Deep Sequencing of Small RNAs in the F1 Hybrid B73 x Mo17 and its Parents

Wesley T. Barber*, Wei Zhang, Kranthi K. Varala, Hlaing Win, Stephen P. Moose, Matthew E. Hudson

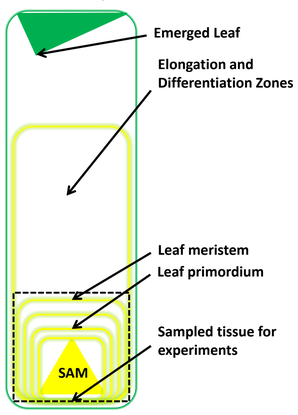
Department of Crop Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, 61801, *contact: barber4@illinois.edu

Rationale:

- Allelic diversity and changes in dosage for key regulatory genes are hypothesized to contribute to heterosis in plants.
- The highly repetitive maize genome generates large amounts of allelic diversity and is believed to be responsible for the crop's high manifestation of heterosis.
- Small RNAs (sRNAs) are a new class of regulatory factors that control plant growth and development, function in plant immunity, and maintain the structure of the genome.
- In *Arabidopsis* and maize, the activities of miRNA156 and miRNA172 control shoot development and maturation. Recently, miRNA396 has been shown to regulate overall leaf size in *Arabidopsis* through its control of cellular proliferation.
- Changes in the duration or level of expression of key miRNA genes may play a role in the heterosis observed early on in maize hybrid seedlings compared to their parents.
- In *Arabidopsis,* repeat derived siRNAs have been shown to not only silence repeat elements from which they are derived, but also genes containing pieces of elements remaining from transposition.
- Maize produces an additional abundant class of repeat derived siRNAs, generated through a pathway independent of *MEDIATOR OF PARAMUTATION1* (*mop1-1*), an orthologue of the *RNA-DEPENDENT RNA POLYMERASE2* (*RDR2*) in *Arabidopsis*.
- In an F1 hybrid, the activity of these repeat derived siRNAs in *trans* could produce novel gene expression that is not present in either parent.
- Are the patterns of sRNA accumulation between inbreds and their hybrids consistent with many of the properties of heterosis defined by previous studies?

Shoot apices correspond to a stage of rapid leaf initiation and cellular proliferation when heterosis is easily observed, and are expected to be enriched for sRNAs with key regulatory functions for development.

A. Tissue Sampled



B. Mid-parent and best parent heterosis observed for seedling biomass

Genotype	Above Ground Dry Weight (g)		
B73	0.078		
Mo17	0.044	BPH	MPH
B73 x Mo17	0.111	42.10%	81.70%
Mo17 x B73	0.157	101%	157%

C. Summary Statistics of Small RNA Libraries

Genotype	Unfiltered	Distinct Sequences (Filtered ≥5 reads)	
	# Reads	# Seqs	# Reads
B73	8,635,395	47,436	749,105
Mo17	6,132,828	28,172	475,588
B73 x Mo17	8,526,014	57,115	1,011,801
Mo17 x B73	8,539,768	63,059	1,073,780
All Libraries	31,834,005	107,768	3,310,274

What modes of gene action do different classes of small RNAs exhibit?

A. Filter criteria based on sequence abundance

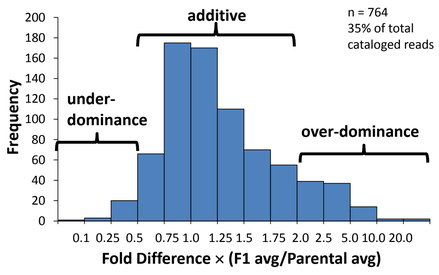
107,768 sequences (100% of total reads)

At least 4.6 reads per million (rpm) in:
1) All four libraries (B)
2) 1 parent and both hybrid libraries* (C)

* At most 0.65 rpm in the other parents' library

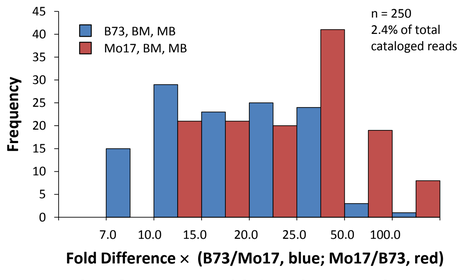
1,014 sequences (37.4% of totals reads)

B. Filter 1: F1 deviation from mid-parent abundance



- 22 miRNAs fall between 0.5 and 2.0
- 1 miRNA is 2-fold higher in parents
- 3 miRNAs are 2-fold higher in the hybrids
- 1 MITE sequence is 2-fold higher in the hybrids
- 55 rDNA sequences are 2-fold higher in the hybrids

C. Filter 2: Extreme levels of dominance



- 87 retrotransposon sequences
- 69 sequences that do not map perfectly to other parent's genome
- 8 Sequences with SNPs between parents (below)

D. Examples of sequences with SNPs between the parents

21-nt siRNA that perfectly matching only one parent's genome and were not cataloged in other parent's library	Abundance (rpm)			
Signature	B73	Mo17	BM	MB
TGGGAGGATTGATAGGTGCGA	33	0	13	19
CGGGAAGATTGATAGGTGCGA	0	51	44	53
**** **********				
AGTGATGAAACAAAGACTGCG	15	0	13	12
AGCGATGAAACAAAGACTGCG	0	75	58	57
** **********				
AGAACAGATAGAAAGGTGTGA	5	0	5	6
AGAACGGAGAGAAAGGTGTGA	0	10	8	11
***** ** **********				

SNPs present in seed sequence.

Divergent regulatory functions?

B73 sequences map to same cluster.

Summary:

- sRNAs abundantly generated in each parent frequently accumulate to similar levels in the hybrids.
- Repeat associated siRNAs (ra-siRNAs) appear to exhibit large relative differences between the parents and between the parents and the hybrids more frequently than miRNA genes.

Are core shoot development miRNAs non-additively expressed between the parents and offspring?

A. Deep Sequencing Experiment

	B73	Mo17	B73 x Mo17	Mo17 x B73
miR156	2830	4375	2243	1267 ←
miR159	17	9	37	34
miR160	8	5	12	14
miR162	2	3	5	2
miR164	4	15	10	8
miR166	7359	4716	9188	6577
miR167	162	252	341	310
miR168	3617	4830	6517	7636
miR169	6	17	8	13
miR171	13	11	15	11
miR172	72	101	148	178 ←
miR319	9	15	12	11
miR390	3	4	7	7
miR393	2	3	2	1
miR394	4	9	14	13
miR396	6	8	6	2
miR528	9	11	18	23
miR529	1	1	1	1
miR827	6	22	3	2

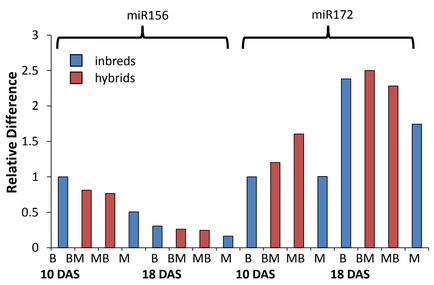
Normalized abundance of sequences matching mature miRNAs from each family.

←Tested in two similar experiments; data for one experiment shown at the right.

B. Additional Greenhouse Experiment

Genotype	DAS, stage	Above Ground Biomass	
		BPH (%)	MPH (%)
B73 x Mo17	10 DAS, 3 leaves	8.42	18.32
Mo17 x B73	10 DAS, 3 leaves	27.54	39.18
B73 x Mo17	18 DAS, 5 leaves	17.91	24.55
Mo17 x B73	18 DAS, 5 leaves	57.75	66.63

*Inbreds sampled also had 3 or 5 leaves emerged.



ΔΔCt values for miR156 and miR172 as the target genes and miR168 as the reference gene and B73 as reference sample. Expression was assessed using Taqman miRNA q-RT-PCR (triplicate reactions). Shoot apices were isolated from four plants grown in the green house. Tissues were pooled prior to RNA isolation.

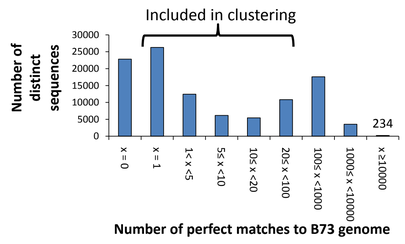
Summary:

- For this cross, the rate of miR156 decay for the hybrids falls between the range of their parents. miR172 expression levels rise earlier in hybrids, suggesting the hybrids become adults slightly earlier than their parents.
- Other potential candidates to test for non-additive expression between the parents and hybrids are miR159, miR167, miR528, and miR827.

Clustering of sequences to B73 genome

A. Clustering scheme

- Sequences within 500 bp of each other based on their B73 location were clustered together.
- GFF intersect perl program was used to merge cluster data with B73 GFF files (WGS, repeats).



B. Clustering of sequences mapping once B73 genome

	# clusters ≥5 sequences	avg size of cluster (bp)	avg sum of reads (rpm)	# of clusters overlapping w/ WGS GFF	# of clusters overlapping w/ MIPs repeat GFF
21-nt majority	10	169	402	6	1
22-nt majority	112	780	71	35	75
24-nt majority	668	114	16	76	272

A few clusters contain sequences with SNPs between the parents. Example shown below.

```
                    CAGTCTTTGTTTCATCaCTCA
                    CAGTCTTTGTTTCATCgCTCA
B73:  TATGTGACGGCGCTATCCTTACCGAGCTTTCATGCATGTACCACGCCCTCTTTCCTCTTCGCAGTCTTTGTTTCATCACTCA
B73:  ATACACTGCCGCGATAGGAATGGCTCGAAAGTACGTACATGGTGCGGGAGAAAGGAGAAGCGTCAGAAACAAAGTAGTGAGT
Mo17: TATGTGACGGCGCTATCCTTACCGAGCTTTCATG----TACCACGTCCTCTTTTCTCTCCGCAGTCTTTGTTTCATCGCTCA
Mo17: ATACACTGCCGCGATAGGAATGGCTCGAAAGTAC----ATGGTGCAGGAGAAAAGAGAGGCGTCAGAAACAAAGTAGCGAGT
      ********************************----********.********.*****.******************.****
                      GAGAgGCGTCAGAAACAAAGT
                      GAGAaGCGTCAGAAACAAAGT
                       AgGCGTCAGAAACAAAGTAGc
                       AaGCGTCAGAAACAAAGTAGt
                        gGCGTCAGAAACAAAGTAGcG
                        aGCGTCAGAAACAAAGTAGtG
                         GCGTCAGAAACAAAGTAGcGA
                         GCGTCAGAAACAAAGTAGtGA
                          CGTCAGAAACAAAGTAGcGAG
                          CGTCAGAAACAAAGTAGtGAG
```

←sense

←antisense

←5′ nt of sRNA is in **bold**, SNP in lower case

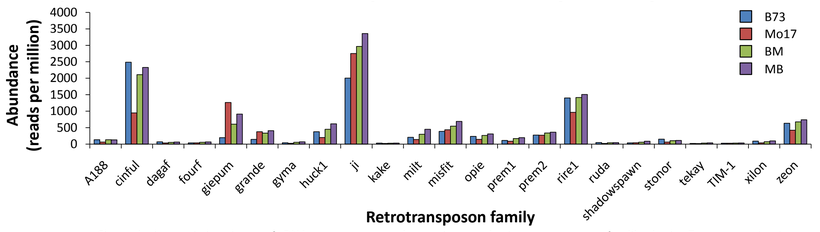
C. Clustering of sequences mapping to B73 genome 1-100 times

	# of clusters overlapping w/ WGS GFF and MIPs repeat GFF			
	Total	avg size of cluster (bp)	avg sum of reads (rpm)	w/ reads only in M, BM, MB libraries
22-nt majority	520	1426	105	45 ←
24-nt majority	3280	382	56	48 ←

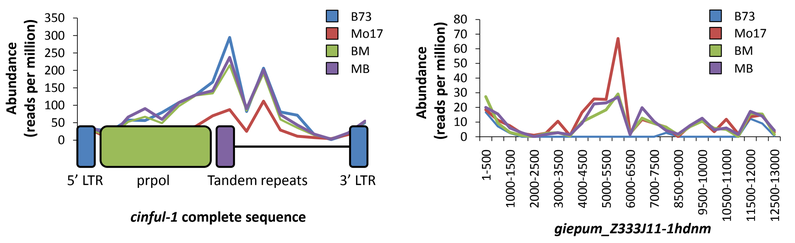
ra-siRNAs generated exclusively from one parent have the potential to act in *trans* on the other parent's genome.

Acknowledgements:
We thank the **National Science Foundation Plant Genome Research Program** and the **University of Illinois Critical Research Initiatives Program** for supporting this project. We also thank other members of the Hudson and Moose Labs for thoughtful discussion.

Parents pass on differences in retrotransposon siRNA activity to the hybrids.



Shown is the total abundance of sRNA sequences mapping to characterized retrotransposon families in the *Zea* repeats database. Matching sequences were uncovered using Patman pattern matching program allowing for 1 mismatch with no gaps. Redundant sequences matching multiple family members were reduced to one occurrence.

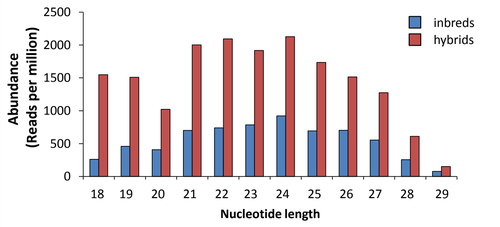


Summary:

- *cinful* retrotransposons are more active in generating siRNAs B73, whereas *giepum* retrotransposons are more active in Mo17.
- Depending on the genomic structure of each parent, parental differences in ra-siRNAs could lead to novel interactions in *trans* in the hybrid.

Do hybrids have a higher abundance of sRNAs mapping to rDNA?

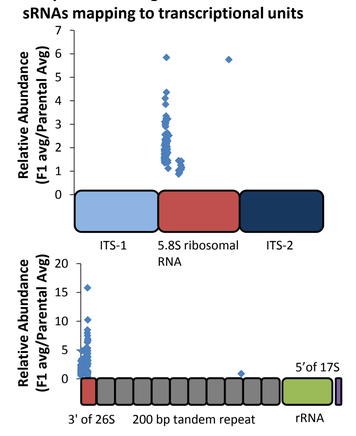
A. Abundance of sRNAs mapping to rDNA in *Zea* repeats database



B. sRNAs map to transcriptional units of rDNA

Zea Repeats DB ID	Abundance (rpm)				Features where sRNAs are concentrated
	B73	Mo17	BMo	MoB	
ZRSiRGRR0000009	3037	3077	6844	7582	5.8S rRNA
ZRSiRGRR0000020	579	713	2633	3202	3' terminus of 17S rRNA
ZRSiRGRR0000019	404	494	2120	2706	3' terminus of 25S rRNA
ZRSiRGRR0000004	831	827	1454	1658	17S rRNA
ZRSiRGRR0000002	350	336	775	973	18S rRNA
ZRSiRGRR0000026	73	66	224	308	28S rRNA
ZRSiRGRR0000029	25	22	92	114	28S rRNA
ZRSiRGRR0000028	2	2	52	88	28S rRNA
ZRSiRGRR0050005	27	27	50	64	5S rRNA, intergenic spacer

C. Hybrids have a greater abundance of sRNAs mapping to transcriptional units



Summary:

- Hybrids have a higher abundance of sRNAs mapping to the transcriptional units of 5.8S, 17S, 25S, 18S, and 28S rDNA.
- Are these sequences degradation products? Do they have a regulatory function?

Heterosis for potential kernel number is not reduced in *mop1-1* mutant hybrids

Genotype	mop1-1/?	plants sampled	total potential kernel #	Std error	% change due to mop1-1/-	% MPH	% BPH
BMo*	+	36	725	20.30		52.20%	39.75%
BMo*	-	22	347	43.55	52.14	118.03%	101.11%
B73	+	19	519	15.01			
B73	-	11	173	18.99	66.74		
Mo17	+	25	434	12.05			
Mo17	-	12	146	17.31	66.41		

*Compared to WT, mutant dry stover biomass was reduced by 10%

WT (left) and mutant hybrids (right) in same row; summer 2009

Summary:

- The penetrance of the mutation may be altered in hybrid, or 24-nt siRNAs may not be important for heterosis.

Future Work:

- Profiling the expression of other key miRNAs during seedling development for BMo.
- Investigating patterns of small RNA accumulation in other tissues for BMo, and for other inbred combinations.
- Confirming observed differences in abundance of ra-siRNAs between B73 and Mo17, and determining if they have any functional significance in an F1 hybrid.
- Confirming observed differences in abundance of siRNAs with SNPs between B73 and Mo17, and determining if they have divergent targets.