



# Machine Learning for Identifying Plant- Microbiome Interactions

Anastasiia Kim (CCS-3) , Eric Moore (B-11),  
Nicholas Lubbers (CCS-3), Sanna Sevanto (EES-14)

09/14/2022

## OFFICIAL USE ONLY

May be exempt from public release under the Freedom of Information Act  
(5 U.S.C. 522), exemption number and category: \_\_\_\_\_

Department of Energy review required before public release

Name/Org: \_\_\_\_\_ Date: \_\_\_\_\_

Guidance (if applicable): \_\_\_\_\_

LA-CP-XX-XXXXX

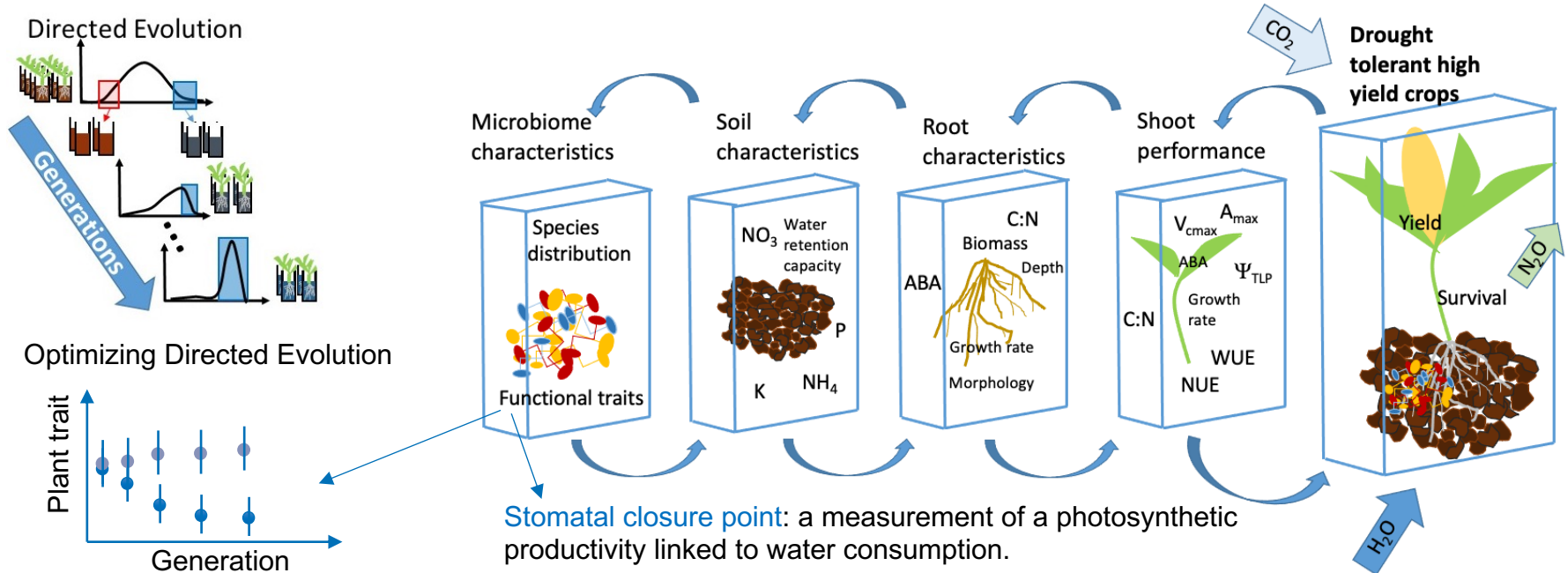


Managed by Triad National Security, LLC, for the U.S. Department of Energy's NNSA.

OFFICIAL USE ONLY

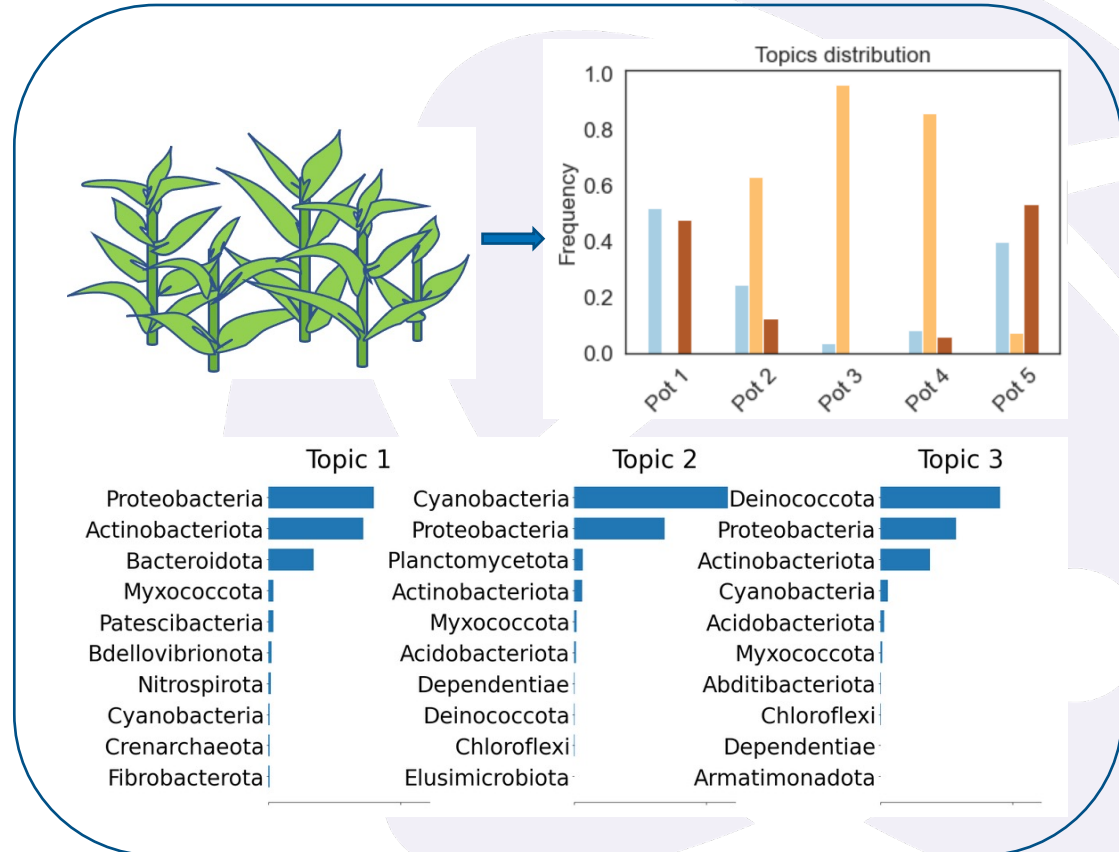
# Improving plant drought tolerance is essential for matching the future food and biofuel needs

**Goal:** find the links between microbiome and experimental setup variables and plant traits and chemistry.



# Probabilistic topic modeling

- Analyzed microbiomes using Latent Dirichlet Allocation, probabilistic generative model developed for language modeling of a set of documents
- Found strong associations between topics and experimental conditions and plant traits



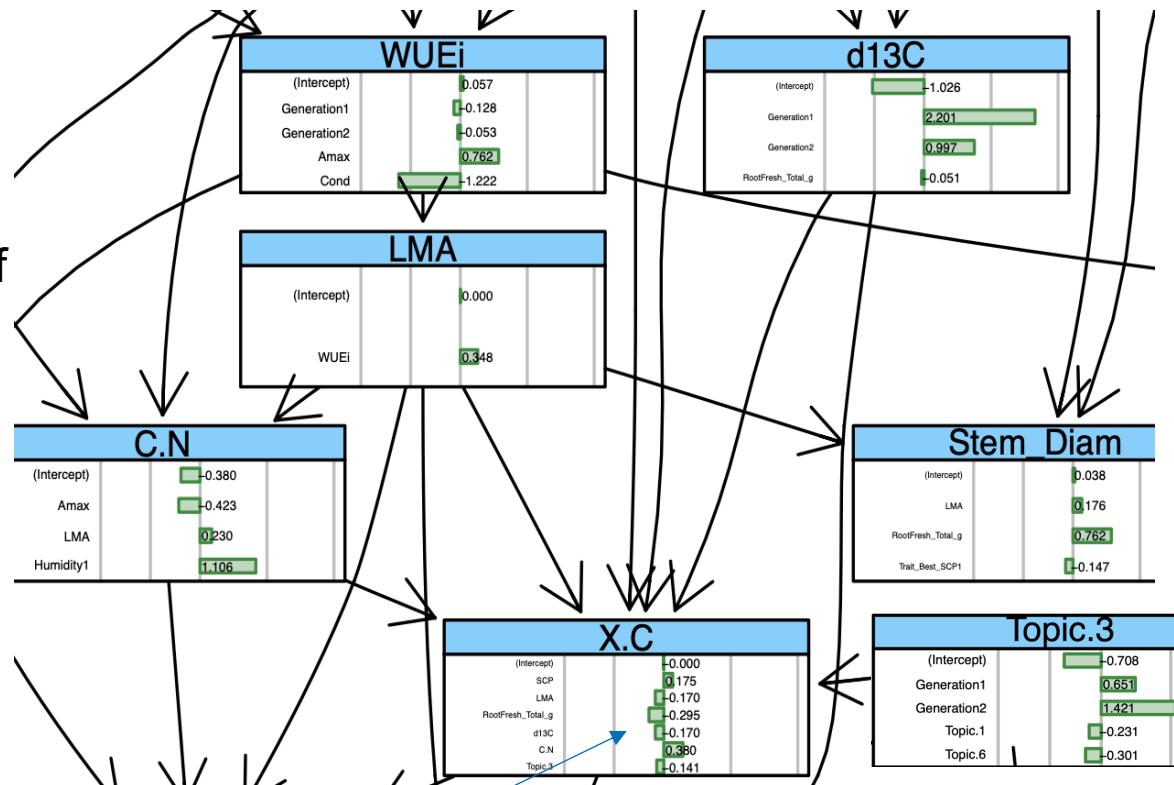
# Establishing links via Bayesian networks

- Unsupervised framework
- Handles uncertainty
- Constrainable
- Can answer the questions of interest:

*How do certain soil source and abundances of bacteria affect a Stomatal closure point (SCP)?*

Probability (SCP < average given that Soil type is Forest and Topic 1 > average) = 0.32

Probability (SCP < average given that Soil type is Agricultural and Topic 1 > average) = 0.49.



regression coefficients

Topic 3: Proteobacteria, Deinococcota, Actinobacteriota bacteria