PNAS RESEARCH ARTICLE | PLANT BIOLOGY

An adaptive teosinte *mexicana* introgression modulates phosphatidylcholine levels and is associated with maize flowering time

Allison C. Barnes[a,1], Fausto Rodríguez-Zapata[a,b,1], Karla A. Juárez-Núñez[b,1], Daniel J. Gates[c], Garrett M. Janzen[d,e], Andi Kur[a], Li Wang[d], Sarah E. Jensen[f], Juan M. Estévez-Palmas[b], Taylor M. Crow[e], Heli S. Kavi[a], Hannah D. Pil[a], Ruthie L. Stokes[a], Kevan T. Knizner[h], Maria R. Aguilar-Rangel[b], Edgar Demesa-Arévalo[i], Tara Skopelitis[i], Sergio Pérez-Limón[b], Whitney L. Stutts[a,j], Peter Thompson[a,j], Yu-Chun Chiu[j], David Jackson[i], David C. Muddiman[h,j], Oliver Fiehn[k], Daniel Runcie[g], Edward S. Buckler[f], Jeffrey Ross-Ibarra[c], Matthew B. Hufford[d], Ruairidh J. H. Sawers[b,l], and Rubén Rellán-Álvarez[a,b,2]

Edited by Detlef Weigel, Max-Planck-Institut fur Biologie Tübingen, Tübingen, Germany; received February 9, 2021; accepted April 8, 2022

Native Americans domesticated maize (*Zea mays* ssp. *mays*) from lowland teosinte *parviglumis* (*Zea mays* ssp. *parviglumis)* in the warm Mexican southwest and brought it to the highlands of Mexico and South America where it was exposed to lower temperatures that imposed strong selection on flowering time. Phospholipids are important metabolites in plant responses to low-temperature and phosphorus availability and have been suggested to influence flowering time. Here, we combined linkage mapping with genome scans to identify *High PhosphatidylCholine 1* (*HPC1*), a gene that encodes a phospholipase A1 enzyme, as a major driver of phospholipid variation in highland maize. Common garden experiments demonstrated strong genotype-by-environment interactions associated with variation at *HPC1,* with the highland *HPC1* allele leading to higher fitness in highlands, possibly by hastening flowering. The highland maize *HPC1* variant resulted in impaired function of the encoded protein due to a polymorphism in a highly conserved sequence. A meta-analysis across HPC1 orthologs indicated a strong association between the identity of the amino acid at this position and optimal growth in prokaryotes. Mutagenesis of *HPC1* via genome editing validated its role in regulating phospholipid metabolism. Finally, we showed that the highland *HPC1* allele entered cultivated maize by introgression from the wild highland teosinte *Zea mays* ssp. *mexicana* and has been maintained in maize breeding lines from the Northern United States, Canada, and Europe. Thus, *HPC1* introgressed from teosinte *mexicana* underlies a large metabolic QTL that modulates phosphatidylcholine levels and has an adaptive effect at least in part via induction of early flowering time.

phospholipid metabolism | maize genetics | highland adaptation | flowering time | selection

Significance

Despite more than a century of genetic research, our understanding of the genetic basis of the astounding capacity of maize to adapt to new environments is in its infancy. Recent work in many crops has pointed to the potentially important role for introgression in underpinning adaptation, but clear examples of adaptive loci arising via introgression are lacking. Here, we elucidate the evolutionary history of a major metabolic quantitative trait locus (QTL) that we mapped down to a single gene, *HPC1*. Alterations in highland *HPC1* are the result of a teosinte *mexicana* introgression in maize, leading to high phosphatidylcholine levels and improving fitness by accelerating flowering.

Elevation gradients are associated with changes in environmental factors that impose substantial physiological constraints on an organism. Adaptation to highland environments is achieved via the selection of genetic variants that improve their ability to withstand lower oxygen availability (1, 2), increased ultraviolet (UV) radiation (3), and lower temperatures (4). In particular, cold temperatures reduce thermal time accumulation, measured in growing degree days (GDDs) (5), and select for accelerated development and maturity as a compensatory mechanism (6). Following domestication from teosinte *parviglumis* (*Zea mays* ssp. *parviglumis)* (7) in the lowland, subtropical environment of the Balsas River basin (Guerrero, Mexico), cultivated maize (*Zea mays* ssp. *mays*) expanded throughout Mexico and reached the highland valleys of central Mexico around 6,500 y ago (8).

In Mexico, highland adaptation of maize was aided by substantial adaptive introgression from a second teosinte subspecies, teosinte *mexicana* (*Zea mays* ssp. *mexicana)*, that had already adapted to the highlands of Mexico thousands of years after its divergence from teosinte *parviglumis* (9, 10). Adaptation to low temperature and soils with low phosphorus content in highland environments drove *mexicana* genetic divergence from the lowland *parviglumis* (11). Phenotypically, the most evident signs of *mexicana* introgression into maize are the high levels of stem pigmentation and pubescence (12) that are thought to protect against high UV radiation and low temperatures. The ability to withstand low temperatures and efficiently photosynthesize during the early stages of seedling development are key factors in maize highland adaptation (13). Indeed, recent transcriptome deep sequencing (RNA-seq) analysis showed that the inversion *Inv4m*, introgressed from *mexicana*, strongly affects the expression of genes involved in chloroplast physiology and photosynthesis (14). Given the slow accumulation of GDDs in typical highland environments, selection has favored shorter generation times in highland-adapted maize (15).

Author contributions: A.C.B., F.R.-Z., K.A.J.-N., R.J.H.S., and R.R.-A. designed research; A.C.B., F.R.-Z., K.A.J.-N., J.M.E.-P., T.M.C., H.S.K., H.D.P., R.L.S., K.T.K., M.R.A.-R., S.P.-L., W.L.S., P.T., Y.-C.C., R.J.H.S., and R.R.-A. performed research; S.E.J., E.D.-A., T.S., D.J., D.C.M., O.F., and E.S.B. contributed new reagents/analytic tools; A.C.B., F.R.-Z., K.A.J.-N., D.J.G., G.M.J., A.K., L.W., S.E.J., J.M.E.-P., T.M.C., H.S.K., H.D.P., R.L.S., K.T.K., W.L.S., P.T., Y.-C.C., D.C.M., D.R., E.S.B., J.R.-I., M.B.H., R.J.H.S., and R.R.-A. analyzed data; and A.C.B., F.R.-Z., K.A.J.-N., A.K., and R.R.-A. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]A.C.B., F.R.-Z. and K.A.J.-N., contributed equally to this work.

[2]To whom correspondence may be addressed. Email: rrellan@ncsu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2100036119/-/DCSupplemental.

Published June 30, 2022.

By the time maize reached the Mexican highlands, its range had already expanded far to the South, including the colonization of highland environments in the Andes (16, 17). Andean maize adaptation occurred largely independently of *mexicana* introgression (18, 19), and there is no known wild teosinte relative in South America. These multiple events of maize adaptation to highland environments make maize a good system to study the evolutionary and physiological mechanisms of convergent adaptation (18, 19).

In comparison to its southward expansion, the northward migration of maize into the modern-day United States, where summer daylengths are longer, occurred at a much slower pace (20, 21) due to delayed flowering of photoperiod-sensitive tropical maize lines (22). A host of evidence suggests that maize cultivation in northern latitudes was enabled by the selection of allelic variants that led to a reduction in photoperiod sensitivity to allow flowering under longer photoperiods (22–28). Some of the early-flowering alleles that conferred an adaptive advantage in highland environments are the result of *mexicana* introgressions into highland maize (24). Maize first entered into the United States via the Mexican highlands (20), and these early-flowering alleles show further evidence of selection in northern latitudes (24), consistent with a likely role for *mexicana* introgression(s) in maize adaptation to shorter daylength. When introduced into Northern Europe, photoperiod-insensitive maize from the Northern United States and Canada thrived as it was already preadapted to northern latitudes and lower temperatures (29). The genetic, physiological, and phenotypic basis of these adaptations, however, is quite limited.

Plant phospholipids, as well as other glycerolipids such as sulfolipids, galactolipids, and nonpolar lipids such as triacylglycerols, are involved in plant responses to low temperatures. Phospholipid levels are increased in plants exposed to low temperatures (30) and levels of unsaturated fatty acids in glycerolipids are reduced (31, 32), which may help maintain the fluidity of cell membranes. Under stressful conditions, the proportions of differently shaped lipids are modulated to maintain membrane flexibility while preventing membrane leakage. For instance, phosphatidylcholines (PCs) are rectangular polar lipids that are well suited for the formation of bilayer membranes because the size of their glycerol backbone, choline headgroup, and fatty acid tails are similar. By contrast, lyso-phosphatidylcholine (LPC) is a triangular PC with a single acyl group that cannot form a bilayer because its headgroup is much larger than its fatty acid (33). LPCs do allow for some membrane movement, but high LPC concentrations act as a detergent (34) and can facilitate cell leakage and damage at low temperatures, effects that would be prevented by higher PC levels. In cold-adapted maize temperate lines and *Tripsacum* species (a distant maize relative), genes involved in phospholipid biosynthesis show accelerated rates of protein sequence evolution, further supporting an important role for phospholipid metabolism across several species during cold adaptation (35). Finally, multiple phospholipids can bind to *Arabidopsis* (*Arabidopsis thaliana*) FLOWERING LOCUS T (FT) and accelerate flowering. Recent work has shown that phosphatidylglycerol binds and sequesters FT in companion cells in low temperatures, while higher temperatures release it to the shoot apical meristem (36). There, it interacts with certain species of PC, the most abundant phospholipid in plant cells (37), through an unknown mechanism (38). Consistent with this observation, glycerolipid levels in maize have predictive power for flowering time (39).

Here, we identified an adaptive teosinte *mexicana* introgression that alters highland maize phospholipid metabolism and leads to early flowering. Using genome scans and linkage mapping, we identified *High PhosphatidylCholine1* (*HPC1*), a gene encoding a phospholipase A1, as a driver of high PC levels in highland maize. Data from thousands of genotyped landrace test crosses grown in common garden experiments at different elevations in Mexico showed a strong genotype-by-environment effect at the *HPC1* locus, where the highland allele leads to higher fitness in the highlands and reduced fitness at lower elevations. Furthermore, we determined that the highland *HPC1* allele, which was introgressed from teosinte *mexicana*, was carried northward and is now present in maize cultivars grown in the Northern United States and European Flint lines. These results suggest that the *HPC1* highland allele has a beneficial effect in cold, high-latitude environments, where early flowering is advantageous.

2. Results

A. Highland Mesoamerican Maize Shows High PC/LPC Ratios and Selection of Highly Unsaturated PCs. As lipids play an important role in adaptation to adverse environments, we quantified the glycerolipid levels of 120 highland and lowland landraces from Mesoamerica and South America (Fig. 1 *A* and *B* and Dataset S1).* This diversity panel, hereafter referred to as the HiLo diversity panel, was grown in highland (2,650 m above sea level [masl]) and lowland (20 masl) common garden experiments in Mexico. We determined that Mesoamerican highland landraces have high PC/LPC ratios, particularly when grown in the highlands (Fig. 1 *A* and *B*).

The differences observed in phospholipid levels between highland and lowland maize may be the result of adaptive natural selection or random genetic drift during maize colonization of highland environments. To distinguish between these two possibilities, we compared the variance of each phenotype across the population with the genetic variance of neutral markers using a Q_{ST} –F_{ST} comparison (40). We calculated Q_{ST} –F_{ST} using diversity array technology sequencing (DartSeq) genotypes (41) from the same plants phenotyped for glycerolipid levels and calculated the Q_{ST}–F_{ST} values for each glycerolipid species for highland/lowland populations from each continent (*SI Appendix*, Fig. S1). Mean Q_{ST} was greater than mean F_{ST} in both Mesoamerican and South American comparisons. We identified PC and LPC species with higher Q_{ST} values than the neutral F_{ST} in both continents (Fig. 1 *C* and *D*). The species with the highest Q_{ST} value included long-chain PCs with more than one unsaturation such as PC-36:2.

B. Genes Involved in PC/LPC Conversion Show Strong Highland Selection Signals. We selected a set of 597 maize glycerolipid genes from their functional annotations (*Materials and Methods*) to identify selection signals using the population branch excess (PBE) (42) statistic across four highland populations: Southwestern United States (SWUS), Mexican highlands (MH), Guatemalan highlands (GH), and Andes (AN) (Fig. 1*E*) (19). We identified a significant excess of genes that are targets of selection in more than two populations ($P < 3 \times 10^{-5}$, Fig. 1*F*). The most overrepresented intersection of selected glycerolipid genes was between the SWUS, MH, and GH populations ($P = 1 \times 10^{-15}$, Fig. 1*F*), suggesting that genes were specifically selected in these three populations relative to the AN population and/or that there was closer kinship among SWUS, MH, and GH populations than the AN population and thus weaker statistical

*Open-pollinated varieties that have been selected for specific uses and environments by smallholder farmers; they are characterized by similar morphotypes. Landraces typically show as much diversity within individuals from the same landrace group as between groups.

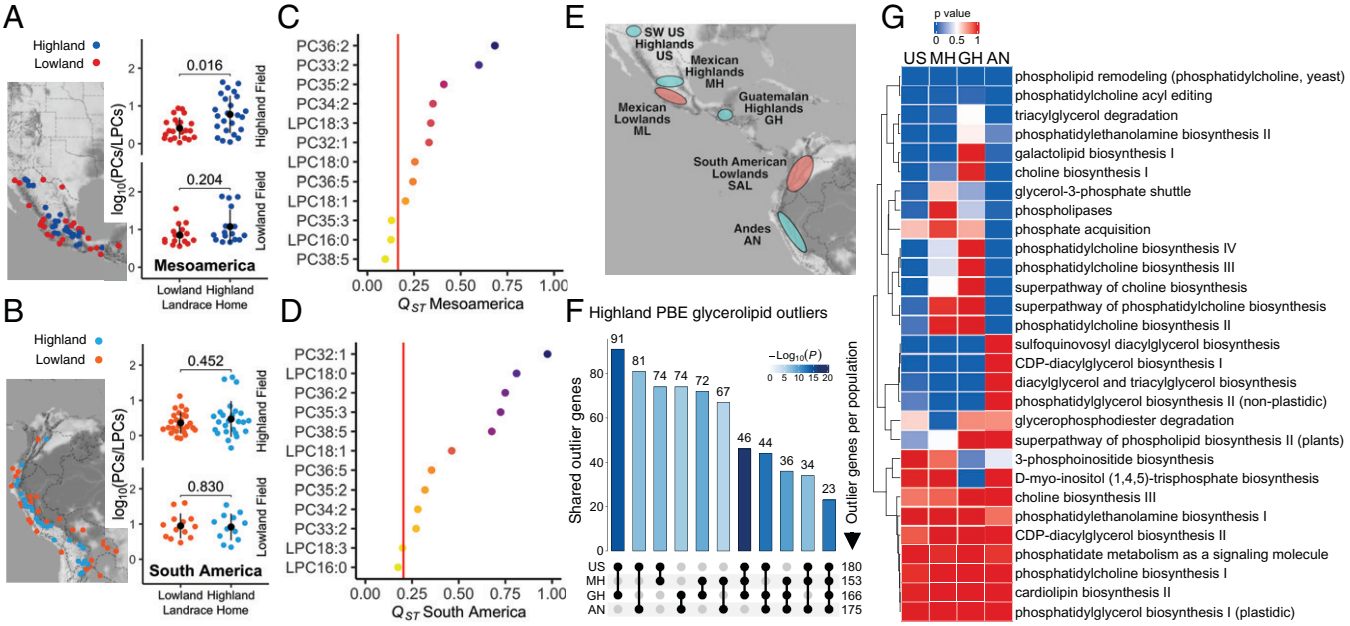

Fig. 1. Phospholipid selection in highland maize. (*A* and *B*) *Left:* Geographical origin of 120 maize accessions from the HiLo diversity panel used in the common garden experiment for glycerolipid quantitation. *Right:* PC/LPC ratio, $\log_{10}$ scaled, for highland and lowland landraces from (*A*) Mesoamerica, $n = 54$. The black circle indicates the mean, and the vertical line is the SD. Significant differences were tested with a false-discovery rate (FDR)-adjusted *t* test; the resulting *P* values are shown. (*C* and *D*) Q_{ST} –F_{ST} analysis of phospholipids between highland and lowland landraces from (*C*) Mesoamerica and (*D*) South America. Red line, neutral F_{ST} + 2 SD. (*E*–*G*) Highland vs. lowland PBE analysis. (*E*) PBE geographical sampling. (*F*) PBE outlier counts for glycerolipid genes in four highland populations. Bar shade indicates Fisher's exact significance test for excess shared outliers. (*G*) Highland selection of glycerolipid-related pathways using PBE. The heatmap shows the probability of randomly sampling gene SNPs with mean PBE greater than the mean for the gene SNPs in each pathway.

independence. From all annotated glycerolipid genes, 23 were consistently PBE outliers across all four highland populations ($P =< 1\times^{-10}$, Fig. 1*F*). We then assigned a PBE value to each of the 30 glycerolipid metabolism pathways (using a 10-kb window around each constituent gene) and compared their average PBE value with a genome-wide random sampling distribution of PBE values within genic regions. From these, we established that "phospholipid remodeling" and "PC acyl editing" exhibit significantly higher PBE values in all four populations, indicating a possible role for phospholipid remodeling in maize highland adaptation (Fig. 1*G*).

We considered two possible explanations for the extent of convergent selection in highland populations (explanations in refs. 19 and 43). Adaptation may be conferred by a small number of genes, thereby imposing a physiological constraint on the sources of adaptation leading to convergence. Alternatively, adaptation may be the result of many genes, but deleterious pleiotropic effects restrict the number of genes that can be targeted by selection, also leading to convergence. Using Yeaman et al.'s (43) C_{hyper} statistic, which quantifies these two modes of convergent adaptation, we determined that the overlap among putative adaptive genes in the four highland populations cannot be explained merely by physiological constraints ($C_{hyper} = 3.96$; *SI Appendix*, Table S1). A certain degree of pleiotropic constraint is therefore likely. Overlap between adaptation candidates was higher for the SWUS, MH, and GH population pairs ($C_{hyper} = 4.79$) than between the Andean and SWUS, MH, and GH pairs ($C_{hyper} = 3.14$).

To further understand selection at the gene level, we used genotyping by sequencing (GBS) data from 2,700 Mexican maize landraces, generated by the SeeD project (15, 44), to run a *pcadapt* (45) analysis to determine how loci might contribute to observed patterns of differentiation along major principal components of genetic variation. The first *pcadapt* principal component separated Mexican landraces based on the elevation of their geographical origin (Fig. 2*B*). Using this first principal component, we identified outlier single-nucleotide polymorphisms (SNPs) across the genome that are significantly associated with genetic variation along elevation and potentially involved in local adaptation (Fig. 2*A*). From the list of ≈600 maize glycerolipid-related genes, 85 contained SNPs that were *pcadapt* outliers for association with the first genetic principal component [top 5% $-log_{10}(P)$] and of which 8 were also PBE outliers for Mexican highlands (Fig. 2*A* and *SI Appendix*, Table S2 and Dataset S2). These eight selection candidates, supported by two different sources of evidence, included two genes coding for putative enzymes whose orthologs are known to directly catalyze PC/LPC interconversion reactions. The first gene, Zm00001d039542, with a $-log_{10}(P)$ of 110.28, encoded a putative phospholipase A1 that we name *High Phosphatidyl-Choline 1* (*HPC1*). The second gene, Zm00001d017584, with a $-log_{10}(P)$ of 99.31, encoded a predicted *Lyso-Phosphatidylcholine Acyl Transferase 1* that we refer to as *ZmLPCAT1* (Fig. 2*C*). Although these two types of enzymes catalyze broadly opposite reactions (degradation vs. biosynthesis of PC) they are unlikely to catalyze strictly reverse reactions in the Lands cycle. A1 phopholipases attack the phosphatidylcholine at the *sn-1* carbon, while LPC acyl transferases usually acylate *sn-2* (46, 47). Instead, plant PLA1 enzymes like *HPC1* are better known for their role in the first step of jasmonic acid biosynthesis (48, 49).

Both genes showed strong changes in elevation-dependent allele frequency (Fig. 2*D*) across Mexican landraces. *HPC1* was not an outlier for branch excess between the Andean and the South American lowland accessions. By contrast, *ZmLPCAT1* was indeed a PBE outlier for all four populations, which may indicate parallel/convergent selection for this gene between Mesoamerican and Andean landraces. Importantly, both *HPC1* and *ZmLPCAT1* are annotated as part of pathways with outlier PBE values in all highland populations for phospholipid remodeling and PC acyl editing (Fig. 1*F* and Dataset S2). Taken together, these two

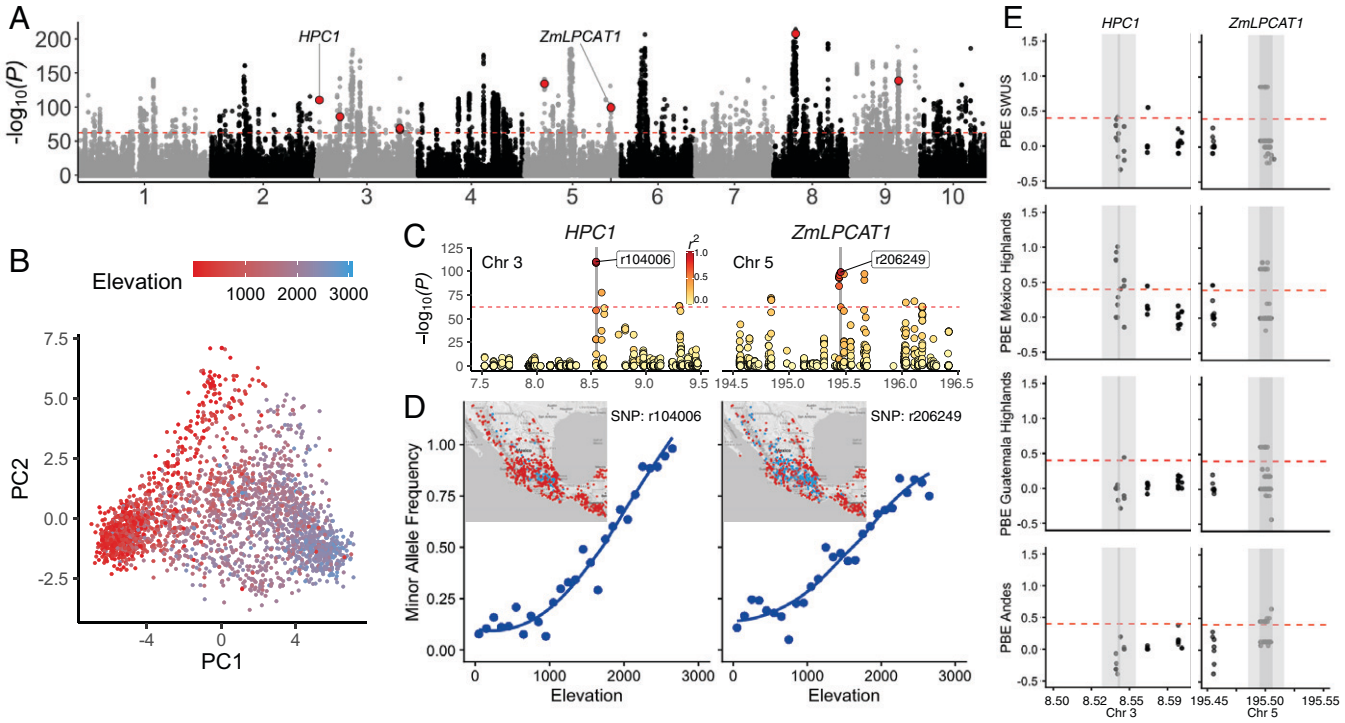

Fig. 2. Evidence of highland selection in genes determining PC/LPC ratios. (*A*) Association with genetic principal component 1 (PC1) from *pcadapt* analysis of Mexican landraces (open-pollinated varieties). Dashed line marks the top 5% $-log_{10}(P)$. Red points show SNPs in glycerolipid metabolism genes that are also PBE outliers for Mexican highlands (*SI Appendix*, Table S2). From these, *HPC1* and *ZmLPCAT1* are the top genes with orthologs catalyzing PC/LPC interconversion. (*B*) Scatter plot of the *pcadapt* first two genetic principal components illustrating that PC1 correlates with elevation. (*C*) Extended region from *A* of the 1-Mb interval around *HPC1* and *ZmLPCAT1*. Linkage disequilibrium (r^2) with the peak SNP for each gene is illustrated by the color scale; both peak SNPs are located in the coding sequence. (*D*) Elevation clines for the peak SNPs from *C*. *Insets* show the geographic distribution of the highland (blue) and lowland (red) alleles. (*F*) Population branch excess of SNPs in *ZmLPCAT1* and *HPC1*. Dark gray, coding sequence; light gray, 10 kb upstream and downstream; dashed line is the threshold for the top 5% PBE value outliers.

independent population genetic approaches show that pathways involved in phospholipid remodeling, including genes controlling the PC/LPC ratio like *HPC1* and *ZmLPCAT1*, show strong selection signals in highland maize. These results indicate that selection on phospholipid metabolite levels (Fig. 1 *B*–*D*) is supported at the gene level by outlier PBE and principal component analysis *pcadapt* values in genes controlling phospholipid biosynthesis and degradation.

C. A Major QTL Explaining PC to LPC Conversion Overlaps *HPC1*.

To further characterize the genetic architecture of phospholipid biosynthesis in highland maize, we developed a recombinant inbred line (RIL) BC1S5 population derived from a cross between the temperate inbred line B73 and the Mexican highland landrace Palomero Toluqueño (PT), using B73 as the recurrent parent (75% B73, 25% PT) (50).

The parental PT accession is a popcorn landrace (*palomero* means popcorn in Spanish) from the Toluca Valley in Mexico (*Mexi5* CIMMYTMA 2233) (Fig. 3*A*). We grew the HiLo diversity panel and the B73 × PT BC1S5 RIL population in the same highland and lowland common gardens and collected samples for glycerolipid analysis. The locally adapted PT landrace displayed higher fitness than B73 in the highland field (Fig. 3*A*), probably due to adaptation to low temperatures in this highland environment. In the Mexican highlands, values of 5 GDDs per day are typical, while 15 to 20 GDDs per day are common in lowland environments.

We detected major quantitative trait loci (QTLs) for the sum of LPC species levels, PC species levels, and the PC/LPC ratio that all mapped to the same locus on chromosome 3, around 8.5 Mb (Fig. 3*B*). We tested for epistatic interactions for LPC levels, PC levels, and the PC/LPC ratio through a combination of R/qtl scantwo and stepwise functions (51). The three QTLs *qLPCs3@8.5*, *qPCs3@8.5*, and *qPCs/LPCs3@8.5* were robust against environmental effects and were detected in both the highland and lowland environments. The additive effect of the PT allele at these QTLs was associated with high levels of PCs, low levels of LPCs, and consequently high PC/LPC ratios, while the B73 allele had the opposite effect (Fig. 3 *C*, *Top*). Individual PC and LPC species QTLs at this locus (*SI Appendix*, Fig. S2) showed the same additive effect for the PT allele as the sum of each class (PCs, LPCs, and PCs/LPCs) of species (Fig. 3 *C*, *Top*). All individual LPC QTLs at the *qLPCs3@8.5* locus corresponded to LPCs that contained at least one double bond in their fatty acid (*SI Appendix*, Fig. S2 and Dataset S3). *qPCs3@8.5* was driven mainly by PC species with more than two fatty acid double bonds, such as PC 36:5 (Fig. 3*D* and *SI Appendix*, Fig. S2 *A*–*D* and File 3). We then sought to identify candidate genes underlying the QTLs on chromosome 3. The PC/LPC ratio QTL had the highest significance, with a logarithm of the odds (LOD) of 24.5, and explained the most phenotypic variance (87%). The underlying QTL interval contained 72 genes within its 1.5-LOD drop CI (7.9 to 10 Mb). We hypothesized that the metabolic phenotypes we observed might be due to a gene involved in PC-LPC conversion. The maize genome encodes 75 putative phospholipases (*SI Appendix*, Fig. S3*A*), of which half are predicted to be phospholipase A1 type (PLA1) (*SI Appendix*, Fig. S3*B*). Notably, *HPC1* mapped within the interval (position on chromosome 3: 8,542,107 to 8,544,078 bp), making it a high-confidence candidate causal gene (Fig. 3*B*). HPC1 was predicted to have phospholipase A1-Igamma1 activity and can be classified as a PC-hydrolyzing PLA1 class I

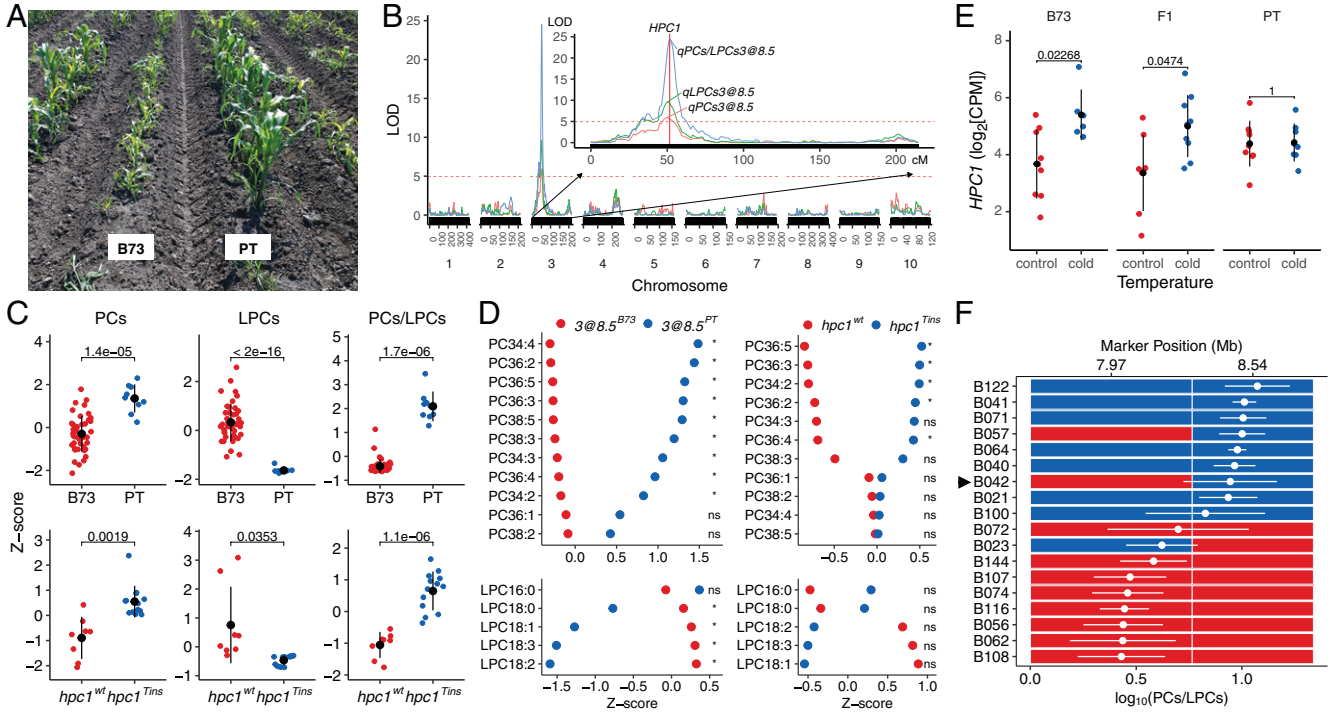

Fig. 3. *HPC1* defines a major QTL explaining PC/LPC conversion. (*A*) PT and B73 plants growing in the highland Metepec field. Picture was taken 64 days after planting. (*B*) QTL analysis identified overlapping major QTLs around 8.5 Mb on chromosome 3 for PC and LPC levels and PC/LPC ratio, using data collected from plants grown in highland and lowland fields. The QTL peaks coincide with the physical location of *HPC1*. (*C*) Effect sizes for PCs, LPCs, and PC/LPCs (z-score normalized) in RILs that are either homozygous for B73 or PT at 8.5 Mb on chromosome 3 (*Top Row*) and CRISPR-Cas9 $hpc1^{Tins}$ mutant and wild-type plants (*Bottom Row*). Significant differences were tested by *t* test; the resulting *P* values are shown. (*D*) Effect sizes for individual PC and LPC species (z-score normalized) in RILs at 8.5 Mb on chromosome 3 (*Left*) or the CRISPR/Cas9 $hpc1^{Tins}$ mutant (*Right*). *Significant difference at $P < 0.05$ (*t* test, after Benjamini and Hochberg correction); ns, not significant. (*E*) *HPC1* expression analysis in B73, PT, and their F_1 hybrids grown in standard and cold temperatures in a growth chamber. Significant differences were tested by *t* test with Benjamini and Hochberg correction; the resulting *P* values are shown. (*F*) PC/LPC ratio for several RILs. RIL B042 (indicated by the black arrow) bears a recombination event 500 bp upstream of the *HPC1* translation start codon. In *C*–*F*, phenotypes associated with the B73 haplotype are in red; the equivalent values for the PT haplotype are in blue.

phospholipase based on its two closest *Arabidopsis* orthologs (encoded by *At1g06800* and *At2g30550*) (52). PLA1-type phospholipases hydrolyze phospholipids in the *sn-1* position and produce a lyso-phospholipid and a free fatty acid (*SI Appendix*, Fig. S3*B*). In B73, *HPC1* was one of the most highly expressed phospholipase genes, with expression almost exclusively restricted to vegetative leaves (V4 to V9) (*SI Appendix*, Fig. S4*A*) (53), which was the biological material we sampled for glycerolipid analysis. In B73 leaves, *HPC1* was also the most highly expressed gene within the QTL interval (*SI Appendix*, Fig. S3*C*) (53). Class I phospholipases are chloroplast-localized proteins; in agreement, we identified a chloroplast transit peptide (CTP) at the beginning of the predicted HPC1 sequence using the online tool ChloroP (54). We validated the chloroplast localization of HPC1 by transiently expressing a construct encoding the HPC1 CTP fragment fused to green fluorescent protein (GFP) in *Nicotiana benthamiana* leaves (*SI Appendix*, Fig. S5).

The effect of *HPC1* on PC/LPC levels may be caused by misregulation of *HPC1* expression in highland landraces and/or by a mutation affecting HPC1 enzymatic activity. To distinguish between these two possibilities, we analyzed *HPC1* expression in B73, PT, and the corresponding F_1 hybrid plants grown at high and low temperatures to simulate highland and lowland conditions, respectively (Fig. 3*E*). Under cold conditions, *HPC1-B73* was up-regulated, but *HPC1-PT* was not (Fig. 3*E*). The lack of up-regulation in cold conditions of *HPC1* may explain the high PC/LPC levels in PT. However, *HPC1* expressed to the same levels in B73 and PT under control conditions. In the F_1 hybrids, *HPC1* expression was consistent with a dominant B73 effect. We also observed a dominant B73 effect at the metabolic level when we analyzed B73 × PT RILs that are heterozygous at the *HPC1* locus (*SI Appendix*, Fig. S3*D*). Variation in *HPC1* may also affect enzymatic activity of the HPC1-PT variant. To test this hypothesis we sequenced three B73 × PT RILs (B021, B042, B122) that are homozygous for the *HPC1-PT* allele. We discovered a recombination point between 493 and 136 bp upstream of the *HPC1* translation start codon (Fig. 3*F* and *SI Appendix*, Fig. S6) in RIL B042, resulting in a chimeric locus with the coding region from PT combined with a promoter segment from B73. PC/LPC levels in RIL B042 were similar to other RILs that are homozygous for the PT haplotype at the 8.54-Mb marker at the QTL peak (Fig. 3*F*). This result supports the hypothesis that the metabolic effect in the B73 × PT RILs is likely due to an impaired function of the HPC1-PT enzyme rather than to changes in the *HPC1-PT* regulatory region. However, regulatory variants in the first 500 bp of the promoter may also impact expression levels in RILB042.

If *HPC1* is the underlying causal gene of this QTL, the observed metabolic phenotypes would be consistent with a reduction or loss of HPC1-PT enzyme function, leading to higher levels of PCs and lower levels of LPCs in the PT background. To test this hypothesis we generated mutants in *HPC1* via CRISPR/CRISPR-Cas9–mediated genome editing (*SI Appendix*, Fig. S7*A*) in the B104 inbred, a temperate stiff-stalk inbred derived from the Iowa Stiff Stalk Synthetic population like B73. We identified two transgenic mutants, hereafter designated $hpc1^{CR\ T\ ins}$ and $hpc1^{CR\ T\ del}$ (*SI Appendix*, Fig. S7*A*). We then measured PC and LPC species in wild-type and mutant plants grown under long day conditions. The phospholipid profiles

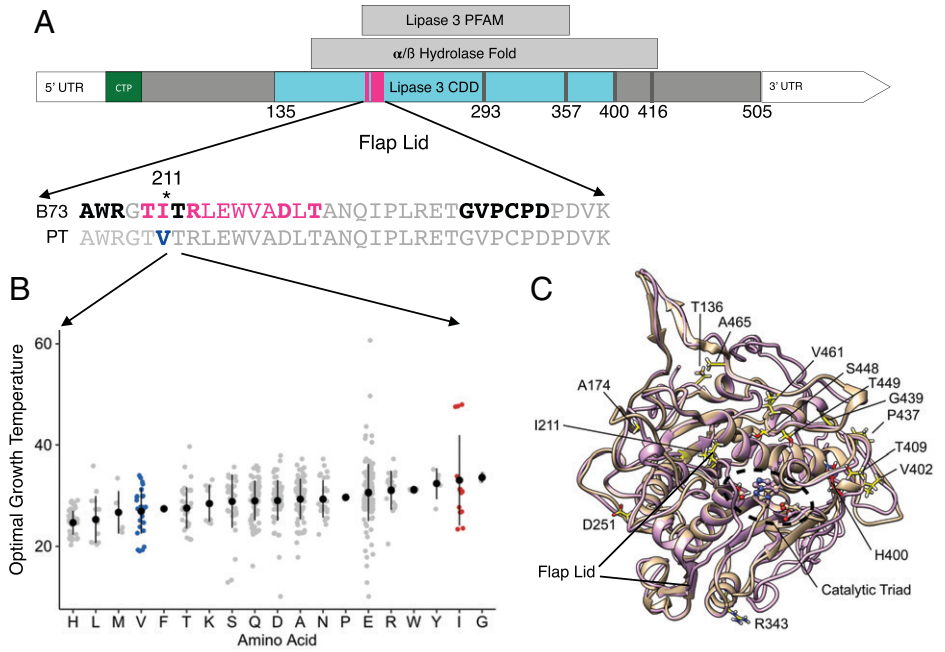

Fig. 4. Analysis of the possible causal SNP affecting the flap-lid domain of HPC1. (*A*) *HPC1* coding sequence showing different encoded features. CTP, chloroplast transit peptide. The coding sequence fragment that encodes the flap-lid domain and the corresponding amino acids are shown in magenta. Residues in boldface type showed significant associations with prokaryote optimal growth temperature (55, 59). (*B*) Optimal prokaryote growth temperature (OGT) test of different amino acids at the residue equivalent to the polymorphic residue 211 in HPC1. (*C*) Top I-TASSER model (purple) for the putative phospholipase encoded by *HPC1-B73*, overlaid on chain B of the crystal structure of phospholipase A1 from *Arabidopsis* (PDB 2YIJ, tan). Residues that differ between the PT and B73 phospholipase are shown in yellow and some are labeled.

of the $hpc1^{CR}$ plants replicated those of the *PT* allele in the RILs (Fig. 3 *C* and *D*, *Bottom* and *SI Appendix*, Fig. S7*B*), confirming that the *HPC1-PT* allele impairs HPC1 function and thus underlies the QTL on chromosome 3 around 8.5 Mb. Finally, we in vitro translated HPC1-B73 and HPC1-PT versions of the protein in a cell-free system and incubated them with various phospholipid substrates. We then measured the amount of phospholipid substrate and lyso-phospholipid product for each compound (*SI Appendix*, Fig. S8). This experiment confirmed that both HPC1 variants have PLA1 activity and suggested that HPC1-B73 may have higher activity on substrates like PC36:4 that show stark differences in abundance between highland and lowland lines, as well as in $hpc1^{CR}$ lines. Together, the RIL and CRISPR mutant results showed that *HPC1* underlies a major metabolic QTL explaining PC/LPC ratio. A mutation in the flap-lid domain of HPC1 affecting substrate accessibility likely leads to impaired function in the highland PT allele.

D. Identification of the Putative Causal SNP in *HPC1*. We Sanger sequenced the *HPC1* locus from several RILs harboring the PT haplotype at *HPC1* and identified several nonsynonymous SNPs within the coding sequence that might influence HPC1 function (*SI Appendix*, Fig. S11).

We focused our attention on SNP 631 affecting the flap-lid domain, which led to a conservative replacement of valine by isoleucine (V211I, Fig. 4*A*). The flap-lid domain is important for phospholipase activity and is located in a lipase class 3 domain (Protein Families (PFAM) database domain PF01764) that is highly conserved across the tree of life (55). We recovered 982 observations of the lipase class 3 PFAM domain from 719 prokaryotic species using PfamScan (56, 57) and estimated optimal growth temperatures from their tRNA sequences (58). We then tested whether genetic variation in the sequence encoding the lipase class 3 domain was significantly associated with optimal growth temperature in bacteria (59). We detected several significant associations, all of which were located in the flap-lid region (Fig. 4*A*, letters in boldface type). Notably, the presence of a valine at residue 211, as observed in the PT allele, was accompanied by lower optimal growth temperatures relative to an isoleucine at residue 211, as observed in B73 (Fig. 4*B*), suggesting that the PT allele may be better adapted to the low temperatures to which highland maize is exposed.

We used the crystal structure of phospholipase A1 from *Arabidopsis* (Protein Data Bank [PDB] 2YIJ) to model the structure of HPC1-B73 by Iterative Threading Assembly Refinement (I-TASSER) (Fig. 4*C*). Residues that differ between HPC1-PT and HPC1-B73 are shown in yellow. The catalytic triad and H400 identified from the conserved domain database (CDD) conservation analysis are also labeled. Our models placed H416 rather than H400 in the catalytic triad. Of the residues that differ between HPC1-PT and HPC1-B73, residues 211, 448, and 449 were the closest to the catalytic triad. I211 was positioned on the N terminus of the flap-lid domain and is well suited to stabilize binding of a lipid substrate through hydrophobic interactions. Replacement of I211 with V211 results in the loss of a methyl group that may influence the strength of these hydrophobic interactions, affect substrate binding, and/or affect the dynamics of the flap-lid domain. These results strongly point to the mutation in the flap-lid domain as the most likely underlying mutation affecting HPC1 activity.

E. *HPC1* Shows Strong Elevation-Dependent Antagonistic Pleiotropy in Mexican Landrace Fitness Phenotypes. Our selection and QTL analysis provided strong evidence that *HPC1* is under selection in highland maize and controls phospholipid metabolism. To evaluate the possible fitness effects of HPC1 variation in locally adapted landraces across Mexico, we reanalyzed phenotypic data from a previously reported F_1 association mapping panel (15, 44) composed of about 2,000 landrace F_1s grown in 23 common garden environments across an elevation

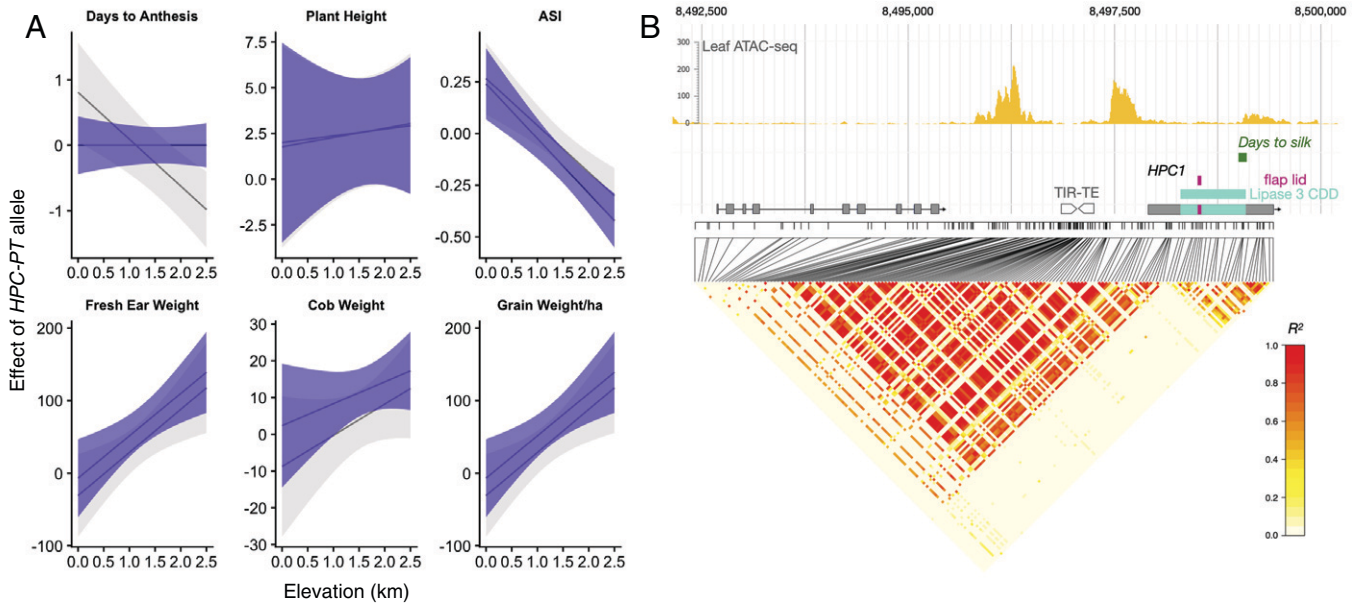

Fig. 5. Fitness effects of the *HPC1-PT* allele and *HPC1* LD analysis. (*A*) We used best linear unbiased predictions (BLUPs) and GBS data from 2,700 landrace topcrosses from ref. 15, evaluated in 23 common gardens at different elevations in Mexico. We modeled each trait as a function of the *HPC1-PT* genotype, trial elevation, and tester line, with controls for main effects and responses to elevation of the genomic background. Gray lines and ribbons show estimates of the effect of the highland allele of *HPC1-PT* as a function of common garden elevation ±2 SEM, using the *GridLMM* package (60). Purple lines show estimates of the *HPC1-PT* effect in a model that also adjusts for effects of days to anthesis. ASI, anthesis to silking interval. (*B*) Linkage disequilibrium and genomic features of *HPC1* and upstream region. LD heatmap was drawn from ref. 83 using HapMap 3 data. *Top* track shows leaf ATAC-seq peaks in B73 (data from ref. 64). The regions coding for the lipase and flap-lid domains are highlighted on the *HPC1* gene model. See Fig. 4*A* for further details. The term days to anthesis indicates the GWAS SNP on the lipase domain from ref. 62. TIR-TE, terminal inverted repeat transposable element. Tracks were obtained from the MaizeGDB B73 V5 browser (61).

gradient. We then fitted a model to estimate the effect of variation at *HPC1-PT* on the relationship between fitness traits and elevation (60). *HPC1* was a clear outlier in a genome-wide association study (GWAS) of genotype-by-elevation fitness traits like flowering time and yield (*SI Appendix*, Fig. S9 *A* and *B*), indicating that elevation-dependent variation at *HPC1* has an effect not only on phospholipid levels but also on fitness traits. Indeed, variation at *HPC1* showed significant genotype × elevation effects for several fitness traits (Fig. 5*A*). The effect of *HPC1* on flowering time revealed antagonistic pleiotropy between highland and lowland environments (Fig. 5*A*). The highland *HPC1-PT* allele was associated in low elevations with delayed flowering, increasing days to anthesis (DTA) by about 1 d. Meanwhile the same allele exhibited accelerated flowering at high elevation (with a decrease in DTA of almost 1 d; Fig. 5*A*). Variation at *HPC1* also displayed conditional neutrality on fresh ear weight and grain weight per hectare traits: The highland allele had no effect in lowland environments but was associated with greater values in highland environments (Fig. 5*A*). Other known domestication and flowering-time genes did not show a clear genotype-by-environment (G × E) effect (*SI Appendix*, Fig. S10). We also checked previous reports for associations between *HPC1* and flowering time in other populations through the MaizeGDB (61) genome browser. We in fact found a significant flowering-time SNP in the *HPC1* coding sequence (Fig. 5*B*) for the nested association mapping (NAM) population (62). This additive flowering-time locus is only 6 bp from the focal SNP we used to test G × E at *HPC1* in the SEEDs panel (Fig. 4). Variation at this SNP correlated with a reduction in flowering time of 8.5 h, relative to B73, and explained 1.12% of the trait variance, which is about one-third of the largest effect observed for flowering-time variation in the NAM population.

We then used genetic marker data from the HapMap 3, which includes the NAM parents (63), to analyze linkage disequilibrium (LD) of the *HPC1* region (Fig. 4*B*). We detected a strong LD block of about 150 bp in length in the coding sequence that includes the focal SNP mentioned above (Fig. 5 *A* and *B*). We identified another LD block covering the 5′ region of *HPC1* and the promoter region up to 2 kb upstream of *HPC1* (Fig. 5*B*). Interestingly, this second LD block on the promoter overlapped with two strong ATAC-seq (assay for transposase-accessible chromatin followed by sequencing) peaks identified in B73 (64) (Fig. 5*B*). These results confirmed that the SNPs associated with fitness traits like flowering time on the *HPC1* coding sequence are not linked to other SNPs upstream of the *HPC1* coding sequence. However, the SEEDs dataset lacks GBS markers for several kilobases upstream of *HPC1*, raising the possibility of a second regulatory variant in the promoter (Fig. 4*B*) that might have an effect on *HPC1* expression. We further evaluated the possible effect of *HPC1* on flowering using both $hpc1^{CR}$ mutants in long-day conditions during the summer of 2021 in Raleigh, NC. Although the mutants showed high PC/LPC ratios, we observed no significant difference in flowering time relative to the wild type (*SI Appendix*, Fig. S7*C*). As shown in Fig. 4*A* the effect of the *HPC1* allele on flowering time has a strong G × E pattern and we observe a significant effect only in very high or low elevations. We speculate that the absence of significant differences in the B104 CRISPR mutants in Raleigh conditions could partly be explained by the strong G × E effect of *HPC1* and/or genetic background effects.

In summary, the SNPs in the short, lipase-domain–encoding LD block of *HPC1* show strong genotype × elevation fitness effects in both Mexican landraces grown across multiple altitudes and the NAM population. The phospholipid changes induced by *HPC1* have physiological effects that may explain the strong selection of *HPC1* in highland environments.

F. *HPC1-PT* Was Introgressed from Teosinte *mexicana* and Is Conserved in Flint Inbred Lines. We explored the segregation of the V211I SNP among other highland maize varieties. We detected the PT allele at high frequencies in highland landraces

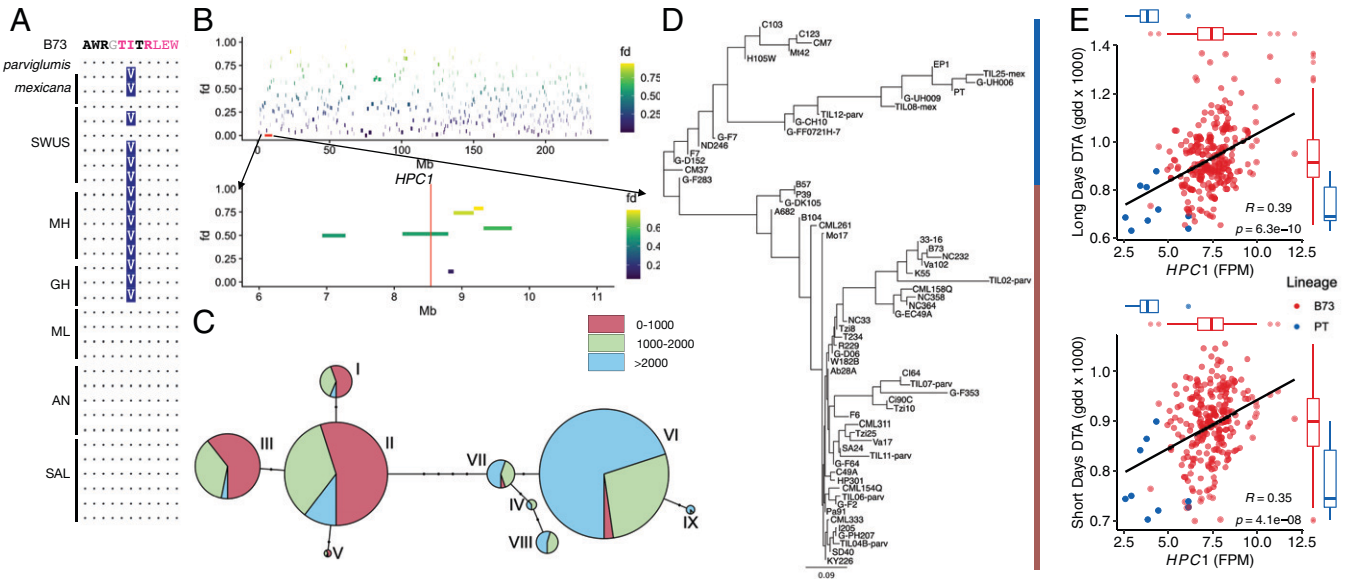

Fig. 6. Introgression of teosinte *Mexicana* into maize *HPC1*. (*A*) Alignments around the V211I polymorphism in the flap-lid domain of HPC1 in B73, *mexicana*, and *parviglumis* and landraces of the SWUS, MH, Mexican lowlands (ML), GH, AN, and South America lowlands (SAL). (*B*) f_d analysis of the *mexicana* introgression. Data were obtained from ref. 10. (*C*) Haplotype network analysis of SNPs within the *HPC1* coding region, using 1,060 Mexican homozygous individuals from the SeeD dataset. Haplotypes are color coded by elevation: red, 0 to 1,000 masl; green, 1,000 to 2,000 masl; blue, >2,000 masl. (*D*) Cluster analysis of the *HPC1* coding region using a sample of HapMap3 inbred lines and the PT landrace. (*E*) Correlation between *HPC1-PT* expression and DTA in plants grown in short or long days. Inbred lines from the PT lineage shown in *C* are indicated in blue and inbred lines from the B73 lineage are in red; data are from ref. 65.

from Mexico and Guatemala. In addition, the PT allele segregated in Southwestern US landraces. The B73 allele was fixed in lowland Mexican, South American, and Andean landraces (Fig. 6*A*). These results were consistent with our PBE results (Fig. 2*F*). The PT allele was also present in one-fourth of all teosinte *parviglumis* accessions tested and in both *mexicana* accessions reported in HapMap 3 (63) (Fig. 6*A*). This observation prompted us to examine whether the PT allele was the result of postdomestication introgression from teosinte *mexicana* during maize highland colonization or whether it was selected from *parviglumis* standing variation. To test for introgression from *mexicana*, we used f_d data from ref. 10 and established that the genomic region containing *HPC1* shows signatures of introgression from *mexicana* into highland maize (Fig. 6*B*). We then performed a haplotype network analysis using SNP data from the *HPC1* coding region of 1,160 Mexican accessions from the SeeD dataset (44) that are homozygous for all SNPS across the coding region and the teosinte inbred lines (TILs) from HapMap 3 (63). We identified nine haplotype groups that cluster mainly based on elevation (Fig. 6*C*). The two major groups, II and VI, contained mainly lowland and highland landraces, respectively. The two *mexicana* teosinte inbred lines (TIL08 and -25) belonged to group IV (Fig. 6*C*) together with highland landraces primarily collected in the Trans-Mexican Volcanic Belt (30/36 from the highlands of Jalisco, Michoacán, México, Puebla, and Veracruz). We then examined whether this *mexicana* haplotype (denoted *ZxHPC1*) that is introgressed into Mesoamerican highland maize was also present in modern maize inbred lines.

To this end, we performed a neighbor-joining cluster analysis using HapMap 3 (63) inbred lines including those from the 282 inbred panel, teosinte inbred lines, German lines, and PT. We identified two main groups, one containing the *HPC1-PT* haplotype and the other containing the *HPC1-B73* haplotype. PT and the teosinte *mexicana* inbred lines TIL08 and TIL25 clustered together with Northern European Flint inbred lines such as EP1, UH008, and UH009 (Fig. 6*D*). Other Northern US flints, e.g., CM7, were also closely related to the *mexicana* *ZxHPC1* haplotype. These data suggest that after introgression into highland maize, the *ZxHPC1* haplotype was maintained in Flint materials adapted to cold environments in the Northern United States, Canada, and Europe. We then used a large gene expression dataset consisting of multiple developmental stages of the 282-maize diversity panel (65) and phenotypic datasets collected from the same panel grown in long days and short days. Notably, *HPC1* expression levels were highly correlated with flowering time in both long and short days. Lines carrying the *HPC1-PT* allele were characterized by lower *HPC1* expression and earlier flowering times relative to lines carrying the *HPC1-B73* allele (Fig. 6*E*). Taken together, these data show that *HPC1* was introgressed from teosinte *mexicana* into highland maize and that this introgression was carried over into high-latitude–adapted modern inbred lines with low *HPC1* expression and early flowering.

G. *HPC1*/Phosphatidylcholine Interactions with Maize Flowering-Time Protein ZCN8. Using the same expression dataset from the 282 panel (65), we discovered that *HPC1* and *ZmLPCAT1* expression levels are inversely correlated in most tissues (*SI Appendix*, Fig. S12), further supporting the idea that these two enzymes are coordinately regulated. We observed significant associations between *HPC1* expression levels in aerial tissues and several flowering-time traits (*SI Appendix*, Fig. S12). The magnitude of these associations was similar to that seen for other well-characterized flowering-time genes (*SI Appendix*, Fig. S12) such as *ZEA CENTRORADIALIS8* (*ZCN8*) and the APETALA2(AP2)/ETHYLENE RESPONSE FACTOR (ERF) transcription factor gene RELATED TO AP2.7 (*ZmRAP2.7*). In *Arabidopsis*, florigen (FT) has recently been shown to interact with both PC and phosphatidylglycerol (PG), depending on the ambient temperature and the cellular location of FT (36, 38). As ZCN8 is a homolog of FT, it may mediate the hastening of flowering time seen in highland maize. Recent work improved upon the crystal structure of *Arabidopsis* FT, onto which we modeled ZCN8 and compared

the PC binding sites (*SI Appendix*, Fig. S13 *A* and *B*) (66). Using AutoDock Vina to simulate docking and the sites from ref. 66, we identified sites 2 and 4 and 1 and 4 as potential binding sites of PC34:2 and PC36:2 with site 4 heavily favored in both cases (*SI Appendix*, Fig. S13 *A–C*). As maize ZCN8 and *Arabidopsis* FT are highly conserved, the predicted binding-site similarity is not surprising (*SI Appendix*, Fig. S13*D*). We next heterologously produced and purified ZCN8 fused to a SPOT tag from yeast (*Saccharomyces cerevisiae*) cells to assess lipid binding to ZCN8 in vivo. We extracted and analyzed all lipids copurifying with the recombinant protein following the same lipidomics pipeline used for maize lipids. We identified PC34:2 in the purified ZCN8 protein in all samples, confirming experimentally that ZCN8 effectively bound PC species (*SI Appendix*, Fig. S14).

3. Discussion

Understanding the genetic, molecular, and physiological basis of crop adaptation to different environments and the role that wild relatives have played in these processes is relevant for the identification of favorable genetic variation that can be used to improve modern crops. The repeated events of maize adaptation to highland environments constitute an excellent natural experiment to study local adaptation. Recent studies (14, 18, 19) have helped expand our understanding of the genetic basis underlying maize highland adaptation. However, the responsible molecular, physiological, and genetic mechanisms underlying maize highland adaptation and the possible role of highland maize traits in modern, commercial varieties remain largely unknown. Phospholipids are key structural components of plant membranes that also function as signaling molecules during adaptation to stresses that would be prevalent in highland environments (52, 67) such as low phosphorus availability (68–70) and low temperatures (30, 31, 71). In *Arabidopsis* and rice (*Oryza sativa*), phospholipid species regulate flowering time via interactions with *Arabidopsis* FT and rice Heading date 3a (Hd3a), respectively (36, 38, 72). Flowering time is a major driver of maize adaptation to highland environments (15, 44, 73) and to northern latitudes (21, 22).

Genes involved in the biosynthesis and degradation of phospholipids appeared to have been repeatedly selected in several highland maize populations of North America, Central America, and South America (Figs. 1 and 2). *ZmLPCAT1* and *HPC1* were two such genes with the strongest, repeated signals of selection, as measured by PBE and *pcadapt* in highland populations (Fig. 2). In a previous study, *ZmLPCAT1* exhibited high F_{ST} values when highland and lowland landraces were compared (18). In temperate inbred lines, *HPC1* was up-regulated while *ZmLPCAT1* was down-regulated by cold stress (*SI Appendix*, Fig. S4 *B* and *C*) (74). These expression patterns are in agreement with the high and low PC/LPC ratios observed in our experiments. Furthermore, we determined here that *HPC1* is up-regulated in B73 and B73 × PT F_1 hybrids after cold exposure. By contrast, the *HPC1-PT* and *HPC1-B73* alleles were expressed to comparable levels in control conditions, and the *HPC1-PT* allele was not induced upon cold conditions (Fig. 3*E*). Selection at these two loci is likely to have driven the high PC/LPC ratio we observed in highland Mexican landraces (Fig. 1).

Our QTL analysis of the PC/LPC ratios in a B73 × PT mapping population and in the $hpc1^{CR}$ mutant alleles demonstrates that the highland *HPC1-PT* allele results in an enzyme with impaired function that alters highland Mexican maize PC metabolism, leading to higher PC/LPC ratios (Fig. 3 and *SI Appendix*, Fig. S7*B*). Adaptive loss-of-function mutations can be an effective way to gain new metabolic functions in new environmental conditions (75). Taking advantage of the conserved lipase domain of HPC1 in bacteria, we used a method that can identify how genetic variation in DNA regions encoding protein domains is correlated with optimal growth temperature of bacteria (55, 59). The probable causal SNP in HPC1 changed an amino acid in the flap-lid domain of HPC1 (Fig. 5*A*) that may affect substrate accessibility and/or substrate binding (Fig. 5*A*). Indeed, the flap-lid domain has been the target of biotechnological modification for lipases (76). The region encoding the flap-lid domain was located at a recombination point that separates two clear LD blocks. The first LD block covered a 2-kb promoter region of *HPC1* and the first 600 bp of *HPC1*, while the second one covered the rest of the *HPC1* coding sequence. In fact, this pattern is characteristic of selective sweeps that leave two LD blocks on either side of the adaptive mutation sweep (77). LD analysis together with the open chromatin detected by ATAC-seq that overlaps the same promoter region (Fig. 4*B*) and the correlation between *HPC1* expression and flowering time (Fig. 6*E*) presented here suggest a possible transcriptional regulation of *HPC1* that may act additively with the coding sequence variation we described.

Why were the metabolic changes induced by HPC1 variation selected for in highland maize? PC metabolism is intimately connected to multiple stress responses and developmental pathways; alterations in PC amounts and PC/LPC ratios affect overall plant fitness. The *qPC/LPC3@8.5* QTL is driven by individual QTLs for PC and LPC species with high levels of unsaturated fatty acids (Fig. 3*D*). Several of these species, like PC 36:5 and LPC 18:1 (Fig. 3*D* and Dataset S3), have been shown to display similar patterns during *Arabidopsis* cold acclimation (31) and sorghum (*Sorghum bicolor*) low-temperature responses (71). PC 36:5 also showed high Q_{ST} values when comparing highland and lowland landraces from both Mesoamerica and South America (Fig. 1 *C* and *D* and Dataset S5). In maize, *HPC1* expression is under the control of the circadian clock (78) with a peak at the end of the day. In *Arabidopsis*, highly unsaturated PC (34:3, 34:4, 36:5, 36:6) species increase in the dark (79). This peak in contents coincides in maize with low *HPC1* expression levels during the same dark hours (78). PC, and lipid metabolism in general, is also intimately connected to flowering time. For instance, PCs were shown to bind to *Arabidopsis* FT in the shoot apical meristem to hasten flowering (38, 66) by unknown cellular mechanisms. Similarly, PG species can sequester FT in phloem companion cells in low temperatures (36) and then release FT into the phloem later after temperatures increase, allowing FT to reach the shoot apical meristem. In agreement with an effect of lipids on flowering time, overexpression of a gene encoding a secretory phospholipase D delayed heading time in rice (72). In line with a role for phospholipids in flowering, we established that genetic variation at SNPs within the region of *HPC1* encoding the lipase domain exhibits a strong interaction with elevation for the highland *HPC1-PT* allele. This variation leads to a delay in flowering time in low elevations and an acceleration at high elevations, both of which are close to 1 day in amplitude. Interestingly, the effect of the highland *HPC1* allele exhibited typical conditional neutrality in yield-related traits, with higher fitness conferred by the *HPC1-PT* allele in highlands (Fig. 4*A*). In the NAM population, we identified another SNP mapping to the region encoding the lipase domain that is associated with a hastening of flowering time by 8 h with respect to B73 (62), further supporting the role of *HPC1* in controlling flowering time. Analysis of *HPC1* CRISPR mutant alleles in the B104 background grown in Raleigh displayed a PC/LPC phenotype that mimicked the highland allele but not a flowering-time phenotype (*SI Appendix*, Fig. S7), probably due to the strong G × E effect of *HPC1* and/or genetic background effects.

The strong G × E effect we observed in *HPC1* is similar to the well-known teosinte *mexicana* introgression *inv4m* (14). In fact, our analysis showed that *HPC1* is indeed another introgression from teosinte *mexicana* (Fig. 6 *A*–*C*). Recent analysis using sympatric teosinte and maize populations across elevation gradients in Mexico further supports the introgression of *mexicana* at *HPC1* and shows that the *mexicana* ancestry of *HPC1* increases at a rate of +0.079 per 100 m of elevation (80). We further demonstrated that the *mexicana* intogression at *HPC1* is conserved in high-latitude–adapted Flint lines from both Europe and the United States (Fig. 6*D*). *HPC1* in inbred lines carrying the highland *ZxHPC1 mexicana* haplotype was expressed at low levels and resulted in earlier flowering (Fig. 6*E*) (65).

Adaptation to higher latitudes involved a reduction of photoperiod sensitivity and flowering time that enabled maize to thrive in longer-day conditions characteristic of the growing season at high latitudes (21, 22, 26, 27). Additive mutations in the regulatory region of the gene *ZCN8* (81), including a teosinte *mexicana* introgression, lead to higher expression of *ZCN8*, which contributes to maize adaptation to long days in temperate conditions (82). *ZCN8* is a close ortholog of *Arabidopsis FT*, whose encoding protein interacts with several species of phospholipids to modulate flowering time (36, 38). A similar interaction was also demonstrated for the rice FT ortholog Hd3a (72), and we hypothesize that the same may be occurring in maize. Comparison of docking simulations of phospholipids with ZCN8 using the *Arabidopsis* FT crystal structures as a model (66) showed similar PC interactions in ZCN8 (*SI Appendix*, Fig. S13). We corroborated this interaction via mass spectrometry analysis of lipids bound to ZCN8 heterologously produced in yeast (*SI Appendix*, Fig. S14).

In summary, we used a combination of genomic scans, linkage mapping, lipidomics, and reverse genetics to identify and clone the adaptive gene *HPC1*, introgressed from teosinte *mexicana*, in highland maize landraces. HPC1 variants lead to a major reorganization of phosphatidylcholine metabolism. We showed that the fitness advantage conferred by the *HPC1* highland *mexicana* allele is due, at least in part, to its association with flowering time. This effect may have contributed to adaptation of maize to colder, higher latitudes.

Our work identifies the important role of a gene controlling phospholipid metabolism in plant local adaptation and further supports the emerging role of phospholipid metabolism in fine-tuning flowering time across different plant species (36, 38, 82). This study highlights the largely underappreciated role of highland maize and highland teosinte *mexicana* in modern maize.

4. Materials and Methods

The diversity panels and mapping populations used in this paper for population genetics measures of selection, QTL mapping, and G × E analysis have been described previously by refs. 15, 19, 41, 44, and 50. Lipidomics analyses were performed with high-resolution mass spectrometry UPLC-MS (Ultra Performance Liquid Chromatography - Mass Spectrometry). Mutant alleles of *HPC1* were obtained using CRISPR-Cas9 editing. Full description and details of all materials and methods are provided in *SI Appendix*. All the data and code are contained within this paper, *SI Appendix*, and the associated GitHub repository of the project.

Data Availability. Code and data have been deposited in https://github.com/sawers-rellan-labs/High-PC1-paper (84). The mass spectrometry data has been deposited in the metabolights repository at www.ebi.ac.uk/metabolights/MTBLS5074 (85). All other data are included in this article and/or supporting information. Previously published data were used for this work (https://hdl.handle.net/11529/10548233) (86).

ACKNOWLEDGMENTS. This work was supported by Conacyt Young Investigator (CB-238101), Consejo Nacional de Ciencia y Tecnología National Problems (APN-2983), University of California-Mexus, and North Carolina State startup funds awarded to R.R.-Á. Fieldwork and mapping population development were supported by NSF-Plant Genome Research (PGR) Award 1546719 to D.R., J.R.-I., M.B.H., and R.J.H.S. A.C.B. was supported by NSF-Plant Genome Research Program, Postdoctoral Research Fellowships in Biology (PGRP PRFB) Grant 2010703. A.K. was supported as an Agricultural Biotechnology in Our Evolving Food, Energy, and Water Systems fellow by NSF Award 1828820. F.R.-Z. was supported by the Science and Technologies for Phosphorus Sustainability Center, a NSF Science and Technology Center (CBET-2019435). This study made use of NMRbox: National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource, which is supported by NIH Grant P41GM111135 (National Institute of General Medical Sciences). This work was performed in part by the Molecular Education, Technology and Research Innovation Center at North Carolina State University, which is supported by the State of North Carolina. We thank Josh Strable, Peter Balint-Kurti, Luis Herrera-Estrella, and James Holland for critical review of the manuscript and feedback. We thank the members of the M.B.H., R.J.H.S., R.R.-Á., and Flint-Garcia laboratories for their help developing and evaluating mapping populations. We thank Cruz Robledo and the Puerto Vallarta Agricultura Invernal team, including members of the local Wixárika community for their work in our lowland field. We thank Fernando Delgado, Denise Costich, and Cristian Gálvez and the staff of the International Maize and Wheat Improvement Center (CIMMYT) Metepec field station and the CIMMYT Germplasm bank for their assistance in genetic nurseries and field evaluation. We thank the CIMMYT and Puerto Vallarta Winter Nursery crews that have helped generate and evaluate the mapping populations used in this paper. We thank the crew of the Central Crops Research Station at North Carolina State for their help and assistance with field experiments. We specially acknowledge the indigenous people of the Americas and the ingenuity through which they domesticated and made the spread and adaptation of maize throughout the continent possible. We acknowledge the painful history of Native American peoples dispossession and honor the ongoing connection of Native American peoples, past and present, to the lands and waters where the field experiments described here were performed. This work would not have been possible without the international maize research community and the willingness of so many colleagues to support the development of new research programs.

Author affiliations: [a]Department of Molecular and Structural Biochemistry, North Carolina State University, Raleigh, NC 27695; [b]National Laboratory of Genomics for Biodiversity 36821, Irapuato, Mexico; [c]Department of Evolution and Ecology, Center for Population Biology, and Genome Center, University of California, Davis, CA 95616; [d]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011; [e]Department of Plant Biology, University of Georgia, Athens, GA 30602; [f]US Department of Agriculture–Agricultural Research Service, Cornell University, Ithaca, NY 14853; [g]Department of Plant Sciences, University of California, Davis, CA 95616; [h]Department of Chemistry, North Carolina State University, Raleigh, NC 27695; [i]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724; [j]Molecular Education, Technology and Research Innovation Center, North Carolina State University, Raleigh, NC 27695; [k]West Coast Metabolomics Center, University of California, Davis, CA 95618; and [l]Department of Plant Science, The Pennsylvania State University, PA 16802

1. X. Yi *et al.*, Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010).
2. A. Bigham *et al.*, Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. *PLoS Genet.* **6**, e1001116 (2010).
3. J. Yang *et al.*, Genetic signatures of high-altitude adaptation in Tibetans. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4189–4194 (2017).
4. F. Cicconardi *et al.*, Genomic signature of shifts in selection in a subalpine ant and its physiological adaptations. *Mol. Biol. Evol.* **37**, 2211–2227 (2020).
5. E. C. Gilmore Jr., J. S. Rogers, Heat units as a method of measuring maturity in corn 1. *Agron. J.* **50**, 611–615 (1958).
6. J. L. Hatfield, J. H. Prueger, Temperature extremes: Effect on plant growth and development. *Weather Clim. Extrem.* **10**, 4–10 (2015).
7. Y. Matsuoka *et al.*, A single domestication for maize shown by multilocus microsatellite genotyping. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 6080–6084 (2002).
8. D. R. Piperno, K. V. Flannery, The earliest archaeological maize (*Zea mays* L.) from highland Mexico: New accelerator mass spectrometry dates and their implications. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 2101–2103 (2001).
9. M. B. Hufford *et al.*, The genomic signature of crop-wild introgression in maize. *PLoS Genet.* **9**, e1003477 (2013).

10. E. Gonzalez-Segovia *et al.*, Characterization of introgression from the teosinte *Zea mays* ssp. *mexicana* to Mexican highland maize. *PeerJ* **7**, e6815 (2019).
11. J. A. Aguirre-Liguori *et al.*, Divergence with gene flow is driven by local adaptation to temperature and soil phosphorus concentration in teosinte subspecies (*Zea mays* parviglumis and *Zea mays* mexicana). *Mol. Ecol.* **28**, 2814–2830 (2019).
12. N. Lauter, C. Gustus, A. Westerbergh, J. Doebley, The inheritance and evolution of leaf pigmentation and pubescence in teosinte. *Genetics* **167**, 1949–1959 (2004).
13. A. K. Hardacre, H. A. Eagles, Comparisons among populations of maize for growth at 13°C. *Crop Sci.* **20**, 780–784 (1980).
14. T. Crow *et al.*, Gene regulatory effects of a large chromosomal inversion in highland maize. *PLoS Genet.* **16**, e1009213 (2020).
15. D. J. Gates *et al.*, Single-gene resolution of locally adaptive genetic variation in Mexican maize. *bioRxiv* [Preprint] (2019). https://doi.org/10.1101/706739. Accessed 8 February 2021.
16. J. Stephen Athens *et al.*, Early prehistoric maize in northern highland Ecuador. *Lat. Am. Antiq.* **27**, 3–21 (2016).
17. A. Grobman *et al.*, Preceramic maize from Paredones and Huaca Prieta, Peru. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1755–1759 (2012).
18. S. Takuno *et al.*, Independent molecular basis of convergent highland adaptation in maize. *Genetics* **200**, 1297–1312 (2015).
19. L. Wang *et al.*, Molecular parallelism underlies convergent highland adaptation of maize landraces. *Mol. Biol. Evol.* **38**, 3567–3580 (2021).
20. R. R. da Fonseca *et al.*, The origin and evolution of maize in the Southwestern United States. *Nat. Plants* **1**, 14003 (2015).
21. K. Swarts *et al.*, Genomic estimation of complex traits reveals ancient maize adaptation to temperate North America. *Science* **357**, 512–515 (2017).
22. H.-Y. Hung *et al.*, ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E1913–E1921 (2012).
23. Z. Liang *et al.*, Conventional and hyperspectral time-series imaging of maize lines widely used in field trials. *Gigascience* **7**, 1–11 (2018).
24. L. Guo *et al.*, Stepwise cis-regulatory changes in ZCN8 contribute to maize Flowering-Time adaptation. *Curr. Biol.* **28**, 3005–3015.e4 (2018).
25. N. D. Coles, M. D. McMullen, P. J. Balint-Kurti, R. C. Pratt, J. B. Holland, Genetic control of photoperiod sensitivity in maize revealed by joint multiple population analysis. *Genetics* **184**, 799–812 (2010).
26. C. Huang *et al.*, *ZmCCT9* enhances maize adaptation to higher latitudes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E334–E341 (2018).
27. Q. Yang *et al.*, CACTA-like transposable element in ZmCCT attenuated photoperiod sensitivity and accelerated the postdomestication spread of maize. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16969–16974 (2013).
28. S. Salvi *et al.*, Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11376–11381 (2007).
29. J. T. Brandenburg *et al.*, Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts. *PLoS Genet.* **13**, e1006666 (2017).
30. T. Degenkolbe *et al.*, Differential remodeling of the lipidome during cold acclimation in natural accessions of *Arabidopsis thaliana*. *Plant J.* **72**, 972–982 (2012).
31. R. Welti *et al.*, Profiling membrane lipids in plant stress responses. Role of phospholipase D alpha in freezing-induced lipid changes in *Arabidopsis*. *J. Biol. Chem.* **277**, 31994–32002 (2002).
32. D. V. Lynch, P. L. Steponkus, Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (*Secale cereale* l. cv Puma). *Plant Physiol.* **83**, 761–767 (1987).
33. J. Jouhet, Importance of the hexagonal lipid phase in biological membrane organization. *Front. Plant Sci.* **4**, 494 (2013).
34. J. R. Henriksen, T. L. Andresen, L. N. Feldborg, L. Duelund, J. H. Ipsen, Understanding detergent effects on lipid membranes: A model study of lysolipids. *Biophys. J.* **98**, 2199–2205 (2010).
35. L. Yan *et al.*, Parallels between natural selection in the cold-adapted crop-wild relative *Tripsacum dactyloides* and artificial selection in temperate adapted maize. *Plant J.* **99**, 965–977 (2019).
36. H. Susila *et al.*, Florigen sequestration in cellular membranes modulates temperature-responsive flowering. *Science* **373**, 1137–1142 (2021).
37. Y. Gu *et al.*, Biochemical and transcriptional regulation of membrane lipid metabolism in maize leaves under low temperature. *Front. Plant Sci.* **8**, 2053 (2017).
38. Y. Nakamura *et al.*, *Arabidopsis* florigen FT binds to diurnally oscillating phospholipids that accelerate flowering. *Nat. Commun.* **5**, 3553 (2014).
39. C. Riedelsheimer, Y. Brotman, M. Méret, A. E. Melchinger, L. Willmitzer, The maize leaf lipidome shows multilevel genetic control and high predictive value for agronomic traits. *Sci. Rep.* **3**, 2479 (2013).
40. T. Leinonen, R. J. S. McCairns, R. B. O'Hara, J. Merilä, Q(ST)-F(ST) comparisons: Evolutionary and ecological insights from genomic heterogeneity. *Nat. Rev. Genet.* **14**, 179–190 (2013).
41. G. M. Janzen *et al.*, Demonstration of local adaptation of maize landraces by reciprocal transplantation. *Evol. Appl.* **15**, 817–837 (2022).
42. J. E. Pool, D. T. Braun, J. B. Lack, Parallel evolution of cold tolerance within Drosophila melanogaster. *Mol. Biol. Evol.* **34**, 349–360 (2017).
43. S. Yeaman, A. C. Gerstein, K. A. Hodgins, M. C. Whitlock, Quantifying how constraints limit the diversity of viable routes to adaptation. *PLoS Genet.* **14**, e1007717 (2018).
44. J. A. Romero Navarro *et al.*, A study of allelic diversity underlying flowering-time adaptation in maize landraces. *Nat. Genet.* **49**, 476–480 (2017).
45. K. Luu, E. Bazin, M. G. B. Blum, *pcadapt*: An R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* **17**, 67–77 (2017).
46. L. Wang *et al.*, Metabolic interactions between the Lands cycle and the Kennedy pathway of glycerolipid synthesis in *Arabidopsis* developing seeds. *Plant Cell* **24**, 4652–4669 (2012).
47. G. S. Richmond, T. K. Smith, Phospholipases A_1. *Int. J. Mol. Sci.* **12**, 588–612 (2011).
48. K. Wang *et al.*, Two abscisic acid-responsive plastid lipase genes involved in jasmonic acid biosynthesis in *Arabidopsis thaliana*. *Plant Cell* **30**, 1006–1022 (2018).
49. S. Ishiguro, A. Kawai-Oda, J. Ueda, I. Nishida, K. Okada, The DEFECTIVE IN ANTHER DEHISCIENCE gene encodes a novel phospholipase A1 catalyzing the initial step of jasmonic acid biosynthesis, which synchronizes pollen maturation, anther dehiscence, and flower opening in *Arabidopsis*. *Plant Cell* **13**, 2191–2209 (2001).
50. S. Perez-Limón *et al.*, A B73 × Palomero Toluqueño mapping population reveals local adaptation in Mexican highland maize. *G3 (Bethesda)* **12**, jkab447 (2022).
51. K. W. Broman, H. Wu, S. Sen, G. A. Churchill, R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890 (2003).
52. S. B. Ryu, Phospholipid-derived signaling mediated by phospholipase A in plants. *Trends Plant Sci.* **9**, 229–235 (2004).
53. S. C. Stelpflug *et al.*, An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. *Plant Genome* **9** (2016).
54. O. Emanuelsson, H. Nielsen, G. von Heijne, ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978–984 (1999).
55. S. E. Jensen, E. S. Buckler, Pfam domain adaptation profiles reflect plant species' evolutionary history. *bioRxiv* [Preprint] (2021).
56. S. C. Potter *et al.*, HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204 (2018).
57. S. El-Gebali *et al.*, The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).
58. E. Cimen, S. E. Jensen, E. S. Buckler, Building a tRNA thermometer to estimate microbial adaptation to temperature. *Nucleic Acids Res.* **48**, 12004–12015 (2020).
59. S. E. Jensen, L. C. Johnson, T. Casstevens, E. S. Buckler, Predicting protein domain temperature adaptation across the prokaryote-eukaryote divide. *bioRxiv* [Preprint] (2021). https://doi.org/10.1101/2021.07.13.452245 (Accessed 1 July 2021).
60. D. E. Runcie, L. Crawford, Fast and flexible linear mixed models for genome-wide genetics. *PLoS Genet.* **15**, e1007978 (2019).
61. M. R. Woodhouse *et al.*, A pan-genomic approach to genome databases using maize as a model system. *BMC Plant Biol.* **21**, 385 (2021).
62. J. G. Wallace *et al.*, Association mapping across numerous traits reveals patterns of functional variation in maize. *PLoS Genet.* **10**, e1004845 (2014).
63. R. Bukowski *et al.*, Construction of the third generation *Zea mays* haplotype map. *Gigascience* **7**, 1–12 (2018).
64. W. A. Ricci *et al.*, Widespread long-range cis-regulatory elements in the maize genome. *Nat. Plants* **5**, 1237–1249 (2019).
65. K. A. G. Kremling *et al.*, Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. *Nature* **555**, 520–523 (2018).
66. Y. Nakamura *et al.*, High-resolution crystal structure of *Arabidopsis* FLOWERING LOCUS T illuminates its phospholipid-binding site in flowering. *iScience* **21**, 577–586 (2019).
67. Y. Nakamura, Plant phospholipid diversity: Emerging functions in metabolism and protein–lipid interactions. *Trends Plant Sci.* **22**, 1027–1040 (2017).
68. E. J. Veneklaas *et al.*, Opportunities for improving phosphorus-use efficiency in crop plants. *New Phytol.* **195**, 306–320 (2012).
69. A. Cruz-Ramírez *et al.*, The xipotl mutant of *Arabidopsis* reveals a critical role for phospholipid metabolism in root system development and epidermal cell integrity. *Plant Cell* **16**, 2020–2034 (2004).
70. H. Lambers *et al.*, Proteaceae from severely phosphorus-impoverished soils extensively replace phospholipids with galactolipids and sulfolipids during leaf development to achieve a high photosynthetic phosphorus-use-efficiency. *New Phytol.* **196**, 1098–1108 (2012).
71. S. R. Marla *et al.*, Comparative transcriptome and lipidome analyses reveal molecular chilling responses in chilling-tolerant sorghums. *Plant Genome* **10**, ●●● (2017).
72. L. Qu, Y. J. Chu, W. H. Lin, H. W. Xue, A secretory phospholipase D hydrolyzes phosphatidylcholine to suppress rice heading time. *PLoS Genet.* **17**, e1009905 (2021).
73. K. L. Mercer, H. Perales, Structure of local adaptation across the landscape: Flowering time and fitness in Mexican maize (*Zea mays* l. subsp. *mays*) landraces. *Genet. Resour. Crop Evol.* **66**, 27–45 (2019).
74. A. J. Waters *et al.*, Natural variation for gene expression responses to abiotic stress in maize. *Plant J.* **89**, 706–717 (2017).
75. A. K. Hottes *et al.*, Bacterial adaptation through loss of function. *PLoS Genet.* **9**, e1003617 (2013).
76. F. I. Khan *et al.*, The lid domain in lipases: Structural and functional determinant of enzymatic properties. *Front. Bioeng. Biotechnol.* **5**, 16 (2017).
77. Y. Kim, R. Nielsen, Linkage disequilibrium as a signature of selective sweeps. *Genetics* **167**, 1513–1524 (2004).
78. S. Khan, S. C. Rowe, F. G. Harmon, Coordination of the maize transcriptome by a conserved circadian clock. *BMC Plant Biol.* **10**, 126 (2010).
79. S. Maatta *et al.*, Levels of *Arabidopsis thaliana* leaf phosphatidic acids, phosphatidylserines, and most trienoate-containing polar lipid molecular species increase during the dark period of the diurnal cycle. *Front. Plant Sci.* **3**, 49 (2012).
80. E. Calfee *et al.*, Selective sorting of ancestral introgression in maize and teosinte along an elevational cline. *PLoS Genet.* **17**, e1009810 (2021).
81. C. M. Lazakis, V. Coneva, J. Colasanti, ZCN8 encodes a potential orthologue of *Arabidopsis* FT florigen that integrates both endogenous and photoperiod flowering signals in maize. *J. Exp. Bot.* **62**, 4833–4842 (2011).
82. T. Guo *et al.*, Optimal designs for genomic selection in hybrid crops. *Mol. Plant* **12**, 390–401 (2019).
83. W. Zhou, L. Wang, W. Zheng, W. Yao, MaizeSNPDB: A comprehensive database for efficient retrieve and analysis of SNPs among 1210 maize lines. *Comput. Struct. Biotechnol. J.* **17**, 1377–1383 (2019).
84. R. Rellán-Álvarez et al., Teosinte mexicana introgression modulates phosphatidylcholine levels and maize flowering time. GitHub. https://github.com/sawers-rellan-labs/High-PC1-paper. Accessed 22 February 2022.
85. A. C. Barnes *et al.*, Mass spectrometry data from "An adaptive teosinte mexicana introgression modulates phosphatidylcholine levels and is associated with maize flowering time." Metabolights Repository. www.ebi.ac.uk/metabolights/MTBLS5074www.ebi.ac.uk/metabolights/MTBLS5074. Accessed 23 June 2022.
86. M. C. Willcox, J. Burgueño, S. J. Hearne, D. J. Gates, Phenotypic data for "Single-gene resolution of locally adaptive genetic variation in Mexican maize." CIMMYT Research Data & Software Repository Network. https://hdl.handle.net/11529/10548233. Deposited 23 January 2020.