CENTRO DE INVESTIGACIÓN Y ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD IRAPUATO
UNIDAD DE GENÓMICA AVANZADA, LANGEBIO

“Patterns of lipid remodeling in maize highland adaptation”

This thesis is presented by

IBT. Karla Azucena Juárez Núñez

Directed by:

Dr. Rubén Rellán Álvarez

Guanajuato, Mexico
June 16th, 2017

CONTENT

CONTENT 1
ACKNOWLEDGEMENTS 2
ABSTRACT 3
RESUMEN 4
BACKGROUND 5
Maize domestication, migration, diversification and local adaptation 5
A single domestication event of maize followed by colonization and diversification in Mexican highlands 5
Highland conditions acting as selective forces on maize leading convergent adaptation 6
Strategies used by plants for adapting to cold temperatures and low phosphorus availability: lipid remodeling 8
Plant response to low temperature 8
Plant strategies to increase phosphorus availability 9
Natural variation in maize 11
HYPOTHESIS 13
GENERAL OBJECTIVE 14
SPECIFIC OBJECTIVES 14
MATERIALS AND METHODS 14
RESULTS AND DISCUSSION 25
Significantly different lipids between high and lowland accessions of the Diversity Panel 25
Genetic basis behind high and lowland biochemical phenotypes 31
CONCLUSION 40
PERSPECTIVES 41
REFERENCES 43

ACKNOWLEDGEMENTS

Al Colegio Nacional de Ciencia y Tecnología (CONACYT) por la beca 587909 (CVU:707666) que se me otorgó para poder cursar este programa de posgrado.

Un agradecimiento al Dr. Rubén Rellán Álvarez por ser un estupendo asesor y amigo; y por siempre tener la puerta de su oficina abierta.

Al Dr. Ruairidh Sawers por siempre escucharme y explicarme genética.

Al Dr. Sean Rovito por hacerme sufrir con preguntas que me hicieron aprender.

Al Dr. Robert Winkler por escuchar mis anécdotas personales y por empujarme a escribir esta tesis lo antes posible.

A todos mis amigos del laboratorio y del instituto, que siempre me recuerdan que existe vida fuera de estas paredes y aun así a veces prefiero quedarme aquí. Y por hacerme reír mucho, lo cual representa el ejercicio de cada día.

A todos mis amigos pre-maestría, que seguirán conmigo apoyándome y viceversa.

Y finalmente a mis padres, que siempre me han apoyado para que siga aprendiendo todo lo que quiera (aunque duden de mi talento musical). Eso es importante y eso es libertad.

ABSTRACT

After domestication from lowland teosinte (*Zea mays ssp. parviglumis*), maize colonized the Trans-Mexican volcanic belt highlands (>2500 masl), which are characterized –among other factors- by low temperatures and lower bioavailability of nutrients, especially phosphorus.

Under low phosphorus, plants increase the level of non-phosphated lipids while decreasing glycerophospholipids; membrane glycerophospholipids are replaced by galactolipids, releasing and transporting phosphate to young tissues, which are energetically more vulnerable to this deficit. Conversely, at low temperatures, plants increase glycerophospholipids and decrease non-phosphated lipids in order to maintain membrane fluidity.

We hypothesize that low temperature and phosphorus bioavailability were major selective forces during the process of maize adaptation to Mexican highlands. Recent population genetics data, in fact, point to this idea as genes with high differentiation between highland and lowland sites are related with phospholipid metabolism and phosphorus acquisition.

In order to identify the genetic architecture of phospholipid metabolism and its role on maize highland adaptation, we grew under highland conditions, on a field located in Metepec (2500 masl), a maize Recombinant Inbred Line (RIL) mapping population developed using Palomero Toluqueño (PT) (a Mexican highland landrace) and B73 an inbred line adapted to temperate conditions; and a landrace diversity panel composed of highland and lowland landraces from Mesoamerica and South America. We we sampled leaves from plants at V3 - V6 developmental stage and performed biochemical phenotyping (UPLC-QTOF MS/MS) on these samples and used the lipid intensities as a phenotype for QTL analysis. We identified a QTL peak at 8.6 Mb in chromosome 3 that seems to be explaining the conversion of phosphatidylcholine (PC) to lysophosphatidylcholine (LPC). At the marker in the QTL peak, PT alleles lead to a higher PC/LPC ratio than the B73 alleles, suggesting that this conversion of PC to LPC was selected for in the highland landrace Palomero Toluqueño. This biochemical phenotype seems to be characteristic of Mexican highland landraces.

RESUMEN

Después de la domesticación del teosinte de tierras bajas (*Zea mays ssp. parviglumis*), el maíz colonizó el cinturón trans-volcánico mexicano (>2500 masl), esta área se caracteriza -entre otros factores- por presentar bajas temperaturas y menor biodisponibilidad de nutrientes, especialmente fósforo.

A bajo fósforo, las plantas incrementan el nivel de lípidos no fosfatados mientras decrementan los fosfolípidos; los fosfolípidos de membrana son reemplazados por galactolípidos, liberando y transportando fosfato a tejidos jóvenes, los cuales son energéticamente más vulnerables ante este déficit. Por el contrario, a bajas temperaturas, las plantas incrementan la cantidad de fosfolípidos y decrementan lípidos no fosfatados con el fin de mantener la fluidez de la membrana.

Nosotros hipotetizamos que la baja temperatura y baja biodisponibilidad de fósforo fueron las principales fuerzas selectivas durante el proceso de adaptación del maíz a tierras altas mexicanas. Datos recientes de genética de poblaciones apuntan de hecho a esta idea, ya que existen genes con alta diferenciación entre tierras altas y tierras bajas, los cuales están relacionados con el metabolismo de fosfolípidos y la adquisición de fósforo. Una mejor comprensión de los cambios metabólicos adaptativos que ocurren en las plantas es muy necesario.

Con el fin de identificar la arquitectura genética del metabolismo fosfolipídico y su rol en la adaptación del maíz a tierras, crecimos una población de mapeo de tipo líneas puras recombinantes (Recombinant Inbred Lines, RILs) la cual se desarrolló empleando Palomero Toluqueño (PT) (una raza mexicana de tierras altas) y B73, siendo éste una línea pura adaptada a condiciones templadas.

Esta población de mapeo se creció a 2600 msnm (Metepec, Estado de México, México); muestreamos hojas de plantas en estadios de desarrollo entre V3 - V6. Posteriormente llevamos a cabo un fenotipado bioquímico (UPLC-QTOF MS/MS) para esas muestras y utilizamos las intensidades de las diferentes especies lipídicas como fenotipo para realizar análisis de QTL. Identificamos un pico de QTL a 8.5 Mb en el cromosoma 3, el cual parece explicar la conversión de fosfatidilcolina (PC) a lisofosfatidilcolina (LPC). En el marcador más cercano al pico del QTL, los alelos de PT conducen a tener una proporción mayor de PC/LPC que los alelos de B73, sugiriendo que esa conversión de PC a LPC fue seleccionada para la raza de tierras altas, Palomero Toluqueño.

BACKGROUND

Maize domestication, migration, diversification and local adaptation

A single domestication event of maize followed by colonization and diversification in Mexican highlands

The wild ancestor of maize, teosinte, belongs the *Poaceae* family (grasses) native from Central America, which is found as perennial or annual grasses. Within the family *Poaceae*, the genus *Zea* contains the species *Zea diploperennis*, *Zea luxurians*, *Zea perennis*, *Zea nicaraguensis* and *Zea mays*; but in general it can be categorized in two groups: *Luxuriantes* and *Zea* (Doebley and Iltis 1980; Iltis and Benz 2000). The *Zea* group is formed by the sole species *Zea mays*, and is composed of four subspecies: *ssp. mays*, *ssp. mexicana*, *ssp. parviglumis* and *ssp. huehuetenangensis* (Strable and Scanlon 2009). In the case of ssp. *mexicana*, there is evidence it diverged from ssp. *parviglumis*, approximately 60, 000 years ago (Ross-Ibarra, Tenaillon, and Gaut 2009).

The great diversity within maize lead to the idea that maize underwent multiple domestication events (Matsuoka et al. 2002). However, the study performed by Matsuoka and collaborators (Matsuoka et al. 2002) provided phylogenetic and population structure analysis, across 193 maize accessions collected from North America to South America and 67 annual teosintes (ssp. *mexicana* and ssp. *parviglumis*), which, along with archeological evidence supports that all maize arose from a single domestication event about 9,000 years ago. As in other studies, Matsuoka and collaborators place ssp. *parviglumis* as the progenitor of maize and due to this single domestication event the native region of teosinte *parviglumis*, central region of the Balsas River Drainage in southwest Mexico (Guerrero state), can be set as a candidate cradle of maize domestication (Matsuoka et al. 2002; Piperno and Flannery 2001).

After domestication, maize colonized the Mexican highlands, possibly aided by significant introgression of adaptive alleles from *Zea mays ssp. mexicana* (Hufford et al. 2013); hybridization in the Mexican highlands (>1,600 meters above sea level) between ssp. *mexicana* and maize is common. It has been estimated that 12% of the Mexican race Palomero Toluqueño (PT) comes from ssp. *mexicana* via introgression.

Highland conditions acting as selective forces on maize leading convergent adaptation

Maize migration from Mexican lowlands to highlands came with the challenge to adapt to several contrasting biotic and abiotic characteristics. Highland conditions can be seen as important selective forces driving maize local adaptation.

In contrast to lowlands, highlands are characterized by drier conditions, lower temperatures, higher radiation levels, lower air pressure, less human land use (Körner 2007), and lower phosphorus availability due to volcanic soils composition. The latter is characteristic of the central highlands in the Trans-Volcanic belt in Mexico (Krasilnikov et al. 2013).

Teosinte *mexicana* is distributed along highlands of northern and central Mexico, and is adapted to drier (~300 mm average less in annual precipitation) and cooler (average anual temperatura 7 °C lower) conditions than lowlands. Highland areas are defined by elevations higher than 1600 m (Hufford et al. 2013, 2012; Matsuoka et al. 2002; Piperno and Flannery 2001).

Teosinte *mexicana* and teosinte *parviglumis* exhibit some contrasting morphological differences. The leaf sheaths of *parviglumis* are green and glabrous while *mexicana* develops dense leaf sheath pubescence (macrohairs) accompanied by dark red-purple pigmentation on the stem. Teosinte *mexicana* also produces larger spikelets and seeds, and fewer tassel branches than *parviglumis* (Doebley and Iltis 1980). Those traits have been tentatively suggested as key elements for adapting to lower temperatures and higher radiation levels (Lauter et al. 2004).

After the maize domestication, lowland maize from southern Mexico and Guatemala also moved to lowlands in South America (~6000 years before present) where it adapted again to highland conditions (~4000 years BP), at the same time that maize colonized and diversified in highlands (Matsuoka et al. 2002; Takuno et al. 2015; Hufford et al. 2013).

Adaptation to the Mexican highlands was possibly aided by teosinte *mexicana* introgression, while in South American highlands (Andes) no inter-fertile wild *Zea* species are found (Matsuoka et al. 2002; Hufford et al. 2013).

Mesoamerican and South American highlands landraces (open pollinated varieties) face very similar environmental gradients (Takuno et al. 2015). Besides their different origin they share some phenotypic similarities also present in *mexicana* (Doebley 1984). For example, the red-purple pigmentation (Wilkes 1977; Wellhausen 1957) caused by anthocyanin presence that is thought to be an adaptation to high radiation levels (Liakopoulos et al. 2006;

Steyn et al. 2002). This trait is absent in lowland varieties and could be pointing at adaptive introgression (Hufford et al., 2013) and convergent adaptation between highlands (Hanson et al. 1996; Takuno et al. 2015).

On the other hand, other traits differ between Mesoamerican and South American highlands such as macrohairs; Mexican highland maize exhibits a very dense pubescence while South American maize is glabrous.

Figure 1. From left to right: Lowland teosinte and lowland maize followed by highland teosinte and highland maize (Courtesy of Ruairidh Sawers).

Convergent adaptation occurs when species or populations exposed to similar environments develop similar phenotypes (Arendt and Reznick 2008; Elmer and Meyer 2011; Wood, Burke, and Rieseberg 2005).

Several evolutionary genetic studies have established the basis for three main routes in which convergent evolution can arise. The first one establishes that when identical mutations appear independently, and are fixed through natural selection in different genetic backgrounds (species or populations), convergent phenotypes can emerge under similar conditions (Currat et al. 2002; Kwiatkowski 2005). Second, convergence occurs when different mutations arise within a single locus associated with a specific trait (Kovach et al. 2009). And finally, similar traits may arise by repeated natural selection acting on standing genetic variation derived from an ancestral population (Colosimo et al. 2005).

Both Mexican highland maize and teosinte *mexicana*, present dense macrohairs, red-purple stem pigmentation (Wilkes 1977), tassel branching, ear husk number (Brewbaker 2014) and high anthocyanin pigmentation as biochemical defenses against ultraviolet radiation (Casati and Walbot 2005). These mechanisms can be considered as a convergent phenotype for highland conditions.

Mesoamerican and South American highland landraces also share the red-purple stem pigmentation. Furthermore, genetic studies have affirmed that maize varieties from both highland subcontinental areas (Mesoamerica and South America) have arisen independently from their respectives lowland populations (Matsuoka et al. 2002; van Heerwaarden et al. 2011; Vigouroux et al. 2008), suggesting adaptive convergence.

With these examples in which the environmental conditions seem to be playing a key role for the expression of certain phenotypes, it is also important to take into account that the external conditions are not the only player. Phenotypic differences can also be due to a distinct genetic background (Fukunaga et al. 2005; Ross-Ibarra, Tenaillon, and Gaut 2009).

Strategies used by plants for adapting to cold temperatures and low phosphorus availability: lipid remodeling

When talking of local adaptation, in addition to the morphological and biochemical changes that maize might present we should also consider that there are specific molecular processes and metabolic profiles that could being involved in some of these adaptative mechanisms.

Plant response to low temperature

As mentioned before, during highland colonization and diversification maize had to adapt to cold and low phosphorus conditions through different strategies. Those conditions can be hypothesized as main selective forces.

Plant membrane integrity is a very important factor for proper growth and development (Lin, Liu, and Nakamura 2015); the severe membrane damage caused by cold temperatures is the main cause of freezing injury (Degenkolbe et al. 2012) through freeze-induced osmotic contraction and dehydration. Compositional changes within the membrane lipid species after cold exposure can be considered as a freeze-induced mechanism to modulate the cryobehavior of membranes (D. V. Lynch and Steponkus 1987).

Within the structural-membrane changes induced by low temperatures, we find the increase

of unsaturated fatty acids and phospholipids content (Nishida and Murata 1996). Through different studies in *Arabidopsis*, it has been shown that there is a lipid remodeling after cold acclimation of different accessions. During cold conditions, glycerophospholipids like phosphatidylcholine (PC) and phosphatidylethanolamine (PE) increase (D. V. Lynch and Steponkus 1987), while nos-phosphorous lipids such as monogalactosyl diacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG) tend to decrease (Degenkolbe et al. 2012).

Plant strategies to increase phosphorus availability

Low bioavailability of phosphorus in soils is a major factor limiting growth and yield of crop plants. Although it is widely distributed in soils, it is commonly present as inorganic orthophosphate (PO_4^{-3}), a very reactive form which tends to join cationic elements such as aluminium, iron, magnesium and calcium, forming insoluble phosphate species unavailable for plant uptake (Hinsinger 2001; Yadav and Verma 2012). A class of soil present in the Trans-Mexican Volcanic Belt and in several areas of South American highlands are andosols, which have low phosphorus bioavailability and are found in all altitudes and climates but are mainly present in volcanic areas (Krasilnikov et al. 2013).

Plants phosphorus sensing under phosphate starvation conditions can lead to two different responses to increase Pi availability: local response or a systemic response.

A local response consists of a precise regulation of root system architecture during a plant's lifetime, dynamically shaping the roots according to changes in the local environmental conditions. Under Pi deficiency increasing root surface exploration area is a common local response (Rellán-Álvarez et al. 2015; J. P. Lynch and Brown 2012).

The systemic response controls phosphate homeostasis, meaning whole-organism responses for distribution and redistribution of phosphorus where it is needed within the plant (Veneklaas et al. 2012; L. Li et al. 2007). Phosphate homeostasis can be further classified in three main strategies: release and uptake of soil phosphorus upregulation of phosphate transporters and Pi recycling, (Veneklaas et al. 2012; Sulpice et al. 2014; Peret et al. 2014).

Phosphorus is found in plants as free inorganic orthophosphate (Pi) or forming phosphate esters. Plants can recycle phosphate from the main P-pools: nucleic acids (RNA), glycerophospholipids, P-esters and Pi. RNA is the largest organic P pool, accounting for about 40% - 60% of phosphorus in leaves with rRNA as the major compound within this reservoir. On the other hand, glycerophospholipids represent around 24% of phosphorus in

leaves (Veneklaas et al. 2012). In membranes, glycerophospholipids can be replaced by non-phosphorus lipids like sulfolipids and galactolipids (Veneklaas et al. 2012; Michaud et al. 2016). For the formation of the two mitochondrial membranes, a high phospholipid exchange between endoplasmic reticulum (ER) and mitochondrion is required in plants. Under Pi-starvation there is an active lipid transfer from plastids to the mitochondrion. In *Arabidopsis thaliana* the transmembrane lipoprotein (MTL) complex has a protein domain named AtMic60. This subunit is located in both mitochondrial membranes and is responsible for the export of PE from mitochondria and galactolipids (such as DGDG) import from plastids under phosphate deficiency (Michaud et al. 2016).

Upregulating purple acid phosphatases (PAPs) and RNases are common mechanisms to hydrolyze glycerophospholipids and RNA (Tran, Hurley, and Plaxton 2010/7; Plaxton and Tran 2011). Phosphate is released and transported to different tissues but Pi recycling mainly takes place in chloroplast membranes (Michaud et al. 2016). The direction of this transport is established out from senescing tissues to younger ones, which have higher energetic requirements to synthesize all components de *novo* (H. Lambers et al. 2015; Veneklaas et al. 2012). Within the angiosperm family *Proteaceae*, we find the species *Hakea prostata*, native to southwestern Australia. This species inhabits one of the most phosphorus-impoverished soils (Hopper 2009; Hans Lambers 2014), where it thrives through diverse strategies like P remobilization from mature organs to younger ones and constant lipid remodeling during leaf development. *Proteaceae* replaces glycerophospholipids in a very selective way, maintaining high glycerophospholipids amount in young leaves and replace them with galactolipids in mature leaves (Lambers et al. 2015). During leaf development PC and PE are present in higher level than phosphatidylglycerol (PG), probably because glycerophospholipids are basic for membrane integrity during cell division and also forming intermediates involved in developmental processes (H. Lambers et al. 2015; Gagne and Clark 2010) whilst PG is a key structural and functional element of photosystems (Hagio et al. 2000; Domonkos et al. 2004).

Metabolomic profiling analyses have shown that *Arabidopsis thaliana* experiences lipidome remodeling when exposed to low phosphorus supply. A decrease in glycerophospholipid levels is the principal adaptative mechanism detected under Pi-starving conditions. This process is coupled with the increase of glycerolipids (non-phosphorus lipids) (Okazaki et al. 2013).

Extreme temperatures are also a keys factor defining plant distribution and local yield. High temperatures produce unsustainable dehydration, while cold inhibits plant metabolism and therefore growth. Fatty acids unsaturation is an important method to maintain membrane

fluidity, in order to prevent liquid crystalline transitions, which can damage plant structure (Degenkolbe et al. 2012; Nishida and Murata 1996)

When *Arabidopsis* is exposed to low temperatures lipid profile changes by accumulating storage lipids and decreasing saturation. When *Arabidopsis thaliana* accessions are exposed to cold temperaturas, levels of PC and PE tend to increase while non-phosphorus lipids like monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG) decrease (Degenkolbe et al. 2012).

Another abiotic factor affecting plant development is day length. Sunlight, whose availability is correlated with day lenght, is required for photosynthesis. Furthermore, light also heats up the soil, increasing processes like evaporation. Hence, this is one of the most important environmental factors exerting selection pressure on organisms. Several studies of different plant species, have showed different strategies of photoperiod and temperature sensing to adjust flowering time to seasonal changes (Bohlenius et al. 2006; Nakamura et al. 2014; Song et al. 2015; Song, Ito, and Imaizumi 2013) in *Arabidopsis* the genes *CONSTANS* (CO) and *FLOWERING LOCUS T* (FT) are key elements regulating flowering. During long days CO mRNA has a its maximum accumulation at the end of the day, while in short days the accumulation occurs at night. Nevertheless, CO protein is labile to darkness, so the only net accumulation happens during long days. CO induces transcription of the FT gene in the companion cells of phloem; here FT mRNA moves to shoot apex where the FT protein induces flowers formation (Bohlenius et al. 2006; Nakamura et al. 2014).

Glycerophospholipids are also important signaling molecules (Lin, Liu, and Nakamura 2015; Nakamura et al. 2014). In plants, specifically *Arabidopsis thaliana*, *in vitro* membrane-lipid overlay assays using His-tagged FT showed how this protein binds exclusively to PC. FT binds different preferentially low unsaturated phosphatidylcholine (PC) species during the day stimulating early flowering. Using a transgenic approach targeting enzyme CTP:phosphorylethanolamine cytidylyltransferase (PECT1) which is key for PE biosynthesis, leaded to the partial suppression of PECT1 and an increase of the PC/PE ratios and significant reduction of flowering time. (Nakamura et al. 2014).

Natural variation in maize

An immense natural phenotypic variation is present in maize, which is key for dissecting complex genetic traits. Quantitative traits, show a continuous range of phenotypic variation and are often under the control of several genes and/or chromosomal regions known as

Quantitative Trait Loci (QTL) (Meihls, Kaur, and Jander 2012; Flint-Garcia et al. 2005; Griffiths et al. 2015) which contribute in different amounts to the phenotype, also by interacting with other genetic and environmental factors (Peleman et al. 2005).

Linkage and association mapping are frequently used to map QTL in plants. Linkage mapping is the most traditional way to identify QTL and requires developing a mapping population. The population is usually the result of a biparental cross between contrasting genotypes and/or phenotypes; some common biparental populations are Recombinant Inbred Lines (RILs), that is an homozygous population, Near Isogenic Lines (NILs), that are backcross inbred lines (Keurentjes et al. 2007) and backcrosses generated by crossing an F1 with one of the parentals.

Individuals within an inbred line population are genetically equal due to their common origin, same parent(s) (Griffiths et al. 2015), and individuals are homozygous for all the alleles. This allows us to analyze the effect of each haplotype on different traits. Inbred lines are created when specific strains/ecotypes of a given biological system are self-fertilized or sib-mated for multiple generations so they become homozygous for the most alleles along the genome.

A common population for QTL mapping is Recombinant Inbred Lines (RILs), which are generated by an initial cross between two parents. The consequent F1 population is auto-pollinated for multiple times for fixing all alleles in each descendent. In selfing species, to increase the number of meiotic events, individuals from F1 can be sib-mated to generate a F2 and then self-pollinated as many times as necessary to reach a high percentage of homozygosity; ending with around a equal proportion of each parent (Pollard 2012; Broman 2005; Eichten et al. 2011).

In RILs populations all individuals must be genotyped with genetic markers (SNPs, SSRs, RFLps, etc) covering all the genome, often with Genotyping By Sequencing (GBS). Individuals are phenotyped to analyze the phenotypic data via linkage mapping so certain markers can be linked to the trait (Flint-Garcia et al. 2005; Griffiths et al. 2015).

A limitation of the linkage mapping is its resolution (precision of QTL position), which is completely dependent on the size of the population and on the few recombination events within these biparental populations, giving a resolution between 10-30 cM (Alpert and Tanksley 1996; Flint-Garcia et al. 2005). On the other hand, association mapping or linkage disequilibrium (LD) mapping allows a higher resolution which makes possible to identify a single polymorphism within a candidate gene for a giving phenotype (Flint-Garcia, Thornsberry, and Buckler 2003; Riedelsheimer et al. 2013; Flint-Garcia et al. 2005; Sorkheh et al. 2008).

Linkage disequilibrium is defined as “the nonrandom association of alleles at different loci. This allelic association can be explained based on the shared history of mutation and recombination within certain alleles, drift, nonrandom mating, etc. A high LD occurs between neighboring bases, because due to their proximity the recombination level between is very low, so they are gained or lost as a unit. Therefore, markers utilized in these studies are the ones associated to the trait of interest through LD, making it tractable to study the genetic basis of the trait (Flint-Garcia, Thornsberry, and Buckler 2003; Flint-Garcia et al. 2005; Griffiths et al. 2015).

Association mapping makes possible the evaluation of genetic diversity within natural populations, without the need to develop a mapping population before (Flint-Garcia, Thornsberry, and Buckler 2003). Beyond this, analyzing genetic diversity within a sample of nonrelatives makes the mapping resolution higher thanks to the increase of recombination events (Ewens and Spielman 2001; Kruglyak 1999; Lander and Schork 1994). Currently, the most commonly used approach to do this is Genome Wide Association Studies (GWAS).

However, association mapping also present weak points like the population size is tipically small and the corresponding LD, nevertheless, the most significant one is presence of population structure (Knowler et al. 1988; Flint-Garcia et al. 2005). Low LD requires more markers to cover the genome (Tenaillon et al. 2001; Flint-Garcia et al. 2005) but it will also allow a higher mapping resolution and these numbers will also differ from outcrossing and auto-pollinated species. In outcrossing species LD decays rapidly; in maize landrace LD decays within 1 kb and 2 kb for diverse inbred lines. For selfing species like rice, LD decays within 100 kb (Xu et al. 2017/4).

Despite the different research aims or the mapping populations needed for each of the above mentioned methods, both, linkage mapping and association mapping approaches allow investigation of the genetic architecture of traits (number of genes and contribution of each) and environmental factors that have influenced the traits.

HYPOTHESIS

As low temperature and low phosphorous bioavailability lead to lipid remodeling, increasing glycerophospholipids and decreasing glycerolipids or the opposite way respectively. I hypothesize that the reorganization of phospholipid metabolism has had a major role in the adaptation of maize to the highlands of Mexico.

GENERAL OBJECTIVE

Characterize the remodeling of highland lipid profiles involved in phospholipid metabolism in order to shed light on the possible role of glycerophospholipids in maize adaptation to highlands.

SPECIFIC OBJECTIVES

- Characterize the lipid profile of 100 B73 X Palomero Toluqueño Recombinant Inbred Lines grown in a common garden under highland conditions.
- Perform a QTL analysis with RILs population.
- Characterize a 120 landrace panel composed of highland and lowland Mesoamerican and South American lines grown in a common garden maize under highland conditions.

MATERIALS AND METHODS

Plant Materials

We sampled 120 landraces (2 plants each) which were evaluated in highland conditions (Metepec field). From those 120 landraces, 60 were from Mesoamerica and 60 from South America and within each subcontinental group the accessions were paired selected, 30 were collected from highlands and the rest from lowland areas.

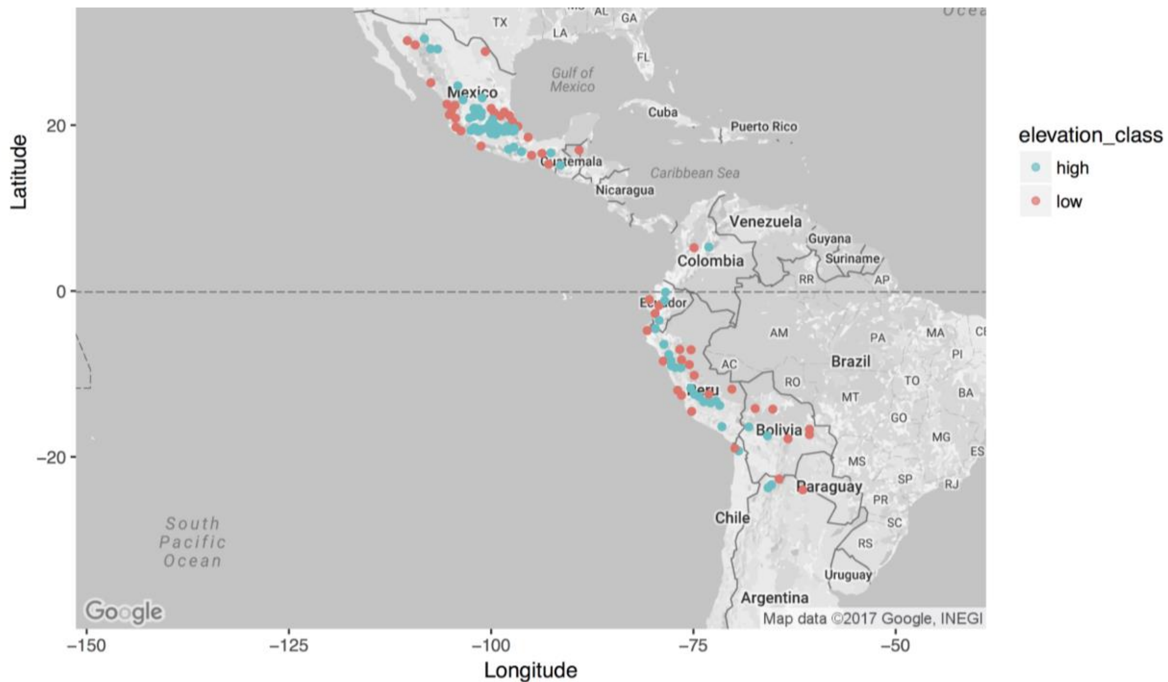

Figure 3. Diversity panel (DP) accessions map. This selection was performed by Daniel Runcie at UC Davis.

Diversity Panel (DP) formed by 60 Mesoamerican landraces and 60 South American landraces. Within each subcontinent, 30 landraces come from highlands, >1600 meters above sea level (masl) and 30 come from lowlands (<1000 masl). The landraces are selected as pairs, so for each highland landrace a lowland landrace located at the same latitude (±0.5 º) was picked.

We also used 100 Recombinant Inbred Lines (RILs) derived from a cross between B73 (inbred line lowland conditions-like adapted) and PT (Mexican Highland race) (Figure 4).

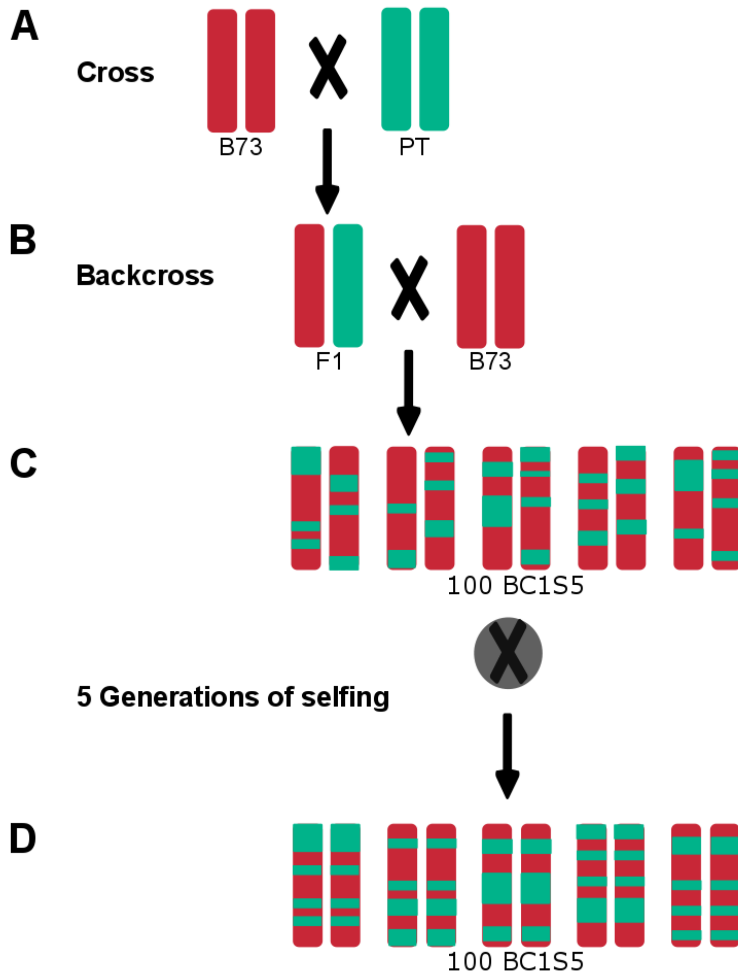

Figure 4. Developing of the B73xPT RIL population (BC1S5). After the 5 self-pollinations of the 100 individuals, we obtained 100 Recombinant Inbred Lines with an estimated 25% of PT on B73 background. This mapping population has been developed by the laboratory of Ruairidh Sawers at Langebio.

Growth Conditions

The following environmental conditions are based on the National Meteorological Service (smn.cna.gob.mx/) database and on historical records from the National Institute of Forestry, Agriculture and Livestock Research (www.inifap.gob.mx/).

Metepec field is localized (19°13'28.7"N 99°32'51.6"W) within the Trans-Mexican Volcanic Belt. The experimental field is 2610 masl, the range of average monthly temperatures along the year go from 5 °C to 21.5 °C with an average annual of 13.6 °C. Its average annual precipitation is around 809 mm. Growing degree days (GDD) is a very useful indicator for correlating temperature with growth and development of any crop.

$$GDD = (Tmax + Tmin)/2 - Tbase$$

The base temperature (T_{base}) for maize is set at 10 °C because it does not grow below that. Taking into account the reported data by INIFAP, the Metepec field has a GDD value of 15, while a lowland site, would have a GDD of around 5, meaning that a site with higher GDD (highlands) will have faster growth.

Both mapping populations, RILs and Diversity Panel, were planted here on 5 April, 2016 and were sampled for lipid analysis by 9 June 2016, when the plants were around V4 -V6 developmental stage.

Making an estimation with the actual data from INIFAP from sowing date to harvest, the average GDD for Metepec was 5.52 and the total GDD (over the length of the growing season) was 364.25. On the other hand, during this period the average daily precipitation (mm) was 1.6 while the total precipitation was 96.5 mm.

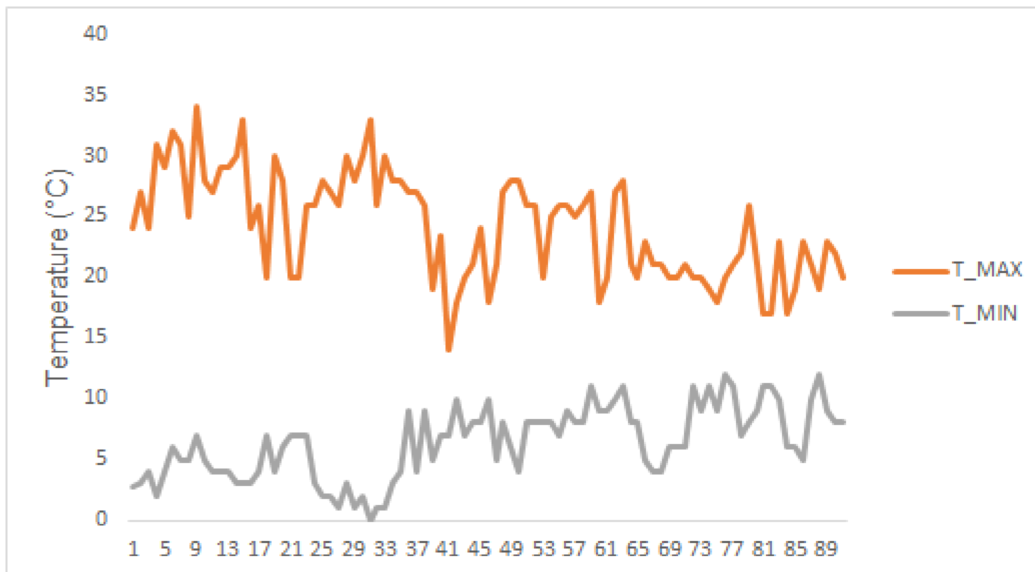

Figure 6. Maximum and minimum temperature from sowing to sampling date.

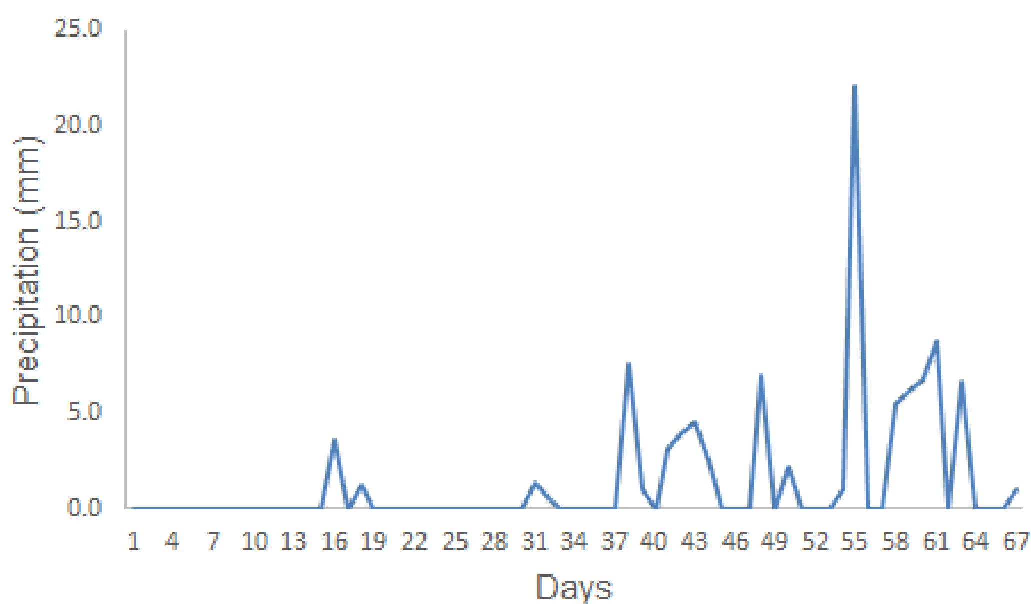

Figure 7. Precipitation level (mm) from sowing to sampling date.

Lipid Extraction

Leaf tissue was collected when the plant was, in average, in the V3 - V6 stage. We tested using 50 mg wet weight leaf tissue for lipidome profiling, as this amount was deemed appropriate for *Arabidopsis* (Degenkolbe et al. 2012). Frozen material was homogenized in a tissue grinder Retsch (Haan, Germany) during 40 seconds at 30 1/s. After grinding all tubes were stored in liquid nitrogen.

Lipid extraction used was reported by Matyash and collaborators (Matyash et al. 2008). First 225 µL of cold methanol (MeOH), previously prepared with a Quality Control (QC) mix, was added to each sample, always keeping MeOH on ice during the extraction. Each sample was vortexed for 10 seconds, keeping the rest of materials on ice. Then 750 µL of cold methyl tert-butyl ether (MTBE) were added, keeping MTBE on ice along all the process. Again each sample had to be vortexed for 10 seconds, followed by shaking each for 6 minutes at 4°C in the orbital mixer. After that, 188 µL of LC/MS grade water at room temperature (RT) were added, vortexing all the samples for 20 seconds.

In order to separate the phases, all samples were centrifuged for 2 min at 14, 000 rcf (12,

300 rpm). The upper phase is the organic one and it was possible to remove around 700 µL of supernatant, so it was split into two aliquots of 350 µL, one serving as back up and for the preparation of pools. Finally, samples were dried using a speed vacuum concentration system and stored at room temperatura (RT) but protected against light. Otherwise, the lower phase (inorganic one) was also dried and stored for future HILIC analysis.

Lipid Profiling

Samples resuspension for analysis

Dry samples were resuspended in 110 µL of MeOH-Toluene 90:10 (with CUDA 50 ng/mL). The rack of samples was vortexed at low speed for 20 seconds and then all tubes were sonicated at RT for 5 min. Aliquots of 50 µL per sample were transferred into an insert within an amber glass vial.

Pre-run sequence

A standard calibration was used before each run of around 300 sample batch. For ESI (+) positive mode the profile of the calibration standard and the intensity of ions of hexamethoxyphosphazene m/z 322.0481, Hexakis (2,2-difluoroethoxy) phosphazene m/z 622.0290 and Hexakis (2,2,3,3-tetrafluoropropoxy) phosphazene m/z 922.0098 were checked, which must be higher than 400k, 500k and 500k respectively. When the intensity of any of the ions was not high enough, the ion sources was cleaned and the instrument tuning repeated.

Reference ions m/z 121.0509 and m/z 922.0098 also had to be checked, and must have intensities around 5-20k. The intensity of reference ions could be adjusted by modulating flow rates.

Before start running samples a new column was set, it is recommended to change the column (Acquity UPLC C18 1.7µm 2.1x100 mm Column) (Waters; Milford, MA, USA) every 1000 samples. After the column change, the new one must be purged for 5 min to take air out. The UPLC column is coupled to a VanGuard pre-column (Acquity UPLC C18 1.7µm 2.1x5 mm Pre-column Waters®), which is replaced after ~330 sample injections. Six "no sample injections" were injected at the beginning of each run to condition the column, followed by ten samples, one pool (made out of the mix of the second aliquot of all the samples contained per UPLC plate) and one blank.

Optimization of lipid extraction and UPLC-QTOF MS/MS lipid analysis

To determine the best mode for lipid identification, 50 mg of tissue from 8 samples of PT, 8 of B73 and 10 RILs samples grown in Metepec during 2015, were extracted and run under postivive and negative ESI modes on Agilent® Ultra-Performance Liquid Chromatography and Quadrupole Time of Flight Mass Spectrometry (UPLC-QTOF MS/MS).

The positive mode showed the presence of 24 different lipids, within the species were CEs, DGs, LPCs, PCs, PE and TGs. On the other side, under negative mode only 16 lipid species were identified, the species were Fas and PCs.

According to the diversity and amount of lipid species identified under each mode, the best method for this project was posItive mode (Table 1). The UPLC-QTOF MS/MS utilized was Agilent 6530 ESI (+) injects into UPLC-QTOF MS/MS 1.67 µL per sample, the running time per sample is 15 min.

Table 1. UPLC-QTOF MS/MS lipid analysis under positive and negative mode, respectively.

Positive mode		Negative mode	
		Fatty acids	14
Cholesterol esters	2		
Diacylglycerols	3		
Lysophosphatidylcholines	2		
Phosphatidylcholines	10	Phosphatidylcholines	2
Phosphatidylethanolamine	1		
Triacylglycerols	6		
Unknowns	1214	Unknowns	939
TOTAL	**1238**	**TOTAL**	**955**

Metabolic data analysis

Retention time correction

In Agilent MassHunter Qualitative Analysis®, B.06.00 version, raw metabolic data was used to extract EIC (extracted-ion chromatogram) for each internal standard within the blanks. The m/z of internal standards (ISTDs) used to obtain the EIC were taken from a ESI (+) library acquired on Agilent 6530 UPLC-QTOFMS (Table 2).

Table 2. M/z and Retention time for the internal standards within the QC mix used during lipid extraction. Values acquired on ESI (+)

ISTD	m/z	Method RT (min)
Cuda	341.2799	0.774
Sphing	286.274	1.040
LPE	466.2928	1.346
LPC	510.3554	1.827
MG	345.2999	3.038
DG 18	416.3365	3.162
PC	636.4599	3.502
DG 12	474.4152	4.248
Chol	376.3953	4.787
SM	717.5905	5.053
Cer	552.535	5.948
PE	720.5538	6.263
TG	869.8323	11.000
CE	724.6966	11.727

Once obtained the EIC for each ISTD, the “current” values for each retention time were used

to calculate a delta of retention time (Current batch retention time – Method Retention time = delta retention time). This delta value (X axis) and the method retention time value (Y axis) were graphed in Excel® to get a polynomial formula out of the trend line.

MS1 lipids database created in Excel®, by the laboratory of Dr. Oliver Fiehn, includes MS1 m/z values and retention times for 409 lipid species and 92 common unidentified unknowns. The equation obtained in the previous step is utilized for calculating a new delta retention time value (Y) through substitution of the X variable with the retention time values included within the MS1 lipids database. This delta retention time value and the retention time registered for each lipidic species in the library, are used to estimate a "current" retention time value for the run (new delta retention time + library retention time = Current retention time).

In this way a new MS1 lipids library is created using the current retention times values after the retention time correction process. This new library is upload in MS-DIAL for identification of lipid species along the run.

Peak identification with MS-DIAL

Raw files from samples are in .d Agilent proprietary format. In order to use MS-DIAL we used Reifycs Abf Converter® to get all the files in Analysis Base File (.abf) format.

MS-DIAL was used under ESI (LC/MS, LC/MS/MS) mode and under positive mode. The method type was Data dependent MS/MS. Data type (MS1) was set as centroid data and for target omics option Lipidomics was selected.

Data collection was acquired with a retention time range from 0.5 – 12.5 min and with a mass range from 50 – 1200 Da. For peak detection, mínimum peak width was set at 5 and minimum peak height at 5000 in amplitude.

Identification was performed through MS-DIAL based on the library generated after the retention time correction, which must be uploaded as .txt format. The retention time tolerance was 0.1 min, accurate mass tolerance was established as 0.01 Da and the identification score cutoff was set as 85%.

For the Adduct ion setting the molecular species selected were [M+H]+, [M+NH4]+, [M+Na]+, [2M+H]+, [2M+NHA]+ and [2M+Na]+. Under Alignment settings, a random sample was selected as reference file. Peak count filter value was defined as 30%, meaning that

aligned peaks which are less than this parameter will be removed from the final result.

Finally, result export format was selected as .txt and spectra type as centroid, the latter being much lighter than profile data.

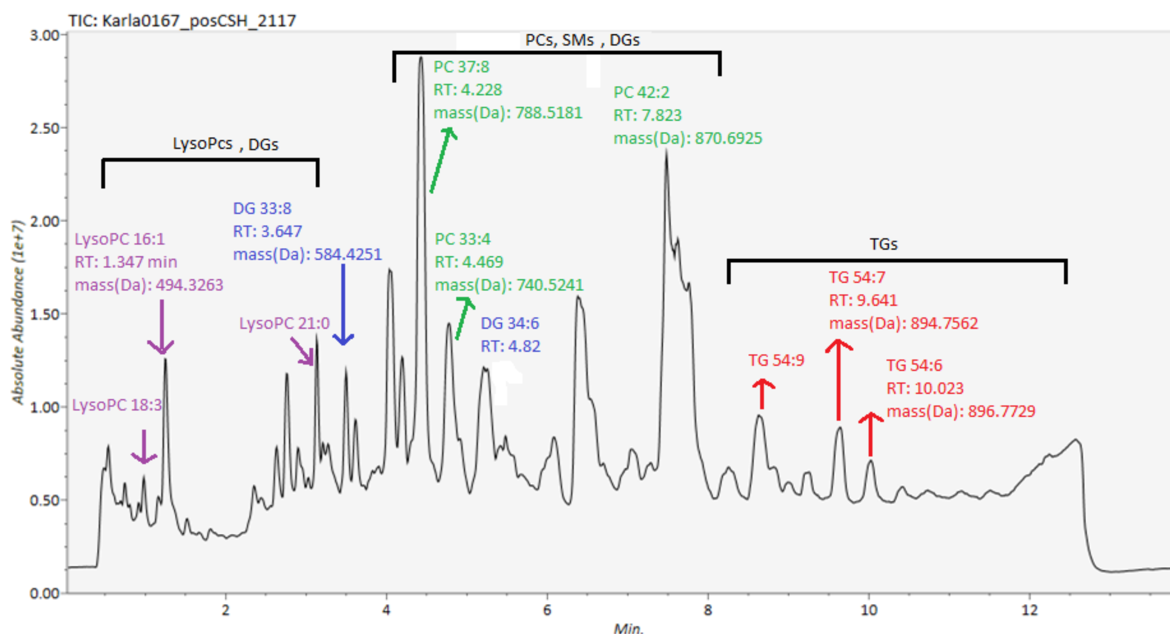

Figure 8. Total Ion Chromatogram (TIC) obtained from MS-DIAL. As shown here, in maize leaf lipid profiles, the separation of lipids occurs in an interval of 0 to 13 min. From 0 to 4 min we find lysophosphatidylcholines (LPCs) and diacylglycerols (DGs), then from 4 to 8 min PCs, DGs and sphingomyelins (SMs) appear and finally from 8 min to 13 min all the species of triacylglycerides (TGs) separate. This TIC corresponds to the LANMLR17B036 RIL sample.

Data processing and normalization

The results from MS-DIAL as .txt format are saved as “Raw” sheet. Contents from “Raw” sheet were pasted in another sheet named Sorted, in which we sorted internal standards (ISTDs) to the top. In this case the only ISTD used along all samples, pools and blanks was CUDA. Three columns were added for calculation of Blank average (Blank avg = average (Blanks)), Maximum of samples (max) and Fold change (FC =max/ Blank avg). All non-ISTDs features were sorted by FC largest to smallest.

All these data were copied into a new sheet named as Reduced. Here, all the non-ISTDs features with a FC < 10 were deleted. All data were then sorted by smallest to largest m/z in order to identify features with the same m/z along different retention times. This was corroborated by looking at the extracted ion chromatogram (EIC) for repeated m/z in MS-DIAL, noise peaks look like straight lines on the graph time vs m/z and the EIC looks as a undefined peak (Supplementary Figure 1). Noise peaks detected were deleted from Reduce sheet.

Reduced data were copied and pasted in a new sheet named as QC (Quality Control), where columns from Blank average to %QC were deleted and QC average (QC avg), QC standard deviation (QC Stdev) and %CV (Coefficient of Variation = Stdev/average x 100) were added.

QC data were sorted by %CV largest to smallest. Features with %CV > 30% and peak height > 5000 were deleted. QC data were copied and pasted into a new sheet “To be processed”, where blanks and columns MS/MS included through Fragment presence % were deleted; also QC avg to %CV were deleted. These data were copied into a new workbook and saved as .txt format for processing with MS-FLO (http://msflo.fiehnlab.ucdavis.edu/#/).

MS-FLO performs duplicate peak removal, isotope detection and adduct joining. The output has a .txt format, so the data were copied into a new sheet named as Processed. The latter were copied and pasted in a new sheet “Duplicates”. Sorted by duplicate flag column, if m/z was different to <0.01 and RT different to <0.05 min, the features were checked in MS-DIAL, when two features are duplicates they appear under the same peak in the EIC. For duplicate features from one peak, the maximum of peak height was calculated and the original values were deleted. Having this, a new “Isotopes” sheet was created with the filtered data, duplicate flags were erased. This was sorted by isotope flag and case by case checked if [M+H+1] had a $R^2 \geq 0.7$ and peak height ratio (PHR) <1, when those requirements are fulfilled peak heights were plotted (scatter plot), for pairs with linear correlation, the row with the bigger average m/z was deleted. Isotopes data were copied to a new sheet, “adducts”, where isotope flags were eliminated. Features were sorted by metabolite name and content for unknowns was cleared. Features were then sorted by adduct flag and checked if m/z differs for each match. Adduct rows were combined for the same compound (m/z1_m/z2, RT1_RT2), summing peak heights. Files with the data processing of the four batches we had can be found as Supplementary Files 1 to 4.

Data Normalization

First, for normalizing the data mTIC method was used. mTIC stands for the sum of all identified metabolites and its formula is:

$$Metabolite_{ij,normalized} = \frac{metabolite_{ij,raw}}{mTIC_j * mTIC_{average}}$$

where i=metabolite and j:=sample.

For mTIC, we calculated the sum of all identified compounds intensities per sample, followed by the total average of the identified compound intensities sum, of all samples. Then, each intensity per compound per sample, is divided by the product of that sample intensities sum and the fixed total average of samples sum. This gave us normalized peak intensities, nevertheless, as we ran the samples in the UPLC-QTOF MS/MS for a month, the differences between samples might be also affected by machine sensitivity, tuning and other parameters, so additional normalizations (batch normalization) must be performed.

Statistical analysis

Some of the statistical analysis were performed using MetaboAnalyst (http://www.metaboanalyst.ca/) software version 3.0. Here, no sample normalization was used because mTIC normalization was already applied before. Logarithm (Log) transformation was applied and for data scaling Pareto was used. Data transformation and scaling make individual features more comparable. Principal component analysis (PCA) was performed with the among sample variance-covariance matrix of lipid intensities.

RESULTS AND DISCUSSION

Significantly different lipids between high and lowland accessions of the Diversity Panel

After the generation of lipid profiles for the Diversity Panel population, the data were analyzed in order to identify any differences between Mesoamerican and South American landraces. Among the 120 landraces, we did not observe a clear difference between Mesoamerican and South American landraces (Figure 9). This may be due to the inclusion of both subcontinental highland and lowland landraces.

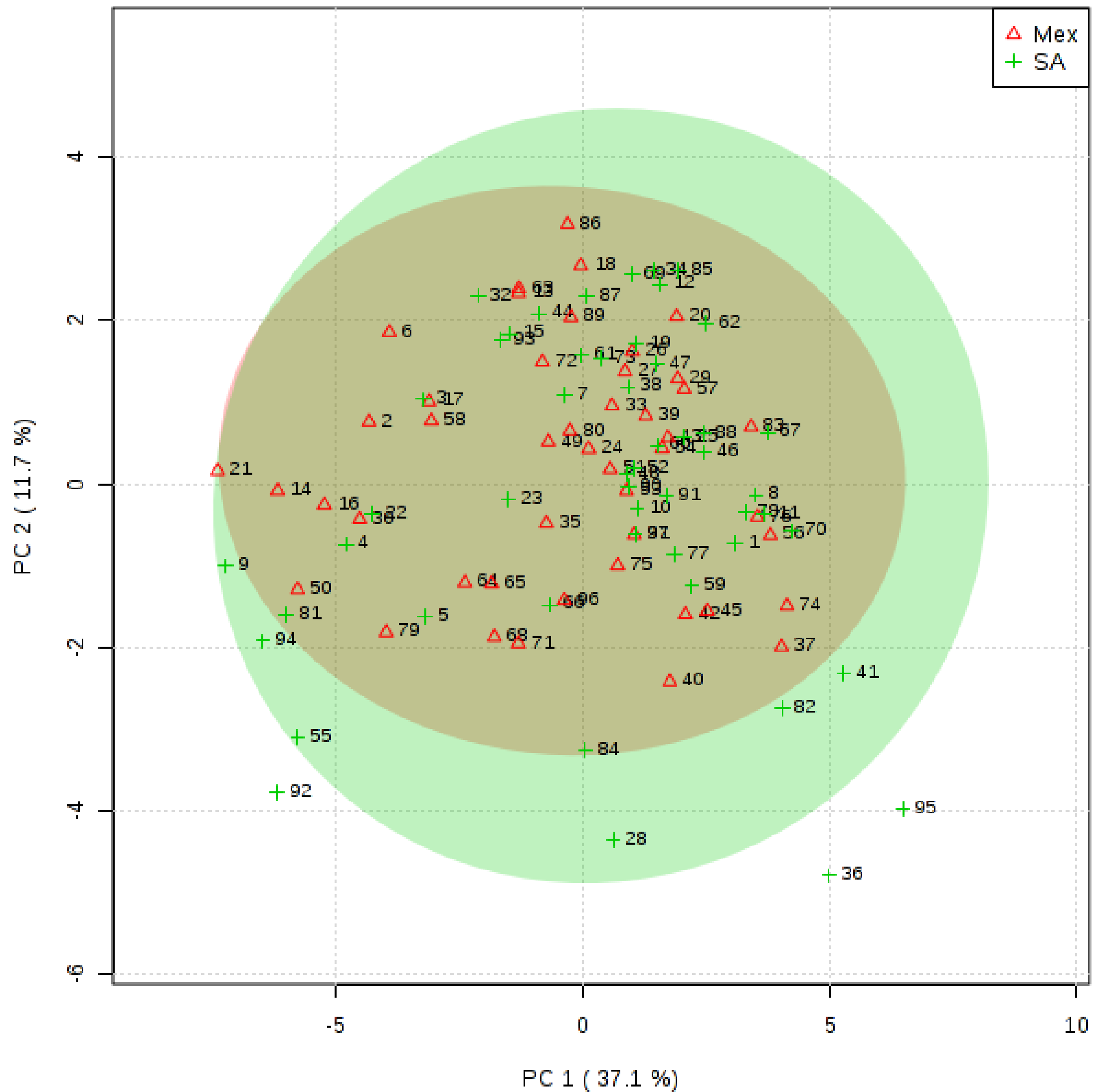

Figure 9. Mesoamerican and South American landraces planted in Metepec field. In this figure Mesoamerican (Mex) and South American (SA) landraces lipid profiles, including highland and lowland ones, are shown. At first sight there is no an obvious separation between the groups, however, for PCA we need to analyze the correlation between the variables and the principal components (Supplementary Table 4).

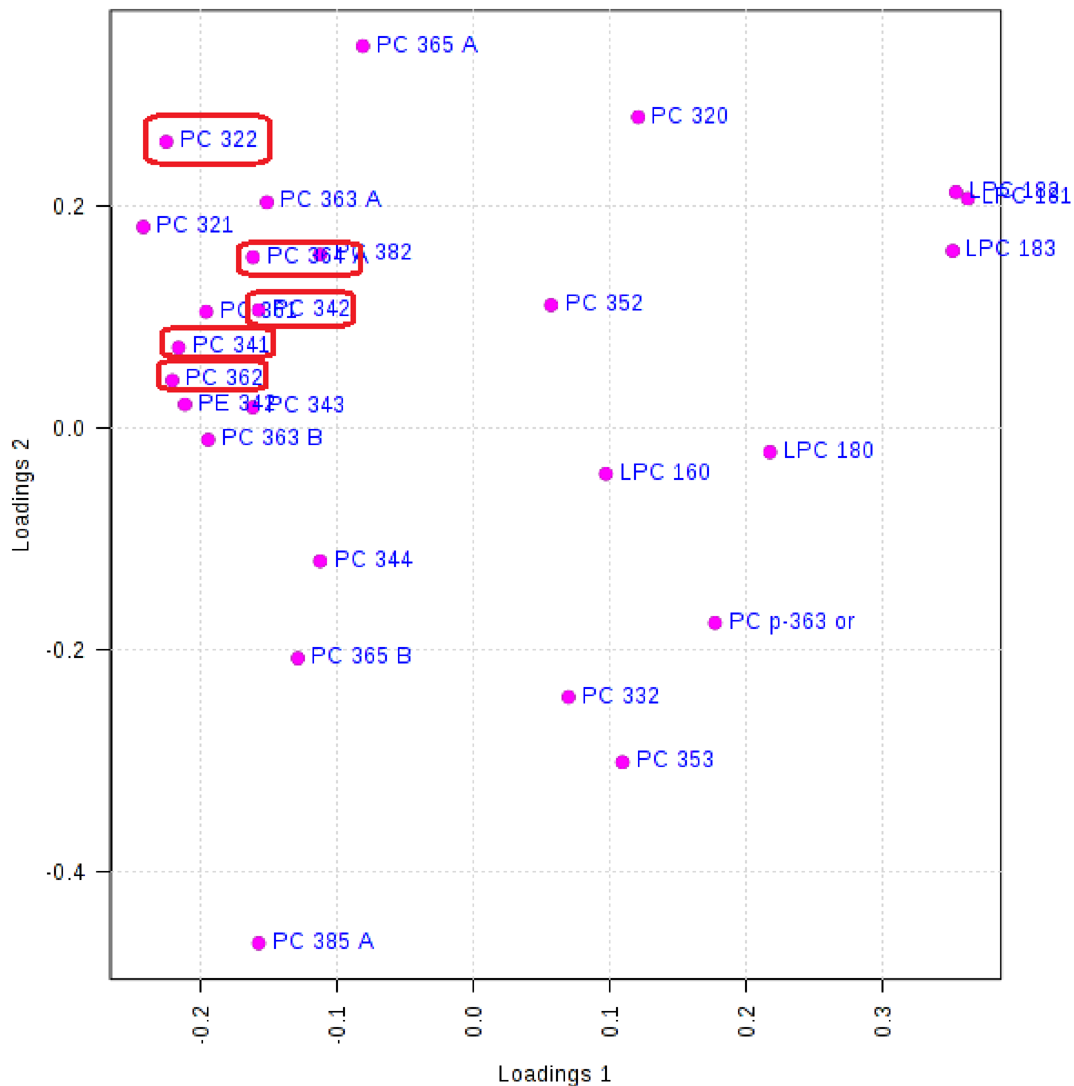

Figure 10. Loadings of PCA for Mesoamerican and South American landraces planted in Metepec field.

In Figure 10 LPCs and PCs load on PC1 in different directions. PC1 is strongly correlated with all LPC species, meaning LPCs vary together while PCs and PE vary together as well (See also Supplementary Table 4).

In order to be able to better appreciate the difference between high and lowland landraces, we separately analyzed Mesoamerican and South American landraces.

PCAs for Mesoamerican high and lowland landraces and for South American high and lowland landraces (Figure 11) show near-complete overlap between both categories within each subcontinental area; this might be because few compounds, compared to total amount of lipid species in the profiles, are significantly different between high and lowland landraces.

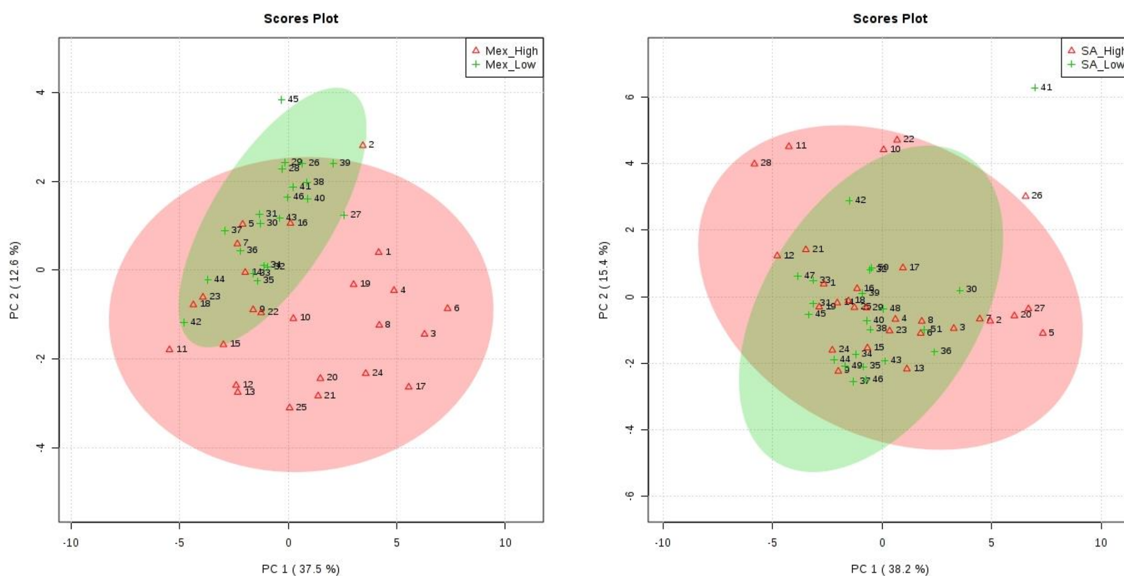

Figure 11. Mesoamerican and South American landraces, respectively, grouped in high and lowland. Highland in red and Lowland landraces in green.

In both subcontinental areas, highland landrace lipid diversity includes almost all of the lowland landraces diversity.

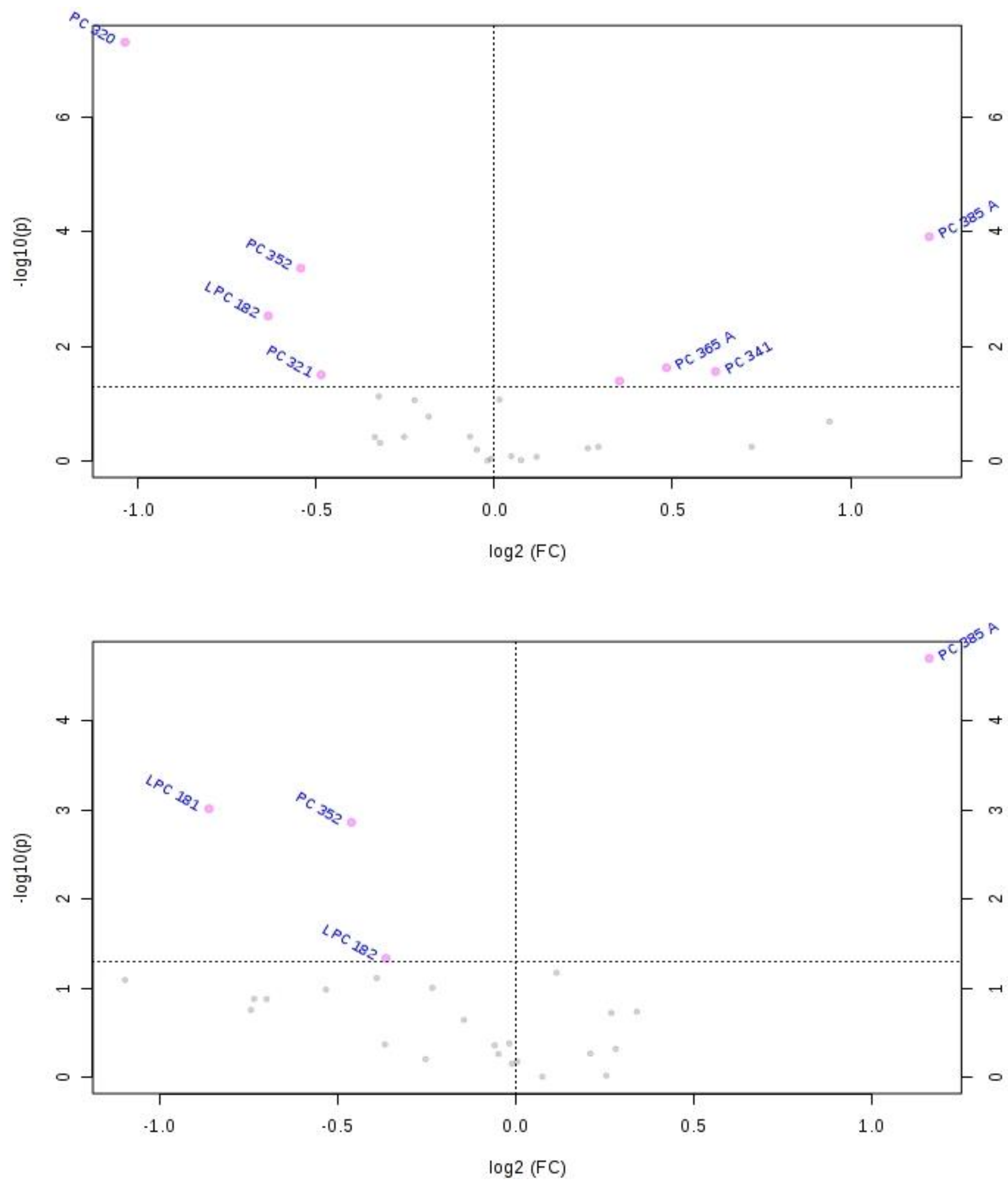

Figure 12. Significantly different compounds between Mesoamerican highland and lowland landraces and South American ones.

In volcano plots on Figure 12 we can see which compounds are significantly different (red spots) between highland and lowland landraces from Mesoamerica (upper figure) and South America (lower figure), with a fold change (FC) of 1 and a p<0.05. Comparison type for each volcano plot: MA_Highland/MA_lowland and SA_highland/SA_lowland respectively.

The compounds which are significantly different between high and lowland groups, in both subcontinental regions, belong to two lipid species: phosphatidylcholines (PCs) and lysophosphatidylcholines (LPCs).
PCs and LPCs could be involved in local adaptation and also, these lipid species could represent a way of molecular convergence between Mesoamerican high and lowland landraces and South American high and lowland landraces.

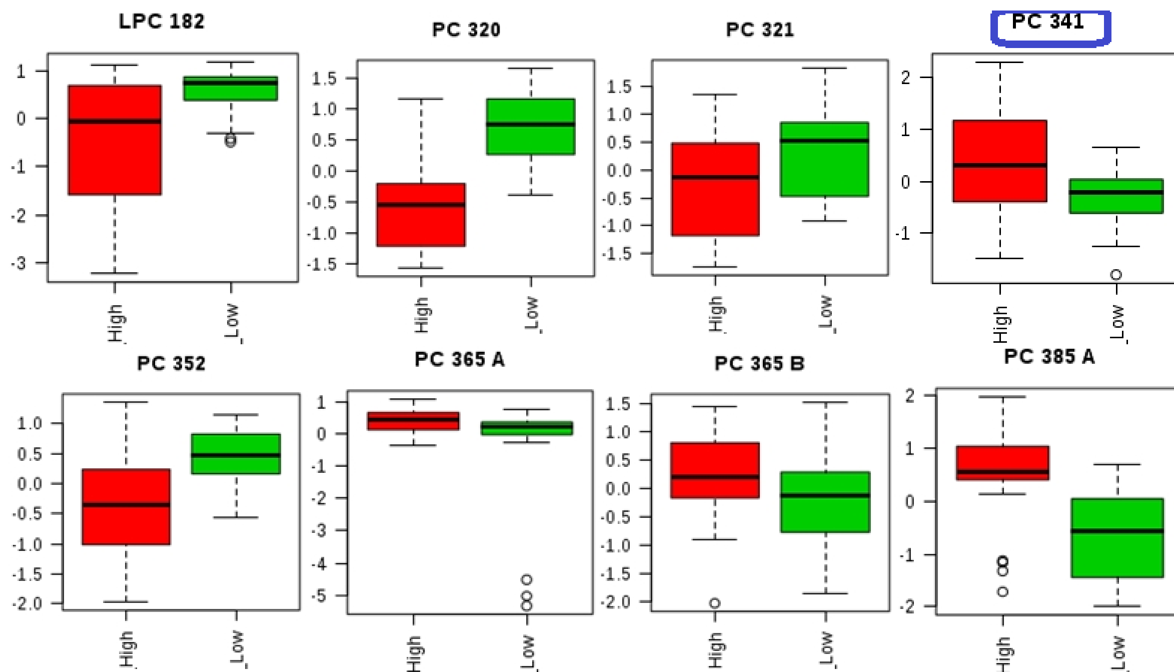

Figure 13. Differentiated lipid species between Mesoamerican highland and lowland landraces.
Box plots show the difference of lipid abundance between Mesoamerican high and lowlands, of the significantly different compounds shown in upper part Figure 13. The Fold change and the p value for each species is shown in Supplementary Table 5.

We can say highland and lowland landraces exhibited a very similar lipid profile when

evaluated in Metepec field, however, there are a few lipid species which are significantly different between both. These species, shown in Figure 12, might have an important role in the local adaptation of high and/or lowland landraces.

Genetic basis behind high and lowland biochemical phenotypes

To shed light on the importance of PCs and LPCs for maize local adaptation, the B73xPT RIL mapping population was used.

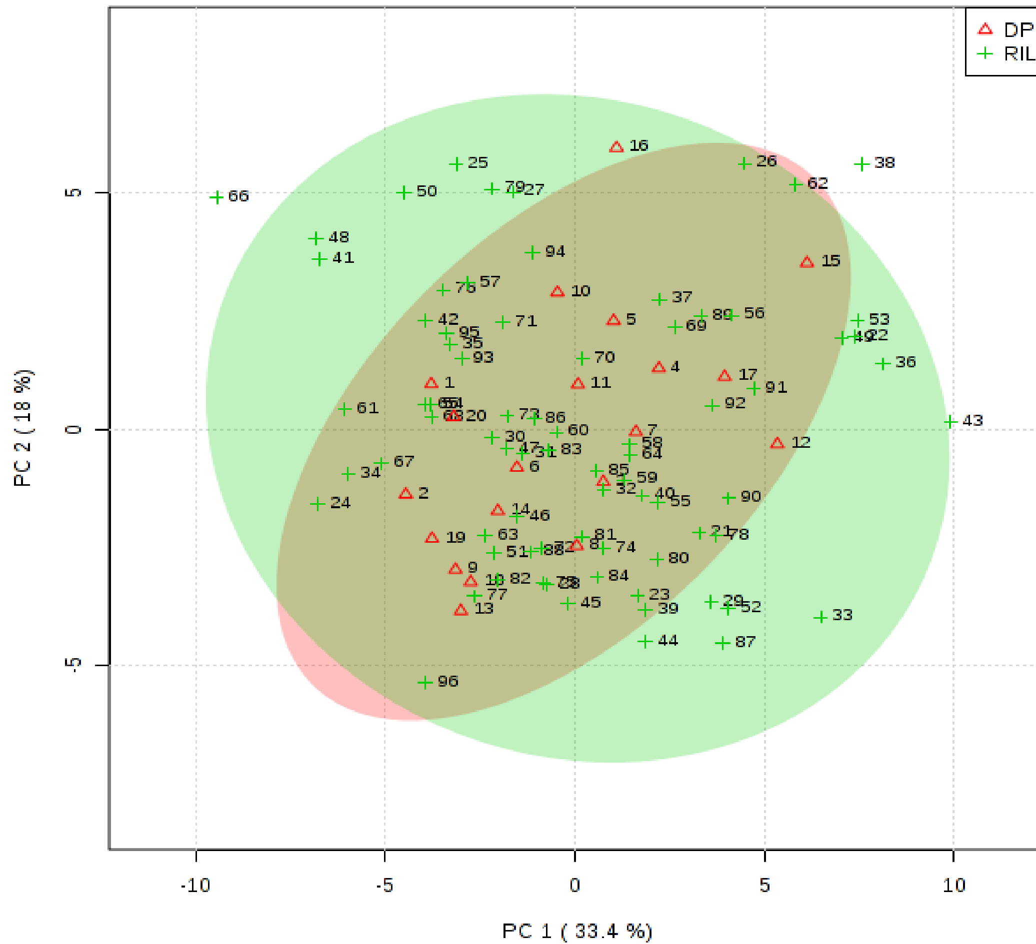

Figure 14. RILs cover the whole lipidomic space of the Diversity Panel landraces. RILs lipid profile is diverse enough to capture all the lipid diversity within the DP landraces.

The advantages of using a biparental population are the breaking up of population structure and the reassortment of alleles by the recombination, producing a range of phenotypic values. In this case, B73 is an inbred line adapted to lowland-like conditions while PT is a mexican highland-adapted landrace. Using all the lipid species identified by UPLC-QTOF MS/MS, the RILs parentals were shown to be completely different for all the lipid species intensities (Figure 15).

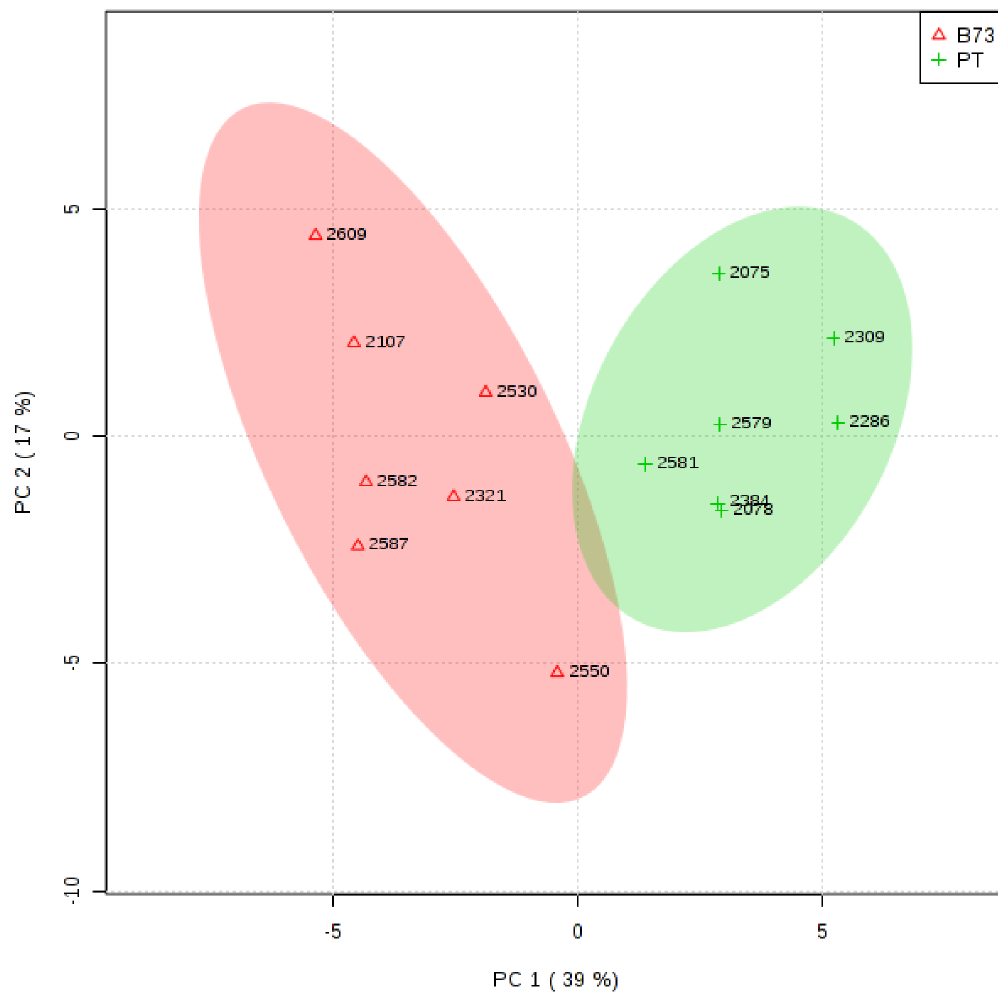

Figure 15. RIL population parentals planted in Metepec field. In Figure 15, PT in green and B72 in red are clearly separated among all the identified compounds. A volcano plot showing significantly different lipid species, their fold change and p-values are shown in Supplementary Figure 3 and Table 6.

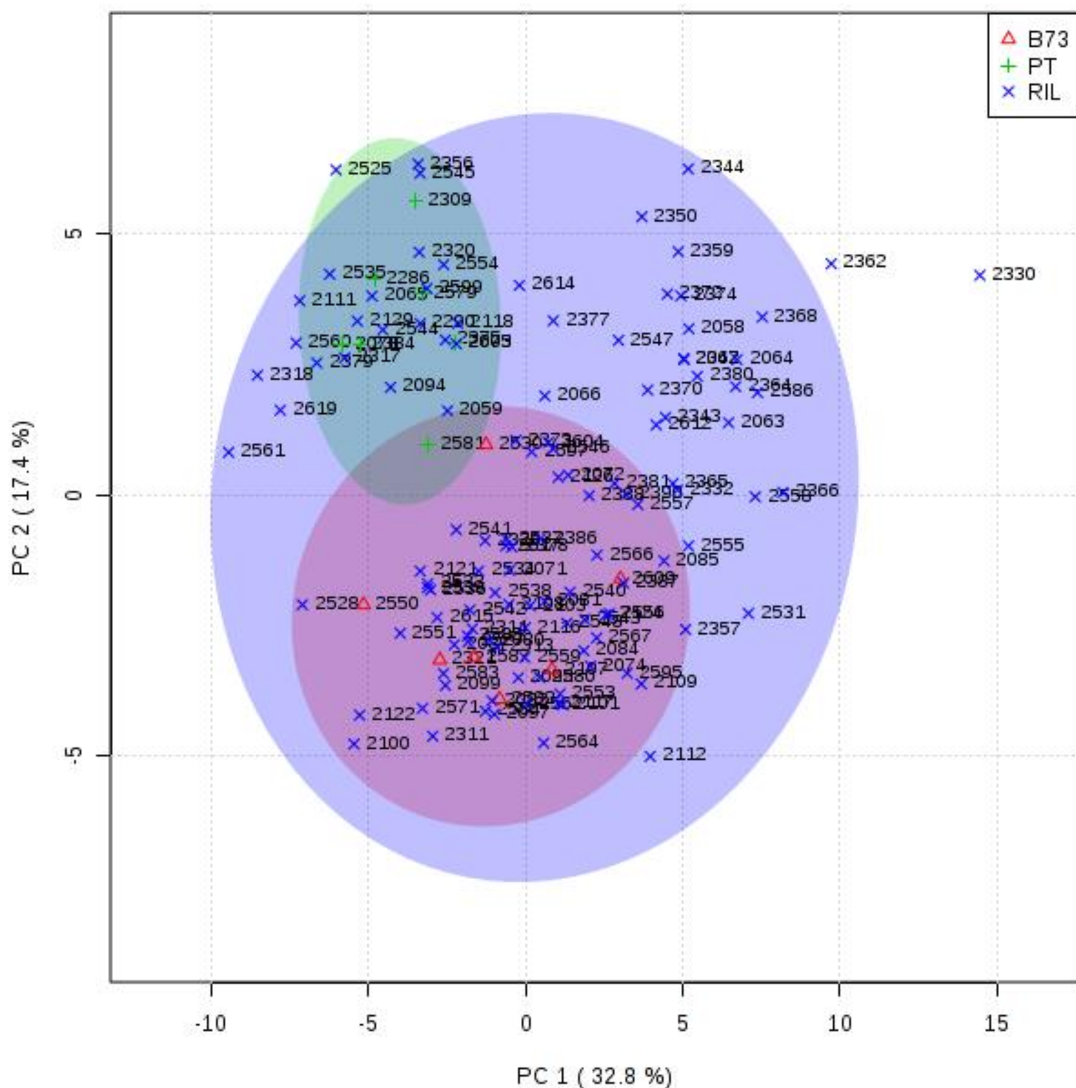

Figure 16. RILs diversity capture the diversity of its parents. B73 in red, PT in green and RILs in blue.

Nevertheless, when including RIL population into the PCA, RILs showed a highly diverse set of lipids, diverse enough to cover the parental lipid profiles.

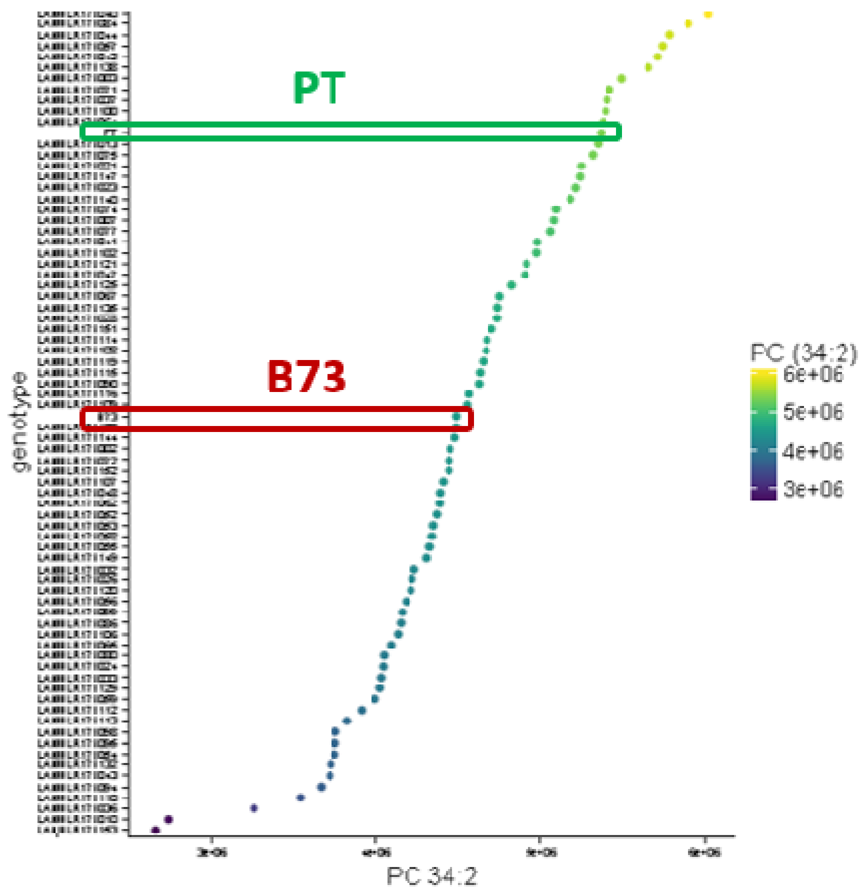

Figure 17. Transgressive segregation along RILs for PC (34:2) (phosphatidylcholine with 34 carbon and 2 unsaturations). Here the PC (34:2) intensities for all the RILs, B73 (in red) and PT (in green).

As shown in Figure 14, the BC1S5 B73xPT RILs have a transgressive segregation, meaning that along the population extreme and intermediate phenotypes are found, respect to the parental phenotypes. This is corroborated in Figure 17, where the intensities for PC (34:2) varying in a range from extreme RILs to intermediate and parental-like phenotypes. For this lipid, PT has a greater intensity than B73, so very probably RILs having PT alleles, rather than B73 alleles, will exhibit a higher level of this metabolite (Li et al. 2016). This transgressive segregation occurs with all the identified lipid species.

In a study conducted by Li et al. (2016), they also showed how through the development of a RIL population, from two rice subspecies, it was possible to generate a metabolite and

growth transgressive segregation. This makes it possible to consider in the creation of novel metabolic phenotypes by the combination of different germplasms (Li et al. 2016).

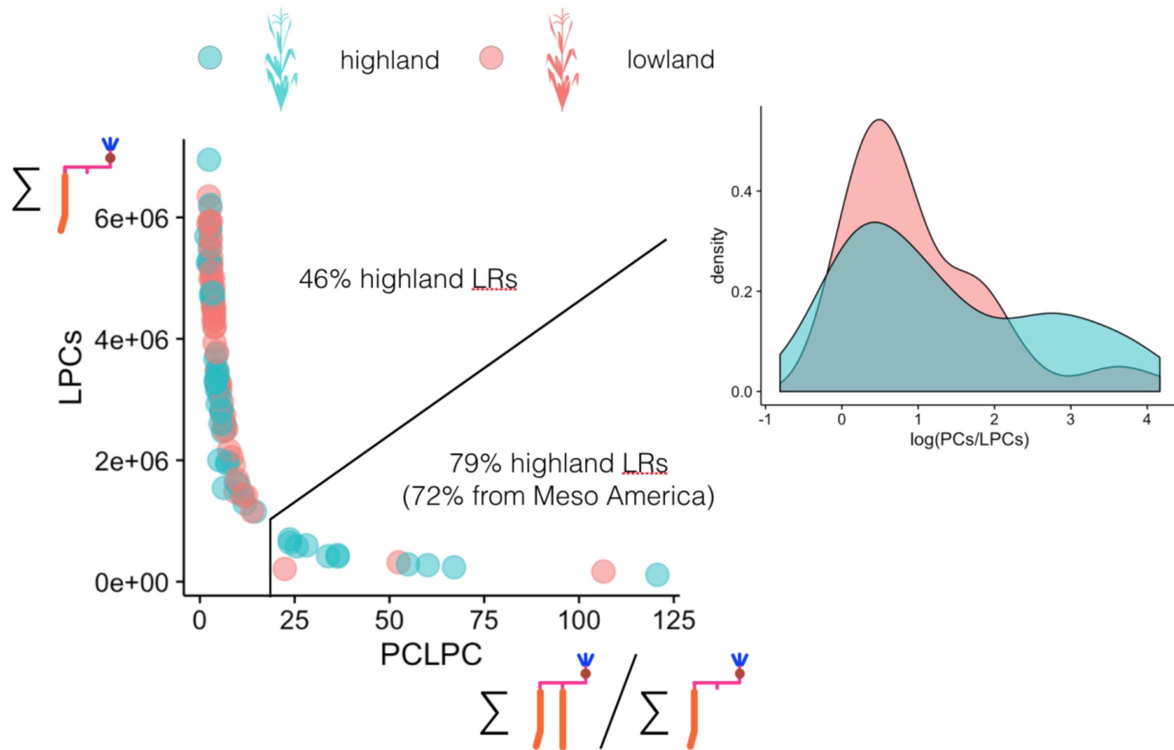

Figure 18. In the Diversity Panel high PC/LPC ratios are due to low levels of LPCs (Courtesy of Rubén Rellán). Correlation between sum of LPC species and the ratio of the sum of PC and the sum of LPC species shows that landraces with high ratios are due to low levels of LPC species Landraces with high PC/LPC ratios are mainly Mexican highland landraces.

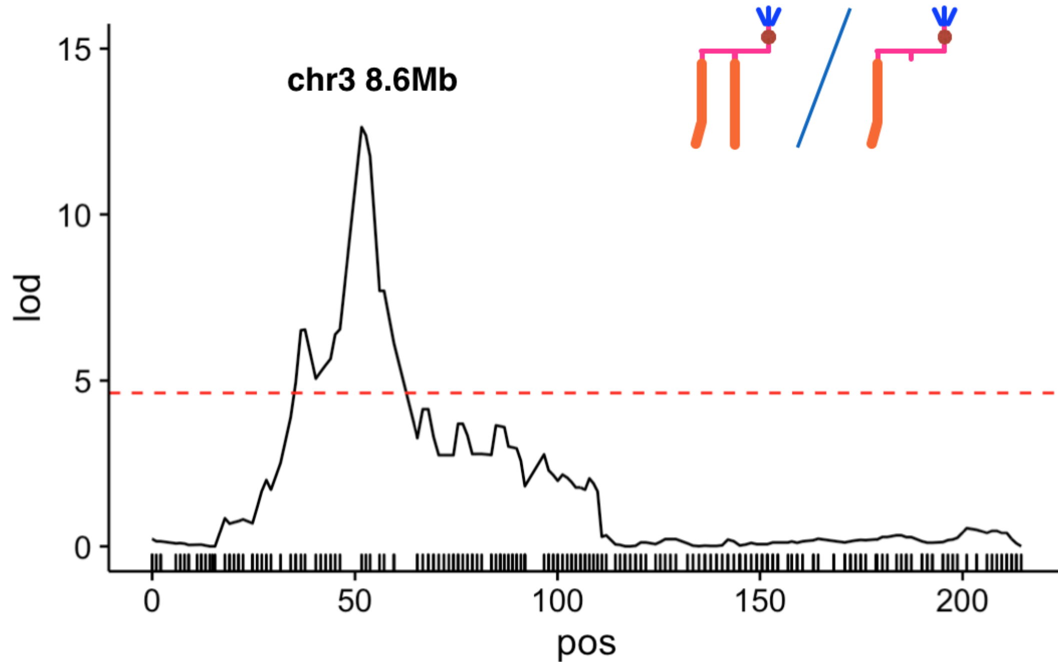

Figure 19. QTL peak on chromosome 3 found for the sum of PCs was also found on the same region of Chr3 for sum of LPCs.
This QTL peak for PCs and LPCs represent solid piece of evidence that there is a genetic basis behind these phenotypes that could be involved in local adaptation. Further studies are needed. For this QTL analysis, the intensities of PCs and LPCs were used as the phenotype. All RILs were previously genotyped by GBS.

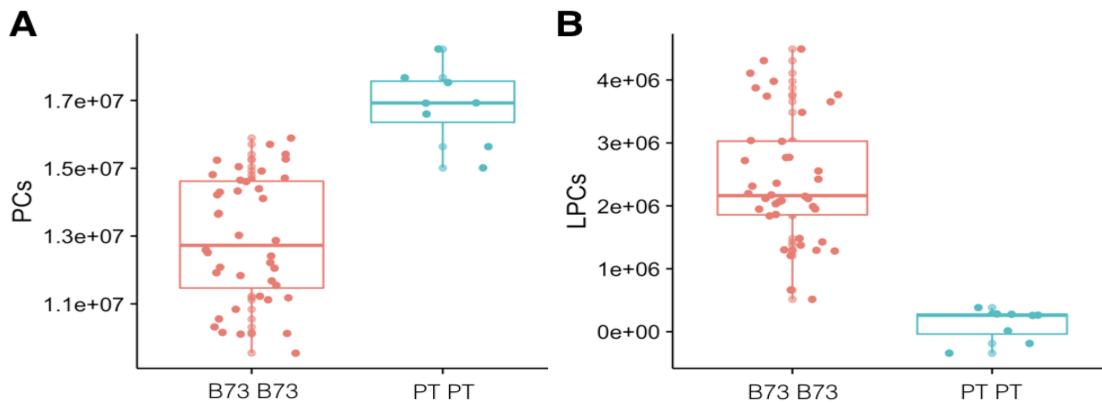

Figure 20. Allelic effect on the phenotype.

In Figure 20, the effect plots correspond to the allelic effect of B73 and PT alleles at the QTL peak found for the RILs. For the PC QTL peak, upper left plot, being homozygous for the B73 allele is reflected in a low level of PCs, while homozygous for PT alleles lead to higher levels of PCs. On the upper right plot the effect plot for LPCs QTL peak. Here, homozygous B73 alleles leads to higher LPCs levels, conversely, when a RIL is homozygous for PT alleles it leads to lower LPCs amount. So PT and B73 alleles at the QTL peak have opposite effects on PCs and LPCs.

After the QTL was identified, we looked for candidate genes in that region (Figure 19) that could explain the lipid phenotype we identified.

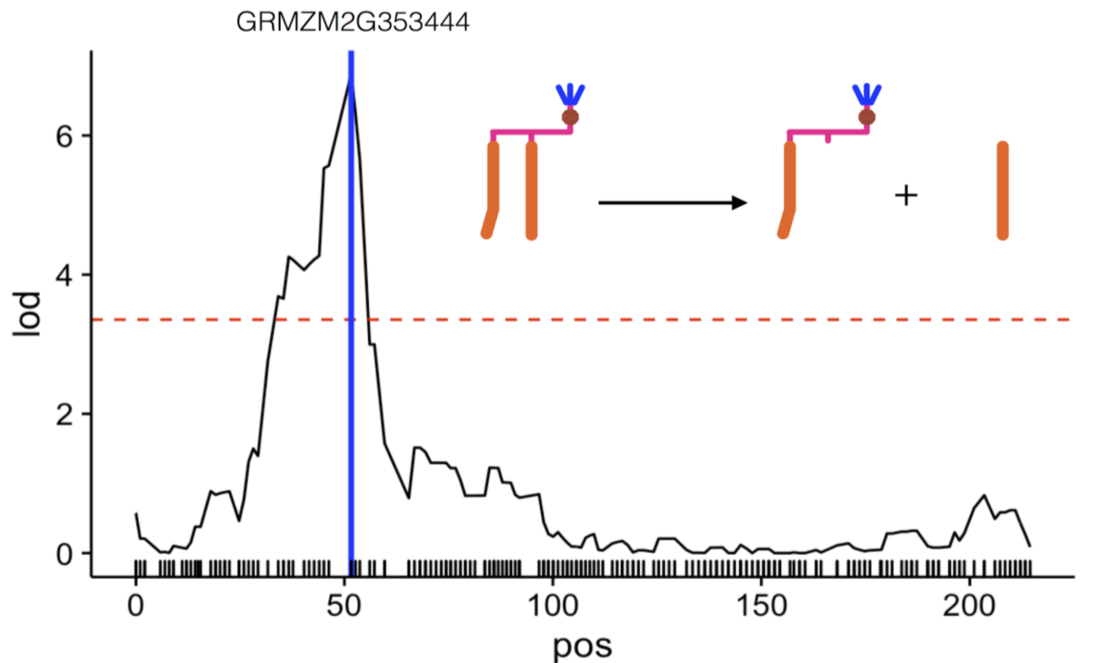

Figure 21. A phospholipase, GRMZM2G353444, lies in the PCs-LPCs QTL. A phospholipase with a predicted PC to LPC activity is lies within the QTL found on chromosome 3. GRMZM2G353444 is located at 7,735,192 - 7,737,952 bp on Chr3 and has a putative phospholipase type A (PLA_1) activity, which cleaves at sn-1 fatty acid (Figure 2) releasing a lysopohospholipid and a fatty acid. This PLA_1 activity could explain the conversion of PCs into LPCs, making GRMZM2G353444 a good candidate gene for explaining this phenotype.

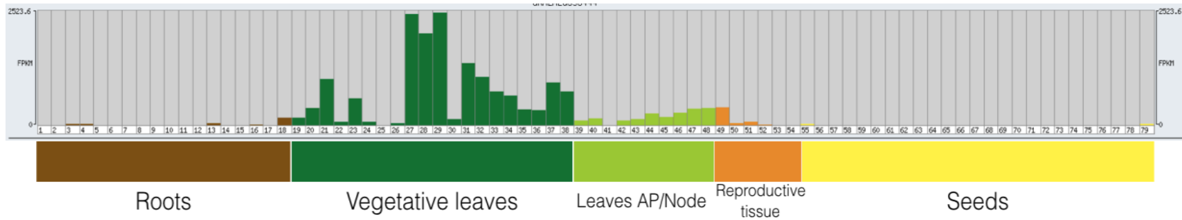

Figure 22. In B73 GRMZM2G353444 is highly expressed in leaves (V3-V9) 2523 fragments per kilobase per million mapped fragments (FPKM) according to http://www.maizegdb.org/.

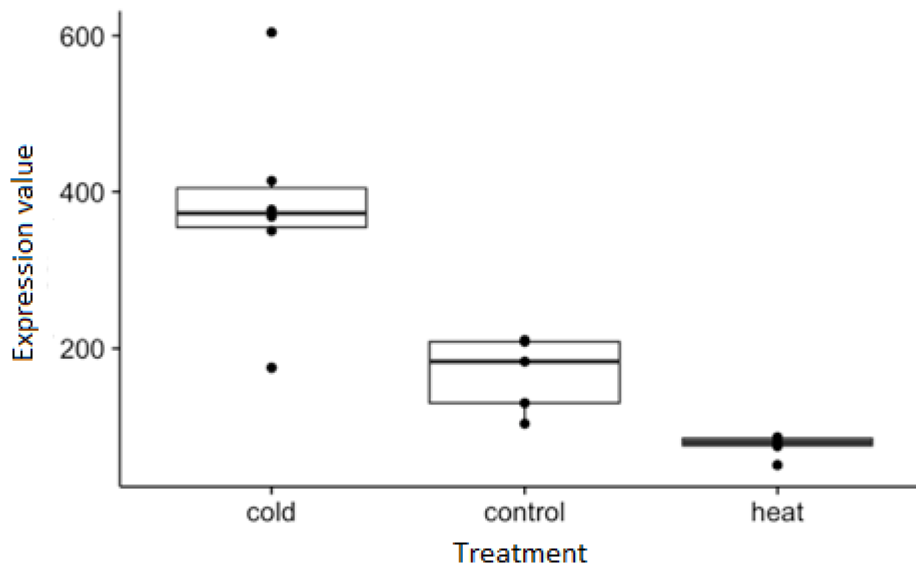

Figure 23. GRMZM2G353444 is upregulated in cold conditions.

All these results together with the upregulation of this PLA_1 under cold conditions suggest that PC to LPC conversion was selected against during maize adaptation to the Mexican highlands.This phenotype might be important for cold and/or phosphorus deficiency tolerance.

As shown in Figure 21 and 22, the candidate gene is more expressed on vegetative leaves than in any other tissue, which could be due to its importance for photoperiod-regulated processes.

When evaluating B73 and PT in Highland (Metepec) and Lowland (Valle de Banderas) fields,

one of the biggest differences between them is the flowering time (Figure 24 A).

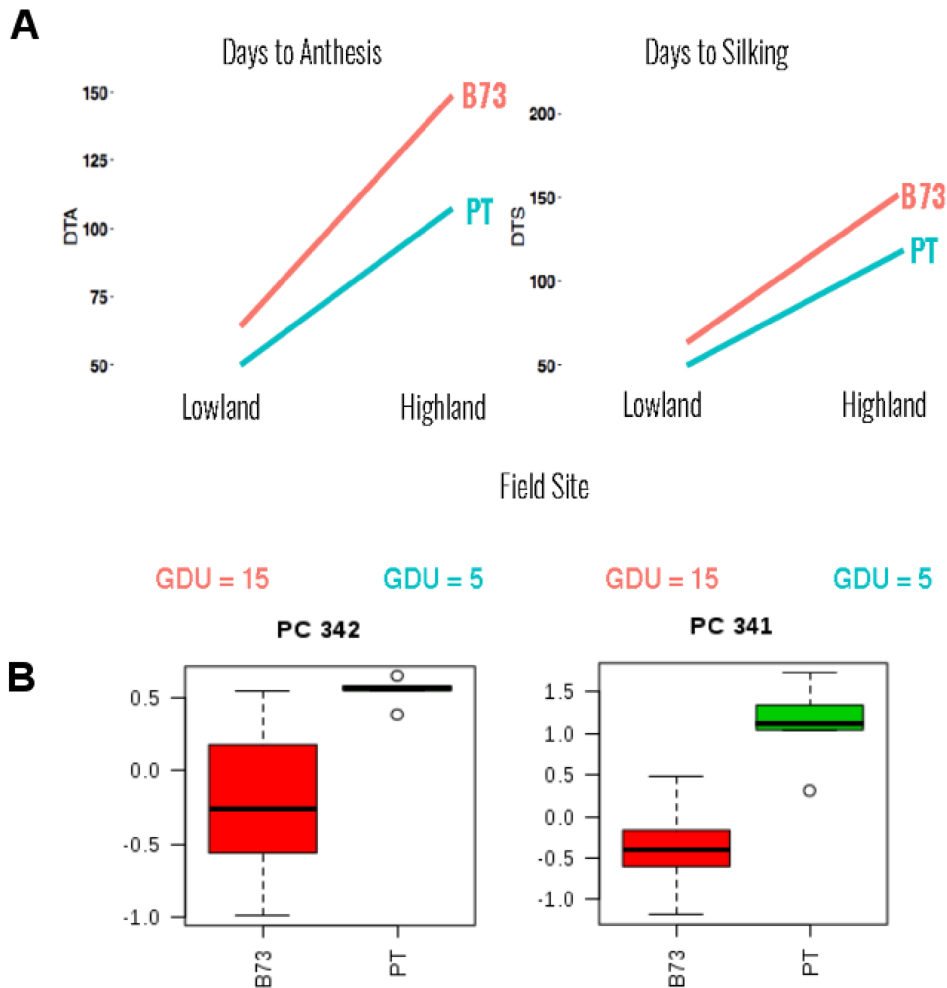

Figure 24. Flowering time for B73 and PT at highland and lowland conditions. **A.** Lowland site is located in Valle de Banderas, Nayarit. While Highland field is located in Metepec, Estado de México. This evaluation was made during the year 2015. (Aguilar, unpublished) **B.** Lipid species that show significant differences between B73 and PT, planted in Metepec.

Highland conditions are more challenging for maize, mainly due to the lower temperatures, phosphorus-deficient soils and high levels of radiation. As mentioned before, for lowland conditions (Valle de Banderas field) there is a GDD value of 5 while in highland conditions

(Metepec field) the GDD is set around 15. This index predicts that at lowland conditions maize would grow and develop three times faster than a highland site and we corroborated that with this figure. Highland landraces (here PT) are fast flowering lines under both, lowland and highland conditions.

In Figure 24 B, we see PC (34:2) and PC (34:1) as two of glycerophospholipid species that show to be more differentiated between RIL parentals, B73 and PT. These lipids are found at higher intensities in highland landrace, PT. Figure 10 (red highlighted) shows that PC (32:2), PC (36:4), PC (34:2), PC (34:1) and PC (36:2) vary together and explain 11.7% of the variance (component 2) between Mesoamerican and South American landraces. Finally, in Figure 13 (blue highlighted), PC (34:1) is one of the lipid species differentiated between Mesoamerican high and lowland landraces, found at higher intensities in highland landraces; comparing PC (34:1) intensities of highland landraces/ lowland landraces the fold change for it is 1.5381 and p value= 0.027101, so it is a significant different lipid species.

All these listed PC species were also found in a study performed by Nakamura and collaborators (Nakamura et al. 2014). According to that study, *Arabidopsis* florigen FT (flowering inducer) binds PC *in vitro*, while *in vivo*, transgenic plants producing an increased PC levels at the shoot meristem accelerates flowering and low PC levels delay this trait. So, PC levels seem to be correlated with flowering time. Moreover, along the day PC species are varying in levels and in saturation. Diurnal PCs are less unsaturated than dark period PCs (36:6 and 36:5). When performed a membrane-lipid overlay assay the found out that FT binds only to diurnal molecular PCs (32:2, 36:2, 36:4, 34:2 and 34:1) accelerating flowering. We found the same diurnal PC species being at higher levels along highland landraces, which according to Figure 24 A, correlates with an earlier flowering time.

Despite of these findings, deeper investigation about this correlation is needed, as our samples were taken (V3 - V6) much earlier than when flowering is initiated.

CONCLUSION

Using lipid profiling for generating biochemical phenotypes is a fast and very informative way to shed light on the genetic architecture behind local adaptation. We identified a QTL peak at 8.6Mb in chromosome 3 for phosphatidylcholine (PC) and lysophosphatidylcholine (LPC). A candidate gene at the QTL peak has a putative PC to LPC enzymatic activity and it is upregulated in cold conditions in B73. Our data suggests that PC to LPC conversion was selected against in highland Mesoamerican maize and to a lesser extent in highland South American maize.

These findings highlight the candidate structural function of this PLA_1 in maize stress response and also provides important hints to the genetic basis of the natural variation in maize cold tolerance. This could be a good start for future improvement of this trait.

PC to LPC conversion was selected against during maize adaptation to Mexican highlands. This phenotype might be caused by a mutation in PT, which could be selected also across other Mexican highland landraces, and might be important for cold and/or phosphorus deficiency tolerance. Further analyses using a landrace diversity panel suggest that this biochemical phenotype might have been selected for in highland Mexican maize but not in South American highland races.

PERSPECTIVES

As a first approach to identify if the differences between high and lowland phenotypes are due to a mutations or due to regulatory processes, the candidate gene (GRMZM2G353444) will be analyzed through Sanger sequencing. This will show if there is a mutation that leads to a non functional protein in genotypes such as PT where there is a higher level of PC than LPC. For this we will use the B73xPT RIL mapping population, which is convenient because all individuals are inbreds, and thus homozygous. For this, we are going to use extreme RILs for PC/LPC ratio; as shown in Figure 20 and Table S8, RILs with the highest PC/LPC ratio turn out to be the ones homozygous for PT alleles, while the ones with the lowest ratio are homozygous for B73 alleles. We will extract DNA from PT homozygous RILs and B73 homozygous RILs, which have the same PT or B73 haplotype since they were generated from an initial B73xPT cross. Based on the reference B73 sequence, we will design primers for amplifying the promoter, 5'UTR, exons, intron and 3'UTR of the candidate gene. In the case a mutation is found, we will design genetic markers to be able to follow this mutation along different materials.

Another suitable option for the development of a genetic marker to detect this possible mutation would be the use of SSRs. Based on SSRs sequences reported for B73, primers will be designed to amplify these fragments along B73xPT RILs which are known homozygous for B73 alleles and RILS homozygous for PT alleles. Once known if there is a mutation on the candidate gene, in collaboration with Michigan State University (MSU) we will proceed with the functional characterization of the gene through an enzyme-substrate assay. Also at MSU, we will do heterologous expression, of GRMZM2G353444, in *E.coli* will be performed to later be able to transform *Arabisopsis thaliana* and evaluate the effect of this gene under different stresses like low Pi and cold.

As shown before, Figures 21 and 22, GRMZM2G353444 is tissue-specific and is highly expressed along V3 - V9 stages in B73. Given that in the present work all the leaf samples were collected through V3 - V6 stages, analyzing gene expression, using RT-PCR, along these developmental range through RILs and parentals would be very informative to compare expression level of GRMZM2G353444 and other enzymes involved in glycerolipid metabolism, through different developmental stages. Along with this, transient expression of the candidate PLA_1 will be performed to know the exact location of the protein.

B73xPT F1 population had been grown in Valle de Banderas (lowland field) and it is currently growing in Metepec (highland field). From this population samples for lipid profiling and samples for RNAseq were collected, only for Valle de Banderas so far, from the same leaf longitudinally along each side of the main vein. WIth this data, it will be possible to analyze allele specific expression and also use GWAs as an approach to correlate genomic variation with our biochemical traits.

Finally, to understand the role of PCs in flowering time, we will be analyzing data we generated from lipid profiling of B73, PT and CML312 in greenhouse conditions. We collected leaf tissue from V3 to V12 stage, and from three different sections of the leaves (Base, middle and tip). PT started flowering with 15 fully expanded leaves, B73 flowered with 13 expanded leaves and 4 emerging leaves and CM312 started flowering at 17 expanded leaves and 1 emerging one, we believe the lipid profiles we acquired will be more representative for trying to correlate PC species with flowering time.

REFERENCES

Alpert, K. B., and S. D. Tanksley. 1996. “High-Resolution Mapping and Isolation of a Yeast Artificial Chromosome Contig Containing fw2.2: A Major Fruit Weight Quantitative Trait Locus in Tomato.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (26): 15503–7.

Arendt, Jeff, and David Reznick. 2008. “Convergence and Parallelism Reconsidered: What Have We Learned about the Genetics of Adaptation?” *Trends in Ecology & Evolution* 23 (1): 26–32.

Bohlenius, Henrik, Tao Huang, Laurence Charbonnel-Campaa, Amy M. Brunner, 4. Stefan Jansson, Steven H. Strauss, and Ove Nilsson. 2006. “CO/FT Regulatory Module Controls Timing of Flowering and Seasonal Growth Cessation in Trees.” *Science* 312.

Brewbaker, James L. 2014. “Diversity and Genetics of Tassel Branch Numbers in Maize.” *Crop Science* 55. Madison, WI: The Crop Science Society of America, Inc.: 65–78.

Broman, Karl W. 2005. “The Genomes of Recombinant Inbred Lines.” *Genetics* 169 (2): 1133–46.

Broun, P., J. Shanklin, E. Whittle, and C. Somerville. 1998. “Catalytic Plasticity of Fatty Acid Modification Enzymes Underlying Chemical Diversity of Plant Lipids.” *Science* 282 (5392): 1315–17.

Casati, Paula, and Virginia Walbot. 2005. “Differential Accumulation of Maysin and Rhamnosylisoorientin in Leaves of High-Altitude Landraces of Maize after UV-B Exposure.” *Plant, Cell & Environment* 28: 788–99.

CEA. 2017. “Biogenesis, Dynamics and Homeostasis of Membrane Lipids,” February. CEA. http://big.cea.fr/drf/big/english/PCV/LiPMB.

Chapman, Kent D. 1998. “Phospholipase Activity during Plant Growth and Development and in Response to Environmental Stress.” *Trends in Plant Science* 3 (11): 419–26.

Colosimo, Pamela F., Kim E. Hosemann, Sarita Balabhadra, Guadalupe Villarreal Jr, Mark Dickson, Jane Grimwood, Jeremy Schmutz, Richard M. Myers, Dolph Schluter, and David M. Kingsley. 2005. “Widespread Parallel Evolution in Sticklebacks by Repeated Fixation of Ectodysplasin Alleles.” *Science* 307 (5717): 1928–33.

Currat, Mathias, Guy Trabuchet, David Rees, Pascale Perrin, Rosalind M. Harding, John B. Clegg, Andre Langaney, and Laurent Excoffier. 2002. “Molecular Analysis of the B-Globin Gene Cluster in the Niokholo.” *American Journal of Human Genetics* 70: 207–23.

Degenkolbe, Thomas, Patrick Giavalisco, Ellen Zuther, Bettina Seiwert, Dirk K. Hincha, and Lothar Willmitzer. 2012. “Differential Remodeling of the Lipidome during Cold Acclimation in Natural Accessions of Arabidopsis Thaliana.” *The Plant Journal: For Cell and Molecular Biology* 72 (6): 972–82.

Doebley, John F. 1984. “Maize Introgression Into Teosinte-A Reappraisal.” *Annals of the Missouri Botanical Garden. Missouri Botanical Garden* 71 (4). Missouri Botanical

Garden Press: 1100–1113.
Doebley, John F., and Hugh H. Iltis. 1980. “Taxonomy of Zea (Gramineae). I. A Subgeneric Classification with Key to Taxa.” *American Journal of Botany* 67 (6). Botanical Society of America: 982–93.
Domonkos, Ildikó, Przemyslaw Malec, Anna Sallai, László Kovács, Kunihiro Itoh, Gaozhong Shen, Bettina Ughy, et al. 2004. “Phosphatidylglycerol Is Essential for Oligomerization of Photosystem I Reaction Center.” *Plant Physiology* 134 (4): 1471–78.
Eichten, Steven R., Jillian M. Foerster, Natalia de Leon, Ying Kai, Cheng-Ting Yeh, Sanzhen Liu, Jeffrey A. Jeddeloh, Patrick S. Schnable, Shawn M. Kaeppler, and Nathan M. Springer. 2011. “B73-Mo17 near-Isogenic Lines Demonstrate Dispersed Structural Variation in Maize.” *Plant Physiology* 156 (4): 1679–90.
Elmer, Kathryn R., and Axel Meyer. 2011. “Adaptation in the Age of Ecological Genomics: Insights from Parallelism and Convergence.” *Trends in Ecology & Evolution* 26 (6): 298–306.
Ewens, W. J., and R. S. Spielman. 2001. “Locating Genes by Linkage and Association.” *Theoretical Population Biology* 60 (3). Elsevier: 135–39.
Fahy, Eoin, Dawn Cotter, Manish Sud, and Shankar Subramaniam. 2011. “Lipid Classification, Structures and Tools.” *Biochimica et Biophysica Acta* 1811 (11): 637–47.
Flint-Garcia, Sherry A., Jeffry M. Thornsberry, and Edward S. Buckler 4th. 2003. “Structure of Linkage Disequilibrium in Plants.” *Annual Review of Plant Biology* 54: 357–74.
Flint-Garcia, Sherry A., Anne-Céline Thuillet, Jianming Yu, Gael Pressoir, Susan M. Romero, Sharon E. Mitchell, John Doebley, Stephen Kresovich, Major M. Goodman, and Edward S. Buckler. 2005. “Maize Association Population: A High-Resolution Platform for Quantitative Trait Locus Dissection.” *The Plant Journal: For Cell and Molecular Biology* 44 (6): 1054–64.
Fukunaga, Kenji, Jason Hill, Yves Vigouroux, Yoshihiro Matsuoka, Jesus Sanchez G, Kejun Liu, Edward S. Buckler, and John Doebley. 2005. “Genetic Diversity and Population Structure of Teosinte.” *Genetics* 169 (4): 2241–54.
Gagne, J. M., and S. E. Clark. 2010. “The Arabidopsis Stem Cell Factor POLTERGEIST Is Membrane Localized and Phospholipid Stimulated.” *The Plant Cell* 22 (3): 729–43.
Griffiths, Anthony J. F., Susan R. Wessler, Sean B. Carroll, and John Doebley. 2015. *An Introduction to Genetic Analysis*. W. H. Freeman.
Hagio, M., Z. Gombos, Z. Várkonyi, K. Masamoto, N. Sato, M. Tsuzuki, and H. Wada. 2000. “Direct Evidence for Requirement of Phosphatidylglycerol in Photosystem II of Photosynthesis.” *Plant Physiology* 124 (2): 795–804.
Hanson, M. A., B. S. Gaut, A. O. Stec, S. I. Fuerstenberg, M. M. Goodman, E. H. Coe, and J. F. Doebley. 1996. “Evolution of Anthocyanin Biosynthesis in Maize Kernels: The Role of Regulatory and Enzymatic Loci.” *Genetics* 143 (3): 1395–1407.
Heerwaarden, Joost van, John Doebley, William H. Briggs, Jeffrey C. Glaubitz, Major M. Goodman, Jose de Jesus Sanchez Gonzalez, and Jeffrey Ross-Ibarra. 2011. “Genetic

Signals of Origin, Spread, and Introgression in a Large Sample of Maize Landraces." *Proceedings of the National Academy of Sciences of the United States of America* 108 (3): 1088–92.
Hinsinger, Philippe. 2001. "Bioavailability of Soil Inorganic P in the Rhizosphere as Affected by Root-Induced Chemical Changes: A Review." *Plant and Soil* 237: 173–95.
Hopper, Stephen D. 2009. "OCBIL Theory: Towards an Integrated Understanding of the Evolution, Ecology and Conservation of Biodiversity on Old, Climatically Buffered, Infertile Landscapes." *Plant and Soil* 322 (1-2). Springer Netherlands: 49–86.
Horn, Patrick J., and Christoph Benning. 2016. "The Plant Lipidome in Human and Environmental Health." *Science* 353 (6305): 1228–32.
Hufford, Matthew B., Pesach Lubinksy, Tanja Pyhäjärvi, Michael T. Devengenzo, Norman C. Ellstrand, and Jeffrey Ross-Ibarra. 2013. "The Genomic Signature of Crop-Wild Introgression in Maize." *PLoS Genetics* 9 (5): e1003477.
Hufford, Matthew B., Enrique Martínez-Meyer, Brandon S. Gaut, Luis E. Eguiarte, and Maud I. Tenaillon. 2012. "Inferences from the Historical Distribution of Wild and Domesticated Maize Provide Ecological and Evolutionary Insight." *PloS One* 7 (11): e47659.
Iltis, Hugh H., and Bruce F. Benz. 2000. "Zea Nicaraguensis (Poaceae), a New Teosinte from Pacific Coastal Nicaragua." *Novon: A Journal for Botanical Nomenclature / Missouri Botanical Garden* 10 (4): 382–90.
Keurentjes, Joost J. B., Leónie Bentsink, Carlos Alonso-Blanco, Corrie J. Hanhart, Hetty Blankestijn-De Vries, Sigi Effgen, Dick Vreugdenhil, and Maarten Koornneef. 2007. "Development of a near-Isogenic Line Population of Arabidopsis Thaliana and Comparison of Mapping Power with a Recombinant Inbred Line Population." *Genetics* 175 (2): 891–905.
Knowler, W. C., R. C. Williams, D. J. Pettitt, and A. G. Steinberg. 1988. "Gm3;5,13,14 and Type 2 Diabetes Mellitus: An Association in American Indians with Genetic Admixture." *American Journal of Human Genetics* 43 (4). ncbi.nlm.nih.gov: 520–26.
Körner, Christian. 2007. "The Use of 'Altitude' in Ecological Research." *Trends in Ecology & Evolution* 22 (11): 569–74.
Kovach, Michael J., Mariafe N. Calingacion, Melissa A. Fitzgerald, and Susan R. McCouch. 2009. "The Origin and Evolution of Fragrance in Rice (Oryza Sativa L.)." *Proceedings of the National Academy of Sciences of the United States of America* 106 (34): 14444–49.
Krasilnikov, Pavel, Ma del Carmen Gutiérrez-Castorena, Robert J. Ahrens, Carlos Omar Cruz-Gaistardo, Sergey Sedov, and Elizabeth Solleiro-Rebolledo. 2013. *The Soils of Mexico*. Springer Science & Business Media.
Kruglyak, L. 1999. "Prospects for Whole-Genome Linkage Disequilibrium Mapping of Common Disease Genes." *Nature Genetics* 22 (2): 139–44.
Kwiatkowski, Dominic P. 2005. "How Malaria Has Affected the Human Genome and What Human Genetics Can Teach Us about Malaria." *American Journal of Human Genetics* 77: 171–92.

Lambers, Hans. 2014. *Plant Life on the Sandplains in Southwest Australia: A Global Biodiversity Hotspot*. Apollo Books.

Lambers, H., P. M. Finnegan, R. Jost, W. C. Plaxton, M. W. Shane, and M. Stitt. 2015. "Phosphorus Nutrition in Proteaceae and beyond." *Nature Publishing Group*, August. Macmillan Publishers Limited, 1–9.

Lander, E. S., and N. J. Schork. 1994. "Genetic Dissection of Complex Traits." *Science* 265 (5181). search.proquest.com: 2037–48.

Lauter, Nick, Charles Gustus, Anna Westerbergh, and John Doebley. 2004. "The Inheritance and Evolution of Leaf Pigmentation and Pubescence in Teosinte." *Genetics* 167 (4): 1949–59.

Liakopoulos, Georgios, Dimosthenis Nikolopoulos, Aspasia Klouvatou, Kornilios-Andrianos Vekkos, Yiannis Manetas, and George Karabourniotis. 2006. "The Photoprotective Role of Epidermal Anthocyanins and Surface Pubescence in Young Leaves of Grapevine (Vitis Vinifera)." *Annals of Botany* 98 (1): 257–65.

Li, Baohua, Yuanyuan Zhang, Seyed Abolghasem Mohammadi, Dongxin Huai, Yongming Zhou, and Daniel J. Kliebenstein. 2016. "An Integrative Genetic Study of Rice Metabolism, Growth and Stochastic Variation Reveals Potential C/N Partitioning Loci." *Scientific Reports* 6 (July): 30143.

Li-Beisson, Yonghua, Basil Shorrosh, Fred Beisson, Mats X. Andersson, Vincent Arondel, Philip D. Bates, Sébastien Baud, et al. 2013. "Acyl-Lipid Metabolism." *The Arabidopsis Book / American Society of Plant Biologists*, January. The American Society of Plant Biologists, e0161.

Li, Long, Shu-Min Li, Jian-Hao Sun, Li-Li Zhou, Xing-Guo Bao, Hong-Gang Zhang, and Fu-Suo Zhang. 2007. "Diversity Enhances Agricultural Productivity via Rhizosphere Phosphorus Facilitation on Phosphorus-Deficient Soils." *Proceedings of the National Academy of Sciences of the United States of America* 104 (27): 11192–96.

Lin, Ying-Chen, Yu-Chi Liu, and Yuki Nakamura. 2015. "The Choline/Ethanolamine Kinase Family in Arabidopsis: Essential Role of CEK4 in Phospholipid Biosynthesis and Embryo Development." *The Plant Cell* 27 (5): 1497–1511.

Lynch, D. V., and P. L. Steponkus. 1987. "Plasma Membrane Lipid Alterations Associated with Cold Acclimation of Winter Rye Seedlings (Secale Cereale L. Cv Puma)." *Plant Physiology* 83 (4): 761–67.

Lynch, Jonathan P., and Kathleen M. Brown. 2012. "New Roots for Agriculture: Exploiting the Root Phenome." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 367 (1595): 1598–1604.

Matyash, Vitali, Liebisch, Gerhard, Kurzchalia, Teymuras, Shevchenko, Andrej, and Schwudke, Dominik. 2008. "Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics." *Journal of lipid research* 49 (5): 1137–46.

Matsuoka, Yoshihiro, Yves Vigouroux, Major M. Goodman, Jesus Sanchez G, Edward Buckler, and John Doebley. 2002. "A Single Domestication for Maize Shown by Multilocus Microsatellite Genotyping." *Proceedings of the National Academy of Sciences of the United States of America* 99 (9): 6080–84.

Meer, Gerrit van, Dennis R. Voelker, and Gerald W. Feigenson. 2008. “Membrane Lipids: Where They Are and How They Behave.” *Nature Reviews. Molecular Cell Biology* 9 (2): 112–24.
Meihls, L. N., H. Kaur, and G. Jander. 2012. “Natural Variation in Maize Defense against Insect Herbivores.” *Cold Spring Harbor Symposia on Quantitative Biology* 77 (December): 269–83.
Michaelson, Louise V., Johnathan A. Napier, Diana Molino, and Jean-Denis Faure. 2016. “Plant Sphingolipids: Their Importance in Cellular Organization and Adaption.” *Biochimica et Biophysica Acta* 1861 (9 Pt B): 1329–35.
Michaud, Morgane, Valérie Gros, Marianne Tardif, Sabine Brugière, Myriam Ferro, William A. Prinz, Alexandre Toulmay, et al. 2016. “AtMic60 Is Involved in Plant Mitochondria Lipid Trafficking and Is Part of a Large Complex.” *Current Biology: CB* 26 (5): 627–39.
Nakamura, Yuki, Fernando Andrés, Kazue Kanehara, Yu-Chi Liu, Peter Dörmann, and George Coupland. 2014. “Arabidopsis Florigen FT Binds to Diurnally Oscillating Phospholipids That Accelerate Flowering.” *Nature Communications* 5 (April). Nature Publishing Group: 3553.
Nishida, I., and N. Murata. 1996. “CHILLING SENSITIVITY IN PLANTS AND CYANOBACTERIA: The Crucial Contribution of Membrane Lipids.” *Annual Review of Plant Physiology and Plant Molecular Biology* 47 (June): 541–68.
Okazaki, Yozo, Hitomi Otsuki, Tomoko Narisawa, Makoto Kobayashi, Satoru Sawai, Yukiko Kamide, Miyako Kusano, Toshio Aoki, Masami Yokota Hirai, and Kazuki Saito. 2013. “A New Class of Plant Lipid Is Essential for Protection against Phosphorus Depletion.” *Nature Communications* 4: 1510.
Peleman, Johan D., Crispin Wye, Jan Zethof, Anker P. Sørensen, Henk Verbakel, Jan van Oeveren, Tom Gerats, and Jeroen Rouppe van der Voort. 2005. “Quantitative Trait Locus (QTL) Isogenic Recombinant Analysis: A Method for High-Resolution Mapping of QTL within a Single Population.” *Genetics* 171 (3): 1341–52.
Peret, B., T. Desnos, R. Jost, S. Kanno, O. Berkowitz, and L. Nussaume. 2014. “Root Architecture Responses: In Search of Phosphate.” *Plant Physiology* 166 (4): 1713–23.
Piperno, D. R., and K. V. Flannery. 2001. “The Earliest Archaeological Maize (Zea Mays L.) from Highland Mexico: New Accelerator Mass Spectrometry Dates and Their Implications.” *Proceedings of the National Academy of Sciences of the United States of America* 98 (4): 2101–3.
Plaxton, William C., and Hue T. Tran. 2011. “Metabolic Adaptations of Phosphate-Starved Plants.” *Plant Physiology* 156 (3): 1006–15.
Pollard, Daniel A. 2012. “Design and Construction of Recombinant Inbred Lines.” *Methods in Molecular Biology* 871: 31–39.
Pritchard, J. K., and N. A. Rosenberg. 1999. “Use of Unlinked Genetic Markers to Detect Population Stratification in Association Studies.” *American Journal of Human Genetics* 65 (1). Elsevier: 220–28.
Rellán-Álvarez, Rubén, Guillaume Lobet, Heike Lindner, Pierre-Luc Pradier, Jose Sebastian, Muh-Ching Yee, Yu Geng, et al. 2015. “GLO-Roots: An Imaging Platform

Enabling Multidimensional Characterization of Soil-Grown Root Systems." *eLife* 4 (August). doi:10.7554/eLife.07597.
Riedelsheimer, Christian, Yariv Brotman, Michaël Méret, Albrecht E. Melchinger, and Lothar Willmitzer. 2013. "The Maize Leaf Lipidome Shows Multilevel Genetic Control and High Predictive Value for Agronomic Traits." *Scientific Reports* 3 (August). Nature Publishing Group: 2479.
Ross-Ibarra, Jeffrey, Maud Tenaillon, and Brandon S. Gaut. 2009. "Historical Divergence and Gene Flow in the Genus Zea." *Genetics* 181 (4): 1399–1413.
Sharbel, T. F., B. Haubold, and T. Mitchell-Olds. 2000. "Genetic Isolation by Distance in Arabidopsis Thaliana: Biogeography and Postglacial Colonization of Europe." *Molecular Ecology* 9 (12). Wiley Online Library: 2109–18.
Smith, Anthony David. 2000. *Oxford Dictionary of Biochemistry and Molecular Biology*. Oxford University Press.
Song, Young Hun, Shogo Ito, and Takato Imaizumi. 2013. "Flowering Time Regulation: Photoperiod- and Temperature-Sensing in Leaves." *Trends in Plant Science* 18 (10): 575–83.
Song, Young Hun, Jae Sung Shim, Hannah A. Kinmonth-Schultz, and Takato Imaizumi. 2015. "Photoperiodic Flowering: Time Measurement Mechanisms in Leaves." *Annual Review of Plant Biology* 66: 441–64.
Sorkheh, Karim, Lyudmyla V. Malysheva-Otto, Michelle G. Wirthensohn, Saeed Tarkesh-Esfahani, and Pedro Martínez-Gómez. 2008. "Linkage Disequilibrium, Genetic Association Mapping and Gene Localization in Crop Plants." *Genetics and Molecular Biology* 31 (4). Sociedade Brasileira de Genética: 805–14.
Steyn, W. J., S. J. E. Wand, D. M. Holcroft, and G. Jacobs. 2002. "Anthocyanins in Vegetative Tissues: A Proposed Unified Function in Photoprotection." *The New Phytologist* 155 (3). Blackwell Science Ltd: 349–61.
Strable, Josh, and Michael J. Scanlon. 2009. "Maize (Zea Mays): A Model Organism for Basic and Applied Research in Plant Biology." *Cold Spring Harbor Protocols* 2009 (10): db.emo132.
Sulpice, Ronan, Anna Flis, Alexander A. Ivakov, Federico Apelt, Nicole Krohn, Beatrice Encke, Christin Abel, Regina Feil, John E. Lunn, and Mark Stitt. 2014. "Arabidopsis Coordinates the Diurnal Regulation of Carbon Allocation and Growth across a Wide Range of Photoperiods." *Molecular Plant* 7 (1): 137–55.
Takuno, Shohei, Peter Ralph, Kelly Swarts, Rob J. Elshire, Jeffrey C. Glaubitz, Edward S. Buckler, Matthew B. Hufford, and Jeffrey Ross-Ibarra. 2015. "Independent Molecular Basis of Convergent Highland Adaptation in Maize." *bioRxiv*. doi:10.1101/013607.
Tenaillon, M. I., M. C. Sawkins, A. D. Long, R. L. Gaut, J. F. Doebley, and B. S. Gaut. 2001. "Patterns of DNA Sequence Polymorphism along Chromosome 1 of Maize (Zea Mays Ssp. Mays L.)." *Proceedings of the National Academy of Sciences of the United States of America* 98 (16). National Acad Sciences: 9161–66.
Tran, Hue T., Brenden A. Hurley, and William C. Plaxton. 2010/7. "Feeding Hungry Plants: The Role of Purple Acid Phosphatases in Phosphate Nutrition." *Plant Science: An*

International Journal of Experimental Plant Biology 179 (1–2): 14–27.
Veneklaas, Erik J., Hans Lambers, Jason Bragg, Patrick M. Finnegan, Catherine E. Lovelock, William C. Plaxton, Charles A. Price, et al. 2012. “Opportunities for Improving Phosphorus-Use Efficiency in Crop Plants.” *The New Phytologist* 195 (2): 306–20.
Vigouroux, Yves, Jeffrey C. Glaubitz, Yoshihiro Matsuoka, Major M. Goodman, Jesús Sánchez G., and John Doebley. 2008. “Population Structure and Genetic Diversity of New World Maize Races Assessed by DNA Microsatellites.” *American Journal of Botany* 95 (10): 1240–53.
Wellhausen, Edwin John. 1957. *Races of Maize in Central America*. National Academies.
Wilkes, H. G. 1977. “Hybridization of Maize and Teosinte, in Mexico and Guatemala and the Improvement of Maize.” *Economic Botany* 31 (3). New York Botanical Garden Press: 254–93.
Wood, Troy E., John M. Burke, and Loren H. Rieseberg. 2005. “Parallel Genotypic Adaptation: When Evolution Repeats Itself.” *Genetica* 123 (1-2): 157–70.
Xu, Yang, Pengcheng Li, Zefeng Yang, and Chenwu Xu. 2017/4. “Genetic Mapping of Quantitative Trait Loci in Crops.” *The Crop Journal* 5 (2): 175–84.
Yadav, B. K., and A. Verma. 2012. “Phosphate Solubilization and Mobilization in Soil through Microorganisms under Arid Ecosystems.” *The Functioning of Ecosystems. InTech,*. researchgate.net. https://www.researchgate.net/profile/Ali_Mahamane/publication/259465664_The_Functioning_of_Ecosystems/links/5537f0ce0cf247b8587d08e1.pdf#page=105.
Zheng, Guowei, Bo Tian, Fujuan Zhang, Faqing Tao, and Weiqi Li. 2011. “Plant Adaptation to Frequent Alterations between High and Low Temperatures: Remodelling of Membrane Lipids and Maintenance of Unsaturation Levels.” *Plant, Cell & Environment* 34 (9): 1431–42.

SUPPLEMENTARY MATERIAL

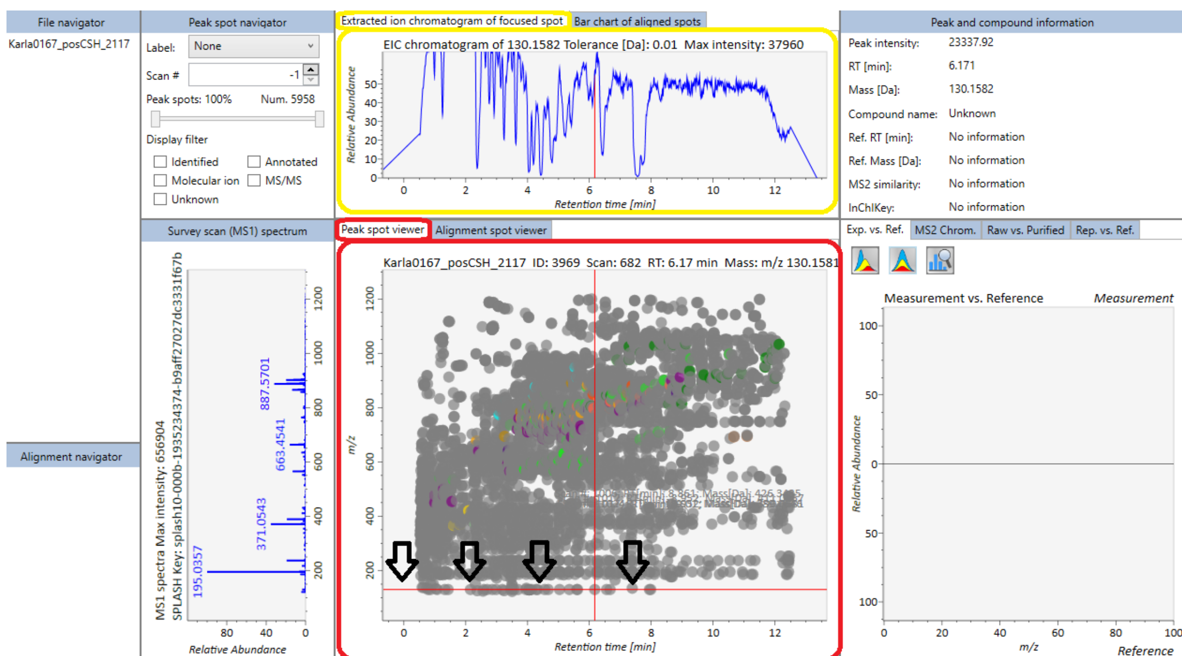

Figure S1. MS-DIAL noise peak detection. Features with m/z of 130.16, 225.209 and 211.19 are often found in LC.MS runs, and they correspond to noise. As seen on the Peak spot viewer (in red), the m/z 130.1582 is found repeatedly along different retention times (pointed with black arrows), which is the first characteristic of noise. Then, on the EIC (in yellow), this m/z does not show any defined peak. This image corresponds to LANMLR17B036 RIL lipid sample analyzed with MS-DIAL.

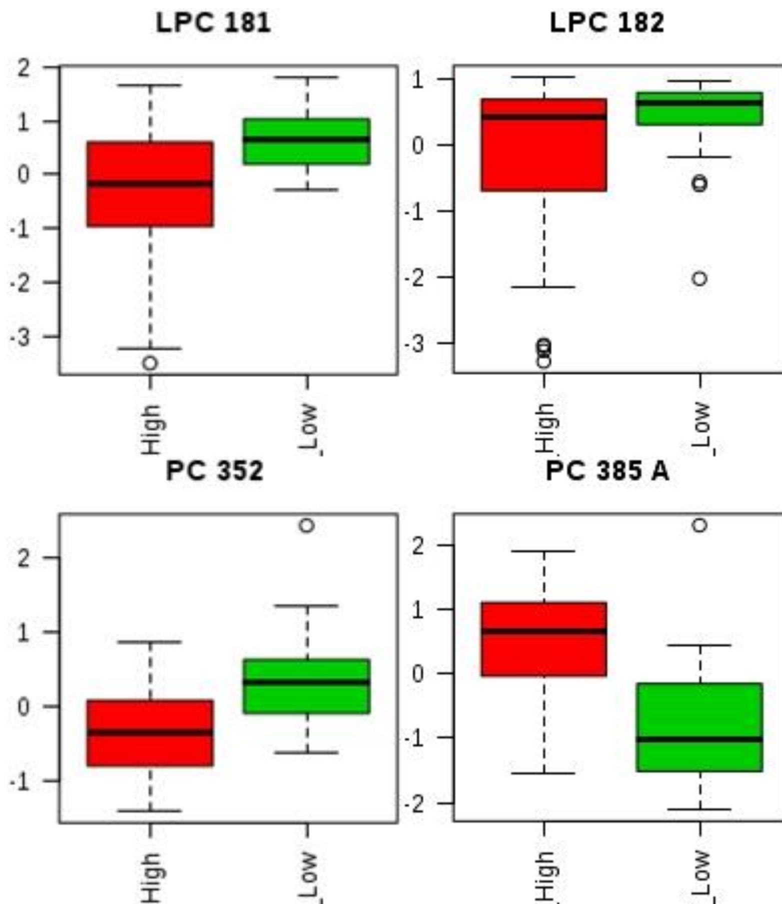

Figure S2. Lipid species that show significant differences between South American highland and lowland landraces. 60 South American landraces were used, 30 of them were highland landraces and 30 lowland ones. Two plants of each landraces were planted in Metepec. Student's t-test was used as statistical test and box plots show the difference of lipid abundance between South American high and lowlands.
Fold change and p values: LPC (18:1) FC= 0.55046 and pvalue = 9.6379E-4, LPC (18:2) FC= 0.7769 and pvalue= 0.046116, PC (35:2) FC= 0.72633 and p value= 0.0013691 and finally for PC (38:5) FC= 2.2376 and p value= 1.9793E-5.

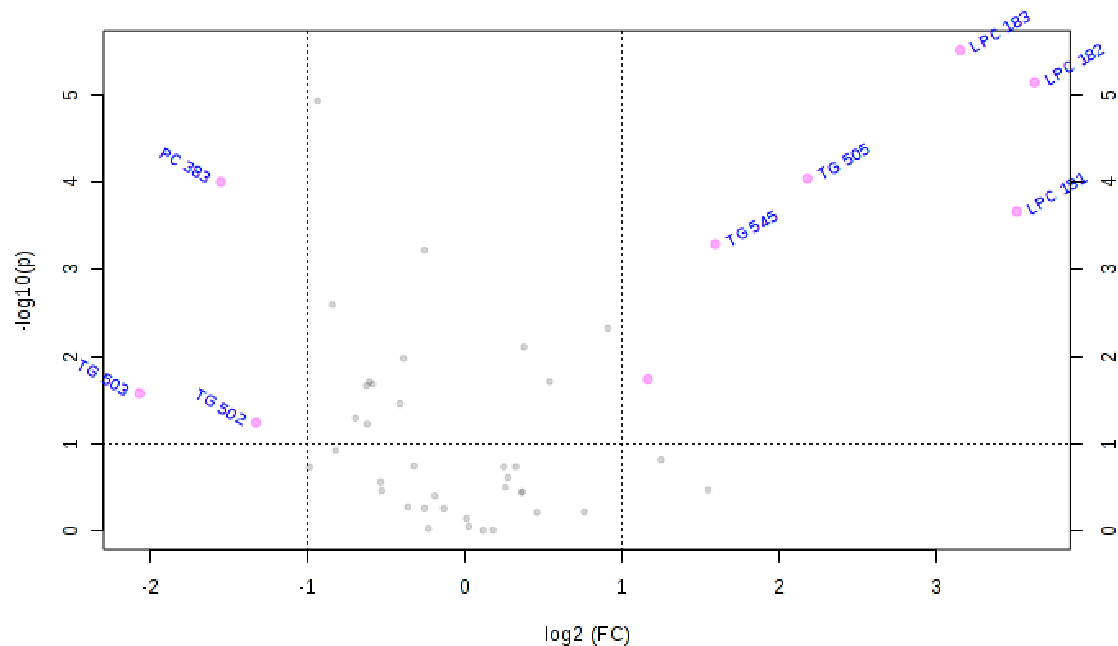

Figure S3. Differentiated lipid species between B73 and PT in Figure 14. Significantly different lipid species shown with red spots, FC of 1 and p=0.05

Tables S1, S2, S3, S4, S7, S8 and Files S1 to S4 can be found in the next site:

http://www.rrlab.org/karla_ms_thesis.html

Table S5. Fold change and p value of differentiated lipid species between Mesoamerican highland and lowland landraces of Figure 13.

Lipid	Fold Change	P-value
LPC (18:2)	0.64431	0.002907
PC (32:0)	0.48763	4.94E-08

PC (32:1)	0.71424	0.030949
PC (34:1)	1.5381	0.027101
PC (35:2)	0.6863	4.27E-04
PC (36:5) A	1.3983	0.023299
PC (36:5) B	1.276	0.039681
PC (38:5) A	2.331	1.21E-04

Fold change is based on the comparison of MA_highlands/MA_lowlands.

Table S6. Fold change and p value of differentiated lipid species between South American highland and lowland landraces of Supplementary Figure 2.

Lipid	Fold change	P-value
LPC (18:1)	0.55046	9.64E-04
LPC (18:2)	0.7769	4.61E-02
PC (35:2)	0.72633	1.37E-03
PC (38:5) A	2.2376	1.98E-05