

CONSERVATION AND CHANGE: A COMPARISON OF *IN-SITU* AND *EX-SITU*  
CONSERVATION OF JALA MAIZE GERMPLASM IN MEXICO

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Elizabeth Barnhart Rice

August 2004

© 2004 Elizabeth Barnhart Rice

CONSERVATION AND CHANGE: A COMPARISON OF *IN-SITU* AND *EX-SITU*  
CONSERVATION OF JALA MAIZE GERMPLASM IN MEXICO

Elizabeth Barnhart Rice, Ph.D.

Cornell University 2004

Conservation of agricultural genetic diversity is necessary as a source of variation for breeding and selection efforts. Traditionally, conservation efforts focus on the maintenance of diversity both in genebanks (*ex-situ*) and on farm (*in-situ*). This study takes an interdisciplinary approach to understanding *in-situ* and *ex-situ* conservation of Jala, a gigantic Mexican variety of maize (*Zea mays*).

To understand the social and policy context of Jala farmers' conservation decisions, a formal, random survey of 79 households was undertaken. In the past, 77% of farmers grew Jala. Today, 19% of farmers plant Jala on only 5% of the cultivated maize area. Jala growers are "generalists," older farmers, seed keepers with larger landholdings, diversified in many income-producing activities, with small Jala plots for household consumption and local sale; or "specialists," younger farmers, less likely to save seed, with smaller landholdings, heavily (75% of area) committed to Jala. To date, Jala has been effectively conserved in farmers' fields; conservation will likely continue as long as farmers have an economic incentive to do so.

To understand the genetic effects of *in-situ* and *ex-situ* conservation, a diversity study was performed using 22 microsatellite (SSR) markers. Populations studied included Jala from farmers' fields, genebank Jala, other maize races, and teosinte (*Zea spp.*). Farmers' Jala populations were highly diverse but not highly differentiated ( $F_{st} < 0.05$ ), indicating the unit of conservation is likely the valley, not an individual farmer. Older Jala genebank populations were less diverse and more

differentiated from recent Jala, perhaps due to genebank collection and regeneration methods. Jala's allelic profiles remained stable from 1944 to 1999. Therefore, Jala appears to have been well-conserved both *in-situ* and *ex-situ*.

By resampling the above data (1000 bootstraps), optimal sample size was evaluated. The number of individuals for accurate measurement of allele number ( $A_n$ ), gene diversity ( $H_e$ ) and population differentiation ( $F_{st}$ ) were small. Unexpectedly, population category—whether Jala from the genebank or teosinte—had little effect on optimal sample size.

Studying populations derived from advanced generations of the hybrid variety *Dekalb* 880 showed gene flow in the valley. These populations represent a reservoir of conserved traditional Jala genetic material.

## BIOGRAPHICAL SKETCH

Elizabeth “Ellie” Rice was born June 13, 1973 in Stanford, CA. She spent her childhood in the San Francisco Bay Area, with a five year detour to the suburbs of Detroit, Michigan. She attended Stanford University as an undergraduate from 1991 to 1995 as an inter-disciplinary major in Human Biology, with an emphasis on environmental design.

As an undergraduate, Ellie became intrigued by international agriculture and began her work with Mexican agriculture. Her undergraduate thesis focused on the effects of trade liberalization on nitrogen dynamics in Sonora’s Yaqui Valley. In 1996, she went to work at the International Maize and Wheat Improvement Center (CIMMYT) in Mexico City. Through work with Melinda Smale in subsistence maize systems of the Sierra de Santa Marta, Veracruz, she became interested in farmer management of genetic diversity. In 1997, she returned to Stanford to work with Roz Naylor and Wally Falcon at the Institute for International Studies.

In 1997, Ellie came to Cornell as a Ph.D. student in crop and soil science. Her interest in farmer management of diversity lead to extensive course work in genetics and plant breeding under the tutelage of Margaret Smith, Stephen Kresovich, Susan McCouch and Erick Fernanades. While at Cornell, Ellie met and married her friend Dan Ardia—perhaps the most important event of her graduate career!

For the men in my life...

--for my beloved husband Dan whose joy brightens every day...

--and for the three doctors Rice who have gone before me.

## ACKNOWLEDGMENTS

This thesis would not have been possible without the extensive support of my husband, Dan Ardia. In addition to being a patient help-mate, he has modeled and provoked clear scientific thinking, improved and encouraged my teaching skills, and stimulated my curiosity about the world around me.

I also owe thanks to my parents Tom Rice and Ann Ingham for the support, emotional and educational, that allowed me to consider the luxury of a Ph.D. I am proud to be the first of the fourth generation of Rice family Ph.D.s—my thanks to my father, grandfather and great-grandfather who are and were three gentle, clear-thinking, inquisitive role models.

Margaret Smith has been a mentor in the truest sense of the word. Her guidance, support and friendship have been the solid foundation of my Cornell experience. Steve Kresovich patiently supported my fledging experience in the laboratory and sharpened this project with his keen questions. Susan McCouch has been a constant source of new and intriguing ideas. Erick Fernandes has given a steady stream of enthusiastic encouragement.

This project has been funded and supported by many organizations. The Saltonstall fellowship generously funded my first three years at Cornell. The NSF research training grant in conservation and sustainable development (RTG-CSD) supported both coursework and fieldwork, as well as providing an interesting community of colleagues. Cornell's International Institute for Agriculture and Development (CIIFAD) generously supported several years of fieldwork. Dr. Suketoshi Taba, of the germplasm unit at the International Maize and Wheat Improvement Center (CIMMYT) in Mexico, provided both expertise and functional support with the Mexican fieldwork. Steve Kresovich and the Institute for Genomic

Diversity generously covered laboratory expenses involved in this study. The Einaudi Center at Cornell, funded a scouting trip to Mexico to search for field sites. Margaret Smith and the Corn Project, at Cornell, made field assistance possible in 1999. Finally, I have been supported, in the writing phase by the NSF GK-12 teaching fellowship, under the guidance of Marianne Krasny and Nancy Trautmann.

My transition into the laboratory was eased by Sharon Mitchell, Amanda Garris, Gemma White, Alexandra Casa, Joanne Labate, Martha Hamblin, and Julie Ho. Invaluable assistance, both in the field or the laboratory, was provided by Anna Herforth, Francisco Rivas, Israel Ruvalcalva, Christopher Miller, Xiaoyun Wang, Geronima Martinez and Charlotte Achayara. Josh Izzo rescued my computer more times than I care to count. Constancia Guzman and Catarino Hernandez graciously welcomed me into their home and their hearts. Joel Bernier provided invaluable, immediate help with the mathematics of Chapter 3.

I enjoyed the support of an excellent set of colleagues in the U.S. and Mexico. Roz Naylor and Wally Falcon opened the door to the world of agricultural research for me. Melinda Smale sparked my interest in farmer management of genetic diversity. Arahon Hernandez Guzman and Aquiles Carballo Carballo at the Colegio de Postgraduados provided insight and help both in Mexico City and Jala. Mauricio Bellon, Michael Morris and Ericka Meng, all at CIMMYT, have also provided insight.

Finally, I would like to thank my family and friends who have provided a deep and generous network of support: particularly my husband Dan Ardia, my sisters Molly and Amy Rice, my parents and step-parents, my grandparents and the Ardia family. For me, working towards this degree has been a tumultuous path, and I would not have made it without your support. In particular, Dan's love, laughter, patience, and wisdom have been a daily steadying force. Thank you to all those who have made this possible.



## TABLE OF CONTENTS

|  |           |
|--|-----------|
| <b>CHAPTER 1 INTRODUCTION.....</b>                                   | <b>1</b>  |
| <i>Overview and context: .....</i>                                   | <i>1</i>  |
| <i>Definition of terms to explain partitioning of diversity.....</i> | <i>2</i>  |
| <i>Major conservation strategies and research efforts.....</i>       | <i>2</i>  |
| <i>In-situ conservation .....</i>                                    | <i>3</i>  |
| <i>Ex-situ conservation .....</i>                                    | <i>4</i>  |
| <i>Genetic studies of conservation efforts .....</i>                 | <i>5</i>  |
| <i>Diversity studies .....</i>                                       | <i>7</i>  |
| <i>Project goals and specific objectives: .....</i>                  | <i>8</i>  |
| <b>CHAPTER 2 CONSERVATION IN A CHANGING WORLD: <i>IN-SITU</i></b>    |           |
| <b>CONSERVATION OF THE GIANT MAIZE OF JALA.....</b>                  | <b>9</b>  |
| <i>Introduction .....</i>  | <i>9</i>  |
| <i>Social and political context .....</i>                            | <i>10</i> |
| <i>Definition of terms and research questions .....</i>              | <i>11</i> |
| <i>Methods and study site.....</i>                                   | <i>12</i> |
| <i>Study site.....</i>   | <i>12</i> |
| <i>Methods .....</i>   | <i>14</i> |
| <i>History of cultivation in Jala.....</i>                           | <i>15</i> |
| <i>Historical patterns.....</i>                                      | <i>15</i> |
| <i>The CONASUPO years.....</i>                                       | <i>17</i> |
| <i>Current Jala maize varieties.....</i>                             | <i>19</i> |
| <i>Improved varieties .....</i>                                      | <i>19</i> |
| <i>Criollo varieties.....</i>  | <i>19</i> |
| <i>Traditional varieties .....</i>                                   | <i>20</i> |

|   |           |
|---|-----------|
| <b><i>Jala variety</i></b> .....  | <b>21</b> |
| <b><i>Farmer adaptations to changing circumstances</i></b> .....  | <b>25</b> |
| <i>Migration and farmer age in Jala</i> .....   | 25        |
| <i>Niche markets</i> .....  | 27        |
| <i>Consolidation and land size</i> .....  | 28        |
| <i>Fragmentation of the landscape</i> .....   | 28        |
| <i>Animal production</i> .....  | 29        |
| <i>High-value, maize-alternative crops</i> .....  | 30        |
| <b><i>Implications and discussion</i></b> .....   | <b>30</b> |
| <i>Current situation</i> .....  | 30        |
| <i>Future of on-farm conservation in Jala</i> .....   | 32        |
| <i>Is Jala a good candidate for targeted in-situ conservation efforts?</i> .....                                | 35        |
| <b>CHAPTER 3 – EXPLORING ISSUES OF OPTIMAL SAMPLE SIZE FOR</b>  |           |
| <b>DIVERSITY STUDIES: A CASE FROM MAIZE</b> .....   | <b>37</b> |
| <b><i>Introduction</i></b> .....  | <b>37</b> |
| <b><i>Materials and methods</i></b> .....   | <b>39</b> |
| <i>Populations studied</i> .....  | 39        |
| <i>Microsatellite data and DNA extraction</i> .....   | 41        |
| <i>Resampling program</i> .....   | 41        |
| <i>Mathematical analysis</i> .....  | 42        |
| <i>Calculation of actual sample values for complete data set</i> .....  | 43        |
| <b><i>Results</i></b> .....   | <b>45</b> |
| <i>Bootstrapped data for <math>F_{st}</math>, <math>H_e</math> and <math>A_n</math> for four datasets</i> ..... | 45        |
| <i>Comparison between bootstrapped and complete datasets</i> .....  | 46        |
| <i>Issues of sample size</i> .....  | 49        |
| <b><i>Discussion</i></b> .....  | <b>51</b> |

|  |                  |
|--|------------------|
| <i>Differences for the different measures: <math>F_{st}</math>, <math>H_e</math> and <math>A_n</math>.....</i>                         | <i>51</i>        |
| <i>Differences by categories of populations.....</i>   | <i>52</i>        |
| <i>Conclusion.....</i>   | <i>53</i>        |
| <b>CHAPTER 4 CONSERVATION AND CHANGE: A COMPARISON OF <i>IN-SITU</i> AND <i>EX-SITU</i> CONSERVATION OF JALA MAIZE GERMPLASM .....</b> | <b>54</b>        |
| <b><i>Introduction .....</i></b>   | <b><i>54</i></b> |
| <b><i>Materials and methods .....</i></b>  | <b><i>57</i></b> |
| <i>Plant material.....</i>   | <i>57</i>        |
| <i>DNA extraction and microsatellite markers .....</i>   | <i>58</i>        |
| <i>PCR and genotyping.....</i>   | <i>60</i>        |
| <i>Calculation of population genetic values.....</i>   | <i>61</i>        |
| <b><i>Results.....</i></b>   | <b><i>61</i></b> |
| <i>Descriptive statistics.....</i>   | <i>61</i>        |
| <i>Population differentiation.....</i>   | <i>63</i>        |
| <i>Molecular variance.....</i>   | <i>63</i>        |
| <i>Regression of allele number and gene diversity with year of collection .....</i>  | <i>65</i>        |
| <i>Allele frequency shifts with time .....</i>   | <i>65</i>        |
| <b><i>Discussion.....</i></b>  | <b><i>68</i></b> |
| <i>Jala populations compared to teosinte and maize race populations.....</i>   | <i>68</i>        |
| <i>Variation between individuals within farmers' fields .....</i>  | <i>69</i>        |
| <i>Comparison of Jala in genebank, over time.....</i>  | <i>71</i>        |
| <i>Jala in farmers' fields versus Jala in the genebank.....</i>  | <i>72</i>        |
| <i>Conservation target.....</i>  | <i>73</i>        |
| <i>Conservation implications .....</i>   | <i>75</i>        |
| <b>CHAPTER 5 GENETIC CHANGE IN “<i>CRIOLO</i>” IMPROVED VARIETY<br/>SEED RECYCLED BY FARMERS .....</b>                                 | <b>77</b>        |

|  |            |
|--|------------|
| <b><i>Introduction</i></b> .....   | <b>77</b>  |
| <b><i>Materials and methods</i></b> .....                                    | <b>80</b>  |
| <i>Genetic study: populations, markers and statistical analysis</i> .....    | 80         |
| <i>Social Survey</i> .....   | 81         |
| <b><i>Results and discussion</i></b> .....                                   | <b>81</b>  |
| <i>Results from social survey</i> .....                                      | 81         |
| <i>Results from genetic study</i> .....                                      | 83         |
| <i>Population differentiation</i> .....                                      | 84         |
| <i>Allele frequency comparisons</i> .....                                    | 86         |
| <b><i>Conclusions</i></b> .....  | <b>87</b>  |
| <i>Evidence of gene flow</i> .....   | 89         |
| <i>Implications for conservation and varietal introduction</i> .....         | 89         |
| <i>Future research</i> .....   | 89         |
| <b>CHAPTER 6 CONCLUSIONS</b> .....   | <b>92</b>  |
| <i>Social context for conservation in the field</i> .....                    | 93         |
| <i>Distribution of diversity between species and among maize races</i> ..... | 95         |
| <i>Diversity in farmers' fields</i> .....                                    | 97         |
| <i>In-situ and ex-situ conservation of Jala maize</i> .....                  | 98         |
| <i>Implications for conservation</i> .....                                   | 99         |
| <i>Questions for further research</i> .....                                  | 100        |
| <b>APPENDIX I SURVEY FOR JALA STUDY 2001</b> .....                           | <b>103</b> |
| <b>APPENDIX II SAMPLE SIZE FOR DIVERSITY STUDIES</b> .....                   | <b>106</b> |
| <b>APPENDIX III DOES SAMPLING METHOD MATTER?</b> .....                       | <b>108</b> |
| <b>APPENDIX IV POPULATIONS AVAILABLE FOR STUDY</b> .....                     | <b>111</b> |
| <b>REFERENCES</b> .....  | <b>112</b> |

## LIST OF FIGURES

|   |     |
|---|-----|
| Figure 2.1 Location of Jala, Nayarit .....  | 13  |
| Figure 2.2 Temperature, rainfall and key activities in the maize planting cycle in Jala,<br>Nayarit .....   | 13  |
| Figure 2.3 Reconstructed maize areas planted in Jala from 1925 to 2001 based on 36<br>informal interviews in 1999, and a formal survey of 79 households in 2001 ..... | 16  |
| Figure 2.4 Age of surveyed farmers versus age of population in Jala, Mexico, 2001..   | 26  |
| Figure 2.5 Ages of farm heads of households in Jala (2001 survey) compared to<br>average for Mexican ejidatario heads of household (de Janvry et al. 1997) .....      | 27  |
| Figure 3.1 Values and standard deviations for $A_n$ , $H_e$ and $F_{st}$ for the four datasets.....   | 47  |
| Figure 3.2 Normalized data for $F_{st}$ , $H_e$ and $A_n$ for the four datasets .....   | 48  |
| Figure 4.1 Pair-wise $F_{st}$ values for Jala populations from the genebank and from<br>farmers' fields, with history of collection and regeneration .....            | 64  |
| Figure 4.2 Allele frequency profiles over time for 2 randomly selected loci: phi011<br>and phi033 .....   | 67  |
| Figure 4.3 Evaluation of the target for conservation: common and rare allele<br>frequencies compared to widespread and local distribution for Jala data .....         | 74  |
| Figure 5.1 Comparison of allele frequencies for D880 and Jala populations at 6<br>randomly selected loci .....  | 88  |
| Figure III.1 Comparison of $N_e$ sizes by sampling method .....   | 110 |

## LIST OF TABLES

|   |    |
|---|----|
| Table 2.1 Distribution and area of different classes of maize in Jala 2001 .....  | 18 |
| Table 2.2 Subjective estimates of maize yields for different varieties and categories of<br>maize grown in Jala 2001, for good, normal and bad years .....  | 20 |
| Table 2.3 Distribution, size and longevity of Jala variety plantings in Jala, Mexico,<br>2001 .....   | 22 |
| Table 2.4 Reasons surveyed farmers stopped growing Jala variety .....   | 22 |
| Table 2.5 Profiles of Jala growers in Jala, Mexico 2001 .....   | 23 |
| Table 2.6 Prices for Jala variety sold as <i>elote</i> (corn-on-the-cob), grain and seed,<br>compared to average maize grain price, 2001 .....  | 25 |
| Table 3.1 Farmers' field Jala, genebank Jala, maize races and teosinte populations<br>studied, with their seed source and catalog numbers .....   | 40 |
| Table 3.2 SSR markers used to evaluate Jala, maize race and teosinte populations,<br>with associated bin location, core repeats, and allele size ranges, PIC values,<br>numbers of alleles and missing data ..... | 42 |
| Table 3.3 Fitted equations for $F_{st}$ , $H_e$ , $A_n$ data, based on 1000 bootstrapped samples of<br>33 populations, with $R^2$ values to reflect quality of fit .....  | 44 |
| Table 3.4 Comparison of $F_{st}$ , $H_e$ , and $A_n$ values between bootstrapped and complete<br>datasets for four categories of populations .....  | 49 |
| Table 3.5 Sample size thresholds (numbers of individuals) for $F_{st}$ , $H_e$ and $A_n$<br>calculations for the four population categories studied.....  | 50 |
| Table 4.1 Farmers' field Jala, genebank Jala, maize races and teosinte populations<br>studied, with seed source and catalog numbers .....   | 59 |

|   |     |
|---|-----|
| Table 4.2 SSR markers used to evaluate Jala, maize race and teosinte populations,<br>with associated bin location, core repeats, and allele size ranges, PIC values,<br>numbers of alleles and missing data. .... | 60  |
| Table 4.3 Descriptive information for genotyped populations, organized by category,<br>with category averages and totals, averaged across 22 SSR loci .....   | 62  |
| Table 4.4 AMOVA for Jala from genebank and farmers' fields, with average and range<br>over 22 SSR loci .....  | 65  |
| Table 4.5 Regression of allele number ( $A_n$ ) and gene diversity ( $H_e$ ), by year, for Jala<br>populations from the genebank (1944 to 1988), for 22 SSR loci .....  | 66  |
| Table 5.1 Subjective response maize yields for different varieties and categories of<br>maize in good, normal and bad production years .....  | 82  |
| Table 5.2 Farmer descriptions of phenotypic changes with <i>criollo</i> varieties .....   | 84  |
| Table 5.3 Descriptive information for genotyped populations across 22 SSR loci .....  | 84  |
| Table 5.4 Pairwise $F_{st}$ values for Jala from genebank and farmers' fields, 22 loci .....  | 85  |
| Table II.1 Sample sizes for recent diversity studies using SSR markers .....  | 106 |
| Table III.1 Population and collection method .....  | 108 |
| Table IV.1 Advanced generation D880 populations available for study .....   | 111 |

## CHAPTER 1 INTRODUCTION

### ***Overview and context:***

Conservation of genetic resources, together with plant breeding, has been critical to the agricultural gains that have fed our burgeoning world population (Hoisington et al. 1999). Genetic diversity is necessary as a source of variation for plant breeding and selection. Because future trait requirements, whether sources of resistance to pathogens or quality enhancing traits, are unknown at the time of collection, conservation efforts strive to preserve useful genes and gene complexes by protecting broad samples of genetic diversity. Long ago, people suggested we look in centers of diversity to find useful agricultural variation (Vavilov 1926; 1997).

Conservation involves a series of complex social, political, biological and genetic issues. Therefore, studies have recently begun to take a more integrated, interdisciplinary approach to conservation. [See Bellon et al (2003), together with Pressoir and Berthaud (2004b) for an integrated approach to maize (*Zea mays* L.) conservation in Oaxaca, Mexico.] To implement effective conservation strategies, it is critical to understand the social, economic and political environment that influences the conservation decisions of farmers and genebank managers. It is also important to evaluate the genetic impacts of different conservation approaches. Finally, we should think beyond the current approaches to conservation and look for alternate reservoirs of genetic information embodied in traditional varieties.

This thesis is a detailed case study of the giant maize of Jala, Nayarit, Mexico. First, the social, economic and policy context influencing Jala conservation in the field are discussed and analyzed. Next, the genetic effects of Jala conservation, both in the field and in the genebank are examined. Finally, advanced-generation improved materials that farmers manage as “*criollos*” are evaluated as a potential reservoir of Jala’s traditional genetic material.



### *Definition of terms to explain partitioning of diversity*

Many studies use different terms relating to the structure of diversity in slightly different ways. In this study, “seed lot” described the batch of seed a farmer managed and planted (Louette et al. 1997). A “variety” was a composite of seed lots with the same name, and a common, perceived set of characteristics. A “traditional” farmer’s variety was more variable and less tightly constrained than a commercial “improved” variety, but both were distinguished by a unique set of characteristics. Many studies have used the term “landrace” to describe a unit similar to the “traditional varieties” used here. [See Zeven (1998) for a review.] The term “race” was used to group varieties according to physical, usually ear-related characteristics (Anderson and Cutler 1942). The term “wild relative” was used interchangeably with “teosinte” (*Zea spp.*) to describe related wild populations.

### ***Major conservation strategies and research efforts***

Currently there are two important approaches to conservation of agricultural genetic diversity; both have a storage and a management component (Wood and Lenne 1997). *Ex-situ*, or off-site conservation, involves the collection of resources in the field and their subsequent preservation, usually in a cold-storage genebank facility. Critics often focus on the storage phase of genebank conservation, where few changes occur, and misperceive *ex-situ* conservation as a static process (Wood and Lenne 1997). Indeed, few genetic changes occur in cold storage (Roberts 1975). However, the management part of genebank conservation can be the source of genetic bottlenecks at the time of collection or during regeneration. In contrast, *in-situ* conservation protects varieties in their natural environment or on the farm, under

farmer management. In this paper, the terms on-farm and *in-situ* conservation will be used interchangeably, as will genebank and *ex-situ* conservation. Maxted et al. (2002) emphasize that *in-situ* conservation can be used more broadly, as it is in the Convention on Biodiversity, to include protected areas for the preservation of wild relatives of crop species, but here it is used interchangeably with “on-farm” to describe farmer-managed conservation of agricultural diversity.

The *in-situ* and *ex-situ* conservation approaches are complementary (Maxted et al. 1997) and may be best targeted to different types of varieties. *In-situ* approaches are best focused on those varieties that have high value to farmers as well as high genetic diversity value (Smale and Bellon 1999). Farmers cannot and should not be expected to conserve varieties for which they must forfeit productivity or the potential to adopt new varieties. Nor should farmers be expected to enter situations that work against their own economic best interest (Smale et al. 2004). *In-situ* approaches should focus on areas where farmers are least likely to substitute traditional varieties for improved ones, and where conserving adapted, local varieties has present and future economic value, likely due to socio-economic, ecological or development reasons (Smale et al. 2004). Varieties with low economic value to farmers, on the other hand, are good targets for *ex-situ* collection and conservation. On-farm conservation is becoming much more targeted with recent work to establish frameworks for designing and implementing *in-situ* conservation projects—a more intentional strategy than the preceding *de facto* on-farm conservation efforts (Maxted et al. 2002; Bellon et al. 2003).

### *In-situ conservation*

Discussions about *in-situ* conservation tend to focus on the farmer management aspect as natural, dynamic and evolving with farmer needs (Cleveland et

al. 1994). Often these discussions neglect the risk that the storage aspect of on-farm conservation is potentially subject to accidental loss of diversity due to environmental, political or socio-economic considerations (Wood and Lenne 1997).

Field management of varieties obviously has genetic importance: planted areas determine the size of genetic populations, neighboring crops are potential sources of gene flow, and the timing of agronomic activities influences the plants' receptivity to genetic exchange. Farmers' fields have capacity to support more rare alleles and different genotypes than genebanks (Brown 2000). Equally critical to good on-farm conservation is a clear understanding of the process of selecting seed. Selection is often performed by both men and women, at harvest and during grain processing. Farmers often have an ideal varietal image, or ideotype, towards which they select (Rice et al. 1997; Cleveland et al. 2000; Louette and Smale 2000; Soleri et al. 2000). Thus, there are a variety of ways in which both the management and storage phases of on-farm conservation can have genetic impacts.

#### *Ex-situ conservation*

Genebank conservation is often undertaken by public actors—governments and international non-governmental organizations—for the public good. Recent efforts to establish an endowment for international genetic resource protection have gone so far as to put a price tag on each type of accession in the international genebank system. Regardless of the precise costs, which are sensitive to geographic location and costs of labor and energy, maize conservation costs more than other seed crops because its regenerations require hand pollinations and its larger grain size requires more space in a genebank. Koo et al. (2003) give an annual cost of US\$2.16 per accession for maize, compared with other seed based crops at US\$1.50. A different estimate places maize storage costs at US\$0.93 per maize accession (Pardey

et al. 2001). Furthermore, because maize is an out-crossing species, accession regeneration requires labor-intensive hand-pollinations to ensure genetic integrity, estimated by Pardey et al.(2001) to cost US\$103 per accession over time. Wild relatives, due to difficulties in germination, collection and regeneration, are most costly to maintain.

Though the storage phase of genebank conservation is considered very stable, there are chances for genetic change during collection and regeneration of seed. Today, theoretical models based on population genetics establish clear guidelines about how to collect diversity (Crossa 1989). A genebank accession can never be more diverse than the genetic material originally collected in the field. Similar models (Crossa et al. 1993; Crossa et al. 1994) also guide regeneration methods for maintenance of maximum diversity. In the early years of *ex-situ* conservation, the emphasis was on collecting and categorizing as much phenotypically diverse material as possible—conserving varieties, without as strong a focus on conserving the genetic diversity within populations. By the 1970s, collection guidelines began to explicitly use population genetic theory to obtain maximum genetic diversity for the effort expended (Frankel and Hawkes 1975). Rare alleles were considered too difficult to capture, and widespread alleles would be captured as a conservation by-product. Therefore collection efforts began to target pockets of locally common alleles (Marshall and Brown 1975).

#### *Genetic studies of conservation efforts*

As molecular genetic tools have grown easier and less expensive to use, they have been applied to an ever-widening circle of conservation questions. Most efforts have focused on either genebank or on-farm conservation. Few studies have tried to directly compare one method to the other.

One important use of molecular markers has been to characterize genebank populations to understand distribution of diversity, particularly between races and agronomic groups. [For example see Rodriguez et al. (1999) for capsicum (*Capsicum spp.*), Bhattacharjee et al. (2002) for pearl millet (*Pennisetum glaucum* L.), and Ni et al. (2002) for rice (*Oryza sativa* L.)]. Other molecular studies showed evidence of duplication and redundancy in collections. [See Virk et al. (1995) for rice, Phippen et al. (1997) for cabbage (*Brassica oleracea* L. var. *capitata*), Zeven et al. (1998) for kale (*Brassica oleracea* L. var. *ramosa*), Dean et al. (1999) for sorghum (*Sorghum bicolor* L. *moench*), Verma et al. (1999) for basmati rice, and McGregor et al. (2002) for wild potato (*Solanum acuale* ssp.).] Several of these in-depth genetic studies turned up evidence of mix-ups and misclassifications within the genebank system [Steiner et al. (1997) for oat collections (*Avena sativa* ssp.), Dean et al. (1999) for sorghum, and McGregor et al. (2002) for wild potato].

One important, but little understood, aspect of on-farm conservation is the partitioning of genetic diversity within farmers' fields, between varieties, and between communities (Wood and Lenne 1997). Recently, several studies have used molecular markers of various types to look at farmer-managed diversity at the field level. All of these studies have found high, often surprisingly high, levels of genetic diversity in farmers' fields, regardless of breeding system. Inbreeding sorghum populations have lower than expected differentiation between farmers' fields, and higher diversity within populations (Dje et al. 1999). From this, Dje et al. (1999) conclude that sorghum genebank collection strategies need to sample more individuals in farmers' fields to better assay genetic diversity. Cassava (*Manihot esculenta* C.), a species propagated by cuttings, has surprisingly high levels of genetic diversity within named varieties (Elias et al. 2000). Phenotypically distinct, differently named varieties can not be distinguished at the genetic level in either out-crossing pearl millet populations

in Nigeria (Busso et al. 2000) or out-crossing maize populations in Oaxaca (Pressoir and Berthaud 2004a; 2004b). In other words, farmers' variety names do not reflect a discreet genetic unit. Geography may (Busso et al. 2000) or may not (Pressoir and Berthaud 2004b) lead to genetic differentiation between varieties of outcrossing species from different villages in a region.

In spite of all the attention focused on the genetic effects of both *in-situ* and *ex-situ* conservation, few studies have directly compared the approaches of one to another. In one of two exceptions, Soleri and Smith (1995) found phenotypic differences between Hopi maize varieties conserved on-farm and in a genebank. Their analysis, however, was confined to morphological characteristics and did not consider molecular genetic data. In the other exception, Parzies et al. (2000) looked at historical collections of barley (*Hordeum vulgare*), an inbred crop, and compared them to the same variety in farmers' fields. There was a clear decline of *ex-situ* genetic diversity with storage time and regeneration. Older genebank populations became more genetically distinct from one another; an effect the authors attributed to genetic drift during genebank management.

### *Diversity studies*

Beginning with the maize race classifications of Anderson and Cutler (1942) and refined by Wellhausen et al. (1952), scientists have struggled to organize the impressive physical array of maize diversity. The early classification schemes were predominantly based on a series of ear characteristics. Subsequent studies looked at plant morphological characteristics independently or with genetic tools. Isozymes, genetically-based variants of enzymes, were the first tool used (Doebley et al. 1985; Doebley et al. 1986; Doebley et al. 1988). These studies looked at a wide range of maize diversity—races from Mexico, South America, the United States, the

Caribbean—and helped to explain the ancestral and evolutionary relationships of one to another. Later studies used RFLPs, AFLPs and most recently microsatellites (SSRs) to target or reexamine increasingly specific subsets of maize diversity—Mexican, Caribbean, and South American races, European open pollinated varieties and US historical landraces (Dubreuil and Charcosset 1999; Rebourg et al. 2001; Matsuoka et al. 2002; Labate et al. 2003). All of these studies focus at or above the level of the maize race. A few, like Labate et al. (2003), begin to look at depth and variation within a maize race (for example the Northern Flints of the United States) or the distribution of that race in farmers’ fields (Pressoir and Berthaud 2004a; 2004b).

***Project goals and specific objectives:***

In this thesis, a detailed case study of conservation of Mexican maize is developed, specifically for the giant maize of Jala. The following objectives were addressed in four studies:

- 1) Describe the social and policy context in which Jala farmers make their on-farm conservation decisions using social survey techniques.
- 2) Investigate issues impacting sample size for diversity measures using empirical data about different types of populations.
- 3) Determine how diversity is partitioned within and between farmers’ fields and between genebank populations of the variety Jala to evaluate the success and field dynamics of Jala conservation.
- 4) Examine farmer-managed “*criollos*” grown in Jala as a potential reservoir of Jala’s traditional genetic material.

## CHAPTER 2 CONSERVATION IN A CHANGING WORLD: *IN-SITU* CONSERVATION OF THE GIANT MAIZE OF JALA.

### *Introduction*

Conservation of agricultural genetic resources provides the genetic building blocks for plant breeders to use in constructing better plant varieties. New plant varieties are needed to meet the demands of a growing world population and ever-changing disease, insect and environmental stresses. Conservation of genes, gene complexes, physical traits and varieties has value not only to plant breeders, but to farmers in developed and developing nations. While there is a social value to conserving unique genes, gene complexes and varieties, decisions about what to grow, and therefore conserve, are often made by economic criteria (Bellon et al. 2003).

There are two important strategies for conserving agricultural diversity. *Ex-situ* conservation, also referred to as genebank conservation, collects diversity in the field and preserves it in a genebank, usually a cold-storage facility. *In-situ*, or on-farm, conservation conserves agricultural diversity in the field, under farmer management. The two strategies are viewed as complementary (Maxted et al. 1997) and may be best targeted to different varieties and environments. *In-situ* approaches are best focused on varieties with high economic value to farmers as well as high genetic diversity (Smale and Bellon 1999). Areas where farmers are least likely to substitute traditional varieties for improved ones, and where conserving adapted local varieties has present and future economic value are the best candidates for *in-situ* conservation efforts. Varieties with low economic value to farmers make good targets for *ex-situ*, or genebank, conservation approaches (Smale et al. 2004).

Long ago, investigations suggested we look in centers of diversity to find useful agricultural variation (Vavilov 1926; 1997). Mexico is considered to be the center of origin for maize and has high levels of maize diversity (Doebley 1990;



Matsuoka et al. 2002). Therefore, Mexico has long been the focus of conservation efforts for maize [e.g., Anderson and Cutler (1942); Wellhausen et al. (1952)]. This paper focuses on a case study of the Jala race of maize from the town of Jala, state of Nayarit, Mexico.

### *Social and political context*

In Mexico, recent decades have seen many important political and economic changes that will have an impact on conservation of maize genetic diversity. Since the Mexican revolution in 1917, Mexican agriculture has been characterized by large amounts of land in collective land holdings called *ejidos*. These small, relatively inefficient landholdings were in turn supported by heavy government subsidies of credit, inputs and price supports. In recent years, with reform to Article 17 of the Mexican Constitution, *ejido* lands have been privatized and subsidies removed (de Janvry et al. 1997).

The net impact of all the liberalizing policies led many to predict consolidation of land because of tenure changes (e.g., Collier 1994; Stanford 1994) and substantial changes in cropping patterns, shifting away maize and towards high value crops like vegetables. Early studies of the *ejido* sector changes (Murphy 1994; Cornelius and Myhre 1998) did not see the anticipated consolidation of lands with farmers freer to rent and sell land. Recently, in the highly developed agriculture of the Yaqui Valley, Sonora—often used as a bellwether for agricultural change (Naylor et al. 2001)—Lewis (2002) found that though few *ejidatarios* had officially privatized their land, many unofficial ‘sales’ and a high rate of rental were part of a move towards consolidation. All of these changes have implications for land size, crop choice and how farmers grow maize. Therefore, they also have implications for traditional varieties, and Jala in particular.

*Definition of terms and research questions*

The term “seed lot” is used to describe the unit of seed a farmer plants and manages (Louette et al. 1997). A “variety” is composed of seed lots with the same name, distinguished by a common, perceived set of characteristics. A “traditional” farmer’s variety is more variable and less tightly constrained than a commercial “improved” variety. Both are distinguished by a unique set of characteristics. A “*criollo*” variety is initially a commercially-produced, uniform, improved variety that farmers recycled and maintained as a traditional variety, keeping seed from one cycle to the next. The term “race” of maize (Anderson and Cutler 1942) is used to group varieties according to physical, usually ear-related, characteristics. Jala is unusual in that it is both a farmer’s traditional variety and a race of maize—so unique are its characteristics. Reference is also made to “production” characteristics of maize, meaning traits relating to the growth and management of a variety (e.g., tolerates sandy soils, emerges from depth). “Consumption” characteristics refer to traits like taste, texture and color which influence food preferences.

Conservation, especially on-farm conservation, takes place in a social context. The decisions farmers make when they choose varieties and select seed for the next cycle have critical genetic effects, especially for conservation of a unique variety like Jala. This paper focused on survey data collected in Jala by interviewing farmers between 1999 and 2001. Questions focused on the present and past patterns of use for the Jala variety. Patterns of utilization were compared for different groups of farmers to understand how Jala use has changed and may change in the future. Finally, social and policy influences on farmer conservation decisions and their economic consequences are examined.

## ***Methods and study site***

### *Study site*

Choice of a study population of maize presented a challenge. Many populations in genebanks are simply labeled “*blanco criollo*,” local white, and thus it is impossible to know whether today’s “local white” maize is the same variety as the “local white” of 50 years ago. Therefore, this study sought a maize variety that was identifiable in both the past and present for a set of unique characteristics.

The town of Jala, in the state of Nayarit (see Figure 2.1) has long been known for its giant maize. Its August *Feria de Elote* (Corn-on-the-cob Festival) is well-attended and well-known for its giant ears of corn-on-the-cob. As early as 1924, reports appear in the literature documenting the variety’s unusual height, up to 6 m, and surprisingly large ear size, up to 50 or 60 cm (Kempton 1924). Farmers tell stories of harvesting on horseback to reach the tall ears, and carrying the ears back, bundled like logs on the back of *burros*. Today plant heights are about 5 m, and ear lengths measure up to 35 or 40 cm of grain.

Today, Jala is a town of about 9,000 inhabitants<sup>1</sup> (Jala 1996). The town is located at the foot of the active Ceboruco Volcano, approximately 1,100 meters above sea level, in the Sierra Madre mountain range that divides Mexico’s interior from its Pacific Coast. The area has one primary growing cycle, in the wet season from June to September, with substantial, regular rainfall [737mm of the annual 860mm (86%) fall between June and September] and temperatures average around 24 C, as shown in Figure 2.2 (Jala 1996). Traditional and improved maize varieties are planted at different times of the year and have overlapping flowering periods, which could lead to gene flow (Figure 2.2). The dry season is too dry for most crops.

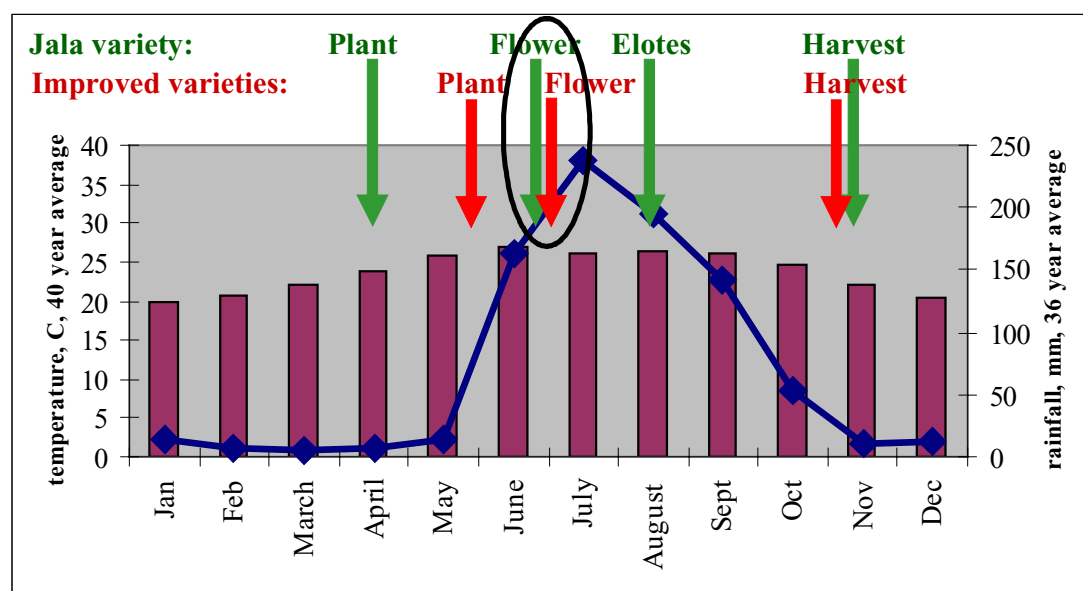
---

<sup>1</sup> The towns of Jala, Jomulco and Coapan are considered together and have 9,793 inhabitants. This study covers only Jala and Jomulco.

Farmers in Jala described themselves as “*maiceros*” or maize growers, and maize was the dominant crop in the valley, comprising 77% of the valley’s planted area. Increasingly, farmers were planting tobacco (*Nicotiana tabacum*) as a cash crop. Other alternative crops included small areas of sugar cane (*Saccharum officinarum*), sorghum, pasture and peanut (*Arachis hypogaea*).



**Figure 2.1 Location of Jala, Nayarit**



**Figure 2.2 Temperature, rainfall and key activities in the maize planting cycle in Jala, Nayarit**

### *Methods*

To better understand the cultivated Jala variety in its unique context, this study began with a series of informal surveys in the towns of Jala and Jomulco—home to the producers of the Jala valley. Early, informal interviews included 30 farmers, most of whom planted the Jala variety, known locally as ‘*maiz de húmedo*’<sup>2</sup>. In 2001, a formal, random survey was undertaken. Lists of farmers from the *ejido* and *pequeño propietario* associations for both Jala and Jomulco were revised with the help of key informants to remove repetitions and names of people who were no longer farming or who were otherwise unavailable. In total, there were 444 farmers in the valley, of whom 79 were randomly selected for interviews (17.8%)<sup>3</sup>. Questions for the formal survey were a mixture of open-ended and objective questions. A copy of the questionnaire is available in Appendix I.

Analysis of the survey data divides the Jala data by the size of the farmer landholding. Several studies have shown differences in patterns of cultivation and economic strategies that vary with size of landholdings in Mexico (de Janvry et al. 1997; de Janvry and Sadoulet 2001; Lewis 2002). Furthermore, with the predicted consolidation of land in Mexico, understanding the interaction between landholding size, patterns of cultivation and conservation of traditional varieties on-farm becomes increasingly important. The landholdings of farmers in Jala, as in much of Mexico, are small: an average of 5.9 hectares (range from 0.25 to 20 hectares). The Jala land

---

<sup>2</sup> The term “*maiz de húmedo*” stems from the way the Jala landrace is traditionally planted—early, in April, in the residual soil moisture (*de humedad*). Other types of maize can be planted “*de humedad*,” but there is only one “*maiz de húmedo*.”

<sup>3</sup> In total, 110 farmers were randomly selected. Of these, 14 were no longer planting and 17 could not be reached. On the original lists, there were 171 farmers from Jala (43.8%) and 219 from Jomulco (56.2%). In the survey, 39 (49.4%) are from Jala and 40 (50.6%) are from Jomulco. The difference in frequencies can be attributed, in part, to the fact that more people in Jomulco are no longer planting (2 from Jala and 12 from Jomulco.)

sizes are similar to the distribution found by de Janvry et al. (1997) for the *ejido* sector over all of Mexico. Therefore, the same categories for analysis are used, with larger land holders defined as those who have at least ten hectares and small landholders defined as those who have five hectares or less.

### ***History of cultivation in Jala***

#### *Historical patterns*

Many years ago, there were fewer types of maize grown in Jala. The majority of the maize planted in Jala was the variety Jala<sup>4</sup> (Figure 2.3). The interviewed farmers believed the Jala variety had always been in the valley, from before the time of their grandparents' grandparents. "*Fue el maíz de los abuelos de mis abuelos.*" Kempton (1924) refers to several centuries of continuous cultivation preceding his 1924 visit. In that era, as today, there were other specialty types of maize grown. Yellow maize was cultivated in drier, sandier areas where other types of maize do not grow as well. Small areas of purple maize were grown for specialty uses and tradition.<sup>5</sup>

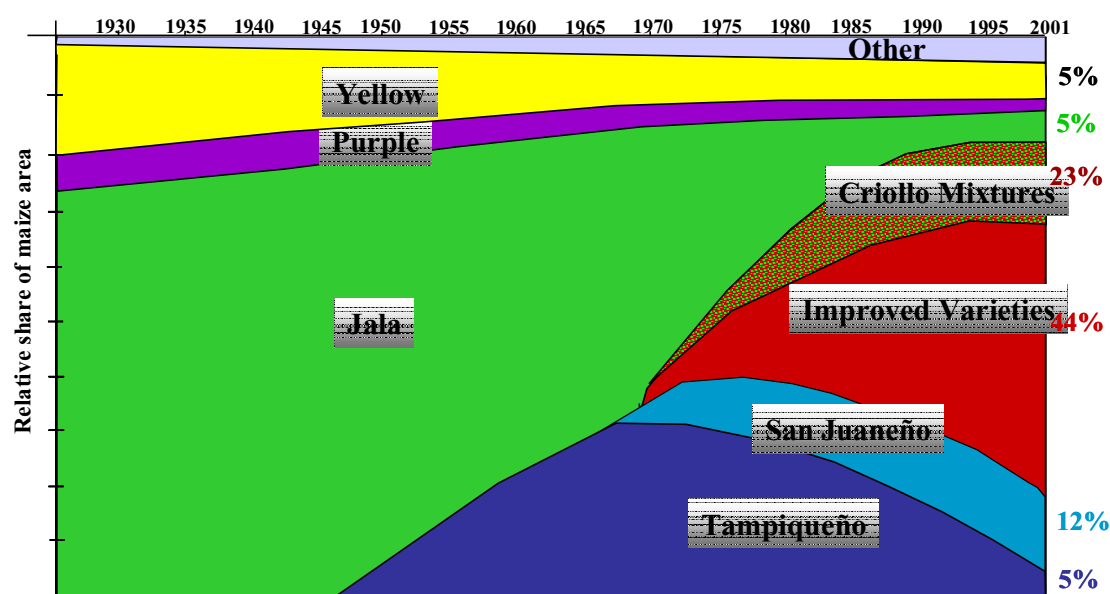
Maize cultivation in the valley changed very little for many years, due at least in part to poor roads and few transport vehicles. It wasn't until the 1950s that another variety began to infringe upon the Jala variety's dominance. Tampiqueño, a variety from the state of Tampico, shared many of Jala's characteristics and filled the same

---

<sup>4</sup> Farmers sometimes make a distinction between *maíz de húmedo* and *maíz chino* as two varieties from the past which together constitute *maíz de húmedo* today. The *maíz chino* had a much fatter cob (about 16 rows) and the *maíz de húmedo* (as evidenced by ear impressions set in concrete of building foundation) was 10-12 rows with a long, thin cob.

<sup>5</sup> One eloquent farmer explained that they grow red and purple maize in memory of the year with the red tortillas. In the first year of the new *ejido*, the *ejidatarios* had trouble finding seed. The *hacienda* owners had to surrender their land, but refused to co-operate in other ways, withholding animals and seed. The only seed the *ejidatarios* could find for that first year, according to the stories, was purple/red seed from the next valley. That year, he told me with a smile, they ate red tortillas with their hand over their eyes to ward off any bad spirits from what must be the work of the devil.

niche as a staple grain. Initially, Tampiqueño was shorter than the Jala variety, with smaller ears, whiter grain and a denser texture. However, years of planting these varieties in small adjoining plots all over the valley have mixed them to a great extent. Physically distinct varieties in farmers' fields often cannot be distinguished from one another genetically (Busso et al. 2000; Pressoir and Berthaud 2004b). In much the same way, genetic analysis using 22 microsatellite (SSR) markers on 24 individuals of a Tampiqueño and a Jala population collected for the genebank in 1988 show no detectable genetic differentiation between the populations (Chapter 4).



**Figure 2.3 Reconstructed maize areas planted in Jala from 1925 to 2001 based on 36 informal interviews in 1999, and a formal survey of 79 households in 2001<sup>†</sup>**

<sup>†</sup>The figure is based on 36 informal interviews from 1999. The reconstruction is intended only to illustrate gross proportions of area cultivated and is only as accurate as the recollections of the interviewees

<sup>‡</sup>Colored percentages on the right side of the figure reflect actual proportions of each variety found in 2001 random survey of 79 households.

*The CONASUPO years*

Today, the San Juaneño variety is more popular than Tampiqueño and Jala combined. The San Juaneño variety is also tall, with large ears. It has dense, very white grain. Density is not usually valued by Mexican farmers processing their harvest for home use and local sale. A softer grain, like that of Jala or Tampiqueño, is easier to grind by hand and makes fluffier flour, good for local delicacies.

In the 1970s, according to farmer accounts, CONASUPO arrived in the valley,<sup>6</sup> radically shifting the way farmers in Jala grew maize. CONASUPO was the buying agent for the government, buying maize at pre-established prices to promote stability in the countryside (Ochoa 2000). Selling one's harvest and buying one's tortillas represented a fundamental change for Mexico's farmers.<sup>7</sup> Even without the subsidized tortilla price often provided by the Mexican government, buying one's tortillas represents a tremendous savings in labor, while selling one's harvest carries a reduced risk of post-harvest losses. Most farmers agree that these benefits come at a slight sacrifice of tortilla flavor and quality.

Not surprisingly, the ideal qualities of grain quickly shifted as well. CONASUPO bought grain by weight, instead of by the traditional volume measure, and was only interested in white grain. Flavor and texture were no longer as important to farmers. The Jala variety, with its large, light-weight grains and creamy color

---

<sup>6</sup> CONASUPO was formed in 1961, replacing the State Food Agency, founded in the mid-1930s which had the same goal of buying from farmers at guaranteed prices (Ochoa 2000, p. 39). Farmer recollections that there was no extension of this agency in Jala until the 1970s are consistent with CONASUPO's push in the mid-1970s, under the López-Portillo administration, to increase their presence in small rural communities with populations between 500 and 10,000 (Flores 1999, p. 190).

<sup>7</sup> Many in Mexico still have not made this shift. De Janvry et al. (1995, 1997) pointed out that the majority of producers in Mexico are not producing maize for the market. However, in Jala, only 6.4% of the surveyed population was using their maize purely for household consumption. A few were grinding their entire harvest and feeding it to their animals (12.8%), implying that they were buying their food maize, even though they were not selling their harvest. Another 28.2% of surveyed farmers were consuming some of their harvest and selling the rest. The remainder was selling their entire harvest.



quickly fell out of favor. It took farmers several years to convince the CONASUPO buyers that the grain, once processed, yielded white flour since the creamy color is only in the pericarp, which is washed away in processing. Dense varieties, like San Juaneño became increasingly popular.

At about the same time, farmers began to experiment with improved varieties. Quickly these primarily white, dense varieties came to occupy the primary maize niche. The improved varieties were fundamentally different from the traditional varieties in Jala: usually short plants, with dense white grain, a short growing cycle and tolerance to dense planting. Many farmers saw improved varieties' yield advantages and ease of sale; their areas have steadily increased to 44% of maize area today (Table 2.1). To this day, some farmers dislike improved varieties complaining that they are so short that they hurt one's back to harvest, they do not produce enough silage for animals and they require too many inputs.

Many farmers chose not to buy improved seed every year, but instead kept it from one year to the next, managing it in much the same way as traditional varieties. Farmers use the word *criollo*, meaning local or one's own, to describe these varieties. These *criollo* varieties accounted for 23% of the maize area in the valley (Table 2.1), and occupied an intriguing genetic niche. Their genetic foundation was that of the improved varieties, but quickly they began to take on the genetic and physical characteristics of the local varieties. These varieties may be important reservoirs of the genetic diversity of traditional varieties, incorporated through pollen movement (see Chapter 5).

**Table 2.1 Distribution and area of different classes of maize in Jala 2001**

|                        | % of farmers | % of total area | Avg. area per farmer(ha) | SD area per farmer (ha) |
|------------------------|--------------|-----------------|--------------------------|-------------------------|
| Traditional variety    | 62.3%        | 32.9%           | 2.5                      | 2.0                     |
| Improved variety       | 39.0%        | 44.1%           | 5.1                      | 4.3                     |
| <i>Criollo</i> variety | 32.5%        | 23.0%           | 3.0                      | 2.3                     |
| Total                  | 77 farmers   | 365.6 ha        |                          |                         |

### ***Current Jala maize varieties***

#### *Improved varieties*

Many improved varieties have come and gone in Jala. The emergence of a cottage industry in processing and packaging maize husks has increased the value of a few varieties over the others. Currently a hybrid called *Dekalb 880*, known locally just as “880”, has risen to prominence because it produces prolific husks as well as good grain on robust plants. It alone accounted for half the area planted with improved varieties. The yields from improved varieties, as described by the farmers who plant them, were better than traditional varieties under all circumstances, though the advantage was much greater in good years than in normal or bad years (Table 2.2). In a normal year, the yield difference between the best improved variety (880) and the best traditional variety (San Juaneño) was only 0.4T/ha, about 11% of the total; not enough to justify the cost of a bag of new seed. ( $0.4 \text{ T} * 1200\text{pesos/T} = 480 \text{ pesos}$ ; a new bag of seed costs 650 pesos). However, in a good year, the difference was 1.1T/ha, which more than justifies the seed cost.

In other parts of Mexico, these improved varieties are considered the provenance of the rich (Bellon and Risopoulos 2001). In Jala, they were the provenance of the larger landholders, with 79% of larger farmers growing them compared with 26% of small landholders. The improved varieties in the valley were hybrids, and these farmers said they were buying new seed every year.

#### *Criollo varieties*

*Criollo* varieties are sometimes seen as a “middle way” by which the poor can access technology and adapt varieties to their local conditions without the cost of buying seed every year (Bellon and Risopoulos 2001). Not surprisingly, most farmers

growing *criollo* varieties grew a *criollo* version of the popular *Dekalb* 880, though it showed no appreciable yield advantage over *criollo* versions of other improved varieties. *Criollo* varieties accounted for 23% of the valley's maize area, and 880 *criollos* accounted for more than half of the *criollo* area (Table 2.1). Interestingly though, *criollo* maize was the lowest yielding of all categories (Table 2.2). *Criollos* were considerably lower yielding than their new, improved counterparts, and had no yield advantage over traditional varieties.

**Table 2.2 Subjective estimates of maize yields for different varieties and categories of maize grown in Jala 2001, for good, normal and bad years<sup>†</sup>**

|                      | -----Good year----- |      |     | -----Normal year---- |      |     | -----Bad year----- |      |     |
|----------------------|---------------------|------|-----|----------------------|------|-----|--------------------|------|-----|
|                      | T/ha                |      |     | T/ha                 |      |     | T/ha               |      |     |
|                      | average             | SD   | (n) | average              | SD   | (n) | average            | SD   | (n) |
| <u>New Improved</u>  | 5.15                | 1.74 | 23  | 3.52                 | 1.22 | 20  | 2.29               | 1.38 | 21  |
| 880 new              | 5.40                | 1.70 | 18  | 3.77                 | 1.07 | 16  | 2.37               | 1.50 | 17  |
| Other new            | 4.22                | 1.72 | 5   | 2.54                 | 1.42 | 4   | 1.94               | 0.72 | 4   |
| <u>Criollo</u>       | 3.32                | 1.67 | 16  | 2.43                 | 1.19 | 18  | 1.56               | 0.79 | 17  |
| 880 <i>criollo</i>   | 3.62                | 1.52 | 7   | 2.63                 | 1.13 | 10  | 1.68               | 0.72 | 9   |
| Other <i>criollo</i> | 3.09                | 1.83 | 6   | 2.17                 | 1.19 | 5   | 1.42               | 0.79 | 5   |
| <u>Traditional</u>   | 3.44                | 1.30 | 32  | 2.68                 | 1.12 | 32  | 1.67               | 0.97 | 29  |
| SanJuaneño           | 4.30                | 1.35 | 10  | 3.39                 | 1.07 | 11  | 2.19               | 0.86 | 9   |
| Jala                 | 3.50                | 1.48 | 6   | 2.61                 | 1.19 | 5   | 1.75               | 1.44 | 5   |
| Tampiqueño           | 3.25                | 1.21 | 6   | 2.74                 | 1.26 | 6   | 1.49               | 1.14 | 5   |
| Perla                | 3.25                | 0.35 | 2   | 2.50                 | 0.71 | 2   | 2.00               | 0.71 | 2   |
| Tabloncillo          | 3.00                |      | 1   | 2.00                 |      | 1   | 1.00               |      | 1   |
| Mixture              | 2.74                | 0.73 | 5   | 1.86                 | 0.63 | 5   | 1.30               | 0.65 | 5   |
| Trad. yellow         | 2.08                | 0.65 | 6   | 1.42                 | 0.45 | 5   | 0.80               | 0.57 | 5   |

<sup>†</sup>Subjective yield data (Bellon and Risopoulos 2001) consists of farmers' answers to questions about yield. It is not objective data from controlled yield trials. It is intended only as a relative indicator of yield, as it is subject to farmers' perceptions and in no way controls for differences between fields or between farmers' management strategies.

### *Traditional varieties*

Most farmers in Jala (62%) grow traditional varieties (Table 2.1). The area planted to various varieties corresponds with the yields farmers describe for each. San Juaneño was the most prevalent traditional variety, accounting for 36% of the area

planted to traditional varieties, and also had the highest described yield [at 3.4T/ha, it is 0.8T/ha greater than the average for traditional varieties (Table 2.2)]. Tampiqueño and a traditional yellow occupied similar areas, at 16% of traditional area. The yellow had the lowest yield of the traditional maize types, because it was grown in dry, sandy, much less productive land. The Tampiqueño, however, was the second most productive of the traditional varieties (2.7 T/ha). The Jala variety accounted for only 14% of the traditional area, and had yields of about 2.6 T/ha. The remainder was comprised of other, lower-yielding varieties like Perla and mixtures between different traditional varieties.

Many of Jala's farmers grow only traditional varieties (24 of the 48 farmers who grow traditional varieties). In general, those growing only traditional varieties commit more of their land to those varieties, but also tend to be the farmers in the survey with smaller land holdings (22/24 have five hectares or less). The other half of farmers with traditional varieties grows them in conjunction with either improved varieties or *criollo* varieties. No one surveyed grew all three categories of maize.

### ***Jala variety***

Though two-thirds of farmers in Jala grew some sort of traditional variety, far fewer farmers (only 19%) are planting the traditional Jala variety (Table 2.3). They also planted Jala in smaller plots than their other varieties, and so it accounted for only about 5% of the maize area in the valley. Most of the farmers (77%) had grown this variety in the past, a story consistent with the oral histories represented in Figure 2.3. They stopped growing Jala an average of 18 years ago, for the set of reasons shown in Table 2.4. Not surprisingly, many of the responses were related to sale-related characteristics like density, grain color and husk quality. The other set of reasons for abandonment had to do with plant characteristics associated with production: the

plants grow too tall and lodge in high winds, they did not resist insects well, and they should be planted early.<sup>8</sup>

**Table 2.3 Distribution, size and longevity of Jala variety plantings in Jala, Mexico, 2001**

| Descriptor   | n          |    |
|--|------------|----|
| Share of valley's maize area planted with Jala (%) | 4.7%       | 79 |
| Farmers who plant Jala (%)                         | 19.0%      | 79 |
| Farmers who planted Jala in the past (%)           | 77.2%      | 61 |
| Stopped growing Jala how many years ago? (years)   | 18.4 years | 61 |
| Average area of Jala (ha)                          | 1.1 ha     | 15 |

**Table 2.4 Reasons surveyed farmers stopped growing Jala variety<sup>†</sup>**

| Reason given                            | percentage | n  |
|---|------------|----|
| lightweight grain                       | 62%        | 28 |
| lost the seed                           | 36%        | 16 |
| plant height too tall/lodging           | 18%        | 8  |
| insect problems                         | 16%        | 7  |
| not good for husk industry <sup>‡</sup> | 9%         | 4  |
| yellowish color                         | 4%         | 2  |
| early planting                          | 2%         | 1  |
| total                                   | 147%       | 45 |

<sup>†</sup>These were responses to an open-ended question about why farmers had stopped growing the Jala variety. If a farmer gave more than one reason, it is counted under all relevant categories, thus the total sums to more than 100%.

<sup>‡</sup>The Jala variety has thin husk coverage, which contributes to its insect problems. However, the husks are very long. Many farmers cited this as an advantage of the variety, saying they could fit two tamales in one husk. However, there is no market for such long husks outside of Jala because they are so unusual.

Farmers grew Jala in two distinctly different ways. One group of farmers, the “generalists,” grew a small area of Jala (defined as less than 1.5 ha, with an average of 0.4 ha) that represented a small fraction of their total cultivated area (an average of 6.6% of their total land) (Table 2.5). These farmers were generally older, had grown the variety for many years, and had kept their seed lots for most of their lifetime. Nearly all planted their own seed. They were diversified farmers: they had sorghum,

<sup>8</sup> Far more farmers saw early planting as an advantage of Jala, allowing them to distribute the work of planting over a longer period of time. They also added that the variety could be planted later, but *elotes* wouldn't be ready in time for the August 15<sup>th</sup> festival week.

sugar cane, tobacco, peanut and pasture grass. Most of them had cattle. They grew Jala predominantly for consumption characteristics because it was good for *pozole* (a corn soup), had good texture for *marquezote* (a sweet corn bread) and *gorditas* (a sweet corn cookie), and made excellent *elotes* (corn-on-the-cob). They ground it for their animals. Several of them also cited tradition as a primary reason for growing Jala.

**Table 2.5 Profiles of Jala growers in Jala, Mexico 2001**

|  | “Generalists”<br>(<1.5 ha Jala) |       | “Specialists”<br>( 1.5 ha Jala) |       |
|--|---------------------------------|-------|---------------------------------|-------|
|  | average                         | SD    | average                         | SD    |
| Farmers growing Jala (no.)                         | 9                               |       | 6                               |       |
| Jala area (ha)                                     | 0.41                            | 0.42  | 2.08                            | 0.49  |
| Total land area (ha)                               | 8.96                            | 6.90  | 3.17                            | 1.17  |
| Share of total land in Jala variety (%)            | 6.6                             | 6     | 73.3                            | 29    |
| Farmers’ age (years)                               | 73.00                           | 13.13 | 59.83                           | 8.66  |
| Time planting variety (years)                      | 60.75                           | 11.78 | 16.50                           | 19.50 |
| Time planting own seed lot (years)                 | 37.63                           | 30.13 | 6.33                            | 11.62 |
| Farmers who plant their own seed (no.)             | 8 of 9 (89%)                    |       | 3 of 6 (50%)                    |       |
| Cattle per farm (no.)                              | 50.375                          | 65.7  | 0                               | 0     |
| Farmers with cattle (no.)                          | 7 of 9                          |       | 0 of 5                          |       |
| <u>Reasons for growing Jala</u>                    |                                 |       |                                 |       |
| Consumption characteristics                        | 5 of 8 (63%)                    |       | 2 of 5 (40%)                    |       |
| Tradition  | 2 of 8 (25%)                    |       | 1 of 5 (20%)                    |       |
| Production characteristics (ease, low input, etc.) | 1 of 8 (13%)                    |       | 2 of 5 (40%)                    |       |
| <u>Marketing of Jala</u>                           |                                 |       |                                 |       |
| yes, there is a market for Jala                    | 4 of 7 (57%)                    |       | 4 of 4 (100%)                   |       |
| sell <i>elote</i>                                  | 1 of 7 (14%)                    |       | 3 of 5 (60%)                    |       |
| sell grain   | 4 of 7 (57%)                    |       | 2 of 3 (67%)                    |       |
| sell seed  | 4 of 7 (57%)                    |       | 1 of 2 (50%)                    |       |
| sell forage  | 1 of 6 (17%)                    |       | 1 of 2 (50%)                    |       |

The second group of Jala growers was younger (by an average of 13 years) and more market driven. They planted larger areas of Jala (average 2 ha) and it represented a much higher percentage of their total land (73%, as opposed to 7% for the smaller growers). They were the Jala “specialists.” They had no other crops, few

other maize varieties and no cattle. In general, they had been growing the variety for less time (even after accounting for the difference in their average age) and renewed their seed more often. Only half of them were planting their own seed. The majority (3 of 5) of them sold the vast majority of their crop as *elote* during the festival week (a lucrative strategy, but one incompatible with keeping one's own seed).

Every August 15, Jala hosts the annual *Feria de Elote* (Corn-on-the-cob Festival), associated with the traditional, religious festival week for Jala. In 1981, Jala's mayor Carlos Carilo Santana created the contest and the week long *Feria de Elote* expressly to promote the preservation of Jala maize (Listman and Estrada 1992). The festival alone is responsible for selling many people's entire Jala crop as *elote* (corn-on-the-cob). It is sponsored jointly by the municipal government and by the Secretary of Agriculture and Water Resource (SARH) and gives a cash prize equivalent to about 10 days worth of wage labor (Listman and Estrada 1992). Family members return from abroad, the town's population swells and the town takes on a carnival atmosphere. It all culminates in the *Elote Más Grande del Mundo* (biggest ear of corn in the world) contest on the last day of the festival. Many of the farmers interviewed were growing small areas of the traditional variety, just to be able to compete in the contest.

Most of those who continue to grow the Jala variety found a way to make it profitable. Regardless of the amount they plant, the Jala growers were known within the community and most sell grain or seed to their neighbors. Sales of Jala grain were by volume, not by weight. Maize sold as *elote* fetched an average price of 1.2 pesos/ear (Table 2.6). Using a conversion factor of 500g dry weight per ear, farmers selling maize as *elote* earned twice as much as they could selling the grain by weight to a buyer like CONASUPO. Jala sold as grain fetched a price of 9.8 pesos/ *medida* (*medida* is a volume measure equivalent to 3.29 kg Jala grain, according to the survey

farmers) – a premium of 2.5 times the general maize grain price given by surveyed farmers. Farmers received an even larger premium (of 3.5%) for Jala seed (though that price fell far short of the price of commercial seed (650pesos/35kg bag=18,570pesos/T, a multiplier of 15.5).

**Table 2.6 Prices for Jala variety sold as *elote* (corn-on-the-cob), grain and seed, compared to average maize grain price, 2001**

| Sells as                      | Price (pesos)             | Conversion factor      | Price/ T (pesos) | Premium multiplier |
|-------------------------------|---------------------------|------------------------|------------------|--------------------|
| <i>Elote</i>                  | 1.2 pesos/ ear            | dry wt/ ear = 500g     | 2400             | 2.0                |
| Grain                         | 9.8 pesos/ <i>medida</i>  | <i>medida</i> = 3.29kg | 2979             | 2.5                |
| Seed                          | 13.8 pesos/ <i>medida</i> | <i>medida</i> = 3.29kg | 4195             | 3.5                |
| Maize grain price from survey | 1,200 pesos/ T            |                        | 1200             |                    |

### ***Farmer adaptations to changing circumstances***

#### ***Migration and farmer age in Jala***

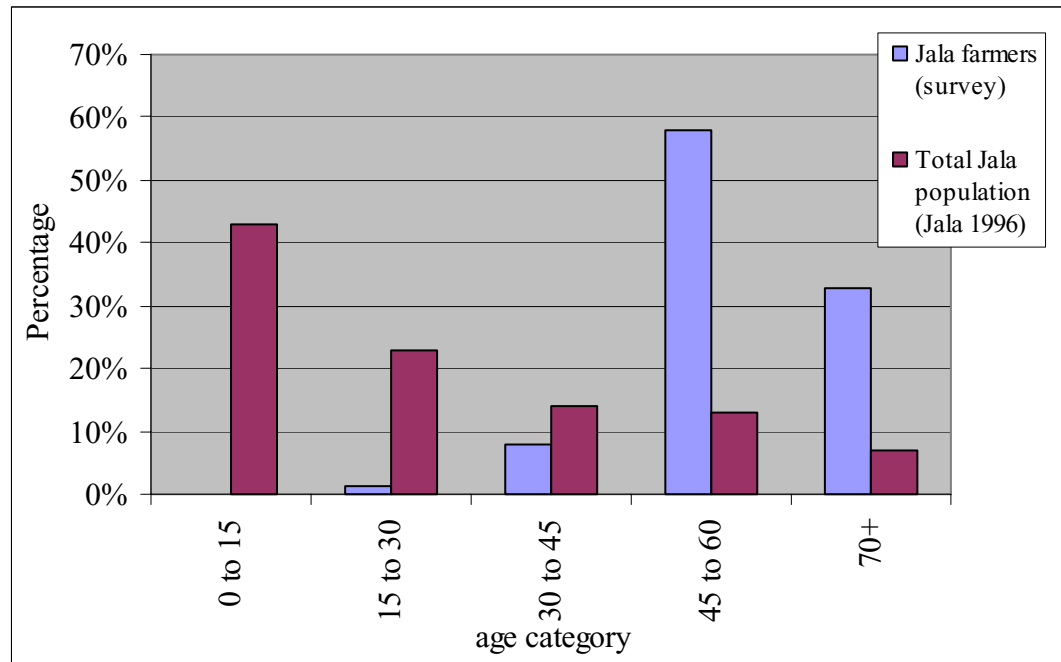
Like many villages in Mexico, Jala has seen the migration of generations of its young people to other cities in Mexico or to the United States—places where they have higher earning potential and where more jobs are available. Remittances from family members who have migrated away, on average, represent about 7% of household income in Mexico as a whole (de Janvry and Sadoulet 2001). Of the 25.6% of adults who leave the state of Nayarit, the vast majority (84.2%) go to the United States, for a total of 21.6% of the adult population.

Perhaps as a result, the average age of farmers had a very different profile than that of the general populace in Jala (Figure 2.4). Average age in the nationwide *ejido* study (de Janvry et al. 1997) for the head of household<sup>9</sup> was 49.6, a marked difference

<sup>9</sup> The difference in ages could be attributed to differences in the sampled populations: all producers were sampled in Jala, while the de Janvry study only sampled *ejidatarios*. However, the *pequeño propietarios* in Jala are younger than the *ejidatarios*, thus their inclusion cannot be the reason for the agedness of Jala farmers. The farmers interviewed in this survey were the people making the production decisions and working the land. In many cases, this was a widow's son, even though she



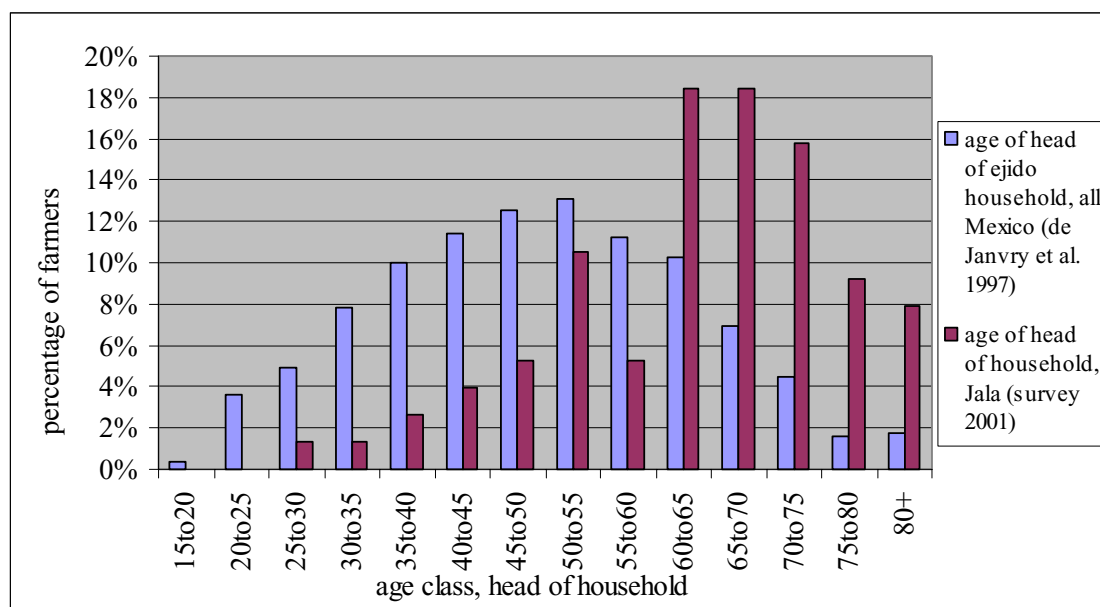
from Jala where the average head of household age was 64.3 (Figure 2.4). The difference was likely due to regional differences; de Janvry et al. (1997) created an average for *ejidatarios* over all of Mexico, while Jala is clearly in a migratory region (de Janvry and Sadoulet 2001) and had an older population. The rising age of farmers could also reflect a change between 1994 and 2001. Farmers may give up land to their children less readily in a mechanized age, where farming is less physically demanding. This is likely to be especially pronounced now that *ejidatarios* are not required to personally work the land because of changes to Article 27 of the Mexican Constitution.



**Figure 2.4 Age of surveyed farmers versus age of population in Jala, Mexico, 2001**

---

was the title holder. The sampling method used in this study would reduce head of household age if anything.



**Figure 2.5** Ages of farm heads of households in Jala (2001 survey) compared to average for Mexican ejidatario heads of household (de Janvry et al. 1997)

#### *Niche markets*

Niche markets for local use (silage, *elotes*, *pozole* and floury-textured maize for baking) tend to favor traditional varieties, in particular the Jala variety. Many farmers growing the Jala variety were explicitly producing for these niches. *Elote* and *pozole* markets are likely to remain active as the village population ages—people of all ages prepare these foods. However, the niche for floury textured maize for grinding and baking is likely to decrease as a new generation that is more accustomed to purchasing prepared foods, replaces the older generation.

The local market for husks has influenced the desirable maize production characteristics heavily towards improved varieties with abundant husk coverage. Farmers typically enter into an agreement with a husk processor before planting. Usually the farmer received a bag of improved seed (value 650 pesos) and the husk processor harvested the field and removed the husks. Some farmers used *criollo* seed and received a cash payment roughly equivalent to the cost of seed. Others chose to

sell their husks, by the kilogram, after the harvest. The husk purchasing market encouraged farmers to invest in new, improved seed where they might not otherwise. Though the husks of the Jala variety were prized for their size, the variety was also known for its poor husk coverage (few husks, not tightly closed) and therefore it was not considered desirable for husk production.

### *Consolidation and land size*

The consolidation of Mexican agriculture predicted to accompany trade liberalization has not happened rapidly in Jala: the two largest farmers in the study still only planted 20 hectares, and only eight farmers (10% of those surveyed) planted 15 or more hectares. However, the farmers with the most land owned much more than those with the least and also rented much more, implying that they were making concerted efforts to increase their landholdings. The farmers with large landholdings tended to be diversified and to have animals. Many had small areas of Jala they planted for niche markets. Most had several types of maize, and often they had other crops as well.

Many of Jala's farmers with larger landholdings (>10ha) adopted new, improved, higher-yielding varieties of maize. Increasingly, these farmers expected to buy seed every year and to use chemical inputs on it. Farmers expected to see yields on improved varieties about 35 percent higher than their traditional maize in normal years and bad years, and nearly 50% higher in good years. However, almost half of these large-landholding farmers also had traditional varieties.

### *Fragmentation of the landscape*

Land size in Jala was small; farmers had an average of 5.9 hectares. Those hectares were further subdivided into several small parcels of land for each farmer.

The fragmentation of the agricultural landscape all over Mexico is the legacy of the *ejido* system (Bellon and Brush 1994). Jala was planted by 19% of farmers, but it accounted for only 5 % of the maize area; thus, it was planted in very small plots scattered across the valley. Such small areas of planted Jala, combined with the annual selection pressure when farmers save small numbers of ears for their future small plots of Jala, could result in reduced genetic diversity. However, the genetic data presented in Chapter 3 showed that Jala populations from farmers' fields are highly diverse.

The small size of farmers' fields, coupled with the overlapping flowering times between the Jala variety and improved varieties in the area (Figure 2.2), also suggests the possibility of gene flow between the different types of maize. Preliminary genetic evidence (Chapter 5) was consistent with gene flow from the improved varieties and into Jala—a type of genetic movement of particular interest with the increasing prevalence of genetically engineered maize (Bellon and Berthaud 2004). The evidence of gene flow into *criollo* varieties was clear, raising the potential of *criollo* varieties as overlooked reservoirs of traditional genetic diversity (Chapter 5).

### *Animal production*

Animal production, particularly beef production, leads to an increased demand for feed. Many farmers pointed out that their maize was worth more locally as feed than as food. Beef, pork and poultry production facilities in the Jala valley were growing. Several large feeding operations opened within the village itself, as well as in the surrounding hillsides. Cattle were the most common animal, but some farmers were venturing into hog and chicken production. The growth of animal production has been common throughout Mexico in the 1990s (de Janvry et al. 1997). Several surveyed farmers mentioned that their maize was worth more as silage than as grain.

The Jala variety produces large amounts of biomass and thus is favored by some farmers for silage production.

### *High-value, maize-alternative crops*

In Mexico, there has been a strong trend towards production of vegetables as high-labor, high-value alternatives to traditional maize and bean cultivation (de Janvry et al. 1997). In 1999, a high value, high labor alternative arrived in Jala: tobacco. Ironically, Jala's maize and maize farmers drew the tobacco companies to the Jala valley. Maize farmers are accustomed to tending their crops carefully, as tobacco needs to be tended, in contrast to sugar cane farmers found in nearby communities, whose crop requires less careful attention. Many farmers in Jala began contracting to grow tobacco. Some farmers were convinced that they could make a great deal more money growing tobacco than they would with maize. Other farmers remained skeptical, pointing towards the high costs and debt incurred. Nevertheless, the area of tobacco was increasing rapidly in Jala.

## ***Implications and discussion***

### *Current situation*

The Jala variety has been well-conserved on farms in the Jala valley. Farmers continued to grow the traditional variety for reasons ranging from sentimental to purely economic. For the loci assayed, the genetic data showed remarkable stability in Jala's genetic profiles across time and space in the valley (Chapter 4). As long as farmers continue to grow Jala in large enough areas, conservation of its genetic diversity will continue. The question of how large is large enough remains open, especially in light of emerging evidence that the unit of conservation may be the valley, not the individual household (Chapter 4). However, Jala from farmers' fields

appeared to be highly diverse (given the methods used in Chapter 4), implying that areas of Jala currently grown in the valley are sufficient to maintain high levels of genetic diversity. High levels of genetic diversity could also be caused by high rates of gene flow from other varieties. However, the stability of Jala's allelic profile over time suggests that it has been well conserved in the field, maintaining both genetic integrity and diversity.

Central to many of these conclusions is the concept that farmers with different size plots of land manage their agriculture, in particular their maize, differently. The small landholders who grew Jala tended to be younger and grow larger areas of it. In essence, they “specialized” in Jala, usually targeted for a niche like *elote* production. They were less likely to save seed, relying on others in the community to do so for them. The producers with large amounts of land tended to have smaller areas of Jala—they were “generalists”. They grew Jala for household consumption and tradition and kept small amounts for local sale. Though the areas they plant are much smaller, they have usually kept their seed lots for their entire lifetimes.

Though about 50% of farmers were growing traditional varieties, the large land holders were growing them together with improved varieties or growing improved varieties only. The small land holders, on the other hand, were growing traditional varieties only or growing them with *criollos*—both important repositories of traditional genetic information. The large land holders tended to be more diversified with animals, other crops and different maize varieties, while the small land holders tended to specialize in niches. This trend could be explained in part by the fact that large land holders had more plots of land to work with, but also because they were likely to devote more of their household earning power to agriculture while smaller land holders depended more on outside sources of income (de Janvry and Sadoulet 2001).

In general, Mexico's past policies legislated a national agricultural sector composed of many small farms and then subsidized the sector to allow the rural poor to eke a living from the land. This relatively economically inefficient agricultural sector may have contributed to the surprising finding that instead of replacing their traditional varieties with improved ones, Mexican farmers have continued to grow the traditional alongside the improved (Rice et al. 1997; Aguirre-Gomez et al. 2000; Louette and Smale 2000; Bellon and Risopoulos 2001). This implies that the other characteristics like flavor, low input demand, low investment and tradition have had value equivalent to the potential gains with improved varieties. With the *ejido* reform and removal of subsidy and price supports, many have argued that Mexico's agriculture is on the verge of becoming larger-scale and more economically efficient, at least in areas like Jala where economies of scale are possible. This consolidation of agriculture is unlikely to favor traditional varieties, as larger land holders are far more likely to be growing improved seed. However, at the moment in Jala, the larger farmers are also more likely to be diversified and are growing more varieties.

#### *Future of on-farm conservation in Jala*

Over time, the market for maize has shaped the perception of desirable characteristics of maize. Traditionally, as maize was used for local consumption, consumption characteristics like flavor, texture and ease of processing held primary importance. The Tampiqueño variety came to share Jala's place as the primary grain because it shared many of the same desirable consumption characteristics. However those desirable characteristics shifted when farmers began to sell their harvest by weight to the national CONASUPO buyer. CONASUPO was only interested in white grain, texture was no longer important and weight became the primary desirable characteristic. The traditional Jala variety fared poorly on this scale and fell into

disfavor, yielding its primary niche to the denser, whiter San Juaneño and improved varieties. However, with the end to nationalized buying in recent years, desirable characteristics could change again, perhaps tipping the balance back in favor of the traditional Jala.

The current shift in the valley towards more animal production, coupled with the low price for maize grain means that more and more farmers are selling the entire biomass from their fields to grind as silage. Some farmers prefer the traditional Jala variety for silage because the plants are so large and produce so much biomass. The local market for the Jala variety continues to pay a premium: twice the average maize grain price for Jala *elotes*, 2.5 times that price for Jala grain and 3.5 times that price for Jala seed. The *elote* demand is driven largely by the week long festival, and is likely to stay the same or increase as long as the festival continues. The local market for seed will be driven by farmer's desire to plant the variety. The market for grain for making specialty baked items, where floury texture is prized, may change as a younger generation, more accustomed to buying their food than making it, becomes the village's primary consumers. *Pozole* demand is likely to stay the same, as people of all ages make this dish.

Much hinges on the decisions of the future generation of farmers. Will they be interested in continuing to grow the traditional variety? They will continue only if there is incentive, economic or otherwise, to do so. The economic and policy context will play a critical role in determining whether or not they will continue to grow the Jala variety.

There are several forces working against Jala's future, as well. The thriving husk industry does not value the traditional variety. Though some farmers prize the husks from the Jala variety because they are so large, the variety does not have good



husk coverage (contributing to problems of insect damage in the field) and thus would require some breeding efforts before it produced a desirable volume of husks.

The Mexican program to target poor households with subsidized tortillas may also work against conservation of the traditional varieties. Poorer households with smaller areas of cultivated land were often growing traditional varieties. The program creates an indirect incentive to sell one's harvest when one receives free tortillas, thus devaluing production characteristics and nudging farmers away from traditional varieties. Though this may not be a bad strategy from a development perspective, it could be a harmful one from a conservation perspective. Finally, the shift all over Mexico away from maize and towards higher value crops like tobacco bodes poorly for maize in general, and the Jala variety in particular. Farmers who no longer grow maize are far less likely to plant a few rows of a traditional variety on the side.

The formation in 2002 of a producers group to promote the Jala variety is a positive sign, but much will depend on the continuation and hopefully the expansion of the niche markets for the Jala variety. Perhaps the group will be able to find a national *pozole* market or make inroads in the local silage markets. If there is a market for the traditional variety, farmers will continue to grow it. To date, the contest and festival have provided critical incentive, both economic and sentimental, for farmers to continue growing the variety. There is room for improvement in the variety to reduce lodging, lower ear height, improve uniformity, increase husk coverage and improve tolerance of denser planting. Finally, there is need for education and outreach about conservation, so that conservation effects become not hidden consequences of policy, but instead factors to be considered in economic and political decision making.

*Is Jala a good candidate for targeted in-situ conservation efforts?*

Conservation in farmers' fields has been occurring, *de facto*, for thousands of years. Recently, however, on-farm conservation efforts have become more explicitly targeted with frameworks (Maxted et al. 2002; Bellon et al. 2003) and guided by economic principles (Smale et al. 2004). Farmers will not, and should not be expected to conserve if it works against their economic best interest—for example, if it requires forgoing higher yielding varieties or opportunities to diversify into other crops. Instead, farmers must have incentive to conserve (Smale et al. 2004) and conservation efforts should be targeted to varieties that have high value to farmers, perhaps due to consumption or production characteristics, as well as high value in terms of genetic diversity (Bellon et al. 2003).

Jala has high value in terms of genetic diversity: it is unique among other races of maize, it has valuable characteristics in its height and ear size, and it has unique gene complexes (Chapter 4). Therefore it is a good target for conservation, whether in farmers' fields or in genebanks. Farmers continue to grow Jala and have found ways to make it profitable, suggesting that it has long been a good candidate for conservation efforts on the farm. Today, the rising interest in tobacco cultivation and the increasing utilization of improved varieties both suggest that there may be other, economically attractive options for farmers. However, many studies have shown that farmers tend to incorporate new technologies and varieties rather than displacing their older ones [see Wood and Lenne (1997) for a review]. Therefore, if farmers incorporate new crops like they incorporate new varieties, increased tobacco cultivation may not make *in-situ* conservation of the Jala variety infeasible.

Venues like the contest and the week-long *Feria*, which provide economic incentives for *in-situ* conservation, are not likely to disappear. Local markets for specialty grain may change with an aging population; however those changes may be

offset by other alternatives like increases in the local forage market for Jala. It is heartening that younger farmers are specializing in Jala production, implying that they find it profitable and suggesting that they are likely to continue growing the traditional variety. However, these younger farmers are sensitive to policy and economic changes, and on-farm conservation of the Jala variety could be displaced by changing market conditions.

In Jala, activities like the *elote* contest and the week-long *Feria* have been important to on-farm conservation of the traditional variety. Farmers cite tradition and the contest as compelling reasons to continue growing small areas of the Jala variety. However, it is less clear how to replicate these conservation incentives in places where the maize is not as distinctive, or where genetic diversity does not lend itself to contests.

Historically, conservation efforts have focused on preserving phenotypic differences that are obvious, but may not be genetically meaningful. Recent studies have shown that phenotypically distinct traditional varieties are difficult to distinguish genetically (Busso et al. 2000; Pressoir and Berthaud 2004b, Chapter 4). Instead, how do we identify and conserve differences that are genetically meaningful, but not obvious?

## **CHAPTER 3 – EXPLORING ISSUES OF OPTIMAL SAMPLE SIZE FOR DIVERSITY STUDIES: A CASE FROM MAIZE**

### ***Introduction***

Optimal sample size for molecular diversity studies depends on the biological question asked, the populations studied, the markers assayed, the allele frequencies sampled, and the level of error tolerated. Estimating allele frequencies with two decimal accuracy requires one thousand to ten thousand individuals, depending on the loci and allele frequencies sampled (B-Rao 2001). Capturing rare alleles in a population requires fewer individuals than determining allele frequencies: to capture alleles present at frequencies of 0.05 and higher at 150 loci with 95% accuracy requires 210 individuals (Crossa et al. 1993). Diversity studies, however, require only accurate estimates of diversity and population differentiation parameters, rather than highly accurate estimates of allele frequencies or an assurance of capturing nearly all rare alleles. Therefore, diversity studies might not require sampling as many individuals, thereby increasing the ability of such studies to assay more populations in a cost-effective manner.

As sample size increases, the marginal gain from each additional sample diminishes. Therefore, an optimal size should provide a maximum amount of data for a minimum number of samples and will be specific to a particular data set, measurement and question. Several studies use population genetics and statistics to evaluate minimum or optimum sample sizes to capture rare alleles (Crossa 1989; Crossa et al. 1993; Crossa et al. 1994) or to estimate allele frequency distributions (Ott 1992). However, guidance from empirical diversity study data is needed to approximate optimal sample sizes for diversity studies.

There is wide variation in diversity study sample sizes—no obvious optimum sample size could be predicted from the literature. Using fewer individuals and more

markers has generally been the preferred approach to sampling for population genetic studies (Baverstock and Moritz 1996). Isozyme studies used about 6-24 individuals per population, with about 20 markers (Doebley et al. 1985; 1988). RFLP studies had slightly larger sample sizes, from 25-30 individuals per population, often bulked together, in conjunction with 16-29 RFLP markers (Rebourg et al. 1999; Bhattacharjee et al. 2002). As studies have recently shifted to using microsatellite (SSR) markers, there was a slight tendency towards larger sample sizes, with a range from 1-48 individuals per population studied (Matsuoka et al. 2002; Warburton et al. 2002; Labate et al. 2003). Appendix I gives a tabular review of sample sizes, markers and measures used in diversity studies of out-crossing agricultural species. Independent of marker type, there remains a large amount of variation in sample sizes even among studies with similar goals, markers and populations.

Even within the category of diversity studies, optimal sample size will depend on the biological question asked, the populations studied and the number of markers used. For example, to differentiate two very distinct populations like maize and its wild relative teosinte, one would expect to need fewer individuals. The two populations are intrinsically different due to evolutionary history, and therefore few individuals would be required to distinguish the two populations. Matsuoka et al. (2002) clearly distinguished between diverse maize races and wild relatives with only one individual per population, though they used nearly 100 microsatellite markers. On the other hand, very similar populations, like two populations of the same variety from neighboring fields, are likely to require much larger numbers of individuals for differentiation (Pressoir and Berthaud 2004b).

Furthermore, optimal sample size will vary depending on the measure of interest. For example, one would expect to need larger numbers of individuals to accurately determine population differentiation ( $F_{st}$ ) values than for gene diversity ( $H_e$ )

values.  $F_{st}$  values are based upon comparing allele frequencies in different populations, so the number of individuals that carry each specific allele at a specific locus becomes important. However,  $H_e$  values are population-specific and reflect different allele frequencies but without as much importance attached to the individual allele, and therefore estimates  $H_e$  of should be less sensitive to sample size.

This study investigated optimal sample size for four different populations and three different diversity measures: population differentiation ( $F_{st}$ ), gene diversity ( $H_e$ ) and allele number ( $A_n$ ). To do so, we used a resampling program to examine the additional information gained with each additional individual sampled. For population differentiation and gene diversity, the amount of variance around the measure was used as an indication of accuracy of measure for the population. For allele number, the measure itself was examined. Though conclusions about ideal sample size will be specific to a particular population and biological question, empirical data from a range of genetic materials can give an indication of appropriate sample sizes for diversity studies.

## ***Materials and methods***

### *Populations studied*

The populations studied represented a balanced hierarchical sampling of genetic diversity in maize (Table 3.1). At the bottom of the hierarchy were the eight populations of the race Jala, collected from farmers' fields in the valley of Jala, Nayarit, Mexico. Farmers' fields were diverse, open-pollinating populations. Eight populations of the race Jala, collected from 1944 to 1988 in the valley of Jala, were obtained from the International Maize and Wheat Improvement Center (CIMMYT) genebank in Mexico City. These represent depth within one maize race. Representing breadth across maize races was a set of populations each representing a different

maize race. To form the teosinte and maize race subsets, twelve accessions were suggested by experts (John Doebley, Major Goodman, and Suketoshi Taba). The lists were compared to Matsuoka et al. (2002) to choose the most diverse eight populations.

**Table 3.1 Farmers' field Jala, genebank Jala, maize races and teosinte populations studied, with their seed source and catalog numbers**

| Population   | Seed Source | Catalog Number |
|--|-------------|----------------|
| <u>Jala Maize from Farmers' Fields</u>               |             |                |
| FF1  | 1999-1-CP   |                |
| FF2  | 1999-2-CU   |                |
| FF3  | 1999-3-CP   |                |
| FF4  | 1999-4-CP   |                |
| FF5  | 1999-5-CU   |                |
| FF6  | 1999-6-CU   |                |
| FF7  | 1999-7-CU   |                |
| FF8  | 1999-8-CP   |                |
| <u>Jala Maize from Genebank (year of collection)</u> |             |                |
| 1944a  | CIMMYT      | CIMMYT 2246    |
| 1944b  | CIMMYT      | CIMMYT 16037   |
| 1951a  | CIMMYT      | CIMMYT 10500   |
| 1951b  | CIMMYT      | CIMMYT 11223   |
| 1952   | CIMMYT      | CIMMYT 16038   |
| 1968   | CIMMYT      | CIMMYT 7068    |
| 1988a  | CIMMYT      | CIMMYT 11417   |
| 1988b  | CIMMYT      | CIMMYT 11420   |
| 1988c  | CIMMYT      | CIMMYT 11421   |
| <u>Representatives of Maize Races</u>                |             |                |
| Chalqueño (Mexico)                                   | USDA        | PI 483613      |
| Coroico (Bolivia)                                    | USDA        | PI 485399      |
| Guirua (Venezuela)                                   | USDA        | PI 445002      |
| Olotón (Guatemala)                                   | CIMMYT      | CIMMYT 2510    |
| Pira (Colombia)                                      | USDA        | CIMMYT 444165  |
| Tepecintle (Guatemala)                               | CIMMYT      | CIMMYT1842     |
| Tuxpeño (Mexico)                                     | USDA        | CIMMYT 488974  |
| Uchuguilla (Peru)                                    | USDA        | PI 571462      |
| <u>Teosinte (<i>Zea</i> spp.)</u>                    |             |                |
| <i>Z. diploperennis</i>                              | Doebley     | m5             |
| <i>Z. mays hueheutenangensis</i>                     | CIMMYT      | CIMMYT 9479    |
| <i>Z. luxurians</i>                                  | CIMMYT      | CIMMYT 9478    |
| <i>Z. mays mexicana a</i>                            | CIMMYT      | CIMMYT 8771    |
| <i>Z. mays mexicana b</i>                            | CIMMYT      | CIMMYT 11369   |
| <i>Z. mays mexicana c</i>                            | CIMMYT      | CIMMYT 11400   |
| <i>Z. mays parviglumis a</i>                         | CIMMYT      | CIMMYT 8756    |
| <i>Z. mays parviglumis b</i>                         | CIMMYT      | CIMMYT 8781    |

For each population assayed in Table 3.1, 24 individuals were sampled. Crossa et al. (1993) showed that the sample size ( $n$ ) required to retain at least one copy of  $k$  alleles at  $m$  loci with probability  $P$  should be larger than:

$$n > \frac{\log[1 - (P)^{1/m}] - \log(k - 1)}{\log(1 - p)}$$

With 24 individuals, for  $m=22$  loci with  $k=5$  alleles, alleles with frequency  $p > .144$  can be detected with  $P=95\%$  confidence. At  $m=1$  locus with  $k=5$  alleles, alleles with frequency  $p > .087$  can be detected with  $P=95\%$  confidence. Rare alleles of  $p=0.05$  can be detected at  $m=1$  locus with  $k=5$  alleles at  $P=65.8\%$  confidence.

#### *Microsatellite data and DNA extraction*

DNA was extracted using a CTAB method described in Mitchell et al. (1997). The 22 fluorescently labeled PCR primers used to amplify microsatellite (SSR) loci are described in Table 3.2. The 22 SSR markers were widely distributed on maize's ten chromosomes with approximately one marker per chromosome arm. For further details, see Chapter 4.

#### *Resampling program*

To examine the effect of sample size on important measures of genetic diversity, we created a resampling program to sample each of the data sets (teosinte, maize races, genebank accessions, and farmers' fields). The program 1) sampled the data set with replacement, choosing 1000 bootstrapped samples at each sample size from 2 to 24 individuals; 2) computed allele number ( $A_n$ ) and gene diversity ( $H_e$ ) for each bootstrapped population; 3) calculated an  $F_{st}$  value (overall  $F_{st} = \theta$ ) using each bootstrapped sample as described in Weir (1996) as a measure of differentiation between the eight populations within the data set; 4) averaged across the bootstraps



and computed a standard deviation for  $A_n$ ,  $H_e$  and  $F_{st}$ . We tested the program using a test data set and comparing results to output from GDA 1.1 (Lewis and Zaykin 2002).

**Table 3.2 SSR markers used to evaluate Jala, maize race and teosinte populations, with associated bin location, core repeats, and allele size ranges, PIC values, numbers of alleles and missing data**

| Marker    | Bin location | Core Repeat | Allele Size Range | PIC value | Allele No. | Missing Data |
|-----------|--------------|-------------|-------------------|-----------|------------|--------------|
| phi056    | 1.01         | GCC         | 230-277           | 0.77      | 14         | 6.2%         |
| phi011    | 1.1          | GCT         | 206-292           | 0.67      | 11         | 13.7%        |
| phi064    | 1.11         | ATCC        | 76-123            | 0.89      | 20         | 4.6%         |
| phi083    | 2.04         | CTAG        | 126-150           | 0.76      | 8          | 2.5%         |
| phi127    | 2.07         | GTCT        | 97-132            | 0.68      | 14         | 2.7%         |
| phi453121 | 3.01         | ACC         | 207-230           | 0.79      | 20         | 0.8%         |
| phi053    | 3.05         | ATGT        | 169-272           | 0.68      | 12         | 10.3%        |
| phi213984 | 4.01         | ACC         | 284-319           | 0.10      | 9          | 1.8%         |
| phi093    | 4.08         | CTAG        | 261-293           | 0.84      | 23         | 2.4%         |
| phi333597 | 5.05         | AAG         | 204-226           | 0.53      | 6          | 1.0%         |
| phi085    | 5.06         | GCGTT       | 224-268           | 0.77      | 13         | 6.7%         |
| phi159819 | 6            | CCG         | 117-145           | 0.66      | 13         | 1.8%         |
| phi389203 | 6.03         | AGC         | 291-315           | 0.61      | 9          | 4.4%         |
| phi034    | 7.02         | CCT         | 123-164           | 0.81      | 14         | 2.8%         |
| phi051    | 7.05         | AGG         | 133-151           | 0.73      | 14         | 3.7%         |
| phi115    | 8.03         | ATAC        | 292-304           | 0.36      | 2          | 4.2%         |
| phi121    | 8.04         | CCG         | 93-107            | 0.15      | 5          | 1.9%         |
| phi015    | 8.08         | TTTG        | 74-118            | 0.77      | 24         | 2.9%         |
| phi033    | 9.02         | CCT         | 230-269           | 0.65      | 14         | 2.5%         |
| phi032    | 9.04         | TTTC        | 230-244           | 0.76      | 14         | 1.6%         |
| phi050    | 10.03        | AAGC        | 78-100            | 0.42      | 8          | 1.9%         |
| phi062    | 10.04        | GAC         | 154-177           | 0.45      | 7          | 1.0%         |
| Total     |              |             |                   | 0.63      | 12.45      | 3.7%         |

### *Mathematical analysis*

The data was fitted with equations to describe their behavior using MatLab (Mathworks), as well as an  $R^2$  value to assess the quality of fit (Table 3.3).  $H_e$  and  $F_{st}$  values were fit with power function equations that approach a y-offset value, germane to each population. The data were normalized so they could be compared on the same scale. Each value was divided by the y-offset, so all equations approach  $y=1$ . At  $y=1$ , the slope of the equation (the rate of change with each additional data point) was zero.

However, there was still variance because the resampling program generated 1000 datasets for each sample size, even though the amount of variance was no longer changing with additional data points. The lines for each population or set of populations were compared at  $y=2$ , where the lines representing the standard deviation of  $F_{st}$  and  $H_e$  contain twice the amount of variance they did at zero. The populations approach their minimum variance (at  $y=1$ ) slowly. The goal in sampling was not to eliminate variance, but merely to find a point where the variance was acceptably small. By comparing sets of populations at a consistent point, inferences about optimal sample sizes can be drawn.

For the allele number data, the measure itself, rather than the standard deviation, was evaluated because the standard deviation behaved in the same way as the allele number. As the allele number rose, so did the variance of the increasingly large numbers. The allele number data was also fit with a power function equation. These allele number equations, however, did not approach an off-set or asymptote, but instead increased slowly to infinity. However, the slope, or rate of change, declined with each additional individual decreasing slowly so that each additional individual contributes less and less new information. To be able to compare the different equations on the same scale, the allele number data was normalized by the arbitrary point where the each 100 new individuals assayed contribute one new allele, at an average locus (slope of the equation equal 0.01).

#### *Calculation of actual sample values for complete data set*

The  $A_n$ ,  $H_e$  and  $F_{st}$  values for the bootstrapped data sets were not the same as the values for the complete data set of 24 individuals. For the complete, non-bootstrapped data for each set of populations, an overall  $F_{st}$  (also known as  $\theta$ ) was calculated as above, and then jackknifed over populations and loci to produce a

confidence interval according to Weir (1996). GDA 1.1 was used to obtain  $F_{st}$ , allele number, gene diversity ( $H_e$ ), and homozygosity ( $H_o$ ) results (Lewis and Zaykin 2002).

**Table 3.3 Fitted equations for  $F_{st}$ ,  $H_e$ ,  $A_n$  data, based on 1000 bootstrapped samples of 33 populations, with  $R^2$  values to reflect quality of fit**

| Population | $F_{st}$ equation                    | $F_{st} R^2$ | $H_e$ Formula                                       | $H_e R^2$ | $A_n$ equation                               | $A_n R^2$ |
|------------|--------------------------------------|--------------|---|-----------|--|-----------|
| FFpop0     | $y = 0.11(x - 1.12)^{-0.64} + 0.004$ | 0.9989       | $y = 0.0851259(x - 1.84901)^{-0.201944} + 0.18188$  | 0.9865    | $y = 2.536987(x - 1.5840412)^{0.19178002}$   | 0.9953    |
| FFpop1     |                                      |              | $y = 0.0904638(x - 1.72772)^{-0.287443} + 0.183582$ | 0.9904    | $y = 2.7663185(x - 1.7119641)^{0.17029254}$  | 0.9904    |
| FFpop2     |                                      |              | $y = 0.0815682(x - 1.49249)^{-0.475302} + 0.197886$ | 0.9944    | $y = 2.6413252(x - 1.6418696)^{0.19228424}$  | 0.9920    |
| FFpop3     |                                      |              | $y = 0.0856742(x - 1.75118)^{-0.263293} + 0.183284$ | 0.9924    | $y = 2.632461(x - 1.7039167)^{0.16087767}$   | 0.9914    |
| FFpop4     |                                      |              | $y = 0.076477(x - 1.51673)^{-0.52223} + 0.192358$   | 0.9957    | $y = 2.63877(x - 1.6345048)^{0.18290599}$    | 0.9947    |
| FFpop5     |                                      |              | $y = 0.0934199(x - 1.92849)^{-0.130519} + 0.169981$ | 0.9747    | $y = 2.5798154(x - 1.6515683)^{0.17773018}$  | 0.9922    |
| FFpop6     |                                      |              | $y = 0.109666(x - 1.34051)^{-0.477578} + 0.174687$  | 0.9951    | $y = 2.5193389(x - 1.6050447)^{0.18256408}$  | 0.9943    |
| FFpop7     |                                      |              | $y = 0.0855462(x - 1.84897)^{-0.208119} + 0.184024$ | 0.9866    | $y = 2.6405675(x - 1.6941012)^{0.16810992}$  | 0.9930    |
| GBpop0     | $y = 0.17(x - 0.40)^{-1.11} + 0.03$  | 0.9997       | $y = 0.0745745(x - 1.67414)^{-0.355799} + 0.196582$ | 0.9913    | $y = 2.6195658(x - 1.7034255)^{0.15738611}$  | 0.9917    |
| GBpop1     |                                      |              | $y = 0.111474(x - 1.92629)^{-0.117615} + 0.183646$  | 0.9833    | $y = 2.5236029(x - 1.6940359)^{0.1643907}$   | 0.9910    |
| GBpop2     |                                      |              | $y = 0.103012(x - 1.40413)^{-0.491341} + 0.177026$  | 0.9976    | $y = 2.2524461(x - 1.6862473)^{0.12463569}$  | 0.9929    |
| GBpop3     |                                      |              | $y = 0.0933965(x - 1.54947)^{-0.38266} + 0.188927$  | 0.9953    | $y = 2.2967683(x - 1.6210594)^{0.14813001}$  | 0.9949    |
| GBpop4     |                                      |              | $y = 0.0979903(x - 1.88627)^{-0.14717} + 0.177211$  | 0.9852    | $y = 2.4938466(x - 1.6900462)^{0.13223359}$  | 0.9916    |
| GBpop5     |                                      |              | $y = 0.0926278(x - 1.67956)^{-0.301168} + 0.190654$ | 0.9934    | $y = 2.6778203(x - 1.7535896)^{0.14754102}$  | 0.9875    |
| GBpop6     |                                      |              | $y = 0.0905335(x - 1.76744)^{-0.263492} + 0.189739$ | 0.9867    | $y = 2.6048182(x - 1.6221546)^{0.18778194}$  | 0.9935    |
| GBpop7     |                                      |              | $y = 0.0795049(x - 1.69352)^{-0.347386} + 0.182503$ | 0.9914    | $y = 2.7641675(x - 1.7328582)^{0.14997091}$  | 0.9905    |
| GBpop8     |                                      |              | $y = 0.079587(x - 1.77156)^{-0.286764} + 0.186833$  | 0.9886    | $y = 2.7114825(x - 1.6378847)^{0.18714555}$  | 0.9927    |
| MRpop0     | $y = 0.10(x - 0.90)^{-0.81} + 0.065$ | 0.9992       | $y = 0.0911068(x - 1.84062)^{-0.204258} + 0.190303$ | 0.9886    | $y = 2.4334165(x - 1.7840018)^{0.11012292}$  | 0.9889    |
| MRpop1     |                                      |              | $y = 0.0981508(x - 1.51321)^{-0.340021} + 0.187004$ | 0.9953    | $y = 1.6866343(x - 1.4493356)^{0.11222294}$  | 0.9956    |
| MRpop2     |                                      |              | $y = 0.104795(x - 1.86502)^{-0.154366} + 0.18399$   | 0.9821    | $y = 2.1591115(x - 1.6196255)^{0.14547288}$  | 0.9944    |
| MRpop3     |                                      |              | $y = 0.106895(x - 1.33288)^{-0.48296} + 0.169786$   | 0.9967    | $y = 1.8607047(x - 1.4522318)^{0.10570618}$  | 0.9971    |
| MRpop4     |                                      |              | $y = 0.107978(x - 1.78976)^{-0.178197} + 0.194735$  | 0.9882    | $y = 2.0364555(x - 1.5114866)^{0.14105979}$  | 0.9971    |
| MRpop5     |                                      |              | $y = 0.0926285(x - 1.54525)^{-0.399968} + 0.187396$ | 0.9951    | $y = 2.333311(x - 1.6477034)^{0.14757844}$   | 0.9914    |
| MRpop6     |                                      |              | $y = 0.112885(x - 1.84537)^{-0.150972} + 0.193824$  | 0.9889    | $y = 1.9107052(x - 1.5332234)^{0.11347368}$  | 0.9981    |
| MRpop7     |                                      |              | $y = 0.100398(x - 1.79281)^{-0.202093} + 0.191414$  | 0.9910    | $y = 1.6750287(x - 0.93895017)^{0.15912154}$ | 0.9989    |
| Tpop0      | $y = 0.11(x - 0.98)^{-0.96} + 0.07$  | 0.9991       | $y = 0.137427(x - 1.82871)^{-0.113576} + 0.176146$  | 0.9860    | $y = 1.8718697(x - 1.566004)^{0.14794183}$   | 0.9955    |
| Tpop1      |                                      |              | $y = 0.106287(x - 1.38113)^{-0.468352} + 0.182124$  | 0.9942    | $y = 2.3709464(x - 1.6672473)^{0.14329967}$  | 0.9920    |
| Tpop2      |                                      |              | $y = 0.115601(x - 1.94112)^{-0.0965234} + 0.172712$ | 0.9797    | $y = 1.9081078(x - 1.6310466)^{0.12583403}$  | 0.9926    |
| Tpop3      |                                      |              | $y = 0.0973764(x - 1.55698)^{-0.357713} + 0.202178$ | 0.9957    | $y = 2.3676755(x - 1.650659)^{0.15136079}$   | 0.9948    |
| Tpop4      |                                      |              | $y = 0.0901499(x - 1.66289)^{-0.294461} + 0.181843$ | 0.9914    | $y = 2.6530855(x - 1.6382429)^{0.1952875}$   | 0.9935    |
| Tpop5      |                                      |              | $y = 0.0928235(x - 1.76949)^{-0.239163} + 0.188444$ | 0.9895    | $y = 2.5133317(x - 1.6605321)^{0.18332811}$  | 0.9908    |
| Tpop6      |                                      |              | $y = 0.0893149(x - 1.73897)^{-0.28648} + 0.19956$   | 0.9926    | $y = 2.7008447(x - 1.6814924)^{0.1830269}$   | 0.9928    |
| Tpop7      |                                      |              | $y = 0.0732254(x - 1.76217)^{-0.317876} + 0.177712$ | 0.9874    | $y = 2.642344(x - 1.5613835)^{0.22379012}$   | 0.9961    |

For the bootstrapped data, the best approximation of the population value is the value at 24 bootstrapped individuals. For  $H_e$  and  $A_n$  the approximated bootstrapped values are averages of the values for the eight populations within each data set.

## **Results**

### *Bootstrapped data for $F_{st}$ , $H_e$ and $A_n$ for four datasets*

With increasing sample size, the bootstrapped  $F_{st}$  data approached a characteristic value for each population and the variance around each value declined (Figure 3.1). The maize race populations showed the greatest degree of differentiation, followed by the teosinte populations and the Jala accessions from the genebank. For the normalized  $F_{st}$  values shown in, the differentiated populations (teosintes, maize races and Jala from the genebank) all approached the minimum variance quickly, usually with sampling only four or five individuals. The Jala accessions from farmers' fields showed very low levels of differentiation; they were well below the threshold of 0.05 conventionally used to mildly differentiate out-crossing populations (Hartl and Clark 1997). The populations from the farmers' fields took much longer to reach the minimum variance, and until well beyond the number of individuals shown in Figure 3.2. This likely reflects the fact that the  $F_{st}$  values from farmers' fields were small and the standard deviations were comparatively large (Figure 3.1). Therefore, the variation due to resampling was important and a relatively large sample size was required to approach the minimum variance.

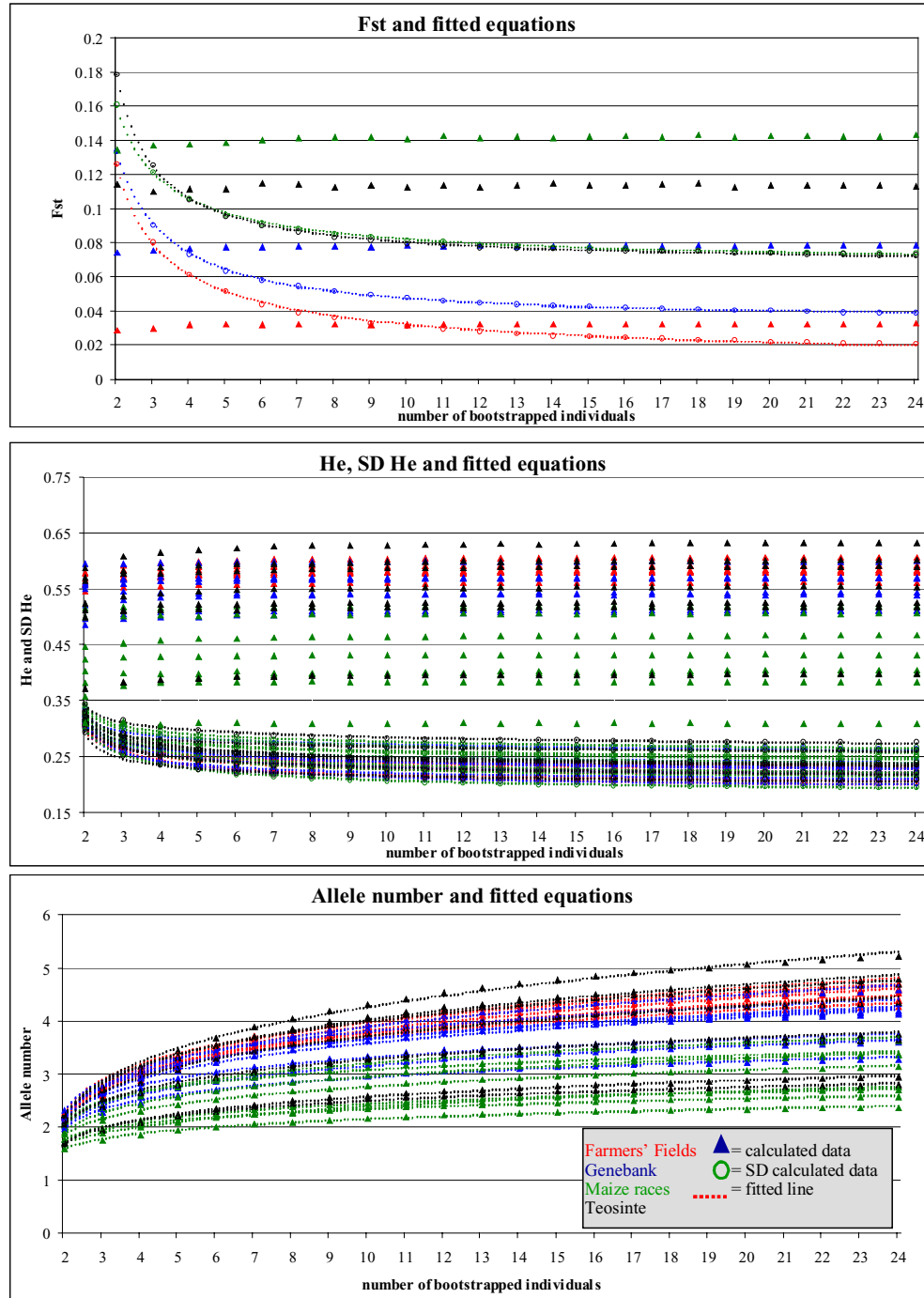
Like  $F_{st}$  values, the gene diversity ( $H_e$ ) values approach the minimum variance quickly (Figure 3.1). The teosinte populations, Jala populations from the genebank and Jala populations from farmers' fields all have similar  $H_e$  values. Only the dataset of maize race populations, with lower gene diversity values, is not intercalated with the others. As will be discussed in the next section, the maize race populations are well differentiated, but not highly diverse. Regardless of the data set,  $H_e$  values of bootstrapped populations as small as two individuals had low levels of variance and were already close to the ultimate value of  $H_e$  (Figure 3.1 and 3.2).

In contrast to  $F_{st}$  and  $H_e$  values, allele numbers ( $A_n$ ) continued to ascend with sample size. Increased numbers of sampled individuals sampled rarer and rarer alleles, and so  $A_n$  continued to increase. However, the rate of acquisition of new alleles slows, as sample size increases. The curves of the maize race populations flattened more quickly than the other sets of populations, indicating that they acquired many alleles with the first few individuals sampled, and then few alleles with subsequent sampling. The representatives of maize races are the least diverse set of populations with the fewest total alleles to sample. The teosinte populations were highly variable: some acquired their alleles quickly and others much more slowly.

*Comparison between bootstrapped and complete datasets*

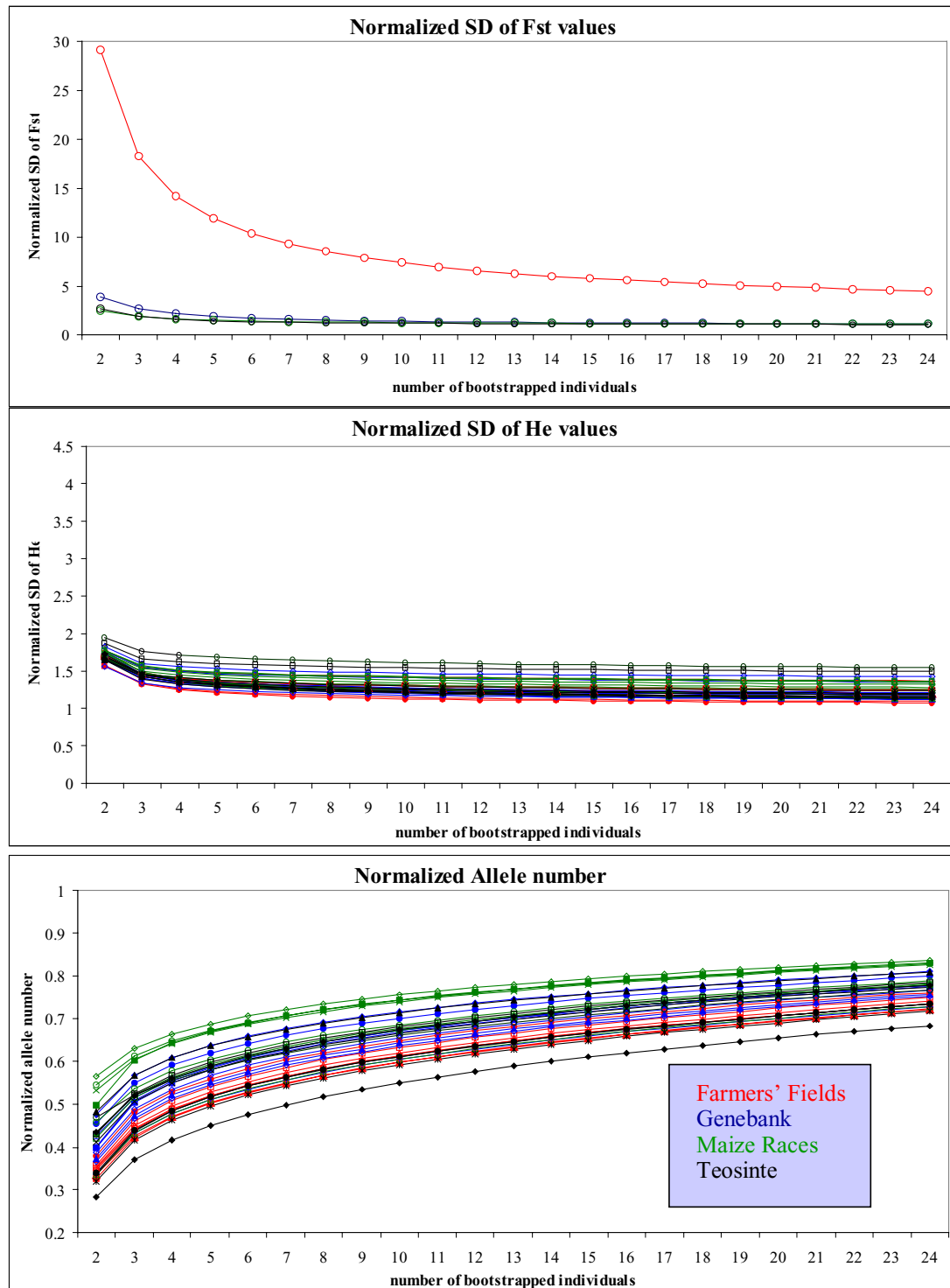
The bootstrapped values, shown in Figure 3.1, clearly approached a single value for each measure, listed in Table 3.4. In nearly all cases, the value for the bootstrapped data was smaller than for the complete data. Bootstrapped data were resampled and therefore some individuals were represented more than once. Thus it is not surprising that a smaller number of alleles were captured and gene diversity was reduced. Populations containing multiple representations of the same individuals would be less differentiated, with the exception of farmers' field populations, which were not genetically distinct.

In farmers' fields, resampling the same individual led to more differentiation, because a resampled individual was more similar to itself than to other members of its population (Table 3.4). Oversampling the same individuals over-represented certain genotypes, and led to shifts in allele frequencies that reduced gene diversity.



**Figure 3.1** Values and standard deviations for  $A_n$ ,  $H_e$  and  $F_{st}$  for the four datasets

Figure 3.1 notes: Farmers' field populations of Jala are shown in red, genebank Jala populations in blue, maize races in green and teosinte populations in black. Solid triangles represent the calculated data points ( $F_{st}$ ,  $H_e$  and  $A_n$ ). Open circles represent the calculated standard deviations of  $F_{st}$ , and  $H_e$ . The dotted lines represent the fitted equations shown in Table 3.3.



**Figure 3.2 Normalized data for  $F_{st}$ ,  $H_e$  and  $A_n$  for the four datasets**

Figure 3.2 notes: Farmers' field populations are represented in red, Jala populations from the genebank in blue, maize race populations in green, teosinte populations in black.

The bootstrapped  $F_{st}$  values were highest for the maize races, and not the teosinte populations, as expected, though both bootstrapped  $F_{st}$  values fell within the confidence interval for the complete data set. The maize races were clearly distinct as shown by the  $F_{st}$  value, but not diverse as shown by their low  $H_e$  values. Individual populations within the maize race set had relatively few alleles, but their allele profiles were distinct from other maize race populations. The teosinte populations, in contrast, were truly more highly differentiated, as shown by their high  $F_{st}$  value for the complete data set. However, the bootstrapped  $F_{st}$  value was much lower than the complete value, while the bootstrapped  $H_e$  and  $A_n$  values were comparable to the other data sets. This indicates that there were substantial differences between the allele profiles of individuals within data sets, so the level of differentiation between the populations was sensitive to which individuals were sampled.

**Table 3.4 Comparison of  $F_{st}$ ,  $H_e$ , and  $A_n$  values between bootstrapped and complete datasets for four categories of populations**

| Data Set               | $F_{st}$ (95% CI)    | $H_e$ | $A_n$ |
|------------------------|----------------------|-------|-------|
| <u>Farmers' Fields</u> |                      |       |       |
| bootstrapped           | 0.03                 | 0.58  | 4.38  |
| complete               | 0.02 (-0.008, 0.045) | 0.60  | 5.03  |
| <u>Genebank</u>        |                      |       |       |
| bootstrapped           | 0.08                 | 0.56  | 3.91  |
| complete               | 0.09 (0.021, 0.148)  | 0.57  | 4.45  |
| <u>Maize races</u>     |                      |       |       |
| bootstrapped           | 0.14                 | 0.42  | 2.93  |
| complete               | 0.20 (0.069, 0.329)  | 0.44  | 3.34  |
| <u>Teosinte</u>        |                      |       |       |
| bootstrapped           | 0.11                 | 0.55  | 3.93  |
| complete               | 0.26 (0.083, 0.441)  | 0.56  | 4.97  |

#### *Issues of sample size*

Very few individuals were needed to capture most of the information in the data sets. To compare  $F_{st}$  and  $H_e$  values between data sets, a threshold was set as the



samples size that would result in a standard deviation that was twice its minimum value (Table 3.5). This threshold did not represent an ideal sample size, but instead an acceptable level of variance around the measures of interest. With the exception of  $F_{st}$  values for farmers' fields, very few individuals (fewer than six) were required to achieve the threshold value. The  $F_{st}$  values for farmers' fields behaved differently because there was very little differentiation between the populations, and a great deal of variation between individuals. Therefore,  $F_{st}$  values for farmers' fields were very small, and highly sensitive to the particular individuals sampled.

The allele number ( $A_n$ ) thresholds were set at 50% of the number of alleles where each new 100 individuals resulted in one allele at an average locus—a threshold beyond which continued sampling does not seem worthwhile. Here, maize race populations reach the threshold most quickly, by three bootstrapped individuals. The Jala accessions from the genebank and teosinte accessions reach the threshold within four individuals, and the farmer's field collections, within five individuals.

**Table 3.5 Sample size thresholds (numbers of individuals) for  $F_{st}$ ,  $H_e$  and  $A_n$  calculations for the four population categories studied**

| Population      | $F_{st}$ | $H_e$ | $A_n$ |
|-----------------|----------|-------|-------|
| Farmers' Fields | 161.77   | 1.79  | 4.19  |
| Genebank        | 4.58     | 1.80  | 3.23  |
| Maize races     | 2.65     | 1.80  | 2.19  |
| Teosinte        | 2.75     | 1.81  | 3.94  |

Table 3.5 notes: For  $F_{st}$  and  $H_e$ , the threshold value represents the number of bootstrapped individuals where standard deviation is twice the standard deviation at the y-offset minimum. For  $A_n$ , the value represents the number of bootstrapped individuals where  $A_n$  is 50% of  $A_n$  when slope is 0.01 (1 new allele gained at an average locus for each 100 new individuals).

## ***Discussion***

*Differences for the different measures:  $F_{st}$ ,  $H_e$  and  $A_n$ .*

Though the different measures ( $F_{st}$ ,  $H_e$  and  $A_n$ ) all reach the designated thresholds with different numbers of bootstrapped individuals, most of the information in the data sets is captured with very few individuals. Gene diversity,  $H_e$ , requires the fewest individuals; the variation around gene diversity values was small with as few as two individuals. Gene diversity values are a function of allele frequencies and are therefore more influenced by common alleles than by rare ones. The common alleles are likely to be sampled with few individuals, and thus it is logical that gene diversity values can be reasonably estimated with few individuals.

Like gene diversity ( $H_e$ ) values, bootstrapped  $F_{st}$  measures of population differentiation rapidly approach the population value. However, the variation around  $F_{st}$  values did not decrease as rapidly. For very different populations like maize races and teosintes, only a few individuals (three) were required to reach the designated thresholds. For more similar populations like different accessions of the maize race Jala from the genebank, a few more individuals (five) were necessary. However, for very similar populations, like Jala from farmers' fields, much higher numbers of individuals (over 160) were required before the variance approached its minimum.

Allele number can theoretically increase infinitely with increasing sample size; practically, to capture all the alleles would require prohibitive numbers of individuals. However, in these data sets most of the alleles were captured with a few individuals (fewer than five). Taken together, these patterns suggest that much of the information and variation was captured within the first few individuals sampled.

*Differences by categories of populations.*

Given the different nature of each data set of populations, we anticipated different optimal sample sizes for each data set. Individuals of the same variety collected from farmers' fields in the same valley and year would be expected to be diverse but not distinct, because of the cross-pollinating nature of maize. However individuals from different races of maize or species of teosinte collected in different places around the world at different times and maintained in genebanks would be expected to be distinct. We expected the distinct populations to require fewer individuals to capture diversity and differentiate between them.

Surprisingly, the different sets of populations reached the thresholds at similar numbers of bootstrapped individuals. The only exception to this was the  $F_{st}$  value for the farmers' fields—more an issue of resolution than one of data set differences. To be sure, more similar populations like farmers' fields required more individuals to achieve the same resolution than did more distinct populations like the maize races or teosintes. However, the differences in numbers of individuals for measures other than  $F_{st}$  were small.

The bootstrapping approach, like any sampling approach, was sensitive to issues of allele distribution within the population sampled. Additionally, because bootstrapping resamples some individuals more than once and neglects to sample others, the approach is particularly sensitive to distribution of alleles between the individuals of a population. For this reason, the bootstrapped teosinte populations appeared less differentiated than the maize races, even though the complete teosinte data set was more differentiated than the complete maize race data set.

*Conclusion*

Clearly, optimal sample size depends on the biological question asked and the measures of interest. Estimating allele frequencies accurately requires thousands of individuals (B-Rao 2001). Capturing rare alleles in a population requires hundreds of individuals (Crossa et al. 1993). Diversity studies, however, need fewer individuals, by one or perhaps two orders of magnitude, to capture most of the diversity and to distinguish between differentiated populations. Obviously the optimal sample size will depend heavily upon the number of markers used. This type of bootstrapping analysis could be extended to evaluate optimal numbers of markers, as well.

This bootstrapping analysis gave a way to visualize rates of change in the data—an approach that should help set thresholds of acceptable variation. Those thresholds can, in turn, help choose reasonable sample sizes. Choosing smaller sample sizes will allow characterization of more populations for the same investment.

## CHAPTER 4 CONSERVATION AND CHANGE: A COMPARISON OF *IN-SITU* AND *EX-SITU* CONSERVATION OF JALA MAIZE GERMPLASM

### *Introduction*

In a world of increasing population, changing cultures, globalization, and migration out of indigenous communities, loss of genetic diversity from farmer's fields in centers of origin is a subject of major concern (FAO 1997; Brush 1999). In order to preserve the genetic resources found in agricultural systems, two strategies are employed. *Ex-situ* conservation captures genetic material, usually seed, and protects it in a genebank, often a cold storage facility. In contrast, *in-situ* conservation allows adaptive evolutionary processes to continue by preserving varieties in farmers' fields, under farmer management. It has been suggested that the two strategies are highly complimentary (Altieri and Merrick 1987; Brush 1991).

*Ex-situ*, or genebank, conservation is frequently misperceived as static, but genetic change can occur (Wood and Lenne 1997). Cold storage of genebank accessions is relatively stable, with very little, if any, genetic change occurring (Roberts 1975). However, seed is only viable for 20-50 years, and eventually must be regenerated or re-grown. In regeneration, a population is planted and grown with controlled pollination to produce new seed true to its parents and as genetically diverse as possible, given the sample size. Theoretical models based on population genetics establish clear guidelines about how to manage regeneration for maximum diversity conservation (Crossa 1989). Similar models (Crossa et al. 1993; 1994) also guide collection methods; a genebank population can never be more diverse than the genetic material originally collected in the field. In the early years of *ex-situ* conservation, the emphasis was on collecting and categorizing as much phenotypically diverse material as possible—conserving varieties, and perhaps less critically, the genetic diversity within them. As our knowledge of molecular genetics of individuals and

populations has grown, the goals and methods of *ex-situ* conservation have shifted subtly to focus on capturing and maintaining maximum allelic diversity within varieties and populations.

*In-situ*, or on-farm, conservation allows crops to evolve dynamically with farmers' needs and the changing environment. However, it is relatively risky, as crops may be lost to a host of environmental and economic influences. Increasingly, studies use interdisciplinary approaches to understand the complex socioeconomic, agronomic and genetic decisions involved in farmers' seed management. [See Bellon et al. (2003), together with Pressoir and Berthaud (2004b) for an integrated approach to maize conservation in Oaxaca, Mexico.]

To date, very little research has compared the effects of genebank conservation to on-farm conservation. In one of two exceptions, Soleri and Smith (1995) found phenotypic differences between Hopi maize varieties conserved on-farm and in a genebank. However, their analysis was confined to morphological characteristics and did not consider molecular genetic data. In the other exception, Parzies et al. (2000) looked at historical collections of barley, an inbred crop, and compared them to the same variety in farmers' fields. There was a clear decline of *ex-situ* genetic diversity with storage time and regeneration. Older genebank populations became more genetically distinct from one another; an effect the authors attributed to genetic drift during genebank management.

Various genetic technologies have been applied to understanding genetic diversity of conserved maize races (for example: Doebley et al. 1985; Rebourg et al. 2001; Matsuoka et al. 2002). Other studies have used molecular markers to characterize genebank populations to understand distribution of diversity, particularly between races and agronomic groups. [For example see Rodriguez et al. (1999) for capsicum, Bhattacharjee et al. (2002) for pearl millet, and Ni et al. (2002) for rice].

Molecular evidence from other studies showed duplication and redundancy in collections. [See Virk et al. (1995) for rice, Phippen et al. (1997) for cabbage, Zeven et al. (1998) for kale, Dean et al. (1999) for sorghum, Verma et al. (1999) for basmati rice, and McGregor et al. (2002) for wild potato.] Several of these in-depth genetic studies turned up evidence of misclassifications within the genebank system [Steiner et al. (1997) for oat collections, Dean et al. (1999) for sorghum, and McGregor et al. (2002) for wild potato].

Recently, several studies used molecular markers of various types to look at farmer-managed diversity at the field level. These studies found surprisingly high levels of genetic diversity in farmers' fields, regardless of breeding system. Inbreeding sorghum populations had lower than expected differentiation between farmers' fields, and higher diversity within populations (Dje et al. 1999). Cassava, a species propagated by cuttings, has surprisingly high levels of genetic diversity within named varieties (Elias et al. 2000). Phenotypically distinct, differently named varieties can not be distinguished at the genetic level in either out-crossing pearl millet populations in Nigeria (Busso et al. 2000) or out-crossing maize populations in Oaxaca (Pressoir and Berthaud 2004a; 2004b). In other words, farmers' variety names do not reflect a discreet genetic unit. Geography may (Busso et al. 2000), or may not (Pressoir and Berthaud 2004b) lead to genetic differentiation between villages in a region for out-crossing species.

The range of diversity in maize can be organized, hierarchically, into maize races, varieties and seed lots. A race will contain many varieties with similar traits. The seed a farmer plants is called a seed lot and is one representation of a variety (Louette 1999). A variety has a set of distinguishable characteristics: e.g., tall stalk, blue grain, floury texture, red silks. "Traditional" varieties are those grown by farmers

for many generations, usually from saved seed. “Improved” varieties are usually produced by a seed company for commercial release.

To ensure comparison of the same variety, past to present, this study focuses on a very unique type of maize called Jala, from Jala, Nayarit, Mexico. Jala is extremely tall (up to 5 m) and bears very long ears (up to 45 cm). It appeared in the literature as early as 1924 (Kempton 1924) and has been the target of a promotional campaign to promote on-farm conservation (Listman and Estrada 1992). Jala is both a variety, planted by farmers, and a unique race of maize.

To conserve varieties like Jala effectively, it is important to first understand genetic dynamics within and between farmers’ fields. Though poorly understood, the field level is where decisions about what to conserve are made, whether they are decisions about farmer selection or decisions about what materials to collect for genebank management. This study investigates the partitioning of diversity in maize and its implications for conservation of agricultural diversity. To do this, we use a set of populations of Jala from farmers’ fields, historical collections of Jala from a genebank, as well as reference populations of other maize races and teosinte. As such, we highlight a specific case study of conservation of a maize race, both in the genebank and in farmers’ fields.

## ***Materials and methods***

### ***Plant material***

The populations studied represented a balanced hierarchical sampling of genetic diversity in maize (Table 4.1). At the bottom of the hierarchy were eight populations of the race Jala, collected from farmers’ fields in the valley of Jala, Nayarit, Mexico. The samples of Jala from farmers’ fields were collected in Jala, Nayarit from farmers at their harvest in November 1999, by taking one kernel from



each of 20-60 ears. (Different sampling methods between farmers' fields and genebank populations had no effect on the effective population size, see Appendix III for details.) Eight populations of the race Jala, collected from 1944 to 1988 in the valley of Jala, were obtained from the International Maize and Wheat Improvement Center (CIMMYT) genebank in Mexico City. These represent depth within one maize race. (For analyses where balanced sampling was not important, three populations from 1988 were used, for a total of nine Jala genebank populations.) Representing breadth across maize races was a set of populations each representing a different maize race. To form the teosinte and maize race subsets, twelve populations were suggested by experts (John Doebley, Major Goodman, and Suketoshi Taba). The twelve populations were compared to microsatellite data Matsuoka et al. (2002) to choose the most diverse eight populations.

For each population assayed, 24 individuals were sampled. Crossa et al (1993) showed that the sample size ( $n$ ) required to retain at least one copy of  $k$  alleles at  $m$  loci with probability  $P$  should be larger than:

$$n > \frac{\log[1 - (P)^{1/m}] - \log(k - 1)}{\log(1 - p)}$$

With 24 individuals, for  $m=22$  loci with  $k=5$  alleles, alleles with frequency  $p > .144$  can be detected with  $P=95\%$  confidence. At  $m=1$  locus, with  $k= 5$  alleles, alleles with frequency  $p > .087$  can be detected with  $P=95\%$  confidence. Rare alleles of  $p=0.05$  can be detected at  $m=1$  locus, with  $k= 5$  alleles, with  $P= 65.8\%$  confidence.

#### *DNA extraction and microsatellite markers*

For each population in Table 4.1, we extracted DNA from 24 individual plants, 10-20 days post-germination, using a CTAB method described by Mitchell et al. (1997). We used 22 fluorescently labeled PCR primer pairs to amplify microsatellite

(SSR) loci described in Table 4.2. Further marker information is currently available from MaizeDB (<http://www.agron.missouri.edu/ssr.html>). In the future, it will be available from MaizeGDB. The 22 SSR markers are widely distributed throughout maize's ten chromosomes with approximately one marker per chromosome arm.

**Table 4.1 Farmers' field Jala, genebank Jala, maize races and teosinte populations studied, with seed source and catalog numbers**

| Population   | Seed Source | Catalog Number |
|--|-------------|----------------|
| <u>Jala Maize from Farmers' Fields</u>               |             |                |
| FF1  | 1999-1-CP   |                |
| FF2  | 1999-2-CU   |                |
| FF3  | 1999-3-CP   |                |
| FF4  | 1999-4-CP   |                |
| FF5  | 1999-5-CU   |                |
| FF6  | 1999-6-CU   |                |
| FF7  | 1999-7-CU   |                |
| FF8  | 1999-8-CP   |                |
| <u>Jala Maize from Genebank (year of collection)</u> |             |                |
| 1944a  | CIMMYT      | CIMMYT 2246    |
| 1944b  | CIMMYT      | CIMMYT 16037   |
| 1951a  | CIMMYT      | CIMMYT 10500   |
| 1951b  | CIMMYT      | CIMMYT 11223   |
| 1952   | CIMMYT      | CIMMYT 16038   |
| 1968   | CIMMYT      | CIMMYT 7068    |
| 1988a  | CIMMYT      | CIMMYT 11417   |
| 1988b  | CIMMYT      | CIMMYT 11420   |
| 1988c  | CIMMYT      | CIMMYT 11421   |
| <u>Representatives of Maize Races</u>                |             |                |
| Chalqueño (Mexico)                                   | USDA        | PI 483613      |
| Coroico (Bolivia)                                    | USDA        | PI 485399      |
| Guirua (Venezuela)                                   | USDA        | PI 445002      |
| Olotón (Guatemala)                                   | CIMMYT      | CIMMYT 2510    |
| Pira (Colombia)                                      | USDA        | CIMMYT 444165  |
| Tepecintle (Guatemala)                               | CIMMYT      | CIMMYT1842     |
| Tuxpeño (Mexico)                                     | USDA        | CIMMYT 488974  |
| Uchuguilla (Peru)                                    | USDA        | PI 571462      |
| <u>Teosinte (<i>Zea</i> spp.)</u>                    |             |                |
| <i>Z. diploperennis</i>                              | Doebly      | m5             |
| <i>Z. mays hueheutenangensis</i>                     | CIMMYT      | CIMMYT 9479    |
| <i>Z. luxurians</i>                                  | CIMMYT      | CIMMYT 9478    |
| <i>Z. mays mexicana a</i>                            | CIMMYT      | CIMMYT 8771    |
| <i>Z. mays mexicana b</i>                            | CIMMYT      | CIMMYT 11369   |
| <i>Z. mays mexicana c</i>                            | CIMMYT      | CIMMYT 11400   |
| <i>Z. mays parviglumis a</i>                         | CIMMYT      | CIMMYT 8756    |
| <i>Z. mays parviglumis b</i>                         | CIMMYT      | CIMMYT 8781    |

**Table 4.2 SSR markers used to evaluate Jala, maize race and teosinte populations, with associated bin location, core repeats, and allele size ranges, PIC values, numbers of alleles and missing data.**

| Marker    | Bin location | Core Repeat | Allele Size Range | PIC value | Allele No. | Missing Data |
|-----------|--------------|-------------|-------------------|-----------|------------|--------------|
| phi056    | 1.01         | GCC         | 230-277           | 0.77      | 14         | 6.2%         |
| phi011    | 1.1          | GCT         | 206-292           | 0.67      | 11         | 13.7%        |
| phi064    | 1.11         | ATCC        | 76-123            | 0.89      | 20         | 4.6%         |
| phi083    | 2.04         | CTAG        | 126-150           | 0.76      | 8          | 2.5%         |
| phi127    | 2.07         | GTCT        | 97-132            | 0.68      | 14         | 2.7%         |
| phi453121 | 3.01         | ACC         | 207-230           | 0.79      | 20         | 0.8%         |
| phi053    | 3.05         | ATGT        | 169-272           | 0.68      | 12         | 10.3%        |
| phi213984 | 4.01         | ACC         | 284-319           | 0.10      | 9          | 1.8%         |
| phi093    | 4.08         | CTAG        | 261-293           | 0.84      | 23         | 2.4%         |
| phi333597 | 5.05         | AAG         | 204-226           | 0.53      | 6          | 1.0%         |
| phi085    | 5.06         | GCGTT       | 224-268           | 0.77      | 13         | 6.7%         |
| phi159819 | 6            | CCG         | 117-145           | 0.66      | 13         | 1.8%         |
| phi389203 | 6.03         | AGC         | 291-315           | 0.61      | 9          | 4.4%         |
| phi034    | 7.02         | CCT         | 123-164           | 0.81      | 14         | 2.8%         |
| phi051    | 7.05         | AGG         | 133-151           | 0.73      | 14         | 3.7%         |
| phi115    | 8.03         | ATAC        | 292-304           | 0.36      | 2          | 4.2%         |
| phi121    | 8.04         | CCG         | 93-107            | 0.15      | 5          | 1.9%         |
| phi015    | 8.08         | TTTG        | 74-118            | 0.77      | 24         | 2.9%         |
| phi033    | 9.02         | CCT         | 230-269           | 0.65      | 14         | 2.5%         |
| phi032    | 9.04         | TTTC        | 230-244           | 0.76      | 14         | 1.6%         |
| phi050    | 10.03        | AAGC        | 78-100            | 0.42      | 8          | 1.9%         |
| phi062    | 10.04        | GAC         | 154-177           | 0.45      | 7          | 1.0%         |
| Total     |              |             |                   | 0.63      | 12.45      | 3.7%         |

### *PCR and genotyping*

Twenty microliter PCR reactions were set up as described by Mitchell et al. (1997), with one modification. To promote the 3' plus A addition to PCR products, extension times at 72C were increased to one hour in the last PCR cycle. Genotyping for 12 markers was done on an ABI 377 sequencer and data was analyzed with the computer program Genotyper (Applied Biosystems). Ten markers were run on an ABI 3700 sequencer and data analyzed with Genemapper (Applied Biosystems).

### *Calculation of population genetic values*

For the calculation of pair-wise  $F_{st}$  values between populations and population specific allele frequencies, the Genepop 3.2 program was used (Raymond and Rousset 1995). GDA 1.1 was used to obtain population description values: allele number ( $A_n$ ), gene diversity ( $H_e$ ), homozygosity ( $H_o$ ), and inbreeding co-efficient ( $f$  or  $F_{is}$ ) (Lewis and Zaykin 2002; Ni et al. 2002). The analysis of molecular variation (AMOVA) was performed using Arlequin 2.0 (Schneider et al. 2000). Linear regressions were performed in Excel.

## **Results**

### *Descriptive statistics*

For the 22 SSR loci surveyed, 274 total alleles were present in the 790 genotypes (Table 4.3). All 22 SSR loci were polymorphic, with the number of alleles per locus ranging from 2 to 24, with an average of 12.5 alleles per locus. The average proportion of missing data across all loci and individuals was 3.7%, with smaller proportions of missing data in maize populations (average 1.5%) and higher proportions in teosinte populations (average 10.5%). Average allele number per locus ranged from a maize race population with 3.1 alleles per locus to a teosinte population with 6.5 alleles per locus. On average, the Jala populations from farmers' fields and the teosintes had the highest number of alleles per locus, followed closely by the Jala populations from the genebank. The maize races had the fewest alleles per locus as well as the lowest gene diversity values. The Jala populations from farmers' fields had the highest gene diversity values. The observed heterozygosity was highest in the Jala populations, both from farmers' fields and from the genebank. The fixation index was highest in the teosinte populations. Nearly all of the unique alleles were found in the teosinte populations.

**Table 4.3 Descriptive information for genotyped populations, organized by category, with category averages and totals, averaged across 22 SSR loci**

| Population                                     | Sample Size | Missing Data (%) | Allele number | Gene diversity ( $H_e$ ) | Observed Heterozygosity ( $H_o$ ) | Fixation index ( $F_{is}$ ) | # unique alleles |
|--|-------------|------------------|---------------|--------------------------|-----------------------------------|-----------------------------|------------------|
| <u>Jala from farmers' fields</u>               |             |                  |               |                          |                                   |                             |                  |
| FF1  | 24          | 0.6%             | 5.0           | 0.60                     | 0.49                              | 0.19                        | 0                |
| FF2  | 24          | 1.0%             | 5.2           | 0.60                     | 0.48                              | 0.21                        | 1                |
| FF3  | 24          | 1.9%             | 5.1           | 0.61                     | 0.54                              | 0.13                        | 1                |
| FF4  | 24          | 2.5%             | 5.3           | 0.60                     | 0.50                              | 0.18                        | 0                |
| FF5  | 24          | 0.7%             | 4.7           | 0.60                     | 0.50                              | 0.17                        | 0                |
| FF6  | 24          | 3.2%             | 5.2           | 0.62                     | 0.49                              | 0.21                        | 1                |
| FF7  | 24          | 2.5%             | 4.8           | 0.61                     | 0.48                              | 0.21                        | 0                |
| FF8  | 24          | 2.5%             | 5.0           | 0.58                     | 0.50                              | 0.15                        | 1                |
| <i>Average</i>                                 |             | 1.8%             | 5.0           | 0.60                     | 0.50                              | 0.18                        | 0.5              |
| <i>Total</i>                                   | 192         |                  | 7.2           | 0.61                     |                                   |                             | 4                |
| <u>Jala from Genebank (year of collection)</u> |             |                  |               |                          |                                   |                             |                  |
| 1944a  | 24          | 0.2%             | 3.6           | 0.52                     | 0.39                              | 0.25                        | 1                |
| 1944b  | 24          | 0.7%             | 4.1           | 0.53                     | 0.46                              | 0.14                        | 0                |
| 1951a  | 24          | 1.5%             | 4.1           | 0.55                     | 0.52                              | 0.05                        | 0                |
| 1951b  | 24          | 0.4%             | 4.6           | 0.59                     | 0.50                              | 0.15                        | 0                |
| 1952   | 24          | 0.0%             | 4.6           | 0.59                     | 0.52                              | 0.11                        | 0                |
| 1968   | 24          | 2.8%             | 4.6           | 0.56                     | 0.46                              | 0.19                        | 0                |
| 1988a  | 24          | 1.7%             | 5.1           | 0.58                     | 0.52                              | 0.11                        | 0                |
| 1988b  | 24          | 1.7%             | 5.1           | 0.62                     | 0.55                              | 0.12                        | 0                |
| 1988c  | 24          | 1.7%             | 5.4           | 0.63                     | 0.53                              | 0.17                        | 0                |
| <i>Average</i>                                 |             | 1.2%             | 4.6           | 0.57                     | 0.49                              | 0.14                        | 0.1              |
| <i>Total</i>                                   | 216         |                  | 7.3           | 0.62                     |                                   |                             | 1                |
| <u>Representatives of Maize Races</u>          |             |                  |               |                          |                                   |                             |                  |
| Chalqueño                                      | 24          | 0.0%             | 3.6           | 0.53                     | 0.48                              | 0.10                        | 0                |
| Coroico  | 24          | 0.6%             | 2.5           | 0.31                     | 0.30                              | 0.04                        | 0                |
| Guirua   | 24          | 2.8%             | 3.7           | 0.48                     | 0.39                              | 0.20                        | 0                |
| Olotón   | 24          | 3.0%             | 3.1           | 0.42                     | 0.37                              | 0.12                        | 1                |
| Pira   | 24          | 1.0%             | 3.5           | 0.44                     | 0.37                              | 0.18                        | 0                |
| Tepecintle                                     | 24          | 1.1%             | 3.9           | 0.52                     | 0.47                              | 0.11                        | 2                |
| Tuxpeño  | 24          | 3.0%             | 3.2           | 0.41                     | 0.36                              | 0.12                        | 0                |
| Uchuguilla                                     | 24          | 1.3%             | 3.2           | 0.37                     | 0.31                              | 0.17                        | 0                |
| <i>Average</i>                                 |             | 1.6%             | 3.3           | 0.44                     | 0.38                              | 0.13                        | 0.4              |
| <i>Total</i>                                   | 192         |                  | 6.7           | 0.55                     |                                   |                             | 3                |
| <u>Teosinte (<i>Zea spp.</i>)</u>              |             |                  |               |                          |                                   |                             |                  |
| <i>Z. diploperennis</i>                        | 24          | 19.3%            | 3.9           | 0.46                     | 0.23                              | 0.50                        | 5                |
| <i>Z. mays hueheutenangensis</i>               | 24          | 9.1%             | 5.0           | 0.57                     | 0.42                              | 0.27                        | 15               |

**Table 4.3 (Continued)**

|                              |     |       |      |      |       |      |     |
|------------------------------|-----|-------|------|------|-------|------|-----|
| <i>Z. luxurians</i>          | 24  | 10.6% | 3.2  | 0.40 | 0.27  | 0.33 | 9   |
| <i>Z. mays mexicana a</i>    | 24  | 5.9%  | 4.6  | 0.55 | 0.44  | 0.21 | 2   |
| <i>Z. mays mexicana b</i>    | 24  | 4.8%  | 5.6  | 0.62 | 0.44  | 0.29 | 4   |
| <i>Z. mays mexicana c</i>    | 22  | 9.3%  | 5.1  | 0.59 | 0.41  | 0.32 | 4   |
| <i>Z. mays parviglumis a</i> | 24  | 22.3% | 6.0  | 0.65 | 0.41  | 0.38 | 5   |
| <i>Z. mays parviglumis b</i> | 24  | 2.7%  | 6.5  | 0.66 | 0.47  | 0.28 | 8   |
| <i>Average</i>               |     | 10.5% | 5.0  | 0.56 | 0.39  | 0.32 | 6.5 |
| <i>Total</i>                 | 190 |       | 11.4 | 0.73 |       |      | 52  |
| <i>Grand total</i>           | 790 |       | 12.5 | 0.66 | 0.446 | 0.34 | 60  |

*Population differentiation*

Jala populations from farmers' fields showed very little differentiation from one another, by the conventional distinctions described by Hartl and Clark (1997) (Figure 4.1). One of the populations collected from farmers' fields in 1999 (FF5) and one of the genebank collections (1988a) were collected from the same farmer. Though the farmer had kept his seed for the interval, without replacing or adding to it, the two populations were no more similar (less differentiated) than other farmers' fields and 1988 genebank populations. Most Jala populations from the genebank showed moderate differentiation from one another, with more differentiation between the older populations (1944 in particular) and less differentiation between more recently collected populations. Interestingly, the most highly differentiated populations were the two collected in 1944.

*Molecular variance*

Nearly all (94%) of the variation in the Jala accessions (farmers' fields and genebank together) was between individuals (Table 4.4). The year accounted for very little of the variation. Fields within a year accounted for more variation, though the value was heavily influenced by the two greatly differentiated fields from 1944.

| Type of sample collected  | Number of regenerations | Pair-wise Fst values from the genebank and from farmers' fields, with history of collection and regeneration |       |       |       |      |      |       |       |      |      |      |      |      |      |      |  |
|---------------------------|-------------------------|--|-------|-------|-------|------|------|-------|-------|------|------|------|------|------|------|------|--|
|                           |                         | Pair-wise Fst values from the genebank and from farmers' fields, with history of collection and regeneration |       |       |       |      |      |       |       |      |      |      |      |      |      |      |  |
|                           |                         | Pair-wise Fst values from the genebank and from farmers' fields, with history of collection and regeneration |       |       |       |      |      |       |       |      |      |      |      |      |      |      |  |
| Jala from genebank        |                         | 1944a  | 1944b | 1951a | 1951b | 1952 | 1968 | 1988a | 1988b | ff1  | ff2  | ff3  | ff4  | ff5  | ff6  | ff7  |  |
| 1944a                     | ?                       | 3  |       |       |       |      |      |       |       |      |      |      |      |      |      |      |  |
| 1944b                     | ?                       | 2  | 0.17  |       |       |      |      |       |       |      |      |      |      |      |      |      |  |
| 1951a                     | ?                       | 3  | 0.10  | 0.13  |       |      |      |       |       |      |      |      |      |      |      |      |  |
| 1951b                     | ?                       | 1  | 0.12  | 0.11  | 0.07  |      |      |       |       |      |      |      |      |      |      |      |  |
| 1952                      | ?                       | 2  | 0.13  | 0.10  | 0.10  | 0.09 |      |       |       |      |      |      |      |      |      |      |  |
| 1968                      | ?                       | 1  | 0.13  | 0.11  | 0.07  | 0.06 | 0.06 |       |       |      |      |      |      |      |      |      |  |
| 1988a                     | 14 ears                 | 2  | 0.10  | 0.09  | 0.09  | 0.08 | 0.06 | 0.07  |       |      |      |      |      |      |      |      |  |
| 1988b                     | 16 ears                 | 2  | 0.11  | 0.07  | 0.07  | 0.06 | 0.04 | 0.05  | 0.03  |      |      |      |      |      |      |      |  |
| Jala from farmers' fields |                         |  |       |       |       |      |      |       |       |      |      |      |      |      |      |      |  |
| ff1                       | 20 ears                 | 0  | 0.09  | 0.09  | 0.05  | 0.05 | 0.06 | 0.05  | 0.04  | 0.02 |      |      |      |      |      |      |  |
| ff2                       | 30 ears                 | 0  | 0.10  | 0.08  | 0.05  | 0.04 | 0.05 | 0.04  | 0.02  | 0.02 | 0.01 |      |      |      |      |      |  |
| ff3                       | 20 ears                 | 0  | 0.11  | 0.10  | 0.08  | 0.06 | 0.05 | 0.04  | 0.05  | 0.03 | 0.02 | 0.02 |      |      |      |      |  |
| ff4                       | 20 ears                 | 0  | 0.10  | 0.09  | 0.04  | 0.04 | 0.06 | 0.05  | 0.04  | 0.02 | 0.01 | 0.00 | 0.02 |      |      |      |  |
| ff5                       | 20 ears                 | 0  | 0.12  | 0.10  | 0.09  | 0.08 | 0.04 | 0.08  | 0.03  | 0.03 | 0.03 | 0.02 | 0.02 | 0.03 |      |      |  |
| ff6                       | 30 ears                 | 0  | 0.08  | 0.09  | 0.04  | 0.05 | 0.03 | 0.05  | 0.03  | 0.02 | 0.02 | 0.00 | 0.02 | 0.01 | 0.02 |      |  |
| ff7                       | 30 ears                 | 0  | 0.11  | 0.11  | 0.05  | 0.06 | 0.06 | 0.06  | 0.05  | 0.03 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 |  |
| ff8                       | 20 ears                 | 0  | 0.12  | 0.11  | 0.04  | 0.06 | 0.06 | 0.05  | 0.04  | 0.03 | 0.01 | 0.02 | 0.03 | 0.01 | 0.04 | 0.03 |  |

**Figure 4.1 Pair-wise Fst values for Jala populations from the genebank and from farmers' fields, with history of collection and regeneration**

Populations that are little differentiated [by convention where  $F_{st} < 0.05$  (Hartl and Clark 1997)] are not highlighted. Moderately differentiated populations ( $F_{st} 0.05$  to  $0.15$ ) are highlighted in light gray. Greatly differentiated populations ( $F_{st} 0.15$  to  $0.25$ ) are highlighted in medium gray (1944a v 1944b only). Very greatly differentiated populations ( $F_{st} > 0.25$ ) do not appear in this figure. Comparisons between populations collected in the same year are in bold typeface.

**Table 4.4 AMOVA for Jala from genebank and farmers' fields, with average and range over 22 SSR loci**

| Source of Variation | d.f. | Sum of squares | Variance components | Percentage of variation | Range of variation over 22 loci |
|---------------------|------|----------------|---------------------|-------------------------|---------------------------------|
| Year                | 5    | 162.22         | 0.125               | 2.07                    | 0% to 9.0%                      |
| Fields within year  | 11   | 196.71         | 0.254               | 4.20                    | 0.9% to 7.2%                    |
| Individuals         | 799  | 4534.44        | 5.675               | 93.73                   | 89.6% to 98.4%                  |
| Total               | 815  | 4893.38        | 6.055               |                         |                                 |

*Regression of allele number and gene diversity with year of collection*

A linear regression analysis (Table 4.5) showed significant relationships between year of collection and both allele number and gene diversity for the Jala populations from the genebank. The older populations had fewer alleles and lower gene diversity, while the newer populations had more alleles and higher gene diversity. The relationships were significant only 4 to 8 loci, but the overall effect across all 22 loci was strong.

*Allele frequency shifts with time*

Though there was substantial variation in allele frequencies from farmers' fields even in the same year, the allele profile of Jala populations remained relatively constant over time. (Figure 4.1 shows examples of the allele frequencies at two randomly selected loci for each of the Jala populations from the genebank and from farmers' fields.) A few rare alleles were detected only in recent populations or in older populations. The most common alleles generally occurred in all populations at relatively similar, high frequencies.



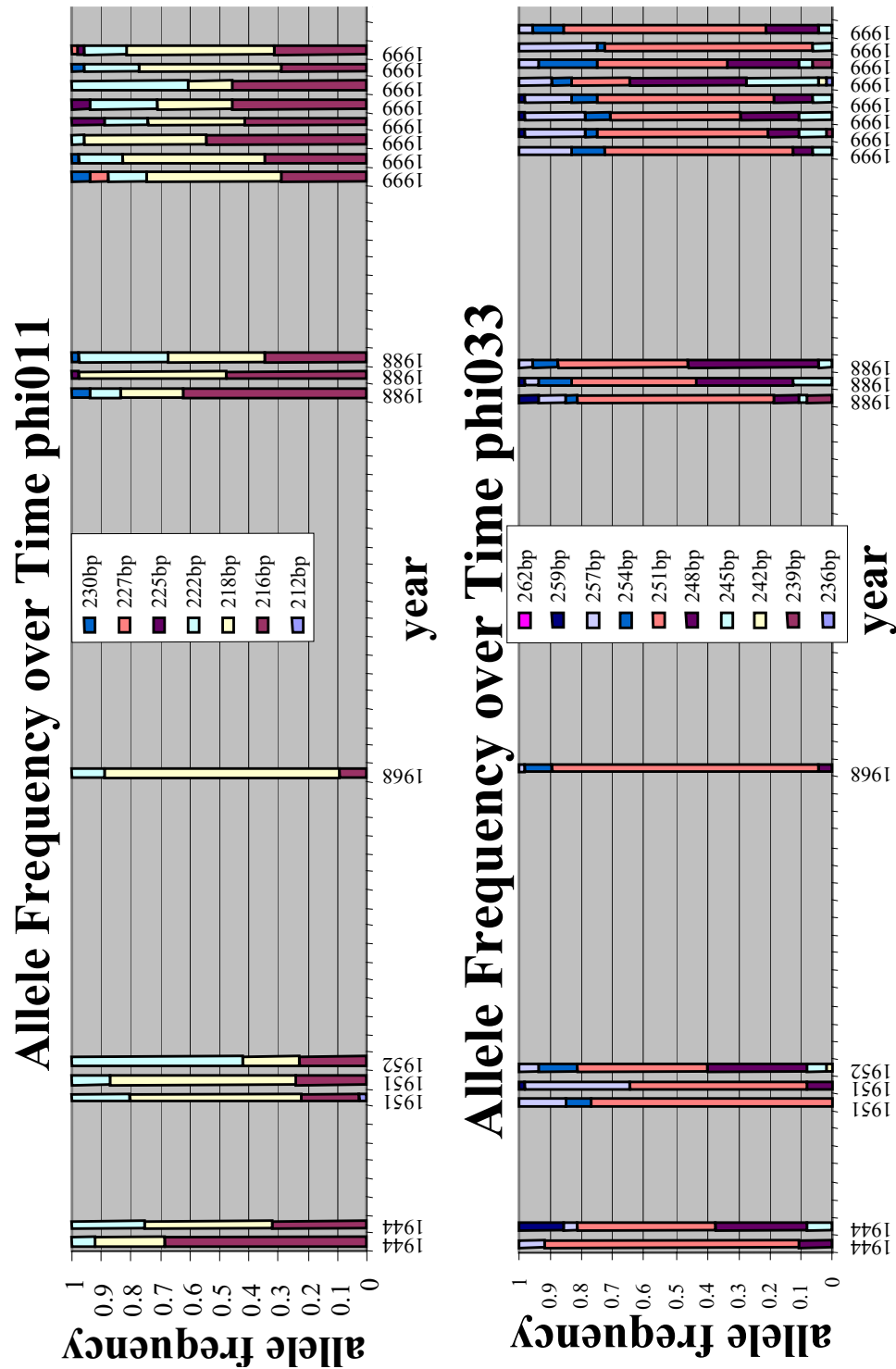
**Table 4.5 Regression of allele number ( $A_n$ ) and gene diversity ( $H_e$ ), by year, for Jala populations from the genebank (1944 to 1988), for 22 SSR loci**

|           | Allele Number  |                | Gene Diversity ( $H_e$ ) |                |
|-----------|----------------|----------------|--------------------------|----------------|
|           | Significance F | R <sup>2</sup> | Significance F           | R <sup>2</sup> |
| overall   | 0.000***       | 0.720          | 0.000***                 | 0.624          |
| phi011    | 0.029*         | 0.280          | 0.080                    | 0.190          |
| phi015    | 0.069          | 0.203          | 0.007**                  | 0.398          |
| phi032    | 0.007**        | 0.397          | 0.074                    | 0.198          |
| phi033    | 0.013*         | 0.349          | 0.094                    | 0.176          |
| phi034    | 0.123          | 0.151          | 0.001***                 | 0.539          |
| phi050    | 0.007**        | 0.397          | 0.226                    | 0.096          |
| phi051    | 0.097          | 0.173          | 0.054                    | 0.226          |
| phi053    | 0.192          | 0.110          | 0.645                    | 0.014          |
| phi056    | 0.158          | 0.128          | 0.332                    | 0.063          |
| phi062    | 0.564          | 0.023          | 0.728                    | 0.008          |
| phi064    | 0.032*         | 0.272          | 0.025*                   | 0.292          |
| phi083    | 0.852          | 0.002          | 0.277                    | 0.078          |
| phi085    | 0.003**        | 0.450          | 0.004**                  | 0.442          |
| phi093    | 0.001***       | 0.542          | 0.246                    | 0.088          |
| phi115    | n/a            | 1.000          | 0.128                    | 0.148          |
| phi121    | 0.352          | 0.058          | 0.784                    | 0.005          |
| phi127    | 0.885          | 0.001          | 0.623                    | 0.017          |
| phi159819 | 0.174          | 0.120          | 0.618                    | 0.000          |
| phi213984 | 0.219          | 0.099          | 0.256                    | 0.085          |
| phi333597 | 0.817          | 0.001          | 0.277                    | 0.078          |
| phi389203 | 0.044*         | 0.019          | 0.107                    | 0.588          |
| phi453121 | 0.256          | 0.085          | 0.015*                   | 0.334          |

\* indicates significance at the 0.05 level.

\*\* indicates significance at the 0.01 level.

\*\*\* indicates significance at the 0.001 level.



**Figure 4.2 Allele frequency profiles over time for 2 randomly selected loci: phi011 and phi033**  
Figure 4.2 notes: In each drawing, the alleles are stacked in proportion to their frequencies, organized by size, and measured in base pairs.

## ***Discussion***

### *Jala populations compared to teosinte and maize race populations*

Most likely, maize was domesticated from teosinte (Doebley 1990; Matsuoka et al. 2002). The strong selection pressures an organism would experience during domestication lead to a reduction in genetic diversity between the domesticate (maize) and the wild relative (teosinte). Therefore it was not surprising that the teosinte populations have 90% of the total alleles in the study and the greatest gene diversity (Table 4.3). The maize populations had only about 55% of the total alleles (6.7 to 7.3 out of a total of 12.4 alleles). Nearly all of the unique alleles (52 out of 60, or 87%) were specific to teosinte populations. Though the teosinte populations were the most diverse as a category, the individual populations were not substantially more diverse than the Jala populations from the farmers' fields in terms of allele number and gene diversity. Average number of alleles is about 5 for both categories and average gene diversity was about 0.60 for both categories.

The populations of the maize race category were distinct from one another, with high  $F_{st}$  values (Chapter 3). However, with the lowest number of alleles and gene diversity of any category, the maize race populations were not highly diverse individually, compared to the diversity found in Jala, at least for at the markers used for this study. Therefore the maize races had higher identity between the individuals of the population, and strong differences between the populations, but not large numbers of alleles or high gene diversity within a population (see Chapter 3 for a discussion). The maize race populations are less diverse even than the oldest Jala genebank populations, to which their collection and management might have been similar. Whether compared to the three least diverse Jala populations (1944a, 1944b,

1951a) or to the five oldest accessions (1944 to 1952), t-tests were significant [p-values for gene diversity ( $H_e$ ) 0.002 for 1944-1952, and  $p=0.007$  for three least diverse populations; p-values for allele number 0.007 for 1944-1952 and  $p=0.06$  for three least diverse populations], indicating the maize races are indeed less diverse than even the least diverse Jala populations.

There were two alternate hypotheses to explain the lack of maize race population diversity relative to Jala populations: 1) the Jala populations may be unusually diverse, or 2) the maize race populations may be less diverse than the Jala race populations because of the way they were collected, maintained or utilized. The first hypothesis, that the Jala populations are unusually diverse, seems unlikely; most maize races are composed of many varieties and come from wide geographic areas, but the Jala race is just one variety and comes from just one valley. Therefore one would expect Jala to be less diverse than other races. However, the maize race populations used in this study sampled only one accession as a representation of the race, and therefore may represent only a narrow slice of the race's diversity. Sampling of multiple populations of other maize races would better address the first hypothesis. Though it seems likely that genebank history may play a role, without documentation about how each of the maize race populations were collected, managed and regenerated it is impossible to test the latter hypothesis.

#### *Variation between individuals within farmers' fields*

To understand conservation of crop genetic diversity, first we must understand something about genetic dynamics at the level of farmer's fields—where decisions were made about conservation. Pressoir and Berthaud (2004b) found very low differentiation between varieties in the central valleys of Oaxaca, Mexico, both within and between villages; a condition they attributed to both pollen and seed flow within

and between communities. Similarly, the  $F_{st}$  data for Jala accessions in farmers' fields in the valley of Jala also shows very little differentiation between farmers' fields (Figure 4.1). The vast majority of the variation in the data set can be explained by differences between individuals within fields, rather than by differences between fields or years of collection (Table 4.4). Together, these two pieces of information suggest that farmers' fields planted with traditional varieties like Jala are dynamic, diverse populations with high levels of variation between individuals, and high levels of genetic interchange. These phenomena might be explained by pollen transfer during overlapping flowering times, as well as cultural and seed management practices.

Today's similarity between different farmers' Jala could be attributed to 1) the limited area of Jala planted resulting in bottlenecks and reductions genetic diversity, or 2) pollen or seed mixing. Though there are very limited areas of Jala planted today (Chapter 2), Jala itself was highly diverse (Table 4.3). This genetic situation is inconsistent with reductions in genetic diversity associated with bottlenecks from limited planting of materials with a reduced genetic base—thus the first hypothesis seems unlikely. The Jala farmers surveyed were planting the traditional variety in small areas (average 1.14 ha, range 0.06-3.00 hectares), widely scattered across the valley (Chapter 2). Gene transfer into advanced generations of improved varieties indicated strong pollen flow in the valley (Chapter 5). Farmers' accounts described long periods of time managing the same Jala seed lot without replacement or supplementation (average of 43 years, Chapter 2). In contrast to farmers in other areas of Mexico who have a culture of “refreshing” seed through purchase or exchange every few years (Louette et al. 1997; Aguirre-Gomez et al. 2000), Jala farmers took pride in managing the same seed lot for generations. Therefore, it seems likely that pollen mixing was primarily responsible for the lack of genetic differentiation between Jala farmers' seed lots.

*Comparison of Jala in genebank, over time*

Based upon the evidence from farmers' fields, one might hypothesize that the valley functions as a united genepool, and predict that Jala populations collected in the same year would be more like one another than Jala collected in other years. The pair-wise  $F_{st}$  values in Figure 4.1 show the two populations collected in 1988 were the two most similar genebank populations ( $F_{st} = 0.03$ ). The two populations collected in 1951 were moderately differentiated ( $F_{st} = 0.07$ ). Surprisingly, however, the two populations collected in 1944 were the two most highly differentiated populations studied.

Genebank management, particularly during the regeneration process, is potentially a source of genetic change for accessions in the system, due to bottlenecks, inbreeding, random genetic drift, and unintentional mixing (Crossa et al. 1994). The 1988, 1951 and 1944 accessions were each regenerated two or three times (Figure 4.1). Thus the number of regenerations alone could not explain the differences between the older populations. Today genebank regenerations are carefully managed to maintain and minimize genetic drift and bottlenecks (Crossa 1989; Crossa et al. 1993; Crossa et al. 1994). However, very little is known about how accessions were regenerated in the past. The number of regenerations, as shown in Figure 4.1, did not appear to be a good predictor of allele number, gene diversity, homozygosity or year of collection (regressions all not significant, data not shown). So perhaps the regeneration method rather than the number of regenerations was important. Furthermore, we know nothing about how most of the genebank populations were collected. Even referring to the original notebooks kept by the collectors on their journeys does not shed light on whether they were collecting one ear or many, a handful of seeds or a bagful.

More recently collected populations are more diverse than older populations, as shown by regression analysis in Table 4.5. These differences could reflect inherent differences in Jala field populations in the 1940s and 1950s or the method of collection or regeneration. Though we had no way of ruling out inherent differences between fields in the past, Jala's history makes it seem an unlikely option. In the 1940s and 1950s, Jala was planted on nearly all the land in the Jala valley (Chapter 2). Today it represents only 5% of maize area in the valley. Farmers appear to manage their seed today in the same way they did in the past, keeping seed lots for generations without switching, buying or replenishing seed from neighbors. With more Jala planted in the past, there would have been more Jala pollen and more mixing of Jala, leading to genetic similarity, not genetic distinction (Chapter 2). Additionally, there would have been more people growing Jala on larger areas, another factor likely to lead to large, heterogeneous, indistinct populations, rather than the very distinct populations with reduced diversity we see in 1944. Therefore, it seems likely that method of collection or method of regeneration affected these populations to make them more distinct from one another and less diverse than the populations they were collected from in the field.

*Jala in farmers' fields versus Jala in the genebank*

The populations currently in farmers' fields showed little differentiation from the most recently collected Jala genebank populations from 1988 (Figure 4.1). Populations from the 1950s and 1960s showed some relationships with moderate levels of differentiation and some with little differentiation. The populations from 1944 were consistently the most differentiated from other populations, a situation that would be consistent with changing patterns of cultivation in the valley (discussed in Chapter 2), or non-representative sampling, discussed above. In 1944, Jala was nearly

the only variety grown in the valley. In the 1950s and 1960s, other varieties began to arrive from other parts of Mexico. By 1988, the maize environment looked much like it does today, with many improved varieties, many traditional varieties from other parts of Mexico and reduced areas of the traditional Jala variety. However, the fact that the two 1944 populations were even more different from one another than they were from anything else suggests again that collection or management history are likely to be responsible for much of the difference.

The mild temporal differences in the Jala genebank populations raise two important issues: 1) loss of alleles that were present in the 1940s and 1950s, and 2) “contamination” by new alleles from new varieties in the valley. However, the common alleles of the past remained the common alleles in the present (Figure 4.2). There was a high level of variation in allele frequency between fields sampled in the same year. However, the variation in allele frequencies over years did not appear to greatly exceed this background level of field-to-field variation. There are a few rare alleles present in the older populations that are not present in more modern populations. Also there are rare alleles present in modern populations that were not present in older populations. As these are rare alleles, however, it was impossible to know if their appearance or disappearance reflects limited sampling or true changes in the population. There do not appear to be systematic shifts in the major alleles at the loci studied.

#### *Conservation target*

Collection efforts should focus on alleles that are common, but locally distributed as important for plant breeding, because of local adaptation to biotic and abiotic stresses (Marshall and Brown 1975). There is no need to focus conservation efforts on common, widespread alleles because they will be conserved in varieties



collected at many locations. The numbers of individuals required to target rare alleles quickly become prohibitive [see Crossa et al. (1993) for calculations of sample size and probabilities]. Therefore, it makes sense to target pockets of diversity –alleles common in local areas, but rare elsewhere—to capture maximum genetic diversity with minimum efforts. By using the populations found in farmers’ fields in Jala as an example of locally distributed alleles and the alleles found in the maize race populations as a proxy for widespread alleles, we can quantify the target for conservation (Figure 4.3). These results are limited by the precision of our sampling method: rare alleles, given these sampling methods, are only captured with 66% confidence, the common alleles, the target of interest, will be captured with higher confidence. Therefore, this estimate is likely to represent an upper limit of the local common alleles of interest.

| Allele<br>Distribution | Allele frequency    |                     | total         |
|------------------------|---------------------|---------------------|---------------|
|                        | common              | rare                |               |
|                        | (allele freq > .05) | (allele freq < .05) |               |
| widespread             | 44.4%               | 29.2%               | 73.6%         |
| local                  | 10.1%               | 16.3%               | 26.4%         |
| total                  | 54.5%               | 45.5%               | n=178 alleles |

**Figure 4.3 Evaluation of the target for conservation: common and rare<sup>†</sup> allele frequencies compared to widespread<sup>‡</sup> and local<sup>§</sup> distribution for Jala data**

<sup>†</sup>The threshold for rare alleles is set at 5% (Marshall 1975).

<sup>‡</sup>Widespread alleles are approximated by the set of populations representing races.

<sup>§</sup>Local alleles are approximated by the Jala populations from farmers’ fields.

Of the 178 alleles sampled in farmers’ fields and populations representing races of maize, using these 22 microsatellite markers, about three-quarters were widespread while about one quarter were locally distributed (Figure 4.3). About 46% of the alleles were rare, defined as having a frequency below 5% (Marshall and Brown 1975). Fifty-five percent of alleles sampled in this study were common alleles. Most

common alleles (70.1%) were present in both the races and Jala, showing that the common alleles of maize are often widely dispersed, even in local pockets of diversity like Jala. The target for conservation—the local, common alleles—constituted less than 10% of the alleles in this study. As a by-product of conservation targeted to common locally-distributed alleles, many of the rare, local alleles would be also captured, as well as widespread alleles, whether common or rare.

### *Conservation implications*

In general, maize contained only about 55% to 60% of the number of alleles found in the teosinte populations (Table 4.3). Wild relatives can be important sources of diversity and economically important traits in cultivated species (Tanksley and McCouch 1997). Therefore it is critical to maintain good collections of wild relatives.

The Jala maize found in farmers' fields was highly diverse at the neutral markers assayed: one farmers' field contains 40% of the alleles found in the entire data set. A similar study focused on markers associated with functional alleles—a future direction for research—might show different sorts of genetic relationships due to selection history. Jala in one farmer's field showed little differentiation from the Jala found in another farmer's fields. Recent collections, like the 1988 populations, for which we have documentation of collection and regeneration methods, also support the notion of similarity between fields in a given year. This implies, at least for Jala maize, that the valley represents the unit of conservation and not the individual farmer. Therefore, it is far more important to sample many individuals from one field than to sample a few individuals from each of many fields in the valley.

There were some small to moderate differences between the Jala populations in the genebank, which parallel changes in planting patterns in the valley, but also may reflect genebank collection and management. The two populations from 1944 that

were so strongly differentiated from one another hint that collection and early management of older accessions may contribute to genetic narrowing and drift. It is important to recall that in the early days of genebank collection, people were focused on collecting visibly different varieties without our current focus on the molecular and population genetics underlying varietal characteristics.

The common alleles in Jala populations from 1944 to the present have remained the same, implying that Jala has been consistently maintained, at least for these loci, both in the genebank and in farmers' fields. Though farmers in Jala today plant much smaller areas of the variety Jala than in the past, the areas appear to be sufficient to not have caused bottlenecks and genetic narrowing in the field populations studied.

## CHAPTER 5 GENETIC CHANGE IN “*CRIOULLO*” IMPROVED VARIETY SEED RECYCLED BY FARMERS

### *Introduction*

Developing world farmers often save seed derived from improved hybrid varieties for many generations, calling them “*criollo*” and managing them as traditional varieties. Published studies refer to these varieties as “*creole*” (Bellon and Risopoulous 2001), “recycled” (Morris et al. 1999), and “rusticated” (Prain 1993). Though 21 to 54% of developing-world maize farmers surveyed in Central America, Guatemala, Nicaragua, India and Malawi are growing *criollo* varieties, little work has explored the reasons farmers keep them, or the genetic effects of farmers’ seed management. [See Morris et al. (1999) for a review of studies.] Farmers in Jala gave two explanations for keeping *criollo* varieties: 1) hybrid seed is expensive and there is still benefit from growing improved materials in subsequent seasons, and 2) after several years of growing in the valley, local adaptation leads to an improved, improved variety.

These two *criollo* stories are explanations of phenomena that are a result of plant genetics. A plant adapting to local conditions would show some improved characteristics, such as increased ability to emerge from deep soil, higher yield, larger ears, increased pest or disease resistance. The adaptation would be driven by farmer selection combined with natural selection. On the other hand, hybrid varieties begin to segregate in subsequent generations, and thus variation increases. This leads to expected yield declines and increased variation in subsequent generations.

*Criollo* varieties provide a case where one can document evidence of gene flow and perhaps even rates of gene flow in an agricultural ecosystem. Because a hybrid has a known genetic constitution, one can predict what the allelic profiles of its progeny should be. Deviations from the profile can be attributed to gene flow, and

thus the gene flow will be easily detected. If the genes flowing in are from traditional varieties, these *criollos* could be an important, overlooked, reservoir of genetic diversity from traditional varieties. Moreover, gene flow from these improved varieties is of concern today, in light of the growing prevalence of genetically engineered maize (Bellon and Berthaud 2004)

The decision to replant hybrid varieties into *criollo* varieties could be a purely economic one (Heisey et al. 1997). Farmers will buy new seed only when they expect the benefits of new seed to outweigh the costs. If farmers feel they can gain the benefits from the improved variety without having to buy seed every year, it makes sense economically to recycle seed into *criollo* varieties. *Criollo* varieties are sometimes seen as a “middle way” by which the poor can access technology and adapt varieties to their local conditions without the cost of buying seed every year (Bellon and Risopoulou 2001).

Improved maize varieties, a term used in this study to denote commercial varieties improved by professional plant breeders, are often hybrid varieties. Hybrids are produced by crossing two highly inbred lines that are homozygous, or nearly so, due to multiple generations of inbreeding. The hybrid, or  $F_1$ , generation has a boost in yield and vigor, known as hybrid vigor or heterosis. The subsequent generation, the  $F_2$  generation, is formed from crosses between members of the  $F_1$  generation. Genetic theory predicts yield declines in the segregating generations ( $F_2$ ,  $F_3$ , etc.) due to inbreeding. The original formulas for yield loss in the  $F_2$  and subsequent generations were derived by Wright (1922), with modifications by Mather and Jinks (1971) and Hallauer and Miranda (1981). These formulas were then applied by Morris et al. (1999) in a simulation exercise to examine the yield effects of recycling *criollo* varieties. All hybrid varieties show their largest yield losses between the  $F_1$  and  $F_2$  generations, both in the modeling exercise and in field trials (Ramirez-Vallejo et al.

1986; Espinosa-Calderon et al. 1990; Espinosa-Calderon et al. 1993). In field trials, there were not statistically significant yield changes between the F<sub>2</sub> and the F<sub>6</sub> generations and Ramirez Vallejo et al. (1986) emphasize that with positive selection pressures, like those exerted by farmers, yields indeed might have increased.

The town of Jala in Nayarit, Mexico, is known for its remarkably tall local maize variety. The traditional Jala variety is only grown on 5% of the maize area in the valley, while an improved variety, *Dekalb* D880, locally known as “880” has risen to prominence and accounts for nearly half the valley’s maize area. This three-way cross hybrid (Morris 2000, personal communication), is preferred because it produces prolific husks for the local tamale husk packaging cottage industry. Not surprisingly, most farmers growing *criollo* varieties are growing a *criollo* version of the popular *Dekalb* D880.

In Jala, traditional varieties are often planted in April, before the rains arrive. Improved varieties are planted in June with the onset of the rains (see Chapter 2). The improved and traditional plants, in spite of their different planting dates, flower within the same window of time in late July and early August. Pollen can travel great distances (Luna et al. 2001) and therefore there is tremendous potential for mixing between the varieties, especially because farmers often have many small plots in different locations. Plots in Jala range from 0.06 to about 2 hectares. The small plot size and fragmented landscape are relics of the *ejido* system and lead to high levels of gene flow in farmers’ fields across Mexico (Bellon and Brush 1994; Louette et al. 1997) Farmers counteract this gene flow with the selection they exert at the end of the season. However, many of the ear characteristics upon which farmers select are pre-determined by the maternal plant, and thus there is a one year delay in farmers’ ability to identify off-type plants. Thus, even with farmer selection, traditional and *criollo* varieties are variable.

The objectives of this study were to understand whether farmers keep advanced generations of *criollo* seed because of economic motivation, driven by the cost of new seed, or because of ecological motivation, driven by the process of local adaptation. In this study, we examine a combination of genetic and social evidence to better understand why farmers grow *criollo* varieties, how *criollo* varieties change genetically and physically under farmer management, and whether *criollo* varieties are an important reservoir of genetic diversity from traditional varieties.

### ***Materials and methods***

#### *Genetic study: populations, markers and statistical analysis*

For this preliminary study, microsatellite markers were used to genotype 24 individuals from each of three different populations: an improved variety, *Dekalb* 880, D880 after one year in farmers' fields (D880 F<sub>2</sub>) and D880 after seven years under farmer management (D880 F<sub>8</sub>). The samples of these populations, as well as additional materials listed in Appendix III but not analyzed, were collected in Jala, Nayarit from farmers at their harvest in November 2000 and 2001, by taking one kernel from each of 30 ears. The F<sub>2</sub> and F<sub>3</sub> materials not analyzed in this study were managed and selected by collaborating farmers. Wherever possible, both a selected and an unselected sample of seed were obtained. Additional advanced generation materials were identified through a series of formal interviews and associated seed collections conducted in the summer of 2001.

For each population, DNA was extracted from 24 individual plants using a CTAB method described in Mitchell et al. (1997). For PCR, 22 fluorescently labeled primer pairs were used to amplify 22 widely-distributed microsatellite (SSR) loci described in Chapter 4. Further information about PCR, genetic analysis and statistical methods is available in Chapter 4.

Sample size was not chosen to optimize precision of allele frequencies, though they are used in analysis. Estimating allele frequencies with two decimal accuracy requires an impractical one thousand to ten thousand individuals, depending on the loci and allele frequencies sampled (B-Rao 2001). Crossa et al. (1993) showed that the sample size ( $n$ ) required to retain at least one copy of  $k$  alleles at  $m$  loci with probability  $P$  should be larger than:

$$n > \frac{\log[1 - (P)^{1/m}] - \log(k - 1)}{\log(1 - p)}$$

With 24 individuals, for  $m=22$  loci with  $k=5$  alleles, alleles with frequency  $p > 0.144$  can be detected with  $P=95\%$  confidence. At  $m=1$  locus, with  $k=5$  alleles, alleles with frequency  $p > 0.087$  can be detected with  $P=95\%$  confidence. Rare alleles of  $p=0.05$  can be detected at  $m=1$  locus, with  $k=5$  alleles, with  $P=65.8\%$  confidence.

### *Social Survey*

In summer 2001, a formal, random survey was implemented. Lists of farmers from the *ejido* and *pequeño propietario* associations were obtained for both Jala and Jomulco. The lists were revised with the help of key informants to remove repetitions and names of people who were no longer farming or who were otherwise unavailable. In total, there were 444 farmers in the valley, of whom I randomly selected 79 to interview (17.8%). (See Chapter 2 for further details.)

## **Results and discussion**

### *Results from social survey*

A majority (55%) of the farmers surveyed had experience with *criollo* varieties. Of them, most (61%) felt that reusing improved seed was a bad idea—that it declines in yield, loses its strength, and is not the same after the first year. Several farmers felt that they could keep seed for two to four years without declines in yield or



quality. Many of them mentioned the cost of a bag of seed as a factor in their decision. A few vocal farmers (7%) told a story of local adaptation—a variety is good in the first year, poorer in the second year, and then gets better and better with time.

Yield was obviously an important factor in a farmer's decision about whether or not to grow a variety. To assess yield, subjective yield descriptions based on farmer recollection were solicited for a good year, a bad year and a normal year. The data is not based on trials and is intended only as a relative indicator of yield, as it is subject to farmers' perceptions and in no way controls for differences between fields or between farmer management strategies (Bellon and Risopoulos 2001). Table 2.2 shows a subset of the subjective yields (see Chapter 3 for complete data and explanation). Regardless of the type of year, *criollo* varieties are reported to have lower yields than both improved and the best traditional varieties (Table 2.2). Thus farmers do not perceive that they are gaining a yield advantage by keeping D880 *criollos*.

**Table 5.1 Subjective response maize yields for different varieties and categories of maize in good, normal and bad production years**

|                      | ---Good year--- |      |     | ---Normal year--- |      |     | -----Bad year---- |      |     |
|----------------------|-----------------|------|-----|-------------------|------|-----|-------------------|------|-----|
|                      | T/ha            |      |     | T/ha              |      |     | T/ha              |      |     |
|                      | average         | SD   | (n) | average           | SD   | (n) | average           | SD   | (n) |
| <u>New Improved</u>  | 5.15            | 1.74 | 23  | 3.52              | 1.22 | 20  | 2.29              | 1.38 | 21  |
| 880 new              | 5.40            | 1.70 | 18  | 3.77              | 1.07 | 16  | 2.37              | 1.50 | 17  |
| Other new            | 4.22            | 1.72 | 5   | 2.54              | 1.42 | 4   | 1.94              | 0.72 | 4   |
| <u>Criollo</u>       | 3.32            | 1.67 | 16  | 2.43              | 1.19 | 18  | 1.56              | 0.79 | 17  |
| 880 <i>criollo</i>   | 3.62            | 1.52 | 7   | 2.63              | 1.13 | 10  | 1.68              | 0.72 | 9   |
| Other <i>criollo</i> | 3.09            | 1.83 | 6   | 2.17              | 1.19 | 5   | 1.42              | 0.79 | 5   |
| <u>Traditional</u>   | 3.44            | 1.30 | 32  | 2.68              | 1.12 | 32  | 1.67              | 0.97 | 29  |
| SanJuaneño           | 4.30            | 1.35 | 10  | 3.39              | 1.07 | 11  | 2.19              | 0.86 | 9   |
| Jala                 | 3.50            | 1.48 | 6   | 2.61              | 1.19 | 5   | 1.75              | 1.44 | 5   |
| Tampiqueño           | 3.25            | 1.21 | 6   | 2.74              | 1.26 | 6   | 1.49              | 1.14 | 5   |

Subjective yield data (Bellon and Risopoulos 2001) consists of farmers' answers to questions about yield. It is not objective data from controlled yield trial. It is intended only as a relative indicator of yield, as it is subject to farmers' perceptions and in no way controls for differences between fields or between farmers management strategies.

Using the yield estimates from Table 2.2, the trade-off farmers make when they keep *criollo* variety seed can be roughly quantified. On a straight cost basis, admittedly a simplistic analysis, the investment appeared to pay for itself. A bag of improved seed for one hectare cost 650 pesos in 2001. The yield difference between new D880 and *criollo* D880 was 1.1T/ha. Farmers reported 1,200 pesos/T as the local price for maize. Therefore, the benefit of new seed in terms of yield off-sets the direct cost of purchase. However, this does not account for risk (particularly of planting failure, a common problem with machine planted maize), availability of cash, cost of credit or other economic factors.

To understand the biology behind farmers' reasoning about reusing *criollo* varieties, farmers were surveyed about changes in a series of plant characteristics shown in Table 5.2. Farmers' responses contained some elements of both local adaptation and hybrid segregation strategies. Most respondents felt with *criolloization*, their plants became taller and slower to mature—both characteristics associated with the traditional, local varieties. Surprisingly though, most farmers did not think that ear size was increasing, another trait associated with the traditional varieties. Farmers consistently reported a decrease in yield and an increase in variation of both plant height and ear size. However, most did not perceive any change in mean ear size or variation in ear height or germination time.

#### *Results from genetic study*

The commercial D880 variety is a three-way cross hybrid. Therefore, the average of 2.5 alleles per locus, a low number compared to traditional varieties, was not surprising. The allele number in advanced, creolized generations rose, with most of the increase occurring in the first year (Table 5.3). Gene diversity also increased.

The observed heterozygosity appeared to be stable, around 0.35 in all populations. The fixation index, also known as the inbreeding co-efficient, was understandably negative in the commercial D880F<sub>1</sub> hybrid and increased in subsequent generations.

**Table 5.2 Farmer descriptions of phenotypic changes with *criollo* varieties**

|                 | yield        | plant<br>height | ear<br>size  | cane<br>size | drought<br>resistance | pest<br>resistance | days to<br>maturity | variation<br>in ear<br>size | variation<br>in plant<br>height | variation<br>in germin-<br>ation time | variation<br>in ear<br>height |
|-----------------|--------------|-----------------|--------------|--------------|-----------------------|--------------------|---------------------|-----------------------------|---------------------------------|---------------------------------------|-------------------------------|
| 1=increases     | 13.5%        | <u>51.4%</u>    | 24.3%        | 26.5%        | 17.9%                 | 15.8%              | <u>60.0%</u>        | <b>44.1%</b>                | <b>41.2%</b>                    | 8.8%                                  | 16.7%                         |
| 2=stays<br>same | 24.3%        | 24.3%           | 27.0%        | 44.1%        | 39.3%                 | 52.6%              | 40.0%               | 50.0%                       | 58.8%                           | 76.5%                                 | 66.7%                         |
| 3=decreases     | <b>62.2%</b> | 24.3%           | <i>48.6%</i> | 29.4%        | 42.9%                 | 31.6%              | 0.0%                | 5.9%                        | 0.0%                            | 14.7%                                 | 16.7%                         |
| n               | 37           | 37              | 37           | 34           | 28                    | 19                 | 5                   | 34                          | 34                              | 34                                    | 6                             |

Bold text indicates responses consistent with traditional expectations about how hybrid segregation and yield degradation occur.

Underlined text is consistent with improved varieties adopting the characteristics of local, traditional varieties.

Gray, italicized text runs counter to local adaptation expectations.

**Table 5.3 Descriptive information for genotyped populations across 22 SSR loci**

| Population         | Sample<br>size | Actual<br>sample size† | Allele<br>number<br>per locus | Gene<br>diversity | Observed<br>heterozygosity | Fixation<br>index (Fis) |
|--------------------|----------------|------------------------|-------------------------------|-------------------|----------------------------|-------------------------|
| D880F <sub>1</sub> | 24             | 23.36                  | 2.50                          | 0.32              | 0.38                       | -0.19                   |
| D880F <sub>2</sub> | 24             | 23.82                  | 3.41                          | 0.40              | 0.31                       | 0.22                    |
| D880F <sub>8</sub> | 24             | 22.73                  | 4.05                          | 0.55              | 0.34                       | 0.39                    |

† Actual sample size is adjusted for null data.

### *Population differentiation*

The commercial F<sub>1</sub> hybrid variety was “very greatly” differentiated from the F<sub>2</sub> of the same variety by the conventional measures of F<sub>st</sub> differentiation detailed in Hartl and Clark (1997) and shown in Table 5.4. This differentiation was consistent with the yield declines predicted and observed between the F<sub>1</sub> and F<sub>2</sub> generations (Ramirez-Vallejo et al. 1986; Espinosa-Calderon et al. 1990; Espinosa-Calderon et al. 1993; Morris et al. 1999). Obviously, there is a great deal of change that occurs between the first and second years of growing a hybrid variety. These differences are due both to the allele frequency shifts from the hybrid state, and new alleles arriving via pollen

flow. Surprisingly, the D880 hybrid was less differentiated from the  $F_8$  generation variety planted by one farmer than from the  $F_2$  generation variety planted by another. Further investigation with more individuals and more populations will be required to determine if this type of effect is seen in all fields or if it can be explained by field-to-field variation.

D880 was greatly differentiated from populations of the traditional variety Jala, whether in the field in 1999 or from the genebank (represented respectively by the FF and GB categories in Table 5.4). As farmers kept increasingly advanced generations, the D880 *criollo* variety began to resemble the Jala variety, with  $F_{st}$  values declining for the  $F_2$  and  $F_8$  generations. The implication is that D880 acquired alleles from the surrounding area through pollen flow. Farmer selection perhaps subsequently enhanced allele frequency shifts.

**Table 5.4 Pairwise  $F_{st}^1$  values for Jala from genebank and farmers' fields, 22 loci**

| Population  | DD880 | DD880 $F_2$ | FF <sup>2</sup> | GB <sup>3</sup> | GB old <sup>4</sup> | GB 1988 <sup>5</sup> |
|-------------|-------|-------------|-----------------|-----------------|---------------------|----------------------|
| DD880       | -     | -           | 0.18            | 0.22            | 0.23                | 0.22                 |
| DD880 $F_2$ | 0.34  | -           | 0.14            | 0.15            | 0.16                | 0.18                 |
| DD880 $F_8$ | 0.23  | 0.10        | 0.06            | 0.07            | 0.08                | 0.08                 |

<sup>1</sup> According to (Wright 1978)(Hartl and Clark 1997), little differentiated populations, where  $F_{st}$  is  $< 0.05$ , are not highlighted. Moderately differentiated populations,  $F_{st}$  is between 0.05 and 0.15 are highlighted in light gray. Greatly differentiated populations,  $F_{st}$  between 0.15 and 0.25 are highlighted in medium gray. Very greatly differentiated populations,  $F_{st} > 0.25$  are highlighted in dark gray.

<sup>2</sup> FF represents a composite of Jala from 8 farmers' fields (collected in 1999).

<sup>3</sup> GB represents a composite of 8 genebank populations of Jala.

<sup>4</sup> GBold represents a composite of 5 Jala genebank populations 1944 to 1952, before improved varieties were introduced in Jala.

<sup>5</sup> GB1988 represents a composite of the three populations from 1988, after improved varieties were introduced in Jala.

Remarkably, Jala, a variety only planted on 5% of the valley's maize area, appeared to heavily influence the genetics of D880 *criollo* varieties, even though D880 and D880-derived *criollo* varieties were planted on more than 30% of the valley maize

area. However, the Jala variety probably reflected the allele profile of many traditional varieties in the area. Traditional varieties are planted on about one-third of the valley's maize area. Several recent studies have found it impossible to differentiate farmers' varieties genetically, even though they may be physically different (Busso et al. 2000; Pressoir and Berthaud 2004b). The flowering times of D880 and Jala overlap, so pollen transfer is possible (Chapter 2). Furthermore, Jala plants are very tall, with large tassels and prolific amounts of pollen. Because of variation between plants and the large tassel size, Jala flowers for more than one month. The D880 plants are shorter, with a much briefer window of flowering times, as well as smaller tassels and much reduced quantities of pollen.

Gene flow in the other direction, from D880 into Jala, is another type of dynamic change that has become increasingly important to understand, given increasing prevalence of genetically engineered maize varieties (Bellon and Berthaud 2004). A paired t-test comparing  $F_{st}$  values between D880 and the Jala genebank composite (at all loci) against  $F_{st}$  values between D880 and the older (1944 to 1952) Jala populations from the genebank (from before the introduction of improved varieties in the valley) is highly significant ( $p=0.001$ ). This indicates that the older populations from the genebank are more different from D880 than are the more recent ones. However, the 1944 populations are more highly differentiated from everything, including each other, and may be influencing the significance levels. (Comparing to a composite of 1950s genebank populations gives  $p=0.095$ ).

#### *Allele frequency comparisons*

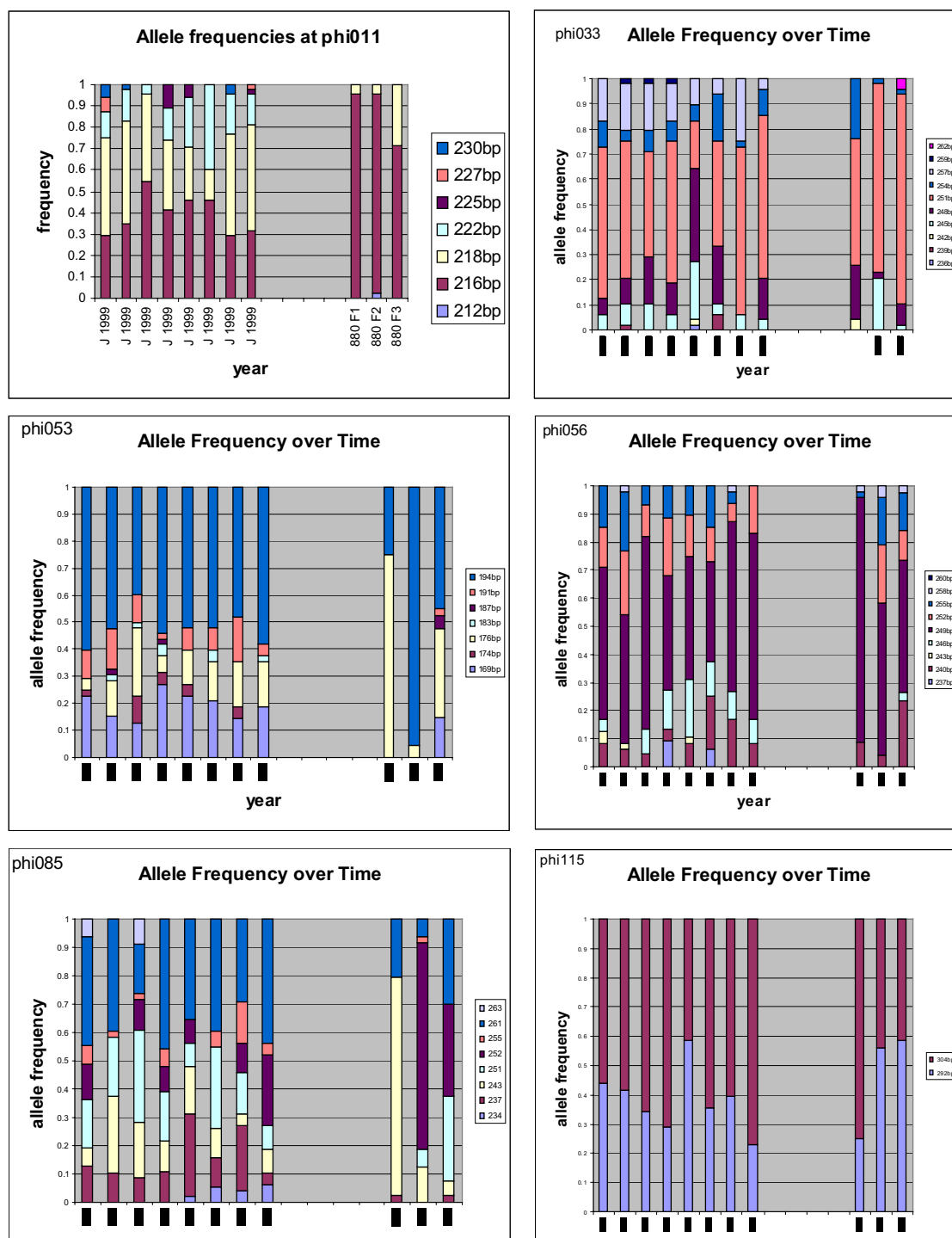
Theoretically, a three-way cross like D880 has a maximum of three alleles in a 1:2:1 ratio, 2 alleles in 3:1 or 1:1 ratios, or just one allele at a locus. Deviations from these proportions reflect error caused by sampling or detection methods or could be

the result of seed impurities (tolerated at a 5% level). Figure 5.1 shows the allele frequencies for six randomly selected loci in the Jala populations collected from farmers' fields in 1999 and the D880-derived populations collected in 2001. It is important to note that the sample sizes for this analysis were not selected to maximize accuracy of allele frequencies. Therefore, substantial variation in allele frequencies is to be expected from sampling alone. Compared with Jala at the 22 loci studied, D880 contained only one different allele, and that was at very low frequency (2%, likely the result of seed impurities). Thirty percent of the D880 alleles (16 of 54) were outside the range of frequencies found in Jala populations. Quickly, often within the first generation, the allele frequencies shift to proportions much more similar to the Jala profile.

### ***Conclusions***

For many, if not most, Jala farmers, the primary rationale for keeping *criollo* varieties was economic. Many believed that yields remain acceptable for a few years but eventually declined. Farmers in Jala were keeping their *criollo* varieties for two to four years before renewing seed. A few other farmers kept seed because they believed the variety “improves”—a local adaptation story.

The surveyed farmers described plant changes in *criollo* varieties that were consistent with both hybrid degeneration and local adaptation. *Criollo* plants take on characteristics of the local varieties, especially plant height. However, only a few people believe that yield increases. Most describe yield decreases and increased variation, consistent with the loss of hybrid vigor.



**Figure 5.1 Comparison of allele frequencies for D880 and Jala populations at 6 randomly selected loci**

In each drawing, the alleles are stacked in proportion to their frequencies, organized by size and measured in base pairs.

### *Evidence of gene flow*

The genetic data shows evidence of gene flow. Given the patchwork of small fields in Jala, it was not surprising that pollen mixes and genes move. The *criollo* version of D880 was clearly picking up alleles consistent with the profile of the traditional Jala variety. Older accessions from the genebank are less similar to D880 than are recent genebank collections—perhaps indicating gene flow from improved varieties into the traditional variety as well. Much of the movement from the traditional into the *criollo* varieties seems to be happening quite quickly, probably driven by pollen mixing in the first year. The allelic profile of a 7-year *criollo* is similar in numbers of alleles (and often in allele frequencies) to that of a traditional variety.

### *Implications for conservation and varietal introduction*

The increasing similarity to the Jala variety implies that *criollo* varieties (and perhaps other traditionally managed varieties) could be potentially overlooked reservoirs of traditional genetic information. If the goal of conservation is to maintain adaptive genes or gene complexes, we may find them incorporated into more than just the targeted varieties.

### *Future research*

These results are preliminary, but indicate clear and interesting avenues for further pursuit. A carefully executed study using these materials, supplemented by those described in Appendix III, could provide important information about field-level gene flow, local adaptation, effects of farmer selection, a quantification of gene



transfer due to pollen flow, and information about how *criollo* varieties change under farmer management.

Further efforts should include both an agronomic study and a genetic study. The agronomic study should be replicated across several environments, including Jala, to evaluate the question of local adaptation. It should focus on physical traits such as seedling vigor, emergence from depth, drought resistance, plant height, ear size and other characteristics mentioned by farmers as involved in local adaptation. To evaluate the issues of hybrid degeneration, the agronomic study should also focus on yield evaluations and measures of variability in characteristics like plant height, ear height, flowering time, germination date and days to maturity which will enable the field data to be compared to the modeled rates of change for three-way cross hybrids. A modeling approach can be used to account for the variation expected with hybrid degeneration, but will require knowledge of the degree of inbreeding in and genetic distance between the inbred line parentage of D880.

The molecular component of the study should include both more populations and larger numbers of individuals evaluated per population. Larger sample sizes will bring down sampling error and help to look for shifts in allele frequencies. Unfortunately there are indications that numbers much larger than 100 are not sufficient to accurately estimate gene frequencies (B-Rao 2001). Since there are other measures of population differentiation and diversity available to distinguish populations, allele frequencies are not the only tool available.

By using selected versus unselected samples, the effects of farmer selection versus pollen flow combined with farmer management can be separated. Using other  $F_2$  and multiple intermediate ( $F_3$ ,  $F_4$ ,  $F_5$ , etc.) populations will allow evaluation of how much change is field-specific and how much is generalizable, as well as addressing rates and directions of change in  $F_2$  and  $F_3$  generations.

This study has exciting potential, especially in light of recent concerns about gene flow from transgenic maize into traditional populations of Mexican maize landraces. Furthermore, it will help us to better understand little-studied, field-level genetic dynamics, and help to quantify the effects of farmer selection, farmers' seed management and pollen flow.

## CHAPTER 6 CONCLUSIONS

Conservation occurs in a social, policy and biological context. The decisions of farmers and policy makers not only affect conservation but also have implicit genetic effects. The case study of the giant maize of Jala examined the social and biological context of Jala's conservation, as well as the genetic effects of *in-situ* and *ex-situ* conservation efforts. The social and policy context in which the farmers of Jala make their conservation decisions, described in Chapter 2, raised genetic questions about Jala conservation. Sampling—a critical issue not only in genetic studies, but in conservation as well—was evaluated for molecular genetic measures of diversity in Chapter 3, using different types of diverse maize populations. The sample size requirement for accurate population diversity and differentiation measures was smaller than anticipated, and population history was less important to sample size than anticipated. However, the distribution of diversity within and between farmers' fields and between genebank populations of the variety Jala, described in Chapter 4, showed consistent conservation of the Jala variety *in-situ* and *ex-situ*. Finally, farmer-managed “*criollo*” varieties grown in Jala, examined in Chapter 5, were a potential reservoir of Jala's traditional genetic material that also presented a possibly intriguing approach to evaluating gene flow in farmers' fields.

The following sections summarize the important points of the four studies in an integrated, problem-oriented fashion. First, the social context for conservation will be addressed together with the biological questions raised, and the genetic answers found. Next, the distribution of diversity between wild relatives, maize races and the Jala race will be discussed, followed by a section on diversity in farmers' fields. Then the *in-situ* and *ex-situ* conservation efforts for Jala will be compared, followed by a discussion of implications for conservation.

### *Social context for conservation in the field*

Current patterns of cultivation of the variety Jala in farmers' fields raised concerns about bottlenecked, genetically narrow, populations in the field. In the survey, Jala was grown on only 5% of the valley's maize area, yet nearly 20% of Jala's farmers grew it. Therefore the average plot size of the Jala variety was very small. Many farmers kept Jala seed for long periods of time—decades, if not generations—a situation that could further reduce genetic diversity, when farmers save only small amounts of seed for their small planted areas.

Contrary to the concerns raised, the Jala variety's genetic profile proved to be highly diverse. The average farmer's field of Jala contained 40% of the alleles found in the entire dataset—including wild relatives and varieties representing a broad spectrum of maize races from around the world. The Jala populations did not appear to have experienced a recent bottleneck or genetic drift, the genetic consequences of a small, fragmented agricultural landscape. That they did not was likely due to gene flow between farmers' fields and will be discussed in the section about diversity in farmers' fields. To understand whether this diversity is unique to Jala or is present in many races of maize found in farmers' fields, further studies are needed.

For consistent conservation to occur on-farm, two conditions must apply: 1) a variety should have high value to farmers and 2) it should make an important contribution to genetic diversity. In other words, farmers must have an incentive to conserve (Smale and Bellon 1999; Smale et al. 2004). Those who continued to grow Jala have found a way to make it economically viable. Jala producers fell into two categories. The “generalists” typically had relatively large amounts of land, of which they plant a very small proportion (about 6%) with the Jala variety. They tended to be older and they were “seed keepers”; most have had their seed lots for their entire lives. Typically they used the variety for home use, cutting *elotes* for family occasions,

selling grain to neighbors for specialty uses, maintaining seed for sale, or grinding the whole plant as silage for their animals. The Jala “specialists” had no animals and relatively small total amounts of land. They planted most (about 75%) of their land with the Jala variety. They were younger, had not had their seed lots for as long, sold most of their harvest as *elotes*, and often purchased their seed from the generalists. Furthermore, both the generalists and the specialists grew Jala because of the contest and the *Feria de Elote*, during which they could make a nice profit by selling their gigantic ears of corn-on-the-cob. The contest was started in the early 1990s to promote pride in Jala’s unique maize and to provide an incentive for conservation. The strategy has worked. Many of the surveyed farmers kept tiny plots of maize “*para concursar*,” to compete in the contest.

Together, all these pieces of information suggested that farmers in Jala currently find several incentives to grow the traditional Jala variety. However, farmers in Jala were substantially older than the farming population in Mexico [compared to de Janvry et al. (1997)]. Many of their children had emigrated to the cities of Mexico or the United States, and whether they will return to take over their parents’ farms is an open question—not just in Jala but all over Mexico (Taylor et al. 1999; Winters et al. 2001). The older population in Jala was driving the demand for specialty use of the grain in *pozole*, *marquezote* (a sweet bread), and *gorditas* (locally, a cornmeal cookie). Will local demand for Jala continue with the next generation? As people prepare fewer specialty foods at home, it seems unlikely. Other incentives, like sale of corn-on-the-cob and the contest, are likely to continue.

For the past 30 years in Jala, the Mexican government, through its buying agency CONASUPO, has set the desired production characteristics of maize. They bought only white grain, by weight. Therefore, the Jala variety, with its large, light-weight, cream-colored kernels was less than desirable. Now, farmers must find their

own buyers for maize, and the desirable characteristics may change. In Jala, as elsewhere in Mexico, maize is not highly profitable. Increasingly, the farmers of Jala are turning to tobacco and animal husbandry as potentially more profitable activities. Locations where farmers are likely to specialize in products other than maize are poor targets for *in-situ* conservation efforts (Smale et al. 2004). If Jala farmers specialize in tobacco to the exclusion of maize, Jala will become a poor choice for *in-situ* conservation. However, many studies have shown that farmers incorporate new varieties rather than completely replacing them [see Wood and Lenne (1997) for a discussion]. If farmers incorporate new species as they incorporate new varieties, tobacco may diversify farmers' production systems without necessarily displacing maize. Furthermore, if enough farmers specialize in animal production, there may be a bright future for on-farm conservation of Jala. Many people preferred the Jala variety for silage production, because of the tremendous amounts of biomass its 5m tall plants produce.

### ***Distribution of diversity between species and among maize races***

Maize is widely believed to have been domesticated from teosinte (Doebley 1990; Matsuoka et al. 2002). [See MacNeish and Eubanks (2000) for a different perspective.] The categories of maize populations examined in this study, whether diverse races, depth within one race, or populations from farmers' fields, each contained about 55% of the alleles in the entire dataset. In contrast, the set of teosinte populations held more than 90% of the alleles in the dataset. The reduction of diversity in maize populations was consistent with a bottleneck, as would be experienced in domestication. Interestingly, the populations representing maize races were distinct, as shown by high  $F_{st}$  values, but not very diverse, as shown by their relatively low allele number and gene diversity ( $H_e$ ) values. The lack of diversity in

the representative populations of maize races is likely due to *ex-situ* management factors, like the initial collection unit and regeneration history of the population. Without documentation of this history, it is impossible to rule out coincidental choice of non-diverse materials (in spite of explicit efforts not to do so) or unusual diversity of the Jala race. The unique physical characteristics and unusually narrow geographic distribution of the Jala race suggest that Jala would be less diverse than other maize races, rather than more—reinforcing the notion that *ex-situ* history is likely involved in the relatively low diversity of the maize race populations.

The different genetic histories of the teosinte populations, maize races, Jala from the genebank and Jala from farmers' fields suggested that different sampling techniques would be needed to assay their genetic diversity. Population genetic theory and modeling give good guidelines for the number of individuals that should be sampled in order to capture alleles in the field (Crossa 1989; Crossa et al. 1993; Crossa et al. 1994). To capture alleles present at frequencies of 0.05 and higher at 150 loci with 95% accuracy, 210 individuals are required (Crossa et al. 1993). However, estimating allele frequencies with two decimal accuracy requires one thousand to ten thousand individuals, depending on the loci and allele frequencies sampled (B-Rao 2001). For molecular genetic studies, neither capturing all alleles in a population nor estimating allele frequencies extremely accurately are necessary. Instead, accurate measures of population diversity and differentiation are needed. The resampling program detailed in Chapter 3 showed that for  $F_{st}$  measures of differentiation as well as allele number ( $A_n$ ) and gene diversity ( $H_e$ ), the variance around the statistic narrowed quickly; in some cases within a few individuals, and in most cases within ten individuals. Interestingly, there were mild differences for each measure used but few differences whether looking at highly differentiated populations like teosinte or very similar populations like Jala accessions out of the genebank. The only case in

which many individuals were required was that of  $F_{st}$  values for farmers' fields, which were undifferentiated from one another. In contrast to techniques for capturing alleles in a population or accurately estimating allele frequencies, the sample sizes required for diversity measures of populations were on the order of ten individuals.

### ***Diversity in farmers' fields***

In interviews, the farmers of Jala explained that one person's seed was much the same as another's—Jala is Jala, they said, referring to the variety. However, most of these same farmers also said they had kept their seed for decades, if not generations. With decades of different management, selection and genetic drift might separate their seed lots into genetically distinct entities, but the genetic data did not support this hypothesis. Instead, the farmers' statements were upheld: using the genetic markers and methods described here, Jala populations from different farmers' fields could not be differentiated using these 22 microsatellite markers. Other recent studies have shown it is difficult to genetically differentiate even differently named, physically distinct varieties from farmers' fields (Busso et al. 2000; Pressoir and Berthaud 2004b).

The farmers of Jala had many, very small plots of land scattered across the valley. The land distribution is a legacy of the *ejido* system, and the Jala valley, at least in this respect, was similar to other parts of Mexico (Bellon and Brush 1994; Louette et al. 1997; Rice et al. 1997; Bellon and Risopoulos 2001). Farmers grew many varieties of maize, both traditional and improved. Planting dates were staggered in such a way that flowering times overlap. Therefore, there was high potential for genetic interchange—*gene flow*—between varieties.

Farmers also kept improved varieties, sometimes for many generations. Because of the prevalence of the variety D880, a time sequence of D880 was obtained



from farmers who had grown it for one to eight years. D880 is a three-way cross hybrid with a distinct allelic profile. With a known genetic beginning point, and genetic theory to predict what the genetic profiles of subsequent generations should look like, any new alleles and frequency deviations from the model could be attributed to gene flow. A hybrid variety undergoes dramatic genetic changes between the  $F_1$  and  $F_2$  generations [see Morris et al. (1999) for a review]. D880 also picked up a large number of new alleles through gene flow in this generation. The D880 data showed evidence for gene flow into *criollo* varieties. The “new” alleles arriving in the *criollo* varieties correspond to alleles in the traditional Jala variety—thus ruling out mutation or transposable element activity as the source of this “novel” genetic information. There was also some evidence for gene flow out of improved varieties into the traditional Jala variety; a type of genetic exchange that has become increasingly important to understand, in light of the growing prevalence of transgenic maize (Bellon and Berthaud 2004).

### ***In-situ and ex-situ conservation of Jala maize***

The strong similarity of Jala in farmers’ fields could lead to the inference that the valley is the unit of conservation, perhaps functioning as a unit for pollen exchange. Logically then, *ex-situ* populations of Jala collected in the same year should be more similar to one another than to other populations in the genebank. For recent collections, like 1988, this prediction held true. However, the two oldest Jala populations in the genebank, collected in 1944, were the two most differentiated from one another, from all the genebank accession, and from the Jala materials in farmer’s fields. The older populations were also less diverse than recently collected

populations. Because the 1944 populations were more like other farmers' fields and genebank populations than they were like one another, contamination was ruled out.

The processes of collection and regeneration in genebanks can be sources of genetic narrowing. Recent data from a comparison of barley accessions conserved both in a genebank for up to 72 years and in farmers' fields showed a significant genetic narrowing and separation of older genebank populations (Parzies et al. 2000). Given that the goal of genebank collections was to find as much phenotypically diverse material as possible, without the emphasis on individual and population genetics we now have, it is not surprising that early collections were not as genetically diverse. Furthermore early management was probably not as carefully governed by population genetic theory as it is today.

Regardless of the relative genetic narrowness and distinction of the oldest Jala populations, Jala appears, for the loci assayed, to have been well conserved both in the genebank and in farmers' fields. The same alleles were present in all populations, within relatively similar frequencies. Jala in the field probably had some genetic material introgressed from improved varieties, but the overarching genetic profile has changed little in 55 years.

### ***Implications for conservation***

Admittedly, this was but one case study, from one community, examining only one race of maize. Generalization from these results must be made with caution, and should be verified by other studies. However, these results have some interesting implications for conservation both in genebanks and in the field. Even small plots of land may be large enough to preserve the genetic diversity of a species like Jala, at least to the resolution that one can easily assay with a genetic study like this. Farmers' fields of Jala were anything but genetically depauperate. Farmers who can find

incentives, be they economic or related to food quality, will continue to grow traditional varieties.

If there is little distinction between farmers' fields and a substantial amount of diversity within them, genebank collection of maize should focus on sampling many individuals from one field, rather than a few individuals from many fields. The critical element is the number of individuals collected from a population. In this case, the population is Jala over the valley, rather than one farmers' field. Together with studies like the barley genebank study (Parzies et al. 2000), these results point to the importance of the dynamic nature of genebank management, rather than the static nature of genebank storage commonly referred to in the literature. Someone screening genebank accessions for a desired trait in maize will likely find more diversity in recent accessions than in older ones.

### ***Questions for further research***

As mentioned before, this study is one specific case study that raises interesting questions about genebank and on-farm conservation of agricultural genetic diversity. Similar studies are needed in other species and with other maize varieties to understand if these results are unique to Jala or similar to other situations. For example, is Jala unusually diverse? Do other races of maize show the same consistency of diversity across farmers' fields? Do other older genebank accessions show the same genetic distinction from one another and reductions in diversity as this study and Parzies et al. (2000) found in barley? Do other varieties and species show the same constancy of allele profiles over time?

This approach used neutral markers to look at diversity across the maize genome. Use of markers that tag functional polymorphism, regulatory or structural genes, might reveal very different diversity relationships and histories. Farmers select

upon the functional genes, and thus they are likely to have a different evolutionary history.

Because these results are observational not experimental, it is impossible to completely disentangle the effects of genebank management from the dynamics of change in farmers' fields. If they could be separated experimentally, proportions of relative change could be assigned to each—a result that would be valuable in helping us to understand the differences in two dynamic processes. We attempted such an experiment, taking a seed lot and putting it under both genebank and farmer management to evaluate change. The approach suffered from two problems. The first was logistical: we had difficulty with the genebank regenerations and eventually lost the continuity of the experiment. The second had to do with resolution of the markers for detecting difference. If we could detect no difference between farmers' fields, it is difficult to have confidence that we could detect subtle differences due to selection and management in just a few generations. However, over more generations than the two thesis timing would allow, better resolution might be feasible. If one took selected and unselected samples, such an experiment would also be a way to evaluate the effect of farmer selection against the genetic influx resulting from pollen and gene flow.

The pilot study with D880 opens up some very interesting questions about gene flow, effect of farmer selection, local adaptation, as well as physical and genetic changes in advanced generations of improved materials that I hope to pursue in the future. These *criollo* materials, and probably other traditional varieties, are intriguing reservoirs of Jala diversity, even though they are physically distinct from the Jala variety.

Finally, the town of Jala will be an interesting location to continue to track on-farm diversity. How will conservation in the valley change with time? Will farmers

move out of maize production? Will the next generation continue to value the traditional variety? With this study as a baseline, a full economic analysis would shed much brighter light on the process of farmer decision making. Recently, the Mexican government has taken an interest in Jala as a case study of on-farm conservation and a candidate for further efforts. Hopefully, this national-level interest in the maize of Jala will help to keep it a viable, well-conserved variety for many more generations.

## APPENDIX I SURVEY FOR JALA STUDY 2001

Nombre \_\_\_\_\_ Esposo(a) \_\_\_\_\_  
 Edad \_\_\_\_\_ Encuestado por: \_\_\_\_\_  
 Comunidad \_\_\_\_\_ Fecha: \_\_\_\_\_

| Cultivos en este temporal? | Dueño | Superficie? | Superficie en total |
|----------------------------|-------|-------------|---------------------|
|                            |       |             |                     |
|                            |       |             |                     |
|                            |       |             |                     |
|                            |       |             |                     |
|                            |       |             |                     |
|                            |       |             |                     |

Tiene Ud. ganado? S/N                      Cuantas? \_\_\_\_\_  
 Tiene Ud. puercos? S/N                    Cuantas? \_\_\_\_\_  
 Tiene Ud. gallinas? S/N                   Cuantas? \_\_\_\_\_  
 Tiene Ud. caballos? S/N                   Cuantas? \_\_\_\_\_

| Maíz: Variedad | Tipo | Season | Superficie | Porque le gusta esa variedad? |
|----------------|------|--------|------------|-------------------------------|
|                |      |        |            |                               |
|                |      |        |            |                               |
|                |      |        |            |                               |
|                |      |        |            |                               |
|                |      |        |            |                               |

| Maíz: variedad | anos con<br>variedad semilla | anos con<br>semilla bueno | fuelle | rend.año<br>malo | rend.año<br>normal | rend.año<br>forraje | rend. |
|----------------|------------------------------|---------------------------|--------|------------------|--------------------|---------------------|-------|
|                |                              |                           |        |                  |                    |                     |       |
|                |                              |                           |        |                  |                    |                     |       |
|                |                              |                           |        |                  |                    |                     |       |
|                |                              |                           |        |                  |                    |                     |       |
|                |                              |                           |        |                  |                    |                     |       |

Que le gusta Ud. Hacer con este maíz?

\_\_\_\_\_

\_\_\_\_\_

**Maíces Mejorados:**

Variedades que Ud. Ha probado:

Variedad    años que tenia    características    porque deajo de sembrar?

|  |  |  |  |
|--|--|--|--|
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Maíz acriollado:**

Como cambia el maíz: entre el primer año y el \_\_\_\_\_ año?

Característica    Primer año    Despues de \_\_\_\_\_ años    Como ha cambiado

|                                       |  |  |  |
|---------------------------------------|--|--|--|
| Rendimiento                           |  |  |  |
| Altura de la Planta                   |  |  |  |
| Tamano de la Mazorca                  |  |  |  |
| Guesor del tallo                      |  |  |  |
| Variación de las mazorcas             |  |  |  |
| Variación de la altura de las plantas |  |  |  |
| Nacimiento (germinacion)              |  |  |  |
| Resiste sequia                        |  |  |  |
| Resiste plagas                        |  |  |  |
|                                       |  |  |  |
|                                       |  |  |  |

Cuantos anos es mejor para acriollar? \_\_\_\_\_ Porque? \_\_\_\_\_

Variedad \_\_\_\_\_ Anos que ha tendió? \_\_\_\_\_ De donde vino? \_\_\_\_\_

Fue semilla nueva? \_\_\_\_\_ Ud. Ha añadido otra semilla? \_\_\_\_\_

**Maíz de Húmedo:**

Tiene Ud. Maíz de Húmedo? Si/No Si si, por cuanto tiempo? \_\_\_\_\_

Porque siga Ud.

cultivándola? \_\_\_\_\_

Si no, Tenia Ud. MDH? Si/No Si si, hace cuanto tiempo? \_\_\_\_\_-----\_\_\_\_\_

Porque deajo Ud. de cultivarla? \_\_\_\_\_

Hay **mercado** para MDH? Si/NoVende en elote? Si/No    precio por mazorca \_\_\_\_\_    precio por costal \_\_\_\_\_  
precio por ha \_\_\_\_\_Vende en grano? Si/No    precio por medida \_\_\_\_\_    precio por kg \_\_\_\_\_  
precio por hL \_\_\_\_\_

Vende para semilla? Si/No    precio por medida \_\_\_\_\_    precio por kg \_\_\_\_\_

Vende en forraje? Yes/No    precio por ton \_\_\_\_\_

Cuántas kgs/medida (MDH)?    medidas/hL de MDH? \_\_\_\_    Kgs/hL de MDH? \_\_\_\_

Cuántas kgs/medida de mejorado?    medidas/hL de mejorado? \_\_\_\_    Kgs/hL de mejorado \_\_\_\_

**Fuente de semilla** para este año? \_\_\_\_\_    El año pasado? \_\_\_\_\_

Hace tres años? \_\_\_\_\_    Hace 5 años? \_\_\_\_\_

Que le hace si no tiene semilla para el próximo ciclo? \_\_\_\_\_

Ud. Le gusta añadir semilla de otra fuente? \_\_\_\_\_

De donde/quien? \_\_\_\_\_    Cuando? \_\_\_\_\_



## APPENDIX II SAMPLE SIZE FOR DIVERSITY STUDIES

**Table II.1 Sample sizes for recent diversity studies using SSR markers**

| Study                               | crop         | populations  | samp.size                    | # markers                       | measures  | goal  |
|-------------------------------------|--------------|--|------------------------------|---------------------------------|---|---|
| Pressoir and Berhaud (2004a; 2004b) | maize        | 31 farmers' field maize populations from 6 villages                  | 20                           | 11 SSR<br>1 chloro-plast<br>SSR | 1. $F_{is}$ , $H_e$<br>2. $F_{st}$  | GD within populations<br>GD among populations         |
| Reif et al. (2004)                  | maize        | 23 maize breeding populations adapted to different environments.     | 21-48                        | 83 SSR                          | $A_n$ , $H_e$<br>$F_{st}$ and $G_{st}$  | GD within populations<br>GD among populations         |
| Barcaccia et al. (2003)             | maize        | 10 farmers' fields, 3 synthetics/landraces, 1 inbred line            | 20                           | 83 SSR<br>53 I-SSR              | 1. $A_n$ , $H_e$ , $F_{is}$<br>2. $D_{st}$ , $G_{st}$ , $F_{st}$ , $\Delta T$ | 1. GD within populations<br>2. GD among populations   |
| Budak et al. (2003)                 | pearl millet | 53 pearl millet lines  | 5                            | 18 SSR                          | 1. genetic distance   | 1. GD in populations                                  |
| Fu et al. (2003)                    | oat          | 96 released varieties  | 10 bulked                    | 30 SSR                          | 1. $A_n$ , $S_{ij}$ (similarity)  | 1. GD in populations                                  |
| Labate et al. (2003)                | maize        | 56 open pollinated populations from 4 racial groups, 12 inbred lines | 5-25/<br>pop<br>1/<br>inbred | 20 SSR                          | 1. $A_n$ , $H_e$ ,<br>2. $F_{st}$ ,<br>genetic distance                       | 1. GD within populations<br>2. GD between populations |
| Vom Brocke et al. (2003)            | pearl millet | 27 farmer varieties<br>12 control varieties                          | 19-30                        | AFLP                            | 1. AMOVA<br>2. AMOVA<br>GD, <i>gene flow</i>                                  | 1. GD in populations<br>2. GD between populations     |
| Bhattaharjee et al. (2002)          | pearl millet | 10 landrace/genebank accessions                                      | 25                           | 16 (58) RFLP                    | 1. AMOVA<br>2. AMOVA  | 1. GD within populations<br>2. GD between populations |
| Gethi et al. (2002)                 | maize        | 6 inbred lines   | 8/inbred                     | 44 SSR                          | 1. $A_n$ , $H_e$ , AMOVA<br>2. genetic distance, AMOVA                        | 1. GD within populations<br>2. GD between populations |
| Matsuoka et al. (2002)              | maize        | 193 maize landraces<br>71 teosinte                                   | 1                            | 99 SSR                          | 1. genetic distance, structure  | 1. GD between populations                             |
| Warburton et al. (2002)             | maize        | 7 breeding populations, 57 inbreds                                   | 48/pop                       | 85 SSR                          | 1. $A_n$<br>2. genetic distance   | 1. GD within populations<br>2. GD between populations |
| Rebourg et al. (2001)               | maize        | 130 European maize populations                                       | 30 bulked                    | 29 RFLP                         | 1. $A_n$ , $H_e$<br>2. $H_e$ , $G_{st}$ , MRD                                 | 1. GD within populations<br>2. GD among populations   |
| Sanchez et al. (2000)               | maize        |  |                              | isozyme                         |   |   |

**Table II.1 (Continued)**

|                                |       |  |            |   |                                     |                              |
|--------------------------------|-------|--|------------|---|-------------------------------------|------------------------------|
| Dubreuil and Charcosset (1999) | maize | 10 European OPVs<br>62 Inbred lines            | 30         | 28<br>RFLP                                |                                     |                              |
| Rebourg et al. (1999)          | maize | 65 European OPVs<br>US synthetics              | 30<br>bulk | 28<br>RFLP                                |                                     |                              |
| Senior et al. (1998)           | maize | 94 inbred lines                                | 1/line     | 70 SSR                                    | 1. A <sub>ns</sub> gen.<br>distance | 1. GD between<br>populations |
| Doebley et al. (1988)          | maize | 94 maize landraces<br>37 teosinte pops         | 12         | 21<br>isozyme                             | 1. allele<br>frequency              | GD between<br>populations    |
| Bretting, et al. (1987)        | maize | 6 West Indian<br>landraces<br>(19 populations) | 12         | 23 isozy<br>me<br>21 chro<br>mo.<br>knobs |                                     |                              |
| Doebley et al. (1986)          | maize | 18 Northern Flint<br>populations               | 12         | 21<br>isozyme                             |                                     |                              |
| Smith (1986)                   | maize | 47 cornbelt dent<br>OPVs                       | 25         | 21<br>isozyme                             |                                     |                              |
| Doebley et al. (1985)          | maize | 34 mexican landraces<br>(94 populations)       | 6-12       | 23<br>isozyme                             |                                     |                              |
| Smith et al. (1985)            | maize | 43 populations<br>teosinte and maize           | 12-24      | 21<br>isozyme                             |                                     |                              |
| Smith et al. (1984)            | maize | 78 Mexican landraces                           | 12-24      | 21<br>isozyme                             |                                     |                              |
| Goodman and Stuber (1983)      | maize | 31 Bolivian landraces<br>(101 populations)     | 12         | 23<br>isozyme                             |                                     |                              |

## APPENDIX III DOES SAMPLING METHOD MATTER?

### *Introduction*

For this study, we collected seed in a variety of different ways (Table III.1). Some seed was collected from farmers by taking a balanced number of kernels from a particular number of ears (20, 30 or 60 ears). Other seed came from the genebank and therefore was a random sampling of seed from a jar, with a very different management history. We were concerned that the different sampling methods could have an impact on our ability to determine critical measures of genetic diversity or population differentiation like  $F_{st}$ ,  $A_n$  or  $H_e$ .

Population genetic theory evaluates the potential impact of sampling in terms of effective population size, or the number of breeding individuals within a population (Hedrick 2000). For sampling or germplasm regeneration, Crossa et al. (1994) point out that the number of gametes each individual contributes is important to determining the expected sample size of the next generation.

**Table III.1 Population and collection method**

| Name         | # in use | Kernels / ear     | # ears collected |
|--------------|----------|-------------------|------------------|
| FF1          | 24       | 1 kernel / ear    | 30               |
| FF2          | 24       | 1-2 kernels / ear | 20               |
| FF3          | 24       | 1-2 kernels / ear | 20               |
| FF4          | 24       | 1-2 kernels / ear | 20               |
| FF5          | 24       | 1-2 kernels / ear | 20               |
| FF6          | 24       | 1 kernel / ear    | 60               |
| FF7          | 24       | 1-2 kernels / ear | 30               |
| FF8          | 24       | 1 kernel / ear    | 20               |
| all GB pops  | 24       | 24 seeds from bag |                  |
| all LR pops  | 24       | 24 seeds from bag |                  |
| all Teo pops | 24       | 24 seeds from bag |                  |

This section evaluates whether different sampling methods would have a theoretical or measured effect on the effective sampling size of the maize populations sampled.

### ***Methods***

In essence, there are two ways to evaluate effective population size for our populations. The first is the theoretical effective population size and is based upon the number of male and female parents contributing to the next generation (Crossa and Vencovsky 1994). The second uses actual data, particularly the inbreeding coefficient (f), to compare effective population size to that of an ideal population. The inbreeding coefficient will depend on the number of individuals in the parental populations and the resultant probability that alleles are identical by descent (Crow and Kimura 1970).

1) “theoretical” effective population size (Crossa and Vencovsky 1994):

$$N_e = \frac{N_m}{\frac{N_m + 3}{4N_f} \frac{1}{4}}$$

where  $N_e$ = effective population size,  $N_m$ =number male gametes and  $N_f$  = number of female gametes.

2) “actual” effective population size (Vencovsky and Crossa 1999):

$$N_e = \frac{N}{1 + f}$$

where  $N_e$ =effective population size,  $N$ = number of seeds sampled and  $f$ =inbreeding coefficient calculated from genetic data.

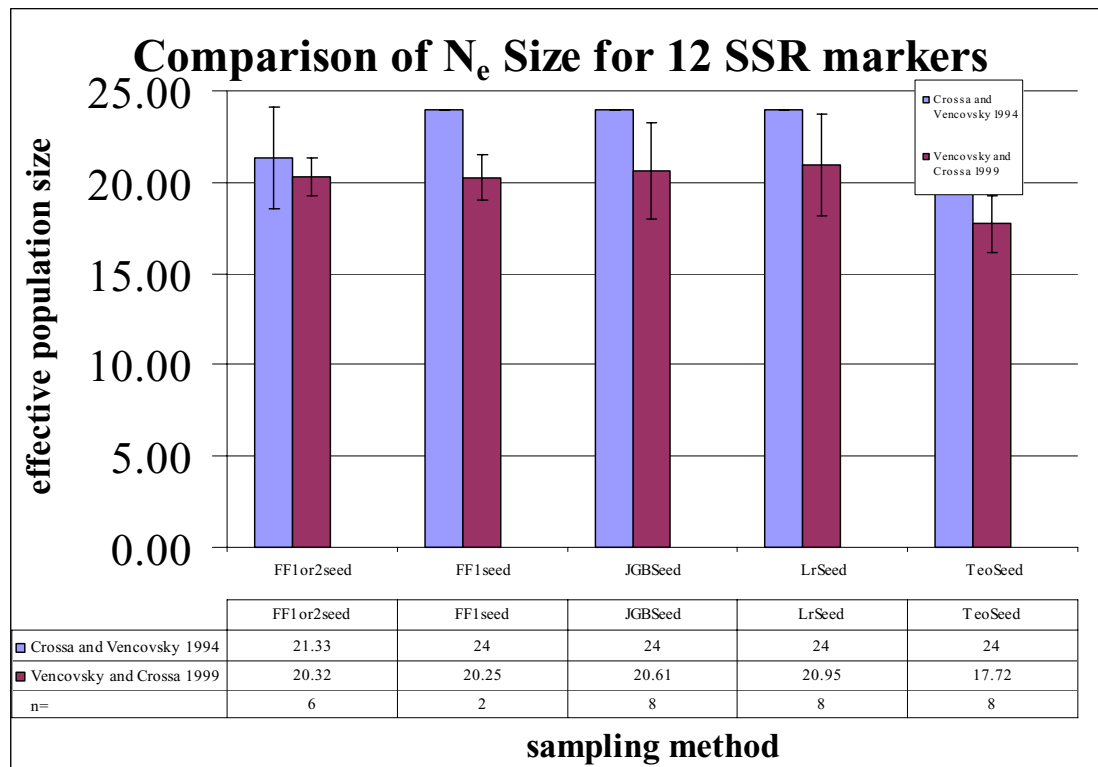
By using both these equations applied to the data from the populations of Jala collected from farmers’ fields (see Chapter 3) the effect of initial sample size on the effective populations size ( $N_e$ ) was evaluated.

### ***Results and conclusions***

Figure III.1 shows the different population sizes predicted for each of the different sampling methods. The light gray bars, showing the “theoretical” effective

population size based on male and female parents, do show some difference on the verge of being significant between the farmers' fields sampled using one seed per ear versus those where one to two seeds per ear were sampled. However, the dark gray bars show clearly that there is no difference in actual effective population size between farmers' fields sampled with one or one to two seed sampling.

Interestingly, the only populations that show any actual differences are the teosinte populations, which show some slight difference from farmer's fields in terms of the effective population size sampled. This difference is likely due to the management of the populations in the genebank. The seeds were quite old and there were germination problems, which could potentially reduce the diversity sampled.



**Figure III.1 Comparison of  $N_e$  sizes by sampling method**

## APPENDIX IV POPULATIONS AVAILABLE FOR STUDY

**Table IV.1 Advanced generation D880 populations available for study**

| Working name                                | Seed name                        | Seed ID | Amount seed                     |
|---|----------------------------------|---------|---------------------------------|
| <u>Populations in use</u>                   |                                  |         |                                 |
|   | D880 (commercial hybrid)         | d1      | 1400 seeds                      |
| D880F <sub>1</sub>                          |                                  |         |                                 |
| D880F <sub>2</sub>                          | D880F <sub>2</sub> C sel 2000    | d2      | 3 sets x 1 seed / ear x 30 ears |
| D880F <sub>8</sub>                          | D880F <sub>8</sub> N sel 2001    | d16     | 3 sets x 1 seed / ear x 30 ears |
| <u>Populations collected but not in use</u> |                                  |         |                                 |
|   | D880F <sub>2</sub> C unsel 2000  | d3      | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>3</sub> C sel 2001    | d8      | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>2</sub> S sel 2000    | d4      | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>2</sub> S unsel 2000  | d5      | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>3</sub> S sel 2001    | d9      | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>3</sub> S unsel2 2001 | d10     | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>2</sub> R sel 2001    | d6      | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>2</sub> P sel 2001    | d7      | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>4</sub> H sel 2001    | d11     | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>4</sub> A sel 2001    | d12     | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>5</sub> O sel 2001    | d13     | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>6</sub> Z sel 2001    | d14     | 3 sets x 1 seed / ear x 30 ears |
|   | D880F <sub>7</sub> ? V sel 2001  | d15     | 3 sets x 1 seed / ear x 30 ears |

## REFERENCES

- Aguirre-Gomez, J.A., M.R. Bellon, and M. Smale. 2000. A regional analysis of maize biological diversity in Southeastern Guanajuato, Mexico. *Economic Botany* 54:60-72.
- Altieri, M.A., and L.C. Merrick. 1987. *In-situ* conservation of crop genetic resources through maintenance of traditional farming systems. *Economic Botany* 41:86-96.
- Anderson, E., and H. Cutler. 1942. Races of *Zea mays* I: Their recognition and classification. *Annals of the Missouri Botanical Garden* 21:69-88.
- Barcaccia, G., M. Lucchin, and P. Parrini. 2003. Characterization of a flint maize (*Zea mays* var. *indurata*) Italian landrace, II. Genetic diversity and relatedness assessed by SSR and Inter-SSR molecular markers. *Genetic Resources and Crop Evolution* 50:253-271.
- Baverstock, P.R., and C. Moritz. 1996. Chapter 2: Project Design, *In* D. M. Hillis, ed. *Molecular systematics*. Sinauer Associates, Sunderland, MA.
- Bellon, M., and S.B. Brush. 1994. Keepers of the maize in Chiapas, Mexico. *Economic Botany* 48:196-209.
- Bellon, M., and J. Risopoulos. 2001. Small-scale farmers expand the benefits of improved maize germplasm: a case study from Chiapas, Mexico. *World Development* 29:799-811.
- Bellon, M., and J. Berthaud. 2004. Transgenic maize and the evolution of landrace diversity in Mexico: The importance of farmers' behavior. *Plant physiology* 134:883-888.
- Bellon, M.R., J. Berthaud, M. Smale, J.A. Aguirre, S. Taba, F. Aragon, *et al.* 2003. Participatory landrace selection for on-farm conservation: An example from the Central Valleys of Oaxaca, Mexico. *Genetic Resources and Crop Evolution* 50:401-416.
- Bhattacharjee, R., P.J. Bramel, C.T. Hash, M.A. Kolesnikova-Allen, and I.S. Khairwal. 2002. Assessment of genetic diversity within and between pearl millet landraces. *Theoretical and Applied Genetics* 105:666-673.
- B-Rao, C. 2001. Sample size considerations in genetic polymorphism studies. *Human Heredity* 52:191-200.

- Brown, A.D.H. 2000. The genetic structure of crop landraces and the challenge to conserve them *in-situ* on farms, *In* N. Maxted, et al., eds. Plant genetics conservation: the in-situ approach. Chapman and Hall, London.
- Brush, S.B. 1991. A farmer-based approach to conserving crop germplasm. *Economic Botany* 39:310-325.
- Brush, S.B., (ed.) 1999. Genes in the Field: On-Farm Conservation of Crop Diversity, pp. 1-300. IDRC/IPGRI/Lewis Publishers 1999.
- Budak, H., F. Pedraza, P.B. Cregan, P.S. Baenziger, and I. Dweikat. 2003. Development and utilization of SSRs to estimate the degree of genetic relationships in a collection of pearl millet germplasm. *Crop Science* 43:2284-2291.
- Busso, C.S., K.M. Devos, G. Ross, M. Mortimore, W.M. Adams, M.J. Ambrose, *et al.* 2000. Genetic diversity within and among landraces of pearl millet (*Pennisetum glaucum*) under farmer management in West Africa. *Genetic Resources and Crop Evolution* 47:561-568.
- Cleveland, D.A., D. Soleri, and S. Smith. 1994. Do folk crop varieties have a role in sustainable agriculture? *Bioscience* 44:740-751.
- Cleveland, D.A., Daniela Soleri, and S. Smith. 2000. A biological framework for understanding farmers' plant breeding. *Economic Botany* 54:377-394.
- Collier, G.A. 1994. Reforms of Mexico's agrarian code: impacts on the peasantry. *Research in Economic Anthropology* 15:105-128.
- Cornelius, W., and D. Myhre. 1998. Introduction, *In* W. Cornelius and D. Myhre, eds. Transformation of Rural Mexico: Reforming the Ejido Sector. Center for US-Mexican Studies, UCSD, San Diego.
- Crossa, J. 1989. Methodologies for estimating the sample size required for genetic conservation of outbreeding crops. *Theoretical and Applied Genetics* 77:153-161.
- Crossa, J., and R. Vencovsky. 1994. Implications of the variance effective population size on the genetic conservation of monoecious species. *Theoretical and Applied Genetics* 89:936-942.
- Crossa, J., C.M. Hernandez, P. Bretting, S.A. Eberhart, and S. Taba. 1993. Statistical considerations for maintaining germplasm collections. *Theoretical and Applied Genetics* 86:673-678.



- Crossa, J., S. Taba, S.A. Eberhart, P. Bretting, and R. Vencovsky. 1994. Practical considerations for maintaining germplasm in maize. *Theoretical and Applied Genetics* 89:89-95.
- Crow, J.F., and M. Kimura. 1970. *An introduction to population genetics theory*. Burgess Publishing, Minneapolis, MN.
- de Janvry, A., and E. Sadoulet. 2001. Income strategies among rural households in Mexico: the role of off-farm activities. *World Development* 29:467-480.
- de Janvry, A., G. Gordillo, and E. Sadoulet. 1997. Mexico's second agrarian reform: household and community responses, 1990-1994. Center for US-Mexican Studies; University of California, San Diego., San Diego, California.
- Dean, R.E., J.A. Dahlberg, M.S. Hopkins, S.E. Mitchell, and S. Kresovich. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. *Crop Science* 39:1215-1222.
- Dje, Y., D. Forcioli, M. Ater, C. Lefebvre, and X. Vekemans. 1999. Assessing population genetic structure of sorghum landraces from North-western Morocco using allozyme and microsatellite markers. *Theoretical and Applied Genetics* 99:157-163.
- Doebley, J. 1990. Molecular evidence and the evolution of maize. *Economic Botany* 44:6-27.
- Doebley, J., M.M. Goodman, and C.W. Stuber. 1985. Historic variation in the races of maize from Mexico. *American Journal of Agricultural Economics* 72:629-639.
- Doebley, J., M.M. Goodman, and C.W. Stuber. 1986. Exceptional genetic divergence of Northern flint corn. *American Journal of Botany* 73:64-69.
- Doebley, J., M.M. Goodman, and C.W. Stuber. 1988. Patterns of isozyme variation between maize and Mexican annual teosinte. *Economic Botany* 41:234 -246.
- Dubreuil, P., and A. Charcosset. 1999. Relationships among maize inbred lines and populations from European and North-American origins as estimated using RFLP markers. *Theoretical and Applied Genetics* 99:473-480.
- Elias, M., O. Panaud, and T. Robert. 2000. Assessment of genetic variability in a traditional cassava (*Manihot esculenta* Crantz) farming system, using AFLP markers. *Heredity* 85:219-230.
- Espinosa-Calderon, A.E., M. Tadeo Robledo, and R. Valdivia Bernal. 1990. Capacidad productiva de semilla F1, F2 y F3 del hibrido simplede maiz H-34. Ciudad Juarez, Mexico.

- Espinosa-Calderon, A.E., M.A. Lopez-Pereira, and M. Tadeo Robledo. 1993. Analisis agro-economico del uso de generaciones avanzadas de semilla mejorada de maiz en los valles altos de Mexico. San Jose, Costa Rica.
- FAO. 1997. Report on the state of the world's plant genetic resources for food and agriculture. Prepared for the international technical conference on plant genetic resources, Leipzig, Germany.
- Frankel, O.H., and J.G. Hawkes. 1975. Crop Genetic Resources for Today and Tomorrow. Cambridge University Press, Cambridge, England.
- Fu, Y.-B., G.W. Peterson, G. Scoles, B. Rossnagel, D.J. Schoen, and K.W. Richards. 2003. Allelic diversity changes in 96 Canadian oat cultivars released from 1886 to 2001. *Crop Science* 43:1989-1996.
- Gethi, J.G., J.A. Labate, K.R. Lamkey, M.E. Smith, and S. Kresovich. 2002. SSR variation in important U.S. maize inbred lines. *Crop Science* 42:951- 957.
- Hallauer, A.R., and J.B. Miranda. 1981. Quantitative genetics in maize breeding. Iowa State University Press, Ames, Iowa.
- Hartl, D.L., and A.G. Clark. 1997. Principles of Population Genetics. Sinauer Associates, Inc., Sunderland, MA.
- Hedrick, P.W. 2000. Genetics of populations. Jones and Bartlett Publishers, Sudbury, MA.
- Heisey, P., M.L. Morris, D. Byerlee, and M.A. Lopez-Pereira. 1997. Economics of hybrid maize adoption, *In* M. L. Morris, ed. Maize seed industries in developing countries. Lynne Rienner Publishers, Boulder, Colorado.
- Hoisington, D., M. Khairallah, T. Reeves, J.-M. Ribaut, B. Skovmand, S. Taba, *et al.* 1999. Plant genetic resources: What can they contribute toward increased crop productivity? *Proceedings of the National Academy of the United States* 96:5937 - 5943.
- Jala. 1996. Jala: Plan Municipal de Desarrollo, 1996-1999. Oficina del President Municipal, Jala, Nayarit, Mexico.
- Kempton, J.H. 1924. Jala maize: A giant variety from Mexico. *Journal of Heredity* 15:337-344.
- Koo, B., P. Pardey, and B. Wright. 2003. The price of conserving agricultural biodiversity. *Nature Biotechnology* 21:126-129.

- Labate, J.A., K.R. Lamkey, S.E. Mitchell, S. Kresovich, H. Sullivan, and J.S.C. Smith. 2003. Molecular and historical aspects of corn belt dent diversity. *Crop Science* 43:80 - 91.
- Lewis, J. 2002. Agrarian change and privatization of *ejido* land in Northern Mexico. *Journal of Agrarian Change* 2:p. 401-419.
- Lewis, P.O., and D. Zaykin. 2002. GDA for Windows. Release 1.1 ed. University of Connecticut, Department of Ecology and Evolution, Storrs, CT.
- Listman, G.M., and F.P. Estrada. 1992. Mexican prize for the giant maize of Jala: source of community pride and genetic resources conservation. *Diversity* 8:14-15.
- Louette, D. 1999. Traditional management of seed and genetic diversity: what is a landrace?, *In* S. B. Brush, ed. *Genes in the field: On-farm conservation of crop diversity*. IDRC/IPGRI/Lewis Publishers.
- Louette, D., and M. Smale. 2000. Farmers' seed selection practices and traditional maize varieties in Cuzalapa, Mexico. *Euphytica* 113:25-41.
- Louette, D., A. Charrier, and J. Berthaud. 1997. *In-situ* conservation of maize in Mexico: genetic diversity and maize seed management in a traditional community. *Economic Botany* 51:20-38.
- Luna, S., J.F. M., B.B. M., R.G. L., R. Townsend, and J.B. Schoper. 2001. Maize Pollen Longevity and Distance Isolation Requirements for Effective Pollen Control. *Crop Science* 41:1551-1557.
- Macneish, R.S., and M.W. Eubanks. 2000. Comparative analysis of the Rio Balsas and Tehuacan models for the origin of maize. *Journal of Latin American Antiquity* 11:3-20.
- Marshall, D.R., and A.H.D. Brown. 1975. Optimum sampling strategies in genetic conservation, *In* O. H. Frankel, J G Hawkes, ed. *Crop genetic resources for today and tomorrow*. Cambridge University Press, Cambridge, England.
- Mather, K., and J.L. Jinks. 1971. *Biometrical genetics*. Chapman and Hall, London, UK.
- Matsuoka, Y., Y. Vigouroux, M.M. Goodman, J. Sanchez, E. Buckler, and J. Doebley. 2002. A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of the United States* 99:6080–6084.
- Maxted, N., B.V. Ford-Lloyd, and J.G. Hawkes. 1997. *Plant genetic conservation: the in-situ approach*. Chapman and Hall, London, UK.

- Maxted, N., L. Guarino, L. Myer, and E.A. Chiwona. 2002. Towards a methodology for on-farm conservation of plant genetic resources. *Genetic Resources and Crop Evolution* 49:31-46.
- Mcgregor, C.E., R.V. Treuren, R. Hoekstra, and T.J.L.V. Hintum. 2002. Analysis of the wild potato germplasm of the species *Acaulia* with AFLPs: implications for *ex-situ* conservation. *Theoretical and Applied Genetics* 104:146-156.
- Mitchell, S.E., S. Kresovich, C. Jester, C.J. Hernandez, and A.K. Szewck-McFadden. 1997. Application of multiplex PCR and florescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. *Crop Science* 37:617-624.
- Morris, M. 2000. Personal communication.
- Morris, M.L., J. Risopoulos, and D. Beck. 1999. Genetic change in farmer recycled maize seed: a review of the evidence. CIMMYT Economics Working Paper 99-07.
- Murphy, A. 1994. To title or not to title: Article 27 and Mexico's Urban Ejidos. *Urban Anthropology* 23:209-232.
- Naylor, R.L., W.P. Falcon, and A. Puente-Gonzalez. 2001. Policy reforms in Mexican agriculture: Views from the Yaqui Valley. CIMMYT, Mexico City.
- Ni, J., P.M. Colowit, and D.J. Mackill. 2002. Evaluation of genetic diversity in rice subspecies using microsatellite markers. *Crop Science* 42:601-607.
- Ochoa, E.C. 2000. Feeding Mexico: the political uses of food since 1910. Scholarly Resources, Inc, Wilmington, Delaware.
- Ott, J. 1992. Strategies for characterizing highly polymorphic markers in human gene mapping. *American Journal of Human Genetics* 51:283-290.
- Pardey, P.G., B. Koo, B.D. Wright, M Eric Van Dusen, and E. Al. 2001. Costing the conservation of genetic resources: CIMMYT's *ex-situ* maize and wheat collection. *Crop Science* 41:1286-1299.
- Parzies, H.K., W. Spoor, and R.A. Ennos. 2000. Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. *vulgare*) conserved for different lengths of time in *ex-situ* gene banks. *Heredity* 84:476-486.
- Phippen, W.B., S. Kresovich, F.G. Candelas, and J.R. McFerson. 1997. Molecular characterization can quantify and partition variation among genebank holdings: a case study with phenotypically similar accessions of *Brassica oleracea* var. *capitata* L. (cabbage) 'Golden Acre'. *Theoretical and Applied Genetics* 94:227-234.

- Prain, G. 1993. Mobilizing local expertise in plant genetic resources research, p. 102-110, *In* W. d. Boef, et al., eds. Cultivating knowledge: Genetic diversity, farmer experimentation and crop research. Intermediate Technology Publications, London, UK.
- Pressoir, G., and J. Berthaud. 2004a. Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity* 92:88–94.
- Pressoir, G., and J. Berthaud. 2004b. Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity* 92:95-101.
- Ramirez-Vallejo, P., M.M. Balderas, and F. Geron-Xavier. 1986. Potencial productivo de las generaciones avanzadas de los hibridos tropicales de maiz H-503, H-507 y H-510. *Fitotecnica* 8:20-34.
- Raymond, M., and F. Rousset. 1995. Genepop (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86:248-249.
- Rebourg, C., P. Dubreuil, and A. Charcosset. 1999. Genetic diversity among maize populations: bulk RFLP analysis of 65 accessions. *Maydica* 44:237-249.
- Rebourg, C., B. Gouesnard, and A. Charcosset. 2001. Large scale molecular analysis of traditional European maize populations. Relationships with morphological variation. *Heredity* 86:574-587.
- Reif, J.C., X.C. Xia, A.E. Melchinger, M.L. Warburton, D.A. Hoisington, D. Beck, *et al.* 2004. Genetic Diversity Determined within and among CIMMYT Maize Populations of Tropical, Subtropical and Temperate Germplasm by SSR Markers. *Crop Science* 44:326-334.
- Rice, E.B., J.L. Blanco, and M. Smale. 1997. Adoption and impact of improved seed selection practices in Mexican maize: evidence and issues from the Sierra de Santa Marta. *World Development* 36:1625-1640.
- Roberts, E.H. 1975. Problems of long-term storage of seed and pollen for genetic resources conservation, *In* O. H. Frankel, J G Hawkes, ed. Crop genetic resources for today and tomorrow. Cambridge University Press, Cambridge, England.
- Rodriguez, J.M., T. Berke, L. Engle, and J. Nienhuis. 1999. Variation among and within *Capsicum* species revealed by RAPD markers. *Theoretical and Applied Genetics* 99:147-156.
- Sanchez, G., M. Goodman, and C. Stuber. 2000. Isoenzymatic and morphological diversity in the races of maize in Mexico. *Economic Botany* 54:43-59.

- Genetics and Biometry Laboratory, University of Geneva, Switzerland. 2000. Arlequin 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland, Geneva, Switzerland.
- Senior, M.L., J.P. Murphy, M.M. Goodman, and C.W. Stuber. 1998. Utility of SSRs for determining genetic similarities and relationships in maize using an agarose gel system. *Crop Science* 38:1088-1098.
- Smale, M., and M. Bellon. 1999. A conceptual framework for valuing on-farm genetic resources, p. 387-408, *In* D. Wood and J. Lenne, eds. *Agrobiodiversity: characterization, utilization and management*. CABI Publishing, Wallingford, UK.
- Smale, M., M. Bellon, D. Jarvis, and B. Sthapit. 2004. Economic concepts for designing policies to conserve crop genetic resources on farms. *Genetic Resources and Crop Evolution* 51:121-135.
- Soleri, D., and S. Smith. 1995. Morphological and phenological comparisons of two Hopi maize varieties conserved *in-situ* and *ex-situ*. *Economic Botany* 49:56-77.
- Soleri, D., S. Smith, and D. Cleveland. 2000. Evaluating the potential for farmer and plant breeder collaboration: A case study of farmer maize selection in Oaxaca, Mexico. *Euphytica* 116:41-57.
- Stanford, L. 1994. *Ejidal Organization and the Mexican State: Confrontation and Crisis in Michoacan*. The Institute, Inc, New Mexico.
- Steiner, A.M., P. Ruckebauer, and E. Goecke. 1997. Maintenance in genebanks, a case study: contaminations observed in the Nurnberg oats of 1831. *Genetic Resources and Crop Evolution* 44:533-538.
- Tanksley, S.D., and S.R. McCouch. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277:1063-1066.
- Taylor, J.E., A. Yunez-Naude, and G. Dyer. 1999. Agricultural price policy, employment and migration in a diversified rural economy: a village-town CGE analysis from Mexico. *American Journal of Agricultural Economics* 81:653-662.
- Vavilov, N.I. 1926. Studies on the origin of cultivated plants. *Bulletin of Applied Botany* 16.
- Vavilov, N.I. 1997. *Five Continents*. International Plant Genetic Resources Institute, Rome, Italy.

- Vencovsky, R., and J. Crossa. 1999. Variance effective population size under mixed self and random mating with application to genetic conservation of species. *Crop Science* 39:1282-1294.
- Verma, S., V. Khanna, and N. Singh. 1999. Random amplified polymorphic DNA analysis of Indian scented basmati rice (*Oryza sativa* L.) germplasm for identification of variability and duplicate accessions, if any. *Electrophoresis* 20:1786– 1789.
- Virk, P., H. Newbury, M. Jackson, and B. Ford-Lloyd. 1995. The identification of duplicate accessions within a rice germplasm collection using RAPD analysis. *Theoretical and Applied Genetics* 90:1049–1055.
- Vom Brocke, K., A. Christinck, E.W. R., T. Presterl, and H.H. Geiger. 2003. Farmers' seed systems and management practices determine pearl millet genetic diversity patterns in semiarid regions of India. *Crop Science* 43:1680-1689.
- Warburton, M.L., X. Xianchun, J. Crossa, J. Franco, A. Melchinger, M. Frisch, *et al.* 2002. Genetic characterization of CIMMYT inbred lines and open pollinated populations using large scale fingerprinting methods. *Crop Science* 42:1832-1840.
- Weir, B.S. 1996. *Genetic Data Analysis II*. Sinauer Associates Inc., Sunderland.
- Wellhausen, E., L. Roberts, E. Hernandez, and P. Mangelsdorf. 1952. *Races of Maize in Mexico: Their Origin, Characteristics and Distribution*. The Bussey Institute of Harvard University, Cambridge, Ma.
- Winters, P., A.D. Janvry, and E. Sadoulet. 2001. Family and community networks in Mexico-U.S. migration. *Journal of Human Resources* 36:159-184.
- Wood, D., and J. Lenne. 1997. The conservation of agrobiodiversity on-farm: questioning the emerging paradigm. *Biodiversity and Conservation* 6:109-129.
- Wright, S. 1922. The effects of inbreeding and crossbreeding on guniea pigs. III. Crosses between highly inbred families USDA Bulletin No. 1121. USDA, Washington DC.
- Zeven, A.C. 1998. Landraces: A review of definitions and classifications. *Euphytica* 104:127-139.
- Zeven, A.C., K.J. Dehmer, T. Gladis, K. Hammer, and H. Lux. 1998. Are the duplicates of perennial kale (*Brassica oleracea* L. var. *ramosa* DC.) true duplicates as determined by RAPD analysis? *Genetic Resources and Crop Evolution* 45:105-111.