Instituto de Microbiologia
Paulo de Goés UFRJ

GLOBAL ALTERATION OF MICRORNAS AND TRANSPOSON-DERIVED SMALL RNAS IN COTTON (*GOSSYPIUM HIRSUTUM*) DURING COTTON LEAFROLL DWARF POLEROVIRUS (CLRDV) INFECTION

Elisson Romanel[1], Tatiane da Franca Silva[1], Régis L. Corrêa[1], Laurent Farinelli[2], Jennifer Hawkins[3], Carlos E. G. Schrago[1] & Maite F. S. Vaslin[1]
[1]-UFRJ. Rio de Janeiro. Brazil,
[2]-Fasteris Co., Plan-les-Ouates. Switzerland,
[3]-West Virginia University. Morgantown. United States

Laboratório de Virologia Molecular Vegetal

Depto. Virologia, IMPPG - Universidade Federal do Rio de Janeiro
Rio de Janeiro, Brasil

Profa. Maïté Vaslin F. Silva (maite@micro.ufrj.br)

Cotton:

Gossypium hirsutum L.
Family: Malvaceae

- **The most widely produced natural fiber in the world;**

Represents about 40 per cent of the world textile market

Cotton Diseases

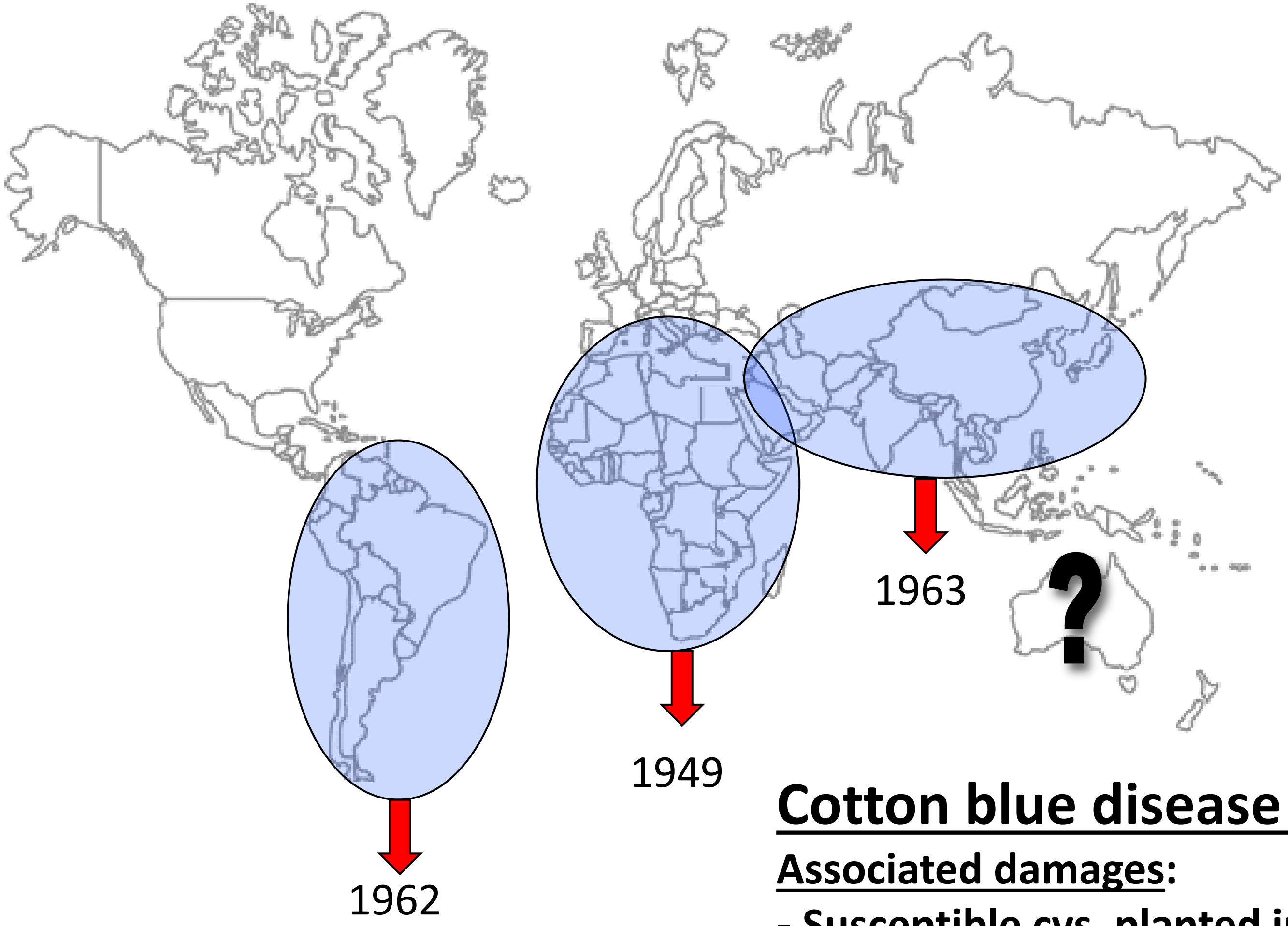
Caused by Viruses:

Leaf crumple/Leaf curl - *Cotton leaf curl virus (CLCuV), Geminivirus*

Anthocyanosis

Blue disease

Blue Disease in the World



Cotton blue disease in Brazil

Associated damages:

- Susceptible cvs. planted in cerrado until end of 90s
- Losses around 1500 kg.ha-1 in MT at GO
- 91% of productivity per plant reduction
    - -Cotton weight (-53,3%) e fiber (-43,0%).
    - -In soft infections: reduction of 36%

Symptoms of CBD are leaf rolling, intense green foliage, vein yellowing and a severe to moderate stunting caused by internodal shortening.

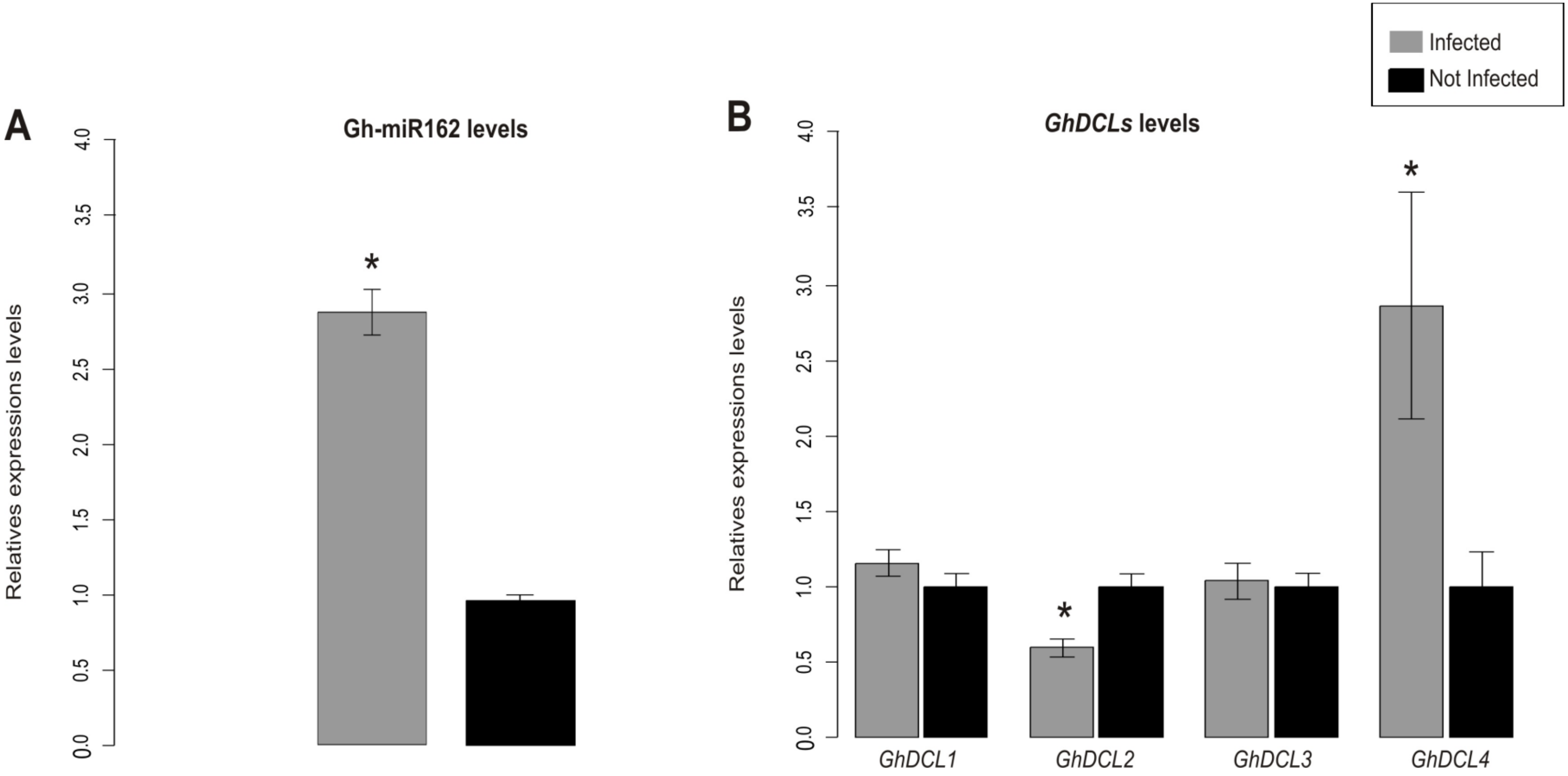
How CLRDV (casual agent of CBD) can led to
the strong developmental alterations observed in infected plants?

Each mechanisms are involved?

miRNA are an important class of gene expression regulation RNA working in several developmental genes downregulation

So, how are miRNAs profiles during virus infection??

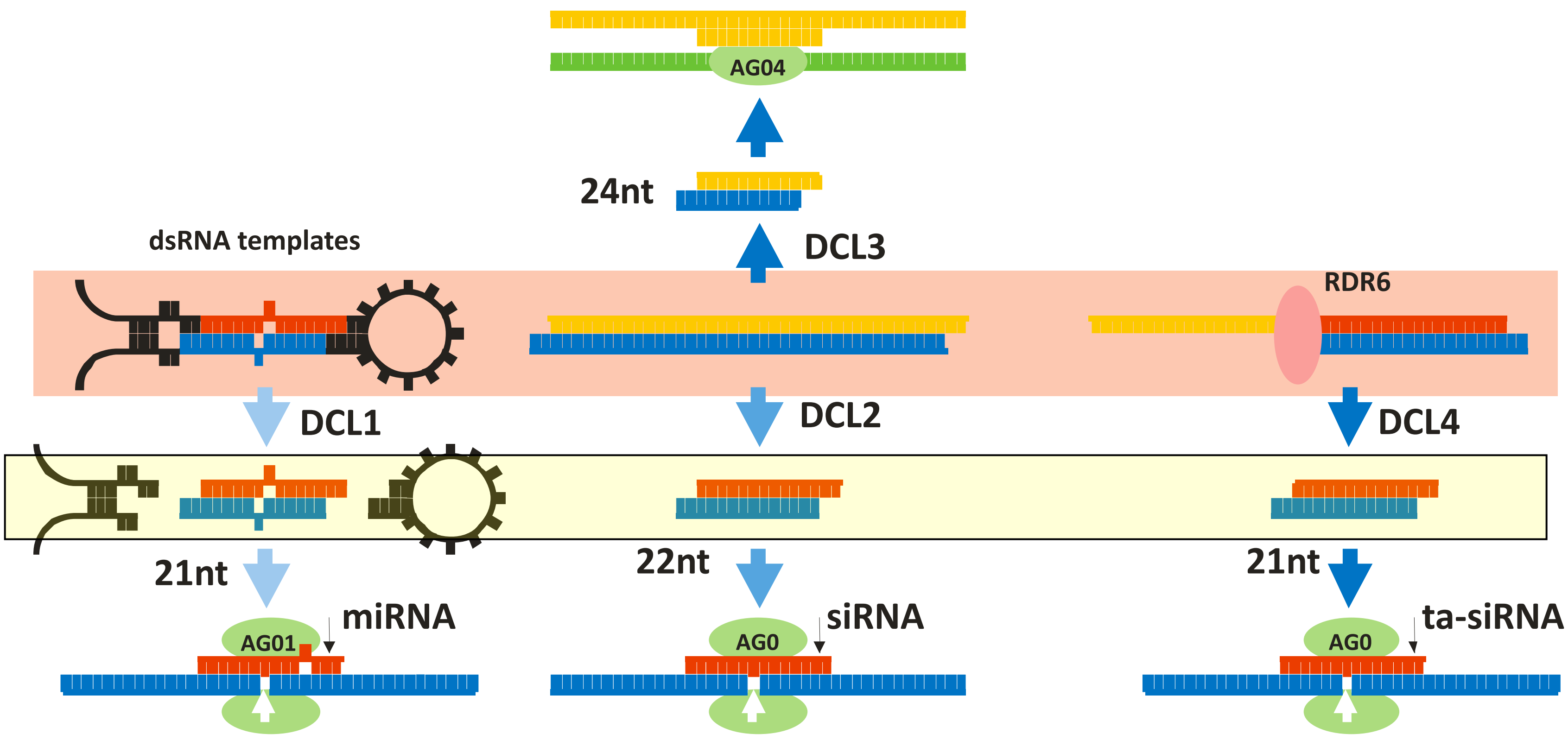
Expression of Cotton DCL ribonucleases during infection



Quantitative RT-PCR analysis. Expression levels of Gh-miR162 (**A**) and cotton Dcl mRNAs (**B**) in infected and uninfected plants.

Silva et al. BMC Molecular Biology 2011, 12:40

The silencing pathways



Heterochromatin formation

AG04

24nt

DCL3

dsRNA templates

RDR6

DCL1

DCL2

DCL4

21nt

22nt

21nt

miRNA

AG01

siRNA

AG0

ta-siRNA

AG0

mRNA cleavage

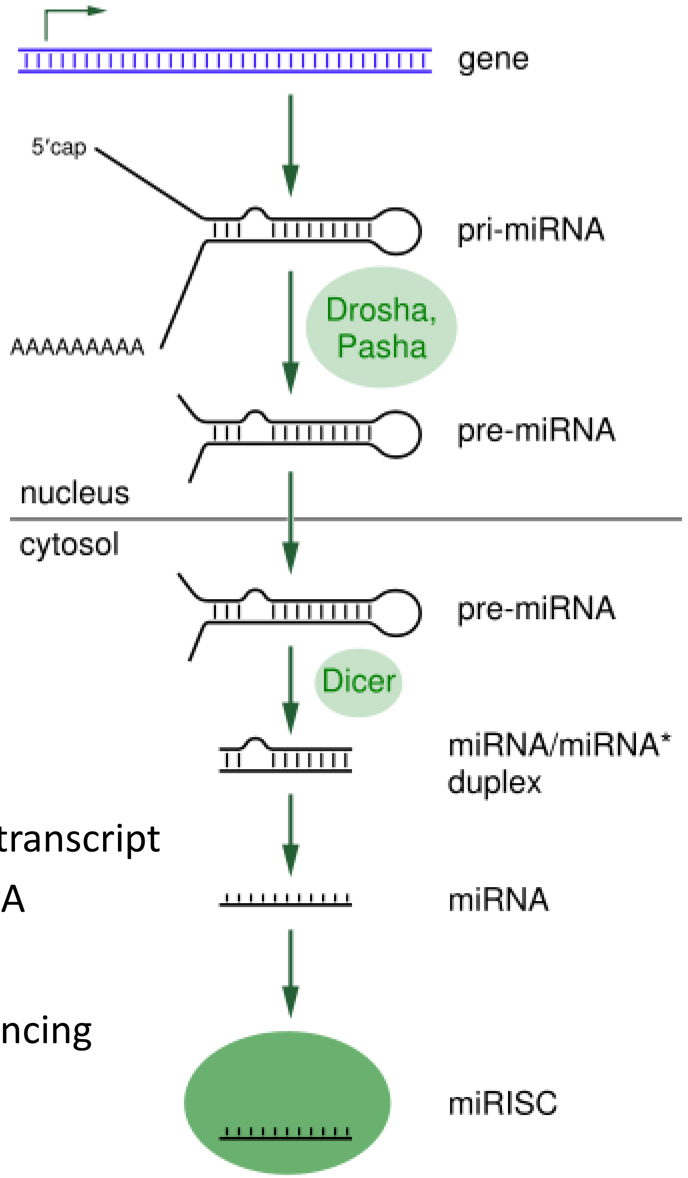
Virus RNA cleavage

mRNA cleavage

Developmental regulation

Viral defence

Overview of microRNA processing

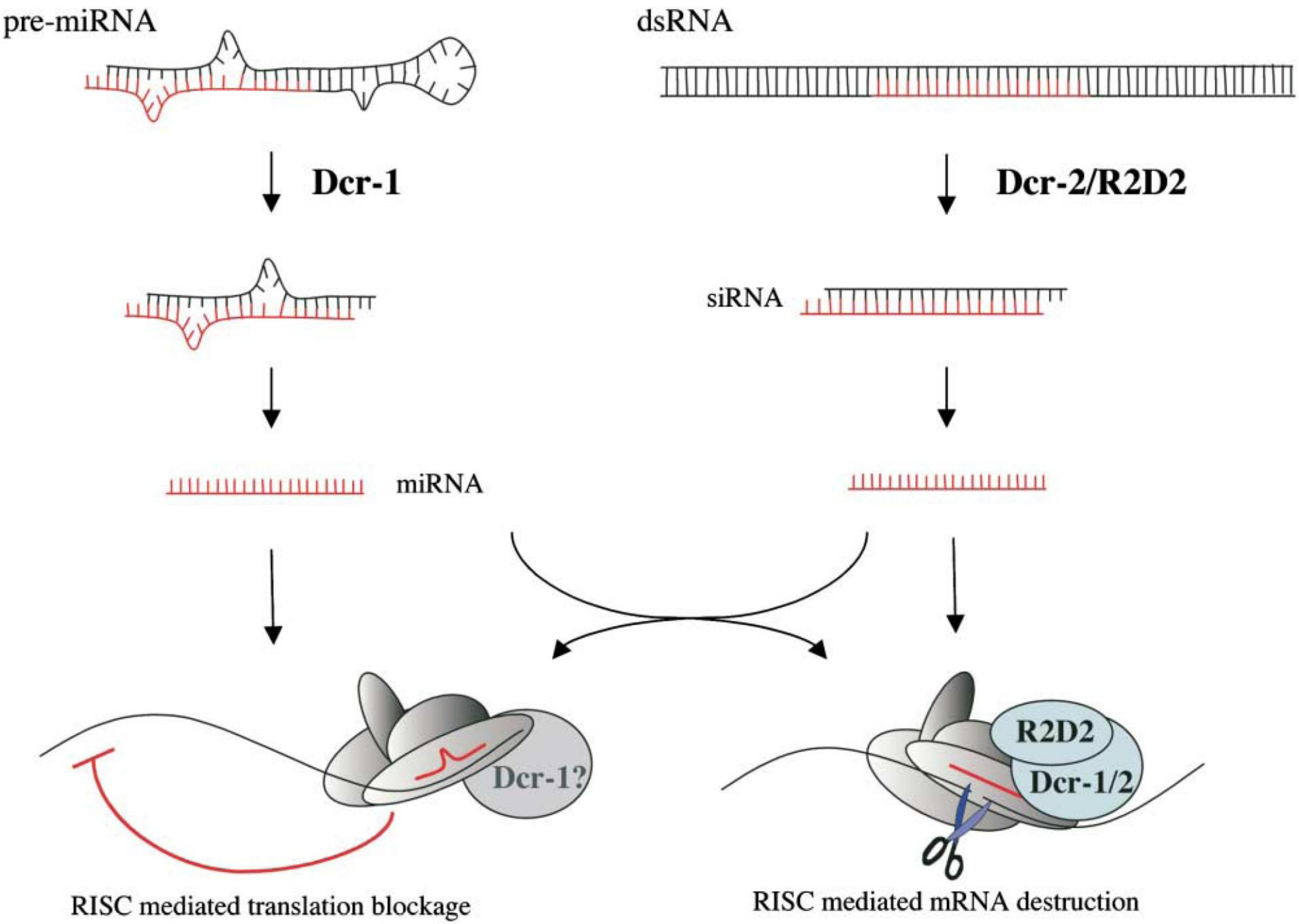


pri-miRNA = primary microRNA transcript

pre-miRNA = precursor microRNA

miRNA* = antisense microRNA

miRISC = microRNA-induced silencing complex


pre-miRNA
dsRNA
Dcr-1
Dcr-2/R2D2
siRNA
miRNA
Dcr-1?
R2D2
Dcr-1/2
RISC mediated translation blockage
RISC mediated mRNA destruction

Characterization of cotton sRNAs

Uninfected small-RNA library

Infected small-RNA library

Total of small RNA sequenced

9,480,917

10,566,377

3`and 5`adaptors removed, < 18-nt and >26-nt

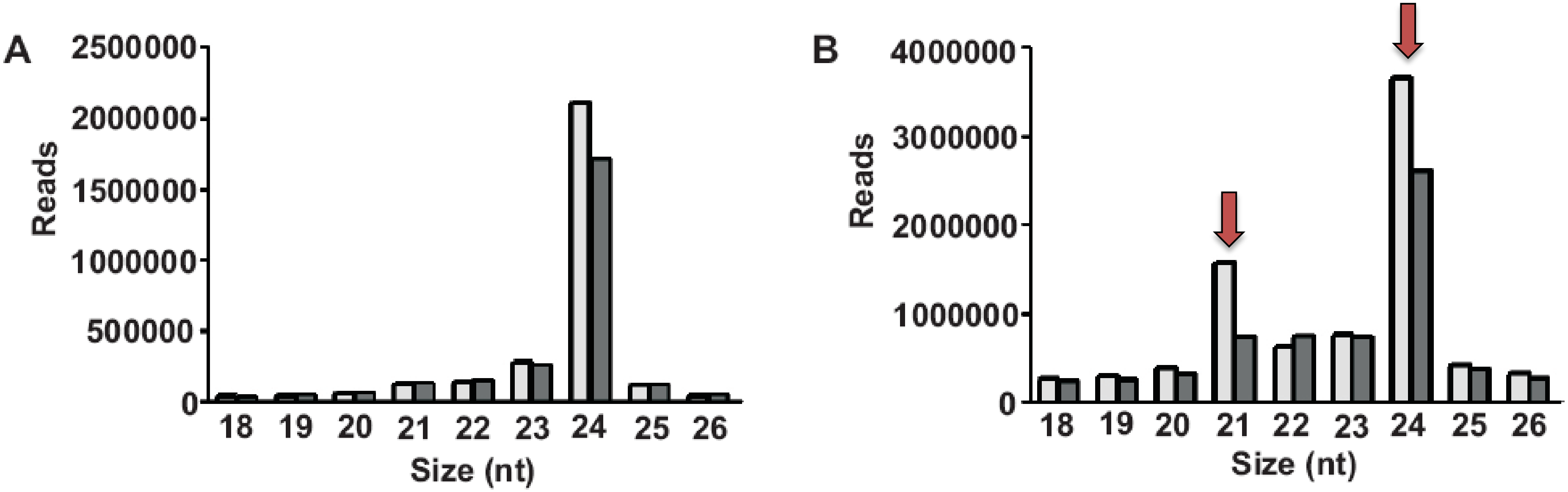
7,852,992

6,569,610

BLAST against the CLRDV genome

Profile of all small RNAs from the infected and the uninfected libraries

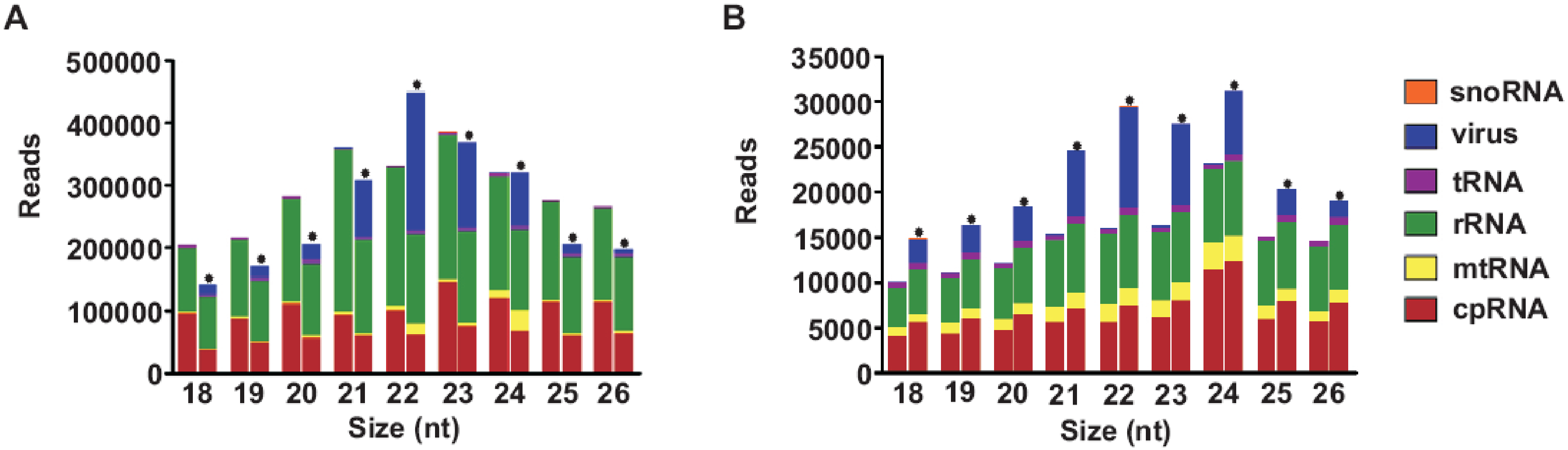
24-nt 21-nt
predominant classes



Size distribution of sequenced *Gossypium hirsutum* small RNAs (sRNAs) in uninfected and infected leaves. Size distribution of unique (only one read) (A) and redundant (B) cotton sRNAs from the uninfected (light gray bars) and infected (dark gray bars) libraries. Histograms represent the number sRNA reads.

Is the profile of cellular non-coding RNA altered by virus infection?

Identification of distinct non-coding cellular and viral small RNA (sRNA) sequences in the *Gossypium hirsutum* uninfected (UIL) and infected (IL) libraries.



redundant cotton sRNAs **unique cotton sRNAs**

the number of redundant cpRNA and rRNA reads was drastically decreased in the IL compared with the UIL

Histograms represent the number of reads within each length class mapping to known cellular and viral RNAs in the uninfected (left) and infected libraries (right) (asterisk)

Supplemental Table S1: Distribution of small RNAs in the uninfected (UIL) and CLRDV-infected (IL) libraries

	UIL		IL	
Library	**Unique**	**Redundant**	**Unique**	**Redundant**
input	2,937,808	8,282,945	2,593,589	6,217,467
cpRNA	53,033	970,863	68,218	529,876
mtRNA	14,307	43,215	14,162	68,002
rRNA	59,972	1,588,517	62,407	1,096,668
tRNA	4,548	33,452	7,130	62,284
snoRNA	282	2,165	205	701
virus	1,064	1,976	51,607	640,325
Other	2,804,544	5,642,646	2,392,627	3,853,932

Note: The numbers represent the number of reads expressed as RP10M, reads per 10 million

deep sequencing the small RNAs

Small RNAs
Uninfected cotton

Small RNAs
Infected cotton

Elimination of cellular non-coding RNAs
rRNA, tRNA, snoRNA, mtRNA e cpRNA

BACs sequences (Cotton DB)

Cotton EST (CGI-TIGR)

G. raimondii WGS (500 Mb)

miRNA
identification(PMRD)

Identification of miRNA
precursors (mfold)

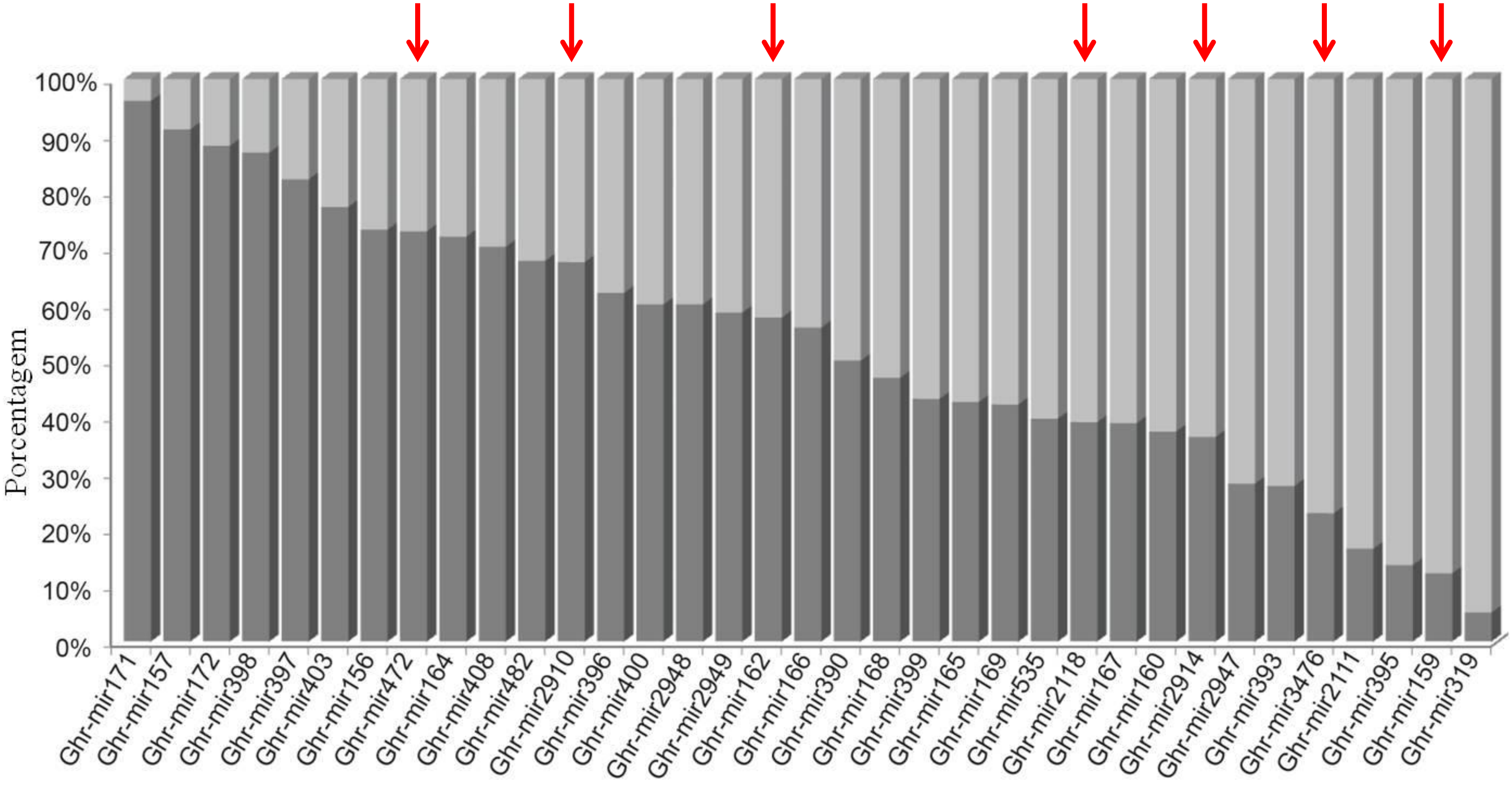
miRNAs tagets

Table 1: **Novel conserved miRNAs identified in the *Gossypium hirsutum* uninfected (UIL) and CLRDV-infected (IL) libraries, their potential targets and their evolutionary relationship with other plants.**

60 miRNA families/ 19 new/novel

Ghr-miRNA family	Sequence 5`to 3'	Size	Reads UIL	Reads IL	Targeted genes or ESTs[a]	Targeted protein[b]	Conserved Ghr[c]	Ath	Pth	Vvi	Osa	Ppt
Ghr-miR158a*[1]	UCCCAAAUGUAGACAAAGCA	20	30.58	37.85	TC238771	Unknown	-	2	-	-	-	1
Ghr-miR161	UGAAAGUGACUACAUCGGGG	20	4.21	-	TC255407 TC238121	Adenine phosphoribosyltransferase; P-glycoprotein	-	2	-	-	-	-
Ghr-miR161	UGAAAGUGACUACAUCGGGGU	21	1.05	0.94	TC234712; TC237985; TC255407	P glycoprotein; Dioxygenase; Adenine phosphoribosyltransferase 1						
Ghr-miR163	UUGAAGAGGACUUGGAACUUCGA	23	1.05	-	TC276513	Unknown	-	1	-	-	1	-
Ghr-miR163	UUGAAGAGGACUUGGAACUUCGAU	24	25.31	1.89	TC276513	Unknown						
Ghr-miR173	UUCGCUUGCAGAGAGAAAUCAC	22	3.16	-	CO111583	glycosyltransferase	-	2	-	-	-	-
Ghr-miR400a	UAUGAGAGUAUUAUAAGUCAC	21	6.32	8.51	TC276035	Pentatricopeptide repeat (PPR) proteins	1	1	-	-	-	-
Ghr-miR781	UUAGAGUUUUCUGGAUACUUA	21	-	0.94	TC237955	GDRP-19 gland development related protein	-	1	-	-	-	-
Ghr-miR822	UGCGGGAAGCAUUUGCACAUG	21	1.05	0.94	CO075409	CDP-diacylglycerol synthase	-	1	-	-	-	-
Ghr-miR824	UAGACCAUUUGUGAGAAGGGA	21	2.10	-	TC233512	1-deoxy-D-xylulose 5-phosphate reductoisomerase	-	1	-	-	-	-
Ghr-miR841	AUUUCUAGUGGGUCGUAUUCA	21	-	0.94	TC278476	RAV-like B3 domain transcription factor	-	1	-	-	-	-
Ghr-miR843	UUUAGGUCGAGCUUCAUUGGA	21	1.05	-	ES808840	Unknown	-	1	-	-	-	-
Ghr-miR844	UGGUAAGAUUGCUUAUAAGC	20	1.05	-	DR461934 ES799013	Alpha-1,4-glucan-protein synthase	-	1	-	-	-	-
Ghr-miR844	UGGUAAGAUUGCUUAUAAGCU	21	-	0.94	DR461934	Alpha-1,4-glucan-protein synthase						
Ghr-miR847	UCACUCCUCUUCUUCUUGAUG	21	-	0.94	TC278151	60S ribosomal protein	-	1	-	-	-	-
Ghr-miR853	UCCCCUCUUUAGCUUGGAGAAG	22	-	0.94	TC234146	Malate dehydrogenase	-	1	-	-	-	-
Ghr-miR863	UUGAGAGCAACAAGACAUAAU	21	1.05	-	DR454972	S-adenosyl-L-methionine-dependent	-	2	-	-	-	-
Ghr-miR896	GUCAAUAUGUCCGAGUGGUUAAGG	24	-	0.94	TC259806	Unknown	-	-	-	-	-	1
Ghr-miR2111a	UCCUUGGGGUGCAGAUUACC	20	2.10	-	TC242916 BG440518	Beta-tubulin; Unknown	-	2	-	1	-	-
Ghr-miR2111b	UAAUCUGCAUCCUGAGGUUU	20	-	4.73	CO099476	F-box family protein						
Ghr-miR2111a	GUCCUUGGGGUGCAGAUUACC	21	95.98	15.14	CA992878	Alphahydrolase						
Ghr-miR2118b	GCCGAUUCCACCCAUGCCUA	20	5.27	0.94	TC270900	Unknown	-	-	-	-	18	-
Ghr-miR2118a	UGCCGAUUCCACCCAUGCCUA	21	76.99	48.26	TC239469 TC270900	Quinone oxidoreductase Unknown						
Ghr-miR2118a	UUGCCGAUUCCACCCAUGCCUA	22	1031.54	1548.30	TC235376 ES807202	NBS-LRR resistance protein						
Ghr-miR2910*[2]	AGUUGGUGGAGGAUUUGUCU	20	1.05	-	TC235132 TC263389	18S ribosomal	-	-	-	-	-	-
Ghr-miR2910*[2]	UAGUUGGUGGAGCGAUUUGUC	21	27.42	56.78	TC235132	18S ribosomal						
Ghr-miR2914a*[3]	UGGUGGUGACGGGUGACGGAG	21	14.76	0.94	TC232414 TC264817	18S ribosomal	-	-	-	-	-	-

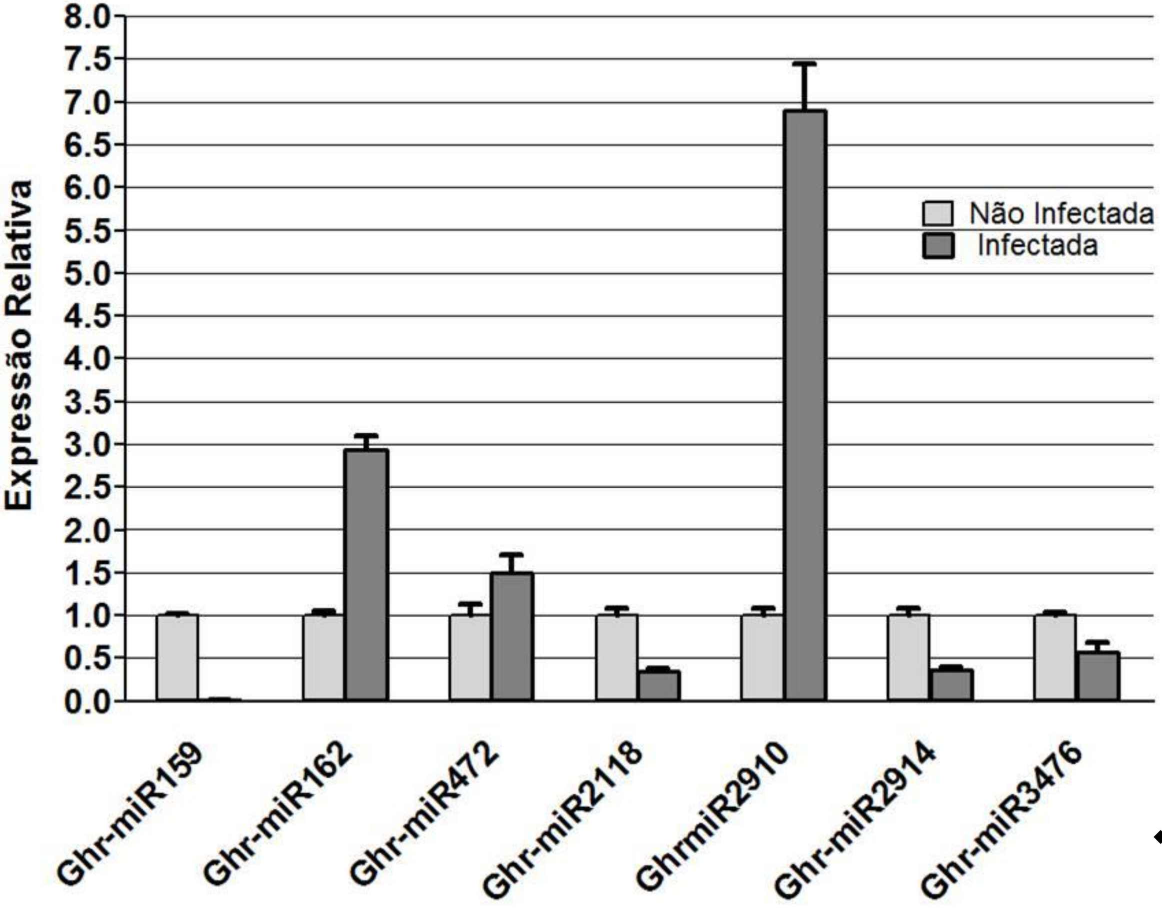
Differential expression of 21-nt Ghr-miRNAs during virus infection



Ghr-miR159 (74,8% of total miRNAs) and Ghr-miR3476 (10%) families are overexpressed in UIL libraries

Differential expression of 21-nt Ghr-miRNAs during virus infection

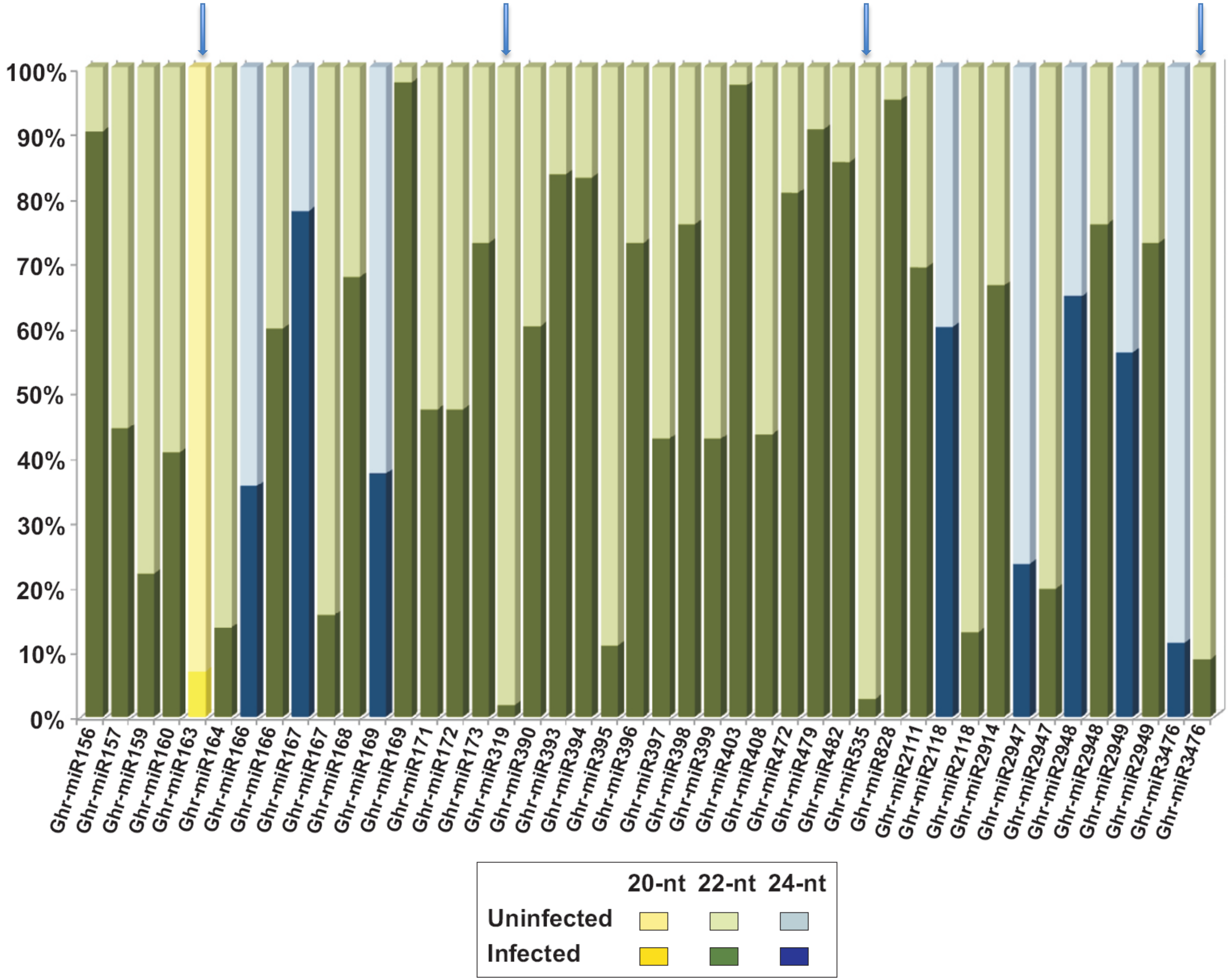
- *deep sequencing* results validation by RT-qPCR



- **miR162**: DCL1 regulation
- **miR159**: leaf development
- **miR472**: NBS-LRR ptns

- **miR2118, 2910, 2914**: Pela primeira vez descritos

❖*miR163 (20nt) and miR319, 525 and 3476 (of 22nt) are drastically reduced*

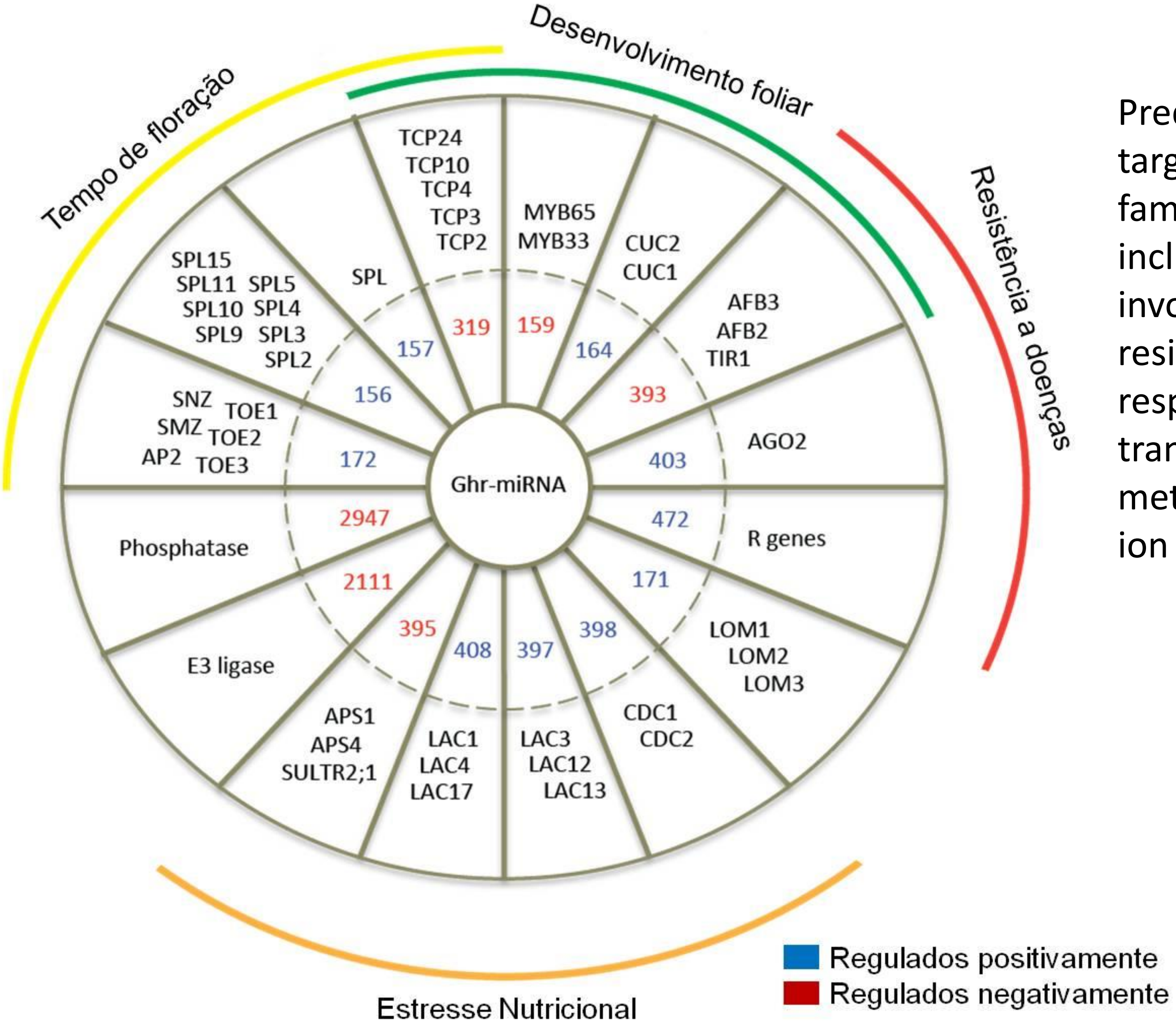


Target prediction of the miRNAs differentially expressed during infection

- Target prediction: psRNATarget

Ghr-miRNAs	Sequences 5`- 3'	*Number of reads*		Gene ID	Putative target functions
		Uninfected	Infected		
156[a]	UUGACAGAAGAGAGUGAGCAC	37,97	110,72	TC239557	Proteína de ligação ao promotor do gene Squamosa [C,D,E]
157b	UUGACAGAAGAUAGAGAGCAC	67,50	690,87	TC239557	Proteína de ligação ao promotor do gene Squamosa [B,C,D,E]
159[a]	UUUGGAUUGAAGGGAGCUCUA	798534,62	109335,57	TC279858	ATP sintase F0 [C]
160[a]	UGCCUGGCUCCCUGUAUGCCA	978,80	893,39	TC246307	Fator de resposta a Auxina (ARF) [B,C,D,E]
162[a]	UCGAUAAACCUCUGCAUCCAG	374,43	500,64	TC237985; TC267453	Proteína LETM1; Proteína de transporte de Auxina
164[a]	UGGAGAAGCAGGGCACGUGCA	50,62	125,87	CO109521	Fator transcricional NAC [C,D,E]
165[a]	UCGGACCAGGCUUCAUCCCCC	23,20	16,08	TC267202	Fator transcricional HD-ZIP [C,D,E]
166h	UCGGACCAGGCUUCAUUCCCG	10751,07	18055,38	ES808067	Fator transcricional HD-ZIP [B,C,D,E]
167c	UGAAGCUGCCAGCAUGAUCUC	24,25	304,74	DW226555	Proteína com domínio LIM
167g	GAAGCUGCCAGCAUGAUCUGG	974,58	207,26	ES834701 TC243185	Fator de resposta a Auxina [C,D]

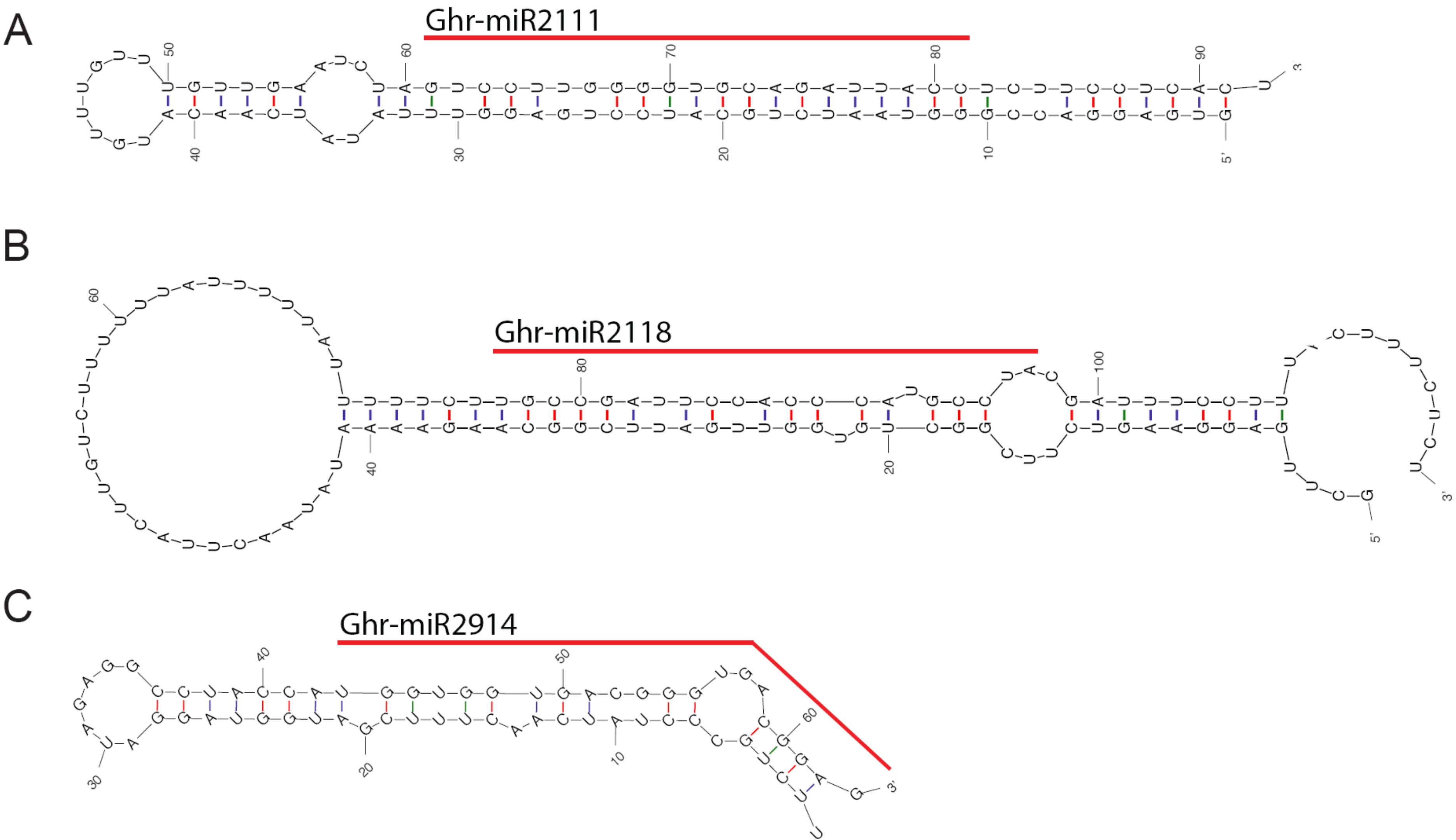
Identification of putative targets



Predicted Ghr-miRNA targets for 50 miRNA families, which include genes involved in disease resistance, auxin response, transcription factors, metabolism and metal ion transport.

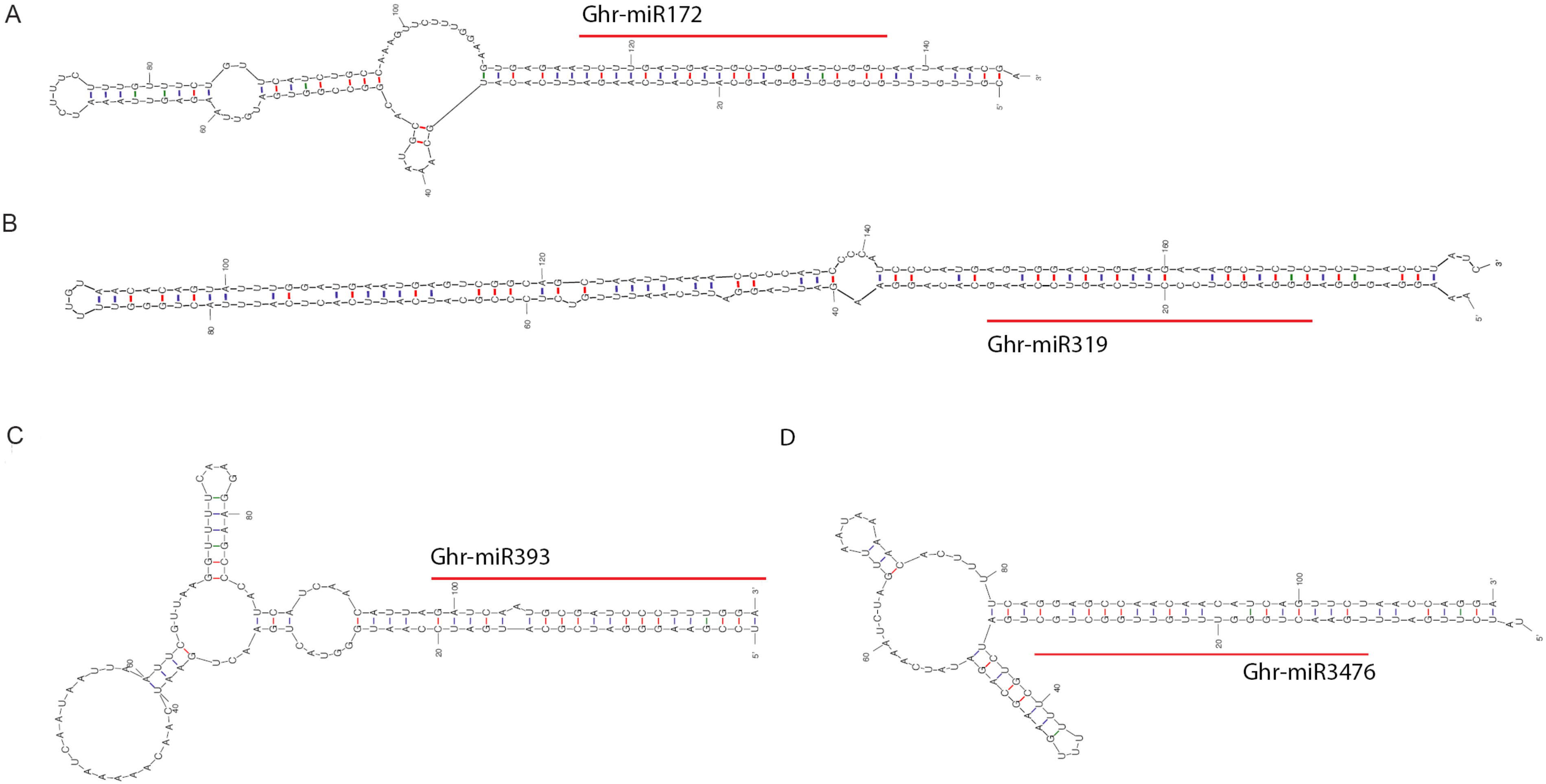
IDENTIFICATION OF PRE-MiCRO RNAs

Predicted stem-loop hairpin secondary structures of three new cotton miRNAs identified



Predição da estrutura secundária mfold: 21 Pré-miRNAs

Predicted stem-loop hairpin secondary structures of four novel cotton miRNAs



reduced frequency of 24-nt sRNAs in the IL

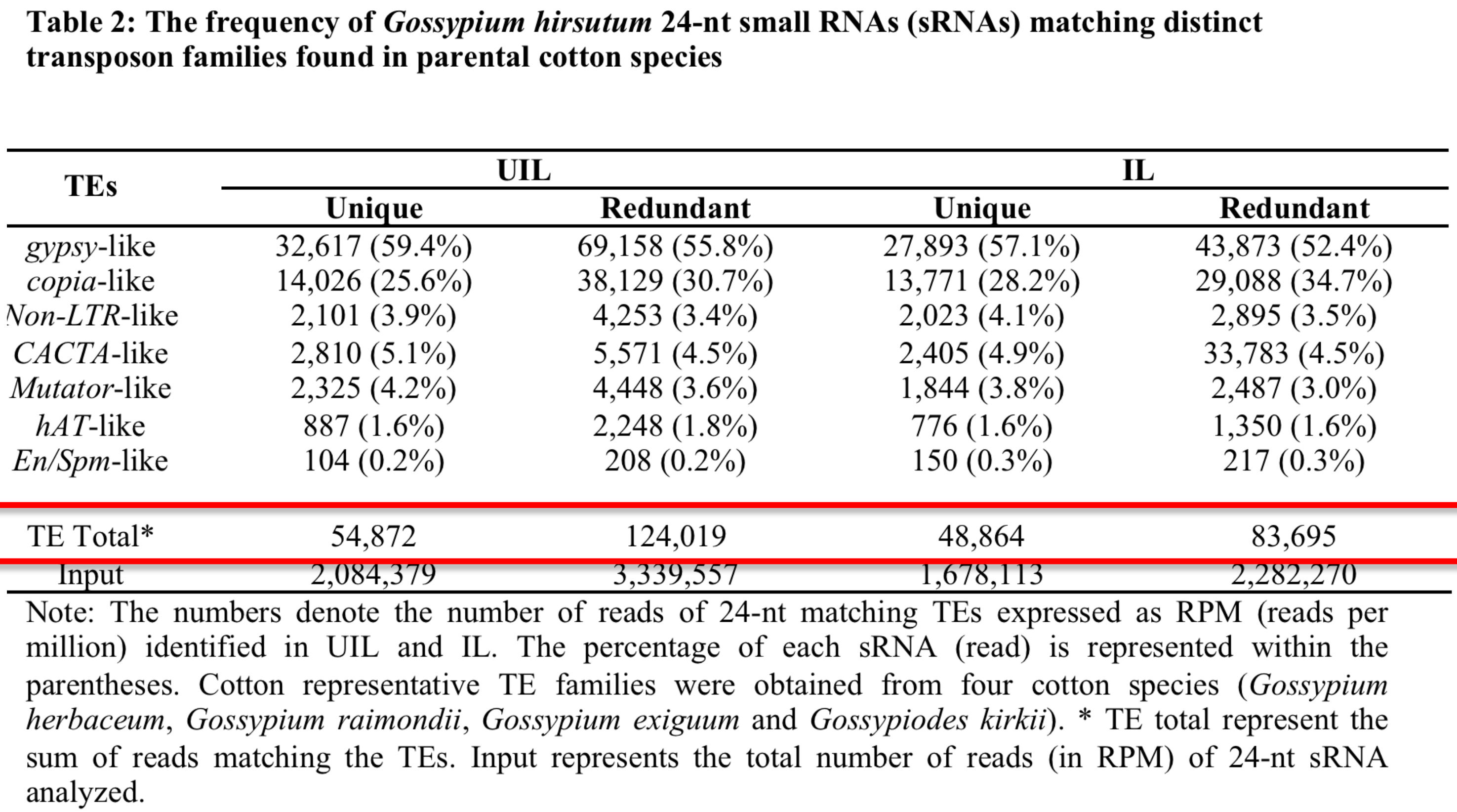


Table 2: The frequency of *Gossypium hirsutum* 24-nt small RNAs (sRNAs) matching distinct transposon families found in parental cotton species

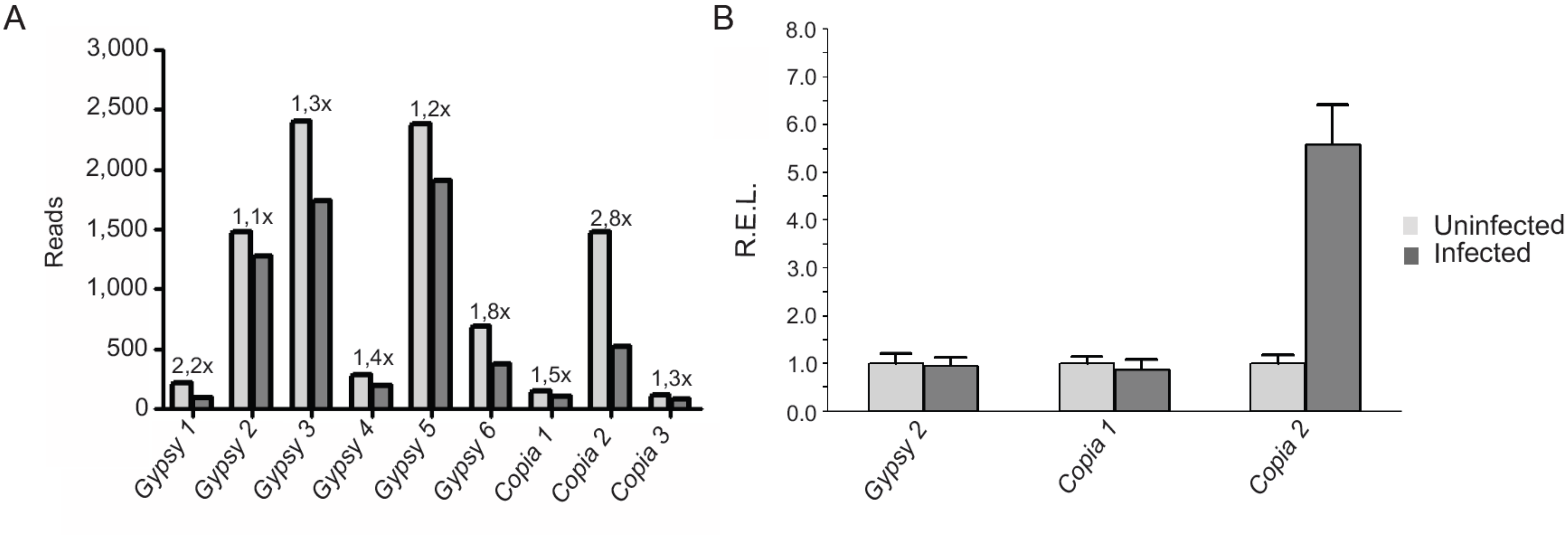
TEs	UIL		IL	
	Unique	Redundant	Unique	Redundant
gypsy-like	32,617 (59.4%)	69,158 (55.8%)	27,893 (57.1%)	43,873 (52.4%)
copia-like	14,026 (25.6%)	38,129 (30.7%)	13,771 (28.2%)	29,088 (34.7%)
Non-LTR-like	2,101 (3.9%)	4,253 (3.4%)	2,023 (4.1%)	2,895 (3.5%)
CACTA-like	2,810 (5.1%)	5,571 (4.5%)	2,405 (4.9%)	33,783 (4.5%)
Mutator-like	2,325 (4.2%)	4,448 (3.6%)	1,844 (3.8%)	2,487 (3.0%)
hAT-like	887 (1.6%)	2,248 (1.8%)	776 (1.6%)	1,350 (1.6%)
En/Spm-like	104 (0.2%)	208 (0.2%)	150 (0.3%)	217 (0.3%)
TE Total*	54,872	124,019	48,864	83,695
Input	2,084,379	3,339,557	1,678,113	2,282,270

Note: The numbers denote the number of reads of 24-nt matching TEs expressed as RPM (reads per million) identified in UIL and IL. The percentage of each sRNA (read) is represented within the parentheses. Cotton representative TE families were obtained from four cotton species (*Gossypium herbaceum*, *Gossypium raimondii*, *Gossypium exiguum* and *Gossypiodes kirkii*). * TE total represent the sum of reads matching the TEs. Input represents the total number of reads (in RPM) of 24-nt sRNA analyzed.

GSS TE libraries from Dr. Jennifer Hawkins, West Virginia University, USA

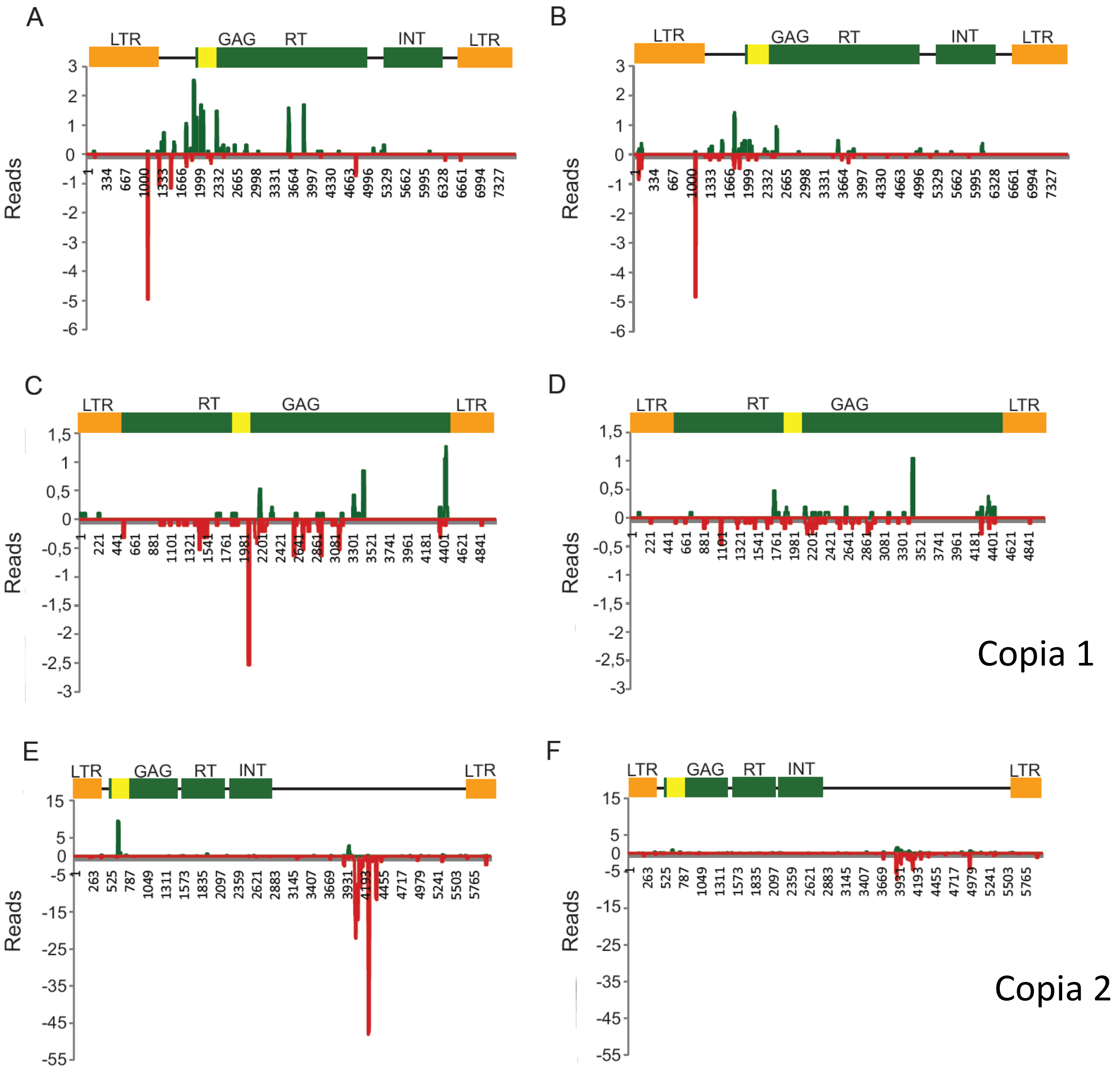
Analysis of cotton 24-nt sRNAs matching TEs

RNAi is necessary to specifically silence TEs through DNA methylation
24-nt sRNA - methylation



A – deep-sequencing: 24-nt sRNAs mapping retrotransposable elements in the uninfected and CLRDV-infected sRNA libraries. Reads are expressed as reads per 10 million. B – Detection of three TEs (*Gypsy 2*, *Copia 1* and *Copia 2*) transcripts in 5 dpi CLRDV-infected and uninfected cotton leaves by RT-qPCR

Copia 2 element was up-regulated in the infected leaf, which correlates well with the observed reduction in the total number of siRNAs mapping to this retrotransposon



Gypsy-like 2

Copia 1

Copia 2

Fig. 6 The spatial distribution and frequency of cotton 24-nt small RNAs (sRNAs) along *gypsy*-like and *copia*-like retrotransposons in CLRDV-infected and uninfected libraries The distribution of perfectly matching 24-nt sRNAs along the *Gossypium hirsutum Gypsy 2* full-length sequence in the uninfected (A) and CLRDV-infected (B) libraries.

Lab. de Virologia Molecular Vegetal

Instituto de Microbiologia
Paulo de Goés UFRJ

Depto. Genética/I.Biologia

Dr. Tatiane da Franca Silva

Prof. Régis Lopes Corrêa

Dr. Jennifer Hawkins, Wesr Virginia University, USA

FASTERIS

Laurent Farinelli
Cécile Deluen
Magne Osteras

Dr. Elisson Romannel

Prof. Carlos Guerra Schrago

FAPERJ
Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro

CAPES

CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico