Improving Accuracy for Genes Involved in Phosphorus Use Efficiency in Corn

Anika Chowdhury, Marzea Akter, Rebecca Seipelt-Thiemann, Biology Department and Honors College

Background

Zea mays (corn) is one of the most important crops grown worldwide (Asim et al. 2017). Because of its importance, it is important to grow corn efficiently.

One of the nutrients that is extremely important for plant growth is phosphorus (P). Usually small amounts of P are available for plant uptake (Battini et al. 2017). P deficiency is a major problem affecting plant growth (Wu et al. 2016). Therefore, it is important to study genes involved in P pathway with the aim of understanding the uptake and use pathways to be able to modify it. Twenty-one genes that play a role in P pathway will be examined in collaboration with Cold Spring Harbor Laboratories to generate accurate gene and transcript models based on actual RNA evidence with the hope of being able to target P-related genes for gene editing on selective breeding and generate more sustainable corn varieties.

Methods

Genome model

- Identify gene models using (using gramene and apollo version 5)
- Constructed a theoretical gene model using a curated version 4 gene model, RNA evidence from Mikado (collapsed RNA sequencing data from six tissues), full-length cDNAs and isoseq RNA evidence
- A region of interest was determined
- Primers were designed to be contained within constitutive regions to determined the splicing events

Test Gene

- RNA was extracted from corn leaf using trizol
- Reverse transcription was performed on RNA using M-MLV RT'ase
- cDNA was tested using Tubulin primer
- The annealing temperature from 44.8° C to 57.0 °C was tested for the best temperature for primer
- PCR on cDNA using phusion Master Mix (three tissues: leaf, root, shoot)

Protein Analysis

- Protein was analyzed using gramene

Results

Table 1: Comparison table for *Zea mays* genes involved in phosphorus uptake. Summary table displays list of 21 genes of phosphorus pathway. The table includes information of genome ID, name, location, number of known transcripts and number of exons for reference sequence.

Maize Genome ID	Gene name
Zm00001d011498	MYB-transcription factor 41
Zm00001d004305	Phosphate transporter protein6
Zm00001d004301	phosphate transporter protein3
Zm00001d004306	phosphate transporter protein8
Zm00001d004308	phosphate transporter protein5
Zm00001d003924	Glycerophosphodiester phosphodiesterase GDPD2
Zm00001d039920	PR5-like receptor kinase
Zm00001d039921	PR5-like receptor kinase
Zm00001d051945	phosphate transporter2
Zm00001d016938	sucrose transporter GRMZM2G081589
Zm00001d017069	Inorganic phosphate transporter 2-1 chloroplastic
Zm00001d040019	MYB-transcription factor 137 (MaizeGDB)
Zm00001d038451	WRKY-transcription factor 57 (MaizeGDB)
Zm00001d040034	Multicopper oxidase LPR2
Zm00001d040035	Multicopper oxidase LPR2
Zm00001d039923	PR5-like receptor kinase
Zm00001d040094	G-type lectin S-receptor-like serine/threonine-protein kinase SD1-8
Zm00001d040035	Multicopper oxidase LPR2
Zm00001d018406	MGD2
Zm00001d038389	Phosphatidate phosphatase PAH1
Zm00001d025343	Purple acid phosphatase 22

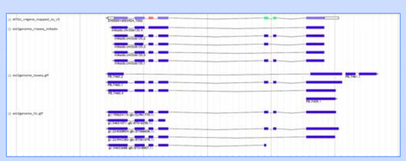

Figure 1: Transcripts Evidence for Gene Zm00001d003924. The transcripts were used to generate a collapsed map. The evidence includes RNA evidence from Mikado (collapsed RNA sequencing data from six tissues), full-length cDNAs and isoseq RNA evidence. The figure was taken from Apollo Instance of Corn Genome, version 5 (Ware, pess.comm).

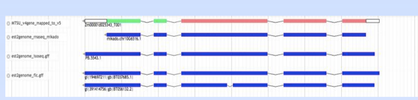

Figure 2: Transcript evidence for Zm00001d025343 used to collapsed the gene map (Apollo version 4)

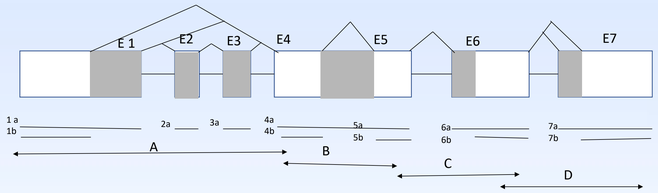

Figure 3: Proposed Gene Structure of Gene Zm00001d003924. The boxes depict exons and the lines depict introns. The exons are numbered from E1 to E7 and the lines below each exons (eg. 1a, 1b) are different transcriptions of a particular exon. The white areas of the boxes are constitutive exons which are exons that are always included in mRNA. the grey areas of the boxes are alternative regions not always included in mRNA. There are four regions of interest (A,B,C,D). the structure was designed using RNA evidence from Apollo Instance of Corn Genome, version 5 (Ware, pess.comm).

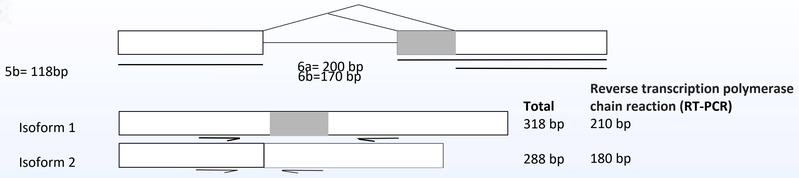

Figure 4: Transcript Comparison for Region of interest. The boxes are exons and the lines are introns. The region of interest includes exon 5 and 6 and the lines below each exon (eg. 6a, 6b) are different versions of a particular exon. The numbers are number of base pairs (bp) in an exon. Splicing patterns above can be used to generate two transcripts (isoform 1 and isoform 2). Primers to detect both transcripts were designed using Primer3 Plus. The half arrows indicate where primers bind in both transcripts.

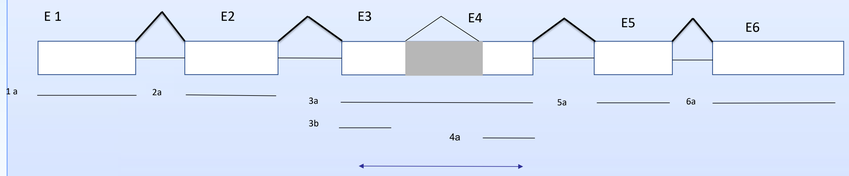

Figure 5: Proposed Gene Structure for gene Zm00001d025343. The boxes indicate exons and the lines indicate introns. The exons are numbered from E1 to E6 and the lines below each exon (eg. 1a) indicate different version of a particular exon. The white boxes are constitutive exons which always included in mRNA. The grey boxes are alternative regions which are not always included in mRNA. The region of interest is noted as a double headed arrow.

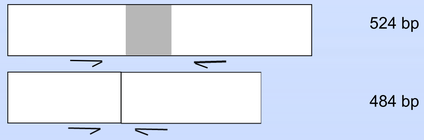

Figure 6: The region of interest includes exon 3 and 4 and the lines below each exon (eg. 4, 3b) are different version of a particular exon. The numbers are number of base pairs in an exon. Splicing patters above can be generated two transcripts (isoform 1 and isoform 2). Primers to detect both were transcripts were designed using Primer3Plus (Rozen et 2000) The half arrows are where set the primers bind in both transcripts.

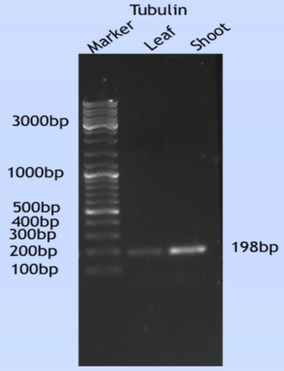

Figure 7 : Reverse Transcription Polymerase Chain Reaction Test of RNA using Tubulin Primers. Agarose gel electrophoresis was used to separate products of RT-PCR using corn RNA samples as noted above each well. Marker sizes (OGene Ruler; Thermo Fisher) are noted at left while the observe size is noted at right. Expected size is 195 bp. Observe size is 198 bp.

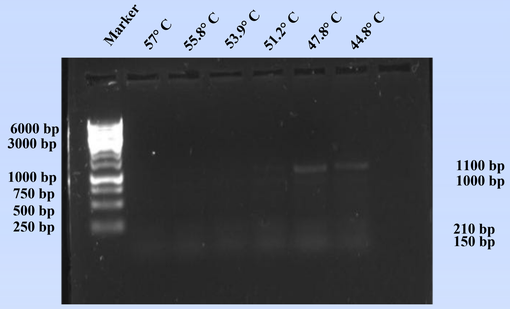

Figure 8: The Gradient Gel to Find the Best Annealing Temperature for RT-PCR Gel Run. Six temperatures (57° C, 55.8C, 53.9° C, 51.2° C, 47.8° C, and 44.8 ° C) were tested to find the best annealing temperature for primer. Marker sizes (OGene Ruler; Thermo Fisher) are noted at left while the observe size is noted at right. Each of the temperature displayed one to four bands at either some or all base pair noted: 150 bp, 210 bp, 1000 bp, and 1100 bp.

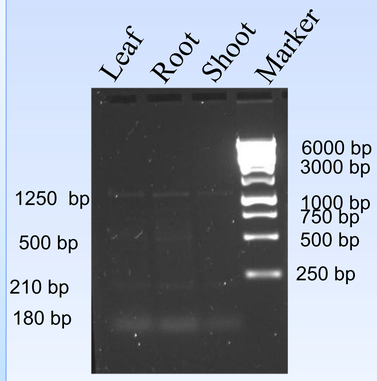

Figure 9: Reverse Transcription Polymerase Chain Reaction of Corn Tissue: Leaf, Root, Shoot. Agarose gel electrophoresis was used to separate products of RT-PCR using corn RNA from root, shoot and leaf. Marker sizes (OGene Ruler; Thermo Fisher) are noted at right while the observe size is noted at left. Observed sizes are 180, 210,500, and 1250 base pairs (bp) for leaf and root and 180, 210, and 1250 base pairs (bp) for shoot.

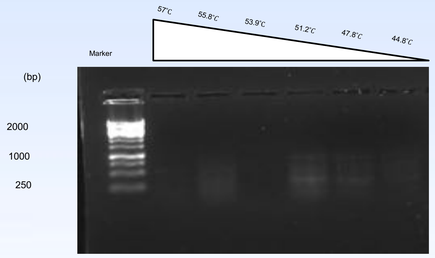

Figure 10: Reverse transcription- polymerase chain reaction (RT-PCR) Results indicate Zm00001d025343 Primers are visible at 250 bp for most of Anneal Temperatures. Best anneal temperature at 51.2°C

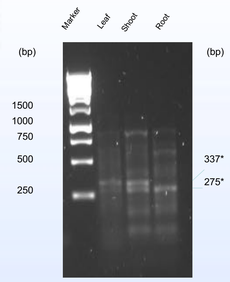

Figure 11: RT-PCR test of RNA using Zm00001d025343 primers at 51.2°C. Agarose gel electrophoresis was used to separate RT-PCR products produced using corn leaf, root, and shoot RNA. Expected isoforms didn't show up, but two new isoforms showed up.

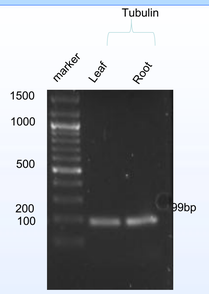

Figure 12: Reverse- transcription PCR Test of RNA using Tubulin Primers. Agarose gel electrophoresis was used to separate products of RT-PCR using corn RNA samples as noted above the wells. Marker rings (OGemeRuler Thermofishar) are noted at left in base pairs (bp) while the observed tubulin fragment size is noted at right. Expected size 195bp.

Table 2: Isoform Comparison Table for Four isoforms Observed During Reverse Transcription Polymerase Chain Reaction of Corn RNA from Leaf, Root, Shoot.

RT-PCR size (bp)	Transcript	Provides evidence for revision to gene model?	Expected and/or observed ?	Domains likely missing	Likely functional?
180	isoform 2	no	E,O	no	yes
210	isoform 1	no	E,O	no	yes
500	isoform 3	yes	O	Unknown	Unknown
1250	isoform 4	yes	O	Unknown	Unknown

Conclusions

- Region of interest is contained within a protein domain.
- Theoretical gene model predicts two expected gene structure of 210 bp and 180 bp.
- Both RNA extraction and reverse transcription was successful.
- The best annealing temperature for the primer was 51.2 C.
- Both of the expected isoform was observed. The expected isoform are not missing any domain and therefore both are functional.
- PCR results shows two larger isoforms that was not expected. The functionality of proteins made from these isoform is unknown. This provides evidence for revision to gene models. One possible explanation to get larger isoforms than expected is that the isoforms might include introns.
- Both expected isoforms were observed. These domains did not lose any domain therefore it is functional.
- The expected isoforms (524 bp & 484 bp) didn't show up, but two new isoforms (337 bp & 275 bp) showed up. The best annealing temperature is at 51.2°C. The gene structure should be revised
- The leaf RNA extraction (199bp) and reverse transcription was successful.
- Predicted and observed RNAs affect (Pfam, PANTHER, Gene3D, CDD)
- The alternative protein that are made won't be functional for Purple acid phosphatase-like, N-terminal

Future Direction

- The PCR tests should be repeated several times for more accurate result and to have more clear bands.
- Possible alternative splicing patterns that were not expected from the theoretical gene model can be used to modify the gene model.
- Any student who takes over the project can work on the isoforms that was not expected and include it in a modified gene model.

References

Asim M, Hussain Q, Ali A, Farooq S, Khan R, Shah S. 2017. Responses of maize to different levels and sources of phosphorus application. Pure and Applied Biology [Internet]. [cited 2019 Sept 20]; 6(3): 1030-1036. Available from: https://www.thepab.org/index.php/journal/article/view/204.

Battini F, Grønlund M, Agnolucci M, Giovannetti M, Jakobsen I. 2017. Facilitation of phosphorus uptake in maize plants by mycorrhizosphere bacteria. Scientific Reports [Internet]. [cited 2019 Sept 2]; Available from: https://www.nature.com/articles/s41598-017-04959-0.

Gupta P, Naithani S, Tello-Ruiz MK, Chougule K, D'Eustachio P, Fabregat A, Jiao Y, Keays M, Lee YK, Kumari S, Mulvaney J, Olson A, Preece J, Stein J, Wei S, Weiser J, Huerta L, Petryszak R, Kersey P, Stein LD, Ware D, Jaiswal P. (2016). *Gramene Database: Navigating Plant Comparative Genomics Resources.* Curr Plant Biol. 2016 Nov;7-8:10-15. PMID: 28713666. doi: 10.1016/j.cpb.2016.12.005.

Untergasser A, Nijveen H, Rao X, Bisseling T, Guerts R, and Leunissen J. [Internet]. 2007. Primer3Plus, an enhanced web interface to Primer3 Nucleic Acids Research. [cited 2019 Nov 19]. Available from: http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi.

Wu F, Liu Z, Xu J, Gao S, Lin H, Liu L, Liu Y, Lu Y. 2016. Molecular Evolution and association of natural variation in ZmARF31 with low phosphorus tolerance in maize. Frontiers in Plant Science [Internet]. [cited 2019 Sept 20]; 7(1076). Available from: ttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC4955381/.