Research Approach: Evolutionary genomics of maize

My research program applies population genetic approaches to understand the genomic basis of evolution change in maize (*Zea mays* ssp. mays) and its wild relatives the teosintes. Agricultural plants provide a particularly excellent system for basic discovery, with their rich history of basic genetic research, ample modern genetic resources, and the opportunity to sample or replicate genotypes across a wide array of environments. Among crop plants, maize provides perhaps the best opportunities for basic research and discovery. In addition to its status as the country's most economically valuable crop, maize is also one of the oldest model organisms, and its rich phenotypic and genetic diversity have enabled a number of basic scientific discoveries, from hybrid vigor (Shull 1908) to the relationship between crossing over and recombination (Creighton and McClintock 1931), meiotic drive (Rhoades 1942), and transposable elements (McClintock 1950). Maize continues to be an important model in the genomics era (Nannas and Dawe 2015), and the publicly available genomic and phenotypic resources available are unparalleled among other crops: more than 1,000 genomes have been resequenced, more than 20,000 individuals genotyped at ≈ 1 million SNPs, and association mapping studies now approach scales rarely seen outside of biomedical studies (e.g. 65,000 plots in Peiffer et al. 2014).

Maize is particularly well-suited as a model to study plant evolution. Both domestication and modern breeding can be seen as long-term experimental evolution, and both population genomic (Hufford et al. 2012b) and archaeological (Purugganan and Fuller 2011) data suggest that these are likely relevant and representative of patterns of evolution in natural populations. Maize has a large genome compared to many other model plants — the maize reference is 2.3Gb (Schnable et al. 2009) — but is in fact close to the median (and below the mean) genome size for most flowering plants (Leitch and Leitch 2013), and any understanding of plant evolution would be incomplete without consdiering the structural variation and repeat evolution inherent in such complex genomes. Finally, the wild relatives of maize, collectively referred to as teosinte, inhabiting a diverse array of environments, and include both endangered species and taxa with large, stable, randomly-mating populations (Hufford et al. 2012a). Because all of these species are closely related, genetic resources developed in maize are often transferred seemlessly to these wild taxa as well (e.g. Pyhäjärvi et al. 2013, Fang et al. 2012). Though much effort has been made to apply the genomic resources of maize to accelerate plant breeding, there are surprisingly few efforts to take advantage of maize as a model to more broadly understand plant adaptation and evolution. Below I describe current and planned population genetic work in my lab which aims to use maize and teosinte as a model to understand plant adaptation and genome evolution.

Domestication

Plant domestication represents an opportunity to understand not only the genetic basis of evolutionary change, but also how the complex interplay of demography and selection shape modern genetic and phenotypic diversity. My group has worked to understand the process of selection during domestication, documenting its polygenic nature (Hufford et al. 2012b) and the importance of selection on multiple beneficial mutations (Wills et al. 2013) and standing genetic variation (Wills et al. 2013, Studer et al. 2011). Our current work focuses not on finding the targets of selection,

but instead on how plant domestication has shaped geenetic and phenotypic diversity. Maize underwent a demographic bottleneck during domestication, reducing its effective population size and thus the efficacy of purifying selection to remove deleterious variants (Figure 1A). Following its domestication, however, maize spread rapidly across the Americas, and this explosive population growth — our estimates of the current effective size of open-pollinated maize exceed 10^12 — has lead to a higher efficacy of selection on new mutations, and patterns of diversity near genes suggest that new mutations in maize are more strongly selected than those in its wild-ancestor (Figure 1B). Population demography also has the potential to change the size, number, and dominance of loci underlying phenotypic traits (Lohmueller 2014, Gazave et al. 2013). Our future work will utilize data from large-scale association mapping studies in both domesticated and wild populations, coupled with population genetic inference of demographic change and the fitness consequences of new mutations, to study how the process of domestication may have shaped the genetic architecture of phenotypic traits. Results from these studies will help us better understand the functional consequences of new mutations as well as improve our utilization of genetic diversity to predict phenotype.
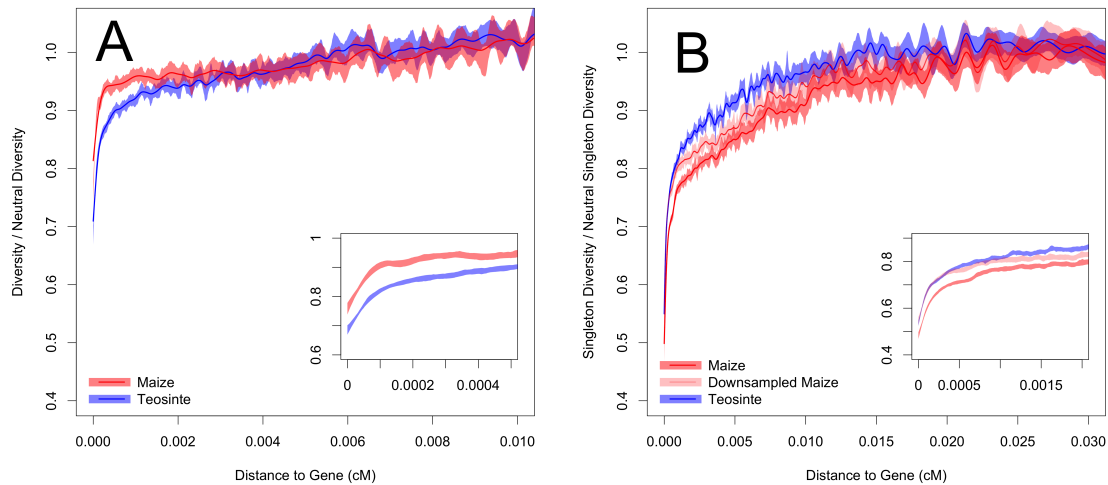

Figure 1: Relative diversity as a function of distance to nearest gene. A) Pairwise nucleotide diversity π is most influenced by intermediate frequency alleles and reflective of long-term demography. The stronger effect of purifying selection on π in teosinte is consistent with its larger long-term effective population size. B) For very recent variants such as those found in a single individuals, purifying selection appears stronger in maize, reflecting its exponential growth following domestication.
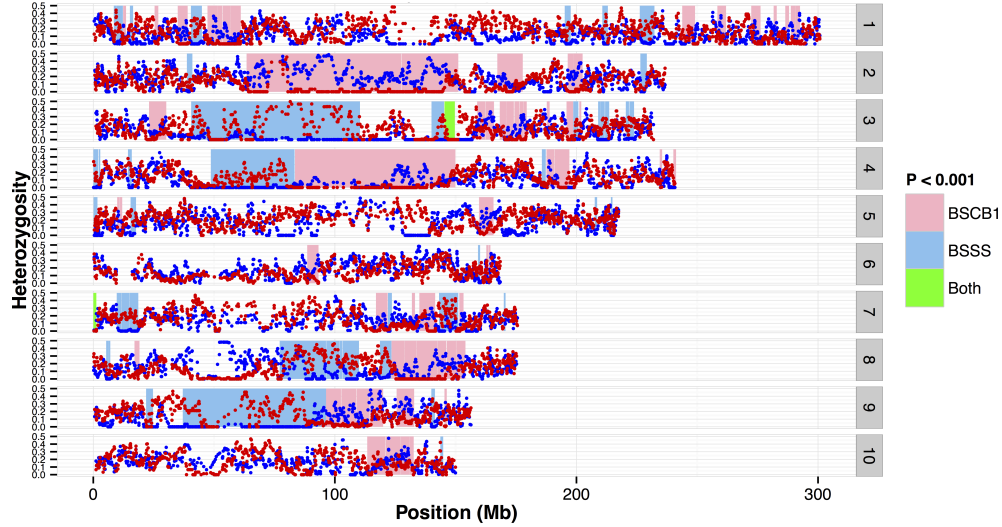
STOP HERE

2

Figure 2: Heterozygosity across all 10 chromosomes of maize from cycle 16 of the Iowa Reciprocal Recurrent Selection experiment. Heterozygosity values in the BSSS (blue dots) and BSCB1 (red dots) populations are superimposed in one panel. 2cM windows of heterozygosity lower than 10 of 10,000 simulations ($P < 0.001$) are shaded in light blue (BSSS) or pink (BSCB1). Two regions genome-wide show significantly low heterozygosity in both populations and are shaded green.

Local adaptation

After domestication, rapidly spread across the Americas, adapting to a wide range of environments. Today maize is found growing from sea level to altitudes of $> 4,000$m, and from deserts to near-flooded conditions; maize is grown across a broader geographic breadth than any of the world's other staple crops (**?**). In addition, the wild relatives of maize have adapted to environments varying widely in eleveation, temperature, and moisture gradients. Our prevous work has shown that adaptation in teosinte is often restricted to individual local populations, and that selection has taken advantage of both regulatory and structural variation (Pyhäjärvi et al. 2013). Our recent efforts have investigated convergent phenotypic evolution to high elevation habitats in maize and teosinte. This work has shown that rapid adaptation was made possible by extensive standing variation, leading to virtually no overlap in genes showing evidence of selection in spite of strong phenotypic simiarilites (**?**). We have also shown the importance of adaptive introgression, demonstraintg that gene flow from teosinte enabled maize in Mexico to rapidly adapt to novel conditions (**?**).

Experimental evolution

heterosis. artificial selection

Genome evolution

inversion polymorphisms. likely more common because no underdominance (Maguire 1966) TE insertions (Studer et al. 2011, Makarevitch et al. 2015) and regulation. genome size too (Tenaillon
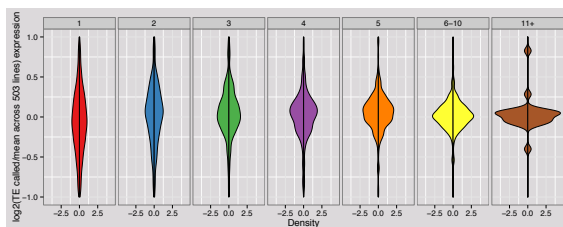
3

Figure 3: The effect of *de novo* non-reference transposable element insertion on gene expression. Panels show the impact of insertions identified in 1 (far left) to > 10 (far right) inbred lines in a panel of 23 maize inbreds. Shown in each panel is the relative expression of genes in inbred lines with the insertion within the annotated gene region to the mean expression of that gene across 500 maize lines (Hirsch et al. 2014). Recent, low frequency insertions tend to decrease gene expression compared to older, high-frequency insertions.

Figure 4: The effect of copy number variation (CNV) on estimates of Tajima's D, a measure of the allele freqeuncy spectrum. In regions with no CNV, Tajima's D is negative consistent with population expansion. Tajima's D is inferred to be strongly postiive in regions with high-frequency CNVs, however, which could be (erroneously) interpreted as population decline or balancing selection.
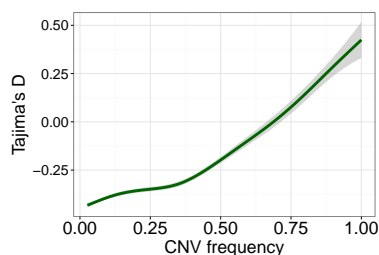
et al. 2011). paul.

michelle stats. 344K insertions, 2728 in genes!

TEs, CNVs, genome size evolution

missing data vs. allele frequencies. TEs

References

H. B. Creighton and B. McClintock. A correlation of cytological and genetical crossing-over in zea mays. *Proceedings of the National Academy of Sciences of the United States of America*, 17(8): 492–497, 1931.

Z. Fang, T. Pyhäjärvi, A. L. Weber, R. K. Dawe, J. C. Glaubitz, J. d. J. S. González, C. Ross-Ibarra, J. Doebley, P. L. Morrell, and J. Ross-Ibarra. Megabase-scale inversion polymorphism in the wild ancestor of maize. *Genetics*, 191(3):883–894, 2012.

E. Gazave, D. Chang, A. G. Clark, and A. Keinan. Population growth inflates the per-individual number of deleterious mutations and reduces their mean effect. *Genetics*, 195(3):969–978, 2013.

S. Hake and J. Ross-Ibarra. Genetic, evolutionary and plant breeding insights from the domestication of maize. *eLife*, 4:e05861, 2015.

C. N. Hirsch, J. M. Foerster, J. M. Johnson, R. S. Sekhon, G. Muttoni, B. Vaillancourt, F. Peñagaricano, E. Lindquist, M. A. Pedraza, K. Barry, et al. Insights into the maize pan-genome and pan-transcriptome. *The Plant Cell Online*, 26(1):121–135, 2014.

M. Hufford, P. Bilinski, T. Pyhäjärvi, and J. Ross-Ibarra. Teosinte as a model system for population and ecological genomics. *Trends in Genetics*, 12:606–615, 2012a.

M. Hufford, P. Lubinksy, T. Pyhäjärvi, M. Devengenzo, N. Ellstrand, and J. Ross-Ibarra. The genomic signature of crop-wild introgression in maize. *PLoS genetics*, 9(5):e1003477–e1003477, 2013.

M. B. Hufford, X. Xu, J. Van Heerwaarden, T. Pyhäjärvi, J.-M. Chia, R. A. Cartwright, R. J. Elshire, J. C. Glaubitz, K. E. Guill, S. M. Kaeppler, et al. Comparative population genomics of maize domestication and improvement. *Nature Genetics*, 44(7):808–811, 2012b.

I. J. Leitch and A. R. Leitch. *Genome size diversity and evolution in land plants.* Springer, 2013.

K. E. Lohmueller. The impact of population demography and selection on the genetic architecture of complex traits. *PLoS Genetics*, 10(5), 2014.

M. P. Maguire. The relationship of crossing over to chromosome synapsis in a short paracentric inversion. *Genetics*, 53(6):1071, 1966.

I. Makarevitch, A. J. Waters, P. T. West, M. Stitzer, C. N. Hirsch, J. Ross-Ibarra, and N. M. Springer. Transposable elements contribute to activation of maize genes in response to abiotic stress. *PLoS genetics*, 11(1):e1004915, 2015.

B. McClintock. The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences*, 36(6):344–355, 1950.

N. J. Nannas and R. K. Dawe. Genetic and genomic toolbox of zea mays. *Genetics*, 199(3):655–669, 2015.

J. A. Peiffer, M. C. Romay, M. A. Gore, S. A. Flint-Garcia, Z. Zhang, M. J. Millard, C. A. Gardner, M. D. McMullen, J. B. Holland, P. J. Bradbury, et al. The genetic architecture of maize height. *Genetics*, 196(4):1337–1356, 2014.

M. D. Purugganan and D. Q. Fuller. Archaeological data reveal slow rates of evolution during plant domestication. *Evolution*, 65(1):171–183, 2011.

T. Pyhäjärvi, M. B. Hufford, S. Mezmouk, and J. Ross-Ibarra. Complex patterns of local adaptation in teosinte. *Genome biology and evolution*, 5(9):1594–1609, 2013.

M. Rhoades. Preferential segregation in maize. *Genetics*, 27(4):395–407, 1942.

P. S. Schnable, D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T. A. Graves, P. Minx, A. Reily, L. Courtney, S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S. Rock, E. Belter, F. Du, K. Kim, R. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, B. Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar, T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D. Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W. Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M. Levy, L. McMahan, P. Van Buren, M. Vaughn, K. Ying, C. Yeh, S. Emrich, Y. Jia, A. Kalyanaraman, A. Hsia, W. Barbazuk, R. Baucom, T. Brutnell, N. Carpita, C. Chaparro, J. Chia, J. Deragon, J. Estill, Y. Fu, J. Jeddeloh, Y. Han, H. Lee, P. Li, D. Lisch, S. Liu, Z. Liu, D. Nagel, M. McCann, P. SanMiguel, A. Myers, D. Nettleton, J. Nguyen, B. Penning, L. Ponnala, K. Schneider, D. Schwartz, A. Sharma, C. Soderlund, N. Springer, Q. Sun, H. Wang, M. Waterman, R. Westerman, T. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J. Bennetzen, R. Dawe, J. Jiang, N. Jiang, G. Presting, S. Wessler, S. Aluru, R. Martienssen, S. Clifton, W. McCombie, R. Wing, and R. Wilson. The B73 maize genome: complexity, diversity, and dynamics. *Science*, 326(5956): 1112–1115, 2009.

G. H. Shull. The composition of a field of maize. *Journal of Heredity*, (1):296–301, 1908.

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

S. Takuno, P. Ralph, K. Swarts, R. J. Elshire, J. C. Glaubitz, E. S. Buckler, M. B. Hufford, and J. Ross-Ibarra. Independent molecular basis of convergent highland adaptation in maize. *Genetics*, 2015. doi: 10.1534/genetics.115.178327. URL http://www.genetics.org/content/early/2015/06/15/genetics.115.178327.abstract.

M. I. Tenaillon, M. B. Hufford, B. S. Gaut, and J. Ross-Ibarra. Genome size and transposable element content as determined by high-throughput sequencing in maize and zea luxurians. *Genome biology and evolution*, 3:219–229, 2011.

D. M. Wills, C. J. Whipple, S. Takuno, L. E. Kursel, L. M. Shannon, J. Ross-Ibarra, and J. F. Doebley. From many, one: genetic control of prolificacy during maize domestication. *PLoS Genetics*, 9(6):e1003604, 2013.