**Characterization of the Spaetzle Protein Family in the Deafferented Mediterranean field cricket (*Gryllus bimaculatus*)**

Researcher: Anthea L. Bell

Advisor: Hadley Horch

Bowdoin College

Biology/Neuroscience Department

**Abstract**

Neural plasticity describes how neurons change in structure and function to adapt to changes in their environment. The Horch lab investigates cricket neural plasticity by studying how the cricket auditory system recovers after an injury in a mechanism called compensatory growth. To injure the cricket, the lab deafferents or disrupts the cricket’s auditory neurons by removing one of its ears located on the first forelegs. In the days after deafferentation (injury of the auditory neurons), neurons from the afflicted side grow across to the unafflicted side to make connections that restore the function of the auditory neurons (*1*). Thus, the cricket regains its ability to detect important sounds such as mating calls from other crickets. The growth of neurons to the unafflicted side to restore hearing is known as the compensatory growth mechanism. To investigate the proteins responsible for the compensatory growth of neurons in crickets, the Horch lab created a database of all the transcripts—RNA instructions to make particular proteins—present in injured and non-injured cricket (*2*). The lab compared the transcripts present in injured versus non-injured crickets so the lab could observe which genes were regulated up or down and, thus, correlated with the compensatory growth of neurons. Up-regulation of a gene occurs when signals tell the cell to increase its production of a specific protein, while down-regulation of a gene is when signals tell a cell to decrease its production of a specific protein. Based on the predictions of differential expression data, we have hypothesized that the Toll receptor protein was downregulated one day following foreleg removal. One of the key functions of the Toll signaling pathway is to serve as an immune response against gram-positive bacteria and fungi (*3, 4*). However, research in fruit flies (*Drosophila* *melanogaster*) has shown that the Toll pathway may also be associated with the development of neurons in fruit fly embryos (*5, 6*). Furthermore, the Spaetzle protein that activates Toll receptors bears resemblance to mammalian neurotrophins such as the nerve growth factor (NGF) (*7*). Neurotrophins make up a family of proteins in vertebrates that regulate neuronal survival, proliferation, and differentiation during development and modulate learning and memory in adults (*8*). Overall, we predict that Spaetzle may be involved in the compensatory growth mechanism in deafferented crickets because previous literature has shown that the Spaetzle family is associated with neurotrophins and neuronal development. As a first step in investigating the possible role of Spaetzle in cricket compensatory growth, we characterized the Spaetzle family in the cricket (*Gryllus bimaculatus*) by comparing it to Spaetzles in other species. We found four variations of the Spaetzle protein in the cricket: Spz-1, Spz-3, Spz-5, and Spz-6 as compared to the six Spaetzle variations (Spz-1 to Spz-6) (*9*) found in *Drosophila*, providing information about the ancestral family of Spaetzles in insects.

**Projects Objectives**

To find Spaetzle protein sequences in our database of RNA transcripts found in injured crickets and then make phylogenetic trees comparing cricket Spaetzles to those found in other species. From the phylogenetic trees, we hope to differentiate the different Spaetzle protein variants (also known as homologs) found in injured crickets.

**Methods**

Spaetzle protein sequences found in a variety of insect species were taken from the GenBank database (*10*). We used Spaetzle protein sequences from fruit flies (*Drosophila* *melanogaster*) to search for similar protein sequences in the cricket using the tblastn program in Geneious Prime, version 2020.1.2 (*11*). The two databases we searched in were the database the Horch lab created using RNA extracted from the prothoracic ganglion of injured and non-injured crickets as well as a database from the web that contained RNA transcripts extracted from cricket embryos (http://asgard.rc.fas.harvard.edu/blast.html) (*12*). Two phylogenetic trees were made based on different multiple sequence alignment and tree-building programs in Geneious Prime (*11*). One tree was based on a MUSCLE alignment, version 3.8.425 (*13, 14*), and the tree was made using the Jukes-Cantor genetic distance model and the neighbor-joining tree build method. The Spaetzle protein found in the brine shrimp (ADQ43816.1) was used as the outgroup. For the second tree, we tried to replicate the process published by Wang and Zhu (*15*). Although Wang and Zhu used ClustalX to align their sequences, we used the updated Clustal Omega program, version 1.2.3 (*16, 17*), instead. Wang and Zhu made a UPGMA phylogenetic tree, so we used the Jukes-Cantor genetic distance model and the UPGMA tree build method to make our tree. An NCBI conserved domain search (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml)(*18*) was also conducted on the protein sequences we found.

**Results and Significance/Interpretation of Results**

Characterization of the Spaetzle family in the deafferented Mediterranean field cricket (*Gryllus bimaculatus*) has yielded four Spaetzle variants that most closely resemble Spz-1, Spz-3, Spz-5, and Spz-6 in *Drosophila*. Out of the four variants, we predict that Spz-1 and Spz-5 may be involved in the compensatory growth mechanism of the cricket because previous literature has shown that the overexpression of Spz-2 (DNT-1) and Spz-5 (DNT-2) decreases the instances of cell death in the developing embryonic central nervous system (CNS), which suggests that Spz-2 and Spz-5 are involved in neuronal survival during development (*19*). The same study showed that Spz-1 is also important for neuronal survival in the developing CNS, but to a lesser extent than Spz-2 or Spz-5. Spz-1, Spz-2, and Spz-5 appeared to be more similar to each other than to Spz-3, Spz-4, and Spz-6. This corroborates the findings of another study (*20*) and our phylogenetic trees (Fig. 1) that suggests that Spz-1, Spz-2, and Spz-5 are more closely related to each other than with the other three members of the Spaetzle family.

Our NCBI conserved domain searches also showed that the sequences from the adult prothoracic ganglion and the embryonic transcriptomes contained the cystine knot domain, which is a conserved motif among notable neurotrophin molecules (*21, 22*). Neurotrophins are proteins found in vertebrates that influence the development of neurons in embryos by regulating neuronal survival and proliferation (*8*). Thus, the presence of a cystine knot domain further supports how the Spaetzle proteins we found may serve a similar function as neurotrophins by influencing neuron growth in the cricket. However, a study has observed that the cystine knot structure of Spz-1, Spz-2, and Spz-5 are more similar to that of neurotrophins than Spz-3, Spz-4, and Spz-5 are (*19*). This further suggests that Spz-1, Spz-2, and Spz-5 are more likely to be involved in axon guidance and neuron growth than Spz-3, Spz-4, and Spz-6.

Based on our phylogenetic trees, it appears that Spz-2 was missing from the adult prothoracic ganglion and the embryonic transcriptomes. One possible explanation is that Spz-2 evolved more recently among insects. Although we found suspected Spz-2 members in Isoptera (termites), Thysanoptera (thrips), and Hymenoptera (bees) orders, we found a greater variety of suspected Spz-2 sequences in the Lepidoptera (moths & butterflies) and Diptera (true flies) orders (Fig. 1). Since the order Mediterranean field crickets belong to (Orthoptera) is much older than the Lepidoptera and Diptera orders (*23*), we predict that Spz-2 evolved from another Spaetzle member—possibly Spz-1 or Spz-5 due to the similarities between these three Spaetzle members—sometime after the evolution of crickets. However, we do not have a clear prediction of why Spz-4 was not found in the Mediterranean field cricket because it is found among a variety of orders and appears to be well conserved (Fig. 1). The lack of Spz-4 in our findings may suggest Spz-4 might just not be expressed in the embryo during development or in prothoracic ganglion during the compensatory growth mechanism after deafferentation. However, Spz-4 may be expressed in other tissue of the cricket under different conditions. Thus, searching the genome in a future study would be necessary to help us understand if Spz-4 is present in the cricket.

When looking back at which proteins in our database were upregulated versus downregulated, we noticed that two Spz-1 proteins in the cricket (TRINITY27\_DN134778\_c0\_g1\_i3.p1 and TRINITY30\_DN137643\_c0\_g2\_i1.p1) were upregulated one day post-deafferentation. This supports our hypothesis that Spz-1 may be involved in the compensatory growth mechanism in injured crickets.

Further wet lab experiments are necessary to confirm the upregulation of Spz-1 in the deafferented cricket—as well as to check if other Spaetzle members were up- or downregulated. If we find that Spz-1, and possibly Spz-5, are indeed upregulated in deafferented crickets, Spaetzle may play a role in dendrite guidance across the midline of the injured cricket.

**Figures/Charts**

|  |  |
| --- | --- |
| **TRINITY Sequence**  | **Expected Spz Member**  |
| isotig09642 translation 1 | Spz-1  |
| TRINITY27\_DN134778\_c0\_g1\_i3.p1 translation 1 | Spz-1 |
| TRINITY30\_DN137643\_c0\_g2\_i1.p1 translation 1 | Spz-1 |
| TRINITY25\_DN129355\_c0\_g1\_i1.p1 translation 3 | Spz-3 |
| TRINITY25\_DN129355\_c0\_g1\_i4.p1 (reversed) translation 1 | Spz-3 |
| TRINITY25\_DN129355\_c0\_g1\_i6.p1 translation 1 | Spz-3 |
| TRINITY27\_DN127802\_c3\_g1\_i5.p1 (reversed) translation 1 | Spz-3 |
| TRINITY30\_DN132994\_c0\_g3\_i1.p1 (reversed) translation 1 | Spz-3 |
| TRINITY21\_DN56706\_c10\_g1\_i1.p1 translation 1 | Spz-5 |
| TRINITY25\_DN124009\_c0\_g1\_i3.p1 translation 1 | Spz-5 |
| TRINITY25\_DN124009\_c0\_g1\_i4.p1 translation 1 | Spz-5 |
| TRINITY21\_DN134183\_c0\_g1\_i1.p1 translation 1 | Spz-6 |
| TRINITY25\_DN42601\_c0\_g1\_i1.p1 translation 1 | Spz-6 |

**Table 1.** Sequences found in the Mediterranean field cricket (*Gryllus bimaculatus*) prothoracic ganglia (TRINITY) and embryonic (isotig09642) transcriptomes and their corresponding Spaetzle family member designation. Spz-1, Spz-3, Spz-5, and Spz-6 were expected to be found in the prothoracic ganglia of male crickets in the days post-deafferentation. Spz-1 was expected to be found in the developing cricket. Spz-2 and Spz-4 appeared to be missing in both the adult cricket prothoracic ganglia and in cricket embryos.

****

**Acknowledgments**

I would like to thank Professor Horch and our lab technician, Justin Beckman, for all their expertise and guidance through this project this summer.

**Literature Cited**

1. H. W. Horch, A. Pfister, O. Ellers, A. S. Johnson, in *The Cricket as a Model Organism: Development, Regeneration, and Behavior,* H. W. Horch, T. Mito, A. Popadic, H. Ohuchi, S. Noji, Eds. (Springer Japan, Tokyo, Japan, 2017), chap. 8, pp. 105-128.

2. H. P. Fisher *et al.*, De novo assembly of a transcriptome for the cricket Gryllus bimaculatus prothoracic ganglion: An invertebrate model for investigating adult central nervous system compensatory plasticity. *PloS one* **13**, (2018).

3. D. Hultmark, Drosophila immunity: paths and patterns. *Current opinion in immunology* **15**, 12-19 (2003).

4. M. P. Belvin, K. V. Anderson, A conserved signaling pathway: the Drosophila toll-dorsal pathway. *Annual review of cell and developmental biology* **12**, 393-416 (1996).

5. N. Anthoney, I. Foldi, A. Hidalgo, Toll and Toll-like receptor signalling in development. *Development* **145**, dev156018 (2018).

6. S. Roth, T. Schupbach, The relationship between ovarian and embryonic dorsoventral patterning in Drosophila. *Development* **120**, 2245-2257 (1994).

7. M. Lewis *et al.*, Cytokine Spätzle binds to the Drosophila immunoreceptor Toll with a neurotrophin-like specificity and couples receptor activation. *Proceedings of the National Academy of Sciences* **110**, 20461-20466 (2013).

8. L. F. Reichardt, Neurotrophin-regulated signalling pathways. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 1545-1564 (2006).

9. J. S. Parker, K. Mizuguchi, N. J. Gay, A family of proteins related to Spätzle, the toll receptor ligand, are encoded in the Drosophila genome. *Proteins: Structure, Function, and Bioinformatics* **45**, 71-80 (2001).

10. K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, E. W. Sayers, GenBank. *Nucleic acids research* **44**, D67-D72 (2016).

11. M. Kearse *et al.*, Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649 (2012).

12. V. Zeng, C. G. Extavour, ASGARD: an open-access database of annotated transcriptomes for emerging model arthropod species. *Database* **2012**, (2012).

13. R. C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics* **5**, 113 (2004).

14. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* **32**, 1792-1797 (2004).

15. Y. Wang, S. Zhu, Evolutionary and functional epitopes of the Spätzle protein: New insights into activation of the Toll receptor. *Cellular and molecular life sciences* **66**, 1595-1602 (2009).

16. F. Sievers, D. G. Higgins, Clustal Omega for making accurate alignments of many protein sequences. *Protein Science* **27**, 135-145 (2018).

17. F. Sievers *et al.*, Fast, scalable generation of high‐quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* **7**, 539 (2011).

18. S. Lu *et al.*, CDD/SPARCLE: the conserved domain database in 2020. *Nucleic acids research* **48**, D265-D268 (2020).

19. B. Zhu *et al.*, Drosophila neurotrophins reveal a common mechanism for nervous system formation. *PLoS biology* **6**, (2008).

20. M. Chowdhury *et al.*, Toll family members bind multiple Spätzle proteins and activate antimicrobial peptide gene expression in Drosophila. *Journal of Biological Chemistry* **294**, 10172-10181 (2019).

21. J. Murray-Rust *et al.*, Topological similarities in TGF-β2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure* **1**, 153-159 (1993).

22. S. Iyer, K. R. Acharya, Tying the knot: The cystine signature and molecular‐recognition processes of the vascular endothelial growth factor family of angiogenic cytokines. *The FEBS journal* **278**, 4304-4322 (2011).

23. B. Misof *et al.*, Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763-767 (2014).