Crop species coupled with their wild relatives represent ideal study systems for evolutionary and ecological genomics. The substantial genomic resources developed specifically for crops often transfer to their wild relatives, allowing researchers to apply cutting edge technology to empirical investigations in natural plant populations. This has certainly been the case for maize (*Zea mays* ssp. *mays*) and related wild taxa (*Zea* spp.). My research has applied the resources of maize for comparative genomic investigations of multiple epochs of selection and for analyses of evolution through adaptive introgression.

1) Selection during maize domestication, improvement and local adaptation:

Within the *Zea* genus, many compelling evolutionary comparisons can be made. I am lead author of a manuscript (1) resulting from an international collaboration in which data from 75 re-sequenced genomes and full transcriptomes of wild maize (*Z. mays* ssp. *parviglumis* (hereafter, *parviglumis*)), maize landraces (*i.e.*, local farmer varieties), and improved maize were analyzed to assess patterns of selection during both maize domestication and improvement. This study represents the most comprehensive look at maize evolution and plant domestication to date and has produced several novel findings: i) a much wider array of genes appear to have been selected than those identified based on previous QTL studies; ii) selection was stronger during initial maize domestication than during subsequent improvement; and iii) selection during modern improvement has resulted in complementation in gene expression between heterotic groups and has targeted highly expressed genes. As an extension of this work, I have collaborated on a separate manuscript to more fully document differentially expressed genes and rewiring of gene expression networks between maize and *parviglumis* during the process of domestication (2).

Maize landraces represent a fascinating case study of local adaptation: cultivated across wide latitudes (from the tip of South America to Canada) and thousands of meters of elevation, landraces have adapted to numerous climes during diffusion from the maize center of origin in Mesoamerica. I co-wrote (and am co-PI) on a recently funded USDA-NIFA Plant Genome grant ($300,000) to identify loci involved in independent adaptation of landraces to highland environments in Mexico and South America. Since temperature is a main variable distinguishing these highland and lowland regions, our hope is loci identified in scans for selection will be useful targets for plant breeding in the face of current climate change (Figure 1A & 1B). My first step in this investigation has been to genotype 94 Mexican and South American landraces from both the lowlands and the highlands for ~55,000 SNPs. My preliminary analysis of these data has identified two non-synonymous SNPs with very high levels of differentiation between highland and lowland environments. Strikingly, the same genotypes predominate in both highland regions (Figure 1C). Moving forward, our grant includes funding to re-sequence genomes of a subset of 40 of these highland and lowland landraces to high depth in order to carry out population genetic analyses and more fully address the extent to which independent adaptation to highland environments has involved the same loci and alleles. Candidate loci related to temperature adaptation will

be confirmed through genome-wide association studies based on phenotypes obtained in large-scale growth chamber experiments. Additionally, we are currently developing a mapping population of highland-lowland crosses from inbred landrace lines in order to define QTL for highland adaptation based on phenotyping in highland and lowland environments. I am collaborating with researchers at multiple universities to fund and carry out this latter portion of the project.

Figure 1. Identifying loci for temperature adaptation in maize. (A) Current and predicted future mean August temperatures in the United States. (B) Comparison of current temperatures in the Corn Belt (Iowa) and the highlands of South America/Mexico to predicted future Iowa temperatures and current South American/Mexico lowland temperatures. (C) Mean growing season temperatures for landraces with various genotypes of two SNPs found to differentiate highland and lowland maize.
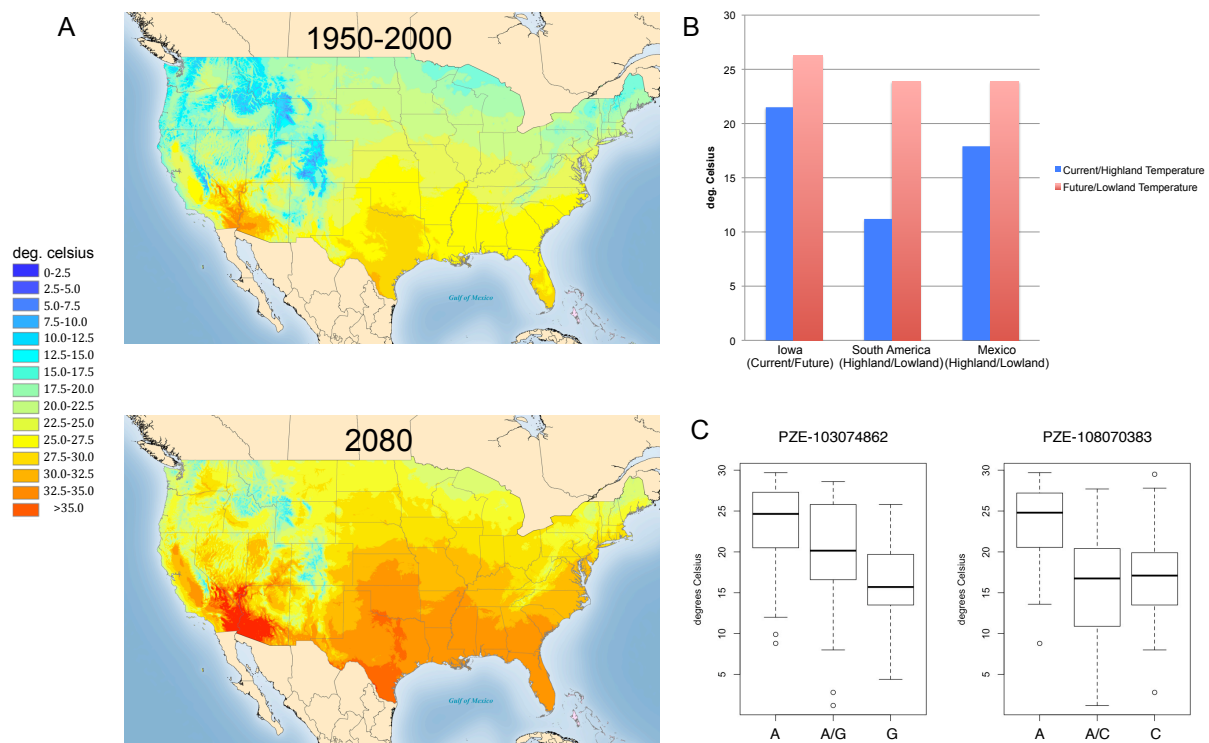

2) Evolution through introgressive hybridization:

I began my studies of gene flow in the genus *Zea* by estimating mating system parameters and patterns of pollen flow in *parviglumis* from southwest Mexico. Though practical constraints prevented direct estimates of gene flow, I was able to apply variance-component, or two-generation methods, to estimate paternal contributions in progeny arrays from across a population and subsequently model relatedness of pollen pools across the landscape. I applied these methods to a *parviglumis* population of several million individuals in the first such study in a high-density, annual species (3). By also assessing cryptic sub-population genetic structure through spatially-explicit

Bayesian assignment, I was able to demonstrate that failure to account for non-random mating while using two-generation methods can lead to underestimates of the extent of pollen flow, a significant contribution to this field. While carrying out this work, I observed numerous recent hybrids between *parviglumis* and sympatric maize and became curious about the role of hybridization in the evolution of maize landraces and wild relatives.

My current work is utilizing genomic resources to characterize introgression between maize landraces and the highland teosinte taxon *Z. mays* ssp. *mexicana* (hereafter, *mexicana*). Despite widespread evidence for introgression between these taxa, they maintain their distinct morphologies and have done so in sympatry for thousands of years. Two fascinating unanswered questions are the extent to which the genomes of maize and *mexicana* remain differentiated in hybrid swarms and how reciprocal gene flow has adapted these taxa to their current agronomic setting. The potential for adaptive introgression is particularly compelling given the domestication history of maize: maize was domesticated from *parviglumis* in the lowlands of southwest Mexico but then expanded into the highlands of the Mexican Central Plateau where it came into sympatry with the highland-adapted *mexicana*. Conserved regions of *mexicana* introgression into maize could indicate maize received alleles for local adaptation from *mexicana* during its expansion into the highlands. To address these questions from a genome-wide perspective, I have genotyped individuals from nine pairs of sympatric populations of maize and *mexicana* for ~55,000 SNPs. By assessing the extent of shared haplotypes between putatively admixed individuals and reference allopatric populations of maize and *mexicana*, I have discovered multiple genomic regions that show parallel evidence of introgression (*mexicana* into maize) across several populations (Figure 2). To evaluate the potential phenotypic effects of this introgression, I conducted a growth chamber experiment under highland conditions with maize populations that contained and lacked *mexicana* introgression at multiple key loci. Those containing introgression showed increased plant height (a proxy for fitness) and displayed phenotypes known to be important for highland adaptation. Another compelling finding emerging from this study is resistance to introgression in either direction of gene flow around known incompatibility loci (*e.g.*, *ga1* and *tcb1*; Figure 2) and loci I have previously shown to be important during domestication (1). This study represents the most extensive look at introgression across the genome in plant populations to date. My findings should spur further work into adaptation through introgression and suggest that multiple teosinte taxa may have contributed to modern maize. In the future, I hope to continue my work with introgression by focusing analysis on specific genomic regions. I am currently pursuing collaborations to develop near-isogenic introgression lines of *mexicana* in a lowland maize background. These introgression lines can be evaluated under highland and lowland conditions to definitively demonstrate the adaptive nature of introgression and to assess the individual effects of introgressed alleles.

Figure 2. Introgression in sympatric populations of *mexicana* (left) and maize (right) based on a hidden Markov model implemented in the program HapMix. The y axis for each population ranges from 0-2 copies of alleles from *mexicana* (maroon) or maize (gold) and the x axis indicates physical position along chromosome 4. Coordinates of the maize-teosinte incompatibility loci *ga1* and *tcb1* are indicated by dashed lines. Localities of sympatric populations are listed between plots.
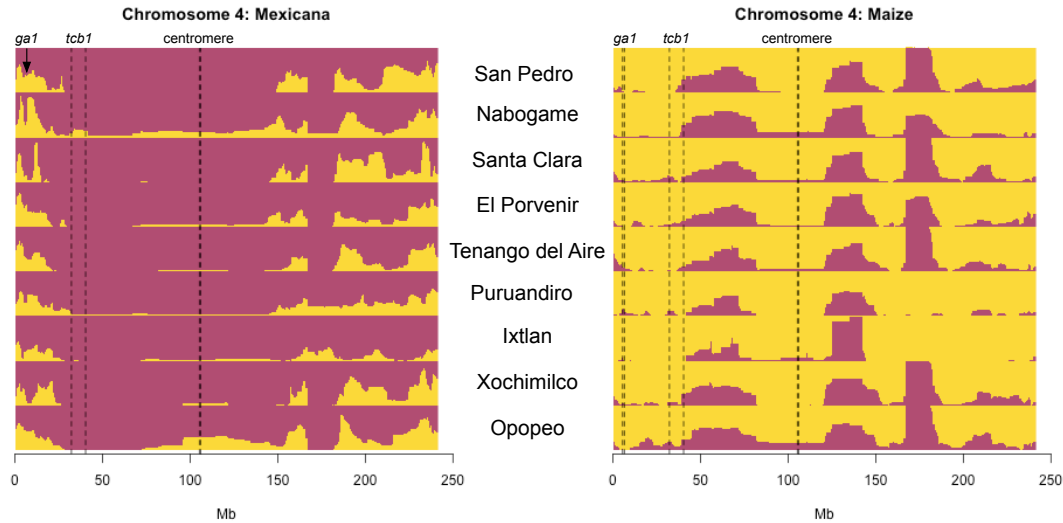
Further investigations and collaborations:

In addition to the work I have led on selection and introgression in *Zea*, I have also collaborated in several recent and ongoing plant evolutionary studies. For instance, by mapping paired-end, short-read data from maize and *Zea luxurians* to a maize transposable element (TE) database, collaborators and myself were able to assess how genome-wide abundance of TEs has changed since divergence of these two species (5). More recently I have collaborated on manuscripts introducing the second version of the maize haplotype map (6), describing selection during different periods of improvement of North American maize (7), assessing the diversity and distribution of the Abnormal chromosome 10 meiotic drive complex in *Zea* (8) and analyzing gene flow at range limits in *Mimulus laciniatus* (9). I have also participated in investigations related to the species niche (10) and local adaptation (11) of the wild subspecies of *Zea mays* and am first author on a recently published review on teosinte as a model system for population and ecological genomics (12). Continuing on in my postdoc, I will be using a diverse panel I have developed of individuals from across the *Zea* (temperate and tropical improved maize lines, landraces, *parviglumis*, *mexicana*, *Zea luxurians*) in order to assess evidence for evolution in centromeres. I will use this panel to evaluate evidence for rapid evolution of kinetochore binding proteins and the centromere drive hypothesis. I have previously found evidence for rapid evolution in *Zea* centromeres (1) and these new investigations will help resolve the paradox of how this occurs in genomic regions with such a conserved and important function.

In conclusion, I believe it is an exciting time for plant evolutionary genomics and that maize and teosinte present an excellent study system for making major advances in our understanding of domestication, local adaptation, and introgression. Surprisingly,

despite teosinte's close evolutionary proximity to economically important maize, it is quite understudied. Iowa State University would be a perfect setting to continue my research due to the strength of the Department of Ecology, Evolution, and Organismal Biology, the availability of impressive genomic resources through the Center for Plant Genomics, and its eminent reputation in crop genetics and genomics. Moreover, I believe my skill set and expertise in plant domestication and genomic studies of crop wild relatives nicely complement and expand upon the strengths of the Department. I can already see the potential for much fruitful collaboration!

References Cited:

1. Hufford, M. B., X. Xu, J. van Heerwaarden, T. Pyhäjärvi, J.-M. Chia, R. A. Cartwright, R. J. Elshire, J. C. Glaubitz, K. E. Guill, S. Kaeppler, J. Lai, P. L. Morrell, L. M. Shannon, C. Song, N. M. Springer, R. A. Swanson-Wagner, P. Tiffin, J. Wang, G. Zhang, J. Doebley, M. D. McMullen, E. S. Buckler, D. Ware, S. Yang, J. Ross-Ibarra. 2012. Comparative population genomics of maize domestication and improvement. NATURE GENETICS **44**:808-811.
2. Swanson-Wagner, R., R. Briskine, R. Schaefer, M. B. Hufford, J. Ross-Ibarra, C. L. Myers, P. Tiffin, N. M. Springer. 2012. Reshaping of the maize transcriptome by domestication. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **109**:11878-11883.
3. Hufford, M. B., P. Gepts, and J. Ross-Ibarra. 2011. Influence of cryptic population structure on observed mating patterns in the wild progenitor of maize (*Zea mays* ssp. *parviglumis*). MOLECULAR ECOLOGY **20**:46-55.
4. Hufford, M. B., P. Lubinsky, T. Pyhäjärvi, M. T. Devengenzo, N. C. Ellstrand, J. Ross-Ibarra. The genomic signature of crop-wild introgression in maize. (Submitted) Preprint archived at: http://arxiv.org/abs/1208.3894
5. Tenaillon, M. I., M. B. Hufford, B. S. Gaut, and J. Ross-Ibarra. 2011. Genome size and transposable element content as determined by high-throughput sequencing in maize and *Zea luxurians*. GENOME BIOLOGY AND EVOLUTION **3**:219-229.
6. Chia J.-M., C. Song, P. Bradbury, D. Costich, N. de Leon, J. Doebley, R. Elshire, B. Gaut, L. Geller, J. Glaubitz, M. Gore, K. Guill, J. Holland, M. B. Hufford, J. Lai, M. Li, X. Liu, Y. Lu, W. McCombie, R. Nelson, J. Poland, B. M. Prasanna, T. Pyhäjärvi, T. Rong, R. Sekhon, Q. Sun, M. Tenaillon, F. Tian, J. Wang, X. Xu, Z. Zhang, S. Kaeppler, J. Ross-Ibarra, M. D. McMullen, E. Buckler, G. Zhang, Y. Xu, D. Ware. 2012. Maize HapMap2 identifies extant variation from a genome in flux. NATURE GENETICS **44**:803-807.
7. van Heerwaarden, M. B. Hufford, J. Ross-Ibarra. 2012. Historical genomics of North American maize. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA **109**:12420-12425.
8. Kanizay, L. B., T. Pyhäjärvi, E. Lowry, M. B. Hufford, D. G. Peterson, J. Ross-Ibarra, R. K. Dawe. Diversity and abundance of the Abnormal chromosome 10 meiotic drive complex in *Zea mays*. (Submitted) Preprint archived at: http://arxiv.org/abs/1209.5466
9. Sexton, J. P., M. B. Hufford, A. Bateman, D. B. Lowry, H. Meimberg, S. Y. Strauss, K. J. Rice. Mechanisms generating species range limits: Inference from patterns of gene flow and abundance across a species' range. (Submitted)
10. Hufford, M. B., E. Martinez-Meyer, B. S. Gaut, L. E. Eguiarte, and M. I. Tenaillon. Inferences from the historical distribution of wild and domesticated maize provide ecological and evolutionary insight. PLOS ONE (In Press)
11. Pyhäjärvi, T., M. B. Hufford, S. Mezmouk, and J. Ross-Ibarra. Complex patterns of local adaptation in teosinte. (Submitted) Preprint archived at: http://arxiv.org/abs/1208.0634
12. Hufford, M. B., P. Bilinski, T. Pyhäjärvi, J. Ross-Ibarra. Teosinte as a model system for population and ecological genomics. TRENDS IN GENETICS (In Press)