



UNIVERSIDADE ESTADUAL DE LONDRINA
CENTRO DE CIÊNCIAS BIOLÓGICAS
Colegiado do CURSO DE CIÊNCIAS BIOLÓGICAS



**Ciências
Biológicas**
UEL

TRABALHO DE CONCLUSÃO DO CURSO DE GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS

LUCAS GUTHIERRES DE SOUZA SOARES

**IDENTIFICAÇÃO DE ELEMENTOS TRANSPONÍVEIS NAS
MONTAGENS *DE NOVO* E GUIADA DE LAGARTAS DE
BOMBYX MORI SUBMETIDAS AO ESTRESSE TÉRMICO**

Londrina – Paraná
2025

**TRABALHO DE CONCLUSÃO DO CURSO DE GRADUAÇÃO
EM CIÊNCIAS BIOLÓGICAS**

LUCAS GUTHIERRES DE SOUZA SOARES

**IDENTIFICAÇÃO DE ELEMENTOS TRANSPONÍVEIS NAS
MONTAGENS *DE NOVO* E GUIADA DE LAGARTAS DE
BOMBYX MORI SUBMETIDAS AO ESTRESSE TÉRMICO**

Monografia apresentada ao Curso de Graduação em Ciências Biológicas da Universidade Estadual de Londrina como um dos requisitos à obtenção do título de Bacharel em Ciências Biológicas.

Orientador: Rogério Fernandes de Souza

**Londrina – Paraná
2025**

S676 Soares, Lucas Guthierres de Souza.
IDENTIFICAÇÃO DE ELEMENTOS TRANSPONÍVEIS NAS MONTAGENS
DE NOVO E GUIADA DE LAGARTAS DE *BOMBYX MORI* SUBMETIDAS
AO ESTRESSE TÉRMICO / Lucas Guthierres de Souza Soares. - Londrina,
2025.
86 f.

Orientador: Rogério Fernandes de Souza.
Trabalho de Conclusão de Curso (Graduação em Ciências Biológicas) -
Universidade Estadual de Londrina, Centro de Ciências Biológicas, Graduação
em Ciências Biológicas, 2025.
Inclui bibliografia.

1. Bioinformática - TCC. 2. *Bombyx mori* - TCC. 3. Transcriptoma - TCC. 4.
Elementos transponíveis - TCC. I. Souza, Rogério Fernandes de. II.
Universidade Estadual de Londrina. Centro de Ciências Biológicas. Graduação
em Ciências Biológicas. III. Título.

CDU 574

BANCA EXAMINADORA

Prof. Dr. Rogério Fernandes de Souza

Dr Laurival A. Vilas-Boas

Dr^a Jaqueline Fernanda Dionisio

SUPLENTE:

Dr^a Larissa Forim Pezenti

Londrina, 27 de novembro de 2025.

DEDICATÓRIA

Dedico este trabalho a minha família, por todo apoio incondicional inclusive nos momentos difíceis, durante todos esses anos de graduação.

AGRADECIMENTOS

Gostaria de agradecer, primeiramente, à Universidade Estadual de Londrina (UEL), pelo acolhimento, e ao Laboratório de Bioinformática (LBI) pelas oportunidades, pelos recursos disponibilizados e pelo ambiente de aprendizado proporcionado. Em especial, agradeço ao meu orientador, Rogério Fernandes de Souza, por acreditar neste projeto e por estar sempre à disposição para me auxiliar.

Agradeço a minha mãe, ao meu pai e aos meus avós, pelo apoio incondicional ao longo deste trabalho, bem como por todo o investimento necessário para os deslocamentos e a permanência na universidade, além de sempre acreditarem na minha capacidade. À minha irmã, por me aconselhar, acompanhar de perto e comemorar comigo cada conquista ao longo da graduação. Aos amigos e companheiros de sala, que foram fundamentais para tornar o ambiente mais leve e acolhedor, ajudando a lidar da melhor forma possível com avaliações, projetos, trabalhos, sempre com o apoio mútuo.

SOARES, Lucas Guthierres de Souza. **Identificação de elementos transponíveis nas montagens *de novo* e guiada de lagartas de *Bombyx mori* submetidas ao estresse térmico**. 2025. 86 p. Trabalho de Conclusão de Curso (Graduação em Ciências Biológicas) – Universidade Estadual de Londrina, Londrina. 2025.

RESUMO

Bombyx mori Linnaeus, 1758 (Lepidoptera: Bombycidae) possui grande relevância para a atividade agrícola da sericicultura no Brasil, devido à sua capacidade de produzir seda durante o estágio larval, finalizando seu casulo ao final do quinto instar de seu ciclo de vida. O estado do Paraná consolidou-se como o maior produtor nacional de casulos de seda, atividade desenvolvida principalmente por pequenos produtores rurais, constituindo uma importante fonte de geração de renda. Entre os vários fatores que afetam o processo produtivo desse organismo, a temperatura se destaca por influenciar o desenvolvimento, expressão gênica e a qualidade da seda. Além disso, sabe-se que elementos transponíveis, podem ser ativados em condições de estresse térmico, contribuindo com mudanças genômicas nesse organismo. O presente trabalho teve como objetivo realizar a análise transcriptômica de *Bombyx mori*, submetida ao estresse térmico, com foco na identificação de elementos móveis, a partir das montagens *de novo* e guiada por referência. As duas montagens resultaram em quantidades distintas de transcritos únicos identificados, sendo 4.205 na montagem *de novo* e 3.529 na guiada. Destes, 70,7% foram comuns às duas montagens, 22,1% exclusivos à montagem *de novo* e 7,2% à montagem guiada. O critério utilizado na seleção dos TEs, baseado no maior escore e viés, obtidos pela hmmerScan, resultou em alterações na classificação geral dos diferentes grupos de TEs identificados.

Palavras-chave: Sericicultura, elementos móveis, transcriptoma, quinto instar.

SOUZA SOARES, Lucas Guthierres. **Identification of transposable elements in *de novo* and reference-guided assemblies of *Bombyx mori* larvae exposed to thermal stress**. 2025. 86 p. Final Dissertation (Biological Sciences Undergraduation) – Londrina State University. Londrina. 2025.

ABSTRACT

Bombyx mori Linnaeus, 1758 (Lepidoptera: Bombycidae) is highly relevant to sericulture in Brazil due to its ability to produce silk during the larval stage, completing its cocoon at the end of the fifth instar of its life cycle. The state of Paraná has established itself as the largest national producer of silk cocoons, an activity carried out mainly by small rural producers, constituting an important source of income. Among the various factors that affect the production process of this organism, temperature stands out for influencing development, gene expression, and silk quality. In addition, it is known that transposable elements can be activated under conditions of thermal stress, contributing to genomic changes in this organism. The objective of this study was to perform transcriptomic analysis of *Bombyx mori* subjected to thermal stress, focusing on the identification of mobile elements from *de novo* and reference-guided assemblies. The two assemblies resulted in different numbers of unique transcripts identified, with 4,205 in the *de novo* assembly and 3,529 in the guided assembly. Of these, 70.7% were common to both assemblies, 22.1% were exclusive to the *de novo* assembly, and 7.2% were exclusive to the reference-guided assembly. The criterion used in the selection of TEs, based on the highest score and bias obtained by hmmerScan, resulted in changes in the overall classification of the different groups of TEs identified.

Keywords: Silk farming, mobile elements, transcriptome, fifth instar.

SUMÁRIO

	Pág.
1. INTRODUÇÃO.....	9
2. REVISÃO BIBLIOGRÁFICA	10
3. MATERIAL E MÉTODOS	21
4. RESULTADOS E DISCUSSÃO	21
5. CONCLUSÕES	38
REFERÊNCIAS	40
6. APÊNDICES	48

1. INTRODUÇÃO

Os insetos da espécie *Bombyx mori* Linnaeus, 1758 pertencem à ordem Lepidoptera, família Bombycidae, e são conhecidos como bichos-da-seda devido à sua capacidade de produzir fibras utilizadas na sericicultura. Esta é uma atividade agrícola milenar, originada na China há cerca de cinco mil anos, sendo considerada uma das práticas agroindustriais mais antigas já desenvolvidas. A seda produzida por *B. mori* é considerada uma fibra ecológica de alta resistência, sendo empregada não apenas na confecção de tecidos, mas também em aplicações nas áreas eletrônica, aeronáutica e médica (BABU, 2013).

De acordo com o prognóstico do Instituto de Desenvolvimento Rural do Paraná (IDR-Paraná), o Brasil é o único produtor de fio de seda em escala comercial no Ocidente, e o sétimo produtor mundial de seda crua. O estado do Paraná consolidou-se como o maior produtor nacional de casulos de seda, sendo responsável por 85% da produção nacional de casulos verdes. Das 3,025 mil toneladas de casulos produzidas nacionalmente, o Paraná responde por uma produção aproximadamente de 2,505 mil toneladas. Destaca-se a participação de pequenos produtores rurais, envolvendo cerca de 153 municípios e 1,3 mil produtores no Estado (IDR-Paraná, 2024).

Contudo, por se tratar de um inseto poiquilotérmico, são altamente sensíveis a flutuações de temperatura, que influenciam significativamente seus processos fisiológicos e desenvolvimento geral. As variações sazonais de temperatura impactam consideravelmente a expressão fenotípica de características genotípicas, afetando parâmetros críticos da produção, como peso do casulo, peso da casca e proporção entre casca e casulo. A exposição a temperaturas extremas

leva ao crescimento retardado, maior suscetibilidade a doenças e redução da qualidade do casulo (THRILEKHA, 2024).

Sendo assim, o estresse térmico tem sido associado à modulação de elementos transponíveis em insetos, tanto em nível de expressão gênica quanto de pequenos RNAs derivados desses elementos, como observado em *Drosophila subobscura* após choque térmico (BODELÓN, 2023).

Os elementos genéticos transponíveis constituem um grupo de sequências de DNA capazes de se mover de uma posição para outra dentro do genoma, podendo se dividir em duas classes principais com base em seus intermediários de transposição: Classe I – retrotransposons e Classe II – transposons de DNA (WELCH; LINQUIST; MALIK, 2021).

Neste contexto, o presente trabalho teve como foco a identificação de elementos transponíveis expressos em lagartas de *B. mori* do quinto instar, submetidas a duas temperaturas, 26°C e 34°C, utilizando as abordagens de montagem transcriptômica *de novo* e guiada, bem como antes e depois da escolha de um protocolo de filtragem dos resultados que determina os alinhamentos com maior escore e menor viés, ou seja, os hits mais confiáveis em acordo com o Dfam.

2. REVISÃO BIBLIOGRÁFICA

2.1 Aspectos biológicos de *B. mori*

Bombyx mori é um inseto holometábolo pertencente à ordem Lepidoptera e à família Bombycidae. Acredita-se que tenha sido domesticado a partir do bicho-da-seda da amoreira selvagem, *B. mandarina*, há cerca de 5.000 a 10.000 anos. *B. mori* e *B. mandarina* são morfologicamente e fisiologicamente semelhantes (Yu-Ping Li, 2009).

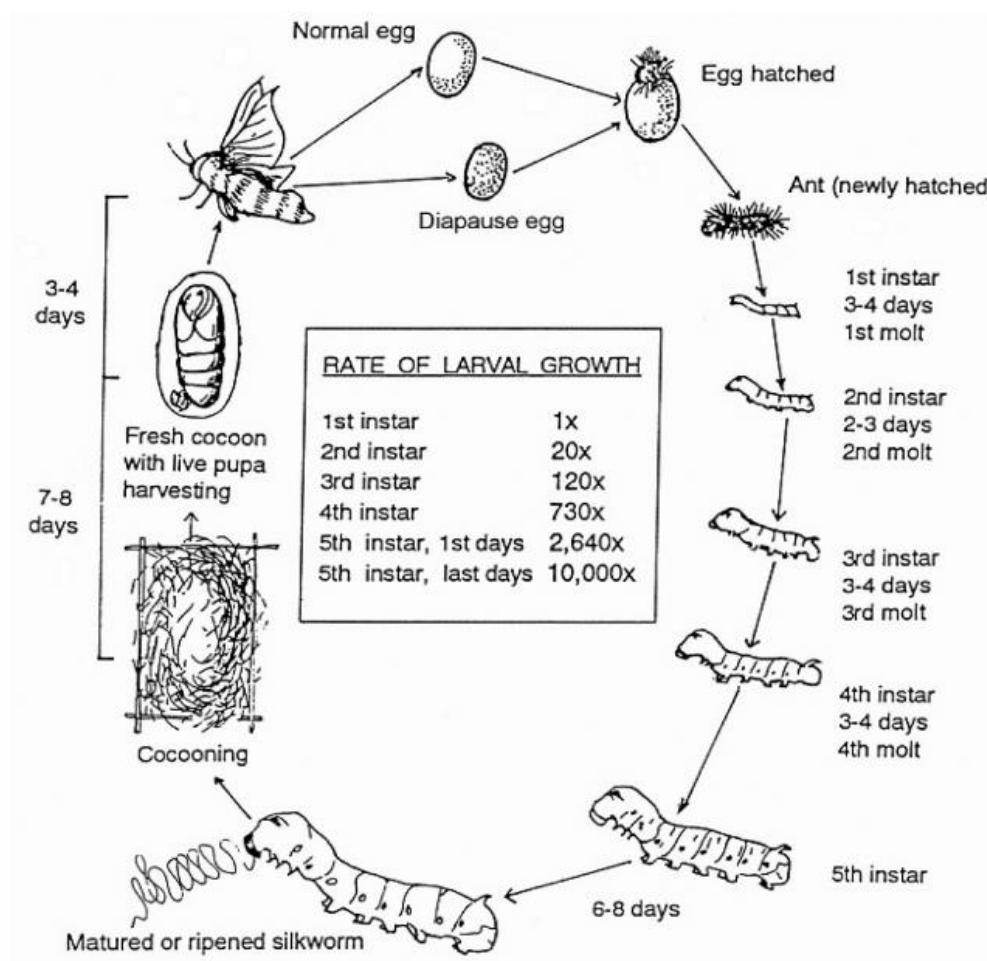
Sua grande importância econômica está ligada a produção de fios de seda através de uma glândula exócrina especializada, responsável pela síntese e secreção das proteínas que formarão o fio de seda do casulo. Ela é um órgão par, constituído por glândulas labiais/salivares, localizado lateralmente ao canal alimentar. Cada glândula é um tubo de epitélio formado por epitélio glandular, com duas fileiras de células que circundam o lúmen. Este órgão secreta duas principais proteínas responsáveis pelos fios de seda: a fibroína e a sericina (TRIVEDY, 2007).

A fibroína é a proteína fibrosa que forma o núcleo do fio de seda, inicialmente secretada como massas esféricas no lúmen da glândula posterior, que se reorganizam e se orientam ao longo da glândula da seda até originar o filamento compacto que constituirá o casulo. Já a sericina é uma outra proteína de matriz que envolve e “cimenta” o fio de fibroína, apresentando uma estrutura agregada que pode formar filmes e fibras orientadas, porém com orientação estrutural instável e reversível, especialmente na presença de água quente (TRIVEDY, 2007).

A respeito de seu ciclo de vida, é compreendido por quatro estágios principais: ovo, larva, pupa (no interior do casulo) e adulto (imago). Os adultos são mariposas domesticadas, ápteras funcionalmente (praticamente não voam) e sem peças bucais funcionais, vivendo apenas para reprodução; fêmeas apresentam abdome mais volumoso e machos antenas mais desenvolvidas. O ciclo inicia-se com os ovos, muito pequenos e de casca rígida, postos no verão/início do outono, que entram em diapausa até a primavera, quando a elevação da temperatura estimula a eclosão. A fase larval é o estágio vegetativo de crescimento, no qual as lagartas, monófagas, alimentam-se de folhas de amoreira (*Morus* spp.). Durante aproximadamente 27 dias, as larvas sofrem quatro mudas, passando por cinco instares; logo após a eclosão são pequenas e pilosas, mas, após a primeira muda,

adquirem tegumento liso. Ao final do quinto ínstar, a larva inicia a formação do casulo, fiando ao seu redor um fio contínuo de seda que pode atingir até cerca de 1,5 km de comprimento, dentro do qual se transforma em pupa. No interior do casulo ocorrem as transformações metamórficas que culminam no surgimento da mariposa adulta, que então emerge, acasala e reinicia o ciclo por meio da postura de novos ovos (GAUTAM et al., 2022).

Figura 1: O ciclo de vida do bicho-da-seda em estágios e duração de cada estágio de desenvolvimento (ANJANA, 2021).



O bicho-da-seda, é um organismo altamente sensível a variações de temperatura, o que pode comprometer em vários processos fisiológicos, como o funcionamento da glândula de seda, balanço hídrico, respiração, absorção de nutrientes e metabolismo em geral. Diferentes fases do ciclo de vida apresentam faixas ótimas distintas: de modo geral, o melhor desempenho produtivo ocorre entre 22 °C e 28 °C, com destaque para 25–26 °C associados a maior crescimento, bom desenvolvimento da glândula de seda e casulos de melhor qualidade, especialmente quando combinados a umidade relativa em torno de 80–85%. Temperaturas mais altas podem acelerar o crescimento e encurtar o período larval, porém, acima de 30 °C passam a prejudicar traços de casulo, aumentar a mortalidade (sobretudo em estágios não alimentares) e comprometer reprodução, reduzindo oviposição, fertilidade e viabilidade embrionária (THRILEKHA, 2024).

Em condições de estresse térmico, há desorganização da síntese proteica, produção de proteínas mal enoveladas, alterações hormonais, metabólicas e imunológicas, além de desequilíbrios na microbiota intestinal, resultando em menor produtividade e maior suscetibilidade a patógenos. Por outro lado, linhagens bivoltinas e híbridos termotolerantes podem responder ao calor com aumento da expressão de proteínas de choque térmico (HSPs) e maior atividade de enzimas antioxidantes, o que confere maior tolerância a temperaturas elevadas (em torno de 35–40 °C), embora muitas vezes à custa de desempenho reprodutivo levemente reduzido (THRILEKHA, 2024).

2.2 Importância da sericicultura para o Estado do Paraná

De acordo com o Boletim de Conjuntura Agropecuária do Paraná, sericicultura é uma atividade agrícola de importante valor comercial que movimenta o

mercado internacional. Segundo o Governo do Estado do Paraná (2021), o Brasil é o único produtor de fio de seda em escala comercial no Ocidente e o sétimo produtor mundial de seda crua. O estado do Paraná consolidou-se como o maior produtor nacional de casulos de seda, sendo responsável por 85% da produção nacional de casulos verdes. Das 3,025 mil toneladas de casulos produzidas nacionalmente, o Paraná responde por uma produção aproximadamente de 2,505 mil toneladas. Destaca-se a participação de pequenos produtores rurais, envolvendo cerca de 153 municípios e 1,3 mil produtores no Estado (IDR-Paraná, 2024).

A seda produzida é um biopolímero com notável resistência mecânica, sendo importante para o mercado com aplicações biomédicas, como estruturas para engenharia de tecidos e liberação de fármacos. A proteína responsável por essa estabilidade é a fibroína. Por conta dela, há estudos envolta da sua estabilidade em condições de temperatura e pressão adversa. Como por exemplo, em Nettey-Oppong et al. (2024), que ao realizar simulações de dinâmica molecular sob diferentes condições de temperatura e pressão, chegaram à conclusão que ambas variáveis impactam significativamente na estabilidade estrutural das conformações solvatadas da seda, evidenciando sua sensibilidade a mudanças ambientais.

2.3 Influência das mudanças de temperatura no desenvolvimento de *B. mori*

Além dessa relevância econômica, *B. mori* possui um genoma bem caracterizado e um ciclo de vida curto e previsível, portanto, tem sido usado bastante em experimentos envolvendo temperaturas variáveis, por se tratar de uma espécie poiquilotérmico. Um desses experimentos, investiga alterações na microbiota intestinal após 48, 96 e 144 horas de tratamento transitório com alta temperatura (TTAT) de 37

°C por 8 horas, chegando à conclusão que *B. mori* em altas temperaturas tende a ter uma menor resistência a patógenos, alterando consideravelmente o equilíbrio de sua flora intestinal, levando a ocorrência frequente de doenças bacterianas, portanto, resultando em menor produção e qualidade de casulos (SUN et al., 2017). Em um segundo trabalho, SUN et al. (2022) observaram que, em nível biomolecular, a atividade enzimática digestiva relacionada nos intestinos é afetada em condição de estresse térmico, conseqüentemente, levando a uma baixa digestão e absorção de nutrientes.

Segundo Muhammad et al. (2025), o bicho-da-seda e a sericicultura enfrentam múltiplos desafios no Antropoceno, incluindo poluição ambiental, mudanças climáticas, doenças, perda de diversidade genética e instabilidades econômicas, o que ameaça a sustentabilidade da produção de seda. Sobre a temperatura, ele destaca que temperaturas mais baixas (abaixo de 20 °C) diminuem a atividade metabólica e prejudicam o crescimento e o desenvolvimento, enquanto temperaturas mais altas aceleram as taxas metabólicas, aumentam a demanda de oxigênio, induzem estresse oxidativo e afetam negativamente os ovários e a reprodução (acima de 42 °C). Além disso, este estresse térmico reduz a taxa de fotossíntese e a qualidade nutricional das folhas de amoreiras, comprometendo sua dieta (MUHAMMAD, 2025).

2.4 Elementos transponíveis

Elementos transponíveis (TEs) são conhecidos por muitos termos como: "transposons", "repetições intercaladas", "elementos genéticos egoístas", "genes saltadores" e "DNA parasitário" por conta da sua capacidade de transposição dentro do genoma, podendo contribuir de forma positiva ou negativa para a evolução

de diversas características do hospedeiro, como gestação interna, memória, coloração e imunidade adaptativa (HAYWARD;GILBERT, 2022). A aptidão dos elementos transponíveis depende de sua capacidade de escapar das defesas do hospedeiro, proliferar e se espalhar para outros indivíduos (HAYWARD; GILBERT, 2022). Assim, por definição, são sequências móveis do genoma que podem se mover ou copiar-se para diferentes regiões, podendo influenciar a regulação gênica e a estabilidade do genoma. Além de seu potencial mutagênico, os TEs atuam como fontes de inovação genômica, influenciando a estrutura e a regulação de genes próximos, bem como a evolução de novos elementos regulatórios (CHUONG; ELGIN; FESCHOTTE, 2017).

Existem, no entanto, estratégias celulares para controlar a atividade dos TEs, que visam inibir a transposição do DNA móvel. Como por exemplo pequenos RNAs não codificantes, que atuam como guias para as proteínas efetoras da família Argonaute, considerada uma característica fundamental das vias de interferência de RNA (RNAi). Essas vias podem suprimir a atividade dos transposons nos níveis transcricional e pós-transcricional (GEBERT; ROSENKRANZ, 2015).

Os retroelementos, ou retrotransposons, são um tipo de TE pertencentes a classe I que realizam sua transposição por meio de um intermediário de RNA, através do mecanismo de “copiar e colar”. Essa classe não precisa de subclasses, nenhum de seus membros cliva ou transfere fitas de DNA no sítio do doador. Em vez disso, o intermediário de RNA é transcrito a partir de uma cópia genômica e, depois, retrotranscrito em DNA por uma transcriptase reversa (RT) codificada pelo próprio TE. Cada ciclo completo de replicação produz uma nova cópia. Consequentemente, os retrotransposons são, com frequência, os maiores contribuintes da fração repetitiva em genomas grandes (WICKER, 2007). Eles

podem ser divididos em 5 ordens com base em suas características de transposição, organização e filogenia da transcriptase reversa: LTR, DIRS, Penelope (PLEs), LINE e SINE (WICKER, 2007).

Há também elementos transponíveis conhecidos como DNA transposons, pertencentes a classe II. Dentro dessa classe, temos duas subclasses.

A subclasse 1 compreende os TEs clássicos de “cortar e colar” da ordem TIR, caracterizados por suas repetições terminais invertidas (TIRs) de comprimento variável. As nove superfamílias conhecidas são diferenciadas pelas sequências das TIRs e pelo tamanho da duplicação do sítio alvo (TSD). A transposição é mediada por uma transposase, enzima que reconhece as TIRs e corta ambas as fitas de DNA em cada extremidade (WICKER, 2007).

A subclasse 2 reúne TEs de DNA que passam por um processo de transposição que envolve replicação sem clivagem de dupla fita, bem diferente daquele da subclasse 1. Esses TEs de “copiar e colar” mudam de posição por meio de replicação envolvendo o deslocamento de apenas uma das fitas. Sua inclusão na classe II reflete o fato de, em comum, não utilizarem intermediário de RNA, mas isso não implica necessariamente ancestralidade comum. Seus representantes são as ordens Helitron e Maverick (WICKER, 2007).

Dentro de uma análise genômica comparativa, em insetos da ordem Diptera os retrotransposons LTR dominam o conteúdo de TEs, em alguns casos contribuindo com cerca de 50% do total, como observado em *Drosophila simulans*. Em Hymenoptera, por outro lado, prevalecem os transposons de DNA, como em *Harpegnathos saltator* (formiga-saltadora), em que esse grupo corresponde a aproximadamente 35,25% do conteúdo de TEs. Já em Hemiptera e Psocodea, os LINEs podem alcançar até 39,3% do conteúdo de elementos transponíveis, como

descrito para *Acyrtosiphon pisum* e *Cimex lectularius*, com exceção do piolho humano *Pediculus humanus*, em que os transposons de DNA contribuem com 44,43% do conteúdo de TEs conhecido. Os SINEs, por sua vez, estão presentes em todas as ordens de insetos, mas em geral contribuem com menos de 10% do genoma em TEs, exceto em alguns táxons específicos, como *Helicoverpa punctigera* (18,48%), *Acyrtosiphon pisum* (27,11%) e *Bombyx mori* (26,38%). Os autores ressaltam ainda que esses valores provavelmente são conservadores, uma vez que grande parte do conteúdo de DNA, LTR, LINE e SINE ainda pode estar “escondida” na fração de TEs não classificados (PETERSEN, 2019).

Em *Bombyx mori*, utilizando de uma montagem genômica de alta qualidade, Kawamoto et al. (2019) realizaram uma triagem mais precisa desses elementos utilizando o RepeatMasker em conjunto com uma biblioteca específica de TEs de *B. mori* e, adicionalmente, com a biblioteca Repbase. Nessa análise destacam-se as ordens SINE, LINE, LTR com o maior número de elementos detectados (Kawamoto, 2019).

Segundo Fujiwara et al. (2005) em *B. mori*, mais de 1.000 cópias de retrotransposons não-LTR, designadas famílias *TRAS* e *SART*, ocorrem entre as repetições teloméricas e se acumulam, o que compensaria a baixa atividade da telomerase por conta da perda das repetições (TTAGG). Ou seja, esses retrotransposons não-LTR, mostram-se eficazes na manutenção telomérica, que é importante para replicação completa e a estabilidade dos cromossomos. Essa associação entre retrotransposons e regiões teloméricas tem sido observada em diversos insetos, como por exemplo no gênero *Drosophila*, que não possui telomerase; em vez disso, três retrotransposons específicos do telômero mantêm as extremidades dos cromossomos, sendo assim, mostrando-se um exemplo notável

de coevolução entre elementos móveis e o genoma hospedeiro (PARDUE; DEBARYSHE, 2011).

2.5 Montagem de transcriptoma: estratégias *de novo* e guiada

Um transcriptoma é o conjunto completo de transcrições em uma célula. O conteúdo de um transcriptoma varia entre os tipos de células e estágios de desenvolvimento. Entender o conteúdo dos transcriptomas e rastrear sua diferenciação espacial e temporal é essencial para interpretar os elementos funcionais do genoma, revelar os constituintes moleculares de células entender processos de desenvolvimento e doenças (ZHU et al., 2016). O sequenciamento de RNA (RNA-seq) é uma tecnologia de criação de perfil de transcriptoma que utiliza sequenciamento de última geração de alto rendimento. A maioria dos dados de RNA-seq é gerada a partir de DNAs complementares (cDNAs) convertidos de RNAs mensageiros (mRNAs) usando a plataforma de sequenciamento de leitura curta Illumina. (WANG, 2009).

A análise de expressão gênica baseada em RNA-seq pode ser conduzida por duas estratégias principais de montagem: *de novo* e guiada por genoma de referência. A montagem *de novo* constrói transcritos a partir das leituras brutas, sem utilizar um genoma de referência, sendo útil para detectar transcritos não anotados ou variantes específicas. Essa estratégia é frequentemente implementada por ferramentas como Trinity (Grabherr et al., 2011) Por outro lado, a montagem guiada utiliza um genoma de referência, oferecendo maior precisão para transcritos previamente identificados, mas com risco de omitir novos elementos. A comparação entre essas abordagens é fundamental em estudos que visam identificar sequências repetitivas ou divergentes, como é o caso dos TEs (SCHATZ

MC 2010; HAAS et al., 2013).

Na abordagem guiada por genoma, a montagem do transcriptoma utiliza uma sequência genômica de referência como base para reconstruir os transcritos. Inicialmente, as leituras de RNA-Seq são alinhadas ao genoma por meio de programas que reconhecem junções de *splicing*, como TopHat2, HISAT2 ou STAR. A partir desses alinhamentos, é construído um grafo de *splicing*, que representa as conexões entre éxons. Os transcritos finais são extraídos por meio da travessia desse grafo, que busca os caminhos mais prováveis entre os éxons mapeados (HOLZER; MARZ, 2019). Ferramentas como TransDecoder ou Cufflinks são exemplos de softwares que executam etapas específicas na reconstrução e anotação dos transcritos, auxiliando na predição de regiões codificantes e na quantificação de isoformas gênicas (HAAS et al., 2013; TRAPNELL et al., 2010).

Por sua vez, na abordagem *de novo*, a montagem do transcriptoma reconstrói sequências de transcritos a partir de reads curtos sem usar um genoma de referência. Ruído transcricional, artefatos de sequenciamento e isoformas de transcrições originárias de *splicing* alternativo também são representados nesses dados (HOLZER; MARZ, 2019). Nesse método, a maior parte dos montadores modernos adota uma abordagem baseada em grafos de De Bruijn, em que as leituras são decompostas em unidades menores chamadas k-mers — subsequências de comprimento k derivadas das leituras. Cada k-mer é representado como um nó (ou vértice) no grafo, e duas k-mers são conectadas por uma aresta se tiverem uma sobreposição de k - 1 nucleotídeos. A travessia desses caminhos no grafo permite reconstruir os transcritos possíveis (HOLZER; MARZ, 2019).

Diferente da montagem genômica, que resulta em um grande grafo interconectado, a montagem *de novo* do transcriptoma gera diversos subgrafos

independentes, cada um representando famílias de isoformas ou genes semelhantes. Contudo, a precisão da montagem depende fortemente da escolha do tamanho do k-mer, que afeta a sensibilidade e a especificidade da montagem (HOLZER; MARZ, 2019).

3. MATERIAL E MÉTODOS

Esse estudo se utilizou de dados desenvolvidos por Lopes (2023), cuja pesquisa envolveu a comparação de transcriptomas de larvas de 5^o instar de *B. mori* mantidas sob as temperaturas de 26 °C (grupo controle) e 34 °C (grupo experimental). Neste caso, foram utilizados os transcriptomas oriundos das montagens *de novo*, realizada com o programa Trinity v2.12.0, e guiada por referência, realizada com o programa StringTie V2.1.1.

O atual estudo contribuiu para a anotação funcional dos elementos transponíveis foi utilizado o banco de dado Dfam (versão curada), que reúne uma coleção aberta de alinhamentos de sequências de DNA de elementos transponíveis, modelos de Markov ocultos (HMMs), sequências de consenso e anotações de genoma (<https://www.dfam.org/home>). A identificação dos elementos transponíveis foi baseada no programa hmmscan, a partir da ferramenta HMMER (<http://hmmer.org/>). A manipulação e organização dos dados foi realizada em planilha eletrônica, bem como por meio de interpretador via linha de comando e linguagem de scripting em Linux.

4. RESULTADOS E DISCUSSÃO

Na montagem *de novo*, usando apenas o limite de 1×10^{-10} para o E-value da sequência completa, observou-se 96.435 anotações para um total de 4.205 acessos Dfam únicos e 4.802 transcritos. Já na montagem guiada, obteve-se 89.292 anotações para um total de 3.529 acessos únicos em 2.850 transcritos. Em ambos os casos, foi realizada a filtragem considerando apenas a anotação de maior escore e o menor viés para as sequências completas de cada transcrito anotado, ou seja, dentre os múltiplos alinhamentos reportados pelo hmmscan para um mesmo transcrito, manteve-se apenas o hit com maior bit score (alinhamento mais significativo) e menor valor de bias (menor influência de composição enviesada da sequência), descartando anotações redundantes ou menos confiáveis. Portanto, isso resultou em 687 acessos Dfam distintos para a montagem *de novo* e 546 para a montagem guiada. A Figura 2 apresenta o total de acessos Dfam comuns e exclusivos para cada montagem antes e depois da filtragem de maior escore e menor viés.

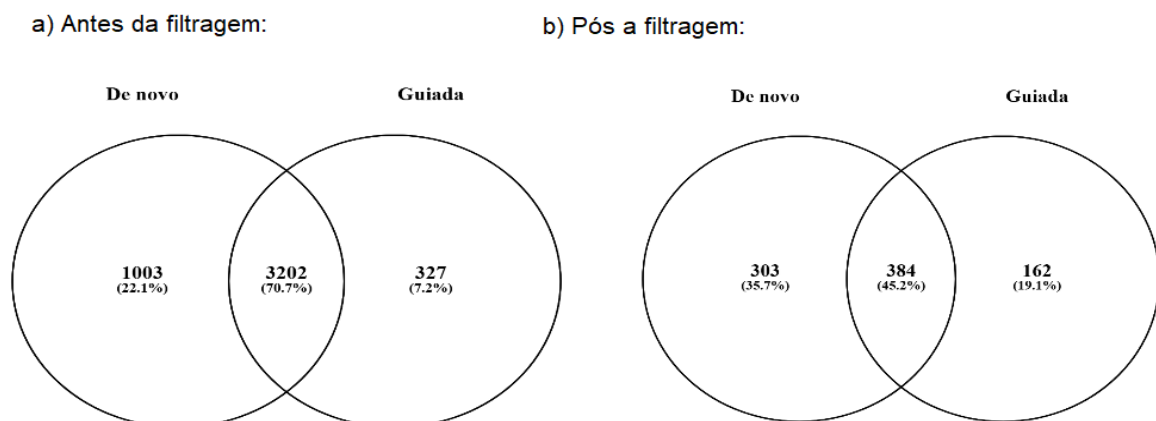
Dos 687 acessos Dfam da montagem *de novo*, 592 foram classificados como elementos transponíveis e 95 como não-TEs ou desconhecidos. Por sua vez, na montagem guiada, dos 546 acessos Dfam, 449 foram identificados como elementos transponíveis e 97 como não-TEs. Proporcionalmente, na montagem *de novo*, antes da filtragem de dados, os TEs mais abundantes foram LTR (33,03%), LINE (28,89%), TIR (10,66%). Já na montagem guiada, os TEs mais abundantes foram LTR (30,99%), LINE (28,73%) e TIR (9,15%) (Figura 3a). Por sua vez, pós a filtragem (Figura 3b), obteve-se na montagem *de novo* a ordem LINE (30,09%), TIR (17,83%) e LTR (14,81%). Já, na guiada, LINE (34,18%), LTR (15,54%) e TIR (13,54%). Essa inversão de ranking demonstra que a filtragem por maior escore e menor viés foi fundamental para remoção de muitos transcritos de

baixa qualidade de alinhamento. A queda acentuada dos LTR sugere que a maioria dos transcritos inicialmente anotados para essa ordem representava cópias redundantes. Em contraste, a ascensão dos LINEs e, notavelmente, do TIR (Classe II), indica que esses elementos são representados por transcritos de alta integridade e menor redundância.

A proeminência do TIR sugere também que os DNA transposons estão ativamente expressos e podem ter um papel funcional importante na resposta do *B. mori* ao estresse térmico. Porém, para que possa concluir a relação de ascensão desses elementos com o estresse térmico, é importante que aja mais revisões utilizando das mesmas metodologias aplicadas no atual trabalho com o *B. mori* em condições de controle para que aja outras comparações.

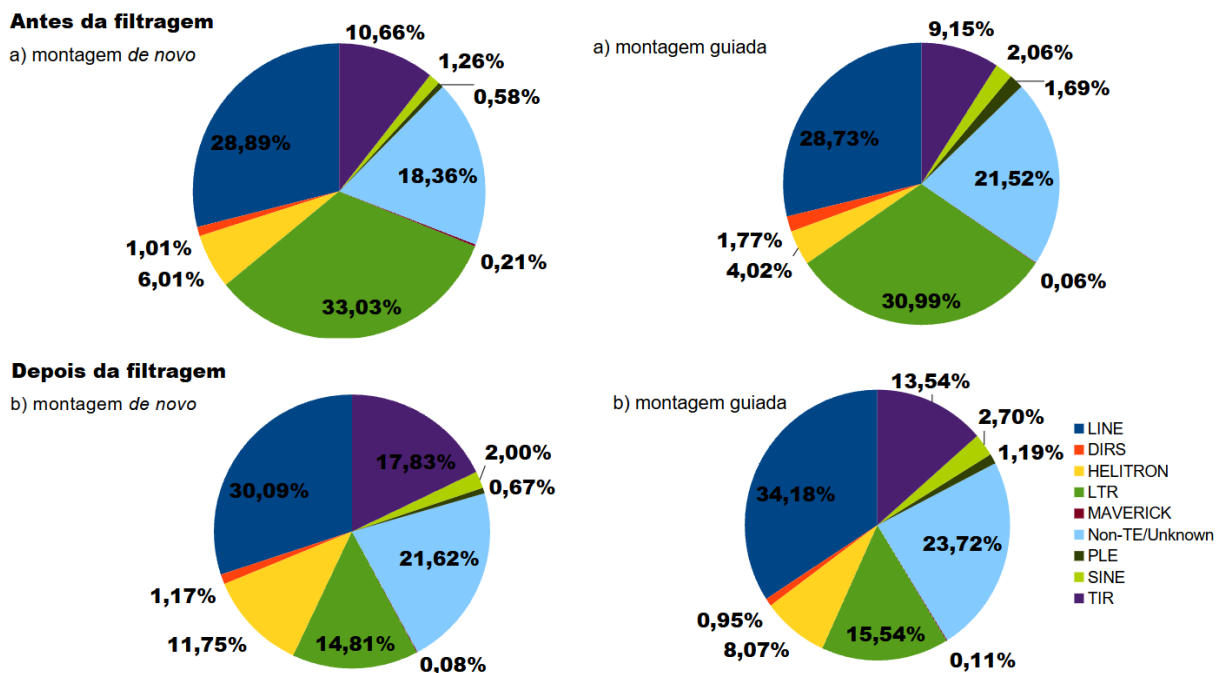
Notou-se um padrão na ordem de abundancia entre ambas as montagens antes da filtragem, porém após o filtro a ordem TIR é mais abundante que o LTR na *de novo* e na guiada essa ordem inverte. O TEs da ordem Maverick, foram os menos identificada nas duas montagens analisadas, independente da filtragem.

Figura 2. Gráfico de Venn. Acessos Dfam comuns e exclusivos das montagens *de novo* e guiada.



Considerando o total de acessos únicos Dfam para as duas montagens combinadas, 55,12% pertencem à Classe I (Retrotransposons), 28,98% pertencem à Classe II (DNA Transposons) e 15,90% foram considerados Não-TEs (rRNA, tRNA, satélites, etc.) ou classificados como acesso desconhecido. A Figura 4 apresenta os 10 acessos Dfam de maior escore e menor viés mais abundantes identificados, considerando todos os transcritos das montagens combinadas antes (Figura 4a) e depois (Figura 4b) da filtragem de dados. Antes da filtragem, os elementos mais abundantes foram MMSAT4 (Não-TE) com 2.218 ocorrências, Acol_Jockey_Ele1 (Não-TE) com 1.590 ocorrências e CR1-Zenon-2_Diul (LINE) com 873 ocorrências. Após a filtragem, observou-se mudanças no ranking, com CR1-Zenon-2_Diul (LINE) apresentando 460 elementos com 382 elementos e Acol_Jockey_Ele1 (Não-TE) com 206 elementos (Figura 3).

Figura 3. Distribuição dos transcritos relacionados aos 849 acessos Dfam únicos de maior escore e menor viés classificados pela ordem ou Non-TE / Unknown da montagem *de novo* e guiada antes (a) e depois (b) do filtro.



A comparação dos resultados da presente análise transcriptômica com estudos genômicos prévios, como o de Osanai-Futahashi et al. (2008), revela um contraste significativo no repertório de elementos transponíveis mais abundantes. Osanai-Futahashi et al. (2008), utilizando DNA genômico de *B. mori* e a ferramenta RepeatMasker, encontraram SINE (12,8%) e não-LTR (13,8%) como as ordens mais abundantes. Em contrapartida, na presente análise transcriptômica sob estresse térmico (34°C), as ordens LINE, TIR e LTR foram as mais numerosas.

O fato de o elemento TIR (Classe II) ter se destacado em nosso transcriptoma, enquanto o SINE foi proeminente no estudo genômico, sugere uma possibilidade de ativação diferencial de TEs sob estresse. A baixa abundância de SINE em nosso transcriptoma, apesar de sua alta representação genômica, reforça

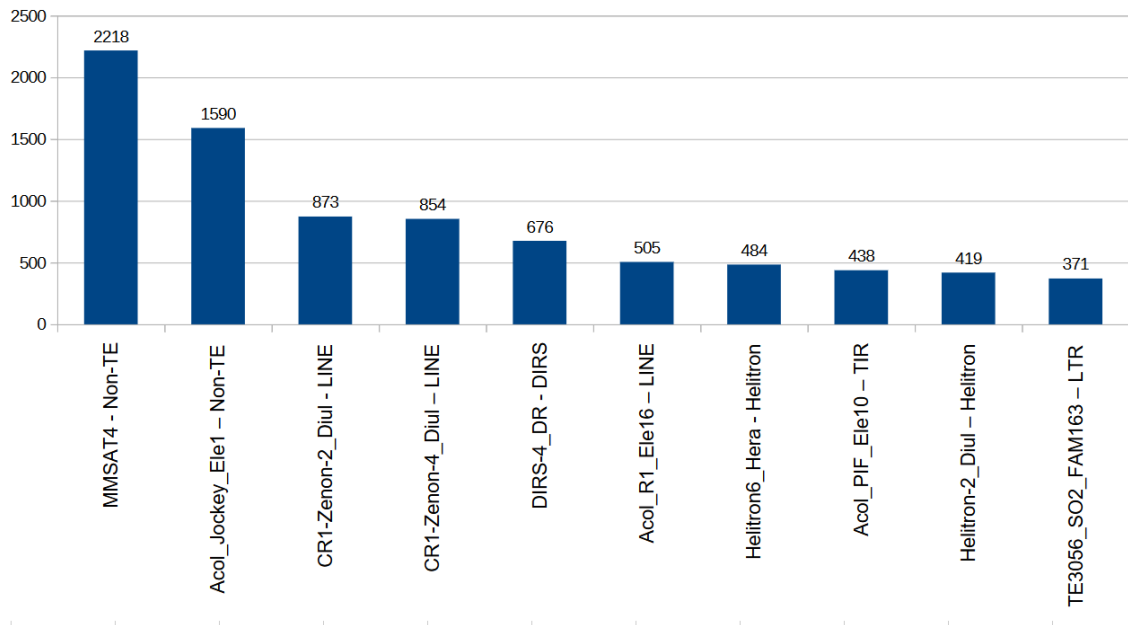
a hipótese de que o TIR pode atuar como um marcador de resposta ao estresse térmico específico no *B. mori*, indicando que esta ordem está sendo mobilizada ou transcrita ativamente como parte da adaptação do inseto ao calor. Porém, é necessário mais estudos utilizando da mesma metodologia adotada no atual trabalho para chegar nessa conclusão, já que é preciso fazer comparações em temperatura controle de *B. mori* para que isso seja de fato comprovado.

Temos exemplos de estudos em *Drosophila melanogaster* e *D. simulans* que mostra a relação entre TEs, estresse e expressão gênica, afirmando ser complexa e altamente contexto-dependente. Embora algumas inserções específicas de TEs já tenham sido associadas à modulação da expressão de genes vizinhos em diferentes tipos de estresse. Do mesmo modo, nem fome nem frio induziram uma ativação global do conteúdo de TEs em *D. melanogaster* e *D. simulans*; em vez disso, observaram-se alterações discretas e família-específicas, com alguns TEs sendo ativados e outros reprimidos, reforçando que a resposta dos TEs ao estresse é dinâmica, dependente do tipo de estressor, da família de TE, do genótipo e da espécie (MERENCIANO, 2025).

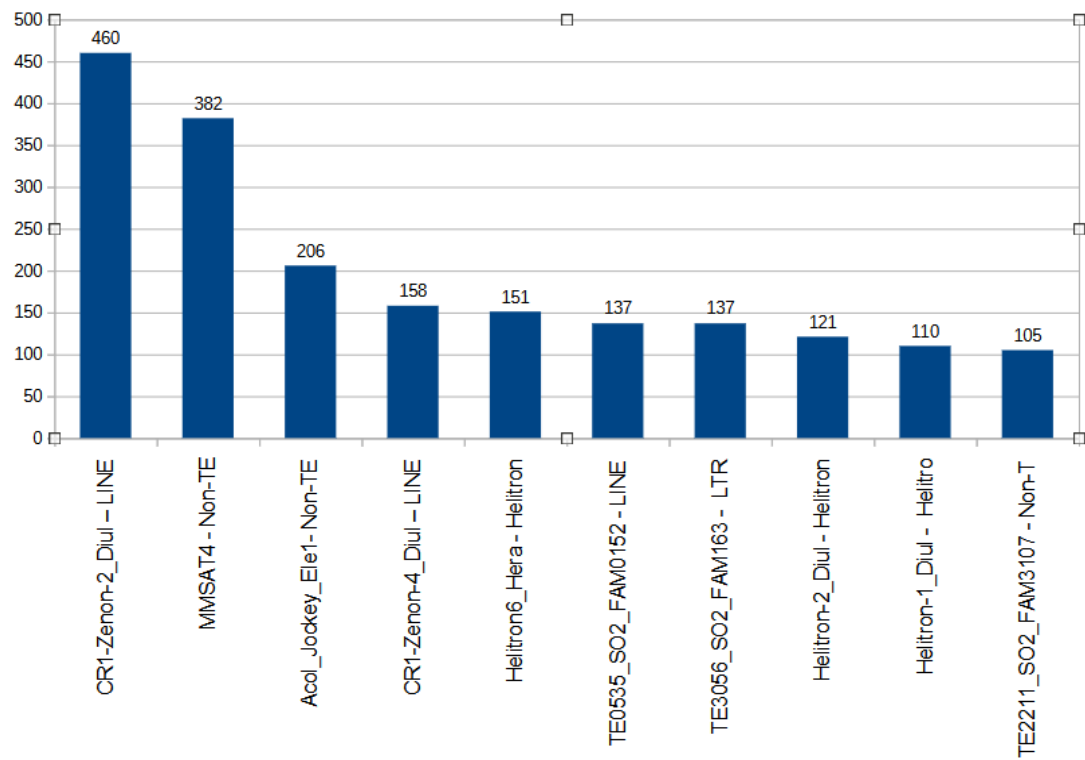
Em um outro estudo, (XU et al. 2013) ampliaram essa caracterização ao construir o banco BmTEdb, por meio da combinação de abordagens de homologia, estrutura e detecção *de novo*, no qual 1.308 famílias de elementos transponíveis foram classificadas. Nesse banco, as ordens LINE, TIR e LTR também representam a maior parte da diversidade de TEs descritos para o genoma de *B. mori* corroborando com o presente trabalho, principalmente na montagem *de novo* após a filtragem na qual observamos essa exata ordem de abundância LINE, TIR e LTR.

Figura 4. Os 10 acessos Dfam identificados mais abundantes somados de ambas montagens antes da filtragem (a) e depois da filtragem (b).

a) Antes da filtragem:



b) Depois da filtragem:



A análise comparativa dos elementos mais abundantes antes e depois da filtragem mostra uma inversão na composição do ranking. O elemento CR1-Zenon-2_Diul (LINE), que ocupava a terceira posição antes da filtragem, ascendeu à primeira colocação após o processo de seleção por maior escore e menor viés, enquanto os Não-TEs MMSAT4 e Acol_Jockey_Ele1, anteriormente nas duas primeiras posições, foram deslocados para a segunda e terceira posições, respectivamente.

Após a filtragem três elementos foram excluídos do Top 10: DIRS-4_DR (DIRS, 5ª posição), Acol_RL_Ele16 (LINE, 6ª posição) e Acol_PLE_Ele10 (PLE, 8ª posição), sendo substituídos por TE0535_SO2_FAM0152 (LINE), Helitron-1_Diul (Helitron) e TE2211_SO2_FAM3107 (Non-TE). Entre os elementos que permaneceram, destaca-se o comportamento diferencial nas mudanças de posição: CR1-Zenon-2_Diul (LINE) foi o único elemento que ascendeu, seguido por TE3056_SO2_FAM163 (LTR), que subiu três posições, e Helitron6_Hera (Helitron), que avançou duas posições. Isso demonstra que esses elementos que ascenderam principalmente apresentaram alinhamentos mais robustos e menos enviesados, sendo, portanto, candidatos mais confiáveis a TEs efetivamente expressos sob estresse térmico.

Antes da filtragem, o Top 10 incluía elementos LINE, DIRS, Helitron, LTR e Não-TEs; após a filtragem, permaneceram LINE, Helitron, LTR e Não-TEs, com a ausência notável de DIRS.

A Figura 5 apresenta os 10 acessos Dfam mais abundantes em ambas as montagens, antes (Figura 5a) e após (Figura 5b) a filtragem. Antes da filtragem, observou-se captura diferencial entre as montagens, com alguns elementos sendo preferencialmente detectados na montagem *de novo* e outros na

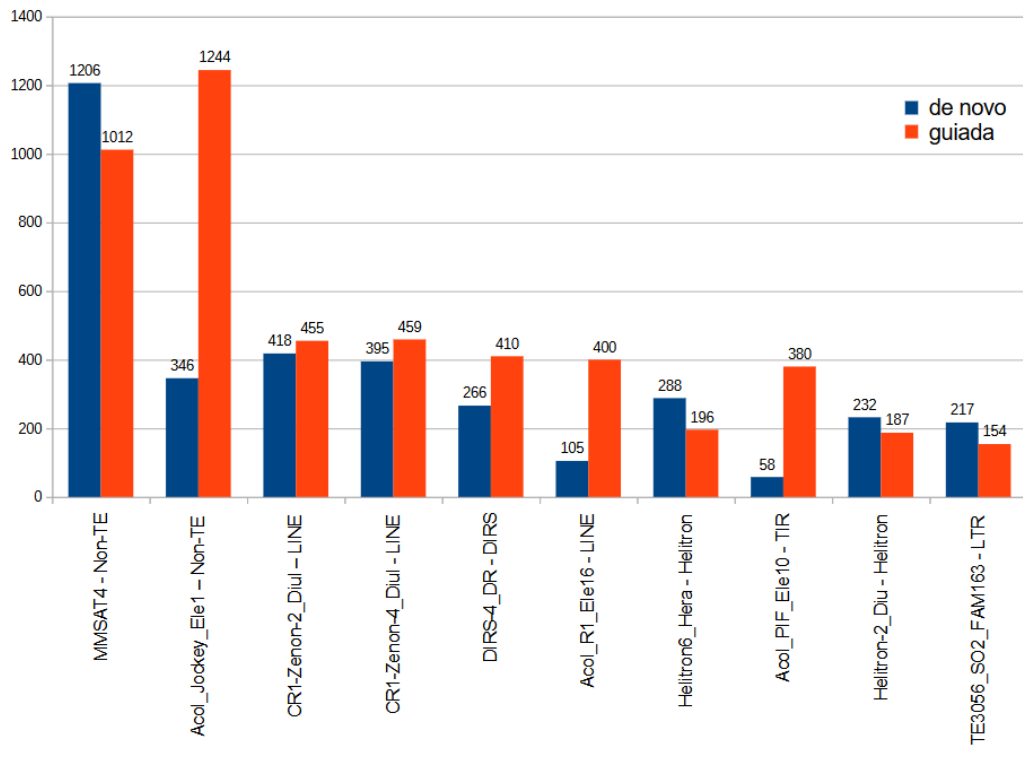
guiada. Como os elementos Acol_Jockey_Ele1 (Non-TE), CR1-Zenon-2_Diul (LINE) e CR1-Zenon-4_Diul (LINE) foram mais abundantes na montagem guiada. Em contrapartida, elementos como MMSAT4 (Non-TE), Helitron6_Hera, Helitron-2_Diul e TE3056_SO2_FAM163 (LTR) foram mais abundantes na montagem *de novo*. Portanto, dos 10 elementos mais abundantes antes da filtragem, 6 foram preferencialmente capturados pela montagem guiada (60%), enquanto 4 foram mais abundantes na *de novo* (40%) (Figura 5a).

Depois da filtragem (Figura 5b), revela-se alterações, sendo uma delas o elemento CR1-Zenon-2_Diul (LINE), que dessa vez foi mais abundante na montagem *de novo*, revelando que a seleção de maior escore e menor viés, removeu uma maior parte dos seus transcritos dentro da montagem guiada. Em seguida, temos o MMSAT4 (Non-TE) e o Helitron6_Hera (Helitron) como ainda destaques da montagem *de novo*. Agora, na montagem guiada pós filtragem o Acol_Jockey_Ele1 (Non-TE) permaneceu como destaque, apesar de ter perdido uma posição e passado a ser o terceiro do ranking total, seguido pelo CR1-Zenon-4_Diul (LINE) e agora o TE0535_SO2_FAM0152 (LINE) como os mais abundantes em comparação com a *de novo*.

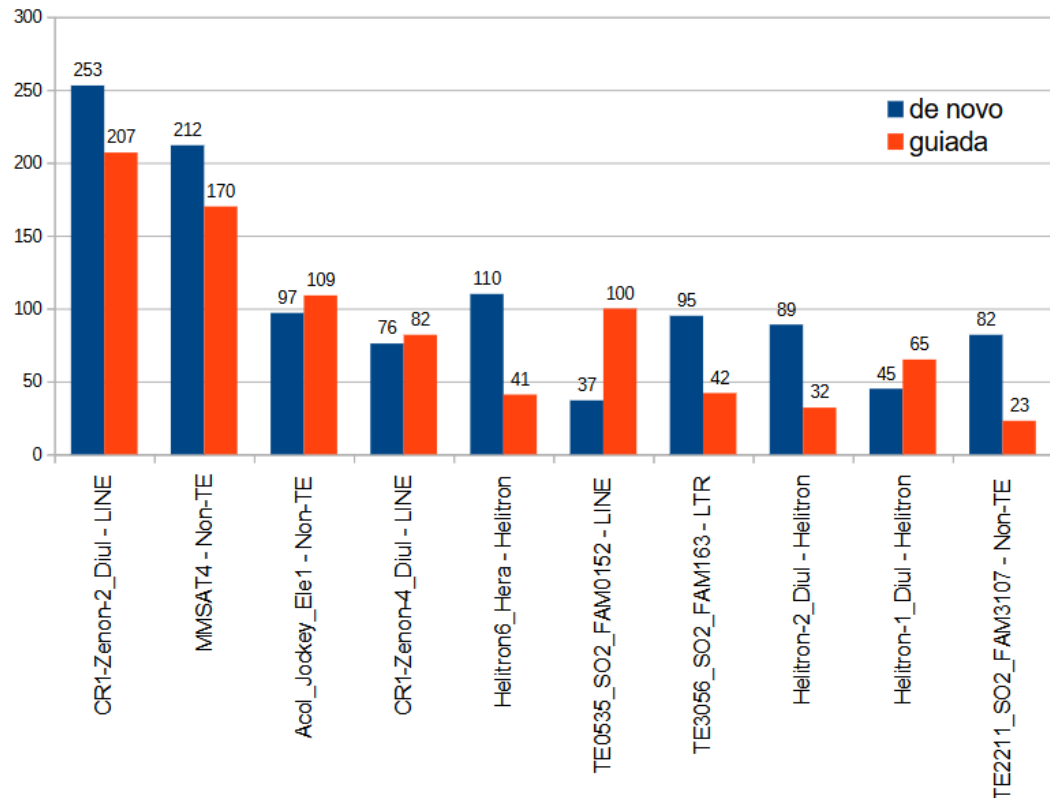
Após filtragem, 60% dos elementos do top 10 são mais abundantes na montagem *de novo* e 40% são mais abundantes na montagem guiada. No total, 525 elementos (61,8%) dos 849 acessos Dfam únicos são mais abundantes na montagem *de novo*, 53 elementos (6,2%) estão igualmente em ambas as montagens e 272 (32%) apenas são mais abundantes na montagem guiada.

Figura 5. Os 10 acessos Dfam mais abundantes nas diferentes estratégias de montagens antes da filtragem (a) e pós a filtragem (b).

a) Antes da filtragem:



b) Depois da filtragem:



Destaca-se também os elementos que saíram do ranking depois da

filtragem, como o DIRS4_DR – DIRS (22°), Acol_R1_Ele16 – LINE (58°) e Acol_PIF_Ele10 – TIR (115°). Ou seja, uma considerável parte dos transcritos desses acessos Dfam foram removidos com a seleção imposta. Entretanto, temos outros três acessos que entraram neste ranking: TE0535_SO2_FAM0152 – LINE, Helitron-1_Diul – Helitron e TE2211_SO2_FAM3107 – Non-TE que antes da filtragem estavam nas posições 24°, 30° e 36°.

A figura 6 (a e b) mostra os 10 acessos Dfam mais abundantes exclusivos de cada montagem.

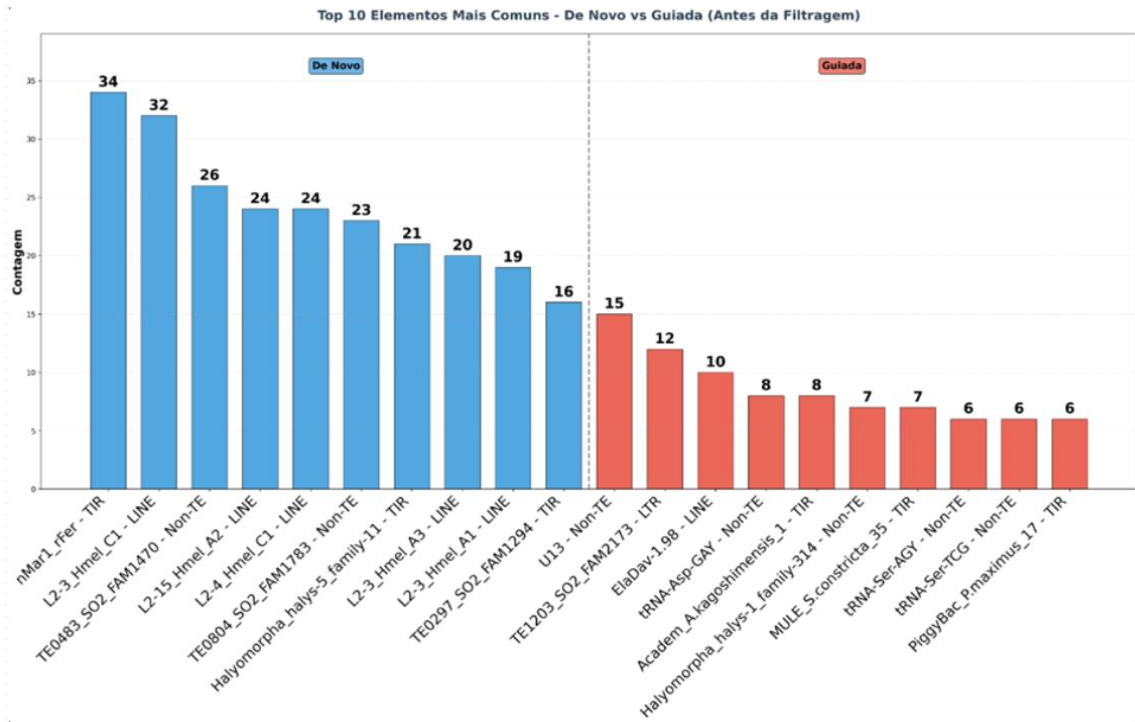
Portanto, antes da filtragem tivemos ao todo 17% de acessos que eram exclusivos em uma das montagens, já pós filtragem essa variável aumentou para aproximadamente 54%. Isso revela a grande seleção imposta quanto a contabilidade de transcritos, considerando o maior escore e menor viés.

Os acessos exclusivos mais abundantes na montagem *de novo* foram nMar1_rFer (TIR) com 34 elementos, L2-3_Hmel_C1 (LINE) com 32 elementos e TE0483_SO2_FAM1470 (Não-TE) com 26 elementos. Observa-se uma predominância de elementos TIR e LINE entre os acessos exclusivos mais abundantes, com alguns Não-TEs também presentes.

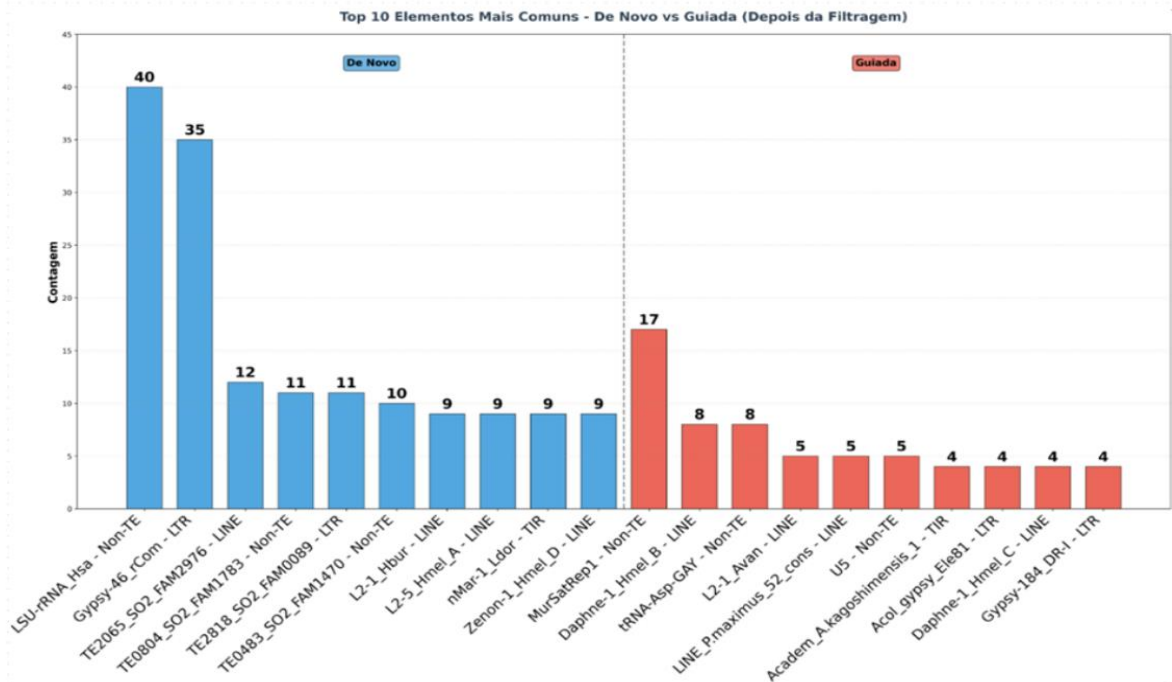
Após a filtragem (Figura 6b), houve uma redução no número de alguns elementos exclusivos. Porém, destaca-se LSU-rRNA_Hsa (Não-TE) com 40 elementos, Gypsy-46_rCom (LTR) com 35 elementos e TE2065_SO2_FAM2976 (RTE / LINE) com 12 elementos. Aqui também se observa uma maior predominância das ordens LINE e TIR.

Figura 6. Os 10 acessos Dfam mais abundantes identificados exclusivamente na montagem *de novo* e na guiada antes (a) e depois (b) da filtragem.

a) Antes da filtragem



b) Depois da filtragem



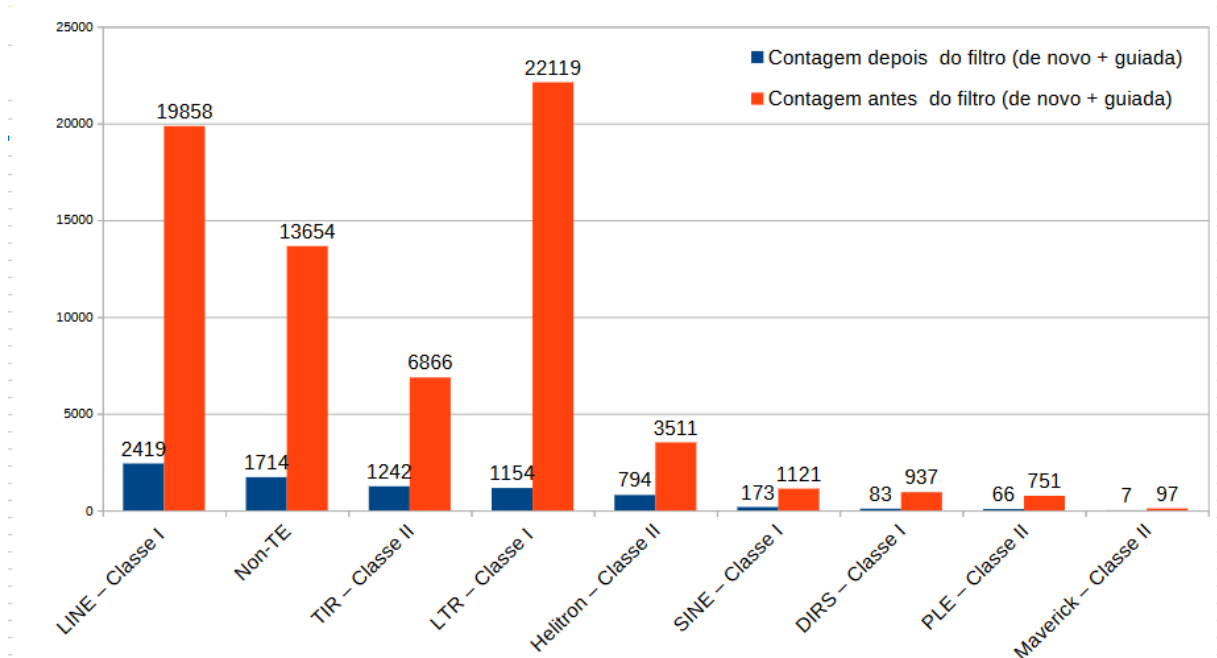
A Figura 7 apresenta a quantidade total de elementos identificados

para cada classe e ordem, somando as contribuições das montagens *de novo* e guiada, tanto antes quanto depois da aplicação do filtro de maior escore e menor viés.

Nota-se uma inversão no ranking de abundância: enquanto os LTR eram os mais abundantes antes da filtragem (32,1%), eles caíram para a quarta posição após a filtragem (15,1%). Por outro lado, os LINEs, que eram a segunda ordem mais abundante antes da filtragem (28,8%), assumiram a primeira posição após a filtragem (31,6%).

No artigo de Petersen et al., (2019), analisa-se dezenas de genomas de artrópodes, incluindo o de *B. mori*, a partir de uma pipeline padronizada que utiliza o RepeatModeler e o RepeatMasker. Nesse panorama comparativo, apesar do presente trabalho ser uma análise de transcriptoma, observar-se, de forma semelhante, uma maior presença de TEs da Classe I, bem representados principalmente pela ordem SINE, LINE e LTR reforçando a concordância com os resultados.

Figura 7. Elementos transponíveis de Classe I e II distribuídos em suas respectivas ordens, bem como os não-TEs mais abundantes encontrados com as abordagens combinadas das montagens *de novo* e guiada, antes e depois do filtro.



Por último, obteve-se a análise que permite avaliar as diferenças na capacidade de detecção de cada método de montagem e o impacto diferencial da filtragem sobre cada abordagem (Figura 8a e 8b).

Observa-se que se manteve o mesmo padrão de contabilidade de transcritos em ambas as abordagens, e a proporção caiu de forma que se manteve a mesma regularidade pós a filtragem.

Antes da filtragem, ambas as montagens apresentaram perfil similar de abundância relativa. Na montagem *de novo*, as ordens mais abundantes foram LTR (12.300 transcritos, 33,0%), LINE (10.757, 28,9%), Non-TE (6.836, 18,4%) e TIR (3.968, 10,7%), enquanto na guiada foram LTR (9.819, 31,0%), LINE (9.101, 28,7%), Non-TE (6.818, 21,5%) e TIR (2.898, 9,1%). Diferenças notáveis incluíram maior proporção de Helitrons na *de novo* (2.236, 6,0% vs 1.275, 4,0% na guiada),

enquanto a guiada detectou mais PLE (534, 1,7% vs 217, 0,6%) e DIRS (562, 1,8% vs 375, 1,0%).

Em termos absolutos, a montagem *de novo* identificou 37.234 transcritos, enquanto a guiada detectou 31.680, evidenciando 17,5% maior capacidade de detecção da abordagem *de novo*. Segundo LU (2013), observou-se que ambas as abordagens de montagem apresentam vantagens e limitações. No caso da montagem guiada por genoma, os autores destacam que ela tende a ser particularmente sensível a transcritos pouco expressos, uma vez que mesmo um número reduzido de leituras pode ser mapeado ao genoma de referência. Além disso, o alinhamento prévio das leituras ao genoma contribui para atenuar erros de sequenciamento que complicariam o grafo de montagem, e leituras alinhadas em regiões próximas podem ser agrupadas em éxons, sendo que os sítios de splicing auxiliam na delimitação mais precisa de seus limites (LU et al., 2013).

Por outro lado, a montagem *de novo* também apresenta méritos importantes. Com cobertura de leitura suficiente, ela é capaz de reconstruir de forma eficiente um transcriptoma praticamente completo. Além disso, por não depender do alinhamento por splicing, evita que erros de mapeamento e artefatos introduzidos nessa etapa levem à inferência de isoformas falso-positivas, problema mais comum nas abordagens guiadas. O fato de ser independente do genoma de referência permite ainda a detecção de novos transcritos não anotados, como vistos no presente trabalho, seja por ausência de genes na referência, por variantes estruturais ou outros motivos, bem como a identificação de transcritos com íntrons longos e de eventos particulares, como trans-splicing e rearranjos cromossômicos (LU, 2013).

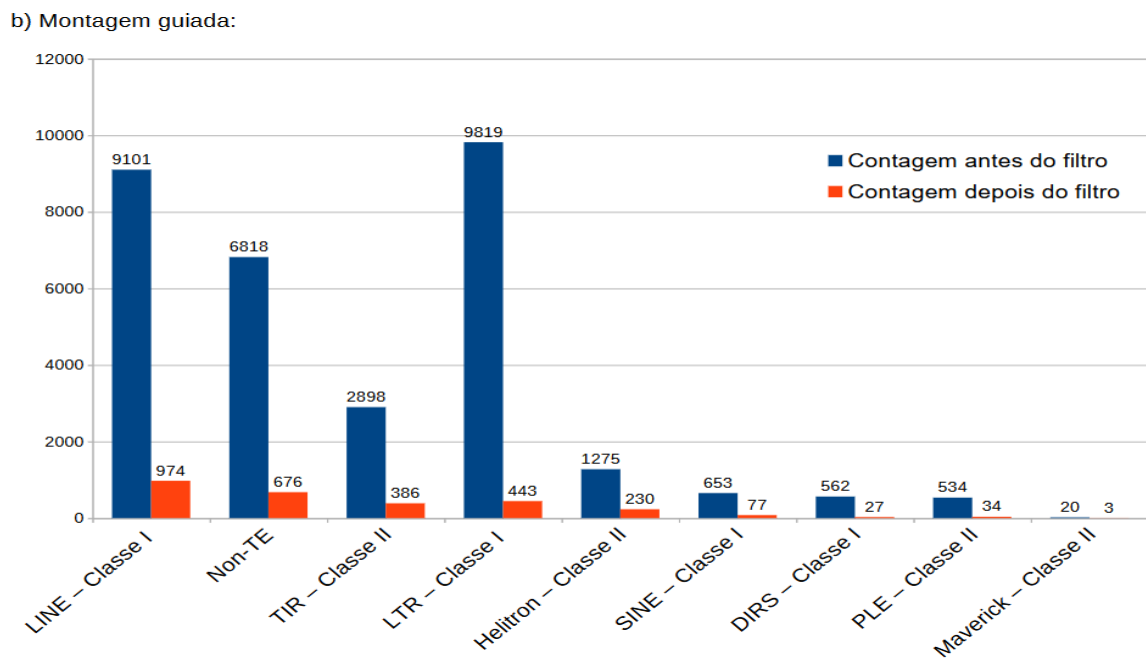
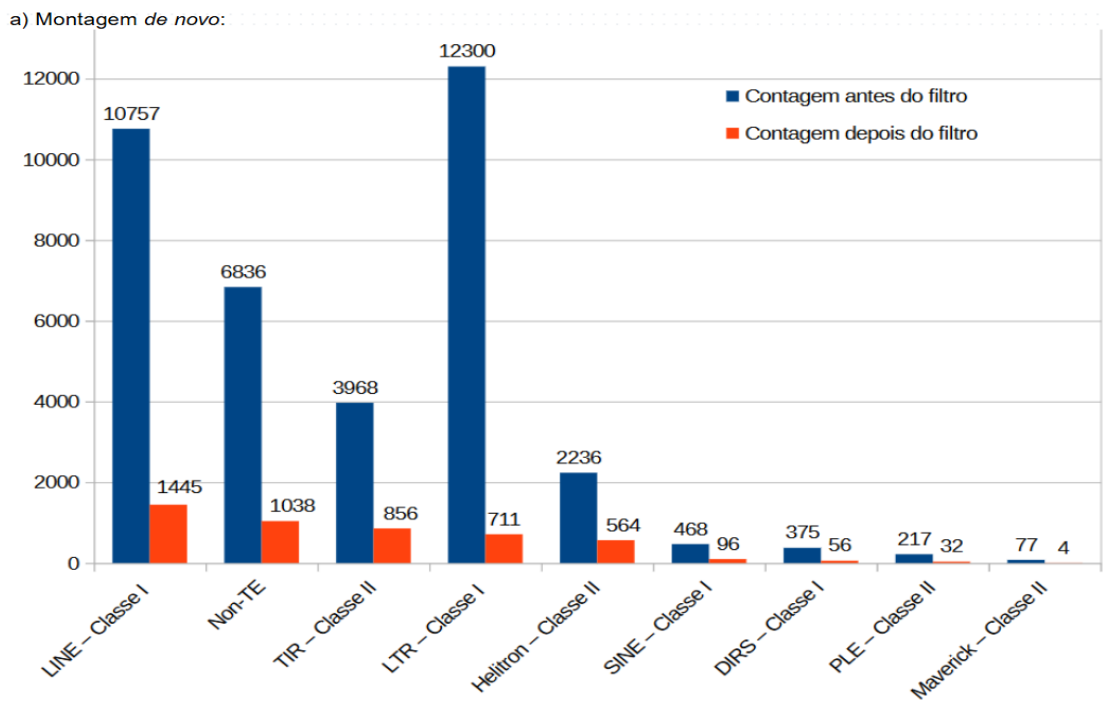
Portanto, como visto em LU (2013), é necessário combinar a

montagem guiada por genoma e a montagem *de novo* para obter um transcriptoma mais abrangente. Essa abordagem combinada pode aproveitar as vantagens de cada estratégia e complementar as limitações de cada uma.

Destaca-se também que após a filtragem por maior escore e menor viés, observou-se inversão no ranking de abundância em ambas as montagens. Na *de novo*, LINE assumiu a primeira posição (1.445, 30,1%), seguido por Non-TE (1.038, 21,6%), TIR (856, 17,8%), LTR (711, 14,8%) e Helitron (564, 11,7%). Na guiada, o ranking foi LINE (974, 34,2%), Non-TE (676, 23,7%), LTR (443, 15,5%), TIR (386, 13,5%) e Helitron (230, 8,1%). Ou seja, destaca-se a inversão do ranking entre a ordem TIR e LTR.

Elementos da Classe II (DNA transposons) ganharam destaque proporcional: TIR aumentou de 10,7% para 17,8% na *de novo* e de 9,1% para 13,5% na guiada, enquanto Helitron aumentou de 6,0% para 11,7% na *de novo* e de 4,0% para 8,1% na guiada. Em contrapartida, LTR sofreu queda drástica de 33,0% para 14,8% na *de novo* e de 31,0% para 15,5% na guiada.

Figura 8. TEs mais abundantes distribuídos em classe (I e II) e suas respectivas ordens, antes e depois do filtro das montagens *de novo* (a) e guiada (b).



Estudos como o Zhang et al. (2023), em insetos da família Noctuidae, que, assim como *B. mori*, também pertencem à ordem Lepidoptera, verificaram que LINEs e os transposons de DNA contribuíram mais para o tamanho do genoma, ao contrário dos SINEs, que não apresentaram correlação significativa em Noctuidae, embora algumas subfamílias, como SINE/5S e SINE/B2, exibam expansões recentes e específicas de linhagem.

Outro estudo relevante, é o Wu & Lu (2019) baseado em uma análise comparativa em larga escala de 14 espécies artrópodes, inclusive o *B. mori* que afirma que entre as espécies *Melitaea cinxia*, *Papilio xuthus* e *Spodoptera frugiperda* da ordem Lepidoptera, *B. mori* apresenta as subfamílias de TEs mais abundantes e a maior carga de TEs. Mais da metade de seu genoma é ocupada por TEs. Afirma também que LTRs, TIRs e LINEs se encontram na faixa de 1–10%, especialmente os LTRs, que estão quase todos localizados na faixa de 1–5%.

Tais resultados corroboram com o presente trabalho, mesmo este tratando-se de uma abordagem transcriptômica, e não genômica, o predomínio de retrotransposons de Classe I como os LTRs, LINE e da Classe II TIR são significativos, reforçando a importância desses elementos em Lepidoptera.

5. CONCLUSÕES

As duas montagens resultaram em quantidades distintas de transcritos totais. A montagem *de novo* apresentou aproximadamente 68,5% mais transcritos e 19,16% mais acessos Dfam únicos em relação à montagem guiada, sendo que 70,7% dos acessos Dfam foram compartilhados entre ambas, antes da aplicação do filtro. Também foram observados acessos Dfam exclusivos, com a

montagem *de novo* apresentando 1.003 acessos e a guiada apenas 327. Quanto as ordens de TEs identificadas e contadas, proporcionalmente, as duas montagens mantiveram-se semelhante, o que demonstra uma padronização e confiabilidade dos dados, uma vez que tal variação entre abordagens é esperada.

Após a filtragem, a abordagem *de novo* também se destacou na detecção de elementos transponíveis, cerca de 31,8% a mais de TEs em comparação com a guiada. Em relação aos acessos Dfam, 45,2% foram compartilhados entre as duas abordagens. Quanto aos acessos exclusivos, houve uma diferença menor entre as montagens, sendo 303 acessos Dfam apenas na montagem *de novo* e 162 na montagem guiada.

Observou-se ainda uma inversão no ranking de abundância pós-filtragem, na qual, em ambas as abordagens, a ordem LINE passou a ser mais abundante que a ordem LTR, indicando que o processo de seleção imposto pelo filtro atuou de forma semelhante nas duas montagens e manteve uma proporção consistente entre as ordens.

Contudo, visto que para um melhor aproveitamento dos dados de TEs, é de maior proveito combinar ambas as montagens para obter um transcriptoma mais abrangente. Essa abordagem combinada pode aproveitar as vantagens de cada estratégia e complementar as limitações de cada uma.

Por fim, espera-se que este trabalho contribua de forma significativa para comunidade científica, por se tratar da compreensão transcriptômica e adaptativa de *B. mori* em uma situação adversa de temperatura, além de fornecer como base a identificação e comparação desses elementos transponíveis que podem ser utilizados em pesquisas futuras.

REFERÊNCIAS

Anjana Baburaj & Subrata Das (2021): Dye Fed Silkworms to Produce Naturally Coloured Silk Cocoons, *Journal of Natural Fibers*, DOI: 10.1080/15440478.2021.1889430

BABU, K. M. *Silk: Processing, Properties and Applications*. Cambridge: Woodhead Publishing, 2013. ISBN 978-1-78242-155-9.

Bourque, G., Burns, KH, Gehring, M. et al. Dez coisas que você deve saber sobre elementos transponíveis. *Genome Biol* 19, 199 (2018). <https://doi.org/10.1186/s13059-018-1577-z>

Chénais B. Transposable Elements and Human Diseases: Mechanisms and Implication in the Response to Environmental Pollutants. *Int J Mol Sci*. 2022 Feb 25;23(5):2551. doi: 10.3390/ijms23052551. PMID: 35269693; PMCID: PMC8910135.

CHUONG, E. B.; ELGIN, S. C. R.; FESCHOTTE, C. Regulatory activities of transposable elements: from conflicts to benefits. *Nature Reviews Genetics*, v. 18, n.2, p. 71-86, 2017. DOI: <https://doi.org/10.1038/nrg.2016.139>

DFAM. Database of repetitive DNA families. Disponível em: <https://www.dfam.org/home>. Acesso em: 7 jul. 2025.

DFAM. Dfam Families Release. Disponível em:
<https://www.dfam.org/releases/current/families/>. Acesso em: 7 jul. 2025.

FUJIWARA, H., OSANAI, M., MATSUMOTO, T. et al. Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, *Bombyx mori*. *Chromosome Res* 13, 455–467 (2005). <https://doi.org/10.1007/s10577-005-0990-9>

GAUTAM, M. et al. A review on silkworm (*Bombyx mori* Linn.) – an economic important insect. *Biological Forum – An International Journal*, v. 14, n. 4a, p. 482–491, 2022.

Gebert, D., Rosenkranz, D. RNA-based regulation of transposon expression, 2015. *WIREs RNA* 2015, 6:687–708. doi: 10.1002/wrna.1310

GOVERNO DO ESTADO DO PARANÁ. Prognóstico agropecuário: sericicultura – v. 13, n. 40. Curitiba: Secretaria de Estado da Agricultura e do Abastecimento do Paraná, 2021. Disponível em:
https://www.agricultura.pr.gov.br/sites/default/arquivos_restritos/files/documento/2022-01/vol_13_n_40_2021_prognostico_agropecuaria_sericicultura.pdf. Acesso em: 10 maio 2025.

HAAS, B., PAPANICOLAOU, A., YASSOUR, M. et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8, 1494–1512 (2013). <https://doi.org/10.1038/nprot.2013.084>

HAYWARD A, GILBERT C. Transposable elements. *Curr Biol.* 2022 Sep 12;32(17):R904-R909. doi: 10.1016/j.cub.2022.07.044. PMID: 36099891.

Hong-En Xu, Hua-Hao Zhang, Tian Xia, Min-Jin Han, Yi-Hong Shen, Ze Zhang, BmTEdb: um banco de dados coletivo de elementos transponíveis no genoma do bicho-da-seda, Banco de dados, Volume 2013, 2013, bat055, <https://doi.org/10.1093/database/bat055>

KAWAMOTO, M. et al. High-quality genome assembly of the silkworm, *Bombyx mori*. *Insect Biochemistry and Molecular Biology*, v. 107, p. 53–62, 2019. DOI: 10.1016/j.ibmb.2019.02.002.

Kunz RI, Brancalhão RM, Ribeiro LF, Natali MR. Silkworm Sericin: Properties and Biomedical Applications. *Biomed Res Int.* 2016;2016:8175701. doi: 10.1155/2016/8175701. Epub 2016 Nov 14. PMID: 27965981; PMCID: PMC5124675.

LI, Y.-P. et al. Mitochondrial genome nucleotide substitution pattern between domesticated silkworm, *Bombyx mori*, and its wild ancestors, Chinese *Bombyx*

mandarina and Japanese Bombyx mandarina. *Genetics and Molecular Biology*, v. 33, n. 1, p. 186–189, 2009.

LISCHER, H.E.L., SHIMIZU, K.K. Reference-guided *de novo* assembly approach improves genome reconstruction for related species. *BMC Bioinformatics* 18, 474 (2017). DOI: <https://doi.org/10.1186/s12859-017-1911-6>

LOPES, THAYNÁ BISSON FERRAZ. Influência da variação de temperatura na expressão gênica de *Bombyx mori* Linnaeus, 1758 (Lepidoptera: Bombycidae). 2023. 88 f. Tese (Doutorado em Genética e Biologia Molecular) – Universidade Estadual de Londrina, Londrina, 2023.

Lu B, Zeng Z, Shi T. Comparative study of *de novo* assembly and genome-guided assembly strategies for transcriptome reconstruction based on RNA-Seq. *Sci China Life Sci.* 2013 Feb;56(2):143-55. doi: 10.1007/s11427-013-4442-z.

MARTIN HÖLZER, MANJA MARZ, *De novo* transcriptome assembly: A comprehensive cross-species comparison of short-read RNA-Seq assemblers, *GigaScience*, Volume 8, Issue 5, May 2019, giz039, <https://doi.org/10.1093/gigascience/giz039>

Merenciano, M., Oliveira, DS, Salces-Ortiz, J. et al. Expressão de genes e elementos transponíveis em resposta ao estresse em populações temperadas e tropicais de *Drosophila* . *Mobile DNA* 16 , 35 (2025). <https://doi.org/10.1186/s13100-025-00372-x>

MONDAL, M.; TRIVEDI, K.; KUMAR, S. N. The silk proteins, sericin and fibroin in silkworm, *Bombyx mori* Linn. – a review. *Caspian Journal of Environmental Sciences*, v. 5, n. 2, p. 63–76, 2007.

MUHAMMAD, A.; IQBAL, J.; SUN, C.; SHAO, Y. The Anthropocene presents significant challenges for silkworms and sericulture. *Journal of Resource Insects*, v. 97, art. 100002, 2025. DOI: 10.1016/j.jori.2025.100002.

Nettey-Oppong, E.E.; Muhammad, R.; Ali, A.; Jeong, H.-W.; Seok, Y.-S.; Kim, S.-W.; Choi, S.H. The Impact of Temperature and Pressure on the Structural Stability of Solvated Solid-State Conformations of *Bombyx mori* Silk Fibroins: Insights from Molecular Dynamics Simulations. *Materials* 2024, 17, 5686. <https://doi.org/10.3390/ma17235686>

Osanai-Futahashi M, Suetsugu Y, Mita K, Fujiwara H. Genome-wide screening and characterization of transposable elements and their distribution analysis in the silkworm, *Bombyx mori*. *Insect Biochem Mol Biol*. 2008 Dec;38(12):1046-57. doi: 10.1016/j.ibmb.2008.05.012. PMID: 19280695.

Petersen, M., Armisén, D., Gibbs, R.A. et al. Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects. *BMC Evol Biol* 19, 11 (2019). <https://doi.org/10.1186/s12862-018-1324-9>

REVISTA DE CIÊNCIA ELEMENTAR. *Bombyx mori*. Casa das Ciências, 2022. Disponível em: <https://rce.casadasciencias.org/art/2022/018/>. Acesso em: 12 maio 2025.

SCHATZ MC, DELCHER AL, SALZBERG SL. Assembly of large genomes using second-generation sequencing. *Genome Res.* 2010 Sep; 20(9):1165-73. doi: 10.1101/gr.101360.109. Epub 2010 May 27. PMID: 20508146; PMCID: PMC2928494.

SI, H., CAO, Y., ZHU, H. et al. Transposable Element Bm1645 is a Source of BmAGO2-associated Small RNAs that affect its expression in *Bombyx mori*. *BMC Genomics* 18, 201 (2017). <https://doi.org/10.1186/s12864-017-3598-5>

Sun, X.; Yuan, Q.; Du, B.; Jin, X.; Huang, X.; Li, Q.; Zhong, Y.; Pan, Z.; Xu, S.; Sima, Y. Relação entre alterações nos microrganismos intestinais e o efeito da alta temperatura no crescimento e desenvolvimento de larvas de *Bombyx mori*. *Int. J. Mol. Sci.* 2022, 23, 10289. <https://doi.org/10.3390/ijms231810289>

Sun, Z., Kumar, D., Cao, G. et al. Efeitos do tratamento transitório com alta temperatura na flora intestinal do bicho-da-seda *Bombyx mori*. *Sci Rep* 7, 3349 (2017). <https://doi.org/10.1038/s41598-017-03565-4>

Thrilekha D, P Harshita Mala, B Gowthami, Chethan Kumar DS. Influence of temperature on growth and development of bivoltine silkworm breeds and hybrids: A

comprehensive review. *Int J Res Agron* 2024;7(6S):590-597. DOI: 10.33545/2618060X.2024.v7.i6Sh.2072.

THRILEKHA, D. et al. Influence of temperature on growth and development of bivoltine silkworm breeds and hybrids: a comprehensive review. *International Journal of Research in Agronomy*, v. 7, n. 6 (Supl.), 2024.

TRAPNELL, C., WILLIAMS, B., PERTEA, G. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28, 511–515 (2010). <https://doi.org/10.1038/nbt.1621>

Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009 Jan; 10(1):57-63. doi: 10.1038/nrg2484. PMID: 19015660; PMCID: PMC2949280.

WELCH, M. M.; LINQUIST, S.; MALIK, H. S. A field guide to eukaryotic transposable elements. *Annual Review of Genetics*, v. 55, p. 581–612, 2021. DOI: 10.1146/annurev-genet-040620-022145.

WELLS JN, FESCHOTTE C. A Field Guide to Eukaryotic Transposable Elements. *Annu Rev Genet.* 2020 Nov 23; 54:539-561. doi: 10.1146/annurev-genet-040620-022145. Epub 2020 Sep 21. PMID: 32955944; PMCID: PMC8293684.

WICKER, T., SABOT, F., HUA-VAN, A. et al. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8, 973–982 (2007). <https://doi.org/10.1038/nrg2165>

WICKER, THOMAS. So many repeats and so little time: how to classify transposable elements. In: GRANDBASTIEN, Marie-Angèle; CASACUBERTA, Josep M. (ed.). *Plant transposable elements: impact on genome structure and function*. Berlin Heidelberg: Springer-Verlag, 2012. (Topics in Current Genetics, v. 24). cap.1, p. 1–15. DOI: https://doi.org/10.1007/978-3-642-31842-9_1

Wu, C.; Lu, J. Diversification of Transposable Elements in Arthropods and Its Impact on Genome Evolution. *Genes* 2019, 10, 338. <https://doi.org/10.3390/genes10050338>

ZHANG, Q. et al. Effect of temperature stress on expression of heat shock protein genes in the silkworm, *Bombyx mori*. *Journal of Insect Physiology*, v. 55, n. 8, p. 777–782, 2009. DOI: <https://doi.org/10.1016/j.jinsphys.2009.04.013>

Zhang, C.; Wang, L.; Dou, L.; Yue, B.; Xing, J.; Li, J. Transposable Elements Shape the Genome Diversity and the Evolution of Noctuidae Species. *Genes* 2023, 14, 1244. <https://doi.org/10.3390/genes14061244>

Zhu, J., Chen, G., Zhu, S. et al. Identification of Tissue-Specific Protein-Coding and Noncoding Transcripts across 14 Human Tissues Using RNA-seq. *Sci Rep* 6, 28400 (2016). <https://doi.org/10.1038/srep28400>

6. APÊNDICES

Tabela 1. Acessos Dfam identificados nas montagens *de novo* e guiada.

Acessos	Total
5S_DM	64
7SK	1
7SLRNA	7
Academ_A.kagoshimensis_1	4
Academ_S.constricta_1	3
AcoCah-1.76	1
Acol_copia_Ele12	4
Acol_copia_Ele13	2
Acol_copia_Ele2	1
Acol_copia_Ele22	1
Acol_copia_Ele6	11
Acol_copia_Ele9	1
Acol_CR1_Ele5	1
Acol_gypsy_Ele18	1
Acol_gypsy_Ele20	5
Acol_gypsy_Ele21	3
Acol_gypsy_Ele23	4
Acol_gypsy_Ele25	2
Acol_gypsy_Ele33	2
Acol_gypsy_Ele34	8
Acol_gypsy_Ele39	1
Acol_gypsy_Ele5	1
Acol_gypsy_Ele58	1
Acol_gypsy_Ele69	2
Acol_gypsy_Ele74	1
Acol_gypsy_Ele81	4
Acol_gypsy_Ele90	10
Acol_gypsy_Ele91	2
Acol_gypsy_Ele95	54
Acol_hAT_Ele6	2
Acol_Jockey_Ele1	206
Acol_Jockey_Ele7	1
Acol_L1_Ele2	2
Acol_mar_Ele15	4
Acol_Pao_Bel_Ele16	1
Acol_Pao_Bel_Ele22	1
Acol_Pao_Bel_Ele25	3
Acol_Pao_Bel_Ele26	2
Acol_Pao_Bel_Ele29	1
Acol_Pao_Bel_Ele33	4
Acol_Pao_Bel_Ele37	1
Acol_Pao_Bel_Ele39	2
Acol_Pao_Bel_Ele4	1
Acol_Pao_Bel_Ele41	4
Acol_Pao_Bel_Ele45	1
Acol_Pao_Bel_Ele47	16
Acol_Pao_Bel_Ele51	8
Acol_Pao_Bel_Ele58	3

Acol_Pao_Bel_Ele6	1
Acol_Pao_Bel_Ele61	3
Acol_Pao_Bel_Ele62	4
Acol_Pao_Bel_Ele64	1
Acol_Pao_Bel_Ele67	1
Acol_Pao_Bel_Ele72	3
Acol_Pao_Bel_Ele75	3
Acol_Pao_Bel_Ele81	2
Acol_PIF_Ele1	39
Acol_PIF_Ele10	13
Acol_R1_Ele1	4
Acol_R1_Ele16	28
Acol_RTE_Ele1	2
Acol_Tc1_Ele15	1
Acol_Tc1_Ele20	1
ALR	1
AluSc8	1
AluSp	1
AluY	1
BEL11-I_DR	1
BEL-65_DRe-I	1
BGLII_B2	7
Bica_I	1
BisBis-1.8	1
Blastopia_I	10
CfSE1	22
CheMed-1.218_int	26
Chimpo_I	1
Chlamys-1_bBra	5
Chlamys-4_cRei	15
Chlamys-5_cSch	7
Chlamys-6_cBra	14
CMC_M.nervosa_1	1
CMC_M.nervosa_3	2
COMP-subunit_TAF11_rnd-6_family-27360	2
Copia-105_rCom	7
Copia-1_Dmel-I	5
Copia-2_Hm	1
Copia-2_rCom	2
Copia-3_DM-I	2
Copia-3_INT_AstCal	6
Copia-7-I_DR	19
Copia-8-I_DR	7
CR1-1_Hmel_A	2
CR1-1_Hmel_B2	1
CR1-4_Hmel_A	3
CR1-4_Hmel_B	1
CR1-8_Hmel_A	2
CR1-8_Hmel_B	5
CR1-8_Hmel_C	4
CR1-9_Hmel_A	1
CR1-9_Hmel_B	3
CR1-Zenon-1a_Avan	1
CR1-Zenon-1b_Avan	1
CR1-Zenon-1_Etal	80

CR1-Zenon-1_Ldor	4
CR1-Zenon-2_Diul	460
CR1-Zenon-3_Diul	2
CR1-Zenon-4_Diul	158
CraTho-1.264	1
CraTho-1.281	1
CraTho-6.8701a	19
Daphne-1_Hmel_A	2
Daphne-1_Hmel_B	8
Daphne-1_Hmel_C	4
Daphne-2_Hmel_A	3
Daphne-2_Hmel_B	1
DF003447270	2
DF003447444	3
DF003447511	13
DF003447527	1
DF003447577	1
DF003447704	1
DF003447766	2
DF003447959	1
DF003448021	1
DF003448133	1
DF003448229	1
DF003448318	31
DF003448407	2
DF003448584	1
DF003448600	37
DF003448838	6
DF003448853	1
DF003448865	3
DF003449051	1
DF003449069	1
DF003449136	1
DF003449138	1
DF003449172	13
DF003449224	28
DF003449493	15
DF003449516	7
DF003449572	1
DF003449575	4
DF003449741	3
DF003449821	2
DF003449851	1
DF003449868	8
DF003449931	1
DF003449993	4
DF003450110	3
DF003450149	1
DF003450178	2
DF003450287	1
DIRS-15B_DR	1
DIRS-4_DR	60
DIRS-8_DR	22
DMSAT6	3
DNA16_Hera	5

DNA16_Htel	7
DNA-1_Hbes	1
DNA-1_Hmel	1
DNA1_Htel	1
DNA20_Hera	7
DNA-2-15_DR	22
DNA-2-24_DR	2
DNA25_Hera	9
DNA-3-5_DR	2
DNA6-2_DR	1
DNA8-25_DR	7
DNA-like-11_Hmel	1
DNA-like-1_Hmel	1
DNA-like-6_Hmel	1
DOC3_DM	1
DOC5_DM	9
Dong-R4-1_Hcyd	3
Dong-R4-1_Ldor	8
Dong-R4-2_AstCal	3
ElaDav-1.98	2
EnSpm-9_DR	1
EubJap-5.255	32
EulFla-1.126	3
FelNig-1.17	2
FelNig-1.24	1
FelNig-1.34	2
FW_DM	1
GalVar-1.127	1
Gypsy-100_INT_AstCal	3
Gypsy-107_INT_AstCal	3
Gypsy-11_Hmel_I	7
Gypsy-13_Hmel_I	1
Gypsy-14_Hmel_I	5
Gypsy158-I_DR	4
Gypsy163-I_DR	1
Gypsy-184_DR-I	2
Gypsy-187_DR-I	14
Gypsy-206_DR-I	5
Gypsy-208_DR-I	1
Gypsy-217_DR-I	3
Gypsy-24_INT_AstCal	2
Gypsy-25_INT_AstCal	6
Gypsy-26_DMel-I	2
Gypsy-27_INT_AstCal	5
Gypsy-31_INT_AstCal	2
Gypsy-31-LTR_DR	5
Gypsy-3_DMel-I	35
Gypsy-45_INT_AstCal	3
Gypsy-46_rCom	1
Gypsy-48_rCom	3
Gypsy-49_INT_AstCal	3
Gypsy56-I_DR	6
Gypsy58-I_DR	1
Gypsy-78_INT_AstCal	1
Gypsy86-I_DR	5

Gypsy-87_INT_AstCal	4
Gypsy-88_INT_AstCal	5
Gypsy98-LTR_Dr	3
Gypsy99-I_DR	4
Halyomorpha_halys-1_family-129	1
Halyomorpha_halys-1_family-166	2
Halyomorpha_halys-1_family-177	1
Halyomorpha_halys-1_family-260	1
Halyomorpha_halys-1_family-314	1
Halyomorpha_halys-1_family-321	1
Halyomorpha_halys-1_family-349	3
Halyomorpha_halys-1_family-43	3
Halyomorpha_halys-1_family-449	7
Halyomorpha_halys-1_family-494	2
Halyomorpha_halys-1_family-520	3
Halyomorpha_halys-1_family-724	1
Halyomorpha_halys-1_family-733	1
Halyomorpha_halys-4_family-1342	1
Halyomorpha_halys-4_family-638	55
Halyomorpha_halys-5_family-1188	2
Halyomorpha_halys-5_family-1934	2
Halyomorpha_halys-5_family-2362	3
Halyomorpha_halys-5_family-3015	3
Halyomorpha_halys-5_family-466	3
Halyomorpha_halys-5_family-682	3
Halyomorpha_halys-6_family-1101	1
Halyomorpha_halys-6_family-2394	12
Halyomorpha_halys-6_family-713	1
Halyomorpha_halys-6_family-945	2
hAT_C.sinensis_2	1
hAT_M.nervosa_33	2
hAT-N22_DR	5
hAT-N85_DR	1
Helitron-10_Diul	110
Helitron10_Hera	65
Helitron-13_Diul	5
Helitron-1_Diul	3
Helitron-1_Etal	40
Helitron-1_Hdem	28
Helitron-2_Avan	121
Helitron-2_Diul	3
Helitron2_Hera	15
Helitron-2_Hsar	22
Helitron-3_Etal	2
Helitron-3_Hnum	1
Helitron-4_Etal	2
Helitron-4_Hbur	4
Helitron-4_Hera	5
Helitron-6_Diul	1
Helitron6_Hera	2
Helitron-7_Etal	5
Helitron7_Hera	4
Helitron-7_Hpar	17
Helitron8_Hera	151
Helitron-like-1_Ldor	32

Helitron-like-8_Hmel	17
HemHyl-2.7	3
HipAmp-1.111	1
HSATII	1
I-1_Avan	17
I-1_Hmel_D	4
I-1_Hmel_E	4
I-1_Hmel_F	11
I-1_Hmel_G	2
I-1_Hmel_H	12
I-1_Hmel_I	4
I-1_Hmel_N	5
I-1_Hmel_O	1
IMPB_01	3
IndInd-1.118	1
I-Nimb-1a_Hcyd	3
I-Nimb-1b_Hcyd	47
Invader1_I	3
Invader4f_I	4
Invader7_I	15
Jockey-1_Hbur	1
Jockey-2_Hbur	32
Jockey-2_Hmel_B	6
Jockey-3_Hmel_D	8
Jockey-3_Hmel_E	32
L1-14_AstCal	1
L1-22_rCom	4
L1-Tx1-1_Diul	37
L2-11_Hmel_A1	5
L2-11_Hmel_A2	1
L2-14_Hmel_B2	4
L2-15_Hmel_A2	2
L2-15_Hmel_B1	2
L2-1a_Hnum	9
L2-1_Avan	1
L2-1b_Hnum	17
L2-1_Diul	13
L2-1_Etal	23
L2-1_Etal_B	29
L2-1_Hbes	3
L2-1_Hbur	2
L2-1_Hmel_B	1
L2-1_Ldor	8
L2-1_Ldor_D	3
L2-1_Ldor_E	1
L2-1_Ldor_F	1
L2-1_Ldor_G	1
L2-2_Etal	1
L2-2_Hbes	1
L2-3_Etal	1
L2-3_Hmel_A1	2
L2-3_Hmel_A3	8
L2-3_Hmel_C1	1
L2-4_Hmel_A	1
L2-4_Hmel_C1	2

L2-5_Hmel_A	9
LasBor-1.134	3
LasBor-1.248	1
LasBor-4.3117	2
LasBor-5.38	6
LasBor-5.398	1
LaSE1	1
LINE_A.kagoshimensis_13_cons	4
LINE_A.kagoshimensis_32_cons	1
LINE_B.platifrons_159_cons	1
LINE_C.sinensis_1_cons	1
LINE_C.sinensis_4_cons	1
LINE_C.sinensis_71_cons	31
LINE_C.virginica_54_cons	2
LINE_D.rostriformis_104_cons	2
LINE_D.rostriformis_4_cons	1
LINE_D.rostriformis_99_cons	4
LINE_M.mercenaria_187_cons	3
LINE_M.mercenaria_45_cons	1
LINE_M.nervosa_11_cons	1
LINE_M.nervosa_12_cons	1
LINE_M.nervosa_19_cons	11
LINE_M.phylippinarum_13_cons	1
LINE_M.phylippinarum_93_cons	2
LINE_M.yessoensis_57_cons	60
LINE_M.yessoensis_67_cons	1
LINE_P.fucata_18_cons	1
LINE_P.maximus_108_cons	7
LINE_P.maximus_28_cons	1
LINE_P.maximus_52_cons	5
LINE_S.broughtoni_10_cons	2
LINE_S.broughtoni_19_cons	10
LINE_S.constricta_1_cons	1
LINE_S.grandis_18_cons	1
LINE_S.grandis_40_cons	1
LINE_T.granosa_24_cons	1
LSU-rRNA_Cel	81
LSU-rRNA_Dme	72
LSU-rRNA_Hsa	40
LTR-4_Hm	5
LTR-5_Hm	1
LTR75B	1
MARINCE1	10
Mariner-12_DR	3
Mariner1_Hera	1
Mariner-1_Hmel	1
Mariner-1_Hpar	1
Mariner-4_DR	1
Mariner5_Hera	42
Mariner-N5B_DR	3
MDG3_I	1
MegLyr-5.207b	2
MegLyr-5.2125	1
Merlin_M.phylippinarum_6	1
Merrow_Cpb	4

MesAur-1.193	1
Micropia2_DM-I	5
Micropia6_Dros-I	1
Micropia7_Dmel_I	5
MinSch-5.2265	1
MirCoq-1.117	10
MMSAT4	382
MsSE2	11
MULE_S.constricta_35	1
MurFea-5.2046	2
MurFea-6.1933	24
MurSatRep1	17
MyoBra-1.237	5
MyoBra-1.328	3
MyoBra-1.435	20
NA-DNA-15_AstCal	27
nhAT-1_Diul	14
nhAT-1_Hdem	4
nhAT-1_Hhim	21
nhat-1_Hmel	4
nhAT3_mMyo	3
Nimb-12_DR	1
nMar-11_Hmel	10
nMar-14_Hmel	5
nMar-16_Hele	9
nMar-16_Hpar	3
nMar-1_Etal	1
nMar-1_Hhim	4
nMar-1_Ldor	4
nMar1_rFer	3
nMar-21_Hmel	1
nMar-4_Hmel	3
nMar5_rFer	3
nMar-7_Hmel	2
nMar-8_Hmel	29
nMar-8_Hnum	7
nMar-9_Hmel	1
NomLeu-5.736_int	1
npiggyBac-5_Hmel	8
nTc3-3_Hcyd	4
ObSE1	1
OnyTor-1.356_int	3
PanOnc-1.30	1
PanPar-1.29	1
PanTig-1.13	4
Pao-32_INT_AstCal	7
Pao-41_INT_AstCal	4
Pao-42_INT_AstCal	3
Pao-4_INT_AstCal	1
ParHer-1.13	1
Penelope-1_Avan	1
Penelope-1_Hhim	1
Penelope-3_Avan	20
PIF-Harbinger_C.gigas_17	1
PIF-Harbinger_C.gigas_20	1

PIF-Harbinger_C.gigas_21	2
PIF-Harbinger_D.rostriformis_14	2
PIF-Harbinger_M.coruscus_15	4
PIF-Harbinger_M.edulis_34	1
PIF-Harbinger_S.constricta_37	3
PIF-Harbinger_S.glomerata_8	31
PiggyBac_A.i.concentricus_4	1
PiggyBac_A.marissinica_178	1
PiggyBac_B.platifrons_29	9
PiggyBac_C.ariakensis_10	9
PiggyBac_P.maximus_17	3
PipPip-5.1542	10
PipPip-6.1914	2
PLE-X1_cBra	1
PLE-X2_cBra	1
Pman_ALU_5.1881	1
Pman_ALU_6.2822	5
Pman_ALU_6.4482	5
Pman_ERVL_5.247	6
Pman_ERVL-MALR_6.16429	6
Pman_MIR_6.141	6
Pman_SNRNA_5.818	9
Pman_SNRNA_6.7824	12
Pman_UNKNOWN_6.2013	5
Polinton-1_DR	7
Proto2-1_AstCal	12
PumCon-1.11	1
PxSE2	32
R1-1_Hmel_C	21
R1-2_Hmel_A	30
R4-1_HMe_A	6
R4-1_HMe_B	4
R4-1_HMe_C	1
RhiPru-5.1045	1
RhiSin-1.199	1
RhiSin-1.210	1
RLTR20A4	10
RLTR31D_MM	3
RLTR51A_Mm	1
RooA_I	1
RTE-13_Hmel_A	71
RTE-13_Hmel_B	16
RTE-16_Hmel_A	5
RTE-16_Hmel_B	17
RTE-16_Hmel_C	18
RTE-16_Hmel_D	42
RTE-16_Hmel_E	6
RTE-16_Hmel_F	10
RTE-18_Hmel_A	5
RTE-1_Avan	11
RTE-1_Heca	8
RTE-1_HMe_A1	8
RTE-1_HMe_A2	2
RTE-1_HMe_B	10
RTE-1_Htim	2

RTE-20_Hmel_B	1
RTE-20_Hmel_C	7
RTE-20_Hmel_D	1
RTE-20_Hmel_E	1
RTE-20_Hmel_F	2
RTE-20_Hmel_G	2
RTE-2_Hmel_A	19
RTE-2_Hmel_A2	9
RTE-2_Hmel_B	1
RTE-3_Hcyd	7
RTE-3_Hmel_A	3
RTE-3_Hmel_C	6
RTE-4_Hmel_B	3
RTE-4_Hmel_C	3
RTE-4_Hmel_D	1
RTE-5_Hmel_C	1
RTE-6_Hmel_A3	1
RTE-6_Hmel_B	8
RTE-6_Hmel_C1	2
RTE-6_Hmel_C2	12
RTE-6_Hmel_D	11
RTE-6_Hmel_E	14
RTE-6_Hmel_F	1
RTE-6_Hmel_G	4
RTE-6_Hmel_H1	2
RTE-6_Hmel_H2	2
RTE-6_Hmel_I	11
RTE-7_Hmel_A	3
RTE-7_Hmel_B	6
RTE-7_Hmel_C	11
RTE-7_Hmel_E	4
RTE-BovB-1_Heca	1
RTE-BovB-4_AstCal	26
SART1_R1_A	8
SAT-15_DR	1
SAT-16_DR	1
ScaAqu-1.29	1
Sola1_B.platifrons_33	12
Sola2_A.marissinica_189	10
SorAra-1.36	2
Ssa_0070	12
Ssa_0264	58
Ssa_0972	1
Ssa_1250	3
Ssa_1657	3
Ssa_1781	1
Ssa_1959	10
Ssa_2003	1
SSU-rRNA_Dme	30
SSU-rRNA_Hsa	74
SVA_A	3
TadBra-1.184	1
TadBra-6.2216_int	26
TcMar_A.kagoshimensis_27	7
TcMar_A.kagoshimensis_32	1

TcMar_A.kagoshimensis_36	5
TcMar_B.platifrons_48	2
TcMar_C.sinensis_15	1
TcMar-ISRm11-4_AstCal	2
TcMar_M.mercenaria_15	29
TcMar_M.mercenaria_38	2
TcMar_M.philippinarum_21	3
TcMar_P.maximus_23	1
TcMar_R.philippinarum_24	3
TcMar_S.broughtoni_19	1
TcMar_S.grandis_51	1
TcMar_T.granosa_20	2
TE0030_SO2_FAM1029	1
TE0042_SO2_FAM1041	1
TE0050_SO2_FAM0105	3
TE0053_SO2_FAM1052	11
TE0075_SO2_FAM1072	1
TE0096_SO2_FAM1094	8
TE0102_SO2_FAM0011	13
TE0105_SO2_FAM1103	4
TE0111_SO2_FAM1109	4
TE0118_SO2_FAM1117	2
TE0121_SO2_FAM0112	2
TE0123_SO2_FAM1121	4
TE0169_SO2_FAM1167	7
TE0172_SO2_FAM0117	1
TE0181_SO2_FAM1178	2
TE0183_SO2_FAM0118	20
TE0188_SO2_FAM1184	4
TE0200_SO2_FAM1198	7
TE0203_SO2_FAM0120	2
TE0251_SO2_FAM0125	8
TE0262_SO2_FAM0126	31
TE0281_SO2_FAM1278	1
TE0288_SO2_FAM1285	3
TE0297_SO2_FAM1294	5
TE0298_SO2_FAM1295	4
TE0300_SO2_FAM1297	3
TE0303_SO2_FAM0013	2
TE0314_SO2_FAM0131	3
TE0320_SO2_FAM1315	1
TE0331_SO2_FAM1325	4
TE0346_SO2_FAM0134	9
TE0366_SO2_FAM1359	4
TE0367_SO2_FAM0136	9
TE0376_SO2_FAM1368	30
TE0389_SO2_FAM0138	2
TE0399_SO2_FAM0139	3
TE0410_SO2_FAM0140	1
TE0415_SO2_FAM1404	5
TE0430_SO2_FAM0142	5
TE0440_SO2_FAM0143	4
TE0451_SO2_FAM1440	12
TE0458_SO2_FAM1447	5
TE0459_SO2_FAM1448	15

TE0461_SO2_FAM0145	2
TE0483_SO2_FAM1470	2
TE0519_SO2_FAM1504	10
TE0533_SO2_FAM1518	1
TE0535_SO2_FAM0152	1
TE0553_SO2_FAM0154	137
TE0555_SO2_FAM1541	2
TE0564_SO2_FAM0155	7
TE0573_SO2_FAM1559	2
TE0578_SO2_FAM1563	9
TE0585_SO2_FAM0157	1
TE0595_SO2_FAM0158	1
TE0597_SO2_FAM1581	3
TE0605_SO2_FAM0159	2
TE0616_SO2_FAM0016	3
TE0624_SO2_FAM1606	3
TE0628_SO2_FAM0161	4
TE0639_SO2_FAM0162	1
TE0733_SO2_FAM1714	1
TE0749_SO2_FAM0173	1
TE0753_SO2_FAM1733	4
TE0770_SO2_FAM0175	21
TE0790_SO2_FAM0177	16
TE0804_SO2_FAM1783	10
TE0910_SO2_FAM1889	11
TE0917_SO2_FAM1896	1
TE0991_SO2_FAM1969	2
TE1008_SO2_FAM1986	1
TE1045_SO2_FAM0202	8
TE1178_SO2_FAM0215	10
TE1203_SO2_FAM2173	17
TE1210_SO2_FAM0218	2
TE1214_SO2_FAM2183	4
TE1233_SO2_FAM0220	3
TE1243_SO2_FAM2209	10
TE1274_SO2_FAM0224	1
TE1275_SO2_FAM2240	22
TE1296_SO2_FAM0226	2
TE1299_SO2_FAM2262	20
TE1307_SO2_FAM0227	1
TE1316_SO2_FAM2278	6
TE1329_SO2_FAM2291	22
TE1346_SO2_FAM2309	1
TE1356_SO2_FAM2319	4
TE1373_SO2_FAM2336	11
TE1377_SO2_FAM0234	1
TE1379_SO2_FAM2341	4
TE1382_SO2_FAM2344	1
TE1397_SO2_FAM2358	2
TE1409_SO2_FAM0237	1
TE1416_SO2_FAM2376	18
TE1452_SO2_FAM2410	4
TE1499_SO2_FAM2455	4
TE1521_SO2_FAM2475	1
TE1592_SO2_FAM2540	2

TE1645_SO2_FAM0259	7
TE1671_SO2_FAM2614	6
TE1677_SO2_FAM0262	5
TE1699_SO2_FAM0264	3
TE1721_SO2_FAM0266	17
TE1742_SO2_FAM2679	7
TE1755_SO2_FAM2690	1
TE1763_SO2_FAM2698	54
TE1776_SO2_FAM0271	6
TE1826_SO2_FAM2757	2
TE1829_SO2_FAM0276	41
TE1846_SO2_FAM2776	5
TE1861_SO2_FAM0279	14
TE1871_SO2_FAM2799	1
TE1922_SO2_FAM2846	7
TE1928_SO2_FAM2851	8
TE1970_SO2_FAM0289	5
TE1982_SO2_FAM0290	4
TE2015_SO2_FAM0293	10
TE2026_SO2_FAM0294	3
TE2048_SO2_FAM0296	1
TE2058_SO2_FAM0297	1
TE2065_SO2_FAM2976	1
TE2080_SO2_FAM0299	12
TE2092_SO2_FAM0300	8
TE2125_SO2_FAM0303	1
TE2136_SO2_FAM0304	4
TE2147_SO2_FAM0305	2
TE2160_SO2_FAM3061	1
TE2169_SO2_FAM0307	5
TE2180_SO2_FAM0308	1
TE2211_SO2_FAM3107	9
TE2241_SO2_FAM3134	105
TE2247_SO2_FAM0314	1
TE2257_SO2_FAM0318	7
TE2260_SO2_FAM0320	9
TE2262_SO2_FAM0322	8
TE2266_SO2_FAM0327	1
TE2270_SO2_FAM0331	1
TE2282_SO2_FAM0346	6
TE2283_SO2_FAM0347	3
TE2294_SO2_FAM0359	35
TE2300_SO2_FAM0364	1
TE2308_SO2_FAM0372	3
TE2328_SO2_FAM0391	1
TE2332_SO2_FAM0395	1
TE2343_SO2_FAM0405	2
TE2348_SO2_FAM0041	1
TE2354_SO2_FAM0415	1
TE2358_SO2_FAM0042	20
TE2364_SO2_FAM0426	8
TE2365_SO2_FAM0427	7
TE2381_SO2_FAM0441	13
TE2384_SO2_FAM0446	8
TE2385_SO2_FAM0447	1

TE2387_SO2_FAM0449	6
TE2388_SO2_FAM0045	4
TE2389_SO2_FAM0450	1
TE2392_SO2_FAM0456	17
TE2395_SO2_FAM0459	3
TE2397_SO2_FAM0460	1
TE2410_SO2_FAM0473	2
TE2412_SO2_FAM0475	1
TE2413_SO2_FAM0476	8
TE2419_SO2_FAM0482	5
TE2425_SO2_FAM0488	1
TE2427_SO2_FAM0490	11
TE2434_SO2_FAM0498	2
TE2441_SO2_FAM0503	2
TE2453_SO2_FAM0514	1
TE2456_SO2_FAM0517	2
TE2463_SO2_FAM0525	4
TE2465_SO2_FAM0527	2
TE2475_SO2_FAM0537	2
TE2479_SO2_FAM0540	2
TE2482_SO2_FAM0543	1
TE2485_SO2_FAM0546	2
TE2486_SO2_FAM0547	9
TE2492_SO2_FAM0553	8
TE2509_SO2_FAM0569	1
TE2510_SO2_FAM0057	15
TE2513_SO2_FAM0573	4
TE2519_SO2_FAM0579	6
TE2520_SO2_FAM0058	5
TE2524_SO2_FAM0584	2
TE2525_SO2_FAM0585	6
TE2528_SO2_FAM0588	2
TE2531_SO2_FAM0592	1
TE2543_SO2_FAM0603	73
TE2549_SO2_FAM0610	1
TE2550_SO2_FAM0611	2
TE2552_SO2_FAM0613	3
TE2566_SO2_FAM0628	1
TE2567_SO2_FAM0629	4
TE2576_SO2_FAM0064	15
TE2577_SO2_FAM0640	1
TE2589_SO2_FAM0651	1
TE2607_SO2_FAM0668	2
TE2644_SO2_FAM0707	6
TE2659_SO2_FAM0722	1
TE2660_SO2_FAM0724	1
TE2666_SO2_FAM0073	1
TE2685_SO2_FAM0752	3
TE2699_SO2_FAM0771	3
TE2711_SO2_FAM0782	2
TE2735_SO2_FAM0807	1
TE2739_SO2_FAM0810	4
TE2776_SO2_FAM0850	5
TE2783_SO2_FAM0857	1
TE2785_SO2_FAM0859	3

TE2812_SO2_FAM0884	7
TE2818_SO2_FAM0089	3
TE2831_SO2_FAM0901	11
TE2840_SO2_FAM0091	4
TE2843_SO2_FAM0913	8
TE2850_SO2_FAM0092	1
TE2861_SO2_FAM0930	2
TE2863_SO2_FAM0933	10
TE2877_SO2_FAM0948	1
TE2887_SO2_FAM0957	4
TE2912_SO2_FAM0982	36
TE2920_SO2_FAM0099	6
TE2940_SO2_FAM1014	10
TE2941_SO2_FAM1014	2
TE2987_SO2_FAM121	12
TE2988_SO2_FAM121	5
TE2996_SO2_FAM123	35
TE2999_SO2_FAM127	4
TE3000_SO2_FAM127	1
TE3025_SO2_FAM147	1
TE3056_SO2_FAM163	1
TE3064_SO2_FAM1643	137
TE3088_SO2_FAM179	1
TE3136_SO2_FAM201	1
TE3137_SO2_FAM201	2
TE3164_SO2_FAM221	2
TE3174_SO2_FAM023	4
TE3184_SO2_FAM233	1
TE3197_SO2_FAM025	5
TE3213_SO2_FAM275	1
TE3232_SO2_FAM333	3
TE3246_SO2_FAM390	1
TE3250_SO2_FAM419	7
TE3260_SO2_FAM454	6
TE3264_SO2_FAM465	4
TE3317_SO2_FAM701	1
TE3341_SO2_FAM076	3
TE3379_SO2_FAM088	53
TE3380_SO2_FAM088	1
TE3381_SO2_FAM088	18
TE3391_SO2_FAM094	6
TE-X-12_DR	9
TraJav-1.74	1
tRNA-3_DRe	1
tRNA-Ala-GCG	3
tRNA-Ala-GCY	8
tRNA-Arg-AGA	1
tRNA-Asn-AAC	5
tRNA-Asp-GAY	8
tRNA-Cys-TGY	1
tRNA-Gly-GGA	1
tRNA-Ile-ATT	2
tRNA-Lys-AAG	5
tRNA-Met_v	3
tRNA-Pro-CCA	1

tRNA-Pro-CCY	2
tRNA-Ser-AGY	4
tRNA-Ser-TCG	2
tRNA-Thr-ACA	1
tRNA-Val-GTG	4
U1	26
U13	1
U17	2
U2	2
U3	8
U4	3
U5	5
U6	2
Zaphod	1
ZapHud-1.7	1
Zator-1_Etal	2
Zenon-1_Hmel_A	8
Zenon-1_Hmel_B	5
Zenon-1_Hmel_C	1
Zenon-1_Hmel_D	9
ZenoSINE1_Avan	4
ZenoSINE1_eTal	4
ZenoSINE1_hTel	15
ZenoSINE1_IDor	5
ZymTri_2023_family_141	88
ZymTri_2023_family_1487	4
ZymTri_2023_family_246	2
ZymTri_2023_family_365	1
ZymTri_2023_family_445	3
ZymTri_2023_family_532	61
ZymTri_2023_family_717	1

Tabela 2. Acessos Dfam identificados apenas na montagem guiada.

Acessos	Total
5S_DM	53
7SLRNA	3
Academ_A.kagoshimensis_1	4
Academ_S.constricta_1	1
Acol_copia_Ele12	3
Acol_copia_Ele13	1
Acol_copia_Ele2	1
Acol_copia_Ele6	2
Acol_gypsy_Ele18	1
Acol_gypsy_Ele20	4
Acol_gypsy_Ele21	1
Acol_gypsy_Ele23	3
Acol_gypsy_Ele33	2
Acol_gypsy_Ele34	2
Acol_gypsy_Ele39	1
Acol_gypsy_Ele5	1
Acol_gypsy_Ele58	1

Acol_gypsy_Ele69	2
Acol_gypsy_Ele81	4
Acol_gypsy_Ele90	6
Acol_gypsy_Ele95	10
Acol_Jockey_Ele1	109
Acol_L1_Ele2	1
Acol_Pao_Bel_Ele16	1
Acol_Pao_Bel_Ele22	1
Acol_Pao_Bel_Ele25	3
Acol_Pao_Bel_Ele26	2
Acol_Pao_Bel_Ele33	1
Acol_Pao_Bel_Ele39	1
Acol_Pao_Bel_Ele4	1
Acol_Pao_Bel_Ele41	1
Acol_Pao_Bel_Ele45	1
Acol_Pao_Bel_Ele47	3
Acol_Pao_Bel_Ele51	6
Acol_Pao_Bel_Ele58	3
Acol_Pao_Bel_Ele6	1
Acol_Pao_Bel_Ele61	3
Acol_Pao_Bel_Ele62	3
Acol_Pao_Bel_Ele67	1
Acol_Pao_Bel_Ele72	3
Acol_Pao_Bel_Ele75	2
Acol_PIF_Ele1	6
Acol_PIF_Ele10	9
Acol_R1_Ele1	1
Acol_R1_Ele16	26
Acol_Tc1_Ele15	1
Acol_Tc1_Ele20	1
BGLII_B2	6
BisBis-1.8	1
Blastopia_I	2
CfSE1	9
CheMed-1.218_int	17
Chimpo_I	1
Chlamys-1_bBra	4
Chlamys-4_cRei	10
Chlamys-5_cSch	6
Chlamys-6_cBra	6
CMC_M.nervosa_1	1
CMC_M.nervosa_3	1
COMP-subunit_TAF11_rnd-6_family-27360	1
Copia-1_Dmel-I	3
Copia-105_rCom	3
Copia-2_rCom	2
Copia-3_INT_AstCal	1
Copia-7-I_DR	7
Copia-8-I_DR	1
CR1-1_Hmel_A	1
CR1-4_Hmel_B	1
CR1-8_Hmel_A	1
CR1-8_Hmel_B	3
CR1-8_Hmel_C	2
CR1-Zenon-1a_Avan	29

CR1-Zenon-1b_Avan	1
CR1-Zenon-2_Diul	207
CR1-Zenon-3_Diul	1
CR1-Zenon-4_Diul	82
CraTho-6.8701a	5
Daphne-1_Hmel_A	2
Daphne-1_Hmel_B	8
Daphne-1_Hmel_C	4
Daphne-2_Hmel_A	3
Daphne-2_Hmel_B	1
DF003447270	1
DF003447444	3
DF003447511	7
DF003447766	1
DF003447959	1
DF003448021	1
DF003448133	1
DF003448318	7
DF003448407	2
DF003448600	7
DF003448838	3
DF003448865	2
DF003449172	3
DF003449224	5
DF003449493	6
DF003449516	4
DF003449572	1
DF003449575	2
DF003449741	1
DF003449821	2
DF003450110	3
DF003450178	2
DF003450287	1
DIRS-4_DR	20
DIRS-8_DR	7
DMSAT6	1
DNA-1_Hbes	1
DNA-1_Hmel	2
DNA-like-11_Hmel	14
DNA1_Htel	6
DNA16_Hera	1
DNA20_Hera	1
DOC5_DM	5
Dong-R4-1_Hcyd	1
Dong-R4-2_AstCal	1
ElaDav-1.98	2
EubJap-5.255	20
EulFla-1.126	2
FelNig-1.17	2
FelNig-1.34	2
Gypsy-100_INT_AstCal	1
Gypsy-107_INT_AstCal	1
Gypsy-11_Hmel_I	2
Gypsy-14_Hmel_I	1
Gypsy-184_DR-I	4

Gypsy-208_DR-I	2
Gypsy-217_DR-I	4
Gypsy-24_INT_AstCal	1
Gypsy-25_INT_AstCal	1
Gypsy-26_DMeI-I	2
Gypsy-3_DMeI-I	2
Gypsy-31_INT_AstCal	1
Gypsy-45_INT_AstCal	3
Gypsy-78_INT_AstCal	3
Gypsy-88_INT_AstCal	4
Gypsy56-I_DR	2
Gypsy58-I_DR	3
Gypsy86-I_DR	4
Gypsy98-LTR_Dr	1
Halyomorpha_halys-1_family-129	1
Halyomorpha_halys-1_family-314	1
Halyomorpha_halys-1_family-349	1
Halyomorpha_halys-1_family-449	3
Halyomorpha_halys-1_family-494	2
Halyomorpha_halys-1_family-520	1
Halyomorpha_halys-4_family-1342	1
Halyomorpha_halys-4_family-638	12
Halyomorpha_halys-5_family-2362	1
Halyomorpha_halys-5_family-3015	1
Halyomorpha_halys-6_family-2394	2
Halyomorpha_halys-6_family-945	1
hAT-N22_DR	3
Helitron-1_Diul	65
Helitron-1_Etal	14
Helitron-1_Hdem	3
Helitron-10_Diul	1
Helitron-13_Diul	5
Helitron-2_Avan	3
Helitron-2_Diul	32
Helitron-2_Hsar	1
Helitron-3_Etal	1
Helitron-3_Hnum	4
Helitron-4_Etal	1
Helitron-6_Diul	1
Helitron-like-8_Hmel	1
Helitron10_Hera	2
Helitron2_Hera	1
Helitron6_Hera	41
Helitron7_Hera	10
Helitron8_Hera	1
HemHyl-2.7	1
I-1_Avan	2
I-1_Hmel_F	1
I-1_Hmel_G	1
I-1_Hmel_H	1
I-1_Hmel_I	3
I-1_Hmel_N	1
I-Nimb-1a_Hcyd	2
I-Nimb-1b_Hcyd	1
IndInd-1.118	12

Invader1_I	2
Invader4f_I	2
Invader7_I	11
Jockey-1_Hbur	1
Jockey-2_Hbur	5
Jockey-3_Hmel_D	1
Jockey-3_Hmel_E	6
L1-Tx1-1_Diul	6
L2-1_Avan	5
L2-1_Etal	3
L2-1_Etal_B	2
L2-1_Hbes	2
L2-1_Ldor	11
L2-1_Ldor_D	10
L2-1_Ldor_E	15
L2-1_Ldor_F	13
L2-1_Ldor_G	2
L2-11_Hmel_A1	1
L2-14_Hmel_B2	1
L2-2_Etal	1
L2-3_Etal	1
LasBor-1.134	1
LasBor-1.248	1
LasBor-5.38	1
LasBor-5.398	1
LINE_A.kagoshimensis_32_cons	1
LINE_B.platifrons_159_cons	1
LINE_C.sinensis_71_cons	7
LINE_C.virginica_54_cons	2
LINE_D.rostriformis_104_cons	1
LINE_M.mercenaria_45_cons	1
LINE_M.nervosa_11_cons	1
LINE_M.nervosa_19_cons	4
LINE_M.yessoensis_57_cons	19
LINE_M.yessoensis_67_cons	1
LINE_P.maximus_108_cons	4
LINE_P.maximus_52_cons	5
LINE_S.broughtoni_19_cons	4
LSU-rRNA_Cel	1
LSU-rRNA_Dme	3
LTR-4_Hm	1
LTR-5_Hm	1
MARINCE1	2
Mariner1_Hera	4
MDG3_I	1
Merrow_Cpb	1
Micropia2_DM-I	3
Micropia6_Dros-I	1
MirCoq-1.117	4
MMSAT4	170
MsSE2	8
MULE_S.constricta_35	1
MurFea-5.2046	1
MurFea-6.1933	17
MurSatRep1	17

MyoBra-1.237	1
MyoBra-1.435	6
NA-DNA-15_AstCal	11
nhAT-1_Diul	2
nhAT-1_Hdem	1
nhAT-1_Hhim	4
nhat-1_Hmel	2
nhAT3_mMyo	1
nMar-1_Etal	1
nMar-1_Hhim	4
nMar-11_Hmel	2
nMar-14_Hmel	1
nMar-16_Hpar	1
nMar-21_Hmel	1
nMar-9_Hmel	4
npiggyBac-5_Hmel	3
Pao-32_INT_AstCal	1
Pao-4_INT_AstCal	3
Penelope-1_Hhim	1
Penelope-3_Avan	7
PIF-Harbinger_C.gigas_17	1
PIF-Harbinger_C.gigas_21	1
PIF-Harbinger_M.coruscus_15	1
PIF-Harbinger_M.edulis_34	1
PIF-Harbinger_S.glomerata_8	11
PiggyBac_B.platifrons_29	2
PiggyBac_C.ariakensis_10	7
PiggyBac_P.maximus_17	3
PipPip-5.1542	2
PipPip-6.1914	2
Pman_ALU_6.2822	1
Pman_ALU_6.4482	1
Pman_ERVL_5.247	3
Pman_ERVL-MALR_6.16429	2
Pman_SNRNA_5.818	4
Pman_SNRNA_6.7824	10
Pman_UNKNOWN_6.2013	1
Polinton-1_DR	3
Proto2-1_AstCal	9
PxSE2	12
R1-1_Hmel_C	1
R1-2_Hmel_A	4
R4-1_HMe_A	1
RhiSin-1.199	1
RLTR20A4	8
RLTR51A_Mm	1
RooA_I	1
RTE-1_Avan	11
RTE-1_Heca	4
RTE-1_HMe_A1	3
RTE-1_HMe_A2	3
RTE-1_HMe_B	7
RTE-1_Htim	13
RTE-13_Hmel_A	2
RTE-13_Hmel_B	2

RTE-16_Hmel_A	2
RTE-16_Hmel_B	5
RTE-16_Hmel_C	5
RTE-16_Hmel_D	1
RTE-16_Hmel_E	1
RTE-16_Hmel_F	5
RTE-2_Hmel_A2	1
RTE-20_Hmel_B	1
RTE-20_Hmel_E	6
RTE-20_Hmel_F	2
RTE-20_Hmel_G	1
RTE-3_Hcyd	1
RTE-3_Hmel_A	3
RTE-3_Hmel_C	5
RTE-4_Hmel_B	3
RTE-6_Hmel_A3	1
RTE-6_Hmel_B	5
RTE-6_Hmel_C1	2
RTE-6_Hmel_C2	5
RTE-6_Hmel_D	5
RTE-6_Hmel_E	1
RTE-6_Hmel_F	1
RTE-6_Hmel_G	4
RTE-6_Hmel_H1	1
RTE-6_Hmel_H2	1
RTE-6_Hmel_I	4
RTE-7_Hmel_A	3
RTE-7_Hmel_B	1
RTE-7_Hmel_C	2
RTE-7_Hmel_E	1
RTE-BovB-4_AstCal	17
SART1_R1_A	4
SAT-16_DR	1
ScaAqu-1.29	1
Sola1_B.platifrons_33	2
Sola2_A.marissinica_189	7
SorAra-1.36	1
Ssa_0070	4
Ssa_0264	9
Ssa_1250	2
Ssa_1959	8
Ssa_2003	1
SSU-rRNA_Dme	26
SSU-rRNA_Hsa	2
TadBra-6.2216_int	23
TcMar_A.kagoshimensis_27	3
TcMar_A.kagoshimensis_36	3
TcMar_B.platifrons_48	1
TcMar_M.mercenaria_38	15
TcMar_S.broughtoni_19	1
TcMar_S.grandis_51	1
TcMar-ISRm11-4_AstCal	1
TE0042_SO2_FAM1041	1
TE0050_SO2_FAM0105	3
TE0075_SO2_FAM1072	5

TE0096_SO2_FAM1094	7
TE0102_SO2_FAM0011	3
TE0105_SO2_FAM1103	1
TE0111_SO2_FAM1109	2
TE0123_SO2_FAM1121	1
TE0172_SO2_FAM0117	2
TE0181_SO2_FAM1178	8
TE0203_SO2_FAM0120	6
TE0251_SO2_FAM0125	17
TE0262_SO2_FAM0126	1
TE0281_SO2_FAM1278	1
TE0303_SO2_FAM0013	2
TE0320_SO2_FAM1315	1
TE0331_SO2_FAM1325	3
TE0346_SO2_FAM0134	4
TE0366_SO2_FAM1359	2
TE0367_SO2_FAM0136	2
TE0389_SO2_FAM0138	3
TE0410_SO2_FAM0140	3
TE0415_SO2_FAM1404	2
TE0430_SO2_FAM0142	4
TE0440_SO2_FAM0143	5
TE0451_SO2_FAM1440	4
TE0458_SO2_FAM1447	4
TE0459_SO2_FAM1448	1
TE0461_SO2_FAM0145	2
TE0535_SO2_FAM0152	100
TE0553_SO2_FAM0154	1
TE0555_SO2_FAM1541	5
TE0564_SO2_FAM0155	2
TE0573_SO2_FAM1559	2
TE0585_SO2_FAM0157	1
TE0595_SO2_FAM0158	2
TE0597_SO2_FAM1581	1
TE0616_SO2_FAM0016	2
TE0628_SO2_FAM0161	1
TE0733_SO2_FAM1714	1
TE0749_SO2_FAM0173	1
TE0753_SO2_FAM1733	7
TE0770_SO2_FAM0175	3
TE0790_SO2_FAM0177	3
TE0910_SO2_FAM1889	1
TE0917_SO2_FAM1896	1
TE0991_SO2_FAM1969	1
TE1008_SO2_FAM1986	5
TE1045_SO2_FAM0202	3
TE1178_SO2_FAM0215	12
TE1203_SO2_FAM2173	2
TE1233_SO2_FAM0220	2
TE1243_SO2_FAM2209	1
TE1274_SO2_FAM0224	1
TE1296_SO2_FAM0226	5
TE1307_SO2_FAM0227	3
TE1316_SO2_FAM2278	10
TE1329_SO2_FAM2291	1

TE1346_SO2_FAM2309	1
TE1356_SO2_FAM2319	3
TE1377_SO2_FAM0234	1
TE1409_SO2_FAM0237	3
TE1452_SO2_FAM2410	1
TE1645_SO2_FAM0259	3
TE1677_SO2_FAM0262	1
TE1699_SO2_FAM0264	3
TE1755_SO2_FAM2690	17
TE1763_SO2_FAM2698	4
TE1776_SO2_FAM0271	2
TE1826_SO2_FAM2757	13
TE1846_SO2_FAM2776	5
TE1871_SO2_FAM2799	1
TE1922_SO2_FAM2846	1
TE1928_SO2_FAM2851	1
TE1970_SO2_FAM0289	3
TE1982_SO2_FAM0290	2
TE2015_SO2_FAM0293	1
TE2026_SO2_FAM0294	1
TE2048_SO2_FAM0296	1
TE2080_SO2_FAM0299	3
TE2092_SO2_FAM0300	1
TE2125_SO2_FAM0303	3
TE2180_SO2_FAM0308	3
TE2211_SO2_FAM3107	23
TE2247_SO2_FAM0314	6
TE2257_SO2_FAM0318	2
TE2260_SO2_FAM0320	3
TE2283_SO2_FAM0347	25
TE2300_SO2_FAM0364	3
TE2343_SO2_FAM0405	1
TE2348_SO2_FAM0041	1
TE2354_SO2_FAM0415	11
TE2358_SO2_FAM0042	1
TE2364_SO2_FAM0426	1
TE2365_SO2_FAM0427	7
TE2381_SO2_FAM0441	1
TE2384_SO2_FAM0446	1
TE2385_SO2_FAM0447	3
TE2387_SO2_FAM0449	1
TE2389_SO2_FAM0450	5
TE2392_SO2_FAM0456	1
TE2395_SO2_FAM0459	1
TE2410_SO2_FAM0473	1
TE2413_SO2_FAM0476	3
TE2425_SO2_FAM0488	5
TE2427_SO2_FAM0490	2
TE2434_SO2_FAM0498	1
TE2441_SO2_FAM0503	1
TE2456_SO2_FAM0517	2
TE2482_SO2_FAM0543	2
TE2485_SO2_FAM0546	8
TE2486_SO2_FAM0547	2
TE2509_SO2_FAM0569	7

TE2510_SO2_FAM0057	3
TE2513_SO2_FAM0573	1
TE2519_SO2_FAM0579	1
TE2524_SO2_FAM0584	4
TE2525_SO2_FAM0585	2
TE2528_SO2_FAM0588	1
TE2531_SO2_FAM0592	35
TE2543_SO2_FAM0603	1
TE2550_SO2_FAM0611	1
TE2567_SO2_FAM0629	4
TE2576_SO2_FAM0064	1
TE2589_SO2_FAM0651	2
TE2607_SO2_FAM0668	4
TE2644_SO2_FAM0707	1
TE2659_SO2_FAM0722	1
TE2660_SO2_FAM0724	1
TE2666_SO2_FAM0073	1
TE2685_SO2_FAM0752	1
TE2699_SO2_FAM0771	1
TE2735_SO2_FAM0807	3
TE2739_SO2_FAM0810	4
TE2783_SO2_FAM0857	1
TE2785_SO2_FAM0859	1
TE2812_SO2_FAM0884	1
TE2840_SO2_FAM0091	5
TE2861_SO2_FAM0930	6
TE2877_SO2_FAM0948	2
TE2887_SO2_FAM0957	5
TE2920_SO2_FAM0099	2
TE2940_SO2_FAM1014	1
TE2941_SO2_FAM1014	5
TE2987_SO2_FAM121	4
TE2988_SO2_FAM121	21
TE2996_SO2_FAM123	3
TE2999_SO2_FAM127	1
TE3000_SO2_FAM127	1
TE3056_SO2_FAM163	42
TE3088_SO2_FAM179	1
TE3136_SO2_FAM201	2
TE3137_SO2_FAM201	1
TE3164_SO2_FAM221	1
TE3174_SO2_FAM023	1
TE3197_SO2_FAM025	1
TE3213_SO2_FAM275	3
TE3232_SO2_FAM333	1
TE3246_SO2_FAM390	3
TE3250_SO2_FAM419	4
TE3260_SO2_FAM454	2
TE3341_SO2_FAM076	16
TE3380_SO2_FAM088	4
TE3381_SO2_FAM088	1
TE3391_SO2_FAM094	7
TraJav-1.74	1
tRNA-3_DRe	1
tRNA-Ala-GCG	3

tRNA-Ala-GCY	7
tRNA-Arg-AGA	1
tRNA-Asn-AAC	4
tRNA-Asp-GAY	8
tRNA-Cys-TGY	1
tRNA-Gly-GGA	1
tRNA-Ile-ATT	2
tRNA-Lys-AAG	3
tRNA-Met_v	2
tRNA-Pro-CCA	1
tRNA-Pro-CCY	2
tRNA-Ser-AGY	4
tRNA-Ser-TCG	2
tRNA-Thr-ACA	1
tRNA-Val-GTG	4
U1	23
U13	1
U2	2
U3	4
U4	3
U5	5
U6	1
Zator-1_Etal	1
Zenon-1_Hmel_A	1
Zenon-1_Hmel_C	1
ZenoSINE1_Avan	3
ZenoSINE1_eTal	2
ZenoSINE1_hTel	10
ZenoSINE1_IDor	3
ZymTri_2023_family_141	40
ZymTri_2023_family_1487	2
ZymTri_2023_family_246	1
ZymTri_2023_family_445	3
ZymTri_2023_family_532	17

Tabela 3. Acessos Dfam identificados apenas na montagem *de novo*.

Acesso	Total
5S_DM	11
7SK	1
7SLRNA	4
Academ_S.constricta_1	2
AcoCah-1.76	1
Acol_copia_Ele12	1
Acol_copia_Ele13	1
Acol_copia_Ele22	1
Acol_copia_Ele6	9
Acol_copia_Ele9	1
Acol_CR1_Ele5	1
Acol_gypsy_Ele20	1
Acol_gypsy_Ele21	2
Acol_gypsy_Ele23	1
Acol_gypsy_Ele25	2

Acol_gypsy_Ele34	6
Acol_gypsy_Ele74	1
Acol_gypsy_Ele90	4
Acol_gypsy_Ele91	2
Acol_gypsy_Ele95	44
Acol_hAT_Ele6	2
Acol_Jockey_Ele1	97
Acol_Jockey_Ele7	1
Acol_L1_Ele2	1
Acol_mar_Ele15	4
Acol_Pao_Bel_Ele29	1
Acol_Pao_Bel_Ele33	3
Acol_Pao_Bel_Ele37	1
Acol_Pao_Bel_Ele39	1
Acol_Pao_Bel_Ele41	3
Acol_Pao_Bel_Ele47	13
Acol_Pao_Bel_Ele51	2
Acol_Pao_Bel_Ele62	1
Acol_Pao_Bel_Ele64	1
Acol_Pao_Bel_Ele75	1
Acol_Pao_Bel_Ele81	2
Acol_PIF_Ele1	33
Acol_PIF_Ele10	4
Acol_R1_Ele1	3
Acol_R1_Ele16	2
Acol_RTE_Ele1	2
ALR	1
AluSc8	1
AluSp	1
AluY	1
BEL-65_DRe-I	1
BEL11-I_DR	1
BGLII_B2	1
Bica_I	1
Blastopia_I	8
CfSE1	13
CheMed-1.218_int	9
Chlamys-1_bBra	1
Chlamys-4_cRei	5
Chlamys-5_cSch	
Chlamys-6_cBra	8
CMC_M.nervosa_3	1
COMP-subunit_TAF11_rnd-6_family-27360	1
Copia-1_Dmel-I	4
Copia-105_rCom	2
Copia-2_Hm	1
Copia-3_DM-I	2
Copia-3_INT_AstCal	5
Copia-7-I_DR	12
Copia-8-I_DR	6
CR1-1_Hmel_A	1
CR1-1_Hmel_B2	1
CR1-4_Hmel_A	3
CR1-8_Hmel_A	1

CR1-8_Hmel_B	2
CR1-8_Hmel_C	2
CR1-9_Hmel_A	1
CR1-9_Hmel_B	3
CR1-Zenon-1_Etal	1
CR1-Zenon-1_Ldor	1
CR1-Zenon-1a_Avan	51
CR1-Zenon-1b_Avan	3
CR1-Zenon-2_Diul	253
CR1-Zenon-3_Diul	1
CR1-Zenon-4_Diul	76
CraTho-1.264	1
CraTho-1.281	1
CraTho-6.8701a	14
DF003447270	1
DF003447511	6
DF003447527	1
DF003447577	1
DF003447704	1
DF003447766	1
DF003448229	1
DF003448318	24
DF003448584	1
DF003448600	30
DF003448838	3
DF003448853	1
DF003448865	1
DF003449051	1
DF003449069	1
DF003449136	1
DF003449138	1
DF003449172	10
DF003449224	23
DF003449493	9
DF003449516	3
DF003449575	2
DF003449741	2
DF003449851	1
DF003449868	8
DF003449931	1
DF003449993	4
DF003450149	1
DIRS-15B_DR	1
DIRS-4_DR	40
DIRS-8_DR	15
DMSAT6	2
DNA-1_Hbes	4
DNA-1_Hmel	5
DNA-2-15_DR	1
DNA-2-24_DR	1
DNA-3-5_DR	1
DNA-like-1_Hmel	7
DNA-like-11_Hmel	8
DNA-like-6_Hmel	2

DNA1_Htel	3
DNA16_Hera	1
DNA16_Htel	1
DNA20_Hera	6
DNA25_Hera	1
DNA6-2_DR	1
DNA8-25_DR	1
DOC3_DM	1
DOC5_DM	4
Dong-R4-1_Hcyd	2
Dong-R4-1_Ldor	8
Dong-R4-2_AstCal	2
EnSpm-9_DR	1
EubJap-5.255	12
EulFla-1.126	1
FelNig-1.24	1
FW_DM	1
GalVar-1.127	1
Gypsy-100_INT_AstCal	2
Gypsy-107_INT_AstCal	2
Gypsy-11_Hmel_I	5
Gypsy-13_Hmel_I	1
Gypsy-14_Hmel_I	4
Gypsy-187_DR-I	1
Gypsy-206_DR-I	2
Gypsy-208_DR-I	12
Gypsy-217_DR-I	1
Gypsy-25_INT_AstCal	2
Gypsy-27_INT_AstCal	6
Gypsy-31_INT_AstCal	4
Gypsy-31-LTR_DR	2
Gypsy-45_INT_AstCal	2
Gypsy-46_rCom	35
Gypsy-48_rCom	3
Gypsy-49_INT_AstCal	1
Gypsy-87_INT_AstCal	3
Gypsy-88_INT_AstCal	2
Gypsy158-I_DR	1
Gypsy163-I_DR	1
Gypsy56-I_DR	3
Gypsy58-I_DR	1
Gypsy86-I_DR	1
Gypsy98-LTR_Dr	2
Gypsy99-I_DR	4
Halyomorpha_halys-1_family-166	2
Halyomorpha_halys-1_family-177	1
Halyomorpha_halys-1_family-260	1
Halyomorpha_halys-1_family-321	1
Halyomorpha_halys-1_family-349	2
Halyomorpha_halys-1_family-43	3
Halyomorpha_halys-1_family-449	4
Halyomorpha_halys-1_family-520	2
Halyomorpha_halys-1_family-724	1
Halyomorpha_halys-1_family-733	1

Halyomorpha_halys-4_family-638	43
Halyomorpha_halys-5_family-1188	2
Halyomorpha_halys-5_family-1934	2
Halyomorpha_halys-5_family-2362	2
Halyomorpha_halys-5_family-3015	2
Halyomorpha_halys-5_family-466	3
Halyomorpha_halys-5_family-682	3
Halyomorpha_halys-6_family-1101	1
Halyomorpha_halys-6_family-2394	10
Halyomorpha_halys-6_family-713	1
Halyomorpha_halys-6_family-945	1
hAT_C.sinensis_2	1
hAT_M.nervosa_33	2
hAT-N22_DR	2
hAT-N85_DR	1
Helitron-1_Diul	45
Helitron-1_Etal	51
Helitron-1_Hdem	2
Helitron-10_Diul	2
Helitron-13_Diul	35
Helitron-2_Avan	25
Helitron-2_Diul	89
Helitron-2_Hsar	2
Helitron-3_Etal	14
Helitron-3_Hnum	18
Helitron-4_Etal	1
Helitron-4_Hbur	1
Helitron-4_Hera	2
Helitron-6_Diul	3
Helitron-7_Etal	5
Helitron-7_Hpar	1
Helitron-like-1_Ldor	2
Helitron-like-8_Hmel	4
Helitron10_Hera	2
Helitron2_Hera	16
Helitron6_Hera	110
Helitron7_Hera	22
Helitron8_Hera	16
HemHyl-2.7	2
HipAmp-1.111	1
HSATII	1
I-1_Avan	15
I-1_Hmel_D	4
I-1_Hmel_E	4
I-1_Hmel_F	10
I-1_Hmel_G	1
I-1_Hmel_H	11
I-1_Hmel_I	1
I-1_Hmel_N	4
I-1_Hmel_O	1
I-Nimb-1a_Hcyd	1
IMPB_01	3
IndInd-1.118	35
Invader1_I	1

Invader4f_I	2
Invader7_I	4
Jockey-2_Hbur	27
Jockey-2_Hmel_B	6
Jockey-3_Hmel_D	7
Jockey-3_Hmel_E	26
L1-14_AstCal	1
L1-22_rCom	4
L1-Tx1-1_Diul	31
L2-1_Diul	1
L2-1_Etal	1
L2-1_Hbur	9
L2-1_Hmel_B	1
L2-1_Ldor	6
L2-1_Ldor_D	3
L2-1_Ldor_E	8
L2-1_Ldor_F	16
L2-1_Ldor_G	1
L2-11_Hmel_A1	1
L2-11_Hmel_A2	1
L2-14_Hmel_B2	7
L2-15_Hmel_A2	3
L2-15_Hmel_B1	1
L2-1a_Hnum	1
L2-1b_Hnum	1
L2-2_Hbes	1
L2-3_Hmel_A1	2
L2-3_Hmel_A3	8
L2-3_Hmel_C1	1
L2-4_Hmel_A	1
L2-4_Hmel_C1	2
L2-5_Hmel_A	9
LasBor-1.134	2
LasBor-4.3117	2
LasBor-5.38	5
LaSE1	1
LINE_A.kagoshimensis_13_cons	4
LINE_C.sinensis_1_cons	1
LINE_C.sinensis_4_cons	1
LINE_C.sinensis_71_cons	24
LINE_D.rostriformis_104_cons	1
LINE_D.rostriformis_4_cons	1
LINE_D.rostriformis_99_cons	4
LINE_M.mercenaria_187_cons	3
LINE_M.nervosa_12_cons	1
LINE_M.nervosa_19_cons	7
LINE_M.phylippinarum_13_cons	1
LINE_M.phylippinarum_93_cons	2
LINE_M.yessoensis_57_cons	41
LINE_P.fucata_18_cons	1
LINE_P.maximus_108_cons	3
LINE_P.maximus_28_cons	1
LINE_S.broughtoni_10_cons	2
LINE_S.broughtoni_19_cons	6

LINE_S.constricta_1_cons	1
LINE_S.grandis_18_cons	1
LINE_S.grandis_40_cons	1
LINE_T.granosa_24_cons	1
LSU-rRNA_Cel	80
LSU-rRNA_Dme	69
LSU-rRNA_Hsa	40
LTR-4_Hm	4
LTR75B	1
MARINCE1	8
Mariner-1_Hmel	3
Mariner-1_Hpar	1
Mariner-12_DR	1
Mariner-4_DR	1
Mariner-N5B_DR	1
Mariner1_Hera	38
Mariner5_Hera	3
MegLyr-5.207b	2
MegLyr-5.2125	1
Merlin_M.phylippinarum_6	1
Merrow_Cpb	3
MesAur-1.193	1
Micropia2_DM-I	2
Micropia7_Dmel_I	5
MinSch-5.2265	1
MirCoq-1.117	6
MMSAT4	212
MsSE2	3
MurFea-5.2046	1
MurFea-6.1933	7
MyoBra-1.237	4
MyoBra-1.328	3
MyoBra-1.435	14
NA-DNA-15_AstCal	16
nhAT-1_Diul	12
nhAT-1_Hdem	3
nhAT-1_Hhim	17
nhat-1_Hmel	2
nhAT3_mMyo	2
Nimb-12_DR	1
nMar-1_Etal	9
nMar-1_Hhim	1
nMar-1_Ldor	9
nMar-11_Hmel	1
nMar-16_Hele	4
nMar-16_Hpar	3
nMar-21_Hmel	2
nMar-4_Hmel	1
nMar-7_Hmel	3
nMar-8_Hmel	3
nMar-8_Hnum	2
nMar-9_Hmel	25
nMar1_rFer	7
nMar5_rFer	1

NomLeu-5.736_int	1
npiggyBac-5_Hmel	5
nTc3-3_Hcyd	4
ObSE1	1
OnyTor-1.356_int	3
PanOnc-1.30	1
PanPar-1.29	1
PanTig-1.13	4
Pao-32_INT_AstCal	6
Pao-4_INT_AstCal	1
Pao-41_INT_AstCal	3
Pao-42_INT_AstCal	1
ParHer-1.13	1
Penelope-1_Avan	1
Penelope-3_Avan	13
PIF-Harbinger_C.gigas_20	1
PIF-Harbinger_C.gigas_21	1
PIF-Harbinger_D.rostriformis_14	2
PIF-Harbinger_M.coruscus_15	3
PIF-Harbinger_S.constricta_37	3
PIF-Harbinger_S.glomerata_8	20
PiggyBac_A.i.concentricus_4	1
PiggyBac_A.marissinica_178	1
PiggyBac_B.platifrons_29	7
PiggyBac_C.ariakensis_10	2
PipPip-5.1542	8
PLE-X1_cBra	1
PLE-X2_cBra	1
Pman_ALU_5.1881	1
Pman_ALU_6.2822	4
Pman_ALU_6.4482	4
Pman_ERVL_5.247	3
Pman_ERVL-MALR_6.16429	4
Pman_MIR_6.141	6
Pman_SNRNA_5.818	5
Pman_SNRNA_6.7824	2
Pman_UNKNOWN_6.2013	4
Polinton-1_DR	4
Proto2-1_AstCal	3
PumCon-1.11	1
PxSE2	20
R1-1_Hmel_C	20
R1-2_Hmel_A	26
R4-1_HMe_A	5
R4-1_HMe_B	4
R4-1_HMe_C	1
RhiPru-5.1045	1
RhiSin-1.210	1
RLTR20A4	2
RLTR31D_MM	3
RTE-1_Avan	60
RTE-1_Heca	12
RTE-1_HMe_A1	2
RTE-1_HMe_A2	14

RTE-1_HMe_B	11
RTE-1_Htim	29
RTE-13_Hmel_A	4
RTE-13_Hmel_B	8
RTE-16_Hmel_A	3
RTE-16_Hmel_B	6
RTE-16_Hmel_C	3
RTE-16_Hmel_D	7
RTE-16_Hmel_E	1
RTE-16_Hmel_F	5
RTE-18_Hmel_A	2
RTE-2_Hmel_A	1
RTE-2_Hmel_A2	6
RTE-2_Hmel_B	1
RTE-20_Hmel_C	2
RTE-20_Hmel_D	2
RTE-20_Hmel_E	13
RTE-20_Hmel_F	7
RTE-3_Hcyd	6
RTE-3_Hmel_C	1
RTE-4_Hmel_C	3
RTE-4_Hmel_D	1
RTE-5_Hmel_C	1
RTE-6_Hmel_B	3
RTE-6_Hmel_C2	7
RTE-6_Hmel_D	6
RTE-6_Hmel_E	13
RTE-6_Hmel_H1	1
RTE-6_Hmel_H2	1
RTE-6_Hmel_I	7
RTE-7_Hmel_B	5
RTE-7_Hmel_C	9
RTE-7_Hmel_E	3
RTE-BovB-1_Heca	1
RTE-BovB-4_AstCal	9
SART1_R1_A	4
SAT-15_DR	1
Sola1_B.platifrons_33	10
Sola2_A.marissinica_189	3
SorAra-1.36	1
Ssa_0070	8
Ssa_0264	49
Ssa_0972	1
Ssa_1250	1
Ssa_1657	3
Ssa_1781	1
Ssa_1959	2
SSU-rRNA_Dme	4
SSU-rRNA_Hsa	72
SVA_A	3
TadBra-1.184	1
TadBra-6.2216_int	3
TcMar_A.kagoshimensis_27	4
TcMar_A.kagoshimensis_32	1

TcMar_A.kagoshimensis_36	2
TcMar_B.platifrons_48	1
TcMar_C.sinensis_15	1
TcMar_M.mercenaria_15	2
TcMar_M.mercenaria_38	14
TcMar_M.phylippinarum_21	2
TcMar_P.maximus_23	3
TcMar_R.philippinarum_24	1
TcMar_S.broughtoni_19	2
TcMar_T.granosa_20	1
TcMar-ISRm11-4_AstCal	1
TE-X-12_DR	1
TE0030_SO2_FAM1029	1
TE0042_SO2_FAM1041	2
TE0050_SO2_FAM0105	8
TE0053_SO2_FAM1052	1
TE0075_SO2_FAM1072	3
TE0096_SO2_FAM1094	6
TE0102_SO2_FAM0011	1
TE0105_SO2_FAM1103	3
TE0118_SO2_FAM1117	2
TE0121_SO2_FAM0112	4
TE0123_SO2_FAM1121	6
TE0169_SO2_FAM1167	1
TE0181_SO2_FAM1178	12
TE0183_SO2_FAM0118	4
TE0188_SO2_FAM1184	7
TE0200_SO2_FAM1198	2
TE0203_SO2_FAM0120	2
TE0251_SO2_FAM0125	14
TE0281_SO2_FAM1278	2
TE0288_SO2_FAM1285	5
TE0297_SO2_FAM1294	4
TE0298_SO2_FAM1295	3
TE0300_SO2_FAM1297	2
TE0303_SO2_FAM0013	1
TE0314_SO2_FAM0131	1
TE0320_SO2_FAM1315	3
TE0331_SO2_FAM1325	6
TE0366_SO2_FAM1359	7
TE0367_SO2_FAM0136	28
TE0376_SO2_FAM1368	2
TE0399_SO2_FAM0139	1
TE0410_SO2_FAM0140	2
TE0415_SO2_FAM1404	3
TE0440_SO2_FAM0143	7
TE0451_SO2_FAM1440	1
TE0458_SO2_FAM1447	11
TE0459_SO2_FAM1448	1
TE0483_SO2_FAM1470	10
TE0519_SO2_FAM1504	1
TE0533_SO2_FAM1518	1
TE0535_SO2_FAM0152	37
TE0553_SO2_FAM0154	1

TE0555_SO2_FAM1541	2
TE0573_SO2_FAM1559	7
TE0578_SO2_FAM1563	1
TE0595_SO2_FAM0158	1
TE0597_SO2_FAM1581	1
TE0605_SO2_FAM0159	3
TE0616_SO2_FAM0016	1
TE0624_SO2_FAM1606	4
TE0639_SO2_FAM0162	1
TE0749_SO2_FAM0173	3
TE0753_SO2_FAM1733	14
TE0770_SO2_FAM0175	13
TE0790_SO2_FAM0177	7
TE0804_SO2_FAM1783	11
TE0917_SO2_FAM1896	1
TE1008_SO2_FAM1986	3
TE1045_SO2_FAM0202	7
TE1178_SO2_FAM0215	5
TE1210_SO2_FAM0218	4
TE1214_SO2_FAM2183	3
TE1233_SO2_FAM0220	8
TE1274_SO2_FAM0224	21
TE1275_SO2_FAM2240	2
TE1296_SO2_FAM0226	15
TE1299_SO2_FAM2262	1
TE1307_SO2_FAM0227	3
TE1316_SO2_FAM2278	12
TE1346_SO2_FAM2309	3
TE1356_SO2_FAM2319	8
TE1373_SO2_FAM2336	1
TE1377_SO2_FAM0234	3
TE1379_SO2_FAM2341	1
TE1382_SO2_FAM2344	2
TE1397_SO2_FAM2358	1
TE1409_SO2_FAM0237	15
TE1416_SO2_FAM2376	4
TE1452_SO2_FAM2410	3
TE1499_SO2_FAM2455	1
TE1521_SO2_FAM2475	2
TE1592_SO2_FAM2540	7
TE1645_SO2_FAM0259	3
TE1671_SO2_FAM2614	5
TE1677_SO2_FAM0262	2
TE1699_SO2_FAM0264	14
TE1721_SO2_FAM0266	7
TE1742_SO2_FAM2679	1
TE1755_SO2_FAM2690	37
TE1763_SO2_FAM2698	2
TE1826_SO2_FAM2757	28
TE1829_SO2_FAM0276	5
TE1846_SO2_FAM2776	9
TE1861_SO2_FAM0279	1
TE1871_SO2_FAM2799	6
TE1922_SO2_FAM2846	7

TE1928_SO2_FAM2851	4
TE1970_SO2_FAM0289	1
TE1982_SO2_FAM0290	8
TE2015_SO2_FAM0293	2
TE2058_SO2_FAM0297	1
TE2065_SO2_FAM2976	12
TE2080_SO2_FAM0299	5
TE2125_SO2_FAM0303	1
TE2136_SO2_FAM0304	2
TE2147_SO2_FAM0305	1
TE2160_SO2_FAM3061	5
TE2169_SO2_FAM0307	1
TE2180_SO2_FAM0308	6
TE2211_SO2_FAM3107	82
TE2241_SO2_FAM3134	1
TE2247_SO2_FAM0314	1
TE2257_SO2_FAM0318	7
TE2260_SO2_FAM0320	5
TE2262_SO2_FAM0322	1
TE2266_SO2_FAM0327	1
TE2270_SO2_FAM0331	6
TE2282_SO2_FAM0346	3
TE2283_SO2_FAM0347	10
TE2294_SO2_FAM0359	1
TE2308_SO2_FAM0372	1
TE2328_SO2_FAM0391	1
TE2332_SO2_FAM0395	2
TE2354_SO2_FAM0415	9
TE2358_SO2_FAM0042	7
TE2364_SO2_FAM0426	6
TE2365_SO2_FAM0427	6
TE2381_SO2_FAM0441	7
TE2385_SO2_FAM0447	3
TE2387_SO2_FAM0449	3
TE2388_SO2_FAM0045	1
TE2389_SO2_FAM0450	12
TE2392_SO2_FAM0456	2
TE2397_SO2_FAM0460	2
TE2412_SO2_FAM0475	8
TE2413_SO2_FAM0476	2
TE2419_SO2_FAM0482	1
TE2425_SO2_FAM0488	6
TE2434_SO2_FAM0498	1
TE2453_SO2_FAM0514	2
TE2456_SO2_FAM0517	2
TE2463_SO2_FAM0525	2
TE2465_SO2_FAM0527	2
TE2475_SO2_FAM0537	2
TE2479_SO2_FAM0540	1
TE2485_SO2_FAM0546	1
TE2486_SO2_FAM0547	6
TE2492_SO2_FAM0553	1
TE2509_SO2_FAM0569	8
TE2510_SO2_FAM0057	1

TE2513_SO2_FAM0573	5
TE2519_SO2_FAM0579	4
TE2520_SO2_FAM0058	2
TE2524_SO2_FAM0584	2
TE2531_SO2_FAM0592	38
TE2549_SO2_FAM0610	2
TE2550_SO2_FAM0611	2
TE2552_SO2_FAM0613	1
TE2566_SO2_FAM0628	4
TE2567_SO2_FAM0629	11
TE2577_SO2_FAM0640	1
TE2607_SO2_FAM0668	2
TE2666_SO2_FAM0073	2
TE2685_SO2_FAM0752	2
TE2699_SO2_FAM0771	1
TE2711_SO2_FAM0782	1
TE2735_SO2_FAM0807	1
TE2739_SO2_FAM0810	1
TE2776_SO2_FAM0850	1
TE2783_SO2_FAM0857	2
TE2785_SO2_FAM0859	6
TE2812_SO2_FAM0884	2
TE2818_SO2_FAM0089	11
TE2831_SO2_FAM0901	4
TE2840_SO2_FAM0091	3
TE2843_SO2_FAM0913	1
TE2850_SO2_FAM0092	2
TE2861_SO2_FAM0930	4
TE2863_SO2_FAM0933	1
TE2877_SO2_FAM0948	2
TE2887_SO2_FAM0957	31
TE2912_SO2_FAM0982	6
TE2920_SO2_FAM0099	8
TE2940_SO2_FAM1014	1
TE2941_SO2_FAM1014	7
TE2987_SO2_FAM121	1
TE2988_SO2_FAM121	14
TE2996_SO2_FAM123	1
TE3025_SO2_FAM147	1
TE3056_SO2_FAM163	95
TE3064_SO2_FAM1643	1
TE3137_SO2_FAM201	1
TE3164_SO2_FAM221	3
TE3184_SO2_FAM233	5
TE3246_SO2_FAM390	4
TE3250_SO2_FAM419	2
TE3260_SO2_FAM454	2
TE3264_SO2_FAM465	1
TE3317_SO2_FAM701	3
TE3341_SO2_FAM076	37
TE3379_SO2_FAM088	1
TE3380_SO2_FAM088	14
TE3381_SO2_FAM088	5
TE3391_SO2_FAM094	2

tRNA-Ala-GCY	1
tRNA-Asn-AAC	1
tRNA-Lys-AAG	2
tRNA-Met_v	1
U1	3
U17	2
U3	4
U6	1
Zaphod	1
ZapHud-1.7	1
Zator-1_Etal	1
Zenon-1_Hmel_A	7
Zenon-1_Hmel_B	5
Zenon-1_Hmel_D	9
ZenoSINE1_Avan	1
ZenoSINE1_eTal	2
ZenoSINE1_hTel	5
ZenoSINE1_IDor	2
ZymTri_2023_family_141	48
ZymTri_2023_family_1487	2
ZymTri_2023_family_246	1
ZymTri_2023_family_365	1
ZymTri_2023_family_532	44
ZymTri_2023_family_717	1
