Wnt Signaling Genes in *Diaphorina citri*

Chad Vosburg[1], Tom D'Elia[1], Teresa Shippy[2], Surya Saha[3]

[1]Indian River State College, Fort Pierce, FL 34981; [2]Division of Biology, Kansas State University, Manhattan, KS 66506; [3]Boyce Thompson Institute, Ithaca, NY 14853

Introduction:

In metazoans, Wnt signaling is involved in many biological processes such as patterning, cell polarity, tissue generation, and stem cell maintenance [17][18]. Wnt signaling falls into two categories: beta-catenin-dependent and beta-catenin-independent. These are also referred to as canonical and noncanonical pathways, respectively. The secretion of Wnt from a cell is facilitated by two proteins, Wntless and Porcupine. Once outside of the cell, Wnt acts as a morphogen by binding to the receptors on neighboring cells. In canonical signaling, the binding of Wnt to Frizzled receptors inhibits the Beta-catenin destruction complex (composed of APC, Axin, CK1, and Shaggy) by recruiting Dishevelled to the membrane. Inhibiting the complex allows Beta-catenin (the mammalian ortholog of Armadillo) to build up in the cytoplasm and subsequently migrate into the nucleus, where it binds to the transcription factor TCF/Pangolin and regulates the transcription of target genes.

Non-canonical pathways do not utilize beta-catenin in their cascade. These pathways can differ in their downstream components and produce different cellular responses and outcomes. Different co-receptors are associated with different functions of Wnt signaling. Low-density lipoprotein (LDL) receptor-related proteins (LRP) serve as co-receptors for Beta-catenin-dependent signaling, whereas co-receptors Doughnut, Derailed, and ROR2 are associated with Beta-catenin-independent signaling. The most studied functions of non-canonical Wnt signal transduction include regulation of the cytoskeleton and intracellular calcium concentrations, but many more functions exist [22].

Many aspects of even the most well studied Wnt pathways remain an enigma due to the complex molecular interactions that direct Wnt signaling into specific cascades. Such direction can be affected by the type of Wnt ligand encountered, the type of receptor, the extracellular and intracellular environments, and the cross reactivity of ligands, receptors, coreceptors, and other modulators [22]. For example, Derailed and Doughnut are two Wnt co-receptors that are typically associated with beta-catenin-independent signaling in *Drosophila*, however, a study by Lu et al. suggests the mammalian homolog of these genes, Ryk, binds directly to Frizzled and Wnt1 to initiate the beta-catenin-dependent cascade in mammalian cells [21]. Similarly, Wnt5a is suspected of eliciting both a canonical and non-canonical response in the presence of Frizzled and ROR2 receptors, respectively [22].

We have curated a complete repertoire of Wnt signaling genes in *D. citri.* In total, 24 genes associated with canonical and noncanonical Wnt signaling have been annotated including seven Wnt ligands, three *frizzled* homologs, *arrow*, and several receptor tyrosine kinases such as *ROR* and *doughnut.* The mechanisms of canonical Wnt signaling appear to be mostly conserved and comparable to that which is found in the model organism, *Drosophila melanogaster*. This is an important first step for understanding critical biological processes that may be targeted to control the spread of *D. citri* and possibly provide a broader insight into the mechanisms of Wnt signaling.

Methods:

The *D. citri* genome was manually annotated by a student-driven community as part of a collaboration between multiple institutions described by Hosmani *et al.* [26]. Protein sequences orthologous to Wnt pathway genes were collected from the NCBI protein database and used for a BLAST search of the *Diaphorina citri* MCOT protein database available on citrusgreening.org to find predicted protein models. NCBI BLAST was used to confirm the accuracy of the predicted protein models. The top scoring MCOT IDs were used to search the *Diaphorina citri* version 2.0 genome, and regions of high sequence identity were

investigated. Gene models were manually annotated in WebApollo, using DNA-Seq, RNA-Seq, and Iso-Seq gene model predictions as evidence to support the annotated model's gene structure. The gene models were analyzed with NCBI BLAST to assess their completeness. MUSCLE multiple sequence alignments of the *D. citri* gene model sequences and orthologous sequences were created through MEGA7 [28]. Neighbor-joining trees were constructed using MEGA7 with p-distance for determining branch length and one thousand bootstrapping replications to measure the precision of branch placement. In special cases, phylogenetic analysis in conjunction with NCBI BLAST scores was used to properly name and characterize the manually annotated gene models.

Results/Discussion

Orthologs of twenty-four genes involved in the canonical and non-canonical Wnt signaling pathways have been manually annotated in the *Diaphorina citri* genome (Table 1). All curated genes have strong support from multiple sources of evidence, including *de novo* transcriptome. In addition, three Wnt ligands that are found in other insects could not be located in the current genome assembly or the available *de novo* transcriptomes and may have been lost in the evolution of *D. citri* (Table 2). However, future revisions of the genome assembly could reveal the presence of these missing orthologs. A model for canonical Wnt signaling in *D. citri* based on curated genes is shown in Figure 1.

		Evidence supporting annotation				
Gene	Identifier	MCOT	*de novo transcriptome*	Iso-Seq	RNA-Seq	Ortholog
Wnt1	DcitrG031975.2 DcitrG031995.2	X	X		X	X
Wnt5	DcitrG051230.1	X	X		X	
Wnt6	DcitrG031970.2	X	X	X	X	
Wnt7†	DcitrG094025.2	X	X		X	X
Wnt10	DcitrG039730.1	X	X		X	X
Wnt11	DcitrG075660.1	X	X		X	
WntA	DcitrG056455.1	X	X		X	X
pangolin†	DcitrG006563.1(2)	X	X		X	
armadillo	DcitrG093465.2	X		X	X	X

wntless	DcitrG053870.2	X	X	X	X	X
porcupine	DcitrG042445.1	X	X	X	X	
derailed	DcitrG053130.2	X	X	X	X	
doughnut	DcitrG054000.1	X	X	X	X	X
arrow	DcitrG022215.2	X	X	X	X	X
frizzled	DcitrG015150.2	X	X		X	
frizzled 2	DcitrG039580.2 DcitrG039545.1	X	X	X	X	
frizzled 3	DcitrG053187.1	X	X	X		
ROR	No identifier yet (2)	X	X	X	X	X
ROR2	DcitrG043210.2	X	X	X	X	X
dishevelled	DcitrG039265.1	X	X		X	X
shaggy	DcitrG082860.2	X	X	X	X	X
Axin	DcitrG054460.2	X	X		X	
ck1-gamma	DcitrG033745.1	X	X	X	X	X
Apc†	DcitrG065410.2	X	X		X	
† Gene is manually annotated as a partial model in Genome v3.0. A complete representation of the gene and protein sequence can be determined from MCOT transcriptome data.						

Table 1: Table of evidence supporting gene annotation. Manually annotated Wnt pathway genes in *Diaphorina citri*. Number of isoforms noted in parentheses if there are more than one. There are 24 genes in total. Each gene has been assigned an identifier, and the evidence used to validate or modify the structure of the gene model has been listed. Table is marked with 'X' when supporting evidence of MCOT, *de novo* transcriptome, Iso-Seq, RNA-Seq and ortholog support is present. MCOT is comprehensive transcriptome based on genome independent and MAKER gene predictions. *De novo* transcriptome is an independent transcriptome using Iso-Seq long-reads and RNA-Seq data. Iso-Seq transcripts are full-length transcripts generated with Pacific Biosciences technology. RNA-Seq reads mapped to genome are also used as supporting evidence for splice junctions. Ortholog evidence is comprised of proteins from related hemipteran species and Drosophila melanogaster.

Gene	*Drosophila melanogaster*	*Apis melifera*	*Tribolium castaneum*	*Acyrthosiphon pisum*	*Diaphorina cit*ri v3
Wnt1	1	1	1	1	1
Wnt5	1	1	1	1	1
Wnt6	1	1	1	0	1
Wnt7	1	1	1	1	1
Wnt8/D	1	0	1	0	0
Wnt9	1	0	1	0	0

Wnt10	1	1	1	0	1
Wnt11	0	1	1	1	1
Wnt16	0	0	0	1	0
WntA	0	1	1	1	1
pangolin	1	1	1	1	1
armadillo	1	1	2	2	1
wntless	1	1	1	1	1
porcupine	1	1	1	1	1
derailed	2	1	0	1	1
doughnut	1	1	1	1	1
arrow	1	1	1	1	1
frizzled	4	2	3	2	3
ROR	2	2	3	2	2
dishevelled	1	1	1	1	1
shaggy	1	1	1	2	1
Axin	1	1	1	1	1
ck1-gamma	1	1	1	1	1
Apc	2	1	1	1	1

Table 2: Gene copy table. Wnt pathway ortholog numbers in five different insect species. *Drosophila melanogaster, Apis mellifera, Tribolium castaneum,* and *Acyrthosiphon pisum* numbers were determined using Flybase, OrthoDB, NCBI Genbank, Uniprot, and several other publications [1][3][4][5][6]. *Diaphorina citri* numbers represent the number of manually annotated genes in the *D. citri* v3.0 genome.

The loss of Wnt ligand genes is more common in insects than in other metazoans which leads to a highly variable array of *Wnts* from species to species [1]. This may be facilitated by the promiscuous nature of Wnt proteins that may allow certain subfamilies to compensate for the loss of others by sharing receptors [12]. This presents some challenges when characterizing *Wnt* genes, and further sequence analysis was performed to describe which *Wnt* genes are present in the *D. citri* genome. Seven different *D. citri Wnt* genes were identified and classified as *Wnt1*, *Wnt5*, *Wnt6*, *Wnt7*, *Wnt10*, *Wnt11*, and *WntA* (table 1). In comparison, seven *Wnt* genes have been identified in *Drosophila melanogaster*, nine in *Tribolium castaneum*, and six in *Acyrthosiphon pisum* [7][6]. The collection of *Wnt* genes in the *D. citri* is largely conserved among insects, and there have been no *Wnt* subfamilies identified that are unique to *D. citri*. Contrary to what has been previously reported, *D. citri* does in fact appear to possess a *Wnt6* gene

[19]. This finding refutes the notion that the loss of maxillary palps in Hemipteran evolution is correlated to the loss of *Wnt6* [19].

Wnt1, *Wnt6*, and *Wnt10* typically occur in very close proximity to one another in a highly conserved gene cluster [7][13][14][16]. Accordingly, it is believed that this cluster is also conserved in *D. citri* and is supported by the chromosomal length assembly in v3.0. The close phylogenetic relationship of *Wnt1*, *Wnt6*, and *Wnt10* in *D. citri* (Fig 3) supports the hypothesis that this cluster is the result of an ancient duplication event that may predate the divergence of cnidarians and bilaterians [16]. The orientation of these clustered *D. citri Wnt* genes is similar to that found in *D. melanogaster* and differs from what may be a basal organization of *Wnts* found in species of Coleoptera, Hymenoptera, and Cladocera (Fig 4). When present, *Wnt9* is also associated with this gene cluster, as seen in *D. melanogaster* and *T. castaneum*. However, as in *Acyrthosiphon pisum*, *Wnt9* was not found in the *D. citri* genome and appears to have been lost. A second *Wnt* cluster, *Wnt5* and *Wnt7*, is also common among non-insect metazoans. This cluster is not seen in *D. citri,* but phylogenetic analysis indicates a close relationship between *Wnt5* and *Wnt7,* suggesting that these genes might also be the result of a duplication event.

The mechanisms that act to conserve these *Wnt* gene clusters are not well understood. Clustered *Wnt* genes do not exhibit similar expression patterns or *Hox*-like collinearity [13]. Data obtained from Psyllid Expression Network (PEN) also shows varying level of expressions amongst the clustered genes (Fig 5). Our annotation findings support the hypothesis that natural selection is preserving *Wnt* clusters. Furthermore, gene orientations within the clusters are subject to rearrangement (Fig 4). This may indicate that directionality or shared transcriptional regulators are not the mechanisms responsible for conserving this cluster, and there may be some other unknown selective pressure at work. A better understanding of the regulative hierarchy that controls *Wnt* expression might shed light on the significance of *Wnt* gene associations in the genome.

The organization of the genomic reference sequence into chromosomal length scaffolds was essential to revealing this *D. citri* gene clustering. The previous genome assemblies were often unsupportive in confirming the proximity of genes due to the shorter scaffold lengths, and genome v2.0 assembly errors had likely misrepresented the location of *Wnt10,* making it appear to be separated from *Wnt1* and *Wnt6.* The presence of this expected 1-6-10 cluster in the v3.0 assembly suggests an improvement in the quality of the v3.0 assembly over previous versions. This also exemplifies how the quality of the reference genome should be considered when performing phylogenetic studies.

Orthologs for *Wnt2*, *Wnt3*, *Wnt4*, *Wnt9*, *Wnt8/D*, and *Wnt16* were unable to be located the *D. citri* genome. The close identity of certain *Wnt* subfamilies makes distinguishing between them difficult, however, the loss of *Wnt2–4* is expected as they are absent in all insects [1]. *Apis mellifera* and the hemipteran *Acyrthosiphon pisum* have been reported to lack *Wnt8/D* which suggests that perhaps this *Wnt* subfamily has been lost in the divergence from other insect groups [6][11]. Additionally, *Wnt16* was not found *D. citri* v3.0*.* This finding contrasts with the gene predictions of other hemipteran genomes, namely *A. pisum, Sipha flava,* and *Nilaparvata lugens,* which appear to have *Wnt16* orthologs (Fig 3).

Several receptors and co-receptions that are associated canonical and non-canonical signaling have been identified (Table 1). Three homologs for the Wnt receptor frizzled have been found – one of which is closely related to frizzled 3 in *D. melanogaster* (Fig 6). A homolog for this third frizzled has not been reported to be in the *A. pisum* genome [6], but it is predicted to be in other hemipterans such as *Halymorpha halys* and *N. lugens* (Fig 3). A homolog for the frizzled co-receptor Arrow (LRP5/6 in mammals) has also been identified, and it shares 3 of the 5 conserved PPPSP amino acid sequence motifs found in *D. melanogaster* (Fig 7)[27]. These intracellular PPPSP motifs are recognized and phosphorylated by the Beta-catenin destruction complex in the presence of Wnt, thus allowing canonical signaling to continue. Both *ROR* and *ROR2* have been identified. Interestingly, *ROR* has two isoforms; one containing an immunoglobulin (IG) domain and the other lacking it. RNA-Seq reads suggest the isoform containing the

IG domain is much more expressed in the nymph and egg stages than in the adult and may be involved in the early developmental stages of *D. citri*.

It appears that the mechanisms for Wnt signaling in *D. citri* remain heavily conserved as all the major downstream components have been found in the genome with some notable variations in gene copy numbers. This is an important first step for understanding critical biological processes that may be targeted to control the spread of *D. citri* and possibly provide a broader insight into the mechanisms of Wnt signaling.

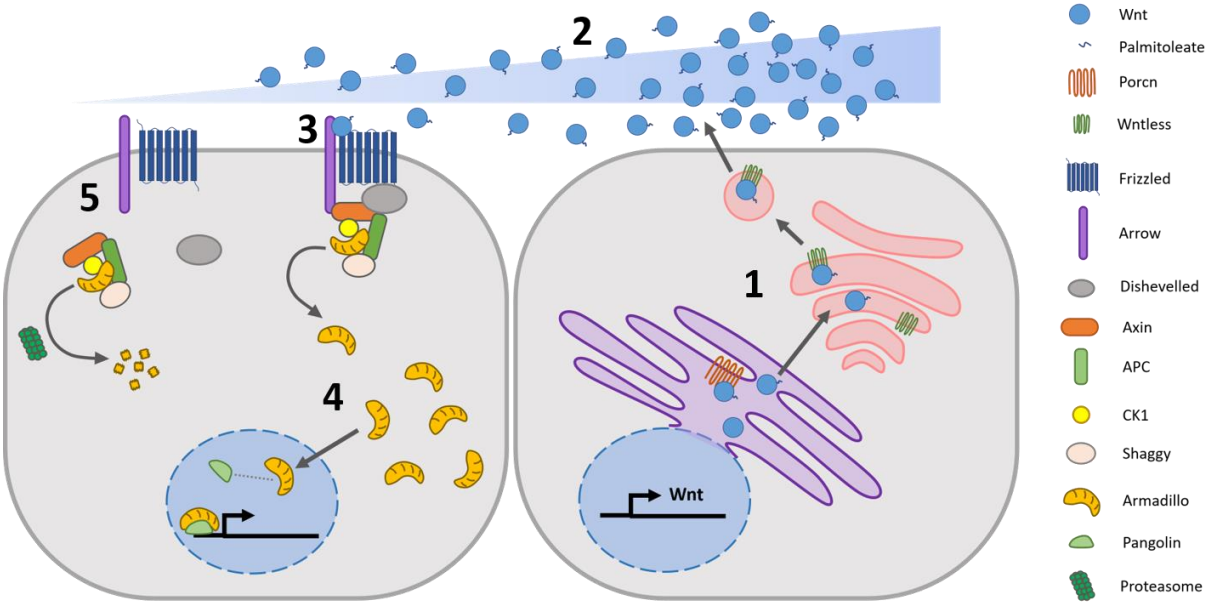

Figure 1: Canonical Wnt signaling cascade in D. citri based on curated genes. 1) Wnt is secreted; 2) Concentration gradient forms; 3) Wnt binds and releases armadillo; 4) Arm migrates into the nucleus and regulates transcription; 5) Arm is degraded in the absence of Wnt.

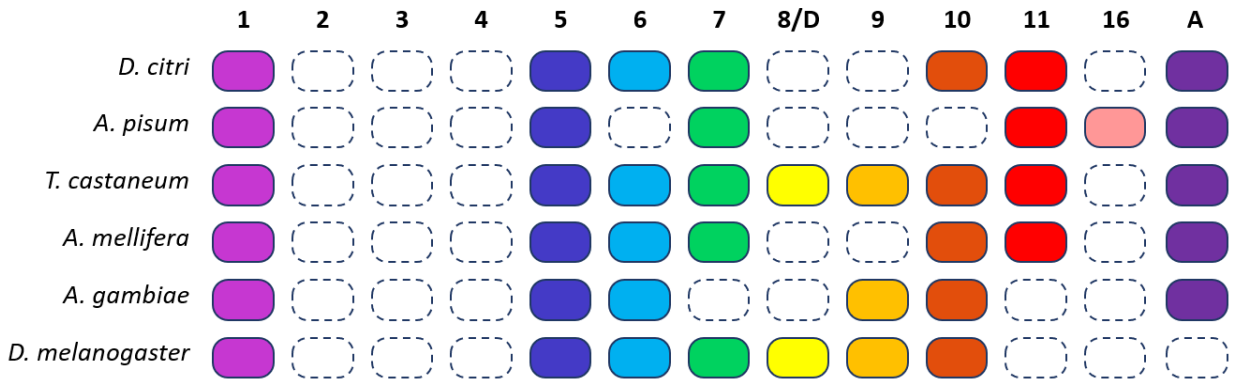

Figure 2: *Wnt* genes in six insects. A colored box indicates the presence of a *Wnt* subfamily (1 to 11, 16, and A) in that insect, while a white box indicates the loss of a subfamily. For example, all six species have *Wnt1* and *Wnt5*, none have *Wnt2-4*, and only *A. pisum* has *Wnt16*. Homologs of *Wnt8* in *T. castaneum* and *D. melanogaster* are also referred to as *WntD*.

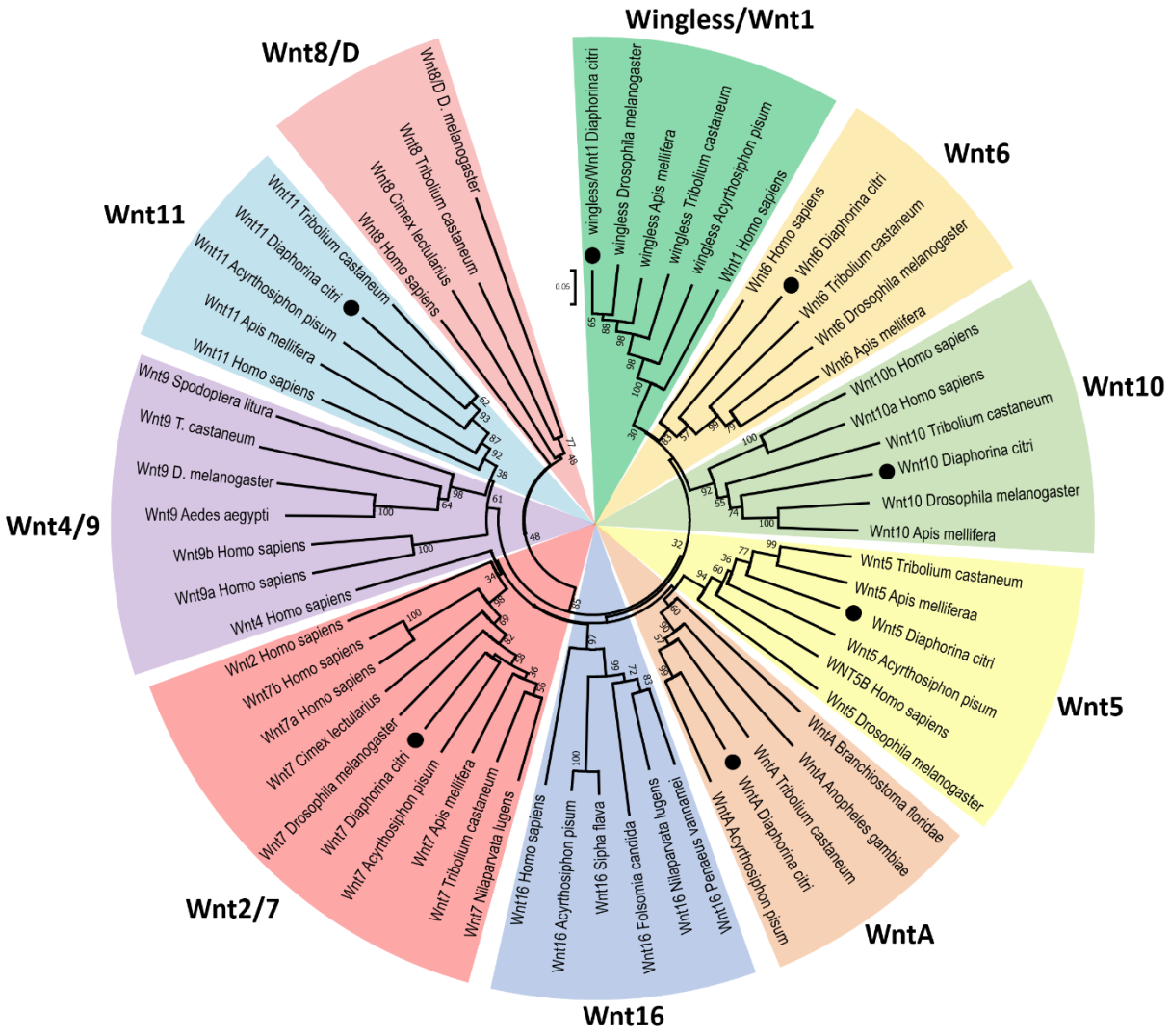

Figure 3: Neighbor-joining tree of Wnt protein sequences. Phylogenetic analysis was performed to categorize the seven *D. citri Wnt* genes (signified by dots). Wnt families are distinguished by clades and are color coded. Bootstrap values under 25 were removed. Ortholog sequences collected from NCBI protein database. Analysis performed by MEGA7.

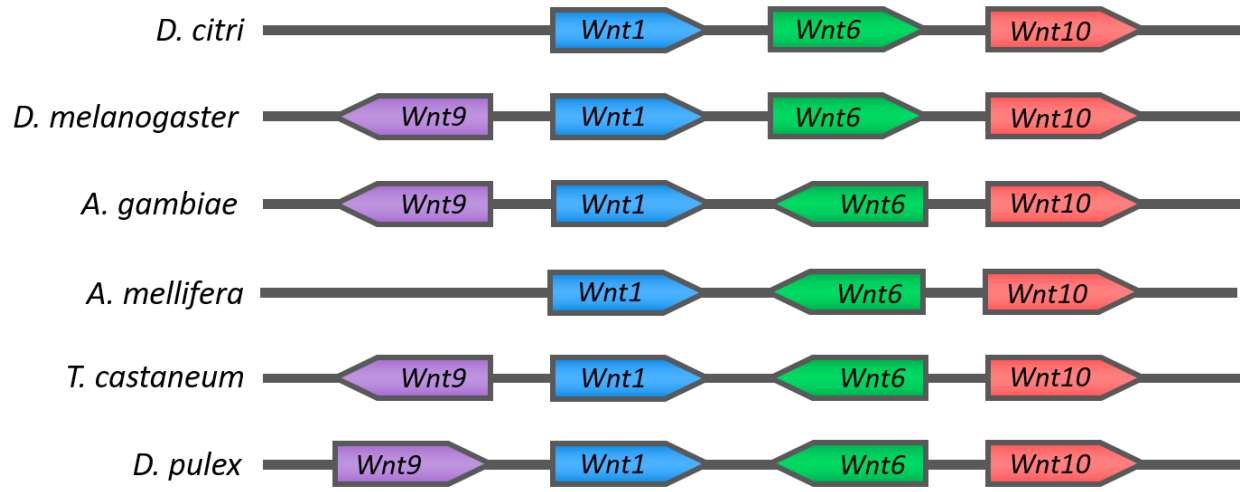

Figure 4: *Wnt1-6-10* Cluster comparison. Gene lengths are not to scale.

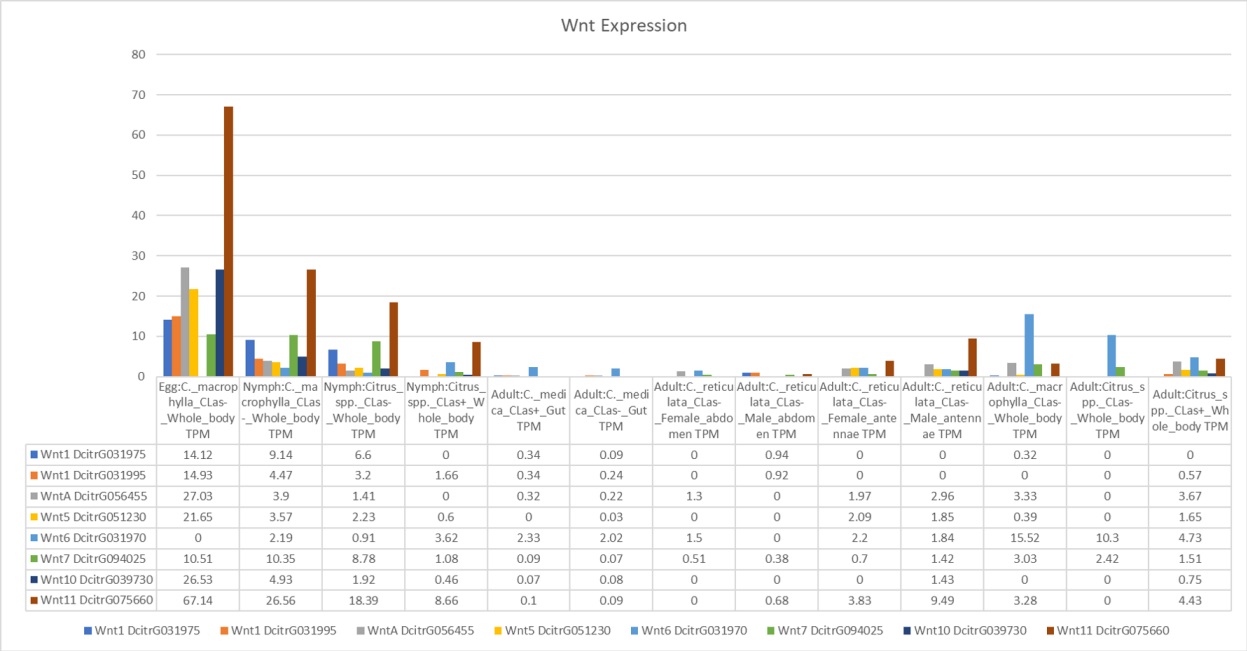

Figure 5: Wnt ligand expression from Psyllid Expression Network data available on citrusgreening.org. Expression values shown in transcripts per million (TPM).

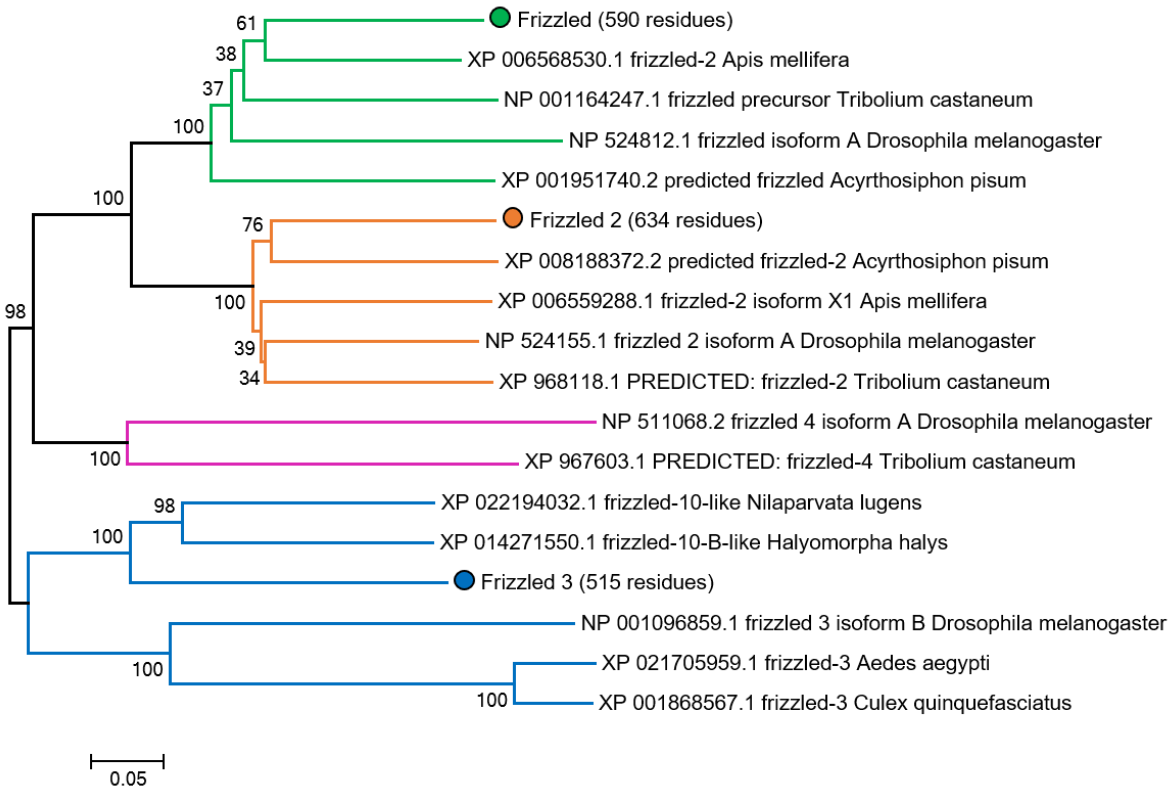

Figure 6: Neighbor-joining tree of frizzled protein sequences. Frizzled subfamilies (1-4) are colored and circles indicate the *D. citri* homologs.

```
Drosophila      1509 SVADAVRMSTLNSRNSMNSYDRNHITGASSSTTNGSSMVAYPIPPPSPA   1558
                     .|....|:::| ..:|.:||..      .|||    |::.||:|||||| 
d.citri         1434 HVIPPPRLTSL-PPSSSSSYQH------QSST----SLLCYPLPPPSPA   1472

Drosophila      1559 TRSRRP------YRHYKIINQPPPPTPCSTDICDESDSNYTSKSNSNNSN   1602
                     |..:|.      |       |||||||||:|||||.||
d.citri         1473 TTVQRDEYCCPRY-------QPPPPTPCSTDVCDESDCNY----------   1505

Drosophila      1603 GGATKHSSSSAAACLQYGYDSEPY-PPPPTPRSHYHSDVRIVPESSCPPS   1651
                                          :|| ||||||.|  ||::         :
d.citri         1506 --------------------DPYPPPPPTPTS--HSEL----------N   1522

Drosophila      1652 PSSRSSTYFSPLPPPPSPVQSPSRGFT   1678
                     .||.:|:|.|...
d.citri         1523 SSSSNSSYSSYRT-------------   1535
```

Figure 7: Conserved PPPSP motifs. Pairwise alignment of arrow (LRP5/6) intracellular domain sequences in *Drosophila* and *D. citri* sequences. Conserved PPPSP motifs are outlined in red.

References

1. Janssen R, Le Gouar M, Pechmann M, Poulin F, Bolognesi R, Schwager EE, Hopfen C, Colbourne JK, Budd GE, Brown SJ, et al. 2010. Conservation, loss, and redeployment of Wnt ligands in protostomes: Implications for understanding the evolution of segment formation. BMC Evol Biol. 10(1):1–21. doi:10.1186/1471-2148-10-374.

2. https://www.ncbi.nlm.nih.gov/books/NBK6212/

3. Bao R, Fischer T, Bolognesi R, Brown SJ, Friedrich M. 2012. Parallel duplication and partial subfunctionalization of β-Catenin/armadillo during insect evolution. Mol Biol Evol. 29(2):647–662. doi:10.1093/molbev/msr219.

4. Huang HC, Klein PS. 2004. The frizzled family: Receptor for multiple signal transduction pathways. Genome Biol. 5(7):1–7. doi:10.1186/gb-2004-5-7-234.

5. A context-dependent combination of Wnt receptors controls axis elongation and leg development in a short germ insect https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3188592/

6. Shigenobu S, Bickel RD, Brisson JA, Butts T, Chang CC, Christiaens O, Davis GK, Duncan EJ, Ferrier DEK, Iga M, et al. 2010. Comprehensive survey of developmental genes in the pea aphid, Acyrthosiphon pisum: Frequent lineage-specific duplications and losses of developmental genes. Insect Mol Biol. 19(SUPPL. 2):47–62. doi:10.1111/j.1365-2583.2009.00944.x.

7. Manley G. 2013. Public Access NIH Public Access. 2008:193–202. doi:10.1038/mp.2011.182.doi.

8. Somorjai IML, Martí-Solans J, Diaz-Gracia M, Nishida H, Imai KS, Escrivà H, Cañestro C, Albalat R. 2018. Wnt evolution and function shuffling in liberal and conservative chordate genomes. Genome Biol. 19(1):1–17. doi:10.1186/s13059-018-1468-3.

9. He X. 2004. LDL receptor-related proteins 5 and 6 in Wnt/ -catenin signaling: Arrows point the way. Development. 131(8):1663–1677. doi:10.1242/dev.01117.

10. Price MA, Price MA. 2006. CKI , there ' s more than one : casein kinase I family members in Wnt and Hedgehog signaling. :399–410. doi:10.1101/gad.1394306.

11. Patterns of conservation and change in honey bee developmental genes

12. Rapp J, Jaromi L, Kvell K, Miskei G, Pongracz JE. 2017. WNT signaling - lung cancer is no exception. Respir Res. 18(1):1–16. doi:10.1186/s12931-017-0650-6.

13. Sullivan JC, Ryan JF, Mullikin JC, Finnerty JR. 2007. Conserved and novel Wnt clusters in the basal eumetazoan Nematostella vectensis. Dev Genes Evol. 217(3):235–239. doi:10.1007/s00427-007-0136-5.

14. Cho SJ, Vallès Y, Giani VC, Seaver EC, Weisblat DA. 2010. Evolutionary dynamics of the wnt gene family: A lophotrochozoan perspective. Mol Biol Evol. 27(7):1645–1658. doi:10.1093/molbev/msq052.

15. Prühs R, Beermann A, Schröder R. 2017. The Roles of the Wnt-Antagonists Axin and Lrp4 during Embryogenesis of the Red Flour Beetle Tribolium castaneum. J Dev Biol. 5(4):10. doi:10.3390/jdb5040010.

16. An ancient cluster of Wnt paralogues

17. Holstein TW, Zimmerman ZF, Moon RT, Chien J, Whyte JL, Smith A a, Jill a. 2014. The Evolution of the Wnt Pathway:1–18. doi:10.1101/cshperspect.a007922.

18. Nusse R, Fuerer C, Ching W, Harnish K, Logan C, Zeng A, Ten Berge D, Kalani Y. 2008. Wnt signaling and stem cell control. Cold Spring Harb Symp Quant Biol. 73:59–66. doi:10.1101/sqb.2008.73.035.

19. Doumpas N, Jékely G, Teleman AA. 2013. Wnt6 is required for maxillary palp formation in Drosophila. BMC Biol. 11. doi:10.1186/1741-7007-11-104.

20. Flores M, Saha S, Hosmani PS, Brown S, A Mueller L. Community portal for the Citrusgreening disease resources [Internet]. figshare; 2017. doi:10.6084/m9.figshare.5307391.v1

21. Lu W, Yamamoto V, Ortega B, Baltimore D. 2004. Mammalian Ryk is a Wnt coreceptor required for stimulation of neurite outgrowth. Cell. 119(1):97–108. doi:10.1016/j.cell.2004.09.019.

22. De A. 2011. Wnt/Ca 2 signaling pathway: A brief overview. Acta Biochim Biophys Sin (Shanghai). 43(10):745–756. doi:10.1093/abbs/gmr079.

23. Zhai L, Chaturvedi D, Cumberledge S. 2004. Drosophila Wnt-1 undergoes a hydrophobic modification and is targeted to lipid rafts, a process that requires porcupine. J Biol Chem. 279(32):33220–33227. doi:10.1074/jbc.M403407200.

24. Galli LM, Barnes TL, Secrest SS, Kadowaki T, Burrus LW. 2007. Porcupine-mediated lipid-modification regulates the activity and distribution of Wnt proteins in the chick neural tube. Development. 134(18):3339–3348. doi:10.1242/dev.02881.

25. Bänziger C, Soldini D, Schütt C, Zipperlen P, Hausmann G, Basler K. 2006. Wntless, a Conserved Membrane Protein Dedicated to the Secretion of Wnt Proteins from Signaling Cells. Cell. 125(3):509–522. doi:10.1016/j.cell.2006.02.049.

26. Hosmani PS, Shippy T, Miller S, Benoit JB, Munoz-Torres M, Flores-Gonzalez M, et al. 2019. A quick guide for student-driven community genome annotation. PLoS Comput Biol 15(4):e1006682. https://doi.org/10.1371/journal.pcbi.1006682

27. Tamai K, Zeng X, Liu C, Zhang X, Harada Y, Chang Z, and He X. A Mechanism for Wnt Coreceptor Activation. Mol Cell. 13:149–156.

28. Sudhir Kumar, Glen Stecher, Koichiro Tamura, MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets, Molecular Biology and Evolution, Volume 33, Issue 7, July 2016, Pages 1870–1874, https://doi.org/10.1093/molbev/msw054