In Silico Characterization of Retinoic Acid Pathway Proteins Using *Gryllus bimaculatus* Transcriptome

Student
Researcher: Alicia G. Edwards Advisor: Hadley Horch
Bowdoin College
Neuroscience

Abstract

Neural plasticity, or the ability of neurons in the central nervous system (CNS) to adapt in response to environmental changes, is common during developmental nervous systems but seen significantly less in adults. The auditory system of the cricket *Gryllus bimaculatus* is known to have an unusual level of neural plasticity (Horch et al. 2017). The amputation of a cricket's leg (deafferentation) removes the ear and severs the auditory afferents, or nerve fibers. Upon deafferentation, there is extensive branching of dendrites and axons across the midline of the CNS, a normally respected barrier to neuronal growth. This reorganization can help restore essential behavioral responses in the cricket (Horch et al. 2009). In this study, key proteins involved in developmental processes are explored to provide a deeper understanding of neural plasticity, as well as regenerative biological processes. Specifically, the retinoic acid (RA) pathway proteins were observed and characterized due to their essential role in sustaining the life of many different organisms. Using accessions from *Drosophila melanogaster* and other invertebrates, we were able to perform BLAST procedures in coordination with a previously mined *Gryllus bimaculatus* transcriptome to reveal genetic similarities. This *de novo* cricket transcriptome was mined for select proteins in the RA pathway: alcohol dehydrogenases (Adhs), retinol/retinal dehydrogenase (Rdhs), cytochrome p450s (CYP450s), and 17-beta hydroxysteroid dehydrogenase (17BHD). By generating phylogenetic trees of these particular protein families, we were able to visualize connections and characterizations, which provides a solid foundation for future investigations on the RA pathway and its potential involvement in the neuroplasticity of the cricket auditory system.

Project Objectives

This study explores and characterizes protein members of the RA pathway to provide insight into their potential role in the neuroplasticity of the cricket *Gryllus bimaculatus*. We aimed to perform bioinformatic analyses and produce phylogenetic trees that characterize four protein families—Adhs, Rdh, CYP450s, and 17BHD. These findings will provide the Horch lab future opportunities for investigation on retinoids acid pathway proteins involved in the neuroplasticity of the cricket CNS.

Methodology Used

Initially, a literature search was conducted to select proteins involved in the RA pathway. Using a cricket, *Gryllus bimaculatus,* transcriptome of the prothoracic ganglion assembled previously in the Horch lab (Wang 2020), additional proteins with both an observed regulation after deafferentation and connection to the RA pathway were included in this study. For every protein selected, *Drosophila melanogaster* accessions and FASTA sequences were obtained from the NCBI Protein database and imported into *Geneious*, a bioinformatics software program that can be used for alignment, assembly, and analysis of proteins. These sequences were then BLASTed, an algorithm used for comparing proteins, against the *de novo* cricket transcriptome produced previously in the Horch lab. The produced TRINITY sequences were then translated and edited to find the correct frame as well as remove stop codons. Reciprocal blasts through the NCBI database were performed to confirm the identity and reading frame of each cricket TRINITY sequence. Duplicate sequences were identified and removed using phylogenetic trees and pairwise alignments. Sequences found to be 95% or more identical with another sequence were deemed duplicates and removed from the study. An NCBI domain search was used to do a final confirmation that the identification of each trinity sequence matched its predicted

protein. In *Geneious*, "Identity" cost matrix phylogenetic trees were created for each of the four query protein families—Adhs, Rdh, CYP450s, and 17BHD—with the initial *Drosophila melanogaster* accessions and the blasted TRINITY sequences. The phylogenetic trees were expanded to include the sequences of orthologues from other insect species. The TRINITY sequences were termed potential proteins based on location within the tree.

Results Obtained

In this study, select proteins involved in the RA pathway were identified using the Horch lab database for regulated proteins, as well as an extended literature search for potential proteins not documented in the initial research. The proteins collected from Wang 2020's research were found to be related to Adh, Rdh, CYP450, and 17BHSD. Within these protein families, individual proteins were identified in coordination with the RA pathway and their regulation status during deafferentation of the cricket, *Gryllus bimaculatus*. All proteins identified were upregulated. A literature search found additional proteins connected to the RA pathway outside the initial identification in Wang's research. The additional proteins were retinal dehydrogenase 12 and aldehyde dehydrogenase. After selection of both protein collections, a search in NCBI was conducted to find their *Drosophila melanogaster* counterpart. If the protein was uncharacterized, an additional similar protein was also included in the study from another *Drosophila* species or a similar insect (Table 1).

BLAST procedures produced cricket TRINITY sequences that were later used to form phylogenetic trees in coordination with the original Drosophila accessions (Table 1) and additional insect accessions (Table 3), within each individual protein family—Adh, Rdh, Cytochrome P450, and 17-beta (Figure 1). The formation of these trees enabled the potential identification of proteins in the cricket transcriptome (Figure 1) as well as uncharacterized proteins in *Drosophila melanogaster* (Table 1). Uncharacterized protein (NP_610235.1) was found to be 82% identical to *Drosophila arizonae*'s aldose reductase. Uncharacterized protein (NP_647840.1) was found to be 95% *Drosophila serrata*'s aldose reductase-like. Uncharacterized protein (NP_733183.1) was found to be 65% identical to *Cryptotermes secundus*'s retinal dehydrogenase 1. Uncharacterized protein (NP_651111.1) was found to be 72% identical to *Drosophila suzukii's* 17-beta-hydroxysteroid dehydrogenase 13-like. These additional similar proteins were used to potentially identify the unknown proteins in the *Drosophila* transcriptome.

Significance and Interpretation of Results

The literature review conducted on the RA pathway proteins revealed important information about the process, its effects on neuronal plasticity and differentiation, and the protein domains that can be useful in further studies of these proteins with respect to synaptic plasticity in the cricket. Retinoic acid is a derivative of retinol, or Vitamin A, and may be involved in differentiation, central nervous system patterning, and synaptic plasticity. The wide distribution of different RA receptors throughout the adult brain strongly suggests that retinoid signaling is necessary for neuronal plasticity, neurogenesis, and for cognitive and motor controlling functions (Das et al. 2014). In this study, *Drosophila melanogaster* was predominately used as a model organism because of its suspected role in the visual development of the fruit fly (Wang et al. 2005). Another advantage of *Drosophila* is its highly characterized and studied transcriptome, aiding mostly in the identification of alcohol dehydrogenase-related proteins in *Drosophila* (Benach et al. 2005 and Winberg et al. 1998). The enzymes involved in synthesizing RA were the focus of this paper. Alcohol dehydrogenase and other short-chain dehydrogenase/reductase enzymes catalyze the oxidation of retinol to the aldehyde, retinal, which is subsequently oxidized to RA by the action of retinal dehydrogenases, aldehyde dehydrogenases, and cytochrome p450 enzymes. Retinol dehydrogenase catalyzes the reduction of retinal to retinol to promote retinoid storage (Duester 1996). The 17-beta-hydroxysteroid dehydrogenase (17BHSD) protein was found to be active later in the pathway after the synthesis of RA. 17BHSD controls the last step in

the formation of all androgens and all estrogens, including 17-beta estradiol, which has shown to play a role in modulating synaptic plasticity in the hippocampus (Hojo et al. 2011 and Labrie et al. 1997).

The formation of phylogenetic trees helps visualize the connections within protein families as well as characterize the proteins (Figure 1). Using pairwise alignments, uncharacterized *Drosophila melanogaster* proteins were compared to *Drosophila* organisms and similar insects to potentially characterize these unidentified proteins. Statistically, this study supports the identification of the uncharacterized protein NP_647840.1 as aldose reductase-like, because it was found to be 95% identical to *Drosophila serrata's* aldose reductase-like. The generation of these trees aids the identification of unknown proteins and provides opportunity for future research. In future studies, research on more proteins involved in the RA pathway, such as its receptors, could provide vast information on the insect nervous system and would be incredibly valuable in increasing knowledge on neural plasticity.

Figures/Charts

Table 1. *Drosophila melanogaster* retinoic acid pathway protein accessions

Family	Drosophila Counterpart to Wang 2020 Observed Proteins	Drosophila Melanogaster Alternative, If Necessary	Most Similar Drosophila Counterpart Accession Number	% identical to uncharacterized
Alcohol Dehydrogenase	alcohol dehydrogenase	-	AAA28347	-
	aldo-keto reductase 1B, isoform D	-	AGB94411.1	-
	aldehyde dehydrogenase, isoform B	-	AGB92825.1	-
	aldose reductase	*Drosophila arizonae*	XP_017867367	82%
	uncharacterized protein Dmel_CG9436	-	NP_610235.1	
	aldose reductase-like	*Drosophila serrata*	XP_020807626.1	95%
	uncharacterized protein Dmel_CG10863	-	NP_647840.1	
Retinal and Retinol Dehydrogenase	Retinal dehydrogenase 1	*Cryptotermes secundus*	XP_023718025.1	65%
	uncharacterized protein, isoform A	-	NP_733183.1	
	retinol dehydrogenase 12	-	NP_610308.2	-
	retinol dehydrogenase B	-	AGB96211.1	-
Cytochrome P450	cytochrome P450	-	AAC47424.1	-
	Cyp6g1, isoform B	-	NP_001286335.1	40%
	CP450 6j1	-	*Not found in Drosophila organisms*	
	Cytochrome P450 9f2	-	NP_650189.1	47%
	CP450 9e2	-	*Not found in Drosophila organisms*	
	CP450 6a13, isoform E	-	AAF59076.1	-
	CP450 6a14	-	NP_001286199.1	-
	cytochrome P450-4c3	-	AAF57098.1	-
	cytochrome p450 monooxygenase	-	AAB05550.1	-
17-beta-hydroxysteroid dehydrogenase	uncharacterized protein Dmel_CG13833	-	NP_651111.1	72%
	17-beta-hydroxysteroid dehydrogenase 13-like	*Drosophila suzukii*	XP_016931784.1	

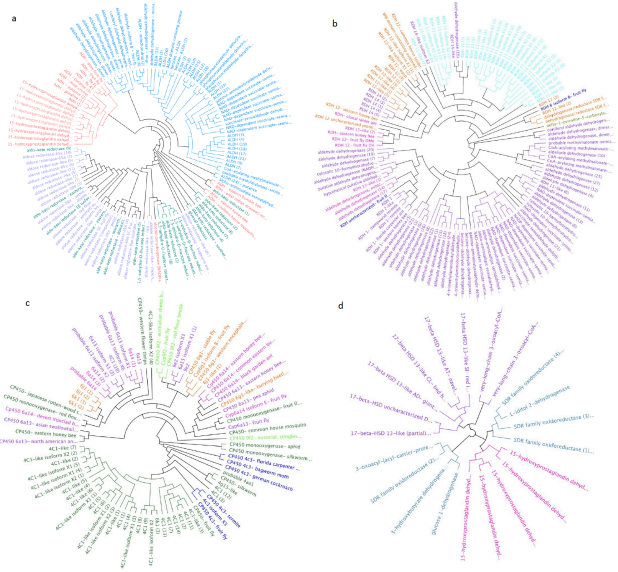

Figure 1. Phylogenetic tree of TRINITY sequences for proteins related to (a) Alcohol dehydrogenases, (b) Retinol/al dehydrogenases, (c) Cytochrome p450s, and (d) 17-beta hydroxysteroid dehydrogenase in the cricket transcriptome along with reference proteins in *Drosophila melanogaster* and other insects.

Acknowledgements and References

We would like to thank the Maine Space Grant Consortium for funding this project. The author would also like to thank Professor Hadley Horch and Justin Beckman for their continued support and guidance throughout the study.

Benach, Jordi, Jan Wolof Winberg, John Sigurd Svendsen, Silvia Atrian, Rose Gonzalez-Duarte, and Rudolf Ladenstein. "*Drosophila* Alcohol Dehydrogenase: Acetate–Enzyme Interactions and Novel Insights into the Effects of Electrostatics on Catalysis." Journal of Molecular Biology 345, no 3 (February 1998): 561–570.

Das, Bhaskar C., Pritam Thapa, Radha Karki, Sasmita Das, Sweta Mahapatra, Ting-Chun Liu, Ingrid Torregroza, et al. "Retinoic Acid Signaling Pathways in Development and Diseases." *Bioorganic & Medicinal Chemistry* 22, no. 2 (January 15, 2014): 673–83.

Duester, Gregg. "Involvement of Alcohol Dehydrogenase, Short-Chain Dehydrogenase/Reductase, Aldehyde Dehydrogenase, and Cytochrome P450 in the Control of Retinoid Signaling by Activation of Retinoic Acid Synthesis" *Biochemistry* 35, no. 38 (September 1996): 12222-12227.

Hojo, Yashushi, Shimpei Higo, Suguru Kawato, Yusuke Hatanaka, Yuuki Ooishi, Gen Murakami, Hirotaka Ishii, Yoshimasa Komatsuzaki, Mari Ogiue-Ikeda, Hideo Mukai, and Tetsuya Kimo. Hippocampal synthesis of sex steroids and corticosteroids: essential for modulation of synaptic plasticity. *Front Endocrinol (Lausanne)* 2, no. 43. (October 10, 2011) doi:10.3389/fendo.2011.00043

Horch, Hadley W., Sarah S. McCarthy, Susan L. Johansen, and James M. Harris. "Differential Gene Expression During Compensatory Sprouting of Dendrites in the Auditory System of the Cricket Gryllus Bimaculatus." *Insect Molecular Biology* 18, no. 4 (August 2009): 483–96.

Horch, Hadley Wilson, Alexandra Pfister, Olaf Ellers, and Amy S. Johnson. "Plasticity in the Cricket Central Nervous System." In *The Cricket as a Model Organism: Development, Regeneration, and Behavior*, edited by Hadley Wilson Horch, Taro Mito, Aleksandar Popadić, Hideyo Ohuchi, and Sumihare Noji, 105–28. Tokyo: Springer Japan, 2017.

Labrie, Fernand, Van Luu-The, Sheng-Xiang Lin, Claude Labrie, Jacques Simard, Roch Breton, and Alain Bélanger. "The key role of 17β-hydroxysteroid dehydrogenases in sex steroid biology." Steroids 62, no. 1, (January 1997): 148-158.

Wang, Felicia. "Differential gene expression during compensatory plasticity in the prothoracic ganglion of the cricket, *Gryllus bimaculatus*." (2020)

Wang, Tao, and Craig Montell. "Rhodopsin Formation in *Drosophila* Is Dependent on the PINTA Retinoid-Binding Protein." *Journal of Neuroscience* 25, no 21 (May 2005): 5187-5194.

Winberg, Jan Olof, and John S. McKinley-McKee. "Drosophila melanogaster alcohol dehydrogenase: mechanism of aldehyde oxidation and dismutation." *Biochemical Journal* 329 (February 1998): 561–570.