

**Seed and seedling characteristics of chile pepper (*Capsicum annuum* L.) across a
domestication gradient**

Katelyn Fulcher

Advised by Dr. Kristin Mercer

Department of Horticulture and Crop Science

The Ohio State University

November 2023

Table of Contents

CHAPTER 1: Introduction: Effects of domestication and location of origin on chile pepper phenotypes	3
CHAPTER 2: Seed and seedling characteristics of chile pepper (<i>Capsicum annuum</i> L.) across a domestication gradient.....	13
Abstract	13
Introduction	13
Methods and Materials	17
Results	25
Discussion	30
Acknowledgements	38
Literature Cited	39
Tables	47
Figures	55
Supplementary Materials.....	62

CHAPTER 1: Introduction: Effects of domestication and location of origin on chile pepper phenotypes

Capsicum

Chile peppers (*Capsicum annuum* L.) are of great significance to people around the world. First domesticated in approximately 5000-6000 B.C. and used as early as 10,000 years ago, this crop is both an invaluable food source and component of many important cultural practices (Bridgemohan et al., 2017; Kraft et al., 2014; Aguilar-Meléndez, 2021). During the several thousand years of cultivation of this crop, chile peppers were distributed around the world through extensive trading and are now grown far beyond their geographical center of origin in Mexico. Five domesticated species of the genus *Capsicum* exist, with *Capsicum annuum* maintaining the highest amount of diversity in structural morphology (Aguilar-Meléndez et al., 2009). Seed traits vary widely across *Capsicum*, with traits being so closely characteristic of each species that they can be used for positive identification to a species level (Chiou & Hastorf, 2014). The extensive cultivation of chiles has led to a great diversity in plant characteristics. The evolution of this crop is human-mediated and is ongoing, with many varieties and landraces preserved for local preferences or traditions (Perez-Martinez et al., 2022). Chile peppers are grown for a wide variety of purposes, from culinary, to medicinal, to ornamental, among countless others. Chiles are cultivated across a wide diversity of regions and ecozones, with landrace varieties of *Capsicum annuum* expected to be adapted to their local environment. Landraces are local varieties of crops that have been cultivated over time through traditional agricultural practices and are often selected for by farmers for these local adaptations (Villa et al., 2005). The genetic diversity of certain landraces can be exceptionally high due to wild and semiwild plants being present within the agroecosystem, allowing for cross-breeding

(Taitano et al., 2018). This is true for other crop species as well, with such introgression of wild genetic material into crop species introducing new genetic material and increasing variation (Khoury et al., 2021).

Domestication

The process of domestication of wild plants to cultivated crops changes morphology and growth habit (Flint-Garcia, 2013). The domestication process consists of stages, with wild plants being the starting point with a progression over time to landraces and improved varieties (Meyer and Purugganan, 2013). Typical domestication traits include increased fruit size, ease of harvest, and improved flavor, among many others (Sun et al., 2017). Wild chile peppers are typically characterized by their smaller fruit size, which are deciduous and fall to the ground after ripening. This is a trait bred out of cultivated chile fruits, which remain on the peduncle after maturity (Aguilar-Meléndez et al., 2009). Seed characteristics such as size, shape, and mass may correlate to the level of domestication as well, with many crops producing larger seed sizes than their wild ancestors (Spengler, 2020). Landrace accessions vary widely in fruit color, size, and shape, depending on the desired culinary use (Naegele et al., 2016). Differences in height and branching patterns also vary widely across accessions of *C. annuum*, though this may be closer tied to the cultivation setting rather than the level of domestication itself (McCoy et al., 2022).

The genetic diversity of crops can be greatly affected by domestication. As a whole, the domestication syndrome in plants often leads to reduced genetic diversity when compared to wild ancestors of the species, as in the case of many major crops such as wheat, soybean, and tomato (Shi & Lai, 2015). These genetic bottlenecks can be problematic as overlooked beneficial traits become lost in commercial varieties, such as the abiotic and biotic stress resistance that

wild ancestors evolved. This may leave crops vulnerable to pests, pathogens, extreme temperatures, drought, flooding, and more (Esquinas-Alcázar, 2005).

Cultivation system

The cultivation type of chile peppers can be divided into three main categories. In Mexico, wild chiles (*C. annuum* var. *glabriusculum*) can be found in forest settings and are uncultivated, and, though they may be foraged occasionally, there is no intentional planting or care for these plants by humans (Aguilar-Meléndez et al., 2009). The second type, semi-wild chiles, can be found in backyard settings, where they may be “let-standing” or experiencing some degree of cultivation, in which they are likely to be naturally dispersed and then tended for harvesting of the fruits. They can also be commonly found on edges such as roadsides (Aguilar-Meléndez et al., 2009). The third type, cultivated chiles, are typically grown in two cultivation settings, which are milpa and plantation. Milpas are cultivated fields that are often smaller in size and include a variety of different crop species that may be grown together. Landraces in milpa systems are likely to have encountered little formal breeding. By contrast, plantations are farm fields with higher levels of mechanization and inputs where farmers often use improved varieties rather than landraces (Bernau et al., 2020). While these categorizations are very straightforward, many chile peppers can have characteristics that are indicative of a combination of these cultivation settings, rather than a single clear-cut category. Cultivation settings can be related to the level of domestication although they do not always correspond. While plantations may use improved varieties of crops in many cases, this is not the only setting in which improved varieties are cultivated. In addition, landrace varieties may be cultivated in settings such as backyards (Bernau et al., 2020).

Cultivation systems can play a large role in the maintenance of crop genetic diversity. With chile peppers, landrace crops are an important source of genetic variation for farmers. This can be due to the presence of wild chile peppers growing often together or alongside the landraces, allowing for cross-pollination between both wild and landrace populations (Perez-Martinez et al., 2022). In addition, the exchange of germplasm of chile cultivars between farmers allows even further genetic diversity to be maintained despite cultivation (González-Jara, 2011). Compared to landraces, more intensive cultivation systems with greater artificial selection for desired genes can lead to an even greater loss in genetic diversity of crops (Flint-Garcia, 2013). Unlike the monocultures of many major crops seen in the United States, this close proximity of wild chiles to the cultivated crops allows for maintained genetic diversity in many cases (Aguilar-Meléndez et al., 2009).

Geographic origin and genetic diversity

Characteristics of *C. annuum* also differ by geographic origin. Plants are often highly adapted to the environment in which they live, oftentimes with adaptation to local factors such as moisture, elevation, soil types, and local biotic communities. This can be especially evident in crop landraces, which develop adaptations due to natural and human-mediated selection over time (Villa et al., 2007). Associations between a plant's genome and its environment were shown in sorghum landraces. Alleles responsible for tannin presence were found in plants living in cool, wet climates that favor mold growth, functioning as a resistance for grain mold. This shows that the distribution of this allele is altered by geographical conditions (Lasky et al., 2015). Maize landraces from highland locations were shown to be especially sensitive to elevation, with a much higher chance of higher quality seed being produced when grown at higher elevation compared to lower elevation. Maize adapted to lowland regions showed the opposite, with a 25%

reduction in seed quality when grown in higher elevations (Mercer et al., 2008). In another study, *C. annuum* accessions sourced from drier ecozones germinated slower than those from wetter ecozones, with the delayed germination possibly functioning as an avoidance response to drought injury (Bernau et al., 2020). Geographic origin may also have an impact on overall genetic diversity. Genetic diversity of chile peppers was shown to be higher in some geographic regions than others, with the Yucatan Peninsula region having the highest mean genetic diversity and a significant drop in diversity across western Mexico (Aguilar-Melendez et al., 2009).

Due to such local adaptations, the genetic diversity within a single crop species can be vast. This is typically the highest within a species' geographic center of origin. These geographic centers typically harbor extensive gene pools in which natural selection and adaptation are ongoing, which is further amplified in regions where commercial agriculture is not the main method used and, instead, diversity in cultivation practices is common (Engels et al., 2006). Mexico is one such region, with multiple independent domestication events occurring for *C. annuum* from widespread populations (Aguilar-Melendez et al., 2009). Central America is also known as a center of crop diversity, with over 50 separate crop species having been domesticated from this region (Engels et al., 2006). This increased genetic diversity allows for more extensive local adaptation and increased presence of novel alleles within the gene pool.

During a time of climate change, it is especially important that we maintain or reintroduce these sources of resistance from landraces and crop-wild ancestors into our crops (Khoury et al., 2019; McCoy et al., 2022). Plant breeding that reintegrates sources of genetic diversity into crops has been shown to increase the adaptation of many crop species to climate-related abiotic stressors such as higher temperatures and drought (Temesgen, 2017). Some successful examples of this application into field studies include the use of heat and drought

tolerant cultivars of cotton showing increased population numbers (Rahman et al., 2018) and the use of heat tolerant cultivars of winter wheat showing increased yield (Bergkamp et al., 2018), both despite exposure to heat stressors.

Crop landraces contain the bulk of genetic diversity in domesticated species and are considered to be a highly threatened form of genetic resources (Illa et al., 2005; Brush and Meng, 1998). Re-integration of useful alleles into crop species holds great value in agriculture, with integration allowing for the creation of improved varieties of crops that hold a number of uses. In the Mediterranean region, a tomato landrace known for extensive shelf-life and drought tolerance has been identified as an extremely important genetic resource to be integrated into breeding programs (Conesa et al., 2020). In rice, bacterial blight resistance genes were found in landraces of basmati rice, which were successfully incorporated into development of resistant cultivars (Ullah et al., 2012).

While landrace crops are known for their local genetic adaptations and overall increased genetic diversity when compared to plants originating from modern breeding systems, crop wild relatives are also important sources of genetic material (Benlioglu and Adak, 2019). Wild crop ancestors are known for their resistances to abiotic and biotic factors, with many of our main crop species being unable to be cultivated in a commercial setting without these resistances from wild ancestors being incorporated (Benlioglu and Adak, 2019). During a time of climate change and rapidly increasing populations to feed, the ability of our crops to withstand these changes is essential. Resistances to pests and pathogens are especially valued, making up 80% of traits bred into crops from wild ancestors (Bohra et al., 2022). Rapid advancements in plant breeding and molecular biology have allowed for integration of these traits despite challenges (Bohra et al., 2022).

Yet, in order for us to make use of the landrace and wild genetic resources, we must ensure that they remain available for us to use. Conservation of genetic resources can be broken up into two types. *In situ* conservation involves the conservation of these plants within their habitat, whether that be on a farm or in a wild setting, while *ex situ* conservation occurs once the genetic material is collected as germplasm to be maintained in a collection or a gene bank (Engels et al., 2006). Each of these methods are needed for adequate conservation of these valuable resources.

Ecologically relevant traits in seeds

Seed morphology is closely tied to its ecology in a number of ways. Seed size plays a large role in seedling survival, with higher seed mass correlating to high germination success in tropical rain forest tree species (Uft, 2004). Larger seed size may also be correlated to the survival of the plant throughout its life cycle, not just during the seedling stage (Metz et al., 2010). Across multiple species of New Zealand herbaceous flora, species with smaller and more rounded seeds were shown to have a higher persistence in the soil (Moles et al., 2003). Seed traits can also be an important indicator of their quality and composition. This is important for agriculture as the quality of a seed plays an important role in the ability of the seed to germinate and withstand both abiotic and biotic stressors (ElMasry et al., 2019). Seed quality monitoring with a focus on traits such as size, mass, or other qualities may help to ensure the survival of crop plants as a result (Saatkamp et al., 2018).

Distinguishable qualities include seed size, shape, color, and mass. Seed size can be quantified as seed area, perimeter, length, width, and diameter. Such quantitative traits are easily measured by use of rulers, calipers, and scales. Traits such as color or roughness are more difficult to describe, and such descriptive methods often differ from one observer to the next.

Seed shape and color can be especially difficult to describe in a way that is understandable across species, as the terms are chosen for the seeds from the species being measured only and are highly subjective (Baek et al., 2020). Other highly subjective descriptors to characterize the angle, curvature, and feeling of the seed coat as rough or smooth are also used in such traditional phenotyping (Robocker, 1970).

With the implementation of digital techniques into seed quality assessments, many of the more ambiguous traits can be quantified. Traditional measurements of characteristics such as length and width of the seed often leave much room for error in hand measurements; it is also hard to increase throughput when using hand measurements. Digital analyses software and imaging technology allow for a method of measuring seeds for such characteristics with higher accuracy, detail, and speed (Tanabata, et al., 2012). Shape qualities that can be measured with digital tools include sphericity, angularity, aspect ratio, beak angle, and length, width, and area of the attachment scar (Chiou & Hastorf, 2014). Color can be quantified as values of hue (e.g., red) and lightness intensity (e.g., bright vs. dark), and qualities of the seed texture can even be measured as a quantified surface roughness value (Laddo et al., 2022).

Ecologically relevant traits in seedlings

The seedling stage of a plant is arguably the most vulnerable point of its life cycle. Seedling mortality can occur for a variety of reasons, both biotic and abiotic. Moisture stress, extreme temperature, herbivory, and pathogens are just a few of the primary causes. Some seedlings have adaptive responses that allow for increased survival in challenging environments, such as greater shade acclimation, increased chlorophyll content, greater root biomass, and more (Khurana and Singh, 2001; Zipperlen and Press, 1996). Larger cotyledon size of seedlings has also been associated with the survival of seedlings in nutrient poor soils (Milberg and Lamont,

1997). Such adaptations are extremely important for the survival of seedlings, especially in harsh environments.

Ecologically relevant traits in roots

The root architecture is integral to plant functionality, with the differences in ratio of biomass across the roots and shoots being an indicator of where resources are most heavily allocated. In some species, domestication resulted in greater root mass and a large root to shoot ratio (Weeden, 2007; Grebenstein et al., 2012). The opposite holds true for other species such as *Triticum turgidum* L., as modern cultivars of wheat plants showed more shallow root systems with greater allocation of biomass to shoots when compared to the ancestral form (Roucou et al., 2017). Abiotic stressors also have a large impact on root:shoot ratio, with a higher ratio being an adaptation of drought tolerance in many species (Mo et al., 2015). This is shown not only for mature plants but at the seedling stage as well (Zimmer & Grose, 1958; Paz, 2003). Nutrient deficiency can also lead to a plant allocating more biomass to the roots, as in the case of nitrogen deficiency in *Triticum turgidum* spp. *durum* (Gioia et al., 2015). Root:shoot ratio is also affected by level of shading, with understory species that are competing for sunlight having greater root:shoot ratios and altered root morphology as seedlings (Paz, 2003).

Conclusion

In this study, we aim to elaborate upon the effects of domestication on seed characteristics, seedling growth, and biomass allocation. While phenotypic changes resulting from domestication syndrome are well documented in scientific literature, the focus typically lies on mature plants. For crops that are harvested for their vegetables or fruits and not for their seeds, there is less knowledge existing surrounding the characteristics of the earliest life stages of

the crop. In addition, we aim to further investigate the phenotypic changes differentiating domestication stages and see if these patterns differ across locations within the geographic center of origin of chile pepper. We may see significant changes in both seed and seedling characteristics across locations, potentially in part due to local environmental adaptation.

Chile peppers are a crop with an extensive history of domestication across varying regions. This work will cast further light on the adaptations that may have been selected for during domestication. Understanding the genetic variation present across the single species of *C. annuum* will deepen our understanding of this species and the diversity it may hold. With the earliest life stages of a plant being when it is the most vulnerable to abiotic and biotic factors, this focus holds a particular importance to chiles ability to survive in conditions with extreme shading, competition, or other adverse ecological factors. A steady decline of genetic diversity is visible across many crop species due to intensive agricultural practices, but by studying the genetic diversity of this species and identifying genes of interest we may be able to further mitigate this loss by reintroducing such diversity back into our crop plants. Identifying and conserving these pockets of diversity in our wild chiles is increasingly important, and the vitality of the crops in our landraces, plantations, and home gardens depends on their ability to withstand the adverse conditions imposed upon them.

CHAPTER 2: Seed and seedling characteristics of chile pepper (*Capsicum annuum* L.) across a domestication gradient

Abstract

Chile peppers (*Capsicum annuum* L.) are a crop with great importance both economically and culturally across Mesoamerica and beyond. During a time of climate change, the use of this crop species is threatened by rising temperatures and droughts. As such, the sources of genetic diversity found in landrace and wild (*C. annuum glabriusculum*) accessions are incredibly valuable. Identification of the natural adaptations in landraces and the wild relatives of chiles can facilitate the breeding of biotic and abiotic stress resistance into our crops, better securing their continued use. Seedlings, the most vulnerable stage of the plant life cycle, are an important focus area. Seed and seedling characteristics can be adaptations that assist in the survival of the plant. In this study, we examined seed and seedling characteristics across a domestication gradient in order to better understand the responses in phenotype resulting from the domestication process and the degree to which those changes may be similar across localities of origin. In order to investigate this, *C. annuum* seeds representing multiple domestication levels from across Mexico were measured using both digital phenotyping methods and physical measurements to quantify traits regarding size and shape of the seeds. Seeds were then planted, and seedlings grown in the greenhouse; they were measured from the time of emergence for phenotypes such as height, cotyledon size, leaf size, and number of leaves. We predicted that both seed and seedling phenotypes would significantly differ across level of domestication and across location of origin. We found that this was true for the majority of phenotypes examined, with certain phenotypes such as cotyledon area and height being especially impacted by domestication. We also found

that seed phenotypes were significantly affected by both domestication and location of origin but were more strongly influenced by origin overall. These results inform our understanding of the effects of domestication and local environments on phenotypes, indicating that characteristics influencing seedling vigor may have been selected for during domestication. In addition, this research uncovered landrace and wild accessions of interest that show beneficial phenotypes that are recommended for further examination.

Introduction

Chile peppers are a staple crop playing an important economic role around the world. Due to the extensive cultivation and repeated domestication events of *C. annuum*, genetic diversity of this crop has been greatly altered over time (Aguilar-Melendez et al., 2009). Chiles are grown in a variety of agricultural systems that include both modern and traditional farming methods. Landrace varieties of *C. annuum* often have exceptionally high genetic diversity due to the ongoing introgression of genetic material from wild and semiwild chiles in these cropping systems (Taitano et al., 2018). In addition, wild chiles (*C. annuum* var. *glabriusculum*) and semiwild chile populations may be untapped sources of genetic diversity. Chile peppers, like many other crops, are increasingly vulnerable to the effects of climate change such as drought and high temperatures. These sources of genetic material within landrace and wild accessions may be valuable tools that can be reintegrated into our crops through plant breeding, providing resistance to both biotic and abiotic stressors that were lost over time due to cultivation. With Mexico being the geographic center of origin for *C. annuum*, genetic diversity is expected to be highest within this region.

The domestication of crops has resulted in significant changes in morphology and growth over time, known as domestication syndrome (Flint-Garcia, 2013). While many of the changes resulting from domestication syndrome are beneficial to the plants use as a crop, the genetic diversity of the species within cultivated varieties can be greatly reduced when compared to its wild ancestors (Shi and Lai, 2015). This can lead to a decrease in crops' adaptive ability to withstand environmental changes and can further increase the risk of genetic bottlenecks in response to extreme climatic events. Maintaining the genetic diversity of our crops is increasingly important as a result (McCoy et. al., 2022).

Phenotypes of seeds, seedlings, and roots play a vital role in the survival and fitness of plants. Seed morphology in particular has large effects on its ecological interactions. Characteristics of seeds such as size and shape may have a large impact on the ability of the seed to persist in the environment for extended periods of time and on the ability of the seed to germinate. Size of the seed is often correlated with quality and can impact not just the germination of the seed but the ability of the plant to survive throughout its entire life cycle (Metz et al., 2010). Seedling characteristics such as height and cotyledon size can also play an important role in the ability of the plant to survive to maturity (Grossnickle, 2012; Milberg and Lamont, 1997). Seedlings in harsh environments that experience drought or have little available nutrients often have adaptive characteristics that assist in their survival (Khurana and Singh, 2001; Zipperlen and Press, 1996). One such adaptation includes changes in biomass allocation. Increased seedling and mature plant root mass is an adaptive trait often seen in environments that are heavily shaded, nutrient devoid, or as a response to drought (Paz, 2003; Gioia et al., 2015; Mo et al., 2015).

The purpose of this study was to quantify the variation in seed and seedling characteristics across a domestication gradient, allowing us to see how domestication has affected the variation within *C. annuum*. We hypothesize that traits such as faster development, greater heights, and larger leaves may be characteristic of more domesticated plants. Seedling traits will be visualized as by growth over time, and this information will allow us to better understand how the domestication process affects the early seedling characteristics of chiles, rather than the characteristics of the later life stages. With root to shoot ratios being an indicator of nutrient allocation and a response to a variety of biotic and abiotic conditions, we will also be looking at the root to shoot ratios across the domestication spectrum. As this ratio varies widely across species and environments, there is still much to be learned about how domestication has affected this nutrient allocation. By comparing the root and shoot biomass of *C. annuum*, we will be able to better to understand how domestication of this species affects nutrient allocation and whether semi-wild species are intermediate between wild and landrace accessions in the root to shoot ratio.

In this addition to the seedling stage, our goal is also to quantify the variation in phenotypes of seeds. With seed characteristics playing an indisputably vital role in the growth of plants, the phenotypic expression of seeds within *C. annuum* may have ecological impacts on the germination of the seeds and fitness of the emerged seedlings. We hypothesize that, as with many other crop species, larger seed size will be seen across the more heavily domesticated plants and smaller seeds will be more common for plants on the less domesticated end of the spectrum. This in-depth phenotyping of seeds using both physical and digital measurements tracked to the individual plant will allow us to see how these measurements play a role in the timing of seedling emergence and the seedling characteristics after germination.

The following objective will be addressed in this study: to determine if phenotypic variation in *C. annuum* seed characteristics, seedling growth, and biomass allocation patterns are affected by the level of domestication or geographic origin in accessions from southern Mexico. This will deepen our understanding of the effects of domestication syndrome and local environmental adaptation on *C. annuum*.

Methods and Materials

Germplasm

All germplasm was initially collected from Mexico and then subsequently grown in uniform greenhouse environment. Collections were made in collaboration with Mexican researchers from The National Autonomous University of Mexico (UNAM) at the Jardón-Barbolla lab. Collections were made from a range of environments, primarily coastal regions, from the Mexican states of Oaxaca, Chiapas, and Yucatan. In addition, select accessions sourced from the U.S. were included from the states of New Mexico and California. Accessions were also collected across environmental gradients for temperature, precipitation, and elevation. The accessions from Mexico used in this study are comprised of coastal, central valley, and the Península Yucatán regions. Collections made were from four different cultivation systems: forest, backyard garden, milpa, and plantation (Bernau, 2019).

In order to compare phenotypes across a domestication spectrum, we selected 21 accessions of *C. annuum* to represent seven locations in Southern Mexico and four levels of domestication (Table 1). The seven locations were divided by the state and municipality with Pochutla, Rosedal, San Pablo Coatlan, Copalita, and Saachilac within the state of Oaxaca, Nuevo Celosio and Arriaga within the state of Chiapas, and Xpujil being within the state of Campeche.

A minimum of three accessions were selected from each location, with one accession representing wild and the remaining two accessions both representing landrace. Select additional accessions were included that were not a part of this location grouping. Due to limited availability of semiwild accessions, these were among the accessions not included in the location grouping. In total, 12 of the accessions were categorized as landrace, five were categorized as wild, two were categorized as semiwild, and two were categorized as commercial varieties.

This study doubled as a seed increase for another study, so some accessions came from different seed increases. The majority of the seed was increased in 2019 by Vivian Bernau, with four accessions (Ca0060, Ca0020, Ca0345, and Ca0305) having been increased in 2014 and two accessions (Ca0437 and Ca0447) in 2021 (Bernau, 2019). Two rounds of seeds were observed. Firstly, seeds to be used in the seed increase were selected in order to test seed characteristics as covariates for seedling phenotypes. After this, extra seed was used to do a more thorough phenotyping. For this second round of seeds, the majority of the seeds used were also from the same seed increase.

Experimental Design

This study consisted of two experiments in summer 2022. The objective of the first experiment was to quantify seed traits and determine if the level of domestication or geographic origin affected these traits. To assess this, an in-depth phenotyping was performed using both physical measurements and digital image analyses. The seeds of 21 accessions were used, with groups of 10 seeds constituting each experimental unit. Two blocks were used, for a total of 420 seeds measured.

The objective of the second experiment was to determine if phenotypic variation in *C. annuum* seedlings correlates to domestication level or location of origin. In this second experiment, we assessed seedling characteristics from seeds that we had also phenotyped for physical traits in the lab prior to planting. The seeds used were from the same seed lot as those used in the grow-out for the greenhouse experiment. We used a randomized complete block design with 21 accessions and four blocks. Each of our 21 accessions randomized within a block represented an experimental unit, with six individual pots used as sampling unit within each experimental unit. In total, with four blocks, there were a total of 504 pots. Seeds were planted at a depth of 1cm directly into ½ quart pots that were placed on heating mats that were set to 28 °C to encourage germination (Monroy-Barbosa & Bosland, 2008). Irrigation for the seedling pots consisted of irrigation lines suspended above the pots that were set to mist at hourly intervals between 8:00 am and 5:00 pm. Seedling pots were removed from heating and transplanted into larger pots approximately four weeks after planting, in which they were watered weekly by hand. Blocks were spaced out over time due to space limitations, with blocks 1 and 2 planted initially and blocks 3 and 4 planted once the block 1 and 2 seedlings were transferred into larger pots. The experiment was terminated once the majority of plants reached maturity. Greenhouse climate temperatures throughout the experimental period reached a low of 12.5 °C at night and a high of 38.44 °C during the day. The average daily temperature was 24.9 °C.

Seed characteristics

For both experiments, seed phenotyping included physical measurements followed by imaging. For the first experiment on seedling growth, phenotyped seeds were eventually planted, with sample identity kept throughout seed and seedlings. Physical measurements included the seed mass, with an analytical balance to the nearest 0.000g, and seed thickness, with digital

calipers that measured to the nearest 0.00mm. Thickness measurements were taken at the widest point of the seed across the embryonic axis. In addition to single seed measurements, each experimental unit containing six seeds (subsamples) was measured for the total weight of each experimental unit. Length and width of the seeds were not collected due to extreme variability and curved edges of seeds, making longest and widest points very difficult to determine, as calipers could not be properly positioned.

The seed morphology study was conducted using the same physical measurements as noted above, as well as digital image measurements. After physical measurements were completed, an imaging box was built for the digital portion of measurements. This box was constructed of solid hardwood that measured 19 inches wide, 13 inches long, and 12 inches tall, with all edges and openings sealed to prevent light entering or exiting the box. A lid was made for the top, which could be removed for arranging of the seeds, with a round hole cut into the top to fit the lens of a camera. The front of the imaging box contained an open slot at the bottom that was used to slide trays of seeds into the box. LED lights were installed on the inside of the box, fully extending around the entirety of the walls to provide even lighting. Once seeds were positioned inside, the open slot was covered fully. The positioning was the same for each subsample, with the 10 seeds organized into two rows of five and numbered 1-10 based on their position. A metric ruler was placed in view of the camera for each image to be used as a scale comparison. The camera used for imaging was a Canon EOS Rebel T6 digital camera. The lens attached for all imaging was an 18-55 mm lens, set to 1/320" shutter speed, an ISO of 400, framerate of 5.6, with a 2 second delay timer. These settings allowed for the clearest images given the lighting conditions of the box, with the delay timer ensuring full stabilization of the

camera before each image was taken. The camera was set to autofocus on the center-most seed in each image.

Image processing was conducted using the open-source software ImageJ (Abramoff, 2007) to collect measurements on area, perimeter, minimum length, maximum length, and circularity. Imaging methods were modified from Baek et. al (2020) and Tanabata et. al (2012). Each image was processed individually without the use of automation. Once an image was uploaded into the program, the scale was first set using the standard metric ruler included within the view of each image. The scale was set to 1 cm on the ruler, which corresponded to a range of 384-385 pixels per cm in each image. After the scale was set, the image was transformed and rotated 90 degrees to allow for easier counting of each seed in its position. The image was cropped using the rectangle select tool to only include the seeds and the background was subtracted to allow for separation of colors. Subtracting the background functioned to remove any background abnormalities, such as any uneven spots of color or lighting that may have appeared on the white background. The process of removing the background uses a “rolling ball” of a designated pixel size that subtracts any abnormalities across the background. Pixel abnormalities are determined based on deviations from the mean background value, and similar values were replaced with this mean value. A rolling ball radius was determined independently for each photo by selecting a value matching the radius of the largest seed. Due to some seeds being much larger than others, it was necessary to specify this radius individually. To determine the best radius for the rolling ball tool, the largest seed in the photo was measured manually using the polygon select tool, with the width measured in pixels being used as the rolling ball radius. This ensured that the rolling ball was at least the size of the largest seed, mitigating any distortion of this object. In most cases, this corresponded to a value near 300 pixels. The image

type was then set to 8-bit grayscale, limiting the total number of colors possible per pixel to 8 bits each, for use in analysis.

Next, an image threshold was set. Thresholding is the process of separating the pixels in an image and categorizing them based on color. This allows for the separation of the objects from the rest of the image. Our use of a solid white background following by removal of potential background abnormalities allowed for a more accurate thresholding process with clear outlines of the seeds. Auto-detection of the threshold was used first and then altered manually to allow for selection of only the seeds without any background interference and minimize the selection of any shadowing around seed edges. After this, the particles were analyzed. The object size was manually set to the minimum seed size in each image, with options selected to exclude on edges and not include any holes. This floods the inside of the object and fills any extreme changes in color due to seed variation that may be perceived as a hole in the object, ensuring that these areas of variation are still measured as part of the seed. At this point, the measurements were deemed sound and could be entered into a spreadsheet.

Seedling characteristics

The seedling pots were monitored for emergence every 1-2 days, with the date of emergence being measured as a number of days after planting (DAP). One week after planting, regular measurements began for the emerged seedlings. Measurements included height, cotyledon width and length, true leaf width and length, and the total number of true leaves. Pots were continuously monitored for newly emerged seedlings throughout the experiment. Height consisted of two separate measurements: to the first node and to the top node. Height to the first node was measured with a standard metric ruler from the soil surface to first node on the stem, which contained the cotyledons. Height to the top node was measured from the soil surface to the

highest node on the plant. A single cotyledon on each plant was chosen at random for measurements, with lengths and widths of the lamina being measured only when the cotyledon was fully extended, and the edges were uncurled. True leaves were measured in the same manner, but the true leaf used as a representative was the most recently emerged true leaf that was still fully extended. The total number of true leaves was also counted for each plant. Measurements were taken at 3–5-day intervals throughout the entire experiment. After roughly 30 days of seedling measurements, all but one seedling from each accession in each block was harvested for biomass measurements, and the remaining seedling was transferred into a larger, 6-liter pot for continued measurements. Continued measurements of seedling characteristics were taken using the same methods and ceased once the majority of plants reached maturity, at 60 DAP. Plants were determined to have reached maturity once the reproductive stage of development was entered, distinguished by the formation of flower buds on the plant.

Seedling data was first binned into 3–5-day intervals by averaging the values across time periods, for each sub-sample, in order to conduct analyses on the irregularly timed measurements. First, we looked at data by a single point in time for each phenotype. We individually looked at the phenotypes of height to the first node, cotyledon area, true leaf area, and total number of true leaves. For true leaf area and cotyledon area, a length-width-proxy was used to determine each value with values in cm^2 . The time point in DAP was chosen based on the binned period that contained the most complete set of data across all accessions so that sufficient comparisons could be made. Due to differences in emergence timing and physiological stages across accessions, the selections for the DAP to use for analyses were made as follows: For cotyledon area this was 23 DAP, for height to the first node this was chosen as 27 DAP, for true leaf area this was 28 DAP, and for total leaf number this was 29 DAP.

Biomass characteristics

After 30 DAP, all plants that were not kept for continued measurements were harvested for biomass. Seedlings were removed from the pots by hand and the roots were washed completely clean of soil, with any broken segments collected and cleaned as well. Roots and shoots were then separated and dried in an oven at about 55°C for 48 hours before weighing.

Statistical analysis

All analyses were performed using R Studio version 4.2.2. Models for all seed data and for single time point seedling and biomass data were fitted using general linear mixed models (GLMMs) using the R package “lme4”. For preliminary analyses, accessions were tested in a GLMM to ensure that they had a significant effect on phenotypes before further analysis was conducted. For these models, the response variable corresponded to each phenotype with accession being the fixed factor and with block as a random factor. This was conducted for seed, seedling, and biomass allocation characteristics. When these tests were significant, we further explored the differences between accessions comparing least squares means using the Tukey method in the “emmeans” R package. Compact letter displays from the `clld` function in this package were used to visualize mean separation of least squared means with error bars depicting standard error. Post-hoc analyses were conducted for all models to assess the normality assumption of our residuals by creating histograms of residuals, Q-Q plots of residuals, and residuals vs. fitted values plots. All post-hoc analyses showed a normal distribution of data.

Further models were built using each phenotype as a response variable with location grouping, domestication levels, and the interaction between location and domestication as predictor variables with block as a random effect. For seed, seedling, root, and biomass allocation characteristics, the response variables used corresponded to each phenotype with the

fixed factors being domestication and location. A fixed interaction term between domestication and location was also used with block as a random factor. Mean separation was conducted and visualized in the same manner as noted above.

Repeated measures data for seedling phenotypes (height to first node, true leaf area at time of leaf emergence, and total true leaf number) was analyzed by first averaging the values across subsamples per experimental unit. Data for each phenotype was fitted to a general linear mixed model (LMM) using the “lme4” package in R. Multiple models were fitted with different covariance structures and their AIC, BIC, and log-likelihood values were compared to determine the best fit. The best fitting covariance-variance structure was “unstructured” for height to the first node and total number of true leaves and “diagonal” for the true leaf area phenotype (Table 2). The model included domestication, location, and DAP and their two- and three-way interaction terms as fixed predictor variables and block was included as a random effect. Due to the repeated measures nature of the data (i.e., a within-subjects design), error was accounted for by including both DAP and Block as random factors in the model statement. The LSmeans function from the “lsmeans” package in R was then used to obtain least-squared means for each DAP value. Lsmeans \pm SE was visualized using the “ggplot” package in R with separation by both domestication level and location with error bars depicting standard error.

Results

Seed Phenotypes

In a simple ANOVA, accession had a significant effect on all seed phenotypes (Table 3; Supplementary Figures 1-5).

Analyses of seed phenotypes by location, domestication, and their interaction showed that seed area, perimeter, and circularity were affected by domestication; seed area, weight, perimeter, and circularity were affected by location; and only seed thickness was affected by the interaction of location and domestication (although seed circularity had marginal effects from the interaction, with $p < 0.01$; Table 4). While seed circularity is affected by domestication and location, subsequent mean separation tests showed that factor level effects (of domestication and location) could not be elucidated for seed circularity.

Mean separation tests of the phenotypes across locations showed that the largest seeds by area and perimeter, as well as the heaviest seeds, were seen in U.S. accessions (Table 5; Figure 1). Across the nine locations' phenotypic means, seed area ranged from 0.013 ± 0.0069 to $0.0978 \pm 0.01239\text{cm}^2$; seed weight ranged from 0.0017 ± 0.0009 to $0.0082 \pm 0.0005\text{g}$; seed perimeter ranged from 0.714 ± 0.0969 to $1.557 \pm 0.0485\text{cm}$; and, although means for seed circularity did not differ across locations, circularity ranged from 75.39 ± 1.24 to $83.65 \pm 2.47\%$ (Table 5).

Mean separation revealed significant changes across the domestication gradient in seed area and perimeter. Across the four domestication levels (wild, semiwild, landrace and commercial), average seed area ranged from 0.0479 ± 0.0123 to $0.1173 \pm 0.0123\text{cm}^2$, and seed perimeter ranged from 0.8411 ± 0.0831 to $1.3893 \pm 0.0831\text{cm}$, and although means for seed circularity did not differ across domestication levels, circularity ranged from 75.27 ± 1.75 to $83.03 \pm 1.75\%$ (Table 6). Commercial and landrace accessions were the largest by area and perimeter, although wild seed areas were not significantly different from, yet were intermediate between, both landrace and semi-wild (Table 6). Although seed thickness showed significant effects from the interaction between location and domestication (Table 4), further mean separation found little differences across domestication levels (Table 6).

Further inspection of the domestication-by-location interaction showed that seed percent circularity was the most uniform across locations and domestication levels (Figure 1A), whereas area, weight, thickness, and perimeter of seeds from US tended to be larger (Figure 1B-1D), although bins for means of all phenotypes never completely separated (Figure 1).

Seedling Phenotypes

We found a significant effect of accession on the variance of all seedling phenotypes except for true leaf area, which was nearly significant (Table 7). Subsequent mean separation of seedling phenotypes by accession showed no significant differences between any pair of accessions for number of true leaves and true leaf area (Supplementary Figures 6-7) and some significant differences in both height and cotyledon area (Supplementary Figures 8-9),

Height to the first node (27 DAP) and total number of true leaves (29 DAP) were each significantly affected by location (Table 8). The strongest effect from location was on height to the first node, by far (Table 8). Mean separation across locations showed that average height to the first node ranged from 1.0376 ± 0.3413 to 1.7425 ± 0.1846 cm, cotyledon area ranged from 0.1486 ± 0.5997 to 1.7321 ± 0.2129 cm², true leaf area ranged from 0.2308 ± 0.1421 to 0.4491 ± 0.1173 cm², and total number of true leaves from $1.2852 \pm 2.9515 \pm 0.4082$, while only height to first node and number of true leaves had significant differences across locations (Table 9). The tallest plants were the U.S. accessions, while all other locations had significantly shorter plants (Table 9).

Out of the seedling phenotypes, domestication was shown to have a significant effect on the variance of cotyledon area (23 DAP) only, with near-significant effects on height to the first node (27 DAP) and true leaf area (28 DAP) (Table 8). Further mean separation tests showed that,

although not statistically significant, height to the first node and true leaf area both showed a trend with the levels of domestication. Mean separation also revealed that commercial accessions had significantly larger cotyledons (Table 10). Means across domestication levels ranged from 0.9059 ± 0.2912 to 2.5592 ± 0.1869 cm for height to the first node, 0.1431 ± 0.3953 to 2.2349 ± 0.2489 cm² for cotyledon area, -0.1972 ± 0.2241 to 0.5517 ± 0.1331 cm² for true leaf area, and 1.0378 ± 0.4214 to 2.9093 ± 0.5635 for total number of true leaves (Table 10). From the repeated measures analyses, ANOVAs of seedlings across location, domestication, time (DAP), and their multifactorial interactions, showed that height to the first node and the total number of true leaves both were affected by two interactions: the interaction of domestication and time, and that of location and time (Table 11). Mean separation of the effects of domestication on seedling phenotypes over time showed that differences across some domestication levels were amplified over time for both the total number of true leaves (Figure 2) and height to the first node (Figure 3). For total number of true leaves, semiwild remained not significantly different than either wild or landraces over time (Figure 2).

Domestication levels began to show significant differentiation in the total number of true leaves at 17 DAP (Figure 2), while differentiation began at 7 DAP for height to the first node (Figure 3). Commercial accessions were significantly larger than other domestication levels in both of these phenotypes across later time intervals. Landrace accessions became taller than wild accessions at 17 DAP (Figure 2) and leafier than wild and semiwild accessions by 13 DAP (Figure 3). Semiwilds remained intermediate between wilds and landraces for true leaf count, while fluctuating between intermediate and grouping with wilds for height to first node after 7 DAP.

Mean separation across locations, in the repeated measures analysis, showed that accessions from the U.S. were significantly taller than all other locations, with differentiation beginning near 7 DAP and amplifying over time. All other locations had the same heights over time (Figure 4). Accessions from the U.S. had significantly more true leaves than those in Mexico starting at 13 DAP (Figure 5). Differentiation across other locations was observed to begin at 17 DAP, when accessions from Saachilac started having more true leaves than the other Mexican accessions except for those from Nuevo Colosio. Those differences amplified over time (Figure 5). The three-way interaction of location, domestication, and time did not significantly affect seedling phenotypes (Supplementary Figures 7-11).

Root and Shoot Phenotypes

Accession had a significant effect on the variance in both root mass and shoot mass but not root-to-shoot ratio (R:S; Table 12). Mean separation of biomass by accession showed some differences across accessions for root mass and shoot mass. CM334, a New Mexico variety derived from a Mexican landrace, had higher root and shoot biomass than some of the Mexican accessions (Supplementary Figures 15-16).

Among locations, mean separation showed significant differences for root mass and shoot mass, with accessions from the U.S. having the highest root and shoot mass (Table 15, Figure 6). R:S exhibited a significant interaction between domestication and location (Table 13); however, mean separation could not resolve this clearly across locations (Figure 6).

The variance in R:S was significantly affected by domestication while the variance in both root and shoot biomass were significantly affected by location (Table 13). While means did not significantly differ across domestication levels, the R:S in semiwild accessions trended to be

the highest, ranging from 0.5162 ± 0.0772 to 1.1492 ± 0.2388 g (Table 14). By location, root mass ranged from 0.0007 ± 0.0028 to 0.0135 ± 0.0016 g and shoot mass ranged from 0.0014 ± 0.0046 to 0.0201 ± 0.0026 g, in both cases with U.S. accessions being the largest (Table 15).

Discussion

Seed Phenotypes

Overall, seeds of *C. annuum* were more strongly influenced by their local environment than by their level of domestication. The seed phenotypes for weight, perimeter, area, and circularity were influenced by location, particularly in US versus Mexico accessions for the former three. This phenotypic differentiation across locations may be due to local environmental adaptations in chiles although we were unable to test the adaptive nature of these traits. Higher seed circularity (e.g., a more rounded seed shape) and smaller seed size have been associated with greater persistence in the soil (Moles et al., 2003), which may be enabling chiles to persist in the seed bank for extended periods. This increased persistence could enhance the seeds' chances of survival during unfavorable germination conditions. Although not significant, seed size metrics such as weight, perimeter, and area were relatively low for seeds from Xpujil compared to all other locations, with the addition of seeds from Xpujil having relatively higher circularity. Elevation and temperature data for this location are well within the range of our other locations, but Xpujil is located much further away from our other locations, being in the Yucatan Peninsula. This greater geographic separation and likely higher genetic differentiation may explain the distinct characteristics observed with our accession from this location.

While seed perimeter and area were significantly affected by domestication level, seed weight was not. Seed area had especially strong separation across domestication levels. Seed perimeter and area were each observed to be higher across commercial and landrace accessions. *C. annuum* seeds show improvement in both the number and size of seeds produced during cultivation over time (Jarret et al., 2018), which could explain why higher seed perimeter and area were observed in cultivated accessions. Additionally, seed thickness exhibited an interaction between location and domestication level. This interaction may indicate that the effect of domestication is different across locations. Chiles may be able to better adapt to unfavorable environmental conditions when being of a certain level of domestication, such as wild or landrace. Seed size, which can include the metric of thickness, is often correlated with larger energy stores within the seed (Milberg and Lamont, 1997). This could mean that larger energy stores within the seed is a trait developed during domestication but that is only expressed within certain environmental conditions. Commercial accessions had the thickest seeds as well as the largest seeds by area and weight. This indicates that domestication may indirectly contribute to larger, heavier, and thicker seed size over time.

Seedling Phenotypes

Seedling phenotypes were influenced by both location and domestication, but not the interaction of location and domestication. Commercial accessions also exhibited higher mean values for cotyledon area. These findings suggest that chile seedling phenotypes are altered through domestication.

Location also significantly impacted seedling phenotypes, with accessions from the U.S. being differentiated from all others by increased height and larger leaf numbers. This observation could be an effect of formal breeding or could indicate that environmental conditions for chiles

in this region are different from those in our other locations, possibly leading to specific adaptations in these phenotypes. The height of *C. baccatum* was shown to be significantly affected by geographic distribution, with shorter heights being characteristic of western accessions (Albrecht et al., 2012). This may be the same case for *C. annuum*. Our results showed that location had a stronger effect on height to the first node than on the number of true leaves, with height to the first node having a much stronger p-value. This could mean that the adaptation of increased height has greater benefit in some environments compared to increased leafiness. Leaf area and cotyledon area each showed no effect from location, indicating that leaf size may not play a large role in environmental adaptation of chiles.

Interactions of both domestication with time and location with time were observed for seedling height to the first node and the total number of true leaves. Competition with nearby plants, which can occur in both wild settings and within agricultural settings, may be a driver for chile seedlings to grow taller as a shade avoidance response, which likely has a geographic pattern as the canopy changes throughout the landscape. In addition, faster growth is often selected for in domesticated plants, with improved pepper varieties such as Early Jalapeno being bred specifically for this trait to allow for an earlier harvest (Pickersgill, 2016). This trend was also observed in our commercial accessions, with significantly faster increases in height and leaf development. Furthermore, height and leaf development were similar across domestication levels near the time of emergence, with differentiation occurring as late as 17 DAP for the number of true leaves. This finding might imply that domestication has a minimal effect on phenotype during the earliest stage of chile development. Instead, changes in development due to domestication occur later in the seedling stage and become more pronounced over time.

Root and Shoot Phenotypes

Variation in biomass allocation of *C. annuum* was significantly affected by variation in location, for both root mass and shoot mass. However, biomass allocation was not significantly affected by domestication for either root mass, shoot mass, or R:S. There was also a significant interaction between domestication and location for R:S. While such interactions can be an indication of differential effects of domestication across locations (Aguilar-Melendez et al., 2009), the nature of our data cannot be interpreted that way due to incomplete combinations of domestication and location and lack of mean separation.

Both root and shoot mass individually were observed to be significantly different among the individual locations, but not among domestication levels. Mean separation showed higher mean values in both shoot mass and root mass for accessions from the U.S. However, these higher values may be due to differences in emergence timing, which can affect biomass at a given number of days after planting. As time of emergence was not taken into account for the analyses, an addition of an emergence time predictor variable in future models might disentangle the effects of location on biomass.

Effects of domestication on chile seeds and seedlings

Domestication had significant effects on the variance in phenotypes, including seed area and perimeter, and when accounting for days after planting, cotyledon area was affected by domestication. Commercial seeds had higher area than semiwild and wild seeds, while landrace seeds were neither significantly different than commercials nor wilds in terms of area. Commercial and landrace seeds had higher parameters than semiwild and wilds, too. These differences among seed characteristics in particular may be attributed to the process of domestication and selection for healthier seeds with larger energy stores, which is typically associated with larger seed size. This could also be related to the larger cotyledon sizes that were

seen in commercial and landrace accessions. Larger seeds have been associated with more vigorous seedlings due to slower development, which may be more favorable in stable environments, such as agricultural settings, while smaller seeds germinate quicker and more reliably, which is better suited to the initial colonization of new habitat (Souza and Fagundes, 2014).

In several cases, the effects of domestication were not revealed in a simple model, but instead emerged when multifactorial interactions of domestication, location, and time were accounted for. While not observed in our simpler models, true leaf area and number of true leaves showed significant effects of domestication in the more complex models accounting for the interactions of location, domestication, and time together (Tables 9 and 11). The development of true leaves over time showed significant differences across domestication, with commercial accessions having a greater number of leaves at 17 DAP (Figure 2). Landrace accessions were also distinct from wild accessions at this point in time, although not from semiwild accessions. Height to the first node also showed a similar pattern, with commercial accessions becoming significantly taller from all other domestication levels over time (Figure 3). In addition, landrace accessions remained distinct from both wild and semiwild chiles after 12 DAP. However, between 7 and 12 DAP, semiwild accessions were observed to resemble both landrace and wild accessions, reverting back to being indistinguishable from wild accessions at 14 DAP.

Interestingly, semiwild accessions often fell between wild and landrace groupings for seedling phenotypes. Semi-wild chiles, while closely resembling their wild counterparts, can experience a degree of cultivation that may result in morphological differences over time. As a result, semi-wild chiles may hold an intermediary position of domestication between wild and

domesticated chiles. In addition, the ongoing gene flow between landrace and wild accessions as seen in more traditional agricultural systems could be contributing to these characteristics (Perez-Martinez et al., 2022). Our results from Figures 2 and 3 support the position of semi-wild chiles as intermediary, though this is only observed at specific developmental stages and not throughout the entire seedling stage.

Biomass allocation, however, did not alter across the domestication gradient. The allocation of biomass as root to shoot ratios has been associated with level of domestication in some species but not in others (Weeden, 2007; Grebenstein et al., 2012; Roucou et al., 2017). More often, biomass allocation is affected by abiotic factors such as drought, nutrient deficiency, or shading (Mo et al., 2015; Gioia et al., 2015; Paz, 2003). Biomass allocation was observed to be consistent across domestication levels in this study, while other seedling phenotypes showed an effect of domestication. For chiles, alterations in biomass allocation, especially involving higher root and shoot masses, seem to be more strongly influenced by location, rather than domestication itself. This could be due to variation in local environmental conditions, such as the cultivation setting. Landraces of *C. annuum* were shown to respond to a water deficit with altered biomass, among other morphological characteristics (McCoy et al., 2022). While the effects of domestication on root and shoot masses were not observed in this study, perhaps the phenotypic plasticity of certain accessions has been impacted by domestication. If so, alterations in biomass allocation may be revealed if exposed to a drought stress treatment, as in McCoy's work.

Effects of location of origin on chile seeds and seedlings

For several phenotypes, location of origin had a stronger effect than domestication. While domestication had a slight effect on seed area, perimeter and circularity, location had a much

stronger effect on the three seed phenotypes. Similarly, location had a stronger effect on height to first node (27 DAP) and total number of true leaves (29 DAP) than domestication. Increased plant height is an adaptation associated with the presence of nearby competition as a shade-avoidance response. Increased leaf number as well as faster leaf development in plants can also be an adaptation to a variety of environmental conditions including temperature, moisture, and nutrient levels (Atkinson and Porter, 1996). *C. annuum* L. var. *grossum* was shown to have high phenotypic responses in plant height and LAI (leaf area index) in response to varying irrigation treatments (Lodhi et al., 2014). A genetic component to these environmental responses in *C. annuum* is likely, with variability in traits such as leaf number or plant height being associated with specific landraces as a result of adaptive or neutral evolution towards differentiation by geography.

Conclusion and implications

Our hypothesis that larger seed size would be correlated with domesticated chiles was confirmed, and we discovered that altered seed size and shape could possibly be an adaptation for chiles in specific environments, whether influenced by geography (location) or domestication. Our expectation that domesticated seedlings would be taller, with phenotypes such as larger leaves was partially met. In our simple models, while the size of true leaves at the time of leaf emergence was not affected by domestication or location, cotyledon area was shown to be significantly affected by domestication. However, in our complex models, true leaf area did show effects from domestication. In addition, faster leaf development was shown for domesticated chiles. Biomass allocation was shown to significantly differ across locations, though our mean separation showed no significant effects of domestication. Consequently, seed phenotypes, seedling phenotypes, and seedling development appear to be the most impacted by

domestication. Furthermore, location of origin significantly affected these traits, with interactions between domestication and location present in a select few of the phenotypes we looked at. Surprisingly, cotyledon area was especially affected by domestication alone. The significant differences seen across domestication level may indicate that cotyledon size plays an essential role in seedling vigor and is indirectly selected for during domestication.

This knowledge may be valuable for the conservation of genetically distinct chile accessions. The preservation of genetic material in our crops is becoming increasingly important due to a rapidly changing climate and a quickly growing demand for food (Zhang et al., 2017). As plants are more vulnerable to both biotic and abiotic factors during the seedling stage, selecting for seedling traits associated with increased vigor and survival could prove to be beneficial in agricultural systems. Landrace accessions such as CM334, ca0300, and ca0260 might be particularly important to maintain, given their increased size and vigor as demonstrated by the phenotypes examined here. Furthermore, it could be advantageous to further investigate wild accessions such as ca0303 for similar reasons. These accessions could be especially valuable for genetic investigations in pepper and may warrant a higher focus in conservation efforts. The increased vigor observed in some of these accessions could be particularly beneficial for integration in agricultural systems facing a greater threat from changing environmental conditions. Although seed phenotypes might not be directly selected for in the majority of cultivated chiles, they may still influence chile survival and vigor. We recommend further examination of seed phenotypes and the potential benefits of traits involving seed shape and size for chile crops. Additionally, we recommend further examination of cotyledon size in relation to the domestication of chiles.

Acknowledgements

I give my deepest thanks to Dr. Kristin Mercer for being an incredible advisor that was a source of endless inspiration, knowledge, guidance, and patience throughout this opportunity. I also extend my deepest gratitude to Hannah Scheppler, for guiding me throughout this entire process and being endlessly supportive of my efforts within this project as well as my future beyond. I further thank Jack McCoy, Layne Connolly, Chee Ngui, and Edison Cigany for their support and guidance. I would also like to thank Dr. Lev Jardón-Barbolla for his guidance during this project and for the germplasm collection that made all of this possible. I am also grateful to Mr. Mike Anderson for his inexhaustible guidance, knowledge, and efforts in coordinating the inner workings of Howlett Greenhouse during this experiment. I am deeply grateful to Matthew Lincicome for his guidance, knowledge, and moral support from start to end. Finally, thank you to my friends and family for their endless support and encouragement throughout this journey.

Literature Cited

- Abramoff, M. (2007). ImageJ as an Image Processing Tool and Library. *Microscopy and Microanalysis*, 13(S02). <https://doi.org/10.1017/s1431927607079652>
- Aguilar-Melendez, A., Morrell, P. L., Roose, M. L., & Kim, S.-C. (2009). Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annuum*; Solanaceae) from Mexico. *American Journal of Botany*, 96(6), 1190–1202. <https://doi.org/10.3732/ajb.0800155>
- Aguilar-Meléndez, A., Vásquez-Dávila, M. A., Manzanero-Medina, G. I., & Katz, E. (2021). Chile (*Capsicum* spp.) as Food-Medicine Continuum in Multiethnic Mexico. *Foods*, 10(10), 2502. <https://doi.org/10.3390/foods10102502>
- Albrecht, E., Zhang, D., Mays, A. D., Saftner, R. A., & Stommel, J. R. (2012). Genetic diversity in *Capsicum baccatum* significantly influenced by its ecogeographical distribution. *BMC Genetics*, 13(1) doi:10.1186/1471-2156-13-68
- Atkinson, D., & Porter, J. R. (1996). Temperature, plant development and crop yields. *Trends in Plant Science*, 1(4), 119-124. doi:10.1016/S1360-1385(96)90006-0
- Baek, J., Lee, E., Kim, N., Kim, S. L., Choi, I., Ji, H., Chung, Y. S., Choi, M.-S., Moon, J.-K., & Kim, K.-H. (2020). High Throughput Phenotyping for Various Traits on Soybean Seeds Using Image Analysis. *Sensors*, 20(1), 248. <https://doi.org/10.3390/s20010248>
- Benlioğlu, B., & Adak, M. S. (2019). Importance of Crop Wild Relatives and Landraces Genetic Resources in Plant Breeding Programmes. *Journal of Experimental Agriculture International*, 1–8. <https://doi.org/10.9734/jeai/2019/v37i330268>
- Bergkamp, B., Impa, S. M., Asebedo, A. R., Fritz, A. K., & Jagadish, S. V. K. (2018). Prominent winter wheat varieties response to post-flowering heat stress under controlled chambers and field based heat tents. *Field Crops Research*, 222, 143–152. <https://doi.org/10.1016/j.fcr.2018.03.009>
- Bernau, V. (2019) Exploring local adaptation and drought tolerance in chile peppers (*Capsicum* spp.) of southern mexico. The Ohio State University.

- Bernau, V. M., Jardón Barbolla, L., McHale, L. K., & Mercer, K. L. (2020). Germination response of diverse wild and landrace chile peppers (*Capsicum* spp.) under drought stress simulated with polyethylene glycol. *PLOS ONE*, 15(11), e0236001. <https://doi.org/10.1371/journal.pone.0236001>
- Bohra, A., Kilian, B., Sivasankar, S., Caccamo, M., Mba, C., McCouch, S. R., & Varshney, R. K. (2022). Reap the crop wild relatives for breeding future crops. *Trends in Biotechnology*, 40(4), 412–431. <https://doi.org/10.1016/j.tibtech.2021.08.009>
- Bridgemohan, P., Mohammed, M. and Bridgemohan, R.S. (2017). Capsicums. In *Fruit and Vegetable Phytochemicals*, E.M. Yahia (Ed.). <https://doi.org/10.1002/9781119158042.ch45>
- Brush, S. B., & Meng, E. (1998). Farmers' valuation and conservation of crop genetic resources. *Genetic resources and crop evolution*, 45, 139-150. <https://doi.org/10.1023/A:1008650819946>
- Chiou, K. L., & Hastorf, C. A. (2014). A Systematic Approach to Species–Level Identification of Chile Pepper (*Capsicum* spp.) Seeds: Establishing the Groundwork for Tracking the Domestication and Movement of Chile Peppers through the Americas and Beyond. *Economic Botany*, 68(3), 316–336. <https://doi.org/10.1007/s12231-014-9279-2>
- Conesa, M. À., Fullana-Pericàs, M., Granell, A., & Galmés, J. (2020). Mediterranean Long Shelf-Life Landraces: An Untapped Genetic Resource for Tomato Improvement. *Frontiers in Plant Science*, 10. <https://doi.org/10.3389/fpls.2019.01651>
- ElMasry, G., Mandour, N., Al-Rejaie, S., Belin, E., & Rousseau, D. (2019). Recent Applications of Multispectral Imaging in Seed Phenotyping and Quality Monitoring—An Overview. *Sensors*, 19(5), 1090. <https://doi.org/10.3390/s19051090>
- Engels, J. M. M., Ebert, A. W., Thormann, I., & de Vicente, M. C. (2006). Centres of Crop Diversity and/or Origin, Genetically Modified Crops and Implications for Plant Genetic Resources Conservation. *Genetic Resources and Crop Evolution*, 53(8), 1675–1688. <https://doi.org/10.1007/s10722-005-1215-y>

- Esquinas-Alcázar, J. Protecting crop genetic diversity for food security: political, ethical and technical challenges. *Nat Rev Genet* **6**, 946–953 (2005). <https://doi.org/10.1038/nrg1729>
- Flint-Garcia, S. A. (2013). Genetics and Consequences of Crop Domestication. *Journal of Agricultural and Food Chemistry*, *61*(35), 8267–8276. <https://doi.org/10.1021/jf305511d>
- Gioia, T., Nagel, K. A., Beleggia, R., Fragasso, M., Ficco, D. B. M., Pieruschka, R., De Vita, P., Fiorani, F., & Papa, R. (2015). Impact of domestication on the phenotypic architecture of durum wheat under contrasting nitrogen fertilization. *Journal of Experimental Botany*, *66*(18), 5519–5530. <https://doi.org/10.1093/jxb/erv289>
- González-Jara, P., Moreno-Letelier, A., Fraile, A., Piñero, D., & García-Arenal, F. (2011). Impact of Human Management on the Genetic Variation of Wild Pepper, *Capsicum annum* var. *glabriusculum*. *PLoS ONE*, *6*(12), e28715. <https://doi.org/10.1371/journal.pone.0028715>
- Grebenstein, C., Kos, S. P., de Jong, T. J., Tamis, W. L. M., & de Snoo, G. R. (2012). Morphological markers for the detection of introgression from cultivated into wild carrot (*Daucus carota* L.) reveal dominant domestication traits. *Plant Biology*, *15*(3), 531–540. <https://doi.org/10.1111/j.1438-8677.2012.00662.x>
- Grossnickle, S. C. (2012). Why seedlings survive: influence of plant attributes. *New Forests*, *43*(5-6), 711–738. <https://doi.org/10.1007/s11056-012-9336-6>
- Illa, T. C. C., Maxted, N., Scholten, M., & Ford-Lloyd, B. (2005). Defining and identifying crop landraces. *Plant Genetic Resources*, *3*(3), 373–384. <https://doi.org/10.1079/pgr200591>
- Jarret, R. L., Levy, I. J., Potter, T. L., & Cermak, S. C. (2013). Seed oil and fatty acid composition in *Capsicum* spp. *Journal of Food Composition and Analysis*, *30*(2), 102–108. <https://doi.org/10.1016/j.jfca.2013.02.005>
- Khoury, C. K., Brush, S., Costich, D. E., Curry, H. A., Haan, S., Engels, J. M. M., Guarino, L., Hoban, S., Mercer, K. L., Miller, A. J., Nabhan, G. P., Perales, H. R., Richards, C., Riggins, C., & Thormann, I. (2021). Crop genetic erosion: understanding and responding to loss of crop diversity. *New Phytologist*, *233*(1), 84–118. <https://doi.org/10.1111/nph.17733>

- Khoury, C. K., Carver, D., Barchenger, D. W., Barboza, G. E., Zonneveld, M., Jarret, R., Bohs, L., Kantar, M., Uchanski, M., Mercer, K., Nabhan, G. P., Bosland, P. W., & Greene, S. L. (2019). Modelled distributions and conservation status of the wild relatives of chile peppers (*Capsicum* L.). *Diversity and Distributions*, 26(2), 209–225.
<https://doi.org/10.1111/ddi.13008>
- Khurana, E., & Singh, J. (2001). Ecology of seed and seedling growth for conservation and restoration of tropical dry forest: A review. *Environmental Conservation*, 28(1), 39-52.
[doi:10.1017/S0376892901000042](https://doi.org/10.1017/S0376892901000042)
- Kraft, K. H., Brown, C. H., Nabhan, G. P., Luedeling, E., Luna Ruiz, J. d. J., Coppens d'Eeckenbrugge, G., Hijmans, R. J., & Gepts, P. (2014). Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico. *Proceedings of the National Academy of Sciences*, 111(17), 6165–6170.
<https://doi.org/10.1073/pnas.1308933111>
- Lasky, J. R., Upadhyaya, H. D., Ramu, P., Deshpande, S., Hash, C. T., Bonnette, J., Juenger, T. E., Hyma, K., Acharya, C., Mitchell, S. E., Buckler, E. S., Brenton, Z., Kresovich, S., & Morris, G. P. (2015). Genome-environment associations in sorghum landraces predict adaptive traits. *Science Advances*, 1(6), e1400218.
<https://doi.org/10.1126/sciadv.1400218>
- Loddo, A., Di Ruberto, C., Vale, A. M. P. G., Uccesu, M., Soares, J. M., & Bacchetta, G. (2022). An effective and friendly tool for seed image analysis. *The Visual Computer*, 39(1), 335–352. <https://doi.org/10.1007/s00371-021-02333-w>
- Lodhi, A., Kaushal, A., & Singh, K. (2014). Impact of irrigation regimes on growth, yield and water use efficiency of sweet pepper. *Indian Journal of Science and Technology*, 7(6), 790. [doi:10.17485/ijst/2014/v7i6.2](https://doi.org/10.17485/ijst/2014/v7i6.2)
- McCoy, J. E., McHale, L. K., Kantar, M., Jardón-Barbolla, L., & Mercer, K. L. (2022). Environment of origin and domestication affect morphological, physiological, and agronomic response to water deficit in chile pepper (*Capsicum* sp.). *PLOS ONE*, 17(6), e0260684. <https://doi.org/10.1371/journal.pone.0260684>

- Mercer, K. L., & Perales, H. R. (2010). Evolutionary response of landraces to climate change in centers of crop diversity. *Evolutionary Applications*, 3(5-6), 480–493. <https://doi.org/10.1111/j.1752-4571.2010.00137.x>
- Mercer, K. L., Shaw, R. G., & Wyse, D. L. (2006). Increased Germination Of Diverse Crop–Wild Hybrid Sunflower Seeds. *Ecological Applications*, 16(3), 845–854. [https://doi.org/10.1890/1051-0761\(2006\)016\[0845:igodch\]2.0.co;2](https://doi.org/10.1890/1051-0761(2006)016[0845:igodch]2.0.co;2)
- Mercer, K., Martínez-Vásquez, Á., & Perales, H. R. (2008). Asymmetrical local adaptation of maize landraces along an altitudinal gradient. *Evolutionary Applications*, 1(3), 489–500. <https://doi.org/10.1111/j.1752-4571.2008.00038.x>
- Metz, J., Liancourt, P., Kigel, J., Harel, D., Sternberg, M., & Tielbörger, K. (2010). Plant survival in relation to seed size along environmental gradients: a long-term study from semi-arid and Mediterranean annual plant communities. *Journal of Ecology*, 98(3), 697–704. <https://doi.org/10.1111/j.1365-2745.2010.01652.x>
- Meyer, R., Purugganan, M. (2013). Evolution of crop species: genetics of domestication and diversification. *Nat Rev Genet* 14, 840–852. <https://doi.org/10.1038/nrg3605>
- Milberg, P., & Lamont, B. (1997). Seed/cotyledon size and nutrient content play a major role in early performance of species on nutrient-poor soils. *The New Phytologist*, 137(4), 665–672. [doi:10.1046/j.1469-8137.1997.00870.x](https://doi.org/10.1046/j.1469-8137.1997.00870.x)
- Mo, Y., Yang, R., Liu, L., Gu, X., Yang, X., Wang, Y., Zhang, X., & Li, H. (2015). Growth, photosynthesis and adaptive responses of wild and domesticated watermelon genotypes to drought stress and subsequent re-watering. *Plant Growth Regulation*, 79(2), 229–241. <https://doi.org/10.1007/s10725-015-0128-9>
- Moles, A. T., Hodson, D. W., & Webb, C. J. (2000). Seed size and shape and persistence in the soil in the New Zealand flora. *Oikos*, 89(3), 541–545. <https://doi.org/10.1034/j.1600-0706.2000.890313.x>
- Monroy-Barbosa, A., & Bosland, P. W. (2008). Genetic Analysis of Phytophthora Root Rot Race-specific Resistance in Chile Pepper. *Journal of the American Society for Horticultural Science*, 133(6), 825–829. <https://doi.org/10.21273/jashs.133.6.825>

- Naegele, R. P., Mitchell, J., & Hausbeck, M. K. (2016). Genetic Diversity, Population Structure, and Heritability of Fruit Traits in *Capsicum annum*. *PLOS ONE*, 11(7), e0156969. <https://doi.org/10.1371/journal.pone.0156969>
- Paz, H. (2003). Root/Shoot Allocation and Root Architecture in Seedlings: Variation among Forest Sites, Microhabitats, and Ecological Groups¹. *BIOTROPICA*, 35(3), 318. <https://doi.org/10.1646/02120>
- Pérez-Martínez, A. L., Eguiarte, L. E., Mercer, K. L., Martínez-Ainsworth, N. E., McHale, L., van der Knaap, E., & Jardón-Barbolla, L. (2022). Genetic diversity, gene flow, and differentiation among wild, semiwild, and landrace chile pepper (*Capsicum annum*) populations in Oaxaca, Mexico. *American Journal of Botany*, 109(7), 1157–1176. <https://doi.org/10.1002/ajb2.16019>
- Pickersgill, B. (2016). Chile Peppers (*Capsicum* spp.). In: Lira, R., Casas, A., Blancas, J. (eds) *Ethnobotany of Mexico. Ethnobiology*. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-6669-7_17
- Rahman, M. H. ur, Ahmad, A., Wang, X., Wajid, A., Nasim, W., Hussain, M., Ahmad, B., Ahmad, I., Ali, Z., Ishaque, W., Awais, M., Shelia, V., Ahmad, S., Fahd, S., Alam, M., Ullah, H., & Hoogenboom, G. (2018). Multi-model projections of future climate and climate change impacts uncertainty assessment for cotton production in Pakistan. *Agricultural and Forest Meteorology*, 253-254, 94–113. <https://doi.org/10.1016/j.agrformet.2018.02.008>
- Robocker, W. C. (1970). Seed Characteristics and Seedling Emergence of Dalmatian Toadflax. *Weed Science*, 18(6), 720–725. <http://www.jstor.org/stable/4041641>
- Roucou, A., Violle, C., Fort, F., Roumet, P., Ecarnot, M., & Vile, D. (2017). Shifts in plant functional strategies over the course of wheat domestication. *Journal of Applied Ecology*, 55(1), 25–37. <https://doi.org/10.1111/1365-2664.13029>
- Saatkamp, A., Cochrane, A., Commander, L., Guja, Lydia K., Jimenez-Alfaro, B., Larson, J., Nicotra, A., Poschlod, P., Silveira, F. A. O., Cross, Adam T., Dalziell, E. L., Dickie, J., Erickson, T. E., Fidelis, A., Fuchs, A., Golos, P. J., Hope, M., Lewandrowski, W.,

- Merritt, D. J., & Miller, B. P. (2018). A research agenda for seed-trait functional ecology. *New Phytologist*, 221(4), 1764–1775. <https://doi.org/10.1111/nph.15502>
- Shi, J., & Lai, J. (2015). Patterns of genomic changes with crop domestication and breeding. *Current Opinion in Plant Biology*, 24, 47–53. <https://doi.org/10.1016/j.pbi.2015.01.008>
- Souza, M. and Fagundes, M. (2014). Seed Size as Key Factor in Germination and Seedling Development of *Copaifera langsdorffii* (Fabaceae). *American Journal of Plant Sciences*, 5, 2566-2573. doi: 10.4236/ajps.2014.517270.
- Spengler, R. N. (2020). Anthropogenic Seed Dispersal: Rethinking the Origins of Plant Domestication. *Trends in Plant Science*, 25(4), 340–348. <https://doi.org/10.1016/j.tplants.2020.01.005>
- Sun, L., Chen, J., Xiao, K., & Yang, W. (2017). Origin of the Domesticated Horticultural Species and Molecular Bases of Fruit Shape and Size Changes during the Domestication, Taking Tomato as an Example. *Horticultural Plant Journal*, 3(3), 125–132. <https://doi.org/10.1016/j.hpj.2017.07.007>
- Taitano, N., Bernau, V., Jardón-Barbolla, L., Leckie, B., Mazourek, M., Mercer, K., McHale, L., Michel, A., Baumler, D., Kantar, M., & Knaap, E. (2018). Genome-wide genotyping of a novel Mexican Chile Pepper collection illuminates the history of landrace differentiation after *Capsicum annum* L. domestication. *Evolutionary Applications*, 12(1), 78–92. <https://doi.org/10.1111/eva.12651>
- Tanabata, T., Shibaya, T., Hori, K., Ebana, K., & Yano, M. (2012). SmartGrain: High-Throughput Phenotyping Software for Measuring Seed Shape through Image Analysis. *Plant Physiology*, 160(4), 1871–1880. <https://doi.org/10.1104/pp.112.205120>
- Temesgen, B. (2017). Global role of plant breeding in tackling climate change. *International Journal of Agricultural Science and Food Technology*, 223–229. <https://doi.org/10.17352/2455-815x.000111>
- Ulft, L. H. (2004). The Effect of Seed Mass and Gap Size on Seed Fate of Tropical Rain Forest Tree Species in Guyana. *Plant Biology*, 6(2), 214–221. <https://doi.org/10.1055/s-2004-815736>

- Ullah, I., Jamil, S., Iqbal, M. Z., Shaheen, H. L., Hasni, S. M., Jabeen, S., Mehmood, A., & Akhter, M. (2012). Detection of bacterial blight resistance genes in basmati rice landraces. *Genetics and Molecular Research*, 11(3), 1960–1966.
<https://doi.org/10.4238/2012.july.20.1>
- Villa, T., Maxted, N., Scholten, M., & Ford-Lloyd, B. (2005). Defining and identifying crop landraces. *Plant Genetic Resources*, 3(3), 373-384. doi:10.1079/PGR200591
- Weeden, N. F. (2007). Genetic Changes Accompanying the Domestication of *Pisum sativum*: Is there a Common Genetic Basis to the “Domestication Syndrome” for Legumes?. *Annals of Botany*, 100(5), 1017–1025. <https://doi.org/10.1093/aob/mcm122>
- Zhang, H., Mittal, N., Leamy, L., Barazani, O., Song, B. (2017). Back into the wild—Apply untapped genetic diversity of wild relatives for crop improvement. *Evolutionary Applications*, 10(1), 5-24. <https://doi.org/10.1111/eva.12434>
- Zimmer, W. J., & Grose, R. J. (1958). ROOT SYSTEMS AND ROOT/SHOOT RATIOS OF SEEDLINGS OF SOME VICTORIAN EUCALYPTS. *Australian Forestry*, 22(1), 13–18.
<https://doi.org/10.1080/00049158.1958.10675840>
- Zipperlen, S. W., & Malcolm C. Press. (1996). Photosynthesis in Relation to Growth and Seedling Ecology of Two Dipterocarp Rain Forest Tree Species. *Journal of Ecology*, 84(6), 863–876. <https://doi.org/10.2307/2960558>

Tables

Table 1. Accessions of *C. annuum* from locations within the state of Oaxaca, Mexico and the U.S. with corresponding domestication, location, cultivation type, and elevation data.

Accession	Domestication	Location	Region	Cultivation	Elevation (m)	Annual Precipitation (mm)	Annual Mean Temperature (°C)
Early Jalapeno	Commercial	New Mexico, U.S.	NA	NA	NA	NA	NA
PI586666	Commercial	California, U.S.	California, U.S.	NA	NA	NA	NA
CM334	Landrace	New Mexico, U.S. ^a	New Mexico, U.S.	NA	NA	NA	NA
ca0437	Landrace	Arriaga, Chiapas	Chiapas Coast, MX	Backyard	65	1623	24.46
ca0060	Landrace	Copalita, Oaxaca	Oaxaca Coast, MX	Backyard	20	1203	27.19
ca0300	Landrace	Saachilac, Santiago Astata, Oaxaca	Oaxaca Coast, MX	Backyard	20	892	27.61
ca0302	Landrace	Saachilac, Santiago Astata, Oaxaca	Oaxaca Coast, MX	Backyard	20	892	27.61
ca0345	Landrace	San Pablo Coatlan, Oaxaca	Oaxaca, Central valleys, MX	Backyard	1440	1054	25.56
ca0167	Landrace	Los Reyes, Pochutla, Chiapas	Chiapas Coast, MX	Backyard	222	1082	25.87
ca0166	Landrace	Los Reyes, Pochutla, Chiapas	Chiapas Coast, MX	Backyard	222	1082	25.87
ca0168	Landrace	Los Reyes, Pochutla, Chiapas	Chiapas Coast, MX	Milpa	222	1082	25.87
ca0447	Landrace	Nuevo Colosio, Chiapas	Chiapas Coast, MX	Backyard	103	1490	27.94
ca0260	Landrace	Rosedal, close to Tomatal, Oaxaca	Oaxaca Coast, MX	Plantation	26	934	26.48
ca0285	Landrace	Rosedal, close to Tomatal, Oaxaca	Oaxaca Coast, MX	Plantation	26	934	26.48
canrose2b	Semiwild	Rosedal, close to Tomatal, Oaxaca	Oaxaca Coast, MX	NA	26	934	26.48
ca0592	Semiwild	Xpujil, Campeche	Peninsula Yucatan, MX	Backyard	257	1040	24.75
ca0436	Wild	Arriaga, Chiapas	Chiapas Coast, MX	Backyard	65	1623	24.46
ca0305	Wild	Saachilac, Santiago Astata, Oaxaca	Oaxaca Coast, MX	Forest	20	892	27.61
ca0303	Wild	Saachilac, Santiago Astata, Oaxaca	Oaxaca Coast, MX	Forest	20	892	27.61
ca0220	Wild	Polvorin, Pochutla, Chiapas	Chiapas Coast, MX	Milpa	260	1007	26.30
ca0293	Wild	Rosedal, close to Tomatal, Oaxaca	Oaxaca Coast, MX	Backyard	48	934	26.48

^aCM334 was originally a landrace from Morelos, Mexico but has been maintained in New Mexico, U.S. for breeding.

Table 2. AIC table of linear mixed-effects models for repeated measures seedling phenotypes of *C. annuum* by domestication, location, and time. Models fit using “lmer” package in R.

Model	Height to the first node	True leaf area	Total number of true leaves
Unstructured	384.0905*	138.0361	720.5512*
Diagonal	385.4989	136.0851*	723.0037

Note: * indicates best fitting covariance-variance structure in model

Table 3. ANOVA results from preliminary analyses of significant effects of accession on seed phenotypes for *C. annuum* using Satterhwaite’s method (F value_{numerator df, denominator df}).

Source	Area	Weight (g)	Thickness (mm)	Perimeter	Circularity
Accession	85.275 _{16, 17} ****	40.692 ₁₆ , 16****	4.5652 ₁₆ , 16**	42.3830 ₁₆ , 17****	4.8322 ₁₆ , 16**

NS P > 0.10, + P < 0.10, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

Table 4. ANOVA results showing the effect of domestication and location on seed phenotypes of *C. annuum* using Satterhwaite’s method (F value_{numerator df, denominator df}).

Source	Area	Weight (g)	Thickness (mm)	Perimeter	Circularity
Domestication	4.2749 _{3, 29} *	0.1115 ₃ , 29 NS	2.2810 _{3, 28} NS	6.1296 ₃ , 29**	3.9940 _{3, 28} *
Location	6.6446 _{7, 29} ****	4.6701 ₇ , 21**	1.4846 _{7, 28} NS	7.8694 ₇ , 29****	2.9901 _{7, 28} *
Domestication:Location	0.7132 _{3, 29} NS	0.1681 ₃ , 29 NS	3.9477 _{3, 28} *	1.8449 _{3, 29} NS	2.6860 _{3, 28} +

NS P > 0.10, + P < 0.10, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

Table 5. Mean separation using Tukey method of seed phenotypes of *C. annuum* by location. Means with the same letter are not significant. The Ohio State University, 2023.

	Xpujil	Pochutla	Rosedal	Arriaga	Saachilac	Nuevo Celosio	Copalita	U.S.
Area (cm²)	0.0343 ± 0.0139 a	0.0734 ± 0.0069 a	0.0827 ± 0.0069 a	0.0748 ± 0.0089 a	0.0809 ± 0.0069 a	0.0919 ± 0.0139 ab	0.0978 ± 0.0139 ab	0.1 312 ± 0.0 069 b
Weight (g)	0.0017 ± 0.0009 a	0.0039 ± 0.0005 a	0.0036 ± 0.0005 a	0.0042 ± 0.0006 a	0.0039 ± 0.0046 a	0.0049 ± 0.0091 a	0.0053 ± 0.0009 ab	0.0 082 ± 0.0 005 b
Thickness (mm)	0.6720 ± 0.0852	0.9865 ± 0.0430	0.9013 ± 0.0430	0.8673 ± 0.0546	0.9129 ± 0.0430	1.0788 ± 0.0852	0.9339 ± 0.0852	1.0 063 ± 0.0 430
Circularity (%)	83.65 ± 2.47 a	79.09 ± 1.24 a	77.16 ± 1.24 a	81.13 ± 1.58 a	79.28 ± 1.24 a	79.69 ± 2.47 a	81.23 ± 2.47 a	75. 39 ± 1.2 4 a
Perimeter (cm)	0.7149 ± 0.0969 a	1.0768 ± 0.0485 b	1.1557 ± 0.0485 b	1.074 ± 0.0619 ab	1.1304 ± 0.0485 b	1.2042 ± 0.0969 bc	1.2296 ± 0.0969 bc	1.4 699 ± 0.0 485 c

Table 6. Mean separation using Tukey method of seed phenotypes of *C. annuum* by domestication. Means with the same letter are not significant. The Ohio State University, 2023.

	Commercial	Landrace	Semiwild	Wild
Area (cm²)	0.1173 ± 0.0123 c	0.0957 ± 0.0049 bc	0.0479 ± 0.0123 a	0.0714 ± 0.0078 ab
Weight (g)	0.0083 ± 0.0009	0.0049 ± 0.0004	0.0028 ± 0.0009	0.0037 ± 0.0006
Thickness (mm)	1.0575 ± 0.0605	0.9681 ± 0.0268	0.7898 ± 0.0605	0.8575 ± 0.0395
Circularity (%)	75.27 ± 1.75	77.96 ± 0.70	83.03 ± 1.75	79.98 ± 1.11
Perimeter (cm)	1.3893 ± 0.0831 b	1.2335 1.3893 ± 0.0333 b	0.8411 ± 0.0831 a	1.0573 ± 0.0526 a

Table 7. ANOVA results from preliminary analyses of significant effects of accession on seedling phenotypes for *C. annuum* using Satterhwaite's method (F value_{numerator df, denominator df}).

Source	Height to first node (27 DAP) (cm)	Cotyledon area (23 DAP) (cm²)	True leaf area (28 DAP) (cm²)	Total number of true leaves (29 DAP)
Accession	9.9994 _{19, 42.2} ****	19.675 _{19, 42.2} ****	1.6952 _{17, 38.1} +	3.2574 _{18, 40.129} ***

NS P > 0.10, + P < 0.10, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

Table 8. ANOVA results showing the effect of domestication and location on seedling phenotypes of *C. annuum* using Satterhwaite's method (F value_{numerator df, denominator df}).

Source	Height to first node (27 DAP) (cm)	Cotyledon area (23 DAP) (cm²)	True leaf area (28 DAP) (cm²)	Total number of true leaves (29 DAP)
Domestication	2.7783 _{3, 48.5} +	3.5043 _{3, 48.8} *	2.5740 _{3, 44.2} +	2.4696 _{3, 46.6} NS
Location	4.84298 _{48.2} ***	0.5482 _{8, 48} NS	0.4303 _{7, 44} NS	2.4936 _{7, 46.1} *
Domestication:Location	1.6469 _{2, 48.4} NS	0.1106 _{2, 48.5} NS	0.2966 _{1, 44} NS	0.4991 _{2, 46.3} NS

NS P > 0.10, + P < 0.10, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

Table 9 Mean separation using Tukey method of seedling phenotypes of *C. annuum* by location. Means with the same letter are not significant. The Ohio State University, 2023.

	Xpujil	Pochutla	Rosedal	Arriaga	Saachilac	Nuevo Celoso	Copaliata	San Pablo Coatlan	U.S.
Height to first node (27 DAP) (cm)	1.0376 ± 0.3413 a	1.3583 ± 0.1274 a	1.2070 ± 0.1463 1 a	1.3004 ± 0.1455 a	1.3505 ± 0.1232 a	1.325 6 ± 0.208 2 a	1.2333 ± 0.1846 a	1.7425 ± 0.1846 a	2.415 9 ± 0.140 7 b
Cotyledon area (23 DAP) (cm²)	0.1486 ± 0.5997 3	0.5160 ± 0.1901	0.5254 ± 0.2128	0.5043 ± 0.2368	0.5249 ± 0.1741	0.481 6 ± 0.300 8	0.4119 ± 0.3024	0.5372 ± 0.3024	1.732 1 ± 0.212 9
True leaf area (28 DAP) (cm²)	NA	0.2888 ± 0.1132	0.2467 ± 0.1173	0.2787 ± 0.1340	0.3567 ± 0.1103	0.442 7 ± 0.142 1	0.2308 ± 0.1421	0.3178 ± 0.1421	0.449 1 ± 0.117 3
Total number of true leaves (29 DAP)	NA	1.5778 ± 0.3774 a	2.7422 ± 0.4081 ab	1.2852 ± 0.4655 a	1.6814 ± 0.3625 ab	2.274 2 ± 0.537 0 ab	1.9167 ± 0.5370 ab	1.6417 ± 0.5370 ab	2.951 5 ± 0.408 2 b

Table 10. Mean separation using Tukey method of seedling phenotypes of *C. annuum* by domestication. Means with the same letter are not significant. The Ohio State University, 2023.

	Commercial	Landrace	Semiwild	Wild
Height to first node (27 DAP) (cm)	2.5592 ± 0.1869	1.4666 ± 0.0958	0.9059 ± 0.2912	1.2049 ± 0.1345
Cotyledon area (23 DAP) (cm²)	2.2349 ± 0.2489 b	0.5909 ± 0.1153 a	0.1431 ± 0.3953 a	0.4075 ± 0.1812 a
True leaf area (28 DAP) (cm²)	0.5517 ± 0.1331	0.3078 ± 0.1040	-0.1972 ± 0.2241	0.3999 ± 0.1282
Total number of true leaves (29 DAP)	2.9093 ± 0.5635	2.1246 ± 0.2936	1.5261 ± 1.0411	1.0378 ± 0.4214

Table 11. ANOVA results showing the effect of domestication, location, and time (DAP) on seedling phenotypes of *C. annuum* using Satterhwaite's method (F value_{numerator df, denominator df}).

Source	Height to first node (cm)	True leaf area (cm²)	Total number of true leaves
Domestication	1.4387 _{3, 343.3} NS	11.3504 _{3, 149.3} ****	3.7117 _{3, 318.7} *
DAP	135.0548 _{1, 8.2} ****	0.3325 _{1, 81.4} NS	45.5849 _{1, 5.9} ****
Location	0.6405 _{8, 345.1} NS	0.3662 _{8, 148.7} NS	1.0597 _{7, 318.99} NS
Domestication:DAP	5.8127 _{3, 343.2} ***	1.5980 _{2, 149.2} NS	10.7767 _{3, 317.9} ****
Domestication:Location	0.5640 _{2, 344.8} NS	0.3400 _{1, 149.1} NS	0.7263 _{2, 318.9} NS
DAP:Location	2.8576 _{8, 345.26} **	0.3400 _{1, 149.9} NS	4.7456 _{7, 318.9} ****
Domestication:DAP:Location	1.2444 _{2, 345.2} NS	NA	1.5466 _{2, 318.9} NS

NS P > 0.10, + P < 0.10, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

Table 12. ANOVA results from preliminary analyses of significant effects of accession on root and shoot biomass for *C. annuum* using Satterhwaite’s method (F value_{numerator df, denominator df}).

Source	Root mass (g)	Shoot mass (g)	Root:Shoot
Accession	3.2347 _{16, 35.7} **	2.9853 _{16, 35.5} **	1.6147 _{16, 34.9} NS

NS P > 0.10, + P < 0.10, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

Table 13. ANOVA results from analyses of significant effects of domestication and location on root and shoot biomass for *C. annuum* using Satterhwaite’s method (F value_{numerator df, denominator df}).

Source	Root mass (g)	Shoot mass (g)	Root:Shoot
Domestication	0.3350 _{3, 34.9} NS	0.9566 _{3, 34.8} NS	5.7351 _{3, 33.7} **
Location	4.8831 _{5, 34.7} **	5.4772 _{5, 34.6} ***	1.8122 _{5, 33.6} NS
Domestication:Location	0.1070 _{1, 34.8} NS	0.1271 _{1, 34.6} NS	6.5472 _{1, 33.6} *

NS P > 0.10, + P < 0.10, * P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001

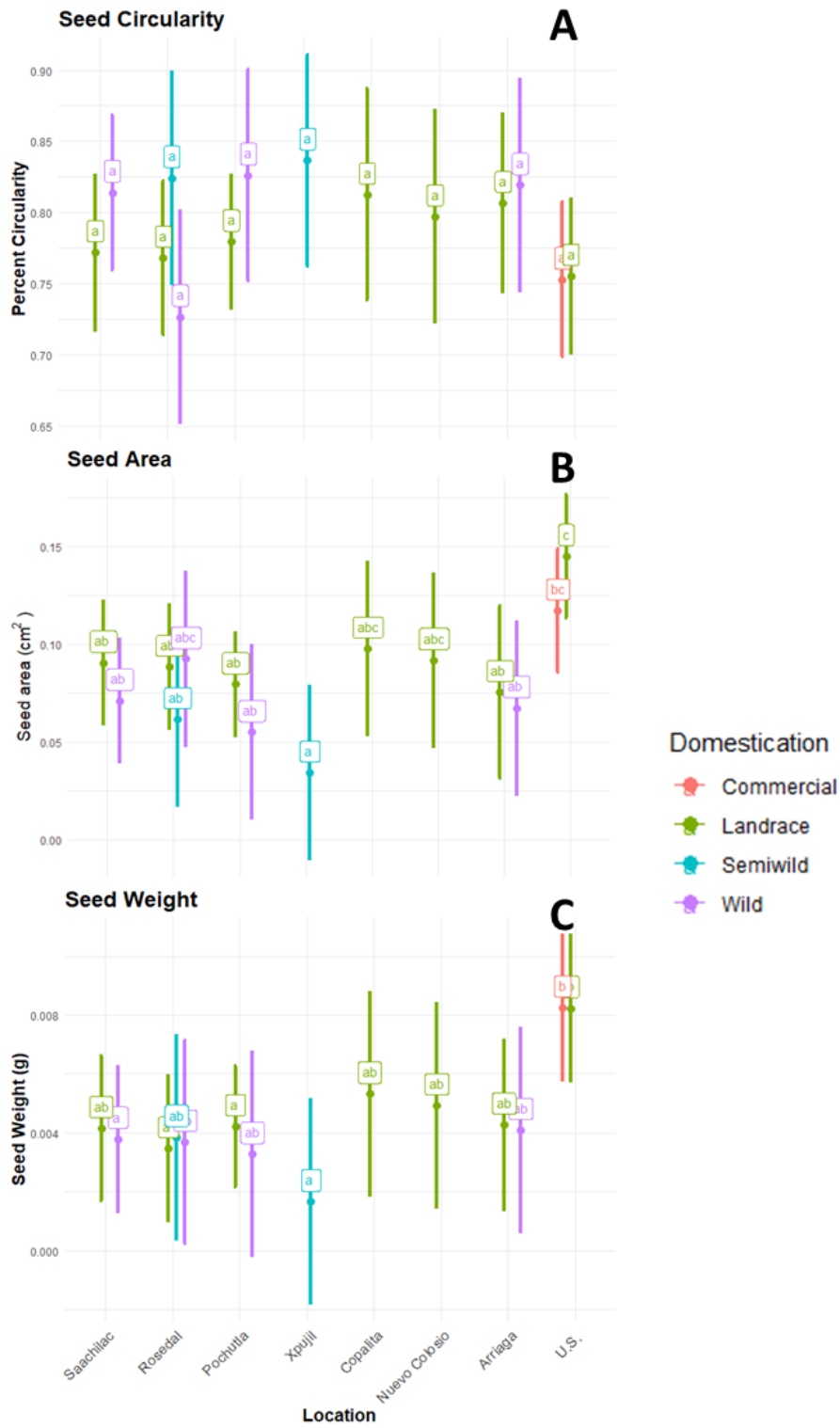
Table 14. Mean separation using Tukey method of root and shoot biomass allocation of *C. annuum* by domestication. Means with the same letter are not significant. The Ohio State University, 2023.

	Commercial	Landrace	Semiwild	Wild
Root mass (g)	0.0136 ± 0.0028	0.0039 ± 0.0009	0.0019 ± 0.0049	0.0017 ± 0.0016
Shoot mass (g)	0.0165 ± 0.0046	0.0068 ± 0.0016	0.0011 ± 0.0079	0.0037 ± 0.0028
R:S	0.8227 ± 0.1341 a	0.6452 ± 0.0374 a	1.1492 ± 0.2388 a	0.5162 ± 0.0772 a

Table 15. Mean separation using Tukey method of root and shoot biomass allocation of *C. annuum* by location. Means with the same letter are not significant. The Ohio State University, 2023.

	Pochutla	Rosedal	Arriaga	Saachilac	Nuevo Celosio	Copalita	San Pablo Coatlan	U.S.
Root mass (g)	0.0021 ± 0.0014 a	0.0034 ± 0.0015 a	0.0018 ± 0.0016 a	0.0034 ± 0.0012 a	0.0022 ± 0.0020 a	0.0007 ± 0.0028 a	0.0035 ± 0.0020 a	0.0135 ± 0.0016 b
Shoot mass (g)	0.0041 ± 0.0023 a	0.0064 ± 0.0025 a	0.0034 ± 0.0026 a	0.0058 ± 0.0021 a	0.0039 ± 0.0033 a	0.0014 ± 0.0046 a	0.0043 ± 0.0033 a	0.0201 ± 0.0026 b
R:S	0.5730 ± 0.0809	0.6259 ± 0.0901	0.6990 ± 0.0955	0.6261 ± 0.0707	0.6052 ± 0.1249	0.4685 ± 0.1802	0.7936 ± 0.1249	0.7277 ± 0.1045

Figures



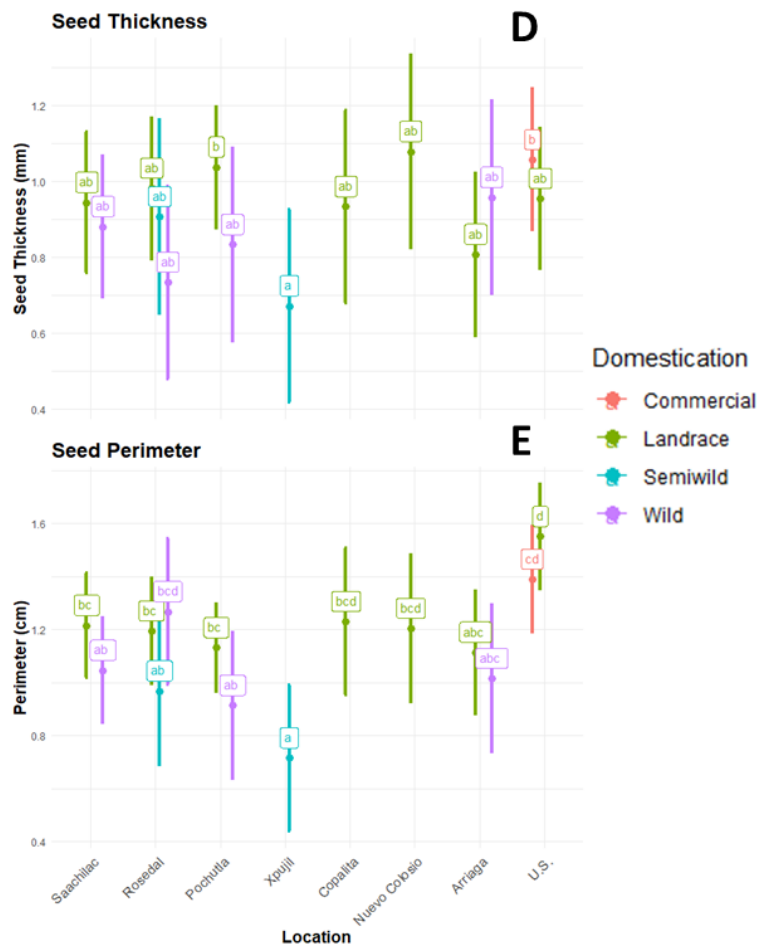


Figure 1. Mean separation using the Tukey test of seed phenotypes of *C. annuum* by location and domestication. Means with the same letter are not significant. (A) Seed circularity. Values were extracted from seed imaging and analyzed using ImageJ software in percent circularity (0-100%). (B) Seed area. Values were extracted from seed imaging and analyzed using ImageJ software in cm^2 . (C) Seed weight. Measured in g. (D) Seed Thickness. Measured in mm. (E) Seed perimeter. Values were extracted from seed imaging and analyzed using ImageJ software in cm^2 . Separation is denoted by both domestication level and location grouping, with locations along the x axis in order of lowest to highest annual precipitation from left to right. Points indicate estimated marginal means and error bars represent standard error. The Ohio State University, 2023.

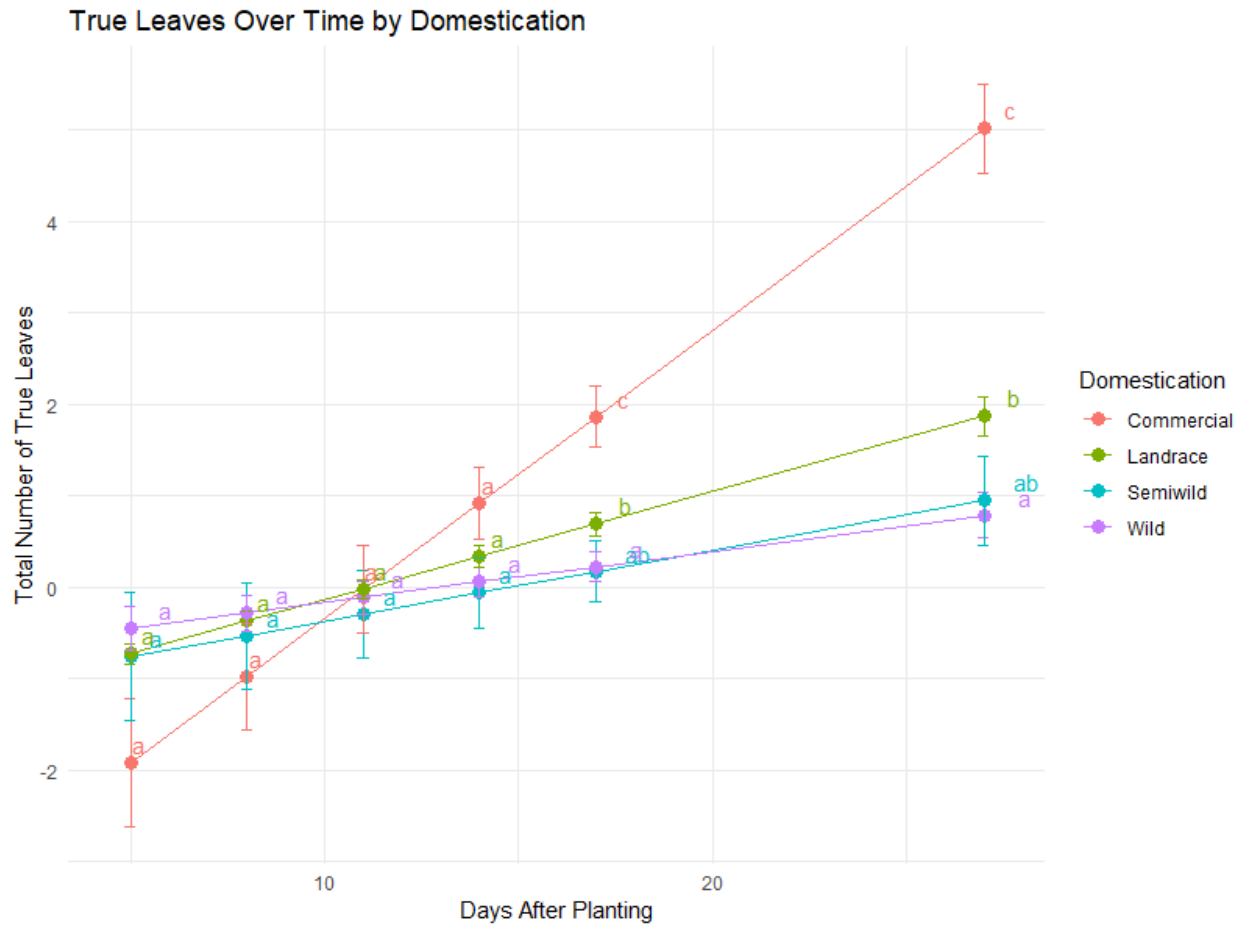


Figure 2. Mean separation using the Tukey test of total number of true leaves of *C. annuum* over time by domestication. Means with the same letter are not significant. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

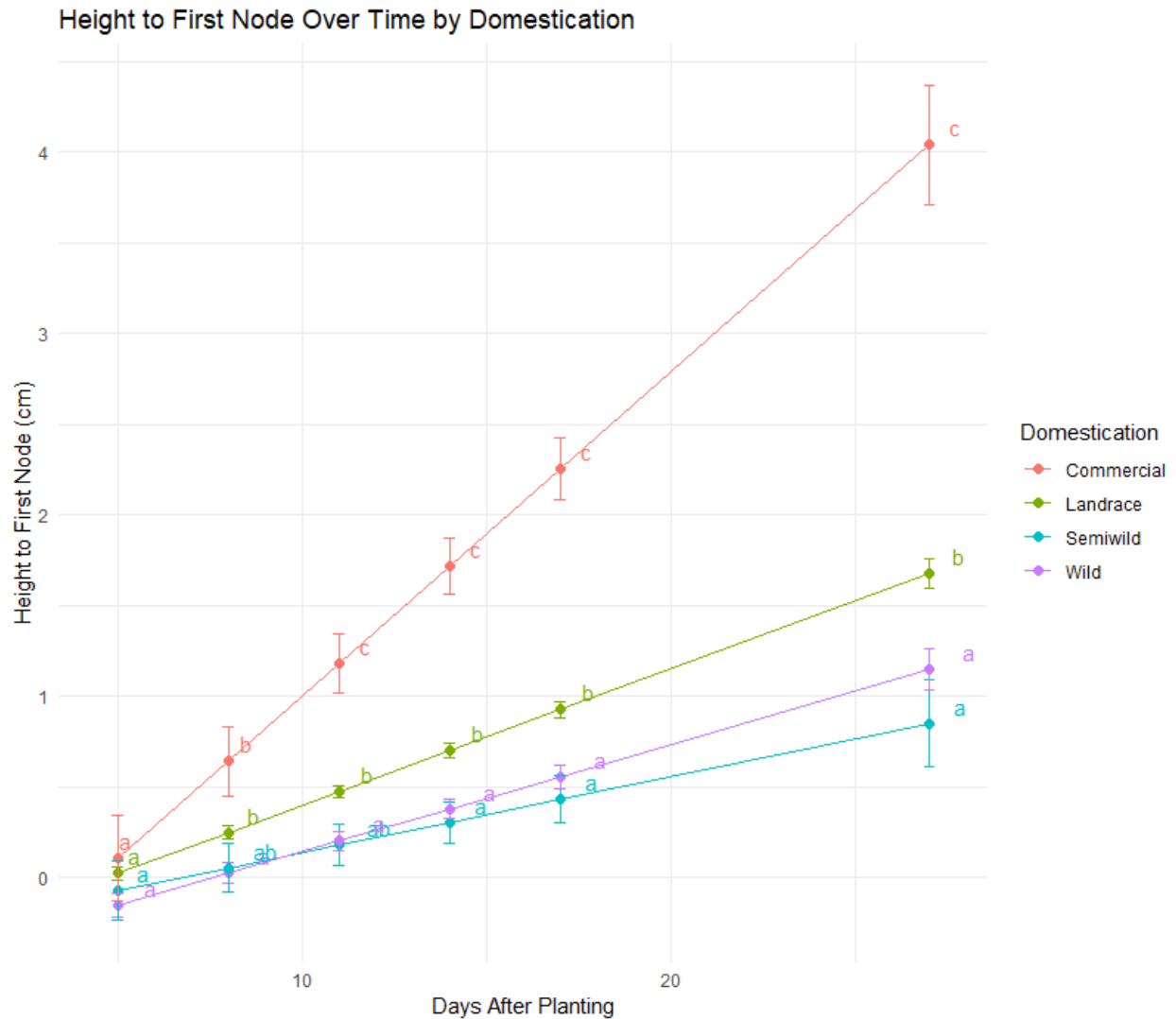


Figure 3. Mean separation using the Tukey test of height to the first node of *C. annuum* over time by domestication. Means with the same letter are not significant. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

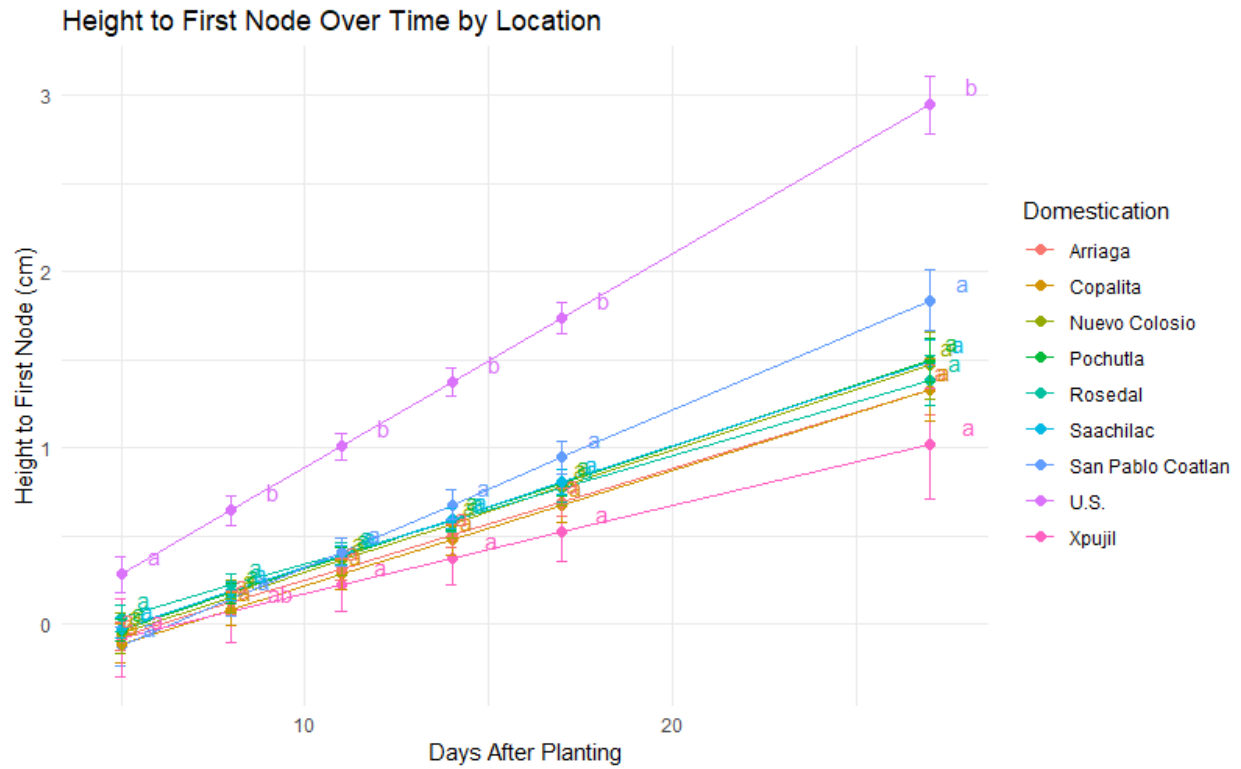


Figure 4. Mean separation using the Tukey test of height to the first node of *C. annuum* over time by location. Means with the same letter are not significant. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

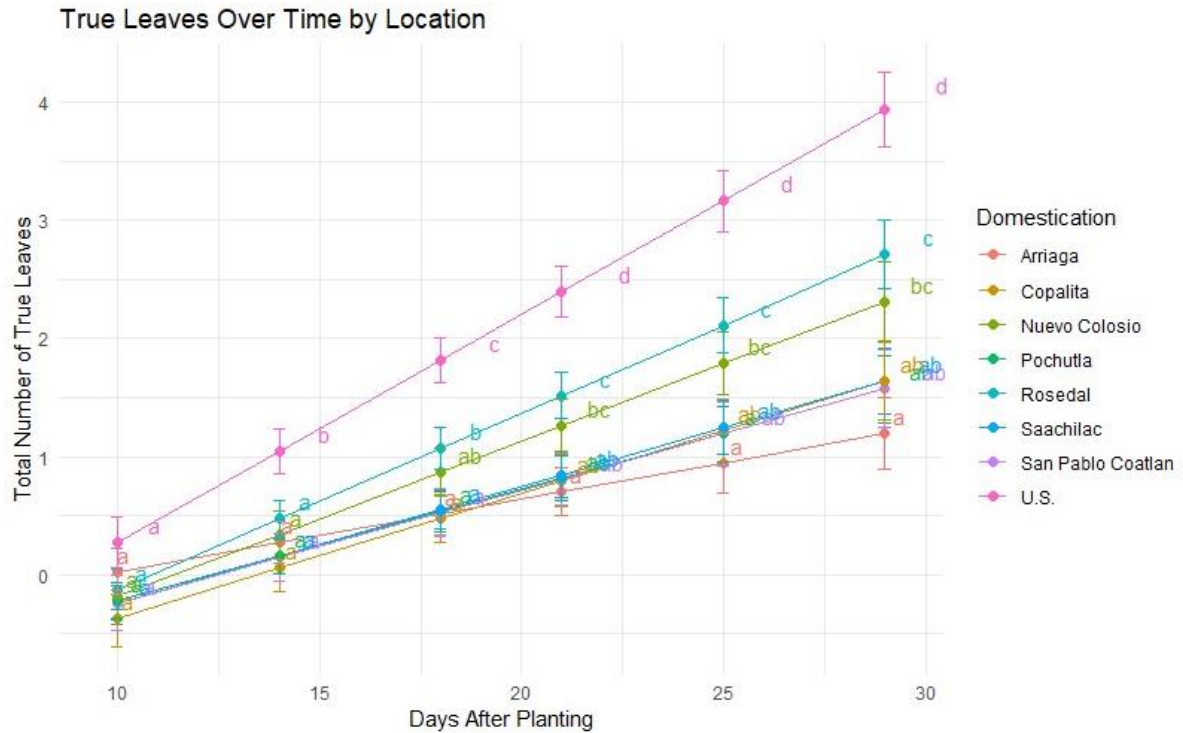


Figure 5. Mean separation using the Tukey test of total number of true leaves of *C. annuum* over time by location. Means with the same letter are not significant. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

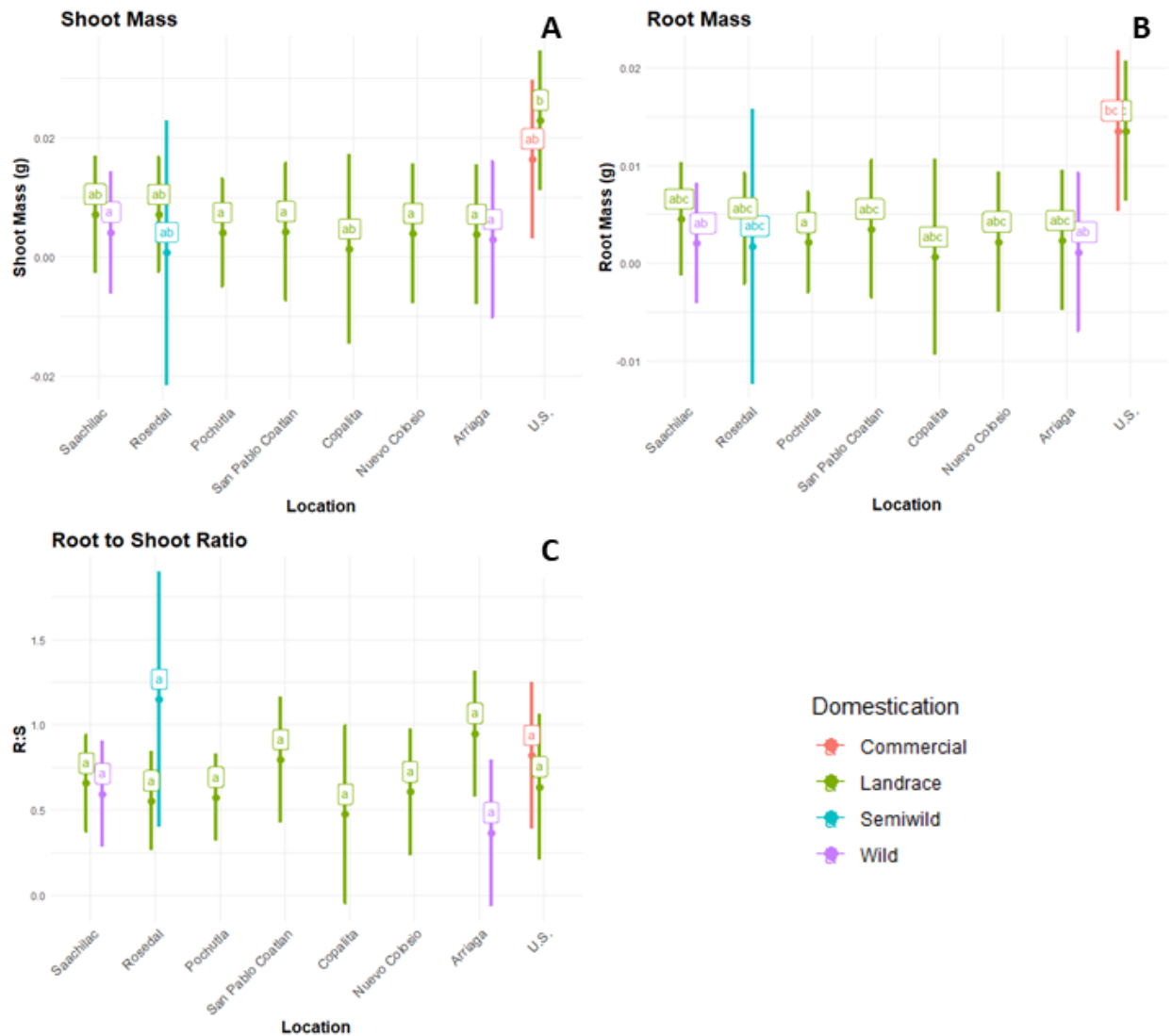
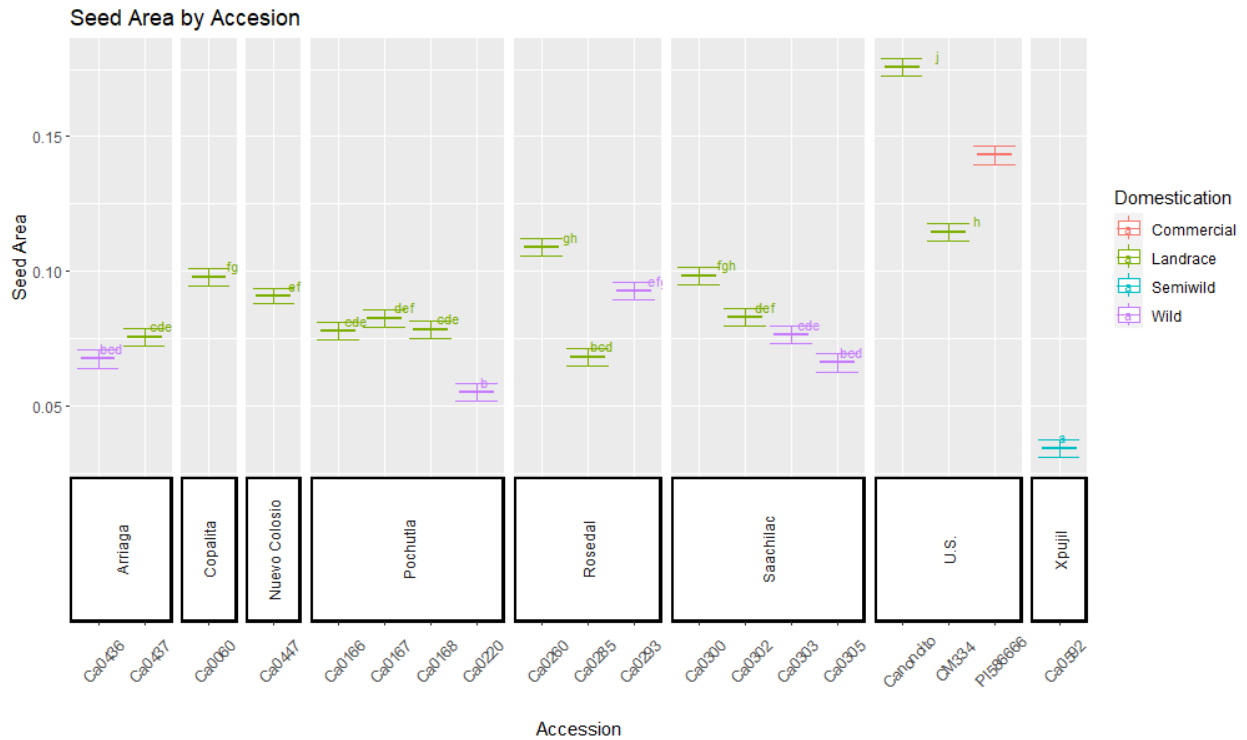
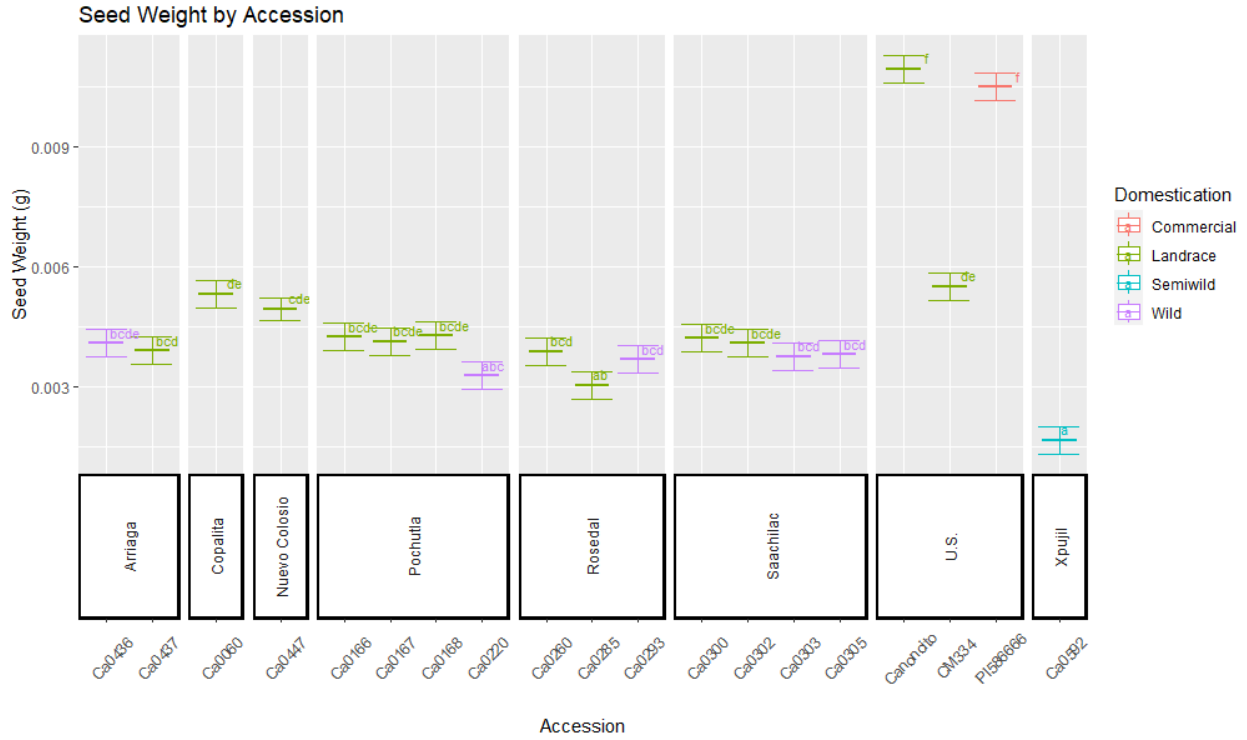


Figure 6. Mean separation using the Tukey test of root and shoot dry weight measurements of *C. annuum* by location and domestication. Means with the same letter are not significant. (A) Shoot mass measured in g. (B) Root mass measured in g. (C) Root to shoot ratio as R:S. Separation is denoted by both domestication level and location grouping, with locations along the x axis in order of lowest to highest annual precipitation from left to right. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

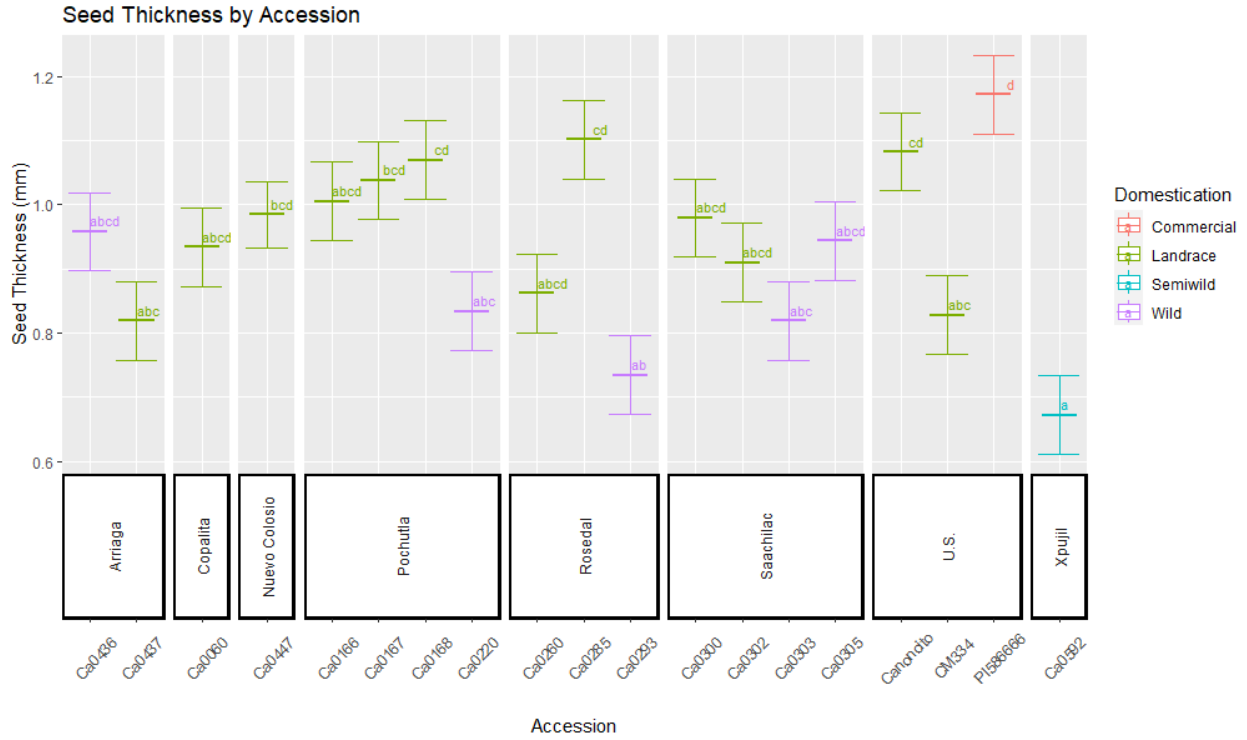
Supplementary Materials



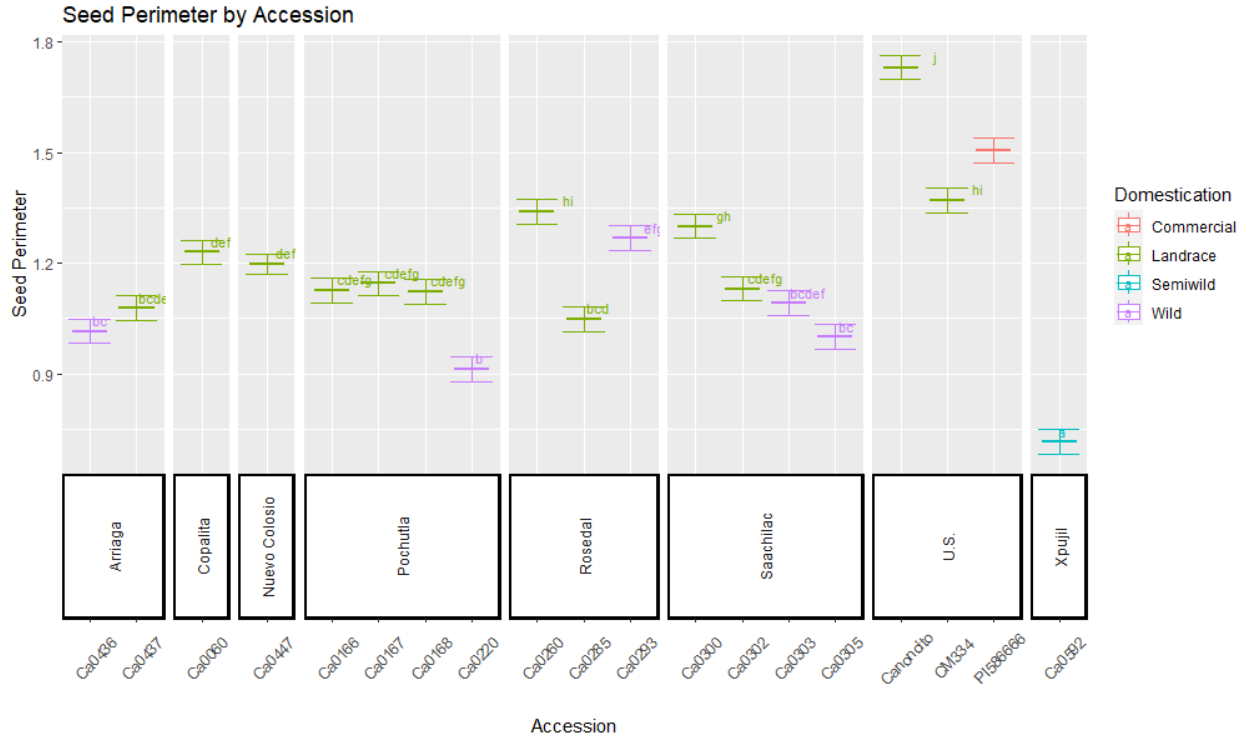
Supplementary Figure 1. Mean separation using Tukey's test of *C. annuum* seed area across accessions from Oaxaca, Mexico using Tukey's test. Means with the same letter are not significant. Values were extracted from seed imaging and analyzed using ImageJ software. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. The Ohio State University, 2023.



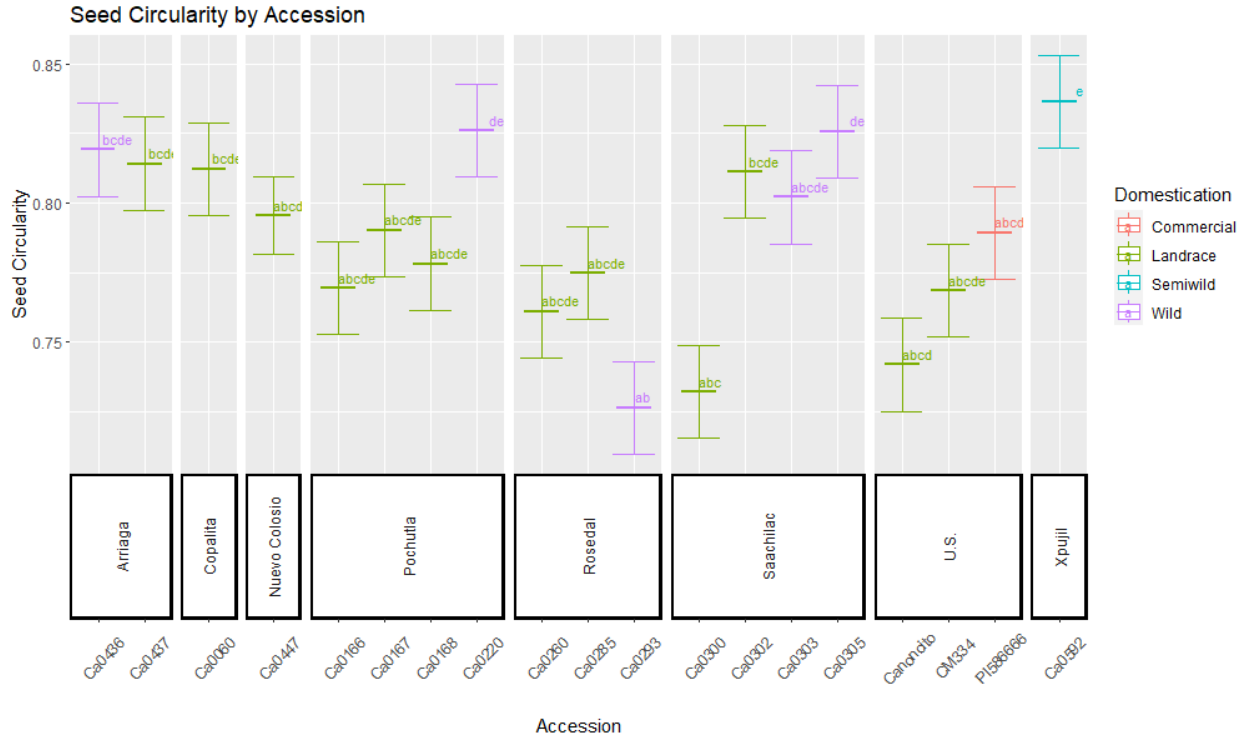
Supplementary Figure 2. Mean separation using Tukey’s test of *C. annuum* seed weight (g) across accessions from Oaxaca, Mexico using Tukey’s test. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. The Ohio State University, 2023.



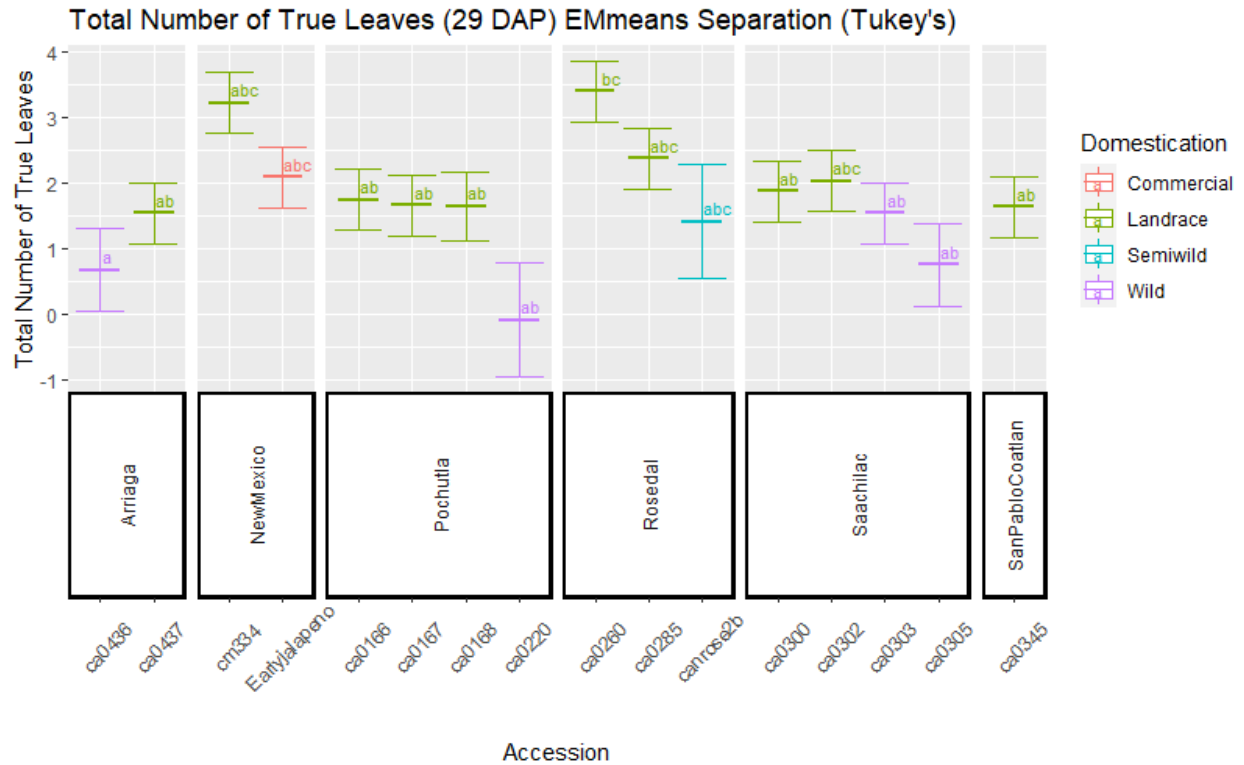
Supplementary Figure 3. Mean separation using Tukey’s test of *C. annuum* seed thickness across accessions from Oaxaca, Mexico using Tukey’s test. Means with the same letter are not significant. Values were extracted from seed imaging and analyzed using ImageJ software. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. The Ohio State University, 2023.



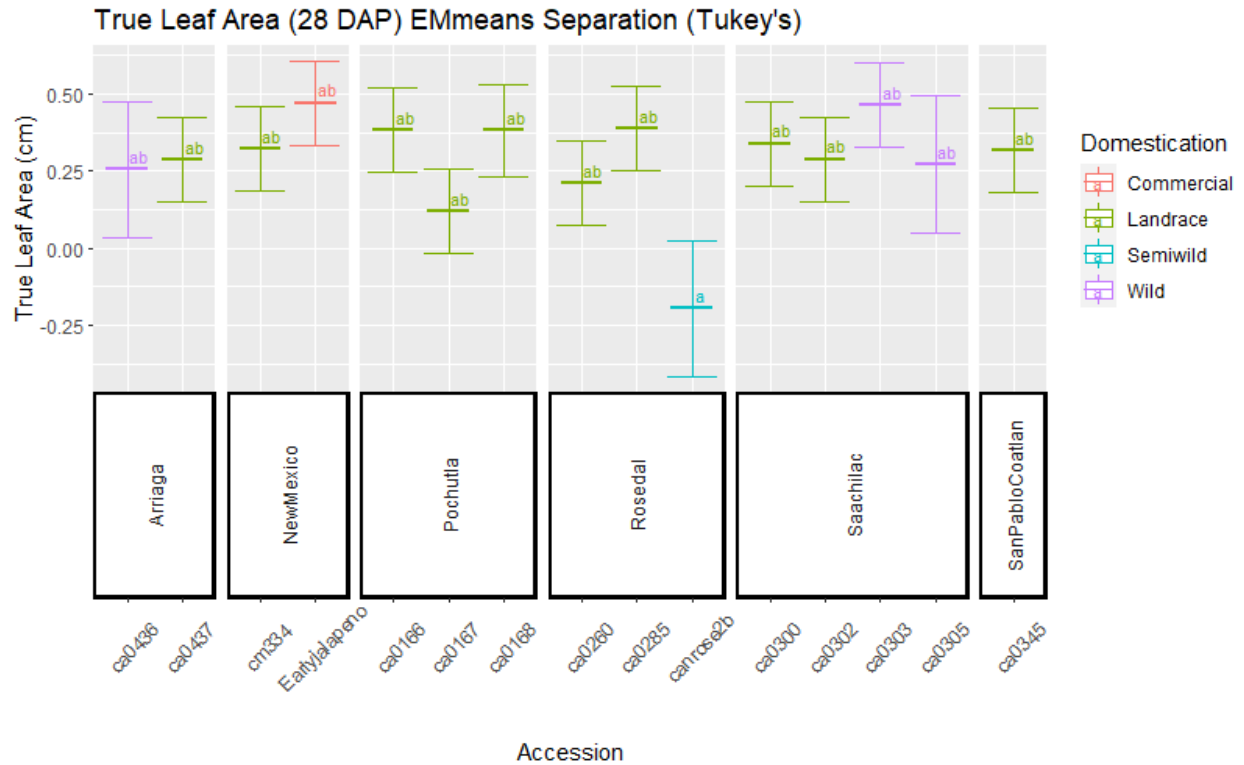
Supplementary Figure 4. Mean separation using Tukey’s test of *C. annuum* seed perimeter across accessions from Oaxaca, Mexico using Tukey’s test. Means with the same letter are not significant. Values were extracted from seed imaging and analyzed using ImageJ software. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. The Ohio State University, 2023.



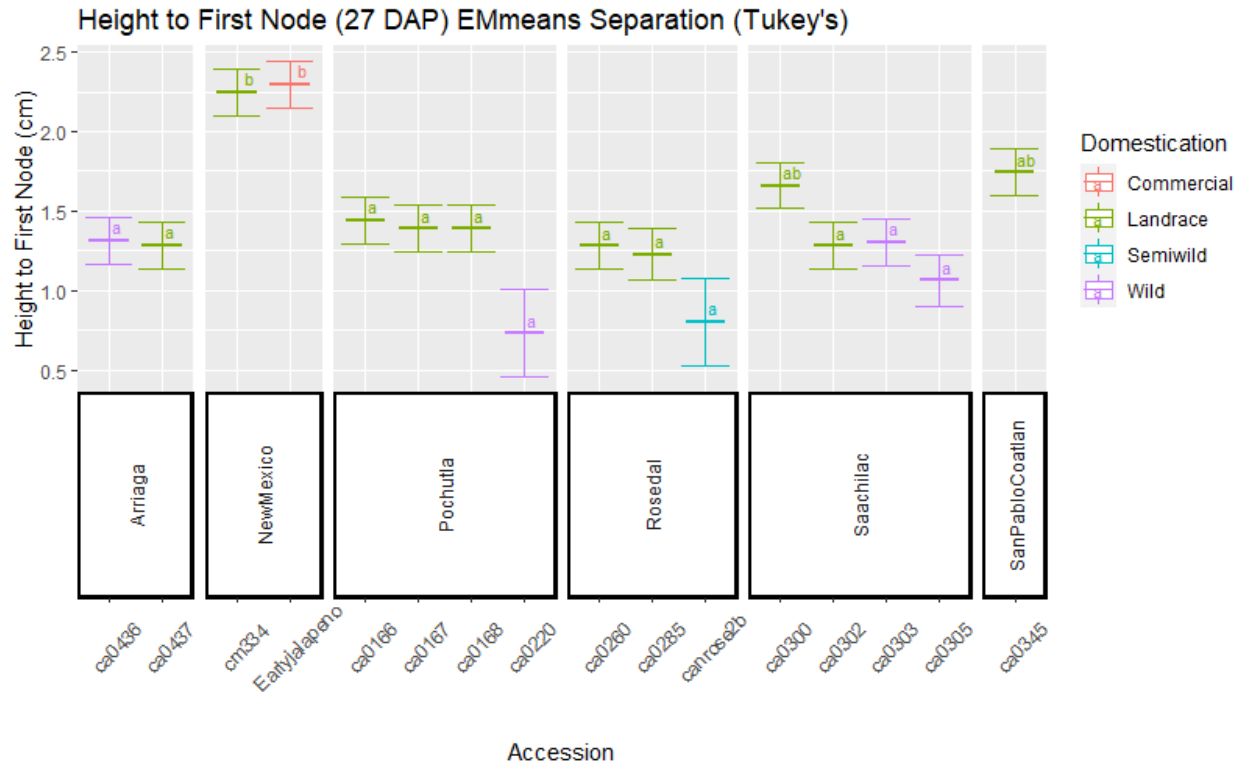
Supplementary Figure 5. Mean separation using Tukey's test of *C. annuum* seed circularity across accessions from Oaxaca, Mexico using Tukey's test. Means with the same letter are not significant. Values were extracted from seed imaging and analyzed using ImageJ software. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. The Ohio State University, 2023.



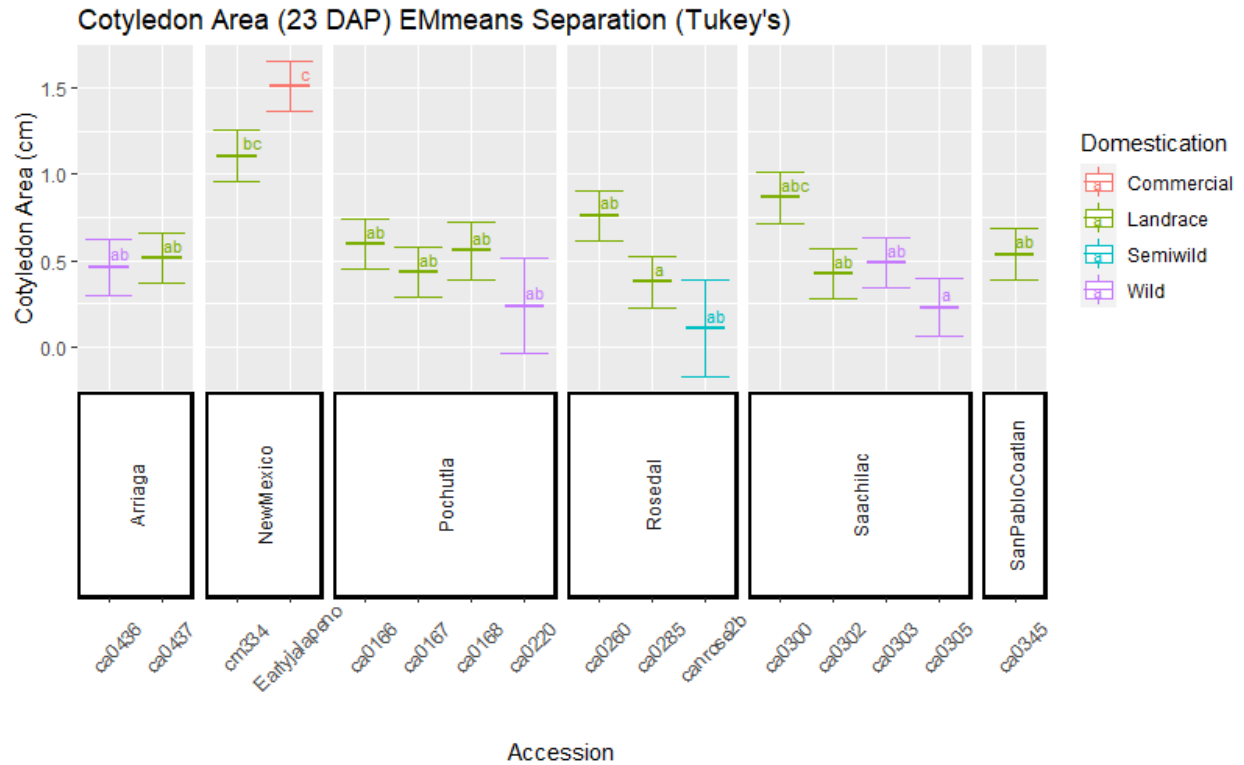
Supplementary Figure 6. Mean separation using Tukey's test of the total number of true leaves at 29 DAP for *C. annuum* across accessions from Oaxaca, Mexico using Tukey's test. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



Supplementary Figure 7. Mean separation using Tukey’s test of true leaf area (cm²) at 28 DAP for *C. annuum* across accessions from Oaxaca, Mexico using Tukey’s test. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

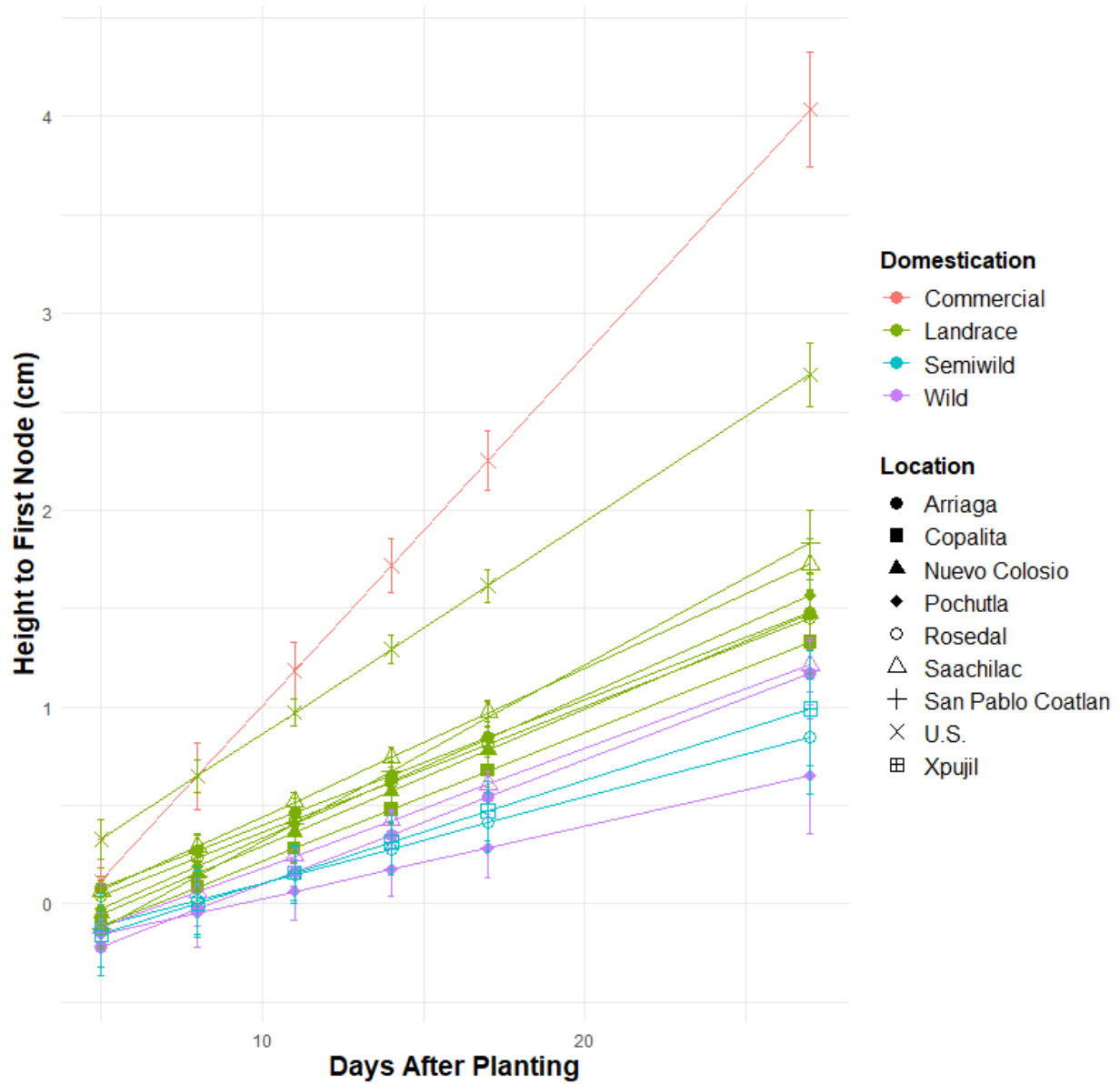


Supplementary Figure 8. Mean separation using Tukey’s test of the height to the first node (cm) at 27 DAP for *C. annuum* across accessions from Oaxaca, Mexico using Tukey’s test. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

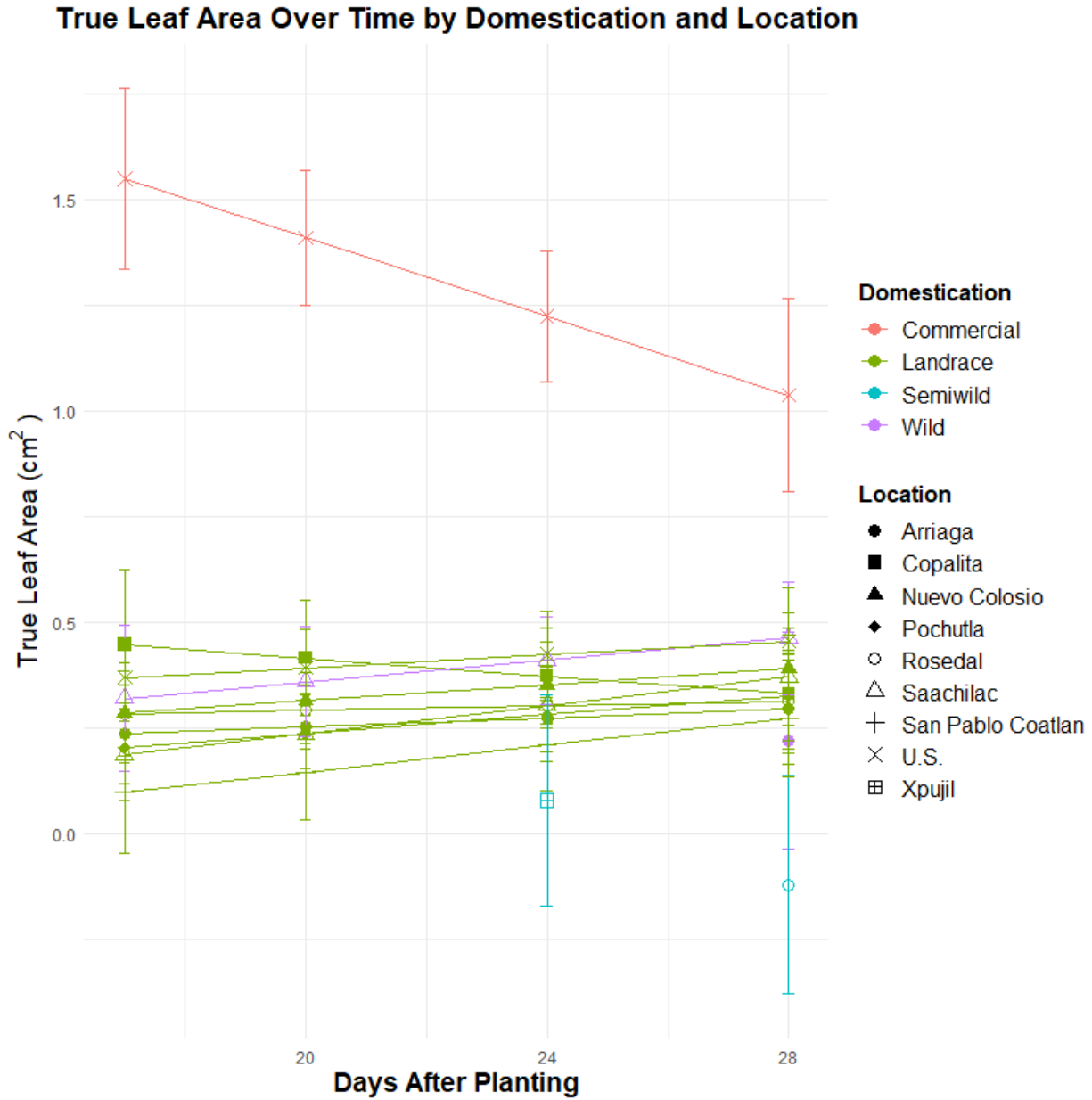


Supplementary Figure 9. Mean separation using Tukey's test of cotyledon area (cm^2) at 23 DAP for *C. annuum* across accessions from Oaxaca, Mexico using Tukey's test. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.

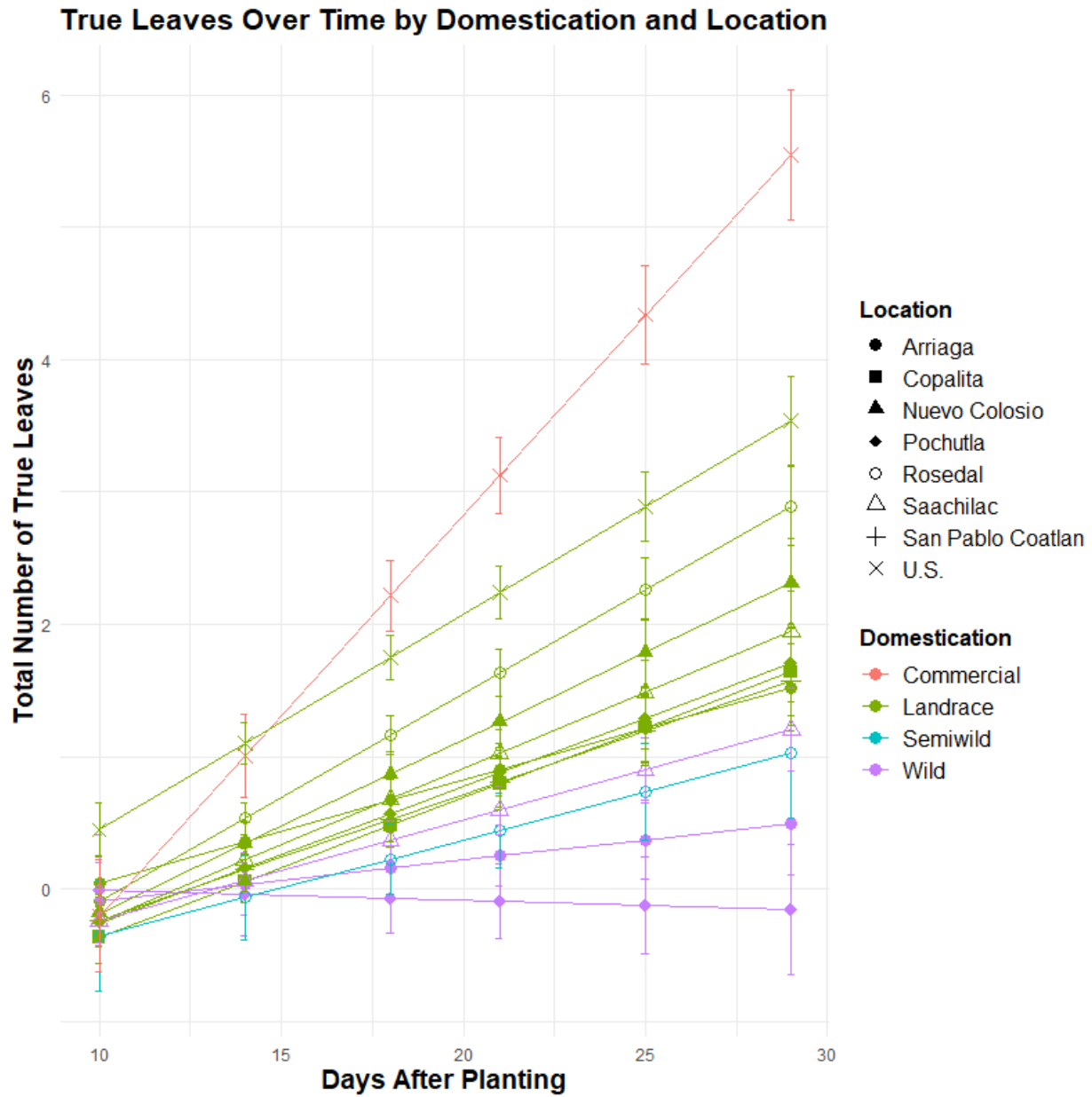
Height to First Node Over Time by Domestication and Location



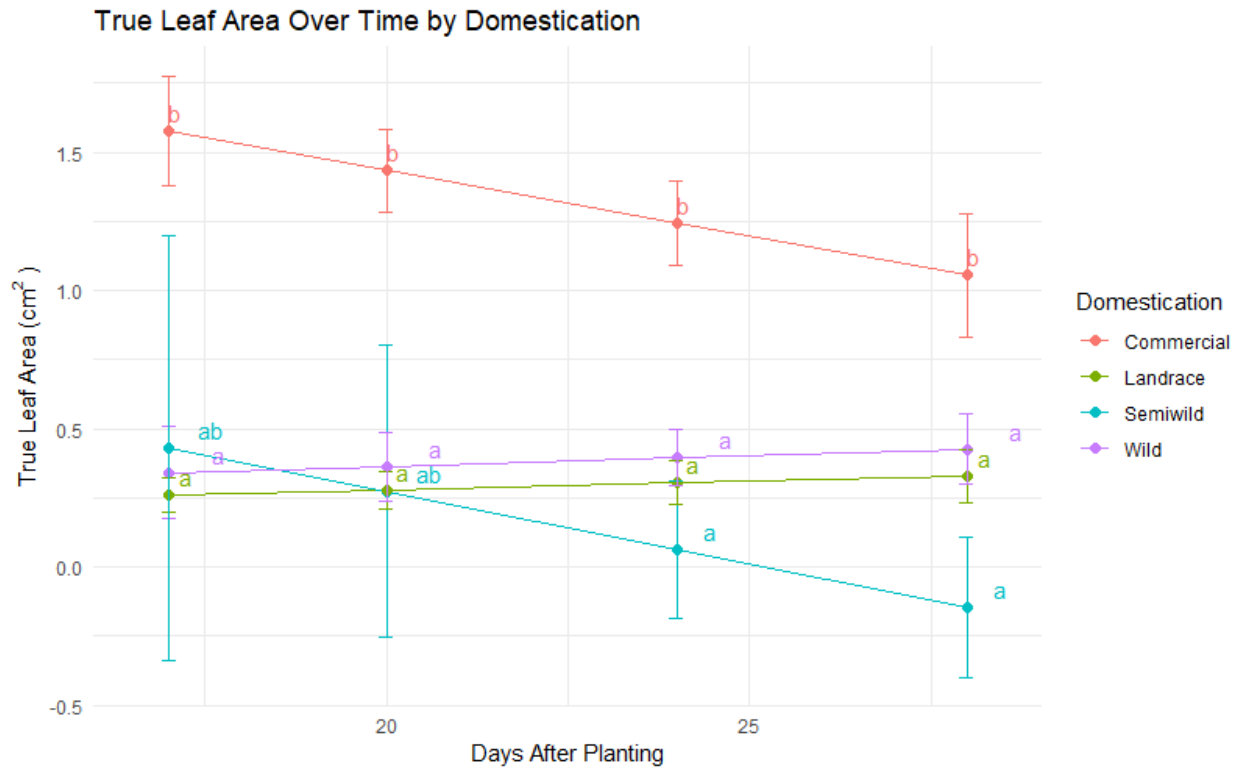
Supplementary Figure 10. LSmeans of height to the first node of *C. annuum* over time by domestication and location. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



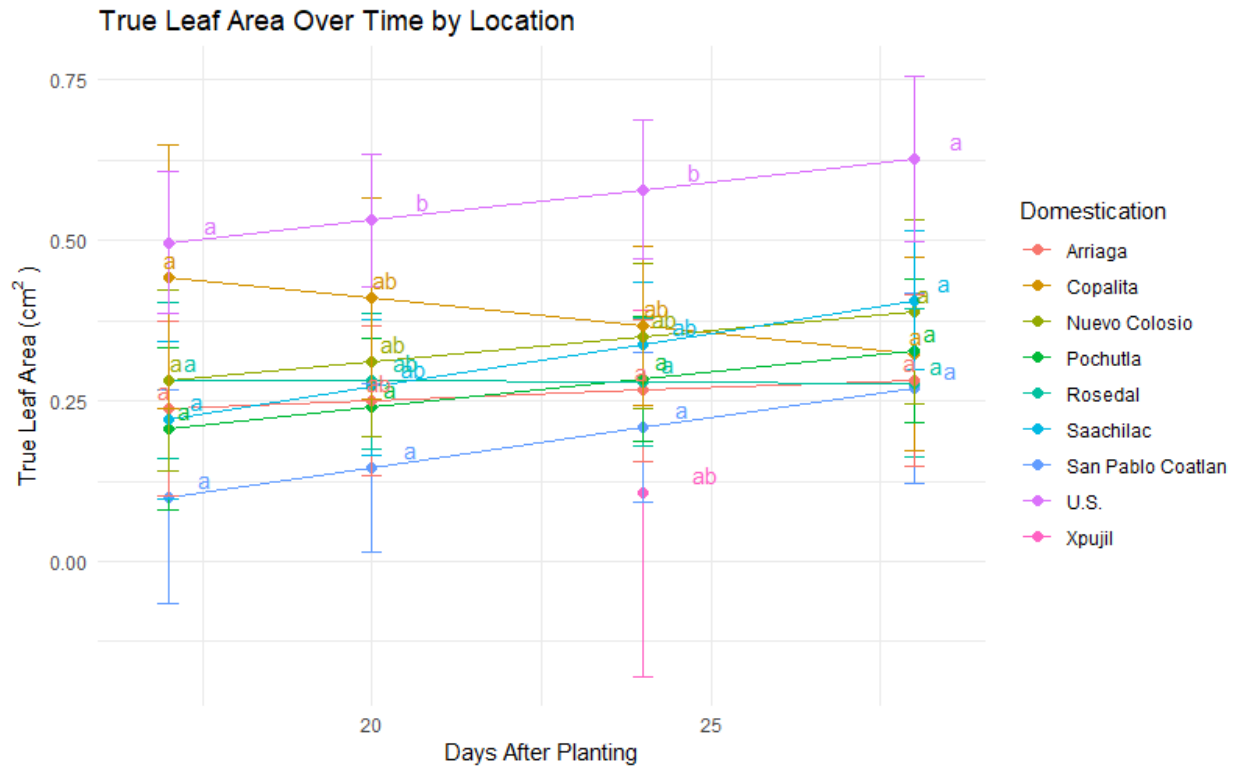
Supplementary Figure 11. LSmeans of true leaf area of *C. annuum* over time by domestication and location. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



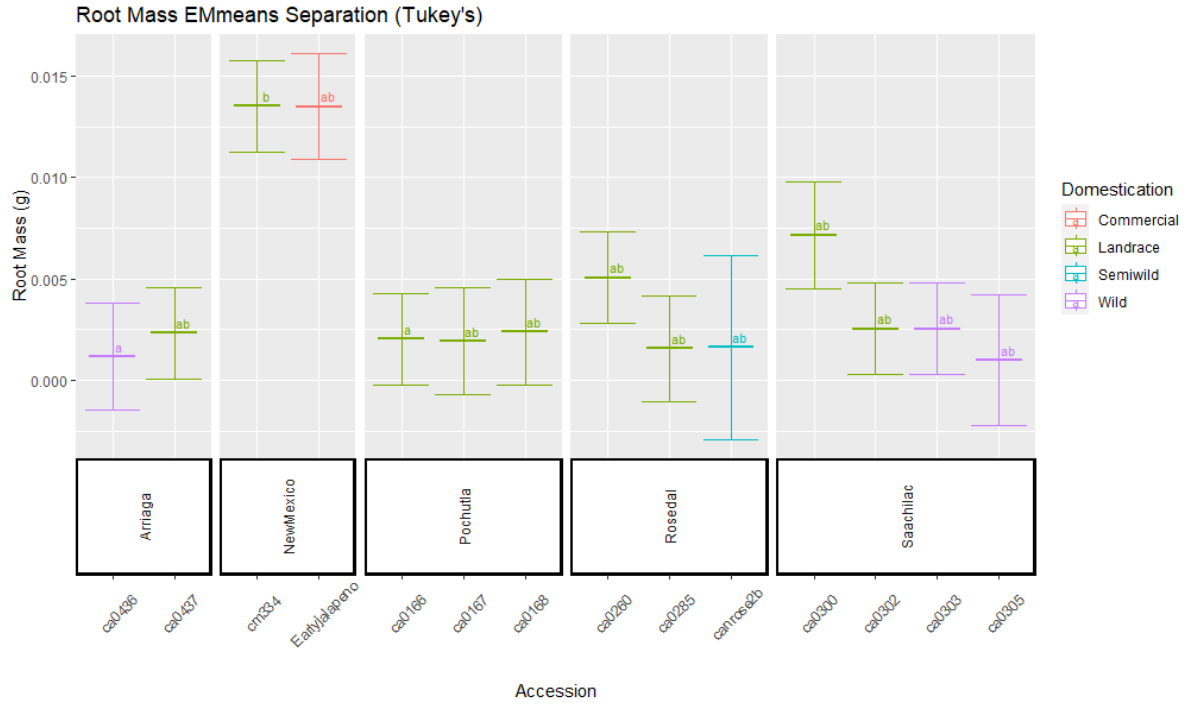
Supplementary Figure 12. LSmeans of total number of true leaves of *C. annuum* over time by domestication and location. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



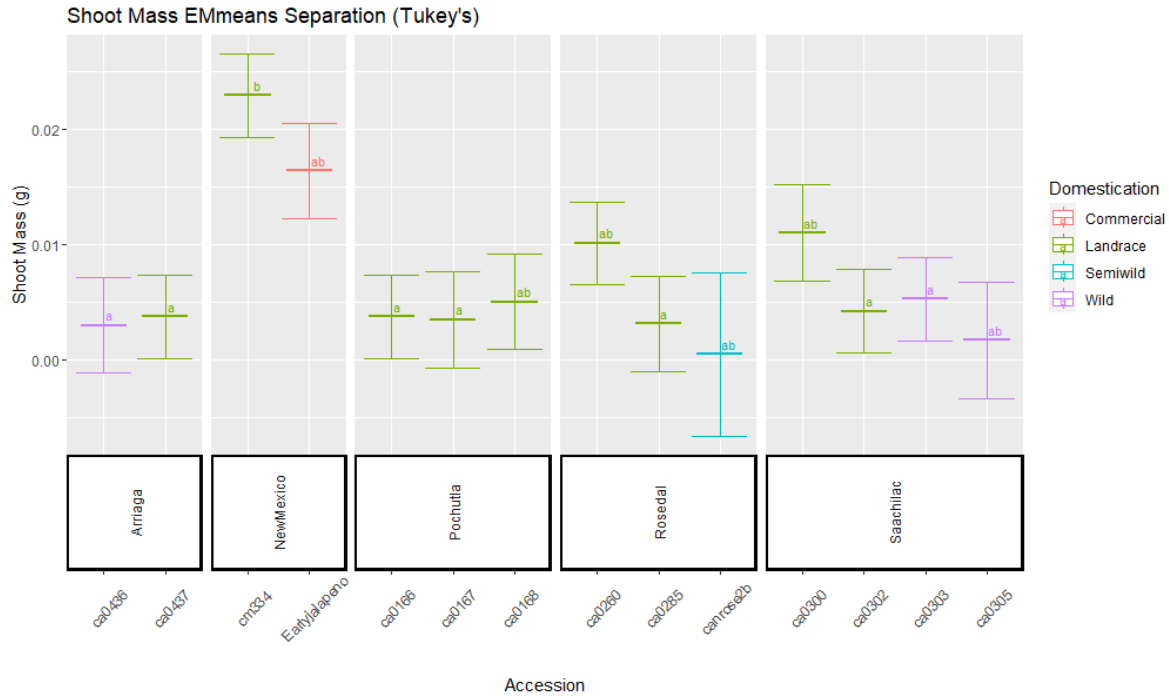
Supplementary Figure 13. Mean separation using the Tukey test of true leaf area at the time of leaf emergence of *C. annuum* over time by domestication. Means with the same letter are not significant. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



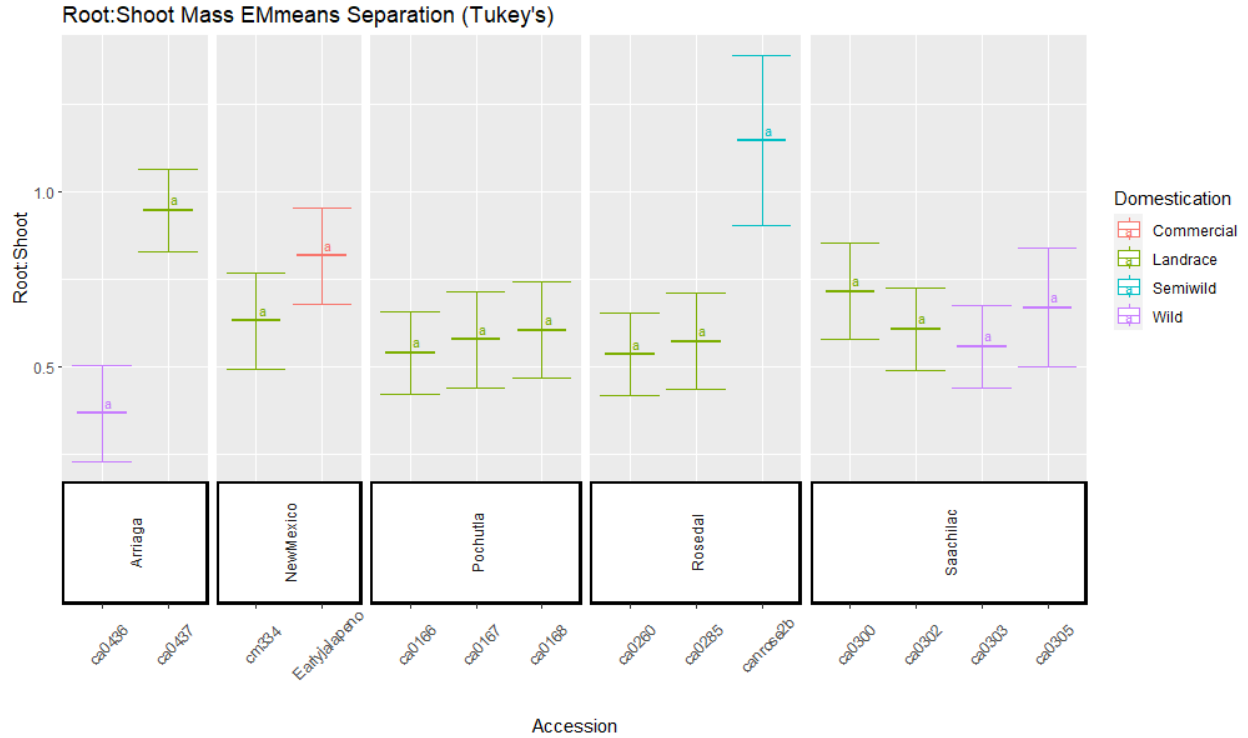
Supplementary Figure 14. Mean separation using the Tukey test of true leaf area at the time of leaf emergence of *C. annuum* over time by location. Means with the same letter are not significant. Repeated measures linear mixed-effects model was fit using “lme4” package prior to mean separation. Points indicate LSmeans and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



Supplementary Figure 15. Mean separation using Tukey’s test of root mass (g) for *C. annuum* across accessions from Oaxaca, Mexico. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



Supplementary Figure 16. Mean separation using Tukey's test of shoot mass (g) for *C. annuum* across accessions from Oaxaca, Mexico. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.



Supplementary Figure 17. Mean separation using Tukey's test of R:S for *C. annuum* across accessions from Oaxaca, Mexico. Means with the same letter are not significant. Separation is denoted by both domestication level and location grouping. Points indicate estimated marginal means and error bars represent standard error. Howlett Greenhouse, The Ohio State University, 2023.