

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

Gil Eshel,<sup>1\*</sup> ([ge30@nyu.edu](mailto:ge30@nyu.edu)), Kranthi Varala,<sup>2</sup> Rodrigo Gutiérrez,<sup>3</sup> Dennis Stevenson,<sup>4</sup> Robert DeSalle,<sup>5</sup> W. Richard McCombie,<sup>6</sup> Jean-Michel Ané,<sup>7</sup> Heidi Kaeppeler<sup>7</sup> and **Gloria Coruzzi**<sup>1</sup>

<sup>1</sup>New York University, New York; <sup>2</sup>Purdue University, West Lafayette; <sup>3</sup>Pontificia Universidad Católica de Chile, Santiago, Chile; <sup>4</sup>New York Botanical Garden, New York; <sup>5</sup>American Museum of Natural History, New York; <sup>6</sup> Cold Spring Harbor Laboratory, New York and <sup>7</sup>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 28 “extreme survivor” Chilean species and compare their genomes 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 “triplicate species” will be used to fuel a novel phylogenomic analysis that 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 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, extremely Nitrogen (N) poor soils in the arid Chilean Andes. It uses a previously validated phylogenomic pipeline we developed called PhyloGeneious [1], which can identify genes that provide positive support to species divergence. By applying this phylogenomic pipeline to the gene sequences of these “triplicate species”, we can identify the genes that distinguish these “extreme survivors” in Chile from their related species adapted to similarly dry regions in California (CA) not constrained by N and/or water availability. 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 “triplicate species” sampling strategy. To maximize our ability to separate the trait-relevant signature from overall speciation events, our “triplicate species” sampling will cover 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 now extend this phylogenomic approach to the study of “marginal survivor strategies” as follows:

**Aim 1. Species collection and deep transcriptome sequencing:** Sample deep transcriptomes of 28 triplicate sister species pairs. (Chile “marginal survivors” and their closest relatives from CA) (NYU, NYBG, Chile). **Progress:** We performed deep RNA sequencing on all 28 Chilean species and 39 of their CA relatives (Table 1), with resulting average gene coverage of 87%, based on the BUSCO conserved single-copy orthologs existence assessment.

**Aim 2. Phylogenomic Analysis:** Phylogenomic analysis of 82 “triplicate 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 optimized our PhyloGeneious phylogenomic pipeline for volume and speed to enable the analysis of this large taxa set (90 species, including the outgroups and the model plant species for functional annotations).

**Aim 3. Network Analysis:** 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 currently developing a new module called **PhyloExpress** that extends the PhyloGeneious pipeline to include gene expression data.

**Aim 4. Functional Validation:** 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 most promising candidates from our preliminary analysis using our Chilean set and their closest sequence available sister species.

**Table 1. Extreme survivor species in Chile (green) and their “sister” species in CA (Drought - yellow; Mesic - blue).** Our project studies 28 triplicates of species from Marginal (Dry +low-N, Chile), Dry (Dry, California), and moist (Mesic, California) soils. All the Chilean species and 39/54 of the California species have already been sequenced while 15/54 California species are being collected.

| Chile Marginal Survivor Species (Drought + low-N) | California Species (Drought )  | California Species (Mesic)        |
|---|--------------------------------|-----------------------------------|
| <i>Atriplex imbricata</i>                         | <i>Atriplex lentiformis</i>    | <i>Atriplex watsonii</i>          |
| <i>Mulinum crassifolium</i>                       | <i>Sanicula crassicaulis</i>   | <i>Conium maculatum</i>           |
| <i>Ambrosia artemisioides</i>                     | <i>Ambrosia chamissonis</i>    | <i>Ambrosia psilostachya</i>      |
| <i>Baccharis boliviensis and Baccharis tola</i>   | <i>Baccharis glutinosa</i>     | <i>Baccharis salicifolia</i>      |
| <i>Trichoclina caulescens</i>                     | <i>Cirsium occidentale</i>     | <i>Cirsium fontinale</i>          |
| <i>Chuquiraga atacamensis</i>                     | <i>Coreopsis douglasii</i>     | <i>Rudbeckia californica</i>      |
| <i>Parastrephia quadrangularis</i>                | <i>Grindelia hirsutula</i>     | <i>Eriophyllum confertiflorum</i> |
| <i>Senecio puchii</i>                             | <i>Senecio californicus</i>    | <i>Senecio mikanioides</i>        |
| <i>Tagetes multiflora</i>                         | <i>Pectis papposa</i>          | <i>Pluchea sericea</i>            |
| <i>Phacelia pinnatifida</i>                       | <i>Eriodictyon tomentosum</i>  | <i>Phacelia nemoralis</i>         |
| <i>Pycnophyllum bryoides</i>                      | <i>Cerastium viride</i>        | <i>Cerastium beeringianum</i>     |
| <i>Adesmia spinosissima</i>                       | <i>Amorpha californica</i>     | <i>Amorpha fruticosa</i>          |
| <i>Lupinus oreophilus</i>                         | <i>Lupinus nana</i>            | <i>Lupinus arboreus</i>           |
| <i>Lupinus subinflatus</i>                        | <i>Lupinus hirsutissimus</i>   | <i>Lupinus latifolius</i>         |
| <i>Aristida adscensionis</i>                      | <i>Danthonia unispicata</i>    | <i>Danthonia californica</i>      |
| <i>Bouteloua simplex</i>                          | <i>Bouteloua curtipendula</i>  | <i>Muhlenbergia filiformis</i>    |
| <i>Calamagrostis crispata</i>                     | <i>Calamagrostis rubescens</i> | <i>Calamagrostis breweri</i>      |
| <i>Calamagrostis cabreriae</i>                    | <i>Festuca californica</i>     | <i>Festuca subuliflora</i>        |
| <i>Munroa decumbens</i>                           | <i>Munroa squarrosa</i>        | <i>Munroa utilis</i>              |
| <i>Nassella nardoides</i>                         | <i>Nassella cernua</i>         | <i>Nassella manicata</i>          |
| <i>Jarava frigida</i>                             | <i>Stipa coronata</i>          | <i>Stipa kingii</i>               |
| <i>Chorizanthe commisuralis</i>                   | <i>Chorizanthe palmeri</i>     | <i>Rumex crispus</i>              |
| <i>Exodeconus integrifolius</i>                   | <i>Lycium cooperi</i>          | <i>Physalis lancifolia</i>        |
| <i>Fabiana denudata</i>                           | <i>Nicotiana glauca</i>        | <i>Petunia parviflora</i>         |
| <i>Solanum chilense</i>                           | <i>Datura wrightii</i>         | <i>Solanum douglasii</i>          |
| <i>Acantholippia deserticola</i>                  | <i>Aloysia wrightii</i>        | <i>Phyla nodiflora</i>            |
| <i>Fagonia chilensis</i>                          | <i>Fagonia laevis</i>          | <i>Tribulus terrestris</i>        |

## 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.

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