Elucidating the patterns of pleiotropy and its biological relevance in maize

Merritt Khaipho-Burch [1]; Taylor Ferebee [2]; Anju Giri [3]; Guillaume Ramstein [3,4]; Brandon Monier [3]; Emily Yi [3]; M. Cinta Romay [3]; Edward Buckler [1,3,5]

mbb262@cornell.edu

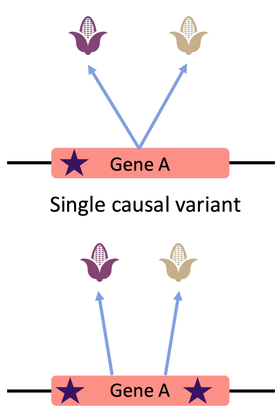
@MerKhaiBurch

1) Section of Plant Breeding and Genetics, Cornell University, Ithaca, NY; 2) Department of Computational Biology, Cornell University, Ithaca, NY; 3) Institute for Genomic Diversity, Cornell University, Ithaca, NY; 4) Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus, Denmark; 5) USDA-ARS; Ithaca, NY

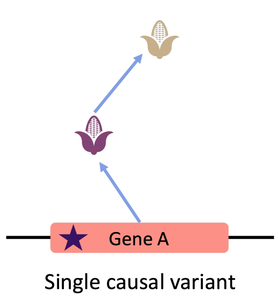
Pleiotropy: a single locus that affects multiple traits. How pleiotropic is any maize locus?

Many forms of pleiotropy exist:

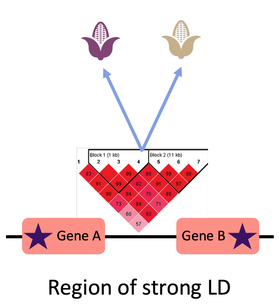
Biological/Horizontal



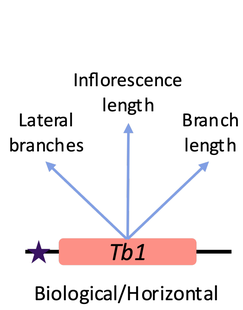
Mediated/Vertical

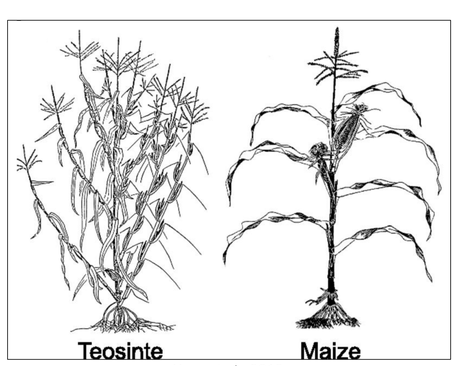


Spurious



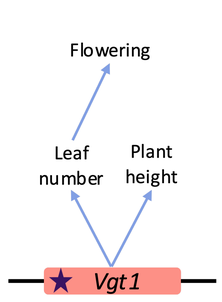
Few examples of pleiotropy have been described and verified in maize:





Yang et al., 2019

Vegetative to generative transition 1



Salvi et al., 2007

GWAS on 120,597 field, metabolite, and expression traits uncovers the patterns of pleiotropy.

Split by population and trait-type
(Population-Trait Categories)

Pop.	Trait-type	N Traits	N Taxa	Perms.
NAM	Field	121	5,186	10
Goodman	Field	222	271	10
Goodman	Metabolite	3,873	271	10
Goodman	Expression	116,381	271	5

Perform GWA
(Filter p < 0.00001)

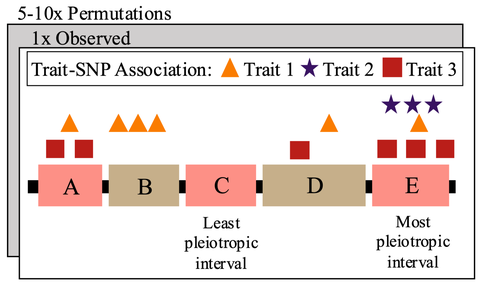
NAM: 15 million SNPs
$y = X\beta + 3\ global\ PCs + main\ window\ PCs + e$
$y = X\beta + 3\ global\ PCs + mid\ window\ PCs + e$

Goodman Panel: 17 million SNPs
$y = X\beta + 3\ global\ PCs + e$

Goodman Panel Expression: 17 million SNPs
$y = X\beta + 3\ global\ PCs + 25\ PEERs + e$

Aggregate GWA results by interval

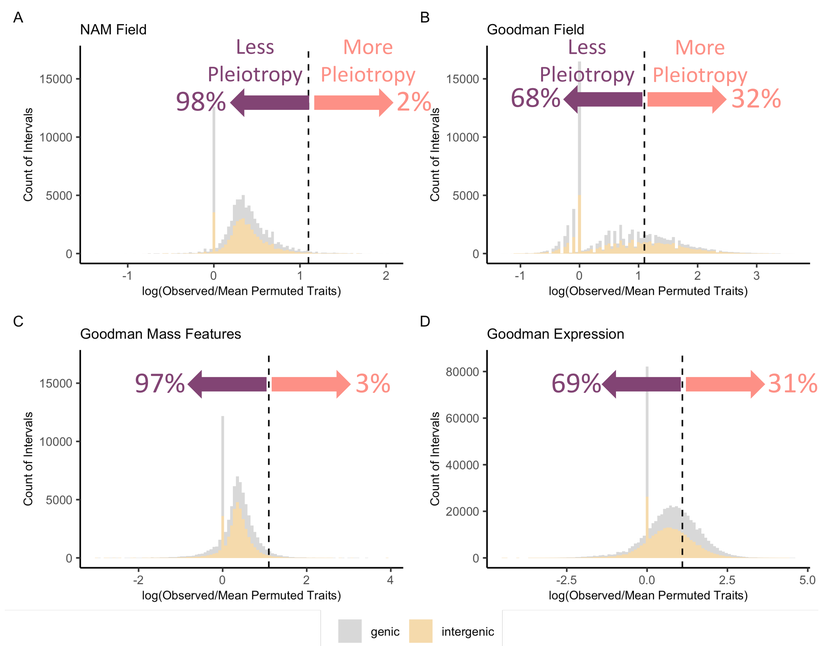
5-10x Permutations
1x Observed
Trait-SNP Association: Trait 1 Trait 2 Trait 3



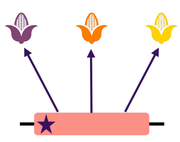
Count unique traits across intervals

Interval	Observed	$Perm_1$	…	$Perm_N$
A	2	0		0
B	1	1		0
C	0	0		1
D	2	2		1
E	3	1		0

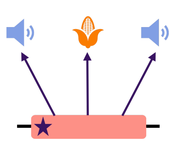
n = 74,490 intervals



68-98% of intervals are involved with 0-1 traits.

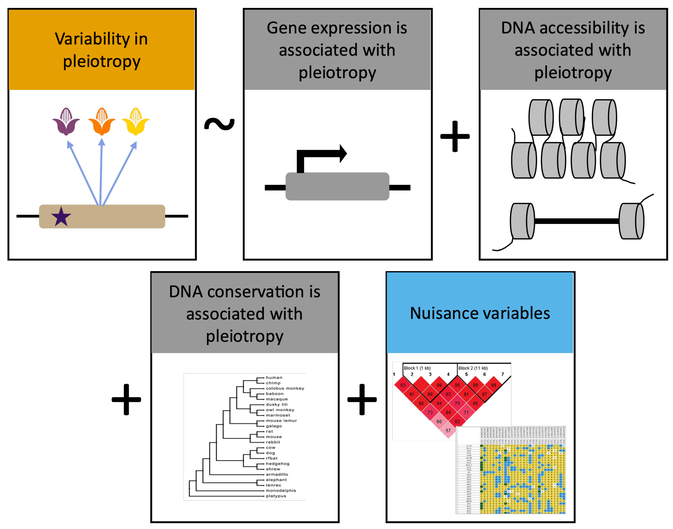


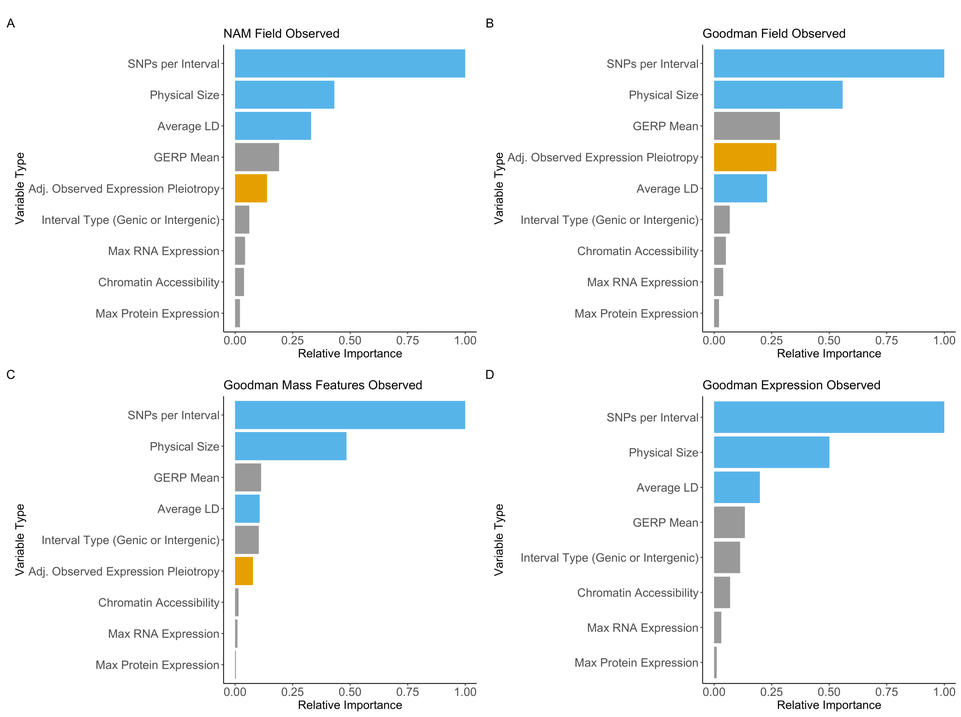
2-31% of intervals are involved with 2+ traits.



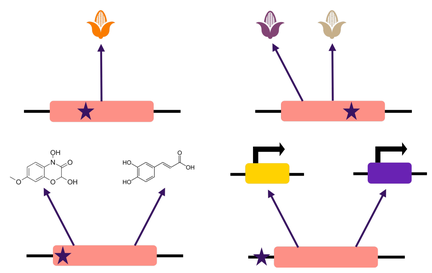
Much of this observed pleiotropy may be spurious.

Random forest models show high importance for nuisance terms, low importance for biological terms.

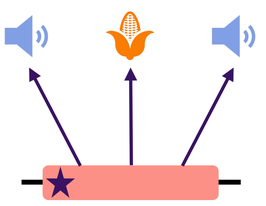




The distribution of pleiotropy differs by population and trait.



Pleiotropy estimates are highly biased towards nuisance terms and investigator bias. Caution & validation should be applied.



Want to know more? Take a FREE zine about this poster!

Thank you to our funding sources: NSF PGRP: IOS#1822330, the USDA ARS, and the USDA NIFA AFRI Pre-doctoral fellowship: 2022-67011-36458.

National Institute of Food and Agriculture
U.S. DEPARTMENT OF AGRICULTURE

Preprint coming soon! bioRxiv THE PREPRINT SERVER FOR BIOLOGY