CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD DE GENÓMICA AVANZADA

UNIDAD IRAPUATO

Caracterización funcional de la familia de genes *Xipotl* (fosfoetanolamina N-metiltransferasa, PEAMT) de *Zea mays*

A thesis submitted by

Estefany Sofia Sánchez Martínez

In fulfillment of the Requirements for the Degree of

Master in Science

With The Specialty in

Plant Biotechnology

Directors:
Dr. Rubén Rellán Álvarez
Dr. Ruairidh J. H. Sawers

Irapuato, Guanajuato Diciembre 2018

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL

UNIDAD DE GENÓMICA AVANZADA

UNIDAD IRAPUATO

Caracterización funcional de la familia de genes *Xipotl* (fosfoetanolamina N-metiltransferasa, PEAMT) de *Zea mays*

Tesis que presenta

Estefany Sofia Sánchez Martínez

Para obtener el grado de

Maestría en Ciencias

En la Especialidad de

Biotecnología de Plantas

Directores de Tesis:
Dr. Rubén Rellán Álvarez
Dr. Ruairidh J. H. Sawers

Irapuato, Guanajuato Diciembre 2018

El presente trabajo de investigación se desarolló en los Grupos de **Nutrición Mineral de Plantas** y **Genómica y Genética del Maíz** a cargo del **Dr. Rubén Rellán Álvarez** y el **Dr. James Hay Ruairidh Sawers**, en la Unidad de Genómica Avanzada (UGA) del CINVESTAV-IPN Irapuato. Se agradece el apoyo a CONACyT a través del programa de becas para estudios de posgrado (CVU/Becario): **711680/153600044**.

For everyone who, once again, believed in me.

"*... And that thou less may wonder at my word, behold the sun's heat, wich becometh wine, joined to the juice that from the vine distils."*

Dante Alighieri. *Divine Comedy*

"*Well Penny, like anything worth writing, it came inexplicably and without method."*

Karen Eiffel. *Stranger Than Fiction*

Acknowledgements

To CINVESTAV and LANGEBIO for the support and all the personnel who may help me in different times

To my advisors: Ruben Rellán and Ruairidh Sawers. First of all, for their patience (I really thank you for this). For all the things you (aware or not) taught me through this years. For absolutely everything, you have my whole admiration and gratitude, because, you let me stay and keep panding around.

To my committee: Alfredo Cruz and Alex de Luna, for your observations during the tutorials and teaching me different points of view of my work.

To my ”sensei” and friend: Juan Estevez. Thank you for listening, or being mad with me, for teaching me in and outside the lab that ”there’s always a solution”.

Mom and Dad, after all this time, I don’t know how to thank you. All your courage, your support, your patience, helped me everytime. Flaquita and Carola, each of your hugs helped me to put myself together. Thank you for the laughs, the hashtags, the silly things that made me happy. My loved family, as always, this one is for you!

Meow, no matter what, we’ll always have those nights (if you know what I mean). Thank you for the beach, Sweden, the rooftops, the low key therapy sessions, the heights and the worst. I think I wouldn’t be able to finish this without you.

Thank you Monkeys! For making each of my days, for your help when I needed it, for the talks, the coffees and every little single detail that made my days.

Bitches, this years were tough as f*ck, so... Thank you for being there for me, and no matter what, I loved every moment we spent together. We’ll always have Laguna!

To each of the fabulous karaoke rock-stars, billiard masters, happy campers, adventure seekers, after-parties survivors, fatty lovely friends. For listening, leading and following me, and never leaving me alone. Thank you!

I found you in the end, and it doesn’t mean the least. My penguin knight, thank you for dissipate the little clouds.

Agradecimientos

A CINVESTAV Unidad Irapuato y a la UGA por todo el apoyo, así como a todo el personal que me ayudó durante todo este tiempo.

A mis asesores: Rubén Rellán y Ruairidh Sawers. Antes que nada, por su paciencia (de verdad les agradezco). Por todas las cosas que me enseñaron en estos años sin saberlo, porque me dejaron quedarme, tienen mi completa admiración y gratitud.

A mi comité tutorial: Alfredo Cruz y Alex de Luna, por sus observaciones durante las reuniones y tutoriales, y por enseñarme una perspectiva diferente sobre mi trabajo.

A mi ”sensei” y amigo: Juan Estevez. Gracias por escucharme (o jalarme las orejas), por enseñarme que incluso ahora, todo tiene solución.

A mis padres, después de todo este tiempo ¡gracias! Porque su fortaleza, su apoyo y su paciencia, me han ayudado a seguir adelante. Flaquita y Carola, por los abrazos, por las risas, los hashtags y todas las ridiculeces que me hacen feliz. Mi amada familia, igual que siempre, ¡ésto es para ustedes!

Meow, esas noches siempre serán nuestras. Gracias por la playa, Suecia, la terraza, las terapias, la comida, las subidas y bajadas. Sabes que no hubiera podido terminar ésto sin ti.

Monkeys! Gracias por hacer mis días, por toda su ayuda, las pláticas, los cafés, la sobremesa y por cada mínimo e insignificante detalle.

Bitches, sabemos lo díficil que fue todo ésto y aún así... Gracias por estar ahí siempre para mí, y no importa que pasó, he amado cada momento que pasamos juntos. ¡Siempre tendremos Laguna!

A cada uno de los cantantes de karaoke, maestros del billar, campistas, aventureros, mis gorditos y amados amigos. Por estar ahí y nunca dejarme: ¡Gracias totales!

Al final (pero no menos importante), encontré a un pingüino que me ayudó a disipar las nubes ¡Gracias Walter!

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional

Abstract

Unidad de Genómica Avanzada

Master in Science

Functional characterization of *Zea mays Xipotl* (phosphoethanolamine N-methyltransferase, PEAMT) family genes

by Estefany Sofia Sánchez Martínez

Phosphatidylcholine (PtCho) is the most abundant phospholipid in eukaryotic cell membranes. PtCho polar head is synthesized via either the Kennedy Pathway or the triple methylation of phosphoethanolamine methyl transferase (PEAMT). In *Arabidopsis* an insertional mutant in the AT3G18000 gene (*PEAMT, xipotl, xpl1*) presented a short root phenotype, this was the first gene reported of the 3 coding loci in Arabidopsis for PEAMT: AT1G48600 and AT1G73600. The *Atxpl2* and *xpl3* mutants, doesn't show short root phenotype, instead, the *xpl2* root seems larger than the wild type and the triple mutant root, *xpl 1,2,3*, shorter than the *xpl1* mutant. In addition, the xpl mutants exhibit affected embryo development. So far, in *Zea mays*, 4 genes are predicted to encode PEAMT enzyme. Nevertheless, only 3 genes conserve the MT1 and MT2 methyltransferase domains. The genes had been named: *xplA, xplB and xplC*.

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional

Resumen

Unidad de Genómica Avanzada

Maestría en Ciencias

Caracterización funcional de la familia de genes *Xipotl* (fosfoetanolamina N-metiltransferasa, PEAMT) de *Zea mays*

Por Estefany Sofia Sánchez Martínez

En las células eucariotas, las fosfatidilcolina es el fosfolípido más abundante que compone a las membranas, puede ser sintetizada por la ruta de Kennedy o por la ruta de la triple metilación realizada por la enzyma fosfoetnolamina N-metiltransferasa (PEAMT, por sus siglas en inglés, *xipotl*, *xpl*). En la planta modelo *Arabidopsis thaliana*, la mutación realizada al gen que codifica a esta enzima (AT3G18000, *xpl1*) generó un fenotipo de raíz corta debido a muerte celular en el meristemo. Este fue el primer gen reportado para PEAMT, después se reportaron 2 genes también codificantes para la enzima: AT1G48600 y AT1G73600. Sim embargo, las mutantes para los genes *xpl2 y xpl3*, respectivamente, no mostraron raíces cortas, y la triple mutante *xpl1,2,3*, mostró una raíz incluso más corta que la mutante sencilla,*xpl1*. Desde entonces, se ha demostrado que PEAMT, se encuentra altamente relacionada a condiciones de estrés, y su secuencia genetica, proteica y mecanismo de regulación se encuentran altamente conservados. Por otro lado, en maíz, se han reportado 4 genes codificantes para PEAMT y debido a la importancia de este cultivo y a la conservación que ha mantenido PEAMT entre las especies, este trabajo se basó en la comparación de las secuencias genéticas y proteicas de PEAMT, para mantener un indicio en su función. También se continuo con trabajo previo realizado en el aislamiento de mutantes loss-of-function para los genes de maíz por medio del sistema de transposones: *Activator/Dissociator* (*Ac/Ds*) y Mutator (*Mu*).

Contents

List of Figures

List of Abbreviations

Ac	***Ac***tivator
Cho	**Cho**line
cDNA	**c**omplementary **D**eoxiribo**n**ucleic **A**cid strand
CDS	**C**oding **D**NA **S**equence
DNA	**D**eoxiribo**n**ucleic **A**cid
Ds	***D***i***s***sociation
Mu	***Mu***tator
RNA	**R**ibo**n**ucleic **A**cid
PCho	**P**hospho**cho**line
PCR	**P**olymerase **C**hain **R**eaction
PDME	**d**i**m**ethyl-**p**hospho**e**thanol**a**mine
PE	**P**hospho**e**thanolamine
PEAMT	**P**hospho**e**thanol**a**mine *N*-**m**ethyl**t**ransferase
PMME	**m**ono**m**ethyl-**p**hospho**e**thanol**a**mine
PtCho	**P**hospha**t**ydil**cho**line
TE	**T**ransposable **E**lement
xpl	***x****ipotl*
Xpl	***X****ipotl* (loci annotation, MaizeGDB[1])

[1]MaizeGDB Nomenclature, http:www.maizegdb.orgnomenclature#LOCI

Introduction

Phosphatidylcholine is the main structural membrane phospholipid

Eukaryotic cells are constituted by four classes of molecules: carbohydrates, lipids, proteins and nucleic acids. Each one is an essential source of energy, structure and function (Cooper 2000). The lipids provide the basis for the cells structure membrane bilayer due to their amphipathic composition (Schwertner and Biale 1973).

The membrane lipids are structured by: 1.a polar head group formed by a hydrophilic region with a residual group esterified to a 3 carbon backbone, and 2.a hydrophobic region conformed by fatty acyl side chain(s). The membrane lipids classification depends on the residual group and the fatty acyl chains quantity and length. The *sn3* backbone can be esterified to different residual groups like: galactolipids (galactose), sulfolipids (sulfoquinovose) and phospholipids (phosphate-containing group). The hydrophobic region may have one or two fatty acyl chains, in the *sn1* and/or *sn2* positions and their length varies between 16:0 or 18:0/18:1 (carbon/insaturations), as shown in figure 1.

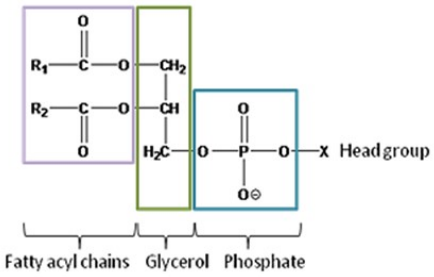


Figure 1: The phospholipids are composed by a phosphate group and 2 fatty acyl chains esterified to a glycerol backbone.

The most abundant phospholipid in membranes bilayer is phosphatidylcholine (PtCho), which accounts approximately for 55% of total lipids (Schwertner and Biale 1973). It's structured by a phosphocholine (PCho) polar head and two fatty acyl tails esterified to the *sn1* and *sn2* glycerol backbone positions and variable insaturations. Like other phospholipids, the PtCho polar head is mainly synthesized in the Endoplasmic Reticulum, then, exported to plastids outer membranes where is esterified to a diacylglycerol (Ohlrogge and Browse 1995; Bishop and Bell 1988; Benning 2008). In plants, PtCho breakdown can lead either to the synthesis of phosphatidic acid (PA), which is employed as a signaling molecule, to the PCho recovery during phosphorous starvation, or choline recovery for osmoprotectants synthesis (Mcneil et al. 2001; Tjellström et al. 2008; Munnik and Testerink 2009).

PEAMT perform the triple methylations from PE to PCho in the PtCho biosynthesis pathway

The PCho biosynthesis follows different routes among eukaryotes. The first pathway described for phosphocholine biosynthesis was the Kennedy pathway (or CDP-choline pathway), which resides in the phosphorylation of choline (Cho) by the choline kinase to synthesize PCho (Gibellini and Smith 2010). The second pathway is the triple methylation pathway which is the major route supply for PCho biosynthesis in plants. It follows the triple sequential methylation of the PE phosphogroup by the enzyme phosphoethanolamine N-methyltransferase (PEAMT) with S-adenosylmethionine as the donor of methyl groups, and monomethyl-phosphoethanolamine (PMME) and dime thyl-phosphoethanolamine (PDME) as intermediate substrates to finally synthesize PCho, shown in figure 2 (BeGora et al. 2010; Cruz-Ramírez et al. 2004; Bolognese and McGraw 2000; Nuccio et al. 2000; Mcneil et al. 2001).

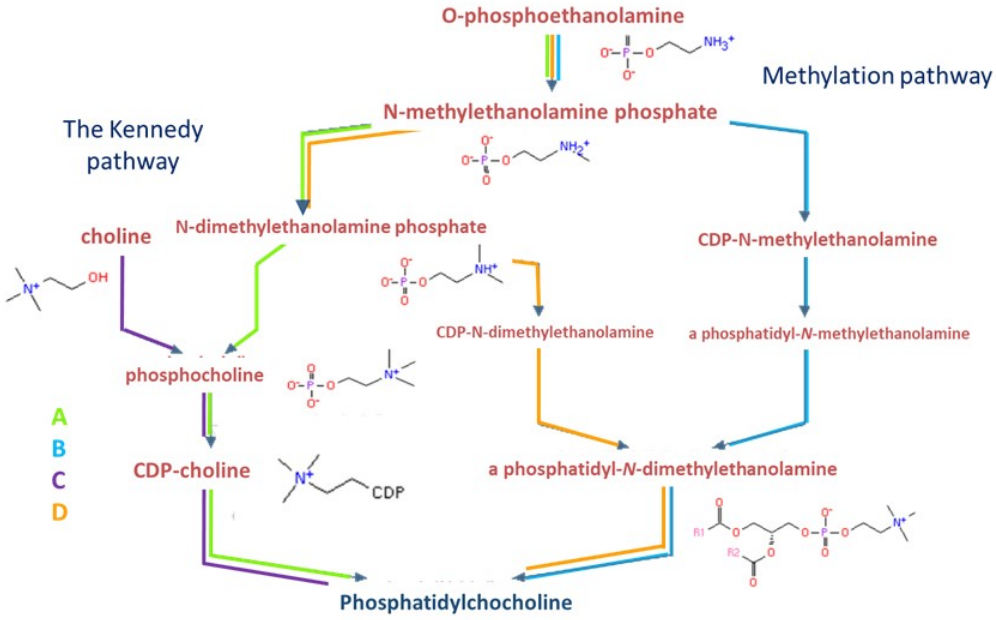


Figure 2: Phosphocholine biosynthesis: the Kennedy pathway (purple), the triple PE methylation (green), the phosphorylation, diacyl-glycerol transfer, and PMME (blue) or PDME (orange) methylation.

In both pathways, the phosphorilcholine transferase phosphorilate PCho to form cytidine di-phosphocholine (CDP-choline), with cytidine triphosphate (CTP) as phosphate donor. Finally, CDP-choline is transferred to the 1,2-diacyl-*sn*-glycerol backbone and the acyl chains to synthesize PtCho. Alternative routes in PtCho synthesis consist in the phosphogroups phosphorylations, then the 1,2-diacyl-*sn*-glycerol transferences and finally the methylation of PMME and PDME, shown in figure 2 (Thompson 1980; Datko

and Mudd 1988; Ohlrogge and Browse 1995; Hocquellet et al. 2005; Lee and Jez 2017).

The PEAMT enzyme conserves structural transmembranal domains (EE) and two methyltransferase functional domains (MT1, MT2) with 4 SAM binding motifs (GxGxG; I, post I, II and III). The PEAMT function among eukariotes is determined by the MT domains presence, in plants PEAMTs are considered as a type I PEAMT enzyme, where both domains perform different methylation steps and seem to function independently (Nuccio et al. 2000; Cruz-Ramírez et al. 2004; BeGora et al. 2010; Lee and Jez 2013; Lee and Jez 2017). Meanwhile the *Plasmodium* type II PEAMT has one functional MT2-like domain able to perform the tree methylations, and nematodes type III PEAMT has one functional MT domain and a vestigial variant of the other MT domain (Lee and Jez 2013).

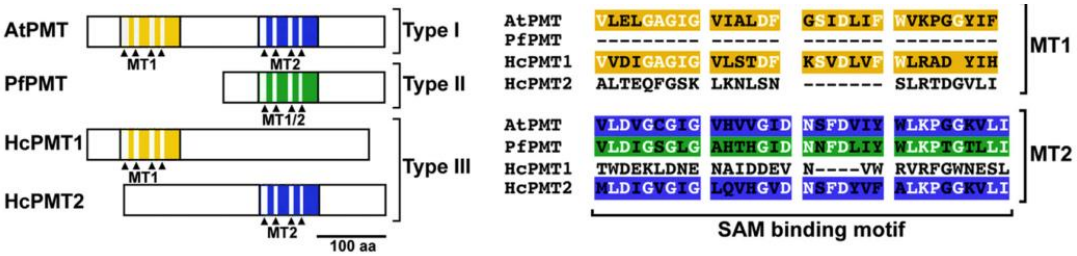


Figure 3: PEAMT MT1 and MT2 domains in *Arabidopsis thaliana* (AtPMT), *Plasmodium falciparum* (PfPMT) and *Haemonchus contortus* (HcPMT) (Lee and Jez 2013)

.

The MT domains are independent due to differences in the protein SAM binding sites in catalytic residues, leading to changes in participation in the methylation steps. The enzyme with a functional MT1 domain is able to perform the tree sequential methylation steps starting with PE; whereas, the one with the MT2 domain, catalyzes the 2nd and 3rd methylations with PMMEA and PDMEA to PCho (Nuccio et al. 2000; Charron et al. 2002; Lee and Jez 2017). The plant PEAMT MT domains structure, function and conservation is diverse between land plants and algae, however, the MT domains sequence are highly conserved (Datko and Mudd 1988; Nuccio et al. 2000; Charron et al. 2002; Wu et al. 2007; BeGora et al. 2010; Lee and Jez 2013; Sato et al. 2016; Lee and Jez 2017). Instead, the MT2 may had appeared after a gene duplication and then fusion with the first gene containing the MT1 domain leading to one di-domain PEAMT protein (Sato et al. 2016; Hirashima et al. 2017; Lee and Jez 2017).

The first gene described in plants related to the triple methylation pathway is the AT3G18000 (PEAMT, *xpl1*) (Nuccio et al. 2000; Cruz-Ramírez et al. 2004). The mutants are named *xipotl* (”to increase in size”, *xpl*) after the swollen root epidermal cells

phenotype shown in the mutants initial growth. The mutation is generated by a *xpl1* 7th exon insertion and causes apoptosis in the root meristematic cells, leading to the primary root development disruption (Cruz-Ramírez et al. 2004). Two paralogs genes are reported in Arabidopsis: AT1G48600 and AT1G73600, which enzymes show preference for the 2nd and 3rd methylation steps from PMME to PDME in PCho biosynthesis. This genes mutants don't present a notably short root phenotype, like xpl1, as shown in figure 4 (BeGora et al. 2010; Nuccio et al. 2000; Lee and Jez 2017).

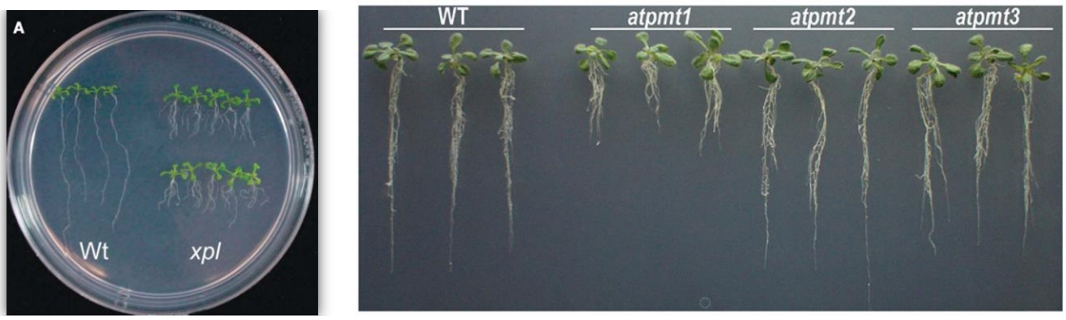


Figure 4: The *At xpl/atpmt1* mutant shows short root phenotype in comparison with the two homologous *atpmt* mutants (Cruz-Ramírez et al. 2004; Lee and Jez 2017).

The major AtPEAMT ORF is highly expressed in roots than in leaves and may lead to tissue specific expression (Alatorre-Cobos et al. 2012; Craddock et al. 2015). Even though not all the *xpl* ORFS expression has been described, *xpl2* is more expressed in rosettes leaves and *xpl1* in roots meristem (Cruz-Ramírez et al. 2004; Craddock et al. 2015; Lee and Jez 2017). Data from The Arabidopsis Information Resources (TAIR[2], 2017) also shows different pattern expression between *xpl* genes during ontogeny.

Also in *Arabidopsis*, the major PEAMT mRNA has an open reading frame (uORF30) which represses the PEAMT transcription in the presence of exogenous choline (Tabuchi et al. 2006; Alatorre-Cobos et al. 2012). This post-transcriptional mechanism is not fully understood, so far, it's known that the uORf presence mediates the PEAMT repression by PA and a mutated uORF leads to a wt PEAMT transcript level recovery (Tabuchi et al. 2006; Eastmond et al. 2010; Alatorre-Cobos et al. 2012; Craddock et al. 2015).

Beside the post-transcriptional uORF regulation, the PtCho synthesis pathway has different known regulatory mechanisms. The rate-limiting steps are the synthesis of Cho, PCho, and PMME, which allosterically regulates PEAMT and enzymes involved in the pathway, change their expression and lead to translational repression (Thompson 1980; Nuccio et al. 2000; Alatorre-Cobos et al. 2012). The immediate PEAMT

[2]TAIR, https:seqviewer.arabidopsis.org

metabolites regulate the PEAMT translation in presence of exogenous Cho and PCho without affecting the PEAMT mRNA transcript levels (Tabuchi et al. 2006; Craddock et al. 2015). Additionally, Cho and PCho repress the Phosphatydilserine synthase required in the synthesis of Serine, the precursor of PE (Thompson 1980).

Most of this regulatory mechanisms are highly conserved in bryophytes, dicots and monocots (Alatorre-Cobos et al. 2012).

Zea mays Xipotl/PEAMT family is constituted by 4 gene members

In maize, 4 genes are predicted to codify a PEAMT enzyme related to the PtCho biosynthesis pathway, according to the Plant Metabolic Network (PMN[3], 2017), bioinformatic analyses and experimental evidence: *xplA* (GRMZM2G060886), *xplB* (GRMZM2G170400), *xplC* (GRMZM2G122296, PEAMT1) and *xplD* (GRMZM2G045249), the last one with less homology than the first three (BeGora et al. 2010; Wu et al. 2007; Rodriguez-Gomez 2013).

The three *Zm* PEAMT protein sequences conserve the SAM-binding motifs in the MT1 and MT2 domains (Alatorre-Cobos et al. 2012; Rodriguez-Gomez 2013). From the *Zm* PEAMT sequences, *xplA* and *xplB* PEAMT MT1 domains are more similar to the *At* MT1 domain SAM binding motifs, which is the one able to perform the triple methylation steps. In the case of *xplC* PEAMT, it seems the less similar to the other *At* PEAMTs and even to the *Zm* PEAMTs (Lee and Jez 2017).

There's only experimental evidence for *Zm xplC* (PEAMT1) to codify a MT enzyme. When the *xplC* CDS is over-expressed in At Columbia ecotype plants, root length, shoot length and silique number are significantly higher in the over-expressed than the wild type, under a salt stress treatment (Wu et al. 2007). In addition, the *Zm xplC* promoter (PEAMT1) presents *cis* regulatory elements related to stress (Niu et al. 2018)

Like the *Atxpl* tissue expression, there's variation between the *Zm Xpl* genes expression within tissues and during ontogeny (Stelpflug et al. 2016). In general, in maize *xplA* and *xplB* are highly expressed in comparison with *xplC*. Moreover, *xplA* is more expressed in roots while *xplB* is in leafs (Lee and Jez 2017). However, in specific root tissues *xplA*, is not always the highest: in the differentiation zone *xplB* is more expressed even at different growth stages, meanwhile, in the meristematic zone, the

[3]PMN, https:www.plantcyc.org

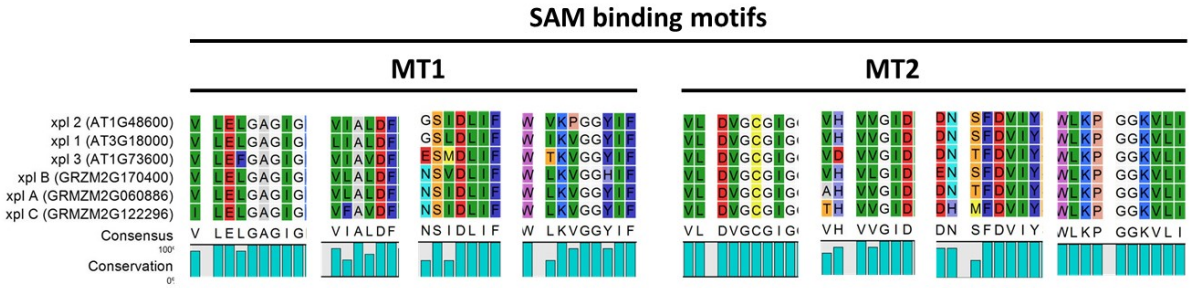


Figure 5: MT domains alignment between the At and Zm SAM binding motifs.

transcripts are higher for *xplA* (Data obtained from Maize GDB [4]), figure 6(Stelpflug et al. 2016).

The regulatory mechanisms of *ZmXpl* genes, like in other plants, is not fully understood. What's confirmed is the uORF conservation upstream the *xplA* and *xplB*, but, for *xplC* it's not quite clear if the peptide sequence may act as a regulatory element beside the upstream predicted TF binding motifs (Wu et al. 2007; Alatorre-Cobos et al. 2012; Rodriguez-Gomez 2013). And, the (*ZmPEAMT1*) promoter sequence presents *cis*-regulatory elements like transcription factors binding motifs differentially expressed among tissues during stress (Wu et al. 2007; Niu et al. 2018).

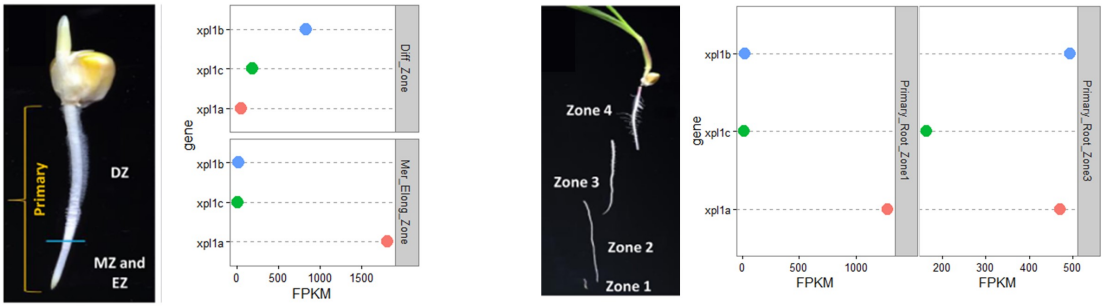


Figure 6: The *Zm* PEAMT family genes expression varies between root tissues (Data obtained from Maize GDB) (Stelpflug et al. 2016)

Due to the known PEAMT biochemistry, the different tissue pattern expression and the number of *Zm Xpl* homologous genes, the *Zm xpl* genes seem to codify more than 1 functional PEAMT able to perform the three methylations from PEA to PCho, leaving two scenarios for the *Zm xpl* genes: it could've been either gene redundancy or a tissue-specific PEAMT expression.

[4] MaizeGDB, https:www.maizegdb.orgexpression

Maize reverse genetics using the transposons system to generate *XplI* mutants

In order to elucidate the PEAMT activity in maize plants, insertional mutants were isolated with the transposons systems *Activator/Dissociation* (*Ac/Ds*) and Robertson's *Mutator (Mu)* previously in the lab. The *xplA* mutants, were isolated with the *Ac/Ds* system and *xplB* with the Mutator system.

The transposons are DNA fragments that are able to insert itself, copies of it in the chromosome or in different chromosomes (Griffiths et al. 1989). Transposons were first described by Barbara McClintock, when she observed a chromosome 9 breakage in maize. In the experiment, they observed the excision of the chromosome leaded to losing one of its extremes where 3 marker genes were located. One of those was the C gene, which participated in the anthocyanin biosynthesis in the aleurone causing the "lack of" or spotted pigmentation over different kernels in one ear(Mcclintock 1941). The excision was due to an element she called *Disociation (Ds)*, further research showed the *Ds* needed another element to break or "jump in" the chromosome, called *Activator(Ac)*. Both elements integrate the *Ac/Ds* transposons system, figure 7 (McClintock 1950).

The transposons system consist in an autonomous and a nonautonomous element. Mostly, an autonomous transposon is composed by a transposase coding region, inverted repeated sequences (11bp) and regulatory elements at their extremes. The autonomous elements codifies for a transposase which allows the TE excision and insertion during cellular division, meanwhile the nonautonomous element lack of transposase activity due to changes or deletions in the derivative sequences, or due to variations in the regulatory elements and coding sequences, therefore, it requires an active transposase to move over the DNA (Du et al. 2011).

When the TE system mobilizes, the transposase cohesively cuts 8bp for *Ac/Ds* and 8bp for *Mu* DNA strands during insertion, leaving the inserted TE element with inverted repeated sequences at the end of both extremes. When the nonautonomous element is mobilized, the result is a footprint that causes DNA damage (Zhang et al. 2009). The footprint may knock-out a gene expression, lead to the translation of a protein variant or to gene derivatives with a reduced number of transcripts (Alleman and Kermicle 1993).

There are collections for *Ac/Ds* and *Mu* tagging, these collections consist in useful features to identify and mobilize TE in the maize genome, specifically in one or more

genes of interest and be able to characterize them (Conrad and Brutnell 2005; Mccarty et al. 2013).

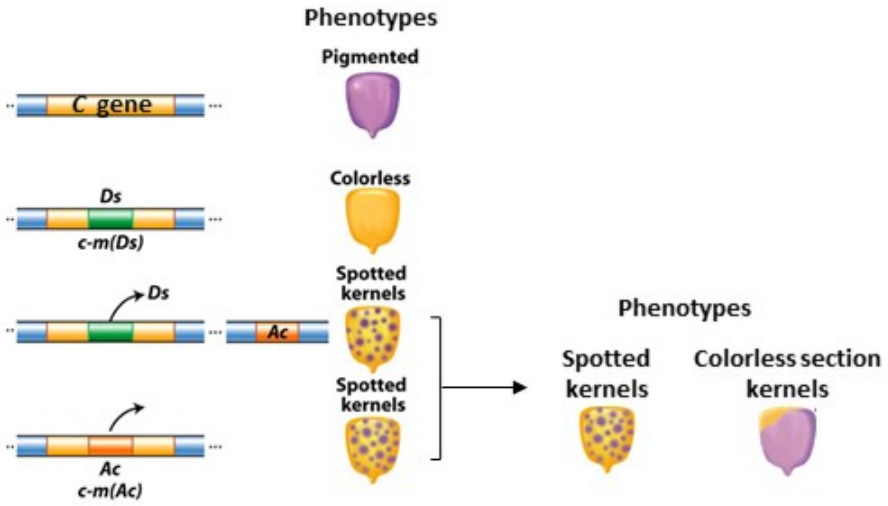


Figure 7: The TE presence/absence lead to spotted, pigmented or colorless kernels. Modified from (Griffiths et al. 1989)

Ds mobilization in the *ZmxplA* mutants

A resource for maize mutant generation is the maize *Ac/Ds* collection, based in maize lines containing *Ds* stuck insertions in the genome, only activated by an *Ac* endogenous transposase, which can be follow by and an stable immobilized *Ac(Ac-im)* insertion in the *R1* gene responsible of the anthocyanin synthesis in the aleurone. Each line is homozygous for a single *Ds* element. The lines are generated in a W22 background and the *Ds* insertions locations are predicted by alignment to the B73 reference Maize Genome (Conrad and Brutnell 2005; Vollbrecht et al. 2010). This collection database is available on the Ac/Ds Tagging Database[5].

According with the Ac/Ds Tagging Database, the closest *Ds* spanning the *xplA* gene is 1.3kb downstream with a 3'-5' orientation, the *Ds* barcode is I.S07.2153 and is inserted in the chromosome 8 from 152591422..152597794. Taking advantage of this genetic tools for reverse mutagenesis, the crosses with this line may promote a *Ds* insertion in the different gene locations and different possible mutated alleles, in order to understand the role of *xplA* in the PtCho synthesis pathway and in the Xpl tissue specific pattern expression.

[5]Ac/Ds Tagging:Genome-wide mutagenesis of Maize using Ac/Ds transposons database, http://acdstagging.org/

ZmxplB and the 5 lines with different *Mu* insertions

Different to the *Ac/Ds* system, *Mutator* has a higher number of copies and transposition rate resulting in a major advantage for the pursuit of the gen of interest. The Mutator elements are the autonomous element *MuDR* and the nonautonomous element *Mu*. The insertions are identified also by seed selection and spotting phenotype (Mccarty et al. 2013).

There're two collections for *Mutator* tagging: TUSC and *Uniform Mu* (McCarty and Meeley 2009). The UniformMu collection consist in populations of B73 and W22 with an partially deleted *MuDR* element, which can be mobilized by an exogenous *MuDR* (McCarty et al. 2013).

According with the Maize B73 genome v3, from the Maize Genetics and Genomics Database[6], there're 5 homozigous lines carrying a *Mu* insertion in the *xplB* gene and their location in the genome, shown in the table 1. The generation of the different *xplB* mutants alleles, may give an insight in differences between transcripts, protein structure and catalytic activity, or gene regulation. Beside, taken together with the *XplA Ds* mutant we could have a bigger picture of the *ZmXpl* family genes.

Table 1: Mu insertions in *xplB* (GRMZM2G170400)

Mu insertion	**Position**
mu1015640	162503509 - 162503517
mu1039328	162505952 - 162505960
mu1017886	162507456 - 162507464
mu1015641	162507628 - 162507636
mu1081615	162507659 - 162507667

[6]MaizeGDB, https:https:maizegdb.orggbrowsemaize_v3name□RMZM2G170400_T01

Justification

The triple methylation pathway is the main source for phospholipids biosynthesis, where the PEAMT enzymes perform three methylations from PE, PMME, PDME to PCho. In plants, these methylations depend on the 2 PEAMT methyltransferase domains structure, sequence and arrangement. Although, the PEAMT genes sequence and regulation also influences the PCho biosynthesis.

So far, in *Arabidopsis thaliana* the *PEAMT/xpl* genes and proteins are well characterized. The identification and gene expression evidence for *xpl* genes support the PEAMT methyltransferase known activity and their relationship with salt, drought, oxidative and cold stress. Also, the PEAMT didomain enzymes structure is highly conserved in different plant species from mosses, monocots and dicots.

Based on this evidence, the *AtPEAMT* sequences characterization will give an insight in the PEAMT activity in other plants, for example: *Zea mays*. It's well know the relevance of this crop added to the importance of the phospholipid biosynthesis. In maize, there're 4 genes predicted as *xpl* genes with methyltransferase domains. However, the investigations have been directed to 1 *xpl* gene (*PEAMT1*) overexpression and transgenes generations to approach their use. This gene from the *At* and *Zm PEAMT* sequence analysis, is the less like from *AtPEAMT1* and show degradation in the regulatory elements on the promoter.

For these reasons, in this work, I will characterize the *ZmXpl* genes and protein sequences in comparison with the *Atxpl* evidence to approach their activity, regulation and conservation, altogether with the generation of *Zm* loss of function mutants and the *At::Zm* transgenes generation for future PEAMT characterizations.

Hipotesis

The *AtPEAMT* family genes and proteins characterization will give an insight in the *Zm-PEAMT* genes regulation, conservation and their phosphoetanolamine n-methyltransferase enzymes function.

Objectives

Aim

To characterize the *Zea mays* phosphoetanolamine n-methyltransferase family genes (*Xipotl*, PEAMT) and protein sequences in comparison with the *Atxpl* family genes, in order to approach the *ZmXpl* family genes and proteins function, regulation and conservation in maize.

Objectives

The objectives will establish an approach in the Zm PEAMT enzymes characterization:

Analyze the *ZmXpl* and *Atxpl* family genes and protein sequences.

Construct transgenes with the *Atxpl1* promoter and the *ZmXpl* coding sequences.

Identify *ZmXpl* family genes mutants with the transposons system.

Materials and Methods

Identification of *Zm* PEAMT genes

Based on latest studies and my confirmation, 3 *Zmxpl* genes: GRMZM2G06886 (xplA), GRMZM2G170 400 (xplB) and GRMZM2G122296 (xplC) were identified as homologous of the *Atxpl* (PEAMT) family genes based in the Plant Metabolic Network (PMN, 2016)[7] (Rodriguez-Gomez 2013). Following this selection and to confirm the recent sequences versions, I verified the sequences and gene annotations from The Maize Genetics and Genomics Database with the B73_reference genome V3 [8] and download them from Gramene [9].

I confirmed the *At xpl1* (PEAMT) gene paralogous sequence based in Cruz-Ramirez (2004) work, also the *xpl1* sequence and gene annotations from The Arabidopsis Information Resources [10] and downloaded the sequences from The National Center for Biotechnology Information Database [11]. I aligned the AtXPL2, AtXPL3, ZmXPLA, ZMXPLB and ZmXPLC sequences with the AtXPL1 sequence using the CLC Genomics Workbench Version 9.0 [12] with the default parameters, in order to confirm the MT domains and SAM binding motifs conservation, shown in Results (Lee and Jez 2013).

At xpl mutants complementation

Plant growth and Tissue collection

To obtain the *At xpl 1* promoter, sterilized and treated *Arabidopsis thaliana* Columbia 0 seeds as described below in AppendixB. Then, the seeds were placed in MS media A on Petry dishes and were maintained in a growth room for 10 days. The leaves and roots were collected separately in 1.5 ml tubes and frozen with liquid Nitrogen for DNA extraction.

For the *Zm Xpl* family genes CDS cloning, *Zea mays* B73 seeds were hydrated in 50ml Falcon tubes with distilled water for 48 h. After that, the seeds were grown under

[7] PMN, https:pmn.plantcyc.org
[8] MaizeGDB, https:www.maizegdb.org
[9] Gramene, http:ensembl.gramene.orggenome_browserindex.html
[10] TAIR, https:seqviewer.arabidopsis.org
[11] NCBI, https:www.ncbi.nlm.nih.gov
[12] QIAGEN Aarhus A/S www.clcbio.com

normal and low P conditions. The seeds were planted in pots filled with sand and vermiculite and were grown in a growth chamber (14h light/$20\,^{\circ}\mathrm{C}$, 10h dark/$10\,^{\circ}\mathrm{C}$, Percival Scientific) until the tissue collection. The low P condition was carried with deferential irrigation with Hoagland nutritional solution adjusted with 1mM (normal conditons) or 0mM (lowP). At the V4 stage, the leaf and root tissues were collected in aluminum folds, frozen with liquid Nitrogen and kept at -$68\,^{\circ}\mathrm{C}$ for RNA and DNA extraction.

DNA extraction

The DNA was extracted from the collected *At* Columbia 0 frozen tissue following a Phenol:Chloro- form:IAA extraction protocol B. To obtain RNA-free DNA, the samples were treated with 0.5 μl of RNase A (20mg/ml, Invitrogen) and incubated for 30 min at $37\,^{\circ}\mathrm{C}$. DNA concentration was quantified as described below.

RNA extraction and cDNA synthesis

The RNA was extracted from the collected *Zm* B73 frozen tissue following a TRizol (Thermo Fisher Scientific) extraction protocol modified from the technical sheet B. To ensure DNA-free RNA, the samples were treated with the Deoxyribonuclease I (Amplification Grade, Invitrogen) to remove possible remained DNA fragments. The protocol was performed as described by the enzyme technical sheet. After that, to analyze the integrity of the DNA-free RNA, the protocol was followed using formaldehyde to denature the RNA, ethidium bromide to stain it, and an electrophoresis through agarose gels to separate the resulting formaldehyde-RNA-ethidium adducts (Sambrook, Fritsch, and Maniatis 1989) .

DNA and RNA quantification was determined by spectrophotometry by Nanodrop, with 50 nm for DNA and 40 nm for RNA. The first strand cDNA was synthesized according to the SuperScript™ II RT Thermo Fisher protocol, starting with 5 μl of DNA-free RNA.

At promoter and *Zm* CDS amplification

For each of the components, *At* promoter and *Zm* CDS, the PCR product was amplified with a High Fidelity Accuprime Taq polymerase (Sigma-Aldrich) as described in the technical sheet.

Table 2: Sequences added to the PCR primers with the recognition sites and overhangs for the restriction enzyme BsaI

Name	RECOGNITION / OVERHANG / Sequence	Description
B_At318000F	GGTCTCCa**GTAG**gcagcgatccttacgtcgtcgt	Prom *At xpl 1*
B_At318000R	GGTCTCa**CCAT**ttcggaaatgtcgtttgtcggga	Prom *At xpl1*
B_Zm060886F	GGTCTCa**ATCA**cgcaacgacctaccgtgatg	CDS *Zm xplA*
B_Zm060886R	GGTCTCg**TCC**Atcatttggtggcgatgaac	CDS *Zm xplA*
B_Zm170400F	GGTCTCt**ATGG**acaccgtcggcgtccccgtggt	CDS *Zm xplB*
B_Zm170400R	GGTCTCg**TCCA**tcacttggtggcaatgaacaaccccc	CDS *Zm xplB*
B_Zm122296F	GGTCTCt**ATGG**ccgccgccgccgccgctgtgaat	CDS *Zm xplC*
B_Zm122296R	GGTCTCg**TCCA**tcacttggtggcgacgaacagccccca	CDS *Zm xplC*

The *At xpl1* promoter was selected 2kb upstream from the initial codon, ATG, considering the Upstream Open Reading Frame 30 as part of the promoter sequence. For the CDS, the enzymes recognition sites and the overhangs were taken as part of the primer sequences together with ATG's, prior to Golden Gate two-parts cloning as described below. The primers sequences signaling the recognition site, overhangs and specific *Zm Xpl* CDS sequences are shown in the Table 2.

Ligation and *E. coli* transformation

The ligation reaction was realized as described on the Promega protocol for the pGEMt-easy cloning vector, the incubation was overnight at $4\,^{\circ}\mathrm{C}$. The transformation was through a modified heat-shock protocol and selected by antibiotic resistance (carbenicillin). The ligation product was added to an aliquot with top 10 *Escherichia coli* competent cells (Storage at -$80\,^{\circ}\mathrm{C}$ in 1.5 ml tubes), the tubes were incubated 10 min on ice and heat shocked at $42\,^{\circ}\mathrm{C}$ for 45 s. After that, the tubes were incubated for 5 min on ice and 200 μl of LB medium were added to the tubes. The transformation product was incubated and shaken on the Thermomixer for 1 hr at $37\,^{\circ}\mathrm{C}$. On a petry dish with LB medium and carbenicillin 100 μg/ml, 200 μl of the transformation product were plumbed and the dishes were incubated overnight at $37\,^{\circ}\mathrm{C}$.

Isolation of Plasmid DNA

The grown colonies resistant to carbenicillin were selected for plasmid isolation, following the procedure by Non-Ion Detergents miniprep modified from (Lezin et al. 2011). To analyze the integrity of the plasmid DNA, quantification was determined by spectrophotometry at 50 nm with a Nanodrop. For the confirmation of the ligation, A PCR with specific primers for the pGEMt-easy vector was performed for each of the *At and*

Table 3: Golden Gate ligation construct primers

Name	Sequence	Description
pSE7F	GCGCGCAAACTAGGATAAATT	pSE7 backbone forward primer
pSE7R	ACCCTAATTTCCCTTATCGGG	pSE7 backbone reverse primer
B_At18000ZmF	TGATGCAGCAGAGAGGACG	Fusion site Atprom:Zm/AtCDS forward primer

Zm fragments and a diagnostic agarose gel was run to check successful assembly. The vectors confirmed to carry a fragment with the *Atprom* or *ZmCDS* lenght, were sequenced by Sanger at LABSERGEN in Langebio.

Golden-Gate ligation

The DNA Assembly Protocol is based on the Golden-gate protocol described in appendix B (Emami, Yee, and Dinneny 2013). The assembly product was transformed into 100 μl of competent *E. coli*. The cells were plumbed on petry dishes with 20 ml LB media/agar, 100 μl/ml Spectinomycin, 40 μl/ml Xgal and 100μl/ml IPTG, for colony selection through color screening. Beside, plasmids were isolated as described above and the constructs ligation was confirmed by PCR. The strategy consisted in verify by PCR the fusion between the At promoter and the Zm CDs with the At promoter and the vector specific reverse primer. The primers were designed flanking the BsaI restriction sites in the pSE7 vector and the fusion site with an specific *At xpl1* promoter primer 130 bp up the *Zm Xpl* ATG 3.

Agrobacterium tumefaciens transformation with the *At::Zm* constructions

The plasmids confirmed with the construct ligation, were transformed in *Agrobacterium tumefaciens* and selected by antibiotic resistance (rifampicin, carbenicillin, and spectinomycin) and colony PCR, shown in AppendixB. The transformation was performed through a modified heat-shock protocol B.

ZmxplA::Ds and *ZmxplB::Mu* mutants genotyping

The *Zm Xpl* mutants genotyping consisted in the PCR based identification of the TE elements: *Ds* for *xplA* and *Mu* for *xplB*. Lately in the lab the *Ds* was mobilized in *xplA*, and for *xplB*, the insertions were identified and started with the seed stock generation and genotyping (Rodriguez-Gomez 2013).

Plant growth and Tissue collection

For *XplA::Ds*, the kernels were obtained from 2 different stocks, the first with kernels of the following rows (individuals): RS46.2 (.1, .2, .3, .4, .5, .6, .7, .8, .9, .10); RS46.3 (.1, .2, .3, .4, .5, .6, .7, .8, .9) and RS46.4 (.1, .2, .3, .4); and the second stock with kernels of self-pollinated individuals: RS16 543.4, RS16 543.8, RS16 543.11 and RS16 545.2. For the seed selection, 25 kernels were taken from each individual of each stock *XplA::Ds*. For *XplB::Mu*, the kernels were obtained from the stock with the mu1039328 insertion, with kernels of the following rows: RS15 4259.1 (.1, .2, .3, .4, .5), RS15 4259.2 (.1, .2, .3, .4, .5), RS15 4259.3 (.1, .2, .3, .4, .5), RS15 4259.4 (.1, .2, .3, .4, .5), 4260.1 (.1, .2, .3, .4, .5). This individuals are the parents for the following rows: RS16 641, RS16 642, RS16 643, RS16 644.

The kernels were hydrated for 48h and grown in 2 lt pots soil under greenhouse conditions (16 h day with 25-29.9 C, 8 h night with 25-23 C, 75-80% humidity). After 10 days of seedling, for each individual, 0.8 cm rounds pools were collected from the cotyledon and the first leaf, in 2 ml Eppendorf tubes with liquid nitrogen. Then, the screening protocol was followed as described above with DNA extraction and PCR genotyping. If one the pools was confirmed for carrying the TE insertion in the cotyledon and the leaf, then, each of the plants of the individual was analyzed. For a single plant screening, half of the cotyledon and half of the first leaf were collected and processed in the same way.

PCR based strategy for TE identification

Following the strategy already established, first, DNA was obtained from the leaf tissue collected and the extraction was performed as described above (Salazar-Vidal et al. 2016). Then, the TEs were identified by PCR using specific *xpl* gene primers and specific primers flanking the TE extremes in each PCR reaction. The spatial positions for the TE primers designed are shown in the figure 8. For the *ZmxplA::Ds* mutants genotyping two possible orientations, 3'- 5' and 5' -3', were considered for the *Ds* insertion.

Table 4: Primers used to flank the TE: *Ds* or *Mu* elements and the respective specific *Xpl* gene where is the insertion

	Name	**Sequence**
Ds	JGp3	ACCGATACGATCCGGTCGGGT
	JRS03	CGATCGGGATAAAACTAACAAAATC
	JRS01	GTTCGAAATCGATCGGGATA
xplA	RS443	CGGTTCCAGCCTATAAATGCCACT
	RS440	TTCAAGCAACCAGTGGACATAGCA
Mu	Tir 6	AGAGAAGCCAACGCCACGCCTCATTTCGTC
xplB	ES001	GACCAATAAGCACAAGATCCACGA
	ES002	TTGTATGTGAAGGATGGTGTCACG
	ES003	GTTGTGGCACTCTTTCTCCTTGTA
	ES004	CCTGTCTTTGCTTCCCTCTTACAA

The PCR programs were performed separately for the *Ds* and *Mu* identification (Settles et al. 2007; Salazar-Vidal et al. 2016). For the confirmation of *Ds* insertion integrity, the PCR bands were extracted from the agarose gel following the Quick-Start Protocol with the Gel Extraction Kit (Quiagen), then, sequenced by Sanger at LABSERGEN.

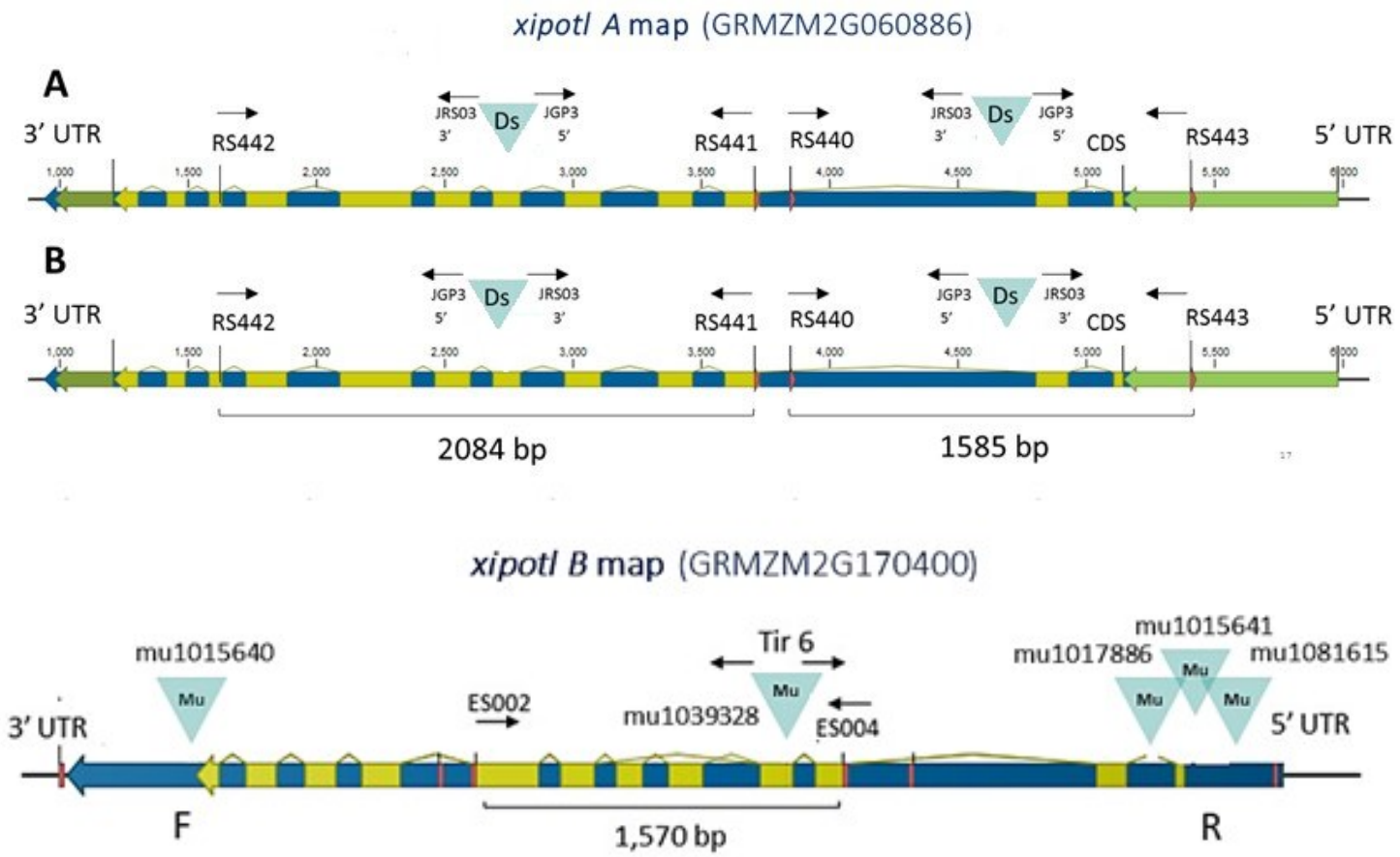


Figure 8: The TE primers locations and orientations in the genes: **A** 3' - 5' and **B** 5'- 3' *Ds* orientations for *xplA* (on the top), and *xplB* (on the bottom).

Results

The At and Zm PEAMTs conserve MT domains

Based on latest studies and my confirmation, 3 *Zmxpl* genes: GRMZM2G06886 (xplA), GRMZM2G170 400 (xplB) and GRMZM2G122296 (xplC) were identified as homologous of the *Atxpl*(PEAMT) family genes, and selected for the work. The *xplD* sequence was predicted only as a methyltransferase enzyme without the 2 MT conserved domains so it was not considered (Rodriguez-Gomez 2013).

The maize *ZmXpl* sequences were found to encode proteins similar to AtPEAMT1, including the conserved MT domain, trans-membrane domain (EE) and SAM-binding motifs (GXGXG). The presence of MT domains in the ZmPEAMT sequences suggested that the maize proteins have the capacity to perform the three methylation steps. I examined MT1 and MT2 domains in the maize proteins for evidence of potentially functionally important amino-acid changes in the catalytic domains. I aligned *At* and *Zm* PEAMT MT domains, finding ∼70% similarity between the *Atxpl* and *ZmXpl* MT1 domains and ∼80% for the *Atxpl* and *ZmXpl* MT2 domains. The *ZmxplA* and *ZmxplB* resulted in a good alignment with a 93.5% similitude in comparison with the *Atxpl1* MT1 domain with just one none conservative mutation H164Y, and 100% with conservative mutations for the MT2 domain. Similar with the *Atxpl1* and *Atxpl2* di-domain with a 100% of similitude with conservative mutations between Valine, Leucine or Isoleucine. Disregarding the fact that *xpl3* and *XplC* aa's sequences had fewer similitudes and pointed mutations in both domains, showed in figure 9.

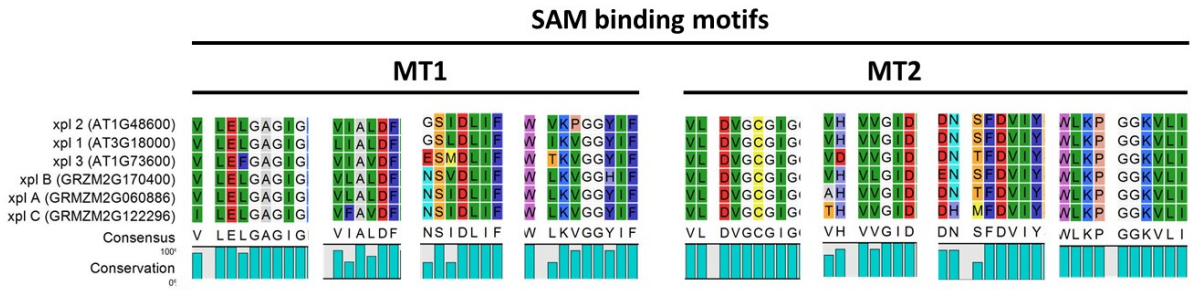


Figure 9: The *At* and *Zm* PEAMT family genes conserved domains.

Initially, as described above, *xpl D* was predicted by the PMN to be another *Zmxpl* homologous gene. However, from 23 transcripts that I obtained from MaizeGDB for *xplD* coding sequences, none of the translated ORFs resulted on a PEAMT sequence,

neither to a conserved SAM binding motif. So, after a nucleotide blast, this gene resulted in a probable methyltransferase for maize, rice, and sorghum and it was discarded for the experiments. I uploaded the sequences and alignments to a GitHub repository [13].

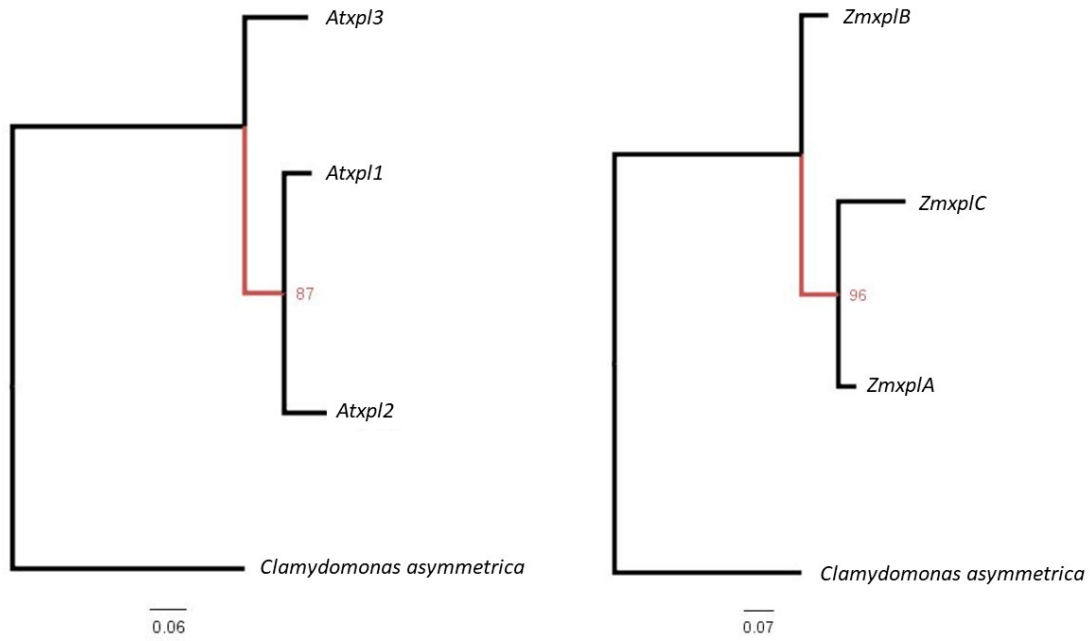


Figure 10: The *At* and *Zm* PEAMT phylogeny.

After the confirmation of the PEAMT MT domains conservation in maize, we decided to investigate the presence of uORFs in maize due to their conservation and possible divergence in ZmPEAMT regulation. With this in mind, I aligned the uORFs to search for differences in the sequence of this regulatory element, shown in the figure 11. The uORFs conserve the the aa's reported previously for *Atxpl1*, and the more similar is the *ZmxplA* uORF, but for *ZmxplB* the R6L, S7A, R10C mutations varies from the *ZmxplA*, *Atxpl1* and *Atxpl2*; and the *ZmxplC* uORF appeared not related (Alatorre-Cobos et al. 2012).

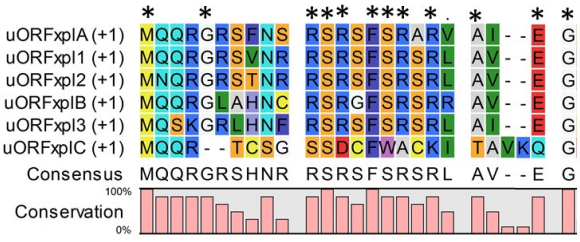


Figure 11: The *At* and *Zm* PEAMT family genes conserved uORFs peptides.

[13] Sawers-Rellan Labs, GitHub organization: https://github.com/sawers-rellan-labs/Sofia-Thesis

The results showed a few synonym mutations in the Zm XPLA and Zm XPLB MT domains, which either those mutations could've affected the protein and its regulation, or could not affect the PEAMT function and resulted in both enzymes capable of performing the three methylations. This scenario opens the possibility of *ZmXpl* redundancy or the tissue-specific expression and modifications in the regulatory mechanism for both genes. This could be probed with the complementation experiment, if both PEAMTs are able to perform the three methylations like *At xpl1* then, the *At xpl1* mutants could recover the wt phenotype, and if the enzymes activities are enough to catalyze the follow methylations, then the *Atxpl1,2,3* triple mutant may recover the phenotype.

At xpl 1,2,3 mutant phenotype in contrast with *xpl 1*

To understand the role of each gene in the PtCho biosynthesis pathway, the *Atxpl* genes were mutated and it resulted in differences between *xpl* mutants root length, this result also was shown in the Lee and Jez (2017) work. In order to analyze the *At xpl* mutants before the complementation experiment, I grew the other *xpl* mutants to observe their phenotype. For the simple mutants besides the characterized *xpl1*, I observed slight differences in comparison with the Wt, the *xpl2* had a larger root and the *xpl3* had a shorter root than the Wt. Meanwhile, the triple mutant had a smaller root than *xpl1*, shown in figure 12. With this characterization, I had a scale for comparison with the transgenic resulted plants.

Figure 12: WT plants, At *xpl1*, *xpl2*, *xpl3*, and *xpl1,2,3* mutants. Scale bar 0.5cm.

The triple mutants also appeared to have a slower growth in comparison with the simple mutants (data not shown) and they also showed aborted or aberrant embryos, this phenotype seemed to be more severe in the *xpl1* and *xpl1,2,3* mutants in comparison with the *xpl2* and *xpl3* mutants. In addition to the aborted *xpl1,2,3* embryos, different abnormalities were observed in the embryos, like a smaller ratio, lack of color (albinos) and amorphous ones, as shown in the figure 13. Further work, must be done to have quantitative data for the *xpl* mutants.

Figure 13: WT, At *xpl1*, *xpl2*, *xpl3*, and *xpl1,2,3* mutants; siliques and embryos.

The generation of *At::Zm xipotl A* and *B* constructions to analyze their PEAMT function

I amplified the *xpl 1* promoter from *At* genomic DNA, I designed the primers to amplify 2,000 bp length upstream the initial codon. And for the CDS, due to the variation in the length of the coding regions, I selected and amplified the transcripts with $\sim$ 1,500 bp, based on the At PEAMT CDS length, 1,476bp. Also, I amplified the *Zm Xpl* genes coding sequences (CDS) from RNA extracted from B73 leaves and roots tissues, according to the gene expression during development data at the Gene Atlas, available in Maize GDB. For the tissue collection, I grew the plants under low P conditions as described in Materials and Methods. I amplified the *Zm xplB* CDS from B73 leaf tissue RNA and the *Zm xplC* CDS from root tissue RNA. Even though I found *Zm xplA* CDS in both tissues, it was more expressed in root tissue. I verified the RNA and cDNA integrity by running agarose gels, shown in the figure 14.

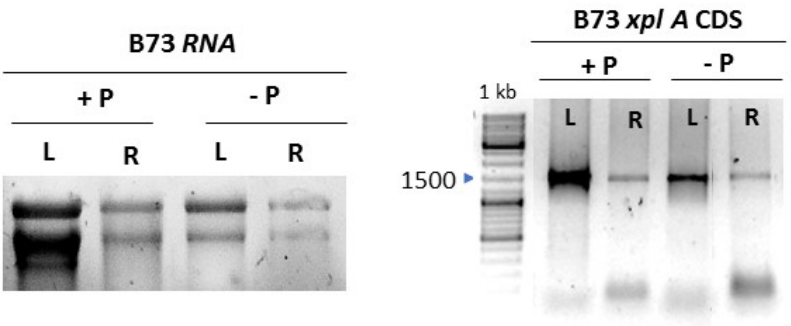
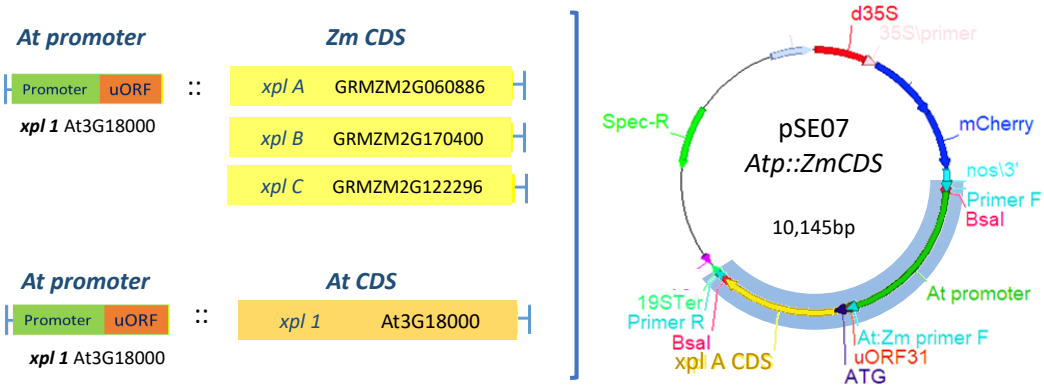


Figure 14: *Zm* B73 RNA and *xplA* CDS expressed from B73 cDNA in leaf and root tissue from plants grown in standard and low P conditions (10mM)

I ligated each of the PCR products and cloned them in E. coli competent cells. I analyzed the clones obtained for *ZmxplB* by PCR to confirm the ligation, and send them to sequencing by Sanger to confirm the CDS integrity; both procedures were described in MnM. For technical reasons, I named the colonies obtained from ligation and E. coli transformation after sequencing. From *ZmxplB* 2 clones (Zmxplb1_F03 and Zmxplb2_H03) were analyzed by PCR giving a ∼ 1,750bp length fragment and then sequenced, due to the alignment % similitudes with B73 *xplB* CDS sequence, I selected the ZmxplB2_H03 to continue the experiments. The ZmxplB2_H03 plasmid sequence alignment resulted in 1,488bp matches and 12 mismatches (with 6 deletions and 6 bp changes). In order to verify that the mismatches didn't change the PEAMT codon sequence, I translated the CDS ORF and the conserved MT domains didn't change.



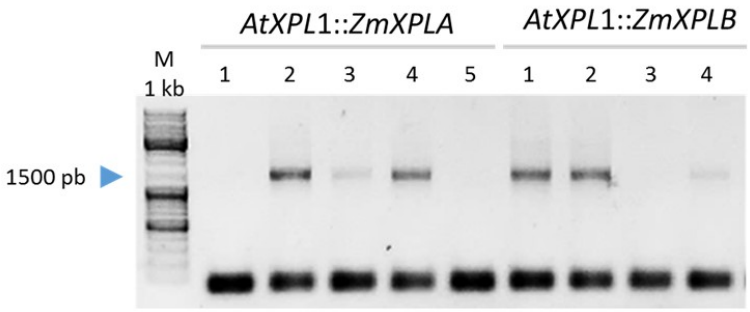


Figure 15: The constructions were built with 2 blocks with the At prom and the *Zm* CDS.

To follow the experiment, I assembled constructions with two-blocks containing the *At xpl1* promoter and the CDS of each of the *ZmxplA* and *ZmxplB*. The constructions were assembled with the Golden Gate cloning strategy. For the control, I also assembled constructions with two-blocks containing the *At xpl1* promoter and CDS, this to confirm that the assembly system didn't affect the constructs or the PEAMT resulting function. In the figure 15, there's a representation of the constructions with At and Zm blocks, and the cloning vector I ligated with the *xplA* CDS. The vector has a mcherry

cassette that gives red fluorescence in the transformed seeds (blue), the At *xpl1* promoter (green), the xplA CDS (yellow) and the representation of the primers I designed to verify the ligation sites and the *At::Zm* fusion sites (Emami, Yee, and Dinneny 2013).

To verify the vector-blocks assembly, I designed vector-specific primers to flank the At promoter beginning and the end of the *Zm* CDS, and to verify the correct fusion between the At promoter and Zm CDS, I designed a specific primer in the At promoter, including the *Atxpl1* uORF, 130bp up the *Zm Xpl* ATG. The E.coli transformed colonies were confirmed to carry the At and Zm blocks by PCR. The fragments amplified within the *At* promoter and the Zm CDS had a $\sim$1,700 bp length, and as a negative control, the specific pSE07 primers amplified a $\sim$700bp fragment if the fusion didn't success. So far, I transformed Agrobacterium competent cells with the following constructions: *Atp::ZmxplA, Atp::ZmplB* and *Atp::Atxpl1*. The *Atxpl* mutants transformation will remain as a perspective.

The *ZmXpl* mutants will give an insight in the tissue specific expression

Beside the Zm PEAMT function, it's still missing the different pattern expression for these genes. For this reason, the generation of individual *Zm Xpl* mutants is so important, first, we could've different phenotypes and genotypes of each gene and understand their role and the plant. So far in the laboratory, we had the stocks described in the Introduction and I continued with the *ZmxplA* and *ZmxplB* genotyping.

An insertional *ZmxplA::Ds* mutant

For the genotyping, I grew seedlings from the seed stocks for *Zm xplA* mutants identification by the *Ac/Ds* transposon system. I confirmed plants from 10 individuals analyzed from RS16 46.2, 9 from RS46.3 and 4 from RS46.4, to carry have a Ds insertion, all from Cornell. From each stock I selected the seed from its dosage and discarded the possible revertants, I hydrated them for 2 days, grew the seedlings for 10 days, and extracted DNA from the cotyledon (**A**) and the first leaf (**B**) to confirm by PCR if the insertion was in both. Then I performed the identification by PCR flanking the *Ds* as described in MnM.

I identified the *Ds* in each of the Cornell seed stocks, in the 2nd intron (1075bp length) after the PCR resulted in complementary fragments with the *xplA* specific primers and the *DS* primers: JRS03-RS440, $\sim$1300bp, and JGP3-RS443, $\sim$600bp. Due to the

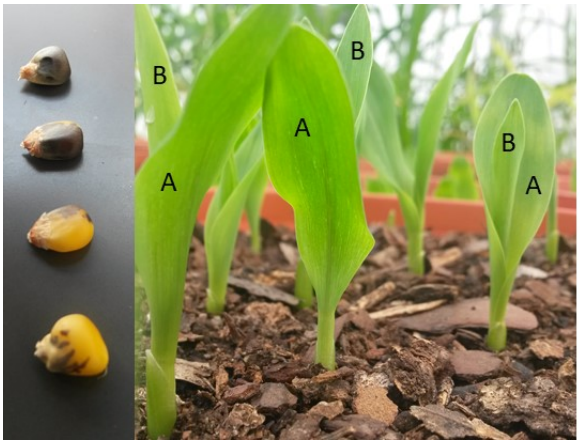


Figure 16: For the DS identification, we selected leaf (B) and the cotyledon (A) tissue from each seedling of each stock grown under greenhouse conditions.

amplification with the JRS03 primer, the *Ds* was confirmed for a 5' to 3' direction and it I track its location in the second intron as shown in figure 17. The map shows a proximal *Ds* location in the xplA gene, the 5' and 3' UTR are represented with green arrows, CDS (exons) with a truncated yellow arrow, introns with blue and the primers with red.

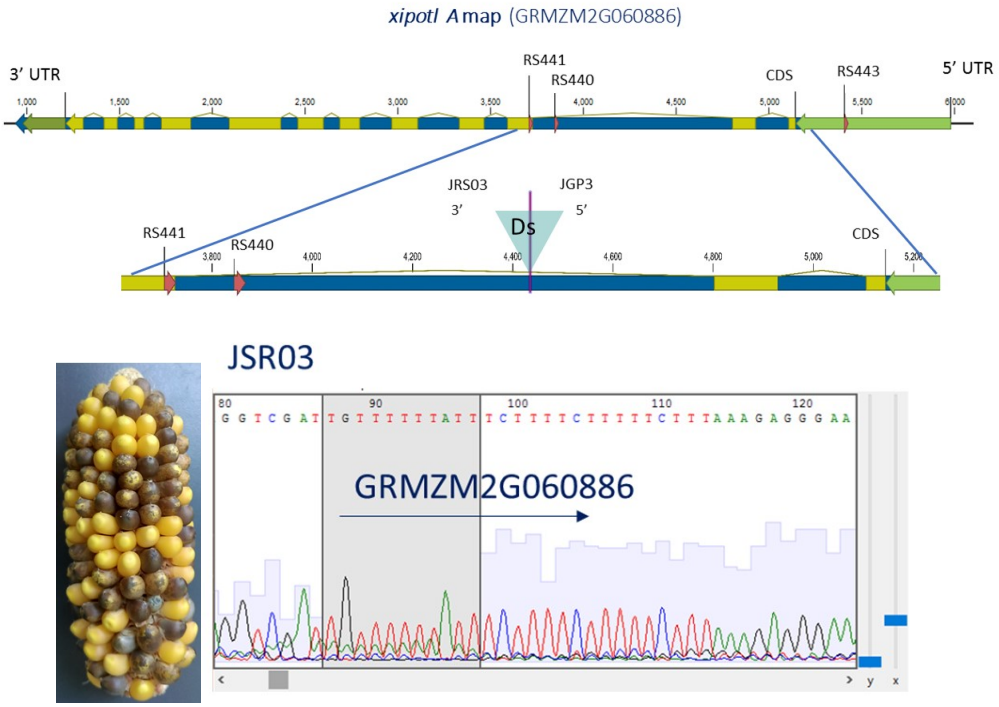


Figure 17: **Up.** The Ds insertion was identified in the 2nd intron. **Down.***xplA::Ds* ear and sequencing

I corroborated the insertion by sequencing the PCR product with the specific *Ds* and *xplA* primers. First, I blasted the out-come sequences to confirm the amplification of the *xplA* gene, and then, to identify the *Ds* location in the gene, I aligned the *Ds*

sequences with the *xplA* sequencing print out, with both senses. Here I present the JRS03 sequencing print out and the signpost of the *xplA* gene. From the *ZmxplA::Ds* mutants to carry the insertion, 3 plants were grown in greenhouse conditions till flowering and back-crossed with T43 as male, to continue the allele segregation. The figure 17 shows an ear obtained from that crosses, which still show the spotted purple/yellow pattern given by transposase doses.

I also genotyped 25 self-pollinated plants obtained from the second stock rows: RS16 543.4, RS16 543.8, RS16 543.11 and RS16 545.2 and were confirmed to carry the same *Ds* insertion. The next image shows the electrophoresis gel with the amplified fragments proposed to carry the same insertion as the Cornell stock. Further work would be needed to confirm by sequencing the insertion in the other seeds stocks, their *Ds* direction and to analyze if there'd more alleles for this *xplA::Ds* mutants. Beside the *Ds* mobilization in the gene to get an insertion in a codifying region or in the uORF sequence.

An insertional *Zm xplB* mutant

Five alleles with *Mu* insertions were recently identified for the *Zmxpl::Mu* mutants, Rodriguez-Gomez (2013) started with the mu1039328 insertion stock genotyping and I continued with it (Rodriguez-Gomez 2013). For the genotyping, I selected seeds from the mentioned stock and I selected 5 individuals from each row: RS15 4259.1, RS15 4259.2, RS15 4259.3, RS15 4259.4, 4260.1. I followed the same protocol from *ZmxplA::Ds* genotyping for the seed selection, tissue collection, DNA extraction, and PCR based identification.

I confirmed an insertion in the individuals analyzed from the stock segregating for the mu1039328 insertion genotyped. I identified the insertion in the 4th exon (138bp). If the footprint resulted in a stop codon, the protein would've 111 aa's length, with half of the MT1 domain leaving it with the I and post-I SAM binding motifs.

These individuals are the parents of the following individuals: RS16 640, RS16 641, RS16 642, RS16 643. As the insertion was confirmed in the stocks, 15 seeds from these individuals were sown to propagate seeds and were out-crossed with CML312 background (Winter 2016, Puerto Vallarta, Mx).

Further work must be done for the generation of the remaining *Mu* alleles. Interestingly, the mu1015641 insertion is located in the uORF, there're seed stocks for this mu insertion: RS15 990, RS15 991, RS15 992, RS15 993, RS15 994, which weren't

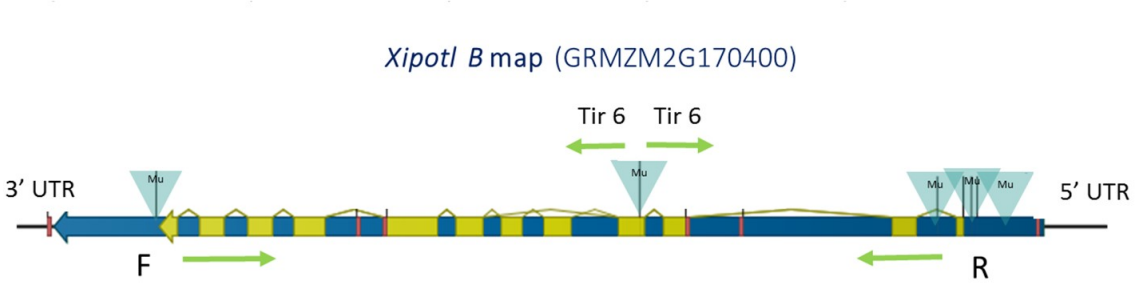


Figure 18: The *Mu* insertion genotyped is located in the 4th exon.

included in this work. However, this mutation could lead to the understanding of the post-transcriptional regulatory element.

Discussion

Summarizing, PtCho could be synthesized by the methylation pathway where PE is triple methylated by PEAMT (*XIPOTL, xpl*) to PCho and then is transferred to diacylglycerol resulting in PtCho. The PEAMT enzyme is highly conserved in plants and its biochemistry is well known; with this in mind, this work was focused in the analysis of the *Zm Xpl* orthologous genes related to the PtCho biosynthesis, the enzymes sequences analysis between the ZmMT similitudes with the AtMT domains and their possible role in the plant. Owing to the fact of the uORF conservancy for 2 *Zmxpl* genes and the tissue-specific expression patterns.

In Arabidopsis, there'd 3 PEAMT homologous genes and, for *Zm*, 4 genes were predicted by homology for a PEAMT enzyme. From the 4 genes, 3 of them (*xplA, xplB*, and *xplC*) present the 2 MT domains sequences and a complete/partial uORF, so the 4th gene was discarded as a PEAMT enzyme, nevertheless it was cataloged as a possible methyltransferase and it could have a different role in the PtCho metabolism.

.

To understand the role of each gene in the PtCho biosynthesis pathway, as described above, the *Atxpl* mutations resulted in differences between *xpl* mutants root length (Cruz-Ramírez et al. 2004; Alatorre-Cobos et al. 2012). This result also was shown in the Lee and Jez (2017) work. Even though this differences didn't appear significant, this may help us to determinate the *ZmXpl* genes complementation effect on the plants.

Surprisingly, the *xpl2* mutant showed less abnormal siliques and embryos in contrast with the *xpl1*, besides the fact that this gene is highly expressed in the leaves, the regulation between vegetative and reproductive tissues is different for PEAMTs, and *xpl1* could be more related in development (Lee and Jez 2017). The triple mutants also appeared to have a slower growth in comparison with the single mutants and they also shown more aborted or aberrant embryos. Even though the mutants were able to grow till flowering, they appeared to be more sensitive to stresses, this shows the importance of PEAMT in the early development and how the lack of the PEAMTs enzymes lead to alternatives routes that compensate the PtCho synthesis, but still the lack of PtCho ends up delaying flowering (Gibellini and Smith 2010; Nakamura et al. 2014).

The Zm and At MT domains similitudes may approach the Zm PEAMT biochemistry and methylation activity

From the PEAMT sequence analysis, first it was hypothesized that the PEAMT function resided in the MT domains, however, the similitudes between the xpl1 and xpl2 domains followed to propose that xpl2 could also perform the first methylation step (Cruz-Ramírez et al. 2004; BeGora et al. 2010; BeGora et al. 2010; Nuccio et al. 2000). Then, recently in an independent work, the 3 At PEAMTs were functionally characterized and it showed that the three At PEAMTs are able to use pEA, pMME, or pDME and SAM as substrates and there's no difference in the preference between them; so, the methylation activity resides in the protein structure. In the AtPEAMT1 the MT domains are oriented side-by-side, whereas in the AtPEAMT2 they have a crossed orientation removing a substrate binding site, as an approach in understanding why *xpl1* PEAMT1 performs the first step in the methylation, and *xpl2* and *xpl3* the following (Lee and Jez 2017).

The PEAMT function resides in the Tyr131 and Tyr254 however, the mutagenesis of this residues didn't change the AtPEAMT2 preference for PEA (Lee and Jez 2013; Lee and Jez 2017). In previous work, the *xpl2* protein showed a lack of 16 aa's in the N-terminal region and it was the reason attributed to the AtPEAMT2 lack of first methylation step, which differs from the sequences selected for this work (BeGora et al. 2010; Lee and Jez 2017). So far, there's still research to do in order to address the protein structure involved in the methylation activity between PEAMTs, starting with a protein prediction for the Zm PEAMT sequences.

From the At and Zm PEAMT aa's analysis, the *xplA* and *xplB* MT1 domains were highly similar with AtPEAMT1 MT1 domain, which could lead to both enzymes able to perform the triple methylations and even though both genes have similar MT domains, the *ZmxplB* uORF had punctual mutations compared with the *ZmxplA*. The out frame mutation in the *xpl B* MT1 domain may have caused a variation in the PEAMT function leaving a partial enzyme activity with a functional MT2 domain, besides it could have a modification like AtXPL2 (Lee and Jez 2017). This could be partially probed with the heterologous complementation experiment, where, if the *xplA* and *xplB* CDS complete the phenotype of the *xpl1* mutant, then there would be 2 enzymes with functional di-domains and probably they could have diverged between the regulatory mechanisms, out-coming in two functional PEAMT spatially differentiated in the plant or leaded to gene redundancy.

In order to complete a whole perspective of the *ZmXpl* role in the plant, the generation of the homozygous mutants will probably show a similar mutant phenotype to the *Atxpl* mutants. In the case of *ZmxplA::Ds* mutants there'd be expected a short root phenotype like the At mutant, and for *ZmxplB::Mu*, the leaves and shoot would be more affected. If *xplA* and *xplB* PEAMTs had tissue-specific expression this will leave *xplC* as the only PEAMT able to methylate PMME to PDME in maize. It showed the less similar MT domains sequences as the other ones, which could lead to a possible scenario like *xpl3* with a possible PEAMT subfunctionalization. Even though *xplC* showed the lowest expression at the Maize Atlas, it appears to get overexpressed under NaCl stress (Wu et al. 2007; Panchy, Lehti-Shiu, and Shiu 2016). The role of this gene would be demonstrated by the generation of *ZmxplA;B* double mutants, either in the PtCho biosynthesis pathway, the uORF regulatory mechanism or in the tissue-specific scenario.

The uORFs could lead the tissue-specific expression

This enzyme its related in the lipids biosynthesis, it has different regulatory levels and due to it's relevance and conservation in plants, several investigation explored PEAMT regulation (Cruz-Ramírez et al. 2004; Nuccio et al. 2000; Charron et al. 2002; Be-Gora et al. 2010; Tabuchi et al. 2006; Jost et al. 2009; Wu et al. 2007; Lee and Jez 2013). The *Atxpl1* uORF30 is highly conserved including the two ZmxplA and ZmxplB genes (Alatorre-Cobos et al. 2012; Rodriguez-Gomez 2013). It's in charge of the post-transcriptional xpl1 mORF repression and its biochemical regulated by Cho, PCho and the PtCho break-down metabolites, like PA (Thompson 1980; Craddock et al. 2015; Eastmond et al. 2010).

The uORF sequence alignment showed that the *Zm Xpl* uORFs share a 50% similitude with the At uORF30. From the Atxpl uORFs, the xpl1 and xpl2 only differ in 2 aa's, and as mentioned above the PEAMT structure doesn't seem to fully affect the MT1 function in the Atxpl1 PEAMT; so, this slight difference could lead to changes in the PEAMT regulation. Also for the ZmXpl uORFs, as expected, the *ZmxplA* uORF is more similar with *Atxpl1* than with *ZmxplB*, discarding *xplC* uORF due to the minimum homology percentage. Even if each of the *At xpl* and *Zm Xpl* genes can't be confirmed as orthologous, this could give an insight into their regulation besides the uORF in the *Zm Xpl* genes.

Lately, *xpl1* uORF was studied with transient expression and the mutated uORF resulted with a GUS $\sim$ fold higher expression than the Wt PEAMT promoter construction. Also, besides the uORF mutation, the GUS construction was more expressed in

roots than in leaves (Craddock et al. 2015). This supports the tissue-specific regulation and expression for the *xpl* family genes, not only for At but for Zm genes, whereas the Xpl diverse expression within tissues may rely on the uORF regulation. Also, Cho and PMME were reported as rate-limiting steps in the PCho synthesis pathway, showing that the synthesis of this substrates allosterically regulates PEAMT expression causing the decrease in the mRNA (Nuccio et al. 2000).

It's not completely understood the *xpl* uORF regulatory mechanism. However, as an approach in a heterologous system, the Zm AtHB1 homologous uORF represses the AtHB1mORF (Ribone et al. 2017). This uORF is conserved and tissue-specific expressed upstream the AtHB1, which represses the mORF by stalling ribosomes. Considering this, leading to the possibility of the same scenario for the ZmXpl uORFs regulatory, there's a chance to predict a similar mechanism for the *Zm xplA* uORF post-transcriptional regulation and evidence to support the *At::ZmXpl* heterologous system (Alatorre-Cobos et al. 2012; Ribone et al. 2017).

To give an approach of the uORFs and its specificity between genes, taking advantage of the *Ac/Ds* and Uniform Mu stocks, the generation of the 2 ZmxplB::Mu mutant alleles will provide the basis to observe if it affects the same way the phenotype like a loss-of-function mutant. Also, with the generation of loss-of-function ZmxplA::Ds mutants, the crosses between this lines with a functional ZmxplA uORF and a functional ZmxplB mORF, would probably show the uORFs specificity for each gene, and if there are variations in the phenotype due to the probable tissue-specific expression.

The *Xpl* genes could result from MT, gene or whole genome duplication events

There're 3 types of PEAMT enzymes and due to its ancestry and function it's hypothesized plant type I, with 2 functional domains, is the ancestor of the 2 other types; the type III found in nematodes share 2 domains however one is functional meanwhile it conserves a vestigial form of the other, and the *Plasmodium falsciparium* type II with 1 functional domain able to perform the 3 methylation steps (Lee and Jez 2013). The evolutionary path between the plant type I and the nematodes type II is not fully understood, however, the comparison between this PEAMTs could give an insight in the structural and catalytic domains involved in the first methylation step. With this, the aa's mutation in the plant PEAMTs would discard the unrelated functional motifs.

The evolution from the plant PEAMT to the nematodes PEAMT also would give an insight in the genes divergence, somehow the *At xpl* genes codify a PEAMT with 2 functional domains (*xpl1*, MT1 and Mt2,), and 2 with one ”vestigial” domain (*xpl2* and *xpl3*,a ”vestigial” MT1 and a functional MT2). The amino acid changes in the *xpl* genes suggest the loss-of-function for the first methylation step, and this would resulted in gene redundancy between the *xpl* genes with a functional MT2 domain, for most plants (Alatorre-Cobos et al. 2012; Sato et al. 2016; Lee and Jez 2017). The gene redundancy for *xpl2* could’ve resulted after the genome duplication and the *xpl3* from a gene duplication of *xpl2* or *xpl1* (Nuccio et al. 2000; Ganko, Meyers, and Vision 2007; Panchy, Lehti-Shiu, and Shiu 2016).

Due to the PEAMT conservancy and probed existence in algae, the comparison with *Clamydomonas asymmetrica* could give us an approach to the ancestry between the At and Zm PEAMTs, where the phylogenetic trees resulted in different events for the At and the Zm PEAMTs, with 2 separated clusters for both At and Zm (Sato et al. 2016). Even though the distance between them varies, they seemed to share a pattern besides the differences reside in the divergence events. For *At*, the ”nearest” gene to *C. asymmetrica* is *xpl3*, meanwhile for *Zm* is *xplB*. From the trees, in maize *xplB* is the one that diverged equally than *xplA* and *xplC*. In terms of structure and sequence, *xplA* and *xplB* are more similar and so far beside regulation, theoretically, the *xplA* and *xplB* PEAMTs may have a similar function as result of a whole genome duplication (Lee and Jez 2013; Panchy, Lehti-Shiu, and Shiu 2016). Both genes somehow were homologous, nevertheless, the key to understand both enzymes reside in the regulation. Compared with *xplA* and *xplB*, *xplC* has a PMEAMT function, its uORF is highly ”degraded” and is less expressed in the plant, therefore, *xplC* could’ve been a duplication from any of the *Xpl* genes.

Supposing, the *Zm xplA* and *xplB* PEAMTs could achieve the 3 methylations would result in gene function equivalency then, gene redundancy could be proposed once the mutation rates are probed to be equivalent for both genes (Vision, Brown, and Tanksley 2000; Panchy, Lehti-Shiu, and Shiu 2016). Nevertheless, the sequence relatedness may not be sufficient, like the *At xpl2* can’t perform the methylation from PEA even though the Mt1 is highly similar to *At xpl1,* so as in *Zea mays*, the PEAMTs could’ve different regulation within leaves and roots (Lee and Jez 2017). Related to genes products with the same function, some proteins from the phospholipid synthesis pathways act redundantly like the phosphatidic acid phosphohydrolases: PAH1 and PAH2, which represses the phospholipid biosynthesis in the ER (Eastmond et al. 2010).

By predicting the PEAMTs isoforms in maize, we could come close to the proteins structures where the PEAMT mutations may elucidate the ancestry of the proteins considering maize had duplication events earlier than *At* (Panchy, Lehti-Shiu, and Shiu 2016). So, with the protein model we could have an approach of its structural and (perhaps) the catalytic domains, in comparison with *At* PEAMTs, followed by the plant PEAMT species analysis from algae to angiosperms, this could result in different scenarios to understand the presence of the 2 MT domains, whether the divergence between the evolutionary paths was a result of an MT duplication, a gene duplication or a whole genome duplication event (Vogel, Teichmann, and Pereira-Leal 2005; Nacher, Hayashida, and Akutsu 2010; Sato et al. 2016; Lee and Jez 2017). Finally, how the plant PEAMTs diverged in 2 different types with 1 functional domain, product of 2 merged domains, able to perform the 3 methylation steps, and another where an MT with loss-of-function resulted in a vestigial PEAMT domain.

Perspectives

The motivation for this work was to analyze the *Zea mays Xipotl* family genes and have an approach to their role in the plant.

In order to elucidate the Zm PEAMTs biochemistry, an approach could be done by analyzing the protein sequence and predicting the isoforms structure and catalytic domains in comparison with the At PEAMT. Additionally, enzymatic assays and go further with the complementation experiments At::Zm for each of the *Xpl* genes will give an insight in their methylation activity.

However, their role in maize would be exclusively determinate by further work with the generation of the single and double *Zm Xpl* mutants.The generation of a double *ZmxplA::Ac* and *ZmxplB::Mu* will lead to understand this genes role in the plant, in the methylation pathway and, to elucidate the *ZmxplC* gene function and performance. Along with the identification of uORF *XplB* and *XplA* mutants with the transposons systems, to analyze post-translational regulation. Beside, the generation of loss-of-function *ZmxplA mORF* mutants, loss-of-function *ZmxplB uORF* mutants and the crosses between them, would probably elucidate the uORFs specificity and regulation, and if there's variations in the phenotype due to the probable tissue specific expression.

Additionally, to understand the tissue specific expression, the *Xpl* genes expression might be analyzed in the *Zm* mutants, joined with the analysis of different phospholipids species and cell imaging to analyze membrane damage; in different developmental and vegetative stages, in leafs and root tissues. The generation of post-translational mutants would also support tissue specific expression (at least in an heterozygous system At::Zm) and will approach the PEAMT location in the plant.

Beside the PEAMT biochemistry and regulation analyses, understand the evolutionary path for the MT domains generation and PEAMT divergence, could outcome with the PEAMT sequences phylogenetic analysis including the groups involved in the evolution of plants: algae, bryophytes, seedless vascular plants, gymnosperms, angiosperms, monocots and dycots.

A Murashige and Skoog medium

The Ms media was obtained from Sigma Aldrich, the preparation was followed as described in the protocol with 4.4 g of powder required to prepare 1 L of medium.

Table A.1: Murashige and Skoog Media

Components	**mg/L**
Ammonium nitrate	1,650.0
Boric acid	6.20
Calcium chloride (anhydrous)	332.20
Cobalt chloride hexahydrate	0.0250
Cupric sulfate pentahydrate	0.0250
Disodium EDTA dihydrate	37.260
Ferrous sulfate heptahydrate	27.80
Glycine	2.0
Magnesium sulfate (anhydrous)	180.70
Manganese sulfate monohydrate	16.90
myo-Inositol	100.0
Nicotinic acid	0.50
Potassium iodide	0.830
Potassium nitrate	1,900.0
Potassium phosphate monobasic	170.0
Pyridoxine hydrochloride	0.50
Sodium molybdate dihydrate	0.250
Thiamine hydrochloride	0.10
Zinc sulfate heptahydrate	8.60

B Described protocols employed

Arabidopsis seeds esterilization

First, hydrate them for 30 min in 1.5 ml tubes and maintained in absolute EtOH for 1 min. After that, the seeds were washed with a detergent solution (Tween 0.001% and Sodium hypoclorithe 30%, absolute EtOH) and rinsed 6 times with 1 ml of sterile MQ water, discarding the residues between each rinse.

To break the dormancy process, the seeds were kept at $4\,^{\circ}\mathrm{C}$ for 48 h.

DNA extraction

The collected *Zm* and *At* frozen tissue (50mg) was ground in 2ml tubes with 5mm stainless steel balls with the Retsch tissue lyser, for 30 s at 30 rrpm. Forward, 500 μl of UEB1 buffer (Urea, 5M NaCl, 1M Tris HCL pH 8, 20% sarkosyl, 0.5M EDTA) were added the the tubes containing the ground frozen tissue, and were heated at $68\,^{\circ}\mathrm{C}$ for 10 min on a Thermomixer. The supernatant was transfered to a 1,5 ml Eppendorf tube and 400 μl of equilibrated Phenol:Chloroform:IAA (...) were added, the tubes were shaken a for 30 min at 1000rpm at room temperature. After that, the tubes were centrifuged at 13krpm at room temperature for 10 min. The aqueous upper phase was transfered to a new tube and the were added 40 μl of 3M NaOAc pH 5.2 and 450 μl of 100% isopropanol. The samples were mixed gently and stored for 1 h at -$20\,^{\circ}\mathrm{C}$, then centrifuged at 13krpm at room temperature for 10 min. After discarding the supernatant, 250 μl of 70% EtOH were added and once again centrifuged at 13krpm at room temperature for 5 min. The EtOH was removed and the samples were left air dry until any dopplet could be seen. The DNA was resuspended in 45 μl TE (Tris, EDTA).

RNA extraction

For the RNA extraction, 50 mg of frozen leaf/root tissue were ground with sterilized mortars and collected in 1.5 ml tubes. Then, it was added 1 ml of TRIzol (Thermo Fisher Scientific) and was incubated at room temperature for 10 min. After that, 200 μl of chloroform were added to the tubes and mixed by immersion at room temperature for 3 min. The tubes were taken to centrifuge at 13 krpm for 10 min at $4\,^{\circ}\mathrm{C}$, removing the aqueous (upper) phase and placing it in a new tube. For the RNA precipitation, 300

μl of 100% isopropanol were added to the tubes and were placed for 20 min at room temp. Following, the tubes were taken to centrifuge at 13 krpm for 10 min at $4\,^{\circ}\mathrm{C}$. The supernatant was discarded and for the wash, 1 ml of 70% EtOH was added and the tubes were taken to centrifuge at 13krpm at $4\,^{\circ}\mathrm{C}$ for 5 min. Finally, after removing the EtOH and air dry, the RNA was resuspended in 100 μl of sterile MQ water. For RNA precipitation, the tubes were stored overnight at $4\,^{\circ}\mathrm{C}$. After that, they're centrifuged at 13krpm at $4\,^{\circ}\mathrm{C}$ for 15 min and the RNA was resuspended in 100 μl of sterile MQ water with 10 μl of 3M sodium acetate pH 5.2 and 250 μl of absolute EtOH. The tubes were stored for 2 h at -$20\,^{\circ}\mathrm{C}$ and centrifuged at 13krpm at $4\,^{\circ}\mathrm{C}$ for 5 min. For the washes, 1 ml of 70% EtOH was added and taken to centrifuge at 13 krpm for 5 min. After removing the EtOH, the RNA was resuspended in 40-50 μl of sterile MQ water.

Super Script II Retrotranscriptase (Thermo Fisher) for first strand cDNA synthesis protocol

The following components were added to a nuclease-free PCR tube: Oligo(dT) (500 μg/mL), 5 μg of total RNA, 1 μL dNTP Mix (10 mM each) and sterile distilled water to 12 μL. The mixture was heated to $65\,^{\circ}\mathrm{C}$ for 5 min and quick chilled on ice. After that, 5X First-Strand Buffer and 4 μL, 0.1 M DTT were added to the tubes and mixed gently. The tubes were incubated at $42\,^{\circ}\mathrm{C}$ for 2 min. Then, 1 μL (200 units) of SuperScript II RT was added and mixed gently by pipetting. The tubes were incubated at $42\,^{\circ}\mathrm{C}$ for 50 min. Finally, the reaction was inactivated by heating at $70\,^{\circ}\mathrm{C}$ for 15 min. For better results, a thermal cycler program was used for the heating, incubation and inactivation steps.

Isolation of plasmid DNA

A bacterial pre-culture with selected colonies in LB medium with carbenicillin 100 μg/ml was incubated for 3 hrs at $37\,^{\circ}\mathrm{C}$, then 1.5-2 ml of bacterial cultures were pelleted at 6000-7000 rpm for 1 min. After discarding the supernatant, 150 μl extraction buffer (5% sucrose, 20–50 mM EDTA, 50 mM Tris pH 8, 0.75 M NH4Cl, 0.5% IGEPAL CA-630 (or Triton X-100), lysozyme 100 μg/ml, and RNase A 25 μg/ml) were added and the pellet resuspended. The bacterial suspension was incubated at $65\,^{\circ}\mathrm{C}$ for 5 min and centrifuged at 14 krpm for 10 min, the pellet was removed with a toothpick. To the tubes with the supernatant, 100-120 μl of isopropanol was added, followed by mixing

Table B.1: Golden Gate assembly reaction program

	Temp $^{\circ}\mathrm{C}$	**Time**
	$37\,^{\circ}\mathrm{C}$	3:00
25 cycles	$16\,^{\circ}\mathrm{C}$	4:00
	$50\,^{\circ}\mathrm{C}$	5:00
1 cycle	$80\,^{\circ}\mathrm{C}$	5:00

and centrifugation of the solution at 7000 rpm for 10 min at RT. After discarding the supernatant, the DNA was centrifuged after adding 70% ethanol. Ethanol was removed, and the DNA pellet was dissolved in 20-50 μl TE buffer.

Golden gate assembly protocol

The following components were added to a nuclease-free PCR tube: 100 ng of the linearized vector backbone (pSE7 for BsaI, pMO42 for SapI), 1.5 μl 10X NEB T4 Buffer, 0.15 μl 100X BSA, 1 μl BsaI, 1 μl NEB T4 Ligase (2 million cohesive end units / mL) and sterile distiled H20 to 15 μl. The assembly reaction was performed in a thermocycler as described in Engler 2009:

Agro transformation

Add 1 μg of plasmid DNA constructions to a 100 μl Agrobacterium aliquot Freeze in liquid N2 for 5 min Heat-shock at $37\,^{\circ}\mathrm{C}$ for 25 min Add the Agros to 1 ml of YEB medium in a 15 ml falcon Incubate at $28\,^{\circ}\mathrm{C}$ for 3 h Plumb 100 μl the Agros on petry dishes with LB medium and rifampicin, carbinicillin and spectinomycin at 100 μg/ml Incubate at $28\,^{\circ}\mathrm{C}$ for 48 h

C Arabidopsis promoter and Zea mays sequences

Arabidopsis

xpl 1 (At3G18000) promoter sequence

xpl A (GRMZM2G886060) CDS sequence

xpl B (GRMZM2G170400) CDS sequence

xpl C (GRMZM2G122296) CDS sequence

D *Zm Xpl* gene expression

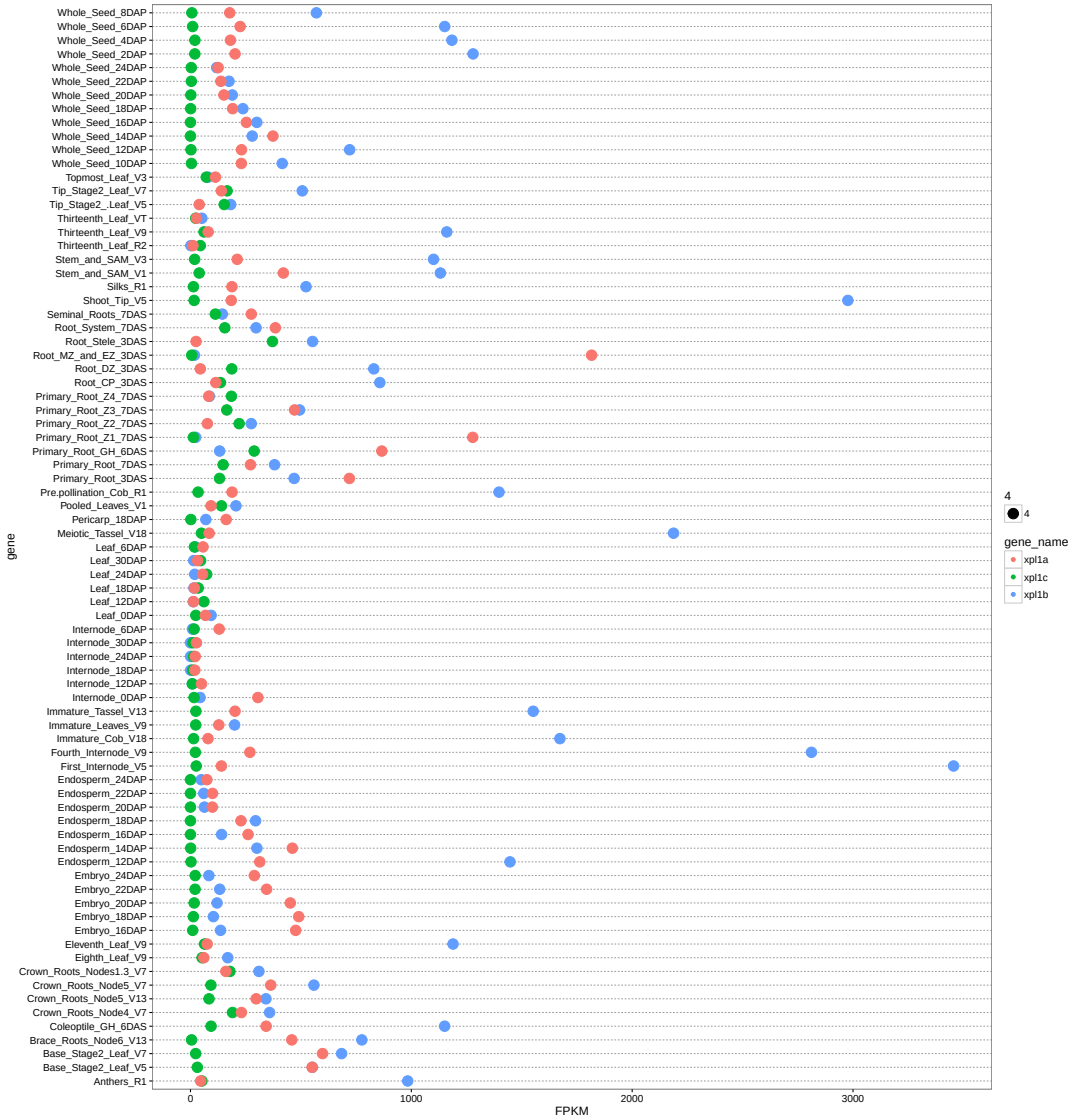


Figure D.1: The RNA-seq gene atlas of maize inbred B73 includes 79 distinct replicated samples. This includes 50 samples from the original array-based gene atlas (Sekhon et al. 2011), a time-course of 12 stalk and leaf samples post-flowering (Sekhon et al. 2012, Plant Phys. 159:1730-1744), and a novel set of 17 samples from the maize seedling and adult root system. The entire dataset contains 4.6 billion mapped reads, with an average of 20.5 million mapped reads per biological replicate, allowing for detection of genes with lower transcript abundance (Stelpflug et al. 2016).

E Genotyping gels for *ZmxplA::Ds* and *ZmxplB::Mu* insertions

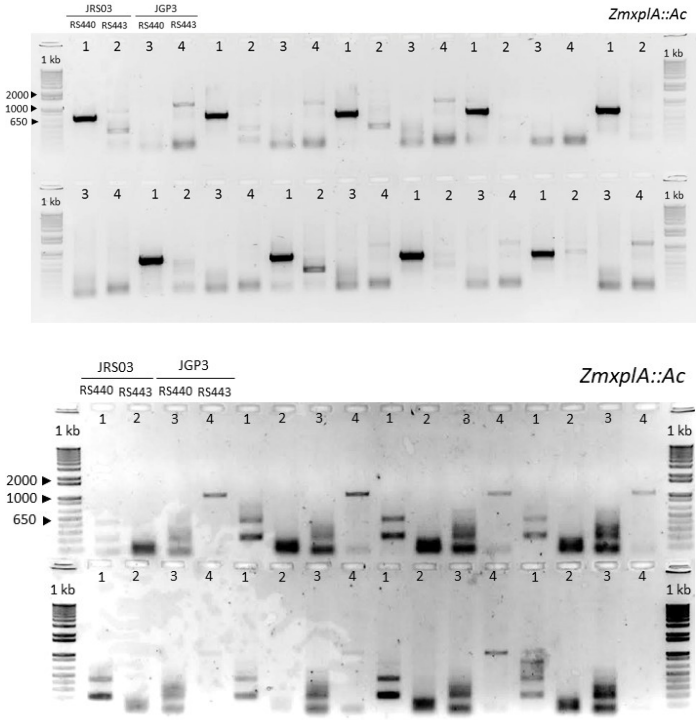


Figure E.1: **Up.** Cornell stocks **Down.** self-pollinated stocks. The electrophoresis gels show the complementary reactions for the *XplA::DS* in the lane **1** (JRS03-RS440, ∼600bp) and in the **4** (JGP3-RS443, ∼1300bp fragment).

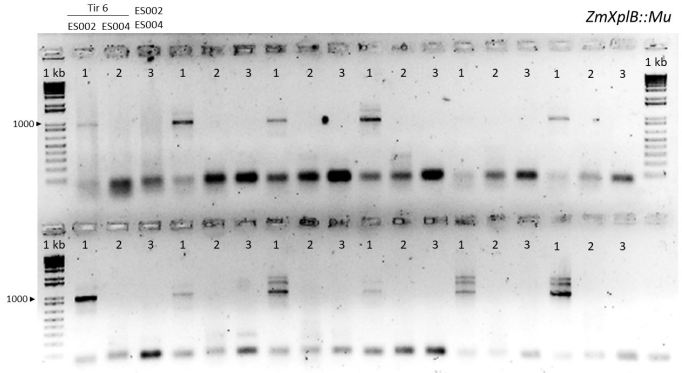


Figure E.2: Complementary reactions with *XplB* primers and Tir6, for the *ZmxplB::Mu* insertion.

Bibliography

Alatorre-Cobos, Fulgencio et al. (2012). “Translational regulation of Arabidopsis XIPOTL1 is modulated by phosphocholine levels via the phylogenetically conserved upstream open reading frame 30”. In: *Journal of Experimental Botany* 63.14, pp. 5203–5221. issn: 1460-2431. doi: 10.1093/jxb/ers180. url: https://academic.oup.com/jxb/article-lookup/doi/10.1093/jxb/err180.

Alleman, M. and J. L. Kermicle (1993). “Somatic variegation and germinal mutability reflect the position of transposable element dissociation within the maize R gene”. In: *Genetics* 135.1, pp. 189–203. issn: 00166731.

BeGora, Michael D. et al. (2010). “Identification of phosphomethylethanolamine N-methyltransferas from Arabidopsis and its role in choline and phospholipid metabolism”. In: *Journal of Biological Chemistry* 285.38, pp. 29147–29155. issn: 1083351X. doi: 10.1074/jbc.M110.112151.

Benning, Christoph (2008). “A role for lipid trafficking in chloroplast biogenesis”. In: *Progress in Lipid Research* 47.5, pp. 381–389. issn: 01637827. doi: 10.1016/j.plipres.2008.04.001.

Bishop, W R and R M Bell (1988). “Assembly of phospholipids into cellular membranes: biosynthesis, transmembrane movement and intracellular translocation.” In: *Annual review of cell biology* 4, pp. 579–610. issn: 10810706. doi: 10.1146/annurev.cellbio.4.1.579.

Bolognese, C P and P McGraw (2000). “The isolation and characterization in yeast of a gene for Arabidopsis S-adenosylmethionine:phospho-ethanolamine N-methyltransferase.” In: *Plant physiology* 124.4, pp. 1800–1813. issn: 0032-0889. url: http://www.plantphysiol.org/content/124/4/1800.short.

Charron, Jean Benoit Frenette et al. (2002). “Molecular and biochemical characterization of a N-methyltransferase from wheat”. In: *Plant Physiology* 129.May, pp. 363–373. issn: 00320889. doi: 10.1104/pp.001776.1.

Conrad, Liza J. and Thomas P. Brutnell (2005). “Ac-Immobilized, a stable source of Activator transposase that mediates sporophytic and gametophytic excision of Dissociation elements in maize”. In: *Genetics* 171.4, pp. 1999–2012. issn: 00166731. doi: 10.1534/genetics.105.046623.

Cooper, Geoffrey M. (2000). *The Cell: A Molecular Approach*. 2nd editio. Sunderland (MA): SinauerAssociates. url: https://www.ncbi.nlm.nih.gov/books/NBK9839/.

Craddock, Christian P et al. (2015). “PHOSPHATIDIC ACID PHOSPHOHYDROLASE Regulates Phosphatidylcholine Biosynthesis in Arabidopsis by Phosphatidic Acid-Mediated Activation of CTP:PHOSPHOCHOLINE CYTIDYLYLTRANSFERASE Activity.” In: *The Plant cell* 27.4, pp. 1251–1264. issn: 1532-298X. doi: 10.1105/tpc.15.00037. url: http://www.plantcell.org/content/early/2015/04/10/tpc.15.00037.

Cruz-Ramírez, Alfredo et al. (2004). “The xipotl mutant of Arabidopsis reveals a critical role for phospholipid metabolism in root system development and epidermal cell integrity”. In: *The Plant Cell …* 16.August, pp. 2020–2034. doi: 10.1105/tpc.103.018648.on. url: http://www.plantcell.org/content/16/8/2020.short.

Datko, A H and S H Mudd (1988). “Enzymes of phosphatidylcholine synthesis in lemna, soybean, and carrot”. In: *Plant physiology* 88.4, pp. 1338–1348. issn: 0032-0889. doi: 10.1104/pp.88.4.1338.

Du, Chunguang et al. (2011). “The complete Ac/Ds transposon family of maize”. In: *BMC Genomics* 12.1, p. 588. issn: 1471-2164. doi: 10.1186/1471-2164-12-588. url: http://www.biomedcentral.com/1471-2164/12/588.

Eastmond, P. J. et al. (2010). “PHOSPHATIDIC ACID PHOSPHOHYDROLASE1 and 2 Regulate Phospholipid Synthesis at the Endoplasmic Reticulum in Arabidopsis”. In: *The Plant Cell* 22.8, pp. 2796–2811. issn: 1040-4651. doi: 10.1105/tpc.109.071423. url: http://www.plantcell.org/cgi/doi/10.1105/tpc.109.071423.

Emami, Shahram, Muh-Ching Yee, and José R Dinneny (2013). “A robust family of Golden Gate Agrobacterium vectors for plant synthetic biology.” In: *Frontiers in plant science* 4.September, p. 339. issn: 1664-462X. doi: 10.3389/fpls.2013.00339. url: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3759027&tool=pmcentrez&rendertype=abstract.

Ganko, Eric W., Blake C. Meyers, and Todd J. Vision (2007). “Divergence in expression between duplicated genes in arabidopsis”. In: *Molecular Biology and Evolution* 24.10, pp. 2298–2309. issn: 07374038. doi: 10.1093/molbev/msm158.

Gibellini, Federica and Terry K Smith (2010). “Invited Critical Review The Kennedy Pathway — De Novo Synthesis of Phosphatidylethanolamine and Phosphatidylcholine”. In: 62.June, pp. 414–428. doi: 10.1002/iub.337.

Griffiths, Anthony J F et al. (1989). *An introduction to genetic analysis*. 10th. Vol. 4. W. H. Freeman and Company, pp. –. isbn: 1429229438.

Hirashima, Takashi et al. (2017). “Characterization of phosphoethanolamine- N -methyltransferases in green algae”. In: *Biochemical and Biophysical Research Communications* 488.1, pp. 141–146. issn: 0006291X. doi: 10.1016/j.bbrc.2017.05.026. url: http://linkinghub.elsevier.com/retrieve/pii/S0006291X17308720.

Hocquellet, Agnès et al. (2005). “Evidence for a different metabolism of PC and PE in shoots and roots”. In: *Plant Physiology and Biochemistry* 43.10-11, pp. 938–946. issn: 09819428. doi: 10.1016/j.plaphy.2005.10.002.

Jost, Ricarda et al. (2009). “Biochemical characterization of two wheat phosphoethanolamine N-methyltransferase isoforms with different sensitivities to inhibition by phosphatidic acid”. In: *Journal of Biological Chemistry* 284.46, pp. 31962–31971. issn: 00219258. doi: 10.1074/jbc.M109.022657.

Lee, Soon Goo and Joseph M. Jez (2013). “Evolution of structure and mechanistic divergence in di-domain methyltransferases from nematode phosphocholine biosynthesis”. In: *Structure* 21.10, pp. 1778–1787. issn: 09692126. doi: 10.1016/j.str.2013.07.023. url: http://dx.doi.org/10.1016/j.str.2013.07.023http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3797223/.

— (2017). “Conformational changes in the di-domain structure of Arabidopsis phosphoethanolamine methyltransferase leads to active site formation”. In: *Journal of Biological Chemistry* 1, jbc.RA117.000106. issn: 0021-9258. doi: 10.1074/jbc.RA117.000106. url: http://www.jbc.org/lookup/doi/10.1074/jbc.RA117.000106.

Lezin, George et al. (2011). “A one-step miniprep for the isolation of plasmid DNA and lambda phage particles”. In: *PLoS ONE* 6.8. issn: 19326203. doi: 10.1371/journal.pone.0023457.

McCarty, Donald and Robert Meeley (2009). *Transposon Resources for Forward and Reverse Genetics in Maize*.

McCarty, Donald R et al. (2013). “Genetic and molecular analyses of UniformMu transposon insertion lines.” In: *Methods in molecular biology (Clifton, N.J.)* 1057, pp. 157–66. issn: 1940-6029. doi: 10.1007/978-1-62703-568-2{_}11. url: http://www.ncbi.nlm.nih.gov/pubmed/23918427.

Mccarty, Donald R et al. (2013). “Plant Transposable Elements”. In: 1057, pp. 157–166. doi: 10.1007/978-1-62703-568-2. url: http://link.springer.com/10.1007/978-1-62703-568-2.

Mcclintock, Barbara (1941). “Chomosome organizaion and genic expression”. In: *Cold Spring Harbor Laboratory Press*, pp. 13–47.

McClintock, Barbara (1950). “The Origin and Behavior of Mutable Loci in Maize”. In: *Genetics* 36.6, pp. 344–355. issn: 0027-8424. doi: 10.1073/pnas.36.6.344.

Mcneil, Scott D et al. (2001). “Enhanced synthesis of choline and glycine betaine in transgenic tobacco plants that overexpress phosphoethanolamine N-methyltransferase”. In: *Proceedings of the National Academy of Sciences* 98.17, pp. 10001 –10005. issn: 00278424. doi: 10.1073/pnas.171228998.

Munnik, Teun and Christa Testerink (2009). “Plant phospholipid signaling: ”in a nutshell”.” In: *Journal of lipid research* 50 Suppl.December 2008, S260–5. issn: 0022-2275. doi: 10.1194/jlr.R800098-JLR200.

Nacher, J. C., M. Hayashida, and T. Akutsu (2010). “The role of internal duplication in the evolution of multi-domain proteins”. In: *BioSystems* 101.2, pp. 127–135. issn: 03032647. doi: 10.1016/j.biosystems.2010.05.005. url: http://dx.doi.org/10.1016/j.biosystems.2010.05.005.

Nakamura, Yuki et al. (2014). “Arabidopsis florigen FT binds to diurnally oscillating phospholipids that accelerate flowering.” In: *Nature communications* 5, p. 3553. issn: 2041-1723. doi: 10.1038/ncomms4553. url: http://www.nature.com/ncomms/2014/140404/ncomms4553/full/ncomms4553.html.

Niu, Gai Li et al. (2018). “Cloning and functional analysis of phosphoethanolamine methyltransferase promoter from maize (Zea mays L.)” In: *International Journal of Molecular Sciences* 19.1. issn: 14220067. doi: 10.3390/ijms19010191.

Nuccio, Michael L. et al. (2000). “cDNA cloning of phosphoethanolamine N-methyltransferase from spinach by complementation in Schizosaccharomyces pombe and characterization of the recombinant enzyme”. In: *Journal of Biological Chemistry* 275.19, pp. 14095–14101. issn: 00219258. doi: 10.1074/jbc.275.19.14095.

Ohlrogge, John and J Browse (1995). “Lipid biosynthesis.” In: *The Plant cell* 7.7, pp. 957–70. issn: 1040-4651. doi: 10.1105/tpc.7.7.957.

Panchy, Nicholas, Melissa D. Lehti-Shiu, and Shin-Han Shiu (2016). “Evolution of gene duplication in plants”. In: *Plant Physiology* 171.August, pp.00523.2016. issn: 0032-0889. doi: 10.1104/pp.16.00523. url: http://www.plantphysiol.org/lookup/doi/10.1104/pp.16.00523.

Ribone, P.A. et al. (2017). “A uORF represses the transcription factor AtHB1 in aerial tissues to avoid a deleterious phenotype”. In: *Plant Physiology* 175.3, pp. 1238–1253. doi: 10.1104/pp.17.01060.

Rodriguez-Gomez, Gustavo (2013). “Reverse genetic analysis of maize genes involved in the response to phosphorous starvation”. PhD thesis. Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional.

Salazar-Vidal, M. Nancy et al. (2016). “Characterization and transposon mutagenesis of the maize (Zea mays) Pho1 Gene Family”. In: *PLoS ONE* 11.9, pp. 1–19. issn: 19326203. doi: 10.1371/journal.pone.0161882.

Sambrook, J, E F Fritsch, and T Maniatis (1989). *Molecular cloning: a laboratory manual.* English. Ed. 2. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, xxxviii + 1546 pp. isbn: 0879693096.

Sato, Naoki et al. (2016). “Diverse pathways of phosphatidylcholine biosynthesis in algae as estimated by labeling studies and genomic sequence analysis”. In: *Plant Journal* 87.3, pp. 281–292. issn: 1365313X. doi: 10.1111/tpj.13199.

Schwertner, H A and J B Biale (1973). “Lipid composition of plant mitochondria and of chloroplasts.” In: *Journal of lipid research* 14.2, pp. 235–42. issn: 0022-2275. url: http://www.ncbi.nlm.nih.gov/pubmed/4698270.

Sekhon, Rajandeep S. et al. (2011). “Genome-wide atlas of transcription during maize development”. In: *Plant Journal* 66.4, pp. 553–563. issn: 09607412. doi: 10.1111/j.1365-313X.2011.04527.x.

Settles, A Mark et al. (2007). “Sequence-indexed mutations in maize using the UniformMu transposon-tagging population.” In: *BMC genomics* 8.1, p. 116. issn: 1471-2164. doi: 10.1186/1471-2164-8-116. url: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-8-116.

Stelpflug, Scott C. et al. (2016). “An Expanded Maize Gene Expression Atlas based on RNA Sequencing and its Use to Explore Root Development”. In: *The Plant Genome* 9.1. issn: 1940-3372. doi: 10.3835/plantgenome2015.04.0025.

Tabuchi, T et al. (2006). “Posttranscriptional regulation by the upstream open reading frame of the phosphoethanolamine N-methyltransferase gene”. In: *Biosci Biotechnol Biochem* 70.9, pp. 2330–2334. issn: 0916-8451. doi: 10.1271/bbb.60309.

Thompson, G A (1980). *The Regulation of Membrane Lipid Metabolism*. CRC Press. isbn: 9780849354274. url: https://books.google.com.mx/books?id=Yp5qAAAAMAAJ.

Tjellström, Henrik et al. (2008). “Membrane phospholipids as a phosphate reserve: The dynamic nature of phospholipid-to-digalactosyl diacylglycerol exchange in higher plants”. In: *Plant, Cell and Environment* 31.10, pp. 1388–1398. issn: 01407791. doi: 10.1111/j.1365-3040.2008.01851.x.

Vision, T J, D G Brown, and S D Tanksley (2000). “The origins of genomic duplication in Arabidopsis”. In: *Science* 290.December, 2114–2117.

Vogel, Christine, Sarah A. Teichmann, and Jose Pereira-Leal (2005). “The relationship between domain duplication and recombination”. In: *Journal of Molecular Biology* 346.1, pp. 355–365. issn: 00222836. doi: 10.1016/j.jmb.2004.11.050.

Vollbrecht, Erik et al. (2010). “Genome-Wide Distribution of Transposed Dissociation Elements in Maize”. In: *The Plant Cell* 22.6, pp. 1667–1685. issn: 1040-4651. doi: 10.1105/tpc.109.073452. url: http://www.plantcell.org/lookup/doi/10.1105/tpc.109.073452.

Wu, Suowei et al. (2007). “Cloning, characterization, and transformation of the phosphoethanolamine N-methyltransferase gene (ZmPEAMT1) in maize (Zea mays L.)” In: *Molecular Biotechnology* 36.2, pp. 102–112. issn: 10736085. doi: 10.1007/s12033-007-0009-1.

Zhang, Jianbo et al. (2009). “Alternative Ac/Ds transposition induces major chromosomal rearrangements in maize”. In: *Genes and Development* 23.6, pp. 755–765. issn: 08909369. doi: 10.1101/gad.1776909.