

***EvoNet*: A Phylogenomic and Systems Biology approach to identify genes underlying plant survival in marginal, low-Nitrogen soils**

Gil Eshel,^{1*} (ge30@nyu.edu), Kranthi Varala,² Rodrigo Gutiérrez,³ Dennis Stevenson,⁴ Robert DeSalle,⁵ W. Richard McCombie,⁶ Jean-Michel Ané,⁷ Heidi Kaepler⁷ and **Gloria Coruzzi**¹

¹New York University, New York; ²Purdue University, West Lafayette; ³Pontificia Universidad Católica de Chile, Santiago, Chile; ⁴New York Botanical Garden, New York; ⁵American Museum of Natural History, New York; ⁶ Cold Spring Harbor Laboratory, New York and ⁷University of Wisconsin, Madison

<http://evonet.org>

Project Goals: This DOE BER sustainability project aims to identify the key genes and gene regulatory networks that enable “extreme survivor” plants to adapt and grow in marginal, nitrogen (N)-poor soils in the hyper-arid Atacama Desert in Chile. These “extreme survivor” species cover the main branches in flowering plants and include 7 grass species of particular interest for biofuels. We focus on 28 “extreme survivor” Atacama species and compare their encoded genes to Californian “sister” species that live in a N-replete conditions in arid (27 species) or mesic (27 species) environments. Deep RNA-sequencing of these “triplet species” was used to fuel a phylogenomic analysis that can identify genes that support the evolutionary divergence of the extreme survivors in Atacama Desert from their sister species in California. The genes thus identified will help to discover the mechanisms underlying physiological and developmental processes that allow plant survival in nitrogen-poor, dry soils. The genes and network modules so uncovered can potentially be translated to biofuel crops to greatly increase biomass and nitrogen use efficiency in marginal, low-fertility soils.

This collaborative project exploits the genomes of “extreme survivor” plants adapted to thrive in marginal, Nitrogen (N) poor soils in the hyper-arid Atacama Desert in the Chilean Andes. It uses a previously validated phylogenomic pipeline we developed called PhyloGeneious [1], which can identify genes that provide support to species divergence. By applying this phylogenomic pipeline to the gene sequences of “triplet species”, we can identify the genes that distinguish these “extreme survivors” in Atacama from their related “sister” species adapted to similarly dry regions in California (CA) not constrained by N and/or water availability and to mesic “sister” species growing in N and water (W) replete conditions. These “extreme survivor” species from the Atacama desert broadly cover the main branches in flowering plants, and therefore offer a wide range of genomic backgrounds within which the survival traits repeatedly arose i.e., multiple independent origins of trait.

To maximize our ability to separate the trait-relevant signature from overall speciation events, our “triplet species” sampling covers multiple independent origins of the low-N adaptive trait. In published studies, we showed that our phylogenomic pipeline could; i) identify genes that underlie convergent evolution of antioxidant synthesis in Rosids in a study of 150 plant genomes [1]; and ii) identify 100+ genes associated with the loss of Arbuscular Mycorrhizal (AM) symbiosis in the *Brassicaceae* [2]. We further extend this phylogenomic approach to the study of “extreme survivor strategies” as follows:

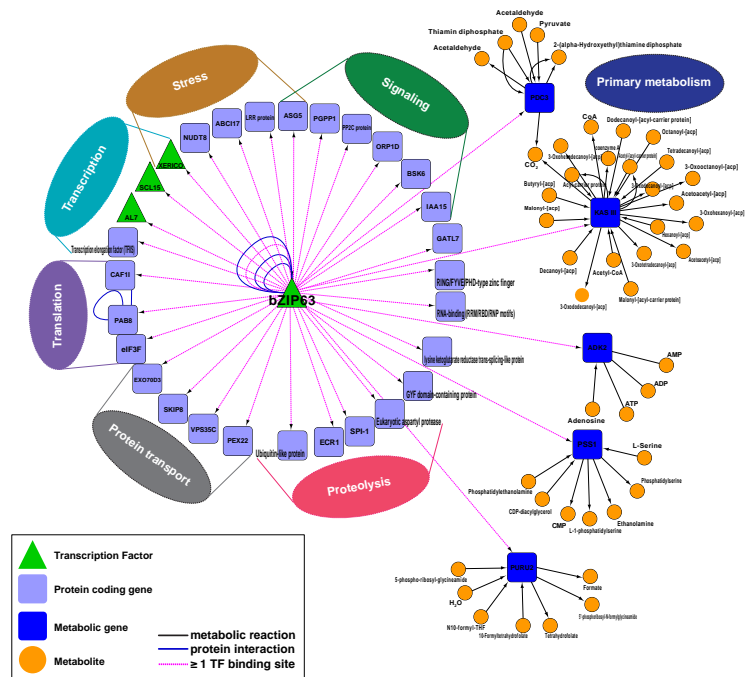
Aim 1. Species collection and deep transcriptome sequencing: (NYU, NYBG, Chile). **Progress:** We sequenced all 28 species collected in the Atacama Desert and 50 of their CA “sister” species, with a resulting average gene coverage of 87%, based on the BUSCO single-copy orthologs assessment.

Aim 2. Phylogenomic Analysis: Perform phylogenomic analysis of 82 “triplet species” to identify genes that repeatedly support nodes that distinguish the “extreme survivors” in the Atacama Desert from their sister species in CA (AMNH, NYU). **Progress:** We performed phylogenomic analysis that includes each of the major plant lineages (*Poaceae* - 13 taxa; *Caryophyllales* – 10 taxa; *Lamiids* – 12 taxa; *Campanulids* – 17 taxa; *Fabaceae* – 12 taxa). This analysis identified 1022 genes, with a strong exclusive signal of positive selection in multiple Atacama “extreme survivors” species.

Aim 3. Network Analysis: We used the Arabidopsis gene-network tool from VirtualPlant (virtualplant.bio.nyu.edu), and revealed a sub-network of 35 positively selected genes, that are connected through a transcription factor (TF) hub, bZIP63 (Figure 1). This transcription factor is known to regulate primary metabolism in order to boost cellular energy in response to starvation, as part of the SnRK1-TOR signaling pathway [3]. We are currently using the *TARGET* system, developed in our lab, to validate the targets of bZIP63, to complement our in planta studies in Aim 4.

Aim 4. Functional Validation: To functionally validate top-ranked candidate genes for low-N adaptation in Arabidopsis and Brachypodium (NYU, Chile, U Wisconsin). **Progress:** We have begun to transform Brachypodium with the 28 most promising candidates from our phylogenetic analysis using our Atacama set and their closest sequence available sister species.

Figure 1. Sub-network of 35 positively selected genes found exclusively in multiple Atacama “extreme survivors” species. Based on the Arabidopsis orthologous genes, we found a set of positively selected genes in Atacama that are connected through the transcription factor bZIP63, known to regulate primary metabolism in order to boost cellular energy in response to starvation, as part of the SnRK1-TOR signaling pathway [3].



References

1. Lee E et al. (2011) A functional phylogenomics view of the seed plants. *PLoS Genet* 7(12):e1002411.
2. Delaux PM et al. (2014) Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. *PloS Genet* 10(7):e1004487.
3. Mair A et al. (2015) SnRK1-triggered switch of bZIP63 dimerization mediates the low-energy response in plants. *eLife*. 4:e05828.

The EvoNet grant is supported by the Genomics Science program within the Office of Biological and Environment Research in the DOE Office of Science.