Project Summary

Maximum of 1 page

Intellectual Merit

Broader Impacts

Project Description

Rationale and Significance

While future climates will undoubtedly differ in a number of important ways, we focus in this proposal on adaptation to changing temperature. Neither of the other major changes predicted as a result of climate change elevated CO2 and changes in rainfall are likely to be as important for maize yields. Though elevated CO2 may actually prove beneficial for yields of some crops, models suggest the effect on maize, a C4 plant, to be relatively minor (Lobell et al. 2011a). Similarly, analyses of past climate change find significant negative effects of temperature on maize yield but much weaker effects of precipitation (Lobell et al. 2003, Roberts and Schlenker 2009, 2010). Increasing temperature, on the other hand, is likely to have strong, nonlinear effects on yield (Lobell et al. 2011b, Schlenker and Roberts 2009), as temperature extremes have an inordinate impact on overall yield. Improved breeding methods Novel approaches to molecular plant breeding have been proposed as a solution to changing climes (Takeda and Matsuoka 2008), and some models suggest that breeding or other technological advances could meaningfully mitigate yield loss (Li et al. 2011). Both marker-assisted selection (MAS) and transgenic approaches have become important tools of modern plant breeding, allowing breeders to combine traits of interest without the need for additional, costly phenotyping, pedigree analysis, or (in the case of MAS) a detailed functional understanding of the molecular basis of a trait. The success of both approaches, however, depend on our ability to identify useful markers and alleles. Genetic mapping has been effective at identifying markers for use in MAS and candidates for transgenic methods. But traditional mapping approaches such as quantitative trait locus (QTL) and association mapping have a number of drawbacks that may limit their utility or generality (see below). Selection mapping can circumvent some of these problems by utilizing changes in allele frequency to identify markers that have been, or are tightly linked to, the target of historical selection. We propose to extend our previous work on selection mapping for yield to identify candidate agronomic loci (CAL) that will prove useful for molecular breeding approaches to adapt maize to changing climates. Our approach takes advantage of several millennia of adaptation of traditional maize varieties to different climate regimes, natural experiments of grand scale, using selection mapping methods to identify the targets of selection in the independent adaptation of maize to highland environments in both Mexico and South America. Importantly, differences between highland and lowland environments mirror the predicted effects of future climate change in the US (Fig. 1 C-E). The resulting list of CAL, along with the bioinformatic tools we will develop, will provide new opportunities for molecular breeding, accelerating the progress of adaptation to climate change and thus ensuring continued yield increases. Moreover, thanks to rapid advances in sequencing technology, the selection mapping approach proposed here should be easily extendible to virtually any crop species with sufficient germplasm resources; a modified method would be possible even for crops without a reference genome. We therefore expect that our approach will prove to be an important advance in breeding methods for several crop species.

1

Maize diffusion and adaptation

Despite its humble beginnings as a wild grass, maize spread rapidly across the globe to become the worlds top-producing crop. Wild members of the genus Zea, which diverged relatively recently (Ross-Ibarra et al. 2009), are endemic to a small region stretching from northern Mexico to Central America. The direct wild ancestor of maize, Zea mays ssp. parviglumis, occurs only along the low-elevation slopes of the Sierra Madre Occidental in the southwestern corner of Mexico (Snchez-Gonzlez and Ruiz-Corral 1997). After its domestication from ssp. parviglumis 9,000 before present (BP; Matsuoka et al. 2002, Piperno et al. 2009), maize spread from the lowlands of Southwest Mexico (van Heerwaarden et al. 2011), rapidly diffusing across the Americas. By 6,000 BP, maize had adapted to the high elevations of central Mexico and spread to the lowlands of South America (Piperno 2006); by 4,000 BP maize was being grown at high altitudes in the Andes (Perry et al. 2006). After European contact with the New World, maize continued its world-wide diffusion, spreading quickly across Europe and subsequently to Asia and Africa. Maize is now the worlds most broadly cultivated crop, currently grown on six continents in 162 countries and territories (FAOSTAT, 2009) in a distribution spanning 90 of latitude (from Chile to Canada) and more than 3000m of elevation (Tenaillon and Charcosset 2011). During expansion to such varied regions, maize encountered and adapted to extremes of temperature, day length, precipitation, and soil types (Troyer 2004). There is clearly tremendous potential to harness the allelic variation present among maize populations for modern breeding, taking advantage of genotypes molded by selection to adapt cultivated maize to new and changing environments. In particular, the fact that maize has independently adapted multiple times to similarly extreme environments (e.g., the highlands of the Mexican Central Plateau and the Andes) presents a unique opportunity to identify adaptive loci for use in maize breeding. While these environments represent cooler, rather than warmer climates, identifying the loci underlying such adaptation in multiple populations should nonetheless be a powerful approach to finding CAL (see Approach below).

While variation in several well-known maize phenotypes (branching, glume architecture, endosperm color) is largely controlled by single loci (tb1, tga1, and y1 respectively), most complex traits have been found to be highly polygenic. For instance, recent genome-wide association studies (GWAS) of leaf morphology, flowering time, and disease resistance have all found tens of QTL of small effect (Buckler et al. 2009, Tian et al. 2011, Kump et al. 2011). These authors suggested that the outcrossing mating system of maize may explain the highly quantitative genetic architecture of adaptive traits in contrast to selfing species like rice and Arabidopsis where fewer loci are involved. Adaptation to varying temperature during maize diffusion post-domestication is thus likely to have also involved a diffuse genetic architecture. For example, tassel blasting and leaf firing, traits related to heat tolerance, have previously been shown to be polygenic in nature (Frova and Sari-Gorla 1994, Bai 2003) and temperature is known to affect many aspects of plant growth and development (e.g., seed germination, photosynthesis, respiration; Wahid et al. 2007).

Preliminary results
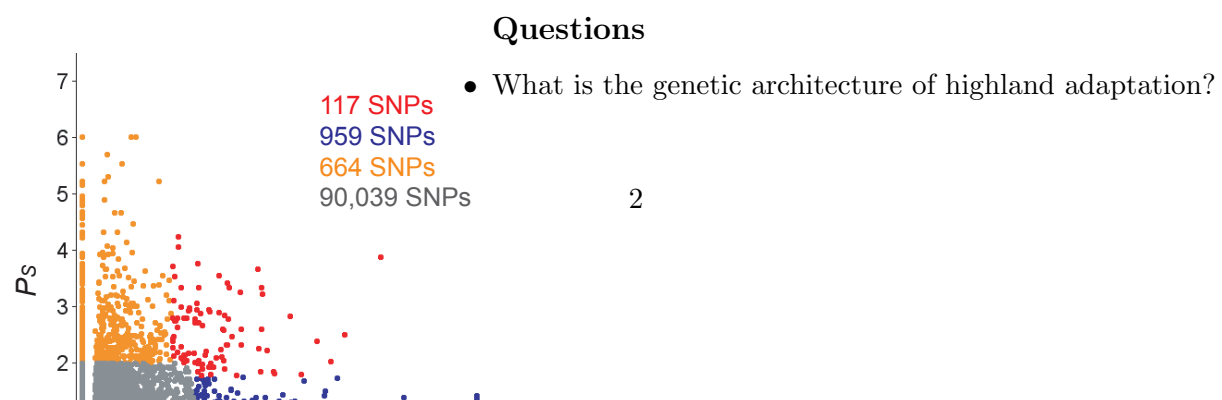
Aim 1   Genetic architecture of highland traits

Questions

- What is the genetic architecture of highland adaptation?

117 SNPs
959 SNPs
664 SNPs
90,039 SNPs

2

Table 1: Parental lines for QTL

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

- 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?

- Are highland QTL/loci widespread in highland climes?

Aim 1.1   QTL mapping of highland adaptation traits

In order to identify the 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 an inbred landrace adapted to lowland conditions with a landrace adapted to highland conditions 1. We make use of specially-inbred landrace lines created by John Doebley (U. Wisconsin) and Seth Murray (Texas A&M), thus simplifying downstream mapping applications and allowing replication of alleles in the our functional studies (see Aim 3).

Because we plan to evaluate these populations in replicated trials over multiple years, we will self-pollinate the F2 plants to create 500 F2:3 families from each population. DNA will be extracted from each ofhte F2 plants and used for genotypic analysis. The parents of the population will be sequenced to 20-30X depth on two lanes of Illumina 150bp paired-end, providing genome-scale SNP data similar in scale (tens millions of markers) to our previous work in maize (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 and allowing simple imputation of full-genome sequence. The genetic map will be created using standard methods (e.g. Broman

Table 2: Parental lines for QTL

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
Ames, Iowa	X	X	X	X

et al., 2003).

The 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 (2). In each location, the experiment will consist of two replications where the 500 entries and parental checks will be arranged in a randomized complete block design or an augmented alpha lattice design. The parental checks will be used to control for field variation. We will collect a large number of phenotypes (??) using our in-house barcode-based data collection program. Germination at different planting depths and temperatures will be evaluated in controlled conditions 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, as expected. Each location will then be analyzed separately to derive least squares means to be used as the phenotypic data in QTL analyses. QTL analysis will be conducted using standard, publicly available software (e.g. SAS; R/qtl Broman et al., 2003). Several iterations of QTL analysis will be con-

Table 3: Phenotypes measured

Abbreviation	Phenotype
MH	leaf sheath macrohairs
FT	flowering time
PH	plant height
BM	total plant biomass
EN	ear number
FK	fifty kernel weight
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

ducted: on individual traits, individual traits adjusted for covariates such as flowering time, and multiple traits simultaneously. The QTL profiles will be compared across populations (Mexican vs South American) and across field sites (elevations) to determine how elevation affects putatively adaptive traits. We expect very different QTL profiles from the highland and lowland evaluation sites, and from the Mexican and South American populations.

Aim 1.2   Admixture mapping in a teosinte hybrid zone

Aim 1.3   Global analysis of highland haplotypes

- Occurrence of highland haplotypes/QTL/SNPs in global pops (MBH)

- 500 worldwide accessions GBS (MBH)

- Case study in Chihuahua (ACJ)

Aim 2   Adaptive introgression of highland alleles

Questions

- Are introgressed loci adaptive?

- Does evidence of introgression and natural selection correspond to QTL?

Aim 2.1   Population genetics of maize-teosinte introgression

GBS of 18 inds x 10 pops x 2 subspecies (mex & maize) (JRI) 384 at 48 plex * $60 = $23040

Popgen on maize/mexicana introgression (JRI, GC) Identification of fine-scaled introgressed regions. Evidence of selection against introgression (recombination) Evidence of selection for introgressed regions (sweep signals, deviation of genome-wide admixture signal) Introgressed regions correlate with admix mapping signals? with QTL? Berg approach

Aim 2.2   Population genetics of hybridization in teosinte

make use of admixed pop sampled by matt Berg approach of phenotypes mapped there for selection in parentals and hybrid

GBS Additional 5 admixed parv/mex populations (50 inds. each) (JRI) also 4 new pops x (12 parv + 12 mex + 12 hybrids) = $8160 Ahuacatitlan and 3 more Introgression and adaptation in additional admix pops (JRI, GC) Selection for/against regions in parv/mex/admix Parallelism across pops in hybrid zone

Aim 3   Functional characterization of QTL

Aim 3.1   Questions

- What do QTL/selected loci/introgressed loci do?

Aim 3.2   Fine map pigmentation

- PT x T43 NIL population (RS)

- GBS genotyping (MBH)

Aim 3.3   Allelic series for QTL of interest

- 10 parents:

  - 4 parents F2:3
  - mexicana TIL18
  - Palomero Toluqueo
  - 2 lowland landraces
  - 2 highland landraces

- Cross into 3 parents for phenotyping

  - B73 (SFG)
  - T43 (MBH)
  - CML457 (RS)

Aim 3.4   RNAseq

- Time series analysis of plants in field (ACJ):

- 15 Lines

  - 4 F2:3 parents
  - 1 NIL chr4,
  - 1 each mex & parv TIL
  - B73, CML457, T43, PT
  - highland/lowland landraces used in allelic series (MBH to inbreed)

- 12 plants per line per environment (2 pools of 6)

- 4 stages/tissues per plant

- 2 environments (high/low fields)

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

- R code for novel analysis for admixed populations

- Software & pipelines on github

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 and pertinent - data. We will develop undergraduate teaching modules in population and quantitative genetics using publicly available human data. These modules will be tested and integrated into the large undergraduate teaching courses at UC Davis. We have already begun to develop and distribute some of these resources, e.g. genome-scale demonstrations of Hardy Weinberg Equilibrium (HWE) using the human HapMap data, see Figure 4. 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. They also offer the chance to more easily teach more complex population genetics concepts. For example, in a mixed population sample the heterozygosity is reduced below the HWE expectation by a factor (1-FST); this can easily be seen and explained for a combined sample of Europeans and Africans (right-side of Figure 4) Other examples will include: using height association data to demonstrate quantitative genetics models; and explaining concepts of genetic and genealogical ancestry using genomic identity by descent. We will also develop software to examine signals of selection around human adaptations, to allow students to examine the patterns of diversity around these loci in the common visual framework. 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 (see Data Management Plan), and advertised via evoldir and other venues. We will be develop and archive the modules on github.com 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. We will regularly deposit updated versions of the modules into figshare and data dryad in order to ensure that a permanent resource is maintained.

Germplasm resources

The highland environment remains an important niche for global maize production, in terms of acreage and farmer involvement, if not overall yield. The highland environment presents a number

of important abiotic challenges beyond high elevation per se, including periods of drought and cold, and highland adapted material has great potential to provide an important source of stress tolerance; the rapid development and inherent earliness of highland material, for example, may have general application in marginal environments.

In Mexico, 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. This project will move this important germplasm a step closer to breeding programs, and in doing so strengthen ties between the broader academic community (represented here by both US and Mexican institutions) and the Mexican private sector (letter of support UNISEM).

This project will also build further the collaboration between US institutions and LANGEBIO, with a number of Mexican graduate students being directly involved. While Mexican plant science retains a traditionally strength in biochemistry and molecular biology, genetics has been less well represented in recent years. As a consequence, while there is ready access to genomics technologies, the human resources to make the most of the opportunities they present may be lacking. This project provides an excellent opportunity for capacity building in a Mexican institution, with the expectation of a lasting impact through future collaboration.

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

Broader impacts   Hufford et al. 2012b Ross-Ibarra et al. 2009 Studer et al. 2011 Shi et al. 2010 Van Heerwaarden et al. 2010 van Heerwaarden et al. 2010 van Heerwaarden et al. 2011 Morrell et al. 2011 Cook et al. 2012 Fang et al. 2012 Chia et al. 2012 Gore et al. 2009 Hufford et al. 2012a Kanizay et al. 2013 Hufford et al. 2013 Melters et al. 2013 Mezmouk and Ross-Ibarra 2014 Pyhajarvi et al. 2013 Tian et al. 2011 Hung et al. 2012b Buckler et al. 2009 Brown et al. 2011 McMullen et al. 2009 Peiffer et al. 2014 Romay et al. 2013 Hung et al. 2012a Bottoms et al. 2010 Dubois et al. 2010 Zhang et al. 2010 Flint-Garcia et al. 2009a Flint-Garcia et al. 2009b Flint-Garcia et al. 2009c

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 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).

Publications Shi et al. (2010); Chia et al. (2012); Fang et al. (2012); Hufford et al. (2012a,b, 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 (Bradburd et al., 2013), 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.

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. Over the last three years, the current iteration of the project (DBI-0820619) has 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. Several positional cloning projects that will facilitate refinement of association mapping approaches are nearing completion. We have also been collaborating with colleagues around the world to resequence more than 100 Zea genomes and skim sequence 20,000 maize varieties. Overall, the project has published 58 scientific papers in the last three years, including 15 in Science, Nature Genetics, and PNAS.

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. (2009); 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. (2012b); Hung et al. (2012b,a); Romay et al. (2013)

Table 4: Proposed timeline of activities and responsibilities

Year	2015	2016	2017	2018	2019
Objective ?? Allelic series	2015	2016	2017	2018	2019
Objective ?? Fine mapping	–	RS, AC	AC,JRI	AC,JRI	–
Objective ?? RNA-seq	2015	2016	2017	2018	2019
Objective Aim 2.1 Maize/mexicana introgression	2015	2016	2017	2018	2019
Objective ?? Admix mapping	2015	2016	2017	2018	2019
Objective ?? QTL mapping	2015	2016	2017	2018	2019
Objective ?? Highland haplotypes	2015	2016	2017	2018	2019
Objective Aim 2.2 Admixture population genetics	2015	2016	2017	2018	2019

References Cited

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.

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 Genet*, 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, 2009.

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.

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.

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.

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, 2012a.

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, 2012b.

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

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, 2012a.

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, 2012b.

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.

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.

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

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.

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.

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

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.

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.

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.

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.

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.

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

Maximum of 3 pages

Personnel

Other Personnel

Fringe Benefits

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

- Faculty - % in FYs 2012, 2013, and 2014

- Postdocs - % in FYs 2012, 2013, and 2014

- Graduate students - % in FYs 2012, 2013, and 2014

- Undergraduate students - % in FYs 2012, 2013, and 2014

- Staff - % in FYs 2012, 2013, and 2014

- Part time staff - % in FYs 2012, 2013, and 2014

 Additionally, the university applies a risk management charge of 1% on all personnel salaries

Travel

Other Direct Costs

Materials and Supplies:
    Consultant Services:
    Graduate Student Tuition

Indirect Costs

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

Maximum of 2 pages

Prior experience – shared data on iPlant, used figshare, slideshare

Sequence data – backed up on drobo at RILab and iPlant, submitted to SRA, genotypes on data dryad or figshare

Germplasm – germplasm sent to stock center, smaller lots kept in database. mysql database accessible by all labs.

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 or biweekly) 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.