



# Alien Index identifies potential horizontal gene transfer events in the *Selaginella moellendorffii* genome

Alexander T. Borowsky, Jennifer H. Wisecaver, and Antonis Rokas

Department of Biological Sciences, Vanderbilt University, Nashville, TN

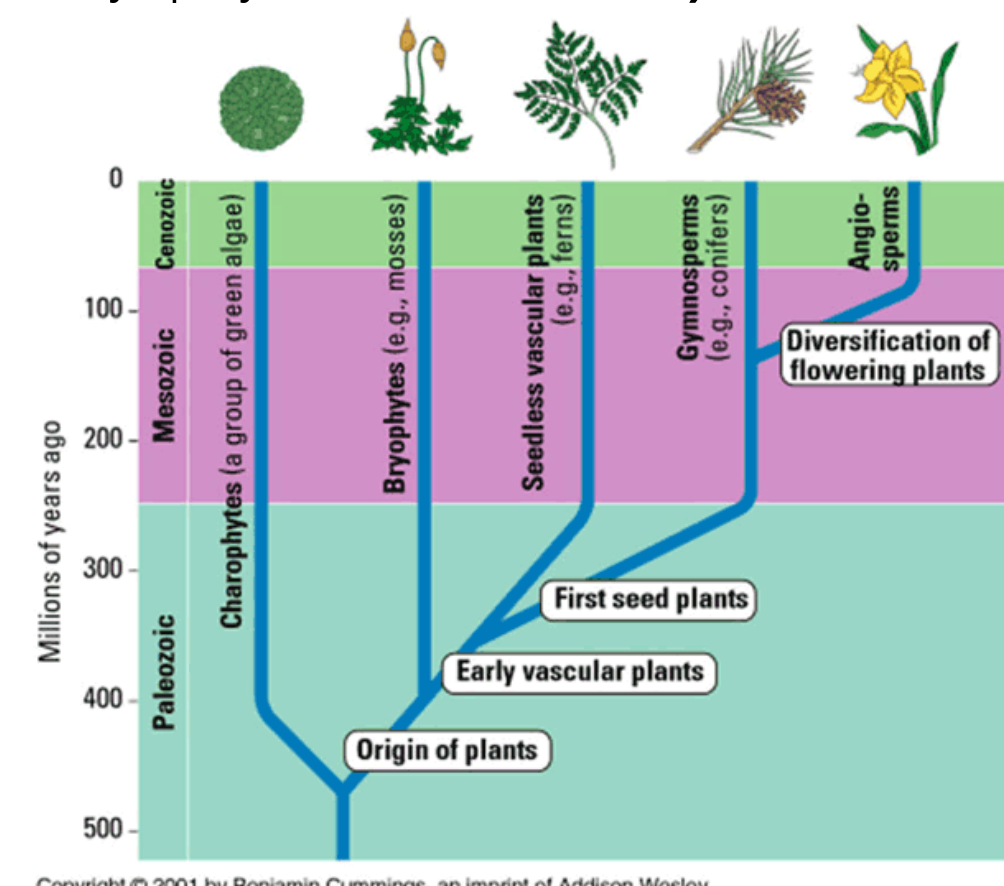
## Introduction

- Selaginella moellendorffii* is a vascular moss that sits on the cusp of a breakthrough in the evolution of plants, the development of seeds.

- The phenomenon of horizontal gene transfer (HGT), wherein DNA is passed across the tree of life, has been identified in *S. moellendorffii*: terpene synthase<sup>1</sup> and glycosyl hydrolase<sup>2</sup> genes from bacteria, and siderophore<sup>3</sup> genes from fungi.

- Modern genomics methods allow for a high-throughput assessment of whole genomes for HGT and other genomic intricacies.

**Figure 1. Evolutionary tree showing the major developments in major plant lineages** (<http://ohioplants.org/bryophytes-introduction/>)

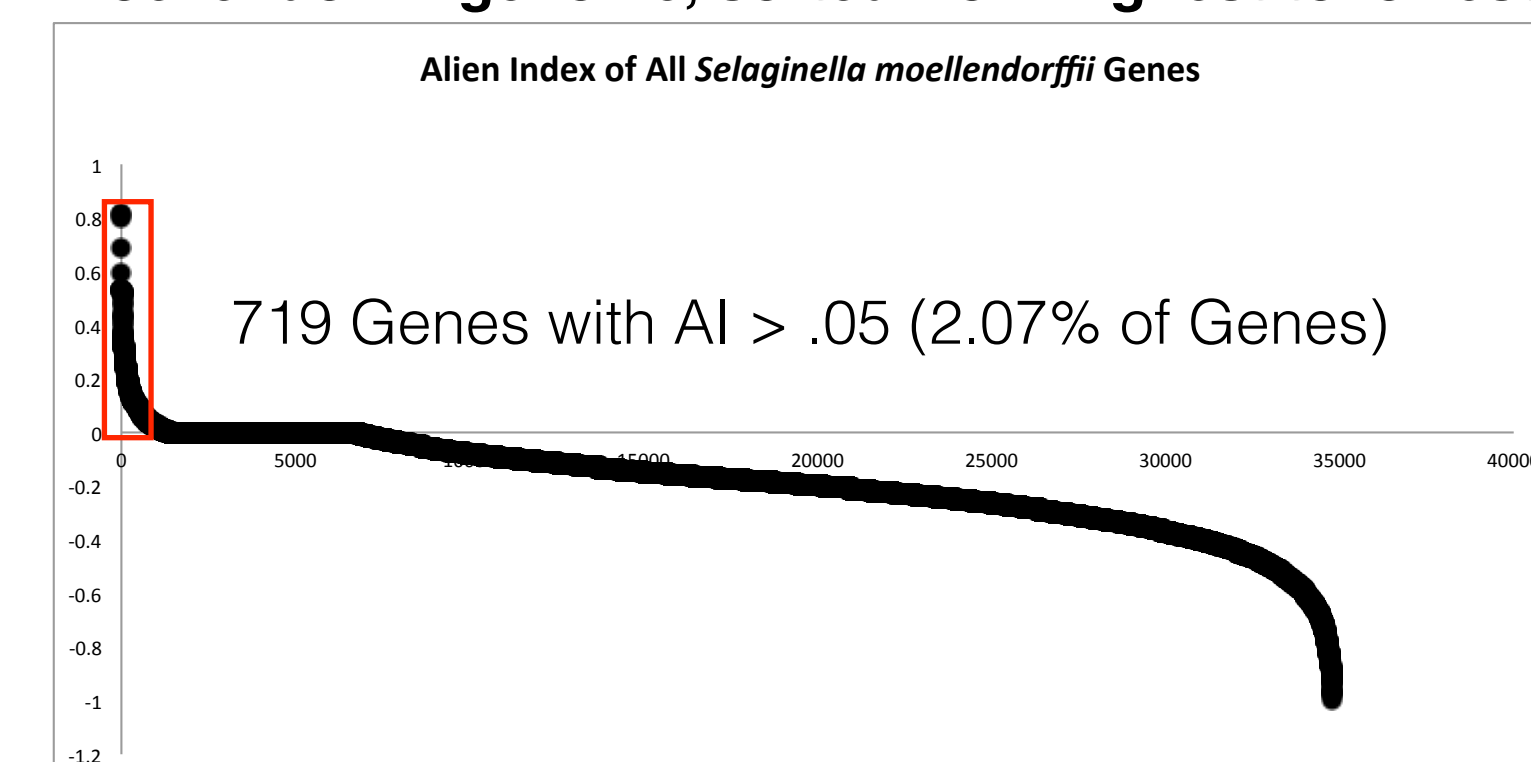


Copyright © 2001 by Benjamin Cummings, an imprint of Addison Wesley.

## Results

The AI screen recovered 719 genes in the *Selaginella* genome with AI > 0.05, as seen in Figure 2.

**Figure 2. AI Values for all genes in the *S. moellendorffii* genome, sorted from highest to lowest.**

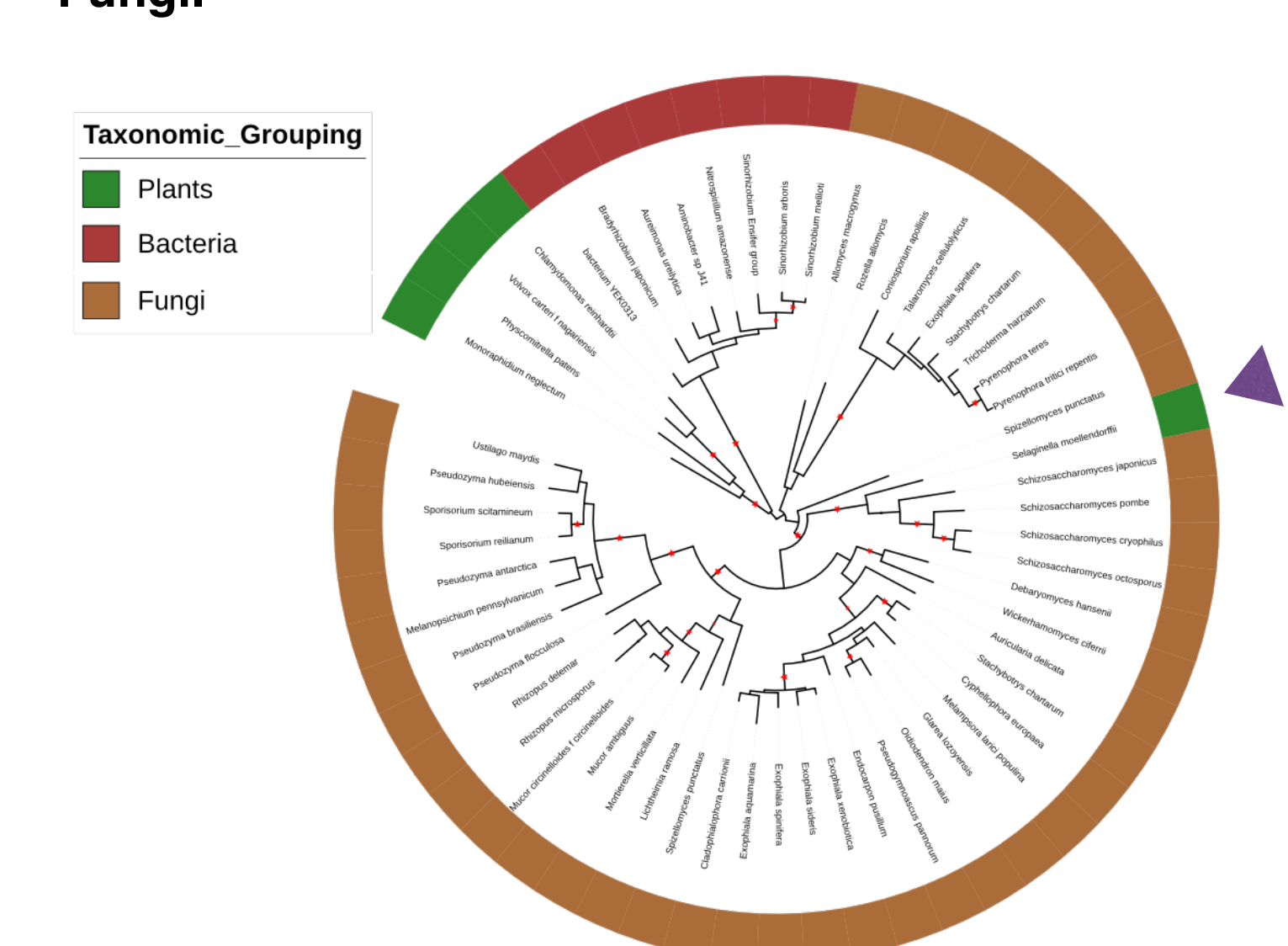


Contamination is ruled out by mapping the AI of genes to their positions on the contig. Genes adjacent to non-alien genes and not on the very edge of the contig are likely not due to contamination in the genome sequence, for example, the contig in Figure 3.

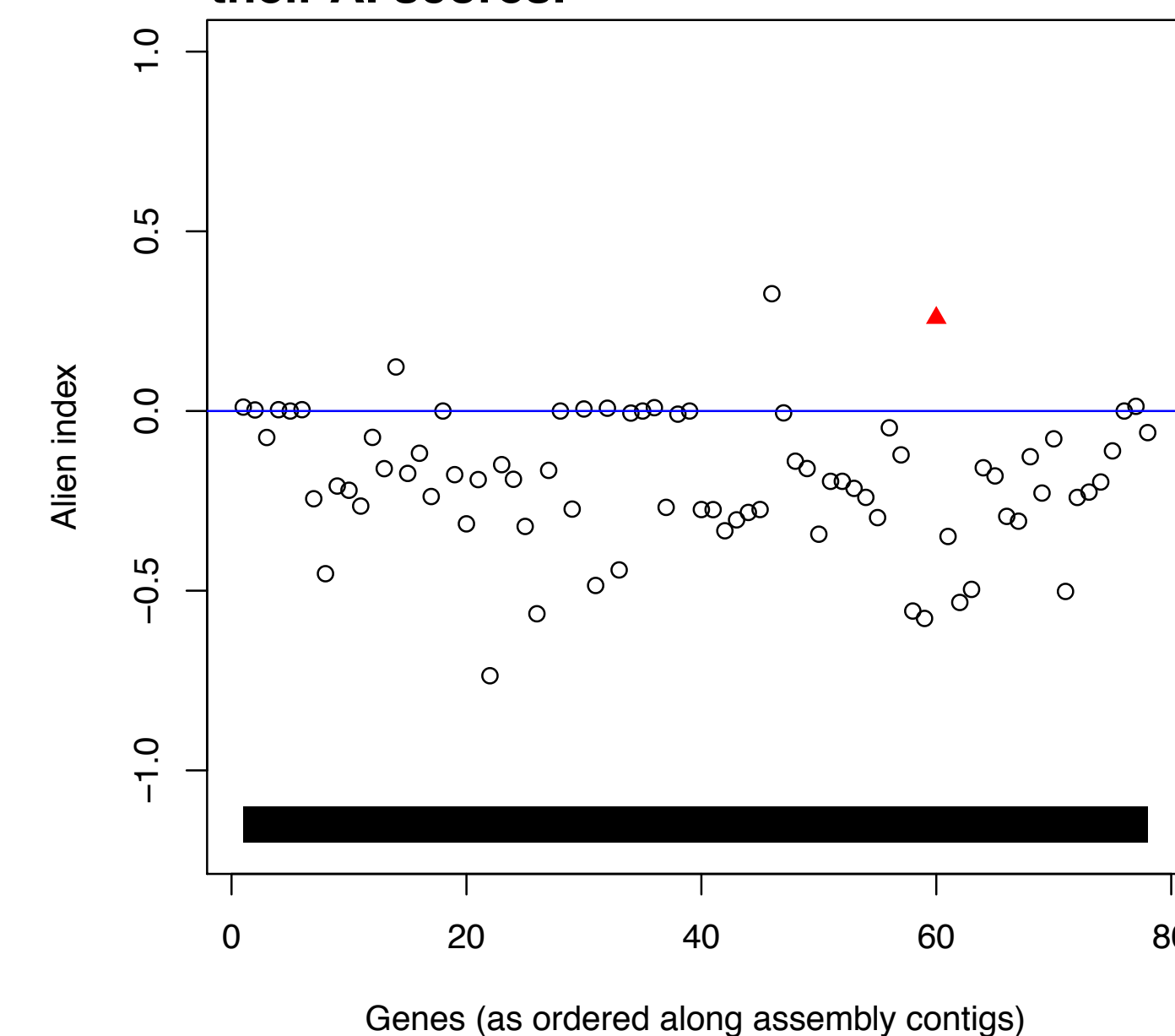
Previously reported HGT events were recovered in the AI screen, with high AI values as seen in Table 1.

Three genes showed fully supported HGT after SH test, two of which are shown Figures 4 and 5.

**Figure 4. RAxML generated tree for Dehydrogenase gene. Branches with bootstrap support > 95 are starred. The *Selaginella* Gene, indicated with a triangle, show strong support for grouping with Fungi.**



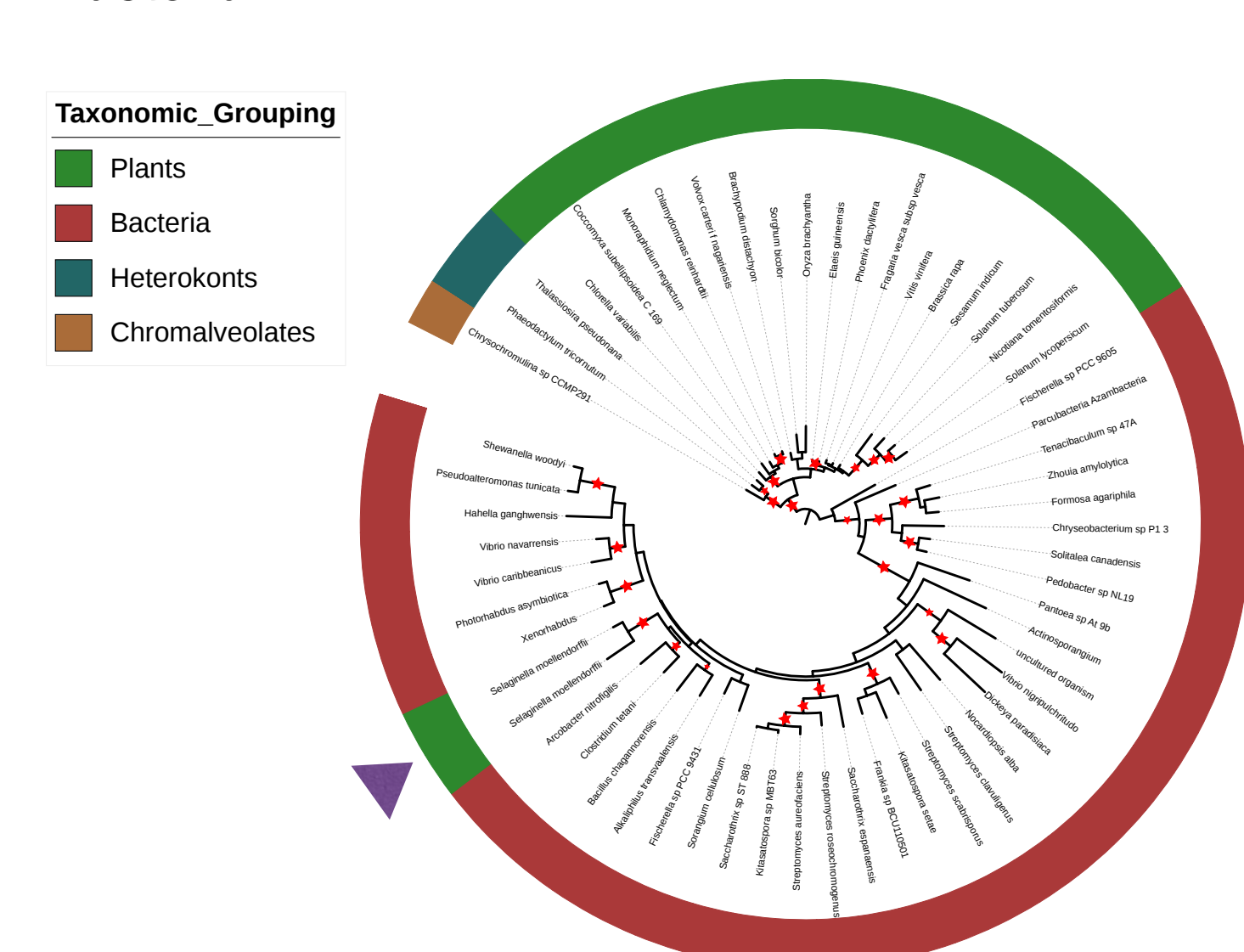
**Figure 3. Ordered genes from one contig from the *S. moellendorffii* genome, with their AI scores.**



**Table 1. Genes previously reported as potential HGT candidates that show high AI scores.**

Gene	Function	Publication	Alien Index
EFJ14547	Glycosyl Hydrolase	Metcalf et al. 2014	0.59
EFJ10553	Aegerolysin	Berne et al. 2009	0.41
EFJ20816	Siderophore	Richards et al. 2009	0.29
EFJ05410	Perforin	Hoang et al. 2009	0.25
EFJ26774	Terpene synthase	Li et al. 2012	0.20

**Figure 5. RAxML generated tree for Decarboxylase gene. Branches with bootstrap support > 95 are starred. The *Selaginella* Genes, indicated with a triangle, show strong support for grouping with Bacteria.**



## Discussion

- Some genes showed strong evidence of HGT, including those likely involved in secondary metabolism, a dehydrogenase and a decarboxylase. Secondary metabolism genes are often horizontally transferred in fungi.
- The activity of these genes in *S. moellendorffii* has only been bioinformatically predicted, not well characterized.
- The evolutionary history of the majority of alien genes was ambiguous.

## Future Work

- Determine more precise biological activity of horizontally transferred genes
- Refine phylogenetic pipeline to gather a wider distribution of sequences
- Address gene loss in seed plants by screening multiple moss lineages in additional AI screens.

## Acknowledgments

Thanks to the other members of the Rokas Lab for providing feedback on this research. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University. This work was partially supported by the National Science Foundation (<http://www.nsf.gov>) under Grants IOS-1401682 to J.H.W. and DEB-1442113 to A.R..

## References

- Li, G., Köllner, T. G., Yin, Y., Jiang, Y., Chen, H., Xu, Y., ... Chen, F. (2012). Nonseed plant *Selaginella moellendorffii* [corrected] has both seed plant and microbial types of terpene synthases. *Proceedings of the National Academy of Sciences of the United States of America*, 109(36), 14711–5. <http://doi.org/10.1073/pnas.1204300109>
- Metcalf, J. A., Funkhouser-Jones, L. J., Brileya, K., Reysenbach, A.-L., Bordenstein, S. R., Abascal, F., ... Doolittle, W. (2014). Antibacterial gene transfer across the tree of life. *eLife*, 3, 2104–2105. <http://doi.org/10.7554/eLife.04266>
- Richards, T. A., Soanes, D. M., Foster, P. G., Leonard, G., Thornton, C. R., & Talbot, N. J. (2009). Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *The Plant Cell*, 21(7), 1897–911. <http://doi.org/10.1105/tpc.109.065805>

## Objectives

- Conduct a high-throughput screen of the *S. moellendorffii* genome for genes that show HGT, gene loss in seed plants, or are moss-specific.
- Further explain the evolution of mosses and seed plants based on the genes identified.

## Methods

**Alien Index to identify candidates**  
AI is a value between 1 and -1 for a single gene, calculated from BLAST hit bit scores from a wide range of species. A positive score means that the gene is more similar to a gene outside of plants than any gene in plants other than mosses. A perl script was used to calculate this value for every gene in the *S. moellendorffii* genome.

**U-Search to identify homologues**  
For genes with an AI score > 0.05:  
• UCLUST to cluster very similar genes  
• U-Search each cluster against NR for 200 similar sequences from across the tree of life  
• CD-HIT to collapse highly similar sequences from the U-Search results

**Building trees to support HGT**  
• Align with MAFFT  
• Trim with TRIMAL  
• Build tree with FastTree and RAxML.  
• For those showing support for HGT, build constrained tree and perform SH Test with RAxML.