



## Diversity and Phylogenetic Distribution of Extracellular Microbial Peptidases

Trang Nguyen (1), Ryan Mueller (2), and David Myrold (1)

(1) Oregon State University, Department of Crop and Soil Science, Corvallis, Oregon, United States

(trang.nguyen@oregonstate.edu), (2) Oregon State University, Department of Microbiology, Corvallis, Oregon, United States

Depolymerization of proteinaceous compounds by extracellular proteolytic enzymes is a bottleneck in the nitrogen cycle, limiting the rate of the nitrogen turnover in soils. Protein degradation is accomplished by a diverse range of extracellular (secreted) peptidases. Our objective was to better understand the evolution of these enzymes and how their functional diversity corresponds to known phylogenetic diversity. Peptidase subfamilies from 110 archaeal, 1,860 bacterial, and 97 fungal genomes were extracted from the MEROPS database along with corresponding SSU sequences for each genome from the SILVA database, resulting in 43,177 secreted peptidases belonging to 34 microbial phyla and 149 peptidase subfamilies. We compared the distribution of each peptidase subfamily across all taxa to the phylogenetic relationships of these organisms based on their SSU gene sequences. The occurrence and abundance of genes coding for secreted peptidases varied across microbial taxa, distinguishing the peptidase complement of the three microbial kingdoms. Bacteria had the highest frequency of secreted peptidase coding genes per 1,000 genes and contributed from 1% to 6% of the gene content. Fungi only had a slightly higher number of secreted peptidase gene content than archaea, standardized by the total genes. The relative abundance profiles of secreted peptidases in each microbial kingdom also varied, in which aspartic family was found to be the greatest in fungi (25%), whereas it was only 12% in archaea and 4% in bacteria. Serine, metallo, and cysteine families consistently contributed widely up to 75% of the secreted peptidase abundance across the three kingdoms. Overall, bacteria had a much wider collection of secreted peptidases, whereas fungi and archaea shared most of their secreted peptidase families. Principle coordinate analysis of the peptidase subfamily-based dissimilarities showed distinguishable clusters for different groups of microorganisms. The distribution of secreted peptidases was found to be significantly correlated with phylogenetic relationships within kingdoms (archaea  $r_{\text{Mantel}}=0.364$ ,  $p=0.001$ ; bacteria  $r_{\text{Mantel}}=0.257$ ,  $p=0.001$ , and fungi  $r_{\text{Mantel}}=0.281$ ,  $p=0.005$ ), inferring an evolutionary relationship where subsets of phylogenetically related organisms share similar types of secreted peptidases. We also tested the phylogenetic signal strength of each peptidase subfamily for each microbial kingdom based on the binary traits of the distribution (presence or absence of secreted peptidase subfamilies in individual species). About one-third of the peptidase subfamilies displayed a strong evolutionary signal; the rest were phylogenetically overdispersed, suggesting that these subfamilies are randomly distributed across the tree of life or the result of events such as horizontal gene transfer. Study of the diversity and phylogenetic distribution of secreted peptidases offered a mechanistic basis to anticipate the proteolytic potential function of microbial communities.