

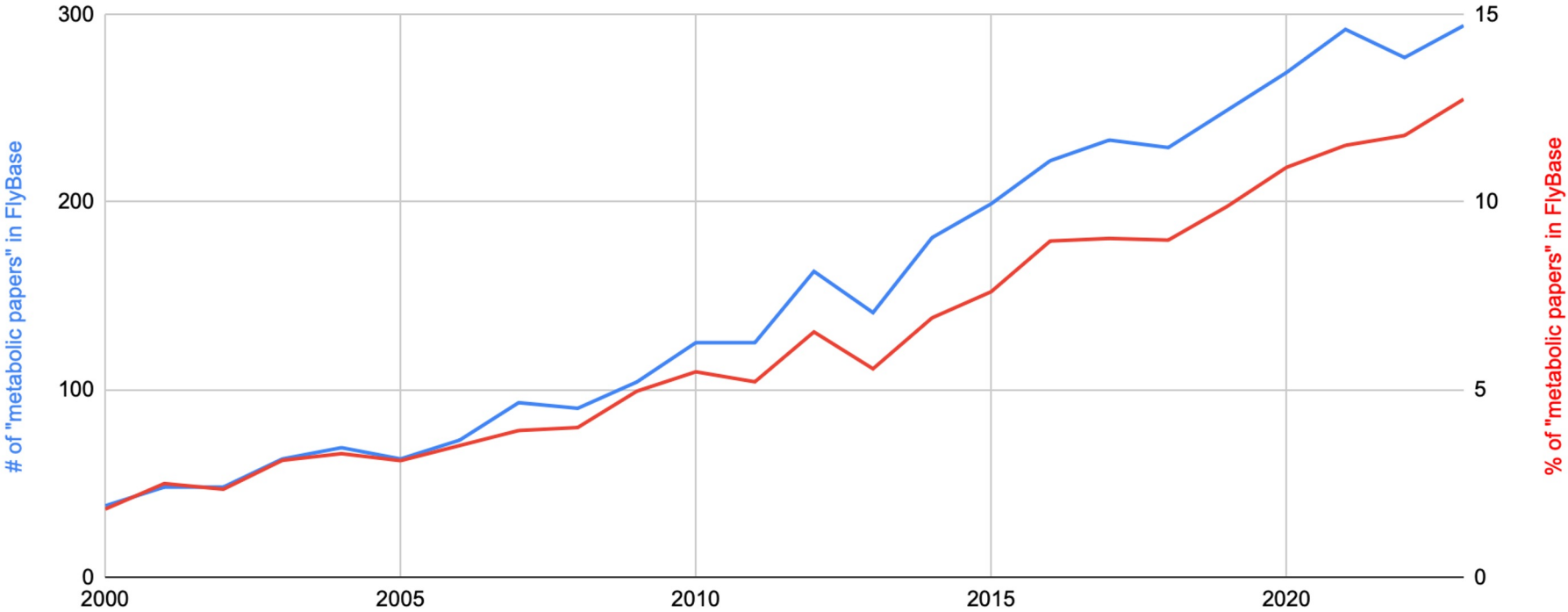


Building better metabolic pathway resources for *Drosophila melanogaster*

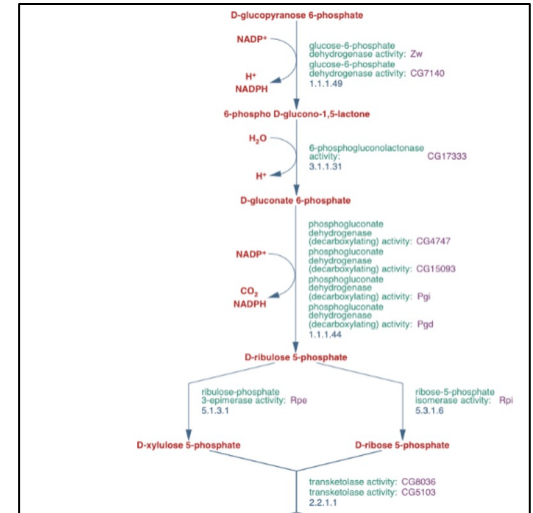
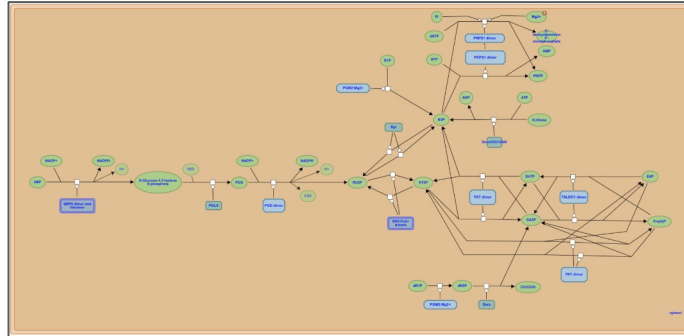
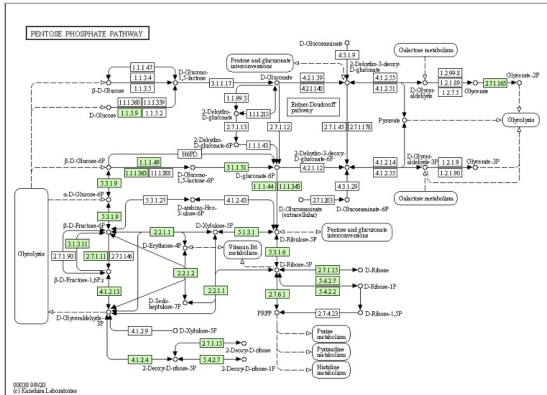
Steven Marygold
sjm41@cam.ac.uk

TAGC - March 2024

Metabolic research in Drosophila

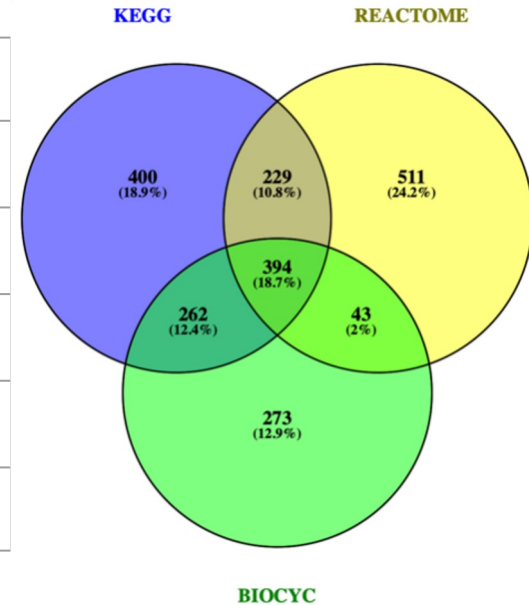


Metabolic pathway resources for Drosophila



Metabolic pathway resources for Drosophila (Nov 2023)

| | KEGG | Reactome | BioCyc (FlyCyc) |
|--------------------------|-----------|-----------|-----------------------|
| # fly pathways | 88 | ~200 | 227 |
| # fly genes | 1,285 | 1,177 | 972 |
| Method | Orthology | Orthology | Functional annotation |
| Last updated | n/a | n/a | ~15 years ago |
| Integrated with FlyBase? | No | No | No |



- Difficult to compare
- When do so, there are significant discrepancies
- None are based on current functional (GO) data in FlyBase

Metabolic annotation using the Gene Ontology (GO)

Three aspects of the GO:

- Molecular Function (MF)

e.g. *6-phosphogluconolactonase activity (GO:0017057)*

- Biological Process (BP)

e.g. *pentose-phosphate shunt (GO:0006098)*

- Cellular Component (CC)

e.g. *cytosol (GO:0005829)*

GO terms include cross-references:

id: GO:0017057

name: 6-phosphogluconolactonase activity

xref: EC:3.1.1.31

xref: RHEA:12556

id: GO:0006098

name: pentose-phosphate shunt

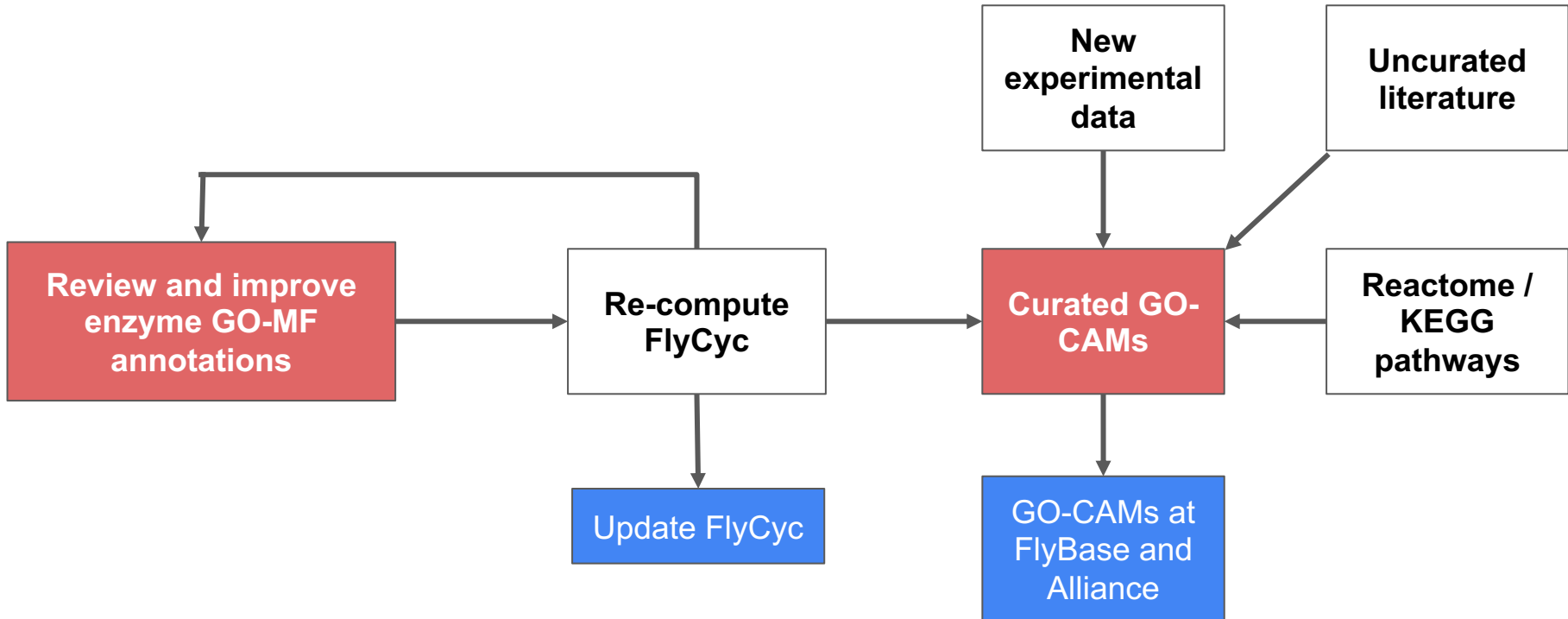
xref: KEGG_PATHWAY:map00030

xref: Reactome:R-HSA-71336

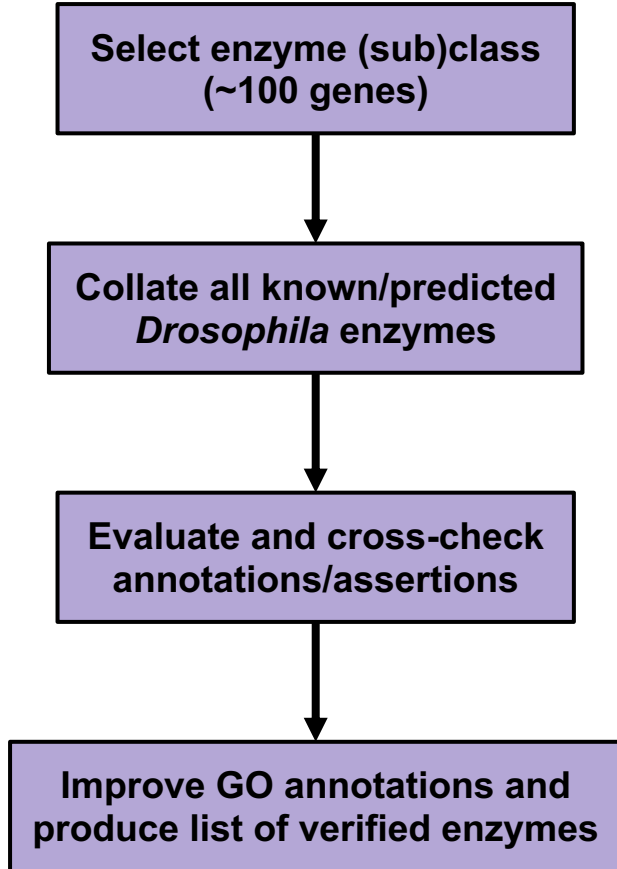
xref: MetaCyc:PENTOSE-P-PWY

Therefore, we can use the GO to annotate metabolic genes, build pathways, and improve & inter-relate existing resources.

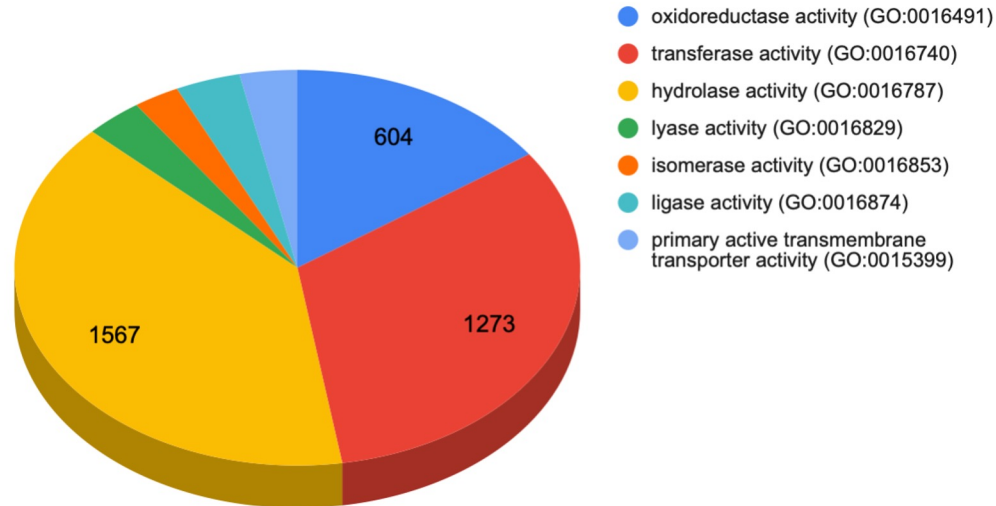
Aim: Produce a high-quality, manually reviewed metabolic pathway resource for Drosophila



Summary of enzyme annotation review (FB2023_06)



| Summary statistics | |
|--------------------------------------|--------|
| # genes reviewed | ~4,500 |
| # verified as enzyme-encoding | ~3,750 |
| # non-enzyme-encoding | ~700 |
| # newly annotated enzymes in FlyBase | ~500 |



Enzyme improvements in FlyBase

| General Information | | |
|--|--|--|
| Symbol | Dmel\Pfk | |
| Name | Phosphofructokinase | |
| Feature Type | protein_coding_gene | |
| Gene Model Status | Current | |
| Enzyme Name (EC) | 6-phosphofructokinase (2.7.1.11) | |
| Species | <i>D. melanogaster</i> | |
| Annotation Symbol | CG4001 | |
| FlyBase ID | FBgn0003071 | |
| Stock Availability | 13 publicly available | |
| Function | | |
| Gene Ontology (GO) Annotations (14 terms) | | |
| Molecular Function (6 terms) | | |
| Terms Based on Experimental Evidence (1 term) | | |
| CV Term | Evidence | References |
| enables 6-phosphofructokinase activity | inferred from mutant phenotype | (Currie and Sullivan, 1994) |
| Biological Process (6 terms) | | |
| Terms Based on Predictions or Assertions (4 terms) | | |
| CV Term | Evidence | References |
| involved_in canonical glycolysis | inferred from biological aspect of ancestor with PANTHER:PTN000344216 | (Gaudet et al., 2011) |
| Cellular Component (2 terms) | | |
| Terms Based on Predictions or Assertions (2 terms) | | |
| CV Term | Evidence | References |
| part_of 6-phosphofructokinase complex | inferred by curator from GO:0003872 inferred from biological aspect of ancestor with PANTHER:PTN000344216 | (Currie and Sullivan, 1994) (Gaudet et al., 2011) |
| located_in cytoplasm | inferred from electronic annotation with InterPro:IPR009161 | (InterPro Project Members, 2004-) |
| Gene Group (FlyBase) | CARBOHYDRATE KINASES | |
| Catalytic Activity (EC/Rhea) | <p>6-phosphofructokinase activity</p> <p>ATP + beta-D-fructose 6-phosphate = ADP + beta-D-fructose 1,6- bisphosphate + H(+) (2.7.1.11)</p> <p>RHEA 16109:</p> | |

Enzyme improvements in FlyBase

QuickSearch

Human Disease Protein Domains **Gene Groups** Pathways GO Data Class

Search FlyBase Homologs GAL4 etc Expression Phenotype References

Search using a gene or Gene Group symbol, name, synonym or ID: ? Search

Enter text: ENZYMES YouTube

Alternatively, browse all [Gene Group reports](#)

Note: Wild cards (*) can be added to your search term

- FB2024_01, released February 22, 2024 → Gene Group List
- ENZYMES (ENZ)
 - HYDROLASES (HYD)
 - ACID ANHYDRIDE HYDROLASES (AAH)
 - ACYLPHOSPHATASES (ACYP)
 - ADP-RIBOSE DIPHOSPHATASES (ADPRASE)
 - DINUCLEOTIDE PHOSPHATASES (NT-PP)
 - NUCLEOSIDE DIPHOSPHATE PHOSPHATASES (NDPASE)
 - NUCLEOSIDE TRIPHOSPHATE DIPHOSPHATASES (NTPDPASE)
 - NUCLEOSIDE TRIPHOSPHATE PHOSPHATASES (NTPASE)
 - ATPASES (ATPASE)
 - ATP-BINDING CASSETTE TRANSPORTER-LIKE - NON-TRANSPORTING (ABC-NT)
 - ABC2 ATP-BINDING CASSETTE TRANSPORTER SUBFAMILY (ABC2)
 - ABCF ATP-BINDING CASSETTE TRANSPORTER SUBFAMILY (ABCF)
 - ATP-DEPENDENT DNA DAMAGE SENSORS (DDS)
 - MUTL ATPASES (MUTL)
 - MUTS ATPASES (MUTS)
 - RAD51-LIKE ATPASES (RAD51)
 - ATP-DEPENDENT FE-S CLUSTER CHAPERONES (FCC)
 - ATP-DEPENDENT PROTEIN FOLDING CHAPERONES (PFC)
 - ATYPICAL HEAT SHOCK PROTEIN 70 CHAPERONES (HSP70-A)
 - HEAT SHOCK PROTEIN 60 CHAPERONINS (HSP60)
 - CHAPERONIN CONTAINING TCP-1 COMPLEX (CCT)
 - HEAT SHOCK PROTEIN 60 CHAPERONINS - GROUP I (HSP60-I)
 - HEAT SHOCK PROTEIN 70 CHAPERONES (HSP70)
 - HEAT SHOCK PROTEIN 90 CHAPERONES (HSP90)
 - HEAT SHOCK PROTEIN 100 CHAPERONES (HSP100)
 - DISLOCASES (DISL)
 - N-ETHYLMALAIMIDE-SENSITIVE FACTORS (NSF)
 - RFC COMPLEX ATPASES (RFC-ATPASE)
 - SNF2-LIKE CHROMATIN REMODELERS (SNF2)
 - STRUCTURAL MAINTENANCE OF CHROMOSOMES GENE FAMILY (SMC)
 - UNCLASSIFIED ATPASES (ATPASE-U)
 - GTPASES (GTPASE)

| General Information | | | |
|------------------------------|--|-----------------------------------|---|
| Name | PYRUVATE KINASES | Species | <i>D. melanogaster</i> |
| Symbol | PK | FlyBase ID | FBgg0001264 |
| Date last reviewed | 2023-09-18 | Number of members | 6 |
| Description | | | |
| Description | Pyruvate kinases catalyze the transfer of phosphate group, usually from ATP, to a pyruvate substrate molecule. | | |
| Notes on Group | There are six paralogous genes encoding pyruvate kinase in <i>D. melanogaster</i> : expression of <i>PyK</i> is ubiquitous, whereas expression of <i>CG7362</i> , <i>CG2964</i> , <i>CG11249</i> , <i>CG12229</i> and <i>CG7069</i> is restricted to the testis. | | |
| Source Material | The PYRUVATE KINASES Gene Group has been compiled using the following publication(s): <i>Heidarian et al., 2023</i> . | | |
| Key Gene Ontology (GO) terms | | | |
| Molecular Function | pyruvate kinase activity | | |
| Biological Process | | | |
| Cellular Component | | | |
| Enzymatic activity | | | |
| Enzyme name (EC) | pyruvate kinase (2.7.1.40) | | |
| Related Gene Groups | | | |
| Parent group(s) | PHOSPHORUS-CONTAINING GROUP TRANSFERASES | | |
| Members (6) | | | |
| For all members: | View Orthologs | Export to HitList | Export to Batch Download |
| GO ribbon stack | | | |
| Gene Symbol | Gene Name | Also Known As | Source Material for Membership |
| CG2964 | | | <i>(Heidarian et al., 2023, FlyBase, 2017-)</i> |
| CG7069 | | | <i>(Heidarian et al., 2023, FlyBase, 2017-)</i> |
| CG7362 | | | <i>(Heidarian et al., 2023, FlyBase, 2017-)</i> |
| CG11249 | | | <i>(Heidarian et al., 2023, FlyBase, 2017-)</i> |
| CG12229 | | | <i>(Heidarian et al., 2023, FlyBase, 2017-)</i> |
| Pyk | Pyruvate kinase | PK | <i>(Heidarian et al., 2023, FlyBase, 2017-)</i> |
| External Data | | | |
| Equivalent Group(s) | | | |
| Other resource(s) | | | |
| Synonyms and Secondary IDs | | | |
| References (3) | | | |

~750 enzyme gene group reports

Update FlyCyc



Change Current Database

Current Database: *Drosophila melanogaster*

Search in Current Database:

Enter a gene, protein, metabolite or pathway

Summary of *Drosophila melanogaster*, version 27.5 Tier 2 Curated Database

Database Authors: Steven J Marygold¹

¹FlyBase

Summary:

FlyCyc is a Pathway/Genome Database (PGDB) of the model organism *Drosophila melanogaster*, based on data from FlyBase, release FB2023_06 (December 2023). Functional information is derived from Gene Ontology (GO) annotation of *D. melanogaster* enzymes and transporters. The genome sequence is the Release 6 assembly provided by the BDGP, with gene model annotations provided by FlyBase. The PGDB was generated by the PathoLogic [1, Karp11, 2] component of Pathway Tools software version 27.5 and MetaCyc version 27.0 on 26-Nov-2023. The FlyBase work is primarily supported by NIH awards U41HG000739 and 1R01DK136945-01.

Taxonomic Lineage: cellular organisms, Eukaryota, Opisthokonta, Metazoa, Eumetazoa, Bilateria, Protostomia, Ecdysozoa, Panarthropoda, Arthropoda, Mandibulata, Pancrustacea, Hexapoda, Insecta, Dicondylia, Pterygota <insects>, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha, Eremoneura, Cyclorhapha, Schizophora, Acalyptratae, Ephydroidea, Drosophilidae, Drosophilinae, Drosophilini, Drosophila <flies>, genus>, Sophophora, melanogaster group, melanogaster subgroup, *Drosophila melanogaster*

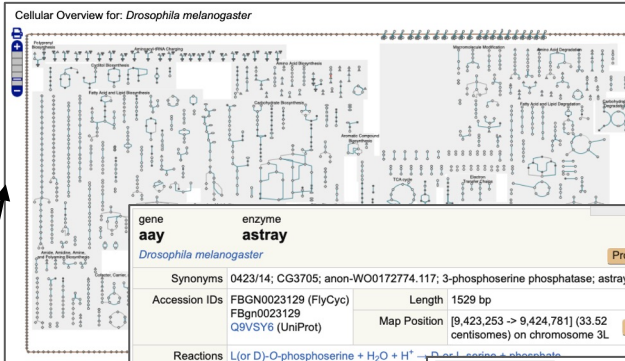
Synonyms: vinegar fly, pomace fly, *Drosophila*

Unification Links: NCBI-Taxonomy:7227

| Replicon | Total Genes | Protein Genes | RNA Genes | Pseudogenes | Size (bp) | NCBI Link |
|----------------------|---------------|---------------|--------------|-------------|--------------------|-----------------------|
| chromosome 3R | 4,226 | 3,421 | 761 | 44 | 32,079,331 | RefSeq:NT_033777.3 |
| chromosome 3L | 3,482 | 2,739 | 702 | 41 | 28,110,227 | RefSeq:NT_037436.4 |
| chromosome 2R | 3,649 | 2,852 | 747 | 50 | 25,286,936 | RefSeq:NT_033778.4 |
| chromosome X | 2,704 | 2,196 | 434 | 74 | 23,542,271 | RefSeq:NC_004354.4 |
| chromosome 2L | 3,508 | 2,657 | 799 | 52 | 23,513,712 | RefSeq:NT_033779.5 |
| chromosome Y | 113 | 23 | 28 | 62 | 3,667,352 | RefSeq:NC_024512.1 |
| chromosome 4 | 114 | 80 | 27 | 7 | 1,348,131 | RefSeq:NC_004353.4 |
| unmapped scaffold 8 | 2 | 2 | 0 | 0 | 88,768 | RefSeq:NW_007931083.1 |
| rDNA | 21 | 0 | 16 | 5 | 76,973 | RefSeq:NW_007931121.1 |
| mitochondrion genome | 37 | 13 | 24 | 0 | 19,524 | RefSeq:NC_024511.2 |
| Total: | 17,856 | 13,983 | 3,538 | 335 | 137,733,225 | |

Ortholog data available? No

| | | | |
|----------------------|--------|------------|----------|
| Genes: | 17,856 | SmartTable | Ontology |
| Pathways: | 290 | SmartTable | Ontology |
| Enzymatic Reactions: | 2,118 | | |
| Transport Reactions: | 30 | | Ontology |
| Polypeptides: | 13,983 | SmartTable | Ontology |
| Protein Complexes: | 0 | SmartTable | Ontology |
| Enzymes: | 2,318 | SmartTable | |
| Transporters: | 247 | SmartTable | |
| Compounds: | 1,475 | SmartTable | Ontology |
| IRNAs: | 312 | | Ontology |
| GO Terms: | 91,040 | | Ontology |



gene
enzyme
astray

Drosophila melanogaster

Synonyms: 0423/14; CG3705; anon-WO0172774.117; 3-phosphoserine phosphatase; astray; 24661601; Aay

Accession IDs: FBGN0023129 (FlyCyc) Length: 1529 bp
FBgn0023129
Q9VSV6 (UniProt) Map Position: [9,423,253 -> 9,424,781] (33.52 centisomes) on chromosome 3L

Reactions: L-(or D)-O-phosphoserine + H₂O + H⁺ → O-phospho-L-serine + H₂O → L-serine

Pathway: L-serine biosynthesis I

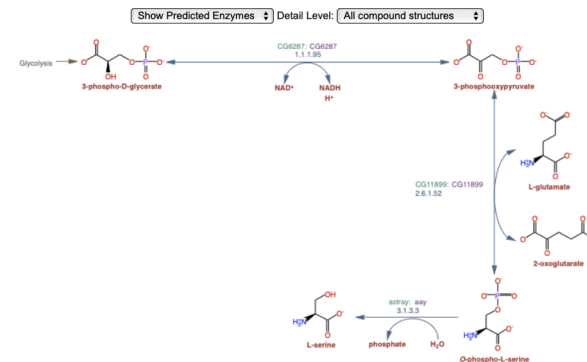
Summary GO Terms (7) Reactions (2) Gene

Gene-Reaction Schematic

3.1.3.3 : L-(or D)-O-phosphoserine + H₂O → L-serine + phosphate

3.1.3.3 : O-phospho-L-serine + H₂O → L-serine + phosphate

Drosophila melanogaster Pathway: L-serine biosynthesis I



If an enzyme name is shown in bold, there is experimental evidence for this enzymatic activity.

BioCyc ID: SERSYN-PWY

Superclasses: Biosynthesis → Amino Acid Biosynthesis → Proteinogenic Amino Acid Biosynthesis → L-serine Biosynthesis

Pathway Summary from MetaCyc:

About This Pathway

L-serine biosynthesis is a major metabolic pathway in many organisms. Its end product, **L-serine**, is not only used in protein synthesis, but also as a precursor for the biosynthesis of glycine, cysteine, tryptophan, and phospholipids. In addition, it directly or indirectly serves as a source of one-carbon units for the biosynthesis of various compounds.

In the pathway described here, which operates in archaea [*Halgodottr07*], bacteria [*H001*], yeast, some protists, plant plastids and animals, serine is derived from **3-phospho-D-glycerate**, an intermediate of glycolysis.

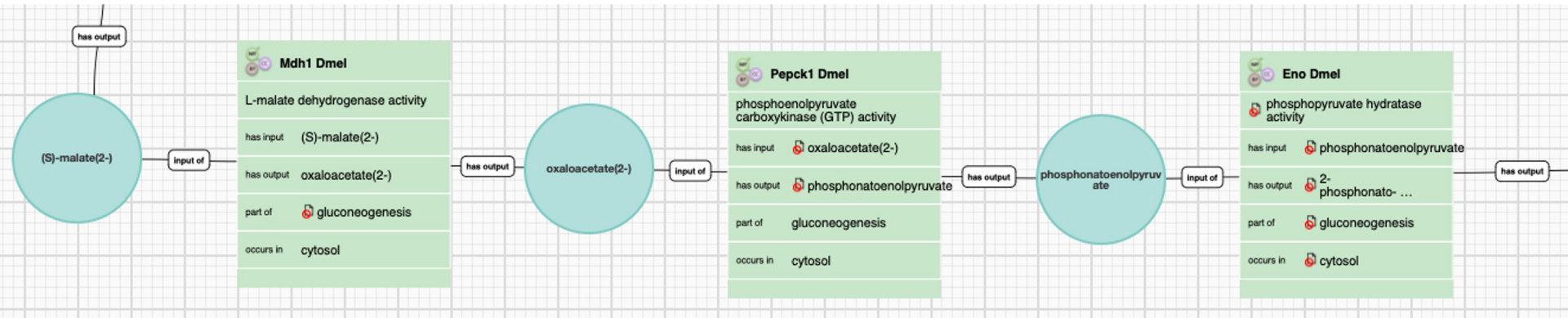
Regulation of the pathway is mainly accomplished by feedback inhibition of the enzyme that catalyzes the first committed step, EC 1.1.1.95, **3-phosphoglycerate dehydrogenase**. The second enzyme in the pathway, EC 2.6.1.52, **phosphoserine transaminase**, requires pyridoxal 5'-phosphate (PLP) as a cofactor and is also required for the biosynthesis of PLP itself. Thus the cell must ensure that the supply of PLP is adequate. Little biochemical work has been done on the final enzyme of the pathway, EC 3.1.3.3, **3-phosphoserine phosphatase**.

Updated FlyCyc links at FlyBase

| | | | |
|---|--|---|--|
| FB2024_01, released February 22, 2024 | | Gene: Dmel\Pfk | |
| [-] Pathways | | | |
| [+] Signaling Pathways (FlyBase) | | | |
| [-] Metabolic Pathways | | | |
| | | FlyCyc Pathways - Pathways from a BioCyc PGDB for Dmel glycolysis III (from glucose) homolactic fermentation | |

GO-Causal Activity Models (GO-CAMs)

- Manually curated models connecting individual GO annotations to form pathways
- Tissue/context-specific pathways
- Actively in use by Model Organism Databases and UniProt
- Already available at the Alliance of Genome Resources website



Metabolic Pathways at the Alliance and FlyBase

Pathways ?

Reactome Pathway (1) Reactome Reactions (1) **GO-CAMs (1)**

Available GO-CAMs:
Pentose-phosphate shunt 1 (Mouse)

Pentose-phosphate shunt 1 (Mouse) Expand Protein Complexes Reset View

Processes and Activities

- Small Molecules
- beta-D-fructofuranose 6-phosphate(2-)
- NADPH(4-)
- pentose-phosphate shunt, oxidative branch

G6pd2 Mmus

glucose-6-phosphate dehydrogenase activity

occurs in cytosol

has output NADPH(4-)

G6pd2 Mmus E

part of pentose-phosphate shunt, oxidative branch

part of pentose-phosphate shunt

G6pdx Mmus

glucose-6-phosphate dehydrogenase activity

G6pdx Mmus EE

occurs in cytosol EE

part of pentose-phosphate shunt, oxidative branch EE

part of pentose-phosphate shunt

has output NADPH(4-)

Pgd Mmus

Relation Types

- Direct regulation
- - - Indirect regulation
- ➔ Positive regulation
- ⊖ Negative regulation
- Positive or negative effect unknown
- Provides substrate for

General Information

| | | | |
|------------------|--|-----------|----|
| Name | Gluconeogenesis | # members | 27 |
| Description | Gluconeogenesis results in the formation of glucose from noncarbohydrate precursors, such as pyruvate, amino acids and glycerol. | | |
| References | | | |
| Key GO term | gluconeogenesis | | |
| Related pathways | Trehalogenesis | | |

Pathway Diagram(s)

Members

| Gene symbol | GO Molecular Function | EC# | Refs |
|---------------------|--|-----------------------------|-------------------------|
| Fbp | fructose 1,6-bisphosphate 1-phosphatase activity | EC:3.1.3.11 | Refs... |
| Pcb | pyruvate carboxylase activity | EC:6.4.1.1 | Refs... |
| etc | | | |

Metabolites

| Name | CHEBI ID | Other pathways |
|------------------------------|-----------------------------|---------------------------|
| pyruvate | CHEBI:15361 | TCA cycle |
| oxaloacetate | CHEBI:16452 | TCA cycle |
| etc | | |

External links

[GO-CAM](#) [FlyCyc](#) [Reactome](#) [KEGG](#)

References

Acknowledgements

- FlyBase team, especially: Rossana Zaru, Phani Garapati, Helen Attrill
 - GO consortium, especially: Pascale Gaudet, Harold Drabkin, Raymond Lee, Marc Feuermann, Sara Chuguransky, Beatriz Pinto, Hsin-Yu Chang, David Hill, Kimberley Van Auken, Paul Thomas, Chris Mungall
 - Reactome: Peter D'Eustachio
 - EC/Rhea: Kristian Axelson, Anne Morgat
 - BioCyc: Peter Karp, Ron Caspi
 - NIDDK grant co-PIs: Jason Tennessen, Norbert Perrimon, Angelo D'Alessandro
-
- Funding: NHGRI (U41HG000739) and NIDDK (1R01DK136945-01)
 - Infrastructure: Department of Physiology, Development and Neuroscience, University of Cambridge



National Institute of
Diabetes and Digestive
and Kidney Diseases



National Human Genome
Research Institute