

# Comparative Genomic Analysis of *Talaromyces verruculosus*



Sarah Entwistle, Dr. Yanbin Yin

Department of Biological Sciences, College of Liberal Arts and Sciences, Northern Illinois University

Northern Illinois University

## Introduction

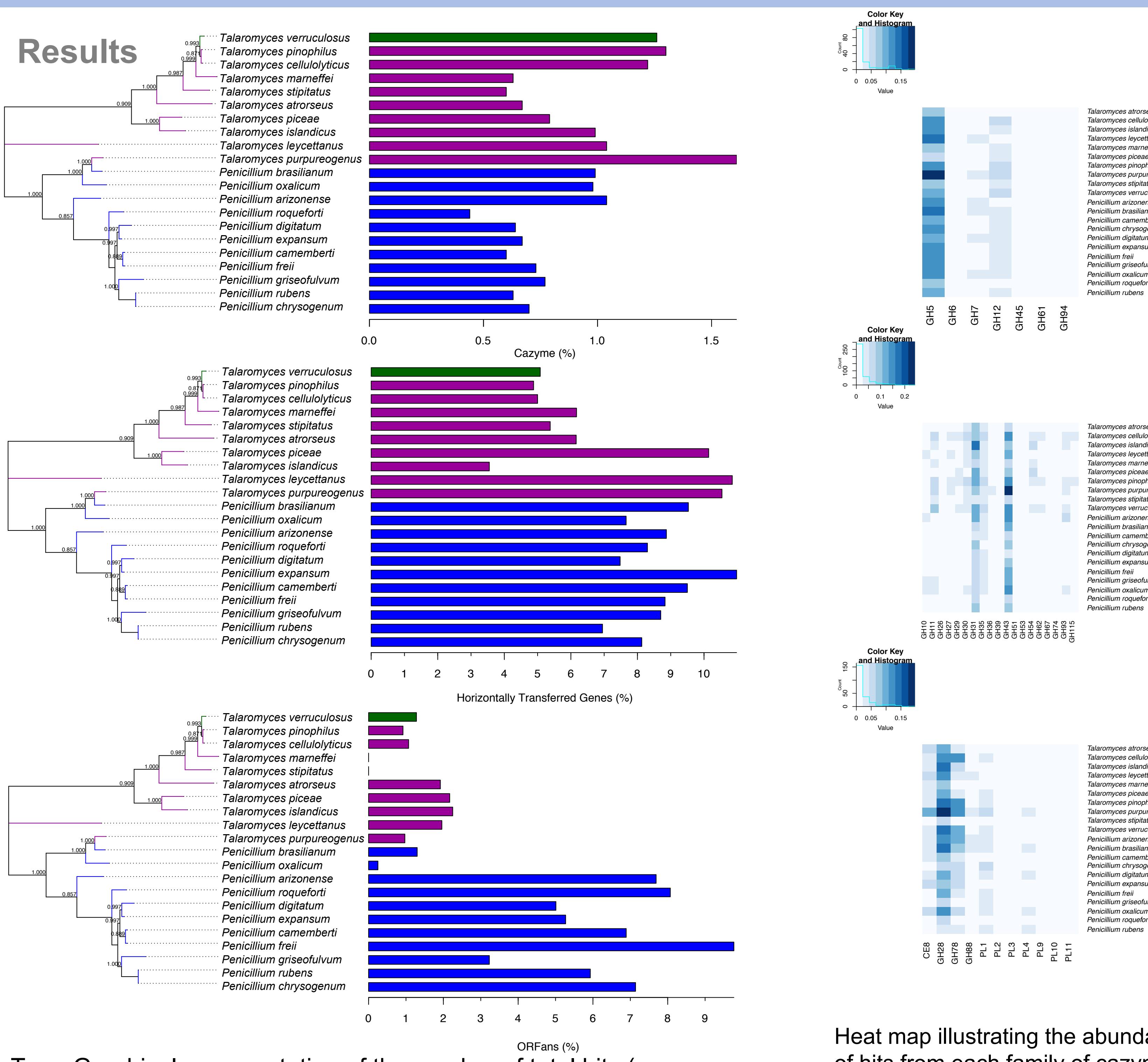
*Talaromyces verruculosus* is a lignocellulosic degrading fungi found primarily on tobacco plants. Lignocellulosic biomass is used in the production of biofuels and one of the most abundant raw materials available. *T. verruculosus* has a unique capability to degrade this biomass, therefore a better understanding of the processes involved would be beneficial to all.

A comparison of *T. verruculosus* along with other *Talaromyces* species, and selected *Penicillium* species is an excellent starting point to understanding more about this fungal species. Using computational methods, the phylogenetic relationship of all 21 species was determined. Then certain genomic sequence features of interest were selected for further examination.

- Carbohydrate enzyme (cazyme) classes
- Horizontally transferred genes (HTGs)
- ORFans – orphan genes not found in other species

## Materials & Methods

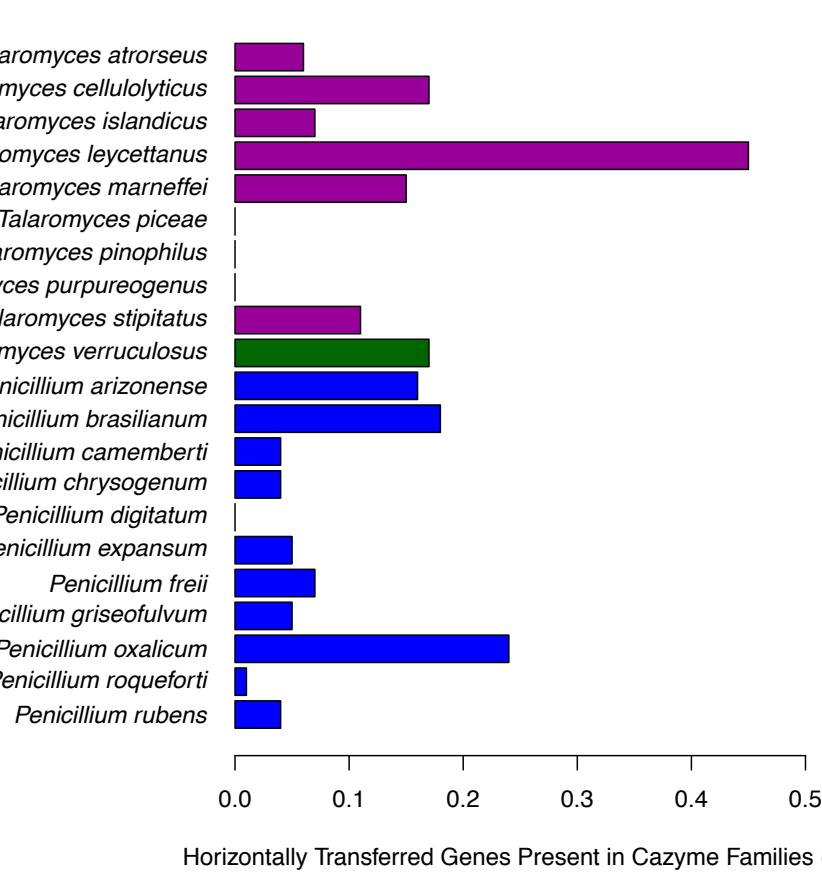
- All *Talaromyces* and *Penicillium* data gathered from NCBI
- Fungal database compiled from JGI
- MAKER – protein annotation
- HMMER – identify cazyme families
- MAFFT – multiple sequence alignment
- FastTree – infer phylogenetic tree
- iTOL – visualize tree
- DIAMOND – similar gene identification
- HGT-Finder – horizontally transferred gene prediction
- ORFanFinder – ORFan prediction
- R – data analysis and visualization



Top – Graphical representation of the number of total hits (as a percentage) from 37 identified cazyme families in each genome. Middle – Number (as a percentage) of horizontally transferred genes identified in each genome. Bottom – Number (as a percentage) of ORFans identified in each genome

## Results

To the right, genes both identified as horizontally transferred and found in one of the 37 cazyme families



## Conclusions

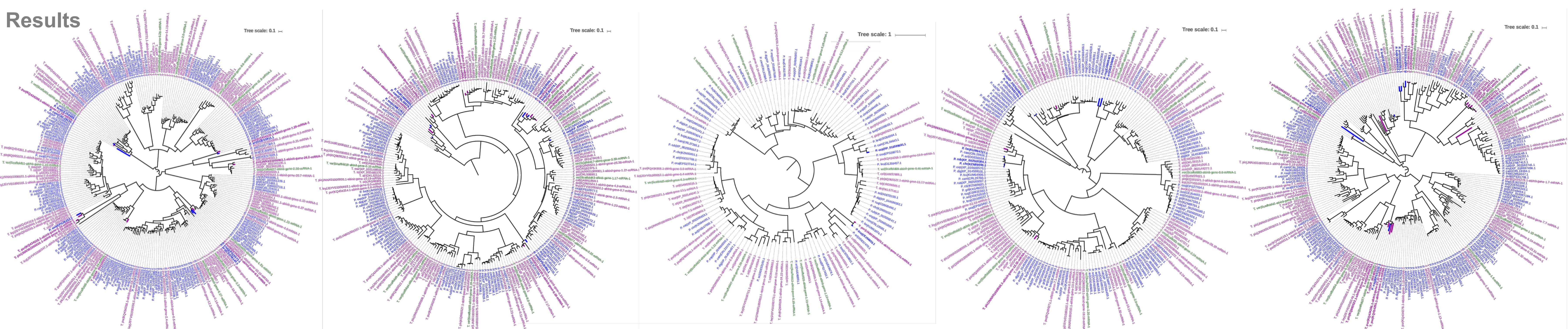
A phylogenetic analysis of *Talaromyces verruculosus* reveals this fungal species to be similarly related to other *Talaromyces* species. *T. verruculosus* shows a slightly higher percentage of cazymes (1.26%) than most of the other species in this analysis. This may begin to explain its lignocellulosic capabilities. More horizontally transferred genes were identified in the *Penicillium* species than the *Talaromyces* species. *T. verruculosus* has a median value (1.28%) in comparison to other *Talaromyces* species. Significantly more ORFans were identified in *Penicillium* species than in *Talaromyces* species, although once again *T. verruculosus* had a median value of 5.08% in comparison to other *Talaromyces* species.

Perhaps the more interesting discovery in this analysis is the determination that *T. purpureogenus* is an outlier among the other *Talaromyces* species. *T. purpureogenus* displays a significantly higher percentage of cazymes (1.61%) than all the other species. All three heat maps show the most abundant values belonging to *T. purpureogenus*. The percentage of horizontally identified genes is also more in accordance with *Penicillium* species as opposed to the other *Talaromyces* species. According to NCBI, this genome is unpublished. Many species of *Talaromyces* have been misclassified as *Penicillium* species. In fact, *T. verruculosus* was originally identified as a *Penicillium* species. *T. purpureogenus* has been rarely reported in a clinical setting and may be a similar case.

## Citations

- Hu, Liwei, Rahil Taujale, Fang Liu, Jizhen Song, Qisheng Yin, Yanling Zhang, Jianhua Guo, and Yanbin Yin. "Draft genome sequence of *Talaromyces verruculosus* ("*Penicillium verruculosum*") strain TS63-9, a fungus with great potential for industrial production of polysaccharide-degrading enzymes." *Journal of Biotechnology* 219 (2016): 5-6. Web.
- ORFanfinder - <http://cys.bios.niu.edu/orfanfinder/>
- HGT-Finder - <http://cys.bios.niu.edu/HGTFinder/>

## Results



Gene trees representing cazyme families of interest. From left to right, GH5 (cellulose degrading), GH28 (pectin degrading), GH78 (pectin degrading), GH31 (hemicellulose degrading), GH43 (hemicellulose degrading). Genes highlighted in bold are identified as horizontally transferred.