

# 1 Summary

How does the genetic architecture of quantitative traits change as a result of demography and selection during the maize domestication bottleneck and the further bottleneck that some maize populations underwent from Central America to South America?

If we compare maize and teosinte (which are locally adapted in their respective pops), do we see evidence of:

- many genes of small effect or few genes of large effect underlying important traits?
- do these differences match our predictions based on their differing demographic histories?

Broadly relevant because these results could inform on maintaining diversity in crops for the future in ways that may have much larger long term impacts in the maintenance of diversity as well as fitness/adaptation in the face of climate change and future adaptation to changing environmental conditions.

## Intro

genetic architecture underlying traits affects how easy/hard and quick/slow local adaptation can occur. important for breeding, conservation, predicting response to climate change, etc. useful info for crops/domesticated species (and also things like disease in humans?)

selection interacts with demography.  $N_e s > 1$  for selection to win. as demography changes,  $N_e s$  changes. so demography is known to directly affects the strength of selection, but what is less studied is how this may restructure the genome in terms of how evolution proceeds for adaptation to new or changing conditions.

Different demographic histories such as bottlenecks, repeated founder effects during expansion, or rapid population growth can have effects that interact with selection. Small population sizes can lead to purging (recessives become homozygous and removed in smaller pops), but also increase in deleterious alleles (surfing phenomenon due to random genetic drift). mutations that affect a trait related to fitness will then be impacted by demography x selection interaction. in annual plants, nearly ALL traits related to fitness we thus predict that genetic architecture (number and size of mutations) should be impacted by demographic change. this is not understood well in any system. controversial in humans (lohmueLLer vs. pritchard etc.). in plants, summaries are on gross overall  $N_e$  (small vs. big) and ignore recent demographic change.

## Proposed Research

The maize/teosinte system, has a well described demographic history, supported by archaeological and other human records of its use and domestication in the Americas.

Maize was domesticated  $\approx 9000$  years ago in SW Mexico which bottlenecked populations to a small size, but was subsequently followed by a large population expansion more recently. This demographic event can be detected genetically, as shown by Beissinger in prep (github repo), where populations were bottlenecked to 5% of their previous size followed by spread across central and southern America and into highland and lowland environments. Populations of maize have since recovered to much larger effective population sizes than even before the domestication bottleneck (humongous growth to at least 300K but maybe as much as 1E9, citation).

This system thus serves as a great resource to compare the effects of this history on the genome in maize versus its extant, wild ancestor teosinte.

Being a crop species that is incredibly useful to humans (lots of citations and examples of usage), there is a wealth of knowledge on quantitative traits. will need to explain rare alleles pops and experiment some (PGRP15 grant)

## 1.1 Obj 1 - Estimate the DFE

from polymorphism/divergence data. use HapMap2 or current teosinte genomes (I think I'd vote for latter) <http://goo.gl/CLmsmX> and <http://www.genetics.org/content/177/4/2251.short>) can either use estimated demographic model or use noncoding sites to normalize SFS

can compare to e.g. GERP distribution perhaps as validation?

this gets at magnitude of effect, but also want to get at how many loci may be contributing to any given trait?

## 1.2 Obj 2 - Simulate scenarios of different traits.

simulate using estimated demography and DFE. simulate traits w/ varying correlation with fitness new mutation effects on fitness determined by DFE, effect on trait by correlation between trait and fitness

evaluate: how many loci contribute to important traits? how strong are these effects? how do details of demography impact outcome? test against theory e.g <http://arxiv.org/abs/1312.3028>

(could also do some broader simulated examples just to stand alone and see if other various outcomes may occur - just don't plan on comparing these to any real data)

nope. we just need DFE. see obj 2

yes, once we can show we can recapitulate real data, i think this is useful

## 1.3 Obj 3 - compare simulation results to modern maize genomes, known to have undergone the same demographies simulated in objective 2

compare to GWAS for maize/teo. do we recapitulate observations? if not, why? are there differences between central and southern American pops?

Other larger picture impacts/intellectual merit?

it is thought, and shown in some human pops, that demog. history such as expansion leads to an increase in delet alleles, and of larger effects - b/c of continued inferred expansions and bottlenecks. is there any evidence of this in maize?

we have no GWAS data for S.Amer. pops, but do have genomes and GERP. we could get freq. etc. of del. mutations from sims and compare to GERP

see ideas and text in "service\_award.tex" that I uploaded too.