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

Kranthi Varala,[1]* (kranthi.varala@nyu.edu), Rodrigo Gutiérrez,[2] Dennis Stevenson,[3] Robert DeSalle,[4] W. Richard McCombie,[5] Jean-Michel Ané,[6] Heidi Kaeppler[6] and **Gloria Coruzzi**[1]

[1]New York University, New York; [2]Pontifica Universidad Católica de Chile, Santiago, Chile; [3]New York Botanical Garden, New York; [4]American Museum of Natural History, New York; [5] Cold Spring Harbor Laboratory, New York and [6]University 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, extremely nitrogen (N) -poor soils in the arid Chilean Andes. These extreme survivor species cover the main branches in flowering plants and include 7 grass species of particular interest for biofuels. We focus on 24 "extreme survivor" Chilean species that are relevant to biofuels and compare their genomes to 24 Californian "sister" species that live in a N replete arid environment. Exploiting a novel phylogenomic pipeline, a "paired species" sampling strategy will identify the genes that support the evolutionary divergence of the extreme survivors in Chile from their sister species in California. The genes thus identified will help to discover the mechanisms underlying physiological and developmental processes that allow efficient assimilation of nitrogen 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, extremely Nitrogen (N) poor soils in the arid Chilean Andes. It uses a previously validated phylogenomic pipeline we developed called BigPlant [1], and a "paired species" sampling strategy, to identify the genes that distinguish these "extreme survivors" in Chile from their related species adapted to similarly dry regions in California (CA) that are not constrained by N. These "extreme survivor" species 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.

Key to our phylogenomic approach is the "paired species" sampling strategy. This will identify the genes responsible for the divergence of the "extreme survivors" adapted to the marginal low-N soils of Chile, from their most closely related species growing on arid but N-replete soils in CA. To maximize our ability to separate the trait-relevant signature from overall speciation events, our "paired species" sampling will cover multiple independent origins of the low-N adaptive trait. In published studies, we showed that this phylogenomic approach could; i) identify genes that underlie convergent evolution of antioxidant synthesis in Rosids in a study of 150 genomes [1]; and ii) identify 100+ genes associated with the loss of Arbuscular Mycorrhizal

Chile Species (Drought + low-N)	California Species (Drought)
Acantholippia deserticola	*Aloysia wrightii*
Adesmia spinosissima	*Amorpha californica*
Allionia incarnata	*Anulocaulis annulatus*
Ambrosia artemisioides	*Dicoria canescens*
Aristida adscensionis	*Danthonia unispicata*
Bouteloua simplex	*Hilaria jamesii*
Calandrinia sp.	*Montia dichotoma*
Chorizanthe conmisuralis	*Pterostegia drymarioides*
Cistanthe sp.	*Calptridium monospermum*
Cristaria sp.	*Malvella leprosa*
Deyeuxia curvula	*Calamagrostis rubescens*
Euphorbia amandi	*Chamaesyce vallis-mortae*
Festuca sp (chrysophylla)	*Vulpia microstachys*
Junellia seriphioides	*Glandularia gooddingii*
Munroa decumbens	*Dasychloa pulchella*
Nassella nardoides	*Achnantherum occidentale*
Nototriche sps.	*Sphaeralcea rusbyi*
Senecio puchii	*Pluchea sdericea*
Sisymbrium sp.	*Stanleya pinnata*
Stipa frigida	*Jarva speciosa*
Tagetes multiflora	*Pectis papposa*
Tarasa operculata	*Malacothamnus fremontii*
Urbania pappigera	*Verbena menthifolia*
Werneria sp.	*Lepidospartum squamatum*

Table 1. Extreme survivor species in Chile (green) and paired "sister" species in CA (yellow). *Our project studies 24 pairs of species from Marginal (Dry +low-N, Chile) and Dry soils (Dry, California). All the Chilean species have already been sequenced while California species are being collected.*

(AM) symbiosis in the Brassicaceae [2]. We now extend this approach to the study of "Marginal survivor strategies" as follows:

Aim 1. Sample transcriptomes of 24 "extreme survivors" & paired species (NYU, NYBG, Chile). ***Progress:*** We sequenced all 24 species from Chile (Table 1) and are currently sampling the sister species collected in the field in California.

Aim 2. Phylogenomic analysis of 48 "paired species" to identify genes that repeatedly support nodes that distinguish the extreme survivors in Chile from their sister species in CA (AMNH, NYU). ***Progress:*** We have adapted our BigPlant phylogenomic pipeline to improve speed, while reducing the memory footprint.

Aim 3. Combine phylogenomics (protein sequence) and gene networks (gene expression) to identify genes and network modules associated with adaptations to marginal, low-N soils (NYU, Chile). ***Progress:*** To exploit a comparative analysis of gene regulatory networks, we are developing a new module **PhyloExpress** that extends the BigPlant pipeline to include gene expression data.

Aim 4. Functionally validate top-ranked candidate genes for low-N adaptation in Arabidopsis and Brachypodium (NYU, Chile, U Wisconsin). ***Progress:*** We are transforming Brachypodium with the most promising candidate from our preliminary analysis of 5 species pairs.

References

1. Lee E et. al., A functional phylogenomics view of the seed plants. PLoS Genet 7(12):e1002411.
2. Delaux et. Al., Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. PloS Genet 10(7):e1004487.