

University of Vermont

ScholarWorks @ UVM

---

Graduate College Dissertations and Theses

Dissertations and Theses

---

2020

# The Potential of Refugee Seed Systems to Promote Contemporary Adaptation in Traditional Crops: A Case Study of African Maize in New England

Andi Kur  
*University of Vermont*

Follow this and additional works at: <https://scholarworks.uvm.edu/graddis>



Part of the [Agriculture Commons](#), and the [Genetics and Genomics Commons](#)

---

## Recommended Citation

Kur, Andi, "The Potential of Refugee Seed Systems to Promote Contemporary Adaptation in Traditional Crops: A Case Study of African Maize in New England" (2020). *Graduate College Dissertations and Theses*. 1280.

<https://scholarworks.uvm.edu/graddis/1280>

This Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact [donna.omalley@uvm.edu](mailto:donna.omalley@uvm.edu).

THE POTENTIAL OF REFUGEE SEED SYSTEMS TO PROMOTE CONTEMPORARY  
ADAPTATION IN TRADITIONAL CROPS:  
A CASE STUDY OF AFRICAN MAIZE IN NEW ENGLAND

A Thesis Presented

by

Andi Kur

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements  
for the Degree of Master of Science  
Specializing in Plant and Soil Science

August, 2020

Defense Date: June 2, 2020  
Thesis Examination Committee:

Eric Bishop-von Wettberg, Ph.D., Advisor  
Daniel Tobin, Ph.D., Chairperson  
Yolanda Chen, Ph.D.  
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

## ABSTRACT

There are many mechanisms by which landraces evolve in a contemporary agricultural setting; however, the influence of forced human migration on landrace redistribution and evolution has received little attention in comparison to the stochastic effects of drift, mutation, and gene flow in the centers of origin. Although the seed systems of forcibly-displaced people remain poorly understood, evidence suggests that refugees often continue to grow traditional crops after resettlement. From a genetics perspective, the crops that are transported to highly disparate environments provide an interesting opportunity to study adaptation.

This research addresses how forced human migration has impacted contemporary landrace evolution in a specific case study of African maize being grown by new American farmers in Vermont and New Hampshire. We utilize a whole genome sequencing approach and methods in population genetics to investigate the origin, genetic diversity, and potential adaptation in these crops. Our findings suggest that maize grown by new American farmers in the study does in fact originate from Africa and that each farmer is growing a genetically distinct crop, although we are unable to link origin of the crop to immigration history of the farmer. We also found that genetic diversity is remarkably high across all samples, even compared to landrace panels assembled from the U.S. National Plant Germplasm System. Lastly, we found numerous signatures of positive selection across all farmer samples, and through Gene Ontology analysis, we identified two significant biological processes enriched by positive selection, (1) cinnamoyl-CoA reductase biosynthesis and (2) glutathione synthase activity, that may indicate recent adaptation and be correlated to increased cold tolerance. Overall, through this case study we show a specific example of how forced human migration has affected landrace redistribution, subsequent evolution, and potential adaptation to a disparate environment. We believe these findings hold interesting implications towards agrobiodiversity conservation and suggest the potential of refugee seed systems to promote contemporary adaptation in traditional crops.

# TABLE OF CONTENTS

	Page
LIST OF TABLES .....	iv
LIST OF FIGURES .....	v
INTRODUCTION .....	1
CHAPTER 1: COMPREHENSIVE LITERATURE REVIEW .....	8
1.1 Landrace Conservation .....	8
1.2 Mechanisms of Landrace Evolution .....	9
1.2.1 Mutation .....	10
1.2.2 Genetic Drift .....	10
1.2.3 Gene Flow.....	12
1.2.3.1 Gene Flow From Modern Varieties.....	12
1.2.3.2 Gene Flow Between Landraces.....	14
1.2.3.3 Gene Flow and Human Migration.....	15
1.3 Forced Migration, Seed Systems, and Food Security .....	16
1.3.1 Forced Migration and Smallholder Farmers .....	16
1.3.2 Implications on Seed Exchange Networks .....	17
1.3.3 Food Insecurity in Resettled Populations .....	18
1.4 Connecting the Gaps: Diaspora and Contemporary Landrace Evolution .....	20

CHAPTER 2: A CASE STUDY OF AFRICAN MAIZE IN NEW ENGLAND.....	22
2.1 Materials & Methods .....	26
2.1.1 Samples and Accessions .....	26
2.1.2 Plant Material, DNA Extraction, and Sequencing .....	29
2.2.3 Data Analysis .....	30
2.2 Results .....	36
2.2.1 Principal Components Analysis .....	36
2.2.2 Population Admixture and Ancestry .....	37
2.2.3 Pairwise Diversity and Tajima’s D .....	40
2.2.3 Selective Sweeps and Gene Ontology.....	42
2.3 Discussion .....	44
2.3.1 Origin of Seed .....	44
2.3.2 On-Farm Diversity .....	46
2.3.3 Selection and Potential Adaptation .....	47
2.4 Conclusion .....	49
COMPREHENSIVE BIBLIOGRAPHY .....	51

## LIST OF TABLES

Table	Page
Table 1. Sample Information .....	29

## LIST OF FIGURES

Figure		Page
Figure 1.	Graphical representation of the distribution of New Farms and Fresh Start maize samples based on farmer-reported nationality (purple) as compared to the distribution of the Africa/Asia Landrace Panel (red). The AA landrace panel was constructed from accessions in the U.S. National Plant Germplasm System to best represent possible origins of new American farmer samples, considering both immigration history and ethnicity .....	28
Figure 2.	Principal Components Analyses plotting the covariance matrices output from PCAngsd, showing: (Left) variation in all samples, including New Farms and Fresh Start, and both landrace panels and (Right) variation between samples from both farm locations. Plotted in ggplot with added ellipses to show major groupings by sample type.....	37
Figure 3.	Admixture analysis to investigate shared ancestry between New Farms and Fresh Start samples and the Africa/Asia and North America landrace panels. Number of subpopulations was increased from K=3, 4, 6, listed next to admixture proportions on the left of the structure plots.....	39
Figure 4.	Admixture analysis comparing samples from each farm (New Farms for New Americans in Burlington, VT and Fresh Start Farm in Dunbarton, NH) to investigate how distinct one farm’s crop is from another. For each farm location, number of subpopulations was set to K = number of farmers (listed above each plot) to see if samples grouped by farmer. Each subpopulation is a different color, and admixture proportions are labeled on the Y axis. Numbers on the X axis represent farmer ID’s and correspond to Table 1 .....	40
Figure 5.	Estimates of genome-wide Pairwise Diversity (top) and Tajima's D (bottom) across sample groups (left) and samples per farmer (right). All sample groups are labeled along the X-axis, and the Y-axis shows the range of values for pi and Tajima’s D, respectively. Values were achieved using the Thetas method in Angsd.....	41
Figure 6.	Selective sweeps output from RAiSD, plotted at 99.9% mu significance threshold. The top plot shows sweeps analyzes for the full sample groups (AA panel, NF, and FS) and the bottom plot shows sweeps from each farmer individually, labeled as New Farms or Fresh Starts. Each chromosome is represented by the horizontal grey box, with chromosome number labeled to the right and the position along the chromosome labeled along the X-axis.....	43

## INTRODUCTION

Conservation of agrobiodiversity is increasingly recognized as an imperative measure to ensure future agricultural stability and food security (Maxted *et al.*, 2012). Agrobiodiversity refers to all genetic resources that promote sustainability in agricultural production ecosystems, which includes diversity found across the full spectrum of food crops- from wild-harvested plants to the formally bred modern varieties. In the center of this spectrum are landrace crops, also known as traditional farmer varieties. Landraces are considered valuable sources of agrobiodiversity for being more diverse than modern varieties and generally more resilient to biotic and abiotic change in low-input systems (Harlan, 1975, Ceccarelli, 2012). Landrace crops play a significant role in constituting the informal seed networks of smallholder farmers in their centers of origin (Almekinders, Louwaars, and Debruijn, 1994), and inherent in their connection with traditional farm systems, landraces are also regarded as valuable in their own right as cultural heritage.

### *Conservation & Contemporary Adaptation in Landraces*

Conservation of landraces is achieved by both *ex situ* and *in situ* strategies. *Ex situ* conservation refers to the preservation of germplasm in gene banks, which provide a controlled storage environment that is readily accessible to breeders and scientists. Landraces tend to be underrepresented in *ex situ* collections, however, and by removing germplasm from the field, *ex situ* conservation also effectively halts evolutionary response (Hammer, 2004). Thus, *ex situ* conservation is often complemented by *in situ* conservation, where farmers are incentivized to maintain dynamic populations of landraces in their center of origin, allowing for continued adaptive evolution and connection to traditional growing



systems. Unfortunately, *in situ* conservation can be difficult to sustain in its idealized sense, as it often tasks smallholder farmers with growing restrictions that may be untenable given their personal socioeconomic realities (Wood and Lenné, 1997). A more nuanced issue prevalent in both conservation strategies is their tendency to treat landraces as if they are resources to be archived and “frozen-in-time” (Iltis, 1974; Peres, 2006), and even *in situ* conservation, which promotes continued adaptive evolutionary response, tends to allow genetic change within restricted limits (for instance, denying the influence of formal breeding or adaptation outside of the center of origin).

New definitions challenge these older ideologies and argue for the importance of recognizing contemporary evolution in landrace populations (Casañas *et al.*, 2017). An important point in the argument is that contemporary evolution in landraces does not have to contradict their conservation, so long as genetic changes are distinguished into those that are beneficial to agrobiodiversity (e.g. adaptation and diversification) and those that are negative (e.g. genetic erosion and swamping).

In general, there are many mechanisms by which landraces evolve in a contemporary agricultural setting. Basic principles of population genetics (i.e. mutation, drift, and gene flow) all continually act on landrace populations, along with human-mediated and natural selection. Especially in outcrossing crop species such as maize, there are a significant number of documented cases of landraces experiencing a change in genetic composition via gene flow from sympatric modern varieties (Bitocchi *et al.*, 2009; van Heerwaarden *et al.*, 2012), from co-occurring wild relatives (Rojas-Barrera *et al.*, 2019), and gene flow from other landraces (van Heerwaarden, van Eeuwijk, and Ross-Ibarra 2010). As DNA sequencing technology has become more affordable and data processing

pipelines more accessible, increased attention has been given to how these mechanisms affect landrace gene pools.

### *Human Migration and Landrace Transportation*

The influence of human migration on contemporary landrace evolution has received little attention in comparison to the effects of drift, mutation, and gene flow. Primarily, there is a distinct gap in the literature addressing landrace redistribution through human migration networks. On one hand, seed system analysts have established that landrace crops are significant constituents of informal seed systems, which are heavily relied upon by agrarian communities in less developed countries (Almekinders, Louwaars, and Debruijn, 1994). Agrarian communities are considered to be at greater risk of experiencing forced migration due to environmental disaster, conflict, or economic collapse (FAO *et al.*, 2018). Forced migration is tightly coupled with food insecurity, whereby food availability, access, and stability operate as both push and pull factors driving direction, timing, and magnitude of migration (FAO, 2017). One consequence of forced migration that exacerbates food insecurity is cultural displacement, which results in severed access to traditional and familiar food markets. Lastly, as landrace varieties are often available only through informal seed networks in a particular region, migration outside of said region would also sever continued access to seed.

Thus, there are many observed incentives that would drive displaced people, and especially displaced farmers, to bring seeds with them as they migrate; however, at the time of this writing, there is little available literature addressing the effects of forced migration on landrace transportation and redistribution. There are, however, many

examples of refugee and immigrant farmers who grow traditional vegetables upon resettlement into the United States (hereafter, new Americans). These examples can be found in the numerous news articles that highlight the contributions of refugee and immigrant communities to the diversity of vegetables found at local farmers markets and in exposés written about the growing network of new American community gardens (see: “Gardening brings taste of home to refugee families in Chicago,” 2017; “Gardens offer refuge to refugees,” 2018; “Putting down roots: Refugee farms and gardens growing strong,” 2019).

While there are many examples where new American farmers continue to grow traditional crops after resettlement, little is actually known about the seed systems of U.S.-based immigrant and refugee populations themselves. Anecdotally, these seed systems are formed via informal sharing networks between farmers, whereby seeds are originally sourced from food purchased at supermarkets, brought over during immigration, or imported after resettlement. This arrangement often results in farmers planting seed that is unsuitable or poorly adapted to the new growing environment, especially in instances where international migration causes resettlement to substantially different climate zones. For example, new American farmers in Vermont have introduced several landrace varieties of African Eggplant into the local crop market with considerable success, but they are unable to grow the plant to its typical 2-year maturity because the winters create an untenable climate for the tropical vegetable. In another example in northern Kansas, new American farmers from Vietnam have been growing jicama, a long-season crop that requires especially hot environments and, in Vietnam, is traditionally intercropped under the canopy of banana plantations. According to Kindscher (2009), gardeners report

substantially lower yields in northern Kansas but are motivated to continue growing the crop because of its cultural significance and high price point in local markets. These are just two examples out of many, but they highlight an interesting theme in new American gardens- that is, the challenge and potential of adapting to new environments.

### *Potential for Adaptation*

Although the means by which landrace crops are redistributed via contemporary human migration is poorly described, it is evident that landrace varieties of different crops are often transported to new environments and maintained in modified traditional farming systems. Landrace performance in these modified systems spans from failure to success and, in between, the crops that experience changes in growth habits, phenology, and production in response to new selective pressures. From a genetics perspective, landrace crops that are transported to highly disparate environments provide an interesting opportunity to study adaptation. For poorly adapted crops that survive to be replanted from season to season, knowledge of the mechanism(s) by which they adapt to the new environment, such as selection on standing variation or introgression from modern varieties, may provide insight into the adaptive capacity of different landraces and potentially help identify functional genes to target in breeding populations. Heraty and Ellstrand (2016) further suggest that resettled refugee and immigrant gardens may serve as a third form of agrobiodiversity conservation that combines both *in situ* and *ex situ* approaches. They suggest that while such “crops of diaspora” may exhibit reduced diversity as a result of genetic drift, these crop populations may also be uniquely diverse as a result of multiple introductions, cross pollination between other introduced varieties,

and new selective pressures. Consequently, cropping systems managed by refugee and immigrant farmers may catalyze contemporary adaptation in landrace populations outside of their centers of origin and promote global agrobiodiversity.

### *Purpose of Study*

This research addresses how contemporary human migration has impacted crop evolution and adaptation in a specific case study of maize being grown by new American farmers in Vermont and New Hampshire. For this study, we partnered with two nonprofit refugee and immigrant support organizations, New Farms for New Americans (NF, located in Burlington, Vermont) and Fresh Start Farms (FS, located in Dunbarton, New Hampshire). Both programs support immigrant communities by providing access to land, equipment, and translation services via a community garden (NF) and farm incubator (FS) approach.

We chose to focus on maize because, in this system, it exemplifies many of the conditions that occur as a result of recent transportation of seeds through human migration networks and that may enact genetic change in the crop. The three primary conditions of which we are concerned with are: (1) transportation to a new climate zone or environment that the crop may be poorly adapted to; (2) transportation of a limited number of seeds, which may cause decreases in landrace genetic diversity; and (3) cultivation in community garden environments, which allows interaction between previously separated landraces.

At both farm locations, farmers grow tropical, open-pollinated maize that they have reportedly brought with them from Africa when they migrated to the United States. We do not have precise data on individual farmer's migration history, but we do know that farmers

in both programs have been growing African maize in their respective gardens for 5-7 years and saving seed from their harvests to plant for the next season. In both gardens, farmer's garden plots are separated by a ~2 foot walking path, meaning that maize grown by different farmers are also in range to cross-pollinate with one another. Thus, while individual farmer's seed stocks are distinct in their origins, the past 5-7 years have created many opportunities for cross pollination and interaction between previously separated crop varieties, providing an interesting opportunity to study how the farmer's crops have changed in diversity and genetic composition as a result of being grown in a community garden. Lastly, we are interested in maize grown in these gardens because it provides an interesting case in which to investigate adaptation as a result of transportation via human migration. Theoretically, tropical maize transported to a temperate environment should be maladapted and perform poorly, especially in terms of flowering time and cold tolerance; however, farmers in our study are able to obtain a reasonable yield from their crop- enough to provide food for home, sell at local farmer's markets, and save seed for the next season. We suspect that as farmers save seed between seasons, some level of selection is taking place on traits that aids in temperate adaptation.

Ultimately, knowledge of if and how maize has adapted within this system will add to our understanding of how human migration affects contemporary adaptation in landraces and, as a result, help inform landrace conservation strategies and guide the formation of participatory breeding programs for migrant communities.

## CHAPTER 1: COMPREHENSIVE LITERATURE REVIEW

### 1.1 Landrace Conservation

As defined by Camacho-Villa *et al.* (2005), landraces are “dynamic population(s) of a cultivated plant that has historical origin, distinct identity and lacks formal crop improvement, as well as being genetically diverse, locally adapted and associated with traditional farming systems.” The conservation of landraces is increasingly recognized as an imperative measure to ensure future agricultural stability and food security because landraces are generally more resilient to biotic and abiotic change in low-input systems and are valuable sources of agricultural biodiversity (Harlan, 1975; Ceccarelli, 2012). This noted, it is prudent to recognize that conservation directives and justifications for these directives are framed by two different (but, still compatible) motivations. One views conservation of landraces as important for the preservation of genetic resources to be utilized for public and private breeding sectors, driven by concerns regarding the instability of current monocultures to maintain yields under increasing biotic (e.g. pests, pathogens) and abiotic (e.g. drought, soil depletion) stressors (Hufford, Berny Mier y Teran, and Gepts, 2019). The other motivation for landrace conservation considers the significant role landraces hold in constituting local seed systems (Almekinders, Louwaars, and Debruijn, 1994). Unregulated by industry, landraces provide a measure of food sovereignty to smallholder farmers who need not rely on market fluctuations for seed access and variety availability. Lastly, inherent in their connection to a people and a place, landraces are considered valuable as a people’s cultural heritage (Graddy-Lovelace, 2013).

When discussed by conservationists, the definition of a landrace is often framed around this connection to tradition and heritage, which is sometimes conflated with the ideology that landraces are “pure” as static resources (Thomas *et al.*, 2011). Casañas *et al.* (2017) make the poignant statement that “most conservative definitions doom landraces to become artifacts in museums,” and they go on to promote a new, inclusive definition of landraces: as “plant materials consisting of cultivated varieties that have evolved and may continue evolving, using conventional or modern breeding techniques, in traditional or new agricultural environments with a defined ecogeographic area and under the influence of local human culture.” Thus, new definitions of landraces include contemporary change and adaptation, which can conditionally occur outside of the traditional setting of a landrace without negating their conservation value.

## **1.2 Mechanisms of Landrace Evolution**

There are many mechanisms by which landrace crops change in a contemporary agricultural setting. Primarily, there is continual artificial selection, the process by which farmers affect genetic change in a crop population by choosing preferred phenotypes to replant from season to season. While artificial selection is the most primary means by which landraces evolve, farmers are only able to select on variation that already exists within their crop. Thus, the other mechanisms that are important to landrace evolution are those that affect the presence or absence of this variation in a crop population. The most important mechanisms to discuss within this context are mutation, genetic drift, and gene flow. These are some of the same mechanisms that govern evolution in natural systems; however, in an agricultural system, these processes are highly influenced by human actions.



### 1.2.1 Mutation

In the short term, mutation alone is unlikely to drive large-scale, contemporary evolution of landraces because the majority of mutations are (1) in the non-coding regions of DNA and are therefore of no phenotypic effect, or (2) deleterious and not passed on in the population. The potential for new mutations to shape landrace evolution, however, should not be overlooked. It is evident from the crop domestication literature that mutations have played a significant role in generating the genes that were selected upon by farmers throughout the ~12,000 years of global crop domestication (Meyer and Purugganan, 2013). For example, a single amino acid substitution (Lysine to Asparagine) in the transcription factor of the *Teosinte glume architecture1* (Tga1) gene of maize resulted in a major phenotypic transition during the crop's initial domestication ~6,200 years ago, causing the hardened kernels of teosinte to become exposed and remain attached to the cob (Dorweiler and Doebley, 1997; Wang *et al.*, 2005). Contemporary examples where mutations affect large phenotypic change are less common; however, one excellent example from recent literature describes an incident where a *de novo* mutation in a landrace crop generated a novel and desirable trait. In this study, Ootsuka *et al.* (2014) found a splicing mutation in the *badh2.10* gene of Japanese rice landraces that results in a novel fragrance allele in the rice. Their evidence suggests that this mutation occurred quite recently relative to the history of rice cultivation and has increased in frequency due to conscious selection by Japanese farmers.

Ultimately, mutations continue to create new genetic diversity in all living populations. In landrace populations, where farmers consciously select their seeds for performance and quality from season to season, a beneficial mutation, such as one that

would result in a uniquely fragrant rice, is more likely to be noticed and selected for than it would in a commercial agricultural setting.

### **1.2.2 Genetic Drift**

Like mutation, genetic drift is also a stochastic (random, systematic) process; however, in contrast to mutation, drift generally results in the removal of diversity from a population. Genetic drift describes the change in allele frequencies in a population due to chance sampling of alleles between generations. In an on-farm setting, genetic drift is of particular concern when a landrace crop is grown on a small scale or in fragmented populations (Ellstrand and Elam, 1993). This is because populations with small demographic sizes are more susceptible to genetic drift: in terms of probability, the range of sampling error is larger when there are less samples to choose from (Kliman, Sheehy, and Schultz, 2008). In other words, it is easier in smaller populations for rare alleles to be lost and for recessive, deleterious alleles to increase in frequency, purely by chance. Thus, as seeds are saved between generations, drift can cause an overall decrease in genetic diversity and an increase in genetic load over time. The effects of genetic drift are frequently reported in the literature for many groups of crop landraces, for example in maize (van Heerwaarden *et al.*, 2012), emmer wheat (Rafeipour *et al.*, 2016), bread wheat (Thomas *et al.*, 2012), Swedish field pea (Leino *et al.*, 2013), and Brassica (Takuno *et al.*, 2010). Ultimately, conservation intervention is aimed towards monitoring and mitigating the effects of drift on landrace populations to prevent genetic erosion.

### **1.2.3 Gene Flow**

As the name implies, gene flow is the transfer of genetic material between populations; but the process itself is not as straight-forward in an agricultural setting as it is in naturally reproducing populations. For instance, gene flow between crops can be both intentional, such as when a breeder purposefully transfers pollen between populations, or unintentional, such as when pollen from one farmer's field makes its way over to another. Furthermore, the potential for gene flow is heavily influenced by migration, which in an agricultural system is achieved by the transfer of seed to new locations; therefore, seed management and sharing networks are also important determining factors in the rate and amount of gene flow between populations. Lastly, it should be noted that judging the significance of gene flow on landrace populations can be fairly complex, as its significance is heavily nuanced by opinion of how landraces should or should not change under different conservation paradigms.

#### **1.2.3.1 Gene Flow From Modern Varieties**

Since the release of modern crop varieties in the 1970's and 80's, the effects of gene flow on landrace populations has received increasing attention from breeders, geneticists, and social scientists, escalating in the 90's over concern of the potential escape of transgenes from genetically modified (GM) varieties (Dale, 1994; Bruce, 2003; Kwit *et al.*, 2011). Occurrences of gene flow from modern varieties into landrace populations has since been documented in rice (Ishikawa *et al.*, 2006), barley (Belluci *et al.*, 2013), and most pronouncedly in maize (Bellon and Berthaud, 2004).

As modern crop varieties are increasingly assimilated into traditional farming systems, concern has also arisen regarding the potential for genetic swamping, where local genotypes are replaced with alleles of modern varieties via gene flow. Genetic swamping can be expedited by a process called creolization, in which farmers incorporate modern varieties into the informal seed system and manage them in the same way as landraces (Bellon and Risopoulos, 2001). As noted by van Heerwaarden *et al.* (2009), however, depending on the dynamics of the system, creolization may replace local diversity or provide a source of new genetic material for farmer selection. Furthermore, it is important to recognize that gene flow between populations is not necessarily followed by introgression. Introgression occurs when genes received via gene flow continue to segregate in a population, and the possible results of introgression are not limited to genetic swamping. In addition to reducing local adaptation (often measured as  $F_{st}$ , or proportion of alleles specific to a subpopulation), introgressed genes can also promote increased nucleotide diversity (also measured as pairwise diversity) or even confer an adaptive advantage.

For example, in a study on contemporary evolution in maize landraces, Rojas-Barrera *et al.* (2019) found evidence that suggests gene flow from modern varieties into Mexican maize landraces resulted in decreased  $F_{st}$  values compared between landraces collected from 1970 to those collected after 2000. They also report, however, that levels of genome-wide pairwise diversity ( $\pi$ ) increased over the same timeframe, probably as a result of gene flow from modern varieties. In a study of maize landraces in central Italy, Bitocchi *et al.* (2009) report that although they detected low levels of introgression from dent hybrid corn into Italian flint landraces, there was no evidence to suggest genetic

erosion as a result. In a follow-up investigation, Bitocchi *et al.* (2015) report evidence of adaptive introgression from modern varieties into Italian maize landraces, where conscious selection on new alleles positively affected starch content in the landrace populations.

### **1.2.3.2 Gene Flow Between Landraces**

While much of the literature focuses on how landraces are affected by gene flow from modern varieties, genes can also be transferred between landrace populations themselves and cause contemporary evolution. Shifting cultivation practices, both spatially and temporally, can reintroduce gene flow between previously separated landrace populations. For example, in response to changing rainfall patterns, farmers in Sahel of Africa have modified their cultivation practices by planting early and late landraces of pearl millet in adjacent fields, promoting gene flow between previously separate landraces (Lakis *et al.*, 2012). To an even greater effect on gene flow are the seed exchange networks themselves, because they are the most important means by which landraces “migrate” throughout their range. Unlike natural plant populations, interactions between landraces are rarely structured by a simple isolation-by-distance model: rather, because landraces depend on seed exchange networks to be spread, their genetic diversity and population structure is also influenced by the interplay of more complex factors, such as ethnolinguistic regions, access to markets and agricultural extension services, road infrastructure, and marriage networks (Jensen *et al.*, 2013; Hodgkin *et al.*, 2007; Teshome, Brown, and Hodgkin 2001). The effect of gene flow via seed exchange networks can be dichotomous, resulting in both decreases and increases in landrace diversity and population structure.

A good example of this dichotomy can be found in barley. In Morocco, Jensen *et al.* (2013) found that seed exchange facilitated by traveling market vendors has resulted in recent increases in gene flow between the barley landraces grown by each commune and subsequent decreases in local diversity within communes. In contrast, Liu *et al.* (2019) found that seed exchange customs in Northwest Yunnan, China, especially those driven by religious use of hulless barley, have had a significant influence on generating and maintaining genetic diversity between villages.

### **1.2.3.3 Gene Flow and Human Migration**

Thus, seed exchange networks perhaps have the greatest potential to affect the evolution of traditional crop varieties, because the way that seeds are distributed across agroecosystems ultimately influences the direction of gene flow, degree of genetic drift, and amount of variation that can be selected upon in a population. These networks are highly complex and woven by a myriad of factors that are deeply influenced by a region's history, societal construction, and environment. Perhaps as a result of this complexity, seed system analyses of informal exchange networks are often regionally bound, and notably, there is a distinct gap in the literature regarding how seed exchange networks are affected by global human migration. Furthermore, few to no studies have been done to investigate how human migration has affected the contemporary redistribution of landraces to new agroecosystems. Ultimately, in order to discuss how migration might affect landrace redistribution and postulate its effects on contemporary landrace evolution, it is critical to review how forced migration affects smallholder farmers first.

## 1.3 Forced Migration, Seed Systems, and Food Security

### 1.3.1 Forced migration and Smallholder Farmers

The FAO defines forced migration as “a reactive move of last resort to survive an event or situation that severely challenges safety, security, or livelihoods in the place of origin” (FAO *et al.*, 2018). Forced migration falls on the extreme end of a continuum of migration decisions, and position on this continuum is greatly influenced by an individual or family’s capacity and resources to adapt to situational stress (FAO *et al.*, 2018). People living in rural regions of less developed countries (LDC) are less likely to have the capacity to adapt to severe situations, and are therefore more likely to experience *forced* migration over *voluntary* migration. Furthermore, rural and agricultural areas in LDCs notably bear the brunt of protracted global conflicts (FAO, 2017) and are more vulnerable to disaster and famine driven by climate change (Morton, 2007).

Thus, a large share of internal and international migrants come from rural, agrarian areas. Exact numerical estimates are challenging because there is no mandate for countries to document their internally displaced people, and numbers of international migrants are confuscated by differences in country-to-country definitions and data collection strategies. According to estimates made by the FAO (2017), there are ~1 billion internal migrants within LDCs and 68.5 million forcibly displaced persons internationally, of which ~40% come from rural areas (approximated by remittances). Reports do not specifically address the proportion of forcibly displaced persons who make their living as smallholder farmers; however, given that smallholder farmers make up >50% of the rural population in LDCs (>75% in Sub-Saharan Africa) and that smallholder farmers are especially vulnerable on

the migration-decision continuum, it is likely that they make up a significant proportion of the forcibly-displaced populations (Jazairy, Alamgir, and Pannuccio, 1992).

### **1.3.2 Implications on Seed Exchange Networks**

Little is explicitly documented about how migration affects the construction and viability of local seeds systems, and yet it is clear that forced migration may significantly interrupt the livelihoods of smallholder farmers who are the drivers and wardens of local seed exchanges. Consequently, the disruption of these networks may have serious implications for landrace conservation.

Two main outcomes are worth considering in this context. Primarily, an extreme disruption in a seed network (such as one caused by protracted conflict or natural disaster) may result in landrace varieties being lost or replaced by high-yielding varieties proffered from international food aid stores. Second, an intriguing but perhaps more nebulous consequence, is the potential for landrace redistribution. While informal seed networks are regionally circumscribed, the network itself exists when seed is exchanged (Thomas *et al.*, 2011). Thus, if displaced farmers decide to bring seeds with them, the result may be the alteration and expansion of a seed network, rather than its dissolution. One BBC News article expands on the desire for seeds in refugee camps, quoting “one of the things they [refugees] pack when they leave is seeds... They want seeds from their own plants, from their own spaces, from their own families” (“Seeds of Hope: The Gardens Springing up in Refugee Camps,” May 21, 2018). The article goes to describe seed access and gardening dynamics in a Syrian refugee camp in Domiz, Kurdistan: “gardeners in the camp grow seeds and plants brought to them from Syria by brothers, uncles and cousins. The



community has come together to build raised beds, which are planted with vegetables and flowers.”

While formal documentation on seed transportation through refugee camps is limited, patterns of migration can be informative in predicting how a seed network could change in response to forced migration. In most situations, forced migration tends to follow a fragmented stepwise pattern, where displaced people migrate between refugee camps and host countries for years before achieving stable resettlement in another country or are able to return home (FAO *et al.*, 2018). In this way, migration patterns tend to expand outward from the origin of displacement, and if we are to superimpose hypothetical seed networks on top of migration patterns, host countries and refugee camps become fascinating nodes of potential exchange, where seed from previously disjunct regions can interact, such as in the communal garden beds at Domiz. In the most extreme extension of this network, forcibly-displaced people may transport seed across continents and between disparate environments, effectively introducing landraces to entirely new agroecosystems that would consequently affect their genetic composition and potential to evolve.

### **1.3.3 Food Insecurity in Resettled Populations**

In order to understand the factors that might incentivize forcibly displaced people to transport seeds with them from their home countries, it is important to consider the connection between forced migration and food security.

Forced migration and food security are complexly linked global issues, and the relationship between the two is often cyclic, one feeding into the other over time. This is especially true for smallholder farmers, for whom food insecurity is both a major cause and

result of migration (FAO, 2017). Rates of food insecurity in refugee populations resettled in industrialized nations have been reported as high as 85% (Wang *et al.*, 2016). Common barriers to food security upon resettlement include foreign language, unfamiliar foods, lack of financial resources, and limited access to markets, reflecting issues in all four dimensions of food security: availability, access, stability, and utilization (FAO, 2008). In response to increasing levels of food insecurity for migrant populations, many non-profit organizations have led initiatives in establishing community garden networks to provide support to resettled migrant populations. Survey studies of resettled refugees who participate in gardening programs report that the gardens significantly promote (1) mental and physical health, (2) self-sufficiency, and (3) a sense of identity and community that can ease some stressors of the resettlement process (Jean, 2015; Patil *et al.*, 2010; Hartwig and Mason, 2016; Gichunge and Kidwaro, 2014).

An underlying theme to food insecurity that is significantly tied to international migration is availability of familiar and culturally appropriate foods. Remarkd by Gichunge, Somerset and Harris (2016), “culture is a determinant of food habits, and food is central to individual identity.” Thus, severed access to familiar crops and cuisines can inflict injury to both food security and to sense of self. Upon resettlement, forcibly displaced people are often faced with unfamiliar markets and the unavailability of foods that are central to their home cuisines. Several studies have shown that many resettled refugees and immigrants will resort to growing traditional vegetables in community gardens and in their own home gardens if the vegetables are unavailable or too expensive in stores (Gichunge, Somerset and Harris, 2016; Hartwig and Mason, 2016). To this end, Jean (2006) suggests that farming traditional vegetables has an important impact the

“place-making process” for resettled refugees, stating that: “Although refugees are unquestionably displaced people, they are also *emplaced* people who make choices, negotiate, and effectively create new places as they bring the weight of their historical sense of place into new spaces.”

#### **1.3.4 Connecting the Gaps: Diaspora and Contemporary Landrace Evolution**

In conclusion, there are many challenges in estimating the effects of forced human migration on the current distribution and potential redistribution of traditional crop varieties. Specific documentation of internal migration flows are lacking, and there are even greater information gaps regarding how migration of smallholder farmers affects local seed networks. Documentation does exist for the transfer of seeds through refugee camps, which create informal nodes of seed exchange that can extend well outside of the variety’s original range; however, this information is not consolidated and exists largely in the informal literature. With so much uncertainty surrounding seed dispersal through human migration networks, it is subsequently difficult to postulate on the specific effects of diaspora on the genetic composition of landrace crops. Knowledge of the basic mechanisms of crop evolution (i.e. selection, mutation, drift, and gene flow) can shed light on hypothetical changes that may occur, such as genetic erosion, swamping, diversification, or adaptation. These changes are especially interesting from an agrobiodiversity conservation perspective, and they also present interesting scenarios in which to study mechanisms of short term crop adaptation.

Ultimately, filling in these information gaps is challenging and currently only possible through informed conjecture. Noting this, it may be more feasible to study these

networks by starting with the traditional crops being grown by resettled refugees rather than by tracking the passage of these seeds through migration networks. Using methods in population genetics, it is possible to investigate the contemporary evolution of landrace crops in specific case studies, which would ultimately inform the system as a whole.

## CHAPTER 2: A CASE STUDY OF AFRICAN MAIZE IN NEW ENGLAND

### 2.1 Introduction

The influence of contemporary human migration on landrace redistribution and evolution remains poorly described and has received little attention compared to the effects of drift, mutation, and gene flow. Of particular interest are events of forced migration and internal displacement, which often affect rural agricultural areas in less developed countries and can disrupt seed exchange networks (FAO *et al.* 2008). The structure of seed exchange networks perhaps have the greatest potential to affect contemporary evolution in landraces, because the distribution of seeds across agroecosystems ultimately influences the direction of gene flow, degree of genetic drift, and amount of variation that can be selected upon in a population (Jensen *et al.*, 2013; Hodgkin *et al.*, 2007; Liu *et al.*, 2013; Teshome, Brown, and Hodgkin 2001).

Although the way in which these seed exchange networks are affected by events of human migration remains poorly described in the literature, it is evident that people who are forcibly displaced often transport landraces as they migrate and continue to cultivate them in modified traditional farming systems upon resettlement. Examples can be found in the numerous news articles that highlight the contributions of refugee and immigrant communities to the diversity of vegetables found at local farmers markets and in exposés written about the growing network of new American community gardens (see: “Gardening brings taste of home to refugee families in Chicago,” 2017; “Gardens offer refuge to refugees,” 2018; “Putting down roots: Refugee farms and gardens growing strong,” 2019). Landrace performance in these modified systems spans from failure to success and, in

between, the crops that experience changes in growth habits, phenology, and production in response to new selective pressures. From a genetics perspective, landrace crops that are transported to highly disparate environments provide an interesting opportunity to study adaptation. For poorly adapted crops that survive to be replanted from season to season, knowledge of the mechanism(s) by which they adapt to the new environment, such as selection on standing variation or introgression from modern varieties, may provide insight into the adaptive capacity of different landraces and potentially help identify functional genes to target in breeding populations.

This research addresses how contemporary human migration has impacted crop adaptation through a specific case study of maize being grown by new American farmers in Vermont and New Hampshire. For this study, we partnered with two nonprofit refugee and immigrant support organizations, New Farms for New Americans (NF, located in Burlington, Vermont) and Fresh Start Farms (FS, located in Dunbarton, New Hampshire). Programs support migrant communities by providing access to land, equipment, and translation services via a community garden (NF) and farm incubator (FS) approach. We chose to focus on maize (*Zea mays* spp. *mays* (L.)) in this system for several reasons:

Our primary motivation for focusing on maize grown by new American farmers at NF and FS is because farmers in both programs grow seed that is reportedly of African origin and suspected to be landrace-derived. This information was initially obtained through personal communication with the respective program directors, but there are several lines of evidence that corroborate reports that the crop is not derived from American seed stocks. First, new American farmers verbally recognize this crop as distinct from the maize varieties available in North America and often refer to it by name in their native

languages (e.g. *nafaka* in Kiswahili, *bu kay tha* in Karen, *ikigori* in Kirundi). Furthermore, the crop is vegetatively and culinarily distinct from American sweet corn. African maize varieties grow more vigorously than commercially available sweet corn and the difference is easily to visually discern. Furthermore, African maize has a markedly different taste and texture, being both tougher and less-sweet, and is prepared differently, either roasted fresh on the cob, milled into flour, or sometimes boiled into porridge. The crop grown by new American farmers at NF and FS fit the visual and culinary expectations of traditional African maize; thus, both anecdotal and physical evidence support that the crop is at least initially derived from traditional African varieties, although we cannot be sure of the path it has taken in transport to gardens in New England.

Our second motivation for focusing on the maize grown by new American farmers is our interest in studying the potential for adaptation in this crop. Theoretically, tropical maize transported to a temperate environment should be maladapted and perform poorly, especially in terms of flowering time and cold tolerance. Flowering time plays a critical role in maize adaptation across different climate zones (Holland, 2018; Camus-Kulandaivelu *et al.*, 2006) because the timing of vegetative and reproductive growth phases determines how and when the plant interacts with different environmental stressors (e.g. drought, frost) and stimulants (e.g. rain, sunlight). Variation in grain yield in maize is highly sensitive to the timing of stressing periods in relation to phenological events (Millet *et al.*, 2019), and thus, flowering time is a key determinant of yield in maize, ultimately affecting when and if the plant produces a cob. In addition to flowering time, cold tolerance also plays a role in affecting grain yield, especially for maize planted in temperate regions that often experience late-season frosts. Chilling temperatures during germination and

establishment can hinder seedling development, which can have consequences for plant productivity throughout the rest of the growing season (Greaves, 1996; Farooqi and Lee, 2016).

In short, tropical maize from Africa should be poorly adapted to the temperate climates in VT and NH; however, farmers in our study are able to obtain a reasonable yield from their crop- enough to provide food for home, sell at local farmer's markets, and save seed for the next season. We suspect that as farmers save seed between seasons, some level of selection is taking place on traits that aids in temperate adaptation. Knowledge of if and how maize has adapted within this system will add to our understanding of how forced human migration affects contemporary adaptation in landraces.

*Research Questions:*

We use DNA sequencing and methods in population genetics to answer questions related to the origin of and potential contemporary evolution in maize grown by new American farmers in New England. Specifically, our questions are: (1) what is the origin of the crop and does it reflect farmer ethnicity and/or immigration history; (2a) Are the farmer's crops genetically distinct from one another and (2b) how does the diversity of the gene pool compare to other North American and African landraces; (3a) is there evidence of recent positive selection and lastly, (3b) if yes, is selection evident on genes that may confer an adaptive advantage for temperate adaptation?



## 2.2 Materials and Methods

### 2.2.1 Samples and Accessions

Primarily, this study was done in partnership with the New Farms for New Americans Initiative (NF), a nonprofit community-based agriculture program supported by the Association of Africans Living in Vermont (AALV). NF serves refugee and immigrant communities in Burlington, VT through a community garden-style program, where families rent seasonal plots ranging in size from 1/16 - 1/8<sup>th</sup> acre in which they grow a combination of traditional and commercial crops, largely for self-consumption. Because some families rent more than one plot per season, this study treats the individual or family unit who has rented garden space as the sample unit, hereafter “farmer”, and this distinction is used to delineate maize accessions per study site. To broaden our sample size, we also partnered with Fresh Start Farms, a farm incubator program located in Dunbarton, New Hampshire that also serves immigrant and refugee populations. Crop production at Fresh Start Farms is primarily market-driven rather than for self-consumption, and plot sizes are planted at a larger scale than those of NF, averaging 2 acres per farmer, of which maize makes up the majority of the planted area (personal communication with Jameson Small, FS Program Manager). In total, 9 farmers from NF and 5 farmers from FS participated in this study, all of whom were growing maize during the 2019 growing season and were chosen based on convenience sampling. Participant identification information was not collected; however, participants were able to self-report their ethnicity and/or home country, following our IRB approval (UVM IRB #STUDY000144). This information was used to inform the selection of accessions for the reference panels (see below).

Secondly, because one of the goals of this study is to determine the origin and genetic composition of maize being grown by new American farmers, we also generated two reference panels: one comprised of maize landraces from Africa and Asia, and the other comprised of maize landraces from Northeast North America. Seeds for these two reference panels were obtained from the U.S. National Plant Germplasm System (US-NPGS). Accessions for the Africa/Asia panel were selected to best reflect the reported ethnic composition of refugees and immigrants participating in the NF and FS programs; however given gaps in the GRIN collection and our partial information on farmer origin, this was only obtainable to a certain extent. For example, while the FS farmers all identify as Somali Bantu, there was no available seed of Somalian origin in the US-NPGS collection. In total, the Africa/Asia panel (AA) includes 49 accessions from 13 countries. The Northeastern North America panel (NA) was chosen to maximize the geographic distribution and includes 30 accessions from 7 US states and Quebec. (Table 1, Figure 1). Accessions information for both panels is documented in supplemental materials.



*Figure 1* Graphical representation of the distribution of New Farms and Fresh Start maize samples based on farmer-reported nationality (purple) as compared to the distribution of the Africa/Asia Landrace Panel (red). The AA landrace panel was constructed from accessions in the U.S. National Plant Germplasm System to best represent possible origins of new American farmer samples, considering both immigration history and ethnicity.

Table 1. Sample Information

<b>SAMPLE GROUP</b>	<b>LOCATION</b>	<b>FARMER ACCESSION ID</b>	<b>FARMER ETHNICITY</b>	<b>TOTAL # OF SAMPLES</b>
<b>NEW FARMS FOR NEW AMERICANS</b>	Burlington, Vermont	NF1	Burundian	6
		NF2	Congolese	7
		NF3	Congolese	7
		NF4	Bhutanese	6
		NF5	Bhutanese	7
		NF6	Bhutanese	7
		NF7	Burundian	6
		NF8	Somali Bantu	7
		NF9	Burundian	7
				<b>Total: 58</b>
<b>FRESH STARTS FARM</b>	Manchester, New Hampshire	FS1	Somali Bantu	7
		FS2	Somali Bantu	6
		FS3	Somali Bantu	7
		FS4	Somali Bantu	7
		FS5	Somali Bantu	7
				<b>Total: 34</b>

### 2.2.2 Plant Material, DNA Extraction, & Sequencing

For the new farmer samples, leaf tissue was collected *in situ* from mature maize plants growing in farmer's plots during the summer of 2019. Due to a delayed field season, tissue was collected from plants that had already produced cobs, and flowering time data was not able to be collected. Plant tissue was collected by cutting a 10 cm leaf segment from each plant, which was then stored on dry ice. 10 leaf samples were collected per farmer at both NF and FS locations resulting in 140 samples in total. For the reference panels, seeds were first germinated in petri dishes and then transferred into transplant trays.

Leaf tissue was collected at the v5 growth stage (~ 2 weeks after germination). For all samples, 100 mg of tissue was then transferred to a 1.5 mL microcentrifuge tube for DNA extraction, and the remaining tissue was frozen at -80 C for long term storage. All leaf tissue samples were then lyophilized and DNA was extracted following the E.Z.N.A. HP Plant DNA Kit protocol (Omega bio-tek). DNA quality was quantified by a NanoDrop Spectrophotometer (Thermo Fisher Scientific). Libraries were prepped at UC Davis using a modified Nextera LITE protocol, described in Rowan et al. (2019). Samples were submitted on two 96-well plates for Whole Genome Sequencing on a Novaseq S4 at IDseq.

### **2.2.3 Data Analysis**

#### **2.2.3.1 Data Cleaning & Alignment**

FastQ files were downloaded from IDseq and FastQC (<https://github.com/s-andrews/FastQC>) reports were generated for each file to inform subsequent data cleaning. Initial reports showed 2-3x coverage across most lines, with 2 samples failing. 6 additional samples were dropped due to <1x coverage. The remaining 172 samples were trimmed with Trimmomatic 0.36 (Bolger *et al.* 2014) to remove adapter contamination and overrepresented leading and trailing kmers, using the parameters ILLUMINACLIP:NexteraPE-PE.fa:2:30:10, CROP:140, and HEADCROP:5. Reads were then aligned to the B73 v4 reference genome (Jiao *et al.* 2017) with the Burrows-Wheeler Aligner (Li and Durbin, 2009) using the *bwa mem* algorithm and all aligned files were further converted to binary files and sorted in SAMtools v.1.9 (Li *et al.* 2009). Lastly, duplicates were marked and removed with Mark Duplicates in Picardtools v.1.70 (<http://broadinstitute>).

github.io/picard). Raw FastQ files are available on the Sequence Read Archive under the BioProject ID: PRJNA606801.

### **2.2.3.2 Variant Calling**

Genotype likelihoods were called with the program ANGSD (Korneliussen, Albrechtsen, and Nielsen, 2014) from the aligned BAM files. Filters were applied to assure a minimum mapping quality of 30, a minimum of 80% non-missing individuals per SNP, and a SNP p-value of  $1e6$ . The resulting Beagle file of genotype likelihoods returned 411,198,270 variant and invariant sites. To reduce computation time for downstream analyses, two randomized 100K and 200K site subsets were taken of the full genotype likelihood dataset using shuf (GNU Coreutils). Because samples were collected *in situ* from farmer's fields and the relationship between samples unknown, relatedness was then calculated with PCAngsd (<https://github.com/Rosemeis/pcangsd>) using the kinship option to check for full- and half-sibs. From this kinship matrix, only two samples had kinship coefficients  $>0.05$ , and these were NF2B and NF8B which shared a kinship coefficient of 0.26, indicating that they are full-sibs. Interestingly, these two samples come from different farmers. Only NF2B was removed from the dataset, however, because a preliminary structure analysis showed it grouping with samples from NF8 (see below for full description of structure analysis).

### **2.2.3.3 Principal Components Analysis**

A principal components analysis (PCA) was computed using PCAngsd. PCAngsd utilizes an approach based on iterative population structure inference and estimates a

covariance matrix based on genotype likelihoods. PCAngsd was chosen for its efficacy in dealing with low-depth sequencing data that varies in coverage across samples. Two PCA's were performed. The first PCA was performed on the full 100K dataset to explore how the new American farmer samples relate to the AA and NA panels. Subsequently, the second PCA was performed on a subset of the 100K dataset that included only farmer samples from NF and FS. The PCA including all samples was also re-performed with the 200K dataset to check that the results were not overly sensitive to the number of SNPs included. The covariance matrices output from PCAngsd were graphed in R with ggbiplot (<https://github.com/vqv/ggbiplot>).

#### **2.2.3.4 Admixture Analysis**

A series of structure analyses were performed using NGSadmix v.0.931 (Skotte, Korneliussen, and Albrechtsen, 2013) in order to further investigate ancestry of the new American farmer maize samples, as well as to obtain information on population substructure within the farmers' crops. NGSadmix estimates admixture proportions based on genotype likelihoods and was chosen for its compatibility with ANGSD and for its ability to work well with low coverage data.

First, NgsAdmix was run on the full 100K subset including all 171 samples at  $K=3$ ,  $K=4$ , and  $K=6$ . Because structure analyses are sensitive to relatedness between samples, we also ran the analysis for the 100K subset containing only samples from NF and FS at  $K=2$  to see if structure was correlated to farm location. Lastly, we ran a separate analysis for each farm location with  $K$  equal to the number of farmers in the sample group (NF at

K=9 and FS at K=5) to investigate if ancestry of samples was correlated by farmer and farmer-reported ethnicity.

### **2.2.3.5 Pairwise Diversity**

To investigate the diversity of farmer samples and gain insight on relative diversity as compared to the panels, we estimated genome-wide pairwise diversity ( $\pi$ ), which is a statistical measure of the degree of differences between samples per called site.  $\pi$  was calculated in Angsd using the Thetas method ([http://www.popgen.dk/angsd/index.php/Thetas,Tajima,Neutrality\\_tests](http://www.popgen.dk/angsd/index.php/Thetas,Tajima,Neutrality_tests)), that is based on an empirical Bayes method described in Korneliussen *et al.* (2013). This method first estimates the site frequency spectrum from BAM files, then calculates per-site thetas, and lastly outputs neutrality statistics, including a measure of  $\pi$  for each window. The method also requires an ancestral state input, for which we used the B73 v4 reference genome. Because this method is based on BAM files, we employed the same filters that were set for calling genotype likelihoods (minimum mapping quality of 30, a minimum of 80% non-missing individuals per SNP, and a SNP p-value of  $1e6$ ) to maintain consistency and quality.

$\pi$  was calculated per farmer (FS1-5 and NF1-9), per farm location (FS and NF), and per landrace panel (NA and AA) using 10Kb windows and 10Kb steps. The results were then filtered for windows where the number of sites called was at least 5% of sites called within the window. Finally, the results were divided by the number of sites within that window to achieve average pairwise diversity per site per window. Results were graphed using R boxplot to compare (1) between farmers and (2) between farm locations



and the panels. Extreme outliers were removed while plotting using the command *outline=FALSE* to better visualize the spread about the mean for each group.

### **2.2.3.6 Measures of Selection**

One overarching goal of this study was to determine if there is evidence that the New American farmers' crops have undergone selection as a result of translocation. To address this question, we first calculated Tajima's D to test against the null hypothesis of neutrality, and we then scanned for selective sweeps across the genome using the software RAI<sub>SD</sub> (Raised Accuracy in Sweep Detection).

Tajima's D is a neutrality test based on allele frequencies comparing the average number of pairwise differences to the number of segregating sites. Like  $\pi$ , Tajima's D can be calculated across a genome or in specific regions to determine if a region is neutrally evolving or appears to be under selection. Initially, we calculated Tajima's D across the genome, again using 10Kb windows and 10 Kb and following the same Angsd procedure as described above in the calculation of pairwise diversity. Likewise, Tajima's D was calculated per farmer (FS1-5 and NF1-9), per farm location (FS and NF), and per landrace panel (NA and AA); windows were removed that contained less than 5% of sites called.

Tajima's D is sensitive to demography and recent changes in population sizes, such as bottlenecks and migration. Thus, to get a clearer sense of how selection may be acting on the genome, we tested for selective sweeps using RAI<sub>SD</sub> (Alachiotis and Pavlidis, 2018). A selective sweep occurs when a new mutation quickly becomes fixed in a population and, due to linkage, the adjacent region surrounding the mutation experiences a significant reduction in polymorphism. RAI<sub>SD</sub> is designed to detect these regions of

reduced diversity by estimating a term they define as the *mu statistic*, a composite score that quantifies changes in the site frequency spectrum, levels of LD, and localized genetic diversity. In order to evaluate selective sweeps at the farmer and farm level, we ran RAiSD and plotted the outputs separately for two groups: first for each farmer (FS1-5 and NF1-9), and second for both farms and the African/Asian panel (NF, FS, and AA). To run RAiSD, we used default parameters and the missing data strategy 2, which creates a mask for valid alleles and treats N as a third state. Additionally, we filtered the results for alignment regions of sufficient quality using a combination of mop (<https://github.com/RILAB/mop>) and the Bedtools v.2.29.2 *intersect* command (Quinlan and Hall, 2010), for correcting the RAiSD reports to include only the quality regions identified by mop. RAiSD reports were then plotted with GenomeRanges in R using the top 0.1% *mu* statistic threshold.

### **2.2.3.7 Gene Ontology Analysis**

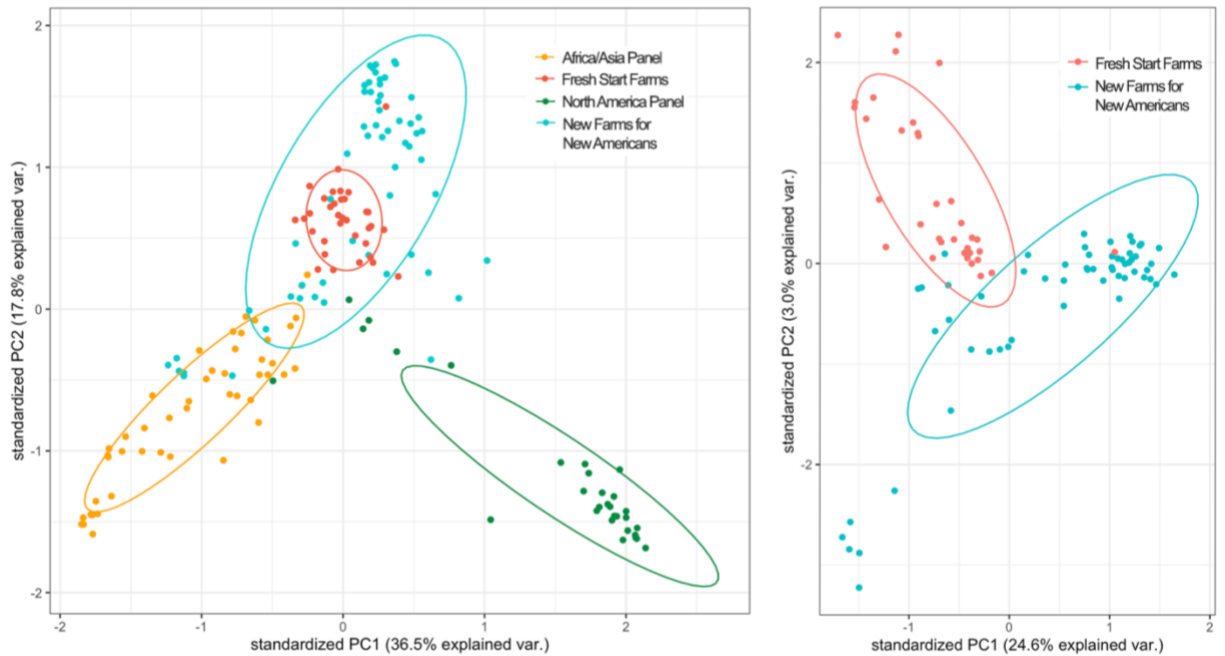
One interest of this study is to determine if selection has acted on traits that confer an adaptive advantage in the African maize grown by new American farmers in New England, such as genes related to flowering time and cold tolerance. While RAiSD identifies locations of positive selection across the genome, it does not output genes at those locations or inferred function. Because our interest is in selection that has occurred as a result of translocation, we used the RAiSD results to identify selective signatures unique to New Farms and Fresh Start as compared to sweeps that are in common with the Africa/Asia panel. It should be noted that this approach is limited to the extent that NF and FS are represented by the panel, such that sweeps that are unique to NF and FS as compared to AA are not necessarily the result of selection caused by translocation. We used Bedtools

intersect to generate three datasets from the RAI<sub>i</sub>SD reports considering windows that contain the top 0.1% of the mu statistics. These three datasets (NF-only, FS-only, and NFFS-only) were then cross-referenced to the B73 v4 annotated genome to return all gene models within these windows. Lastly, we performed a GO term analysis on the gene model lists for each dataset using the Singular Enrichment Analysis (SEA) tool in the program AgriGO v.2.0 (Tian *et al.* 2017), using the maize B73v4 (Maize-GAMER) annotated genome as the background reference.

## **2.3 Results**

### **2.3.1 Principal Components Analysis**

The principal components analysis including all sample groups (NF, FS, AA, NA) shows distinct separation and clustering by group (Figure 2). The first PC explains 36.5% of the variation and clearly separates the two panels (AA & NA) by geographic origin. Within this separation, samples from Fresh Start Farms and New Farms for New Americans group with the Africa/Asia panel; however, they are separated by the second PC (explaining 17.8% of the variance) with only minimal overlap between the farmer samples and the AA panel. With the farmer sample groups, samples from FS form a smaller cluster within the larger grouping of samples from NF, possibly reflecting differing levels of cultural diversity between farmers at the two farm locations. To further investigate the relationship between samples from the two farms, we ran another PCA including only the farmer samples. In this PCA, the first PC explained 24.6% of the variance and the second PC 3.0%. Samples generally group by farm location, but are spread obliquely across the 2 PCs.



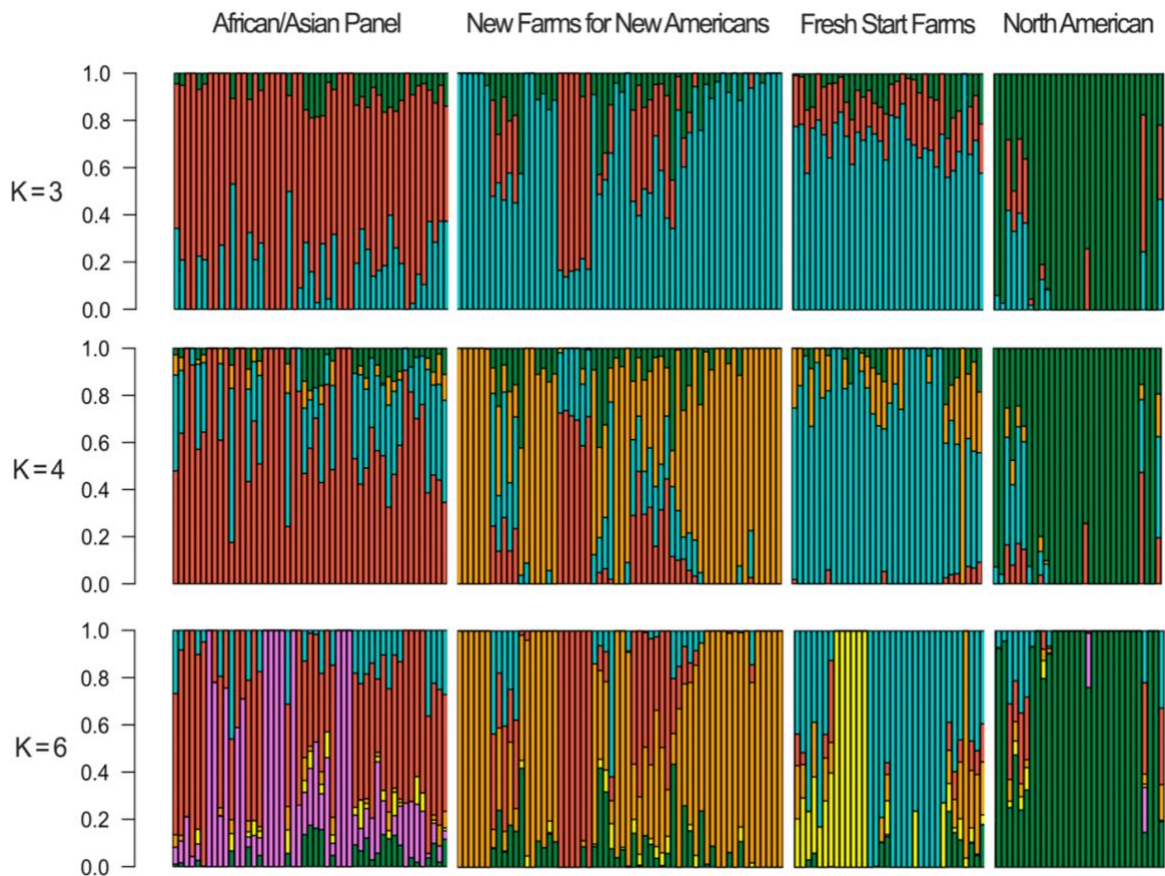
**Figure 2** Principal Components Analyses plotting the covariance matrices output from PCAngsd, showing: (Left) variation in all samples, including New Farms and Fresh Start, and both landrace panels and (Right) variation between samples from both farm locations. Plotted in ggplot with added ellipses to show major groupings by sample type.

### 2.3.2 Population Admixture and Ancestry

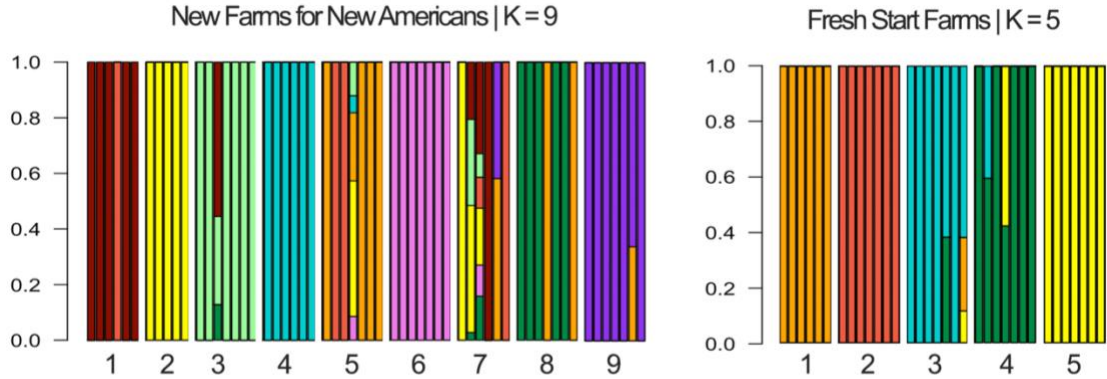
A series of structure analyses were performed with varying levels of  $K$  to investigate shared ancestry between groups (Figure 3). When all sample groups (AA, NA, NF, FS) were included in the analysis at  $K=3$ , we see that the population subdivision evenly defines the two panels as separate populations and the new American samples as a third, more heterogeneous population that shares slightly higher proportions of admixture with the AA panel than with the NA panel. When  $K$  is increased to 4, the two farm locations are separated from each other, indicated by the orange coloring; however, it is clear that NF is

much more heterogeneous than FS. Lastly, when K is increased to 6, we observe that FS and AA both are further subdivided, indicated by the pink coloring in AA and the yellow in FS, while NF and NA reflect the same structure as in K=4 and K=3, respectively.

Lastly, we investigated the relationship between maize samples collected from each farmer at the two farm locations (FS in Dunbarton, NH & NF in Burlington, VT) to obtain a clearer idea of potential seed-sharing dynamics between farmers (Figure 4). When K is set to the number of farmers sampled at each farm (5 for FS, 9 for NF), we see that individual samples are mostly subdivided by farmer. For samples from FS in particular, we see hardly any signature of admixture between farmers' crops, with the exception of 2 samples from FS4 and 2 samples from FS5 that show ~50% admixture. In samples from New Farms, we also observe a similar division by farmer, with samples from 3 farmers showing no signature of admixture and 4 farmers with admixture in only 1-2 samples. The remaining 2 farmers, NF5 and NF7, appear to be highly admixed and do not separate as any distinct group.



**Figure 3** Admixture analysis to investigate shared ancestry between New Farms and Fresh Starts samples and the Africa/Asia and North America landrace panels. Number of subpopulations was increased from  $K=3, 4, 6$ , listed next to admixture proportions on the left of the structure plots.



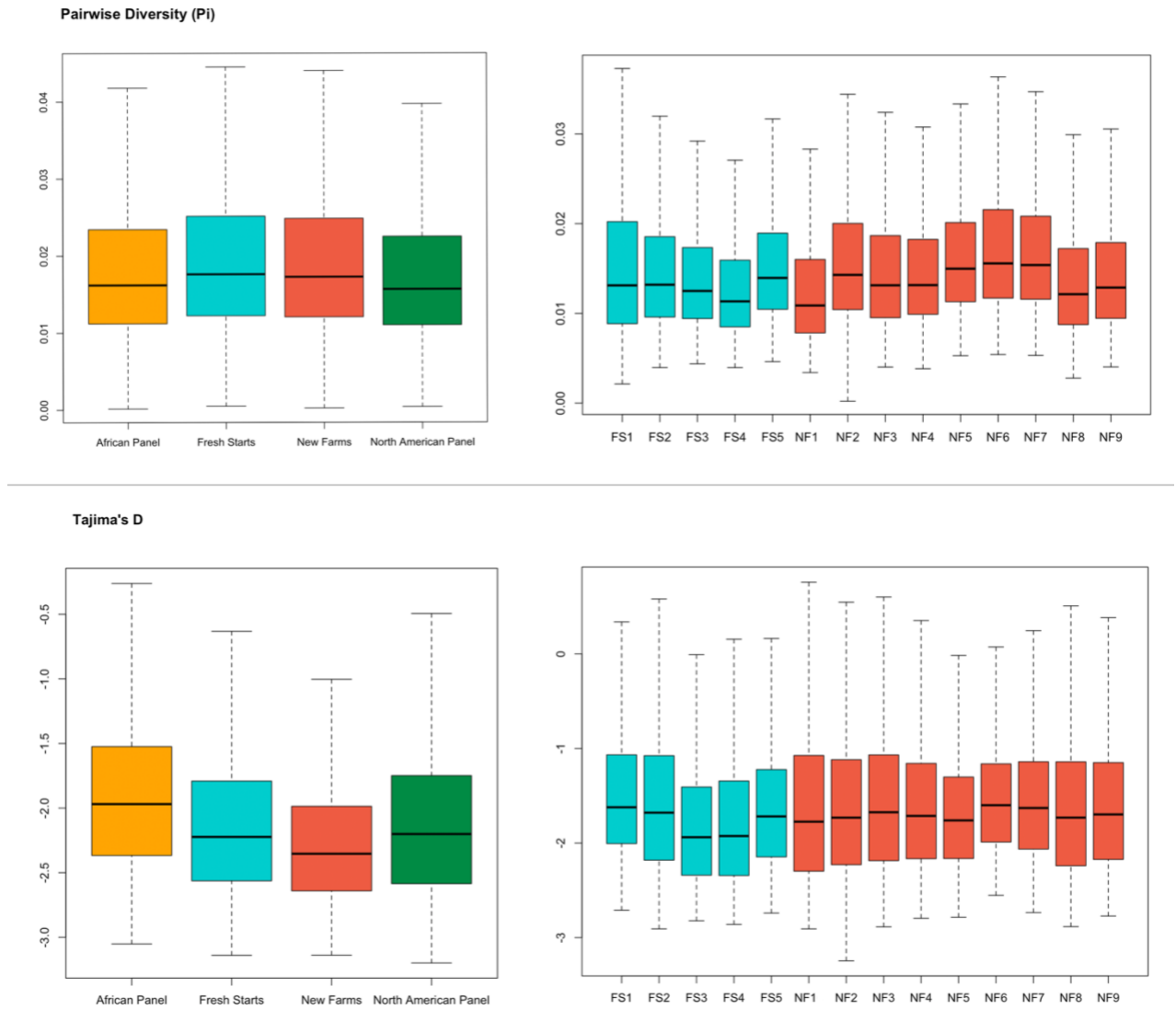
**Figure 4** Admixture analysis comparing samples from each farm (New Farms for New Americans in Burlington, VT and Fresh Start Farms in Dunbarton, NH) to investigate how distinct one farm’s crop is from another. For each farm location, number of subpopulations was set to  $K =$  number of farmers (listed above each plot) to see if samples grouped by farmer. Each subpopulation is a different color, and admixture proportions are labeled on the Y axis. Numbers on the X axis represent farmer ID’s and correspond to Table 1.

### 2.3.3 Pairwise Diversity and Tajima’s D

We calculated genome-wide pairwise diversity ( $\pi$ ) estimates to compare levels of genetic diversity between the sample groups and between farmers (Figure 5, top). When  $\pi$  was calculated for each sample group (AA, NA, NF, FS), we did not observe significant variation between the farmer samples and the landrace panels, and the mean estimate of  $\pi$  for each group is AA = 0.019, NA = 0.018, NF = 0.020, FS = 0.020. We observe more variation at the farmer-to-farmer level, but there is no significant difference in comparing between farm locations.

We also calculated Tajima’s D as a starting point to better understand if and how selection is acting on the farmer’s crops (Figure 5, bottom). Genome-wide estimates of Tajima’s D are negative for both farmer groups and for the 2 landrace panels. As with  $\pi$ ,

we see no significant difference between the average estimates per sample group (AA = -1.90, NA = -2.12, NF = -2.27, FS = -2.13) and only minor variation between individual farmers. Overall, Tajima's D is negative, consistent with recent positive selection or demographic shifts.

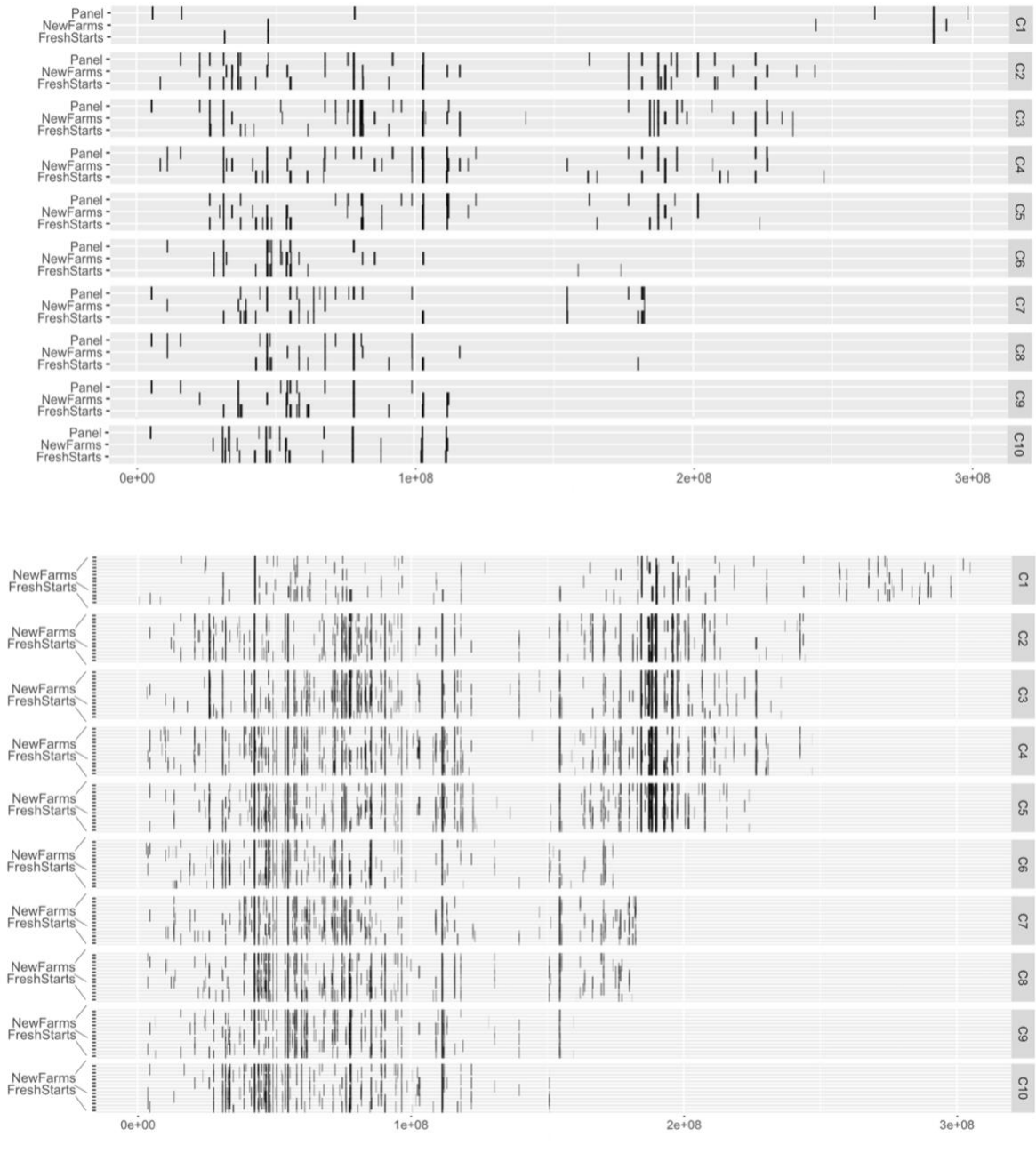


**Figure 5** Estimates of genome-wide Pairwise Diversity (top) and Tajima's D (bottom) across sample groups (left) and samples per farmer (right). All sample groups are labeled along the X-axis, and the Y-axis shows the range of values for  $\pi$  and Tajima's D, respectively. Values were achieved using the Thetas method in Angsd.



### 2.3.4 Selective Sweeps and Gene Ontology

Using the top 0.1% of mu scores output by RAiSD for each sample group, we observe a total of 154 selective sweeps in the Africa/Asia panel, 145 in the New Farms sample group, and 148 in Fresh Start. Windows containing sweeps averaged 516 Kb in length for all samples (Figure 6), which were parsed to generate three datasets: (1) for sweeps that are exclusive to NF, (2) for sweeps exclusive to FS, and (3) for sweeps exclusively shared between NF & FS. Gene models were acquired for each dataset, returning 639 models for NF only, 603 for FS only, and 205 for NF & FS only. Gene Ontology analysis for each group revealed 7 significant GO terms for NF & FS, 3 for NF only, and 0 for FS only. For the NF & FS dataset, the GO term with the highest p-value was cinnamoyl-CoA reductase activity (GO: 0016621) with  $p=4.6e-07$ , which plays a critical role in lignin production and is a precursor in anthocyanin biosynthesis. In the NF dataset, the GO term with the highest p-value was Glutathione synthase activity (GO: 0004364) with  $p=9.9e-06$ , which is a strong nonenzymatic antioxidant thought to play a significant role in cell membrane damage prevention and confer abiotic stress tolerance in plants (Hasanuzzaman *et al.* 2017).



**Figure 6** Selective sweeps output from RAI SD, plotted at 99.9%  $\mu$  significance threshold. The top plot shows sweeps analyzes for the full sample groups (AA panel, NF, and FS) and the bottom plot shows sweeps from each farmer individually, labeled as New Farms or Fresh Starts. Each chromosome is represented by the horizontal grey box, with chromosome number labeled to the right and the position along the chromosome labeled along the X-axis.

## **2.4 Discussion**

In this study, we set out to determine the origin of and potential contemporary evolution in maize grown by new American farmers in two community gardens in New England. Specifically, our questions were: (1) what are the origins of the crops and do they reflect farmer ethnicity and/or immigration history; (2a) Are the farmer's crops genetically distinct from one another and (2b) how does the diversity of the gene pool compare to other North American and African landraces; and lastly, (3a) is there evidence of recent positive selection and if yes, (3b) is selection evident on genes that may confer an adaptive advantage for temperate adaptation?

### **2.4.1 Origins of the Seed**

The initial question on which our investigation builds regards the origin of maize grown by new American farmers in the Northeast U.S. and whether or not the origin of the seed reflects farmer ethnicity, immigration history, or seed sharing networks between resettled farmers. Anecdotally, we have been told by the program managers for both farms that most farmers plant seed they have saved between seasons, having initially sourced their seed by (1) bringing it with them during immigration, (2) acquiring it from a neighbor in the farming community, or in a few cases, (3) attempting to replant seed sourced from commercial grocery markets. Based on field observations, the maize fits the vegetative and culinary expectations of African maize, as compared to commercial varieties available in the United States. By utilizing both PCA and admixture analyses, we can confirm that the samples are likely derived from African maize varieties; however, we do not see evidence to suggest an obvious correlation between farmer-reported ethnicity and seed ancestry.

Primarily, the results of the PCA (Figure 2) suggest that maize samples from New Farms and Fresh Start are more similar to the Africa/Asia panel than they are to the North America panel, and also suggest that geographic origin underlies these groupings. The first 2 PCs explain 54.3% of the total variation, and samples are clearly separated by geographic origin, where the Africa/Asia and North America panels form distinctly separate clusters. The farmer samples situate between the two panels, but cluster more closely to and overlap with the AA panel. The admixture analysis (Figure 3) produces similar results, suggesting that the new American farmer samples share greater ancestry with the Africa/Asia panel than they do with the North American panel. When we introduce higher population subdivision by increasing K from 3 to 6, however, we observe that maize samples from both NF and FS farms are fairly distinct from one another, which is also reflected in the second PCA, and that they tend to separate more from the AA panel as number of subpopulations is increased. Lastly, while both farms share ancestry with the AA panel, it appears that ancestry varies from farmer to farmer, and that not all farmer samples can be said to share a significant relationship with the AA panel.

One major limitation in answering the question of origin in this study is the dependency of the PCA and admixture analyses on the representativeness of the landrace panel in order to interpret origin. Both analyses can only interpret shared variation between the samples given, and thus the results are dependent on the extent to which the landrace panels accurately represent all possible origins of the farmer samples. The AA landrace panel was formed *a priori*, based on relatively limited information of farmer immigration history and the availability of appropriate landraces in the U.S. National Plant Germplasm System. The structure analysis in particular suggests that the AA panel is only partially

representative of the farmer samples, and thus we do not have enough evidence to draw conclusions on the exact origin of the seed based on the panels.

#### **2.4.2 On-Farm Diversity**

The second goal of this study was to quantify the genetic diversity of maize grown by new American farmers at NF and FS, measured and compared at the farm- and farmer-level. Initially, we hypothesized that genetic diversity within each farm would be relatively high, based on the assumption that the seed is derived from landraces and that farmers brought them from many different countries and provinces. We also recognized an alternative hypothesis, however, in which genetic diversity would be quite low based on the founder's effect, where farmers may have initially brought only a few seeds with them to derive their seed stock from, resulting in less variation than would be expected in a landrace crop. Based on measures of genome-wide pairwise diversity (Figure 5), we confirm that genetic diversity is relatively high for each farm as compared to estimates of  $\pi$  for both panels. While there is not a significant difference in levels of  $\pi$  across the four sample groups, we find it important to note that both farms show as much diversity as each landrace panel, which were chosen to represent the maximum diversity of landraces available for their respective regions. Estimates of  $\pi$  between farmers showed minor variation, but overall diversity was not significantly different between farmer's crops.

Recognizing that farm-level diversity is higher than expected, we wanted to investigate whether this diversity was also reflected in differences between the farmers' crops and, furthermore, if we would be able to use this information to make inferences about farmer seed-sharing dynamics. When we re-ran the admixture analysis for both farms

independently and set  $K$  equal to the number of farmers from each location ( $NF = 9$ ,  $FS = 5$ ), we see that individual samples subdivide almost exactly by farmer (Figure 4). These results indicate that in addition to measures of pairwise diversity being high for each farm, the farmer's crops are also genetically distinct from one another, suggesting that individual farmers grow and save their own seed stocks and do not engage in substantial seed exchange with other farmers in the garden.

### **2.4.3 Selection and Potential Adaptation**

Noting that the farmers' crops are most likely derived from Africa, we are presented with an interesting question: how is African maize able to produce harvests in the temperate climate of Vermont and New Hampshire? As we have described, phenology plays a critical role in adaptation of maize to different climate zones (Holland, 2018, Camus-Kulandaivelu *et al.*, 2006), and flowering time is one of the most important determinants of grain yield (Millet *et al.*, 2019) along with chilling tolerance in temperate zones (Greaves, 1996; Farooqi and Lee, 2016). Ultimately, the temperate growing season in VT and NH is substantially shorter and cooler than should be tenable for tropically-adapted maize; however, despite this, farmers in our study are still able to obtain a reasonable yield from their crop each season. Thus, we suspect that some level of recent positive selection has occurred that has contributed to temperate adaptation of the crop, and we set out to investigate these signatures of selection.

We began by estimating Tajima's  $D$  to test the null hypothesis of neutrality, and genome-wide estimates of Tajima's  $D$  were negative for all farmers, indicating a higher number of rare alleles present than would be expected in a neutrally evolving population

(Figure 5). These values could be indicative of two different processes: (1) recent positive selection and/or (2) recent population expansion after a bottleneck, which is consistent with what we know of the demographic history of the crop. Tajima's D alone cannot parse the effects of these two processes, so to infer specifically how selection is acting on the genome, we then scanned for selective sweeps using RAI<sub>SD</sub> (Figure 6) and conducted a Gene Ontology (GO) analysis on the gene models indicated under positive selection.

From the GO analysis, we discovered two significantly enriched biological processes that may indicate recent adaptation and be correlated specifically to temperate adaption. These are cinnamoyl-CoA reductase activity (common to both NF and FS samples) and glutathione synthase activity (exclusive to NF). Cinnamoyl-CoA reductase plays a critical role in lignin production (Pichon *et al.*, 1998). Because the lignin production pathway is highly complex and its products are utilized in a myriad of plant responses, it is difficult to draw any singular conclusion from this result. Relevant in this context, however, it is interesting to note that increases in lignin composition have been hypothesized as a general mechanism for adaptation to cold stress by reducing damage to cellular membranes during freezing (Moura *et al.*, 2010) and facilitating renewed growth in roots after abiotic stress (Fan *et al.*, 2006). The second significant process revealed by the GO analysis is glutathione synthase activity. Glutathione is a strong nonenzymatic antioxidant that is thought to play an important role in cell membrane damage prevention and has been shown to confer abiotic stress tolerance in many different plant species, including tolerance to salinity, drought, toxic metals, and extreme temperatures (reviewed by Hasanuzzaman *et al.*, 2017). Specifically in maize, increases in glutathione concentration have been linked to protection against chilling-induced injury in cold-

sensitive maize lines (Kocsy *et al.*, 2001), lending credence to our hypothesis that the maize grown by new American farmers has experienced some level of adaptation to New England's colder climate.

In conclusion, we found evidence of positive selection acting in numerous places across the genome and revealed significant GO-terms that indicate selection on possible cold tolerance mechanisms in maize. We did not find substantial evidence to suggest enrichment for genes controlling flowering time; however, we note that flowering time is an especially complex trait that is controlled by hundreds of small-effect QTL (Buckler *et al.*, 2009) and may require higher-depth sequencing or more samples to identify. We also note that the results of GO analyses are dependent on the current state of knowledge for any particular species (in this case, the B73v4 annotated reference), and the absence of evidence for a function does not equate to evidence for the absence of the function itself (Gaudet and Dessimoz, 2017). Ultimately, in terms of studying potential adaptation, we think it would be interesting to expand this investigation by sampling from a broader range of new America farms and gardens in New England, as well as collecting samples from the same populations over time, which would allow us to investigate how signatures of selection shift in the populations as time since original introduction increases.

## **2.5 Conclusion**

Ultimately, this research presents a case study in which forced-human migration and resettlement to a disparate environment has affected genetic change in a traditional crop. We have shown in this study that methods in population genetics can specify the degree and direction of these genetic changes, and our results provide novel insight into



how the farms and gardens that are tended by resettled refugees can support unique sources of agrobiodiversity and potentially catalyze contemporary adaptation in landraces outside of traditional farming systems. To this end, it has been previously suggested that resettled refugee and immigrant gardens serve as a hybrid form of *in situ* and *ex situ* conservation (Heraty and Ellstrand, 2016). We conclude by recognizing that refugee seed systems are valuable sources of agrobiodiversity; however, we refrain from identifying resettled refugee communities as “stewards” of this diversity. We do not believe that refugee seed systems reflect or benefit from the same standards employed by *in situ* conservation paradigms, whereby responsibility is placed on farmers for maintaining traditional crop resources. Rather, noting the potential of these systems, we encourage interdisciplinary collaboration and the establishment of participatory breeding programs to empower resettled refugees with resources to more effectively adapt their traditional crops to their new environment, ultimately supporting increased food security, food sovereignty, and the generation of new agrobiodiversity.

## COMPREHENSIVE BIBLIOGRAPHY

- Alachiotis N, and Pavlidis, P. (2018) RAiSD detects positive selection based on multiple signatures of a selective sweep and SNP vectors. *Communications Biology* 1. <https://doi.org/10.1038/s42003-018-0085-8>.
- Almekinders CJM, Louwaars NP, Debruijn GH. (1994) Local seed systems and their importance for an improved seed supply in developing countries. *Euphytica* 78(3): 207-216. Doi: 10.1007/BF00027519.

- Andrews S (2019) FastQC, v0.11.9,  
<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Belluci E, Bitocchi E, Rau D, Nanni L, Ferradini N, Giardini A, Rodriguez M, Attene G, Papa R (2013) Population structure of barley landrace populations and gene-flow with modern varieites. *Plos One* 8(12). Doi: 10.1371/journal.pone.0083891.
- Bellon MR and Berthaud J (2004) Transgenic maize and the evolution of landrace diversity in Mexico: the importance of farmer's behavior. *Plant Physiology* 134(3): 883-888. Doi: 10.1104/pp.103.038331.
- Bellon MR and Risopoulos J (2001) Small-scale farmers expand the benefits of improved maize germplasm: A case study from Chiapas, Mexico. *World Development* 29(5):799-811. Doi: 10.1016/S0305-750X(01)00013-4.
- Bitocchi E, Bellucci E, Rau D, Albertini E, Rodriguez M, et al. (2015) European flint landraces grown in situ reveal adaptive introgression from modern maize. *PLoS One* 10(4). Doi: 10.1371/journal.pone.0121381.
- Bitocchi E, Nanni L, Rossi M, Rau D, Bellucci E, Giardini A, Buonamici A, Vendramin GG, Papa R (2009) Introgression from modern hybrid varieties into landrace populations of maize (*Zea mays* spp. *Mays* L.) in central Italy. *Molecular Ecology* 18(4): 603-621. Doi: 10.1111/j.1365-294X.2008.04064.x.
- Bolger AM, Lohse M, and Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114-2120.  
<https://doi.org/10.1093/bioinformatics/btu170>.
- Bruce D (2003) Contamination, crop trials, and compatibility. *Journal of Agricultural and Environmental Ethics* 16: 595-604. <https://doi-org.ezproxy.uvm.edu/10.1023/B:JAGE.0000004961.84390.ee>
- Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, et al. (2009) The genetic architecture of maize flowering time." *Science* 325(5941): 714-718.
- Camus-Kulandaivelu L, Veyrieras JB, Madur D, Combes V, Fourmann M, Barraud S, Dubreuil P, Gouesnard B, Manicacci D, Charcosset A (2006) Maize adaptation to temperate climate: Relationship between population structure and polymorphism in the *Drawf8* gene. *Genetics* 172(4): 2449-2463. Doi: 10.1534/genetics.105.048603.
- Casañas F, Simo J, Casals J, Prohens J (2017) Toward and Evolved Concept of Landrace. *Frontiers in Plant Science* 8:145. Doi: 10.3389/fpls.2017.00145.

- Ceccarelli S (2012) Landraces: Importance and use in breeding and environmentally friendly agronomic systems. In *Agrobiodiversity Conservation: Securing the Diversity of Crop Wild Relatives and Landraces*; pp. 103-117. CAB International, Wallingford, UK.
- Dale PJ (1994) Impact of hybrids between genetically-modified crop plants and their related species: general considerations. *Molecular Ecology* 3(1): 31-36. Doi: 10.1111/j.1365-294X.1994.tb00039.x
- Dorweiler J and Doebley J (1997) Developmental analysis of teosinte glume architecture1: a key locus in the evolution of maize (Poaceae). *American Journal of Botany* 84.
- Ellstrand NC and Elam DR (1993) Population genetic consequences of small population size: Implications for plant conservation. *Annual Review of Ecology and Systematics* 24: 217-242.
- Farooqi MQU, Lee JK (2016) Cold stress evaluation among maize (*Zea mays* L.) inbred lines in different temperature conditions. *Plant Breed. Biotech.* 4(3): 352-361. Doi: 10.9787/PBB.2016.4.3.352
- Fan L, Linker R, Gepstein S, Tanimoto E, Yamamoto R, Neumann PM (2006) Progressive inhibition by water deficit of cell wall extensibility and growth along the elongation zone of maize roots is related to increased lignin metabolism and progressive stelar accumulation of wall phenolics. *Plant Physiol.* 140, 603–612.
- FAO (2008) An introduction to the basic concepts of food security. Rome.
- FAO (2017) Conflict, migration and food security: the role of agriculture and rural development. Rome.
- FAO IFAD IOM WFP (2018). The linkages between migration, agriculture, food security, and rural development. Rome. 80pp. (<http://www.fao.org/3/CA0922EN/CA0922EN.pdf>). Licence: CC BY-NC-SA 3.0 IGO
- Frary A, Nesbitt TC, Frary A, Grandillo S, van der Knaap E, Cong B, Liu J, Meller J, Elber R, Alpert KB, Tanksley SD (2000) *Science* 289(5476): 85-88.
- “Gardening brings taste of home to refugee families in Chicago” (2017, May 30), *Chicago Tribune*. Retrieved from <https://www.chicagotribune.com>
- “Gardens offer refuge to refugees” (2018, April 5). *Public Health Post*. Retrieved from <http://publichealthpost.org>

- Gaudet P and Dessimoz C (2017) Gene Ontology: Pitfalls, Biases, and Remedies. In: Dessimoz C., Škunca N. (eds) *The Gene Ontology Handbook. Methods in Molecular Biology*, vol 1446. Humana Press, New York, NY
- Graddy-Lovelace, G (2013) Regarding biocultural heritage: In situ political ecology of agricultural biodiversity in the Peruvian Andes. *Agric Hum Values* 30:587-604. Doi: 10.1007/s10460-013-9428-8.
- Gichunge C, Kidwaro F (2014) *Utamu waAfrika* (the sweet taste of Africa): The vegetable garden as part of resettled African refugees' food environment. *Nutrition and Dietetics* 71: 270-275. Doi: 10.1111/1747-008-.12143.
- Gichunge C, Somerset S, Harris N (2016) Using a household food inventory to assess the availability of traditional vegetables among resettled African refugees. *International Journal of Environmental Research and Public Health* 13(1): 1-12. Doi: 10.3390/ijerph13010137.
- Greaves JA (1996) Improving suboptimal temperature tolerance in maize - the search for variation. *Journal of Experimental Botany* 47(296): 307-323.
- Hammer K (2004) Resolving the challenge posed by agrobiodiversity and plant genetic resources - An attempt. *Journal of Agriculture and Rural Development in the Tropics and Subtropics* 76.
- Harlan J (1975) *Crops and Man*. American Society of Agronomy. Madison, WI.
- Hartwig K and Mason M (2016) Community gardens for refugee and immigrant Communities as a means of health promotion. *J Community Health* 41:1153-1159. Doi: 10.1007/s10900-016-0195-5.
- Heraty JM and Ellstrand NC (2016) Maize germplasm conservation in southern California's urban gardens: Introduced diversity beyond *ex situ* and *in situ* management. *Economic Botany* 70(1): 37-48.
- Hodgkin T, Rana R, Tuxill J, Balma D, Subedi A, Mar I, Karamura D, Valdivia R, Collado L et al. (2007) Seed systems and crop genetic diversity in agroecosystems. In: Jarvis DI, Padoch C, Cooper HD (eds) *Managing biodiversity in agricultural ecosystems*. Columbia University Press, New York, pp 77–116
- Holland JB (2018) Plant genetics: Two steps on the path to maize adaptation. *Current Biology* 28(18):R1098-R1101. Doi: 10.1016/j.cub.2018.07.049
- Hufford M, Berny Mier y Teran JC, Gepts P (2019) Crop biodiversity: An unfinished magnum opus of nature. *Annual Reviews of Plant Biology* 70: 727-51. Doi: 10.1146/annurev-arplant-042817-040240.

- Iltis HH (1974) Freezing the genetic landscape: the preservation of diversity in cultivated plants as an urgent social responsibility of plant geneticist and plant taxonomist. *Maize Genet Coop News Lett* 48: 199-200.
- Ishikawa R, Yamanaka S, Fukuta Y, Chitrakon S, Bounphanousay C, Kanyavong K, Tang LH, Nakamura I, Sato T, Sato YI (2006) Genetic Erosion from modern varieties into traditional upland rice cultivars (*Oryza sativa* L.) in Northern Thailand. *Genetic Resources and Crop Evolution* 53: 245-252.
- Jazairy I, Alamgir M, Pannuccio T (1992) *The state of world rural poverty: An inquiry into its causes and consequences*. New York University Press, New York.
- Jean M (2015) The role of farming in place-making processes of resettled refugees. *Refugee Survey Quarterly* 34: 46-69. Doi: 10.1093/rsq/hdv007.
- Jensen HR, Belqadi L, De Santis P, Sadiki M, Jarvis DI, Schoen DJ (2013) A case study of seed exchange networks and gene flow for barley (*Hordeum vulgare* subsp. *vulgare*) in Morocco. *Genetic Resources and Crop Evolution* 60: 1119-1138.
- Jiao Y, Peluso P, Shi J, *et al.* (2017) Improved maize reference genome with single-molecule technologies. *Nature* 546: 524–527. <https://doi.org/10.1038/nature22971>
- Kindscher, K 2009 (1986). *Gardening With Southeast Asian Refugees*. Community Garden Coalition, Inc., Columbia, Missouri. Updated and revised in 2009. Retrieved from <http://kindscher.ku.edu/>
- Kliman R, Sheehy B, Schultz J (2008). Genetic drift and effective population size. *Nature Education* 1(3):3.
- Kocsy G, von Ballmoos P, Rueggsegger A, Szalai G, Galiba G, Brunold C (2001) Increasing the glutathione content in a chilling-sensitive maize genotype using safeners increased protection against chilling-induced injury. *Plant Physiology* 127(3): 1147-56.
- Korneliussen, T. S., Albrechtsen, A., and Nielsen, R. (2014) ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* 15: 1-13. <https://doi.org/10.1186/s12859-014-0356-4>.
- Kwit C, Moon HS, Warwick SI, Stewart CN (2011) Transgene introgression in crop relatives: molecular evidence and mitigation strategies. *Trends in Biotechnology* 26(6): 284-293. Doi: 10.1016/j.tibtech.2011.02.003.

- Laramee, A; Waterman, B (n.d.) Global food, local food: Guide to growing, harvesting & preserving African & Asian crops in the Northeast. *UVM Extension*. Burlington, VT.
- Lakis G, Ousmane AM, Sanoussi D, Habibou A, Badamassi M, Lamy F, Jika N, Sidikou R, Adam T, Sarr A, Luxereau A, Robert T (2012) Evolutionary dynamics of cycle length in pearl millet: the role of farmer's practices and gene flow. *Genetica* 139: 1367-1380.
- Leino MW, Bostrom A, Hagenblad J (2013) Twentieth-century changes in the genetic composition of Swedish field pea metapopulations. *Heredity* 110(4): 338-346. Doi: 10.1038/hyd.2012.93.
- Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, 25: 1754-60. [PMID: 19451168]
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., et al. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25: 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Liu B, Ma G, Bussmann R, Bai K, Li J, Cao W, Long C (2019) Determining factors for the diversity of hulless barley agroecosystem in the himalaya region - A case study from Northwest Yunnan, China. *Global Ecology and Conservation* 18. Doi: 10.1016/j.gecco.2019.e00600.
- Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez J, Buckler E, Doebley J (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *PNAS* 99(9): 6080-6084. Doi: 10.1073/pnas.052125199.
- Maxted N, Dulloo ME, Ford-Lloyd BV, Frese L, Iriando JM, Pinheiro de Carvalho MMA (eds) (2012) Agrobiodiversity conservation: securing the diversity of crop wild relatives and landraces. CAB International, Wallingford.
- Meena HS, Mishra U, Gadag RN, Pathak H (2015) Cold tolerance in maize (*Zea mays*): Physiological and morphological traits. *IJTA* 33(4): 2603-2607.
- Meyer RS, Purugganan MD (2013) Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics* 14(12): 840-852. Doi: 10.1038/nrg3605.
- Millet EJ, Kruijer W, Coupel-Ledru A, Alvarez Prado S, Cabrera-Bosquet L, Lacube S, Charcosset A, Welcher C, van Eeuwijk F, Tardieu F (2019) Genomic prediction of maize yield across European environmental conditions. *Nature Genetics* 51(6): 952-956. Doi: 10.1038/s41588-019-0414-y.

- Morton JF (2007) The impact of climate change on smallholder and subsistence agriculture. *PNAS* 104(50) 19680-19855. Doi: 10.1073/pnas.0701855104.
- Moura JCMS, Bonine CAV, De Oliveira J, Viana F, Dornelas MC, Mazzafera P (2010) Abiotic and biotic stresses and changes in the lignin content and composition in plants. *JIPB* 52(4) 360-376. Doi: 10.1111/j.1744/7909.2010.00892.x
- Ootsuka K, Takahashi I, Tanaka K, Itani T, Tabuchi H, Yoshihashi T, Tonouchi A, Ishikawa R (2014) Genetic polymorphisms in Japanese fragrant landraces and novel fragrant allele domestication in northern Japan. *Breeding Science* 64: 115-124. Doi: 10.1270/jsbbs.64.115.
- Patil CL, McGown M, Nahayo PD, Hadley C (2010) Forced migration: Complexities in food and health for refugees resettled in the United States. *Napa Bulletin* 34: 141-160. Doi: 10.1111/j.1556-4797.2010.01056.x.
- Peres, S. (2016) Saving the gene pool for the future: Seed banks as archives. *Studies in History and Philosophy of Biological and Biomedical Sciences* 55: 96-104. Doi: 10.1016/j.shpsc.2015.09.002.
- Pichon M, Courbou I, Beckert M, Boudet A, Grima-Pettenati J (1997) Cloning and characterization of two maize cDNAs encoding Cinnamoyl-CoD Reductase (CCR) and differential expression of the corresponding genes. *Plant Molecular Biology* 38: 671-676.
- “Putting down roots: Refugee farms and gardens growing strong” (2019, August 24) *NCSL Blog*. Retrieved from <http://ncsl.org/blog>
- Quinlan, A. R. and Hall, I. M. (2010) Bedtools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841-842.
- Rafeipour M, Mirzaghaderi G, Shaaf S, Badakhshan H. (2016) SSR assessment of the genetic diversity of emmer wheat with emphasis on Iranian landraces (*Triticum dicoccon* Schrank). *Genetic Resources and Crop Evolution* 63: 595-600. Doi: 10.1007/s10722-016-0379-y.
- Rojas-Barrera IC, Wegier A, de Jesus Sanchez Gonzalez J, Ownes GL, Rieseberg LH, Piñero D (2019) Contemporary evolution of maize landraces and their wild relatives influenced by gene flow with modern maize varieties. *PNAS* 116(42) 21302-21311. Doi: 10.1073/pnas.1817664116.
- Rowan, B. A., Heavens, D., Feuerborn, T. R., Tock, A. J., Henderson, I. R., Weigel, D. (2019) An ultra high-density *Arabidopsis thaliana* crossover map that refines the influences of structural variation and epigenetic features. *Genetics* 213:771-787. <http://doi.org/10.1534/genetics.119.302406>.

- “Seeds of Hope: The Gardens Springing up in Refugee Camps” (2018, May 21) *BBC News: Science & Environment*. Retrieved from <http://bbc.com/news/>
- Skotte, L., Korneliussen, T. S., and Albrechtsen, A (2013) Estimating individual admixture proportions from Next Generation sequencing data. *Genetics* 195(3): 693-702. <http://doi.org/10.1534/genetics.113.154138>
- Takuno S, Oikawa E, Kitashiba H, Nishio T (2010) Assessment of genetic diversity of accessions in Brassicaceae genetic resources by frequency distribution analysis of S haplotypes. *Theoretical and Applied Genetics* 120(6): 1129-1138. Doi: 10.1007/s00122-009-1240-x.
- Teshome A, Brown ADH, Hodgkin T (2001) Diversity in landraces of cereal and legume crops. *Plant Breed Rev* 21:221-261.
- Thomas M, Dawson JC, Goldringer I, Bonneuil C (2011) Seed exchanges, a key to analyze crop diversity dynamics in farmer-led on-farm conservation. *Genet Resour Crop Evol* 58:321-338.
- Thomas M, Demeulenaere E, Dawson JC, Khan AR, Galic N, Jouanne-Pin S, Remoue C, Bonneuil C, Goldringer I. (2012) On-farm dynamic management of genetic diversity: the impact of seed diffusions and seed saving practices on a population-variety of bread wheat. *Evolutionary Applications* 5(8): 779-795. Doi: 10.1111/j.1752-4571.2012.00257.x.
- Tian, T., Yue, L., Hengyu, Y., Qi, Y., Xin, Y., Zhou, D., Wenying, X., Zhen, Su; (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res.* gkx382.
- Wang H, Nussbaum-Wagler T, Li B, Zhao Q, Vigouroux Y, Faller M, Bomblier-Yant K, Lukens L, Doebley J (2005) The origins of the naked grains of maize. *Nature* 436(7051): 714-719. Doi: 10.1038/nature03863
- Wang Y, Min J, Harris K, Khuri J, Anderson LM (2016) A systematic examination of food intake and adaptation to the food environment by refugees settled in the United States. *Advances in Nutrition* 7(6): 1066-1079. Doi: 10.3945/an.115.011452.
- Wood D and Lenné JM (1997) The conservation of agrobiodiversity on-farm: questioning the emerging paradigm. *Biodiversity Conservation* 6: 109-129.
- van Heerwaarden J, Hellin J, Visser RF, van Eeuwijk FA. (2009) Estimating maize genetic erosion in modernized smallholder agriculture. *Theoretical and Applied Genetics* 119(5): 875-888. Doi: 10.1007/s00122-009-1096-0.



- van Heerwaarden J, van Eeuwijk FA, Ross-Ibarra J (2010) Genetic diversity in a crop metapopulation. *Heredity* 104: 28-39. Doi: 10.1038/hdy.2009.110.
- van Heerwaarden J, Del Vecchio DO, Alvarez-Buylla ER, Bellon MR. (2012). New genes in traditional seed systems: Diffusion, detectability and persistence of transgenes in a maize metapopulation. *PLoS One* 7(10). Doi: 10.1371/journal.pone.0046123
- Camacho-Villa TC, Scholten M, Maxted N, Ford-Lloyd B. (2005) Defining and identifying crop landraces. *Plant Genetic Resources* 3(3): 373-384
- Xiaoyang W, Dan C, Yuqiug L, Weihua L, Xinming Y, Xiuquan L, Juan D, Lihui L (2017) Molecular characteristics of two new waxy mutations in China waxy maize. *Molecular Breeding* 37(3). Doi: 10,1007/s11032-016-0612-6.