



Characterization of Glycoside Hydrolase Families 13 and 31 Reveals Expansion and Diversification of α -Amylase Genes in the Phlebotomine *Lutzomyia longipalpis* and Modulation of Sandfly Glycosidase Activities by *Leishmania* Infection

OPEN ACCESS

Samara Graciane da Costa-Latgé¹, Paul Bates², Rod Dillon^{2,3} and Fernando Ariel Genta^{1,3*}

Edited by:

Jose Eduardo Serrão,
Universidade Federal de Viçosa, Brazil

Reviewed by:

Marie Sofie Moeller,
Technical University of Denmark,
Denmark
Ali R. Bandani,
University of Tehran, Iran

*Correspondence:

Fernando Ariel Genta
genta@ioc.fiocruz.br;
gentafernando@gmail.com

Specialty section:

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

Received: 30 November 2020

Accepted: 26 February 2021

Published: 09 April 2021

Citation:

da Costa-Latgé SG, Bates P,
Dillon R and Genta FA (2021)
Characterization of Glycoside
Hydrolase Families 13 and 31 Reveals
Expansion and Diversification
of α -Amylase Genes
in the Phlebotomine *Lutzomyia*
longipalpis and Modulation of Sandfly
Glycosidase Activities by *Leishmania*
Infection. *Front. Physiol.* 12:635633.
doi: 10.3389/fphys.2021.635633

¹ Laboratory of Insect Biochemistry and Physiology, Oswaldo Cruz Institute, Fundação Oswaldo Cruz, Rio de Janeiro, 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, Brazil

Sugar-rich food sources are essential for sandflies to meet their energy demands, achieving more prolonged survival. The digestion of carbohydrates from food is mainly realized by glycoside hydrolases (GH). To identify genes coding for α -glycosidases and α -amylases belonging to Glycoside Hydrolase Family 13 (GH13) and Glycoside Hydrolase Family 31 (GH31) in *Lutzomyia longipalpis*, we performed an HMMER search against its genome using known sequences from other dipteran species. The sequences retrieved were classified based on BLASTP best hit, analysis of conserved regions by alignment with sequences of proteins with known structure, and phylogenetic analysis comparing with orthologous proteins from other dipteran species. Using RT-PCR analysis, we evaluated the expression of GH13 and GH31 genes, in the gut and rest of the body of females, in four different conditions: non-fed, sugar-fed, blood-fed, and *Leishmania mexicana* infected females. *L. longipalpis* has GH13/31 genes that code for enzymes involved in various aspects of sugar metabolism, as carbohydrate digestion, storage, and mobilization of glycogen reserves, proteins involved in transport, control of N-glycosylation quality, as well as others with a putative function in the regulation of myogenesis. These proteins are representatives of GH13 and GH31 families, and their roles seem to be conserved. Most of the enzymes seem to be active with conserved consense sequences, including the expected catalytic residues. α -amylases also demonstrated the presence of calcium and chloride binding sites. *L. longipalpis* genome shows an expansion in the α -amylase gene family, with two clusters. In contrast, a retraction in the number of α -glucosidases occurred. The expansion of

α -amylases is probably related to the specialization of these proteins for different substrates or inhibitors, which might correlate with the higher diversity of plant foods available in the natural habitat of *L. longipalpis*. The expression of α -glucosidase genes is higher in blood-fed females, suggesting their role in blood digestion. Besides that, in blood-fed females infected with the parasite *Leishmania mexicana*, these genes were also modulated. Glycoside Hydrolases from families 13 and 31 are essential for the metabolism of *L. longipalpis*, and GH13 enzymes seem to be involved in the interaction between sandflies and *Leishmania*.

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

INTRODUCTION

For adult sandflies, the energy requirements are provided primarily by sugar meals obtained from different sources, for both females and males. These sugar sources have been described for different sandfly species, depending on their preferred site for development. Desert and savanna regions are the preferred habitats for the genus *Phlebotomus* (Old World) and tropical forests for the genus *Lutzomyia* (New World). Sandflies commonly feed on plant tissues, fruit juices, nectars from flowers, honeydews excreted by aphids and coccids, among other sources, with different sugar compositions (Killick-Kendrick and Killick-Kendrick, 1987; Macvicker et al., 1990; Cameron et al., 1995). For females feeding on blood for egg maturation, 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 of plant and animal cells (Moraes et al., 2012; Vale et al., 2012). Thus, glycosidases play a crucial role in the digestion of larvae and adult sandflies.

Alpha-amylases and α -glucosidases are vital for sand fly adults since their substrates are common components of plant tissues and secretions. Starch (substrate for amylase) occurs in leaves and fruits, and sucrose (substrate for α -glucosidase) is the main component of sap and nectar, and, besides other sugars, is also found in honeydew (Killick-Kendrick and Killick-Kendrick, 1987; Macvicker et al., 1990). In larvae, these enzymes may play an essential role in glycogen digestion, the reserve carbohydrate in fungi (Moraes et al., 2012). Based on similarity, gene sequences of these enzymes have been grouped into families. In insects,

these enzymes are described in glycoside hydrolases families 13 (GH13) and 31 (GH31) (Cantarel et al., 2009).

Alpha-amylases play an important role in the breakdown of internal α -1,4-glucan bonds of starch and related polysaccharides, such as amylose and amylopectin (Janeček et al., 2014). They are soluble enzymes participating in the process of initial digestion in insects. GH13 α -glucosidases are exoglycosidases acting generally upon α -glucosidic bonds, including α -1,4 bonds, and participate in hydrolysis of aryl glycosides, disaccharides and oligosaccharides with varying efficiency (Terra and Ferreira, 1994). GH31 α -glucosidases are intracellular proteins participating in the glycogen metabolism and glycoconjugate biosynthesis and processing (Lovering et al., 2005).

In insects, the glucose obtained after hydrolysis of sugar-rich food is uptaken by enterocytes through specific membrane transporters. Then, glucose is used to produce energy through glycolysis and related metabolic pathways or converted to trehalose, the circulating sugar in the hemolymph. When circulating trehalose is in excess, glucose is stored as glycogen or converted into triacyl-glycerides in several tissues. Insects have to expend energy continuously, and adaptation to food availability changes is a central challenge for survival. In a starving condition, they must live on reserves (glycogen or triacyl-glycerides) accumulated in periods of food abundance. The flight, for example, is a condition of extreme energy demand (Thompson, 2003; Arrese and Soulages, 2010). In insects, some of the enzymes involved in the storage and mobilization of glycogen, as the 1,4- α -glucan branching enzyme, glycogen debranching enzyme (GDE), and lysosomal α -glucosidase are also members of GH13 or GH31 families.

The GH13 is the largest family of glycoside hydrolases, presenting a wide variety of enzymes with more than 30 different specificities. These enzymes may act on several substrates, by hydrolyzing α -glycosidic bonds, generating α -anomeric mono/oligosaccharides, or catalyzing α -glycosidic bond formation by transglycosylation (Kuriki and Imanaka, 1999). GH13 family includes representatives of α -amylases, α -glucosidases, α -1,4-glucan branching enzymes, pullulanases, cyclodextrin glucanotransferases, 4- α -glucanotransferases (the domain of glycogen debranching enzyme), oligo- α -1,6-glucosidases, amino acid transporters, besides others.

Abbreviations: GH, glycoside hydrolases; CSRs, conserved regions; glycosidase NET37, nuclear envelope transmembrane glycosidase 37; SF, sugar-fed; BF, blood-fed; NF, newly hatched females; RB, rest of the body; NBATHc, neutral and basic amino acid transport protein heavy chain; CD98hc, cluster of differentiation 98 heavy chain; GT, gluconotransferase; GDE, glycogen debranching enzymes; LIAmy, *Lutzomyia longipalpis* α -amylase; LIAGlu, *Lutzomyia longipalpis* α -glucosidase; LINBATHc, *Lutzomyia longipalpis* neutral and basic amino acid transport protein heavy chain; LICD98hc, *Lutzomyia longipalpis* cluster of differentiation 98 heavy chain; LIAGB, *Lutzomyia longipalpis* 1,4-alpha-glucan-branching enzyme; LIGDE, *Lutzomyia longipalpis* glycogen debranching enzyme; LIglyMyo, *Lutzomyia longipalpis* glycosidase NET37; LILysAglu, *Lutzomyia longipalpis* lysosomal alpha-glucosidase; LINAGlu, *Lutzomyia longipalpis* neutral alpha-glucosidase; PpAgluA2, *Phlebotomus papatasi* α -glucosidase A2; CqMal1, *Culex quinquefasciatus* α -glucosidase 1.

The GH31 is also a diverse group, with a range of different hydrolytic activities. Members of this family are capable of cleaving a terminal carbohydrate moiety from substrates varying considerably in size, from disaccharides to complex reserve polysaccharides such as starch, glycogen, or glycoproteins (Ernst et al., 2006). α -glucosidases, α -xylosidases, α -galactosidases, α -mannosidases, 6- α -glucosyltransferases, and 3- α -isomaltosyltransferases, besides others, represent the hydrolytic activities that are 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) (Hoefsloot et al., 1988), and in glycoprotein processing (neutral α -glucosidase) in the endoplasmic reticulum (Trombetta et al., 1996, 2001).

Enzymes classified in the GH13 or GH31 families share a typical tertiary structure in their catalytic domain, the $(\beta/\alpha)_8$ -barrel fold. Members from the GH13 family present a conserved central nucleus composed by the catalytic domain A, domain B, which is a variable-length loop located between sheet β_3 and α_3 helix of the $(\beta/\alpha)_8$ -barrel, and domain C, a Greek key motif in the C-terminal position (Ramasubbu et al., 1996; Janecek, 1997; Kanai et al., 2001; Abad et al., 2002). The conserved regions (CSRs) described in sheets β_2 , β_3 , β_4 , β_5 , β_7 , β_8 and in the loop 3, were defined to help in the assignment of the correct enzymatic specificity to the members of this protein family. The catalytic triad aspartate, glutamate, and aspartate are located in the β_4 , β_5 , and β_7 sheets, respectively (Janecek, 1997). Also, a fourth conserved residue was determined as an arginine positioned two residues preceding the catalytic nucleophile (MacGregor et al., 2001). In TAKA-amylase A from *Aspergillus oryzae* (Matsuura et al., 1984), one of the first GH13 structures to be elucidated, the Asp206 inside the GLRID domain (β_4) acts as a catalytic nucleophile and Glu230 (β_5) as the general acid/base. A second aspartate (Asp297, β_7) contributes to stabilizing the transition state of the substrate during catalysis. It also maintains the glutamate with the correct protonation state for activity (Matsuura et al., 1984; Uitdehaag et al., 1999; Svensson and Macgregor, 2001). Insects α -amylases are typically calcium-dependent, and chloride may also work as an activator (Terra and Ferreira, 2012). The calcium-binding site has one aspartate residue in the loop 3 region, and asparagine and histidine positioned at the β_3 and β_4 sheets, respectively (Boel et al., 1990; Machius et al., 1995; Janecek, 1997). The chloride-binding site has a conserved arginine residue preceding the catalytic nucleophile, one asparagine residue placed two residues prior to the catalytic aspartate in the β_7 sheet, and an Arg/Lys residue localized inside the variable domain RVMSSY (Janecek, 1997; D'Amico et al., 2000).

In the GH31 family, catalytic residues consist in one aspartate (catalytic nucleophile) inside the conserved domain WIDMNE (β_4 sheet), and a second aspartate in the β_6 sheet (general acid/base residue) (Lee et al., 2001; Lovering et al., 2005; Ernst et al., 2006; Sim et al., 2008). The sequence similarity is not evident between GH13 and GH31 members, but both families present a conserved aspartate as a catalytic nucleophile in the β_4 sheet (Janecek et al., 2007). Most members of the GH31 family are multi-domain proteins, but the specific function of the accessory domains is generally unknown.

Phlebotomine sandflies are considered of medical importance, mainly because about 10% of their known species are vectors of some pathogen-caused disease (Akhoundi et al., 2016). Among the diseases caused by pathogens transmitted by phlebotomine sandflies, we can highlight the leishmaniasis. These diseases 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 (Killick-Kendrick, 1999; Sharma and Singh, 2008; Pace, 2014). In the Americas, *Lutzomyia longipalpis* is the natural vector of *Leishmania infantum*, the causative agent of visceral leishmaniasis, the most severe form of the disease (Soares and Turco, 2003; Murray et al., 2005; Akhoundi et al., 2016). In the insect vector, the *Leishmania* develops exclusively inside the gut. It is in direct contact with the digestive enzymes secreted by phlebotomines, especially proteases but also glycosidases, as alpha-amylases and alpha-glucosidases from GH13. *L. longipalpis* is a permissive vector, supporting the development of different *Leishmania* species as *L. mexicana*, *L. tropica*, *L. major*, and *L. amazonensis*, besides *L. infantum* (Soares and Turco, 2003; Kamhawi, 2006). *L. longipalpis* permissivity allows the study under laboratory conditions of the interaction of different *Leishmania* species with the digestive system of this vector.

Recently, we demonstrated the presence of at least four different α -glucosidase activities in *L. longipalpis*, with different specificities and putative roles in the digestion of plant or blood sugars (da Costa et al., 2019). However, the molecular identities of these enzymes are still unknown. In an initial effort to clarify this question, we revealed the presence of genes coding for proteins belonging to GH13 and GH31 in *L. longipalpis*. We compared *L. longipalpis* protein sequences available in Vector Base with the proteins described for other Dipterans from suborder Brachycera (genus *Drosophila*), suborder Nematocera (genus *Aedes*, *Culex*, *Anopheles*), and other members of the family Psychodidae (genus *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 21 genes that belong to the GH13, described as α -amylases, α -glucosidases, amino acid transport proteins (heavy chain), 1,4- α -glucan branching enzyme and glycogen debranching enzyme; and six genes in the GH31 family, described as glycosidases NET37 (Nuclear envelope transmembrane glycosidase 37), lysosomal α -glucosidase and neutral α -glucosidase (α subunit). These enzymes are involved in sugar metabolism, storage, and mobilization of glycogen, protein transport, N-glycosylation quality control, and myogenesis regulation.

MATERIALS AND METHODS

Chemicals

The TRI Reagent® (Cat: T9424) was purchased from Sigma-Aldrich Company (St. Louis, MO, United States). Qubit ssDNA Assay Kit (cat: Q10212) and SYBR™ Green PCR Master Mix (cat: 4309155) were obtained from Thermo Fisher Scientific (Waltham, MA, 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.

Search for GH13 and GH31 Sequences in *L. longipalpis* Genome

Sequences from GH13 and GH31 families from five different insect species – *Aedes aegypti* (AAEL006719, AAEL000647, AAEL008502, AAEL010602, AAEL009838, AAEL017226, AAEL022548), *Culex quinquefasciatus* (CPIJ005064, CPIJ007333, CPIJ011854, CPIJ001201, CPIJ009306), *Anopheles gambiae* (AGAP012230, AGAP010428, AGAP001200, AGAP001534, AGAP000862), *Drosophila melanogaster* (FBpp0311600, FBpp0071525, FBpp0070061, FBpp0084221), and *Drosophila ananassae* (FBpp0115773, FBpp0115136, FBpp0345570) – were retrieved from Vector Base¹, Fly base², and NCBI³ (Cantarel et al., 2009; Gabriško, 2013). These sequences were employed to perform ClustalW alignment and HMMER search in the Vector Base database to find similar sequences in the *L. longipalpis* genome (conserved domains on *Lutzomyia longipalpis*, LlonJ1.4, last updated 27 June 2017). Sequences recovered from *L. longipalpis* 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 Apollo annotation tool (Dunn et al., 2019) and ClustalW (Thompson et al., 1994). Sequences were considered complete when initial methionine, correct exon/intron junction, stop codon were identified, and exon structures were complete based on the alignment with orthologous genes.

After sequences identification, characteristics like putative signal peptide (Signalp) (Petersen et al., 2011), transmembrane domain (THMM) (Krogh et al., 2001), GPI-anchor (Big-PI predictor) (Eisenhaber et al., 1998), protein subcellular localization (DeepLoc) (Almagro Armenteros et al., 2017), N-glycosylation (NetNGlyc) (Gupta et al., 2002), O-glycosylation (NetoGlyc) (Steenftoft et al., 2013), molecular mass and isoelectric point (pI) (Compute pI/Mw) (Bjellqvist et al., 1994; Gasteiger et al., 2005) were predicted *in silico*. The N-glycosylation sites were not considered if predicted inside transmembrane or cytoplasmic domains of membrane proteins. Genes and proteins representative structures were drawn using PROSITE MyDomains tool⁴ (Hulo et al., 2008).

The InterPro tool (Finn et al., 2017) was used to predict catalytic sites and, for α -amylases, also the calcium-binding site. The catalytic sites predicted by InterPro were confirmed by local alignment using ClustalW tool by comparing the *L. longipalpis* sequences with sequences of crystallized proteins deposited in the UniProt databank, and the analysis was based on conserved regions (CSRs) (Janecek, 1997; Kashiwabara et al., 2000; Lovering et al., 2005). This comparison was also used to predict the chloride binding sites in α -amylases. For GH13,

the templates were the sequences of α -amylase from *Aspergillus oryzae* (Uniprot P0C1B3) (Toda et al., 1982), maltase from *Apis cerana japonica* (Uniprot A1IHL0) (Wongchawalit et al., 2006), amino acid transport protein (UniProt Q07837) (Bertran et al., 1993) from *Homo sapiens*, glucan branching enzyme from *E. coli* (UniProt P07762) (Baecker et al., 1986), and glycogen debranching enzyme from *Candida glabrata* (UniProt Q6FSK0) (Zhai et al., 2016). For GH31, the templates were the sequences of neutral α -glucosidase from *Homo sapiens* (UniProt O43451), lysosomal α -glucosidase from *Homo sapiens* (UniProt P10253) (Roig-Zamboni et al., 2017), and glycosidase NET37 from *Homo sapiens* (UniProt Q6NSJ0).

Phylogenetic Tree Construction of GHs Proteins

Using each of the GH13 and GH31 sequences retrieved from *L. longipalpis* genome as a query, we searched for orthologous proteins in six other insect species. 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 Base⁵. Sequences from *Drosophila melanogaster* and *Drosophila ananassae* were retrieved from Fly base⁶. For *Phlebotomus papatasi*, sequences were searched and annotated as described above at “Search for GH13 and GH31 sequences in *L. longipalpis* genome.” The orthologous sequences used in this analysis are described in **Supplementary Tables 1, 2**.

For phylogenetic analysis, sequences were aligned using Clustal Omega algorithm (version 1.2.2) (Sievers et al., 2011; Sievers and Higgins, 2018), after removing the signal peptide. The maximum likelihood phylogenetic tree was obtained using MEGAX software (Kumar et al., 2018) using partial deletion (90% site coverage cutoff). The bootstrap consensus tree was inferred from 500 replicates with cutoff of 50% for bootstrap values for the nodes.

Maintenance of Insects

The experiments were performed using *L. longipalpis* (from Jacobina, Bahia, Brazil), maintained at Lancaster University (United Kingdom). Insects were kept under standard laboratory conditions of temperature ($24 \pm 2^\circ\text{C}$), and a photoperiod of 8 h of light/16 h of darkness. Adult sandflies were fed with 70% (w/v) autoclaved sucrose solution in cotton wool *ad libitum*, unless stated differently in the experiments. For oviposition, females were fed with sheep blood containing Alsever's anticoagulant. For this, an artificial apparatus (Hemotek – Discovery Workshops, United Kingdom) at 37°C was used for 1 h 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 and reared to preserve the colony. For experiments, recently emerged females (0–3 h) were fed with 1.2 M sucrose for 2 days (SF females), with a piece of cotton wool soaked in the sugar solution, or were maintained for 3 days with a cotton wool soaked in water,

¹<http://www.vectorbase.org>

²<http://flybase.org/>

³<https://www.ncbi.nlm.nih.gov/>

⁴<https://prosite.expasy.org/mydomains/>

⁵<http://www.vectorbase.org>

⁶<http://flybase.org/>

before feeding on sheep blood. After blood-feeding (infected or not), they were kept 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 parasites in the amastigote form, they were incubated at 32°C with Grace's insect medium supplemented with 20% of fetal bovine serum, BME₁ vitamins, 2% human urine, and 25 µg/mL gentamycin sulfate. Parasites with a maximum of 26 passages in the amastigote maintaining medium, after isolation from the vertebrate host, were used for sandfly infections (Moraes et al., 2018). Infections were performed using a parasite concentration of 2×10^6 parasites/mL, estimated using Neubauer chamber. Briefly, after centrifugation for 5 min at $2000 \times g$, the supernatant was removed, and parasites were mixed on sheep blood and offered to 4-day old unfed females as described in section "Maintenance of Insects." Unfed females were discarded, and control insects were fed with uninfected blood. After blood-feeding, infected and control females were maintained as described above with water for 2 days.

RNA Extraction and cDNA Preparation

Newly hatched females (0–3 h) (NF), females fed with 1.2 M sucrose for 2 days (SF females), control females fed with uninfected blood (BF females) and infected females (BF_INF females), 2 days after blood feeding, were dissected in sterile phosphate-buffered saline. Pools with 10 guts or rest of the body (RB) were transferred to polypropylene vials containing 50 µL of TRI Reagent® (Cat. No. T9424, Sigma-Aldrich, Darmstadt, Germany) and kept at -80°C until further RNA extraction. For infected females, midgut was checked with light microscopy for parasite presence and only midguts containing parasites were used.

Total RNA was extracted following the supplier's protocol and quantified using NanoDrop® (NanoDrop Technologies, Wilmington, United States). 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

GH13 and GH31 coding sequences identified in the *L. longipalpis* genome were used to design specific oligonucleotides (Thornton and Basu, 2011), using the Primer3 program (Rozen and Skaletsky, 2000), Beacon Designer™ Free Edition⁷, and mFOLD Software⁸. The list of primers is described in **Supplementary Table 3**. Gene expression was initially analyzed by RT-PCR, and at least one gene representative of each cluster (A, B, or C) of the GH13 α -amylase group (see section Results, **Figure 3**) was analyzed. Genes LLOJ000566/LLOJ008156 (LLAglu1/LLAglu2), LLOJ002257 (LLAglu3), LLOJ004841

(LLAamyA4) and LLOJ003489 (LINAglu2), were additionally analyzed by RT-qPCR. Samples obtained from the gut and the rest of the body of recently emerged, sugar, and blood-fed females were used for RT-PCR or RT-qPCR analysis. Samples from females infected with *Leishmania mexicana* parasites were used only for the RT-qPCR analysis.

For RT-PCR, we carried duplicate reactions for each cDNA sample obtained from a tissue pool of 10 sand flies. Three biological replicates were carried for each group, with two cDNA samples obtained for each condition tested. Each amplification reaction was performed in a volume of 25 µL containing $2 \times$ Biomix Red (Bioline, London, United Kingdom), 100 ng cDNA, and 0.2 µM of each primer. The PCR parameters were: incubation at 95°C for 5 min, followed by 26 cycles of 94°C for 30 s, 55°C for 40 s, 72°C for 1 min, and a final incubation of 72°C for 5 min. Relative expression was normalized using a housekeeping gene (LLOJ001891, Glyceraldehyde 3-phosphate dehydrogenase, GAPDH) as control. RT-PCR products were analyzed by electrophoresis in a 1% (w/v) agarose gel containing gel red® (Biotium, Fremont, United States). Expression level was determined by densitometry of bands using the software ImageJ (Schneider et al., 2012).

Reverse transcriptions followed by quantitative polymerase chain reactions (RT-qPCR) were conducted using a CFX96 Touch™ Real-Time PCR Detection System (BioRad). We carried duplicate reactions for each cDNA sample obtained from a tissue pool of 10 sand flies. Three biological replicates were carried for each group, with two cDNA samples collected for each condition tested. The reaction was performed in 10 µL volume containing $2 \times$ of SYBR™ Green Master Mix (Thermo Fisher Scientific), 20 ng cDNA, and 0.5 µM of each primer. The parameters for PCR were: incubation at 95°C for 15 min, followed by 39 cycles of 94°C for 15 s, 56°C for 30 s, and 72°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, an analysis of the melting curve was performed. The expression of target genes in each tissue was quantified by the comparative Ct (Δ Ct) method (Schmittgen and Livak, 2008) normalized with the GAPDH housekeeping gene as an internal control.

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 \pm SEM. All statistical analysis were executed using the Software GraphPad Prism 6.0 (San Diego, CA, United States).

RESULTS

Sequence and Domain Analysis of the GH13 and GH31 Members in the Genome of *Lutzomyia longipalpis*

Proteins belonging to GH13 and GH31 are evolutionary conserved, being found in many classes of organisms. A total

⁷<http://www.premierbiosoft.com/qpcr/index.html>

⁸<http://www.idtdna.com/UNAFold?>

of 15 sequences belonging to the GH13 and four sequences from GH31 were found in the *L. longipalpis* genome, based on HMMER search in the Vector Base platform. For this, we used GH13 and GH31 know sequences of 5 dipteran species (*Aedes aegypti*, *Culex quinquefasciatus*, *Anopheles gambiae*, *Drosophila melanogaster*, *Drosophila ananassae*) as queries (**Supplementary Tables 1, 2**).

Sequences of proteins found were classified based on similarity to other protein sequences and 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, considering the presence of initial methionine, stop codon, correct exon/intron junctions, and similar size of the exons when compared to orthologous proteins. Truncated sequences were analyzed and annotated for missing parts using the Apollo annotation tool and alignment by ClustalW. The transcriptomic data present in the Vector Base was also used to predict the correct size of exons. Modifications for each of the sequences retrieved are described in **Supplementary Table 4**. 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. The nucleotide and protein sequences after annotation are presented in **Supplementary Files 1 and 2**, respectively.

Some nucleotide sequences retrieved were coding for more than one protein belonging to GH13 or GH31. These sequences were split up, so the number of genes and, consequently, proteins increased. From sequences initially retrieved from Vector Base, genes LLOJ004838, LLOJ004880, LLOJ004881, LLOJ005909 codified for four different alpha amylases. After annotation, it was possible to determine that these genes actually codify for ten different alpha-amylases. Gene LLOJ004838, for example, codified for two different alpha-amylases that seem to be transcribed at different moments (corroborated by transcriptomic data from Vector Base). Protein1 (LlAamyA1) was transcribed in both larvae and adults and Protein 2 (LlAamyA2) was transcribed only in immature forms. Annotation resulted in 21 complete sequences for GH13 and 6 for GH31. With these analyses, we identified 14 α -amylases, tree α -glucosidases, two amino acid transport proteins (heavy chain), one 1,4- α -glucan branching enzyme, and one glycogen debranching enzyme from GH13 (**Supplementary Table 5**). For GH31, we identified four glycosidases NET37 (nuclear envelope transmembrane glycosidase 37), one lysosomal α -glucosidase, and one neutral α -glucosidase (α -subunit) (**Supplementary Table 6**). Gene structures are demonstrated in **Supplementary Figure 1**, and protein structures are represented in **Figure 1**.

According to the genomic map available in the Vector Base, the α -amylases genes were organized in two clusters. LlAamyA1 to LlAamyA4 (4 α -amylases) composed the cluster A, and the cluster B was formed by LlAamyB1 to LlAamyB8 (8 α -amylases). Besides that, the α -amylases C1 and C2 were in tandem in the genome. All the putative α -amylases were 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 **Supplementary Table 5**). α -glucosidases were predicted as membrane proteins. LlAglu1 and LlAglu2 were putative transmembrane proteins, and for LlAglu3, both transmembrane domain and GPI anchor were predicted. These enzymes had molecular masses around 70 kDa and estimated isoelectric points from 4.7 to 5.6 (**Supplementary Table 5**). 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 (**Supplementary Table 6**).

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 (**Tables 1, 2**). The active site containing the catalytic residues, the calcium-binding site, and the chloride-binding site (**Supplementary Table 5 and Table 1**) were correctly identified for α -amylases sequences belonging to cluster A and also for LlAamyB7 and LlAamyC2. The α -amylase B1 (LlAamyB1/LLOJ004880_1) had not a typical catalytic triad and presented only two residues that might serve as the nucleophile and the acid/base catalyst (D204 and D241, respectively). Moreover, this enzyme also showed two substitutions: i) one in the calcium-binding site, a conserved histidine residue by asparagine, and ii) one in the chloride-binding site at the β 4 sheet, a conserved arginine by leucine (**Supplementary Table 5 and Table 1**). Some of the α -amylases might not function as active enzymes, since they lacked the correct nucleophile and the acid/base catalyst, as α -amylases B4 and B6 (LlAamyB4/LLOJ004881_1 and LlAamyB6/LLOJ004881_3).

The catalytic triad was D216, E290, D357 for both α -glucosidases 1 and 2 (LlAglu1, LLOJ000566 and LlAglu2, LLOJ0008156), and D236, E304, D375 for α -glucosidase 3 (LlAglu3, LLOJ0002257) with all highly conserved regions preserved (**Supplementary Table 5 and Table 1**).

The amino acid transport proteins (heavy chains) found in our screening shared a domain similarity with alpha-glucosidases, but the general protein primary structure was different (**Figure 1**), and they lacked the conserved catalytic residues (**Table 1**). By similarity, LLOJ008629 was described as “neutral and basic amino acid transport protein” (NBAThc), while LLOJ006803, due to the presence of the SLC3A2 domain, was classified as CD98hc (cluster of differentiation 98) (**Supplementary Table 5 and Table 1**).

The sequence of a 1,4- α -glucan branching enzyme (LlAGB1, LLOJ005533) had a central domain containing the active site, and the catalytic residues similar to the other enzymes from GH13. In the putative sequence of a glycogen debranching enzyme (LlGDE1, LLOJ008312), we identified both α -1,4 glucanotransferase (GT) and α -1,6 glucosidase (GC) domains, this enzyme was a soluble cytoplasm enzyme with an estimated size of 172.6 kDa. The GT domain at the N-terminal region of this protein, despite the low overall sequence identity, had similarity with five short conserved regions of the GH13; the catalytic triad was conserved with the D548 as the catalytic nucleophile, E577 as the proton donor, and the conserved residue D649.

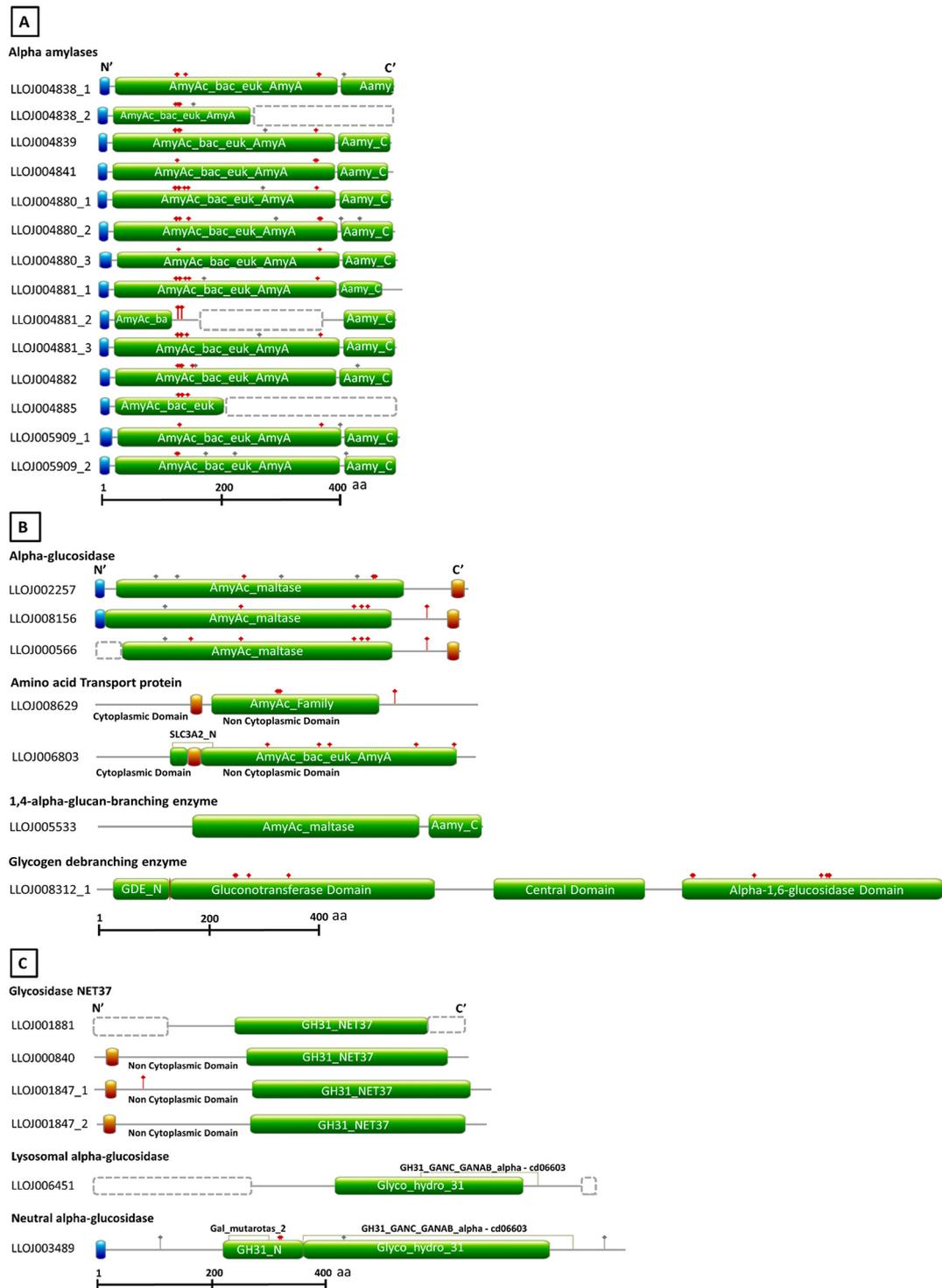


FIGURE 1 | Schematic diagram of the domain architecture of *L. longipalpis* proteins from GH13 and GH31. Sequences were retrieved from Vector Base (Jacobina strain, LlonJ1.4 geneset, June 2017). Signal peptide, transmembrane domains, and domains used to identify the different catalytic activities are boxed with the blue, red and green background, respectively. Gray and red dots above structures represent the N- and O- glycosylation sites. 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 α -glucosidases, 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.

TABLE 1 | Conserved amino acid sequences in proteins belonging to GH13.

Enzyme	$\beta 2$ (CSR VI)	$\beta 3$ (CSR I)	Loop3 (CSR V)	$\beta 4$ (CSR II)	$\beta 5$ (CSR III)	$\beta 7$ (CSR IV)	$\beta 8$ (CSR VII)
GH13							
Alpha-amylase							
<i>A. Orizae</i> (TAKA Amylase)	56-GFTAIWITP	117-DVANH	173-LPDL	202-GLIDTVKH	230-EVLD	323-FVENHD	323-GIPVYAGQE
LLOJ004838_1/LIAmyA1	53-GFAGVQVSP	111-DIVNH	172-LPDLN	200-GFVDAAKH	241-EVID	301-FVENHD	340-GIPVIMSSFA
LLOJ004838_2/LIAmyA2	53-GFAGVQVSA	113-DAIFNN	174-APDL	202-GFVIDSAKH	243-EVID	-----	-----
LLOJ004839/LIAmyA3	51-GFGGVQVSS	111-DAIFNH	171-APDL	199-GFVIDSAKH	240-EVID	303-FVENHD	338-GYPVLMSSDFN
LLOJ004841/LIAmyA4	52-GYAGVQLSP	110-DVLLNH	171-LHDLN	199-GFVDAAKH	240-EVID	299-FVDVHD	338-GTIVLMSSFA
LLOJ004880_1/LIAmyB1	52-GYAGVQVSS	112-FTIFNN	172-TPDL	200-GFVIDSAKH	241-DVIE	302-FVENEQ	340-GNPVIMSDFD
LLOJ004880_2/LIAmyB2	52-GFGGVQVSS	112-DTVNH	172-APDLN	200-GFVIDAAKH	244-EVID	304-FVENHD	343-GIPVIMSSFA
LLOJ004880_3/LIAmyB3	57-GFGGVQVSP	115-DVIFNH	176-LPDLN	204-GFVADACKH	245-EVID	304-FVENHD	343-GIPVIMSSYD
LLOJ004881_1/LIAmyB4	53-GFAGVQISS	113-EIIMNN	173-LPDLN	201-GFVINSACKH	242-DVAD	302-FVDVSPD	339-GHPVIMSSYE
LLOJ004881_2/LIAmyB5	52-GFGGVQVSS	113-DAIFNE	-----	-----	-----	-----	-----
LLOJ004881_3/LIAmyB6	52-GFGGVQVSS	113-DAVFNN	176-VPDL	204-GFVIFAACKH	245-DVID	306-FVENEQ	344-GNPVIMSSDYE
LLOJ004882/LIAmyB7	53-GFAGVQVSA	113-DAIFNN	174-APDL	202-GFVIDSAKH	243-EVID	300-FVENHD	370-GNLVIMSDFD
LLOJ004885/LIAmyB8	52-GFGGVQVSS	113-DAIFNE	176-VPDL	204-GFVILH---	-----	-----	-----
LLOJ005909_1/LIAmyC1	57-GFGGVQVSP	115-DVIFNH	176-LPDLN	204-GFVADACKH	-----	304-FVENHD	305-GIPVIMSSYD
LLOJ005909_2/LIAmyC2	54-GFGGVQVSP	113-DIVVNH	175-LRDLN	203-GFVIDAAKH	244-EVID	304-FVENHD	343-GVPVIMSSYE
Alpha-glucosidase							
<i>A. japonica</i> (JHGase I)	63-GITAIWLSP	117-DLVPNH	191-QPDLN	219-GFRIDAIPH	292-EAYT	349-VKGNHD	378-GVAVTYGEE
LLOJ000566/LIAglu1	58-GMSGTWLSP	112-DFVPNH	184-QPDLN	212-GFRIDAVPY	290-EAYT	352-VLGNHD	382-GMAVTYNGEE
LLOJ008156/LIAglu2	58-GMSGTWLSP	112-DFVPNH	184-QPDLN	212-GFRIDAVPY	290-EAYT	352-VLGNHD	382-GMAVTYNGEE
LLOJ002257/LIAglu3	78-GIFATWLSP	132-DFVPNH	204-QPDLN	232-GFRVDAINH	304-EAYT	370-VLGNHD	400-GVAVTYGEE
Amino acid transport protein							
<i>Homo sapiens</i> (rBAT)	156-NIKTWWITS	210-DFIPNH	282-QPDLN	310-GFSLDAVKF	384-EAYA	---	474-GTPITYGEE
LLOJ006803/LICD98hc	202-NIQIEGKIG	255-DLTPNF	318-RFDLQ	346-GIRLANTKH	---	---	501-GVPVFNVDAS
LLOJ008629/LINBATHc	201-SFQDSNGDG	254-DIDEIL	-----	338-GFYVKGLEN	383-EKVR	---	462-GTPSIFYGDE
1,4-alpha-glucan-branching enzyme							
<i>E. coli</i>	280-GFTHLELLP	335-DWVPGH	---	401-ALRVDAVAS	458-EEST	519-LPLSHD	562-GKKLLFMGNE
LLOJ005533/LIAGB1	221-GYNAIQVMA	276-DWVHSH	---	343-GYRFDGVTS	402-EDVS	466-YAESH	---
Glycogen debranching enzyme (α-1, 4 glucanotransferase)							
<i>Candida glabrata</i> (CgGDE)	105-GTYCFYLSF	230-DIVFNH	---	531-GFRIDNCHS	564-ELFS	622-MDITHD	---
LLOJ008312/LIGDE1	100-GTFHFYFSF	228-DIVLNH	---	544-GVRLDNCHS	577-ELFT	644-LDLTHD	---

Identification based on conserved regions described in Janecek (1997). The catalytic conserved residues, calcium binding and chloride binding residues are with gray, purple and green background, respectively. Non-canonical residues are marked in red.

TABLE 2 | Conserved amino acid sequences in proteins belonging to GH31.

Enzyme	(CR I)	(CR II)	β4 (CR III)	β6 (CR IV)	β7 (CR V)	α7 (CR VI)	α8 (CRS VII)
GH31							
Neutral alpha-glucosidase							
<i>Homo sapiens</i> (GANAB_HUMAN)	427-WLDIEHA	501-WCWPG	538-FWVNDIMNEP	616-TGDNTAE	647-DVGGF	667-GAYQFFRAH	704-RYSLLP
LLOJ003489/LINAglu1	407-WLDIEYT	481-WCWPG	518-MLWNDIMNEP	596-TGDNTAE	627-DVGGF	647-AAFQFFRAH	684-RYSYLP
Lysosomal alpha-glucosidase							
<i>Homo sapiens</i> (LYAG_HUMAN)	402-WNDLDYM	479-KVWPG	514-GMWIDIMNEP	614-TGDVWSS	645-DVCGF	665-GAFYPMRHH	702-RYALLP
LLOJ006451/LLysAglu1	325-L-PFDTH	381-RIGKT	479-GYMWQANWM	568-RRQVPST	537-NICGD	557-ASLSPFYR--	589-RYSLYL
Glycosidase NET37							
<i>Homo sapiens</i> (MYORG_HUMAN)	—?	—?	459-SFKFDAGEV	518-LVD RDSV	—?	—?	—?
LLOJ001847_1/LIGlyMyo1	—?	—?	437-GFYLD FGT S	484-LP VNSS	—?	—?	—?
LLOJ001847_2/LIGlyMyo2	—?	—?	427-GFYLD FGT S	485-LP VNSS	—?	—?	—?
LLOJ001881/LIGlyMyo3	—?	—?	397-SFKFDAGES	458-MIDKDSL	—?	—?	—?
LLOJ00840/LIGlyMyo4	—?	—?	404-NFYFEGGEF	465-LPDMIDNR	—?	—?	—?

Identification based on conserved regions described in Kashiwabara et al. (2000) and Lovering et al. (2005). The conserved catalytic residues are with gray background. Non-canonical residues are marked in red.

The C-terminal GC domain was similar to the catalytic domain of glucoamylases and other members from GH15 family. The general acid and general base for GC activity were identified as D1282 and E1515, respectively (**Supplementary Table 5 and Table 1**).

Regarding the members of GH31 family, the neutral α -glucosidase (α -subunit) (LINAglu1, LLOJ003489) was described with both Galactose mutarotase domain at the N-terminal end and the GANAB_GANAC domain (**Figure 1C**). 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 D522 and D598 (**Supplementary Table 6 and Table 2**). The lysosomal α -glucosidase (LLysAglu1, LLOJ006451) presented a substitution of both catalytic aspartate residues by glutamine (**Supplementary Table 6 and Table 2**). The glycosidases NET37 were predicted to have a single transmembrane domain at N-terminus and the GH31_NET37 domain at its C-terminus. The glucosidase domain was facing the non-cytoplasmic region (**Figure 1C**). Since these proteins are dependent on the catalytic residues to be active, glycosidase NET37 1 (LIGlyMyo1, LLOJ001847_1) and NET37 2 (LIGlyMyo2, LLOJ001847_2), seemed to lack hydrolytic activity. The proton donor residue was substituted for proline (**Supplementary Table 6 and Table 2**). Although the sequence for glycosidase NET37 3 (LIGlyMyo3, LLOJ001881) was not complete, it was possible to predict the two catalytic residues, 401D and 460D. Glycosidase NET37 4 (LIGlyMyo4, LLOJ00840) had glutamate as the conserved catalytic nucleophile, E408, and D467 as a proton donor.

Gene Numbers and Phylogenetic Analysis of GH13 and GH31 Sequences

A phylogenetic tree was constructed, aiming to support the classification and prediction of the functional role of the proteins. Using the *L. longipalpis* GH13 and GH31 protein sequences, we identified orthologs in other dipteran species and verified if gene amplification might have occurred in these GH families in the *Phlebotominae* branch. Phylogenetic analysis of GH13 protein sequences demonstrated the occurrence of 5 clades, each of them representing one of the activities described in this family (**Figure 2**). Within each clade, *Phlebotominae* proteins are more closely related to each other than to their orthologs in other dipteran species (**Figure 2**).

Clade I represented the α -amylases, and one of the subgroups of this clade had representatives only from the *Phlebotominae* subfamily. This subgroup was composed by 10 of 14 (71%) α -amylases of the *L. longipalpis* genome, suggesting a gene expansion in this protein family that included three sequences from cluster A (LLAamyA1, LLAamyA2, LLAamyA3) and seven sequences from cluster B (LLAamyB1, LLAamyB2, LLAamyB4, LLAamyB5, LLAamyB6, LLAamyB7, LLAamyB8) (**Figure 2**). The glucan branching enzymes were represented by a monophyletic group in clade II, demonstrating the sequence conservation among dipteran species. Enzymes from *L. longipalpis* and

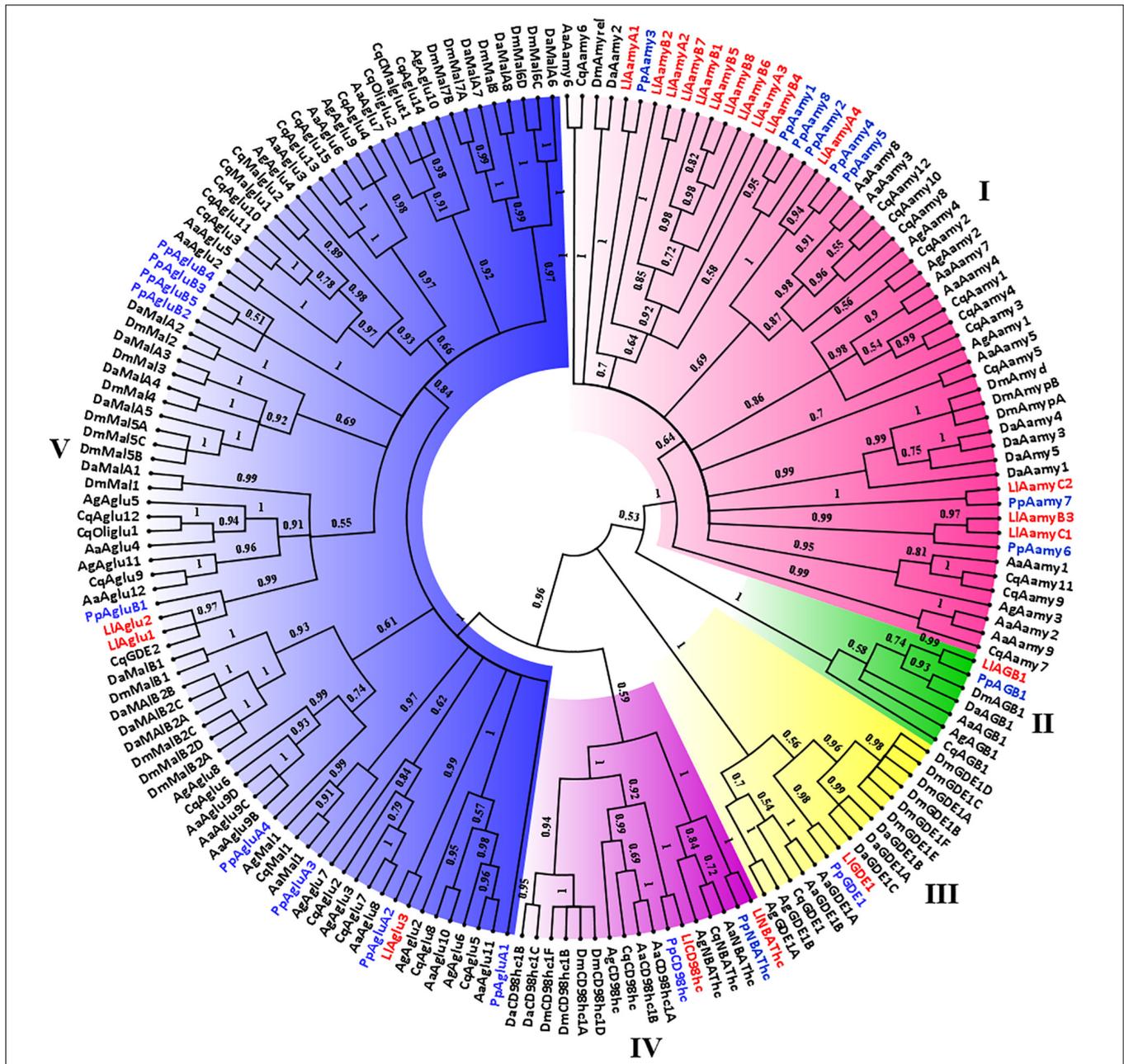


FIGURE 2 | Phylogenetic analysis of *L. longipalpis* GH13 proteins. Radial cladogram represents the alignment of GH13 protein sequences, α -amylases (Aamy, red), α -glucosidases (Aglu or Mal, blue), amino acid transport proteins (CD98hc and NBATHc, purple), 1,4- α -branching enzymes (AGB, green), glycogen debranching enzymes (GDE, yellow), of *L. longipalpis* and other dipterans species. *L. longipalpis* sequences are highlighted in red and *P. papatasi* in blue. *L. longipalpis* (L), *P. papatasi* (Pp), *A. aegypti* (Aa), *A. gambiae* (Ag), *C. quinquefasciatus* (Cq), *D. melanogaster* (Dm) and *D. ananassae* (Da). The tree represents the bootstrap consensus of 500 replicates, visualized with FigTree v1.4.3. Bootstrap values for the nodes are represented. The identifiers and corresponding genes are specified in Additional file 1.

P. papatasi are grouped with the *D. melanogaster* and *D. ananassae* (Figure 2). Clade III represented the debranching enzyme, and was formed by a paraphyletic group, since the Phlebotominae subfamily, Drosophilidae, and Culicidae families established different branches (Figure 2). Clade IV was composed by the amino acid transport proteins, and formed a paraphyletic group divided into two subgroups. One subgroup represents

the neutral and basic amino acid transporters (NBATs), and the other subgroup, the CD98 heavy chain, comprised the large neutral amino acid transport (LAT1). NBATs proteins missed the SLC3A2 domain, which is characteristic of CD98hc. The LICD98hc grouped with orthologs from the Drosophilidae family, already classified as different isoforms of the CD98 heavy chain (Figure 2).

Clade V represented the α -glucosidase proteins, and *L. longipalpis* α -glucosidases were grouped with *P. papatasi* orthologs. Two subgroups had representatives only from the Drosophilidae family (Figure 2). As described in Table 3, *P. papatasi* seemed to have an expansion in the α -glucosidase genes when compared to *L. longipalpis*. One of the subgroups of this clade had representatives only from *P. papatasi*. Interestingly, the PpAgluA3 is ortholog to the CqMal1 (Figure 2), an enzyme from *C. quinquefasciatus* characterized as GPI membrane-bound α -glucosidase with toxin binding properties (do Nascimento et al., 2017).

Phylogenetic analysis of GH31 protein sequences demonstrated the occurrence of 4 clades representing each activity described in this family (Figure 3). Once more, within each clade, *Phlebotominae* proteins related more closely to each other when compared to their orthologs in other dipteran species, except for clade IV, because it had no representative from *P. papatasi* (Figure 3). Clade I and II represented the lysosomal and the neutral alpha-glucosidases, respectively. As described in Table 2, neutral alpha-glucosidases maintained their catalytic residues. The LLLysAglu1 sequence had no conserved catalytic residue, and the phylogeny analysis grouped it with lysosomal enzymes from other dipteran species that also had not catalytic residues (Table 2 and Figure 3). The other two clades, III and IV, contained the glycosidases NET37. For NET37 glycosidases, clade III represented proteins containing the conserved catalytic domain, while clade IV represented glycosidases without the catalytic residues (Table 2 and Figure 3).

When we compare the number of genes for GH13 and GH31 in the *L. longipalpis* genome with their orthologs in 6 dipteran species, a potential expansion of α -amylases genes was observed in *L. longipalpis* (Table 3). 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. Interestingly, a contraction occurred for α -glucosidases genes in *L. longipalpis*, when compared to the other six species (Table 3). The number of genes coding for α -1,4-glucan branching and glycogen debranching enzymes (GDE) was conserved among the species analyzed (Table 3).

For GH31 family, the number of lysosomal and neutral α -glucosidase was conserved in phlebotomine and mosquitoes. *Drosophila* species had no lysosomal α -glucosidases. Glycosidase NET37 demonstrated a variable number of copies among the species analyzed and seemed to be duplicated in *L. longipalpis* when compared to *P. papatasi* (Table 3).

Expression of *L. longipalpis* GH13 and GH31 Genes in Different Physiological Conditions

After analyzing the sequences using bioinformatics tools, primers were designed for selected sequences in the GH13 family and all sequences of GH31. In the case of α -amylases, we analyzed at least one representative for each phylogenetic cluster of Figure 2. Also, it was not possible to design specific primer pairs to independently analyze expression of genes coding for LIAglu1/LIAglu2 or LIMyGly1/LIMyGly2, due to the high sequence identity between these sequences.

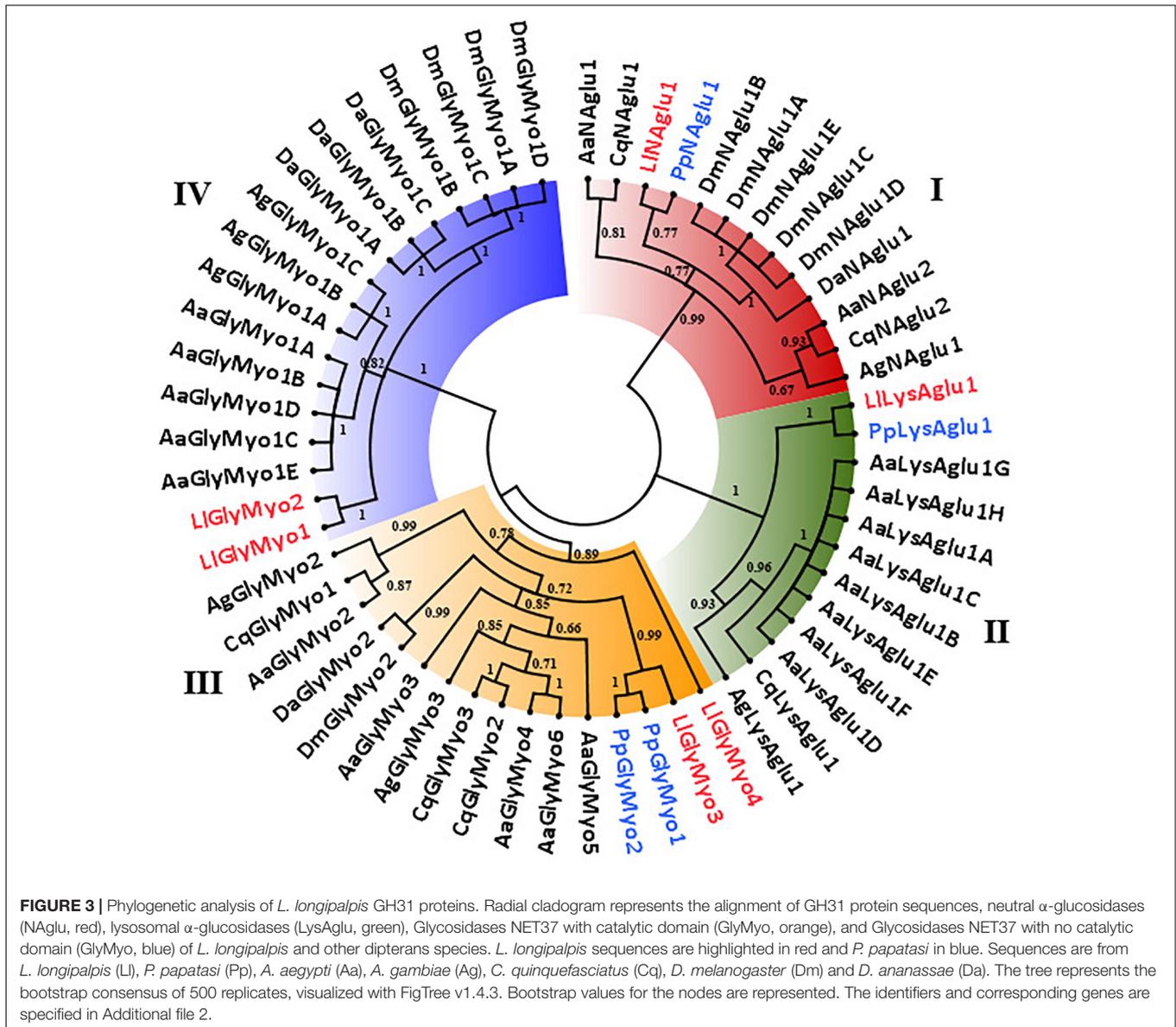
Primers amplified all sequences producing only one amplicon with the expected size from genomic DNA (data not shown). The expression of the *L. longipalpis* GH transcripts was evaluated by RT-PCR using different tissues (rest of body or gut) as sources of RNA, and at different dietary conditions, namely newly hatched females, sugar or blood-fed females.

All tested tissues and conditions poorly expressed both LIGlyMyo4 and the lysosomal α -glucosidase (data not show). The LIAmyA1 and LIAmyA4 were the α -amylases with the highest levels of expression in the gut when compared to other α -amylases genes independently of dietary condition tested (Figure 4A). No significant difference was found for LIAmyA1 expression levels when comparing different dietary conditions inside the same tissue. However, when comparing different tissues, LIAmy1 was significantly more expressed in the gut of sugar-fed females (one way ANOVA, $**p < 0.01$) and blood-fed females (One way ANOVA, $*p < 0.05$) than in rest of the body (Figure 4A). LIAmyA4 expression was higher in the gut for all dietary conditions tested (One way ANOVA, $****p < 0.0001$) and unaltered considering the dietary condition inside the same tissue (Figure 4A). Interestingly, blood-fed females did not express LIAmyB2.

TABLE 3 | Comparison of GH13 and GH31 gene numbers retrieved from Vector Base, Fly Base and NCBI in different dipteran species.

	<i>L. longipalpis</i>	<i>P. papatasi</i>	<i>A. aegypti</i>	<i>A. gambiae</i>	<i>C. quinquefasciatus</i>	<i>D. melanogaster</i>	<i>D. ananassae</i>
GH13							
Alpha-Amylase	14	8	9	4	12	3 (4)	5
Alpha-glucosidase	3	9	12 (14)	11	19	10 (16)	10 (12)
1,4-alpha-glucan-branching enzyme	1	1	1	1	1	1	1
Amino acid transport protein heavy chain	2	2	2 (3)	2	2	1 (4)	1 (2)
Glycogen debranching enzyme	1	1	1 (2)	1 (2)	2	1 (6)	1 (3)
GH31							
Lysosomal alpha-glucosidase	1	1	1 (8)	1	1	0	0
Neutral alpha-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 between brackets represents the total number of transcripts.

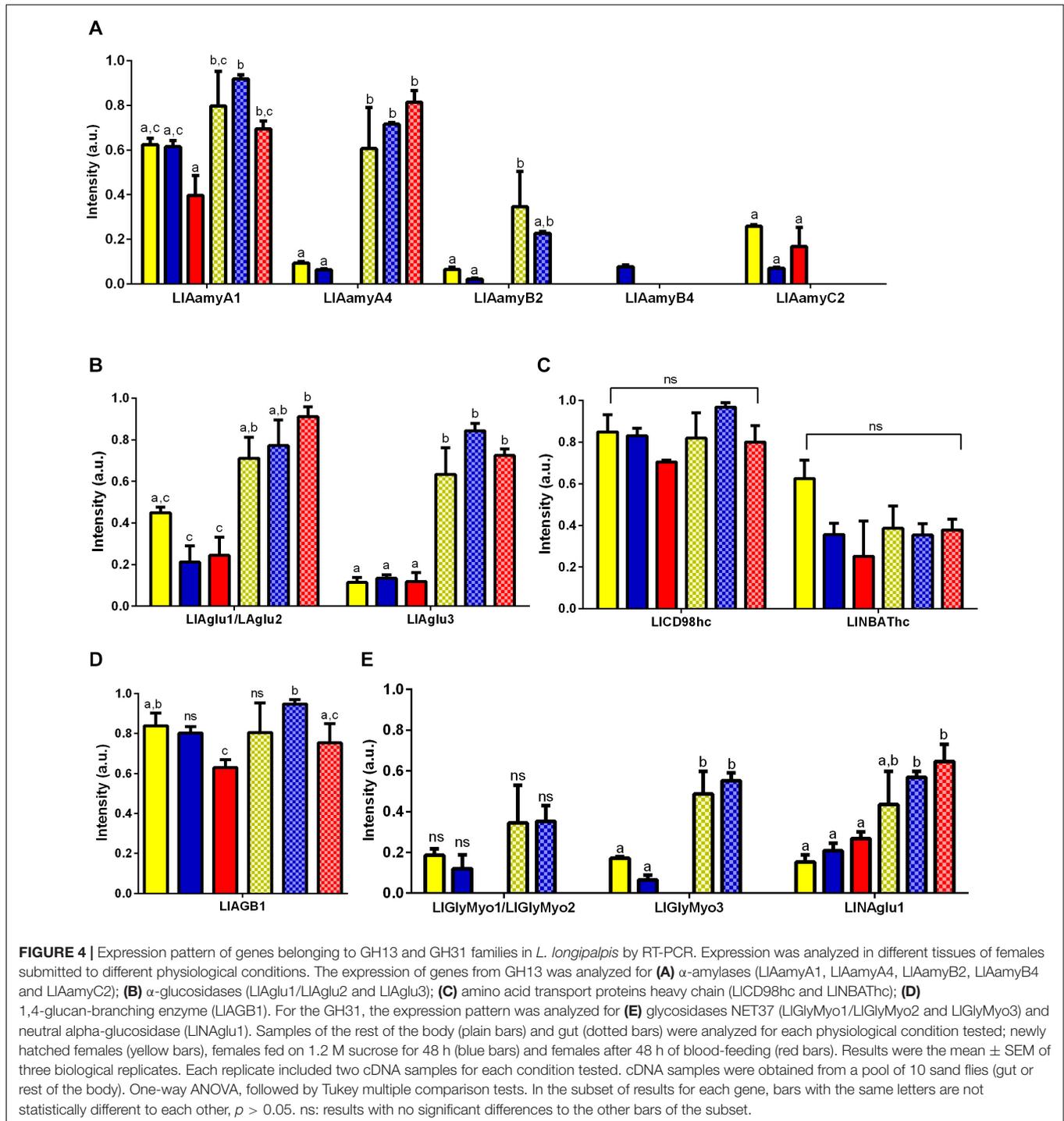


LIAmyB4 was poorly expressed in all tested conditions, while the expression of LIAmyC2 only occurred in the rest of the body (Figure 4A). α -glucosidase transcripts (LIAglu1/LIAglu2 and LIAglu3; Figure 4B) presented an expression pattern similar to LIAmyA1 and LIAmyA4, and no significant differences were found in expression levels when comparing different dietary conditions inside the same tissue. For LIAglu1/LIAglu2, the level of expression in the gut was higher for sugar-fed (One way ANOVA, $**p < 0.01$) and blood-fed females (One way ANOVA, $****p < 0.0001$), when compared to the rest of the body (Figure 4B). LIAglu3 expression was higher in the gut for all dietary conditions tested (One way ANOVA, $NF **p < 0.01$ and SF/BF $****p < 0.001$) (Figure 4B).

The expression of the α -1,4-glucan branching enzyme and the amino acid transporters is not linked to a specific tissue, and LINBATHc and LICD98hc presented no significant changes

in expression (Figure 4C) in different feeding conditions. The expression of LIAGB1 (Figure 4D) was reduced when females feed on blood.

For the GH31, blood-fed females did not express glycosidases NET37 in any tissue. LIglyMyo1/LIglyMyo2 were expressed in the gut and carcass of newly emerged and sugar fed insects, but with no significant differences among these samples (Figure 4E). LIglyMyo3 was also expressed in these same tissues and conditions, but it was more expressed in the gut when compared to the rest of body ($p < 0.05$, Figure 4E). There were no significant differences in the expression of LIglyMyo3 in the gut and rest of body when comparing newly emerged flies with sugar-fed ones (Figure 4E). LINAglu1 was significantly more expressed in the gut when compared to the rest of body, and its expression in each tissue was not affected by the feeding conditions (Figure 4E).

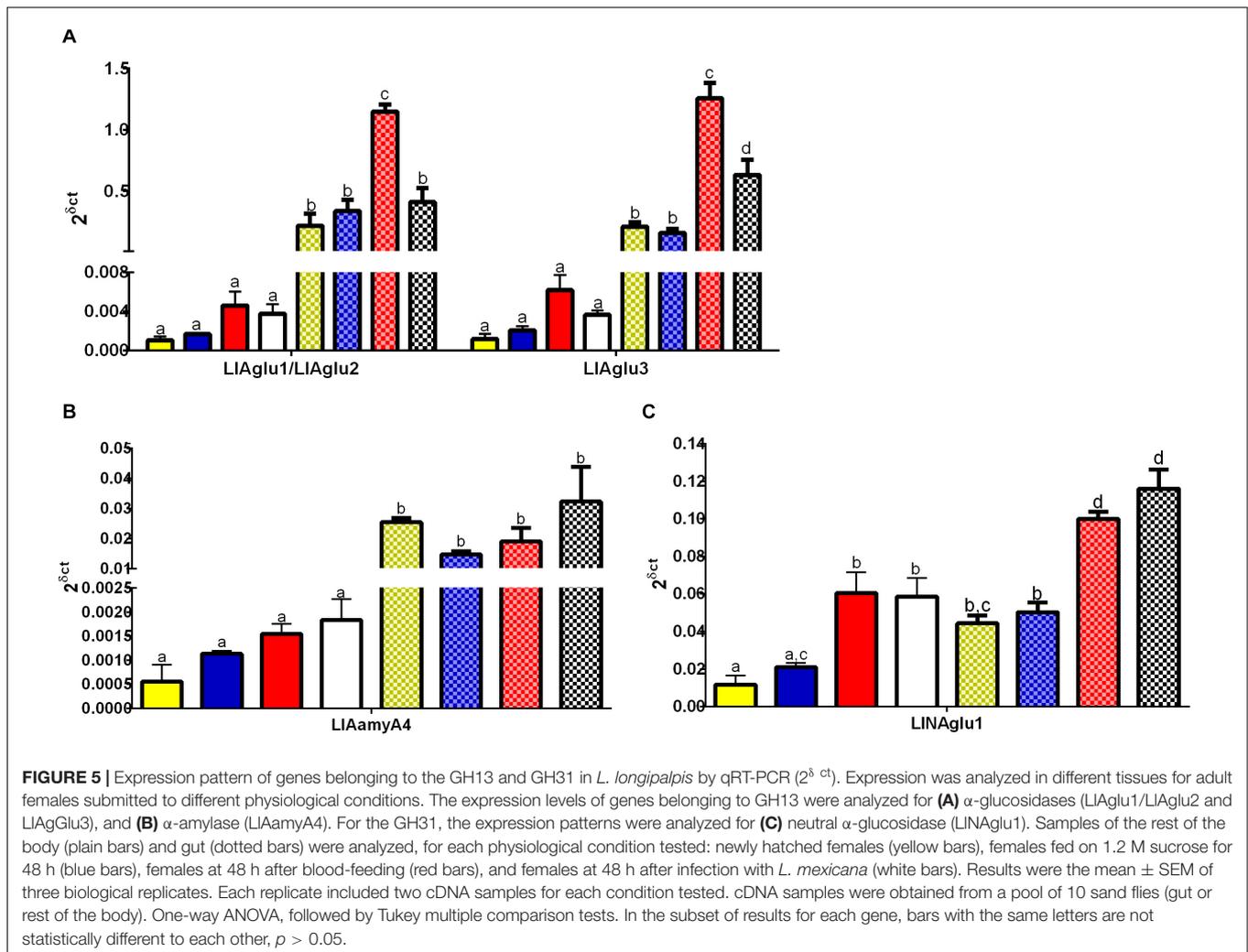


Gene Expression of LIAAmyA4, LIAglu1/LIAglu2, LIAglu3, and LINAglu2 in *L. longipalpis* Infected With *Leishmania mexicana*

Since we have an interest in sandfly proteins that might be affected by *Leishmania*, we selected the genes coding for LIAAmyA4, LIAglu1/LIAglu2, LIAglu3 (GH13) and

LINAglu1 (GH31) for further quantitative expression analysis after *L. mexicana* infections. These genes were chosen considering their preferential expression in the gut, the organ which was in direct contact with the parasite during the experimental infections.

RT-qPCR analysis of the genes LIAglu1/LIAglu2, LIAglu3, LIAAmyA4 (GH13), and LINAglu1 (GH31) confirmed higher gene expressions in gut tissues when compared to the rest of



the body, in all conditions tested (One way ANOVA, $*p < 0.05$) (Figure 5). LIAglu1/LIAglu2 and LIAglu3 genes showed higher quantitative expression values when compared to the genes LIAamyA4 and LINAglu1. Expression values for LIAglu1/2 and LIAglu3 were about 40 X higher than LIAamyA4 and 10 X higher than those observed for LINAglu1 (Figure 5).

Interestingly, the expression levels of the putative α -glucosidases LIAglu1/2 and LIAglu3 were significantly affected by the blood diet and *Leishmania* infection (Figure 5). We observed an induction in the expression levels of LIAglu1/LIAglu2 (1.14 ± 0.06) and LIAglu3 (1.2 ± 0.1) in the gut of blood-fed females in comparison to non-fed or sugar-fed females (One way ANOVA, $****p < 0.0001$; Figure 5A). Also, infection with *L. mexicana* parasite modulated negatively both the expression of LIAglu1/LIAglu2 (0.4 ± 0.1) and LIAglu3 (0.6 ± 0.1) (Figure 5A) (One way ANOVA, $****p < 0.0001$). The sugar diet did not significantly affect expression of these genes when compared to the non-fed condition.

LIAamyA4 had the lowest level of expression among the genes studied and presented the same pattern obtained previously using semi-quantitative RT-PCR, with no significant changes

after sugar or blood feeding, when comparing to newly emerged flies (Figure 5B). Besides that, LIAamyA4 expression was not affected by parasites in the blood diet (Figure 5B).

LINAglu1 is not highly expressed in the gut tissue compared to the rest of the body, but there is a significant difference between these tissues for all dietary conditions tested (Figure 5C). The expression level of LINAglu1 was also induced by the blood diet in both tissues in comparison to non-fed condition or sugar diet (One way ANOVA, $****p < 0.0001$; Figure 5C), but in this case the expression was not modulated by infection with *L. mexicana* parasites (Figure 5C).

DISCUSSION

Glycosidases play a significant role in the digestion of plant and blood sugars in sand flies, especially α -amylases and α -glucosidases. In this work, the majority of the 21 genes belonging to the GH13 family, in the genome of *L. longipalpis*, were classified as enzymes typically involved in carbohydrate digestion, especially 14 α -amylases, and three alpha-glucosidases.

Proteins involved in 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 endoplasmic reticulum (neutral α -glucosidase α subunit) and myogenesis regulation (NET37), were also described in GH13 or GH31 families.

We described potentially extracellular soluble α -amylases with predicted molecular masses around 55 kDa, and isoelectric points (pIs) between 4.5 and 6.5. We also described putative membrane-anchored alpha-glucosidases with molecular masses around 70 kDa and pIs between 4.5 and 5.5. In general, insect α -amylases have molecular weights about 48–60 kDa, pI 3.5–4.0, are calcium-dependent enzymes, and in some cases, they are also activated by chloride (Terra and Ferreira, 2012). The insect α -glucosidases have molecular weights ranging from 60 to 80 kDa (or multiple of these), with pI of 5.0–7.2 (Terra and Ferreira, 2012).

For insects in general, α -amylases are found as soluble enzymes in the midgut lumen, and α -glucosidases can also be seen as soluble enzymes, or anchored to the microvillar or perimicrovillar membranes of midgut cells (Silva and Terra, 1995; Nishimoto et al., 2001; Cristofolletti et al., 2003). The LLAGlu1, LLAGlu2, and LLAGlu3 described for *L. longipalpis* are probably membrane-anchored proteins, and their corresponding genes showed higher expression in the gut, when compared to the rest of the body in adult sandflies (Figures 4B, 5A). Their expression was higher when females fed with blood but did not change in sugar-fed females compared to non-fed newly emerged flies (Figure 5A). 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 (Gontijo et al., 1998; da Costa et al., 2019).

Our expression analysis suggests that the increase of α -glucosidase activity previously reported in sugar-fed females (da Costa et al., 2019) was not due to transcriptional gene regulation. No increase in expression was observed for any gene encoding α -glucosidase when females were sugar-fed compared to newly emerged flies, which contrasts with the observation that the gut sucrose activity increases 3 \times after sugar feeding for 48 h (da Costa et al., 2019). Besides that, sucrose induction is dose-dependent and responds preferentially to sucrose (da Costa et al., 2019). Another possibility is that transcription is regulated transiently, and after 48 h, it was not possible to determine any gene induction or repression. In adults and larvae of *D. melanogaster*, it has been described that some genes encoding carbohydrate digestive enzymes, as MAL-A1, are repressed in response to dietary sugars, due to a mechanism of regulation by negative feedback. The induction of MAL-A4 gene transiently occurs when maltose is offered to larvae and repressed when fed with glucose (Chng et al., 2017; Yamada et al., 2018).

A soluble form of α -glucosidase was described for *L. longipalpis* and *P. langeroni* when females were blood-fed (Dillon and El Kordy, 1997; da Costa et al., 2019). Biochemical assays using different tissues as enzyme source, in different conditions, and chromatographic analysis demonstrated the presence of at least four different α -glucosidases in *L. longipalpis*

(da Costa et al., 2019). In this work, we identified three genes coding for membrane-anchored α -glucosidases, that are positively regulated 48 h after blood feeding. These genes may undergo alternative splicing, producing different isoforms of α -glucosidases. Removal of the last exon might eliminate the transmembrane region and allow these enzymes to be secreted in soluble form, changing the biochemical characteristics of the enzymes. Additionally, the protein LLAGlu3 is putatively anchored to the membrane by a GPI anchor, and this might undergo cleavage by phospholipases releasing the protein from the midgut membrane. The soluble form of these α -glucosidases may pass through the peritrophic membrane and act synergistically with trypsin and other enzymes that release glycopeptides or glycolipids from blood glycoconjugates in the endoperitrophic space. Also, the membrane-bound or soluble α -glucosidases might play a role in the process of heme detoxification during blood digestion, during the hemozoin formation. This physiological function was described for the α -glucosidases of *R. prolixus* (Mury et al., 2009). In this respect, *L. longipalpis* α -glucosidases may play different roles in the hydrolysis of glycosides, during the digestion process, which may be associated with different enzyme specificities for sugars or glyco-derivatives, as glycolipids or glycoproteins.

In insects challenged with *Leishmania*, transcriptomic studies have demonstrated changes in the expression of several immune-related genes, oxidative stress-related molecules, and some digestive enzymes, mainly proteases (Dillon et al., 2006; Ramalho-Ortigão et al., 2007; Jochim et al., 2008; Abrudan et al., 2014). These works indicate that inside the vector, a systemic response is activated to defeat the *Leishmania* infection. However, a recent report showed that, after infection of *L. longipalpis* with *L. infantum*, only a small subset of genes (less than 1% of the transcriptome) showed substantial changes in expression (Coutinho-Abreu et al., 2020). Interestingly, at the same time window used in our experiments (48 h after infection), these authors observed downregulation of upstream factors related to nutrient transport, digestive enzymes, and peritrophic matrix proteins, which is coherent with our findings. In this context, *Leishmania* may likely modulate the physiological responses of phlebotomines, including changes in their microbiota, to overcome several biochemical and molecular challenges for the establishment of infection. The result may be the reduction of an enzymatic activity that may directly hydrolyze parasite surface molecules, as lipophosphoglycans (Ilg et al., 1992) or glycoinositolphospholipids (GIPLs; McConville et al., 1993), or may compete with *Leishmania* glycosidases or receptors for the sucrose present in the anterior parts of the insect's gut (Lyda et al., 2015).

Modulation of different sandfly enzymes by *Leishmania* parasites has been described, emphasizing proteases (Schlein and Romano, 1986; Dillon and Lane, 1993; Telleria et al., 2010; Santos et al., 2014). However, to our knowledge, this was the first time that the modulation of sandfly carbohydrases, caused by *Leishmania* infection, was specifically demonstrated. The modulation of α -glucosidase genes in *L. longipalpis* during *L. mexicana* infection may be related to a systemic response of the organism to the infection, which may cause the activation

of different components of the immune response pathways, consequently affecting the carbohydrate metabolism as a whole. Furthermore, GalNAc-containing glycoproteins present in the microvillar membranes of sandflies midgut are involved in the mechanism of attachment of *Leishmania* in permissive sandfly species (Myskova et al., 2007).

Alpha-glucosidases are enzymes abundantly found in the microvillar membrane of *L. longipalpis* midgut, and as described in this work, they have some potential N-glycosylation sites. In another hypothesis, these enzymes might be acting as a site for the attachment of *L. mexicana* in the midgut of *L. longipalpis*. As a response to the infection, the vector might downregulate the expression of the α -glucosidase genes. However, it remains elusive if these changes in the expression of *L. longipalpis* glucosidase genes are related to specific regulation mechanisms or just results of a general shift in the digestive process caused by the pathogen. Considering that we assessed just one-time point (48 h after feeding with an infected blood meal), changes in the whole temporal digestion pattern might explain our results. It is important to consider, in this context, that our previous reports in the same model showed an increase in trypsin activity but no change in the general course of protein degradation (Moraes et al., 2018). This may suggest differences in the regulation of protease and carbohydrase genes in the gut of *L. longipalpis*.

We described three α -glucosidase coding genes in *L. longipalpis*, suggesting a retraction compared to other dipterans, as *Drosophila* and even *P. papatasi*. This may be related to a strong specialization of these enzymes for the hydrolysis of sucrose, in an environment for *Lutzomyia* with higher offer of sucrose-rich nectar as the tropical forest, when compared to the savanna or desert associated to *P. papatasi*. Supporting this hypothesis, there is previous biochemical evidence that *L. longipalpis* α -glucosidases have strong specificity for sucrose as a substrate (da Costa et al., 2019).

Although the primary role of α -glucosidases is the metabolism of sugars, in the process of evolution, these proteins might have acquired additional functions. In some mosquitoes species, like *C. pipiens*, *C. quinquefasciatus*, and *A. gambiae*, the α -glucosidase is also described as the receptor of the Bin toxin from *Lysinibacillus sphaericus* (Darboux et al., 2001; Romão et al., 2006; Opota et al., 2008). The expression levels of maltase/ α -glucosidase genes for *D. melanogaster* are different when considering tissues and phases of development, suggesting a non-enzymatic regulatory function for these proteins (Gabriško, 2013). There are some examples where enzymes have lost their catalytic activity, developing new functions, but still have high sequence similarity with α -glucosidases. This phenomenon occurs for amino acid transporters (neutral and basic amino acid transport protein (NBAT) and the cluster of differentiation 98 (CD98) (Figure 2). We described two coding genes for these amino acid transport proteins (heavy chain) in *L. longipalpis*. The amino acid transporter is a heterodimeric protein, and the heavy chain has the function of correctly locating and folding the light chain on the plasma membrane. The light chain subunit has the purpose of transporting amino acids across the cell membrane (Gabriško and Janeček, 2009).

Interestingly, the α -amylase genes described for *L. longipalpis* are organized in two clusters (A and B), and the α -amylases C1 and C2 are in tandem in the genome in a different region. LIAmyC2 corresponds to the protein AF132512, previously described as a salivary α -amylase (Charlab et al., 1999). These clustered genes have the same putative enzymatic specificity, are closely spatially localized, and have high sequence similarity to each other, so probably these sequences originated from multiple subsequent genomic duplications. *Phlebotomus papatasi* and *C. quinquefasciatus* also present the clustering of some α -amylases. Our results suggest a gene expansion of *L. longipalpis* α -amylase family compared to *P. papatasi*, mosquitoes, and the genus *Drosophila* (Figure 2 and Table 3).

As a natural defense system, plants produce many aprotic and proteinaceous inhibitors such as acarbose, cyclodextrins, acarviosine-glucose, lectin-like, and knottin type to prevent the digestion by herbivorous insect's midgut alpha-amylases (Franco et al., 2002). Insects feeding on plant sap, as phlebotomines, might also be affected by these inhibitors. The expansion of the α -amylase gene family in *L. longipalpis* might reflect a diversification of these activities to overcome plant inhibitors or even a to account for tissue expression specificity. In the natural habitat of *L. longipalpis*, tropical forest, there is a greater diversity of available food sources when compared to the habitat of *P. papatasi*, savanna and desert (Akhoundi et al., 2016). 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 obtained in the dry-season desert (Jacobson et al., 2007). However, this hypothesis involves an adaptation that depends on the insect development stage. In this respect, it would be interesting to investigate if these α -amylases genes are differentially expressed throughout the entire sand fly life cycle, since amylase activity has been described both in adults and larvae (Ribeiro et al., 2000; Vale et al., 2012).

The seven CSRs and the catalytic residues were found in most *L. longipalpis* α -amylases and α -glucosidases. We also identified three conserved calcium-binding sites in the β 3 sheet (CSR I), loop 3 (CSR V), and β 4 sheet (CSR II) in the complete α -amylases sequences (Supplementary Table 5 and Table 1). In general, four residues function as the calcium-binding site, and three of them are conserved (Machius et al., 1995). Calcium preserves the structural integrity of the active site, maintaining proper folding (Buisson et al., 1987). Although some of the α -amylases identified did not demonstrate the typical catalytic residues (LIAmyB1, LIAmyB4, LIAmyB6), 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.

The chloride binding site is a characteristic of some α -amylases, and chloride works as an allosteric activator of the enzyme. In most *L. longipalpis* α -amylases, the chloride binding residues were identified. The dependence of *L. longipalpis* larvae α -amylases on chloride ions was demonstrated (Vale et al., 2012). However, LIAmyB1 and LIAmyB4 have a substitution in the asparagine residue (Table 1). Some α -amylases are described as

chloride-independent, as α -amylases from plants, fungi and most bacteria (Terra and Ferreira, 1994). In this respect, *L. longipalpis* amylases deserve a more detailed biochemical characterization to clarify their kinetical properties.

Regarding GH31, the phylogenetic analysis of GH31 sequences together with the expression analysis suggest that, at least for the neutral α -glucosidase, this enzyme may play its canonical role in the metabolism of *L. longipalpis*. The neutral α -glucosidases, described as glucosidase II in mammals, are soluble enzymes present in the endoplasmic reticulum. They participate in the quality control of glycoprotein folding, catalyzing the hydrolysis of glucose residues of oligosaccharides bound to peptides (Trombetta et al., 1996, 2001; Porath et al., 2016). These proteins are composed of one catalytic α -subunit and the non-catalytic β -subunit, comprising an ER retention signal (HDEL) at the C-terminal position. In *L. longipalpis*, the LlnAglu1 represents the catalytic α -subunit of the neutral α -glucosidase, being homologous to the α -glucosidase II in mammals. The correct catalytic residues were identified for LlnAglu1, with the nucleophile catalytic inside the domain WNDMNE and the general acid/base aspartate positioned on the β 6 sheet (Lee et al., 2001; Lovering et al., 2005; Ernst et al., 2006; Sim et al., 2008). In insects, neutral α -glucosidases have been identified by bioinformatics approaches, but no biochemical or structural studies have been carried until now.

The *L. longipalpis* lysosomal α -glucosidase showed a truncated sequence but was predicted based on the grouping in the phylogenetic tree with other enzymes from mosquito species. Proteins from mosquitoes present the TREFOIL and the NtCtMGAM domains that are typical of acid α -glucosidases (Roig-Zamboni et al., 2017). Since the classical catalytic residues were not identified in the *L. longipalpis* sequence, and the expression levels for this enzyme were low, it is plausible to assume that this represents a protein that has lost its enzymatic activity during evolution. The family Drosophilidae (*D. melanogaster* and *D. ananassae*) has no lysosomal α -glucosidase, while these enzymes in mosquitoes lacked the catalytic residues. This pattern might indicate the absence of glycogen degradation inside lysosomes in these insects, suggesting that the glycogen debranching enzyme mainly performs the mobilization of glycogen reserves in the cytosol, in association with other enzymes.

The NET37 proteins are not well described, especially in insects. They were recently characterized in mammals as membrane proteins anchored in the nuclear envelope with the catalytic domain facing the endoplasmic reticulum (Chen et al., 2006; Datta et al., 2009). These proteins have a transmembrane domain in the N-terminal region and a catalytic C-terminal region belonging to the GH31. *Lutzomyia longipalpis* NET37 proteins present the same structural organization (Figure 1C). NET proteins with different functions have been described. NET37 is required for myogenic differentiation of C2C12 cells (mouse myoblast cell line), with higher expression in the skeletal muscle of adult mouse. This activity is dependent on the catalytic site of the glycosidase motif (Datta et al., 2009). Due to the general lack of knowledge about these proteins, it is not possible yet to suppose

why their expression is affected by blood-feeding in adult female sandflies.

In summary, the genome of *L. longipalpis* presented several genes with sequences belonging to GH13 and GH31. The putative function of these genes seemed to be conserved, including enzymes involved in sugar metabolism, storage and mobilization of energetic glycogen reserves, and also proteins involved in amino acid transport, N-glycosylation quality control and myogenesis regulation. The canonical role of α -amylases and α -glucosidases is the digestion of plant carbohydrates, but in *L. longipalpis*, these enzymes seemed to be recruited to recognize a wider variety of dietary carbohydrates, including molecules that are present in blood. Our data suggested that genes LLOJ000566 (LlAglu1), LLOJ008156 (LlAglu2), and LLOJ002257 (LlAglu3) are the most probable candidates to be responsible for the previously described α -glucosidase activities produced by *L. longipalpis* adult females after sugar or blood meals. Besides that, it is noteworthy that different α -glucosidase activities, in this insect, are probably regulated by different mechanisms, transcriptional and post-transcriptional, after blood or sugar feeding. It is also important to realize that GH13 and GH31 genes showed a very diverse array of tissue and nutritional expression patterns, emphasizing the complexity of carbohydrate metabolism in insects.

Additionally, new functions may have arisen for these glycosidases, in coherence with an expansion in α -amylase gene family, induction of α -glucosidases after blood-feeding, and downregulation in *Leishmania mexicana* exposed females. To elucidate the roles of α -amylases and α -glucosidases, including the interaction of sandflies with the *Leishmania* parasite, a more detailed analysis of their biochemical features and mechanisms of expression must be pursued.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SC, PB, RD, and FG: conception and design of the work. SC: obtainment of experimental data. SC and FG: data analysis. SC, PB, RD, and FG: writing and revision of the manuscript. All authors have read and approved the final version.

FUNDING

This research was supported by the Brazilian Research Agencies FAPERJ, CNPq, CAPES, and Fiocruz.

ACKNOWLEDGMENTS

The authors thank Michelle Bates from Lancaster University for technical assistance, and Bruno Gomes,

and Maiara do Valle Faria Gama for the critical reading of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2021.635633/full#supplementary-material>

Supplementary Figure 1 | Schematic diagram of exon and intron composition of *L. longipalpis* genes belonging to GH13 and GH31. 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.

Supplementary Table 1 | Identifiers and accession numbers of insect GH13 protein sequences used for the phylogenetic tree of **Figure 2**.

Supplementary Table 2 | Identifiers and accession numbers of insect GH31 protein sequences used for the phylogenetic tree of **Figure 3**.

Supplementary Table 3 | List and sequences of the RT-qPCR primers used for PCR-amplification of cDNA sequences codifying for proteins belonging to GH13 and GH31 of *L. longipalpis*.

Supplementary Table 4 | Modifications on sequences from GH13 and GH31 retrieved from vector base.

Supplementary Table 5 | Manual annotation of *L. longipalpis* GH13 sequences retrieved from Vector Base (*L. longipalpis*, Jacobina strain, LlonJ1.4 geneset, June 2017). Sequences were considered complete when initial methionine, correct exon/intron junction, stop codon were identified, and exon structures were complete based on homology with orthologous genes. For incomplete proteins some N and O-glycosylation sites are missing. In transmembrane proteins, only extracellular domains were considered to be N-glycosylated. Non-canonical catalytic residues are highlighted. ND, Not Determined.

Supplementary Table 6 | Manual annotation of *L. longipalpis* GH31 sequences retrieved from Vector Base (*L. longipalpis*, Jacobina strain, LlonJ1.4 gene set, June 2017). Sequences were considered complete when initial methionine, correct exon/intron junction, stop codon were identified, and exon structures were complete based on homology with orthologous genes. For incomplete proteins some N and O-glycosylation sites are missing. In transmembrane proteins, only extracellular domains were considered as N-glycosylated. Non-canonical catalytic residues are highlighted. ND, Not Determined.

Supplementary File 1 | Exon and introns sequences of *L. longipalpis* genes belonging to GH13 and GH31 after sequence curation.

Supplementary File 2 | Protein sequences of *L. longipalpis* and *P. papatasi* belonging to GH13 and GH31 after sequence curation.

REFERENCES

- Abad, M. C., Binderup, K., Rios-Steiner, J., Arni, R. K., Preiss, J., and Geiger, J. H. (2002). The x-ray crystallographic structure of *Escherichia coli* branching enzyme. *J. Biol. Chem.* 277, 42164–42170. doi: 10.1074/jbc.M205746200
- Abrudan, J., Ramalho-ortigão, M., Neil, S. O., Stayback, G., Bernard, M., Shoue, D., et al. (2014). The characterization of the *Phlebotomus papatasi* transcriptome. *Insect Mol. Biol.* 22, 211–232. doi: 10.1111/imb.12015
- Akhoundi, M., Kuhls, K., Cannet, A., Votýpka, J., Marty, P., Delaunay, P., et al. (2016). A historical overview of the classification, evolution, and dispersion of *Leishmania* parasites and sandflies. *PLoS Negl. Trop. Dis.* 10:e0004349. doi: 10.1371/journal.pntd.0004349
- Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., and Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. doi: 10.1093/bioinformatics/btx431
- Arrese, E. L., and Soulages, J. L. (2010). Insect fat body: energy, metabolism, and regulation. *Annu. Rev. Entomol.* 55, 207–225. doi: 10.1146/annurev-ento-112408-085356
- Baecker, P. A., Greenberg, E., and Preiss, J. (1986). Biosynthesis of bacterial glycogen. Primary structure of *Escherichia coli* 1,4-alpha-D-glucan:1,4-alpha-D-glucan 6-alpha-D-(1, 4-alpha-D-glucano)-transferase as deduced from the nucleotide sequence of the GLG B gene. *J. Biol. Chem.* 261, 8738–8743. doi: 10.1016/S0021-9258(19)84443-5
- Bertran, J., Werner, A., Chillaron, J., Nunes, V., Biber, J., Testar, X., et al. (1993). Expression cloning of a human renal cDNA that induces high affinity transport of L-cystine shared with dibasic amino acids in *Xenopus oocytes*. *J. Biol. Chem.* 268, 14842–14849. doi: 10.1016/S0021-9258(18)82410-3
- Bjellqvist, B., Basse, B., Olsen, E., and Celis, J. E. (1994). Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. *Electrophoresis* 15, 529–539. doi: 10.1002/elps.1150150171
- Boel, E., Brady, L., Brzozowski, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., et al. (1990). Calcium binding in alpha-amylases: an X-ray diffraction study at 2.1-Å resolution of two enzymes from *Aspergillus*. *Biochemistry* 29, 6244–6249. doi: 10.1021/bi00478a019
- Buisson, G., Duée, E., Haser, R., and Payan, F. (1987). Three dimensional structure of porcine pancreatic alpha-amylase at 2.9 Å resolution. Role of calcium in structure and activity. *EMBO J.* 6, 3909–3916. doi: 10.1002/j.1460-2075.1987.tb02731.x
- Cameron, M. M., Pessoa, F. A. C., Vasconcelos, A. W., and Ward, R. D. (1995). Sugar meal sources for the phlebotomine sandfly *Lutzomyia longipalpis* in Ceará State, Brazil. *Med. Vet. Entomol.* 9, 263–272. doi: 10.1111/j.1365-2915.1995.tb00132.x
- Cantarel, B., Coutinho, P., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 37, D233–D238. doi: 10.1093/nar/gkn663
- Charlab, R., Valenzuela, J. G., Rowton, E. D., and Ribeiro, J. M. (1999). 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.* 96, 15155–15160. doi: 10.1073/pnas.96.26.15155
- Chen, I. H. B., Huber, M., Guan, T., Bubeck, A., and Gerace, L. (2006). Nuclear envelope transmembrane proteins (NETs) that are up-regulated during myogenesis. *BMC Cell Biol.* 7:38. doi: 10.1186/1471-2121-7-38
- Chng, W., Bin, A., Hietakangas, V., and Lemaitre, B. (2017). Physiological adaptations to sugar intake: new paradigms from *Drosophila melanogaster*. *Trends Endocrinol. Metab.* 28, 131–142. doi: 10.1016/j.tem.2016.11.003
- Coutinho-Abreu, I. V., Serafim, T. D., Meneses, C., Kamhawi, S., Oliveira, F., and Valenzuela, J. G. (2020). *Leishmania* infection induces a limited differential gene expression in the sand fly midgut. *BMC Genomics* 21:608. doi: 10.1186/s12864-020-07025-8
- Cristofolletti, P. T., Ribeiro, A. F., Deraison, C., Rahbé, Y., and Terra, W. R. (2003). Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrtosiphon pisum*. *J. Insect Physiol.* 49, 11–24. doi: 10.1016/S0022-1910(02)00222-6
- da Costa, S. G., Bates, P. A., Dillon, R. J., and Genta, F. A. (2019). Characterization of α -glucosidases from *Lutzomyia longipalpis* reveals independent hydrolysis systems for plant or blood sugars. *Front. Physiol.* 10:248. doi: 10.3389/fphys.2019.00248
- D'Amico, S., Gerday, C., and Feller, G. (2000). Structural similarities and evolutionary relationships in chloride-dependent alpha-amylases. *Gene* 253, 95–105. doi: 10.1016/S0378-1119(00)00229-8

- Darboux, I., Nielsen-LeRoux, C., Charles, J. F., and Pauron, D. (2001). The receptor of *Bacillus sphaericus* binary toxin in *Culex pipiens* (Diptera: Culicidae) midgut: molecular cloning and expression. *Insect Biochem. Mol. Biol.* 31, 981–990. doi: 10.1016/S0965-1748(01)00046-7
- Datta, K., Guan, T., and Gerace, L. (2009). NET37, a nuclear envelope transmembrane protein with glycosidase homology, is involved in myoblast differentiation. *J. Biol. Chem.* 284, 29666–29676. doi: 10.1074/jbc.M109.034041
- 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
- Dillon, R. J., Ivens, A. C., Churcher, C., Holroyd, N., Quail, M. A., Rogers, M. E., et al. (2006). Analysis of ESTs from *Lutzomyia longipalpis* sand flies and their contribution toward understanding the insect-parasite relationship. *Genomics* 88, 831–840. doi: 10.1016/j.ygeno.2006.06.011
- Dillon, R. J., and Lane, R. P. (1993). Influence of *Leishmania* infection on blood-meal digestion in the sandflies *Phlebotomus papatasi* and *P. langeroni*. *Parasitol. Res.* 79, 492–496. doi: 10.1007/BF00931590
- do Nascimento, N. A., Ferreira, L. M., Romão, T. P., da Correia, D. M. C., dos Vasconcelos, C. R. S., Rezende, A. M., et al. (2017). N-glycosylation influences the catalytic activity of mosquito α -glucosidases associated with susceptibility or refractoriness to *Lysinibacillus sphaericus*. *Insect Biochem. Mol. Biol.* 81, 62–71. doi: 10.1016/j.ibmb.2016.12.009
- Dunn, N. A., Unni, D. R., Diesh, C., Munoz-Torres, M., Harris, N. L., Yao, E., et al. (2019). Apollo: democratizing genome annotation. *PLoS Comput. Biol.* 15:1–14. doi: 10.1371/journal.pcbi.1006790
- Eisenhaber, B., Bork, P., and Eisenhaber, F. (1998). Sequence properties of GPI-anchored proteins near the omega-site: constraints for the polypeptide binding site of the putative transamidase. *Protein Eng.* 11, 1155–1161. doi: 10.1093/protein/11.12.1155
- Ernst, H. A., Lo Leggio, L., Willemoës, M., Leonard, G., Blum, P., and Larsen, S. (2006). Structure of the *Sulfolobus solfataricus* α -Glucosidase: implications for domain conservation and substrate recognition in GH31. *J. Mol. Biol.* 358, 1106–1124. doi: 10.1016/j.jmb.2006.02.056
- Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., et al. (2017). InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res.* 45, D190–D199. doi: 10.1093/nar/gkw1107
- Franco, O. L., Rigden, D. J., Melo, F. R., and Grossi-de-Sá, M. F. (2002). Plant α -amylase inhibitors and their interaction with insect α -amylases: structure, function and potential for crop protection. *Eur. J. Biochem.* 269, 397–412. doi: 10.1046/j.0014-2956.2001.02656.x
- Gabriško, M. (2013). Evolutionary history of eukaryotic α -glucosidases from the α -amylase family. *J. Mol. Evol.* 76, 129–145. doi: 10.1007/s00239-013-9545-4
- Gabriško, M., and Janeček, Š. (2009). 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.* 276, 7265–7278. doi: 10.1111/j.1742-4658.2009.07434.x
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., et al. (2005). “Protein Identification and analysis tools on the ExPASy Server,” in *The Proteomics Protocols Handbook*. Springer Protocols Handbooks, ed. J. M. Walker (Totowa, NJ: Humana Press), 571–607. doi: 10.1385/1-59259-890-0:571
- Gontijo, N. F., Almeida-Silva, S., Cio, F., Costa, F., Mares-Guia, M. L., Williams, P., et al. (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
- Gupta, R., Jung, E., and Brunak, S. (2002). Prediction of N-glycosylation sites in human proteins. *Pacific Symp. Biocomput.* 7, 310–322.
- Hoefslout, L. H., Hoogeveen-Westerveld, M., Kroos, M. A., van Beeumen, J., Reuser, A. J., and Oostra, B. A. (1988). Primary structure and processing of lysosomal alpha-glucosidase; homology with the intestinal sucrase-isomaltase complex. *EMBO J.* 7, 1697–1704. doi: 10.1002/j.1460-2075.1988.tb02998.x
- Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., Cuče, B. A., De castro, E., et al. (2008). The 20 years of PROSITE. *Nucleic Acids Res.* 36, D245–D249. doi: 10.1093/nar/gkm977
- Ilg, T., Etges, R., Overath, P., McConville, M. J., Thomas-Oates, J., Thomas, J., et al. (1992). Structure of *Leishmania mexicana* lipophosphoglycan. *J. Biol. Chem.* 267, 6834–6840. doi: 10.1016/s0021-9258(19)50502-6
- 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
- Janeček, S. (1997). Alpha-amylase family: molecular biology and evolution. *J. Prog. Biophys. Mol. Biol.* 67, 67–97. doi: 10.1016/s0079-6107(97)00015-1
- Janeček, Š, Svensson, B., and MacGregor, E. A. (2007). A remote but significant sequence homology between glycoside hydrolase clan GH-H and family GH31. *FEBS Lett.* 581, 1261–1268. doi: 10.1016/j.febslet.2007.02.036
- Janeček, Š, Svensson, B., and MacGregor, E. A. (2014). α -Amylase: an enzyme specificity found in various families of glycoside hydrolases. *Cell. Mol. Life Sci.* 71, 1149–1170. doi: 10.1007/s00018-013-1388-z
- Jochim, R. C., Teixeira, C. R., Laughinghouse, A., Mu, J., Oliveira, F., Gomes, R. B., et al. (2008). The midgut transcriptome of *Lutzomyia longipalpis*: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and *Leishmania infantum* chagasi-infected sand flies. *BMC Genomics* 9:15. doi: 10.1186/1471-2164-9-15
- Kamhawi, S. (2006). Phlebotomine sand flies and *Leishmania* parasites: friends or foes? *Trends Parasitol.* 22, 439–445. doi: 10.1016/j.pt.2006.06.012
- Kanai, R., Haga, K., Yamane, K., and Harata, K. (2001). Crystal structure of cyclodextrin glucanotransferase from alkalophilic *Bacillus* sp. 1011 Complexed with 1-deoxynojirimycin at 2.0 Å resolution. *J. Biochem.* 129, 593–598. doi: 10.1093/oxfordjournals.jbchem.a002895
- Kashiwabara, S. I., Azuma, S., Tsuduki, M., and Suzuki, Y. (2000). The primary structure of the subunit in *Bacillus thermoamyloliquefaciens* kp1071 molecular weight 540,000 homohexameric α -glucosidase ii belonging to the glycosyl hydrolase family 31. *Biosci. Biotechnol. Biochem.* 64, 1379–1393. doi: 10.1271/bbb.64.1379
- Killick-Kendrick, R. (1999). The biology and control of Phlebotomine sand flies. *Clin. Dermatol.* 17, 279–289. doi: 10.1016/S0738-081X(99)00046-2
- Killick-Kendrick, R., and Killick-Kendrick, M. (1987). Honeydew of aphids as a source of sugar for *Phlebotomus ariasi*. *Med. Vet. Entomol.* 1, 297–302. doi: 10.1111/j.1365-2915.1987.tb00358.x
- Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315
- Kumar, S., Stecher, G., Li, M., Nnyaz, C., and Tamura, K. (2018). MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35, 1547–1549. doi: 10.1093/molbev/msy096
- Kuriki, T., and Imanaka, T. (1999). The concept of the alpha-amylase family: structural similarity and common catalytic mechanism. *J. Biosci. Bioeng.* 87, 557–565. doi: 10.1016/S1389-1723(99)80114-5
- Lee, S. S., He, S., and Withers, S. G. (2001). Identification of the catalytic nucleophile of the Family 31 alpha-glucosidase from *Aspergillus niger* via trapping of a 5-fluoroglycosyl-enzyme intermediate. *Biochem. J.* 359, 381–386. doi: 10.1042/0264-6021:3590381
- Lovering, A. L., Seung, S. L., Kim, Y. W., Withers, S. G., and Strynadka, N. C. J. (2005). Mechanistic and structural analysis of a family 31 α -glycosidase and its glycosyl-enzyme intermediate. *J. Biol. Chem.* 280, 2105–2115. doi: 10.1074/jbc.M410468200
- Lyda, T. A., Joshi, M. B., Andersen, J. F., Kelada, A. Y., Owings, J. P., Bates, P. A., et al. (2015). A unique, highly conserved secretory invertase is differentially expressed by promastigote developmental forms of all species of the human pathogen. *Leishmania. Mol. Cell. Biochem.* 404, 53–77. doi: 10.1007/s11010-015-2366-6
- MacGregor, E. A., Janeček, Š, and Svensson, B. (2001). Relationship of sequence and structure to specificity in the α -amylase family of enzymes. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* 1546, 1–20.
- Machius, M., Wiegand, G., and Huber, R. (1995). Crystal structure of calcium-depleted *Bacillus licheniformis* α -amylase at 2.2 Å resolution. *J. Mol. Biol.* 246, 545–559. doi: 10.1006/jmbi.1994.0106
- Macvicter, J. A. K., Moore, J. S., Molyneux, D. H., and Maroli, M. (1990). Honeydew sugars in wild-caught Italian phlebotomine sandflies (Diptera: Psychodidae) as detected by high performance liquid

- chromatography. *Bull. Entomol. Res.* 80, 339–344. doi: 10.1017/S0007485300050537
- Matsuura, Y., Kusunoki, M., Harada, W., and Kakudo, M. (1984). Structure and possible catalytic residues of Taka-amylase A. *J. Biochem.* 95, 697–702. doi: 10.1093/oxfordjournals.jbchem.a134659
- McConville, M. J., Collidge, T. A. C., Ferguson, M. A. J., and Schneider, P. (1993). The glycoinositol phospholipids of *Leishmania mexicana* promastigotes. Evidence for the presence of three distinct pathways of glycolipid biosynthesis. *J. Biol. Chem.* 268, 15595–15604. doi: 10.1016/s0021-9258(18)82298-0
- Moraes, C. S., Aguiar-Martins, K., Costa, S. G., Bates, P. A., Dillon, R. J., and 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., Lucena, S. A., Moreira, B. H. S., 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
- Murray, H. W., Berman, J. D., Davies, C. R., and Saravia, N. G. (2005). Advances in leishmaniasis. *Lancet* 366, 1561–1577. doi: 10.1016/S0140-6736(05)67629-5
- Mury, F. B., da Silva, J. R., Ferreira, L. S., dos Santos, Ferreira, B., de Souza-Filho, G. 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
- Myskova, J., Svobodova, M., Beverley, S. M., and Volf, P. (2007). A lipophosphoglycan-independent development of *Leishmania* in permissive sand flies. *Microbes Infect.* 9, 317–324. doi: 10.1016/j.micinf.2006.12.010
- Nishimoto, M., Kubota, M., Tsuji, M., Mori, H., Kimura, A., Matsui, H., et al. (2001). Purification and substrate specificity of honeybee, *Apis mellifera* L., alpha-glucosidase III. *Biosci. Biotechnol. Biochem.* 65, 1610–1616. doi: 10.1271/bbb.65.1610
- Opota, O., Charles, J. F., Warot, S., Pauron, D., and Darboux, I. (2008). 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.* 149, 419–427. doi: 10.1016/j.cbpb.2007.11.002
- Pace, D. (2014). Leishmaniasis. *J. Infect.* 69, S10–S18. doi: 10.1016/j.jinf.2014.07.016
- Petersen, T. N., Brunak, S., Von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–486. doi: 10.1038/nmeth.1701
- Porath, B., Gainullin, V. G., Cornec-Le Gall, E., Dillinger, E. K., Heyer, C. M., Hopp, K., et al. (2016). Mutations in GANAB, encoding the glucosidase II α subunit, cause autosomal-dominant polycystic kidney and liver disease. *Am. J. Hum. Genet.* 98, 1193–1207. doi: 10.1016/j.ajhg.2016.05.004
- Ramalhó-Ortigão, M., Jochim, R. C., Anderson, J. M., Lawyer, P. G., Pham, V. M., Kamhawi, S., et al. (2007). Exploring the midgut transcriptome of *Phlebotomus papatasi*: comparative analysis of expression profiles of sugar-fed, blood-fed and *Leishmania major*-infected sandflies. *BMC Genomics* 8:300. doi: 10.1186/1471-2164-8-300
- Ramasubbu, N., Paloth, V., Luo, Y., Brayer, G. D., and Levine, M. J. (1996). Structure of human salivary α -Amylase at 1.6 Å resolution: implications for its role in the oral cavity. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 52, 435–446. doi: 10.1107/S0907444995014119
- Ribeiro, J. M. C., Rowton, E. D., and Charlab, R. (2000). Salivary amylase activity of the phlebotomine sand fly, *Lutzomyia longipalpis*. *Insect Biochem. Mol. Biol.* 30, 271–277. doi: 10.1016/S0965-1748(99)00119-8
- Roig-Zamboni, V., Cobucci-Ponzano, B., Iacono, R., Ferrara, M. C., Germany, S., Bourne, Y., et al. (2017). Structure of human lysosomal acid α -glucosidase-A guide for the treatment of Pompe disease. *Nat. Commun.* 8:1111. doi: 10.1038/s41467-017-01263-3
- Romão, T. P., De Melo Chalegre, K. D., Key, S., Junqueira Ayres, C. F., Fontes, De Oliveira, C. M., et al. (2006). A second independent resistance mechanism to *Bacillus sphaericus* binary toxin targets its α -glucosidase receptor in *Culex quinquefasciatus*. *FEBS J.* 273, 1556–1568. doi: 10.1111/j.1742-4658.2006.05177.x
- Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365–386. doi: 10.1385/1-59259-192-2:365
- Santos, V. C., Vale, V. F., Silva, S. M., Nascimento, A. A. S., Saab, N. A. A., Soares, R. P. P., et al. (2014). Host modulation by a parasite: how *Leishmania infantum* modifies the intestinal environment of *Lutzomyia longipalpis* to favor its development. *PLoS One* 9:e111241. doi: 10.1371/journal.pone.0111241
- Schlein, Y., and Romano, H. (1986). *Leishmania major* and *L. donovani*: effects on proteolytic enzymes of *Phlebotomus papatasi* (Diptera, Psychodidae). *Exp. Parasitol.* 62, 376–380. doi: 10.1016/0014-4894(86)90045-7
- Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089
- Sharma, U., and Singh, S. (2008). Insect vectors of *Leishmania*: distribution, physiology and their control. *J. Vector Borne Dis.* 45, 255–272.
- Sievers, F., and Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* 27, 135–145. doi: 10.1002/pro.3290
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75
- Silva, C. P., and Terra, W. R. (1995). An alpha-glucosidase from perimicrovillar membranes of *Dysdercus peruvianus* (Hemiptera: Pyrrhocoridae) midgut cells. Purification and properties. *Insect Biochem. Mol. Biol.* 25, 487–494. doi: 10.1016/0965-1748(94)00088-G
- Sim, L., Quezada-Calvillo, R., Sterchi, E. E., Nichols, B. L., and Rose, D. R. (2008). Human intestinal maltase-glucoamylase: crystal structure of the N-Terminal catalytic subunit and basis of inhibition and substrate specificity. *J. Mol. Biol.* 375, 782–792. doi: 10.1016/j.jmb.2007.10.069
- Soares, R. P. P., and Turco, S. J. (2003). *Lutzomyia longipalpis* (Diptera: Psychodidae: Phlebotominae): a review. *Ann. Brazilian Acad. Sci.* 75, 301–330. doi: 10.1590/S0001-37652003000300005
- Stentoft, C., Vakhrujev, S. Y., Joshi, H. J., Kong, Y., Vester-Christensen, M. B., Schjoldager, K. T. B. G., et al. (2013). Precision mapping of the human O-GalNAc glycoproteome through simplecell technology. *EMBO J.* 32, 1478–1488. doi: 10.1038/emboj.2013.79
- Svensson, B., and Macgregor, E. A. (2001). Relationship of sequence and structure to specificity in the α -amylase family of enzymes. *Biochim. Biophys. Acta* 1546, 1–20. doi: 10.1016/S0167-4838(00)00302-2
- Telleria, E. L., de Araújo, A. P. O., Secundino, N. F., D'Avila-Levy, C. M., and Traub-Csekö, Y. M. (2010). Trypsin-like serine proteases in *Lutzomyia longipalpis* - expression, activity and possible modulation by *Leishmania infantum chagasi*. *PLoS One* 5:e10697. doi: 10.1371/journal.pone.0010697
- Terra, W. R., and Ferreira, C. (1994). Insect digestive enzymes?: properties, compartmentalization and function. *Comp. Biochem. Physiol. Part B Comp. Biochem.* 109, 1–62. doi: 10.1016/0305-0491(94)90141-4
- Terra, W. R., and Ferreira, C. (2012). “Biochemistry and molecular biology of digestion,” in *Insect Molecular Biology and Biochemistry*, (Amsterdam: Elsevier B.V.), 365–418. doi: 10.1016/b978-0-12-384747-8.10011-x
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673
- Thompson, S. N. (2003). “Trehalose - The Insect ‘Blood’ Sugar,” in *Advances in Insect Physiology*, (Amsterdam: Elsevier Ltd), 205–285. doi: 10.1016/s0065-2806(03)31004-5
- Thornton, B., and Basu, C. (2011). Real-time PCR (qPCR) primer design using free online software. *Biochem. Mol. Biol. Educ.* 39, 145–154. doi: 10.1002/bmb.20461
- Toda, H., Kondo, K., and Narita, K. (1982). The complete amino acid sequence of taka-amylase A. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* 58, 208–212. doi: 10.2183/pjab.58.208
- Trombetta, E. S., Fleming, K. G., and Helenius, A. (2001). Quaternary and domain structure of glycoprotein processing glucosidase II. *Biochemistry* 40, 10717–10722. doi: 10.1021/bi010629u

- Trombetta, E. S., Simons, J. F., and Helenius, A. (1996). 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.* 271, 27509–27516. doi: 10.1074/jbc.271.44.27509
- Uitdehaag, J. C. M., Mosi, R., Kalk, K. H., Van der Veen, B. A., Dijkhuizen, L., Withers, S. G., et al. (1999). X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the α -amylase family. *Nat. Struct. Biol.* 6, 432–436. doi: 10.1038/8235
- 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
- Wongchawalit, J., Yamamoto, T., Nakai, H., Kim, Y.-M., Sato, N., Nishimoto, M., et al. (2006). Purification and characterization of alpha-glucosidase I from Japanese honeybee (*Apis cerana japonica*) and molecular cloning of its cDNA. *Biosci. Biotechnol. Biochem.* 70, 2889–2898. doi: 10.1271/bbb.60302
- Yamada, T., Habara, O., Kubo, H., and Nishimura, T. (2018). Fat body glycogen serves as a metabolic safeguard for the maintenance of sugar levels in *Drosophila*. *Development* 145:dev158865. doi: 10.1242/dev.158865
- Zhai, L., Feng, L., Xia, L., Yin, H., and Xiang, S. (2016). Crystal structure of glycogen debranching enzyme and insights into its catalysis and disease-causing mutations. *Nat. Commun.* 7, 1–11. doi: 10.1038/ncomms11229

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.

Copyright © 2021 da Costa-Latgé, Bates, Dillon and Genta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.