



Molecular evolution of sperm proteins in *Danaus* butterflies

T. Jeffrey Cole¹ and Jamie R. Walters², Ph.D.

Dept. of Biological and Environmental Science, Samford University, Birmingham, AL 35229-2234.¹
Dept. of Ecology and Evolutionary Biology, University of Kansas, KS 66045.²



Introduction

Reproductive proteins tend to evolve rapidly.^{1,2} Sperm morphology is also highly diverse across taxa.³ These observations likely reflect pressures of sexual selection. High rates of evolution are therefore expected among sperm proteins in *Danaus* butterflies.

We investigated the evolutionary rates of sperm proteins in *Danaus* butterflies. A comprehensive sperm proteome exists for *Danaus plexippus*, including both types of dimorphic sperm.

We hypothesize that sperm proteins evolved more rapidly than the genomic average.

Gene origin	Total genes	Total sampled	Percent sampled
Genome	15,130	11,524	76.2%
Autosome	13,955	10,556	75.6%
Z linked	1,175	928	79.0%
Sperm	717	599	83.5%
Apyrene	50	44	88.0%
Eupyrene	319	259	81.2%
Shared	348	296	85.1%

Table 1 Gene types present in *D. plexippus* genome, and portion that Ka/Ks ratios were obtained

Sperm dimorphism

We have the sperm proteome for both sperm types

Eupyrene: fertilizing component

Apyrene: incompetent sperm type

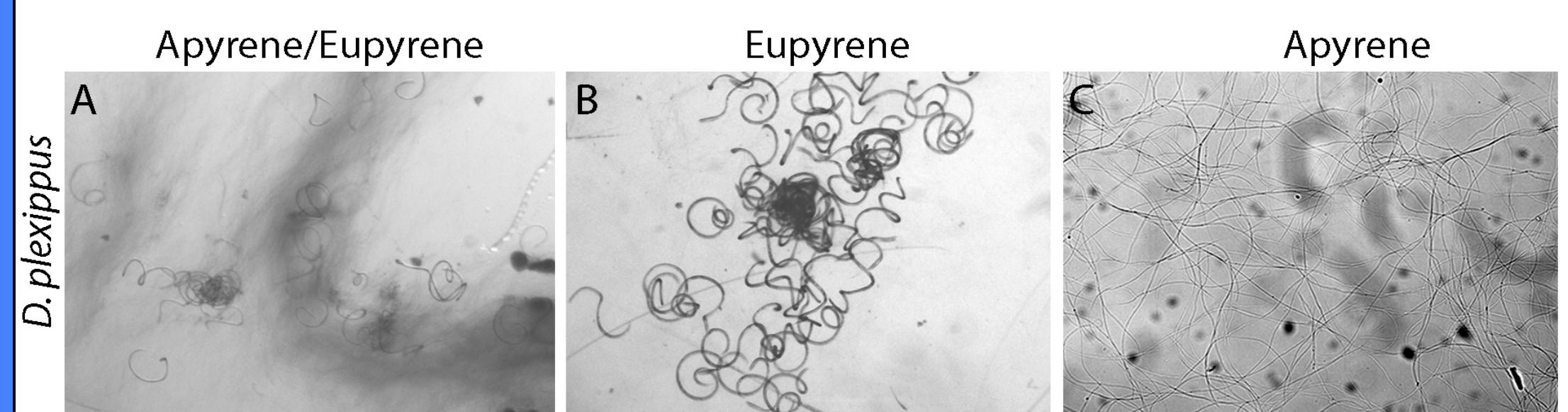


Fig 1 Micrographs of both sperm types in *Danaus plexippus*

Methods

- We measured evolutionary rates using Ka/Ks ratios
- We used *Danaus plexippus* and *Danaus chrysippus* genome assemblies for pairwise Ka/Ks ratio calculation of orthologous genes with gKaKs pipeline.⁴ We used apyrene and eupyrene proteome for *D. plexippus* to calculate Ka/Ks ratios of sperm proteins.

Ka/Ks = rate of functional evolution

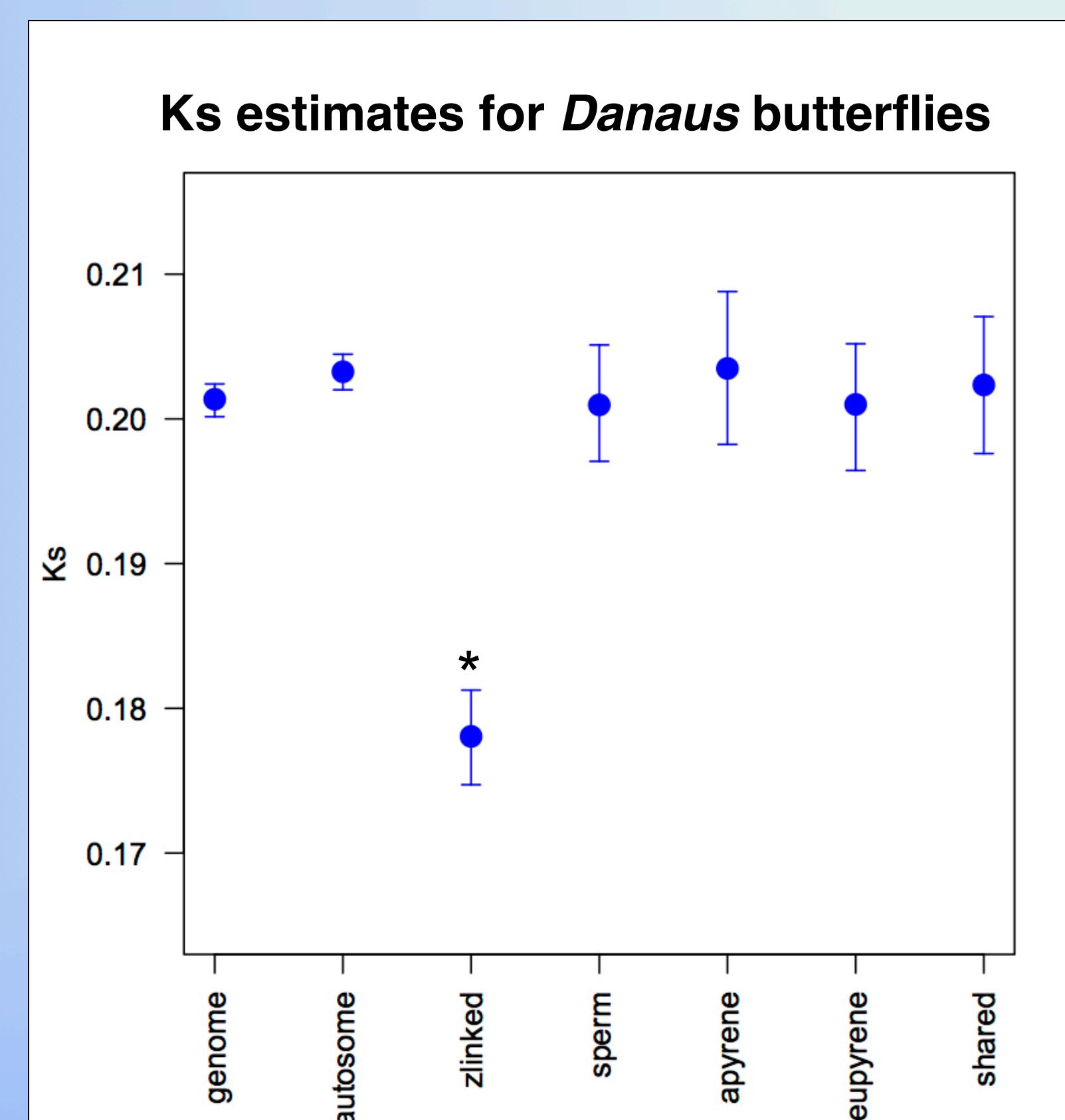
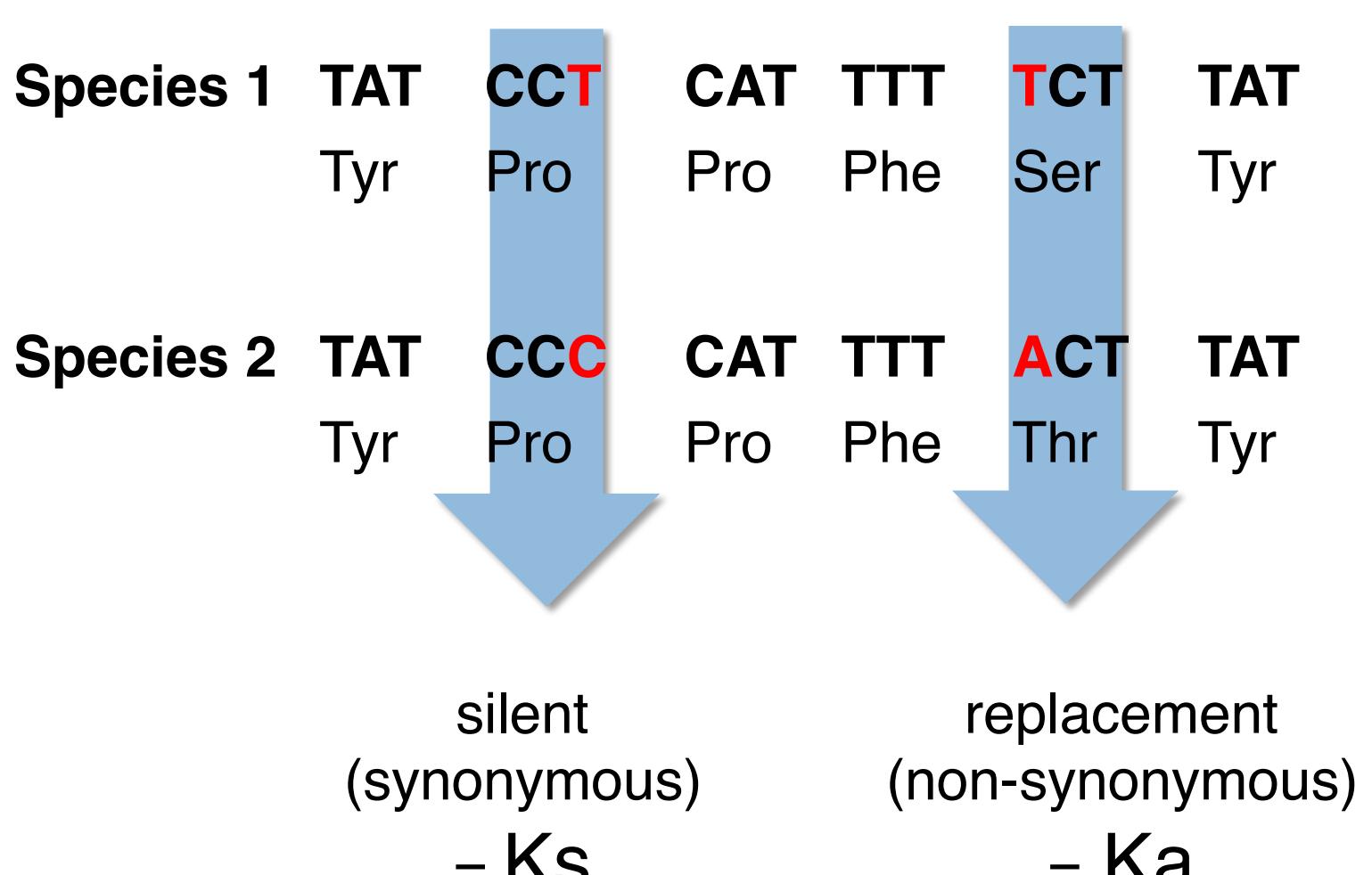


Fig 3 Mean Ks values for each gene type from *D. plexippus* and *D. Chrysippus* alignment. Error bars indicate 95% confidence intervals obtained from 1,000 bootstrapped replicates. Asterisks indicate significance ($p < 0.05$).

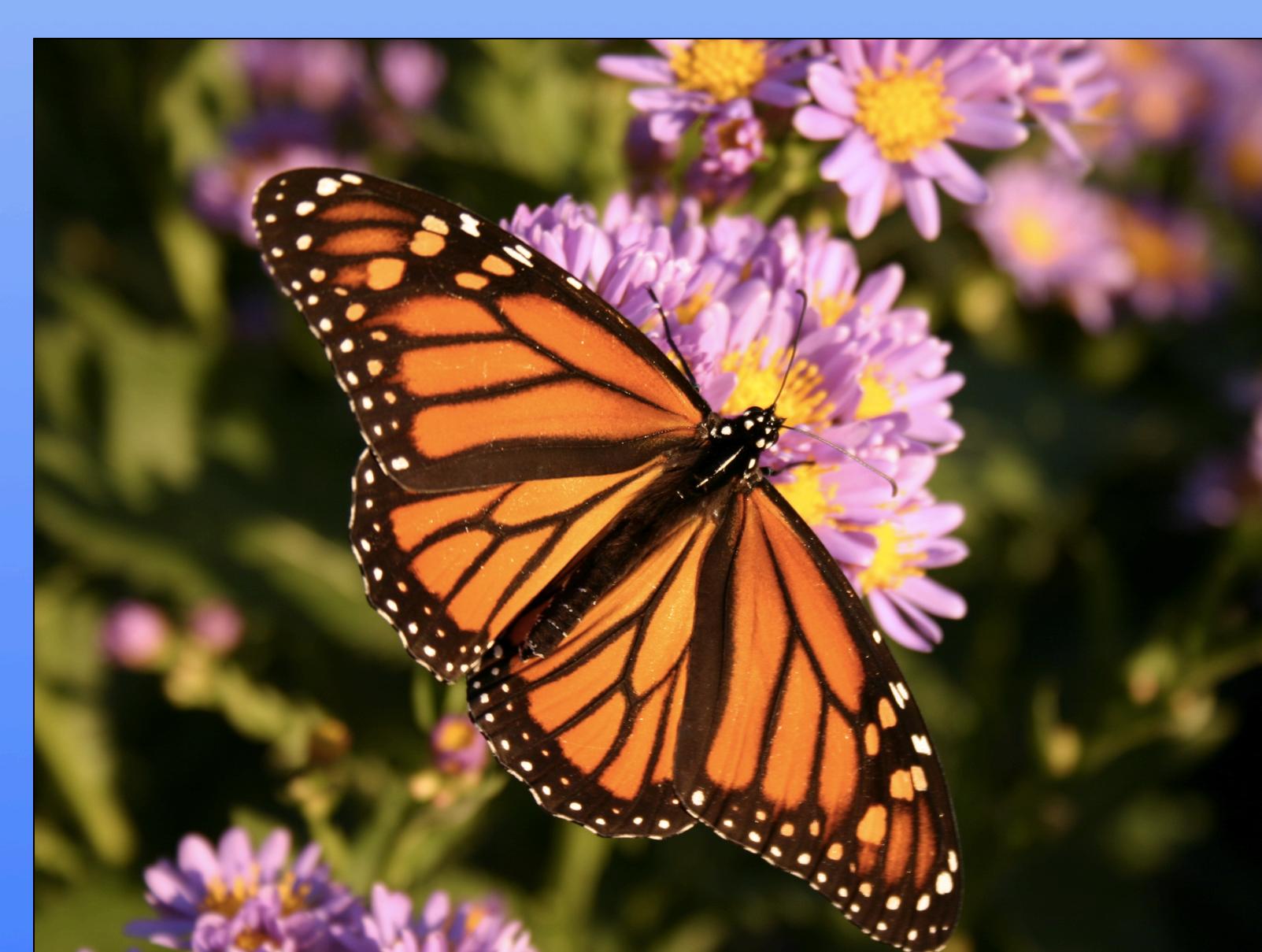


Fig 5 *Danaus plexippus* adult. Photo credit: <https://commons.wikimedia.org>

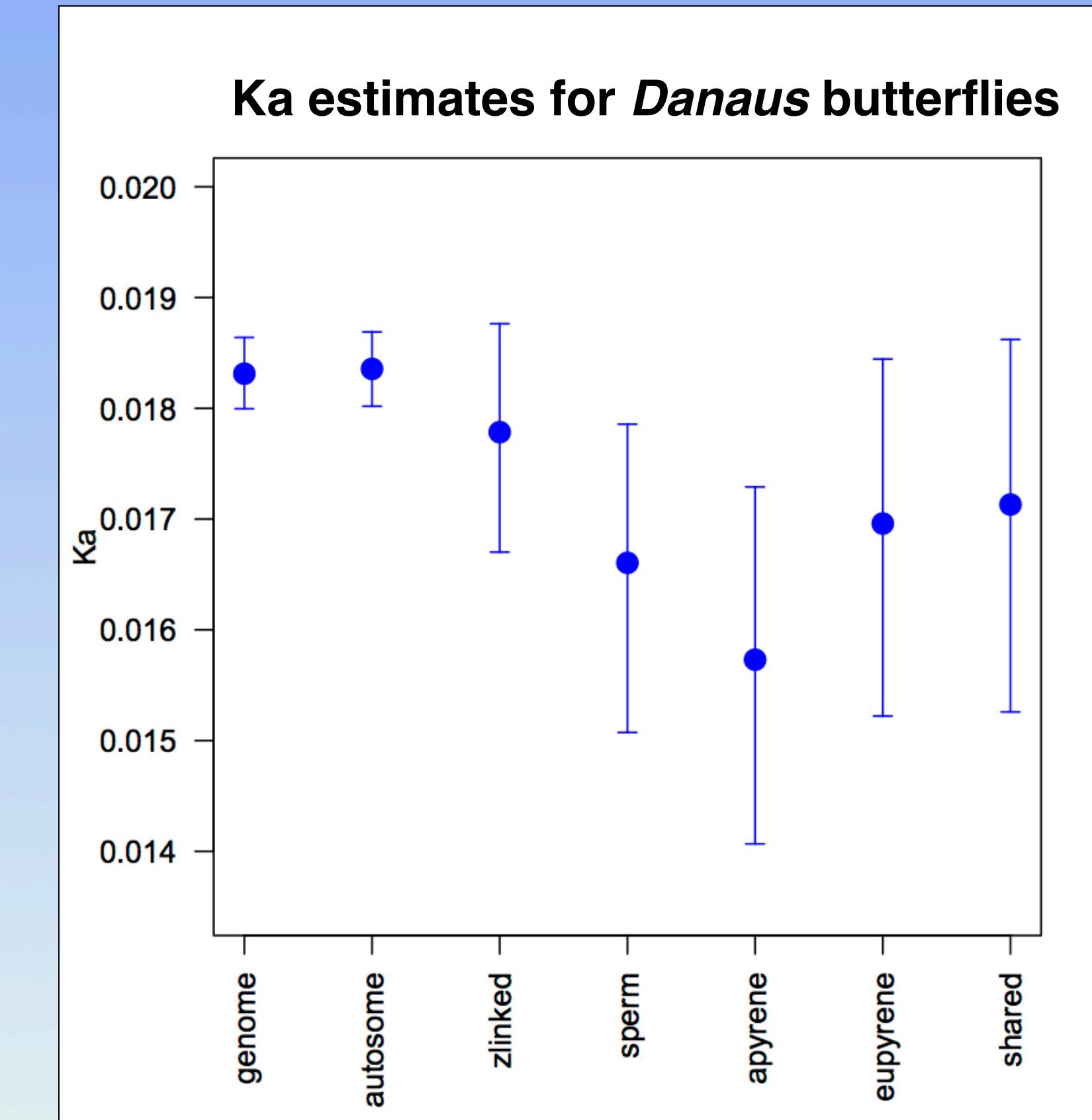


Fig 2 Mean Ka values for each gene type from *D. plexippus* and *D. Chrysippus* alignment. Error bars indicate 95% confidence intervals obtained from 1,000 bootstrapped replicates.

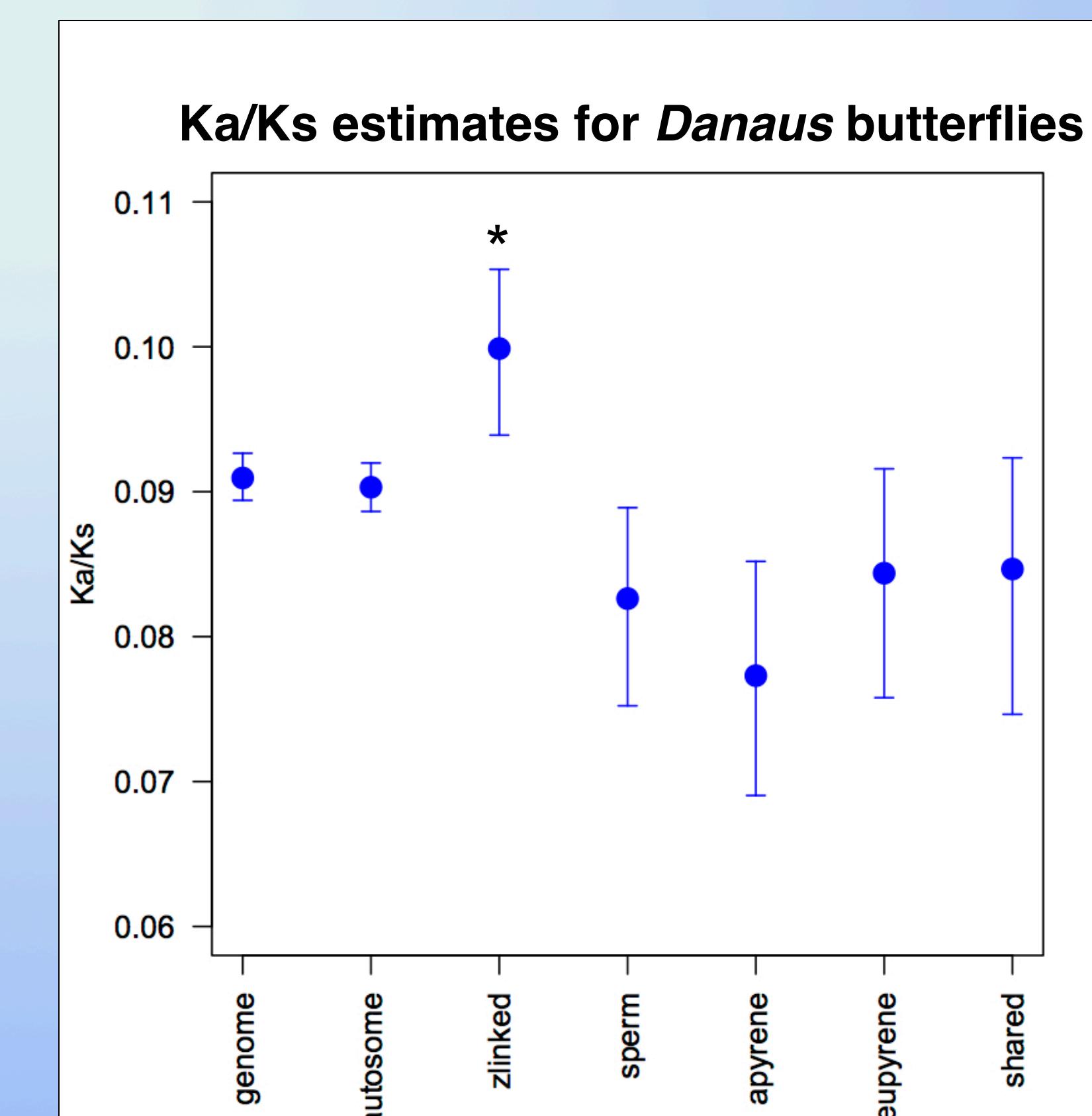


Fig 4 Mean Ka/Ks values for each gene type from *D. plexippus* and *D. Chrysippus* alignment. Error bars indicate 95% confidence intervals obtained from 1,000 bootstrapped replicates. Asterisks indicate significance ($p < 0.05$).

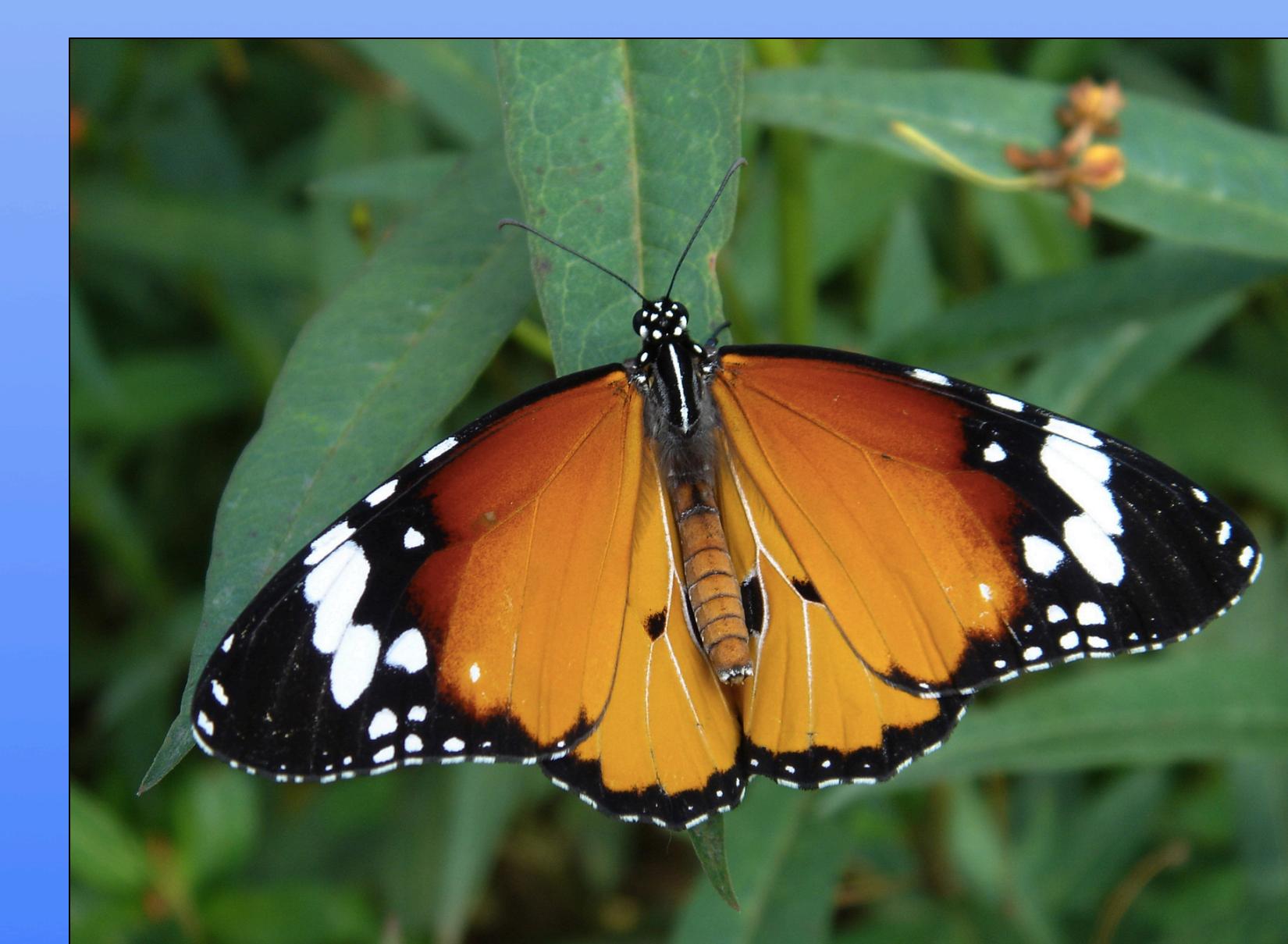


Fig 6 *Danaus chrysippus* adul. Photo credit: <https://www.ufz.de>

Results

- There is no significant difference in Ka/Ks ratios between sperm proteins and the genomic average (Fig. 4).
- There is no significant difference between the Ka/Ks ratios between the sperm types.
- Z linked genes contained Ks values significantly less than all other gene types (Fig. 3) ($p < 0.05$), while also containing Ka/Ks ratios significantly greater than all other gene types (Fig. 4) ($p < 0.05$).

Discussion

- We were unable to detect a difference between evolutionary rates of sperm proteins and the genomic average in *Danaus* butterflies
- Future studies should investigate the processes behind Z linked genes with drastically low Ks values
- Further exploration across different species and using more computationally extensive methods is necessary to better understand the molecular evolution of Lepidoptera sperm.

Acknowledgements

- This research was funded by NSF REU grant DBI-1262795 to University of Kansas Department of Ecology and Evolutionary Biology
- Thanks to Desiree Harpel for sperm proteomic data

Literature Cited

- Clark, N. L., Aagaard, J. E. & Swanson, W. J. Evolution of reproductive proteins from animals and plants. *Reproduction* 131, 11–22 (2006).
- Swanson, W. J. & Vacquier, V. D. Reproductive Protein Evolution. *Annu. Rev. Ecol. Syst.* 33, 161–179 (2002).
- Prakash, S., Prithviraj, E., Suresh, S., Lakshmi, N. V., Ganesh, M. K., Anuradha, M., ... Dinesh, P. Morphological diversity of sperm: A mini review. *Iran. J. Reprod. Med.* 12, 239–42 (2014).
- Zhang, C., Wang, J., Long, M. & Fan, C. gKaKs: the pipeline for genome-level Ka/Ks calculation. *Bioinformatics* 29, 645–6 (2013).

Join the effort!

All necessary shell scripts, R scripts, and relevant data can be found at the following GitHub repository for replication purposes. Feel free to fork the repository if you have further ideas for the project!

