A functional definition of "microRNA targets": current issues, biological implications, perspectives for improvement

Hervé Seitz (herve.seitz@igh.cnrs.fr)

IGH (CNRS and Montpellier university), Montpellier, France

November 8, 2021

This slideshow is available at:
http://www.igh.cnrs.fr/equip/Seitz/SRBC2021.pdf

The principles of microRNA target prediction

Inconsequential microRNA/target interactions

Comparative genomics yield false positives

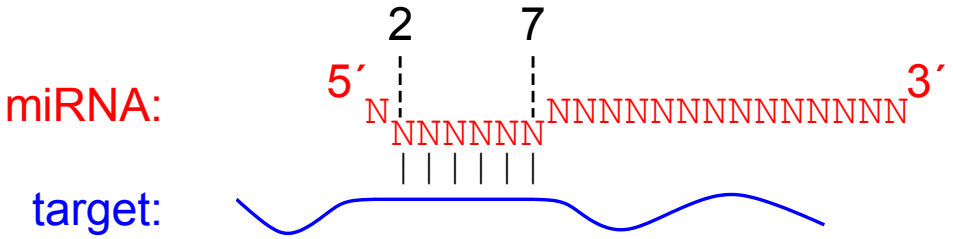
Practical implications: an example

Conclusion

Supplementary data

The principles of microRNA target prediction

The principles of microRNA target prediction



$\longrightarrow$ the "seed"

Functional definition of microRNA targets

The principles of microRNA target prediction

Inconsequential microRNA/target interactions

Comparative genomics yield false positives

Practical implications: an example

Conclusion

Supplementary data

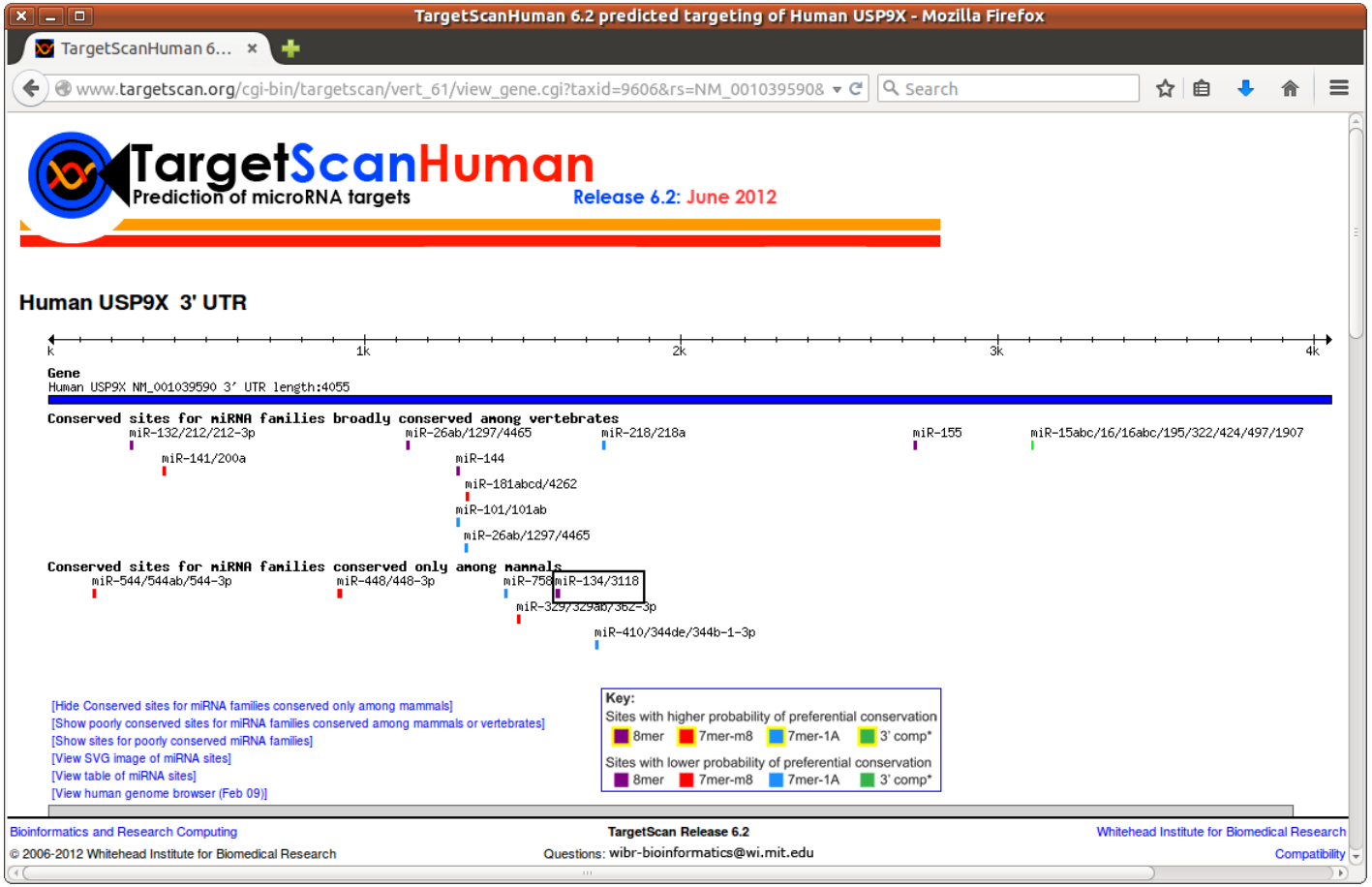
microRNA target prediction

Computational programs for target prediction: look for seed matches in 3´ UTRs, select the ones that were conserved in evolution.

microRNA target prediction

Computational programs for target prediction: look for seed matches in 3´ UTRs, select the ones that were conserved in evolution. Refinements: local AU content, position within UTR, ... (Agarwal *et al.*, 2015).

microRNA target prediction



microRNA target prediction

Computational programs for target prediction: look for seed matches in 3´ UTRs, select the ones that were conserved in evolution. Refinements: local AU content, position within UTR, ... (Agarwal *et al.*, 2015).

Such short matches are very frequent (60 % of human coding genes seem to be targeted: Friedman *et al.*, 2009).

microRNA target prediction

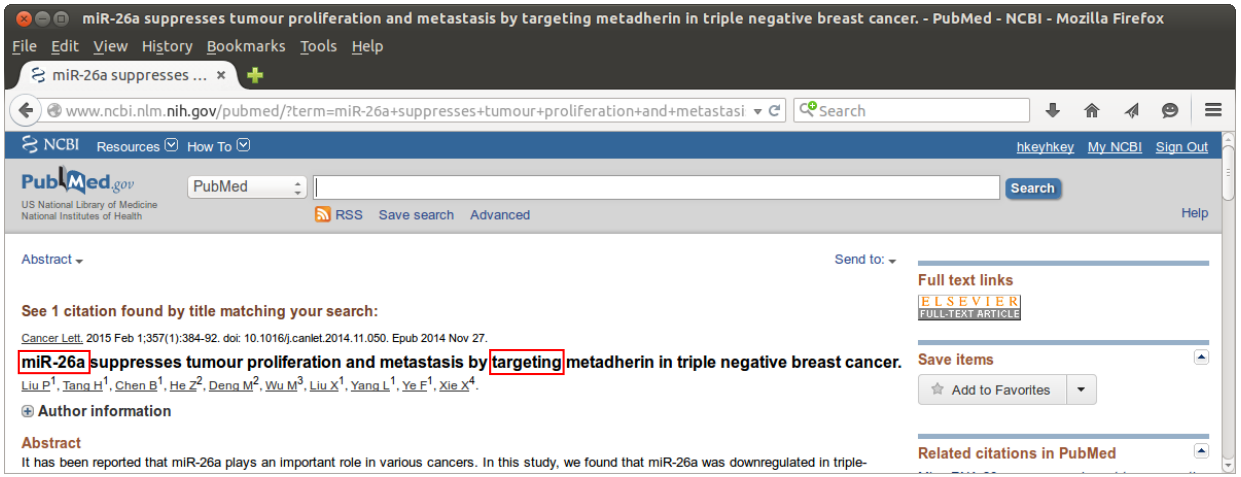
Computational programs for target prediction: look for seed matches in 3´ UTRs, select the ones that were conserved in evolution. Refinements: local AU content, position within UTR, ... (Agarwal *et al.*, 2015).

Such short matches are very frequent (60 % of human coding genes seem to be targeted: Friedman *et al.*, 2009).

microRNA target prediction

Computational programs for target prediction: look for seed matches in 3´ UTRs, select the ones that were conserved in evolution. Refinements: local AU content, position within UTR, ... (Agarwal *et al.*, 2015).

Such short matches are very frequent (60 % of human coding genes seem to be targeted: Friedman *et al.*, 2009).



microRNA target prediction

Computational programs for target prediction: look for seed matches in 3´ UTRs, select the ones that were conserved in evolution. Refinements: local AU content, position within UTR, ... (Agarwal *et al.*, 2015).

Such short matches are very frequent (60 % of human coding genes seem to be targeted: Friedman *et al.*, 2009).

$\Longrightarrow$ miRNAs are implicated in every physiological process in animals.

Inconsequential microRNA/target interactions

Inconsequential microRNA/target interactions

- miRNA-mediated repression is very modest (usually $<$ 2-fold; Baek *et al.*, 2008; Selbach *et al.*, 2008): lower than well tolerated fluctuations in gene expression (*e.g.*, haplosufficiency).

Inconsequential microRNA/target interactions

- miRNA-mediated repression is very modest (usually $<$ 2-fold; Baek *et al.*, 2008; Selbach *et al.*, 2008): lower than well tolerated fluctuations in gene expression (*e.g.*, haplosufficiency).
- For miR-223 in mouse neutrophils: inter-individual fluctuations in gene expression typically exceed miR-223-guided repression (Pinzón *et al.*, 2017).

Inconsequential microRNA/target interactions

- miRNA-mediated repression is very modest (usually $<$ 2-fold; Baek *et al.*, 2008; Selbach *et al.*, 2008): lower than well tolerated fluctuations in gene expression (*e.g.*, haplosufficiency).
- For miR-223 in mouse neutrophils: inter-individual fluctuations in gene expression typically exceed miR-223-guided repression (Pinzón *et al.*, 2017).
- In the best studied cases, miRNA-controlled phenotypes appear to be due to a single target each (Ecsedi *et al.*, 2015; Title *et al.*, 2018; Aeschimann *et al.*, 2019; Garaulet *et al.*, 2020; Yang *et al.*, 2020).

Inconsequential microRNA/target interactions

- miRNA-mediated repression is very modest (usually $<$ 2-fold; Baek *et al.*, 2008; Selbach *et al.*, 2008): lower than well tolerated fluctuations in gene expression (*e.g.*, haplosufficiency).
- For miR-223 in mouse neutrophils: inter-individual fluctuations in gene expression typically exceed miR-223-guided repression (Pinzón *et al.*, 2017).
- In the best studied cases, miRNA-controlled phenotypes appear to be due to a single target each (Ecsedi *et al.*, 2015; Title *et al.*, 2018; Aeschimann *et al.*, 2019; Garaulet *et al.*, 2020; Yang *et al.*, 2020).

$\longrightarrow$ Why have all predicted sites been conserved if they are not functional ?

Comparative genomics yield false positives

The principles of microRNA target prediction

Inconsequential microRNA/target interactions

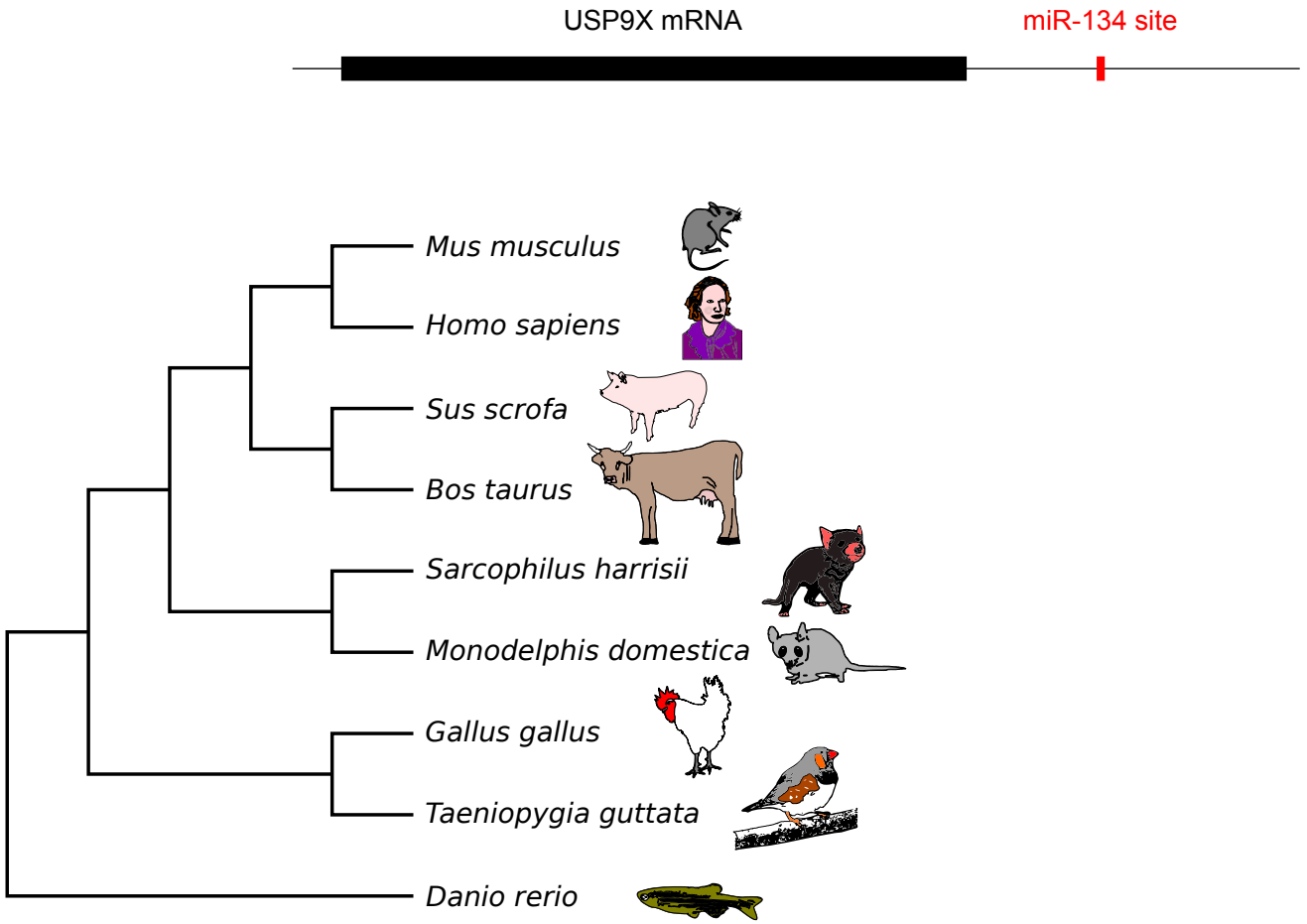
Comparative genomics yield false positives

Practical implications: an example

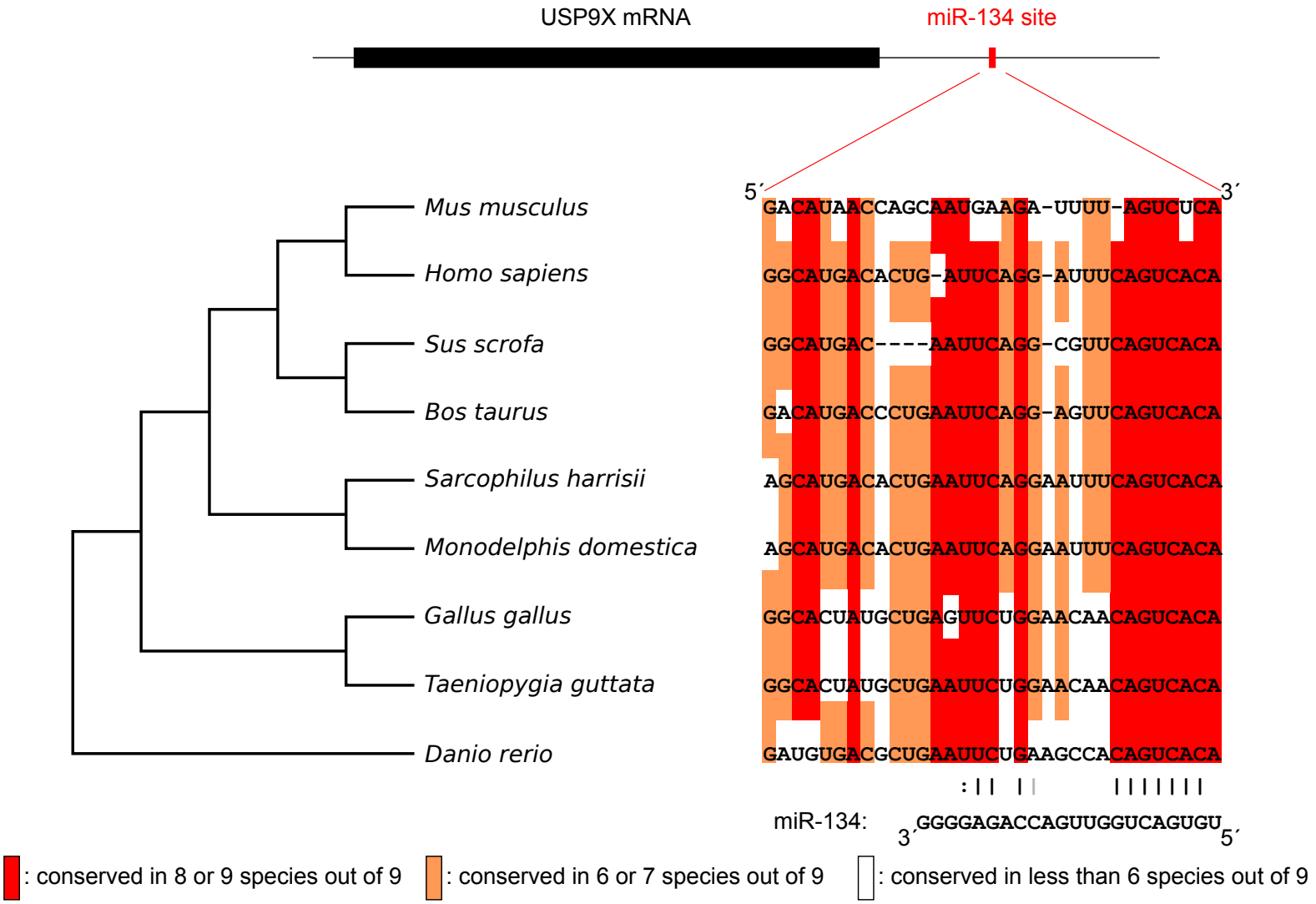
Conclusion

Supplementary data

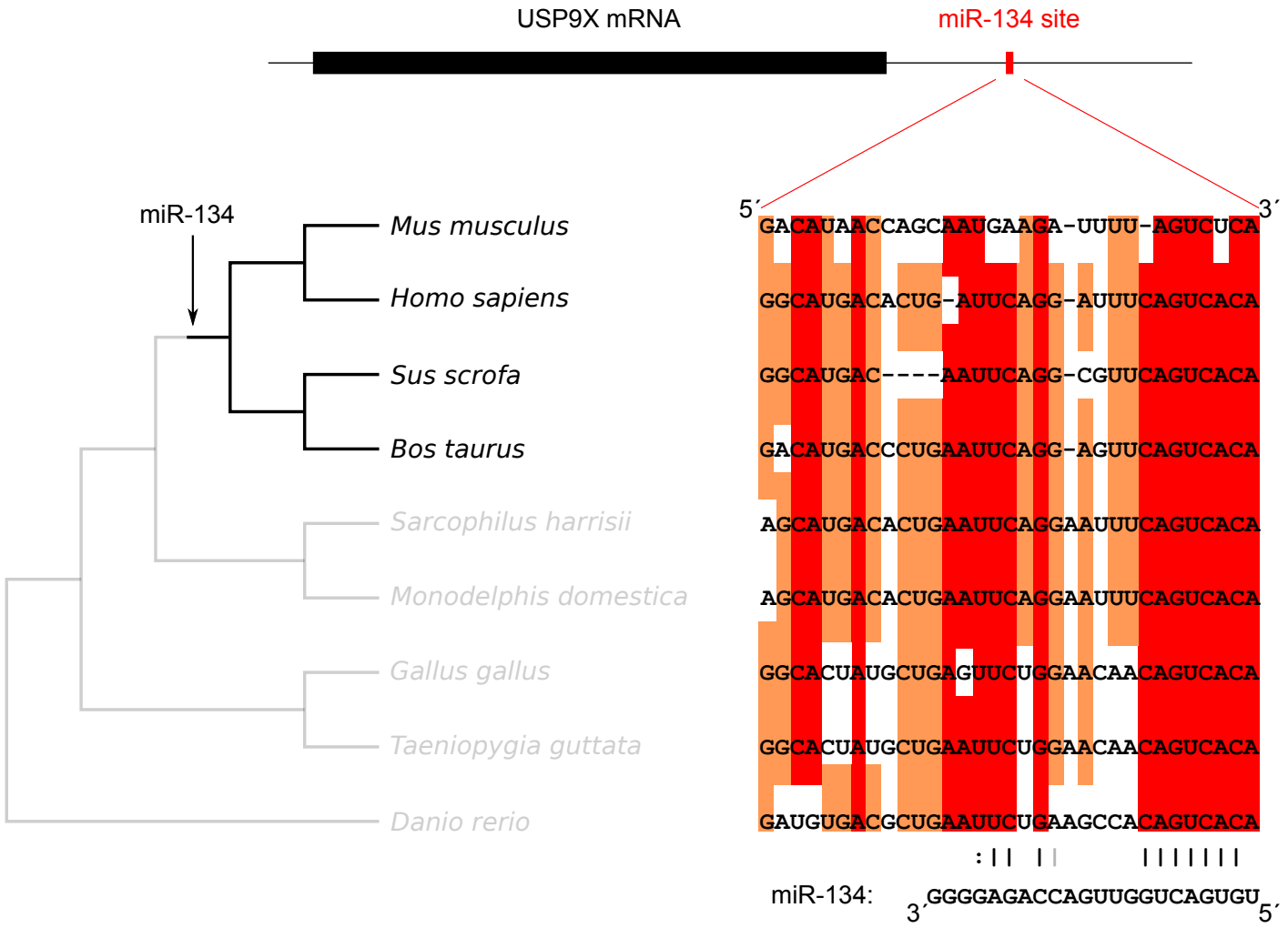
Comparative genomics yield false positives



Comparative genomics yield false positives

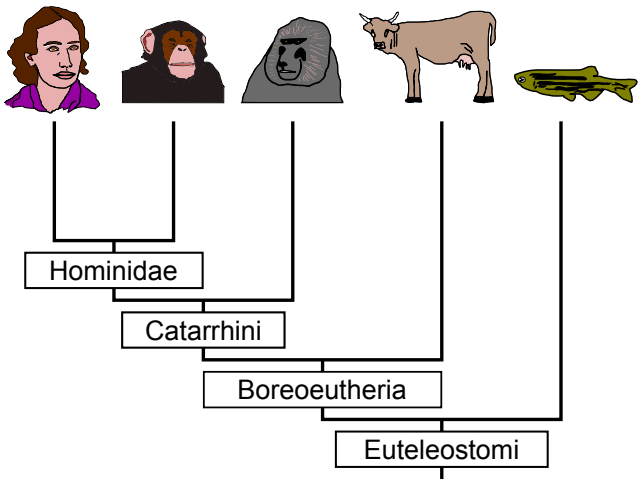


Comparative genomics yield false positives



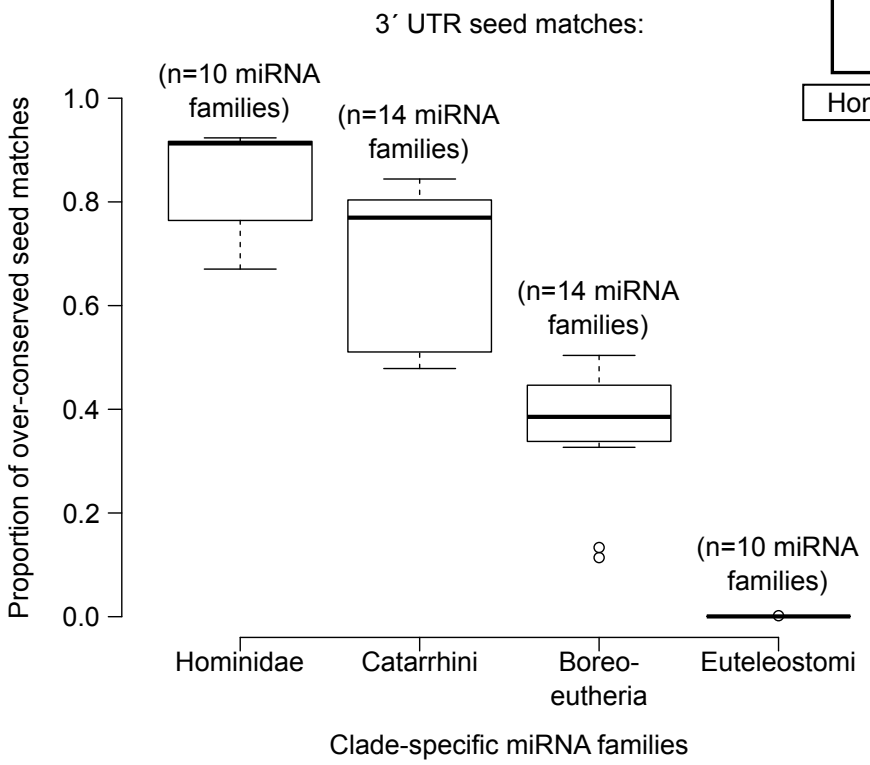
: conserved in 8 or 9 species out of 9 : conserved in 6 or 7 species out of 9 : conserved in less than 6 species out of 9

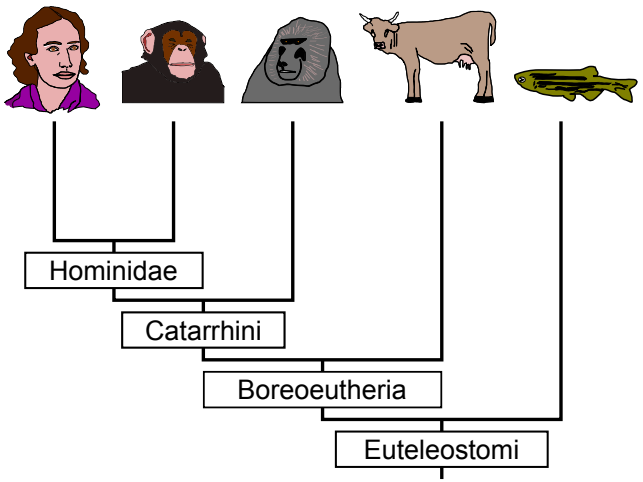
Comparative genomics yield false positives



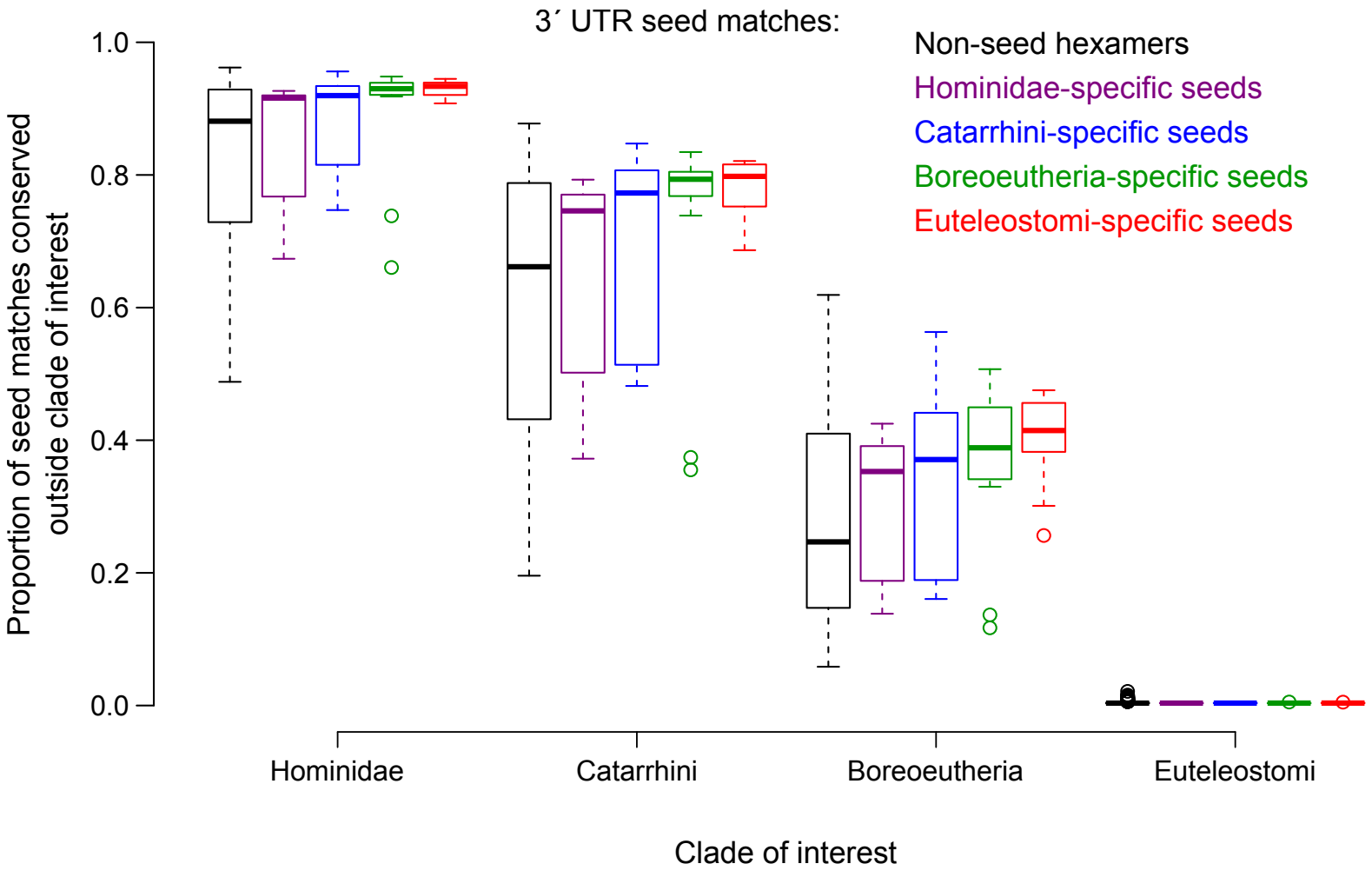
The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

Comparative genomics yield false positives

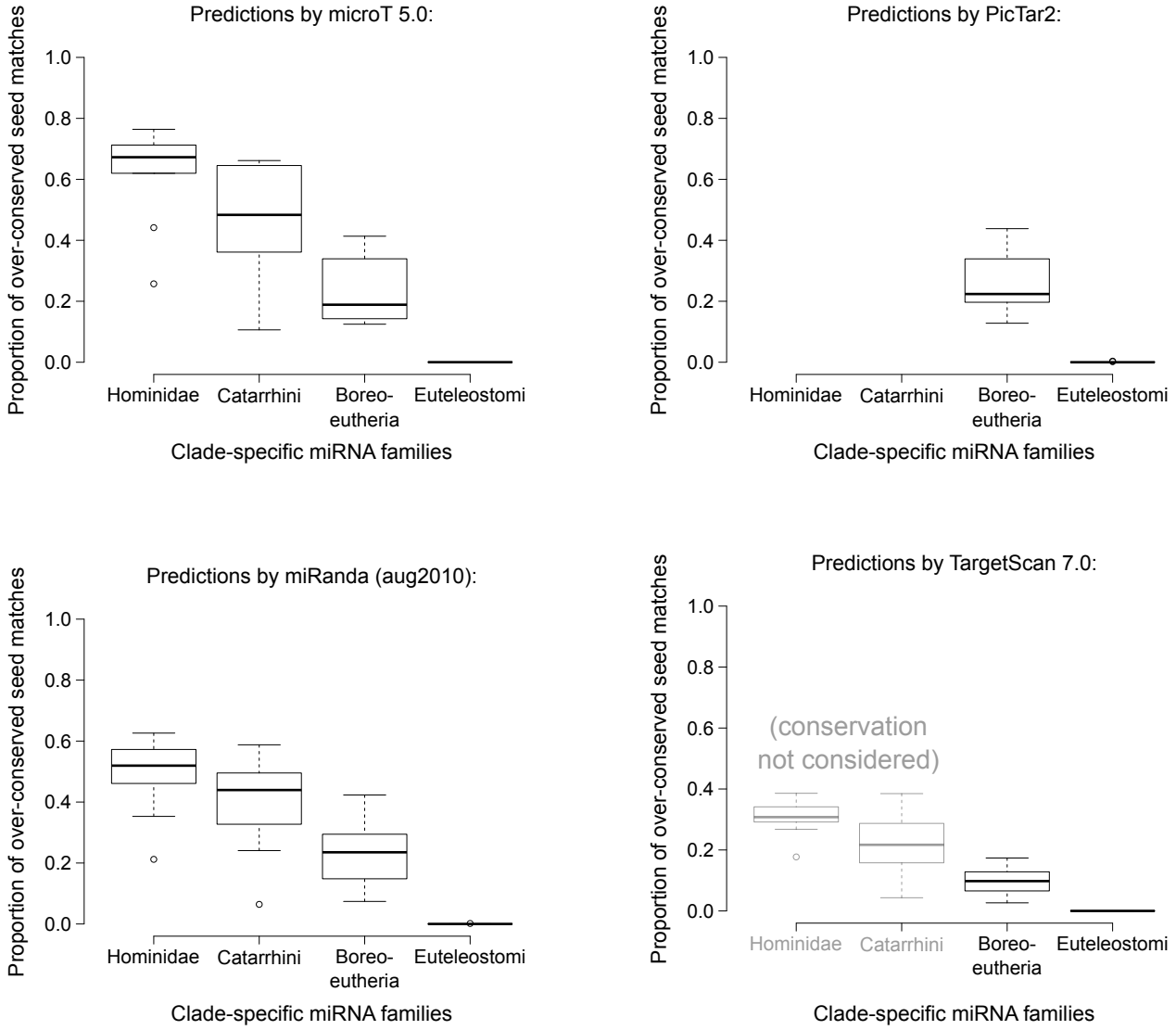




Comparative genomics yield false positives



Comparative genomics yield false positives



The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

Comparative genomics yield false positives

The principles of microRNA target prediction

Inconsequential microRNA/target interactions

Comparative genomics yield false positives

Practical implications: an example

Conclusion

Supplementary data

Comparative genomics yield false positives

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

Comparative genomics yield false positives

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

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

Practical implications: an example

The principles of microRNA target prediction

Inconsequential microRNA/target interactions

Comparative genomics yield false positives

Practical implications: an example

Conclusion

Supplementary data

Practical implications: an example

The miR-34 family is widely perceived as a tumor suppressor (*e.g.*, Slack and Chinnaiyan, 2019) ...

Practical implications: an example

The miR-34 family is widely perceived as a tumor suppressor (*e.g.*, Slack and Chinnaiyan, 2019) ...

... probably by mistake (Mockly *et al.*, in review).

Practical implications: an example

The miR-34 family is widely perceived as a tumor suppressor (*e.g.*, Slack and Chinnaiyan, 2019) ...

... probably by mistake (Mockly *et al.*, in review).

TargetscanHuman 8.0: predicted miRNA targets of miR-34-5p/449-5p — Mozilla Firefox

www.targetscan.org/cgi-bin/targetscan/vert_80/targetscan.cgi?species=Human&gid=&mir_sc=miR-34-5p%2F449-5p&mir_c=&mir_nc=&mir_vnc=

Human | miR-34-5p/449-5p

754 transcripts with conserved sites, containing a total of **838** conserved sites and **389** poorly conserved sites.

Please note that these predicted targets include some false positives. [Read more]

Genes with only poorly conserved sites are not shown. [View top predicted targets, irrespective of site conservation]

Table sorted by cumulative weighted context++ score [Sort table by predicted occupancy] [Sort table by aggregate P_{CT}]

The table shows at most one transcript per gene, selected for being the most prevalent, based on 3P-seq tags. [Download table]

Target gene	Representative transcript ENST0000...	Gene name	Number of 3P-seq tags supporting UTR + 5	Link to sites in UTRs	Conserved sites				Poorly conserved sites				6mer sites	Representative miRNA	Predicted occupancy			Cumulative weighted context++ score	Total context++ score	Aggregate P_{CT}	Previous TargetScan publication(s)
					total	8mer	7mer-m8	7mer-A1	total	8mer	7mer-m8	7mer-A1			mod miRNA	high miRNA	trans-fected miRNA				
MDM4	0391947.2	Mdm4 p53 binding protein homolog (mouse)	443	Sites in UTR	**4**	3	0	1	**3**	2	0	1	1	hsa-miR-34c-5p	0.1129	0.7383	3.1800	-2.38	-2.56	0.94	2011, '15
HCN3	0368358.3	hyperpolarization activated cyclic nucleotide-gated potassium channel 3	41	Sites in UTR	**4**	4	0	0	**2**	1	0	1	0	hsa-miR-449b-5p	0.1128	0.6855	2.4728	-1.57	-1.57	> 0.99	2011, '15
FAM76A	0373954.6	family with sequence similarity 76, member A	173	Sites in UTR	**3**	2	0	1	**0**	0	0	0	0	hsa-miR-449b-5p	0.0862	0.5203	1.7444	-1.47	-1.48	> 0.99	2007, '09, '11, '15
SCN2B	0278947.5	sodium channel, voltage-gated, type II, beta subunit	5	Sites in UTR	**3**	2	0	1	**1**	0	0	1	1	hsa-miR-34c-5p	0.1257	0.7301	2.3678	-1.29	-1.29	> 0.99	2005, '07, '09, '11, '15
SYT1	0457153.2	synaptotagmin I	13	Sites in UTR	**4**	1	2	1	**0**	0	0	0	2	hsa-miR-449b-5p	0.1040	0.6559	2.5358	-1.24	-1.24	> 0.99	2005, '07, '09, '11, '15
FAM167A	0284486.4	family with sequence similarity 167, member A	85	Sites in UTR	**1**	1	0	0	**3**	3	0	0	1	hsa-miR-449b-5p	0.0910	0.5730	2.2638	-1.23	-1.23	0.67	2005, '07, '09, '11, '15
RRAS	0246792.3	related RAS viral (r-ras) oncogene homolog	267	Sites in UTR	**1**	1	0	0	**1**	0	0	1	0	hsa-miR-34c-5p	0.0461	0.2752	0.9159	-1.16	-1.16	0.88	2005, '07, '09, '11, '15
VAMP2	0404970.3	vesicle-associated membrane protein 2 (synaptobrevin 2)	1840	Sites in UTR	**3**	1	2	0	**0**	0	0	0	1	hsa-miR-34a-5p	0.0822	0.4989	1.8318	-1.08	-1.11	0.98	2003, '05, '07, '09, '11,

Bioinformatics and Research Computing
© 2006-2021 Whitehead Institute for Biomedical Research

TargetScan Release 8.0
Questions: barc@wi.mit.edu

Whitehead Institute for Biomedical Research
Compatibility

Conclusion: revisiting miRNA target definition

The principles of microRNA target prediction

Inconsequential microRNA/target interactions

Comparative genomics yield false positives

Practical implications: an example

Conclusion

Supplementary data

Conclusion: revisiting miRNA target definition

Every measurable change in gene expression does not translate into a macroscopic, evolutionarily selectable phenotype.

The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

Conclusion: revisiting miRNA target definition

Every measurable change in gene expression does not translate into a macroscopic, evolutionarily selectable phenotype.

Even phylogenetic conservation is not a great help.

The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

Conclusion: revisiting miRNA target definition

Every measurable change in gene expression does not translate into a macroscopic, evolutionarily selectable phenotype.

Even phylogenetic conservation is not a great help: many RNA-binding protein recognition elements will fortuitously match miRNA seeds:

- $\approx$ 1,500 annotated human miRNA families
- out of $4^6 = 4096$ possible hexamers.

Conclusion: revisiting miRNA target definition

Every measurable change in gene expression does not translate into a macroscopic, evolutionarily selectable phenotype.

Even phylogenetic conservation is not a great help: many RNA-binding protein recognition elements will fortuitously match miRNA seeds:

- $\approx$ 1,500 annotated human miRNA families
- out of $4^6 = 4096$ possible hexamers.

Potential improvements:

- excluding sites more conserved than the cognate miRNA family;
- evaluating phenotypic robustness to expression fluctuations;
- "experimental validation" is a genetic knockout, not a reporter assay.

Conclusion: revisiting miRNA target definition

A central feature of biological systems: their robustness to external insults. Hard to reconcile with the extreme sensitivity required for fine-tuning (the "butterfly effect" has probably been counter-selected).

Acknowledgements

Isabelle Busseau
Élisabeth Houbron
Laura Martinez
Natalia Pinzón
Germain Busto
Blaise Li
Sophie Mockly
Anna Sergeeva

The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

Acknowledgements

Isabelle Busseau
Élisabeth Houbron
Laura Martinez
Natalia Pinzón
Germain Busto
Blaise Li
Sophie Mockly
Anna Sergeeva

Supplementary data

- Robustness of biological pathways
- miR-223-guided repression vs. inter-individual variability
- miR-34 and cell proliferation

Biological robustness

$$\text{Substrate} \xrightarrow{\text{enzyme 1}} \text{Product 1} \xrightarrow{\text{enzyme 2}} \text{Product 2} \xrightarrow{\text{enzyme 3}} \text{Product 3}$$

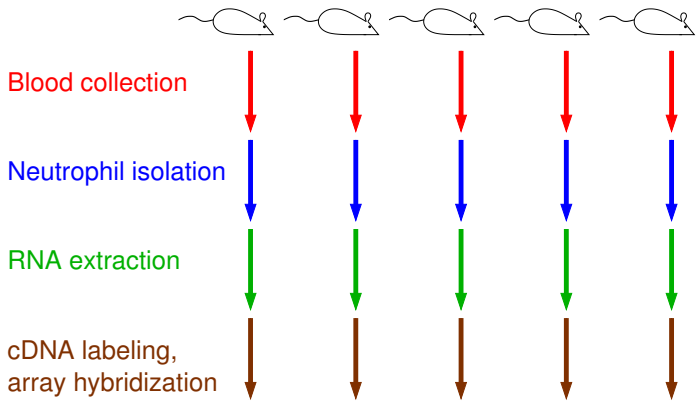
miR-223-guided repression vs. inter-individual variability

miR-223-guided repression vs. inter-individual variability

Baek *et al.*, 2008: quantification of miR-223-mediated repression in mouse neutrophils.

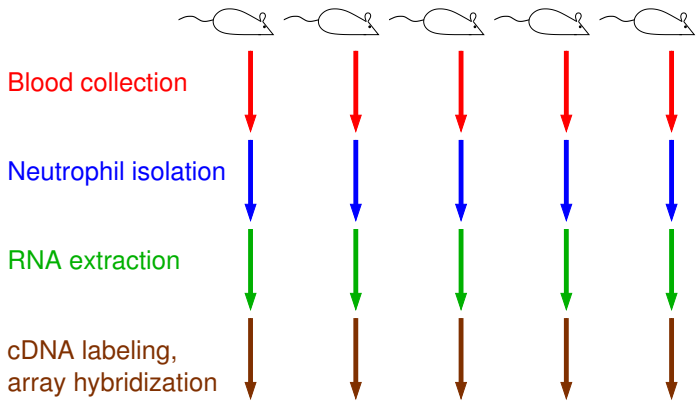
miR-223-guided repression vs. inter-individual variability

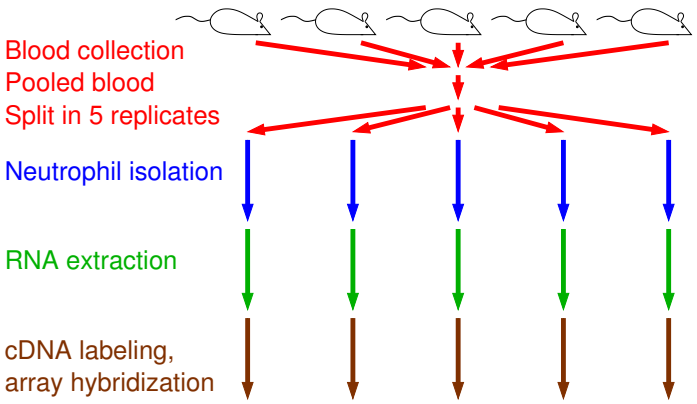
Baek *et al.*, 2008: quantification of miR-223-mediated repression in mouse neutrophils.



miR-223-guided repression vs. inter-individual variability

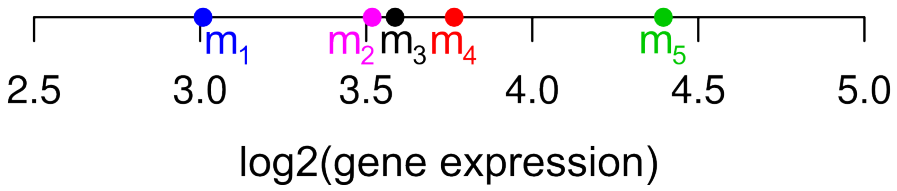
Baek *et al.*, 2008: quantification of miR-223-mediated repression in mouse neutrophils.



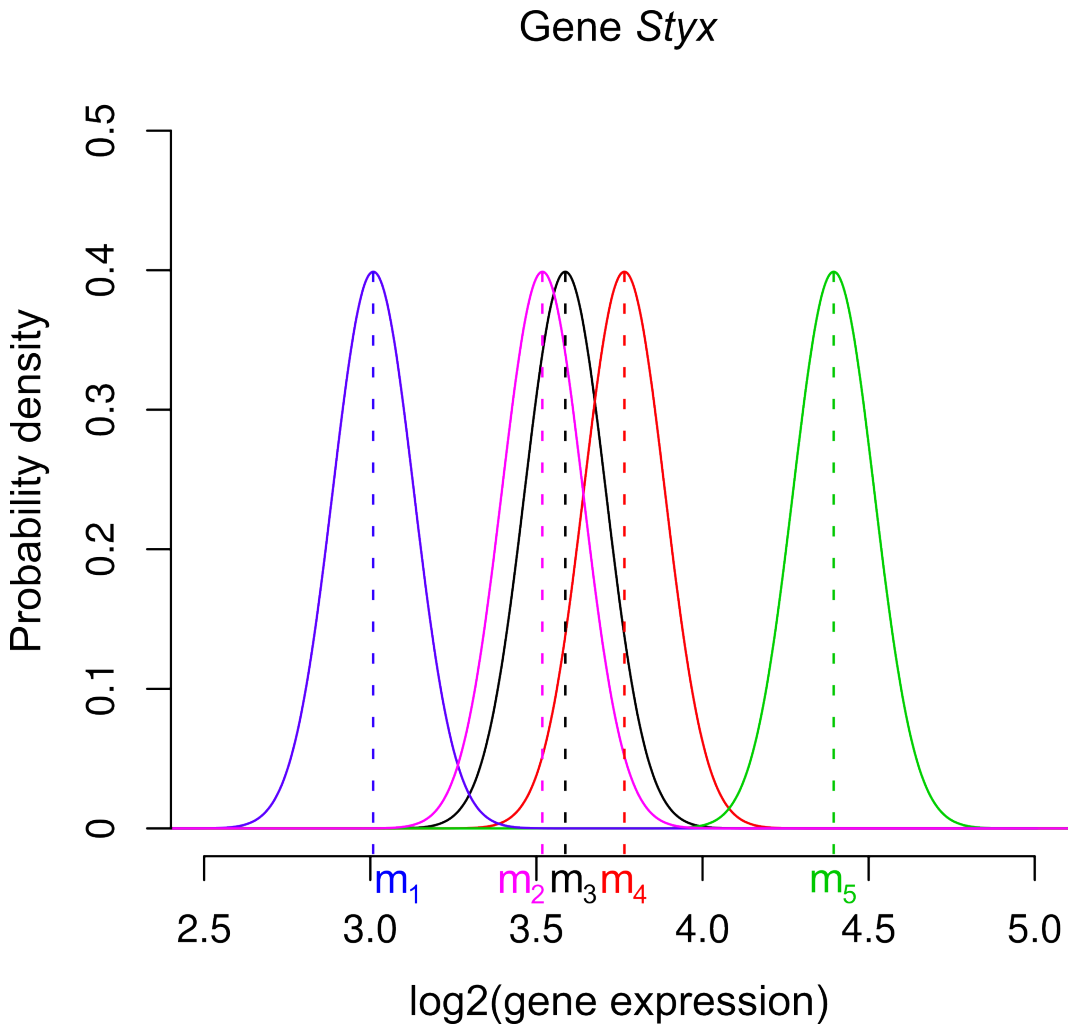


miR-223-guided repression vs. inter-individual variability

Gene *Styx*

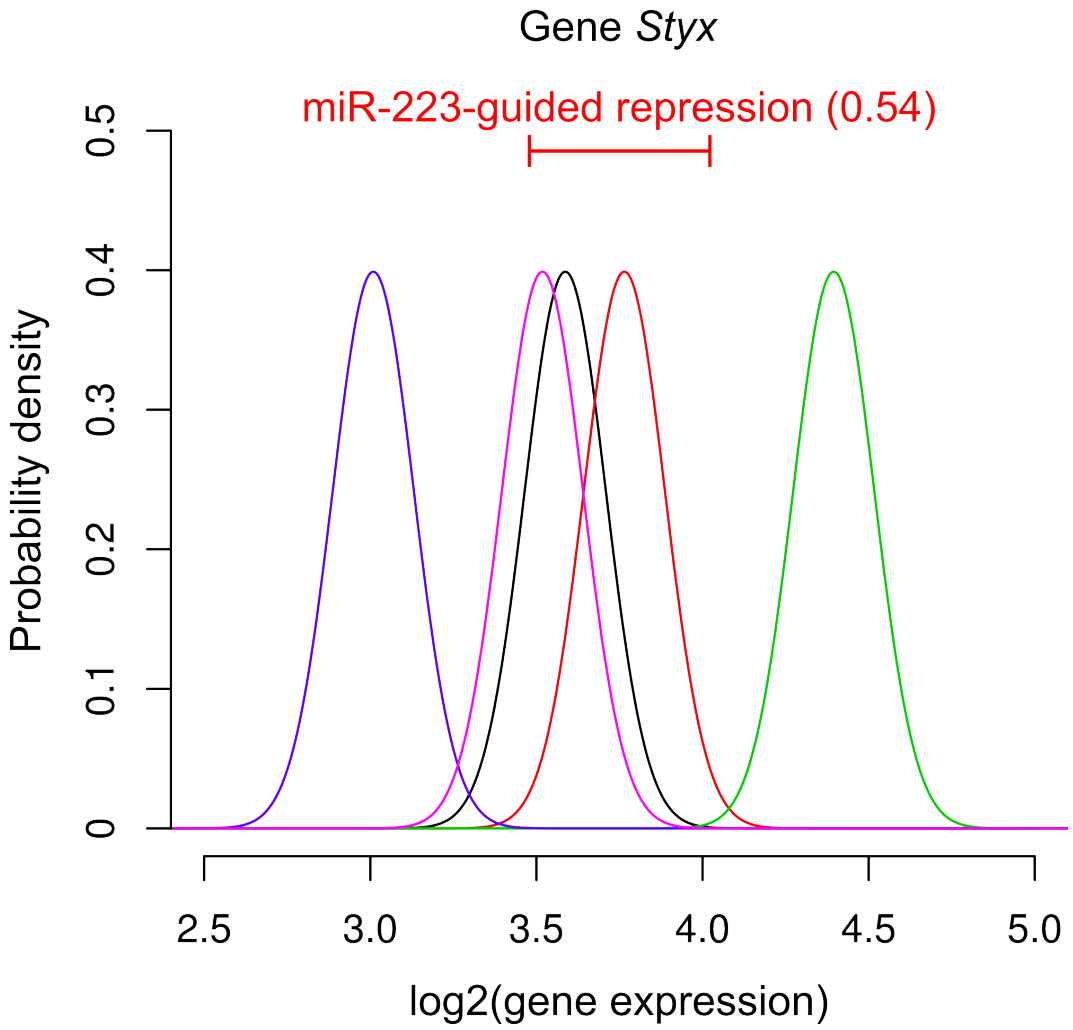


miR-223-guided repression vs. inter-individual variability



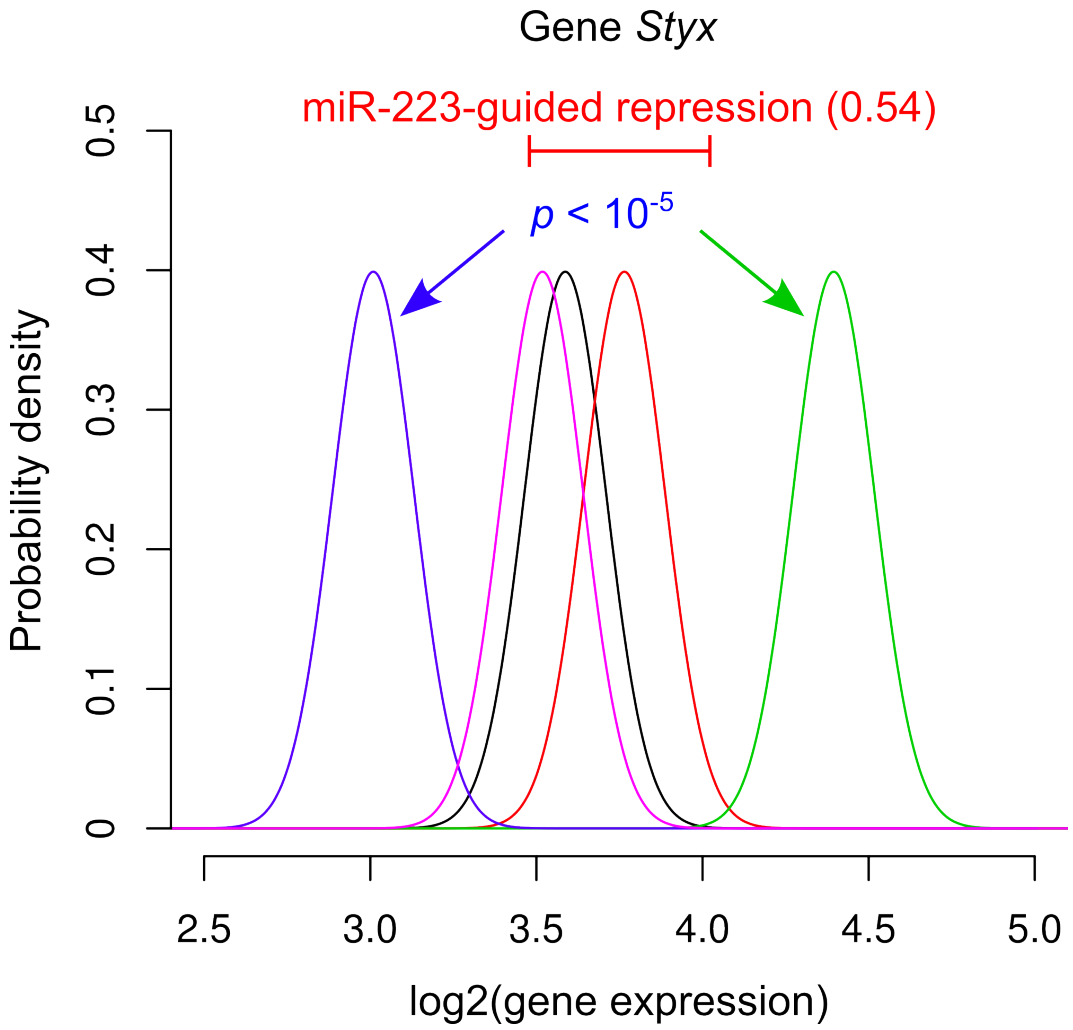
The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

miR-223-guided repression vs. inter-individual variability



The principles of microRNA target prediction
Inconsequential microRNA/target interactions
Comparative genomics yield false positives
Practical implications: an example
Conclusion
Supplementary data

miR-223-guided repression vs. inter-individual variability



p: probability that the difference between two individual mice is smaller than miRNA-guided repression

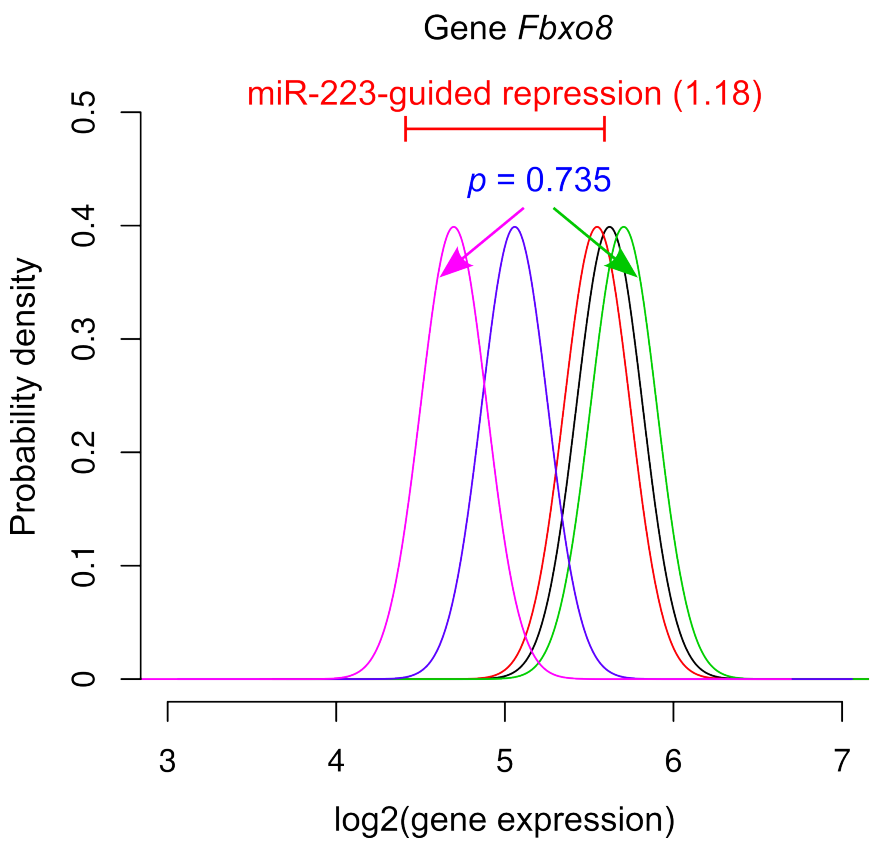
miR-223-guided repression vs. inter-individual variability

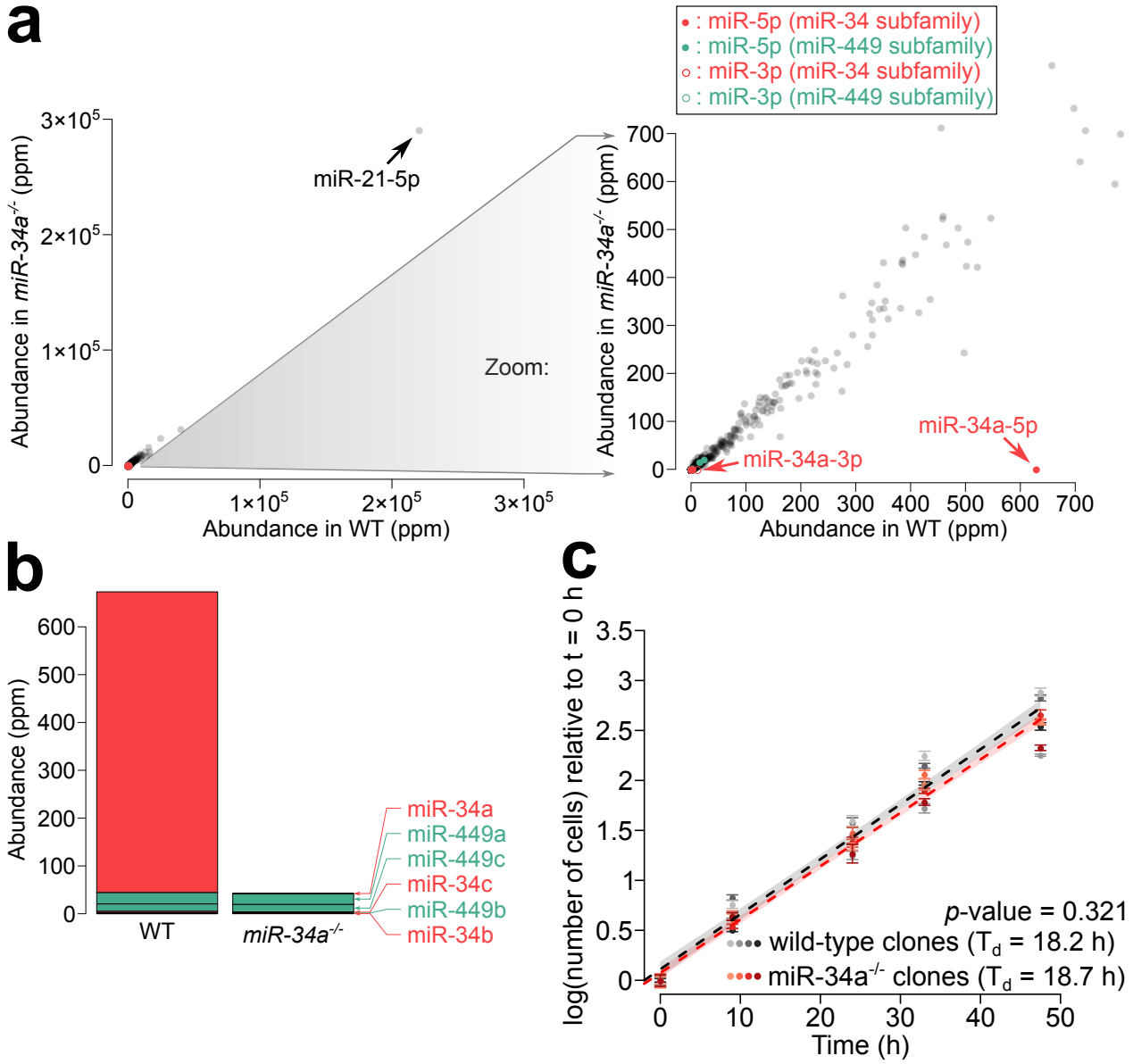
miR-223-guided repression vs. inter-individual variability

For 150 predicted targets out of 192:
inter-individual fluctuations across 5 wild-type mice exceeds miRNA-mediated regulation (p-value < 0.05).

miR-223-guided repression vs. inter-individual variability

For 150 predicted targets out of 192:
inter-individual fluctuations across 5 wild-type mice exceeds miRNA-mediated regulation (p-value < 0.05).



miR-34 and cell proliferation

In HCT-116 cells: