Review: Redefining Domestication: Adaptive Introgression during Crop Expansion

Authors: Garrett M. Janzen[1], Li Wang[1], and Matthew B. Hufford[1,*]

[1]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa, USA

[*]Correspondence: mhufford@iastate.edu (M.B. Hufford)

The process of domestication was once thought to be rapid and geographically constrained, with crops originating from a wild progenitor within one or more defined centers followed by expansion to the modern-day extent of cultivation. *will need citation for this* However, archaeological and genetic evidence are beginning to reveal that, in many cases, domestication has been temporally protracted and more diffuse [1, 2, 3, 4]. This new conception of domestication emphasizes the role of beneficial gene flow (*i.e.*, adaptive introgression) from locally adapted wild relatives during crop expansion after initial domestication.

Adaptive introgression has three components: hybridization between two genomes, backcrossing to one of the parents, and selection on different recombinant genotypes with progressively diminished linkage drag [5, 6]. In domesticated species, adaptive introgression would consist of crop/wild hybrids backcrossing to a crop, retention and increase in frequency of adaptive wild haplotypes in the crop, and removal of undesirable wild background. To date, literature on crop-wild gene flow has focused on the risk of transgene introgression from domesticated crops into wild relatives (for a review, [7]) and on modern plant breeding efforts to introgress desired traits from wild relatives (for a review, [8, 9, 10]). *cite one paper here...the best, most recent and comprehensive review* The history of natural introgression of wild alleles into domesticated crops over evolutionary timescales has received considerably less attention. However, new tools and methods have recently been employed to detect genome-wide patterns of introgression, granting new insights into the prevalence of adaptive introgression in crop histories. Emerging results suggest a need for reevaluating the existing domestication paradigm.

In this review, we will: 1) briefly describe recently developed methods for detecting adaptive introgression and provide a summary of their application for detecting crop-wild introgression, 2) review evidence suggesting wild-to-crop introgression has conferred local adaptation, 3) consider how adaptive introgression alters traditional concepts of domestication, and 4) describe future advances in both basic and applied genetics that can be made through the study of introgression in agroecosystems.

Introgression methods and their application

In this section, I think the overall content is good, but we need to edit to make it more accessible and more explicit about how methods are implemented to detect adaptive introgression

The decreasing cost of genome-wide resequencing and availability of reduced-representation genotyping (*e.g.*, GBS and RAD-Seq), combined with new analytical methods, has facilitated comprehensive study of introgression across a number of species (**Table 1**). High-density marker data can be used with haplotype-based and other methods to assign specific genomic regions to a taxon of origin and identify introgression across taxa [11, 12, 13, 14, 15, 16]. The methods reviewed here do not include those marginally estimating introgression/migration rate as a component of demographic history (*e.g.*, Approximate Bayesian Computation (ABC) [17], diffusion approximations for demographic inference ($\delta a \delta i$) [18], isolation with migration models [19], and a series of methods utilizing the sequentially Markovian coalescent (PSMC, MSMC and SMC++) [20, 21, 22]). Rather, we focus on methods that explicitly identify introgressed genomic segments based on the extent of differentiation, on patterns of nucleotide/haplotype sharing, and phylogenetic relationships.

First, introgressed segments are expected to show low differentiation from their source population. The F_{st} and d_{XY} statistics and their derivates including G_{min} [16] and RND_{min}[15] gauge differentiation. The former two statistics are insensitive to rare migrants and therefore lack power to detect recent introgression, while the latter two overcome this limitation. Additionally, RND_{min} accounts for variable mutation rate, which is detected based on branch length to an outgroup:

$$RND_{min} = \frac{d_{min}}{d_{out}} \tag{1}$$

where d_{min} is the minimum sequence distance between haplotypes in species X and Y and d_{out} equals $(d_{XO} + d_{YO})/2$, the average sequence distance between each species and the outgroup (O).

These statistics have recently been further developed by adding differentiation between both non-admixed (A) and admixed populations (B) and a source population (C) [23]. For example, the $U_{A,B,C(w,x,y)}$ statistic summarizes number of sites where an allele at frequency y in the source population (C) has a frequency higher than x in the admixed population (B) and lower than w in the non-admixed population (C). A similar statistic, $Q95_{A,B,C(w,y)}$, sets a hard cutoff at the 95^{th} percentile of allele frequencies in the admixed population (B) [23]. Further modifications have allowed specification of more than one source population (see details in [23]).

Second, local ancestry deconvolution (also known as chromosome painting) assigns genomic regions to various source populations based on patterns of allele/haplotype sharing [24]. One form of chromosome painting utilizes hidden Markov models to evaluate ancestry across admixed genomes through comparison to reference, non-admixed individuals (*e.g.*, HAPMIX [12]). Another clusters admixed populations with reference samples using a sliding-window approach (*e.g.*, PCAdmix [25] and LAMP [26]). And finally, introgression can be detected through chromosome painting by using a Bayesian model [27] in which deviations from Hardy-Weinberg equilibrium are minimized through creation of genetic groups (*e.g.*, fineSTRUCTURE [13]).

Third, the ABBA-BABA statistic (also known as the D-statistic) and its derivatives are widely applied to introgression detection. These statistics make inferences regarding introgression based on genomic patterns of derived variants that are shared between populations or species. Patterns of allele sharing are interpreted in a phylogenetic context and the method is best suited to detection of introgression at the genome level. Elaborations of the D-statistic capable of localizing introgression to specific genomic regions include $\hat{f}_d$ [11] and the five-taxon D-statistic [14]. The former is quite similar to the D-statistic but uses allele frequencies from each population/species, and the latter detects introgression based on the localized phylogenetic pattern and is capable of determining introgression directionality.

Application of these approaches across a number of plant and animal species suggests introgression can play an adaptive role. For example, introgression from ancient hominins (*e.g.*, Neanderthals and Denisovans) to humans has been detected at loci controlling skin pigmentation, defense against pathogens, and toleration of high altitude (reviewed in [28]); introgression has conferred Müllerian mimicry *I would explain a bit more here...wing coloration loci, protects against predation...* across butterfly species

[29]; introgression has spread insecticide resistance across mosquito species [30], and introgression across *Mimulus* (*i.e.*, monkeyflower) species has resulted in adaptation to pollinator preference and contributed to speciation [31].

Crop adaptation through introgression

Over the last few years, several high-profile publications based on genome-wide data have documented introgression between crops and their wild relatives outside putative domestication centers. Recent empirical studies have revealed that introgression has occurred in many of the world's most important crops (**Table 2**).

1. Maize:

   The relationship between maize (*Zea mays* ssp. *mays*) and the teosinte *Zea mays* ssp. *mexicana* (hereafter referred to as *mexicana*) offers a prime case study of adaptive wild-to-crop introgression. Maize was domesticated from (*Zea mays* ssp. *parviglumis*) approximately 9,000 years ago in the lowlands of the Balsas River Valley in Mexico [32]. From this domestication center, maize spread into the highlands of the Mexican Central Plateau, where it came into sympatry with *mexicana*. Introgression from *mexicana* to maize in the Central Plateau has long been reported based on both morphological [33, 34, 35] and molecular [36, 37, 38, 39, 40] data. However, [41] first localized *mexicana* introgression to chromosomal regions and provided evidence that it was likely adaptive. The authors identified nine genomic regions in several maize populations which consistently showed evidence of *mexicana* introgression based on chromosome painting using both HAPMIX and the linkage model of STRUCTURE (Figure 1). These introgressed segments showed low diversity and overlapped QTL that had previously been found to control anthocyanin content and leaf macrohairs [34], traits known to be adaptive at high elevation. In a growth chamber experiment, the authors demonstrated that maize populations with *mexicana* introgression showed greater plant height (a proxy for fitness) under highland environmental conditions than populations that lacked introgression. Height differences were not detected under lowland conditions.

   Populations of *mexicana* cannot be found outside of Mexico, yet maize has colonized and adapted to high elevation in a number of other regions. A recent study [3] employed the ABBA-BABA and $\hat{f}_d$ statistics to evaluate whether maize with *mexicana* introgression was transferred to other highland regions or whether highland adaptation was obtained *de novo* outside of Mexico. Overall, analyses revealed that maize landraces with *mexicana* introgression were transferred to nearby high elevation regions in Guatemala and the southwestern United States, but more distant high elevation regions (*e.g.,* the Andes) showed no *mexicana* ancestry.

2. Barley:

   Barley (*Hordeum vulgare* subsp. *vulgare*) was domesticated at least twice roughly 10,000 BP: once from the wild subsp. *spontaneum* in the Fertile Crescent and once from subsp. *spontaneum* var. *agriocrithon* in Tibet [42, 43, 44, 45, 46]. *the source morrell2007genetic will be used to update this paragraph once we've worked out for sure that the second domestication of barley was in the Zargos mountains, rather than Tibet. Also, this paper puts the earliest archaeological samples of barley at closer to 8500 calibrated years ago.* Presently, the distribution of subsp *spontaneum* stretches from the eastern Mediterranean through the Middle-East to west-central Asia spanning clines in temperature, precipitation, soil type, and altitude [47]. Barley/*spontaneum* hybrids are known to be fertile and are found spontaneously when wild and domesticated barleys co-occur. In some cases, wild-to-crop introgression has been shown to occur over distances of more than a kilometer [48].

I think here a more focused discussion of potential for adaptive introgression based on the Poets paper is needed The authors of [**?**] used STRUCTURE to look for patterns of introgression from wild relatives in a dataset of 803 landraces, and found a high amount variability in the amount of contribution from wild relatives, as well as its location in the genome, within barley populations. This is indicative of contribution from numerous wild populations. Furthermore, the authors found that wild introgression contribution is generally greatest from geographically-proximate populations, and that introgressed regions might be combined from geographically-separate wild populations. Low linkage disequilibrium and small blocks of identity by state indicate that these introgressed regions are old, perhaps dating back to the beginning of barley domestication. As landraces and nearby wild relatives share similar genomic sequences, the introgressed regions that are exclusive to that landrace are more likely to contain adaptive alleles. Such alleles were not identified specifically, though wild-domesticate breeding experiments have shown that wild barleys have alleles for several important agronomic phenotypes, including powdery mildew resistance, brittleness, flowering time, plant height, lodging, and yield [49, 50**?**].

3. Sunflower:

   The common sunflower (*Helianthus annuus*) shows evidence of domestication in the eastern United States [51, 52] with potential for a second domestication center in Mexico [53]. The pre-Columbian *H. annuus* distribution of cultivated sunflower spanned much of the Great Plains, from what is now north-central Texas to Montana and North Dakota (see figure 1 of [**?**]). Domesticated sunflower has long lived in sympatry with wild relatives such as *H. petiolaris* and *H. bolanderi* and forms stable hybrid populations with these taxa [54, 55, 56]. Wild sunflowers are known to be locally-adapted, and weedy hybrid populations often share these adaptations [57]. However, the most striking example of adaptive introgression within *Helianthus* is that of the cucumberleaf sunflower, *H. debilis* ssp. *cucumerifolius*. *I believe* H. annuus *ssp.* texanus *is a wild sunflower, right? If so, this is more an example of hybrid speciation rather then adaptive wild-to-crop gene flow* Cucumberleaf sunflower is endemic to south-central Texas, and exhibits several adaptations to the region. Introgressive hybridization imparted locally-adapted alleles from *H. debilis* to *H. annuus* via introgressive hybridization [58]. These introgressed hybrids formed a new lineage of sunflower (*H. annuus* ssp. *texanus*, *H. a. texanus* hereafter) which displays *H. debilis*-like traits adaptive to south-central Texas climate and ecology. These adaptive *debilis*-like traits include resistance to herbivorous pests and an increased branching plant architecture, as well as higher overall fitness than *H. annuus* (as measured by higher seed production [**?**]). Although H. annuus and *H. a. texanus* are interfertile, *H. a. texanus* displays persistent phenotypic differences from *H. annuus* [**?**].

   The genome of the common sunflower has been greatly influenced by introgression from wild relatives, due to both natural outcrossing events and concerted breeding efforts in crop improvement. *Helianthus* has several genes for downy mildew resistance, and each imparts resistance to one or more races of *Plasmopara halstedii*, one of the most agronomically important diseases in sunflower cultivation [59]. Some of these downy mildew resistence genes were found in wild relatives (including *H. argophyllus*, *H. tuberosus*, and *H. praecox*) and have been successfully bred into modern *H. annuus* [60]. PlArg, an allele found in wild silverleaf sunflowers (*H. argophyllus*, inbred line Arg1575-2), confers resistance to all known (20 or more) races of downey mildew [61] while others (Pl1-Pl11) are effective for one or more types [62]. Silverleaf sunflower has also been the focus of drought resistance breeding efforts [63] and *Phomopsis* resistance breeding efforts [64]. *H. annuus* shows signs of persistent introgressive hybridization with *H. petiolaris* with evidence of positive selection driving some of the genetic differentiation between the two species [65].

Recent investigations into the history of *Helianthus* introgression have implemented genomic methods. [**?**] analyzed transcriptome sequence variation on cultivated and wild *H. annuus*, *H. petiolaris*, and *H. argophyllus*. Using STRUCTURE, these authors found that introgressions from wild relatives exist on every chromosome in at least one modern line, covering over 10% of the genome. Of particular note is the modern line RHA 274, a modern line which was bred with *H. a. texanus* in the 1970s to restore a branching plant body architecture, which allows the plant to produce pollen for a longer period of time, increasing seed production. RHA 274 has several large introgression from *H. a. texanus*, including one at the site of HaGNAT, the domestication gene associated with branching. These introgressed regions are not found in the non-branching lines Sunrise and VNIIMK8931, further suggesting that the *H. a. texanus* introgressed regions are causative.

4. Asian Rice:

The story of Asian rice (*Oryza sativa*) domestication is still debated. A recent investigation pointed to a single domestication occurring 8,200-13,500 BP in the Yangtze Basin in China from the wild species *Oryza rufipogon* (*rufipogon* hereafter), with later divergence of the two prominent varietal groups, *japonica* and *indica* [66]. [**?**] support this view, noting that present-day patterns of rice variation could be explained by a single domestication event in a population with high standing variation, followed by dispersal into sympatry with locally-adapted wild relatives in diverse environments, genetic admixture, and selection for adaptive alleles. On the other hand, both genetic and archaeobotanical evidence point toward independent domestications of *japonica* and *indica* in the Yangtze Basin and the Indian Ganges plain, respectively [67]. An alternative explanation, as described by [**?**], suggests an initial domestication of *japonica* in China followed by formation of the *indica* group through hybridization of *japonica* and local *rufipogon* populations in Southern and South-eastern Asia. Adaptive introgression may in fact be the signal interpreted as independent domestications.

The high genetic diversity within domesticated rice is likely due to introgression from wild relatives both within the domestication center(s) and in new environments where rice has dispersed [68], and gene flow between domesticated Asian rice and its wild relatives outside of the Yangtze Basin could have imparted locally adaptive traits [**?**]. The wild relatives *rufipogon* and *nivara* both maintain high phenotypic diversity and exhibit locally-adaptive traits. *Rufipogon* reproduction can be either primarily sexual or vegetative, and whereas *rufipogon* is adapted to forested wetland environments, *nivara* is adapted to dryer conditions and has life cycle adaptations to survive grazing pressure. Likewise, domesticated rice varieties display patterns of local adaptation. For examples, two of the domesticated rice deepwater varieties (*rayada* and *ashwina*) are said to be selected for the environment along the Ganges river, the *japonicas* are split into temperate and tropical subgroups, and the *indicas* are best suited for lowland environments. Gene flow between these wild and cultivated rices, though asymmetrical, is exceedingly common, producing nuisance weedy hybrids when the two are grown in sympatry.

To date, research into adaptive introgression in the domestication of rice has been insufficient to detect clearly-supported examples. [69] used a SNP panel and STRUCTURE analysis to

Table 1: List and brief description of recently developed methods and examples of empirical studies employing these methods.

methods	data type	reference
chromosome paiting		
Hapmix	phased haplotype; reference panel	Price et al. 2009
RASPberry	phased haplotype	Wegmann et al. 2011
MultiMix	phased/unphased genotype; reference panel	Churchhouse and Marchini 201
PCAdmix	phased haplotype	Brisbin et al. 2012
LAMP	phased haplotypes; reference panel	Sankararaman et al. 2008
phylogenetic relationship		
ABBA-BABA/D-statistics	biallelic SNP	Durand et al. 2011
fd statistic	biallelic SNP	Martin et al. 2015
five taxon D statistics	biallelic SNP	Pease and Hahn 2015
divergence		
Gmin	biallelic SNP	Geneva et al. 2015
RNDmin	phased haplotype	Rosenzweig et al. 2016
(see .tex file for comment)	biallelic SNP	Racimo et al. 2016
population structure related		
fineStructure	phased haplotype	Lawson et al. 2012
Globetrotter	phased haplotype	Hellenthal et al. 2014

uncover patterns of population structure, admixture, and introgrogression within domesticated rices, and the authors emphasize the importance of similar research that includes wild rice accessions. There are perhaps some practical reasons why research has not yet been devoted to this inquiry. As with many other domesticated crops, gene flow between wild and domesticated rices is highly asymmetric (estimates of wild rice admixture in domesticated rice are less than 5 percent [**?**]). This asymmetry is due in part to the closed floret architecture of the domesticated rice, which hinders outcrossing. However, during early domestication, introgression may have been more prevalent than at present because barriers to crop-to-wild introgression may have been less severe and because the inbreeding reproductive system of rice would not have been as firmly established [**?**]. Furthermore, the contemporary distribution of wild rice does not capture the range and diversity of wild rice during early domestication and range expansion of rice.

Re-evaluating concepts of domestication

A framework in which crops were domesticated from a single population or even a single species is, in several instances, an oversimplification. An history of introgression during diffusion appears to be the rule for crops rather than the exception. Theory suggests that colonizing species will overwhelmingly be recipients of introgression from locally-adapted native species [**?**]. Crops, given their frequent history of diffusion from defined centers of origin, are therefore potential recipients of adaptive introgression.

Crop	Compatible Wild Relatives	Hybrids and/or Hybridization	Evidence of Crop Introgression	Evidence of Adaptiveness	Source
Maize (*Zea mays* subsp. *mays*)	*Z. m.* subsp. *mexicana*, *Z. m.* subsp. *parviglumis*	X	X	X	[70]
Asian Rice (*Oryza sativa*)	*O. rufipogon*	X	X	X	[?]
Barley (*Hordeum vulgare*)	*H. v.* subsp. *spontaneum*	X	X	X	[?]
Sunflower (*Helianthus annuus*)	*H. argophyllus*, *H. bolanderi*, *H. debilis*, *H. petiolaris*	X	X	X	[?]
Cassava (*Manihot esculenta*)	*M. glaziovii*	X	X	X	[?]
Potato (*Solanum tuberosum*)	many	X	X	X	[71]
Tomato (*Solanum lycopersicum*)	*S. pimpinellifolium*	X	X	X	[72]
Olive (*Olea europaea* ssp. *europaea* var. *sativa*)	*O. e.* ssp. *europaea* var. *sylvestris*	X	X		[?]
Soybeans (*Glycine max*)	*G. soja*	X	X		[73]
Common Bean (*Phaseolus vulgaris*)	*P. v.* var. *aborigineus, P. v.* var. *mexicanus* [[not in this source]]	X	X		[74]
Grapes (*Vitis vinifera* subsp. *vinifera*)	*V. v.* subsp. *sylvestris*	X	X		[?]
Sorghum (*Sorghum bicolor* subsp. *bicolor*)	*S. b.* subsp. *arundinaceum, S. b.* subsp. drummondii	X	X		[75]
Wheat (*Tritium monococcum, T. dicoccum, T. aestivum*)	*T. m. boeoticum, T. dioccoides, T. urartu, Aegilops speltoides, A. tauschii*	X	X		[76]
Apple (*Malus domesticus*)	*M. sylvestris*, *M. orientalis*, *M. baccata*, *M. sieversii*	X	X		[?]

With this in mind, certain aspects of crop evolution must be re-evaluated:

* Estimates of the initial domestication bottleneck may be skewed when introgression is not considered. Chromosomal regions experiencing introgression may have an altered effective population size (N_e) relative to non-introgressed regions depending on diversity within the donor taxon. For example, introgression from wild taxa with historically high N_e will lead to underestimates of the strength of the domestication bottleneck. Conversely, if the donor population has a relatively small effective population size, the opposite bias may be imposed upon bottleneck estimations. In most cases, the effective population size of the domesticated crop will be lower than that of the wild progenitor and wild relative populations.

* Estimates of the timing of domestication based on levels of sequence divergence may be affected when introgressed haplotypes are included. The directionality of this effect is likewise dependent on N_e of the donor population.

* Loci under selection during domestication are often identified based on signatures of substantially-reduced nucleotide diversity in the domesticated taxon relative to the wild progenitor and high allele frequency differentiation between these taxa. Introgression may alter these signatures and confound detection of domestication loci.

* When highly-introgressed or hybrid populations are selected for domestication (as in the case of potato and tomato), identification of original progenitor(s) and domestication centers is difficult. Determining whether a crop Crops arising from highly-interfertile compleces of wild relatives

Future studies in crop-wild introgression

Research has so far shown that adaptive crop-wild introgression has played a significant role in the domestication histories of many agronomically-important crops. However, the dynamics of the process in these cases are not yet fully understood. To what extent does the level of introgression across taxa depend on divergence time and/or mutation load between donor and recipient taxa? Can colonizing species and/or hybrid swarms serve as bridges for gene flow between previously allopatric taxa? At what geographic scale does adaptive introgression occur? Is introgression frequently restricted to very local populations, or is it often seen over broad geographic ranges? To what extent does this depend on the slope of environmental gradients such as temperature, precipitation, and elevation? How did the conscious, subconscious, and unconscious decisions of early farmers facilitate or hinder adaptive introgression into their crops during early domestication? How do the practices of contemporary farmers affect the process of adaptive introgression today?

Additional study of introgression in agroecosystems could lead to advances in both basic and applied genetics, and specifically the continued improvement of modern crops. Loci underlying the domesticated phenotype can be more clearly identified by removing the confounding population genetic signal of introgression. These loci are potentially beneficial targets for crop improvement. Furthermore, adaptive introgression that is clearly tied to a specific environment may include beneficial alleles that can be utilized in crop breeding.

Conclusions

The study of crop domestication has been revolutionized by the advent and application of genomic tools. The genomes of crops and their wild relatives tell a story of give-and-take that extends well beyond the initial stages of domestication. Likewise, population genetic theory reinforces the proclivity of wild relatives to provide advantageous, locally-adapted alleles to crops as they disperse beyond their domestication centers into new geographies with new ecological pressures and niches.

References

[1] T. A. Brown, M. K. Jones, W. Powell, and R. G. Allaby, "The complex origins of domesticated crops in the fertile crescent," *Trends in Ecology & Evolution*, vol. 24, no. 2, pp. 103–109, 2009.

[2] R. S. Meyer, J. Y. Choi, M. Sanches, A. Plessis, J. M. Flowers, J. Amas, K. Dorph, A. Barretto, B. Gross, D. Q. Fuller, I. K. Bimpong, M.-N. Ndjiondjop, K. M. Hazzouri, G. B. Gregorio, and M. D. Purugganan, "Domestication history and geographical adaptation inferred from a snp map of african rice," *Nat Genet*, vol. 48, pp. 1083–1088, 09 2016.

[3] L. Wang, T. M. Beissinger, A. Lorant, C. Ross-Ibarra, J. Ross-Ibarra, and M. Hufford, "The interplay of demography and selection during maize domestication and expansion," *bioRxiv*, p. 114579, 2017.

[4] Y. Zhou, M. Massonnet, J. Sanjak, D. Cantu, and B. S. Gaut, "The evolutionary genomics of grape (vitis vinifera ssp. vinifera) domestication," *bioRxiv*, p. 146373, 2017.

[5] N. Barton, "The role of hybridization in evolution," *Molecular Ecology*, vol. 10, no. 3, pp. 551–568, 2001.

[6] C. Feuillet, P. Langridge, and R. Waugh, "Cereal breeding takes a walk on the wild side," *Trends in Genetics*, vol. 24, no. 1, pp. 24 – 32, 2008.

[7] C. N. Stewart, M. D. Halfhill, and S. I. Warwick, "Transgene introgression from genetically modified crops to their wild relatives," *Nature Reviews Genetics*, vol. 4, no. 10, pp. 806–817, 2003.

[8] D. Zamir, "Improving plant breeding with exotic genetic libraries," *Nature reviews genetics*, vol. 2, no. 12, pp. 983–989, 2001.

[9] S. D. Tanksley and S. R. McCouch, “Seed banks and molecular maps: unlocking genetic potential from the wild,” *Science*, vol. 277, no. 5329, pp. 1063–1066, 1997.

[10] R. Hajjar and T. Hodgkin, “The use of wild relatives in crop improvement: a survey of developments over the last 20 years,” *Euphytica*, vol. 156, no. 1-2, pp. 1–13, 2007.

[11] S. H. Martin, J. W. Davey, and C. D. Jiggins, “Evaluating the use of abba–baba statistics to locate introgressed loci,” *Molecular biology and evolution*, vol. 32, no. 1, pp. 244–257, 2015.

[12] A. L. Price, A. Tandon, N. Patterson, K. C. Barnes, N. Rafaels, I. Ruczinski, T. H. Beaty, R. Mathias, D. Reich, and S. Myers, “Sensitive detection of chromosomal segments of distinct ancestry in admixed populations,” *PLoS Genet*, vol. 5, no. 6, pp. 1–18, 2009.

[13] D. J. Lawson, G. Hellenthal, S. Myers, and D. Falush, “Inference of population structure using dense haplotype data,” *PLoS Genet*, vol. 8, no. 1, pp. 1–16, 2012.

[14] J. B. Pease and M. W. Hahn, “Detection and polarization of introgression in a five-taxon phylogeny,” *Systematic biology*, vol. 64, no. 4, pp. 651–662, 2015.

[15] B. K. Rosenzweig, J. B. Pease, N. J. Besansky, and M. W. Hahn, “Powerful methods for detecting introgressed regions from population genomic data,” *Molecular ecology*, 2016.

[16] A. J. Geneva, C. A. Muirhead, S. B. Kingan, and D. Garrigan, “A new method to scan genomes for introgression in a secondary contact model,” *PloS one*, vol. 10, no. 4, p. e0118621, 2015.

[17] M. A. Beaumont, W. Zhang, and D. J. Balding, “Approximate bayesian computation in population genetics,” *Genetics*, vol. 162, no. 4, pp. 2025–2035, 2002.

[18] R. N. Gutenkunst, R. D. Hernandez, S. H. Williamson, and C. D. Bustamante, “Inferring the joint demographic history of multiple populations from multidimensional snp frequency data,” *PLoS Genet*, vol. 5, no. 10, p. e1000695, 2009.

[19] J. Hey and R. Nielsen, “Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of drosophila pseudoobscura and d. persimilis,” *Genetics*, vol. 167, no. 2, pp. 747–760, 2004.

[20] H. Li and R. Durbin, “Inference of human population history from individual whole-genome sequences,” *Nature*, vol. 475, no. 7357, pp. 493–496, 2011.

[21] S. Schiffels and R. Durbin, “Inferring human population size and separation history from multiple genome sequences,” *Nature genetics*, vol. 46, no. 8, pp. 919–925, 2014.

[22] J. Terhorst, J. A. Kamm, and Y. S. Song, “Robust and scalable inference of population history from hundreds of unphased whole genomes,” tech. rep., 2017.

[23] F. Racimo, D. Marnetto, and E. Huerta-Sánchez, “Signatures of archaic adaptive introgression in present-day human populations,” *Molecular Biology and Evolution*, p. msw216, 2016.

[24] J. G. Schraiber and J. M. Akey, “Methods and models for unravelling human evolutionary history,” *Nature Reviews Genetics*, 2015.

[25] A. Brisbin, K. Bryc, J. Byrnes, F. Zakharia, L. Omberg, J. Degenhardt, A. Reynolds, H. Ostrer, J. G. Mezey, and C. D. Bustamante, “Pcadmix: principal components-based assignment of ancestry along each chromosome in individuals with admixed ancestry from two or more populations,” *Human biology*, vol. 84, no. 4, p. 343, 2012.

[26] S. Sankararaman, S. Sridhar, G. Kimmel, and E. Halperin, "Estimating local ancestry in admixed populations," *The American Journal of Human Genetics*, vol. 82, no. 2, pp. 290–303, 2008.

[27] J. K. Pritchard, M. Stephens, and P. Donnelly, "Inference of population structure using multilocus genotype data," *Genetics*, vol. 155, no. 2, pp. 945–959, 2000.

[28] F. Racimo, S. Sankararaman, R. Nielsen, and E. Huerta-Sánchez, "Evidence for archaic adaptive introgression in humans," *Nature Reviews Genetics*, vol. 16, no. 6, pp. 359–371, 2015.

[29] H. G. Consortium *et al.*, "Butterfly genome reveals promiscuous exchange of mimicry adaptations among species," *Nature*, vol. 487, no. 7405, pp. 94–98, 2012.

[30] L. C. Norris, B. J. Main, Y. Lee, T. C. Collier, A. Fofana, A. J. Cornel, and G. C. Lanzaro, "Adaptive introgression in an african malaria mosquito coincident with the increased usage of insecticide-treated bed nets," *Proceedings of the National Academy of Sciences*, vol. 112, no. 3, pp. 815–820, 2015.

[31] S. Stankowski and M. A. Streisfeld, "Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 282, no. 1814, 2015.

[32] Y. Matsuoka, Y. Vigouroux, M. M. Goodman, J. Sanchez, E. Buckler, and J. Doebley, "A single domestication for maize shown by multilocus microsatellite genotyping," *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 6080–6084, 2002.

[33] H. Wilkes, "Hybridization of maize and teosinte, in mexico and guatemala and the improvement of maize," *Economic Botany*, pp. 254–293, 1977.

[34] N. Lauter, C. Gustus, A. Westerbergh, and J. Doebley, "The inheritance and evolution of leaf pigmentation and pubescence in teosinte," *Genetics*, vol. 167, no. 4, pp. 1949–1959, 2004.

[35] J. F. Doebley, "Maize introgression into teosinte-a reappraisal," *Annals of the Missouri Botanical Garden*, pp. 1100–1113, 1984.

[36] Y. Matsuoka, Y. Vigouroux, M. M. Goodman, J. Sanchez, E. Buckler, and J. Doebley, "A single domestication for maize shown by multilocus microsatellite genotyping," *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 6080–6084, 2002.

[37] J. van Heerwaarden, J. Doebley, W. H. Briggs, J. C. Glaubitz, M. M. Goodman, J. d. J. S. Gonzalez, and J. Ross-Ibarra, "Genetic signals of origin, spread, and introgression in a large sample of maize landraces," *Proceedings of the National Academy of Sciences*, vol. 108, no. 3, pp. 1088–1092, 2011.

[38] J. Doebley, M. M. Goodman, and C. W. Stuber, "Patterns of isozyme variation between maize and mexican annual teosinte," *Economic Botany*, vol. 41, no. 2, pp. 234–246, 1987.

[39] M. L. Warburton, G. Wilkes, S. Taba, A. Charcosset, C. Mir, F. Dumas, D. Madur, S. Dreisigacker, C. Bedoya, B. Prasanna, *et al.*, "Gene flow among different teosinte taxa and into the domesticated maize gene pool," *Genetic resources and crop evolution*, vol. 58, no. 8, pp. 1243–1261, 2011.

[40] K. Fukunaga, J. Hill, Y. Vigouroux, Y. Matsuoka, J. Sanchez, K. Liu, E. S. Buckler, and J. Doebley, “Genetic diversity and population structure of teosinte,” *Genetics*, vol. 169, no. 4, pp. 2241–2254, 2005.

[41] M. B. Hufford, P. Lubinksy, T. Pyhäjärvi, M. T. Devengenzo, N. C. Ellstrand, and J. Ross-Ibarra, “The genomic signature of crop-wild introgression in maize,” *Plos genetics*, 2013.

[42] R. Takahashi, “The origin and evolution of cultivated barley,” *Advances in genetics*, vol. 7, pp. 227–266, 1955.

[43] A. Badr, R. Sch, H. El Rabey, S. Effgen, H. Ibrahim, C. Pozzi, W. Rohde, F. Salamini, *et al.*, “On the origin and domestication history of barley (hordeum vulgare),” *Molecular Biology and Evolution*, vol. 17, no. 4, pp. 499–510, 2000.

[44] H.-I. Oka, *Origin of cultivated rice*. Elsevier, 2012.

[45] P. Azhaguvel and T. Komatsuda, “A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley,” *Annals of botany*, vol. 100, no. 5, pp. 1009–1015, 2007.

[46] G. Haberer and K. F. Mayer, “Barley: From brittle to stable harvest,” *Cell*, vol. 162, no. 3, pp. 469–471, 2015.

[47] E. Nevo and G. Chen, “Drought and salt tolerances in wild relatives for wheat and barley improvement,” *Plant, cell & environment*, vol. 33, no. 4, pp. 670–685, 2010.

[48] G. Hillman, R. Hedges, A. Moore, S. Colledge, and P. Pettitt, “New evidence of lateglacial cereal cultivation at abu hureyra on the euphrates,” *The Holocene*, vol. 11, no. 4, pp. 383–393, 2001.

[49] M. Von Korff, H. Wang, J. Léon, and K. Pillen, “Ab-qtl analysis in spring barley: Ii. detection of favourable exotic alleles for agronomic traits introgressed from wild barley (h. vulgare ssp. spontaneum),” *Theoretical and Applied Genetics*, vol. 112, no. 7, pp. 1221–1231, 2006.

[50] L. L. Handley, E. Nevo, J. A. Raven, R. MartInez-Carrasco, C. M. Scrimgeour, H. Pakniyat, and B. P. Forster, “Chromosome 4 controls potential water use efficiency (δ13c) in barley,” *Journal of Experimental Botany*, vol. 45, no. 11, pp. 1661–1663, 1994.

[51] A. V. Harter, K. A. Gardner, D. Falush, D. L. Lentz, R. A. Bye, and L. H. Rieseberg, “Origin of extant domesticated sunflowers in eastern north america,” *Nature*, vol. 430, no. 6996, pp. 201–205, 2004.

[52] D. M. Wills and J. M. Burke, “Chloroplast dna variation confirms a single origin of domesticated sunflower (helianthus annuus l.),” *Journal of Heredity*, vol. 97, no. 4, pp. 403–408, 2006.

[53] D. L. Lentz, M. D. Pohl, J. L. Alvarado, S. Tarighat, and R. Bye, “Sunflower (helianthus annuus l.) as a pre-columbian domesticate in mexico,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 17, pp. 6232–6237, 2008.

[54] A. Schwarzbach and L. Rieseberg, “Likely multiple origins of a diploid hybrid sunflower species,” *Molecular Ecology*, vol. 11, no. 9, pp. 1703–1715, 2002.

[55] L. H. Rieseberg, D. E. Soltis, and J. D. Palmer, “A molecular reexamination of introgression between helianthus annuus and h bolanderi (compositae),” *Evolution*, pp. 227–238, 1988.

[56] M. E. Welch and L. H. Rieseberg, “Patterns of genetic variation suggest a single, ancient origin for the diploid hybrid species helianthus paradoxus,” *Evolution*, vol. 56, no. 11, pp. 2126–2137, 2002.

[57] N. C. Kane and L. H. Rieseberg, “Genetics and evolution of weedy helianthus annuus populations: adaptation of an agricultural weed,” *Molecular Ecology*, vol. 17, no. 1, 2008.

[58] C. B. Heiser Jr, “Hybridization in the annual sunflowers: Helianthus annuus$\times$ h. debilis var. cucumerifolius,” *Evolution*, pp. 42–51, 1951.

[59] Y. Cohen and W. Sackston, “Factors affecting infection of sunflowers by plasmopara halstedii,” *Canadian Journal of Botany*, vol. 51, no. 1, pp. 15–22, 1973.

[60] J. Miller and T. Gulya, “Inheritance of resistance to race 4 of downy mildew derived from interspecific crosses in sunflower,” *Crop science*, vol. 31, no. 1, pp. 40–43, 1991.

[61] C. Dussle, V. Hahn, S. Knapp, and E. Bauer, “Pl arg from helianthus argophyllus is unlinked to other known downy mildew resistance genes in sunflower,” *Theoretical and applied genetics*, vol. 109, no. 5, pp. 1083–1086, 2004.

[62] M. Rahim, C. Jan, and T. Gulya, “Inheritance of resistance to sunflower downy mildew races 1, 2 and 3 in cultivated sunflower,” *Plant Breeding*, vol. 121, no. 1, pp. 57–60, 2002.

[63] F. Sauca *et al.*, “Introgression of drought-resistance gene (s) from helianthus argophyllus to helianthus annuus specie, using embryo rescue techniques.,” *Romanian Agricultural Research*, no. 27, pp. 47–51, 2010.

[64] G. Besnard, Y. Griveau, M. Quillet, H. Serieys, P. Lambert, D. Vares, and A. Bervillé, “Specifying the introgressed regions from h. argophyllus in cultivated sunflower (helianthus annuus l.) to mark phomopsis resistance genes,” *Theoretical and applied genetics*, vol. 94, no. 1, pp. 131–138, 1997.

[65] Y. Yatabe, N. C. Kane, C. Scotti-Saintagne, and L. H. Rieseberg, “Rampant gene exchange across a strong reproductive barrier between the annual sunflowers, helianthus annuus and h. petiolaris,” *Genetics*, vol. 175, no. 4, pp. 1883–1893, 2007.

[66] J. Molina, M. Sikora, N. Garud, J. M. Flowers, S. Rubinstein, A. Reynolds, P. Huang, S. Jackson, B. A. Schaal, C. D. Bustamante, *et al.*, “Molecular evidence for a single evolutionary origin of domesticated rice,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 20, pp. 8351–8356, 2011.

[67] D. Q. Fuller, Y.-I. Sato, C. Castillo, L. Qin, A. R. Weisskopf, E. J. Kingwell-Banham, J. Song, S.-M. Ahn, and J. Van Etten, “Consilience of genetics and archaeobotany in the entangled history of rice,” *Archaeological and Anthropological Sciences*, vol. 2, no. 2, pp. 115–131, 2010.

[68] G. Second, “Origin of the genic diversity of cultivated rice (oryza spp.): study of the polymorphism scored at 40 isozyme loci.,” , vol. 57, no. 1, pp. 25–57, 1982.

[69] K. Zhao, M. Wright, J. Kimball, G. Eizenga, A. McClung, M. Kovach, W. Tyagi, M. L. Ali, C.-W. Tung, A. Reynolds, *et al.*, "Genomic diversity and introgression in o. sativa reveal the impact of domestication and breeding on the rice genome," *PloS one*, vol. 5, no. 5, p. e10780, 2010.

[70] M. B. Hufford, P. Lubinksy, T. Pyhäjärvi, M. T. Devengenzo, N. C. Ellstrand, and J. Ross-Ibarra, "The genomic signature of crop-wild introgression in maize," 2013.

[71] T. Johns and S. L. Keen, "Ongoing evolution of the potato on the altiplano of western bolivia," *Economic Botany*, vol. 40, no. 4, pp. 409–424, 1986.

[72] C. M. Rick, "The role of natural hybridization in the derivation of cultivated tomatoes of western south america," *Economic Botany*, vol. 12, no. 4, pp. 346–367, 1958.

[73] H.-M. Lam, X. Xu, X. Liu, W. Chen, G. Yang, F.-L. Wong, M.-W. Li, W. He, N. Qin, B. Wang, *et al.*, "Resequencing of 31 wild and cultivated soybean genomes identifies patterns of genetic diversity and selection," *Nature genetics*, vol. 42, no. 12, pp. 1053–1059, 2010.

[74] R. Papa and P. Gepts, "Asymmetry of gene flow and differential geographical structure of molecular diversity in wild and domesticated common bean (phaseolus vulgaris l.) from mesoamerica," *Theoretical and Applied Genetics*, vol. 106, no. 2, pp. 239–250, 2003.

[75] P. Aldrich, J. Doebley, K. Schertz, and A. Stec, "Patterns of allozyme variation in cultivated and wild sorghum bicolor," *Theoretical and Applied Genetics*, vol. 85, no. 4, pp. 451–460, 1992.

[76] D. Zohary, J. R. Harlan, and A. Vardi, "The wild diploid progenitors of wheat and their breeding value," *Euphytica*, vol. 18, no. 1, pp. 58–65, 1969.

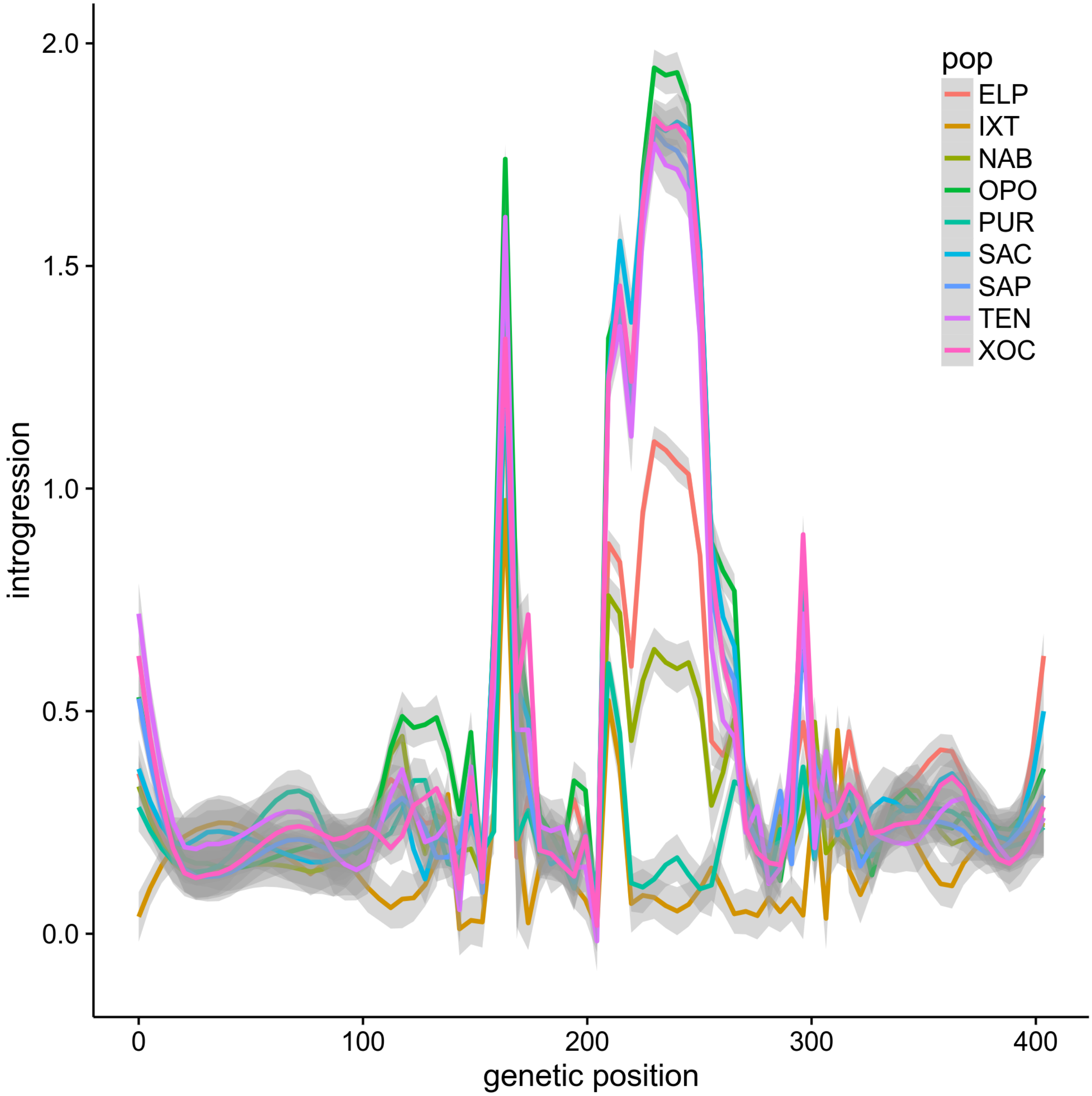

Figure 1: Li's caption here.

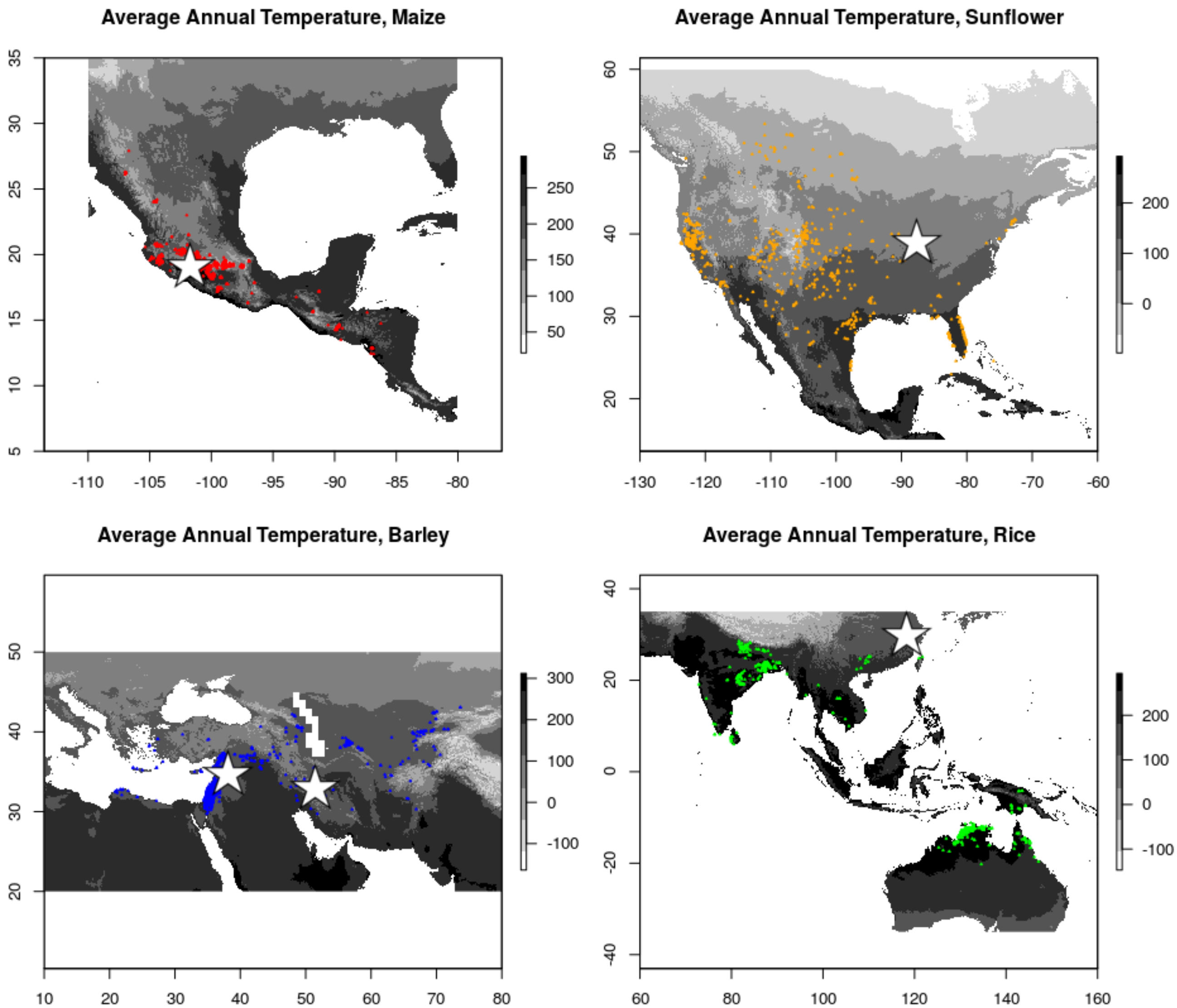

Figure 2: Map of the natural ranges of wild relatives of four domesticated crops, overlayed with average annual temperature.

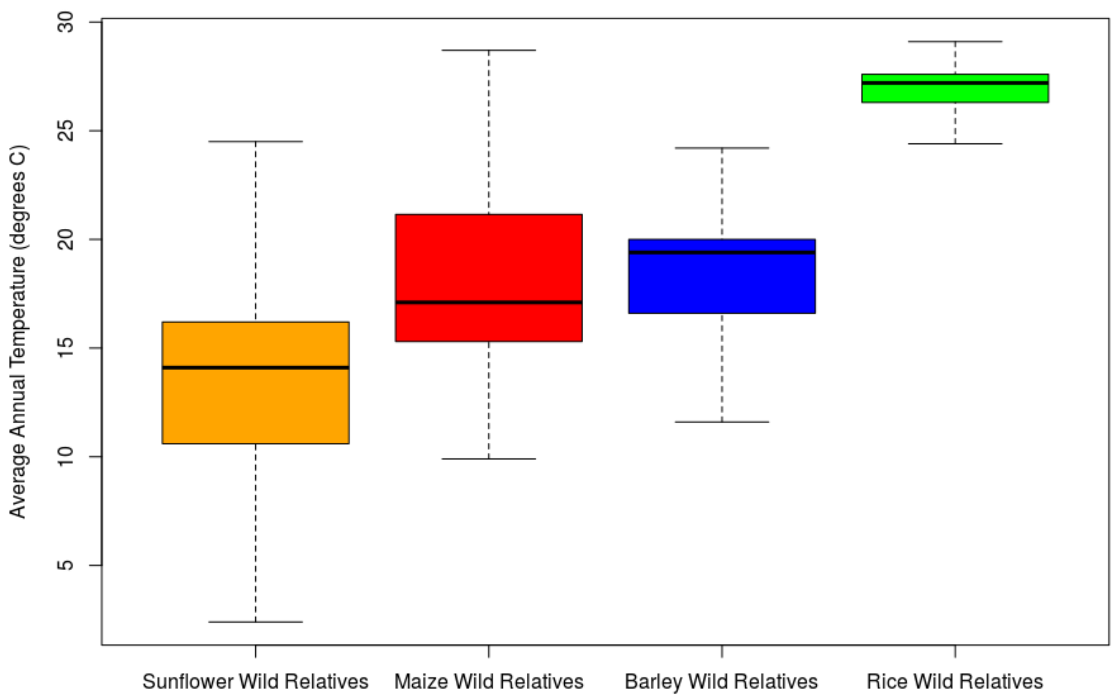

Figure 3: The distribution of average annual temperature experienced in the geographic home ranges of wild relatives interfertile with four crops