



# New Compositae specific probe set reduces paralogs in complex systems

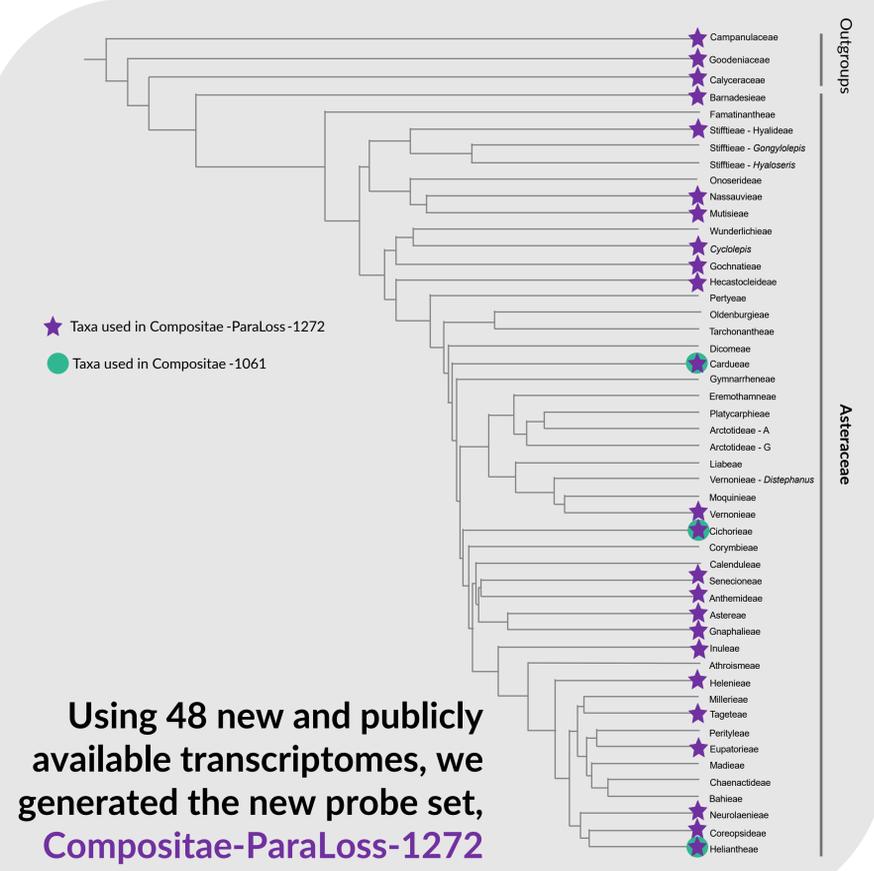
Erika R. Moore-Pollard<sup>1</sup>, Daniel S. Jones<sup>2</sup>, Jennifer R. Mandel<sup>1</sup>  
<sup>1</sup>University of Memphis, Memphis, TN, USA; <sup>2</sup>Auburn University, Auburn, AL, USA

The Asteraceae specific probe set, **Compositae-1061**<sup>1</sup>, has become popular among researchers studying members of Asteraceae, and has paved the way for investigations at lower taxonomic levels and non-model organisms. Though **Compositae-1061** has shown to be highly efficient at higher- and some lower-taxonomic levels within the family, it generally lacks resolution at the genus to species level, especially in groups with complex evolutionary histories including polyploidy and hybridization.

In this study, we developed a new probe set, **Compositae-ParaLoss-1272**<sup>2</sup>, designed to target known single-copy orthologous loci across Asteraceae family members using the tool MarkerMiner<sup>3</sup>. We tested its effectiveness at lower taxonomic levels using 19 taxa within the genus *Packera*, which has a complex evolutionary and taxonomic history, and was previously studied using **Compositae-1061**<sup>4</sup>. Sequences from both probe sets, along with a double-capture approach, were used to generate phylogenies, compare topologies, and assess node support.

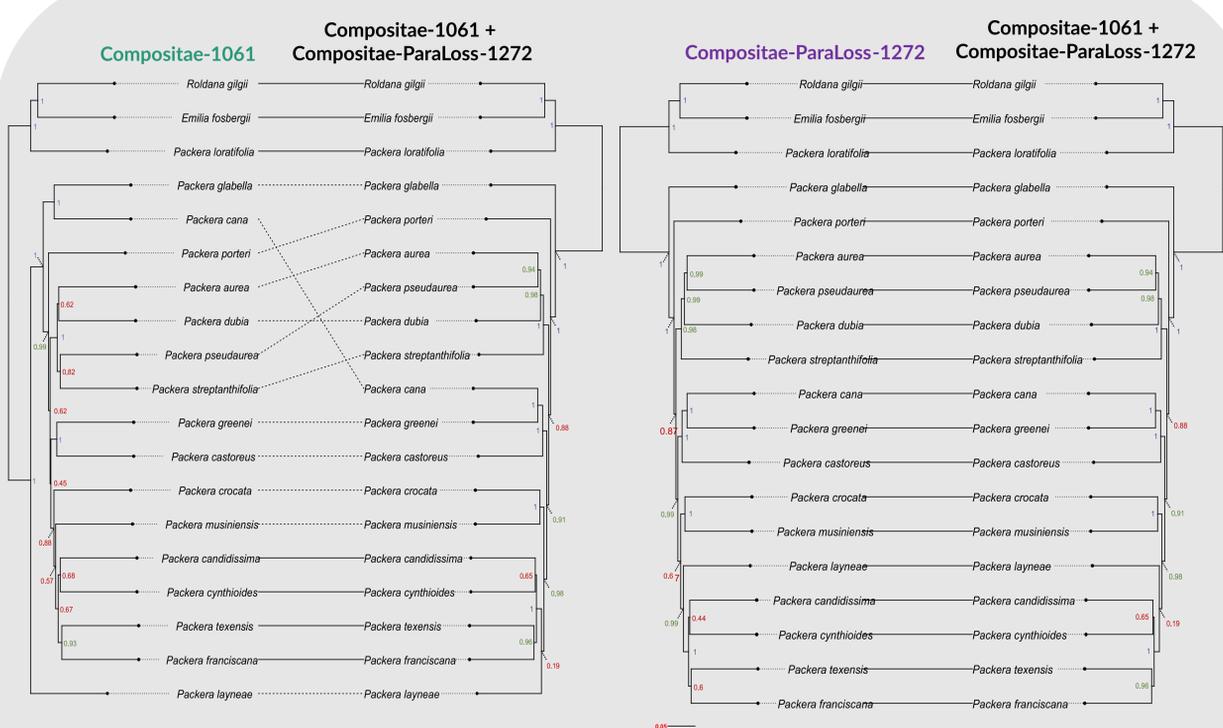
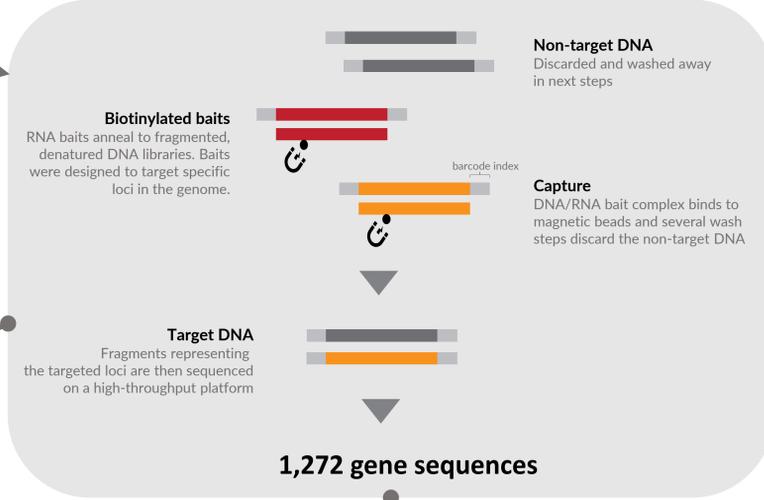
## Differences between **Compositae-ParaLoss-1272**, **Compositae-1061**, and the angiosperm-wide probe set, **Angiosperm-353**<sup>5</sup>

	<b>Compositae-ParaLoss-1272</b>	<b>Compositae-1061</b>	<b>Angiosperm-353</b>
# loci	1,272	1,061	353
# baits	60,158	9,678	75,151
# loci overlap	0	30 (with Angiosperm-353)	30 (with Compositae-1061)
# species as input	48	3	42
Input data	Transcriptomes	Expressed sequence tags (EST)	Transcriptomes
Tool	MarkerMiner	BLAST	k-medoid clustering

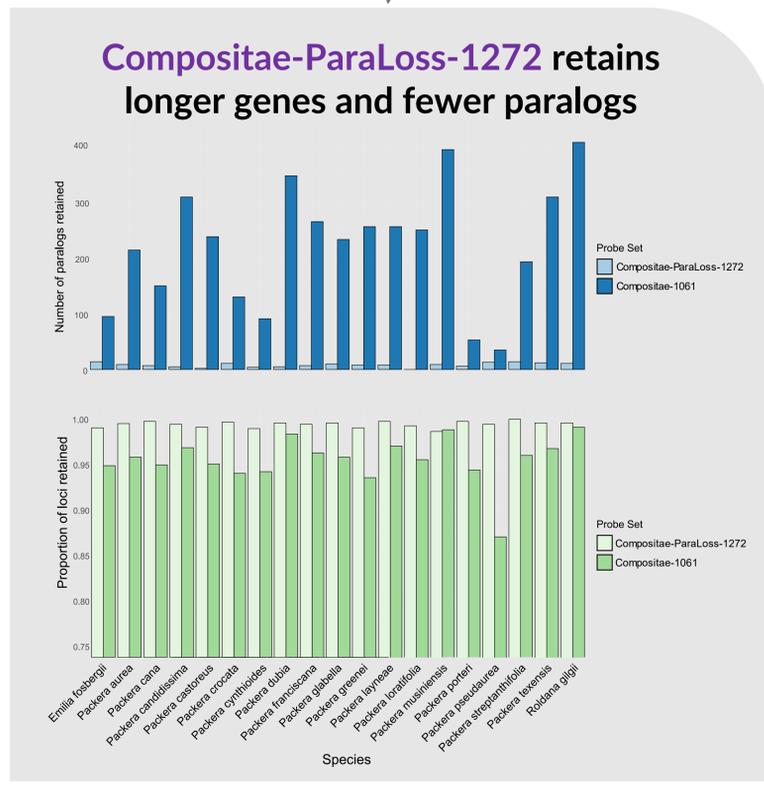


Using 48 new and publicly available transcriptomes, we generated the new probe set, **Compositae-ParaLoss-1272**

**Compositae-ParaLoss-1272** provides improved resolution and utility for studies of complex groups and lower-taxonomic levels in the sunflower family.



Node support is higher when probe sets are combined & relationships remain the same to **Compositae-ParaLoss-1272**



## **Compositae-ParaLoss-1272** retains longer genes and fewer paralogs

Questions? Contact me!  
[moore.erika.r@gmail.com](mailto:moore.erika.r@gmail.com)  
erika-moore.com  
@erikarmore11



CITATIONS:  
<sup>1</sup>Mandel et al. (2014), *APPS*, 2(2), 1300085; <sup>2</sup>Moore-Pollard et al. (2023), *APPS*, in press; <sup>3</sup>Chamala et al. (2015), *APPS*, 3(4), 1400115; <sup>4</sup>Moore-Pollard & Mandel (2023), *bioRxiv*; <sup>5</sup>Johnson et al. (2016), *APPS*, 4(7), 1600016