

microRNA target identification: insights from the evolutionary history of microRNA families

Hervé Seitz

March 9, 2023

This slideshow is available at:

<http://www.igh.cnrs.fr/equip/Seitz/srRNA2023.pdf>

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target identification

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target identification

- Low-throughput genetics (measures *in vivo* phenotypes).

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target identification

- ▶ Low-throughput genetics (measures *in vivo* phenotypes).
- ▶ High-throughput molecular biology (finding mRNAs bound, or repressed, by a miRNA).

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target identification

- ▶ Low-throughput genetics (measures *in vivo* phenotypes).
- ▶ High-throughput molecular biology (finding mRNAs bound, or repressed, by a miRNA).
- ▶ Computational predictions.

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

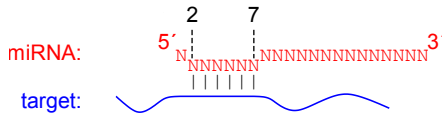
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target prediction



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

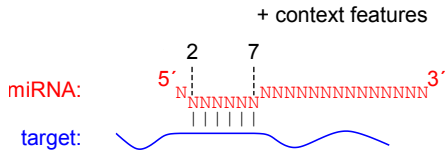
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target prediction



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

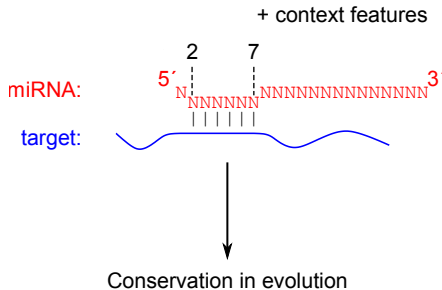
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target prediction



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

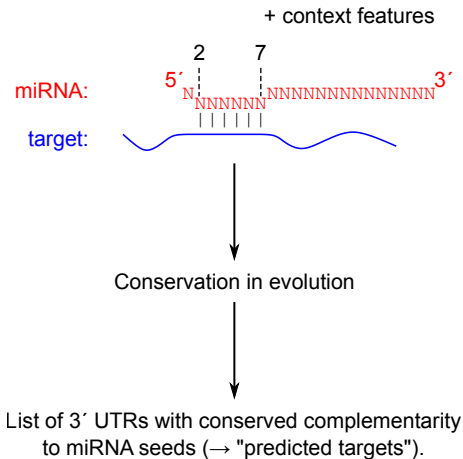
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target prediction



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

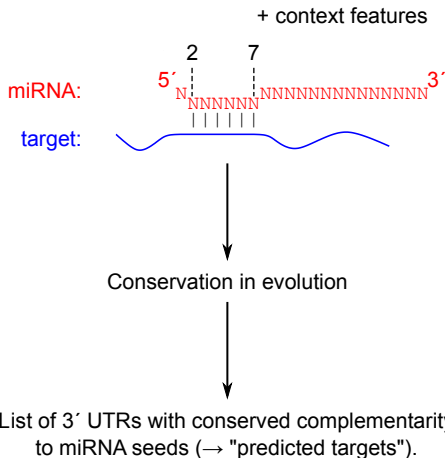
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

miRNA target prediction



Even high-throughput identification needs to assess conservation.

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

False positives in target prediction

Some 3' UTR seed matches are conserved for miRNA-independent reasons.

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

False positives in target prediction

Some 3' UTR seed matches are conserved for miRNA-independent reasons.

Seed match is more conserved than the miRNA itself!
(Pinzón *et al.*, 2017)

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

False positives in target prediction

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

False positives in target prediction

Measuring miRNA-independent conservation with control k -mers (with same G/C content, dinucleotide composition, ...; [Friedman *et al.*, 2009](#)).

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

False positives in target prediction

Measuring miRNA-independent conservation with control k -mers (with same G/C content, dinucleotide composition, ...; [Friedman *et al.*, 2009](#)).

→ not submitted to the same selective pressure than actual miRNA seed matches (e.g., protein binding sites).

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Binding site evolution after miRNA loss

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Binding site evolution after miRNA loss

After loss of a *complete* miRNA family (*), how long do its seed matches persist ?

(*): family: all the miRNAs sharing the same seed sequence

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Binding site evolution after miRNA loss

After loss of a *complete* miRNA family, how long do its seed matches persist ?

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Binding site evolution after miRNA loss

After loss of a *complete* miRNA family, how long do its seed matches persist ?

We need to be lucky: loss of complete miRNA families a long time ago (many species with miRNA loss).

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Strategy

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

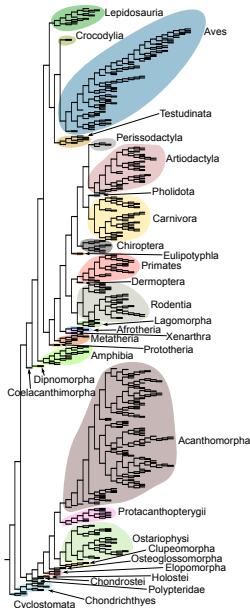
Evolution of seed match
count after loss

Conclusion

Supplementary
data

Strategy

522 vertebrate genomes:



miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Strategy

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

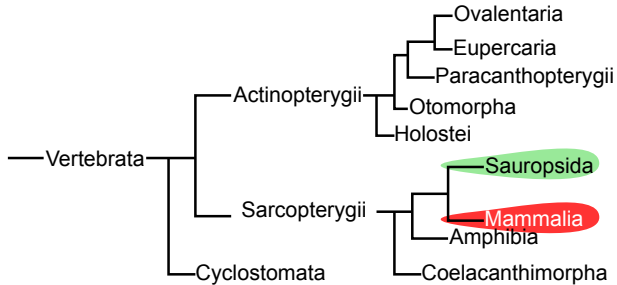
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Strategy



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

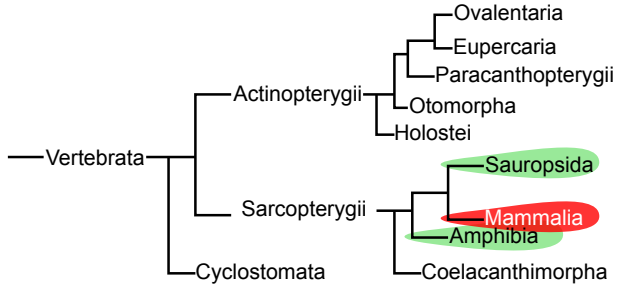
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Strategy



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

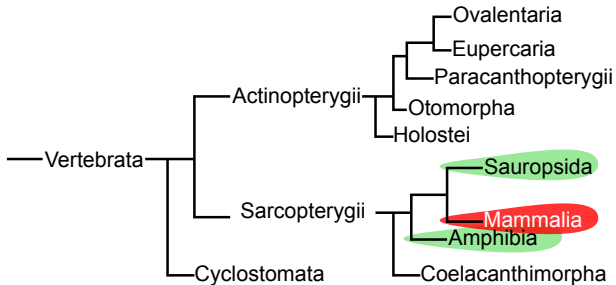
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Strategy



More convincing if many species in clade of loss.

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

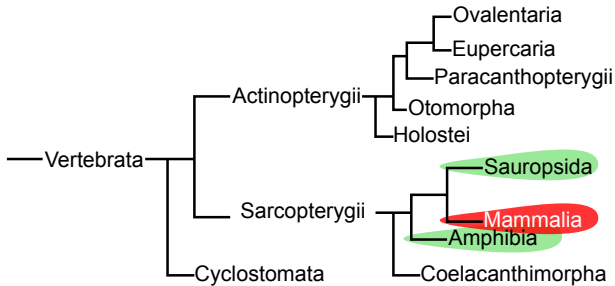
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Strategy



More convincing if many species in clade of loss.

≥ 10 species in clade of loss; presence in sister clade + 3 closest cousin clades.

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Observed losses

Clade	miRNA family
Feliformia	miR-541
Ferae	miR-598
Mammalia	miR-456
Eutheria	miR-2188
Myomorpha	miR-656
Mammalia	miR-460b
Feloidea	miR-676
Cervidae+Bovidae	miR-676
Whippomorpha	miR-769
Cervidae+Bovidae	miR-509
Protacanthopterygii	miR-31
Myomorpha	miR-885
Myomorpha	miR-3959
Theria	miR-458
Pinnipedia+Musteloidea	miR-297
Bovidae	miR-325-3p
Cervidae+Bovidae	miR-514

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Observed losses

Clade	miRNA family
Feliformia	miR-541
Ferae	miR-598
Mammalia	miR-456
Eutheria	miR-2188
Myomorpha	miR-656
Mammalia	miR-460b
Feloidea	miR-676
Cervidae+Bovidae	miR-676
Whippomorpha	miR-769
Cervidae+Bovidae	miR-509
Protacanthopterygii	miR-31
Myomorpha	miR-885
Myomorpha	miR-3959
Theria	miR-458
Pinnipedia+Musteloidea	miR-297
Bovidae	miR-325-3p
Cervidae+Bovidae	miR-514

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

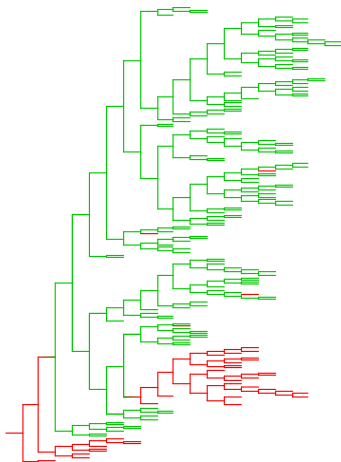
Evolution of seed match
count after loss

Conclusion

Supplementary
data

miR-656 family:

Observed losses



miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

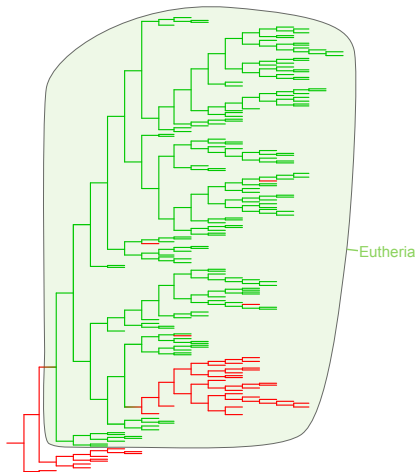
Evolution of seed match
count after loss

Conclusion

Supplementary
data

miR-656 family:

Observed losses



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

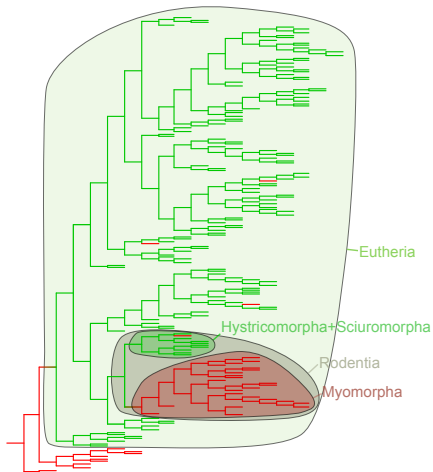
Evolution of seed match
count after loss

Conclusion

Supplementary
data

miR-656 family:

Observed losses



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

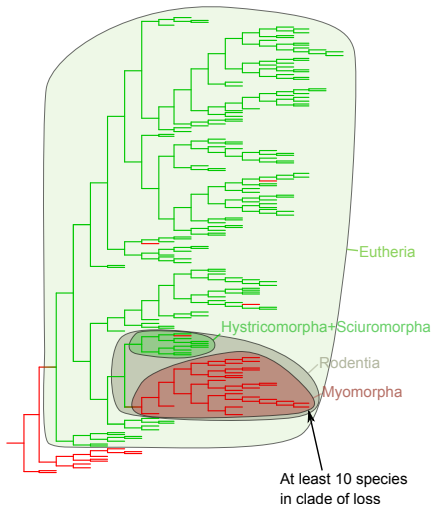
Evolution of seed match
count after loss

Conclusion

Supplementary
data

miR-656 family:

Observed losses



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

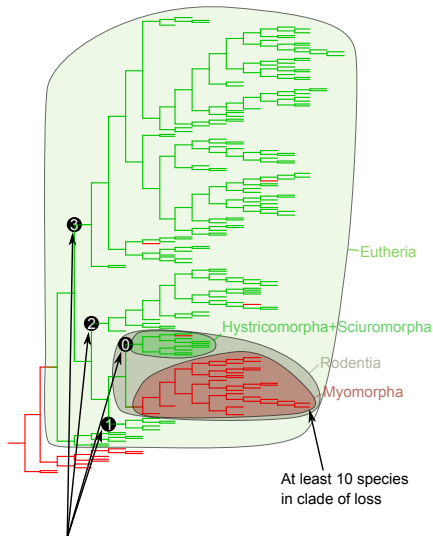
Evolution of seed match
count after loss

Conclusion

Supplementary
data

miR-656 family:

Observed losses



In sister clade + at least 3 degrees of cousinhood: miRNA family with certified presence in >50% of species at each degree

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Evolution of seed match count after loss

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

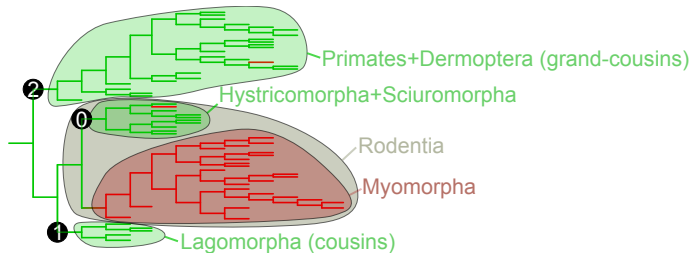
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Evolution of seed match count after loss



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

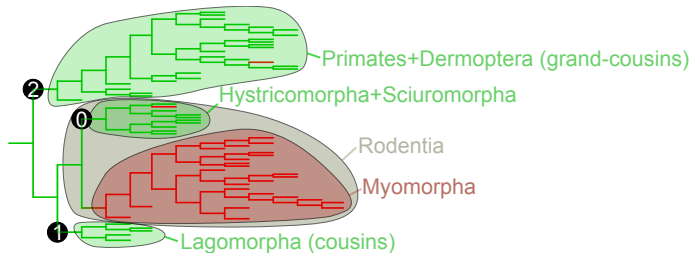
Observed losses

Evolution of seed match
count after loss

Conclusion

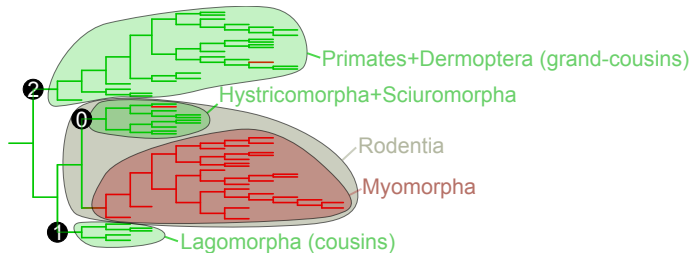
Supplementary
data

Evolution of seed match count after loss



Align 3' UTR sequences in clade of loss and sister clade.

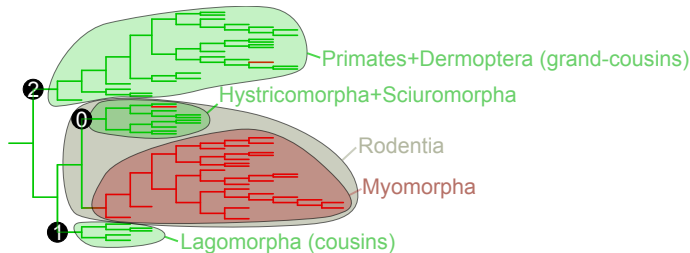
Evolution of seed match count after loss



Align 3' UTR sequences in clade of loss and sister clade.

PhyloP score: measures probability of selective conservation of seed match vs. rest of UTR.

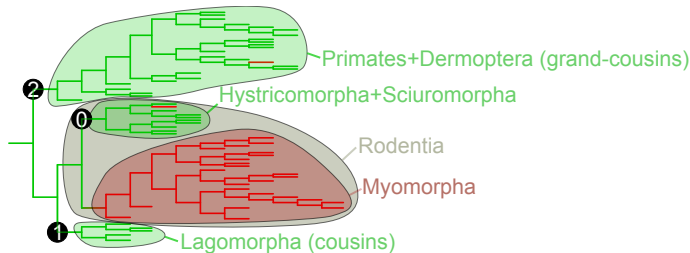
Evolution of seed match count after loss



Align 3' UTR sequences in clade of loss and sister clade.

PhyloP score: measures probability of selective conservation of seed match vs. rest of UTR in a clade vs. rest of tree.

Evolution of seed match count after loss



Align 3' UTR sequences in clade of loss and sister clade.

PhyloP score: measures probability of selective conservation of seed match vs. rest of UTR in a clade vs. rest of tree.

→ Selects seed matches with selective conservation in rodents, then compare their conservation in rodent subclades (clade of loss vs. sister clade).

Evolution of seed match count after loss

After miR-656 family loss:

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

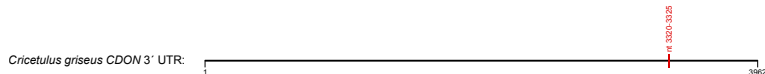
Evolution of seed match
count after loss

Conclusion

Supplementary
data

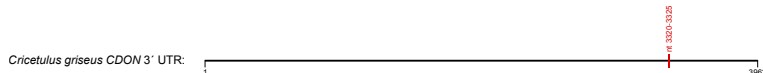
Evolution of seed match count after loss

After miR-656 family loss:



Evolution of seed match count after loss

After miR-656 family loss:

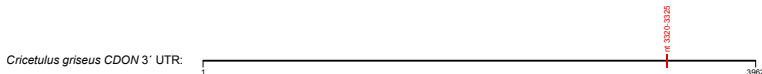


In Rodentia:

Site	Conservation vs. rest of UTR
	<i>p</i> -value
nt 3320–3325	0.02815

Evolution of seed match count after loss

After miR-656 family loss:

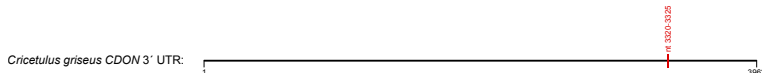


Acceleration of evolution in subclades:

Site	Clade of loss <i>p</i> -value
nt 3320–3325	0.03585

Evolution of seed match count after loss

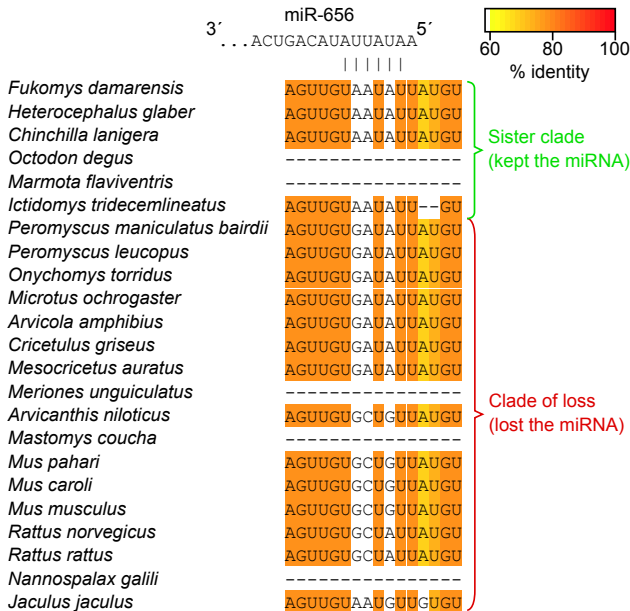
After miR-656 family loss:



Acceleration of evolution in subclades:

Site	Clade of loss <i>p</i> -value	Sister clade <i>p</i> -value
nt 3320–3325	0.03585	0.50000

Evolution of seed match count after loss



miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Evolution of seed match count after loss

After miR-656 family loss:

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

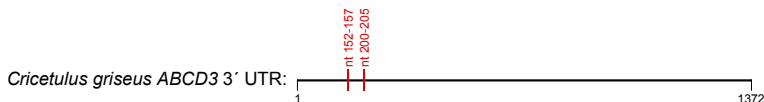
Evolution of seed match
count after loss

Conclusion

Supplementary
data

Evolution of seed match count after loss

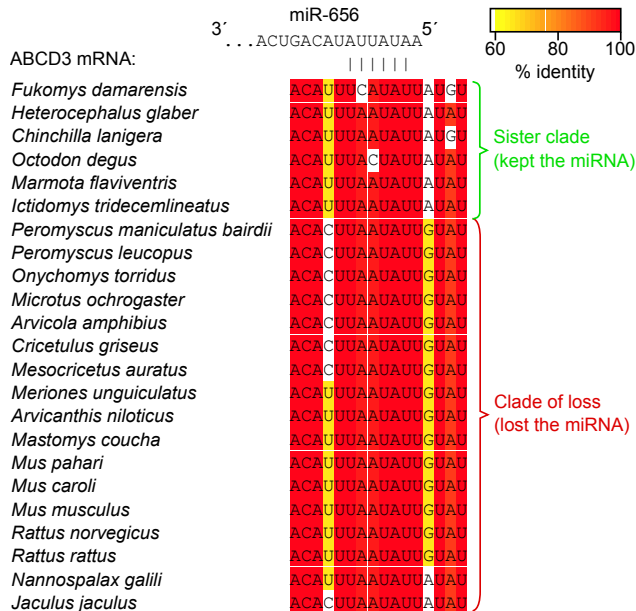
After miR-656 family loss:



Acceleration of evolution in subclades:

Site	Clade of loss <i>p</i> -value	Sister clade <i>p</i> -value
nt 152–157	0.50000	0.00864
nt 200–205	0.23280	0.50000

Evolution of seed match count after loss



miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

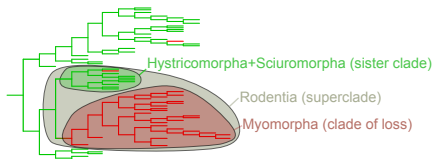
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Evolution of seed match count after loss



miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

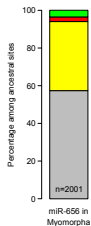
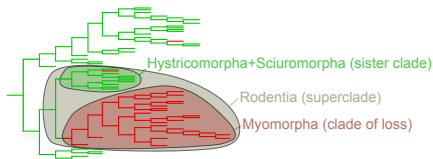
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Evolution of seed match count after loss



- : Accelerated evolution in sister clade
- : Accelerated evolution in clade of loss
- : Equally conserved in clade of loss and sister clade
- : Not significantly conserved in superclade

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

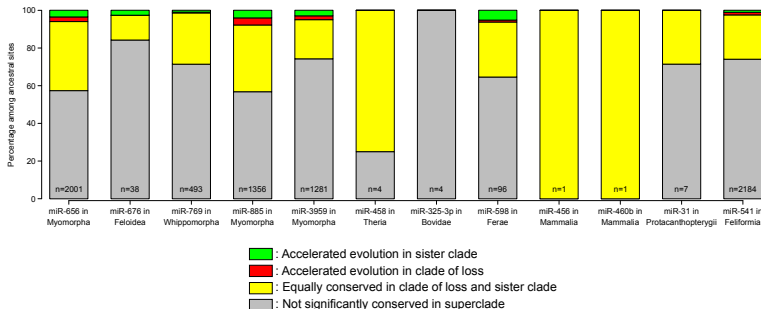
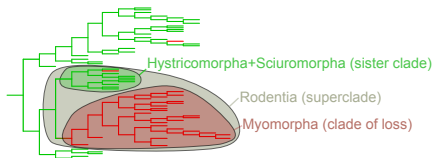
Observed losses

Evolution of seed match
count after loss

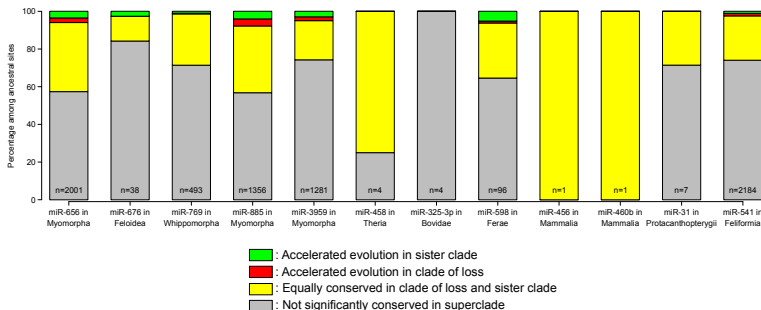
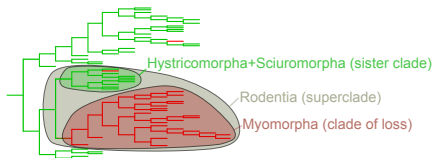
Conclusion

Supplementary
data

Evolution of seed match count after loss



Evolution of seed match count after loss



→ Most of these seed matches are conserved in a miRNA-independent manner.

Conclusion

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Conclusion

High-throughput experimental target identification captures many false positives...

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Conclusion

High-throughput experimental target identification captures many false positives...

... but comparative genomics too!

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Conclusion

High-throughput experimental target identification captures many false positives...

... but comparative genomics too!

Improvement: identification of miRNA-independent conservation of binding sites.

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Conclusion

High-throughput experimental target identification captures many false positives...

... but comparative genomics too!

Improvement: identification of miRNA-independent conservation of binding sites.

On the analyzed examples: most conserved seed matches are not conserved because of the miRNA!

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Conclusion

High-throughput experimental target identification captures many false positives...

... but comparative genomics too!

Improvement: identification of miRNA-independent conservation of binding sites.

On the analyzed examples: most conserved seed matches are not conserved because of the miRNA!

→ A large part of the miRNA literature probably needs a cleanup.

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

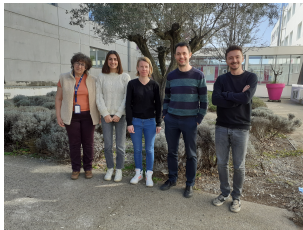
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Acknowledgments



Current members:

Isabelle Busseau
Rémy Rodriguez

Élisabeth Houbbron
Estelle Rebecq

Former members:

Sophie Mockly
Natalia Pinzón
Laura Martinez

Germain Busto
Blaise Li
Anna Sergeeva

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

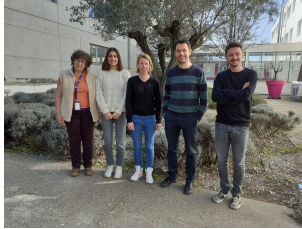
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Acknowledgments



Current members:

Isabelle Busseau
Rémy Rodriguez

Élisabeth Houbbron
Estelle Rebecq

Former members:

Sophie Mockly
Natalia Pinzón
Laura Martinez

Germain Busto
Blaise Li
Anna Sergeeva

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

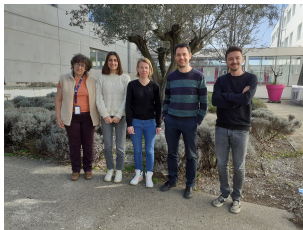
Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Acknowledgments



Current members:

Isabelle Busseau
Rémy Rodriguez

Élisabeth Houbbron
Estelle Rebecq

Former members:

Sophie Mockly
Natalia Pinzón
Laura Martinez

Germain Busto
Blaise Li
Anna Sergeeva

miRNA targets
and evolutionary
history

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

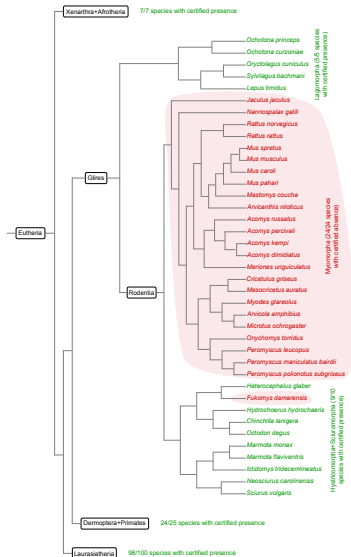
Evolution of seed match
count after loss

Conclusion

Supplementary
data

Observed losses

miR-656 family:



miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Conclusion

Supplementary
data

Observed losses

miRNA target
identification

False positives in target prediction

Binding site
evolution after
miRNA loss

Observed losses

Conclusion

Supplementary data

