Project Summary

Maximum of 1 page

Intellectual Merit

Broader Impacts

Project Description

Introduction

Due to their sessile nature, natural selection will act to adapt plants to their local environments (Stebbins, 1950). Understanding the genetic basis of how plants adapt to local conditions – how many loci are involved, what are their effect sizes, and how similar are they among populations and species – will thus allow for improved breeding and conservation strategies. This is particularly pressing given current issues of climate change, habitat loss, and population growth. These pressures will require adaptation of crops and wild plants to changing local conditions and cultivation of crops in new locales to meet growing demand.

Agricultural species represent promising systems for ongoing research on local adaptation. Most crops were domesticated in narrow geographic centers but encountered and adapted to a wide range of novel environments as agriculture expanded across the globe. In many instances traits important for local adaptation have already been identified. These systems therefore represent compelling opportunities for investigating the genetic architecture of local adaption. Moreover, insights gained regarding adaptive loci can feed back into modern crop improvement, yielding valuable benefits in the face of rapid environmental change.

Here we propose to use maize adaptation (*Zea mays* ssp. *mays*) to high elevation environments as a model for understanding the genetic basis of local adaptation in plants. Maize was domesticated in the lowlands of southwest Mexico from the narrowly distributed teosinte *Zea mays* ssp. *parviglumis* (hereafter, *parviglumis*; Matsuoka et al., 2002). Since domestication, maize has spread worldwide: analysis of cultivation area data indicates maize has the greatest global geographic breadth of 16 staple crops (Figure 1) and is now cultivated on six continents, ranging from southern Chile to Canada (Tenaillon and Charcosset, 2011) and from sea level to well over 3000m in altitude. In addition to maize and *parviglumis*, a related teosinte *Zea mays* ssp. *mexicana* (hereafter, *mexicana*) is found only in highland environments, having adapted to highland environments thousands of years prior to maize domestication (Ross-Ibarra et al., 2009; Hufford et al., 2012a). Maize and teosinte thus form an ideal system in which multiple replicated evolutionary experiments will allow us to dissect the genetic architecture of highland adaptation as a model for understanding plant local adaptation.

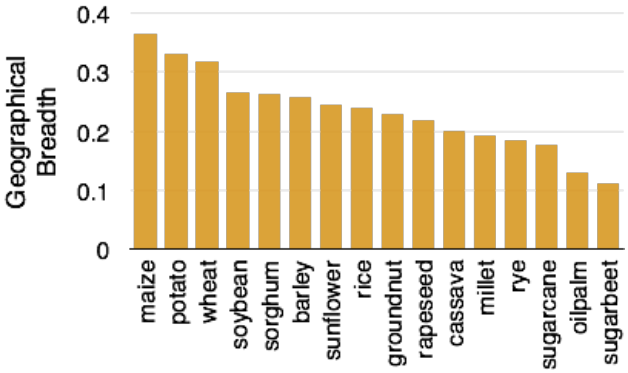


Figure 1: Geographic breadth of the world's 16 staple crops, expressed in percent of land surface area in which each crop is cultivated. Data are from Monfreda et al. (2008).

Aims

We will investigate the genetic basis of highland adaptation in maize by achieving three aims:

1. **Dissect the genetic architecture of highland traits**
2. **Investigate population genetic signatures of highland adaptation**
3. **Characterize functional variation at adaptive quantitative trait loci**

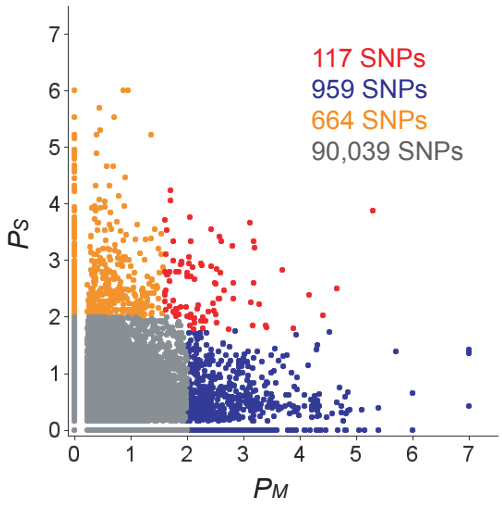
Rationale and Significance

While the genetic basis of local adaptation is generally not well understood, the declining cost of genotyping has enabled a handful of genome-wide studies across populations of model species. For example, Fournier-Level et al. (2011) demonstrated that alleles associated with high fitness in *Arabidopsis thaliana* have a tendency to be both local and linked to climate. Likewise, a recent study across hundreds of accessions of *Medicago truncatula* identified candidate loci for local adaptation and found them to be predictive of growth rate under temperature and soil moisture treatments (Yoder et al., 2014). Finally, our own genome-wide study of teosinte (the wild relatives of maize) revealed an important role for inversion polymorphisms and – in contrast to results from *Arabidopsis* (Hancock et al., 2011) – an enrichment of regulatory variants among loci showing evidence of selection (Pyhajarvi et al., 2013). While much of local adaptation may involve complex quantitative traits (Le Corre and Kremer, 2012), the genetic architecture of these traits does not necessarily mirror results from mapping studies in other populations. In maize, for example, although genome-wide association in the NAM panel suggests that flowering time is largely controlled by many loci of small effect (Buckler et al., 2009b), adaptive change in flowering time across latitudes has involved loci of large effect on photoperiod (Hung et al., 2012a). The strength and timing of selection on a trait also plays a role: while ear and tassel traits in maize share a number of QTL, those underlying ear morphology, which underwent recent strong selection during domestication, are of larger mean effect size (Brown et al., 2011). Though initial genomic studies are beginning to yield valuable insights regarding local adaptation, clearly much remains to be discovered.

Highland adaptation in maize and teosinte are an excellent system in which to study local adaptation. Following domestication, maize spread to the highlands of the Central Plateau, a migration across more than 1000m of increasing elevation. Colonization of the highlands required adaptation to a number of novel abiotic conditions, including gradients of temperature, precipitation, and elevation. Highland landraces have distinct morphologies (e.g., highly pigmented and hairy leaves and stems) that are believed to confer adaptation to this cooler region (Doebley, 1984). Our previous genetic analyses (van Heerwaarden et al., 2011) show that maize has independently adapted to highland environments multiple times, including the southwest US and the Andes of South America, where landraces (i.e., local farmer varieties) are commonly grown above 3000m. Multiple independent instances of highland adaptation in maize and teosinte provide replicated evolutionary experiments, providing power to identify and validate loci common candidate loci as well as discover multiple potential mechanisms for highland adaptation unique to each population.

Study of the genetic architecture of maize adaptation will provide both basic evolutionary insight and essential information to help increase or sustain yield in the face of human population growth and climate change. Historical analyses suggest that climate change over the last 30 years

Figure 2: Little overlap of adaptive loci between continents. Shown is a scatter plot of $-log_{10}$ empirical p-values of genetic differentiation (F_{ST}) in Mexico (P_M on x-axis) and S. America (P_S on y-axis). SNPs showing evidence of selection are highlighted in blue (Mexico), orange (S. America), or red (both Mexico and S. America), along with the number of SNPs in each category.



has already dramatically impacted maize yields worldwide, slowing gains from breeding and management (Lobell et al., 2011b). Lobell et al. (2011a) further determined that future temperature increases will likely decrease yield across 65% of African maize-growing regions, while all of Africa will see diminished maize yield if increased temperature is accompanied by drought. An understanding of how maize has adapted to challenging environmental conditions in the past will help breeders to mitigate yield loss due to future changes.

Preliminary Results

Preliminary work from project members positions us to make excellent progress on our proposed aims. Co-PIs Ross-Ibarra and Hufford have worked extensively on the population genetics of highland adaptation. Pyhajarvi et al. (2013) explored local adaptation in *parviglumis* and *mexicana* populations, finding a large number of loci showing association with altitude and evidence of selection, as well as highlighting the potential importance of regulatory variants and large inversion polymorphisms. This study identified a putatively adaptive inversion on chromosome four that distinguishes the lowland *parviglumis* from the highland *mexicana* and coincides with a quantitative trait locus associated with traits linked to highland adaptation (Lauter et al., 2004). This *Inv4m* inversion is the subject of our functional characterization in Aim 3. Pyhajarvi et al. (2013) also identified populations of *parviglumis* showing extensive admixture with the highland *mexicana* which are the subject of proposed analysis in **??**. Hufford et al. (2013) identified genomic regions in highland maize that have introgressed from the highland *mexicana*. They showed that plants with *mexicana* alleles showed highland phenotypes and superior growth under cold conditions, suggesting an adaptive role for introgression and motivating our population genetic analyses in Aim 2. Finally, recent analyses of selection in genotyping data from a wide collection of landraces from the highlands of Mexico and S. America finds little overlap in the genes important for adaptation (Takuno *et al.* In Prep; Figure 2), motivating the QTL analysis in **??**.

Co-Pi Flint-Garcia and Sr. Personnel Sawers have made important progress on the development of populations for the project. For Aim 1.1 our Mexican cross is already at the F2 generation, and one potential S. American cross is now at the F1 generation (Table 1). Back-crosses of the reference

genome inbred B73 to a highland Mexican landrace Palomero Toloqueño have been made and selfed to generate a BC1S1 population that will be further developed in Aim 3.2 to dissect the function of the *Inv4m* polymorphism. Both SSRs and SNPs that distinguish the *mexicana* alleles at *Inv4m* have also been developed.

Co-PI Coop has worked to develop analytical approaches to understanding local adaptation, including methods that allow genome-wide association with environmental variables (Coop et al., 2010; Günther and Coop, 2013), detection of selection in introgressed populations (Brandvain et al., 2014), and powerful approaches to identify phenotypic selection on quantitative traits Berg and Coop (2014). His group is currently working on methods for mapping and studying adaptation in admixed populations.

Specific Objectives

Aim 1 Genetic architecture of highland traits

One of the primary goals of this proposal is to determine the genetic architecture of highland adaptation. Ultimately, this knowledge will be useful for determining the genes underlying these loci (Aim 3) and the pathways involved in adaptation (Aim 2). These loci can also be used in maize improvement via marker assisted selection. In this aim we wish to determine how many genomic regions control adaptive phenotypes, where these regions are located, and the distribution of allelic effects at these loci. We first perform comparative QTL analysis in two highland x lowland crosses (Aim 1.1), then take advantage of historical recombination and greater resolution to map loci in an admixed population of highland and lowland teosinte (Aim 1.2).

Questions

- What is the genetic architecture of highland adaptation?
- How much of the genetic architecture is shared between Mexico and South America?
- How much of the genetic architecture is shared between maize and teosinte?

Aim 1.1 QTL mapping of highland adaptation

Our first objective is to identify genomic regions controlling highland adaptation in maize. We will conduct QTL mapping studies of one Mexican and one South American population, each derived by crossing a landrace adapted to lowland conditions with a landrace adapted to highland conditions (Table 1). We make use of specially-inbred landrace lines created by John Doebley (U. Wisconsin) and Seth Murray (Texas A&M), thus simplifying downstream applications and allowing replication of alleles in our functional studies (see Aim 3).

We will self-pollinate the F2 plants to create 500 F2:3 families from each population. DNA will be extracted from each of the F2 plants and sequenced to 20-30X depth on two lanes of Illumina (150bp paired-end reads on a HiSeq 2500 at the UC Davis Genome Center), providing genome-scale SNP data similar to our previous work (HapMap.v2; Chia et al., 2012). F2 plants will be genotyped using genotyping-by-sequencing (GBS; Elshire et al., 2011) and run through the standard maize GBS pipeline (Glaubitz et al., 2014) resulting in approximately ˜1M SNPs, allowing straightforward

Table 1: Parental lines for QTL

Population	**Parent**	**Origin (masl)**	**Status**
Mexico	Zapalote Chico Palomero de Jalisco	Oaxaca (46) Jalisco (2520)	F2
S. America	Araguito Sal Prieta	Venezuela (183) Ecuador (2948)	F1

Table 2: Common garden locations

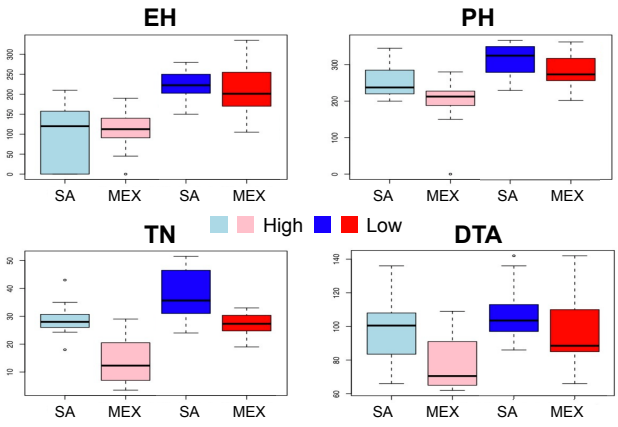
Field Sites	**Lat/Lon**	**Elevation (m)**	**Min/Mean/Max °C**	**Precip (mm)**
Valle de Banderas, Nayarit	20.8, -105.2	54	15.3/25.8/33.7	1184
Irapuato, Guanajuato	20.7, -101.3	1729.0	7.3/20.2/31.7	693
Amealco, Querétaro	19.5, -99.1	2240.0	2.3/15.6/27.0	626
Columbia, Missouri	28.9, -92.2	266.1	-17.8/36.0/40.5	914

imputation of their full-genome sequence. The genetic map will be created using standard methods (e.g. Broman et al., 2003).

Populations will be phenotyped at 3 field locations, including one lowland site (Valle de Banderas in Mexico), one highland site (Irapuato or Queretaro, Mexico), and one temperate site near Columbia, Missouri (Table 2). At each field location, best local practices will be used including fertilizers and pest and weed control.

At each site, the experiment will consist of two replicates in which the 500 entries will be arranged in an augmented alpha lattice design. Parental checks will be included to control for field variation. We will collect a number of phenotypes (Figure 3) using our in-house barcode-based data collection program. Germination assays in controlled conditions will be conducted in Ames, Iowa, and root chilling will be evaluated using a custom hydroponic system at the University of California, Davis (see letter of support from Dr. Arnold Bloom).

Raw data from each plot will be analyzed using mixed-models incorporating replications and environments. Data will be analyzed across environments to determine whether location (elevation) affects the various phenotypes. Each location will then be analyzed separately to derive least squares means to be used as phenotypic data in QTL analyses. QTL analysis will be conducted using standard software (e.g. SAS; R/qtl Broman et al., 2003). Several iterations of QTL analysis will be conducted: on individual traits, individual traits adjusted for covariates such as flowering time, and



Trait	Phenotype
MH	leaf sheath macrohairs
DTS	days to silking
DTA	days to anthesis
PH	plant height
BM	total plant biomass
EH	ear height
EN	ear number
FK	fifty kernel weight
TN	tassel number
TBN	tassel branch number
TL	tassel length
SM	total kernel mass
RC	root chilling response
GDP	germination depth
SC	stem anthocyanin content
GDP	germination depth
GDT	germination temperature

Figure 3: Phenotypic differences between a sampling of highland and lowland landraces from Mexico and S. America, grown in common garden in Puerto Rico (left). List of the pehnotypes to be measured in the field (right)

multiple traits simultaneously. QTL profiles will be compared across populations (Mexico vs South America) and field sites (elevation) to determine differences in how elevation affects putatively adaptive traits. Comparison of the genetic architecture among traits will inform us of the lability of these traits and their amenability to selection via breeding. Finally, the contrast of each Mexican location to the Missouri location will account for daylength differences and agronomic value in the Midwest.

The expected outcomes of this objective will be 1) A map of QTL underlying phenotypic differences between highland and lowland maize in Mexico and South America, detailing the number and effect size of each QTL and differences between crosses, and 2) Estimates of fitness differences (PH, BM, SM, and FK (Figure 3) of highland and lowland plants, as well as F2 with various combinations of QTL, in both environments.

Aim 1.2 Admixture mapping in a teosinte hybrid zone

While *mexicana* and *parviglumis* are largely allopatric, the subspecies overlap in two regions of Mexico, eastern Jalisco state and the eastern Balsas River Basin (Hufford et al., 2012a), and a number of hybrid populations have been documented in these regions (Fukunaga et al., 2005). We have previously documented near equal proportions of ancestry from the two subspecies admixture in one of these populations near the town of Ahuacatitlan in the eastern Balsas (Pyhajarvi et al., 2013). Growth chamber experiments also suggest plants in this population have higher fitness

in cold conditions than other *parviglumis* populations. Moreover, the relatively short length of haplotypes Ahuacatitlan shares with other populations suggests that there has been extensive recombination since the initial admixture event, providing an ideal population for high-resolution admixture mapping of *mexicana* highland adaptation traits.

In November of year one of the project we will travel to Ahuacatitlan and collect seed from 500 individuals drawn randomly from the population. Seed samples will be transported to Langebio in Irapuato, Mexico for cold storage. A single seed per individual (500 total) will be germinated on filter paper and transplanted into our two Mexican field sites (Table 2). Phenotypes detailed in Figure 3 will be collected for admixture mapping. The majority of these traits are known to differ considerably between *parviglumis* and *mexicana* (CITE). Leaflet samples will be collected from plants in the field at the seven-leaf stage, and extracted DNA will be genotyped using GBS. Several computational methods for admixture mapping have already been developed (Winkler et al. (2010)), but we will augment these with methods currently under development by Co-PI Coop.

However, these methods are not well suited to admixture mapping when there are differentially related individuals in the sample, and when natural selection may have systematically distorted admixture at some loci. In natural admixed populations these issues can be expected to occur, and will potentially result in false positives due to the non-independence of individuals (a fact accounted for in plant GWAS, but not in admixture mapping). We will implement novel methods currently under development by Co-PI Coop in our analysis of the Ahuacatitlan population that incorporate this non-independence into admixture association tests, while accounting for uncertainly in admixture calls along the genome.

Aim 2 Adaptive value of highland alleles

In aim Aim 1 we will map loci corresponding to traits differing between highland and lowland maize and teosinte. In this section we will test the adaptive significance of QTL identified in Aim 1 in three natural introgression experiments: gene flow from *mexicana* into highland maize Aim 2.3 and admixture between *mexicana* and *parvgilumis* Aim 2.3.

Questions

- Are highland QTL/loci widespread in highland climes?
- Are loci controlling phenotypic differences between highland and lowland populations adaptive?
- Does natural selection favor introgression from adapted populations?

Aim 2.1 Global analysis of highland haplotypes

Full genome resequencing underway will increase the resolution of this study and expand its scope to include maize from the highlands of Guatemala and the Southwestern United States.

And though quantitative genetic theory suggests that adaptive phenotypic change may not correlate with strong evidence for selection on individual loci (Le Corre and Kremer, 2012), recently developed methods from Co-PI Coop (Berg and Coop, 2014) provide a powerful statistical framework to identify coordinated shifts in allele frequencies at causative QTL (from Aim 1 to look for

weak selection on alleles underlying highly quantitative traits.

Aim 2.2 Population genetics of maize-teosinte introgression

Our previous work (Hufford et al., 2013) we documented extensive introgression between *mexicana* teosinte and highland maize landraces, demonstrating an overlap with teosinte QTL for macrohairs and stem pigmentation (Lauter et al., 2004). Because of the relatively low-density genotyping used, however, we were limited to identifying large regions of ancient introgression present in most populations. We were also unable to investigate evidence of selection for any of the introgressed regions. Here we propose to revisit these populations with higher-density genotyping that will allow identification of ongoing gene flow in individual populations and ask whether introgressed regions have been targeted by natural selection.

We propose to resample 18 individuals from each of the same 9 sympatric population pairs and two allopatric populations studied in Hufford et al. (2013). Each individual will be genotyped by GBS using greater than normal depth (48 plex) to improve genotyping heterozygous sites. These data will provide ˜1M SNPs across the genome (compared to 40K SNPs in Hufford et al. (2013)). We will use both haplotype (Price et al., 2009) and heterozygosity-based (Geneva et al., 2014) methods to identify introgressed segments in individual populations. Genomic regions showing evidence of introgression will be tested for selection using population genetic approaches which utilize evidence from the site frequency spectrum (Nielsen et al., 2005) and haplotype structure (Voight et al., 2006). Correlations between genetic differentiation and recombination will allow us to investigate selection against introgression (Brandvain et al., 2014), quantifying the "linkage drag" associated with introgression of potentially beneficial adaptive alleles. Finally, we will again apply the approach of Berg and Coop (2014) to evaluate selection on individual phenotypes across highland maize landrace populations.

The expected outcomes of this objective are 1) a fine-scale dissection of both ancient and ongoing introgression 2) identification of introgressed regions showing evidence of positive selection, identifying loci important for highland adaptation, 3) evidence for selection on specific phenotypic traits, 4) quantification of the potential "linkage drag" or evidence against introgression across other regions of the genome.

Aim 2.3 Population genetics of hybridization in teosinte

In this objective, we will capitalize on the long history of gene flow between *mexicana* and *parviglumis* (Ross-Ibarra et al., 2009) to identify loci that show evidence of selection in admixed populations. To complement the Ahuacatitlan population from Aim 1.2, we will sample four additional admixed populations. We have already identified multiple admixed populations using small-scale SNP data from across the range of each taxon (van Heerwaarden et al., 2011). We will revisit each of these populations to sample seed, collecting 50 individuals per population. Samples will be genotyped via high-coverage (48plex) GBS to ensure accurate identification of heterozygous sites. The extensive gene flow between these taxa should lead to signatures of admixture throughout the genomes (Pyhajarvi et al., 2013). Population genetic theory predicts, however, that adaptive loci which have introgressed due to natural selection should show distinct signals of elevated admixture, and our preliminary simulation results bear our this prediction (Figure 4). We will use these patterns of admixture, as well as more classical signals of selection (see) to identify loci showing evidence of adaptive introgression in these admixed populations. Replicated admix populations, combined

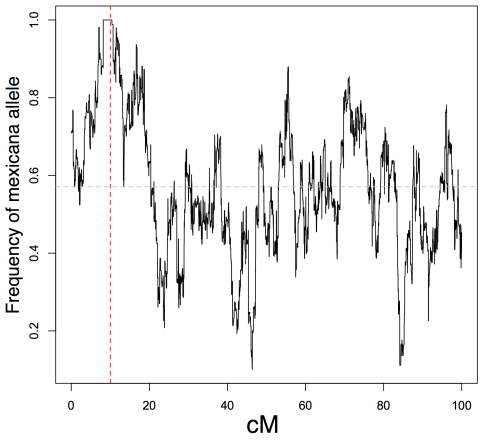


Figure 4: Analysis of 100 generations of simulated admixture between *mexicana* and *parviglumis* across a 100cM chromosome. A beneficial *mexicana* allele with selection strength $s = 0.1$ is introgressed at position 10cM (red vertical line), showing that deviation from background variation in ancestry (horizontal gray line) can be used to detect selection in admixed populations.

with the high resolution afforded by recombination in teosinte, will allow us to distinguish variants selected in individual populations vs. those showing parallel evidence of selection (repeated evolution) across all populations. Signatures of repeated evolution are indicative of standing genetic variation or multiple pathways (a larger mutational target) to achieve a similar pehnotypic outcome (Ralph and Coop, 2010). Finally, the large sample sizes used will also allow us to identify coordinated shifts in allele frequencies Berg and Coop (2014) of even rare QTL identified in **??**.

The expected outcomes of this objective will be 1) identification of adaptive loci in multiple admixed teosinte populations 2) comparison of selection on phenotypic traits in teosinte vs. maize 3) an improved understanding of the role of repeated evolution during the process of local adaptation

Aim 3 Functional characterization of adaptive QTL

After mapping QTL for highland adaptation (Aim 1) and studying their adaptive significance (Aim 2), in this aim we will aim to better understand the functional genetic basis of adpative regions. First, we will study the phenotypic effects of a chromosome 4 QTL introgressed into a maize background (Aim 3.2). Then we will evaluate the effects of an allelic series from highland and lowland maize and teosinte at the chromosome 4 QTL (Aim 3.3). Finally, we will use RNA sequencing data to investigate plasticity, differences in expression, and identify potential candidate loci within QTL (Aim 3.4).

Aim 3.1 Questions

- What are the phenotypic consequences of introgressing a single adaptive QTL?
- Do adaptive QTL contain allelic series with differing functional consequences?
- How do maize and teosinte data differ in expression response to higland and lowland environments?
- Can RNA-seq help refine QTL to identify candidate genes?

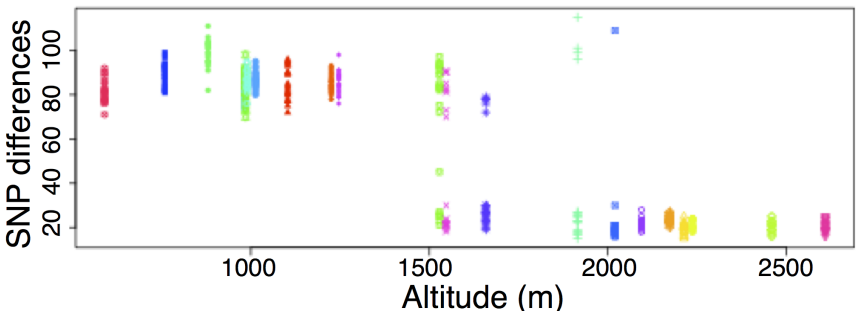


Figure 5: Clinal variation at the Chr4 inversion. Genetic distance (# of SNPs) from the canonical highland haplotype is plotted against elevation for 20 teosinte populations (shown as different colors). Low elevation (¡1500m) populations lack the inversion completely, while it is fixed in populations abouve 2000m. Data from Pyhajarvi et al. (2013).

Aim 3.2 Functional evaluation of a Chr 4 QTL

In this objective we propose to functionally characterize the genetics of a large *mexicana*-maize introgression block located on chromosomes 4 (Chr4: 169-180Mb). Our previous analysis (Hufford et al., 2013) indicates that this region is supported by a robust signature of introgression, shows broad distribution among highland races, and overlaps with a QTL identified in a *parviglumis* x *mexicana* cross (Lauter et al., 2004) associated with leaf pigmentation and pubescence. Our previous population genetic analysis in teosinte (Pyhajarvi et al., 2013) suggests that the region represents an inversion polymorphism that is under selection along an altitudinal cline 5. This suggesting the possibility of characterizing the region as a block by generation of introgression stocks.

We will generate heterogeneous inbred families (HIFS; Tuinstra et al., 1997) from a cross of the landrace Palomero Toluqueño (PT) and the reference genome inbred B73. PT is popcorn originating from the highland valleys of central Mexico that is considered basal to the Mexican highland landrace radiation (Reif et al., 2006); it also exhibits the highest level of *mexicana* introgression among characterized material (Matsuoka et al., 2002). Furthermore, inspection of the PT genome sequence (Vielle-Calzada et al., 2009) shows that PT carries *mexicana* alleles at two SNPs shown previously to exhibit a fixed difference between *mexicana* and maize (Hufford et al., 2013). We will screen an existing collection of ˜150 B73 x PT BC1S3 (three generations of selfing after 1 generation of back-cross) families to identify HIFs segregating for B73 and PT haplotypes using microsatellite makers we have shown distinguish B73 and PT alleles in this region. HIFs will be self-pollinated to generate pairs of near-isogenic lines (NILs) homozygous for the B73 or PT haplotype. While different in the candidate region, NIL pairs will share a common genetic background outside this region, including a sizable (25%) contribution of PT, capturing potential epistatic effects important to expression of the candidate phenotype. NILs will be genotyped by GBS (**?**Glaubitz et al., 2014) both to confirm the extent of introgression at the Chr4 candidate locus and to characterize this shared background. A total of 6 HIF derived NIL pairs (i.e. 12 lines), will be characterized in our three field sites (Table 2) and evaluated for phenotypes described in (Figure 3). In each site, we will plant 3 replicate rows of our NILs and B73 and PT parents. Data will be analyzed broadly as described

in **??**, both treating the introgression region as a single block, or considering individual markers. f The expected outcomes of this objective are 1) estimation of the contribution of differences at the Chr4 candidate locus to variation in a number of important phenotypes, 2) determination of the degree of phenotypic plasticity with respect to highland and lowland environments, 3) identification of differences in phenotypic effect among NIL pairs, indicative of background dependent epistatic interaction among genes.

Aim 3.3 Allelic series for QTL of interest

As described above, the Chr4 candidate region is hypothesised to be an inversion polymorphism. While facilitating generation of test materials and assessment of the region as a block, the predicted lack of recombination will hamper downstream efforts to dissect phenotypic effects and fine map the loci involved. Consequently, we will generate a series of NILs by marker-assisted recurrent backcross to B73 using a collection of seven diverse donor varieties: 3 lowland haplotypes represented by the 2 lowland parents of our mapping populations Aim 1.1 and an inbred *parviglumis*; 4 highland haplotypes represented by the 2 highland parents of our mapping populations Aim 1.1 , an inbred *mexicana* and the Palomero Toluqueño haplotype segregating in our HIFs Aim 3.2. All 3 highland maize varieties are predicted to carry the *mexicana* inverted haplotype at the Chr4 candidate locus. Each of these parents either have resequenced genomes (Vielle-Calzada et al., 2009; Chia et al., 2012) or will be sequenced as part of this project in Aim 1. It is anticipated that this material will be phenotyped selectively in light of initial results generated in the early part of the project.

The expected outcomes of this objective are 1) Estimation of functional variation in the Chr4 candidate region among lowland and highland teosinte and landrace maize; 2) Dissection of the highland haplotype on the basis of phenotypic variation among NILs carrying the inverted form; 3) Generation and identification of material suitable for future fine mapping through crossing of genetically/functionally divergent inversions from NILs.

Aim 3.4 Gene expression

We will first assess the effects of high and low elevation environments on genome-wide expression differences to identify genes responsive to these environments. We will grow the 8 inbred lines which serve as parents of our allelic series analysis in Aim 3.3 in a highland and lowland field site Table 2. From each inbred we will sample leaf and root tissue from three plants at each of two time stages (seedling and flowering adult). Tissue will be flash frozen and sent to UC Davis for extraction and sequencing (multiplexed 12 individuals per lane of an Illumina HiSeq 2500) at the UC Davis Genome Center. Each individual will be barcoded, providing 3 biological replicates for each tissue/time/environment combination. We will assess differences in expression across environments and identify overlap between differentially expressed (DE) genes and QTL from Aim 1, loci showing selection identified in Aim 2, and introgressed regions showing phenotypic differences in **??**. These results will help narrow down potential candidate genes in QTL and serve as functional validation of loci showing population genetic evidence of selection in introgressed and admixed populations. The data will also allow investigation of the relationship between phenotypic plasticity and adaptive change (c.f. Rosas et al., 2013) via comparison of DE genes among environments for a single inbred to differences in DE genes among inbreds to ask whether genes showing a plastic response in unadapted material (lowland landraces, *parviglumis* B73) show constitutive response in adapted lines (*mexicana*,highland landraces).

Our second approach will be a targeted analysis of transcriptomic changes in the Chr4 NIL lines from Aim 3.3. Using NILs generated from each of the same 7 inbred donors (alongside an additional replicate of B73), we will evaluate shoot tissues of three plants sampled at seedling and flowering stage for each of the two genotypes (homozygous B73, homozygous donor). Samples will be extracted and sequenced as described above. These analyses will allow us to refine potential candidate loci within introgressed segments of our NILs, moving us closer to a functional characterization of observed phenotypic differences. Whole-transcriptome comparison to the donor transcriptomes will allow us to differentiate between cis and trans regulation of expression within the Chr4 region, and analysis of co-expression networks (c.f. Swanson-Wagner et al., 2012), will highlight the effects of introgressed genes on expression patterns in the rest of the genome, enabling us to begin to dissect the genetic pathways involved in adaptive highland traits.

The expected outcomes of this objective will be 1) Identification of candidate genes showing plastic differential expression within lines across environments 2) Identification of candidate genes showing differential expression among lines from different environments 3) Detailed information on the effects of introgressed segments on genome-wide expression.

Broader Impacts

Exchange Program

We propose an international student exchange program between the PIs in the U.S. and Senior Personnel at LANGEBIO in Mexico. Over the course of the grant, we propose to fund 10 graduate or undergraduate students for 3-month research internships in one of the collaborating laboratories. Students involved will participate in research projects directly relating to the research focus of the grant, including developing mapping populations, mapping traits, population genetic analysis, or analysis of next-generation data. The expectation is that such research will often lead to co-authorship on publications. Students will be asked to give two presentations, one to the host lab upon arrival, talking about the lab/university they came from and research there, and another to their host lab detailing their work over the 3-month period. Each of the PIs will participate, sending students to Mexico and/or accepting students from Mexico for internships. PI Ross-Ibarra will manage the program, as he is fluent in Spanish and has past experience with a similar exchange program (NSF 0922703). Over the last four years, his lab has hosted 6 Mexican students who have worked on various computational aspects of centromere evolution. Two of those students have earned authorship on a paper to be submitted later this year and one has gone on to a PhD program in the U.S.

Our goal is to involve students directly in research while at the same time fostering intercultural exchange and promoting future international research opportunities. It is particularly appropriate for the study of maize, a crop with significant cultural and economic impact in both Mexico and the U.S. Participating Mexican students will learn new analytical methods – especially computational management of large datasets – that can be introduced to their respective laboratories and peers. American exchange students will similarly benefit from experience with large field experiments and efforts to functionally characterize individual loci. The hope is that Mexican undergraduate students involved may be recruited to graduate programs in the U.S., ideally to work in the lab of one of the PIs, and that American undergraduate students will be exposed to international opportunities for research, graduate education, and collaboration.

Phenotyping workshop

The USDA-ARS group in Columbia has developed a streamlined phenotypic data collection system utilizing a handheld barcode device, barcoded plant tags, and barcoded phenotyping tools in order to maximize efficiency. We will host a phenotyping workshop in Columbia during each year of the grant. Through this workshop, Dr. Flint-Garcias state-of-the-art system will be transferred to other research institutions to aid in large-scale data collection. The phenotyping workshop will include topics on Experimental Design, setting up the FieldBook database, and Data Collection. Experimental design topics include understanding where variation comes from, how to control for environmental/field variability and experimental error; heritability and repeatability. The need for consistent data collection and high-throughput will be emphasized. FieldBook database setup topics include setting up Palm handheld users, locations, traits, projects, assigning plots to projects, assigning traits and measurements to projects, generating barcoded plant tags, and loading the program and trait groups to the Palm to prepare for data collection. Topics to be covered in Data Collection include data collection for specific traits related to local adaptation of interest to our group, synchronizing data from the palm with the desktop/laptop database, managing data conflicts between the palm and the database, running reports, and exporting data. This proposal will provide travel support for instructors. The workshop will be free but participants will be expected to purchase their own Palm handheld and pay for their own travel. The workshop will be held each year in late summer so that the participants can gain hands-on experience in data collection in the corn field.

Software

A good understanding of population and quantitative genetics is key to a students understanding of genetics and evolution, but these subjects are often conceptually quite difficult. An understanding of genetic variation and its phenotypic effects is also an increasingly important part of being an informed citizen, due to the rise of personal genomics and genomic medicine (e.g. Redfield, 2012). The large amount of population genetic and association data being generated offers a superb chance to motivate these subjects using real data. We will develop undergraduate teaching modules in population and quantitative genetics using data from this project. These modules will be tested and integrated into large undergraduate teaching courses (introductory evolutionary biology and genetics) at UC Davis and graduate courses at UC Davis and Iowa State (ecological genomics). We have already begun to develop and distribute some of these resources, e.g. genome-scale demonstrations of Hardy Weinberg Equilibrium (HWE) using human HapMap data. Such demonstrations underscore the usefulness of basic population genetics in describing real world patterns, and begin to expose students to the wealth of genomics data being collected. Other examples will include: using association data from our admixed populations to demonstrate quantitative genetics models; and explaining concepts of genetic and genealogical ancestry using genomic identity by descent. These modules will be prepared in the open source statistical program R, to ensure that they are easily used, modified, and distributed, and to expose students to programming in biology. The modules will be designed so that they can be tailored for use at a variety of levels from teaching basic concepts to large undergraduate classes to providing the raw data for programming exercises for upper division courses.

The modules will be publicly distributed via Github (see Data Management Plan) in a fully open manner. The use of github will allow others to modify and extend the modules and to share

and track these modifications.

Germplasm resources

This project will generate multiple germplasm resources. Seed from the F2 parents will allow additional use of this mapping population to study additional phenotypes of interest (e.g. root morphology and growth). Seed from our NIL populations will allowing investigation of genome-wide introgressions from a variety of exotic lines. Such material could be of interest to the Germplasm Enhancement of Maize (http://www.public.iastate.edu/ usda-gem/) project as well as to public and private breeders both in the US and abroad. In Mexico, for example, the highland niche represents a key target market for an emerging private sector of small breeding companies established following deregulation in 1990s. While highland adapted hybrids are available, these are largely derived from lowland sub-tropical material with little or no contribution of the highland landraces and the germplasm developed here could be an important contribution to furthering such programs. Finally, seed from our collections of teosinte will enhance the sampling of these subspecies and provide additional diversity not currently present in germplasm banks. Seed from our mapping populations will be deposited in the USDA-ARS Maize Stock Center at the University of Illinois, and backups will be kept at Iowa State and Missouri.

Results From Prior NSF Support

Ross-Ibarra, Flint-Garcia: #1238014: Biology of Rare Alleles in Maize and Its Wild Relatives

$13,311,185 ($2,368,767 to Ross-Ibarra and $1,206,211 to Flint-Garcia), 05/15/13-04/30/18. PI Edward Buckler, co-PIs J. Doebley, J. Holland, S. Flint-Garcia, Q. Sun, P. Bradbury, S. Mitchell, J. Ross-Ibarra

Intellectual merit In the first year we have developed accurate imputation approaches, found evidence for the importance of deleterious variants and non-genic polymorphisms in heterosis and GWAS, documented differences in recombination among the parents of the NAM population, and found population genetic evidence suggesting the importance of demography and purifying selection across the genome. The grant has produced 18 total publications in its first year (only publications involving PIs Flint-Garcia and Ross-Ibarra are shown below).

Broader impacts In the first year this project has included 10 postdoctoral and 12 graduate trainees. The GBS workshop and traveling maize exhibit continue to be popular and successful. A new version of the teacher-friendly guide to the evolution of maize has been revised and published online.

Publications Peiffer et al. (2013); Romay et al. (2013); Wills et al. (2013); Mezmouk and Ross-Ibarra (2014); Peiffer et al. (2014); Sood et al. (2014)

Ross-Ibarra: #0922703: Functional Genomics of Maize Centromeres

$5,008,031 ($754,409 to Ross-Ibarra). 09/01/09-08/31/14. PI Kelly Dawe, co-PIs J. Birchler, J. Jiang, G. Presting, J. Birchler, J. Ross-Ibarra

Intellectual merit Centromeres are regions of the genome that organize and regulate chromosome movement, yet the biology of centromeres remains poorly understood. Co-PI Ross-Ibarra's group

has focused in particular on the evolutionary genetics of centromeres. This work has demonstrated the remarkable evolutionary lability of centromere tandem repeats, but has shown that there is little evidence in maize for coevolution between centromere sequence and kinetochore proteins. Ongoing work from the Ross-Ibarra lab seeks to characterize kinetochore proteins, assess the phylogenetic evidence for longer-term coevolution, and understand patterns of centromere and genome size variation in natural populations.
Broader impacts Co-PI Ross-Ibarra has established and currently runs an international student exchange program as part of this grant. Data and result of this project have been disseminated via publications and presentations as well as deposited in the maize genetics community database `www.maizegdb.org`. Former trainees on the grant include Dr. Matthew Hufford (Co-PI on the current grant).
PublicationsShi et al. (2010); Chia et al. (2012); Fang et al. (2012); Hufford et al. (2012b,c, 2013); Melters et al. (2013); Kanizay et al. (2013); Pyhajarvi et al. (2013)

Coop: #1262645: Collaborative Research: ABI Innovation: Visualization And Statistics For Spatial Population Genomic Analysis.

$314,260, with an effective date of 05/01/13. Award Duration: 36 months.
Intellectual merit We are developing a set of spatial statistics methods based on Gaussian random fields for the analysis of geographic population genomics data. The first method based on this approach has just been published, allowing a sound statistical framework to distinguish the effects of geographic and ecological distance on genetic isolation.
Broader impacts The R package of the software has been released online, and has already been used by many molecular ecologists.
Publications Bradburd et al. (2013)

Flint-Garcia: #0820619: Genetic Architecture of Maize and Teosinte

$ 9,823,000. 3/1/2009-2/28/2013. PI Edward Buckler, co-PIs J. Doebley, T. Fulton, S. Flint-Garcia, J. Holland, S. Kresovich, M. McMullen, Qi Sun.
Intellectual merit This project extends over more than a decade, and has pioneered the characterization of population genetic and evolutionary parameters of maize diversity, developed resources to connect this genetic diversity to phenotype through both association and joint linkage-association mapping, conducted fine scale analysis of domestication and agronomic QTL, and recently expanded to whole-genome analysis of diversity, evolution, and phenotype. Overall, the maize diversity project has developed a wide range of approaches and broadened understanding of the maize genome, evolution and adaptation, genetic mapping, and the agricultural improvement of maize. The project successfully released and analyzed the maize Nested Association Mapping (NAM) population, collaborated on making first and second generation haplotype maps for maize, resolved domestication traits, developed a range of novel statistical approaches for association mapping, and dissected complex traits such as flowering time, kernel composition, disease resistance, height, and inflorescence and leaf morphology.
Broader impacts The outreach program included a traveling science museum exhibit on maize diversity, evolution and genetics (seen by at least 300,000 people at five venues to date, including the famous Corn Palace in South Dakota), online Teacher Friendly Guide to the Evolution of Maize, seven Genotyping-By-Sequencing (GBS) workshops (held at primarily at Cornell but has also been

held in Kenya), and training of postdocs, graduate students and undergraduates, the vast majority of which have continued in scientific careers. Former trainees on this grant include Dr. Flint-Garcia and Dr. Ross-Ibarra (PIs of the current grant), only their publications are shown below.
Publications Buckler et al. (2009a); Flint-Garcia et al. (2009a,b,c); Gore et al. (2009); McMullen et al. (2009); Ross-Ibarra et al. (2009); Bottoms et al. (2010); Dubois et al. (2010); Zhang et al. (2010); Van Heerwaarden et al. (2010); van Heerwaarden et al. (2010); Brown et al. (2011); Morrell et al. (2011); Studer et al. (2011); van Heerwaarden et al. (2011); Tian et al. (2011); Chia et al. (2012); Cook et al. (2012); Fang et al. (2012); Hufford et al. (2012c); Hung et al. (2012c,b); Romay et al. (2013)

Table 3: Proposed timeline of activities and responsibilities

Year	2015	2016	2017	2018	2019
Objective Aim 3.3 Allelic series	2015	2016	2017	2018	2019
Objective Aim 3.2 Fine mapping	–	RS, AC	AC,JRI	AC,JRI	–
Objective Aim 3.4 RNA-seq	2015	2016	2017	2018	2019
Objective Aim 2.3 Maize/mexicana introgression	2015	2016	2017	2018	2019
Objective Aim 1.2 Admix mapping	2015	2016	2017	2018	2019
Objective Aim 1.1 QTL mapping	2015	2016	2017	2018	2019
Objective Aim 2.1 Highland haplotypes	2015	2016	2017	2018	2019
Objective Aim 2.3 Admixture population genetics	2015	2016	2017	2018	2019

References Cited

JJ Berg and G Coop. The population genetic signature of polygenic local adaptation. *PLoS Genetics*, In press, 2014.

CA Bottoms, S Flint-Garcia, and MD McMullen. Iview: introgression library visualization and query tool. *BMC Bioinformatics*, 11 Suppl 6:S28, 2010.

GS Bradburd, PL Ralph, and GM Coop. Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution*, 67(11):3258–3273, 2013.

Y Brandvain, AM Kenney, L Flagel, G Coop, and A Sweigart. Speciation and introgression between mimulus nasutus and mimulus guttatus. *PLoS Genetics*, In press, 2014.

KW Broman, H Wu, S Sen, and GA Churchill. R/qtl: Qtl mapping in experimental crosses. *Bioinformatics*, 19(7):889–890, 2003.

PJ Brown, N Upadyayula, GS Mahone, F Tian, PJ Bradbury, S Myles, JB Holland, S Flint-Garcia, MD McMullen, ES Buckler, and TR Rocheford. Distinct genetic architectures for male and female inflorescence traits of maize. *PLoS Genetics*, 7(11):e1002383, 2011.

ES Buckler, JB Holland, PJ Bradbury, CB Acharya, PJ Brown, C Browne, E Ersoz, S Flint-Garcia, A Garcia, JC Glaubitz, MM Goodman, C Harjes, K Guill, DE Kroon, S Larsson, NK Lepak, H Li, SE Mitchell, G Pressoir, JA Peiffer, MO Rosas, TR Rocheford, MC Romay, S Romero, S Salvo, H Sanchez Villeda, HS da Silva, Q Sun, F Tian, N Upadyayula, D Ware, H Yates, J Yu, Z Zhang, S Kresovich, and MD McMullen. The genetic architecture of maize flowering time. *Science*, 325(5941):714–718, 2009a.

ES Buckler, JB Holland, PJ Bradbury, CB Acharya, PJ Brown, C Browne, E Ersoz, S Flint-Garcia, A Garcia, JC Glaubitz, et al. The genetic architecture of maize flowering time. *Science*, 325 (5941):714–718, 2009b.

JM Chia, C Song, PJ Bradbury, D Costich, N de Leon, J Doebley, RJ Elshire, B Gaut, L Geller, JC Glaubitz, M Gore, KE Guill, J Holland, MB Hufford, J Lai, M Li, X Liu, Y Lu, R McCombie, R Nelson, J Poland, BM Prasanna, T Pyhajarvi, T Rong, RS Sekhon, Q Sun, MI Tenaillon, F Tian, J Wang, X Xu, Z Zhang, SM Kaeppler, J Ross-Ibarra, MD McMullen, ES Buckler, G Zhang, Y Xu, and D Ware. Maize hapmap2 identifies extant variation from a genome in flux. *Nat Genet*, 44(7):803–807, 2012.

JP Cook, MD McMullen, JB Holland, F Tian, P Bradbury, J Ross-Ibarra, ES Buckler, and SA Flint-Garcia. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol*, 158(2):824–834, 2012.

G Coop, D Witonsky, A Di Rienzo, and JK Pritchard. Using environmental correlations to identify loci underlying local adaptation. *Genetics*, 185(4):1411–1423, 2010.

JF Doebley. Maize introgression into teosinte-a reappraisal. *Annals of the Missouri Botanical Garden*, pages 1100–1113, 1984.

PG Dubois, GT Olsefski, S Flint-Garcia, TL Setter, OA Hoekenga, and TP Brutnell. Physiological and genetic characterization of end-of-day far-red light response in maize seedlings. *Plant Physiol*, 154(1):173–186, 2010.

RJ Elshire, JC Glaubitz, Q Sun, JA Poland, K Kawamoto, ES Buckler, and SE Mitchell. A robust, simple genotyping-by-sequencing (gbs) approach for high diversity species. *PLoS One*, 6(5): e19379, 2011.

Z Fang, T Pyhajarvi, AL Weber, RK Dawe, JC Glaubitz, J Gonzalez Jde, C Ross-Ibarra, J Doebley, PL Morrell, and J Ross-Ibarra. Megabase-scale inversion polymorphism in the wild ancestor of maize. *Genetics*, 191(3):883–894, 2012.

SA Flint-Garcia, AL Bodnar, and MP Scott. Wide variability in kernel composition, seed characteristics, and zein profiles among diverse maize inbreds, landraces, and teosinte. *Theor Appl Genet*, 119(6):1129–1142, 2009a.

SA Flint-Garcia, ES Buckler, P Tiffin, E Ersoz, and NM Springer. Heterosis is prevalent for multiple traits in diverse maize germplasm. *PLoS One*, 4(10):e7433, 2009b.

SA Flint-Garcia, KE Dashiell, DA Prischmann, MO Bohn, and BE Hibbard. Conventional screening overlooks resistance sources: rootworm damage of diverse inbred lines and their b73 hybrids is unrelated. *J Econ Entomol*, 102(3):1317–1324, 2009c.

A Fournier-Level, A Korte, MD Cooper, M Nordborg, J Schmitt, and AM Wilczek. A map of local adaptation in arabidopsis thaliana. *Science*, 334(6052):86–89, 2011.

K Fukunaga, J Hill, Y Vigouroux, Y Matsuoka, J Sanchez, KJ Liu, ES Buckler, and J Doebley. Genetic diversity and population structure of teosinte. *GENETICS*, 169(4):2241–2254, 2005.

AJ Geneva, CA Muirhead, and LAM Lovato. An improved sequence measure used to scan genomes for regions of recent gene flow. *arXiv preprint: 1403.1552*, 2014.

JC Glaubitz, TM Casstevens, F Lu, J Harriman, RJ Elshire, Q Sun, and ES Buckler. Tassel-gbs: a high capacity genotyping by sequencing analysis pipeline. *PLoS One*, 9(2):e90346, 2014.

MA Gore, JM Chia, RJ Elshire, Q Sun, ES Ersoz, BL Hurwitz, JA Peiffer, MD McMullen, GS Grills, J Ross-Ibarra, DH Ware, and ES Buckler. A first-generation haplotype map of maize. *Science*, 326(5956):1115–1117, 2009.

T Günther and G Coop. Robust identification of local adaptation from allele frequencies. *Genetics*, 195(1):205–220, 2013.

AM Hancock, B Brachi, N Faure, MW Horton, LB Jarymowycz, FG Sperone, C Toomajian, F Roux, and J Bergelson. Adaptation to climate across the arabidopsis thaliana genome. *Science*, 334 (6052):83–86, 2011.

Matthew B Hufford, Enrique Martínez-Meyer, Brandon S Gaut, Luis E Eguiarte, and Maud I Tenaillon. Inferences from the historical distribution of wild and domesticated maize provide ecological and evolutionary insight. *PloS one*, 7(11):e47659, 2012a.

MB Hufford, P Bilinski, T Pyhajarvi, and J Ross-Ibarra. Teosinte as a model system for population and ecological genomics. *Trends Genet*, 28(12):606–615, 2012b.

MB Hufford, X Xu, J van Heerwaarden, T Pyhajarvi, JM Chia, RA Cartwright, RJ Elshire, JC Glaubitz, KE Guill, SM Kaeppler, J Lai, PL Morrell, LM Shannon, C Song, NM Springer, RA Swanson-Wagner, P Tiffin, J Wang, G Zhang, J Doebley, MD McMullen, D Ware, ES Buckler, S Yang, and J Ross-Ibarra. Comparative population genomics of maize domestication and improvement. *Nat Genet*, 44(7):808–811, 2012c.

MB Hufford, P Lubinksy, T Pyhajarvi, MT Devengenzo, NC Ellstrand, and J Ross-Ibarra. The genomic signature of crop-wild introgression in maize. *PLoS Genetics*, 9(5):e1003477, 2013.

H-Y Hung, LM Shannon, F Tian, PJ Bradbury, C Chen, SA Flint-Garcia, MD McMullen, Doreen Ware, ES Buckler, JF Doebley, et al. Zmcct and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proceedings of the National Academy of Sciences*, 109(28):E1913–E1921, 2012a.

HY Hung, C Browne, K Guill, N Coles, M Eller, A Garcia, N Lepak, S Melia-Hancock, M Oropeza-Rosas, S Salvo, N Upadyayula, ES Buckler, S Flint-Garcia, MD McMullen, TR Rocheford, and JB Holland. The relationship between parental genetic or phenotypic divergence and progeny variation in the maize nested association mapping population. *Heredity (Edinb)*, 108(5):490–499, 2012b.

HY Hung, LM Shannon, F Tian, PJ Bradbury, C Chen, SA Flint-Garcia, MD McMullen, D Ware, ES Buckler, JF Doebley, and JB Holland. Zmcct and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proc Natl Acad Sci U S A*, 109(28):E1913–E1921, 2012c.

LB Kanizay, T Pyhajarvi, EG Lowry, MB Hufford, DG Peterson, J Ross-Ibarra, and RK Dawe. Diversity and abundance of the abnormal chromosome 10 meiotic drive complex in zea mays. *Heredity (Edinb)*, 110(6):570–577, 2013.

N Lauter, C Gustus, A Westerbergh, and J Doebley. The inheritance and evolution of leaf pigmentation and pubescence in teosinte. *Genetics*, 167(4):1949–1959, 2004.

V Le Corre and A Kremer. The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, 21(7):1548–1566, 2012.

David B. Lobell, Marianne Bnziger, Cosmos Magorokosho, and Bindiganavile Vivek. Nonlinear heat effects on african maize as evidenced by historical yield trials. *Nature Climate change*, 1(1): 42–45, 2011a.

DB Lobell, W Schlenker, and J Costa-Roberts. Climate trends and global crop production since 1980. *Science*, 333(6042):616–620, 2011b.

Y Matsuoka, Y Vigouroux, MM Goodman, G J Sanchez, E Buckler, and J Doebley. A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci U S A*, 99(9):6080–6084, 2002.

MD McMullen, S Kresovich, HS Villeda, P Bradbury, H Li, Q Sun, S Flint-Garcia, J Thornsberry, C Acharya, C Bottoms, P Brown, C Browne, M Eller, K Guill, C Harjes, D Kroon, N Lepak, SE Mitchell, B Peterson, G Pressoir, S Romero, M Oropeza Rosas, S Salvo, H Yates, M Hanson, E Jones, S Smith, JC Glaubitz, M Goodman, D Ware, JB Holland, and ES Buckler. Genetic properties of the maize nested association mapping population. *Science*, 325(5941):737–740, 2009.

DP Melters, KR Bradnam, HA Young, N Telis, MR May, JG Ruby, R Sebra, P Peluso, J Eid, D Rank, JF Garcia, JL Derisi, T Smith, C Tobias, J Ross-Ibarra, I Korf, and SW Chan. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol*, 14(1):R10, 2013.

S Mezmouk and J Ross-Ibarra. The pattern and distribution of deleterious mutations in maize. *G3 (Bethesda)*, 4(1):163–171, 2014.

C Monfreda, N Ramankutty, and JA Foley. Farming the planet: 2. geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global biogeochemical cycles*, 22(1), 2008.

PL Morrell, ES Buckler, and J Ross-Ibarra. Crop genomics: advances and applications. *Nat Rev Genet*, 13(2):85–96, 2011.

R Nielsen, S Williamson, Y Kim, MJ Hubisz, AG Clark, and C Bustamante. Genomic scans for selective sweeps using snp data. *Genome research*, 15(11):1566–1575, 2005.

JA Peiffer, SA Flint-Garcia, N De Leon, MD McMullen, SM Kaeppler, and ES Buckler. The genetic architecture of maize stalk strength. *PloS one*, 8(6):e67066, 2013.

JA Peiffer, MC Romay, MA Gore, SA Flint-Garcia, Z Zhang, MJ Millard, CA Gardner, MD McMullen, JB Holland, PJ Bradbury, and ES Buckler. The genetic architecture of maize height. *Genetics*, 2014.

AL Price, A Tandon, N Patterson, KC Barnes, N Rafaels, I Ruczinski, TH Beaty, R Mathias, D Reich, and S Myers. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genetics*, 5(6):e1000519, 2009.

T Pyhajarvi, MB Hufford, S Mezmouk, and J Ross-Ibarra. Complex patterns of local adaptation in teosinte. *Genome Biol Evol*, 5(9):1594–1609, 2013.

P Ralph and G Coop. Parallel adaptation: one or many waves of advance of an advantageous allele? *Genetics*, 186(2):647–668, 2010.

RJ Redfield. Why do we have to learn this stuff?a new genetics for 21st century students. *PLoS Biology*, 10(7):e1001356, 07 2012. doi: 10.1371/journal.pbio.1001356. URL `http://dx.doi.org/10.1371%2Fjournal.pbio.1001356`.

JC Reif, ML Warburton, XC Xia, DA Hoisington, J Crossa, S Taba, J Muminović, M Bohn, M Frisch, and AE Melchinger. Grouping of accessions of mexican races of maize revisited with ssr markers. *Theoretical and Applied Genetics*, 113(2):177–185, 2006.

MC Romay, MJ Millard, JC Glaubitz, JA Peiffer, KL Swarts, TM Casstevens, RJ Elshire, CB Acharya, SE Mitchell, SA Flint-Garcia, MD McMullen, JB Holland, ES Buckler, and CA Gardner. Comprehensive genotyping of the usa national maize inbred seed bank. *Genome Biol*, 14(6):R55, 2013.

U Rosas, A Cibrian-Jaramillo, D Ristova, JA Banta, ML Gifford, AH Fan, RW Zhou, GJ Kim, G Krouk, KD Birnbaum, MD Purugganan, and GM Coruzzi. Integration of responses within and across arabidopsis natural accessions uncovers loci controlling root systems architecture. *Proceedings of the National Academy of Sciences*, 2013. doi: 10.1073/pnas.1305883110. URL `http://www.pnas.org/content/early/2013/08/22/1305883110.abstract`.

J Ross-Ibarra, M Tenaillon, and BS Gaut. Historical divergence and gene flow in the genus zea. *Genetics*, 181(4):1399–1413, 2009.

J Shi, SE Wolf, JM Burke, GG Presting, J Ross-Ibarra, and RK Dawe. Widespread gene conversion in centromere cores. *PLoS Biol*, 8(3):e1000327, 2010.

S Sood, S Flint-Garcia, MC Willcox, and JB Holland. Mining natural variation for maize improvement: Selection on phenotypes and genes. In *Genomics of Plant Genetic Resources*, pages 615–649. Springer, 2014.

GL Stebbins. Variation and evolution in plants. *Variation and evolution in plants.*, 1950.

A Studer, Q Zhao, J Ross-Ibarra, and J Doebley. Identification of a functional transposon insertion in the maize domestication gene tb1. *Nat Genet*, 43(11):1160–1163, 2011.

R Swanson-Wagner, R Briskine, R Schaefer, MB Hufford, J Ross-Ibarra, CL Myers, P Tiffin, and NM Springer. Reshaping of the maize transcriptome by domestication. *Proceedings of the National Academy of Sciences*, 2012. doi: 10.1073/pnas.1201961109. URL `http://www.pnas.org/content/early/2012/06/28/1201961109.abstract`.

MI Tenaillon and A Charcosset. A european perspective on maize history. *Comptes rendus biologies*, 334(3):221–228, 2011.

F Tian, PJ Bradbury, PJ Brown, H Hung, Q Sun, S Flint-Garcia, TR Rocheford, MD McMullen, JB Holland, and ES Buckler. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet*, 43(2):159–162, 2011.

MR Tuinstra, G Ejeta, and PB Goldsbrough. Heterogeneous inbred family (hif) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. *Theoretical and Applied Genetics*, 95(5-6):1005–1011, 1997.

J Van Heerwaarden, J Ross-Ibarra, J Doebley, JC Glaubitz, J Gonzalez Jde, BS Gaut, and LE Eguiarte. Fine scale genetic structure in the wild ancestor of maize (zea mays ssp. parviglumis). *Mol Ecol*, 19(6):1162–1173, 2010.

J van Heerwaarden, FA van Eeuwijk, and J Ross-Ibarra. Genetic diversity in a crop metapopulation. *Heredity (Edinb)*, 104(1):28–39, 2010.

J van Heerwaarden, J Doebley, WH Briggs, JC Glaubitz, MM Goodman, J de Jesus Sanchez Gonzalez, and J Ross-Ibarra. Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc Natl Acad Sci U S A*, 108(3):1088–1092, 2011.

JP Vielle-Calzada, O Martinez de la Vega, G Hernandez-Guzman, E Ibarra-Laclette, C Alvarez-Mejia, JC Vega-Arreguin, B Jimenez-Moraila, A Fernandez-Cortes, G Corona-Armenta, L Herrera-Estrella, and A Herrera-Estrella. The palomero genome suggests metal effects on domestication. *Science*, 326(5956):1078, 2009.

BF Voight, S Kudaravalli, X Wen, and JK Pritchard. A map of recent positive selection in the human genome. *PLoS Biology*, 4(3):e72, 2006.

DM Wills, CJ Whipple, S Takuno, LE Kursel, LM Shannon, J Ross-Ibarra, and JF Doebley. From many, one: Genetic control of prolificacy during maize domestication. *PLoS Genetics*, 9(6): e1003604, 2013.

CA Winkler, GW Nelson, and MW Smith. Admixture mapping comes of age*. *Annual review of genomics and human genetics*, 11:65–89, 2010.

JB Yoder, J Stanton-Geddes, P Zhou, R Briskine, ND Young, and P Tiffin. Genomic signature of adaptation to climate in medicago truncatula. *Genetics*, 2014. doi: 10.1534/genetics.113.159319. URL `http://www.genetics.org/content/early/2014/01/15/genetics.113.159319.abstract`.

N Zhang, A Gur, Y Gibon, R Sulpice, S Flint-Garcia, MD McMullen, M Stitt, and ES Buckler. Genetic analysis of central carbon metabolism unveils an amino acid substitution that alters maize nad-dependent isocitrate dehydrogenase activity. *PLoS One*, 5(4):e9991, 2010.

Biographical Sketch: Your Name

Maximum of 2 pages

Biographical Sketch

Professional Preparation

Undergraduate Institution(s)	Major	Degree	Year
Graduate Institution(s)	Major	Degree	Year
Postdoctoral Institution(s)	Area		Year
Year-present	Position, Department, Institution		
Year(s)	Position, Department, Institution		

Publications

Five Publications Most Closely Related to the Proposed Project

1. Author(s): Article Title, *Journal Title* **Volume Number**, Page Numbers, Year of Publication.
2.
3.
4.
5.

Ten Other Significant Publications

1.
2.
3.
4.
5.
6.
7.
8.
9.

10.

1.

2.

3.

4.

5.

Collaborators & Other Affiliations

Collaborators:

Graduate and Postdoctoral Advisors:

Thesis Advisor and Postgraduate-Scholar Sponsor:

Budget Justification

Personnel

No funding is requested for the PI, Co-PIs, or any Senior Personnel.

Other Personnel

Graduate students Funds are requested to support two graduate students each for 6 months during the academic year for each year of the project. At UC Davis, the current pay rate for doctoral students at 50% FTE is $27,319 during the academic year. Included is the estimated annual salary increase of 3%. The two students will be working on analysis of GBS data in the introgression and admix population genetic sections of AIM2, and will likely help with QTL analysis and sequencing in Aim1, and potentially RNA-seq analysis in Aim 3.

Technician Funds are requested for the first three years of the grant for a 50% time technician (Laboratory Assistant III) to extract DNA and RNA, prepare genomic and transcriptomic sequencing libraries, and perform root chilling experiments. The salary for this positions is set at $36,000 ($18,000 for 50% time), with an annual increase of 5%.

Fringe Benefits

Fringe benefits are applied to personnel salaries using the university approved rates:

- Graduate students - 1.3% for all years.
- Technician - 50.4%(1/1/2015-6/31/2015), 53.4%(6/31/2015-6/31/2016), 55.7%(6/31/2016-6/31/2017), 57.3%(6/31/2017-12/31/2017)

Equipment

No equipment funds are requested.

Travel

Travel for the PI and Co-PI Coop and one student the postdoc to 1 domestic conference each year is budgeted at $3,000. Travel for one of the Senior Personnel or CoPIs to participate in the field workshop is budgeted at $1,000 each year.

Travel for Senior Personnel and members of their group to manage field experiments and phenotype is budgeted at $12,000 each of the first 3 years. Travel for both Senior Personnel to 1 international conference each year is budgeted at $3,000 per year.

Participant Support

Our exchange program proposes to exchange two students per year between the US and Mexico. We are requesting funds to pay for 2 exchange students per year of the grant. These funds will cover student subsistence ($1,800 a month to include housing and subsistence) for 3 months, visa costs ($500), and round-trip travel to Mexico ($2,000).

Other Direct Costs

Materials and Supplies: In each of the first three years of the grant, $15,000 is requested in materials and supplies. $10,000 of this is for laboratory supplies for PI Ross-Ibarra for library prep for whole genome sequencing, RNA sequencing, and DNA extraction and preparation for GBS. This also includes funds for supplies for root chilling experiments to be done at UC Davis. In each of the five years, $2,500 is budgeted for standard office supplies, computer supplies (extra storage for our cluster, backup drives for lab members), and other miscellaneous expenses for Co-PI Coop and PI Ross-Ibarra.

Whole genome sequencing : The genomes of each of the four parental lines of our QTL mapping populations will be resequenced to a depth of 20-30X using 2 lanes of paired end 150bp reads on an Illumina HiSeq 2500. Current lane costs are approximately $2,200 per lane, and library preparations costs are approximately $100, for a total cost of $18,000.

GBS : Genotyping-by-sequencing will be performed for our introgression and admixture population genetic analyses. GBS will be performed at the Institute for Genomic Diversity at Cornell. Current prices are $60 per sample to run samples at 48-plex. We will genotype 360 individuals for our introgression analysis in year 1 for a cost of $21,600, and 144 individuals in year 2 for a cost of $8,640.

RNA sequencing : In total, RNA sequencing will be performed on 192 individuals (8 inbreds x 2 stages x 2 tissues x 2 environments x 3 replicates + 8 NILs x 2 genotypes x 2 stages x 2 environs x 3 replicates). Cost to prepare RNA libraries in our lab are approximately $100 per library, and sequencing costs for single-end 50bp reads at the UCD Genome Center are approximately $1,000 per lane. Multiplexing 12 barcodes per lane, this comes out to 32 lanes of sequence and a total cost of $70,400.

Field fees: Fees for the field experiments in our highland and lowland field sites 2 are approximately $60,000 the first three years of the experiment to allow development of the mapping populations and two replicates of the phenotyping. These fees include land rental and basic management (planting, watering, weeding, fertilizing), as well as station fees to hire manual labor for phenotyping. These fees decrease to $10,000 in the last two years of the proposal as subsequent field experiments including evaluation of NILs and RNA-seq lines, will be considerably smaller. Field fees total $200,000 across the five years of the grant.

Graduate Student Tuition: Tuition for graduate students is charged to the project in proportion to the amount of effort the graduate student will work on the project. For a graduate student employed on the project for 9 academic months at 50% FTE, the tuition charge is $31,546 in FY 2015 to account for out-of-state tuition, $17,266 in FY 2016 and increasing 5% each subsequent year.

Publication Costs: In year two $1,500 is requested for publication fees to an open access journal. In subsequent years $3,000 is requested annually.

Total Direct Costs

Total direct costs for UCD come to $874,643. Subawards to USDA-ARS and Iowa State su to $1,218,560.

Indirect Costs

Indirect costs are calculated on Modified Total Direct Costs (Total Direct costs less graduate student fees and participant support and subaward funding beyond the first $25,000) using F&A rates approved by US Department of Health and Human Services. For this project, F&A rates of 55.5% were used from Jan. 1, 2015 through June 30, 2015, 56.5% from July 1, 2015 through June 30, 2016, and 57% from July 1, 2016 until the end of the project.

Facilities, Equipment, and Other Resources

Facilities, Equipment & Other Resources

UC Davis

Dr. Ross-Ibarra has four standard laboratory benches as part of a shared lab space at UCD. The shared space is the single largest lab space on campus, and provides for seamless interaction between the labs housed there. The space currently houses three other PIs, all working on the genetics and genomics of economically important plant taxa (Dubcovsky, Neale, Dandekar). The lab is equipped with standard equipment and tools for molecular biology, including freezers and refrigeration, a shared liquid handling robot, thermal cyclers, centrifuges, gel rigs, balances, and standard molecular biology supplies. A dedicated low-humidity refrigerator for seed storage is available through the university, and low-humidity storage cabinets for tissues and temporary seed storage are in the laboratory. Dr. Ross-Ibnarraoccupies half of a large office suite that includes a conference room and cubicle space for 25 people. Both macintosh and PC workstations are available for student and postdoc employees. The PI is a contributing partner in a large computer cluster, giving the lab dedicated access to 192 processors, with the opportunity for use of nearly 800 additional CPU as resources allow. Recent (2013) additions to the cluster have provided it with additional CPU as well as six new shared high-memory (512Gb RAM) nodes, one of which is dedicated to the Ross-Ibarra lab. Dr. Ross-Ibarra is a faculty member of the UC Davis Genome Center, a large facility that includes bioinformatics, genotyping, metabolomics, proteomics, and expression analysis cores able to perform a variety of genomics analyses at cost for UC Davis faculty. The Genome Center also rents time on its equipment, including a bioanlyzer and library preparation robots. As a member of the Genome Center, Dr. Ross-Ibarra also has access to their additional computational facilities. UC Davis has also entered into a recent partnership with BGI (the Beijing Genomics Institute) to provide additional high-throughput sequencing services via a new Sacramento-based sequencing facility.

Dr. Coops dry space is located on the 3rd floor of the Storer building, which houses the Department of Evolution and Ecology. The space is newly renovated space and consists of 3 offices that can seat a 8 total of people, and a conference room. In addition members of the lab have access to an additional conference room and other offices shared with the Begun, Langley, Lott, Kopp and Turelli groups. This group is part of the larger Center and Graduate Group for Population Biology, one of the leading graduate training programs in ecology and evolution in the world. Each current member of Dr. Coops group has a quad-core Mac pro. The computers are loaded with all the necessary software (Word, R, Mathematica etc) and are connected to the university network as well as to color and black and white printers. The Coop lab has access to the genome center computational facilities: http://www.genomecenter.ucdavis.edu/core-facilities/.

Iowa State

Project components completed in the Hufford Laboratory will include mapping population development, DNA isolation and PCR, and population genetic analysis of genotyping data. Population development will be carried out in field space available at the Curtiss Farm of Iowa State University (ISU). This facility is equipped with irrigation, tractors, tillage equipment, planters, and combines. Seed processing and cold storage facilities are also available on the ISU campus. The Hufford Laboratory has all equipment necessary for DNA isolation and PCR including centrifuges, thermal cyclers, an ultra-low freezer, water baths, a pH meter, balances, and an electrophoresis system. A gel imaging system and a NanoDrop spectrophotometer for DNA quantification are

accessible through the Center for Plant Responses to Environmental Stresses at ISU. The DNA Facility at ISU provides access to cutting-edge genomic technology including HiSeq and MiSeq Illumina sequencing and library preparation for both paired-end and mate-pair approaches. Data analyses will be carried out using the High Performance Computing clusters available at ISU. Dr. Hufford currently has access to the Lightning3 cluster which has a mix of Opteron based servers, consisting of 18 SuperMicro servers with core counts ranging from 32 to 64 and 256 to 512 GB of memory.

U. Missouri Dr. Flint-Garcia has 600 sq ft of laboratory space in Curtis Hall, on the University of Missouri campus. The laboratory is fully equipped for molecular genetics, including a chemical hood, a Beckman table top centrifuge with multiple tube buckets, a Tetrad four plate thermalcycler, several freezers, ultra-low freezers and refrigerators, water baths, a pH meter, and balances. In the building, laboratory personnel have ready access to ultracentrifuges and rotors, growth chambers, an autoclave, lyophilizers, a Sorvall high speed preparative centrifuge with four rotors, a shaker-incubator for bacterial cultures, a chromatography cabinet, electrophoresis equipment for DNA, RNA protein and DNA sequence analysis, a plate reading spectrophotometer/flourometer, a pulse-field electrophoresis system, six Thermolyne thermalcyclers, and four Tetrad four plate thermalcyclers. Dr. Flint-Garcia has multiple personal computers, and computing resources including weekly data backups, direct access to a Sun Ultra10 Unix Workstation and NT server for data sharing, and IT support from USDA-ARS. In addition, the co-PI has access to the Lewis bioinformatics cluster (over 180 compute nodes with more than 1200 processor cores and 5400 GB of memory) via the University of Missouri Bioinformatics Core Facility. Dr. Flint-Garcia has 120 sq ft of office space and ample office and desk space for postdocs, technicians and graduate students. Dr. Flint-Garcia shares two ABI 3100 DNA sequencers, an ABI 7900HT RTPCR machine, and a Beckman NxP robot used primarily for DNA extractions with Mel Oliver and Mike McMullen, and other USDA scientists in the unit. Dr. Flint-Garcia has access to greenhouse and field space (with irrigation capability; University of Missouri South Farm and Bradford Research Center), seed processing and cold storage space, and use of winter nursery facilities in Puerto Rico. The co-PI has access to a complete set of field equipment including multiple tractors, tillage equipment, a 4-row plot planter, and a 2-row plot combine.

LANGEBIO

Langebios mandate is to conduct top-ranked research while promoting genomic knowledge for the protection and sustainable use of Mexican biodiversity. Its unique location in the agricultural center of Mexico facilitates field sampling and field experimentation. We have ample experience growing maize in nurseries located on the West Coast (Valle de Banderas, Nayarit), in Central Mexico (Irapuato; Celaya, Guanajuato), and have begun to establish additional sites in the high valleys of Central Mexico (Queretero; Estado de Mexico). We regularly conduct field expeditions to collect plants in both the dry regions of Northern Mexico (maize collections in Chihuahua, Lamiaceae throughout the Northeast) and the lower valleys of the Eje Volcanico and Costa del Pacifico (Teocintle and maize, Solanaceae, and Cucurbitaceae). Research at Langebio is supported by greenhouse facilities and two service units: Genomics and Mass Spectrometry, both of them equipped with state-of-the-art instrumentation, including several next-generation sequencing machines and diverse mass spectrometry equipments. Other facilities include a computation cluster and a specialized clean room for ancient DNA analysis.

Supplementary Documentation

Data Management Plan

Data Types

This proposal will generate sequence data, genotype, phenotype data, analytical software, teaching resources, germplasm, and publications.

Data Access, Sharing

Sequence data of the parental lines will be deposited to NCBI sequence read archive (SRA) along with passport information on each parent.

Phenotypic data and genotypes from sequencing and GBS will be uploaded to Figshare, where it can be associated with other data (publications, links to germplasm, SRA, code). Data will be grouped into projects, and each project is associated with a unique digital object identifier (DOI). PIs Ross-Ibarra and Coop have already used figshare extensively to share and archive data, preprints, and code (see `http://figshare.com/authors/Jeffrey_Ross-Ibarra/98899` and `http://figshare.com/authors/Graham_Coop/101524`). Data on figshare is publicly available and searchable.

Analytical software and code from this project will be hosted on github, a version-controlled public git repository. Upon submission of papers all code will be made publicly available. PIs Ross-Ibarra and Coop have already done this extensively (see `https://github.com/rossibarra`, `https://github.com/rilab`, and `https://github.com/cooplab`). Publication of all code will ensure reproducibility of all analyses conducted.

All appropriate metadata including plant ID, data collector, sequence run, field location, etc. will be associated with genotype and phenotype data deposited to figshare.

Presentations and teaching resources from our field workshop will also be made publicly available via the Slideshare website.

All publications resulting from this project will be submitted to one or more preprint servers (e.g. arXiv, bioRxiv, PeerJ) such that they will be publicly available immediately upon submission of the paper for publication.

All data, code, and presentations will be made publicly available via a creative commons CC by 2.0 license (http://creativecommons.org/licenses/by/2.0/) allowing free access to reuse, redistribute, and modify, requiring only citation of the license and the original source.

Data Archiving

All data, code, presentations, and publications will be made publicly available online (see above). Prior to public release, all data will be hosted locally. PI Ross-Ibarra will maintain a backup of all raw genotyping, sequence, and phenotyping data. His lab maintains a DROBO distributed backup server (currently ¿8Tb of free space) which is robust to single disk failure. All analytical code will be hosted on github, which maintains version-controlled backups, as private repositories until release.

Seed will be maintained in climate-controlled conditions at Iowa State. International agreements prohibit some of the maize and teosinte germplasm collected from being stored by USDA. We will

deposit small quantities of seed from all our collections with the CIMMYT germplasm bank in Mexico, and deposit samples of our mapping populations in the USDA-ARS Maize Stock Center at the University of Illinois. Both centers provide public access to seed.

Postdoctoral Researcher Mentoring Plan:

The current proposal requests funding for two postdoctoral researchers, one each at Iowa State and USDA-ARS in Columbia. Nonetheless, we expect additional postdocs to join the group via alternative funding opportunities (fellowships, etc.) and anticipate that postdocs in the labs of all the PIs may collaborate to a greater or lesser degree on this project. Much of our thinking on postdoctoral mentoring comes directly from our own mentorship experience – PIs Flint-Garcia, Hufford, and Ross-Ibarra were all postdoctoral scholars on funded NSF programs. For this project, the PI at each institution will act as mentor and supervisor for each postdoc, holding regular weekly meetings to assess progress and set goals. One clear goal will be first authorship on submitted papers, with the expectation of approximately one first author paper per year of duration of the postdoc.

Interaction and experience presenting and discussing science will be highly encouraged. All groups will have internal lab meetings (the Coop and Ross-Ibarra labs at UC Davis already hold joint lab meetings) at which postdocs and graduate students will be given numerous opportunities to hone their presentation skills. The Coop, Ross-Ibarra and Hufford labs currently host weekly journal clubs in which postdocs gain additional training in reading, presenting, and dissecting scientific literature. Members of the Ross-Ibarra and Flint-Garcia labs also attend a weekly journal club as part of another collaborative project (NSF #1238014). In addition, we will organize a monthly group meeting via web-conference in which one lab member presents on their research progress. UC Davis has a ReadyTalk license allowing inexpensive web-conference hosting. Finally, all of our institutions have seminar series specifically for postdoctoral and graduate students to practice presentation skills; members of our labs will be encouraged to attend these.

Postdocs will be encouraged to write and apply for external funding, including fellowships and grant proposals. Both the Ross-Ibarra and Coop labs have a documented history of successful funding with postdoctoral scholars as Co-PIs, providing valuable training (and even initial funding) for the scholars' future academic careers.

Postdocs in the Hufford and Flint-Garcia labs will take part as trainers in the annual phenotyping workshop under supervision of CoPI Flint-Garcia. This will provide additional training in high-throughput phenotyping as well as valuable teaching experience.

Finally, postdocs will be encouraged to take advantage of professional development programs offered by their local institutions. All of our institutions have infrastructure in place for professional development of postdocs and offer training in responsible conduct of research, grantsmanship, mentoring, career development, authorship of journal papers, and teaching.