

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

Lic. Uciel Chorostecki

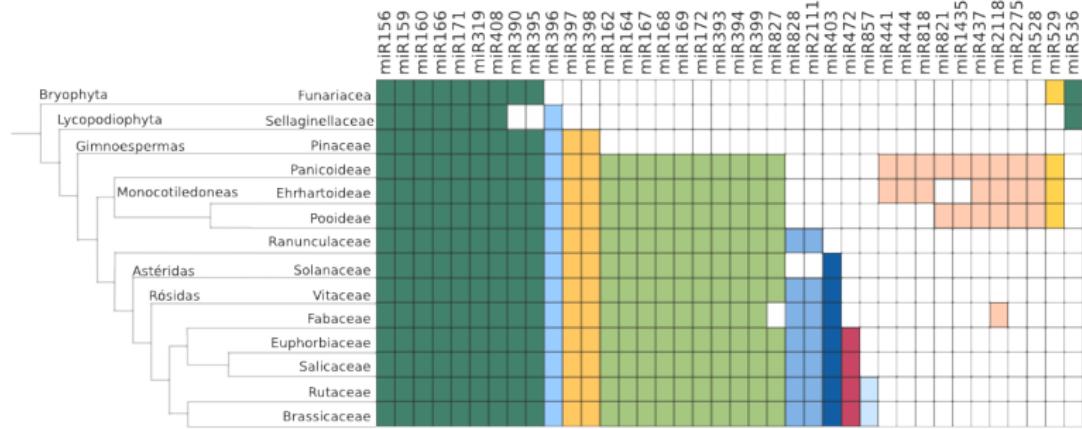
Director Dr. Javier Palatnik

Tesis de Doctorado
Facultad de Ciencias Bioquímicas y Farmacéuticas
Universidad Nacional de Rosario

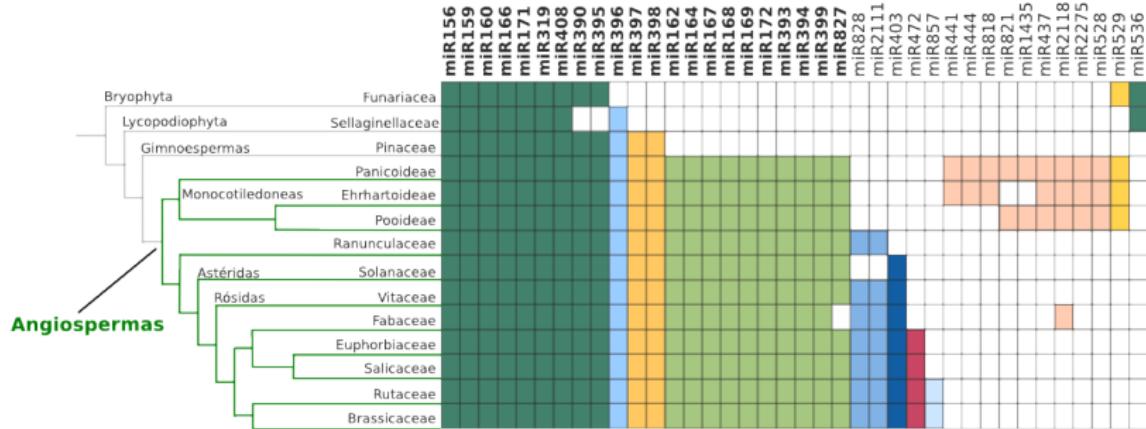
Los microARNs (miARNs) son ARN pequeños de 20-22 nt que regulan la expresión génica en animales y plantas.

- ▶ Controlan una gran variedad de procesos biológicos, como el desarrollo, la diferenciación y proliferación celular, y respuesta a estrés.
- ▶ Están codificados por familias de genes de 1 a 32 miembros que dan lugar a miARNs maduros idénticos o muy similares.
- ▶ Muchos de ellas han aparecido recientemente en la evolución y por lo tanto aparecen en un número pequeño de especies.

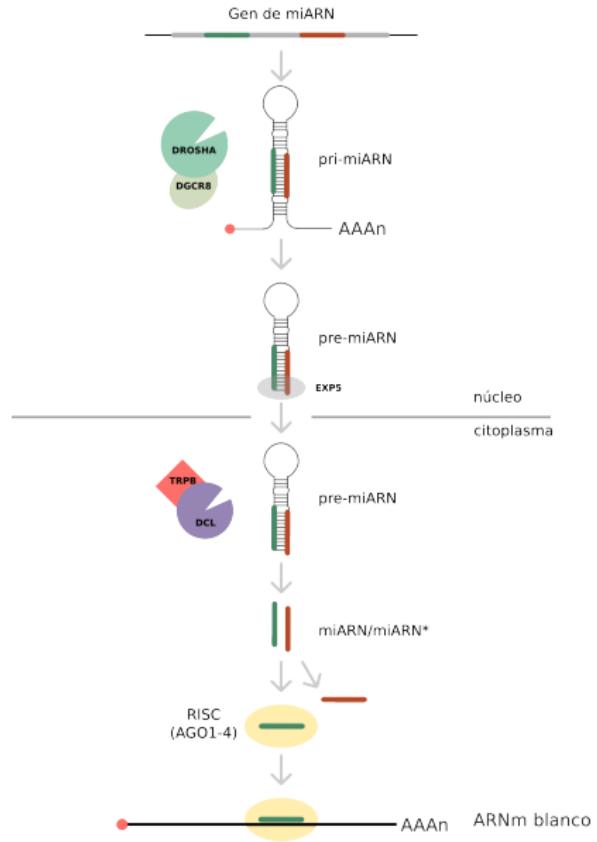
Existen familias de miARNs conservadas en las plantas



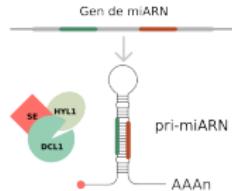
22 familias de miARNs que están altamente conservadas en Angiospermas



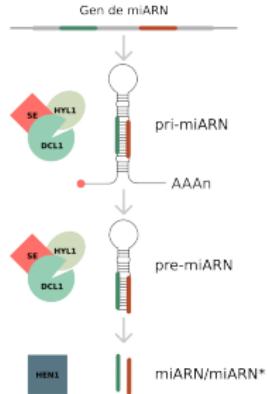
Procesamiento de miARNs en animales



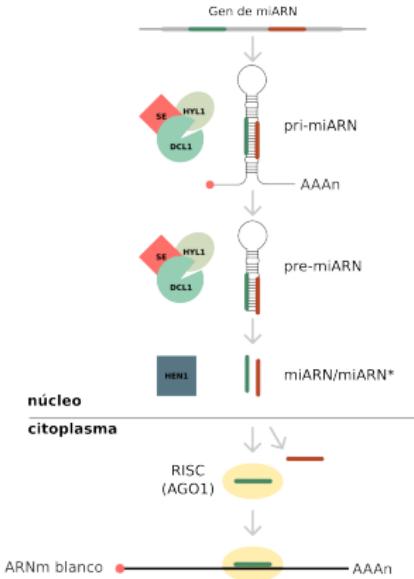
Biogénesis y actividad de miARNs en plantas



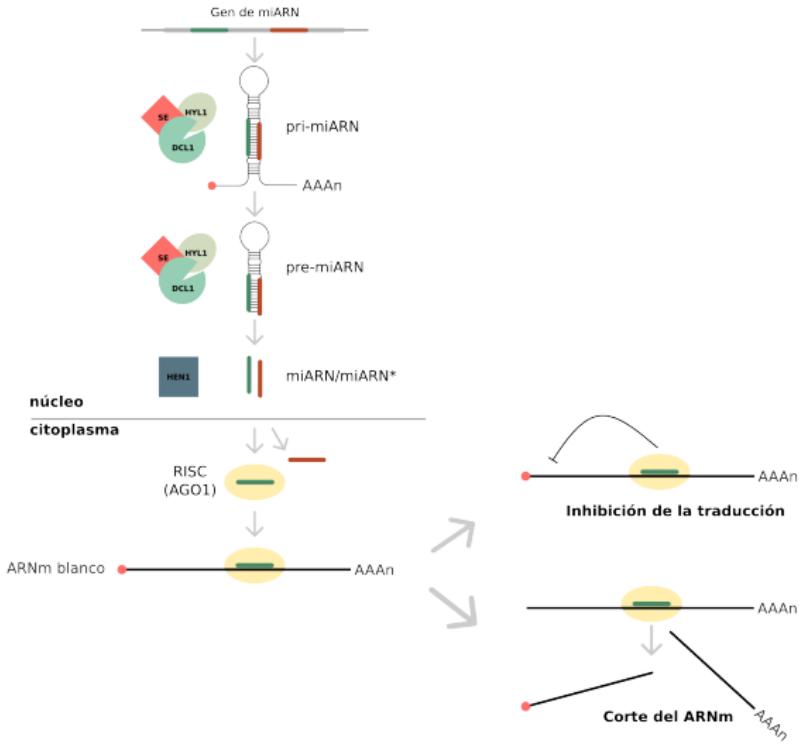
Biogénesis y actividad de miARNs en plantas



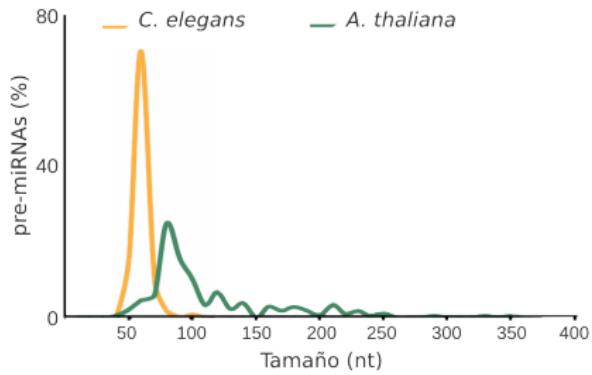
Biogénesis y actividad de miARNs en plantas



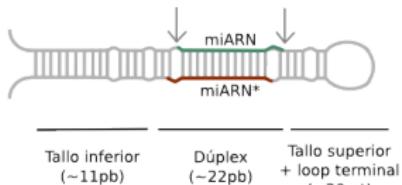
Biogénesis y actividad de miARNs en plantas



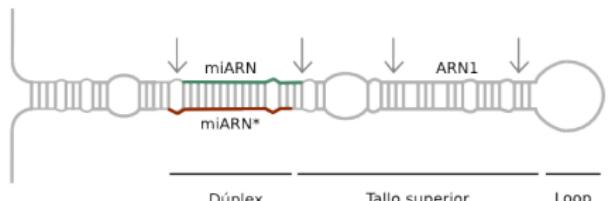
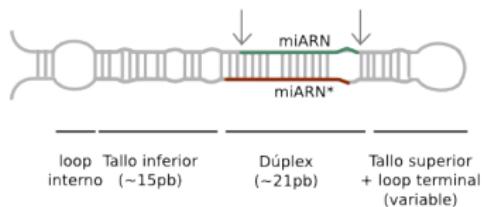
El tamaño de los precursores es muy variado en plantas



Estructuras secundarias de precursores de miARNs



pre-miARNs de animales



pre-miARNs de plantas

Regulación de la expresión génica por miARNs

Seres humanos

Genes: ~25.000

Arabidopsis thaliana

Genes: ~30.000

Regulación de la expresión génica por miARNs

Seres humanos

Genes: ~25.000

miARNS: ~500-1.000

Arabidopsis thaliana

Genes: ~30.000

miARNs: ~300

Regulación de la expresión génica por miARNs

Seres humanos

Genes: ~25.000

miARNS: ~500-1.000

Genes regulados por miARNs:
~40% genes (miles)

Arabidopsis thaliana

Genes: ~30.000

miARNs: ~300

Genes regulados por miARNs:
pocos (cientos)

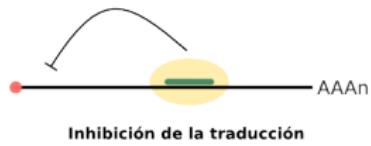
Regulación de la expresión génica por miARNs

Seres humanos

Genes: ~25.000

miARNS: ~500-1.000

Genes regulados por miARNs:
~40% genes (miles)



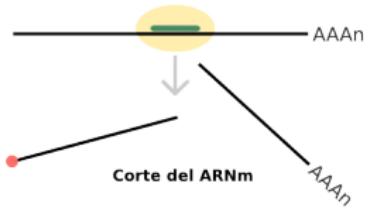
Inhibición de la traducción

Arabidopsis thaliana

Genes: ~30.000

miARNS: ~300

Genes regulados por miARNs:
pocos (cientos)

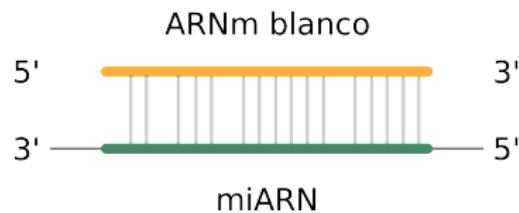


Corte del ARNm

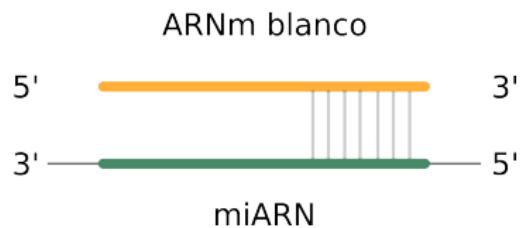
AAAn

Predicción de genes blancos de miARNs

Plantas



Animales



Diseñar estrategias computacionales para:

- ▶ identificar redes regulatorias de miARNs en plantas.
- ▶ comprender la biogénesis de los miARNs en plantas.

Diseñar estrategias computacionales para:

- ▶ identificar redes regulatorias de miARNs en plantas.
- ▶ comprender la biogénesis de los miARNs en plantas.

Objetivos específicos

- ▶ Diseñar una estrategia para la identificación de genes blancos regulados por miARNs.
- ▶ 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 para la identificación de genes blancos regulados por miARNs.
- ▶ 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.

- ▶ Estrategias computacionales tienen en cuenta la complementariedad con sus mensajeros blanco.
- ▶ Uno de los mayores desafíos es predecir los genes regulados por estos ARN pequeños con una baja frecuencia de predicciones falsas.

Variaciones en los núcletodos 1,20 y 21 dentro de cada familia de miARNs

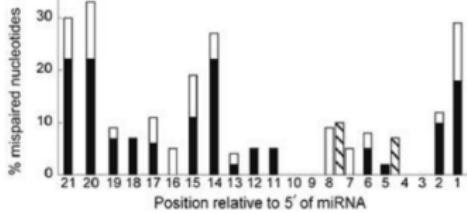
Especie	miR ID	Secuencia maduro
Arabidopsis thaliana	ath-miR159a	UUUGGAUUGAAGGGAGCUCUA
	ath-miR159b	UUUGGAUUGAAGGGAGCUCU U
	ath-miR159c	UUUGGAUUGAAGGGAGCUCU C
Carica papaya	cpa-miR159b	C UUGGAUUGAAGGGAGCUC U

Oryza sativa	osa-miR159b	UUUGGAUUGAAGGGAGCUC U
	osa-miR159c	A UUGGAUUGAAGGGAGCUC A
	osa-miR159f	C UUGGAUUGAAGGGAGCUCUA
Populus trichocarpa	ptc-miR159a	UUUGGAUUGAAGGGAGCUCUA
	sof-miR159a	UUUGGAUUGAAGGGAGCUC U
	Sorghum bicolor	S bi-miR159a
Vitis vinifera	vvi-miR159a	UUUGGAUUGAAGGGAGCUC U
	vvi-miR159c	C UUGGA G UGAAGGGAGCUC U
	Zea mays	UUUGGAUUGAAGGGAGCUC G

Mismatches en targets validados

Especie	miR ID	Secuencia maduro
Arabidopsis thaliana	ath-miR159a	UUUGGAUUGAAGGGAGCUCUA
	ath-miR159b	UUUGGAUUGAAGGGAGCUCU U
	ath-miR159c	UUUGGAUUGAAGGGAGCUCU C
Carica papaya	cpa-miR159b	C UUGGAUUGAAGGGAGCUCU C -

Oryza sativa	osa-miR159b	UUUGGAUUGAAGGGAGCUCU G
	osa-miR159c	A UUGGAUUGAAGGGAGCUCU A
	osa-miR159f	C UUGGAUUGAAGGGAGCUCUA
Populus trichocarpa	ptc-miR159a	UUUGGAUUGAAGGGAGCUCUA
	sof-miR159a	UUUGGAUUGAAGGGAGCUCU G
	sbi-miR159a	C UUGGA <u>G</u> UAGGGAGCUCU G
Sorghum bicolor	vvi-miR159a	C UUGGA <u>G</u> UAGGGAGCUCU C
	vvi-miR159c	UUUGGAUUGAAGGGAGCUCUA
	Zea mays	UUUGGAUUGAAGGGAGCUCU G

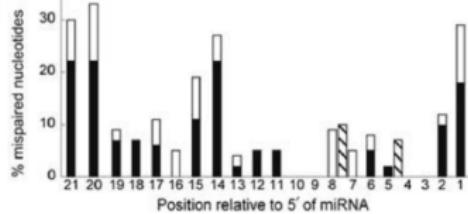


Mallory et al. EMBO J. 2004

Consenso

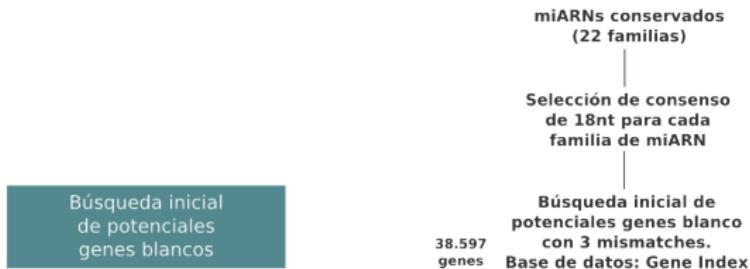
Especie	miR ID	Secuencia maduro
Arabidopsis thaliana	ath-miR159a	UUUGGAUUGAAGGGAGCUCUA
	ath-miR159b	UUUGGAUUGAAGGGAGCUCU U
	ath-miR159c	UUUGGAUUGAAGGGAGCUCU C
Carica papaya	cpa-miR159b	C UUGGAUUGAAGGGAGCUCU C -

Oryza sativa	osa-miR159b	UUUGGAUUGAAGGGAGCUCUG
	osa-miR159c	A UUGGAUUGAAGGGAGCUCCA
	osa-miR159f	C UUGGAUUGAAGGGAGCUCUA
Populus trichocarpa	ptc-miR159a	UUUGGAUUGAAGGGAGCUCUA
	sof-miR159a	UUUGGAUUGAAGGGAGCUCUG
	sbi-miR159a	UUUGGAUUGAAGGGAGCUCUG
Sorghum bicolor	vvi-miR159a	C UUGGA <u>GUAGGGAGCUCUC</u>
	vvi-miR159c	UUUGGAUUGAAGGGAGCUCUA
	Zea mays	UUUGGAUUGAAGGGAGCUCUG
consenso	miR159	<u>UUGGAUUGAAGGGAGCUC</u>

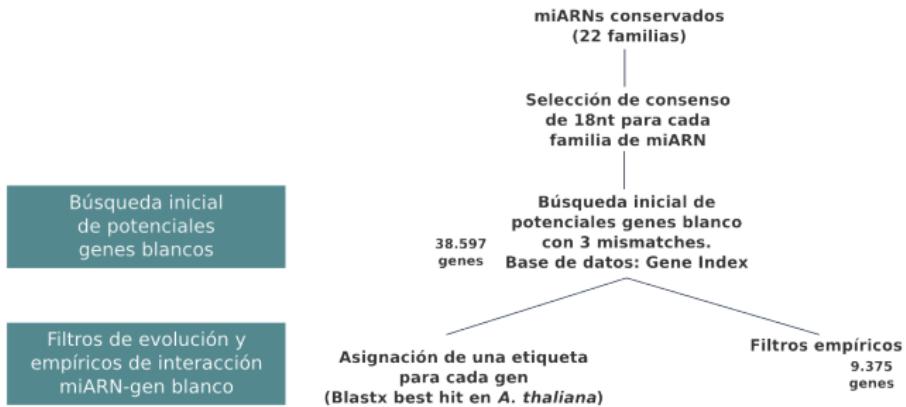


Mallory et al. EMBO J. 2004

Esquema de la estrategia para la identificación de nuevos genes blancos



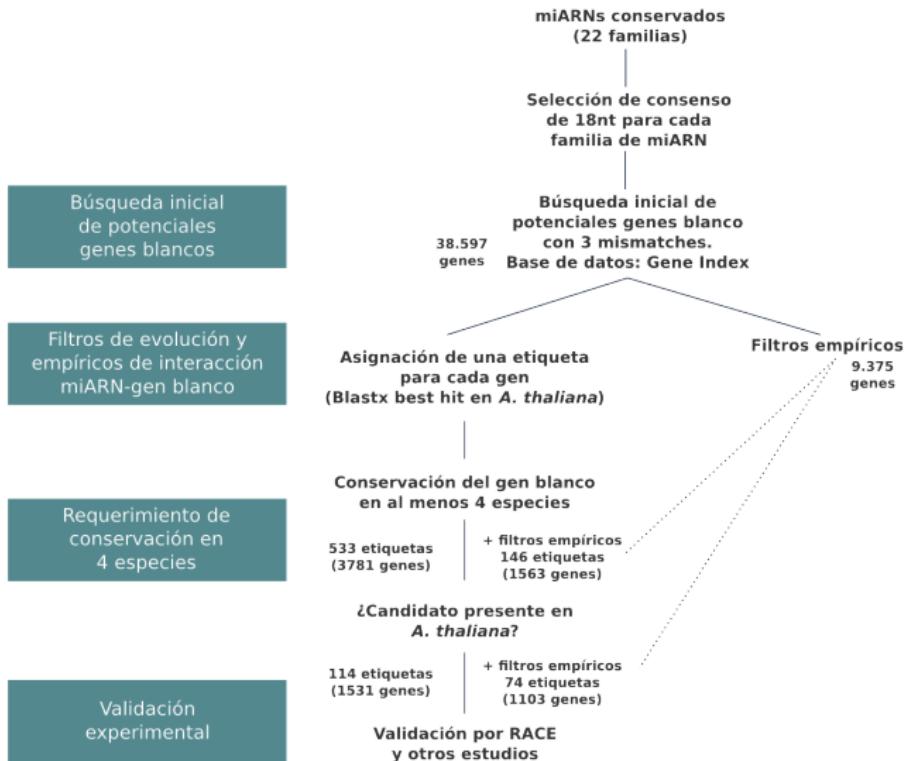
Esquema de la estrategia para la identificación de nuevos genes blancos



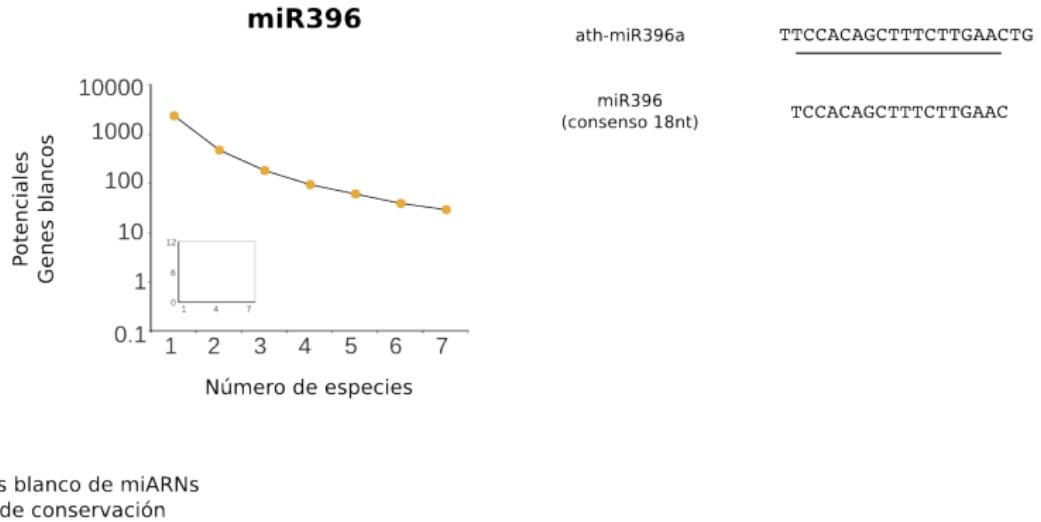
Esquema de la estrategia para la identificación de nuevos genes blancos



Esquema de la estrategia para la identificación de nuevos genes blancos



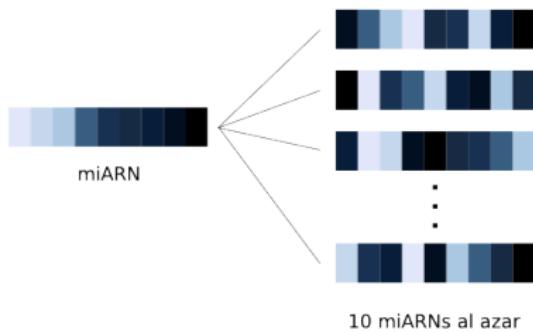
Conservación de la interacción en distintas especies



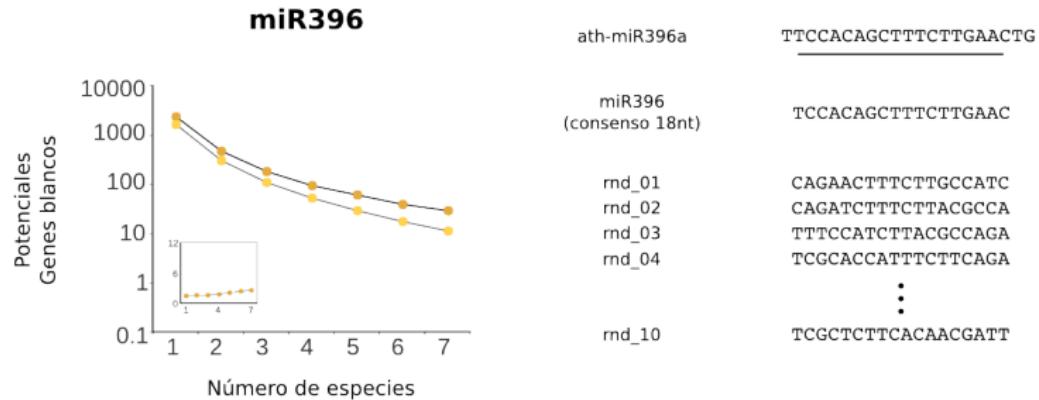
miARN consenso de 18 nt



Control: miARN al azar



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



- genes blanco de miARNs
filtro de conservación
- genes blanco de miARNs al azar
filtro de conservación

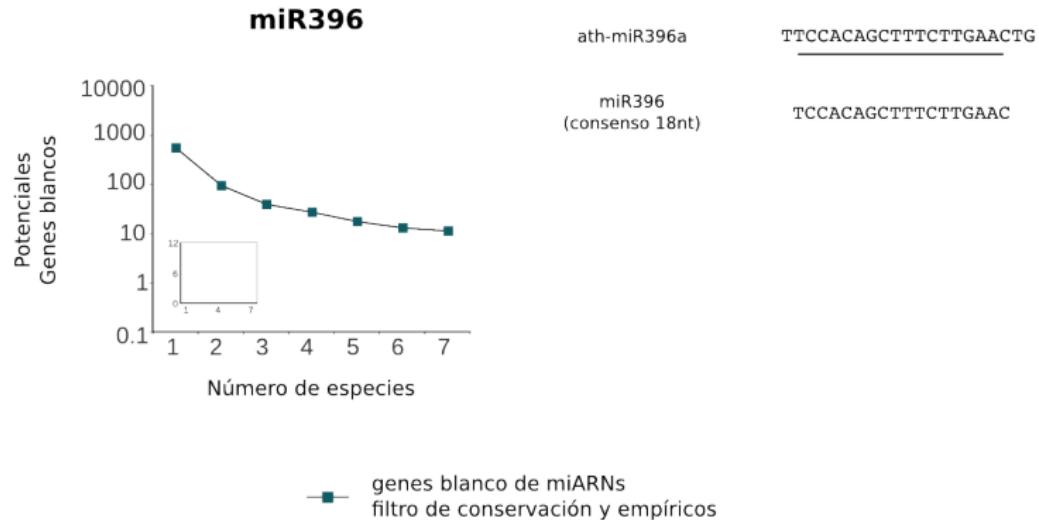
Parámetros empíricos deducidos de interacciones miARN-gen blanco validadas experimentalmente.



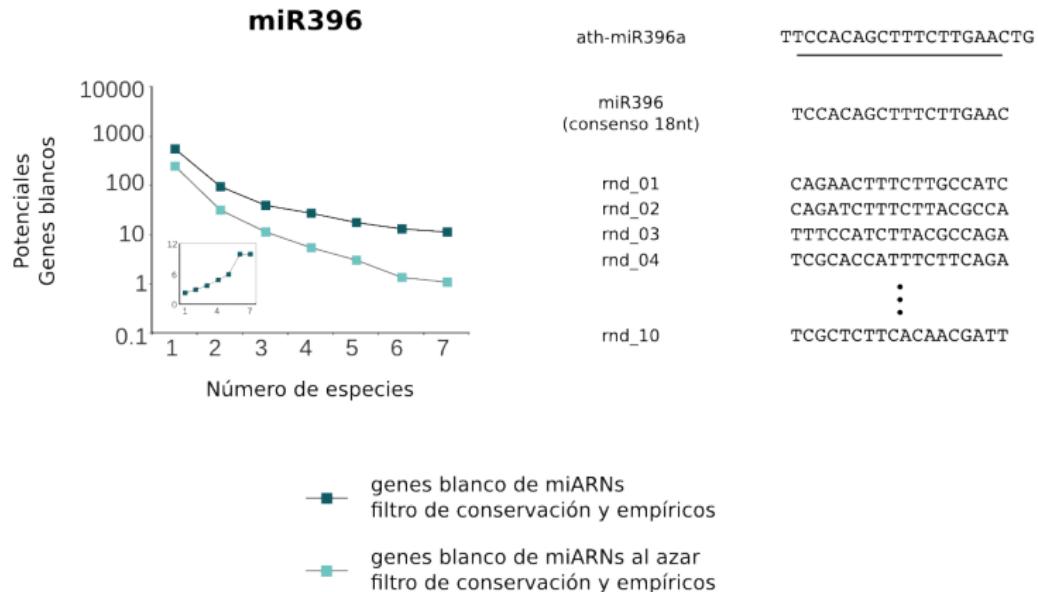
Parámetros empíricos deducidos de interacciones miARN-gen blanco validadas experimentalmente. Filtros aplicados en nuestra estrategia.



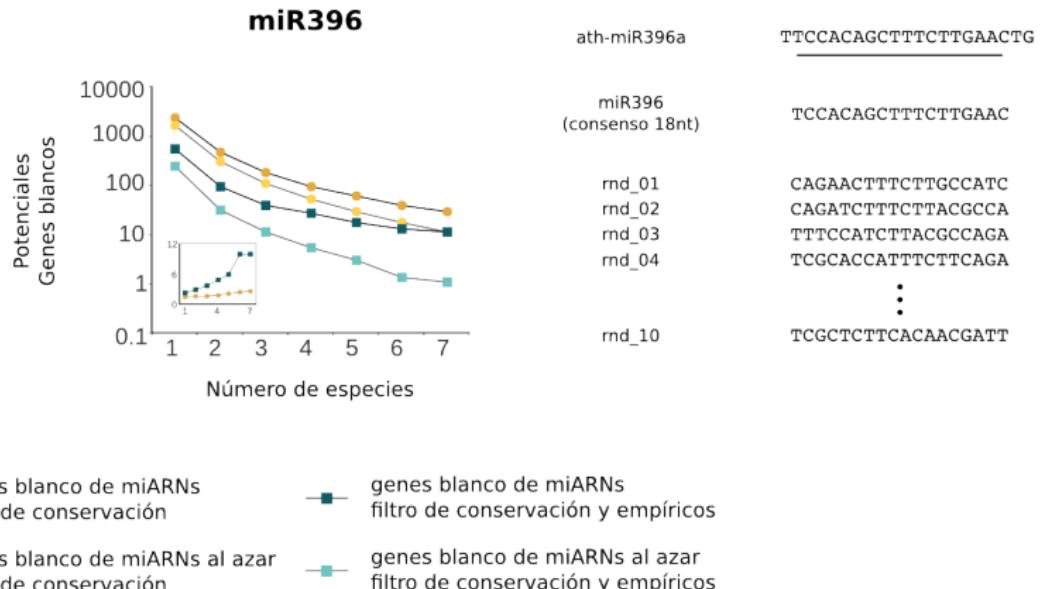
Slección de candidatos teniendo en cuenta los filtros empíricos



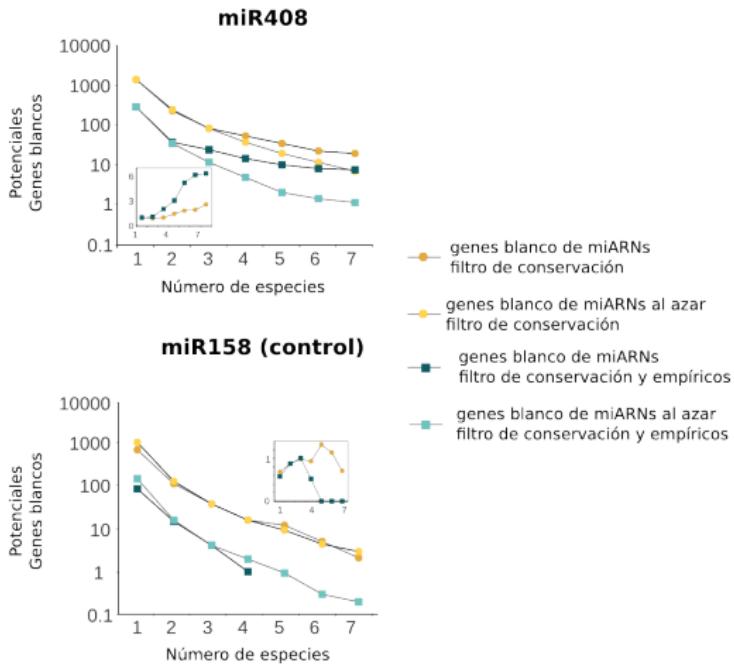
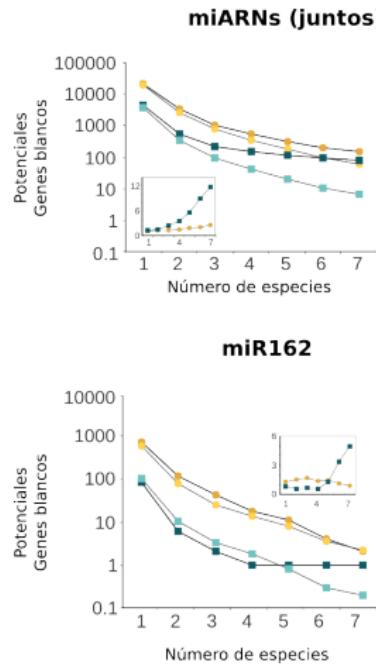
Al aplicar filtros empíricos y de conservación la relación señal/ruido aumenta



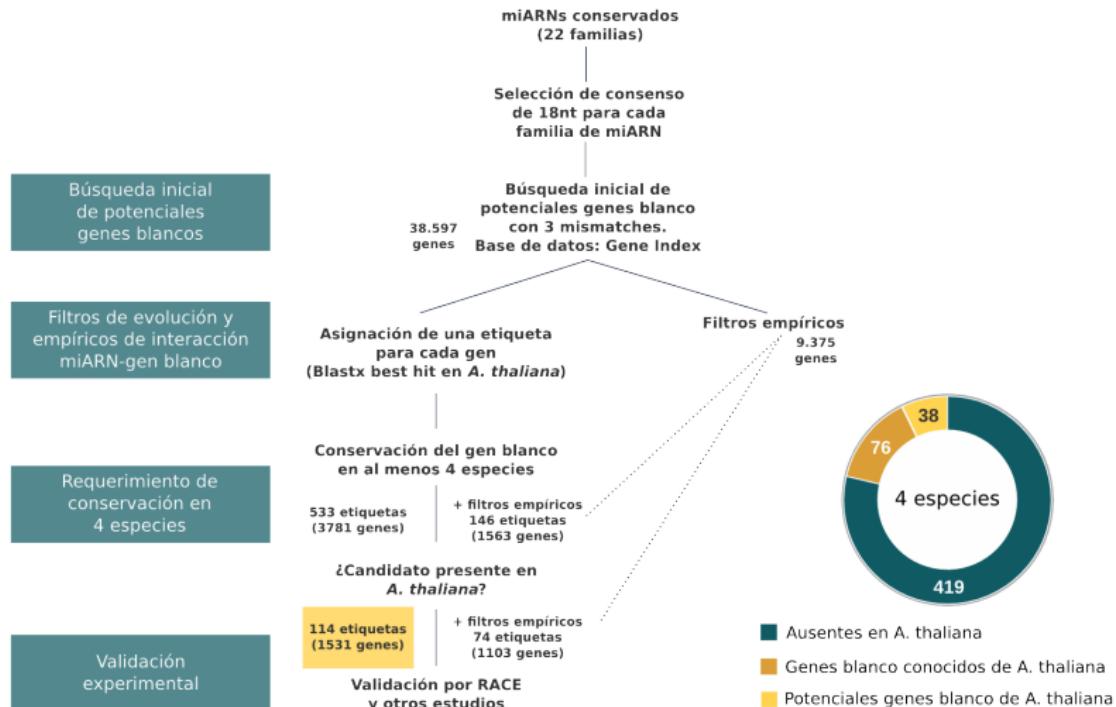
Efecto sinérgico al combinar filtro de conservación evolutiva y empíricos



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

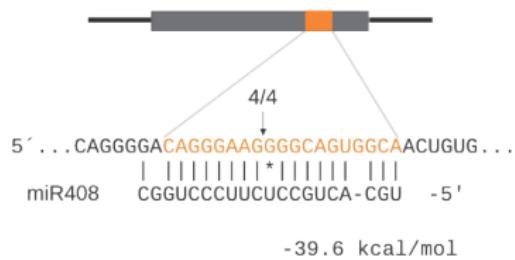


Potenciales genes blancos utilizando solo conservación evolutiva



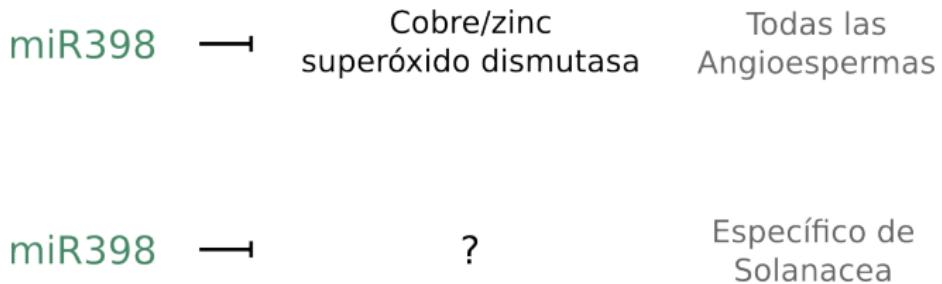
Se validaron 6 nuevos genes blancos en *A. thaliana*

PAA2 (At5g21930)



<i>Arabidopsis thaliana</i>	CAGGGAAAGGGGCAGUGGCA
<i>Capsicum annuum</i>	CAGCGAAGAACAGUUGCA
<i>Glycine max</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Gossypium raimondii</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Haseolus vulgaris</i>	UAGGGAAAGAGGGCAGUUGCA
<i>Ipomoea nil</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Lactuca sativa</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Medicago truncatula</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Saccharum officinarum</i>	UAGGGAAAGAACAGUGACA
<i>Solanum lycopersicum</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Solanum tuberosum</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Sorghum bicolor</i>	UAGGGAAAGAACAGUGACA
<i>Theobroma cacao</i>	CAGGGAAAGAGGGCAGUUGCA
<i>Vitis vinifera</i>	CAGGGAAAGAGGGCAGUUGCA

¿Pueden los miARNs en Angiospermas regular genes específicos de Solanaceae?



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 Solanaceae?
41 genes

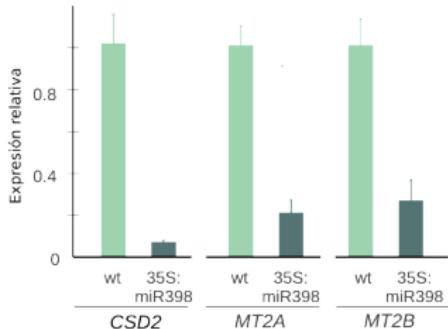
MT2A



MT2B



Validación experimental



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

Herramienta web para la predicción de genes blancos regulados por miARNs en plantas

Desarrollamos ComTAR, que permite realizar la búsqueda de:

Desarrollamos ComTAR, que 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.

Herramienta web para la predicción de genes blancos regulados por miARNs en plantas

Desarrollamos ComTAR, que 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



conserved plant miRNA target prediction tool

Find potential
miRNA targets

Find potential
miRNA target
families

Is this gene a
potential
miRNA target?

Targets

miRNA: miR398_GTTGTCAGGTACCCCC

MFE cutoff: -29.232 kcal/mol

MM Filter: Yes

Species: All

Go Back

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
AT1G80230	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 [?]
Aqua_013_00504.1	Aquilegia coerulea	TGGGCGACCTGGGAACAT * * * * CCCCACTGGACTCTTG	-31.7
471402	Arabidopsis lyrata	TGGGAGACCTGGGAACAC * * * * CCCCACTGGACTCTTG	-32.1
AT1G12520.1	Arabidopsis thaliana	TGGGAGACCTGGGAACAC * * * * CCCCACTGGACTCTTG	-32.1
Bradi5g18900.3	Brachypodium distachyon	TTGGTGACCTGGGAACGC ** * * * CCCCACTGGACTCTTG	-33.5
Bra026968	Brassica rapa	TGGGCGACCTGGGAACAC * * * * CCCCACTGGACTCTTG	-32.5
Carubv10011816m	Capsella rubella	TGGGAGACCTGGGAACAC * * * * CCCCACTGGACTCTTG	-32.1
evm.model.supercontig_29.47	Carica papaya	TAGGTGACCTGAGAACAT ** * * CCCCACTGGACTCTTG	-34.2
Ciclev10021134m	Citrus clementina	TTGGTGACCTGGGAACAC ** * * * CCCCACTGGACTCTTG	-33.9
orange1.1g020436m	Citrus sinensis	TTGGTGACCTGGGAACAg ** * * * CCCCACTGGACTCTTG	-32.7

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

Conclusiones I

- ▶ 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.

miARN	Consenso	Targets conocidos Validados por nosotros
mir156	GACAGAAAGAGAGTGAGCA	fatores de transcripción SPL
mir159	TTCGATTTGAAGGGAGCTC	fatores de transcripción MYB, NOZZLE (NZL)
mir160	GCCTGGCTCCCTGTATGC	fatores de transcripción ARF
mir162	CGATAAACCTCTGATCC	DCL1
mir164	GGAGAGACCAAGGGCACGTC	fatores de transcripción NAC
mir166	CGGACCAAGGCCATTATCC	fatores de transcripción HDZip
mir167	GAAGCTGCCAGCAATGATC	fatores de transcripción ARF, IAA-ALANINE RESISTANT 3 (IAR3)
mir168	CGCTTGGTCCAGGTGGGG	AGO1
mir169	AGCCAAGAGTAGCTTGCC	fatores de transcripción CCAAT-HAP2
mir171	TTGACCCGTGCCAATATC	fatores de transcripción GRAS
mir172	GAATCTTGATGATCTGC	fatores de transcripción AP2
mir319	TGGACTGAAGGGAGCTCC	fatores de transcripción TCP
mir390	AGCTCAGGAGGGATAGCG	TAS RNA
mir393	CCAANGGGATCGCATTTGA	TIR1 proteins, F-BOX proteins
mir394	TGGCATTCTGTCCACCTC	proteínas F-BOX
mir395	TGAAGTGTGTTGGGGGAC	ATP-sulfurilasas, transportadores de sulfato
mir396	TCCACACGTTCTTGAC	fatores de transcripción GRF, MMG4.7, FLUORESCENT IN BLUE LIGHT (FLU)
mir397	CATTGACTGCAGCTTGA	Laccases
mir398	GTGTTTCAGGTCAACCCC	Cu/Zn SODs, Cyt C oxidase protein subunit, Chaperona de cobre (CCS)
mir399	GCCAAGGGAGATTGGCCC	Enzima E2 de conjugación de ubiquitina
mir408	TGCACTGCCCTTCCCCG	Blue copper proteins, Laccases, P-TYPE ATPase (PAA2), PAC1
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.
- ▶ 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.

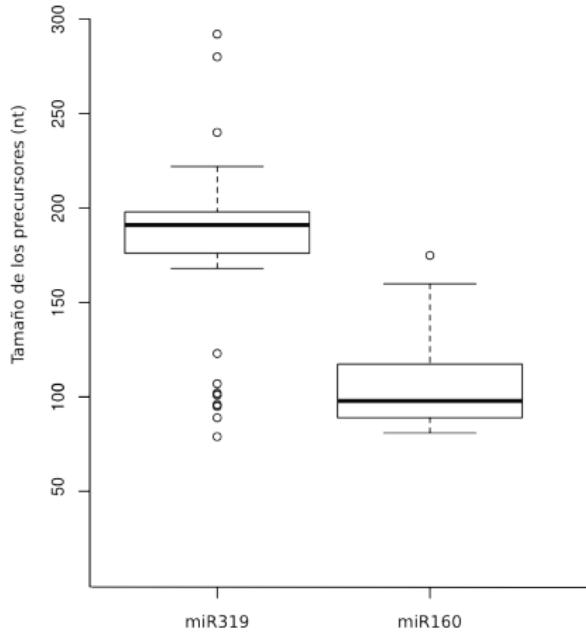
Objetivos específicos

- ▶ Diseñar una estrategia para la identificación de genes blancos regulados por miARNs.
- ▶ 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.

Objetivos específicos

- ▶ Diseñar una estrategia para la identificación de genes blancos regulados por miARNs.
- ▶ 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.

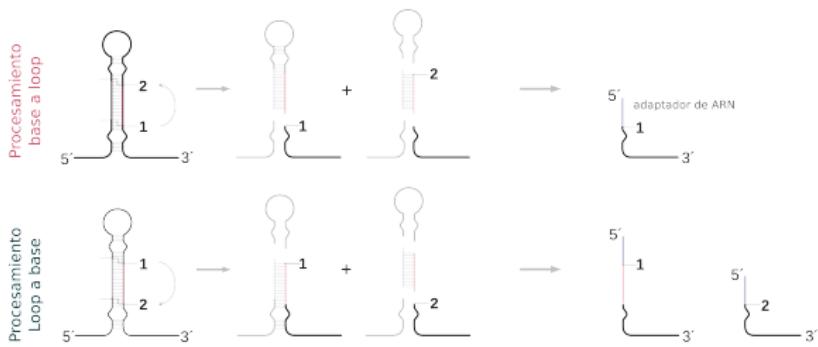
Precursos en plantas son muy variables en tamaño y forma



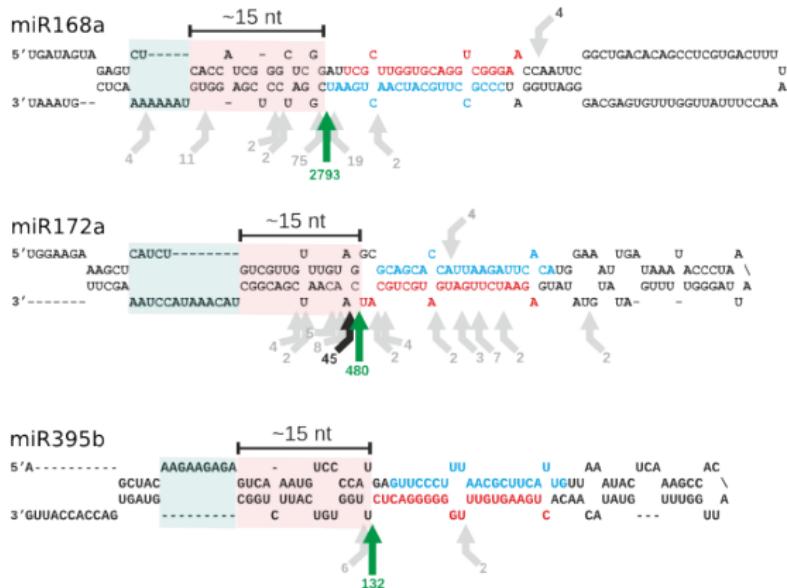
Bibliotecas SPARE para estudios genómicos de biogénesis de miARNs en plantas



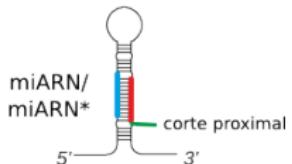
Bibliotecas SPARE para estudios genómicos de biogénesis de miARNs en plantas



Los precursores que se procesan desde la base tienen un sólo pico de señal en las bibliotecas de SPARE



Detección solamente del corte proximal

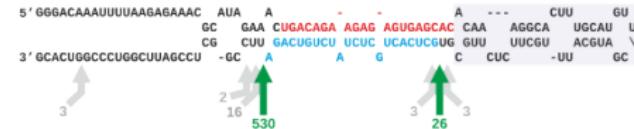


Los precursores que se procesan desde el loop tienen al menos dos picos de señal en las bibliotecas de SPARE

miR156a



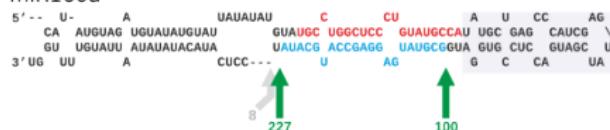
miR156c



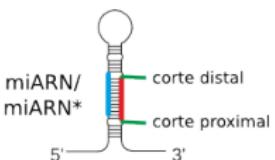
miR156d



miR160a



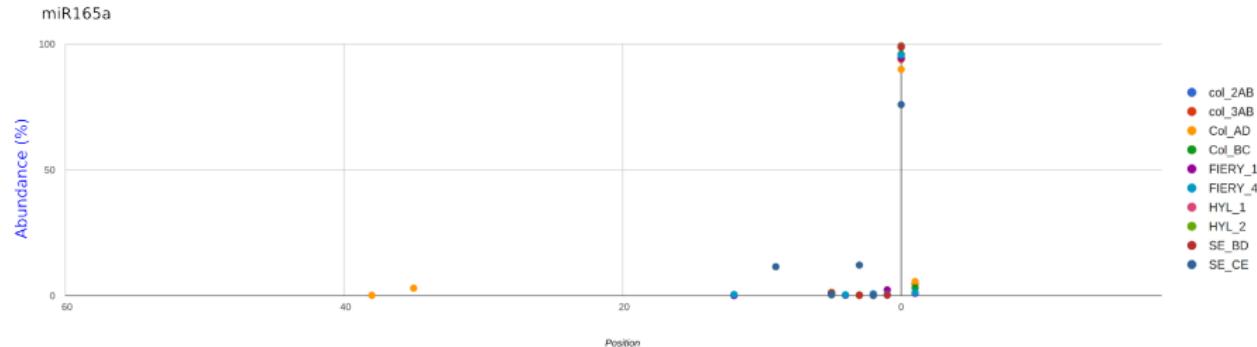
Detección de ambos cortes, proximal y distal



Bibliotecas de SPARE

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

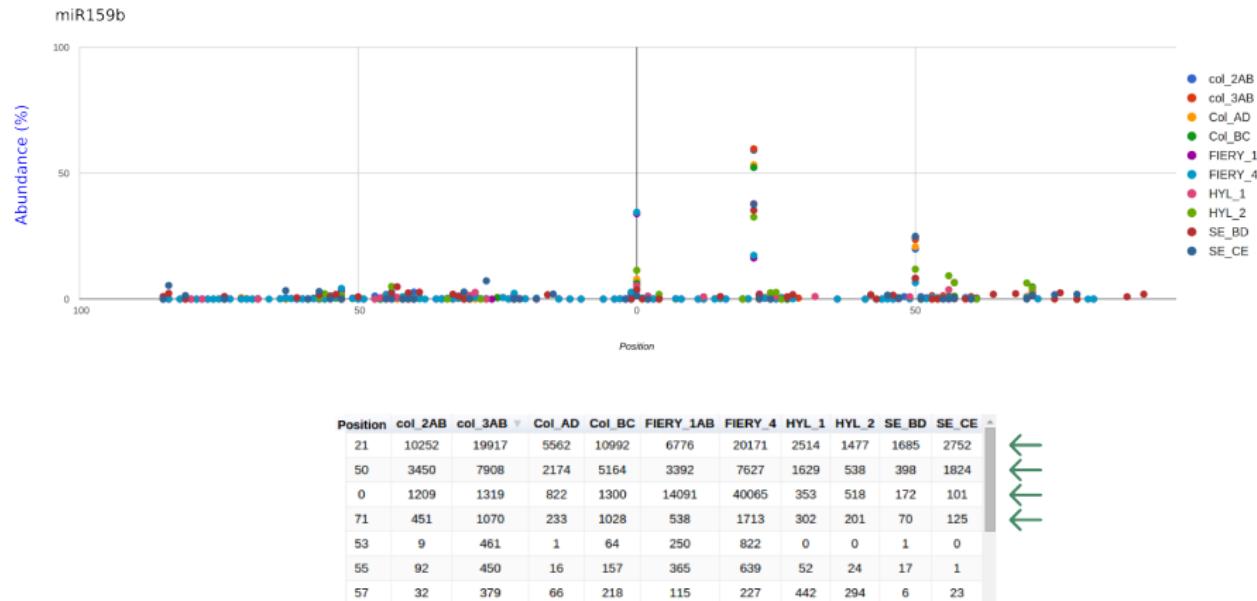
Visualización de precursores procesados desde la base



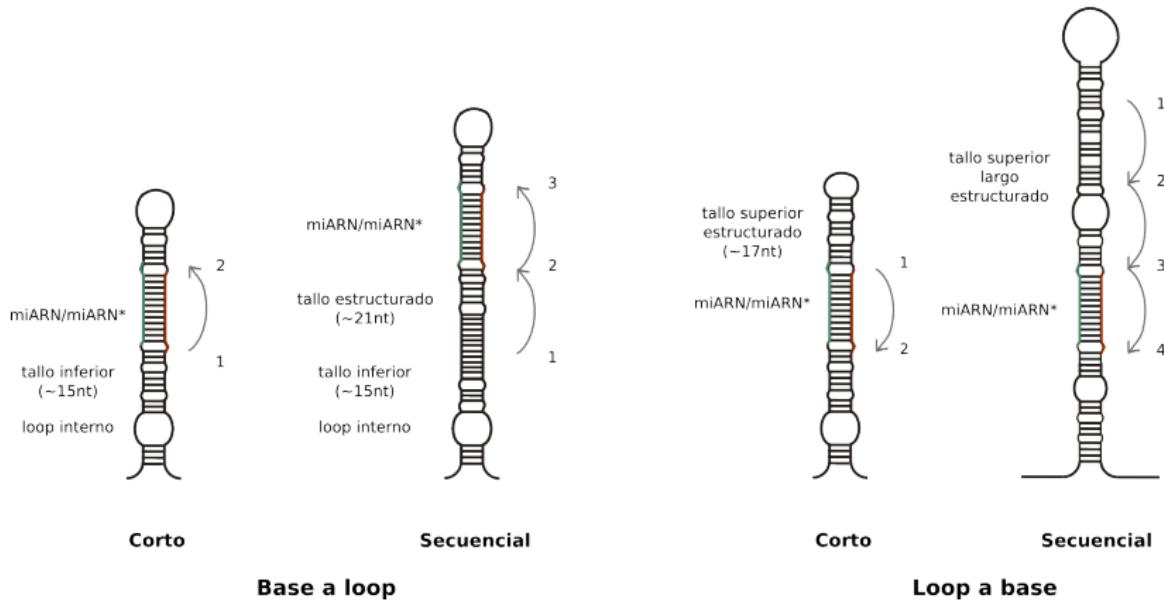
Position	col_2AB	col_3AB	Col_AD	Col_BC	FIERY_1AB	FIERY_4	HYL_1	HYL_2	SE_BD	SE_CE
-38	0	0	4	0	0	0	0	0	0	0
-35	0	0	89	0	0	0	0	0	0	0
-12	1	0	0	0	10	147	0	0	0	0
-9	0	0	0	0	0	0	0	0	0	270
-5	29	67	40	26	116	326	2	1	5	9
-4	0	0	0	0	15	111	0	0	0	0
-3	0	0	1	0	0	0	0	1	1	285
-2	0	0	0	0	77	189	0	0	0	1
-1	0	0	0	13	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



Visualización de precursores procesados desde el loop



Conclusiones II

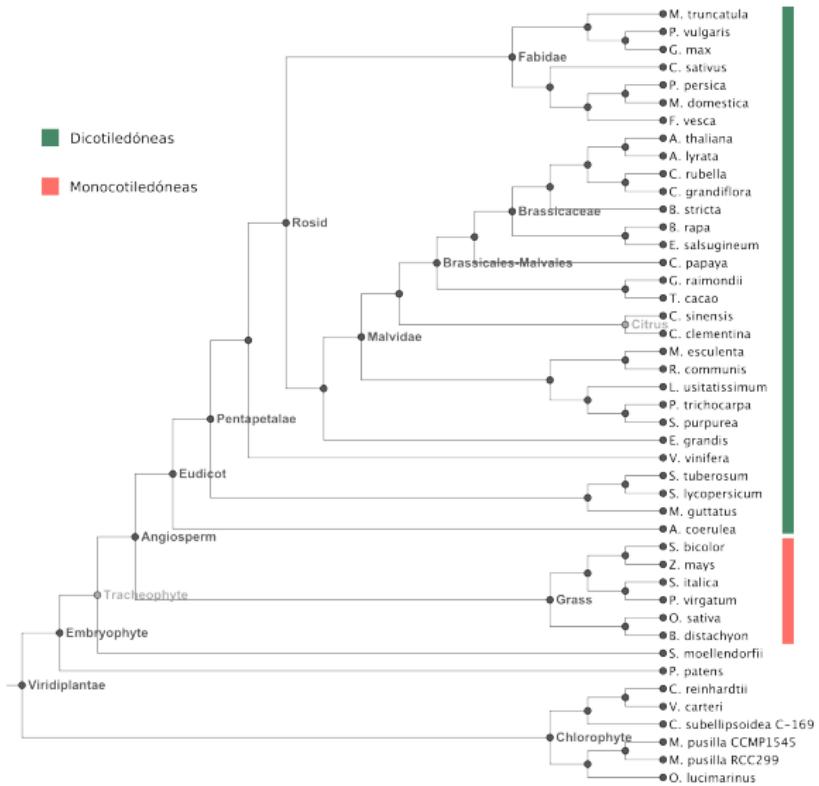


Objetivos específicos

- ▶ Diseñar una estrategia para la identificación de genes blancos regulados por miARNs.
- ▶ 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 para la identificación de genes blancos regulados por miARNs.
- ▶ 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.

Especies utilizadas de Phytozome (30 dicotiledóneas y 6 monocotiledóneas)



Precursors con distintos mecanismos de procesamiento deberían tener elementos estructurales distintos.

Precursores con distintos mecanismos de procesamiento deberían tener elementos estructurales distintos.

Anotación arbitraria en miRBASE.

- ▶ Longitud de precursores.
- ▶ Definición de ortólogos

¿Cuál es el ortólogo en otras especies?

Arabidopsis

>**ath-MIR160a**
GUUAGCCUGGUCCCCUGUAUGCCAUUAGCUGAGCCCCAUUCGAGUAUCAUGACCUCCGUGGAUGGGCGUAUGAGGAGCCAUGCAU
>**ath-MIR160b**
GUCGUUGCCUGGUCCCCUGUAUGCCACAAAGAAAACAUCGAUUUAGUUCACAUAGUGGCCUACAGAGUAUGCAAGCAUGAC
>**ath-MIR160c**
GUUAUJGCCUGGUCCCCUGUAUGCCACAGAGUACCGAUUJGGGUUAAAUCGGCUGCCGUGGGCGUACAAGGAGUCAAGCAUGAC

Ajroz

>**osa-MIR160a**
GUGUGCCUGGUCCCCUGUAUGCCACACAUGUAGACCAACCCAUUGGUUGCCUACUGGGUGGCCUGCAAGGAGCCAAGCAUGC
>**osa-MIR160b**
AAGGUUJUGGUUJUGAUCCGGCUUJUGAGAGCGUGCCUCCUGUAUGCCACUCAUGGCCAACUCCAUUGGUUGGUUJUGGUJUGCUGUGGGUGGCCUGCAAGGAGCCAAGCAUGCGUGCCAUJAUCLJCUCUUCUGCCUU
>**osa-MIR160c**
GGAAUUGGCCUGGUCCCCUGUAUGCCACACUCAUAGAGCAACAAACUUCUGCGAGGGUUGCCUAUAGGUAGGGCGUACCGGAGCCAAGCAUJUCC
>**osa-MIR160d**
GGGGAUUJUGGUCCCCUGUAUGCCACUCGCGUAGCUGCCAAACUCAUGUAAAACACUUGCCUUCCCCGCGAGAUUACGGCAUJUGGUJUGCUACUGCGGAUGGCUGCGAGGAAGCAUGACC
>**osa-MIR160e**
GUAGGGGAUAGCGUGGUCCCCUGUAUGCCACUCCUGCAUCCGCGUCCGCAACCAAAGUACUCGAUCUJUGGUJUGCCGUGCGUACCGGCGAGGUUGCCAAAGCAUGGUJUGCCU
>**osa-MIR160f**
GGAUUACCGCUGGUCCCCUGUAUGCCACUCCUGCAUCCGAGAACCGUGCCGUGUGGCCGGCUGCUUCCUJUGGUJUGCCAAAGGAGUCAUGCGAGGUJUGCCU

Tomate

>**sly-MIR160a**
CAAGUCGUACACGUAAUAGCCUGGUCCCCUGUAUGCCAUUJUGCAAAACUACCGUAAUUAUCGAUGGGCCUJUGUJUGAAUJGGCUAUGAGGAGCCAAGCAUJUUCGUAUCGUAAUAGUJUG

- ▶ Definición de ortólogos para cada miembro de cada familia conservada partiendo de los precursores de *A. thaliana*.
- ▶ Definición arbitraria de los precursores incluyendo 150nt fuera del par miARN/miARN*.

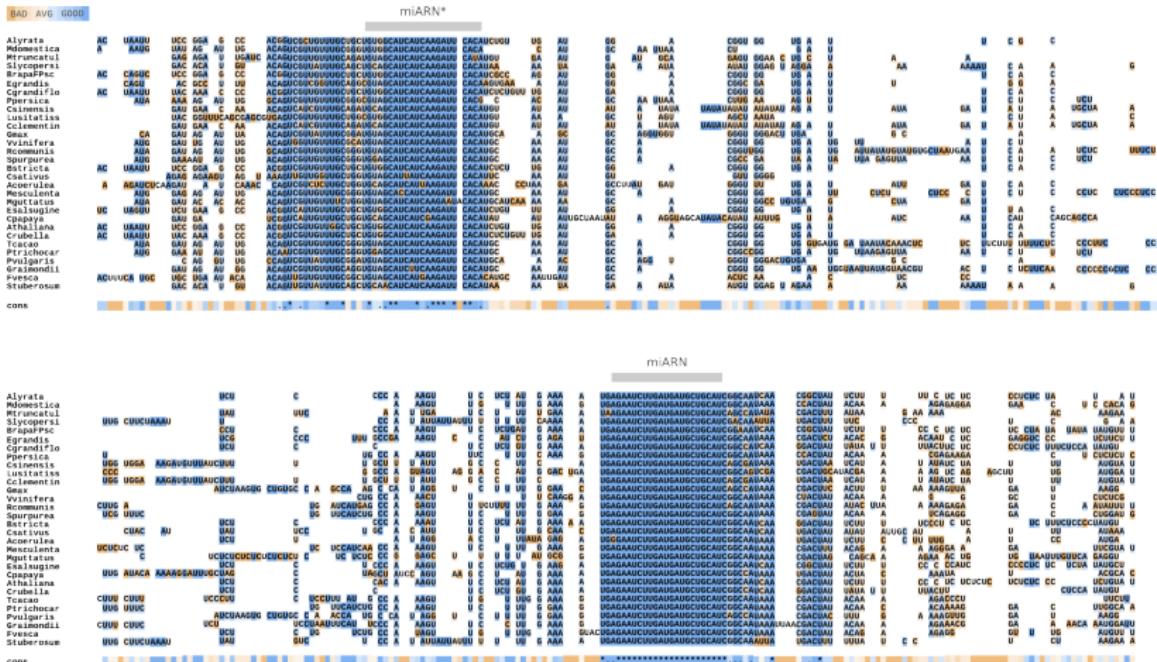
Conservación de la secuencia primaria del miR172a en distintas especies

	BAD	AVG	GDD0	U	C	S	C
Alvareza	AG	UABAU	UCC	GGG	G	US	
Motomestica	BR	AUAG	UAB	AG	RD	US	
Urtica	BR	UAB	AG	RD	US	GRD	
Silicoperis							
Bryopippsc	AG	CBAC	GGG	AA	AA	ECA	
Urticaria	BR	GGC	ACA	AA	AA	ECA	
Cypradiflo	AG	UABUJ	UAC	AA	CC	ECA	
Polygonis	BR	UABUJ	UAC	AA	CC	ECA	
Lutistis							
Urticaria	BR	GGG	ACA	AA	AA	ECA	
Gmax	CA	DAU	AG	AA	AA	ECA	
Vvinifera							
Urticaria	BR	GGG	ACA	AA	AA	ECA	
Spurpurea	AG	DAABA	AG	AA	AA	ECA	
Urticaria	BR	GGG	ACA	AA	AA	ECA	
Casutis	AG	DAABA	AG	AA	AA	ECA	
Acervaria	A	ACACUCA	AG	AA	AA	ECA	
Urticaria	BR	GGG	ACA	AA	AA	ECA	
Mutipatula	AG	DAU	AG	AC	AC	ECA	
Urticaria	BR	GGG	ACA	AA	AA	ECA	
Capapay	AG	UUDU	AG	GG	GG	ECA	
Athelia	AG	UABUJ	AG	GG	GG	ECA	
Tecocao	AG	UABUJ	AG	GG	GG	ECA	
Chacarac	AG	UABUJ	AG	GG	GG	ECA	
Praulinaris	AG	UABUJ	AG	GG	GG	ECA	
Dramonid	AG	UABUJ	AG	GG	GG	ECA	
Strobosoma	AG	UABUJ	AG	GG	GG	ECA	
cons							

ANSWER

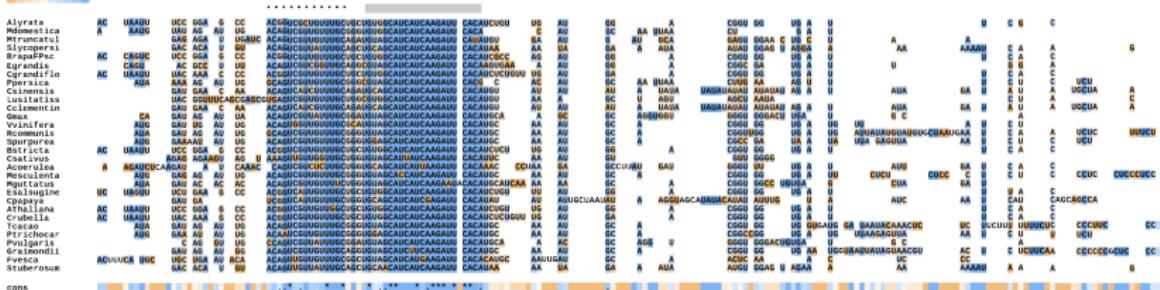
cons

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



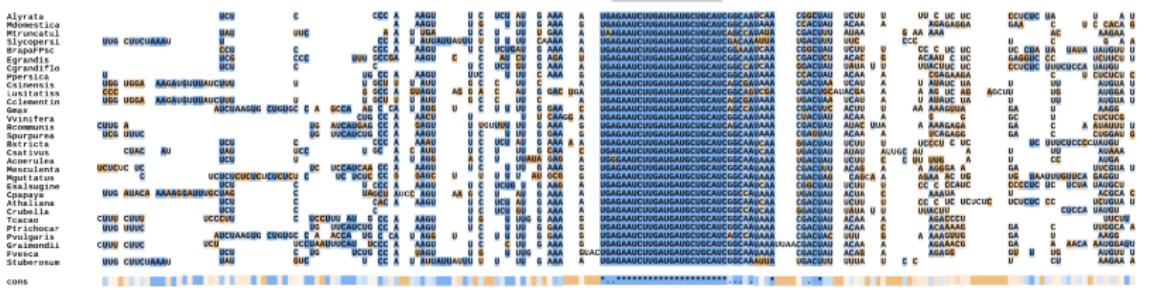
Cola de conservación hacia la izquierda del miARN* y hacia la derecha del miARN

BAD AVG GOOD



cons

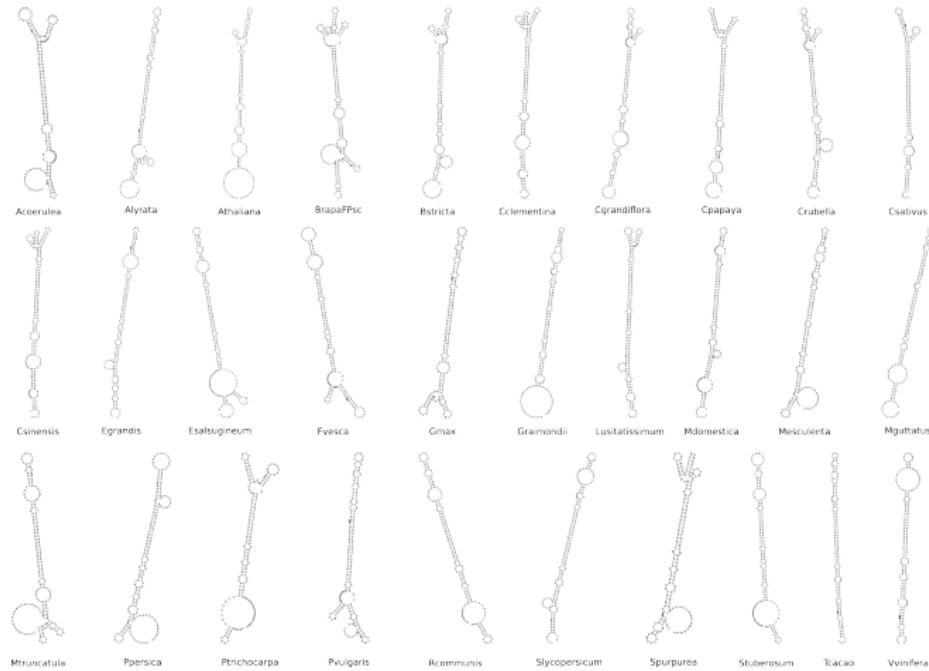
miARN



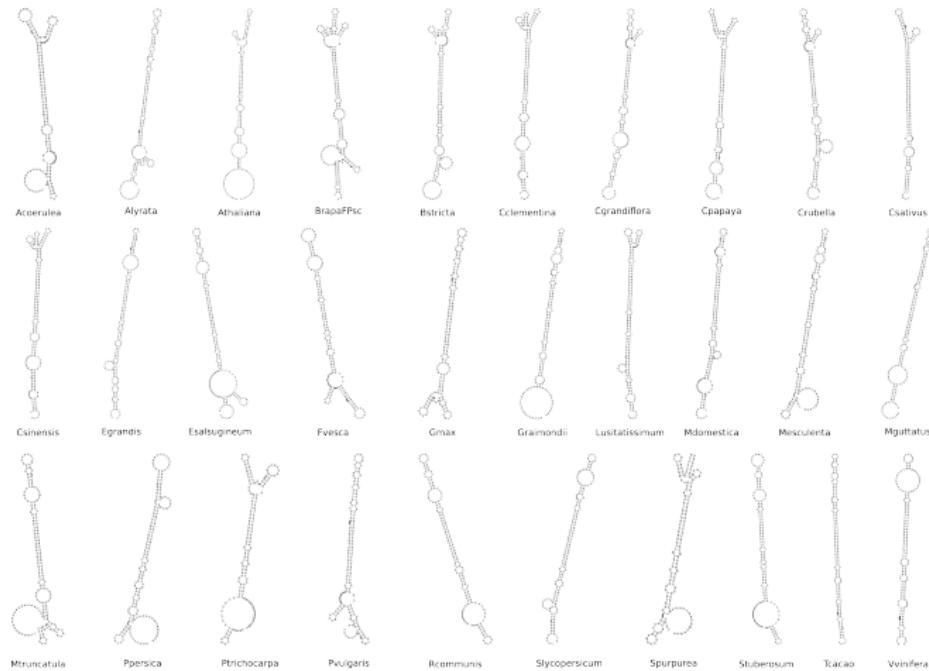
cons



Existe un patrón estructural que comparten los precursores, en la región inmediata por debajo del dúplex miARN/miARN* (miR172a)

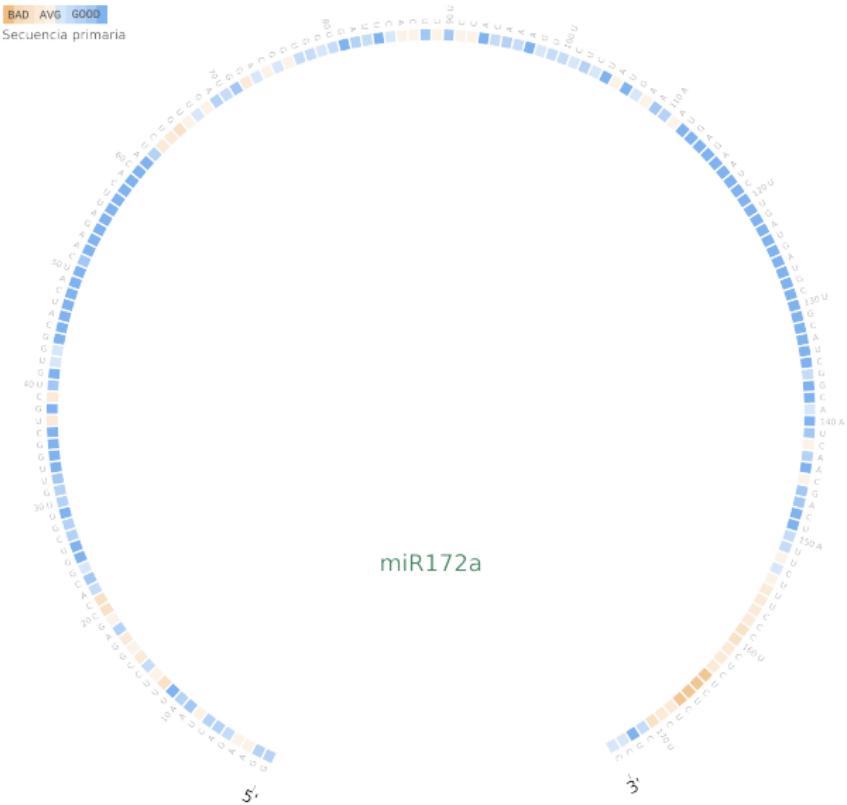


Existe un patrón estructural que comparten los precursores, en la región inmediata por debajo del dúplex miARN/miARN* (miR172a)

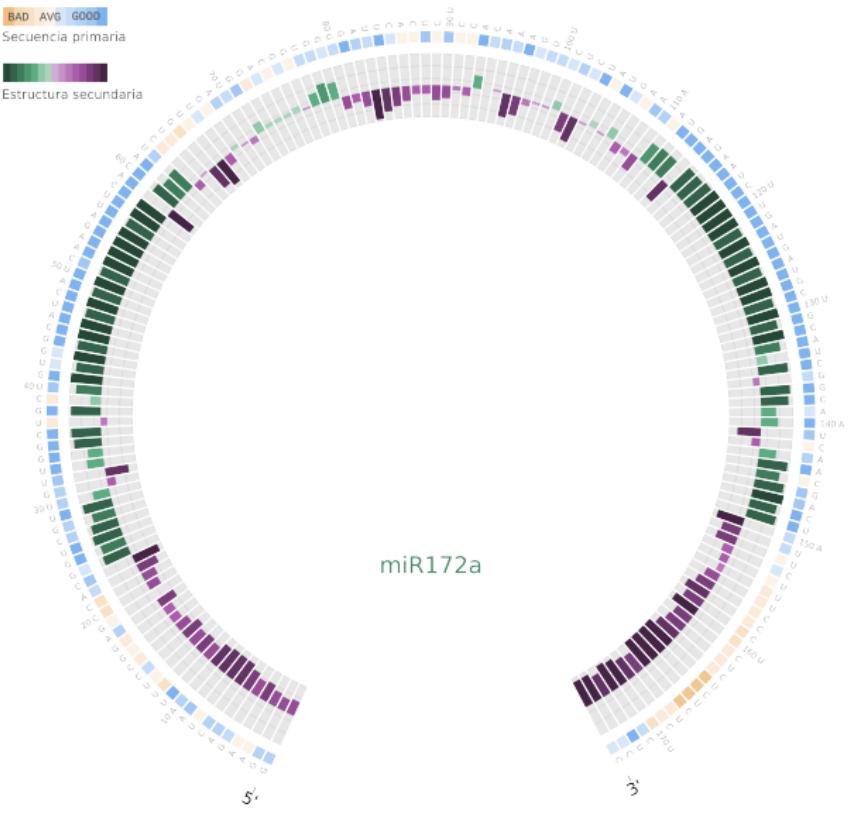


No es trivial deducir información concreta a partir de estas figuras.

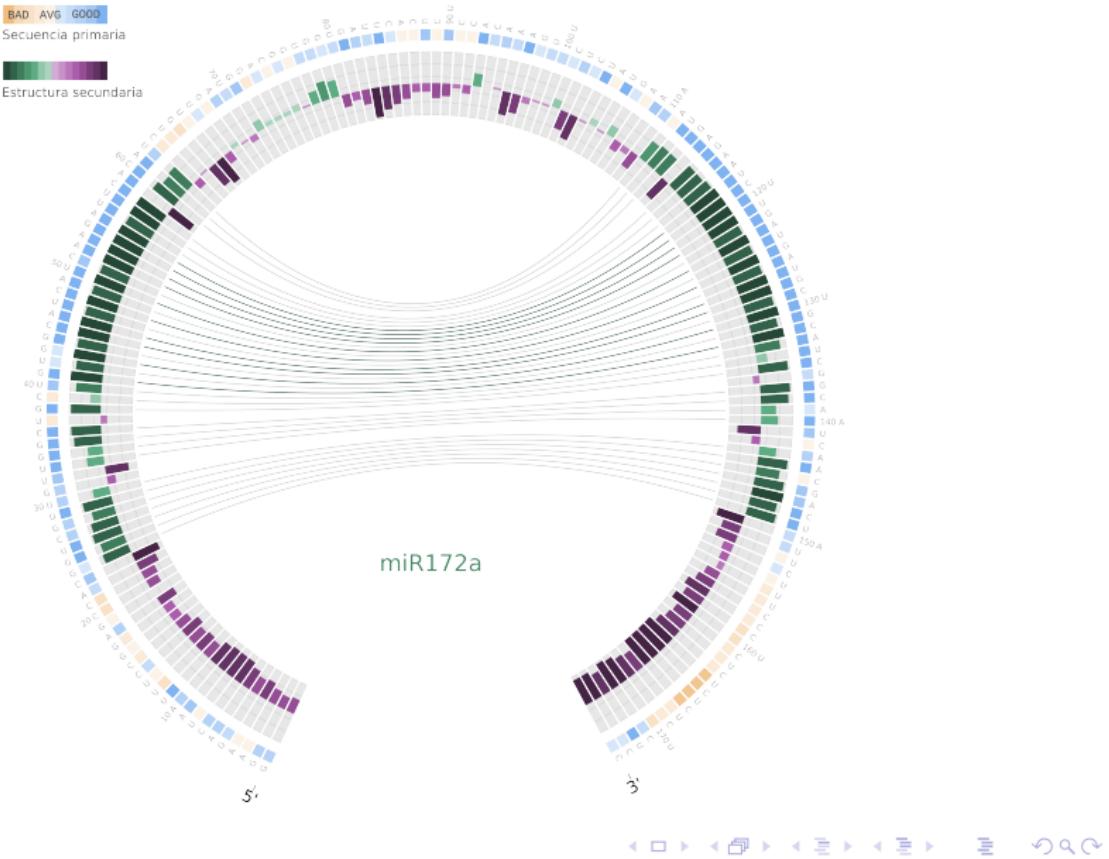
Conservación del consenso en base al alineamiento de secuencia primaria



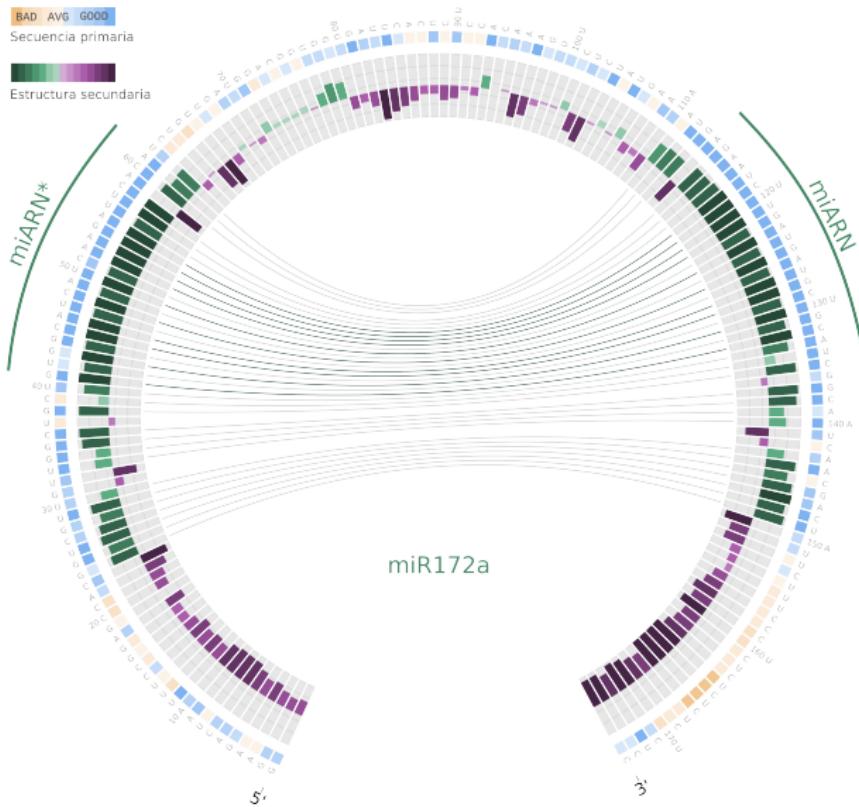
Frecuencia de bases apareadas y desapareadas



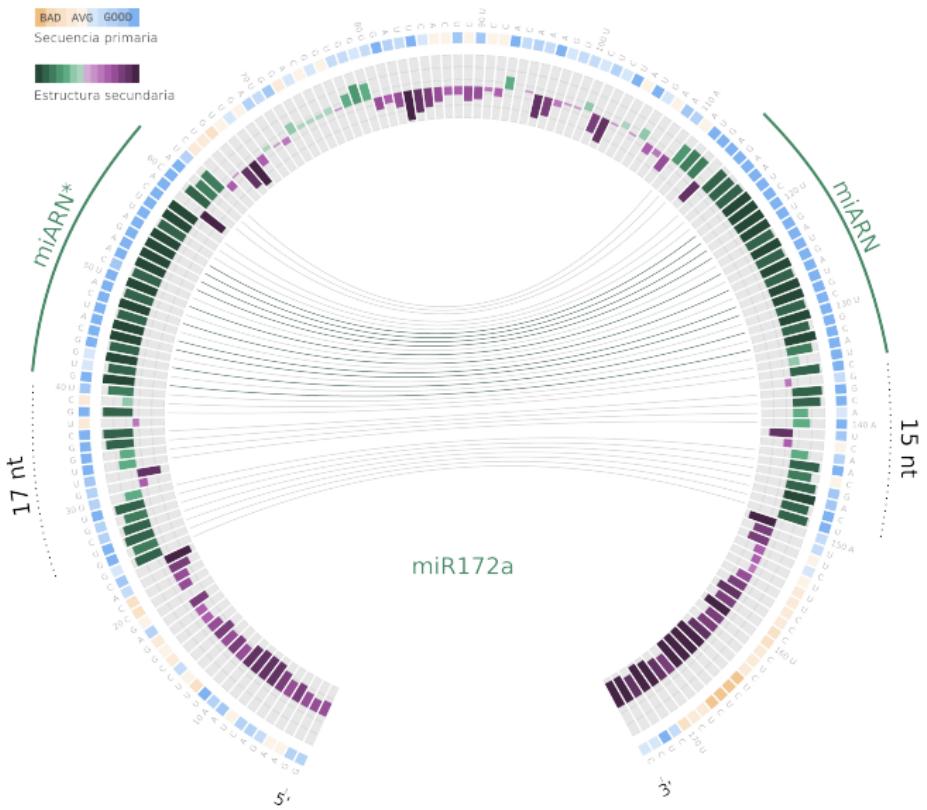
Interacción entre pares de bases considerando estructura secundaria



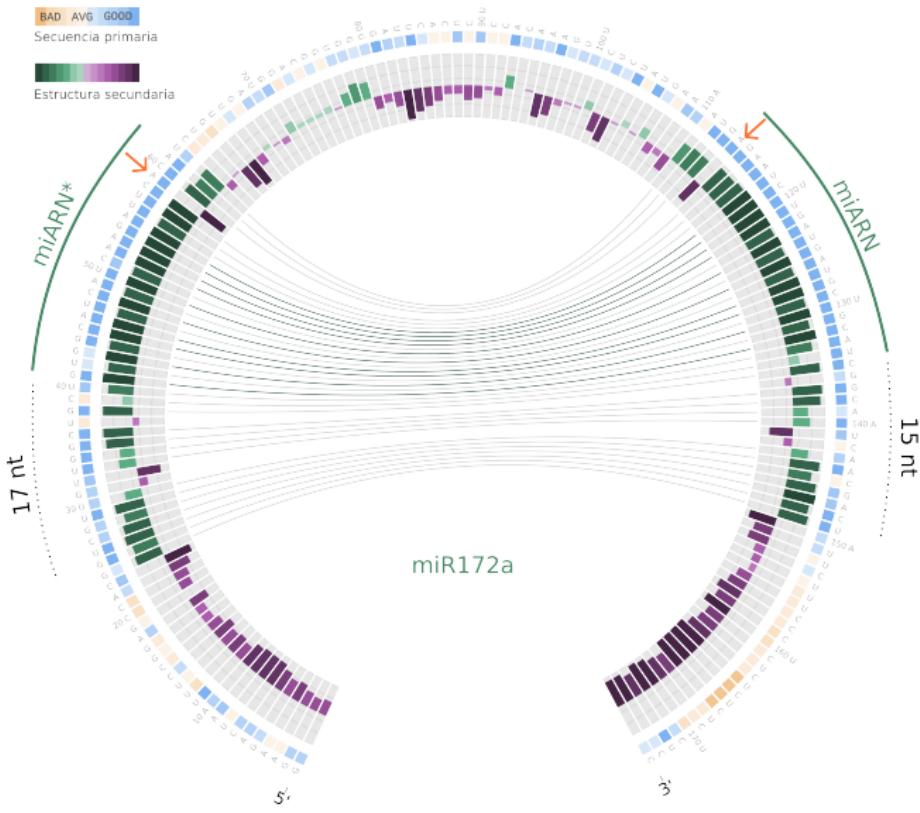
miARN y miARN* conservados en secuencia primaria y estructura



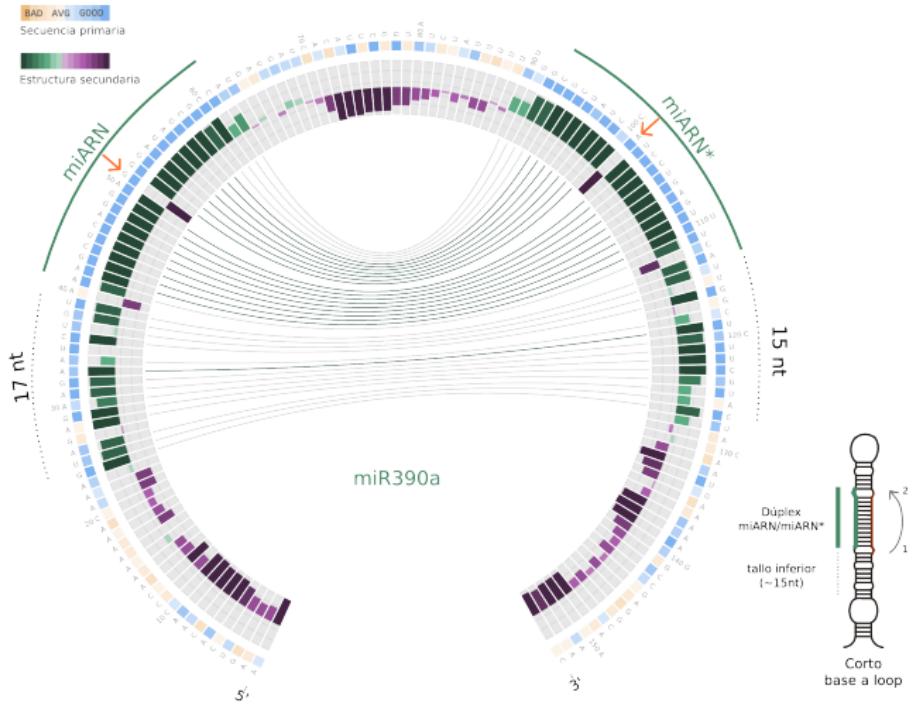
Región conservada por debajo del dúplex que coincide con el tallo inferior



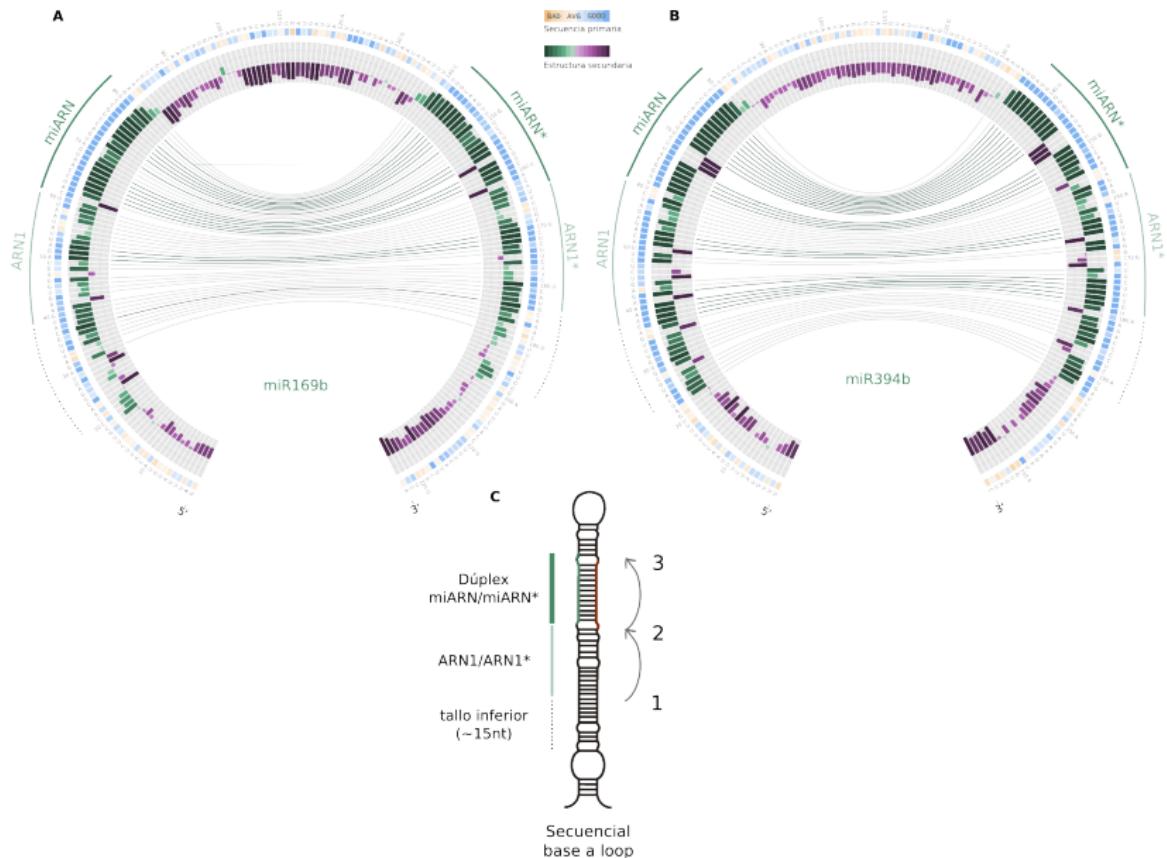
Mismatches conservados



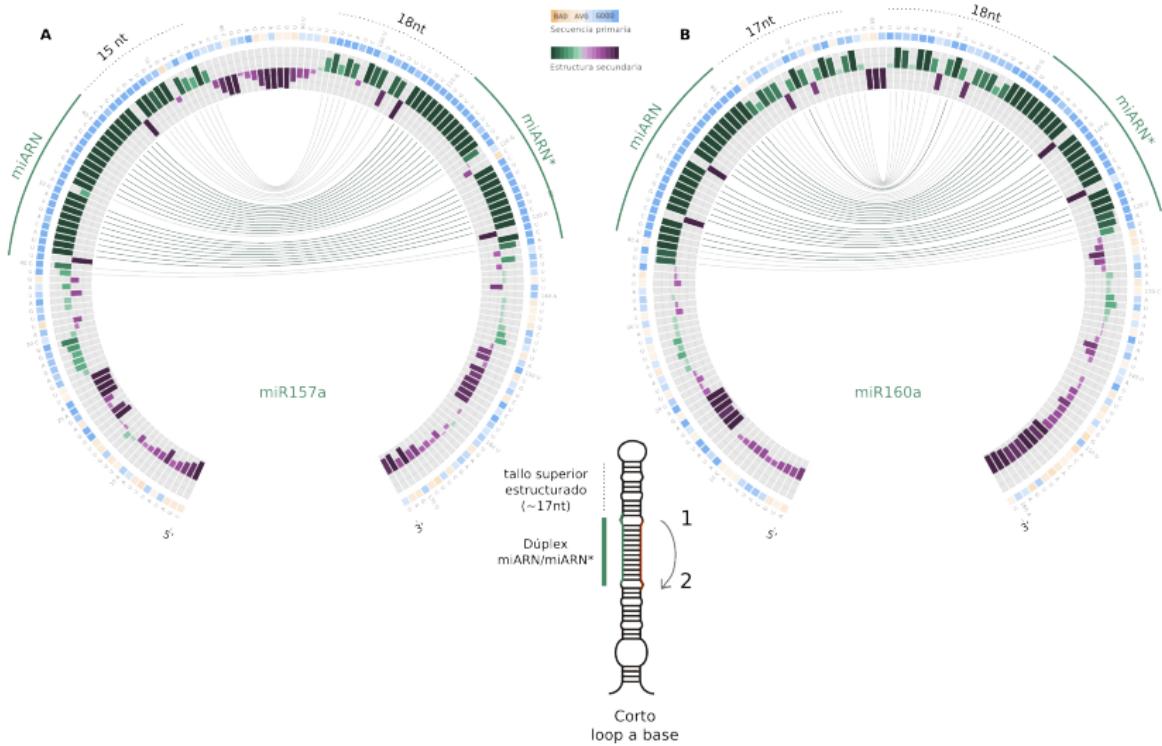
Mismo patrón de conservación en otros precursores que se procesan desde la base



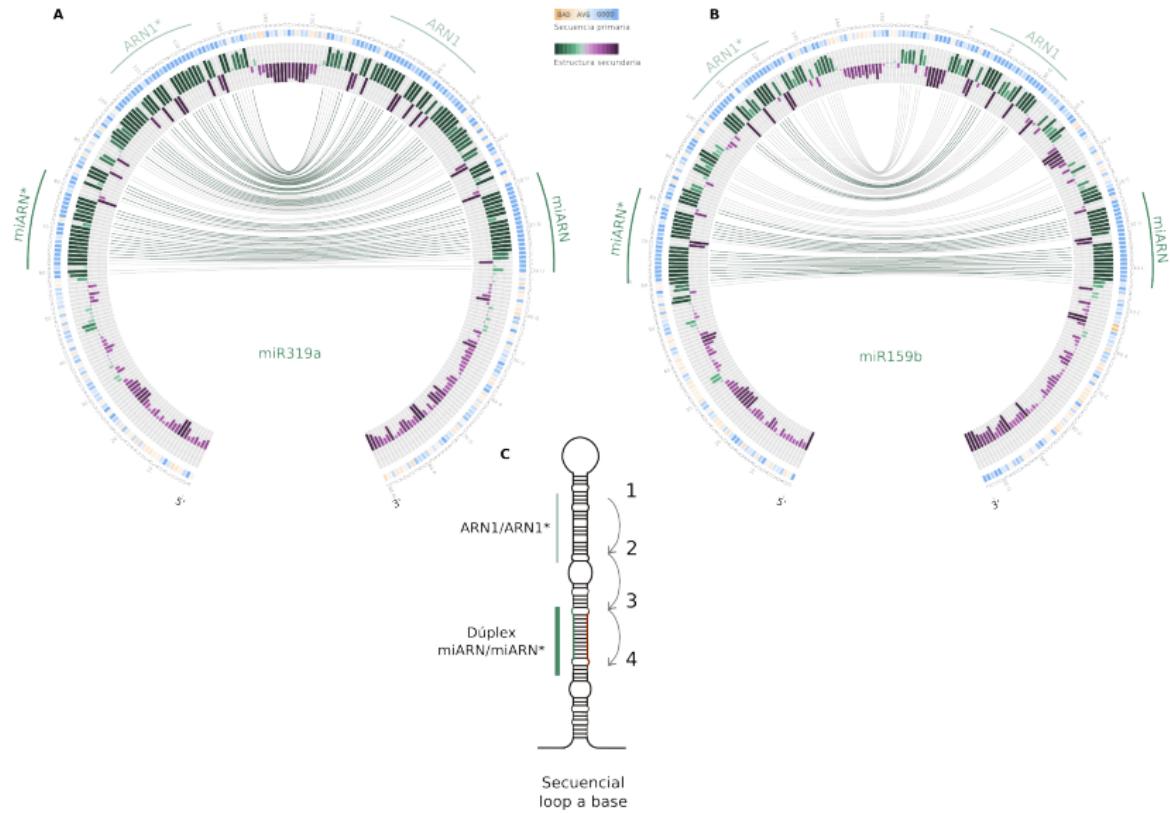
Precursos que se procesan desde la base en forma secuencial



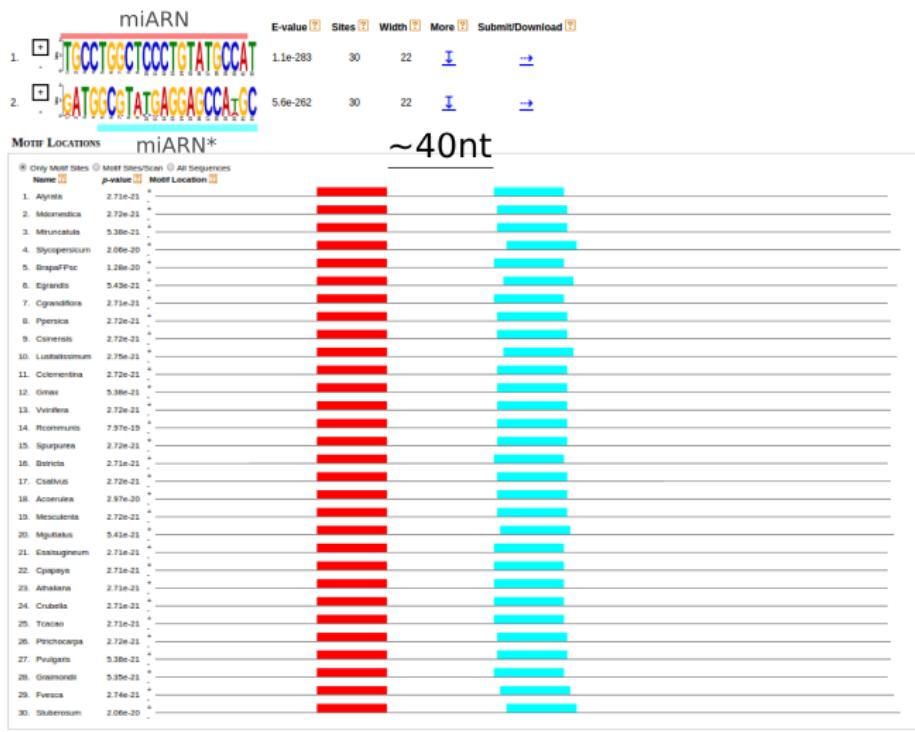
Precursors que se procesan desde el loop corto



Precursors que se procesan desde el loop en forma secuencial



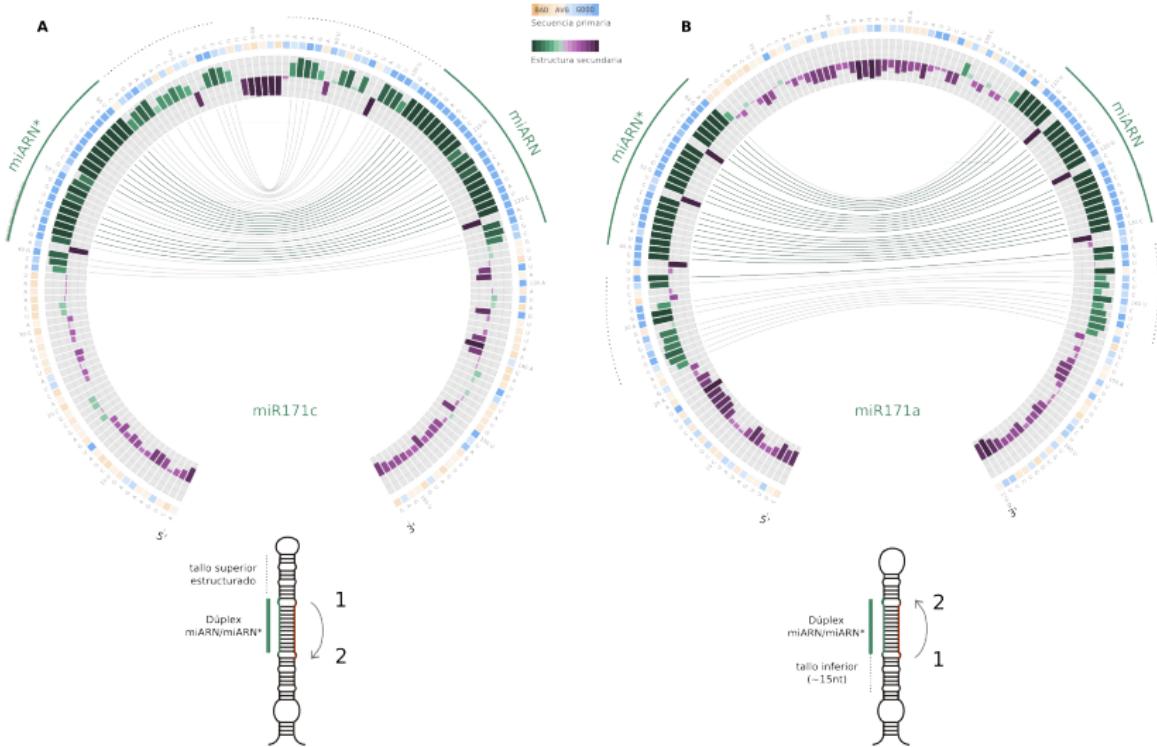
En precursores que se procesan desde el loop, el tamaño de la región que comprende al tallo superior y al loop no varía en distintas especies



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



Procesamiento mixto de miembros de la familia del miR170/miR171

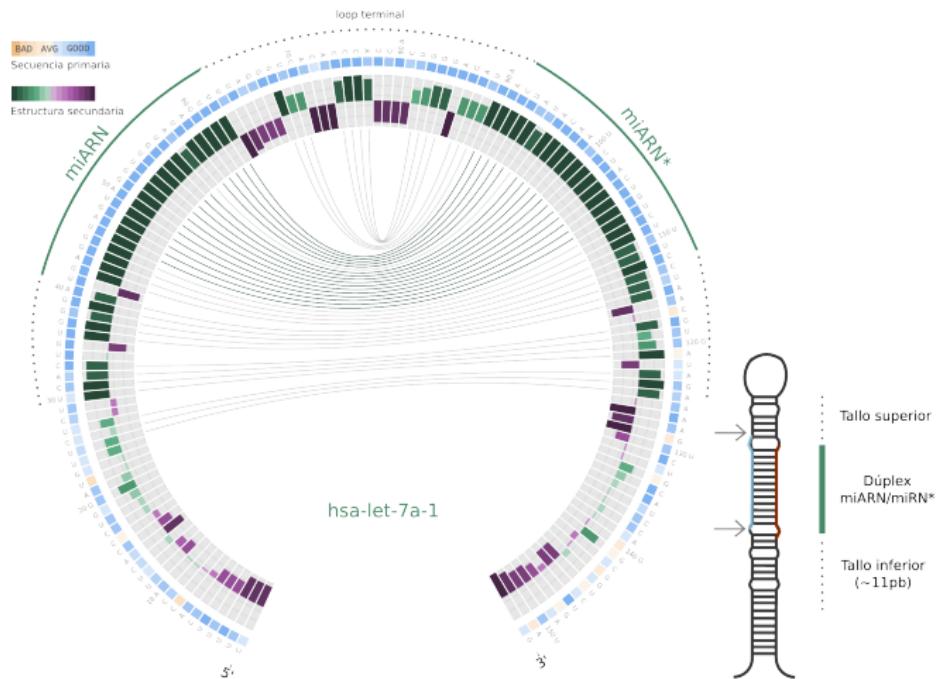


¿Qué sucede con los precursores de miARNs conservados en animales?

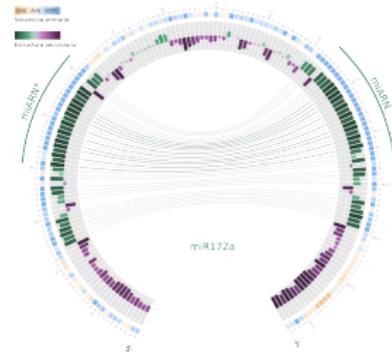
Animales

Bos taurus
Canis familiaris
Equus caballus
Gallus gallus
Gorilla gorilla
Homo sapiens
Macaca mulatta
Monodelphis domestica
Mus musculus
Ornithorhynchus anatinus
Petromyzon marinus
Sus scrofa
Xenopus tropicalis

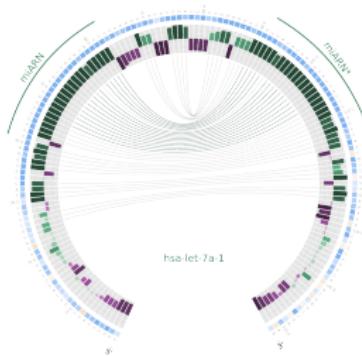
El loop terminal está conservado en la mayoría de los precursores de animales estudiados.



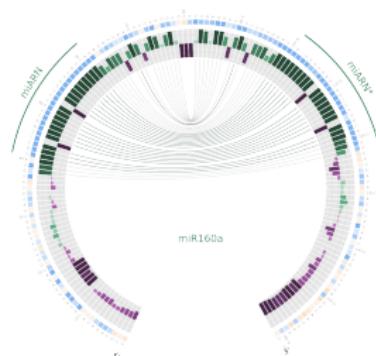
Circos animales vs plantas



Precursors de plantas.
Cortos de base a loop



Precursors de animales.



Precursors de plantas.
Cortos de loop a base

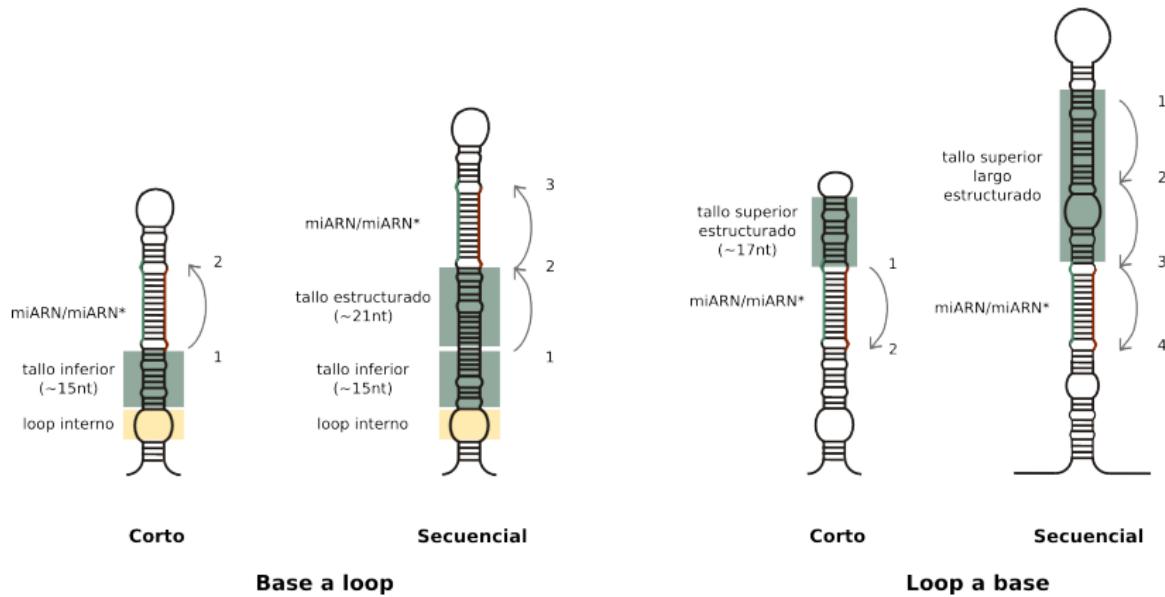
- ▶ Presentamos un enfoque bioinformático para el estudio de la evolución y biogénesis de miARNs en plantas.

- ▶ Presentamos un enfoque bioinformático para el estudio de la evolución y biogénesis de miARNs en plantas.
- ▶ Desarrollamos una implementación gráfica para visualizar de manera simple los precursores de miARNs en distintas especies de plantas.

- ▶ Presentamos un enfoque bioinformático para el estudio de la evolución y biogénesis de miARNs en plantas.
- ▶ Desarrollamos una implementación gráfica para visualizar de manera simple los precursores de miARNs en distintas especies de plantas.
 - ▶ Lo utilizamos para caracterizar la evolución de precursores de miARNs en plantas con distintos mecanismos de procesamiento.

- ▶ Presentamos un enfoque bioinformático para el estudio de la evolución y biogénesis de miARNs en plantas.
- ▶ Desarrollamos una implementación gráfica para visualizar de manera simple los precursores de miARNs en distintas especies de plantas.
 - ▶ Lo utilizamos para caracterizar la evolución de precursores de miARNs en plantas con distintos mecanismos de procesamiento.
 - ▶ Pudimos utilizar este mismo enfoque para estudiar precursores de miARNs en animales.

Conclusiones III



- ▶ Desarrollamos aplicaciones bioinformáticas para el estudio de interacciones miARN-gen blanco.

- ▶ Desarrollamos aplicaciones bioinformáticas para el estudio de interacciones miARN-gen blanco.
- ▶ Desarrollamos una herramienta para el análisis de bibliotecas de SPARE incluyendo una interfaz gráfica.

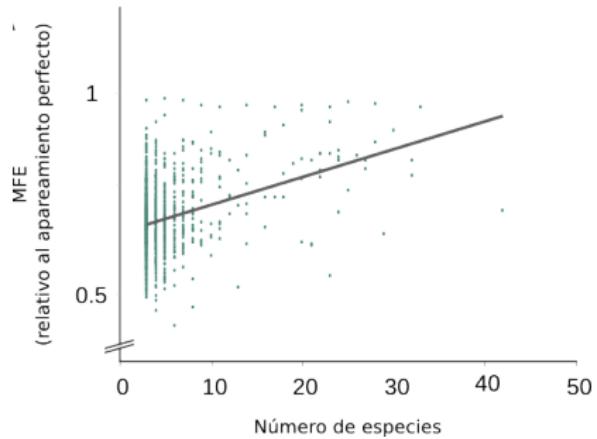
- ▶ Desarrollamos aplicaciones bioinformáticas para el estudio de interacciones miARN-gen blanco.
- ▶ Desarrollamos una herramienta para el análisis de bibliotecas de SPARE incluyendo una interfaz gráfica.
- ▶ Analizamos las estructuras de los precursores y su evolución.

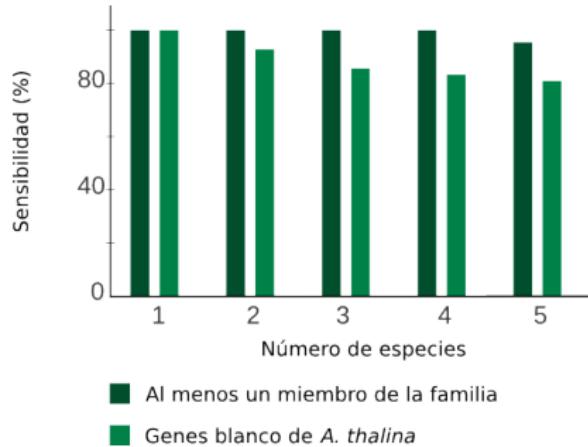
- ▶ Desarrollamos aplicaciones bioinformáticas para el estudio de interacciones miARN-gen blanco.
- ▶ Desarrollamos una herramienta para el análisis de bibliotecas de SPARE incluyendo una interfaz gráfica.
- ▶ Analizamos las estructuras de los precursores y su evolución.
- ▶ Encontramos determinantes mecánisticos del procesamiento de miARNs en plantas.

- ▶ Desarrollamos aplicaciones bioinformáticas para el estudio de interacciones miARN-gen blanco.
- ▶ Desarrollamos una herramienta para el análisis de bibliotecas de SPARE incluyendo una interfaz gráfica.
- ▶ Analizamos las estructuras de los precursores y su evolución.
- ▶ Encontramos determinantes mecánisticos del procesamiento de miARNs en plantas.
- ▶ Realizamos una forma de representación Visualización de información compleja por adaptación de una herramienta Circos.

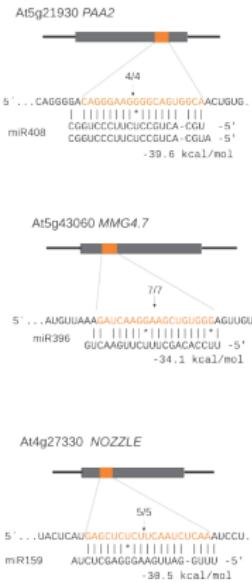
Muchas
gracias







Nuevos genes blancos validados en *A. thaliana*



Arabidopsis thaliana CAGGGAAAGGGCAGUGCA
Capsicum annuum CAGCGGAAGGAAGCAGUUGCA
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

At3g22110 PAC1

5' ... AUCCAAAC~~CAGGGAAACCCACCCACUCA~~GCUGGCC...
miR408 CGGUCCUCUUCUGCGUACCU -5'
-32.9 kcal/mol

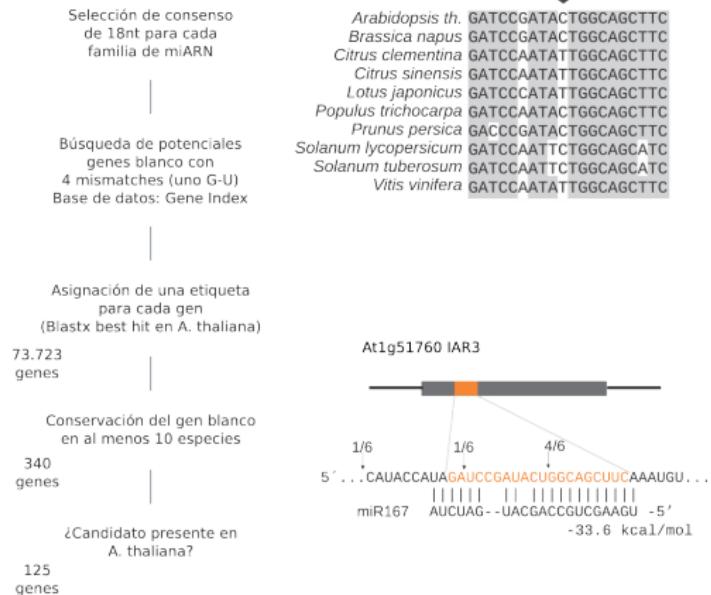
Arabidopsis thaliana CAGGGAAAGAGCAGUCGA
Capiscum annuum AAGGGAAAGGGCAGUUC
Citrus clementina AAGGGAAAGACAGUUC
Glycine max AAGGGAAAGAGCAGUUC
Helianthus annuus CAGAAAAGAGCAGUUC
Lotus japonicus UAGGAAGAGAAGCAGUUC
Medicago truncatula AAGGGAAAGAGCAGUUC
Nicotiana tabacum AAGGGAAAGAGCAGUUC
Petunia hybrida AAGGGAAAGAGCAGUUC
Phaseolus coccineus AAGGGAAAGAGCAGUUC
Prunus persica AAGGAAGAGGCGAGUUC
Solanum tuberosum AAGGGAAAGAGGCGAGUUC
Theobroma cacao AAGGGAAAGAGCAGUUC
Triphysaria versicolor CAGGGAGGAGCGUGUUC

At3g14110 FLU

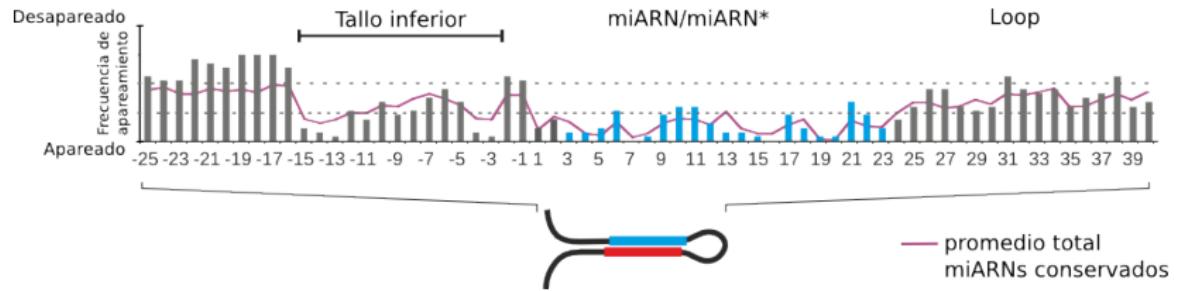
5' ... AUACAGA~~GUUCAAGAUACGUUUGGACUUGC~~...
miR396 GUCAAGUUCUUUCGACCCU -5'
-39.1 kcal/mol

Arabidopsis thaliana GUUCAAGAUACGUUUGGA
Brassica napus GUUCAAGACAGCUUUGGA
Festuca arundinacea GUUACGACAGCUUUGGA
Hordeum vulgare GUUCAAGAUAGCUCUUGGA
Ipomoea nil GUUCAAGAAAGCCUUGA
Panicum virgatum GUUCAGGACAGCUCUUGGA
Triticum aestivum GUUCAAGACGCGCUUGGA

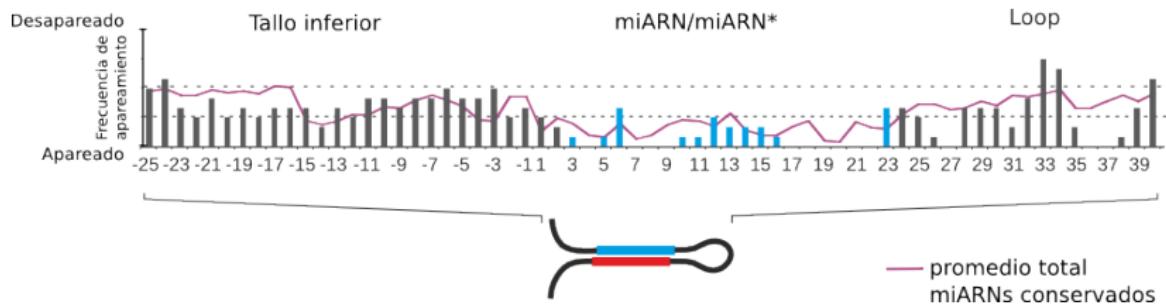
Nuevos genes blancos con interacciones G-U



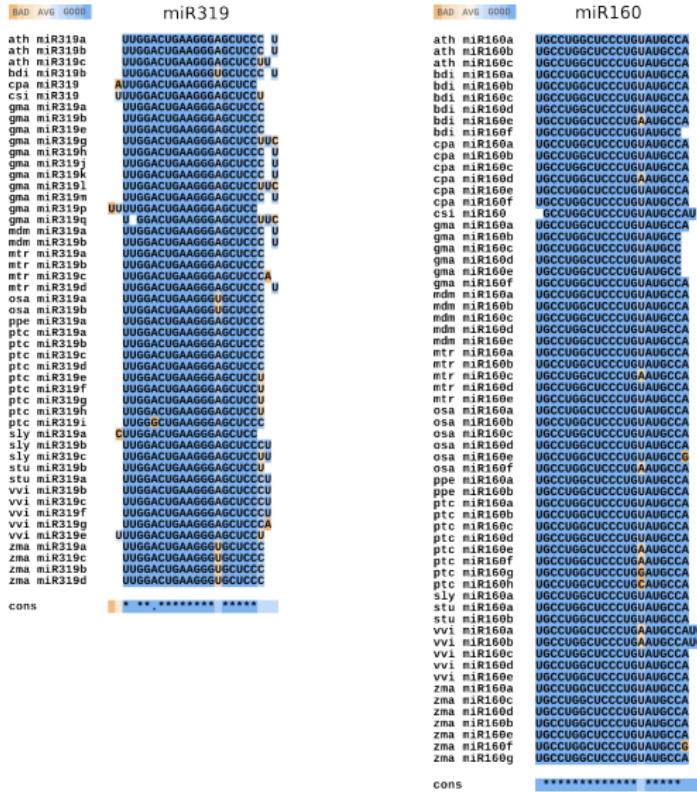
Tallo inferior de 15 nt en precursores procesados desde la base



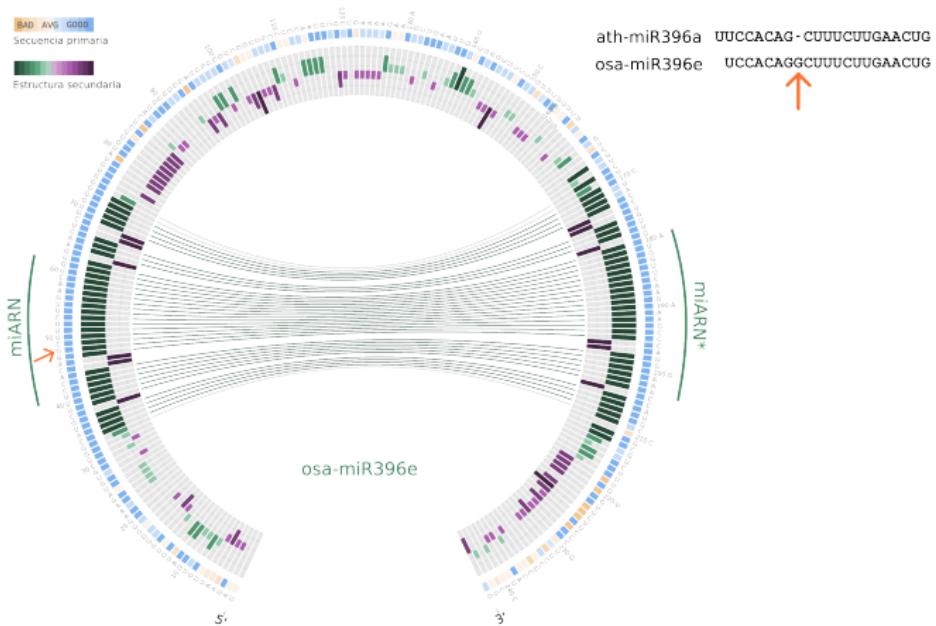
Región terminal estructurada en precursores procesados desde el loop



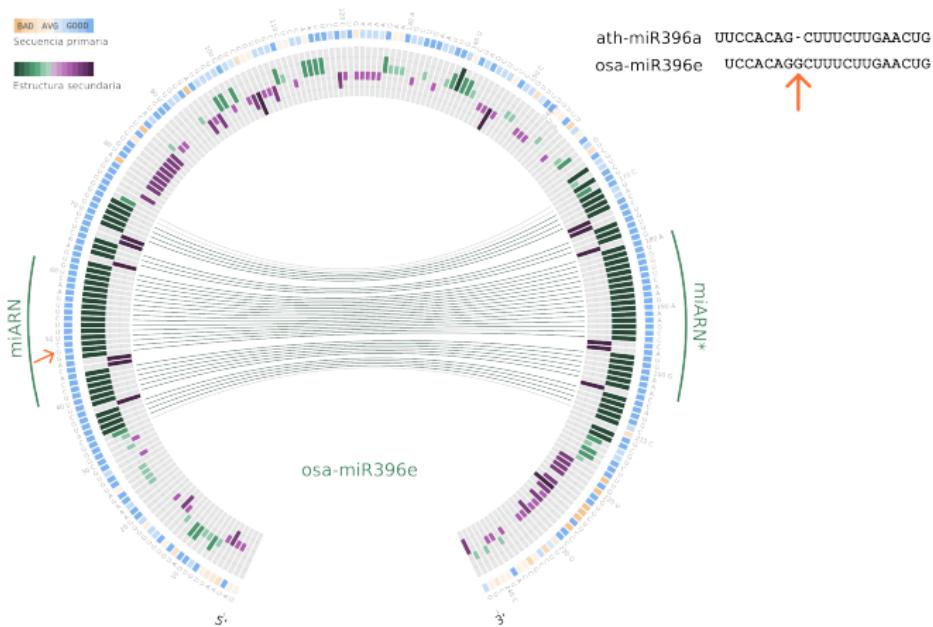
Conservación y divergencia de miARNs en distintas especies



Variante del miR396 específica de monocotiledóneas

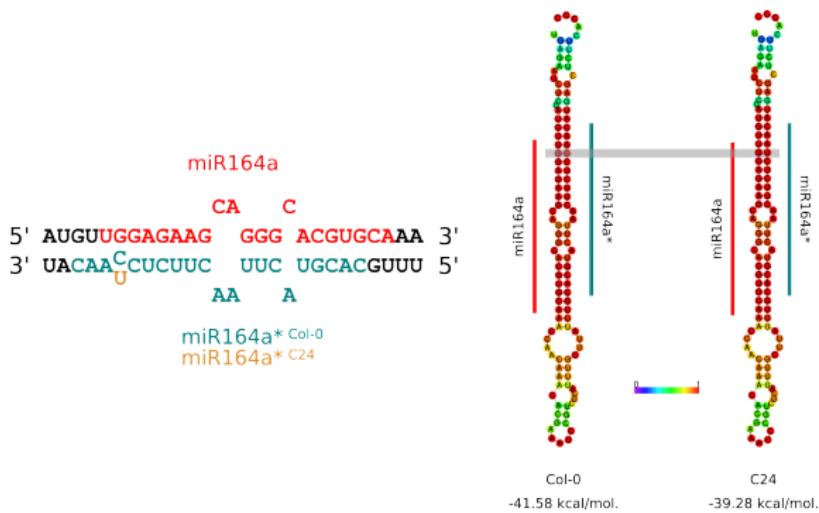


Variante del miR396 específica de monocotiledóneas

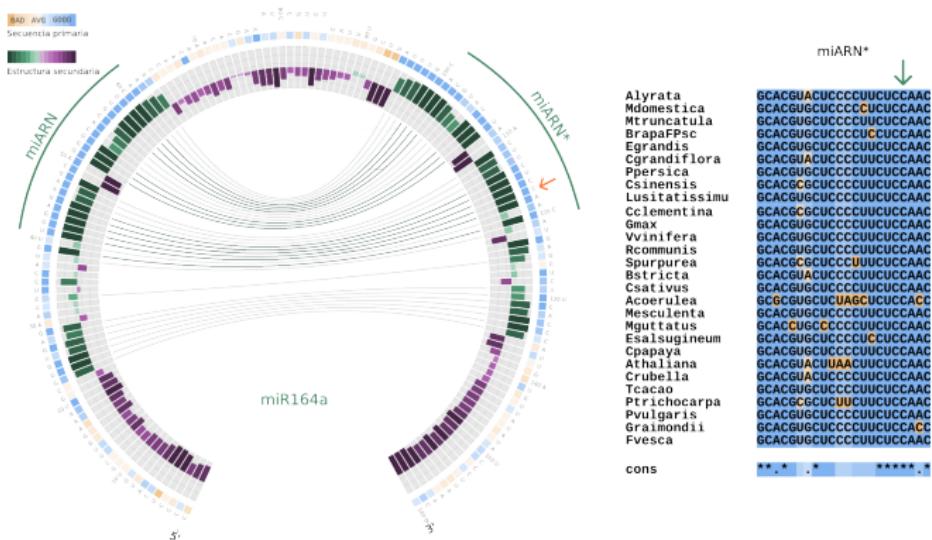


El nucleótido extra, que le da identidad a la variante de monocotiledóneas, está conservado.

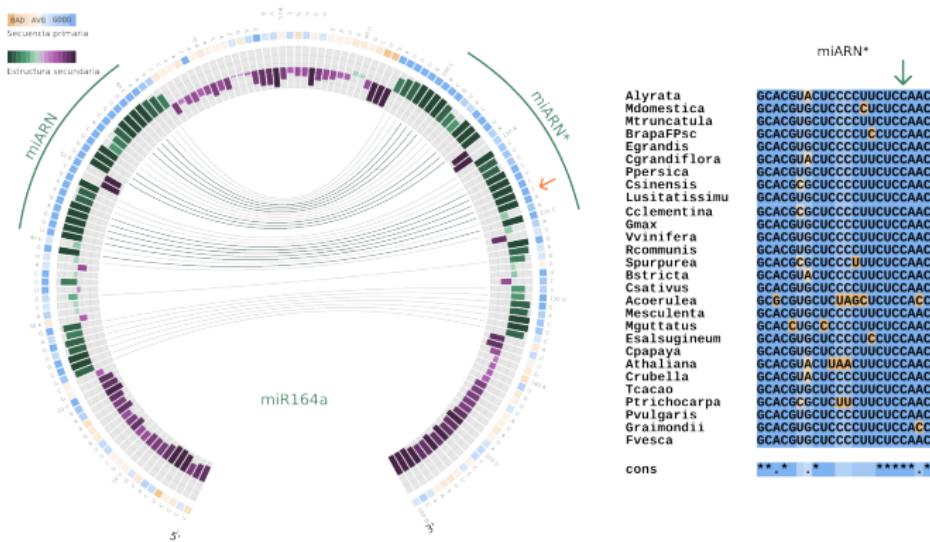
Mutaciones puntuales que afectan el procesamiento de miARNs en plantas



Posición *2 del miR164a* está conservada en dicotiledóneas

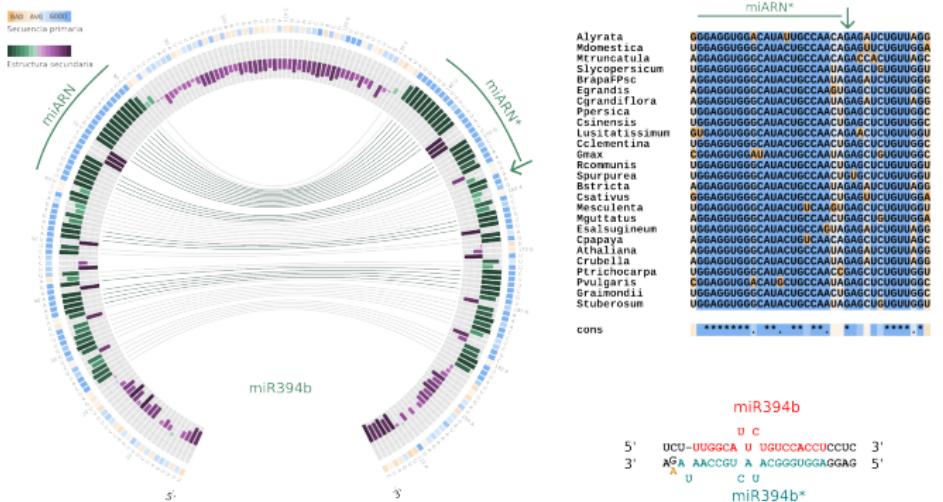


Posición *2 del miR164a* está conservada en dicotiledóneas

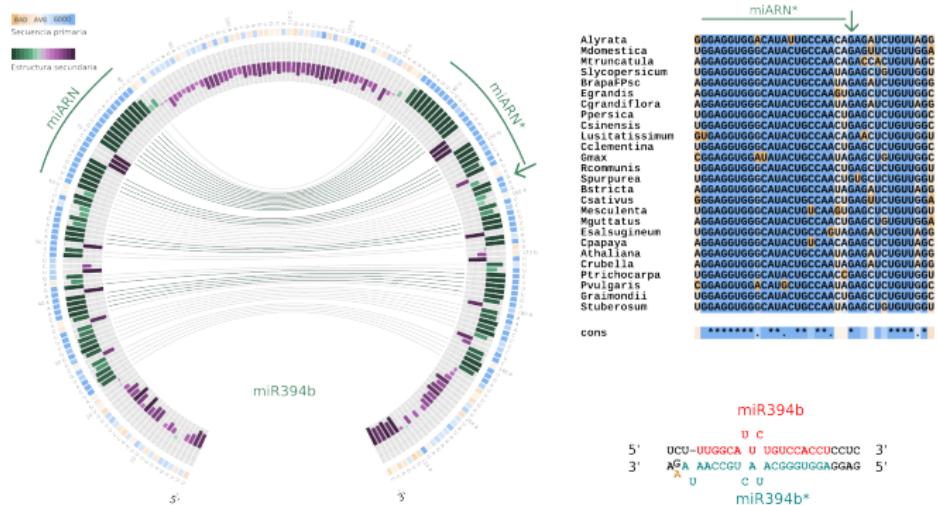


La posición *2 es importante para la estabilidad del precursor y su buen procesamiento.

Alelo mir394b-1 con un “mismatch” en el tallo inferior del precursor del miR394b



Alelo mir394b-1 con un “mismatch” en el tallo inferior del precursor del miR394b



Mutaciones simples en el precursor (fuera del dúplex) puede afectar el reconocimiento de DICER.