

## *Drosophila simulans* –*Ilp2*

Leon F. Laskowski<sup>1</sup>, Cole A. Kiser<sup>2</sup>, Robyn Huber<sup>3</sup>, Sean Kusche<sup>4</sup>, Andrew M Arsham<sup>3</sup>, Thomas C. Giarla<sup>4</sup>, Chinmay P. Rele<sup>2§</sup>

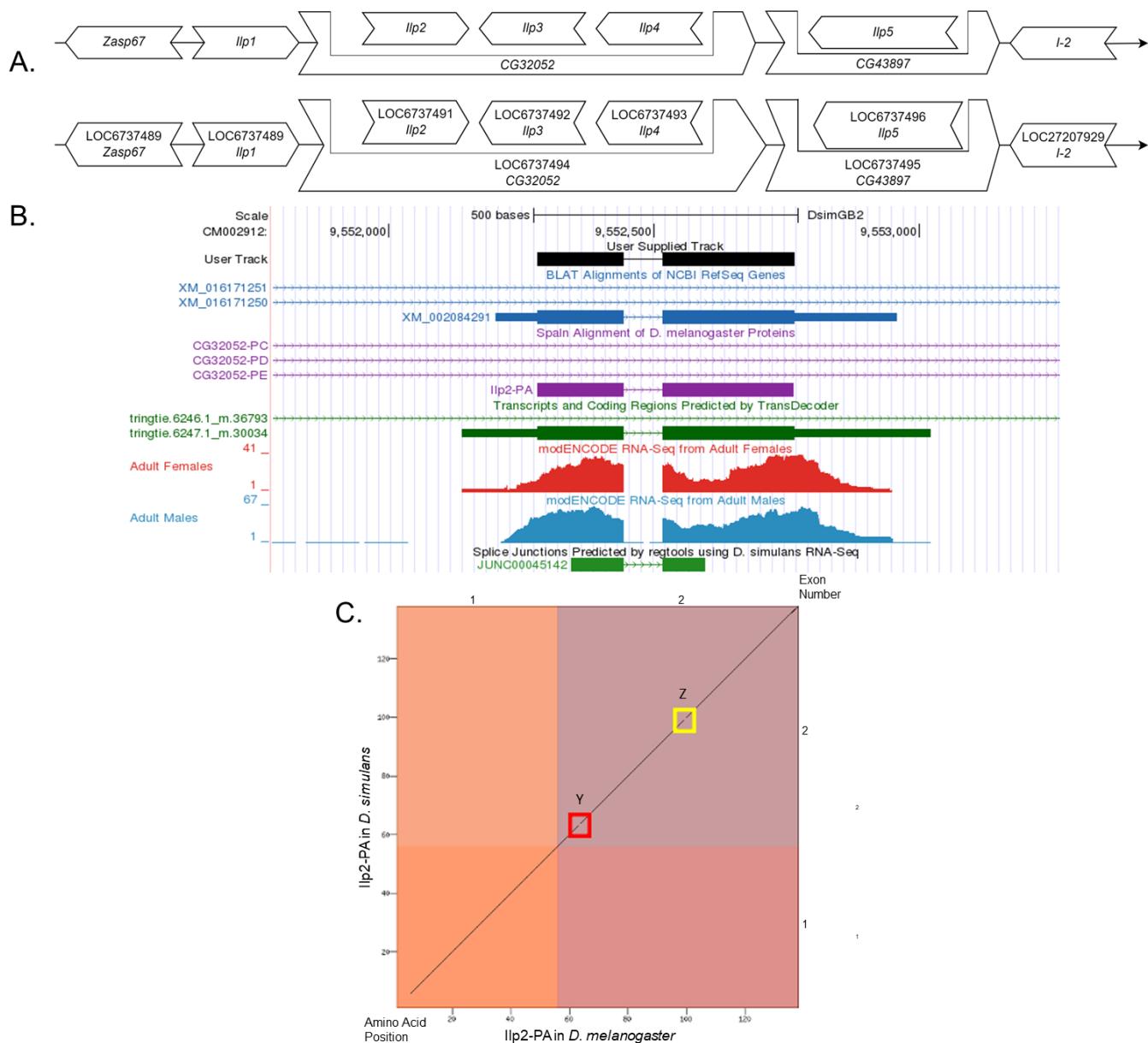
<sup>1</sup>University of St. Francis, Joliet, IL, USA

<sup>2</sup>University of Alabama, Tuscaloosa, AL, USA

<sup>3</sup>Bemidji State University, Bemidji, MN, USA

<sup>4</sup>Siena College, Loudonville, NY, USA

§To whom correspondence should be addressed: cprele@ua.edu



**Figure 1.**

**(A) Synteny comparison of the genomic neighborhoods for *Ilp2* in *Drosophila melanogaster* and *D. simulans*.** Thin underlying arrows indicate the DNA strand within which the target gene—*Ilp2*—is located in *D. melanogaster* (top) and *D. simulans* (bottom).

11/15/2022 - Open Access

*simulans* (bottom). Thin arrow(s) pointing to the right indicate(s) that *Ilp2* is on the positive (+) strand in *D. simulans* and *D. melanogaster*. The wide gene arrows pointing in the same direction as *Ilp2* are on the same strand relative to the thin underlying arrows, while wide gene arrows pointing in the opposite direction of *Ilp2* are on the opposite strand relative to the thin underlying arrows. White gene arrows in *D. simulans* indicate orthology to the corresponding gene in *D. melanogaster*. Gene symbols given in the *D. simulans* gene arrows indicate the orthologous gene in *D. melanogaster*, while the locus identifiers are specific to *D. simulans*. **(B) Gene Model in GEP UCSC Track Data Hub** (Raney *et al.*, 2014). The coding-regions of *Ilp2* in *D. simulans* are displayed in the User Supplied Track (black); coding exons are depicted by thick rectangles and introns by thin lines with arrows indicating the direction of transcription. Subsequent evidence tracks include BLAT Alignments of NCBI RefSeq Genes (dark blue, alignment of Ref-Seq genes for *D. simulans*), Spaln of *D. melanogaster* Proteins (purple, alignment of Ref-Seq proteins from *D. melanogaster*), Transcripts and Coding Regions Predicted by TransDecoder (dark green), RNA-Seq from Adult Females and Adult Males (red and light blue, respectively; alignment of Illumina RNA-Seq reads from *D. simulans*), and Splice Junctions Predicted by regtools using *D. simulans* RNA-Seq (Graveley *et al.*, 2011; SRP006203). Splice junctions shown have a read-depth of 50-99 supporting reads in green. **(C) Dot Plot of Ilp2-PA in *D. melanogaster* (x-axis) vs. the orthologous peptide in *D. simulans* (y-axis)**. Amino acid number is indicated along the left and bottom; coding-exon number is indicated along the top and right, and exons are also highlighted with alternating colors. The boxes highlighted in red and yellow (Box Y; Box Z) indicate lack of sequence similarity between amino acids.

## Description

**Introduction** The Insulin Signaling Pathway is a highly conserved signaling pathway in animals, and is central to nutrient uptake (Hietakangas and Cohen 2009, Grewal 2009). *Ilp2* (aka *dilp2*, *Dilp 2*, *DILP*, *dilp-2*, *Ilp-2*), a member of the Insulin Signaling Pathway, mediates growth by acting as a ligand for the insulin receptor and transducing a signal via the Chico/PI3K/Akt(PKB) pathway (Brogliolo *et al.*, 2001, Park *et al.*, 2014). *Ilp2* plays a role in regulating body size by increasing cell size and cell number of individual organs (Ren *et al.*, 2017). Evolutionary studies showed that loss of *Ilp2* increased lifespan, and changes in *Ilp2* expression may have contributed to the evolution of body size in the Hawaiian *Drosophila* species (Grönke *et al.*, 2010). In the absence of *Ilp2*, over-expression of *Ilp1* and *Ilp3-7* is enough to promote growth in *Drosophila* (Ikeya *et al.*, 2002). *Ilp2* mutants also seem to have severe developmental delay (Grönke *et al.*, 2010).

We propose a gene model for the *D. simulans* ortholog of the *D. melanogaster* Insulin-like peptide 2 (*Ilp2*) gene. The genomic region of the ortholog corresponds to the uncharacterized protein LOC6737491 (RefSeq accession XP\_002084327.1) in the ASM75419v3 Genome Assembly of *D. simulans* (GenBank Accession: GCA\_000754195.3 - Graveley *et al.*, 2011; SRP006203). This model is based on RNA-Seq data from *D. simulans* (Graveley *et al.*, 2011; SRP006203) and *Ilp2* in *D. melanogaster* using FlyBase release FB2022\_04 (GCA\_000001215.4; Larkin *et al.*, 2021). *D. simulans* is part of the melanogaster species group within the subgenus *Sophophora* of the genus *Drosophila* (Sturtevant, 1939; Bock and Wheeler, 1972). It was first described by Sturtevant (1919). *D. simulans* is a sibling species to *D. melanogaster*, thus extensively studied in the context of speciation genetics and evolutionary ecology (Powell, 1990). Historically, *D. simulans* was a tropical species native to sub-Saharan Africa (Lemeunier *et al.*, 1986) where figs served as a primary host (Lachaise and Tsacas, 1983). However, *D. simulans*'s range has expanded world-wide within the last century as a human commensal using a broad range of rotting fruits as breeding sites (<https://www.taxodros.uzh.ch>). The Genomics Education Partnership maintains a mirror of the UCSC Genome Browser (Kent WJ *et al.*, 2002; Gonzalez *et al.*, 2021), which is available at <https://gander.wustl.edu>.

## Synteny

The target gene, *Ilp2*, occurs on chromosome 3L in *D. melanogaster* and is flanked upstream by Z band alternatively spliced PDZ-motif protein 67 (*Zasp67*), Insulin-like peptide (*Ilp1*) and *CG32052*, which nests *Ilp2*, *Ilp3* and *Ilp4*. *Ilp2* is flanked downstream by Insulin-like peptide 3 (*Ilp3*), Insulin-like peptide 4 (*Ilp4*), Insulin-like peptide 5 (*Ilp5*) (nested within *CG43897*), and Inhibitor-2 (*I-2*). The *tblastn* search of *D. melanogaster* *Ilp2*-PA (query) against the *D. simulans* (GenBank Accession: GCA\_000754195.3) Genome Assembly (database) placed the putative ortholog of *Ilp2* within scaffold chromosome 3L (CM002912.1) at locus LOC6737491 (XP\_002084327.1)— with an E-value of 5e-97 and a percent identity of 97.81%. Furthermore, the putative ortholog is flanked upstream by LOC6737489 (XP\_016031285.1), LOC6737490 (XP\_002084326.1) and LOC6737494 (XP\_016031282.1), which correspond to *Zasp67*, *Ilp1* and *CG32052* in *D. melanogaster* (E-value: 0.0, 3e-78 and 0.0; identity: 95.60%, 91.67% and 98.62%, respectively, as determined by *blastp*; Figure 1A, Altschul *et al.*, 1990). The putative ortholog of *Ilp2* is flanked downstream by LOC6737492 (XP\_002084328.1), LOC6737493 (XP\_002084329.1), LOC6737496 (XP\_002084332.2) nested by LOC6737495 (XP\_039149023.1), and LOC27207929 (XP\_016031291.1), which correspond to *Ilp3*, *Ilp4*, *Ilp5*, *CG43897*, and *I-2* in *D. melanogaster* (E-value: 4e-59, 2e-58, 5e-52, 0.0, and 1e-128; identity: 98.33%, 97.35%, 97.22%, 92.64%, and 91.32% respectively, as determined by *blastp*). The putative ortholog assignment for *Ilp2* in *D. simulans* is supported by the following evidence: The genes surrounding the *Ilp2* ortholog

11/15/2022 - Open Access

are orthologous to the genes at the same locus in *D. melanogaster* and local synteny is completely conserved, supported by e-values and percent identities, so we conclude that LOC6737491 is the correct ortholog of *Ilp2* in *D. simulans* (Figure 1A).

### Protein Model

Consistent with the *blastp* search result which shows 97.81% identity between *D. melanogaster* *Ilp2*-PA and the *D. simulans* gene model as well as the low sensitivity parameters used to generate the dot plot (i.e., word size = 3; neighborhood threshold = 11), the dot plot of the two protein sequences contain multiple short gaps along the diagonal (Figure 1C). The red and yellow boxes in the dot plot (Box Y, Box Z; Figure 1C) show lack of sequence similarity between single amino acids near the beginning and middle of coding exon two. *Ilp2* in *D. simulans* has one protein-coding isoform (*Ilp2*-PA; Figure 1B). Isoform (*Ilp2*-PA) contains two protein-coding exons. Relative to the ortholog in *D. melanogaster*, the coding-exon number is conserved. The sequence of *Ilp2*-PA in *D. simulans* has 97.81% identity (E-value: 5e-97) with the protein-coding isoform *Ilp2*-PA in *D. melanogaster*, as determined by *blastp* (Figure 1C). Coordinates of this curated gene model are stored by NCBI at GenBank/BankIt (accession BK059551). These data are also archived in the CaltechDATA repository (see “Extended Data” section below).

**Acknowledgements:** We would like to thank Wilson Leung for developing and maintaining the technological infrastructure that was used to create this gene model, Madeline Gruys for retrofitting this model and Laura K. Reed for overseeing the project.

### Extended Data

Description: FASTA, PEP, GFF. Resource Type: Model. File: [DsimGB2\\_Ilp2.zip](#). DOI: [10.22002/qs6vt-d5d09](https://doi.org/10.22002/qs6vt-d5d09)

### References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). "Basic local alignment search tool." *J. Mol. Biol.* 215:403-410. PMID: 2231712 PubMed ID: [2231712](#)

Bock, I.R., Wheeler, M.R. (1972). The *Drosophila melanogaster* species group. *Univ. Texas Publs Stud. Genet.* 7(7213): 1-102. FBrf0024428 PubMed ID: [null](#)

Brogiolo W, Stocker H, Ikeya T, Rintelen F, Fernandez R, Hafen E. An evolutionarily conserved function of the *Drosophila* insulin receptor and insulin-like peptides in growth control. *Curr Biol.* 2001 Feb 20;11(4):213-21. doi: 10.1016/s0960-9822(01)00068-9. PMID: 11250149. PubMed ID: [11250149](#)

*Drosophila* 12 Genomes Consortium, Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., ..., MacCallum, I. (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450(7167): 203--218. PMID: 17994087 PubMed ID: [17994087](#)

Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, Yang L, Artieri CG, van Baren MJ, Boley N, Booth BW, Brown JB, Cherbas L, Davis CA, Dobin A, Li R, Lin W, Malone JH, Mattiuzzo NR, Miller D, Sturgill D, Tuch BB, Zaleski C, Zhang D, Blanchette M, Dudoit S, Eads B, Green RE, Hammonds A, Jiang L, Kapranov P, Langton L, Perrimon N, Sandler JE, Wan KH, Willingham A, Zhang Y, Zou Y, Andrews J, Bickel PJ, Brenner SE, Brent MR, Cherbas P, Gingeras TR, Hoskins RA, Kaufman TC, Oliver B, Celniker SE. The developmental transcriptome of *Drosophila melanogaster*. *Nature*. 2011 Mar 24;471(7339):473-9. doi: 10.1038/nature09715. Epub 2010 Dec 22. PMID: 21179090; PMCID: PMC3075879. PubMed ID: [21179090](#)

Grewal SS. (2009) Insulin/TOR signaling in growth and homeostasis: a view from the fly world. FBrf0207364 PMID: 18992839 PubMed ID: [18992839](#)

Grönke S, Clarke DF, Broughton S, Andrews TD, Partridge L. Molecular evolution and functional characterization of *Drosophila* insulin-like peptides. *PLoS Genet.* 2010 Feb 26;6(2):e1000857. doi: 10.1371/journal.pgen.1000857. PMID: 20195512; PMCID: PMC2829060. PubMed ID: [20195512](#)

Hietakangas V, Cohen SM. (2009) Regulation of tissue growth through nutrient sensing. PMID: 19694515 PubMed ID: [19694515](#)

Ikeya T, Galic M, Belawat P, Nairz K, Hafen E. Nutrient-dependent expression of insulin-like peptides from neuroendocrine cells in the CNS contributes to growth regulation in *Drosophila*. *Curr Biol.* 2002 Aug 6;12(15):1293-300. doi: 10.1016/s0960-9822(02)01043-6. PMID: 12176357. PubMed ID: [12176357](#)

Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M., Haussler, D. (2002). The Human Genome Browser at UCSC. *Genome Res.* 12, 996-1006. PMID: 12045153 PubMed ID: [12045153](#)

11/15/2022 - Open Access

Lachaise, D., Tsacas, L. (1983). Breeding-sites of tropical African Drosophilids. Ashburner, Carson, Thompson, 1981-1986 d: 221--332. FBrf0038884 PubMed ID: [null](#)

Larkin A, Marygold SJ, Antonazzo G, Attrill H, dos Santos G, Garapati PV, Goodman JL, Gramates LS, Millburn G, Strelets VB, Tabone CJ, and Thurmond J and the FlyBase Consortium (2021). FlyBase: updates to the *Drosophila melanogaster* knowledge base. *Nucleic Acids Res.* 49(D1) D899–D907. PMID: 33219682 PubMed ID: [33219682](#)

Lemeunier, F., David, J., Tsacas, L., Ashburner, M. (1986). The melanogaster species group. Ashburner, Carson, Thompson, 1981-1986 e: 147–256. FBrf0043749 PubMed ID: [null](#)

Navarro Gonzalez, J., Zweig, A. S., Speir, M. L., Schmelter, D., Rosenbloom, K. R., Raney, B. J., et al. (2021). The UCSC Genome Browser database: 2021 update. *Nucleic Acids Research*. 49(1), 1046-1057. PMID: 33221922 PubMed ID: [33221922](#)

Park S, Alfa RW, Topper SM, Kim GE, Kockel L, Kim SK. A genetic strategy to measure circulating *Drosophila* insulin reveals genes regulating insulin production and secretion. *PLoS Genet.* 2014 Aug 7;10(8):e1004555. doi: 10.1371/journal.pgen.1004555. PMCID: PMC4125106; PMID: 25101872 PubMed ID: [25101872](#)

Powell, Jeffrey R. "Progress and prospects in evolutionary biology: the *Drosophila* model." (1997).I PubMed ID: [null](#)

Raney BJ, Dreszer TR, Barber GP, Clawson H, Fujita PA, Wang T, Nguyen N, Paten B, Zweig AS, Karolchik D, Kent WJ. Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC Genome Browser. *Bioinformatics*. 2014 Apr 1;30(7):1003-5. PMID: 24227676 PubMed ID: [24227676](#)

Rele CP, Sandlin KM, Leung W, Reed LK. (2020). Manual Annotation of Genes within *Drosophila* Species: the Genomics Education Partnership protocol. *bioRxiv* 2020.12.10.420521 doi: 10.1101/2020.12.12.420521 PubMed ID: [null](#)

Ren S, Huang Z, Jiang Y, Wang T. dTBC1D7 regulates systemic growth independently of TSC through insulin signaling. *J Cell Biol.* 2018 Feb 5;217(2):517-526. doi: 10.1083/jcb.201706027. Epub 2017 Nov 29. PMCID: PMC5800808; PMID: 29187524 PubMed ID: [29187524](#)

Sturtevant, A.H., 1919. A new species closely resembling to *Drosophila melanogaster*. *Psyche* 26: 488–500. FBrf0000977 PubMed ID: [null](#)

**Funding:** This material is based upon work supported by the National Science Foundation (1915544) and the National Institute of General Medical Sciences of the National Institutes of Health (R25GM130517) to the Genomics Education Partnership (GEP; <https://thegep.org/>; PI-LKR). Any opinions, findings, and conclusions or recommendations expressed in this material are solely those of the author(s) and do not necessarily reflect the official views of the National Science Foundation nor the National Institutes of Health.

**Author Contributions:** Leon F. Laskowski: formal analysis, validation, writing - original draft, writing - review editing. Cole A. Kiser: formal analysis, writing - review editing. Robyn Huber: formal analysis, writing - review editing. Sean Kusche: formal analysis, writing - review editing. Andrew M Arsham: supervision, writing - review editing. Thomas C. Giarla: supervision, writing - review editing. Chinmay P. Rele: data curation, formal analysis, methodology, project, software, supervision, validation, visualization, writing - review editing.

**Reviewed By:** GEP Review Panel

**History:** Received July 26, 2021 Accepted November 10, 2022 Published Online November 15, 2022 Indexed November 29, 2022

**Copyright:** © 2022 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Citation:** Laskowski, LF; Kiser, CA; Huber, R; Kusche, S; Arsham, AM; Giarla, TC; Rele, CP (2022). *Drosophila simulans – Ilp2*. microPublication Biology. [10.17912/micropub.biology.000679](https://doi.org/10.17912/micropub.biology.000679)