Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ
Pós-Graduação em Biologia Celular e Molecular

SAMARA GRACIANE DA COSTA LATGÉ

Caracterização de α-glicosidases de *Lutzomyia longipalpis*: atividades enzimáticas, expressão gênica e modulação por *Leishmania mexicana*

Orientador: Prof. Dr. Fernando Ariel Genta

RIO DE JANEIRO
2018

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ
Pós-Graduação em Biologia Celular e Molecular

SAMARA GRACIANE DA COSTA LATGÉ

Caracterização de alfa-glicosidases de *Lutzomyia longipalpis*: atividades enzimáticas, expressão gênica e modulação por *Leishmania mexicana*

Tese apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Doutor em Biologia Celular e Molecular

Orientador: Prof. Dr. Fernando Ariel Genta

RIO DE JANEIRO
2018

Latgé, Samara Graciane da Costa.

Caracterização de alfa-glicosidases de *Lutzomyia longipalpis*: atividades enzimáticas, expressão gênica e modulação por *Leishmania mexicana* / Samara Graciane da Costa Latgé. - Rio de janeiro, 2018.
XIII, 218f f.; il.

Tese (Doutorado) – Instituto Oswaldo Cruz, Pós-Graduação em Biologia Celular e Molecular, 2018.

Orientador: Fernando Ariel Genta.

Bibliografia: f. 207-218

1. *Lutzomyia longipalpis*. 2. Alfa-glicosidase. 3. Digestão. 4. *Leishmania mexicana*. 5. Flebotomíneo. I. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da Biblioteca de Manguinhos/ICICT com os dados fornecidos pelo(a) autor(a).

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ
Pós-Graduação em Biologia Celular e Molecular

SAMARA GRACIANE DA COSTA LATGÉ

Caracterização de alfa-glicosidases de *Lutzomyia longipalpis*: atividades enzimáticas, expressão gênica e modulação por *Leishmania mexicana*

Orientador: Prof. Dr. Fernando Ariel Genta

Aprovada em: 09/11/2018.

EXAMINADORES:
Prof. Dr. Reginaldo Peçanha Brazil – Presidente (IOC/FIOCRUZ)
Prof. Dra. Maria Helena Neves Lobo Silva Filha (CPqIAM/FIOCRUZ)
Prof. Dr. Nelder de Figueiredo Gontijo (UFMG)
Prof. Dra. Luciana Ordunha Araripe (IOC/FIOCRUZ)
Prof. Dra. Patrícia Fampa Negreiros Lima (UFRRJ)

Rio de Janeiro, 09 de novembro de 2018

Dedico esta tese com todo carinho
aos meus pais e especialmente ao
meu marido Lucas

Agradecimentos

Agradeço à Deus por ter iluminado meus passos durante toda essa jornada e por ter trazido clareza aos meus pensamentos diante de todas as dificuldades.

Ao meu orientador, Dr. Fernando Genta, pela confiança em meu trabalho, pelas discussões científicas, ensinamentos, troca de idéias e por todo apoio. Obrigada pela independência, isso possibilitou o meu desenvolvimento como profissional e obrigada pelo tempo dedicado a este trabalho, sem você a realização não seria possível.

A Dr. Luciana Ordunha Araripe que se disponibilizou para ajudar nas correções deste trabalho e aos componentes da banca que generosamente diponibilizaram seu tempo e aceitaram o convite de avaliar a minha tese.

Ao Dr. Rod Dillon e Dr. Paul Bates que me receberam na Universidade de Lancaster/UK, durante o período de doutorado sanduíche, “You are amazing”.

Ao Roberto Silva, Michele Catunda, Caroline Moraes, Lígia Barizon e Kel Martins pela amizade e apoio durante o tempo que estive longe da minha família. A amizade de vocês tornaram os dias mais felizes.

Ao Roberto Silva, pela troca de conhecimentos e discussão de idéias durante o doutorado sanduíche.

Aos colegas do LABFISI, vocês são sensacionais, obrigada por toda força, apoio nas horas difíceis, e por toda amizade. Obrigada por todas as discussões de trabalho e idéias.

Em especial gostaria de agradecer a Carol Moraes, Tainá Neves e Joyce Ferreira pelo suporte e dedicação na colônia de flebotomíneos. Sem vocês esse trabalho não seria possível, são muitas horas semanais de trabalho cuidadoso.

Agradeço a todos os professores que passaram por minha vida acadêmica, pelos ensinamentos.

A minha mãe e a meu pai, Mauren e Antônio, e a meu irmão Marcelo, vocês são os pilares da minha vida, sem o apoio de vocês eu não teria chegado até aqui. Obrigado por terem acreditado no meu sonho e por terem apoiado minha decisões, mesmo quando cada uma delas me levava para mais longe de vocês.

Ao Lucas, meu marido, pelo apoio incondicional, pelas palavaras de carinho, pela paciência, pelo amor, carinho, compreensão. Você tornou esse caminho muito mais ameno, obrigada pelas palavras de apoio nos momentos difíceis. Obrigada por existir em minha vida!

Agradeço a Capes, CNPq, FAPERJ, FIOCRUZ pelo suporte financeiro.

Enfim agradeço a todos que de alguma forma contribuíram para meu crescimento profissional e pessoal.

“O sucesso é a capacidade de ir de um fracasso ao outro sem perder o entusiasmo”

Winston Churchill

Índice

LISTA DE FIGURAS

LISTA DE TABELAS

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Caracterização de α-glicosidases de *Lutzomyia longipalpis*: atividades enzimáticas, expressão gênica e modulação por *Leishmania mexicana*

RESUMO

TESE DE DOUTORADO

Samara Graciane da Costa Latgé

Lutzomyia longipalpis é o vetor de *Leishmania infantum*, um dos agentes causadores da leishmaniose visceral. Os flebotomíneos adultos possuem uma dieta rica em açúcares, essencial para atender às demandas energéticas necessárias ao desenvolvimento; as fêmeas, além de açúcares, também se alimentam de sangue. As α-glicosidases e α-amilases estão envolvidas na digestão de carboidratos adquiridos na dieta, e são classificadas nas famílias 13 e 31 das glicosídeo hidrolases. Para determinação da atividade de α-glicosidase utilizando o substrato MUαGlu padronizamos uma técnica de ensaio contínuo para realização de ensaios em amostras teciduais individuais. Quantificamos a atividade de α-glicosidase em diferentes tecidos, ressaltando a atividade presente no divertículo e no intestino médio de *L. longipalpis*, sendo essa enzima mais ativa sobre sacarose que sobre o substrato MUαGlu. Atividades basais foram observadas em insetos não alimentados; a alimentação sanguínea induz a atividade no conteúdo do intestino médio e a alimentação açucarada modula a atividade nos tecidos do intestino médio. A exposição a diferentes concentrações ou moléculas de açúcares também alterou a atividade. As α-glicosidases de diferentes tecidos apresentaram diferentes propriedades bioquímicas, como pH ótimo entre 7,0 - 8,0, K_M entre 0,37 - 4,7 mM (MUαGlu como substrato), pH ótimo de 6,0 e K_M entre 11 - 800 mM (sacarose como substrato). Enzimas do divertículo e do tecido do intestino médio apresentaram inibição em altas concentrações de substrato (sacarose), o que explica os altos valores de K_M encontrados. No genoma de *L. longipalpis* foram encontradas nas famílias GH13 e GH31 proteinas envolvidas no metabolismo de açúcar, transporte de aminoácidos, armazenamento e mobilização das reservas de glicogênio e regulação da miogênese. Descrevemos também uma α-glicosidase neutra (controle de N-glicosilação) e uma α-glicosidase lisossomal inativa (sem resíduos catalícos). A estrutura e as funções dessas proteínas identificadas são conservadas. Uma análise comparativa demonstrou retração no número de genes de maltases e expansão das α-amilases, com organização em dois grandes clusters. Maltases são mais expressas no intestino médio de fêmeas alimentadas com sangue, e a infecção por *L. mexicana* modulou negativamente a sua expressão gênica. Em fêmeas alimentadas com sangue (sem sacarose), as taxas de infecção por *L. mexicana* e migração para a região da cárdia não foram afetadas, sugerindo que a migração não é inteiramente baseada nos estímulos de quimiotaxia pelas moléculas de açúcar. Contudo, a ausência de açúcar na dieta do inseto resultou no desenvovimento de *L. mexicana* no intestino posterior. Todos esses resultados sugerem que as α-amilases evoluíram para assumir diferentes funções além da hidrólise de açúcares primários e que as α-glicosidases também estão envolvidas em diferentes processos metabólicos, como digestão de açúcares vegetais, digestão de glicoproteínas ou glicolipídios sanguíneos e mobilização de estoques energéticos.

Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Caracterização de alfa-glicosidases de *Lutzomyia longipalpis*: atividades enzimáticas, expressão gênica e modulação por *Leishmania mexicana*

RESUMO

TESE DE DOUTORADO

Samara Graciane da Costa Latgé

Lutzomyia longipalpis is the vector of the parasite *Leishmania infantum*, the causing agent of visceral leishmaniasis. Adult sand flies have a diet rich in sugars, essential to meet the energy demands for their development and females, besides of sugar, also feed on blood. α-glycosidases and α-amylases are involved in the digestion of dietary carbohydrates and are classified in families 13 and 31 of glycoside hydrolases. For determination of α-glucosidase activity using MUαGlu as substrate, we standardized a continuous assay technique for performing assays on individual tissue samples. α-glucosidase activity was quantified in different tissues. We described the activity present in the crop and midgut of *L. longipalpis*, being more active against sucrose than MUαGlu. Basal activities were observed in non-fed insects; blood feeding induced activity in the midgut contents, and sugar feeding modulated activity in the midgut tissues. Exposure to different sugar concentrations or moieties also changed α-glucosidase activity. α-glucosidases from different tissues showed different biochemical properties, with an optimum pH around 7.0 - 8.0, K_M between 0.37 - 4.7 mM (MUαGlu as substrate), optimum pH around 6.0, and K_M between 11 - 800 mM (sucrose as substrate). Enzymes from crop and midgut tissues showed inhibition in high substrate concentrations (sucrose), which explains the high K_M values found. In the *L. longipalpis* genome we identified proteins in the GH13 and GH31 families that are involved in sugar metabolism, amino acid transport, storage and mobilization of glycogen reserves, and myogenesis regulation. We also described a neutral α-glycosidase (control of N-glycosylation) and an inactive lysosomal α-glycosidase (no catalytic residues). Structure and functions of identified proteins are conserved. A comparative analysis showed a retraction in the number of maltases genes and expansion of α-amylases, with organization in two large clusters. This expansion is probably related to the specialization of these proteins to hydrolyze different substrates. α-amylases have different expression patterns depending on feeding condition and tissue analyzed. Maltases are higher expressed in midgut tissues of blood-fed females, and their gene expression was negatively modulated with *L. mexicana* infection. In blood-fed females (no sugar), the infections rates by *L. mexicana* and migration to cardia region were not affected, suggesting that migration is not entirely based on the chemotaxis stimuli by sugar molecules. However, sugar deprivation resulted in the development of *L. mexicana* in the sand fly hindgut. All these results suggest that α-amylases evolved to assume different functions other than hydrolysis of primary sugars and that α-glucosidases are also involved in different metabolic processes, like digestion of plant sugars, digestion of blood glycoproteins or glycolipids, and mobilization of energetic storages during starvation.

1. Introdução

1.1. Leishmanioses

As leishmanioses estão entre as principais doenças negligenciadas em todo mundo, sendo estas doenças infecto-parasitárias causadas por tripanossomatídeos do gênero *Leishmania*, transmitidos através da picada de fêmeas de flebotomíneos do gênero *Phlebotomus* e *Lutzomyia* durante a alimentação sanguínea (ANTINORI; SCHIFANELLA; CORBELLINO, 2012; COLWELL; DANTAS-TORRES; OTRANTO, 2011; KAYE; SCOTT, 2011). Atualmente mais de 98 espécies de flebotomíneos são consideradas de importância médica (WHO, 2017).

Mais de 20 espécies de *Leishmania* podem causar as leishmanioses. A infecção resultante depende da espécie infectante e também da resposta imune do hospedeiro. As leishmanioses manifestam-se em quatro formas principais, a leishmaniose cutânea (LC), mucocutânea (LMC), visceral ou calazar (LV) e leishmaniose dermal pós calazar (LDPC) (WHO, 2017). Enquanto a leishmaniose cutânea é a forma mais comum da doença, a LV é a forma mais grave e pode levar à morte se não for tratada. Além disso as leishmanioses ainda podem ser classificadas como zoonóticas ou antroponóticas dependendo se o reservatório é um animal ou o homem.

No Velho Mundo (Europa, Ásia e África), a maioria das transmissões do parasito *Leishmania* ocorre peridomesticamente em áreas semi-áridas modificadas por humanos. No Novo Mundo (Américas), a transmissão da leishmaniose cutânea inicialmente associada a habitats silváticos, ocorre em ambientes domésticos. A transmissão de parasitos causadores da leishmaniose visceral além de ocorrer em áreas rurais, também está associado a ambientes urbanos e periurbanos (AKHOUNDI et al., 2016; BRAZIL, 2013). A transmissão do parasito por um vetor pode ocorrer entre animais silvestres, de animais para o homem ou entre pessoas (AKHOUNDI et al., 2016).

Existem mais de 1 billhão de pessoas vivendo em áreas de risco (616 milhões para LV e 431 milhões para LC). Em 2015, foram mais de 20.000 mortes causadas por leishmaniose visceral em todo mundo, sendo estimados mais de 300.000 casos. De 2010 a 2015 foram reportados mais de 1 milhão de casos de leishmaniose cutânea. Em 2015, 87 países foram considerados endêmicos para LC e 75 endêmicos para LV (WHO 2017). Nas Américas, estas doenças possui ampla distribuição geográfica e elevada incidência.

Diversos esforços são realizados para que, nos países nos quais as leishmanioses são um problema substancial de saúde pública, sejam adotados programas nacionais de controle, elaborando orientações, sistemas de vigilância, coleta e análise de dados, tratamento, reabilitação e prevenção, como por exemplo, o Plano de Ação de Leishmanioses nas Américas

2017-2022 . Na Ásia, desde 2008, devido à implementação do programa de eliminação de calazar, o número de casos de LV em 5 países da região sudeste tem diminuído pronunciadamente. No Brasil os casos de LC diminuíram 34 % em 2016, no entanto nas Américas houve um aumento de 5 %, o que pode ter ocorrido devido a melhora no sistema de monitoramento e reportagem dos dados (PAHO, 2018; WHO 2017).

1.1.1. Leishmaniose Cutânea

A leishmaniose cutânea é a forma menos severa e mais comum da doença, ela afeta a camada mais externa da pele causando diversos tipos de lesões. De acordo com as manifestações clínicas e lesões, a LC pode ser classificada em leishmaniose cutânea localizada, com formas típicas e atípicas e também classificada como leishmaniose cutânea disseminada, que pode se apresentar na forma de nódulos ou difusa, com recidiva na cútis ou região mucocutânea (AKILOV; KHACHEMOUNE; HASAN, 2007; ANTINORI; SCHIFANELLA; CORBELLINO, 2012; MCGWIRE; SATOSKAR, 2014; Figura 1). A forma mais simples de leishmaniose cutânea se apresenta na forma de úlceras ou lesões nodulares onde ocorreu a picada do inseto, e geralmente estão localizadas em áreas expostas do corpo como a face, os braços e as porções inferiores das pernas. Esta forma da doença pode, em alguns casos, curar-se sozinha ou evoluir para a leishmaniose mucocutânea, causando lesões destrutivas nas mucosas do nariz, boca e faringe. Na leishmaniose cutânea difusa, lesões nodulares de diferentes tamanhos podem aparecer em diferentes regiões distantes do sítio de exposição ao inseto (GONTIJO; DE CARVALHO, 2003; MCGWIRE; SATOSKAR, 2014; MURRAY et al., 2005). A forma recidiva da leishmaniose ocorre pelo aparecimento da doença nos locais de LC previamente cicatrizadas (MCGWIRE; SATOSKAR, 2014).

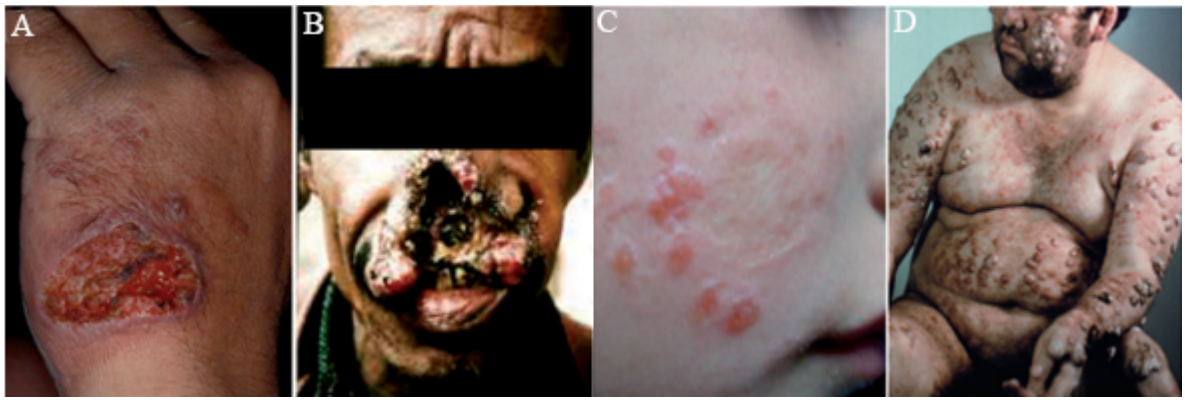


Figura 1: Representação clínica dos diferentes tipos de leishmaniose cutânea. **A**: Lesão leishmaniose cutânea na mão, **B**: Lesão mucocutânea na mão e no nariz, **C**: Lesões emergentes em uma cicatriz antiga de LC na leishmaniose recidiva, **D**: Lesões na leishmaniose cutânea difusa. Fonte: MCGWIRE; SATOSKAR, 2014 (adaptado).

A leishmaniose cutânea pode ser causada por várias espécies de leishmanias, no Velho Mundo espécies como *Leishmania major*, *Leishmania tropica*, *L. arabica* e *L. aethiopica*, são reponsáveis pela LC principalmente no norte e leste da África, parte central da Ásia e Índia. Em várias regiões da América do Sul e América Central pelo menos 12 espécies já foram descritas como causadoras de LC como *Leishmania mexicana*, *Leishmania amazonensis*, *Leishmania guyanensis*, *Leishmania venezuelensis, Leishmania lainsoni, Leishmania naiffi, Leishmania shawi, Leishmania colombiensis, Leishmania lindenberg Leishmania panamensis, Leishmania peruviana* and *Leishmania braziliensis* (AKHOUNDI et al., 2016; BRAZIL; RODRIGUES; FILHO, 2015).

No Brasil já foram identificadas oito espécies causadoras da doença, sendo sete do subgênero *Viannia (V)* e uma do subgênero *Leishmania (L).* As três principais espécies são: *L. (V.) braziliensis, L.(V.) guyanensis* e *L.(L.) amazonensis*. Mais recentemente, as espécies *L. (V.) lainsoni, L. (V.) naiffi, L. (V.) lindenberg, L. (V.) shawi* foram identificadas em estados das regiões Norte e Nordeste (AKHOUNDI et al., 2016; BRAZIL; RODRIGUES; FILHO, 2015; GONTIJO; DE CARVALHO, 2003).

L. braziliensis é a principal espécie prevalente no homem, causando lesões cutâneas e mucocutâneas, e é encontrada na América do Sul e Central, com a maioria dos casos ocorrendo no Brasil, Peru, Guatemala e Bolívia. No Brasil, este parasito é transmitido por diferentes espécies de flebotomíneos sendo os principais *Lutzomyia (N) whitmani*, *L. (N) intermedia*, com distribuição nas áreas periurbanas e urbanas do nordeste, sudoeste, sul e central. Infecções causadas por *Leishmania amazonensis* podem apresentar-se na forma simples da LC, como leishmaniose cutânea difusa ou ainda LMC, e no Brasil é transmitida por *Bi. flaviscutellata*, *Pi. nuneztova*. *Leishmania guyanensis* é encontrada principalmente no norte do Brasil, Bolívia, Guiana Francesa e Suriname, causando as formas LC e LMC, sendo transmitida por *L. (N) umbratilis* (AKHOUNDI et al., 2016; BRAZIL; RODRIGUES; FILHO, 2015; GONTIJO; DE CARVALHO, 2003; MCGWIRE; SATOSKAR, 2014).

Em 2015, quase 200.000 novos casos de LC foram reportados à OMS, sendo mais de 90 % dos casos pertencentes às regiões leste do Mediterrâneo e das Américas. Dentre todos os países registrados como endêmicos para LC, Afeganistão, Brasil, Irã, Iraque e Síria agrupam mais de 100.000 casos, representando 75 % dos casos globais. A leishmaniose cutânea e mucocutânea é endêmica em 18 países das Américas (PAHO, 2018). No Brasil, Colômbia e Peru juntos já foram reportados mais de 30.000 casos em 2015. No Brasil entre 2010 e 2016 foram registrados mais de 120.000 casos, sendo prevalentes nas regiões Norte e Centro Oeste do país (WHO 2017; MINISTÉRIO DA SÁUDE, 2018 (1)).

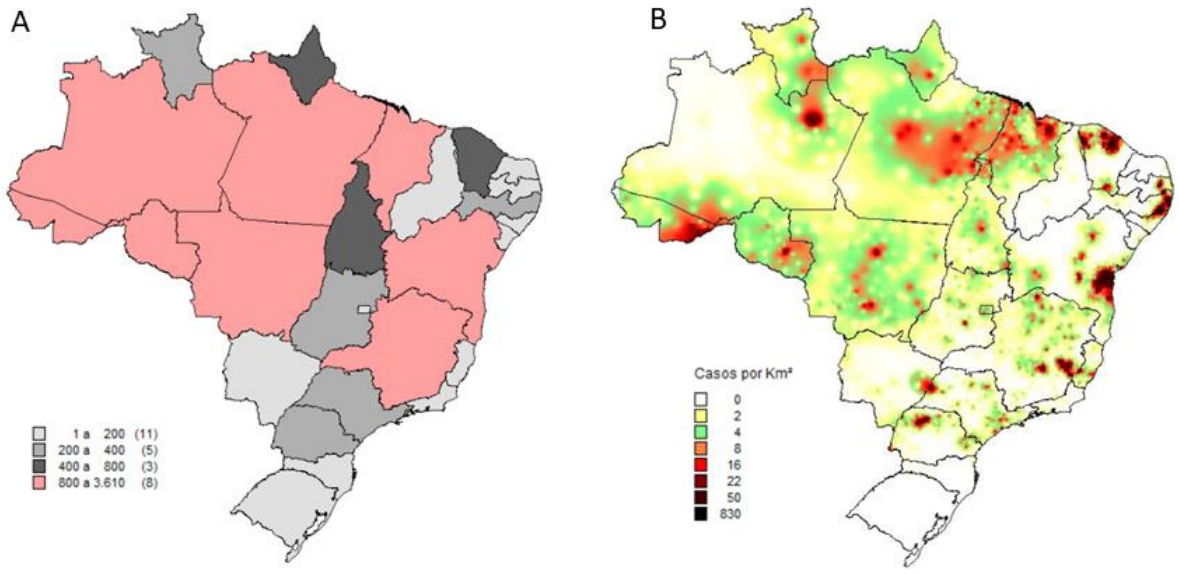


Figura 2: Distribuição de casos de leishmaniose cutânea no Brasil em 2015. A: Casos por estado da federação. **B:** Áreas com maior concentração de casos. Fonte: MINISTÉRIO DA SAÚDE, 2018 (1).

1.1.2. Leishmaniose Visceral

A leishmaniose visceral ou calazar é a manifestação mais grave de leishmaniose, podendo ser fatal se não tratada. Esta infecção é resultante da disseminação dos parasitos através de macrófagos infectados a partir do sítio de infecção, induzindo o acúmulo e infecção de células fagocíticas mononucleares dentro do sistema reticuloendotelial, o que leva a uma hiperplasia secundária dos orgãos invadidos e manifesta-se como esplenomegalia, hepatomegalia, palidez e imunosupressão, devido à invasão da medula óssea. Sintomas comuns são apresentados na forma de febre, anorexia, perda de peso, distenção abdominal e fraqueza, sendo que o tempo de incubação do parasito pode variar de 1 mês a mais de 1 ano (MCGWIRE; SATOSKAR, 2014; PACE, 2014; VAN GRIENSVEN; DIRO, 2012).

Em alguns casos de leishmaniose cutânea, pode ocorrer a visceralização da doença. Casos de visceralização foram reportados para *L. tropica* no Oriente Médio e *L. amazonensis* na América do Sul (VAN GRIENSVEN; DIRO, 2012), em pessoas portadoras de HIV.

Alguns pacientes tratados para leishmaniose visceral e que permanecem assintomáticos de meses a anos, podem desenvolver uma progressiva e rápida proliferação de parasitos na pele levando a diferentes tipos de lesão (nodulares, macular e máculo-papular), causando a chamada leishmaniose dermal pós calazar (PKDL). Esse tipo de apresentação da doença ocorre

principalmente na Índia e no Sudão em pacientes infectados por *L. donovani* (ZIJLSTRA et al., 2003).

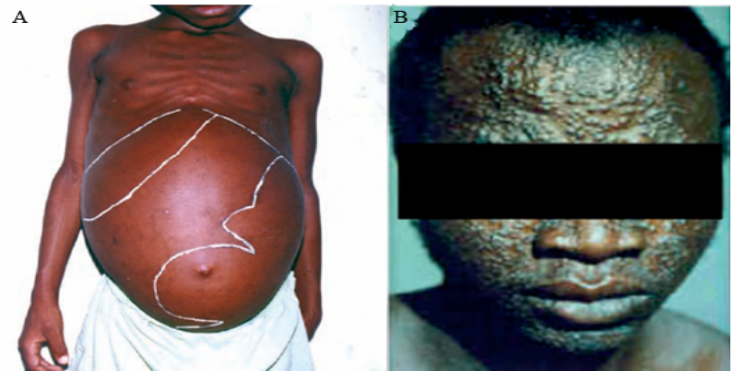


Figura 3:Representação clínica de leishmaniose visceral. A: Distensão abdominal causada pelo aumento dos orgãos em um caso de LV, **B:** Lesão facial em uma caso de PDKL. Fonte: MINISTÉRIO DA SAÚDE, 2006 e MCGWIRE; SATOSKAR, 2014 (adaptado).

A leishmaniose visceral é causada por protozoários pertencentes ao complexo *L. donovani* (VAN GRIENSVEN; DIRO, 2012). De maneira geral, no Velho Mundo a LV é causada pelo parasito *Leishmania donovani* na região do Oriente Médio, Índia, China, Sul da Ásia e África Central e por *Leishmania infantum* em países do mediterrâneo, Norte da África, Sudeste da Europa, Oriente Médio e Ásia Central. No Novo Mundo, *Leishmania infantum* foi identificada como agente causador de LV em países da América do sul e Central (AKHOUNDI et al., 2016; MURRAY et al., 2005; PACE, 2014; VAN GRIENSVEN; DIRO, 2012).

A leishmaniose visceral é trasmitida através da picada de fêmeas hematófagas de flebotomíneos do gênero *Phlebotomus* no Velho Mundo e *Lutzomyia* no Novo Mundo. Existem mais de 10 espécies de flebotomíneos responsáveis pela transmissão de LV (VAN GRIENSVEN; DIRO, 2012). O vetor primário de *L. infantum* na América do Sul e Central é a espécie *L. longipalpis* (BRAZIL; RODRIGUES; FILHO, 2015; DANTAS-TORRES; BRANDÃO-FILHO, 2006). Espécies de *L. cruzi* e *L. forattinii* infectadas com *L. infantum* foram encontradas em áreas endêmicas de LV do município de Corumbá, Mato Grosso do Sul (DE PITA-PEREIRA et al., 2008). Sendo que a espécie *L. cruzi* foi incriminada como vetor na região oeste do Brasil e em algumas regiões da Bolívia. Na venezuela e costa caribenha da Colômbia a espécie *Pintomyia (Pifanomyia) evansi* também é vetor de *L. infamtum* (BRAZIL; RODRIGUES; FILHO, 2015).

Em 2015, mais de 20.000 novos casos foram reportados à OMS, dos quais 73 % estão localizados no Brasil, Índia, Sudão do Sul e Sudão. Nas Américas 12 países são endêmicos para LV. Em 2016, a maioria dos casos foram registrados para crianças menores de 10 anos, sendo que na Colômbia, Honduras e Venezuela os mais afetados foram crianças menores de 5 anos

(PAHO, 2018). Somente o Brasil representa 14 % dos casos reportados em todo mundo em 2015 e 96 % dos casos nas Américas. A leishmaniose visceral no Brasil é uma doença endêmica e, inicialmente, sua ocorrência estava limitada a áreas rurais e pequenas localidades urbanas, mas atualmente os dados epidemológicos revelam a periurbanização e urbanização da doença com expansão para os grande centros. Entre os anos de 2010 e 2016 foram registardos no Brasil mais de 23.000 casos, com a maioria ocorrendo na região Nordeste, com destaque para o estado do Maranhão, e pouco mais de 700 óbitos neste mesmo período (WHO, 2017; MINISTÉRIO DA SAÚDE, 2018 (2)).

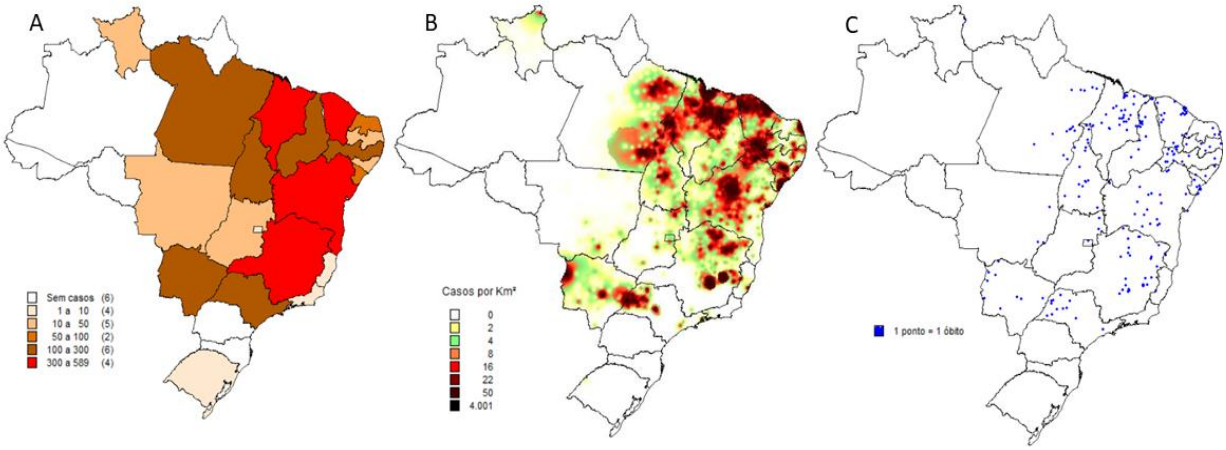


Figura 4: Distribuição de casos de leishmaniose visceral no Brasil em 2015. A: Casos por estado da federação. **B:** Áreas com maior concentração de casos. **C:** Óbitos por município de infecção. Fonte: MINISTÉRIO DA SAÚDE, 2018 (2).

1.2. Parasitos e ciclo de desenvolvimento

Nas descrições anteriores, vimos que o parasito *Leishmania* é o agente etiológico causador das leishmanioses e, dependendo da espécie, pode causar desde úlceras cutâneas até uma doença visceral com o aumento de órgãos, o que pode ser fatal. De maneira geral, as leishmanias são classificadas em quatro subgêneros diferentes *L. (Sauroleishmania), L. (Leishmania), L. (Viannia)* (LAISON; SHAW 1987; KILLICK-KENDRICK et al., 1986) e *L. (Mundinia)* (ESPINOSA et al., 2016). Quanto a diferenças biológicas, *Sauroleishmania* é encontrado em répteis e *Leishmania* e *Viannia* são encontrados em mamíferos. Baseados em diferenças anatômicas e local de desenvolvimento dentro do inseto vetor, o subgênero *Leishmania* possui seu desenvolvimento restrito ao intestino médio do vetor, sendo classificado como suprapilariano e encontrado no Novo e Velho mundo. O subgênero *Viannia* se desenvolve no intestino médio e posterior (peripilariano) e é endêmica apenas no Novo Mundo (AKHOUNDI et al., 2016; PACE, 2014; KAMHAWI, 2006). Exemplos destes dois tipos de

parasitos são *L. (Leishmania) infantum* e *L. (Viannia) braziliensis.* O subgênero *Mundinia* foi recentemente criado baseado na filogenia de sequências de GAPDH e HSP70 de *L. Enriettii* e *L. Martiniquensi*. Estas espécies inicialmente eram classificadas como *Leishmania*-like, mas na filogenia as sequências destes organismos (GAPDH e HSP70) compoem um clado basal que se agrupa com o gênero *Leishmania,* portanto um novo subgênero foi criado (ESPINOSA et al., 2016). Este novo subgênero possui tanto parasitas que são patógenos em humanos como *L. Martiniquensis, L. Orientalis*, e também parasitas não patogênicos para humanos como *L. Enriettii* e *L. macropodum* (JARIYAPAN et al., 2018).

Durante o ciclo de vida do protozoário *Leishmania*, duas formas morfológicas principais são descritas, promastigota e amastigota. A forma extracelular flagelada promastigota, com tamanho de 15-20 µm, é encontrada no vetor, e a forma intracelular obrigatória amastigota (não flagelada), com tamanho de 3-5 µm, se desenvolve intracelularmente no hospedeiro mamífero suscetível. A infecção em mamíferos ocorre com a inoculação de promastigotas na pele após a picada de fêmeas de flebotomíneos infectadas. Durante a alimentação sanguínea o inseto injeta no vertebrado compostos anti-inflamatórios, anti-histamínicos, peptídeos vasodilatadores, entre outros, e minimiza a percepção do vertebrado e a homeostase, facilitando a alimentação (SOARES; TURCO, 2003). Juntamente com estes compostos, a saliva também possui um efeito imunomodulatório que potencializa a infecção por *Leishmania* e estimula a produção de diversas citocinas (SACKS; KAMHAWI, 2001).

Os promastigotas injetados na pele, associados ao dano vascular levam a liberação de fatores quimiotáticos, provenientes do parasita (PSG e LCF), do hospedeiro (IL-8, C3, IL-17, GCP-2) ou do vetor (saliva). Estes fatores, recrutam os neutrófilos de maneira rápida e transitória para o local da infecção. Os parasitas são então fagocitados por neutrófilos e induzem a formação das NETs (armadilhas extracelulares de neutrófilos). Após infecção dos neutrófilos, os promastigotas infectam os macrófagos de forma indireta, induzindo a apoptose de neutrófilos o que leva à fagocitose, ou de forma direta, quando são liberados dos neutrófilas e são diretamente fagocitados pelos macrófagos (HURRELL; REGLI; TACCHINI-COTTIER, 2016). Dentro dos macrófagos, os promastigotas se desenvolvem em formas amastigotas e se multiplicam dentro dos fagolisossomos. Após a lise da célula estes parasitos são liberados e infectam outros macrófagos. No caso da LV, os parasitos podem infectar células do sistema reticuloendotelial em vários tecidos, infiltrando predominantemente o baço, a medula óssea, o fígado, intestino e os linfonodos (PACE, 2014; VAN GRIENSVEN; DIRO, 2012).

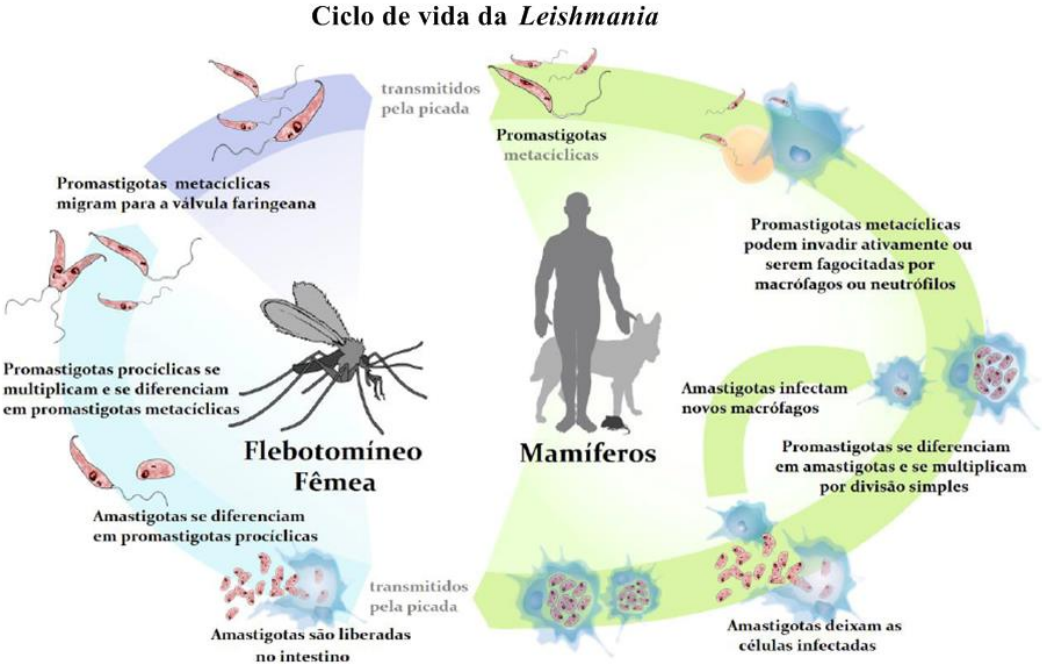


Figura 5: Ciclo de vida dimórfico do protozoário *Leishmania*. Desenvolvimento da forma promastigota dentro do hospedeiro invertebrado e amastigota dentro dos macrófagos de mamíferos. Fonte: CANAL CIÊNCIA, 2018.

O ciclo de vida do parasito *Leishmania* no vetor inicia-se com a ingestão de sangue contendo macrófagos infectados com amastigotas de *Leishmania*, que passam diretamente para o intestino médio abdominal. Durante a migração dos parasitos da porção posterior do intestino médio para a porção anterior, ocorre a diferenciação em formas distintas ao longo do trato digestivo de flebotomíneos: amastigota, promastigota procíclica, promastigotas nectomonadas, promastigotas leptomonadas, promastigotas haptomonadas e formas metacíclicas infecciosas (KAMHAWI, 2006). Em cada estágio de desenvolvimento os parasitos passam por alterações morfológicas e funcionais de forma que consigam ultrapassar as barreiras de desenvolvimento dentro do flebotomíneo.

Categoria Morfológica	Critério	Ilustração
Amastigota	Forma do corpo ovóide, nenhum flagelo saliente do bolso flagelar	
Promastigota Procíclica	Comprimento do corpo 6,5 - 11,5 µm, flagelo <comprimento do corpo (largura do corpo variável)	
Promastigota Nectomonada	Comprimento do corpo> 12 µm, (largura do corpo e comprimento do flagelo é variável)	
Promastigota Leptomonada	Comprimento do corpo 6,5 - 11,5 µm, flagelo ³ comprimento do corpo (largura do corpo variável)	
Promastigota Haptomonada	Expansão do flagelo em forma de disco (Forma do corpo e tamanho do flagelo variável)	
Promastigota Metacíclica	Comprimento do corpo <8 µm, largura do corpo <1µm, flagelo> comprimento do corpo	
Paramastigota	Cinetoplasto adjacente ao núcleo, flagelo externo presente	

Figura 6: Diferentes morfologias presentes durante o ciclo de desevolvimento de *L. mexicana* dentro do inseto vetor. Fonte: ROGERS; CHANCE; BATES, 2002 (Traduzido para o português).

A forma promastigota procíclica surge algumas horas após a infecção e possui uma forma mais arredondada e flagelo pequeno, se multiplica rapidamente estabelecendo a infecção inicial devido ao ambiente nutricionalmente rico e é resistente à ação de enzimas presentes no bolo alimentar. Seu desenvolvimento está confinado dentro do bolo alimentar circundado pela matriz peritrófica. Estas então se diferenciam em promastigotas nectomonadas, formas não replicativas, grandes e delgadas com tamanho de flagelo variável, que escapam da matriz peritrófica e são capazes de se ancorar na parede da porção anterior do intestino médio abdominal, iniciando o processo de migração em direção ao intestino médio torácico e estabelecendo a infecção fora do bolo alimentar (GOSSAGE; ROGERS; BATES, 2003; ROGERS; CHANCE; BATES, 2002; SACKS; KAMHAWI, 2001). O escape da matriz é facilitado pela secreção de quitinases pelo parasito e também pela ação de quitinases endógenas do flebotomíneo, que atuam sobre a matriz peritrófica quitinoproteica (RAMALHO-ORTIGÃO et al., 2005; SCHLEIN; JACOBSON; SHLOMAI, 1991). As formas promastigotas nectomonadas transformam-se em promastigotas leptomonadas, formas menores de parasitos que iniciam um novo ciclo de multiplicação. As leptomonadas juntamente com as haptomonadas são responsáveis pela formação de um *plug* de parasitos bloqueando a válvula de estomodeu e secretam uma matriz composta por proteofosfoglicanos filamentosos (PSG – *Promastigote Secretory Gel*), que restringe a mobilidade dos protozoários (GOSSAGE; ROGERS; BATES, 2003). As haptomonadas se aderem à superfície cutilar do intestino e dependendo da espécie de *Leishmania* pode ser encontrada ancorada a diferentes porções do

intestino, como intestino posterior e válvula de estomodeu, através de uma estrutura similar a hemidesmossomos presente no final de seus flagelos. A origem das formas haptomonadas não está bem definida (nectomonadas ou leptomonadas) (ROGERS; CHANCE; BATES, 2002).

As formas promastigotas leptomonadas se diferenciam em promastigotas metacíclicos que se acumulam na região da cárdia e são altamente infectivas, estas possuem um pequeno corpo celular e flagelo mais longo que o corpo, o que permite uma rápida mobilidade. A presença de parasitas e o bloqueio da válvula de estomodeu pela secreção de PSG, leva à degeneração dessa região. A válvula de estomodeu danificada permanece aberta e durante a alimentação sanguínea ocorre o refluxo de parasitas e PSG, e as fêmeas inoculam o parasito no hospedeiro vertebrado (GOSSAGE; ROGERS; BATES, 2003; KAMHAWI, 2006; ROGERS; CHANCE; BATES, 2002; SACKS; KAMHAWI, 2001). Foi demonstrado que um tipo de proteofosfoglicano filamentoso (fPPG), componente do gel de secreção de promastigotas (PSG), é regurgitado juntamente com os parasitas e é um dos fatores responsáveis pela exarcebação da infecção (ROGERS et al., 2004). Durante a alimentação sanguínea, devido ao bloqueio da probóscide pela presença de parasitos o flebotomíneo ingere uma quantidade menor de sangue e é levado a realizar o *probing* diversas vezes, o que também causa o aumento da transmissão (PACE, 2014; ROGERS; BATES, 2007).

Recentemente foi demonstrado que a ingestão de uma segunda alimentação sanguínea não infectante por flebotomíneos infectados, leva à dediferenciação dos promastigotas metacíclicos em uma forma similar a forma leptomonada, chamada promastigota retroleptomonada, esta forma é capaz de se diferenciar e multiplicar rapidamente em formas infectivas, o que leva ao aumento da infecção no caso de uma nova alimentação sanguínea (SERAFIM et al., 2018). Após uma segunda alimentação sanguínea as formas haptomonadas formam uma estrutura esférica chamada HPS (*Haptomonad Parasite Sphere*), esta juntamente com o gel de secreção de promastigotas (PSG), atuam no bloqueio de parasitas na válvula de estomodeu, promovendo a regurgitação e aumento da transmissão (SERAFIM et al., 2018).

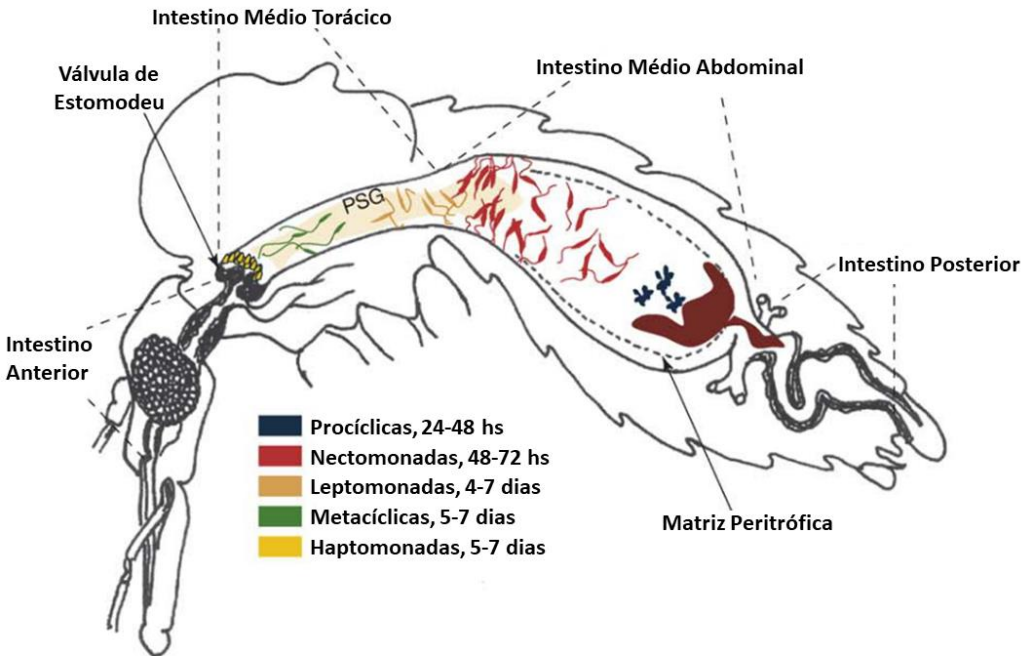


Figura 7: Ciclo de vida do parasito *Leishmania* (suprapilárico) dentro do vetor, demonstrando o aparecimento de diferentes formas morfológicas de promastigotas durante os estágios de desenvolvimento dentro do intestino médio do flebotomíneo. Fonte: KAMHAWI, 2006 (adaptado).

Durante o ciclo de vida da *Leishmania* em flebótomos, vários mecanismos fisiológicos, bioquímicos e moleculares são importantes para a manutenção da infecção. Dentre eles, a modulação de enzimas digestivas por parasitos, ajudando na resistência à digestão por enzimas proteolíticas durante a digestão sanguínea. Ainda podemos destacar mecanismos, como o escape da matriz peritrófica e a ancoragem às galectinas presentes no epitélio intestinal do vetor através de lipofosfoglicanos (LPGs), para que o parasito não seja excretado com as fezes (KAMHAWI et al., 2004; OLIVEIRA et al., 2009; PIMENTA et al., 1992; SECUNDINO et al., 2010; SOARES et al., 2010). Alguns trabalhos também demonstraram que em algumas espécies a ligação do parasito à superfície do epitélio é dependente de LPGs (ROGERS; CHANCE; BATES, 2002). A composição dos LPGs na superfície do parasito é variável, depende da espécie e é considerado um fator crucial na competência vetorial para as espécies de *Leishmania*. Os polimorfismos dos LPGs estão associados com a capacidade das *Leishmanias* em completar seu desenvolvimento em diferentes espécies de vetores. Dessa forma, existem vetores que são restritivos, permitindo o desenvolvimento de apenas uma ou duas espécies de *Leishmania*, e vetores permissivos como *L. longipalpis*, que permite o desenvolvimento de *Leishmania tropica*, *Leishmania mexicana*, *Leishmania infantum*, *L. major* e *L. amazonensis* (KAMHAWI, 2006; SOARES; TURCO, 2003). No entanto, na natureza *L. longipalpis* foi incriminada apenas como vetor de *Leishmania infantum.*

1.3. Flebotomíneos

Inicialmente a taxonomia dos flebotomíneos (Diptera: Psychodidae) era exclusivamente baseada em aspectos morfológicos, mas o desenvolvimento de novas técnicas tem levado a uma melhor classificação e identificação das espécies. Atualmente, dentre as mais de 800 espécies identificadas, 464 são encontradas no Novo Mundo e 375 no Velho Mundo (AKHOUNDI et al., 2016; GALATI, 2003; SECCOMBE, A. K.; READY, P. D.; HUDDLESTON, 1993). Pouco mais de 10 % das espécies conhecidas são vetores de doença e possuem importância médica (WHO, 2017). Dentre as doenças transmitidas por flebotomíneos podemos destacar a bartonelose na América do Sul, causada pela bactéria *Bartonella bacilliformis* e transmitida por *Lutzomyia verrucarum* (SANCHEZ CLEMENTE et al., 2012), algumas arboviroses no norte da África e partes da Ásia com a transmissão do vírus pappataci causador da febre pappataci (Ex. *Phlebotomus papatasi*, *P. perfiliewi*, *P. perniciosus*) (TESH, 1988) e as leishmanioses. As leishmanioses são transmitidas através da picada de fêmeas de flebotomíneos do gênero *Phlebotomus* no Velho Mundo e *Lutzomyia* no Novo Mundo (PACE, 2014; SHARMA; SINGH, 2008; KILLICK-KENDRICK, 1999).

Os flebotomíneos são insetos pequenos, medindo de 2 a 3 mm, apresentam um corpo com muita pilosidade e, quando em repouso, mantêm suas asas em um ângulo de 45 graus acima do corpo. Possuem o corpo e asas pequenas, com pernas longas e delgadas. Estes insetos possuem uma capacidade de vôo baixa, de apenas alguns quilômetros, com os vôos consistindo em saltos curtos, e são mais ativos durante o amanhecer e o anoitecer. Estes insetos possuem o desenvolvimento estritamente terrestre, são holometábolos e passam pelos estágios de ovo, larva, subdivididos em 4 estádios (L1-L4), pupa e adultos (KILLICK-KENDRICK, 1999; SOARES; TURCO, 2003). Seu desenvolvimento é relativamente longo e depende de vários fatores como temperatura, umidade e nutrição. O tempo de uma geração pode variar de 1 a 3 meses. Para *L. longipalpis*, em condições de laboratório, foi demonstrado que o ciclo de vida completo entre os estágios de ovo até adultos, pode variar entre 28 e 36 dias (RANGEL et al., 1986). A maioria dos flebotomíneos prefere locais úmidos e quentes para desenvolvimento. No entanto, algumas espécies como *Lutzomyia verucarrum* e *Lutzomyia peruensis* encontradas na região dos Andes no Peru, e *Phlebotomus orientalis*, encontrada no Sudão, se desenvolvem em condições de baixa umidade e temperatura, respectivamente (LAWYER et al., 2017). Os adultos apresentam dimorfismo sexual, com diferenças em suas estruturas corporais e também comportamentais. Os machos, por exemplo, possuem na extremidade do abdomem estruturas de fixação que são utilizadas para o acasalamento. Os adultos machos e fêmeas se alimentam de uma dieta rica em açúcares obtidos da seiva de plantas, néctar de flores e também de

"honeydews" que são excretas de afídeos e coccídeos (CAMERON et al., 1995; KILLICK-KENDRICK; KILLICK-KENDRICK, 1987; MACVICKER et al., 1990; MOORE et al., 1987). As fêmeas adultas também se alimentam de sangue para obter a quantidade de proteína necessária para maturação dos ovos (BRAZIL, R.P.; BRAZIL, 2003).

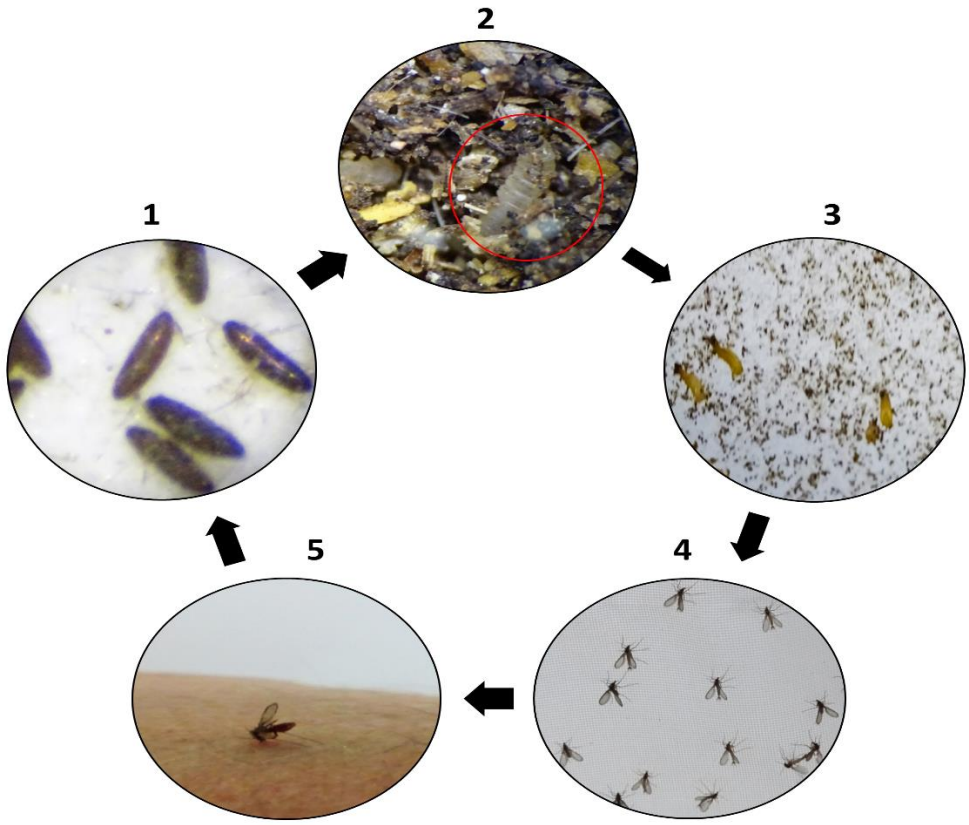


Figura 8: Ciclo de vida *Lutzomyia longipalpis*. 1- Ovos. **2**- Larva (destacada círculo vermelho)**. 3**- Pupas**. 4**- Machos e fêmeas adultos alimentados com sacarose. 5- Fêmea realizando alimentação sanguínea. Fonte: Autor.

Os flebotomíneos podem ser encontrados em torno de habitações humanas e se reproduzem em resíduos orgânicos e cantos escuros, nas fendas das paredes, com alta umidade e temperatura, por exemplo. Os machos encontram o hospedeiro vertebrado, liberam feromônios e atraem as fêmeas para alimentação sanguínea e cópula (SOARES; TURCO, 2003). As fêmeas iniciam então o processo de maturação dos ovos, que se completa em torno de 5 dias. A quantidade de ovos depositada depende da quantidade de sangue ingerido. Alguns trabalhos relatam que o tipo de sangue (espécie vertebrada) também pode afetar a quantidade de ovos (READY, 1979), e no geral entre 15 e 80 ovos são depositados. Os ovos de flebotomíneos são pequenos e possuem uma forma elíptica com 0,3-0,5 mm de comprimento e 0,1 a 0,15 mm de largura (LAWYER et al., 2017). Os ovos são depositados isoladamente ou agrupados, em locais úmidos e protegidos, como fendas de rochas, bases de árvores, buracos de árvores, sob folhas

no chão da floresta, tocas de animais, cupinzeiros, entre outros, e se aproveitam do microclima local para o seu desenvolvimento. Geralmente, os ovos eclodem 6 a 11 dias após a alimentação sanguínea, mas para algumas espécies de flebotomíneos, dependendo das condições este período pode se estender até 30 dias (VOLF; VOLFOVA, 2011). As larvas se desenvolvem em material orgânico, calor e umidade adequados, de forma lenta, levando até 3 semanas para chegar à fase de pupa. As larvas variam de cor, branco a cinza, e tamanho dependendo da espécie. De maneira geral, as larvas possuem 4 estádios de desenvolvimento e variam de <1 a < 4 mm, durante o desenvolvimento. No primeiro instar possuem cabeça escura, minúsculas cerdas laterais e duas cerdas caudais. No segundo e terceiro instar possuem quatro cerdas caudais e são diferenciadas pelo tamanho. No quarto instar elas são maiores, também possuem 4 cerdas caudais; mas as cerdas laterais são mais pronunciadas e possuem uma placa anal dorsal fortemente esclerotizada (LAWYER et al., 2017; SHARMA; SINGH, 2008). Quando as larvas atingem o fim do quarto instar, param de se alimentar e esvaziam o conteúdo instestinal, se tornando esbranquiçadas. Em aproximadamente 24 horas as larvas tornam-se pupas. As pupas no início do desenvolvimento parecem esbranquiçadas e, em seguida, adquirem um tom laranja ou avermelhado a preto, à medida que a eclosão se aproxima (LAWYER et al., 2017).

Flebotomíneos do gênero *Phlebotomus* são encontrados nas áreas mais temperadas do Velho Mundo e localizam-se em áreas de savana, deserto e semiárido. Eles podem se reproduzir em microcriadouros no peridomicílio e assim entram em contato com os humanos. Foram identificados em regiões que se estendem desde o Mediterrâneo, Oriente Médio até as regiões orientais da Ásia Central. Nas regiões mais tropicais apenas algumas espécies desse gênero foram descritas. Os flebotomíneos pertencentes ao gênero *Lutzomyia* apresentam uma maior diversidade que os flebotomíneos do Velho Mundo e são abundantes nas regiões tropicais da América, apresentando uma grande área de dispersão. São encontrados principalmente nas regiões de floresta nas Américas do Sul e Central (AKHOUNDI et al., 2016).

1.4. Sistema digestório em insetos

Os insetos apresentam diferentes estruturas bucais e essa diversidade pemite que possam se alimentar de uma grande variedade de materiais. Além disso, a sua habilidade de adquirir energia a partir de fontes alimentares variadas permitiu que ocupassem diversos nichos ecológicos. Assim, o trato digestivo dos insetos possui uma grande diversidade estrutural, o que permite a especialização na digestão de um tipo específico de dieta (KLOWDEN, 2013). A digestão é um processo no qual as moléculas de alimentos são quebradas em moléculas menores permitindo sua absorção pelas células do intestino. As enzimas digestivas participam deste

processo e as etapas da digestão estão relacionadas a localização dessas enzimas dentro do trato digestivo do inseto (TERRA; FERREIRA, 2012).

1.4.1. Anatomia do tubo digestivo de insetos

O tubo digestivo de insetos é dividido em três regiões principais: intestino anterior (IA), intestino médio (IM) e intestino posterior (IP) (DOUGLAS, 2013). A anatomia do tubo digestivo varia de acordo com a espécie ou até mesmo com os diferentes estágios de desenvolvimento de um determinado organismo, dentro de uma mesma espécie. Assim podem ocorrer variações anatômicas, fisiológicas e bioquímicas. Na figura 9 está representada a estrutra geral do trato digestivo em insetos. Em *L. longipalpis* o trato digestivo de adultos segue o mesmo modelo, onde o intestino médio é anatomicamente dividido em intestino médio torácico (IMT), região mais estreita e o intestino médio abdominal (IMA), região expandida (Figura 10).

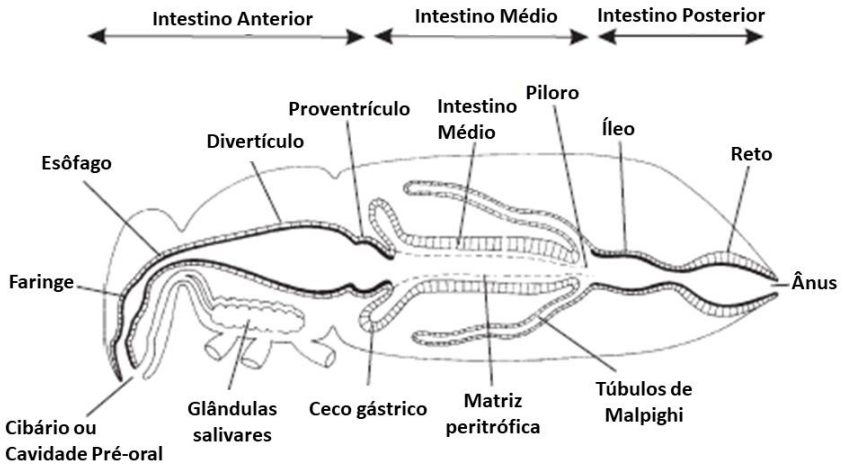


Figura 9: Figura representativa do tubo digestivo em insetos, subdividido em suas principais regiões: intestino anterior (IA), intestino médio (IM) e intestino posterior (IP). Fonte: KLOWDEN, 2013 (adaptado).

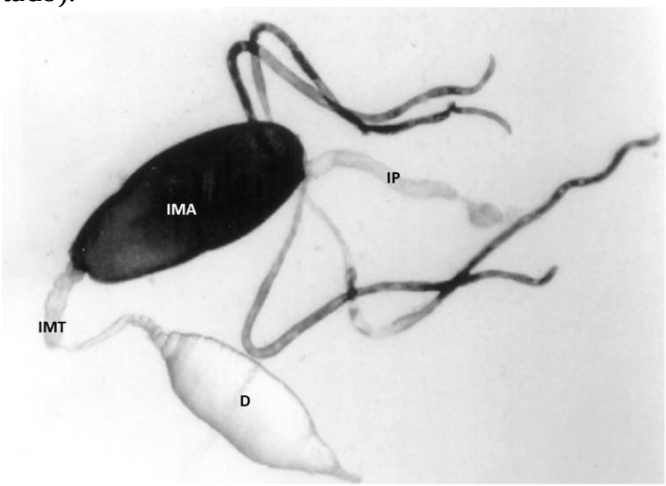


Figura 10: Anatomia do tubo disgestivo de fêmeas de *L. longipalpis* após alimentação sanguínea. D: divertículo preenchido com sacarose, **IMT**: intestino médio torácico, **IMA**:

Intestino médio abdominal preenchido com sangue e **IP**: intestino posterior. Fonte: SANTOS et al., 2008 (adaptado).

Devido à sua origem ectodermal, o intestino anterior e posterior secretam uma cutícula que não permite a secreção de enzimas. O intestino médio não é coberto por esta cutícula. No intestino médio é secretada uma matriz que envolve o alimento, denominada membrana peritrófica. O intestino anterior não está envolvido em secreção ou absorção e pode ser dividido em faringe, esôfago, divertículo e proventrículo. No intestino anterior de insetos mandibulares, existe a cavidade pré-oral, que está dividida em cibário e salivário pela hipofaringe (DOUGLAS, 2013; KLOWDEN, 2013). O cibário em algumas espécies possui espinhos cuticulares que são capazes de romper os glóbulos vermelhos antes de estes passarem para o intestino médio (CHADEE; BEIER; MARTINEZ, 1996). Em *Phlebotomus papatasi* por exemplo, o cibário está localizado na região posterior do canal alimentar e é separado da faringe pela válvula cibarial (WARBURG, 2008).

De maneira geral, o par de glândulas salivares esvaziam seu conteúdo no salivário através dos dutos salivares. A saliva é composta principalmente por água e outros componentes proteícos que variam de acordo com a espécie de inseto. Em alguns casos amilases e invertases presentes na saliva pré-digerem o alimento antes que este entre em contato com as enzimas do intestino médio. Em insetos hematófagos a saliva contém substâncias anti-hemostáticas, vasoativas e imunomoduladoras que modificam a resposta imune do hospedeiro. No caso de insetos vetores de doenças, como os flebotomíneos, o efeito imunosupressor da saliva auxilia no estabelecimento dos patógenos no hospedeiro vertebrado (VOLF; TESAR, 2000). Foi demonstrado que a saliva de flebotomíneos é responsável por exarcebar a infecção por *Leishmania major* em camundongos e que esta é necessária para o estabelecimento das infecções (THEODOS; RIBEIRO; TITUS, 1991; TITUS; RIBEIRO, 1988). Além deste aspecto, diferentes composições salivares de componentes imunossupressores e anti-hemostáticos em diferentes populações de *L. longipalpis*, podem determinar se a infecção por *Leishmania infantum* em humanos se manifesta como forma cutânea ou na forma visceral da doença (WARBURG et al., 1994).

O esôfago é a região após a faringe que leva ao intestino, no entanto em alguns insetos, como é o caso de *L. longipalpis*, esta região pode ser modificada, possuindo um divertículo extensível reponsável pelo armazenamento de soluções açucaradas. O divertículo é ligado por um duto ao esôfago e os açúcares armazenados são liberados lentamente para o intestino médio. Devido ao revestimento cuticular este compartimento não absorve açúcares. O proventrículo é uma válvula presente no início do intestino médio que permite que os alimentos sólidos não

retornem ao intestino anterior e ao mesmo tempo permite o movimento da água em ambas as direções. Em algumas espécies essa região é mais especializada e pode assumir outras funções (DOUGLAS, 2013). Em flebotomíneos, o intestino médio torácico se liga ao intestino anterior através da válvula de estomodeu, ela é revestida por cutícula e em insetos infectados por *Leishmania* esta região possui uma grande concentração de parasitos (WARBURG, 2008).

O principal sítio de digestão nos insetos é o intestino médio, de origem endodérmica é revestido com um epitélio microvilar, que produz e secreta as enzimas digestivas e absorve os nutrientes (GEMETCHU, 1974; RUDIN; HECKER, 1982). O intestino médio se constitui num tubo simples de onde podem se expandir os cecos gástricos, usualmente na sua extremidade proximal. Os cecos gástricos aumenta a área de secreção e absorção do intestino médio e também cria um contrafluxo como resultado de sua absorção diferencial de água. Esse contrafluxo pode ser criado também pela absorção de água pela região proximal do intestino médio, na ausência de cecos gástricos. A água é secretada no lúmen pelo intestino posterior ou pela região posterior do intestino médio, e é reabsorvida na região cecal ou anterior do intestino médio. O contrafluxo da água para a região anterior faz com que os produtos de digestão passem pelo intestino enquanto o que não foi digerido se move para trás para ser eliminado (DOUGLAS, 2013; KLOWDEN, 2013).

Dentre os principais tipos celulares encontrados no intestino médio estão as células colunares. Elas possuem inúmeras microvilosidades voltadas para o lúmen, aumentando a superfície de absorção e secreção do IM. A maioria dos nutrientes portanto, são absorvidos através destes tipos celulares. Essas células possuem uma complexa rede de retículo endoplasmático, necessária para produção das enzimas digestivas. As microvilosidades são recobertas por uma camada filamentosa de glicoproteínas, o glicocálix (KLOWDEN, 2013).

Na maior parte dos insetos, o intestino médio contém no seu lúmen uma matriz quitino protéica, a membrana peritrófica (MP). Isto ocorre porque o intestino médio não possui um revestimento cuticular e pode sofrer abrasão pelo alimento ingerido, mesmo em insetos que ingerem alimentos não particulados (fluidos) (SECUNDINO et al., 2005). A matriz também exerce um papel importante na compartimentalização da digestão, funcionando como uma barreira permeável para enzimas digestivas (HEGEDUS et al., 2009; KLOWDEN, 2013; TERRA; FERREIRA, 2012). Além disto esta matriz auxilia na proteção contra toxinas ingeridas (ABEDI; BROWN, 1961; BARBEHENN, 1999), contra invasão de microorganismos com a produção de muco (TERRA; FERREIRA, 2012) e também sequestra o heme produzido durante a digestão do sangue (PASCOA et al., 2002).

A matriz peritrófica é formada por uma rede de microfibrilas de quitina imersa em uma matriz de carboidratos e proteínas (mucinas e peritrofinas). Ela pode ser formada continuamente a partir da secreção de células epiteliais do intestino médio ou em reposta à alimentação. A MP pode ser produzida ao longo de todo o intestino médio, em camadas (MP tipo I) ou a partir de células da cárdia (MP tipo II) (TERRA; FERREIRA, 2012). Insetos pertencentes às ordens Coleoptera, Dictyoptera, Ephemeroptera, Hymenoptera, Odonata, Orthoptera, Phasmida, Lepidoptera secretam MP do tipo I, assim como adultos hematófagos da ordem Diptera (formação em resposta à alimentação sanguínea) (HEGEDUS et al. 2009; WATERHOUSE, 1957).

A MP tipo II é sintetizada em insetos de outras ordens, como Dermaptera, Isoptera, Embiodea, algumas espécies da ordem Lepidoptera e larvas pertencentes à ordem Diptera (PETERS, 1992; WIGGLESWORTH, 1930). A matriz peritrófica pode ser produzida diferentemente em diferentes estágios do mesmo inseto, por exemplo, em *L. longipalpis* a MP tipo II pode ser encontrada em larvas e a MP tipo I pode ser encontrada em fêmeas adultas após alimentação sanguínea. A cinética de formação e degradação da matriz peritrófica está relacionada à progressão da digestão sanguínea, que varia de acordo com a espécie de inseto. Em *L. longipalpis* a síntese dos componentes da MP é iniciada 1h após a alimentação sanguínea e entre 6 h e 24 h ela está totalmente formada (SECUNDINO et al., 2005). Em *P. Papatasi*, a formação da MP inicia-se 4 h após a alimentação sanguínea (BLACKBURN et al., 1988).

O intestino posterior juntamente com os túbulos de Malpighi estão envolvidos na osmoregulação. O intestino posterior é capaz de reabsorver aminoácidos, água e íons da urina primária iso-osmótica produzida pelos túbulos de Malpighi, assim como recuperar diversas subtâncias importantes de alimentos não digeridos ou bolo fecal antes de este ser descartado (KLOWDEN, 2013).

1.4.2. Etapas da digestão: inicial, intermediária, final

Como dito anteriormente, a membrana peritrófica é uma importante matriz quitino proteica que possui com principal função auxiliar na compartimentalização da digestão. Ela envolve o alimento ingerido e separa o conteúdo luminal em dois compartimentos, o espaço endoperitrófico (dentro da membrana) e o espaço ectoperitrófico (fora da membrana). A MP contém poros que são permeáveis a algumas enzimas digestivas e produtos da digestão, mas não às grandes moléculas presentes no alimento não digerido, como polissacarídeos e proteínas de alto peso molecular. Essa permeabilidade é reponsável pela compartimentalização do processo digestivo (TERRA et al., 1994).

A maioria das moléculas ingeridas durante a alimentação são polímeros que passam por diversas etapas de digestão. A digestão inicial ocorre dentro do espaço endoperitrófico, onde estes polímeros são quebrados em oligômeros por enzimas que conseguem atravessar a MP. A digestão intermediária ocorre no espaço ectoperitrófico, onde estes oligômeros são então clivados em dímeros. As enzimas que participam desta etapa da digestão possuem massas moleculares maiores e não conseguem atravessar o poro da MP. Na digestão final os dímeros são quebrados por enzimas ancoradas à membrana das microvilosidades das células do intestino, gerando monômeros que são absorvidos pelas células (TERRA et al., 1994; TERRA; FERREIRA, 2005, 2012).

A compartimentalização da digestão aumenta a eficiência desse processo, pois previne ligações não específicas do material não digerido à superfície das células do intestino médio, previne a excreção das enzimas digestivas no contrafluxo de água e concentra os monômeros produzidos junto à superfície celular, colocando-os próximos aos carreadores responsáveis pela absorção (TERRA; FERREIRA, 2005, 2012). Além disso, a compartimentalização aumenta a eficiência na digestão de alimentos poliméricos e oligoméricos. Os oligômeros podem ser substratos ou inibidores para algumas hidrolases que atuam em polímeros assim como os substratos poliméricos podem fazer ligações inespecíficas com as "oligômero hidrolases". No entanto, a compartimentalização permite que, após a hidrólise dos substratos poliméricos, os oligômeros sejam transferidos para o espaço ectoperitrófico onde as oligômero hidrolases estão restritas. Assim, a degradação rápida dos polímeros garante que estes não sejam eliminados e que a digestão dos oligômeros ocorra na ausência de uma inibição parcial (ligações não específicas), aumentando a digestibilidade (TERRA et al., 1996).

Além da compartimentalização, outro fator importante para o processo digestivo é o pH do conteúdo intestinal. Aparentemente este fator não se correlaciona com o tipo de dieta, mas sim com a filogenia do inseto (CLARK, 1999). Dentro do intestino do mesmo inseto existem diferentes faixas de tamponamento nas diferentes regiões ao longo do intestino médio. Em Dictyoptera, Orthoptera e na maioria das famílias de Coleoptera, por exemplo, o conteúdo intestinal é mais acídico no intestino médio anterior e mais neutro ou básico no intestino médio posterior (TERRA et al., 1994). No geral, estudos têm demonstrado que enzimas digestivas possuem um pH ótimo de acordo com a região em que atuam dentro do intestino (DILLON; EL KORDY, 1997; GONTIJO et al., 1998; SANTOS et al., 2008).

Microorganismos presentes no intestino também ajudam a aumentar a eficiência do processo digestivo, podendo fornecer enzimas e vitaminas que auxiliam os insetos quando estes apresentam dietas sub-ótimas. Por exemplo, os microorganismos podem fornecer vitaminas do

complexo B para insetos hematófagos e aminoácidos essenciais para insetos que se alimentam da seiva elaborada de plantas (DILLON; DILLON, 2004).

1.4.3. Habitos alimentares x digestão

Para entender o processo digestivo em insetos, assim como as variações morfológicas e fisiológicas do trato digestivo é necessário levar em consideração fatores como a dieta e a posição filogenética dos insetos. Dois fatores principais determinaram a evolução morfológica do trato digestivo, a alimentação sólida ou líquida e a ingestão de alimentos de origem animal ou vegetal. Insetos que se alimentam de plantas por exemplo, precisam ingerir um quantidade maior de alimento para que possam cumprir suas demandas nutricionais. Além disso, os insetos que se alimentam principalmente da seiva de plantas geralmente possuem um intestino mais longo que o corpo para que possam processar uma grande quantidade de fluidos diluídos, como em homopteros e hemípteros (KLOWDEN, 2013).

Em insetos holometábolos por exemplo, existem diferenças no processo digestivo, onde as fases imaturas e adulta exploram diferentes fontes alimentares. Por exemplo, *L. longipalpis* durante o desenvolvimento larval se alimenta principalmente de material em decomposição como folhas de plantas e fezes de animais (FELICIANGELI, 2004). β-1,3-glucanases, quitinases, amilases e algumas glicosidases foram descritas como as principais enzimas envolvidas na digestão larval (MORAES et al., 2012; VALE et al., 2012). Na fase adulta, os flebotomíneos possuem uma dieta rica em açúcares, glicolipídeos e glicoproteínas. Açúcares como sacarose, maltose, trehalose e melizitose são obtidos a partir da seiva de plantas, néctar de flores e também das excretas açucaradas (honeydew) de afídeos e coccídeos (KILLICK-KENDRICK; KILLICK-KENDRICK, 1987; MACVICKER et al., 1990; SCHLEIN; MULLER, 1995). Esses açúcares são clivados em monômeros pela ação das α-glicosidases, enquanto o amido, encontrado principalmente nas folhas de plantas é hidrolisado pelas α-amilases (JACOBSON; SCHLEIN, 2001). As fêmeas adultas também se alimentam de sangue para maturação dos ovos, sendo esta a principal fonte de proteínas e lipídeos para o seu desenvolvimento. Nesses insetos, as proteases atuam na digestão de proteínas do sangue, mas glicolipídeos e glicoproteínas podem estar sujeitos à clivagem pela ação de exoglicosidases. Sendo assim, as carboidrases e proteases são as principais enzimas envolvidas no processo digestivo em adultos de flebotomíneos.

Alguns trabalhos já descreveram a atuação de carboidrases e proteases em adultos de flebotomíneos (DILLON; EL KORDY, 1997; GONTIJO et al., 1998; JACOBSON; STUDENTSKY; SCHLEIN, 2007; MORAES et al., 2018; SANTOS et al., 2014; TELLERIA

et al., 2007, 2010), no entanto, a maioria destes estão focados no estudo de proteases. Além das proteínas serem importantes nutrientes para estes insetos, é durante a alimentação sanguínea de fêmeas adultas que o parasito *Leishmania* é ingerido (amastigota) ou transmitido (promastogotas metacíclicas). No entanto, estes parasitos, após ingestão, se desenvolvem no intestino médio até atingirem sua forma metacíclica e estão em contato direto com as enzimas digestivas, não somente proteases mas também as carboidrases. Devido a este fato, e também à importância dos carboidratos na processo de obtenção de energia destes insetos, é essencial investigar o papel fisiológico e bioquímico das carboidrases.

1.4.4. Principais enzimas envolvidas na digestão em flebotomíneos adultos: proteases, lipases e carboidrases

Durante o processo de digestão em insetos, enzimas capazes de atuar na clivagem de proteínas, carboidratos e lipídeos devem ser secretadas, independente do hábito alimentar dos insetos. De maneira geral, os insetos possuem o mesmo espectro enzimático, no entanto estas enzimas podem ser produzidas em maior ou menor quantidade, refletindo o tipo de alimentação ingerida e o estágio de cada espécie (TERRA; FERREIRA, 2012).

As proteases atuam na digestão de proteínas, clivando as ligações peptídicas e liberando aminoácidos, os quais são absorvidos pelas células do intestino médio. Insetos produzem proteases em abundância e estas atuam principalmente na digestão de proteínas adquiridas na dieta, como as proteínas do sangue no caso de insetos hematófagos. Os insetos também produzem diversas proteases não-digestivas, por exemplo, as que possuem diferentes papéis em cascatas de sinalização, ativação de zimógeno, defesa contra infecções, apoptose, entre outras funções (KANOST; CLEM, 2012). No geral as proteases são classificadas de acordo com a natureza do seu sítio ativo ou de acordo com a posição na qual elas clivam a cadeia peptídica. Sendo assim, dentre as proteases estão as carboxipeptidases e aminopeptidases, que removem os aminoácidos terminais a partir das extremidades C- e N-terminal das cadeias polipeptídicas, respectivamente. Estas proteases são chamadas exopeptidases. As endopeptidases clivam as ligações peptidicas internas das cadeias proteicas (KANOST; CLEM, 2012). Dentre as principais endopeptidases estão as serino proteases. Nestas proteínas, o grupo hidroxila da cadeia lateral do resíduo de serina presente no sítio catalítico atua como nucleófilo na reação de hidrólise da ligação peptídica. Os principais representantes deste grupo são as tripsinas e as quimiotripsinas (KANOST; CLEM, 2012; TERRA; FERREIRA, 2005, 2012).

As lipases (EC 3.1.1.3) são as enzimas menos estudadas dentre os três grupos de hidrolases envolvidos na digestão de insetos. De maneira geral, elas clivam as ligações carboxil-

ésters em triacilgliceróis, diacilgliceróis, galactolipídeos e fosfolipídeos (DOUGLAS, 2013). As fosfolipases removem a porção de ácido graxo dos fosfolipídeos. As triacilglicerol lipases atuam sobre as ligações éster externas dos triacilglicerois e as esterases atuam em moléculas que estão completamente dissolvidas na água, hidrolisando esteres carboxílicos em álcool e carboxilato, e atuam sobre moléculas de colesterol (TERRA; FERREIRA, 2005, 2012). A lipólise ocorre na região proximal do instetino médio dos insetos, liberando ácidos graxos livres, glicerol, acilglicerois parciais e lisofosfolipídeos que são então absorvidos pelas células do intestino (DOUGLAS, 2013).

Os carboidratos adquiridos na dieta são absorvidos no intestino médio dos insetos na forma de monossacarídeos (glicose, frutose, etc). Sendo assim, os dissacarídeos e polissacarídeos ingeridos precisam ser hidrolisados em moléculas menores. As carboidrases ou glicosidases secretadas no IM hidrolisam as ligações glicosídicas em resíduos de açúcares ou polissacarídeos complexos. A especificidade dessas enzimas depende do tipo de ligação a ser clivada e do tipo de substrato que reconhecem. Dentre as enzimas que participam da digestão inicial e que agem sobre polissacarídeos complexos estão as amilases, quitinases, celulases, hemicelulases e pectinases (TERRA; FERREIRA, 2005).

Os polissacarídeos amido e glicogênio são as principais reservas energéticas em plantas e insetos e constituem uma grande fonte de energia. Estes são principalmente digeridos pelas α-amilases que hidrolisam as ligações α-1,4-glicosídicas, mas várias enzimas estão envolvidas no processo. De maneira geral, o amido é constituído por dois tipos de polímeros de glicose, amilose e amilopectina. A amilose é uma grande cadeia linear composta por resíduos de glicose em ligações α-1,4, enquanto a amilopectina consiste em pequenas cadeias de glicose em ligações α-1,4 com cadeias laterais de glicose ligadas à cadeia principal por ligações α-1,6. Dessa forma, existem 3 grupos de enzimas responsáveis pela degradação do amido, as endoamilases, as exoamilases e as enzimas desramificadoras. Enquanto as α-amilases (endoamilase) clivam ligações α-1,4, as exoamilases como as glicoamilases podem clivar tanto as ligações α-1,4 como α-1,6. As enzimas desramificadoras como as pululanase tipo I e isoamylase hidrolisam apenas as ligações α-1,6 (VAN DER MAAREL et al., 2002). O glicogênio também é constituído por uma cadeia de resíduos de glicose em ligações α-1,4 e ramificações de resíduos de glicose ligados à cadeia principal por ligações α-1,6. Este polímero também contem uma pequena quantidade de glicosamina e fosfato (ADEVA-ANDANY et al., 2016). Durante a digestão em insetos este polímero sofre a ação das mesmas enzimas descritas na degradação do amido. As α-amilases, de maneira geral, já foram descritas e caracterizadas

em vários insetos, incluindo a elucidação da estrutura 3D de uma α-amilase de larvas de *Tenebrio molitor* (TERRA; FERREIRA, 2012).

Dentre as enzimas que atuam na digestão final e que estão envolvidas na hidrólise de oligo e dissacarídeos estão as α-glicosidases, que hidrolisam ligações em α-glicosídeos como sacarose, maltose, trealose, melezitose; β-glicosidases, que hidrolisam ligações em beta-glicosídeos como celobiose e gentiobiose; e α-galactosidases, que hidrolisam ligações em melibiose e rafinose. Alguns destes dissacarídeos são obtidos diretamente a partir da alimentação açucarada e outros são resultados da hidrólise dos polissacarídeos. A celobiose, por exemplo, é obtida a partir da hidrólise de celulose e a maltose é obtida como resultado da hidrólise do amido (TERRA; FERREIRA, 2005).

As α-glicosidases são o objeto de estudo deste trabalho. Portanto, suas especificidades, estrutura, função e suas características em insetos serão apresentadas em maior profundidade.

1.4.5. Alfa-glicosidases

As α-glicosidases (EC 3.2.1.20) são exoglicosidases que catalisam a hidrólise de ligações α-1,4, liberando um resíduo de glicose a partir da extremidade não-redutora de aril-glicosídeos, dissacarídeos e oligossacarídeos com eficiência variada (TERRA et al., 1994).

De maneira geral, as proteínas são agrupadas em famílias com base na similaridade da sequência de aminoácidos (estrutura primária) e no seu dobramento (estrutura secundária e terciária). Em uma mesma família de enzimas, α-hélices e folhas–β são localizadas em posições similares na conformação enovelada. Comumente inúmeros resíduos de aminoácidos conservados são observados nas imediações e no sítio ativo, no modelo proposto por Henrissat e colaboradores em 1996, utilizando a ferramenta HCA (Análise de Clusters Hidrofóbicos).

As α-glicosidases pertencem ao grupo das glicosídeo hidrolases. Este grupo de enzimas desempenha papéis críticos na biologia da digestão e decomposição de polissacarídeos, até a biossíntese de glicoproteínas. As α-glicosidases são um subconjunto importante dentro das glicosídeo hidrolases, e participam tanto no metabolismo primário (hidrólise de açúcares) como no processamento e biosíntese de glicoconjugados (LOVERING et al., 2005). Essas enzimas são principalmente encontradas nas famílias 13 e 31 e, em menor proporção, nas famílias 4, 63, 97 e 122 (CANTAREL et al., 2009). As α-glicosidases de insetos são classificadas nas famílias 13 e 31. A hidrólise enzimática das ligações glicosídicas ocorre por um mecanismo geral de catálise ácida que requer a ação da cadeia lateral de dois resíduos principais de aminoácidos, um doador de prótons e um nucleófilo catalítico.

α-glicosidases são amplamente distribuídas em microorganismos e em diferentes tecidos de animais e plantas, executando uma enorme gama de reações a partir de uma grande diversidade de especificidade de substratos. Algumas α-glicosidases, além de hidrolisar α-glicosídeos sintéticos e oligossacarídeos, são capazes de atuar em polissacarídeos complexos como amido e glicogênio (CHIBA, 1997), mas de maneira geral, estas enzimas possuem maior especifidade e atuam com uma velocidade maior sobre oligossacarídeos em relação aos polissacarídeos. Sendo assim, essas enzimas podem receber diferentes denominações, de acordo com sua especificidade ou tecido em que são encontradas em maior abundância. Por exemplo, α-glicosidases neutras ou lisossomais, maltase, maltase ácida, maltase-glicoamilase, exo-1,4-α-glicosidase, sacarase-α-glicosidase (E.C. 3.2.1.48), sacarase isomaltase, sacarase, ou oligo-1,6-glicosidase (E.C. 3.2.1.10). Em alguns casos, o número de classificação (E.C. number) pode ser diferente, mas essencialmente todas as atividades citadas se encaixam na descrição de α-glicosidases. A nomenclatura sacarase é um termo que deve ser evitado, visto que a sacarose pode ser hidrolisada tanto por α-glucosidases como por β-frutosidases. (GONTIJO et al., 1998).

1.4.5.1. Glicosídeo hidrolases pertencentes à família 13

Enzimas pertencentes à família GH13 foram descritas em vários organismos pertencentes a todos os reinos. Esta é a maior dentre as famílias das glicosídeo hidrolases e possui uma grande variedade de enzimas. Estas atuam em diversos substratos hidrolisando ligações α-glicosídicas, gerando mono ou oligosacarídeos α-anoméricos, ou formam ligações α-glicosídicas por transglicosilação (KURIKI; IMANAKA, 1999). É interessante notar a presença de transferases e isomerases classificadas nesta família, assim como transportadores de aminoácidos, que não possuem atividade catalítica. As atividades descritas nesta família estão indicadas na Tabela 1.

Tabela 1: Atividades descritas na família GH13

Enzima	**Número de classificação (E.C.)**
α-amilase	3.2.1.1
Oligo-1,6-glicosidase	3.2.1.10
α-glicosidase	3.2.1.20
Amilo-α-1,6-glicosidase	3.2.1.33
Pululanase	3.2.1.41

Ciclomaltodextrinase	3.2.1.54
Trealose-6-fosfato hidrolase	3.2.1.93
Isoamilase	3.2.1.68
Amilase Maltogênica	3.2.1.133
Neopululanase	3.2.1.135
Malto-oligosiltrealose trealohidrolase	3.2.1.141
Limit dextrinase	3.2.1.142
Glicodextranase	3.2.1.70
Amilosacarase	2.4.1.4
Sacarose fosforilase	2.4.1.7
Maltotriose-forming α-amylase	3.2.1.116
Maltotetraose-forming α-amylase	3.2.1.60
Maltopentaose-forming α-amylase	3.2.1.-
Maltohexaose-forming α-amylase	3.2.1.98
Branching enzyme	2.4.1.18
Enzima desramificadora de glicogênio	2.4.1.25/3.2.1.33
α-1,4-glican: fosfato α-maltosiltransferase	2.4.99.16
Ciclomaltodextrin glicanotransferase (CGTase)	2.4.1.19
4-α-glicanotransferase	2.4.1.25
Malto-oligosiltrehalose sintase	5.4.99.15
Isomaltulose sintase	5.4.99.11
Trealose sintase	5.4.99.16
Transportadores de aminoácidos	Não Possui

Dados obtidos do banco de dados CAZY (CANTAREL et al., 2009).

Devido à grande quantidade de sequências e diversidade de especificidades para estas enzimas, esta família foi subdividida em quase 40 subfamílias (STAM et al., 2006). As subfamílias possuem sequências com maior similaridade e, portanto, propriedades bioquímicas em comum. Por exemplo, as atividades de α-amilase e α-glicosidase de metazoários e *alpha-1,4-glucan branching* de eucariotos foram classificadas nas subfamílias GH13_15, GH13_17 e GH13_8, respectivamente. Os transportadores de aminoácidos foram classificados nas subfamílias GH13_34 e GH13_35 (STAM et al., 2006).

Como descrito acima, as enzimas GH13 atuam sobre uma diversa gama de substratos e possuem diferentes especificidades. No entanto, esta família recebe a descrição geral de

"família das α-amilases", pois as principais enzimas envolvidas na hidrólise e síntese do amido estão classificadas nesta família, sendo estas as α-amilases que clivam as ligações α-1,4 glicosídicas, pululanase ou isoamilase clivando as ligações α-1,6-glicosídicas, cyclodextrin glucano transferases as que fazem a transglicosilação de ligações α-1,4 (CGTases) e a *alpha-1,4-glucan branching enzymes* que realizam a transglicosilação das ligações α-1,6 (KURIKI; IMANAKA, 1999).

As α-amilases, de maneira geral, catalisam a hidrólise de ligações do tipo α-1,4-glicosídicas em polissacarídeos de alto peso molecular como amido e glicogênio, ou de outros polissacarídeos relacionados, como amilose e amilopectina, e são consideradas como endo enzimas. No entanto, existem algumas enzimas que atuam na hidrólise de amido, liberando resíduos de mono a hexassacarídeos a partir da extremidade, que são classificadas como exo enzimas e recebem outro número de classificação, apesar de relizarem a hidrólise do mesmo tipo de ligação, como a maltotriohidrolase (3.2.1.116), maltotetrahidrolase (3.2.1.60), e maltohexahidrolase (3.2.1.98) produzindo maltotriose, maltotetraose e maltohexaose, respectivamente (MACGREGOR; JANEČEK; SVENSSON, 2001). Além disso, as α-amilases possuem um sítio de ligação a cálcio e, em alguns casos, são ativadas pelo íon cloreto (TERRA et al., 1994).

Em plantas e bactérias, as pululanases e isoamilases participam na desramificação do glicogênio e do amido, possuindo apenas a atividade de α-1,6-glucosidase. Em animais e fungos, a GDE (*glycogen debranching enzyme*), uma enzima altamente conservada, participa na desramificação do glicogênio. Este é um polímero de reserva e possui resíduos de glicose em ligação α-1,4 e diversas ramificações de glicose ancoradas à cadeia principal por ligaçãoes α-1,6. A GDE é uma proteína monomérica, possuindo a atividade de α-1,4-glucanotransferase em sua região N-terminal e de α-1,6-glicosidase na região C-terminal (ZHAI et al., 2016). Primeiro a atividade de tranferase remove três resíduos de glicose da ramificação (que possui quatro resíduos de glicose), para a cadeia linear em ligação α-1,4. Com a atividade de glicosidase, retira então o resíduo de glicose final em ligação α-1,6, permitindo o acesso de outra enzima, a glicogênio fosforilase, à cadeia linear com resíduos em ligação α-1,4 (ADEVA-ANDANY et al., 2016). As *alpha-1,4-glucan branching* geralmente hidrolisam uma ligação α-1,4 glicosídica e sintetizam uma ligação α-1,6 glicosídica, criando ramificações no glicogênio ou amilopectina. Essas enzimas promovem a remoção de um oligossacarídeo em ligação α-1,4 da extremidade não-redutora de uma cadeia de glucana linear e a ligação deste oligossacarídeo a uma outra cadeia de oligossacarídeo através de ligação α-1,6, formando uma ramificação (MACGREGOR; JANEČEK; SVENSSON, 2001).

Outro membro desta família, são as cadeias pesadas dos transportadores de aminoácidos (rBAT ou antígeno de superfície 4F2hc) que perderam sua atividade catalítica mas possuem uma alta similaridade de sequência com as α-glicosidases e com as oligo-1,6-α-glicosidases (GABRIŠKO; JANEČEK, 2009; JANECEK, 1997). As cadeias pesadas dos transportadores de aminoácidos possuem a função de localizar e dobrar corretamente a cadeia leve na membrana plasmática, com 12 domínios transmembranares. A cadeia leve é que possui propriamente a função de transportar aminoácidos (GABRIŠKO; JANEČEK, 2009).

As enzimas presentes na família GH13 catalizam a hidrólise de seus substratos através do mecanismo de retenção de configuração anomérica da glicose liberada. Em geral, uma tríade catalítica (resíduos Asp, Glu e Asp) presente no sítio ativo participa do mecanismo de hidrólise.

1.4.5.2. Glicosídeo Hidrolases pertencentes à família 31

Enzimas pertencentes à esta família foram descritas em vários organismos, incluindo arqueas, bactérias, plantas e animais, e são representadas por diferentes atividades hidrolíticas. Essas atividades estão descritas na Tabela 2.

Tabela 2: Atividades descritas na família GH31

Enzima	**Número de classificação (E.C.)**
α-glicosidase	3.2.1.20
α-1,3-glicosidase	3.2.1.84
Sacarose-isomaltase	3.2.1.48 e 3.2.1.10
α-glucan liase	4.2.2.13
α-xilosidase	3.2.1.177
α-galactosidase	3.2.1.22
α-manosidase	3.2.1.24
Isomaltosiltransferase	2.4.1.-
Oligosacarídeo α-1,4-glucosiltransferase	2.4.1.161
Sulfoquinovosidase	3.2.1.-

Dados obtidos do banco de dados CAZY (CANTAREL et al., 2009).

Enzimas pertencentes à família GH31 possuem o mecanismo catalítico comum de clivagem da unidade de carboidrato terminal de substratos que variam consideravelmente em tamanho, desde dissacarídeos a polissacarídeos complexos de reserva como amido, glicogênio ou glicoproteínas (ERNST et al., 2006). Na maioria dos casos, a porção reconhecida no sítio de ligação ao substrato é uma α-D-glicose, mas alguns membros desta família possuem preferência por α-D-xilose.

Dentre as enzimas mais bem caracterizadas dessa família estão as α-glicosidases. Estas estão implicadas em pelo menos três processos biológicos distintos. No catabolismo, a α-glicosidase lisossomal é essencial para a degradação do glicogênio e em humanos, deficiência ou mau funcionamento dessa enzima causa a doença de armazenamento de glicogênio II, também conhecida como doença de Pompe (HOEFSLOOT et al., 1988). No processamento de glicoproteínas, a α-glicosidase II participa no controle de qualidade para dobramento e maturação de glicoproteínas no retículo endoplasmático (TROMBETTA; FLEMING; HELENIUS, 2001; TROMBETTA; SIMONS; HELENIUS, 1996). No metabolismo primário, as enzimas intestinais sucrase-isomaltase e maltase-glucoamilase (MGAM) desempenham papéis-chave no estágio final da digestão de carboidratos (NICHOLS et al., 2003), atuando na hidrólise de ligações α-1,6 em oligossacarídeos produzidos durante a hidrólise do amido e do glicogênio.

Além das atividades citadas, uma proteína integral de membrana localizada no envelope nuclear foi recentemene caracterizada, a NET37 (CHEN et al., 2006). Essa proteína possui um domínio transmembrana na região N-terminal e na região C-terminal um domínio pertencente à família GH31. Estudos demonstraram que a porção pertencente à atividade de glicosidase está localizada na região do envelope nuclear voltada para o retículo endoplasmático. Foi demonstradado que esta proteína é necessária para a diferenciação miogênica de células C2C12 e que esta atividade é dependente do sítio catalítico da região glicosídica (DATTA; GUAN; GERACE, 2009).

As enzimas da família GH31 catalizam a hidrólise de seus substratos através do mecanismo de retenção da configuração anomérica da glicose. Em geral, dois resíduos de ácido aspártico presentes no sítio ativo participam como o nucleófilo catalítico e o grupo ácido/base, respectivamente.

1.4.5.3. Mecanismo de hidrólise de enzimas GH13 e GH31

As glicosídeo hidrolases pertencentes às famílias GH13 e GH31 hidrolisam seus substratos pelo mecanismo de retenção da configuração anomérica, também conhecido como mecanismo clássico (KOSHLAND, 1953). Esta forma de hidrólise permite que estas enzimas realizem transglicosilação. A retenção de configuração ocorre através de duas etapas, glicosilação e deglicosilação, sendo este conhecido como mecanismo de deslocamento duplo involvendo um intermédiário glicosil-enzima. Esta reação ocorre com a assistência ácido/básica e nucleofílica da cadeia lateral de dois aminoácidos, geralmente um glutamato e um aspartato

para enzimas da família GH13 e dois aspartatos para enzimas da família GH31. Na primeira etapa da reação, o nucleófilo ataca o centro anomérico causando o deslocamento da porção aglicona do substrato e ocorre a formação do intermediário glicosil-enzima. Ao mesmo tempo o resíduo que funciona como ácido/base catalítico protona o oxigênio catalítico enquanto a ligação é clivada. Na etapa de deglicosilação o processo é revertido e uma molécula de água ataca com a assistência básica do ácido/base catalítico para liberar o nucleófilo catalítico da molécula de glicose. As duas etapas deste mecanismo passam por um estado de transição com a formação de íon tipo oxocarbênio.

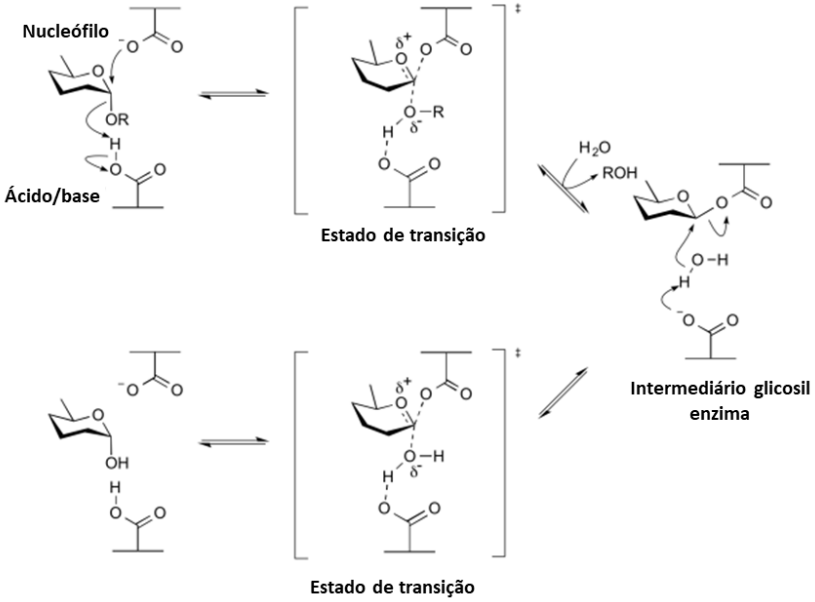


Figura 11: Mecanismo de hidrólise de ligações glicosídicas com retenção de configuração por α-glicosidases. Fonte: CAZypedia, 2018 (adaptado).

1.4.5.4. Estrutura e resíduos catalíticos em enzimas GH13 e GH31

A família GH13, juntamente com as famílias GH70 e GH77, forma o clã GH-H. O clã é um nível de organização hierárquico acima das classificações em famílias do CAZy. Famílias de um mesmo clã possuem um ancestral e uma maquinaria catalítica comum, no entanto possuem baixa similaridade de sequência (DAVIES; HENRISSAT, 1995). Devido à importância das enzimas classificadas nessa família, como α-amilases e α-glicosidases, por exemplo, várias estruturas foram resolvidas por cristalografia e até agora mais de 110 membros desta família possuem sua estrutura 3D elucidada (CAZy, 2018). As enzimas GH13 possuem um núcleo central conservado composto por três domínios, A, B e C (RAMASUBBU et al., 1996). O domínio A é o domínio catalítico, com uma estrutura barril $(\beta/\alpha)_8$ (ABAD et al., 2002;

BRAYER; LUO; WITHERS, 1995; KANAI et al., 2001); o domínio B é um loop de tamanho variável localizado entre a folha β3 e a hélice α3 do barril $(\beta/\alpha)_8$ (JANECEK, 1997) e o domínio C está localizado na porção C-terminal, sendo caracterizado por um motivo chave grega (JANECEK, 1997; RAMASUBBU et al., 1996). Além do núcleo conservado, alguns membros da família GH13 possuem extensões N- ou C-terminal, como os CBMs (módulos de ligação a carboidratos) (JANECEK, 1997).

A família das α-amilases, devido à grande quantidade de atividades envolvidas, possui uma baixa similaridade de sequências. No entanto, proteínas classificadas nessa família possuem várias regiões com sequências de aminoácidos bem definidas e altamente conservadas. Este é o caso das regiões conservadas que carregam os aminoácidos catalíticos e os sítios de ligação a cálcio e cloreto, além de outras regiões que demarcam o sítio ativo (JANECEK, 1997; JANEČEK, 1995). Estas regiões conservadas foram descritas nas folhas β2, β3, β4, β5, β7, β8 e loop 3 da estruttura do barril $(\beta/\alpha)_8$. Devido ao aumento significativo de sequências sendo designadas à família GH13, estas sete regiões foram definidas para ajudar a atribuir a especificidade enzimática correta aos membros da família α-amilase.

Os aminoácidos catalíticos estão localizados em uma fenda entre os domínios A e B (BRZOZOWSKI; DAVIES, 1997). O sítio catalítico é formado pelas extensões C-terminal das folhas β4, β5 e β7, que carregam a tríade catalítica aspartato, glutamato e aspartato, respectivamente (KADZIOLA et al., 1994; MATSUURA et al., 1984; QIAN; HASER; PAYAN, 1993). O primeiro aspartato funciona como nucleófilo catalítico e o glutamato como ácido/base geral. O segundo aspartato participa na estabilização do estado de transição e também mantendo o glutamato com a correta protonação para atividade (MACGREGOR; JANEČEK; SVENSSON, 2001; UITDEHAAG et al., 1999). Um quarto resíduo conservado foi determinado como uma arginina posicionada dois resíduos precedentes ao nucleófilo catalítico (MACGREGOR; JANEČEK; SVENSSON, 2001). O domínio B possui uma região conservada (loop 3) contendo um resíduo de aspartato, sendo este um dos sítios de ligação a cálcio. Este é o caso do Asp175 da TAKA-amylase A (BOEL et al., 1990). Os outros resíduos conservados são uma asparagina localizada na folha β3 e uma histidina na folha β4, estes estão próximos a fenda de ligação ao substrato. Estes resíduos podem também estar envolvidos na ligação de cálcio (BOEL et al., 1990; JANECEK, 1997; MACHIUS; WIEGAND; HUBER, 1995). A ligação do cálcio preserva a integridade estrutural do sítio ativo através da ligação de dois fragmentos do domínio A e do domínio B (BUISSON et al., 1987). Além disso, algumas α-amilases, como a α-amilase humana, possuem um sítio de ligação a cloreto (BRAYER; LUO; WITHERS, 1995). As α- amilases de insetos são dependentes de cálcio e apenas algumas delas

são ativadas por cloreto. Uma exceção aplica-se para algumas espécies de lepidópteros (TERRA et al., 1994; TERRA; FERREIRA, 2012). O sítio de ligação a cloreto é localizado próximo ao centro da estrutura barril $(\beta/\alpha)_8$, caracterizado pela presença de um resíduo de arginina conservado posicionado dois resíduos precedendo o nucleófilo catalítico na folha β4, um resíduo de asparagina posicionado dois resíduos precedendo o aspartato catalítico na folha β7 e um resíduo Arg/Lys presente na região conservada RVMSSY, na folha β8 (D'AMICO; GERDAY; FELLER, 2000; JANECEK, 1997). O cloreto funciona como um ativador alostérico. A mudança conformacional da enzima com a ligação ao cloreto causa uma variação no ambiente do doador de prótons, levando ao aumento do pKa da cadeia lateral deste resíduo de aminoácido. Portanto, a ligação de cloreto provoca o deslocamento do pH ótimo, alterando-o de ácido para ligeiramente básico (D'AMICO; GERDAY; FELLER, 2000; FELLER et al., 1996; TERRA; FERREIRA, 2012). A vantagem evolutiva desse mecanismo de regulação, pode estar relacionada à ampliação do perfil de atuação dessas enzimas em diferentes valores de pH.

Dentre as primeiras estruturas cristalizadas na família GH13 estão a TAKA-amylase A do fungo *Aspergillus oryzae* (MATSUURA et al., 1984) e a α-amilase pancreática suína (BUISSON et al., 1987). Estas auxiliaram na elucidação dos aminoácidos envolvidos na catálise. Na Taka amylase A os resíduos Asp206 dentro da região conservada GLRIDTVKH, Glu230 e Asp 297 formam a tríade catalítica. Em insetos, a α-amilase do tubo digestivo de larvas do coleóptero *Tenebrio molitor* teve sua estrutura 3D elucidada, sendo os aminoácidos Asp185, Glu222 e Asp 287 os resíduos que formam a tríade catalítica (STROBL et al., 1998).

A família GH31, juntamente com as famílias GH27 e GH36, compõem o clã GH-D. Das estruturas elucidadas, foi demonstrado que uma estrutura barril $(\beta/\alpha)_8$ é comum no domínio catalítico dos representantes desta família (ERNST et al., 2006; LOVERING et al., 2005; SIM et al., 2008). A primeira estrutura cristalizada a ser elucidada foi a da α-xilosidase YicL da bactéria *Escherichia coli* (LOVERING et al., 2005). A elucidação desta estrutura permitiu também a confirmação do envolvimento de um resíduo de aspartato como ácido/base catalítico no processo de hidrólise. Anteriormente o resíduo Asp647 da α-glicosidase da levedura *Schizosacharomyces pombe* havia sido classificado como o possível ácido/base catalítico, através da comparação de sequências e estudos cinéticos de mutantes (OKUYAMA et al., 2001). O nucleófilo catalítico de enzimas GH31 foi definido como o aspartato presente na região conservada WIDMNE. O primeiro nucleófilo a ser identificado foi o resíduo Asp224 da α-glicosidase de *Aspergillus niger* (LEE; HE; WITHERS, 2001). Em bactérias, o resíduo Asp416 dentro da região conservada FKTDFG foi identificado na α-xilosidase YicL de *Escherichia coli* (LOVERING et al., 2005). A estrutura da α-glicosidase da arquea *Sulfolobus*

solfataricus (ERNST et al., 2006) foi elucidada como um monômero composto por 4 domínios estruturais, sendo o domínio catalítico central com a topologia barril $(\beta/\alpha)_8$ flanqueado por três domínios sanduíche-β (N, C e D). Os resíduos catalíticos foram definidos como Asp320, sendo o nucleófilo catalítico e Asp 416 como doador de prótons.

A maioria dos membros da família GH31 são proteínas multi-domínios, mas a função específica destes domínios acessórios é geralmente desconhecida. Por exemplo, nas α-glucosidases neutras um dos domínios acessórios tem atividade de galactose mutarotase. A α-xilosidase *Cj*Xyl31A da bactéria *Cellvibrio japonicus*, por exemplo, possui um domínio acessório (PA-14) e este confere a esta enzima uma especificidade catalítica aumentada na hidrólise de oligossacarídeos de cadeia longa (LARSBRINK et al., 2011).

Em geral, as famílias GH13 e GH31 utilizam o mecanismo de retenção para hidrólise de seus substratos, e possuem enzimas com especificidade de α-glicosidases. No que tange às sequências de aminoácidos, a similaridade não é óbvia, mas existe a conservação do nucleófilo catalítico aspartato na folha β4 em ambas as famílias (JANEČEK; SVENSSON; MACGREGOR, 2007). No entanto, na família GH13 o resíduo ácido/base geral glutamato está localizado na folha β5 (JANECEK, 1997) e na família GH31 o resíduo ácido/base geral é um aspartato localizado na folha β6 (LOVERING et al., 2005).

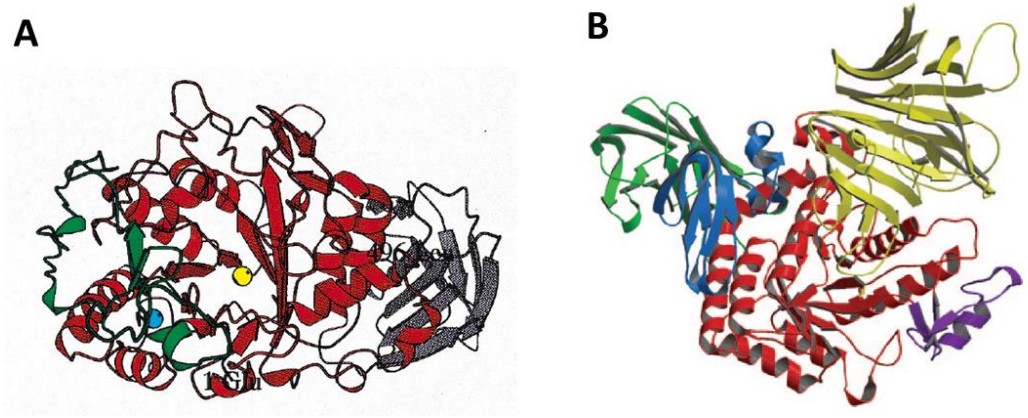


Figura 12: Estrutura 3D de representantes das famílias GH13 e GH31. As α-hélices estão representadas por espirais e as folhas β por setas. **A:** Taka amilase A de *Aspergillus oryzae* (Matsura et al., 1984). Domínios A, B e C estão representados pelas cores vermelho, verde e azul claro respectivamente. **B:** Monômero α-xilosidase YicL de *Escherichia coli* (Lovering et al., 2005). O domínio A está representado em vermelho, o domínio B está representado em roxo e o domínio C está representado nas cores verde e azul. Fonte: LOVERING et al., 2005; MACHIUS; WIEGAND; HUBER, 1995 (adaptado).

1.4.5.5. α-glicosidases em insetos

Em insetos, essas enzimas são agrupadas nas famílias 13 e 31 (CANTAREL et al., 2009) e estão especialmente envolvidas na hidrólise da sacarose para obtenção de energia através da

absorção de monômeros de glicose. As α-glicosidases mais bem descritas em insetos pertencem à família GH13. Estas foram descritas como enzimas solúveis ou ligadas à membrana no intestino médio de flebotomíneos (DILLON; EL KORDY, 1997; GONTIJO et al., 1998; JACOBSON; SCHLEIN, 2001) e mosquitos (BILLINGSLEY; HECKER, 1991; SOUZA-NETO et al., 2007; ZHENG et al., 1995) ou também encontradas nas glândulas salivares de flebotomíneos (JACOBSON; SCHLEIN, 2001) e mosquitos (MARINOTTI; DE B&O; MOREIRU, 1996; MARINOTTI; JAMES, 1990; MOREIRA-FERRO; MARINOTTI; BIJOVSKY, 1999). α-glicosidases estão associadas às membranas perimicrovilares de *Dysdercus peruvianus* (SILVA; TERRA, 1995), *Acyrtosiphon pisum* (CRISTOFOLETTI et al., 2003) e *Rhodnius prolixus* (SILVA et al., 2004), por exemplo.

Devido à abundância destas proteínas em membranas, elas têm sido utilizadas como marcadores moleculares em hemípteros (SILVA et al., 2007) e também estão associadas à ligação de endotoxinas. As α-glicosidase (Cpm1) microvilar de *Culex pipiens* e (Cqm1) de *Culex quinquefasciatus* são receptores da toxina binária de *Lysinibacillus sphaericus*, estas proteínas possuem 97 % de identidade e estão ligadas à membrana por uma âncora de glicosil fosfatidil inositol (GPI) (DARBOUX et al., 2001; GUO et al., 2013). No entanto, o ortólogo destas enzimas (Aam1) em *Aedes aegypti* não possui a capacidade de se ligar a esta toxina e portanto, esta espécie é refratária ao *L. sphaericus* (DO NASCIMENTO et al., 2017; FERREIRA et al., 2010). A α-glicosidase (Agm3) de *Anopheles gambiae* também caracterizada como receptora da toxina binária de *Lysinibacillus sphaericus*, possui também alta afinidade pela toxina Cry11Ba de *Bacillus thurigiensis* (OPOTA et al., 2008; ZHANG et al., 2013).

Em insetos hematófagos, durante a digestão do sangue esta enzima pode também participar nas etapas finais da digestão de açúcares presentes nos gliconconjugados do sangue. O sangue possui predominantemente proteínas como componentes e uma alta proporção são glicoproteínas. É possível que os fragmentos de glicopeptídeos, produzidos após clivagem proteolítica por tripsina e aminopeptidase, sejam suscetíveis à ação de exoglicosidases, assim como os glicolípideos. Alguns trabalhos descreveram, por exemplo, o aumento de α-glicosidases após a ingestão sanguínea em insetos como *P. langeroni* (DILLON; EL KORDY, 1997). O aumento da atividade de α-glicosidase na membrana perimicrovilar de *Rhodnius prolixus* após alimentação sanguínea também foi demonstrado (SILVA et al., 2007). Nesse caso, o aumento da atividade de α-glicosidase está relacionado à formação de hemozoína, mecanismo importante no processo de detoxificação do heme (MURY et al., 2009).

Como descrito anteriormente, devido ao mecanismo de hidrólise utilizado (retenção da configuração anomérica), enzimas pertencentes às famílias GH13 e GH31 são capazes de

realizar transglicosilação. α-glicosidases com atividade de transglicosilação já foram descritas em diferentes insetos, por exemplo, *Apis mellifera* (HUBER; THOMPSON, 1973) e *Acyrthosiphon pisum* (CRISTOFOLETTI et al., 2003). A α-glicosidase purificada a partir do intestino médio de *Acyrthosiphon pisum* catalisa reações de transglicosilação na presença de excesso de sacarose, liberando glicose da sacarose sem aumentar a osmolaridade do meio (CRISTOFOLETTI et al., 2003). Esse fenômeno associado à rápida absorção de frutose (ASHFORD; SMITH; DOUGLAS, 2000) explica porque a osmolaridade luminal do intestino médio diminui à medida que a seiva de floema contendo sacarose ingerida passa ao longo do intestino médio do afídeo.

Existem diversas α-glicosidases descritas baseando-se apenas na comparação de sequências. No entanto, apenas uma fração dessas enzimas foram caracterizadas bioquimicamente. Propriedades bioquímicas de algumas enzimas purificadas ou não já foram determinadas. Em geral, a massa molecular destas proteínas varia entre 60-80 kDa, ou é múltiplo desses valores no caso de proteínas oligoméricas. O pH ótimo varia entre 5-6,5, dependendo do compartimento em que estas enzimas são encontradas (TERRA et al., 1994).

Em *L. longipalpis*, a atividade de α-glicosidase foi caracterizada com pH ótimo de 5,8, consistente com os resultados encontrados para o pH do intestino médio de aproximadamente 6,0. Um ponto interessante a se notar, é que durante a digestão sanguínea o pH do intestino médio abdominal torna-se alcalino. No entanto o intestino médio torácico, onde localiza-se a maior parte das α-glicosidases, permanece ácido (5,5-6,0), de acordo com o pH ótimo encontrado para essa enzima (GONTIJO et al., 1998; SANTOS et al., 2008). Para *P. langeroni*, foi descrito uma α-glicosidase alcalina com pH ótimo em torno de 7,5 e induzida em fêmeas alimentadas com sangue (DILLON; EL KORDY, 1997). O pH ótimo determinado para α-glicosidase salivar e do intestino médio de *Glyphodes pyloalis* (lepidóptero) foi de 7,5 e 8-9, respectivamente, e de *Chilo supressalis* foi de aproximadamente 8,0 para ambos os tecidos (GHADAMYARI; HOSSEININAVEH; SHARIFI, 2010; ZIBAEE; BANDANI; RAMZI, 2009). Estes resultados indicam que α-glicosidases de lepidópteras atuam em um ambiente mais básico. No lado oposto, uma α-glicosidase atuando com pH ótimo de 4,0-6,0, foi descrita para o hemíptera *Eurygaster integriceps* (MEHRABADI et al., 2009).

Em insetos a especificidade destas enzimas sobre diferentes dissacarídeos e oligossacarídeos é muito variada e provavelmente está associada ao tipo de alimentação. Atividades de α-glicosidases que hidrolisam sacarose melhor que maltose, estão relacionadas aos altos níveis de sacarose encontrado nas dietas, como por exemplo em abelhas e flebotomíneos. No caso de *Apis melifera*, três α-glicosidases (I, II e III) foram estudadas. Elas

são codificadas pelos genes hbg1, hbg2, and hbg3, respectivamente, e diferem entre si quanto à sua especificidade de substrato e pH ótimo (KIMURA et al., 1990; KUBOTA et al., 2004; NISHIMOTO et al., 2001; TAKEWAKI et al., 1993). Em *L. longipalpis* a atividade de α-glicosidase utilizando o homogenato do intestino médio de fêmeas não alimentadas apresentou maior atividade sobre sacarose em comparação com outros di e trissacarídeos, como maltose, melibiose, trealose e rafinose (GONTIJO et al., 1998).

Os aminoácidos catalíticos Asp206, Glu271 e Asp297 em *Culex pipiens* foram descritos após clonagem e sequenciamento de Cpm1 (DARBOUX et al., 2001), sendo esta sequência relacionada à atividade caracterizada bioquimicamente de α-glicosidase. Mas várias α-glicosidases foram descritas apenas por comparação de sequências, através da identificação de domínios e regiões conservadas e aminoácidos catalíticos. Através de ferramentas de bioinformática, por exemplo, todas as α-gicosidases da família GH13 foram identificadas em três espécies de mosquitos: *Aedes aegypti*, *Anopheles gambiae* e *Culex quinquefasciatus* (GABRIŠKO, 2013). Até agora nenhuma α-glicosidase de inseto teve sua estrutura 3D obtida por cristalização.

2. Objetivos

2.1. Objetivo Geral

Investigar os aspectos moleculares e o papel fisiológico de glicosidases digestivas, com enfoque para α-glicosidases, em fêmeas de *Lutzomyia longipalpis*, avaliando o efeito dos diferentes tipos de alimentação sobre a atividade enzimática e infecção por *Leishmania*, e também avaliar o papel dessa enzima na interação vetor-parasito.

2.2. Objetivos Específicos

- Estabelecer protocolo para quantificação da atividade de α-glicosidases utilizando ensaios contínuos para substratos fluorescentes.

- Caracterizar a atividade de α-glicosidase nos tecidos de *L. longipalpis* e avaliar o perfil de atividade quando fêmeas adultas são submetidas a diferentes condições alimentares (ex. jejum, dieta rica em açúcares e dieta rica em proteínas).

- Determinar as características bioquímicas das α-glicosidases encontradas nos tecidos de *L. longipalpis* utilizando diferentes substratos (ex. substrato sintético fluorescente, substrato natural), para diferenciar as especificidades das enzimas descritas.

- Identificar no genoma de *L. longipalpis* os genes codificante para proteínas pertencentes as famílias 13 e 31 das glicosídeo hidrolases, realizando anotação e identificação das atividades relacionadas a cada um dos genes encontrados, com enfoque para as α-amilases e maltases.

- Estudar a expressão tecidual dos genes codificantes para as proteínas identificadas em fêmeas de *L. longipalpis* mantidas em diferentes regimes alimentares permitindo uma melhor caracterização das mesmas. Avaliar também o efeito da infecção por *L. mexicana* sobre a expressão das proteínas identificadas.

- Avaliar o efeito de diferentes alimentações, no processo de infecção e desenvolvimento de *L. mexicana*.

3. Estrutura do trabalho

Este trabalho é composto por um artigo sobre aprefeiçoamento metodológico, dois artigos principais e uma comunicação. O primeiro trabalho descreve a padronização da técnica de quantificação da atividade de glicosidases por ensaios contínuos utilizando substratos fluorescentes. O modelo utilizado para a maior parte dos estudos foi *Rhodnius prolixus*, no entanto também padronizamos a técnica para utilização em amostras de *L. longipalpis*. Esses resultados estão descritos no trabalho 1, e foram essenciais para a realização dos ensaios bioquímicos descritos no trabalho 2. Ensaios contínuos permitiram a utilização de uma menor quantidade de amostra, facilitando o estudo em insetos pequenos e permitindo a quantificação da atividade enzimática em amostras individuais e também em tecidos separados. No segundo trabalho, realizamos a caracterização bioquímica da atividade de α-glicosidase utilizando dois substratos, MUαGlu e sacarose. A atividade foi caracterizada nos tecidos de *L. longipalpis* sob diferentes dietas, foram observados diferentes padrões de atividade frente às condições testadas e substratos utilizados. Dados cromatográficos demonstraram que diferentes α-glicosidases estão envolvidas na hidrólise dos substratos utilizados, e também nos diferentes tecidos utilizados para análise. O terceiro trabalho descreve a identificação de genes pertencentes às famílias GH13 e GH31 no genoma de *L. longipalpis*, inclusive genes codificantes para α-amilase e maltase. Foram analisadas características básicas de suas estruturas, como a presença de aminoácidos catalíticos, domínios de membrana, massa molecular, composição de éxons e introns, entre outras. Uma análise comparativa de bioinformática com outros insetos foi possível identificar a expansão da família das α-amilases e retração no número de genes de maltases. Além disso, durante a infecção com *Leishmania mexicana*, observamos que os genes de maltases, identificados na análise de bioinformática, foram regulados negativamente. No quarto trabalho, apresentamos uma comunicação, demonstrando que para *L. mexicana*, o desenvolvimento dentro do vetor e o padrão de migração em direção a porção anterior do intestino médio são parcialmente dependentes da alimentação açucarada ingerida pelo flebotomíneo.

Artigo publicado

TRABALHO 1

Standardization of a continuous assay for glycosidases and its use for screening insect gut samples at individual and populational levels

Gerson S. Profeta, Jessica A. S. Pereira, Samara G. Costa, Patricia Azambuja, Eloi S. Garcia, Caroline da Silva Moraes and Fernando A. Genta

Front. Physiol. 8:308. doi: 10.3389/fphys.2017.00308

Artigos submetidos para publicação

TRABALHO 2

Characterization of α-glucosidases from *Lutzomyia longipalpis* reveals independent hydrolysis systems for plant or blood sugars

Samara G. da Costa, Paul Bates, Rod Dillon and Fernando Ariel Genta

Submetido ao periódico Frontiers in Physiology

TRABALHO 3

Study of glycoside hydrolase families 13 and 31 reveals expansion and diversification of α-amylase genes in the phlebotomine *Lutzomyia longipalpis*, and modulation of sand fly glycosidase activities by *Leishmania* infection

Samara G. da Costa, Paul Bates, Rod Dillon and Fernando Ariel Genta

Submetido ao periódico PLOS Neglected Tropical Diseases

TRABALHO 4

Development of *Leishmania mexicana* in *Lutzomyia longipalpis* in the absence of sugar feeding

Samara G. da Costa, Caroline da Silva Moraes, Paul Bates, Rod Dillon and Fernando Ariel Genta

Submetido ao periódico Memórias do Instituto Oswaldo Cruz

TRABALHO 1

Standardization of a continuous assay for glycosidases and its use for screening insect gut samples at individual and populational levels

Gerson S. Profeta, Jessica A. S. Pereira, Samara G. Costa, Patricia Azambuja, Eloi S. Garcia, Caroline da Silva Moraes and Fernando A. Genta

Front. Physiol. 8:308. doi: 10.3389/fphys.2017.00308

frontiers in Physiology

ORIGINAL RESEARCH
published: 12 May 2017
doi: 10.3389/fphys.2017.00308

Standardization of a Continuous Assay for Glycosidases and Its Use for Screening Insect Gut Samples at Individual and Populational Levels

Gerson S. Profeta[1], Jessica A. S. Pereira[1], Samara G. Costa[1], Patricia Azambuja[1,2], Eloi S. Garcia[1,2], Caroline da Silva Moraes[1] and Fernando A. Genta[1,2]*

[1] Instituto Oswaldo Cruz, Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil, [2] Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular, Rio de Janeiro, Brazil

OPEN ACCESS

Edited by:
Xanthe Vafopoulou, York University, Canada

Reviewed by:
Jalal Jalali Sendi, University of Gilan, Iran
Márcio Galvão Pavan, Oswaldo Cruz Foundation, Brazil

***Correspondence:**
Caroline da Silva Moraes
carolinemoraes83@gmail.com

Specialty section:
This article was submitted to Invertebrate Physiology, a section of the journal Frontiers in Physiology

Received: *21 February 2017*
Accepted: *28 April 2017*
Published: *12 May 2017*

Citation:
Profeta GS, Pereira JAS, Costa SG, Azambuja P, Garcia ES, Moraes CS and Genta FA (2017) Standardization of a Continuous Assay for Glycosidases and Its Use for Screening Insect Gut Samples at Individual and Populational Levels. Front. Physiol. 8:308. doi: 10.3389/fphys.2017.00308

Glycoside Hydrolases (GHs) are enzymes able to recognize and cleave glycosidic bonds. Insect GHs play decisive roles in digestion, in plant-herbivore, and host-pathogen interactions. GH activity is normally measured by the detection of a release from the substrate of products as sugars units, colored, or fluorescent groups. In most cases, the conditions for product release and detection differ, resulting in discontinuous assays. The current protocols result in using large amounts of reaction mixtures for the obtainment of time points in each experimental replica. These procedures restrain the analysis of biological materials with limited amounts of protein and, in the case of studies regarding small insects, implies in the pooling of samples from several individuals. In this respect, most studies do not assess the variability of GH activities across the population of individuals from the same species. The aim of this work is to approach this technical problem and have a deeper understanding of the variation of GH activities in insect populations, using as models the disease vectors *Rhodnius prolixus* (Hemiptera: Triatominae) and *Lutzomyia longipalpis* (Diptera: Phlebotominae). Here we standardized continuous assays using 4-methylumbelliferyl derived substrates for the detection of α-Glucosidase, β-Glucosidase, α-Mannosidase, N-acetyl-hexosaminidase, β-Galactosidase, and α-Fucosidase in the midgut of *R. prolixus* and *L. longipalpis* with results similar to the traditional discontinuous protocol. The continuous assays allowed us to measure GH activities using minimal sample amounts with a higher number of measurements, resulting in data that are more reliable and less time and reagent consumption. The continuous assay also allows the high-throughput screening of GH activities in small insect samples, which would be not applicable to the previous discontinuous protocol. We applied continuous GH measurements to 90 individual samples of *R. prolixus* anterior midgut homogenates using a high-throughput protocol. α-Glucosidase and α-Mannosidase activities showed the normal distribution in the population. β-Glucosidase, β-Galactosidase, N-acetyl-hexosaminidase, and

α-Fucosidase activities showed non-normal distributions. These results indicate that GHs fluorescent-based high-throughput assays apply to insect samples and that the frequency distribution of digestive activities should be considered in data analysis, especially if a small number of samples is used.

Keywords: glycoside hydrolase, glycosidase, fluorescent assay, insect digestion, *Rhodnius prolixus, Lutzomyia longipalpis*

INTRODUCTION

Glycosidases are key enzymes in central and intermediate metabolic pathways. They release terminal sugars from glycosides as oligo- or disaccharides, aryl, and alkyl glycosides. They may also have activity against a wide array of glycoconjugates, as polysaccharides, glycoproteins, or glycolipids (Terra and Ferreira, 2005).

The enzymatic detection of glycosidases is generally based on the secondary detection of released monosaccharides (as glucose). Alternatively, colorimetric or fluorogenic groups from synthetic substrates as *p*-nitrophenol or methyl umbelliferyl derived glycosides are measured after release (Gontijo et al., 1998; Jacobson et al., 2007; Moraes et al., 2012; Vale et al., 2012; Moreira et al., 2015). In most cases, this results in obligatory discontinuous protocols. The indirect detection of sugars may need further incubation with reagents as Glucose Oxidase (Baker, 1991) or conditions for the detection of absorbance/fluorescence of released groups may significantly differ from the conditions used for the reaction of glycosidase with its substrate (Baker and Woo, 1992).

Insect digestive glycosidases are essential enzymes for nutrition, regardless of taxonomic order or diet. These enzymes are essential for final digestion of sugars and have important roles in the interaction of insects with plants and pathogens (Terra and Ferreira, 2005). GH activities have been extensively studied in different insect orders, including agricultural pests and disease vectors. They are responsible for the breaking down of molecules of plant and animal sources (Dillon and el-Kordy, 1997; Gontijo et al., 1998; Jacobson and Schlein, 2001; Terra and Ferreira, 2005; Jacobson et al., 2007; Mury et al., 2009; Fonseca et al., 2010; Ghadamyari et al., 2010; Moraes et al., 2012; Vale et al., 2012; Sellami and Jamoussi, 2016).

Glycosidase activities vary according to the insect diet. Dillon and el-Kordy (1997) have shown differences in profiles of midgut α-glucosidase in sugar and blood fed *Phlebotomus langeroni* sand flies. According to these authors, the results may be related to the variation in the digestion of these two types of meals. Sugar is kept into the crop and released continuously to the midgut, and blood passes directly to this compartment, the place where the digestions of these diets are accomplished. Digestive glycosidases were also found in *L. longipalpis* larvae, and in this case, they are probably involved in degradation of bacterial and fungal cell walls (Moraes et al., 2012; Vale et al., 2012).

Furthermore, *R. prolixus* α-glucosidases have an important role in the detoxification of heme after a blood meal. Mury et al. (2009) have verified an increase in α-glucosidase activity and release of heme in the midgut of *R. prolixus* fed with hemin-enriched diet. Previous works (Ribeiro and Pereira, 1984; Ferreira et al., 1988a,b; Terra et al., 1988) have also shown the presence of several GHs on different nymphal stages induced by blood meal through discontinuous assays.

Discontinuous assays are routinely used for insect gut glycosidases mainly because their acidic optimum pH range is not compatible with the direct measurement of released *p*-nitrophenyl or methyl umbelliferyl groups, which give best signals in alkaline conditions. Due to the small amounts of protein that are obtained from most insect samples, coupled with the high amounts of enzyme needed for discontinuous enzyme assays, high-throughput screening is not available for the study, and characterization of many insect glycosidases. One consequence of these limitations is that GH activities in insects are commonly obtained from pools of samples from several individuals, and described as means ± SEM of few experimental measurements. Because of that, most studies do not assess the variability of GH activities across the population, and the described values do not account for subsets or extreme variations which can occur naturally in a biological species or as a consequence of different physiological or pathological individual conditions.

The aim of this work is to approach this technical problem and have a deeper understanding of the variation of GH activities in insect populations, using as models the disease vectors *R. prolixus* (Hemiptera: Triatominae) and *Lutzomyia longipalpis* (Diptera: Phlebotominae). We tested the use of continuous assays for the detection and measurement of acidic digestive glycosidases in these two insect species. Standardization of assay conditions allowed us to obtain continuous activity measurements similar to the discontinuous protocol. We applied those techniques for the measurement of six glycosidases (α-Glucosidase, β-Glucosidase, α-Mannosidase, N-acetyl-hexosaminidase, β-Galactosidase, and α-Fucosidase) in *R. prolixus* and α-Glucosidase in *L. longipalpis*. Additionally, we conducted high-throughput measurements of these enzymes in *R. prolixus*, which allowed us to study simultaneously the population distribution of all these activities using individual insect samples.

MATERIALS AND METHODS

Chemicals

The substrates used in this study were all from Sigma-Aldrich, namely 4-methylumbelliferyl α-D-glucopyranoside

(cat. No. M9766, for α-glucosidase); 4-methylumbelliferyl α-D-mannopyranoside (cat. No. M3657, for α-mannosidase); 4-methylumbelliferyl β-D-glucopyranoside (cat. No. M3633, for β-glucosidase); 4-methylumbelliferyl β-D-galactopyranoside (cat. No. M1633, for β-galactosidase); 4-methylumbelliferyl N-acetyl-β-glucosaminide (cat. No. M2133, for N-acetyl- β-hexosaminidase), and 4-methylumbelliferyl α-L-fucopyranoside (cat. No M8527, for α-fucosidase). Stock solutions were prepared in dimethyl sulfoxide (DMSO) at 10 mM, α-glucosidase was from *Saccharomyces cerevisiae* (cat. No. G5003). All reagents used in this work are analytical grade (Sigma-Aldrich).

Insects

We used adult males of *R. prolixus* (Hemiptera: Reduviidae), obtained from our insectary at FIOCRUZ (Rio de Janeiro), maintained at 28°C and 75% relative humidity. Insects were regularly fed with defibrinated rabbit blood (Azambuja and Garcia, 1997). For standardization of continuous assays, insects were dissected 5 days after blood feeding. For populational high-throughput measurements, adults were synchronized by collecting 5th instar nymphs that molted to adults in a 1–2 days time range. These adults were starved and kept without any source of food or water for 40 days until dissection and activity measurements.

L. longipalpis originally from Jacobina (Bahia state, Brazil) were kept as described in Moraes et al. (2014). We used adult males fed with sucrose 70% (w/v) for all the experiments described. We used recently emerged adults with ages ranging from 4 to 7 days, with sucrose meals offered *ad libitum* during 4–5 days.

Effect of pH on Fluorescence of 4-Methylumbelliferone

Fluorescence of 4-methylumbelliferone was measured in different 0.2 M buffers ranging from pH 2 to 11, namely Glycine (pH 2–3), Sodium Acetate (pH 4), MES (pH 5–7), Cacodylate (pH 8), Tris (pH 9), and Sodium Carbonate (pH 10–11), in the amount range from 62.4 to 624 picomoles.

Preparation of Samples

Insects were immobilized by placing them on the ice and dissected in cold 0.9% (w/v) NaCl. For *R. prolixus*, we separated the Anterior Midgut (AM) and homogenized the whole organ in 0.6 mL 0.9% NaCl. After that, contents and tissues were separated by centrifugation at 4°C, 5,000 X g, for 5 min (Houseman and Downe, 1981). For *L. longipalpis*, we separated the whole gut from the rest of body (head plus carcass) and homogenized them in 50 μL of citrate phosphate buffer 50 mM pH 6 containing Triton X-100 1% (v/v) for better solubilization.

Enzyme Assays

We mixed 25 μL of the sample at appropriate dilution with 74 μL of 0.2 M sodium acetate buffer pH 4.5 and 1 μL of MU-glycoside stock solution (item 2.1). Substrate hydrolysis was monitored at 30°C in a spectrofluorometer Spectramax Gemini XPS (Molecular Devices) on $\lambda_{Ex} = 355$ nm and $\lambda_{Em} = 460$ nm. Glycosidase activities were expressed in micro units (μU) per insect. One unit (U) correspond to the enzyme quantity that process 1 μmol of the substrate for a minute (IUBMB, 1992). In continuous assays, we measured fluorescence intensity for 60 min with reading each minute. In discontinuous assays, reaction mixtures were incubated at 30°C for 40 min, and 200 μL 0.75 M sodium carbonate buffer pH 10 were added to stop reaction every 10 min, and end point measurements were performed for all time points.

Statistical Analysis

Graphs and data analysis were performed with the software GraphPad Prism version 5.01 for Windows. Lilliefors-van Soest and Shapiro–Wilks normality tests were employed for verification of Gaussian distributions. Data for activities that showed a normal distribution in the population were submitted to parametric unpaired Student's *t*-test, and data for enzymes that did not present Gaussian distribution were submitted to non-parametric Mann–Whitney tests.

RESULTS

Initially, we investigated if fluorescence of the 4-methylumbelliferone (MU) group might be detected in acidic pHs, because glycosidase assays for *R. prolixus* and *L. longipalpis* are routinely done at pHs 4.5 and 6, respectively. For that, we performed standard curves for MU in pHs ranging from 2 to 11 (**Figure 1A**). We observed a detectable and significant response in all tested pHs, with smaller fluorescence values in acidic pHs (below 6) and maximum fluorescence in strongly alkaline conditions (above 9). The change in fluorescence response along pH 6–9 suggests that the MU group has a pKa around 8 (**Figure 1B**). Nevertheless, even in acidic conditions, the fluorescence response of MU was linear with concentration, which allowed us to measure the release of this group at the optimum conditions for the desired target glycosidases above.

To confirm if the release of the MU group by a glycosidase might be followed continuously in acidic conditions, we performed test assays with a commercial α-glucosidase (Sigma) in pH 6 (**Figure 2**). All assays performed showed proportionality of fluorescence with time with significant linear correlations (**Figure 2A**). Additionally, we observed proportionality of the measured activity with increasing concentrations of enzyme (**Figure 2B**).

We applied the continuous protocol for the detection of diverse glycosidases in *R. prolixus* anterior midgut tissue homogenates (**Figure 3**). We obtained linear assays for α-glucosidase (**Figure 3A**), α-mannosidase (**Figure 3B**), β-glucosidase (**Figure 3C**), β-galactosidase (**Figure 3D**), β-N-acetyl-hexosaminidase (**Figure 3E**), and α-fucosidase activities (**Figure 3F**). We observed similar results with assays for *L. longipalpis* α-glucosidase (data not shown). We observed in the assays for *R. prolixus* β-glucosidase, β-galactosidase, and α-fucosidase an initial lag phase in the first 5 min but, after that, linearity was restored (**Figures 3C,D,F**).

Similarly to what was observed with the commercial α-glucosidase, we tested if assays using intestinal homogenates of *R. prolixus* show proportionality between the measured activity

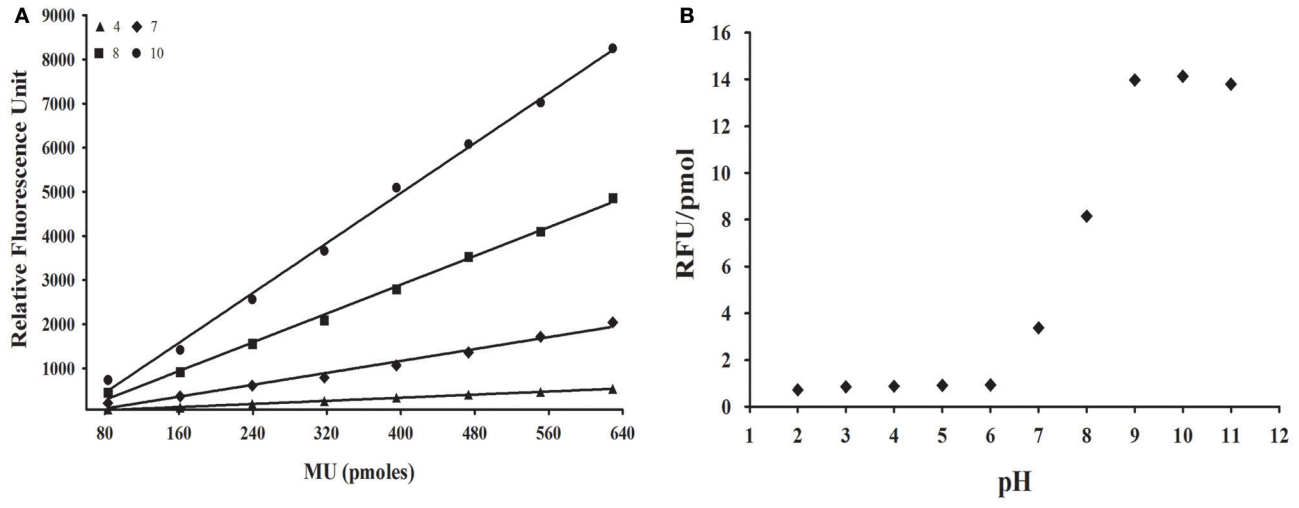


FIGURE 1 | Effect of pH on the fluorescence of methylumbelliferone. (A) Illustrative standard curves of methylumbelliferone vs. fluorescence in four different pH values. **(B)** Dependence of fluorescence response [slope of standard curves as depicted in **(A)**, in RFU/pmol] of methylumbelliferone in the pH range 2–11. MU, Methylumbelliferone.

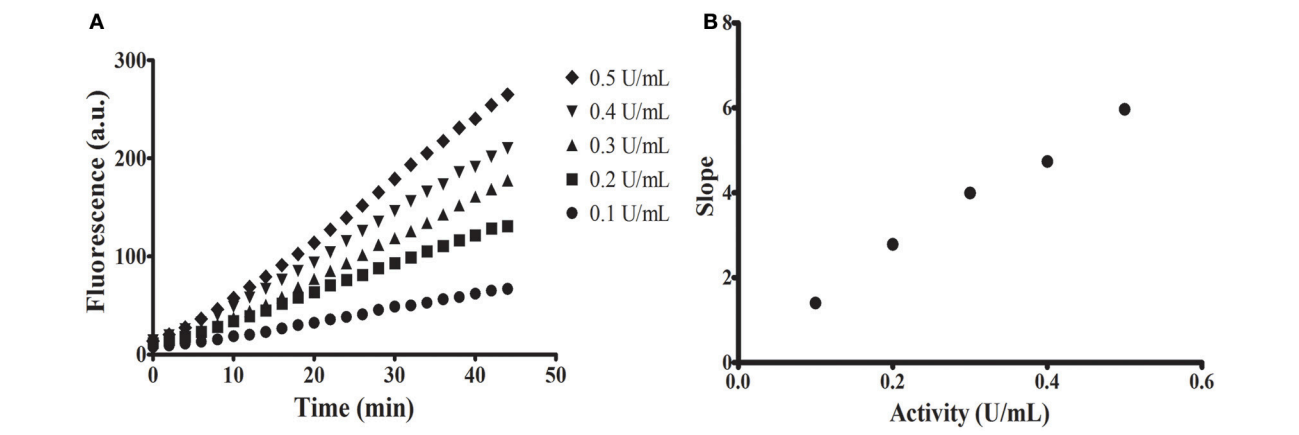


FIGURE 2 | Continuous assays using commercial α-glucosidase from Sigma. (A) Increase in fluorescence with time using different concentrations of enzyme. **(B)** Linear relationship between measured activity [slope of assays in **(A)**] and concentration of enzyme (in units as defined by the fabricant).

and the protein content of the assay (**Figure 4**). In the case of α-glucosidase, a strong deviation from linearity was observed in higher concentrations of enzyme (**Figure 4A**). However, we observed linearity for in a wide range of concentrations for α-mannosidase (**Figure 4B**), β-glucosidase (**Figure 4C**), β-galactosidase (**Figure 4D**), β-N-acetyl-glucosaminidase (**Figure 4E**), and α-fucosidase (**Figure 4F**). Similar results were obtained for *L. longipalpis* α-glucosidase (data not shown).

After that, we compared the results obtained with the continuous assay to measurements using the traditional discontinuous protocol (interruption by adding carbonate buffer pH 10 in excess). For doing that, we assayed the same samples of anterior midgut tissue homogenates of *Rhodnius prolixus* for the six glycosidases anteriorly standardized using both protocols. The detected activity amounts of α-glucosidase, α-mannosidase, β-glucosidase, β-galactosidase, β-N-acetyl-hexosaminidase and α-fucosidase activities obtained with the continuous assay were statistically similar to the activities detected using the discontinuous protocol (**Figure 5**). A similar result was obtained when comparing the α-glucosidase activities of *L. longipalpis* tissues measured with both techniques. Surprisingly, the continuous technique allowed us to measure confidently the activity of only one isolated sand fly midgut, which would be impossible to perform with the discontinuous technique (**Figure 6**). In fact, using the same experimental conditions, discontinuous assays with single fly samples failed to detect (data not shown) the α-glucosidase activity which is observed using samples from pooled insects (**Figure 6**).

After validation of the continuous protocol with our biological samples, we decided to apply the continuous technique to the

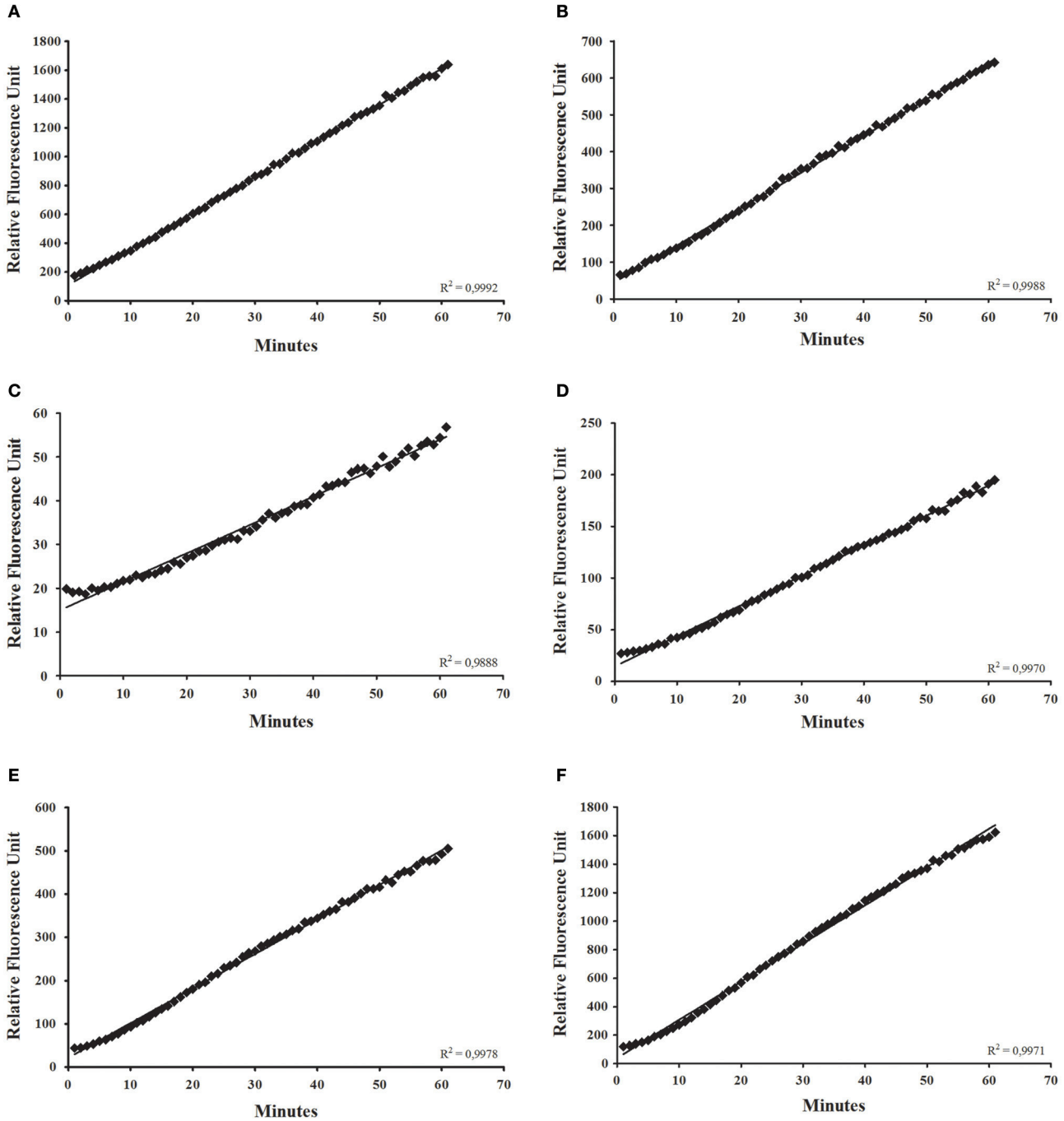


FIGURE 3 | Continuous glycosidase assays using the anterior midgut epithelium of *Rhodnius prolixus* as an enzyme source, showing the initial velocity measurements for (A) α-glucosidase; **(B)** α-mannosidase; **(C)** β-glucosidase; **(D)** β-galactosidase; **(E)** β-N-acetyl-glucosaminidase; **(F)** α-fucosidase.

high-throughput screening of glycosidase activities in individual insect samples. For that, we dissected 90 *R. prolixus* anterior midguts and submitted individual tissue homogenates to the continuous glycosidase assays in the conditions established above. To avoid variations in activity due to different ages or nutritional status, we used synchronized insects that were kept starving for 40 days. The starvation was also performed with the intention of simulating the poor nutritional status that is common in field-collected triatomines. That allowed us to measure the activities of the different six glycosidases in all individual samples in few hours. Results are summarized in **Figure 7**. We observed a huge variation of activity among

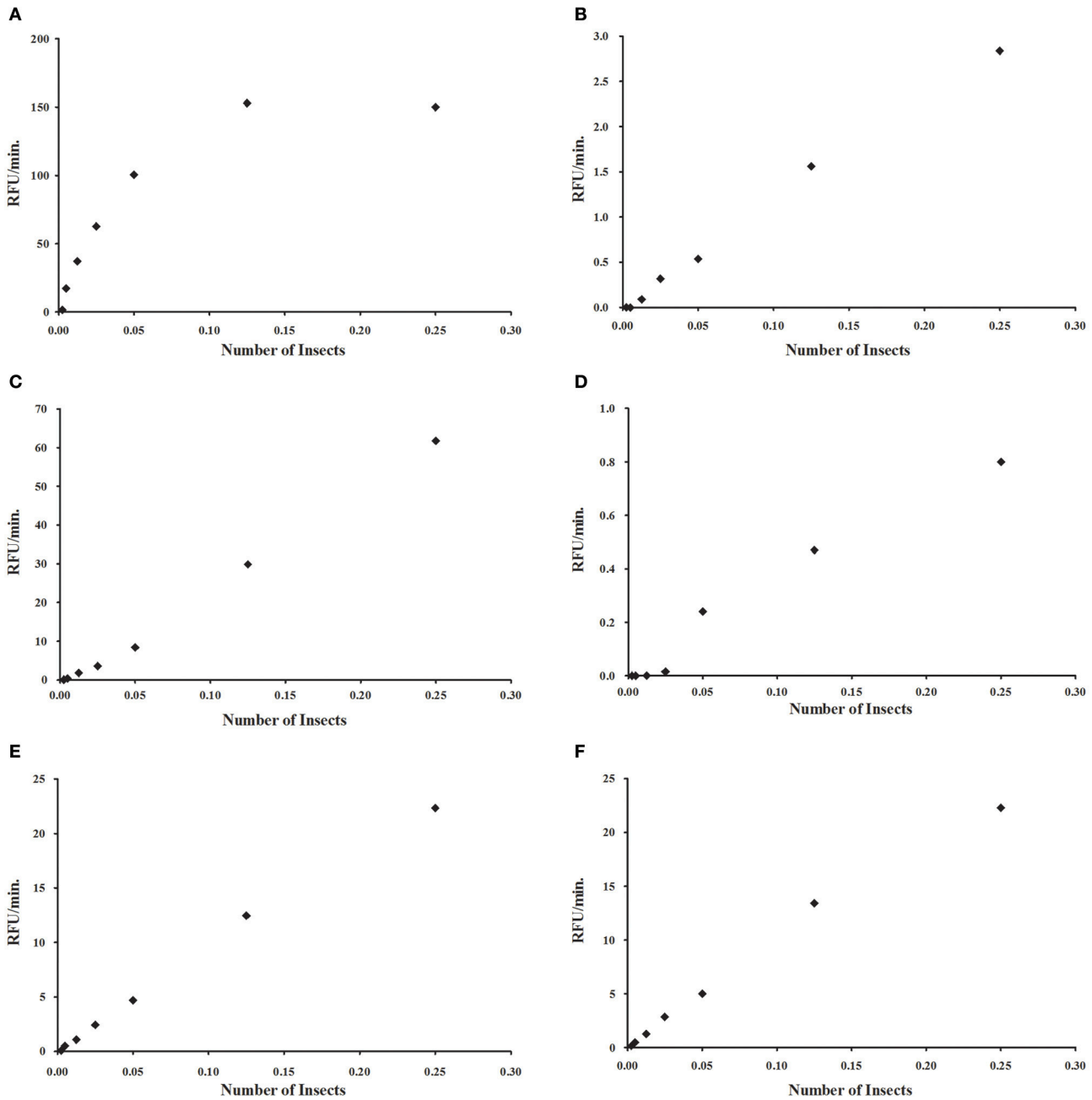


FIGURE 4 | Relationship between measured activity and protein concentration for different glycosidases, obtained by serial dilution of a sample prepared to include 0.25 insect equivalent per reaction mixture. (A) α-glucosidase; **(B)** α-mannosidase; **(C)** β-glucosidase; **(D)** β-galactosidase; **(E)** β-N-acetyl-glucosaminidase; **(F)** α-fucosidase. For **(A)**, due to the hyperbolic profile, samples were further diluted to obtain a linear relationship.

individual samples, with more than two orders of magnitude of variation in some cases (**Figure 7**). However, for all activities tested the distribution of activities in the population allowed us to observe a typical range of activities for each enzyme, as well as a more frequent activity range value, which is typical for each enzyme tested.

We decided to analyze the distribution of activities for each enzyme, to test if these activities have a normal distribution across the population. Frequency histograms are presented in **Figure 8**, and a resume of statistical data analysis is presented in Supplementary Table 1. Normality was observed only for α-glucosidase (**Figure 8A**) and α-mannosidase (**Figure 8B**). β-Glucosidase (**Figure 8C**), β-galactosidase (**Figure 8D**), β-N-acetyl-glucosaminidase (**Figure 8E**), and α-fucosidase (**Figure 8F**) presented non-normal behaviors, with right-skewed and heavy-tailed distributions (Supplementary Table 1).

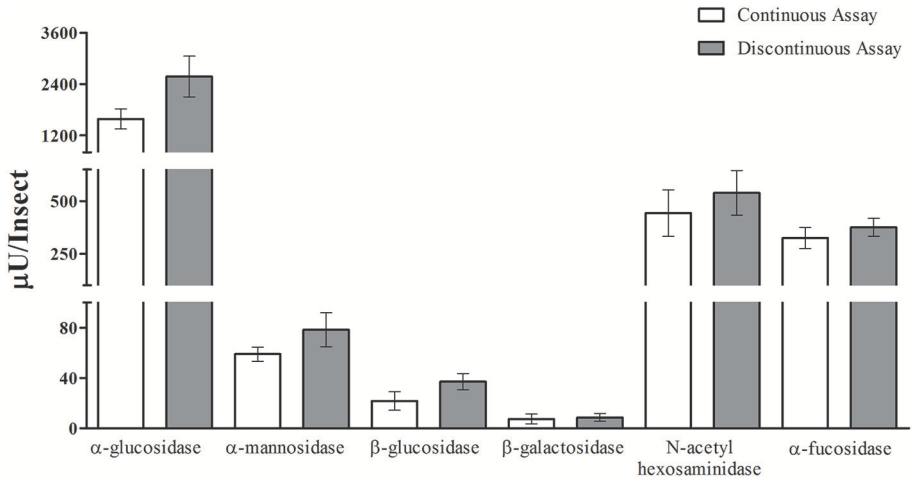


FIGURE 5 | Comparison of activity measurements obtained with continuous or discontinuous protocols. Activity of six different glycosidases was measured in five different samples from the *Rhodnius prolixus* anterior midgut tissues of six insects each. No statistically significant difference was observed between the results obtained with both techniques.

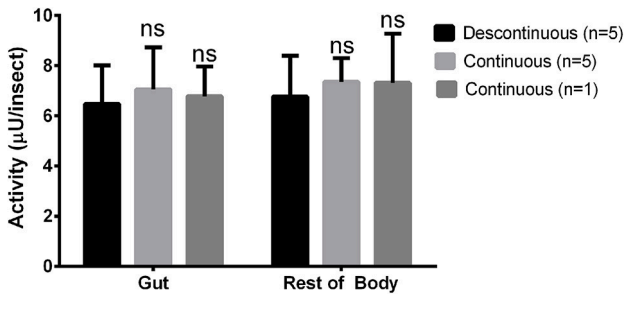


FIGURE 6 | Comparison of activity measurements obtained with continuous or discontinuous protocols in samples obtained from the sand fly *Lutzomyia longipalpis*. Activity of α-glucosidase was measured in different samples from guts or rest of bodies of one or five insects each. No statistically significant difference was observed between the results obtained with both techniques.

Another type of analysis that the high throughput assay allows is to observe the correlation between enzyme activities using a high number of individuals for the comparison. In general, all glycosidase activities showed a significant correlation to the other glycosidases assessed, based on Spearman correlation coefficients (Supplementary Table 2). Nevertheless, some activities showed strong correlations between each other, while other activity comparisons showed poor correlations, as can be observed case by case in scatter plots (**Figures 9A,B**, respectively) or by comparison of the Pearson correlation coefficients for each comparison (**Figure 9C**). In this respect, the best correlation observed was between α-fucosidase and β-glucosidase.

DISCUSSION

The aim of this work was to evaluate the possibility of following glycosidase activities by the continuous release of the methylumbelliferone group in acidic pHs. Methylumbelliferyl-based substrates have been used for a wide range of enzyme activities, but the dependence of the fluorescence on ionization of the group at high pHs has led to obligatory discontinuous assays for the majority of cases (Yang and Hamaguchi, 1980; Hoppe, 1983). Particularly, most insect glycosidases have optimal activities in acidic pHs (Gontijo et al., 1998; Terra and Ferreira, 2005; Jacobson et al., 2007; Cançado et al., 2008; Moraes et al., 2012; Tamaki et al., 2014; Moreira et al., 2015). In this way, the development of substrates with groups that fluoresce in acidic pHs would be greatly beneficial for the study of those enzymes. Despite that, we were able to detect methylumbelliferone even at pHs as acidic as 2 and standardize glycosidase assays at pH 4.5 for *R. prolixus* enzymes. It is important to notice that the application of an enzyme assay will always depend on the specificity of the target enzyme. For example, several insect glycosidases have no activity against synthetic substrates, being able to hydrolyze only their natural substrates, as cellobiose or lactose (Terra and Ferreira, 1994).

In general, linear assays were obtained for all enzymes tested, both in *R. prolixus* and *L. longipalpis* samples. Proportionality of the amount of product released with time is a key factor for the correct measurement of enzyme activities and kinetic parameters, as the main kinetic approaches rely on the measurement of initial velocities of reaction between enzyme and substrate. For discontinuous assays, this implies in the necessity of obtaining several data points and testing for linear correlation of the data. This result in using higher amounts of sample, reagents and time for confirmation of this basic assumption, and result in most cases in a higher occurrence of experimental errors and deviations. The continuous protocol used here to bypass all these difficulties, and in most cases, we were able to confidently measure the initial velocity of reaction in <10 min.

Another important factor in the measurement of enzyme activities is a proportionality of product release with protein

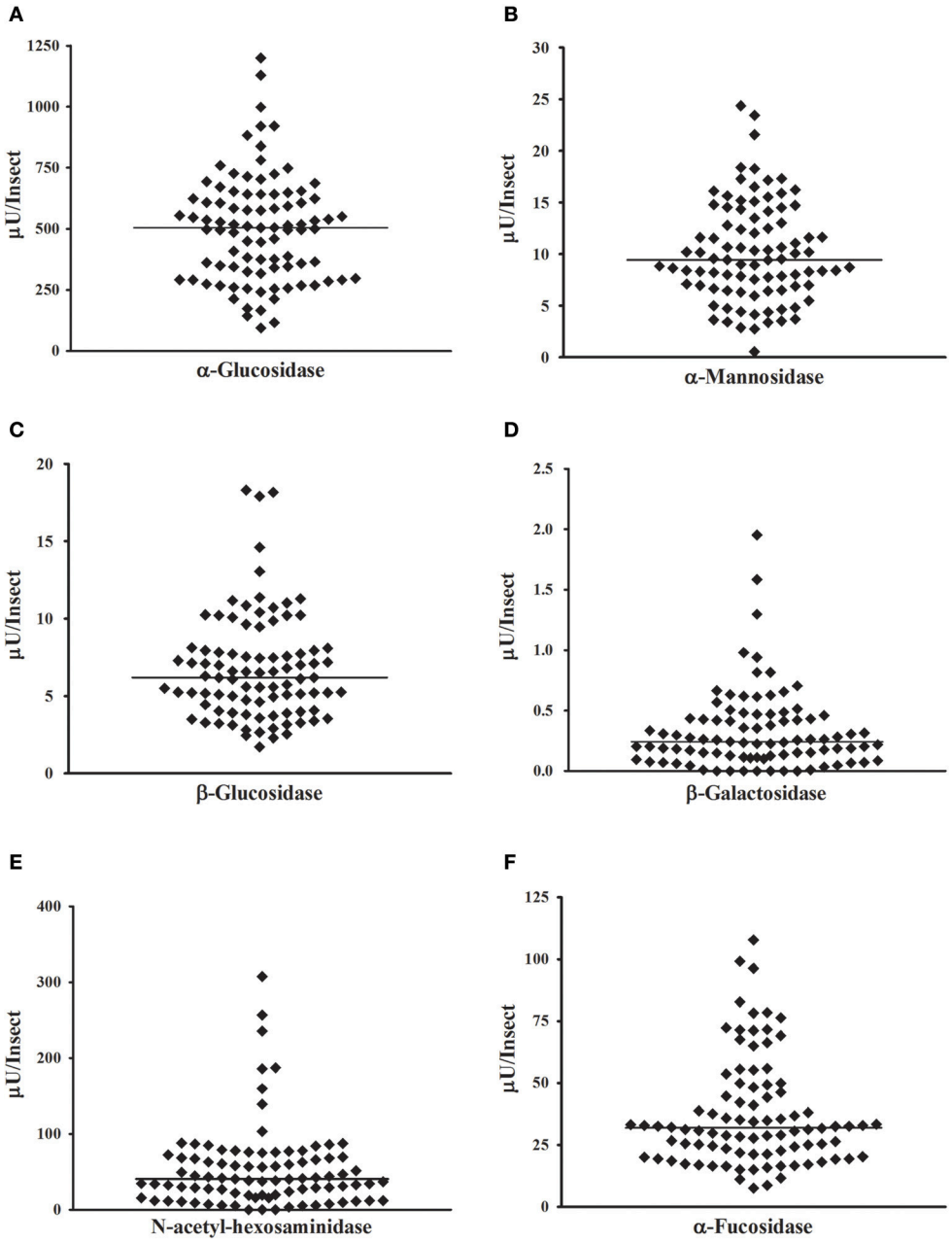


FIGURE 7 | High-throughput measurement of glycosidase activities in samples obtained from one insect each. Samples were prepared from 90 *Rhodnius prolixus* anterior midgut tissues, and all determinations were performed in the same day. **(A)** α-glucosidase; **(B)** α-mannosidase; **(C)** β-glucosidase; **(D)** β-galactosidase; **(E)** β-N-acetyl-glucosaminidase; **(F)** α-fucosidase.

concentration. In this paper, the appropriate assay conditions regarding sample concentration for all enzymes studied were established. For most cases, the observation of a linear correlation between product release and sample concentration was straightforward, with the sole exception of *R. prolixus* α-glucosidase. For this enzyme, inhibition was observed when using high concentrations of the sample. This result might be caused by the presence of inhibitors in the sample, degradation by proteases or by retroinhibition from products released at high concentrations of enzyme.

Physiological inhibitors derived from food molecules are common in insect gut homogenates and, additionally, some glycosidase inhibitors might have important roles in insect-plant and insect-microbial interactions (Terra and Ferreira, 1994). Digestion of bacteria seems to be a necessary step to cope with the paramount bacterial growth which occurs in the

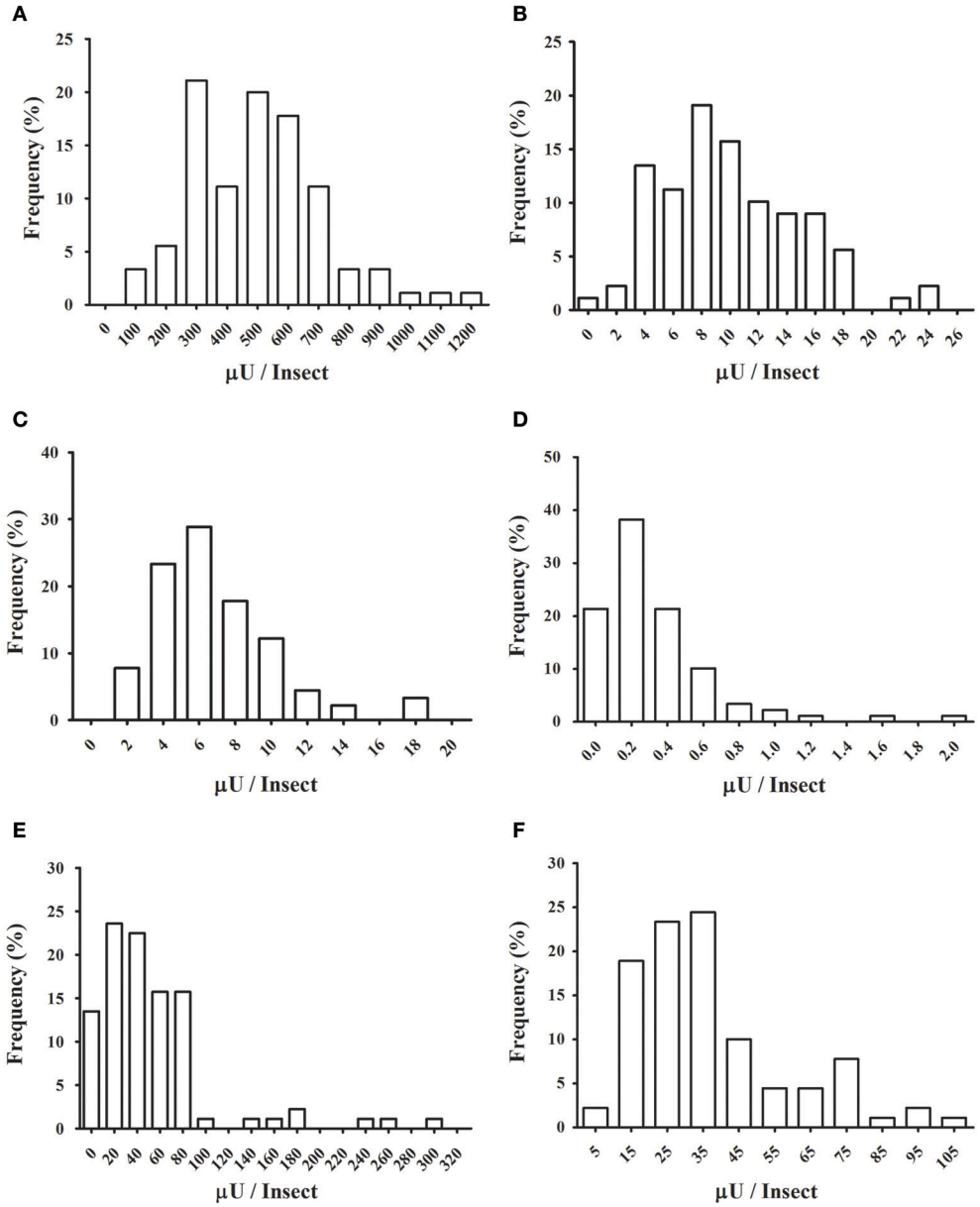


FIGURE 8 | Frequency distributions of glycosidase activities in the anterior midgut tissue across the population of *Rhodnius prolixus* assayed in Figure 7. (A) α-glucosidase; **(B)** α-mannosidase; **(C)** β-glucosidase; **(D)** β-galactosidase; **(E)** β-N-acetyl-glucosaminidase; **(F)** α-fucosidase.

anterior midgut after a blood meal. Therefore, it is possible that these glycosidases are being naturally exposed to alternative substrates of bacterial origin. Importantly, α-glucosidase in *R. prolixus* is a particular enzyme of the perimicrovillar membrane with the additional catalytic role of production of hemozoin, an important physiological step in the detoxification of the pro-oxidant heme molecule which is abundant in the blood meal (Silva et al., 2007; Mury et al., 2009). In this way, *R. prolixus* α-glucosidase might have the capacity of binding to different molecules beyond substrates with sugar moieties, as the heme group (Mury et al., 2009), and these molecules might behave as the putative inhibitors that are present in our homogenates.

Another possible source of non-linearity is the proteolytic attack by endogenous proteases. Proteolysis might seem unlikely in the case of samples from *R. prolixus* anterior midgut because this organ has been traditionally considered as a storage compartment and devoided of proteases. However, recent work described important transcription of several proteases in this tissue (Ribeiro et al., 2014). In this case, it was postulated that those enzymes might be secreted in that organ as zymogens, and it is possible that those enzymes are being activated in our assay

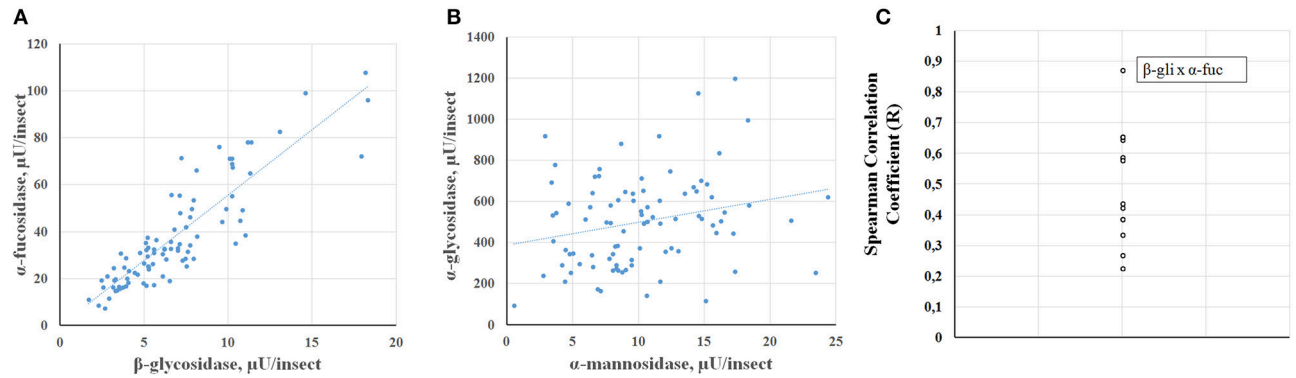


FIGURE 9 | Correlation of different glycosidase activities obtained in *Rhodnius prolixus* samples from Figure 7. (A) Scatter chart of α-fucosidase vs. β-glucosidase data ($R = 0.87$). **(B)** Scatter chart of α-glucosidase vs. α-mannosidase data ($R = 0.22$). **(C)** Spearman correlation coefficients for all comparisons (see Supplementary Table 2).

conditions. In this case, it should be considered that nonlinearity was observed only for α-glucosidase, and it seems unlikely that proteolysis would affect this enzyme preferentially and not the other glycosidases studied here.

Considering retroinhibitions, the two products generated in our α-glucosidase assay are glucose and methylumbelliferone. Inhibition by glucose has been described in a variety of α-glucosidases (Schomburg et al., 2002). No studies regarding the effects of methylumbelliferone on α-glucosidases were performed yet, but some insect glycosidases are inhibited by hydrophobic compounds, mainly aglycones derived from plant glycosides (Hong et al., 2013; Tan et al., 2013). All these possibilities should be considered in future studies of *R. prolixus* α-glucosidase. Despite that, the continuous assay allowed us to find an appropriate sample dilution quickly, by serial dilution and concomitant measurement of a wide range of concentrations of the sample. In this respect, the continuous assay might be valuable for the standardization of assay conditions and characterization of glycosidase activities from unknown samples.

Importantly, the continuous measurement resulted in activity values which are statistically similar to those obtained with the discontinuous protocol from the same samples. Activity values obtained using both techniques were indistinguishable for most of the enzymes studied, with the sole exception of *R. prolixus* α-glucosidase. In this case, values obtained with the continuous assays were somewhat lower than those obtained with the discontinuous assay, but the difference was not statistically significant. This minor artifact does not seem to be related to the specificity of the enzyme because *L. longipalpis* α-glucosidase did not show this behavior. Regardless, values obtained with the continuous assay seem to be reliable, and that encouraged us to use this technique for further studies of these enzyme activities.

Insect glycosidases have been extensively described, studied, and characterized, but the traditional outcome of glycosidase measurements consists in mean activities from samples obtained from pools of several individuals. Sample pooling is firstly a consequence of the limited amount of protein that is recovered from small insects like flies, ants, or beetles. Our approach allowed us for the first time to take a glance in the individual variability that might be hidden in a mean value from a pooled sample. In this case, a wide range of values might correspond to the activities of specific glycosidases, with a considerable presence of outliers, which might distort the mean that is obtained in a pooled sample.

Besides that, the determination of glycosidase activities in samples from individuals allowed us to study the populational distribution of these activities. All activities studied in that regard (six different glycosidases in *R. prolixus* anterior midgut) showed mono-modal frequency distributions, suggesting that they were recorded in a relatively homogenous population. We had the concern of analyzing insects that were synchronized and starved for a long period, to avoid interferences due to different ages or nutritional status, which can severely affect the fitness and physiological characteristics of *R. prolixus* (Díaz-Albiter et al., 2016). However, normal frequency distributions only for two of the activities studied (namely α-glucosidase and α-mannosidase) were observed. The other four activities (β-glucosidase, α-fucosidase, N-acetyl-hexosaminidase, and β-galactosidase) showed non-normal distributions with significant right skewness and kurtosis. Skewed distributions in enzymatic activities might arise from the fact that they are lower-bounded, as there are no negative values. However, a correlation between higher ranges of activity values and normality was not observed, which means that skewness in the non-normal distributions might have other reasons, probably of biological nature. It is important to observe that the non-normal curves presented significant high values for both right skewness and kurtosis. This fact results from the presence of insects with very high activity values. Those exceptional or unusual data points are easily observed in scatter plots like **Figure 7** or frequency histograms as **Figure 8**.

A straightforward consequence of the distributions observed is that for future studies, parametric tests for comparisons between samples and further analysis should be used only for the study of *R. prolixus* α-glucosidase and α-mannosidase. For

the other four *R. prolixus* glycosidases non-parametric statistics is recommended. It has been argued that for studies with a high number of samples (above 100) the distinction between parametric or non-parametric is not critical for most analysis (Ghasemi and Zahediasl, 2012). Nevertheless, this is not the case for the majority of insect glycosidase screenings, and the description of non-normal activity distributions might pave the way for more rigorous studies of insect glycosidases. Another consequence of non-normality is that for the distributions with high levels of skewness and heavy tails, the median better reflects the observations than the mean. However, as most observations in the literature use the mean as a reference, this could lead to misinterpretation of results. Non-normal distributions might be an additional reason for the large errors of the mean (∼20%) which are common in the descriptions of gut enzymatic activities in insects because they tend to have higher standard deviations than normal ones. It is important to consider that the distributions we have observed came from a limited data set ($N = 90$) and that it might be very interesting to corroborate them in the future with more extensive samplings.

An interesting feature that the individual measurement of glycosidase activities allowed us to assess is the possibility of a correlation between enzymes when looking at different insects across the population. This approach allowed the observation of a positive correlation between all the measured glycosidase activities. We expect a basal level of correlation in all comparisons because all activities would be dependent on the protein content of samples. Nevertheless, unusually high levels of correlation between some glycosidases were registered, which might reflect physiological functional groups of enzymes working in concert against the same substrate. Of special interest is the grouping between β-glucosidase and α-fucosidase. We need further studies to understand the relevance of these correlations, both at the molecular and physiological levels.

To our knowledge, this work is the first high-throughput screening of insect glycosidases performed at the individual level. It is important to notice that these enzymes are key factors in pathogen-vector and insect-plant interactions. Some glycosidases have been described as markers for insect resistance against plant defenses (Terra and Ferreira, 2005). We believe that the standardization of such high-throughput glycosidase assays will benefit not only the theoretical studies about the physiological role of these enzymes but also the monitoring of populations of sylvatic insects, agricultural pests, and disease vectors with a more powerful biochemical toolbox.

AUTHOR CONTRIBUTIONS

Conception and design of the work: GP, CM, FG. Obtainment of experimental data: GP, JP, SC. Data analysis: GP, JP, SC, FG. Writing and revision of the manuscript: GP, JP, SC, PA, EG, CM, FG.

ACKNOWLEDGMENTS

This research was supported by the Research Agencies FAPERJ, CNPq, CAPES, and FIOCRUZ (Brazil). CM received a post-doc grant from CAPES/FAPERJ. GP (MSc), JP (MSc), and SC (Ph.D.) are graduated students of Cellular and Molecular Biology Postgraduation Program of Oswaldo Cruz Institute. EG, PA, and FG are staff members of their respective departments. The authors thank Zildo Alves da Cruz and José Carlos Nascimento de Oliveira (IOC-FIOCRUZ) for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fphys.2017.00308/full#supplementary-material

REFERENCES

Azambuja, P., and Garcia, E. S. (1997). "Care and maintenance of triatomine colonies," in *The Molecular Biology of Insect Disease Vectors: A Methods Manual*, eds J. M. Crampton, C. B. Beard, and C. Louis (London: Chapman and Hall), 56–64.

Baker, W. L. (1991). Glucose oxidase reaction for estimation of glucose without horseradish peroxidase. Some microbiological and fermentation applications. *J. Inst. Brew.* 97, 457–462. doi: 10.1002/j.2050-0416.1991.tb01086.x

Baker, W. L., and Woo, S. M. (1992). Beta-glucosidases in the rice weevil, *Sitophilus oryzae* – purification, properties, and activity levels in wheat-feeding and legume-feeding strains. *Insect Biochem. Mol. Biol.* 22, 495–504. doi: 10.1016/0965-1748(92)90146-6

Cançado, F. C., Chimoy Effio, P., Terra, W. R., and Marana, S. R. (2008). Cloning, purification and comparative characterization of two digestive lysozymes from *Musca domestica* larvae. *Braz. J. Med. Biol. Res.* 41, 969–977. doi: 10.1590/S0100-879X2008001100005

Díaz-Albiter, H. M., Ferreira, T. N., Costa, S. G., Rivas, G. B., Gumiel, M., Cavalcante, D. R., et al. (2016). Everybody loves sugar: first report of plant feeding in triatomines. *Parasit. Vectors* 9:114. doi: 10.1186/s13071-016-1401-0

Dillon, R. J., and el-Kordy, E. (1997). Carbohydrate digestion in sandflies: alpha-glucosidase activity in the midgut of *Phlebotomus langeroni*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 116, 35–40. doi: 10.1016/S0305-0491(96)00195-2

Ferreira, C., Garcia, E. S., and Terra, W. R. (1988a). Properties of midgut hydrolases from nymphs and adults of the hematophagous bug *Rhodnius prolixus*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 90, 433–437. doi: 10.1016/0305-0491(88)90100-9

Ferreira, C., Ribeiro, A. F., Garcia, E. S., and Terra, W. R. (1988b). Digestive enzymes trapped between and associated with the double plasma membranes of *Rhodnius-prolixus* posterior midgut cells. *Insect Biochem.* 18, 521–530. doi: 10.1016/0020-1790(88)90003-0

Fonseca, F. V., Silva, J. R., Samuels, R. I., DaMatta, R. A., Terra, W. R., and Silva, C. P. (2010). Purification and partial characterization of a midgut membrane-bound alpha-glucosidase from Quesada gigas (Hemiptera: Cicadidae). *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 155, 20–25. doi: 10.1016/j.cbpb.2009.09.004

Ghadamyari, M., Hosseininaveh, V., and Sharifi, M. (2010). Partial biochemical characterization of alpha- and beta-glucosidases of lesser mulberry pyralid, Glyphodes pyloalis Walker (Lep.: Pyralidae). *C. R. Biol.* 333, 197–204. doi: 10.1016/j.crvi.2009.12.011

Ghasemi, A., and Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *Int. J. Endocrinol. Metab.* 10, 486–489. doi: 10.5812/ijem.3505

Gontijo, N. F., Almeida-Silva, S., Costa, F. F., Mares-Guia, M. L., Williams, P., and Melo, M. N. (1998). *Lutzomyia longipalpis*: pH in the gut, digestive glycosidases, and some speculations upon Leishmania development. *Exp. Parasitol.* 90, 212–219. doi: 10.1006/expr.1998.4336

Hong, H. C., Li, S. L., Zhang, X. Q., Ye, W. C., and Zhang, Q. W. (2013). Flavonoids with α-glucosidase inhibitory activities and their contents in the leaves of Morus atropurpurea. *Chin. Med.* 8, 1–7. doi: 10.1186/1749-8546-8-19

Hoppe, H.-G. (1983). Significance of exoenzymatic activities in the ecology of brackish water: measurements by means of methylumbelliferyl-substrates. *Mar. Ecol. Prog. Ser.* 11, 299–308. doi: 10.3354/meps011299

Houseman, J. G., and Downe, A. E. R. (1981). Exoproteinase activity in the posterior midgut of *Rhodnius prolixus* stal (Hemiptera: Reduviidae). *Insect Biochem.* 11, 579–582. doi: 10.1016/0020-1790(81)90026-3

International Union of Biochemistry and Molecular Biology (IUBMB) (1992). *Enzyme Nomenclature*. San Diego, CA: Academic Press. Available online at: http://www.chem.qmul.ac.uk/iubmb/

Jacobson, R. L., and Schlein, Y. (2001). Phlebotomus papatasi and Leishmania major parasites express alpha-amylase and alpha-glucosidase. *Acta Trop.* 78, 41–49. doi: 10.1016/S0001-706X(00)00164-9

Jacobson, R. L., Studentsky, L., and Schlein, Y. (2007). Glycolytic and chitinolytic activities of *Phlebotomus papatasi* (Diptera: Psychodidae) from diverse ecological habitats. *Folia Parasitol.* 54, 301–309. doi: 10.14411/fp.2007.039

Moraes, C. S., Diaz-Albiter, H. M., Faria, M. V., Sant'Anna, M. R., Dillon, R. J., and Genta, F. A. (2014). Expression pattern of glycoside hydrolase genes in *Lutzomyia longipalpis* reveals key enzymes involved in larval digestion. *Front. Physiol.* 5:276. doi: 10.3389/fphys.2014.00276

Moraes, C. S., Lucena, S. A., Moreira, B. H., Brazil, R. P., Gontijo, N. F., and Genta, F. A. (2012). Relationship between digestive enzymes and food habit of *Lutzomyia longipalpis* (Diptera: Psychodidae) larvae: characterization of carbohydrases and digestion of microorganisms. *J. Insect Physiol.* 58, 1136–1145. doi: 10.1016/j.jinsphys.2012.05.015

Moreira, N. R., Cardoso, C., Ribeiro, A. F., Ferreira, C., and Terra, W. R. (2015). Insect midgut a-mannosidase from family 38 and 47 with emphasis on those of *Tenebrio molitor*. *Insect Biochem. Mol. Biol.* 67, 94–104. doi: 10.1016/j.ibmb.2015.07.012

Mury, F. B., da Silva, J. R., Ferreira, L. S., dos Santos Ferreira, B., de Souza-Filho, G. A., de Souza-Neto, J. A., et al. (2009). Alpha-glucosidase promotes hemozoin formation in a blood-sucking bug: an evolutionary history. *PLoS ONE* 4:e6966. doi: 10.1371/journal.pone.0006966

Ribeiro, J. M., Genta, F. A., Sorgine, M. H., Logullo, R., Mesquita, R. D., Paiva-Silva, G. O., et al. (2014). An insight into the transcriptome of the digestive tract of the bloodsucking bug, *Rhodnius prolixus*. *PLoS Negl. Trop. Dis.* 8:e2594. doi: 10.1371/journal.pntd.0002594

Ribeiro, J. M. C., and Pereira, M. E. A. (1984). Midgut glycosidases of *Rhodnius prolixus*. *Insect Biochem.* 14, 103–108. doi: 10.1016/0020-1790(84)90089-1

Schomburg, I., Chang, A., and Schomburg, D. (2002). BRENDA, enzyme data and metabolic information. *Nucleic Acids Res.* 30, 47–49. doi: 10.1093/nar/30.1.47

Sellami, S., and Jamoussi, K. (2016). Investigation of larvae digestive β-glucosidase and proteases of the tomato pest Tuta absoluta for inhibiting the insect development. *Bull. Entomol. Res.* 106, 406–414. doi: 10.1017/s0007485316000079

Silva, J. R., Mury, F. B., Oliveira, M. F., Oliveira, P. L., Silva, C. P., and Dansa-Petretski, M. (2007). Perimicrovillar membranes promote hemozoin formation into *Rhodnius prolixus* midgut. *Insect Biochem. Mol. Biol.* 37, 523–531. doi: 10.1016/j.ibmb.2007.01.001

Tamaki, F. K., Pimentel, A. C., Dias, A. B., Cardoso, C., Ribeiro, A. F., Ferreira, C., et al. (2014). Physiology of digestion and the molecular characterization of the major digestive enzymes from Periplaneta americana. *J. Insect Physiol.* 70, 22–35. doi: 10.1016/j.jinsphys.2014.08.007

Tan, C., Wang, Q., Luo, C., Chen, S., Li, Q., and Li, P. (2013). Yeast α-glucosidase inhibitory phenolic compounds isolated from Gynura medica leaf. *Int. J. Mol. Sci.* 14, 2551–2558. doi: 10.3390/ijms14022551

Terra, W. R., and Ferreira, C. (1994). Insect digestive enzymes: properties, compartmentalization and function. *Comp. Biochem. Physiol.* 109B, 1–62. doi: 10.1016/0305-0491(94)90141-4

Terra, W. R., and Ferreira, C. (2005). "Biochemistry of digestion," in *Comprehensive Molecular Insect Science, Vol.4: Biochemistry and Molecular Biology*, eds L. I. Gilbert, K. Iatrou, and S. S. Gill (Amsterdam: Elsevier), 171–224.

Terra, W. R., Ferreira, C., and Garcia, E. S. (1988). Origin, distribution, properties and functions of the major *Rhodnius prolixus* midgut hydrolases. *Insect Biochem.* 18, 423–434. doi: 10.1016/0020-1790(88)90058-3

Vale, V. F., Moreira, B. H., Moraes, C. S., Pereira, M. H., Genta, F. A., and Gontijo, N. F. (2012). Carbohydrate digestion in *Lutzomyia longipalpis* larvae (Diptera - Psychodidae). *J. Insect Physiol.* 58, 1314–1324. doi: 10.1016/j.jinsphys.2012.07.005

Yang, Y., and Hamaguchi, K. (1980). Hydrolysis of 4-methylumbelliferyl N-acetyl-chitotrioside catalyzed by hen and turkey lysozymes. pH dependence of the kinetics constants. *J. Biochem.* 87, 1003–1014.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer MP declared a past co-authorship with several of the authors SC, PA, and EG to the handling Editor, who ensured that the process met the standards of a fair and objective review.

1 Additional Files

Supplementary Table 1. Statistical parameters of activity distributions for *Rhodnius prolixus* glycosidases

Glycosidase	Relative Skewness	Relative Kurtosis
α-fucosidase	1.21	1.00
α-glucosidase	0.63	0.60
β-glucosidase	1.33	2.26
β-galactosidase	2.34	7.62
α-mannosidase	0.60	0.21
N-acetyl- β-hexosaminidase	2.45	7.33

Supplementary Table 2. Coefficients and *p*-values for Spearman correlation tests between *Rhodnius prolixus* glycosidase activities (white and grey diagonals, respectively). NAH – N-acetyl-β-hexosaminidase.

Glycosidase	α -fucosidase	α-glucosidase	β-glucosidase	β-galactosidase	α-mannosidase	NAH1
α-fucosidase		< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001
α-glucosidase	0.6538	-	< 0.0001	< 0.0001	0.0338	0.0002
β-glucosidase	0.8724	0.6546	-	< 0.0001	0.0002	< 0.0001
β-galactosidase	0.6441	0.4361	0.6547	-	0.0108	< 0.0001
α-mannosidase	0.4231	0.2253	0.3866	0.2689	-	0.0014
NAH	0.5843	0.3855	0.5879	0.578	0.3345	-

TRABALHO 2

Characterization of α-glucosidases from *Lutzomyia longipalpis* reveals independent hydrolysis systems for plant or blood sugars

Samara G. da Costa, Paul Bates, Rod Dillon and Fernando Ariel Genta

Submetido ao periódico Frontiers in Physiology

Characterization of α-glucosidases from Lutzomyia longipalpis reveals independent hydrolysis systems for plant or blood sugars

Samara G. Da Costa[1], Paul Bates[2], Rod Dillon[2], Fernando A. Genta[1*]

[1]Fundação Oswaldo Cruz (Fiocruz), Brazil, [2]Lancaster University, United Kingdom

Submitted to Journal:
Frontiers in Physiology

Specialty Section:
Invertebrate Physiology

Article type:
Original Research Article

Manuscript ID:
428977

Received on:
04 Oct 2018

Frontiers website link:
www.frontiersin.org

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Author contribution statement

Conception and design of the work: SGC, PB, RJD, and FG. Obtaining of experimental data: SGC. Data analysis: SGC and FG. Writing and revision of the manuscript: SGC, PB, RJD, and FG. All authors read and approved the final version.

Keywords

Lutzomyia longipalpis, Digestion, sugar, Sucrase, α-glucosidase, Glucoside Hidrolase

Abstract

Word count: 281

Lutzomyia longipalpis is the main vector of Leishmania infantum and exploits different food sources during development. Adults have a diet rich in sugars, and females also feed on blood. The sugar diet is essential for maintaining longevity, infection and Leishmaniasis transmission. Carbohydrases, including α-glucosidases, are the main enzymes involved in the digestion of sugars. In this context, we studied the modulation of α-glucosidase activities in different feeding conditions and compartments of Lutzomyia longipalpis females, in order to characterize in detail their roles in the physiology of this insect. All tissues showed activity against MUαGlu and sucrose, with highest activities in the midgut and crop. Activity was 1000 times higher on sucrose than on MUαGlu. Basal activities were observed in non-fed insects; blood feeding induced activity in the midgut contents, and sugar feeding modulated activity in midgut tissues. α-glucosidase activity changed after female exposure to different sugar concentrations or moieties. α-glucosidases from different tissues showed different biochemical properties, with an optimum pH around 7.0 - 8.0 and KM between 0.37 - 4.7 mM, when MUαGlu was used as substrate. Using sucrose as substrate, the optimum pH was around 6.0, and KM ranges between 11 - 800 mM. Enzymes from the crop and midgut tissues showed inhibition in high substrate concentrations (sucrose), with KI ranging from 39 to 400 mM, which explains the high KM values found. Chromatographic profiles confirmed that different α-glucosidases are been produced in L. longipalpis in different physiological contexts, with the distinction of at least four α-glucosidases. The results suggest that some of these enzymes are involved in different metabolic processes, like digestion of plant sugars, digestion of blood glycoproteins or glycolipids, and mobilization of energetic storages during starvation.

Funding statement

This research was supported by the Research Agencies FAPERJ, CNPq, CAPES, and FIOCRUZ (Brazil).

Ethics statements

(Authors are required to state the ethical considerations of their study in the manuscript, including for cases where the study was exempt from ethical approval procedures)

Does the study presented in the manuscript involve human or animal subjects: Yes

Please provide the complete ethics statement for your manuscript. Note that the statement will be directly added to the manuscript file for peer-review, and should include the following information:

- Full name of the ethics committee that approved the study
- Consent procedure used for human participants or for animal owners
- Any additional considerations of the study in cases where vulnerable populations were involved, for example minors, persons with disabilities or endangered animal species

As per the Frontiers authors guidelines, you are required to use the following format for statements involving human subjects: This study was carried out in accordance with the recommendations of [name of guidelines], [name of committee]. The protocol was approved by the [name of committee]. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

For statements involving animal subjects, please use:
This study was carried out in accordance with the recommendations of 'name of guidelines, name of committee'. The protocol was approved by the 'name of committee'.

If the study was exempt from one or more of the above requirements, please provide a statement with the reason for the exemption(s).
Ensure that your statement is phrased in a complete way, with clear and concise sentences.

This study was carried out in accordance with the principles of the Brazilian Society of Science in Laboratory Animals (SBCAL) and recommendations of Animal Research Ethics Board of Oswaldo Cruz Institute (CEUA L-029/2016). The protocol was approved by the Animal Research Ethics Board of Oswaldo Cruz Institute.

Characterization of α-glucosidases from *Lutzomyia longipalpis* reveals independent hydrolysis systems for plant or blood sugars

Samara G. da Costa[1], Paul Bates[2], Rod Dillon[2,3] and Fernando Ariel Genta[1,3,*]

[1]Laboratory of Insect Biochemistry and Physiology, Oswaldo Cruz Institute, FIOCRUZ, Rio de Janeiro, RJ, Brazil

[2]Faculty of Health and Medicine, Division of Biomedical and Life Sciences, Lancaster University, Lancaster, United Kingdom.

[3]National Institute of Science and Technology for Molecular Entomology, Rio de Janeiro, RJ, Brazil

***Correspondence:**
Fernando Ariel Genta
genta@ioc.fiocruz.br or gentafernando@gmail.com

Keywords: *Lutzomyia longipalpis*, digestion, sugar, sucrase, α-glucosidase, glucoside hydrolase

Abstract

Lutzomyia longipalpis is the main vector of *Leishmania infantum* and exploits different food sources during development. Adults have a diet rich in sugars, and females also feed on blood. The sugar diet is essential for maintaining longevity, infection and Leishmaniasis transmission. Carbohydrases, including α-glucosidases, are the main enzymes involved in the digestion of sugars. In this context, we studied the modulation of α-glucosidase activities in different feeding conditions and compartments of *Lutzomyia longipalpis* females, in order to characterize in detail their roles in the physiology of this insect. All tissues showed activity against MUαGlu and sucrose, with highest activities in the midgut and crop. Activity was 1000 times higher on sucrose than on MUαGlu. Basal activities were observed in non-fed insects; blood feeding induced activity in the midgut contents, and sugar feeding modulated activity in midgut tissues. α-glucosidase activity changed after female exposure to different sugar concentrations or moieties. α-glucosidases from different tissues showed different biochemical properties, with an optimum pH around 7.0 - 8.0 and K_M between 0.37 - 4.7 mM, when MUαGlu was used as substrate. Using sucrose as substrate, the optimum pH was around 6.0, and K_M ranges between 11 - 800 mM. Enzymes from the crop and midgut tissues showed inhibition in high substrate concentrations (sucrose), with K_I ranging from 39 to 400 mM, which explains the high K_M values found. Chromatographic profiles confirmed that different α-glucosidases are been produced in *L. longipalpis* in different physiological contexts, with the distinction of at least four α-glucosidases. The results suggest that some of these enzymes are involved in different metabolic processes, like digestion of plant sugars, digestion of blood glycoproteins or glycolipids, and mobilization of energetic storages during starvation.

Introduction

There are approximately 600 species of phlebotomines (Diptera: Psychodidae: Phlebotominae), most belonging to the genus *Phlebotomus* in the old world and *Lutzomyia* in the new world. About 98 phlebotomine species are considered of medical importance, being vectors of diseases such as visceral and cutaneous leishmaniasis, bartonellosis, besides transmitting other trypanosomatids and arboviruses (Sherlock *et al.*, 2003; Oryan and Akbari, 2016; WHO 2017).

Leishmaniasis are caused by more than 20 species of *Leishmania* (Akhoundi et al. 2016). The resulting type of infection depends on the infecting *Leishmania* species. The most common form of the disease is cutaneous leishmaniasis. In Brazil *Leishmania braziliensis* is the prevalent parasite in man causing cutaneous infections (Vianna 1911, 1914, Gontijo and Carvalho 2003) and *Leishmania Infantum* is the etiologic agent of visceral leishmaniasis, the most severe form (Penna 1934, Lainson and Shaw, 1992). The several *Leishmania* species are transmitted by different phlebotomines. The main vector of *Leishmania infantum*, the etiologic agent of visceral leishmaniasis in the Americas, is the sand fly *Lutzomyia longipalpis* (Soares and Turco, 2003).

Phlebotomines explore different food sources in their larval or adult phase. Larvae are detritivore animals and grow in decaying materials, mainly plant tissues or animal feces (Moraes *et al.*, 2012 and Moraes *et al.*, 2014). In adults, sugar meals are essential for their energy requirements, and they feed on plant tissues, nectar of flowers and secretions produced by aphids and coccids (Killick-Kendrick *et al.*, 1987; Moore *et al.*, 1987; MacVicker *et al.*, 1990; Cameron *et al.*, 1995; Junilla *et al.*, 2011). Females also feed on blood for egg maturation (Schlein, 1986). In this respect, these insects have a diet rich in sugars such as sucrose, maltose, trehalose, and melezitose, obtained from sugar meals and glycolipids, and glycoproteins, obtained from blood meals, which reinforces the importance the study of carbohydrases in their digestion process. The adult feeding habit of *L. longipalpis* has been the focus of several studies, since the parasite *Leishmania* is ingested in its amastigote form and transmitted during the blood supply of adult females in the metacyclic form. Besides this, carbohydrases were described as important during the development of the *Leishmania* parasite inside the phlebotomine gut (Pimenta et al., 1992; Dillon and El-Kordy, 1997). The sugars ingested by sand flies are hydrolyzed by glycosidases in small sugars units', which can be absorbed by the parasites and assist their development (Gontijo *et al.*, 1996).

α-glucosidases are an important group of enzymes that are specialized in sugar digestion. The starch found in leaves is hydrolyzed by α-amylase to maltose, which is then cleaved to glucose by α-glucosidases (Dixon and Webb, 1979). In the same way, these enzymes are involved in the hydrolysis of sucrose and maltose acquired from the sap. The energy obtained by the hydrolysis of these sugars is a crucial factor for insect development and maintenance of a sufficient lifetime for infection with *Leishmania* and its subsequent transmission. This enzyme may also participate in the final digestion of blood glycoproteins and glycolipids.

In general, α-glucosidases (EC 3.2.1.20) are typical exo-type hydrolases and catalyze the hydrolysis of α-glucosidic linkages releasing D-glucose residues from the non-reducing end of α aryl glycosides, disaccharides, and oligosaccharides with varying efficiency (Chiba et al., 1997). Based on substrate specificity, α-glucosidases can be classified into three types and also can be divided in families I and II based on the primary structure. Insects α-glucosidases are classified in Family I and they have no similarity with α-glucosidases from mammals, plants and fungi, classified in family II (Chiba et al., 1997). α-glucosidases from family I have four conserved regions, including the active site. These conserved regions are also found in α-amylases, but they do not share any sequence similarity.

Glycoside Hydrolases (GH), the enzyme group to which α-glucosidases belong, are also classified based on the secondary and tertiary structure in more than 100 different families. α-glucosidases were already described as members of families 4, 13, 31, 63, 97, or 122 of the glycoside hydrolases (GHF). In insects, these enzymes are found in families 13 and 31 (Cantarel et al., 2009).

The detailed description of the specificity, function, and structure of the α-glucosidases of *L. longipalpis* may provide the basis for the development of new strategies to control these insects and block the transmission of the disease, due to the importance of this enzyme in the digestion process in these insects and in the relationship between host and parasite.

A few works have described α-glucosidases activities in Phlebotomine sand flies (Samie *et al.*, 1990; Dillon and El Kordy 1997; Jacobson *et al.*, 2001 and Jacobson *et al.*, 2007) and an α-glucosidase in *L. longipalpis* was previously described (Gontijo *et al.*, 1998). In this work, we studied the modulation of α-glucosidase activities in different feeding conditions and compartments of *L. longipalpis* females, in order to characterize in greater detail its importance for the physiology of this insect. The α-glucosidase assays

were adjusted to the characteristics of sand fly enzymes, allowing us to detect activities that were not previously described. At least four different α-glucosidases with distinct biochemical properties were found, as constitutive or induced, in different feeding conditions. They appear to be involved in different metabolic processes, like digestion of plant sugars, digestion of blood glycoproteins or glycolipids, and mobilization of energetic storages during starvation.

Material and methods

Abbreviations

MUαGlu: methylumbelliferyl-α-glucopyranoside; **pNPαGlu:** 4-Nitrophenyl α-D-glucopyranoside; **PMSF:** phenylmethylsulfonyl fluoride, **E64:** (2S,3S)-3-(*N*-{(S)-1-[*N*-(4-guanidinobutyl)carbamoyl]3-methylbutyl}carbamoyl)oxirane-2-carboxylic acid. **EPPS:** 4-(2-hydroxyethyl)-1 piperazinepropanesulfonic acid; **AMPSO:** N-(1,1-dimethyl-2-hydroxyethyl)-3-amino-2-hydroxypropanesulfonic acid; **CAPS:** 3-(cyclohexylamino)-1-propanesulfonic acid; **MOPS:** 4 morpholinepropanesulfonic acid. **BCA:** bicinchoninic acid.

Chemicals

The reagents sucrose, fructose, maltose, glucose and methylumbelliferyl-α-glucopyranoside (MUαGlu) were purchased from Sigma-Aldrich Company (St. Louis, Missouri, USA). Pierce™ BCA Protein Assay Kit was obtained from Thermo Fisher Scientific (Waltham, Massachusetts, USA). Glucose oxidase (GOD) was acquired from Bioclin (Minas Gerais, Brazil). Other reagents used in this work were analytical grade.

Insects

All experiments were performed using insectary-reared *L. longipalpis* Jacobina strain (from Jacobina, Bahia, Brazil). Insects were kept under standard laboratory conditions under a temperature of 26 °C (± 2) and a relative humidity of ≥ 80 % inside the rearing cages (Moraes *et al.*, 2012, 2014, 2018). Adults sand flies were fed *ad libitum* with 70 % (w/v) autoclaved sucrose solution in cotton wool (unless stated differently in experiments). For oviposition females were fed on hamsters (*Mesocricetus auratus*) anesthetized with xylazine (10 mg/kg) plus ketamine (200 mg/kg). Engorged females were transferred to rearing containers with a piece of cotton wool soaked in sugar solution on it. The eggs were separated from dead females' bodies after oviposition. All procedures involving animals were approved by the Animal Research Ethics Board of

Oswaldo Cruz Institute (CEUA L-029/2016) and all experimental procedures were conducted following the biosecurity and institutional safety rules (CDC 2013).

Recently emerged females (0-3 hours) were fed with water for 5 days, or fed with 1.2 M sucrose for 10 days, or fed with sucrose in different concentrations (0.3 M, 0.6 M and 1.2 M) for 2 days, or fed with mono or disaccharides (glucose, fructose, sucrose and maltose) 1.2 M for 2 days. In some experiments, females were transferred from rearing containers to cages containing a Petri dish covered with parafilm and small drops of the offered sugar mixed with 50 % (v/v) of blue food dye (Arcolor, São Paulo, Brazil).

Females fed with blood for experiments had no previous contact with sugar. Recently emerged females (0-3 hours) were maintained for 3 days with cotton wool soaked in water before feeding on hamsters. After blood feeding, they were maintained with a piece of cotton wool soaked in water for 4 days.

Quantification of the volume ingested

When different concentrations of sucrose or different sugars were offered to females, a blue food dye (Arcolor, São Paulo, Brazil) 50 % (v/v) was mixed with sugar to determine the volume ingested by females (Ferreira et al., 2018). Briefly, two days after feeding on sugars, females were individually dissected in saline phosphate buffer (PBS) and the whole gut was homogenized in 2-10 µL of water for proper dilution. Samples were centrifuged at 10000 x g for 5 min. An aliquot of 1 µL from the supernatant was used to read the absorbance at 630 nm using a NanoDrop® (Thermo Fischer scientific). The volume ingested was calculated from a standard curve assembled and read under the same conditions as the samples.

Sample preparation

Females were dissected on PBS and the midgut was individually homogenized in 20 µL of 50 mM citrate-phosphate buffer pH 6.0 and centrifuged at 20.000 g for 10 min at 4 °C. The supernatant was saved and the pellet was washed again with 20 µL of 50 mM citrate-phosphate buffer pH 6, being recentrifuged at 20000 g for 10 min at 4 °C. Supernatants were pooled resulting and named ¨midgut contents¨ fraction. The pellet was suspended in 40 µL of 50 mM citrate-phosphate buffer pH 6 containing 1 % (v/v) Triton X-100. After 2 h of incubation on ice, it was centrifuged at 20.000 g for 10 min at 4 °C. This supernatant was named ¨midgut tissue¨ fraction. When head, carcass, crop or hindgut were used for α-glucosidase assays they were individually homogenized in 40 µL of 50

mM citrate-phosphate buffer pH 6 containing 1 % (v/v) Triton X-100 and incubated on ice for 2 hours. These samples were centrifuged at 20.000 g for 10 min at 4 °C and the supernatant was used as a source of enzymes. For all samples the following protease inhibitors were mixed (final concentrations): PMSF 4 mM, E64 0.02 mM, and pepstatin 0.02 mM. When samples were used for studies of the effect of pH, substrate concentration or chromatography the samples consisted of pools obtained from 100 females.

α-glucosidase assay and protein determination

α-glucosidase activity was measured in individual freshly prepared samples, by continuous assay using MUαGlu as substrate and discontinuous assay using sucrose as substrate. Briefly, samples were incubated for 1 h at 30 ^{0}C using 200 mM AMPSO pH 8.0 buffer and 0.1 mM MUαGlu (final concentration) (Sigma cat. no. M9766), the final volume reaction was 100 µL. The amount of 4-methylumbelliferone released was determined continuously at 355 nm excitation and 460 nm emission (Profeta et al., 2017) in a 96-well SpectraMax Gemini XPS Microplate Reader (Molecular Devices). The amount of product was calculated from a standard curve of 4-methylumbelliferone assembled and read under the same conditions as samples.

In assays with sucrose as substrate, samples were incubated for 30 min at 30 ^{0}C using 200 mM citrate-phosphate pH 6.0 buffer and 200 mM sucrose (final concentration), the final volume reaction was 50 µL. Reactions were interrupted at different time intervals (0, 10, 20 and 30 min) by incubating the mixture at 99 ^{0}C for 5 min. The amount of glucose released was determined according to the specifications provided by the manufacturer of the glucose mono-reagent kit (Bioclin, Brazil), based on the method described by Raabo and Terkildsen, 1960. Adjustments were done for the small volumes used. Briefly, 200 µL of glucose oxidase reagent was added to 50 µL reactions and incubated for 15 min at 37 ^{0}C. Plates were read at 505 nm in a 96-well SpectraMax 190 Microplate Reader (Molecular Devices). The amount of product was calculated from a standard curve of glucose assembled and read under the same conditions as samples.

Controls without substrate or samples were incubated at the same conditions used for the assay of each substrate. One unit of enzymatic activity (U) is defined as the amount of enzyme which releases 1µmol of product/min.

Protein concentration was determined using Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) (Smith et al., 1985), following the protocol provided by the manufacturer for microplate procedures, using bovine serum ovalbumin as a standard.

Effect of substrate concentration and pH on α-glucosidase activity

The effect of pH on enzyme activity was studied at 30 ^{0}C using 0.1 mM MUαGlu or 200 mM sucrose (final concentrations) as substrates and the following buffers (200 mM): sodium citrate-citric acid (pH 3-4), sodium acetate (pH 3.7-6), citrate-phosphate (pH 3-7), MOPS (pH 6-8), EPPS (pH 7-9), AMPSO (pH 8-10), CAPS (pH 10-11). Assays were made as described in the section "α-glucosidase assay and protein determination".

To determinate the optimum pH in tested conditions a curve was fitted according to the equation (1) described below (Segel, 1975):

$$V_{\mathbf{max\ app}} = \frac{V_{\mathbf{max}}}{1 + \frac{[\mathrm{H}^+]}{K_{es1}} + \frac{K_{es2}}{[\mathrm{H}^+]})} \quad \mathbf{(1)}$$

The effect of substrate concentrations on the activity was determined at 30 ^{0}C using 200 mM AMPSO pH 8 and (0.01-2.3 mM) MUαGlu or using 200 mM citrate-phosphate pH 6.0 buffer and (10-1400 mM) sucrose. Activity was determined depending on the substrate as described in the item "α-glucosidase assay and protein determination". Values for apparent $K_{\mathbf{M}}$ ($K_{\mathbf{M\ app}}$), apparent $V_{\mathbf{max}}$ ($V_{\mathbf{max\ app}}$) and K_I were determined using GraFit Software (GraFit version_7.0.3, Erithacus Software Limited) and the Michaelis-Menten equation as described in Segel, 1975. When inhibition by high substrate concentration was detected, kinetic parameters were calculated taking into account the reaction (Scheme 1) and the equation (2) below (Cleland, 1979):

$$\mathbf{V} = \frac{V_{\mathbf{max}}[S]}{K_{\mathbf{M}} + [S]\,(1 + \frac{[S]}{K_i})} \quad \mathbf{(2)}$$

Gel filtration chromatography

Midgut soluble and midgut tissue fractions were applied (sample volume 0.5 mL) to a Superdex 200 10/300 GL (1 cm x 30 cm) (AKTA Purifier, GE, Uppsala, Sweden). The column was equilibrated with 50 mL of 20 mM citrate-phosphate buffer pH 6.0 for midgut content sample and 20 mM citrate-phosphate buffer pH 6.0 containing 1% Triton X-100 for midgut tissue sample. Proteins were eluted with the same buffer used to equilibrate column at 0.5 mL.min^{-1}. Fractions of 0.5 mL were collected and assayed with MUαGlu and sucrose as described in the section "α-glucosidase assay and protein determination".

Statistical analysis

All statistical analysis were performed using GraphPad Prism 6.0 for Windows (San Diego, California, USA). To identify outliers the ROUT method based on the False Discovery Rate (FDR) was used and Q was established as 1 %. To determine how far the distribution is from Gaussian in terms of asymmetry and shape the D'Agostino-Pearson Omnibus K2 normality test was used. For comparison of normally distributed data, one-way ANOVA was used followed by Tukey's multiple comparison tests, and significance was considered when $p < 0.05$. Results are expressed as the means ± S.E.M.

Results

Distribution of α-glucosidase activity in tissues using different substrates

Different parts of *L. longipalpis* females, collected at 0 to 3 hours after emergence, were individually assayed using MUαGlu and sucrose as substrates to see the anatomical distribution of α-glucosidase activities. Activity was found in all tissues, but the distribution was different depending on the substrate used (Figure 1). Using MuαGlu as substrate, most activity was found in the carcass (111 ± 5 µU/insect), but the enzymes present in the midgut also showed significant activity. Using sucrose as substrate, highest activities were localized in the midgut contents (8.1 ± 0.7 mU/insect) and midgut tissues (16 ± 2 mU/insect). The results also show that the activity is much higher on sucrose than on MUαGlu, showing the preference of these enzymes for the natural substrate. In the case of crop and midgut contents, activity was 2500 and 3700 times, respectively, higher on sucrose than MUαGlu.

Detection of α-glucosidase activities in the midgut of females fed with water or sucrose

The activity of α-glucosidase was measured in the midgut contents and midgut tissues of females fed with water or sucrose for 5 or 10 days, respectively. Females were initially fed with water only to verify if there was some sucrase induction associated with age (Figure 2a, b). Interestingly, an induction happened for both midgut contents and midgut tissues after 2 days for the activity against MUαGlu (Figure 2a). However, no induction

was identified for activity against sucrose in females fed only with water. Due to the increasing mortality of insects, we were able to follow these activities only until 5 days.

When feeding females with 1.2 M sucrose for 10 days, we observed different expression profiles for the α-glucosidase activity when we compare the two substrates (Figure 2c,d). In the midgut contents no induction was observed during the 10 days of feeding with sucrose for both substrates (Figure 2c, d), although for midgut tissue there was an induction of activity against sucrose, with a peak after 2 days of feeding with sugar (Figure 2d). A significant increase of 2.8 times in the activity can be observed when comparing recently emerged females (17 ± 2 mU/midgut) to females fed with 1.2 M sucrose (49 ± 2 mU/midgut). After 3 days, the activity decreases but does not return to basal levels, suggesting that the sucrase activity induction occurred due to the presence of sugar meal (Figure 2b). However, there was no induction in the midgut tissue when we followed activity against MUαGlu (Figure 2c).

Females were also fed with different concentrations of sucrose (0.3 M, 0.6 M and 1.2 M) for 2 days and the activity against sucrose was measured, to verify if the expression observed has any relation with the sugar concentration in the meal (Figure 3a). The activity in the midgut contents was not induced by any of the sucrose concentrations tested (Figure 3a). Nevertheless, in the midgut tissue we observed a dose-response relationship between the sucrase activities and sugar concentrations in the sugar meal, from 0.3 to 1.2 M. This is not a consequence of higher ingestion rates, as an equal meal volume was ingested for all sugars concentrations (Figure 3b).

Induction of α-glucosidase activity in the midgut of females fed with mono or disaccharides

Females were fed with glucose, fructose, sucrose or maltose for 2 days. α-glucosidase activity was measured with sucrose, to understand if the sucrase activity induction relies on the type of sugar provided. Results demonstrated that after 2 days of feeding on glucose, fructose or maltose, the activity of α-glucosidase was higher than the observed in recently emerged females, but not statistically different from the observed in females that were fed on water for 2 days (Figure 4a). However, feeding with sucrose resulted in a significant higher sucrase activity when compared to all groups mentioned above (Figure 4a). Figure 4b demonstrates that there is no significant difference in the volume ingested by females feeding of fructose, sucrose, and maltose (32 ± 2; 34 ± 2; 30 ± 2 ηL

per individual, respectively), neither in the percentage of insect that ingests sugar (100% in all these cases, Figure 4c). However, just a small percentage of females (29 %) can feed on glucose (Figure 4c), with a significantly smaller mean volume in the ingested sugar meal (9 ± 1 ηL; Figure 4b).

Detection of α-glucosidase activity in the midgut of females fed with blood

We decided to test if α-glucosidase activities are also involved in blood digestion. For this, activity was measured with MUαGlu and sucrose as substrates, in females starved for 3 days (water-mantained), and in females fed with blood, every 24 hours following feeding during 4 days (24, 48, 72 and 96 h). Activities were measured in midgut contents and midgut tissues (Figure 5a, b).

Results showed induction of both activities, against sucrose or MuαGlu, in all fractions (midgut contents and midgut tissues) 24 h after blood feeding. The activity observed in the midgut contents after blood feeding seems to be different from the previous activity observed after sugar feeding, as we observed induction by blood and no induction when females are fed with 1.2 M sucrose (Figure 2 and 5). In the case of midgut tissues, the pattern of induction is also different, as we observed an increase in activity against MUαGlu after 24 h, wich was absent in females fed with sucrose (Figure 5a and 2a).

Determination of α-glucosidases kinetic properties: effect of pH and substrate concentration on enzyme activities

Because we observed different expression patterns for the α-glucosidase activities when comparing the midgut contents to the midgut tissue using different substrates, we decided to measure the kinetic parameters of each activity to observe if these activities also differ in their molecular properties. The effect of pH and substrate concentration in α-glucosidase activity was determined in samples obtained from recently emerged females (nonfed, NF), females fed with 1.2 M sucrose for 2 days (SF) and females 2 days after feeding on blood (BF), using sucrose and MUαGlu as substrates.

The results are summarized in Table 1. Enzymes acting on MUαGlu have maximum activity in neutral to basic pH, with optimum pH around 7.5 and significant activities between pH 6 and 8.0. Using sucrose as substrate a more acidic optimum pH was found, around 6.5, with significant activities between pH 6 -7 with a substantial decrease in activity at pH 8.0. When comparing samples from the same type of sample, similar

profiles were obtained under different physiological conditions (Supplementary Figures 1 and 2).

The kinetic properties of α-glucosidases from *L. longipalpis* in different physiological conditions are described in Tables 2 and 3. Activities from the same tissue under different physiological conditions showed different substrate affinities. The kinetic parameters were also different when comparing the samples to each other in the same physiological condition.

When sucrose was used as the substrate, the $K_{m\ app}$ values for all samples varied according to the condition tested. The highest affinity was found for samples of the crop from sugar-fed females and midgut content from blood-fed females. $K_{m\ app}$ for midgut content changed with a significant reduction when females were fed with blood compared to nonfed females or females fed with sucrose. In the opposite side, it is important to note the high $K_{m\ app}$ values found for midgut tissue for sugar and blood-fed females. As expected $V_{max\ app}$ values were highest for the midgut tissue, with specific activities of 4 mU/µg protein for nonfed females and 18.7 mU/µg protein for sugar-fed females.

For MUαGlu, the lowest $K_{m\ app}$ was found for the midgut contents in sugar- and blood-fed insects. In the midgut contents, the $V_{max\ app}$ was 3X higher in blood-fed females when comparing to sugar-fed females, but the affinity for the substrate was practically unaffected. For the midgut tissue, the $V_{max\ app}$ was the highest found for nonfed females (300 µU/µg protein), with a specific activity 2X higher than the observed in midgut content (141.30 µU/µg protein). The crop had the lowest $V_{max\ app}$ in all conditions, the lowest activity being found in the crop of blood-fed females (2.8 µU/µg protein).

We also observed that enzymes for all tissues and feeding conditions showed classical Michaelis-Menten kinetics when MUαGlu was used as substrate but, the sucrase activity present in the crop and midgut tissue was inhibited in high sucrose concentrations, except in the crop for sucrose-fed females (Table 2 and Figure 6). Using midgut tissue preparations, the sucrase activity started to be inhibited in the presence of small concentrations of sugar, and the K_i value found is similar or smaller than the $K_{m\ app}$ for the substrate.

Chromatographic analysis

To further clarify if the different induction profiles and biochemical properties found in different tissues and physiological conditions correspond to α-glucosidases with

different molecular properties, we subjected samples of midgut contents and tissues to gel filtration chromatographies. Representative chromatograms for all samples tested are displayed in Figure 7.

For the midgut contents we observed different peak profiles when considering the feeding conditions and substrates used (Figure 7a, b, and c). Two major peaks with activity against MUαGlu can be identified for nonfed (7 and 13.5 mL) and blood-fed females (7 and 12.5 mL), and the activity of the first peak almost doubles in samples from blood-fed females. It is important to point the presence of a third peak in samples from sugar-fed females (7, 11 and 14 mL). When the chromatographic fractions were analyzed with sucrose as substrate different profiles are obtained, with three activity peaks for non-fed and blood-fed females, with higher activity in the last condition. For sugar-fed females only the first peak (7 mL) could be identified. The first peak had the same retention time as the ones identified with MUαGlu, so this probably corresponds to an enzyme with activity on both substrates.

For midgut tissue samples (Figure 7d, 7e and 7f), a major peak was identified with retention volume around 11 mL in all conditions, with significant activity against both substrates used. Using MUαGlu a minor peak appears around 13.5 mL retention volume and for sucrose at 17.5 mL. For sugar-fed and blood-fed females the activity was higher than the activities found for non-fed females, as expected. For the midgut tissue samples Triton X-100 was added to the running buffer as we were probably working with proteins solubilized from the membrane fraction. This certainly resulted in a lower resolution, since the Triton X-100 theoretically adds a molecular mass of about 90 kDa to the solubilized proteins, corresponding to the detergent micelle.

The activity recovered after chromatography was calculated (Table 4) and surprisingly activity against MUαGlu after chromatography was 1.5 (midgut tissue from blood-fed females) to almost 5 x (midgut content from sugar-fed females) higher than the activity quantified in samples initially applied on gel filtration. This probably suggests that an inhibitor was removed during chromatography. For sucrase activity, the recovery was higher than initial activity only for midgut tissue samples from blood-fed females.

Discussion

Sugar meals are essential for the energy requirements of sand flies. These insects have a diet rich in sugars, glycolipids, and glycoproteins, which makes important to study

carbohydrases involved in their digestion process. In this work, we demonstrate the presence of α-glucosidase activity in all tissues (head, carcass, crop, midgut, and hindgut) of *L. longipalpis* unfed females, using two different substrates (Figure 1). The distribution of the activity varied with the type of substrate used. Higher activities were found for the natural substrate sucrose, with more than 1000 X higher activities when compared to MuαGlu. This demonstrates the specialization of these enzymes for sucrose digestion. Jacobson et al., 2007 showed that the composition of sugar-degrading enzymes in the midgut of *Phlebotomus papatasi* sand flies depends on the ecological habitat they were collected, corresponding to different sugar sources. This correlates with the specialization of α-glucosidase described here in the digestion of sucrose.

The presence of carbohydrases has been reported by different groups in different sand fly species. Samie et al., (1990) analyzed the incubation of head and thorax homogenates from *P. papatasi* with sugar solutions and detected monosaccharides by HPLC analysis. They suggested these products were generated by the action of glycosidases from salivary glands and that they are mixed with the sugar meal in the crop. For mosquitoes, it is accepted that sugar digestion occurs in the crop by α-glucosidases produced by the salivary glands (James et al., 1989; Marinotti and James, 1990; Marinotti et al., 1996). Jacobson et al., 2001 described the α-glucosidase activity in the gut and salivary glands of unfed *P. papatasi*. α-glucosidase activity was found in the midgut of blood-fed *Phlebotomus langeroni* females, and a small activity at homogenates of crop, but it was considered contamination from the midgut (Dillon and El-Kordy, 1997). Gontijo et al., 1998, described α-glucosidase activity only in the midgut of nonfed *L. longipalpis*. They used the substrate pNPαGlu (p-nitrophenyl-α-glucopyranoside) and sucrose at 1 mM and 25 mM final concentrations, respectively, to characterize this activity. One difference between our work and Gontijo et al., 1998, relies on the type of substrates and their final concentrations in the assays. For sucrose the $\boldsymbol{K_{m\ app}}$ for these enzymes is very high. In this way, to saturate the enzyme and have a measurement of activity directly proportional to the quantity of enzyme, a very high final concentration of sucrose in the assays is necessary. When we used a small final concentration of sucrose in the assays (data not shown), we were not able to identify activity in the crop and hindgut. The concentration of the sucrose as substrate used also explains the difference in the midgut activity of unfed females of 2.3 mU/midgut (Gontijo et al., 1998) to 16 mU/midgut (our work).

Midgut α-glucosidase activity was found in the soluble and tissue-associated form in non-fed, sugar-fed and blood-fed females, in accordance with what has been described by Gontijo and coworkers (1998). They were the first to characterize a membrane-bound α-glucosidase attached to the epithelial microvilli from the midgut of *L. longipalpis*. Enzymes involved in the late stages of digestion such as exoglycosidases are usually associated with the membrane of the midgut (Terra and Ferreira, 1994).

For nonfed females, activity against MUαGlu in the midgut content and tissue was induced after 2 days of starvation (Figure 2a), so the synthesis of this α-glucosidase do not seem to be directly related to the ingestion of sucrose. Factors like hormones can regulate the production of α-glucosidases in recently emerged females (Souza-Neto et al., 2007). This enzyme may be involved in the intracellular metabolism, somehow in the mobilization of energetic reserves.

In sugar-fed females, sucrase activity in the midgut tissue was also induced after 2 days (Figure 2d). Besides that, this induction was variable depending on the sugar concentration and the type of sugar used to feed females (Figure 3). In this way, these α-glucosidases seem to be induced by sugar ingestion. After induction, this activity does not return to basal levels, when compared to nonfed females. This might reflect the dinamics of sugar digestion inside the sand fly, where part of the sugar meal is directed to the empty midgut and the rest is stored inside the crop. After the meal, small amounts of sugars are released into the midgut during the subsequent days. Jacobson and Schlein (2001) described the increase of α-glucosidase activity in the homogenates of females of *P. papatasi* that fed during the night on the plant *Capparis spinosa*, rich in sucrose and starch. Jacobson et al. (2007) analyzed α-glucosidase activity in females of *P. papatasi* from different ecological habitats, and they showed an increase in the activity measured with sucrose as substrate after feeding with 1 M sucrose and a decrease in activity when pNPαGlu was used. Dillon and El-Kordy (1997) described an activity increase in the midgut of *P. papatasi* after feeding insects with 30% (w/v) sucrose solution.

When we used different sugars to feed *L. longipalpis* females, an increase in the activity was found in the midgut tissue of insects fed with 1.2 M sucrose (Figure 4). This effect can be due to the fact that we used sucrose itself to measure the activities. If maltose had been used as substrate, we might have found a different profile, with the higher activity being detected in females that were fed on 1.2 M maltose. Different works demonstrated that the repertory of glycosidases secreted seems to match the source of sugar present in the diet of these insects (Souza-Neto et al., 2007; Jacobson et al., 2007).

Interestingly, when females were fed with blood an induction was observed in both midgut contents and tissue after 24 h, and this induction was found using MUαGlu and sucrose as substrates (Figure 5). These females had no previous contact with sugar suggesting that this induction was specifically triggered by blood ingestion. Since the induction of activity in the midgut contents was detected using both substrates, this demonstrates that an α-glucosidase with high activity against sucrose might be acting in the digestion of blood. In *P. langeroni*, a soluble α-glucosidase activity was described in the midgut after blood ingestion, being rapidly produced (after 1 h) with maximum activity 24 h after blood feeding (Dillon and El-Kordy, 1997).

It is possible that these enzymes cross the peritrophic matrix and act in the endoperitrophic space, together with other glycosidases, in the hydrolysis of blood glycoconjugates. The increase in this activity (24 h) suggests that these enzymes may be stored in the cells of the gut being secreted in response to the blood supply, with more enzymes being synthesized and secreted as needed. The blood has predominantly protein in their composition and a high proportion of glycoproteins, and their carbohydrate portions along with glycolipids are theoretically susceptible to cleavage by exoglycosidases including α-glucosidases. Fragments of glycopeptides produced by the proteolytic cleavage of trypsin and aminopeptidases are more susceptible to the action of exoglycosidases, so the increase of α-glucosidase activity may follow the pattern of protease activity. Another plausible role for soluble α-glucosidase induction is the involvement in the detoxification of heme during blood digestion, auxiliating the nucleation of hemozoin cristals, as described for α-glucosidases from *Rhodnius prolixus* (Mury et al., 2009).

In sand flies, the induction of protease activity after the blood feeding seems to reach a peak between 24 and 48 h with some variation depending on species and type of blood ingested (Borovsky and Schlein, 1987; Dillon and Lane, 1993; Ramalho-Ortigão et al., 2003). For *L. longipalpis*, according to Sant'Anna et al. (2009) and Moares et al. (2018), trypsin activity starts to be induced after 24 h after blood feeding with maximum activity after 48 h, which is consistent with the α-glucosidase activity observed in our experiments. An additional issue that remains to be investigated is that part of the activity present in the midgut contents might be derived from microbial glycosidases.

The characterization of the activities present in different tissues of *L. longipalpis* showed that the pH of optimal activities is only slightly affected by the feeding condition or tissue localization of enzymes (Table 1). Most differences in the effect of pH on

activity are related to the type of substrate used. Activity against MUαGlu showed a wider range of optimal activity, between 7-8 (Supplementary 1). Activity against sucrose is more affected by pH changes with a more sharp profile in the adjustment curve, and optimal activity around 6.0 (Supplementary 2). These results are consistent with the pH value determined in the gut of sucrose-fed females, which is approximately 6 (Gontijo et al., 1998). It is interesting to observe that when we measured the activity with sucrose, even for females digesting blood, the enzymes display a better performance under an acidic environment. In 2008, Santos and coworkers demonstrated that the pH in the abdominal midgut becomes alkaline during blood digestion but the thoracic midgut remained acid (pH 5.5-6.0). They also demonstrated a higher activity of α-glucosidase in the thoracic midgut in comparison to the abdominal midgut in nonfed females. All these results together suggest that different α-glucosidases are acting in the thoracic and anterior midgut in the sugar and blood digestion. For *P. langeroni* an α-glucosidase with an optimum pH of 7.5 was detected in blood-fed females, suggesting an alkaline activity induced only in blood-fed females (Dillon and El-Kordy, 1997). In this case, the substrate used was pNPαGlu, which may mymic the synthetic substrate used in our study (MUαGlu).

The results showed in our work demonstrates that *L. longipalpis* have different α-glycosidases/sucrases with different affinities and behaviors depending on substrate (Table 2 and Table 3). Some of them did not follow the classical Michaelis-Menten kinetics and were inhibited by high concentrations of sucrose. We showed a $\boldsymbol{K}_{\mathbf{M\ app}}$ variation from 0.37 to 4.7 mM when MUαGlu was used as substrate and from 11 to 800 mM when sucrose was the substrate. These high $\boldsymbol{K}_{\mathbf{M\ app}}$ values described for sucrose substrate suggest that these enzymes may show a huge variation in activity in a broad range of sucrose concentrations, which may be compatible with the feeding of nectars presenting sucrose in concentrations as high as 1.2 M (Heil, 2011). However, this high $\boldsymbol{K}_{\mathbf{M\ app}}$ values may not actually reflect the activity range of the enzyme before saturation of the active site, due to the inhibition caused by high substrate concentrations, when it was the case. The substrate inhibition causes an artifact where the $\boldsymbol{V}_{\mathbf{max}}$ calculated is much higher than any activity measured, and the calculated $\boldsymbol{V}_{\mathbf{max\ app}}$ will never be reached. Correspondingly, the $\boldsymbol{K}_{\mathbf{m\ app}}$, as the enzyme begins to be inhibited before the catalytic site is completely saturated (see the values of $\boldsymbol{K}_{\mathbf{m\ app}}$ and $\boldsymbol{K}_{\mathbf{i}}$), does not correspond to the substrate concentration where the enzyme shows 50% of its maximal rate.

The inhibition by high substrate concentrations can be explained by transglycosylation reactions, by the binding of a second molecule of the substrate in the active site, or even by an osmotic effect (high substrate concentration) leading to a decrease in the reaction rate. Many insect α-glucosidases are classified in family 13 of glycosyl-hydrolases (GH13) and such enzymes are known to catalyze the glycoside linkages by retaining the anomeric configuration of the substrate. This type of mechanism allows these enzymes to do transglycosylation (Chiba et al., 1997). It is noteworthy that similar kinetic behavior was described for the α-glucosidase of *Acyrtosiphon pisum* (Cristofoletti et al., 2003). In this report, the authors consider that transglycosylation might be an adaptation for the obtention of monosaccharides from sucrose without the increase in osmolarity, as a simple hydrolytic reaction of a 0.7 M sucrose solution (phloem concentration) might result in osmotic shock for the midgut epithelial cells. In this respect, *L. longipalpis* enzymes might have the same biochemical adaptation. It would be very interesting to observe if this is a common trait of α-glucosidase from insects feeding on nectar or phloem sap, and verify if this is a case of evolutionary divergence or convergence.

Two main mechanisms of transglycosylation reactions using sucrose as substrate have been described. In the first, after the hydrolysis of the glycosidic bond, the glucosyl residue is retained in an intermediate linked to one of the enzyme catalytic carboxylates, and the fructose residue is released from the active site. After that, the glucosyl moiety is transferred to another sucrose molecule, resulting in a trisaccharide with two glucoses. After several rounds of this reaction, we expect the preferential release of the fructose and the generation of polyglucose chains linked to the initial sucrose molecule (Trincone, 2011). The second known mechanism for transglycosylation with sucrose as donor and acceptor consists in the retention of the fructosyl moiety as an intermediate linked to the catalytic glutamate, with the release of a glucose molecule after the initial hydrolysis of the glycosydic bond. After that, the fructosyl is transferred to an entering sucrose molecule, resulting in a trisaccharide with two fructoses. After several rounds, we expect releasing of most glucose from the substrate, and production of a polyfructose linked to an initial sucrose molecule (Han et al., 2006).

Our data strongly suggest that, in the case of transglycosylation in the presence of high sucrose concentrations, the mechanism preferred by *L. longipalpis* α-glucosidases is the first one described above. The first evidence that points to that is the avidity and preference of the female sand fly for fructose over glucose, which might be coherent with

569 the release of fructose and the correspondent absorption systems for this sugar. The
570 second is that our activity measurements are based on the measurement of released
571 glucose, and the release of glucose is hindered mainly by the first transglycosylation
572 mechanism, as in the second type of transglycosylation this monosaccharide is released
573 after the first hydrolysis step. A similar mechanism was proposed for the *A. pisum* enzyme
574 (Cristofoletti et al., 2003), suggesting that this property may be present in a common
575 ancestor of Eumetabola. Furthermore, it is important to note that additional experiments
576 to confirm whether the inhibition by high substrate concentration was caused by
577 transglycosylation and which mechanism is involved, are necessary.

578 Besides that, the high $K_{m\ app}$ values and inhibition by substrate described for some
579 enzymes of *L. longipalpis* in sucrose hydrolysis might be related to their involvement in
580 metabolic control. Different enzymes are known to have a high K_m value as a form of
581 regulation of specific metabolic pathways. As the concentration of sucrose on sap varies
582 from 0.5 to 1.2 M (Gontijo et al., 1996), it could saturate the enzymes involved in sugar
583 digestion, but with a high $K_{m\ app}$ value these enzymes will need higher amounts of
584 substrate (sucrose) to reach their maximum activity. This effect is increased by substrate
585 inhibition, and so the digestion of sugars will occur at a lower rate than would occur if
586 the $K_{m\ app}$ were low. In this way, the availability of glucose for cell uptake might be strictly
587 controlled.

588 Kinetic parameters for some α-glucosidases from insects have been determined
589 using different substrates and their K_m values show a lot of variation. *L. longipalpis* α-
590 glucosidase from midgut tissue showed the highest $K_{m\ app}$ when compared to others
591 insects. Souza-Neto *et al*. described different isoforms of α-glucosidases from *Anopheles*
592 *aquasalis* with a K_m variation of 1.31 to 8.26 mM for the substrate pNPαGlu. The α-
593 glucosidase III from *Apis mellifera* shows a **Km** value of 30 mM for sucrose and 13 mM
594 for pNPαGlu (Nishimoto *et al*., 2001) and the α-glucosidase I from *Apis cerana japonica*
595 2.4 mM for sucrose (Wongchawalit *et al*, 2006).

596 The profiles obtained after gel filtration chromatography, especially for the
597 midgut contents, and the biochemical parameters determined in different conditions,
598 suggest the presence of different α-glycosidases acting in the metabolism of sand flies
599 (Figure 7). We demonstrated at least three different α-glycosidases/sucrases that are
600 secreted as soluble enzymes, and a major membrane-bound activity in the midgut of *L.*
601 *longipalpis* females. To distinguish the specific roles of each of these α-glucosidases in
602 the digestive process, purification and additional characterization are desired to

603 understand how this enzymatic complex acts under different conditions and if the purified
604 enzymes have different specificities or act sinergycally in the hydrolysis of substrates
605 during digestion.

606 The compartmentalization of the insect digestive process is a hallmark of the
607 understanding of insect physiology that was obtained during the XX century. According
608 to that, in several insects the digestive process is divided among three basic
609 compartments: the endoperitrophic lumen, the ectoperitrophic space, and the surface of
610 the enterocytes (Terra and Ferreira, 2005). In this respect, glycoside hydrolases that act
611 on the initial steps of digestion tend to be soluble enzymes, acting on large molecules as
612 polysaccharides inside the endoperitrophic space. The most common enzymes related to
613 the initial steps of carbohydrate digestion are amylases, cellulases, β-1,3-glucanases,
614 xylanases, pectinases, chitinases, and lysozymes (Terra and Ferreira, 2005). Conversely,
615 glycoside hydrolases acting in the intermediate and final stages of carbohydrate digestion,
616 namely the hydrolysis of oligo- and disaccharides, respectively, tend to be confined in the
617 ectoperithrophic space or bound to the apical cellular membrane of the enterocytes, as
618 proteins anchored to the lipid bilayer or associated to the glycocalyx. In this category, we
619 found α-glucosidases, β-glucosidases, β-galactosidases, trehalases, acetyl-
620 hexosaminidases, β-fructosidases and α-galactosidases (Terra and Ferreira, 2005).

621 Dipterans are considered as an advanced Order in respect to the evolutionary
622 acquisition of derived traits in the midgut, with the description of a well-
623 compartmentalized distribution of enzymes related to the digestive process (Terra and
624 Ferreira 1994; 2005). In this respect, the presence of soluble α-glucosidases, secreted to
625 the lumen and induced preferentially by blood, suggest that these enzymes might be
626 acting on a large substrate molecule, during the initial stages of digestion. As the blood
627 has insignificant amounts of polysaccharides, and α-glucosidases have low activity
628 against this type of substrate, it is possible that the natural substrate of *L. longipalpis*
629 soluble α-glucosidases consists in the carbohydrate moiety of blood glycoproteins or
630 glycolipids. The tissue bound enzyme, which is induced by the ingestion of sucrose, might
631 be participating in the canonical role of membrane bound glucosidases, the final digestion
632 of small sugars as sucrose and other disaccharides. In this respect, the involvement of *L.*
633 *longipalpis* adult α-glucosidases in physically separated stages of hydrolysis of very
634 different food molecules might be the physiological and evolutionary basis for the
635 diversification of this type of digestive enzyme activity.

These results together show that there are different α-glucosidases involved in the energetic metabolism of *L. longipalpis* females. Some of these enzymes might be specialized and some might act synergically in the hydrolysis of different substrates, being induced according to the feeding regime. Some of these α-glucosidases appear to be specialized in the sugar digestion, some may participate in the digestion of blood glycoproteins or glycolipids, and some might be involved in the mobilization of energetic storages during starvation. Our work increases the knowledge about the biology of phlebotomines, providing a detailed description of the induction of α-glucosidase activity in different physiological conditions, demonstrating the presence and especialyzation of α-glucosidases in different compartments of adult females of *L. longipalpis*.

Acknowledgments

This research was supported by the Research Agencies FAPERJ, CNPq, CAPES, and FIOCRUZ (Brazil). Samara Graciane da Costa is a Ph.D. student of Cellular and Molecular Biology Post Graduation Program of Oswaldo Cruz and received a sandwich doctorate grant from Science without Borders (CNPq/CAPES). Fernando A. Genta and Rod J. Dillon are staff members of their respective departments. We would like to acknowledge Dr. Caroline da Silva Moraes and M.Sc. Tainá Neves Ferreira for helping with maintenance of the sand fly colony.

Author's Contributions

Conception and design of the work: SGC, PB, RJD, and FG. Obtaining experimental data: SGC. Data analysis: SGC and FG. Writing and revision of the manuscript: SGC, PB, RJD, and FG. All authors read and approved the final version.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Akhoundi M, Kuhls K, Cannet A, Votýpka J, Marty P, Delaunay P, Sereno, D. (2016). A historical overview of the classification, evolution, and dispersion of *Leishmania* Parasites and Sandflies. PLoS Negl Trop Dis. 2016;10 (3):1–40.

Borovsky, D., Schlein, Y. (1987). Trypsin and chymotrypsin-like enzymes of the sandfly *Phlebotomus papatasi* infected with *Leishmania* and their possible role in vector competence. Med Vet Entomol. 1:235. doi: 10.1111/j.1365-2915.1987.tb00349.x

Cameron, M. M., Pessoa, F. A. C, Vasconcelos, A. W., Ward, R.D. (1995). Sugar meal sources for the phlebotomine sandfly *Lutzomyia longipalpis* in Ceara State, Brazil. Med Vet Entomol. 9:263. doi: 10.1111/j.1365-2915.1995.tb00132.x

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 37:233. doi: 10.1093/nar/gkn663

CDC. 2013. Biosafety in microbiological and biomedical laboratories. Edited by L. C. Chosewood and D. E. Wilson. Fifth edit. USA.

Chiba, S. (1997). Molecular mechanism in alpha-glucosidase and glucoamylase. Biosci Biotechnol Biochem. 61:1233. doi: 10.1271/bbb.61.1233

Cleland, W. W. (1979). "Substrate inhibition" in Methods in Enzymology, ed. D. L. Purich (New York: Academic Press.), 500–513.

Cristofoletti PT, Ribeiro AF, Deraison C, Rahbe Y, Terra WR. (2003) Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrthosiphon pisum*. J Insect Phisiol 49: 11-24.

Dillon, R. J., Lane, R. P. (1993). Bloodmeal digestion in the midgut of *Phlebotomus papatasi* and *Phlebotomus langeroni*. Med Vet Entomol. 7:235. Doi: 10.1111/j.1365-2915.1993.tb00681.x

Dillon, R. J., El-Kordy, E. (1997). Carbohydrate digestion in sandflies: alpha-glucosidase activity in the midgut of *Phlebotomus langeroni*. Comp Biochem Physiol B Biochem Mol Biol. 116:35. doi: 10.1016/S0305-0491(96)00195-2

Dixon, M., Webb, E. C. (1979). Enzymes. New York: Academic Press.

Ferreira, T.N., Pita-Pereira, D., da Costa, S.G., Brazil, R.P., Moraes, C.S., Díaz-Albiter, H., Genta, F.A. (2018). Transmission blocking sugar baits for the control of *Leishmania* development inside sand flies using environmentally friendly beta glycosides and their aglycones. Parasites and Vectors. *In press.*

Gontijo, B., de Carvalho, M. L. R. (2003). Leishmaniose tegumentar americana. Revista da Sociedade Brasileira de Medicina Tropical 36: 71–80.

Gontijo, N. F., Almeida-Silva, S., Costa, F. F., Mares-Guia, M. L., Williams, P., Melo, M. N. (1998). *Lutzomyia longipalpis*: pH in the gut, digestive glycosidases, and some

speculations upon *Leishmania* development. Exp Parasitol. 90:212. doi: 10.1006/expr.1998.4336

Gontijo, N. F., Melo, M. N., Riani, E. B., Almeida-Silva, S., Mares-Guia, M. L. (1996). Glycosidases in L*eishmania* and their importance for *Leishmania* in phlebotomine sandflies with special reference to purification and characterization of a sucrase. Exp Parasitol. 83: 117.doi: 10.1006/expr.1996.0055

Han J, Xu X, Gao C, Liu Z, Wu Z. (2016) Levan-producing *Leuconostoc citreum* strain BD1707 and its growth in tomato juice supplemented with sucrose. Appl Environ Microbiol 82: 1383-1390.

Heil, M. (2011) Nectar: generation, regulation and ecological functions. Trends inPlant Sciences 16, 191-200.

Jacobson, R. L., Studentsky, L., Schlein, Y. (2007). Glycolytic and chitinolytic activities of *Phlebotomus papatasi* (Diptera: Psychodidae) from diverse ecological habitats. Folia Parasitol (Praha). 54:301. doi: 10.14411/fp.2007.039

Jacobson, R. L., Schlein, Y. (2001). *Phlebotomus papatasi* and *Leishmania major* parasites express alpha-amylase and alpha-glucosidase. Acta Trop. 78:4. doi: 10.1016/S0001-706X(00)00164-9

Jacobson, R. L., Schlein, Y., Eisenberger, C. L. (2001). The biological function of sand fly and *Leishmania* glycosidases. Med Microbiol Immunol. 190: 51. doi: 10.1007/s004300100079

James, A. A., Blackmer, K., Racioppi, J. V. (1989). A salivary glandspecific, maltase-like gene of the vector mosquito, *Aedes aegypti*. Gene. 75:73. doi: 10.1016/0378-1119(89)90384-3

Junnila, A., Müller, G. C., Schlein, Y. (2011). Identification of plant tissues from the gut of *Phlebotomus papatasi* by DNA analysis. Acta Trop. 117:14. doi 10.1016/j.actatropica.2010.08.019

Kamhawi, S. (2006). Phlebotomine sand flies and *Leishmania* parasites: friends or foes? Trends Parasitol. 22:439. doi: 10.1016/j.pt.2006.06.012

Killick-Kendrick, R., Killick-Kendrick, M. (1987). Honeydew of aphids as a source of sugar for *Phlebotomus ariasi*. Med Vet Entomol. 1:297. doi.org/10.1111/j.1365-2915.1987.tb00358.x

Lainson R., Shaw J.J. (1992). A brief history of the genus *Leishmania* (Protozoa: Kinetoplastida) in the Americas with particular reference to Amazonian Brazil. Ciência e Cultura 44: 94-106.

MacVicker, J. A. K., Moore, J. S., Molyneux, D. H., Maroli, M. (1990). Honeydew sugars in wild-caught Italian phlebotomine sandflies (Diptera: Psychodidae) as detected by high performance liquid chromatography. Bull Entomol Res. 80:339. doi: 10.1017/S0007485300050537

Marinotti, O., deBrito, M., Moreira, C. K. (1996). Apyrase and alpha-glucosidase in the salivary glands of *Aedes albopictus*. Comp Biochem Physiol. 113B:675. doi: 10.1016/0305 0491(95)02035-7

Marinotti, O., James, A. A. (1990). An α-glucosidase in the salivary glands of the vector mosquito, *Aedes aegypti*. Insect Biochem. 20: 619. doi: 10.1016/0020-1790(90)90074-5

Moore, J. S., Kelly, T. B., Killick-Kendrick, R., Killick-Kendrick, M., Wallbanks, K. R., Molyneux, D. H. (1987). Honeydew sugars in wild-caught *Phlebotomus ariasi* detected by high performance liquid chromatography (HPLC) and gas chromatography (GC). Med Vet Entomol. 1:427. doi: 10.1111/j.1365-2915.1987.tb00373.x

Moraes, C. S., Aguiar-Martins, K., Costa, S. G., Bates, P. A., Dillon, R. J., Genta, F. A. (2018). Second blood meal by female *Lutzomyia longipalpis*: enhancement by oviposition and its effects on digestion, longevity, and *Leishmania* infection. Biomed Res Int. 2018: 2472508. doi: 10.1155/2018/2472508.

Moraes, C. S., Diaz-Albiter, H. M., Faria, M. V., Sant'Anna, M.R., Dillon, R.J., Genta, F. A. (2014). Expression pattern of glycoside hydrolase genes in *Lutzomyia longipalpis* reveals key enzymes involved in larval digestion. Front Physiol. 5:276. doi: 10.3389/fphys.2014.00276.

Moraes, C. S., Lucena, S. A., Moreira, B.H., Brazil, R. P., Gontijo, N.F., Genta, F.A. (2012). Relationship between digestive enzymes and food habit of *Lutzomyia longipalpis* (Diptera: Psychodidae) larvae: Characterization of carbohydrases and digestion of microorganisms. J Insect Physiol. 58:1136. doi: 10.1016/j.jinsphys.2012.05.015.

Mury, F.B., da Silva, J. R, Ferreira, L.S., dos Santos, F. B., de Souza-Filho, G.A., de Souza-Neto, J.A., Ribolla, P. E. M., Silva, C. P., do Nascimento, V.V., Machado, O. L. T., Berbert-Molina, M. A. Dansa-Petretski. (2009). Alpha-glucosidase promotes hemozoin formation in a blood-sucking bug: An evolutionary history. PLoS One. 2009;4(9):e6966.

Nishimoto, M., Kubota, M., Tsuji, M., Mori, H., Kimura, A., Matsui, H., Chiba, S. (2001). Purification and substrate specificity of honeybee, *Apis mellifera* L., alpha-glucosidase III. Biosci Biotechnol Biochem. 65:1610. doi: 10.1271/bbb.65.1610

Oryan, A., Akbari, M. (2016). Worldwide risk factors in leishmaniasis. Asian Pac J Trop Med. 9:925. doi: 10.1016/j.apjtm.2016.06.021

Penna H. A. (1934). Leishmaniose visceral no Brasil. Bras Med 48: 949-950.

Pimenta, P. F. P., Turco, S. J., McConville, M. J., Lawyer, P. G., Perkins, P. V.; Sacks, D. L. (1992). Stage-specific adhesion of *Leishmania* promastigotes to the sandfly midgut. Science. 256: 1812. doi: 10.1126/science.1615326

776 Profeta, G. S., Pereira, J. A. S., Costa, S.G., Azambuja, P., Garcia, E. S., Moraes, C. S.,
777 Genta, F.A. (2017). Standardization of a continuous assay for glycosidases and its use
778 for screening insect gut samples at individual and populational levels. Front Physiol.
779 8:308. doi: 10.3389/fphys.2017.00308

780 Raabo, E., Terkildsen, T. C. (1960). On the enzymatic determination of blood glucose.
781 Scand J Clin Lab Invest. 12:402. doi: 10.3109/00365516009065404

782 Ramalho-Ortigão, J. M., Kamhawi, S., Rowton, E. D., Ribeiro, J. M., Valenzuela, J. G.
783 (2003). Cloning and characterization of trypsin- and chymotrypsin-like proteases from
784 the midgut of the sand fly vector *Phlebotomus papatasi*. Insect Biochem Mol Biol.
785 33:163. doi: 10.1016/S0965-1748(02)00187-X

786 Samie, M., Wallbanks, K. R., Moore, J. S., Molyneux, D. H. (1990). Glycosidase activity
787 in the sandfly *Phlebotomus papatasi*. Comp Biochem Physiol B – Biochem Mol Biol. 96:
788 577. doi: 10.1016/0305-0491(90)90059-3

789 Sant'anna, M. R., Diaz-Albiter, H., Mubaraki, M., Dillon, R. J., Bates, P. A. (2009).
790 Inhibition of trypsin expression in *Lutzomyia longipalpis* using RNAi enhances the
791 survival of Leishmania. Parasit Vectors. 2:62. doi: 10.1186/1756-3305-2-62

792 Santos, V. C., Araujo, R. N., Machado, L. A., Pereira, M. H., Gontijo, N. F. (2008). The
793 physiology of the midgut of *Lutzomyia longipalpis* (Lutz and Neiva 1912): pH in different
794 physiological conditions and mechanisms involved in its control. J Exp Biol. 211:2792.
795 doi: 10.1242/jeb.019836.

796 Schlein, Y. (1986). Sandfly diet and *Leishmania*. Parasitology Today. 2:175-177.
797 doi: 10.1016/0169-4758(86)90150-X.

798 Segel, I. H. (1975). Enzyme kinetics: Behavior and analysis of rapid equilibrium and
799 steady-state enzyme systems. New York: John Willey & Sons.

800 Sherlock, I. A. (2003). "A importância dos flebotomíneos" in Flebotomíneos do Brasil,
801 eds E. F. Rangel, R. Lainson (Rio de Janeiro: Editora Fiocruz), 15–22.

802 Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano,
803 M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., Klenk, D. C. (1985). Measurement of
804 protein using bicinchoninic acid. Anal Biochem. 150:76. doi: 10.1016/0003-
805 2697(85)90442-7

806 Soares, R. P. P., Turco, S. J. (2003). *Lutzomyia longipalpis* (Diptera: Psychodidae:
807 Phlebotominae): a review. An Acad Bras Cienc. 75:301. doi: 10.1590/S0001-
808 37652003000300005.

809 Souza-Neto, J. A., Machado, F. P., Lima, J. B., Valle, D., Ribolla, P. E. (2007). Sugar
810 digestion in mosquitoes: identification and characterization of three midgut alpha-
811 glucosidases of the neo-tropical malaria vector Anopheles aquasalis (Diptera: Culicidae).
812 Comp Biochem Physiol A Mol Integr Physiol. 147:993. doi: 10.1016/j.cbpa.2007.03.008

Terra, W. R., Ferreira, C. (1994). Insect digestive enzymes: properties, compartmentalization and function. Comp Biochem Physiol. 109:1. doi: 10.1016/0305-0491(94)90141-4

Terra, W. R., and Ferreira, C. (2005). "Biochemistry of digestion," in Comprehensive Molecular Insect Science. Biochemistry and Molecular Biology, Vol. 4, eds L. I. Gilbert, K. Iatrov, and S. S. Gill (Oxford: Elsevier), 171–224.

Trincone A. (2011) Marine Biocatalysts: Enzymatic Features and Applications. Mar Drugs. 2011; 9(4): 478–499.

Vianna, G. (1911). Söbre uma nova especie de *Leishmania* (Nota Preliminar). O Brazil Médico 25, 411

Vianna, G. (1914). Parasitismo da celula muscular lisa pela *Leishmania brasiliensis*. Memórias do Instituto Oswaldo Cruz. 6: 40-42.

Wongchawalit, J., Yamamoto, T., Nakai, H., Kim, Y. M., Sato, N., Nishimoto, M., Okuyama, M., Mori, H., Saji, O., Chanchao, C., Wongsiri, S., Surarit, R., Svasti, J., Chiba, S., Kimura, A. (2006) Purification and characterization of alpha-glucosidase I from Japanese honeybee (*Apis cerana japonica*) and molecular cloning of its cDNA. Biosci Biotechnol Biochem. 70:2889. doi: 10.1271/bbb.60302

WHO. (2017) WHO: Weekly epidemiological record: Global leishmaniasis update, 2006-2015, a turning point in leishmaniasis surveillance. World Health organization, v. 92, n. 38, p. 557–572, 2017.

Table 1: pH optimum of α-glucosidases obtained from different samples of *L. longipalpis* females under different physiological conditions, using MUαGlu or sucrose as substrates.

	MUαGlu			Sucrose		
Sample	**Non Fed**	**Sucrose Fed**	**Blood Fed**	**Non Fed**	**Sucrose Fed**	**Blood Fed**
Midgut Content	7.5 ± 0.2	7.4 ± 0.3	7.4 ± 0.2	6.4 ± 0.9	6.3 ± 0.6	6.7 ± 0.2
Midgut Tissue	7.0 ± 0.2	6.8 ± 0.1	7.3 ± 0.2	6.3 ± 0.4	6.5 ± 0.1	6.5 ± 0.2
Carcass	7.4 ± 0.2	7.4 ± 0.1	7.7 ± 0.2	6.2 ± 0.2	6.4 ± 0.2	6.3 ± 0.1

Table 2: Kinetic parameters of α-glucosidase activities obtained from different samples of *L. longipalpis* females under different physiological conditions, using sucrose as substrate.

	Sucrose											
	Non Fed				Sucrose Fed				Blood Fed			
Sample	$V_{max\ app}$ **(mU/Fleb)**	**Protein (µg/Fleb)**	$K_{M\ app}$ **(mM)**	K_I **(mM)**	$V_{max\ app}$ **(mU/Fleb)**	**Protein (µg/Fleb)**	$K_{M\ app}$ **(mM)**	K_I **(mM)**	$V_{max\ app}$ **(mU/Fleb)**	**Protein (µg/Fleb)**	$K_{M\ app}$ **(mM)**	K_I **(mM)**
Crop	4.6 ± 0.6	2.5 ± 0.5	210 ± 40	150 ± 30	0.359 ± 0.007	2.7 ± 0.2	11 ± 1	-	0.58 ± 0.06	5.6 ± 0.7	110 ± 20	400 ± 100
Midgut content	5.5 ± 0.1	5.3 ± 0	44 ± 7	-	2.22 ± 0.07	5.7 ± 0.6	51 ± 6	-	11.1 ± 0.1	258 ± 7	20 ± 1	-
Midgut Tissue	40 ± 10	10 ± 1	170 ± 60	180 ± 60	150 ± 20	8 ± 1	800 ± 100	39 ± 7	50 ± 10	337 ± 3	400 ± 100	70 ± 20

Table 3: Kinetic parameters of α-glucosidase activities obtained from different samples of *L. longipalpis* females under different physiological conditions, using MUαGlu as substrate.

Sample	MUαGlu								
	Non Fed			Sucrose Fed			Blood Fed		
	$V_{max\ app}$ (µU/Fleb)	Protein (µg/Fleb)	$K_{M\ app}$ (mM)	$V_{max\ app}$ (µU/Fleb)	Protein (µg/Fleb)	$K_{M\ app}$ (mM)	$V_{max\ app}$ (µU/Fleb)	Protein (µg/Fleb)	$K_{M\ app}$ (mM)
Crop	90 ± 7	2.5 ± 0.5	2.1 ± 0.2	52 ± 2	2.7 ± 0.2	0.48 ± 0.06	16 ± 1	5.6 ± 0.7	0.9 ± 0.1
Midgut content	390 ± 20	2.76 ± 0.04	0.7 ± 0.1	266 ± 7	3.8 ± 0.3	0.37 ± 0.03	830 ± 10	258 ± 7	0.41 ± 0.02
Midgut Tissue	3000 ± 200	10 ± 1	4.7 ± 0.4	1400 ± 100	6.2 ± 0.5	1.1 ± 0.1	1150 ± 40	337 ± 3	0.50 ± 0.05

Table 4: Recovery of activity from peaks obtained after Gel filtration chromatographies (Superdex 200 column, AKTA-FPLC) of the midgut contents or midgut tissue samples of *L. longipalpis* females. Note the elution volumes of different peaks depending on tissue and feeding condition tested.

			Elution Volume (mL)	Recovery MUαGlu Activity (%)	Recovery Sucrose Activity (%)
Midgut Content	**Non Fed**	Peak 1	7	49	15
		Peak 2	13	-	6
		Peak 3	13,5	351	-
		Peak 4	14,5	-	5
	Sugar Fed	Peak 1	7,5	48	13
		Peak 2	11	277	-
		Peak 3	14	173	-
	Blood Fed	Peak 1	7	30	16
		Peak 2	12,5	123	19
		Peak 3	14,5	-	10
Midgut Tissue	**Non Fed**	Peak 1	11,5	198	80
		Peak 2	13,5	35	-
		Peak 3	17,5	-	9
	Sugar Fed	Peak 1	11	201	-
		Peak 2	11,5	-	47
		Peak 3	13,5	38	-
		Peak 4	17,5	-	4
	Blood Fed	Peak 1	11	115	291
		Peak 2	12,5	35	-
		Peak 3	17	-	43

Figure 1.TIF

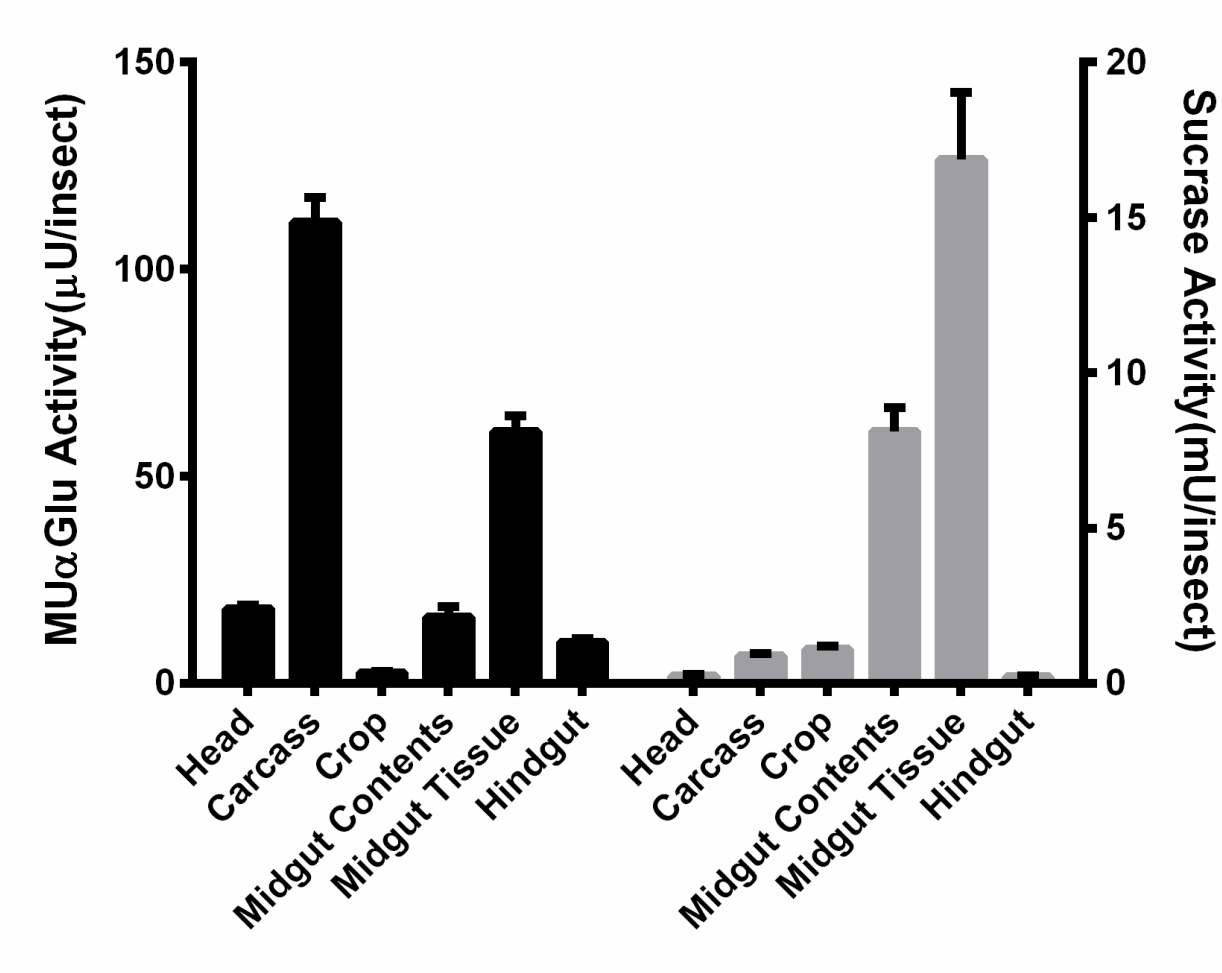


Figure 1: Comparison of α-glucosidase activity using as substrate MUαGlu (black bars) or Sucrose (grey bars) in different tissues of recently emerged (0-3 h) *L. longipalpis* females. The results are the mean ± SEM of three biological replicates each with n = 12

Figure 2.TIF

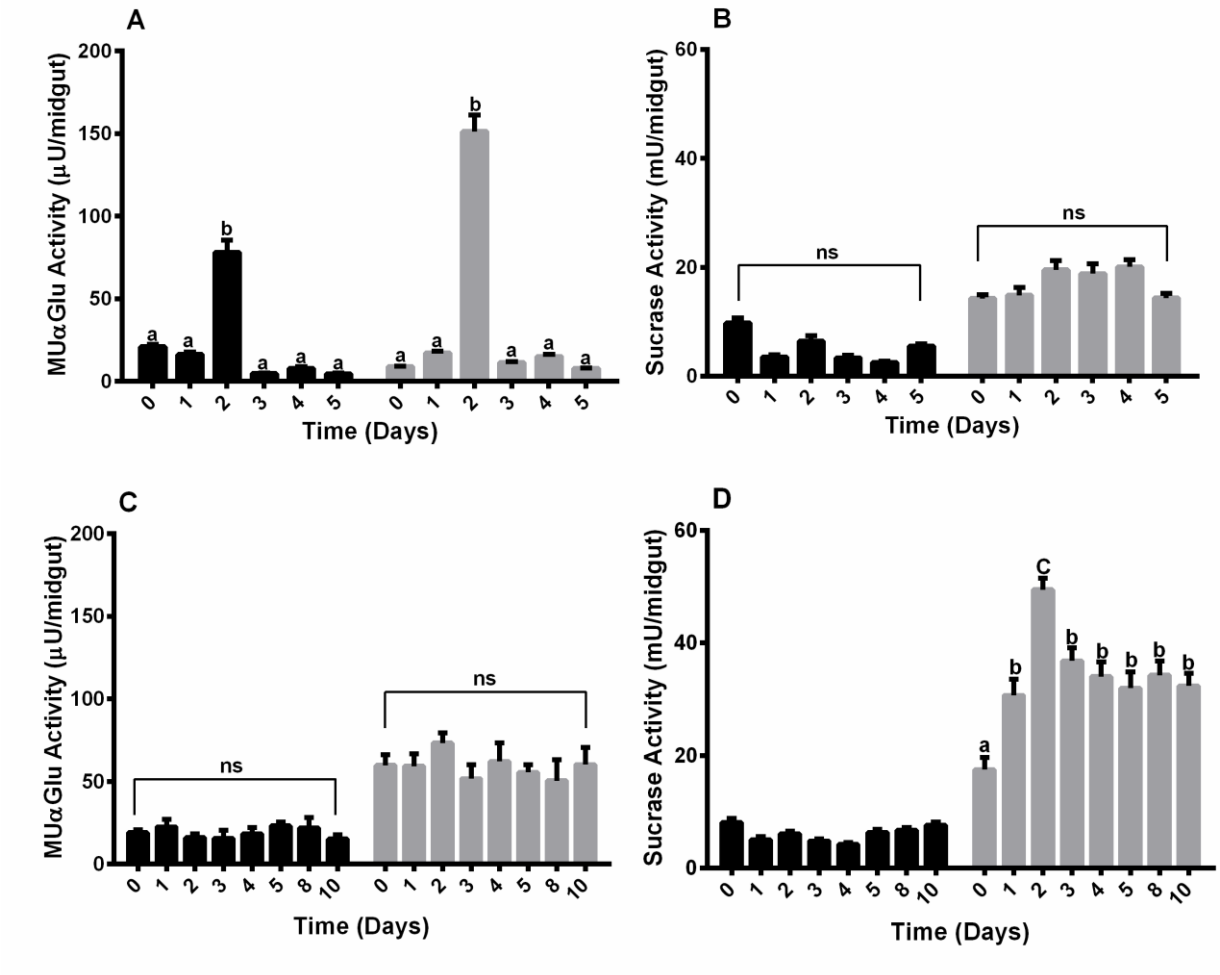


Figure 2: α-glucosidase activity in different tissues of *L. longipalpis* females feeding on water (**A, B**) or 1.2 M sucrose (**C, D**). For each feeding condition, the midgut contents (**black bars**) and midgut tissues (**grey bars**) were analyzed. Activity was determined using MUαGlu (**A, C**) or sucrose (**B, D**) as substrates. Numbers in the *x*-axis represent the time of exposure to water or sucrose after adult emergence. The results are the mean ± SEM of three biological replicates each with n = 15. One-way ANOVA, followed by Tukey multiple comparison test. Different letters indicate statistically significant different groups, $p < 0.05$. **ns**: non-significant difference.

Figure 3.TIF

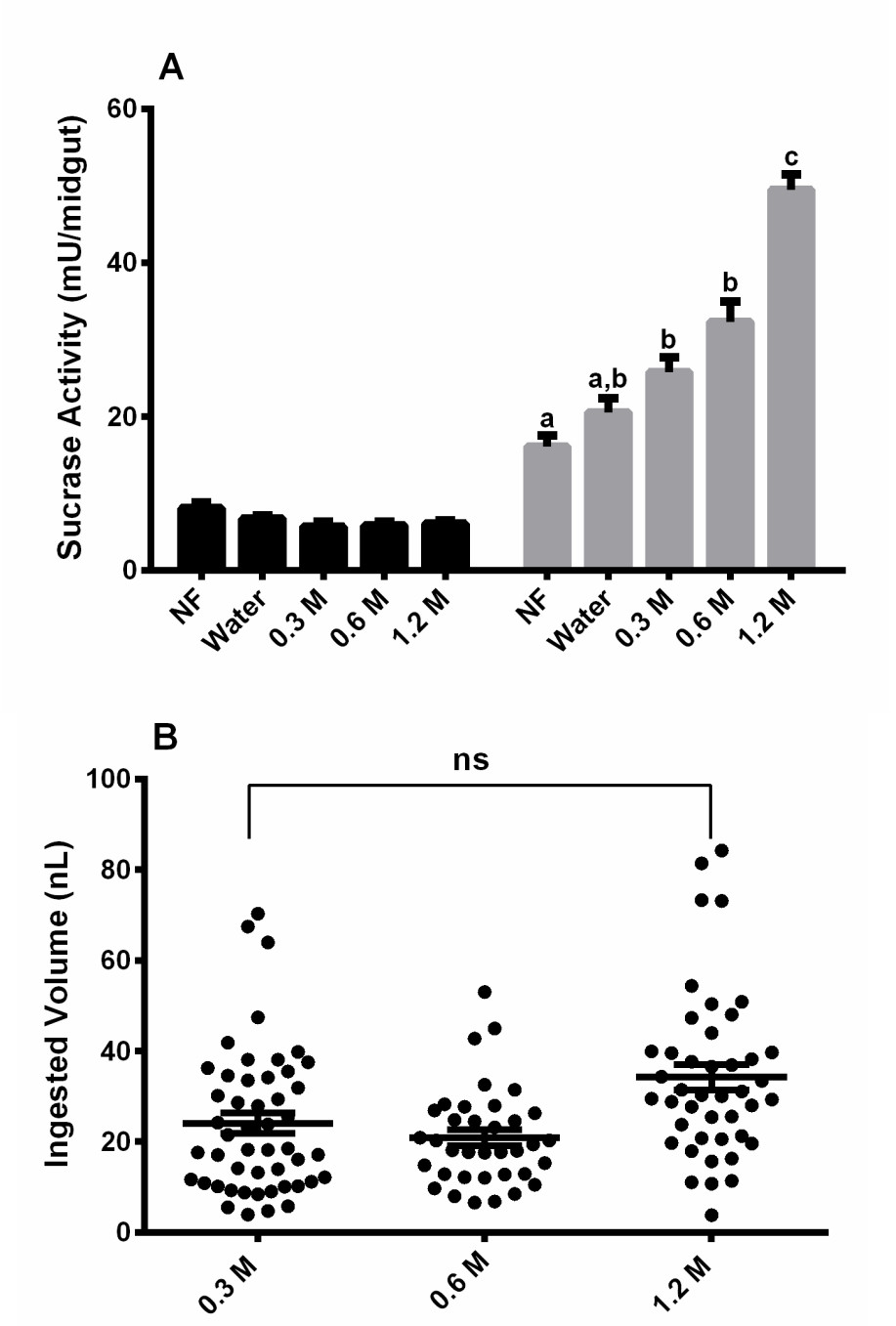


Figure 3: Effect of feeding *L. longipalpis* females with different sucrose concentrations for 2 days, on the α-glucosidase activity. For each feeding condition midgut contents (**black bars**) and midgut tissues (**grey bars**) were analyzed (**A**). The volume ingested in each sucrose concentration was also examined (**B**). The *x*-axis represents the concentration of sugar used to feed the females. Activity was measured using sucrose as substrate. **NF**- Recently emerged females (0 -3 h), **Water** – females fed with only water for 2 days. The results are the mean ± SEM of three biological replicates each with n = 15. One-way ANOVA, followed by Tukey multiple comparison test. Different letters indicate statistically significant different groups, $p < 0.05$. **ns**: non-significant difference.

Figure 4.TIF

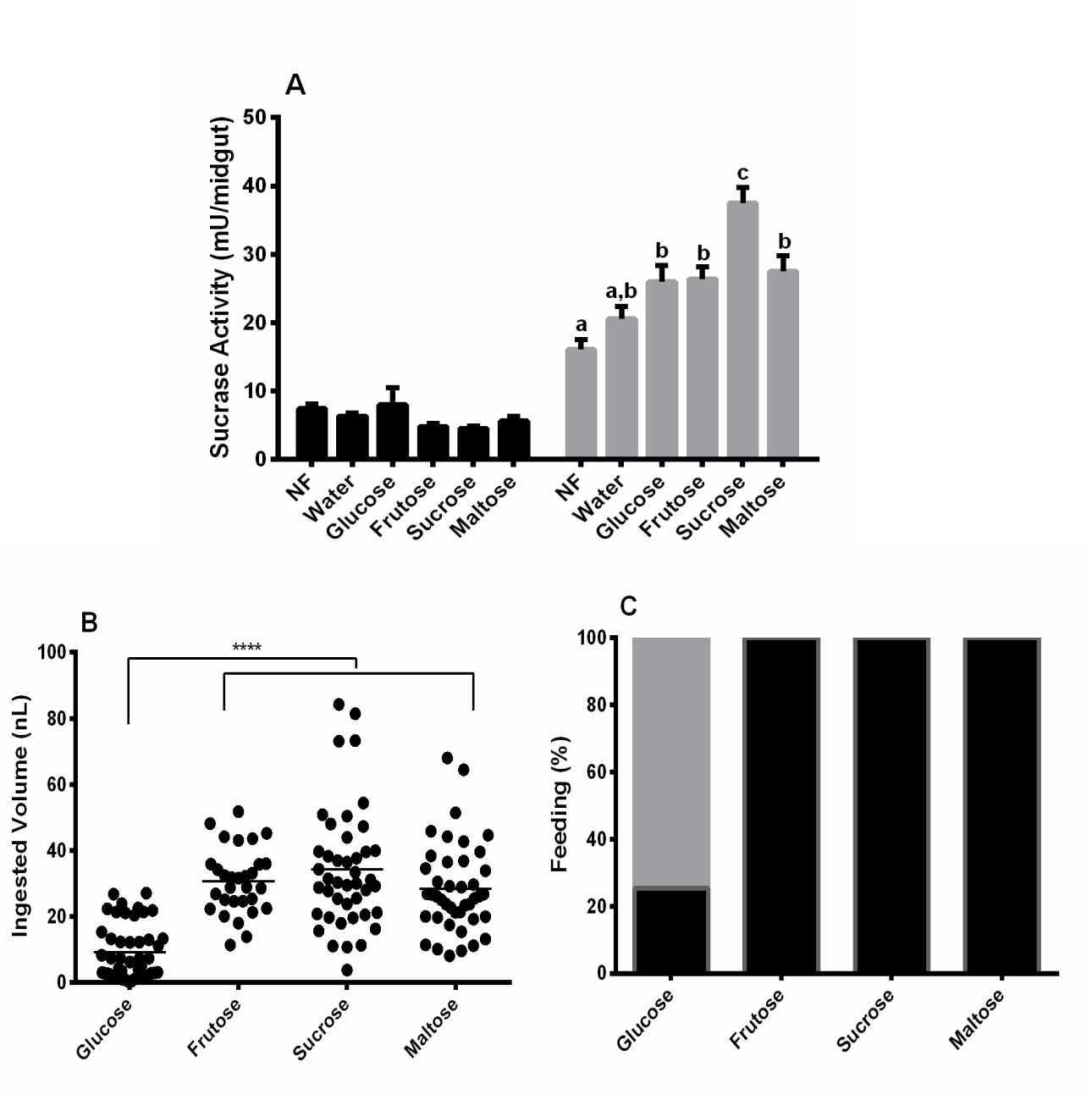


Figure 4: Effect of feeding *L. longipalpis* females with different 1.2 M mono and disaccharides for 2 days on α-glucosidase activity. For each feeding condition midgut contents (**black bars**) and midgut tissues (**grey bars**) were analyzed (**A**). The volume of mono or disaccharide solution ingested was also examined (**B**). Representation of the percentage of females that ingested the mono or disaccharide solution. Fed females **(black),** unfed females **(Grey)** (**C**). Activity was measured using sucrose as substrate. **NF**- Recently emerged females (0 -3 h), **Water** – females fed only with water for 2 days. The results are the mean ± SEM of three biological replicates each with n = 15. One-way ANOVA, followed by Tukey multiple comparison test. Different letters indicate statistically significant different groups, $p < 0.05$. **** indicates statistically significant differences in ingested volumes for different sugars, $p < 0.001$.

Figure 5.TIF

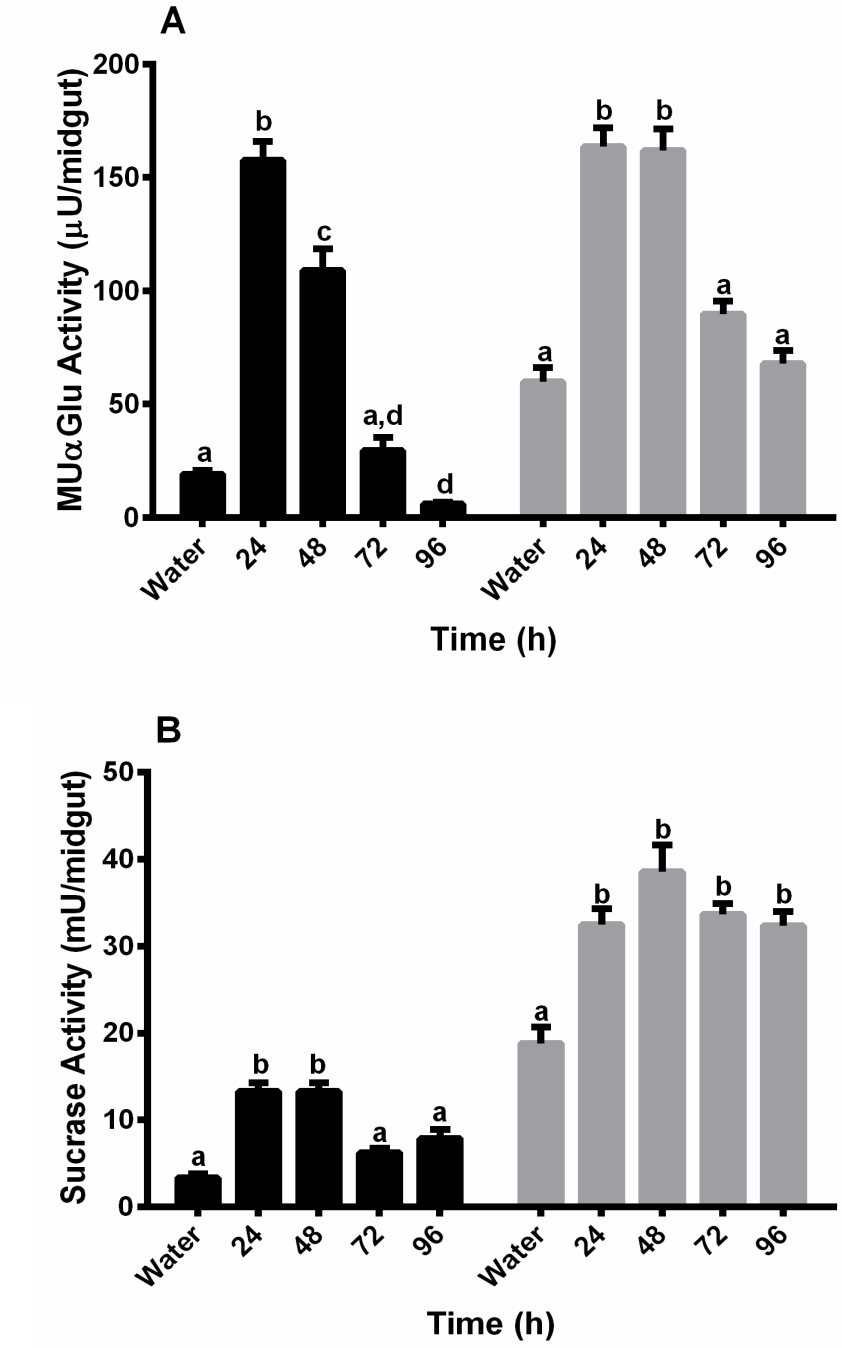


Figure 5: α-glucosidase activity in different tissues of *L. longipalpis* females after feeding with blood. Activity was determined in midgut contents (**black bars**) and midgut tissues (**grey bars**) using MUαGlu (**A**) or sucrose (**B)** as substrates. Numbers in the *x*-axis represent the time elapsed after ingestion of blood. **Water** – females fed only with water for 3 days. The results are the mean ± SEM of three biological replicates each with n = 15. One-way ANOVA, followed by Tukey multiple comparison test. Different letters indicate statistically significant different groups, $p < 0.05$.

Figure 6.TIF

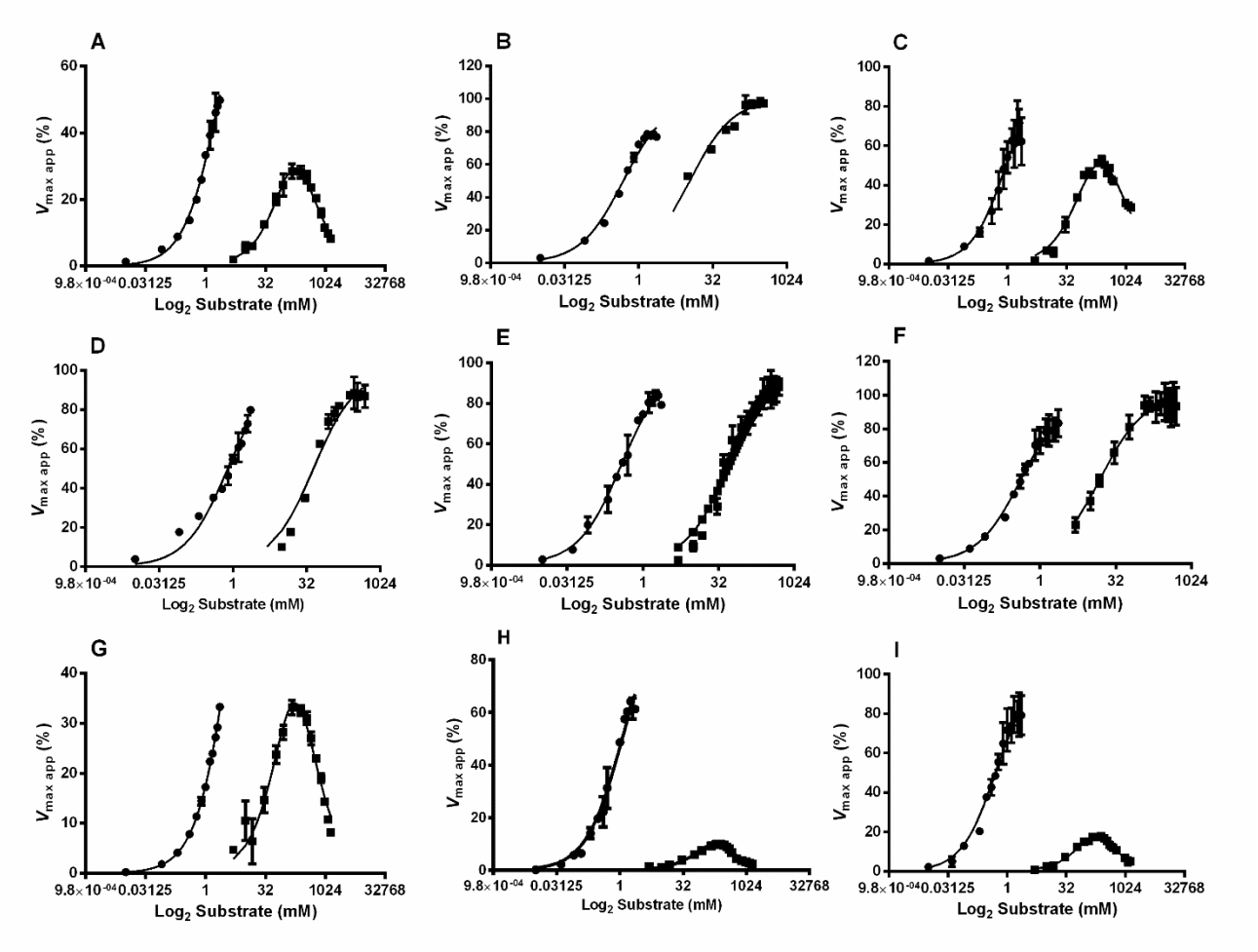


Figure 6: Effect of the substrate concentration on α-glucosidase activity in different feeding conditions and tissues of *L. longipalpis* females. Samples were obtained from recently emerged females (0-3 h) (**A, D, G**), females fed with 1.2 M sucrose for 2 days (**B, E, H**) and females after 2 days of blood feeding (**C, F, I**). For each feeding condition, crop (**A, B, C**), midgut contents (**D, E, F**) and midgut tissue homogenates (**G, H, I**) were analyzed. MUαGlu (dark circles) and Sucrose (dark squares) were used as substrates. The points are experimental results and lines were calculated based on constants computed with the software Graphit using the least-squares method. Note the inhibition by excess of sucrose as substrate. $\boldsymbol{V_{max\ app}}$ (%) was calculated using experimental results compared to values found using the Graphit software (described in Table 2 and 3).

Figure 7.TIF

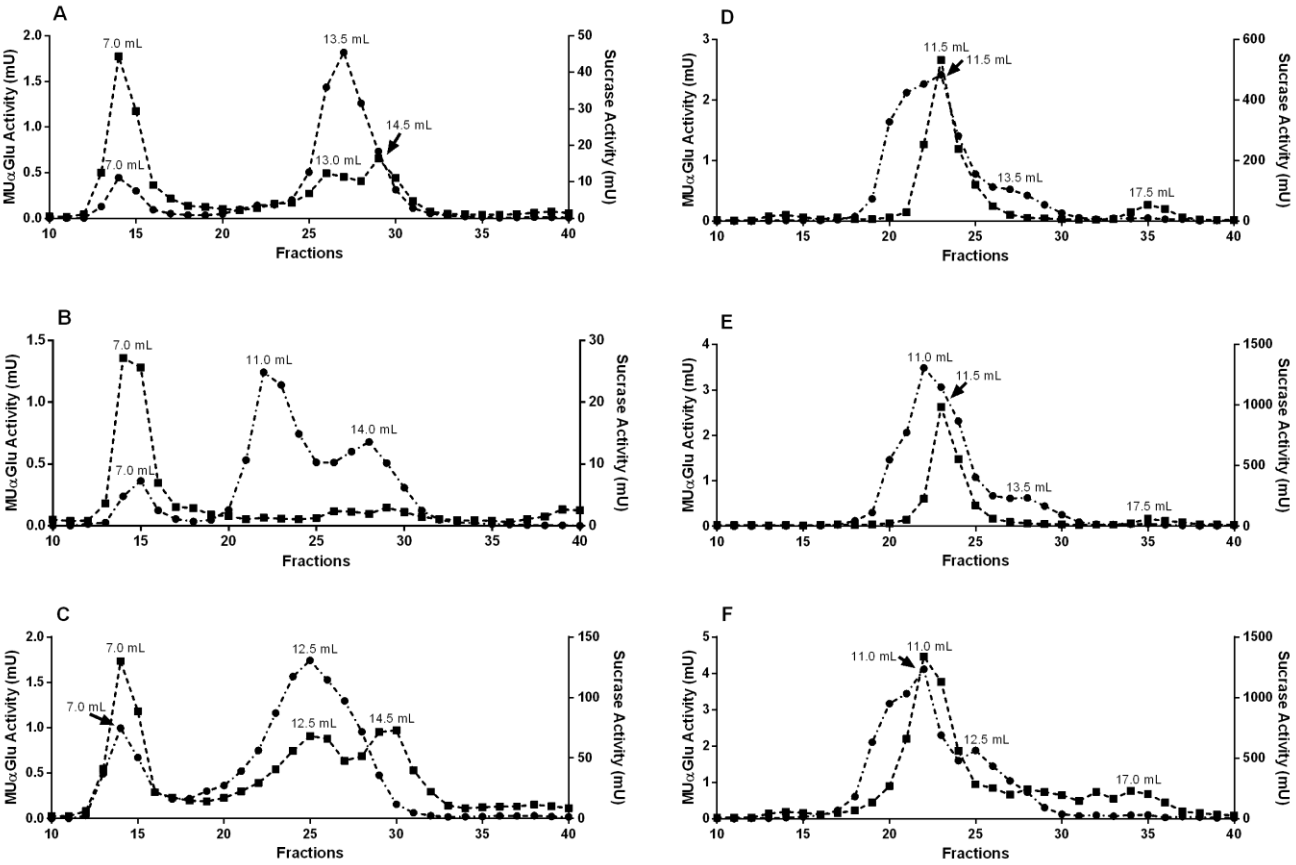
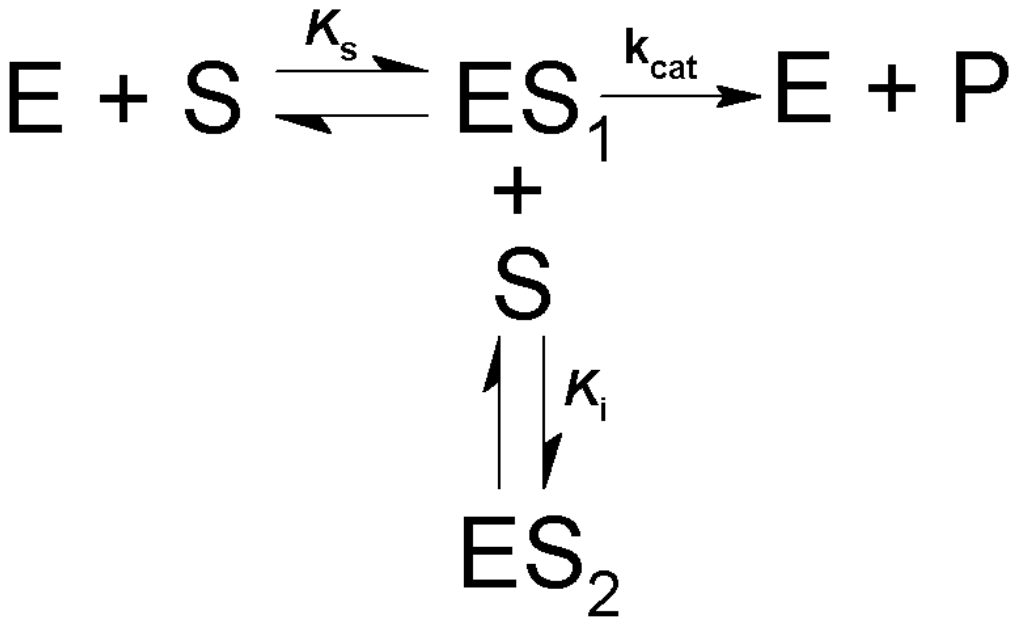


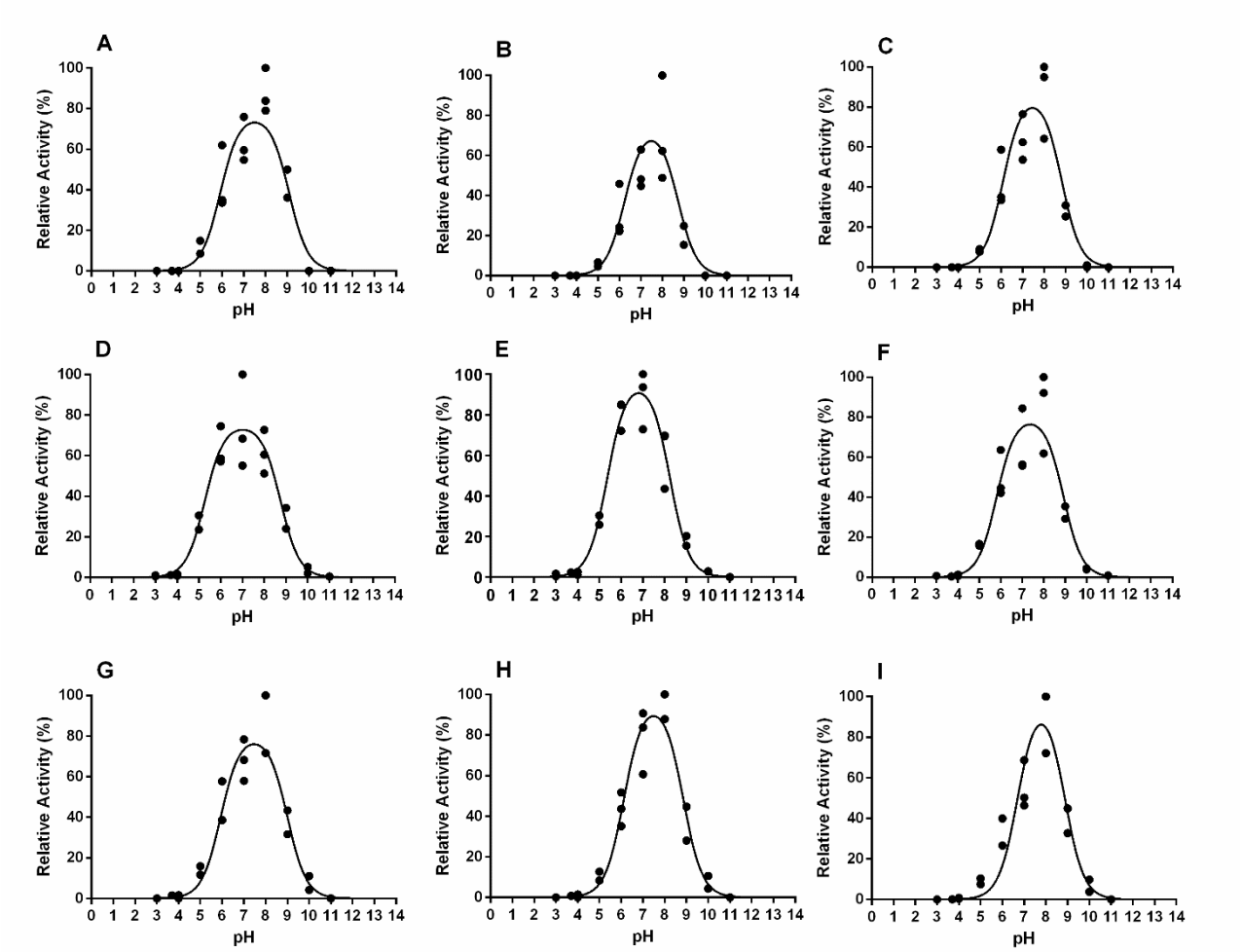
Figure 7: Gel filtration chromatographies (Superdex 200 column, AKTA-FPLC) of the midgut contents (**A, B, C**) or midgut tissue samples (**D, E, F**) of *L. longipalpis* females. Samples were obtained from recently emerged females (0 – 3 h) (**A, D**), females fed with 1.2 M sucrose for 2 days (**B, E**) and females after 2 days of blood feeding (**C, F**). Activity was monitored using MUαGlu (dark circles) and sucrose (dark squares) as substrates. The elution volumes of the activity peaks are indicated in the figure.

Scheme 1

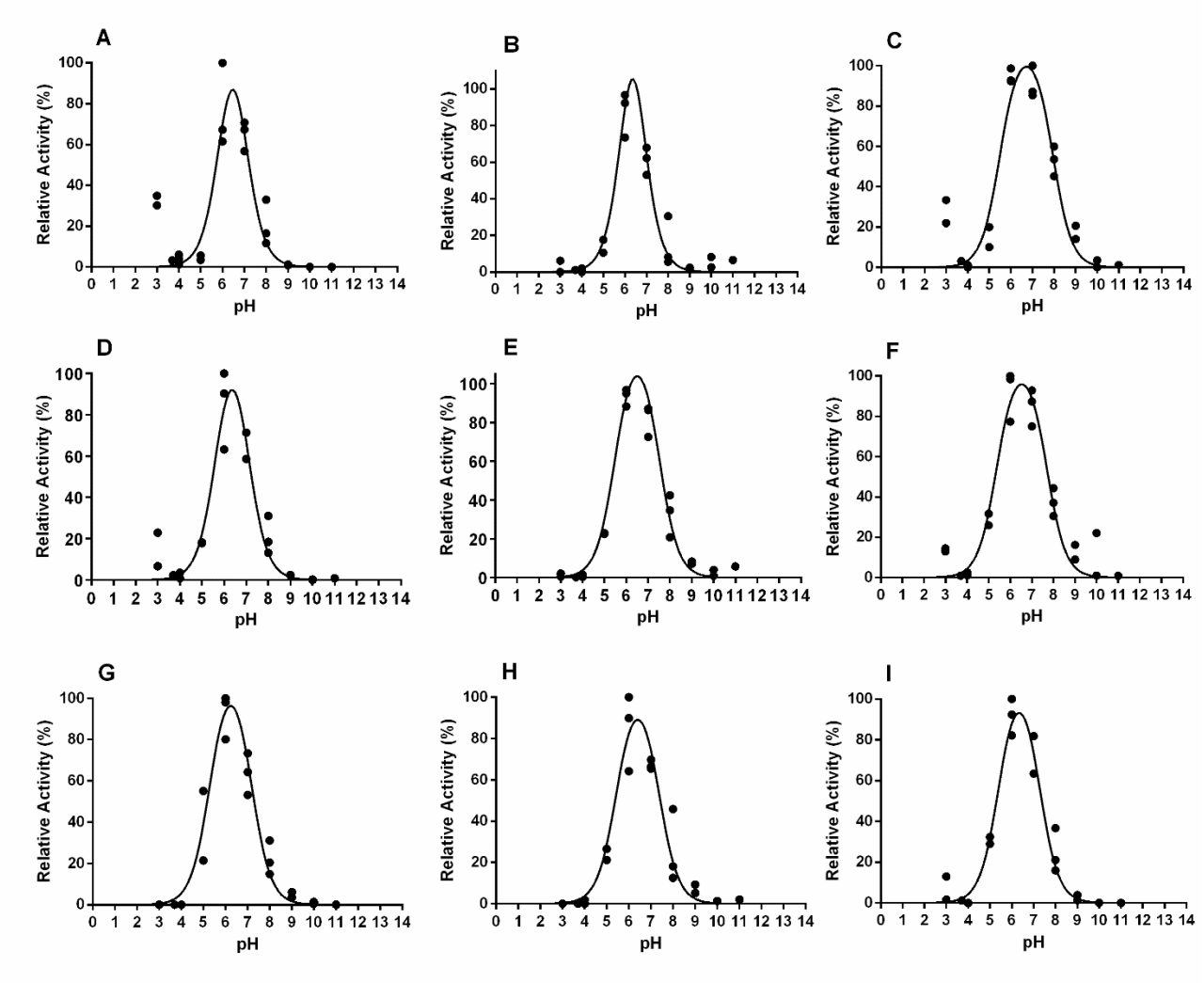


Scheme 1 Model for simple competitive substrate inhibition. According to this, at high concentrations a second molecule of substrate (S) binds to the enzyme-substrate complex (ES_1), forming an inactive ternary complex (ES_2).

1 Additional Files



Supplementary Figure 1: Effect of pH on alpha-glucosidase activity using MUαGlu as substrate in different feeding conditions and tissues of *L. longipalpis* females. Samples were obtained from recently emerged females (0-3 h) (**A, D, G**), females fed with sucrose 1.2 M for 2 days (**B, E, H**) and females after 2 days of blood feeding (**C, F, I**). For each feeding condition midgut content fraction (**A, B, C**), midgut tissue fraction (**D, E, F**) and carcass (**G, H, I**) were analyzed. The points are experimental results and curves were calculated based on constants computed with the software Graphit using equations assuming two ionizable groups.



Supplementary Figure 2: Effect of pH on alpha-glucosidase activity using sucrose as substrate in different feeding conditions and tissues of *L. longipalpis* females. Samples were obtained from recently emerged females (0-3 h) (**A, D, G**), females fed with sucrose 1.2 M for 2 days (**B, E, H**) and females after 2 days of blood feeding (**C, F, I**). For each feeding condition, midgut content fraction (**A, B, C**), midgut tissue fraction (**D, E, F**) and carcass (**G, H, I**) were analyzed. The points are experimental results and curves were calculated based on constants computed with the software Graphit using equations assuming two ionizable groups.

TRABALHO 3

Study of glycoside hydrolase families 13 and 31 reveals expansion and diversification of α-amylase genes in the phlebotomine *Lutzomyia longipalpis*, and modulation of sand fly glycosidase activities by *Leishmania* infection

Samara G. da Costa, Paul Bates, Rod Dillon and Fernando Ariel Genta

Submetido ao periódico PLOS Neglected Tropical Diseases

PLOS Neglected Tropical Diseases

Study of glycoside hydrolase families 13 and 31 reveals expansion and diversification of α-amylase genes in the phlebotomine Lutzomyia longipalpis, and modulation of sand fly glycosidase activities by Leishmania infection

--Manuscript Draft--

Manuscript Number:	PNTD-D-18-01598
Full Title:	Study of glycoside hydrolase families 13 and 31 reveals expansion and diversification of α-amylase genes in the phlebotomine Lutzomyia longipalpis, and modulation of sand fly glycosidase activities by Leishmania infection
Short Title:	Study of glycoside hydrolase families 13 and 31 in Lutzomyia longipalpis
Article Type:	Research Article
Keywords:	Lutzomyia longipalpis; sugar; sucrase; maltase; α-glucosidase; α-amylase; glycoside hydrolase
Corresponding Author:	Fernando Ariel Genta, Ph.D. FIOCRUZ Rio de Janeiro, RJ BRAZIL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	FIOCRUZ
Corresponding Author's Secondary Institution:	
First Author:	Samara Graciane da Costa
First Author Secondary Information:	
Order of Authors:	Samara Graciane da Costa
	Paul Bates
	Rod Dillon
	Fernando Ariel Genta, Ph.D.
Order of Authors Secondary Information:	
Abstract:	Background. Sugar-rich food sources are essential for sand flies to meet their energy demands achieving longer survival. The digestion of carbohydrates from food is mainly realized by glycoside hydrolases (GH). Methodology/principal findings. In this work, we describe genes coding for enzymes involved in various aspects of sugar metabolism, as carbohydrate digestion, storage and mobilization of glycogen reserves, proteins involved in transport, N-glycosylation quality control, and other members with a putative function in the regulation of myogenesis. All these proteins are representatives of GH13 and GH31 families in L. longipalpis, and their functions seem to be conserved in this species. Most of the enzymes seem to be active with conserved consense sequences, including the expected catalytic residues. α-amylases also demonstrated presence of calcium and chloride binding sites. We observed that in L. longipalpis there was an expansion in the alpha-amylase genes with an organization in two clusters, while for maltases a retraction in the number of genes occurred. This expansion is probably related to the specialization of these proteins for different substrates, which might be correlated to the diversity of plant food available in the natural habitat of L. longipalpis. Maltase genes expression is higher in blood-fed females, suggesting a role in blood digestion. Besides that, in blood-fed females infected with L. mexicana, these genes were also modulated. Conclusions/significance. Glycoside Hydrolases from families 13 and 31 are essential for the metabolism of L. longipalpis, and GH13 enzymes seem to be involved in the interactions between sand flies and Leishmania.
Suggested Reviewers:	Dr. Matt Rogers London School of Hygiene and Tropical Medicine

	matthew.rogers@ishtm.ac.uk
	Dr. Omar Hamarsheh Al-Quds University ohamarsheh@gmail.com
	Dr. Nelder Gontijo Universidade Federal de Minas Gerais nelder@icb.ufmg.br
	Dra. Mary Ann MacDowell Notre Dame University mcdowell.11@nd.edu
	Dr. Petr Volf Charles University volf@cesnet.cz
	Dr. Alvaro Acosta Serrano Liverpool School of Tropical Medicine alvaro.acosta-serrano@istmed.ac.uk
Opposed Reviewers:	
Additional Information:	
Question	**Response**
Financial Disclosure Enter a financial disclosure statement that describes the sources of funding for the work included in this submission. Review the submission guidelines for detailed requirements. View published research articles from *PLOS NTDs* for specific examples. This statement is required for submission and **will appear in the published article** if the submission is accepted. Please make sure it is accurate.	This work was supported by the National Counsel of Technological and Scientific Development (CNPq), the Research Support Foundation of the State of Rio de Janeiro (FAPERJ), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Oswaldo Cruz Foundation (FIOCRUZ). SGC received a sandwich Ph.D. grant from Science without Borders (CNPq/CAPES).The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Study of glycoside hydrolase families 13 and 31 reveals expansion and diversification of α-amylase genes in the phlebotomine *Lutzomyia longipalpis*, and modulation of sand fly glycosidase activities by *Leishmania* infection

Samara G. da Costa[1], Paul Bates[2], Rod Dillon[2,3] and Fernando Ariel Genta[1,3,*]

[1]Laboratory of Insect Biochemistry and Physiology, Oswaldo Cruz Institute, FIOCRUZ, Rio de Janeiro, RJ, Brazil

[2]Faculty of Health and Medicine, Division of Biomedical and Life Sciences, Lancaster University, Lancaster, United Kingdom

[3]National Institute of Science and Technology for Molecular Entomology, Rio de Janeiro, RJ, Brazil

*Corresponding Author:

E-mail: genta@ioc.fiocruz.br or gentafernando@gmail.com

Abstract

Background. Sugar-rich food sources are essential for sand flies to meet their energy demands achieving longer survival. The digestion of carbohydrates from food is mainly accomplished by glycoside hydrolases (GH). **Methodology/principal findings.** In this work, we describe genes coding for enzymes involved in various aspects of sugar metabolism, as carbohydrate digestion, storage and mobilization of glycogen reserves, proteins involved in transport, N-glycosylation quality control, and other members with a putative function in the regulation of myogenesis. All these proteins are representatives of GH13 and GH31 families in *L. longipalpis*, and their functions seem to be conserved in this species. Most of the enzymes might to be active since they display conserved consense sequences, including the expected catalytic residues. α-amylases also demonstrated presence of calcium and chloride binding sites. We observed that in *L. longipalpis* there was an expansion in the α-amylase genes with an organization in two clusters, while for maltases a retraction in the number of genes occurred. This expansion is probably related to the specialization of these proteins for different substrates, which might be correlated to the diversity of plant food available in the natural habitats of *L. longipalpis*. Maltase gene expression in blood-fed females is higher, suggesting a role in blood digestion. In addition, in blood-fed females infected with *L. mexicana*, these genes were also modulated. **Conclusions/significance.** Glycoside Hydrolases from families 13 and 31 are essential for the metabolism of *L. longipalpis*, and GH13 enzymes seem to be involved in the interactions between sand flies and *Leishmania*.

Keywords: *Lutzomyia longipalpis*, sugar, sucrase, maltase, α-glucosidase, α-amylase, glycoside hydrolase

Author Summary

α-amylases and maltases are key enzymes in the digestion of carbohydrates. In insects, they belong to glycoside hydrolases families 13 (GHF13) and 31 (GHF31). We screened the genome of the phlebotomine sand fly *Lutzomyia longipalpis* and 21 and 6 genes belonging to GHF13 and GHF31 were found, respectively. These genes showed specific sequence signatures and expression patterns, suggesting their involvement in starch, sucrose and blood digestion, glycogen metabolism, amino acid transport and myogenesis regulation. We observed an amplification in the α-amylase family that might be related

to a higher diversification of plant food sources in sand flies from New World. Besides that, we described for the first time the modulation of insect carbohydrase activities during a *Leishmania* infection.

Introduction

Phlebotomine sand flies are considered of medical importance, mainly because about 10 % of their known species are vectors of some pathogen-caused disease. Currently, more than 800 sand flies species were identified, 464 are found in the New World and 375 in the Old World (1–3). Among the diseases transmitted by phlebotomines, we can highlight the bartonellosis in South America, caused by the bacterium *Bartonella bacilliformis* and transmitted by *Lutzomyia verrucarum* (4), some arboviruses in North Africa and parts of Asia with the transmission of pappataci virus causing the fever pappataci (*Phlebotomus papatasi*, *P. perfiliewi*, *P. perniciosus*) (5), and the leishmaniasis, a parasitic disease, endemic in 87 countries, with more than 200,000 new cases of the cutaneous form and 20,000 cases of the visceral form reported in 2015 (6). The leishmaniasis are caused by protozoan parasites belonging to the genus *Leishmania*, transmitted through the bite of female sandflies of the genus *Phlebotomus* in the Old World and *Lutzomyia* in the New World (7–9). *Lutzomyia longipalpis* is the natural vector of *Leishmania infantum*, the causative agent of visceral leishmaniasis in Americas (1,10,11).

Adult sandflies have a sugar-rich diet obtained routinely from feeding on plant tissues, nectar flowers and aphid sugars, and females also feed on blood for egg development. Thus, glycosidases play a significant role in the digestion of these insects. Especially the α-amylases and α-glucosidases, since their substrates are important components of plant tissues and secretions. Starch (substrate for amylase) is found in leaves, and sucrose (substrate for α-glucosidase) is the main component of sap, besides other sugars found in honeydew (12,13). These enzymes are an important subset inside the glycoside hydrolase families. Gene sequences of thousands of these enzymes have been described and grouped into families based on sequence similarity. α-glucosidases act upon α-1,4-glucosidic bonds and participates in both sugar hydrolysis, glycoconjugate processing and biosynthesis (14). According to the classification system CAZy (Carbohydrate-Active Enzymes), these are classified in the families GH4, GH13, GH31, GH97, and GH122. In insects, these enzymes are described in families 13 and 31 (15). α-amylases play an

important role in the digestion of polysaccharides, and insects' α-amylases belong to GH13, also called the "α-amylase family".

Enzymes classified in the GH13 and GH31 families share a common tertiary structure in their catalytic domain, the $(\beta/\alpha)_8$-barrel fold. The structure of GH13 members is represented by a conserved central nucleus composed by the catalytic domain A, domain B, which is a variable-length loop located between sheet β3 and α3 helix of the $(\beta/\alpha)_8$-barrel, and domain C, a Greek key motif in the C-terminal position (16–19). The conserved regions (CSRs) described in sheets β2, β3, β4, β5, β7, β8 and in the loop 3, were defined to help in the assignment of correct enzymatic specificity to the members of the α-amylase family. The catalytic triad aspartate, glutamate, and aspartate are localized in the β4, β5, and β7 sheets, respectively (17). Also, a fourth conserved residue was determined as an arginine positioned two residues preceding the catalytic nucleophile (20). Aspartate acts as a catalytic nucleophile and glutamate as the general acid/base, and the second aspartate contributes for the stabilization of the transition state and maintaining the glutamate with the correct protonation for activity (21,22). Insects α-amylases were described as calcium-dependent, and some of them are activated by chloride (23). The conserved region loop3 (localized in Domain B) contains an aspartate residue, which is one of the calcium binding sites (24), and the other conserved residues in this pocket are the asparagine positioned at the β3 sheet and a histidine in the β4 sheet (17,25). The chloride binding site is characterized by the presence of a conserved arginine residue positioned two residues preceding the catalytic nucleophile in the β4 sheet, an asparagine residue positioned two residues preceding the catalytic aspartate in the β7 sheet and an Arg/Lys residue localized inside the variable domain RVMSSY (17,26).

In the GH31, the identified catalytic residues are the catalytic nucleophile aspartate inside the conserved domain WIDMNE, localized in the β4 sheet and the general acid/base residue is an aspartate positioned on the β6 sheet (14,27–29). At the sequencing level, the similarity is not evident between GH13 and GH31 members, but there is the conservation of the catalytic nucleophile aspartate in the β4 sheet, in both families (30). Most members of the GH31 family are multi-domain proteins, but the specific function of the accessory domains is generally unknown. As examples, the galactose mutarotase domain found in neutral α-glucosidases and the TREFOIL domain described in lysosomal α-glucosidases.

The GH13 family is the largest of the glycoside hydrolases families and has a wide variety of enzymes. In general, they act on several substrates by hydrolyzing α-glycosidic

bonds and generating α-anomeric mono- or oligosaccharides or forming α-glycosidic bonds by transglycosylation (Kuriki et al., 1999). There are more than 30 different enzyme specificities described in this family, including representatives of α-1,4-glucan branching enzymes, pullulanases, cyclodextrin glucano transferases, 4-α-gluconotransferases (domain of glycogen debranching enzyme), oligo-α-1,6-glucosidase, amino acid transporters, besides others. Due to the large number of sequences and diversity of specificities for these enzymes, this family has been subdivided into almost 40 subfamilies (Stam et al., 2006). The metazoans α-amylase and α-glucosidase activities were classified in the subfamilies GH13_15 and GH13_17, respectively (Stam et al., 2006).

GH31 family is a diverse family with a range of different hydrolytic activities. Members from this family are capable of cleaving a terminal carbohydrate moiety from substrates that vary considerably in size, from disaccharides to reserve complex polysaccharides such as starch, glycogen or glycoproteins (27). α-glucosidases, α-xylosidases, α-galactosidases, α-mannosidases, 6-α-glucosyltransferases, and 3-α-isomaltosyltransferases, besides others, represent the hydrolytic activities present in the GH31 family. The best-described activities in this group are the α-glucosidases that are involved in the glycogen degradation (lysosomal α-glucosidase) (31), in glycoprotein processing (neutral α-glucosidase) in the endoplasmic reticulum (32,33), and in primary metabolism. During starch and glycogen digestion, the enzymes sucrase-isomaltase and maltase-glucoamylase (MGAM) participate in the hydrolysis of α-1,6 bonds in the oligosaccharides produced (NICHOLS et al., 2003).

In this work, we reveal the presence of genes coding for proteins belonging to the α-amylase family GH13, and to the GH31 family in *L. longipalpis*. Using the genome information available in Vector Base, we compared *L. longipalpis* protein sequences with the proteins described for Diptera from suborder Brachycera (genus *Drosophila*), Nematocera (*Aedes*, *Culex*, *Anopheles*) and *Phlebotomus*. We also evaluated the expression pattern of GH13 and GH31 candidate genes in different tissues under different feeding conditions, and the effect of *L. mexicana* infection in their expression. We identified, after annotation, a total of 21 genes belonging to GH13 family and described as α-amylases, maltases, amino acid transport proteins (heavy chain), 1,4-α-glucan branching enzyme and glycogen debranching enzyme, and GH31 genes described as glycosidases NET37, lysosomal α-glucosidase and neutral α-glucosidase (α subunit). The

enzymes described are involved in sugar metabolism, storage, and mobilization of glycogen, protein transport, N-glycosylation quality control and myogenesis regulation.

Material and Methods

Chemicals

The TRI Reagent® (Cat: T9424) was purchased from Sigma-Aldrich Company (St. Louis, Missouri, USA). SYBR™ Green PCR Master Mix (cat: 4309155) was obtained from ThermoFisher Scientific (Waltham, Massachusetts, United States). QuantiTect Reverse Transcription Kit (Cat: 205310) was obtained from Qiagen (Hilden, Germany). Sheep blood with Alsever's anticoagulant (Cat: SB068) was purchased from TCS Biosciences (Buckingham, United Kingdom). Other reagents used in this work were analytical grade.

Identification and analysis of glycoside hydrolases sequences present in *L. longipalpis* genome

To identify sequences of glycoside hydrolases families 13 and 31 from *L. longipalpis*, known sequences from GH13 and GH31 from different insect species genome were retrieved from Vector Base (http://www.vectorbase.org), Fly base (http://flybase.org/) and NCBI (https://www.ncbi.nlm.nih.gov/). Subsequently, GH sequences obtained were employed to perform ClustalW alignment and HMMER search in the Vector Base database to find similar sequences in *L. longipalpis* (conserved domains on *Lutzomyia longipalpis*, LlonJ1.4, last updated 27 June 2017). Sequences obtained with an E-value smaller than e^{-10} were considered for analysis. The retrieved *L. longipalpis* sequences were annotated by similarity with proteins and conserved domains using the BLASTp (ref-seq protein) and Swiss-Prot/UniProt databases. Manual annotation was performed with the Artemis annotation tool (34) and ClustalW (35). After sequences identification, characteristics like signal peptide (SignalIp) (36), transmembrane domain (THMM) (37), GPI-anchor (Big-PI predictor) (38), protein subcellular localization (DeepLoc) (39), N-glycosylation (NetNGlyc) (40), O-glycosylation (NetoGlyc) (41), molecular mass and isoelectric point (pI) (Compute pI/Mw) (42,43) were predicted. The N-glycosylation sites were not considered if predicted inside transmembrane or cytoplasmic domains of membrane proteins. Genes and proteins representative structures were designed using PROSITE MyDomains tool (https://prosite.expasy.org/mydomains/) (44).

The InterPro tool (45) was used to provide a functional analysis of proteins, by predicting catalytic sites and the calcium binding site in α-amylases. The catalytic sites predicted by InterPro were confirmed based on the conserved regions (CSRs) (17), by comparing the *L. longipalpis* sequences with α-amylase from *Aspergillus oryzae* (Uniprot P0C1B3) (46), maltase from *Apis cerana japonica* (Uniprot A1IHL0) (47), amino acid transport proteins (UniProt Q07837) (48), glucan branching enzyme from *E. coli* (UniProt P07762) (49), glycogen debranching enzyme from *Candida glabrata* (Uniprot Q6FSK0) (50), neutral α-glucosidase from *Homo sapiens* (UniProt O43451), lysosomal α-glucosidase from *Homo sapiens* (UniProt P10253) (51) and glucosidase NET37 from *Homo sapiens* (Uniprot Q6NSJ0) by local alignment using ClustalW tool. This comparison was also used to predict the chloride binding sites in α-amylases.

Phylogenetic Tree

The phylogenetic trees were constructed using GH13 or GH31 family protein sequences from different insect species genome. Sequences from *Aedes aegypti* (LVP_AGWG strain, AaegL5.1 geneset), *Anopheles gambiae* (PEST strain, AgamP4.9 geneset), *Culex quinquefasciatus* (Johannesburg strain, CpipJ2.4 geneset) and *Phlebotomus papatasi* (Israel strain, PpapI1.4 geneset) were retrieved from Vector Base (http://www.vectorbase.org) and sequences from *Drosophila melanogaster* and *Drosophila ananassae* retrieved from Fly base (http://flybase.org/). For *Phlebotomus papatasi*, sequences were annotated as described above at item "Identification and Analysis of glycoside hydrolases sequences present *in L. longipalpis* genome" for *L. longipalpis* sequences. The sequence name, species of origin and identifier are shown in S1 Table and S2 Table.

Unrooted distance neighbor-joining trees showing the phylogeny of the GH13 or GH31 families were constructed using MEGAX software (52), and protein sequences were aligned with Muscle algorithm. The pairwise deletion and p-distance functions were selected, and bootstrap values were set at 10000 replications.

Insects Maintenance

The experiments of gene expression were performed using *L. longipalpis* (from Jacobina, Bahia, Brazil), maintained at Lancaster University (United Kingdom). Insects were kept under standard laboratory conditions of temperature (24 ± 2 °C) and a photoperiod of 8 h of light/16 h of darkness. Adults' sandflies were fed with 70 % (w/v)

autoclaved sucrose solution in cotton wool (unless stated differently in experiments). For oviposition, females were fed with sheep blood containing Alsever's anticoagulant via an artificial apparatus (Hemotek - Discovery Workshops, United Kingdom), at 37 °C for 1 hour using chicken skin as a membrane. Engorged females were transferred to rearing containers with a piece of cotton wool soaked in sugar solution. The eggs were separated from dead females after oviposition. Other rearing conditions were performed as described in (53). For experiments, recently emerged females (0 - 3 hours) were fed with 1.2 M sucrose for 2 days (SF females) with a piece of cotton wool soaked in sugar solution, or were maintained for 3 days with cotton wool soaked in water before feeding on sheep blood. After blood feeding (infected or not), they were maintained with a piece of cotton wool soaked in water for 2 days.

Leishmania mexicana Cultivation and Sandflies Infection

For infections, axenic cultures of amastigotes from *L. mexicana* strain M379 were used. For maintaining the parasites in the amastigote form, they were incubated at 32 °C with Grace's Insect Medium supplemented with 20 % of fetal bovine serum (FBS), BME_1 vitamins, 2 % urine and 25 µg/mL gentamycin sulfate. Parasites with a maximum of 26 passages in the amastigote maintaining medium, were used for sandfly infections (54). Infections were performed with a concentration of 2×10^6 parasites/mL, estimated using Neubauer chambers. Briefly, after centrifugation for 5 min at 2000 X g, the supernatant was removed, and parasites were mixed on sheep blood offered to 4-day old unfed females, as described in section "Insects Maintenance". Unfed females were discarded. Control insects were fed with uninfected blood. After blood feeding, infected and control females were maintained with water for 2 days.

RNA Extraction and cDNA Preparation

Newly hatched females (0-3 hours), females fed with 1.2 M sucrose for 2 days, females fed with blood and infected females were dissected in sterile phosphate buffered saline (PBS) and a pool of 10 midguts or rest of the body (RB) were transferred to polypropylene vials containing 50 µL of TRI Reagent® (Cat. No. T9424, Sigma-Aldrich, Darmstad, Germany) and kept at −80 ^{0}C until further RNA extraction. For infected females, gut was checked with light microscopy for parasite presence.

Total RNA was extracted following the supplier's protocol and quantified using NanoDrop® (NanoDrop Technologies, Wilmington, USA). The cDNA was synthesized

by the Reverse Transcriptase (RT) reaction using the QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. cDNA was quantified using Nanodrop® and normalized to a concentration of 50 ng/µl.

Analysis of gene expression by RT-PCR or RT-qPCR

GHF13 and GHF31 coding sequences identified in the *L. longipalpis* genome were used to design specific oligonucleotides (55), using the Primer3 program (56), Beacon Designer™ Free Edition (http://www.premierbiosoft.com/qpcr/index.html) and mFOLD Software (HTTP://www.idtdna.com/UNAFold?). The list of primers is described in S3 Table. Gene expression was initially analyzed by RT-PCR, and only one gene representative of each subgroup in clade I of the α-amylase group (Fig 3) was analyzed. From other clades, genes LLOJ000566/LLOJ008156 (LlMal1/ LlMal2), LLOJ002257 (LlMal3), LLOJ004841 (LlAamy4) and LLOJ003489 (LlNAglu2), were also analyzed by RT-qPCR. Samples obtained from gut and rest of the body (RB) of recently emerged (NF), sugar and blood-fed females were used for RT-PCR or RT-qPCR analysis, and samples from females infected with *Leishmania mexicana* parasites were used only for the RT-qPCR analysis.

For RT-PCR, each reaction was run in duplicate for each pool of tissues (n=10) and the cDNA samples were used in amplification reactions performed in a total volume of 25 µL containing 2X Biomix Red (Bioline, London, UK), 100 ηg cDNA, and 0.2 µM of each primer. The parameters for PCR were: incubation at 95 ^{0}C for 5 min, followed by 26 cycles of 94 ^{0}C for 30 s, 55 ^{0}C for 40 s, 72 ^{0}C for 1 min, and a final incubation (extension) of 72 ^{0}C for 5 min. Relative expression of glycoside hydrolases was normalized using a housekeeping gene (LLOJ001891, Glyceraldehyde 3-phosphate dehydrogenase, GAPDH). RT-PCR products were analyzed by electrophoresis in a 1 % (w/v) agarose gel containing gel red® (Biotium, Fremont, USA) and glycoside hydrolases expression level was determined by densitometry measurement of bands using the software ImageJ (57).

Real-time quantitative polymerase chain reactions (RT-qPCR) were conducted using a CFX96 Touch™ Real-Time PCR Detection System (BioRad). Each reaction was run in duplicate for each pool of tissues (n=10) and performed in a total volume of 10 µL, containing 2X of SYBR™ Green Master Mix (ThermoFisher Scientific), 20 ηg cDNA and 0.5 µM of each primer. The parameters for PCR were: incubation at 95 ^{0}C for 15 min, followed by 39 cycles of 94 ^{0}C for 15 s, 56 ^{0}C for 30 s, and 72 ^{0}C for 30 sec. As a negative

control, PCR reactions were carried out without cDNA template to assess primer dimer formation or contamination in the reactions. To ensure that only a single PCR product was amplified, analysis of the melting curve was performed. The expression of glycoside hydrolase genes in *L. longipalpis* tissues was quantified by the comparative Ct (ΔCt) method (58) normalized with the GAPDH housekeeping gene.

Statistical Analysis

For Multiple comparisons, one-way ANOVA was used followed by Tukey's multiple comparison tests, and significance was considered when $p < 0.05$. Results are expressed as the means ± S.E.M. All statistical analysis was executed using the Software GraphPad Prism 6.0 (San Diego, California, USA).

Results

Sequence and domain analysis of the GH13 and GH31 family members in the genome of *Lutzomyia longipalpis*

The sequences of glycoside hydrolases (GH13 and GH31) from *Aedes aegypti*, *Culex quinquefasciatus, Anopheles gambiae, Drosophila melanogaster, Drosophila ananassae* were employed to carry out a HMMER search against the *L. longipalpis* genome in the Vector Base platform. A total of 16 sequences belonging to GH13 family (identifiers: LLOJ004838, LLOJ004839, LLOJ004841, LLOJ004880, LLOJ004881, LLOJ004882, LLOJ004885, LLOJ005909, LLOJ000566, LLOJ002257, LLOJ008156, LLOJ006803, LLOJ008629, LLOJ005533, LLOJ008212) and 6 sequences from GH31 (identifiers: LLOJ001847, LLOJ000840, LLOJ001881, LLOJ001882, LLOJ006451, LLOJ003489) were retrieved. Proteins were identified by similarity to other protein sequences and to conserved domains using the BLASTp (ref-seq protein) and Swiss-Prot/UniProt databases according to their best hit.

Results indicated that only 3 sequences from GH13 (LLOJ004841, LLOJ008156, LLOJ008629) and 1 from GH31 (LLOJ003489) were complete. Truncated sequences were analyzed and annotated for missing parts using the Artemis annotation tool and alignment by ClustalW. Some sequences could not be completed because they were in a region of poor sequencing quality or at the beginning or end of the scaffold. Annotation resulted in 15 complete sequences for GH13 and 4 for GH31. Gene structures are demonstrated in Fig 1, and protein structures are represented in Fig 2.

We identified that some nucleotide sequences retrieved were codifying for more than one protein. These sequences were splitted, so the number of genes and, consequently, proteins increased. The nucleotide and protein sequences, after annotation, are presented in S1 File and S2 File, respectively. For example, the sequence LLOJ008212 codes for two proteins. The first is a glycoside hydrolase from family 31 and the second is the light chain of a sodium-coupled neutral amino acid transporter. These two sequences were not used for subsequent analysis, since the part coding for the GH31 protein is poorly assembled, and it is not possible to identify what kind of activity it represents.

After correction of the amino acid sequences, the similarity search results allowed us to identify the putative presence of fourteen α-amylases, tree maltases, two amino acid transport proteins (heavy chain), one 1,4-α-glucan branching enzyme and one glycogen debranching enzyme belonging to the GH13 family (Table 1) and four glycosidases NET37, one lysosomal α-glucosidase and the α-subunit of one neutral α-glucosidase belonging to GH31 (Table 2) in the genome of *L. longipalpis*.

338 **Table 1. Manual annotation of *L. longipalpis* GH13 sequences retrieved from vector base.**

Gene	Strand	Status	Best Hit	E value	Signal Peptide	TM Domain	N-Glycosilation Position	O-Glycosilation Position	Mw (kDa)	pI	DeepLoc	Catalytic site (InterPro)	Ca Binding Site (InterPro)
LLOJ004838_1 (LlAamyA1)	Reverse	Complete	Alpha-amylase A (AMYA_DROMA - P54215) - *Drosophila mauritiana*	0.00E+00	1-17	No	410	130, 131, 145, 367, 369	54.6	6.43	Extracellular - Soluble	204D, 241E, 306D	115N, 174D, 208H
LLOJ004838_2 (LlAamyA2)	Reverse	Incomplete (Truncated 5')	Alpha-amylase B (AMYB_DROYA - Q9BN01) - *Drosophila yakuba*	4.40E-92	1-17	No	161	131, 135, 137, 139	N.D.	N.D.	Extracellular - Soluble	206D, 243E, N.D.	N.D.
LLOJ004839 (LlAamyA3)	Forward	Complete	Alpha-amylase 1 (AM4N_DROAN - Q23835) - *Drosophila ananassae*	0.00E+00	1-15	No	280	129, 130, 135, 137, 365	55.2	5.06	Extracellular - Soluble	203D, 240E, 308D	115N, 173D, 207H
LLOJ004841 (LlAamyA4)	Reverse	Complete	Alpha-amilase 2 (AMY2_DROAN - O18345) – *Drosophila ananassae*	0.00E+00	1-16	No	No	134, 365, 367	56.0	5.51	Extracellular - Soluble	203D, 240E, 304D	114N, 173D, 207H
LLOJ004880_1 (LlAamyB1)	Reverse	Complete	Alpha-amylase 2 (AMY2_DROAN - O18345) *Drosophila ananassae*	1.00E-176	1-16	No	277	130, 131, 136, 138, 146, 152, 367	55.3	4.61	Extracellular - Soluble	204D, 241D, **307Q**	116N, 174D, **208N**

LLOJ004880_2 (LlAamyB2)	Reverse	Complete	α-amylase 1 (AMY1_DROAN - Q23835) - *Drosophila ananassae*	0.00E+00	1-16	No	297, 406, 438	130, 131, 136, 151, 370, 372	56.2	5.89	Extracellular - Soluble	204D, 244E, 309D	116N, 174D, 208H
LLOJ004880_3 (LlAamyB3)	Reverse	Complete	α-amylase A (AMYA_DROMA - P54215) - *Drosophila mauritiana*	0.00E+00	1-21	No	No	134	55.9	5.09	Extracellular - Soluble	208D, 245E, 309D	119N, 178D, 212H
LLOJ004881_1 (LlAamyB4)	Forward	Complete	α-amylase A (AMYA_DROME - P08144) - *Drosophila melanogaster*	1.50E-169	1-17	No	177	131, 132, 137, 139, 146, 147, 153, 366, 368	55.3	4.98	Extracellular - Soluble	**205N**, 242D, 307D	117N, 175D, **209F**
LLOJ004881_2 (LlAamyB5)	Forward	Incomplete	α-amylase-related protein (AMYR_DROAN - O18344) - *Drosophila ananassae*	9.90E-50	1-16	No	N.D.	131,132, 137, 139	N.D.	N.D.	Extracellular - Soluble	N. D.	N. D.
LLOJ004881_3 (LlAamyB6)	Forward	Complete	α-amylase 4N (AM4N_DROAN - Q23834) - *Drosophila ananassae*	4.00E-164	1-16	No	269	131, 132, 137, 139, 147, 371	55.9	4.47	Extracellular - Soluble	**208F**, 245D, **311Q**	117N, 178D, 212H
LLOJ004882 (LlAamyB7)	Reverse	Incomplete	α-amylase 1 (AM4N_DROAN - Q23835) - *Drosophila ananassae*	0.00E+00	1-17	No	161, 432	131, 135, 137, 139, 136	N.D.	N.D.	Extracellular - Soluble	206D, 243E, 305D	117N, 176D, 210H

LLOJ004885 (LlAamyB8)	Forward	Incomplete (Truncated 3')	α-amylase-related protein (AMYR_DROAN - O18344) - *Drosophila ananassae*	2.60E-63	1-16	No	N.D.	131, 132, 137, 139, 147	N.D.	N.D.	Extracellular - Soluble	N.D.	N.D.
LLOJ005909_1 (LlAamyC1)	Forward	Incomplete	α-amylase A (AMYA_DROMA - P54215) - *Drosophila mauritiana*	9.70E-174	1-21	No	401	134, 370	N.D.	N.D.	Extracellular - Soluble	208D, N.D., 309D	119N, 178D, 212H
LLOJ005909_2 (LlAamyC2)	Forward	Complete	α-amylase-related protein (AMYR_DROAN - O18344) - *Drosophila ananassae*	0.00E+00	1-18	No	179, 228, 414	130, 133	55.9	6.18	Extracellular - Soluble	207D, 244E, 309D	116N, 177D, 211H
LLOJ000566 (LlMal1)	Reverse	Incomplete (Truncated 3')	Maltase A1 (MAL1_DROME - P07190) – *D. melanogaster*	0.00E+00	N.D.	Yes Position 581 - 600	N. D.	159, 241, 427, 440, 450, 547	N.D.	N.D.	N.D.	216D, 290E, 357D	No
LLOJ008156 (LlMal2)	Reverse	Complete	Maltase A1 (MAL1_DROME - P07190) – *D. melanogaster*	0.00E+00	1-17	Yes Position 581 - 600	116	241, 427, 440, 450, 547	69.6	5.6	Cell Membrane - Membrane	216D, 290E, 357D	No
LLOJ002257 (LlMal3)	Forward	Complete	Maltase 2 (MAL2_DROVI - O16099) - *Drosophila virilis*	0.00E+00	1-16	Yes Position 588 - 610	101, 136, 308, 432	246, 459, 460, 463	70.9	4.7	Cell Membrane - Membrane	236D, 304E, 375D	No

LLOJ008629 (LlNBAThc)	Forward	Complete	Neutral and basic amino acid transport protein rBAT (SLC31_RAT-Q64319) - *Rattus norvegicus*	9.60E-21	No	Yes Position 157 - 176	No	300, 303, 305, 494	69.7	4.93	Endoplasmic reticulum - Membrane	**345K,** 384E, N.D.	No
LLOJ006803 (LlCD98hc)	Forward	Complete	4F2 cell-surface antigen heavy chain (4F2_RAT-Q64319) - *Rattus norvegicus*	5.40E-19	No	Yes Position 151 - 173	No	283, 367, 385, 528, 590	69.9	5.60	Endoplasmic reticulum, Membrane	**350A,** N.D., N.D.	No
LLOJ005533 (LlAGB1)	Forward	Complete	1,4-α-glucan-branching enzyme - XP_017959984.1 - *Drosophila navojoa*	0.00E+00	No	No	No	No	79.8	5.82	Cytoplasm, Soluble	347D, 402E, 471D	No
LLOJ008312 (LlGlyD1)	Forward	Complete	Glycogen debranching enzyme - XP_012157586.1 - *Ceratitis capitata*	0.00E+00	No	No	No	248, 250, 272, 344, 1066, 1069, 1177, 1298, 1307, 1311, 1312, 1316	172.6	6.26	Cytoplasm, Soluble	548D, 577E, 649D,1282 D,1515E	No

For incomplete proteins some N and O- glycosylation sites are missing (N.D.- Not Determined)
In transmembrane proteins, only extracellular domains were considered as N-glycosylated
Sequences retrieved from Vector Base (*L. longipalpis*, Jacobina strain, LlonJ1.4 geneset, June 2017)

346

347 **Table 2. Manual annotation of *L. longipalpis* GH31 sequences retrieved from vector base.**

Gene	Strand	Status	Best Hit	E value	Signal Peptide	TM Domain	N-Glycosilation Position	O-Glycosilation Position	Mw (kDa)	pI	DeepLoc	Catalytic site
LLOJ001847_1 (LlGlyMyo1)	Reverse	Complete	Myogenesis-regulating glycosidase (MYORG_MOUSE - Q69ZQ1) - *Mus musculus*	3.60E-75	No	Yes Position 20-39	No	86	79.8	8.88	Golgi apparatus, Membrane	437D, **484P**
LLOJ001847_2 (LlGlyMyo2)	Reverse	Complete	Myogenesis-regulating glycosidase (MYORG_MOUSE - Q69ZQ1) - *Mus musculus*	2.40E-72	No	Yes Position 13-35	No	No	79.3	8.18	Golgi apparatus, Membrane	427D, **485P**
LLOJ001881 (LlGlyMyo3)	Reverse	Incomplete (Truncated 5' and 3')	Myogenesis-regulating glycosidase (MYORG_HUMAN - Q6NSJ0) Homo sapiens	4.40E-83	N.D.	No	N.D.	No	N.D.	N.D.	N.D.	401D, 460D
LLOJ000840 (LlGlyMyo4)	Forward	Complete	Myogenesis-regulating glycosidase (MYORG_HUMAN - Q6NSJ0) Homo sapiens	4.50E-65	No	Yes Position 21 - 43	No	No	76.3	7.13	Lysosome/Vacuole, Membrane	408E, 467D
LLOJ006451 (LlLysAglu1)	Reverse	Incomplete (Truncated 5' and 3')	Lysosomal α-glucosidase (LYAG_RAT-	2.00E-17	N.D.	No	N.D.	No	N.D.	N.D.	Lysosome/Vacuole, Soluble	492D, 570Q

			Q6P7A9) - *Rattus norvegicus*									
LLOJ003489 (LlNAglu1)	Forward	Complete	Neutral α-glucosidase AB(B0WQR9_CULQU) - *Culex quinquefasciatus*	0.00E+00	1-18	No	114, 435, 891	323, 325	105.8	5.62	Endoplasmic reticulum, Soluble	522D, 598D

348 For incomplete proteins some N and O- glycosylation sites are missing (N.D.- Not Determined)
349 In transmembrane proteins, only extracellular domains were considered as N-glycosylated
350 Sequences retrieved from Vector Base (*L. longipalpis*, Jacobina strain, LlonJ1.4 geneset, June 2017)

According to the genomic map available in the Vector Base, the α-amylases genes are organized in two clusters. The cluster A is composed by LlAamyA1 to LlAamyA4 (four α-amylases), and the cluster B is formed by LlAamyB1 to LlAamyB8 (eight α-amylases). Besides that, the α-amylases C1 and C2 are in tandem in the genome. All the coded α-amylases are putative soluble extracellular proteins, with a predicted molecular mass from 54.6 to 56 kDa, and estimated isoelectric points from 4.4 to 6.4 (only complete sequences were analyzed, see Table 1). The coded maltases were predicted as membrane proteins. LlMal1 and LlMal2 are putative transmembrane proteins, and for LlMal3 both transmembrane domain and GPI anchor were predicted. These enzymes have molecular masses around 70 kDa and estimated isoelectric points from 4.7 to 5.6. A lysosomal membrane α-glucosidase (LlLysAglu1, LLOJ000840) with 76.3 kDa and a soluble α-glucosidase (LlNAglu1, LLOJ003489) putatively active in the endoplasmic reticulum were identified in the GH31 family (Table 2).

Analysis with InterPro and amino acid sequences alignments of *L. longipalpis* with GH13 and GH31 representative sequences from different organisms allowed the identification of highly-conserved regions, catalytic residues and the calcium and chloride binding sites for α-amylases (Table 3). The active site containing the catalytic residues, the calcium binding site and chloride binding site (Tables 1 and 3) were correctly identified for 6 complete protein sequences of α-amylases. The α-amylase B1 (LlAamyB1/ LLOJ004880_1) instead of presenting the typical catalytic triad, has only 2 residues that might serve as the nucleophile and the acid/base catalyst (204D and 241D, respectively). LlAamyB1 has also a substitution of a conserved histidine residue by asparagine in the calcium binding site, and of a conserved arginine by a leucine in the chloride binding site in the β4 sheet (Tables 1 and 3). The α-amylases B4 and B6 (LlAamyB4/LLOJ004881_1 and LlAamyB6/LLOJ004881_3) might not function as active enzymes, since their nucleophile catalytic residue was substituted by an asparagine (N) and phenylalanine (F), respectively. These enzymes also have substitutions in the calcium and chloride binding sites. Maltases also demonstrated the catalytic triad (D216, E290, D357) for both maltase 1 (LlMal1, LLOJ000566) and maltase 2 (LlMal2, LLOJ0008156), and 236D, 304E, 375D for maltase 3 (LlMal3, LLOJ0002257) with all highly conserved regions preserved.

382 **Table 3. Conserved amino acid sequences in proteins belonging to GH13 and GH31 family.**

Enzyme GH13	β2 (CSR VI)	β3 (CSR I)	Loop3 (CSR V)	β4 (CSR II)	β5 (CSR III)	β6	β7 (CRS IV)	β8 (CSR VII)
α-Amylase								
A. Orizae (TAKA Amylase)	56-GFTAIWITP	117-DVVANH	173-LPDLD	202-GLRIDTVKH	230-EVLD	-	323-FVEYHD	323-GIPIIYAGQE
LLOJ004838_1/LlAamyA1	53-GFAGVQVSP	111-DIVINH	172-LPDLN	200-GFRVDAAKH	241-EVID	-	301-FVEYHD	340-GIPRIMSSFA
LLOJ004838_2/LlAamyA2	53-GFAGVQVSA	113-DAIFNN	174-APDLD	202-GFRIDSAKH	243-EVID	-	------	----------
LLOJ004839/LlAamyA3	51-GFGGVQVSS	111-DAIFNH	171-APDLD	199-GFRVDSAKH	240-EVID	-	303-FVEYHD	338-GYPYLMSDFN
LLOJ004841/LlAamyA4	52-GYAGVQLSP	110-DVLLNH	171-LHDLN	199-GFRVDAAKH	240-EVID	-	299-FVDYHD	338-GTIFLMSSFA
LLOJ004880_1/LlAamyB1	52-GYAGVQVSS	112-FTIFNN	172-TPDLD	200-GFLIDSAKH	241-DVIE	-	302-FVEYEQ	340-GNPFIMSDFD
LLOJ004880_2/LlAamyB2	52-GFGGVQVSS	112-DTVINH	172-APDLN	200-GFRIDAAKH	244-EVID	-	304-FVEYHD	343-GIPRIMSSFA
LLOJ004880_3/LlAamyB3	57-GFGGVQVSP	115-DVIFNH	176-LPDLN	204-GFRADACKH	245-EVID	-	304-FVEYHD	343-GIPRIMSSYD
LLOJ004881_1/LlAamyB4	53-GFAGVQISS	113-EIIMNN	173-LPDLN	201-GFRINSAKH	242-DVAD	-	302-FVDSPD	339-GHPFIMSSYE
LLOJ004881_2/LlAamyB5	52-GFGGVQVSS	113-DAIFNE	-----	---------	----	-	------	----------
LLOJ004881_3/LlAamyB6	52-GFGGVQVSS	113-DAVFNN	176-VPDLD	204-GFRIFAAKH	245-DVID	-	306-FVEYEQ	344-GNPLLNSDYE
LLOJ004882/LlAamyB7	53-GFAGVQVSA	113-DAIFNN	174-APDLD	202-GFRIDSAKH	243-EVID	-	300-FVEYHD	370-GNLFIMSDFD
LLOJ004885/LlAamyB8	52-GFGGVQVSS	113-DAIFNE	176-VPDLD	204-GFRILH---	----	-	------	----------
LLOJ005909_1/LlAamyC1	57-GFGGVQVSP	115-DVIFNH	176-LPDLN	204-GFRADACKH	----	-	304-FVEYHD	305-GIPRIMSSYD
LLOJ005909_2/LlAamyC2	54-GFGGVQVSP	113-DIVVNH	175-LRDLN	203-GFRIDAAKH	244-EVID	-	304-FVEYHD	343-GVPRMMSSYE
Maltase								
A. japonica (JHGase I)	63-GITAIWLSP	117-DLVPNH	191-QPDLN	219-GFRIDAIPH	292-EAYT	-	349-VKGNHD	378-GVAVTYYGEE
LLOJ000566/LlMal1	58-GMSGTWLSP	112-DFVPNH	184-QPDLN	212-GFRIDAVPY	290-EAYT	-	352-VLGNHD	382-GMAVTYNGEE
LLOJ008156/LlMal2	58-GMSGTWLSP	112-DFVPNH	184-QPDLN	212-GFRIDAVPY	290-EAYT	-	352-VLGNHD	382-GMAVTYNGEE
LLOJ002257/LlMal3	78-GIFATWLSP	132-DFVPNH	204-QPDLN	232-GFRVDAINH	304-EAYT	-	370-VLGNHD	400-GVAVTYYGEE
Amino Acid Transport Protein								
Homo Sapiens (rBAT)	156-NIKTVWITS	210-DFIPNH	282-QPDLN	310-GFSLDAVKF	384-EAYA	-	-?	474-GTPITYYGEE
LLOJ006803/LlCD98hc	202-NIQIEGIKG	255-DLTPNF	318-RFDLQ	346-GIRLANTKH	----	-	-?	501-GVPVFNVDAS
LLOJ008629/LlNBAThc	201-SFQDSNGDG	254-DIDEIL	-----	338-GFYVKGLEN	383-EKVR	-	-?	462-GTPSIFYGDE
1,4-α-glucan-branching enzyme								
E. coli	280-GFTHLELLP	335-DWVPGH	-?	401-ALRVDAVAS	458-EEST	-	519-LPLSHD	562-GKKLLFMGNE
LLOJ005533/LlAGB1	221-GYNAIQVMA	276-DVVHSH	-?	343-GYRFDGVTS	402-EDVS	-	466-YAESHD	-
Glycogen debranching enzyme (α-1, 4 glucanotransferase)								
Candida glabrata (CgGDE)	105-GTYCFYLSF	230-DIVFNH	-?	531-GFRIDNCHS	564-ELFS	-	622-MDITHD	-
LLOJ008312/LlGlyD1	100-GTFHFYFSF	228-DIVLNH	-?	544-GVRLDNCHS	577-ELFT	-	644-LDLTHD	-
GH31								
Neutral α-glucosidase								
Homo sapiens (GANAB_HUMAN)				538-FVWNDMNEP		616-TGDNTAE		
LLOJ003489/LlNAglu1				518-MLWNDMNEP		596-TGDNTAE		
Lysosomal α-glucosidase								

Homo sapiens(LYAG_HUMAN)	514-GMWIDMNEP	614-TGDVWSS
LLOJ006451/LlLysAglu1	479-GYMVQANWM	568-RRQVPST
Glycosidase NET37		
Homo sapiens (MYORG_HUMAN)	459-SFKFDAGEV	518-LVDRDSV
LLOJ001847_1/LlGlyMyo1	437-GFYLDFGTS	484-LPPVNSS
LLOJ001847_2/LlGlyMyo2	427-GFYLDFGTS	485-LPPVNSS
LLOJ001881/LlGlyMyo3	397-SFKFDAGES	458-MIDKDSL
LLOJ00840/LlGlyMyo4	404-NFYFEGGEF	465-LPDMDNR

383 The catalytic conserved residues, calcium binding and chloride binding residues are with grey, purple and green background, respectively.

The amino acid transport proteins found in our screening share a domain similarity with maltases, but the general protein primary structure is different (Fig 2), and they lack the conserved catalytic residues (Table 3). By similarity, the LLOJ008629 was described as a neutral and basic amino acid transport protein (NBAThc) and the LLOJ006803 due to the presence of the SLC3A2 domain, was classified as CD98hc (cluster of differentiation 98).

The 1,4-α-glucan branching enzyme (LlAGB1, LLOJ005533) has a central domain containing the active site and the catalytic residues similarly to the other enzymes from GH13. In the glycogen debranching enzyme (LLOJ008312), both α-1,4 glucanotransferase (GT) and α-1,6glucosidase (GC) domains were identified, and the enzyme was characterized as a soluble cytoplasm enzyme with an estimated size of 172.6 kDa. The GT domain at N-terminal region of this protein, despite the low overall sequence identity, has similarity with five short conserved regions of the GH13 family; the catalytic triad is conserevd, with the D548 as the catalytic nucleophile, 577E as the proton donor, and the conserved residue 649D. The C-terminal GC domain is similar to the catalytic domain of glucoamylases, and to other members from GH15 family. The general acid and general base for GC activity were identified as 1282D and 1515E, respectively.

Regarding the members of GH31, initially the neutral α-glucosidase (α-subunit) (LlNAglu1, LLOJ003489) was described with both Galactose mutarotase domain at the n-terminal end and the GANAB_GANAC domain (Fig 2C). This enzyme shared a 46.8 % identity with the human neutral α-glucosidase (GANB_HUMAN) and was predicted to be localized at the endoplasmic reticulum. The two aspartic residues involved in the catalytic domain were identified as 522D and 598D (Table 2). The lysosomal α-glucosidase (LlLysAglu1, LLOJ006451) presented a substitution of both catalytic aspartate residues by glutamine. The glycosidases NET37 were predicted to have a single transmembrane domain at its amino terminus and the GH31_NET37 domain at its C-terminus, with the glucosidase domain facing the non-cytoplasmic region (Fig 2C). Although the sequence for glycosidase NET37 3 (LlGlyMyo3, LLOJ001881) was not complete it was possible to predict the two catalytic residues, 401D and 460D. Glycosidase NET37 4 (LlGlyMyo4, LLOJ000840) has a glutamate as the conserved catalytic nucleophile, 408E, and 467 D as a proton donor. Since these proteins are dependent on the catalytic residues to be active, glycosidases NET37 1 (LlGlyMyo1, LLOJ001847_1) and 2 (LlGlyMyo2, LLOJ001847_1), seem to lack hydrolytic activity,

because of the presence of a proline in the conserved region of the proton donor residue (Tables 2 and 3).

Gene numbers and phylogenetic analysis of GH13 and GH31 families

We compared the sequences from GH13 and GH31 from *L. longipalpis* genome with sequences from other dipteran species to verify if gene amplification might have occurred in GH13 and GH31 families in the *Phlebotominae* branch. Phylogenetic trees were constructed using a neighbor-joining algorithm, aiming to help with the classification and prediction of the functional role of the proteins. The sequence name, species of origin and identifier are shown in S1 Table for GH13 proteins and S2 Table for GH31.

After the functional annotation of 21 GH13 genes and 6 GH31 genes in the *Lutzomyia longipalpis* genome, the results suggest an expansion of the α-amylases genes in *L. longipalpis*, when compared to the other 6 dipteran species considered (Table 4). The total of 14 α-amylase sequences found was 75 % and 16 % higher when compared to *Phlebotomus papatasi* and *Culex quinquefasciatus*, respectively, and almost 3-fold the number of α-amylases found in the representatives of the *Drosophilidae* family. *C. quinquefasciatus* was the organism with the largest number of α-amylase genes inside the dipteran order. Interestingly, for maltase genes, a contraction occurred in comparison to all 6 species present in Table 4. *P. papatasi* has 3-fold the number of maltase genes than *L. longipalpis*, and there was a reduction of about 84 % in the number of maltase genes in *L. longipalpis* when compared to *C. quinquefasciatus*. The number of α-1,4-glucan branching and glycogen debranching enzymes (GDE) is conserved among the species analyzed. For GH31 the number of lysosomal and neutral α-glucosidase is conserved in phlebotomine and mosquitoes, but in *Drosophila* the lysosomal α-glucosidase was not found. Glycosidase NET37 demonstrated a variable number of copies among the species analyzed, and seem to be expanded in *L. longipalpis* when compared to *P. papatasi*.

Table 4. Comparison of genes numbers from GH13 and GH31 retrieved from Vector Base, Fly Base and NCBI for different Diptera species.

	L. longipalpis	*P. papatasi*	*A. Aegypti*	*A. gambiae*	*C. quinquefasciatus*	*D. melanogaster*	*D. ananassae*
GH13							
α-Amylase	14	8	9	4	12	3 (4)	5
Maltase/α-glucosidase	3	9	12 (14)	11	19	10 (16)	10 (12)
Amino acid transport protein heavy chain	2	2	2(3)	1	1	1 (4)	1 (2)
1,4-α-glucan-branching enzyme	1	1	1	1	1	1	1
Glycogen debranching enzyme	1	1	1(2)	1(2)	2	1 (6)	1 (3)
GH31							
Lysosomal α-glucosidase	1	1	1 (8)	1	1	0	0
Neutral α-glucosidase α subunit	1	1	2	1	2	1 (5)	1
Glycosidase NET37	4	2	6 (10)	3 (5)	3	2 (5)	2 (4)

The number of genes between brackets represents the total number of transcripts

Phylogenetic analysis of GHF13 protein sequences demonstrated the occurrence of 5 clades, each of them representing one of the activities that were already described in this family (Fig 3). Clade I represents the α-amylases, and we divided this clade into nine different subgroups. The subgroups V and VIII have representatives only from the *Phlebotominae* subfamily. According to our subgroup classification, the α-amylases from *L. longipalpis* are distributed in subgroups I, II, V, VIII. The results showed that 8 α-amylases in the *L. longipalpis* genome, corresponding to 57 % of *L. longipalpis* α-amylases, group in the subgroup VIII, suggesting a gene expansion in this protein family. The other α-amylases in subgroups I, II, and IV, group with *P. papatasi* orthologs.

Clade V represents the maltases and is also divided into nine different subgroups. The *L. longipalpis* maltases are classified in subgroups III and VI. In both subgroups, they group with *P. papatasi* orthologs. The subgroups VIII and IX have representatives only from the *Drosophilidae* family. As described above, *P. papatasi* seems to have an expansion in the maltase genes when compared to *L. longipalpis*. Maltases from *P. papatasi* have representatives in subgroups II, III, IV, VI and VII, with 44 % belonging to the subgroup VII. Interestingly, the LlMal3 and PpMalA2 are orthologous to the mosquitos' maltases in subgroup III. These enzymes are characterized as GPI membrane-bound α-glucosidases with toxin binding properties.

Clade II represents the debranching enzymes and forms a paraphyletic group, as the *C. quinquefasciatus* sequences formed a different branch. The glucan branching enzymes are represented by a monophyletic group in clade IV, demonstrating the

sequence conservation among dipteran species. Both enzymes from *L. longipalpis* and *P. papatasi* are located in a branch grouped with the *D. melanogaster* and *D. ananassae*.

The clade IV, representing the amino acid transport proteins, formed a paraphyletic group, and it was divided into two subgroups. Subgroup I represent the neutral and basic amino acid transporters (NBATs), and subgroup II the CD98 heavy chain, which comprises the large neutral amino acid transport (LAT1). The LlNBAThc form a subgroup with the ortholog proteins from *P. papatasi* (PpNBAThc) and *A. aegypti* (AaNBAThc). Both proteins miss the SLC3A2 domain, which is characteristic from CD98hc. The LlCD98hc grouped with the orthologous in the *Drosophilidae* family, and these proteins are already classified as different isoforms of the CD98 heavy chain.

Phylogenetic analysis of protein sequences from GHF31 revealed that *L. longipalpis* GHF31 genes group with the four major clades found in this family. Two clades contain the glycosidases NET37 (clades I and II) and the other two contain the lysosomal and the neutral α-glucosidases, clades III and IV, respectively (Fig 4). For NET37 glycosidases, clade I represent proteins containing the conserved catalytic domain and clade II represent glycosidases without the catalytic residues.

Expression of *L. longipalpis* GH13 and GH31 proteins in different physiological conditions

After analyzing the sequences using bioinformatics tools, primers were designed for selected sequences in the GH13 family and all sequences from GHF31. In the case of α-amylases, one representative for each subgroup in clade I (Fig 3) was analyzed. Also, due to the high sequence similarity, it was not possible to design specific primer pairs to analyze the expression of genes coding for LlMal1/LlMal2 or LlMyoGly1/LlMyoGly2 independently.

The designed primers were tested for amplification of genomic DNA and establishment of PCR conditions. Using genomic DNA as a template, all sequences were amplified, with detection of only one amplicon with the expected size (data not shown). The expression of the *L. longipalpis* GH transcripts was evaluated by RT-PCR using different tissues (rest of body x gut) as sources for RNA, and in different dietary conditions, namely newly hatched females, sugar or blood-fed females.

Both LlGlyMyo4 and the lysosomal α-glucosidase were poorly expressed in all tested tissues and conditions (data not show). The LlAamyA1 and LlAamyA4 were the α-amylases with the highest levels of expression (Fig 5A). While the LlAamyA1 seems to

be expressed at the same level in both tissues, LlAamyA4 is more expressed in the gut and was not altered considering the dietary condition. The same pattern observed for LlAamyA4 was encountered for the expression of maltase transcripts (LlMal1/LlMal2 and LlMal3; Fig 5B). Interestingly, the LlAamy6 was not expressed when females were fed with blood (data not shown). The LlAamyB4 is poorly expressed in the tested conditions, and LlAamyC2 was detected in the rest of body but not in the gut.

The expression of the 1,4-α-glucan branching enzyme and the amino acid transporters is not linked to a specific tissue, and no significant change in expression was noted for LlNBAThc and LlCD98hc (Fig 5C) in different feeding condition. The expression of LlAGB1 (Fig 5D) is reduced when females feed on blood.

For GH31 family the genes coding for the glycosidases NET37 were not expressed when females were fed with blood, neither in the carcass nor in the gut. The LlNAglu1 was significantly more expressed in the gut and was not affected by the feeding condition.

Expression of LlAamy4, LlMal1/LlMal2, LlMal3, and LlNAglu2 in *L. longipalpis* infected with *Leishmania mexicana*.

Since we have interest in sand fly proteins that might be affected by *Leishmania*, we selected the LlAamyA4, LlMal1/LlMal2, LlMal3 (GH13) and LlNAglu1 (GH31) genes for further quantitative expression analysis during *L. mexicana* infection. In addtion, these genes were selected considering their higher expression in the gut, the tissue that is in direct contact with the parasite.

The results of RT-qPCR for these GH transcripts confirmed that they are more expressed in the gut tissue when compared to the rest of the body. LlMal1/LlMal2 and LlMal3 genes showed higher quantitative expression values compared to the other groups analyzed. Expression values for LlMal1/2 and 3 were about 40 X higher than results obtained for LlAamyA4, and 10 X higher than the values observed for LlNAglu1 (Fig 6). Interestingly, the expression of these putative maltases was significantly affected by the blood diet and *Leishmania* infection. We observed an induction with blood diet of LlMal1/LlMal2 (1.14 ± 0.06) and LlMal3 (1.2 ± 0.1), but infection with *L. mexicana* parasite modulated negatively both the expression of LlMal1/LlMal2 (0.4 ± 0.1) and LlMal3 (0.6 ± 0.1) (Fig 6A). The sugar diet surprisingly did not affect expression compared to the nonfed condition.

LlAamyA4 had the lowest level of expression among the genes studied and presented the same pattern obtained previously using semiquantitative RT-PCR. Besides that,

LlAamyA4 expression was not affected by the presence of parasites in the blood diet (Fig 6B). LlNAglu1 is not highly expressed in the gut tissue compared to RB, but there is a significant difference between tissues for all dietary conditions tested. The blood diet induces the expression of LlNAglu1 in both tissues, but it is not affected by the presence of *L. mexicana* parasites.

Discussion

For sandflies, the energy requirements are provided primarily by sugar meals obtained from different sources, for both female and male adults. These sugar sources have been described for sand fly species, depending on their preferred site for development. Desert and savanna regions are the preferred habitats for the genus *Phlebotomus* (Old World), and forests for the genus *Lutzomyia* (New World). Sand flies commonly feed on plant tissues, fruit juices, nectars from flowers, honeydews excreted by aphids and coccids, among other sources, with different sugar compositions (12,13,59). In the case of females that feed on blood for egg maturation, their meal is also rich in sugar and glyco-derivatives, as glycoproteins and glycolipids. The immature forms of sandflies are detritivores, although there is no definitive proof, they develop by ingesting soil rich in bacteria, fungi and organic molecules derived from decomposition (60,61). Thus, the glycosidases play a crucial role in the digestion of larvae and adults sand flies. α-amylases and α-glucosidases, are especially important for adults, since starch is found in leaves and fruits, and sucrose is the main component of sap and nectar. In larvae these enzymes play an important role in the digestion of glycogen, the reserve carbohydrate in fungi. In insects, these enzymes have been described in families GH13 and GH31 (15).

The glucose obtained from food is up taken by enterocytes through specific membrane transporters, and used to produce energy through glycolysis and related metabolic pathways. The glucose might be also converted to trehalose, the circulating sugar in the hemolymph. When circulating trehalose is in excess, glucose is stored as glycogen or converted in triacyl-glycerides in several tissues. Insects have to expend energy continuously, and adaptation to changes in food availability is a central challenge for survival. In a starving condition, they must live on reserves (glycogen and triacyl-glycerides) accumulated in periods of food abundance. The flight, for example, is a condition of extreme energy demand (62,63). In insects, some of the enzymes involved in the process of storage and mobilization of glycogen, as the 1,4-α-glucan branching

enzyme, glycogen debranching enzyme, and lysosomal α-glucosidase are also members of families GH13 and GH31.

In this work, we have investigated the presence of genes coding for proteins belonging to the α-amylase family GH13, and GH31 family in *L. longipalpis*, using the genome information available in the Vector Base. We identified after annotation a total of 21 genes belonging to GH13 family and described as α-amylases (14), maltases (3), amino acid transporters (2), 1,4-α-glucan branching enzyme (1) and glycogen debranching enzyme (1), and GHF31 genes that were described as glycosidases NET37 (4), lysosomal α-glucosidase (1) and neutral α-glucosidase (1) (α-subunit). These proteins are involved in carbohydrate digestion (α-amylases and maltases), glycogen metabolism (1,4-α-glucan branching enzyme and glycogen debranching enzyme, lysosomal α-glucosidase), amino acid transport (NBAThc and CD98hc), quality control of N-glycosylation in the plasmatic reticulum (neutral α-glucosidase α subunit) and myogenesis regulation (NET37).

We described potentially extracellular soluble α-amylases and membrane-anchored maltases with predicted molecular masses and isoelectric points around 55 kDa, 4.5-6.5 and 70 kDa, 4.5-5.5, respectively. Insect α-amylases have molecular weights about 48-60 kDa, pI 3.5-4.0, are calcium-dependent enzymes, and in some cases activated by chloride. α-glucosidases (including maltases) have molecular weights ranging from 60-80 kDa (or multiple of these), with isoeletric points from 5.0-7.2 (23). In general, α-amylases are found as soluble enzymes in the midgut lumen. α-glucosidases are soluble, secreted to the midgut lumen or restricted to the ecto-peritrophic space, or anchored to the midgut cells, being anchored in the microvillar or perimicrovillar membranes (64–66). The LlMal1, LlMal2, and LlMal3 described for *L. longipalpis* are membrane-anchored proteins with higher expression in the midgut compared to the rest of body in adults sandflies (Figs 5B and 7A). Their expression was higher when females feed with blood, but did not change in sugar-fed females (Fig 6A). Besides that, infection with the parasite *L. mexicana* modulates their expression. α-glucosidases were described as membrane proteins anchored to the midgut glycocalyx in *L. longipalpis*, with a basal activity in non-fed females and activity induction starting 24 h after sugar feeding (67,68). Our expression analysis suggests that the increase of α-glucosidase activity previously reported in sugar-fed females is not due to a transcriptional gene regulation, because no expression induction was observed for any gene encoding maltase when females were sugar-fed. Another possibility is that transcription is regulated transiently and after 48 hours it was not possible to determine any gene induction or repression. In adults and larvae of *D.*

609 *melanogaster*, it has been described that some genes encoding carbohydrate digestive
610 enzymes, as MAL-A1, are repressed in response to dietary sugars, due to a mechanism of
611 regulation by negative feedback. The MAL-A4 gene is transiently induced when maltose
612 is offered to larvae and repressed when they are fed with glucose (69,70).
613 A soluble form of α-glucosidase was described for *L. longipalpis* and *P. langeroni*
614 when females were blood fed, (67,71). Biochemical assays using different tissues as
615 enzyme source, in different conditions, and chromatographic analysis demonstrated the
616 presence of at least 4 different α-glucosidases in *L. longipalpis* (67). In this work, we
617 identified 3 genes coding for membrane anchored maltases/α-glucosidases. It is possible
618 that these genes undergo alternative splicing, producing different isoforms of α-
619 glucosidases. Removal of the last exon eliminates the transmembrane region, and allows
620 these enzymes to be secreted in soluble form, changing some biochemical characteristics
621 of the enzymes. Additionally, the protein LlMal3 is putatively anchored to the membrane
622 by a GPI anchor, and this might undergo cleavage by phospholipases releasing the protein
623 from the midgut membrane. The soluble form of these α-glucosidases/maltases may pass
624 through the peritrophic membrane and act sinergistically with trypsin and other enzymes
625 that release glycopeptides or glycolipids from blood glycoconjugates in the
626 endoperitrophic space. Also the soluble α-glucosidases might play a role in the process
627 of heme detoxification during blood digestion, in the hemozoin formation. This
628 mechanism was described for α-glucosidases form *R. prolixus* (72). In this respect, *L.*
629 *longipalpis* maltases may play different roles in the hydrolysis of glycosides during the
630 digestion process associated with different enzyme specificities for sugars or glyco
631 derivatives.
632 During its development inside the sand fly midgut, the *Leishmania* parasite has to
633 overcome several physiological, biochemical, and molecular changes for the
634 establishment of infection. Modulation of different sand fly enzymes by the parasites has
635 been described, with emphasis in the modulation of proteases (73–76). The modulation
636 of proteases helps in the resistance of the parasite against proteolytic cleavage during the
637 blood digestion. However, to our knowledge, this is the first time that the modulation of
638 a sand fly carbohydrase by *Leishmania* is demonstrated. *Leishmania* species have
639 membrane glycoconjugates as lipophosphoglycan (LPG) and glycoinositolphospholipids
640 (GIPLs) (77), that are complex structures with polymorphisms related to different sugar
641 compositions of its glycan structures. These surface molecules are essential during
642 *Leishmania* development and infection maintenance in the gut of the sand fly vectors (77–

643 81), so a possible role for this modulation of sand fly carbohydrases might be to avoid the
644 damage to surface glycoconjugates.
645 Membrane glycoconjugates are well described for some *Leishmania* species. In *L.*
646 *mexicana*, for example, the non-reducing end the lipophosphoglycan is capped with
647 Manα1-2Man, Manα1-2Manα1-2Man or Manα1-2(Galβ1-4)Man (82). In
648 glycoinositolphospholipids (GIPLs), there is a high degree of galactosylation in the
649 structures linked to the glycan core of Manα1-4GlcN (83). The mannose and galactose
650 are glucose epimers, so its possible that sand fly α-glucosidases act upon these linkages
651 breaking these molecules.
652 We described three maltase coding genes in *L. longipalpis*, suggestion a retraction
653 when compared to other dipterans, as *Drosophila* and even *P. papatasi*. For *P. papatasi*
654 and the three mosquitoes species used for comparison, the maltase genes form two
655 clusters; each one consisting of three to five genes. This data suggests that, although the
656 primary role of α-glucosidases locate in sugar metabolism, in the process of evolution
657 these proteins might have acquired additional functions in these organisms. In some
658 mosquitoes species, like *C. pipiens, C. quinquefasciatus*, and *A. gambiae*, the α-
659 glucosidase is also described as the receptors of the Bin toxin from *Lysinibacillus*
660 *sphaericus* (84–86). The maltase genes of *D. melanogaster* demonstrate different
661 expression levels in different tissues, with temporal changes depending on the phase of
662 development, and a non-enzymatic regulatory function for these proteins was suggested
663 (87). In *R. prolixus*, α-glucosidases are involved in the detoxification of heme during the
664 blood digestion (72). There are some examples where enzymes have lost their catalytic
665 activity, developing new functions, as the heavy chain of the amino acid transporters
666 (neutral and basic amino acid transport protein (hcNBAT) and the cluster of
667 differentiation 98 (CD98)), but they still have high sequence similarity with α-
668 glucosidases (Fig 3). The amino acid transporter is a heterodimeric protein, and the heavy
669 chain, with 12 transmembrane domains, has the function of correctly locating and folding
670 the light chain on the plasma membrane. The light chain has the function of transporting
671 amino acids across the cell membrane (88).
672 Interestingly, the α-amylase genes described for *L. longipalpis* are organized in two
673 clusters (A and B) and the α-amylases C1 and C2 are in tandem in the genome in a
674 different region. LlAamyC2 is correspondent to the protein AF132512, previously
675 described as a salivary α-amylase (89). These clustered genes have the same putative
676 enzymatic specificity, are closely spatially localized and have high sequence similarity to

each other, so probably these sequences originated from multiple subsequent genomic duplications. *P. papatasi* and *C. quinquefasciatus* also present the clustering of some α-amylases. Our results demonstrate a gene expansion of *L. longipalpis* α-amylase family, when compared to *P. papatasi*, mosquitos' species and the genus *Drosophila* (Fig 3, Table 4).

The expansion of the α-amylase gene family in *L. longipalpis* might reflect a tissue expression specificity, adaptation to different reaction conditions or might represent the specialization of these activities to overcome different plant inhibitors. This might be related to the diversity of available food sources that are present in the natural habitat of *L. longipalpis*, when compared to the habitat of *P. papatasi*. *Phlebotomus papatasi* is mainly found in the temperate areas of the Old World, living in savanna, desert, and semi-arid areas, while *L. longipalpis* is found in tropical areas, mainly in forest regions with a higher diversity of plants (1). It was demonstrated that even inside the *P. papatasi* species, there is a variation in the spectrum of glycosidase activities, when comparing phlebotomines collected in an oasis (sugar-rich) to those collected in the dry-season desert (90). In addition, α-amylases may be differentially expressed depending on the insect stage of development, since they were described as active enzymes both in adults and larvae (61,91).

The catalytic domain (Domain A) of GH13 members is formed by a $(\beta/\alpha)_8$ barrel structure (18,19,92) with seven conserved sequence regions (CSRs). The catalytic triad aspartate (catalytic nucleophile), glutamate (general acid/base) and aspartate (transition-state stabilizer) is localized in the β4 (CSR II), β5 (CSR III) and β7 (CSR IV) sheets, respectively (17,20,21). The CSRs and the catalytic residues were found in most *L. longipalpis* α-amylases and maltases, and we also identified three conserved calcium binding sites in the β3 sheet (CSRI, Tables 3 and 1), loop 3 (CSR V) and β4 sheet (CSR II, Table 3) in the complete α-amylases sequences. In general, 4 residues function as the calcium binding site, and three of them are conserved (25). Calcium preserves the structural integrity of the active site, maintaining proper folding (93). Although, the LlAamyB1 did not demonstrate the typical catalytic residues. In LlAamyB4 and LlAamyB6, the nucleophile catalytic residue was substituted by an asparagine (N) and phenylalanine (F), respectively. It is not possible to confirm whether these proteins are catalytically inactive or function with a distinct catalytic mechanism from the classical one described for GH13 proteins.

710 The chloride binding site is a characteristic of some α-amylases. It is situated close to
711 the center of the $(\beta/\alpha)_8$ barrel, and described as the residues Arg195, Asn298 and Arg/Lys
712 337 based on the sequence of PPA (porcine pancreatic α-amylase) (26). In most *L.*
713 *longipalpis* α-amylases, the chloride binding residues were identified. The dependence of
714 *L. longipalpis* larvae α-amylases on chloride ions was demonstrated (61). However,
715 LlAamyB1 and LlAamyB4 have a substitution in the asparagine residue (Table 3). The
716 chloride works as an allosteric activator of the enzyme, leading to a conformational
717 change causing a variation in the proton donor environment, but the evolutionary
718 advantage of this regulation mechanism remains unclear (23,26). The chloride binding
719 causes the displacement of the optimum pH, from acidic to slightly basic, by increasing
720 the pKa of the proton donor (94). In this way, the pH activity profile is broadened and
721 this might be a physiological advantage, especially for alkalophilic species. Some α-
722 amylases do not have the chloride binding site, and work as chloride-independent, as α-
723 amylases from plants, fungal and most bacteria (95).
724 We described in the *L. longipalpis* genome, belonging to the GH31 family, the soluble
725 neutral α-glucosidase α-subunit, a lysosomal α-glucosidase without catalytic residues and
726 a protein with a putative glycosidase activity anchored to the nuclear envelope by a
727 transmembrane domain (NET37). The neighbor-joining analysis of GHF31 sequences
728 demonstrated (Fig 4) that these different activities were grouped in IV different clades,
729 and the NET37 proteins were divided into two different subclades depending whether the
730 catalytic amino acids are present or not. Together with the expression analysis, these
731 results suggest that at least for the neutral α-glucosidase this enzyme may play its
732 canonical role in the metabolism of *L. longipalpis*. The *L. longipalpis* lysosomal α-
733 glucosidase has shown a truncated sequence, but was predicted based in the grouping in
734 the phylogenetic tree with other enzymes from mosquitoes species. Proteins from
735 mosquitoes present the TREFOIL and the NtCtMGAM domains, that are typical of acid
736 α-glucosidases. The best characterized enzyme in this group is the human lysosomal
737 glucosidase (51). Since the classical catalytic residues were not identified in the *L.*
738 *longipalpis* sequence, and the expression levels for this enzyme were low, it is plausible
739 to assume that this represents a protein that has lost its enzymatic activity during
740 evolution. In the *Drosophilidae* family (*D. melanogaster* and *D. ananassaea*), the
741 lysosomal α-glucosidase was not identified, and in the mosquitoes species used for
742 comparison in this work, the representatives found also lack the catalytic residues. This
743 might suggests that, in these insects, the glycogen degradation inside lysosomes was lost

744 and that the glycogen debranching enzyme mainly performs the mobilization of glycogen
745 reserves in the cytosol, in association with other enzymes.
746 The neutral α-glucosidases, described as glucosidase II in mammals, are soluble
747 enzymes present in the endoplasmic reticulum, and participate in the quality control of
748 glycoprotein folding, catalyzing the hydrolysis of glucose residues of oligosaccharides
749 bound to peptides (32,33,96). These proteins are composed of one catalytic α-subunit and
750 the noncatalytic β-subunit, which has an ER retention signal (HDEL) at the C-terminal
751 position. In mammals, the GANAB gene encodes the α-subunit, and the PRKCSH gene
752 encodes the β subunit. In *L. longipalpis*, the LlNAglu1 represents the catalytic α-subunit
753 of the neutral α-glucosidase, being homologous to the α-glucosidase II in mammals. The
754 correct catalytic residues were identified for LlNAglu1, with the nucleophile catalytic
755 inside the domain WNDMNE and the general acid/base aspartate positioned on the β6
756 sheet (14,27–29). In insects, neutral α-glucosidases have been identified by
757 bioinformatics approaches, but no biochemical or structural studies have been carried
758 until now.
759 The NET 37 proteins are not well described, especially in insects. Recently, they were
760 characterized in mammals as membrane proteins anchored in the nuclear envelope with
761 the catalytic domain facing the endoplasmic reticulum (97,98). This protein has a
762 transmembrane domain in the N-terminal region and a catalytic C-terminal region
763 belonging to GH31 family. *Lutzomyia longipalpis* NET37 proteins (Fig 2C) present the
764 same structural organization. Different NET proteins with different functions have been
765 described. It has been demonstrated that NET37 is required for myogenic differentiation
766 of C2C12 cells (mouse myoblast cell line) with higher expression in the skeletal muscle
767 of adult mouse and that this activity is dependent on the catalytic site of the glycosidic
768 region (98). Due to the general lack of knowledge about these proteins, it is not possible
769 to suppose why their expression is affected by blood feeding in adult female sand flies.
770 In summary, our results demonstrate that proteins from GH13 and GH31 families are
771 conserved in *L. longipalpis*, with enzymes involved in sugar metabolism, storage and
772 mobilization of energetic reserves (glycogen), and also proteins putatively involved in
773 amino acid transport, N-glycosylation quality control and myogenesis regulation. The
774 canonical role of α-amylases and maltases is digestion of carbohydrates, but in *L.*
775 *longipalpis*, these enzymes seem to be recruited to recognize a higher variety of dietary
776 carbohydrates. New functions might also have arisen for these glycosidases, in coherence
777 with an expansion in α-amylase gene family, and with the induction of maltases after

778 blood-feeding in females. To elucidate the roles of α-amylases and maltases, including
779 the interaction of sand flies with the *Leishmania* parasite, a more detailed analysis of their
780 biochemistry and mechanisms of expression should be pursued.
781
782 **Acknowledgments**
783 Authors thank Michelle Bates from Lancaster University for technical assistance.
784
785 **References**
786 1. Akhoundi M, Kuhls K, Cannet A, Votýpka J, Marty P, Delaunay P, et al. A
787 Historical Overview of the Classification, Evolution, and Dispersion of
788 Leishmania Parasites and Sandflies. PLoS Negl Trop Dis. 2016;10(3):1–40.
789 2. Galati E. Classificação de Phlebotominae. In: Rangel EF, Lainson R, editors.
790 Flebotomíneos do Brasil. Rio de Janeiro: Fiocruz; 2003. p. 23–51.
791 3. Seccombe, A. K.; Ready, P. D.; Huddleston LM. A Catalogue of Old World
792 Phlebotomine Sandflies (Diptera: Psychodidae, Phlebotominae). Occas Pap Syst
793 Entomol. 1993;no. 8:717.
794 4. Sanchez Clemente N, Ugarte-Gil CA, Solórzano N, Maguiña C, Pachas P, Blazes
795 D, et al. Bartonella bacilliformis: a systematic review of the literature to guide the
796 research agenda for elimination. PLoS Negl Trop Dis [Internet]. 2012 [cited 2018
797 Jul 24];6(10):1819. Available from:
798 http://www.ncbi.nlm.nih.gov/pubmed/23145188
799 5. Tesh RB. The Genus Phlebovirus and its Vectors. Annu Rev Entomol [Internet].
800 1988 Jan [cited 2018 Jul 24];33(1):169–81. Available from:
801 http://www.annualreviews.org/doi/10.1146/annurev.en.33.010188.001125
802 6. WHO. WHO: Weekly epidemiological record: Global leishmaniasis update, 2006-
803 2015, a turning point in leishmaniasis surveillance. World Heal Organ.
804 2017;92(38):557–72.
805 7. Pace D. Leishmaniasis. J Infect. 2014;69(S1):S10–8.
806 8. Sharma U, Singh S. Insect vectors of Leishmania: Distribution, physiology and
807 their control. J Vector Borne Dis. 2008;45:255–72.
808 9. Killick-Kendrick R. The biology and control of Phlebotomine sand flies. Clin
809 Dermatol. 1999;17(3):279–89.
810 10. Soares RPP, Turco SJ. Lutzomyia longipalpis (Diptera: Psychodidae:
811 Phlebotominae): a review. Ann Brazilian Acad Sci [Internet]. 2003;75(3):301–30.
812 Available from: www.scielo.br/aabc
813 11. Murray HW, Berman JD, Davies CR, Saravia NG. Advances in leishmaniasis. In:
814 Lancet. 2005. p. 1561–77.
815 12. Killick-Kendrick R, Killick-Kendrick M. Honeydew of aphids as a source of sugar
816 for Phlebotomus ariasi. Med Vet Entomol. 1987;1(3):297–302.

817 13. Macvicker JA k., Moore JS, Molyneux DH, Maroli M. Honeydew sugars in wild-
818 caught Italian phlebotomine sandflies (Diptera: Psychodidae) as detected by high
819 performance liquid chromatography. Bull Entomol Res. 1990;80(3):339–44.

820 14. Lovering AL, Seung SL, Kim YW, Withers SG, Strynadka NCJ. Mechanistic and
821 structural analysis of a family 31 a-glycosidase and its glycosyl-enzyme
822 intermediate. J Biol Chem. 2005;280(3):2105–15.

823 15. Cantarel B, Coutinho P, Rancurel C, Bernard T, Lombard V, Henrissat B. The
824 Carbohydrate-Active EnZymes database (CAZy): an expert resource for
825 Glycogenomics. Nucleic Acids Res. 2009;37:D233–8.

826 16. Ramasubbu N, Paloth V, Luo Y, Brayer GD, Levine MJ. Structure of Human
827 Salivary α-Amylase at 1.6 Å Resolution: Implications for its Role in the Oral
828 Cavity. Acta Crystallogr Sect D Biol Crystallogr. 1996;52(3):435–46.

829 17. Janecek S. Alpha-Amylase Family: Molecular Biology and Evolution. J Prog
830 Biophys Mol Biol. 1997;67(1):67–97.

831 18. Kanai R, Haga K, Yamane K, Harata K. Crystal structure of cyclodextrin
832 glucanotransferase from alkalophilic Bacillus sp. 1011 Complexed with 1-
833 deoxynojirimycin at 2.0 Å resolution. J Biochem. 2001;129(4):593–8.

834 19. Abad MC, Binderup K, Rios-Steiner J, Arni RK, Preiss J, Geiger JH. The x-ray
835 crystallographic structure of Escherichia coli branching enzyme. J Biol Chem.
836 2002;277(44):42164–70.

837 20. MacGregor EA, Janeček Š, Svensson B. Relationship of sequence and structure to
838 specificity in the α-amylase family of enzymes. Biochim Biophys Acta - Protein
839 Struct Mol Enzymol. 2001;1546(1):1–20.

840 21. Uitdehaag JCM, Mosi R, Kalk KH, Van der Veen BA, Dijkhuizen L, Withers SG,
841 et al. X-ray structures along the reaction pathway of cyclodextrin
842 glycosyltransferase elucidate catalysis in the α-amylase family. Nat Struct Biol.
843 1999;6(5):432–6.

844 22. Svensson B, Macgregor EA. Relationship of sequence and structure to speci ¢ city
845 in the K -amylase family of enzymes. 2001;1546:1–20.

846 23. Terra WR, Ferreira C. Biochemistry and Molecular Biology of Digestion. In:
847 Insect Molecular Biology and Biochemistry [Internet]. Elsevier B.V.; 2012. p.
848 365–418. Available from: http://dx.doi.org/10.1016/B978-0-12-384747-8.10011-
849 X

850 24. Boel E, Brady L, Brzozowski AM, Derewenda Z, Dodson GG, Jensen VJ, et al.
851 Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-A resolution
852 of two enzymes from Aspergillus. Biochemistry. 1990;29(26):6244–9.

853 25. Machius M, Wiegand G, Huber R. Crystal Structure of Calcium-depletedBacillus
854 licheniformisα-amylase at 2.2 Å Resolution. J Mol Biol [Internet].
855 1995;246(4):545–59. Available from:
856 http://linkinghub.elsevier.com/retrieve/pii/S0022283684701069

857 26. D'Amico S, Gerday C, Feller G. Structural similarities and evolutionary
858 relationships in chloride-dependent alpha-amylases. Gene [Internet].
859 2000;253(1):95–105. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/10925206

27. Ernst HA, Lo Leggio L, Willemoës M, Leonard G, Blum P, Larsen S. Structure of the Sulfolobus solfataricus α-Glucosidase: Implications for Domain Conservation and Substrate Recognition in GH31. J Mol Biol. 2006;358(4):1106–24.

28. Sim L, Quezada-Calvillo R, Sterchi EE, Nichols BL, Rose DR. Human Intestinal Maltase-Glucoamylase: Crystal Structure of the N-Terminal Catalytic Subunit and Basis of Inhibition and Substrate Specificity. J Mol Biol. 2008;375(3):782–92.

29. Lee SS, He S, Withers SG. Identification of the catalytic nucleophile of the Family 31 alpha-glucosidase from Aspergillus niger via trapping of a 5-fluoroglycosyl-enzyme intermediate. Biochem J. 2001;359:381–6.

30. Janeček Š, Svensson B, MacGregor EA. A remote but significant sequence homology between glycoside hydrolase clan GH-H and family GH31. FEBS Lett. 2007;581(7):1261–8.

31. Hoefsloot LH, Hoogeveen-Westerveld M, Kroos M a, van Beeumen J, Reuser a J, Oostra B a. Primary structure and processing of lysosomal alpha-glucosidase; homology with the intestinal sucrase-isomaltase complex. EMBO J. 1988;7(6):1697–704.

32. Trombetta ES, Fleming KG, Helenius A. Quaternary and domain structure of glycoprotein processing glucosidase II. Biochemistry. 2001;40(35):10717–22.

33. Trombetta ES, Simons JF, Helenius A. Endoplasmic reticulum glucosidase II is composed of a catalytic subunit, conserved from yeast to mammals, and a tightly bound noncatalytic HDEL- containing subunit. J Biol Chem. 1996;271(44):27509–16.

34. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: Sequence visualization and annotation. Bioinformatics. 2000;16(10):944–5.

35. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–80.

36. Petersen TN, Brunak S, Von Heijne G, Nielsen H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8:785–486.

37. Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol. 2001;305(3):567–80.

38. Eisenhaber B, Bork P, Eisenhaber F. Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase. Protein Eng Des Sel. 1998;11(12):1155–61.

39. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017;

40. Gupta R, Jung E, Brunak S. Prediction of N-glycosylation sites in human proteins.

902 Pacific Symp Biocomput. 2002;7:310–22.

903 41. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB,
904 Schjoldager KTBG, et al. Precision mapping of the human O-GalNAc
905 glycoproteome through SimpleCell technology. EMBO J. 2013;32(10):1478–88.

906 42. Bjellqvist B, Basse B, Olsen E, Celis JE. Reference points for comparisons of two-
907 dimensional maps of proteins from different human cell types defined in a pH scale
908 where isoelectric points correlate with polypeptide compositions. Electrophoresis.
909 1994;15(1):529–39.

910 43. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al.
911 Protein Identification and Analysis Tools on the ExPASy Server. Proteomics
912 Protoc Handb. 2005;571–607.

913 44. Hulo N, Bairoch A, Bulliard V, Cerutti L, Cuche BA, De castro E, et al. The 20
914 years of PROSITE. Nucleic Acids Res. 2008;36(S1):D245–D249.

915 45. Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, et al. InterPro
916 in 2017-beyond protein family and domain annotations. Nucleic Acids Res.
917 2017;45(D1):D190–9.

918 46. Toda H, Kondo K, Narita K. The Complete Amino Acid Sequence of Taka-
919 Amilase A. Proc Jpn Acad Ser B Phys Biol Sci. 1982;58:208–12.

920 47. Wongchawalit J, Yamamoto T, Nakai H, Kim Y-M, Sato N, Nishimoto M, et al.
921 Purification and characterization of alpha-glucosidase I from Japanese honeybee
922 (Apis cerana japonica) and molecular cloning of its cDNA. Biosci Biotechnol
923 Biochem. 2006;70(12):2889–98.

924 48. Bertran J, Werner A, Chillaron J, Nunes V, Biber J, Testar X, et al. Expression
925 cloning of a human renal cDNA that induces high affinity transport of L-cystine
926 shared with dibasic amino acids in Xenopus oocytes. J Biol Chem.
927 1993;268(20):14842–9.

928 49. Baecker P a, Greenberg E, Preiss J. Biosynthesis of bacterial glycogen. Primary
929 structure of Escherichia coli 1,4-alpha-D-glucan:1,4-alpha-D-glucan 6-alpha-D-
930 (1, 4-alpha-D-glucano)-transferase as deduced from the nucleotide sequence of the
931 glg B gene. J Biol Chem. 1986;261(19):8738–43.

932 50. Zhai L, Feng L, Xia L, Yin H, Xiang S. Crystal structure of glycogen debranching
933 enzyme and insights into its catalysis and disease-causing mutations. Nat Commun
934 [Internet]. 2016;7:1–11. Available from: http://dx.doi.org/10.1038/ncomms11229

935 51. Roig-Zamboni V, Cobucci-Ponzano B, Iacono R, Ferrara MC, Germany S, Bourne
936 Y, et al. Structure of human lysosomal acid α-glucosidase-A guide for the
937 treatment of Pompe disease. Nat Commun [Internet]. 2017;8(1). Available from:
938 http://dx.doi.org/10.1038/s41467-017-01263-3

939 52. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular
940 Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol.
941 2018;35(6):1547–9.

942 53. Moraes CS, Diaz-Albiter HM, Faria M do V, Sant'Anna MRV, Dillon RJ, Genta
943 FA. Expression pattern of glycoside hydrolase genes in Lutzomyia longipalpis
944 reveals key enzymes involved in larval digestion. Front Physiol. 2014;5

945 JUL(August):1–16.

946 54. Moraes CS, Aguiar-Martins K, Costa SG, Bates PA, Dillon RJ, Genta FA. Second
947 Blood Meal by Female Lutzomyia longipalpis: Enhancement by Oviposition and
948 Its Effects on Digestion, Longevity, and Leishmania Infection. Biomed Res Int.
949 2018;

950 55. Thornton B, Basu C. Real-time PCR (qPCR) primer design using free online
951 software. Biochem Mol Biol Educ. 2011;39(2):145–54.

952 56. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist
953 programmers. Methods Mol Biol [Internet]. 2000 [cited 2018 Aug 14];132:365–
954 86. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10547847

955 57. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image
956 analysis. Nat Methods. 2012;9(7):671–5.

957 58. Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT
958 method. Nat Protoc [Internet]. 2008;3(6):1101–8. Available from:
959 http://www.nature.com/doifinder/10.1038/nprot.2008.73

960 59. Cameron MM, Pessoa FAC, Vasconcelos AW, Ward RD. Sugar meal sources for
961 the phlebotomine sandfly Lutzomyia longipalpis in Ceará State, Brazil. Med Vet
962 Entomol. 1995;9(3):263–72.

963 60. Moraes CS, Lucena SA, Moreira BHS, Brazil RP, Gontijo NF, Genta FA.
964 Relationship between digestive enzymes and food habit of Lutzomyia longipalpis
965 (Diptera: Psychodidae) larvae: Characterization of carbohydrases and digestion of
966 microorganisms. J Insect Physiol. 2012;58:1136–45.

967 61. Vale VF, Moreira BH, Moraes CS, Pereira MH, Genta FA, Gontijo NF.
968 Carbohydrate digestion in Lutzomyia longipalpis' larvae (Diptera - Psychodidae).
969 J Insect Physiol [Internet]. 2012;58(10):1314–24. Available from:
970 http://dx.doi.org/10.1016/j.jinsphys.2012.07.005

971 62. Thompson SN. Trehalose - The Insect "Blood" Sugar. In: Advances in Insect
972 Physiology. Elsevier Ltd; 2003. p. 205–85.

973 63. Arrese EL, Soulages JL. Insect Fat Body: Energy, Metabolism, and Regulation.
974 Annu Rev Entomol [Internet]. 2010;55(1):207–25. Available from:
975 http://www.annualreviews.org/doi/10.1146/annurev-ento-112408-085356

976 64. Silva CP, Terra WR. An alpha-glucosidase from perimicrovillar membranes of
977 Dysdercus peruvianus (Hemiptera: Pyrrhocoridae) midgut cells. Purification and
978 properties. Insect Biochem Mol Biol. 1995;25(4):487–94.

979 65. Cristofoletti PT, Ribeiro AF, Deraison C, Rahbé Y, Terra WR. Midgut adaptation
980 and digestive enzyme distribution in a phloem feeding insect, the pea aphid
981 Acyrthosiphon pisum. J Insect Physiol. 2003;49:11–24.

982 66. Nishimoto M, Kubota M, Tsuji M, Mori H, Kimura A, Matsui H, et al. Purification
983 and substrate specificity of honeybee, Apis mellifera L., alpha-glucosidase III.
984 Biosci Biotechnol Biochem. 2001;65(7):1610–6.

985 67. da Costa SG, Bates PA, Dillon RJ, Genta FA. Characterization of α-glucosidases
986 from Lutzomyia longipalpis reveals independent hydrolysis systems for plant or

blood sugars. Front Physiol.

68. Gontijo NF, Almeida-Silva S, Cio F, Costa F, Mares-Guia ML, Williams P, et al. Lutzomyia longipalpis: pH in the Gut, Digestive Glycosidases, and Some Speculations upon Leishmania Development. Exp Parasitol. 1998;90:212–9.

69. Chng W bin A, Hietakangas V, Lemaitre B. Physiological Adaptations to Sugar Intake: New Paradigms from Drosophila melanogaster. Trends Endocrinol Metab [Internet]. 2017;28(2):131–42. Available from: http://dx.doi.org/10.1016/j.tem.2016.11.003

70. Yamada T, Habara O, Kubo H, Nishimura T. Fat body glycogen serves as a metabolic safeguard for the maintenance of sugar levels in *Drosophila*. Development [Internet]. 2018;145(6):dev158865. Available from: http://dev.biologists.org/lookup/doi/10.1242/dev.158865

71. Dillon RJ, El Kordy E. Carbohydrate digestion in sandflies: alpha-glucosidase activity in the midgut of Phlebotomus langeroni. Comp Biochem Physiol - B Biochem Mol Biol. 1997;116(1):35–40.

72. Mury FB, da Silva JR, Ferreira LS, dos Santos Ferreira B, de Souza-Filho GA, de Souza-Neto JA, et al. Alpha-glucosidase promotes hemozoin formation in a blood-sucking bug: An evolutionary history. PLoS One. 2009;4(9):e6966.

73. Telleria EL, de Araújo APO, Secundino NF, D'Avila-Levy CM, Traub-Csekö YM. Trypsin-like serine proteases in Lutzomyia longipalpis - expression, activity and possible modulation by Leishmania infantum chagasi. PLoS One. 2010;5(5):e10697.

74. Schlein Y, Romano H. Leishmania major and L. donovani: Effects on proteolytic enzymes of Phlebotomus papatasi (Diptera, Psychodidae). Exp Parasitol. 1986;62(3):376–80.

75. Dillon RJ, Lane RP. Influence of Leishmania infection on blood-meal digestion in the sandflies Phlebotomus papatasi and P. langeroni. Parasitol Res. 1993;79(6):492–6.

76. Santos VC, Vale VF, Silva SM, Nascimento AAS, Saab NAA, Soares RPP, et al. Host modulation by a parasite: How Leishmania infantum modifies the intestinal environment of Lutzomyia longipalpis to favor its development. PLoS One. 2014;9(11).

77. De Assis RR, Ibraim IC, Nogueira PM, Soares RP, Turco SJ. Glycoconjugates in New World species of Leishmania: Polymorphisms in lipophosphoglycan and glycoinositolphospholipids and interaction with hosts. Biochim Biophys Acta - Gen Subj. 2012;1820(9):1354–65.

78. Soares RPP, Macedo ME, Ropert C, Gontijo NF, Almeida IC, Gazzinelli RT, et al. Leishmania chagasi: Lipophosphoglycan characterization and binding to the midgut of the sand fly vector Lutzomyia longipalpis. Mol Biochem Parasitol. 2002;121(2):213–24.

79. Coelho-Finamore JM, Freitas VC, Assis RR, Melo MN, Novozhilova N, Secundino NF, et al. Leishmania infantum: Lipophosphoglycan intraspecific variation and interaction with vertebrate and invertebrate hosts. Int J Parasitol.

2011;41:333–42.

80. Sacks DL. Stage-specific binding of Leishmania donovani to the sand fly vector midgut is regulated by conformational changes in the abundant surface lipophosphoglycan. J Exp Med. 1995;181(2):685–97.

81. McConville MJ, Schnur LF, Jaffe C SP, McConville MJ, Schnur LF, Jaffe C, Schneider P, McConville MJ, Schnur LF, Jaffe C SP. Structure of Leishmania lipophosphoglycan: inter- and intra-specific polymorphism in Old World species. Biochem. 1995;310:807–18.

82. Ilg T, Etges R, Overath P, McConville MJ, Thomas-Oates J, Thomas J, et al. Structure of Leishmania mexicana lipophosphoglycan. J Biol Chem. 1992;267(10):6834–40.

83. McConville MJ, Collidge TAC, Ferguson MAJ, Schneider P. The glycoinositol phospholipids of Leishmania mexicana promastigotes. Evidence for the presence of three distinct pathways of glycolipid biosynthesis. J Biol Chem. 1993;268(21):15595–604.

84. Romão TP, De Melo Chalegre KD, Key S, Junqueira Ayres CF, Fontes De Oliveira CM, De-Melo-Neto OP, et al. A second independent resistance mechanism to Bacillus sphaericus binary toxin targets its α-glucosidase receptor in Culex quinquefasciatus. FEBS J. 2006;273(7):1556–68.

85. Darboux I, Nielsen-LeRoux C, Charles JF, Pauron D. The receptor of Bacillus sphaericus binary toxin in Culex pipiens (Diptera: Culicidae) midgut: Molecular cloning and expression. Insect Biochem Mol Biol. 2001;31(10):981–90.

86. Opota O, Charles JF, Warot S, Pauron D, Darboux I. Identification and characterization of the receptor for the Bacillus sphaericus binary toxin in the malaria vector mosquito, Anopheles gambiae. Comp Biochem Physiol - B Biochem Mol Biol. 2008;149(3):419–27.

87. Gabriško M. Evolutionary history of eukaryotic α-glucosidases from the α-amylase family. J Mol Evol. 2013;76(3):129–45.

88. Gabriško M, Janeček Š. Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 α-amylase family. FEBS J. 2009;276(24):7265–78.

89. Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly Lutzomyia longipalpis. Proc Natl Acad Sci U S A. 1999;96(26):15155–60.

90. Jacobson RL, Studentsky L, Schlein Y. Glycolytic and chitinolytic activities of Phlebotomus papatasi (Diptera: Psychodidae) from diverse ecological habitats. Folia Parasitol (Praha). 2007;54(4):301–9.

91. Ribeiro JMC, Rowton ED, Charlab R. Salivary amylase activity of the phlebotomine sand fly, Lutzomyia longipalpis. Insect Biochem Mol Biol. 2000;

92. Brayer GD, Luo Y, Withers SG. The structure of human pancreatic a-amylase at 1.8 A resolution and comparisons with related enzymes. Protein Sci. 1995;4(9):1730–42.

93. Buisson G, Duée E, Haser R, Payan F. Three dimensional structure of porcine pancreatic alpha-amylase at 2.9 A resolution. Role of calcium in structure and activity. EMBO J. 1987;6(13):3909–16.

94. Feller G, Bussy O Le, Houssier C, Gerday C. Structural and functional aspects of chloride binding to Alteromonas haloplanctis α-amylase. J Biol Chem. 1996;271(39):23836–41.

95. Terra WR, Ferreira C, Bioquimica D De, Quimica I De, Paulo UDS. Insect digestive enzymes : properties , compartmentalization and function. Comp Biochem Physiol Part B Comp Biochem. 1994;109(1):1–62.

96. Porath B, Gainullin VG, Cornec-Le Gall E, Dillinger EK, Heyer CM, Hopp K, et al. Mutations in GANAB, Encoding the Glucosidase IIα Subunit, Cause Autosomal-Dominant Polycystic Kidney and Liver Disease. Am J Hum Genet. 2016;98(6):1193–207.

97. Chen IHB, Huber M, Guan T, Bubeck A, Gerace L. Nuclear envelope transmembrane proteins (NETs) that are up-regulated during myogenesis. BMC Cell Biol. 2006;7:1–16.

98. Datta K, Guan T, Gerace L. NET37, a nuclear envelope transmembrane protein with glycosidase homology, is involved in myoblast differentiation. J Biol Chem. 2009;284(43):29666–76.

Supporting Information

S1 Table. Identifiers and accession numbers of Insect GH13 protein sequences used for the phylogenetic tree of Fig 3.

S2 Table. Identifiers and accession numbers of Insect GH31 protein sequences used for the phylogenetic tree of Fig 4.

S3 Table. List and sequences of the RT-qPCR primers used for PCR-amplification of cDNA sequences codifying for proteins belonging to GH13 and GH31 families of *L. longipalpis*.

S1 File. Protein sequences of *L. longipalpis* and *P. papatasi* from GH13 and GH31 families after sequence curation. The Xs represent missing amino acid in incomplete sequences.

S2 File. Exon and introns sequences of *L. longipalpis* proteins from GH13 and GH31 families after sequence curation. The Xs represent missing nucleotides in incomplete sequences. For missing introns sequences, size was estimated based on homology with complete orthologous genes.

Fig1. tif

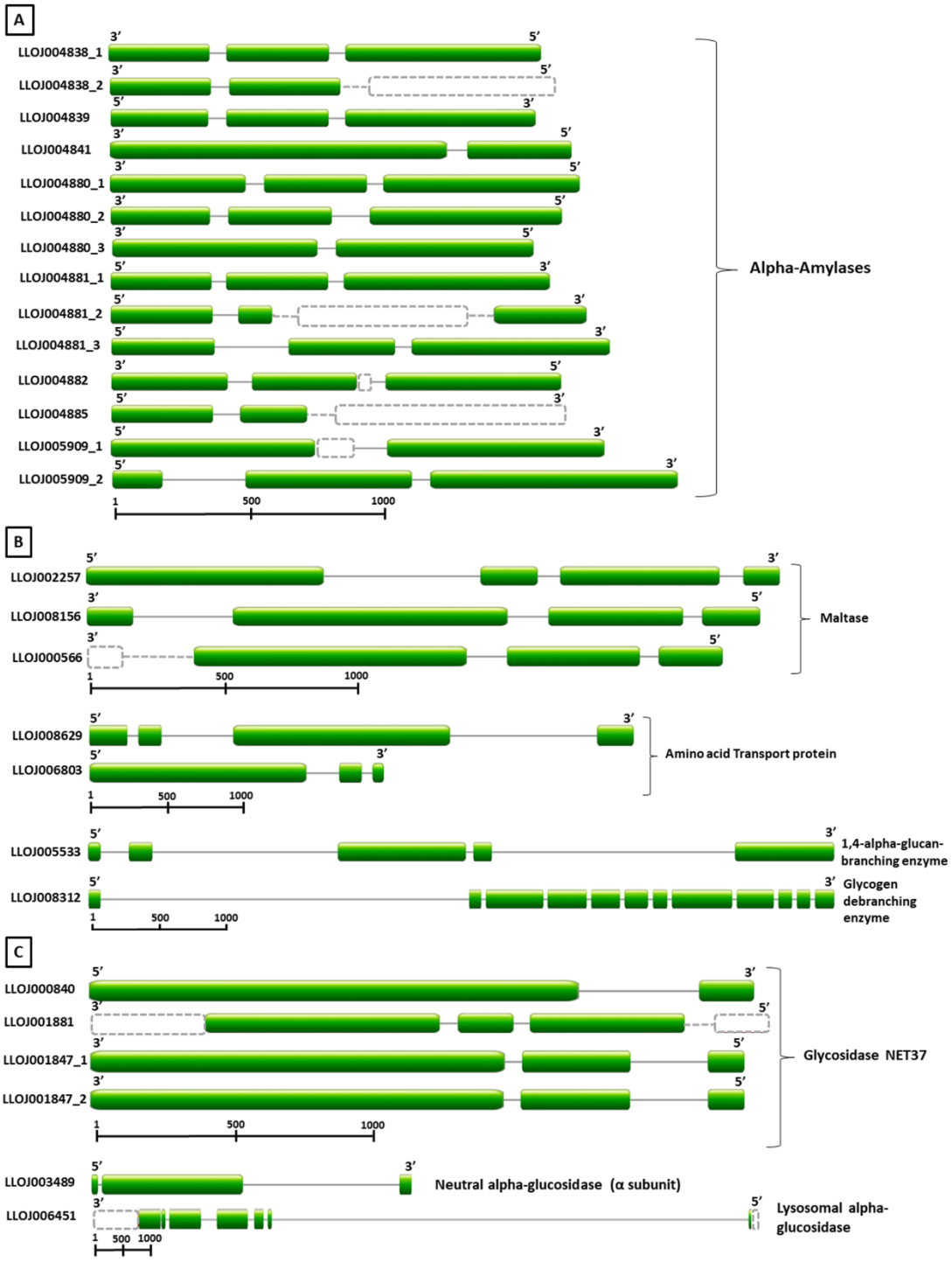


Fig 1. Schematic diagram of exon and intron composition of *L. longipalpis* genes from GH13 and GH31 families. Sequences were retrieved from Vector Base (Jacobina strain, LlonJ1.4 geneset, June 2017). Green boxes and lines represent exons and introns, dotted lines and boxes represent missing parts in incomplete sequences. For incomplete sequences, structures were designed based on homology with orthologous sequences. (A) Exon and intron composition of GH13 alpha-amylases. (B) Exon and intron composition of GH13 maltases, amino acid transport proteins, 1,4-alpha-glucan-branching-enzyme and glycogen debranching enzyme. (C) Exon and intron composition of GH31 glycosidases NET37, neutral alpha-glucosidase and lysosomal alpha-glucosidase.

Fig2.tif

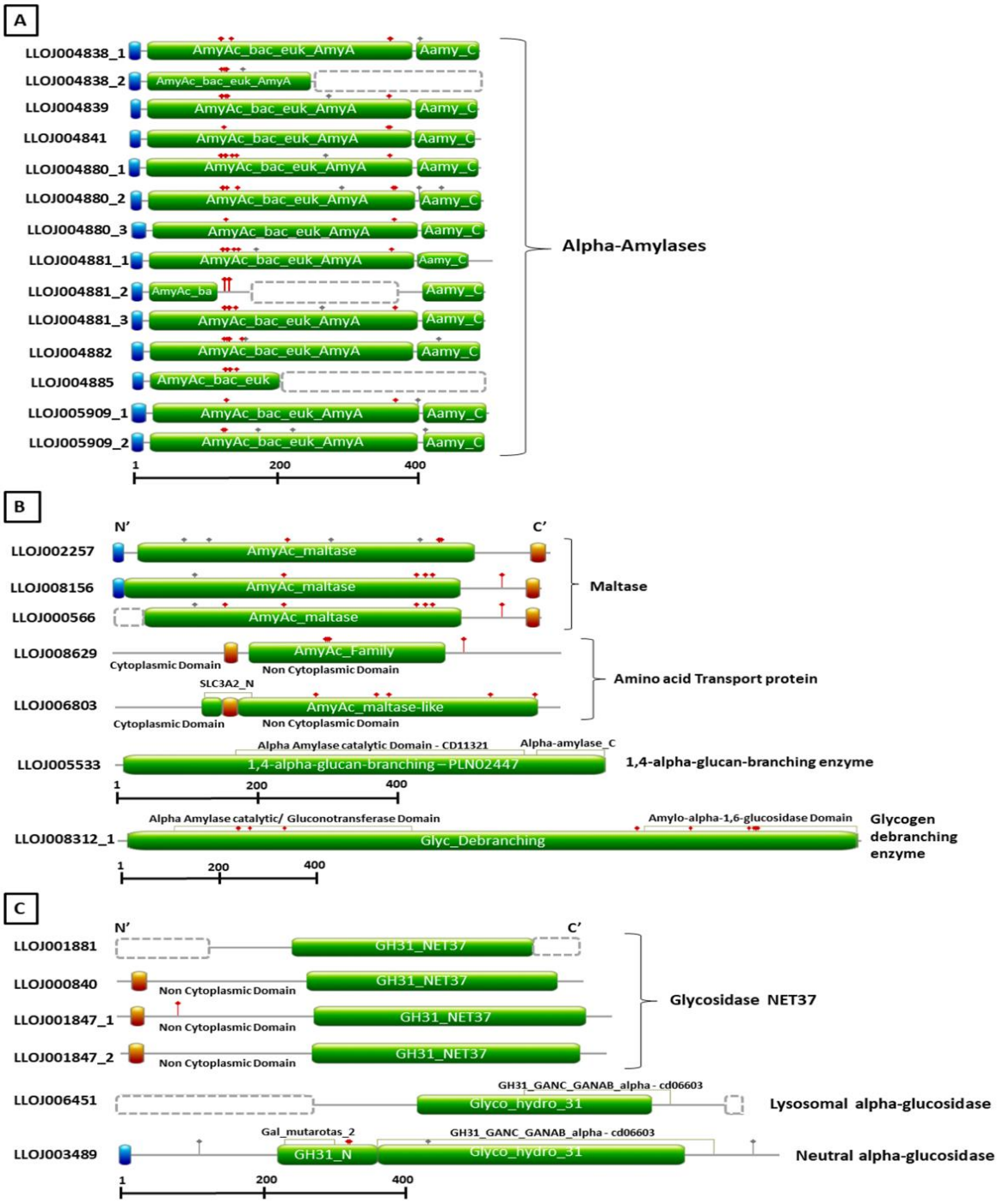


Fig 2. Schematic diagram of domain architecture of *L. longipalpis* proteins from GH13 and GH31 families. Sequences were retrieved from Vector Base (Jacobina strain, LlonJ1.4 geneset, June 2017). Signal peptide, transmembrane domains, and domains used to identify the different catalytic activities are boxed with the blue, red and green background, respectively. Grey and red dots above structures represent the N- and O- glycosylation. Dotted boxes represent missing parts in incomplete protein sequences. For incomplete protein structures, models were designed based on homology with orthologous protein sequences. The structures were outlined using as reference the best hits obtained with Blastp (refseq_protein) and Conserved Domain Database. (A) Domain composition of GH13 α-amylases. (B) Domain composition of GH13 maltases, amino acid transport proteins, 1,4-α-glucan-branching-enzyme and glycogen debranching enzyme. (C) Domain composition of GH31 glycosidases NET37, neutral α-glucosidase and lysosomal α-glucosidase.

Fig3.tif

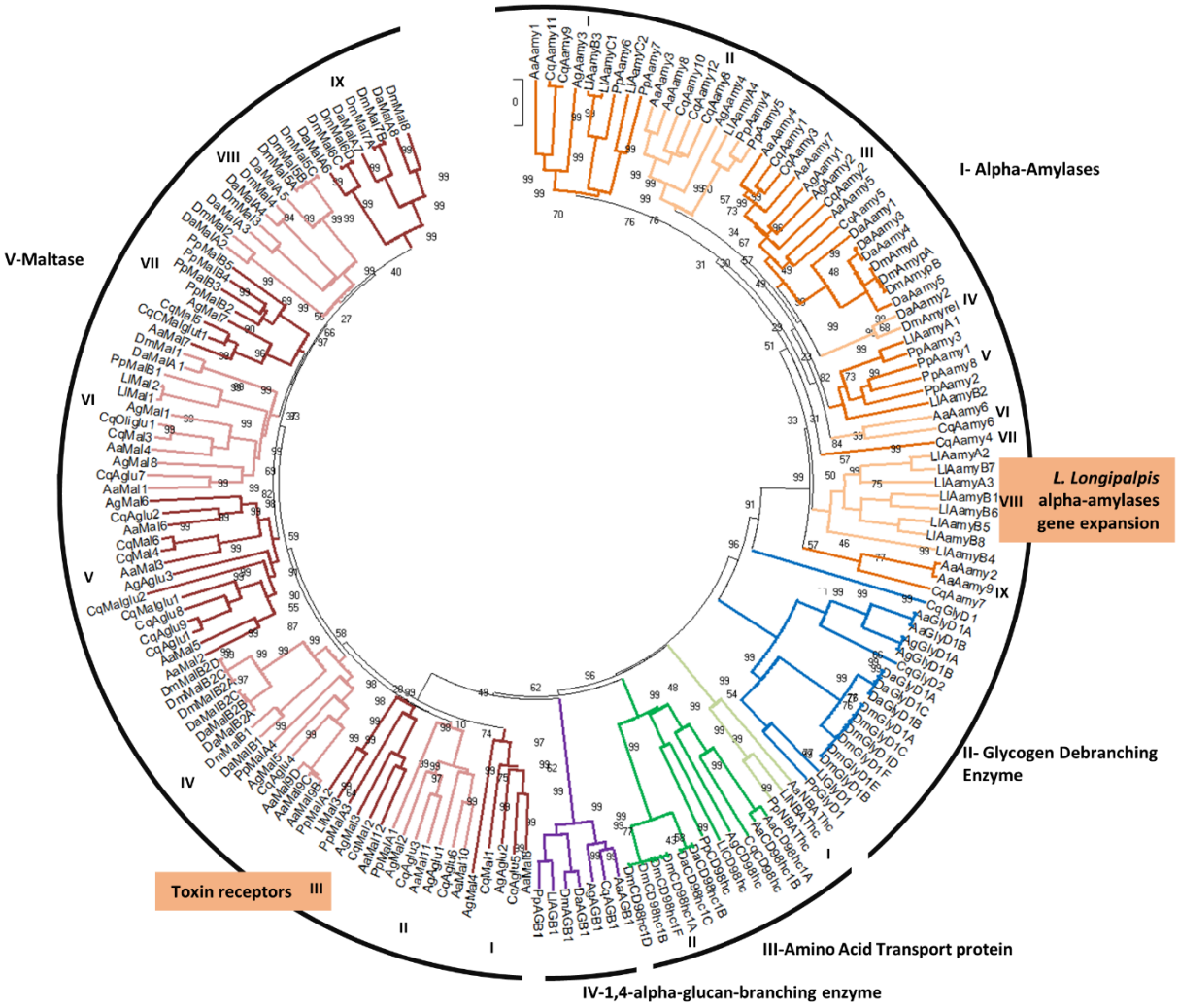


Fig 3. Phylogenetic tree of GH13 proteins from *L. longipalpis*. The Neighbor-Joining method was used to infer the evolutionary history using the amino acid sequences of GH13 family proteins retrieved from Vector Base, Fly Base and NCBI from the Dipteran species: *L. longipalpis*, *P. papatasi*, *A. aegypti*, *A. gambiae*, *C. quinquefasciatus*, *D. melanogaster* and *D. ananassae*. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) are demonstrated next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The p-distance method was used to compute the evolutionary distances, and these are represented in units of the number of amino acid differences per site. The analysis involved 157 amino acid sequences. For each sequence pair, all ambiguous positions were removed. There were a total of 1468 positions in the final dataset. A-amylase (Aamy), maltase (Mal), amino acid transport proteins (CD98hc and NBAThc), 1,4-α-branching enzyme (AGB), glycogen debranching enzyme (GDE). The identifiers and corresponding genes are specified in S1 Table.

Fig4.tif

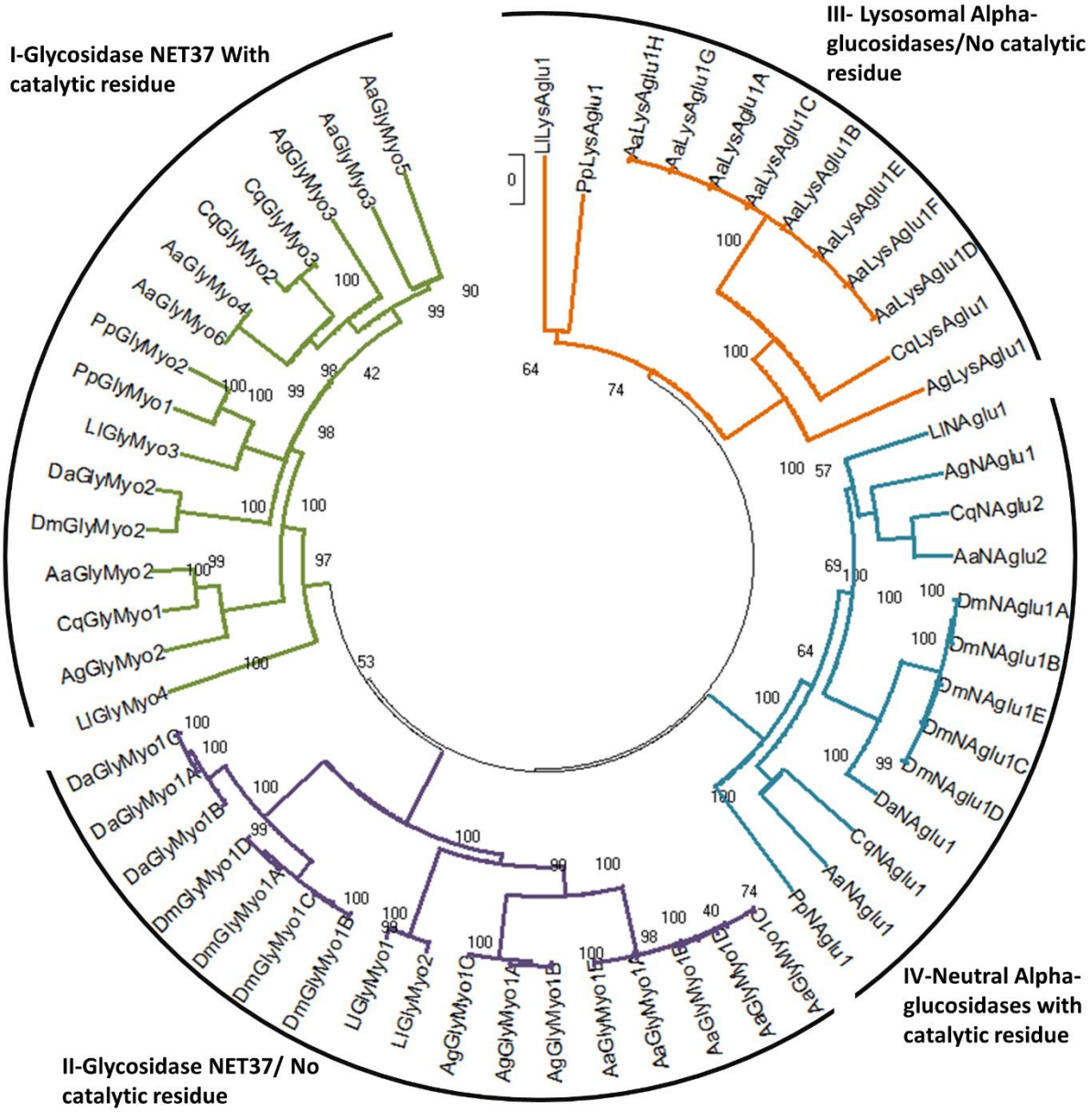


Fig 4. Phylogenetic tree of GH31 proteins from *L. longipalpis*. The Neighbor-Joining method was used to infer the evolutionary history using the amino acid sequences of GH31 family proteins retrieved from Vector Base, Fly Base and NCBI from Dipteran species: *L. longipalpis*, *P. papatasi*, *A. aegypti*, *A. gambiae*, *C. quinquefasciatus*, *D. melanogaster* and *D. ananassae*. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (10000 replicates) are demonstrated next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The p-distance method was used to compute the evolutionary distances, and these are represented in units of the number of amino acid differences per site. The analysis involved 58 amino acid sequences. For each sequence pair, all ambiguous positions were removed. There were a total of 1567 positions in the final dataset. Neutral α-glucosidase (NAglu), lysosomal α-glucosidase (LysAglu) and Glycosidase NET37 (GLyMyo). The identifiers and corresponding genes are specified in S2 Table.

Fig5.tif

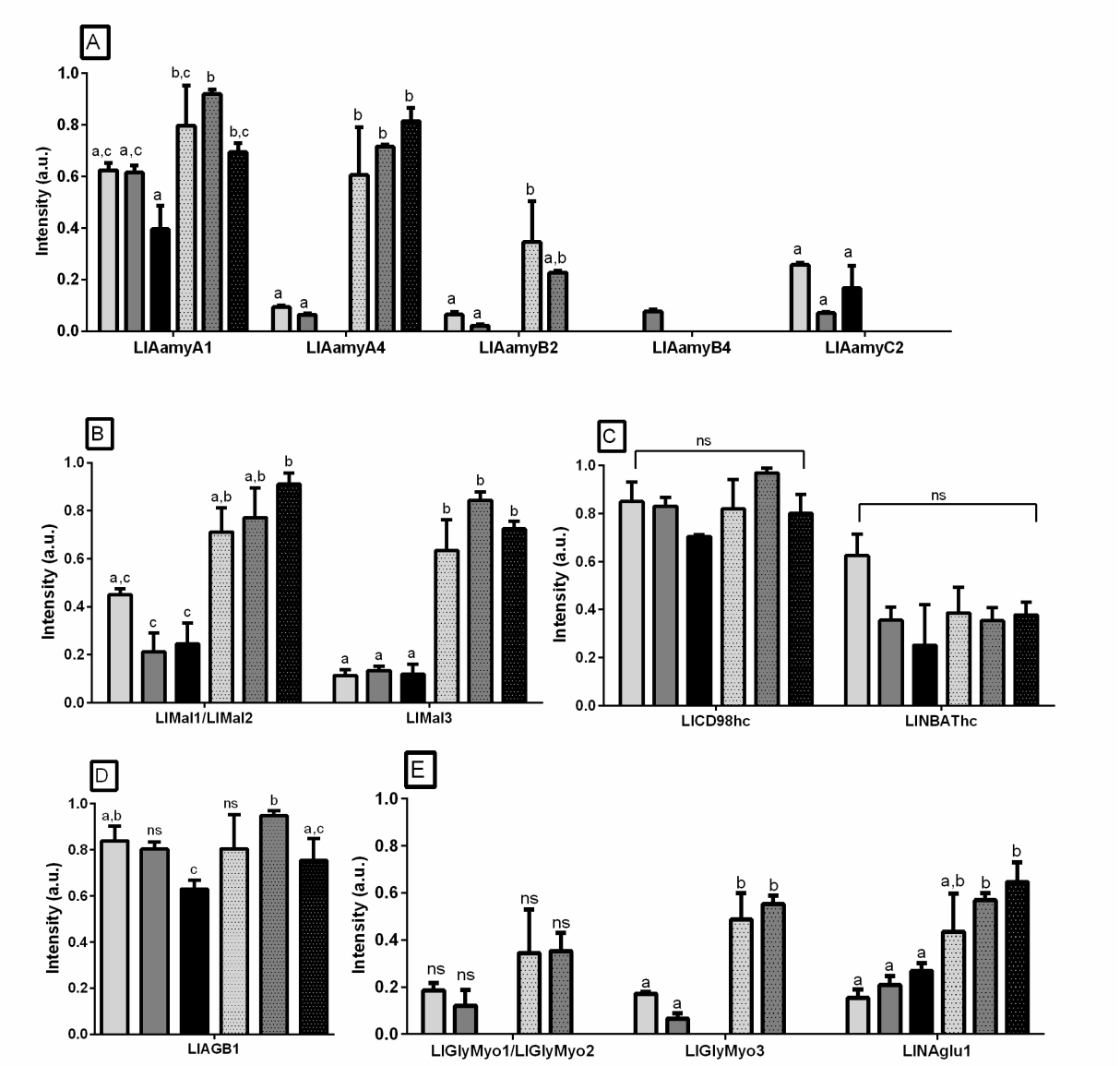


Fig 5. Expression pattern of genes belonging to the GH13 and GH31 families in *L. longipalpis* by RT-PCR. Expression was analyzed in different tissues with females submitted to different physiological conditions. The expression of genes from GH13 family was analyzed for (A) α-amylases 1, 4, 6, 8 and 14; (B) maltases 1/2 and 3; (C) neutral amino acid transport proteins 1 and 2 and (D) 1,4-glucan-branching enzyme. For GH31 family the expression pattern was analyzed for (E) glycosidase NET37 1-3 and neutral alpha-glucosidase 2. Samples from the rest of the body (plain bars) and gut (dotted bars) were analyzed for each physiological condition tested; newly hatched females (light grey bars), females fed on 1.2 M sucrose for 48 h (dark grey bars) and females after 48 h of blood feeding (black bars). Results are the mean ± SEM of three biological replicates each with n = 6. One-way ANOVA, followed by Tukey multiple comparison tests. Different letters indicate statistically significant differences in expression pattern, $p < 0.05$. **ns:** nonsignificant results.

Fig6.tif

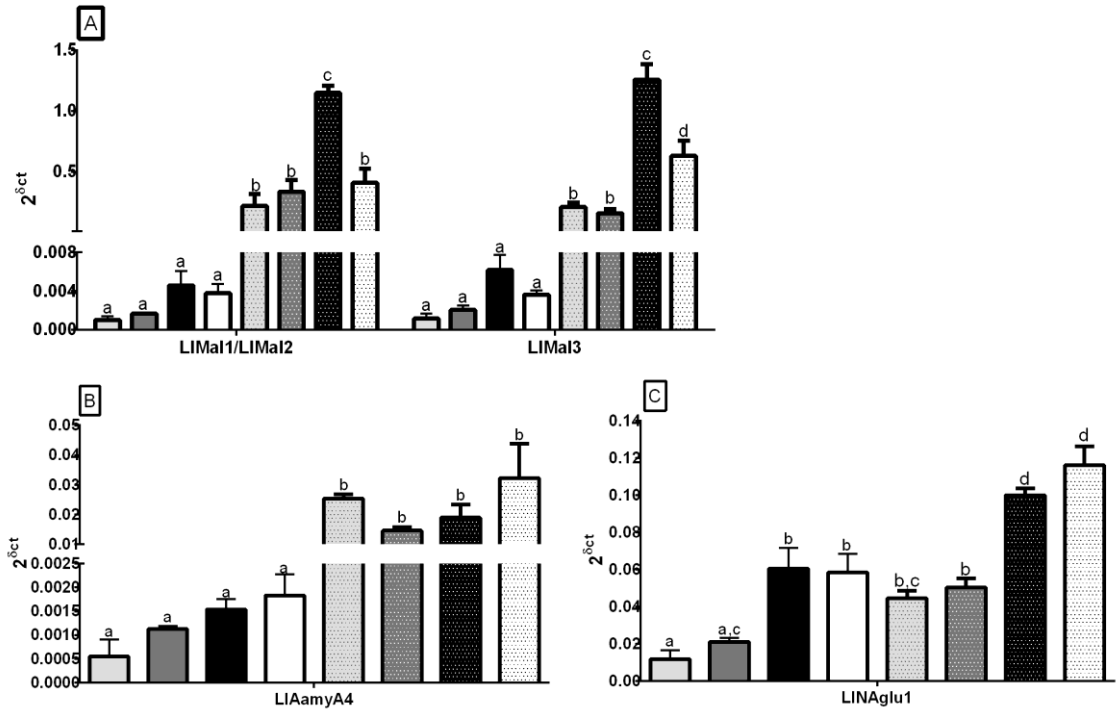


Fig 6. Expression pattern of genes belonging to the GH13 and GH31 families in *L. longipalpis* by qRT-PCR ($2^{\delta ct}$). Expression was analyzed in different tissues with females submitted to different physiological conditions. The expression of genes from GH13 family was analyzed for (A) maltases 1/2 and 3; (B) α-amylases 4. For GH31 family the expression pattern was analyzed for (C) neutral α-glucosidase 2. Samples from the rest of the body (plain bars) and gut (dotted bars) were analyzed, for each physiological condition tested; newly hatched females (light grey bars), females fed on 1.2 M sucrose for 48 h (dark grey bars), females after 48 h of blood feeding (black bars) and females after 48 h infection with *L. mexicana* (white bars). Results are the mean ± SEM of three biological replicates each with n = 6. One-way ANOVA, followed by Tukey multiple comparison tests. Different letters indicate statistically significant differences in expression pattern, $p < 0.05$. ns: nonsignificant results.

1 Additional File

S1 Table. Identifiers and accession numbers of Insect GH13 protein sequences used for the phylogenetic tree of Figure 3.

Specie	Vector base/ Fly base acession numbers	This Work	Current Status	Tree Code Identifier
Lutzomyia longipalpis	LLOJ004838_1	Alpha-Amylase	Alpha-Amylase	LlAamyA1
Lutzomyia longipalpis	LLOJ004838_2	Alpha-Amylase	Alpha-Amylase	LlAamyA2
Lutzomyia longipalpis	LLOJ004839	Alpha-Amylase	Alpha-Amylase	LlAamyA3
Lutzomyia longipalpis	LLOJ004841	Alpha-Amylase	Alpha-Amylase	LlAamyA4
Lutzomyia longipalpis	LLOJ004880_1	Alpha-Amylase	Alpha-Amylase	LlAamyB1
Lutzomyia longipalpis	LLOJ004880_2	Alpha-Amylase	Alpha-Amylase	LlAamyB2
Lutzomyia longipalpis	LLOJ004880_3	Alpha-Amylase	Alpha-Amylase	LlAamyB3
Lutzomyia longipalpis	LLOJ004881_1	Alpha-Amylase	Alpha-Amylase	LlAamyB4
Lutzomyia longipalpis	LLOJ004881_2	Alpha-Amylase	Alpha-Amylase	LlAamyB5
Lutzomyia longipalpis	LLOJ004881_3	Alpha-Amylase	Alpha-Amylase	LlAamyB6
Lutzomyia longipalpis	LLOJ004882	Alpha-Amylase	Alpha-Amylase	LlAamyB7
Lutzomyia longipalpis	LLOJ004885	Alpha-Amylase	Not Especified	LlAamyB8
Lutzomyia longipalpis	LLOJ005909_1	Alpha-Amylase	Alpha-Amylase	LlAamyC1
Lutzomyia longipalpis	LLOJ005909_2	Alpha-Amylase	Alpha-Amylase	LlAamyC2
Lutzomyia longipalpis	LLOJ000566	Maltase	Not Especified	LlMal1
Lutzomyia longipalpis	LLOJ008156	Maltase	Not Especified	LlMal2
Lutzomyia longipalpis	LLOJ002257	Maltase	Not Especified	LlMal3
Lutzomyia longipalpis	LLOJ008629	NBAT protein	Not Especified	LlNBAThc
Lutzomyia longipalpis	LLOJ006803	CD98 heavy chain	Not Especified	LlCD98hc
Lutzomyia longipalpis	LLOJ005533	1,4-alpha-glucan-branching enzyme	Not Especified	LlAGB1
Lutzomyia longipalpis	LLOJ008212	Glycogen debranching enzyme	Not Especified	LlGlyD1
Phlebotomus papatasi	PPAI0010148	Alpha-Amylase	Alpha-Amylase	PpAamy1
Phlebotomus papatasi	PPAI0010150_1	Alpha-Amylase	Alpha-Amylase	PpAamy2
Phlebotomus papatasi	PPAI0010150_2	Alpha-Amylase	Alpha-Amylase	PpAamy3
Phlebotomus papatasi	PPAI0010150_3	Alpha-Amylase	Alpha-Amylase	PpAamy4
Phlebotomus papatasi	PPAI0010150_4	Alpha-Amylase	Alpha-Amylase	PpAamy5

Phlebotomus papatasi	PPAI0010151_1	Alpha-Amylase	Alpha-Amylase	PpAamy6
Phlebotomus papatasi	PPAI0010151_2	Alpha-Amylase	Alpha-Amylase	PpAamy7
Phlebotomus papatasi	PPAI0010151_3	Alpha-Amylase	Alpha-Amylase	PpAamy8
Phlebotomus papatasi	PPAI0010255	Maltase	Not Especified	PpMalA1
Phlebotomus papatasi	PPAI0010256	Maltase	Not Especified	PpMalA2
Phlebotomus papatasi	PPAI0010257	Maltase	Not Especified	PpMalA3
Phlebotomus papatasi	PPAI0010258	Maltase	Not Especified	PpMalA4
Phlebotomus papatasi	PPAI0010759	Maltase	Not Especified	PpMalB1
Phlebotomus papatasi	PPAI0010760_1	Maltase	Not Especified	PpMalB2
Phlebotomus papatasi	PPAI0010760_2	Maltase	Not Especified	PpMalB3
Phlebotomus papatasi	PPAI0010760_3	Maltase	Not Especified	PpMalB4
Phlebotomus papatasi	PPAI0010760_4	Maltase	Not Especified	PpMalB5
Phlebotomus papatasi	PPAI001789	NBAT protein	Not Especified	PpNBAThc
Phlebotomus papatasi	PPAI004854	CD98 heavy chain	Not Especified	PpCD98hc
Phlebotomus papatasi	PPAI008963	1,4-alpha-glucan-branching enzyme	Alpha-Amylase	PpAGB1
Phlebotomus papatasi	PPAI009397	Glycogen debranching enzyme	Not Especified	PpGlyD1
Aedes aegypti	AAEL006719-PA	Alpha-Amylase	Alpha-Amylase (Amy1)	AaAamy1
Aedes aegypti	AAEL001130-PA	Alpha-Amylase	Alpha-Amylase	AaAamy2
Aedes aegypti	AAEL007673-PA	Alpha-Amylase	Alpha-Amylase	AaAamy3
Aedes aegypti	AAEL008451-PA	Alpha-Amylase	Alpha-Amylase	AaAamy4
Aedes aegypti	AAEL008452-PA	Alpha-Amylase	Alpha-Amylase	AaAamy5
Aedes aegypti	AAEL008456-PA	Alpha-Amylase	Alpha-Amylase	AaAamy6
Aedes aegypti	AAEL013421-PA	Alpha-Amylase	Alpha-Amylase	AaAamy7
Aedes aegypti	AAEL026011-PA	Alpha-Amylase	Not Especified	AaAamy8
Aedes aegypti	AAEL026503-PA	Alpha-Amylase	Not Especified	AaAamy9
Aedes aegypti	AAEL009524-PA	Maltase	Alpha-Glycosidase (Mal1)[a]	AaMal1
Aedes aegypti	AAEL000647-PA	Maltase	Alpha-Amylase	AaMal2
Aedes aegypti	AAEL000651-PA	Maltase	Alpha-Amylase	AaMal3
Aedes aegypti	AAEL000667-PB	Maltase	Alpha-Amylase	AaMal4
Aedes aegypti	AAEL000678-PA	Maltase	Alpha-Amylase	AaMal5

Aedes aegypti	AAEL003434-PA	Maltase	Alpha-Amylase	AaMal6
Aedes aegypti	AAEL000642-PA	Maltase	Alpha-Amylase	AaMal7
Aedes aegypti	AAEL010532-PA	Maltase	Alpha-Amylase	AaMal8
Aedes aegypti	AAEL010536-PB	Maltase	Alpha-Amylase	AaMal9B
Aedes aegypti	AAEL010536-PC	Maltase	Alpha-Amylase	AaMal9C
Aedes aegypti	AAEL010536-PD	Maltase	Alpha-Amylase	AaMal9D
Aedes aegypti	AAEL010540-PA	Maltase	Alpha-Amylase	AaMal10
Aedes aegypti	AAEL017128-PB	Maltase	Not Especified	AaMal11
Aedes aegypti	AAEL010537-PA	Maltase	Maltase[b]	AaMal12
Aedes aegypti	AAEL008502-PA	NBAT protein	Not Especified	AaNBAThc
Aedes aegypti	AAEL002527-PA	CD98 heavy chain	Not Especified	AaCD98hc1A
Aedes aegypti	AAEL002527-PB	CD98 heavy chain	Not Especified	AaCD98hc1B
Aedes aegypti	AAEL010602-PA	1,4-alpha-glucan-branching enzyme	Starch branching enzyme	AaAGB1
Aedes aegypti	AAEL009838-PA	Glycogen debranching enzyme	Glycogen debranching enzyme	AaGlyD1A
Aedes aegypti	AAEL009838-PB	Glycogen debranching enzyme	Glycogen debranching enzyme	AaGlyD1B
Anopheles gambiae	AGAP002317-PA	Alpha-Amylase	Alpha-Amylase	AgAamy1
Anopheles gambiae	AGAP002318-PA	Alpha-Amylase	Alpha-Amylase	AgAamy2
Anopheles gambiae	AGAP006371-PA	Alpha-Amylase	Alpha-Amylase (Amyrel)[a]	AgAamy3
Anopheles gambiae	AGAP012230-PA	Alpha-Amylase	Alpha-Amylase (Amyrel)	AgAamy4
Anopheles gambiae	AGAP008962-PA	Alpha-glucosidase	Alpha-glucosidase	AgAglu1
Anopheles gambiae	AGAP008963-PA	Alpha-glucosidase	Not Especified	AgAglu2
Anopheles gambiae	AGAP012400-PA	Alpha-glucosidase	Alpha-glucosidase	AgAglu3
Anopheles gambiae	AGAP012401-PA	Maltase	Maltase (AGM1)	AgMal1
Anopheles gambiae	AGAP006019-PA	Maltase	Maltase	AgMal2
Anopheles gambiae	AGAP008961-PA	Maltase	Maltase[b]	AgMal3
Anopheles gambiae	AGAP008964-PA	Maltase	Maltase	AgMal4
Anopheles gambiae	AGAP008965-PA	Maltase	Maltase	AgMal5
Anopheles gambiae	AGAP011939-PA	Maltase	Not Especified	AgMal6
Anopheles gambiae	AGAP012399-PA	Maltase	Maltase	AgMal7
Anopheles gambiae	AGAP002102-PA	Maltase	Alpha-glucosidase[a]	AgMal8
Anopheles gambiae	AGAP000022-PA	CD98 heavy chain	Not Especified	AgCD98hc
Anopheles gambiae	AGAP010428-PA	1,4-alpha-glucan-branching enzyme	1,4-alpha-glucan-branching enzyme	AgAGB1

Anopheles gambiae	AGAP001200-PA	Glycogen debranching enzyme	Glycogen debranching enzyme	AgGlyD1A
Anopheles gambiae	AGAP001200-PB	Glycogen debranching enzyme	Glycogen debranching enzyme	AgGlyD1B
Culex quinquefasciatus	CPIJ008079-PA	Alpha-Amylase	Alpha-Amylase 1	CqAamy1
Culex quinquefasciatus	CPIJ005060-PA	Alpha-Amylase	Alpha-Amylase B	CqAamy2
Culex quinquefasciatus	CPIJ005061-PA	Alpha-Amylase	Alpha-Amylase B	CqAamy3
Culex quinquefasciatus	CPIJ005062-PA	Alpha-Amylase	Alpha-Amylase	CqAamy4
Culex quinquefasciatus	CPIJ005064-PA	Alpha-Amylase	Alpha-Amylase	CqAamy5
Culex quinquefasciatus	CPIJ005065-PA	Alpha-Amylase	Alpha-Amylase 2	CqAamy6
Culex quinquefasciatus	CPIJ005700-PA	Alpha-Amylase	Alpha-Amylase B	CqAamy7
Culex quinquefasciatus	CPIJ005725-PA	Alpha-Amylase	Alpha-Amylase A	CqAamy8
Culex quinquefasciatus	CPIJ007333-PA	Alpha-Amylase	Amylase	CqAamy9
Culex quinquefasciatus	CPIJ001464-PA	Alpha-Amylase	Alpha-Amylase A	CqAamy10
Culex quinquefasciatus	CPIJ017521-PA	Alpha-Amylase	Alpha-Amylase I	CqAamy11
Culex quinquefasciatus	CPIJ018222-PA	Alpha-Amylase	Alpha-Amylase B	CqAamy12
Culex quinquefasciatus	CPIJ005213-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu1
Culex quinquefasciatus	CPIJ010128-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu2
Culex quinquefasciatus	CPIJ012204-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu3
Culex quinquefasciatus	CPIJ013169-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu4
Culex quinquefasciatus	CPIJ013171-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu5
Culex quinquefasciatus	CPIJ013172-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu6
Culex quinquefasciatus	CPIJ016362-PA	Alpha-glucosidase	Alpha-glucosidase[a]	CqAglu7
Culex quinquefasciatus	CPIJ019691-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu8
Culex quinquefasciatus	CPIJ019693-PA	Alpha-glucosidase	Alpha-glucosidase	CqAglu9
Culex quinquefasciatus	CPIJ013170-PA	Maltase	Maltase 1	CqMal1
Culex quinquefasciatus	CPIJ013173-PA	Maltase	NBAT protein[b]	CqMal2
Culex quinquefasciatus	CPIJ019692-PA	Maltase	Alpha-amylase	CqMal3
Culex quinquefasciatus	CPIJ005208-PA	Maltase	Alpha-Amylase	CqMal4
Culex quinquefasciatus	CPIJ005210-PA	Maltase	Alpha-Amylase	CqMal5
Culex quinquefasciatus	CPIJ018570-PA	Maltase	Alpha-Amylase	CqMal6
Culex quinquefasciatus	CPIJ005212-PA	Oligo-1,6-glucosidase	Oligo-1,6-glucosidase	CqOliglu1
Culex quinquefasciatus	CPIJ019689-PA	Cyclomaltodextrin glucanotransferase	Cyclomaltodextrin glucanotransferase	CqCMalglut1

Culex quinquefasciatus	CPIJ019690-PA	Maltodextrin glucosidase	Maltodextrin glucosidase	CqMalglu1
Culex quinquefasciatus	CPIJ005211-PA	Maltodextrin glucosidase	Alpha-Amylase 2	CqMalglu2
Culex quinquefasciatus	CPIJ011854-PA	CD98 heavy chain	CD98hc amino acid transporter protein	CqCD98hc
Culex quinquefasciatus	CPIJ006166-PA	1,4-alpha-glucan-branching enzyme	Deltamethrin resistance-associated NYD-GBE	CqAGB1
Culex quinquefasciatus	CPIJ013040-PA	Glycogen debranching enzyme	Glycogen debranching enzyme	CqGlyD1
Culex quinquefasciatus	CPIJ020026-PA	Glycogen debranching enzyme	Glycogen debranching enzyme	CqGlyD2
Drosophila melanogaster	FBpp0086263	Alpha-Amylase	Amyrel-PA	DmAmyrel
Drosophila melanogaster	FBpp0086136	Alpha-Amylase	Amy-d-PA	DmAmyd
Drosophila melanogaster	FBpp0086155	Alpha-Amylase	Amy-p-PA	DmAmypA
Drosophila melanogaster	FBpp0311627	Alpha-Amylase	Amy-p-PB	DmAmypB
Drosophila melanogaster	FBpp0087838	Maltase	Mal-A1-PA	DmMal1
Drosophila melanogaster	FBpp0087826	Maltase	Mal-A2-PA	DmMal2
Drosophila melanogaster	FBpp0087837	Maltase	Mal-A3-PA	DmMal3
Drosophila melanogaster	FBpp0087827	Maltase	Mal-A4-PA	DmMal4
Drosophila melanogaster	FBpp0087828	Maltase	Mal-A5-PA	DmMal5A
Drosophila melanogaster	FBpp0290175	Maltase	Mal-A5-PB	DmMal5B
Drosophila melanogaster	FBpp0308388	Maltase	Mal-A5-PC	DmMal5C
Drosophila melanogaster	FBpp0271830	Maltase	Mal-A6-PC	DmMal6C
Drosophila melanogaster	FBpp0312227	Maltase	Mal-A6-PD	DmMal6D
Drosophila melanogaster	FBpp0087836	Maltase	Mal-A7-PA	DmMal7A
Drosophila melanogaster	FBpp0311600	Maltase	Mal-A7-PB	DmMal7B
Drosophila melanogaster	FBpp0087831	Maltase	Mal-A8-PA	DmMal8
Drosophila melanogaster	FBpp0079859	Maltase	Mal-B1-PA[a]	DmMalB1
Drosophila melanogaster	FBpp0079860	Maltase	Mal-B2-PA	DmMalB2A
Drosophila melanogaster	FBpp0292312	Maltase	Mal-B2-PC	DmMalB2C
Drosophila melanogaster	FBpp0292313	Maltase	Mal-B2-PD	DmMalB2D
Drosophila melanogaster	FBpp0081314	CD98 heavy chain	CD98 heavy chain, isoform A	DmCD98hc1A
Drosophila melanogaster	FBpp0303760	CD98 heavy chain	CD98 heavy chain, isoform B	DmCD98hc1B
Drosophila melanogaster	FBpp0303762	CD98 heavy chain	CD98 heavy chain, isoform D	DmCD98hc1D
Drosophila melanogaster	FBpp0304466	CD98 heavy chain	CD98 heavy chain, isoform F	DmCD98hc1F
Drosophila melanogaster	FBpp0086845	1,4-alpha-glucan-branching enzyme	AGBE-PA	DmAGB1
Drosophila melanogaster	FBpp0071525	Glycogen debranching enzyme	CG9485-PA	DmGlyD1A

Drosophila melanogaster	FBpp0071526	Glycogen debranching enzyme	CG9485-PB	DmGlyD1B
Drosophila melanogaster	FBpp0071527	Glycogen debranching enzyme	CG9485-PC	DmGlyD1C
Drosophila melanogaster	FBpp0289623	Glycogen debranching enzyme	CG9485-PD	DmGlyD1D
Drosophila melanogaster	FBpp0289624	Glycogen debranching enzyme	CG9485-PE	DmGlyD1E
Drosophila melanogaster	FBpp0302010	Glycogen debranching enzyme	CG9485-PF	DmGlyD1F
Drosophila ananassae	FBpp0115773	Alpha-Amylase	Dana\Amyc1-PA	DaAamy1
Drosophila ananassae	FBpp0116650	Alpha-Amylase	Dana\Amyc6-PA	DaAamy2
Drosophila ananassae	FBpp0122036	Alpha-Amylase	Dana\Amy35-PA	DaAamy3
Drosophila ananassae	FBpp0122035	Alpha-Amylase	Dana\Amy58-PA	DaAamy4
Drosophila ananassae	FBpp0344793	Alpha-Amylase	Dana\GF27675-PA	DaAmy5
Drosophila ananassae	FBpp0115779	Maltase	Dana\Mal-A1-PA	DaMalA1
Drosophila ananassae	FBpp0115552	Maltase	Dana\Mal-A2-PA	DaMalA2
Drosophila ananassae	FBpp0115778	Maltase	Dana\Mal-A3-PA	DaMalA3
Drosophila ananassae	FBpp0115553	Maltase	Dana\Mal-A4-PA	DaMalA4
Drosophila ananassae	FBpp0115554	Maltase	Dana\Mal-A5-PA	DaMalA5
Drosophila ananassae	FBpp0115555	Maltase	Dana\Mal-A6-PA	DaMalA6
Drosophila ananassae	FBpp0115777	Maltase	Dana\Mal-A7-PA	DaMalA7
Drosophila ananassae	FBpp0115556	Maltase	Dana\Mal-A8-PA	DaMalA8
Drosophila ananassae	FBpp0124798	Maltase	Dana\Mal-B1-PA	DaMalB1
Drosophila ananassae	FBpp0124799	Maltase	Dana\Mal-B2-PA	DaMAlB2A
Drosophila ananassae	FBpp0344084	Maltase	Dana\Mal-B2-PB	DaMAlB2B
Drosophila ananassae	FBpp0349800	Maltase	Dana\Mal-B2-PC	DaMAlB2C
Drosophila ananassae	FBpp0342698	CD98 heavy chain	Dana\GF16561-PB	DaCD98hc1B
Drosophila ananassae	FBpp0342739	CD98 heavy chain	Dana\GF16561-PC	DaCD98hc1C
Drosophila ananassae	FBpp0115136	1,4-alpha-glucan-branching enzyme	Dana\GF11944-PA	DaAGB1
Drosophila ananassae	FBpp0116003	Glycogen debranching enzyme	Dana\GF12811-PA	DaGlyD1A
Drosophila ananassae	FBpp0346541	Glycogen debranching enzyme	Dana\GF12811-PB	DaGlyD1B
Drosophila ananassae	FBpp0351281	Glycogen debranching enzyme	Dana\GF12811-PC	DaGlyD1C

NBAT -Neutral and basic amino acid transport protein

a- Salivary proteins

b- Toxin receptors

S2 Table. Identifiers and accession numbers of Insect GH31 protein sequences used for the phylogenetic tree of Figure 4.

Specie	Vector base/ Fly base acession Number	This Work	Current Status	Tree Code Identifiers	Catalytic Residues
Lutzomyia longipalpis	LLOJ001847_1	Glycosidase NET37	Not Especified	LlGlyMyo1	No
Lutzomyia longipalpis	LLOJ001847_2	Glycosidase NET37	Not Especified	LlGlyMyo2	No
Lutzomyia longipalpis	LLOJ001881	Glycosidase NET37	Not Especified	LlGlyMyo3	Yes
Lutzomyia longipalpis	LLOJ000840	Glycosidase NET37	Not Especified	LlGlyMyo4	Yes
Lutzomyia longipalpis	LLOJ006451	Lysosomal alpha-glucosidase	Not Especified	LlLysAglu1	No
Lutzomyia longipalpis	LLOJ003489	Neutral alpha-glucosidase	Not Especified	LlNAglu1	Yes
Phlebotomus papatasi	PPAI005397	Glycosidase NET37	Not Especified	PpGlyMyo1	Yes
Phlebotomus papatasi	PPAI007193	Glycosidase NET37	Not Especified	PpGlyMyo2	Yes
Phlebotomus papatasi	PPAI002935	Lysosomal alpha-glucosidase	Not Especified	PpLysAglu1	No
Phlebotomus papatasi	PPAI000884	Neutral alpha-glucosidase	Not Especified	PpNAglu1	Yes
Aedes Aegypti	AAEL005481-PA	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo1A	No
Aedes Aegypti	AAEL005481-PB	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo1B	No
Aedes Aegypti	AAEL005481-PC	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo1C	No
Aedes Aegypti	AAEL005481-PD	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo1D	No
Aedes Aegypti	AAEL005481-PE	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo1E	No
Aedes Aegypti	AAEL000223-PA	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo5	Yes
Aedes Aegypti	AAEL004361-PB	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo2	Yes
Aedes Aegypti	AAEL004367-PA	Glycosidase NET37	Not Especified	AaGlyMyo3	Yes
Aedes Aegypti	AAEL004369-PA	Glycosidase NET37	Alpha-glucosidase	AaGlyMyo4	Yes
Aedes Aegypti	AAEL020371-PA	Glycosidase NET37	Not Especified	AaGlyMyo6	Yes
Aedes Aegypti	AAEL017226-PA	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1A	No
Aedes Aegypti	AAEL017226-PB	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1B	No
Aedes Aegypti	AAEL017226-PC	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1C	No
Aedes Aegypti	AAEL017226-PD	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1D	No
Aedes Aegypti	AAEL017226-PE	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1E	No
Aedes Aegypti	AAEL017226-PF	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1F	No
Aedes Aegypti	AAEL017226-PG	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1G	No

Aedes Aegypti	AAEL017226-PH	Lysosomal alpha-glucosidase	Not Especified	AaLysAglu1H	No
Aedes Aegypti	AAEL022548-PA	Neutral alpha-glucosidase	Not Especified	AaNAglu1	Yes
Aedes Aegypti	AAEL022931-PA	Neutral alpha-glucosidase	Not Especified	AaNAglu2	Yes
Anopheles gambiae	AGAP001111-PA	Glycosidase NET37	Not Especified	AgGlyMyo1A	No
Anopheles gambiae	AGAP001111-PB	Glycosidase NET37	Not Especified	AgGlyMyo1B	No
Anopheles gambiae	AGAP001111-PC	Glycosidase NET37	Not Especified	AgGlyMyo1C	No
Anopheles gambiae	AGAP003993-PA	Glycosidase NET37	Not Especified	AgGlyMyo2	Yes
Anopheles gambiae	AGAP003995-PA	Glycosidase NET37	Not Especified	AgGlyMyo3	Yes
Anopheles gambiae	AGAP001534-PA	Lysosomal alpha-glucosidase	Not Especified	AgLysAglu1	No
Anopheles gambiae	AGAP000862-PA	Neutral alpha-glucosidase	Alpha 1,3-glucosidase	AgNAglu1	Yes
Culex quinquefasciatus	CPIJ008904-PA	Glycosidase NET37	Alpha-glucosidase, putative	CqGlyMyo1	Yes
Culex quinquefasciatus	CPIJ016015-PA	Glycosidase NET37	Alpha-glucosidase, putative	CqGlyMyo2	Yes
Culex quinquefasciatus	CPIJ019915-PA	Glycosidase NET37	Alpha-glucosidase, putative	CqGlyMyo3	Yes
Culex quinquefasciatus	CPIJ001201-PA	Lysosomal alpha-glucosidase	Neutral alpha-glucosidase ab precursor	CqLysAglu1	No
Culex quinquefasciatus	CPIJ009306-PA	Neutral alpha-glucosidase	Neutral alpha-glucosidase ab precursor	CqNAglu1	Yes
Culex quinquefasciatus	CPIJ009308-PA	Neutral alpha-glucosidase	Neutral alpha-glucosidase ab precursor	CqNAglu2	Yes
Drosophila melanogaster	FBpp0070809	Glycosidase NET37	CG33080-PA	DmGlyMyo1A	No
Drosophila melanogaster	FBpp0070810	Glycosidase NET37	CG33080-PB	DmGlyMyo1B	No
Drosophila melanogaster	FBpp0309736	Glycosidase NET37	CG33080-PC	DmGlyMyo1C	No
Drosophila melanogaster	FBpp0309737	Glycosidase NET37	CG33080-PD	DmGlyMyo1D	No
Drosophila melanogaster	FBpp0084221	Glycosidase NET37	tobi-PA	DmGlyMyo2	Yes
Drosophila melanogaster	FBpp0070059	Neutral alpha-glucosidase	GCS2alpha-PA	DmNAglu1A	Yes
Drosophila melanogaster	FBpp0070057	Neutral alpha-glucosidase	GCS2alpha-PB	DmNAglu1B	Yes
Drosophila melanogaster	FBpp0070058	Neutral alpha-glucosidase	GCS2alpha-PC	DmNAglu1C	Yes
Drosophila melanogaster	FBpp0070061	Neutral alpha-glucosidase	GCS2alpha-PD	DmNAglu1D	Yes
Drosophila melanogaster	FBpp0070060	Neutral alpha-glucosidase	GCS2alpha-PE	DmNAglu1E	Yes
Drosophila ananassae	FBpp0124661	Glycosidase NET37	Dana\GF21469-PA	DaGlyMyo1A	No
Drosophila ananassae	FBpp0349530	Glycosidase NET37	Dana\GF21469-PB	DaGlyMyo1B	No
Drosophila ananassae	FBpp0350541	Glycosidase NET37	Dana\GF21469-PC	DaGlyMyo1C	No
Drosophila ananassae	FBpp0121278	Glycosidase NET37	Dana\GF18086-PA	DaGlyMyo2	Yes
Drosophila ananassae	FBpp0345570	Neutral alpha-glucosidase	Dana\GF21346-PB	DaNAglu1	Yes

S3 Table. List and sequences of the RT-qPCR primers used for PCR-amplification of cDNA sequences codifying for proteins belonging to GH13 and GH31 family of *L. longipalpis*

Gene	Sense primer (5' to 3')	Antisense primer (5' to 3')
LLOTMP004838_1 (LlAamy1)	GGGCTTTGGTGTTGTGACTGAA	GAGCACGATTGGATGGAAGGAAG
LLOTMP004841 (LlAamy4)	ATGTTGCTCTGGTATTTGTTG	TGTTGCCATCTTGTATTGTCT
LLOTMP004880_2 (LlAamy6)	ATGTCCTCTTTCGCCTTCAC	TTGACTACCTGCCTCCATCT
LLOTMP004881_1 (LlAamy8)	ATAACTACAATGCTCCTGCCTTTCC	TATCTCTGGTATGTGCTCCGCT
LLOTMP005909_2 (LlAamy14)	CCAACGGATTTACCAGACGG	TTCCTCTTGCTCACAATAACTACC
LLOTMP006803 (LlCD98hc)	TCCTCTTCATTTGCTTCTGGG	TGGTGTGGGTGTATCAATCTTC
LLOTMP008629 (LlNBAThc)	TCCACTCCTCCCAACAACACC	CATCATCACCTTGAAACTCACCACC
LLOTMP005533 (LlAGB1)	AGATAGATGGGAATGAGGATGGG	GGTGAGGTAGAGTTGTTGCG
LLOTMP000566/LLOTMP008156 (LlMal1/LlMal2)	GCAACCCGTAGTCACCTCAA	CAGGATGTCCGTTCCAACAA
LLOTMP002257 (LlMal3)	GCTTCTCCACAAATCCCTCC	TACTCGTTTCAGCCACCTTCT
LLOTMP001847_1/LLOTMP001847_2 (LlGlyMyo1/LlGlyMyo2)	GCTTGCTCGTTCCTACTTTG	GTATTATTCGTGAGGCAGTGGT
LLOTMP001881 (LlGlyMyo3)	GATTGGAGGCAACGGCTA	TATGAGAACTGAAGGCTGGG
LLOTMP0000840 (LlGlyMyo4)	TGATGAAGATGGTTTGGCTGTTT	ATGCTGGTAGTATTTATGCGGAC
LLOTMP006451 (LlLysAglu1)	CTGAGAGCATCAACGGAGAG	ATAGAATAAAGGTGTGAGAAGAGGT
LLOTMP003489 (LlNAglu1)	TGCTGAAGAGGGTGGAGAAGG	ATTTGGTTTACTGTCGTGATGCGA
LLOJ001891 (GAPDH)	TTCGCAGAAGACAGTGATGG	CCCTTCATCGGTCTGGACTA

Lutzomyia longipalpis genome, Jacobina strain, LlonJ1.4 geneset, June 2017, was used to design primers.

S1 File. Protein sequences of *L. longipalpis* and *P. papatasi* from GH13 and GH31 families after sequence curation. The Xs represent missing amino acid in incomplete sequences.

Alpha-Amylase

> LLOJ004838_1/ LlAamyA1

MKFCVFLLLTFVALARGQFDPHFVSGRSGIVHLFEWKWEDIAAECERFLGPNGFAGVQVSPPNENVIVTNRPWWERYQPISYRIVT
RSGNEQQFANMVYRCNNVGVRIYPDIVINHMAASHPVMMGTGGSTANVGNRDYPAVPYSIHDFNPSCPITNYNDRYQVRNCELV
GLPDLNQGVAWVRDRIVEYLDNLVALGVAGFRVDAAKHMWPGDLEIIYGRVRELNVVHGFPVHARPFFFQEVIDLGGEAISKNEY
TGFGVVTEFKFSAEIGRVFRGNDLLHHLSNWGEAWGFLPSNRALVFVENHDNERGHGAGGDQILTYKNAKQYKMAVAFTLAHPF
GIPRIMSSFAFTNTEIGPPMDGNQNIISPSINADGTCGNGWVCQHRWRQIFNMIRFRNEAGTAGLTNWWSNGSQQIAFARAGRAFVA
FNNQGSNMNVNLQTSLPGGTYCDVISGNVSGGSCSGKTVTVNGDGTANIVIGAAEEDGVLAIHVGARL#

> LLOJ004838_2/ LlAamyA2

MKILWIFFLAIVATVRGQFDPHYVAGRSVMVFLMDWKYEDVALECERFLGPMGFAGVQVSAASENDIVTSPLRPWWERYQVVSY
KIGSRSGDADSFADMCTRCNNVGVRVYVDAIFNNMGGMGIMKGTAGSLGDSITRTFPAVPYTILNFHNPPCSITDYNNTQQVRNCD
LVGAPDLDQSQPYVRDRIVGYLNHLTSLGCAGFRIDSAKHMWPTDLQAIYSSVIDLNTAQGFPAGSRSFIFQEVIDSGHEAIKKYT L
KKIVYIXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXX

>LLOTMP004839/ LlAamyA3

MFWNLFLTTLAIARGQFNPNYVPGRTVMVFLMDWKYQDVALECERFLGPKGFGGVQVSSPAENDIVTDPLRPWWERYQIISYIIGS
RSGNEDDFADMCYRCNKVGVRIYVDAIFNHMGSLPNMVGTAGSTADGPTKSYPAVPYDASNFHPSCPLTTFQVTFIVRNCELFGAP
DLDQAQPYVREKIKGYLDHLITLGIAGFRVDSAKHMWPKDLKAIYSSLRDLNTDFGFPPKSRPFVFQEVIDAGGGSEPINKYEYSPIG
VVTEFLYSYILTDFFRGNRNLSDLATWGQSEWKFLPSKAALVFVENHDNERATPPVLNYKDGKPYWMAVAFELAHPYGYPYLMS
DFNWTTYDQGPPMDEDQNIISPTINKDGSCGSGWVCQHRWRQIYSMVGFRNQVGYARFTNFWSNGKNQIAFARQGCGFVAFNND
DVDLNEVLQTTLPAGTYCDVSTGYAKAGECTGRTVVVGDDGTATIEISSSDDIGFIAIHIGQML#

>LLOTMP004841/ LlAamyA4

MRIVVLILCLVALGSGQHWPNFWDGRTGIVHLFEWKFRDVADECERFLAPKGYAGVQLSPVNENVIIPNRPWWERYQPISYVLTTR
SGNEQDFLDMSRRCNAVGIRIYVDVLLNHMSADNYPSDGTAGNWADTSQRWWPAVPYGPNDFNPRCEIHNWDDPWEIRQCELV
GLHDLNQGTEHVRSMLVNFLNHLVELGVAGFRVDAAKHVWPHDLEVIYNRVNNLNTNFGFAPGSRPYIYQEVIDLGGGGIRYEYT
PLGAITEFKFSAEIGGAFRRYNDLRWLRNFGPEWGFVPSDVALVFVDNHDNQRGHGAGGEAVLTYKQPRQYKMATAFTMAWNY
GTIRLMSSFAFDHGDQGPPQDANENIISPSINPDNTCGNGWVCEHRWRQIYNMIGFANQVKGTSVNDWWDNGHNQIAFCRGDQGF
IAFNNEDHNFSHTLQTCLPAGTYCDVISGENTGNGCSGVTITVNWDQTAHIVIPVDHYDGVIAIHRGPTSRIG#

>LLOTMP004880_1/LlAamyB1

MKVLVIFLALVALAKGQFNPYYEDGRTVMAFLMDWKYSDIALECERFLGPKGYAGVQVSSPSENAIVTDPLRPWWERYHVVSYII
GSRSGTQDDFADMSARCNKCGVRVYVFTIFNNMGAFPDMVGTAGSTADGPSKSYPAVPYDSTNFHPSCPIQSYQHAFIVRNCENE
GTPDLDQSQDYVREQIVAYLNNLISLGAAGFLIDSAKNMWPTDLKAIYSSLSNLNEDYGFSPNCRPFIYQDVIELGKEAVNKYEYLP
LGDVTDYKVSYSLNSIFNGNKSLSVLATWGESLWNFIPSKSAVVFVENEQNEREGFYFGQVLNYKDGKPYIMAIAFLLAHPYGNPR
IMSDFDWSDSDQGPPMDDNQNIISPTINKDGSCGDGWVCQHRWPQVCSMVTFRNQVGNSNFCNFWSNGNNQIAFAREGHGFIVFN
NNDGKLWQTLPTTLPAGTYCDVITGYVKNNECTGNSIEVADDGTAEFIIYSDDPNGVIAIHIGQMLE

>LLOTMP004880_2/LlAamyB2

MRSALIFLAFVVLARGQHNPNFVPGHTGIVHLFEWKWVDVAAECERFLGPMGFGGVQVSSPAENAIVRDPLRPWWERYHIISYKL
VTRSGNEEEFADMCTRCNKVGVRIYVDTVINHMGSLPNMIGTGGSTADGPSKYYPAVPYNVDHFHTSCPIENWQDEYEIRYCELFG
APDLNQGMGYVRTTIAEYMNHLIRLGVAGFRIDAAKHISSSHLYFTTIDLTYLNDLNPDHGFPARTKPFIYQEVIDFGWDAVSKYDY
TYIGTVTEFMYSYYISRAFRGNDLLRNLATWGPTWGFLNSSYALVFVENHDNERGDGAGGQEILSYKDGKQYRMAVAFALAHTY
GIPRIMSSFAFTERDIGPPMDSNENIISPSIYPDGSCGNGWVCQHRWRQVVNMIAFRNISGNENITNWWTNDKNQIAFARHQRGFIAF
NNDAMNMNVSLQTSLPAGIYCDVISGNLVDKLCTGKNITVRGDGVARILIMKDEEDQVVAFHVVK
QG#

>LLOTMP004880_3/ LlAamyB3

MHFISLGGGILLVLVLHVADGQFDPHFLPGRSVIVHLFEWKFSDIAAECENYLGPNGFGGVQVSPINECLVSPERAWWERYQPVSY
AIVSRSGDEKEFAEMVKRCYEAGVRVYVDVIFNHMASGEGEVVGTGGSLVYPEERLYPHVPYGPEDFNPHCVIEDYQDVDQVRNC
ALVSLPDLNQKSDNVKRSVIEFLDRLIDHGVAGFRADACKHMWPEDIKFLFGSTKNLSPEFGFPDKARPFLYQEVIDLGQEPISKNE
YTSIGVVTEFLFSAEIGNIFRYKKLKEFMKWGTDERFLRSDRALVFVENHDNERGHGAGGKDILTYKDGKRYLLAVLFTIAHPYGIP
RIMSSYDFSNSDEGPPANAKGEIISPVFNDRGLCTNGWICQHRWLGVASMVQFRNAVAGSGIVNWMDNGEQQFAFCRGDLGFVAF
NGYTMSHLNQQLRLTYHLASTDDVISGEVTLDGCTGLEVVVVKNDGYANIFIPGDSPTG VLAVHLGSAYIID#

>LLOTMP004881_1/ LlAamyB4

MKATIFALLSLLAQTKGYFNPNFVPGRSGIVWLFEWKFADIALECERFLGPNGFAGVQISSPAENAIVTNPLRPWWERYNPVSYIIAS
RSGNEQDFSDMVTRCNNVGVRTYVEIIMNNMGGYPNMVGTGGSTADGPSRSFPAVPYTSADFHATCSITTMLLPFLIRNCAVEGLP
DLNQSGAHTRDMIVGYLNHLISLGVAGFRINSAKFVWPSDLKVIYDRTNNLNTNYGFAPNSRPFIFQDVADYGLDAISKYEYSSMA
VVSEYKFLLNLGRIFRHLDVLSVLTNWVQTFGLLPSDAALVFVDSPDSERGSQGSEILTYKDGKAYILALIYMLAVPYGHPIIMSSYE
FKDVNSGPPMDA
NQNIVSPTINPDGSCGSGYVCQHRWTPVSSMVAFSNNVGDAPVTNVWTNGQNQLAFARDGVGFVAINSNAETLDVSVQTGLVGG
VYCNVLGGPIVDGTCAGQTVVVGDDGIAAITITTQDNLGAIVLDVNQQLLNGEAKKFKQEHPSLY#

>LLOTMP004881_2/ LlAamyB5

MKFFMILLALVASTRGQFNPYFWPNHNTIVYLLDWKYSDVALECENFLGPKGFGGVQVSSPSENAIITDPYNRPWWERFHVLSYIIG
SRSGNQDDFADMTARCNKCGVRVFVDAIFNEMGGLPDMVGTAGSTGDGPSRTFPAVPYDSSNFHPPCPIQSWLHAFIVRNXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXKKGTVCDSPYVCQHRWPQVCSMVTFRNQVGDA

NFCNFWSNGNNQIAFAREGQGFIVFNNNDKKLWQTLQTTLPEGTYCDVITGYVKNDECTGRSIEVADDGTAEFIIYTDDPNGVIAIH
TGQML#

>LLOTMP004881_3/ LlAamyB6
MLKFVVFLALVAFTRGQFNPYFWPGHNTIVYLLDWKYSDVALECENFLGPKGFGGVQVSSPSENAIITDPYNRPWWERFHVLSYII
GSRSGNQDDFADMTARCNKCGVRVFVDAVFNNMGGLPDMVGTAGSTGDGPSMTFPAVPYDSSNFHPPCPIQSWLHGFIIRNCEDP
DFLPVPDLDQSQDYVREQIVTYLNNLISLGAAGFRIFAAKHMWPTDLEAIYSSLSNLNEDYGFSPNCRPYIYQDVIDLGGEAVNKYE
YIPLGLVSEYNASQSLADIFRGNRPLS
VLATWGEAEWNFVPSANAIVFVENEQNELDTWYFGRVLNYKDGKPYIMAIAFLLAHPFGNPLLNSDYEWTDPNEGPPMDDDQNIL
SPVINKKGTCDSPYVCQHRWPQVCSMVTFRNQVGDANFCNFWSNGNNQIAFAREGQGFIVFNNNDKKLWQTLQTTLPEGTYCDVI
TGYVKNDECTGRSIEVADDGTAEFIIYTDDPNGVIAIHTGQML#

>LLOTMP004882/ LlAamyB7
MKIFWIFFLATVATVRGQFDPHYVSGRSVMVFLMDWKYEDVALECERFLGPMGFAGVQVSAASENDIVTSPLRPWWERYQVVSY
KIGSRSGDADSFADMCTRCNNVGVRVYVDAIFNNMGGMGIMHGTAGSLGDSITRTFPAVPYTILNFHNPPCSITDYNNTQQVRNCD
LVGAPDLDQSQPYVRDKIVGYLNHLTSLGCAGFRIDSAKHMWPTDLQAIYSSVIDLNTAQGFPAGSRSFIYQEVIDSGHEAIKKYTX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXCNEPLSTLSTWGQSNWGFLPSDRALVFVENHD
NERGQYESAALNYKDGKPYKMAVAFALAHPFGNLRIMSDFDWTTYDQGPPMDANQDIISPTINADGSCGSGWVCQHRWRQIYGM
VGFRNTAGTTEVTNFWSNGNNQIGFARQGRAFVAFNNNDADMNQTLQTTLPAGTYCDVITGVAENGACTGRTIVVANDGTANFV
ILTSDADGVIAIHVDQKL#

>LLOJ004885/ LlAamyB8
MKLFVVFLALVAFTRGQFNPYFWPNHNTIVYLLDWKYSDVALECENFLGPKGFGGVQVSSPSENAIITDPYNRPWWERFHVLSYII
GSRSGNQDDFADMTARCNKCGVRVFVDAIFNEMGGLPDMVGTAGSTGDGPSRTFPAVPYDSSNFHPPCPIQSWLHAFIVRNCENA
DYLPVPDLDQSQDYVREQIVTYLNNLISLGAAGFRILHXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXX

>LLOTMP005909_1/ LlAamyC1
MHFISFGGGILLVLVLHVADGQFDPHFLPGRSVIVHLFEWKFSDIAAECENYLGPNGFGGVQVSPINECLVSPERAWWERYQPVSY
AIVSRSGDEKEFAEMVKRCYEAGVRVYVDVIFNHMASGEGEVVGTGGSLVYPEERLYPHVPYGPEDFNPHCVIEDYQDVDQVRNC
ALVSLPDLNQKSDNVKRSVIEFLDRLIDHGVAGFRADACKHMWPEDINEYTSIGVVTEFLFSAEIGNIFRYKKLKEFMKWGTDERFL
RSDRALVFVENHDNERGHGAGGKDILTYKDGKRYLLAVLFTIAHPYGIPRIMSSYDFSNSDEGPPANAKGEIISPVFNDRGLCTNGW
lCQHRWLGVASMVQFRNAVAGSGIVNWMDNGEQQFAFCRGDLGFVAFNGYTMSHLNSTVKVCLPPGIYCDVISGEVTLDGCTGL
EVVVKDDGYANIFIPGDSPTGVLAVHLGSAYIID3

>LLOTMP005909_2/ LlAamyC2
MKFLIGFGVVLILVAVAAQYDPHYLDKRTVMVHLFEWKWNDIAKECQDYLGPNGFGGVQVSPVNENWVSGERAWFERYQPISY
KLTTRSGTEDEFAAMVKTCRQNGVRIFVDIVVNHMASGALEDTIYGTGGSEAHPGPFDYPAVPYEKNDFHPDCSISDYQNVYQVR
NCQLSSLRDLNQTIPYVREKILDFLNHLVDLGVAGFRIDAAKHMDPKDLRYIYNHIKKLNRTLSFRAGDKAFIFQEVIDLGGEAVSS
REYISLGVVTEFKASDDLGKLFRGQVALSTLERWGPQYGLLPSNRALAFVENHDNERGHGAGGTNILTYKDGKIYTMAVVFNLAH
SYGVPRMMSSYEFNDPSQGPPHDDNNNILTPEFSADGNSCTNGWVCQHRWRPMRNMVKFRNIVGRKPVRKWYDNGSNQIAFSRG
NQGFVAFNMDIVDFNQQVPTDLPDGVYCDVITGEKNGNECTGKVVIVSKRKAAVILRADDDYGVLAIHSESKL#

Maltase
>LLOTMP000566/ LlMal1
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXNIGITQRLQYVKDLGMSGTWLSPIFKSPMADFGYDTAD
YTAIQPEYGTMEDFEALIKKANEIGIKIILDFVPNHSSDQHEWFKKSVDRVPGYEDFYVWHPGKIVNGQRQPPNNWVSVFRGSAWT
WNEKRQEYYFHAFLKEQPDLNYRNQATVDAMKETLRFWMQKGVFGFRIDAVPYLFEVPADSNGNIPDEPLTGPSCPDPTHDCYTQ
HIYTQNLDETFDMVYQWREVVDEFAKQTDNVPRILMIEAYTPLENMKRLYEDGHGREGAQLPFNFELISKLNGKSTAKDFKDVID
GWISRLPQGKENNWVLGNHDNKRIASRFGIDRADLINILLQTLPGMAVTYNGEELALKDVYISWEDTIDPAGCNTSPDVYEQYSRD
PVRTPFPWNANKNAGFSNSDKTWLPVSSDYKEVNVEAQEKATRSHLKSFRKLTKMHQHEKAFTQGQLLMKLVGTDILAYERRVP
GTKPEENFVMILNFSNNTHKVNIPQLFPQMSQMYRIEVVSMHSLSHKEGDTIDGKNFEVKPNEAFVLRGGAPRLYSYLILCFTMFSA
FFNINLTRA#

>LLOTMP008156/LlMal2
MRVIIFVFLVISATTEGVWYNNGNFYQIYPRSFMDSNNDGVGDLKGITQRLQYVKDLGMSGTWLSPIFKSPMADFGYDTADYTAIQ
PEYGTMEDFEALIKKANEIGIKIILDFVPNHSSDQHEWFKKSVDRVPGYEDFYVWHPGKIVNGQRQPPNNWVSVFRGSAWTWNEK
RQEYYFHAFLKEQPDLNYRNQATVDAMKETLRFWMQKGVFGFRIDAVPYLFEVPADSNGNIPDEPLTGPSCPDPTHDCYTQHIYT
QNLDETFDMVYQWREVVDEFAKQTDNVPRILMIEAYTPLENMKRLYEDGHGREGAQLPFNFELISKLNGKSTAKDFKDVIDGWIS
RLPQGKENNWVLGNHDNKRIASRFGIDRADLINILLQTLPGMAVTYNGEELALKDVYISWEDTIDPAGCNTSPDVYEQYSRDPVRT
PFPWNANKNAGFSNSDKTWLPVSSDYKEVNVEAQEKATRSHLKSFRKLTKMHQHEKAFTQGQLLMKLVGTDILAYERRVPGTKP
EENFVMILNFSNNTHKVNIPQLFPQMSQMYRIEVVSMHSLSHKEGDTIDGKNFEVKPNEAFVLRGGAPRLYSYLILCFTMFSAFFNI
NLTRA#

>LLOTMP002257/ LlMal3
MKLLGLLLVTVLTVEGHLLGERQSSEKADLRHDHPELDWYERANLYQIYPRSFMDSDGDGIGDLNGITAKLEHLQESGIFATWLSP
IFSSPMRDFGYDISNFTEINSDYGTMEDFLQLLHEARRLNIRLILDFVPNHTSDLHEWLVKSEDNDPEYRDYYVWRDARYVGEERHP
PNNWLSVFHGPAWTWSEKRQQYYLHQFAKEQPDLNFENEKVVEEMTNVLKFWLDIGVDGFRVDAINHLFEDPSFADNPPNREIYD
PMDYSHFDTIYSKDLPKSYQQVYDWRKYLDSYAEEHGTDTRVLLTEAYTNITNTMKWQMSEDGTQKGAHFSFNFNLIMELKTLP
DELNAADVKATIDEWLEHLPDGVPPNWVLGNHDRPRVASRFGVDLVDNMNILVQTLPGVAVTYYGEEIGMEDFRGISWEDTQDP
QACGSNETVYQLYTRDPVRTPMQWDDTQNAGFSTNPSPWLPVHPDYVQNNLAQQKVAETSTYKLFLDLVKLRIDHALEYGDFKS
EALSNGVYAYSRNMEGHESLVVALNFASTEATVDLTPVLLENSHSEGLVELATTKSIHKRDDVVNLRTLTLQQYDGVILRVNSSAA
TIFISSFALLLVVLKTLLY#

Amino acid transport protein

>LLOTMP006803/ LlCD98hc

MTDKTERNSTKLNLEAQRLSASDSPFPKETYKTIPDLDHHQKAEKKLEDGLCDLLSHCVFCMVISDAVKNSNMVREKVDMSEEGA
DEKMLGAGEEEQQKLAKKEEVKFISGDQQNGDAKIDIGAIDKTFTGLTKEELMKYANDPFWIRLRWFLFICFWGLWAGMLVGAV
LIIIYAPKCAAPEPLSWWKQGPLVKIDTPTPDNIQIEGIKGLGVKGVIYEVPADETYNIGSDPAISDALKKLAASFAAYGINLIVDLTPN
FVSRGDPLFQKALNNPDVRSAFVTREGHQVPNNWLSLVNGSAWTNEGNLIFLQQFGHDRFDLQLNDPVALTKLKGVLKEVVSLG
AKGIRLANTKHFIVDRELKDNVPSNNPNYDHTQYGFWTHAHTTYQDGLGDLLYDLKAYVHNITNSEGFLAASDDIDRPEVYMSSR
GVQSLDMIQLGRVTNLLANSTSADHIYRELQNVWKSTVETGKHFWLQWNYVKQELLQQITISEYNIFLQLIPGVPVFNVDASKAKA
EDYSTEAFHKIRTSPSYMHGSFDLYRDVNSTVIGYSRTKSGNPGFFVVFNPTEQYVNANFSNVVGIAEELTVHTLSTNYNATDVAV
KGKVSSSSIPVSRYSAIILTYVPKA#

>LLOTMP008629/ LlNBAT1

MNNLGISTDPALLGVDLPSSLTTSPSVSTFMPEEDASICPLLPTTPSPPPMDFIHPLTPSTGIDEGNNEDLGGEFQGDDAQADSSSSGSS
SGIGIQVTANGPGSLFNKHAYQHLGSKNGDITQDNGVTQTGVFSIPMTKDTPSFVNWNWPLIRKCTFFLFMSGLFAMCAIVVAMMF
KLPKSCNPEVPWYKGAVFYEIFPASFQDSNGDGLGDLKGLASRIDYLESLGVGAVRLNSIFPAKHYPDHFQNVTTLLDIDEILGNPR
DLTYLARILQRRNISLILDLPIYPLVKHLAVIRENFTEATNSTDTEDEEVFQDSDTSEDPVLTAMRFWLSMGVDGFYVKGLENYAGD
PYLIENLQEWKFALGSDRILMVSKLVFDSVDDDTAEKVRHCVDLVDVFVDVSNGTKVIAEKVQSILNSRFLAPGFGPWIHWSLGGV
TERRLAQGTSSNISLAATLMQLMLPGTPSIFYGDEVALQEVYDPLGEHEESKHLHHLSTMVWDSEIQFTVKEHLPWLPRGAFVAFH
HFEHVAGMIRLRGISPSIYQNAVIKEHQSVLNTAVRYSKNDILILERWYPRRNTFTSISNFGSKSLVLDLSGMFYAGEIMLGSLKGEK
VFFSSIQVKPMETIIVKLDK#

1,4-alpha-glucan-branching enzyme

>LLOTMP005533/ LlAGB1

MDPMKVEVPEIDKLFAIDGYLRPQERELRRRHGVIKDWLKKIDGNEDGGVDGFSQAYKYYGIHIQPDNAVIAREWAPGAQQLYLT
GDFNDWQWEATPYTKLEFGKWELKIPPREDGSCAIRHLSEIKVIVRTQEGTLVDRLSPWAKYVLQPPKEANQGTNFKQYLWHPPP
HEKYMFRYGRPQKPQSLRIYECHVGIATEELGVGKYRDFGDKIIPRIVKQGYNAIQVMAIMEHAYYASFGYQVTSFYAASSRYGTP
DELKYMIDMAHKAGLYVLLDVVHSHASKNVADGLNRFDGTNTGFFHDGARGEHPLWDSRLFNYTEYEVLRFLVSNLRWWHDEY
NFDGYRFDGVTSMLYHSRGVGEGFSGDYNEYFGLNVDTDALVYLAIANEILHKLDKDIITIAEDVSGMPTMCRPVSEGGIGFDYRL
GMAIPDKWIEYLKEKRDDDWNMGNIVHTLTNRRWMEKTVAYAESHDQALVGDKTLAFWLMDKEMYTHMSTLSEPPLIIDRGIAL
HKMIRLITHSLGGEAYLNFMGNEFGHPEWLDFPRVGNNDSYHYARRQWHLVDDEMLKYKYLNEFDRAMNGLEMKYGWLACDP
AYVSCKHEDDKVIVFERAGLLFAFNFHPSKSFTDYRLGVEVGGLYRIVLSTDDPLFGGMNRVDVKCDHLSTPEGYAGRRNFIQAYL
PARTAFVFAKFD#

Glycogen debranching enzyme

>LLOJ008312/ LlGlyD1

MGHQDVLSLAISHGEDKEGTLFRLKRGTILHIVPGPNLLGRHIALYCNYPQNGEKGFERTKYSTLIWHTSSGQKLTNAAQPFVEVND
LDIYCEIVANKSGTFHFYFSFQESPKEREGSLYVQVEPKIKVGPKNAEREIPLDSIRCQTVMAKCLGHLETWEAKLRVTHESGYNLL
HFTPIQKLGNSRSGYSLCDQLSVNPDFGPSANFDKVAKIVKKCREEWGIASICDIVLNHTANESPWLKEHPETTYSCSTCPHLRPAFL
LDSVLAMASSDTGNGLLETFGVPVEIDREDHINALRHQLHSNYLPKARIHELYQCNVEEYVKKFSEEMRKRPPPKAPKESQSKETIV
LKQDPAYRRLACTIDFESAFEIFNVFRNDAFDEDTRLRKCAEAFRQHLEELNEVVRCEIAEHLRSAVENCLSGLRYERIQDDGPRLR
EISIKHPLFCRYFTDGDVTCKTIEDIEALMYGDSAKYLMAHNGWVMNGDPLKDFARPQPTTANVYIRRELIAWGDSVKLKYGDKP
EDCPYLWDHMKKYVEITAKIFDGVRLDNCHSTPLHVAEYLLDAARKINPELYVVAELFTNSDHTDNIFVNRLGITSLIREALSAWDS
HEEGRLVYRYGGAPVGAFFSSPKRLLAPCIAHALFLDLTHDNPSPIQKRSVFDLLPSAGLVSMACCATGSTRGYDELVPHHIHVVDE
EREYQEWGKAVDAETGIISGKKAINLLHGELAEKGFSEVFVDQMDYNVVAVTRSCPTTRESVILVAHTCFSYPDPYSGPTNVRSLRF
EGHLEEIILEAGVTHKSSKPYDCPYRYEKDEKYINGLTEYQLKIVEHIPLEKSTIFSTEMVKDGNITQLNFKNLTPGSIVAVRVSLHER
TRPHFANVQKLVDAFHFQKGQIFSDLQKIVSKLNLIDLNRALYRCDEEEKDMGMGSGTYDIPGFGRLVYSGTQGFASALSVIGPNN
DLGHPFCDNLRRGDWMIDYIRDRLARWSGTKALSEWWATNTASLKEMPRYLIPSYFDVILTGVNALLLEQSVTLMSDFVRKGSSF
VQSLAMGSVQCVAECPSANLPALSPQTKAPKPPNQCATMSAGLPHFSTGYMRCWGRDTFISLRGLLLLTGRFDEARYMILGFGSCL
RHGLIPNLLDGGYKARFNCRDAIWWWLYSIKYFVEEAPKGVEILTEKVSRIFPMDDSEARKAGECDQALCDVIQEAITVHFQGLVY
RERNAGPAIDAHMTEKGFNNQIGVHPDTGFVFGGNDANCGTWMDKMGSSEKARNRGVPSTPRDGSAVELVGLQMATLRFLQKM
STEGKFPHKSVERTSKNGTKTVWTYKEWADRIKANFEKNFFVGQGEIPLANKRGIYKDSCGATQTWQDFQLRPNFPIAMVAAPELF
DPKHAWEALEQARKYLLGPLGMKTLDPQDWGYHGDYDNSNDSEDPKVAHGANYHQGPEWVWPIGFYLRARLIFAAQNNRLKD
TVAETWAILTAHLKEVKSSWWRGLPELTNSNGSFCSGSCTTQAWSMATVLEVLYDLERLKCKS#

***L. longipalpis* Glycoside Hydrolases from family 31**

Glycosidase NET37

>LLOJ001847_1/ LlGlyMyo1

MFPSMIKKKKKSKDFKIKAFVACLFLIIVFLVGYAYIMYNQQVLARSYFDRVKLNKGQRVIGIYNEKGVLMVTGVWVRLYIQRKPT
TASRIIRWKMEVCVWNGMARRGCISIFQELSPSVPCYTIRWQSLSPDVLPTDCYESFMGQGHWYGGGLTRGGNWPLETESFPFAPFI
TGDAKRQQWGNVLKRYFISSRGVAIQVDEKSPLYVSMNDNKSGEMCFRAKHDHFAFVNRLTPLPELKYRICTSDNMRQLHQQMT
QQSLWDGLKEHDINVVHSLLEEPVWQIPSSGGKGITGDSIHNYTERVIALGFLRLGHVLVNEFWQRHIGDFTLETERFPNLEETVTIL
HRRGFRIVFSVQPYISTDSDNFAQSVAQKLLVYERQSERSIPALTRYKSVASAGVLDITNNASIPWLIEKLEKIQKTYKIDGFYLDFG
TSQNMPHYYQCNKTLYNPDQYKTIFTTALEGVISVIGVSSAVTVPRPPAFVSLPPVNSSWEGLRTVVTSALTYGIIGYPFIMPGPIGGD
YLLPPSASNETVSFYFMEEPPLPDQELYLRWMQLATFLPVIRFTHLPSEYKSELIMEAVKELTLIRQKAVIPLLKKYLSDAMNEGLPLI
RPLWMLDAQDTACLYVNDEFSIGEDLIVAPILEKGQLQREVYLPQGVWKDGIDGSLRKGSRWIHNYRVPEDKVAYFMKMPDNTR
F#

>LLOJ001847_2/ LlGlyMyo2

MIKKKKKSKDFKIKAFVACLFLIIVFLVGYAYIMYNQQVLARSYFDRVKLNKGQRVIGIYNEKGVLMVTGRLGTTIHSEKAYHCLT
NNTLEDGSVCLEWDGKARMYLNFQELSPSVPCYTIRWQSLSPDVLPTDCYESFMGQGHWYGGGLTRGGNWPLETESFPFAPFITG
DAKRQQWGNVLKRYFISSRGVAIQVDEKSPLYVSMNDNKSGEMCFRAKHDHFAFVNRLTPLPELKYRICTSDNMRQLHQQMTQQ
SLWDGLKEHDINVVHSLLEEPVWQIPSSGGKGITGDSIHNYTERVIALGFLRLGHVLVNEFWQRHIGDFTLETERFPNLEETVTILHR
RGFRIVFSVQPYIS
TDSDNFAQSVAQKLLVYERQSERSIPALTRYKSVASAGVLDITNNASIPWLIEKLEKIQKTYKIDGFYLDFGTSQNMPHYYQCNKTL
YNPDQYKTIFTTALEGVISVIGVSSAVTVPRPPAFVSLPPVNSSWEGLRTVVTSALTYGIIGYPFIMPGPIGGDYLLPPSASNETVSFYF
MEEPPLPDQELYLRWMQLATFLPVIRFTHLPSEYKSELIMEAVKELTLIRQKAVIPLLKKYLSDAMNEGLPLIRPLWMLDAQDTGLL
YVNDVIFNLVEDLIVAPILEKGQLQREVYLPQGVWKDGIDGSLRKGSRWIHNYRVPEDKVAYFMKMPDNTRF#

>LLOTMP001881/ LlGlyMyo3
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXGPQQKYQYWPVEKLVLSNYSYVTKE
ADSAAIAERYWLNSEGAFIYLSDRAPLFIDQNSMMDNHLCFVASLALPYNPRAQTYDFEYSIGVASDPRKAHMEAVHRFLGKPSGV
PDRLMAARPIWSTWARYKRDVDQYVVMEFAKEIIDNRFTNSQFELDDDWEICYGALTFSTTKFPNIRKTVQDLKSMGFRVTLWIHP
FINKGCNPWYNEALTNRYFVLDHNGKEDTQWWNSEQGEAAYIDFTNPEAANWYTSRLHKLLEDTGIDSFKFDAGESSWQPDDPV
LPSTSLDQHPLKIQTDYIMHVAQFGPIVEVRSGFRNQFQQIYMRMIDKDSLWSWNNGLPTLITTLLQLNMNGYPLVLPDMIGGNGY
EGAPSREMFIRWLQANVFMPSLQFSYVPWDYSDGDGIKILDLCRHFVKLHEDYTDLIMERFQKAVDHGEPVNPPIWWVDPKNPVA
QSIYDXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

>LLOTMP000840/ LlGlyMyo4
MEKKRVKGNPKYKSAFKWGTAGIIATVIFVLVEVFIIHWVTPFDNFLRVHFPKAGVHVQFECEHIEDGVAFSIHKNRTLVQIVRMG
DNLGWDLRVEDHGGGKYVLRSREGSVSFNTAVDEDGLAVFHINQTIRTTDFVEHCFDLVNPDGNNWFAGPHKYYQHWPSQVLQF
TDDAYLPKESSHSSISERYWLTSKGMFIYFSDRTPLFLNRKPNHLCFAAKKQNPYYTYDTVFVFNYTIGIAADARKAHKAAVGRYL
RKPLHYPDRRVIQHPKWILRNHNMSQVQEYIDNLLYYNFSTSSIILDKSWETCTGALEFDPIKFPNVGSMIRYFRRHRTHLTLTVTPY
IHEDCNPYFQDALSRGFLVKTHQNTYLNFDRTATVDFTQVAARRWFLARLKDLQALGVDNFYFEGGEFDSAPNDPNFEGVSTSLH
PIQYTIDSLRALAEFDHNTIVRTGFCTQDLPLLLRLPDMDNRWDIQNGLDSLIPKILQTNLNGYYFLMAPIGGTTSEGLTKELYIRWM
QAVVFLPSMAFSTPPWEFDDETIELAQKFIRLHMRHVSLFVDLFKLAKSDGDPVNLPIWWLYPQNWRAQETHDQYLLGEDIIVAPV
LKPNSLERGIFLPYGIWRDGNDNSSLYQGPRWLPSYRAPLDVLPYFVRARNKN#

Lysosomal alpha-glucosidase
>LLOTMP006451/LlLysAglu1
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXNVFKFACFNFDFRISIILDATTEMFHGIYINAPYPVEVEITESYLVIVR
AIFDTSFHIDVLCGPTIGNLMDQLMEQTEQNLHPLPNWFFGFHICDTNITRNLTDSVIEVLELLSSDLPFDTHCIREQLLWLGDDIKGL
PSEVLPVDLKSAAYHMAKAKGLLMRHPTLEEPYIGRIGKTNVVYVDWVGSNQENLYEWSNQFMPNNLSADGYMVQANWMWDQ
SDGEKIPHSFPYISMDMINASKNIVPWFIRQSGTDGKASIFSHNFAASAQIDVIRKVLGNESFILSESSLFANVPISRRQVPSTWTELRN
LVRRTMGQSISGIYFSSANICGDGETWTEELCIRWYQFASLSPFYRVASAKAPTSFSKFAQSLLRASTERRYSLYLYIHTVVTERRPLL
TPLFYEYPDLIALMDNLTHQALVGPSLLFAPVLLPGVQHINIFFPEVYYEIGGGQQLIANTWAELPVVETDAPLFIRAGHIVPIQHTKN
VRSITVMQMKPIHLYVALGPLENGISTAIGKVQFDSEYKVTFNAAYTTSSEQTQLKLSSYSKLTCRTTTNLTIESRCRVEEGFLPAIFE
KFQPLSYXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Neutral alpha-glucosidase
>LLOTMP003489/LlNAglu1
MRFLLLGLLGLLITLTGSVDKNNFKTCDQSSFCRRCRKVDPGIPSPYEVQLGTLKTFPDHITVDVLNKNNEQVFNLKVVGLKGNKF
HVEVDEKSPLKPRYRVVDALIAPPVTENVSVTKEEEGSVVVKCGENRAVVVASPFRIDFYHGEVLVVSANAKGLMKFEHLRKKSV
PVPAEAAEEGGEGAAAAENEVENPQTNEDDDPGAWEENFKSHHDSKPNGPEAVALDFTFPQAEVLFGIPEHADSFALKPTLGSEPY
RLYNLDVFEYELDSPMALYGSVPVIYGHGEGNTAGVYWQNAAETWVDVYNSNAKKNVMSSIVNFVSRSRQADPPAAHFMSESGI
MDYYVLLGPTPMETFMQYADLTGPAPLPQMFALGYHQSRWNYNDEADVAGVHDKFDEHDIPMDTIWLDIEYTDAKKYFTWDGH
KFPHPLEMIRNLTERGRHLTIIIDPHIKRDGGYFFHNDCTDRGYYVKNKDGRDYEGWCWPGAASYADFFSPDVRQYYADQYQLDK
FDTTTADVMLWNDMNEPSVFNGPEVTMLKDNIHFGGWEHRDVHNLYGHMHVLGTFEGLVRRSGGVQRPFILTRGHFAGTQRYA
AIWTGDNTAEWGHLQASIKMCLSESVAGFSFCGADVGGFFGNPDAELFERWYQAAAFQPFFRAHSHIDTRRREPWLWPEATMLV
VRDALRRRYSYLPFWYTLFYEHERTGRPVMRPLLTHYPLDKETFTIDYEYLLGDILLVRPVLQQGVSKVDVYFPAVDGKKEGDIW
YDVDDHRKITAAGFESVPVDNYKIPVYQRGGTILTKKERIRRAATLMANDPVTFVVALDKEHSARGTLYVDDEKTFEYRKGKYIY
LNLEFKDNVISCRKNDEKANYDTKSWVERIVIVGLEHVPKSATLNVSGEPSVTLEVYKHGESVVVRKPKVVLSRKWSIQLNY#

***P. papatasi* Glycoside Hydrolases from family 13**
Alpha-Amylase
>PPATMP010150_1/ PpAamy2
MEFCIILFIALVPNVLGQFDPHYAAGHNVIVHLFDWKWVDIAAECERFLGPNGFGGVQVSSPAENAIVKEPNRPWWERYQIISYILQ
TRSGSQEEFADMVRRCNNARVRIYVDAIFNHMGGTGADMRGTGGSTANSVIRSYPAVPYDILDFHVPCPIFDYQNKDNVRNCDLV
GAPDLNQALEDVREKIVDYLSRLIALGVAGFRIDSAKHIWPNDLKVIFGRLPNLMTDHGFPPDAKPFIYQEVIDFGGEAISKYEYIEL
GAVTEFRYSFNISGVFRGLARLDTLQNWGEPWGFLPSDLALVFVENHDNERNHGAGGAQILTYKDGKRYIMAVTFGLAHPYGISRI
MSSFFFTDTEAGPPMDANEAIISPSINPDGSCGNGWVCQHRWRQIYNMVKFKNVAGTAPLLNWWSNGNNQIAFARSGKGFVAFNN
EGSDMDVILETSLPAGTYCDVISGSVSGTTCSGKTVIVQTDGTANIVIGNAEEDGVLAIHIEAKL#

>PPATMP010150_2/ PpAamy3
MNLSFFVFLALVAVAHGQFNPNFVSGRNGIVHLFEWKWEDIAAECERFLGPNGFAGVQVSPPNENVIVNNRPWWERYQPMSYRL
VTRSGNEQQFANMVYRCNNVGVRIYPDIVINHMAAGHPNMIGTGGSTANVGNRDFPGVPYSIHDFNPSCPINNYQDRYQVRNCEL
VGLPDLNQGVAWVRDRIVEYLDNLVNLGVAGFRIDAAKHMWPGDLEIIYGRVRNLNTVHGFPANARPFFAQEVIDLGGEAVSKNE
YTGFGVVTEFKFSAEIGRVFNGNDRLTHLSNWGEAWGFLPSNRALVFVENHDNERGHGAGGDQILTYKNAKPYKMAVAFTLAHP
FGIPRIMSSFAFTNTDIGPPMDGNQNIISPSINADGTCGNGWVCQHRWRQIFNMIRFRNEAGSAALTNWWSNGNNQIAFARSGRGFV
AFNNEGSNMNVNLQTSLPAGTYCDVISGNVSGGSCTGKTVVVQGNGNANIVIGAAEYDGVLAIHVGARL#

>PPATMP010150_3/ PpAamy4
MRIASAVLVSFFLALASGQHNPNFWGGRTGIVHLFEWKFSDVADECERFLAPRGYAGVQLSPVNENVIIPNRPWWERYQPISYVLT
TRSGNENDFLDMSRRCNAVGIRIYVDVILNHMSADNNPSDGTAGNYADTSARLWPAVPFGPNDFNPRCSIYNWDDPVEIRNCELV
GLHDLNQATEHVRSMLVNFLNHLVQLGAAGFRVDAAKHVWPHDLEEIYNRVNNLNTDFGFAPGSRPYIYQEVIELGGSPIRYEYTP
LGSITEFKFSAEIGRVFRGYDQLRWLRNWGPEWGFVPSNVALVFVDNHDNQRGHGAGGENILTYKQPRQYKMATAFTLAWDYGT
VRLMSSFAFNHGDQGPPQDANENIISPSINPDNTCGNGWVCEHRWRQIYNMIGFANQVQGTGVNDFWDDGGNQIAFCRGDQGFIA
FNNDQHNFSHTLQTCLPAGTYCDIISGENTGQGCSGVTITVNPDQTAHIVIPVDHYDGMIAIHRGAASRVN#

>PPATMP010150_4/ PpAamy5
MRIASAVLVLCCLALASGHHWPNFWGGRTGIVHLFEWKFKDVADECERFLAPKGYAGVQLSPVNENVI

IPNRPWWERYQPISYVLTTRSGNENDFLDMSRRCNAVGIRIYIDVLLNHMSADNHPSDGTAGNWADTSQRWWPAVPFGPNDFNPR
CSIYNWDDPVEIRNCELVGLHDLNQGSEHVRSMLVNFLNHLVQLGAAGFRVDAAKHVWPHDLEEIYNRVNNLNTDFGFAPGSRP
YIYQEVIDLGGRPFRHEYTPLGAITEFRFSAEIGRVFRGYDQLRWLRNWGPEWGFVDSRVALVFVDNHDNQRGHGAGGENVLTYK
QPRQYKMATAFTLAWDYGTVRLMSSFAFNHGDQGPPQDANENIISPSINPDNTCGNGWVCEHRWRQIYNMIGFANQVQGTGVND
FWDNGGNQIAFCRGDQGFIAFNNDQHNFSHTLQTCLPAGTYCDIISGENTGQGCSGLTITVNWDQTAHIVIPVDHYDGVIAIHRGPA
SRLN

>PPATMP010151_1/ PpAamy6
MLEAQFDPYFLPGRSVIVHLFEWKFADIAKECEEYLGPNGFGGVQVSPVNEVLVTPLGAWWERYQPVSYAIIGRSGDELDFMDMV
KRCYDAGVRIYVDVVFNHMAAGPGEIIGTGGSVVYPDERLYPHVPYGPEDFNDDCLIEDYQDTVQVRDCALVSLPDLNQKSDNVK
RNVIEFLDRLIDYGVAGFRADACKHMWPDDIKFLFGNTKNLSIEFGFPPNARPFLYQEVIDLGQEPISKNEYITNGAITEFLFSAEIGDI
FRNKKLKELMNWGTDERFLRSDRALVFVENHDNERGHGAGGENILTYKDGKRYMMAVLFAIAHPYGIVRIMSSYDFANSDQGPP
AKPDLEIISPVFNEKGLCTNGWICQHRWPEVAAMVKFRNIVGERGVMHWADNGEQQFAFCRGDLGFVAFNGYTRSNLESDVKVC
LPPGIYCDVISGSVGELGCTGLEVVVNDAGYAKVFIPADSPSGVLAIHLGSVFIPSK#

>PPATMP010151_2/ PpAamy7
MNILIALAVIVTSLSVRAQFNPHYIDNRNVIVHLFEWKWNDIASECKNFLGPNGFAGVQVSPVNENWVSPERAWYERYQPISYKLT
TRSGTQEKFASMVETCRQSGVRIYVDVVFNHMASGSLDTILVGTAGSEAYPQVFDYPAVPFTKADFHPDCTITDYQDVVQIRNCQL
SSLRDLNQTSPYVREKILGFLNQLVDLGVAGFRVDAAKHMDPQDLRYIYRTIKRLNKDSGFRGTDKAFIVQEVIDLGGEALSSREYT
SFGVVTEFKASDELGKAFRGKYPLADLRHWGPLYGLLPPNRALVFVENHDNERGHGAGGENILTYKNGKIYTMAVAFNLAHSYGI
PRIMSSYEFNDPSQGPPHDDSNNILSPEFSSNGESCTNGWVCQHRWLAIRNMVKFRNVVRKTSVTLWYDNGSNQIAFSRGTRGFIAF
NGDKVDFSAKVRTYLRSGTYCDVISGEKNGNSCTGKTVNVRRSEVAVIIEADDPYGVIAIHAESKL#

>PPATMP010151_3/ PpAamy8
MNLSAILIFTFLTLAHGQFDPNYAPGRNVIVHLFEWKWVDIAEECERFLGPNGFAGVQVSSPAENAIVTVLGFNRPWWERYHIMSYI
LETRSGNELEFADMVRRCNAVGVRIYPDAVFNHMGSLPNMTGTGGSTADGVIRTYPEIPYNATHFHIPCPIVDYQNSDNVRNCQLL
GAPDLNHTLEYVRNKIAGYLDHLITLGVAGFRVDSAKHIWPSDSQDIYGRVSNLNTDHGFADNSKPYFYHEVIDVGGEAVSKYDY
TGLGGVTEFLLPYYIGITFRGNRSLDTLQNWGEPWGFLPSDRGIAFVENHDNERGTSGRDILTYKDGKPYRMAVAFNLAHPYGITRI
MSSFYFSYDNIDQSPPMDDNQNIISPSINPDGSCGNGWVCQHRWHQIYNMVKFKNVAGTAPLSNWWSNGNNQIAFARSGKGFVAF
NNEGSDMDVILQTSLPAGTYCDVISGSVSGTTCSGKTVIVQVDGMANIVIGSAEEDGVLAIHVGAKL#
Maltase
>PPATMP010760_1/ PpMal6
MRIFSVTVLCLCYSEILGDLDWWKTANIYQIYPRSFKDSNGDGIGDLNGITERLPYLKDLGIHAVWLSPIFKSPMKDFGYDISDFYDI
QPEYGTIQDFEHLLKVAHDLNIRVLLDFVPNHSSDENEWFVKSANRDPDFEDFYVWHPGKEDPSDATKKLPPSNWLSVFRGSAWR
LHEGRGEYYLHQFVYQQPDLNYRNPKVVDQMKDVLRFWLSKGVDGFRIDAVPHLYEIAPENGQYPDEPKSGNSDDPDNSAYLDH
IYTQNQPETVQMVYQWREVLDEFTNRDQNTRLMLTEAYSDIDILMTYYGDGVKNGSHVPFNFFLITQLNNDSNAIDFHKTIRLWLD
KLPQDCIANWVIGNHDQRRVASRFGTYRIDVINMILNTLPGVSITYNGEEIGMTDVFLTWEETVDPAACNSDPQHYQEFSRDPERTP
FXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXX

> PPATMP010760_2/ PpMal7
MRIFSVTILCLCYSGILGDLDWWKTANVYQIYPRSFKDSNGDGIGDLNGITERLPYLKDLGIHAAWLSPIFKSPMVDFGYDISDFRDI
QPEYGTLQDFDRLLKIAHDLNIKILLDLVPNHSSDQHEWFVKSANRDPDFVDYYVWHPGKDDPNDPTKKLPPSNWLSDFRGSAWK
WHEERGEYYLHQFAYQQPDLNYRNAKVVDEMKDIIRFWLNRGVDGFRIDAVLCLFEAEPVDGVYPDEPKSGNTDDPDNPKYLTHI
YTQNQPETVEMVYQWREVFEEFTNADQNTRLMLTEAYSDIDILMTYYGDGVRKGSHVPFNFFLLTHIKNDSNAIDFRKAIQIWLDK
LPEDCVANWVLGNHDQRRVASRFGTHRIDIINMILNTLPGISITYNGEEIGMTDVFLTWEETVDPPACNSNSEIYDKFSRDPARTPFQ
WDDSTSAGFSTNTKTWLPVSPLYKEVNVKVERSSRRSHYKVYRKLLQLRRTPTLQKGNVETRTHGQNVLSITRSHPGEDTFITVVN
IGPDHEMIDLRELFSGRLTFHIVSVYSKRHEGDPIRGPILDLDPHEGVILKGFPXXXXXXXXXXXX

> PPATMP010760_3/ PpMal8
MRIFSVTVLCLCYSGILGDLDWWKTANIYQIYPRSFKDSNGDGIGDLNGITERLPYLKDLGIHAVWLSPIFKSPMADFGYDISDFYDI
QPEYGTIQDFEHLLKVAHDLNVRVLLDFVPNHSSDQNEWFIKSANRDPDFVDYYVWRDWKGDPSDPTKRLPPTNWISLFRGSAWK
WHEGREQFYLHQFAYQQPDLNYRNPKVVAQMKDVLRFWLNKGVDGFRIDAVWTVFEIAPDSDGNFPDEPLSGNSNDPDDHGYL
NHIYTQNQPETVDMVYQWRELLEEFTSKDQNTRLMLTEAYADVDIVMTYYGDGVRKGSHVPFNFNMITDLHNASNAVDFHRLIR
VWLDKLPKDCVANWVIGNHDQKRVGSRFGTHRIDVINMMLNTLPGISITYNGEEIGMTDVFMTWETVDPAACNSNPQIYEGLSRD
PARTPFQWDDSTSAGFSTNPKTWLPVSPLYKEVNVKVERSSPRSHYKIYRKLLQLRRTPTLQKGNVETRTHGENVLSITRSLPGEDT
YITVVNIGPDREVVDLRELFSGRLTFHIISLTSKRREGVAVRGPILDLDPHEGVILKGSAFHXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

> PPATMP010760_4/ PpMal9
MKTFLLWVLCWNHICGSQSDLDWWRTANFYQIYPRSFKDSNGDGIGDLNGITERLPYLKDLGIHAVWLSPIFKSPMKDFGYDISDF
YDIQPEYGTLQDFDRLIRIAHDLNLKILLDFVPNHSSNENEWFEKSANRDPDFEDFYVWHPGKDDPSDPIKKLPPSNWRSVFRGSAW
TYHNKRGEYYLHQFAHFQPDLNYRNPKVVDTMKGVLRYWLDKGVDGFRVDAVPNLFEILPQNGHYPDEPLSSLTEDEDSFDYLQ
HIYTMDQPETVEMLYQWREVLDVYTARDGNMRTMLSESDSEISILMTYYGDGIRNGSHVPFNFKLLFNLNKNSNAIDFRNTIHIWL
DNLPEGCVANWVIGNHDRNRVASRFGSFRIDLINIIINTLPGISITYYGEEIGMMDVFVTWEDSVDPAACNSDPDHYLEFSRDPARTP
FQWDDSTSSGFSTNPKTWLPVSPLYKEVNVKVERSSPRSHYKVYRKLLQLRRTPTLQKGNVETRTHGQNVLSITRSHPGEDTFITVV
NIGPDHEMVDLRELFSGRLIFYIVCVSSTRREGDVVRAPILNLDPFEAIILKGAIQYXXXXXXXXXX

S2. File Exon and introns sequences of *L. longipalpis* proteins from GH13 and GH31 families after sequence curation. The Xs represent missing nucleotides in incomplete sequences. For

missing introns sequences, size was estimated based on homology with complete orthologous genes.

***L. longipalpis* Glycoside Hydrolases from family 13**
Alpha-Amylase
> LLOJ004838_1/ LlAamyA1

```
>LLOTMP004838_1-E1 exon: PUTATIVE_protein_coding
ATGAAGTTCTGTGTTTTCCTCCTATTGACCTTTGTGGCCCTGGCCAGGGGTCAATTTGATCCGCATTTTGTTTCCGGACGCAGT
GGAATTGTTCACCTTTTTGAGTGGAAATGGGAAGATATTGCAGCTGAGTGCGAAAGATTCTTGGGTCCAAATGGTTTTGCTGG
AGTCCAAGTTTCACCACCAAATGAGAATGTTATCGTAACGAATCGTCCATGGTGGGAACGTTATCAGCCCATATCTTATAGGA
TTGTAACGAGATCAGGGAATGAGCAACAATTCGCAAACATGGTCTACAGATGCAACAATGTTGGTGTTCGCATCTACCCTGAT
ATTGTGATCAATCACATGGCTGCTTCTCATCCAGTTATGATG
>LLOTMP004838_1-E2 exon: PUTATIVE_protein_coding
GGAACTGGAGGGTCTACTGCTAACGTTGGAAACCGAGACTACCCAGCTGTGCCCTACAGCATCCATGATTTCAACCCTTCATG
CCCTATTACCAACTACAATGATCGCTATCAAGTGAGGAATTGTGAGCTAGTAGGCTTGCCAGATCTCAATCAAGGAGTTGCCT
GGGTCAGGGATAGAATTGTTGAGTATTTGGACAATCTTGTGGCACTGGGTGTTGCTGGATTCCGCGTTGATGCAGCCAAGCAC
ATGTGGCCAGGTGATCTGGAGATCATCTATGGACGCGTAAGAGAACTGAATGTAGTTCACGGCTTCCCAGTTCATGCTAGACC
ATTCTTCTTCCAAGAAGTCATTGACTTGGGCGGTGAAGCTATATCCAAG
>LLOTMP004838_1-E3 exon: PUTATIVE_protein_coding
AACGAATACACGGGCTTTGGTGTTGTGACTGAATTCAAATTTTCCGCTGAAATTGGTCGTGTCTTCCGTGGCAATGATCTTCTG
CATCACCTGTCCAATTGGGGTGAAGCATGGGGCTTCCTTCCATCCAATCGTGCTCTGGTGTTTGTTGAGAATCACGACAATGA
ACGCGGACACGGAGCTGGAGGAGACCAAATCCTAACCTACAAAAATGCTAAACAATACAAAATGGCCGTTGCTTTTACACTT
GCGCATCCATTCGGAATTCCACGTATCATGTCCTCCTTTGCGTTCACCAACACCGAAATTGGTCCACCAATGGATGGGAATCA
GAATATTATTTCTCCCAGTATTAATGCCGATGGTACGTGTGGCAATGGATGGGTGTGCCAGCACAGATGGAGGCAGATCTTCA
ATATGATTCGCTTCCGGAATGAAGCAGGAACAGCTGGACTTACAAATTGGTGGTCAAATGGAAGCCAGCAGATTGCTTTTGC
ACGTGCCGGTCGGGCTTTTGTTGCCTTCAACAATCAAGGATCCAACATGAATGTAAATCTTCAGACTTCCCTTCCTGGTGGAA
CGTACTGCGATGTTATCTCGGGGAATGTTTCCGGAGGATCATGTTCCGGAAAGACCGTTACAGTCAATGGAGACGGAACTGC
CAATATAGTCATTGGAGCAGCTGAAGAAGATGGCGTGCTTGCTATTCACGTTGGTGCAAGATTGTA
>LLOTMP004838_1-Intron 1: PUTATIVE_protein_coding
TGAGTGTTTGTTTACGCATAAGAAAATGAGGAAGGAGATTAAAAGTTATTGTTTTTTTTGTAGT
>LLOTMP004838_1-Intron 2: PUTATIVE_protein_coding
TAAGGACAACAACGTCTTTCCTCCTTCCTTGCTCCTAATTCTTTGACTTTCTTTTCCTAGG
```

> LLOJ004838_2/ LlAamyA2

```
>LLOTMP004838_2-E1 exon: PUTATIVE_protein_coding
ATGAAAATCTTATGGATTTTCTTCTTGGCCATAGTGGCCACTGTTCGTGGTCAATTTGATCCGCACTACGTTGCAGGTCGTAGT
GTGATGGTCTTCCTTATGGACTGGAAGTATGAGGACGTTGCTCTGGAGTGTGAGAGATTCCTCGGACCAATGGGCTTTGCTGG
TGTTCAGGTTTCAGCTGCTTCTGAGAATGATATTGTTACGAGTCCACTTCGTCCTTGGTGGGAACGCTATCAAGTTGTTTCCTA
CAAGATCGGTTCGAGATCTGGAGATGCAGACAGTTTTGCAGACATGTGCACCAGATGCAATAACGTTGGGGTACGCGTCTAC
GTTGATGCCATCTTCAACAACATGGGGGGTATGGGGATTATGAA
>LLOTMP004838_2-E2 exon: PUTATIVE_protein_coding
AGGAACAGCTGGATCATTGGGAGATTCAATAACAAGAACCTTCCCTGCTGTCCCCTACACTATTCTAAATTTCCACAATCCAC
CATGTTCGATAACAGACTACAACAACACGCAACAGGTTCGCAATTGCGATCTTGTAGGTGCTCCGGATCTTGATCAATCTCAA
CCCTATGTACGAGATAGAATTGTAGGTTATCTGAATCATCTAACTTCCCTTGGGTGCGCTGGTTTTCGCATTGATTCAGCCAAG
CATATGTGGCCAACGGATCTCCAGGCCATCTATAGCAGTGTCATAGACTTAAATACAGCCCAAGGATTTCCAGCTGGTAGCAG
ATCATTCATCTTCCAAGAAGTCATTGATTCTGGCCATGAAGCAATTAAGAAGTATACTCTGAAAAAAATAGTTTATATTTGA
>LLOTMP004838_2-E3 exon: PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
>LLOTMP004838_2-Intron 1: PUTATIVE_protein_coding
GTAAGTGCAGTCATGCAACATCAAGATGATCTGGAGTGTCTGGTCTTTTTTGACAATCTTTCATTACAG
>LLOTMP004838_2-Intron 2: PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```

>LLOTMP004839/ LlAamyA3

```
>LLOTMP004839-E1 exon :PUTATIVE_protein_coding
ATGTTCTGGAATCTTTTTTTGACAACGTTGGCCATAGCCCGTGGTCAATTTAACCCAAACTACGTACCAGGACGTACGGTAAT
GGTCTTCCTTATGGACTGGAAGTATCAGGATGTCGCATTGGAGTGTGAGAGATTCCTGGGACCGAAAGGCTTTGGTGGAGTCC
AAGTTTCATCTCCAGCTGAGAATGATATTGTTACGGATCCTTTGCGTCCTTGGTGGGAACGCTATCAAATTATTTCCTACATCA
TCGGATCGAGGTCTGGAAACGAAGATGATTTTGCAGACATGTGTTACAGATGCAATAAAGTCGGAGTTCGTATCTACGTTGAC
GCCATTTTCAATCATATGGGGTCTCTACCAAACATG
>LLOTMP004839-E2 exon :PUTATIVE_protein_coding
TTGGAACTGCTGGATCAACAGCAGATGGACCAACTAAGTCCTATCCAGCTGTACCGTATGACGCATCAAACTTCCATCCTTCC
TGCCCTCTTACTACATTCCAAGTTACCTTCATTGTTCGCAACTGTGAACTGTTTGGTGCTCCTGATCTTGATCAAGCGCAACCC
TATGTCCGTGAGAAAATCAAAGGATATCTCGATCATTTGATTACTCTCGGAATAGCTGGTTTTCGCGTTGATTCCGCCAAGCA
TATGTGGCCTAAAGACCTCAAGGCTATTTACAGCAGCCTAAGAGATCTTAATACAGATTTTGGTTTCCCGCCAAAATCTAGGC
CTTTCGTTTTTCAAGAAGTTATCGATGCAGGCGGTGGCTCTGAACCCATTAACAA
>LLOTMP004839-E3 exon :PUTATIVE_protein_coding
```

GTATGAATATTCACCAATTGGAGTTGTTACTGAATTCCTCTACTCATACATTTTGACTGACTTCTTCCGTGGCAATCGAAATTT
GAGCGATTTAGCCACGTGGGGTCAATCTGAGTGGAAGTTCCTTCCATCGAAGGCTGCTCTGGTGTTTGTTGAGAACCACGACA
ACGAGCGGGCAACACCTCCTGTGCTTAACTACAAGGATGGCAAACCGTACTGGATGGCTGTTGCTTTTGAATTAGCTCATCCA
TACGGCTACCCTTATCTCATGTCAGATTTTAACTGGACAACCTACGATCAGGGACCACCAATGGACGAAGATCAGAACATAAT
CTCACCAACAATCAATAAAGATGGCAGCTGCGGAAGTGGATGGGTTTGCCAGCATCGTTGGAGGCAGATCTACAGCATGGTT
GGCTTTAGGAACCAAGTAGGCTACGCACGGTTTACTAACTTCTGGTCTAATGGCAAGAATCAGATTGCTTTTGCTCGTCAAGG
ATGCGGTTTTGTTGCCTTCAACAATGATGATGTTGATCTGAATGAAGTGCTCCAGACTACTCTTCCAGCAGGCACTTACTGTGA
CGTTTCCACAGGATACGCCAAAGCAGGTGAATGCACTGGAAGGACTGTCGTTGTGGGAGATGATGGAACAGCCACCATTGAA
ATTTCTTCATCAGATGATATTGGATTCATTGCTATCCATATTGGGCAGATGTTGTAA
>LLOTMP004839-Intron 1: PUTATIVE_protein_coding
GTGTAAGTTGAATCAGTTTAAAATGAAATCCAGGATCAAATCTTCTTTCCTGATTTCTTTTTTTTTTAG
>LLOTMP004839-Iintron 2: PUTATIVE_protein_coding
GTATAAACCTTCGACTTTTAAGTCTGATCGTTTTGACAGAATTTCTTTAAACTTTTTTACAG

>LLOTMP004841/ LlAamyA4
>LLOTMP004841-E1 exon: PUTATIVE_protein_coding
TGGGGAATAAAATGCGGATAGTTGTGTTGATTTTGTGCTTGGTTGCCCTGGGAAGTGGTCAGCATTGCCAAACTTCTGGGATG
GCCGAACGGGTATTGTGCATTTGTTTGAGTGGAAATTCCGTGATGTGGCTGATGAGTGTGAACGCTTCCTGGCACCCAAAGGC
TATGCCGGTGTTCAGTTGTCGCCAGTCAATGAGAACGTAATCATTCCCAATCGTCCATGGTGGGAACGCTACCAACCAATTTC
CTACGTCCTCACAACACGATCTGGAAATGAGCAGGATTTCTTAGATATGTCCAGAAGGTGCAATGCGGTGGGAATTCGAATTT
ACGTTGATGTTCTCCTGAATCACATGTCTGCAGATAATTATCCATCAGATGGAACTGCTGGTAACTGGGCAGATACATCACAA
CGTTGGTGGCCAGCTGTTCCATATGGTCCCAATGACTTCAATCCTCGCTGCGAAATTCACAACTGGGACGATCCATGGGAGAT
CAGGCAATGCGAACTAGTTGGTCTGCATGATCTCAATCAGGGCACTGAACACGTACGTAGTATGTTGGTGAACTTCCTGAATC
ACCTTGTCGAACTGGGCGTTGCTGGTTTCCGCGTTGATGCTGCAAAGCACGTCTGGCCTCATGATCTAGAAGTAATCTACAAT
CGCGTCAATAATCTGAATACTAACTTCGGTTTTGCTCCTGGAAGTCGCCCTTACATCTACCAGGAAGTTATTGATCTAGGTGG
AGGTGGAATTCGGTATGAATACACACCTCTAGGAGCTATAACAGAGTTCAAGTTCTCCGCCGAAATTGGAGGAGCCTTCAGA
CGTTACAACGATCTGCGCTGGTTGCGGAACTTCGGTCCTGAATGGGGCTTTGTTCCATCTGATGTTGCTCTGGTATTTGTTGAC
AATCACGACAACCAGCGTGGACATGGAGCAGGTGGTGAAGCCGTTCTGACCTACAAACAGCCTAGACAATACAAGATGGCA
ACAGCCTTCACAATGGCCTGGAACTACGGAACAATTCGTCTAATGTCGTCATTTGCCTTCGATCATGGAGATCAGGGACCACC
TCAGGATGCCAATGAGAACATCATTTCCCCATCAATCAATCCTGACAACACTTGTGGCAATGGATGGGTATGCGAACATCGCT
GGAGGCAGATCTACAACATGATTGGTTTTGCTAACCAAGTCAAGGGAACATCAGTTAACGACTGGTGGGACAACGGCCATAA
CCAGATAGCCTTCTGCCGAGGAGATCAAG
>LLOTMP004841-E2 exon: PUTATIVE_protein_coding
GCTTCATTGCTTTCAACAATGAGGATCACAACTTCAGTCATACACTGCAAACCTGCTTACCAGCTGGAACGTACTGCGACGTC
ATTTCCGGTGAAAATACCGGAAATGGTTGCTCAGGCGTCACAATCACAGTAAACTGGGATCAAACAGCTCACATTGTCATCCC
TGTTGATCACTACGACGGTGTTATTGCCATTCATCGTGGTCCAACATCGAGGATCGGTTAAACAACAGCGTCAGTAAACTTCC
CAGTCAATTTCGTGATACATATGACGGAATTTACTCATCCAAAAGTCTGCGCAGTCTTTTGAGCAGAAGATCGTACTGAATAC
TACTCCACTTGGATTCTGAAGTCTGAAAATTAAAGAAAATGATGAAATAATTTAAAAGAAA
>LLOTMP004841-Intron 1: PUTATIVE_protein_coding
GTATGAAAATGGCTTGAAAATTGGCTAAACTCTGACTTGTAACTTTAATGAACTTTAATCCTAAACTTATTTACAG

>LLOTMP004880_1/LlAamyB1
>LLOTMP004880_1-E1 exon: PUTATIVE_protein_coding
CGTATTTACCCAAAATTATCATTAAAAGTGAATGACCAAAGAATTTTTTGGCACCTTTCCTCGGAGTATAAAAGACGGTAATC
ACGTTTCATTTCTGCAGTATGACTTTTTGATTCCATTAAATCAGTCATGAAAGTTTTAGTGATATTCTTAGCCCTTGTGGCCTTA
GCTAAGGGTCAGTTTAATCCTTACTACGAGGATGGTCGAACTGTGATGGCTTTCCTTATGGACTGGAAGTACTCAGACATTGC
TCTGGAGTGTGAGAGATTCCTTGGACCGAAGGGGTACGCCGGTGTCCAGGTTTCATCCCCATCTGAGAATGCAATTGTAACAG
ATCCTCTGCGTCCTTGGTGGGAACGCTATCATGTTGTATCGTACATCATTGGATCCAGATCAGGGACTCAGGATGATTTTGCC
GACATGAGTGCAAGGTGCAACAAATGCGGAGTTCGTGTTTACGTCTTTACCATCTTCAATAATATGGGAGCGTTTCCAGATAT
GGT
>LLOTMP004880_1-RA-E2 exon: PUTATIVE_protein_coding
TGGAACTGCTGGATCTACAGCAGATGGACCATCTAAATCCTACCCAGCTGTACCTTATGACAGTACAAACTTCCATCCATCCT
GTCCCATTCAATCTTACCAGCATGCCTTCATTGTACGTAACTGCGAAAATGAAGGTACTCCTGATCTCGATCAATCCCAAGAC
TACGTCAGGGAACAAATTGTTGCCTATCTCAACAATTTGATAAGCCTCGGAGCTGCTGGCTTTCTCATTGATTCAGCCAAGAA
TATGTGGCCAACAGATCTGAAAGCTATCTACAGTAGCCTCTCGAATTTGAATGAAGACTACGGATTTTCTCCCAACTGCCGGC
CATTCATCTACCAAGATGTCATTGAGTTAGGCAAGGAAGCTGTTAACAA
>LLOTMP004880_1-E3 exon: PUTATIVE_protein_coding
ATATGAGTACCTTCCCTTAGGTGATGTGACTGATTACAAAGTTTCATATAGCCTTAATAGTATCTTCAACGGCAATAAATCGCT
CAGTGTTTTAGCCACTTGGGGAGAATCTCTCTGGAACTTTATTCCATCGAAAAGTGCTGTAGTTTTCGTAGAAAATGAACAAA
ATGAACGCGAAGGATTCTACTTTGGACAAGTCCTTAACTACAAAGATGGAAAACCCTACATAATGGCTATTGCTTTCCTCTTG
GCTCATCCATACGGTAATCCACGCATCATGTCAGACTTTGACTGGTCAGACTCCGATCAAGGACCTCCAATGGATGATAATCA
GAACATCATCTCACCGACTATCAATAAAGATGGTAGTTGCGGTGATGGATGGGTTTGTCAGCATAGATGGCCTCAGGTTTGCA
GCATGGTTACATTCAGGAATCAAGTGGGTAATTCCAATTTCTGCAACTTCTGGTCCAACGGCAACAATCAGATCGCCTTTGCG
CGTGAAGGACATGGCTTCATTGTCTTCAACAATAACGACGGGAAACTCTGGCAGACTCTGCCGACAACCCTTCCGGCTGGGA
CCTACTGCGACGTCATCACAGGTTATGTTAAAAATAATGAATGCACTGGAAACAGCATTGAGGTAGCAGACGATGGAACAGC
AGAATTCATCATCTACTCTGATGATCCAAATGGCGTGATCGCCATTCACATTGGACAAATGTTGGAATGA
>LLOTMP004880_1-Intron 1:PUTATIVE_protein_coding
GTAAGTTGAACAGCTTTAAAAGAGTCCTTGTTTTCCTGAAACAGTCTTCATGGTTTTTTCAAACTTACAG
>LLOTMP004880_1-Intron 2: PUTATIVE_protein_coding
GTAATTAAGAGTAGTTCTTTTGAAAACAATGAGGGAAAACTAAAATGCCTTCTCTTTCAG

>LLOTMP004880_2/LlAamyB2
>LLOTMP004880_2-E1 exon: PUTATIVE_protein_coding
ATGAGAAGTGCGTTAATTTTCCTTGCCTTTGTGGTCTTAGCTCGGGGACAGCATAACCCAAATTTCGTTCCTGGACACACTGGT
ATAGTTCATCTATTCGAGTGGAAATGGGTGGATGTTGCTGCTGAATGTGAAAGATTTCTCGGACCTATGGGCTTTGGAGGAGT
TCAAGTATCCTCTCCAGCTGAAAATGCCATTGTGAGGGATCCTCTGCGTCCTTGGTGGGAACGCTATCACATCATTTCGTACA

AACTTGTAACGCGATCTGGGAATGAGGAAGAATTTGCCGATATGTGTACAAGATGCAATAAAGTCGGAGTTCGTATCTACGT
TGATACGGTTATCAACCACATGGGTTCTCTTCCGAATATGAT
>LLOTMP004880_2-E2 exon: PUTATIVE_protein_coding
TGGAACTGGAGGATCAACAGCAGATGGACCTAGCAAGTACTACCCAGCTGTACCCTACAACGTTGATCATTTTCATACATCTT
GCCCAATAGAGAACTGGCAAGATGAATATGAAATTCGTTATTGTGAATTATTCGGTGCTCCTGACTTGAATCAGGGAATGGGC
TATGTCAGAACTACAATTGCGGAGTACATGAATCACCTGATACGTCTCGGAGTAGCTGGTTTCCGTATTGATGCAGCTAAGCA
CATTAGTTCTTCTCATTTATATTTTACAACTATTGATTTAACATACTTAAACGACTTAAATCCAGATCACGGATTTCCCGCCAG
AACAAAACCTTTCATCTACCAAGAAGTCATAGACTTTGGCTGGGATGCTGTTTCAAA
>LLOTMP004880_2-E3 exon: PUTATIVE_protein_coding
GTACGATTACACCTACATAGGTACGGTAACGGAATTTATGTATTCATACTACATCAGCCGAGCCTTCCGTGGTAACGATTTAC
TGAGGAATTTGGCAACTTGGGGACCAACTTGGGGTTTCCTAAACTCCTCATATGCTCTTGTTTTTGTTGAGAATCACGACAATG
AACGCGGTGATGGAGCTGGAGGGCAGGAGATTCTCTCCTACAAGGATGGCAAGCAATACCGTATGGCCGTTGCTTTTGCCCT
AGCTCACACCTATGGTATCCCACGCATCATGTCCTCTTTCGCCTTCACTGAGAGAGATATTGGCCCACCAATGGATTCAAATG
AGAACATTATCTCACCCAGTATTTACCCGGATGGAAGTTGCGGAAACGGTTGGGTGTGTCAGCACAGATGGAGGCAGGTAGT
CAATATGATTGCCTTCAGGAATATTTCGGGGAATGAAAATATCACAAATTGGTGGACCAATGACAAAAATCAGATTGCCTTTG
CGAGGCACCAGCGTGGCTTCATTGCTTTCAATAACGACGCAATGAATATGAATGTGAGCCTGCAAACGTCCCTACCGGCTGG
AATCTACTGCGATGTTATATCAGGGAATTTAGTGGATAAACTCTGCACAGGGAAGAACATAACCGTGAGAGGGGATGGAGTT
GCCAGGATTCTTATAATGAAAGATGAAGAAGATCAAGTTGTCGCTTTTCATGTTGGTCAATGA

>LLOTMP004880_2-Intron 1: PUTATIVE_protein_coding
GTAGGAGAGTTTCTTAATAGTTTTAATATAAGAGCTGCTAAGAGAGTATTTAATTTTGTTTATTTCTAG
>LLOTMP004880_2-Intron 2: PUTATIVE_protein_coding
GTACGTTGATAGTGTATATAAAATCAGAATTAGGATATGGCCCTTAGAGCCTATCTGTCAGTAAACTTTTAGTACTGATCCTA
GAAAAGATCCAAGACATGGTCTGAGACAGATCTCACAACGGCTTTTCTTTAAATTTCAG

>LLOTMP004880_3/ LlAamyB3
>LLOTMP004880_3-E1 exon: PUTATIVE_protein_coding
ATGCATTTTATCAGTCTCGGTGGCGGAATTCTTTTGGTGCTGGTGCTTCATGTTGCGGATGGTCAATTTGATCCACACTTCCTC
CCTGGAAGAAGTGTTATTGTGCATCTGTTTGAGTGGAAATTCTCCGATATAGCCGCGGAGTGTGAGAATTATTTGGGCCCAAA
TGGATTTGGGGGCGTCCAGGTGTCCCCGATAAATGAATGTCTTGTATCACCTGAGAGAGCGTGGTGGGAACGGTATCAACCC
GTGTCGTATGCTATTGTAAGTCGATCAGGGGATGAAAAAGAATTTGCGGAGATGGTTAAACGATGCTACGAGGCTGGGGTAC
GTGTTTACGTTGACGTCATCTTCAATCACATGGCATCTGGGGAGGGTGAAGTTGTTGGAACTGGAGGAAGCTTGGTCTACCCA
GAAGAGCGTCTCTACCCGCACGTACCCTACGGTCCAGAAGACTTCAACCCCCATTGCGTCATTGAGGACTATCAGGATGTAGA
TCAGGTGCGAAATTGCGCTCTTGTCTCCCTTCCAGATCTCAATCAGAAGTCCGATAATGTAAAGAGGAGTGTCATTGAATTCC
TGGATCGTCTCATTGATCACGGTGTTGCGGGATTTCGGGCAGATGCATGCAAGCACATGTGGCCGGAGGATATTAAATTCCTC
TTTGGCAGTACGAAGAATCTAAGTCCTGAATTTGGATTCCCCGACAAGGCTAGACCATTTCTCTACCAGGAAGTTATTGACCT
GGGCCAGGAACCGATATCCAA
>LLOTMP004880_3-E2 exon: PUTATIVE_protein_coding
AAACGAATACACCTCCATAGGCGTTGTTACGGAGTTCCTATTCTCAGCTGAGATTGGAAACATCTTCCGGTACAAGAAACTCA
AGGAATTCATGAAATGGGGTACAGATGAGAGATTCCTGCGATCTGATAGAGCTCTTGTATTTGTTGAGAATCACGACAATGA
GCGTGGGCATGGAGCAGGTGGAAAGGATATTCTAACGTACAAAGATGGCAAACGCTACCTCCTGGCCGTTCTCTTTACAATT
GCCCATCCATACGGCATCCCACGCATCATGAGTTCGTACGACTTCTCGAATAGTGACGAAGGCCCACCGGCCAATGCAAAAG
GTGAAATAATCTCACCTGTTTTCAATGACAGAGGTCTCTGTACGAATGGTTGGATCTGTCAACATCGCTGGTTAGGCGTTGCA
AGTATGGTGCAGTTCCGTAATGCTGTTGCCGGTAGTGGTATCGTCAACTGGATGGACAACGGAGAACAGCAATTTGCCTTCTG
TCGTGGAGATCTGGGGTTCGTCGCCTTCAACGGCTACACCATGTCCCATCTTAATCAACAGTTAAGGCTGACCTACCACCTGG
CATCTACTGACGACGGTAATCAGCGGAGAAGTCACACTAGATGGCTGTACTGGGCTCGAGGTTGTTGTCAAAAATGACGGTT
ATGCCAACATTTTCATTCCAGGCGACTCCCCAACAGGTGTTCTTGCCGTTCACCTGGGCTCTGCCTATATTATCGATTAA
>LLOTMP004880_3-Intron 1: PUTATIVE_protein_coding
GTATGTTCCTACGTCTTAATCTGATGTTCCTTTGTTTCCTAACTTCCTTTAAATTCTCACCCATCAATAG

>LLOTMP004881_1/ LlAamyB4
>LLOTMP004881_1-E1 exon: PUTATIVE_protein_coding
ATGAAAGCCACGATCTTTGCTTTACTTTCCCTGCTAGCCCAAACTAAGGGCTATTTTAACCCCAATTTTGTACCAGGTCGTAGT
GGGATTGTGTGGCTTTTTGAGTGGAAGTTCGCGGATATTGCTTTGGAATGTGAACGATTCCTCGGACCGAATGGCTTTGCAGG
TGTTCAAATTTCATCCCCAGCTGAGAATGCAATCGTTACGAACCCCTTGAGGCCATGGTGGGAGCGCTACAACCCAGTCTCCT
ACATCATTGCCTCGAGATCAGGCAATGAACAGGATTTTAGTGATATGGTTACTCGTTGCAATAATGTCGGAGTACGGACTTAT
GTTGAAATCATCATGAATAACATGGGGGGTTATCCTAATATGGT
>LLOTMP004881_1-E2 exon: PUTATIVE_protein_coding
AGGAACTGGAGGATCAACAGCAGATGGTCCAAGCAGGAGTTTCCCAGCTGTTCCCTACACATCTGCGGATTTTCATGCAACAT
GCTCCATAACTACAATGCTCCTGCCTTTCCTAATACGCAATTGTGCCGTTGAAGGCTTACCAGATCTTAATCAGAGCGGAGCA
CATACCAGAGATATGATTGTGGGTTACCTTAATCATCTAATTTCACTCGGCGTTGCAGGATTTCGTATAAACAGTGCAAAGTT
CGTCTGGCCGTCTGACCTTAAAGTTATCTACGACAGGACTAATAATTTGAATACAAACTACGGCTTTGCTCCCAATAGTCGTC
CGTTTATCTTCCAGGATGTTGCAGACTATGGACTTGATGCGATTTCCAA
>LLOTMP004881_1-E3 exon: PUTATIVE_protein_coding
GTATGAATACTCCTCAATGGCGGTTGTTTCGGAGTACAAGTTTCTCCTCAATCTTGGTCGAATTTTCCGCCATCTGGATGTTTT
GAGTGTCCTAACTAACTGGGTGCAGACTTTTGGACTTCTTCCATCAGATGCGGCTTTAGTCTTCGTTGACAGTCCCGATAGCGA
GCGTGGGAGTCAAGGAAGTGAAATCCTAACCTACAAAGATGGCAAGGCGTACATTTTGGCTCTAATCTACATGTTAGCCGTC
CCCTATGGCCATCCGATAATTATGTCTTCCTACGAATTTAAGGATGTGAATTCTGGTCCACCAATGGATGCAAACCAAAATAT
TGTCTCGCCTACAATAAATCCCGATGGAAGCTGCGGGAGTGGTTATGTGTGTCAGCACAGATGGACACCGGTGTCCAGTATG
GTGGCTTTCAGTAACAATGTTGGCGATGCTCCAGTTACAAATGTGTGGACAAATGGTCAAAATCAATTGGCTTTCGCTCGGGA
TGGAGTTGGTTTTGTTGCTATTAACAGCAACGCAGAGACTTTGGATGTAAGTGTACAGACTGGTCTTGTAGGTGGAGTATACT
GCAACGTCCTCGGAGGACCTATAGTTGATGGAACTTGCGCTGGACAGACTGTTGTTGTTGGAGATGATGGAATTGCTGCAATT
ACCATAACAACACAAGACAACCTCGGAGCTATTGTACTTGATGTTAATCAACAACTACTTAATGGTGAAGCAAAGAAATTTA
AACAAGAACATCCGTCCTTGTATTAA
>LLOTMP004881_1-Intron 1: PUTATIVE_protein_coding
GTAAGTAAGAACCAAGATATTATGTCAGAAGGATACAAAGAAGGCTTCTTTACAG
>LLOTMP004881_1-Intron 2: PUTATIVE_protein_coding

GTAATATTCAGACCAGCTTTTTCAGAAAAATAGAAAAGTCTAGTATTTTATATTTCTAG

>LLOTMP004881_2/ LlAamyB5
>LLOTMP004881_2-E1 exon: PUTATIVE_protein_coding
ATGAAGTTCTTCATGATACTCTTGGCCCTTGTGGCCTCCACTAGGGGCCAGTTTAATCCGTACTTTTGGCCTAACCACAATACT
ATTGTCTACCTCTTGGATTGGAAGTACTCAGACGTTGCTCTGGAATGCGAGAACTTCCTTGGACCGAAAGGCTTTGGTGGAGT
TCAGGTATCGTCTCCTTCAGAAAATGCCATTATTACGGATCCCTACAATCGTCCATGGTGGGAACGCTTCCATGTTTTGTCCTA
CATAATCGGCTCGAGATCTGGCAATCAGGATGACTTTGCAGACATGACTGCTAGATGCAACAAGTGCGGAGTTCGCGTTTTTG
TAGATGCTATCTTTAACGAAATGGGTGGTTTACCAGATATGGT
>LLOTMP004881_2-E2 exon: PUTATIVE_protein_coding
TGGAACAGCTGGATCTACTGGAGACGGCCCATCAAGGACTTTCCCTGCTGTTCCCTATGACAGCTCAAACTTTCATCCACCTT
GCCCTATTCAATCCTGGCTACATGCCTTCATCGTTCGAAATTG
>LLOTMP004881_2-E3 exon: PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
>LLOTMP004881_2-E4 exon: PUTATIVE_protein_coding
ACAAGAAGGGAACCGTGTGACAGTCCTTATGTATGCCAACATAGATGGCCTCAGGTTTGCAGTATGGTTACATTCAGGAATC
AAGTTGGTGACGCCAATTTCTGTAACTTCTGGTCCAACGGCAACAATCAGATCGCCTTTGCGCGTGAAGGACAAGGCTTCATT
GTCTTTAACAATAACGACAAGAAGCTGTGGCAGACTCTGCAGACAACCCTTCCTGAGGGAACCTACTGCGACGTCATCACAG
GTTATGTTAAAAATGATGAATGCACTGGAAGGAGCATTGAGGTTGCAGACGATGGGACAGCTGAATTCATCATCTACACTGA
TGATCCAAATGGG
>LLOTMP004881_2-Intron 1: PUTATIVE_protein_coding
GTAAGTAAAGGAGATTCTTATGTAAGACTCTTTAGGTGCCTTTTGTGATCAGAACGATCGATAGCTCCAAAAACCGCTTATAT
TATACTTGACAG
>LLOTMP004881_2-Intron 2: PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXX
>LLOTMP004881_2-Intron 3: PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXX

>LLOTMP004881_3/ LlAamyB6
>LLOTMP004881_3-E1 exon: PUTATIVE_protein_coding
ATGAAGCTCTTCGTGGTATTCTTGGCCCTTGTAGCCTTCACCAGGGGCCAGTTTAATCCGTACTTTTGGCCTGGACACAATACT
ATTGTTTACCTCCTGGACTGGAAGTACTCAGACGTTGCTCTGGAATGTGAGAACTTCCTTGGACCGAAAGGCTTTGGTGGAGT
TCAAGTATCGTCCCCTTCAGAAAATGCCATTATTACTGATCCCTACAATCGTCCATGGTGGGAACGCTTCCATGTTTTATCCTA
CATAATCGGCTCGAGATCTGGCAATCAAGATGATTTTGCAGACATGACTGCTAGATGCAACAAATGCGGAGTTCGCGTTTTTG
TAGATGCTGTCTTCAACAATATGGGTGGTTTACCAGACATGGT
>LLOTMP004881_3-E2 exon: PUTATIVE_protein_coding
TGGAACAGCTGGATCTACTGGAGACGGCCCATCGATGACTTTCCCTGCTGTTCCCTATGACAGCTCTAACTTTCATCCACCCTG
CCCTATTCAATCCTGGCTACATGGTTTTATTATTCGCAATTGCGAAGATCCCGATTTCTTGCCTGTTCCTGATCTCGATCAATCC
CAAGACTACGTCAGGGAACAAATTGTTACCTATCTCAACAATTTGATAAGCCTTGGAGCAGCTGGTTTCCGTATTTTTGCAGC
TAAGCACATGTGGCCAACGGATTTGGAAGCTATCTATAGTAGTCTTTCGAATCTGAATGAGGACTACGGATTTTCTCCAAACT
GCCGACCATACATCTACCAAGATGTCATTGATTTGGGCGGTGAAGCCGTTAACAA
>LLOTMP004881_3-E3 exon: PUTATIVE_protein_coding
GTATGAATATATCCCCCTAGGGCTTGTGTCTGAATACAATGCGTCACAGAGTCTTGCAGATATCTTCCGTGGCAATCGACCGT
TGAGTGTTTTAGCCACTTGGGGAGAGGCTGAATGGAATTTTGTTCCATCCGCAAATGCAATTGTTTTCGTGGAAAACGAACAA
AATGAGCTTGACACATGGTACTTTGGAAGAGTCCTCAACTACAAAGATGGTAAACCCTACATAATGGCTATTGCTTTCCTCTT
GGCTCACCCATTTGGCAATCCACTCCTTAATTCGGACTACGAGTGGACTGATCCCAACGAAGGTCCCCCAATGGATGATGATC
AGAACATTCTTTCGCCGGTTATTAACAAGAAGGGAACCTGTGACAGTCCTTATGTATGCCAACATAGATGGCCTCAGGTTTGC
AGTATGGTTACATTCAGGAATCAAGTTGGTGACGCCAATTTCTGTAACTTCTGGTCCAACGGCAACAATCAGATCGCCTTTGC
GCGTGAAGGACAAGGCTTCATTGTCTTTAACAATAACGACAAGAAGCTCTGGCAGACTTTGCAGACAACCCTTCCTGAAGGA
ACCTACTGTGACGTCATCACAGGTTATGTTAAAAATGATGAATGCACTGGAAGAAGCATTGAGGTTGCAGACGATGGAACAG
CTGAATTCATCATCTACACTGATGATCCAAATGGGGTTATTGCTATTCATACTGGACAGATGTT
>LLOTMP004881_3-Intron 1: PUTATIVE_protein_coding
GTAAGTACAGATCAATTGATTCCTGTGTAAAATACTGGCGTTGTATTCCGTGATCAGGATGTTCCTCCTCGTTATGTAGAGTCT
AGGAAAGGGTGCTAACTAAGTGTTTTAGACTGTAGGTGTATATATTGTGTATTTTCTATGTGCCCTCCAATATTTAGGTTAGTT
TTTCATTAAGTCTTGTCAGAAGTGAATAGGTTTGTTAAGTGCCCAAGGCTGTAAAGGGGTCTGAACTTATTTTCTGTTAAATTA
TCAATTTTATACTTGACAG
>LLOTMP004881_3-Intron 2: PUTATIVE_protein_coding
GTAAGCTAACAAGCATATCTGAATAGACTTTCAAACAAAGCTAAAATACTCCTACTTGTAG

>LLOTMP004882/ LlAamyB7
>LLOTMP004882-E1 exon: PUTATIVE_protein_coding
AATTCTCATAAGTCTTTAGTATATTTCTGAGTTTTGATCGTTGAATTTCATTTGAAATGAAAATCTTTTGGATTTTCTTCTTGGC
CACAGTGGCCACTGTTCGTGGTCAATTTGATCCGCACTACGTATCAGGTCGTAGTGTCATGGTCTTCCTTATGGACTGGAAGT
ATGAAGACGTTGCTCTGGAGTGTGAGAGATTCCTCGGACCTATGGGCTTTGCTGGTGTTCAGGTTTCAGCTGCTTCTGAGAAT
GATATTGTTACAAGTCCACTTCGTCCTTGGTGGGAACGCTATCAAGTTGTCTCATACAAGATCGGTTCGAGATCTGGAGATGC
AGACAGTTTTGCAGACATGTGCACCAGATGCAATAACGTTGGGGTACGCGTCTACGTTGATGCCATCTTCAACAACATGGGG
GGTATGGGGATTATGCA
>LLOTMP004882-E2 exon: PUTATIVE_protein_coding

TGGAACAGCCGGATCATTGGGAGATTCAATAACAAGAACCTTCCCTGCTGTCCCCTACACTATTCTAAATTTCCATAATCCAC
CATGTTCGATAACAGACTACAACAACACGCAACAGGTTCGCAATTGCGATCTTGTAGGTGCTCCGGATCTTGATCAATCTCAA
CCCTATGTGCGAGATAAAATTGTAGGTTATCTGAATCACCTAACTTCCCTGGGTTGCGCTGGTTTTCGCATCGATTCAGCCAAG
CATATGTGGCCAACGGATCTCCAGGCCATCTATAGCAGTGTCATAGACTAAATACAGCCCAAGGATTTCCAGCTGGCAGCAG
ATCATTCATCTACCAAGAAGTCATTGATTCTGGCCATGAAGCAATTAAGAAGTATACXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXGAT
>LLOTMP004882-E3 exon: PUTATIVE_protein_coding
GCAATGAACCTCTGAGCACCTTGAGTACTTGGGGTCAATCAAACTGGGGATTCCTTCCATCGGATCGTGCTTTGGTGTTTGTTG
AGAATCACGACAACGAACGTGGTCAATACGAATCAGCAGCGCTCAACTATAAAGATGGCAAACCCTACAAAATGGCTGTGGC
CTTCGCTTTAGCACATCCATTTGGAAATTTACGCATCATGTCCGATTTTGACTGGACAACCTACGATCAGGGACCACCAATGG
ATGCCAATCAGGATATTATTTCACCCACTATCAATGCAGATGGTAGCTGTGGAAGTGGATGGGTATGTCAGCATCGTTGGAGG
CAGATTTACGGTATGGTTGGCTTCAGGAATACCGCAGGTACAACAGAAGTCACCAACTTCTGGTCCAACGGCAACAATCAGA
TTGGTTTTGCTCGTCAAGGACGTGCCTTCGTGGCTTTCAACAACAACGACGCGGATATGAATCAGACTCTTCAGACAACCCTT
CCAGCAGGCACGTACTGTGATGTCATAACCGGCGTTGCGGAGAATGGAGCCTGCACAGGAAGAACCATTGTTGTAGCAAATG
ATGGAACGGCTAACTTTGTTATCCTTACATCAGATGCAGATGGTGTGATTGCCATTCATGTTGACCAGAAATTGTAA
>LLOTMP004882-Intron 1: PUTATIVE_protein_coding
GTAAGTGCGGGTATACGGACAGATGGTCTGGAGTGTCTGGGCTGTTTTGGCAATCTTGTCTTGTTATCATATTAACTAAATTTC
TGTTACAG
>LLOTMP004882-Intron 2: PUTATIVE_protein_coding
GTTACTGAATTCCTCTATTCGTACTACCTTACTGATATCTTCAAAG

>LLOJ004885/ LlAamyB8
>LLOJ004885-E1 exon: Putative_protein_coding
ATGAAGCTCTTCGTGGTATTCTTGGCCCTTGTGGCCTTCACTAGAGGACAATTTAATCCGTACTTTTGGCCTAACCACAATACT
ATTGTCTACCTTTTGGACTGGAAGTACTCAGACGTTGCTCTGGAATGCGAGAACTTCCTTGGACCGAAAGGCTTTGGTGGAGT
TCAAGTATCGTCCCCTTCAGAAAATGCCATTATTACGGATCCCTACAATCGTCCATGGTGGGAACGCTTCCATGTTTTGTCCTA
CATAATCGGCTCGAGATCTGGCAATCAAGATGATTTTGCAGACATGACTGCTAGATGCAACAAATGCGGAGTTCGCGTTTTTG
TAGATGCTATCTTTAACGAAATGGGTGGTTTACCAGATATGGT
>LLOJ004885-E2 exon:Putative_protein_coding
TGGAACAGCTGGATCTACTGGAGACGGCCCATCAAGGACTTTCCCTGCTGTTCCCTATGACAGCTCAAACTTTCATCCACCTT
GCCCTATTCAATCCTGGCTACATGCCTTCATCGTTAGAAATTGCGAAAACGCAGATTACTTGCCTGTACCTGATCTGGATCAAT
CTCAAGACTACGTCAGGGAACAAATTGTTACTTATCTAAATAATTTGATAAGTCTCGGAGCTGCTGGTTTCCGTATTCTTCAT
>LLOJ004885-E3 exon: Putative_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XX
>LLOJ004885-Intron 1:Putative_protein_coding
GTAAGTATATAAGGAGTATATTCTTATGCAAGACACTTTAGGTGCCTTTTGTGATCAGAACGTTCGATAGCTCCAGAAACCGC
TTGTATTATACTTGACAG
>LLOJ004885-Intron 2:Putative_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXX

>LLOTMP005909_1/ LlAamyC1
>LLOTMP005909_1-E1 exon:PUTATIVE_protein_coding
ATGCATTTTATCAGCTTCGGTGGTGGAATTCTTTTGGTGCTGGTGCTTCATGTTGCGGATGGTCAATTTGATCCACACTTCCTC
CCTGGAAGAAGTGTTATTGTGCATCTGTTTGAGTGGAAATTCTCCGATATAGCCGCGGAGTGTGAGAATTATTTGGGCCCAAA
TGGTTTTGGGGGCGTTCAGGTGTCCCCGATAAATGAATGTCTTGTTTCCCCTGAACGAGCATGGTGGGAACGGTATCAACCCG
TATCGTATGCTATTGTAAGTCGATCAGGGGATGAAAAAGAATTTGCAGAGATGGTTAAACGTTGCTACGAGGCTGGGGTACG
TGTTTACGTTGACGTCATCTTCAATCACATGGCATCTGGGGAGGGTGAAGTTGTTGGAACTGGAGGAAGCTTGGTCTACCCAG
AGGAGCGTCTCTACCCGCACGTACCCTACGGTCCAGAAGACTTCAACCCCCATTGCGTCATCGAGGACTATCAGGATGTAGAT
CAGGTGCGAAATTGCGCTCTTGTCTCCCTTCCAGATCTCAATCAGAAGTCCGATAATGTAAAGAGGAGTGTCATTGAATTTCT
GGATCGTCTCATTGATCACGGTGTTGCGGGATTTCGGGCAGATGCATGCAAGCACATGTGGCCGGAGGATATTXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXX
>LLOTMP005909_1-E2 exon:PUTATIVE_protein_coding
AAACGAATACACCTCCATTGGCGTTGTTACGGAGTTCCTATTCTCTGCTGAGATCGGAAACATCTTCCGGTACAAGAAACTCA
AAGAATTCATGAAATGGGGTACAGATGAGAGATTCCTACGATCTGATAGAGCTCTTGTATTTGTTGAGAATCATGACAATGA
GCGTGGGCATGGAGCAGGTGGAAAGGATATTCTAACGTACAAAGATGGCAAACGCTACCTCCTGGCCGTTCTCTTTACAATT
GCCCATCCATACGGCATCCCACGCATCATGAGTTCGTACGACTTCTCGAATAGTGACGAAGGCCCACCGGCCAATGCAAAAG
GTGAAATAATCTCACCTGTTTTCAATGACAGAGGTCTCTGTACGAATGGTTGGATCTGTCAACATCGCTGGTTAGGCGTTGCA
AGTATGGTGCAGTTCCGTAATGCTGTTGCCGGTAGTGGTATCGTCAACTGGATGGACAACGGAGAACAGCAATTTGCCTTCTG
TCGTGGAGATCTGGGGTTCGTCGCCTTCAACGGCTACACCATGTCCCATCTTAATTCAACAGTTAAGGTCTGCCTACCACCTG
GCATCTACTGCGACGTAATTAGCGGGGAAGTCACACTAGATGGCTGTACTGGGCTTGAGGTTGTTGTTAAAGATGACGGTTAT
GCCAACATTTTCATTCCAGGCGACTCCCCAACAG
>LLOTMP005909_1-Intron 1:PUTATIVE_protein_coding
AAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXTTCCTACGTCTTTATCTTATGTTCCTTTG
TTTCCTAACTTCCTTTAAATTCTCACCCATCAATAG

>LLOTMP005909_2/ LlAamyC2

>LLOTMP005909_2-E1 exon: PUTATIVE_protein_coding
ATGAAGTTTTTGATTGGATTTGGAGTAGTTCTGATCCTGGTAGCTGTTGCTGCGCAGTACGATCCGCATTATCTTGACAAG AG
AACTGTAATGGTTCATTTATTTGAATGGAAATGGAACGACATTGCCAAGGAGTGTCAGGATTATCTAGGACCAAATGGGTTTG
GTGGAGTTCAG
>LLOTMP005909_2-E2 exon: PUTATIVE_protein_coding
GTTTCACCGGTTAATGAGAATTGGGTTTCAGGAGAACGTGCCTGGTTCGAACGCTATCAACCAATATCATACAAACTAACCAC
ACGATCTGGAACTGAGGACGAATTTGCCGCAATGGTAAAAACCTGCCGACAAAATGGCGTTCGAATCTTCGTTGACATTGTG
GTCAACCACATGGCTTCAGGTGCTCTTGAAGATACAATTTATGGAACAGGTGGATCTGAGGCACATCCTGGTCCTTTTGACTA
TCCAGCTGTTCCGTATGAAAAGAATGACTTTCATCCTGATTGTAGCATCTCAGACTACCAGAATGTGTATCAAGTTAGAAATT
GTCAGTTGTCTAGCTTGAGGGATCTCAATCAAACTATTCCCTATGTCAGAGAGAAGATCCTAGACTTCCTCAATCATTTAGTA
GACTTGGGAGTAGCGGGCTTCAGAATTGACGCAGCCAAGCATATGGATCCGAAGGACTTGCGATACATCTACAATCATATAA
AGAAATTAAATAGGACGCTGAGTTTCAGAGCAGGAGACAAAGCATTTATTTTCCAGGAAGTTATTGATCTAGGAGGAGAAGC
TGTATCATC
>LLOTMP005909_2-E3 exon: PUTATIVE_protein_coding
ACGTGAATATATATCGCTGGGCGTTGTAACAGAGTTCAAGGCATCCGATGATCTTGGAAAGCTTTTCCGGGGACAAGTTGCTC
TGTCGACCTTGGAAAGATGGGGTCCACAATATGGGCTTCTGCCTTCAA ACCGAGCTCTTGCCTTTGTTGAGAATCACGACAAC
GAACGTGGGCACGGAGCTGGCGGAACAAATATCCTAACGTACAAGGATGGGAAAATCTACACAATGGCCGTTGTATTCAACT
TAGCCCACTCCTATGGGGTTCCACGAATGATGAGCAGCTATGAATTTAACGATCCAAGCCAGGGACCTCCGCATGATGACAA
TAACAACATTCTAACTCCTGAATTCTCTGCAGATGGTAATTCCTGCACTAATGGTTGGGTTTGTCAGCATCGTTGGCGTCCTAT
GAGGAATATGGTGAAGTTCCGGAATATCGTCGGTCGGAAGCCAGTTAGGAAGTGGTATGACAACGG
AAGCAATCAGATAGCCTTCTCCAGAGGCAATCAGGGCTTTGTCGCCTTTAACATGGACATTGTTGACTTTAACCAACAAGTTC
CAACGGATTTACCAGACGGGGTATATTGCGACGTTATTACAGGTGAGAAGAATGGCAATGAGTGTACTGGAAAGGTAG TTAT
TGTGAGCAAGAGGAAGGCTGCTGTTATCCTAAGGGCAGATGACGACTACGGAGTTCTGGCGATTCATTCAGAGTCTAAGTTG
TAAGAATTATGCTGAAATGTTTGCTCCAACAATTCAATGACTTCGGTGCTTGGAAGATTAAGATGAAACTTTTTTGTTTGAGAT
AGTTTCAGCTAACCTATTTCTATTATTTTTTTATTTTAATAAAGAAGATGACTTGCTGTAGTA
>LLOTMP005909_2-Intron 1:PUTATIVE_protein_coding
GTAAGTAAAGAGGAAACTAACTTAAAGACTATAATTCACAAGACTGACTTTACAATGAAGACCTTCATGTTCCGCTTTCTTCA
TGTTCAGTAAAGGACTAATATTTCCTAACGAATCCTTATTTAGAGTCTCTGAACTCTGAAGATCAACTCTCTCAACAATACGG
AGTTTGAAGCCCTTTTTAGAGGATCTAGGCCCGTTGTTATGAGACAGACTGTTAAAACAAACTGCTACAACTAGGCCATCATA
ACATGCAACAAAGATTACAGAGAAAGCTTCTCTTCATTCTTCCAG
>LLOTMP005909_2-Intron2:PUTATIVE_protein_coding
GTGAGTAGAAGACTAGTAGCGGTTTAAGTATTCCCGATTTTTGTTCAATTTAATATCATTTCAG

Maltase

>LLOTMP000566/ LlMal1

>LLOTMP000566--E1 exon: PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
>LLOTMP000566-E2 exon:PUTATIVE_protein_coding
AATATAGGTATTACCCAGAGACTCCAATATGTAAAGGATCTAGGAATGAGCGGCACATGGTTGTCTCCTATCTTTAAATCGCC
AATGGCTGATTTTGGCTATGACACAGCTGACTATACGGCGATTCAACCTGAGTACGGAACTATGGAAGACTTCGAAGCACTA
ATTAAAAAAGCCAATGAAATTGGAATAAAGATTATCTTGGACTTTGTACCGAATCATTCGAGTGATCAACATGAATGGTTTAA
AAAGAGCGTTGATAGAGTACCAGGCTATGAGGATTTCTATGTGTGGCATCCTGGTAAGATTGTTAATGGACAGCGCCAGCCA
CCTAATAACTGGGTTTCTGTCTTCCGCGGAAGTGCTTGGACATGGAATGAAAAGCGCCAAGAGTACTACTTCCATGCTTTCTT
AAAAGAACAACCAGATTTGAACTACAGAAATCAAGCGACTGTGGATGCAATGAAGGAGACTCTGCGATTTTGGATGCAAAA
AGGAGTATTTGGATTTCGTATTGATGCTGTGCCATATCTTTTCGAAGTTCCTGCAGATTCAAATGGGAACATTCCTGATGAACC
CTTAACCGGGCCTAGTTGTCCAGACCCTACACATGACTGCTACACGCAGCACATTTACACTCAAAATTTGGATGAAACTTTCG
ATATGGTATACCAATGGCGGGAAGTTGTGGATGAATTTGCAAAGCAAACTGATAACGTACCAAGGATTCTAATGATAGAAGC
TTATACGCCATTGGAAAATATGAAGAGACTCTATGAAGATGGACATGGACGGGAAGGTGCTCAGCTTCCCTTCAATTTTGAGC
TGATCAGCAAACTTAACGGAAAGTCAACCGCTAAGGATTTCAAAGACGTAATTGATGGATGGATTAGTCGTCTGCCTCAAGG
AAAGGAAAACAATTGGGTCCTAGGGAATCATGATAATAAGAGAATTGCATCGAGATTTGGGATTGATCGGGCTGATCTCATA
AATATACTTCTGCAGACACTTCCCGGGATGGCGGTTACCTACAAT
>LLOTMP000566-E3 exon: PUTATIVE_protein_coding
GGTGAAGAACTAGCTTTGAAGGATGTCTATATCAGTTGGGAAGATACCATAGATCCTGCTGGTTGTAATACAAGTCCTGATGT
TTACGAGCAATATTCAAGAGACCCAGTGCGAACTCCTTTTCCCTGGAATGCCAATAAAAATGCTGGCTTCTCTAATTCTGACA
AAACTTGGCTTCCTGTCTCATCTGATTACAAGGAAGTTAATGTTGAAGCACAGGAAAAGGCAACCCGTAGTCACCTCAAGTCC
TTCCGAAAACTGACAAAGATGCATCAACATGAAAAAGCTTTCACCCAGGGGCA ACTTTTGATGAAACTTGTTGGAACGGACA
TCCTGGCGTACGAAAGGAGGGTACCTGGTACTAAACCTGAGGAAAACTTTGTGATGATCTTGAATTTTAGCAACAATACTCAT
AAAGTCAACATCCCACAACTTTTCCCGCAAATGTCACAAATGTATCGCATCGAAGTTGTTTCTATGCACAGCCTCAGTCATAA
GGAAGGG
>LLOTMP000566-E4 exon: PUTATIVE_protein_coding
GACACAATTGACGGGAAGAACTTTGAAGTAAAACCCAATGAAGCTTTCGTACTTCGTGGAGGAGCTCCAAGACTTTATTCAT
ATTTGATTTTATGTTTCACAATGTTCTCCGCATTCTTCAACATCAACTTAACCCGTGCCTGA
>LLOTMP000566-Intron 1:PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
>LLOTMP000566-Intron 2:PUTATIVE_protein_coding
GTGAGTTTTCAGACCTTTATGACCTTCTGATATTCTACCATTTACAATTTCATTTTTAAAGTTAAATCATAGTTTAAAGCTAATT
TTAATCAATTTATAATTAATTTTTCTTTTTTGTTTGCTAAACGTTACATTTTTTTTTTGGACTTTAG
>LLOTMP000566-Intron 3:PUTATIVE_protein_coding
TAATTGAAGAATTTTATACCTTAATCCTAATTTGGAATATTCTAATTTTACTTCACTCTTTCTTTCCAGG

>LLOTMP008156/LlMal2

>LLOTMP008156-E1 exon: PUTATIVE_protein_coding
GTTGTGGGAATATCTCTGGGGACTTCAAAGCTACACAATGAGAGTGATCATTTTTGTTTTTCTGGTGATTTCTGCCACTACTGA
GGGTGTCTGGTACAATAATGGGAATTTCTATCAAATATATCCCAGATCCTTTATGGACTCTAACAATGACGGTGTGGGTGATC
TGAAAG
>LLOTMP008156-E2 exon: PUTATIVE_protein_coding
GTATTACCCAGAGACTCCAATATGTAAAGGATCTAGGAATGAGCGGCACATGGTTGTCTCCTATCTTTAAATCGCCAATGGCT
GATTTTGGCTATGACACAGCTGACTATACGGCGATTCAACCTGAGTACGGAACTATGGAAGACTTCGAAGCACTAATTAAAA
AAGCCAATGAAATTGGAATAAAGATTATCTTGGACTTTGTACCGAATCATTCGAGTGATCAACATGAATGGTTTAAAAAGAG
CGTTGATAGAGTACCAGGCTATGAGGATTTCTATGTGTGGCATCCTGGTAAGATTGTTAATGGACAGCGCCAGCCACCTAATA
ACTGGGTTTCTGTCTTCCGCGGAAGTGCTTGGACATGGAATGAAAAGCGCCAAGAGTACTACTTCCATGCTTTCTTAAAAGAA
CAACCAGATTTGAACTACAGAAATCAAGCGACTGTGGATGCAATGAAGGAGACTCTGCGATTTTGGATGCAAAAAGGAGTAT
TTGGATTTCGTATTGATGCTGTGCCATATCTTTTCGAAGTTCCTGCAGATTCAAATGGGAACATTCCTGATGAACCCTTAACCG
GGCCTAGTTGTCCAGACCCTACACATGACTGCTACACGCAGCACATTTACACTCAAAATTTGGATGAAACTTTCGATATGGTA
TACCAATGGCGGGAAGTTGTGGATGAATTTGCAAAGCAAACTGATAACGTACCAAGGATTCTAATGATAGAAGCTTATACGC
CATTGGAAAATATGAAGAGACTCTATGAAGATGGACATGGACGGGAAGGTGCTCAGCTTCCCTTCAATTTTGAGCTGATCAG
CAAACTTAACGGAAAGTCAACCGCTAAGGATTTCAAAGACGTAATTGATGGATGGATTAGTCGTCTGCCTCAAGGAAAGGAA
AACAATTGGGTCCTAGGGAATCATGATAATAAGAGAATTGCATCGAGATTTGGGATTGATCGGGCTGATCTCATAAATATACT
TCTGCAGACACTTCCCGGGATGGCGGTTACCTACAAT
>LLOTMP008156-E3 exon: PUTATIVE_protein_coding
GGTGAAGAACTAGCTTTGAAGGATGTCTATATCAGTTGGGAAGATACCATAGATCCTGCTGGTTGTAATACAAGTCCTGATGT
TTACGAGCAATATTCAAGAGACCCAGTGCGAACTCCTTTTCCCTGGAATGCCAATAAAAATGCTGGCTTCTCTAATTCTGACA
AAACTTGGCTTCCTGTCTCATCTGATTACAAGGAAGTTAATGTTGAAGCACAGGAAAAGGCAACCCGTAGTCACCTCAAGTCC
TTCCGAAAACTGACAAAGATGCATCAACATGAAAAAGCTTTCACCCAGGGGCAACTTTTGATGAAACTTGTTGGAACGGACA
TCCTGGCGTACGAAAGGAGGGTACCTGGTACTAAACCTGAGGAAAACTTTGTGATGATCTTGAATTTTAGCAACAATACTCAT
AAAGTCAACATCCCACAACTTTTCCCGCAAATGTCACAAATGTATCGCATCGAAGTTGTTTCTATGCACAGCCTCAGTCATAA
GGAAGG
>LLOTMP008156-E4 exon: PUTATIVE_protein_coding
GACACAATTGACGGGAAGAACTTTGAAGTAAAACCCAATGAAGCTTTCGTACTTCGTGGAGGAGCTCCAAGACTTTATTCAT
ATTTGATTTTATGTTTCACAATGTTCTCCGCATTCTTCAACATCAACTTAACCCGTGCCTGA
>LLOTMP008156-Intron 1:PUTATIVE_protein_coding
GTAATTCAGTGCATGTGACTTTGTGTAATTTAGATGTGGTTCAGTCAATATAAAAGTCAAAAGAAGAAATTATTAATCATCAA
GTCTCCGGTTTAATTTCGCGGGGATTTGTTGCGCTTCGAGCAAATTTAAGATACAAAAAAATTATGATGATTTATAAGTATAT
ATAGAAGAAAGATAGAAGATCTATGCGCTGTTATTCTATAAAAATTAGGCCGTTTTTTCCGAATAGTGAAAACTTTTTTTTAA
GAAATAATGGGGGATTTGGGGATGGAAGTTTTTAGTTTTTTTCATATTATTTTATGGTACTCACCATAAAATACACACCAAATT
TAAATTGGATTACAACGACTTGTATCTTTTAAACAATATAG
>LLOTMP008156-Intron 2:PUTATIVE_protein_coding
GTGAGTTTTCAGACCTTTATGACCTTCTGATATTCTACCATTTACAATTTCATTTTTAAAGTTAAATCATAGTTTAAAGCTAATT
TTAATCAATTTATAATTAATTTTTCTTTTTTGTTTGCTAAACGTTACATTTTTTTTTTGGACTTTAG
>LLOTMP008156-Intron 3:PUTATIVE_protein_coding
GTAATTGAAGAATTTTATACCTTAATCCTAATTTGGAATATTCTAATTTTACTTCACTCTTTCTTTCCAG

>LLOTMP002257/ LlMal3
>LLOTMP002257-E1 exon: PUTATIVE_protein_coding
ATGAAGCTTCTTGGACTTTTATTGGTCACTGTGTTGACGGTGGAAGGGCATTTGCTGGGTGAACGGCAGTCTTCAGAGAAAGC
CGATTTACGGCATGATCATCCAGAATTGGATTGGTATGAACGAGCAAATCTCTACCAAATCTACCCGAGATCCTTCATGGATA
GCGATGGGGATGGAATTGGAGATTTAAATGGCATCACAGCGAAATTGGAGCATCTTCAGGAGTCTGGGATCTTTGCAACATG
GCTCAGTCCAATTTTCAGCTCCCCCATGCGAGATTTTGGCTATGATATTTCAAATTTTACAGAAATTAATTCCGACTATGGTAC
CATGGAAGACTTCCTGCAACTTCTGCATGAAGCACGAAGGCTCAATATTCGTCTCATCTTGGACTTTGTGCCCAATCACACGA
GTGATCTGCATGAATGGCTTGTAAAGTCCGAGGACAATGATCCTGAATACCGAGACTACTACGTATGGAGGGATGCACGGTA
TGTGGGGGAAGAGAGACACCCACCAAATAATTGGCTTTCCGTCTTCCATGGACCTGCGTGGACGTGGAGTGAAAAACGGCAA
CAGTACTACCTTCATCAGTTTGCCAAAGAGCAGCCAGATTTGAATTTTGAAAATGAGAAAGTGGTGGAAGAGATGACGAATG
TATTGAAATTTTGGCTGGATATTGGTGTCGACGGCTTCCGTGTGGATGCCATCAATCACCTCTTTGAGGACCCCAGTTTCGCCG
ATAATCCACCTAATCGTGAGATTTATGATCCAATGGACTATAGCCACTTTGACACAATTTACAGTAAAGATCTTCCAAAGTCC
TATCAGCAGGTGTATGATTGGCGCAAATATCTCGATAGCTATGCTGAAGAGCACGGCACGGATACC
>LLOTMP002257-E2 exon: PUTATIVE_protein_coding
AGGGTCCTCCTGACTGAGGCATACACCAACATCACGAATACGATGAAGTGGCAAATGTCAGAAGATGGTACGCAGAAGGGT
GCACATTTCTCATTTAATTTTAACTTAATTATGGAGCTCAAGACTCTCCCCGATGAGCTCAATGCGGCAGATGTGAAGGCAAC
CATTGATGAGTGGCTTGAACACCTCCCAGATGGAGTTCCACCAAATTGGGTG
>LLOTMP002257-E3 exon: PUTATIVE_protein_coding
CTTGGCAATCACGACCGCCCACGCGTTGCATCGCGTTTTGGTGTGGATTTGGTGGACAATATGAATATTCTCGTGCAAACCCT
CCCTGGAGTTGCGGTGACATACTACGGTGAGGAGATTGGAATGGAGGATTTCCGTGGAATCTCCTGGGAGGATACACAGGAT
CCTCAGGCGTGTGGTTCCAATGAAACAGTCTACCAATTGTATACGAGGGACCCTGTACGAACCCCAATGCAGTGGGATGATA
CCCAAAATGCGGGCTTCTCCACAAATCCCTCCCCGTGGCTACCTGTACATCCAGACTATGTGCAAAATAACCTGGCACAGCAG
AAGGTGGCTGAAACGAGTACATATAAATTGTTCCTGGATTTAGTAAAGCTCCGCATTGATCATGCCCTAGAGTATGGAGATTT
TAAGTCAGAAGCTCTAAGCAATGGAGTCTATGCCTACAGTCGGAATATGGAGGGACATGAAAGTCTCGTGGTGGCACTCAAT
TTCGCCAGTACAGAGGCAACAGTTGACCTTACTCCCGTCCTTCTTGAGAATTCACATTCAGAAGGTCTAGTCGAATTGGCTAC
AACCAAGTCCATTCATAAGAGA
>LLOTMP002257-E4 exon: PUTATIVE_protein_coding
AGTGATGTCGTTAACTTGAGGACCCTCACACTGCAGCAGTACGACGGTGTCATTCTACGAGTTAACTCATCCGCTGCAACAAT
CTTCATTTCCTCATTTGCTCTTCTTCTTGTTGTACTGAAAACACTCCTTTACTAA
>LLOTMP002257-Intron 1:PUTATIVE_protein_coding
AAGTAAGCTCGTTGAACTTTCTCAAGTACTCATTTTGCTATATATTAGTTGGTATACCACATATGGGCCTATTTCTTACAAGTC
TGATTAACGAAAGATTGAAAAAAAAAGAATATTTGTTGGATATAAATTTGGTTGATGAATTTTTATTAAAAAAATTTAGCAAA
TTAACTACAAAATTAGTCAAAAAATTATTACTGACTGATTAATTGTTTTTTTCTTCTTTTAGTAAGTTAAAGCTTAACATAAAA
CTTTTCAAATTCAATGCTCTTTCTTGGGTATTACTTAATTAATTTTATTTAAAATAGATTCAAATTAAAAGTCTTCTTCAGCTCT
AAAACTCTCAAGTACTTTGGGAATTTTTAAGTCCATTCTTTGGGGCTCAAAAGTACATAAAAATTATTTTAGAAAATGACGAA
TTTCCCAACAATTTTGTGATTGTTCCAAAAAATATCAATTCTACAATCTACATCCTCTCATACAAAAGATAAATCAAATCACA

ACGCATAGGAAGCTCCGCGTTGTGGTACACCAACTCACATAAGTCGTGGACATATTATCAGTGATTTGTTCAATTGCTTTCAA
TATT
>LLOTMP002257-Intron 2:PUTATIVE_protein_coding
GTATGCTCATGATTGAAATTCAAACAAAATCATTTTAATTAAATTTAAATTCAATCCCAATATCTCTCAAATTTTCCACATAG
>LLOTMP002257-Intron 3:PUTATIVE_protein_coding
GAGTAAGTTTTTTTTAAAAAGAAAATTAATCTCATTAACTCATTGAAATAATTTTCTTTTGCTAATAATTAACTTTTTTTCTAAT
TTTT

Amino acid transport protein
>LLOTMP006803/ LlCD98hc
>LLOTMP006803-E1 exon: PUTATIVE_protein_coding
ATGGTGCGCGAGAAGGTTGACATGTCTGAGGAGGGAGCCGATGAGAAGATGCTTGGCGCCGGTGAGGAGGAACAGCAGAAA
CTGGCAAAGAAGGAGGAGGTAAAATTCATCAGTGGCGATCAGCAGAATGGCGATGCAAAGATCGATATTGGAGCCATTGAC
AAGACATTCACGGGATTGACAAAGGAGGAACTCATGAAGTACGCCAATGATCCCTTCTGGATTCGCTTGCGTTGGTTCCTCTT
CATTTGCTTCTGGGGCCTCTGGGCTGGGATGCTTGTTGGAGCTGTTCTCATTATTATTTATGCACCAAAATGTGCAGCTCCTGA
GCCACTGTCATGGTGGAAGCAGGGACCACTGGTGAAGATTGATACACCCACACCAGATAATATTCAAATTGAAGGCATTAAG
GGACTAGGCGTGAAGGGTGTCATCTACGAAGTTCCCGCTGATGAGACGTACAATATTGGAAGTGATCCTGCTATTTCTGATGC
ACTGAAGAAGTTGGCGGCATCTTTCGCTGCCTATGGCATTAATTTAATTGTTGATCTCACTCCGAATTTCGTTTCGCGTGGAGA
TCCATTGTTTCAGAAAGCTCTCAATAATCCTGATGTAAGAAGTGCATTTGTTACACGCGAAGGACACCAAGTTCCCAATAATT
GGCTCTCATTGGTCAATGGAAGTGCATGGACAAATGAAGGGAATCTCATCTTTCTCCAGCAATTTGGGCATGATCGCTTTGAT
CTTCAACTCAATGATCCCGTTGCTCTGACAAAACTAAAGGGAGTCCTCAAGGAGGTGGTTTCTCTCGGAGCTAAAGGCATTCG
TCTGGCCAATACGAAGCACTTTATTGTTGATCGAGAACTTAAGGACAATGTACCATCGAATAATCCAAACTACGATCATACCC
AGTATGGTTTCTGGACGCACGCTCATACCACATATCAAGATGGATTGGGTGATCTCCTGTATGATTTAAAGGCGTACGTTCAC
AATATTACAAACTCCGAAGGTTTTCTCGCAGCATCCGATGACATTGATCGTCCGGAAGTTTATATGTCGTCCAGAGGAGTACA
ATCTCTGGATATGATTCAGCTTGGACGAGTAACCAATCTTCTGGCCAATAGTACCAGTGCTGATCATATCTACAGGGAGCTAC
AGAATGTGTGGAAGAGTACAGTTGAAACAGGGAAGCATTTCTGGCTTCAATGGAACTATGTGAAGCAGGAACTTCTGCAACA
AATCACAATTTCCGAATACAACATCTTCCTCCAGCTCATCCCTGGTGTTCCAGTGTTCAATGTAGATGCTTCCAAAGCAAAGG
CGGAGGACTACTCTACGGAGGCATTCCACAAAATCCGCACTTCACCATCCTACATGCACGGGAGCTTTGATCTCTATCGAGAT
GTCAACAGCACCGTTATTGGCTATTCCAG
>LLOTMP006803-E2 exon: PUTATIVE_protein_coding
GACAAAGTCCGGAAATCCAGGCTTCTTTGTAGTGTTCAATCCCACTGAGCAGTACGTGAATGCCAACTTTTCCAATGTTGTTG
GGATAGCTGAGGAATTGACTGTGCACACTCTCAGTACCAATTACAATGCCACAGATGTGGCTGTAAA
>LLOTMP006803-E3 exon: PUTATIVE_protein_coding
GGGAAAGGTGTCCAGTTCATCAATTCCTGTTTCCAGATACTCTGCCATAATCCTAACATACGTCCCAAAAGCTTAA
>LLOTMP006803-Intron 1:PUTATIVE_protein_coding
GTAAAACAGAAATCCCGCTTTATATAGTTGACTCATATTTTCTGCTGCTAACAAGACTAATCAAACTAATCATTTCACATGTTC
TTTCCTCTCTGTGGCATTAACCATTTCTCACTAAAACATGGCTAACAGAATTTGGTAAATGATTTCACTAACACTTTTTTTTTAA
CCATCTGGCAATGAAAGTTCTTACTCTATTGTTGTTTTATACTTTTAG
>LLOTMP006803-Intron 2:PUTATIVE_protein_coding
GTGAGTTTCTCTTCTTACTTCTCAAAAAATGATCCATTTTAACATATTTTTTTTTAAAATTTCTTTCAG

>LLOTMP008629/ LlNBAT1
>LLOTMP008629-E1 exon:PUTATIVE_protein_coding
ATGAATAATTTGGGTATTTCGACGGATCCTGCTCTCCTTGGGGTGGATTTGCCCTCATCACTCACGACCTCACCGTCTGTGTCA
ACCTTCATGCCCGAAGAGGATGCATCCATATGTCCACTCCTCCCAACAACACCCAGCCCACCGCCAATGGATTTTATACATCC
CCTCACGCCAAGTACGGGGATTGATGAAGGAAATAATGAAGATCTCGGTGGTGAGTTTCAAG
>LLOTMP008629-E2 exon:PUTATIVE_protein_coding
GTGATGATGCTCAAGCAGATTCAAGCTCTTCAGGAAGTAGCTCTGGCATTGGGATTCAAGTCACTGCCAATGGGCCAGGATCT
CTCTTCAATAAGCACGCATATCAGCATTTGGGCAGCAAAAATGGAGACATTACGCAA
>LLOTMP008629-E3 exon:PUTATIVE_protein_coding
GACAATGGTGTTACCCAAACAGGAGTCTTCAGCATCCCCATGACAAAGGATACACCGTCTTTTGTGAACTGGAACTGGCCTCT
CATCCGCAAATGTACCTTCTTCCTCTTCATGTCCGGTCTCTTTGCAATGTGTGCCATTGTTGTTGCCATGATGTTCAAATTACCC
AAATCCTGCAATCCCGAAGTACCCTGGTACAAGGGAGCTGTCTTCTATGAGATCTTCCCGGCGAGTTTTCAGGATTCCAATGG
GGACGGTTTGGGTGATTTAAAGGGCTTGGCGAGTCGCATTGATTACCTCGAGAGTTTGGGCGTTGGAGCTGTACGGTTGAATT
CAATATTCCCCGCCAAGCACTATCCGGATCACTTCCAAAACGTCACAACTCTCCTGGACATTGATGAAATTCTGGGAAATCCG
AGGGATTTAACGTATTTGGCGCGTATACTGCAAAGACGGAATATTTCATTGATCCTGGACCTTCCCATTTATCCTCTTGTTAAG
CATCTCGCTGTAATTCGTGAGAATTTCACGGAAGCAACAAACTCAACTGACACAGAAGATGAGGAAGTGTTCCAGGATAGTG
ACACTTCGGAAGATCCTGTACTTACAGCTATGCGATTTTGGCTCTCAATGGGTGTGGATGGATTCTATGTGAAAGGACTCGAG
AACTACGCAGGTGATCCGTATTTAATTGAGAATCTTCAAGAGTGGAAATTCGCTTTGGGCTCAGATCGGATTCTCATGGTTAG
TAAACTCGTCTTTGACTCTGTGGATGATGACACGGCGGAGAAGGTGAGGCATTGCGTGGATTTGGTGGATGTTTTCGTGGATG
TCTCCAATGGGACAAAAGTCATTGCTGAGAAGGTTCAATCAATCCTCAATTCTCGCTTCCTGGCACCTGGTTTTGGTCCATGG
ATTCACTGGAGTTTGGGTGGTGTAACCGAGAGAAGATTGGCCCAAGGAACAAGCTCAAATATTTCTCTAGCAGCCACACTCA
TGCAGCTCATGCTTCCAGGAACTCCGAGCATTTTCTACGGAGACGAAGTTGCTCTGCAGGAAGTTTATGATCCTCTCGGGGAG
CACGAAGAGTCCAAGCATCTGCATCATCTCTCAACAATGGTCTGGGATTCGGAGATTCAGTTCACGGTGAAGGAGCATTTGCC
TTGGCTACCTCGGGGAGCCTTTGTGGCTTTCCATCATTTTGAACATGTGGCAGGGATGATAAGACTTCGTGGAATTTCTCCTTC
CATCTACCAGAATGCCGTCATAAAGGAACATCAATCGGTTCTCAATACTGCTGTGAG
>LLOTMP008629-E4 exon:PUTATIVE_protein_coding
ATACAGCAAAAATGACATTTTGATCCTCGAAAGATGGTACCCCAGACGAAATACCTTCACATCAATCTCAAATTTTGGAAGTA
AATCCCTCGTCCTTGATCTCTCTGGGATGTTCTATGCGGGAGAAATAATGCTGGGATCCCTCAAGGGGGAAAAGGTCTTCTTC
AGCAGTATCCAAGTGAAGCCAATGGAGACGATTATTGTAAAGCTGGACAAGTAA
>LLOTMP008629-Intron 1:PUTATIVE_protein_coding
GTGAGTTTGTCTTGATTTTTAACATTTTCTTAACTTTTTAATTAAATTCTCTTTTTTCCATGAAG
>LLOTMP008629-Intron 2:PUTATIVE_protein_coding
GTACGTCTTATGCTAATACATACGCACATAAAATCAATTTTATCAGACCTTTGATCAAAATTTGATGAGAAATGGAAATGTTT
TAGAATTGAATTTTGTAGGAAAGATTTTCTTTGTAAGATTTTTTTTTAAATGACTCCGCGAACAAGATTTTCAAAATTTTGATT
AAAAAATCATAATTTCATAGCAAAGACAAATTTTATTTTAACTATAAAATAGTCAGGACACATGTTGTGCTATAAAAATTCTG
ATATTTTGACTAAATATCTAATGAAATTAAAATAAAAAGGATTTAAAGTACAAATCAAAAGGACGTGTTAGTTAAGGATCTTT

TTCAATATTTTCAAGACCTAATAAACTAATCTAGACAATTACCCAAAGAATATTTCTATAAAAGATAATTAATTAATTAATTA
AATT
TTAATTTTAG
>LLOTMP008629-Intron 3:PUTATIVE_protein_coding
GTAATTATTTTAAAGGATATTTTTGAAGAATAATCAAAGTATTAGAAATAATTTTAAAGAATCTATTCGAAGGATTCGGTATT
ATATAAAATGTTTAAAATATTTTAGAAGAATCTTCAAAACATTCGGAATATTTAAAGTATTCGGAATATATTTGAAGAATATT
CAAGGGATTCGAAATATTTAAAGAATATTCAAAGTATTCAGAATATTTAAGAAGAATATTTTCATTAAAGAGAATATTTATGA
AGAATATTTCCGAAAATCAAATTATTTTAAAGAATATTTTCGCAATATTTTATATAAAACATAATTTTGACCCAACCTGCAGG
AAATTGAAGTTTGTACACTTTTCCTTCCCGGAACAGTTGAATGAGCTTCCCGGAAAAAGATTTAAGTGGCTGGAAATTGTGTA
GTTATCCTGCCCATTGCACCAGAATATTTTTAAGGGATCTTAGGAAACTCCTTTTTAACACCCGGCTGATCCTTGATGTCATCT
TTGGGTCAAAATCTGTCCATTTAGGAAGTCCCTTTAGCAATTTCGTGAAGCATTTTTTTAAACTCAGACTTTGACCGAAAGATA
ACATCAAGAATCAGCCGGGTGTTGAAAAGGAAGTTCCTAAGATCCCTTGAAACTTCCTGATATAATGGGCAGGAAGACTACT
CAACCTCTAGGCACTTACGGCACTTTCCAAGAACATATCAAACCGTTCCGGGAACAAAAAAGTCTATAAACTTTAATTTGCTC
CAACCTGCTTCAATTTCTAATCCATTTTTTTAAAGAATTTAAAGAGAAAAATTAAAATAAAATTCTTAAAGCTTCGTTTAAAA
AAAAAACAAATAAAACACATTAAATGTTTTTTTTTCTAAATTCCATTTCAG

1,4-alpha-glucan-branching enzyme
>LLOTMP005533/ LlAGB1
>LLOTMP005533-E1 exon:PUTATIVE_protein_coding
ATGGATCCCATGAAGGTGGAAGTGCCGGAGATTGACAAATTGTTTGCCATCGATGGTTACCTGAGGCCACAAGAGAGGGAAT
TGCGAAGAAGG
>LLOTMP005533-E2 exon:PUTATIVE_protein_coding
GCATGGCGTGATCAAGGATTGGCTAAAGAAGATAGATGGGAATGAGGATGGGGGTGTGGATGGATTTTCACAGGCCTACAA
ATACTATGGGATTCACATTCAGCCCGATAATGCGGTGATTGCCCGCGAATGGGCCCCTGGGGCGCAACAACTCTACCTCACCG
GTGATTTCA
>LLOTMP005533-E3 exon:PUTATIVE_protein_coding
GATGATTGGCAGTGGGAAGCCACACCGTACACAAAGCTCGAATTTGGGAAGTGGGAATTGAAGATCCCACCACGAGAGGAT
GGCTCGTGTGCCATTCGGCACCTTTCGGAGATTAAGGTGATCGTTCGCACGCAAGAGGGCACCCTCGTTGATCGCCTAAGTCC
TTGGGCAAAGTACGTCCTCCAACCACCGAAGGAGGCTAATCAGGGTACGAACTTCAAGCAATACCTTTGGCATCCGCCGCCA
CATGAGAAGTACATGTTCCGCTATGGACGCCCCCAGAAGCCCCAATCGCTGCGCATCTACGAATGTCACGTTGGCATTGCCAC
GGAGGAGCTTGGTGTGGGGAAGTATCGGGATTTTGGGGATAAAATCATTCCGCGAATTGTCAAGCAGGGCTACAATGCCATT
CAGGTGATGGCAATTATGGAGCATGCGTACTACGCTAGCTTTGGGTACCAAGTGACGAGCTTCTATGCTGCTTCCAGTCGCTA
CGGGACACCTGATGAGCTGAAATACATGATTGATATGGCCCACAAGGCTGGTCTCTACGTTCTCCTTGACGTTGTGCATTCCC
ACGCGAGTAAGAATGTTGCCGATGGGCTCAACCGCTTTGACGGCACCAATACGGGCTTCTTCCATGATGGCGCCCGCGGGGA
GCATCCCCTCTGGGACAGTAGACTCTTCAACTACACCGAATACGAAGTTTTGCGTTTCTTGGTGTCCAATCTCCGGTGGTGGC
ATGACGAGTACAACTTTGATGGGTACAGATTTGACGGCGTTACGTCAATGCTGTACCATTCACGTGGTGTGGGGGAGGGTTTC
AGTGGGGACTACAATGAGTACTTTGGCCTGAATGTGGATACAGATGCTCTCGTTTATTTGGCCATTGCCAATGAGATTCTCCA
CAAACTTGACAAGGATATCATTACAATTGCTGAG
>LLOTMP005533-E4 exon:PUTATIVE_protein_coding
GACGTGTCTGGGATGCCAACAATGTGTCGTCCTGTCTCCGAAGGTGGCATTGGGTTTGACTATCGCCTGGGTATGGCCATTCC
TGACAAGTGGATTGAGTATCTGAAGGAGAAACGTGATGACGACTGGAATATG
>LLOTMP005533-E5 exon:PUTATIVE_protein_coding
GGTAACATCGTTCATACCCTAACTAATAGACGATGGATGGAGAAAACCGTTGCTTATGCTGAATCTCACGATCAGGCACTCGT
TGGGGATAAAACTTTGGCTTTCTGGCTGATGGATAAGGAGATGTACACGCACATGTCGACCTTGTCGGAGCCTCCGTTGATCA
TTGATCGTGGCATTGCTCTGCACAAAATGATCCGTTTGATCACGCACTCCCTCGGCGGGGAGGCGTATTTGAATTTTATGGGA
AATGAATTTGGTCATCCGGAATGGTTGGACTTCCCGCGTGTTGGCAACAATGATTCCTACCACTACGCTCGACGTCAGTGGCA
TCTGGTTGATGATGAAATGCTCAAGTACAAGTACCTCAATGAATTCGATCGTGCCATGAATGGGTTGGAGATGAAGTACGGCT
GGTTGGCATGTGATCCGGCCTACGTGAGTTGCAAGCATGAGGATGACAAGGTGATTGTCTTCGAGAGAGCTGGACTCCTGTTT
GCCTTCAATTTCCATCCATCGAAGAGTTTTACGGACTACAGGTTGGGCGTTGAAGTTGGTGGCTTGTACAGGATTGTCCTGAG
TACGGATGATCCCCTCTTTGGTGGTATGAATCGTGTTGATGTGAAATGTGATCATCTCTCAACGCCCGAAGGGTTATGCTGGA
AGACGGAATTTCATTCAAGCCTACCTTCCAGCACGAACTGCCTTCGTATTCGCCAAGTTTGATTAG
>LLOTMP005533-Intron 1:PUTATIVE_protein_coding
GTGAGCGTAACCGGAAGTCTGTTCTGTTTTCCCCAAAAAAGGATGAAGAATCTCCACTGATTCCACATTCTGTGAGAGATGCT
TCAATTTTTTTCCGTTGCGGGCATTTCCCAATGTCACGGCAAGGTCAATTTTTGGCTGCTATCGCAAAATGAATTTATATCAAT
AAAAAAAAATAGAAAATTGATGTTTTTTCTCTTATTGCCAG
>LLOTMP005533-Intron 2:PUTATIVE_protein_coding
GTGAGTTCCTTTTAGTCGCTCTTAGGTGAATTCTTTATGTGAATTTTGGTTATTTAAAGAATTTGTGTAAGAAAACGGTTAGAC
ATAAATCCAATAAGACTAAATATAGGGGAGAATGGCCTAAATTCGACAATGAGATTTTTTTTTTAATTTAATAATTTGTATAAT
GTAAAATAAACTGCACAAAAAGTTGTGAATCGTAGGGTAATGATAAACGATAAGGGCAAATAAAAAGGCGATAAGTCAATT
ATAAATGATTTAATATAATAACTTAAAATTTTGAACAATTATCCAAAAAAAAAAAGAACTTAAGGTGATTTATTTGCATATTCG
CGCATAAAGGCTTTATATGCTGCATACTTCTCCTTCACTTCCTCCACTTCCCGCTGCAGCTCGAAATTTCTCGCTGGAGGAGGA
GATTCTCCTCCTGAAGACGTGCGTTGATCTCCAGAAACTTATTGTTGCTGTTCTTTATTAAAAAATTAAATTAAAAAATTGATA
CGTCATCTAACAGAGCCAGCAGAGCCTTATTGCACTCCTTCGCCTGGTTGATGGCTTGGTCGAACTCCTTTTGCACCTCTTTTG
GATCCATCTTTCCTCCTTTTTTTGACTAATCAAAATTTTCTTCCTTACAACGGGCCTAATTCAGCGACCAAGAAATCCTTGAAA
TCATGAAATTGCTGTACAAAACTCAAAGATTTCAAGGATTTCTTGGTCGCTGAATAAGGCCCGGATATATCCTTCCTTTTGACT
GTGATTTTTTTTAATTAAATTTTTACTCTGTACTATTTACTGTCATTATAAATGCCAGATACCTGGATGTTATTTATGTAGCCGT
CCAGGAGGTTTTGGATTGTCTGCATCGACAACCGCCTTTTTGTCCTCTGTCACCTCGTCGAAAAAAGCTTCTAGCGCTTCACGA
TTCATGTTTTGCTCTATATCTTTGAAACCGCCACATAAATTTTTAAGATTTTTTTTTTGGCTAAAAGGTGTGTCTGTCAGTTAAG
ATATACTAAAATTTAATCAAAATGTATCGTCAAGTTTTTGAGATATTAATCGAAAAATTGTCGAAAACTCATCGAAAATTTTG
TTTTTGATTGTAGGCCCCCTAGCGGTAACTTTTAAAACTTCCTATGTACAGATAATTGTAGACCTTCTTGAGATCTTCCACTTCT
TAATGTTTTTCCCCACTCTTGATGTTATTCAACACCTTTTGATCTCATTCACCACTTTTTGATTTCTACACACCACATTTCCATTC
AATTAATCTCAAGGCGAAAAATCATAAAATTATGCAAATAGCTTAATAAAAAAGGGTCTAAAATTCGTTCAAATGATTTAAA
TAATGTAAAAATTCCACCTCCA
>LLOTMP005533-Intron 3:PUTATIVE_protein_coding
GTGAGTTTGAGAATCTCATTTGCAGCCCTTCGTTGACTCTCCTTTGTTCCTCAG
>LLOTMP005533-Intron 4:PUTATIVE_protein_coding

GGTAAGTTCCGCAACCTAAACGATTTTTTGTTGTTGTATTCGTTAATTTTTAAATCTTATTCACCACTTCTTGATCTTTTTCACC
GCTTTTTGACCTTCTTCACCACTTTTTGACCTTTTTCACCACGTTTTGATCAATTCACTACATGTTCAATTTATTGTCGATATTCT
TGTCTTGATTTTATTCACAGGCTGTGCGTTAATTATTCAATCTTATTCACCATATTTTAATCTTAATTCACCACGTTTTGATTTA
TTTACCCCTTTTTTAATTTTTTATACTCTTAAATATTCTTGTCTTGACTTTATTCACCACTGTTCGATTCTAGGATAAAAACTAG
ACAGTCTTAAATATCTTGTTTTCTTAATATTAACCTTTGAATGCGGAGGCCTTGGTAGTTACGTAGAATGCGAAGGTGGGGTC
GTTGAACGACCCCAAAAAGAAGGCCAATTTTCTCCCAAACTATACAACCTGGAGTTGTCGGGTTAGTGCCTATAGATTCCTTA
TATAGTCCTCTTGCACCCTGCACCACAAATTCGCCGCCCTAAAATACTTAGAAGGAGATATTAGGGAAAAACATTXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXAAAACTAAGAGAACTTAAGATGAAATCTTTCTTAACT
TTTAAGAGAACTTAAGAGGAAGTCTTTCTTAACTTTTAAGAGAACTGAAGATAAAATTTTTCCTAACTTTTAAGAGAGCTTAA
GATGAAGTCTTTGTTAAGTTTTAAGAGAACTTAAGAGGAAGTCTTTCTTAATTTTTAAGAGAGCTTAGAATGAAGTCTTTCTTT
ACTTTAAAGAGAACTAAAGATACAATTTTTCTTAACTTTTAAGAGAGCTTAAGATGAAATCTTTCTTAACTTTTAAGAGGACTT
AAGATGAAGTCTTTCTTAAGTTTTAAGACAGAGGAATTATCAGCAATGATAGCATCCCCTTTAATAAATTATTAAAACTTATG
ATTCTCTTAATTTTCTTAAGTGAAGCTTAAAATCTTTAATTTTTACAAAATTACTTTAATTTTCTTAAGCAAAAGTTACGAATTT
TAAGATTTCATGAGATTTTTAAAGTTTTTTTTAATTATCTCAAAATTTCTTAAGTTTTTTTAGGCGAAAATTAAGAGTTCTTAAG
ATTTTAAATGTTTCTAAAGATTCTTTTTAAAATTTTCTTTGGAAAAGCTATGAATTTTTAAAAAGTATTCGCGAGAGTTAAGAT
TTATTAAGTCAGGCTGAATCCGAAGATTTTTACGATGAATATTAATTTAACCCAATTCAGTCTGACTAAAGAAATCTTAAATTT
TTCTTAAGAAAACTTAGGAAATCTTTAAAGAATCTTTAAAGAAAACTTAAGAAATATTTTCTAAGTTTTAAGAGAACTTTGAA
CTTTAGAAAACTTAAGAACTTTACTTAATATTCCTTAAATTTTCTTAAGCGAACTTTCTTCAATACTCTTCAGCGAAAGTTAAG
AGTTCTAAGGATTTTTTTTTTAATAATTGTTAATAATTTATAATTGTTTATTCACATTCCATTAAGATTTTACATACTTTATACAA
ATTTTCTCTACAAATTCAATTTGCGATATTTTCGTAAATTTCTCAAAATTTCTTAAGTTTTCTTATAGCAAAAGTAACAAATTTT
AAGAATTTTAAGAATTTTAAGATTAAAAAAAATTCTTAAGAGAAAATTAAGAGTTCTTAAGATTTTTAAAGATTATATCAACC
CTAATCTGATTTTTTTTCTTCTA

Glycogen debranching enzyme
>LLOJ008312/ LlGlyD1
>LLOJ008312-E1 exon:PUTATIVE_protein_coding
ATGGGGCATCAGGATGTATTGAGTTTGGCCATTTCACACGGTGAGGATAAGGAAGGAACGCTCTTCAGGCTCAAAAGGGGTA
CAATTCTGCACATTGTTCCTGGACCTAATCTCCTTGGCAGGCACATCGCATTGTACTGCAACTATCCACAAAATGGTGAGAAA
>LLOJ008312-E2 exon:PUTATIVE_protein_coding
GGTTTTGAAAGAACAAAATATTCAACTCTAATTTGGCACACAAGTTCAGGTCAAAAGCTCACAAATGCGGCTCAACCCTTTGT
CGAAGTTAATGATCTAGATATTTACTGCGAAATAGTGGCTAACAAATCAGGAACTTTTCATTTCTACTTTAGCTTCCAAGAAA
G
>LLOJ008312-E3 exon:PUTATIVE_protein_coding
TCCAAAAGAACGAGAAGGTTCCCTTTACGTTCAAGTTGAACCAAAGATTAAAGTCGGACCGAAGAATGCAGAACGTGAGATT
CCCCTTGACTCCATACGATGCCAGACTGTGATGGCAAAATGTCTGGGGCATTTGGAGACATGGGAAGCTAAGTTGAGGGTTA
CTCATGAGAGTGGCTACAACCTCCTTCACTTCACACCCATTCAGAAGCTAGGAAACTCCCGTTCGGGATATTCTCTTTGCGATC
AACTCAGCGTCAATCCAGACTTCGGTCCTAGTGCAAATTTTGATAAAGTCGCAAAAATCGTTAAAAAATGTCGCGAAGAATG
GGGCATTGCCTCTATCTGTGACATTGTCCTCAATCACACAGCCAATGAATCTCCCTGGCTAAAGGAACATCCAGAGACAACCT
ACTCATGCAGTACTTGTCCACACCTTAGACCTGCTTTCCTGCTCGACAGCGTCCTTGCAATGGCGTCGAGTGATACAGGAAAT
GGTCTTCTGGAAACTTTCGGTGTACCTGTTGAAATTGATCGTGAAGATCATATCAACGCATTGCGTCATCAACTGCATTCAAA
CTACCTTCCCAAGGCTAGAATACATGAACTGTATCAGTGCAATGTGGAGGAGTATGTTAAGAAATTCAGTGAAGAAATGCGC
AAGAGACCTCCGCCAAAAGCTCCAAAGGAATCCCAATCAAAGGAGACAATCGTCTTGAAGCAGGATCCTGCCTATAGACGCC
TTGCTTGTACAATTGATTTTGAGAGTGCTTTTGAGATCTTTAATGTTTTTAG
>LLOJ008312-E4 exon:PUTATIVE_protein_coding
GAATGACGCCTTTGATGAAGATACAAGATTGAGGAAATGCGCTGAAGCCTTTAGGCAACACTTGGAGGAGTTAAATGAAGTT
GTTCGTTGTGAAATAGCTGAACATCTTCGTTCAGCCGTTGAAAATTGTTTATCCGGTCTCCGATATGAAAGGATTCAGGATGA
TGGTCCTAGACTACGGGAGATCAGCATCAAACATCCACTGTTCTGTAGATACTTCACAGATGGTGATGTTACGTGTAAAACAA
TTGAGGATATTGAAGCTCTCATGTACGGGGATTCGGCAAAGTATCTTATGGCTCACAATGGATGGGTTATGAATGGGGATCCC
CTGAAGGACTTCGCTAGGCCCCAGCCAACAACTGCGAATGTCTACATTCGACGGGAGTTGATTGCTTGGGGTGATAGCGTTAA
GCTCAAGTATGGGGATAAACCCGAAGATTGCCCATACCTGTGGGATCACATGAAGAAGTACGTTGA
GATAACTGCTAAAATTTTTGATGGAGTTCGTCTTGATAACTGCCATTCAACTCCGCTGCAT
>LLOJ008312-E5 exon:PUTATIVE_protein_coding
GTTGCAGAGTATCTTCTGGATGCAGCACGAAAGATTAATCCTGAACTCTACGTTGTTGCGGAACTCTTTACAAATTCCGACCA
TACGGATAATATTTTTGTCAATAGATTAGGGATTACATCGCTGATCAGGGAGGCTCTTTCCGCATGGGATTCCCATGAAGAAG
GGAGACTTGTTTATCGTTACGGAGGAGCACCTGTTGGGGCATTTTTCAGCTCTCCCAAGAGACTCCTGGCGCCGTGCATTGCT
CATGCTCTCTTTCTCGATCTTACACACGACAATCCGTCCCCGATTCAGAAACGATCGGTCTTTGATCTGCTCCCCTCAGCAGGA
TTGGTTTCAATGGCATGCTGTGCAACTGGAAGTACAAGAGGATATGACGAACTCGTCCCTCATCAT
>LLOJ008312-E6 exon:PUTATIVE_protein_coding
ATTCACGTCGTTGACGAAGAAAGGGAGTATCAGGAGTGGGGTAAGGCTGTTGATGCGGAAACTGGAATAATTTCTGGCAAGA
AAGCAATTAATCTTCTTCATGGAGAACTAGCTGAGAAGGGTTTCAGTGAGGTATTTGTTGATCAAATGGACTACAACGTTGTT
GCTGTTACCAGGAGTTGTCCTACTACGAGAGAAAGTGTCATCCTGGTTGCACACACCTGCTTCAGTTATCCAGATCCATACAG
TGGACCTACAAATGTTAGATCTCTTAGATTTGAAGGGCATCTTGAGGAGATTATCCTTGAAGCAGGAGTTACTCACAA
>LLOJ008312-E7 exon:PUTATIVE_protein_coding
ATCTTCAAAGCCATATGACTGTCCATATAGGTATGAGAAAGATGAGAAATATATCAATGGACTC ACAGAGTATCAGTTGAAG
ATTGTTGAGCATATTCCTCTGGAGAAATCAACAATTTTCAGCACAGAAATGGTTAAAGATGGCAACATAACGCAATTGAATTT
TAAGAATTTAACTCCCGGATCAATTGTTGCTGTTCG
>LLOJ008312-E8 exon:PUTATIVE_protein_coding
AGTTTCCCTGCATGAGCGAACTCGTCCCCATTTTGCCAATGTTCAGAAACTCGTGGATGCTTTTCACTTCCAAA AAGGCCAGA
TCTTCAGTGATCTCCAGAAAATTGTGTCCAAACTCAATCTCATCGATCTCAATCGTGCTCTCTACCGTTGCGATGAAGAGGAG
AAGGATATGGGAATGGGGTCTGGAACTTATGACATTCCCGGATTTGGGCGGTTGGTTTACAGTGGAACTCAGGGATTTGCATC
AGCTTTGAGTGTTATTGGGCCTAATAATGATTTGGGGCATCCATTCTGCGATAATCTCCGTCGTGGGGATTGGATGATTGA CT
ACATACGCGATAGATTAGCTCGTTGGTCTGGGACTAAAGCTCTTTCAGAATGGTGGGCAACCAATACGGCATCCTTGAAGGA
AATGCCTCGTTATCTGATTCCCAGCTACTTCGATGTTATCCTCACGGGAGTTAATGCTCTTCTCCTGGAACAGAGTGTAACGCT
AATGTCTGATTTTGTCCGGAAGGGAAGTTCGTTTGTGCAATCTCTGGCCATGGGATCAGTTCAGTGTGTTGCTGAATGCCCATC
AGCTAATCTTCCGGCTCTTAGTCCGCAAACAAAGGCTCCCAAGCCGCCAAATCAGTGTGCAACTATGTCAGCGGGATTGCCGC
ATTTTTCAACAGGATACATGAGATGCTGGGGCAGGGATACCTTCATCTCCCTCCGGGGACTTCTTCTCCTAACCGGAAGGTTC

GATGAAGCACGCTACATGATTCTTGGATTTGGTTCCTGTCTAAGGCATGGACTTATCCCGAATCTCCTCGATGGTGGCTACAA
AGCAAG
>LLOJ008312-E9 exon:PUTATIVE_protein_coding
ATTCAACTGTCGCGATGCCATTTGGTGGTGGCTGTACTCCATTAAGTATTTCGTGGAGGAGGCTCCAAAAGGAGTTGAAATCC
TCACGGAAAAAGTCTCCAGGATATTCCCAATGGATGACAGTGAAGCCCGCAAGGCAGGGGAATGTGATCAAGCTCTCTGTGA
TGTAATCCAGGAAGCCATCACTGTACACTTCCAGGGTCTTGTTTACCGTGAAAGGAATGCAGGTCCTGCCATTGATGCGCACA
TGACGGAAAAGGGCTTCAATAATCAAATTGGTGTTCATCCTGACACGGGATTTGTCTTTGGGGGAAATGATGCAAATTGTGGG
ACGTGGATGGATAAAATGGGATCGTCAGAGAAGGCAAGGAATCGCGGAGTTCCCAGCACACCACGAGATGGATCAGCTGTG
GAGTTGGTTGGTCTCCAAATGGCCACCCTCAGGTTCCTGCAGAAGATGAGCACTGAAGGGAAATTCCCCCATAAATCTGTTGA
AAGAACATCCAAGAATG
>LLOJ008312-E10 exon:PUTATIVE_protein_coding
GAACCAAAACCGTTTGGACGTACAAAGAATGGGCTGATAGGATAAAAGCGAATTTTGAGAAGAATTTCTTCGTGGGACAAGG
GGAAATACCTTTGGCTAACAAGAGAGGGATCTACAAGGATTCATGCGGAGCTACGCAGACGTGGCAGGACTTTCAACTTCGT
CCCAATTTTCCAATTGCAATGGTTGCG
>LLOJ008312-E11 exon:PUTATIVE_protein_coding
GCTCCTGAATTATTTGACCCAAAGCACGCATGGGAAGCTCTCGAACAAGCTCGAAAATACCTTCTTGGTCCGCTGGGGATGAA
AACTCTCGATCCTCAGGATTGGGGATATCACGGCGACTACGACAACTCCAATGACTCTGAAGATCCAAAAGTAGCTCATGGA
GCAAACTACCATCAAGGCCCT
>LLOJ008312-E12 exon:PUTATIVE_protein_coding
GAATGGGTTTGGCCAATTGGATTTTACTTGAGGGCTCGACTGATCTTTGCCGCACAGAATAATCGCCTGAAGGATACTGTTGC
TGAAACATGGGCCATCCTCACAGCTCATCTCAAGGAGGTGAAGTCCTCCTGGTGGCGAGGACTCCCTGAATTAACAAACTCC
AACGGAAGCTTCTGCAGTGGTTCCTGCACAACACAAGCATGGAGTATGGCAACTGTCCTGGAAGTTTTGTACGATCTGGAGC
GTCTTAAATGTAAATCTTAG
>LLOTMP008312-Intron 1:PUTATIVE_protein_coding
ATTTCTTTAAATTCATTACCTAAAGCATGAGGCATTTGCTCTAATTTTAGCCAGTATGATGGTACAACATCAACAACAACTAC
AAAATGATCTTGTTGATCATGTGTTGATCAAGTCTAGTCATCAGGGTTGTTTACATTTTTTTTGTCAAGATCTTACTATGTATGT
TGCAACTTGTTATCTATTAATTATAGTTATTTTATTAACATGTGAAATAACAATTTATAGCGTAAATTGGTCGTAACATTTTTT
AAGGCTTTATTTGAAATTAATCTAAACTTTCAATAACAAACTTAGGGTCTTTTTACAAATTAATTATCATTAAATAAAAATTTA
ATTTGTCTGACAGTCACAGCTTTTTCATACAGCTTTTTTTCTTTAAACTGAATTTATTTACATTTCTTCAATAAATTTTAATTATT
CCACGCAAAGCTTTTTGATAGAGACATTTATTTTGATTGCTTTTAATAAAACATGCTATAAAATTCACGTGGTTGTATTCTTGA
CCCAAATAAATACTTAAAAAAAAAGTTATTTCAGTGGAAAGTTTTAGTAAGATTTAGTAAAATTTTAGCGTGAAAGAAGCTTA
GAATCATTTTTTTAATTAAAGATATTCTTATAACTTTTCAAAAAATTATTGAAAAGAGCGATAGAGTTTAGAATAGATTCCATT
TGCATTTATGAGCTCATCGTGAGTTCCCATCTCGGTGATCTTTCCTTGATCCATCACACAAATAAGATTAGCGTTTCTAATGGA
GGATAGTCTGTGAGAAATGATCAAGCAGGTTCTGTTCTCACTTGCTCTATTCAGAGCATCTTGAACTGTCTTCTCACTTTCTGC
ATCCAAAGCTGATGTTGCTTCATCCAGCAGGAGGACTTTTGGCTTCCTGATAAGAGCCCTTGCAATGGCTATCCGTTGTTTTTG
TCCTCCCGATAGAAGAGTTCCTCGTGATCCAAGACGTGTTGCATAGCCCTAAAAATTCGTTTTTTATATCAAATTAAGTCTTAA
ATTATTTTTGTGAATGTTGTAACTTACAAGAGGTAGTTTTGTGACAAATTCCTTATGTATTTGAGCTTCTTTTGCTGCTTCAATA
ACTTCCGTCAGAGGTACATCTCGTGAATTGTCACCGTAGGCTATATTTTCAGCAATAGTACAGTCAAAGAGAACAGGCTCCTG
ACTTACAAGACTTAGGTAGGCACGAAGCTTAGAAAGAGAAAATTCCGTCGTTGCTGCACCATCCACAGCTACTTTTCCCTCAT
TAGGATCGTAGAAGCGAAGAAGCAGTTGAAAACATGTTGATTTTCCACATCCTGAAGGACCCACAATAGCTACGGTTTTCATT
CCATCAACATGGAAGTCTAGACCTTTAAGGATTTCCGTTTCTGGTCGATTTGGATAACTAAATTTTATATTTTTGTATTCAATA
TCTCCAGAGAAAACCTAGAAAATACAAAGTTTTGTTGAAAGATCTTCTTTCTTAAAATTGCTTAGTCTTACCATTTGGCCGTCA
TGCTTCATGCTCTGAGCAGTTTCTGACGACTTTGTTGAGGTGTCTATAAAGTCCATGATATTTGCTGCAGACGTTAGAGCAGA
ATTAACATTTGGAGCAGTATTGAGAACAATTCCCATCATCCATGCTCCATAAATTATAGCTTCGATTACCCTAAAAGGATAAA
AATCTTATTCAAATCTTGCCCTTATTCTTAGATCCTTCTTGTATTTACTTACATTAAAGCATCAATTAAGGATATTTCTCCGGTT
GATATCAAATGTCCAGCGTAGATAAAAGTAACGCCGTATGCAAAGATACTCATTGAGATATTTAGACTATATGCTAGACCTCT
ATATTTGGTACTCTTCCTTGCTTCTCCTTCTGCTATTTCAATCTTTTCAATAAATCTCCGAAGTACGTGTTTCTCCTGACCAAGG
GAGGCTACAGTTCTTACGCTCATAATTGCTTCAATTGCAATCCTTGAAGCATTTTCCATTGTTTCCTTCTCCTTTGCTGCGCTTG
ATTTTGCATAGAGAGACTCAAAAGCAGCAACAAAGATGAGCAAAGGTACGGTTAGGGAGACAACAAGAGTAAGCTTCCAGA
AAACAACTGATCCTACTATAACGCCAATTAGAATGGATGAAGTAGCTTGGAGGATTCCAGCTAAACGAATCCCCGTAGCTCC
TTGAACATTAGCACAGTCAGTGGAGAGTCTAGCAGATAATGCGCCAACGGCATTGCTTGGATGATCATACCAAGCTATATTCT
GCTTTATAGCAGTTTTGAAACACAAATTCCTAAGAAGAGAAGTAAGACGAACTCCAGCTTTCCCGAAAACATACGTTTGGGTT
ACTGAAGCTATTCCAGTTACAGTAGCTAGAATGAGGAAGTTTATTCCACATTGAAATTGAACCACATGAATCCTGTCTAGGTC
TGTTGTTTCAAAAGCTTCAAAAATTTCCGCATTTAATACACCACTTATGGGGTATGATGCCCCAACGATGATTGAGCAGATGC
ATCCAATAGCTAAGAGCTTCCAATCCTTCAAACTTATTCTTATAAGGCGCTTCATTGTATTTTGCTTCATTTTCAGAAGCTCTTC
TTTGTCTTTGACACCTACAATATCCAATTGTCCAAAATCTTTTGATGGAAATTCACTTTCTTCATCAATTTTATCGGTATTATCA
AATTGTAGACCATCATTGAGATCTGAGAATGCAACCTCTTGGCTAGCCATTCGTAGATCATGGTAGATTCCTTTTTGAAGCATT
AATTCATTATGCGTTCCAACTTCTGCGATTTTTCCCTGATCCAGGACTATTATAAAGTCAGCATTTCTTATGGTTGAAAGTCGA
TGGGAAACAATTAGAGTTGTACGACCTTTTAGAGCCTTATCCAGAGCAGTCTGGACTAATTTCTCCGACTGAGCATCCAAGGC
TGAGGTTGCTTCATCAAGAAGCAAAATTGCTGGGTTCCTAACAATGGCTCTTGCAATGGCAATTCTCTGTTTCTGCCCTCCAGA
CAGTTGGGCTCCCTTTTCTCCAACCACAGTGTCGTATCCTTGCGGTAGCTTTGTAATGAAACTGTGACAGTTTGCTATCTTGGC
TGCATCTTCAATTTCACTTTGAGTTGCTTCAGGATTTCCAAAACGAATGTTTTCTGCAATGGTTGTTGAGAAGAGGACTGGTTC
CTGATTAACTACGCCTATTTGGGATCTCAAGAATCCTAAATTGTACTGTTGGATGTTGACCCCATCTACAAGGATTTCCCCCTG
AATTACAAAAAAAATTAAATAAAATGAAAGAAGAGTTTAGCGTCTTCTCCTTAATATTATCTAAGATACTTACATCATTCACA
TCGTAAAGTCGTTGAAGGAGTTGCAAACACGTAGACTTTCCACACCCAGAATGTCCAACAATAGCCACCGTCTGACCACTGTT
GGCACAAAAACTGACTCCTTTCAATACTTTTGTGTCTTTCCGCGATGGAAATTGGAAATGAACCTTTTTAAATTCAATGTTTCC
TTTTAAGTATTCAACCGTCTTTCCTTTTTCACTAAACGGATTAATTTTAGATTTACGATCAATAACTTCATGAATACATTTTATT
GCTCCTTTTGCTGATGCTATAGTTTCCAAGCTAGGAATTGTTAATGTAACAAATTGCGCTCCGTTAACCATGCAGAAGACAAC
AGTTACAATGGAAATTAAGTTGTAGTTCCTCTGATCTTCAGGTAATTCAAATTCATCAATCATTAGCTGTATACCGTAGCAAA
AAGCCGCACCAAAACATGCAAATACCAGAAACCACATAAGGCCACTTTGAAGTCCAGTAAATAGACTCTTCTTAATTCCTGCT
GCTTCTGCTATCTTGAGATTCTCAGCATACCTCTGGGATTCTTTAGTTTCACCATTAAAAGCTTTAACTGTTCTAATTGAATCG
AATACTTGTTCTGCTATTGCTCCAGCTTTTGAGTAGCTCTGTAACTCTTTTTCCGTTAAGCGTTTTGAGGATTTTCCAACATAAA
TGGAAAAACCAATAATTAGAATTGTGTAGACTGCTGTCATAGCAAGAGCTAATTTCCATCCTACAAGAAAAGCTATAACAAC
TTGCCAGATAAAAGCTGAGAAGAGGAAAACTGTTGTTACGAGCTTTTCACCAATAGCATCCTTTATCAAGTCAATATTGCTCG
TTATATTGCTCGCAATATCCATTTCCGTCCTGGTGTCAAACCACGCCATATCCTGACTGAGGATAGATTGAATGAAAAGTTTA
CGGAGTCGTGTAGTTTGTCTAAAAGCTAGAATATTTATCATGTAAACTGCATGCGTTAGGACAACAATGATTGCTACAACAAT
CGCAACATTTAAGTAGAACATAGACATAGCATCAGCACTCATAGCTTCTGAAATTTCCATGCTGGTTGCATTTGTTAGAATTTT

TCCACCTCCAAAGTAATTAAGTAGGTTGATTTTTGTCGTTGTTGTATTGGCTTTTGTACGATCAATTGTAATTCCAAGAACTTC
AGCAAAACCAATCGCAGATAGACCAATTCCCAAGGAAACCACAGACACAAGAAAGAGATTCACAACTGTAGCCAATTTTTCT
GTAAACGTGGCATAGCGAAACTGAAATTTTATATTTAAAAATCTTTGTTGTAAAAAGAAAAATATTTAGAATTTATCAATATT
AAAACACTCACTAAATGATGAAAATCAGAGGAAAAATTTAGAATTTTGGTCTTTTTCATTTTCTCTAAATTTTCACAAGTATCA
CAAGAAGTAAGATATTTTATAAGTTGAATTTCTTCAGTCAACAGCAGATCCACAACTGAATAAACCTCGCAACTTGCAGAGA
GAGCAAAATTTCTTAATTGGTTAATTTACATTTACACTTATAATAAAAAATTTTTGTGCAGTATGTTTAATCTTTTTCGTGTTAA
TATTTTTTACTAAAATCACGCACAGTTGATTCATTTTACTTATTCATTCATACGTCATACGTGCATCTTCTTACGTGTTGAGCTT
TTTTATAAATGAATATTTAAAATTCTACTACGGTTTTCCTTAAAGATTTTCTTTTTATTTTTGTCAGATTCTAAT
>LLOTMP008312-Intron 2:PUTATIVE_protein_coding
GTAAATTATTTCGTTAATTACTCAATTATTCAGTATTTTAAACTGTTGTAACTTTCTTGAATCAG
>LLOTMP008312-Intron 3:PUTATIVE_protein_coding
GTAAGATTCATAATTTTAGGTAGACTAGGATTGTTGTCTGACTATCCTTTTTTACCAG
>LLOTMP008312-Intron 4:PUTATIVE_protein_coding
GTAAGTTTAAAGAAAACTGCTCAAAATCTTTAAGATCATTTTCTAAATTTGATAATTAAAG
>LLOTMP008312-Intron 5:PUTATIVE_protein_coding
GTAAAATATTTAAGAAGACTTCTCTTTTCTTTGTCTTGACTGAAAAAATATTTTTTCATTAG
>LLOTMP008312-Intron 6:PUTATIVE_protein_coding
GTATGAGTCAGATGAAGGATCAATCCTAAACTTTTCTAACATCTTTTTGGGTGATTCTTATTGCAG
>LLOTMP008312-Intron 7:PUTATIVE_protein_coding
GTGAGGAAGAATCTTTACGCATCCTTCTTCCTTCTTGTTCATTTTTTGTGATTCTTTTCAG
>LLOTMP008312-Intron 8:PUTATIVE_protein_coding
GTGAGTTGTGAACAAGAAGAGAATTTAGCAAAGATTTAACATTTTCTTATGAATTTCCAG
>LLOTMP008312-Intron 9:PUTATIVE_protein_coding
GTAAGAGGATCTTTTTTAATTTCTTATTTAAATTGTTCTTTTAATAGAATTAATTTGCTCCAG
>LLOTMP008312-Intron 10:PUTATIVE_protein_coding
GTAAGATTCTTAAAAATTAATTAATTAATTAATTTAAAAACTATTATTTTATTAATTCTTTTAG
>LLOTMP008312-Intron 11:PUTATIVE_protein_coding
GTAAGCCCATAAGCTAAATAATTCTTAAATTAGTAAATTAATCGACTTCCTTTTTCATTTTTAAG

***L. longipalpis* Glycoside Hydrolases from family 31**
Glycosidase XET37
>LLOJ001847_1/ LlGlyMyo1
>LLOTMP001847_1-E1 exon:PUTATIVE_protein_coding
ATGTTCCCATCGATGATAAAGAAAAAGAAGAAATCAAAGGACTTCAAGATAAAGGCCTTTGTTGCTTGTTTGTTCCTGATCAT
CGTCTTCCTCGTGGGATATGCGTACATTATGTACAACCAGCAGGTGCTTGCTCGTTCCTACTTTGATCGCGTAAAATTGAACAA
GGGCCAGCGTGTGATTGGAATCTACAACGAAAAGGGTGTTCTGATGGTAACTGGTGTTTGGGTACGACTATACATTCAGAGA
AAGCCTACCACTGCCTCACGAATAATACGCTGGAAGATGGAAGTGTGTGTCTGGAATGGGATGGCAAGGCGAGGATGTATCT
CAATTTTCCAGGAACTCAGTCCATCTGTTCCCTGCTACACGATCCGTTGGCAATCACTGTCTCCGGATGTTTTACCTACGGATT
GCTATGAGTCTTTTATGGGTCAAGGACATTGGTATGGAGGTGGCCTAACACGTGGTGGGAACTGGCCTCTGGAGACGGAATC
CTTTCCATTTGCACCCTTCATCACGGGTGATGCTAAACGACAGCAGTGGGGAAATGTCCTGAAGAGATACTTTATAAGTTCAC
GCGGTGTGGCAATTCAGGTGGATGAGAAATCCCCTCTATATGTCAGCATGAACGATAACAAATCAGGCGAGATGTGCTTCCG
TGCGAAACATGATCATTTTGCCTTTGTCAATCGCCTTACGCCACTTCCTGAGCTGAAATACAGAATTTGCACGAGCGACAATA
TGCGTCAGTTGCAT
CAGCAGATGACACAGCAGAGTCTCTGGGATGGACTGAAAGAGCATGATATCAATGTGGTTCATTCTCTTCTGGAGGAGCCAG
TTTGGCAAATTCCATCTTCTGGCGGGAAAGGTATCACGGGAGACAGCATCCACAACTACACTGAGCGCGTAATCGCACTGGG
ATTCCTCCGTTTAGGCCATGTGCTGGTCAATGAATTCTGGCAGAGACACATTGGAGACTTTACGCTGGAAACTGAACGATTCC
CAAATCTCGAGGAGACCGTTACAATTCTCCATCGTCGTGGCTTTAGGATTGTCTTCTCAGTTCAGCCATACATAAGCACGGAT
AGTGATAATTTCGCTCAGTCTGTAGCGCAGAAACTTCTAGTCTACGAGAGGCAATCTGAGAGAAGTATTCCTGCCCTGACGAG
GTATAAAAGTGTTGCAAGTGCAGGAGTATTGGATATCACAAATAATGCGTCAATTCCTTGGCTCATTGAAAAGCTGGAGAAG
ATTCAAAAAACTTACAAAATTGATGGATTCTACTTGGATTTTGGAACTTCTCAGAACATGCCCCACTATTATCAGTGCAATAA
AACCCTCTACAATCCGGATCAATATAAAACAATCTTCACAACTGCCCTCGAGGGTGTTATCAGTGTTATTGGAGTATCAAGTG
CAGTAACTGTTCCTCGACCACCTGCCTTTGTTAGCCTACCTCCGGTAAATTCATCCTGGGAAGGACTTCGAACTGTCGTGACCA
GTGCTCTGACATACGGCATCATTGGATATCCATTCATAATGCCA
>LLOTMP001847_1-E2 exon:PUTATIVE_protein_coding
GTCCAATTGGTGGAGATTATCTGCTCCCTCCTTCAGCTTCCAATGAAACTGTATCATTCTACTTCATGGAAGAGCCTCCACTTC
CGGATCAAGAACTCTACTTACGTTGGATGCAACTGGCAACGTTCCTACCTGTGATCAGATTTACTCATTTACCGTCAGAGTAT
AAGAGTGAGCTGATAATGGAAGCAGTAAAGGAATTGACGCTTATTCGGCAGAAGGCAGTCATTCCACTGCTCAAGAAGTACC
TCAGTGATGCCATGAATGAGGGTCTACCACTCATTAGACCTCTCTGGATGTTGGATGCTCAGGACACGGCTTGTCTCTATGTC
AATGACGAATTTTCAATTGGTGAGGATCTCATTGTAGCACCTATCCTCGAGAAGGGACAATTACAGCGAGAAG
>LLOTMP001847_1-E3 exon:PUTATIVE_protein_coding
TTTATCTGCCTCAGGGTGTCTGGAAAGATGGGATTGACGGTTCCCTGCGCAAAGGAAGCCGCTGGATTCACAACTACCGCGTG
CCTGAGGACAAAGTGGCTTATTTTATGAAAATGCCCGACAATACGAGATTCTAA
>LLOTMP001847_1-Intron 1:PUTATIVE_protein_coding
GTAAGTTTAATGAATCTTTCTTTTCATCATCAAACAACACAGTAAAAAATTCTTAATTTATAG
>LLOTMP001847_1-Intron 2:PUTATIVE_protein_coding
GTAAGCAATATATCAAGTTCTGTAACAGACACATACTAAAATAACTTCTTTTCACACATTCCTTAAAATCTAATTCATGTTAAT
TAAGAGCATGATTAACAATCAATCGAGTAATTAAAATATTCGCAGAATATGAATAATTTAGTCTTGCATTAATTTTGAGATCT
TTTCAAGATCAAATGGTAGCTTCCTCTCAATTTTTCAAGAACTCATGAATATCAATATAATCTAATTTGTATGTGACATGAGAA
GACACTCTGCTGATTAATACGCTTGTTTTCTTTGCAG

>LLOJ001847_2/ LlGlyMyo2
>LLOTMP001847_2-E3 exon:PUTATIVE_protein_coding
ATGATAAAGAAAAAGAAGAAATCAAAGGACTTCAAGATAAAGGCCTTTGTTGCTTGCTTGTTCCTGATCATCGTCTTCCTCGT
GGGATATGCGTACATTATGTACAACCAGCAGGTGCTTGCTCGTTCCTACTTTGATCGCGTAAAATTGAACAAGGGCCAGCGTG
TGATTGGAATCTACAACGAAAAGGGTGTTCTGATGGTAACTGGTCGTTTGGGTACGACTATACATTCAGAGAAAGCCTACCAC
TGCCTCACGAATAATACGCTGGAAGATGGAAGTGTGTGTCTGGAATGGGATGGCAAGGCGAGGATGTATCTCAATTTCCAGG
AACTCAGTCCATCTGTTCCCTGCTACACGATCCGTTGGCAATCACTGTCTCCGGATGTTTTACCTACGGATTGCTATGAGTCTT

TTATGGGTCAAGGACATTGGTATGGAGGTGGCCTAACACGTGGTGGGAACTGGCCTCTGGAGACGGAATCCTTTCCATTTGCA
CCCTTCATCACGGGTGATGCTAAACGACAGCAGTGGGGAAATGTCCTGAAGAGATACTTTATAAGTTCACGCGGTGTGGCAA
TTCAGGTGGATGAGAAATCCCCTCTATATGTCAGCATGAACGATAACAAATCAGGCGAGATGTGCTTCCGTGCGAAACATGA
TCATTTTGCCTTTGTCAATCGCCTTACGCCACTTCCTGAGCTGAAATACAGAATTTGCACGAGCGACAATATGCGTCAGTTGCA
TCAGCAGATGACACAGCAGAGTCTCTGGGATGGACTGAAAGAGCATGATATCAATGTGGTTCATTCTCTTCTGGA GGAGCCA
GTTTGGCAAATTCCATCTTCTGGCGGGAAAGGTATCACGGGAGACAGCATCCACAACTACACTGAGCGCGTAATCGCACTGG
GATTCCTCCGTTTAGGCCATGTGCTGGTCAATGAATTCTGGCAGAGACACATTGGAGACTTTACGCTGGAAACTGAACGATTC
CCAAATCTCGAGGAGACCGTTACAATTCTCCATCGTCGTGGCTTTAGGATTGTCTTCTCAGTTCAGCCATACATAAGCACGGA
TAGTGATAATTTCGCTCAGTCTGTAGCGCAGAAACTTCTAGTCTACGAGAGGCAATCTGAGAGAAGTATTCCTGCCCTGACGA
GGTATAAAAGTGTTGCAAGTGCAGGAGTATTGGATATCACAAATAATGCGTCAATTCCTTGGCTCATTGAAAAGCTGGAGAA
GATTCAAAAAACTTACAAAATTGATGGATTCTACTTGGATTTTGGAACTTCTCAGAACATGCCCCACTATTATCAGTGCAATA
AAACCCTCTACAATCCGGATCAATATAAAACAATCTTCACAACTGCCCTCGAGGGTGTTATTAGTGTTATTGGAGTATCAAGT
GCAGTAACTGTTCCTCGACCACCTGCCTTTGTTAGCCTACCTCCGGTAAATTCATCCTGGGAAGGACTTCGAACTGTCGTGACC
AGTGCTCTGACATACGGCATCATTGGATATCCATTCATAATGCCAG
>LLOTMP001847_2-E2 exon:PUTATIVE_protein_coding
GTCCAATTGGTGGAGATTATCTGCTCCCTCCTTCAGCTTCCAATGAAACTGTATCATTCTACTTCATGGAAGAGCCTCCACTTC
CAGATCAAGAACTATACTTACGTTGGATGCAACTGGCAACGTTCCTACCTGTGATCAGATTTACTCATTTACCGTCAGAGTAC
AAGAGTGAGCTGATAATGGAAGCAGTAAAGGAATTGACACTTATTCGGCAGAAGGCAGTCATTCCACTGCTCAAGAAGTACC
TCAGTGATGCCATGAATGAGGGTCTACCACTCATTAGACCTCTCTGGATGTTGGATGCTCAGGACACGGGCTTGCTCTATGTC
AATGACGTAATTTTCAATTTGGTAGAGGATCTCATTGTAGCACCTATCCTCGAGAAGGGACAATTACAGCGAGAAG
>LLOTMP001847_2-E3 exon:PUTATIVE_protein_coding
TTTATCTGCCTCAGGGTGTCTGGAAAGATGGGATTGACGGTTCCCTGCGCAAAGGAA GCCGCTGGATTACAACTACCGCGTGC
CTGAGGACAAAGTGGCTTATTTTATGAAAATGCCCGACAATACGAGATTCTAA
>LLOTMP001847_2-Intron 1:PUTATIVE_protein_coding
GTAAGTTTAATGAATCTTTCTTTTCATCATCAAACAACACAGTAAAAAATTCTTAATTTATAG
>LLOTMP001847_2-Intron 2:PUTATIVE_protein_coding
GTAAGCAATATATCAAGTTCTGTAACAGACACATACTAAAATAACTTCTTTTCACACATTCCTTAAAATCTAATTCATGTTAAT
TAAGAGCATGATTAACAATCAATCGAGTAATTAAAATATTCGCAGAATATGAATAATTTAGTCTTGCATTAATTTTGAGATCT
TTTCAAGATCAAATGGTAGCTTCCTCTCAATTTTTCAAGAACTCATGAATATCAATATAATCTAATTTGTATGTGACATGAGAA
GACACTCTGCTGATTAATACGCTTGTTTTCTTTGCAG

>LLOTMP001881/ LlGlyMyo3
>LLOTMP001881-E1 exon:PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXGGTCCACAGCAAAAGTATCAGTACTGGCCAGTTGAGAAACTCGTCCTATCCAATTATTCCTACGTTACAAAGAAGC
CGACAGTGCGGCTATTGCTGAGAGATATTGGCTCAATTCGGAAGGAGCTTTTATTTATCTCTCAGACCGAGCTCCACTCTTCAT
TGACCAAAACAGCGTGGTCCACAGCAAAAGTATCAGTACTGGCCAGTTGAGAAACTCGTCCTATCCAATTATTCCTACGT TAC
AAAGAAGCCGACAGTGCGGCTATTGCTGAGAGATATTGGCTCAATTCGGAAGGAGCTTTTATTTATCTCTCAGACCGAGCTCC
ACTCTTCATTGACCAAAACAGCATGATGGACAATCATTTGTGCTTCGTTGCTAGCCTTGCACTGCCCTACAATCCACGTGCTCA
GACGTATGATTTTGAGTACAGCATCGGGGTGGCTTCTGATCCGAGAAAAGCTCACATGGAGGCTGTTCATCGGTTTCTGGGAA
AGCCCTCAGGGGTGCCTGATAGACTAATGGCAGCACGACCTATTTGGTCAACATGGGCCAGGTACAAGAGGGACGTTGATCA
ATATGTAGTGATGGAGTTTGCAAAAGAAATCATCGACAACCGCTTCACAAATAGTCAATTTGAGCTTGATGATGATTGGGAA
ATTTGCTACGGAGCTTTGACTTTTAGCACTACAAAATTCCCCAATATTCGCAAAACCGTTCAAGATCTAAAGTCAATGGGATT
TAGGGTAACGCTGTGGATTCATCCGTTCATTAACAAGGGCTGCAACCCATGGTACAATGAAGCTCTCACAAATAG
>LLOTMP001881-E2 exon:PUTATIVE_protein_coding
ATACTTTGTTCTCGATCACAATGGCAAAGAGGATACTCAGTGGTGGAATAGCGAACAAGGTGAGGCTGCGTACATTGATTTTA
CAAATCCTGAAGCTGCAAATTGGTACACATCGAGATTGCATAAGCTCCTCGAAGACACAGGCATTGATAGCTTTAAGTTTGAT
GCTGGTGAAAGTAGTTGGCAACCTGAT
>LLOTMP001881-E3 exon:PUTATIVE_protein_coding
GATCCAGTACTACCCTCTACTTCCCTGGATCAGCATCCATTGAAAATTCAAACAGACTACATAATGCATGTTGCTCAATTTGG
CCCAATTGTGGAAGTTCGATCGGGTTTCCGTAATCAATTCCAGCAGATTTATATGCGAATGATTGACAAGGATTCCCTTTGGA
GCTGGAACAATGGTCTCCCAACTTTGATCACGACGCTCCTTCAGCTCAATATGAATGGATATCCTTTGGTTTTGCCCGACATGA
TTGGAGGCAACGGCTACGAGGGAGCACCTTCGAGGGAAATGTTTATTCGCTGGCTTCAAGCTAATGTTTTCATGCCCAGCCTT
CAGTTCTCATACGTTCCATGGGACTACTCGGATGGGGATGGCATTAAGATCCTCGATCTATGTAGGCATTTTGTTAAACTGCA
CGAAGACTACACTGACTTGATCATGGAACGTTTCCAAAAAGCCGTAGACCATGGGGAGCCTGTTAATCCACCCATTTGGTGG
GTTGATCCTAAGAATCCAGTAGCTCAAAGTATCTACGAT
>LLOTMP001881-E4 exon:PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
>LLOTMP001881-Intron 1:PUTATIVE_protein_coding
GTAATTTTTTGAAGGAGCTTTGCAGCAAAATCAGAATGTTTAAATTCAATTCTTCTCATTCAG
>LLOTMP001881-Intron 2:PUTATIVE_protein_coding
GTGAGTAGTTTAGCTAATAACATTAAGGATCTCGTTAACAGGGTTTACTGCATTGTAG
>>LLOTMP001881-Intron 3:PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXX

>LLOTMP000840/ LlGlyMyo4
>LLOTMP000840-E1 exon:PUTATIVE_protein_coding
ATGGAAAAGAAGCGCGTAAAGGGAAACCCAAAGTATAAGTCCGCTTTTAAGTGGGGTACAGCTGGAATTATAGCCACTGTGA
TTTTCGTTCTTGTGGAGGTTTTTATAATTCACTGGGTGACACCATTTGATAATTTCCTTAGGGTTCACTTCCCTAAAGCCGGAG
TTCATGTGCAATTTGAGTGTGAACACATAGAGGATGGGGTAGCATTTAGTATTCACAAGAATCGCACCCTTGTTCAAATTGTC

CGGATGGGCGACAATTTAGGATGGGATTTGAGGGTTGAGGATCACGGAGGTGGGAAGTATGTTTTAAGGAGTCGCGAGGGAT
CAGTGAGCTTTAATACGGCTGTTGATGAAGATGGTTTGGCTGTTTTTCACATCAATCAAACAATTCGTACAACGGATTTCGTTG
AACATTGCTTTGATCTGGTGAATCCTGATGGGAATAATTGGTTTGCTGGTCCGCATAAATACTACCAGCATTGGCCATCACAG
GTTCTACAATTCACAGATGATGCCTACTTGCCGAAAGAGTCTTCACATTCATCCATATCCGAGAGGTATTGGCTCACATCCAA
AGGGATGTTCATCTACTTCAGTGACCGAACTCCACTATTCCTCAATCGCAAACCCAACCATTTGTGCTTTGCTGCAAAGAAGC
AAAATCCTTACTACACCTACGACACAGTTTTTGTCTTTAATTATACAATTGGAATAGCTGCAGATGCCAGGAAAGCTCACAAG
GCTGCAGTGGGTAGGTATTTGAGGAAACCCCTTCATTATCCTGACAGGAGAGTAATTCAGCATCCAAAATGGATCCTCCGAA
ATCACAACATGTCTCAAGTTCAGGAATATATCGACAATTTGCTGTACTACAATTTTAGCACATCCAGCATCATTCTGGATAAA
TCATGGGAAACGTGCACAGGAGCTTTGGAGTTTGATCCGATCAAATTTCCCAATGTTGGGAGTATGATACGATATTTTAGACG
CCACAGAACCCATCTCACACTCACTGTGACTCCGTACATTCACGAGGATTGCAACCCCTACTTCCAGGATGCTCTCTCACGTG
GATTTTTGGTCAAGACTCACCAGAAT
ACCTATTTAAATTTCGATCGAACTGCTACTGTGGACTTCACACAGGTTGCTGCTAGACGGTGGTTTCTGGCACGACTCAAGGA
TCTGCAAGCTCTAGGCGTTGACAATTTCTACTTTGAAGGTGGAGAGTTCGATTCAGCGCCAAATGATCCGAATTTCGAAGGAG
TCTCAACGAGTTTGCATCCCATTCAGTACACCATTGATTCCCTAAGAGCATTGGCGGAGTTTGATCACAACACAATTGTTCGA
ACGGGATTTTGCACCCAGGATCTTCCTCTTTTGCTCCGACTTCCGGATATGGATAATCGATGGGACATTCAAAATGGTTTAGAT
TCGCTTATTCCGAAAATATTGCAGACAAACCTCAATGGATACTACTTCCTCATGGCTCCAATCGGAGGAACTACTTCAGAGGG
TCTGACAAAGGAGCTCTACATACGCTGGATGCAAGCTGTGGTCTTCTTGCCATCAATGGCCTTTTCGACTCCTCCATGGGAGTT
TGATGATGAAACCATTGAATTAGCTCAAAAATTCATTCGTCTTCATATGCGACATGTATCCTTATTTGTGGATCTGTTTAAGTT
GGCAAAATCGGATGGGGATCCTGTTAATTTACCTATTTGGTGGCTCTACCCGCAAAATTGGCGTGCACAGGAAACTCATGATC
>LLOTMP000840-E1 exon:PUTATIVE_protein_coding
AGTATCTCCTGGGTGAGGACATTATCGTGGCTCCGGTGCTGAAGCCAAATTCTTTGGAAAGAGGAATATTCCTCCCATATGGC
ATATGGCGCGATGGCAATGATAACAGTTCACTCTACCAGGGACCTCGGTGGTTACCTAGCTATAGAGCCCCATTGGATGTCCT
CCCATACTTTGTCCGAGCTCGCAATAAAAATTAA
>LLOTMP000840-Intron 1:PUTATIVE_protein_coding
GTAAGTAAAAGAAAGAAATATTAACGAACGAATTGTCCTTCAATTTACTTGTCTGTTTGTATCCTTGTTTGTGTGGATCTCTGA
AAGATGTTATGAATGCATGAAAATGAATAAAGTCCTTTACAGGAATCTGTTGAACAATAAATCTAATAATTAAAACCAAGTTT
ATTTGTTGAATTTAAATTAAATTAATTTATTGAAACAAGGATCTAATATAATTTGTTTTGCTACACATGGTTTTGCGGATTAAC
ATGGCTACAGAGGAATCATGGATATCTCTTCTAGATAGATGTCCCTAATACATGATGACCTTTGGTTTGATTTCGTTTCTTATT
TTTCAAGGCCAAGAATTCTTAATTTTTTTTTTCGTAGCAAATCCGCAATACATTTTTTTTCTAATTAATTTTAATAATTTTTTTT
GTCTTTTTTTCGTAG

Lysosomal alpha-glucosidase
>LLOTMP006451/LlLysAglu1
>LLOTMP006451-E1 exon:PUTATIVE_protein_coding
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXAATGTTTTCAAGTTTGCATGTTTTAA
TTTTGATTTCCGTATTTCCATCATTTTAGATGCTACAACCGAAATGATGCTACAACCGAAATGTTCCATGGAATTTATATAAAC
GCACCGTATCCAGTTGAAGTTGAAATCACTGAGAGCTATTTAGTTATAGTTAGAGCTATTTTTGATACATCATTTCACATTGAT
GTACTATGTGGTCCAACAATCGGTAATTTGATGGATCAGTTAATGGAGCAAACAGAACAAAATCTACATCCACTGCCAAATT
GGTTTTTTGGTTTTCACATTTGTGATACGAATATTACAAGAAATCTCACAGATTCAGTGATTGAGGTTCTGGAGTTACTTTCCT
CTGATTTG
>LLOTMP006451-E2 exon:PUTATIVE_protein_coding
CCATTTGATACTCATTGCATACGAGAACAGTTGCTGTGGTTGGGTGACGATATAAAGGGATTACCAAGCGAG
>LLOTMP006451-E3 exon:PUTATIVE_protein_coding
TACTTCCAGTTGATCTCAAGAGTGCAGCTTATCATATGGCCAAAGCAAAAGGGATTATTAATGAGGCATCCCACTTTAGAAGA
GCCTTACATTGGCCGTATTGGCAAGACGAATGTGGTCTATGTGGACTGGGTTGGATCAAATCAAGAAAATTTGTATGAATGGT
CAAATCAATTCATGCCCAACAATTTATCTGCAGATGGATATATGGTGCAGGCCAATTGGATGTGGGATCAAAGTGATGGTGA
AAAAATACCCCACTCATTTCCATATATTTCAATGGATATGATAAATGCCAGCAAGAATATAGTACCGTGGTTTATTCGTCAAA
GTGGCACAGATGGAAAGGCCTCAATTTTCTCCCATAATTTTGCAGCATCAGCGCAGATTGACGTAATAAGGAAAGTCCTTGGA
AATGAATCGTTTATTTTATCTGAAAGTTCGCTATTTGCAAATGTACCAATCTCTCGTAGACAAGTACCATCTACATGGACAGA
ATTGAGAAATTTA
>LLOTMP006451-E4 exon:PUTATIVE_protein_coding
GTTCGTCGTACGATGGGCCAATCGATTTCTGGAATTTACTTCAGTTCAGCAAATATTTGTGGAGATGGCGAGACATGGACGGA
AGAGCTATGTATTCGTTGGTATCAATTCGCTTCGCTATCACCTTTCTACCGAGTAGCGTCCGCAAAGGCTCCAACATCATTTTC
GAAATTCGCTCAAAGTTTGCTGAGAGCATCAACGGAGAGACGATACTCACTTTATTTGTACATTCATACTGTTGTGACAGAAA
GAAGACCTCTTCTCACACCTTTATTCTATGAATATCCTGACCTCATAGCATTGATGGACAATTTAACACACCAGGCGCTTGTCG
GACCCTCTCTCCTGTTTGCACCAGTACTGCTGCCTGGTGTTCAACATATCAATATTTTCTTCCCTGAGGTATACTACGAGATTG
GCGGTGGGCAGCAATTAATTGCAAATACGTGGGCTGAACTACCTGTTGTTGAGACAGACGCACCCCTTTTCATTCGTGCTGGA
CATATTGTTCCTATACAGCATACAAAG
>LLOTMP006451-E5 exon:PUTATIVE_protein_coding
AATGTTCGGTCAATAACAGTCATGCAAATGAAACCTATTCATTTATATGTCGCTCTGGGTCCATTGGAAAATGGCATAAGTAC
TGCAATAGGTAAAGTGCAATTTGATTCAGAATACAAAGTTACATTTAACGCGGCTTACACAACCTCATCAGAACAG
>LLOTMP006451-E6 exon:PUTATIVE_protein_coding
ACGCAACTAAAATTATCATCGTACTCTAAACTAACTTGTCGCACCACTACAAATCTCACCATAGAAAGT
>LLOTMP006451-E7 exon:PUTATIVE_protein_coding
CGATGCCGTGTTGAGGAGGGATTTCTCCCGGCAATTTTCGAGAAGTTTCAACCCCTCTCCTACXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
>LLOTMP006451:LLOTMP006451-RA intron 1:PUTATIVE_protein_coding

GTAACTGAGCAG
>LLOTMP006451:LLOTMP006451-RA intron 2:PUTATIVE_protein_coding
GTGAGGAGCGCTATTAGTAAAATGAAATGGCTCGGAAAAAGTTTTATTGCTCCCATGACTGCAG
>LLOTMP006451:LLOTMP006451-RA intron 3:PUTATIVE_protein_coding
GTAAGTATCTATAGAAAATATATTAGACTTATGTACCTATAACGCATATCTTTTTTACGTAAATATTTTGAAATGTGCCGTTTT
ATAATATTAATTTCTAGTTTTTTTTTTTAAAGAAAATATTATATGTATTTAGGTTTACAGGTATATGAATGTACTTCACATGAAT
AGGCTTAGATGTAAAAGTTCATTTAAAAATACACAGCTAAATCCTTTTAGCTGTGATTCTTTTTATATTGTTGTGATGCCTGTG
ATGCCCAGTAAAGTTGGTTGTTTTAAGTAAGAAAATCTTTGATTCAG
>LLOTMP006451:LLOTMP006451-RA intron 4:PUTATIVE_protein_coding
GTATTATTTGCTGCTGTATTATGTATAGAAACATTAACAAATTTCTATACATCATTTTATATTGGAATTTCATATTGAATTGTAT
TGAATGTTAAATTTCATTTTAG
>LLOTMP006451:LLOTMP006451-RA intron 5:PUTATIVE_protein_coding
GTAAATTTATATTTTTGTAAAGATATTTTATCAATCCAATATATATAATATATATATTTAATTTCAATTAG
>LLOTMP006451:LLOTMP006451-RA intron 6:PUTATIVE_protein_coding
GTAAAAATCTATGGTTCTGGTGTTTCGGATGTTGGCGTCTTTACATATTCTATAAAAGTGGGGCATAATTTATGTGGTGGAAAT
TTAAACTTGACAGAAGATTTGAGTTTTCTAAATAAAAGAAATTCAATATAAGGATTACTGATTTTTTCTTTATGTATTAGCAAT
AGTATTTGTCACACGAAAAATTTATTGTATTGTACCTACTGTATACAGATTGTTTGTAATAAATAAAAAAAATGTATTAGTTTT
CTCAATGTACATTGTATATATTCGCTTGGCCCATTCATTATGAACCTATCTCAAAAACGCTACTTTTCAGGAGAGCCGATTTTC
TAAAATGTATGAGAAATTTACGTATTTTGCCTTTTTTCTGGCTCCAACTTACCAATTATCTCACACAATGATGGTATTCACTCA
CTATTACCATTTAGAAAGCACTGGCACCCCAAATTATGTAAAATAATTAGCAAAAAGTGCAAAAATATTGCAAAATATGCAC
AGAAAATACTTTATAGGTTCTCCTTGCACGGAGTTTGAGAAAATTTTTATGTGCCGGAGTGTTCGGCAAAGCTCGGGTCGACC
GTCCTTGAATGTAAGCACTAGGTCTCTCTTTCTTTCCTGCGTGCGCCGATAGATTCGTGTGCATCTTGTATGTGTGTGTGTTAG
CTCCACAATTTCCTTCAGCGAGAATGAATTTTTCGTCCCAAATGGAATTGTTGCTTCTCTGCCACCAAGGGGATTTGGAATCTC
ATCTCCATAACCCAATGGATTATCACATCCTGGACAATTAGGGGGGGCTGGGGCAAAAATGTTTGTTTTTTATCAAATTCTGG
AAAAATAGGGGATTTTGGGTCAAATTGGTGGGTTTTTCGCCGGAATTGTGGTCACTCCTCATTGGTAATGATGGAGTGACCTC
CGGATTGGGGATCTTCATCTCAGGAATATCAACTCCTGGACAATTAGGGGGGCTGGGGCAAAATTTTTTATTTTTAGCCGGAA
TTGTGGTCACTCCTCATTGGTAATGATGGAGTGACCTTCGGATTGGGGATCTTCATCTCAGGAATATCAACTCCTGGACAATT
AGGGGAGGCTGGGGCAATTTTTTATTTTCGATGCAATTGTGGTCACACCTCATTGGTAATGATGGAGTGACCTTCGGATTAGG
GATCTCATCCCAGGAATATCAACTCCTGGACAATTAGGGGAGGCTGGGGCAAATTTTTTGTTTTTTGCCGGATTTGTGGTCACT
CCTCATTGGTAATGATGGAGTGACCTCCGGATTGGGGATCTCACCCCAGGAATATAAAATTCTGGAAAAATAGGGGATGTTG
GGGCAAATTGGTGGGTTTTGCGCCGGAATTGTGCTCACTCCTCAGTGTTAATGATGACTAAAATATTTTATAAAAATTTTTTCA
ATAAACAAGAATATTTTCAAGCGGCTTCCTCAGAAAATCGCACTGCCTATATATAACTTTTCTTAATTAAAAAAAAAATATTGT
AAAGCCCGTAAAGTTTTCGAGCAATACCAAACTTTTTATTTAAAATTTTTCTTCAATATAATAAAAAATCGTCAAGCCCGTAG
AGCCTTCGCAAAGAATCAAACTACTTTTCTTTTCTATTGAGAAGAAATACCTTTTCTTCTCAATAAATTCACTCATTATTCTTA
AAGAGTTCTTTCCTTTTTCCTTTAAGACTCCCGAAAATTCTTTCAAAGAACAAAATAATTCTTAATTAAAGAAAGAATTAAAA
AAAGAAATGCCCTTTCTTCTCAATAAATTCACTAGTTATTCTTGAAGAGTTCTTTCCTTTTTCCTTTAAATTCCCGGAAATTCTT
TCAAAGAACAAAATGATTTTTTATTAAAGAAAGAATTAAAGGAAGAAATACCTTTTCTTCTCAATAAATTCACTAGTTATTCT
TGAAGACTTCTTTCCTTTTTCCTTTAAATTCCCGGAAATTCTTTCAAAGAACAAAATGATTTTTTGTTAAAGAAGGAATTAAAG
GAAGAAATAGCTTTTCTTCTAAATAAATTCATTAGTTATTCTTGAAGAGTTCTTTCCTTTTTCCTTTAAATTCCCGGAAATTCTT
TCAAAGAACAAAATGATTCTTAATTAAAGAAGGAATTAAAAAAAGAAAGGCCTTTTCTTCTAAATAAATTCACTAATTATTCT
TGAAGAGTTCTTTCCTTCAGGATTCCCGAAAATTCTTTAAAAGGACAAAATGATTCCTAATTAAAGAAGGAATTAAAGAAAG
AAAGGCCCTTCTCAATAAATTCACTAATTATTCTTGAGGAGTTCTTTCCTTTTTCCTTCAAGATTCTTTATTAAAGAAAGAAAT
GCCTTTTTCTTAACAAATTCAGTAATTATTCTTAAATAGTTCTTTTTTTTCATTTTTTTTAAATACCCCAATTTTTTTACAAGAAC
AAAATGATTCAAAAAAATTTTGGATCTAAAAAAAATGGTGAATTTTCTAAGAAGAGAAGGCATTTCTTTCTTTAATCCCTTCT
TTTATAAAGAATCATTTTGTTCTTCTAATTTTTCTTTATTTTCCTATTAAAATTAAAATGGTTTCGTTCTTCTAAAGACTTTTCAG
ATACCTATTTTGATAAAAATTAAAAGGGAAGGATCTCTTTAAGAATAATTAGTGAATTTTTTGAGAAAAAAGGGAATTTCTTT
CTTTAGTTCCGTATTACAGATAAAACTACTTTTTTCTTCTTAAAATTTTTAAGGCACTTTGATAAAATTTAAAAGAAAATAACT
CTTTAGGAATAATTAGTGAATTTGTTGACCAAAAAAGTAGATTTCATTCTTTAAATCCTTCTTTGATTAAAAATTATTTTTATTC
AGGCTTGAGAATATTTTTCTATTAAAATGTAGTATAGTTATTACCGAAGGCTCTACGGGCTTTACAATATATATTTTTTTTAAA
TTAAGAGAAGTTATAAATAGGCAGTTTGACTTTCTGCGGAAGCTGCTTGTAAATATTCTTTTTATTGAAAATAATTTTTTTAAA
ATATTTTAGTCACTTCATCATTGTAAATAAAAAGTGACCACAATTTCGGTGCAAAACCCTACAATTTGCCCCAACATCCCCTAT
TTTTCCAGAATTTTATATTCCTGGGATGAGATXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXTTTAAACTTGACAGAAGATTTTGAGTTTTCTAAATAAAAGAAATTCAATATAAGGATTACTGATTTTTTCTTTATGTATTAG
CAATAGTATTTGTCACACGAAAAATTTATTGTATTGTACCTACTGTATACAGATTGTTTGTAATAAATAAAAAAAATGTATTA
GTTTTCTCAATGTACATTGTATATATGTATCGACGCTCCCGAAGATTATGTGTTAAAAGATAGGAATATGGTTCATAATCTTCG
GGCACGTCGGTATTTTTTTACAATGTACATACATATCGAGGGCTGCATTGTXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXCTCTTATACCTTTTGTTGGGAGATAACTA
CAAAATAGATATACATACATACATAACTCTCTCAAGGAGGAGAGTGAACTAGTATAAATGAAGTATTTTCCATAAATTGTGTG
GAAGCCTCGCTCTTTGAATCGCAACATATTGCCATTCTAGCATTCACAATCGAGTTAAATTGAGAGATGCATTGGGAGATGTA
GAAGTGAATTAGATAAATTGAATTTTATCCGTTTCGCTCTGTAGCCTAAACTAAATAATTGCCTTTAATTTGACGATGAGACG
ATGTGAATAATAATCGCAGTGGATGCACTCACAGCACGATTATATTCAACACACAAATGTGAATGACATGCCGCTGCTAGCA
AGCAAAATGCATAATATAGGAAGTTTTCTTCAATGCAGAAATTGTGAACCCCACTCATAGCCCTCCAATTCTCAAACACATCA
CGCTCTTTTCTGCATAGATTGTCCAATTAAGCGGAGGCATATTGTATGTATCTGGTCGTTTTTTTTCATCCTATACAACATCTC
GTATTGTTCATGCCATTGCATCACATTAATTGGACCATAAATGTTCATTAAAATGAAAGCGATTCAGATCACAAAATGCTACG
ATATGAATGAAGTAACACATAAATTGCTTGTATACGGCTCTGTCCCTCTAATCTCAACACTATTTGTCCCATGACAACGCATTC
CGCAGCTTGATGAATACTTTCAGCATTTAGCATTCATAAAATTTCGTATACAACAAGCACAGTTCAATGGAATTGATGTATAA
AATGATTTTGAAGTGGAGGAAGAGATATGCAAAAACCATGATTATATTGTGACTGAAATTGAGGCTTTTGTGTATTTTGGATT
CATTAAAAGCATCCTGTGAATGATGTTGAACAGTGGTGTATTCAGGGGAGGGGGATTTGTGTGTCCAAACACCAAATAACGT
TAGATTGGTTTTTAGAGAGATCAAAAGAATTATTTTTAGGCATAAAATTATACGGAAAATTTTAAAAATTGGATTTACTAATG
TTAATAGAATTTTTTTGACATTAGAAAAAAATTATTTCAAAGGTTTTAAAAAAAGTTAAATGTTGGAAGAAACTTCAAACGTT
GAAAGAAACGTCAAACGTTGAAAGAAACGTAAAAAGATGAAAGAAACGTCAAAAGTTAGAAGAAGCATCAAACCTCGGAA
GAAACGTCAAGCGTCGGAAGAAACGTCAAACGTCGGAAGAAACGTCAAGCGTCGGAAGAAACGTCAAGCGTCGGAAGAAAC
GTCACACGTCGGAACAAATGTCAAATGTTGGAAGAAACGTCAAAATTTGAACGAAACGTCAAACATCAATCAAAAATTAAAA
GAGAAATGTTAACCATTAAAAAAAAAAACTTTTAAAGAAATGTTAATTGTTTGAAAGCA AATTCGAAATTTAAAAAAAAATG
TCAACTTAAAAAACAGATTTGTTTAACATTACAAAAGGAACTTTAAACGTTTGAAGTAACTCCAGTCAATAGAAAATAAACG
TAAAACATCACAGAAACGTTAGAAAATAAATTCTAAACATGAGAAATAATGTCAAACAGTTGTAGGAAAACGTCTAACGTTT
TAAGAGATGTCAAACATTAAAAAAAAAGAATGTCATGAGTTTAAAAATAAACATCAATCGTTATAAAAA TAAAATACTTGGT
ATGCCACGATTGTGGTATAACAATTAATTAGCTTTAAAGAAAAACGTCAAACTTTAAAAATGCAAAAAGGCTCCATAATTTGT
TAATTGTAATACTAACAGTCGTTGTATGATTTAATTAATTAAATTTATTAGGTTGTGCCATAAATTTCCAGATTTCTTACCATA
ATGTACATTTTTTCTTAAAGGTTTTTTTTCTTAAGGAAATATTTTTAATATCTTTTATATTTAAAAATTTCTTTCT ATGATATCAT
CCTTTTGAATAAATAAATAAATTGAAAAAAAAATTTTGCGAATTTATGGAAGAGATGGAAAAGCTTTCAATAGCTGCTCTTGT
CACTAATTTTATATGCTAATAAAGTTTATTCTAGAAATTCCATTCTAGCAGAGAATAAATCATTAATTCATTCAAAATATAAAT
TAATATTTATTTTGGGAATCTTTAGATGAAAATTGACTCATTTATACTAAAAATCTTTGTTCTAGAAGACTTTCACAAG GTCAA
ATACATAAAGAACCTTCTAGAACTGCCCCAGAATCCTCATATTTAGAAACATTCATGCCAAATTCTTATGAATGAATAATTCC
TCACACACGTAAAAAAAAACTCTTCGCAGCATTTCTAGAAATATTATTTTGATTTCCAACACAAAACTCCACGGAAAGAACCA
AAAGAACCCACCTCAATCCACACATTCTTTTCCCATTCCCTCTCAGCACACAGTAGGTTGTTTTGTAATTAAAAGCGAAATGT
ATGGGTTAGAAGGTAAAATGAGTTTTCTTTTCCTTCAACCAATTTTATATGGGAGTGTGAGAGGAGTGCCTAAAAAATACAGA
AAGCTCCTCTGTCTATTTATAAAAACAATTTAAGCCCAAACACGAAATTTCTGAGTTTGCGGAGAATGTTGAATTGTATAAAT
ATGATTTCTCACAG

Neutral alpha-glucosidase
>LLOTMP003489/LlNAglu1
>LLOTMP003489-E1 exon:PUTATIVE_protein_coding
TGGCACTGCTGGTTGTATGGATGACGTTGTGGGAAAAAAAACCATAAATTGATGAAAAATTCAACAAGATAAAAAATTCATT
GGAAAATTACGTGGAAAATAGTGAAAATACTCGACCGGAAAAGTTAAAAAGTGCTCATAGATAAGTCGGAGTAGCTGATAA
G
>LLOTMP003489-E2 exon:PUTATIVE_protein_coding
ATGAGGAAACACAATGAGGTTTCTTCTTTTGGGGCTCCTGGGACTCCTAATTACCCTCACGGGGAGTGTTGATAAGAATAACT
TCAAGACATGCGACCAGAGTAGCTTCTGCAG
>LLOTMP003489-E3 exon:PUTATIVE_protein_coding
GCGCTGCCGGAAAGTTGATCCGGGAATCCCTTCTCCGTACGAAGTTCAATTGGGTACACTTAAGACATTCCCAGATCACATTA
CGGTGGATGTGCTGAACAAAAATAACGAGCAGGTGTTTAATTTGAAAGTTGTTGGGCTGAAGGGGAATAAATTCCACGTGGA
AGTTGATGAGAAATCCCCGCTGAAGCCACGATATCGCGTTGTGGATGCCCTCATAGCACCCCCAGTTACTGAGAATGTCTCCG
TGACGAAGGAAGAAGAAGGATCGGTTGTTGTGAAATGCGGTGAGAATAGAGCTGTCGTCGTGGCATCACCATTCCGCATTGA
TTTCTACCACGGGGAAGTTCTTGTGGTGTCAGCAAATGCCAAGGGGCTAATGAAATTTGAGCATTTGCGCAAGAAATCCGTAC
CTGTTCCCGCCGAGGCTGCTGAAGAGGGTGGAGAAGGTGCCGCTGCGGCAGAGAATGAGGTGGAGAA
TCCCCAAACAAATGAGGATGATGACCCCGGAGCATGGGAGGAGAATTTCAAGTCGCATCACGACAGTAAACCAAATGGCCC
AGAAGCTGTGGCTCTGGACTTTACATTCCCACAAGCTGAGGTACTTTTCGGGATTCCCGAGCATGCAGATTCATTCGCACTTA
AACCCACTCTTGGCTCTGAACCTTACCGGCTGTACAATTTGGATGTGTTTGAGTATGAATTGGATAGCCCAATGGCCCTGTAT
GGATCCGTTCCTGTCATCTATGGGCATGGAGAGGGCAATACAGCTGGAGTTTACTGGCAGAATGCCGCCGAGACGTGGGTGG
ATGTCTACAATTCCAATGCCAAGAAGAACGTTATGTCGTCTATTGTGAATTTCGTATCGAGATCTCGTCAAGCTGATCCTCCG
GCTGCTCATTTTATGTCTGAAAGTGGAATTATGGACTACTACGTCCTTCTGGGGCCAACTCCCATGGAAACCTTCATGCAGTA
CGCAGATCTTACGGGTCCTGCACCACTTCCGCAAATGTTTGCCCTTGGCTACCATCAGTCACGTTGGAATTACAATGATGAGG
CTGACGTAGCGGGGGTACATGATAAATTTGACGAGCATGACATCCCCATGGATACGATTTGGTTGGATATTGAGTACACTGAT
GCGAAGAAATACTTCACGTGGGACGGACACAAGTTCCCCCATCCCCTGGAGATGATCAGGAATCTCACAGAGCGCGGGAGGC
ATCTCACAATCATCATTGATCCCCATATTAAGCGAGACGGTGGGTACTTCTTCCAC AATGACTGCACAGATCGTGGGTACTAT
GTGAAGAATAAAGATGGGCGAGATTATGAGGGATGGTGCTGGCCCGGAGCTGCGAGTTATGCGGACTTTTTCAGCCCCGACG
TGAGGCAATACTACGCAGATCAGTACCAGCTTGATAAGTTTGACACAACAACGGCTGATGTGATGCTGTGGAATGACATGAA
TGAGCCATCTGTGTTTAATGGACCCGAAGTGACAATGCTCAAGGATAATATTCACTTTGGTGGCT GGGAGCATCGTGATGTGC
ACAACCTGTACGGCCATATGCATGTTCTGGGAACATTTGAGGGTCTTGTGCGGCGCAGTGGAGGTGTTCAGAGACCCTTCATC
CTCACTCGTGGCCACTTTGCGGGCACACAGCGCTACGCAGCCATCTGGACGGGTGACAATACAGCCGAATGGGGACATCTCC
AGGCATCCATTAAGATGTGCCTGTCTGAGTCAGTTGCTGGCTTCTCATTCTGTGGCGCTGACGTTGGTGGCTT CTTTGGGAATC
CCGATGCGGAGCTCTTTGAGCGATGGTATCAAGCAGCTGCCTTCCAACCCTTCTTCCGGGCTCACTCCCACATTGACACGCGT
CGTCGTGAACCATGGCTTTGGCCTGAAGCAACCATGCTGGTTGTCCGTGATGCCCTTCGTCGTCGCTACAGCTATCTCCCCTTC
TGGTACACCCTGTTCTATGAGCACGAACGCACAGGGCGCCCTGTGATGCGTCCCCTTCTCACACATTACCCCCTCGAC AAGGA
AACCTTCACCATTGACTACGAATACCTCCTCGGGGATATCCTCCTTGTACGCCCAGTCCTCCAGCAGGGTGTCTCCAAGGTGG
ATGTTTACTTCCCGGCCGTTGATGGGAAGAAGGAAGGTGATATCTGGTACGATGTCGATGATCACCGCAAGATTACCGCAGCT
GGTTTCGAATCCGTTCCAGTGGATAACTACAAGATCCCCGTCTACCAACGTGGAGGAACAATTCTTACGAAGAAGGAACGCA
TTCGACGTGCAGCCACGCTCATGGCAAACGATCCAGTCACCTTTGTTGTGGCACTGGATAAGGAACATTCCGCCCGTGGGACG

CTGTATGTGGATGATGAGAAAACATTCGAGTACCGCAAAGGGAAATACATCTACCTCAATTTGGAATTCAAAGACAATGTTA
TCAGTTGCAG
>LLOTMP003489-E4 exon:PUTATIVE_protein_coding
GAAAAATGATGAGAAGGCCAATTATGATACGAAATCATGGGTGGAACGCATTGTGATTGTCGGTCTGGAGCACGTACCGAAA
TCTGCAACATTGAATGTGAGTGGAGAGCCATCGGTGACACTTGAGGTCTACAAACACGGCGAGAGTGTCGTTGTACGGAAGC
CGAAGGTTGTTTTGTCACGGAAGTGGTCAATTCAGCTCAACTATTAG
>LLOTMP003489-Intron 1:PUTATIVE_protein_coding
GTAAGAATTTGTTGATTTAAAAGGAGTGCCAAAGAGAAAATAACAAAATCATTATTTTTGCAG
>LLOTMP003489-Intron 2:PUTATIVE_protein_coding
GTGAGAATAACTTTATTAACTTTAATTTTTCTCTTGAATTTGACGGGAAAATAATCTCTTTCTAAG
>LLOTMP003489-Intron 3:PUTATIVE_protein_coding
GTGAGTTTTTATTTATTTTTATTTATTGGGAGACATTGAAAAAAGAAGAACTTTCCACGTTTTTTATCTTATAATTCATCAGTTT
TAAGATTTTTGGTACCTATTTTAGAAAATATTTCAAGAATTGGACAGTTTGTTTCCTATTTTAATCGTGTTTTTAAAGGTAAAG
ATTTTTGAAAAAAAATTAATAAAATTATTTTTTGGAGTTCCCAGTAAACAAAATTTGCTCAATATTGAGTCAATGTATATAGG
TCTACATAGATGGTCAATTATTGATCAATTTAATTGATAGCTAATAATTGTTCAAATATTGATCGATTACTATCAAATCAGTGC
TGTGCTGCGCATGGAAACTGATTGAGAAGCAATTGATTAAAAAAAAAATGGAGCAATTTATAAAAGAACGAGGTGAATATTTG
TATAGAAAAGAAGGAACGACGATCACCAACAGACCTTTCTTTATTCATTATATAGATAAATCATAAATCATCTCTTCTATATA
ATAAAGAAAGGTCTGTTTGTTGGTAATCGTCGTTCCTTCTTTTCTATACAAATCCACACCGTTTGACCGATCGCG ATGAAATTT
GGTACTCCCCGCGAAGCTTTATGATTTCGTAAATCATAAAACATAAATCATAAATCATAAATCATAAAACATAAATCATAAAT
TATTAAACATAAATCATAAAACATAAATCATAAATTATAAAACATAATTCATAAAACATAAAACATAAATCATAAAACATAA
ATCATAAAACATAACTCATAAAACATAAATCATAAAACATAAATCATAAAACATTAATCATAAAACATAAATCATAAATCAT
AAATTATTAAACATAAATCATAAAACATTAATCATAAAACATTAATCATAAAACATTAATCATAAATCATAAATCATAAATCA
TAAATTATTAAACATAAATCATAAAACATAAAACATAAATCATAAATCATAAATTATAAAACATAAAACATAAATCATAAAT
CATAAATTATAAAACATAAAACATAAAACATAAATCATAAAACATAACTCATAAAACATAACTCATAAAACATAAATCATAA
AACATAAAACATAAATCATAAAACATAAATCATAAAACATAAAACTTCGCTCCTTCGCGGGTACTTGTTGCGCTTCGCGCTTT
TAGAGAAAAAGTTAATTTGTGATGGAGGTAGATTAGGGCCACTTATCAATGCCGGAAAAAAATTTTGCAAACAGAAGAAATA
ATTTTCATTGAACATTGCCCTTCCTGATCCTATTAGGGATTGTAAAACATAGTCTTGTATTGTGGACTCAATATTTAAAATATA
ACTACCTATACAAAAAAAAAAACATTTTCAAATTCGAGTCTTTGGGTCAGCGTATGACTCAATGAACCTTAAAGGGGACCTTC
ATTTAAATATGTTTCATTTTATGCAGAAAACTATTATTACAATACATTACAGCATACTTTATTGTTTTTACCGACGTGATTATTC
AAATTCAGTTTTCCGGGTTCTTTGTCGGATCGGTATGGATGGAGATTAAAACACCAGAGTXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXTAAAACATAAGCATAAAACATAAATCATAAAACATAAAACATAAATCATAAA
ACATAAATCATAAAATCAAAAATACATAAAACATAAATCAAAAATCAAAAATCATAAATCATAAAACATAAATCATAAAAC
ATAAATCATAAATCATAAATCATAAAACATAAATCATAAATCATAAATCATAAATCATAAAACATAAATCATAAAACATAAA
TCATAAAACATAAATCATAAATCAAAAATCATAAATCATAAAACATAAATCAAAAATCAAAAATCAAAAATCAAAAATCATA
AAACATAAAACATAAATCATAAAACATAAATCATAAAACAAAACATAAATCATAAAACATAAAACATAAAACATAATACAT
AAAACATAATACATAAAACATAATACATAAATCATAAATCATAAAACATAAATCATAAAACATAAATCATAAAACATAAATA
ATAAAACATAAATAATAAAACATAAATCATAAAACATAAATCATAAAACATAAAACTTCGCTCCTTCGCGGGTACCTGTTGC
GCTTCGCGCTTTTAGAGAAAAAATTAATTTGTGATGGAGGTAGATTTGGGCCACTTATCAATCCCGGAAAAAAATTTTGCAAA
CAGAAGAAATAATTTTCATTGAACATTGCCCTTCCTGATTCTATGAGTCCTTGGGGATATCCTTTTGTGATCATAAAACAATTA
GGGATTGTAAAACATAGTCTTGTGGGGGTACCCCCGCTGGCCCCAGGGGTGGGTTTAGCTAATTAAAACACTTATTTATATAG
ATAATTCACTCTGCCTCTTTATTGTGCACTCTACACATAACCCTTCACCTGACACATCTATCTATCTCCCTCACTCCCAACACCC
AGATCACCACATCTTGTATTGTGGACTCAATATTTAAAATATAACTACCTATACAAAAAAAAAACATTTTCAAATTCGAGTCT
TTGGGTCAGCGTATGACTCAATGAACCTTAAAGGGGACCTTCATTTAAATATGTTTCATTTTATGCAGAAAACTATTATTACA
ATACATTACAGCATACATGTTAATGAGACATACATAAAAAATAAGAGAGATTTGATGATCTCCCCCTCTTATTTTTTTTCCTTG
TTCTTTCACAG

TRABALHO 4

Development of *Leishmania mexicana* in *Lutzomyia longipalpis* in the absence of sugar feeding

Samara G. da Costa, Caroline da Silva Moraes, Paul Bates, Rod Dillon and Fernando Ariel Genta

Submetido ao periódico Memórias do Instituto Oswaldo Cruz

Development of *Leishmania mexicana* in *Lutzomyia longipalpis* in the absence of sugar feeding

Journal:	*Memórias do Instituto Oswaldo Cruz*
Manuscript ID	MIOC-2018-0482.R1
Manuscript Type:	Short Communication
Date Submitted by the Author:	n/a
Complete List of Authors:	da Costa, Samara; Instituto Oswaldo Cruz, Laboratório de Bioquímica e Fisiologia de Insetos Moraes, Caroline; Instituto Oswaldo Cruz, Laboratório de Bioquímica e Fisiologia de Insetos Bates, Paul; Lancaster University, United Kingdom, Department: Biomedical and Life Sciences Dillon, Roderick; Lancaster University, Biomedical and Life Sciences Genta, Fernando; Instituto Oswaldo Cruz, Laboratório de Bioquímica e Fisiologia de Insetos
Keyword:	Leishmania mexicana, Lutzomyia longipalpis, Sugar
Theme:	Leishmaniasis, Parasitology, Tropical Diseases
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Supplementary_1.mp4	

Development of *Leishmania mexicana* in *Lutzomyia longipalpis* in the absence of sugar feeding

Samara G. da Costa[1], Caroline da Silva Moraes[1], Paul Bates[2], Rod Dillon[2,3] and Fernando Ariel Genta[1,3,*]

[1]Laboratory of Insect Biochemistry and Physiology, Oswaldo Cruz Institute, FIOCRUZ, Rio de Janeiro, RJ, Brazil

[2]Faculty of Health and Medicine, Division of Biomedical and Life Sciences, Lancaster University, Lancaster, United Kingdom

[3]National Institute of Science and Technology for Molecular Entomology, Rio de Janeiro, RJ, Brazil

*Corresponding Author: *Laboratory of Insect Biochemistry and Physiology, Oswaldo Cruz Institute, FIOCRUZ, Rio de Janeiro, Av Brasil 4365, Pav. Leônidas Deane, room 207. Manguinhos. RJ, CEP21040360, Brazil.* E-mail address: genta@ioc.fiocruz.br or gentafernando@gmail.com

Abstract

Leishmaniasis is a neglected disease which affects thousands of people annually. This disease is caused by *Leishmania* parasites and transmitted through the bite of sand flies. During the parasite development inside the midgut of the vector, promastigotes move toward the anterior region of the midgut, the cardia. It has been reported that the sugar meal acquired by sand flies during feeding is essential for the development of parasites in the insect midgut and that the sugar gradient orientates the migration of parasites to cardia region by chemotaxis and osmotaxis. We demonstrated that the distribution of *L. mexicana* parasites depends on the sugar meals obtained by the phlebotomines. Besides that, sugar meals are essential to supply the energy requirements for the survival of sand flies, especially during blood digestion. The promastigotes migration towards the cardia region, a mechanism crucial for transmission, seems to be only partially based on the chemotaxis stimuli by sugar molecules. In the absence of sugars, significant amounts of parasites develop in the hindgut. As some parasites migrate to the cardia, it is possible that other components inside the vector gut create the chemical ambient to provide an orientation stimulus for *Leishmania* migration.

Keywords: ***Leishmania mexicana, Lutzomyia longipalpis*, sugar**

Sponsorship: This research was supported by FAPERJ, CNPq, CAPES, and FIOCRUZ (Brazil) research agencies. SGC received a sandwich Ph.D. grant from Science without Borders (CNPq/CAPES). CSM is a Postdoc from PNPD (CAPES/FAPERJ) and received a grant from Science without Borders (CNPq/CAPES).

Introduction

The *Leishmania* parasites develop as an extracellular form (promastigotes) in the gut of their sand fly vectors and as obligate intracellular form (amastigotes) inside phagolysosomes of infected macrophages in the vertebrate host. The development of *Leishmania* parasite inside the vector is complex and dynamic. Depending on *Leishmania* subgenera a different pattern of development can be observed inside the gut of the vector, for *Leishmania mexicana* (subgenera *Leishmania*) parasites develop exclusively in the midgut and foregut of their vectors characterizing a suprapylarian development (1).

After ingestion of an infective blood meal by the sand fly, macrophages containing parasites release the amastigotes forms in the blood meal, and the change in pH conditions triggers the transformation of amastigotes parasites into promastigotes (2), a motile and replicative form. The parasites have different developmental stages inside the gut of the vector. For infection establishment, two cycles of multiplication occur during parasites development. The first cycle occurs with the multiplication of procyclic promastigotes inside the peritrophic matrix, in the blood meal phase (3). After the digestion of blood, parasites escape from the peritrophic matrix, attach to the midgut epithelium and migrate

to the anterior midgut region (4). The second cycle of multiplication takes place in the sugar meal phase with the leptomonads, which differentiate in the non-multiplicative infective promastigote metacyclic forms. It is hypothesized that the presence of sugar ingested by females sand fly triggers the multiplication of leptomonads (3).

Between infectious blood meal feeds, sand flies take sucrose-rich meals that are stored in the crop (5). The sugar meal is then released in small quantities to the midgut. After blood meal digestion, the sugar meal rich in sucrose, raffinose, maltose, melezitose, starch, and cellulose (besides other types of sugars) is a source of nutrition for parasites development inside vector gut. It is believed that the sugar ingestion by the vector impacts the developing of promastigote parasite population (6–8). It was described for different *Leishmania* species the secretion of glycosidases, enzymes specialized in the digestion of sugars, like alpha-glucosidase, sucrases, invertases, alpha-amylases, besides others (8–11). For *L. mexicana*, both invertase and sucrase activity were identified as secreted by promastigotes (7,10). In this respect, *L. mexicana* might use the sugar meals as an exogenous source of energy for its development.

Besides this, sugar ingestion by females sand flies create a sugar gradient along the midgut, and it was described that this gradient provides the stimulus for parasite migration toward stomodeal valve region (critical for efficient transmission) by mechanisms of chemo and osmotaxis (12–15). However, studies about the effects of the sugar meal on parasites migration and development using an in vivo model are scarce.

In this work, we demonstrated that the distribution of *Leishmania mexicana* along the gut of *L. longipalpis* is dependent on sugar feeding by phlebotomines. In the absence of sugar meals, though the parasites are capable of reaching the stomodeal valve region, a significant amount of parasites intensely develop in the hindgut of the insect. Although the sugar feeding is not necessary for the complete development of parasites, the *L. longipalpis* survival is drastically affected by the absence of sugar feeding, especially after blood-feeding. In this respect, we emphasize the importance of sugar meals during the life cycle of both sand fly vectors and *Leishmania* parasites.

Material and Methods

Insects Maintenance

Insectary-reared *L. longipalpis* (Jacobina, Bahia, Brazil), maintained at Lancaster University (United Kingdom), were used for experiments. Insects were kept under standard laboratory conditions of temperature (24 ± 2 °C) and a photoperiod of 8 h of

light/16 h of darkness. Adult sand flies were fed with 70 % (w/v) autoclaved sucrose solution in cotton wool (unless stated differently in experiments). For oviposition, females were fed with sheep blood containing Alsever's anticoagulant (Cat: SB068, TCS Biosciences, Buckingham, United Kingdom), via an artificial apparatus (Hemotek - Discovery Workshops, United Kingdom) at 37 °C for 1 hour using chicken skin membranes. Engorged females were transferred to rearing containers with a piece of cotton wool soaked in sugar solution. Eggs separation from dead females was conducted after oviposition. Other details for sand fly colony maintenance were as described in (16).

For experiments, recently emerged females (0 - 3 hours) were separated into small cages and subjected to different feeding conditions: water (non fed) or 1.2 M sucrose, followed by blood feeding or infected blood feeding using the Hemotek as described below.

Leishmania mexicana Cultivation and Sand flies Infection

L. mexicana (World Health Organization strain MNYC/BZ/1962/M379) from an axenic culture of amastigote-like forms was used for infections. Amastigote-like culture and sand fly infections were performed as described in (17). For infections, a concentration of 2 x 10^6 parasites/mL, estimated with Neubauer chambers, was used. Briefly, after centrifugation at 2000 x g for 5 min, the supernatant was removed, and parasites were mixed with sheep blood and offered to 3-day old females maintained with water (unfed). After blood feeding, unfed females were discarded, and the fed ones were kept with water only. Control insects were maintained throughout the experiment with 1.2 M sucrose.

Estimation of *Leishmania mexicana* infection numbers

The whole gut of infected females was dissected and analyzed under light microscopy at 3, 6 and 10 days after blood feeding to check the infection establishment. Dissections were conducted in PBS on microscope slides using needles. Dissected guts were transferred to polypropylene tubes containing 20 µL PBS and 2 % paraformaldehyde, used to slow down parasites. After homogenization and proper dilution, 10 µL sample was transferred to Neubauer chambers, and the total number of parasites were counted according to manufacturers. We also analyzed the number of metacyclic promastigotes on the sixth day's samples. The identification of metacyclic promastigotes followed the characteristics described for the identification of *Leishmania* different developmental

stages (18). On the third and sixth days after the blood feeding, the number of parasites was also estimated in the hindgut and midgut, separately, using the same procedure described above.

Survival

The longevity of *L. longipalpis* was estimated under different conditions. Mortality was evaluated, and dead insects were removed from cages daily. For each biological replicate and condition tested, 100 females were used. As specified above, emerged females (0-3 hours) were separated, and six different feeding conditions were monitored. The following females were analyzed: nonfed maintained with water, fed on 1.2 M sucrose (SF), blood-fed (infective meal or not) maintained with 1.2 M sucrose or blood-fed (infective meal or not) maintained with water.

Statistical Analysis

All statistical analysis was performed using GraphPad Prism 6.0 for Windows (San Diego, California, USA), and the D'Agostino-Pearson Omnibus K2 normality test was used. The outliers were identified with the ROUT method, and Q was established as 1 %. One-way ANOVA (multiple comparisons) followed by Tukey's multiple comparison tests and significance was considered when $p< 0.05$. For survival, results were analyzed using the Kaplan-Meier survival curve obtained with the GraphPad Prism 6.0 for Windows (San Diego, California, USA) and thus the average survival time was determined in each condition. The log-rank Mantel-Cox test was used to compare survival curves. Significance was considered when $p < 0.05$.

Results

Effect of sugar feeding on *Leishmania mexicana* infection

The results demonstrated that the number of parasites present in the whole gut was not affected by the presence of sucrose. Comparisons were performed three and six days after infection (Fig 1). Due to the high mortality, it was not possible to evaluate infection at ten days in water-fed females. Furthermore, the number of parasites did not increase following the days after infection for either water or sucrose fed females. After six days, we also analyzed the number of *Leishmania* metacyclic forms in the midgut (including cardia), and no significant difference was detected in the numbers when comparing water

fed to 1.2 M sucrose fed females, with 1200 ± 200 and 1400 ± 200 metacyclic per midgut, respectively.

Although no differences were found in the total number of parasites after different feeding regimes, we observed a difference in the pattern of parasite distribution. In females fed with water after three and six days of infection, we found a large amount of parasites in the hindgut. In Fig 2, we present images obtained by light microscopy demonstrating the presence of parasites in the hindgut of water-fed females (Figs 2A and 2B), with a large number of parasites in this region. The hindgut of these insects is filled with parasites, but these seem not attached to the cuticle (supplementary video 1). In water-fed females, parasites were also present in the midgut and cardia region (Fig 2C), while for sugar-fed females parasites are not distributed along the hindgut (Fig 2D).

Considering the migration of parasites to the hindgut, we evaluated the number of parasites in this compartment in water and sucrose-fed females. The percentage of infected females presenting parasites in the hindgut was larger for water-maintained, in this case, after three days post-blood feeding: 70 % of analyzed insects had parasites in the hindgut against 20 % in sugar-maintained (Fig 3A). Six days after blood-feeding almost 90 % of water-maintained females had parasites in the hindgut (Fig 3A). For midgut, the number of parasites (Fig 3B) presented the same pattern demonstrated in Fig 1, in both water and sugar-maintained insects a massive number of parasites concentrates in this region. On the other hand, the number of parasites in the hindgut of water-fed females was significantly higher compared to sucrose-fed females (Fig 3C), even though, the parasites number in the hindgut did not increase from 3 to 6 days. Data obtained in these assays suggest that the number of parasites quantified in the hindgut, compared to what we can observe in the images (Fig 2B) was underestimated, due to the limitation of the technique of rupture of gut for separation of midgut and hindgut.

L. longipalpis survival in different feeding conditions

The median survival was drastically reduced from 25 days to 5 days, for sucrose fed females compared to starving (Figs 4A and 4B). The results also demonstrate that blood-feeding affects the survival but not the presence of *L. mexicana* parasite. In sugar-fed females, the mean survival was reduced from 25 to 7 days, almost 70 % reduction, in blood-fed females (infected or not infected) compared to the non-blood fed ones. For

water-maintained females, the median survival was reduced from 5 to 3 days, in blood-fed females (infected or not infected), compared to the non-blood fed females.

Discussion

In this work, we demonstrate that *L. mexicana* is capable of developing inside *L. longipalpis*, even in the absence of sugar feeding by the phlebotomine. There was no significant difference comparing the total number of parasites inside the gut, or the number of metacyclic forms, in water or sugar-fed females. Previous works discuss the importance of sugar feeding by phlebotomine for parasites development (6–8,10). According to them, *Leishmania* promastigotes depend on the diet of the phlebotomines to sustain their growth. Sugars may also prevent the egestion of *Leishmania* during defecation of blood meal remnants (19). In our work, we demonstrated the multiplication and development of promastigote and metacyclic forms in the gut of water-fed flies, although the infectivity of these parasites in a second blood meal is a question that remains to be addressed. Besides that, we do not know if the absence of sugar can affect the appearance of different promastigotes forms during development. The development of *L. mexicana* parasites into different promastigotes forms was described as sugar dependent (3). Protocols for studying the effect of a second blood meal in *L. longipalpis* upon laboratory conditions had been recently developed (20) and it has been recently demonstrated that the ingestion of a second non-infecting blood meal by infected sand flies leads to enhanced disease transmission by amplifying the number of parasites acquired in the infected blood meal. The process occurs by dedifferentiation of the metacyclic promastigotes into the replicative retroleptomonad promastigote form, which leads to increased infection (21). Although components present in plasma triggers dedifferentiation, we cannot rule out the hypothesis that sugar absence might have an effect in the appearance of promastigotes retroleptomonad form or affect the numbers of parasites in a case of a second blood meal. So, we expect to extend these observations in the future.

Our data suggest that nutrients obtained from sugar meals are not strictly necessary for growing and differentiation, as the parasites might nourish from nutrients released from blood hydrolysis in the absence of a sugar meal. The nutrients obtained from blood are especially necessary for the early phases of development when the parasite is trapped inside the peritrophic matrix. In the case of *L. longipalpis*, a soluble α-glucosidase is induced during blood feeding and might be acting in the hydrolysis of small

glycoconjugates released from blood (22), and probably *L. mexicana* might be able to absorb these products released by the enzyme secreted by the vector. Differently from the results demonstrated in our work, in infections of *L. longipalpis* with *L. donovani*, a regular sugar meal was shown to enhance the number of parasites inside the gut of the vector (23) and also for *L. youngi*, the efficiency of infection with *L. amazonensis* was affected by the type of sugar used to feed the sand flies (6). It is possible that, in our conditions, the presence or absence of sugar meals impacts the development and survival of parasites only in longer periods after the blood feed, and further studied must confirm or reject this hypothesis.

During the development of parasites inside the gut of the vector, the movement of promastigotes to the anterior region of the sand fly midgut, with the accumulation of metacyclic promastigotes in the stomodeal valve, is critical, causing a rupture in the valve and transmission to a mammal when a next blood-feeding occurs (2,18). Taxis is a phenomenon where an organism responds to specific stimuli by movement. It was proposed that during development the promastigotes could be attracted by the sugar meals ingested by sand flies, then migrating to the anterior region of the midgut (24). Some works described that *Leishmania* promastigotes undergo chemotaxis in a gradient constituted of different sugars (13,15) and likewise by serum albumin, hemoglobin, besides others (15). The movement is also due to the osmotic gradient generated by the presence of sugars (12). For *L. amazonensis* it was demonstrated that the parasite was able to respond both to chemotaxis and osmotaxis stimuli (14). In this respect, both mechanisms of chemotaxis and osmotaxis play a role in the direction of parasites to the stomodeal valve region. We demonstrated that parasites are more frequently found in the hindgut (not attached) of water-fed compared to sugar-fed females. However, in both conditions, a significantly high number of parasites were also able to reach cardia. In this respect, the presence of sugar, creating an osmotic and chemical gradient, is important but not obligatory to direct the migration of parasites toward the cardia region. In an ambient where the sugar concentration is much higher than the other components, it might function as the central stimulator for parasites migration. Although in an ambient where a large quantity of sugars is not present, the movement toward cardia might be explained by the water flow or by presence of other components inside the vector gut that might also create an orientation stimulus for parasite migration. The midgut of the vector is divided into specialized regions with a variety of chemical and structural features that *Leishmania* parasites might exploit for orientation. Some studies have demonstrated that

chemotaxis in *Leishmania* could be elicited by a wide range of compounds (13,15), so saliva components might also work as taxic agents. It was proposed that the receptors involved in chemotaxis possess low specificity and a wide range of affinity, the same receptor might be able to bind structurally related molecules (14).

Leishmania parasites may be classified in suprapylarian, peripylarian or hipopylarian, based on the preferential region of their development along the gut of the sand fly vector (1). *Leishmania* species that develop in the gut regions before the pylorus are considered suprapylarian, and belong to the subgenera *Leishmania*. *Leishmania* species that colonize the abdominal gut regions, around the pylorus, are named peripylarian, and belong to the New World *Viannia* subgenera. *Leishmania* species that develop mainly in the hindgut are named hypopylarian, belong to the subgenera *Sauroleishmania* and infect reptiles. Interestingly, our data suggest that the distribution of *Leishmania* parasites along the gut of sand flies depends on the sugar meal of the vector, as in our conditions, *L. mexicana*, a suprapylarian parasite from the subgenera *Leishmania*, shows considerable development in the hindgut in the absence of sugars in the phlebotomine diet.

Although parasite does not seem to require sugar for development in our conditions, sugar is essential for phlebotomine survival. Without sugar meals, the mortality of sand flies is drastically enhanced, especially when females are blood-fed. The sugar feeding is vital to the metabolic demands in the phlebotomines. The glucose, for example, obtained from sugar hydrolysis is uptake by enterocytes, and converted to trehalose or stored as glycogen to supply the energetic demands of insects, like flight. In a starving condition, the reserves of glycogen and triglycerides are mobilized (25,26). During blood digestion, nutrients as heme and amino acids are present in excess, and these molecules need to be detoxified by disposal or converted to advantageous derivatives. The release of heme is toxic because it potentiates oxygen-reactive species and can permeate membranes (27).

Moreover, there is an enhanced microbial growth after blood feeding that needs to be controlled (28). Briefly, we suggest that during the blood digestion, there is an energetic demand to maintain the homeostasis in the organism. In a starving phlebotomine, weakened by the lack of energy, the hazardous effects of molecules or pathogens increased during blood digestion is enhanced, and the pathways used for detoxification of these compounds or control of pathogens might be restricted.

In this respect, according to the results reported here, the development and migration of *L. mexicana* towards the cardia region, a mechanism essential for transmission, is not utterly dependent on sugar feeding by phlebotomines, but the sugar meals are necessary

to supply the energy requirements for the survival of sand flies, especially during blood digestion. The survival of phlebotomine for an extended time is crucial for *Leishmania* transmission to the mammalian host since two blood feedings are necessary for this. Thus, even with the viable development of the parasites in the absence of sugar, the transmission cycle might not occur, because the sand flies do not survive long enough to perform two blood feeds.

Acknowledgments

Authors thank Michelle Bates from Lancaster University for technical assistance.

References

1. Lainson R, Shaw JJ. Evolution, classification and geographical distribution. In: Peters W, Killick-Kendrick R, editors. Leishmaniasis in Biology and Medicine. London: Academic Press; 1987. p. 1–120.
2. Bates P, Rogers M. New Insights into the Developmental Biology and Transmission Mechanisms of Leishmania. Curr Mol Med. 2004;4(6):601–9.
3. Gossage SM, Rogers ME, Bates PA. Two separate growth phases during the development of Leishmania in sand flies: Implications for understanding the life cycle. Int J Parasitol. 2003;33(10):1027–34.
4. Kamhawi S. Phlebotomine sand flies and Leishmania parasites: friends or foes? Trends Parasitol. 2006;22(9):439–45.
5. Jacobson RL, Studentsky L, Schlein Y. Glycolytic and chitinolytic activities of Phlebotomus papatasi (Diptera: Psychodidae) from diverse ecological habitats. Folia Parasitol (Praha). 2007;54(4):301–9.
6. Añez N, Nieves E, Cazorla D. The validity of the developmental pattern in the sandfly gut for classification of leishmania. Trans R Soc Trop Med Hyg. 1989;83(5):634–5.
7. Gontijo NF, Melo MN, Riani EB, Almeida-Silva S, Mares-Guia ML. Glycosidases

in Leishmania and Their Importance for Leishmania in Phlebotomine Sandflies with Special Reference to Purification and Characterization of a Sucrase. Exp Parasitol. 1996;83:117–24.

8. Jacobson RL, Schlein Y, Eisenberger CL. The biological function of sand fly and Leishmania glycosidases. Med Microbiol Immunol. 2001;(190):51–5.

9. BLUM JJ, OPPERDOES FR. Secretion of Sucrase by Leishmania donovani. J Eukaryot Microbiol. 1994;41(3):228–31.

10. Lyda TA, Joshi MB, Andersen JF, Kelada AY, Owings JP, Bates PA, et al. A unique, highly conserved secretory invertase is differentially expressed by promastigote developmental forms of all species of the human pathogen, Leishmania. Mol Cell Biochem [Internet]. 2015;404(1–2):53–77. Available from: http://dx.doi.org/10.1007/s11010-015-2366-6

11. Jacobson RL, Schlein Y. Phlebotomus papatasi and Leishmania major parasites express a-amylase and a-glucosidase. Acta Trop. 2001;78:41–9.

12. Leslie G, Barrett M, Burchmore R. Leishmania mexicana: Promastigotes migrate through osmotic gradients. Exp Parasitol. 2002;102(2):117–20.

13. Oliviera JS, Melo MN, Gontijo NF. A sensitive method for assaying chemotaxic responses of Leishmania promastigotes. Exp Parasitol. 2000;96(3):187–9.

14. Barros VC, Oliveira JS, Melo MN, Gontijo NF. Leishmania amazonensis: Chemotaxic and osmotaxic responses in promastigotes and their probable role in development in the phlebotomine gut. Exp Parasitol. 2006;112(3):152–7.

15. BRAY RS. Leishmania: Chemotaxic Responses of Promastigotes and Macrophages In Vitro. J Protozool. 1983;30(2):322–9.

16. Moraes CS, Diaz-Albiter HM, Faria M do V, Sant'Anna MRV, Dillon RJ, Genta FA. Expression pattern of glycoside hydrolase genes in Lutzomyia longipalpis

355 reveals key enzymes involved in larval digestion. Front Physiol. 2014;5
356 JUL(August):1–16.
357 17. da Costa SG, Bates PA, Dillon RJ, Genta FA. Study of glycoside hydrolase
358 families 13 and 31 reveals expansion and diversification of α-amylase genes in the
359 phlebotomine Lutzomyia longipalpis, and modulation of sand fly glycosidase
360 activities by Leishmania infection. PLoS Negl Trop Dis.
361 18. Rogers ME, Chance ML, Bates PA. The role of promastigote secretory gel in the
362 origin and transmission of the infective stage of Leishmania mexicana by the
363 sandfly Lutzomyia longipalpis. Parasitology. 2002;124(5):495–507.
364 19. Warburg A, Schlein Y. The effect of post-bloodmeal nutrition of Phlebotomus
365 papatasi on the transmission of Leishmania major. Am J Trop Med Hyg.
366 1986;35(5):926–30.
367 20. Moraes CS, Aguiar-Martins K, Costa SG, Bates PA, Dillon RJ, Genta FA. Second
368 Blood Meal by Female Lutzomyia longipalpis: Enhancement by Oviposition and
369 Its Effects on Digestion, Longevity, and Leishmania Infection. Biomed Res Int.
370 2018;
371 21. Serafim TD, Coutinho-Abreu I V., Oliveira F, Meneses C, Kamhawi S, Valenzuela
372 JG. Sequential blood meals promote Leishmania replication and reverse
373 metacyclogenesis augmenting vector infectivity. Nat Microbiol [Internet]. 2018
374 May 19 [cited 2018 Jul 24];3(5):548–55. Available from:
375 http://www.nature.com/articles/s41564-018-0125-7
376 22. da Costa SG, Bates PA, Dillon RJ, Genta FA. Characterization of α-glucosidases
377 from Lutzomyia longipalpis reveals independent hydrolysis systems for plant or
378 blood sugars. Front Physiol.
379 23. SHERLOCK IA, SHERLOCK VA. Sôbre a infecção experimental de

Phlebotomus longipalpis pela Leishmania donovani. Rev Bras Biol. 1961;21:409–18.

24. Killick-Kendrick R. Recent advances and outstanding problems in the biology of phlebotomine sandflies. A review. Acta Trop [Internet]. 1978 Dec [cited 2018 Aug 28];35(4):297–313. Available from: http://www.ncbi.nlm.nih.gov/pubmed/32749

25. Thompson SN. Trehalose - The Insect "Blood" Sugar. In: Advances in Insect Physiology. Elsevier Ltd; 2003. p. 205–85.

26. Arrese EL, Soulages JL. Insect Fat Body: Energy, Metabolism, and Regulation. Annu Rev Entomol [Internet]. 2010;55(1):207–25. Available from: http://www.annualreviews.org/doi/10.1146/annurev-ento-112408-085356

27. Graça-Souza A V., Maya-Monteiro C, Paiva-Silva GO, Braz GRC, Paes MC, Sorgine MHF, et al. Adaptations against heme toxicity in blood-feeding arthropods. Insect Biochem Mol Biol. 2006;36:322–35.

28. Sterkel M, Oliveira JHM, Bottino-Rojas V, Paiva-Silva GO, Oliveira PL. The Dose Makes the Poison: Nutritional Overload Determines the Life Traits of Blood-Feeding Arthropods. Trends Parasitol [Internet]. 2017;33(8):633–44. Available from: http://dx.doi.org/10.1016/j.pt.2017.04.008

Additional files

S1 Video: Video demonstrating the hindgut of water-fed *L. longipalpis* infected with *L. mexicana* 6 days post infection.

Fig1.tif

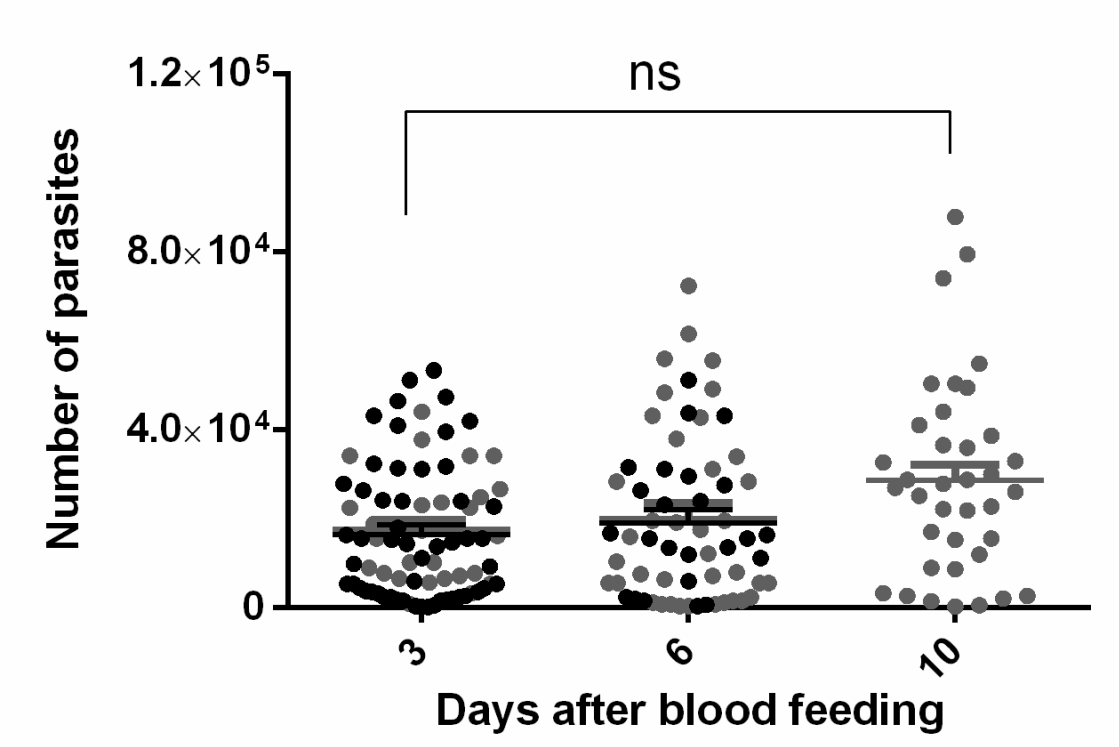


Fig 1: *L. mexicana* parasites quantification in the gut of *L. longipalpis* females at different days following the blood feeding. Recently emerged (0-3 h) females were collected and maintained with water for 3 days before feeding with infected blood meal. After feeding, insects were maintained with sucrose 1.2 M (grey circles) or water (black circles). Circles represent the number of parasites per individual gut. The results are the mean ± SEM of two independent experiments. One-way ANOVA followed by Tukey multiple comparison test was performed. **ns**: non-significative difference.

Fig2. tif

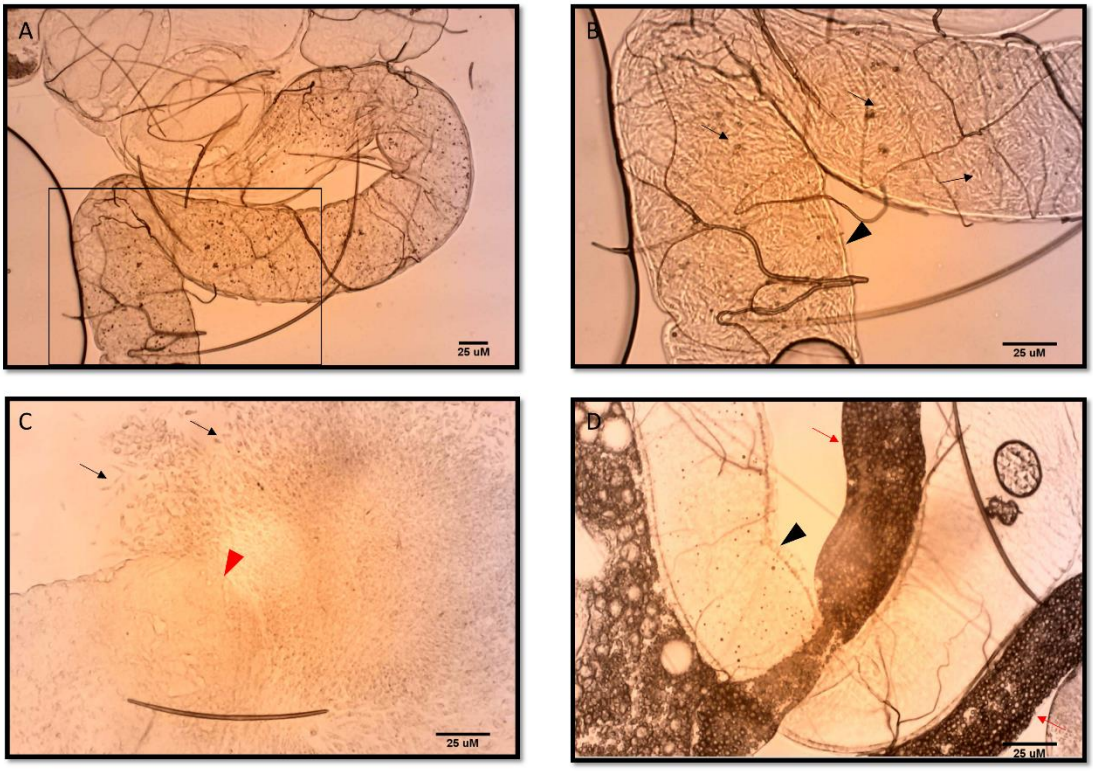


Fig 2: Light microscopy image of infected *L. longipalpis* females 6 days after blood feeding. Recently emerged (0-3 h) females were collected and maintained with water for 3 days before feeding with an infected blood meal. After feeding, insects were maintained with water (**A, B, C**) or sucrose 1.2 M (**D**). **A:** hindgut 20 X magnification of infected water-maintained females. **B:** 40 X magnification of delineated section in figure 2A. **C:** 40 X magnification of cardia of infected water-maintained females. **D:** hindgut 40 X magnification of infected sugar-maintained females. Note gut epithelium (black arrowhead), *Leishmania* parasites (black arrows), malpighian tubules (red arrows) and cardia (red arrowhead).

Fig3. tif

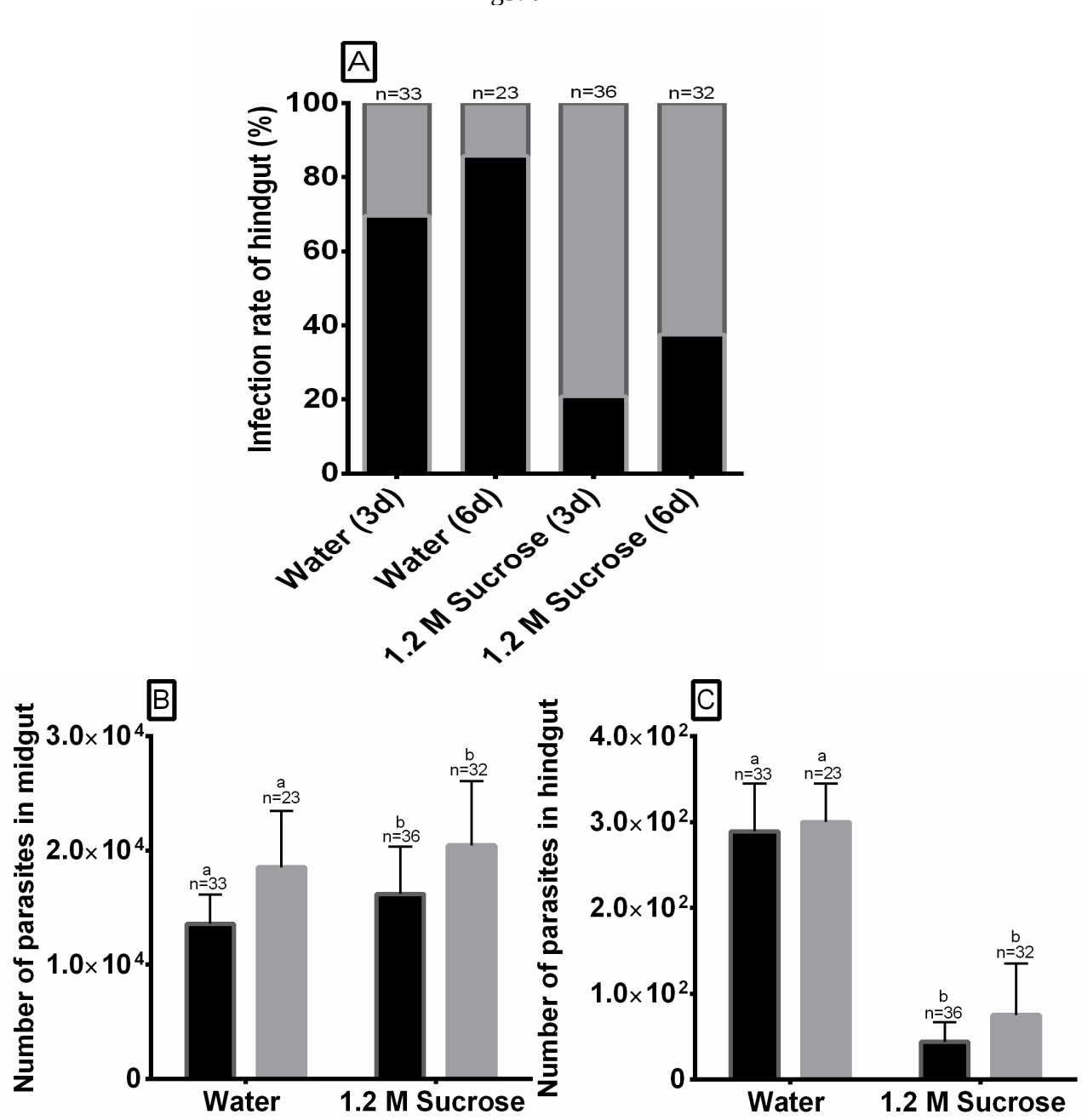


Fig 3: *L. mexicana* parasites quantification in the midgut and hindgut of *L. longipalpis* after 3 and 6 days following the blood feeding. Recently emerged (0-3 h) females were collected and maintained with water for 3 days before feeding with infected blood meal. After feeding, insects were maintained with sucrose 1.2 M or water. **A:** Infection rate of hindgut of water or sugar-maintained females. The black background indicates the percentage of positive females' samples containing parasites in the hindgut and the grey background represents the percentage of negative female samples **B:** Quantification of parasites in the midgut of water or sugar-maintained females after 3 days **(black bars)** and 6 days **(grey bars)**. **C:** Quantification of parasites in the hindgut of water or sugar-maintained females after 3 days **(black bars)** and 6 days **(grey bars)**. One-way ANOVA, followed by Tukey multiple comparison tests. Different letters indicate statistically significant differences in quantification, $p < 0.001$.

Fig4.tif

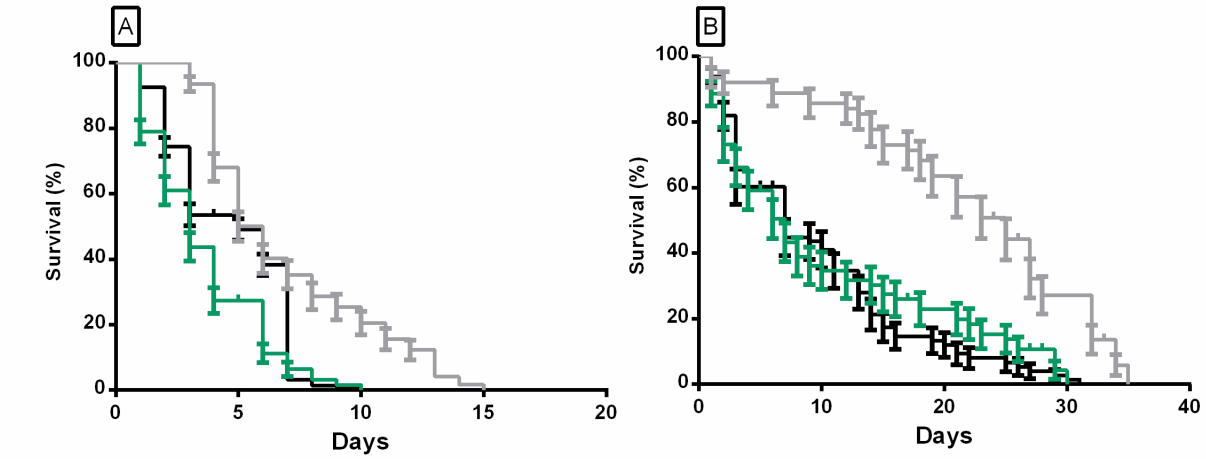


Fig 4: Survival curve of *L. longipalpis* females in different feeding conditions, maintained in controlled humidity and temperature conditions. Recently emerged (0-3 h) females were collected and maintained with water for 3 days before feeding with infected blood meal. After blood feeding (infected or not), insects were maintained with sucrose 1.2 M or water. Control groups were only fed with sucrose 1.2 M or water (no blood meal). **A:** non-fed females maintained with water **(grey line)**, blood fed females maintained with water post blood feeding **(green line)**, infective blood fed females maintained with water after infection **(black line)**. **B:** sucrose 1.2 M fed females **(grey line)**, blood fed females maintained with sucrose 1.2 M post blood feeding **(green line)**, infective blood fed females maintained with sucrose 1.2 M after infection **(black line)**. The results are representative of three independent experiments. For each replicate at least, n=100 females were used. Log rank Mantel-Cox test was performed and the survival curves were significant different at $p<0.0001$.

4. Discussão Geral

Para flebotomíneos, as demandas energéticas são supridas principalmente pela ingestão de dietas açucaradas obtidas a partir de diferentes fontes alimentares, tanto para fêmeas quanto para machos. Açúcares como sacarose, maltose, trealose e melezitose são obtidos principalmente da alimentação direta em tecidos vegetais, néctar de flores, excretas de afídeos e coccídeos, entre outros (CAMERON et al., 1995; KILLICK-KENDRICK; KILLICK-KENDRICK, 1987; MACVICKER et al., 1990). Diferentes espécies de flebotomíneos se desenvolvem em regiões com maior ou menor abundância ou diversidade dessas fontes alimentares. Flebotomíneos do gênero *Phlebotomus* (Velho Mundo) se desenvolvem em regiões de deserto e savana, enquanto insetos do gênero *Lutzomyia* (Novo Mundo) se desenvolvem em áreas úmidas perto de florestas (AKHOUNDI et al., 2016). Assim, diferentes fontes de açúcar podem ser utilizadas para alimentação. De maneira geral, a sacarose e o amido são os principais produtos encontrados nas folhas de plantas (JACOBSON; SCHLEIN, 2001). O amido é hidrolisado pela ação de endoamilases (principalmente a α-amilase), exoamilases e enzimas desramificadoras, liberando os resíduos de maltose que são clivados em monômeros de glicose pela ação das α-glicosidases (VAN DER MAAREL et al., 2002). Do mesmo modo, a sacarose obtida da seiva ou do néctar pode ser hidrolisada por α-glicosidases, liberando monômeros de glicose e frutose.

As fêmeas de flebotomíneos também possuem uma dieta rica em proteínas pois se alimentam de sangue para maturação de seus ovócitos. Portanto, a dieta dos flebotomíneos é rica em carboidratos e proteínas. Os carboidratos também podem ser obtidos a partir da digestão dos glicoderivados do sangue, como glicoproteínas e glicolipídios. Assim, as glicosidases desempenham um papel crucial na digestão desses insetos, especialmente as α-amilases e as α-glicosidases. Nos insetos, essas enzimas são descritas nas famílias 13 e 31 (CANTAREL et al., 2009).

Após a hidrólise dos açúcares, a glicose obtida dos alimentos é absorvida pela células intestinais através de transportadores de membrana específicos. A glicose pode entrar em diferentes vias metabólicas para produção de energia, respiração celular, ainda pode ser transformada em trealose, o açúcar presente na hemolinfa em insetos, ou armazenada como glicogênio ou gordura nas células (ARRESE; SOULAGES, 2010). Os insetos precisam de energia continuamente e, em condições de jejum devem sobreviver com as reservas energéticas (glicogênio e triglicerídeos) acumuladas em períodos de abundância de alimento. Além disso, em atividades como o vôo, por exemplo, condição de extrema demanda energética, as reservas são rapidamente mobilizadas (DOUGLAS, 2013; KLOWDEN, 2013). Algumas das enzimas

envolvidas no processo de armazenamento e mobilização de glicogênio, como a enzima ramificadora de α-1,4-glucana (GBE), a enzima desramificadora de glicogênio (GDE) e a α-glicosidase lisossomal também são descritas nas famílias 13 e 31 em insetos.

No presente trabalho, caracterizamos bioquimicamente as atividades de α-glicosidases em diferentes tecidos e diferentes contextos fisiológicos utilizando um substrato sintético, MUαGlu, e o substrato natural sacarose. Investigamos também a presença de genes codificando para proteínas pertencentes à família da α-amilase GH13 e da família GH31 no genoma de *L. longipalpis*, a partir de dados genômicos da plataforma "Vector Base". Com as informações obtidas foi possível selecionar e avaliar a expressão dos genes em diferentes contextos fisiológicos.

Um dos grandes desafios deste trabalho e também no estudo bioquímico utilizando insetos de pequeno porte, é o estudo da atividade e especificidade de diferentes enzimas, principalmente a caracterização em diferentes tecidos e por indivíduo, devido à pequena quantidade de material disponível. A atividade de glicosidases normalmente é determinada em ensaios descontínuos, com diferentes métodos de detecção dos produtos liberados ao longo do tempo. Os protocolos descontínuos resultam no uso de uma grande quantidade de amostras, para obtenção dos pontos em cada tempo de detecção e réplica experimental. Em casos em que atividade enzimática sobre o substrato estudado é baixa, ensaios longos e laboriosos são utilizados, e devido à pequena quantidade de material disponível para análise, estas necessitam serem feitas a partir de pools de amostras.

Como demonstramos, a atividade de α-glicosidase inicialmente foi detectada em todos os tecidos estudados de *L. longipalpis*, cabeça, carcaça, divertículo, intestino médio e intestino posterior. Atividades maiores foram encontradas para o substrato natural sacarose, com valores mais de 1000 vezes superiores em comparação com o substrato sintético fluorescente MUαGlu. Dessa forma, a padronização de ensaios contínuos de glicosidase utilizando o substrato sintético fluorescente, permitiu neste trabalho a quantificação da atividade de α-glicosidase utilizando o substrato fluorescente em amostras individuais, utilizando uma menor quantidade de amostra por ensaio, o que permitiu um maior número de quantificações por amostra e resultados mais confiáveis. A padronização destes ensaios contínuos abre caminhos para o estudo de atividades enzimáticas utilizando substratos fluorescentes em outros insetos e também em diferentes fases de desenvolvimento, por exemplo em insetos que apresentam estágios larvais, facilitando a execução das análises devido à menor quantidade de amostra biológica. Além disso, possibilita a quantificação de atividades em amostras individuais, fornecendo dados de distribuição e frequência das atividades na população. Esta ferramenta também pode auxiliar os estudos de

silenciamento destas enzimas por RNAi, permitindo a análise de insetos silenciados individualmente.

Como descrito no Trabalho 2, a atividade das α-glicosidases foi quantificada nos diferentes tecidos de *L. longipalpis*. Nesse trabalho identificamos atividade no divertículo, descrita em *L. longipalpis* pela primeira vez, esta atividade pode sugerir a presença de uma α-glicosidase salivar ou estar apenas relacionada a uma atividade intracelular. A maior parte da atividade de α-glicosidase está concentrada no intestino médio e foi encontrada na forma solúvel e associada à membrana em fêmeas não alimentadas, alimentadas com açúcar e alimentadas com sangue. Os dados de anotação do genoma sugeriram (Trabalho 3) que as maltases (prováveis α-glicosidases) são proteínas ancoradas à membrana do intestino médio através de domínios transmembranas ou âncora GPI. Essas sequências são mais expressas no intestino médio em relação ao resto do corpo, o que está em concordância com os resultados de quantificação de atividade. Trabalhos anteriores (GONTIJO et al., 1998) caracterizaram uma α-glicosidase ligada à membrana das microvilosidades epiteliais do intestino médio em fêmeas de *L. longipalpis* não alimentadas, mas no entanto a atividade presente no divertículo não foi descrita. A diferença encontrada em nosso trabalho, quando comparado ao trabalho citado (GONTIJO et al., 1998), provavelmente se baseia no tipo de substrato e na concentração final usada nos ensaios. Quando sacarose foi utilizada como substrato de α-glicosidase, encontramos um alto valor de $\boldsymbol{K_m}$ (11 – 800 mM). Portanto, para saturar os sítios de ligação ao substrato e obter uma determinação da atividade proporcional à quantidade da enzima presente, foi necessário utilizar uma alta concentração final de substrato (sacarose) nos ensaios. Inicialmente, quando utilizamos uma concentração final de sacarose menor (dados não mostrados), não foi possível identificar atividade no divertículo e no intestino posterior. Para a espécie *Phlebotomus papatasi*, a atividade de α-glicosidase foi identificada no intestino e glândulas salivares de insetos não alimentados. Com isso, foi sugerido que, após a ingestão dos açúcares, as α-glicosidases salivares se misturam à alimentação açucarada no divertículo (JACOBSON; SCHLEIN, 2001; SAMIE et al., 1990). Em *Phlebotomus langeroni* a atividade de α-glicosidase também foi identificada no divertículo, no entanto os autores sugeriram que a atividade representava uma contaminação das atividades descritas no intestino médio (DILLON; EL KORDY, 1997), devido à baixa atividade detectada. Diferente de flebotomíneos, em mosquitos a presença de alfa-glicosidases nas glândulas salivares está bem consolidada, e já foi descrita, por exemplo, em espécies como *Aedes albopictus* e *Aedes aegypti* (MARINOTTI; DE B&O; MOREIRU, 1996; MARINOTTI; JAMES, 1990).

Em fêmeas não alimentadas, a atividade sobre MUαGlu no conteúdo (atividade solúvel) e no tecido do intestino médio foi induzida após 2 dias de jejum, de modo que a síntese de α-glicosidases nesta condição fisiológica não está diretamente relacionada à ingestão de sacarose. Para fêmeas alimentadas com açúcar, a atividade sobre sacarose nos tecidos do intestino médio também foi induzida após 2 dias de alimentação e esta indução foi variável dependendo da concentração de açúcar e do tipo de açúcar (mono ou dissacarídeos) utilizado para alimentação dos flebotomíneos. Os resultados demonstram que, neste caso, a indução da atividade de α-glicosidase é dependente da ingestão de açúcar. Durante a ingestão de açúcares, parte da alimentação é direcionada ao intestino médio e parte é armazenada no divertículo. Nos dias seguintes à alimentação, os açúcares são liberados em pequenas quantidades para o intestino médio (DILLON; EL KORDY, 1997). Esse padrão corrobora nossos resultados de indução da atividade de α-glicosidase em que a atividade atinge um pico de indução após 48 h da ingestão dos açúcares e permenace elevada nos dias seguintes à alimentação (Trabalho 2). Quando avaliamos a expressão de genes que codificam proteínas relacionadas à atividade de α-glicosidase, não ocorreu indução da expressão dos genes selecionados (Trabalho 3) no intestino médio de fêmeas em jejum ou alimentadas com sacarose, demonstrando diferenças entre o padrão bioquímico e o padrão de expressão. Portanto, a indução da atividade de α-glicosidase não está sendo regulada transcricionalmente, ou a transcrição foi regulada transitoriamente e após 48 horas não foi possível determinar uma indução ou repressão do gene. Em adultos e larvas de *D. melanogaster* foi descrito que alguns genes que codificam carboidrases digestivas, como MAL-A1, são reprimidos em resposta a uma alimentação açucarada, devido a um mecanismo de regulação por feedback (CHNG; HIETAKANGAS; LEMAITRE, 2017). No entanto, o gene MAL-A4 é transitoriamente induzido quando a maltose é oferecida às larvas e a expressão é reprimida quando as larvas são alimentadas com glicose (YAMADA et al., 2018).

A atividade de α-glicosidase foi afetada pela ingestão de diferentes concentrações de sacarose e induzida na presença de sacarose, mas não dos outros mono ou dissacarídeos testados. Esses resultados indicam um mecanismo de ativação pela ingestão de açúcar, neste caso, sacarose. A indução da atividade de α-glicosidase após a alimentação com fontes ricas em açúcares também foi relatada para *Phlebotomus papatasi e Phlebotomus langeroni* (DILLON; EL KORDY, 1997; JACOBSON; SCHLEIN; EISENBERGER, 2001; JACOBSON; STUDENTSKY; SCHLEIN, 2007), mas nestes trabalhos a expressão dos genes não foi avaliada.

Quando fêmeas foram alimentadas com sangue, foi observada indução tanto na fração do conteúdo do intestino médio (enzimas solúveis) quanto na fração tecidual do intestino médio

(enzimas de membrana) após 24 h. Essa indução foi determinada usando tanto o substrato MUαGlu quanto sacarose. Estas fêmeas não tiveram contato com açúcar, mostrando que esta indução foi desencadeada pela ingestão de sangue. Neste contexto também ocorreu a indução da expressão dos genes de maltase/α-glicosidase. Os resultados indicam a presença de uma α-glicosidase solúvel com afinidade por sacarose atuando na digestão das porções glicosídicas do sangue. Essa enzima pode ultrapassar a membrana peritrófica e agir no espaço endoperitrófico, provavelmente atuando sobre glicolipídeos e glicopeptídeos em sinergia com a clivagem proteolítica (tripsina e aminopeptidases) ou lipídica. Em *P. langeroni*, a atividade solúvel de α-glicosidase também foi descrita no intestino médio após a ingestão de sangue com pico de atividade após 24 h de ingestão (DILLON; EL KORDY, 1997). Portanto, as alfa-glicosidases de *L. longipalpis* podem desempenhar diferentes papéis na hidrólise de cadeias glicosídicas durante o processo de digestão, possuindo uma especificidade enzimática diferenciada para açúcares e possivelmente glicoconjugados do sangue. Outro possível papel das α-glicosidases solúveis a se discutir, seria sua atuação no processo de detoxificação de heme durante a digestão do sangue, promovendo a formação de cristais de hemozoína. Este papel foi descrito para as α-glicosidases de *Rhodnius prolixus* (MURY et al., 2009).

O pH ótimo dessas enzimas é apenas ligeiramente afetado pela condição de alimentação ou localização (tecido) das enzimas. Nossos resultados demonstram que as alfa-glicosidases/sacarases de *L. longipalpis* possuem diferentes afinidades e comportamentos dependendo do substrato utilizado. Algumas das enzimas estudadas não seguem a cinética clássica de Michaelis-Menten, sendo inibidas por altas concentrações de substrato. Essas enzimas apresentam altos valores de $K_{m\ app}$ para o substrato sacarose, o que pode ser uma característica relacionada a uma forma de controle metabólico devida à grande ingestão de açúcar, limitando a quantidade de glicose disponível para ser absorvida. Esses altos valores de $K_{m\ app}$ também sugerem que essas enzimas podem apresentar uma enorme variação de atividade em uma ampla gama de concentrações de sacarose, o que é compatível com a alimentação de néctares, onde a concentração de sacarose pode atingir até 1,2 M (HEIL, 2011). A inibição em altas concentrações de substrato pode ser explicada por dois mecanismos: a ligação de uma segunda molécula de substrato no sítio ativo da enzima, levando ao mecanismo de inibição pelo subtrato ou ainda por reações de transglicosilação. Enzimas da família GH13 catalizam a hidrólise das ligações glicosídicas, retendo a configuração anomérica do substrato, o que permite a realização das reações de transglicosilação (CHIBA, 1997). Para α-glicosidase de *Acyrtosiphon pisum*, um comportamento cinético similar foi descrito (CRISTOFOLETTI et al., 2003). Neste trabalho, os autores consideram que a transglicosilação pode ser uma adaptação

para a obtenção de monossacarídeos a partir de sacarose em alta concentração, sem o aumento da osmolaridade, de forma que as células epiteliais do instestino médio não sofram choque osmótico. Nesse aspecto, as enzimas de *L. longipalpis* podem possuir a mesma adaptação bioquímica. Seria muito interessante observar se esta é uma característica cinética comum das α-glicosidases de insetos que se alimentam de néctar ou seiva de plantas e também avaliar se esta característica representa um caso de divergência ou convergência evolutiva.

Como descrito anteriormente, as análises de bioinformática demonstraram a presença de 3 genes codificantes para maltases e estas mostram sinais de ancoramento à membrana. Os resultados bioquímicos descrevendo a presença de uma atividade solúvel, os padrões de indução frente as condições de alimentação, constantes cinéticas diferentes e um perfil cromatográfico complexo frente aos substratos utilizados, demonstram a presença de pelo menos 4 isoformas de α-glicosidases atuando nas condições descritas. Alguns mecanismos podem levar ao aparecimento das diversas atividades descritas neste trabalho. A maltase LlMal3, por exemplo, é supostamente ancorada à membrana por uma âncora GPI. Esta pode sofrer clivagem pela ação de fosfolipases liberando-a da membrana, o que leva ao aparecimento de uma forma solúvel. Além disso os genes codificantes para maltases podem sofrer "splicing alternativo", produzindo múltiplas isoformas de α-glicosidase. A remoção do último exon elimina a região transmembrana destas proteínas e permite que sejam secretadas na forma solúvel, o que altera as características bioquímicas das enzimas. Diferentes transcritos de maltases foram descritos em *A. aegypti*, *D. melanogaster* e *D. ananassae*, por exemplo.

Enquanto observamos a indução na expressão dos genes de maltase em fêmeas alimentadas exclusivamente com sangue, em fêmeas infectadas com *L. mexicana* a expressão é modulada negativamente. Neste mesmo contexto, a infecção de fêmeas de *L. longipalpis* com *L. mexicana*, nos permitiu avaliar o desenvolvimento dos parasitos em um ambiente nutricionalmente pobre devido à diminuição de carboidratos na dieta dos flebotomíneos (Trabalho 4). Durante o desenvolvimento do parasito *Leishmania* dentro do inseto vetor, ocorrem várias mudanças fisiológicas, bioquímicas e moleculares, estas devem ser ultrapassadas para que ocorra o estabelecimento da infecção. A modulação de diferentes enzimas por parasitos tem sido descrita; como a modulação de proteases (DILLON; LANE, 1993; SANTOS et al., 2014; TELLERIA et al., 2010), auxiliando na resistência do parasito à clivagem proteolítica durante a digestão do sangue. Contudo, é a primeira vez que a modulação de uma carboidrase, frente a infecção é demonstrada. Um possível papel para a modulação da expressão das α-glicosidases é evitar a quebra dos glicoconjugados de superfície, que são essenciais durante o desenvolvimento e manutenção da infecção por *Leishmania* no intestino

dos insetos vetores (COELHO-FINAMORE et al., 2011; DE ASSIS et al., 2012; SACKS, 1995; SOARES et al., 2002). As espécies de *Leishmania* possuem glicoconjugados superficiais como lipofosfoglicanos (GLP) e glicoinositolfosfolipídeos (GIPLs) (DE ASSIS et al., 2012), e estas são estruturas complexas com polimorfismos causados pela composição de açúcar das estruturas de glicana. Para *L. mexicana* estas estruturas estão bem descritas. Por exemplo, na extremidade não redutora o lipofosfoglicano possui um "cap" composto por Manαl-2Man, Manαl-2Manαl-2Man ou Manαl-2(Galβl-4)Man (ILG et al., 1992) e nos glicoinositolfosfolipidios (GIPLs) existe um elevado grau de galactosilação nas estruturas ligadas ao glicano central de Manα1-4GlcN (MCCONVILLE et al., 1993). Manose e galactose são epímeros de glicose, então é possível que as α-glicosidases atuem sobre essas ligações, hidrolisando esses glicoconjugados. A presença de glicosidases solúveis com atividade de N-acetil-β-D-glucosaminidase e N-acetil-β-D-galactosaminidase foi descrita no homogenato do intestino médio de fêmeas de *L. Longipalpis* alimentadas com sangue (GONTIJO et al., 1998), estas atividades podem representar a mesma atividade solúvel descrita neste trabalho.

Quanto ao desenvolvimento de *L. mexicana* em fêmeas que não receberam uma dieta açucarada, os resultados demonstraram que o parasito não é estritamente dependente da ingestão de açúcares pelo flebotomíneo. Enquanto muitos trabalhos relatam a importância da ingestão de uma dieta açucarada pelo flebotomíneo para o desenvolvimento do parasito (AÑEZ; NIEVES; CAZORLA, 1989; GONTIJO et al., 1996; JACOBSON; SCHLEIN, 2001; LYDA et al., 2015), nossos resultados demonstram que o número de parasitos e de formas metacíclicas presentes não é significativamente diferente quando comparamos fêmeas que receberam açúcar ou fêmeas alimentadas exclusivamente com sangue. Durante o desenvolvimento, além de adquirir os nutrientes provenientes da dieta açucarada, os parasitos devem ser eficientes na absorção de nutrientes liberados durante a digestão do sangue, especialmente nas fases iniciais do seu desenvolvimento, quando o parasito fica retido dentro da matriz peritrófica. A absorção de nutrientes provenientes da digestão sanguínea permite que o parasito se desenvolva mesmo em condições de restrição em que moléculas de açúcares não estão disponíveis.

Durante o desenvolvimento de parasitos dentro do intestino do vetor, ocorre a migração das formas promastigotas em direção à cardia, o que leva ao acúmulo de parasitos e rompimento da válvula de estomodeu. Assim, durante uma nova alimentação sanguínea as fêmeas podem transmitir as formas promastigotas metacíclicas infectantes (BATES; ROGERS, 2004; ROGERS; CHANCE; BATES, 2002). Foi proposto que durante o desenvolvimento, o movimento de migração é direcionado pela detecção das moléculas de açúcares ingeridas pelos flebotomíneos (KILLICK-KENDRICK, 1978), estimuladas por um mecanismo de osmotaxia e

quimiotaxia (BRAY, 1983; LESLIE; BARRETT; BURCHMORE, 2002; OLIVIERA; MELO; GONTIJO, 2000). Demonstramos que o movimento de migração do parasito em direção a região anterior do intestino médio, não é totalmente afetado pela falta de dieta açucarada. A presença de açúcar, gerando um gradiente osmótico e também a natureza química do açúcar, é importante para direcionar a migração de parasitos para a região da cárdia, mas não essencial. Em um ambiente com ausência de moléculas de açúcares, outros componentes devem fornecer o ambiente químico necessário para o estímulo de migração, como componentes da saliva, por exemplo. Os resultados encontrados apontam novos caminhos para o estudo do desenvolvimento do parasito dentro do intestino do inseto vetor, como a elucidação dos fatores quimiotáxicos envolvidos na migração dos parasitos e a modificação na expressão de proteínas e glicoconjugados de superfície para sobrevivência em uma ambiente de baixa osmolaridade. Foi proposto que, durante o desenvolvimento do parasito, estes possuem dois ciclos independentes de multiplicação. Inicialmente, com a divisão em promastigotas procíclicas, que ocorre dentro da matriz peritrófica durante a digestão do sangue, e posteriormente na forma de promastigotas leptomonadas (*sugar phase*), que ocorre após a digestão do sangue (GOSSAGE; ROGERS; BATES, 2003). O aparecimento das formas promastigotas leptomonadas pode ser influenciado pela presença de açúcares. Assim, seria interessante avaliar se todas as formas descritas na presença de açúcar também se desenvolvem em um ambiente de baixa osmolaridade (ausência de açúcar). Recentemente, um novo ciclo replicativo foi descrito, gerando promastigotas retroleptomonas após uma segunda alimentação sanguínea (SERAFIM et al., 2018), mas o papel da dieta açucarada nesse fenômeno ainda não foi estudado.

O genoma de *L. longipalpis* possui 21 genes pertencentes à família GH13, descritos como α-amilases, maltases, proteínas de transporte de aminoácidos (cadeia pesada), enzima ramificadora de α-1,4-glucana e enzima desramificadora de glicogênio, e 6 genes GH31 descritos como glicosidases NET37, α-glicosidase lisossomal e α-glicosidase neutra (subunidade α). Dentre as funções canônicas descritas para estas proteínas estão o metabolismo do açúcar (α-amilases, maltases), metabolismo do glicogênio (enzima ramificadora de α-1,4-glucana e enzima desramificadora de glicogênio, α-glicosidase lisossomal), transporte de aminoácidos (NBAThc e CD98hc), controle de qualidade de N-glicosilação no retículo endoplasmático (α-glicosidase neutra) e regulação da miogênese (glicosidase NET37).

Dentre os genes classificados como α-amilases, ocorreu uma expansão no número de genes em comparação com outros insetos, sendo estes organizados em dois “clusters”. Os genes agrupados são espacialmente localizados próximos uns dos outros, e alguns destes genes possuem alta similaridade de seqüência, o que indica que provavelmente se originaram de

duplicações recentes. Dentre as sequências de α-amilases descritas e classificadas no "cluster" B, algumas não apresentaram os resíduos catalíticos nas regiões conservadas presentes no sítio ativo, a exemplo de LlAamyB1, LlAamyB4, LlAamyB6. Dessa forma não foi possível confirmar se estas proteínas representam cópias inativas ou funcionam com um mecanismo catalítico distinto do clássico descrito para as proteínas GH13. Na família GH31 os resíduos catalíticos para a α-glicosidase lisossomal também não foram encontrados, e o gene não foi expresso nas condições analisadas. Isso sugere que em adultos de *L. longipalpis* a via responsável pela degradação de glicogênio ocorre no citosol com a atuação da enzima desramificadora de glicogênio (estrutura conservada e domínios catalíticos conservados) e não nos lisossomos. Em espécies como *A. aegypti*, *A. gambiae*, *C. quinquefasciatus* também não identificamos os resíduos catalíticos nos domínios conservados das α-glicosidases lisossomais e em *D. melanogaster* e *D. ananassae* não identificamos genes codificantes para a proteína lisossomal. A degradação do glicogênio ocorre através de uma cascata enzimática desencadeada por regulação hormonal. Quando os níveis de trealose na hemolinfa estão baixos, os hormônios trealosêmico, octopamina e adipocinético desencadeiam uma cascata enzimática levando à ativação da glicogênio fosforilase. A glicogênio fosforilase libera resíduos de glicose-1-fosfato da cadeia linear de glicogênio. A completa degradação do glicogênio ocorre com a ação da enzima desramificadora de glicogênio que desfaz os pontos de ramificação (ADEVA-ANDANY et al., 2016; KLOWDEN, 2013; THOMPSON, 2003). Em larvas *D. melanogaster* foi demonstrado que o hormônio adipocinético não está envolvido na mobilização de glicogênio no corpo gorduroso em períodos de jejum, e que a expressão de genes de maltase é transientemente induzida quando a larva passa de um período de jejum para uma alimentação contendo maltose, amido ou farelo de milho. Além disso, quando as larvas são alimentadas com glicose a expressão de maltase não é induzida (YAMADA et al., 2018). Quando glicose é ingerida um mecanismo de regulação por "feedback", chamado supressão por glicose, esta atua para limitar a digestão de carboidratos (CHNG; HIETAKANGAS; LEMAITRE, 2017).

No genoma de *L. longipalpis*, além da expansão no número de α-amilases descritas, também ocorreu retração no número de genes de maltases. Em outros insetos utilizados para comparação como *P. papatasi*, *A. aegypti*, *A. gambiae* e *C. quinquefasciatus*, os genes codificantes de maltase formam dois *clusters*, cada um consistindo de três a cinco genes, e no processo de evolução alguns destes genes podem ter adquirido outras funções. Em algumas espécies de mosquitos como *Culex pipiens*, *C. quinquefasciatus* e *A. gambiae*, as α-glicosidases são descritas como proteínas ancoradas à membrana por uma âncora GPI e funcionam como receptores de toxina binária da bactéria entomopatogênica *Lysinibacillus sphaericus*

(DARBOUX et al., 2001; OPOTA et al., 2008; ROMÃO et al., 2006). Em *D. melanogaster*, diferentes níveis de expressão em diferentes tecidos, com alterações temporais dependentes da fase de desenvolvimento foram observados para maltases, e foi sugerida uma função reguladora não enzimática para essas proteínas (GABRIŠKO, 2013). Algumas α-glicosidases de *R. prolixus*, por exemplo, estão envolvidas na desintoxicação do heme durante a digestão sanguínea com a formação de hemozoína (MURY et al., 2009). No caso das maltases de *L. longipalpis*, apesar de ter ocorrido a retração no número de genes, a caracterização bioquímica permitiu demonstrar diferentes padrões de indução, o que sugere a presença de diferentes α-glicosidases. Além de manterem seu papel na hidrólise de açúcares, o padrão de expressão e a indução de uma atividade solúvel em fêmeas alimentadas com sangue sugere a diversificação de função com uma possível especialização na digestão de glicoproteínas e glicolipídios do sangue. A expansão das α-amilases pode refletir uma especificidade tecidual, adaptação a diferentes condições de reação, ou ainda uma adaptação a presença de diferentes inibidores de plantas. Pode também refletir a adaptação às condições fisiológicas variadas, ou além disso as α-amilases podem ter adquirido funções adicionais à digestão de carboidratos. Assim, os resultados encontrados neste trabalho abrem perspectivas para investigação futura de vários aspectos relacionados à caracterização detalhada do papel de cada uma dessas enzimas dentro do metabolismo de açúcares, regulação da expressão de enzimas envolvidas no metabolismo energético de flebotomíneos e também do seu papel na interação com o parasito durante o processo de infecção.

5. Conclusões

- A atividade de α-glicosidase está presente sobretudo no intestino médio, e em outras partes de adultos de *L. longipalpis*, como cabeça, divertículo, carcaça e intestino posterior, sendo as enzimas mais ativas sobre sacarose em comparação com o substrato sintético MUαGlu.
- *L. longipalpis* possui diferentes α-glicosidases com diferentes características bioquímicas. O pH ótimo é ligeiramente afetado pela condição de alimentação ou localização (tecido) das enzimas. Algumas das α-glicosidases descritas não seguem a cinética clássica de Michaelis-Menten, sendo inibidas em altas concentrações de substrato e apresentam altos valores de $\boldsymbol{K_m}$ para o substrato sacarose. Isso caracteriza um tipo de controle metabólico devido à ingestão de altas concentrações de açúcares, limitando a quantidade de monossacarídeos disponíveis para absorção.
- A inibição dessas α-glicosidases em altas concentrações de substrato pode ocorrer através do mecanismo de transglicosilação, o que pode representar uma adaptação para a obtenção de monossacarídeos a partir de sacarose em alta concentração, sem o aumento da osmolaridade, de forma que as células epiteliais do intestino médio não sofram choque osmótico.
- As α-glicosidases/maltases de *L. longipalpis* podem ser proteínas ancoradas à membrana do intestino médio (maior atividade e maior expressão). Estas podem também se apresentar como isoformas solúveis a partir da solubilização das proteínas de membrana ou através de 'splicing" alternativo produzindo proteínas sem domínio transmembrana, alterando assim suas propriedades bioquímicas. Estas enzimas apresentam diferentes padrões de indução de atividade e expressão dos genes diante de diferentes condições fisiológicas. Dependendo da condição fisiológica envolvida o aumento da atividade catalítica pode ou não estar relacionado a um controle transcricional.
- α-glicosidases/maltases de *L. longipalpis* desempenham diferentes papéis na hidrólise de glicosídeos durante o processo de digestão associado a uma especificidade enzimática diferenciada para açúcares e possivelmente para glico derivados do sangue.
- A indução da expressão das maltases durante a alimentação sanguínea é modulada negativamente durante infecção com *L. mexicana*, provavelmente limitando a quebra dos glicoconjugados de superfície, que são essenciais durante o desenvolvimento e manutenção da infecção por *Leishmania* no intestino dos insetos vetores.

- O desenvolvimento de *L. mexicana* em *L. longipalpis* não é estritamente dependente da ingestão de uma dieta açucarada pelo flebotomíneo. A presença de açúcar, gerando um gradiente osmótico e também a natureza química do açúcar, é importante para direcionar a migração de parasitos para a região da cárdia, mas não essencial. Em um ambiente com ausência de moléculas de açúcares, outros componentes podem fornecer o ambiente químico necessário para o estímulo de migração.
- O genoma de *L. longipalpis* possui 21 genes pertencentes à família GH13 e 6 genes da GH31, codificando proteínas envolvidas no metabolismo do açúcar (α-amilases, maltases), metabolismo do glicogênio (enzima ramificadora de α-1,4-glucana e enzima desramificadora de glicogênio, α-glicosidase lisossomal), transporte de aminoácidos (NBAThc e CD98hc), controle de qualidade da N-glicosilação no retículo endoplasmático (α-glicosidase neutra) e regulação da miogênese (glicosidase NET37).
- Em *L. longipalpis* ocorreu retração no número de genes de maltases e uma expansão das α-amilases, em comparação com outros insetos analisados (Diptera), sendo as α-amilases organizadas em dois "clusters". A expansão das α-amilases sugere que estas possam ter adquirido novas funções ou especificidade tecidual, ou se adaptado a diferentes inibidores presentes em plantas que são utilizadas como fontes alimentares (seiva).

6. Outras atividades

6.1. Trabalhos publicados

FERREIRA, T.N. et al. Transmission blocking sugar baits for the control of *Leishmania* development inside sand flies using environmentally friendly beta glycosides and their aglycones. **Parasites and Vectors**, 2018. *In press*

DA COSTA, S. G. et al. Penicillium citrinum UFV1 β-glucosidases: purification, characterization, and application for biomass saccharification. **Biotechnology for Biofuels**, v. 11, n. 1, p. 226, 2018.

MORAES, C. S. et al. Second Blood Meal by Female Lutzomyia longipalpis: Enhancement by Oviposition and Its Effects on Digestion, Longevity, and Leishmania Infection. **BioMed Research International**, 2018.

DO NASCIMENTO, N. A. et al. N-glycosylation influences the catalytic activity of mosquito α-glucosidases associated with susceptibility or refractoriness to Lysinibacillus sphaericus. **Insect Biochemistry and Molecular Biology**, v. 81, n. 2017, p. 62–71, 2017.

HENRIQUES, B. S. et al. Genome wide mapping of peptidases in Rhodnius prolixus: Identification of protease gene duplications, horizontally transferred proteases and analysis of peptidase a1 structures, with considerations on their role in the evolution of hematophagy in triatomine. **Frontiers in Physiology**, v. 8, n. DEC, p. 1–22, 2017.

MESQUITA, R. D. et al. Genome of Rhodnius prolixus, an insect vector of Chagas disease, reveals unique adaptations to hematophagy and parasite infection. **Proceedings of the National Academy of Sciences**, v. 112, n. 48, p. 14936–14941, 2015.

PROFETA, G. S. et al. Standardization of a continuous assay for glycosidases and its use for screening insect gut samples at individual and populational levels. **Frontiers in Physiology**, v. 8, n. MAY, p. 1–12, 2017.

6.2. Trabalhos em preparação

Blood-derived haem is an elicitor for an anti-leishmanial activity in the sand fly model *Lutzomyia longipalpis*. José R. Silva, Samara Costa, Emma Shawcross and Rod Dillon.

Understanding the sugar nectar composition in Christmas Bell flowers using Capillary Electrophoresis and enzymatic assays. Dayanne Mozaner Bordin, Samara Costa, Graham Pyke, John Kalman, Philip Doble, Fernando Ariel Genta, and Lucas Blanes.

Molecular modeling and inhibition kinetics studies of glycoconjugate 1*H*-1,2,3-triazoles on α-glucosidase. Mario Roberto Senger, Samara Costa , Bruno Junior Neves, Rafael Ferreira Dantas, Sabrina Baptista Ferreira, Carolina Horta Andrade, Vitor Francisco Ferreira, Fernando Ariel Genta, Floriano Paes Silva Junior.

7. Referências Bibliográficas

ABAD, M. C. et al. The x-ray crystallographic structure of Escherichia coli branching enzyme. **Journal of Biological Chemistry**, v. 277, n. 44, p. 42164–42170, 2002.

ABEDI, Z. H.; BROWN, A. W. A. Peritrophic Membrane as Vehicle for DDT and DDE Excretion in Aedes aegypti Larvae. **Annals of the Entomological Society of America**, 1961.

ADEVA-ANDANY, M. M. et al. Glycogen metabolism in humans. **BBA Clinical**, v. 5, p. 85–100, 2016.

AKHOUNDI, M. et al. A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies. **PLoS Neglected Tropical Diseases**, v. 10, n. 3, p. 1–40, 2016.

AKILOV, O. E.; KHACHEMOUNE, A.; HASAN, T. Clinical manifestations and classification of Old World cutaneous leishmaniasis. **International Journal of Dermatology**, v. 46, p. 132–142, 2007.

AÑEZ, N.; NIEVES, E.; CAZORLA, D. The validity of the developmental pattern in the sandfly gut for classification of leishmania. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 83, n. 5, p. 634–635, 1989.

ANTINORI, S.; SCHIFANELLA, L.; CORBELLINO, M. Leishmaniasis: New insights from an old and neglected disease. **European Journal of Clinical Microbiology and Infectious Diseases**, v. 31, p. 109–118, 2012.

ARRESE, E. L.; SOULAGES, J. L. Insect Fat Body: Energy, Metabolism, and Regulation. **Annual Review of Entomology**, v. 55, n. 1, p. 207–225, 2010.

ASHFORD, D. A.; SMITH, W. A.; DOUGLAS, A. E. Living on a high sugar diet: the fate of sucrose ingested by a phloem-feeding insect, the pea aphid Acyrthosiphon pisum. **Journal of Insect Physiology**, v. 46, n. 3, p. 335–341, 2000.

BARBEHENN, R. V. Non-Absorption of Ingested Lipophilic and Amphiphilic Allelochemicals by Generalist Grasshoppers: The Role of Extractive Ultrafiltration by the Peritrophic Envelope. **Archives of Insect Biochemistry and Physiology**, v. 42, n. 2, p. 130–137, 1999.

BATES, P.; ROGERS, M. New Insights into the Developmental Biology and Transmission Mechanisms of Leishmania. **Current Molecular Medicine**, v. 4, n. 6, p. 601–609, 2004.

BILLINGSLEY, P. F.; HECKER, H. Blood Digestion in the Mosquito, Anopheles stephensi Liston (Diptera: Culicidae): Activity and Distribution of Trypsin, Aminopeptidase, and a-Glucosidase in the Midgut. **J. Med. Entomol**, v. 28, n. 6, p. 865–871, 1991.

BLACKBURN, K. et al. The peritrophic membrane of the female sandfly Phlebotomus papatasi. **Annals of Tropical Medicine and Parasitology**, v. 82, n. 6, p. 613–619, 1988.

BOEL, E. et al. Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-A resolution of two enzymes from Aspergillus. **Biochemistry**, v. 29, n. 26, p. 6244–6249, 1990.

BRAY, R. S. Leishmania: Chemotaxic Responses of Promastigotes and Macrophages In Vitro. **The Journal of Protozoology**, v. 30, n. 2, p. 322–329, 1983.

BRAYER, G. D.; LUO, Y.; WITHERS, S. G. The structure of human pancreatic a-amylase at 1.8 A resolution and comparisons with related enzymes. **Protein Science**, v. 4, n. 9, p. 1730–

1742, 1995.

BRAZIL, R.P.; BRAZIL, B. . Biologia de Flebotomíneos Neotropicais. In: RANGEL, E.F., LAINSON, R. (Ed.). . **Flebotomíneos do Brasil**. Rio de Janeiro: Fiocruz, 2003. p. 257–274.

BRAZIL, R. P. The dispersion of Lutzomyia longipalpis in urban areas. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 46, n. 3, p. 263–264, 2013.

BRAZIL, R. P.; RODRIGUES, A. A. F.; FILHO, J. D. A. Sand Fly Vectors of Leishmania in the Americas - A Mini Review. **Entomology, Ornithology and Herpetology: Current Research**, v. 04, n. 02, p. 4–7, 2015.

BRZOZOWSKI, A. M.; DAVIES, G. J. Structure of the Aspergillus oryzae α-amylase complexed with the inhibitor acarbose at 2.0 Å resolution. **Biochemistry**, v. 36, n. 36, p. 10837–10845, 1997.

BUISSON, G. et al. Three dimensional structure of porcine pancreatic alpha-amylase at 2.9 A resolution. Role of calcium in structure and activity. **The EMBO journal**, v. 6, n. 13, p. 3909–3916, 1987.

CAMERON, M. M. et al. Sugar meal sources for the phlebotomine sandfly Lutzomyia longipalpis in Ceará State, Brazil. **Medical and Veterinary Entomology**, v. 9, n. 3, p. 263–272, 1995.

CANAL CIÊNCIA, 2018. Dispponível em: < http://www.canalciencia.ibict.br/pesquisa/0295_A_caminho_da_cura_da_leishmaniose_visceral_canina.html >. Acesso em 12 de dezembro de 2018.

CANTAREL, B. et al. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. **Nucleic Acids Res**, v. 37, p. D233–D238, 2009.

CAZy, 2018. Disponível em:< http://www.cazy.org/GH13.html>. Acesso em 19 de junho de 2018.

CAZypedia, 2018. Disponível em : <http://www.cazypedia.org/index.php/Glycoside_hydrolases#Classical_Koshland_retaining_mechanism>. Acesso em 11 julho de 2018.

CDC, 2018. Disponível em: <https://www.cdc.gov/parasites/leishmaniasis/biology.html.>. Acesso em 20 julho 2018.

CHADEE, D. D.; BEIER, J. C.; MARTINEZ, R. The effect of the cibarial armature on blood meal haemolysis of four anopheline mosquitoes. **Bulletin of Entomological Research**, 1996.

CHEN, I. H. B. et al. Nuclear envelope transmembrane proteins (NETs) that are up-regulated during myogenesis. **BMC Cell Biology**, v. 7, p. 1–16, 2006.

CHIBA, S. Molecular mechanism in α-glucosidase and glucoamylase. **Bioscience, Biotechnology, and Biochemistry**, v. 61, n. 8, p. 1233–1239, 1997.

CHNG, W. BIN A.; HIETAKANGAS, V.; LEMAITRE, B. Physiological Adaptations to Sugar Intake: New Paradigms from Drosophila melanogaster. **Trends in Endocrinology and Metabolism**, v. 28, n. 2, p. 131–142, 2017.

CLARK, T. M. Evolution and adaptive significance of larval midgut alkalinization in the insect superoder mecopterida. **Journal of Chemical Ecology**, v. 25, n. 8, p. 1945–1960, 1999.

COELHO-FINAMORE, J. M. et al. Leishmania infantum: Lipophosphoglycan intraspecific variation and interaction with vertebrate and invertebrate hosts. **International Journal for Parasitology**, v. 41, p. 333–342, 2011.

COLWELL, D. D.; DANTAS-TORRES, F.; OTRANTO, D. Vector-borne parasitic zoonoses: Emerging scenarios and new perspectives. **Veterinary Parasitology**, v. 182, n. 1, p. 14–21, 2011.

CRISTOFOLETTI, P. T. et al. Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid Acyrthosiphon pisum. **Journal of Insect Physiology**, v. 49, p. 11–24, 2003.

D'AMICO, S.; GERDAY, C.; FELLER, G. Structural similarities and evolutionary relationships in chloride-dependent alpha-amylases. **Gene**, v. 253, n. 1, p. 95–105, 2000.

DANTAS-TORRES, F.; BRANDÃO-FILHO, S. P. Visceral leishmaniasis in Brazil: Revisiting paradigms of epidemiology and control. **Revista do Instituto de Medicina Tropical de Sao Paulo**, v. 48, n. 3, p. 151–156, 2006.

DARBOUX, I. et al. The receptor of Bacillus sphaericus binary toxin in Culex pipiens (Diptera: Culicidae) midgut: Molecular cloning and expression. **Insect Biochemistry and Molecular Biology**, v. 31, n. 10, p. 981–990, 2001.

DATTA, K.; GUAN, T.; GERACE, L. NET37, a nuclear envelope transmembrane protein with glycosidase homology, is involved in myoblast differentiation. **Journal of Biological Chemistry**, v. 284, n. 43, p. 29666–29676, 2009.

DAVIES, G.; HENRISSAT, B. Structures and mechanisms of glycosyl hydrolases. **Structure**, v. 3, n. 9, p. 853–859, 1995.

DE ASSIS, R. R. et al. Glycoconjugates in New World species of Leishmania: Polymorphisms in lipophosphoglycan and glycoinositolphospholipids and interaction with hosts. **Biochimica et Biophysica Acta - General Subjects**, v. 1820, n. 9, p. 1354–1365, 2012.

DE PITA-PEREIRA, D. et al. Detection of natural infection in Lutzomyia cruzi and Lutzomyia forattinii (Diptera: Psychodidae: Phlebotominae) by Leishmania infantum chagasi in an endemic area of visceral leishmaniasis in Brazil using a PCR multiplex assay. **Acta Tropica**, v. 107, n. 1, p. 66–69, jul. 2008.

DILLON, R. J.; DILLON, V. M. The gut bacteria of insects : Nonpathogenic Interactions. **Annual Review of Entomology**, v. 49, n. 1, p. 71–92, 2004.

DILLON, R. J.; EL KORDY, E. Carbohydrate digestion in sandflies: α-glucosidase activity in the midgut of Phlebotomus langeroni. **Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology**, v. 116, n. 1, p. 35–40, 1997.

DILLON, R. J.; LANE, R. P. Influence of Leishmania infection on blood-meal digestion in the sandflies Phlebotomus papatasi and P. langeroni. **Parasitology Research**, v. 79, n. 6, p. 492–496, 1993.

DO NASCIMENTO, N. A. et al. N-glycosylation influences the catalytic activity of mosquito α-glucosidases associated with susceptibility or refractoriness to Lysinibacillus sphaericus. **Insect Biochemistry and Molecular Biology**, v. 81, n. 2017, p. 62–71, 2017.

DOUGLAS, A. E. Alimentary canal, digestion and absorption. In: SIMPSON, S. J.; DOUGLAS, A. E. (Eds.). . **The Insects: Structure and Function**. Fifth ed. Cambridge, UK: Cambridge University Press, 2013. p. 46–80.

ERNST, H. A. et al. Structure of the Sulfolobus solfataricus α-Glucosidase: Implications for Domain Conservation and Substrate Recognition in GH31. **Journal of Molecular Biology**, v. 358, n. 4, p. 1106–1124, 2006.

ESPINOSA, O. A. et al. An appraisal of the taxonomy and nomenclature of trypanosomatids presently classified as Leishmania and Endotrypanum. **Parasitology**, v. 145, n. 4, p. 430–442, 2016.

FELICIANGELI, M. D. Natural breeding places of phlebotomine sandflies. **Medical and Veterinary Entomology**, v. 18, n. 1, p. 71–80, 2004.

FELLER, G. et al. Structural and functional aspects of chloride binding to Alteromonas haloplanctis α-amylase. **Journal of Biological Chemistry**, v. 271, n. 39, p. 23836–23841, 1996.

FERREIRA, L. M. et al. The orthologue to the Cpm1/Cqm1 receptor in Aedes aegypti is expressed as a midgut GPI-anchored alpha-glucosidase, which does not bind to the insecticidal binary toxin. **Insect Biochemistry and Molecular Biology**, v. 40, n. 8, p. 604–610, 2010.

GABRIŠKO, M. Evolutionary history of eukaryotic α-glucosidases from the α-amylase family. **Journal of Molecular Evolution**, v. 76, n. 3, p. 129–145, 2013.

GABRIŠKO, M.; JANEČEK, Š. Looking for the ancestry of the heavy-chain subunits of heteromeric amino acid transporters rBAT and 4F2hc within the GH13 α-amylase family. **FEBS Journal**, v. 276, n. 24, p. 7265–7278, 2009.

GALATI, E. Classificação de Phlebotominae. In: RANGEL, E. F.; LAINSON, R. (Eds.). . **Flebotomíneos do Brasil**. Rio de Janeiro: Fiocruz, 2003. p. 23–51.

GEMETCHU, T. The morphology and fine structure of the midgut and peritrophic membrane of the adult female, Phlebotomus longipes Parrot and Martin (Diptera: Psychodidae). **Annals of Tropical Medicine and Parasitology**, 1974.

GHADAMYARI, M.; HOSSEININAVEH, V.; SHARIFI, M. Partial biochemical characterization of a and B-glucosidases of lesser mulberry pyralid, Glyphodes pyloalis Walker (Lep.: Pyralidae). **Comptes rendus - Biologies**, v. 333, p. 197–204, 2010.

GONTIJO, B.; DE CARVALHO, M. DE L. R. Leishmaniose tegumentar americana. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 36, n. 1, p. 71–80, 2003.

GONTIJO, N. F. et al. Glycosidases in Leishmania and Their Importance for Leishmania in Phlebotomine Sandflies with Special Reference to Purification and Characterization of a Sucrase. **Experimental Parasitology**, v. 83, p. 117–124, 1996.

GONTIJO, N. F. et al. Lutzomyia longipalpis: pH in the Gut, Digestive Glycosidases, and Some Speculations upon Leishmania Development. **Experimental Parasitology**, v. 90, p. 212–219, 1998.

GOSSAGE, S. M.; ROGERS, M. E.; BATES, P. A. Two separate growth phases during the development of Leishmania in sand flies: Implications for understanding the life cycle. **International Journal for Parasitology**, v. 33, n. 10, p. 1027–1034, 2003.

GUO, Q. et al. Single nucleotide deletion of cqm1 gene results in the development of resistance to Bacillus sphaericus in Culex quinquefasciatus. **Journal of Insect Physiology**, v. 59, n. 9, p. 967–973, 2013.

HEGEDUS, D. et al. New Insights into Peritrophic Matrix Synthesis, Architecture, and Function. **Annual Review of Entomology**, v. 54, n. 1, p. 285–302, 2009.

HEIL, M. Nectar: Generation, regulation and ecological functions. **Trends in Plant Science**, v. 16, p. 191–200, 2011.

HOEFSLOOT, L. H. et al. Primary structure and processing of lysosomal alpha-glucosidase; homology with the intestinal sucrase-isomaltase complex. **The EMBO journal**, v. 7, n. 6, p. 1697–1704, 1988.

HUBER, R. E.; THOMPSON, D. J. Studies on a Honey Bee Sucrase Exhibiting Unusual Kinetics and Transglucolytic Activity. **Biochemistry**, v. 12, n. 20, p. 4011–4020, 1973.

HURRELL, B. P.; REGLI, I. B.; TACCHINI-COTTIER, F. Different Leishmania Species Drive Distinct Neutrophil Functions. **Trends in Parasitology**, v. 32, n. 5, p. 392–401, 2016.

ILG, T. et al. Structure of Leishmania mexicana lipophosphoglycan. **Journal of Biological Chemistry**, v. 267, n. 10, p. 6834–6840, 1992.

JACOBSON, R. L.; SCHLEIN, Y. Phlebotomus papatasi and Leishmania major parasites express a-amylase and a-glucosidase. **Acta Tropica**, v. 78, p. 41–49, 2001.

JACOBSON, R. L.; SCHLEIN, Y.; EISENBERGER, C. L. The biological function of sand fly and Leishmania glycosidases. **Med Microbiol Immunol**, n. 190, p. 51–55, 2001.

JACOBSON, R. L.; STUDENTSKY, L.; SCHLEIN, Y. Glycolytic and chitinolytic activities of Phlebotomus papatasi (Diptera: Psychodidae) from diverse ecological habitats. **Folia Parasitologica**, v. 54, n. 4, p. 301–309, 2007.

JANECEK, S. Alpha-Amylase Family: Molecular Biology and Evolution. **Journal of Progress in Biophysics & Molecular Biology**, v. 67, n. 1, p. 67–97, 1997.

JANEČEK, Š. Close evolutionary relatedness among functionally distantly related members of the (α/β)8-barrel glycosyl hydrolases suggested by the similarity of their fifth conserved sequence region. **FEBS Letters**, v. 377, n. 1, p. 6–8, 1995.

JANEČEK, Š.; SVENSSON, B.; MACGREGOR, E. A. A remote but significant sequence homology between glycoside hydrolase clan GH-H and family GH31. **FEBS Letters**, v. 581, n. 7, p. 1261–1268, 2007.

JARIYAPAN, N. et al. Leishmania (Mundinia) orientalis n. sp. (Trypanosomatidae), a parasite from Thailand responsible for localised cutaneous leishmaniasis. **Parasites and Vectors**, v. 11, n. 1, p. 351, 2018.

KADZIOLA, A. et al. Crystal and molecular structure of barley alpha-amylase. **Journal of molecular biology**, v. 239, n. 1, p. 104–121, 1994.

KAMHAWI, S. et al. A role for insect galectins in parasite survival. **Cell**, v. 119, n. 3, p. 329–341, 2004.

KAMHAWI, S. Phlebotomine sand flies and Leishmania parasites: friends or foes? **Trends in Parasitology**, v. 22, n. 9, p. 439–445, 2006.

KANAI, R. et al. Crystal structure of cyclodextrin glucanotransferase from alkalophilic Bacillus sp. 1011 Complexed with 1-deoxynojirimycin at 2.0 Å resolution. **Journal of Biochemistry**, v. 129, n. 4, p. 593–598, 2001.

KANOST, M. R.; CLEM, R. J. **Insect Proteases**. [s.l.] Elsevier B.V., 2012.

KAYE, P.; SCOTT, P. Leishmaniasis: Complexity at the host-pathogen interface. **Nature Reviews Microbiology**, v. 9, n. 8, p. 604–615, 2011.

KILLICK-KENDRICK, R. Recent advances and outstanding problems in the biology of phlebotomine sandflies. A review. **Acta tropica**, v. 35, n. 4, p. 297–313, dez. 1978.

KILLICK-KENDRICK, R. The biology and control of Phlebotomine sand flies. **Clinics in Dermatology**, v. 17, n. 3, p. 279–289, 1999.

KILLICK-KENDRICK, R.; KILLICK-KENDRICK, M. Honeydew of aphids as a source of sugar for Phlebotomus ariasi. **Medical and Veterinary Entomology**, v. 1, n. 3, p. 297–302, 1987.

KIMURA, A. et al. Allosteric Properties, Substrate Specificity, and Subsite Affinities of Honeybee a-Glucosidase I. **J. Biochem**, v. 107, p. 762–768, 1990.

KLOWDEN, M. J. Metabolic Systems. In: KLOWDEN, M. J. (Ed.). . **Physiological Systems in Insects**. Third ed. [s.l.] Academic Press, 2013. p. 305–364.

KOSHLAND, D. E. Stereochemistry and the Mechanism of Enzymatic Reactions. **Biol. Rev.**, v. 28, n. February, p. 416–436, 1953.

KUBOTA, M. et al. Localization of alpha-glucosidases I, II, and III in organs of European honeybees, Apis mellifera L., and the origin of alpha-glucosidase in honey. **Bioscience, biotechnology, and biochemistry**, v. 68, n. 11, p. 2346–2352, 2004.

KURIKI, T.; IMANAKA, T. The Concept of the Alpha-Amylase Family: Structural Similarity and Common Catalytic Mechanism. **Journal of Bioscience and Bioengineering**, v. 87, n. 5, p. 557–565, 1999.

LARSBRINK, J. et al. Structural and enzymatic characterization of a glycoside hydrolase family 31 α-xylosidase from Cellvibrio japonicus involved in xyloglucan saccharification. **Biochemical Journal**, v. 436, n. 3, p. 567–580, 2011.

LAWYER, P. et al. Laboratory colonization and mass rearing of phlebotomine sand flies (Diptera, Psychodidae). **Parasite**, v. 24, p. 42, 2017.

LEE, S. S.; HE, S.; WITHERS, S. G. Identification of the catalytic nucleophile of the Family 31 alpha-glucosidase from Aspergillus niger via trapping of a 5-fluoroglycosyl-enzyme intermediate. **The Biochemical journal**, v. 359, p. 381–386, 2001.

LESLIE, G.; BARRETT, M.; BURCHMORE, R. Leishmania mexicana: Promastigotes migrate through osmotic gradients. **Experimental Parasitology**, v. 102, n. 2, p. 117–120, 2002.

LOVERING, A. L. et al. Mechanistic and structural analysis of a family 31 a-glycosidase and its glycosyl-enzyme intermediate. **Journal of Biological Chemistry**, v. 280, n. 3, p. 2105–2115, 2005.

LYDA, T. A. et al. A unique, highly conserved secretory invertase is differentially expressed by promastigote developmental forms of all species of the human pathogen, Leishmania. **Molecular and Cellular Biochemistry**, v. 404, n. 1–2, p. 53–77, 2015.

MACGREGOR, E. A.; JANEČEK, Š.; SVENSSON, B. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. **Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology**, v. 1546, n. 1, p. 1–20, 2001.

MACHIUS, M.; WIEGAND, G.; HUBER, R. Crystal Structure of Calcium-depletedBacillus licheniformisα-amylase at 2.2 Å Resolution. **Journal of Molecular Biology**, v. 246, n. 4, p. 545–559, 1995.

MACVICKER, J. A. K. et al. Honeydew sugars in wild-caught Italian phlebotomine sandflies

(Diptera: Psychodidae) as detected by high performance liquid chromatography. **Bulletin of Entomological Research**, v. 80, n. 3, p. 339–344, 1990.

MARINOTTI, O.; DE B&O, M.; MOREIRU, C. K. Apyrase and a-glucosidase in the Salivary Glands of Aedes ulbo~ictus. **Biochem. Physiol**, v. 113, n. 4, p. 675–679, 1996.

MARINOTTI, O.; JAMES, A. A. An a-glucosidase in the salivary glands of the vector mosquito, Aedes aegypti. **Insect Biochemistry**, v. 20, n. 6, p. 619–623, 1990.

MATSUURA, Y. et al. Structure and possible catalytic residues of Taka-amylase A. **Journal of Biochemistry**, v. 95, n. 3, p. 697–702, 1984.

MCCONVILLE, M. J. et al. The glycoinositol phospholipids of Leishmania mexicana promastigotes. Evidence for the presence of three distinct pathways of glycolipid biosynthesis. **Journal of Biological Chemistry**, v. 268, n. 21, p. 15595–15604, 1993.

MCGWIRE, B. S.; SATOSKAR, A. R. Leishmaniasis: Clinical syndromes and treatment. **Qjm**, v. 107, n. 1, p. 7–14, 2014.

MEHRABADI, M. et al. Sunn pest, Eurygaster integriceps Putton (Hemiptera: Scutelleridae), digestive α-amylase, α-glucosidase and β-glucosidase. **Journal of Asia-Pacific Entomology**, v. 12, n. 2, p. 79–83, 2009.

MINISTÉRIO DA SAÚDE, 2006. **Manual de vigilância e controle de leishmaniose visceral.** 1ª. ed. Brasília, 2006. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/manual_vigilancia_controle_leishmaniose_visceral.pdf. Acesso em 20 julho 2018.

MINISTÉRIO DA SAÚDE, 2018 (1). **Leishmaniose Tegumentar - Gráficos e Mapas.** Disponível em: <http://portalms.saude.gov.br/saude-de-a-z/leishmaniose-tegumentar-americana-lta/11328-situacao-epidemiologica-dados>. Acesso em 19 julho 2018.

MINISTÉRIO DA SAÚDE, 2018 (2). **LV - Gráficos e Mapas.** Disponível em: <http://portalms.saude.gov.br/saude-de-a-z/leishmaniose-visceral/11334-situacao-epidemiologica-dados>. Acesso em 19 julho 2018.

MOORE, J. S. et al. Honeydew sugars in wild-caught Phlebotomus ariasi detected by high performance liquid chromatography (HPLC) and gas chromatography (GC). **Medical and Veterinary Entomology**, v. 1, n. 4, p. 427–434, 1987.

MORAES, C. S. et al. Relationship between digestive enzymes and food habit of Lutzomyia longipalpis (Diptera: Psychodidae) larvae: Characterization of carbohydrases and digestion of microorganisms. **Journal of Insect Physiology**, v. 58, p. 1136–1145, 2012.

MORAES, C. S. et al. Second Blood Meal by Female Lutzomyia longipalpis: Enhancement by Oviposition and Its Effects on Digestion, Longevity, and Leishmania Infection. **BioMed Research International**, 2018.

MOREIRA-FERRO, C. K.; MARINOTTI, O.; BIJOVSKY, A T. Morphological and biochemical analyses of the salivary glands of the malaria vector, Anopheles darlingi. **Tissue & cell**, v. 31, n. 3, p. 264–273, 1999.

MURRAY, H. W. et al. **Advances in leishmaniasis**. Lancet. **Anais**...2005

MURY, F. B. et al. Alpha-glucosidase promotes hemozoin formation in a blood-sucking bug: An evolutionary history. **PLoS ONE**, v. 4, n. 9, p. e6966, 2009.

NICHOLS, B. L. et al. The maltase-glucoamylase gene: Common ancestry to sucrase-isomaltase with complementary starch digestion activities. **Proceedings of the National Academy of Sciences**, v. 100, n. 3, p. 1432–1437, 2003.

NISHIMOTO, M. et al. Purification and substrate specificity of honeybee, Apis mellifera L., alpha-glucosidase III. **Bioscience, biotechnology, and biochemistry**, v. 65, n. 7, p. 1610–1616, 2001.

OKUYAMA, M. et al. Carboxyl group of residue Asp647 as possible proton donor in catalytic reaction of α-glucosidase from Schizosaccharomyces pombe. **European Journal of Biochemistry**, v. 268, n. 8, p. 2270–2280, 2001.

OLIVEIRA, F. et al. Sand flies, Leishmania, and transcriptome-borne solutions. **Parasitology International**, v. 58, p. 1–5, 2009.

OLIVIERA, J. S.; MELO, M. N.; GONTIJO, N. F. A sensitive method for assaying chemotaxic responses of Leishmania promastigotes. **Experimental Parasitology**, v. 96, n. 3, p. 187–189, 2000.

OPOTA, O. et al. Identification and characterization of the receptor for the Bacillus sphaericus binary toxin in the malaria vector mosquito, Anopheles gambiae. **Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology**, v. 149, n. 3, p. 419–427, 2008.

PACE, D. Leishmaniasis. **Journal of Infection**, v. 69, n. S1, p. S10–S18, 2014.

PASCOA, V. et al. Aedes aegypti peritrophic matrix and its interaction with heme during blood digestion. **Insect Biochemistry and Molecular Biology**, v. 32, n. 5, p. 517–523, 2002.

PIMENTA, P. F. P. et al. Stage-specific adhesion of Leishmania promastigotes to the sandfly midgut. **Science**, v. 256, n. 5065, p. 1812–1815, 1992.

QIAN, M.; HASER, R.; PAYAN, F. Structure and Molecular Model Refinement of Pig Pancreatic α-Amylase at 2·1 Å Resolution. **Journal of Molecular Biology**, v. 231, n. 3, p. 785–799, 1993.

RAMALHO-ORTIGÃO, J. M. et al. Characterization of a blood activated chitinolytic system in the midgut of the sand fly vectors Lutzomyia longipalpis and Phlebotomus papatasi. **Insect Molecular Biology**, v. 14, n. 6, p. 703–712, 2005.

RAMASUBBU, N. et al. Structure of Human Salivary α-Amylase at 1.6 Å Resolution: Implications for its Role in the Oral Cavity. **Acta Crystallographica Section D Biological Crystallography**, v. 52, n. 3, p. 435–446, 1996.

RANGEL, E. F. et al. Biology of Lutzomyia intermedia Lutz & Neiva, 1912 and Lutzomyia longipalpis Lutz & Neiva, 1912 (Diptera, Psychodidae), under experimental conditions. I. Feeding aspects of larvae and adults. **Memorias do Instituto Oswaldo Cruz**, v. 81, n. 4, p. 431–8, 1986.

READY, P. D. Factors affecting egg production of laboratory-bred Lutzomyia longipalpis (Diptera: Psychodidae). **Journal of Medical Entomology**, v. 16, n. 5, p. 413–423, 1979.

ROGERS, M. E. et al. Transmission of cutaneous leishmaniasis by sand flies is enhanced by regurgitation of fPPG. **Nature**, v. 430, n. 6998, p. 463–467, 2004.

ROGERS, M. E.; BATES, P. A. Leishmania manipulation of sand fly feeding behavior results in enhanced transmission. **PLoS Pathogens**, v. 3, n. 6, p. e91, 2007.

ROGERS, M. E.; CHANCE, M. L.; BATES, P. A. The role of promastigote secretory gel in

the origin and transmission of the infective stage of Leishmania mexicana by the sandfly Lutzomyia longipalpis. **Parasitology**, v. 124, n. 5, p. 495–507, 2002.

ROMÃO, T. P. et al. A second independent resistance mechanism to Bacillus sphaericus binary toxin targets its α-glucosidase receptor in Culex quinquefasciatus. **FEBS Journal**, v. 273, n. 7, p. 1556–1568, 2006.

RUDIN, W.; HECKER, H. Functional morphology of the midgut of a sandfly as compared to other hematophagous nematocera. **Tissue and Cell**, v. 14, n. 4, p. 751–758, 1982.

SACKS, D.; KAMHAWI, S. Molecular Aspects of Parasite-Vector and Vector-Host Interactions in Leishmaniasis. **Annual Review of Microbiology**, v. 55, p. 453–483, 2001.

SACKS, D. L. Stage-specific binding of Leishmania donovani to the sand fly vector midgut is regulated by conformational changes in the abundant surface lipophosphoglycan. **Journal of Experimental Medicine**, v. 181, n. 2, p. 685–697, 1995.

SAMIE, M. et al. GLYCOSIDASE ACTIVITY IN THE SANDFLY PHLEBOTOMUS PAPATASI. **Biochem. Physiol**, v. 96, n. 3, p. 577–579, 1990.

SANCHEZ CLEMENTE, N. et al. Bartonella bacilliformis: a systematic review of the literature to guide the research agenda for elimination. **PLoS neglected tropical diseases**, v. 6, n. 10, p. 1819, 2012.

SANTOS, V. C. et al. The physiology of the midgut of Lutzomyia longipalpis (Lutz and Neiva 1912): pH in different physiological conditions and mechanisms involved in its control. **Journal of Experimental Biology**, v. 211, n. 17, p. 2792–2798, 2008.

SANTOS, V. C. et al. Host modulation by a parasite: How Leishmania infantum modifies the intestinal environment of Lutzomyia longipalpis to favor its development. **PLoS ONE**, v. 9, n. 11, 2014.

SCHLEIN, Y.; JACOBSON, R. L.; SHLOMAI, J. Chitinase secreted by Leishmania functions in the sandfly vector. **Proceedings of the Royal Society B: Biological Sciences**, v. 245, n. 1313, p. 121–126, 1991.

SCHLEIN, Y.; MULLER, G. Assessment of plant tissue feeding by sand flies (Diptera: Psychodidae) and mosquitoes (Diptera: Culicidae). **Journal of Medical Entomology**, v. 32, n. 6, p. 882–887, 1995.

SECCOMBE, A. K.; READY, P. D.; HUDDLESTON, L. M. A Catalogue of Old World Phlebotomine Sandflies (Diptera: Psychodidae, Phlebotominae). **Occasional papers on Systematic Entomology**, v. no. 8, p. 717, 1993.

SECUNDINO, N. et al. Proteophosphoglycan confers resistance of Leishmania major to midgut digestive enzymes induced by blood feeding in vector sand flies. **Cellular Microbiology**, v. 12, n. 7, p. 906–918, 2010.

SECUNDINO, N. F. C. et al. Lutzomyia longipalpis Peritrophic Matrix: Formation, Structure, and Chemical Composition. **EVOLUTION J. Med. Entomol**, v. 42, n. 6, p. 928–938, 2005.

SERAFIM, T. D. et al. Sequential blood meals promote Leishmania replication and reverse metacyclogenesis augmenting vector infectivity. **Nature Microbiology**, v. 3, n. 5, p. 548–555, 19 maio 2018.

SHARMA, U.; SINGH, S. Insect vectors of Leishmania: Distribution, physiology and their control. **Journal of Vector Borne Diseases**, v. 45, p. 255–272, 2008.

SILVA, C. P. et al. Occurrence of midgut perimicrovillar membranes in paraneopteran insect orders with comments on their function and evolutionary significance. **Arthropod Structure and Development**, v. 33, n. 2, p. 139–148, 2004.

SILVA, C. P.; TERRA, W. R. An alpha-glucosidase from perimicrovillar membranes of Dysdercus peruvianus (Hemiptera: Pyrrhocoridae) midgut cells. Purification and properties. **Insect Biochemistry and Molecular Biology**, v. 25, n. 4, p. 487–494, 1995.

SILVA, J. R. et al. Perimicrovillar membranes promote hemozoin formation into Rhodnius prolixus midgut. **Insect Biochemistry and Molecular Biology**, v. 37, n. 6, p. 523–531, 2007.

SIM, L. et al. Human Intestinal Maltase-Glucoamylase: Crystal Structure of the N-Terminal Catalytic Subunit and Basis of Inhibition and Substrate Specificity. **Journal of Molecular Biology**, v. 375, n. 3, p. 782–792, 2008.

SOARES, R. P. et al. Differential midgut attachment of Leishmania (Viannia) braziliensis in the sand flies Lutzomyia (Nyssomyia) whitmani and Lutzomyia (Nyssomyia) intermedia. **Journal of Biomedicine and Biotechnology**, 2010.

SOARES, R. P. P. et al. Leishmania chagasi: Lipophosphoglycan characterization and binding to the midgut of the sand fly vector Lutzomyia longipalpis. **Molecular and Biochemical Parasitology**, v. 121, n. 2, p. 213–224, 2002.

SOARES, R. P. P.; TURCO, S. J. Lutzomyia longipalpis (Diptera: Psychodidae: Phlebotominae): a review. **Annals of the Brazilian Academy of Sciences**, v. 75, n. 3, p. 301–330, 2003.

SOUZA-NETO, J. A. et al. Sugar digestion in mosquitoes: Identification and characterization of three midgut α-glucosidases of the neo-tropical malaria vector Anopheles aquasalis (Diptera: Culicidae). **Comparative Biochemistry and Physiology - A Molecular and Integrative Physiology**, v. 147, n. 4, p. 993–1000, 2007.

STAM, M. R. et al. Dividing the large glycoside hydrolase family 13 into subfamilies: Towards improved functional annotations of α-amylase-related proteins. **Protein Engineering, Design and Selection**, v. 19, n. 12, p. 555–562, 2006.

STROBL, S. et al. Crystal structure of yellow meal worm α-amylase at 1.64 Å resolution. **Journal of Molecular Biology**, v. 878, n. 3, p. 617–628, 1998.

TAKEWAKI, S. et al. Substrate Specificity and Subsite Affinities of Honeybee α -Glucosidase II. **Bioscience, Biotechnology, and Biochemistry**, v. 57, n. 9, p. 1508–1513, 1993.

TELLERIA, E. L. et al. Constitutive and blood meal-induced trypsin genes in Lutzomyia longipalpis. **Archives of Insect Biochemistry and Physiology**, v. 66, n. 2, p. 53–63, 2007.

TELLERIA, E. L. et al. Trypsin-like serine proteases in Lutzomyia longipalpis - expression, activity and possible modulation by Leishmania infantum chagasi. **PLoS ONE**, v. 5, n. 5, p. e10697, 2010.

TERRA, W. R. et al. Insect digestive enzymes : properties , compartmentalization and function. **Comparative Biochemistry and Physiology Part B: Comparative Biochemistry**, v. 109, n. 1, p. 1–62, 1994.

TERRA, W. R. et al. Digestive enzymes. In: LEHANE, M. J.; BILLINGSLEY, P. F. (Eds.). . **Biology of the Insect Midgut**. London: Chapman & Hall, 1996. p. 153–194.

TERRA, W. R.; FERREIRA, C. Biochemistry of Digestion. In: **Biochemistry**. São Paulo: University of São Paulo, 2005.

TERRA, W. R.; FERREIRA, C. Biochemistry and Molecular Biology of Digestion. In: **Insect Molecular Biology and Biochemistry**. [s.l.] Elsevier B.V., 2012. p. 365–418.

TESH, R. B. The Genus Phlebovirus and its Vectors. **Annual Review of Entomology**, v. 33, n. 1, p. 169–181, jan. 1988.

THEODOS, C. M.; RIBEIRO, J. M. C.; TITUS, R. G. Analysis of enhancing effect of sand fly saliva on Leishmania infection in mice. **Infection and Immunity**, 1991.

THOMPSON, S. N. Trehalose - The Insect "Blood" Sugar. In: **Advances in Insect Physiology**. [s.l.] Elsevier Ltd, 2003. v. 31p. 205–285.

TITUS, R.; RIBEIRO, J. Salivary gland lysates from the sand fly Lutzomyia longipalpis enhance Leishmania infectivity. **Science**, v. 239, n. 4845, p. 1306–1308, 1988.

TROMBETTA, E. S.; FLEMING, K. G.; HELENIUS, A. Quaternary and domain structure of glycoprotein processing glucosidase II. **Biochemistry**, v. 40, n. 35, p. 10717–10722, 2001.

TROMBETTA, E. S.; SIMONS, J. F.; HELENIUS, A. Endoplasmic reticulum glucosidase II is composed of a catalytic subunit, conserved from yeast to mammals, and a tightly bound noncatalytic HDEL- containing subunit. **Journal of Biological Chemistry**, v. 271, n. 44, p. 27509–27516, 1996.

UITDEHAAG, J. C. M. et al. X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the α-amylase family. **Nature Structural Biology**, v. 6, n. 5, p. 432–436, 1999.

VALE, V. F. et al. Carbohydrate digestion in Lutzomyia longipalpis' larvae (Diptera - Psychodidae). **Journal of Insect Physiology**, v. 58, n. 10, p. 1314–1324, 2012.

VAN DER MAAREL, M. J. E. C. et al. Properties and applications of starch-converting enzymes of the alpha-amylase family. **Journal of Biotechnology**, v. 94, p. 137–155, 2002.

VAN GRIENSVEN, J.; DIRO, E. Visceral Leishmaniasis. **Infectious Disease Clinics of North America**, v. 26, p. 309–322, 2012.

VOLF, P.; TESAR, P. Salivary proteins and glycoproteins in phlebotomine sandflies of various species , sex and age. **Medical and Veterinary Entomology**, v. 14, p. 251–256, 2000.

VOLF, P.; VOLFOVA, V. Establishment and maintenance of sand fly colonies. **Journal of Vector Ecology**, v. 36, p. S1–S9, 2011.

WARBURG, A. et al. Saliva of Lutzomyia longipalpis sibling species differs in its composition and capacity to enhance leishmaniasis. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 345, n. 1312, p. 223–230, 1994.

WARBURG, A. The structure of the female sand fly (Phlebotomus papatasi) alimentary canal. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 102, n. 2, p. 161–166, 2008.

WATERHOUSE, D. F. Digestion in Insects. **Annual Review of Entomology**, v. 2, p. 1–18, 1957.

WHO. WHO: Weekly epidemiological record: Global leishmaniasis update, 2006-2015, a turning point in leishmaniasis surveillance. **World Health organization**, v. 92, n. 38, p. 557–572, 2017.

WIGGLESWORTH, V. B. Memoirs: The Formation of the Peritrophic Membrane in Insects,

with Special Reference to the Larvae of Mosquitoes. **Journal of Cell Science**, v. s2, p. 593–616, 1930.

YAMADA, T. et al. Fat body glycogen serves as a metabolic safeguard for the maintenance of sugar levels in *Drosophila*. **Development**, v. 145, n. 6, p. dev158865, 2018.

ZHAI, L. et al. Crystal structure of glycogen debranching enzyme and insights into its catalysis and disease-causing mutations. **Nature Communications**, v. 7, p. 1–11, 2016.

ZHANG, Q. et al. Analyses of alpha-amylase and alpha-glucosidase in the malaria vector mosquito, Anopheles gambiae, as receptors of Cry11Ba toxin of Bacillus thuringiensis subsp. jegathesan. **Insect Biochemistry and Molecular Biology**, v. 43, n. 10, p. 907–915, 2013.

ZHENG, L. et al. Two genes encoding midgut-specific maltase-like polypeptides from anopheles gambiae. **Experimental Parasitology**, v. 81, n. 3, p. 272–283, 1995.

ZIBAEE, A.; BANDANI, A. R.; RAMZI, S. Enzymatic properties of α- and β-glucosidases extracted from midgut and salivary glands of rice striped stem borer, Chilo suppressalis Walker (Lepidoptera: Pyralidae). **Comptes Rendus - Biologies**, v. 332, p. 633–641, 2009.

ZIJLSTRA, E. E. et al. Post-kala-azar dermal leishmaniasis. **Lancet Infectious Diseases**, v. 3, n. 2, p. 87–98, 2003.