

Introduction

Adelgids are sap-sucking insects that contain bacterial endosymbionts to help them synthesize essential amino acids. The adelgid life cycle alternates between sexual generations that parasitize spruce as a primary host and asexual generations that parasitize a secondary host plant species. The adelgid family consists of five lineages, each with a different secondary host plant: Douglas fir, fir, hemlock, larch, and pine. Each adelgid lineage has a different pair of symbionts, a primary symbiont that was acquired by the adelgid first and a secondary symbiont that was acquired later. *Vallotia* is a symbiont shared between the Douglas fir lineage, where it is the secondary symbiont, and the larch lineage, where it is the primary symbiont.

Objectives

Determine the role of *Vallotia* in amino acid synthesis in the Douglas fir lineage species *Adelges cooleyi* and the larch lineage species *Adelges abeitis* and *Adelges lariciatus*. Compare the annotated *Vallotia* genomes in the three species to infer evolutionary relationships.

Methods

FastQC was used to evaluate the quality of raw adelgid read data. The Georgia Advanced Computing Research Center (GACRC) cluster was used to assemble and annotate genomes from the raw reads. Low-quality parts of raw reads were trimmed out using Trimmomatic. PEAR was used to merge overlapping reads. SPADES was used to assemble the merged reads, and QAST was used to evaluate the assembly quality. BLAST was used to identify symbiont scaffolds.

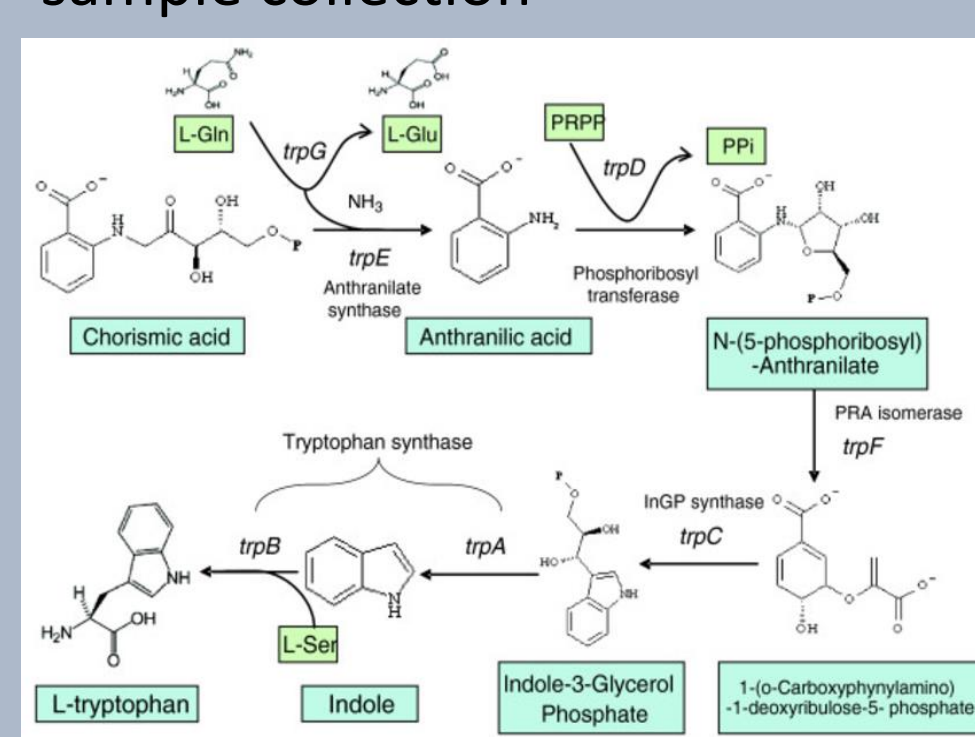
Symbiont genomes were scaffolded using SSPACE and Tablet 1.19.05.28. Genes were annotated with PROKKA and Geneious Prime, and biochemical pathways were reconstructed with help from BioCyc.



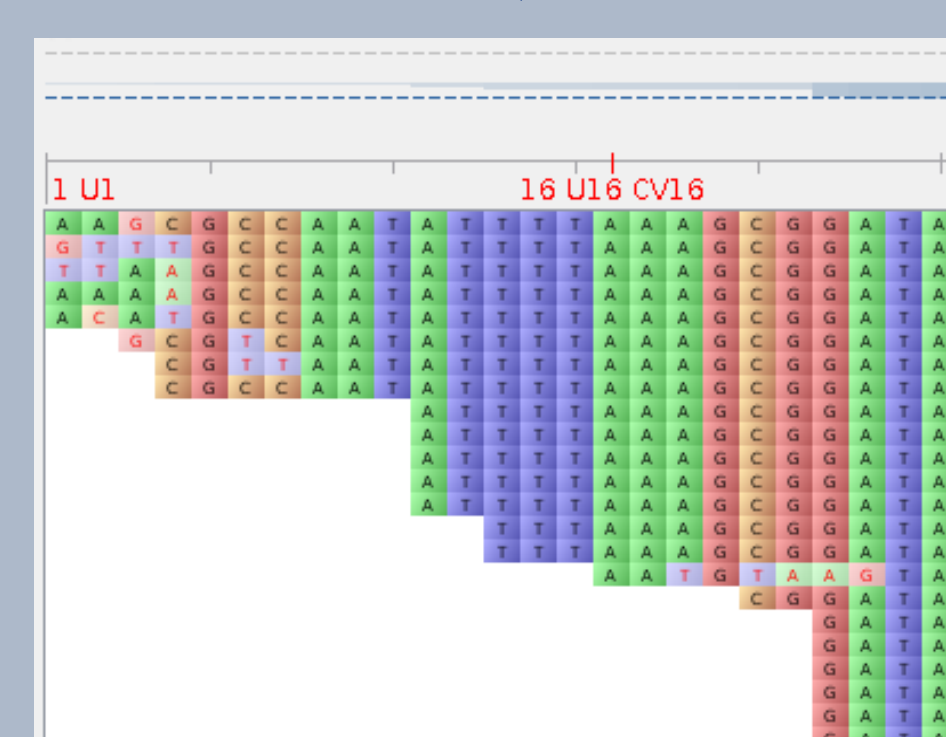
Raw adelgid genome sample collection



Sequencing of host and symbiont DNA



Gene annotation and biochemical pathway reconstruction



Bacterial symbiont genome assembly

Fig. 1: Overview of analysis procedure.

Results

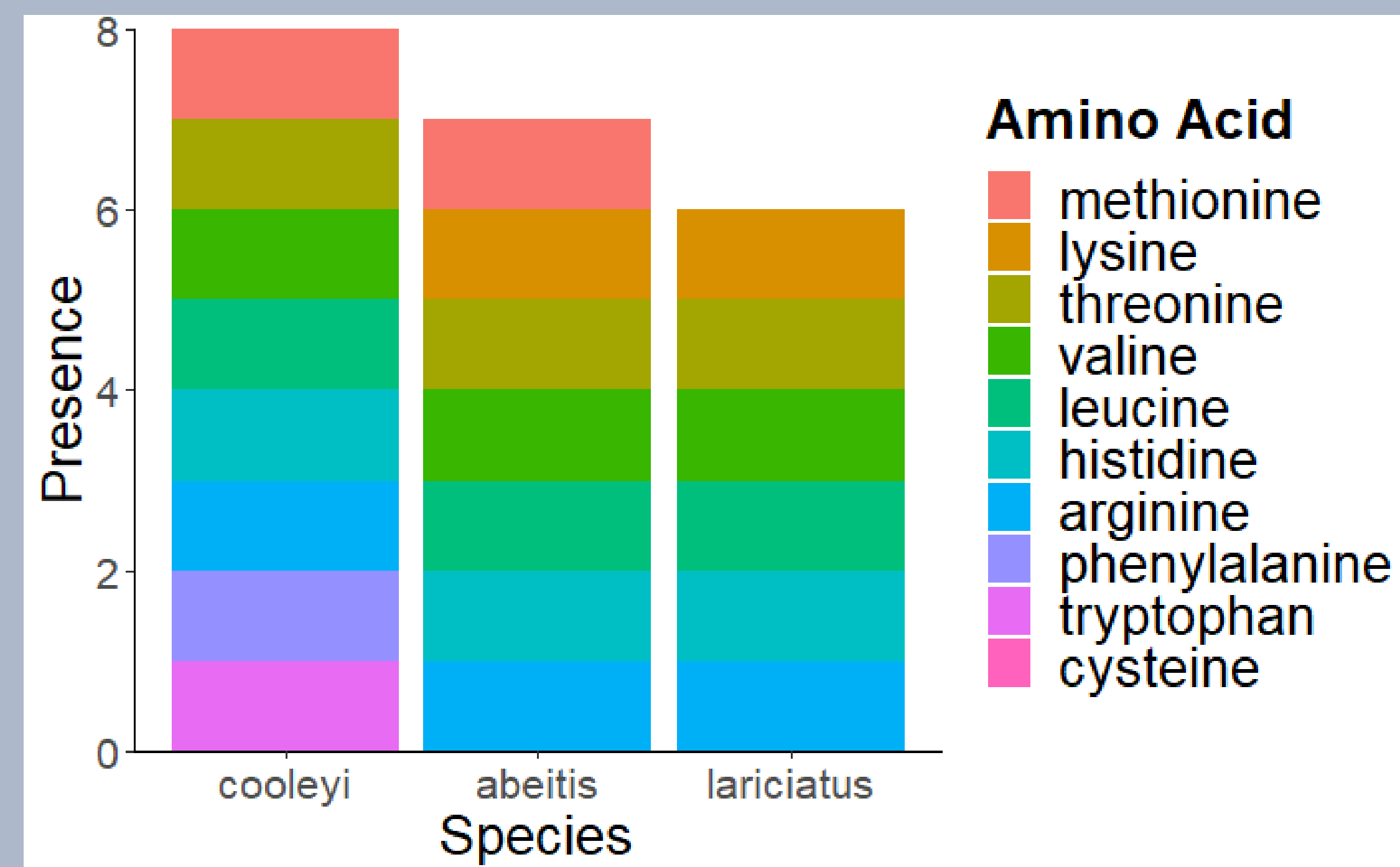


Fig. 2: Amino acids synthesized by *Vallotia* in *A. cooleyi*, *A. abeitis*, and *A. lariciatus*.

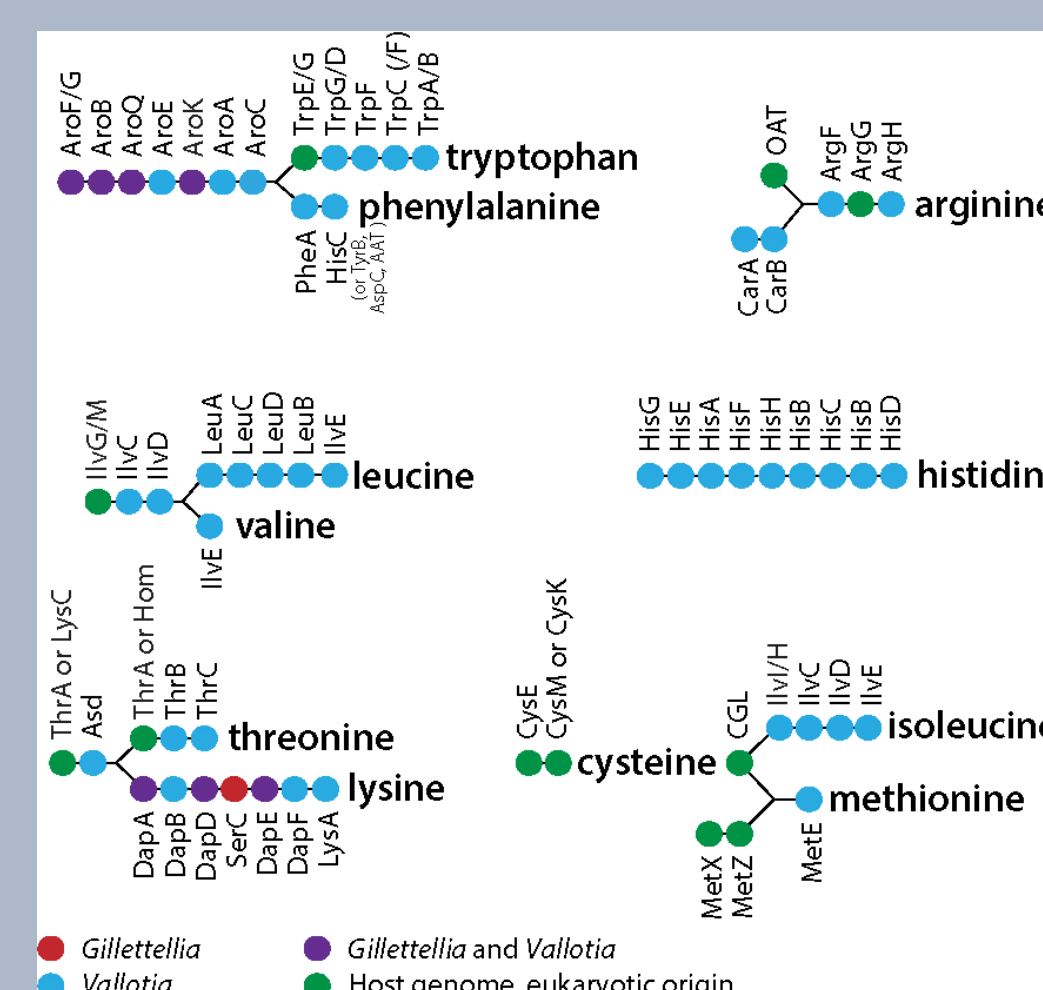


Fig. 3: Amino acid synthesis pathways of *A. cooleyi*. Each circle denotes the gene coding for the protein involved in a particular step in a biosynthetic pathway. *Gillettellia* is the primary symbiont of *A. cooleyi*.

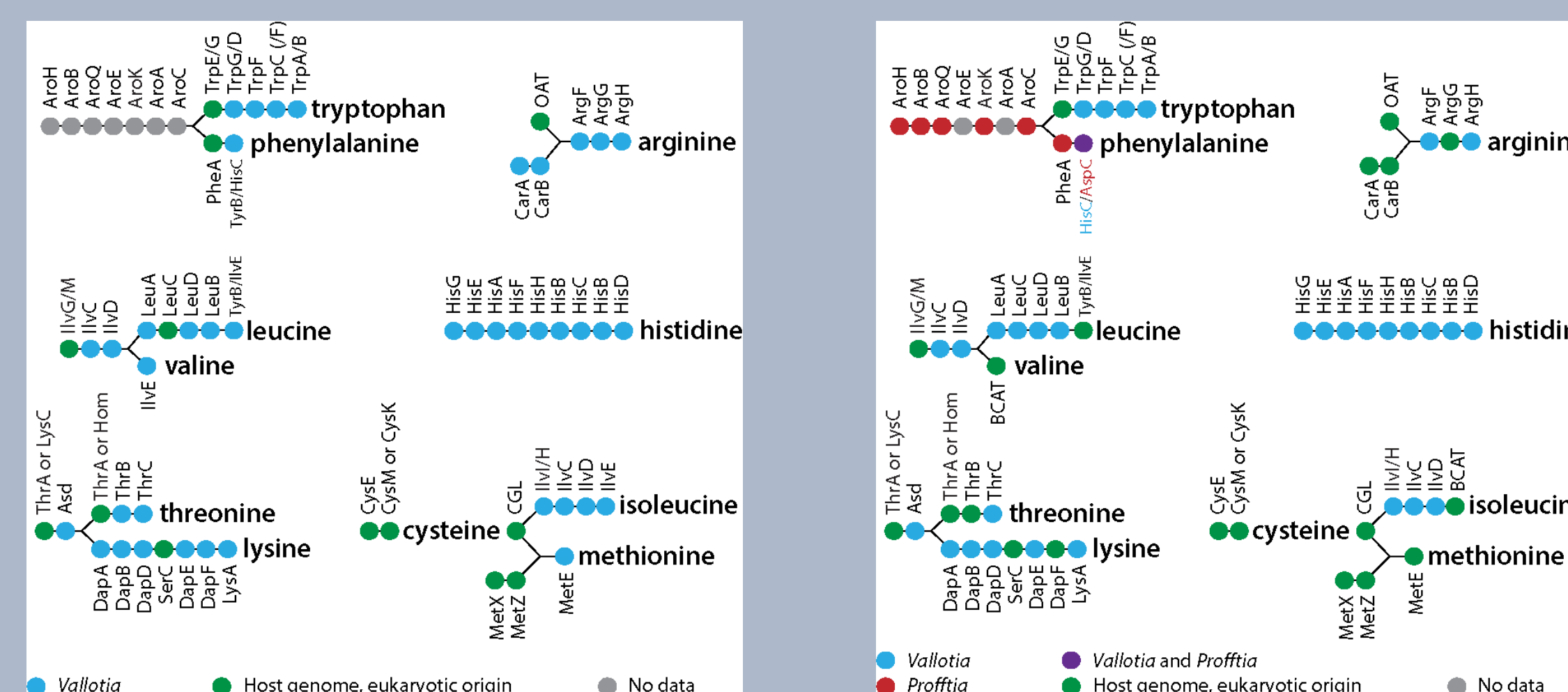


Fig. 4: Amino acid synthesis pathways of *A. abeitis* (left) and *A. lariciatus* (right). *Proffitia* is the secondary symbiont of both *A. abeitis* and *A. lariciatus*.

Lineage	Symbiont	Genome Size	GC	CDS	Coding density
Larch	' <i>Ca. Proffitia lariciata</i> '	1,106,964 bp	22.3%	778	64.3%
Larch	' <i>Ca. Vallotia lariciata</i> '	1,202,503 bp	42.4%	593	47.4%
Larch	' <i>Ca. Proffitia virida</i> '	N/A*	N/A*	N/A*	N/A*
Larch	' <i>Ca. Vallotia virida</i> '	1,193,586	43.0%	839	65.7%
Doug-fir	' <i>Ca. Gillettellia cooleya</i> ' PNW	687,415 bp**	34.0%**	329**	43.5%**
Doug-fir	' <i>Ca. Vallotia cooleya</i> ' PNW	1,157,315 bp	45.8%	862	65.3%
Doug-fir	' <i>Ca. Gillettellia cooleya</i> ' AZ	1,409,977 bp	34.0%	686	43.6%
Doug-fir	' <i>Ca. Vallotia cooleya</i> ' AZ	1,230,406 bp	45.5%	896	66.3%
Pine	' <i>Ca. Annandia pinicola</i> '	348,766 bp	18.7%	306	87.3%
Pine	' <i>Ca. Hartigia pinicola</i> '	2,035,178 bp	25.7%	699	33.0%
Hemlock	' <i>Ca. Annandia adelgestuga</i> '	334,746 bp	17.8%	285	89.3%
Hemlock	' <i>Ca. Pseudomonas adelgestugas</i> '	1,835,598 bp	39.5%	944	45.5%
Fir	' <i>Ca. Ecksteinia adelgicola</i> '	953,517 bp	19.78%	442	45.6%
Fir	' <i>Ca. Steffania adelgicola</i> '	1,369,640 bp	37.1%	825	52.1%

Fig. 5: Symbiont genome summary statistics for each adelgid lineage.

*Read data was too low quality to identify any amino acid synthesis genes
**Does not include entire genome due to low read quality

Conclusions

Vallotia is primarily responsible for synthesis of all essential amino acids except cysteine in *A. cooleyi*. *Gillettellia*, the primary symbiont in the Douglas fir lineage, works together with *Vallotia* in lysine and aromatic amino acid synthesis. In both larch lineage species, *Vallotia*, as the primary symbiont, plays a major role in the synthesis of most essential amino acids but is only responsible for the final steps in tryptophan synthesis and depends on *Proffitia* in *A. lariciatus* and probably *A. abeitis* for most steps in aromatic synthesis.

In *A. cooleyi* the primary symbiont *Gillettellia* is not solely responsible for synthesizing any amino acid. This suggests that *Gillettellia* has a degraded genome from millions of years of dependence on the adelgid host cells. The common ancestor of the Douglas fir and larch lineages likely contained *Gillettellia* and *Vallotia* as symbionts. The larch lineage probably dropped *Gillettellia* due to its gradually decreasing function in amino acid synthesis. Over time, the larch lineage *Vallotia* lost the ability to synthesize aromatics since neither *A. abeitis* nor *A. lariciatus* has *Vallotia* with genes for the early steps of aromatic synthesis. Thus, the larch lineage ancestor obtained *Proffitia* to fix this deficiency, as can be seen in *A. lariciatus*.

References

- Dial, D. T., Burke, G. R., Weglarz, K. M., Havill, N., & von Dohlen, C. D. (2019, June). *Differential division of labor in the ancient symbionts of sap-sucking insects (Sternorrhyncha: Adelgidae)*. Poster session presented at Gordon Research Conference: Animal-Microbe Symbioses as Nested Ecosystems, West Dover, VT.
- von Dohlen, C. D., Spaulding, U., Shields, K., Havill, N. P., Rosa, C., & Hoover, K. (2013). Diversity of proteobacterial endosymbionts in hemlock woolly adelgid (*Adelges tsugae*) (Hemiptera: Adelgidae) from its native and introduced range. *Environmental microbiology*, 15(7), 2043-2062.
- von Dohlen, C. D., Spaulding, U., Patch, K. B., Weglarz, K. M., Footitt, R. G., Havill, N. P., & Burke, G. R. (2017). Dynamic acquisition and loss of dual-obligate symbionts in the plant-sap-feeding Adelgidae (Hemiptera: sternorrhyncha: aphidoidea). *Frontiers in microbiology*, 8, 1037.
- Havill, N. P., & Footitt, R. G. (2007). Biology and evolution of Adelgidae. *Annu. Rev. Entomol.*, 52, 325-349.
- Weglarz, K. M., Havill, N. P., Burke, G. R., & von Dohlen, C. D. (2018).

Acknowledgments

I would like to thank Andrea Silletti and Dr. John Drake for organizing the Population Biology of Infectious Diseases REU, the National Science Foundation for funding the program, Ania Majewska for helping with R programming, and the Burke lab for providing training and guidance on this project.