US-Mexico planning visit and workshop to assess the genomic basis of local adaptation in maize

Objectives of the Proposed Planning Trip and Workshop:

The crop plant maize (*Zea mays* ssp. *mays*) was domesticated from teosinte (*Zea mays* ssp. *parviglumis*), a wild grass species with a narrow ecological distribution in southwest Mexico (Matsuoka et al. 2002; Fukunaga et al. 2005). Despite the limited environments inhabited by its teosinte progenitor, maize has spread globally and is currently cultivated across a wide range of latitudes (from the tip of South America to Canada), altitudes, precipitation gradients, and soil types (Tenaillon & Charcosset 2011). Reciprocal transplants of local farmer varieties of maize (*i.e.*, landraces) from altitudinal extremes have experimentally demonstrated local adaptation (Mercer et al. 2008). Previous work has suggested these adaptations have been obtained through gene flow from locally adapted wild relatives (Hufford et al. 2013) and/or through mutations newly arising during the spread of maize (Takuno et al. In Prep). The genomic architecture of local adaptation in maize can be precisely determined through a combination of high-density genotyping and reciprocal transplant field experiments. Such studies will provide information regarding the basic evolutionary history of maize and will generate valuable data for continued improvement of maize in the face of growing human populations and climate change.

Here we propose a 12-month project that will catalyze a new international collaboration between three universities in the United States and researchers from the National Laboratory of Genomics for Biodiversity (Langebio) in Irapuato, Mexico. The project will bring together the expertise necessary for studying local adaptation in maize and will provide essential preliminary data for a planned NSF Plant Genome Research Program (PGRP) proposal. Over the course of the proposed project, participants will attend an initial planning meeting (September 2014), develop resources for studying local adaptation in maize (October 2014-August 2015), and hold a workshop on high-throughput phenotyping technologies/methods (January 2015).

For the initial planning meeting, the project PI (Matthew Hufford), co-PI's (Graham Coop and Jeffrey Ross-Ibarra), and Senior Personnel (Sherry Flint-Garcia) will travel to Mexico to meet with the Principal Foreign Collaborators (Angelica Cibrián Jaramillo and Ruairidh Sawers). The meeting will begin with project organization at the Langebio campus in Irapuato, Mexico. Participants will tour facilities at Langebio and meet with administrators to discuss a memorandum of understanding between institutions and the data management plan for the project. Project responsibilities will also be clarified and a plan of action will be developed for proposed CNIC activities. Following organizational discussions at Langebio, participants will travel to potential field sites that might be useful in developing and evaluating local adaptation mapping populations. Dr. Sawers has been investigating potential field sites in both the highlands (near the town of Amealco, Querétaro State) and lowlands (near the town of Bucerías, Nayarit State) of Mexico. During field site visits, participants will discuss the potential of sites for meeting long-term project goals and the logistics involved in using sites.

After the initial planning meeting, the project will shift to development of resources that will eventually be used for in-depth analysis of local adaptation. Drs.

Flint-Garcia, Hufford and Sawers are developing mapping populations for analysis of highland adaptation in maize. Two types of populations are currently under development: F2:3 families of highland × lowland maize landrace crosses and near isogenic lines of a highland landrace in the background of two modern inbred lines. These resources have not yet reached maturity and will be advanced two generations (winter 2015 in Mexico and summer 2015 in Mexico, Iowa, and Missouri) during the timeframe of our proposed CNIC project. Fully developed mapping populations will be an essential resource for a planned PGRP proposal. With these resources, we can effectively link phenotype to genotype and elucidate the genomic architecture of highland adaptation in maize. Highland and lowland subspecies of teosinte (*Zea mays* ssp. *mexicana* and *Zea mays* ssp. *parviglumis* respectively) in Mexico also present an excellent opportunity to study the genomic architecture of local adaptation. These subspecies are distributed across a continuous altitudinal gradient with substantial admixture occurring between subspecies at intermediate altitudes. Project co-PI Coop is currently developing methods to detect population genetic signatures of selection using admixture mapping. During the CNIC project, Dr. Hufford will collect samples of both subspecies of teosinte across an altitudinal transect in Mexico for an empirical testing of Dr. Coop's methods that is planned as part of our future PGRP proposal.

Finally, in preparation for the substantial phenotyping that will take place in our planned PGRP project on local adaptation in maize, Dr. Flint-Garcia will lead a phenotyping workshop in Bucerías, Mexico in January of 2015. Dr. Flint-Garcia has extensive experience generating maize phenotypic data and members of her USDA-ARS group have developed a streamlined system utilizing a handheld barcode device, barcoded plant tags, and barcoded phenotyping tools in order to maximize efficiency. We will purchase four handheld devices for use in phenotyping training and data collection for the planned PGRP project. The phenotyping workshop will take place at the project's field site in Bucerías and will include Senior Personnel, Principal Foreign Collaborators, and, additionally, students and young researchers from the University of California, Davis, Iowa State University, the University of Missouri and Langebio. Through this workshop, Dr. Flint-Garcia's state-of-the-art system will be transferred to three institutions.

Background on Project Participants:

The researchers and institutions brought together through this proposed CNIC project have the theoretical (Coop et al. 2010a, b; Gunther & Coop 2013), computational (Hufford et al. 2013; Pyhäjärvi et al. 2013), biological (Cibrian-Jaramillo et al. 2010; Sawers et al. 2010; Hufford et al. 2012), and field-based (Flint-Garcia et al. 2005; Buckler et al. 2009; Cook et al. 2012) expertise necessary to address questions of local adaptation in maize.

From the United States:
1) Matthew Hufford (PI), Assistant Professor, Iowa State University, Department of Ecology, Evolution and Organismal Biology.
Relevant Expertise: Field-based study of evolutionary genetics in maize and teosinte.
Role in Project: Organization of planning meeting and workshop, collection of teosinte in Mexico, maize population development at Iowa State University.

2) Graham Coop (co-PI), Associate Professor, University of California, Davis, Department of Evolution and Ecology.
Relevant Expertise: Theoretical evolutionary genetics.
Role in Project: Guiding data and sample collection for planned empirical testing of evolutionary theory.

3) Jeffrey Ross-Ibarra (co-PI), Associate Professor, University of California, Davis, Department of Plant Sciences.
Relevant Expertise: Computational approaches for evolutionary genetics in maize and teosinte.
Role in Project: Guiding data and sample collection and population development for planned empirical studies.

4) Sherry Flint-Garcia (Senior Personnel), USDA-ARS Research Geneticist and Adjunct Assistant Professor, University of Missouri.
Relevant Expertise: Population development for quantitative genetics in maize and teosinte.
Role in Project: Maize population development at the University of Missouri, leading workshop on collection of phenotypic data in Bucerías, Mexico.

From Mexico:
1) Angelica Cibrián Jaramillo, Assistant Professor, National Laboratory of Genomics for Biodiversity (Langebio).
Relevant Expertise: Local adaptation of agricultural plants in Mexico.
Role in Project: Organization of planning meeting and workshop, coordinating training opportunities for students and young researchers.

2) Ruairidh Sawers, Assistant Professor, National Laboratory of Genomics for Biodiversity (Langebio).
Relevant Expertise: Population development for molecular genetics in maize.
Role in Project: Maize population development in Mexico.

Itinerary for Initial Project Planning Meeting:
        The initial project-planning meeting will take place over the course of five days with the following proposed itinerary:
Day 1: Organizational Meeting at Langebio Campus, Irapuato, Mexico.
        Participants will spend the morning touring the facilities at Langebio and meeting with administrators to discuss a memorandum of understanding and the data management plan. The afternoon and evening will be spent clarifying responsibilities and the plan of action for proposed CNIC activities.

Day 2: Visit to highland field site in Amealco, Mexico; continued project organization.
        During the morning of Day 2, participants will travel to a potential highland field site outside the nearby town of Amealco and discuss its suitability as a long-term highland testing location. The afternoon of Day 2 will consist of additional meetings to

discuss project organization.

Day 3: Travel to lowland field site in Bucerías, Mexico.
　　　　The field site in the lowlands of Mexico is a full day's drive via car or bus from the Langebio campus. Day 3 will therefore be spent traveling and holding informal discussions regarding project organization.

Day 4: Visit to lowland field site in Bucerías, Mexico.
　　　　During the morning of Day 4, participants will tour the lowland Bucerías field site and discuss its potential as a long-term lowland testing location. The afternoon of Day 4 will consist of meetings to discuss specifics of mapping population development and the teosinte collection trip planned as part of the CNIC project.

Day 5: Planning meeting to discuss population development and teosinte collection trip.
　　　　During the final day of the planning trip, discussions will continue regarding the logistics of the maize mapping population development and the teosinte collection. Participants will also discuss organization of the proposed phenotyping workshop to be held in Bucerías.

Itinerary for Phenotyping Workshop:
　　　　The phenotyping workshop will take place at the Bucerías field site over the course of four days with the following proposed itinerary:
Day 1: Experimental design.
　　　　The first day will focus on topics such as 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.

Day 2: Setting up the FieldBook database on the desktop/laptop prior to data collection.
　　　　The second day will focus on the logic needed to ensure high-throughput phenotyping. Topics to be covered include setting up palm 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.

Day 3: Data collection in the field.
　　　　Day three will focus on using Palm-based barcode scanners and barcoded tools to collect phenotypic data. Topics to be covered 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,

Day 4: Overview of phenotypic data analysis.
　　　　The final day will focus on basic statistics including outlier analysis, regression and correlation, and analysis of variance.

History of Proposed Collaboration:

Connections between investigators at the University of California, Davis and Langebio have been developing recently through an NSF-funded student exchange program included in a current PGRP grant on functional genomics of maize centromeres. Informal talks at the annual Maize Genetics Conference and a preliminary meeting of Drs. Cibrián-Jaramillo, Coop, Flint-Garcia, Hufford, Ross-Ibarra, and Sawers at the University of California, Davis have clarified the potential in collaboration between US institutions and Langebio for the study of local adaptation in maize. While each of the researchers included in this CNIC proposal has worked independently on questions of local adaptation and/or maize evolution, no formal collaboration has been established between the US investigators (Drs. Coop, Flint-Garcia, Hufford, and Ross-Ibarra) and Drs. Cibrián-Jaramillo and Sawers from Langebio.

Opening New Scientific Directions:

Previously, US researchers in this project have developed methods for detecting the population genetic signature of local adaptation (Coop et al. 2010a; Ralph & Coop 2010; Gunther & Coop 2013) and have applied these and other methods in empirical studies of maize and its wild relatives (Hufford et al. 2013; Pyhäjärvi et al. 2013). The proposed CNIC project will open a number of new scientific directions for US researchers. The development of mapping populations will allow researchers to assess which loci are associated with traits thought to underlie local adaptation. For example, stem macrohairs and pigment are thought to be important highland adaptations in maize (Lauter et al. 2004; Hufford et al. 2013). Once the loci controlling these traits are identified through the use of mapping populations developed here, researchers can assess their overlap with regions of the genome showing marked population genetic evidence of selection in the highlands. In addition, teosinte represents an ideal study system for investigating population genetic signatures of local adaptation in natural populations (Hufford et al. 2012). As mentioned above, Dr. Coop is developing methods for detecting selection through admixture mapping. The teosinte collections obtained during this project are not currently available and will be ideally suited for empirical testing of Dr. Coop's methods.

Justification for Collaboration with Langebio:

Langebio, and specifically Drs. Cibrián-Jaramillo and Sawers, are ideal collaborators for both the CNIC project proposed here and for the project that will be proposed to PGRP. Langebio is arguably the premiere institution in Mexico for applying genomic tools to evolutionary and ecological questions in plants. Researchers at this institution have also contributed meaningfully to the study of maize. For example, scientists at Langebio led the assembly of the Palomero Toluqueño genome (a highland maize landrace; published in *Science* (Vielle-Calzada et al. 2009)). Genomic resources at Langebio may eventually play a supporting role in our local adaptation project: the Core Facility at Langebio employs a staff of 18 and offers 454, SOLiD, Illumina and Sanger sequencing platforms, RT-PCR instrumentation, high-throughput robotics, and a

proteomics facility. Langebio is also ideally situated for ongoing maize population development and phenotyping in highland and lowland environments. Finally, Drs. Cibrián-Jaramillo and Sawers (and several others at Langebio) are international experts on Mexican agroecosystems and will provide crucial biological insight throughout the course of the project.

Training Students and Early Career Researchers:

The main training opportunity for students and early career researchers in the proposed CNIC project will be the phenotyping workshop to take place in Bucerías, Mexico in January of 2015. Graduate students or postdoctoral scholars from the University of California, Davis, the University of Missouri, Iowa State University and Langebio will be invited to attend. Dr. Flint-Garcia will lead this four-day workshop and will describe methods of both phenotypic data collection and analysis. Ancillary training opportunities on mapping population development will also be available throughout the project.

Future Plans for Collaboration:

Our initial plan for continuing our collaboration will be to use the preliminary data and resources produced by the CNIC project to support a proposal to be submitted to the NSF Plant Genome Research Program. Should we receive funding from PGRP, we will characterize the genomic architecture of highland adaptation using two distinct methodologies. First, we will utilize the mapping populations developed during our CNIC project to characterize the genomic architecture of specific traits thought to underlie highland adaptation (*e.g.*, macrohairs, pigment, flowering time). Once we discover regions of the genome underlying these traits we will proceed with functional characterization of promising candidate genes. Second, we will use our teosinte collections across an altitudinal gradient in Mexico to identify candidate genomic regions under selection in the highlands using admixture-mapping methods developed by Dr. Coop. The advantage of this second approach relative to our first is that relevant traits are not identified *a priori*, thereby limiting our view of potential loci important for highland adaptation. An evaluation of the genetic architecture of highland adaptation using these approaches will greatly inform our understanding of the evolutionary history of maize and will generate valuable data for continued crop improvement.

Broader Impacts:

We will strive to produce societally relevant outcomes in our proposed CNIC project. We have identified three outlets in which we can broaden the impact of our research. First, we will create an educational opportunity through our phenotyping workshop in Bucerías, Mexico. We will encourage minority students and/or postdocs from the US-based groups to attend this workshop, offering them both valuable STEM training and an international scientific experience. Second, we will work with collaborators at the Maize Genetics Cooperation Stock Center to make our mapping populations publicly available, thereby creating new and valuable infrastructure for the

maize research community. Finally, Drs. Coop and Ross-Ibarra have a proven track record of making their population genetic software available to the public and these resources are widely used. The teosinte collections generated as part of this project will be instrumental in developing theory and associated software related to admixture mapping. All software will be made publicly available on the websites of Drs. Coop and Ross-Ibarra. Moreover, all data generated during both the proposed CNIC and PGRP projects will be released to the public with a permanent doi via figshare, datadryad or similar services.

Results from Prior NSF Support:

Co-Principal Investigator Graham Coop:

NSF grant #1262645: Collaborative Research: ABI Innovation: Visualization And Statistics For Spatial Population Genomic Analysis. $314,260. 5/1/2013-5/1/2016.
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 is already being used by many molecular ecologists.
Bradburd, G., Ralph, P., Coop, G. (2013) Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution* 67: 3258–3273

Co-Principal Investigator Jeffrey Ross-Ibarra:

NSF grant #0922703: Plant Genome Research Program: Functional Genomics of Maize Centromeres. $5,014,464. 6/15/2010-5/31/2015
Intellectual Merit: Centromeres are regions of the genome that organize and regulate chromosome movement. They are key elements in the development of next-generation artificial chromosomes for introducing new traits into crop plants. The experiments carried out in this project are completing the sequence and assembly of maize centromeres, addressing known epigenetic limitations associated with employing artificial chromosomes, developing improved strategies for streamlining centromere study and use, and documenting the diversity of centromeres in maize and allied species. The Ross-Ibarra lab has investigated the population genetics and evolution of centromeres in maize and its wild relatives, revealing evidence of selection on maize centromeres, changes in centromere repeat abundance during domestication, and gene conversion within maize centromeres.
Broader Impacts: The grant has established two training programs in plant genetics: a plant cytogenetics workshop at the University of Wisconsin and an international student exchange program at UC Davis. Data and result of this project have been disseminated via publications as well as deposited in the maize genetics community database (www.maizegdb.org).

Publications from co-PI Ross-Ibarra resulting from this award:

1. Pyhäjärvi T, Hufford MB, Mezmouk S, Ross-Ibarra J (2013) Genome Biol & Evol *In Press*
2. Hufford MB, Lubinsky P, Pyhäjärvi T, *et al*. PLoS Genetics 9(5): e1003477.
3. Melters DP, Bradnam KR, Young HA, *et al.* (2013) Genome Biology 14:R10
4. Kanizay LB, Pyhäjärvi T, Lowry E, *et al.* Heredity 110: 570-577.
5. Hufford MB, Bilinski P, Pyhäjärvi T, Ross-Ibarra J (2012) Trends in Genetics 12:606-615
6. Hufford MB, Xun X, van Heerwaarden J, *et al.* (2012) Nature Genetics 44:808-811
7. Chia J-M, Song C, Bradbury P, *et al.* (2012) Nature Genetics 44:803-807
8. Fang Z, Pyhajarvi T, Weber AL, *et al.* (2012) GENETICS 191:883-894
9. Shi J, Wolf S, Burke J, Presting G, *et al.* (2010) PLoS Biology 8: e1000327

Senior Personnel Sherry Flint-Garcia:

NSF Grant #0820619: Plant Genome Research Program: Genetic Architecture of Maize and Teosinte. $9,823,000. 3/1/2009-2/28/2013.

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.

The project 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 travelling 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 primarily at Cornell but also in Kenya), and training of postdocs, graduate students and undergraduates, the vast majority of whom have continued in scientific careers.

A complete list of the project's publications can be found on the project website: http://panzea.org/lit/publication.html