STRUCTURAL ANNOTATION OF CORN INFLORESCENCE GENE *MADS-box 1*

Anna Yuhas, Rebecca L. Seipelt-Thiemann
Biology Department, Middle Tennessee State University

Introduction

Corn, also known as *Zea mays*, is economically and culturally important. Two aspects of corn's biology that can influence its production are the abilities to reproduce and produce seed kernels. Both are affected by the plants ability to flower, inflorescence. In Midwest Maize: How Corn Shaped the U.S. Heartland, it states that the growth of urban markets, manufacturing, and developments in transportation relied on growth in farm countries (Anderson, 2016). There are two types of inflorescence architecture produced by *mays*, the female ear and the male tassel **(**Figure 1; Wei, 2019). These two architectures are controlled by quantitative trait loci and are related to crop yield (Manfei, 2018). Also, tassel size of mays affects crop yield (Gage, 2017). The maize genome has been roughly estimated by looking at related plants and using processes to determine if they possess alike genes. Genome four, which is the focus in this experiment, is estimated and contains rough locations of genes. The problem with genome four is that there is no clear information on the location of the genes and where they are transcribed from. In the experiment, if the locations of inflorescence genes and their sources of transcription can be confirmed within the *Zea mays* genome, it could alter agriculture systems worldwide.

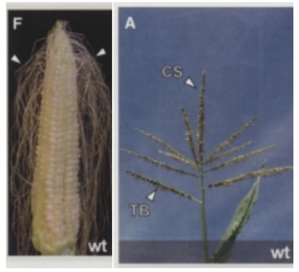

Figure 1. Reproductive Structures of Corn
A. Male flower (tassel) TB indicates a primary branch growing from central spike, CS indicates central spike
B. Female flower (ear) with arrows pointing at fertilized silks (Chatterjee 2014).

Methods

Observation of Apollo versions four and five of the genome *Zea mays* to get a general idea of the gene structure as the focus to be used in the research.

Gene data collected using RNA evidence from Apollo was used to construct gene models which were used to proximate separate isoforms, identify domains, and create primers to focus on in future tests.

The first test was a Reverse-Transcriptase PCR test of RNA using a Tubulin marker, in order to highlight the present isoforms in base pairs (bp) length in the corn leaf sample.

The second test was putting the primers designed through the thermocycler to determine the best temperature for annealing and to make sure that throughout the experiment there is reliable data.

These test results lead to the final experiment of testing the RT-PCR and using the best annealing temperature shown (51.2°C) for inflorescence gene 13—to test the corn leaf, shoot, and root RNA for the expected length (bp).

Results

Feature	Maize Genome ID	Gene name	Chromosome location:	Known transcripts	Number of exons for the REF SEQ
Gene 1	Zm00001d036242	phosphatidylethanolamine-binding protein13	Chromosome 6: 79,177,907-79,179,890 forward strand.	1	4
Gene 2	Zm00001d028905	phytochromeB1	Chromosome 1: 50,443,216-50,455,013 forward strand.	1	4
Gene 3	Zm00001d047632	phytochromeB2	Chromosome 9: 137,491,058-137,497,818 reverse strand.	3	4
Gene 4	Zm00001d017176	C2C2-CO-like-transcription factor 3	Chromosome 5: 188,032,923-188,034,708 forward strand.	1	2
Gene 5	Zm00001d008826	gigantea1	Chromosome 8: 21,787,238-21,795,721 reverse strand.	89	14 (T002)
Gene 6	Zm00001d039589	gigantea2	Chromosome 3: 8,521,898-8,533,430 forward strand.	104	6
Gene 7	Zm00001d032784	G2-like-transcription factor 45	Chromosome 1: 238,041,004-238,043,016 forward strand.	1	5
Gene 8	Zm00001d018667	MADS3	Chromosome 7: 2,191,954-2,203,098 forward strand.	11	???
Gene 9	Zm00001d007949	zea apetala homolog1	Chromosome 2: 243,898,782-243,908,001 forward strand	6	9
Gene 10	Zm00001d007107	Zinc finger protein CONSTANS-LIKE 13	Chromosome 2: 222,393,488-222,401,395 reverse strand.	1	4
Gene 11	Zm00001d027425	agamous-like6	Chromosome 1: 4,979,131-4,994,850 reverse strand	2	7
Gene 12	Zm00001d048474	MADS1	Chromosome 9: 156,960,598-156,980,213 forward strand.	4	7
Gene 13	Zm00001d013259	Zea mays MADS-box 1	Chromosome 5: 7,240,613-7,258,833 forward strand.	7	10
Gene 14	Zm00001d034045	MADS-transcription factor 4	Chromosome 1: 282,136,174-282,165,980 reverse strand	8	2
Gene 15	Zm00001d024909	CO CO-LIKE TIMING OF CAB1 protein domain1	Chromosome 10: 94,430,850-94,433,495 reverse strand.	1	2
Gene 16	Zm00001d011712	ABI3-VP1-transcription factor 19	Chromosome 8: 159,488,658-159,493,031 forward strand	4	7
Gene 17	Zm00001d049485	CCAAT-DR1-transcription factor 5	Chromosome 4: 31,972,821-31,973,600 forward strand	1	1
Gene 18	Zm00001d045735	constans1	Chromosome 9: 36,009,335-36,013,889 forward strand.	4	5

Table 1. Gene List Comparison Table of all genes available in genome four that are estimated to be related to Inflorescence genes in *Zea Mays*. The focus gene for this project is *MADS-box 1* (gene 13).

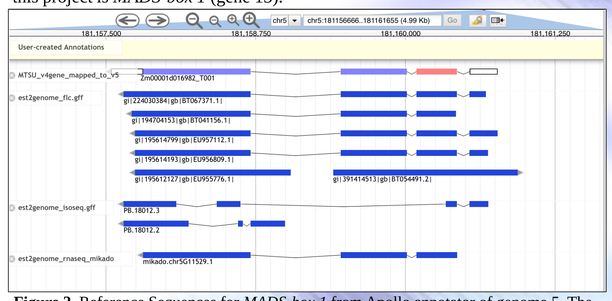

Figure 2. Reference Sequences for *MADS-box 1* from Apollo annotator of genome 5. The lines represent exons and the blue, red, and white blocks represent introns.

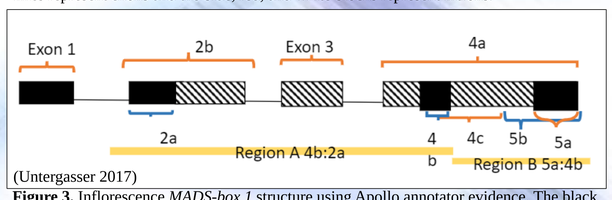

Figure 3. Inflorescence *MADS-box 1* structure using Apollo annotator evidence. The black blocks represent introns among the sequences that are always present. The dashed blocks represent introns among the sequences that are not always present. While the lines connecting the blocks represent exons that are always present among all sequences.

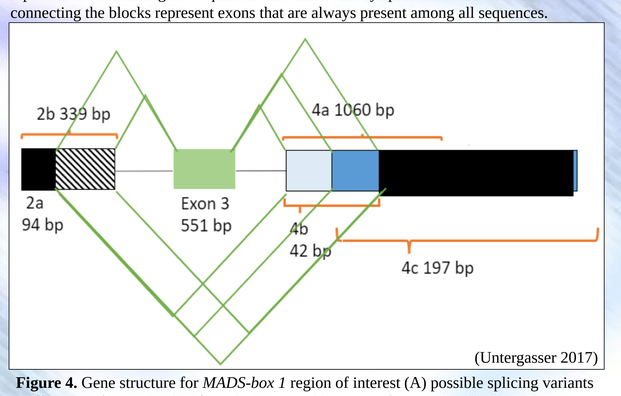

Figure 4. Gene structure for *MADS-box 1* region of interest (A) possible splicing variants (different isoforms possible from the estimated location of primers) with the expected base pair amount for each exon that is represented by a change in block appearance.

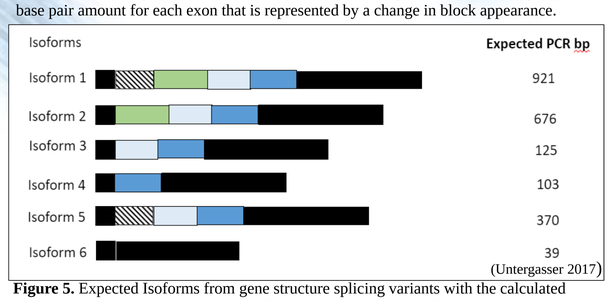

Figure 5. Expected Isoforms from gene structure splicing variants with the calculated base pair numbers that would show on Reverse-Transcriptase PCR results.

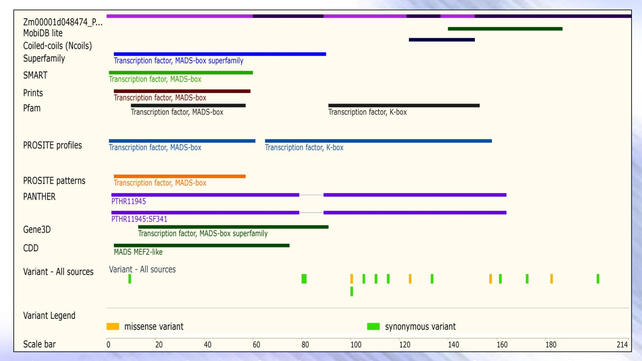

Figure 7. Diagrammatic domain evidence for MADS-box 1. Protein domains are indicated by colored bars and represent the domain location and approximate length. The algorithm in which the domain was identified is noted at the left of the image.

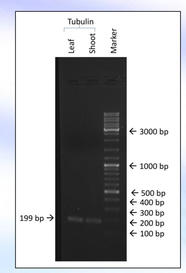

Figure 8. Reverse-Transcription PCR Test of RNA using Tubulin Primers. Agarose gel electrophoresis was used to separate RT-PCR products using corn RNA samples as noted above each sample. Marker OGeneRuler (ThermoFisher) sizes are noted at right while the observed fragment size is noted at left. Expected size 195 bp. This is Cassandra Perrone's sample.

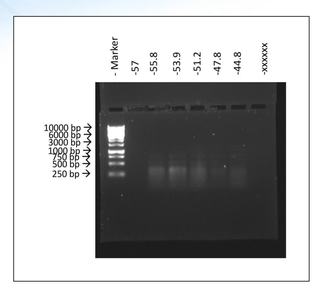

Figure 9. Marker OGeneRuler (ThermoFisher) sizes are noted at left while the observed fragment size is noted at right. Best annealing temperature 51.2 degrees Celsius.

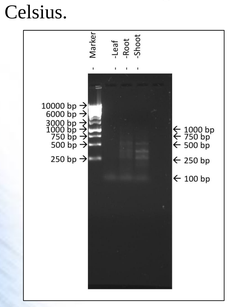

Figure 10. Reverse-Transcription PCR Test of RNA at 51.2 degrees Celsius for annealing temperature. Agarose gel electrophoresis was used to separate RT-PCR products using corn RNA samples as noted above each sample. Marker OGeneRuler (ThermoFisher) sizes are noted at left while the observed fragment size is noted at right. Observed sizes fall between 1000 and 100 bp.

Transcript	RT-PCR size (bp)	Provides evidence for revision to gene model.	Expected or Observed	Domains likely missing?	Likely Functional?
Isoform 1	921	No	E,O	No	Yes
Isoform 2	676	Yes	E	Unknown	Unknown
Isoform 3	125	No	E,O	No	Yes
Isoform 4	103	No	E,O	No	Yes
Isoform 5	370	Yes	E	Unknown	Unknown
Isoform 6	39	Yes	E	Unknown	Unknown
Isoform 7	500	Yes	O	Unknown	Unknown
Isoform 8	250	Yes	O	Unknown	Unknown

Table 2. Comparison Table. This table summarizes the results of each anticipated isoform in comparison to what was expected and how the results effect the future experiments and results for genome five. If the results provide evidence for the revision of the gene model it signifies that the results prove the location of certain exons in the genome and can be used for future testing.

Conclusions & Future Directions

Over the years it has been a topic on many news channels whether genetically modified organisms are harmful to the environment and humans. In Midwest Maize, the author mentions a scientific consensus proving the process of genetically modifying to be more beneficial than harmful (Anderson, 2016). The results of this experiment can be used to further the benefits of genetic engineering. The leaf RNA extraction was successful. The reverse transcription process was also successful, but RT-PCR should be repeated in future experiments to produce clearer results. The best annealing temperature for gene 13 was determined to be 51.2°C. Based on the results from the final RT-PCR, the gene structure should be revised due to unexpected isoforms and unobserved isoforms that were predicted. The region of interest is in a protein domain and the brown and blue bar towards the middle of the diagram (Figure 7) lie in the middle of the domain and their function may be compromised. Their functions relate to the K-box and this will affect folding of alpha-helices in the proteins. It is likely that these bars will be compromised and will affect the genes expression if more proteins are made with the region compromised the functionality of the proteins produced is unknown because you cannot know if the domains are missing or not as shown in the comparison table (Table 2).

Literature Cited

Anderson J.L. 2016. Midwest maize: how corn shaped the U.S. heartland. Canadian Journal of History [Internet]. [Cited 2019 Sep 22];51(1):163-165. Available from: http://web.b.ebscohost.com.ezproxy.mtsu.edu/ehost/pdfviewer/pdfviewer?vid=58&sid=4398c53c-a385-45e8-aee8-0f243b3b0465%40pdc-v-sessmgr05. doi:10.3138/cjh.ach.51.1.rev19.

Chatterjee M, Tabi Z, Galli M, Malcomber S, Buck A, Muszynski M, Gallavotti A. 2014. The boron efflux transporter rotten ear is required for maize inflorescence development and fertility. Plant Cell [Internet].[Cited 2019 Sep 22];26(7):16. Available from http://www.plantcell.org.ezproxy.mtsu.edu/content/26/7/2962.full. doi:https://doi-org.ezproxy.mtsu.edu/10.1105/tpc.114.125963.

Gage JL, Miller ND, Spalding EP, Kaepler SM, De Leon N. 2017. Tip: a system for automated image-based phenotyping of maize tassels. Plant Methods [Internet]. [Cited 2019 Aug 27]; 13(21):1-12. Available from: https://plantmethods.biomedcentral.com/track/pdf/10.1186/s13007-017-0172-8. doi: 10.1186/s13007-017-0172-8.

Li M, Zhong W, Yang F, Zhang Z. 2018. Genetic and molecular mechanisms of quantitative trait loci controlling maize inflorescence architecture- a review. Plant and Cell Physiol. [Internet]. [Cited 2019 Aug 27];59(3):448-457. Available from:https://academic.oup.com/pcp/article-pdf/59/3/448/24246675/pcy022.pdf. doi:10.1093/PCP/pcy022.

Sun W, Chen D, Xue Y, Zhai L, Zhang D, Cao Z, Liu L, Cheng C, Zhang Y, Zhang Z. 2019. Genome-wide identification of AGO18b-bound miRNAs and phasiRNAs in maize by cRIP-seq. GMC Genomics [Internet].[Cited 2019 Aug 27];20(656):1-11. Available from: https://rdcu.be/bPH3V.Doi:https://doi.org/10.1186/s12864-019-6028-z.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen J. 2007. Primer3Plus, an enhanced wed interface to Primer3. Nucleic Acids Research [Internet]. [Cited 2019 Nov 14]; 35(1): W71-W74. Available from: http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plusAbout.cgi. DOI: 10.1093/nar/gkm306.