

Supplementary Material

microRNAs AND OTHER SMALL RNAs ENRICHED IN THE ARABIDOPSIS RNA-DEPENDENT RNA POLYMERASE-2 MUTANT

Cheng Lu ¹, Karthik Kulkarni ¹, Frédéric F. Souret ¹, Ramesh Muthu Valliappan ¹, Shivakundan Singh Tej ¹, R. Scott Poethig ², Ian R. Henderson ³, Steven E. Jacobsen ^{3,4}, Wenzhong Wang ¹, Pamela J. Green ^{1,5}, and Blake C. Meyers ^{1,5}

¹ Department of Plant and Soil Sciences & Delaware Biotechnology Institute, University of Delaware, Newark, DE 19711, USA.

² Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA.

³ Department of Molecular, Cellular and Developmental Biology, University of California, Los Angeles, CA 90095, USA

⁴ Howard Hughes Medical Institute University of California, Los Angeles, CA 90095, USA

⁵ Corresponding authors. Email meyers@dbi.udel.edu or green@dbi.udel.edu; fax (302) 831-4841

This document contains additional data and analyses that relate to the main published text, but some of this material is not directly discussed in that text.

Supplementary Tables

Table S1. Known miRNAs sequences from wildtype and rdr2.

A. Perfect matches to known miRNAs.

Columns from left to right indicate the name and family member of the known miRNA name, the normalized abundance in TPQ in the *rdr2* and wildtype inflorescence (FLR) MPSS libraries, and the raw abundance in the 454 libraries including the *rdr2* mutant, wildtype inflorescence (Col-0), *rdr6*, *dcl2/3/4*, and *dcl1-7*.

<i>mir_id</i>	MPSS		454				
	<i>rdr2</i>	FLR	<i>rdr2</i>	<i>Col0</i>	<i>rdr6</i>	<i>dcl2/3/4</i>	<i>dcl1-7</i>
<i>miR156a</i>	492	45	5	0	10	0	0
<i>miR156b</i>	492	45	6	0	10	0	0
<i>miR156c</i>	492	45	5	0	10	0	0
<i>miR156d</i>	492	45	5	0	10	0	0
<i>miR156e</i>	492	45	5	0	10	0	0
<i>miR156f</i>	492	45	5	0	10	0	0
<i>miR156g</i>	0	0	0	0	0	0	0
<i>miR156h</i>	173	0	0	0	0	1	0
<i>miR157a</i>	531	0	3	1	36	3	0
<i>miR157b</i>	531	0	3	1	36	3	0
<i>miR157c</i>	531	0	3	2	37	3	0
<i>miR157d</i>	4	0	0	0	0	0	0
<i>miR158a</i>	3107	64	7	3	67	6	0
<i>miR158b</i>	10	0	0	0	0	0	0
<i>miR159a</i>	0	0	233	205	322	377	38
<i>miR159b</i>	0	0	61	48	112	103	13
<i>miR159c</i>	0	0	3	4	38	11	10
<i>miR160a</i>	1373	596	5	4	10	15	1
<i>miR160b</i>	1373	596	5	4	10	15	1
<i>miR160c</i>	1373	596	5	4	10	15	1
<i>miR161</i>	2695	16	31	9	133	38	22
<i>miR162a</i>	893	271	5	4	15	20	2
<i>miR162b</i>	893	271	5	4	15	20	2
<i>miR163</i>	14955	45	209	52	82	232	0
<i>miR164a</i>	1560	465	6	3	2	11	0
<i>miR164b</i>	1560	465	6	3	2	8	0
<i>miR164c</i>	1560	465	1	1	0	0	0
<i>miR165a</i>	326	410	25	9	38	53	5
<i>miR165b</i>	326	410	20	8	27	53	5
<i>miR166a</i>	2083	8546	156	123	258	408	14
<i>miR166b</i>	2083	8546	156	123	258	408	14
<i>miR166c</i>	2083	8546	156	123	258	408	14
<i>miR166d</i>	2083	8546	156	123	258	408	14
<i>miR166e</i>	2083	8546	153	126	219	407	12
<i>miR166f</i>	2083	8546	153	126	219	407	12
<i>miR166g</i>	2083	8546	153	126	219	407	12
<i>miR167a</i>	11039	59392	127	156	1235	331	2
<i>miR167b</i>	11039	59392	116	161	1238	253	2

<i>miR167c</i>	12	16	0	0	0	0	0
<i>miR167d</i>	11039	59392	10	10	175	14	0
<i>miR168a</i>	2025	2205	2	8	17	37	36
<i>miR168b</i>	2025	2205	2	8	17	37	36
<i>miR169a</i>	10905	4485	41	18	249	9	0
<i>miR169b</i>	10905	4485	4	2	21	1	0
<i>miR169c</i>	10905	4485	4	2	23	1	0
<i>miR169d</i>	621	611	8	2	2	4	1
<i>miR169e</i>	621	611	8	2	2	4	1
<i>miR169f</i>	621	611	8	2	2	4	1
<i>miR169g</i>	621	611	8	2	2	4	2
<i>miR169h</i>	2653	2091	176	77	193	6	0
<i>miR169i</i>	2653	2091	196	86	237	6	0
<i>miR169j</i>	2653	2091	197	88	237	6	0
<i>miR169k</i>	2653	2091	176	77	193	6	0
<i>miR169l</i>	2653	2091	197	88	237	6	0
<i>miR169m</i>	2653	2091	177	86	203	6	0
<i>miR169n</i>	2653	2091	197	88	237	6	0
<i>miR170</i>	10180	15603	96	51	119	61	0
<i>miR171a</i>	3220	30	66	73	19	72	2
<i>miR171b</i>	3257	84	23	3	7	23	8
<i>miR171c</i>	3257	84	23	3	7	23	8
<i>miR172a</i>	92371	1873	1712	410	324	1516	10
<i>miR172b</i>	92371	1873	1712	410	324	1516	10
<i>miR172c</i>	92371	1873	548	118	7	502	22
<i>miR172d</i>	92371	1873	548	118	7	502	22
<i>miR172e</i>	1101	47	42	25	3	77	1
<i>miR173</i>	4009	509	40	8	43	17	0
<i>miR319a</i>	301	367	8	1	2	7	0
<i>miR319b</i>	301	367	8	1	2	7	0
<i>miR319c</i>	301	367	10	6	3	4	0
<i>miR390a</i>	16637	11038	141	23	6	67	0
<i>miR390b</i>	16637	11038	141	23	6	67	0
<i>miR393a</i>	972	45	8	4	30	10	0
<i>miR393b</i>	972	45	8	4	30	10	0
<i>miR394a</i>	333	69	3	1	3	3	1
<i>miR394b</i>	333	69	3	1	3	3	1
<i>miR395a</i>	2	0	0	1	0	0	0
<i>miR395b</i>	21	13	0	0	0	0	0
<i>miR395c</i>	21	13	0	0	0	0	0
<i>miR395d</i>	2	0	0	1	0	0	0
<i>miR395e</i>	2	0	0	1	0	0	0
<i>miR395f</i>	21	13	0	0	0	0	0
<i>miR396a</i>	1611	819	1	1	7	6	3
<i>miR396b</i>	1611	819	4	9	20	46	1
<i>miR397a</i>	0	0	0	0	0	1	0
<i>miR397b</i>	0	0	0	0	0	1	0
<i>miR398a</i>	228	111	0	0	1	1	0
<i>miR398b</i>	228	111	2	1	18	35	0
<i>miR398c</i>	228	111	2	1	18	35	0

<i>miR399a</i>	3	0	0	0	1	0	0
<i>miR399b</i>	87	9	0	0	3	1	0
<i>miR399c</i>	87	9	0	0	3	1	0
<i>miR399d</i>	3	0	0	0	0	0	0
<i>miR399e</i>	3	0	0	0	0	0	0
<i>miR399f</i>	3	0	0	0	2	0	0
<i>miR400</i>	109	0	0	0	16	0	1
<i>miR401</i>	0	0	0	0	0	0	0
<i>miR402</i>	117	6	0	0	0	0	0
<i>miR403</i>	306	73	2	2	4	2	0
<i>miR404</i>	0	0	0	0	0	0	0
<i>miR405a</i>	0	0	0	0	0	0	0
<i>miR405b</i>	0	0	0	0	0	0	0
<i>miR405d</i>	0	0	0	0	0	0	0
<i>miR406</i>	0	0	0	0	0	0	0
<i>miR407</i>	0	0	0	0	0	0	0
<i>miR408</i>	12	0	1	0	1	6	3
<i>miR413</i>	0	0	0	0	0	0	0
<i>miR414</i>	0	0	0	0	0	0	0
<i>miR415</i>	0	0	0	0	0	0	0
<i>miR416</i>	0	0	0	0	0	0	0
<i>miR417</i>	0	0	0	0	0	0	0
<i>miR418</i>	0	0	0	0	0	0	0
<i>miR419</i>	0	0	0	0	0	0	0
<i>miR420</i>	0	0	0	0	0	0	0
<i>miR426</i>	0	0	0	0	0	0	0
<i>miR447a</i>	0	0	0	0	0	0	0
<i>miR447b</i>	0	0	0	0	0	0	0
<i>miR447c</i>	0	0	0	0	0	0	0

B. Known miRNAs sequences from wildtype and *rdr2* allowing for small differences in start sites. This is a version of the table above in part (A), but allowing small RNAs that match in up to the +2 to -2 positions compared to the annotated miRNA.

<i>miRNA</i>	MPSS		454				
	<i>rdr2</i>	FLR	<i>rdr2</i>	<i>Col0</i>	<i>rdr6</i>	<i>dcl2/3/4</i>	<i>dcl1-7</i>
<i>miR156a</i>	496	45	5	0	10	0	0
<i>miR156b</i>	496	45	6	0	10	0	0
<i>miR156c</i>	496	45	5	0	10	0	0
<i>miR156d</i>	787	45	8	0	11	0	1
<i>miR156e</i>	493	45	5	0	10	0	0
<i>miR156f</i>	493	45	5	0	10	0	0
<i>miR156g</i>	1	0	0	0	0	0	0
<i>miR156h</i>	187	0	0	1	0	1	0
<i>miR157a</i>	535	0	3	1	36	3	0
<i>miR157b</i>	535	0	3	1	36	3	0
<i>miR157c</i>	535	0	3	2	37	3	0
<i>miR157d</i>	153	0	1	0	1	1	0
<i>miR158a</i>	3241	64	8	3	73	8	0

<i>miR158b</i>	10	10	0	0	4	0	0
<i>miR159a</i>	105	64	237	207	324	382	38
<i>miR159b</i>	105	64	61	48	114	103	13
<i>miR159c</i>	54	11	3	4	38	11	10
<i>miR160a</i>	1387	597	5	4	11	16	2
<i>miR160b</i>	1375	596	5	4	10	15	1
<i>miR160c</i>	1387	597	5	4	11	16	2
<i>miR161</i>	4248	913	54	22	212	73	37
<i>miR162a</i>	932	275	6	4	15	21	2
<i>miR162b</i>	932	275	6	4	15	21	2
<i>miR163</i>	15092	83	210	52	82	234	0
<i>miR164a</i>	1560	467	6	3	2	11	0
<i>miR164b</i>	1560	467	6	3	2	8	0
<i>miR164c</i>	1560	467	1	1	0	0	0
<i>miR165a</i>	395	642	25	10	38	55	5
<i>miR165b</i>	997	854	20	8	27	53	5
<i>miR166a</i>	3270	10059	168	131	263	432	14
<i>miR166b</i>	2762	9214	164	124	260	416	15
<i>miR166c</i>	2159	9005	159	123	259	408	14
<i>miR166d</i>	2159	9005	159	123	259	408	14
<i>miR166e</i>	2762	9214	161	127	221	415	13
<i>miR166f</i>	2760	9166	161	127	221	415	13
<i>miR166g</i>	2159	9005	156	126	220	407	12
<i>miR167a</i>	11039	59519	129	157	1244	331	2
<i>miR167b</i>	11039	59519	118	162	1249	253	2
<i>miR167c</i>	12	45	0	0	1	0	0
<i>miR167d</i>	11049	59574	10	10	179	14	0
<i>miR168a</i>	2100	2267	2	8	17	37	36
<i>miR168b</i>	2100	2267	2	8	17	37	36
<i>miR169a</i>	11243	4892	47	19	256	9	0
<i>miR169b</i>	11243	4892	6	2	22	1	0
<i>miR169c</i>	11140	4842	4	2	24	1	0
<i>miR169d</i>	921	1063	13	3	5	4	1
<i>miR169e</i>	921	1063	13	3	5	4	1
<i>miR169f</i>	921	1063	10	3	3	4	1
<i>miR169g</i>	919	1063	10	3	3	4	2
<i>miR169h</i>	2931	2455	181	77	195	6	0
<i>miR169i</i>	2931	2455	202	86	240	6	0
<i>miR169j</i>	2895	2455	201	88	240	6	0
<i>miR169k</i>	2931	2455	181	77	195	6	0
<i>miR169l</i>	2888	2448	201	88	240	6	0
<i>miR169m</i>	2931	2455	182	86	205	6	0
<i>miR169n</i>	2895	2455	201	88	240	6	0
<i>miR170</i>	10180	15704	98	52	122	61	0
<i>miR171a</i>	3220	288	66	73	21	74	2
<i>miR171b</i>	3257	88	23	3	7	23	8

<i>miR171c</i>	3257	88	23	3	7	23	8
<i>miR172a</i>	92487	1894	1732	413	326	1537	10
<i>miR172b</i>	92487	1894	1732	413	326	1537	10
<i>miR172c</i>	92487	1894	551	118	7	504	22
<i>miR172d</i>	92487	1894	551	118	7	504	22
<i>miR172e</i>	1178	68	54	28	5	95	1
<i>miR173</i>	4010	519	44	9	44	19	0
<i>miR319a</i>	395	532	11	2	2	12	0
<i>miR319b</i>	301	372	10	2	2	7	0
<i>miR319c</i>	427	372	16	6	6	9	0
<i>miR390a</i>	17445	11349	158	25	7	84	0
<i>miR390b</i>	17445	11349	158	25	7	84	0
<i>miR393a</i>	972	49	8	4	31	10	0
<i>miR393b</i>	972	49	8	4	31	10	0
<i>miR394a</i>	396	80	3	1	3	3	1
<i>miR394b</i>	396	80	3	1	3	3	1
<i>miR395a</i>	2	0	0	1	0	0	0
<i>miR395b</i>	21	13	0	0	0	0	0
<i>miR395c</i>	21	13	0	0	0	0	0
<i>miR395d</i>	2	0	0	1	0	0	0
<i>miR395e</i>	2	0	0	1	0	0	0
<i>miR395f</i>	21	13	0	0	0	0	0
<i>miR396a</i>	1611	820	1	1	8	6	3
<i>miR396b</i>	1611	819	4	9	20	46	1
<i>miR397a</i>	0	0	0	0	0	1	0
<i>miR397b</i>	0	0	0	0	0	1	0
<i>miR398a</i>	228	111	0	0	1	1	0
<i>miR398b</i>	228	111	2	1	18	35	0
<i>miR398c</i>	228	111	2	1	18	35	0
<i>miR399a</i>	4	0	0	0	1	0	0
<i>miR399b</i>	87	9	0	0	3	1	0
<i>miR399c</i>	87	9	0	0	3	1	0
<i>miR399d</i>	4	0	0	0	0	0	0
<i>miR399e</i>	4	0	0	0	0	0	0
<i>miR399f</i>	4	0	0	0	3	0	0
<i>miR400</i>	109	0	0	0	16	0	1
<i>miR401</i>	0	0	0	0	0	0	0
<i>miR402</i>	123	6	0	0	0	0	0
<i>miR403</i>	307	73	2	2	4	2	0
<i>miR404</i>	0	0	0	0	0	0	0
<i>miR405a</i>	0	0	0	0	0	0	0
<i>miR405b</i>	0	0	0	0	0	0	0
<i>miR405d</i>	0	0	0	0	0	0	0
<i>miR406</i>	0	0	0	0	0	0	0
<i>miR407</i>	0	0	0	0	0	0	0
<i>miR408</i>	115	385	1	1	1	9	3

<i>miR413</i>	0	0	0	0	0	0	0
<i>miR414</i>	0	0	0	0	0	0	0
<i>miR415</i>	0	0	0	0	0	0	0
<i>miR416</i>	0	0	0	0	0	0	0
<i>miR417</i>	0	0	0	0	0	0	0
<i>miR418</i>	0	0	0	0	0	0	0
<i>miR419</i>	0	0	0	0	0	0	0
<i>miR420</i>	0	0	0	0	0	0	0
<i>miR426</i>	0	0	0	0	0	0	0
<i>miR447a</i>	0	0	0	0	0	0	0
<i>miR447b</i>	0	0	0	0	0	0	0
<i>miR447c</i>	0	0	0	0	0	0	0

Table S2. New miRNAs and other rdr2-independent small RNAs identified by deep sequencing.
A. New miRNAs.

miRNA	Sequence	Wildtype MPSS (TPQ)	rdr2 MPSS (TPQ)	RNA gel blot results				Venn position in Fig. 3
				wt	rdr2	rdr6	dc1/1-7	dc12/3/4
miR771 ^a	TGAGCCTCTGTGGTAGCCCTC	225	669	+	+	+	-	+
miR772 ^a	TTTTTCTACTCCGCCCATAC	7	60	+	+	+	-	+
miR773 ^a	TTTGCTTCCAGCTTTTGTCTC	98	432	+	+	+	-	+
miR774	TTGGTTACCCATATGGCCATC	79	242	+	+	+	-	+
miR775	TTCGATGTCTAGCAGTGCCAA	270	1196	-	+	+	-	+
miR776	TCTAAGTCTTCTATTGATGTT	7	456	+ ^a	+	+	-	+
miR777	TACGCATTGAGTTTCGTTGCT	13	62	+	+	+	-	+
miR778	TGGCTTGGTTTATGTACACCG	5	40	+	+	+	-	+
miR779	TTCTGCTATGTTGCTGCTCAT	5	45	+	+	+	-	+

B. Other RDR2-independent small RNAs.

small ID	Sequence	Wildtype MPSS (TPQ)	rdr2 MPSS (TPQ)	RNA gel blot results				Venn position in Fig. 3
				wt	rdr2	rdr6	dc1/1-7	dc12/3/4
small49	AGGACCATTCGGGTTGTGCAA	57	343	+	+	- ^b	-	-
small57	TGCGGGAAGCATTTGCACATG	23	227	+	+	+	+	-
small58	TACCGCAAGATCAAAGTTTAC	0	17	+ ^b	+	-	-	-
small62	CAACTCCAGGATTGGACCAGT	0	47	-	-	-	-	-

See Figure 4 for RNA gel blot analyses of these sequences.

^a Indicates that this small RNA was previously reported as a potential miRNA (Lu et al., 2005), but was not previously confirmed or submitted to the miRNA registry.

^b Indicates the bands for these small RNAs in the indicated background were weak.

Table S3. Predicted targets of new miRNAs and ta-siRNAs.

A. Predicted targets of new miRNAs.

Small RNA	Target Family ^a	Target Gene IDs (score)	# of Targets	Target Site
miR772	NBS-LRR disease resistance genes	At1g51480 (1), At5g43740 (1), At1g12290 (1.5), At1g12210 (1.5), At5g63020 (1.5), At4g14610 (2), At4g10780 (2), At1g12220 (2), At1g15890 (2), At1g12280 (2.5), At5g47260 (2.5), At5g05400 (2.5),	12	ORF
miR773	DNA (cytosine-5-)-methyltransferase	At4g14140 (2), At4g08990 (2.5)	6	ORF
	and others	At4g05390 (2), At3g15330 (2.5), At3g16230 (2.5)		
miR774	F-box family genes	At2g22730 (2)		UTR ?
miR775	galactosyltransferase family gene	At3g19890 (1), At3g17490 (2)	2	ORF
		At1g53290 (2)	1	ORF
miR776		At5g62310 (1.5)	2	ORF
		At1g08760 (1.5)		UTR ?
miR778	SET domain-containing genes	At2g22740 (1.5), At2g35160 (2.5)	2	ORF
miR779	S-locus protein kinase	At2g19130 (2.5)	1	UTR?
miR771		None		
miR777		None		

Score is based on the system described by Jones-Rhoades and Bartel (2004). The number of predicted targets is based on a cut-off score of 2.5.

B. Predicted targets of new ta-siRNAs.

Small RNA	Target Family ^a	Target Gene IDs (score)	# of Targets	Target Site
Small49	n.a.	At4g00600 (3)	2	ORF
		At4g00610 (3)		UTR?
Small58	n.a.	At2g39980 (3)	9	UTR?
		At5g20200 (3)		ORF

As above, the score is based on the system described by Jones-Rhoades and Bartel (2004), but the number of predicted targets is based on a cut-off score of 3.

^a "n.a." indicates "not applicable" because the targets were too diverse to predominantly represent a single family.

Table S4. Genomic loci with features of ta-siRNA loci.

Chr. #	start	end	hits	rdr2	wildtype	rdr2 / wildtype	comments ^a
1	4182124	4182323	11	147	7	21.00	***
1	4354497	4355226	20	188	0	188.00	PPR gene family
1	4368786	4369099	13	1028	58	17.72	
1	5297877	5298129	65	302	22	13.73	
1	23181100	23182270	43	177	1	177.00	PPR gene family
1	23208490	23209751	55	563	51	11.04	PPR gene family
1	23279171	23280268	19	148	6	24.67	PPR gene family
1	23303291	23304571	130	842	70	12.03	PPR gene family
1	23305811	23307450	88	740	35	21.14	PPR gene family
1	23310777	23312267	105	437	34	12.85	PPR gene family
1	23389058	23390321	51	476	81	5.88	PPR gene family
1	23392690	23393912	121	901	107	8.42	PPR gene family
1	23417056	23418359	134	779	89	8.75	PPR gene family
1	23423630	23424830	79	859	50	17.18	PPR gene family
1	23493873	23495043	45	172	1	172.00	PPR gene family
1	23511578	23512642	96	616	27	22.81	PPR gene family
1	23590850	23591523	14	292	6	48.67	PPR gene family
1	25282658	25283382	30	713	105	6.79	*** ^b
2	819173	823134	183	627	34	18.44	***
2	7198149	7198613	61	282	13	21.69	
2	17231588	17231885	26	127	10	12.70	***
4	1318892	1319151	27	133	8	16.63	
4	11383503	11384499	78	164	24	6.83	***
4	13295428	13296124	16	230	14	16.43	*** ^b
5	897027	897335	18	517	35	14.77	***
5	15774898	15775413	50	282	21	13.43	
5	16656600	16658007	36	121	2	60.50	
5	20151669	20151865	42	525	46	11.41	

The filters used to identify these loci are as follows: 1) The sum of abundance in *rdr2* ≥ 100 . 2) The number of distinct small RNAs in *rdr2* ≥ 10 . 3) The ratio of *rdr2* / wt ≥ 5 . 4) The loci do not correspond to miRNAs, known ta-siRNAs, transposons, retrotransposons, or centromeric repeats. "Hits" indicates the number of distinct small RNAs found at each locus in both *rdr2* and wildtype.

^a PPR gene families are noted because they have been described as strong sources of small RNAs (Lu et al., 2005). *** indicates that RNA gel blots were performed using a small RNA sequence selected from this locus (data not shown), which was confirmed to have the expression pattern of a canonical ts-siRNA (present in wildtype, enriched in *rdr2*, absent in *rdr6*, *dcl1-7* and *dcl2/3/4*).

^b These loci also showed phasing similar to known ta-siRNAs, and are shown in more detail, along with the RNA gel blot, in Figure 7.

Supplementary Figures

Figure S1. Comparison of MPSS and 454 sequence data for *rdr2*.

A. Venn diagram representing genome-matched *rdr2* 454 and MPSS sequences from Table 1. To compare the different length 454 and MPSS sequences for the center of the Venn diagram, 454 signatures were counted if an MPSS signature was contained anywhere within the sequence. Because the MPSS signatures are shorter, some match to more than one 454 sequence. “wt” indicates wildtype. **B.** Abundance plot for *rdr2* 454 and MPSS data (genome-matched sequences only). The dotplots indicate the correlation among abundance levels for genome-matching sequences identified by both technologies for both wildtype (wt, on left) and *rdr2* (on right). In order to visualize the distribution at the lower expression levels, a small number of higher abundance data points are not shown. The abundance of each distinct 17 nt MPSS signature was compared to the sum of abundance of all 454 sequences with the same first 17 nt. **C.** Histograms illustrate the number of distinct sequences in each technology for both wildtype and *rdr2* inflorescence libraries. For the two plots of MPSS data, the X-axis indicates a range of the normalized abundance for the distinct signatures (TPQ), whereas for the two 454 plots, the X-axis represents raw values. Compared to the corresponding wildtype libraries, a higher proportion of small RNAs were sequenced multiple times from *rdr2* with both 454 and MPSS indicating that the *rdr2* sequencing is closer to saturation.

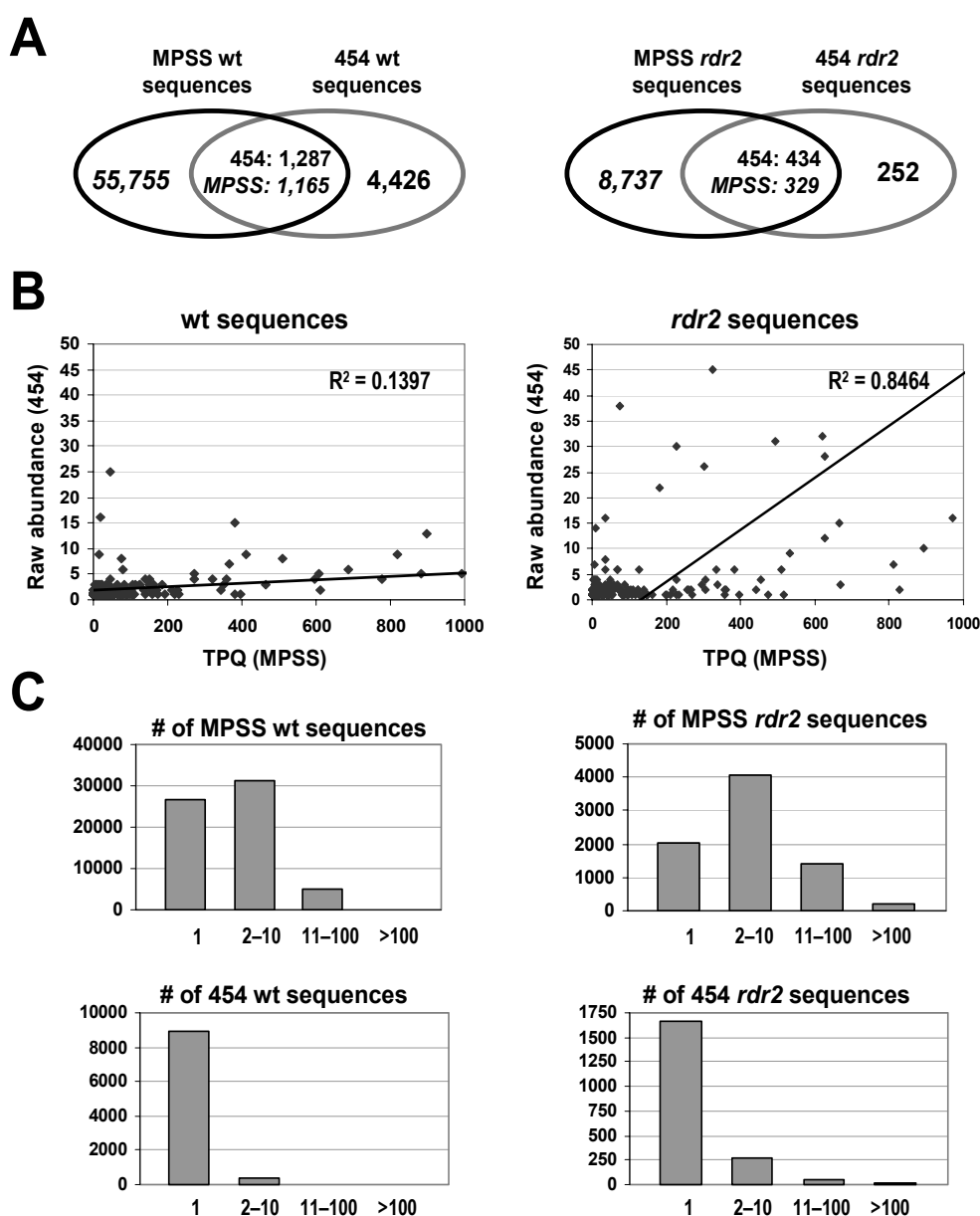


Figure S2. Distribution of *rdr2* and wildtype small RNAs among different genomic features.

Histograms of matches to genomic features for wildtype and *rdr2* MPSS libraries. Wildtype data is indicated by grey bars, *rdr2* data is indicated by black bars. These data are enumerated in Table 3.

A. The number of distinct signatures corresponding to each class of genomic feature.

B. The sum of the abundances (in TPQ) corresponding to the distinct signatures in each class of genomic feature.

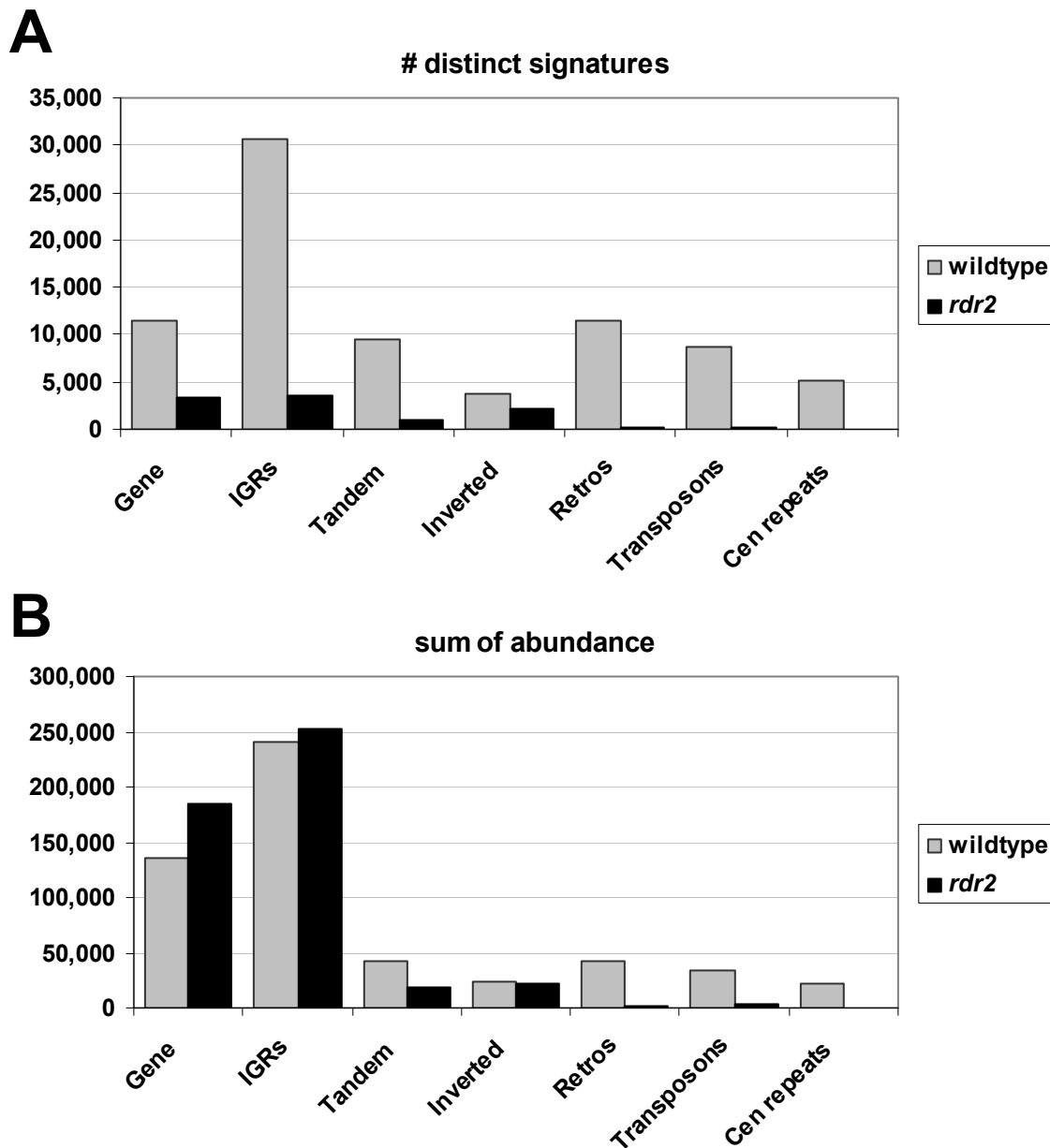


Figure S3. Potential secondary structures of new miRNA precursors.

Secondary structures were predicted for the nine new miRNAs. These structures were predicted using mFOLD (<http://www.bioinfo.rpi.edu/applications/mfold/>). The miRNA sequences identified by MPSS analysis are indicated with curly braces. The RNA gel blots for these small RNAs are shown in Figure 4. **A.** Genomic region encoding miR771. The region is on Chr. 3 between AT3G53010 and AT3G53020. **B.** Genomic region encoding miR772. The region is on Chr.1 between AT1G12290 and AT1G12300. **C.** Genomic region encoding miR773. The region is on Chr.1 between AT1G35500 and AT1G35510. **D.** Genomic region encoding miR774. The region is on Chr.1 between AT1G60070 and AT1G60075.

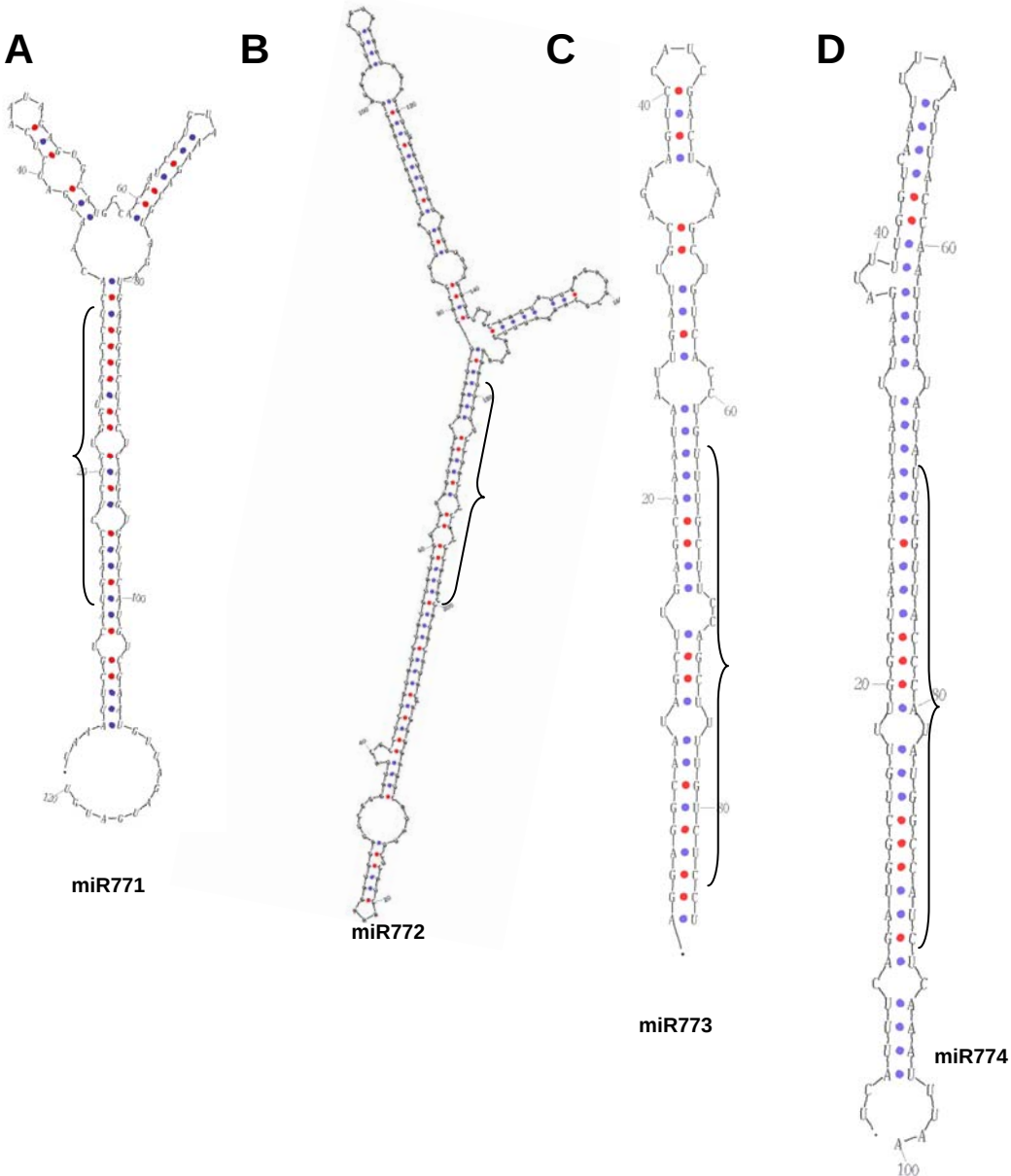


Figure S3. Continued. **E.** Genomic region encoding miR775. The region is on Chr. 1 between AT1G78200 and AT1G78210. **F.** Genomic region encoding miR776. The region is on Chr. 1 between AT1G61730 and AT1G61740. **G.** Genomic region encoding miR777. The region is on Chr. 1 between AT1G70640 and AT1G70650. **H.** Genomic region encoding miR778. The region is on Chr. 2 between AT2G41610 and AT2G41620. **I.** Genomic region encoding miR779. The region is on Chr. 2 between AT2G22490 and AT2G22500.

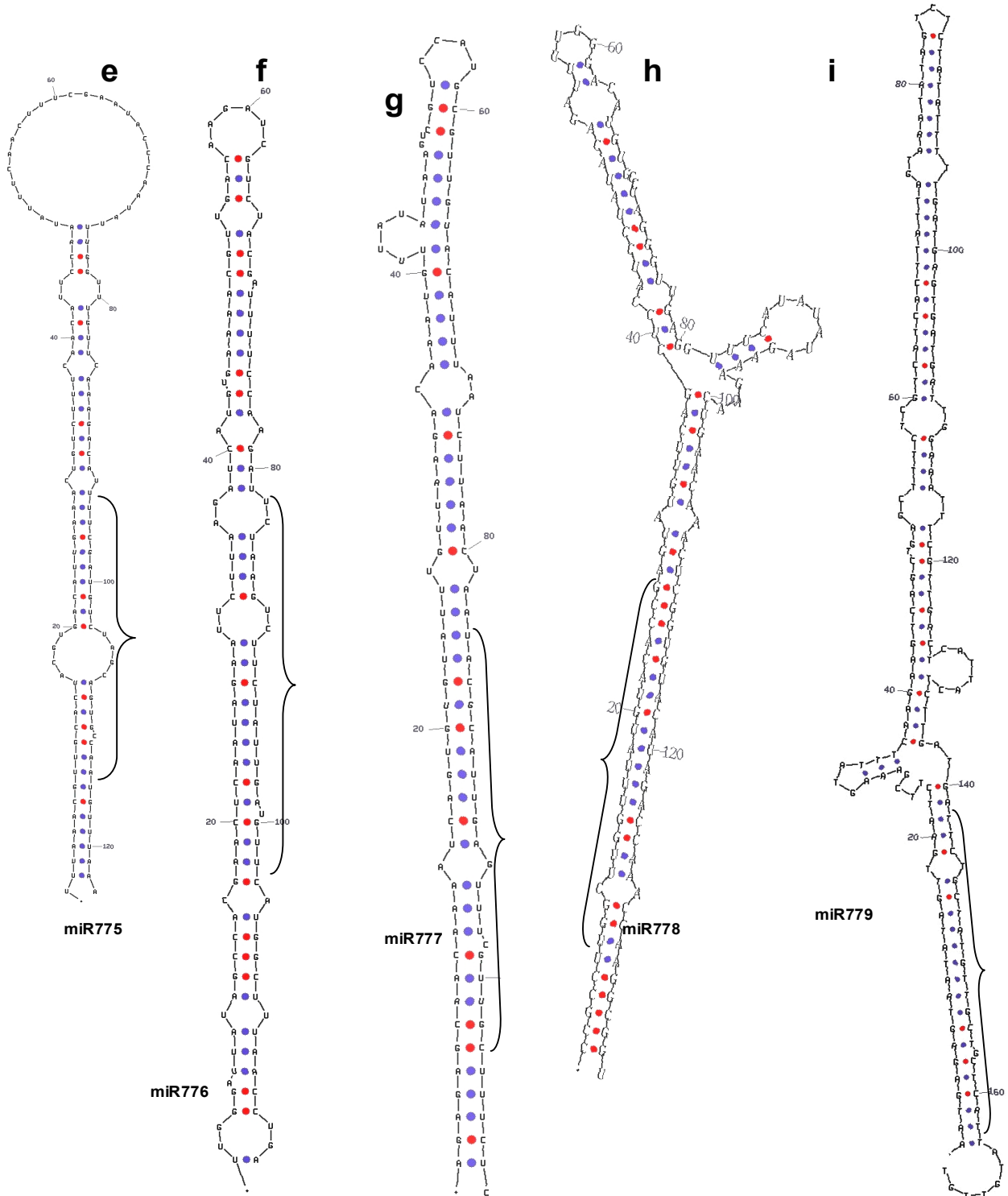
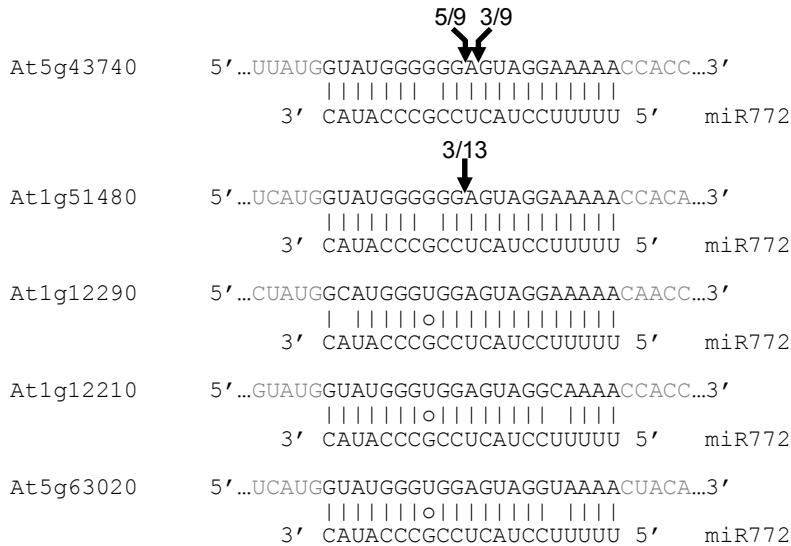


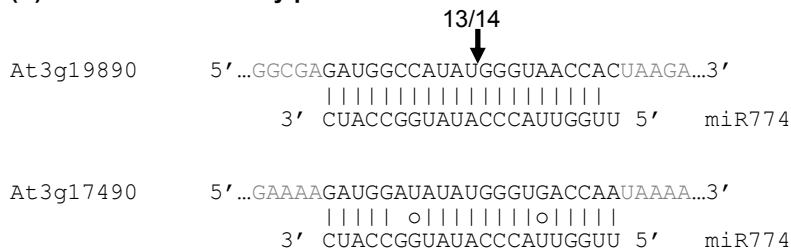
Figure S4. Predicted targets of new miRNAs.

Targets were predicted using the method described by Jones-Rhoades and Bartel (2004). The mRNA target is shown above and the miRNA below in each alignment; matches are indicated with vertical lines, mismatches are unmarked and G-U wobbles are indicated with a circle; grey text indicates nucleotides flanking the target site; for experimentally validated targets, the arrow indicates a site verified by 5' RACE, with the number of cloned RACE products sequenced shown above. In this algorithm, each mismatch is given a score of 1, each wobble (G:U mismatch) is given a score of 0.5, and each bulge is given a score of 2. Only targets with a penalty score of less than or equal to 1.5 are shown in this figure; a complete list of targets scoring 2.5 or less is shown in Table S3.

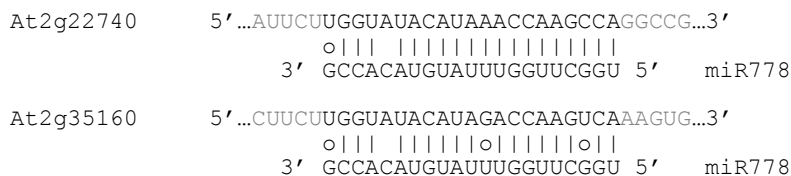
(A) miR772: CC-NBS-LRR class of putative disease resistance proteins



(B) miR774: F-box family proteins



(C) miR778: SET domain-containing proteins (SUVH)

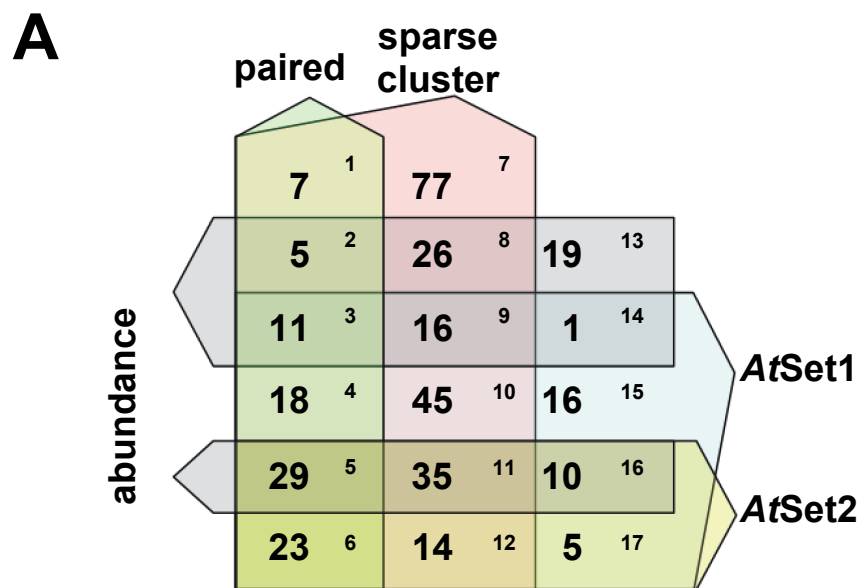


(D) miR773: DNA-methyltransferases



Figure S5. Venn diagram of *rdr2* MPSS sequences.

A. Five-way Venn diagram of selection criteria for miRNAs, using all of the *rdr2* MPSS data. The number of distinct *rdr2* MPSS signatures matching the criteria is indicated in each box numbered in upper right. The figure excludes 13,099 distinct signatures that did not pass any of the criteria. The paired, sparse, abundance filters, and *AtSet1* and *AtSet2* filters are described elsewhere (Lu et al., 2005) but represent potential hairpin structures typical of miRNA precursors, and conservation of those structures in rice, respectively. Includes all perfect matches of small RNA signatures to miRNAs. Some signatures match to multiple genomic locations, so the same known miRNAs may be matched by multiple groups; therefore, the total number of known miRNAs and miRNA families is less than the sum of these columns. **B.** Relationship of sequences in *rdr2* compared to wildtype, from Venn diagrams in part A and in Figure 3.



B

Venn Group	Signatures only in <i>rdr2</i>	Signatures only in wildtype	Signatures in wildtype and <i>rdr2</i>
1	7	825	5
2	5	20	4
3	11	13	9
4	18	54	16
5	29	34	28
6	23	27	16
7	77	19062	65
8	26	101	24
9	16	24	16
10	45	483	36
11	35	45	33
12	14	46	10
13	19	185	14
14	1	20	1
15	16	707	13
16	10	13	10
17	5	31	3

Figure S6. New miRNAs expressed at low levels in wildtype and *rdr2*.

A. Table of new miRNAs identified from among the low abundance *rdr2*-independent small RNAs. See Table 2 for additional details. **B.** Table of the predicted targets of the new miRNAs from (A). As in Table S3, the score is based on the system described by Jones-Rhoades and Bartel (2004). The number of predicted targets is based on a cut-off score of 2.5. **C.** (continued on next page).

A

miRNA	Sequence	Wildtype MPSS (TPQ)	<i>rdr2</i> MPSS (TPQ)	RNA gel blot results				
				wt	<i>rdr2</i>	<i>rdr6</i>	<i>dcl1-7</i>	<i>dcl2/3/4</i>
miR780	TTTCTTCGTGAATATCTGGCA	5	134	+	+	+	-	+
miR781	TTAGAGTTTCTGGATACTTA	0	77	+ ^a	+	+	-	+
miR782	ACAAACACCTTGGATGTTCTT	6	16	+	+	+	-	+
miR783	AAGCTTTGCTCGTTCATGTTT	0	35	+	+	+	-	+

^a Indicates the band for these small RNAs in the indicated background were weak.

B

Small RNA	Target Family ^a	Target Gene IDs (score)	# of Targets	Target Site
miR780		None	1	ORF
miR781	n.a.	At1g26960 (2), At5g23480 (2.5), At1g44900 (2.5)	3	ORF
miR782	n.a.	At5g33405 (2.5)	1	ORF
miR783	Extra-large G-protein-related	At4g01090 (2)	1	ORF

^a "n.a." indicates "not applicable" because the targets were hypothetical proteins or too diverse to predominantly represent a single family.

Figure S6. Continued. C. Folded genomic sequences containing new miRNAs; see Figure S3 for more details on folding. The genomic region encoding miR780 is on Chr. 4 antisense to AT4G14810; miR781 is on Chr. 1 between AT1G21200 and AT1G21210; miR782 is on Chr. 3 between AT3G16220 and AT3G16230; miR783 is on Chr. 1 between AT1G66300 and AT1G66310.

C

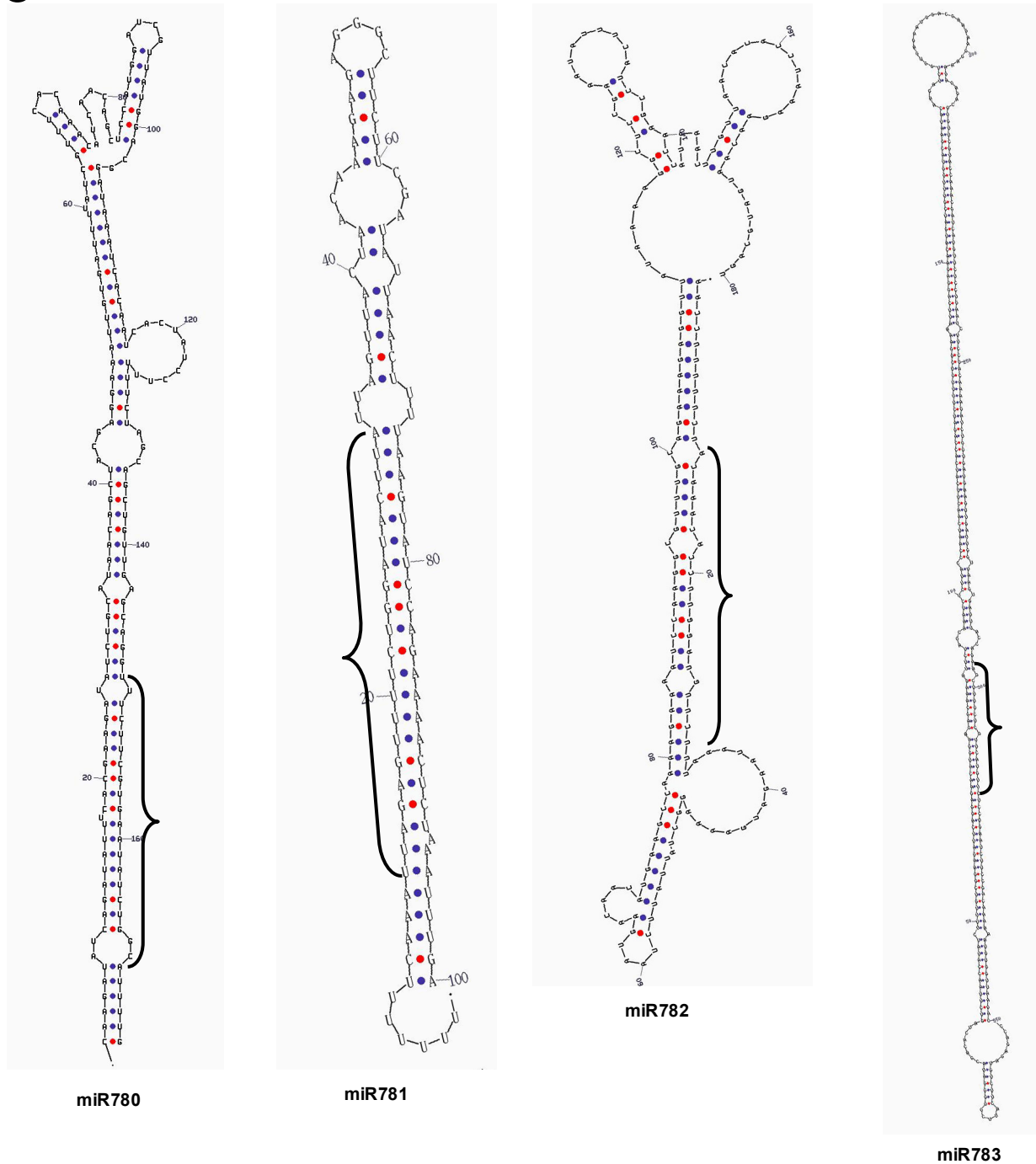


Figure S7. Foldback sequences are sources of numerous rdr2-independent small RNAs.

Inverted repeats are predicted to form “foldback” hairpin structures that are the source of numerous small RNAs in the *rdr2* libraries. Although the difference in the length of the repeat unit is statistically significant between the RDR2-dependent and RDR2-independent sets, some RDR2-independent inverted repeats are quite short (see lower examples). This figure shows views from our website; small RNAs are black triangles, inverted repeats are orange shaded regions. More information for each locus is available at <http://mpss.udel.edu/at>. Open triangles indicate a match to more than one location in the genome; most small RNAs in these inverted repeats match twice, once in each arm of the repeat. Small57 may be an evolving miRNA locus. This locus is the same as ASRP1729 (Allen et al., 2005).

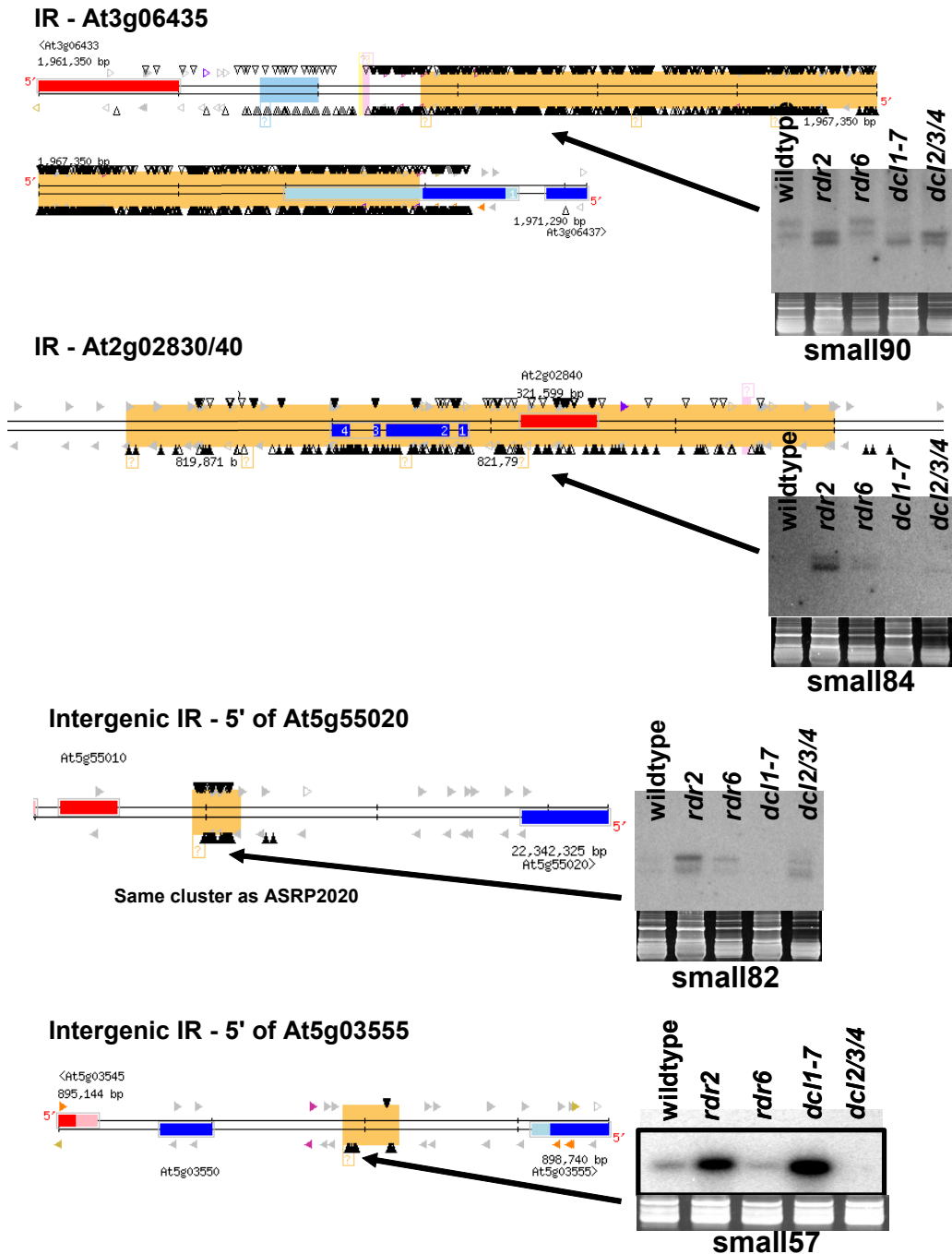


Figure S8. The *A. thaliana* gene encoding SRK contains an inverted repeat that is the source of RDR2-independent small RNAs.

A. An image of the *A. thaliana* SRK locus, with the inverted repeat shown in orange, exons of SRK (At4g21370) indicated as blue boxes, and the annotated adjacent gene (At4g21366) shown in red. **B.** An RNA gel blot of small85 from the SRK locus. **C.** A total of 963 nt of sequence from the inverted repeat spanning At4g21370 and At4g21366 was analyzed using mFold. This sequence is predicted to form a near-perfect double-stranded RNA of 390 bp. Small RNAs were identified by MPSS that matched throughout the stem structure but were absent from the loops.

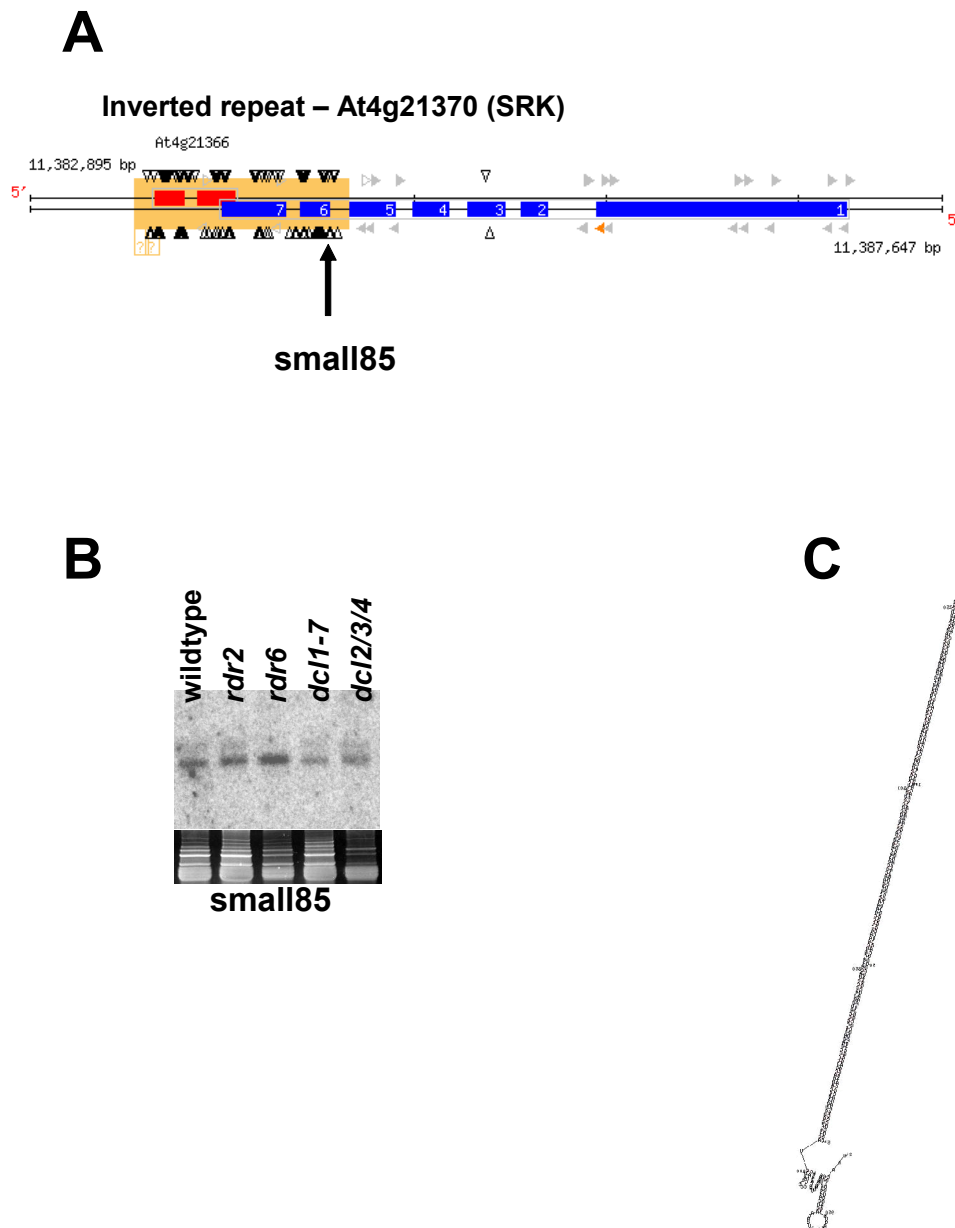


Figure S9. Enrichment of small RNAs at the *TAS1a* locus in *rdr2* compared to wildtype.

Bars indicate the abundance of the small RNAs (MPSS data, in TPQ) found at each position within the locus; bars above the center line indicate the upper strand, bars below the center line indicate the bottom strand. Red bars indicate small RNAs in wildtype and black bars indicate small RNAs in *rdr2*. Due to limited space, non-expressed sites have been removed. The upper and lower boxes are in logarithmic scale to indicate the most abundant small RNAs. The position within the locus is indicated near the bottom, with the zero position indicating the functional ta-siRNA (Peragine *et al.*, 2004; Allen *et al.*, 2005) which is identified by the MPSS signature TTCTAAGTCCAACATAG found at 6169 TPQ in *rdr2*, corresponding to 11,729,063 bp on Chr. 2.

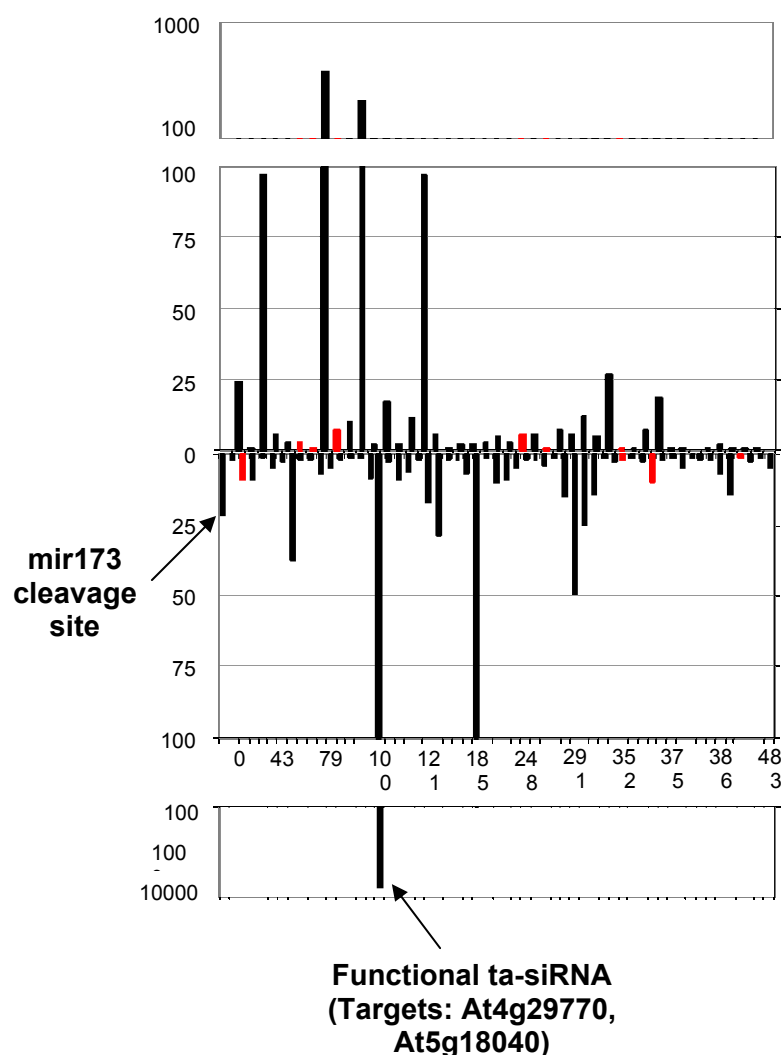
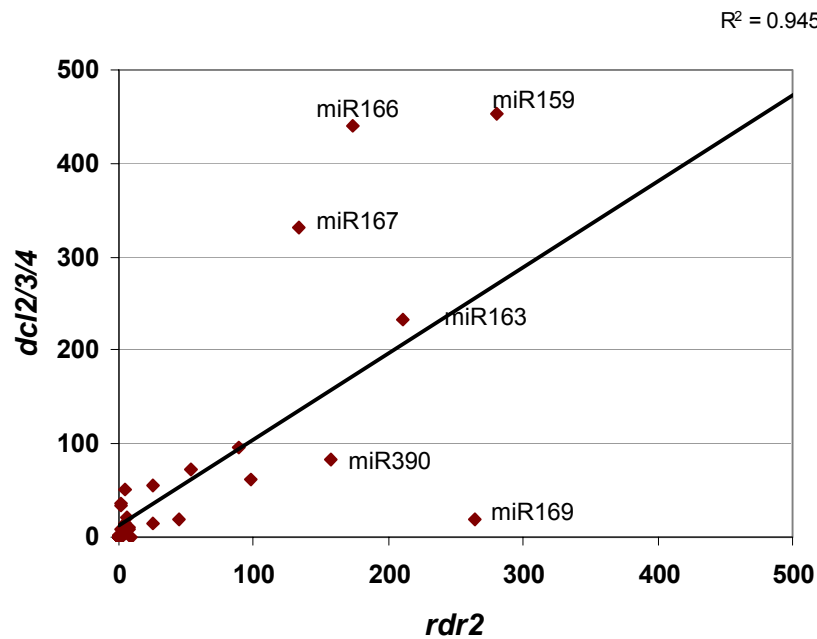


Figure S10. Correlation of miRNA gene abundances in the *rdr2* and the *dcl2/3/4* triple mutant.

The figure is based on the 454 data for these mutant lines shown in Table 2. Due to the plot scale and its abundance, miR172 is not shown. The diagonal line indicates the trend line for the data. The high-abundance miRNA genes are marked for reference. X- and Y-axis values are raw abundances.



Supplementary Files

Supplemental File 1. Inverted and tandem repeats in wildtype and *rdr2* MPSS libraries.

Excel file describing inverted and tandem repeats with matching small RNAs from wildtype and *rdr2* MPSS libraries.