Evolutionary and food supply implications of ongoing maize domestication by Mexican *campesinos*

Mauricio R. Bellon[1], Alicia Mastretta-Yanes[2], Alejandro Ponce-Mendoza[1], Daniel Ortiz-Santamaría[1], Oswaldo Oliveros-Galindo[1], Hugo Perales[3], Francisca Acevedo[1], José Sarukhán[1,4]

[1]Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO). Liga Periférico-Insurgentes Sur No. 4903, Tlalpan, Mexico City, 14010, Mexico.

[2] CONACYT-CONABIO. Liga Periférico-Insurgentes Sur No. 4903, Col. Parques del Pedregal, Del. Tlalpan, 14010, Mexico City, Mexico.

[3]El Colegio de la Frontera Sur (ECOSUR). Departamento de Agricultura, Sociedad y Ambiente, Grupo de Agroecología, San Cristóbal de Las Casas, 29290, Chiapas, Mexico.

[4]Instituto de Ecología, UNAM. Universidad Nacional Autónoma de México, Mexico City 04510, Mexico.

Corresponding author: Mauricio R. Bellon. Email: mrbellon@gmail.com

ORCID: Mauricio R. Bellon https://orcid.org/0000-0003-0642-3402; Alicia Mastretta-Yanes https://orcid.org/0000-0003-2951-6353; Alejandro Ponce-Mendoza https://orcid.org/0000-0002-6220-2746; Hugo Perales https://orcid.org/0000-0003-3431-5759

Abstract

Maize evolution under domestication is a process that continues today in Mexico carried out by smallholder family farmers, known as *campesinos*. Here we examine the evolutionary and food security implications of the scale and scope under which *campesinos* produce maize. We gathered official municipal-level data on maize production under rainfed conditions and identified *campesino* agriculture as occurring in municipalities with average yields of ≤ 3 t/ha. Environmental conditions vary widely in those municipalities and are associated with a great diversity of maize races, representing 85.3% of native maize samples collected in the country. We estimate that in those municipalities around 1.38×10^{11} genetically different individual plants are subjected to evolution under domestication each season. This implies that 5.24×10^{8} mother plants contribute to the next generation with their standing genetic diversity and rare alleles. Such a large breeding population size also increases the total number of adaptive mutations that may appear and be selected for. We also estimate that *campesino* agriculture could potentially feed around 54.7 million people. This analysis provides insights about the importance of smallholder agriculture around the world as both an engine of crop evolution and a crucial source of food security.

1. Introduction

Domestication is not only a fascinating example of evolution, but also a process that changed the course of human history and that continues to influence the fate of humanity. Crop evolution under domestication, which started ca. 10,000 years ago [1], continues today in a range of agricultural systems spanning traditional farming to industrialized large-scale agriculture. However, it is in the centers of origin and/or domestication of crops species where the highest diversity is concentrated, and where this process continues to be driven directly by farmers [2]. Mexico is one such center for many crops, including maize [3-5], one of the most important food crops in the world, providing 30% of food calories to approximately 4.5 billion people in 94 developing countries [6]. It is the most important food crop in Mexico for both urban and rural populations [7].

Maize diversity is represented by a wide array of morphologically distinct crop populations grown across the country, known as native varieties or landraces. These are grouped into 59 maize races [8], containing an impressive level of genetic diversity, even within a single race [5, 9, 10]. *Campesino*s are heirs to, and trustees of, the largest genetic diversity of maize in the world [4, 5, 8-

11] which they maintain in their agricultural systems today. *Campesinos* are smallholder farmers managing family farms producing, at least partially, for self-consumption, relying mostly on family labor and using combinations of animal and mechanical traction, manure and inorganic fertilizers, and planting mostly native varieties. Their farms account for the largest area planted with maize in the country [11, 12]. The role *campesinos* continue to play in maintaining maize landrace diversity has been well documented [13-18]. They maintain traditional knowledge about the performance of their landraces and practices of saving and sharing seed by and among themselves, from one cycle to the next. This allows alleles to pass from one generation to the next, thus continuing the evolutionary processes that sustain and generate crop genetic diversity [3, 8, 14, 19].

In spite of their important role in maintaining maize genetic diversity, there is a common perception that *campesinos* are unproductive, anachronistic and a hindrance on Mexican agriculture [20, 21]. Their demise has been predicted for a long time, particularly after the implementation of the North American Free Trade Agreement (NAFTA), based on the assumption that they are inefficient and at a disadvantage in the face of competition from large-scale commercial farmers in open, globalized markets [22-24]. Current Mexican official agricultural policies are focused on medium and large commercially-oriented farmers, considered to be more economically viable, while *campesinos* are largely ignored [21, 23, 25]. A reduction in the importance of maize as income for rural households has continued since pre-NAFTA times [26, 27], but despite this, plus biased policies and negative predictions about *campesino* production, they continue to exist and produce maize for Mexico [21, 23, 25]. A key reason for this is that, for *campesinos,* maize is a "multi-functional" crop with great cultural importance [13, 28], providing a broad scope of benefits, such as food for self-consumption and food security [25, 29, 30], a diverse gastronomy and multiple products [29, 31], and opportunities to participate in local and regional markets to generate income [23, 25, 32].

In this paper we aim to assess the contributions made by *campesinos* producing maize under rainfed conditions for the evolution of this crop under domestication and for the national maize supply. Both contributions are complex in their own right, but it is worth addressing them together because they are intrinsically linked and reinforce each other. Although there are several regional and case studies characterizing *campesino* agriculture [23-25, 30, 33], to our knowledge these contributions have not been analyzed at a national scale; nor has their relevance to national and global scales been acknowledged. To solve this knowledge gap, we examined the implications of

the scale and scope under which *campesinos* produce maize by employing a multifaceted approach that integrates: (a) maize consumption and human demography, (b) biophysical variables using geographic information systems, and (c) population genetics and evolutionary theory.

2. Methods

(a) Identifying *campesino* maize production at the national level

Our study focuses on the main rainfed agricultural season (May-October) of 2010 (year of the last National Mexican Population Census). Municipal-level data on area planted and harvested with maize, average yields and total production were obtained from the official government statistics of the Sistema de Información Agroalimentaria y Pesquera (SIAP) Annual Statistics of Agricultural Production (http://nube.siap.gob.mx/cierre_agricola/). These data are based on field information collected and validated at the municipality level according to SIAP's Technical Standard for the Generation of Basic Agricultural and Fishery Statistics [34]. Municipalities were grouped into seven classes according to their average maize yield/ha in increments of one t/ha, from zero tons (area was planted with maize but none was harvested) to the maximum observed yield of 8.75 t/ha. Since SIAP's data do not include data on the structure of farm sizes, we estimated it using data from PROCAMPO (a government program providing a subsidy per hectare of land cultivated to farmers) on the number of beneficiaries that received a payment, and the size of the area they planted with rainfed maize. We cross-classified these data with SIAP's data on municipal average yields and area planted, assuming that a beneficiary corresponded to a farm. We adjusted PROCAMPO's and SIAP's data to estimate the overall distribution of area planted to maize by farm and by yield class (supplementary material, 1.1 Estimating farm size structure). Although PROCAMPO's data are not specifically intended for this purpose, these are the only data available with farm size of maize producers by municipality. Demographic data on total and rural population sizes by municipality were obtained from the National Census of 2010. We combined SIAP's and National Census´ data using the municipality as a joining term and keeping only those producing maize. The area planted with improved and native varieties was estimated based on the quantity of improved seed reported by different sources (supplementary material, 1.2 Estimating area planted with improved and native varieties, table S1).

(b) Environmental conditions of areas planted with maize

To analyze the variation in environmental conditions we focused on precipitation, temperature, altitude and slope at 1 km resolution. Precipitation and temperature data from May to October (rainy season) were extracted from Cuervo-Robayo *et al.* [35]. Altitude and slope data come from Guevara & Arroyo-Cruz [36, 37]. Before analyses, data were cropped to agricultural land (without rainfed/irrigation subcategories due to the uncertainty of defining such polygons) according to INEGI [38] for each for the municipalities presenting maize production in 2010 (SIAP's data). These data were then analyzed by: (1) visualizing the data with boxplots and violin plots by yield class; (2) examining differences in the spread of data using Feltz and Miller´s test [39]; (3) analyzing the overlap of the conditions by performing a Principal Component Analysis (PCA) and testing for differences among each yield-class with a distance-based test of homogeneity of multivariate dispersions [40]. For the PCA, ellipsoids were fitted to the components scores associated with each yield class and altitude was excluded because it is highly correlated with temperature; (4) analyzing the distance among the environments of each yield class with a dendrogram constructed from a hierarchical clustering analysis. This was done using Ward's method with Euclidean distances and the suggested partition (the one with higher relative loss of inertia). Analyses were performed using ArcMap and the R [41] packages raster [42], maptools [43], cvequality [44], vegan [45] and FactoMineR [46].

(c) *Campesinos´* contribution to maize genetic diversity

To estimate the contribution of smallholders to maize genetic diversity, we used classic population genetic formulae using as a proxy of effective population size (N_e) the effective number of breeding individuals (N_b [47]), which we assumed to be the number of ears from where seeds are extracted in order to sow 4.68 million hectares (Mha). This is a good proxy for maize because in maize mating is random, sex ratio is equal, generations are not overlapping, and the number of breeding individuals is relatively constant over generations. Due to the lack of empirical data at the national level for the number of ears set aside to generate the seed needed to plant one ha, and because this probably changes from region to region, we used four estimations that range from 114 to 290 ears/ha (supplementary material, 1.3 Estimating the amount of saved seed to plant one hectare).

The effect of genetic drift on heterozygosity (*H*) in the absence of mutation is calculated with the formula $H_{n+1}=(1-1/2N_e)H_n$ [48], where N_e is the effective population size. The decay of heterozygosity can then be estimated as $1/(2N_e)$. We used this formula to estimate the decay of

heterozygosity in the Mexican maize population using N_b instead of N_e. N_b was estimated as the number of mother plants that was used for planting the total rainfed maize area sown in 2010. We then estimated the number of new mutations by the formula $\mu 2 N_b$, where μ is the mutation rate. Here we only considered substitutions, using the mutation rate of $\mu=1.63\times10^{-8}$ [49]. The former number of mutations gives an estimate at the site level, so we next estimated the total number of mutations for the entire coding regions (genes that generate proteins) of the maize genome. For this, we assumed that the amount of DNA bases in coding regions is around 2% of the 2700 Mb of the maize genome [50] and multiplied it for the number of mutations estimated previously. Analyses were performed using R [41].

(d) Assessing the potential contribution of rainfed production to maize demand

The potential population that could be fed from maize production in 2010 was estimated by combining municipal-level data on maize production and rural population, assuming a per capita consumption of 108.5 kg/person. This number was calculated using the per capita consumption of tortillas of 217.9 g/person/day and 70.2 g/person/day of maize grain, employed by the government to develop poverty lines for the country [51] (supplementary material, table S2). We calculated four parameters for each municipality: (a) the size of the total population that could be fed from local maize production, regardless of the population present; (b) the size of the rural population in the municipality that could be fed from local maize production; (c) the size of any additional population that could be fed from surpluses if available beyond (b); (d) the size of the net population that could be or not be fed from local maize production (see details in supplementary material, 1.4 Estimating per capita annual maize consumption). We also calculated the size of the rural population that could be fed by year from 2003-2015 by following the same procedure (supplementary material 1.4 Estimating per capita annual maize consumption). This analysis is restricted to municipalities where there were demographic data in both census years to be able to calculate population rates of change (90.4% of those planted with maize in 2010).

3. Results

(a) Identifying *campesino* rainfed maize production at the national level

During the rainfed season of 2010, maize production took place in 2,271 municipalities (92.4% of the total). Municipalities in yield classes of ≤ 3 t/ha accounted for 78.4% of the area planted with

maize, generating 49.8% of the total production (figure 1). Most of the country's total and rural population was located there. Municipalities in yield classes of > 3 t/ha accounted for only 21.6% of the planted area, producing 50.2% of the total production (figure 1). We estimated that 1.99 million farms planted maize across all municipalities, with 88.6% planting up to 5 ha/farm (supplementary material, table S4). Most rainfed maize production occurred in municipalities with average yields of ≤3 t/ha and in farms planting up to 5 ha of maize, involving 1.68 million farms. Municipalities in yield classes below and above the 3 t/ha threshold produced each about half of the total rainfed maize production. Those below the threshold however, accounted for 4/5 of the total area planted, while those above it for the rest. This contrast is explained by differences in yields (supplementary material, figure S1). The overall average yield for municipalities in yield classes of ≤3 t/ha was 1.3 (±0.7) t/ha, while for municipalities above this threshold was 4.4 (±1.4) t/ha.

While there are no specific data on the type of germplasm used to plant each yield class, we estimated that improved varieties—mostly hybrids, and thus not subject to evolution under domestication *in situ* since seed generated elsewhere is purchased every year—were used in about 20% of the total rainfed area planted with maize, i.e. the majority of the area planted in municipalities in the top four yield classes (>3 t/ha) (see supplementary material 1.2 Estimating area planted with improved and native varieties). The rest of the area, comprising municipalities in the lower three yield classes, was mostly planted with the farmers´ own seed, thus subject to evolution under domestication. The presence of native varieties is widespread in municipalities in yield classes of ≤3 t/ha, as demonstrated by the fact that the Global Maize Project [8] collected 85.3% of their nationwide samples of native landraces in these municipalities (supplementary material, table S5).

Based on these data, we consider that *campesino* maize agriculture corresponds mostly to farms producing this crop in municipalities in yield classes of ≤3 t/ha. The bulk of smallholder farmers (planting ≤ 5 ha with maize) is located there, mostly relying on native landraces and traditional seed management practices, with relatively low average yields (1.3±0.7 t/ha). They are more likely to consume a larger proportion of their production and have smaller surpluses to sell. Commercially-oriented agriculture corresponds mostly to farmers producing in municipalities in yield classes > 3 t/ha relying mostly on improved varieties.
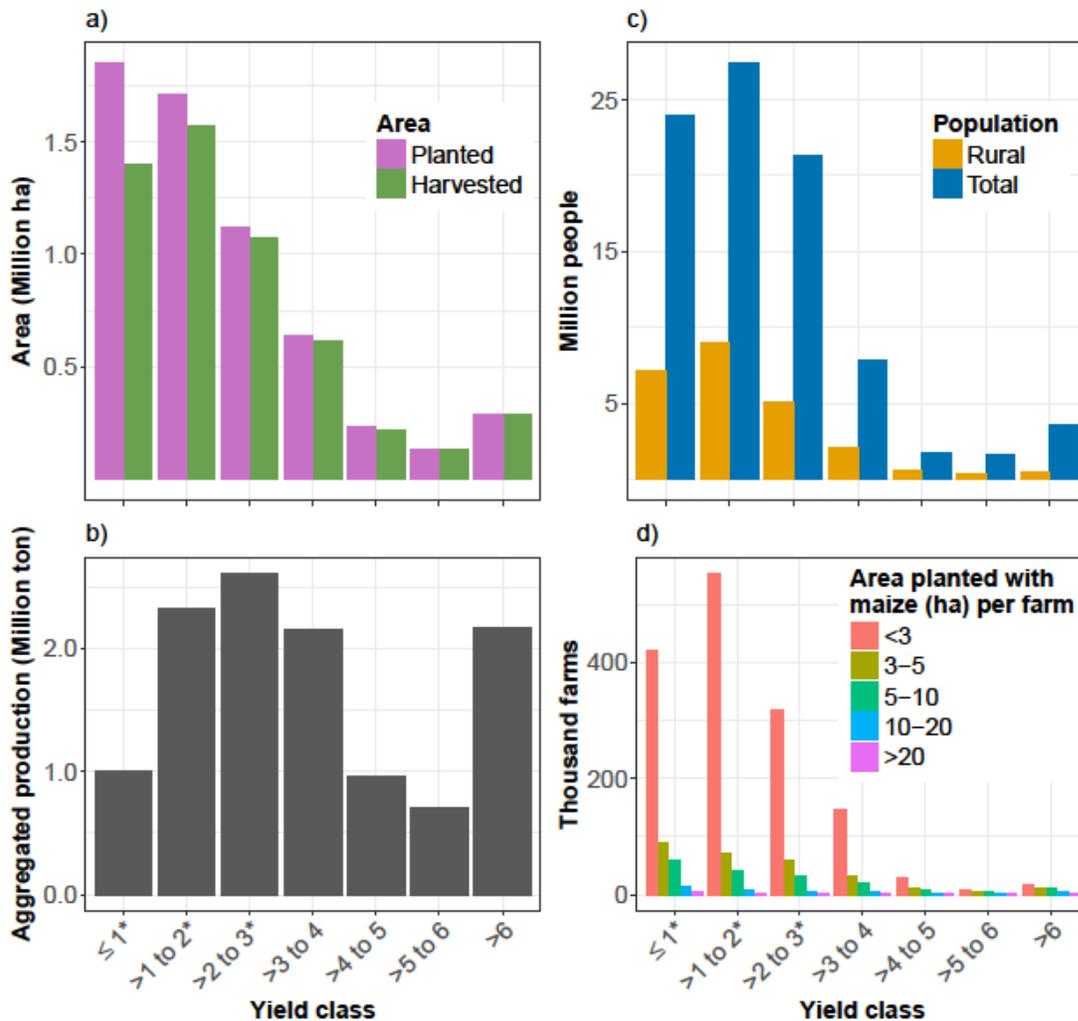
Figure 1. Summary characteristics of each yield class for the year 2010. a) area planted and harvested with rainfed maize and b) rainfed maize production according to SIAP, c) total and rural population according to INEGI, and d) estimated number of farms by area planted with maize per farm. *Yield classes mostly representing *campesino* agriculture.

(b) *Campesino* farmers produce maize under a wide variety of environments

The spread (variance) of the environmental conditions of maize areas differ greatly among yield classes; lower yield classes have a higher spread, that decreases in the higher yield classes (figure 2, supplementary material table S6). Altitude and temperature have a bimodal distribution that reflects the two main environments where maize is cultivated: temperate and tropical, consistent with genetic differentiation among highland and lowland maize [5, 9]. The PCA shows that there is a degree of environmental overlap among the seven yield classes (figure 3a) but significant differences exist among most yield classes (see test for homogeneity in supplementary material,

table S7). The widest ellipsoids include the lowest three yield classes (i.e. *campesino* agriculture), and the narrower ellipsoids contain commercially-oriented agriculture. A dendrogram from a Hierarchical Clustering of the PCA (figure 3b) shows that the municipalities can be grouped in k=3 clusters. Of these, two include most municipalities in yield classes ≤2 t/ha, while municipalities >2 t/ha form the other cluster. Within this cluster, there is a substantial mix of municipalities in different yield classes, showing an overlap of environmental conditions among them, except for municipalities in yield classes > 5 t/ha that are grouped together. Municipalities in higher yield classes are concentrated in relatively few regions of the country, and those in lower yield classes are widely scattered across Mexico, where most native landraces have been collected [8] (supplementary material figure S2).

(c) *Campesinos* contribute to maintain and generate maize genetic diversity

As shown above, 4.68 Mha were planted with rainfed maize in approximately 1.68 million farms in municipalities with average yields of up to 3 t/ha, representative of *campesino* agriculture. Assuming that around 30,000 plants/ha were grown [52], this results in 1.38×10^{11} genetically different individual plants, especially considering that Mexican maize landraces present a high diversity that is structured mostly according to the interaction of latitude and altitude and not race identity [9, 10, 53]. On each of those individual plants, open-pollination increases diversity through recombination. While most of the resulting seed of the 30,000 plants/ha are destined for human consumption or sale, in *campesino* agriculture a subset is set aside for planting in the next cycle. This is about 114-290 ears/ha depending on how it is calculated. Because seeds from the same ear have the same mother this reduces the population to that number of effective families (a sub-estimate, considering that fathers of each kernel would likely be different) passing to the next generation. Assuming our lowest estimate (supplementary material, table S8) of 114 maize families (ears) per ha in 4.68 Mha, this means an N_b of 5.24×10^8 plants contributing to the next generation with their background genetic diversity along with rare alleles. This N_b is of the same order of magnitude as the $N_e = 1.1 \times 10^8$ that has been estimated for Mexican maize (considering together low- and highlands maize estimated N_e [54]).

Selection of the ears constituting the N_b is done independently by around 2.52 million smallholders (considering about 1.68 million farms with 1.5 people per farm involved in the seed selection process), in different environments (figures 2-3), and in a process in which both men and women make selection decisions based on their own multiple criteria. Therefore, there are local

and regional differences regarding which types of ears are looked for and which individual plants would survive to produce seed, thus making the pool of N_b = 5.24x10^8 a diverse sample of the entire population. Considering this lower estimate of N_b, we estimate a decay of heterozygosity of 9.53x10^{-10} per generation (supplementary material, table S8). Depending on the number of maize ears per ha used to plant 4.68 Mha, an upper limit of 9.23x10^8 to 2.35x10^9 new mutations by generation can be estimated to occur in the coding regions of the maize genome (supplementary material, table S8).
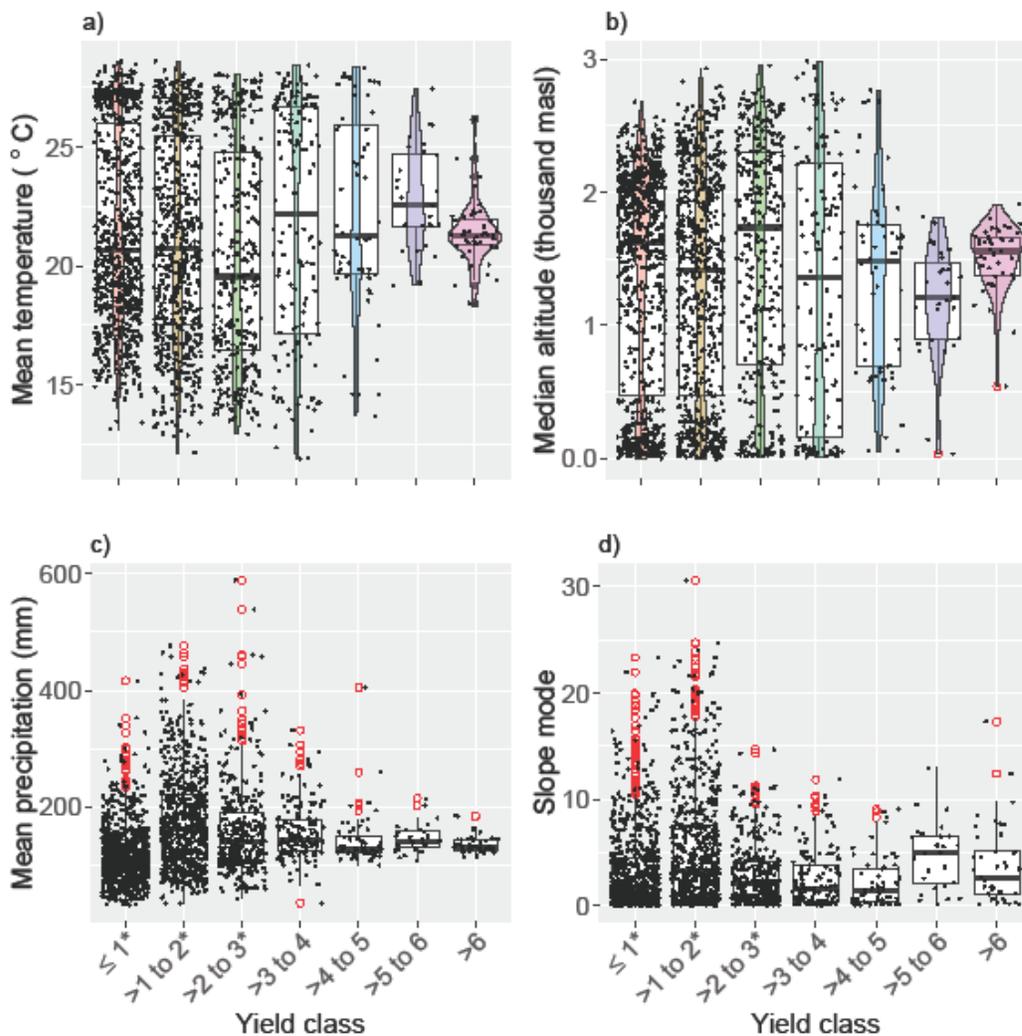

Figure 2. Box plots of the distribution of municipalities planted to rainfed maize in 2010 grouped by yield class for four environmental variables: a) mean temperature during rainy season (May-Oct); b) median altitude; c) mean precipitation during rainy season and d) slope mode. For a) and b) violin plots are shown on top of boxplots, showing that for the levels 1-3 the data has a bimodal distribution. *Yield classes mostly representing *campesino* agriculture. (See supplementary material, table S6 for equality of coefficients of variation test results).
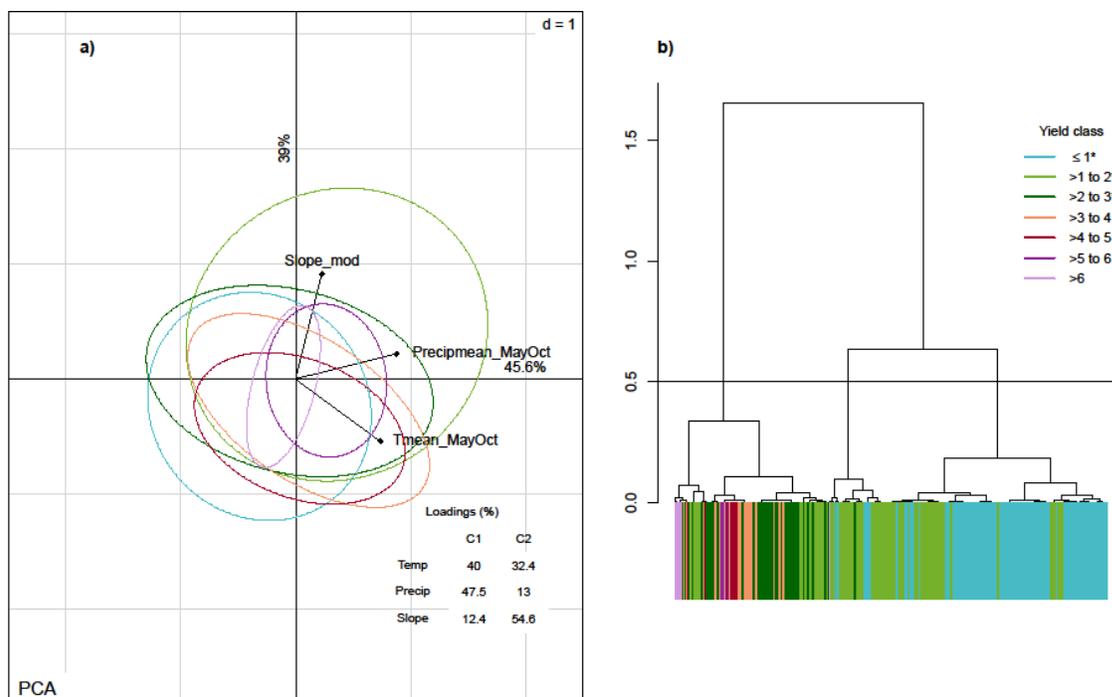
Figure 3. Environmental conditions of maize producing municipalities: a) Principal Component Analysis for the temperature, precipitation and slope environmental variables of municipalities grouped by yield classes (color codes and in b). Loadings are shown for the first two components (see supplementary material, table S7 for test for homogeneity). b) Dendrogram from a Hierarchical Clustering performed on the PCA. The horizontal line cuts in the suggested number of clusters (k=3). Colors represent municipalities grouped by yield classes. *Yield classes mostly representing *campesino* agriculture.

(d) *Campesino* maize farmers contribute importantly to the national maize food supply

Farmers representing mostly commercially-oriented agriculture (yields > 3 t/ha), produced enough to potentially feed 55.3 million people; but farmers representing mostly *campesino* agriculture (yields ≤3 t/ha), could potentially feed an additional 54.7 million people, including all the rural population in municipalities with *campesino* agriculture (21.1 million people), and with surpluses that could be used to feed an extra 33.6 million people (supplementary material, table S3). However, at a disaggregate level important differences appear (figure 4). In certain municipalities local production would be insufficient to feed all the rural population present, while in others surpluses would be available to feed additional people (figure 4a). Results show that within each yield class, the surplus maize produced, at least in theory, could contribute to cover any deficit, and there still could be enough to feed additional people elsewhere. These results also can be presented as number of municipalities with surpluses and deficits (figure 4b), with their spatial

distribution across the country (supplementary material, figure S3). The patterns described above are consistent for more than a decade for which municipal-level data are publicly available (2003 to 2015; supplementary material, figure S4).
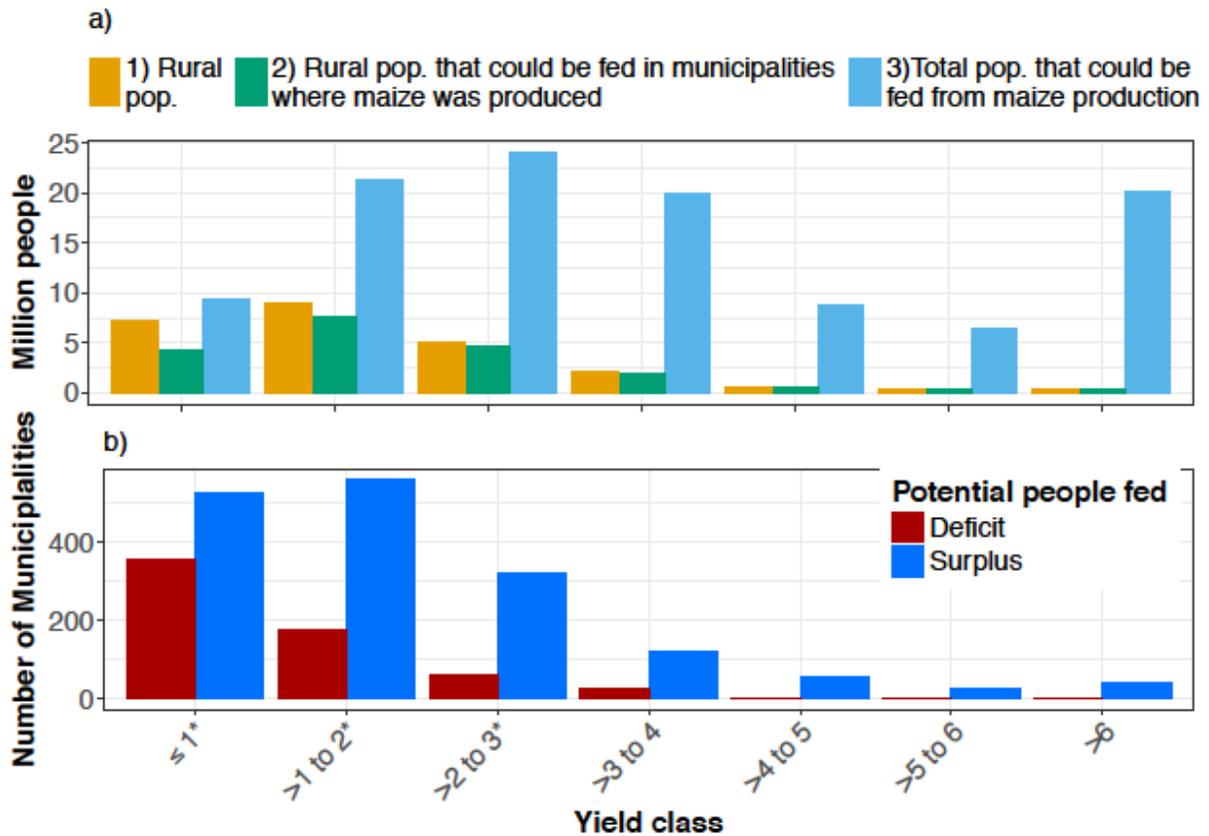

Figure 4. (a) Comparison by yield class, of the rural population present (bar1) with the estimated rural population that could be fed without (bar 2) and with local surpluses (3). The difference between the size of bars 1 and 2 corresponds to the size of the rural population that could not be fed from local production, but due to municipalities with surpluses that could cover any deficit, the estimated total population that could be fed (bar 3) would be larger than the rural population present (difference between bars 1 and 3). (b) Number of surplus and deficit municipalities based on whether local maize production could or could not feed the resident rural population. *Yield classes mostly representing *campesino* agriculture.

4. Discussion

(a) *Campesino* farmers contribute to maintain and generate maize genetic diversity, an "evolutionary service" important as a global public good

Genetic data suggest that maize Ne started to recover from its domestication bottleneck only around 1,000 years ago, with a dramatic expansion in the last 200 years [54]. This is not only congruent with the N_b of 5.24×10^8 we estimate for *campesino* maize, but also shows that the current time may represent the point of highest diversity in maize history [55]. Empirical studies have documented the evolutionary signatures of maize domestication (e.g. [5, 10, 54, 55]). However, those studies have focused on extant diversity, but not on the implications of the scope and extent in which *campesinos* produce maize in Mexico. Here, we suggest that *campesinos* not only contribute to maintaining the evolution of maize, but that the scale at which this is done becomes in itself an irreplaceable evolutionary or *evosystem* service (defined as "the uses or services to humans that are produced from the evolutionary process"[56]) [56, 57]. This *evosystem* service emerges because maize *campesino* agriculture combines in a single system three of the main factors known to positively affect adaptive evolution: large effective population size*,* high standing genetic diversity and environmental change [58-61], as follows:

(i) Our estimated slow decay of heterozygosity (9.53×10^{-10} per generation) causes a small loss of standing genetic diversity and rare alleles (frequency <1-5%), a primary source of adaptive genetic variation for crop domestication and breeding [55]. Conserving this diversity *in situ* thus increases the possibility that beneficial variation may be immediately available to experience new selective pressures [58], including alleles that can suffer selective sweeps [62], as well as several different alleles from which polygenic traits can undergo adaptive change [63].

 (ii) Growing a large amount of genetically diverse maize plants is useful not only for preserving the extant genetic diversity, but also for creating new diversity by means of mutation, as exemplified by the large estimated number of new mutations that could appear in the coding regions of the maize genome (9.23×10^8 to 2.35×10^9 per generation). Most new mutations would be neutral or detrimental to the mutant individual, but a small percentage may be beneficial and under the right selection pressure will increase their frequency rapidly [64-66].

(iii) Natural selection depends on local conditions. Given the wide range of environments where maize landraces are grown in Mexico (particularly in municipalities in yield classes ≤ 3 t/ha, figure 3b), it is likely that different alleles are being favored by different selection pressures, in a process that historically as well as presently has involved seed exchange and farmer's migration. This creates a setting of recurrent adaptation to a changing environment, which is a powerful driver of molecular adaptation [59].

To put these numbers into perspective, in 2010 ca. 41 Mha of maize were harvested in Canada, the US and Mexico together [67]. However, US and Canadian maize agriculture relies almost exclusively on hybrids, largely genetically modified varieties, whose seeds are bought each cycle and sown in large but genetically homogeneous areas. Additionally, although genetic diversity can be sourced into breeding materials from landraces [68], hybrids are reproduced avoiding gene flow among them and with landraces. Thus, the N_e and diversity of maize for the whole North American region comes mostly from the maize grown in the 4.68 million ha under Mexican *campesino* agriculture.

The large maize population maintained by *campesino* agriculture not only preserves genetic diversity, but also provides multiple opportunities for the appearance of potentially beneficial mutations. This is important since we cannot predict the future, and the presence of genetic diversity is fundamental for adaptive evolution in response to changing environmental conditions [68, 69]. Therefore, the larger the scale and the more diverse the scope of this process, the more genetic diversity available; and thus the higher the probability that some alleles, although currently rare or unknown, will become widely adaptive under new, and in many cases unpredictable conditions [19], that is, the higher the option value of genetic diversity under evolution. This constitutes not only an evosystem service, but also a positive externality —an uncompensated positive impact of one agent´s actions (i.e. *campesinos* growing maize) over the welfare of another (i.e. current and future maize consumers worldwide). If this diversity were no longer available, breeding programs would lack an important source of novel genetic variation to tap in order to allow continued genetic gains in a dynamic agricultural environment [68].

The evolutionary service provided by *campesino* agriculture cannot be replaced by the conservation of seeds in gene banks (*ex situ* conservation), because those seeds represent just a snapshot of the genetic diversity present at time of collection and are no longer under evolution [19]. Therefore, although ex situ conservation is a fundamental component of a crop conservation strategy, the capacity to generate and maintain adaptive genetic change in response to dynamic environmental conditions cannot be accomplished by gene banks alone. Similarly, for in situ conservation to be successful we should consider that (a) large maize populations are needed, as suggested by our population genetics calculations, and (b) these populations should be spread along different environments, because the loss environmental range may imply the loss of genetic diversity [69]. Considering the extent and scope of Mexican *campesino* agriculture, this cannot be

achieved by targeting just a handful of farmers in a few locations to conserve maize races based in their phenotypes [70].

(b) *Campesino* maize farmers are crucial for national food security

Campesinos also make a crucial contribution to the maize supply of Mexico. While individual *campesino* households may not produce enough for their needs [23, 33], as a group they produce important surpluses, which have the potential to not only feed themselves, but also to make a substantial contribution to feed others in their communities and regions.

While *campesinos* produce maize due a multiplicity of values they ascribe to this crop, particularly to their native landraces, these values tend to be systematically ignored by markets [71], so their decisions to produce maize may appear uneconomical and inefficient when judged by standard financial profitability criteria. The discrepancy between financial profitability and the willingness of farmers to produce maize reflects a market that fails to incorporate the multiple values these farmers ascribe to this crop into market prices [71]. Hence, the large number of *campesinos* that produce maize under highly diverse, and in many instances adverse environmental conditions, can be seen as producing another positive externality of national relevance: more maize produced and people fed than would have occurred if farmers were purely profit-driven.

Our results should not be interpreted as implying that rainfed maize production is feeding the entire population of rural areas. Available data simply indicate the potential to achieve this. To put these results in perspective, in 2010 Mexico produced 23.3 million tons (MT) of mostly white maize, making it self-sufficient in maize for human consumption, and imported 7.7 MT of mostly yellow maize intended for animal feed. *Campesino* agriculture contributed 25.5% of the total national production, with the potential to feed 48.7% of all the Mexican population (supplementary material, tables S10 and S3).

(c) Strengthening *campesino* agriculture and its positive externalities

Our results show a yield gap between *campesino* and commercially-oriented agriculture. This gap is normally attributed, at least partially, to the extensive reliance of the former on native landraces and on improved varieties by the latter. However many native landraces can provide reasonably high yields with modest amounts of inputs and small changes in agronomic management (supplementary material, figure S5a and table S9) and are competitive or even out-yield hybrids in

many environments ([70]; supplementary material, figure S5b), so the observed low productivity is not inherently related to the use of native landraces. Moreover, the remarkable genetic diversity and adaptability of native landraces to wide-ranging agro-ecological conditions [52, 72] allow *campesinos* to produce maize at the wide scale and in areas where improved varieties are unlikely to be adapted [14, 73].

Despite the negative perception caused by this yield gap and minimal support from agricultural public policies, Mexican *campesinos* generate positive externalities of national and global relevance that should be considered public goods and supported as such, particularly since there is no guarantee that they will continue to provide these goods in the future. No effective policies and programs exist in Mexico to reward and support these farmers´ contributions. Some practices related to local education, development and marketing of products from landraces with added value, and other practices supporting *in situ* conservation have been promoted and need to be supported at a larger scale [74]. However, to truly encompass the extent of maize genetic diversity, what needs to be supported is *campesino* agriculture as a production system that generates conservation outcomes. Some ways of fulfilling this include: (a) exploring new approaches to improve the yields and quality of native varieties under the conditions that *campesinos* face, so as to improve their productivity while still keeping them under evolution. Agronomic interventions such as participatory breeding, or the addition of modest amounts of inputs and slight changes in agronomic management compatible with these farmers´ circumstances, could make a significant contribution to improving their productivity; (b) improving our understanding of maize trading at the local level and how surpluses reach consumers, either through markets or other exchange mechanisms. This can be the basis for identifying interventions to improve trading conditions; (c) understanding the conditions that foster farmers to continue the management of this diversity, as well as the circumstances that hinder it, in order to design better interventions to support them; (d) instead of a public policy that assumes that these farmers should be moved out of production, we should recognize that they will continue to exist, so that research and extension could focus on addressing their needs and improving their wellbeing; (e) refining population genetics models to better quantify *campesinos´* contribution for generating and maintaining maize genetic diversity, as the basis for monitoring and safeguarding these processes.

Besides the former recommendations, the first step towards strengthening *campesino* agriculture is to recognize their positive contributions to crop evolution and food security, as done here. As evolution under domestication continues to occur in other crops in other countries—especially in regions of high agrobiodiversity—smallholders´ contributions in those regions merit closer examination.

References

1. Larson G *et al*. 2014 Current perspectives and the future of domestication studies. *Proc. Natl Acad. Sci.* **111**, 6139–6146. (doi: /10.1073/pnas.1323964111)

2. Brush SB. 2004 *Farmers' bounty. Locating crop diversity in the contemporary world*. New Haven, CT: Yale University Press.

3. Doebley J. 2004 The genetics of maize evolution. *Ann. Rev. Genet.* **38**, 37-59. (doi: 10.1146/annurev.genet.38.072902.092425)

4. Kato TA, Mapes C, Mera ML, Serratos JA, Bye RA, eds. 2009 *Origen y diversificación del maíz: una revisión analítica*. México DF: Universidad Nacional Autónoma de México, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. http://www.biodiversidad.gob.mx/genes/pdf/origen_div_maiz.pdf.

5. Hufford MB *et al*. 2012 Comparative population genomics of maize domestication and improvement. *Nature Genet* **44**, 808–11. (doi:10.1038/ng.2309)

6. Shiferaw B, Prasanna BM, Hellin J, Bänziger M. 2011 Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Secur.* **3**, 307-327. (doi.org/10.1007/s12571-011-0140-5)

7. Fernandez-Suarez R, Morales-Chavez LA, Galvez-Mariscal A. 2013 Importancia de los maíces nativos de México en la dieta nacional: una revisión indispensable. *Rev. Fitotec. Mex.* **36** Suppl 3-A, 275-283.

8. CONABIO. 2011 *Recopilación, generación, actualización y análisis de información acerca de la diversidad genética de maíces y sus parientes silvestres en México*. México DF: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. http://www.biodiversidad.gob.mx/genes/pdf/proyecto/InformedeGestion_V1.pdf)

9. Arteaga MC, Moreno-Letelier A, Mastretta-Yanes A, Vazquez-Lobo A, Breña-Ochoa A, Moreno-Estrada A, Eguiarte LE, Piñero D. 2016 Genomic variation in recently collected maize landraces from Mexico. *Genom. Data* **7**, 38–45. (doi.org/10.1016/j.gdata.2015.11.002)

10. Romero-Navarro JA, Willcox M, Burgueño J, Romay C, Swarts K, Trachsel S, Preciado E, Preciado E, Terron A, Vallejo H et al. 2017 A study of allelic diversity underlying flowering-time adaptation in maize landraces. *Nature Genet.* **49,** 479-480. (doi:10.1038/ng.3784)

11. Mera-Ovando L M, Mapes-Sánchez C. 2009 El maíz. Aspecto biológicos. In *Origen y diversificación del Maíz: una revisión analítica* (eds Kato TA, Mapes C, Mera LM, Serratos JA, Bye RA), pp. 19-32. México DF: Universidad Nacional Autónoma de México, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Editorial Impresora Apolo.

12. SIAP. 2007 Situación actual y perspectivas del maíz en México 1996–2012. México DF: Sistema de Información Agroalimentaria y Pesquera. Available at: http://www.campomexicano.gob.mx/portal_siap/Integracion/EstadisticaDerivada/ComercioExterior/Estudios/Perspectivas/maiz96-12.pdf.

13. Bellon MR, Brush SB. 1994 Keepers of maize in Chiapas, Mexico. *Econ. Bot.* 48, 196-209. (doi.org/10.1007/BF02908218)

14. Perales H, Brush SB, Qualset CO. 2003 Landraces of maize in Central Mexico: an altitudinal transect*. Econ. Bot.* **57**, 7-20. (doi.org/10.1663)

15. Pressoir G, Berthaud J. 2004 Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity* **92**, 88–94. (doi: 10.1038/sj.hdy.6800387)

16. Pressoir G, Berthaud J. 2004 Population structure and strong divergent selection shape phenotypic diversification in maize landraces. *Heredity* **92**, 95–101. (doi: 10.1038/sj.hdy.6800388)

17. Perales HR, Benz BF, Brush SB. 2005 Maize diversity and ethnolinguistic diversity in Chiapas, Mexico. *Proc. Natl. Acad. Sci. USA* **102**, 949–954. (doi.org/10.1073/pnas.0408701102)

18. Orozco-Ramírez Q, Ross-Ibarra J, Santacruz-Varela A, Brush S. 2016 Maize diversity associated with social origin and environmental variation in southern Mexico. *Heredity* **116**, 477–84. (doi: 10.1038/hdy.2016.10)

19. Bellon MR, Dulloo E, Thormann I, Sardos J, Burdon J. 2017 In-situ conservation – harnessing natural and human derived evolutionary forces to ensure future crop adaptation. *Evol. Appl.* **10**, 965-977.DOI: 10.1111/eva.12521.

20. Diario Oficial de la Federación. 2013 DECRETO por el que se aprueba el Programa Sectorial de Desarrollo Agropecuario, Pesquero y Alimentario 2013 - 2018. Available at: http://www.sagarpa.gob.mx/ganaderia/Documents/2015/MANUALES%20Y%20PLANES/Programa_Sectorial_SAGARPA_2013-2018%20(1).pdf.

21. Fox J, Haight L. 2010 Mexican agricultural policy: multiple goals and conflicting interests. In *Subsidizing inequality: Mexican corn policy since NAFTA* (eds. Fox J, Haight L), pp. 9-50. Santa Cruz, CA: Woodrow Wilson International Center for Scholars.

22. Levy S, van Wijnbergen S. 1992 Maize and the Free Trade Agreement between Mexico and the United States. *World Bank Econ. Rev.* **6**, 481-502. (doi: 10.1093/wber/6.3.481)

23. Eakin H, Perales H, Appendini K, Sweeney S. 2014 Selling maize in Mexico: the persistence of peasant farming in an era of global markets. *Dev. Change* **45**, 133-155. (doi: 10.1111/dech.12074)

24. Yuñez Naude A. 2010 Las políticas públicas dirigidas al sector rural: el cáracter de las reformas para el cambio estructural. In *Los grandes problemas de México. XI economía rural* (ed. Yuñez-Naude A), pp. 23-62. México, DF: El Colegio de México.

25. Eakin H, Appendini K, Sweeney S, Perales H. 2015 Correlates of maize land and livelihood change among maize farming households in Mexico. *World Dev.* **70**, 78-91. (doi: 10.1016/j.worlddev.2014.12.012)

26. de Janvry A, Sadoulet E. 2001 Income strategies among rural households in Mexico: the role of off-farm activities. *World Dev.* **29**, 467-480. (doi: 10.1016/S0305-750X(00)00113-3)

27. Yúnez-Naude A, Rivera-Ramírez F, Chávez-Alvarado MA, Mora-Rivera JJ, Taylor JE, eds. 2015 *La Economía del Campo Mexicano: Tendencias y Retos para su Desarrollo*. México DF: El Colegio de México.

28. Bellon MR. 1996 The dynamics of crop infraspecific diversity: conceptual framework at the farmer level. *Econ. Bot.* **50**, 26-39. (doi 10.1007/BF02862110)

29. Appendini K, Quijada MG. 2016 Consumption strategies in Mexican rural households: pursuing food security with quality. *Agric. Hum. Values* **33**, 439-454. (doi.org/10.1007/s10460-015-9614-y)

30. Dyer G, Taylor JE, Boucher S. 2006 Subsistence response to market shocks. *Am. J. Agric. Econ.* **88**, 279–291. ([doi.org/10.1111/j.1467-8276.2006.00858.x](doi.org/10.1111/j.1467-8276.2006.00858.x))

31. Hellin J, Orenstein O, Beuchelt T, Camacho C, Flores D. 2013 Maize stover use and sustainable crop production in mixed crop-livestock systems in Mexico. *Field Crop Res.* **153**, 12-21. ([//doi.org/10.1016/j.fcr.2013.05.014](//doi.org/10.1016/j.fcr.2013.05.014))

32. Keleman A, Hellin J. 2009 Specialty maize varieties in Mexico : a case study in market-driven agro-biodiversity conservation. *J. Lat. Amer. Geogr.* **8**, 147–74. ([doi.org/10.1016/j.fcr.2013.05.014](doi.org/10.1016/j.fcr.2013.05.014))

33. Gordillo de Anda G, de Janvry A, Sadoulet E (1999*) La Segunda Reforma Agraria de México: Respuestas de Familias y Comunidades, 1990-1994*. México DF: Fondo de Cultura Económica.

34. SIAP. 2016 *Norma técnica para la generación de estadística básica agropecuaria y pesquera*. México DF: Secretaria de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación, pp. 348-355. http://www.gob.mx/siap/normatividad-estadistica/.

35. Cuervo-Robayo AP, Tellez-Valdez O, Gomez-Albores MA, Venegas-Barrera CS, Manjarrez J, Martinez-Meyer. 2014 An update of high-resolution monthly climate surfaces for Mexico. *Int. J. Climatol.* **34**, 2427–2437. ([doi.org/10.1002/joc.3848](doi.org/10.1002/joc.3848))

36. Guevara M, Arroyo-Cruz CE. 2016 *Modelo digital de elevaciones corregido resolución 1km, escala: 1:4000000, Edición: 1*. Cd de México: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.

37. Guevara M, Arroyo-Cruz CE. 2016 *'Pendiente', Escala: 1:4000000, Edición: 1*. Cd de México: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.

38. INEGI. 2015 *Conjunto de datos vectoriales de uso de suelo y vegetación Escala 1:250 000, Serie V* (capa unión). México DF: Instituto Nacional de Estadística y Geografía.

39. Feltz CJ, Miller GE. 1996 An asymptotic test for the equality of coefficients of variation from k populations. *Stat. Med. 15*, 647-658. (doi.org/10.1002/(SICI)1097-0258(19960330)15:6<647::AID-SIM184>3.0.CO;2-P)

40. Anderson, MJ. 2006 Distance-based tests for homogeneity of multivariate dispersions. *Biometrics* **62**, 245–253. (doi: 10.1111/j.1541-0420.2005.00440.x)

41. R. Core Team. 2012 R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria) available at: http://www.R-project.org/.

42. Hijmans RJ *et al*. 2014 raster: geographic data analysis and modeling. Available at: http://cran.at.r-project.org/web/packages/raster/index.html.

43. Bivand R *et al*. 2015 maptools: tools for reading and handling spatial objects. Available at: http://cran.r-project.org/web/packages/maptools/index.html.

44. Marwick B, Krishnamoorthy K. 2016 cvequality: tests for the equality of coefficients of variation from multiple groups. Available at: https://cran.r-project.org/web/packages/cvequality/index.html.

45. Oksanen J, *et al*. 2016 vegan: community ecology package. Available at: https://cran.r-project.org/web/packages/vegan/index.html.

46. Husson F, Josse J, Le S, Mazet J. 2017 FactoMineR: multivariate exploratory data analysis and data mining. Available at: https://CRAN.R-project.org/package=FactoMineR.

47. Caballero A. 1994 Developments in the prediction of effective population size. *Heredity* **73**, 657–679. (doi:10.1038/hdy.1994.174)

48. Gillespie JH. 1998 *Population genetics: a concise guide*. Baltimore Md: The Johns Hopkins University Press.

49. Jiao Y, Zhao H, Ren L, Song W, Zeng B, Guo J, Wang B, Liu Z, Chen J, Li W, Zhang M, Xie S, Lai J. 2012 Genome-wide genetic changes during modern breeding of maize. *Nature Genet. 44*, 812-815. (doi: 10.1038/ng0914-1039)

50. Schnable PS *et al*. 2009 The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**, 1112–1115. (doi: 10.1126/science.1178534)

51. CONEVAL. 2012 Construcción de las líneas de bienestar. Documento metodológico. México DF: Consejo Nacional de Evaluación de la Política de Desarrollo Social.

52. Mercer KL, Martínez-Vásquez A, Perales HR. 2008 Asymmetrical local adaptation of maize landraces along an altitudinal gradient. *Evol. Appl.* **1**, 489–500. (doi: 10.1111/j.1752-4571.2008.00038.x)

53. Ruiz Corral JA, Durán Puga N, González Sánchez JJ, Ron Parra J, González Eguiarte DR, Holland, JB, Medina García M. 2008 Climatic adaptation and ecological descriptors of 42 Mexican maize races. *Crop Sci.* **48**, 1502–12. (doi: 10.2135/cropsci2007.09.0518)

54. Wang L, Beissinger TM, Lorant A, Rossi-Ibarra C, Ross-Ibarra J, Hufford MB. 2017 The interplay of demography and selection during maize domestication and expansion. *Genome Biol.* **1**, 215. (doi:10.1186/s13059-017-1346-4)

55. Beissinger TM, Wang L, Crosby K, Durvasula A, Hufford MB, Ross-Ibarra J. 2016 Recent demography drives changes in linked selection across the maize genome. *Nat. Plants* **2**, 16084. (doi:10.1038/nplants.2016.84)

56. Faith DP, Magallon S, Hendry AP, Conti E, Yahara T, Donoghue MJ. 2010 Evosystem services : an evolutionary perspective on the links between biodiversity and human well-being. *Curr. Opin. Environ. Sustain.* **2**, 1–9. (doi 10.1016/j.cosust.2010.04.002)

57. Bellon MR, Gotor E, Caracciolo F. 2015 Conserving landraces and improving livelihoods: how to assess the success of on-farm conservation projects? *Int. J. Agr. Sustain.* **13**, 167-182. (doi: 10.1080/14735903.2014.986363)

58. Barrett R, Schluter D. 2008 Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44. (doi.org/10.1016/j.tree.2007.09.008)

59. Lourenço JM, Glémin S, Galtier N. 2013 The rate of molecular adaptation in a changing environment. *Mol. Biol. Evol.* **30**, 1292–1301. (doi: 10.1093/molbev/mst026)

60. Gossmann TI, Keightley PD, Eyre-Walker A. 2012 The effect of variation in the effective population size on the rate of adaptive molecular evolution in Eukaryotes. *Genome Biol. Evol.* **4**, 658–667. (doi: 10.1093/gbe/evs027)

61. Lanfear R, Kokko H, Eyre-Walker A. 2014 Population size and the rate of evolution. *Trends Ecol. Evol.* **29**, 33–41. (doi.org/10.1016/j.tree.2013.09.009)

62. Dwivedi SL, Ceccarelli S, Blair MW, Upadhyaya HD, Are AK, Ortiz R. 2016 Landrace germplasm for improving yield and abiotic stress adaptation. *Trends Plant Sci.* **21**, 31–42. (doi: 10.1016/j.tplants.2015.10.012)

63. Pritchard JK, Di Rienzo A. 2010 Adaptation – not by sweeps alone. *Nat. Rev. Genet.* **11**, 665–667. (doi: 10.1038/nrg2880)

64. Gillespie JH. 1984 Molecular evolution over the mutational landscape. *Evolution* **38**, 1116–1129. (doi: 10.2307/2408444)

65. Orr HA. 2003 The distribution of fitness effects among beneficial mutations. *Genetics* **163**, 1519–1526. (doi: doi.org/10.1016/j.jtbi.2005.05.001)

66. Eyre-Walker A, Keightley PD. 2007 The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* 8:610–618. (doi: 10.1038/nrg2146)

67. FAOSTAT. 2018 Statistics Division. Available at: http://faostat.fao.org. Accessed July 2018

68. Smith JSC, Gardner CAC, Costich DE. 2017 Ensuring the genetic diversity of maize and its wild relatives. In *Achieving sustainable cultivation of maize. volume 1: from improved varieties to local applications* (ed. Watson D), pp. 3-50. Cambridge, UK: Burleigh Dodds Science Publishing Ltd. DOI 19103/AS.2016.0001.02

69. Bruford MW, Davies N, Dulloo ME, Faith DP, Walters M. 2017 Monitoring changes in genetic diversity. In *The Geo handbook on biodiversity observation networks* (eds. Walters M, Scholes RJ), pp 107-128. Springer Nature: Cham, DOI 101007/978-3-319-27288-7_5.

70. Perales H. 2016 Landrace conservation of maize in Mexico and evolutionary breeding. In *Enhancing crop genepool use: capturing wild relative and landrace diversity for crop improvement* (eds Maxted N, Dulloo ME, Ford-Lloyd BV), pp. 271-281. Wallingford UK: CAB Intl.

71. Arslan A, Taylor JE. 2009 Farmers' subjective valuation of subsistence crops: the case of traditional maize in Mexico. *Am. J. Agric. Econ.* **91**, 956–72. (doi: doi.org/10.1111/j.1467-8276.2009.01323.x)

72. Perales H, Golicher D. 2014 Mapping the diversity of maize races in Mexico. *PLoS One* 9(12): 1–20. doi:10.1371/journal.pone.0114657.

73. Bellon MR, Hodson D, Bergvinson D, Beck D, Martinez-Romero E, Montoya Y. 2005 Targeting agricultural research to benefit poor farmers: relating poverty mapping to maize environments in Mexico. *Food Policy* **30**, 476-492. (doi: 10.1016/j.foodpol.2005.09.003)

74. CONABIO. 2017 *Ecosystems and agro-biodiversity across small and large-scale maize production systems*. Geneva: TEEB Agriculture & Food, UNEP. Available at http://teebweb.org/agrifood/home/maize/

75. Bellon MR, Mastretta-Yanes A, Ponce-Mendoza A, Ortiz-Santamaría D, Oliveros-Galindo O, Perales H, Acevedo F, Sarukhan J. 2018 Data from: Evolutionary and food supply implications of ongoing maize domestication by Mexican *campesinos*. Dryad Digital Repository. (doi:10.5061/dryad.79q870b)