

The SET-Domain Containing Protein and MYB-related Families: Genes AT2G05900 & AT1G17460



Kristin Gill
HC70AL Spring 2008



What is AT2G05900?

- The gene is 939 base pairs in length

5' **EXON** 3'

- AT1G05900 encodes a **SET-Domain Containing Protein** of 312 amino acids

What is a SET-Domain Containing Protein?

- Includes all but one of the proteins that methylate histones on lysine
- Function is to **transfer a methyl group to lysine on the histone**
- Methylation of histone lysine functions as a **post-translational modification**
- Serves as a marker for the recruitment of specific complexes that direct the organization of chromatin
- Important in the regulation of chromatin and gene expression

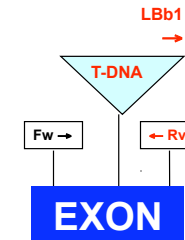
What role do SET-Domain Containing Proteins play in *Arabidopsis thaliana*?

- **Control gene expression during development**
- Important for developmental processes, such as:
 - Embryogenesis
 - Regulation of pollen tube growth
 - Flowering time control

What role does SET-Domain Containing Protein AT2G05900 Gene play in *Arabidopsis* seed development?

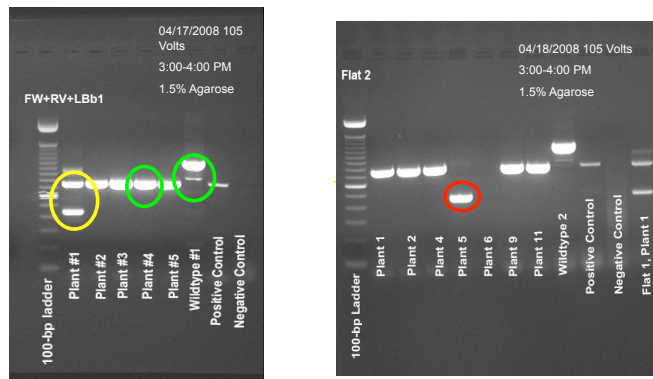
- Study AT2G05900 knockout line

Where is the T-DNA insert in AT2G05900?



According to SALK, the T-DNA inserts at nucleotide **622**
Sequencing results confirmed SALK's predictions!

What were the genotyping results?



Hemizygous
(Plant 1)

Wildtype
(Plants 1, 2, 4, 9, 11)
Expected Size: 750bp

Homozygous
(Plant 5)
Expected Size (T-DNA):
455 bp

What do my results mean?

Genotype	# of Plants
Wildtype/Wildtype	11
Hemizygous	1
Mutant/Mutant	1

Since I found a homozygous mutant, **a knockout of AT2G05900 does NOT cause seed lethality!!!**

Where is AT2G05900 active?

Experiment	Data
ATH1-121501/WT Ws-0/Chalazal Endosperm/Globular Stage	A
ATH1-121501/WT Ws-0/Chalazal Seed Coat/Globular Stage	A
ATH1-121501/WT Ws-0/Embryo Proper/Globular Stage	A
ATH1-121501/WT Ws-0/Floral Bud/Reproductive	A
ATH1-121501/WT Ws-0/General Seed Coat/Globular Stage	A
ATH1-121501/WT Ws-0/Leaf/Vegetative	A
ATH1-121501/WT Ws-0/Micropylar Endosperm/Globular Stage	A
ATH1-121501/WT Ws-0/Ovule/Reproductive	A
ATH1-121501/WT Ws-0/Peripheral Endosperm/Globular Stage	A
ATH1-121501/WT Ws-0/Root/Vegetative	A
ATH1-121501/WT Ws-0/Seed/24-Hr Post-Fertilization	A
ATH1-121501/WT Ws-0/Seed/Cotyledon Stage (7-8DAP)	A
ATH1-121501/WT Ws-0/Seed/Globular Stage (3-4DAP)	A
ATH1-121501/WT Ws-0/Seed/Mature Green Stage (13-14DAP)	A
ATH1-121501/WT Ws-0/Seed/Post-Mature Green Stage (18-19DAP)	A
ATH1-121501/WT Ws-0/Seedling/3 Days After Imbibition (3DAI)	A
ATH1-121501/WT Ws-0/Stem/Vegetative	A
ATH1-121501/WT Ws-0/Suspensor/Globular Stage	A

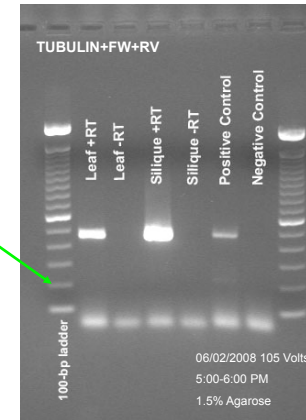
Gene Chip Data

A = Absent

AT2G05900 is absent in these stages of development

Where is AT2G05900 active?

Expected RT-PCR
Product Size: 196 bp
Observed RT-PCR
Product Size: **NONE!**



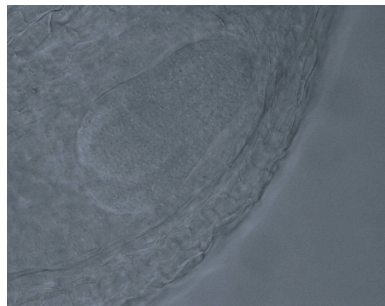
AT2G05900 is absent in the leaf and silique stages of development and **confirms Gene Chip data**

Did the mutant plants show any phenotypic differences?

Knockout of AT2G05900



Wildtype Plant



NO! At the torpedo stage, the **mutant plant appears normal**

What does gene AT1G17460 code for?

A MYB-related Family Transcription Factor

What is a MYB Family Transcription Factor?

- MYeloBlast ---> MYB
- First MYB gene identified was the "oncogene" v-MYB derived from the avian myeloblastosis virus
- 203 MYB Family Transcription Factors identified
- Domain generally contains up to three imperfect repeats, R1, R2 and R3

The MYB Family Transcription Factor in *Arabidopsis thaliana*

- 150 loci code for MYB Family Transcription factors
- In *Arabidopsis*, almost all of the MYB proteins belong to the MYB-R2R3 class (131 members)
- Contain two imperfect repeats of the MYB motif

What does gene AT1G17460 code for?

A MYB-related Family Transcription Factor

What is a MYB-related Family Transcription Factor?

- In addition to the MYB-R2R3 proteins, *Arabidopsis* contains additional transcription factors characterized by a more divergent MYB domain
- This domain is present usually as a **single repeat** and are called "MYB-related"
- In the *Arabidopsis* genome there are 49 loci that code for MYB-related Family Transcription factors

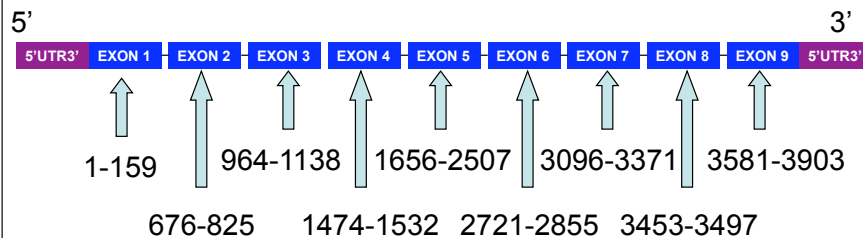
What are the functions of plant MYB proteins?

- Wide diversity of functions
- By no means fully characterized
- In general: regulate many aspects of plant development and metabolism
- Specifically,
 - Aid in pathogen defense
 - Induce response to abiotic stress
 - Involved in tryptophan biosynthesis
 - Important in cellular morphogenesis

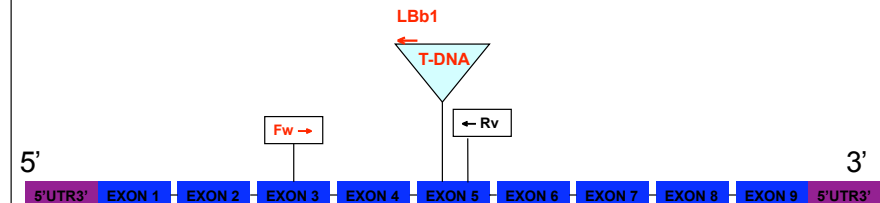
What is the size and orientation of AT1G17460?

- The gene is 3,903 base pairs in length
- The size of its protein is 604 amino acids
- AT2G05900 is oriented in the 5' → 3' direction

What is the structure of AT1G17460?

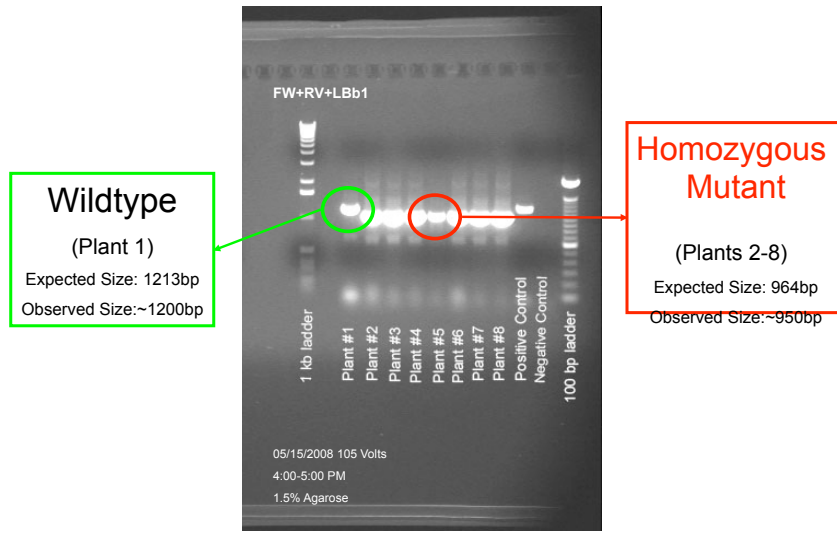


Where is the T-DNA insert in AT1G17460?



According to SALK, the T-DNA inserts at nucleotide 1806
Sequencing results confirmed SALK's predictions!

What were the genotyping results?

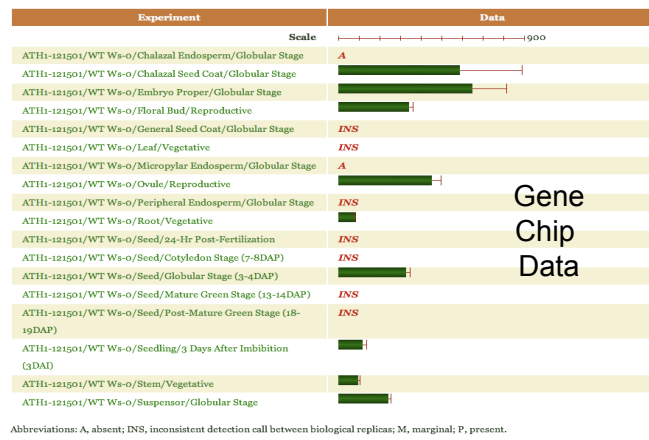


What do my results mean?

Genotype	# of Plants
Wildtype/Wildtype	1
Hemizygous	0
Mutant/Mutant	7

Since I found homozygous mutants, a **knockout of AT1G17460 does NOT cause seed lethality!**

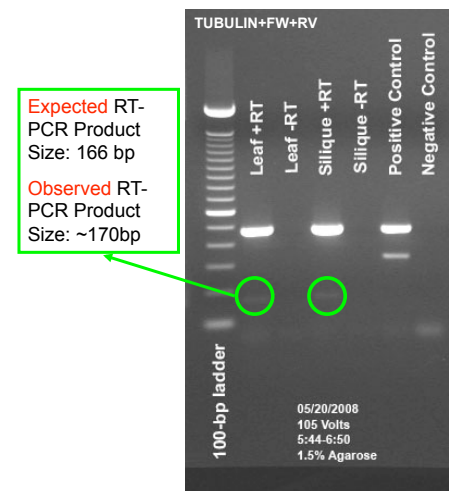
Where is AT1G17460 active?



Abbreviations: A, absent; INS, inconsistent detection call between biological replicates; M, marginal; P, present.

AT1G1460 is expressed in at least one stage of seed development and data is inconsistent as to whether it is expressed in the leaf

Where is AT1G17460 active?

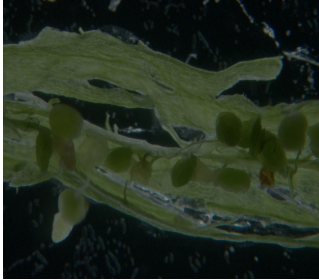


My RT-PCR data supports the gene chip data in that **the gene is active in the seed**

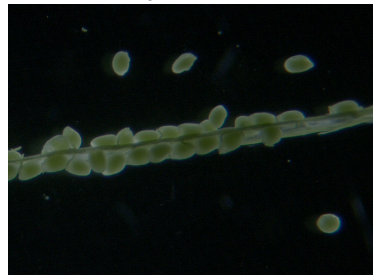
It also shows that **the gene is active in the leaf**

Did the mutant plants show any phenotypic differences?

Mutant Seeds
(Knockout of AT1G17460)



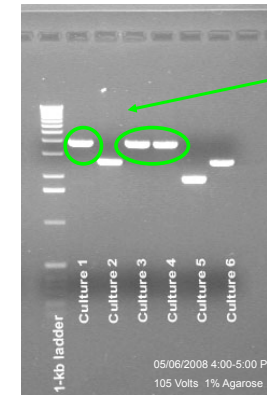
Wildtype Seeds



Mutant plant shows white seed phenotype!

How did I study gene activity in AT1G17460?

- Gene activity is controlled by the upstream region
- Amplified upstream region using iProof Polymerase
- Ligated into the TOPO-vector
- E. coli cells transformed with recombinant plasmids
- Restriction Digest with Asc I
- Isolate and sequence recombinant plasmids



Recombinant Plasmids

- Expected Size of Upstream PCR Product: 1.2 kb
- Expected Size of Recombinant Plasmid: 3.8 kb
- Observed Size of Recombinant Plasmid: ~3.8 kb

Cloned upstream region of Plasmid 1 corresponds to the upstream region of AT1G17460!

What is the significance of my results?

- A knockout of **AT2G05900** does not cause seed lethality
 - There may be redundancy in gene function
- A knockout of **AT1G17460** does not cause seed lethality, but some seeds had a different phenotype
 - The knockout may cause the different seed phenotype
 - OR
 - It is the result of environmental conditions

What further research should be done?

AT2G05900

- Examine more seeds and siliques for phenotypic differences
- Clone the upstream region and use GFP's to determine where the gene is expressed
- Double knockout

AT1G17460

- Use GFP's and cloned upstream region to determine where the gene is expressed
- Examine more seeds and siliques for phenotypic differences
- Double knockout

Acknowledgements

I would like to thank Anhthu, Daisy, Bekah, Tomo, Brandon, the entire HC70AL class, and Dr. Goldberg for all of their help this quarter!

It wouldn't have been the same without you!