Variance of Variables: Shoot vs. Root Phenotypes for Explaining Class Differences in Maize

Urban Halpern, Ruby Noland*, Jessica Araszewski, Keara Botanes, Zhuocheng Gan, Amanda Godfrey, Gunnar Larsen, Malisa Lo, Ken Kiehl, Maya Pimolwatana-Montoya, Michael Kantar, Tai Maaz, Yuriy Mileyko, Michael Muszynski, Nhu Nguyen*

Introduction

- Maize is a globally important food, feed, and biofuel crop that serves as a model organism in agricultural science.
- Roots are laborious to study leaving a gap in our understanding of this important food crop.
- Because **roots are understudied**, it is unclear if you get the same information from using roots as you would using shoots.
- **GOAL**: **Compare variance** in shoot versus root traits that are most correlated with genetic background, in order to **identify significant traits to focus on in future classification studies**

Data Collection

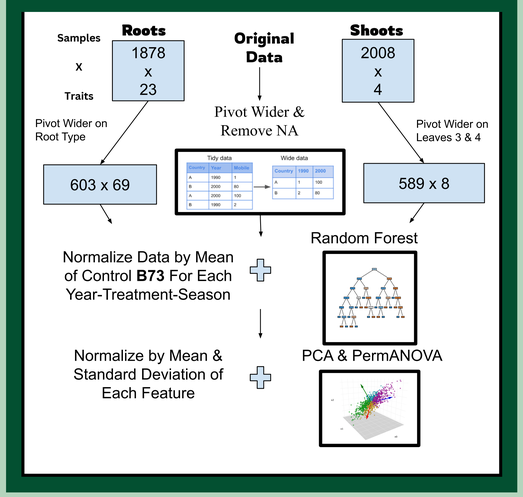
- Collected morphological data from 20 diverse maize genotypes.

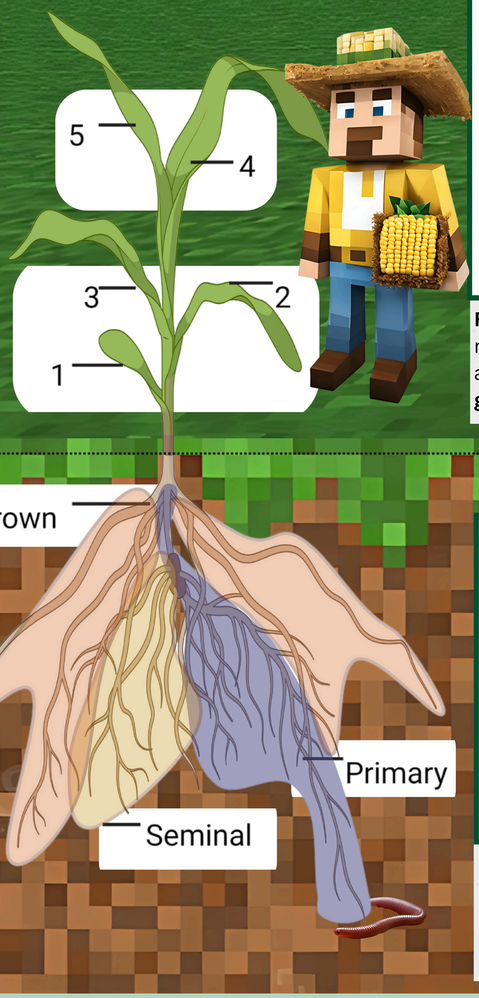
Genotypes Grouped by Genetic Background

Genetic Background	Genotypes
Mixed	Mo18W, Tx303, M37W
Non-Stiff-Stalk	M162W, OH7B, Ky21, Ms71
Popcorn	HP301
Stiff-Stalk	B73, B97
Subtropical	NC350, CML277, CML247, Ki11, CML333, CML103, CML322, CML228, CML52, CML69

- Image analysis was conducted on **root architecture** and **shoot morphological traits.**

Data Analysis





SHOOTS

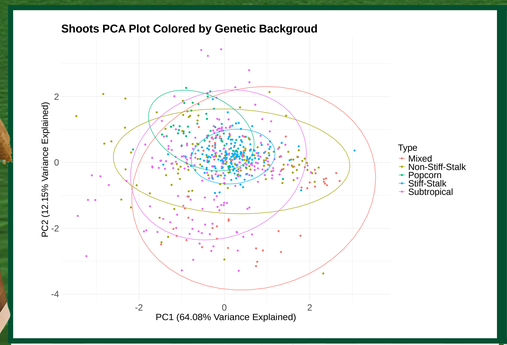


Fig. 1A - PCA results of **shoot data** after Varimax rotation with 95% confidence ellipses. Centroids appear to be **more spread out** and **dispersion is greater** than in the root data.

Pairwise Permanova Results on Shoot Data (95% Confidence Interval P-values)

	Mixed	Non-Stiff-Stalk	Popcorn	Stiff-Stalk
Non-Stiff-Stalk	0.0017*	-	-	-
Popcorn	0.0017*	0.0057*	-	-
Stiff-Stalk	0.0017*	0.012*	0.0017*	-
Subtropical	0.0075*	0.01*	0.0017*	0.0017*

Red cells indicate p-values where the assumption of homogeneity is violated.

Fig. 2A - PermANOVA and Tukey test results on the Principal Components. **Mixed** is the most significantly different centroid-wise. **Stiff-stalk** was the most different **dispersion-wise**.

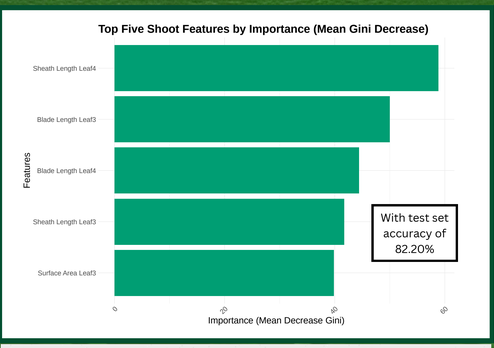


Fig. 3A - Best features to split on in the final model by Mean Gini Decrease (MGD). **Shoots dataset** had a **prediction accuracy** of **82.20%** on Genetic Background which is greater than **roots (70.83%)**

ROOTS

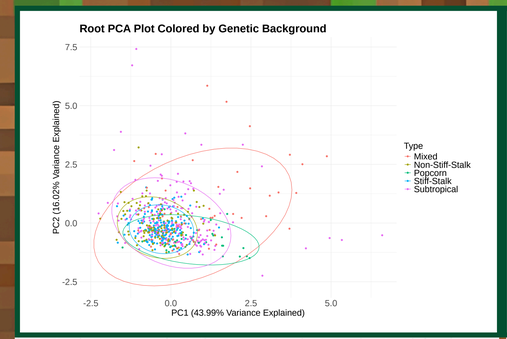


Fig. 1B - Displays PCA results of **root data** after Varimax rotation with 95% confidence ellipses. Centroids appear to be **more clustered** and **dispersion is less** than in the shoot data.

Pairwise Permanova Results on Root Data (95% Confidence Interval P-values)

	Mixed	Non-Stiff-Stalk	Popcorn	Stiff-Stalk
Non-Stiff-Stalk	0.0011*	-	-	-
Popcorn	0.0011*	0.0011*	-	-
Stiff-Stalk	0.0011*	0.002*	0.0011*	-
Subtropical	0.0011*	0.0011*	0.0011*	0.0011*

Red cells indicate p-values where the assumption of homogeneity is violated.

Fig. 2B - PermANOVA and Tukey test results on the Principal Components. **Popcorn** was the most different centroid-wise. **Mixed** was the only class that was not different from any other class centroid-wise.

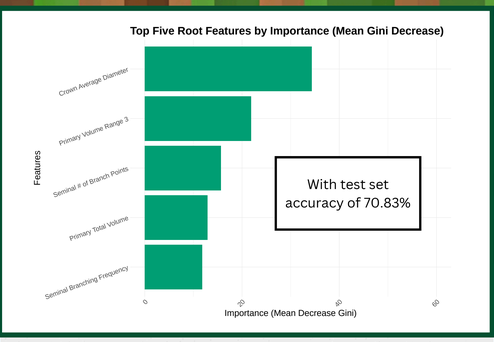


Fig. 3B - Best features to split on in the final model by Mean Gini Decrease (MGD). **Root phenotypes** were **worse at reducing variance** among classes in child nodes possibly due to high number of features in the root dataset.

RF Training Results

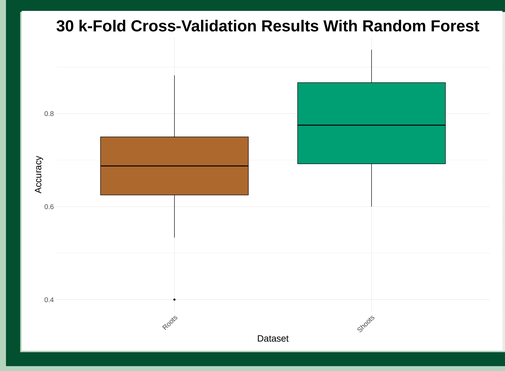


Fig. 4 - Random forest training with 500 trees and 30 k-folds Shoots (n=471) had a higher average accuracy than roots (n=483) along with larger variation.

Conclusions

- **Shoots** are a better discriminator of Genetic Background using **Random Forest**.
- The best features were Leaf 4 sheath length and Crown root average diameter.
- **Mixed** was the standout class in the **shoots** while **Popcorn** was the standout in **roots.**
- Scientists can **leverage this information** to reduce labor in the data collection process by **focusing on** only shoots or only **key traits**.

Future Studies

- More samples needed so that genetic origin classes are balanced.
- Possible interactions between Root and Shoot features were not accounted for.
- Variation during reproductive stage.

Acknowledgements

We would like to thank CTAHR and all of the research centers and farms that provided us with hands on research experience and unique opportunities to mālama ʻaina. Thank you to Jonathon, Ishwora, John, Zuki, Nugget, Drogo, and Xavie for your unending support. Thank you especially to Dr. Tai and Dr. Mikey for taking nervous undergrads and helping us feel like scientists. We were funded by USDA REEU grant 2020-67037-30665.