A metabolic insight into the Asian citrus psyllid: Annotation of glycolysis and gluconeogenesis pathways

Blessy Tamayo, Kyle Kercher, Tom D'Elia, Helen Wiersma-Koch, Surya Saha, Teresa Shippy, Susan J. Brown, and Prashant Hosmani

Introduction

Glycolysis, as in other animals, is the major metabolic pathway that insects use to extract energy from carbohydrates [1]. This process consists of ten reactions that convert one molecule of glucose into two molecules of pyruvate within the cytosol, generating a net gain of two molecules of ATP. Simple carbohydrates are acquired through the insect diet and processed through glycolysis and the central carbohydrate metabolic steps. Comparative genomic analysis of *Apis mellifera*, *Drosophila melanogaster*, and *Anopheles gambiae* has revealed the presence of genes for metabolic pathways that support dietary habits dependent on sugar-rich substrates ([2]; [3]; [4]; [5]). In addition to processing sugars for energy, glycolysis also serves as a central pathway that serves as both a source of important precursors and destination for key intermediates from many metabolic pathways.

Gluconeogenesis is the process through which glucose is synthesized from non-carbohydrate substrates, and is closely associated with glycolysis. Eleven enzymatic reactions occur during gluconeogenesis. Eight of the enzymes involved in the steps also catalyze the reverse reactions in glycolysis and the three remaining enzymes are specific to gluconeogenesis.

Gluconeogenesis generated carbohydrates are required as substrate for anaerobic glycolysis, synthesis of chitin, glycoproteins, polyols and glycoside detoxication products [1]. Gluconeogenesis is essential in insects to maintain sugar homeostasis and serves as the initial process towards the generation of glucose disaccharide trehalose, which is the main circulating sugar in the insect hemolymph ([6]; [7]). Thus, insects use trehaloneogenesis to synthesize trehalose from the last substrate of gluconeogenesis, glucose-6-phosphate. The process requires three additional enzymatic steps [7].

Diaphorina citri (Asian citrus psyllid) is the vector for the citrus greening disease bacterial pathogen *Candidatus* Liberibacter asiaticus. Citrus greening disease has been detrimental to the citrus industry, causing a continued loss of crops and decreasing economic growth. A first step towards development of treatments to prevent the spread of the disease requires a highly curated set of important genes in *D. citri* that could then be used to design molecular therapeutics. *D. citri* feeds on the phloem sap of citrus trees, a nutrient-rich food source containing high concentrations of sugars and providing a source of carbon and energy [8]. Annotation and characterization of the central core pathways of glycolysis and gluconeogenesis provides the genomic basis to understanding these processes in *D. citri* and their potential as molecular targets for treatments such as RNA interference. A community driven annotation strategy [9] was used to annotate the genome of *D. citri*. Through this process, the genes involved in glycolysis and gluconeogenesis were manually curated and further characterized.

Results and Discussion

Carbohydrate metabolism and annotation

The process of glycolysis oxidizing glucose to obtain ATP can be divided into two fundamental stages. The first stage is the energy investment stage, followed by the second stage of energy production. The ten enzymes of glycolysis, three enzymes specific to gluconeogenesis, and associated enzymes in both pathways were manually annotated in version 3.0 of the *D. citri* genome. The following discussion of the annotation and characterization of the carbohydrate metabolism pathways in *D. citri* is separated into three distinct groups: the energy investment stage of glycolysis, the energy production phase of glycolysis, and gluconeogenesis.

Energy investment stage of glycolysis

In this stage of glycolysis, a molecule of glucose is converted into two glyceraldehyde-3-phosphate molecules; this process consumes a total of 2 ATP molecules [10]. In addition, there are two key regulators of glycolysis: hexokinase and phosphofructokinase which are involved in irreversible reactions that consume ATP to synthesize phosphorylated intermediates. The following enzymes are involved in the energy investment stage of glycolysis and the genes were manually curated using Apollo annotation editor on the 3.0 version of the genome of *D. citri*: *hexokinase*, *phosphoglucose isomerase*, *phosphofructokinase*, *aldolase*, and *triosephosphate isomerase*.

Hexokinase

Hexokinase (HK) catalyzes the first step in glycolysis, utilizing ATP to phosphorylate glucose creating glucose-6-phosphate [10]. Insect muscle *HK* activity is inhibited by its product, glucose-6-phosphate, to initiate flight muscle activity [11]. While only one *HK* has been found in *D. citri*, multiple copies are present in other insects. *D. melanogaster* has four duplicated *HK* genes, with *Hex-A* being the most conserved and essential flight muscle *HK* isozyme among *Drosophila* species [12]. *HK* is present in *Tribolium castaneum* as two *Hexokinase A* genes, *Tc-HexA1* and *Tc-HexA2* [13]. Phylogenetic analysis of the single copy of *HK* in *D. citri* in Figure 1 shows that this enzyme is most similar to *Hex-A* orthologs of *T. castaneum, A. mellifera, Acyrthosiphon pisum, D. melanogaster,* and *Drosophila pseudoobscura.*

In *D. melanogaster*, northern blot analysis of *Hex-A* confirmed that it is strongly expressed in the thorax of male and female adult flies with no detectable transcripts in adult abdomens [12]. In addition, this gene was found to be expressed in adult flight muscle [14]. However, this differs in *D. citri* where the highest expression of *HK* was detected in abdomen (80.91 transcripts per million (TPM) male abdomen and 56.05 TPM female abdomen) (Figure 2). Although moderate expression of *HK* was observed in whole body (51.43 TPM). The two genes of *T. castaneum*, *Tc-HexA1* and *Tc-HexA2*, appear to be expressed during embryogenesis and in embryonic tissue [13]. The *HK* gene expression in *D. citri* was detected with lower expression in the egg with 11.28 TPM. Studies using parental RNAi in *T. castaneum* adult female have shown that *HK* is essential for embryonic development by affecting oogenesis, glucose content, and reducing egg lay which indicates its effectiveness as a potential RNAi target in *D. citri* [13].

The curated *HK* model has support from MCOT, de novo transcriptome, Iso-Seq, RNA-seq and orthologous sequence evidence tracks (Table 1). Based on NCBI BLASTp results, it shares reasonably high identity and query coverage with another Hemipteran, *Bemisia tabaci* (69.64%, 10%). Additionally, the annotated model contains the Hexokinase_1 and Hexokinase_2 domains (pfam00349 and pfam03727, respectively).

Phosphoglucose isomerase

Phosphoglucose isomerase (PGI) catalyzes the interconversion of glucose-6-phosphate and fructose-6-phosphate in the second step of glycolysis forming ATP [15]. Genetic variation in *PGI* has been correlated with flight metabolism, dispersal rate and metapopulation dynamics in the Glanville fritillary butterfly [16]; [17] In addition, *PGI* has been highly studied and identified as a key enzyme candidate in insect thermal tolerance to extreme temperatures [18].

A single copy of *PGI* was found in the *D. citri* genome version 3.0 and was supported by *de novo* transcriptome, RNA-seq, and orthologous evidence tracks (Table 1). A single copy of the gene was also found in other orthologs as shown in Table 2. Based on amino acid BLASTp results, this gene contains the PGI superfamily domain and shares high amino acid identity and query coverage with *B. tabaci* (78.10% and 97%, respectively). There is high expression for *PGI* in the antennae (172.17 TPM Adult female antennae and 166.47 TPM male antennae) compared to the gut (157.17 *C*Las positive and 130.37 TPM *C*Las negative) (Figure 2).

Phosphofructokinase

Phosphofructokinase (PFK), which catalyzes the phosphorylation of fructose-6-phosphate using ATP to generate fructose-1,6-bisphosphate and ADP, is the most prominent regulatory enzyme in the control of glycolysis [19]. This step is crucial as one of the major control points for carbohydrate metabolism in insects as it catalyzes one of the pathway's rate-determining reactions [11]. One copy of *PFK* was found and three isoforms were annotated in *D. citri*, compared to two copies found in *A. pisum* and only one copy found in *D. melanogaster*. However, no implications have been reported from differences in copy number for this rate-limiting enzyme in *A. mellifera*, *D. melanogaster*, and *Anopheles gambiae* [5].

PFK is moderately expressed in *D. citri*; there is low expression in the nymph whole body (14.37 TPM) compared to the highest expression in the adult whole body (75.63 TPM) (Figure 2). Moderate expression is detected in the adult gut (38.7 TPM *C*Las positive and 36.65 TPM *C*Las negative).

Aldolase

Fructose 1,6-bisphosphate aldolase (aldolase) catalyzes the reversible aldol cleavage of fructose-1,6-bisphosphate to form two trioses, glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP) [20]. A *UAS-Aldolase-RNAi* construct was used for RNAi knockdown of *aldolase* expression in *D. melanogaster* neurons and glia to observe normal lifespan and age-dependent neuroprotection [21]. The knockdown experiments suggest that the loss of *aldolase* in glia or neurons is sufficient to cause a reduction in lifespan and thus, *aldolase*

is essential in glia for neuronal maintenance as a function of age. As indicated by the results, *aldolase* may pose as a potential target for knockdown in *D. citri.*

The annotated *aldolase* gene model in *D. citri* had supporting evidence from MCOT, *de novo* transcriptome, Iso-Seq, RNA-seq, and orthologous sequence tracks (Table 1). A single copy of this gene was found in *D. citri*, which is comparable to other organisms (Table 2). From analyzing NCBI BLASTp results, the annotated model contains the glycolytic superfamily domain and shares high amino acid identity and query coverage with *T. castaneum* (74.18%, 100%) and *B. tabaci* (73.35%, 100%).

Aldolase transcripts analyzed in *D. melanogaster* were abundant at the earliest stages of development, most likely as a result of storage in the oocyte from transcription during oogenesis [20]. The transcripts, during the larval, pupal, and adult stages, rose, fell, and rose again in a pattern that is evident in intermediary metabolism. However, in *D. citri* the expression in the egg and nymph whole body were substantially low ranging from 1.6 to 5.78 TPM (Figure 2). Based on the expression in the adult stage of *D. citri*, there is relatively high expression in male abdomen (92.19 TPM) and whole body (118.36 TPM *C. macrophylla C*Las negative, 115.59 TPM Citrus spp. *C*Las negative, 40.04 TPM Citrus spp. *C*Las positive). However, there is little to no expression in female antennae (0 TPM), female abdomen (1.24 TPM), and in the gut (2.36 TPM, 5.62 TPM). Overall, the expression in the adult stage of *D. citri* is higher in *C*Las negative compared to *C*Las positive.

Triosephosphate isomerase

In the fifth step of glycolysis, triosephosphate isomerase (TPI) catalyzes the interconversion of DHAP and GAP, which are produced as a result of the aldol cleavage of fructose 1,6-diphosphate by fructose 1,6-bisphosphate aldolase in the previous step [22]. Though the function of this enzyme is essential in proceeding through the glycolytic steps, it is also essential in dipteran insects as TPI interacts with the *sn*-glycerol-3-phosphate dehydrogenase-glycerol phosphate oxidase shuttle to produce NAD^+ for flight muscle activity [23]. TPI in this cycle is used to sustain DHAP to maintain flight muscle activity, contrasting with the function of converting DHAP to GAP in the glycolytic cycle.

The annotated *TPI* model in the *D. citri* genome has support from MCOT, *de novo* transcriptome, RNA-seq, and orthologous sequence evidence tracks (Table 1). A single copy of the gene was found in *D. citri* and also in other orthologs (Table 2). NCBI BLASTp results show that this model shares high identity and query coverage with other Hemipterans, such as *N. lugens* (77.33%, 100%) and *B. tabaci* (81.78%, 100%).

In early embryos of *D. melanogaster*, TPI mRNA is abundant but declines in quantity in the first six hours, and then increases in abundance mid-embryogenesis at twelve hours showing a pattern due to the substantial amount of mRNA produced during oogenesis [22]. TPI mRNA content during the larval and adult stages follows a similar pattern of transcript content with several other glycolytic genes, such as aldolase and phosphoglycerate kinase (*PGK*). This varies in comparison to TPI expression in *D. citri* where it is lower in the egg (45.07 TPM) and in the nymph (31.26 TPM Citrus spp. *C*Las negative and 142.19 TPM *C. macrophylla C*Las negative) (Figure 2). However, there is also higher expression in the nymph (317.29 TPM Citrus spp. *C*Las positive

whole body), and also in the adult (280.84 TPM Citrus spp. *C*Las positive and 309.13 TPM negative). The overall result is relatively higher expression for TPI in the adult stage.

Energy production phase

In the energy production phase of glycolysis, two molecules of GAP are converted into two pyruvate molecules. Two molecules of triose are formed for each molecule of glucose used, a total of four molecules of ATP are synthesized from ADP for each molecule of hexose cleaved [10]. Thus, a net of two molecules of ATP are produced as a result of the overall glycolytic process. PYK, a key regulatory enzyme of glycolysis involved in an irreversible reaction, and PGK both produce ATP from phosphorylated intermediates. The genes for the following enzymes take part in the energy production phase of glycolysis and were annotated in the genome of *D. citri*: *glyceraldehyde phosphate dehydrogenase (GAPDH)*, *PGK*, *phosphoglycerate mutase (PGAM)*, *enolase*, and *PYK*.

Glyceraldehyde phosphate dehydrogenase

Glyceraldehyde phosphate dehydrogenase (GAPDH) catalyzes the conversion of GAP to 1,3-bisphosphoglycerate [24]. Two models of *GAPDH* were annotated in the *D. citri* genome. Two copies of *GAPDH* were found in other orthologous species, except for *A. pisum* (Table 2).

Two copies of *GAPDH*, *Gapdh-1* and *Gapdh-2,* were found in *D. melanogaster* and are located on chromosomes 43E-F and 13F, respectively [25]. In this study, both genes are expressed at higher levels in larval, late pupal and adult stages compared to embryonic, early, and midpupal stages. The *Gapdh-1* transcript is abundant in the thorax section compared to *Gapdh-2* which is

uniformly distributed in the head, thorax, and abdomen in *D. melanogaster*. In 18 hour embryos, the transcript level of *Gapdh-1* remained low and *Gapdh-2* started to increase. In the adult stage, the transcript level of *Gapdh-1* decreased as adults aged and the level of *Gapdh-2* remained constant. This is similar to expression data obtained from *D. citri* in comparison to the two *GAPDH* genes. In *D. citri*, the expression for Dcitr10g11030.1.1 ranges from no expression in antennae (0 TPM) to the highest expression seen in adult male abdomen (77.34 TPM) (Figure 2). The other annotated GAPDH model, Dcitr01g03160.1.1 is highly expressed in *D. citri* ranging from 201.51 TPM in female antennae and up 1106.88 TPM in adult whole body.

BLASTp results from flybase.org indicate that the two curated *GAPDH* genes align closely to the Gapdh-1 form of *D. melanogaster* as opposed to Gapdh-2. One of the annotated *GAPDH* models in *D. citri*, DcitrG027290.1.1, shares high amino acid identity and query coverage with other Hemipterans, such as *B. tabaci* (76.29%, 90%) and *N. lugens* (75.99%, 90%). The other annotated *GAPDH* model, DcitrG085445.2.1, shares high amino acid identity and query coverage with *B. tabaci* (90.36%, 100%) and *N. lugens* (88.86%, 100%). Both of the annotated models have supporting evidence from the following tracks, MCOT, Iso-Seq, RNA-Seq, and orthologous sequence.

Phosphoglycerate kinase

Phosphoglycerate kinase (PGK) catalyzes the conversion of 1,3-bisphosphoglycerate to 3-phosphoglycerate and generates one molecule of ATP [26]. Two protein domains, the C-terminal and N-terminal domains of PGK, may contain the ATP binding and phosphoglycerate binding regions, respectively ([27]; [28]). A single *PGK* gene has been identified in *D. melanogaster* [26].

A single model of the gene was annotated in *D. citri*, and the rest of the orthologs also have a single copy (Table 2).

The expression of PGK mRNA in *D. melanogaster* is similar to TPI; it is abundant in early embryos, declines in quantity through the first 6 hours, and in mid-embryogenesis, PGK mRNA begins to increase in abundance [26]. This pattern indicates a substantial amount of mRNA is produced during oogenesis, as it is stored in the oocyte and used to support translation during early development. PGK mRNA content during larval and adult stages follows a similar pattern for enzymes in related metabolic pathways, such as the glycolytic genes including *TPI* [22], *Aldolase* [20] and *GAPDH* [29].

The expression of *PGK* in *D. citri* is higher in the adult compared to the egg and nymph (Figure 2). Expression in egg and nymph are 158.09 and 88.74 TPM, respectively. However, in female and male antennae, there is high expression of the gene with 321.35 and 272.84 TPM, respectively. This does not correlate with the expression data seen in *D. melanogaster*.

The *PGK* gene model in *D. citri* shares high amino acid identity and query coverage with *B. tabaci* (83.33%, 99%), *N. lugens* (80.19%, 99%), *C. lectularius* (78.50%, 99%) and *S. flava* (78.99%, 99%). The curated model in version 3.0 of the genome does not have any support from other evidence tracks, such as de novo transcriptome, MCOT or Iso-Seq, but has support from RNAseq data.

Phosphoglycerate mutase

Phosphoglycerate mutase (PGAM) is an evolutionarily conserved enzyme that converts 3-phosphoglycerate to 2-phosphoglycerate and is a founding member of the PGAM protein family

[30]. Members of the PGAM family share a common domain, known as the PGAM domain, and function as phosphotransferases and/or phosphohydrolases. A single copy of *PGAM* was annotated in *D. citri*. Similarly, one copy was also found in *A. mellifera* and *A. pisum*, while two copies were found in *D. melanogaster* and *T. castaneum* (Table 2). The glycolytic enzyme phosphoglyceromutase (PGLYM) is found in *D. melanogaster* as two copies, *PGLYM87* and *PGLYM78* [31]. The BLASTp result obtained from flybase.org indicates that *PGAM* in *D. citri* aligns closely to the *PGLYM78* gene in *D. melanogaster*.

The expression of *PGAM* in *D. citri* is higher in the adult stage compared to the egg and nymph (Figure 2). The range in the adult stage has the lowest expression in male abdomen (83.44 TPM) and the highest in the gut (163.19 TPM). When the amino acid sequence of the annotated model was BLAST searched on NCBI, a gap was present in the HP_PGM_like domain. The model shared high identity and query coverage with *N. lugens* (81.50%, 100%), *A. pisum* (79.92%, 100%) and *B. tabaci* (82.68%, 100%).

Enolase

Enolase functions in the glycolytic pathway by catalyzing the conversion of 2-phosphoglycerate to phosphoenolpyruvate [32]. An *α-enolase* has been implicated as an RNAi target in *N. lugens*. Knockdown of the *α-enolase* reduced egg production, offspring, and hatching rate; however, mortality of adults was unaffected [32]. In addition, the *α-enolase* expression in *N. lugens* was higher in female adult than in male. The high level of expression was detected in hemolymph, fat body, gut, ovaries, salivary glands, wings, and eggs, with trace mRNA levels detected in testis. According to the study, a comparison of the α-enolase protein sequences with

other insect *enolase* genes showed a high degree of conservation. In a pairwise alignment between the *N. lugens* and *D. citri* sequences, both contain the hydrophobic domain in the N-terminal (AAVPSGASTGI) at position 31-41, seven amino acids for substrate binding pocket (H^{159}, E^{211}, K^{345}, $HRS^{373-375}$, and K^{396}), metal-binding site (S^{38}, D^{246}, E^{295}, and D^{320}) of the enolase family and the *enolase* signature (LLLKVNQIGSVTES).

The expression of *enolase* is relatively high in *D. citri*. Among the highest expression is seen in the adult stage whole body (*C. macrophylla C*Las negative 483.02 TPM and Citrus spp. *C*Las negative 728.52 TPM, respectively) (Figure 2). The lowest expression is identified in egg with 151.26 TPM. In the adult gut, *enolase* appears to be highly upregulated in *C. medica C*Las positive (413.9 TPM) compared to *C. medica C*Las negative (255.67 TPM). Additionally, expression is higher in the female adult than the male; analogous to enolase expression in *N. lugens.* This similar expression data is seen in female abdomen (251.48 TPM) compared to the male abdomen with (217.63 TPM) and also in female antennae (335.68 TPM) and male antennae (314.65 TPM).

The annotated *enolase* gene in *D. citri* shares high amino acid identity and query coverage with *B. tabaci* (87.53%, 100%) and *N. lugens* (85.25%, 100%). In addition, *enolase* gene copy numbers vary among insects (Table 2). A single *enolase* gene was identified in *D. citri* and in *N. lugens*, but three were found in *T. castaneum* and two in *A. pisum*. The annotated model has supporting evidence from the tracks Iso-Seq, RNA-seq, and orthologous sequence.

Pyruvate kinase

Pyruvate kinase (PYK) catalyzes the transfer of a phosphoryl group from phosphoenolpyruvate to ADP, generating pyruvate and the second ATP in the tenth reaction of

glycolysis [33]. *PYK* has a large difference in copy number between species, such as *A. mellifera* with two copy numbers, *D. melanogaster* with six, and *A. gambiae* with one [5]. There are at least six genes that encode *PYK* in *A. mellifera*, and their activity increases after emergence thus operating in honeybee foraging ([34], [35]). There may be potentially more in *D. melanogaster*, suggesting the enzyme has diversified as it encounters a range of carbohydrates [36]. Furthermore, the activities of PYK on energy metabolism are inferred to be essential in insect embryonic muscle development in *D. melanogaster* and early embryogenesis in the beetle, *Rhynchophorus palmarum* [37], [35].

RNAi treatments silencing an *N. lugens PYK* (*NlPYK*) was conducted by combining triazophos (TZP), an organophosphate insecticide, and dsNlPYK which led to reduced mRNAs encoding PYK*,* ovarian protein content, ovarian and fat body soluble sugar contents, and fecundity [35]. Thus, it is hypothesized that PYK acts directly in egg development and insect fecundity which may pose PYK as a potential RNAi target in *D. citri*. However, TZP treatments led to the upregulation of *NlPYK* transcripts, resulting in increased carbon flow through the citric acid cycle and increased energy. Therefore, exposure to different insecticides can lead to increased NlPYK activity and increased energy generation [35].

The two annotated *D. citri* gene models of *PYK* were located on chromosome 1 and 7, and NCBI BLASTp results indicate that both of the models contain the PYK superfamily domain (cd00288). The model on chromosome 7, Dcitr07g06140.1.1, shares high amino acid identity and query coverage with *S. flava* (70.45%, 93%) and *A. pisum* (69.70%, 96%). Dcitr07g06140.1.1 has supporting evidence from all of the evidence tracks in Table 1. The model on chromosome 1, Dcitr01g11190.1.1, shares reasonably high amino acid identity and query coverage with *C.*

lectularius (35.03%, 99%) and *H. halys* (35.37%, 97%). Dcitr01g11190.1.1 has support from all of the evidence tracks, except for orthologous sequence (Table 1).

Enzymes of gluconeogenesis

The enzymes of gluconeogenesis are characterized by eight enzymes, that also catalyze the reverse reactions in glycolysis, and three enzymes that catalyze the bypass reactions [7]. The eight genes found in both pathways were annotated in the *D. citri* genome: *PGI*, *Aldolase*, *GAPDH*, *TPI*, *PGK*, *PGAM* and *enolase*. The three enzymes specific to gluconeogenesis are phosphoenolpyruvate carboxykinase (PEPCK), fructose-1, 6-bisphosphatase (FBPase), and glucose-6-phosphatase (G6Pase). However, G6Pase was the only enzyme not annotated since it was not found within the *D. citri* genome.

Pyruvate carboxylase

Pyruvate carboxylase (PC) catalyzes the ATP-dependent carboxylation of pyruvate to oxaloacetate [38]; [39]. This enzyme was found in homogenates of bumblebee flight muscle, where the activity was inhibited by ADP and activated by acetyl-coA, and was the first demonstration of PC in muscle [38]. However, the enzyme was restricted to insect flight muscle and was not detected in leg muscles. Enzyme activity of PC is found in the flight muscle of *A. mellifera* and another Hemipteran, *Lethocerus cordofanus* [38]. However, the activity was below the limit of detectability of assays in the foreleg of *L. cordofanus,* supporting the activity of PC in flight muscle. In addition, PC may have another role in insect flight muscle for supplying

oxaloacetate to support the large increase in activity of the tricarboxylic acid cycle when an insect takes flight [38].

The curated *PC* model in *D. citri* has support from all of the evidence tracks in Table 1. BLASTp results show high identity and query coverage in alignment of the *PC* model with Hemipterans such as, *S. flava* (82.96%, 100%) and *A. pisum* (82.79%, 100%). *PC* is moderately expressed in adult female antennae (262.79 TPM), followed by expression in adult male antennae (214.27 TPM) and adult whole body (208.94 TPM) (Figure 2). The gene has lower expression in the gut, abdomen, and whole body. Furthermore, one copy of *PC* was found in the *D. citri* genome as seen in other insects (Table 2).

Phosphoenolpyruvate carboxykinase

Phosphoenolpyruvate carboxykinase (PEPCK) controls the cataplerotic flux and converts oxaloacetate from the TCA cycle to form PEP [40]. PEPCK is essential in maintaining glucose homeostasis and plays a key role in stress and starvation response [41]. *PEPCK* is present in *D. melanogaster* as two genes, *pepck1* (CG17725) and *pepck2* (CG10924), which is homologous to the two isoforms in mammals, PEPCK-C and PEPCK-M ([42]; [40]).

Two gene models of *PEPCK* were annotated and characterized within the 3.0 version of the *D. citri* genome. Based on an NCBI BLASTp alignment, the first annotated model, Dcitr08g02760.1.1, shares reasonably high identity and query coverage with another Hemipteran, *B. tabaci* (65.17%, 94%). The other annotated model, Dcitr05g10240.1.1, also shares reasonably high identity and query coverage with *B. tabaci* (70.10%, 97%). Both of the models contain the PEPCK-HprK superfamily domain. A pairwise alignment between the two models

shows that the identity is 69.05% and query coverage is 96%. The low identity indicates that the two models have no significant similarity and are not duplicates of each other.

Fructose 1,6-bisphosphatase

Fructose 1,6-bisphosphatase, (FBPase), facilitates one of the three bypass reactions occurring in gluconeogenesis [38]. This enzyme in bumble-bee flight muscle differs from the enzyme from other tissues that it is not inhibited by AMP [43]. It was demonstrated that this enzyme in flight muscles of bumble-bees have unusually high maximum catalytic activity, and has been discussed in relation to heat production [43]. However, FBPase activity is not present in some insect flight muscles, including Lepidoptera and Hymenoptera [38].

The annotated *FBPase* shares high identity and query coverage with *B. tabaci* (66.86% and 98%) and *S. flava* (66.57% and 97%). *FBPase* expression in *D. citri* is relatively low (*C. medica* *C*Las positive, 1.82 TPM; and negative adult gut, 3.25 TPM) (Figure 2). However, an increase in expression is observed in *C. macrophylla C*Las negative egg with 93.9 TPM. A single copy of the gene was annotated in the version 3.0 *D. citri* genome comparable to one copy found in *D. melanogaster and T. castaneum* (Table 2). However, two copies are present in *A. mellifera* and *A. pisum*.

Glucose-6-phosphatase

Glucose-6-phosphatase (G6Pase) which is specific to gluconeogenesis, catalyzes the last step of the pathway where glucose-6-phosphate is converted to glucose [7]. However, this enzyme was not present within the *D. citri* genome version 3.0 and was the only enzyme not

annotated in the gluconeogenesis pathway. Glucose-6-phosphatase appears to be completely missing in the assembly of the *A. mellifera* genome [5]. BLAST searches using human and *Drosophila* glucose-6-phosphatase protein sequences showed that this gene was confined to higher animals and many, but not all, arthropods and mollusks [7]. In insects, trehalose is the main sugar constituent of the circulatory system. Glucose 6-phosphate is converted by trehalose-6-phosphate synthase and trehalose-6-phosphatase into trehalose. Trehalose then moves to cells through the hemolymph, where it is converted into glucose by trehalase.

Evidence Table

Gene	**Identifier**	**MCOT**	***de novo* transcripts**	**Iso-Seq**	**RNA-Seq**	**Ortholog**
Glycolysis						
Hexokinase	Dcitr03g19430.1.1	x	x	x	x	x
Phosphoglucose isomerase	Dcitr00g06460.1.1		x		x	x
Glucose-6-phosphate 1-epimerase	Dcitr13g02890.1.1	x	x	x	x	
ATP Dependent 6-Phosphofructokinase isoform 2	Dcitr01g16570.1.1 Dcitr01g16570.1.2 Dcitr01g16570.1.3	x	x	x	x	x
Fructose-bisphosphate aldolase	Dcitr04g02510.1.1	x	x	x	x	x
Triosephosphate isomerase	Dcitr10g08030.1.1	x	x		x	x

Glyceraldehyde 3-phosphate dehydrogenase	Dcitr10g11030.1.1 Dcitr01g03160.1.1	x		x	x	x
Phosphoglycerate kinase	Dcitr00g01740.1.1					x
Phosphoglycerate mutase	Dcitr03g17850.1.1			x	x	x
Enolase	Dcitr02g07600.1.1			x	x	x
Pyruvate kinase-like 1	Dcitr07g06140.1.1	x	x	x	x	x
Pyruvate kinase-like 2	Dcitr01g11190.1.1	x	x	x	x	
Phosphoglucomutase	Dcitr02g10730.1.1			x	x	x
Phosphoglucomutase	Dcitr05g09820.1.1	x		x	x	x
Gluconeogenesis						
Pyruvate carboxylase	Dcitr08g01610.1.1	x	x	x	x	x
Phosphoenolpyruvate carboxykinase-like 1	Dcitr08g02760.1.1	x		x	x	x
Phosphoenolpyruvate carboxykinase-like 2	Dcitr05g10240.1.1	x		x	x	
Aldose 1-epimerase	Dcitr04g09830.1.1	x		x	x	x
Aldose 1-epimerase	Dcitr02g14030.1.1				x	
Glucose-6-phosphate 1-epimerase	Dcitr13g02890.1.1	x	x	x		
Fructose-1,6-bisphosphatase	Dcitr11g08070.1.1		x		x	x

Table 1: List of annotated *D. citri* models along with their evidence tracks. Each manually annotated gene in glycolysis and gluconeogenesis associated with a *D. citri* identifier shows supporting evidence used in the curation of the gene model. Evidence tracks are as follows, RNA-Seq, long-read Iso-Seq, MCOT, *de novo* assembled transcripts and orthologous proteins. A gene marked with an "x" within the table indicate that the gene model is supported by the evidence track.

Copy Number Table

	A. pisum	***T. castaneum***	***A. mellifera***	***D. melanogaster***	***D. citri***
Hexokinase (HK)	3	2	1	4	1
Phosphoglucose isomerase (PGI)	1	1	1	1	1
Glucose-6-phosphate 1-epimerase	1	1	1	1	1
6-Phosphofructokinase (PFK)	2	1	1	1	1*
Fructose bisphosphate-aldolase (ALDA or ALDOA)	1	1	1	1	1
Triosephosphate isomerase (TPI)	1	1	1	1	1
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	1	2	2	2	2

Phosphoglycerate mutase (PGAM)	1	2	1	2	1
Phosphoglycerate kinase (PGK)	1	1	1	1	1
Enolase	1	3	2	1	1
Pyruvate kinase (PYK)	1	4	6†	6†	2
Pyruvate carboxylase (PC)	1	1	1	1	1
Phosphoenolpyruvate carboxykinase (PEPCK)	1	1	1	1	2
Phosphoglucomutase (PGM)	2	2	2	1	2
Aldose 1-epimerase (GALM)	3	2	3	1	2
Fructose 1,6-bisphosphatase (FBPase)	2	1	2	1	1
Glucose-6-phosphatase (G6P)	0	0	0	1	0

Table 2: The number of genes are identified in glycolysis and gluconeogenesis in *D. citri* and related organisms. † indicates that there are possibly more PYK genes in *D. melanogaster* and potentially six in *A. mellifera* [35]. * indicates that three PFK isoforms were annotated in *D. citri*.

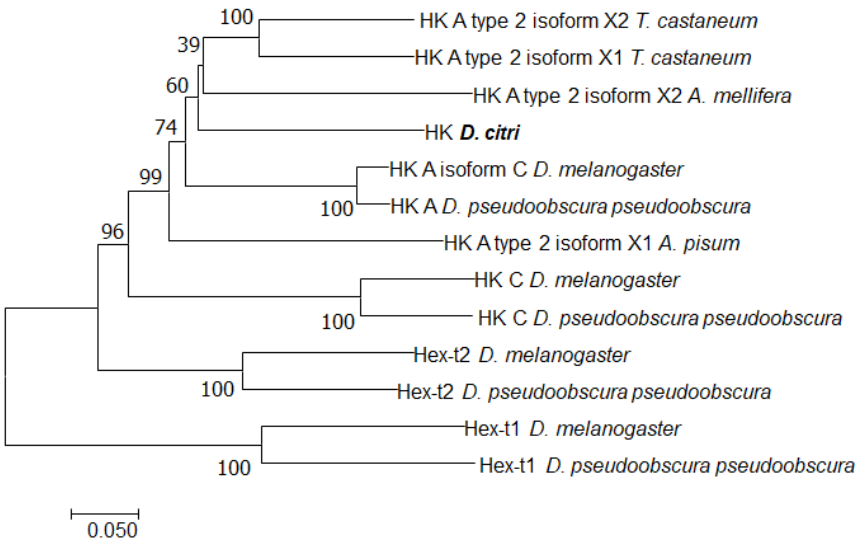


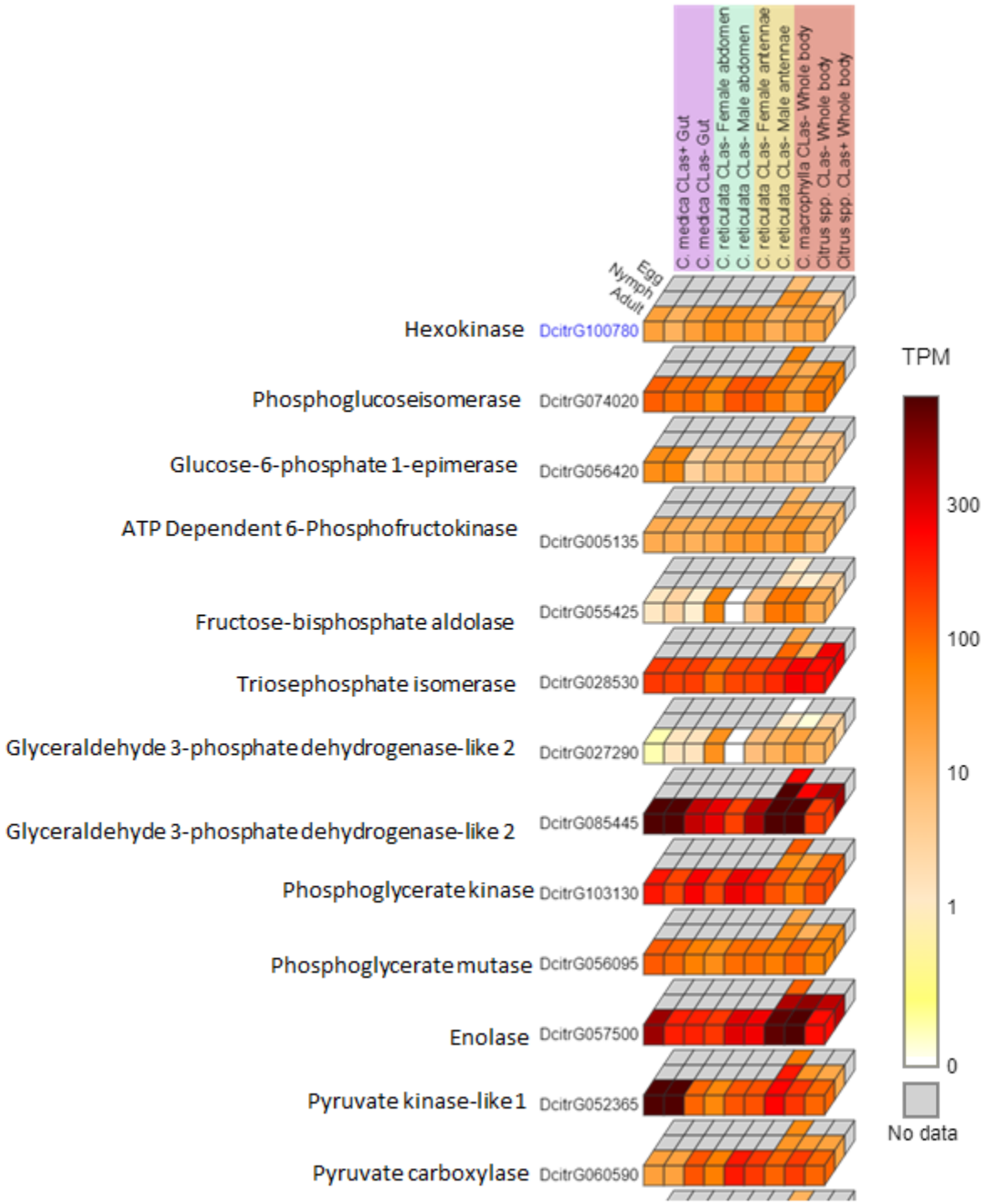
Figure 1: Phylogenetic analysis of Hexokinase. *Hexokinase* of *D. citri* groups closely with the HK A gene of the *Drosophila* species, *T. castaneum*, *A. mellifera* and *A. pisum*. MUSCLE multiple sequence alignments of *HK* in *D. citri* and orthologs were performed on MEGA7 with p-distance for determining evolutionary placement and 1000 bootstrapping replicates to measure the precision of branch placement [44].

Accession Number Table

Gene	*Drosophila melanogaster*	*Tribolium castaneum*	*Apis mellifera*	*Acyrthosiphon pisum*	*Drosophila pseudoobscura*
Hex-A	NP_001259384.1	XP_008201714.1 XP_970645.1	XP_006557646.1	XP_003242238.1	XP_001355083.1
Hex-C	NP_524674.1				XP_001360104.1
Hex-t1	NP_788744.1				XP_001359146.2
Hex-t2	NP_733151.2				XP_002137641.1

Table 3: The accession numbers for the Hexokinase genes in each ortholog were obtained from the NCBI database.

Expression Data Analysis



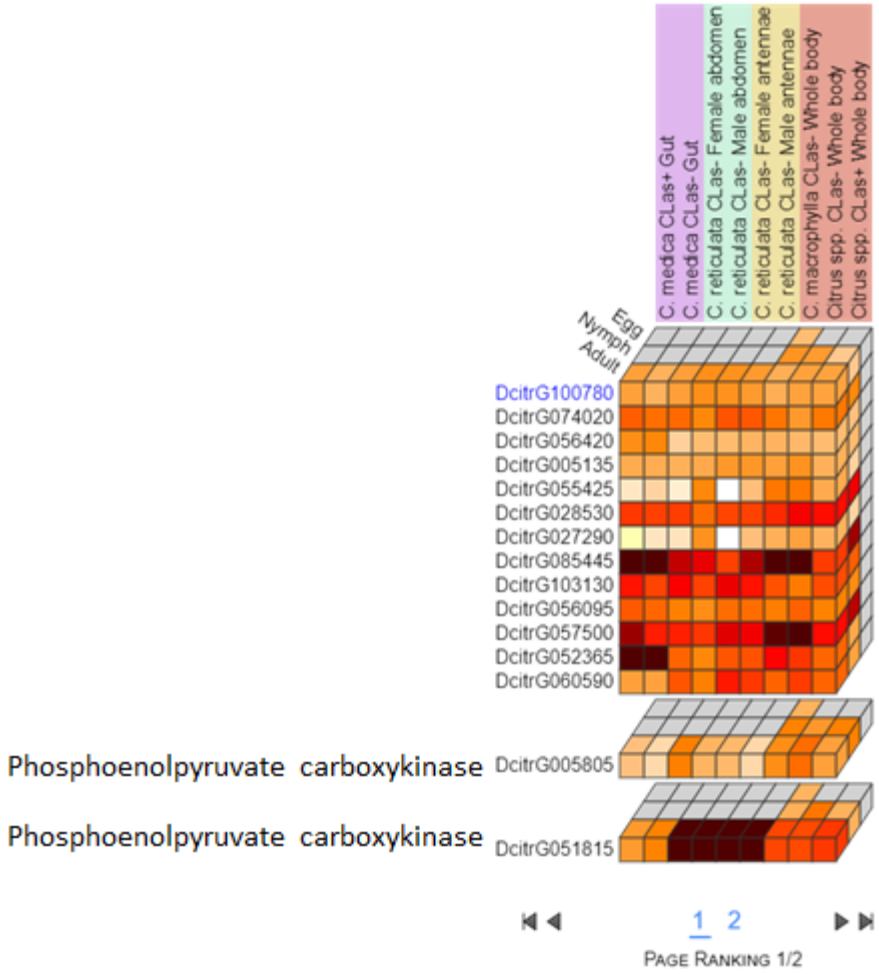


Figure 2: Expression data cube in *D. citri*. Expression of glycolysis and gluconeogenesis genes are shown across three life stages: egg, nymph, and adult. The expression is measured in transcripts per million (TPM) and are measured in different body parts of *D. citri*, such as the gut, abdomen, antennae, and whole body. This is observed in different citrus plant species present with *C*Las (*C*Las +) or without *C*Las (*C*Las -). Each row in the expression cube corresponds to a gene in either glycolysis or gluconeogenesis (not in order to the corresponding pathway).

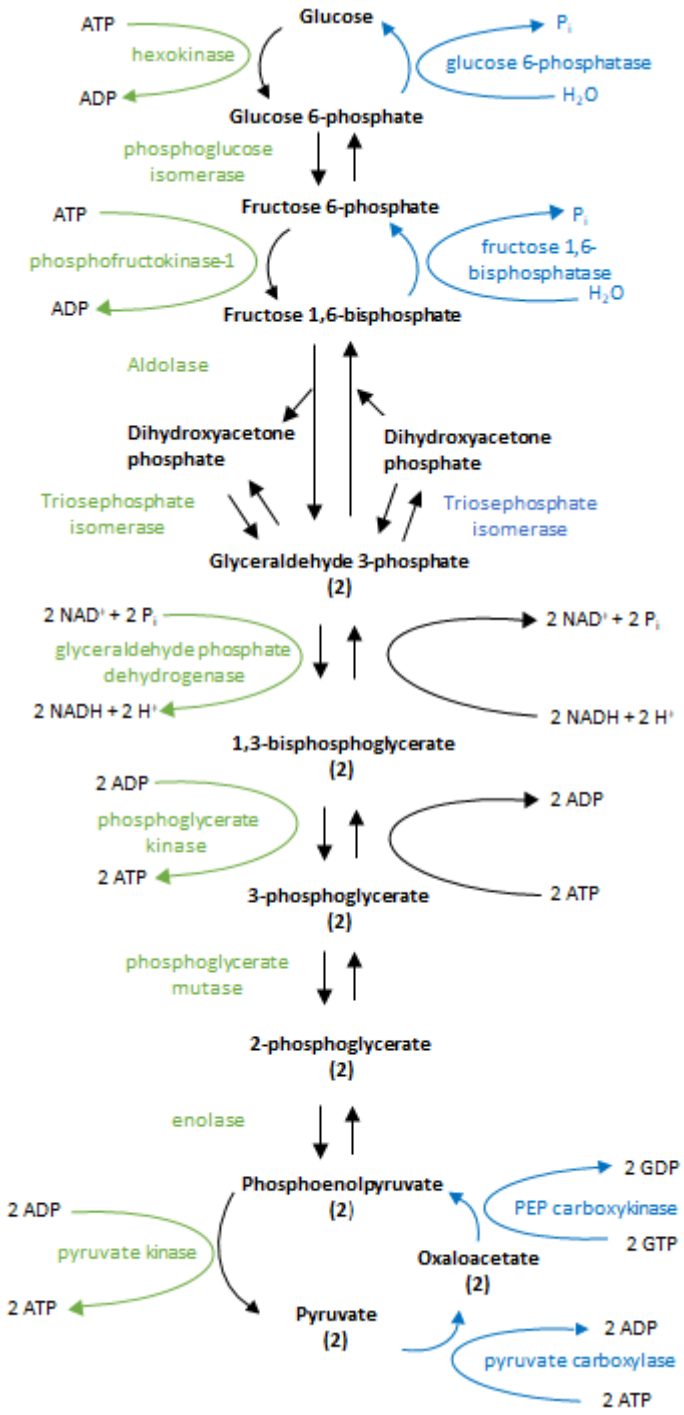

ATP
hexokinase
ADP
Glucose
Glucose 6-phosphate
P_i
glucose 6-phosphatase
H_2O
phosphoglucose isomerase
Fructose 6-phosphate
ATP
phosphofructokinase-1
ADP
Fructose 1,6-bisphosphate
P_i
fructose 1,6-bisphosphatase
H_2O
Aldolase
Dihydroxyacetone phosphate
Dihydroxyacetone phosphate
Triosephosphate isomerase
Triosephosphate isomerase
Glyceraldehyde 3-phosphate (2)
2 NAD^+ + 2 P_i
glyceraldehyde phosphate dehydrogenase
2 NADH + 2 H^+
2 NAD^+ + 2 P_i
2 NADH + 2 H^+
1,3-bisphosphoglycerate (2)
2 ADP
phosphoglycerate kinase
2 ATP
2 ADP
2 ATP
3-phosphoglycerate (2)
phosphoglycerate mutase
2-phosphoglycerate (2)
enolase
Phosphoenolpyruvate (2)
2 ADP
pyruvate kinase
2 ATP
2 GDP
PEP carboxykinase
2 GTP
Oxaloacetate (2)
Pyruvate (2)
2 ADP
pyruvate carboxylase
2 ATP

Figure 3: Glycolysis and gluconeogenesis pathway. The glycolysis pathway consists of ten enzymes, indicated in green, that convert glucose into pyruvate as a final product. The gluconeogenesis pathway consists of eight enzymes, indicated in blue, where three are unique to the pathway that bypass the irreversible reactions in glycolysis to convert non-carbohydrate molecules into glucose.

References:

1. Kerkut GA. Biochemistry [Internet]. Elsevier; 2013. Available from: https://play.google.com/store/books/details?id=SUTgBAAAQBAJ

2. Adams MD. The Genome Sequence of Drosophila melanogaster [Internet]. Science. 2000. p. 2185–95. Available from: http://dx.doi.org/10.1126/science.287.5461.2185

3. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, et al. The genome sequence of the malaria mosquito Anopheles gambiae. Science [Internet]. 2002;298:129–49. Available from: http://dx.doi.org/10.1126/science.1076181

4. Honeybee Genome Sequencing Consortium. Insights into social insects from the genome of the honeybee Apis mellifera. Nature [Internet]. 2006;443:931–49. Available from: http://dx.doi.org/10.1038/nature05260

5. Kunieda T, Fujiyuki T, Kucharski R, Foret S, Ament SA, Toth AL, et al. Carbohydrate metabolism genes and pathways in insects: insights from the honey bee genome. Insect Mol Biol [Internet]. 2006;15:563–76. Available from: http://dx.doi.org/10.1111/j.1365-2583.2006.00677.x

6. Rockstein M. Biochemistry of Insects [Internet]. Elsevier; 2012. Available from: https://play.google.com/store/books/details?id=gNwC1D8deYUC

7. Miyamoto T, Amrein H. Gluconeogenesis: An ancient biochemical pathway with a new twist. Fly [Internet]. 2017;11:218–23. Available from: http://dx.doi.org/10.1080/19336934.2017.1283081

8. Douglas AE. Phloem-sap feeding by animals: problems and solutions. J Exp Bot [Internet]. 2006;57:747–54. Available from: http://dx.doi.org/10.1093/jxb/erj067

9. Hosmani PS, Shippy T, Miller S, Benoit JB, Munoz-Torres M, Flores-Gonzalez M, et al. A quick guide for student-driven community genome annotation. PLoS Comput Biol [Internet]. 2019;15:e1006682. Available from: http://dx.doi.org/10.1371/journal.pcbi.1006682

10. Gilmour D, Others. The biochemistry of insects. The biochemistry of insects [Internet]. New York, NY & London, Academic Press; 1961; Available from: https://www.cabdirect.org/cabdirect/abstract/19610500595

11. Nation, Nation. Insect Physiology and Biochemistry [Internet]. 2001. Available from: http://dx.doi.org/10.1201/9781420058376

12. Duvernell DD, Eanes WF. Contrasting molecular population genetics of four hexokinases in Drosophila melanogaster, D. simulans and D. yakuba. Genetics [Internet]. 2000;156:1191–201. Available from: https://www.ncbi.nlm.nih.gov/pubmed/11063694

13. Fraga A, Ribeiro L, Lobato M, Santos V, Silva JR, Gomes H, et al. Glycogen and glucose metabolism are essential for early embryonic development of the red flour beetle Tribolium castaneum. PLoS One [Internet]. 2013;8:e65125. Available from: http://dx.doi.org/10.1371/journal.pone.0065125

14. Moser D, Johnson L, Lee CY. Multiple forms of Drosophila hexokinase. Purification, biochemical and immunological characterization. J Biol Chem [Internet]. 1980;255:4673–9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/6768723

15. Watt WB. Adaptation at specific loci. I. Natural selection on phosphoglucose isomerase of Colias butterflies: Biochemical and population aspects. Genetics [Internet]. 1977;87:177–94. Available from: https://www.ncbi.nlm.nih.gov/pubmed/914029

16. Haag CR, Saastamoinen M, Marden JH, Hanski I. A candidate locus for variation in dispersal rate in a butterfly metapopulation. Proceedings of the Royal Society B: Biological Sciences [Internet]. Royal Society; 2005;272:2449–56. Available from: https://doi.org/10.1098/rspb.2005.3235

17. Niitepõld K, Smith AD, Osborne JL, Reynolds DR, Carreck NL, Martin AP, et al. Flight metabolic rate and Pgi genotype influence butterfly dispersal rate in the field. Ecology [Internet]. 2009;90:2223–32. Available from: http://dx.doi.org/10.1890/08-1498.1

18. Kallioniemi E, Hanski I. Interactive effects of Pgi genotype and temperature on larval growth and survival in the Glanville fritillary butterfly. Funct Ecol [Internet]. Wiley Online Library; 2011;25:1032–9. Available from: https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2435.2011.01854.x

19. Currie PD, Sullivan DT. Structure and expression of the gene encoding phosphofructokinase (PFK) in Drosophila melanogaster. J Biol Chem [Internet]. 1994;269:24679–87. Available from: https://www.ncbi.nlm.nih.gov/pubmed/7929140

20. Shaw-Lee R, Lissemore JL, Sullivan DT, Tolan DR. Alternative splicing of fructose 1,6-bisphosphate aldolase transcripts in Drosophila melanogaster predicts three isozymes. J Biol Chem [Internet]. 1992;267:3959–67. Available from: https://www.ncbi.nlm.nih.gov/pubmed/1740444

21. Miller D, Hannon C, Ganetzky B. A mutation in Drosophila Aldolase causes temperature-sensitive paralysis, shortened lifespan, and neurodegeneration. J Neurogenet [Internet]. 2012;26:317–27. Available from: http://dx.doi.org/10.3109/01677063.2012.706346

22. Shaw-Lee RL, Lissemore JL, Sullivan DT. Structure and expression of the triose phosphate isomerase (Tpi) gene of Drosophila melanogaster. Mol Gen Genet [Internet]. 1991;230:225–9. Available from: http://dx.doi.org/10.1007/bf00290672

23. O'Brien SJ, MacIntyre RJ. Genetics and biochemistry of enzymes and specific proteins of Drosophila. Genetics and biology of Drosophila [Internet]. 1978;2a. Available from:

http://agris.fao.org/agris-search/search.do?recordID=US201301393035

24. Ercolani L, Florence B, Denaro M, Alexander M. Isolation and complete sequence of a functional human glyceraldehyde-3-phosphate dehydrogenase gene. J Biol Chem [Internet]. 1988;263:15335–41. Available from: https://www.ncbi.nlm.nih.gov/pubmed/3170585

25. Sun XH, Tso JY, Lis J, Wu R. Differential regulation of the two glyceraldehyde-3-phosphate dehydrogenase genes during Drosophila development. Mol Cell Biol [Internet]. 1988;8:5200–5. Available from: http://dx.doi.org/10.1128/mcb.8.12.5200

26. Roselli-Rehfuss L, Ye F, Lissemore JL, Sullivan DT. Structure and expression of the phosphoglycerate kinase (Pgk) gene of Drosophila melanogaster. Mol Gen Genet [Internet]. 1992;235:213–20. Available from: http://dx.doi.org/10.1007/bf00279363

27. Blake CCF, Evans PR, Scopes RK. Structure of Horse-muscle Phosphoglycerate Kinase at 6 Å Resolution [Internet]. Nature New Biology. 1972. p. 195–8. Available from: http://dx.doi.org/10.1038/newbio235195a0

28. Bryant TN, Watson HC, Wendell PL. Structure of yeast phosphoglycerate kinase. Nature [Internet]. 1974;247:14–7. Available from: http://dx.doi.org/10.1038/247014a0

29. Wojtas KM, von Kalm L, Weaver JR, Sullivan DT. The evolution of duplicate glyceraldehyde-3-phosphate dehydrogenase genes in Drosophila. Genetics [Internet]. 1992;132:789–97. Available from: https://www.ncbi.nlm.nih.gov/pubmed/1468631

30. Jedrzejas MJ. Structure, function, and evolution of phosphoglycerate mutases: comparison with fructose-2,6-bisphosphatase, acid phosphatase, and alkaline phosphatase. Prog Biophys Mol Biol [Internet]. 2000;73:263–87. Available from: http://dx.doi.org/10.1016/s0079-6107(00)00007-9

31. Currie PD, Sullivan DT. Structure, expression and duplication of genes which encode phosphoglyceromutase of Drosophila melanogaster. Genetics [Internet]. 1994;138:352–63. Available from: https://www.ncbi.nlm.nih.gov/pubmed/7828819

32. Wang W-X, -X. Wang W, -L. Li K, Chen Y, -X. Lai F, Fu Q. Identification and Function Analysis of enolase Gene NlEno1 from Nilaparvata lugens (Stal) (Hemiptera:Delphacidae) [Internet]. Journal of Insect Science. 2015. p. 66–66. Available from: http://dx.doi.org/10.1093/jisesa/iev046

33. Ohta S, Nishikawa A, Imamura K. Molecular cloning and expression of pyruvate kinase from globefish (Fugu rubripes) skeletal muscle. Comp Biochem Physiol B Biochem Mol Biol [Internet]. 2003;135:397–405. Available from: http://dx.doi.org/10.1016/s1096-4959(03)00095-2

34. Harrison JM. Caste-Specific Changes in Honeybee Flight Capacity. Physiol Zool [Internet]. The University of Chicago Press; 1986;59:175–87. Available from: https://doi.org/10.1086/physzool.59.2.30156031

35. Ge L-Q, Huang B, Li X, Gu H-T, Zheng S, Zhou Z, et al. Silencing pyruvate kinase (NlPYK) leads to reduced fecundity in brown planthoppers, Nilaparvata lugens (Stål) (Hemiptera: Delphacidae). Arch Insect Biochem Physiol [Internet]. 2017;96. Available from: http://dx.doi.org/10.1002/arch.21429

36. Soliman MH, Hani Soliman M, van Herrewege J. Nutrition [Internet]. Drosophila as a Model Organism for Ageing Studies. 1988. p. 46–58. Available from: http://dx.doi.org/10.1007/978-1-4899-2683-8_4

37. Tixier V, Bataillé L, Etard C, Jagla T, Weger M, Daponte JP, et al. Glycolysis supports embryonic muscle growth by promoting myoblast fusion. Proc Natl Acad Sci U S A [Internet]. 2013;110:18982–7. Available from: http://dx.doi.org/10.1073/pnas.1301262110

38. Crabtree B, Higgins SJ, Newsholme EA. The activities of pyruvate carboxylase, phosphoenolpyruvate carboxylase and fructose diphosphatase in muscles from vertebrates and invertebrates. Biochem J [Internet]. 1972;130:391–6. Available from: http://dx.doi.org/10.1042/bj1300391

39. Kola VSR, Renuka P, Madhav MS, Mangrauthia SK. Key enzymes and proteins of crop insects as candidate for RNAi based gene silencing. Front Physiol [Internet]. 2015;6:119. Available from: http://dx.doi.org/10.3389/fphys.2015.00119

40. Bartok O, Teesalu M, Ashwall-Fluss R, Pandey V, Hanan M, Rovenko BM, et al. The transcription factor Cabut coordinates energy metabolism and the circadian clock in response to sugar sensing. EMBO J [Internet]. 2015;34:1538–53. Available from: http://dx.doi.org/10.15252/embj.201591385

41. Hou Y, Wang X-L, Saha TT, Roy S, Zhao B, Raikhel AS, et al. Temporal Coordination of Carbohydrate Metabolism during Mosquito Reproduction [Internet]. PLOS Genetics. 2015. p. e1005309. Available from: http://dx.doi.org/10.1371/journal.pgen.1005309

42. Stark R, Guebre-Egziabher F, Zhao X, Feriod C, Dong J, Alves TC, et al. A role for mitochondrial phosphoenolpyruvate carboxykinase (PEPCK-M) in the regulation of hepatic gluconeogenesis. J Biol Chem [Internet]. 2014;289:7257–63. Available from: http://dx.doi.org/10.1074/jbc.C113.544759

43. Newsholme EA, Crabtree B, Higgins SJ, Thornton SD, Start C. The activities of fructose diphosphatase in flight muscles from the bumble-bee and the role of this enzyme in heat generation. Biochem J [Internet]. 1972;128:89–97. Available from: http://dx.doi.org/10.1042/bj1280089

44. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets [Internet]. Molecular Biology and Evolution. 2016. p. 1870–4. Available from: http://dx.doi.org/10.1093/molbev/msw054