	27.7	27.7
Ce-4	80.8	81.2	80.8	97.7	79.3	70.4	80.8	80.8	80.8	70.4	70.4	71.2	59.3	56.7	96.6	47.6	95.5	88.0	88.0	88.0	27.3	27.3	27.3	27.3	27.3	27.3	27.3
Ce-5	81.6	82.0	81.6	96.2	80.1	71.2	81.6	81.6	81.6	71.5	71.5	71.9	60.4	57.5	95.1	47.6	94.0	91.4	91.4	91.4	26.4	26.4	26.4	26.4	26.4	26.4	26.4
Ce-6	82.0	82.3	82.0	96.5	80.5	71.9	82.0	82.0	82.0	71.9	71.9	72.7	60.4	57.8	96.2	47.2	95.9	96.6	96.6	96.6	26.8	26.8	26.8	26.8	26.8	26.8	26.8
Ce-7	82.3	82.7	82.3	97.4	80.8	71.9	82.3	82.3	82.3	71.9	71.9	72.7	60.8	58.2	96.2	48.0	95.1	96.5	96.5	96.5	27.1	27.1	27.1	27.1	27.1	27.1	27.1
Ce-8	82.7	83.1	82.7	86.1	81.6	70.0	82.7	82.7	82.7	70.0	70.0	70.8	58.6	56.7	96.2	48.0	95.1	96.5	96.5	96.5	27.1	27.1	27.1	27.1	27.1	27.1	27.1
Ce-9	83.1	83.5	83.1	86.5	81.6	70.0	83.1	83.1	83.1	70.0	70.0	70.8	58.6	56.7	96.2	48.0	95.1	96.5	96.5	96.5	27.1	27.1	27.1	27.1	27.1	27.1	27.1
Ce-10	60.7	60.7	60.3	61.4	59.9	62.3	60.7	60.7	60.7	63.1	62.7	62.7	89.1	82.0	60.3	44.2	59.6	41.6	41.6	41.6	27.3	27.3	27.3	27.3	27.3	27.3	27.3
Ce-11	60.2	60.7	60.3	61.4	59.9	62.3	60.7	60.7	60.7	63.1	62.7	62.7	89.1	82.0	60.3	44.2	59.6	41.6	41.6	41.6	27.3	27.3	27.3	27.3	27.3	27.3	27.3
Ce-12	82.0	82.3	82.0	98.9	80.5	71.9	82.0	82.0	82.0	71.9	71.9	72.7	61.2	58.0	96.6	47.2	96.2	96.6	96.6	96.6	26.8	26.8	26.8	26.8	26.8	26.8	26.8
Ce-13	61.4	61.8	61.4	61.8	61.0	60.8	61.4	61.4	61.4	61.6	61.2	61.2	88.0	80.8	61.0	45.3	60.3	62.5	62.5	62.5	44.2	44.2	44.2	44.2	44.2	44.2	44.2
Ce-14	61.0	61.4	61.0	61.8	61.0	60.8	61.0	61.0	61.0	61.6	61.2	61.2	88.3	81.6	61.8	45.7	61.0	62.5	62.5	62.5	44.6	44.6	44.6	44.6	44.6	44.6	44.6
Ce-15	61.4	61.8	61.4	61.8	61.0	60.8	61.4	61.4	61.4	61.6	61.2	61.2	88.0	80.8	61.0	45.3	60.3	62.5	62.5	62.5	44.2	44.2	44.2	44.2	44.2	44.2	44.2
Similarity Score (%)																											

Additional files

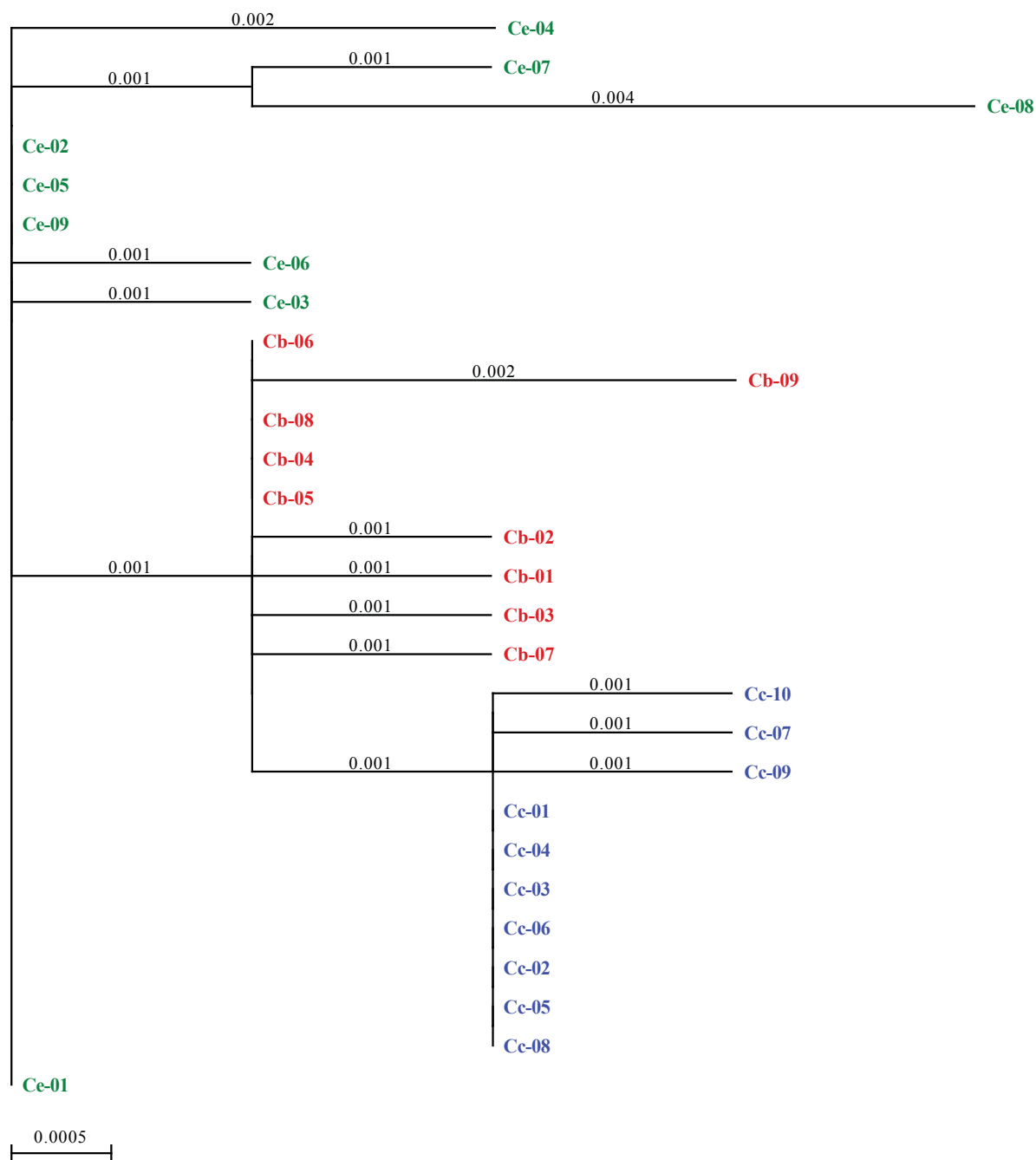
Additional file 1: Multiple sequence alignment of approximately 840 bp of 28 rRNA gene sequences isolated from *Coprophanaeus bellicosus* (Cb), *C. cyanescens* (Cc) and *C. ensifer* (Ce). The analyzed nucleotide sequences were isolated by PCR with the primers 28S-rev (including the R2 insertion region) and R2-up120_for. Asterisks indicate nucleotide position identity among the sequences.

Additional file 2: Maximum likelihood tree based in the analysis of approximately 840 bp of the 28S rRNA gene sequences. The sequences were isolated with the primers 28S-rev and R2-up120_for from *C. bellicosus*, *C. cyanescens* and *C. ensifer* (highlighted in different colors). The values on the nodes and the scale bar indicate the genetic distance.

Additional file 3: Multiple alignments of ITS2 region sequences from *Coprophanaeus cyanescens* (a-b) and *C. ensifer* (c-e). The 5.8S, ITS2 and 28S regions are highlighted in green, orange and blue, respectively. Asterisks indicate nucleotide position identity among the sequences.

Additional file 4: Multiple sequence alignment of approximately 800 bp of 32 R2 sequences isolated by PCR with the primer R2+50_rev and the degenerate primers for the QGDPLS and AFADD regions from *Coprophanaeus bellicosus* (Cb), *C. cyanescens* (Cc) and *C. ensifer* (Ce). Dashes indicate indels.

[illegible]



Additional file 3

a) Cc-4 1 GAACATCGACATTTTGAACGCGACATTGCGGTCTCTCGGATACCTGTTCCCTGGACCACTCCTGGCTGAGGGTCGTTTCAAGGGTCGCGGCCGGAAACGGACGC 100
Cc-5 1 GAACATCGACATTTTGAACGCGACATTGCGGTCTCTCGGATACCTGTTCCCTGGACCACTCCTGGCTGAGGGTCGTTTCAAGGGTCGCGGCCGGAAACGGACGC 100
Cc-6 1 GAACATCGACATTTTGAACGCGACATTGCGGTCTCTCGGATACCTGTTCCCTGGACCACTCCTGGCTGAGGGTCGTTTCAAGGGTCGCGGCCGGAAACGGACGC 100
* * * * *
Cc-4 101 CCGGAAAAATGTGGTGTTAGGGAGGATCCGCCATCCCAGATCGCACGGCGCGCCCAAGTCCTTCTTGAAACGGGGGCCACTCACCCAGAGAGGGTGCCAGGCC 200
Cc-5 101 CCGGAAAAATGTGGTGTTAGGGAGGATCCGCCATCCCAGATCGCACGGCGCGCCCAAGTCCTTCTTGAAACGGGGGCCACTCACCCAGAGAGGGTGCCAGGCC 200
Cc-6 101 CCGGAAAAATGTGGTGTTAGGGAGGATCCGCCATCCCAGATCGCACGGCGCGCCCAAGTCCTTCTTGAAACGGGGGCCACTCACCCAGAGAGGGTGCCAGGCC 200
* * * * *
Cc-4 201 CGTAGAGGCCGCCCGTCGGTAGCGGGAGGATCTCTCTCAGAGTCGGGTTCGTCGAGAGTGCAGCCCTAAGCGGGTGGTAAACTCCATCTTAAGGCTAAATA 300
Cc-5 201 CGTAGAGGCCGCCCGTCGGTAGCGGGAGGATCTCTCTCAGAGTCGGGTTCGTCGAGAGTGCAGCCCTAAGCGGGTGGTAAACTCCATCTTAAGGCTAAATA 300
Cc-6 201 CGTAGAGGCCGCCCGTCGGTAGCGGGAGGATCTCTCTCAGAGTCGGGTTCGTCGAGAGTGCAGCCCTAAGCGGGTGGTAAACTCCATCTTAAGGCTAAATA 300
* * * * *
Cc-4 301 TGACCACGAGACCCGATAGCGAACAAAGTACCGTGAGGGAAAGTT 343
Cc-5 301 TGACCACGAGACCCGATAGCGAACAAAGTACCGTGAGGGAAAGTT 343
Cc-6 301 TGACCACGAGACCCGATAGCGAACAAAGTACCGTGAGGGAAAGTT 343
* * * * *

c) Ce-1 1 GAACATCGACATTTTGAACGCACATTTGCGGTCTCGGATCGTACTACACGCCGGGGCTTCGAACAGTGCCGACTCGTCCCGAACGACGACGACGTCT 100
Ce-6 1 GAACATCGACATTTTGAACGCACATTTGCGGTCTCGGATCGTACTACACGCCGGGGCTTCGAACAGTGCCGACTCGTCCCGAACGACGACGACGTCT 100
Ce-7 1 GAACATCGACATTTTGAACGCACATTTGCGGTCTCGGATCGTACTACACGCCGGGGCTTCGAACAGTGCCGACTCGTCCCGAACGACGACGACGTCT 100

Ce-1 101 CAGACCGAGGTCCTCCCGTACCCGAGAGGTTACGGGCCGGAGCGTTCGCGCGAACGCGCTTATTTACGCTGCAATAACGCGAAAGTAAAAATAATGAA 200
Ce-6 101 CAGACCGAGGTCCTCCCGTACCCGAGAGGTTACGGGCCGGAGCGTTCGCGCGAACGCGCTTATTTACGCTGCAATAACGCGAAAGTAAAAATAATGAA 200
Ce-7 101 CAGACCGAGGTCCTCCCGTACCCGAGAGGTTACGGGCCGGAGCGTTCGCGCGAACGCGCTTATTTACGCTGCAATAACGCGAAAGTAAAAATAATGAA 200

Ce-1 201 CGACACACGCGCTTGGCGGACCTCAGAGCAGGTGAGACTACCCGCGGAATTTAAGCATAATTACTAAGCGGAGGAAAGAAACTAACCAGGATTCCCTA 300
Ce-6 201 CGACACACGCGCTTGGCGGACCTCAGAGCAGGTGAGACTACCCGCGGAATTTAAGCATAATTACTAAGCGGAGGAAAGAAACTAACCAGGATTCCCTA 300
Ce-7 201 CGACACACGCGCTTGGCGGACCTCAGAGCAGGTGAGACTACCCGCGGAATTTAAGCATAATTACTAAGCGGAGGAAAGAAACTAACCAGGATTCCCTA 300

Ce-1 301 GTAGCGGCGAGCGAACAGGGAAGAGCCAGCACCGAATCCCGCGCGGAGACGGACGCGGGAAATGTGTGTAGGGAGGATCCGCCATCCCGAGATC 400
Ce-6 301 GTAGCGGCGAGCGAACAGGGAAGAGCCAGCACCGAATCCCGCGCGGAGAGCGACGCGGGAAATGTGTGTAGGGAGGATCCGCCATCCCGAGATC 400
Ce-7 301 GTAGCGGCGAGCGAACAGGGAAGAGCCAGCACCGAATCCCGCGCGGAGAGCGACGCGGGAAATGTGTGTAGGGAGGATCCGCCATCCCGAGATC 400

Ce-1 401 GCACGGCGCGCCCAAGTCCTTCTTGACGGGGCCACTCACCCAGAGAGGGTGCAGGGCCGCTAGAGCCGCGCTAGCGGGAGGATCTCTCCTCAGA 500
Ce-6 401 GCACGGCGCGCCCAAGTCCTTCTTGACGGGGCCACTCACCCAGAGAGGGTGCAGGGCCGCTAGAGCCGCGCTAGCGGGAGGATCTCTCCTCAGA 500
Ce-7 401 GCACGGCGCGCCCAAGTCCTTCTTGACGGGGCCACTCACCCAGAGAGGGTGCAGGGCCGCTAGAGCCGCGCTAGCGGGAGGATCTCTCCTCAGA 500

Ce-1 501 GTCGGGTTGCTTGAGAGTGCAGCCCTAAGCGGGTGGTAACTCCATCTAAGGCTAAATATGACACGAGACCGATAGCGAACAAAGTACCGTGAGGGAAAG 600
Ce-6 501 GTCGGGTTGCTTGAGAGTGCAGCCCTAAGCGGGTGGTAACTCCATCTAAGGCTAAATATGACACGAGACCGATAGCGAACAAAGTACCGTGAGGGAAAG 600
Ce-7 501 GTCGGGTTGCTTGAGAGTGCAGCCCTAAGCGGGTGGTAACTCCATCTAAGGCTAAATATGACACGAGACCGATAGCGAACAAAGTACCGTGAGGGAAAG 600

Ce-1 601 TT 602
Ce-6 601 TT 602
Ce-7 601 TT 602
**

d) Ce-5 1 GAACATCGACATTTTGAACGCACATTGCGGTCCTCGGATACTGTTTCCCTGGACCACTCCTGGCTGAGGGTCGTTTCAAAATTCACGACTGCCGGGCGTCGA 100
Ce-2 1 GAACATCGACATTTTGAACGCACATTGCGGTCCTCGGATACTGTTTCCCTGGACCACTCCTGGCTGAGGGTCGTTTCAAAATTCACGACTGCCGGGCGTCGA 100

Ce-5 101 TAGCGCGCCGCCACATGTAGCGCGCGCCCGAGCAAGCGATGGCCGTTTCGGGGTAAGCGGCGTGAGTGTTCCTACTACACGCCGGGGCGTTTCGAACAGTG 200
Ce-2 101 TAGCGCGCCGCCACATGTAGCGCGCGCCCGAGCAAGCGATGGCCGTTTCGGGGTAAGCGGCGTGAGTGTTCCTACTACACGCCGGGGCGTTTCGAACAGTG 200

Ce-5 201 CCGACTCGTCCCGAACGACGACGACGAGTCTCAGACCGACGCGGTCTCCCGTACCCGAGAGGTTACGGGCGGAGCGTGCCGCCGGAACGGCGTTATTT 300
Ce-2 201 CCGACTCGTCCCGAACGACGACGACGAGTCTCAGACCGACGCGGTCTCCCGTACCCGAGAGGTTACGGGCGGAGCGTGCCGCCGGAACGGCGTTATTT 300

Ce-5 301 TAGGCTGCAATAACGCGAAGTAAAAATAA TGAACGACACACGCGCTTGGCGGACCTCAGAGCAGGTGAGACTACCCGCCGAATTTAAGCATATTACTAAG 400
Ce-2 301 TAGGCTGCAATAACGCGAAGTAAAAATAA TGAACGACACACGCGCTTGGCGGACCTCAGAGCAGGTGAGACTACCCGCCGAATTTAAGCATATTACTAAG 400

Ce-5 401 CCGAGGAAAGAACTAACCAAGGATTCCTTAGTAGCGGCGAGCGAACAGGGAAGGCCCAGCACCCGAATCCCGCGGCGGAGACGGACGGCCGGGAAATG 500
Ce-2 401 CCGAGGAAAGAACTAACCAAGGATTCCTTAGTAGCGGCGAGCGAACAGGGAAGGCCCAGCACCCGAATCCCGCGGCGGAGACGGACGGCCGGGAAATG 500

Ce-5 501 TGGTGTTAGGGAGGATCCGCGATCCCGAGATCGGACGGGCGCCCAAGTCCTTCTTGAACGGGGCCACTCACCCAGAGAGGTTGCCAGGCCCGTAGAGGC 600
Ce-2 501 TGGTGTTAGGGAGGATCCGCGATCCCGAGATCGGACGGGCGCCCAAGTCCTTCTTGAACGGGGCCACTCACCCAGAGAGGTTGCCAGGCCCGTAGAGGC 600

Ce-5 601 CCGCGGTCGGTAGCGGGAGGATCTCTCTCAGAGTCGGGTTGCTTGAGAGTGCAGCCCTTAAGCGGGTGGTAAACTCCATCTAAGGCTAAATATGACCACGA 700
Ce-2 601 CCGCGGTCGGTAGCGGGAGGATCTCTCTCAGAGTCGGGTTGCTTGAGAGTGCAGCCCTTAAGCGGGTGGTAAACTCCATCTAAGGCTAAATATGACCACGA 700

Ce-5 701 GACCGATAGCGAACAAAGTACCGTGAGGGAAAGTT 734
Ce-2 701 GACCGATAGCGAACAAAGTACCGTGAGGGAAAGTT 734

e) Ce-3 1 GAACATCGACATTTTGAACGCGACATTGCGGTCTCTCGGATACGTGTTCTTGGACCCACTCTCTGGCCGAGGGTCTGTTTCAACATCTACGACTGCGGGGCGCGTC 100
Ce-4 1 GAACATCGACATTTTGAACGCGACATTGCGGTCTCTCGGATACGTGTTCTTGGACCCACTCTCTGGCCGAGGGTCTGTTTCAACATCTACGACTGCGGGGCGCGTC 100

Ce-3 101 AGGCCCGCGACCGCTAGCTAACGGTCCGCGGACGTCGCCGCGCAGGGCAATTCCATGGTCGTTTCGCGCTCGCTTAAACGAATGAGGGGACGTCGACGTAAA 200
Ce-4 101 AGGCCCGCGACCGCTAGCTAACGGTCCGCGGACGTCGCCGCGCAGGGCAATTCCATGGTCGTTTCGCGCTCGCTTAAACGAATGAGGGGACGTCGACGTAAA 200

Ce-3 201 GACGGTTTCGCTAACCGCGCTCGCCCTCTGAGCGCGCTCGACCCGAATCAACGCGCGCTCGGCGCAATGTAAAAATTTGGGCTGCGAAAACGCTCGGGCCGA 300
Ce-4 201 GACGGTTTCGCTAACCGCGCTCGGCCCTCGGAGCGCGTCGACCCGAATCAACGCGCGCTCGGCGCAATGTAAAAATTTGGGCTGCGAAAACGCTCGGGCCGA 300

Ce-3 301 GCGCGGTGACGCTGTGAGAGAAATGCTATCCTACGCTGCGGCCACGCGCGCGTGAAGCAACGCGACGACGAGACGTCAGACGCGGCCGCGCTTCGCT 400
Ce-4 301 GCGCGGTGACGCTGTGAGAGAAATGCTATCCTACGCTGCGGCCACGCGCGCGTGAAGCAACGCGACGACGAGACGTCAGACGCGGCCGCGCTTCGCT 400

Ce-3 401 CTCGAAAGCGCGAGCGTGCAGCGCGGCTAAGCGGAAAAAACGCGGAACGAGCGTAGCTCAGGCGCGCGCTTTTCG 500
Ce-4 401 CTCGAAAGCGCGAGCGTGCAGCGCGGTTAAGCGGAAAAAACGCGGAACGAGCGTAGCTCAGGCGCGCGCTTTTCG 500

Ce-3 501 GCAGGTGAGACTACCCGCTGAACCTAAGCATATTACTAAGCGGAGGAAAAAGAACTAACCCAGGATTCCCTAGTAGCGGCGAGCGAACAGGGAAGAGCCC 600
Ce-4 501 GCAGGTGAGACTACCCGCTGAATTTATGCATATTACTAAGCGGAGGAAAAAGAACTAACCCAGGATTCCCTAGTAGCGGCGAGCGAACAGGGAAGAGCCC 600

Ce-3 601 AGCACCGAATCCCGCGCGGAGACGAGACGCGCGGAAAGTGTGTGTTAGGAGGAGCCCTCCATCCCGAGATCGCACGCGCGGCCCAAGTCTTTGAAC 700
Ce-4 601 AGCACCGAATCCCGCGCGGAGACGAGACGCGCGGAAATGTGTGTTAGGAGGAGCCCTCCATCCCGAGATCGCACGCGCGGCCCAAGTCTTTGAAC 700

Ce-3 701 GGGGCCACTTACCCAGAGAGGGTGCCAGGGCCCGTACGCGGCCGCTCGGTAGCGGGTGGGTCTCTCCTCAGAGTCGGGTGCTTGAGAGTGCAGGCCCTAA 800
Ce-4 701 GGGGCCACTTACCCAGAGAGGGTGCCAGGGCCCGTACGCGGCCGCTCGGTAGCGGGTGGGTCTCTCCTCAGAGTCGGGTGCTTGAGAGTGCAGGCCCTAA 800

Ce-3 801 GCGGGTGGTAAACTCCATCTAAGGCTAAATACGACCGAGACCGGATAGCGAACAAAGTACCGTGAGGGAAAGTT 874
Ce-4 801 GCGGGTGGTAAACTCCATCTAAGGCTAAATATGACCACGAGACCGGATAGCGAACAAAGTACCGTGAGGGAAAGTT 874

Additional file 4

C. bellicosus-1 1 CAAGGCGACCCACTATCTCCAATCCTATTTAACCTTATTTTAGACGAGCTCTTTTGCCAG 60
 C. bellicosus-2 1 CAAGGCGACCCACTATCTCCAATCCTATTTAACCTTATTTTAGACGAGCTCTTTTGCCAG 60
 C. bellicosus-3 1 CAAGGCGACCCACTATCTCCAATCCTATTTAACCTTATTTTAGACGAGCTCTTTTGCCAG 60
 C. bellicosus-4 1 CAAGGCGACCCACTATCGCCAGTCTCTTTAATCTCGTGCTCGATGAATTATTCTGTAAG 60
 C. bellicosus-5 1 CAAGGCGACCCCTATCTCCAATCCTATTTAACTTATTTTAGACGAGCTCTTTTGCCAG 60
 C. bellicosus-6 1 CAGGGTGATCCTTTTTCGCCTGTCCTGTTAACCTGGTGTGGACGAGCTTATTTGTTCT 60
 C. bellicosus-7 1 CAAGGCGACCCACTATCTCCAATCTTATTTAACCTTATTTTAGACGAGCTCTTTTGCCAG 60
 C. bellicosus-8 1 CAAGGCGACCCACTATCTCCAATCCTATTTAACCTTATTTTAGACGAGCTCTCTTGCCAG 60
 C. bellicosus-9 1 CAAGGCGACCCACTATCTCCAATCTTATTTAACCTTATTTTAGACGAGCTCTTTTGCCAG 60
 C. bellicosus-1 1 CAGGGTGATCCTTTTTCGCCTGTCCTGTTAACCTGGTGTGGACGAGCTTATTTGTTCT 60
 C. bellicosus-1 1 CAGGGTGATCCTTTTTCGCCTGTCCTGTTAACCTGGTGTGGACGAGCTTATTTGTTCT 60
 C. bellicosus-1 1 CAGGGTGATCCTTTTTCGCCTGTCCTGTTAACCTGGTGTGGACGAGCTTATTTGTTCT 60
 C. cyanescens-1 1 CAAGGGGATCCGCTCTCACCATACTTGTTAACTTGGTCTGGACGAATTAATGGACCAG 60
 C. cyanescens-2 1 CAGGGGGATCCGCTCTCACCATACTTGTTAACTTGGTCTGGACGAATTAATGGACCAG 60
 C. cyanescens-3 1 CAAGGAGACCCACTATCGCCAGTCTTATTCAACATCGTGCTGGATGAACTGTTTTGCAAA 60
 C. cyanescens-4 1 ----- 0
 C. cyanescens-5 1 CAAGGAGACCCACTATCGCCAGTCTTATTCAACATCGTGCTGGATGAACTGTTTTGTAAA 60
 C. ensifer-1 1 CAGGGGGATCCGCTCTCACCATATTTGTTAACTTGGTCTGGACGAATAATGGACCAA 60
 C. ensifer-2 1 CAGGGGGATCCGCTCTCACCATATCTGTTAACTTGGTCTGGACGAGCTAATGGACCAA 60
 C. ensifer-3 1 CAGGGTGATCCGCTCTCACCATATCTGTTAACTTGGTCTGGACGAGCTAATGGACCAA 60
 C. ensifer-4 1 CAAGGAGACCCACTATCGCCAGTCTCTTTAATCTCGTGCTAGATGAACTATTTTGTAAG 60
 C. ensifer-5 1 CAAGGAGACCCACTATCGCCAGTCTGTTCAATCTCGTGCTAGACGAGCTGTTCTGTAAG 60
 C. ensifer-6 1 CAAGGAGACCCACTATCGCCAGTCTCTTTAATCTTGCTAGATGAGTTATTCTGCAAG 60
 C. ensifer-7 1 CAAGGAGACCCCTATCGCCAGTCTGTTCAATCTCGTGCTAGACGAGCTGTTCTGTAAG 60
 C. ensifer-8 1 CAAGGCGACCCCTATCGCCGGTCTATTTAACTTAGTCTCGACGAGCTGTTTTGCGAT 60
 C. ensifer-9 1 ----- 0
 C. ensifer-10 1 CAGGGTGATCCGTTCTCACCATATCTGTTAACTTGGTCTGGACGAGCTAATGGACCAA 60
 C. ensifer-11 1 CAGGGGGATCCGTTCTCACCATATCTGTTAACTTGGTCTGGACGAGCTAATGGACCAA 60
 C. ensifer-12 1 CAAGGAGACCCACTGTGCGCCAGTCTCTTTAATCTTGCTAGATGAGTTATTCTGCAAG 60
 C. ensifer-13 1 CAGGGTGATCCGTTTTCGCCTTATCTGTTAATTTGGTCTGGATGAACTAATGGACCAA 60
 C. ensifer-14 1 CAGGGTGATCCGTTTTCGCCTTATCTGTTAATTTGGTCTGGATGAATTAATGGACCAA 60
 C. ensifer-15 1 CAGGGGGATCCGTTTTCGCCTTATCTGTTAATTTGGTCTGGATGAATTAATGGACCAA 60

C. bellicosus-1 61 CTAACAAATTCACCAACAGACGGCTTGAAAAATCAATGATACA--AGAATTAATAATCATC 117
 C. bellicosus-2 61 CTAACAAATTCACCAACAGACGGCTTGAAAAATCAATGATACA--AAAATTAATAATCATC 117
 C. bellicosus-3 61 CTAACAAATTCACCAACAGACGGCTTGAAAAATCAATGATACA--AAAATTAATAATCATC 117
 C. bellicosus-4 61 CTCGATTACGAGCAAGCCGGGGAATAACTGTGCAAGAAGAG--CGAGTCAGAATCATC 117
 C. bellicosus-5 61 CTAACAAATTCACCAACAGACGGCTTGAAAAATCAATGATACA--AGAATTAATAATCATC 117
 C. bellicosus-6 61 TTGGAGCAGAACACTCGTTCCGGCATAACCGTCAACGACGACACAAAGGTTACGGCTTTG 120
 C. bellicosus-7 61 CTAACAAATTCACCAACAGACGGCTTGAAAAATCAATGATACA--AAAATTAATAATCATC 117
 C. bellicosus-8 61 CTAACAAATTCACCAACAGACGGCTTGAAAAATCAATGATACA--AAAATTAATAATCATC 117
 C. bellicosus-9 61 CTAACAAATTCACCAACAGACGGCTTGAAAAATCAATGATACA--AAAATTAATAATCATC 117
 C. bellicosus-1 61 TTGGAGCAGAACACTCGTCCGGCATAACCGTCAACGACGACACAAAGGTTACGGCTTTG 120
 C. bellicosus-1 61 TTGGAGCAGAACACTCGTTCCGGCATAACCGTCAACGACGACACAAAGGTTACGGCTTTG 120
 C. bellicosus-1 61 TTGGAGCAGAACACTCGTTCCGGCATAACCGTCAACGACGACACAAAGGTTACGGCTTTG 120
 C. cyanescens-1 61 TTGCCG-----GCAAGGGGATCCGCTCTCACCATACTTGTT--TAACTTGGTCTGGA 111
 C. cyanescens-2 61 TTGCCG-----GCAGGGGGATCCGCTCTCACCATACTTGTT--TAACTTGGTCTGGA 111
 C. cyanescens-3 61 CTCGATTGCGAGTTAACCGGGGAATCACTGTGCAGGATGAG--CGAAGGAGACCCACT 117
 C. cyanescens-4 1 ----- 0
 C. cyanescens-5 61 CTCGATTGCGAGTTAACCGAGGAATCACTGTGCAGGAAGAG--CGAGTCAAAATTATC 117
 C. ensifer-1 61 TTGCCG-----GCAGGGGGATCCGCTCTCACCATATTTGTT--TAACTTGGTCTGGA 111
 C. ensifer-2 61 TTGCCG-----GCGAAGTCGGGGTTGAGTATCGGGGATCAG--AAGGTATGCTGTATG 111
 C. ensifer-3 61 TTGCCG-----GCAGGGTGATCCGCTCTCACCATATCTGTT--TAACTTGGTCTGGA 111
 C. ensifer-4 61 CTCGACTTACGAGCCAGCCGGGGAATAACAGTGCAAGACGAG--CGAGTCAGAATTATC 117
 C. ensifer-5 61 CTCGACACGCGAACAACCTGGGGACTGACAGTGCAAGATGAG--CGAGTCAGAATAATC 117
 C. ensifer-6 61 CTCGATACACGAGCAAACCGGGGACTCACAGTGCAAGACGAG--CGAGTCAGAATAATC 117
 C. ensifer-7 61 CTCGACACGCGAACAACCGGGGACTGACAGTGCAAGATGAG--CGAGTCAGAATAATC 117
 C. ensifer-8 61 TTGGAGAGCCGTACAGGTGGAGGATTAACCGTCAAGGACGAGG--AGAGTCAAAATACTC 117
 C. ensifer-9 1 ----- 0
 C. ensifer-10 61 TTGCCG-----GCGAAGTCGGGGTTGAGTATCGGGGATCAG--AAGGTATGCTGTATG 111
 C. ensifer-11 61 TTGCCG-----GCGAAGTCGGGGTTGAGTATCGGGGATCAG--AAGGTATGCTGTATG 111
 C. ensifer-12 61 CTCGATACACGAGCAAACCGGGGACTCACAGTGCAAGACGAG--CGAGTCAGAATAATC 117
 C. ensifer-13 61 TTGCCG-----GCGACGTCCGGGTTGAGTATCGGGGATCAG--AAGGTATGCTGTATG 111
 C. ensifer-14 61 TTGCCG-----GCGACGTCCGGGTTGAGTATCGGGGATCAG--AAGGTATGCTGTATG 111
 C. ensifer-15 61 TTGCCG-----GCGACGTCCGGGTTGAGTATCGGGGATCAG--AAGGTATGCTGTATG 111

C. bellicosus-1 118 GGATATGCTGACGACTTAGTGATCTTAGATAAATTCTATTACCGGCGCCAGAAAACTTTG 177
 C. bellicosus-2 118 GGATATGCTGACGATTTAGTGATCTTAGATAAATTCTATTACCGGCGCCAGAAAACTTTG 177
 C. bellicosus-3 118 GGATATGCTGACGATTTAGTGATCTTAGATAAATTCTATTACCGGCGCCAGAAAACTTTG 177
 C. bellicosus-4 118 GGCTACGCCGACGATATCTACTACTAGATGACTCGCTCGAAGGCGCCAGAAAGTTATTA 177
 C. bellicosus-5 118 GGATATGCTGACGACTTAGTGATTTTAGATAAATTCTATTACCGGCGCCAGAAAACTTTG 177
 C. bellicosus-6 121 GCATACGCAGACGATCTGCTGATTATGGACGACACCGTGGAGGGTGCAAAGCGGTCCCTT 180
 C. bellicosus-7 118 GGATATGCTGACGACTTAGTGATCTTAGATAAATTCCATTACCGGCGCCAGAAAACTTTG 177
 C. bellicosus-8 118 GGATATGCTGACGACTTAGTAATCTTAGATAAATTCTATCACCGGCGCCAGAAAACTCTG 177
 C. bellicosus-9 118 GGATATGCTGACGACTTAGTGATCTTAGATAAATTCCATTACCGGCGCCAGAAAACTTTG 177
 C. bellicosus-1 121 GCATACGCAGACGATCTGCTGATTATGGACGACACCGTGGAGGGTGCAAAGCGGTCCCTT 180
 C. bellicosus-1 121 GCATACGCAGACGATCTGCTGATTATGGACGACACCGTGGAGGGTGCAAAGCGGTCCCTT 180
 C. bellicosus-1 121 GCATACGCAGACGATCTGCTGATTATGGACGACACCGTGGAGGGTGCAAAGCGGTCCCTT 180
 C. cyanescens-1 112 CGAATTAATGGACCAGTTGCCGGCGGGGTTCGGGGTTGAGTATCGGGGATCAGAAGCTATG 171
 C. cyanescens-2 112 CGAATTAATGGACCAGTTGCCGGCGGGGTTCGGGGTTGAGTATCGGGGATCAGAAGCTATG 171
 C. cyanescens-3 118 ATCGCCAGTCTTATTCAACATCGTGTGGATGAAGTGTTCGAAACTCGATTTGCGAGT 177
 C. cyanescens-4 1 GCCTTCGCGGACGATCTGGTTATCCTGGAGGACAGGGAGGTCAATGTTCCCGCCATCTG 60
 C. cyanescens-5 118 GGCTACGCCGACGATATCTTGCTTATGGATGACTCGCTCGATGGCGCCAGAAAGCTATTA 177
 C. ensifer-1 112 CGAATAATGGACCAATTGCCGGCGAAGTCGGGGTTGAGTATCGGGGATCAGAAGGTATG 171
 C. ensifer-2 112 GCATACGCAGGACGACCTCGTACTCCTTGCTCCTTCGACGTCGGCAATGCGCGAGCTTCTT 171
 C. ensifer-3 112 CGAGCTAATGGACCAATTGCCGGCGAAGTCGGGGTTGAGTATCGGGGATCAGAAGGTATG 171
 C. ensifer-4 118 GGCTATGCCGACGATATCTGCTTTTGGATGACTCGCTCGAGGGCGCCAGAAAATTGCTA 177
 C. ensifer-5 118 GGCTATGCCGATGATATCTACTCATGGACGACTCGCTCGAGGGCGCCAGAAAATTATTA 177
 C. ensifer-6 118 GGCTACGCCGACGATATTATGCTTATGGATGACTCGCTCGAGGGCGCCAGAAAATTACTT 177
 C. ensifer-7 118 GGCTATGCCGATGATATCTACTCATGGACGACTCGCTCGAGGGCGCCAGAAAATTATTA 177
 C. ensifer-8 118 GGCTATGCCGACGATATTTTATTACTCGACGACACTCTCGAGGGCGCCAGGAGATCACTA 177
 C. ensifer-9 1 GCGTTCGCGGACGATCTGGTTATCCTGGAGGACAGGGAGGTCAACGTTCCCGCCATCTG 60
 C. ensifer-10 112 GCATACGCAGGACGACCTCGTACTCCTTGCTCCTTCGACGTCGGCAATGCGCGAGCTTCTT 171
 C. ensifer-11 112 GCATACGCAGGACGACCTCGTACTCCTTGCTCCTTCGACGTCGGCAATGCGCGAGCTTCTT 171
 C. ensifer-12 118 GGCTACGCCGACGATATTATGCTTATGGATGACTCGCTCGAGGGCGCCAGAAAATTACTA 177
 C. ensifer-13 112 GCATACGCAGGACGACCTCGTCTCCTTGCTCCTTCGACGTCGGCAATGCGCGAGCTTCTT 171
 C. ensifer-14 112 GCATACGCAGGACGACCTCGTCTCCTTGCTCCTTCGACGTCGGCAATGCGCGAGCTTCTT 171
 C. ensifer-15 112 GCATACGCAGGACGACCTCGTCTCCTTGCTCCTTCGACGTCGGCAATGCGCGAGCTTCTT 171

C. bellicosus-1 178 AATACTACAACCTGAGTTCCTGAAAGAACGTGGACTAAGTCTGAACCCCTTCCAAATGTATG 237
 C. bellicosus-2 178 AATACTACAACCTGAGTTCCTGAAAGAACGTGGACTAAGTCTGAACCCCTTCCAAATGTATG 237
 C. bellicosus-3 178 AATACTACAACCTGAGTTCCTGAAAGAACGTGGACTAAGTCTGAACCCCTTCCAAATGTATG 237
 C. bellicosus-4 178 AAAGAAGCGGTTGGGTTCTTCGAAAGAACGCGGGTTATGTTTAAACCCAGAAAGTGTGCT 237
 C. bellicosus-5 178 AATACTACAACCTGAGTTCCTGAAAGAACGTGGACTAAGTCTGAACCCCTTCCAAATGTATG 237
 C. bellicosus-6 181 AAGGTCGCGGTAGACCTCTTTGCCCGGAGAGGTCTATCGCTGAATCCTGCCAAATGTATA 240
 C. bellicosus-7 178 AATACTACAACCTGAGTTCCTGAAAGAACGTGGACTAAGTCTGAACCCCTTCCAAATGTATG 237
 C. bellicosus-8 178 AATACTACAACCTGAGTTCCTGAAAGAACGTGGACTAAGTCTGAACCCCTTCCAAATGTATG 237
 C. bellicosus-9 178 AATACTACAACCTGAGTTCCTGAAAGAACGTGGACTAAGTCTGAACCCCTTCCAAATGTATG 237
 C. bellicosus-1 181 AAGGTCGCGGTAGATTCTTTGCCCGGAGAGGTCTATCGCTGAATCCTGCCAAATGTATA 240
 C. bellicosus-1 181 AAGGTCGCGGTAGACTTCTTTGCCCGGAGAGGTCTATCGCTGAATCCTGCCAAATGTATA 240
 C. bellicosus-1 181 AAGGTCGCGGTAGATTCTTTGCCCGGAGAGGTCTATCGCTGAATCCTGCCAAATGTATA 240
 C. cyanescens-1 172 CTGTATGGCATATGCGGACGACCTGTTCTCCTTGCTCCTTCAACGTCGGCAATGCGCGG 231
 C. cyanescens-2 172 CTGTATGGCATATGCGGACGACCTCGTTCTCCTTGCTCCTTCAACGTCGGCAATGCGCGA 231
 C. cyanescens-3 178 TAACCGGGGAATCACTGTGCAGGATGAGCGAGTCAGAATTATCGGCTATGCCGACGATAT 237
 C. cyanescens-4 61 GCCATGACCACCGAGTTTTCCGAGGAGAGGCAATGTCGCTGAACCCCTAGGAAGACTCTC 120
 C. cyanescens-5 178 AGGGAAGCAGTTGGGTTCTTCGATGAACGCGGGTTAAGTCTTAACCCACAAAAGTGTGCT 237
 C. ensifer-1 172 CTGTATGGCATACGCGGATGATCTCGTTCTCCTTGCTCCTTCGACGTCGGCAATGCGCGA 231
 C. ensifer-2 172 CGCAAGTGCACTAAGTTCCTCGAGAAACGCGGGACTGCAACTTAACCCCTTCTAAATGCACG 231
 C. ensifer-3 172 CTGTATGGCATACGCGGATGACCTCGTACTCCTTGCTCCTTCGACGTCGGCAATGCGCGA 231
 C. ensifer-4 178 AAGGAAGCGGTTGGGTTCTTTAAGAAGCAGGGTTATCTTTGAACCCACAAAAGTGTGCT 237
 C. ensifer-5 178 AAGGAAGCAGTTGGGTTTTTCGAAGAAGCAGGGTTAAGCATTAAACCCAAAAAGTGTGCG 237
 C. ensifer-6 178 AAGGAAGCAGTTGGGTTTTTCGAAGAAGCAGGGTTATGTATTAACCCGAAAAAGTGTGCG 237
 C. ensifer-7 178 AAGGAAGCAGTTGGGTTTTTCGAAGAAGCAGGGTTAAGCATTAAACCCAAAAAGTGTGCG 237
 C. ensifer-8 178 AGAGCTGTAGTTGGGTTCTTTGAGAAACGCGGGACTACGCTTAAATCCAGCTAAAAGCGTA 237
 C. ensifer-9 61 GACATGACCACCGAGTTCCTCCGAAGGAGAAGCGTGTCTGCTGAACCCCTAGGAAGACTCTC 120
 C. ensifer-10 172 CGCAAGTGCACTAAGTTCCTCGAGAAACGGGGATTGCAACTTAACCCCTTCTAAATGCACG 231
 C. ensifer-11 172 CGCAAGTGCACTAAGTTCCTCGAGAAACGGGGACTGCAACTTAACCCCTTCTAAATGCACG 231
 C. ensifer-12 178 AAGGAAGCAGTTGGGTTTTTCGAAGAAGCAGGGTTATGTATTAACCCGAAAAAGTGTGCG 237
 C. ensifer-13 172 CGCAAGTGCACTAAGTTCCTCGAGAAACGGGGACTGCAACTTAACCCCTTCTAAATGCACG 231
 C. ensifer-14 172 CGCAAGTGCACTAAGTTCCTCGAGAAACGGGGACTGCAACTTAACCCCTTCTAAATGCACG 231
 C. ensifer-15 172 CGCAAGTGCACTAAGTTCCTCGAGAAACGGGGACTGCAACTTAACCCCTTCTAAATGCACG 231

C. bellicosus-1 238 GCACTTTCCGTGAGCACAGTGCCCGGACGTAAACAGTTATATTCCCATACAACACCATGT 297
 C. bellicosus-2 238 GCACTTTCCGTGAGCACAGTGCCCGGACGTAAACAGTTATATTCCCATACAACACCATGT 297
 C. bellicosus-3 238 GCACTTTCCGTGAGCACAGTGCCCGGACGTAAACAGTTATATTCCCATACAACACCATGT 297
 C. bellicosus-4 238 GCACTAGTATGCAGCACTGTACCCGGGCGTAAGCAATTATATTGCGACACTGTCCCACGG 297
 C. bellicosus-5 238 GCACTTTCCGTGAGCACAGTGCCCGGACGTAAACAGTTATACTCCCATACTACACCATGT 297
 C. bellicosus-6 241 GCTCTCTCGGTGGGGGTTGTGCCGGGTAAGAAGATCTTATATACCCACACAAAGGCACGA 300
 C. bellicosus-7 238 GCACTTTCCGTGAGCACAGTGCCCGGACGTAAACAGTTATATTCCCATACAACACCATGT 297
 C. bellicosus-8 238 GCACTTTCCGTGAGCACAGTGCCCGGACGTAAACAGTTATATTCCCATACAACACCATGT 297
 C. bellicosus-9 238 GCACTTTCCGTGAGCACAGTGCCCGGACGTAAACAGTTATATTCCCATACAACACCATGT 297
 C. bellicosus-1 241 GCTCTCTCGGTGGGGGTTGTGCCGGGTAAGAAGATCTTATATACCCACACAAAGGCACGA 300
 C. bellicosus-1 241 GCTCTCTCGGTGGGGGTTGTGCCGGGTAAGAAGATCTTATATACCCACACAAAGGCACGA 300
 C. bellicosus-1 241 GCTCTCTCGGTGGGGGTTGTGCCGGGTAAGAAGATCTTATATACCCACACAAAGGCACGA 300
 C. cyanescens-1 232 GCTTCTTCGCAAGTGACAGGTTCTTCGAGAAACGGGGACTGCAAATTAACCCCTACTAA 291
 C. cyanescens-2 232 GCTTCTTCGCAAGTGACAGGTTCTTCGAGAAACGGGGACTGCAAATTAACCCCTACTAA 291
 C. cyanescens-3 238 TCTTTTAATGGACGACTCGCTCGATGGCGCCAGAAAGCTACTAAGAGAAGCAGTTGGGTT 297
 C. cyanescens-4 121 TGTTTCTCCGCGGCTTCGGTCGATGGA-----GTCAGTGTCCTCCAGATCCAGATCTAAT 174
 C. cyanescens-5 238 GCACTTGTTGTGCAGCACCGTACCCGGGCGAAAGCAGTTGTAAGTCTCGCACACTGTCTCCCGC 297
 C. ensifer-1 232 GCTTCTTCGCAAGTGACAGGTTCTTCGAGAAACGGGGACTGCAAATTAACCCCTACTAA 291
 C. ensifer-2 232 TCCCTCTCGATGAACACGGTCCCAAAGAAGAAGAACTGTATCCGGTTACCCGCTCTCTG 291
 C. ensifer-3 232 GCTTCTTCGCAAGTGACAGGTTCTTCGAGAAACGGGGACTGCAAATTAACCCCTACTAA 291
 C. ensifer-4 238 GCACTAGTCTGCAGCACGGTACCCGGGCGCAAGCAACTATATTGCGACACCGTCTCACGG 297
 C. ensifer-5 238 GCACTCGCATGTAGTACCGTCTCTGGGCGAAAGCAACTGTACTCGCACACGGTCTCCCGG 297
 C. ensifer-6 238 GCACTAGTGTAGTACCGTACCTGGGCGAAAGCAGCTGTACTCGCACACGGTCTCCCGG 297
 C. ensifer-7 238 GCACTCGCATGTAGTACCGTCTCTGGGCGAAAGCAACTGTACTCGCACACGGTCTCCCGG 297
 C. ensifer-8 238 GCCCTAACTGTTAGTACAGTGCCGGGGAGAAAGCAGCTGTATTCCCACACCGTCCCGAGG 297
 C. ensifer-9 121 TGTTTCTCCGCGGCTTCGGTCGATGGA-----GTCAGTGTCCTCCAGATCCAGATCTAAT 174
 C. ensifer-10 232 TCCCTCTCGATGAACACGGTCCCAAAGAGGAAGAACTGTATCCGGTTACCCGCTCTCTG 291
 C. ensifer-11 232 TCCCTCTCGATGAACACGGTCCCAAAGAAAGAAGAACTGTATCCGGTTACCCGCTCTCTG 291
 C. ensifer-12 238 GCACTAGTGTAGTACCGTACCTGGGCGAAAGCAGCTGTACTCGCACACGGTCTCCCGG 297
 C. ensifer-13 232 TCCCTCTCGATGAACACGGTCCCAAAGAAAGAAGAACTGTATCCGGTAACTCGCTCTCTG 291
 C. ensifer-14 232 TCCCTCTCGATGAACACGGTCCCAAAGAAAGAAGAACTGTATCCGGTAACTCGCTCTCTG 291
 C. ensifer-15 232 TCCCTCTCGATGAACACGGTCCCAAAGAAAGAAGAACTGTATCCGGTAACTCGCTCTCTG 291

C. bellicosus-1 298 TTCAATGTGCAGGGAATAAATATTTTACAAAGTAAATTCAAAGGAATTTTTTAAATACCTC 357
 C. bellicosus-2 298 TTCAATGTGCAGGGAATAAATATTTTACAAAGTAAATTCAAAGGAATTTTTTAAATACCTC 357
 C. bellicosus-3 298 TTCAATGTGCAGGGAATAAATATTTTACAAAGTAAATTCAAAGGAATTTTTTAAATACCTC 357
 C. bellicosus-4 298 TTTACTGTGGGAGGTGTGCAAGTGTCCACAAAGTCTCGCCGGGGGAATCTTTAAATATCTC 357
 C. bellicosus-5 298 TTCAATGTGCAGGGAATAAATATTTTACAAAGTAAATTCAAAGGAATTTTTTAAATACCTC 357
 C. bellicosus-6 301 TTTTTCGTGTCTGGATCGCCCTCTCACAATTGGAGCCCGCGGCTTTTTTTCAGTTACTTG 360
 C. bellicosus-7 298 TTCAATGTGCAGGGAATAAATATTTTACAAAGTAAATTCAAAGGAATTTTTTAAATACCTC 357
 C. bellicosus-8 298 TTCAATGTGCAGGGAATAAATATTTTACAAAGTAAATTCAAAGGAATTTTTTAAATACCTC 357
 C. bellicosus-9 298 TTCAATGTGCAGGGAATAAATATTTTACAAAGTAAATTCAAAGGAATTTTTTAAATACCTC 357
 C. bellicosus-1 301 TTTTTCGTGTCTGGATCGCCCTCTCACAATTGGAGCCCGCGGCTTTTTTTCAGTTACTTG 360
 C. bellicosus-1 301 TTTTTCGTGTCTGGATCGCCCTCTCACAATTGGAGCCCGCGGCTTTTTTTCAGTTACTTG 360
 C. bellicosus-1 301 TTTTTCGTGTCTGGATCGCCCTCTCACAATTGGAGCCCGCGGCTTTTTTTCAGTTACTTG 360
 C. cyanescens-1 292 ATGCACGTCCCTCTCGATGAACACGGTCCCAAAGAAGAAGAACTGTACCCGGTTACCCG 351
 C. cyanescens-2 292 ATGCACGTCCCTCTCGATGAACACGGTCCCAAAGAAGAAGAACTGTACCCGGTTACCCG 351
 C. cyanescens-3 298 CTTTGAAGAACGGGGGTTAAGTCTTAACCCCCAGAAATGTGCTGCACCTGCATGCAGCAC 357
 C. cyanescens-4 175 TTTAAGATCGATGGAATCCCATTCAGATGGTTCGGCAACACAACTTTTAAATACCTA 234
 C. cyanescens-5 298 TTTTGTGTGGGAGGAGTACAGTTGCCCAAGGTTTCGCCCGGGGAATTTTTTAAATATCTC 357
 C. ensifer-1 292 ATGCACGTCCCTCTCGATGAACACGGTCCCAAAAAGAAGAACTGTACCCGGTTACCCG 351
 C. ensifer-2 292 CATTACATCGCGGGCAAGGGAGTGAACAGATCGCGCTGGGGATTTATTTAAGTATCTC 351
 C. ensifer-3 292 ATGCACGTCCCTCTCGATGAACACGGTCCCAAAGAGGAAGAACTGTACCCGGTTACCCG 351
 C. ensifer-4 298 TTTAGTGTGGGAGGTGTGCAAGTGTCCACAAAGTTTCGCCCGGGGAATCTTTAAATATCTC 357
 C. ensifer-5 298 TTCTCTGTGGGAGGCGTACAATTGCCGCAAGTTTCGCCCTGGGGACTTTTTCAAATATCTC 357
 C. ensifer-6 298 TTCAACGTGGGAGGCGTACAGTTGCCACAGGTTTCGCCCGGGGAGTTCTTTCAAATATCTC 357
 C. ensifer-7 298 TTCTCTGTGGGAGGCGTACAGTTGCCGCAAGTTTCGCCCTGGGGACTTTTTCAAATATCTC 357
 C. ensifer-8 298 TTTACTGTGAACGGGACGACGCGCCACAAATTTCCCAGGAGAATTTTAAATATCTG 357
 C. ensifer-9 175 TTTAAGATCGATGGCGTCCCATTCAGATGGTTCGGCAACACTAACACTTTTTAAATACCTG 234
 C. ensifer-10 292 CATTACATCGCGGGCAAGGGAGTGAACAGATCGCGCTGGGGATTTATTTAAGTATCTC 351
 C. ensifer-11 292 CATTACATCGCGGGCAAGGGAGTGAACAGATCGCGCTGGGGATTTATTTAAGTATCTC 351
 C. ensifer-12 298 TTCAACGTGGGAGGCGTACAGTTGCCACAGGTTTCGCCCGGGGAGTTCTTTCAAATATCTC 357
 C. ensifer-13 292 CATTACATAGCGGGTAAGGGAGTGAACAGATCGCGCTGGGGATTTATTTAAGTACCTC 351
 C. ensifer-14 292 CACTACATAGCGGGTAAGGGAGTGAACAGATCGCGCTGGGGATTTATTTAAGTACCTC 351
 C. ensifer-15 292 CATTACATAGCGGGTAAGGGAGTGAACAGATCGCGCTGGGGATTTATTTAAGTACCTC 351

C. bellicosus-1 358 GGCTCACGTTACAACCTTTATGGGGCAAATCGCACCAACATTGGATTCAATTAGAACTCAA 417
 C. bellicosus-2 358 GGCTCACGTTACAACCTTTATGGGGCAAATCGCACCAACATTGGATTCAATTAGAACTCAA 417
 C. bellicosus-3 358 GGCTCACGTTACAACCTTTATGGGGCAAATCGCACCAACATTGGATTCAATTAGAACTCAA 417
 C. bellicosus-4 358 GGGGCCAAATATAATTACATGGGACAAAATTCCTCGGCTCGAGAGCGTTTCAACCCAA 417
 C. bellicosus-5 358 GGCTCACGTTACAACCTTTATGGGGCAAATCGCACCAACATTGGATTCAATTGGAGCTCAA 417
 C. bellicosus-6 361 GGCCGCAGCTATAACTTCATGGGACGTCCATCTGTATCCCATGAGAACACCAAGCTTATG 420
 C. bellicosus-7 358 GGCTCACGTTACAACCTTTATGGGGCAAATCGCACCAACATTGGATTCAATTAGAACTCAA 417
 C. bellicosus-8 358 GGCTCACGTTACAACCTTTATGGGGCAAATCGCACCAACATTGGATTCAATTAGAACTCAA 417
 C. bellicosus-9 358 GGCTCACGTTACAACCTTTATGGGGCAAATCGCACCAACATTGGATTCAATTAGAACTCAA 417
 C. bellicosus-1 361 GGCCGCAGCTATAACTTCATGGGACGTCCATCTGTATCCCATGAGAACACCAAGCTTATG 420
 C. bellicosus-1 361 GGCCGCAGCTATAACTTCATGGGACGTCCATCTGTATCCCATGAGAACACCAAGCTTATG 420
 C. cyanescens-1 352 TTCTCTGCATTACATCGCGGGCAAGGGAGTGAAACAGATCGCGCCTGGGGATTTATTTAA 411
 C. cyanescens-2 352 TTCTCTGCATTACATCGCGGGCAAGGGAGTGAAACAGATCGCGCCTGGGGATTTATTTAA 411
 C. cyanescens-3 358 AGTACCCGGGCGAAAGCAACTGTACTCGCACACAGTCTCCCGTTTCAGCGTGGGAGGAGT 417
 C. cyanescens-4 235 GGACAAGGTTACACACTATCTGGACTATCGAAGCCATCTTTGGCTAACTTATCGTCGTGG 294
 C. cyanescens-5 358 GGGGCCAAATACAACCATATGGGGCAAAAACCCCTGGCCTCGATAGCGTTCGAACCCAG 417
 C. ensifer-1 352 TTCTCTGCATTACATTGCGGGCAAGGGAGTGAAACAGATCGCGCCTGGGGATTTATTTAA 411
 C. ensifer-2 352 GGGCTCCACTACTCTTATGACGGCAGCGTTAATTTTAAAGTTCAAACCACCATCGAAAGC 411
 C. ensifer-3 352 TTCTCTGCATTACATAGCGGGCAAGGGAGTGAAACAGATCGCGCCTGGGGATTTCTTTAA 411
 C. ensifer-4 358 GGGGCCAAATACAACCTACATGGGACAAAATTCCTCGGCTCGAGAGCGTTTCAACCCAA 417
 C. ensifer-5 358 GGGGCCAAATACAACCTACATGGGACAAAATTCCTCGGCTCGAGAACCTACGAACCCAA 417
 C. ensifer-6 358 GGGGCCAAATATCACTACATGGGACAAAATTCCTCGGCTCGAGAACTGTTCAACCCAG 417
 C. ensifer-7 358 GGGGCCAAATACAACCTACATGGGACAAAATTCCTCGGCTCGAGAACCTACGAACCCAA 417
 C. ensifer-8 358 GGCTCTAAATATAATCATATGGGACATATCGGCCAGTTTAGCCAGGTAAGAACTCAG 417
 C. ensifer-9 235 GGACAAGGTTACACACTATCTGGACTATCGAAGCCTTCTTTGGCAAACCTTATCGTCGTGG 294
 C. ensifer-10 352 GGGCTCCACTACTCTTATGACGGCAGCGTTAATTTTAAAGTTCAAACCACCATCGTAAGC 411
 C. ensifer-11 352 GGGCTCCACTACTCTTATGACGGCAGCGTTAATTTTAAAGTTCAAACCACCATCGAAAGC 411
 C. ensifer-12 358 GGGGCCAAATATCACTACATGGGACAAAATTCCTCGGCTCGAGAACTGTTCAACCCAG 417
 C. ensifer-13 352 GGGCTCCACTATTCTATGACGGCAGCGTTAATTTTAAAGTTCAAACCACCATCGAAAGC 411
 C. ensifer-14 352 GGGCTCCACTATTCTATGACGGCAGCGTTAATTTTAAAGTTCAAACCACCATCGAAAGC 411
 C. ensifer-15 352 GGGCTCCACTATTCTATGACGGCAGCGTTAATTTTAAAGTTCAAACCACCATCGAAAGC 411

C. bellicosus-1 418 TTGAATCGAATACAAATCGCACCCCTAAAACCTCATCAGAAAGTAGAACTGGTGAGGTCG 477
 C. bellicosus-2 418 TTGAATCGAATACAAATCGCACCCCTAAAACCTTATCAGAAAGTAGAACTGGTGAGATCG 477
 C. bellicosus-3 418 TTGAATCGAATACAAATCGCACCCCTAAAACCTTATCAGAAAGTAGAACTGGTGAGATCG 477
 C. bellicosus-4 418 CTCGCCCCGATTACGCGAGCTCCACTAAAACCATTTCAAAAAGTGGAGCTGGTTAGATCG 477
 C. bellicosus-5 418 TTGAGTCGAATACAAATCGCACCCCTAAAACCTCATCAGAAAGTAGAACTGGTGAGATCG 477
 C. bellicosus-6 421 CTCTTGCGTATAAGCAAGGCCCACTAAAGCCCCACCAAAAAGTCGAGTTTGTGAGATCG 480
 C. bellicosus-7 418 TTGAATCGAATACAAATCGCACCCCTAAAACCTCATCAGAAAGTAGAACTGGTGAGATCG 477
 C. bellicosus-8 418 TTGAGTCGAATACAAATCGCACCCCTAAAACCTCATCAGAAAGTAGAACTGGTGAGATCG 477
 C. bellicosus-9 418 TTGAATCGAATACAAATCGCACCCCTAAAACCTCATCAGAAAGTAGAACTGGTGAGATCG 477
 C. bellicosus-1 421 CTCTTGCGTATAAGCAAGGCCCACTAAAGCCCCACCAAAAAGTCGAGTTTGTGAGATCG 480
 C. bellicosus-1 421 CTCTTGCGTATAAGCAAGGCCCACTAAAGCCCCACCAAAAAGTCGAGTTTGTGAGATCG 480
 C. bellicosus-1 421 CTCTTGCGTATAAGCAAGGCCCACTAAAGCCCCACCAAAAAGTCGAGTTTGTGAGATCG 480
 C. cyanescens-1 412 GTATCTCGGGTTCCACTACTCTTATGACGGCACCCTTAATTTTAAAGTTCAAACCACCAT 471
 C. cyanescens-2 412 GTATCTCGGGTTCCACTACTCTTATGACGGCACCCTTAATTTTAAAGTTCAAACCACCAT 471
 C. cyanescens-3 418 GCAGTTGCCTCAAGTTTCGCCCCGGGAATTTTCAAGTACCTCGGGGCCAAATATAACCA 477
 C. cyanescens-4 295 GTCTCCAATCTTTCTTCAGCCCCCTTAAAGCCTCAGCAAAAAGTGACTGTTTACGGGAT 354
 C. cyanescens-5 418 CTAGCCCGCATCCACACAGCTCCACTAAAGCCGCATCAAAAAGTTGAGCTGGTAAGGTCG 477
 C. ensifer-1 412 GTATCTCGGGTTCCACTACTCTTATGACGGCACCCTTAATTTTAAAGTTCAAACCACCAT 471
 C. ensifer-2 412 CTGAGGAGAGTGGAAGAGGCCCTCTGAAACCTTTTCAAGAGCTCAATATCCTGAGATAC 471
 C. ensifer-3 412 GTATCTCGGGTTCCACTACTCTTATGACGGCAACGTTAATTTTAAAGTTCAAACCACCAT 471
 C. ensifer-4 418 CTCGCCCCGATTACGCGAGCTCCACTAAAGCCTTTTCAAGAAAGTGGGGCTGGTTAGGTCG 477
 C. ensifer-5 418 CTTGCGCGTGTAAGCCGCTCCACTTAAAGCCTATCAGAAAGTGGAGTTGGTAAGATCA 477
 C. ensifer-6 418 CTTGCCCGTATACATGCAGCTCCACTCAAAACCATATCAAAAAGTGGAGCTCGTAAGATCA 477
 C. ensifer-7 418 CTTGCGCGTGTAAGCCGCTCCACTTAAAGCCTATCAGAAAGTGGAGTTGGTAAGATCA 477
 C. ensifer-8 418 CTCAGCCGAATACAAAATCGCCGCTGAAACCATACCAGAAATTAGAATTGGTCAGGTCG 477
 C. ensifer-9 295 GTCTCCAATCTTTCTTCAGCCCCCTTAAAGCCTCAGCAAAAAGTGACTATTTTACGGGAT 354
 C. ensifer-10 412 CTGAGGAGAGTGGAAGAGGCCCTCTGAAACCTTTTCAAGAGCTCAATATCCTGAGATAC 471
 C. ensifer-11 412 CTGAGGAGAGTGGAAGAGGCCCTCTGAAACCTTTTCAAGAGCTCAATATCCTGAGATAC 471
 C. ensifer-12 418 CTTGCCCGTATACATGCAGCTCCACTCAAAACCATATCAAAAAGTGGAGCTCGTAAGATCA 477
 C. ensifer-13 412 CTGAGGAAAGTGGAAGAGGCCCTCTGAAAGCCTTTTCAAGAGCTCAACATCCTGAGGTAC 471
 C. ensifer-14 412 CTGAGGAAAGTGGAAGAGGCCCTCTGAAAGCCTTTTCAAGAGCTTAACATCCTGAGGTAC 471
 C. ensifer-15 412 CTGAGGAAAGTGGAAGAGGCCCTCTGAAAGCCTTTTCAAGAGCTTAACATCCTGAGGTAC 471

C. bellicosus-1 478 CATCTTATACCACGACTTTTATTCCAATGGCAGACACCCCGTGTAGCAGAGTAGCACTT 537
 C. bellicosus-2 478 CATCTTATACCACGACTTTTATTCCAATGGCAGACACCCCGTGTAGCAGAGTAGCACTT 537
 C. bellicosus-3 478 CATCTTATACCACGACTTTTATTCCAATGGCAGACACCCCGTGTAGCAGAGTAGCACTT 537
 C. bellicosus-4 478 CACCTAATACCCCGCCTACTATTTTCAGCTGCAGACCCCGAGGATAAATCGTAAGATGCTC 537
 C. bellicosus-5 478 CATCTTATACCACGACTTTTATTCCAATGGCAGACACCCCGTGTAGCAGAGTAGCACTT 537
 C. bellicosus-6 481 TTTCTTATCCCTCGACTCCTCCATCTCTTACAGACACCAAGAACGACGAAGGCCATCTTA 540
 C. bellicosus-7 478 CATCTTATACCACGACTTTTATTCCAATGGCAGACACCCCGTGTAGCAGAGTAGCACTT 537
 C. bellicosus-8 478 CATCTTATACCACGACTTTTATTCCAATGGCAGACACCCCGTGTAGCAGAGTAGCACTT 537
 C. bellicosus-9 478 CATCTTATACCACGACTTTTATTCCAATGGCAGACACCCCGTGTAGCAGAGTAGCACTT 537
 C. bellicosus-1 481 TTTCTTATCCCTCGACTCCTCCATCTCTTACAGACACCAAGAACGACGAAGGCCATCTTA 540
 C. bellicosus-1 481 TTTCTTATCCCTCGACTCCTCCATCTCTTACAGACACCAAGAACGACGAAGGCCATCTTA 540
 C. bellicosus-1 481 TTTCTTATCCCTCGACTCCTCCATCTCTTACAGACACCAAGAACGACGAAGGCCATCTTA 540
 C. cyanescens-1 472 TGAAAGCCTGAGGAGAGTGGAAAGAGCACCTCTGAAACCTTTTCAGAAGCTCAATATCCT 531
 C. cyanescens-2 472 TGAAAGCCTGAGGAGAGTGGAAAGAGCACCTCTGAAACCTTTTCAGAAGCTCAATATCCT 531
 C. cyanescens-3 478 TATGGGGCAAAACTCCCCAGGCCTCGACAGCGTTAGAACCCTCGCCCGCATACATGC 537
 C. cyanescens-4 355 TACTTAGTTCCTCGTTTACTCTATGGGCTGCAGACTCCTGCGACCACTGGTACTCTCCTC 414
 C. cyanescens-5 478 CATCTTATACCCCGATTGTTATTTTCAGCTGCAGACCCCTCGGGTTAATCGTGGGATGCTT 537
 C. ensifer-1 472 CGAAAGCCTGAGGAGAGTGGAAAGAGCCCTCTGAAACCTTTTCAGAAGCTCAATATCCT 531
 C. ensifer-2 472 TATACCCTCCCGAGATTACGCACGCATTGCAAAACCCCTCGTGTGGACGGCAAAATACTC 531
 C. ensifer-3 472 CGAAAGCCTGAGGAGAGTGGAAAGAGCTCCTCTGAAACCTTTTCAGAAGCTCAATATCCT 531
 C. ensifer-4 478 CACCTAATTCCCGCTTAATATTCCAGCTGCAGACCCCGAGGATAAATCGTAAGATGCTT 537
 C. ensifer-5 478 CATCTCGTACCCCGCCTTTTGTTCAGCTGCAGACCCCGAGGATTAATCGTAGGATGCTC 537
 C. ensifer-6 478 CATCTTGTACCCCGCCTTTTATTCCAGCTGCAGACCCCAAGGATAAACCGTAGGATGCTG 537
 C. ensifer-7 478 CATCTCGTACCCCGCCTTTTGTTCAGCTGCAGACCCCGAGGATTAATCGTAGGATGCTC 537
 C. ensifer-8 478 CATTTAATTCCCGGACTGATGTTTCAATGGCAGACACCCCGAATGAATAAGAAGGCACTT 537
 C. ensifer-9 355 TACCTAGTTCCTCGTTTACTTTATGGGCTGCAGACTCCTGCGACGACTGGCACTCTCCTC 414
 C. ensifer-10 472 TATACCCTCCCGAGATTACGCACGCATTGCAAAACCCCTCGTGTGGACGGGAAAACACTC 531
 C. ensifer-11 472 TATACCCTCCCGAGATTACGCACGCATTGCAAAACCCCTCGTGTGGACGGGAAAATACTC 531
 C. ensifer-12 478 CATCTTGTACCCCGCCTTTTATTCCAGCTGCAGACCCCAAGGATAAACCGTAGGATGCTG 537
 C. ensifer-13 472 TACACCCTCCCGAGGATAACGCACGCATTGCAAAACCCCTCGAGTGGACGGGAAAACCCCTC 531
 C. ensifer-14 472 TACACCCTCCCGAGGATAACGCACGCATTGCAAAACCCCTCGAGTGGACGGGAAAACCCCTC 531
 C. ensifer-15 472 TACACCCTCCCTAGGATAACGCATGCATTGCAAAACCCCTCGAGTGGACGGGAAAACCCCTC 531

C. bellicosus-1 538 AAAGGGATAGACAGGATTATCCGCTTAAATGTCAGGAAATTTCTCCACTTAAATAAAAGC 597
 C. bellicosus-2 538 AAAGGGATAGACAGGATTATCCGCTTAAATGTCAGGAAATTTCTCCACTTAAATAAAAGC 597
 C. bellicosus-3 538 AAAGGGATAGACAGGATTATCCGCTTAAATTTTCAGGAAATTTCTCCACTTAAATAAAAGC 597
 C. bellicosus-4 538 AAATCGGTTCGATCGGAGTCTCCGCTTATATGTGCGAAAATTTGTGCACCTGAACAAAGAGC 597
 C. bellicosus-5 538 AAAGGGATAGACAGGATTATCCGCTTAAATGTCAGGAAATTTCTCCACTTAGATAAAAGC 597
 C. bellicosus-6 541 TCGTCGGTTGATCGATTAATTCGAATATCTGTTAGGCGTTTCCTTCACTTGAATAAGACG 600
 C. bellicosus-7 538 AAAGGGATAGACAGGATTATCCGCTTAAATGTCAGGAAATTTCTCCACTTAAATAAAAGC 597
 C. bellicosus-8 538 AAAGGGATAGACAGGATTATCCGCTTAAATGTCAGGAAATTTCTCCACTTAAATAAAAGC 597
 C. bellicosus-9 538 AAAGGGATAGACAGGATTATCCGCTTAAATGTCAGGAAATTTCTCCACTTAAATAAAAGC 597
 C. bellicosus-1 541 TCGTCGGTTGATCGATTAATTCGAATATCTGTTAGGCGTTTCCTTCACTTGAATAAGACG 600
 C. bellicosus-1 541 TCGTCGGTTGATCGATTAATTCGAATATCTGTTAGGCGTTTCCTTCACTTGAATAAGACG 600
 C. bellicosus-1 541 TCGTCGGTTGATCGATTAATTCGAATATCTGTTAGGCGTTTCCTTCACTTGAATAAGACG 600
 C. cyanescens-1 532 GAGATACTACACCCTCCCGAGGTTTACGCATGCATTGCAAAACCCCTCGTGTGGACGGGAA 591
 C. cyanescens-2 532 GAGATACTACACCCTCCCGAGGTTTACGCATGCATTGCAAAACCCCTCGTGTGGACGGGAA 591
 C. cyanescens-3 538 AGCTCACTCAAGCTATACCAAAAGGAGTGGAGCTGGTAAGGTGCGATCTTATACCCGATT 597
 C. cyanescens-4 415 AAGGATGCAGATAAGCTTATCAGAGGTTTTGTTAAAGAGCTTTACATCTACACCTTCAT 474
 C. cyanescens-5 538 AAATCGGTGGATAGGAGTCTCCGCTTTTATGTACGCAATTTATACACCTAAATAAGAGC 597
 C. ensifer-1 532 GAGATACTATACCCTCCCGAGGTTTATGCATGCATTGCAAAACCCCTCGTGTGGACGGGAA 591
 C. ensifer-2 532 CGTGCCGTGGACCGTGTAGTGCGGATCTCCGTAAGGAAGATCATGCATTGCAAAAGTCC 591
 C. ensifer-3 532 GAGATACTATACCCTCCCTAGATTACGCATGCATTGCAAAACCCCTCGTGTGGACGGGAA 591
 C. ensifer-4 538 AAATCGGTTCGATCGGAGTCTCCGCTTATGTGCGAAAATTCGTACACTTGAATAAGAGC 597
 C. ensifer-5 538 AAATCGGTGGACAGGATTCTCCGCTGCATATACGAAAGTTTCGTACACTGAACAAGAGC 597
 C. ensifer-6 538 AAGTCGGTGGATCGGAGTCTCCGCTTATATGTACGAAAATTCGTACATTGAATAAGAGC 597
 C. ensifer-7 538 AAATCGGTGGACAGGATTCTCCGCTGCATGTACGAAAGTTTCGTACACTGAACAAGAGC 597
 C. ensifer-8 538 AAGAGCATCGACAGGGTTATCCGCTAAATGTCAGAAAATTTATTCACTTCAACAAGAGC 597
 C. ensifer-9 415 AAGGATGCAGATAAACTTATCAGAGGCTTTGTTAAAGAGCTTTACATTTACATCTTAC 474
 C. ensifer-10 532 CGTGCCGTGGACCGTGTAGTGCGGATCTCCGTAAGGAAGATCATGCATTGCAAAAGTCC 591
 C. ensifer-11 532 CGTGCCGTGGACCGTGTAGTGCGGATCTCCGTAAGGAAGATCATGCATTGCAAAAGTCC 591
 C. ensifer-12 538 AAGTCGGTGGATCGGAGTCTCCGCTTATATGTACGAAAATTCGTACATTGAATAAGAGC 597
 C. ensifer-13 532 CGTGCCGTGGACCGTGTAGTGCGGATCTCCGTTAGGAAGATCCTGCATTGCAAAAGTCC 591
 C. ensifer-14 532 CGTGCCGTGGACCGTGTAGTGCGGATCTCCGTTAGGAAGATCCTGCATTGCAAAAGTCC 591
 C. ensifer-15 532 CGTGCCGTGGACCGTGTAGTGCGGATCTCCGTTAGGAAGATCCTGCATTGCAAAAGTCC 591

C. bellicosus-1 598 AGCCCCGATGCATTTATTCATGCCTCCATCCGTGAGGGCGGTCTTGGAAATCATA---TCC 654
 C. bellicosus-2 598 AGCCCTGATGCATTTATTCATGCCTCCATCCGTGAGGGCGGTCTTGGAAATCATA---TCC 654
 C. bellicosus-3 598 AGCCCTGATGCATTTATTCATGCCTCCATCCGTGAGGGCGGTCTTGGAAATCATA---TCC 654
 C. bellicosus-4 598 AGCCCCGACGCGTTTCATCCATGCGCCCGTAAAAGAAGGCGGGCTTGGAAATTA---TCG 654
 C. bellicosus-5 598 AGCCCCGATGCATTTATTCATGCCTCCATCCGTGAGGGCGGTCTTGGAAATCATA---TCC 654
 C. bellicosus-6 601 ACAGCAGATGCCTTTATCCACGCTCCCATAGGGAGGGAGGATTGGGCATCCTA---TCA 657
 C. bellicosus-7 598 AGCCCCGATGCATTTATTCATGCCTCCATCCGTGAGGGCGGTCTTGGAAATCATA---TCC 654
 C. bellicosus-8 598 AGCCCCGATGCATTTATTCATGCCTCCATCCGTGAGGGCGGTCTAGGAATCATA---TCC 654
 C. bellicosus-9 598 AGCCCCGATGCATTTATTCATGCCTCCATCCGTGAGGGCGGTCTTGGAAATCATA---TCC 654
 C. bellicosus-1 601 ACAGCAGATGCCTTTATCCACGCTCCCATAGGGAGGGAGGATTGGGCATCCTA---TCA 657
 C. bellicosus-1 601 GCAGCAGATGCCTTTATCCACGCTCCCATAGGGAGGGAGGATTGGGCATCCTA---TCA 657
 C. bellicosus-1 601 ACAGCAGATGCCTTTATCCACGCTCCCATAGGGAGGGAGGATTGGGCATCCTA---TCA 657
 C. cyanescens-1 592 AACACTCCGTGCCGTGGACCGGTAGTGCGGATCTCCGTAAGGAAGATCCTGCATTGCA 651
 C. cyanescens-2 592 AACACTCCGTGCCGTGGACCGGTAGTGCGGATCTCCGTAAGGAAGATCCTGCATTGCA 651
 C. cyanescens-3 598 GTTATTTTCAGCTGCAGACCCCTAGGGTAAACCGTGGGATGCTCAAATCGGTGGA---TAG 654
 C. cyanescens-4 475 ACCTCTGATGCTAGCATTTATGCTAGAGAGCGAGACGGTGGTCTGGGGATTTTT---AAC 531
 C. cyanescens-5 598 AGCCCCGACGCGTTTCATCCATGCCCCGTTAAAGACGGTGGGCTGGGCATTTTA---TCG 654
 C. ensifer-1 592 AACACTCCGTGCCGTGGACCGGTAGTGCGGATCTCCGTAAGGAAGATCCTGCA---TTT 648
 C. ensifer-2 592 TCTCCGACGCGTTTTGTTACGCGAGACGTTAAGGGGGCGGTATTGGAGTGGGC---CCG 648
 C. ensifer-3 592 AACACTCCGTGCCGTGGACCGGTAGTGCGGATCTCCGTAAGGAAGATCCTGCA---TTT 648
 C. ensifer-4 598 AGCCCCGACGCGTTTCATCCATGCGCCCATCAGAGAAGGCGGGCTTGGAACTA---TCG 654
 C. ensifer-5 598 AGCCCCGACGCGTTTCATCCACGCGCCCGTCAAAGAAGGCGGGCTCGGAATTCTA---TCA 654
 C. ensifer-6 598 AGCCCCGACGCGTTTCATTCATGCACCCGTCAAAGAAGGCGGGCTTGGAAATTCTA---TCC 654
 C. ensifer-7 598 AGCCCCGACGCGTTTCATCCACGCGCCCGTCAAAGAAGGCGGGCTCGGAATTCTA---TCA 654
 C. ensifer-8 598 AGCCCCGATGCATTTATCCATGCCCCCTCAAGGAGGGTGGGCTTGGAAATAA---TCA 654
 C. ensifer-9 475 ACCTCTGATGCTAGCATTTGCGCTAGAGAGCGAGATGGTGGCTTGGGGATTTTT---AAC 531
 C. ensifer-10 592 TCCCCGACGCGTTTTGTTACGCGAGACGTTAAGGGGGCGGTATTGGAGTGGGC---CCG 648
 C. ensifer-11 592 TCCCCGACGCGTTTTGTTACGCGAGACGTTAAGGGGGCGGTATTGGGGTGGGC---CCG 648
 C. ensifer-12 598 GCCCGACGCGTTTCATCCATGCACCCGTCAAAGAAGGCGGGCTTGGAAATTCTAT---CCA 654
 C. ensifer-13 592 TCCCCGATGGCGTTTTGTTACGCGAGACGTTAAGGGCGGCGGCATCGGAGTAGGC---CCG 648
 C. ensifer-14 592 TCCCCGATGGCGTTTTGTTACGCGAGACGTTAAGGGCGGCGGCATCGGAGTAGGC---CCG 648
 C. ensifer-15 592 TCCCCGATGGCGTTTTGTTACGCGAGACGTTAAGGGCGGCGGCATCGGAGTAGGC---CCG 648

C. bellicosus-1 655 TTAAATTCCCATATACCGTGCATCATGAGACGCCGAATGACAAACATCATATCAACA--- 711
 C. bellicosus-2 655 TTAAATTCCCATATACCGTGCATCATGAGACGCCGAATGACAAACATCATATCAACA--- 711
 C. bellicosus-3 655 TTAAATTCCCATATACCGTGCATCATGAGACGCCGAATGACAAACATCATATCAACA--- 711
 C. bellicosus-4 655 ATGAACGCGCAGCTTCCCTGTATACTAAGACGTCGCATAACAAATCTTGTGTCCACT--- 711
 C. bellicosus-5 655 TTAAATTCCCATATACCGTGCATCATGAGACGCCGAATGACAAACATCATATCAACT--- 711
 C. bellicosus-6 658 CTCAAGGCGCATGTCCCCTGCATCCTGAGTCGCCGTATGACGAATCTGGCTCTGCAG--- 714
 C. bellicosus-7 655 TTAAATTCCCATATACCGTGCATCATGAGACGCCGAATGACAAACATCATATCAACA--- 711
 C. bellicosus-8 655 TTAAATTCCCATATACCGTGCATCATGAGACGCCGAATGACAAACATCATATCAACA--- 711
 C. bellicosus-9 655 TTAAATTCCCATATACCGTGCATCATGAGACGCCGAATGACAAACATCATATCAACA--- 711
 C. bellicosus-1 658 CTCAAGGCGCATGTCCCCTGCATCCTGAGTCGCCGTATGACGAATCTGGCTCTGCAG--- 714
 C. bellicosus-1 658 CTCAAGGCGCATGTCCCCTGCATCCTGAGTCGCCGTATGACGAATCTGGCTCTGCAG--- 714
 C. cyanescens-1 652 CAAGTCCTCCCGACGGCGTTTTGTTACGTCACGTTAAGGGGGGGCGGGATTTGGAAGTG 711
 C. cyanescens-2 652 CAAGTCCTC-----CCC-----GACGGCGTTTTGTTACGCGAGACGTTAAGGGGGGGCGG 699
 C. cyanescens-3 655 GAGTCTCCGCTTTTATGTACGTAAATTCGTACATCTGAATAAAAGCAGCCCCGACGC--- 711
 C. cyanescens-4 532 CTTAGGATACAAATACCCCAAATATATCTCAAGGCTGAAAAAGCTCGCCGATAACGCT 591
 C. cyanescens-5 655 ATGAACGCGCATATCCCCTGCATACTAAGACGTCGCATTACTAACTAGTCTCCAAC--- 711
 C. ensifer-1 649 GCACAAGTCCTCCCGACGGCGTTTTGTTACGCGAGACGTTAAGGGGGGCGGTATTGGAGT 708
 C. ensifer-2 649 CTTGCGCATAACATCCCGCGCATAGTCCTGTGAGGCTGGATAAGCTCGCGTCGACCACT 708
 C. ensifer-3 649 GCACAAGTCCTCCCGACGGCGTTTTGTTACGCGAGACGTTAAGGGGGGCGGTATTGGGGT 708
 C. ensifer-4 655 ATGGACGCACATGTTCCCTTGATCCTAAGACGTCGCATAACTAATCTTGATCCACT--- 711
 C. ensifer-5 655 ATGAACGCGCATGTCCCCTGCATTTTAAGACGTCGCATGACCAACTTAATATCTACT--- 711
 C. ensifer-6 655 ATGAACGCGCAGCTCCCCTGTATTTTAAGACGTCGCATGACAAATCTAGTAGCCAAT--- 711
 C. ensifer-7 655 ATGAACGCGCATGTCCCCTGCATTTTAAGACGTCGCATGACCAACTTAATATCTACT--- 711
 C. ensifer-8 655 TTAAATGCACATTTGCCGGCAATCATGAGACGCCGTATGACCACCATTATCGCAACA--- 711
 C. ensifer-9 532 CTTAGGTTACAAATACCCCAAATTTATCGATCAAGGCTGAAAAAGCTCGCCGATAACGCT 591
 C. ensifer-10 649 CTTGCGCATAACATCCCGCGCATAGTCCTGTGAGGCTGGATAAGCTCGCGTTGACCACT 708
 C. ensifer-11 649 CTTGCGCATAACATCCCGCGCATAGTCCTGTGAGGCTGGATAAGCTCGCGTTGACCACT 708
 C. ensifer-12 655 AGGAGACCCACTGTCGCCAGTCTCTTTAATCTTGTGCTAGATGAGTTATTCTGCAA--- 711
 C. ensifer-13 649 CTAGCGCAGAACATCCCGCGCATAGTCCTCTGAGGCTGGATAAGCTCGCGTTGACCACT 708
 C. ensifer-14 649 CTAGCGCAGAACATCCCGCGCATAGTCCTATCGAGGCTGGATAAGCTCGCGTTGACCACT 708
 C. ensifer-15 649 CTAGCGCAGAACATCCCGCGCATAGTCCTATCGAGGCTGGATAAGCTCGCGTTGACCACT 708

C. bellicosus-1 712 -----GCCGACCCATTAACTAAAACAATATTAGCGTTGCCCTCCTCTATAAAGTTCTAT 765
 C. bellicosus-2 712 -----GCCGACCCATTAACTAAAACAATATTAGCGTTGCCCTCCTCTATAAAGTTCTAT 765
 C. bellicosus-3 712 -----GCCGACCCATTAACTAAAACAATATTAGCGTTGCCCTCCTCTATAAAGTTCTAT 765
 C. bellicosus-4 712 -----GCCGATAGAGTAACCGCCTCCGTTCTTTCACTGCCATCCGCAGTGAAATTATAT 765
 C. bellicosus-5 712 -----GCCGACCCATTAACTAAAACAATATTAGCGTTGCCCTCCTCTATAAAGTTCTAT 765
 C. bellicosus-6 715 -----GCCGACGCACTTACGACGGCCATTTTTCTATCGTCCTACTCGCAGAAATTCTGC 768
 C. bellicosus-7 712 -----GCCGACCCATTAACTAAAACAATATTAGCGTTGCCCTCCTCTATAAAGTTTAT 765
 C. bellicosus-8 712 -----GCCGACCCATTAACTGAAACAATATTAGCGTTGCCCTCCTCTATAAAGTTCTAT 765
 C. bellicosus-9 712 -----GCCGACCCATTAACTAAAACAATATTAGCGTTGCCCTCCTCTATAAAGTTTAT 765
 C. bellicosus-1 715 -----GCCGACGCACTTATGACGGCCATTTTTCTATCGTCCTACTCGCAGAAATTCTGC 768
 C. bellicosus-1 715 -----GCCGACGCACTTACGACGGCCATTTTTCTATCGTCCTACTCGCAGAAATTCTGC 768
 C. bellicosus-1 715 -----GCCGACGCACTTACGACGGCCATTTTTCTATCGTCCTACTCGCAGAAATTCTGC 768
 C. cyanescens-1 712 -----GGGCCCGCTTGCGCATAACATCCCGCGCATAGTCTGTGCGAGGCTGGATAAGCT 765
 C. cyanescens-2 700 -----TATTTGGAGTGGGCCGCTTGCGCATAACATCCCGCGCATAGTCTGTGCGAGGC 753
 C. cyanescens-3 712 -----GTTTCATCCATGCCCGCGTTAAAGACGGTGGGCTCGGCATTCTATCGATGAACGC 765
 C. cyanescens-4 592 GATCCCGATGACAGGCTTCTTACCAACTTTGTTTCTTCGGAGTATTTTATAACTCTCCAA 651
 C. cyanescens-5 712 -----GCCGACCGAAGGAGACCCACTATCGCCAGTCTTATTCAACATCGTGCTGGATGA 765
 C. ensifer-1 709 -----GGGCCCGCTTGCGCATAACATCCCGCGC---CATAGTCTGTGCGAGGCTGGATAA 759
 C. ensifer-2 709 -----TCTGACGCGGCTGTCCAAGCAATGCTG---AAGGAAGGCCCGGTGGTGCCTTTT 759
 C. ensifer-3 709 -----GGGCCCGCTTGCGCATAACATCCCGCGCATAGTCTGTGCGAGGCTGGATAAGCT 762
 C. ensifer-4 712 -----GCCGACAGAGTAACCGCCTCCGTTCTTTCACTGCCGTCCGCAGTGAAATTATAC 765
 C. ensifer-5 712 -----GCCGATAGAGTAACCGCCTCCGTTCTTTCACTACCGTCTGCAGTGAAACTATAC 765
 C. ensifer-6 712 -----GCCGATAGAGTAACCGCCTCCGTTCTTTCACTGCCATCTGCAGTGAAACTGTAT 765
 C. ensifer-7 712 -----GCCGATAGAGTAACCGCCTCCGTTCTTTCACTACCGTCTGCAGTGAAACTATAC 765
 C. ensifer-8 712 -----GCCGACCCAGTCACTGCAAGAACTGCGGCTGCCGTCCGCAGTAAATTTCTAT 765
 C. ensifer-9 592 GACCCAGACGACAGGCTACTTACCAACTTTGTTTCTTCGGAGTATTTTTTAAATCTCCAA 651
 C. ensifer-10 709 -----TCTGACGCGGCTGTCCAAGCAATGCTGAAGGAAGGCCCGGTGGTGCCTTTTAAA 762
 C. ensifer-11 709 -----TCTGACGCGGCTGTCCAAGCAATGCTGAAGGAAGGCCCGGTGGTGCCTTTTAAA 762
 C. ensifer-12 712 -----GCTCGATACACGAGCAAAACCGGGACTCACAGTGCAGGACGAGCGAGTCAGAAT 765
 C. ensifer-13 709 -----TCTGACGCGGCTGTCCAAGCAATGCTGAAGGAAGGCCCGGTGGTGCCTTTTAAA 762
 C. ensifer-14 709 -----TCTGACGCGGCTGTCCAAGCAATGCTGAAGGAAGGCCCGGTGGTGCCTTTTAAA 762
 C. ensifer-15 709 -----TCTGACGCGGCTGTCCAAGCAATGCTGAAGGAAGGCCCGGTGGTGCCTTTTAAA 762

C. bellicosus-1 766 GAAAAAATTACCAAATGGACTACAAATAACGGA 798
 C. bellicosus-2 766 GAAAAAATTACCAAATGGACTACAAATAACGGA 798
 C. bellicosus-3 766 GAAAAAATTACCAAATGGACTACAAATAACGGA 798
 C. bellicosus-4 766 GAAAAGCTCATCTCTTTGGACGAGAGATAACGGA 798
 C. bellicosus-5 766 GAAAAAATTACCAAATGGACTACAAATAACGGA 798
 C. bellicosus-6 769 GATAGGCTTTCACTTTGGTCATCGTCTTATGGA 801
 C. bellicosus-7 766 GAAAAAATTACCAAATGGACTACAAATAACGGA 798
 C. bellicosus-8 766 GAAAAAATTACCAAATGGACTACAAATAACGGA 798
 C. bellicosus-9 766 GAAAAAATTACCAAATGGACTACAAATAACGGA 798
 C. bellicosus-1 769 GATAGGCTTTCACTTTGGTCATCGTCTTATGGA 801
 C. bellicosus-1 769 GATAGGCTTTCACTTCGGTCATCGTCTTATGGA 801
 C. bellicosus-1 769 GATAGGCTTTCACTTTGGTCATCGTCTTATGGA 801
 C. cyanescens-1 766 CGCGTTGACCACTTCTGACGCGGCTGTCCAAGC 798
 C. cyanescens-2 754 TGGATAAGCTCGCGTTGACCACTTCTGACGCGG 786
 C. cyanescens-3 766 GCATATTCCATGTATATTGAGACGTGCGATTAG 798
 C. cyanescens-4 652 AGAAGAACTAGTCTGTTGGCTGGCGACACCGCC 684
 C. cyanescens-5 766 ACTGTTTTGTAACTCGATTGCGGAGTTAACCG 798
 C. ensifer-1 760 GCTCGCGTTGACCACTTCTGACGCGGCTGTCCA 792
 C. ensifer-2 760 AAAGAGCGTATGCTGCATCTCTGAGCGGTGTG 792
 C. ensifer-3 763 CGCGTTGACCACTTCTGACGCGGCTGTCCAAGC 795
 C. ensifer-4 766 GAAAAGCTTATATCTTTGGACAAGAGAAAACGGA 798
 C. ensifer-5 766 GACCACCTACTAGCCTGGACTAGAGAAAACGGA 798
 C. ensifer-6 766 GACAGACTACTCTCGTGGACGAGGAGAACGGA 798
 C. ensifer-7 766 GACCACCTACTAGCCTGGACTAGGGAAAACGGA 798
 C. ensifer-8 766 GATAAACTCGTCAGATGGACGAGTGGTAACGGA 798
 C. ensifer-9 652 AGGAAAACCTAGCCTGTTGGCTGGCGACACCGCT 684
 C. ensifer-10 763 GAGCGTATGCTGCATCTCCTGAGCGGTGTGCCT 795
 C. ensifer-11 763 GAGCGTATGCTGCATCTCCTGAGCGGTGTGCCT 795
 C. ensifer-12 766 AATCGGCTACGCCGACGATATTATGCTTATGGA 798
 C. ensifer-13 763 GAGCGTATGCTGCATCTCCTGGGCGGTGTGCCT 795
 C. ensifer-14 760 GAGCGTATGCTGCATCTCCTGGGCGGTGTGCC 795
 C. ensifer-15 763 GAGCGTATGCTGCATCTCCTGGGCGGTGTGCCT 795

4. CONCLUSÕES

As espécies do gênero *Coprophanaeus* mostraram número e morfologia cromossômica conservados, e diferentes sistemas sexuais (XY, Xy, XY_p). Uma grande quantidade de heterocromatina constitutiva (HC) foi observada em todas as espécies, bem como heterogeneidade para a riqueza de pares de bases. A hibridização *in situ* fluorescente (FISH) revelou sinais de DNAr 5S em um par autossômico de todas as espécies, e sinais de DNAr 18S variando de 1-8 pares autossômicos.

As análises de hibridização da fração *Cot*-1 DNA em *C. cyanescens* revelaram um bloco pericentromérico no pequeno cromossomo B meta-submetacêntrico, e nenhuma marcação com o uso de sondas de DNAr 5S ou 18S. Utilizando sondas do retrotransposon não-LTR *LOA-like* foi evidenciado que oito bivalentes e o cromossomo B possuíam marcações.

O mapeamento citogenético de transposons *Mariner* em *Coprophanaeus cyanescens*, *C. ensifer* e *Diabroctis mimas* mostrou que estas sequências estão localizadas na região pericentromérica heterocromática. As análises das sequências de *Mariner* revelaram a presença de dois grandes grupos de sequências nas três espécies estudadas e em outros insetos, com baixos níveis de similaridade entre as transposases.

Análises comparativas de transposons *Mariner* (*Mariner_Tbel* e *Mariner1_BT*) em diferentes animais mostraram fortes evidências de transferência horizontal (HT) entre insetos e mamíferos. *Mariner_Tbel* foi transferido para musaranho (tree shrew) e ouriço (hedgehog), respectivamente, há aproximadamente 70 e 67 milhões de anos. O transposon *Mariner1_BT*, por sua vez, foi introduzido pela primeira vez no ancestral comum de ruminantes e cetáceos há aproximadamente 60 milhões de anos.

As sequências de retrotransposons não-LTR R2 apresentaram níveis extremamente altos de divergências, e as inserções destes elementos não afetaram a evolução em concerto das sequências de genes ribossomais.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- Aksoy S, Williams S, Chang S, Richards FF (1990) SLACS retrotransposon from *Trypanosoma brucei gambiense* is similar to mammalian lines. *Nucleic Acids Res* 18: 785-792.
- Almeida MC, Goll LG, Artoni RF, Nogaroto V, Matiello RR, Vicari MR (2010) Physical mapping of 18S rDNA cistron in species of the *Omophota* genus (Coleoptera, Alticinae) using fluorescent in situ hybridization. *Micron* 41: 729-734.
- Angelini DR, Jockusch EL (2008) Relationships among pest flour beetles of the genus *Tribolium* (Tenebrionidae) inferred from multiple molecular markers. *Mol Phylogenet Evol* 46: 127-141.
- Angus RB, Wilson CJ, Mann DJ (2007) A chromosomal analysis of 15 species of Gymnopleurini and Coprini (Coleoptera: Scarabaeidae). *Tijdsch Voor Entomol* 150: 201-211.
- Bachtrog D (2005) Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in *Drosophila*. *Genome Res* 15: 1393-1401.
- Balloux F, Ecoffey E, Fumagalli L, Goudet J, Wyttenbach A, Hausser J (1998) Microsatellite conservation, polymorphism, and GC content in shrews of the genus *Sorex* (Insectivora, Mammalia). *Mol Biol Evol* 15: 473-475.
- Barceló F, Gutiérrez F, Barjau I, Portugal J (1998) A theoretical perusal of the satellite DNA curvature in Tenebrionid beetles. *J Biomol Struct Dyn* 16: 41-50.
- Bartolome C, Bello X, Maside X (2009) Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes. *Genome Biol* 10:R22.
- Beauregard A, Curcio MJ, Belfort M (2008) The take and give between retrotransposable elements and their hosts. *Annu Rev Genet* 42: 587-617.
- Bertolotto CEV, Pellegrino KCM, Yonenaga-Yassuda Y (2004) Occurrence of B chromosomes in lizards: a review. *Cytogenet Genome Res* 106: 243-246.
- Beukeboom LW (1994) Bewildering Bs: an impression of the 1st B-chromosome conference. *Heredity* 73: 328-336.
- Biémont C (1994) Dynamic equilibrium between insertion and excision of *P* elements in highly inbred lines from an *M*'strain of *Drosophila melanogaster*. *J Mol Evol* 39: 466-472.
- Biémont C, Vieira C (2006) Genetics: Junk DNA as an evolutionary force. *Nature* 443: 521-524.

- Biet E, Sun J, Dutreix M (1999) Conserved sequence preference in DNA binding among recombination proteins: an effect of ssDNA secondary structure. *Nucleic Acids Res* 27: 596-600.
- Bione EG, Camparoto ML, Simões ZL (2005a) A study of the constitutive heterochromatin and nucleolus organizer regions of *Isocoprins inhiata* and *Diabroctis mimas* (Coleoptera: Scarabaeidae, Scarabaeinae) using C-banding, AgNO₃ staining and FISH techniques. *Gen Mol Biol* 28: 111-116.
- Bione EG, Moura RC, Carvalho R, Souza MJ (2005b) Karyotype, C-banding pattern, NOR location and FISH study of five Scarabaeidae (Coleoptera) species. *Gen Mol Biol* 28: 376-381.
- Birchler JA, Presting GG (2012) Retrotransposon insertion targeting: a mechanism for homogenization of centromere sequences on nonhomologous chromosomes. *Genes Dev* 26: 638-640.
- Böhne A, Brunet F, Galiana-Arnoux D, Schultheis C, Volff J-N (2008) Transposable elements as drivers of genomic and biological diversity in vertebrates. *Chromosome Res* 16: 203-215.
- Borisov YM (2012) The polymorphism and distribution of B chromosomes in germline and somatic cells of *Tscherskia triton* de Winton (Rodentia, Cricetinae). *Russ J Genet* 48: 538-542.
- Brady SG, Larkin L, Danforth BN (2009) Bees, ants, and stinging wasps (Aculeata). In: *The Timetree of Life*. Hedges SB, Kumar S (eds.) Oxford University Press, Huntington beach, USA, pp264-269.
- Braquart C, Royer V, Bouhin H (2001) DEC: a new miniature inverted-repeat transposable element from the genome of the beetle *Tenebrio molitor*. *Insect Mol Biol* 8: 571-574.
- Brookfield JFY, Badge RM (1997) Population genetics model of transposable elements. *Genetica* 52: 281-294.
- Brown DD, Wensink PC, Jordan E (1972) A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: the evolution of tandem genes. *J Mol Biol* 63: 65-73.
- Brown JR (2003) Ancient horizontal gene transfer. *Nat Rev Genet* 4: 121-132.
- Bruvo B, Pons J, Ugarkovic D, Juan C, Peitpierre E, Plohl M (2003) Evolution of low-copy number and major satellite DNA sequences coexisting in two *Pimelia* species-groups (Coleoptera). *Gene* 312: 85-94.
- Bruvo-Madaric B, Plohl M, Ugarkovic D (2007) Wide distribution of related satellite DNA families within the genus *Pimelia* (Tenebrionidae). *Genetica* 130: 35-42.

- Burke WD, Eickbush DG, Xiong Y, Jakubczak J, Eickbush TH (1993) Sequence Relationship of Retrotransposable Elements R1 and R2 Within and Between Divergent Insect Species'. *Mol Biol Evol* 10:163-185.
- Burke WD, Malik HS, Jones JP, Eickbush TH (1999) The domain structure and mechanism of R2 elements are conserved throughout arthropods. *Mol Biol Evol* 16: 502-511.
- Burke WD, Malik HS, Lathe III W C, Eickbush TH (1998) Are retrotransposons long-term hitchhikers? *Nature* 392: 141-142.
- Burke WD, Müller F, Eickbush TH (1995) R4, a non-LTR retrotransposon specific to the large subunit rRNA gene of nematodes. *Nucleic Acids Res* 23: 4628-4634.
- Burke WD, Singh D, Eickbush TH (2003) R5 retrotransposons insert into a family of infrequently transcribed 28S rRNA genes of Planaria. *Mol Biol Evol* 20: 1260-1270.
- Cabral-de-Mello DC, Moura RC, Martins C (2010a) Chromosomal mapping of repetitive DNAs in the beetle *Dichotomius geminatus* provides the first evidence for an association of 5S rRNA and histone H3 genes in insects, and repetitive DNA similarity between the B chromosome and A complement. *Heredity* 104: 393-400.
- Cabral-de-Mello DC, Moura RC, Martins C (2011b) Cytogenetic mapping of rRNAs and histone H3 genes in 14 species of *Dichotomius* (Coleoptera, Scarabaeidae, Scarabaeinae) beetles. *Cytogenet Genome Res* 134: 127-135.
- Cabral-de-Mello DC, Moura RC, Melo AS, Martins C (2011b) Evolutionary dynamics of heterochromatin in the genome of *Dichotomius* beetles based on chromosomal analysis. *Genetica* 139: 315–325.
- Cabral-de-Mello DC, Oliveira SG, Moura RC, Martins C (2011a) Chromosomal organization of the 18S and 5S rRNAs and histone H3 genes in Scarabaeinae coleopterans: insights into the evolutionary dynamics of multigene families and heterochromatin. *BMC Genetics* 12: 88.
- Cabral-de-Mello DC, Oliveira SG, Ramos I, Moura RC (2008) Karyotype differentiation patterns in species of the subfamily Scarabaeinae (Scarabaeidae, Coleoptera). *Micron* 39: 1243-1250.
- Cabral-de-Mello DC, Silva FAB, Moura RC (2007) Karyotype characterization of *Eurysternus caribaeus*: The smallest diploid number among Scarabaeidae (Coleoptera: Scarabaeoidea). *Micron* 38: 323-325.
- Cabral-de-Mello, Moura DC, Carvalho R, Souza MJ (2010c) Cytogenetic analysis of two related *Deltochilum* (Coleoptera, Scarabaeidae) species: diploid number reduction, extensive heterochromatin addition and differentiation. *Micron* 41: 112-117.

- Camacho JP, Sharbel TF, Beukeboom LW (2000) B-chromosome evolution. *Philos T Roy Soc B* 355: 163-178.
- Camacho JPM (2005) B Chromosomes. In: *The Evolution of the Genome*, Gregory TR (Ed.). Elsevier, San Diego, USA, pp223-286.
- Capy P (2005) Classification and nomenclature of retrotransposable elements. *Cytogenet Genome Res* 110: 457-461.
- Capy P, Maisonhaute C (2002) Acquisition/loss of modules: the construction set of transposable elements. *Russ J Genet* 38: 594-601.
- Capy P, Vitalis R, Langin T, Higuete D, Bazin C (1996) Relationships between transposable elements based upon the integrase-transposase domains: is there a common ancestor? *J Mol Evol* 42:359-368.
- Carranza S, Baguña J, Riutort M (1999) Origin and evolution of paralogous rRNA gene clusters within the flatworm family dugesiidae (Platyhelminthes, Tricladida). *J Mol Evol* 49: 250-259.
- Charlesworth B, Snlegowski P, Stephan W (1994) The evolution dynamics of repetitive DNA in eukaryotes. *Nature* 371: 215-220.
- Charlesworth B (1991) Transposable elements in natural populations with a mixture of selected and neutral insertion sites. *Genet Res* 57: 127-134.
- Charlesworth B, Charlesworth D (1983) The population dynamics of transposable elements. *Genet Res* 42: 1-27.
- Charlesworth B, Langley CH (1989) The population genetics of *Drosophila* transposable elements. *Annu Rev Genet* 23: 251-87.
- Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95: 118-128.
- Chen JM, Stenson PD, Cooper DN, Ferec C (2005) A systematic analysis of LINE-1 endonuclease-dependent retrotranspositional events causing human genetic disease. *Hum Genet* 117: 411-427.
- Chen T, Tsujimoto N, Li E (2004) The PWWP Domain of Dnmt3a and Dnmt3b Is Required for Directing DNA Methylation to the Major Satellite Repeats at Pericentric Heterochromatin. *Mol Cell Biol* 24: 9048-9058.
- Chevenet F, Brun C, Banuls AL, Jacq B, Christen R (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 7:439.
- Christensen S, Pont-Kingdon G, Carroll D (2000) Target specificity of the endonuclease from the *Xenopus laevis* non-long terminal repeat retrotransposon, Tx1L. *Mol Cell Biol* 20:

1219-1226.

- Christensen SM, Ye J, Eickbush TH (2006) RNA from the 5' end of the R2 retrotransposon controls R2 protein binding to and cleavage of its DNA target site. *Proc Natl Acad Sci U S A* 103: 17602-17607.
- Coen ES, Strachan T, Dover GA (1982) The dynamics of concerted evolution of rDNA and histone gene families in the melanogaster species subgroup of *Drosophila*. *J Mol Biol* 158: 17-35.
- Colomba M, Vitturi R, Libertini A, Gregorini A, Zunino M (2006) Heterochromatin of the scarab beetle, *Bubas bison* (Coleoptera: Scarabaeidae) II. Evidence for AT-rich compartmentalization and a high amount of rDNA copies. *Micron* 37: 47-51.
- Colomba MS, Monteresino E, Vitturi R, Zunino M (1996) Characterization of mitotic chromosomes of the scarab beetles *Glyphoderus sterquilinus* (Westwood) and *Bubas bison* (L.) (Coleoptera: Scarabaeidae) using conventional and banding techniques. *Biol Zentbl* 115: 58-70.
- Colomba MS, Vitturi R, Zunino M (2000) Karyotype analysis, banding, and fluorescent *in situ* hybridization in the scarab beetle *Gymnopleurus sturmi* McLeay (Coleoptera Scarabaeoidea: Scarabaeidae). *J Hered* 91: 260-264.
- Coluccia E, Cannas R, Cau A, Deiana AM, Salvadori S (2004) B chromosomes in Crustacea Decapoda. *Cytogenet Genome Res* 106: 215-221.
- Conconi A, Widmeret RM, Koller T, Sogo JM (1989) Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. *Cell* 57: 753-761.
- Dammann R, Lucchini R, Koller T, Sogo JM (1995) Transcription in the yeast rRNA gene locus: distribution of the active gene copies and chromatin structure of their flanking regulatory sequences. *Mol Cell Biol* 15: 5294-5303.
- Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A (1990) Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* 124: 339-355.
- Davis C, Wurdack K (2004) Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales. *Science* 305: 676-678.
- Dawe RK (2003) RNA interference, transposons, and the centromere. *Plant Cell* 15: 297-301.
- Dawid IB, Rebbert ML (1981) Nucleotide sequence at the boundaries between gene and insertion regions in the rDNA of *D. melanogaster*. *Nucleic Acids Res* 9: 5011-5020.
- De La Rúa P, Serrano J, Hewitt GM and Galián J (1996) Physical mapping of rDNA genes in the ground beetle *Carabus* and related genera (Coleoptera: Carabidae). *J Zool Syst Evol*

Res 34: 95-101.

- Deininger PL, Moran JV, Batzer MA, Kazazian HH Jr (2003) Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev* 13: 651-658
- Delaunay L, Chénais B, Hardivillier Y, Gauvry L, Casse N (2009) *Mariner* transposons as genetic tools in vertebrate cells. *Genetica* 137: 9-17.
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-F, Guindon S, Lefort V, Lescot M, Claverie J-M, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36: 465-469.
- Dimitri P, Arcà B, Berghella L, Mei E (1997) High genetic instability of heterochromatin after transposition of the LINE-like I factor in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 94: 8052-8057.
- Dimitri P, Junakovic N (1999) Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. *Trends Genet* 4: 123-124.
- Dover G (1986) Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends Genet* 2: 159-165.
- Dover G (2002) Molecular drive. *Trends Genet* 18: 587-589.
- Dover GA (1982) Molecular drive: a cohesive mode of species evolution. *Nature* 299: 111-117.
- Dover GA, Falvell RB (1984) Molecular coevolution: DNA divergence and the maintenance of function. *Cell* 38: 622-623.
- Edmonds WD (1967) The immature stages of *Phanaeus* (Coprophanaeus) *jasius* Oliver and *Phanaeus* (Metallophanaeus) *saphirinus* Sturm (Coleoptera: Scarabaeidae). *Coleopt Bull* 21: 97-105.
- Edmonds WD (1972) Comparative skeletal morphology, systematics and evolution of the Phanaeine dung beetles (Coleoptera: Scarabaeidae). *Kans Univ Sci Bull* 49: 731-874.
- Edmonds WD (1994) Revision of *Phanaeus* MacLeay, a new world genus of Scarabaeine dung beetles (Coleoptera: Scarabaeidae, Scarabaeinae). *Natural History Museum of Los Angeles County, Contributions in Science* 443: 105pp.
- Edmonds WD (2000) Revision of the neotropical dung beetle genus *Sulcophanaeus* (Coleoptera: Scarabaeidae: Scarabaeinae). *Folia Heyrovskyana Supplementum* 6: 60pp.
- Edmonds WD, Zidek J (2004) Revision of the neotropical dung beetle genus *Oxysternon* (Scarabaeidae: Scarabaeinae: Phanaeini). *Folia Heyrovskyana Supplementum* 11, 58pp.
- Edmonds WD, Zidek JA (2010) Taxonomic review of the neotropical genus *Coprophanaeus* Olsoufieff, 1924 (Coleoptera: Scarabaeidae, Scarabaeinae). *Insecta Mundi* 129: 1-111.

- Eickbush DG, Eickbush TH (2003) Transcription of endogenous and exogenous R2 elements in the rRNA gene locus of *Drosophila melanogaster*. *Mol Cell Biol* 23: 3825-3836.
- Eickbush DG, Junqiang Y, Zhang X, Burke WD, Eickbush TH (2008) Epigenetic regulation of retrotransposons within the nucleolus of *Drosophila*. *Mol Cell Biol* 28: 6452-6461.
- Eickbush DG, Lathe WC, Francino MP, Eickbush TH (1995) R1 and R2 Retrotransposable Elements of *Drosophila* Evolve at Rates Similar to those of Nuclear Genes. *Genetics* 139: 685-695.
- Eickbush TH (2002) R2 and related site-specific non-long terminal repeat retrotransposons. In: Mobile DNA II. Craig NL, Craigie R, Gellart M, Lambowitz AM (eds.). Washington DC, American Society of Microbiology, 813-835.
- Eickbush TH, Eickbush DG (2007) Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* 175: 477-485.
- Eirín-López JM, Rebordinos L, Rooney AP, Rozas J (2012) The Birth-and-Death evolution multigene families revisited. *Genome Dyn* 7: 170-196.
- Eymeri A, Callanan M, Vourc'h C (2009) The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *Int J Dev Biol* 53: 259-268.
- Felger I, Hunt JA (1992) A non-LTR retrotransposon from the Hawaiian *Drosophila*: the LOA element. *Genetica* 85: 119-130.
- Feng Q, Schumann G, Boeke JD (1998) Retrotransposon R1Bm endonuclease cleaves the target sequence. *Proc Natl Acad Sci U S A* 95: 2083-2088.
- Fenton B, Malloch G, Germa F (1998) A study of variation in rDNA ITS regions shows that two haplotypes coexist within a single aphid genome. *Genome* 41: 337-345.
- Ferreira A, Cella D, Tardivo JR, Virkki N (1984) Two pairs of chromosomes: a new low record for Coleoptera. *Braz J Genet* 2: 231-239.
- Ferreira IA, Martins C (2008) Physical chromosome mapping of repetitive DNA sequences in Nile tilapia *Oreochromis niloticus*: evidences for a differential distribution of repetitive elements in the sex chromosomes. *Micron* 39: 411-418.
- Feschotte C (2008) Transposable elements and the evolution of regulatory networks. *Nature Rev Genet* 9: 397-405.
- Feschotte C, Pritham EJ (2007) DNA transposons and the evolution of eukaryotic genomes. *Ann Rev of Genet* 41: 331-368.
- Flavell AJ (1995) Retroelements, reverse transcriptase and evolution. *Comp Biochem Phys B: Biochem Mol Biol* 110: 3-15.

- Friz CT (1978) The biochemical composition of the free-living amoebae *Chaos chaos*, *Amoeba dubia*, and *Amoeba proteus*. *Comp Biochem Phys* 26: 81-90.
- Fujiwara M, Inafuku J, Takeda A, Watanabe A, Fujiwara A, et al. (2009) Molecular organization of 5S rDNA in bitterlings (Cyprinidae). *Genetica* 135: 355-365
- Galián J, Hogan JE, Vogler AP (2002) The origin of multiple sex chromosomes in tiger beetles. *Mol Biol Evol* 19: 1792-1796.
- Gálian J, Vogler AP (2003) Evolutionary dynamics of a satellite DNA in the tiger beetle species pair *Cicindela campestris* and *C. maroccana*. *Genome* 46: 213-223.
- Ganley AR, Kobayashi T (2007) Highly efficient concerted evolution in the ribosomal DNA repeats: Total DNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Res* 17: 184-191.
- Gemayel R, Vences MD, Legendre M, Verstrepen KJ (2010) Variable Tandem Repeats Accelerate Evolution of Coding and Regulatory Sequences. *Annu Rev Genet* 44: 445-77.
- Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, et al. (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452: 949-955.
- Gilbert C, Pace JK, Feschotte C (2009) Horizontal SPINning of transposons. *Commun Integr Biol* 2: 1-3.
- Gilbert C, Schaack S, Pace JK, 2nd, Brindley PJ, Feschotte C (2010) A role for host-parasite interactions in the horizontal transfer of transposons across phyla. *Nature* 464: 1347-1350.
- Gilbert W (1986) Origin of life: the RNA world. *Nature* 319: 618.
- Gladyshev EA, Arkhipova IR (2009) Rotifer rDNA-specific R9 retrotransposable elements generate an exceptionally long target site duplication upon insertion. *Gene* 448: 145-150.
- Gómez-Zurita J, Pons J, Petitpierre E (2004) The evolutionary origin of a novel karyotype in *Timarcha* (Coleoptera, Chrysomelidae) and general trends of chromosome evolution in the genus. *J Zool Syst Evol Res* 42: 332-341.
- Gomulski LM, Brogna S, Babaratsas A, Gasperi G, Zacharopoulou A et al. (2004) Molecular basis of the size polymorphism of the first intron of the *Adh-1* gene of the Mediterranean fruit fly, *Ceratitis capitata*. *J Mol Evol* 58: 732-742.
- Goodwin SB, M'Barek SB, Dhillon B, Wittenberg AHJ, Crane CF (2001) Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet* 7:e1002070.

- Gray YHM (2000) It takes two transposons to tango - transposable-element-mediated chromosomal rearrangements. *Trends Genet* 16: 461-468.
- Green DM (2004) Structure and evolution of B chromosomes in amphibians. *Cytogenet Genome Res* 106: 235-242.
- Gregory TR (2005) Genome size evolution in animals. In: Gregory TR (Ed.), *The evolution of the genome*. Elsevier, San Diego, CA, pp. 3-87.
- Grumbach S, Tahi F (1994) A new challenge for compression algorithms: genetic sequences. *Inform Process Manag* 30: 875-886.
- Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59: 307-321.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696-704.
- Hancock JM (1999) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: *Microsatellites* (eds. DB Goldstein and C Schlotterer), pp 1-8. Oxford University Press: Oxford.
- Hanski I, Cambefort Y (1991) *Dung Beetle Ecology*. Princeton University Press, Princeton, USA, 481pp.
- Hartl DL, Lohe AR, Lozovskaya ER (1997b) Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu Rev Genet* 31: 337-358.
- Hartl DL, Lozovskaya ER, Lawrence JG (1992) Nonautonomous transposable elements in prokaryotes and eukaryotes. *Genetica* 86: 47-53.
- Hartl DL, Lozovskaya ER, Nurminsky DI, Lohe AR (1997a) What restricts the activity of mariner-like transposable elements? *Trends Genet* 13: 197-201.
- Hedges SB, Blair JE, Venturi ML, Shoe JL (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol Biol* 4:2.
- Hillis DM, Dixon MT (1991) Ribosomal DNA: Molecular evolution and phylogenetic inference. *Q Rev Biol* 66: 411-453.
- Houben A, Kynast RG, Heim U, Hermann H, Jones RN, Forster JW (1996) Molecular cytogenetic characterization of the terminal heterochromatic segment of the B chromosome of rye (*Secale cereale*). *Chromosoma* 105: 97-103.
- Houben A, Leach CR, Verlin D, Rofe R, Timmis JN (1997) A repetitive DNA sequence common to the different B chromosomes of the genus *Brachycome*. *Chromosoma* 106: 513-519.

- Houben A, Nasuda S, Endo TR (2011) Plant B chromosome. *Methods Mol Biol* 701: 97-111.
- Hua-Van A, Le Rouzic A, Maisonhaute C, Capy P (2005) Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. *Cytogenet Genome Res* 110: 426-440.
- Ingram VM (1961) Gene evolution and the haemoglobins. *Nature* 189: 704-708.
- Jakubczak J, Xiong Y, Eickbush TH (1990) Type I (R1) and Type II (R2) ribosomal DNA insertions of *Drosophila melanogaster* are retrotransposable elements closely related to those of *Bombyx mori*. *J Mol Biol* 212: 37-52.
- Jakubczak JL, Burke WD, Eickbush TH (1991) Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *Proc Natl Acad Sci U S A* 88: 3295-3299.
- Jakubczak JL, Zenni MK, Woodruff RC, Eickbush TH (1992) Turnover of R1 (type I) and R2 (type II) retrotransposable elements in the ribosomal DNA of *Drosophila melanogaster*. *Genetics* 131: 129-142.
- James LV, Angus RB (2007) A chromosomal investigation of some British Cantharidae (Coleoptera). *Genetica* 130: 293-300.
- Jarman AP, Wells RA (1989) Hypervariable minisatellites, recombinators or innocent bystanders? *Trends Genet* 5: 367-371.
- Jeffreys AJ, Neumann R, Wilson V (1990) Repeat unit sequence variation in minisatellites: a novel source of DNA polymorphism for studying variation and mutation by single molecule analysis. *Cell* 60: 473-485.
- Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable minisatellite regions in human DNA. *Nature* 314: 67-74.
- John B, Miklos GL (1979) Functional aspects of satellite DNA and heterochromatin. *Int Rev Cytol* 58: 1-114.
- Jones RN (1991) B-chromosome drive. *Am Nat* 137: 430-442.
- Juan C, Petitpierre E (1989) C-banding and DNA content in seven species of Tenebrionidae (Coleoptera). *Genome* 32: 834-839.
- Juan C, Pons J, Petitpierre E (1993) Localization of tandemly repeated DNA sequences in beetle chromosomes by fluorescent *in situ* hybridization. *Chromosome Res* 1: 167-174.
- Junakovic N, Terrinoni A, Di Franco C, Vieira C, Loevenbruck C (1998) Accumulation of transposable elements in the heterochromatin and on the Y chromosome of *Drosophila simulans* and *Drosophila melanogaster*. *J Mol Evol* 46: 661-668.
- Jurka J (2009) LINE retrotransposons from the parasitic wasp *Nasonia vitripennis*. *RR* 9: 483-483.

- Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110: 462-467.
- Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, Ashburner M, Celniker S (2002) The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol* 3: 1-20.
- Kapitonov VV, Jurka J (2001) Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci USA* 98: 8714-8719.
- Kapitonov VV, Jurka J (2008) A universal classification of eukaryotic transposable elements implemented in Repbase. *Nat Rev Genet* 9: 411-412.
- Kazazian HH (2004) Mobile Elements: Drivers of Genome Evolution. *Science* 303: 1626-1632.
- Kean VM, Fox DP, Faulkner R (1982) The accumulation mechanism of the supernumerary (B-) chromosome in *Piceasitchensis* (Bong.) Carr. and the effect of this chromosome on male and female flowering. *Silvae Genet* 31: 126-131.
- Keller I, Chintauan-Marquier IC, Veltsos P, Nichols RA (2006) Ribosomal DNA in the grasshopper *Podisma pedestris*: escape from concerted evolution. *Genetics* 174: 863-874.
- Kidwell MG (1992) Horizontal transfer. *Curr Opin Genet Develop* 2: 868-873.
- Kidwell MG (1993) Lateral transfer in natural populations of eukaryotes. *Annu Rev Genet* 27: 235-256.
- Kidwell MG (2002) Transposable elements and the evolution of genome size in eukaryotes. *Genetica* 115: 49-63.
- Kidwell MG, Lisch DR (2001) Transposable elements, parasitic DNA, and genome evolution. *Evolution* 55: 1-24.
- Kojima KK, Fujiwara H (2004) Cross-genome screening of novel sequence-specific non-LTR retrotransposons: various multicopy RNA genes and microsatellites are selected as targets. *Mol Biol Evol* 21: 201-217.
- Kojima KK, Kuma K, Toh H, Fujiwara H (2006) Identification of rDNA-specific non-LTR retrotransposons in Cnidaria. *Mol Biol Evol* 23: 1984-1993.
- Kordis D, Gubensek F (1998) Unusual horizontal transfer of a long interspersed nuclear element between distant vertebrate classes. *Proc Natl Acad Sci USA* 95: 10704-10709.
- Kubis S, Schmidt T, Heslop-Harrison JS (1998) Repetitive DNA elements as a major

- component of plant genomes. *Ann Bot* 82: 45-55.
- Kurlang C, Canback B, Berg O (2003) Horizontal gene transfer: critical view. *Proc Natl Acad Sci U S A* 100: 9658-9662.
- Lampe DJ, Churchill MEA, Robertson HM (1996) A purified *mariner* transposase is sufficient to mediate transposition in vitro. *EMBO J* 15: 5470-5479.
- Lampe DJ, Witherspoon DJ, Soto-Adames FN, Robertson HM (2003) Recent horizontal transfer of *Mellifera* subfamily *Mariner* transposons into insect lineages representing four diff. *Mol Biol Evol* 20: 554-562.
- Landaïs I, Chavigny P, Castagnone C, Pizzol J, Abad P, Vanlerberghe-Masutti F (2000) Characterization of a highly conserved satellite DNA from the parasitoid wasp *Trichogramma brassicae*. *Gene* 255: 65-73.
- Langley CH, Montgomery E, Hudson R, Kaplan N, Charlesworth B (1988). On the role of unequal exchange in the containment of transposable element copy number. *Genet Res* 52: 223-235.
- Lankenau D-H, Volff J-N (2009) Transposable and the dynamic genome. Lankenau D-H, Volff J-N, eds. Springer-Verlag Berlin Heidelberg, German, 184pp.
- Lathe WC III, Burke WD, Eickbush DG, Eickbush TH (1995) Evolutionary stability of the R1 retrotransposable element in the genus *Drosophila*. *Mol Biol Evol* 12: 1094- 2005.
- Lathe WD III, Eickbush TH (1997) A single lineage of r2 retrotransposable elements is an active, evolutionarily stable component of the *Drosophila* rDNA locus. *Mol Biol Evol* 14: 1232-1241.
- Leaver M (2001) A Family of Tc1-like transposon from the genomes of fishes and frogs: evidence for horizontal transmission. *Gene* 271: 203-214.
- Lerat E, Rizzon C, Biémont C (2003) Sequence divergence within transposable element families in the *Drosophila melanogaster*. *Genome Res* 13: 1889-1896.
- Li Y, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol Ecol* 11: 2453-2465.
- Liao D (1999) Concerted Evolution: Molecular Mechanism and Biological Implications. *Am J Hum Genet* 64: 24-30.
- Liao D (2000) Gene conversion drives within genic sequences: concerted evolution of ribosomal RNA genes in Bacteria and Archaea. *J Mol Evol* 51: 305-317.
- Lohe AR, Roberts PA (2000) Evolution of DNA in heterochromatin: the *Drosophila melanogaster* sibling species subgroup as a resource. *Genetica* 109: 125-130.
- Long EO, Dawid IB (1980) Repeated genes in eukaryotes. *Annu Rev Biochem* 49: 727-764.

- Long EO, Dawid IB (1980) Repeated genes in eukaryotes. *Annu Rev Biochem* 49: 727-764.
- Loreto EL, Valente VL, Zaha A, Silva JC, Kidwell MG (2001) *Drosophila mediopunctata* P elements: a new example of horizontal transfer. *J Hered* 92: 375-381.
- Loreto V, Cabrero J, López-León MD, Camacho JPM, Souza MJ (2008) Possible autosomal origin of macro B chromosomes in two grasshopper species. *Chromosome Res* 16: 233-241.
- Loreto V, Stadtler E, Melo NF, Souza MJ (2005). A comparative cytogenetic analysis between the grasshopper species *Chromacris nuptialis* and *C. speciosa* (Romaleidae): constitutive heterochromatin variability and rDNA sites. *Genetica* 125: 2530-260.
- Lorite P, Palomeque T, Garnería I, Petitpierre E (2001) Characterization and chromosome location of satellite DNA in the leaf beetle *Chrysolina americana* (Coleoptera, Chrysomelidae). *Genetica* 110: 143-150.
- Louzada JN (2008) Scarabaeinae (Coleoptera: Scarabaeidae) detritívoros em ecossistemas tropicais: biodiversidade e serviços ambientais. In: Moreira FMS, Siqueira JO, Brussaard L (eds): *Biodiversidade do Solo em Ecossistemas Brasileiros*, pp 309-332 (UFLA, Lavras 2008).
- Luchetti A, Mingazzini V, Mantovani B (2012) 28S junctions and chimeric elements of the rDNA targeting non-LTR retrotransposon R2 in crustacean living fossils (Branchiopoda, Notostraca). *Genomics* 100: 51-56.
- Makalowski M (2003) Not junk after all. *Science* 23: 1246-1247.
- Malik HS, Burke WD, Eickbush TH (1999) The age and evolution of non-LTR retrotransposable elements. *Mol Biol Evol* 16: 793-805.
- Mandrioli M (2003). Identification and chromosomal localization of *Mariner*-like elements in the cabbage moth *Mamestra brassicae* (Lepidoptera). *Chromosome Res* 11: 319-322.
- Marschner S, Meister A, Blattner FR, Houben A (2007) Evolution and function of B chromosome 45S rDNA sequences in *Brachycome dichromosomatica*. *Genome* 50: 638-644.
- Martins C, Cabral-de-Mello DC, Valente GT, Mazzuchelli J, Oliveira SG, Pinhal D (2011) *Animal genomes under the focus of cytogenetics*. Hauppauge: Nova Science Publisher, 1 ed., 160p.
- Martins C, Galletti Jr PM (2001) Two rDNA arrays in Neotropical fish species: is it a general rule for fishes? *Genetica* 111: 439-446.
- Martins VG (1994) The chromosome of five species of Scarabaeidae (Polyphaga, Coleoptera). *Naturalia* 19: 89-96.

- Maruyama K, Hartl DL (1991) Evidence for interspecific transfer of the transposable element mariner between *Drosophila* and *Zaprionus*. *J Mol Evol* 33: 514-524.
- McAllister BF, Werren JH (1997) Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*. *Chromosoma* 106: 243-253.
- McDonald JF (1993) Evolution and consequences of transposable elements. *Curr Opin Genet Develop* 3: 855-864.
- Meredith RW, Janecka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC et al. (2011) Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* 334: 521-524.
- Mesa A, Fontanetti C (1985) The chromosomes of a primitive species of beetle: *Ytu zeus* (Coleoptera, Myxophaga, Torridincolidae). *Acad Nat Phil* 137: 102-105.
- Mingazzini V, Luchetti A, Mantovani B (2010) R2 dynamics in *Triops cancriformis* (Bosc, 1801) (Crustacea, Branchiopoda, Notostraca): turnover rate and 28S concerted evolution. *Heredity*: 1-9.
- Miskey C, Izsvák Z, Kawakami K, Ivics Z (2005) DNA transposons in vertebrate functional genomics. *Cell Mol Life Sci* 62: 629-641.
- Montgomery EA, Huang SM, Langley CH, Judd BH (1991) Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. *Genetics* 129: 1085-1098
- Montiel EE, Cabrero J, Camacho JPM, López-León MD (2012) *Gypsy*, *RTE* and *Mariner* transposable elements populate *Eyprepocnemis plorans* genome. *Genetica* 140: 365-374.
- Moreau CS, Bell CD, Vila R, Archibald SB, Pierce NE (2006) Phylogeny of the ants: diversification in the age of angiosperms. *Science* 312:101-104.
- Moura RC, Melo NF, Souza MJ (2008) High levels of chromosomal differentiation in *Euchroma gigantea* L 1735 (Coleoptera, Buprestidae). *Gen Mol Biol* 31: 431-437.
- Moura RC, Souza MJ, Melo NF, Lira-Neto AC (2003) Karyotypic characterization of representatives from Melolonthinae (Coleoptera: Scarabaeidae): karyotypic analysis, banding and fluorescent *in situ* hybridization (FISH). *Hereditas* 138: 200-206.
- Mravinac B, Plohl M (2007) Satellite DNA junctions identify the potential origin of new repetitive elements in the beetle *Tribolium madens*. *Gene* 394: 45-52.
- Mravinac B, Plohl M, Ugarkovic D (2004) Conserved patterns in the evolution of *Tribolium* satellite DNAs. *Gene* 332: 169-177.
- Mravinac B, Ugarkovic D, Franjevic D, Plohl M (2005) Long inversely oriented subunits form

- a complex monomer of *Tribolium brevicornis* satellite DNA. J Mol Evol 60: 513-525.
- Mueller UG, Wolfenbarger LW (1999) AFLP genotyping and fingerprinting. TREE 14: 389-394.
- Mukha DV, Mysina V, Mayropulo V, Schal C (2010). Structure and molecular evolution of the ribosomal DNA external transcribed spacer in the cockroach genus *Blatella*. Genome 54: 222-234.
- Murphy WJ, Pringle TH, Crider TA, Springer MS, Miller W (2007) Using genomic data to unravel the root of the placental mammal phylogeny. Genome Res 17: 413-421.
- Nakakita H, Omura O, Winks RG (1981) Hybridization between *Tribolium freemani* (Hinton) and *Tribolium castaneum* (Herbst) and some preliminary studies on the biology of *Tribolium freemani* (Coleoptera: Tenebrionidae). Appl Entomol Zool 16: 209-215.
- Nei M, Gu X, Sitnikova T (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci U S A 94: 7799-806.
- Nei M, Rooney AP (2005) Concerted and birth-and-death evolution of multigene families. Ann Rev Genet 39:121-152.
- Noller HF, Hoffarth V, Zimniak L (1992) Unusual resistance of peptidyl transferase to protein extraction procedures. Science 256: 1416-1419.
- Nowak R (1994) Mining treasures from junk DNA. Science 263: 608-610.
- O'Donnell KA, Burns KH (2010) Mobilizing diversity: transposable element insertions in genetic variation and disease. Mobile DNA 1:21.
- Okazaki S, Ishikawa H, Fujiwara H (1995) Structural analysis of TRAS1, a novel family of telomeric repeat- associated retrotransposons in the silkworm, *Bombyx mori*. Mol Cell Biol 15: 4545-4552.
- Oliveira SG, Bao W, Martins C, Jurka J (2012c) Horizontal transfers of *Mariner* transposons between mammals and insects. Mobile DNA 3:14.
- Oliveira SG, Cabral-de-Mello DC, Arcanjo AP, Xavier C, Souza MJ, Martins C, Moura RC (2012a) Heterochromatin, sex chromosomes and rRNA gene clusters in *Coprophanaeus* beetles (Coleoptera, Scarabaeidae). Cytogenet Genome Res 138: 46-55.
- Oliveira SG, Moura RC, Martins C (2012b) B chromosome in the beetle *Coprophanaeus cyaneus* (Scarabaeidae): emphasis in the organization of repetitive DNA sequences. BMC Genetics 13:96.
- Oliveira SG, Moura RC, Silva AEB, Souza MJ (2010) Cytogenetic analysis of two *Coprophanaeus* species (Scarabaeidae) revealing wide constitutive heterochromatin variability and the largest number of 45S rDNA sites among Coleoptera. Micron 41:

960-965.

- Pace JK, II, Gilbert C, Clark MS, Feschotte C (2008) Repeated horizontal transfer of a DNA transposon in mammals and other tetrapods. *Proc Natl Acad Sci USA* 105: 17023–17028.
- Palestis, BG, Trivers R, Burt A, Jones RN (2004) The distribution of B chromosomes across species. *Cytogenet Genome Res* 106: 151-158.
- Palomeque T, Lorite P (2008) Satellite DNA in insects: a review. *Heredity* 100: 564-573.
- Palomeque T, Muñoz-López M, Carrillo JA, Lorite P (2005) Characterization and evolutionary dynamics of a complex family of satellite DNA in the leaf beetle *Chrysolina carnifex* (Coleoptera, Chrysomelidae). *Chromosome Res* 13: 795-807.
- Pavlopoulos A, Berghammer AJ, Averof M, Klingler M (2004) Efficient transformation of the beetle *Tribolium castaneum* using the *Minos* transposable element: quantitative and qualitative analysis of genomic integration events. *Genetics* 167:737–746.
- Pavlopoulos A, Oehler S, Kapetanaki MG (2007). The DNA transposon *Minos* as a tool for transgenesis and functional genomic analysis in vertebrates and invertebrates. *Genome Biol* 8(Suppl 1): S2
- Perfectti F, Werren JH (2001) The interspecific origin of B chromosomes: experimental evidence. *Evolution* 55: 1069-1073.
- Petrov DA (2001) Evolution of genome size: new approaches to an old problem. *Opinion, Trends Genet* 17: 23-28.
- Petrov DA, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE (2003) Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. *Mol Biol Evol* 20: 880-892.
- Philips TK, Edmonds WD, Scholtz CH (2004) A phylogenetic analysis of the New World tribe Phanaeini (Coleoptera: Scarabaeidae: Scarabaeinae): hypotheses on relationships and origins. *Insect Syst Evol* 35: 43-63.
- Pikaard C, Pontes O (2007) Heterochromatin: condense or excise. *Nature Cell Biol* 9: 19-20.
- Piskurek O, Okada N (2007) Poxviruses as possible vectors for horizontal transfer of retroposons from reptiles to mammals. *Proc Natl Acad Sci USA* 104: 12046-12051.
- Plasterk RH (1996) The Tc1/mariner transposon family. *Curr Top Microbiol Immunol* 204: 125-143.
- Plasterk RHA, Izsvák Z, Ivics Z (1999) Resident aliens: the Tc1/mariner superfamily of transposable elements. *Trends Genet* 15: 326-332.
- Plohl M, Lucijanac-Justic V, Ugarkovic D, Petitpierre E, Juan C (1993) Satellite DNA and

- heterochromatin of the flour beetle *Tribolium confusum*. *Genome* 36: 467-475.
- Plohl M, Ugarkovic D (1994) Characterization of two abundant satellite DNAs from the Mealworm *Tenebrio obscurus*. *J Mol Evol* 39: 489-495.
- Pons J (2003) Cloning and characterization of a transposable-like repeat in the heterochromatin of the darkling beetle *Misolampus goudoti*. *Genome* 47: 769-774.
- Pons J, Bruvo B, Juan C, Petitpierre E, Plohl M, Ugarkovic D (1997) Conservation of satellite DNA in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). *Gene* 255: 183-190.
- Pons J, Bruvo B, Petitpierre E, Plohl M, Ugarkovic D, Juan C. (2004) Complex structural features of satellite DNA sequences in the genus *Pimelia* (Coleoptera: Tenebrionidae): random differential amplification from a common 'satellite DNA library. *Heredity* 92: 418-427.
- Pons J, Petitpierre E, Juan C (1993) Characterization of the heterochromatin of the darkling beetle *Misolampus goudoti*: cloning of two satellite DNA families and digestion of chromosomes with restriction enzymes. *Hereditas* 119: 179-185.
- Pons J, Petitpierre E, Juan C (2002) Evolutionary dynamics of satellite DNA family PIM357 in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). *Mol Biol Evol* 19: 1329-1340.
- Pritham EJ (2009) Transposable elements and factors influencing their success in eukaryotes. *J Hered* 100: 648-655.
- Proença SJR, Serrano ARM, Collares-Pereira MJ (2002) Cytogenetic variability in genus *Odontocheila* (Coleoptera, Cicindelidae): karyotypes, C-banding, NORs and localization of ribosomal genes of *O. confusa* and *O. nodicornis*. *Genetica* 114: 237-245.
- Puerma E, Acosta MJ, Barragán MJ, Martínez S, Marchal JÁ, et al (2008) The karyotype and 5S rRNA genes from Spanish individuals of the bat species *Rhinolophus hipposideros* (Rhinolophidae; Chiroptera). *Genetica* 134: 287-295.
- Rao SR, Trivedi S, Emmanuel D, Merita K, Hynniewta M (2010) DNA repetitive sequences: types, distribution and function: A review. *JCMB* 7-8: 1-11.
- Raska I, Koberna K, Malinsky J, Fidlerová H, Masata M (2004) The nucleolus and transcription of ribosomal genes. *Biol Cell* 96: 579-594.
- Ray DA, Feschotte C, Pagan HJT, Smith JD, Pritham EJ, et al. (2008). Multiple waves of recent DNA transposon activity in the bat, *Myotis lucifugus*. *Genome Res* 18: 717-728.
- Rho M, Tang H (2009) MGEScan-non-LTR: computational identification and classification of autonomous non-LTR retrotransposons in eukaryotic genomes. *Nucleic Acids Res*

37:e143.

- Ribeiro LF, Fernandez MA (2004) Molecular characterization of the 5S ribosomal gene of the *Bradysia hygida* (Diptera: Sciaridae). *Genetica* 122: 253-260.
- Richards RI, Sutherland GR (1994) Simple repeat DNA is not replicated simply. *Nat Genet* 6: 114-116.
- Robertson HM (1993) The mariner transposable element is widespread in insects. *Nature* 362: 241-245.
- Robertson HM (1995) The Tc1-mariner superfamily of transposons in animals. *J of Insect Physiol* 41: 99-105.
- Robertson HM, Lampe DJ (1995a) Distribution of transposable elements in arthropods. *Annu Rev Entomol* 40: 333-57.
- Robertson HM, Lampe DJ (1995b) Recent horizontal transfer of a mariner transposable element among and between Diptera and Neuroptera. *Mol Biol Evol* 12: 850-862.
- Robertson HM, MacLeod EG (1993) Five major subfamilies of mariner transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect Mol Biol* 2: 125-139.
- Robertson HM, Zumpano KL (1997) Molecular evolution of an ancient mariner transposon, Hsmar1, in the human genome. *Gene* 205: 203-217.
- Röder M, korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A Microsatellite Map of Wheat. *Genetics* 149: 2007-2023.
- Roiha H, Miller R, Woods LC, Glover DM (1981) Arrangements and rearrangements of sequences flanking the two types of rDNA insertion in *D. melanogaster*. *Nature* 290: 749-754.
- Rooney AP (2004) Mechanisms underlying the evolution and maintenance of functionally heterogeneous 18S rRNA genes in Apicomplexans. *Mol Biol Evol* 21: 1704-1711.
- Rooney AP, Ward TJ (2005) Evolution of a large ribosomal RNA multigene family in filamentous fungi: birth and death of a concerted evolution paradigm. *Proc Natl Acad Sci U S A* 102: 5084-5089.
- Roussel P, Hernandez-Verdun D (1994) Identification of Ag-NOR proteins, markers of proliferation related to ribosomal genes activity. *Exp Cell Res* 214: 465-472.
- Rouzik A, Capy P (2009) Theoretical approaches to the dynamics of transposable elements in genomes, populations, and species. In: *Transposable and the Dynamic Genome*. Lankenau D-H, Volff J-N (eds.). Springer-Verlag Berlin Heidelberg, German, 184pp.
- Rozék M, Lachowska D, Petitpierre E, Holecová M (2004) C-bands on chromosomes of 32

- beetles species (Coleoptera: Elateridae, Cantharidae, Oedemeridae, Cerambycidae, Anthicidae, Chrysomelidae, Attelabidae and Curculionidae). *Hereditas* 140: 161-170.
- Sambrook J, Russel DW (2001) *Molecular cloning. A laboratory manual*, Third Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sánchez-Gea JF, Serrano J, Gálian J (2000) Variability in rDNA loci in Iberian species of the genus *Zabrus* (Coleoptera: Carabidae) detected by fluorescence *in situ* hybridization. *Genome* 43: 22-28.
- Sanchez-Gracia A, Maside X, Charlesworth B (2005) High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet* 21: 200-203.
- Schaack S, Gilbert C, Feschotte C (2010) Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* 25: 537-546.
- Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109: 365-371.
- Schmidt T, Heslop-Harrison JS (1998) Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci* 3: 195-199.
- Schneider MC, Rosa SP, Almeida MC, Costa C, Cella DM (2007) Chromosomal similarities and differences among four Neotropical Elateridae (Conoderini and Pyrophorini) and other related species, with comments on the NOR patterns in Coleoptera. *J Zool Syst Evol Res* 45: 308-316.
- Schweizer D (1976) Reverse fluorescent chromosome banding with chromomycin and DAPI. *Chromosoma* 58: 307-324.
- Schweizer D, Loidl J (1987) A model for heterochromatin dispersion and the evolution of C-bands patterns. *Chromosome Today* 9: 61-74.
- Shapiro JA, Sternberg R (2005) Why repetitive DNA is essential to genome function. *Biol Rev* 80: 1-24.
- Sharbel TF, Green DM, Houben A (1998) B chromosome origin in the endemic New Zealand frog *Leiopelma hochstetteri* through sex chromosome evolution. *Genome* 41: 14-22.
- Silva GM, Bione EG, Cabral-de-Mello DC, Moura RC, Simões ZL, Souza MJ (2009) Comparative cytogenetics of three species of *Dichotomius* (Coleoptera, Scarabaeidae). *Genet Mol Biol* 32: 276-280.
- Silva JC, Kidwell MG (2000) Horizontal transfer and selection in the evolution of P elements. *Mol Biol Evol* 17:1542–1557.
- Silva JC, Loreto EL, Clark JB (2004) Factors that affect the horizontal transfer of

- transposable elements. *Curr Issues Mol Biol* 6: 57-71.
- Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nature Rev Genet* 8: 272–285.
- Smit S, Widmann J, Knight R (2007) Evolutionary rates vary among rRNA structural elements. *Nucleic Acids Res* 35: 3339-3354.
- Smith SG, Virkki N (1978) Coleoptera. In: John B (ed) *Animal Cytogenetics*. Borntraeger, Berlin, Stuttgart, 366 pp.
- Sormacheva I, Smyshlyaev G, Mayorov V, Blinov A, Novikov A, Novikovaz O (2012) Vertical evolution and horizontal transfer of CR1 non-LTR Retrotransposons and Tc1/mariner DNA transposons in Lepidoptera species. *Mol Biol Evol* 29: 3685-3702.
- Stage DE, Eickbush TH (2007) Sequence variation within the rRNA gene loci of 12 *Drosophila* species. *Genome Res* 17: 1888-1897.
- Stage DE, Eickbush TH (2010) Maintenance of multiple lineages of R1 and R2 retrotransposable elements in the ribosomal RNA gene loci of *Nasonia*. *Insect Mol Biol* 1: 37-48.
- Steinemann M, Steinemann S (1992). Degenerating Y chromosome of *Drosophila miranda*: a trap for retrotransposons. *Proc Natl Acad Sci USA* 89: 7591-595.
- Sumner AT (1972) A simple technique for demonstrating centromeric heterochromatin. *Exp Cell Res* 75: 304-306.
- Sumner AT (2003) *Chromosomes: Organization and Function*. Blackwell Science Ltd, Norths Berwich, UK, 1st ed., 287pp.
- Syvanen M (1994) Horizontal gene transfer: evidence and possible consequences. *Annu Rev Genet* 28: 237-261.
- Takahashi H, Fujiwara H (2002) Transplantation of target site specificity by swapping the endonuclease domains of two LINEs. *EMBO J* 21: 408-417.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599.
- Tautz D, Hancock JM, Webb DA, Tautz C, Dover GA (1988) Complete sequences of the rRNA genes of *Drosophila melanogaster*. *Mol Biol Evol* 5: 366-376.
- Thomas CA Jr (1971) The genetic organization of chromosomes. *Ann Rev Genet* 5: 237- 256.
- Torti C, Gomulski LM, Moralli D, Raimond E, Robertson HM, et al. (2000) Evolution of different subfamilies of *mariner* elements within the medfly genome inferred from abundance and chromosomal distribution. *Chromosoma* 108: 523-532.

- Treangen TJ, Salzberg SL (2012) Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat Rev Genet* 13: 36-46.
- Ugarkovic D, Durajlija S, Plohl M (1996) Evolution of *Tribolium madens* (Insecta, Coleoptera) satellite DNA through DNA inversion and insertion. *J Mol Evol* 42: 350-358.
- Ugarkovic D, Petitpierre E, Juan C, Plohl M (1995) Satellite DNAs in tenebrionid species: structure, organization and evolution. *Croat Chem Acta* 68: 627-638.
- Ugarkovic D, Plohl M (2002). Variation in satellite DNA profiles-causes and effects. *EMBO J* 21: 5955-5959.
- Ugarkovic Đ, Plohl M, Lucijanic-Justic V, Borstnik B (1992) Detection of satellite DNA in *Palorus ratzeburgii*: analysis of curvature profiles and comparison with *Tenebrio molitor* satellite DNA. *Biochimie* 74: 1075-1082.
- Vaz-de-Mello FZ (2000) Estado atual de conhecimento dos Scarabaeidae s. str. (Coleoptera: Scarabaeoidea) do Brasil, in Martin-Piera F, Morrone JJ, Melic A (eds): Hacia un Proyecto CYTED para el inventario y estimación de la diversidad Entomológica en Iberoamérica: PrIBES-2000, pp. 183-195, 326 (Sociedad Entomológica Aragonesa & CYTED, m3m: Monografias Tercer Milenio, Zaragoza).
- Vergnaud G, Denoeud F (2000) Minisatellites: Mutability and genome architecture. *Genome Res* 10: 899-907.
- Vidal OR (1984) Chromosome numbers of Coleoptera from Argentina. *Genetica* 65: 235-239.
- Vidal OR, Giacomozzi RO (1978) Los cromosomas de la subfamilia Dynastinae (Coleoptera, Scarabaeidae). II. Las bandas C em *Enema pan* (Fabr.). *Physis* 38: 113-119.
- Vidal OR, Nocera CP (1984) Citogenética de la tribo Eucranini (Coleoptera, Scarabaeidae). Estudios convencionales y com bandeio C. *Physis* 42: 83-90.
- Virkki N (1983) Banding of Oedionychina (Coleoptera, Alticinae) chromosomes. C- and Ag-bands. *J Agr U Puerto Rico* 67: 221-255.
- Vitturi R, Colomba MS, Barbieri R, Zunino M (1999) Ribosomal DNA location in the scarab beetle *Thorectes intermedius* (Costa) (Coleoptera: Geotrupidae) using banding and fluorescent *in-situ* hybridization. *Chromosome Res* 7: 255-260.
- Vitturi R, Colomba MS, Pirrone AM, Mandrioli M (2002) rDNA (18S-28S and 5S) colocalization and linkage between ribosomal genes (TTAGGG)_n telomeric sequence in the earthworm, *Octodrilus complanatus* (Annelida: Oligochaeta: Lumbricidae), revealed by single- and double-color FISH. *J Hered* 93: 279-282.
- Vitturi R, Colomba MS, Volpe N, Lannino A, Zunino M (2003) Evidence for male X0 sex-

- chromosome system in *Pentodon bidens punctatum* (Coleoptera: Scarabaeoidea: Scarabaeidae) with X-linked 18S–28S clusters. *Genes Genet Syst* 78: 427-432.
- Vivares CP (1999) On the genome of microsporidia. *J Eukaryot Microbiol* 46: 16A.
- Vogler AP, DeSalle R (1993) Phylogeographic patterns in coastal North American Tiger Beetles, *Cicindela dorsalis*, inferred from mitochondrial DNA sequences. *Evolution* 47: 1192-1202.
- Vogt P (1992) Code domains in tandem repetitive DNA sequence structures. *Chromosoma* 101: 585-589.
- Volff J-N, Lehrach H, Reinhardt R, Chourroutà D (2004) Retroelement dynamics and a novel type of chordate retrovirus-like element in the miniature genome of the tunicate *Oikopleura dioica*. *Mol Biol Evol* 21: 2022-2033.
- von Stenberg RM, Novick GE, Gao G-P, Herrera RJ (1992) Genome canalization: the coevolution of transposable and interspersed repetitive elements with single copy DNA. *Genetica* 86: 106-137.
- Walisko O, Jursch T, Izavák Z, Ivics Z (2009) Transposon-host cell interactions in the regulation of Sleeping Beauty transposition. In: D-H Lankenau, J-N Volff: *Transposons and the Dynamic Genome*, Springer-Verlag Berlin Heidelberg, 109-132.
- Wang S, Lorenzen MD, Beeman RW, Brown SJ (2008) Analysis of repetitive DNA distribution patterns in the *Tribolium castaneum* genome. *Genome Biol* 9: R61.
- White MJD (1973) *Animal Cytology and Evolution*. Cambridge University, London, UK, 3rd ed., 959pp
- Wicker TW, Sabot F, Hua-Van A, Bennetzen JL, Capy P, et al (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev* 8: 973-982.
- Wilson CJ, Angus RB (2004) A chromosomal analysis of ten European species of *Aphodius* Illiger, subgenera *Acrossus* Mulsant, *Nimbus* Muldant & Rey and *Chilothorax* Motschulsky (Coleoptera: Aphodiidae). *Koleopt Rdsch* 74: 367-374.
- Wilson CJ, Angus RB (2005) A chromosomal analysis of 21 species of Oniticeellini and Onthophagini (Coleoptera: Scarabaeidae). *Tijdschr Entomol* 148: 63-76.
- Wilson CJ, Angus RB (2006) A chromosomal analysis of eight species of *Aphodius* Illiger, subgenera *Agiolinus* Schmidt, *Agrilinus* Mulsant & Rey and *Planolinus* Mulsant & Rey (Coleoptera: Aphodiidae). *Proc Russ Entomol Soc* 77: 28-33.
- Wolff RK, Plaeke R, Jeffreys AJ, White R (1991) Unequal crossing over between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci. *Genomics* 5: 382-384.

- Wong LH, Choo KH (2004) Evolutionary dynamics of transposable elements at the centromere. *Trends Genet* 20: 611-616.
- Xiong Y, Eickbush TH (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* 9: 3353-3362.
- Yadav JS, Pillai RK (1979) Evolution of karyotypes and phylogenetic relationships in Scarabaeidae (Coleoptera). *Zool Anz* 202: 105-118.
- Yadav JS, Pillai RK, Karamjeet (1979) Chromosome numbers of Scarabaeidae (Polyphaga: Coleoptera). *Coleopts Bull* 33: 309-318.
- Yadav JS, Pillai RK, Yadav AS (1990) Karyotypic study of some scarab beetles with comments on phylogeny (Coleoptera: Scarabaeoidea). *Elytron* 4: 41-51.
- Yunis JJ, Yasmineh WG (1971) Heterochromatin, satellite DNA, and cell function. *Science* 174: 1200-1209.
- Zimmer EA, Martin SL, Beverley SM, Kan YW, Wilson AC (1980) Rapid duplication and loss of genes coding for the alpha chains of hemoglobin. *Proc Natl Acad Sci U S A* 77:2158-2162.
- Zinic SJ, Ugarjovik D, Cornudella L, Plohl M (2000) A novel interspersed type of organization of satellite DNAs in *Tribolium madens* heterochromatin. *Chromosome Res* 8: 201-212.
- Zwick MS, Hanson RE, McKnight TD, Nurul-Islam-Faridi M, Stelly DM (1997) A rapid procedure for the isolation of *C0t*-1 DNA from plants. *Genome* 40: 138-142.

ANEXOS

Os anexos da presente tese se organizam da seguinte maneira:

Anexos A: Trabalhos da tese publicados

Anexo A1: Heterochromatin, sex chromosomes and rRNA gene clusters in *Coprophanaeus* beetles (Coleoptera, Scarabaeidae)

Anexo A2: B chromosome in the beetle *Coprophanaeus cyanescens* (Scarabaeidae): emphasis in the organization of repetitive DNA sequences

Anexo A3: Horizontal transfers of *Mariner* transposons between mammals and insects

Anexo B: Relação de trabalhos paralelos publicados relacionados ao tema da tese

Heterochromatin, Sex Chromosomes and rRNA Gene Clusters in *Coprophanaeus* Beetles (Coleoptera, Scarabaeidae)

S.G. Oliveira^a D.C. Cabral-de-Mello^b A.P. Arcanjo^c C. Xavier^d M.J. Souza^c
C. Martins^a R.C. Moura^d

^aDepartamento de Morfologia, Instituto de Biociências, Universidade Estadual Paulista/UNESP, Botucatu,

^bDepartamento de Biologia, Instituto de Biociências, Universidade Estadual Paulista/UNESP, Rio Claro,

^cDepartamento de Genética, Centro de Ciências Biológicas/CCB, Universidade Federal de Pernambuco/UFPE, and

^dDepartamento de Biologia, Instituto de Ciências Biológicas/ICB, Universidade de Pernambuco/UPE, Recife, Brazil

Key Words

Chromosomal evolution • Comparative cytogenetics • Karyotype • Meiosis • Multigene family • Repetitive DNA

Abstract

Repetitive DNA sequences constitute a high fraction of eukaryotic genomes and are considered a key component for the chromosome and karyotype evolution. For a better understanding of their evolutionary role in beetles, we examined the chromosomes of 5 species of the genus *Coprophanaeus* by C-banding, fluorochrome staining CMA₃/DA/DAPI, and fluorescence in situ hybridization (FISH) with probes for 18S and 5S rRNA genes. The *Coprophanaeus* species have identical chromosome numbers and a conserved chromosome morphology. However, they show different sex chromosome forms, XY, Xy, XY_p, and heterochromatin seems to be involved in the origin and diversification of these forms. C-banding showed primarily the presence of di-phasic chromosomes in all species examined. After CMA₃/DA/DAPI staining, 1–9 autosomal pairs showed CMA₃-positive blocks depending on the species, while DAPI-positive blocks were detected only in *Coprophanaeus dardanus*. FISH mapping revealed 5S rDNA signals in one autosomal pair in each species, whereas the number of pairs with 18S rDNA

signals varied from 1–8 between the *Coprophanaeus* species. Our results suggest that distinct genetic mechanisms had been involved in the karyotype evolution of *Coprophanaeus* species, i.e. mechanisms maintaining the conserved number of 5S rDNA clusters and those generating variability in the amount of heterochromatin, sex chromosome forms, and distribution of 18S rDNA clusters.

Copyright © 2012 S. Karger AG, Basel

A large fraction of eukaryotic genomes is composed of repetitive DNA sequences. Although these sequences were long considered 'junk' DNA, they are now considered a major player in genomic architecture and function and are thought to play a role in the structural and functional organization of chromosomes [Nowak, 1994; Biémont and Vieira, 2006]. The variation in total DNA content between species is mostly due to variation in the repetitive DNA content which is often present at many chromosomal sites [Sumner, 2003]. Repetitive DNA sequences have been extensively explored as markers for cytogenetic mapping because their reiterated number of copies generates easily visualizable signals on chromosomes. The cytogenetic mapping of repetitive DNA provides useful chromosome markers that can be used in

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

© 2012 S. Karger AG, Basel
1424–8581/12/1381–0046\$38.00/0

Accessible online at:
www.karger.com/cgr

Cesar Martins
UNESP – Univ. Estadual Paulista
Instituto de Biociências/IB, Departamento de Morfologia
CEP 18618-900 Botucatu, SP (Brazil)
Tel. +55 14 3811 6264, E-Mail cmartins@ibb.unesp.br

studies of genome organization, species evolution, and applied genetics and in the identification of specific chromosomes, homologous chromosomes, chromosome rearrangements, and sex chromosomes.

The subfamily Scarabaeinae (Scarabaeidae, Polyphaga) comprises a diverse and cosmopolitan group of Coleoptera that play an important role in the conservation of ecosystems as seed dispersers, pollinators, and recyclers of organic matter [Louzada, 2008]. The group encompasses approximately 250 genera and over 6,000 species, with 618 species belonging to 49 genera in Brazil [Hanski and Cambefort, 1991; Vaz-de-Mello, 2000]. Among them, the tribe Phanaeini comprises approximately 150 species distributed in 12 genera, of which *Gromphas* and *Oruscatus* are considered the most basal taxa, while *Coprophanaeus*, *Dendropaemon*, *Megatharsis*, and *Tetramereia* are more derived [Philips et al., 2004]. The phanaeines are restricted to the Neotropics [Edmonds, 1967, 1972; Philips et al., 2004], and their representatives have been detected in all Brazilian regions [Edmonds and Zidek, 2010].

The karyotype variation observed for Polyphaga coleopterans is clearly represented in the high diversity of sex chromosomes identified in the group that includes the X0, XY, neo-XY, X_y_p, XY_p, XY_r, XY_c, and multiple systems [White, 1973; Smith and Virkki, 1978; Ferreira et al., 1984; Mesa and Fontanetti, 1985; Yadav et al., 1990; Galián et al., 2002; Rozék et al., 2004; Cabral-de-Mello et al., 2010a]. In the different forms of association between X and Y chromosomes the 'y' represents a smaller chromosome than the X, the 'p' refers to a 'parachute' meiotic conformation between the X and Y(y), 'r' refers to a road-shaped configuration, and 'c' to a rare central association of X and Y. The Scarabaeinae species follows the high level of karyotype diversity of the group, with variation in the diploid number from 2n = 8 in *Eurysternus caribaeus* to 2n = 24 in *Oniticellus spinipes* (= *Timiocellus spinipes*) and 7 types of sex chromosomes mechanisms (XY_p, X_y_p, XY, X_y, XY_r, neo-XY, and X0) [Smith and Virkki, 1978; Yadav et al., 1979; Cabral-de-Mello et al., 2008]. Although diverse with respect to their chromosomal characteristics, there is a predominance of the 2n = 20 karyotype, XY_p sex chromosome mechanism, and meta-submetacentric chromosome morphology among Scarabaeinae species. Among the 12 tribes of Scarabaeinae, Phanaeini predominantly present a chromosome number of 2n = 20 and a high variability among the sex chromosomes mechanisms (XY, X_y, XY_p, X_y_p, neo-XY, and X0) [Cabral-de-Mello et al., 2008, 2010c; Oliveira et al., 2010].

To characterize the macro-chromosomal evolution, sex chromosomes, and diversification of repetitive DNA, in-

cluding heterochromatin and rDNAs, the chromosomes of 5 species of *Coprophanaeus* belonging to the 3 subgenera of the group (*Coprophanaeus*, *Megaphanaeus*, and *Metallophanaeus*) were studied with classical cytogenetic techniques (conventional staining, C-banding, and fluorochrome CMA₃/DA/DAPI staining) and fluorescence in situ hybridization (FISH) using probes for the 18S and 5S rRNA genes. With the exception of previously published information concerning the basic karyotype data of *C. (Coprophanaeus) dardanus* [Cabral-de-Mello et al., 2008], this study is the first to describe the basic karyotypes, the heterochromatin composition and distribution, and the rRNA gene organization of *C. (Coprophanaeus) acrisius*, *C. (Coprophanaeus) dardanus*, *C. (Metallophanaeus) pertyi*, *C. (Metallophanaeus) horus*, and *C. (Megaphanaeus) bellicosus*. The results indicated the conservation of macro-chromosomal structures in the genus, the conservation of the number of 5S rDNA sites in contrast to heterochromatin heterogeneity, and variation in the chromosomal distribution of 18S rRNA gene clusters and sex chromosomes.

Materials and Methods

Samples from adult males of *Coprophanaeus (Coprophanaeus) acrisius* (MacLeay, 1819), *C. (Coprophanaeus) dardanus* (MacLeay, 1819), *C. (Metallophanaeus) pertyi* (Olsoufieff, 1924), *C. (Metallophanaeus) horus* (Waterhouse, 1891), and *C. (Megaphanaeus) bellicosus* (Olivier, 1789) were collected from different regions in Pernambuco and Minas Gerais States, Brazil (table 1). The animals were collected in the wild according to Brazilian laws for environmental protection (wild collection permit, MMA/IBAMA/SISBIO no. 2376-1). The experimental research on animals was conducted according to the international guidelines followed by São Paulo State University (Protocol no. 35/08 – CEEA/IBB/UNESP). The testes were fixed in Carnoy solution (3:1 ethanol:acetic acid) and stored in the freezer at -20°C.

Chromosome preparations were obtained by the classical testicular follicle squashing technique. The C-banding technique was performed according to Sumner [1972] with modifications to analyze the male meiotic spreads and heterochromatin regions. Fluorochrome staining using the combination Chromomycin A₃/Distamycin/4'6-diamino-2'-phenylindol (CMA₃/DA/DAPI) method was performed according to Schweizer [1976] with modifications to quantify the heterochromatin in relation to the AT/GC base pair content.

DNA probes for the 18S and 5S rRNA genes were obtained from the dung beetle *Dichotomius semisquamosus* [Cabral-de-Mello et al., 2010b]. The 18S rRNA gene probe was labeled by nick translation using biotin-11-dATP (Invitrogen, San Diego, Calif., USA), and the 5S rRNA gene was labeled with digoxigenin-11-dUTP (Roche, Mannheim, Germany) via PCR. The FISH procedures were performed according to the protocol adapted by Cabral-de-Mello et al. [2010b] for Coleoptera. The chromosome preparations were counterstained with DAPI, and the slides were

Table 1. Diploid numbers, heterochromatin patterns, and chromosomal location of rRNA gene clusters in *Coprophanaeus* species

Species	Chromosomal formula (males)	Heterochromatin distribution	CMA ₃ /DA/DAPI		45S rDNA		5S rDNA		Collection sites in Brazil	References
			auto-some	sex chromosome	auto-some	sex chromosome	auto-some	sex chromosome		
<i>Coprophanaeus</i> (<i>Coprophanaeus</i>) <i>acrisius</i>	20 = 9 + XY	diphasic	4: CMA ₃ ⁺	–	5	–	2	–	Refúgio Ecológico Charles Darwin, Igarassu-PE	this work
<i>C. (Coprophanaeus)</i> <i>cyaneus</i>	20 = 9 + XY _p	diphasic	2: CMA ₃ ⁺ ^a	–	5/4 ^b	–/X ^b	2	–	Refúgio Ecológico Charles Darwin, Igarassu-PE; Andaraí, Chapada Diamantina-BA	Oliveira et al. [2010]; Cabral-de-Mello et al. [2011a]
<i>C. (Coprophanaeus)</i> <i>dardanus</i>	20 = 9 + Xy	diphasic	10: DAPI ⁺ 6: CMA ₃ ⁺	–	4	–	2	–	Refúgio Ecológico Charles Darwin, Igarassu-PE	Cabral-de-Mello et al. [2008]; this work
<i>C. (Metallophanaeus)</i> <i>horus</i>	20 = 9 + XY _p	diphasic	2: CMA ₃ ⁺	–	2	X	2	–	Carrancas-MG	this work
<i>C. (Metallophanaeus)</i> <i>pertyi</i>	20 = 9 + XY _p	diphasic	4: CMA ₃ ⁺	–	2	–	2	–	Brejo Novo, Caruaru-PE	this work
<i>C. (Megaphanaeus)</i> <i>bellicosus</i>	20 = 9 + XY	diphasic	6: CMA ₃ ⁺	–	14	–	2	–	Parque Ecológico João Vasconcelos Sobrinho, Caruaru-PE	this work
<i>C. (Megaphanaeus)</i> <i>ensifer</i>	20 = 9 + XY	diphasic	18: CMA ₃ ⁺	Y: CMA ₃ ⁺	14/10 ^b	X/– ^b	–	X, Y	Aldeia, Paudalho-PE; Andaraí, Chapada Diamantina-BA	Martins [1994]; Oliveira et al. [2010]; Cabral-de-Mello et al. [2011a]

PE = Pernambuco State; BA = Bahia State; MG = Minas Gerais State.

^a Patterns previously misidentified for the sex chromosomes. ^b Presence of a polymorphism in the same species.

mounted in Vectashield mounting medium (Vector, Burlingame, Calif., USA). Images were captured using an Olympus DP71 digital camera coupled to a BX61 Olympus microscope and were optimized for brightness and contrast using Adobe Photoshop CS2.

Results

Karyotypes and Heterochromatin Characterization

Coprophanaeus species showed similar karyotypes consisting of 2n = 20 and meta-submetacentric chromosomes with a gradual reduction in size. The species analyzed, however, differed in sex-chromosome mechanisms: XY in *C. (Coprophanaeus) acrisius* and *C. (Megaphanaeus) bellicosus*, with the X and Y chromosomes similar in size; Xy in *C. (Coprophanaeus) dardanus*, with the y being smaller than the X; and XY_p in *C. (Metallophanaeus) horus* and *C. (Metallophanaeus) pertyi* (fig. 1; table 1) where the 'p' refers to a 'parachute' configuration.

The C-banding technique primarily showed the presence of diphasic chromosomes (with the long arm hetero-

chromatic) in the karyotypes of the 5 analyzed species (fig. 1; table 1). In *C. (Megaphanaeus) bellicosus*, all the autosomal chromosomes were diphasic (fig. 1e), while in *C. (Coprophanaeus) acrisius*, *C. (Metallophanaeus) pertyi*, and *C. (Metallophanaeus) horus* one autosomal pair had a larger pericentromeric block, and the remaining autosomes had a diphasic pattern (fig. 1a, c, d). In *C. (Coprophanaeus) dardanus*, 2 autosomal pairs had large heterochromatic blocks in the pericentromeric regions and 7 pairs were diphasic (fig. 1b). With respect to the sex chromosomes, the X was almost completely heterochromatic in all species analyzed (fig. 1), but the Y chromosome was almost completely heterochromatic in *C. (Coprophanaeus) acrisius* and *C. (Megaphanaeus) bellicosus* (fig. 1a, e) and had a large pericentromeric block in *C. (Metallophanaeus) pertyi* (fig. 1c).

In the 5 analyzed species, CMA₃/DA/DAPI staining revealed positive signals only in the autosomes (fig. 2; table 1). Chromomycin A₃-positive (CMA₃⁺) signals were observed in 2 chromosomal pairs of *C. (Coprophanaeus) acrisius* and *C. (Metallophanaeus) pertyi*, with one block

Fig. 1. C-banding of metaphase I chromosomes of *Coprophanaeus* species. **a** *C. (Coprophanaeus) acrisius*. **b** *C. (Coprophanaeus) dardanus*. **c** *C. (Metallophanaeus) pertyi*. **d** *C. (Metallophanaeus) horus*. **e** *C. (Megaphanaeus) bellicosus*. The arrows indicate the sex chromosome bivalents, the asterisks indicate the autosomal pairs with pericentromeric blocks of heterochromatin, and the arrowhead in **(b)** indicates the autosomal pair with a heterochromatic block in the short arm. In **(d)** the autosomal pair with pericentromeric blocks of heterochromatin is shown in detail in 1 (mitotic pair) and 2 (meiotic pair). Bars = 5 μ m.

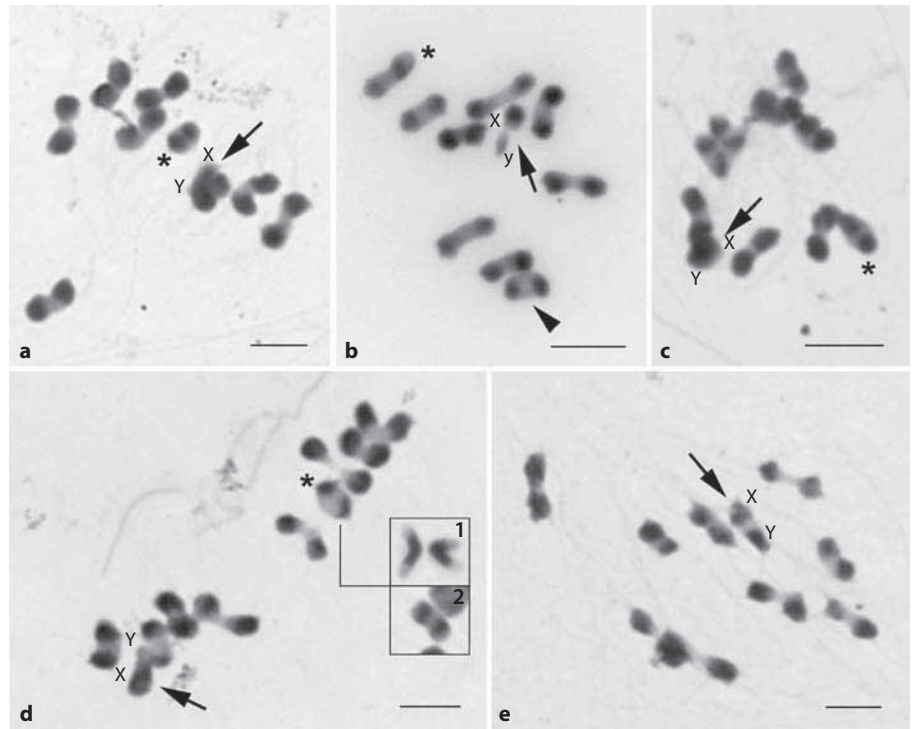
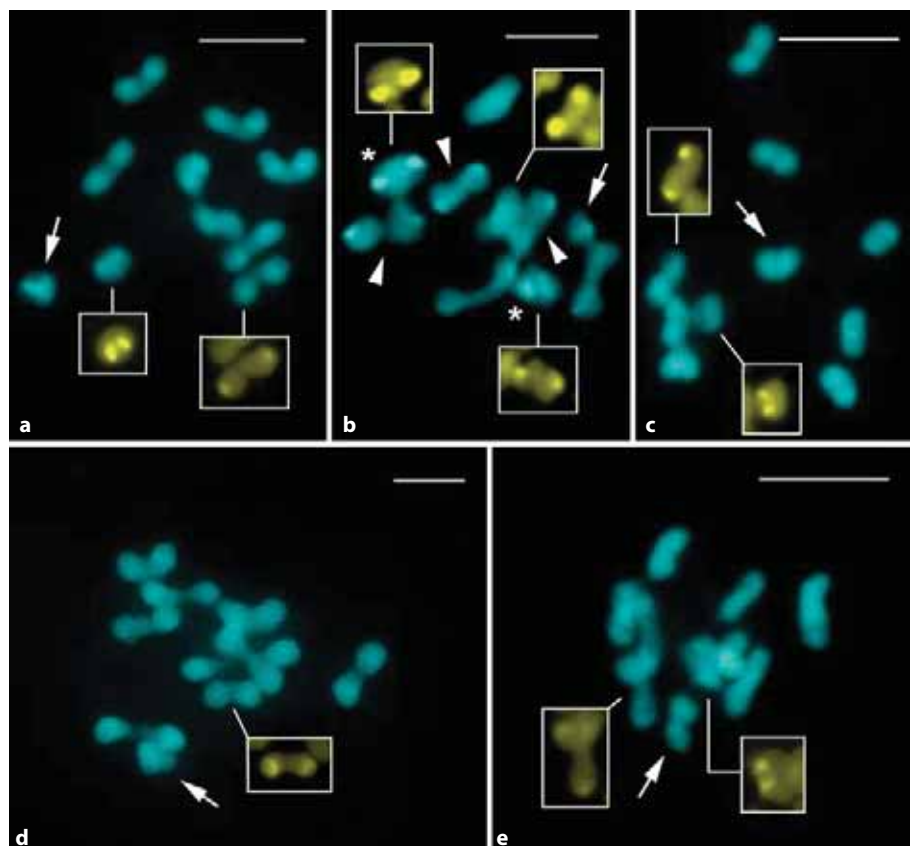


Fig. 2. Fluorochrome staining (CMA₃/DA/DAPI) of metaphases I of *Coprophanaeus* species. **a** *C. (Coprophanaeus) acrisius*. **b** *C. (Coprophanaeus) dardanus*. **c** *C. (Metallophanaeus) pertyi*. **d** *C. (Metallophanaeus) horus*. **e** *C. (Megaphanaeus) bellicosus*. The arrows indicate the sex chromosome bivalents. In **(b)** the DAPI⁺ blocks adjacent to the CMA₃⁺ blocks are indicated with asterisks (pericentromeric blocks) and arrowheads (terminal blocks). Bars = 5 μ m.



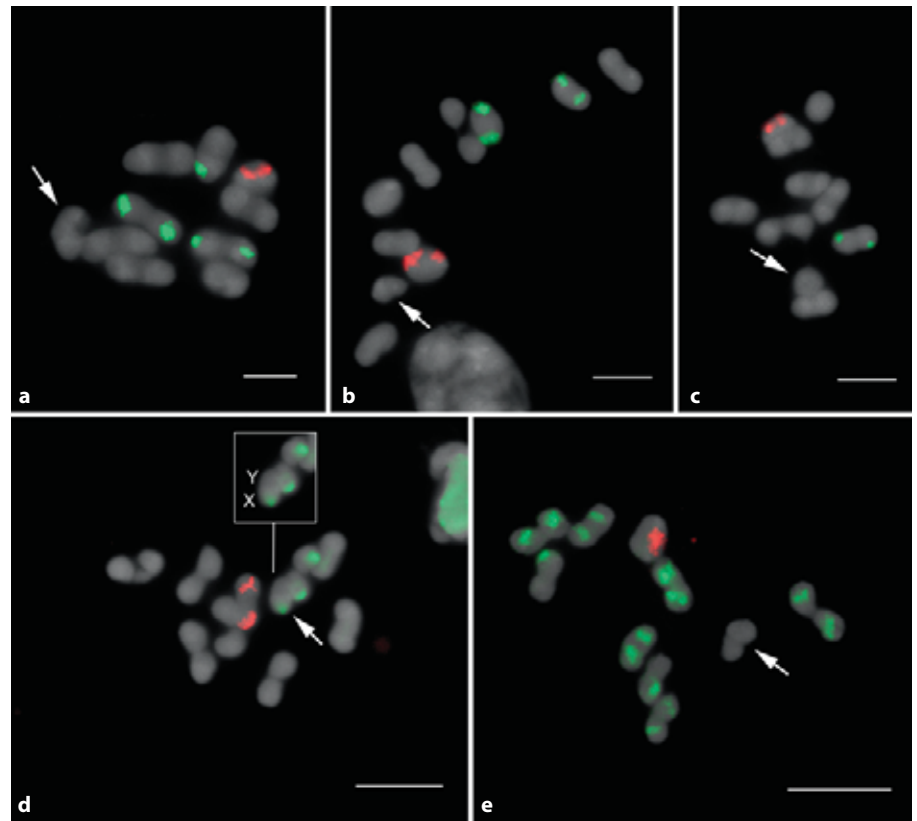


Fig. 3. FISH with 18S (green) and 5S rRNA (red) genes probed in metaphases I of *Coprophanaeus* species. **a** *C. (Coprophanaeus) acrisius*. **b** *C. (Coprophanaeus) dardanus*. **c** *C. (Metallophanaeus) pertyi*. **d** *C. (Metallophanaeus) horus*. **e** *C. (Megaphanaeus) bellicosus*. The arrows indicate the sex chromosome bivalents. The X and Y sex chromosomes of the metaphase plate in (**d**) are indicated in the box. Bars = 5 μ m.

in the terminal region and the other in the pericentromeric region of another pair (fig. 2a, c). Which chromosomes carried the blocks could not be determined due to the high chromosome condensation level. In *C. (Megaphanaeus) bellicosus*, CMA₃⁺ blocks were observed in the pericentromeric region of one pair and in the terminal region of 2 other pairs (fig. 2e). CMA₃⁺ signals were observed only in the pericentromeric region of one pair in *C. (Metallophanaeus) horus* (fig. 2d). In these 4 species, DAPI-positive (DAPI⁺) signals were not visualized (fig. 2a, c–e). In *C. (Coprophanaeus) dardanus*, the CMA₃/DA/DAPI staining identified DAPI⁺ blocks in 5 autosomal pairs, 3 with the blocks in the terminal region and 2 with the blocks in the pericentromeric region. Two autosomal pairs that have pericentromeric DAPI⁺ blocks also showed adjacent CMA₃⁺ sites. CMA₃⁺ blocks were also observed in a third autosomal bivalent region, but this region was DAPI-negative (DAPI[−]) (fig. 2b).

Cytogenetic Mapping of 5S and 18S rDNA

Chromosomal mapping of 5S rDNA revealed signals in the pericentromeric region of one autosomal bivalent with 2 chiasmata in all 5 species analyzed (fig. 3; table 1). In

contrast, the mapping of 18S rDNA revealed variability for this marker among the species studied. These gene clusters were located in the pericentromeric regions of 2 autosomal pairs in *C. (Coprophanaeus) dardanus* (fig. 3b) and in the terminal regions of 3 autosomal pairs in *C. (Coprophanaeus) acrisius* (only one homolog member was labeled for 1 of the 3 chromosomal pairs) (fig. 3a). The 18S rDNA sites were observed in the terminal region of one autosomal pair in *C. (Metallophanaeus) horus* and *C. (Metallophanaeus) pertyi* and in the pericentromeric region of the X chromosome in *C. (Metallophanaeus) horus* (fig. 3c, d). In *C. (Megaphanaeus) bellicosus*, the 18S rDNA sites were identified in the pericentromeric region of 8 autosomal pairs, although 2 pairs were heteromorphic, possessing only one site in one of the homologs (fig. 3e).

Discussion

Heterochromatin Features

The diploid number, $2n = 20$, and the meta-submetacentric morphology observed in *Coprophanaeus* species are in agreement with the karyotype considered modal

for Scarabaeidae, for the suborder Polyphaga, and for Coleoptera [Smith and Virkki, 1978; Wilson and Angus, 2005; Cabral-de-Mello et al., 2008]. The presence of $2n = 20$ is also shared among other distinct Phanaeini species, such as *Bolbites ornitoides*, *Diabroctis mimas*, *Gromphas lacordairei*, and *Sulcophanaeus imperator* [Vidal, 1984; Bione et al., 2005a; Cabral-de-Mello et al., 2008], and in species belonging to closely related groups such as Onitini, Eucraniini, and Dichotomini [Smith and Virkki, 1978; Vidal, 1984; Colomba et al., 1996; Angus et al., 2007; Cabral-de-Mello et al., 2008]. These data indicate that $2n = 20$ could be the ancient condition for this group. In contrast, variability for this characteristic was observed in some species of the Phanaeini tribe, such as *Oxysternon silenensis* and *Phanaeus daphnis* [Smith and Virkki, 1978; Cabral-de-Mello et al., 2008].

The high amount of heterochromatin observed in *Coprophanaeus* species differs from the most common pattern observed in the family Scarabaeidae which generally presents heterochromatic blocks in the pericentromeric region of the autosomes, although variations in the location of the heterochromatin along the sex chromosomes were observed [Moura et al., 2003; Wilson and Angus, 2004, 2005; Bione et al., 2005b; Wilson and Angus, 2006; Angus et al., 2007; Silva et al., 2009]. For the tribe Phanaeini, diphasic autosomes were also observed in *C. (Coprophanaeus) cyaneus*, *C. (Megaphanaeus) ensifer*, and *Diabroctis mimas* [Bione et al., 2005a; Oliveira et al., 2010], and a large amount of heterochromatin in the genome appears to be common in this group. These large heterochromatic blocks indicate that the chromosomes of the species suffered an expansion of repetitive DNA, involving amplification by different spreading processes as described for vertebrates and other insects [Charlesworth et al., 1994; Hancock, 1999; Landais et al., 2000; Li et al., 2002]. All the Phanaeini species analyzed showed a high amount of heterochromatin [Bione et al., 2005a; Oliveira et al., 2010; Cabral-de-Mello et al., 2011a; present work] which could indicate that the amplification of heterochromatin occurred early during group diversification.

The heterochromatic regions have shown a heterogeneous pattern for base pair richness in representatives of the Scarabaeidae family [Colomba et al., 1996; Vitturi et al., 1999; Colomba et al., 2000; Moura et al., 2003; Vitturi et al., 2003; Bione et al., 2005b; Colomba et al., 2006; Cabral-de-Mello et al., 2010b, 2010c; Oliveira et al., 2010; present work]. The presence of CMA₃⁺ blocks is predominant in the genomes of Phanaeini species. DAPI⁺ blocks, however, were observed in *C. (Coprophanaeus) dardanus*

and *Bolbites onitoides* [Virkki, 1983]. Different mechanisms can be involved in the heterogeneous pattern of the base pair richness observed in *Coprophanaeus* species. Bione et al. [2005a] suggested the occurrence of small duplications in tandem, resulting in the additional heterochromatic blocks observed in *Diabroctis mimas*. Indeed, different families of satellite DNA have been previously observed in the genome of some Coleoptera representatives [Juan et al., 1993; Ugarkovic et al., 1995; Zinic et al., 2000]. Lohe and Roberts [2000] proposed that the heterochromatin can undergo different rearrangements, allowing the expansion or reduction of satellite repeats during evolution and may allow the emergence of new satellites. The spread of new satellite DNA can occur through a variety of mechanisms, including gene conversion, unequal crossing-over, slippage replication, transposition, and RNA-mediated exchanges [Dover, 2002; Palomeque and Lorite, 2008]. These events may have occurred during the chromosomal evolution of *Coprophanaeus* species which explains the increase in the quantity and variability of heterochromatin, including the variability observed in the sex chromosomes.

Sex Chromosome Systems

Distinct chromosomal rearrangements involving inversions, autosome fusions and fissions, X-autosome fusions, and y chromosome loss appear to be the most probable events involved in the generation of diploid number and sex chromosome variability observed in Scarabaeinae and in the Phanaeini tribe. The mechanisms of chromosomal evolution differ between different tribes and between different genera in the group. Although high chromosomal variability was observed in Phanaeini, *Coprophanaeus* showed conservation of the diploid number and chromosome morphology as well as of sex chromosome mechanisms. Although the *Coprophanaeus* species analyzed have X and y chromosomes, these chromosomes showed different forms of association. During meiosis in the Xy and XY systems, the sex chromosomes associate in the classical configuration pattern of sex chromosomes [White, 1973; Smith and Virkki, 1978]. In the Xy_p mechanism, however, the X chromosome represents the dorsal of the parachute, and the y chromosome represents the parachutist, connected by 2 thin wires. Usually, the X is a medium metacentric chromosome and the y is an extremely small metacentric chromosome. In cases where the X and Y chromosomes are relatively large and of similar size, the sex mechanism is called XY_p [White, 1973; Mesa and Fontanetti, 1985].

The XY_p mechanism observed in *C. (Metallophanaeus) horus*, *C. (Metallophanaeus) pertyi*, and the previously analyzed *C. (Coprophanaeus) cyanescens* [Oliveira et al., 2010] was also observed in other Scarabaeidae species, such as *Deltochilum (Calhyboma) verruciferum* and *Malagoniella aff astianax*, and *Phyllophaga (Phyllophaga) aff capilata* [Moura et al., 2003; Cabral-de-Mello et al., 2008, 2010c; Oliveira et al., 2010]. The XY_p mechanism is not a common mechanism for the group; it is considered to be derived and is characterized by sex chromosomes with similar sizes, differing from the Xy_p mechanism often observed in Coleoptera. This mechanism is thus considered primitive. From the observation of the XY_p mechanism in *Itu zeus* (Myxophaga), Mesa and Fontanetti [1985] proposed 2 hypotheses to explain the origin of this mechanism: (i) from an achiasmatic XY system, with X and Y chromosomes of similar sizes and (ii) by the addition of heterochromatin in an Xy_p bivalent system. Recently, Cabral-de-Mello et al. [2010c] proposed a third hypothesis based on the analysis of the XY_p mechanism in *Deltochilum (Calhyboma) verruciferum* which allowed the observation of argyrophilic proteins in the lumen of the sex bivalent region and the presence of a large, diphasic Y chromosome, reinforcing the idea that this mechanism could have been derived from a primitive Xy_p, at least in *D. (Calhyboma) verruciferum*.

In the Phanaeini tribe, the chiasmatic Xy mechanism without the formation of the parachute observed in *C. (Coprophanaeus) dardanus* was described only in *Bolbites onitoides* [Vidal, 1984; Bione et al., 2005a; Cabral-de-Mello et al., 2008]. The XY sex mechanism observed in *C. (Coprophanaeus) acrisius* and *C. (Megaphanaeus) bellicosus*, however, was also observed in 3 other Phanaeini species: *C. (Megaphanaeus) ensifer*, *Phanaeus chalcomelas*, and *P. igneus* [Smith and Virkki, 1978; Martins, 1994; Cabral-de-Mello et al., 2008]. The variability of the sex chromosomes in the *Coprophanaeus* species is enhanced by the occurrence of different chiasmatic (XY and Xy) and achiasmatic (XY_p and Xy_p) sex mechanisms. The variations observed relative to the amount and distribution of heterochromatin suggest the involvement of heterochromatin in the origin of the sex chromosome diversity in this group which may have involved the amplification and gain of heterochromatin.

Chromosomal Organization of 5S and 18S rRNA Gene Clusters

The presence of only one chromosome bivalent carrying the 18S rDNA clusters is the ancient condition in Scarabaeinae [Cabral-de-Mello et al., 2011a] and in the

order Coleoptera, at least for Polyphaga representatives [Schneider et al., 2007; Cabral-de-Mello et al., 2011a]. The presence of more than one chromosome bivalent carrying clusters of 18S rDNA was previously observed in *C. (Coprophanaeus) cyanescens* and *C. (Megaphanaeus) ensifer*, with 5 and 15 sites (the largest number of 18S rDNA sites of the order Coleoptera), respectively [Oliveira et al., 2010; Cabral-de-Mello et al., 2011a]. Among the *Coprophanaeus* species studied in this work, only *C. (Metallophanaeus) pertyi* showed one autosomal pair with 18S rDNA sites, while *C. (Coprophanaeus) bellicosus* had 14 sites of 18S rDNA. *Coprophanaeus (Coprophanaeus) cyanescens*, *C. (Metallophanaeus) horus*, and *C. (Megaphanaeus) ensifer* show intraspecific polymorphisms with respect to both the number and/or position of 18S rDNA sites [Oliveira et al., 2010; Cabral-de-Mello et al., 2011a; present work]. Intraspecific polymorphisms, the repositioning of the rDNA sites, increasing numbers of rDNA sites, and the movement of rDNA sites to different autosomes and sex chromosomes were observed [Oliveira et al., 2010; Cabral-de-Mello et al., 2011a].

Correlating the amount of heterochromatin with the number of 18S rDNA sites was not possible because the heterochromatin amount and distribution is very similar in all *Coprophanaeus* analyzed. Considering the subgenera distribution, *Metallophanaeus* appears to present the ancestral condition, showing conservation in the number of sites, while *Megaphanaeus* appears to have undergone amplification of 18S rDNA sites, with *Coprophanaeus* showing an intermediate position. Despite these observations, addressing a correlation between the phylogenetic position and the variation of 18S rDNA sites was not possible.

The data obtained by FISH with the 18S rDNA probes in *Coprophanaeus* species support the hypothesis proposed by Sánchez-Gea et al. [2000] for inter- and intra-specific polymorphisms observed in the genus *Zabrus* (Carabidae) and by Oliveira et al. [2010] for the large number of rDNA sites in *C. (Megaphanaeus) ensifer*. This hypothesis postulates that the association of rDNA with heterochromatic regions, including the presence of transposable elements, could affect the occurrence of these polymorphisms.

In contrast to the variability in the number of 18S rDNA sites, the 5 *Coprophanaeus* species analyzed in this work, as well as the previously analyzed *C. (Coprophanaeus) cyanescens* and *C. (Megaphanaeus) ensifer* [Cabral-de-Mello et al., 2011a], showed only one chromosomal pair carrying 5S rRNA genes. Although defining which autosomal pair carries the 5S rDNA clusters is not possible, it

appears to be the same pair in all species due to its meiotic behavior. In *C. (Megaphanaeus) ensifer*, the 5S rDNA clusters are located on the sex chromosomes [Cabral-de-Mello et al., 2011a], possibly as a derived condition that resulted from chromosomal rearrangements. In addition to the species analyzed here, 14 other previously analyzed Scarabaeinae species showed only one pair of autosomal chromosomes carrying the 5S rDNA sites [Cabral-de-Mello et al., 2011a]. Some species have shown more than one 5S rDNA site, although the presence of 5S rDNA at only one site with a distinct location from the 18S rDNA sequences is commonly observed in diverse eukaryote groups [Martins and Galetti, 2001; Vitturi et al., 2002; Ribeiro and Fernandez, 2004; Loreto et al., 2008; Puerma et al., 2008; Cabral-de-Mello et al., 2011a, b]. The physical separation observed between the 5S and 18S rDNA sites could be the result of a functional advantage for these ribosomal sequences, considering that the 18S rRNA gene is transcribed by RNA polymerase I and that the 5S rRNA gene is transcribed by RNA polymerase III [Roussel and Hernandez-Verdun, 1994; Sumner, 2003; Raska et al., 2004].

Despite the diversification of the species, the presence of one 5S rDNA site in different species of *Coprophanaeus* and in other Coleoptera may indicate that this represents the ancestral condition that has been preserved from major changes, possibly due to a purifying selection that prevented the spread of these sequences to other regions of the genome [Rooney and Ward, 2005; Fujiwara et al., 2009]. In contrast, the clusters of 45S rDNA evolution appear to be under the action of ectopic recombination generated from different mechanisms such as unequal crossing-over and gene conversion and transposable elements [Dover, 1986; Montgomery et al., 1991;

Charlesworth et al., 1994; Petrov et al., 2003], allowing the spreading of this repetitive family to new locations as here observed for *Coprophanaeus*.

Conclusions

The *Coprophanaeus* species showed karyotypes considered to be the possible ancient condition for Coleoptera. The high amount of heterochromatin appears to be common in *Coprophanaeus* species and indicates that the chromosomes of the species underwent an expansion of repetitive DNA early during group diversification. The different sex mechanisms observed suggest that the amplification and distribution of heterochromatin were involved in the origin of these mechanisms. The heterogeneous pattern for base pair richness also indicates the heterogeneity of the heterochromatin. The variation observed for the 18S rDNA in contrast to that of conserved 5S rDNA patterns suggests that distinct evolutionary forces govern the genomic regions that harbor both rRNA gene classes.

Acknowledgements

The authors are grateful to J. Louzada, C.M.Q. Costa, F. França, and F.A.B. Silva for the viability of the collection in Minas Gerais State and the taxonomic identification of the species studied. This study was supported by the Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE – BPD-003.2.02/08 e APQ-0464-2.02/10), the Programa de Fortalecimento Acadêmico da UPE (PFAUPE), and the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP).

References

- Angus RB, Wilson CJ, Mann DJ: A chromosomal analysis of 15 species of Gymnopleurini and Coprini (Coleoptera: Scarabaeidae). *Tijdschr Entomol* 150:201–211 (2007).
- Biéumont C, Vieira C: Genetics: junk DNA as an evolutionary force. *Nature* 443:521–524 (2006).
- Bione EG, Camparoto ML, Simões ZL: A study of the constitutive heterochromatin and nucleolus organizer regions of *Isocoprini inhiata* and *Diabrotica mima* (Coleoptera: Scarabaeidae, Scarabaeinae) using C-banding, AgNO₃ staining and FISH techniques. *Gen Mol Biol* 28:111–116 (2005a).
- Bione EG, Moura RC, Carvalho R, Souza MJ: Karyotype, C-banding pattern, NOR location and FISH study of five Scarabaeidae (Coleoptera) species. *Gen Mol Biol* 28:376–381 (2005b).
- Cabral-de-Mello DC, Oliveira SG, Ramos IC, Moura RC: Karyotype differentiation patterns in species of the subfamily Scarabaeinae (Scarabaeidae, Coleoptera). *Micron* 39:1243–1250 (2008).
- Cabral-de-Mello DC, Moura RC, Martins C: Chromosomal mapping of repetitive DNAs in the beetle *Dichotomius geminatus* provides the first evidence for an association of 5S rRNA and histone H3 genes in insects, and repetitive DNA similarity between the B chromosome and A complement. *Heredity* 104:393–400 (2010a).
- Cabral-de-Mello DC, Moura RC, Souza MJ: Amplification of repetitive DNA and origin of a rare chromosomal sex bivalent in *Deltochilum (Calhyboma) verruciferum* (Coleoptera, Scarabaeidae). *Genetica* 138:191–195 (2010b).

- Cabral-de-Mello, Moura DC, Carvalho R, Souza MJ: Cytogenetic analysis of two related *Deltochilum* (Coleoptera, Scarabaeidae) species: diploid number reduction, extensive heterochromatin addition and differentiation. *Micron* 41:112–117 (2010c).
- Cabral-de-Mello DC, Oliveira SG, Moura RC, Martins C: Chromosomal organization of 18S and 5S rRNA and H3 histone genes in Scarabaeinae coleopterans: insights on the evolutionary dynamics of multigene families and heterochromatin. *BMC Genet* 12:88 (2011a).
- Cabral-de-Mello DC, Moura RC, Martins C: Cytogenetic mapping of rRNAs and histone H3 genes in 14 species of *Dichotomius* (Coleoptera, Scarabaeidae, Scarabaeinae) beetles. *Cytogenet Genome Res* 134:127–135 (2011b).
- Charlesworth B, Snlegowski P, Stephan W: The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215–220 (1994).
- Colomba MS, Monteresino E, Vitturi R, Zunino M: Characterization of mitotic chromosomes of the scarab beetles *Glyphoderus sterquilinus* (Westwood) and *Bubas bison* (L.) (Coleoptera: Scarabaeidae) using conventional and banding techniques. *Biol Zentralbl* 115:58–70 (1996).
- Colomba MS, Vitturi R, Zunino M: Karyotype analysis, banding, and fluorescent in situ hybridization in the scarab beetle *Gymnopleurus sturmi* McLeay (Coleoptera Scarabaeoidea: Scarabaeidae). *J Hered* 91:260–264 (2000).
- Colomba MS, Vitturi R, Libertini A, Gregorini A, Zunino M: Heterochromatin of the scarab beetle, *Bubas bison* (Coleoptera: Scarabaeidae) II. Evidence for AT-rich compartmentalization and a high amount of rDNA copies. *Micron* 37:47–51 (2006).
- Dover G: Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends Genet* 2:159–165 (1986).
- Dover G: Molecular drive. *Trends Genet* 18:587–589 (2002).
- Edmonds WD: The immature stages of *Phanaeus* (*Coprophanaeus*) *jasius* Oliver and *Phanaeus* (*Metallophanaeus*) *saphirinus* Sturm (Coleoptera: Scarabaeidae). *Coleopt Bull* 21:97–105 (1967).
- Edmonds WD: Comparative skeletal morphology, systematics and evolution of the Phanaeinae dung beetles (Coleoptera: Scarabaeidae). *Kans Univ Sci Bull* 49:731–874 (1972).
- Edmonds WD, Zidek JA: Taxonomic review of the neotropical genus *Coprophanaeus* Olsoufieff, 1924 (Coleoptera: Scarabaeidae, Scarabaeinae). *Insecta Mundi* 129:1–111 (2010).
- Ferreira A, Cella D, Tardivo JR, Virkki N: Two pairs of chromosomes: a new low record for Coleoptera. *Braz J Genet* 2:231–239 (1984).
- Fujiwara M, Inafuku J, Takeda A, Watanabe A, Fujiwara A, et al: Molecular organization of 5S rDNA in bitterlings (Cyprinidae). *Genetica* 135:355–365 (2009).
- Galián J, Hogan JE, Vogler AP: The origin of multiple sex chromosomes in tiger beetles. *Mol Biol Evol* 19:1792–1796 (2002).
- Hancock JM: Microsatellites and other simple sequences: genomic context and mutational mechanisms, in Goldstein DB, Schlötterer C (eds): *Microsatellites*, pp 1–8 (Oxford University Press, Oxford 1999).
- Hanski I, Cambefort Y: *Dung Beetle Ecology*, pp. 481 (Princeton University Press, Princeton 1991).
- Juan C, Pons J, Petitpierre E: Localization of tandemly repeated DNA sequences in beetle chromosomes by fluorescent in situ hybridization. *Chromosome Res* 1:167–174 (1993).
- Landais I, Chavigny P, Castagnone C, Pizzol J, Abad P, Vanlerberghe-Masutti F: Characterization of a highly conserved satellite DNA from the parasitoid wasp *Trichogramma brassicae*. *Gene* 255:65–73 (2000).
- Li Y, Korol AB, Fahima T, Beiles A, Nevo E: Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Mol Ecol* 11:2453–2465 (2002).
- Lohe AR, Roberts PA: Evolution of DNA in heterochromatin: the *Drosophila melanogaster* sibling species subgroup as a resource. *Genetica* 109:125–130 (2000).
- Loreto V, Cabrero J, López-León MD, Camacho JPM, Souza MJ: Possible autosomal origin of macro B chromosomes in two grasshopper species. *Chromosome Res* 16:233–241 (2008).
- Louzada JN: Scarabaeinae (Coleoptera: Scarabaeidae) detritívoros em ecossistemas tropicais: biodiversidade e serviços ambientais, in Moreira FMS, Siqueira JO, Brussaard L (eds): *Biodiversidade do Solo em Ecossistemas Brasileiros*, pp 309–332 (UFPA, Lavras 2008).
- Martins C, Galletti Jr PM: Two rDNA arrays in Neotropical fish species: is it a general rule for fishes? *Genetica* 111:439–446 (2001).
- Martins VG: The chromosome of five species of Scarabaeidae (Polyphaga, Coleoptera). *Naturalia* 19:89–96 (1994).
- Mesa A, Fontanetti C: The chromosomes of a primitive species of beetle: *Ytu zeus* (Coleoptera, Myxophaga, Torridincolidae). *Acad Nat Phil* 137:102–105 (1985).
- Montgomery EA, Huang SM, Langley CH, Judd BH: Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. *Genetics* 129:1085–1098 (1991).
- Moura RC, Souza MJ, Melo NF, Lira-Neto AC: Karyotypic characterization of representatives from Melolonthinae (Coleoptera: Scarabaeidae): karyotypic analysis, banding and fluorescent in situ hybridization (FISH). *Hereditas* 138:200–206 (2003).
- Nowak R: Mining treasures from junk DNA. *Science* 263:608–610 (1994).
- Oliveira SG, Moura RC, Silva AEB, Souza MJ: Cytogenetic analysis of two *Coprophanaeus* species (Scarabaeidae) revealing wide constitutive heterochromatin variability and the largest number of 45S rDNA sites among Coleoptera. *Micron* 41:960–965 (2010).
- Palomeque T, Lorite P: Satellite DNA in insects: a review. *Heredity* 100:564–573 (2008).
- Petrov DA, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE: Size matters: non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. *Mol Biol Evol* 20:880–892 (2003).
- Philips TK, Edmonds WD, Scholtz CH: A phylogenetic analysis of the New World tribe Phanaeini (Coleoptera: Scarabaeidae: Scarabaeinae): hypotheses on relationships and origins. *Insect Syst Evol* 35:43–63 (2004).
- Puerma E, Acosta MJ, Barragán MJ, Martínez S, Marchal JÁ, et al: The karyotype and 5S rRNA genes from Spanish individuals of the bat species *Rhinolophus hipposideros* (Rhinolophidae; Chiroptera). *Genetica* 134:287–295 (2008).
- Raska I, Koberna K, Malinsky J, Fidlerová H, Masata M: The nucleolus and transcription of ribosomal genes. *Biol Cell* 96:579–594 (2004).
- Ribeiro LF, Fernandez MA: Molecular characterization of the 5S ribosomal gene of the *Bradysia hygida* (Diptera: Sciaridae). *Genetica* 122:253–260 (2004).
- Rooney AP, Ward TJ: Evolution of a large ribosomal RNA multigene family in filamentous fungi: birth and death of a concerted evolution paradigm. *Proc Natl Acad Sci* 102:5084–5089 (2005).
- Roussel P, Hernandez-Verdun D: Identification of Ag-NOR proteins, markers of proliferation related to ribosomal genes activity. *Exp Cell Res* 214:465–472 (1994).
- Rozék M, Lachowska D, Petitpierre E, Holecová M: C-bands on chromosomes of 32 beetle species (Coleoptera: Elateridae, Cantharidae, Oedemeridae, Cerambycidae, Anthicidae, Chrysomelidae, Attelabidae and Curculionidae). *Hereditas* 140:161–170 (2004).
- Sánchez-Gea JF, Serrano J, Galian J: Variability in rDNA loci in Iberian species of the genus *Zabrus* (Coleoptera: Carabidae) detected by fluorescence in situ hybridization. *Genome* 43:22–28 (2000).
- Schneider MC, Rosa SP, Almeida MC, Costa C, Cella DM: Chromosomal similarities and differences among four Neotropical Elateridae (Conoderini and Pyrophorini) and other related species, with comments on the NOR patterns in Coleoptera. *J Zool Syst Evol Res* 45:308–316 (2007).
- Schweizer D: Reverse fluorescent chromosome banding with chromomycin and DAPI. *Chromosoma* 58:307–324 (1976).
- Silva GM, Bione EG, Cabral-de-Mello DC, Moura RC, Simões ZL, Souza MJ: Comparative cytogenetics of three species of *Dichotomius* (Coleoptera, Scarabaeidae). *Genet Mol Biol* 32:276–280 (2009).

- Smith SG, Virkki N: Coleoptera, in John B (ed): Animal Cytogenetics, pp 366 (Borntraeger, Berlin 1978).
- Sumner AT: A simple technique for demonstrating centromeric heterochromatin. *Exp Cell Res* 75:304–306 (1972).
- Sumner AT: Chromosomes: Organization and Function, pp 287 (Blackwell Publishing, Oxford 2003).
- Ugarkovic D, Petitpierre E, Juan C, Plohl M: Satellite DNAs in tenebrionid species: structure, organization and evolution. *Croat Chem Acta* 68:627–638 (1995).
- Vaz-de-Mello FZ: Estado atual de conhecimento dos Scarabaeidae s. str. (Coleoptera: Scarabaeoidea) do Brasil, in Martin-Piera F, Morrone JJ, Melic A (eds): Hacia un Proyecto CYTED para el inventario y estimación de la diversidad Entomológica en Iberoamérica: PrIBES-2000, pp. 183–195, 326 (Sociedad Entomológica Aragonesa & CYTED, m3m: Monografías Tercer Milenio, Zaragoza 2000).
- Vidal OR: Coleoptera from Argentina. *Genetica* 65:235–239 (1984).
- Virkki N: Banding of Oedionychina (Coleoptera, Alticinae) chromosomes. C- and Ag-bands. *J Agr Univ Puerto Rico* 67:221–255 (1983).
- Vitturi R, Colomba MS, Barbieri R, Zunino M: Ribosomal DNA location in the scarab beetle *Thorectes intermedius* (Costa) (Coleoptera: Geotrupidae) using banding and fluorescent in-situ hybridization. *Chromosome Res* 7: 255–260 (1999).
- Vitturi R, Colomba MS, Pirrone AM, Mandrioli M: rDNA (18S-28S and 5S) colocalization and linkage between ribosomal genes (TTAGGG)_n telomeric sequence in the earthworm, *Octodrilus complanatus* (Annelida: Oligochaeta: Lumbricidae), revealed by single- and double-color FISH. *J Hered* 93: 279–282 (2002).
- Vitturi R, Colomba MS, Volpe N, Lannino A, Zunino M: Evidence for male X0 sex-chromosome system in *Pentodon bidens punctatum* (Coleoptera: Scarabaeoidea: Scarabaeidae) with X-linked 18S–28S clusters. *Genes Genet Syst* 78:427–432 (2003).
- White MJ: Animal Cytology and Evolution, ed 3, (Cambridge University, London 1973).
- Wilson CJ, Angus RB: A chromosomal analysis of ten European species of *Aphodius* Illiger, subgenera *Acrossus* Mulsant, *Nimbus* Muldant & Rey and *Chilo thorax* Motschulsky (Coleoptera: Aphodiidae). *Koleopt Rdsch* 74: 367–374 (2004).
- Wilson CJ, Angus RB: A chromosomal analysis of 21 species of Oniticellini and Onthophagini (Coleoptera: Scarabaeidae). *Tijdschr Entomol* 148:63–76 (2005).
- Wilson CJ, Angus RB: A chromosomal analysis of eight species of *Aphodius* Illiger, subgenera *Agiolinus* Schmidt, *Agrilinus* Mulsant & Rey and *Planolinus* Mulsant & Rey (Coleoptera: Aphodiidae). *Proc Rus Entomol Soc St Petersburg* 77:28–33 (2006).
- Yadav JS, Pillai, RK, Karamje ET: Chromosome numbers of Scarabaeidae (Polyphaga: Coleoptera). *Coleopt Bull* 33:309–318 (1979).
- Yadav JS, Pillai RK, Yadav AS: Karyotypic study of some scarab beetles with comments on phylogeny (Coleoptera: Scarabaeoidea). *Elytron* 4:41–51 (1990).
- Zinic SJ, Ugarkovic D, Cornudella L, Plohl M: A novel interspersed type of organization of satellite DNAs in *Tribolium madens* heterochromatin. *Chromosome Res* 8:201–212 (2000).

RESEARCH ARTICLE

Open Access

B chromosome in the beetle *Coprophanaeus cyanescens* (Scarabaeidae): emphasis in the organization of repetitive DNA sequences

Sarah Gomes de Oliveira¹, Rita Cassia de Moura² and Cesar Martins^{1*}

Abstract

Background: To contribute to the knowledge of coleopteran cytogenetics, especially with respect to the genomic content of B chromosomes, we analyzed the composition and organization of repetitive DNA sequences in the *Coprophanaeus cyanescens* karyotype. We used conventional staining and the application of fluorescence *in situ* hybridization (FISH) mapping using as probes *C₀t-1* DNA fraction, the 18S and 5S rRNA genes, and the *LOA-like* non-LTR transposable element (TE).

Results: The conventional analysis detected 3 individuals (among 50 analyzed) carrying one small metacentric and mitotically unstable B chromosome. The FISH analysis revealed a pericentromeric block of *C₀t-1* DNA in the B chromosome but no 18S or 5S rDNA clusters in this extra element. Using the *LOA-like* TE probe, the FISH analysis revealed large pericentromeric blocks in eight autosomal bivalents and in the B chromosome, and a pericentromeric block extending to the short arm in one autosomal pair. No positive hybridization signal was observed for the *LOA-like* element in the sex chromosomes.

Conclusions: The results indicate that the origin of the B chromosome is associated with the autosomal elements, as demonstrated by the hybridization with *C₀t-1* DNA and the *LOA-like* TE. The present study is the first report on the cytogenetic mapping of a TE in coleopteran chromosomes. These TEs could have been involved in the origin and evolution of the B chromosome in *C. cyanescens*.

Keywords: Chromosomal rearrangements, Heterochromatin, Multigene families, Supernumerary chromosome, Transposable elements

Background

Eukaryote genomes are composed of classical genes and genetic elements, including transposable elements (TEs), B chromosomes and several cytoplasmic factors that do not follow Mendelian laws of inheritance [1]. B chromosomes (also called supernumerary or accessory chromosomes) are not essential for the life of a species and are thus considered “dispensable” additional chromosomes. B chromosomes have been observed in approximately 15% of living species [1-4]. Most B chromosomes are heterochromatic and composed of repetitive DNA sequences, supporting the idea that these chromosomes are non-coding. However, some B chromosomes show the presence of active genes [5-7]. B

chromosomes demonstrate an irregular behavior during mitosis and meiosis that allows them to accumulate in the germ line in a non-Mendelian pattern of inheritance [3,8]. Although B chromosomes have been the focus of intensive work in a diversity of eukaryotic species [9-17], several questions concerning their origin, evolutionary mechanism and function remain unanswered.

In Coleoptera, the presence of B chromosomes has been described in approximately 80 species belonging to several families, including Buprestidae [18], Cantharidae [19], Cicindelidae [20] and Scarabaeidae [21,22]. In general, the studies in Coleoptera have concentrated on the presence or absence of B chromosomes in species, with few reports covering their frequency in populations and/or their molecular content [18,21-23]. There are a few reports on the presence of B chromosomes in the Scarabaeidae family, including species of the Scarabaeinae and Cetoniinae

* Correspondence: cmartins@ibb.unesp.br

¹Department of Morphology, Bioscience Institute, UNESP - Sao Paulo State University, Botucatu, SP 18618-970, Brazil

Full list of author information is available at the end of the article

subfamilies [21,22]. Among scarabaeines, the *Coprophanaeus* species (Phanaeine) showed similar karyotypes consisting of $2n = 20$ and meta-submetacentric chromosomes with a gradual reduction in size, three types of sex chromosomes mechanisms (XY, Xy, XY_p), a high amount of constitutive heterochromatin, and there is no description of B chromosomes for this group until now [24-26]. Besides their karyotype characteristics, the phanaeines are restricted to the Neotropical region and play an important role in the ecosystems including nutrient recycling [27-29].

Although the cytogenetic mapping of repetitive DNA sequences has been performed for several species of coleopterans, the data are limited to the analysis of satellite DNA, rRNA and H3 histone genes e.g. [22,24-26,30-34]. Based in the heterochromatic nature of the B chromosomes and that several families of TEs are particularly enriched in heterochromatin, it is particularly interesting the analysis of TE sequences in relation to their organization in B chromosomes. Considering the gap of knowledge on the genomic content of Coleoptera B chromosomes, the present work performed molecular cytogenetic mapping of repetitive DNAs in the beetle *Coprophanaeus cyaneus*, with emphasis in the investigation of the B chromosome.

Results

The standard karyotype observed in *C. cyaneus* was $2n = 20$, XY_p ("p" refers to a "parachute" meiotic conformation between the X and Y), with meta-submetacentric chromosomes that showed a gradual reduction in size (Figure 1a). In addition, three individuals among the 50 analyzed (6%) carried 1 small-sized B meta-submetacentric chromosome. For each individual carrying the B chromosome, at least 30 metaphase I stages were analyzed, and 13.8% of the cells did not present the extra chromosome, indicating mitotic instability. The B chromosome had a condensation pattern similar to that of the autosomal chromosomes and was easily recognized as a small univalent structure in metaphase I (Figure 1).

The FISH analysis using the *Cot-1* DNA probe revealed positive hybridization in the long arms of all the autosomal chromosomes and the X and Y chromosome and in a pericentromeric block in the B chromosome (Figure 1b). The chromosomal mapping using the 18S and 5S rDNA probes showed clusters on distinct chromosomes (Figure 1c). The 18S rDNA clusters were observed at nine sites (four autosomal pairs plus one single chromosome), and the 5S rDNA clusters were observed at two sites (one autosomal pair) (Figure 1c).

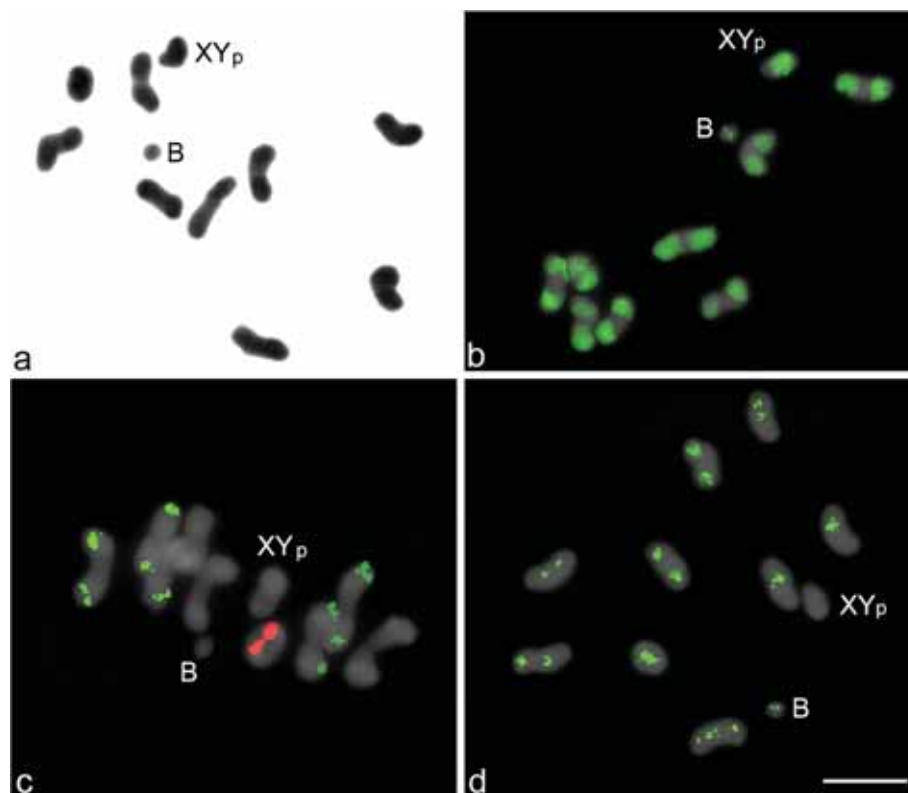


Figure 1 Metaphase I stages of *Coprophanaeus cyaneus* carrying 1 B chromosome. Conventional staining (a), FISH mapping of *Cot-1* DNA (b), 18S (green) and 5S (red) rDNA genes (c) and LOA-like non-LTR retrotransposon (d). The B chromosome and the XY_p sex chromosomes are indicated. Bar = 5 μm.

None of the rDNA probes hybridized with the B chromosome (Figure 1c).

Analysis of the non-LTR retrotransposon sequence (hereafter named the *LOA-like* non-LTR retrotransposon), which was isolated by polymerase chain reaction (PCR) and subsequently cloned, revealed a segment of 223 bp that shared ~65% similarity to the *Baggins-1_Nvi* family previously identified in *Nasonia vitripennis* [35]. The alignment of these sequences is shown in Additional file 1. FISH analysis using probes for the *LOA-like* element revealed large pericentromeric blocks in eight autosomal bivalents and the B chromosome and a pericentromeric block extending to the short arm in one autosomal pair; a positive hybridization signal was not observed in the sex chromosomes (Figure 1d).

Discussion

Basic characteristics of the *C. cyanescens* karyotype

The basic karyotype structure for *C. cyanescens* (composed of $2n = 20, XY_p$, with meta-submetacentric chromosomes) is in concordance with previous karyotype data reported for *Coprophanaeus* species [26,31,32]. However, this is the first study to identify a B chromosome in this species as well as in the Phanaeini tribe. In contrast to the small size of the B chromosome observed in *C. cyanescens*, the B chromosomes were medium- or large-sized in the other Scarabaeidae species [21,22,36]. In *Onthophagus vacca*, the presence of one medium-sized B chromosome was observed with the presence of heterochromatin in its centromeric region, whereas *Onthophagus similis* and *O. gazella* showed respectively medium- and small-sized B chromosomes; however, there was no information about the heterochromatic pattern. Large heterochromatic B chromosomes, ranging in number from three to nine, were detected in all the specimens studied for *Bubas bubalus* [21]. Individuals carrying one heterochromatic B chromosome in two populations of *Dichotomius geminatus*, corresponding to an average prevalence rate of 20.93% and 25.00% in each of the populations, were observed [22].

The frequency with which B chromosomes are detected in natural populations varies widely between populations. B chromosomes can be present in high frequencies based on the degree to which a species can tolerate the extra chromosome and their power of accumulation [23]. It is difficult to determine the factors that are involved in the low frequency of B chromosomes in the population studied, and several mechanisms may be involved, including selection, random transmission, and historical factors.

Among Coleoptera species, the studies reporting the presence of B chromosomes have generally focused on the presence or absence of this element and have not considered their frequency in the population or their molecular

content [18,21,23,36]. The presence of B chromosomes was reported in representatives of the Cetoniinae and Scarabaeinae, subfamilies of Scarabaeidae [21,22]. The evolution of the Scarabaeinae karyotype appears to have occurred under diverse mechanisms of chromosomal rearrangements [37], which could have contributed to the origin of the B chromosome in this group.

Molecular cytogenetic mapping of *C. cyanescens*

The hybridization of the *Cot-1* DNA to the pericentromeric regions extending up to the long arms of *C. cyanescens* chromosomes is in agreement with the heterochromatin distribution pattern observed in this species [26]. Although heterochromatin analyses were not conducted in the present work, the accumulation of repeated DNAs in the pericentromeric region of the B suggests also the compartmentalization of heterochromatin in the same region. The formation of the heterochromatic chromocenters in the Phanaeini species [38,39] indicates that this mechanism of heterochromatin amplification may be involved in the formation of diphasic chromosomes, including the large pericentromeric block of the B chromosome.

The distribution of *Cot-1* DNA in the A complement and the B chromosome suggests an intraspecific origin of the extra element and the occurrence of homogenization mechanisms in the heterochromatic regions between the B and A elements. Generally, B chromosomes of more recent origin are enriched in repetitive DNA sequences when compared with the genome from which they originated [1,23]. This enrichment is indicative of a massive amplification of repetitive sequences over a relatively short time-scale; and, it has also been suggested that repetitive sequences amplification may be a mechanism through which a chromosome fragment (as a neo-B chromosome) may become stabilized and selected [1,23]. This does not appear to be the case for *C. cyanescens*, indicating that the B chromosome may not have been recently established in this species. Although the data obtained indicates an intraspecific origin of the B chromosome, it was not possible to identify which chromosomal A element was involved in the process. However, the chromosomes carrying the 5S and 18S RNA genes are probably not involved in this process, as the B element does not contain rRNA gene sequences.

The cytogenetic mapping of the *LOA-like* non-LTR retrotransposon mostly to the pericentromeric regions, including those of the B chromosome, indicates the exchange of genetic material between the A and B chromosomes, implying that the B chromosome has coexisted with the A chromosomes during the period of transposition. However, it is not possible to reject the hypothesis that the B chromosome originated from a segment without *LOA-like* that was received later, by transposition. According to a previous report [40], B

chromosomes can accumulate DNA from various sources, including transposable elements, and may affect the structure of the genome by ectopic recombination. A study in *Drosophila melanogaster* identified 25 transposon-mediated rearrangements by ectopic recombination in the region flanking the white locus [41]. The B chromosomes could act as a refuge for TEs, which in turn would generate structural variability in the whole genome. The hybridization that occurred in homologous regions, such as the pericentromeric regions, is another indication of recombination between the A complement and the B chromosome, and this recombination event could be explained by the chromocenter formation during the beginning of meiosis [37].

The present study is the first report on the cytogenetic mapping of a transposable element in coleopteran chromosomes. The LOA non-LTR retrotransposon was first isolated from the genome of *Drosophila silvestris*, a species that is endemic to the Hawaiian Islands [42]. These elements belong to evolutionarily younger clades of non-LTR retrotransposons [43], contain very few known elements, and have mostly been identified in *Drosophila*, *Aedes* and *Ciona* genomes [44].

The distribution of LOA-like elements in the chromosomes reinforces an evolutionary relationship between the A complement and the B chromosome at least in the pericentromeric area. Recent work involving the centromere-enriched retrotransposons indicates that these elements preferentially insert into the centromeric regions [45]. The LOA-like elements may have been maintained in the genome of *C. cyaneus* due to a possible functional role they play in the maintenance of the pericentromeric regions. The absence of LOA-like elements in the sex chromosomes suggests that sex differentiation occurs before the distribution of this transposable element into the genome. Subsequently, the suppression of recombination could have produced the differences observed in the distribution of TEs between the A complement and the sex chromosomes. These results suggest that LOA-like element could have been involved in the maintenance of the pericentromeric regions and might contribute to the origin of the B chromosome.

Conclusions

The results obtained by the hybridization of *Cot*-1 DNA and the LOA-like non-LTR retrotransposon indicate that the origin of the B chromosome is associated with autosomal elements. The present study is the first report on the cytogenetic mapping of a transposable element in coleopteran chromosomes. Our work further suggests that TEs could also have been involved in the origin and evolution of the B chromosome in *C. cyaneus*.

Methods

Animal sampling and cytogenetic analysis

Fifty adult specimens of *Coprophanaeus cyaneus* (Olivier, 1789) (Coleoptera: Scarabaeidae: Scarabaeinae: Phanaeini) were obtained from Parque João Vasconcelos Sobrinho, Caruaru, Pernambuco State, Brazil. The specimens were collected in the wild according to Brazilian laws for environmental protection (wild collection permit, MMA/IBAMA/SISBIO n°. 2376-1). The experimental research on animals was conducted according to the international guidelines followed by São Paulo State University (Protocol no. 35/08 – CEEA/IBB/UNESP).

The testes were fixed in Carnoy solution (3:1 ethanol: acetic acid) and later stored at -20°C. The chromosome preparations were obtained by using the classical testicular follicle squashing technique. Conventional chromosome analysis was performed after staining the slides with 5% Giemsa.

Chromosomal probe isolation

The DNA samples were obtained from frozen tissues collected from specimens. The procedure for extraction of genomic DNA followed the protocol previously described [46] with minor modifications. The quality and quantity of purified DNA was evaluated in 0.8% agarose gel and spectrophotometry.

Three sets of DNA sequences were used as probes for fluorescence *in situ* hybridization (FISH) as follow: (i) sequences for the 18S and 5S rRNA genes were obtained from cloned sequences of the dung beetle, *Dichotomius semisquamatus* [22]; (ii) sequences of the LOA-like non-LTR retrotransposon were obtained from *C. cyaneus* by PCR with the RF-Co (5' CGC CTA CTT CAG GAC CAG AG 3') and RR-Co (5' AGA CTG CAG GCC GTA GAA AA 3') primers [47]; (iii) *Cot*-1 DNA sequences were isolated from *C. cyaneus* based on the DNA re-association kinetics [48] with modifications [49].

PCR products from the non-LTR retrotransposons were inserted into the pGEM-T plasmid (Promega) according to the manufacturer's recommendations, and the recombinant plasmids were used to transform competent *Escherichia coli* cells (Invitrogen, San Diego, CA, USA). The presence of the inserts in the recombinant plasmids was analyzed by PCR, and the recombinant clones were stored at -80°C. The recombinant plasmids were subjected to nucleotide sequencing using an Applied Biosystems sequencer (3500 Genetic Analyzer).

Analysis of transposable elements

The LOA-like non-LTR retrotransposon sequences isolated by PCR from *C. cyaneus* were used as queries to detect related TEs in other genomes available from the Repbase (<http://www.girinst.org/repbase/>) and NCBI (National Center for Biotechnology Information - <http://>

www.ncbi.nlm.nih.gov/) databases. The search included whole genome shotgun contigs, nucleotide collections, and high throughput genomic sequences. Analysis of the recovered DNA sequences were performed with the LIRMM software (Laboratoire Le d'Informatique, Robotique et de Microélectronique of Montpellier) available online (<http://www.phylogeny.fr/>) [50-52].

Fluorescence *in situ* hybridization

The DNA probes were labeled by nick translation with biotin-11-dATP (Invitrogen) or digoxigenin-11-dUTP (Roche, Mannheim, Germany) by PCR. The FISH technique was performed according to a protocol adapted for Coleoptera [22]. The chromosome spreads were counterstained with DAPI (4', 6-diamidino-2-phenylindole), and the slides were mounted in Vectashield mounting medium (Vector, Burlingame, CA, USA). The images were captured using an Olympus DP71 digital camera coupled to a BX61 Olympus microscope and were optimized for brightness and contrast using Adobe Photoshop CS2 and Corel Photo-Paint 13.

Additional file

Additional file 1: Alignment of the LOA non-LTR retrotransposon nucleotide sequences from *Nasonia vitripennis* (Baggins-1_NVi) and *Coprophanæus cyaneus* (Cc-1 to Cc-3). The asterisks (*) indicate similarity in sequence, and the dashes (-) indicate indels.

Abbreviations

CEEA: Comissão de Ética em Experimentação Animal; DAPI: 4' 6-Diamidino-2-Phenylindole; FISH: Fluorescence *In Situ* Hybridization; IBAMA: Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis; IBB: Instituto de Biotecnologia de Botucatu; LIRMM: Laboratoire Le d'Informatique, Robotique et de Microélectronique de Montpellier; LTR: Long Terminal Repeat; MMA: Ministério do Meio Ambiente; NCBI: National Center for Biotechnology Information; PCR: Polymerase Chain Reaction; rDNA: ribosomal DNA; rRNA: ribosomal RNA; SISBIO: Sistema de Autorização e Informação em Biodiversidade; UNESP: Universidade Estadual Paulista; TE(s): Transposable Element(s).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SGO, RCM and CM contributed to the development of the hypothesis, specimen collection and preparation, and analysis and interpretation of data. SGO and CM drafted the first version of the manuscript. RCM revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors are grateful to CMQ Costa and FAB Silva for the taxonomic identification of the studied species. The study was supported by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) of Brazil.

Author details

¹Department of Morphology, Bioscience Institute, UNESP - São Paulo State University, Botucatu, SP 18618-970, Brazil. ²Department of Biology, Biological Sciences Institute, UPE - Pernambuco State University, Recife, PE 50100-130, Brazil.

Received: 20 August 2012 Accepted: 4 November 2012

Published: 6 November 2012

References

- Camacho JP, Sharbel TF, Beukeboom LW: **B-chromosome evolution.** *Phil Trans R Soc Lond B* 2000, **355**:163-178.
- White MJD: *Animal Cytology and Evolution*. 3rd edition. London: Cambridge University; 1973.
- Jones RN: **B-chromosome drive.** *Am Nat* 1991, **137**:430-442.
- Palestis BG, Trivers R, Burt A, Jones RN: **The distribution of B chromosomes across species.** *Cytogenet Genome Res* 2004, **106**:151-158.
- Green DM: **Structure and evolution of B chromosomes in amphibians.** *Cytogenet Genome Res* 2004, **106**:235-242.
- Marschner S, Meister A, Blattner FR, Houben A: **Evolution and function of B chromosome 45S rDNA sequences in *Brachycome dichromosomatica*.** *Genome* 2007, **50**:638-644.
- Houben A, Nasuda S, Endo TR: **Plant B chromosome.** *Methods Mol Biol* 2011, **701**:97-111.
- Kean VM, Fox DP, Faulkner R: **The accumulation mechanism of the supernumerary (B-) chromosome in *Picea sitchensis* (Bong.) Carr. and the effect of this chromosome on male and female flowering.** *Silvae Genet* 1982, **31**:126-131.
- Houben A, Kynast RG, Heim U, Hermann H, Jones RN, Forster JW: **Molecular cytogenetic characterization of the terminal heterochromatic segment of the B chromosome of rye (*Secale cereale*).** *Chromosoma* 1996, **105**:97-103.
- Houben A, Leach CR, Verlin D, Rofe R, Timmis JN: **A repetitive DNA sequence common to the different B chromosomes of the genus *Brachycome*.** *Chromosoma* 1997, **106**:513-519.
- McAllister BF, Werren JH: **Hybrid origin of a B chromosome (PSR) in the parasitic wasp *Nasonia vitripennis*.** *Chromosoma* 1997, **106**:243-253.
- Sharbel TF, Green DM, Houben A: **B chromosome origin in the endemic New Zealand frog *Leiopelma hochstetteri* through sex chromosome evolution.** *Genome* 1998, **41**:14-22.
- Goodwin SB, M'Barek SB, Dhillon B, Wittenberg AHJ, Crane CF: **Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis.** *PLoS Genetics* 2001, **7**:e1002070.
- Perfectti F, Werren JH: **The interspecific origin of B chromosomes: experimental evidence.** *Evolution* 2001, **55**:1069-1073.
- Bertolotto CEV, Pellegrino KCM, Yonenaga-Yassuda Y: **Occurrence of B chromosomes in lizards: a review.** *Cytogenet Genome Res* 2004, **106**:243-246.
- Coluccia E, Cannas R, Cau A, Deiana AM, Salvadori S: **B chromosomes in Crustacea Decapoda.** *Cytogenet Genome Res* 2004, **106**:215-221.
- Borisov YM: **The polymorphism and distribution of B chromosomes in germline and somatic cells of *Tscherskia triton* de Winton (Rodentia, Cricetinae).** *Russ J Genet* 2012, **48**:538-542.
- Moura RC, Melo NF, Souza MJ: **High levels of chromosomal differentiation in *Euchroma gigantea* L 1735 (Coleoptera, Buprestidae).** *Gen Mol Biol* 2008, **31**:431-437.
- James LV, Angus RB: **A chromosomal investigation of some British Cantharidae (Coleoptera).** *Genetica* 2007, **130**:293-300.
- Proença SJR, Serrano ARM, Collares-Pereira MJ: **Cytogenetic variability in genus *Odontocheila* (Coleoptera, Cicindelidae): karyotypes, C-banding, NORs and localization of ribosomal genes of *O. confusa* and *O. nodicornis*.** *Genetica* 2002, **114**:237-245.
- Angus RB, Wilson CJ, Mann DJ: **A chromosomal analysis of 15 species of Gymnopleurini and Coprini (Coleoptera: Scarabaeidae).** *Tijdsch Voor Entomol* 2007, **150**:201-211.
- Cabral-de-Mello DC, Moura RC, Martins C: **Chromosomal mapping of repetitive DNAs in the beetle *Dichotomius geminatus* provides the first evidence for an association of 5S rRNA and histone H3 genes in insects, and repetitive DNA similarity between the B chromosome and A complement.** *Heredity* 2010, **104**:393-400.
- Camacho JPM: **B Chromosomes.** In *The Evolution of the Genome*. Edited by Gregory TR. San Diego: Elsevier; 2005:223-286.
- Palomeque T, Muñoz-López M, Carrillo JA, Lorite P: **Characterization and evolutionary dynamics of a complex family of satellite DNA in the leaf beetle *Chrysolina camifex* (Coleoptera, Chrysomelidae).** *Chrom Res* 2005, **13**:795-807.
- Cabral-de-Mello DC, Moura RC, Carvalho R, Souza MJ: **Cytogenetic analysis of two related *Deltochilum* (Coleoptera, Scarabaeidae) species: diploid**

- number reduction, extensive heterochromatin addition and differentiation. *Micron* 2010, **41**:112–117.
26. Oliveira SG, Moura RC, Silva AEB, Souza MJ: Cytogenetic analysis of two *Coprophanaeus* species (Scarabaeidae) revealing wide constitutive heterochromatin variability and the largest number of 45S rDNA sites among Coleoptera. *Micron* 2010, **41**:960–965.
 27. Edmonds WD: Comparative skeletal morphology, systematics and evolution of the Phanaeina dung beetles (Coleoptera: Scarabaeidae). *Kans Univ Sci Bull* 1972, **49**:731–874.
 28. Philips TK, Edmonds WD, Scholtz CH: A phylogenetic analysis of the New World tribe Phanaeini (Coleoptera: Scarabaeidae: Scarabaeinae): Hypotheses on relationships and origins. *Insect Syst Evol* 2004, **35**:43–63.
 29. Edmonds WD, Zidek JA: Taxonomic review of the neotropical genus *Coprophanaeus* Olsoufieff, 1924 (Coleoptera: Scarabaeidae, Scarabaeinae). *Insecta Mundi* 2010, **129**:1–111.
 30. Bione EG, Moura RC, Carvalho R, Souza MJ: Karyotype, C-banding pattern, NOR location and FISH study of five Scarabaeidae (Coleoptera) species. *Gen Mol Biol* 2005, **28**:376–381.
 31. Oliveira SG, Cabral-de-Mello DC, Arcanjo AP, Xavier C, Souza MJ, Martins C, Moura RC: Heterochromatin, sex chromosomes and rRNA gene clusters in *Coprophanaeus* beetles (Coleoptera, Scarabaeidae). *Cytogenet Genome Res* 2012, **138**:46–55.
 32. Cabral-de-Mello DC, Oliveira SG, Moura RC, Martins C: Chromosomal organization of the 18S and 5S rRNAs and histone H3 genes in Scarabaeinae coleopterans: insights into the evolutionary dynamics of multigene families and heterochromatin. *BMC Genet* 2011, **12**:88.
 33. Pons J, Bruvo B, Juan C, Petitpierre E, Plohl M, Ugarkovic D: Conservation of satellite DNA in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). *Gene* 1997, **255**:183–190.
 34. Gálian J, Vogler AP: Evolutionary dynamics of a satellite DNA in the tiger beetle species pair *Cicindela campestris* and *C. maroccana*. *Genome* 2003, **46**:213–223.
 35. Jurka J: LINE retrotransposons from the parasitic wasp *Nasonia vitripennis*. *Repbase Reports* 2009, **9**:483–483.
 36. Wilson CJ, Angus RB: A chromosomal analysis of 21 species of Oniticellini and Onthophagini (Coleoptera: Scarabaeidae). *Tijdschr Entomol* 2005, **148**:63–76.
 37. Cabral-de-Mello DC, Oliveira SG, Ramos IC, Moura RC: Karyotype differentiation patterns in species of the subfamily Scarabaeinae (Scarabaeidae, Coleoptera). *Micron* 2008, **39**:1243–1250.
 38. Bione EG, Camparoto ML, Simões ZL: A study of the constitutive heterochromatin and nucleolus organizer regions of *Isocopris inhiata* and *Diabroctis mimas* (Coleoptera: Scarabaeidae, Scarabaeinae) using C-banding, AgNO₃ staining and FISH techniques. *Gen Mol Biol* 2005, **28**:111–116.
 39. Colomba M, Vitturi R, Libertini A, Gregorini A, Zunino M: Heterochromatin of the scarab beetle, *Bubas bison* (Coleoptera: Scarabaeidae) II. Evidence for AT-rich compartmentalization and a high amount of rDNA copies. *Micron* 2006, **37**:47–51.
 40. Beukeboom LW: Bewildering Bs: an impression of the 1st B-chromosome conference. *Heredity* 1994, **73**:328–336.
 41. Montgomery E-A, Huang S-M, Langley CH, Judd BH: Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. *Genetics* 1991, **129**:1085–1098.
 42. Felger I, Hunt JA: A non-LTR retrotransposon from the Hawaiian *Drosophila*: the LOA element. *Genetica* 1992, **85**:119–130.
 43. Volff J-N, Leirach H, Reinhardt R, Chourroutà D: Retroelement dynamics and a novel type of chordate retrovirus-like element in the miniature genome of the tunicate *Oikopleura dioica*. *Mol Biol Evol* 2004, **21**:2022–2033.
 44. Rho M, Tang H: MGEscan-non-LTR: computational identification and classification of autonomous non-LTR retrotransposons in eukaryotic genomes. *Nuc Acids Res* 2009, **37**:e143.
 45. Birchler JA, Presting GG: Retrotransposon insertion targeting: a mechanism for homogenization of centromere sequences on nonhomologous chromosomes. *Genes Dev* 2012, **26**:638–640.
 46. Sambrook J, Russel DW: *Molecular Cloning. A Laboratory Manual*. Third Editionth edition. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2001.
 47. Burke WD, Eickbush DG, Xiong Y, Jakubczak J, Eickbush TH: Sequence relationship of retrotransposable elements RI and R2 within and between divergent insect species. *Mol Biol Evol* 1993, **10**:163–185.
 48. Zwick MS, Hanson RE, McKnight TD, Nurul-Islam-Faridi M, Stelly DM: A rapid procedure for the isolation of C₀t-1 DNA from plants. *Genome* 1997, **1997**(40):138–142.
 49. Ferreira IA, Martins C: Physical chromosome mapping of repetitive DNA sequences in Nile tilapia *Oreochromis niloticus*: evidences for a differential distribution of repetitive elements in the sex chromosomes. *Micron* 2008, **39**:411–418.
 50. Guindon S, Gascuel O: A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003, **52**:696–704.
 51. Chevenet F, Brun C, Banuls AL, Jacq B, Christen R: TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics* 2006, **7**:439.
 52. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-F, Guindon S, Lefort V, Lescot M, Claverie J-M, Gascuel O: Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 2008, **36**:465–469.

doi:10.1186/1471-2156-13-96

Cite this article as: Gomes de Oliveira et al.: B chromosome in the beetle *Coprophanaeus cyanescens* (Scarabaeidae): emphasis in the organization of repetitive DNA sequences. *BMC Genetics* 2012 **13**:96.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



SHORT REPORT

Open Access

Horizontal transfers of *Mariner* transposons between mammals and insects

Sarah G Oliveira^{1*}, Weidong Bao², Cesar Martins¹ and Jerzy Jurka²

Abstract

Background: Active transposable elements (TEs) can be passed between genomes of different species by horizontal transfer (HT). This may help them to avoid vertical extinction due to elimination by natural selection or silencing. HT is relatively frequent within eukaryotic taxa, but rare between distant species.

Findings: Closely related *Mariner*-type DNA transposon families, collectively named as *Mariner-1_Tbel* families, are present in the genomes of two ants and two mammalian genomes. Consensus sequences of the four families show pairwise identities greater than 95%. In addition, mammalian *Mariner1_BT* family shows a close evolutionary relationship with some insect *Mariner* families. Mammalian *Mariner1_BT* type sequences are present only in species from three groups including ruminants, tooth whales (*Odontoceti*), and New World leaf-nosed bats (*Phyllostomidae*).

Conclusions: Horizontal transfer accounts for the presence of *Mariner_Tbel* and *Mariner1_BT* families in mammals. *Mariner_Tbel* family was introduced into hedgehog and tree shrew genomes approximately 100 to 69 million years ago (MYA). Most likely, these TE families were transferred from insects to mammals, but details of the transfer remain unknown.

Keywords: DNA transposon, Genome evolution, Horizontal transfer, *Mariner*

Findings

In contrast to the vertical transmission of the genetic material from parents to offspring, the horizontal transfer (HT) is a process in which new genetic information is transmitted between different, sometimes distant, species [1,2]. HT is likely to be one of the factors leading to the persistence of transposable elements (TEs) in eukaryotes [3-5], and complicating the evolutionary trees.

The detection of HT is mostly inferential, mainly based on the combination of two types of evidence: unusually high similarity between TE sequences from species that have long diverged from each other, and a limited distribution of one particular TE family within a group of species [6]. To date, numerous HTs have been detected in eukaryotes [6-10], but of particular interest are HTs across distant branches. A recent example of such a rare event is HT of hAT DNA transposon families between vertebrate and invertebrate species [11].

Here, we report two families of *Mariner*-type DNA transposons that have possibly undergone HT from insects to mammals. The first family, called *Mariner_Tbel*, was originally identified in the tree shrew (*Tupaia belangeri*), but families nearly identical to *Mariner_Tbel* were also found in the genome of another mammal, European hedgehog (*Erinaceus europaeus*), and in two ant species: red harvester ant (*Pogonomyrmex barbatus*) and Jerdon's jumping ant (*Harpegnathos saltator*) (Tables 1 and 2). Although the copy numbers and divergence vary between the families, the family consensus sequences reconstructed in each genome show a high level of identity to each other throughout the entire length (approximately 1.3 kb) (Table 1). The lowest identity is found between the two families in *E. europaeus* and ant *P. barbatus* (95.84%), and the highest identity is found between the two ant families (98.45%). Therefore, unless otherwise stated, the four families in the genomes are referred collectively to as *Mariner_Tbel* families. Given the long divergence time between insects and mammals (approximately 1 billion years) [12-14], this high identity strongly indicates that HT took place during the evolutionary history of *Mariner_Tbel* families.

* Correspondence: sarahg.oliveira@gmail.com

¹Morphology Department, Bioscience Institute, UNESP - Sao Paulo State University, Botucatu, Sao Paulo 18618-970, Brazil

Full list of author information is available at the end of the article

Table 1 Pairwise identities (%) between the *Mariner_Tbel* consensus from mammals (*Tupaia belangeri*, *Erinaceus europaeus*) and insects (*Pogonomyrmex barbatus*, *Harpegnathos saltator*)

	<i>T. belangeri</i>	<i>E. europaeus</i>	<i>P. barbatus</i>
<i>E. europaeus</i>	96.55	-	-
<i>P. barbatus</i>	98.05	95.84	-
<i>H. saltator</i>	97.90	95.69	98.45

This notion is consistent with the fact that mammal *Mariner_Tbel* sequences were found only in two distantly related mammalian species, even though over 30 mammalian genomes were sequenced to date.

We then estimated the approximate ages of the four *Mariner_Tbel* families in each genome. In mammals, we compared the sequence divergences of *Mariner_Tbel* to an older *Mariner*-type family (*TIGGER1*), relatively common in the mammalian genomes (Table 2). *TIGGER1* elements are present in multiple copies in eutherian mammals, but only one or two degenerated copies were found in marsupial genomes, including *Macropus eugenii*, *Monodelphis domestica*, and *Sarcophilus harrisii* (Figure 1A). Therefore, mammalian *TIGGER1* families likely expanded after the split of marsupials and placentals (190 million years ago (MYA)), but before the placental radiation (approximately 100 MYA) [15]. In the genome of the tree shrew and European hedgehog the divergence of the *TIGGER1* family is $21.2 \pm 2\%$ and $28.0 \pm 3\%$, respectively (Table 2). Therefore, based on the divergence of *Mariner_Tbel* in the two mammal genomes ($15.0 \pm 2\%$ and $19.4 \pm 3\%$, respectively), the ages of *Mariner_Tbel* in the two mammals were calculated to be approximately 134 to 70 MYA and approximately 131 to 69 MYA, respectively. Because it is unlikely that the mammalian *Mariner_Tbel* expanded in the common

ancestor of placental mammals before 100 MY, we adjusted the ages to be 100 to 70 MYA and 100 to 69 MYA, respectively (see Figure 1A and [15]).

In the ant genomes, no *Mariner* family was yet identified as unambiguously present in the common ancestor of all ant species. Among potential candidates are the oldest known *Mariner* families present in some of the ant genomes (for example, *Mariner-28_SIn* or *Mariner-94_HSa*; Figure 1). These small families may have expanded in the common ancestor of all ant species (140 MYA) [13], assuming that they were lost in some ant species. Alternatively, these old families might have expanded in some ant species after they split from their common ancestor. Under either scenario, the outermost ages when the two ant *Mariner_Tbel* families expanded could be still estimated by comparing their diversities with the diversity of *Mariner-28_SIn* (Table 2). Based on that, *Mariner_Tbel* family in the red harvester ant (*P. barbatus*) and Jerdon's jumping ant (*H. saltator*) expanded at most approximately 43 and approximately 50 million years ago, respectively.

The above age estimates suggest that the two ant *Mariner_Tbel* families are possibly younger than the mammalian *Mariner_Tbel* families. However, the history of *Mariner_Tbel* can be traced further back in ants and their insect relatives than in mammals. Individual *Mariner_Tbel*-like elements from distinct families, such as AEAQ01009575, AEAB01001421 and AFJA01006902 (Figure 1B), were also found in the genomes of two other ants (*Solenopsis invicta* and *Camponotus floridanus*) as well as in the alfalfa leafcutting bee (*Megachile rotundata*). These *Mariner_Tbel*-like sequences and *Mariner_Tbel* sequences form a single lineage in the phylogenetic tree (Figure 1B), with the bee sequences in a more ancestral position (Figure 1B). The topology of this particular lineage mirrors the evolutionary history of

Table 2 Divergence of *Mariner* transposable element (TE) families in mammalian and insect genomes

Family	Length (bp)	Copy no.	Divergence (%) ^a
<i>Mariner-1_Tbel</i> (TBel)	1,279	>400	15.0 ± 2 (183)
<i>Mariner-1_Tbel</i> (EEr)	1,266	>70	19.4 ± 3 (34)
<i>Mariner-1_Tbel</i> (PBa)	1,285	>90	6.3 ± 1 (53)
<i>Mariner-1_Tbel</i> (HSa)	1,285	>30	7.2 ± 2 (16)
<i>TIGGER1</i> (Tbel)	2,413	>80	21.2 ± 2 (27)
<i>TIGGER1</i> (EEu)	2,410	>47	28.0 ± 3 (25)
<i>TIGGER1</i> (BT)	2,408	>500	17.3 ± 2 (102)
<i>TIGGER1</i> (TTr)	2,419	>580	12.1 ± 1 (159)
<i>Mariner-28_SIn</i> (SIn)	1,226	Approximately 14	20.1 ± 3 (12)
<i>Mariner1_BT</i> (BT) ^b	1,277	>400	14.7 ± 2 (95)
<i>Mariner1_BT</i> (TTr)	1,285	>700	7.6 ± 1 (101)

^aThe divergence represents average pairwise k-distance between individual copies and the consensus. Numbers of individual sequences used for the k-distance calculation are indicated in parentheses.

^b*Mariner1_BT* families from other *Bos indicus*, *Bos grunniens*, *Bubalus bubalis* are not shown.

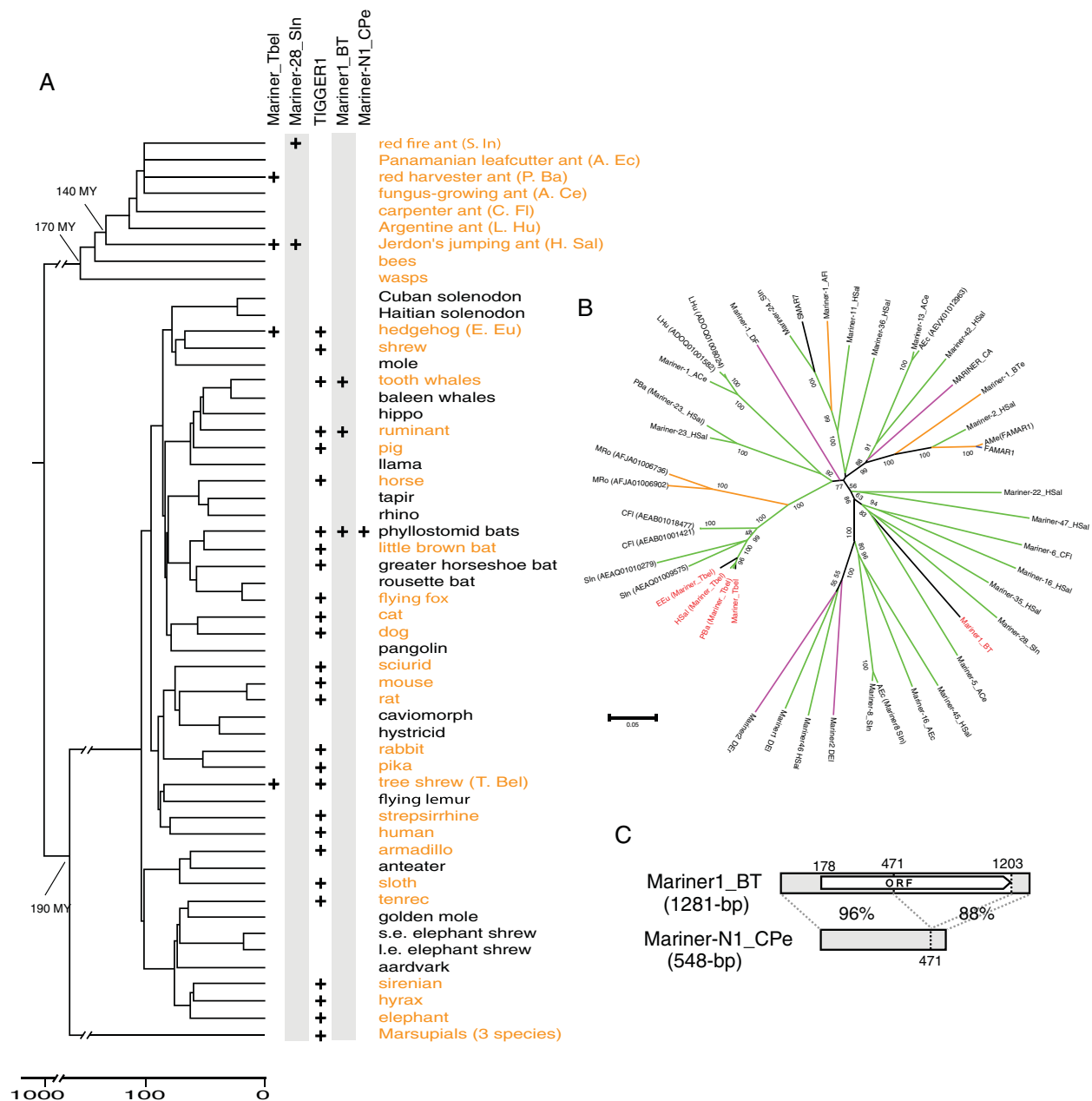


Figure 1 (A) Distribution of five Mariner families (*Mariner_Tbel*, *Mariner1_BT*, *TIGGER1*, *Mariner-28_Sln*, and *Mariner-N1_CPe*) in placental mammals, marsupials, and insects. The phylogeny is adopted from published trees of insects and mammals [12-15]. Marsupials are represented by three species (*Monodelphis domestica*, *Macropus eugenii*, *Sarcophilus harrisii*). The scale bar on the bottom left indicates time of evolutionary branching in the tree. Species with genomic sequences available from public databases are highlighted in orange. **(B)** Phylogenetic position of the two horizontally transferred Mariners (*Mariner_Tbel* and *Mariner1_BT*; red color), relative to other closely related Mariners. Most of the families are from insects with color-coded branches: ants (green), bees (orange), flies (pink) and other insects (blue). The remaining branches (black) are from mammals and planarian (SMAR7). Except for a few individual sequence segments (with accession numbers), all other families are represented by consensus sequences deposited in Repbase (except for highly similar copies). **(C)** Structural relationship between *Mariner1_BT* and *Mariner-N1_CPe* family. The species abbreviations are: ACe (*Atta cephalotes*), AEc (*Acromyrmex echinator*), AFI (*Apis florea*), AMe (*Apis mellifera*), BTe (*Bombus terrestris*), BT (*Bos taurus*), CA (*Chymomyza amoena*), CFI (*Camponotus floridanus*), Del (*Drosophila elegans*), DEr (*Drosophila erecta*), DF (*Drosophila ficusphila*), EEu (*Erinaceus europaeus*), HSal (*Harpegnathos saltator*), LHu (*Linepithema humile*), MRO (*Megachile rotundata*), PBa (*Pogonomyrmex barbatus*), Sln (*Solenopsis invicta*), SMAR7 (*Schmidtea mediterranea*), Tbel (*Tupaia belangeri*), CPe (*Carollia perspicillata*).

the ant and bee species (Figure 1A). Furthermore, Figure 1B indicates that the *Mariner_Tbel* family and many other similar *Mariner* families in ants and other insects shared a common ancestral sequence. These observations suggest the ancestor of ants *Mariner_Tbel* may have been present in some ant or other insect species very long time ago, probably as far back as the common ancestor of bees and ants (approximately 150 MYA) [14]. Thus, the mammalian *Mariner_Tbel* families probably originated from HTs from insects to mammals through some unknown vectors. Given that the two mammals belong to two distinct lineages, *Mariner_Tbel* in tree shrew and hedgehog may represent two independent HTs (Figure 1A). Notably, we cannot rule out the possibility that the *Mariner_Tbel* families in one of the two ant species, or both, also originated by HTs. This possibility is suggested by two facts: (a) the relatively young ages (at most approximately 43 to 50 MYA) of the two families, (b) the high identity (98.5%) between the two family consensus sequences, even *H. saltator* and *P. barbatus* diverged from each other approximately 100 million years ago [13]. Among insect species, frequent HTs have been documented in flies [16]. Alternatively, *Mariner_Tbel* sequences could have survived for a very long time in either of the two ant genomes before the most recent family expansions.

In addition to mammalian *Mariner_Tbel* families, *Mariner1_BT* DNA families might also have originated by HT from insects (Figure 1A). We were able to obtain high quality consensus of *Mariner1_BT* from bovine species (*Bos taurus*, *Bos indicus*, *Bos grunniens*, *Bubalus bubalis*) and bottlenosed dolphin *Tursiops truncatus* (Table 2). All the derived consensus sequences show similar lengths (approximately 1,280 bp), and high pairwise identities throughout the entire length (>98%). Blast screening against National Center for Biotechnology Information (NCBI) databases using *Mariner1_BT* consensus sequence as query also detected this family in several other mammalian species, including one bat species, *Carollia perspicillata* (Seba's short-tailed bat), additional ruminants, and whale (Table 3). These BlastN hits show similar score and query coverage (>80% identity to the

consensus and >90% coverage). In summary, *Mariner1_BT* type TEs were found only in three taxonomic groups to date: ruminants, tooth whales (*Odontoceti*), and New World leaf-nosed bats (*Phyllostomidae*) (Figure 1A). Notably, in *C. perspicillata* (short-tailed fruit bat) and *Desmodus rotundus* (vampire bat), we also detected a family of non-autonomous DNA transposon, called *Mariner-N1_CPe*, which was likely derived from the bat *Mariner1_BT* family (Figure 1C).

Remarkably, the other closest relatives of *Mariner1_BT* are all found in ant species: *Mariner1_BT* coclusters significantly (bootstrap = 83) with three other ant *Mariner* families (*Mariner-5_ACe*, *Mariner-28_SIn* and *Mariner-35_HSal*) (Figure 1B). Given the vast diversity of *Mariners* found in insects (Figure 1B), and the confined distribution of *Mariner1_BT* in mammals, we propose *Mariner1_BT* family could also originate from a horizontally transferred insect-like element. Using a similar method above, that is, based on the family divergence and mammalian phylogeny (Table 2 and Figure 1A), we estimated the ages of bovine *Mariner1_BT* to be 90 to 85 MYA, and 90 to 63 MYA for dophin *T. truncatus* *Mariner1_BT* family. The age of *Mariner1_BT* in bat could not be estimated due to insufficient data. We also could not determine if HT happened in mammals more than once, because the three taxonomic groups that include *Mariner1_BT* are relatively close.

In summary, this is the first report of two cases of horizontally transferred *Mariner* elements (*Mariner_Tbel* and *Mariner1_BT*) between insects and mammals. Previously, four families of DNA transposons from the *hAT* superfamily were also found to be involved in multiple waves of HT between insects and other vertebrates including mammals [11]. This could partially be attributed to the fact that insects are the largest and the most diverse group of invertebrate animals on earth. While insects are the most likely source of the horizontally transferred transposons, the original source or possible intermediaries, such as parasitic insects [11] or viruses, remain unclear. This is complicated by the possibility that recurrent HTs of related *Mariner* elements are likely to take place between different insects [16]. The

Table 3 *Mariner1_BT* sequences detected in mammals

Groups	Species	Accession	Score	Query coverage	E value	Identity	Gaps
Bat	<i>Carollia perspicillata</i>	AC152852.2	1,324	98%	0	1,068/1,291 (83%)	87/1,291 (7%)
Ruminants	<i>Odocoileus hemionus</i>	AY330343.1	1,278	98%	0	1,061/1,294 (82%)	73/1,294 (6%)
	<i>Ovis aries</i>	AC148039.3	1,274	99%	0	1,078/1,309 (82%)	64/1,309 (5%)
	<i>Muntiacus reevesi</i>	AC174385.3	1,265	99%	0	1,074/1,312 (82%)	68/1,312 (5%)
	<i>Muntiacus muntjak vaginalis</i>	AC152844.3	1,242	99%	0	1,071/1,313 (82%)	71/1,313 (5%)
	<i>Capra hircus</i>	EU870890.1	1,142	98%	0	1,031/1,296 (80%)	109/1,296 (8%)
Whale	<i>Pseudorca crassidens</i>	AP011081.1	1,833	98%	0	1,180/1,288 (92%)	22/1,288 (2%)

The table shows only top scores using *Mariner1_BT* as BlastN query.

role of viruses in HT proposed some time ago [17] still remains to be understood. As more genome sequence data become available, more mechanistic details on HT between mammals and insects are likely to emerge.

Methods

Mariner transposable elements from Repbase (<http://www.girinst.org/repbase/>) were used as an initial query to screen *Mariners* in diverse genomes available at NCBI (National Center for Biotechnology Information: <http://www.ncbi.nlm.nih.gov/>). Family consensus sequences were constructed whenever possible. The copy numbers in each family were determined by BLASTN using consensus sequences as queries. Sequence divergence within each family was assessed by the average pairwise k-distance (Kimura two-parameter model) between individual insertions and the corresponding consensus sequences. The k-distance was calculated using the software MEGA 4 [18]. For a given family, individual sequences used in k-distance calculation were randomly chosen from the family members; in most cases individual sequences matched >70% of the consensus length.

We used *Mariner_Tbel* and *Mariner1_BT* as BLASTN queries against Repbase to select top scoring TE entries for phylogeny analysis. Individual sequences selected from GenBank were also used in the tree if Repbase consensus sequences were not available. The sequence alignments are shown in Additional file 1. The alignments were done using the online MAFFT server (<http://mafft.cbrc.jp/alignment/software/>). The phylogeny tree was inferred using MEGA 4 [18], using the neighbor joining (NJ) method and k-distances. Branch support was estimated using 1,000 bootstrap replicates.

Additional file

Additional file 1: Sequence alignments of three *Mariner* families (*Mariner_Tbel*, *Mariner1_BT*, *Mariner-28_Sln*) from eutherian mammals and insects. Except for a few individual sequence segments (with accession numbers), all other families are represented by consensus sequences deposited in Repbase (excluding highly similar copies). The species are: ACe (*Atta cephalotes*), AEc (*Acromyrmex echinator*), AFl (*Apis florea*), AMe (*Apis mellifera*), BTe (*Bombus terrestris*), BT (*Bos taurus*), CA (*Chymomyza amoena*), CFl (*Camponotus floridanus*), Del (*Drosophila elegans*), DEr (*Drosophila erecta*), DF (*Drosophila fuscipennis*), EEu (*Erinaceus europaeus*), HSal (*Harpegnathos saltator*), LHu (*Linepithema humile*), MRo (*Megachile rotundata*), PBa (*Pogonomyrmex barbatus*), Sln (*Solenopsis invicta*), SMAR7 (*Schmidtea mediterranea*), Tbel (*Tupaia belangeri*).

Abbreviations

ACe: *Atta cephalotes*; AEc: *Acromyrmex echinator*; AFl: *Apis florea*; AMe: *Apis mellifera*; BT: *Bos taurus*; BTe: *Bombus terrestris*; CA: *Chymomyza amoena*; CFl: *Camponotus floridanus*; Del: *Drosophila elegans*; DEr: *Drosophila erecta*; DF: *Drosophila fuscipennis*; EEu: *Erinaceus europaeus*; HSal: *Harpegnathos saltator*; HT: Horizontal transfer; LHu: *Linepithema humile*; MRo: *Megachile rotundata*; PBa: *Pogonomyrmex barbatus*; Sln: *Solenopsis invicta*; SMAR7: *Schmidtea mediterranea*; Tbel: *Tupaia belangeri*; TTr: *Tursiops truncatus*.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SGO contributed to development of the hypothesis, collection, preparation, analysis and interpretation of data, wrote the first draft of the manuscript, and revised the text. WB contributed to the analysis and interpretation of data, writing and revising the manuscript. CM contributed to the discussion of data and revisions of the manuscript. JJ contributed to development of the hypothesis, interpretation of data and final revisions. All the authors read and approved the final manuscript.

Acknowledgements

This work was supported by funds from the Sao Paulo Research Foundation (FAPESP), Sao Paulo State University (UNESP) and the National Institutes of Health grant 5 P41 LM006252. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Library of Medicine or the National Institutes of Health.

Author details

¹Morphology Department, Bioscience Institute, UNESP - Sao Paulo State University, Botucatu, Sao Paulo 18618-970, Brazil. ²Genetic Information Research Institute, 1925 Landings Drive, Mountain View, CA 94043, USA.

Received: 11 May 2012 Accepted: 24 August 2012

Published: 26 September 2012

References

- Syvanen M: Horizontal gene transfer: evidence and possible consequences. *Annu Rev Genet* 1994, **28**:237–261.
- Brown JR: Ancient horizontal gene transfer. *Nat Rev Genet* 2003, **4**:121–132.
- Hartl DL, Lohe AR, Lozovskaya ER: Modern thoughts on an ancient mariner: function, evolution, regulation. *Annu Rev Genet* 1997, **31**:337–358.
- Sanchez-Gracia A, Maside X, Charlesworth B: High rate of horizontal transfer of transposable elements in *Drosophila*. *Trends Genet* 2005, **21**:200–203.
- Schaack S, Gilbert C, Feschotte C: Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol Evol* 2010, **25**:537–546.
- Silva JC, Loreto EL, Clark JB: Factors that affect the horizontal transfer of transposable elements. *Curr Issues Mol Biol* 2004, **6**:57–71.
- Daniels SB, Peterson KR, Strausbaugh LD, Kidwell MG, Chovnick A: Evidence for horizontal transmission of the P transposable element between *Drosophila* species. *Genetics* 1990, **124**:339–355.
- Loreto EL, Valente VL, Zaha A, Silva JC, Kidwell MG: *Drosophila mediopunctata* P elements: a new example of horizontal transfer. *J Hered* 2001, **92**:375–381.
- Maruyama K, Hartl DL: Evidence for interspecific transfer of the transposable element mariner between *Drosophila* and *Zaprionus*. *J Mol Evol* 1991, **33**:514–524.
- Robertson HM, Lampe DJ: Recent horizontal transfer of a mariner transposable element among and between *Diptera* and *Neuroptera*. *Mol Biol Evol* 1995, **12**:850–862.
- Gilbert C, Schaack S, Pace JK 2nd, Brindley PJ, Feschotte C: A role for host-parasite interactions in the horizontal transfer of transposons across phyla. *Nature* 2010, **464**:1347–1350.
- Hedges SB, Blair JE, Venturi ML, Shree JL: A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol Biol* 2004, **4**:2.
- Moreau CS, Bell CD, Vila R, Archibald SB, Pierce NE: Phylogeny of the ants: diversification in the age of angiosperms. *Science* 2006, **312**:101–104.
- Brady SG, Larkin L, Danforth BN: Bees, ants, and stinging wasps (Aculeata). In *The Timetree of Life*. Edited by Hedges SB, Kumar S. Oxford, UK: Oxford University Press; 2009:264–269.
- Meredith RW, Janečka JE, Gatesy J, Ryder OA, Fisher CA, Teeling EC, Goodbla A, Eizirik E, Simão TL, Stadler T, Rabosky DL, Honeycutt RL, Flynn JJ, Ingram CM, Steiner C, Williams TL, Robinson TJ, Burk-Herrick A, Westerman M, Ayoub NA, Springer MS, Murphy WJ: Impacts of the Cretaceous terrestrial revolution and KPg extinction on mammal diversification. *Science* 2011, **334**:521–524.

16. Bartolome C, Bello X, Maside X: **Widespread evidence for horizontal transfer of transposable elements across *Drosophila* genomes.** *Genome Biol* 2009, **10**:R22.
17. Kidwell MG: **Lateral transfer in natural populations of eukaryotes.** *Annu Rev Genet* 1993, **27**:235–256.
18. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.** *Mol Biol Evol* 2007, **24**:1596–1599.

doi:10.1186/1759-8753-3-14

Cite this article as: Oliveira *et al.*: Horizontal transfers of *Mariner* transposons between mammals and insects. *Mobile DNA* 2012 3:14.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Anexo B: Relação de trabalhos paralelos publicados relacionados ao tema da tese**Artigo completo publicado em periódico**

Cabral-de-Mello DC, Oliveira SG, Moura RC, Martins (2011) Chromosomal organization of the 18S and 5S rRNAs and histone H3 genes in Scarabaeinae coleopterans: insights into the evolutionary dynamics of multigene families and heterochromatin. BMC Genetics 12:88.

Capítulo de livro publicado

Martins C, Cabral-de-Mello DC, Valente GT, Mazzuchelli J, Oliveira SG (2010) Cytogenetic mapping and its contribution to the knowledge of animal genomes. In: Columbus F. (Org.). Genetic Mapping. 1ed. Hauppauge: Nova Science Publisher, p. 1-82.

Livro publicado

Martins C, Cabral-de-Mello DC, Valente GT, Mazzuchelli J, Oliveira SG, Pinhal D (2011) Animal Genomes Under the Focus of Cytogenetics. 1. ed. Hauppauge: Nova Science Publisher, v. 1.