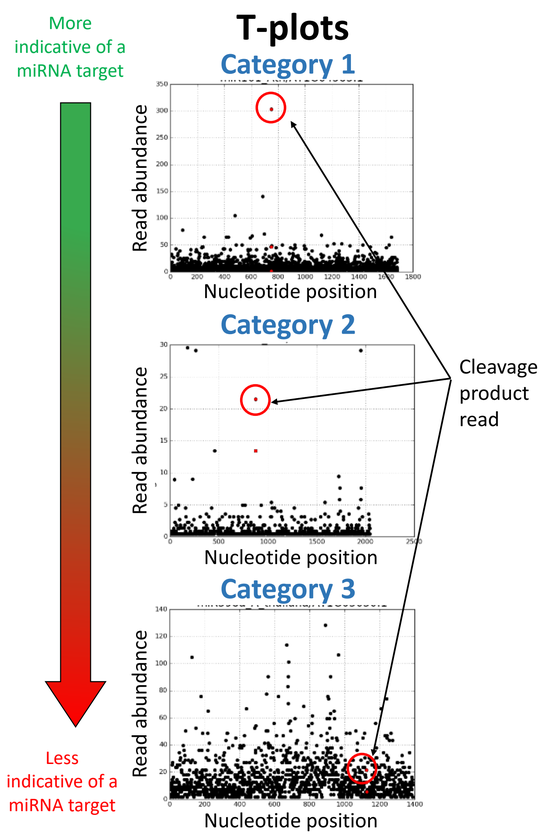
# Defining a microRNA targetome in plants: How many genes are regulated by miRNA in *Arabidopsis thaliana*?

Gigi Wong, Tony Millar
Plant Sciences, Research School of Biology, Australian National University, Canberra, ACT, 2601

Australian National University

## 1. Introduction

- MicroRNA are short non-coding RNA which are essential for plant development and function by down regulating target gene expression.
- In plants, miRNA bind to mRNA sequences of high complementarity leading to cleavage at specific sites of the transcript.
- Therefore, miRNA targets can be experimentally identified *in vivo* using **degradome sequencing** which captures all uncapped transcripts, including miRNA target cleavage products[1].




- Degradome reads are mapped to gene transcripts to form T-plots[1]
- Reads corresponding to the cleavage product are categorized based on the relative abundance to reads at all other positions on the transcript:
  - Category 1: cleavage read is most abundant on transcript
  - Category 2: cleavage read is not the most abundant, but greater than the average abundance of reads per nt position
  - Category 3: cleavage read abundance is less than the average
- Degradome data from many publicly available experiments are accessible via degradome database, WPMIAS[2].
- Many miRNA targets are predicted for *A. thaliana* by multiple miRNA target prediction programs, but few have been validated experimentally suggesting many false positives

**AIM: Are we able to develop a bioinformatic pipe to define the scope of miRNA-target interactions in *A. thaliana*?**

## 2. Bioinformatic Pipeline




MiRNA targets are identified as **Major** and **Minor** targets based on the **Strength** and **Frequency** of their T-plots across multiple degradome experiments

*Transcripts per 10 Million

## 4. Conclusion

- An improved understanding on the scope of the *A. thaliana* targetome – there are 100s of targets and not 1000s as predicted
- Regulation of the *A. thaliana* miRNA targetome is correlated with conservation – Highly conserved miRNA disproportionally contribute to the targetome
- Development of a degradome based bioinformatic pipeline that identifies miRNA target based on strength and frequency of T-plots.

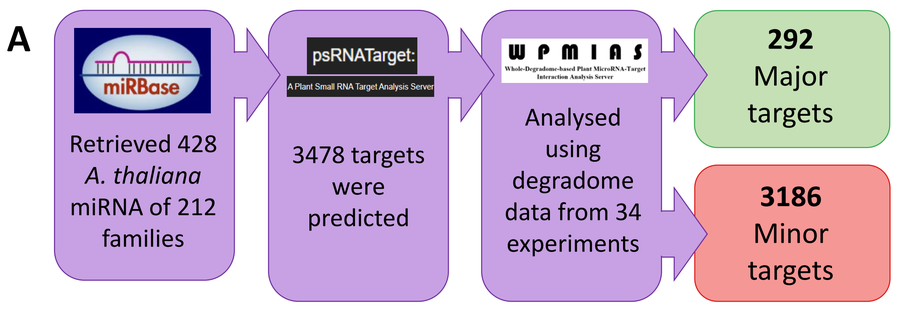
## 3. Results

### Benchmarking

Our pipeline was benchmarked against a set of 106 previously experimentally validated miRNA-target interactions in the model plant species, *Arabidopsis thaliana* (*A. thaliana),* and was able to identify 97 (92%) as Major targets.

### Application to the *A. thaliana* targetome

Our pipeline was applied to study the collection of all genes regulated by miRNA in *A. thaliana* (*A. thaliana* targetome).

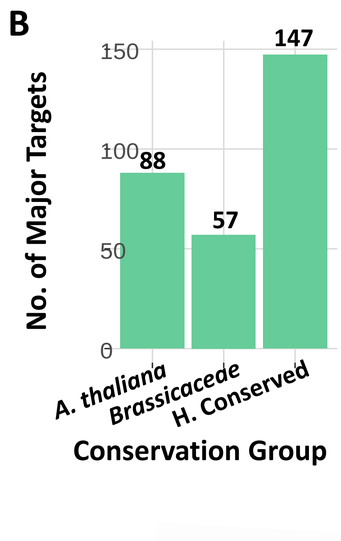



- Much fewer Major targets were identified compared to Minor targets for all *A. thaliana* miRNAs

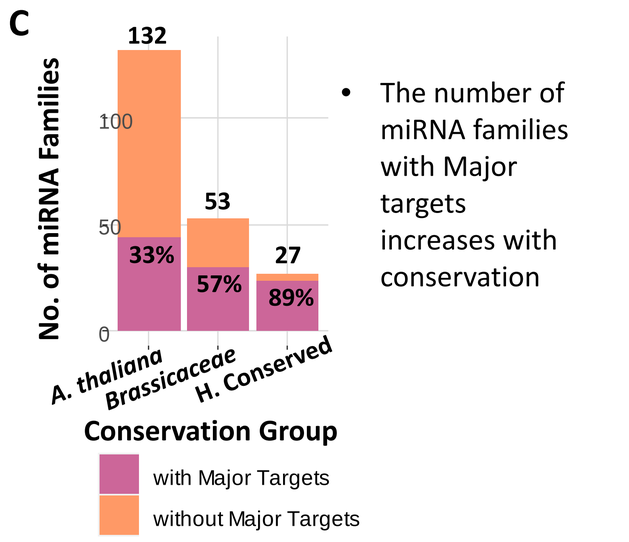
*A. thaliana* miRNA families were further grouped by conservation:

- Highly conserved** (27 families)
- *Brassicaceae* specific (53 families)
- *A. thaliana* specific (132 families)

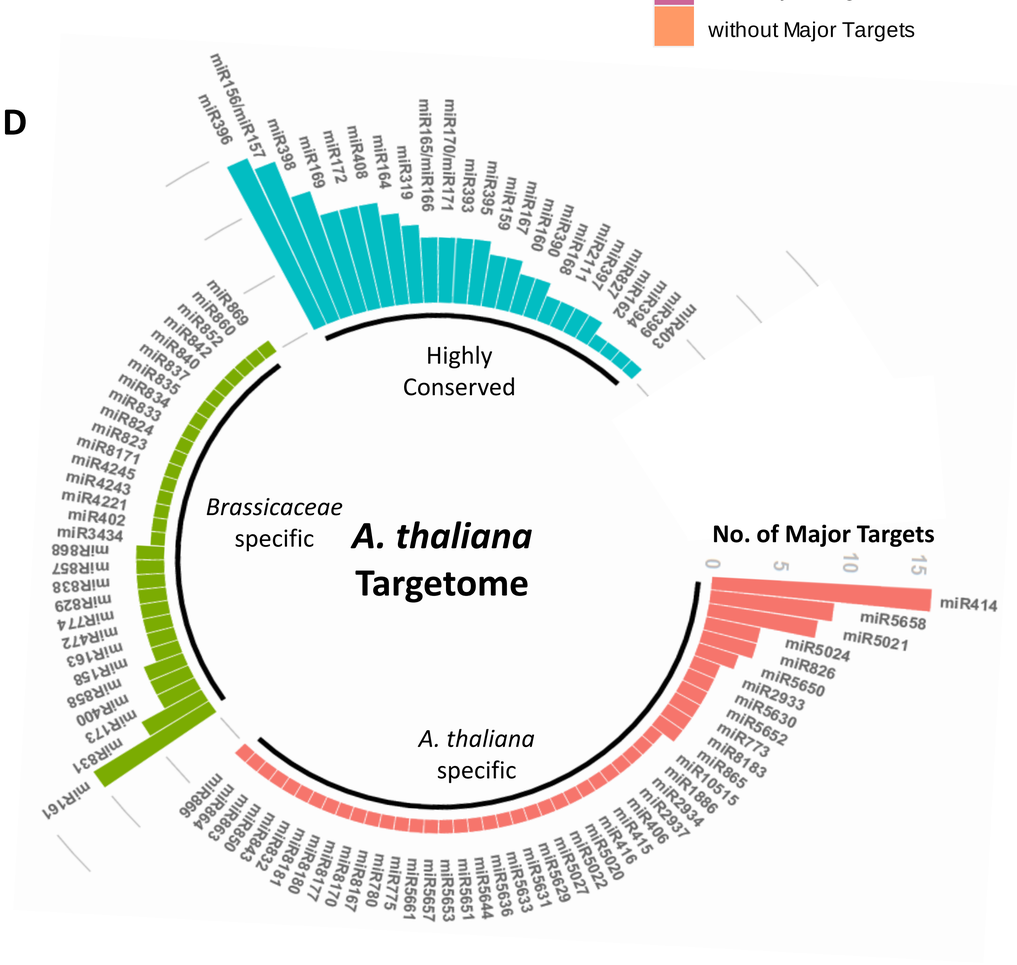
conservation ↓




- Highly conserved miRNA families have the most Major targets despite having the smallest miRNA family group




- The number of miRNA families with Major targets increases with conservation




- Many highly conserved miRNA families were identified with many Major targets per family.
- *Few A. thaliana* and *Brassicaceae* specific miRNA families had many Major targets per miRNA family; many had few targets.
- The more conserved a miRNA is, the more likely it is to have a miRNA target.

**Composition of the *A. thaliana* targetome using our bioinformatic pipeline.** A) The overall number of Major, Minor and predicted targets from all *A. thaliana* miRNAs retrieved from miRBase[3]. B) The distribution of Major targets across conservation groups. C) The proportion of miRNA families with Major targets across conservation groups. D) The number of Major targets per miRNA family across conservation groups.

**Conserved across multiple clades of land plants

## References

1) Addo-Quaye, C., Eshoo, T. W., Bartel, D. P., & Axtell, M. J. (2008). Endogenous siRNA and miRNA Targets Identified by Sequencing of the Arabidopsis Degradome. *Current Biology*, *18*(10), 758–762. https://doi.org/10.1016/j.cub.2008.04.042
2) Fei, Y., Mao, Y., Shen, C., Wang, R., Zhang, H., & Huang, J. (2020). WPMIAS: Whole-degradome-based plant MicroRNA–Target interaction analysis server. *Bioinformatics*, *36*(6), 1937–1939. https://doi.org/10.1093/bioinformatics/btz820
3) Kozomara A, Birgaoanu M, Griffiths-Jones S. (2019). MiRBase: from microRNA sequences to function. Nucleic Acids Research. 47:D155-D162; doi.org/10.1093/nar/gky1141
4) Dai, X., Zhuang, Z., & Zhao, P. X. (2018). PsRNATarget: A plant small RNA target analysis server (2017 release). *Nucleic Acids Research*, *46*(W1), W49–W54. https://doi.org/10.1093/nar/gky316