

Widespread microRNA-independent conservation of presumptive microRNA binding sites in vertebrates

Hervé Seitz

November 20, 2024

Slideshow accessible for download at:

<http://www.igh.cnrs.fr/equip/Seitz/ENSL24.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

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target identification

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

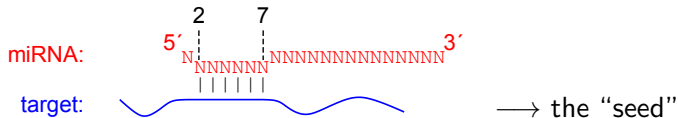
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target identification



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target identification

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target identification

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

Example: [Ecsedi et al., 2015](#)

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

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).

Example: [Chi et al., 2009](#)

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

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.

Example: [Agarwal *et al.*, 2015](#)

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

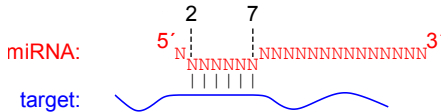
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target prediction



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

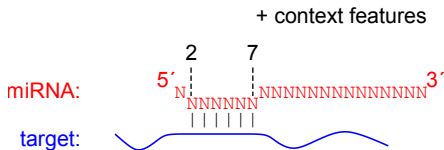
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target prediction



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

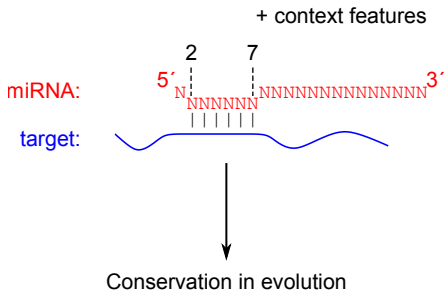
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target prediction



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

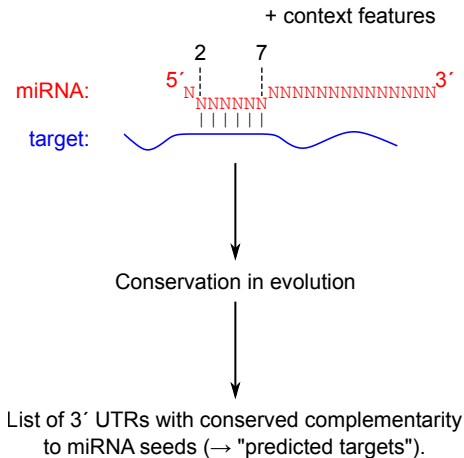
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

miRNA target prediction



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

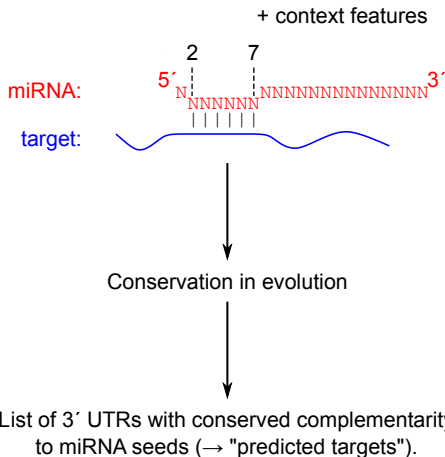
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

miRNA target prediction



Even high-throughput identification needs to assess conservation.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

False positives in target prediction

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

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

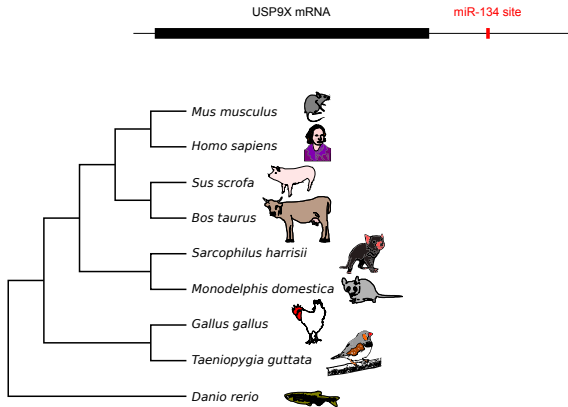
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

False positives in target prediction



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

False positives in target prediction

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

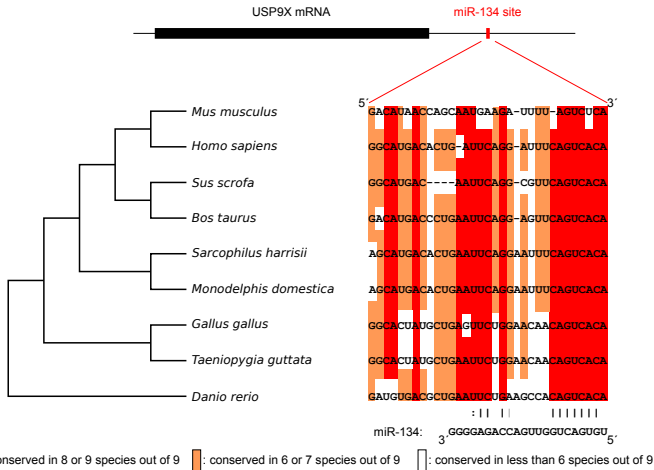
Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data



(Pinzón et al., 2017)

False positives in target prediction

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

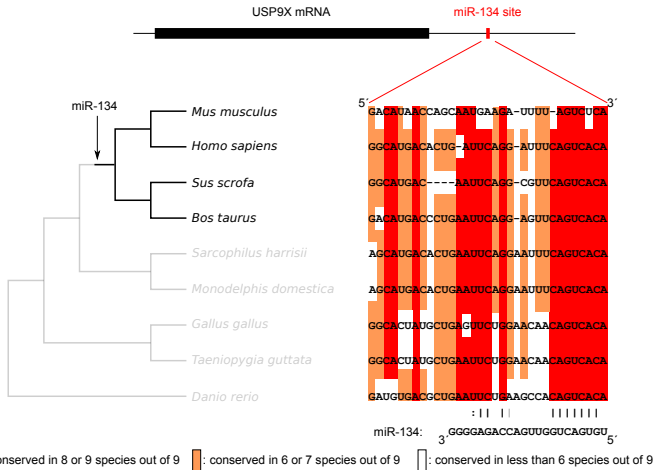
Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data



(Pinzón *et al.*, 2017)

False positives in target prediction

Seed match “overconservation” could only be precisely measured for the most recently emerged miRNAs.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

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](#)).

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

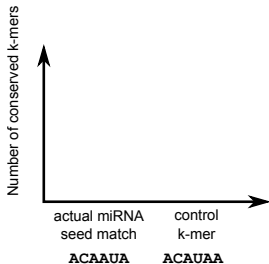
Generalization:
defining regulatory
targets

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](#)).



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

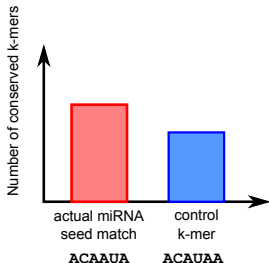
Generalization: defining regulatory targets

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](#)).



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

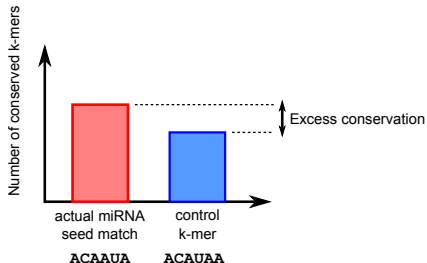
Generalization: defining regulatory targets

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](#)).



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

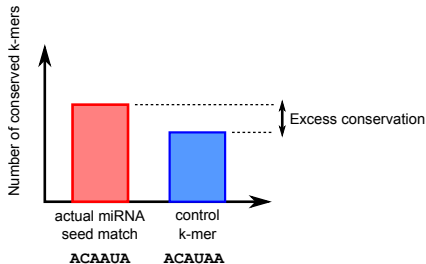
Conclusion

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).



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Binding site evolution after miRNA loss

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

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

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Binding site evolution after miRNA loss

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

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

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).

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Strategy

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

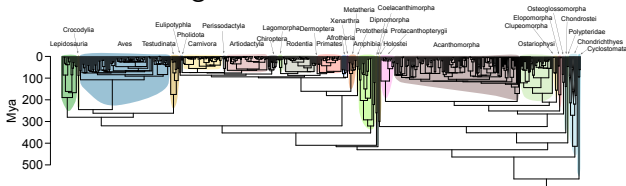
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Strategy

522 vertebrate genomes:



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

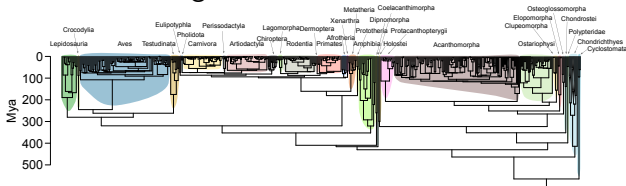
Conclusion

Supplementary
data

Strategy

Conservation of
presumptive
microRNA binding
sites

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

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

848 best annotated vertebrate miRNA families.

Strategy

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

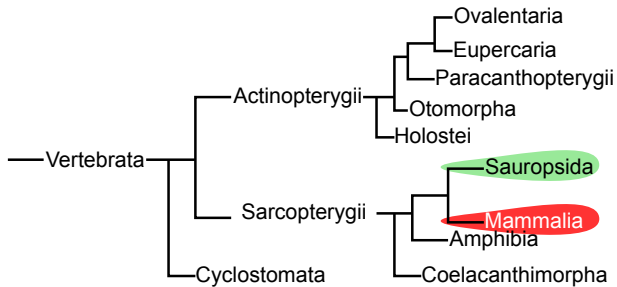
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Strategy



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

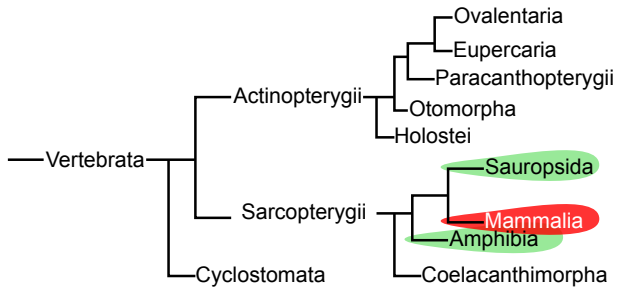
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Strategy



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

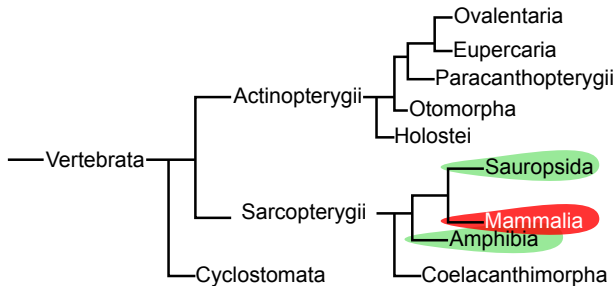
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Strategy



More convincing if many species in clade of loss.

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

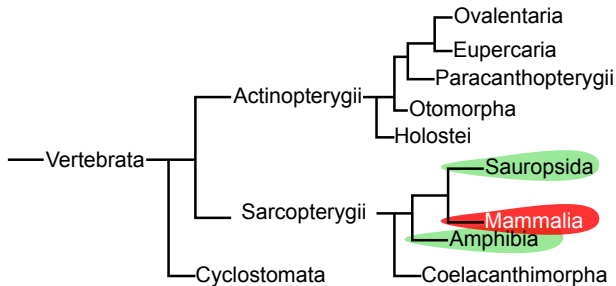
Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Strategy



More convincing if many species in clade of loss.

≥ 5 species in clade of loss; presence in sister clade + 3 closest cousin clades.

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

Conservation of
presumptive
microRNA binding
sites

Family	Clade of loss
miR-509	Cervidae+Bovidae
miR-541	Feliformia
miR-224	Mustelidae
miR-338	Lagomorpha
miR-338	Syngnathidae
miR-460	Serpentes
miR-660	Murinae
miR-880	Deomyinae+Gerbilinae
miR-452	Mustelidae
miR-456	Balistidae+Tetraodontidae
miR-296-3p	Lagomorpha
miR-2188	Eutheria
miR-2114	Phocidae
miR-656	Myomorpha
miR-654	Delphinoida
miR-460b	Mammalia
miR-460b	Serpentes
miR-676	Cervidae+Bovidae
miR-676	Feloidea
miR-187	Serpentes
miR-769	Yangochiroptera
miR-769	Whippomorpha
miR-1185	Yangochiroptera

Family	Clade of loss
miR-451a	Paralichthyidae+Pleuronectidae
miR-509	Yangochiroptera
miR-31	Protacanthopterygii
miR-31	Lophiiformes+Tetraodontiformes
miR-885	Myomorpha
miR-296-5p	Lagomorpha
miR-296-5p	Canidae
miR-412	Lagomorpha
miR-889	Muridae+Cricetidae
miR-889	Bovinae
miR-458	Theria
miR-374	Yangochiroptera
miR-513b	Bos
miR-653	Yangochiroptera
miR-1662	Endoglyptodonta
miR-297p	Pinnipedia+Musteloidea
miR-2131	Psittaciformes
miR-122b	Paralichthyidae+Pleuronectidae
miR-325-3p	Bovidae
miR-599	Batrachia
miR-599	Corvidae+Passerides
miR-599	Galloanserae

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

Conservation of
presumptive
microRNA binding
sites

Family	Clade of loss
miR-509	Cervidae+Bovidae
miR-541	Feliformia
miR-224	Mustelidae
miR-338	Lagomorpha
miR-338	Syngnathidae
miR-460	Serpentes
miR-660	Murinae
miR-880	Deomyinae+Gerbilinae
miR-452	Mustelidae
miR-456	Balistidae+Tetraodontidae
miR-296-3p	Lagomorpha
miR-2188	Eutheria
miR-2114	Phocidae
miR-656	Myomorpha
miR-654	Delphinoida
miR-460b	Mammalia
miR-460b	Serpentes
miR-676	Cervidae+Bovidae
miR-676	Feloidea
miR-187	Serpentes
miR-769	Yangochiroptera
miR-769	Whippomorpha
miR-1185	Yangochiroptera

Family	Clade of loss
miR-451a	Paralichthyidae+Pleuronectidae
miR-509	Yangochiroptera
miR-31	Protacanthopterygii
miR-31	Lophiiformes+Tetraodontiformes
miR-885	Myomorpha
miR-296-5p	Lagomorpha
miR-296-5p	Canidae
miR-412	Lagomorpha
miR-889	Muridae+Cricetidae
miR-889	Bovinae
miR-458	Theria
miR-374	Yangochiroptera
miR-513b	Bos
miR-653	Yangochiroptera
miR-1662	Endoglyptodonta
miR-297p	Pinnipedia+Musteloidea
miR-2131	Psittaciformes
miR-122b	Paralichthyidae+Pleuronectidae
miR-325-3p	Bovidae
miR-599	Batrachia
miR-599	Corvidae+Passerides
miR-599	Galloanserae

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

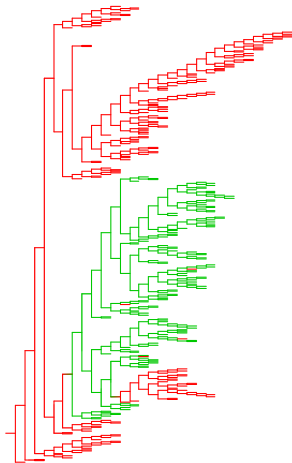
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

miR-656 family:



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

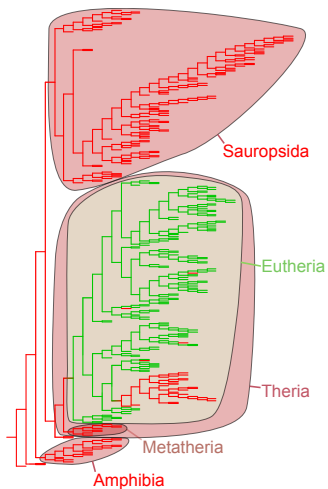
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

miR-656 family:



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

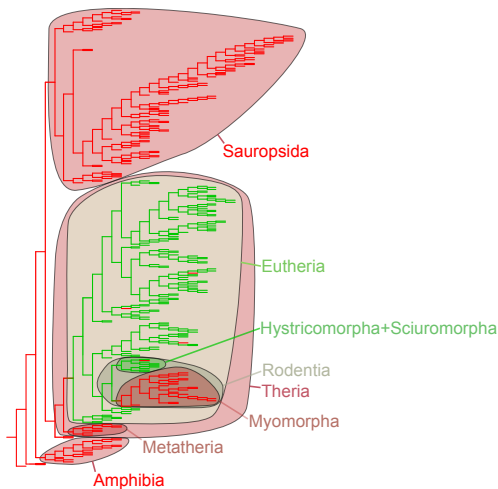
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

miR-656 family:



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

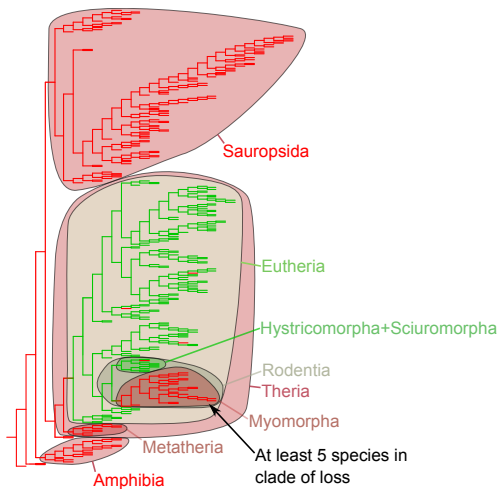
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

miR-656 family:



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

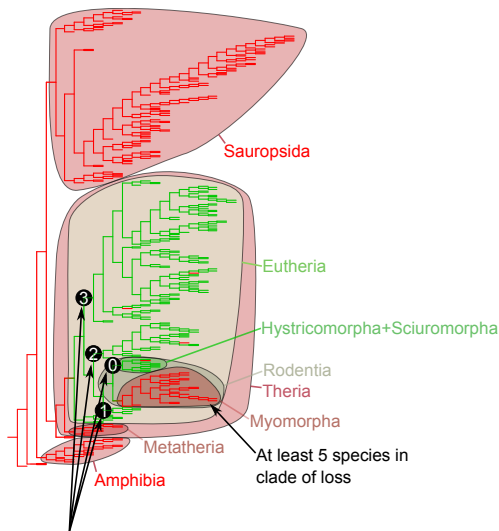
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

miR-656 family:



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

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Observed losses

miR-656 family:

Percentage identity among 470 mammals: 60 80 100



Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

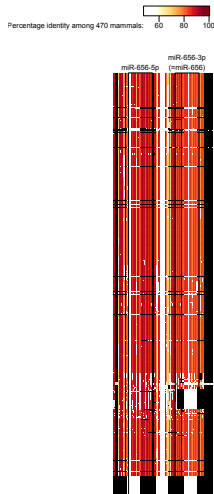
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Observed losses

miR-656 family:



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

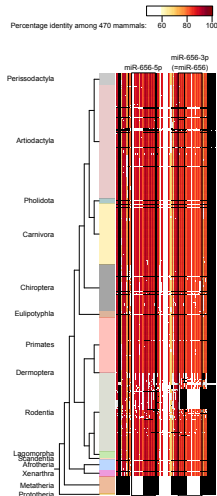
Generalization: defining regulatory targets

Conclusion

Supplementary data

Observed losses

miR-656 family:



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

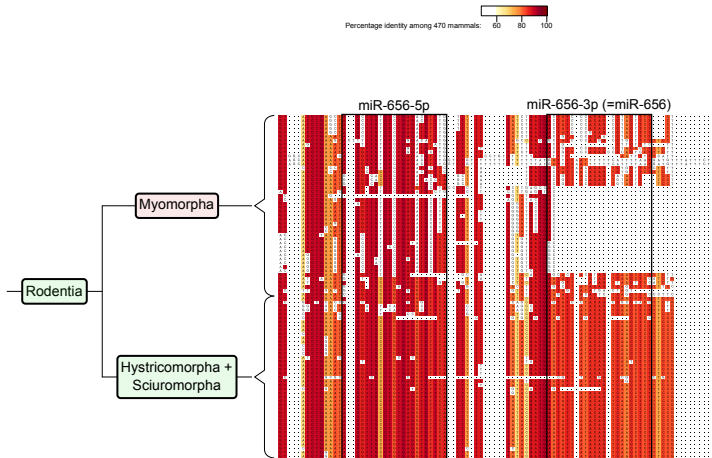
Generalization: defining regulatory targets

Conclusion

Supplementary data

Observed losses

miR-656 family:



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Evolution of seed match count after loss

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

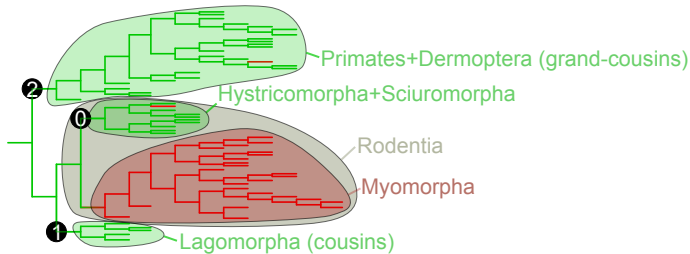
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Evolution of seed match count after loss



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

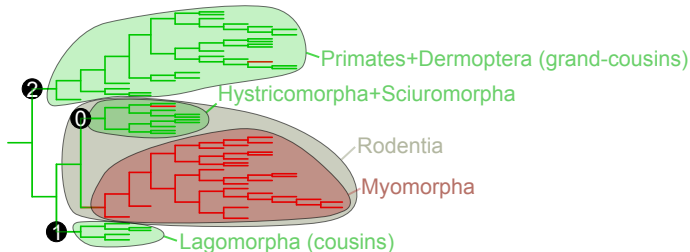
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Evolution of seed match count after loss



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

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

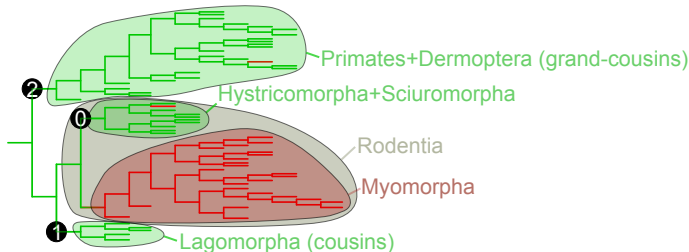
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

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.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

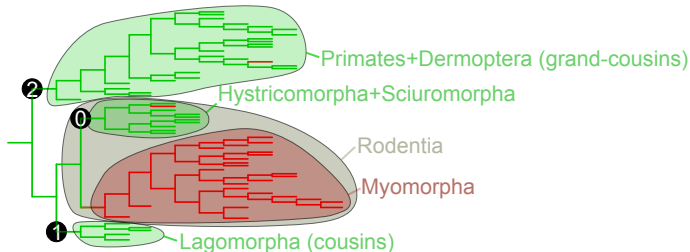
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

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.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

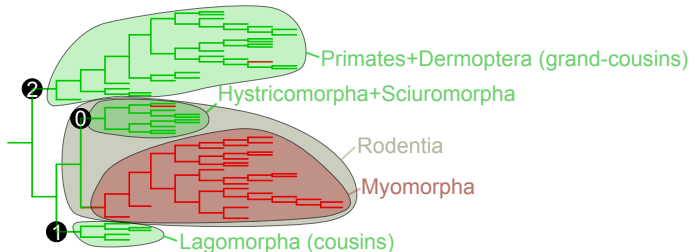
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

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).

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Evolution of seed match count after loss

After miR-656 family loss:

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

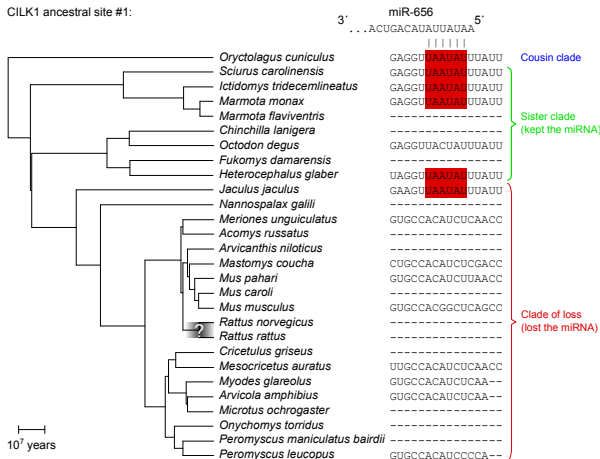
Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Evolution of seed match count after loss

After miR-656 family loss:



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

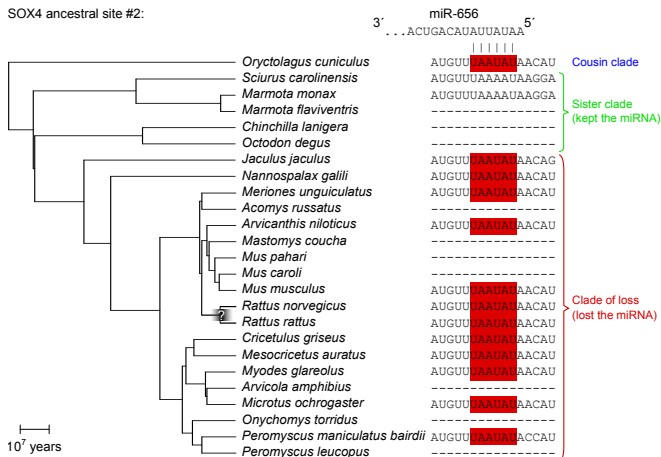
Conclusion

Supplementary data

Evolution of seed match count after loss

Conservation of presumptive microRNA binding sites

After miR-656 family loss:



miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

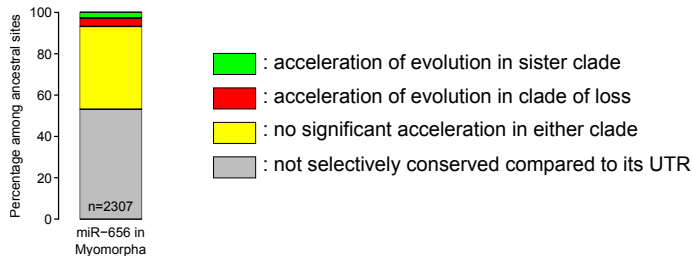
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Evolution of seed match count after loss



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

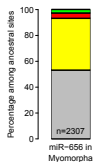
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Evolution of seed match count after loss



- acceleration of evolution in sister clade
- acceleration of evolution in clade of loss
- no significant acceleration in either clade
- not selectively conserved compared to its UTR

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

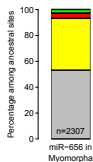
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Evolution of seed match count after loss



- acceleration of evolution in sister clade
- acceleration of evolution in clade of loss
- no significant acceleration in either clade
- not selectively conserved compared to its UTR

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

► Reasons for missing statistics

Evolution of seed match count after loss

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

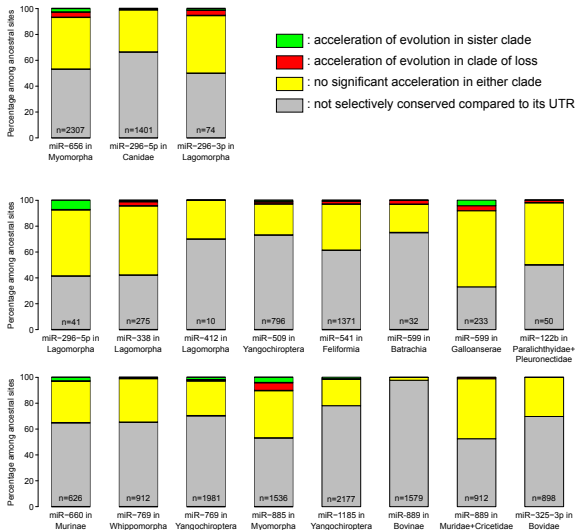
Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data



Evolution of seed match count after loss

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

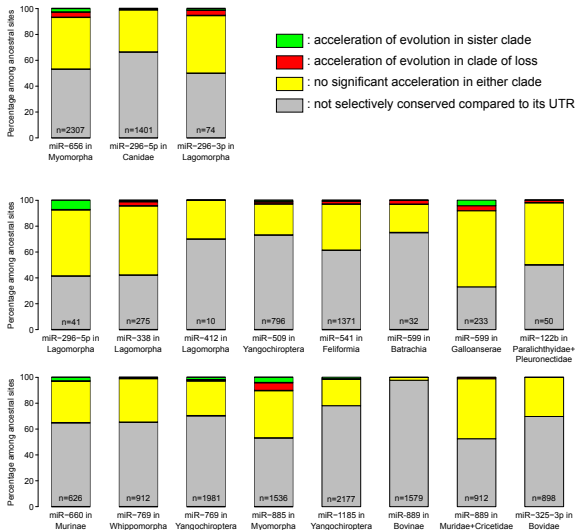
Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data



Most of these sites: conserved independently of the miRNA.

Generalization: defining regulatory targets

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Before the advent of high-throughput molecular biology
(years 2000's):

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

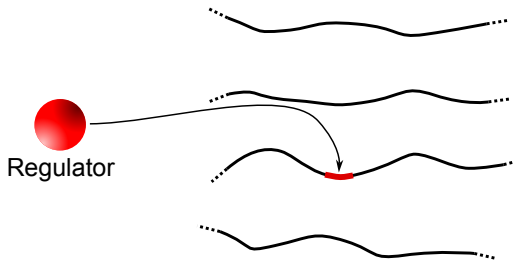
**Generalization:
defining regulatory
targets**

Conclusion

Supplementary
data

Generalization: defining regulatory targets

Before the advent of high-throughput molecular biology (years 2000's):



The “key and lock” model

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

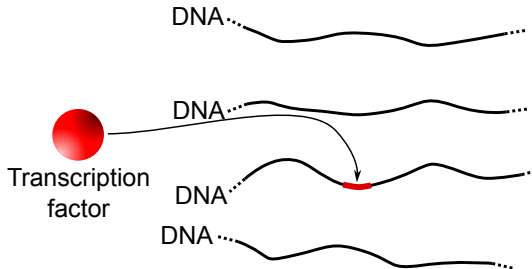
Generalization:
defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Before the advent of high-throughput molecular biology (years 2000's):



The “key and lock” model

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

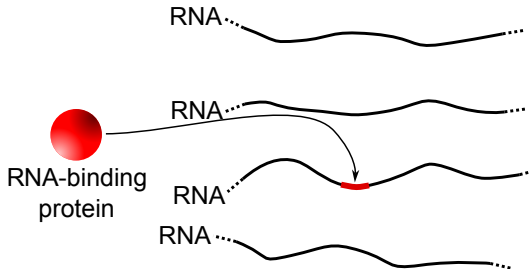
Generalization: defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Before the advent of high-throughput molecular biology (years 2000's):



The “key and lock” model

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization:
defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

**Generalization:
defining regulatory
targets**

Conclusion

Supplementary
data

Generalization: defining regulatory targets

Since the advent of high-throughput molecular biology (years 2000's):

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

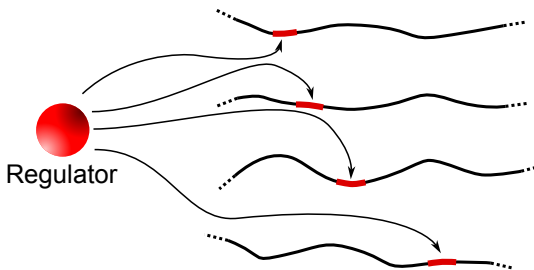
Generalization: defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Since the advent of high-throughput molecular biology (years 2000's):



Poor biochemical specificity

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy
Observed losses

Evolution of seed match count after loss

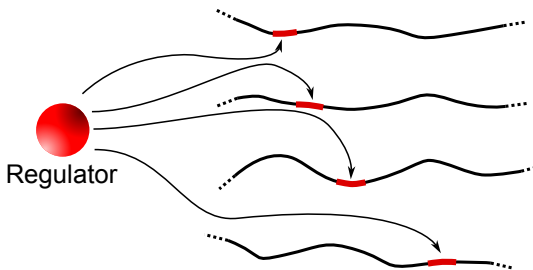
Generalization:
defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Since the advent of high-throughput molecular biology (years 2000's):



Poor biochemical specificity

ChIP-on-chip, ChIP-Seq, CLIP, ...

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy
Observed losses

Evolution of seed match count after loss

Generalization:
defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Each regulator tends to have specific biological roles (→ specific phenotypes when mutated).

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Each regulator tends to have specific biological roles (→ specific phenotypes when mutated):

- ▶ “Key and lock model”: specificity in biological response is due to a high biochemical specificity.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Each regulator tends to have specific biological roles (→ specific phenotypes when mutated):

- ▶ “Key and lock model”: specificity in biological response is due to a high biochemical specificity.
- ▶ Newer model: regulators have a poor biochemical specificity, they control phenotypes through the coordinated regulation of many targets.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization:
defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Each regulator tends to have specific biological roles (→ specific phenotypes when mutated):

- ▶ “Key and lock model”: specificity in biological response is due to a high biochemical specificity.
- ▶ Newer model: regulators have a poor biochemical specificity, they control phenotypes through the coordinated regulation of many targets.

Experimental identification of transcription factor binding sites.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization:
defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Each regulator tends to have specific biological roles (→ specific phenotypes when mutated):

- ▶ “Key and lock model”: specificity in biological response is due to a high biochemical specificity.
- ▶ Newer model: regulators have a poor biochemical specificity, they control phenotypes through the coordinated regulation of many targets.

Experimental identification of transcription factor binding sites → short motifs (≈ 6 to 8 bp), very frequent in genomes.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Generalization: defining regulatory targets

Each regulator tends to have specific biological roles (→ specific phenotypes when mutated):

- ▶ “Key and lock model”: specificity in biological response is due to a high biochemical specificity.
- ▶ Newer model: regulators have a poor biochemical specificity, they control phenotypes through the coordinated regulation of many targets.

Experimental identification of transcription factor binding sites → short motifs (≈ 6 to 8 bp), very frequent in genomes.

The “futility theorem” ([Wasserman and Sandelin, 2004](#)): most occurrences of these motifs are non-functional *in vivo*.

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Generalization: defining regulatory targets

Each regulator tends to have specific biological roles (→ specific phenotypes when mutated):

- ▶ “Key and lock model”: specificity in biological response is due to a high biochemical specificity.
- ▶ Newer model: regulators have a poor biochemical specificity, they control phenotypes through the coordinated regulation of many targets.

Experimental identification of transcription factor binding sites → short motifs (≈ 6 to 8 bp), very frequent in genomes.

The “futility theorem” ([Wasserman and Sandelin, 2004](#)): most occurrences of these motifs are non-functional *in vivo*.
→ Newer model: in fact many of them bind the regulator *in vivo*.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets

Among regulators: miRNAs have a “digital” target recognition rule.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization:
defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets

Conservation of presumptive microRNA binding sites

Among regulators: miRNAs have a “digital” target recognition rule.

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Regulator	Functional target(s)
<i>Mus musculus</i> miR-155	<i>AID</i> and <i>SOCS1</i>
<i>Cænorhabditis elegans</i> <i>let-7</i>	<i>lin-41</i>
<i>C. elegans</i> miR-791	<i>akap-1</i> and <i>cah-3</i>
<i>M. musculus</i> miR-200	<i>Zeb1</i>
<i>C. elegans</i> miR-35	<i>egl-1</i>
<i>Drosophila melanogaster</i> miR-iab-4/8	<i>hth</i>
<i>M. musculus</i> miR-140-5p	<i>Wnt11</i>
<i>Oryzias latipes</i> miR-202	<i>Tead3b</i>

Consequential vs. inconsequential targets

Some evidence suggests that other regulators also have consequential and inconsequential targets:

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization:
defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets

Some evidence suggests that other regulators also have consequential and inconsequential targets:

Many experimentally-identified binding sites are not under a selective pressure:

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets

Some evidence suggests that other regulators also have consequential and inconsequential targets:

Many experimentally-identified binding sites are not under a selective pressure:

- For transcription factors: Odom *et al.*, 2007, Schmidt *et al.*, 2010;

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets

Some evidence suggests that other regulators also have consequential and inconsequential targets:

Many experimentally-identified binding sites are not under a selective pressure:

- ▶ For transcription factors: Odom *et al.*, 2007, Schmidt *et al.*, 2010;
- ▶ For RNA-binding proteins: Litterman *et al.*, 2019.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets

Some evidence suggests that other regulators also have consequential and inconsequential targets:

Many experimentally-identified binding sites are not under a selective pressure:

- ▶ For transcription factors: Odom *et al.*, 2007, Schmidt *et al.*, 2010;
- ▶ For RNA-binding proteins: Litterman *et al.*, 2019.

—→ An even-newer model: regulators are specific, but not just through biochemical recognition.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

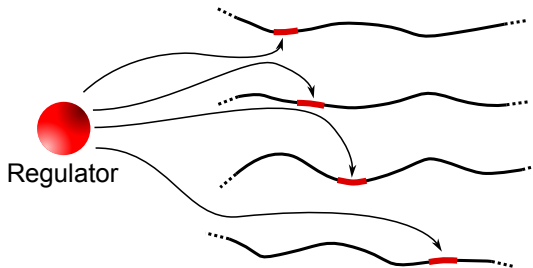
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

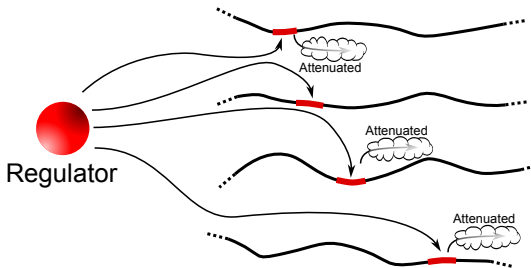
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

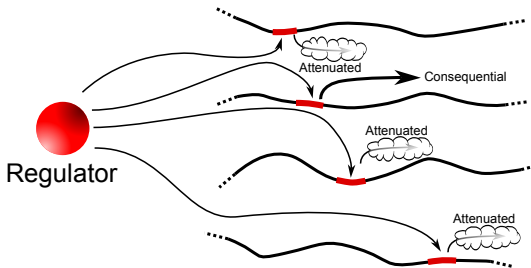
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets



Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

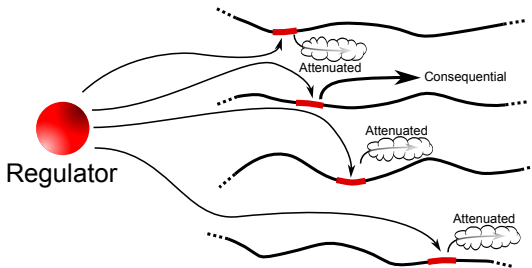
Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Consequential vs. inconsequential targets



Reviewed in [Seitz, 2024](#).

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Conclusion

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Conclusion

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

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Conclusion

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

... but comparative genomics too!

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

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.

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy
Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

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!

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy
Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

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.

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy
Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Conclusion

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Conclusion

May be generalized to other regulators (transcription factors, RNA-binding proteins).

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Conclusion

May be generalized to other regulators (transcription factors, RNA-binding proteins).

In vivo veritas: genome editing may correct a lot of false beliefs.

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Conclusion

May be generalized to other regulators (transcription factors, RNA-binding proteins).

In vivo veritas: genome editing may correct a lot of false beliefs.

A natural antagonism between developmental transitions and homeostasis.

Conservation of presumptive microRNA binding sites

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data

Conclusion

May be generalized to other regulators (transcription factors, RNA-binding proteins).

In vivo veritas: genome editing may correct a lot of false beliefs.

A natural antagonism between developmental transitions and homeostasis. Gene regulation should not only be described at the molecular scale.

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Acknowledgments



Current members:

Isabelle Busseau
Estelle Rebecq

Élisabeth Houbron
Pauline De Sousa

Rémy Rodriguez

Former members:

Sophie Mockly
Natalia Pinzón
Laura Martinez

Germain Busto
Blaise Li
Anna Sergeeva

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Acknowledgments



Current members:

Isabelle Busseau
Estelle Rebecq

Élisabeth Houbron
Pauline De Sousa

Rémy Rodriguez

Former members:

Sophie Mockly
Natalia Pinzón
Laura Martinez

Germain Busto
Blaise Li
Anna Sergeeva

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Supplementary data

► Missing statistics

Conservation of
presumptive
microRNA binding
sites

miRNA target
identification

False positives in
target prediction

Binding site
evolution after
miRNA loss

Strategy

Observed losses

Evolution of seed match
count after loss

Generalization:
defining regulatory
targets

Conclusion

Supplementary
data

Evolution of seed match count after loss

Conservation of presumptive microRNA binding sites

Family	Seed	Clade of loss	Reason
miR-451a	AACCGU	Paralichthyidae+Pleuronectidae	<10 ancestral sites
miR-224	AAGUCA	Mustelidae	Under-studied sister clade
miR-338	ACAAUA	Syngnathidae	<10 ancestral sites
miR-460	ACAGCG	Serpentes	<10 ancestral sites
miR-880	ACUCCA	Deomyinae+Gerbillinae	Under-studied sister clade
miR-452	ACUGUU	Mustelidae	Under-studied sister clade
miR-456	AGGCUG	Balistidae+Tetraodontidae	Under-studied sister clade
miR-2188	AGGUCC	Eutheria	<10 ancestral sites
miR-2114	AGUCCC	Phocidae	Under-studied sister clade
miR-654	AUGUCU	Delphinoida	Under-studied sister clade
miR-460b	CCUCAU	Mammalia	<10 ancestral sites
miR-460b	CCUCAU	Serpentes	<10 ancestral sites
miR-676	CGUCCU	Cervidae+Bovidae	Under-studied sister clade
miR-676	CGUCCU	Feloidea	<10 ancestral sites
miR-187	CGUGUC	Serpentes	<10 ancestral sites
miR-509	GAUUGA	Cervidae+Bovidae	Under-studied sister clade
miR-31	GGCAAG	Protacanthopterygii	<10 ancestral sites
miR-31	GGCAAG	Lophiiformes+Tetraodontiformes	<10 ancestral sites
miR-458	UAGCUC	Theria	<10 ancestral sites
miR-374	UAUAAU	Yangochiroptera	In progress
miR-513b	UCACAA	Bos	<10 ancestral sites
miR-653	UGAAAC	Yangochiroptera	In progress
miR-1662	UGACAU	Endoglyptodonta	Under-studied sister clade
miR-297p	UGUAUG	Pinnipedia+Musteloidea	Under-studied sister clade
miR-2131	UGUUAC	Psittaciformes	Under-studied sister clade
miR-599	UUGAUA	Corvidae+Passerides	Under-studied sister clade

miRNA target identification

False positives in target prediction

Binding site evolution after miRNA loss

Strategy

Observed losses

Evolution of seed match count after loss

Generalization: defining regulatory targets

Conclusion

Supplementary data