Phototransduction in *Diaphorina citri*

Jordan Norus, Chad Vosburg, Thomson Paris, Tom D'Elia, Helen Wiersma-Koch, Teresa Shippy, Prashant Hosmani, Surya Saha and Sue Brown

Introduction

The phototransduction pathway is responsible for transducing energy from light into electrical signals attributed to vision. Accordingly, the pathway acts as a cascade involving several different genes (Scott and Zuker 1998). Insects rely heavily on information generated by visual stimulus, thus the phototransduction pathway contributes to their overall survival and fitness by providing visual cues and spatial detail (Howard et al. 1984). The process of phototransduction is initiated by photonic energy that comes from light. Light is absorbed through rhodopsin, a seven helix, G-protein-coupled-receptor (Lamb and Pugh 2006). Once light has been absorbed, the chromophore ligand of rhodopsin is isomerized to an all-trans configuration (Lamb and Pugh 2006). The subsequent absorption of light activates the rhodopsin molecule and initiates the signaling cascade of phototransduction (Lamb and Pugh 2006).

Photoactivated rhodopsin stimulates a GDP/GTP exchange in the heterotrimeric G-protein which activates the gene PLC encoded by NorpA (Montell 2012). Stimulation of PLC leads to the production of inositol 1,4,5,-tris-phosphate receptor and diacylglycerol. Sn1-specific diacylglycerol lipase beta hydrolyzes the diacylglycerol forming monoacylglycerol and saturated fatty acid products (Montell 2012). This leads to the opening of the transient receptor potential protein channel and transient-receptor-potential-like protein cation channel in the photoreceptor cells (Montell 2012). Following the activation of the channels, Ca^{2+} is extruded from the photoreceptor cell causing depolarization to occur (Montell 2012). The signal cascade stops upon the inactivation of rhodopsin by arrestin 2 (Dolph et al. 1993). The processes described are utilized for the signal propagation that is responsible for vision. Following the inactivation of rhodopsin, signals will be regenerated given visual stimulus.

A total of 18 genes in the phototransduction pathway have been found and annotated in *D. citri.* An understanding of the pathway in *D. citri*, vector of the citrus greening disease, is essential to understanding its role in transmission. Manual annotation of the phototransduction pathway in *D. citri* is vital to ensure the quality and accuracy of gene models. Annotated gene models provide potential targets for RNA interference gene silencing

experiments and other molecular therapeutics. Annotation of the *opsin* genes and *arrestin 2* required additional measures to validate models, however the rest of the pathway was found and annotated as expected.

Materials and Methods

The *D. citri* genome was annotated as part of an annotation community described previously (Hosmani et al 2019) and was used as a guideline for annotation of the phototransduction genes. The NCBI protein database was used to collect orthologous protein sequences. A BLAST search of the orthologous sequences was made to the *D. citri* MCOT protein database to find predicted protein models (Flores et al. 2017). Confirmation of the predicted MCOT sequences was given by NCBI BLAST results. A BLAT search of the above MCOT sequences was conducted on the *D.citri* genome, and the most promising results were investigated. Predicted gene models were annotated manually using WebApollo. Illumina DNAseq, RNAseq, and Stringtie gene predictions as lines of evidence to support gene structure. Analysis of gene models and pairwise alignment tables were carried out using NCBI BLAST. MEGA7 software was used to conduct ClustalW multiple sequence alignments of gene models and orthologous sequences, along with formation of neighbor-joining trees. P-distance determined branch lengths and one thousand bootstrapping replications were used to measure the precision of branch placements. Phylogenetic trees were constructed using MEGA7 software. Protein sequences were obtained from the NCBI protein database. ClustalW multiple sequence alignments was done using the annotated gene models and orthologous sequences. Neighbor-joining trees were generated from the multiple sequence alignments. P-distance determined branch lengths and one thousand bootstrapping replications were used to measure the precision of branch placements.

Results/Discussion

Gene Copy Numbers

A total of 18 genes were found within the *D. citri* genome that were determined to be involved in the phototransduction pathway. Evidence for each manually annotated gene is provided in Table 1. MCOT, Isoseq, and

orthologous sequences, along with RNAseq data were all used to determine the accuracy of gene models. Within the phototransduction pathway, *D. citri* contained four copies of *opsin* (Table 2), a gene responsible for the ability of an organism to detect light (Porter et al. 2011). This is similar to the four copies of *opsin* found in *Apis mellifera,* however, *opsin* copies differ in *Drosophila melanogaster* (Terakita 2005). *D. melanogaster* contains seven copies of *opsin*, all of which are denoted as *rhodopsin* (Şahin and Çelik 2013). Notably, *beta-adrenergic-receptor kinase*, a gene responsible for phosphorylating light-activated rhodopsin in *D. melanogaster* (Koch et al. 1992) could not be found within the *D. citri* genome. Likewise, it is not found in other hemipterans similar to *D. citri* such as *Acyrthosiphon pisum*.

Opsin

Opsin proteins are a diverse group of G-Protein coupled receptors containing seven transmembrane alpha helices (Scheerer et al. 2008). They are photoreceptive due to the chromophore retinal, a derivative of vitamin A, which is bound to a conserved lysine in the seventh transmembrane helix (Collantes-Alegre et al. 2018). Opsins are involved in visual phototransduction providing the photoperiodism system with light and dark cycle information (Senthilan and Helfrich-Förster 2016). Opsin is the ligand-free form of rhodopsin, a G-protein-coupled receptor (Scheerer et al. 2008). In *D. citri*, a total of four *opsin* genes were identified (Table 2). One gene copy of *UV-sensitive opsin, short-wavelength opsin, long-wavelength opsin*, and *rhodopsin 7* were annotated. Phylogenetic analysis of the opsin sequences was performed to confirm the predicted classification of each of the gene models annotated (Figure 1). Rhodopsin 7 of *D. citri* groups closely with rhodopsin 7 of *D. melanogaster* along with sequences of other UV-sensitive opsins. Further support that the *D. citri* sequence for rhodopsin 7 is accurately classified is the absence of a G-protein -activating QAKK motif which is an important motif in the designation of rhodopsin 7. The QAKK motif is found in opsin proteins, however it is lacking in rhodopsin 7 (Grebler et al. 2017). Additionally, UV-sensitive opsin sequences in *Pieris rapae* and *Bombyx mori* lack the QAKK domain providing support that they are rhodopsin 7 sequences.

Arrestin 2

Arrestin 2 is a gene involved in the phototransduction pathway. It plays a vital role in mediating the inactivation of rhodopsin which is required for bringing the phototransduction cascade to a stop (Dolph et al. 1993). Without the function of arrestin 2, photoreceptors would remain in a continuously active state noted as prolonged depolarized

afterpotential (Dolph et al. 993). During the process of annotation, an OGS predicted model could not be found, however an arrestin homolog model was manually annotated. A tree was generated using arrestin 1 and arrestin 2 sequences from an array of organisms. The results show that the arrestin homolog in D. citri most closely resembles arrestin 2, thus providing evidence that the manually annotated model is arrestin 2.

Table 1: Manually annotated phototransduction genes in *Diaphorina citri*. There are 18 genes total. Each gene has been assigned an identifier and denoted as complete or partial based on available evidence. MCOT evidence means a de novo Oases or Trinity model from an independent transcriptome was identified and the sequence from that transcript was used to validate or modify our model. IsoSeq means single, long RNAseq reads generated with Pacific Biosciences technology were available and were used to help validate the exon structure of the model. RNASeq means that individually mapped Illumina RNASeq reads were used to help validate or modify our model. Ortholog means ortholog sequences from other insects and information about conserved motifs and domains had to be used to help determine the final annotation.

Gene	*D. citri Identifier*	*Complete*	*Partial*	*MCOT*	*IsoSeq*	*RNASeq*	*Ortholog*
Myosin IIIA	DcitrG100665.2.1	Yes	No	X		X	X
Actin	DcitrG049900.2.1	Yes	No	X	X	X	
Calmodulin	DcitrG010570.1.1	Yes	No		X	X	
Inactivation-no-after-potential D protein	DcitrG070755.1.1	Yes	No	X		X	
Transient receptor potential protein	DcitrG052775.1.1	Yes	No			X	X
Sn1-specific diacylglycerol lipase beta	DcitrG065620.1.1	Yes	No	X		X	

Calcium/calmodulin-dependent protein kinase II alpha chain	DcitrG057015.2.1	Yes	No	X	X	X	
NorpA	DcitrG005760.2.1	Yes	No	X	X	X	
PLC	DcitrG094315.1.1	Yes	No	X	X	X	X
Jetlag	DcitrG053550.2.1	Yes	No	X	X	X	X
Inositol 1,4,5,-tris-phosphate receptor	DcitrG058685.1.1	Yes	No			X	X
Arrestin 2	DcitrG041805.2.1	Yes	No	X		X	X
Protein kinase C	DcitrG085885.1.1	Yes	No	X	X	X	X
Transient-receptor-potential-like protein	DcitrG027370.2.1	Yes	No			X	X
Long-Wavelength Opsin	DcitrG099955.1.1	Yes	No	X	X	X	X
UV Opsin	DcitrG070285.2.1	Yes	No	X		X	X
Short-Wavelength Opsin	DcitrG085770.2.1	Yes	No	X	X	X	X
Rh7	DcitrG018785.2.1	Yes	No	X	X	X	X
Guanine nucleotide-binding protein G(q) subunit alpha	DcitrG032220.1.1	Yes	No	X	X	X	X
Rhodopsin Kinase	DcitrG026410.2.1	Yes	No	X	X	X	X

Table 2. Phototransduction pathway ortholog numbers in five different insect species. *Drosophila melanogaster, Apis mellifera, Tribolium castaneum, and Acyrthosiphon pisum* numbers were determined using NCBI Genbank, UniProt, OrthoDB, and several publications [1][2][3]. *Diaphorina citri* numbers represent the number of manually annotated genes in the D. citri v2.0 genome.

Gene	*Drosophila melanogaster*	*Apis mellifera*	*Tribolium castaneum*	*Acyrthosiphon pisum*	*Diaphorina citri*
Myosin IIIA	1 *†	1*	N/A	N/A	1
Calmodulin	1*†	2*	2*	1*†	1
Inactivation-no-after-potential D protein	1*†	N/A	1*†	1*†	1
Transient receptor potential protein	1*	1*	1*	1*	1
Sn1-specific diacylglycerol lipase	1*	1*	1*†	1*	1
Calcium/calmodulin-dependent protein kinase II alpha chain	1†	1*	1*	N/A	1
NorpA	3†	N/A	N/A	N/A	1
PLC	2†	1*	7†	5†	1
Jetlag	1*†	N/A	N/A	N/A	1
Inositol 1,4,5,-tris-phosphate receptor	1 *†	1*‡	N/A	N/A	1
Arrestin 2	1*†‡	N/A	1*†‡	N/A	1
Protein kinase C	1*†	1*†	1†	1*	1
Transient-receptor-potential-like protein	1*‡	1*	1*	3*	1

Guanine nucleotide-binding protein G(q) subunit alpha	2‡	1*	1*	1*	1
Rhodopsin Kinase	1*	0*	0*	0*	1
Actin	5‡	7‡	2‡	4‡	3
Opsin	7_1	4*	2_2	5_3	4

*NCBI †UniProt ‡Orthodb (Terakita 2005) $_1$ (Richards et al. 2008) $_2$ (Li et al. 2013) $_3$

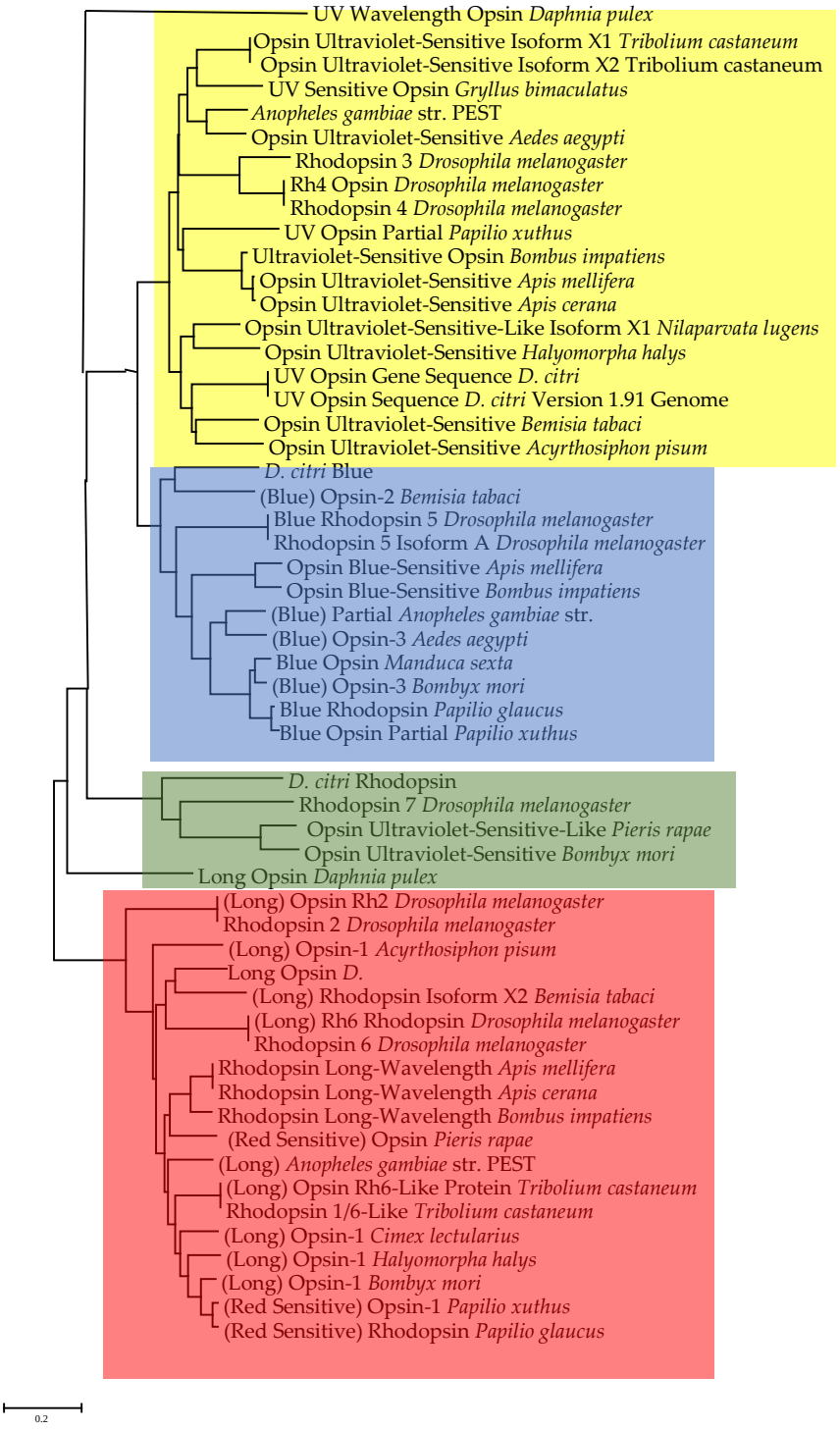

Figure 1. Neighbor-joining tree of opsin homologs. Full length proteins sequences were used in phylogenetic analysis. Colors differentiate between the opsins. Yellow: UV-Sensitive Opsin. Blue: Short-Wavelength Opsin. Red: Long-Wavelength Opsin. Green: Rhodopsin 7.

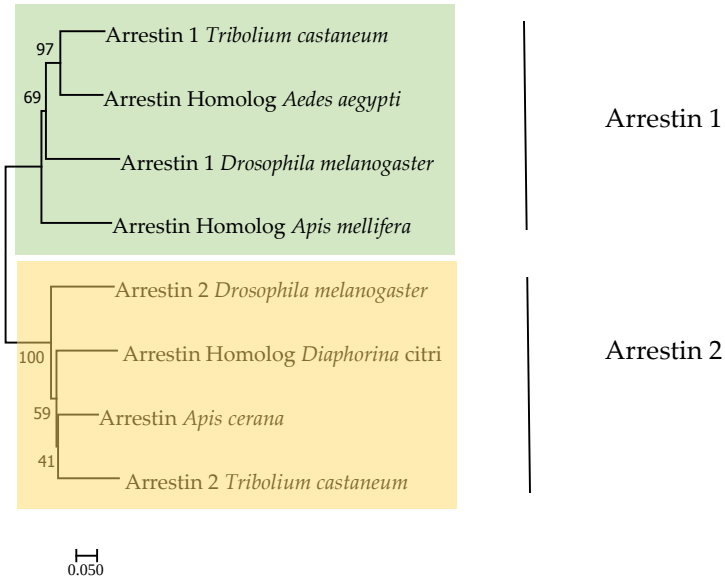

Figure 2. Neighbor-joining tree of opsin homologs. Full length proteins sequences were used in phylogenetic analysis. Colors distinguish between the two arrestins. Green: Arrestin 1. Yellow: Arrestin 2.

Table 3: Accession numbers for genes used in phylogenetic analysis (Figure 1; Figure 2). XP identifiers are computer predicted models.

Organism	UV-Sensitive Opsin	Short-Wavelength Opsin	Long-Wavelength Opsin	Rhodopsin 7	Arrestin 1	Arrestin 2	Arrestin Homolog
Drosophila melanogaster	NP 524411.1 AAA28856.1	AAC47426.1 NP 477096.1	AAA28734.1 NP 524398.1 CAB06821.1 NP 524368.4	NP 524035.2	NP_476681.1	NP_523976.1	
Tribolium castaneum	XP 970344.2 XP 015833491.1		EFA03667.1 NP 001155991.1		EFA04632.1	EFA07498.1	
Bombyx Mori	XP 012545028.1	XP 004932925.1	XP 021205252.1				
Apis mellifera	NP 001011605.1	NP 001011606.1	NP 001011639.2				XP_016772051.1
Acyrthosiphon pisum	XP 001951613.2		XP 008185403.1				
Aedes aegypti	XP 021698798.1	XP 001662982.2					XP_001663732.1
Apis cerana	NP 001315411.1		PBC29422.1				AEY59100.1
Anopheles gambiae	XP 001688790.1	XP 319247.1	XP 003435763.1				
Bombus impatiens	NP 001267052.1	XP 003494923.1	XP 012237087.1				
Bemisia tabaci	XP 018909731.1	XP 018913779.1	XP 018901683.1				
Papilio glaucus		AAD34223.1	AAD34221.1				
Pieris rapae	XP 022114135.1		BAD36899.1				
Daphnia pulex	EFX81332.1		EFX66668.1				
Papilio xuthus	BAA93470.1	BAA93469.1	NP 001299732.1				
Gryllus bimaculatus	BAR40295.1						
Manduca sexta		AAD11966.1					
Cimex lectularius			XP 014240082.1				
Nilaparvata lugens	XP 022201946.1						
Halyomorpha halys	XP 014273911.1		XP 014274077.1				

References

Dolph PJ, Ranganathan R, Colley NJ, Hardy RW, Socolich M, Zuker CS. 1993. Arrestin function in inactivation of G protein-coupled receptor rhodopsin in vivo. Science (80-). doi:10.1126/science.8316831.

Grebler R, Kistenpfennig C, Rieger D, Bentrop J, Schneuwly S, Senthilan PR, Helfrich-Förster C. 2017. Drosophila Rhodopsin 7 can partially replace the structural role of Rhodopsin 1, but not its physiological function. J Comp Physiol A Neuroethol Sensory, Neural, Behav Physiol.

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

Howard J, Dubs A, Payne R. 1984. The dynamics of phototransduction in insects - A comparative study. J Comp Physiol A. doi:10.1007/BF01350224.doi:10.1007/s00359-017-1182-8.

Lamb TD, Pugh EN. 2006. Phototransduction, dark adaptation, and rhodopsin regeneration the proctor lecture. Invest Ophthalmol Vis Sci. doi:10.1167/iovs.06-0187.

Montell C. 2012. Drosophila visual transduction. Trends Neurosci. doi:10.1016/j.tins.2012.03.004.

Pitcher JA, Inglese J, Higgins JB, Arriza JL, Casey PJ, Kim C, Benovic JL, Kwatra MM, Caron MG, Lefkowitz RJ. 1992. Role of βγ subunits of G proteins in targeting the β-adrenergic receptor kinase to membrane-bound receptors. Science (80-). doi:10.1126/science.1325672.

Porter ML, Blasic JR, Bok MJ, Cameron EG, Pringle T, Cronin TW, Robinson PR. 2011. Shedding new light on opsin evolution. Proc R Soc B Biol Sci. doi:10.1098/rspb.2011.1819.

Şahin HB, Çelik A. 2013. Drosophila Eye Development and Photoreceptor Specification. In: eLS.

Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauß N, Choe HW, Hofmann KP, Ernst OP. 2008. Crystal structure of opsin in its G-protein-interacting conformation. Nature. doi:10.1038/nature07330.

Scott K, Zuker CS. 1998. Assembly of the Drosophila phototransduction cascade into a signalling complex shapes elementary responses. Nature. doi:10.1038/27448.

Senthilan PR, Helfrich-Förster C. 2016. Rhodopsin 7–The unusual Rhodopsin in Drosophila. PeerJ. doi:10.7717/peerj.2427.

Terakita A. 2005. The opsins. Genome Biol. doi:10.1186/gb-2005-6-3-213.