

**Wheat *Mds-1* encodes a heat-shock protein and governs susceptibility towards
the Hessian fly gall midge**

Supplementary Material

Xuming Liu¹, Chitvan Khajuria¹, Jiarui Li², Harold N. Trick², Li Huang³, Bikram S. Gill², Gerald
R. Reeck⁴, Ginny Antony², Frank F. White², and Ming-Shun Chen^{1,5†}

¹Department of Entomology, Kansas State University, Manhattan, KS 66506

²Department of Plant Pathology, Kansas State University, Manhattan, KS 66506

³Department of Plant Sciences and Plant Pathology, Montana State University, Bozeman, MT
59717

⁴Department of Biochemistry, Kansas State University, Manhattan, KS 66506

⁵USDA-ARS, Hard Winter Wheat Genetics Research Unit, Kansas State University, Manhattan,
KS 66506

†Corresponding Author: Ming-Shun Chen, USDA-ARS and Department of Entomology, 123
Waters Hall, Kansas State University, Manhattan, KS 66506.

Phone: 785-532-4719, Fax: 785-532-6232, Email: mchen@ksu.edu or [ming-
shun.chen@ars.usda.gov](mailto:ming-shun.chen@ars.usda.gov)

Table S1. Hessian fly populations tested against *Mds-1*-knockdown seedlings of Newton wheat.

Biotype/Population	Biotype	Reference
Biotype <i>GP</i>	<i>GP</i> (avirulent to all known R genes)	48
Biotype <i>vH9</i>	<i>vH9</i>	49
Biotype <i>vH13</i>	<i>vH13</i>	49
Biotype <i>L</i>	<i>vH3, vH5, vH6, vH7H8</i>	48
Scott-KS-2005	<i>vH3, vH4, vH5, vH6, vH7H8, vH9, vH10, vH11, vH12, vH16, vH17, vH18, vH22, vH23, vH24, vH26</i>	50
Fannin-TX-2007	<i>vH3, vH4, vH5, vH6, vH7H8, vH9, vH11, vH16, vH17, vH18, vH24, vH26, vH31</i>	50
Kay-OK-2007	<i>vH3, vH4, vH5, vH6, vH7H8, vH9, vH10, vH11, vH12, vH16, vH17, vH18, vH22, vH23, vH24, vH25, vH26, vHdic</i>	50

Supplementary Table S2. Characterization of transgenic T0 plants with the *Mds-1* knockdown and ectopic expression constructs, respectively*.

Plant	Tube Date	Soil Date	Bar	GOIa	GOIb	SS	PHF	PPM	Seed set
Transgenic Bobwhite lines with the <i>Mds-1</i> knockdown construct:									
150	3/30/2009	4/28/2009	+	+	+	+	R	NA	yes
152	3/23/2009	4/28/2009	+	+	+	+	R	NA	yes
358	3/30/2009	6/8/2009	+	+	+	+	R	NA	yes
432	4/23/2009	6/16/2009	+	+	+	+	R	NA	yes
439	5/19/2009	6/16/2009	+	+	+	-	S	NA	yes
523	5/19/2009	6/29/2009	+	+	+	+	R	NA	yes
1630	2/8/2010	2/22/2010	+	+	+	+	R	MR	yes
1639	2/8/2010	3/2/2010	+	+	+	+	R	MR	yes
1703	3/1/2010	3/23/2010	+	+	+	+	R	MR	yes
2030	5/17/2010	6/1/2010	+	+	+	+	R	MR	yes
2059	5/17/2010	6/1/2010	+	+	+	-	S	S	dead
2089	5/24/2010	6/8/2010	+	+	+	+	R	MR	heat dead
2090	5/24/2010	6/8/2010	+	+	+	+	R	MR	yes
2095	5/24/2010	6/8/2010	+	+	+	+	R	MR	yes
2101	5/10/2010	6/8/2010	+	+	+	+	R	MR	heat dead
2129	5/17/2010	6/16/2010	+	+	+	+	R	MR	yes
2164	6/1/2010	6/22/2010	+	+	+	+	R	MR	yes
2222	5/24/2010	6/29/2010	+	+	+	-	S	S	dead
2270	6/1/2010	7/7/2010	+	+	+	-	S	S	yes
2297	7/28/2010	9/8/2010	+	+	+	+	R	MR	yes
2348	8/4/2010	9/15/2010	+	+	+	+	R	MR	yes
2349	8/11/2010	9/22/2010	+	+	+	+	R	MR	yes
2355	6/9/2010	7/19/2010	+	+	+	+	R	MR	yes
2357	6/9/2010	7/19/2010	+	+	+	+	R	MR	yes
Transgenic Molly lines with the <i>Mds-1</i> over-expression construct:									
640	6/15/2009	6/29/2009	+	NA	+		S		yes
691	6/22/2009	7/6/2009	+	NA	-		R		Disposed
793	6/22/2009	7/13/2009	+	NA	-		R		yes
1111	6/29/2009	8/10/2009	+	NA	+		S		yes
1117	8/10/2009	8/24/2009	+	NA	+		S		yes
1166	8/10/2009	8/24/2009	+	NA	+		S		yes
1238	8/17/2009	8/31/2009	+	NA	+		S		yes
1417	8/31/2009	9/21/2009	+	NA	+		S		yes
1490	8/17/2009	8/31/2009	+	NA	+		S		yes

*Tube date and soil date are the dates when plants were transferred to tubes with culturing media and to pot soil, respectively. Bar, GOIa, and GOIb represent PCR detection results for the presence of the bar gene (herbicide resistance), inserted gene fragment at antisense, and sense orientation, respectively, in transgenic plants. Positive symbols (+) represent the detection of a DNA fragment with the expected size, whereas negative symbols (-) indicate no PCR amplification at all (Supplementary Figs. S3 and S9b). NA means not applicable or not determined. SS represents successful suppression of *Mds-1* induction by Hessian fly infestation based on northern blot analyses. PHF represents resistant (R) or susceptible (S) phenotypes of the transgenic plants to Hessian fly infestation. Hessian fly biotype *GP* was used for testing. PPM represents resistant or susceptible phenotypes of transgenic plants to powdery mildew infection. Seeds set represents transgenic plants that produced seeds successfully (yes) or did not reach the seed production stage (dead). MR represents moderate resistance (between grade 1 and 2 of disease severity on a scale of 0 to 4) (51, also see Methods). 'Heat dead' represents dead plants due to a malfunction of an overheating growth chamber, whereas 'disposed' means plants were disposed before they reached the flowering stage.

Table S3. Primers used in various experiments.

Gene Name	Purpose or target	Primer
Primers for <i>Mds-1</i> cloning		
Adapter primer	cDNA cloning	Forward: 5'-AAGCAGTGGTATCAACGCAGAGT
<i>Mds-1</i> R	5' RACE	Reverse: 5'-CGCAAGAACTCGCACACACATC
<i>Mds-1</i> La	cDNA cloning	Forward: 5'-CGTGAAGAAGGAGGAGGTCAAG
Adapter primer	3' RACE	Reverse: 5'-AAGCAGTGGTATCAACGCAGAGT
<i>Mds-1</i> Lb	<i>Mds-1</i> full length	Forward: 5'-ACCTGCGACCCAATCCAGAAC
<i>Mds-1</i> R	Cloning	Reverse: 5'-CGTGAAGAAGGAGGAGGTCAAG
Primers for construct		
<i>Mds-1a</i> '	<i>Mds-1</i> RNAi	Forward: 5'-cacCGTGAAGAAGGAGGAGGTCAAG
<i>Mds-1</i> R	construct	Reverse: 5'-CGCAAGAACTCGCACACACATC
<i>Mds-1</i> Lc	<i>Mds-1</i> expression	Forward: 5'-CTAGtgatcATGTCGATCGTGCGGCGGAG
<i>Mds-1</i> R	construct	Reverse: 5'-CGCAAGAACTCGCACACACATC
Primers for qPCR		
<i>Mds-1</i>	JN162442	Forward: 5'-ACAGCATCCTACCTGCGA Reverse: 5'-ATCGTCGGTGTACTTTG
<i>Hb</i>	3BS_10441917	Forward: 5'-CTACAACCCAAAGCAACATC Reverse: 5'-TCAGCAAGAACAAGACACAAC
<i>Hd</i>	3DS_2601138	Forward: 5'-CGAGATCTCCGTCTGAGTGA Reverse: 5'-GGCAAGAACAAGATACTTTTCAG
<i>HSP16.9</i>	CD896621	Forward: 5'-GTGATGAAATAAAACCAAGA Reverse: 5'-GTACAGACTGATGAACAA
<i>HSP16.9</i>	BQ804592	Forward: 5'-TGAGTATTTGGGTCAGAG Reverse: 5'-AAGAAACGCAATACAGAG
<i>HSP16.9</i>	EU649679	Forward: 5'-TGAGTATTTGGGTCAGAG Reverse: 5'-AACAAAGACACAACCTGTAG
<i>HSP16.9</i>	X64618	Forward: 5'-TGAGTATTCGGGTCAGAG Reverse: 5'-GCAAGAACACAGACACAA
EST	CA717853	Forward: 5'-CGGTCCGATACAGATTTTCC Reverse: 5'-CGAAGGGGTGCAACACAT
EST	CJ618891	Forward: 5'-AGATTAAGGTTAAGGTTCAATTCCG Reverse: 5'-CTCAACACGGACCCAAAT
HSP17.8	AF350423	Forward: 5'-CGACATCAAGGTGCAGGTG Reverse: 5'-GTACTTGCGTCCTCCTTCTC
Amino acid transporter	BJ275567	Forward: 5'-TCCCAATAAACGAACCCTAA Reverse: 5'-TGCAAAGGACTTCTTTTCATG
Enolase	CK208852	Forward: 5'-GATTAGAAATAAAGGCACCGA Reverse: 5'-GTCCGTCAGCAAAAATGT

Methionine synthase	CN012484	Forward: 5'-TGTGATAGTGCCAGTTGA Reverse: 5'-GAAGGACAGGAGTTCATCT
Lipid transfer protein	AF334185	Forward: 5'-TGCTCTGATATGATCTCCAT Reverse: 5'-ATCCTATATGATAAGCGTACATC
Family II extra cellular lipase 1	BQ171153	Forward: 5'-TGAACCTCTAGCCTTTACC Reverse: 5'-CAATGGACGCATGAACAA
Type III peroxidase	CK200808	Forward: 5'-GAAGCATATACTCCTACTCTTG Reverse: 5'-CATGTGAGAGATCAGTTAGTT
Actin	AF326781	Forward: 5'-AAATCTGGCATCACACTTTCTAC Reverse: 5'-GTCTCAAACATAATCTGGGTCATC

Table S4. Characterization and comparison of *Mds-1*-silenced transgenic wheat lines*.

Wheat Lines	Plant Height (cm)		Weight per 100 seeds (g)		Germination Rate (%)		Coleoptile Length (cm)		Radicle Length (cm)	
Bobwhite	75.7	a	4.84	b	93.7	a	8.9	a	14.2	a
T1630	65.3	b	4.46	c	78.7	b	6.7	a	9.2	a
T1639	72	a	4.52	c	84.3	b	6.8	a	12.0	a
T2095	57.7	c	5.54	a	65	c	5.4	b	8.3	c
T2357	74	a	4.16	d	83	b	6.9	a	12.0	a

*Germination rates were determined by soaking dry seeds in water for 6 hours. Seeds were then transferred to a plastic dish (150x15 mm) with two completely moistened towels at the bottom and two moistened towels on the top of seeds. Dishes were covered to keep moisture and stored under room temperature. Three replicates were carried out with at least 100 seeds for each. Germination data were scored 48 h later. Data were subjected to ANOVA analysis and Tukey's honestly significant difference (HSD) multiple comparison using ProStat Version 5.5(Poly Software International Inc., Pearl River, NY, USA).

Figure S1

AAGCCAGTGTCAACCAAAAGTAGACAGCATCCTA
CCTGCGACCCAATCCAGAACAGATTTTTTCCCGAGCTAGAAACCAAAGTAACACCGACG
ATGTGCATCGTGC GGCGAGCAGCGTTCGACCCCTTCGCCGACCTCTGGGCTGACCCT
M S I V R R S S V F D P F A D L W A D P
TTCGACACCTTCCGCTCCATCATCCC GGCGATCTCAGGCGGCAACAGCGAGACGGCCGCG
F D T F R S I I P A I S G G N S E T A A
TTCGCCAATGCTCGCATGGACTGGAAGGAGACCCCGAGGCGCACGTCTTCAAGGCCGAC
F A N A R M D W K E T P E A H V F K A D
CTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGGCAACGTGCTCGTC
L P G V K K E E V K V E V E D G N V L V
GTCAGCGGCGAGCGCACAAAGGAGAAGGAGGACAAGAACGACAAGTGGCACCGCGTGGAG
V S G E R T K E K E D K N D K W H R V E
CGCAGCAGCGGCAAGTTCGTCAGGCGGTTCCGCCTCCCCGAGGACGCCAAGGTGGAGGAG
R S S G K F V R R F R L P E D A K V E E
GTGAAGGCCGGGCTGGAGAACGGTGTGCTCACCGTCACCGTGCCCAAGGCCCAGGTCAAG
V K A G L E N G V L T V T V P K A Q V K
AAGCCCGAGGTGAAGGCCATCCAGATCTCCGGCT**TGA**GTGGACGCGTCTCGGCGTATGATC
K P E V K A I Q I S G
ATCAGGGATGGAGCCAGTTTGGTTGATGTGTGTGCGAGTTCTTGCGAGTCTGATGAGACA
TCTCTGTATTGTGTTTCTTTCCCCAGTGTTCCTGTACTTGTGTAATCGGCTAATCGCCA
ACAGATTCGGCGATGAATAAATGAGAAATAAATTGTTCTGATTTTGGGTGC

Figure S2

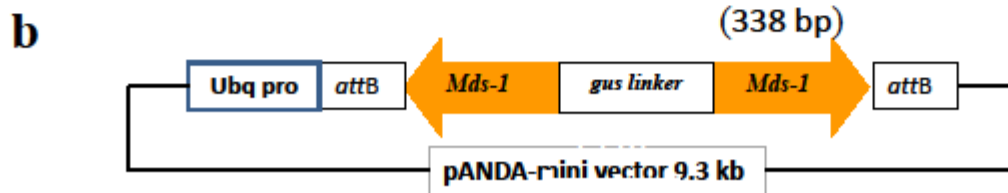
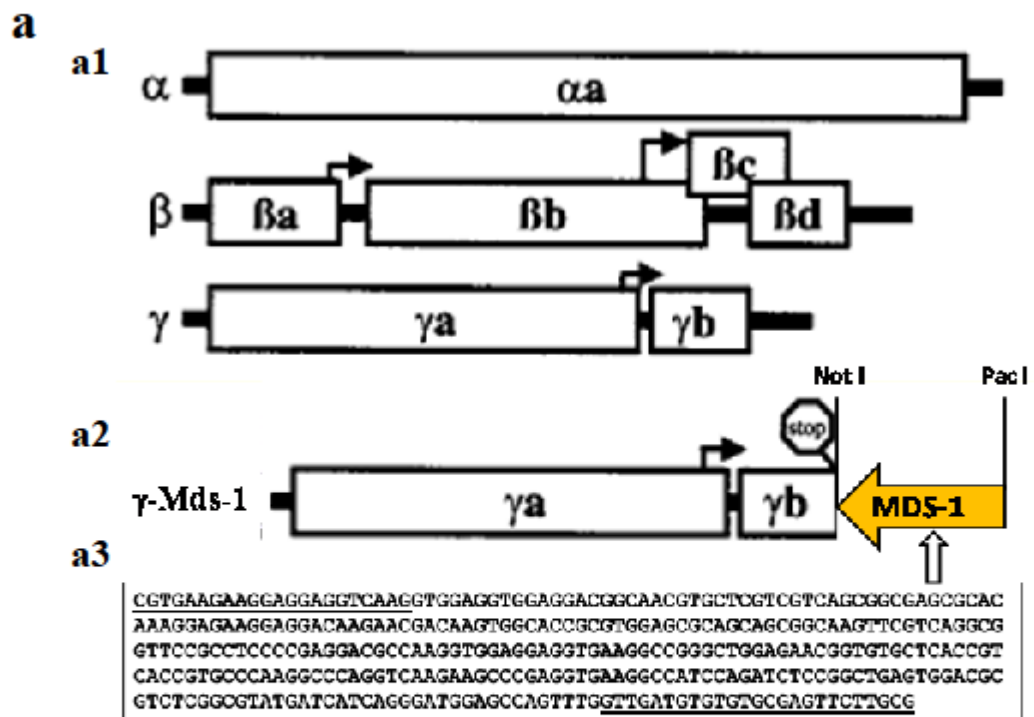


Figure S3

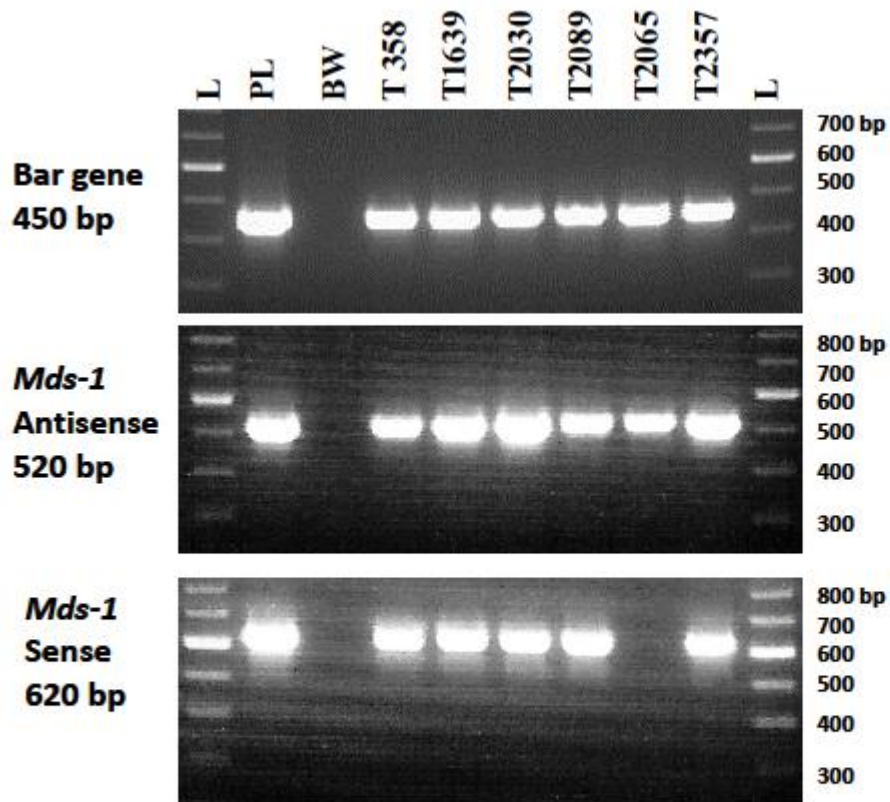


Figure S4

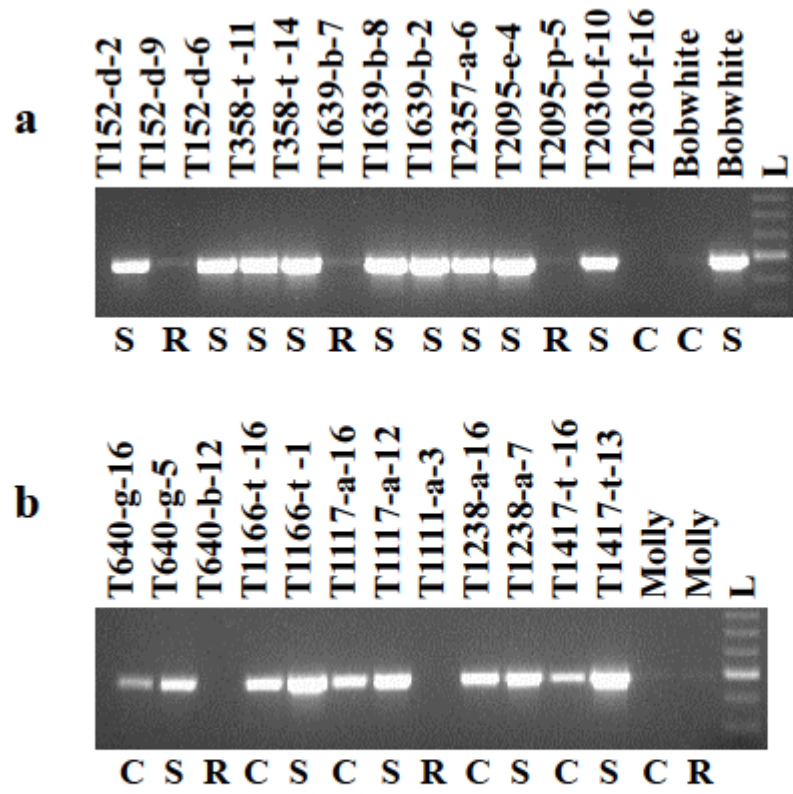


Figure S5

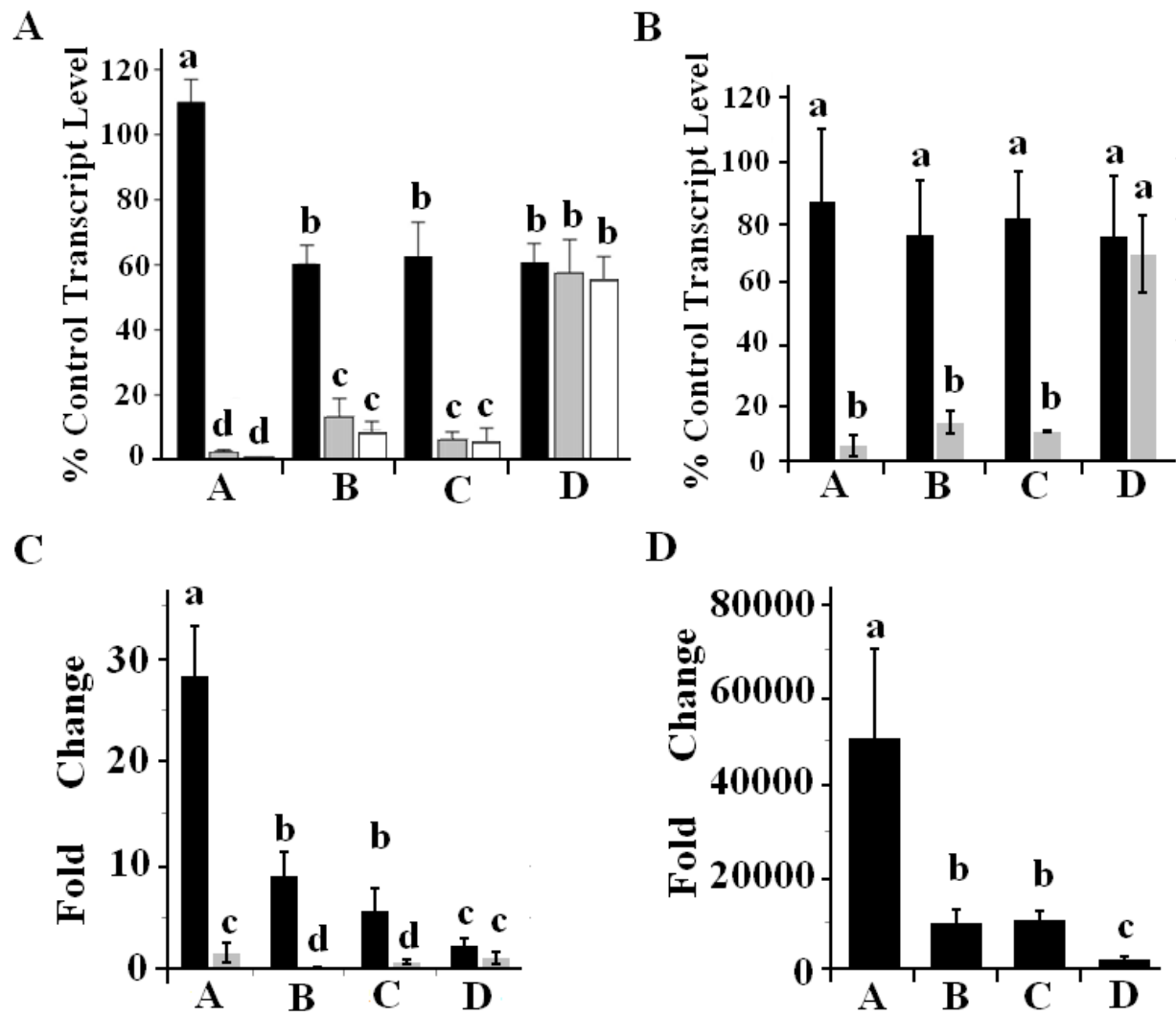


Figure S6

a

MDS-1 AGCCAGTGTCAACCAAAGTAGACAGCATCCTACCTGCGACCCAATCCAGAACAGATTTT
CA717853 -----AGCAAAGCAGACAGCATCCTACCTGCGGTCCGATACAG-----ATTT
BQ804592 -----GCACACAGCATCCTAGTCTACCTACGATC---TACAG-----AATT

MDS-1 TTCCCGAGCTAGAAACAA---AGTAACACCGACGATGTCGATCGTGC GGCGGAGCAGC
CA717853 TCCCGAGCTACAACCAA---AGCAACAACGACATGTCGATCGTACGGCGGAGCAAT
BQ804592 TTCCGAGCGCACAGCGAAACAAATCAACACCGACGATGTCGATCGTGC GGCGGAGCAAC

MDS-1 GTGTTTCGACCCCTTCGCGACCTCTGGGTTGACCCCTTCGACACCTTCCGCTCCATCATC
CA717853 GTGTTTCGACCCCTTCGCGACCTCTGGGTCGACCCCTTCGACACCTTCCGCTCCATCGTC
BQ804592 GTGTTTCGACCCCTTCGCGACCTCTGGGCGACCCCTTCGACACCTTCCGCTCCATCGTC

MDS-1 CCGGCGATCTCAGGCGCAACAGCGAGACGGCCGCGTTCGCCAATGCTCCGATGGACTGG
CA717853 CCGGCGATCTCAGGCGCAACAATGAGACAGCTGCGTTCGCGAACGCCCGGATGGACTGG
BQ804592 CCGGCGATCTCTGGCGGCAACAGCGAGACGGCCGCGTTCGCCAACGCCCGGATGGACTGG

MDS-1 AAGGAGACCCCGAGGCGCACGTCTTCAAGGCCGACCTCCCGCGGTGAAGAAGGAGGAG
CA717853 AAGGAGACCCCTGAGGCAACAGTCTTCAAGGCCGACCTCCCGCGGTGAAGAAGGAGGAG
BQ804592 AAGGAGACCCCGAGGCGCACGTCTTCAAGGCCGACCTCCCGCGGTGAAGAAGGAGGAG

MDS-1 GTCAAGTGGAGGTGGAGGACGGCAACGTGCTCGTCTCAGCGGCGAGCGCACAAAGG-A
CA717853 GTCAAGTGGAGGTGGAGGACGGCAACGTGCTCGTCTCAGTGGCGAGCGCACAAAAGGA
BQ804592 GTCAAGTGGAGGTGGAGGACGGCAACGTGCTCGTCTCAGCGGCGAGCGCACAAAGG-A

MDS-1 GAAGGAGGACAAGAACGACAAGTGGCA-CCGCGTGGAGCGCAGCAGCGCAAGTTCTGCA
CA717853 NAAAGNAGACAAGAACGACAAGTGGCAACCGCGTGGAGCGCAACAACCGNAAATTTGTGC
BQ804592 GAAGGAGGACAAGAACGACAAGTGGCA-CCGCGTGGAGCGCAGCAGCGCAAGTTTGTCA

MDS-1 GCGGTTCCGCTCCCG-AGGACGCCAAGGTGGA-GGAGGTGAAGGCCGGCTGGAGAA
CA717853 GGNCTTCCGGCTNCCCGAAGAACCCTAAGGTGGANGAAGGTNAAGGCCG-CTTGAGAA
BQ804592 GCGGTTCCGGCTCCCG-AGGACGCCAAGGTGGA-GGAGGTGAAGGCCGGCTGGAGAA

MDS-1 CGGTGTGCTCACCGTCACCGTG-CCCAAGGCCAGGTCAAGAAGCCCGAGGTGAAGGCCA
CA717853 AAGGGTGNCAACGTCAACGTNTCCCAAGGCCAGGTAAAAAACCCANNGNAAAGGCAAA
BQ804592 CGGTGTGCTCACCGTCACCGTG-CCCAAGGCCAGGTCAAGAAGCCCGAGGTCAAGGCCA

MDS-1 TCCAGATCTCCGGCTGAGTGGACGCGTCTCGGCGTATGATCATCAGGGATGG-----
CA717853 --CNGATTTCCNNGN-----
BQ804592 TCGAGATATCCGGCTGAGTATTTGGTCAGAGCCAGTCTGGTTCATGTGTGTGCAAGTTC

MDS-1 -----AGCCAGTTGCTGTGATGTGTGTGCGAGTTCTTGG-AGTCTCATGAGACAT--
CA717853 -----
BQ804592 TGCAGCCTAAAAGTATGCTACAGTGTGTCTTTGTTCTCGCCAGTCTGGTGAGGCATTT

MDS-1 ---CTCTGATTGTGTTCTTTCCCAAGTGTCTCTGTAATCGGCTAATCGC
CA717853 ---
BQ804592 GGTCTCTGATTGCTTTCTTTTCCCTGTGCGTCTGTAATCG-----C

MDS-1 CAACAGATTCCGGCATGAATAAATGAGAAATAAATTGTTCTGATTTTGGGTGC
CA717853 -----
BQ804592 CAACAGATTCCGGCATGAATAAATGAGAAATAAATTGTTCTGATTTTCACTG-

b

MDS-1
AF350423
AAGCCAGTGTCAACCAAAGTAGACAGCATCCTACCTGCGACCCAATCCAGAACAGATTTT
-----CGCGGCCGCGAGAA-TGGAGGGCAGGATG-----TTCGGA CTGGAG---

MDS-1
AF350423
T TCCCCGAGCTAGAAACCAAAGTAAACCCGACGATCTCG-ATCGTGCGGCGGAGCAG-CG
- ACCCCGCTGATGACGGCGCTGCAGCACCTGCTG-GACGTACCGGACGGCGAGGCCGGCG

MDS-1
AF350423
T GTTCGACCCCTTCGCCGACCTCTGGCTGACCCCTTCGACACCTTCGCTCCATCATCTC
G ACCCGGCAATGCCGCCGGCGAGAAACAGGGCCCGACGCGCGCCTACG--TCCG-CGACG

MDS-1
AF350423
C GCGGATCTCAGGGCGCAACAGCGAGACGGCCGCGTTCGCCAATGCTCGCATGGACTGGA
C G-CGCGCCATGGCGCCACC-----CCGGCCG-----ACGTGA

MDS-1
AF350423
A GGAGACCCCGAGGCGCACGTCTTCAAGGCCGACCTCCCGGCGTGAAGAAGGAGGAGG
A GGAGCTGCCGGCGCTACGCGTTCGTGGTGGACATGCCGGGCTGGGGTCCGGCGACA

MDS-1
AF350423
T CAAGGTGAGGTGGAGGACGGCAACCGTCTCGTCTCAGCGGCGAGCGCACAAAGGAGA
T CAAGGTGAGGTGAGGACGAGCGCGTGTGGTGTATCAGCGGCGAGCGGCGGAGGAGG

MDS-1
AF350423
A GGAGACAAGAACGCAAGTGGCACCGGTGGAGCGCAACAGCGGCAAGTTCGTACGGC
A GAAGG---AGGACCGCAAGTACTCTCGGATGGAGCGCCGCATGGCAAGATGATGCCCA

MDS-1
AF350423
G GTTCGCGCTCCCCGAGGACGCCAAGGTGGAGGAGCTGAAGGCCGGCTGGAGAACGGTG
A GTTCGTGCTCCCCGAGAACGCCGACATGGAGAAGATCTCCGCCGTGTGCCGCGACGGCG

MDS-1
AF350423
T GCTCACCGTCAACCTGCCAAGGCCAGGTCAAGAAAGCCGAGGTGAAGGCCATCCAGA
T GCTCACCGTCTCCTT-----GGAGAAGCTGCCG-----CCGCCCGAA

MDS-1
AF350423
T CTCGGGCTGAGTGGACGGCTCTCGGGTATGATCATCAGGGATGGAGCCAGTTTGGTGT
A CCAAG-----AAGCCCAAGACCATCCAGGT-----CCAGGTCCCTG

MDS-1
AF350423
A TGTGTGTGGAGTTCTTGGCAGTCTGATGAGACATCTCTGTATTGTGTTCTTTCCCCA
A GATGTAT-CGTG----TGCGCTCG-----AATCGAAGCAGAGAGTAGC-----

MDS-1
AF350423
G TGTCTTCTGTACTTGTGTAATCGGCTAATCGCCAACAG-ATTCGGCGATGAATAAATGA
G AGTTTCTGTG-ATGAGCGATGGTCTGTCTGTCTTTTGGTTTCG-----TTAATGG

MDS-1
AF350423
G AAATAAATTTGTTCTGATTTTGGGTGC-
-----AAATGG-----CGATGCGGCCGCG

Figure S7

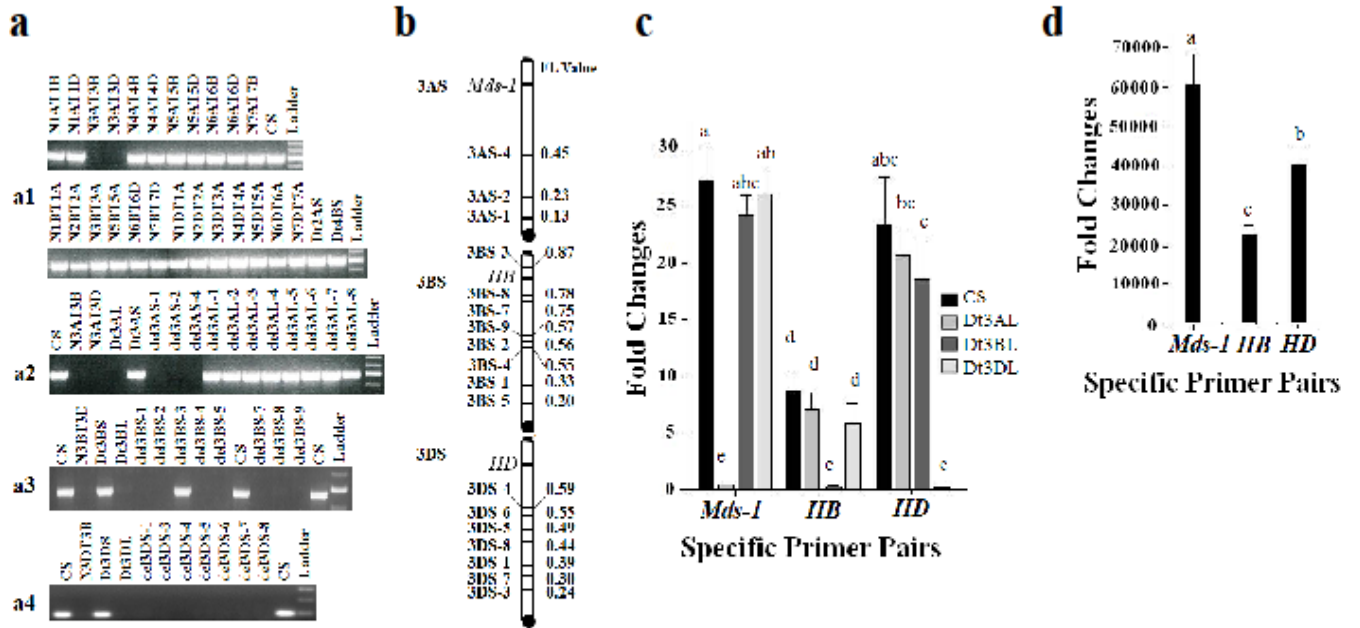


Figure S8

a

```
MDS-1 MSIVRRSSVFDPFADLWADPFDTFRSIPAISSGNS---ETAAFANARMDWKETPEAHVF
HB MSIVRRSNVFDPFADLWADPFDTFRSIVPAISSGNN---ETAAFANARMDWKETPEAHVF
HD -----PFADLWADPFDTFRSIVPAISSGSS---ETAAFANARVDWKETPEAHVF
Contig17 MSIVRRSSVFDPFADLWADPFDTFRSIPAISSGNS---ETAAFANARMDWKETPEAHVF
Contig4a MSIVRRSNVFDPLADLWADPFDTFRSIVPAISSGNS---ETAAFANARMDWKETPEAHVF
Contig1 MSIVRRSNVFDPFADLWADPFDTFRSIVPAISSGNS---ETAAFANAHDWKETPEAHVF
Contig14r MSIVRRSNVFNPFADLWADPFDTFRSIVPAISSGNS---ETAAFANARVDWKESSPEAHVF
Contig15r MSIVRRSNVFDPFADLWADPFDTFRSIVPAISSGNS---ETAAFANARVDWKETPEAHVF
Contig11 MSIVRRSNVFDPFADLWADPFDTFRSIVPAISSGSS---ETAAFANARVDWKETPEAHVF
Contig7 MSIVRRSNVFDPFADLWADPFDTFRSIVPAISSGSSS---ETAAFASARVDWKETPEAHVF
Contig1ar MSIVRRSNVFDPFADLWADPFDTFRSIVPAISSGSS---ETAAFANARVDWKETPEAHVF
Contig10 MSMVRRSNVFDPFADLWADPFDTFRSIVPAMISGNNNDETAAFANARVDWKETPEAHVF
*****:*. * . *****.*:****:*****
```

```
MDS-1 KADLPGVKKEEVKVEVEDGNVLVVS GERTKEKEDKNDKWHRVERSSGKFVRRFRLPEDAK
HB KADLPGVKKEEVKVEVEDGNVLVVS GERTKEKEDRNDKWHRVERSSGKFVRRFRLPEDAK
HD KADLPGVKKEEVKVEVEDGNVLVVS GERTKEEEDKNDKWHRVERSSGKFVRRFRLPEDAK
Contig17 KADLPGVKKEEVKVEVEDGNVLVVS GERTKEKEDKNDKWHRVERSSGKFVRRFRLPEDAK
Contig4a KADLPGVKKEEVKVEVEDGNVLVVS GERTKEKEDKNDKWHRVERSSGKFVRRFRLPEDAK
Contig1 KADLPGVKKEEVKVEVEDGNVLVVS GERTKEKEDKNDKWHRVERSSGKFVRRHFRLPEDAK
Contig14r KADLPGVKKEEVKVELEDGNVLVVS GERTKEKEDKNDKWHRVERSSGKFVRRFRLPEDAK
Contig15r KADLPGVKKEEVKVEVEDGNVLVVS GERTKEKEDKNDKWHRVERSSGKFVRRFRLPEDAK
Contig11 KVDLPGVKKEEVKVEVEDGNVLVVS GER SREKEDKNDKWHRVERSSGKFVRRFRLPEDAK
Contig7 KADLPGVKKEEVKVEVEDGNVLVVS GER SREKEDKNDKWHRVERSSGKFARRFRLPEDAK
Contig1ar KADLPGVKKEEVKVEVEDGNVLVVS GERTKEKEDKNDKWHRVERSSGKFVRRFRLPEDAK
Contig10 KADLPGVKKEEVKVEVEDGNVLVVS GERTKEEEDKNDKWHRVERSSGKFVRRFRLPEDAK
*.*****:*****:*. ** :*****:***** . :*** **
```

```
MDS-1 VEEVKAGLENGVLTVTVPKAVKKEVKAIQISG
HB VEEVKAGLENGVLTVTVPKAEVKKPEVKAIEISG
HD VGEVKAGLENGVLTVTVPKAEVKKPEVKAIEISV
Contig17 VEEVKAGLENGVLTVTVPKAVKKEVKAIQISG
Contig4a VEEVKAGLENGVLTVTVPKAEVKKPVKAIEISG
Contig1 VEEVKAGLENGVLTVTVPKAEVKKPVKAIQISG
Contig14r VEEVKAGLENGVLTVTVPKAEVKKPEVKAIEISG
Contig15r VGEVKAGLENGVLTVTVPKAEVKKPEVKAIEISV
Contig11 VEEVKAGLENGVLTVTVPKAEVKKPEVKAIEISG
Contig7 VEEVKAGLENGVLTVTVPKAHVKKPEVKAIEISG
Contig1ar VEEVKAGLENGVLTVTVPKAEVKKPEVKAIEISG
Contig10 VEEVKAGLENGVLTVTVPKAEVKKPEVKAIEITG
* *****:*****:*. ** :*****:*****
```

b

```
Mds-1 -----AAGCCAGTGTCAACCAAAAAG-TAG----ACAGCATCCTA-----CCTGCG 40
Hb AACTCAAATCAGCGTCAAGCAAAAAG-CAG----ACAGCATCCT-----AC 40
Hd -----
Contig17 -----AAGCCAGTGTCAACCAAAAAG-TAG----ACAGCATCCTA-----CCTGCG 40
Contig11 -----AAGCCAGTGTCAAGCAACAG-CAA----ACAACATCCTTTT-----CCTAC- 41
Contig14r -----AAATCAGTGTCAAGCAAAAAG-TAG----ACGACATCCTATT-----CCTAC- 41
Contig1 -----AAGCCAGTGTCAAGCAAAAAG-CAG----ACAAAATCCCATATTCCCATTCTAC- 49
Contig15r -----AAGCCAGTGTCAAGCAAAAAG-CAG----ACAGCAAGCCCGTT-----CCTAC- 42
```

Contig4a -----ATCCCAAACCTCAAGCCAGCGTCAA-----GCAAAATCCTACG----- 37
Contig10 -----AAGCCAGTGTCAAGCAAAAG-CAAGCAAACAGCATCCTGTT-----CCTAC 45
Contig1ar -----ACAGCATCCTAGT-----CCTAC 18
Contig7 -----AAGCCAGTGTCAACCAAAAG-CAG----ACAGCAACCCCAA-----CCTAC 41

Mds-1 ACCCAATCCAGAACAG----ATTTTTTCCC-CG--AGCTAG--AAACCAAAG-TAACACC 90
Hb CTGCGGTCCGATACAG--ATTTTCCCCCGAGC---TACAACCCAAAGCAA-----CATC 89
Hd -----
Contig17 ACCCAATCCAGAACAG----ATTTTTTCCC-CG--AGCTAG--AAACCAAAG-TAACACC 90
Contig11 CTACGATCCGATACG----AATTTCCGAGCG--CACAAGCCAAACCAAAGCAAACACT 95
Contig14r CTACGATCC---ACAG----AATTTCCGAGCG--CACAAGCCGAAGCAAT--CAGCACC 90
Contig1 CTACGATCCAATACAT----AATCTTTCCCGAGC-TACAAGCCAAAGCAAG--CAATACC 102
Contig15r CTACGATCAAATACAT----AATTTCCCCCAAGCTACAAGCCAAAGCAAG--CAACACC 96
Contig4a -TACGATCCGATACAG----AAT-TTCCCCCCG--AGCTAC--AAACCGAAG-CAACACC 86
Contig10 CTACGATCCGATACAGAATTTTTCTTCCGAGCG--CACAAGCCGAAGCAAT--CAGCACC 101
Contig1ar CTACGATC---TACAG----AATTTCCGAGCG--CACAAGCCGAAGCAAT--CAACACC 67
Contig7 CTACGATCCAATACAG----AATTTCCTAGCG--CAAGAGTCAAAGCAAG--CAACACC 93

▼
Mds-1 GACGATGTCGATCGTGCGGGGAGCAGCGTGTTCGACCCCTTCGCGGACCTCTGGGCTGA 150
Hb GACAATGTCGATCGTGAGGCGGAGCAATGTGTTCGACCCCTTCGCGGACCTCTGGGCGGA 149
Hd -----CCCTTCGCGGACCTCTGGGCGGA 23
Contig17 GACGATGTCGATCGTGCGGGGAGCAGCGTGTTCGACCCCTTCGCGGACCTCTGGGCTGA 150
Contig11 GACGATGTCGATCGTGAGGCGGAGCAACGTGTTCGACCCCTTCGCGGACCTCTGGGCGGA 155
Contig14r GACCAATGTCGATCGTGAGGCGGAGCAACGTGTTCGACCCCTTCGCGGACCTCTGGGCGGA 150
Contig1 GGCAATGTCGATCGTGAGGCGGAGTAACGTGTTCGATCCCTTCGCGGACCTCTGGGCGGA 162
Contig15r GACAATGTCGATCGTGAGGCGGAGCAACGTGTTCGACCCCTTCGCGGACCTCTGGGCGGA 156
Contig4a GAGAATGTCGATCGTGAGGCGGAGCAACGTGTTCGACCCCTTCGCGGACCTCTGGGCGGA 146
Contig10 GACAATGTCGATGGTGAGGCGGAGCAACGTCTTCGATCCCTTCGCGGACCTCTGGGCGGA 161
Contig1ar GACGATGTCGATCGTGAGGCGGAGCAACGTGTTCGACCCCTTCGCGGACCTCTGGGCGGA 127
Contig7 GGCAATGTCGATCGTGCGGGGAGCAACGTGTTCGACCCCTTCGCGGACCTCTGGGCGGA 153
***** ** ***** ***** **

Mds-1 CCCTTTCGACACCTTCCGCTCCATCATCCCGGCGAT---CTCAGGCGGCAACAGC----- 202
Hb CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAAT----- 201
Hd CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAGC----- 75
Contig17 CCCTTTCGACACCTTCCGCTCCATCATCCCGGCGAT---CTCAGGCGGCAACAGC----- 202
Contig11 CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAGC----- 207
Contig14r CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAGC----- 202
Contig1 CCCTTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAGC----- 214
Contig15r CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAGC----- 208
Contig4a CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAGC----- 198
Contig10 CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGATGATCTCAGGCAACAACAACAACGA 221
Contig1ar CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCG---GCAGCAG--- 178
Contig7 CCCCTTCGACACCTTCCGCTCCATCGTCCCGGCGAT---CTCAGGCGGCAACAGC--- 207
*** ***** ***** ***** ***** **

Mds-1 -GAGACGGCCGCGTTTCGCAATGCTCGCATGGACTGGAAGGAGACCCCGAGGCGCACGT 261
Hb -GAGACAGCTGCGTTTCGCAACGCCGGATGGACTGGAAGGAGACGCCTGAGGCACACGT 260
Hd -GAGACGGCCGCGTTTCGCAATGCTCGCATGGACTGGAAGGAGACGCCTGAGGCGCACGT 134
Contig17 -GAGACGGCCGCGTTTCGCAATGCTCGCATGGACTGGAAGGAGACCCCGAGGCGCACGT 261
Contig11 -GAGACGGCCGCGTTTCGCAACGCCGGTGTGGACTGGAAGGAGACCCCGAGGCGCACGT 266
Contig14r -GAGACGGCGGCATTTCGCAACGCCGGTGTGGACTGGAAGGAGACCCCGAGGCGCACGT 261
Contig1 -GAGACGGCCGCGTTTCGCAACGCCCGCATGGACTGGAAGGAGACCCCGAGGCGCACGT 273
Contig15r -GAGACGGCCGCGTTTCGCAACGCCCGTGTGGACTGGAAGGAGACTCCCGAGGCGCACGT 267
Contig4a -GAGACGGCCGCGTTTCGCAACGCCCGCATGGACTGGAAGGAGACCCCGAGGCGCACGT 257
Contig10 CGAGACAGCTGCGTTTCGCAACGCCCGTGTGGACTGGAAGGAGACCCCTGAGGCGCACGT 281
Contig1ar CGAGACGGCCGCGTTTCGCAACGCCCGTGTGGACTGGAAGGAGACCCCGAGGCGCACGT 238
Contig7 CGAGACGGCCGCGTTTCGCAATGCTCGCATGGACTGGAAGGAGACCCCGAGGCGCACGT 267
***** ** ** ***** * ** * ***** ***** ***** *****

Mds-1 CTCAAGGCCGACCTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 321
Hb CTCAAGGCCGACCTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 320
Hd CTCAAGGCCGACCTCCCAGGCGTGAAGAAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 194
Contig17 CTCAAGGCCGACCTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 321
Contig11 CTCAAGGTTCGACCTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 326
Contig14r CTCAAGGCAGATCTTCCAGGGGTGAAAAAGGAGGAGGTCAAGGTGGAAC TGGAGGACGG 321
Contig1 CTCAAGGCCGACCTCCCAGGGGTGAAGAAGGAGGAGGTCAAGGTGGAAGTGGAGGACGG 333
Contig15r CTCAAGGCCGACCTCCCAGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 327
Contig4a CTCAAGGCCGACCTCCCTGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 317
Contig10 CTCAAGGCCGACCTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 341
Contig1ar CTCAAGGCCGACCTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 298
Contig7 CTCAAGGCCGACCTCCCCGGCGTGAAGAAGGAGGAGGTCAAGGTGGAGGTGGAGGACGG 327
***** .** ** ** * *****.*.*****.*****.*****

Mds-1 CAACGTGCTCGTCGTCAGCGGCGAGCGCACAAAGGAGAAGGAGGACAAGAACGACAAGTG 381
Hb CAACGTGCTCGTCGTCAGTGGCGAGCGCACAAAGGAGAAGGAGGACAGGAACGACAAGTG 380
Hd CAACGTGCTCGTCGTCAGCGGCGAGCGCACAAAGGAGGAGGAGGACAAGAACGACAAGTG 254
Contig17 CAACGTGCTCGTCGTCAGCGGCGAGCGCACAAAGGAGAAGGAGGACAAGAACGACAAGTG 381
Contig11 CAACGTGCTCGTCGTCAGCGGCGAGCGGAGCAGGGAGAAGGAGGACAAGAACGACAAGTG 386
Contig14r CAACGTGCTCGTCGTCAGCGGCGAGCGCACAAAAGAGAAGGAGGACAAGAACGACAAGTG 381
Contig1 CAACGTGCTCGTCGTCAGCGGCGAGCGCACAAAAGAGAAGGAGGACAAGAACGACAAGTG 393
Contig15r CAACGTGCTCGTCGTCAGTGGCGAGCGCACAAAGGAGAAGGAGGACAAGAACGACAAGTG 387
Contig4a CAACGTGCTCGTCGTCAGCGGCGAGCGCACCAAGGAGAAGGAGGACAAGAACGACAAGTG 377
Contig10 CAACGTGCTCGTCGTCAGCGGCGAGCGCACGAAGGAGGAGGAGGACAAGAACGACAAGTG 401
Contig1ar CAACGTGCTCGTCGTCAGCGGCGAGCGCACAAAGGAGAAGGAGGACAAGAACGACAAGTG 358
Contig7 CAACGTGCTCGTCGTCAGCGGCGAGCGGAGCAGGGAGAAGGAGGACAAGAACGACAAGTG 387
*****.***** * * .***.*****.*****

Mds-1 GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCAGGCGGTTCCGCCTCCCCGAGGACGC 441
Hb GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCGCGGCGTTCCGCCTCCCCGAGGACGC 440
Hd GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCAGGCGGTTCCGCCTGCCCGAGGACGC 314
Contig17 GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCAGGCGGTTCCGCCTCCCCGAGGACGC 441
Contig11 GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCGCGGCGTTTCCGGCTGCCCGAGGACGC 446
Contig14r GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCGCGGCGTTTCCGGCTGCCCGAGGACGC 441
Contig1 GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCAGGCAC TCCGCCTGCCCGAGGACGC 453
Contig15r GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCAGGCGGTTCCGGCTGCCCGAGGACGC 447
Contig4a GCACCGTGTGGAGCGCAGCAGCGGCAAGTTCGTCAGGCGTTCCGCCTGCCCGAGGACGC 437
Contig10 GCACCGTGTGGAGCGTAGCAGCGGCAAGTTCGTCGAGGCGGTTCCGCCTGCCCGAGGACGC 461
Contig1ar GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGTCAGGCGGTTCCGCCTCCCCGAGGACGC 418
Contig7 GCACCGCGTGGAGCGCAGCAGCGGCAAGTTCGCGAGGCGTTCCGGCTGCCCGAGGACGC 447
***** ***** ***** .***. ***** ** * *****

Mds-1 CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGTGTGCTCACCGTCACCGTGCCCAA 501
Hb CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGCGTGCTCACCGTCACCGTGCCCAA 500
Hd CAAGGTGGGGGAGGTGAAGGCCGGGCTGGAGAACGGCGTGCTCACCGTCACCGTGCCCAA 374
Contig17 CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGTGTGCTCACCGTCACCGTGCCCAA 501
Contig11 CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGCGTGCTCACCGTCACCGTGCCCAA 506
Contig14r CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGCGTGCTCACCGTCACCGTGCCCAA 501
Contig1 CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGTGTGCTCACCGTCACCGTGCCCAA 513
Contig15r CAAGGTGGGGGAGGTGAAGGCCGGGCTGGAGAACGGCGTGCTCACCGTCACCGTGCCCAA 507
Contig4a CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGTGTGCTCACCGTCACCGTGCCCAA 497
Contig10 CAAGGTGGAGGAGGTGAAGGCCGGTCTGGAGAACGGTGTGCTCACTGTACCGTGCCCAA 521
Contig1ar CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGCGTGCTCACCGTCACCGTGCCCAA 478
Contig7 CAAGGTGGAGGAGGTGAAGGCCGGGCTGGAGAACGGCGTGCTCACCGTCACCGTGCCCAA 507
*****.***** ***** ***** ***** ***** ***** .**

Mds-1 GGCCGAGGTCAAGAAGCCCGAGGTGAAGGCCATCCAGATCTCCGGCTGAGTGGACGCGTC 561
Hb GGCCGAGGTCAAGAAGCCCGAGGTGAAGGCCATCCAGATCTCCGGCTGAG----- 550
Hd GGCCGAGGTCAAGAAGCCCGAGGTGAAGGCCATCCAGATCTCCGGCTGAGTGGATGATGCGTC 434
Contig17 GGCCGAGGTCAAGAAGCCCGAGGTGAAGGCCATCCAGATCTCCGGCTGAGTGGACGCGTC 561
Contig11 GGCCGAGGTCAAGAAGCCCGAGGTGAAGGCCATCCAGATCTCCGGCTGAGTGCATGCGTT 566
Contig14r GGCCGAGGTCAAGAAGCCCGAGGTGAAGGCCATCCAGATCTCCGGCTGAGTGCATGCGTT 561

Contig1 GGCAGAGGTCAAGAAGCCCCAGGTGAAGGCCATCCAGATCTCCGGCTGAGTGGACGCGTC 573
 Contig15r GGCCGAGGTCAAGAAGCCCCAGGTGAAGGCCATCCAGATCTCCGTCTGAGTGAATGCGTC 567
 Contig4a GGCAGAGGTCAAGAAGCCCCAGGTGAAGGCCATCGAAATCTCCGGCTGAGTGGACGCGTA 557
 Contig10 GGCCGAGGTGAAGAAGCCTGAGGTGAAGGCCATCGAGATCACCGGCTGAG----- 571
 Contig1ar GGCCGAGGTCAAGAAGCCCCAGGTGAAGGCCATCCAGATCTCCGGCTGAGTGGACGCGTC 538
 Contig7 GGCCACGTCAAGAAGCCCCAGGTGAAGGCCATCCAGATATCCGGCTGAG----- 557
 . * ** ** ********** *.*.* ** ** **

Mds-1 TCGGCG-----TATGATCATCAGGGATGGAGCCAGTTTGGTTGATGTGTGTGCGAG--- 612
Hb -----TATTCGGGTCAGAGCCTGTCTGGTTCATGTGTGTGCGAGCTC 592
Hd TCTGCG-----AATGATCATCGGGTTGGAGCCAGTTTCGTTATGCAACCTGAAAG--- 485
 Contig17 TCGGCG-----TATGATCATCAGGGATGGAGCCAGTTTGGTTAATGTGTGTGCGAG--- 612
 Contig11 TGGACG-----TATGCTCATTCAGGTTGGAGCCAATTTGGTTGATGTGTGTGCGAG--- 617
 Contig14r TGGACG-----TATGCTCATTCAGGTTGGAGCCAATTTGGTTGATGTGTGTGCGAG--- 612
 Contig1 TCGGCG-----TATGATCATCAGGGATGGAGCCAGTTTGGTTGATGTGTGTGCGAGGTC 627
 Contig15r TCTGCG-----AATGATCATCGGGTTGGAGCCAGTTTCGTTATGCAACCTGAAAG--- 618
 Contig4a TCGTCGACGTATGATGATCTTTTGGGCCGGAGCCAGTTTGGTTGATGTGTGTAAGAGTTC 617
 Contig10 -----CATTGG-----AGCCAGTTTAGTTGATGTGTGTTTGGCAGCTC 607
 Contig1ar TCGGCG-----TATGATCATCAGGGATGGAGCCAGTTTGGTTGATGTGTGTGCGAGGTC 592
 Contig7 -----CAT-----TGGAGCCAGTTTGATTGACGTGTGTGCGAGGTC 593
 : * *****:.* * .* : :. *...**

Mds-1 -----TTCTTG-----CGAGTCT- 625
Hb TGCCACCT-----AAAAGTATGCTATAGTTGTGCTCTTGTCTTGTCTGATCGAGTGT- 643
Hd -----TATCTTTGTTCTTGC-----CGAGTCT- 507
 Contig17 -----TTCTGC-----CGAGTCT- 625
 Contig11 -----TTCTTG-----CGAGTCT- 630
 Contig14r -----TTCTTG-----CGAGTCT- 625
 Contig1 TGCAACCTG-----AAAGTATGCTACAGTT-GTGTCTTGTCTTGG-----CGAGTCT- 674
 Contig15r -----TATCTTTGTTCTTGC-----CGAGTCT- 640
 Contig4a TGCAACCTA-----AAATTATGCTATAGTTGTCTTGTCTTGTCTTGT-----CGAGTCTT 666
 Contig10 TGCAACCTCTGCAACCTAAAGTATGGTACGGTTGTGTCTTGTCTTGC-----CGAGTCT- 662
 Contig1ar TGCAACCT-----GAAAGTATGCTACAGTT-----CTTGCTGATCGAGTCT- 633
 Contig7 TGCAACCT-----GAAAGTATGCTACAGTT-----CTTGCTGATCGAGTCT- 634
 ** ***** *

Figure S9

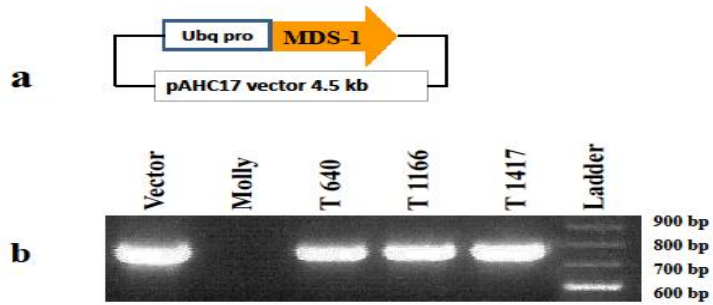


Figure S10N

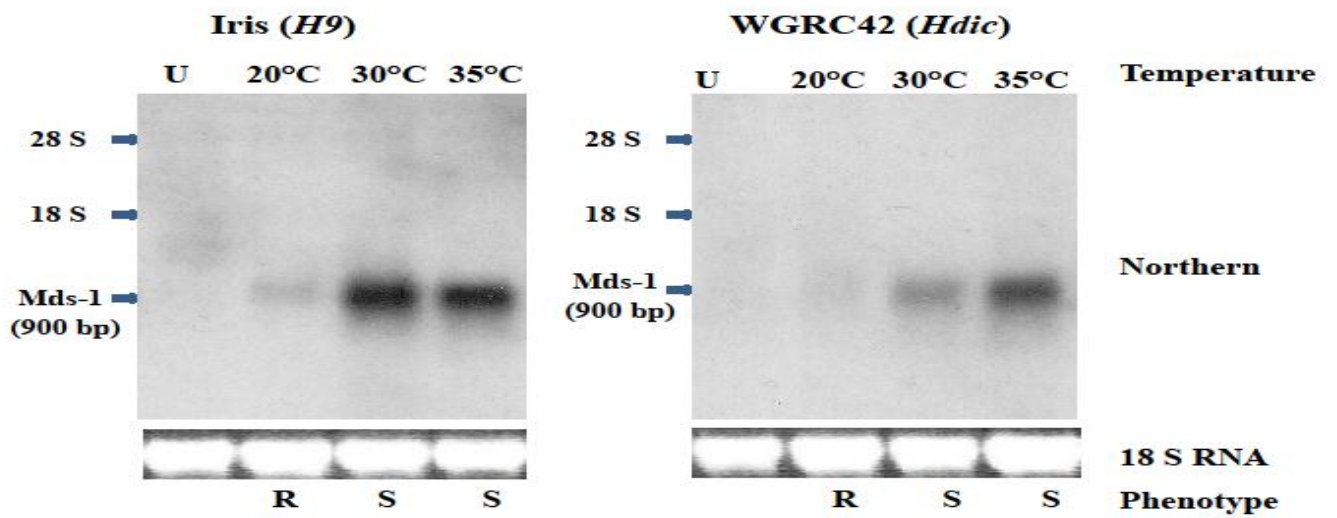


Figure S11

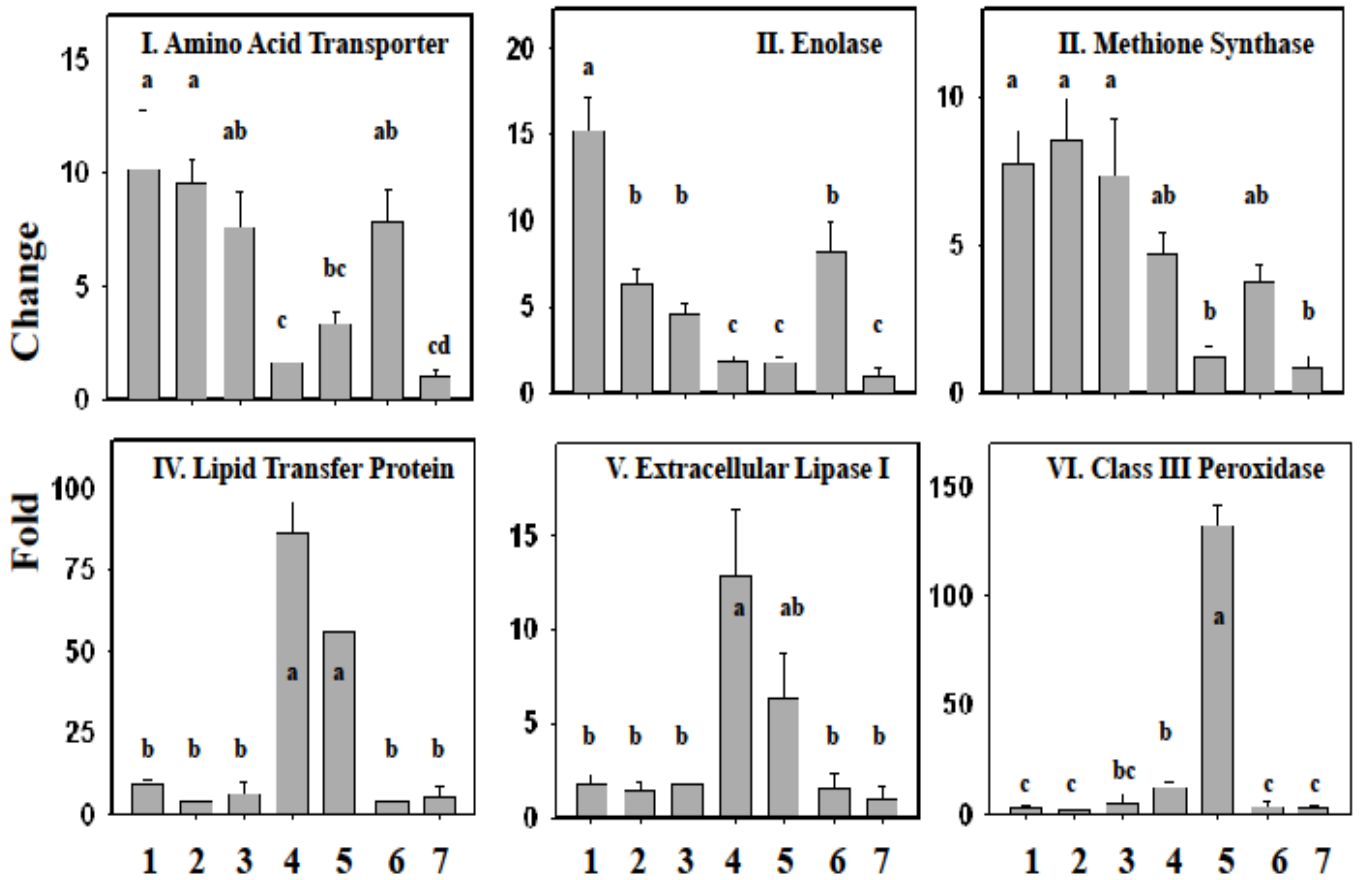


Figure S12

