

## **Project Summary**

*Maximum of 1 page*

**Intellectual Merit**

**Broader Impacts**

# Project Description

## 1 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 CO<sub>2</sub> and changes in rainfall—are likely to be as important for maize yields. Though elevated CO<sub>2</sub> may actually prove beneficial for yields of some crops, models suggest the effect on maize, a C<sub>4</sub> 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 climates (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.

Introduction Maize diffusion and adaptation Despite its humble beginnings as a wild grass,

maize spread rapidly across the globe to become the world's 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-Gonzalez 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 world's 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). Genetics of adaptation 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).

## 2 Genetic architecture of highland adaptation in *Zea mays*

### 2.1 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?

- Are highland QTL/loci widespread in highland climes?

## **2.2 QTL mapping of highland adaptation traits in high x low mapping populations from Mexico and S. America**

- Phenotyping at 4 sites:
  - low Mexico (RS)
  - high Mexico (RS)
  - Missouri (SFG)
  - growth chamber (MBH)
- 960 F2:3 families; 40 checks
- population development (RS, SFG)
- genotyping and mapping (JRI, SFG)
- DNA extraction, genotypes (SFG)
- Sequence parents (JRI)
- Mapping (SFG)

## **2.3 Admixture mapping in mex/parv hybrid zone**

- collection trip (MBH)
- 500 individuals from Ahuacatitlan in Mexico (MBH, RS)
- DNA extractions, genotyping (MBH)
- phenotyping (MBH, RS)
- mapping (MBH/GC)

## **2.4 Phenotypes**

- macrohairs (SFG, RS, MBH)
- flowering time (SFG, RS, MBH)
- tassel morphology (SFG, RS, MBH)
- plant height every 2 weeks & at flowering (SFG, RS, MBH)
- biomass (SFG, RS, MBH)
- # ears (mz), 50k weight (mz/teo), total seed weight (mz) (SFG, RS, MBH)
- stem/plant color (SFG, RS, MBH)

- germination depth, temp in greenhouse (F2:3 only; MBH)
- roots (inquire with Bloom)
- ionomics (Ivan via letter of support)

## 2.5 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)

## 3 Population Genetics of Introgression

### 3.1 Questions

- Are introgressed loci adaptive?
- Does evidence of introgression and natural selection correspond to QTL?

### 3.2 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

### 3.3 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* (.

## 4 Functional characterization of QTL

### 4.1 Questions

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

### 4.2 Fine map pigmentation

- PT x T43 NIL population (RS)
- GBS genotyping (MBH)

### 4.3 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)

### 4.4 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 LANGE BIO 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

- Yearly workshop at UM
- Participants pay to purchase handheld
- 2 day workshop provides software, training on phenotyping in maize

## 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. 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  $(1-F_{ST})$ ; 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 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

# Results From Prior NSF Support

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

*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](http://www.maizegdb.org). Former trainees on the grant include Dr. Matthew Hufford (Co-PI on the current grant).

- Pyhäjärvi T, Hufford MB, Mezouk S, Ross-Ibarra J (2013) *Genome Biol & Evol*: 5:1594-1609
- Hufford MB, Lubinsky P, Pyhäjärvi T, et al. *PLoS Genetics* 9(5): e1003477.
- Melters DP, Bradnam KR, Young HA, et al. (2013) *Genome Biology* 14:R10
- Kanizay LB, Pyhäjärvi T, Lowry E, et al. *Heredity* 110: 570-577.
- Hufford MB, Bilinski P, Pyhäjärvi T, Ross-Ibarra J (2012) *Trends in Genetics* 12:606-615
- Hufford MB, Xun X, van Heerwaarden J, et al. (2012) *Nature Genetics* 44:808-811
- Chia J-M, Song C, Bradbury P, et al. (2012) *Nature Genetics* 44:803-807
- Fang Z, Pyhäjärvi T, Weber AL, et al. (2012) *Genetics* 191:883-894



- Shi J, Wolf S, Burke J, Presting G, et al. (2010) PLoS Biology 8: e1000327

**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: <http://genescape.ucdavis.edu/scripts-and-code/bedassle/>, and has already been used by a many molecular ecologists.

- Bradburd, G., Ralph, P., Coop, G. (2013) Evolution 67: 3258-3273

**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.

- Bottoms, C., Flint-Garcia, S. & McMullen, M.D. BMC Bioinformatics 11 Suppl 6, S28 (2010).
- Brown, P.J. et al. PLoS Genetics 7, e1002383 (2011).

- Buckler, E.S. et al. *Science* 325, 714-8 (2009).
- Chia, J.-M. et al. *Nature Genetics* 44: 803-807 (2012).
- Cook, J.P. et al. *Plant Physiology* 158, 824-834 (2012).
- Dubois, P.G. et al. *Plant Physiology* 154, 173-86 (2010).
- Fang Z, Pyhäjärvi T, Weber AL, Dawe RK, Glaubitz JC, Sanchez Gonzalez JJ, Ross-Ibarra C, Doebley J, Morrell PL, Ross-Ibarra J. *Genetics* 191: 883-894. (2012)
- Flint-Garcia, S.A., Bodnar, A.L. & Scott, M.P. *Theoretical and Applied Genetics*. 119, 1129-42 (2009).
- Flint-Garcia, S.A., Buckler, E.S., Tiffin, P., Ersoz, E. & Springer, N.M. *PLoS ONE* 4, e7433 (2009).
- Flint-Garcia, S.A., Dashiell, K.E., Prischmann, D.A., Bohn, M.O. & Hibbard, B.E. *Journal of Economic Entomology* 102, 1317-1324 (2009).
- Flint-Garcia, S., Guill, K. & Sanchez-Villeda, H. *Maydica* 54, 375-386 (2009).
- Hufford, M. et al. *Nature Genetics* 44: 808-811 (2012).
- Hung, H.-Y. et al. *Heredity* 1-10 (2011).
- Hung H-Y, Shannon LM, Tian F, Bradbury PJ, Chen C, Flint-Garcia SA, McMullen MD, Ware D, Buckler ES, Doebley JF,
- Holland JB. *PNAS* 109: E1913-E1921. (2012)
- McMullen, M.D. et al. *Science* 325, 737-40 (2009).
- Morrell, P.L., Buckler, E.S. & Ross-Ibarra, J. *Nature Reviews Genetics* 13, 85-96 (2011).
- Ross-Ibarra, J., Tenaillon, M. & Gaut, B.S. *Genetics* 181, 1399-413 (2009).
- Studer, A., Zhao, Q., Ross-Ibarra, J. & Doebley, J. *Nature Genetics* 43, 1160-3 (2011).
- Tian, F. et al. *Nature Genetics* 43, 159-62 (2011).
- van Heerwaarden, J. et al. *Molecular Ecology* 19, 1162-73 (2010).
- van Heerwaarden, J., van Eeuwijk, F. a & Ross-Ibarra, J. *Heredity* 104, 28-39 (2010).
- van Heerwaarden, J. et al. *PNAS* 108, 1088-1092 (2011).
- Zhang, N, A Gur, Y Gibon, R Sulpice, S Flint-Garcia, MD McMullen, M Stitt, ES Buckler. *PLoS One* 5: e9991 (2010)

## References Cited

# 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-Ibarra occupies 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 bioanalyzer 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 total of 8 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.

### **LANGE BIO**

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 (Queretaro; 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:** *Maximum of 1 page*

**Documentation of collaborative arrangements** of significance to the proposal through letters of commitment.

**Documentation regarding research involving the use of human subjects,** hazardous materials, vertebrate animals, or endangered species.