

Estudios sobre la regulación de la expresión génica por microARNs en plantas mediante estrategias bioinformáticas

Uciel Chorostecki

Director Dr. Javier Palatnik
Instituto Biología Molecular y Celular Rosario

- ▶ Los microARNs (miARNs) son ARN pequeños de 20-22 nt que regulan la expresión génica en animales y plantas.
- ▶ En plantas controlan procesos vitales como el desarrollo, señalización hormonal y respuestas al estrés

Objetivos

1. Identificar genes regulados por miARNs en plantas.
2. Estudiar la biogénesis de los miARNs en plantas.

Objetivos

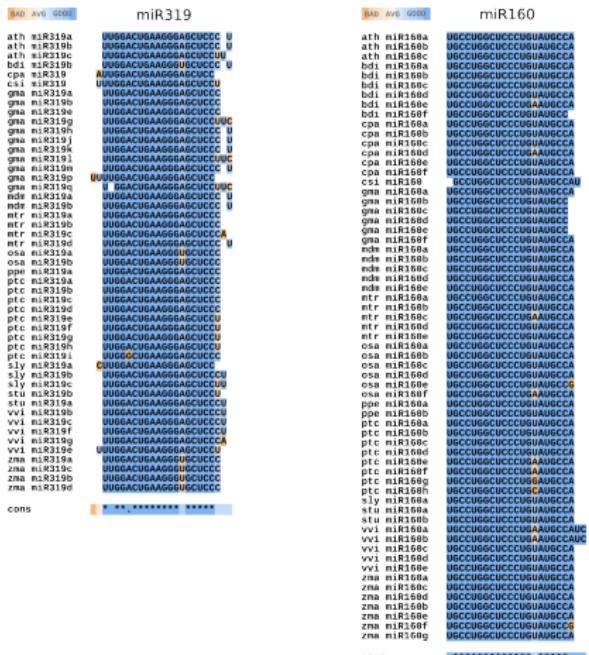
1. Identificar genes regulados por miARNs en plantas.
2. Estudiar la biogénesis de los miARNs en plantas.

- ▶ Diseñar una estrategia y una herramienta web para la identificación de genes blancos regulados por miARNs en plantas.
- ▶ Desarrollar herramientas para el análisis de los intermediarios de procesamiento de miARNs en plantas.
- ▶ Identificar y caracterizar precursores de miARNs en distintas especies que tengan mecanismos de procesamiento distintos.
- ▶ Caracterizar la relación entre la evolución de los precursores de miARNs en plantas y los mecanismos de procesamiento determinados previamente.

- ▶ Diseñar una estrategia y una herramienta web para la identificación de genes blancos regulados por miARNs en plantas.
- ▶ Desarrollar herramientas para el análisis de los intermediarios de procesamiento de miARNs en plantas.
- ▶ Identificar y caracterizar precursores de miARNs en distintas especies que tengan mecanismos de procesamiento distintos.
- ▶ Caracterizar la relación entre la evolución de los precursores de miARNs en plantas y los mecanismos de procesamiento determinados previamente.

Aplicaciones bioinformáticas para el estudio de interacciones miARN-gen blanco

Conservación y divergencia de miARNs en distintas especies



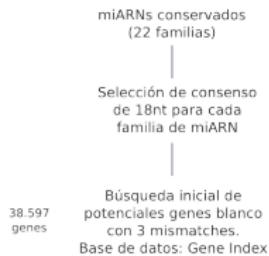
22 miARNs que están conservados en Angiospermas

miARNs conservados
(22 familias)

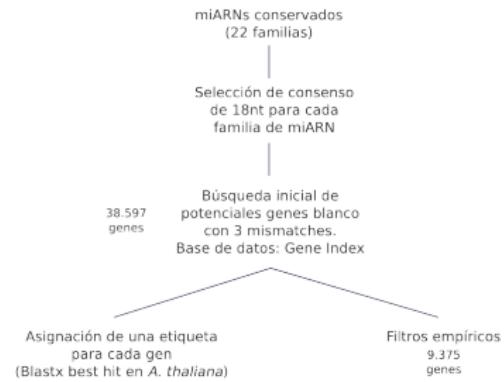
Secuencias consensos de 18 nt (2-19)

miARNs conservados
(22 familias)
|
Selección de consenso
de 18nt para cada
familia de miARN

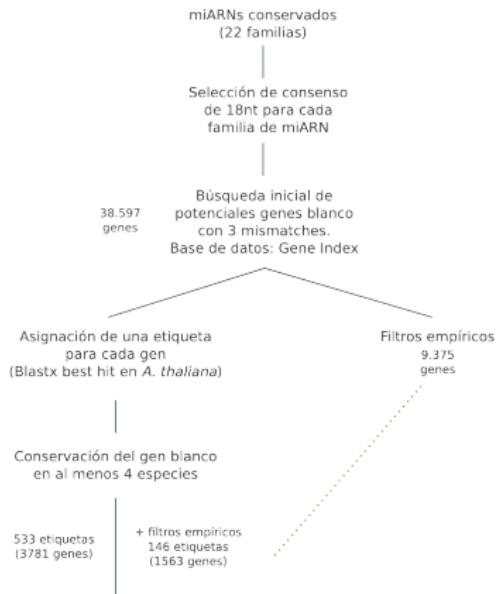
Primera búsqueda general de potenciales genes blancos



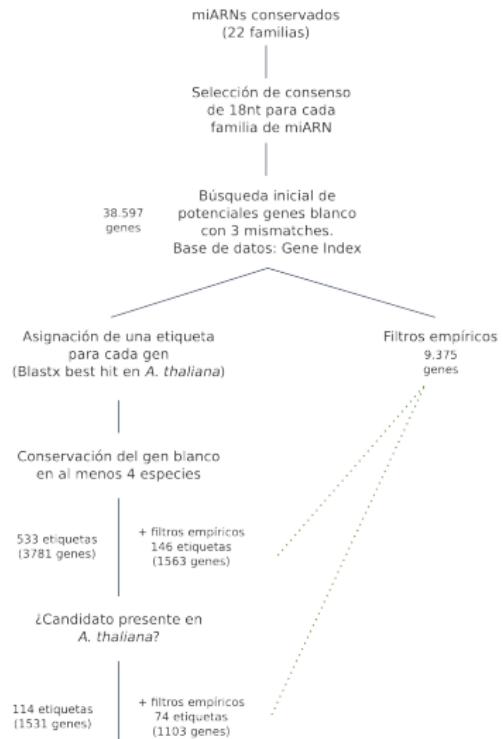
Filtros de evolución y empíricos de interacción miARN-gen blanco



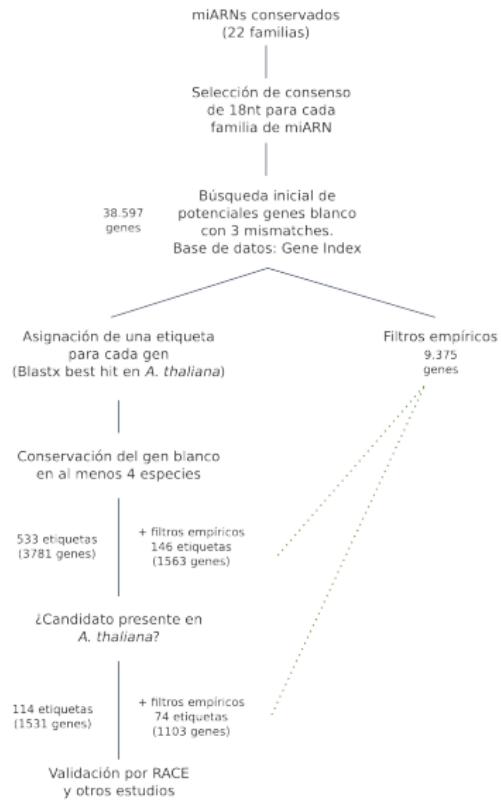
Mínimo de 4 especies requeridas



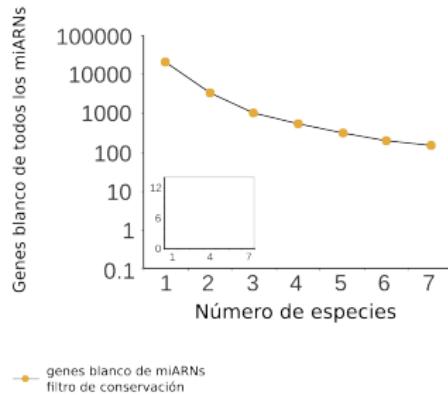
Genes blancos en *A. thaliana*



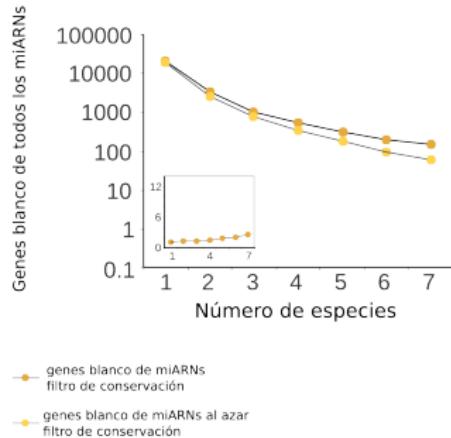
Validación experimental



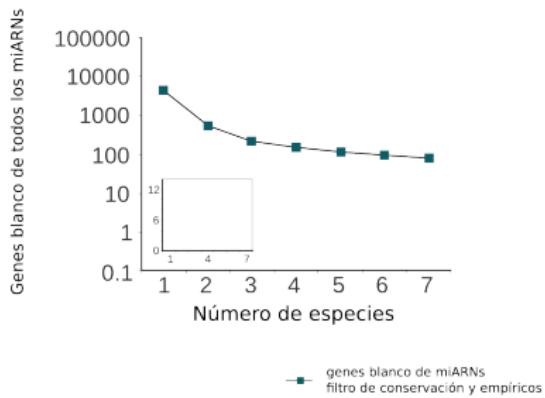
Conservación de la interacción en distintas especie



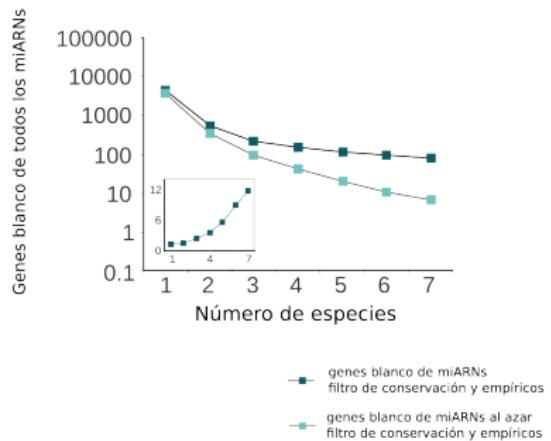
La relación señal/ruido incrementa al aumentar el número de especies



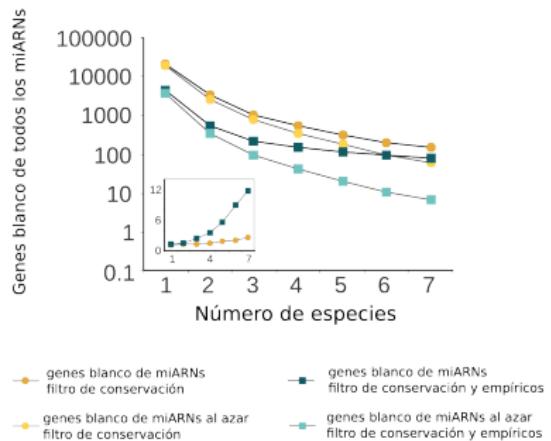
Selección de candidatos teniendo en cuenta los filtros empíricos



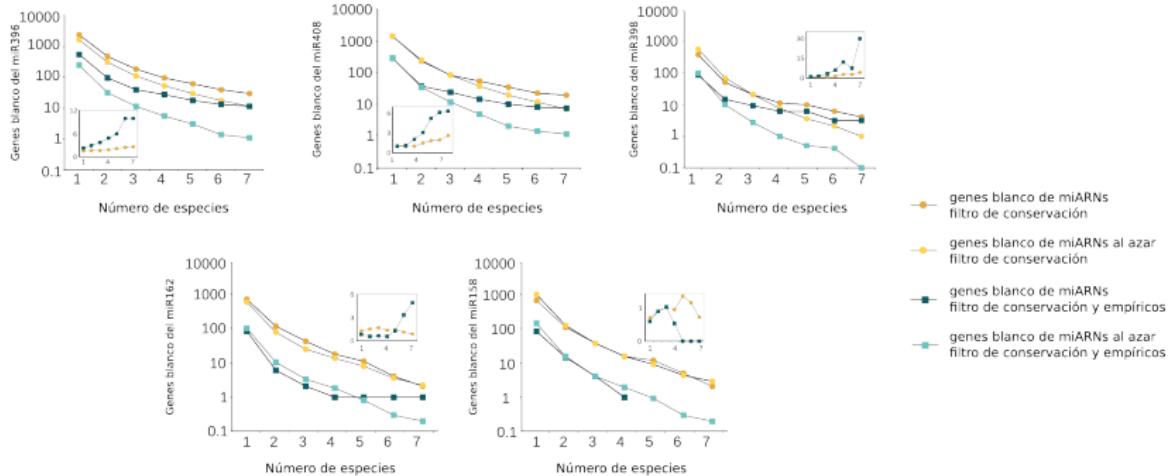
Al aplicar filtros empíricos y de conservación la relación señal/ruido aumenta



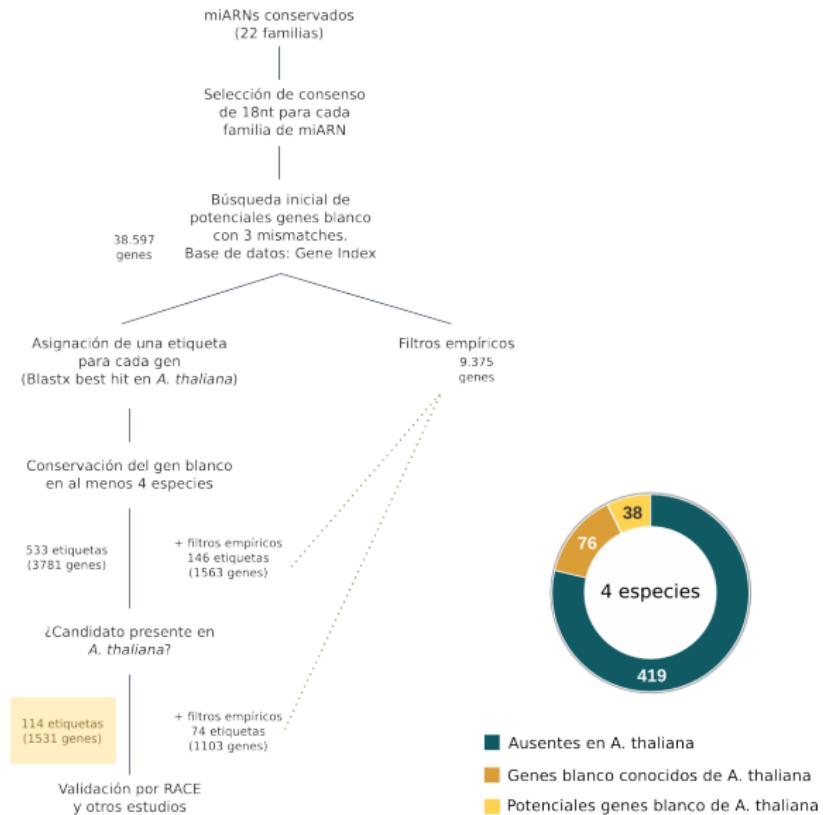
Efecto sinérgico al combinar filtro de conservación evolutiva y empíricos



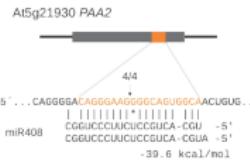
El número de genes blancos candidatos y la relación señal/ruido es variable entre los distintos miARNs



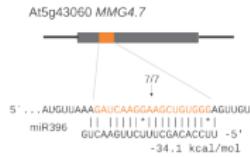
Potenciales genes blancos utilizando solo conservación evolutiva



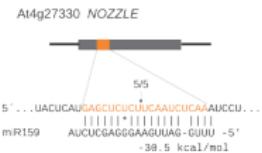
Nuevos genes blancos validados en *A. thaliana*



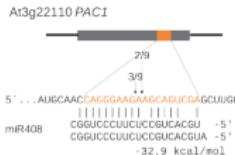
Arabidopsis thaliana CAGGGGAAGGGCAGUGCA
Capsicum annuum CAGGGAAAGGAGCAGUUGCA
Glycine max CAGGGAAAGGGCAGUUGCA
Gossypium raimondii CAGGGAAAGGGCAGUUGCA
Haseoulus vulgaris UAGGGAAAGGGCAGUUGCA
Ipomoea nil CAGGGAAAGGGCAGUUGCA
Lactuca sativa CAGGGAAAGGGCAGUUGCA
Medicago truncatula CAGGGAAAGGGCAGUUGCA
Saccharum officinarum UAGGGAAAGGGCAGUUGCA
Solanum lycopersicum CAGGGAAAGGGCAGUUGCA
Solanum tuberosum CAGGGAAAGGGCAGUUGCA
Sorghum bicolor UAGGGAAAGGGCAGUUGCA
Theobroma cacao CAGGGAAAGGGCAGUUGCA
Vitis vinifera CAGGGAAAGGGCAGUUGCA



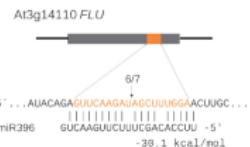
Arabidopsis thaliana GAUCAAGGAAGCUGUGGG
Citrus clementina GAUCAAGGAAGCUGUGGG
Citrus sinensis GAUCAAGGAAGCUGUGGG
Helianthus annuus GAUCAAGGAAGJUGUGGA
Ipomoea nil GAUCAAGGAAGCUGUGGG
Medicago truncatula GAUCAAGGAAGJUGUGGA
Nicotiana benthamiana GAUCAAGGAAGCUGUGGG
Nicotiana tabacum GAUCAAGGAAGCUGUGGG
Petunia hybrida GAUCAAGGAAGCUGUGGG
Triphysaria versicolor AUUGAAAGGAAGCUGUGGC
Vitis vinifera CUIGAAGAAAGCUGUGGG



Arabidopsis thaliana GAGCUDUCUUCAAUUCAA
Brassica napus GAGCCCUCUCAAUC -CCA
Lactuca sativa GAACUCUCUUCAAUU -CCA
Lactuca serriola GACUCUCUUCAAUU -CCA
Oryza sativa GAGCCCCCUUCAAAC -GAA
Vitis vinifera GAGCUCUCUUCAAUUCCAA

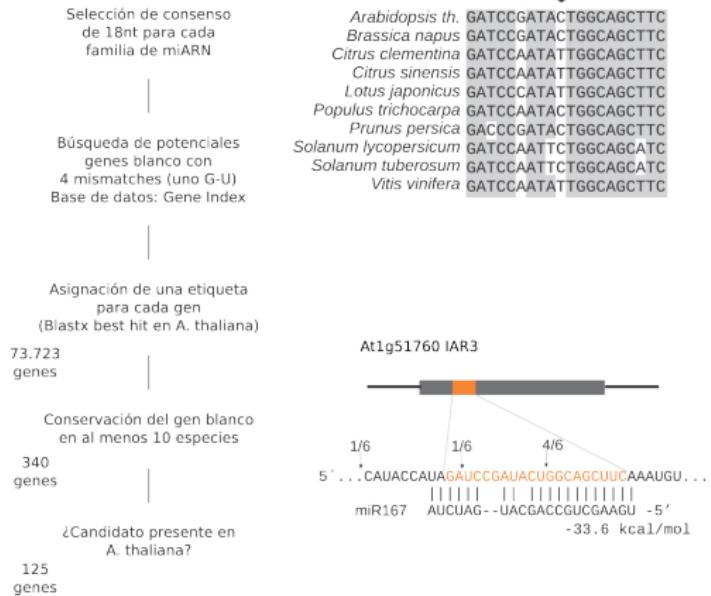


Arabidopsis thaliana CAGGGAAAGAGCAGUUGCA
Capsicum annuum AAGGGAAAGGGCAGUUGCA
Citrus clementina AAGGGAAAGAGCAGUUGCA
Glycine max AAGGGAAAGAGCAGUUGCA
Helianthus annuus CAGAAAAGAGCAGUUGCA
Lotus japonicus UAGAGAAAGAGCAGUUGCA
Medicago truncatula AAGGGAAAGAGCAGUUGCA
Nicotiana tabacum AAGGGAAAGAGCAGUUGCA
Phaseolus coccineus AAGGGAAAGAGCAGUUGCA
Prunus persica AAGAGAAAGAGCAGUUGCA
Solanum tuberosum AAGGGAAAGAGCAGUUGCA
Theobroma cacao AAGGGAAAGAGCAGUUGCA
Triphysaria versicolor CAGGGAGGAGCUGUUC



Arabidopsis thaliana GUUCAAGAUAGCUUJUGG
Brassica napus GUUCAAGACAGCUUJUGG
Festuca arundinacea GUUCAAGACAGCUUJUGG
Hordeum vulgare GUUCAAGAUAGCUCUUGGA
Ipomoea nil GUUCAAGAAAGCCUUGA
Panicum virgatum GUUCAGGACAGCUCUUGGA
Triticum aestivum GUUCAAGACGCGCUUGGA

Nuevos genes blancos con interacciones G-U



Identificación de genes blancos específicos de Solanaceae

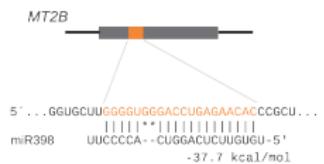
Genes blanco de miARNs candidatos

Genes blanco presente en al menos tres especies de Solanaceae

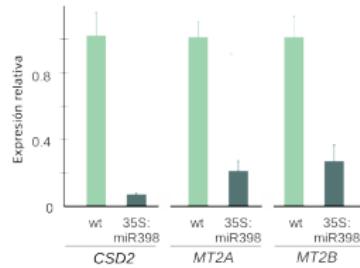
132 genes

Genes blanco específicos de Solanacea?

41 genes



Capsicum annuum GGGGUUGGGACCUUGAGAAAGAC
Nicotiana benthamiana GGGGUUGGGACCUUGAGAAAGAC
Nicotiana tabacum GGGGUUGGGACCUUGAGAAAGAC
Petunia hybrida GGGGUUGGGACCUUGAGAAAGAC
Solanum lycopersicum GGGGUUGGGACCUUGAGAAAGAC
Solanum tuberosum GGGGUUGGGACCUUGAGAAAGAC



ComTAR, una herramienta para predecir potenciales genes blancos regulados por miARNs en plantas basada en la conservación evolutiva del par miARN-gen blanco. Permite realizar la búsqueda de:

- ▶ potenciales genes blancos a partir de un miARN.
- ▶ familias de potenciales genes blancos de un miARN.
- ▶ un gen de interés para ver si es potencial gen blanco de algún miARN conservado
- ▶ nuevos ARNs pequeños

ComTAR, una herramienta para predecir potenciales genes blancos regulados por miARNs en plantas basada en la conservación evolutiva del par miARN-gen blanco. Permite realizar la búsqueda de:

- ▶ potenciales genes blancos a partir de un miARN.
- ▶ familias de potenciales genes blancos de un miARN.
- ▶ un gen de interés para ver si es potencial gen blanco de algún miARN conservado
- ▶ nuevos ARNs pequeños

► <http://rnabiology.ibr-conicet.gov.ar/comtar>

Potenciales genes blancos del miR398



Targets					
Arabidopsis Tag [?]	Count [?]	Species [?]	Target description [?]	Gene family [?]	Alignments [?]
AT1G12520	16	Show/Hide	copper chaperone for SOD1		View
AT1G08830	9	Show/Hide	copper/zinc superoxide dismutase 1		View
AT3G15640	4	Show/Hide	Rubredoxin-like superfamily protein		View
AT1G08230	3	Show/Hide	Rubredoxin-like superfamily protein		View
AT2G26975	3	Show/Hide	Ctr copper transporter family		View
AT2G40340	3	Show/Hide	Integrase-type DNA-binding superfamily protein	AP2-EREBP Transcription Factor Family	View
AT3G27200	3	Show/Hide	Cupredoxin superfamily protein	Miscellaneous Membrane Protein Families	View
AT3G43860	3	Show/Hide	glycosyl hydrolase 9A4	Glycoside Hydrolase Gene Families	View
AT4G00050	3	Show/Hide	basic helix-loop-helix (bHLH) DNA-binding superfamily protein	basic Helix-Loop-Helix (bHLH) Transcription Factor	View
AT5G12220	3	Show/Hide	las1-like family protein		View

ComTAR permite visualizar el alineamiento, energía de hibridación en cada especie

Sequence ID	Species	5'-target-3' Alignment 3'-miRNA-5' [•]	MFE [•]
Aquica_013_00504.1	Aquilegia coerulea	TTGGCGGACCTGGGAACAT * * * * * CCCACTGGACTCTTGTC	-31.7
471402	Arabidopsis lyrata	TGGGAGACCTGGGAACAC * * * * * CCCACTGGACTCTTGTC	-32.1
AT1G12520.1	Arabidopsis thaliana	TGGGAGACCTGGGAACAC * * * * * CCCACTGGACTCTTGTC	-32.1
Brad5g18900.3	Brachypodium distachyon	TTGGTGAACCTGGGAACGC ** * * * * CCCACTGGACTCTTGTC	-33.5
Bra026968	Brassica rapa	TGGGAGACCTGGGAACAC * * * * * CCCACTGGACTCTTGTC	-32.5
Carubv10011818m	Capsella rubella	TGGGAGACCTGGGAACAC * * * * * CCCACTGGACTCTTGTC	-32.1
evm.model.superconig_29.47	Carica papaya	TAGGTGAACCTGGAGACAT ** * * * * CCCACTGGACTCTTGTC	-34.2
Ciclev10021134m	Citrus clementina	TTGGTGAACCTGGGAACAC ** * * * * CCCACTGGACTCTTGTC	-33.9
orange1.1g020436m	Citrus sinensis	TTGGTGAACCTGGGAACAg ** * * * * CCCACTGGACTCTTGTC	-32.7

Conclusiones I

miARN	Consenso (18 nt)	Targets conocidos ^(a,b)
miR156	GACAGAAGAGAGTGAGCA	factores de transcripción SPL
miR159	TTGGATTGAAGGGAGCTC	factores de transcripción MYB, NOZZLE (NZL)
miR160	GCCTGGCTCCCTGTATGC	factores de transcripción ARF
miR162	CGATAAACCTTGTGATCC	DCL1
miR164	GGAGAAAGCAGGGCACGTG	factores de transcripción NAC
miR166	CGGACCAGGCTTCATTC	factores de transcripción HDZip
miR167	GAAGCTGCCAGCATGATC	factores de transcripción ARF, IAA-ALANINE RESISTANT 3 (IAR3)
miR168	CGCTTGGTGCAGGTCGGG	AGO1
mir169	AGCCAAGGATGACTTGCC	factores de transcripción CCAAT-HAP2
mir171	TTGAGCCGTGCCAATATC	factores de transcripción GRAS
miR172	GAATCTTGTATGATGCTGC	factores de transcripción AP2
miR319	TGGACTGAAGGGAGCTCC	factores de transcripción TCP
miR390	AGCTCAGGGAGGGATAGCG	TAS RNA
miR393	CCAAAGGGATCGCATTTGA	TIR1 proteins, F-BOX proteins
miR394	TGGCAATTCTGTCCACCTC	proteínas F-BOX
miR395	TGAAGTTGTTGGGGAAAC	ATP-sulfurilasas, transportadores de sulfato
miR396	TCCACAGCTTCTTGAAAC	factores de transcripción GRF, MMG4.7, FLUORESCENT IN BLUE LIGHT (FLU)
miR397	CATTGAGTGCAGCGTTGA	Laccases
miR398	GTGTTCTCAGGTACCCCC	Cu/Zn SODs, CytC oxidase protein subunit, Chaperona de cobre (CCS)
miR399	GCCAAAGGGAGATTGCC	Enzima E2 de conjugación de ubiquitina
miR408	TGCACTGCCTTCTCCCTG	Blue copper proteins, Laccases, P-TYPE ATPase (PAA2), PAC1 (Proteasome component)
miR827	TAGATGACCATCAGCAAA	SPX proteins

- ▶ Diseñamos una estrategia para identificar genes blancos regulados por miARNs en plantas, basado en la conservación evolutiva del par miARN-gen blanco.

- ▶ Diseñamos una estrategia para identificar genes blancos regulados por miARNs en plantas, basado en la conservación evolutiva del par miARN-gen blanco.
- ▶ Identificamos nuevos genes blancos en *A. thaliana* y se validaron experimentalmente varios de ellos.

- ▶ Diseñamos una estrategia para identificar genes blancos regulados por miARNs en plantas, basado en la conservación evolutiva del par miARN-gen blanco.
- ▶ Identificamos nuevos genes blancos en *A. thaliana* y se validaron experimentalmente varios de ellos.
- ▶ Esta estrategia puede ser utilizada para identificar genes blancos presentes en un grupo específico de especies.

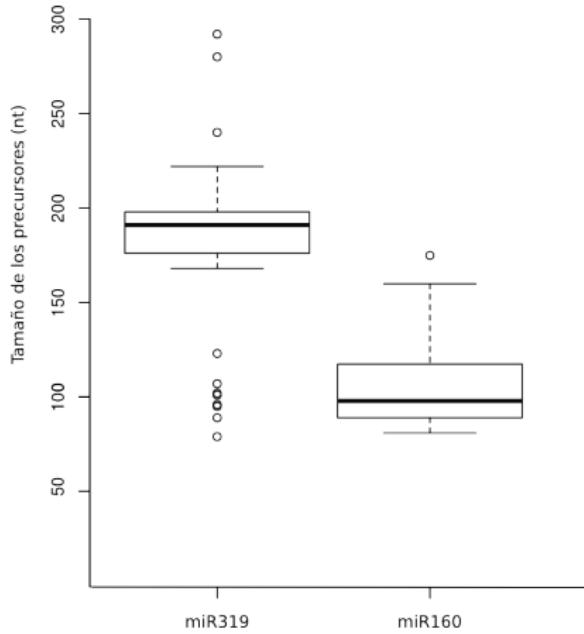
- ▶ Diseñamos una estrategia para identificar genes blancos regulados por miARNs en plantas, basado en la conservación evolutiva del par miARN-gen blanco.
- ▶ Identificamos nuevos genes blancos en *A. thaliana* y se validaron experimentalmente varios de ellos.
- ▶ Esta estrategia puede ser utilizada para identificar genes blancos presentes en un grupo específico de especies.
- ▶ Interacciones miARN-gen blanco conservadas probablemente participen en procesos biológicos relevantes

- ▶ Diseñamos una estrategia para identificar genes blancos regulados por miARNs en plantas, basado en la conservación evolutiva del par miARN-gen blanco.
- ▶ Identificamos nuevos genes blancos en *A. thaliana* y se validaron experimentalmente varios de ellos.
- ▶ Esta estrategia puede ser utilizada para identificar genes blancos presentes en un grupo específico de especies.
- ▶ Interacciones miARN-gen blanco conservadas probablemente participen en procesos biológicos relevantes
- ▶ Desarrollamos una herramienta web denominada comTAR para predecir potenciales genes blancos regulados por miARNs en plantas

- ▶ Diseñar una estrategia y una herramienta web para la identificación de genes blancos regulados por miARNs en plantas.
- ▶ Desarrollar herramientas para el análisis de los intermediarios de procesamiento de miARNs en plantas.
- ▶ Identificar y caracterizar precursores de miARNs en distintas especies que tengan mecanismos de procesamiento distintos.
- ▶ Caracterizar la relación entre la evolución de los precursores de miARNs en plantas y los mecanismos de procesamiento determinados previamente.

- ▶ Diseñar una estrategia y una herramienta web para la identificación de genes blancos regulados por miARNs en plantas.
- ▶ Desarrollar herramientas para el análisis de los intermediarios de procesamiento de miARNs en plantas.
- ▶ Identificar y caracterizar precursores de miARNs en distintas especies que tengan mecanismos de procesamiento distintos.
- ▶ Caracterizar la relación entre la evolución de los precursores de miARNs en plantas y los mecanismos de procesamiento determinados previamente.

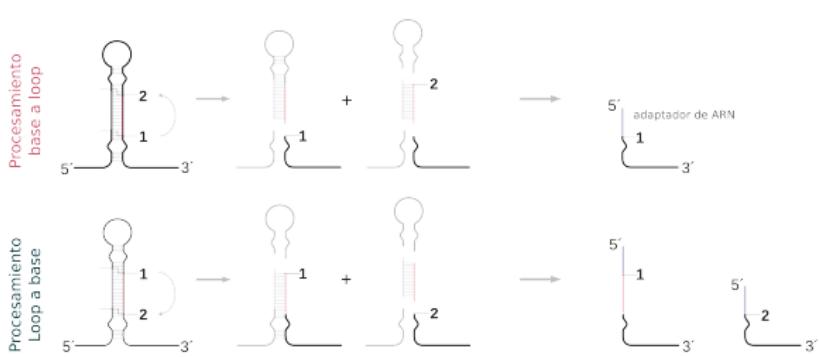
Precursos en plantas son muy variables en tamaño y forma



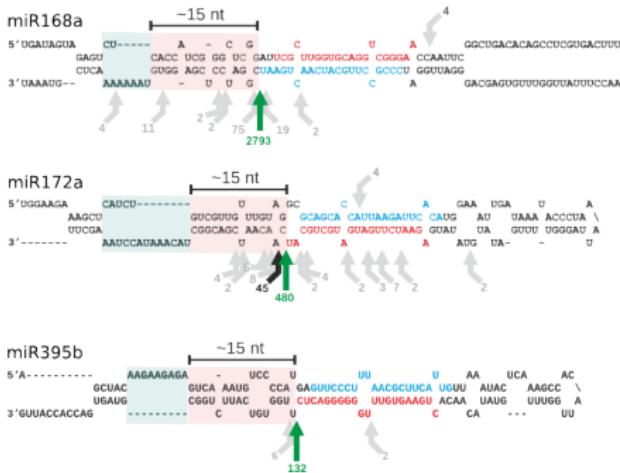
Bibliotecas SPARE para estudios genómicos de biogénesis de miARNs en plantas



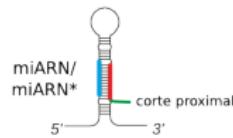
Bibliotecas SPARE para estudios genómicos de biogénesis de miARNs en plantas



Visualización de precursores que se procesan desde la base

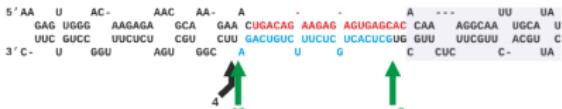


Detección solamente
del corte proximal

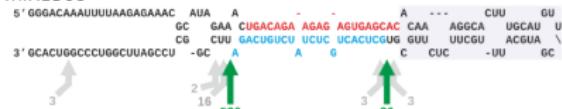


Visualización de precursores que se procesan desde el loop

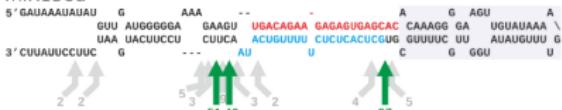
miR156a



miR156c



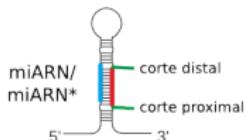
miR156d



miR160a

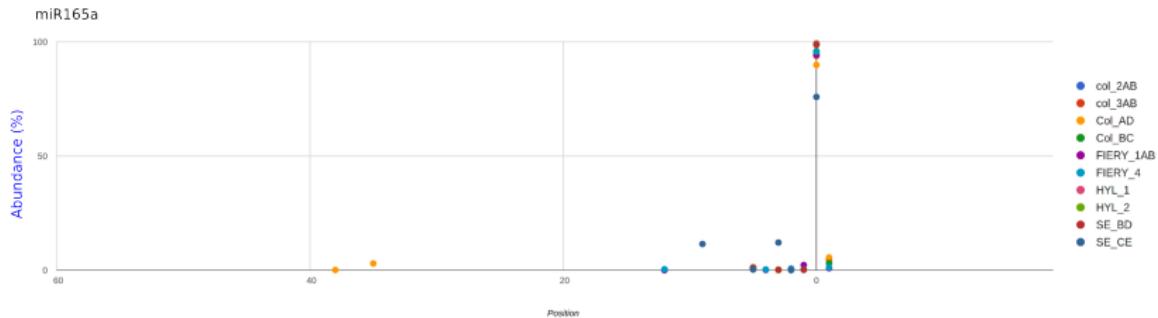


Detección de ambos cortes,
proximal y distal



Bibliotecas	Muestras	Secuencias totales	Secuencias que mapean los precursores	Secuencias únicas que mapean los precursores
col_2AB	Col-0 réplica 1. Control de fiery y hyl1	13911694	80166	308
col_3AB	Col-0 réplica 2. Control de fiery y hyl1	16618008	126556	426
Col_AD	Col-0 réplica 1. Control de se	13758567	119368	496
Col_BC	Col-0 réplica 2. Control de se	14648459	241973	553
FIERY_1AB	fiery réplica 1	9832923	470789	1655
FIERY_4	fiery réplica 2	23529725	821562	1752
HYL_1	hyl1 réplica 1	10171629	45653	316
HYL_2	hyl1 réplica 2	8864406	35860	320
SE_BD	se réplica 1	15291993	299513	639
SE_CE	se réplica 1	25296809	510438	693

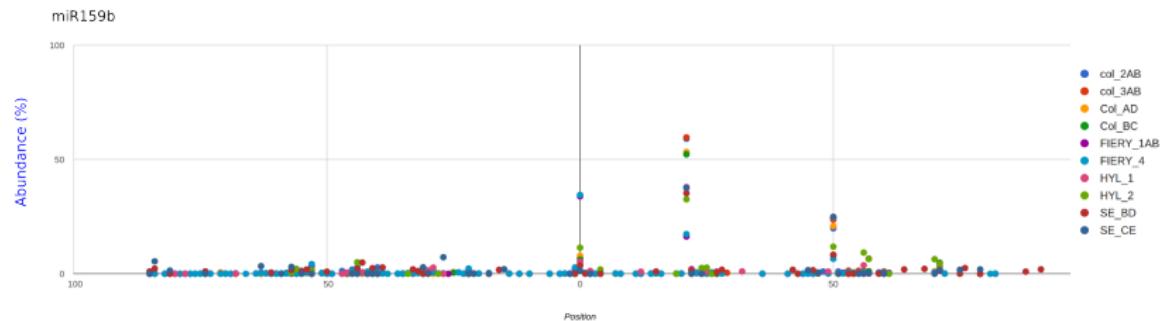
Precursors procesados desde la base



Position	col_2AB	col_3AB	Col_AD	Col_BC	FIERY_1AB	FIERY_4	HYL_1	HYL_2	SE_BD	SE_CE
-38	0	0	4	0	0	0	0	0	0	0
-35	0	0	89	0	0	0	0	0	0	0
-12	1	0	0	0	10	147	0	0	0	0
-9	0	0	0	0	0	0	0	0	0	270
-5	29	67	40	26	116	326	2	1	5	9
-4	0	0	0	0	15	111	0	0	0	0
-3	0	0	1	0	0	0	0	1	1	285
-2	0	0	0	0	77	189	0	0	0	1
-1	0	0	0	13	247	141	0	0	1	0
0	2902	4574	2700	3993	10065	29877	317	314	546	1781
1	0	231	169	129	97	387	0	0	0	0

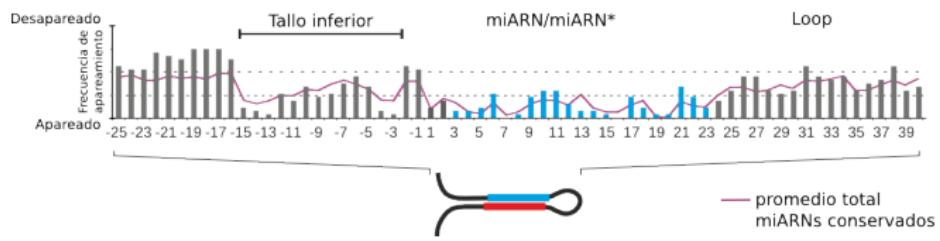


Precursors procesados desde el loop

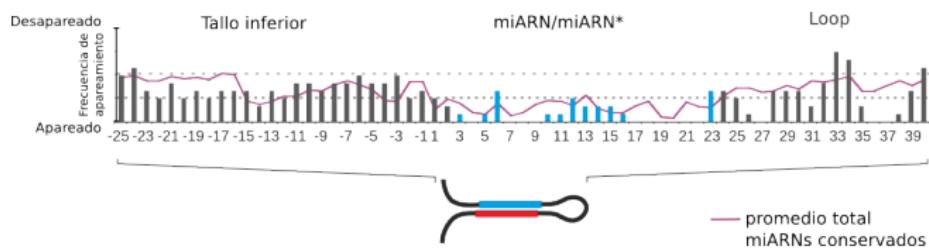


Position	col_2AB	col_3AB	Col_AD	Col_BC	FIERY_1AB	FIERY_4	HYL_1	HYL_2	SE_BD	SE_CE
21	10252	19917	5562	10992	6776	20171	2514	1477	1685	2752
50	3450	7908	2174	5164	3392	7627	1629	538	398	1824
0	1209	1319	822	1300	14091	40065	353	518	172	101
71	451	1070	233	1028	538	1713	302	201	70	125
53	9	461	1	64	250	822	0	0	1	0
55	92	450	16	157	365	639	52	24	17	1
57	32	379	66	218	115	227	442	294	6	23

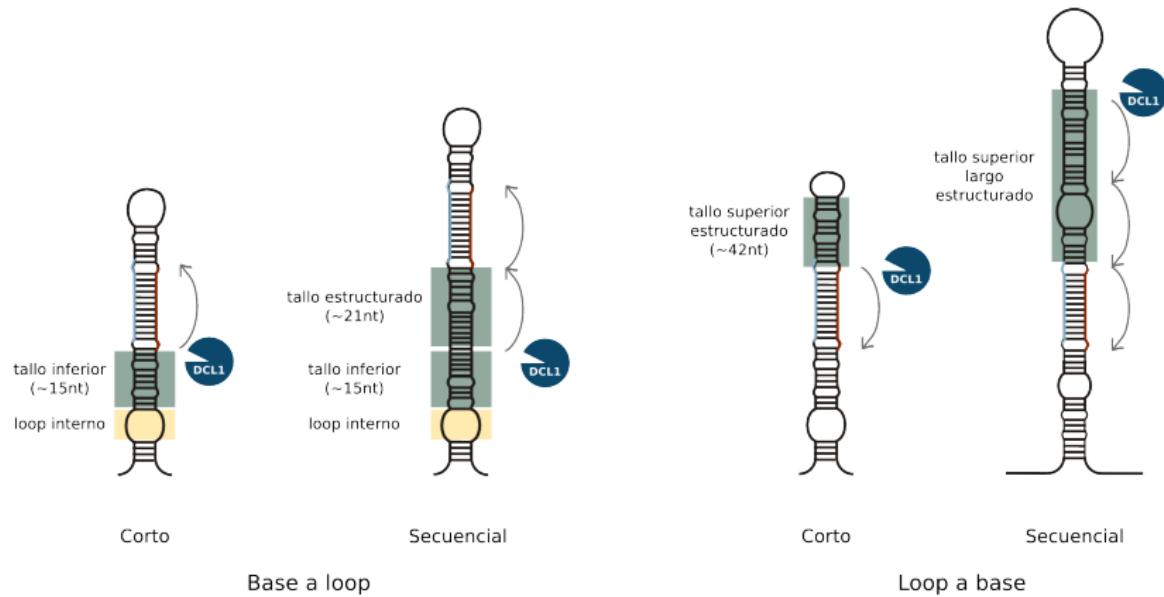
Tallo inferior de 15 nt en precursores procesados desde la base



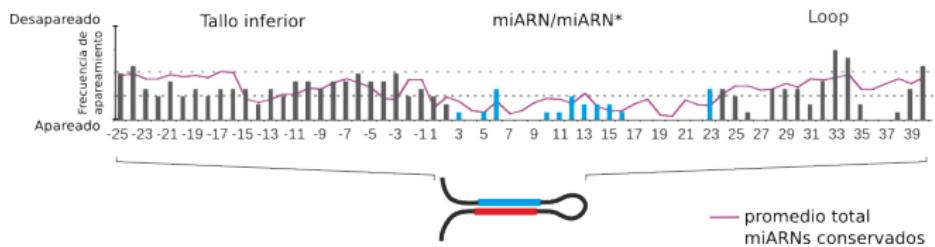
Región terminal estructurada en precursores procesados desde el loop



Conclusiones II



Búsqueda de ortólogos de precursores de miARNs.



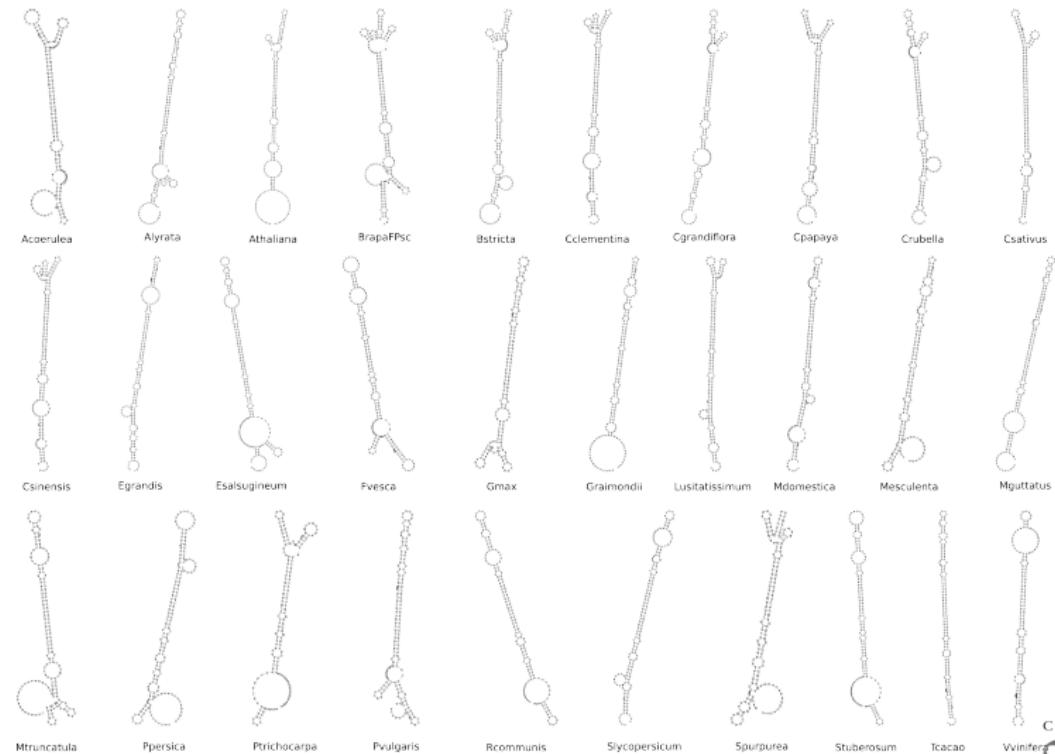
- ▶ Diseñar una estrategia y una herramienta web para la identificación de genes blancos regulados por miARNs en plantas.
- ▶ Desarrollar herramientas para el análisis de los intermediarios de procesamiento de miARNs en plantas.
- ▶ Identificar y caracterizar precursores de miARNs en distintas especies que tengan mecanismos de procesamiento distintos.
- ▶ Caracterizar la relación entre la evolución de los precursores de miARNs en plantas y los mecanismos de procesamiento determinados previamente.

- ▶ Diseñar una estrategia y una herramienta web para la identificación de genes blancos regulados por miARNs en plantas.
- ▶ Desarrollar herramientas para el análisis de los intermediarios de procesamiento de miARNs en plantas.
- ▶ Identificar y caracterizar precursores de miARNs en distintas especies que tengan mecanismos de procesamiento distintos.
- ▶ Caracterizar la relación entre la evolución de los precursores de miARNs en plantas y los mecanismos de procesamiento determinados previamente.

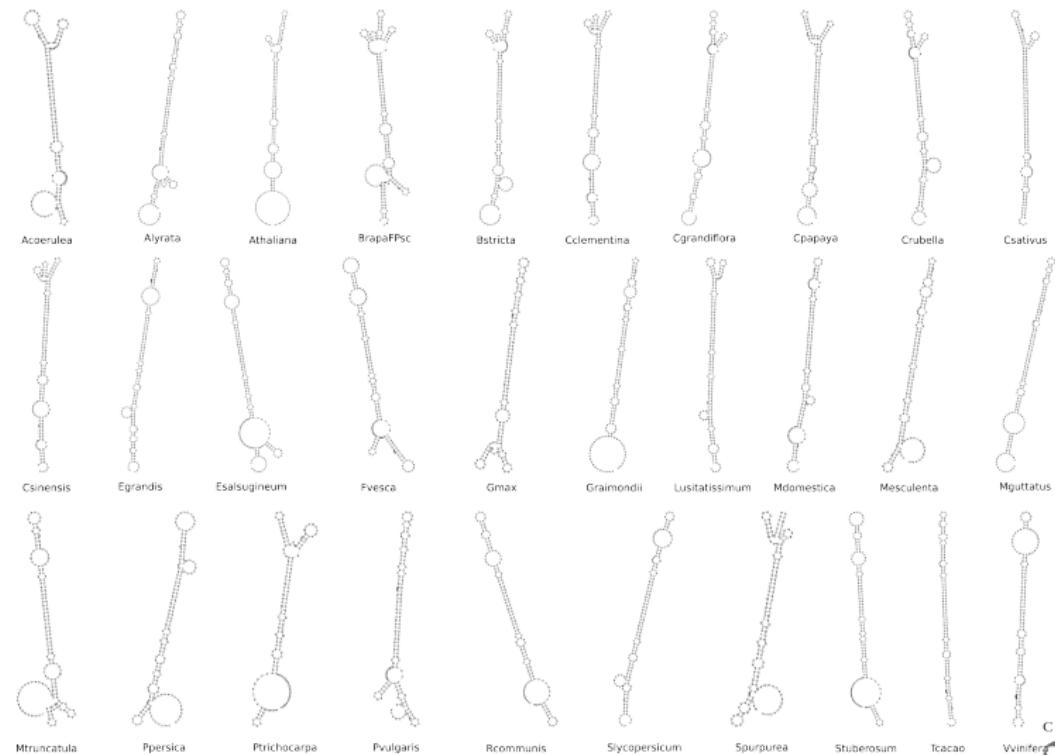
VER LO QUE PUSE EN EL SEMINARIO IBR!!!!

Comenzamos nuestro análisis con una definición arbitraria de los precursores de plantas incluyendo 150 nt fuera del par miARN/miARN*. Para cada miembro de cada familia de *A. thaliana* no es trivial asignarle un ortólogo en otra especie teniendo en cuenta la anotación de miRBase. Por esto, realizamos una búsqueda de ortólogos para cada miembro de cada familia de *A. thaliana* utilizando como criterio la técnica de Blast recíproco.

Existe un patrón estructural que comparten los precursores, en la región inmediata por debajo del dúplex miARN/miARN*



Existe un patrón estructural que comparten los precursores, en la región inmediata por debajo del dúplex miARN/miARN*



Es difícil deducir información concreta a partir de esta figura.

Conservación de la secuencia primaria del miR172a en distintas especies

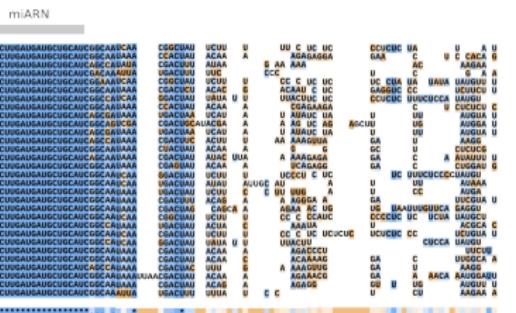
	BAD	AVG	GDD0	U	C	S	C
Alvareza	AG	UABAU	UCC	GGG	G	US	
Motomestica	BR	AUAG	UAB	AG	RD	US	
Urtica	BR	UAB	UAB	AG	RD	US	
Silicopera				ACACI	GU	GU	
BrapaPsc	AG	CBAC	GGG	AA	AA	GU	
Urticaria	BR	GGG	GGG	AA	AA	GU	
Cypradiflo	AG	UABAU	UAC	AA	CC	GU	
Urticaria	BR	GGG	GGG	AA	AA	GU	
Ptilimensis				ACACI	GU	GU	
Lutistatia				ACACI	GU	GU	
Urticaria				ACACI	GU	GU	
Gmax	CA	DAU	AG	AA	AA	GU	
Vvinifera				ACACI	GU	GU	
Urticaria				ACACI	GU	GU	
Spurpurea				ACACI	GU	GU	
Urticaria	AG	UABAU	UAB	AG	RD	GU	
Cstavtus				ACACI	GU	GU	
Acervaria	A	ACAC	GU	GU	GU	GU	
Urticaria				ACACI	GU	GU	
Mgutatus				ACACI	GU	GU	
Urticaria				ACACI	GU	GU	
Capapay	UG	UAB	GU	GG	GU	GU	
Athelia	AG	UABAU	UAB	GU	GU	GU	
Tecocao				ACACI	GU	GU	
Urticaria				ACACI	GU	GU	
Pvalmaris	AG	UABAU	AG	GU	GU	GU	
Urticaria				ACACI	GU	GU	
Urticaria				ACACI	GU	GU	
Strobosoma				ACACI	GU	GU	
cons				ACACI	GU	GU	

ANSWER

[CODE](#)

El miR172a maduro y el miR172a* están conservados en las distintas especies

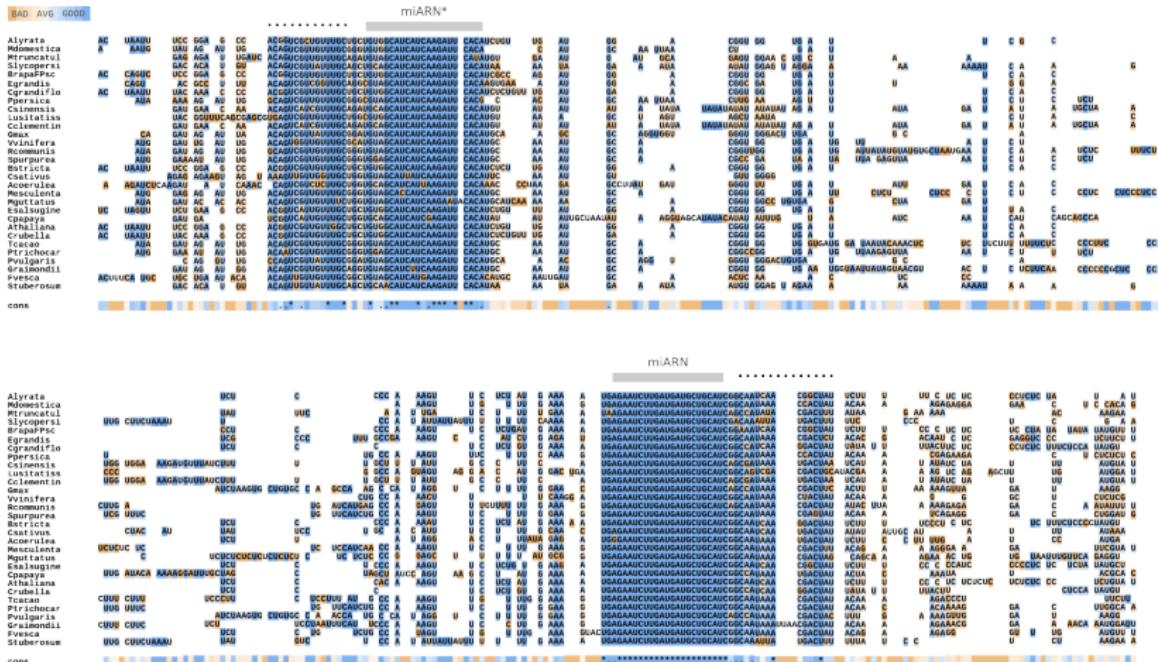
Copyright © 2010 Pearson Education, Inc., publishing as Pearson Benjamin Cummings. All Rights Reserved.



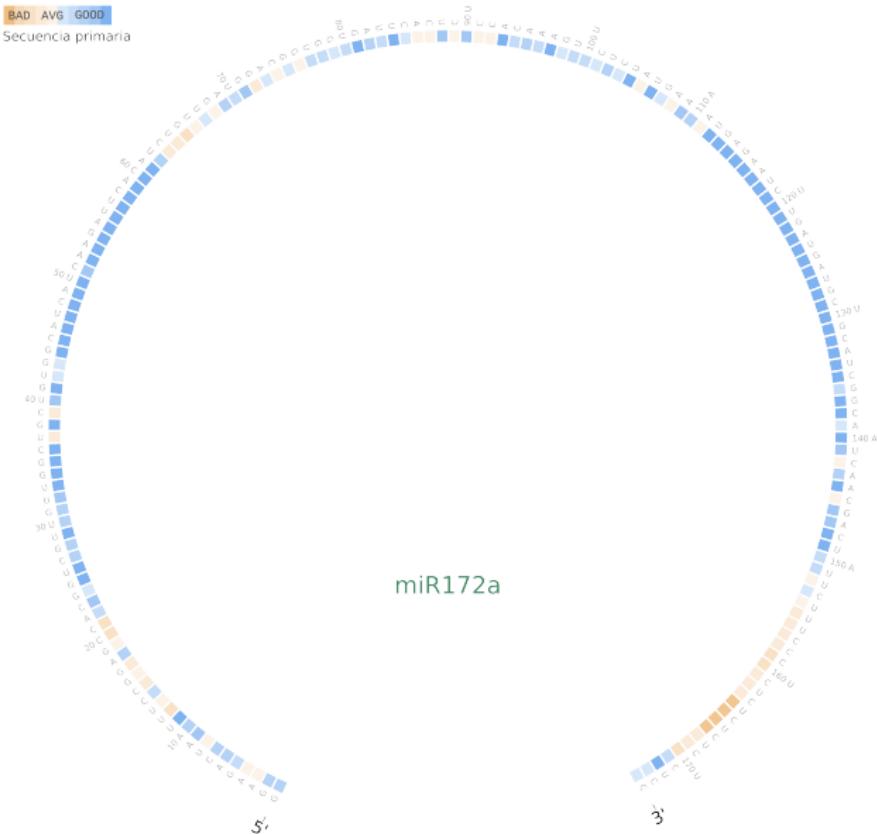
cons



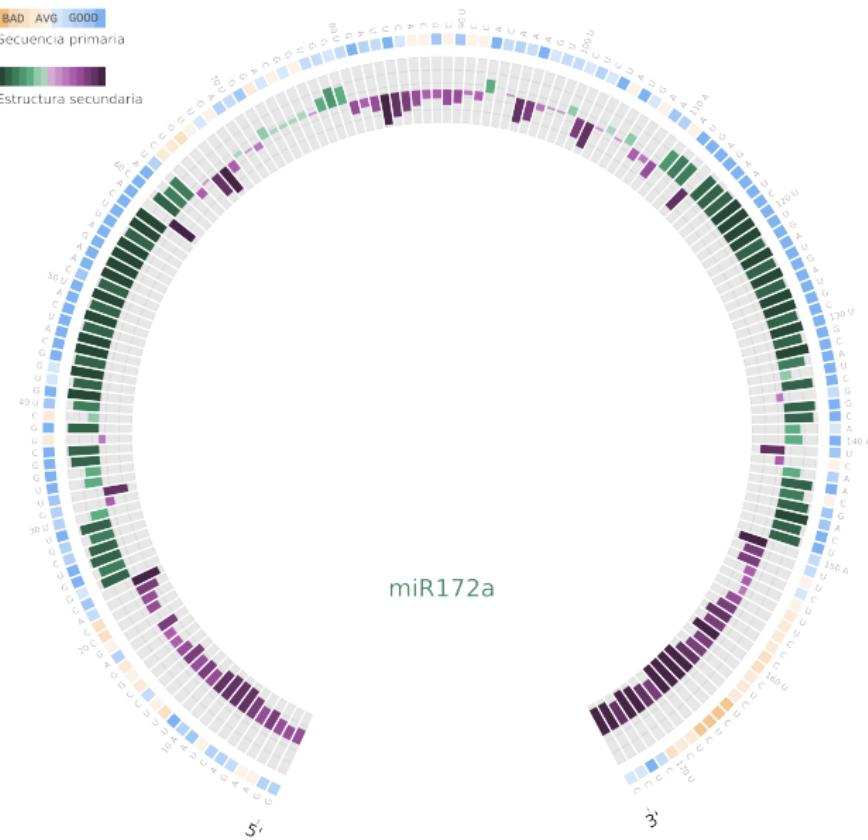
Cola de conservación hacia la izquierda del miARN y hacia la derecha del miARN*



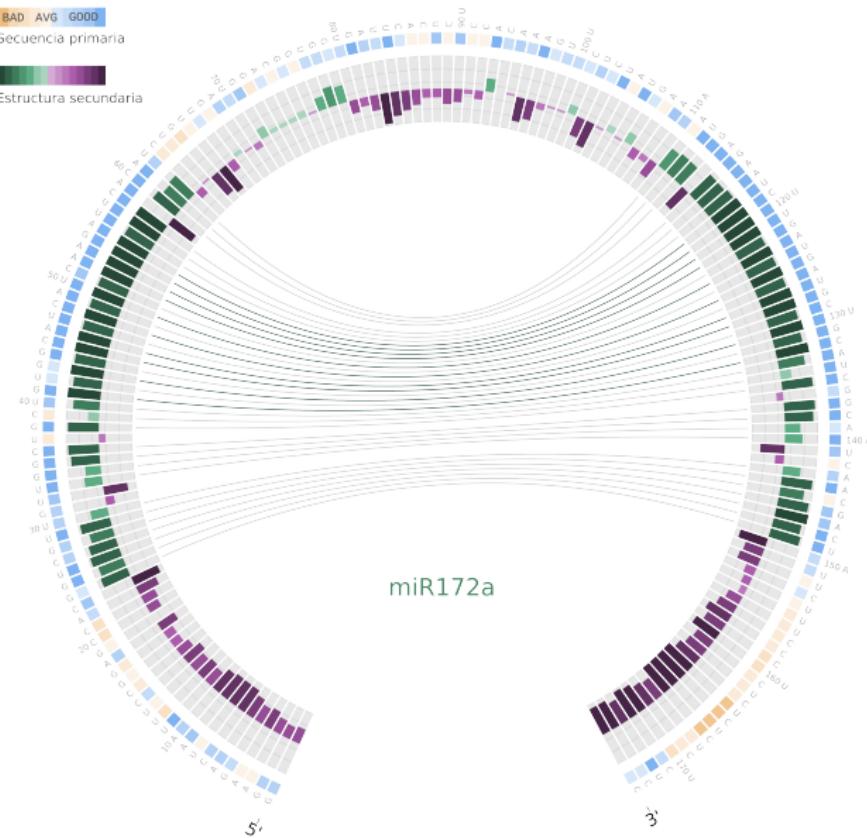
Conservación del consenso en base al alineamiento de secuencia primaria



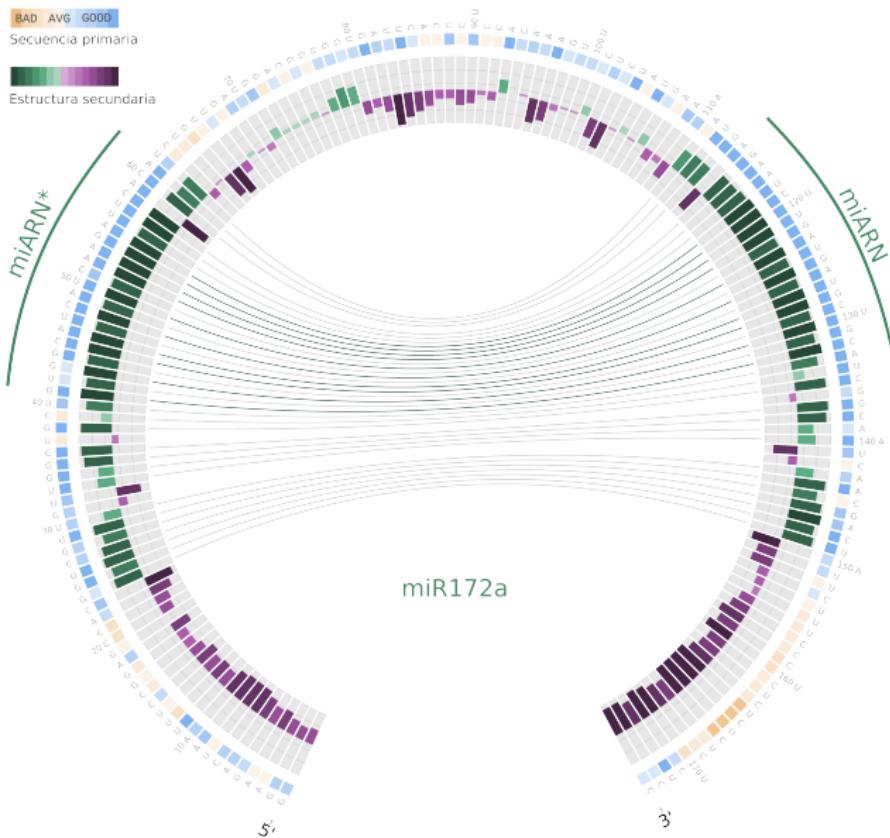
Frecuencia de bases apareadas y desapareadas



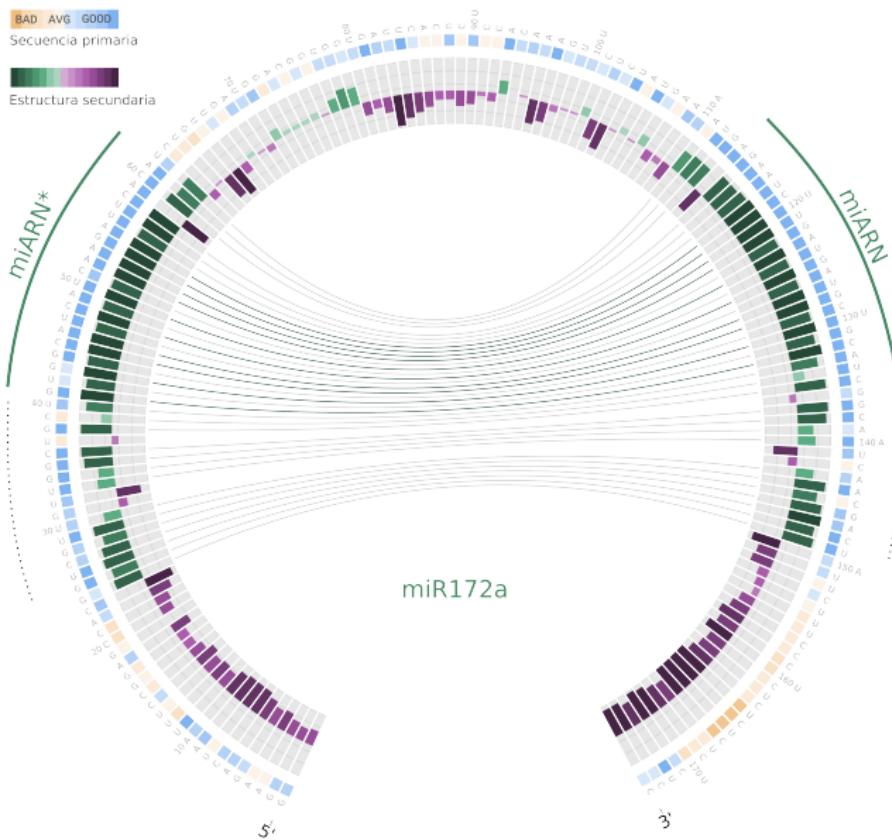
Interacción entre pares de bases considerando estructura secundaria



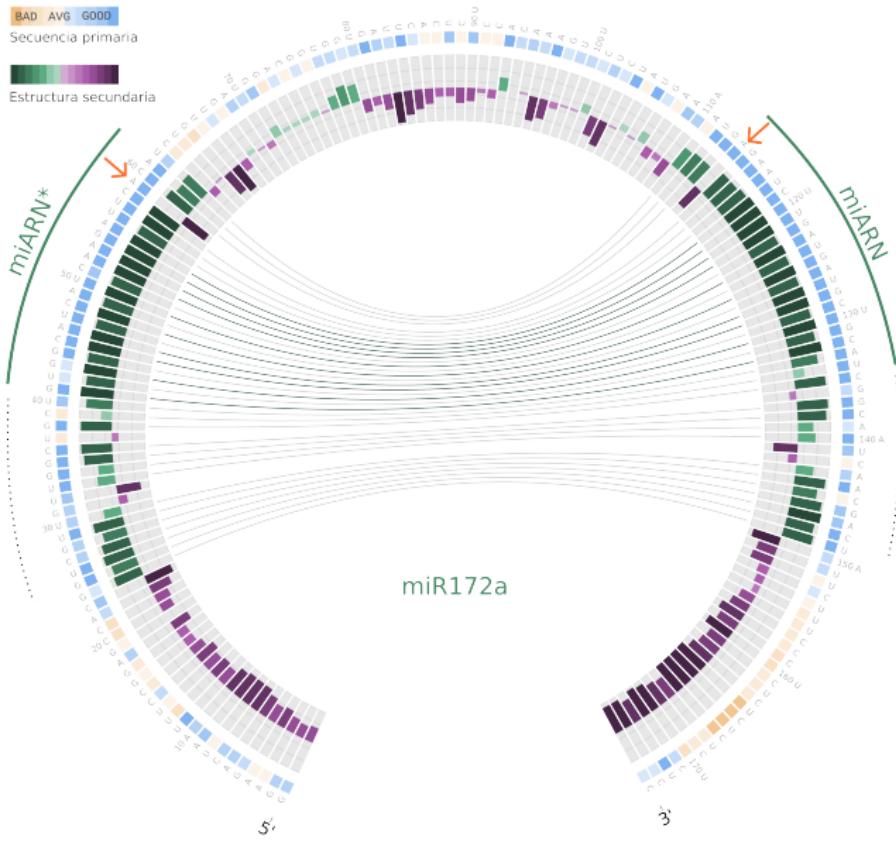
miARN y miARN* conservados en secuencia primaria y estructura



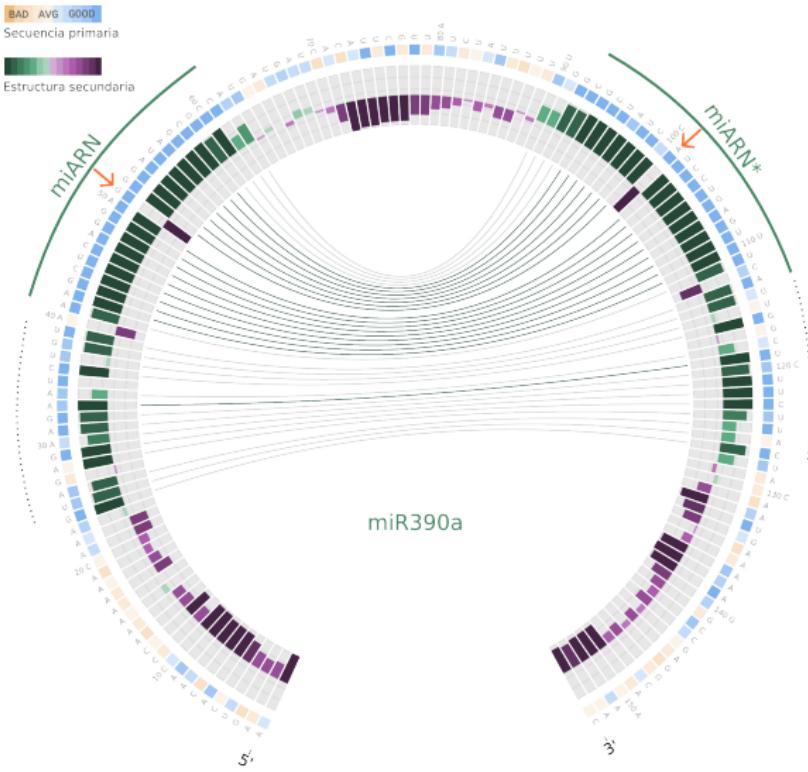
Región conservada por debajo del dúplex que coincide con el tallo inferior



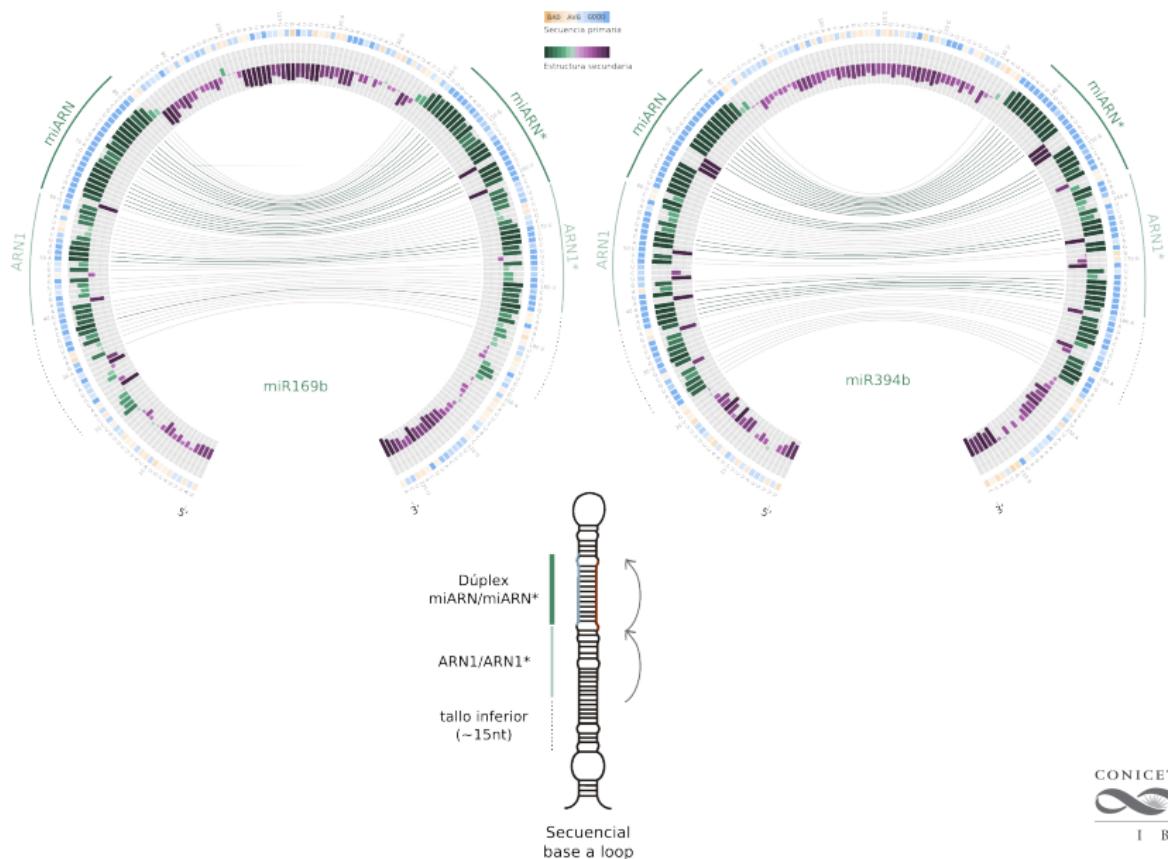
Mismatches conservados

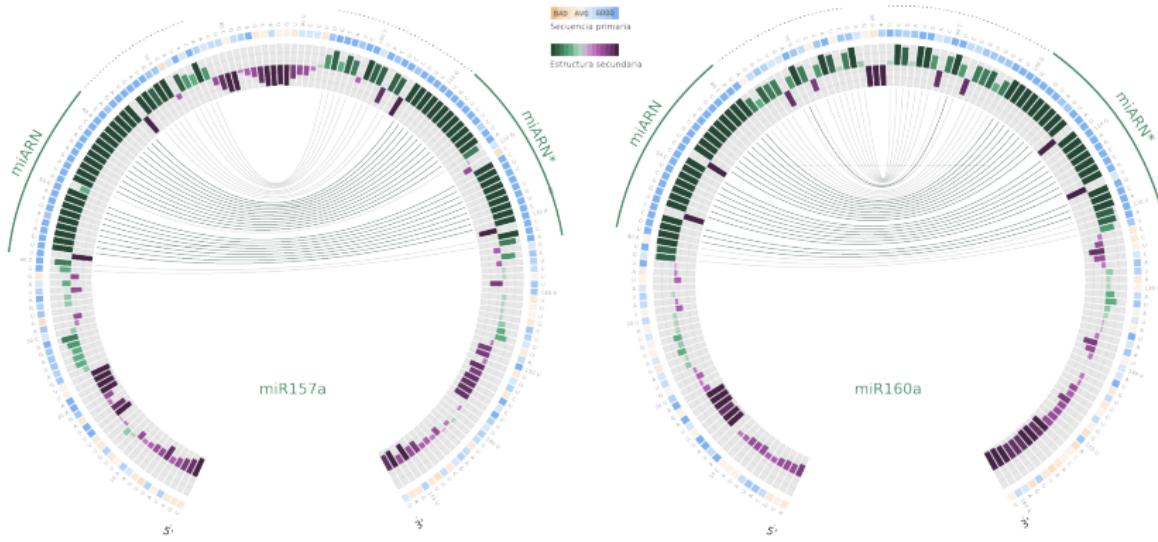


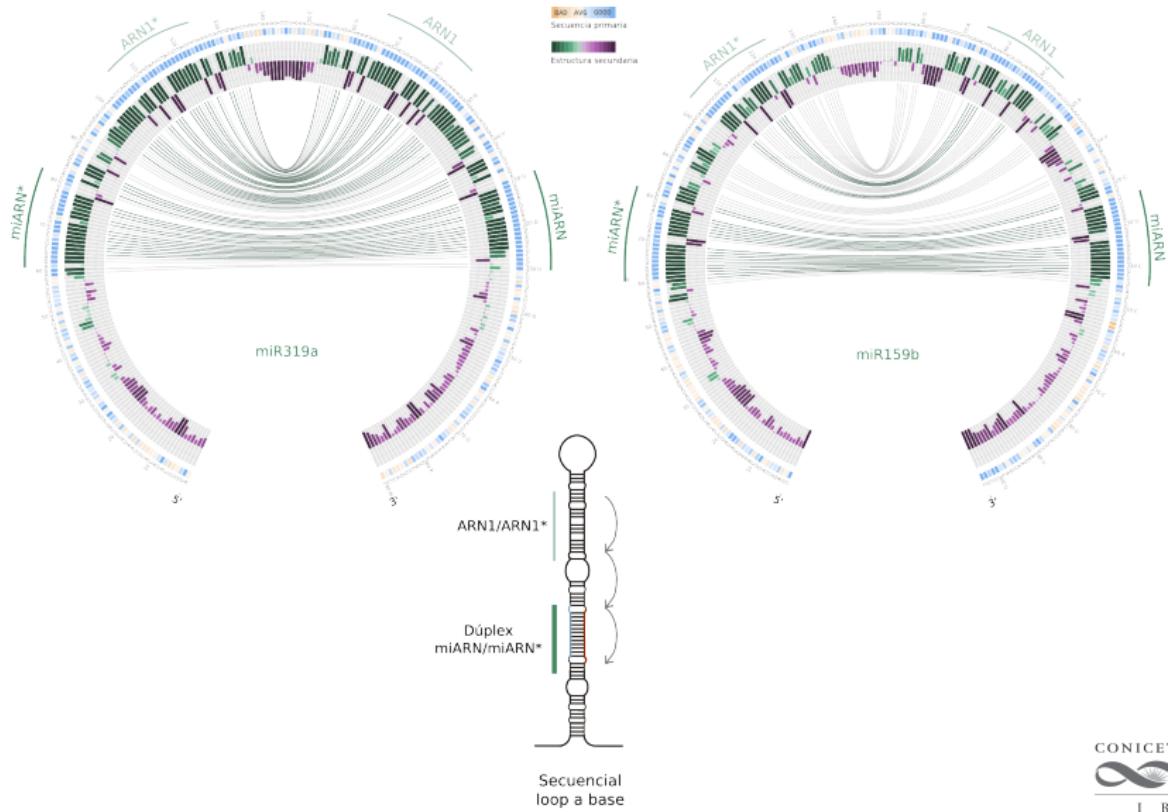
Mismo patrón de conservación en otros precursores que se procesan desde la base



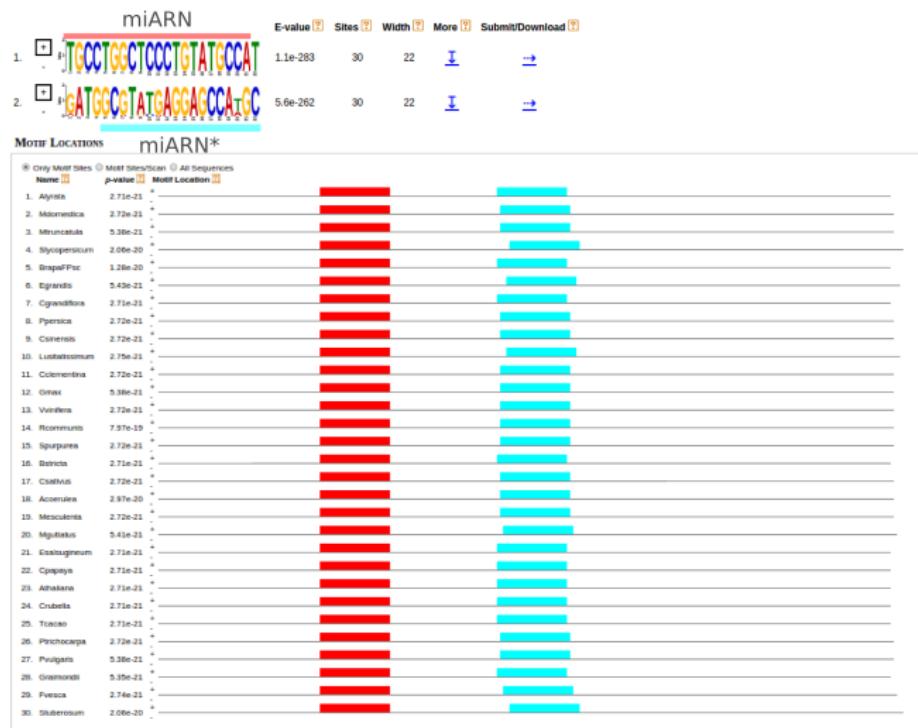
Precursos que se procesan desde la base en forma secuencial







En precursores que se procesan desde el loop, el tamaño del tallo superior el loop no varía en distintas especies



En precursores que se procesan desde la base, el tamaño del tallo superior y al loop es muy variado en distintas especies



Muchas gracias.

