# microRNAs:
# general overview, recent findings

Hervé Seitz

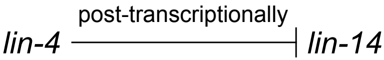
IGH du CNRS, Montpellier, France

June 4, 2015

# Introduction: a brief chronology

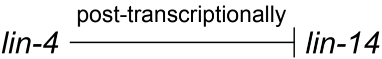
# Introduction: a brief chronology

In *Cænorhabditis elegans* development:



# Introduction: a brief chronology

In *Cænorhabditis elegans* development:



1993: the *lin-4* locus is non-coding.

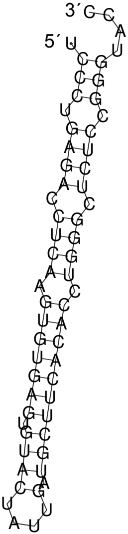
# Introduction: a brief chronology

In *Cænorhabditis elegans* development:

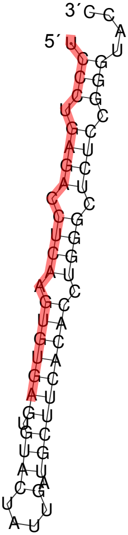
$$\textit{lin-4}\ \overset{\text{post-transcriptionally}}{\longrightarrow\!\!\!\!\dashv}\ \textit{lin-14}$$

1993: the *lin-4* locus is non-coding. Mature product is a 22 nt long RNA (+ a stem-loop-folded 61 nt long precursor).

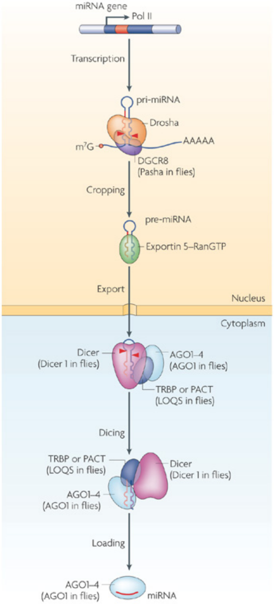
# Introduction: a brief chronology



# Introduction: a brief chronology



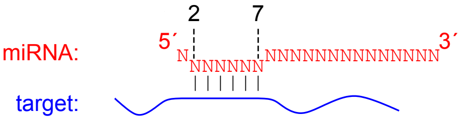
# Introduction: a brief chronology



(Kim *et al*, 2009)

# microRNA target prediction

# microRNA target prediction

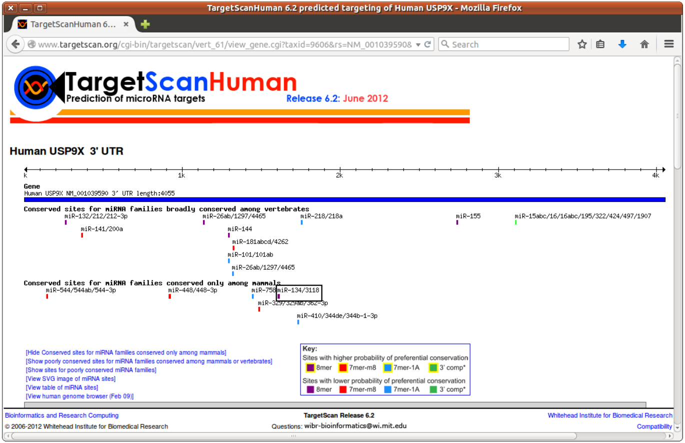


⟶ the “seed”

# 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



TargetScanHuman 6.2 predicted targeting of Human USP9X - Mozilla Firefox

**Human USP9X 3' UTR**

[Hide Conserved sites for miRNA families conserved only among mammals]
[Show poorly conserved sites for miRNA families conserved among mammals or vertebrates]
[Show sites for poorly conserved miRNA families]
[View SVG image of miRNA sites]
[View table of miRNA sites]
[View human genome browser (Feb 09)]

**Key:**
Sites with higher probability of preferential conservation
🟪 8mer 🟥 7mer-m8 🟦 7mer-1A 🟩 3' comp*
Sites with lower probability of preferential conservation
🟪 8mer 🟥 7mer-m8 🟦 7mer-1A 🟩 3' comp*

Bioinformatics and Research Computing
**TargetScan Release 6.2**
Whitehead Institute for Biomedical Research
© 2006-2012 Whitehead Institute for Biomedical Research
Questions: wibr-bioinformatics@wi.mit.edu
Compatibility

# microRNA target prediction

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

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.

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

See 1 citation found by title matching your search:

Cancer Lett. 2015 Feb 1;357(1):384-92. doi: 10.1016/j.canlet.2014.11.050. Epub 2014 Nov 27.

**miR-26a suppresses tumour proliferation and metastasis by targeting metadherin in triple negative breast cancer.**

Liu P[1], Tang H[1], Chen B[1], He Z[2], Deng M[2], Wu M[3], Liu X[1], Yang L[1], Ye F[1], Xie X[4].

⊕ Author information

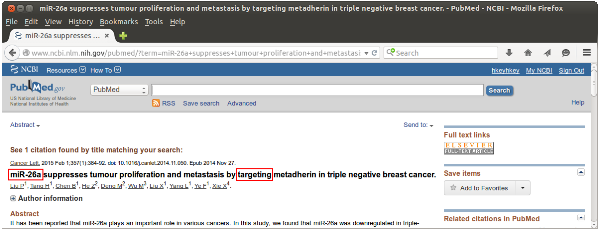
Abstract

It has been reported that miR-26a plays an important role in various cancers. In this study, we found that miR-26a was downregulated in triple-

# microRNA target prediction

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

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



See 1 citation found by title matching your search:

Cancer Lett. 2015 Feb 1;357(1):384-92. doi: 10.1016/j.canlet.2014.11.050. Epub 2014 Nov 27.

**miR-26a suppresses tumour proliferation and metastasis by targeting metadherin in triple negative breast cancer.**

Liu P[1], Tang H[1], Chen B[1], He Z[2], Deng M[2], Wu M[3], Liu X[1], Yang L[1], Ye F[1], Xie X[4].

Author information

Abstract

It has been reported that miR-26a plays an important role in various cancers. In this study, we found that miR-26a was downregulated in triple-

# microRNA target prediction

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

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.

# microRNA target prediction

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

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.

miRNA-mediated repression is very modest (usually $<$ 2-fold): lower than well tolerated fluctuations in gene expression (*e.g.*, haplosufficiency). Why have these sites been conserved if they are not functional?

# Robustness of predicted targets

# Robustness of predicted targets

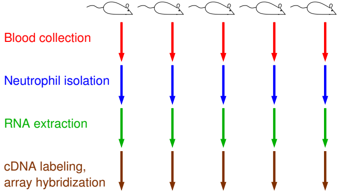
# Robustness of predicted targets

# Robustness of predicted targets

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

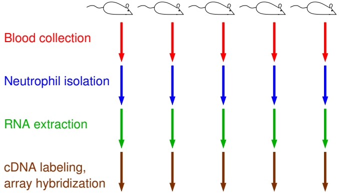
# Robustness of predicted targets

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

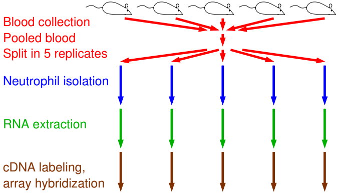


# Robustness of predicted targets

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

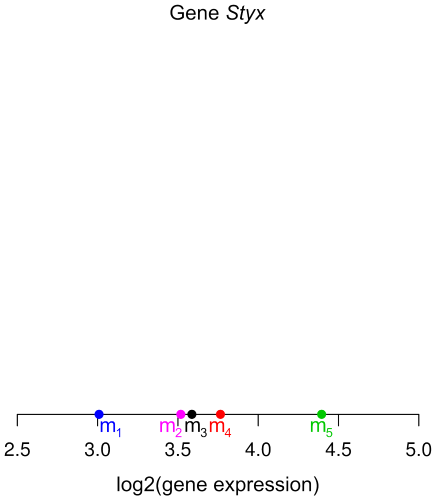


Blood collection
Neutrophil isolation
RNA extraction
cDNA labeling,
array hybridization

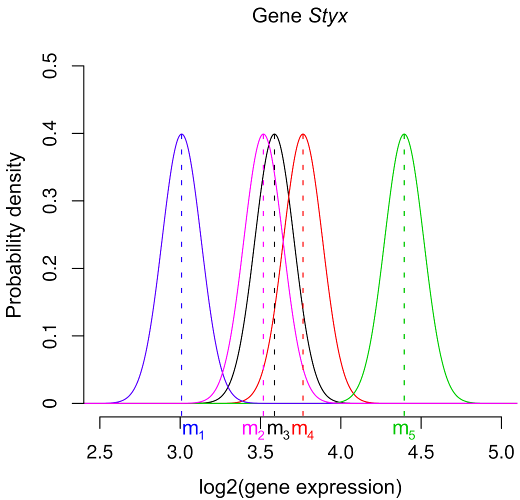


Blood collection
Pooled blood
Split in 5 replicates
Neutrophil isolation
RNA extraction
cDNA labeling,
array hybridization

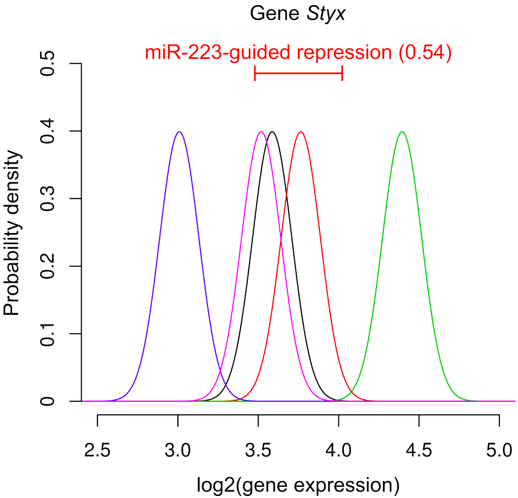
# Robustness of predicted targets



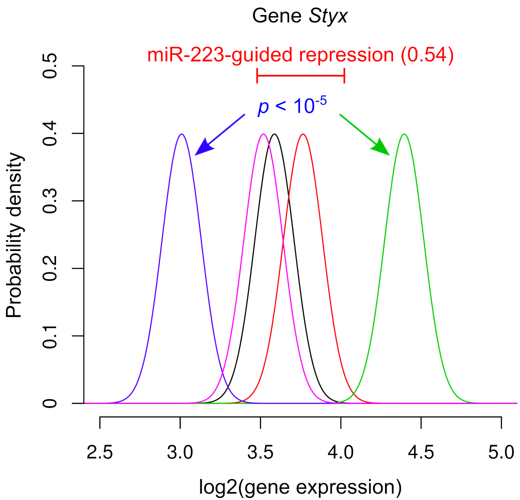
# Robustness of predicted targets



# Robustness of predicted targets



# Robustness of predicted targets



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

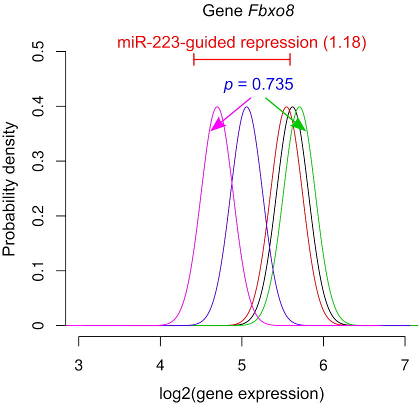
# Robustness of predicted targets

# Robustness of predicted targets

For 168 predicted targets out of 189:
inter-individual fluctuations across 5 wild-type mice exceeds miRNA-mediated regulation ($p$-value $< 0.05$).

# Robustness of predicted targets

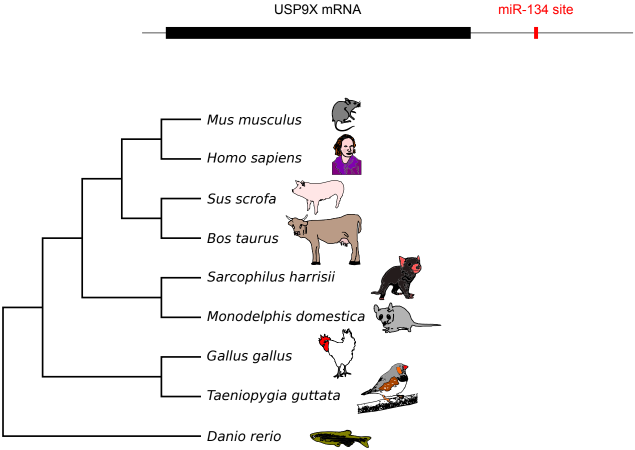
For 168 predicted targets out of 189:
inter-individual fluctuations across 5 wild-type mice exceeds miRNA-mediated regulation ($p$-value $< 0.05$).



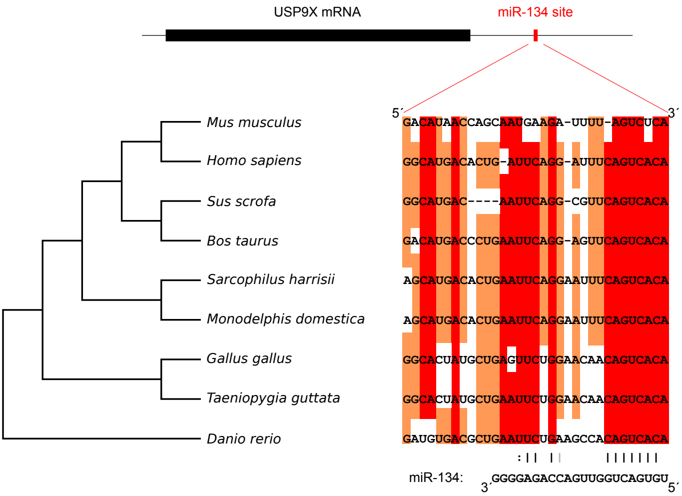
# Comparative genomics yield false positives

# Comparative genomics yield false positives

# Comparative genomics yield false positives

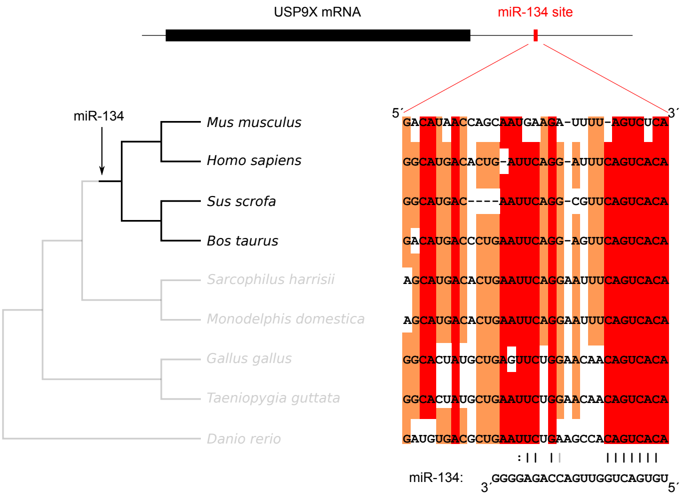


# Comparative genomics yield false positives



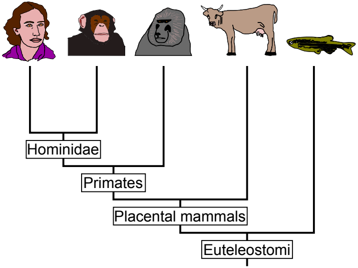
🟥 : conserved in 8 or 9 species out of 9 🟧 : conserved in 6 or 7 species out of 9 ⬜ : conserved in less than 6 species out of 9

# Comparative genomics yield false positives

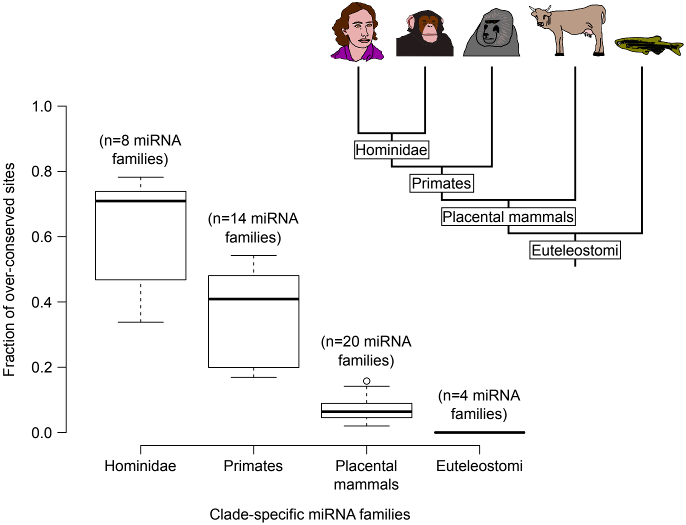


🟥: conserved in 8 or 9 species out of 9 🟧: conserved in 6 or 7 species out of 9 ⬜: conserved in less than 6 species out of 9

# Comparative genomics yield false positives

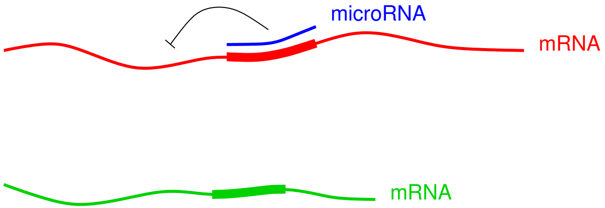


# Comparative genomics yield false positives

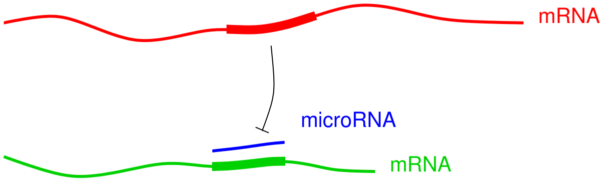


# An alternative hypothesis

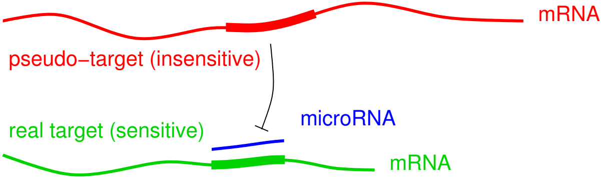
# An alternative hypothesis



# An alternative hypothesis



# An alternative hypothesis



# Absolute RNA quantification

# Absolute RNA quantification

# Absolute RNA quantification

For mRNAs to titrate miRNAs, they need to be abundant enough. ▸ Issues with published examples

# Absolute RNA quantification

For mRNAs to titrate miRNAs, they need to be abundant enough.

▸ Issues with published examples

Murine cell line “C2C12”: differentiates in myotubes; expresses muscle-specific miR-1, miR-206 and miR-133.

# Absolute RNA quantification

For mRNAs to titrate miRNAs, they need to be abundant enough.

▸ Issues with published examples

Murine cell line “C2C12”: differentiates in myotubes; expresses muscle-specific miR-1, miR-206 and miR-133.

miR-1 and miR-206 (at least) control C2C12 differentiation (Goljanek-Whysall *et al.*, 2012).

# Absolute RNA quantification

For mRNAs to titrate miRNAs, they need to be abundant enough.

▸ Issues with published examples

Murine cell line "C2C12": differentiates in myotubes; expresses muscle-specific miR-1, miR-206 and miR-133.

miR-1 and miR-206 (at least) control C2C12 differentiation (Goljanek-Whysall *et al.*, 2012).

$\Longrightarrow$ absolute quantification of miR-1, miR-206, miR-133 and their predicted targets in differentiating C2C12.

# Absolute RNA quantification

For mRNAs to titrate miRNAs, they need to be abundant enough.

▸ Issues with published examples

Murine cell line “C2C12”: differentiates in myotubes; expresses muscle-specific miR-1, miR-206 and miR-133.

miR-1 and miR-206 (at least) control C2C12 differentiation (Goljanek-Whysall *et al.*, 2012).

$\Longrightarrow$ <u>absolute</u> quantification of miR-1, miR-206, miR-133 and their predicted targets in differentiating C2C12.

# Absolute RNA quantification

For mRNAs to titrate miRNAs, they need to be abundant enough.

▸ Issues with published examples

Murine cell line “C2C12”: differentiates in myotubes; expresses muscle-specific miR-1, miR-206 and miR-133.

miR-1 and miR-206 (at least) control C2C12 differentiation (Goljanek-Whysall *et al.*, 2012).

$\Longrightarrow$ absolute quantification of miR-1, miR-206, miR-133 and their predicted targets in differentiating C2C12.

- miRNAs: Northern blot (calibrated with RNA oligos);

▸ Representative results

# Absolute RNA quantification

For mRNAs to titrate miRNAs, they need to be abundant enough.

▸ Issues with published examples

Murine cell line "C2C12": differentiates in myotubes; expresses muscle-specific miR-1, miR-206 and miR-133.

miR-1 and miR-206 (at least) control C2C12 differentiation (Goljanek-Whysall *et al.*, 2012).

$\Longrightarrow$ absolute quantification of miR-1, miR-206, miR-133 and their predicted targets in differentiating C2C12.

- miRNAs: Northern blot (calibrated with RNA oligos);
- mRNAs: RNA-Seq (calibrated with *in vitro* transcripts).

▸ Representative results

# Absolute RNA quantification

| miRNA family | Predicted targets | |
|---|---|---|
| | Total | Above $K_d$ |
| miR-1 and miR-206 | 436 | 115 (26 %) |
| miR-133 | 384 | 84 (22 %) |

($K_d$ value taken from Wee *et al.*, 2012: 26 pM)

# Absolute RNA quantification

Titration effect of individual miRNA binding sites:

| mRNA | Sample | Increase in free miR-1/miR-206 if site is lost |
|---|---|---|
| *Anxa2* | Day 0 | 8 % |
| *Ptma* | Day 0 | 9 % |
| *Tmsb4x* | Day 0 | 10 % |
| *Actb* | Day 3 | 15 % |
| *Tmsb4x* | Day 3 | 9 % |
| *Actb* | Day 6 | 13 % |
| *Tmsb4x* | Day 6 | 11 % |

# Absolute RNA quantification

Titration effect of individual miRNA binding sites:

| mRNA | Sample | Increase in free miR-133 if site is lost |
|---|---|---|
| *Tpm4* | Day 0 | 5 % |
| *Ptma* | Day 0 | 16 % |
| *Eif4a1* | Day 0 | 5 % |
| *Tagln2* | Day 0 | 6 % |
| *Ptma* | Day 3 | 7 % |
| *Eef1a1* | Day 3 | 8 % |
| *Ptma* | Day 6 | 6 % |
| *Eef1a1* | Day 6 | 6 % |

# Conclusion: revisiting miRNA target definition

# Conclusion: revisiting miRNA target definition

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

# Conclusion: revisiting miRNA target definition

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

Among the (many !) predicted targets: finding those that are responsible for visible phenotypes.

# Conclusion: revisiting “gene regulation” definition

High-throughput identification of transcription factor or RNA-binding protein targets: thousands of genes are bound, many of them are not under selective pressure to keep these binding sites. ▸ Supplementary data

# Conclusion: revisiting “gene regulation” definition

High-throughput identification of transcription factor or RNA-binding protein targets: thousands of genes are bound, many of them are not under selective pressure to keep these binding sites. ▸ Supplementary data

Microscopic events which are neutral in evolutionary terms.

# Conclusion: revisiting “gene regulation” definition

High-throughput identification of transcription factor or RNA-binding protein targets: thousands of genes are bound, many of them are not under selective pressure to keep these binding sites. ▸ Supplementary data

Microscopic events which are neutral in evolutionary terms.

Are there pseudo-targets for these other regulators?

# Conclusion: revisiting "gene regulation" definition

High-throughput identification of transcription factor or RNA-binding protein targets: thousands of genes are bound, many of them are not under selective pressure to keep these binding sites. ▸ Supplementary data

Microscopic events which are neutral in evolutionary terms.

Are there pseudo-targets for these other regulators?

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

Anna Sergeeva

Laura Martinez

Blaise Li

Natalia Pinzón

Isabelle Busseau

# Acknowledgements

Anna Sergeeva

Laura Martinez

Blaise Li

Natalia Pinzón

Isabelle Busseau

Jessy Presumey and Florence Apparailly (INM, Montpellier, France)

# Acknowledgements

Anna Sergeeva

Laura Martinez

Blaise Li

Natalia Pinzón

Isabelle Busseau

Jessy Presumey and Florence Apparailly (INM, Montpellier, France)

# Supplementary data

## Introduction:

- Mechanism of target repression

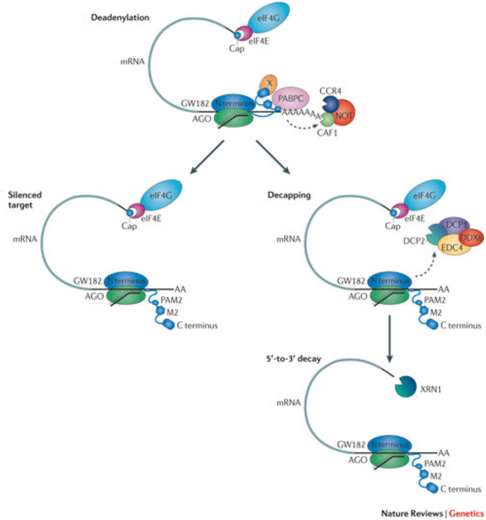
## Results:

- Robustness of biological pathways
- Published evidence for genome-wide targeting
- Issues with published pseudo-targets
- Absolute RNA quantification results
- RNA-Seq statistics
- Candidate titrators

## Conclusion:

- Pseudo-targets for other regulators?
- Propagation of gene expression perturbation

# miRNA target repression



(adapted from Huntzinger and Izaurralde, 2011)

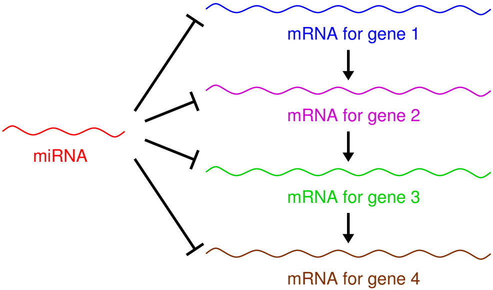
# Biological robustness

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

# Published evidence

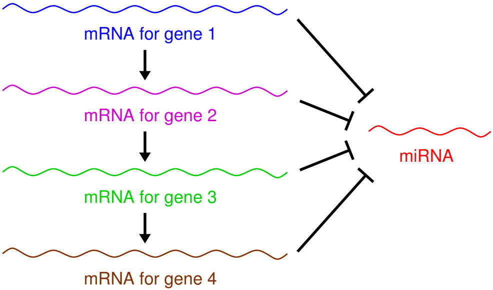
# Published evidence

Targets for a given miRNA often belong to the same biological pathways.



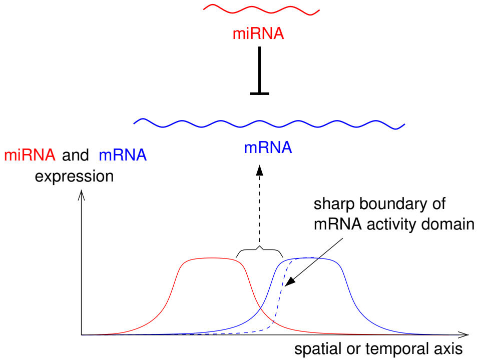
# Published evidence

Targets for a given miRNA often belong to the same biological pathways.



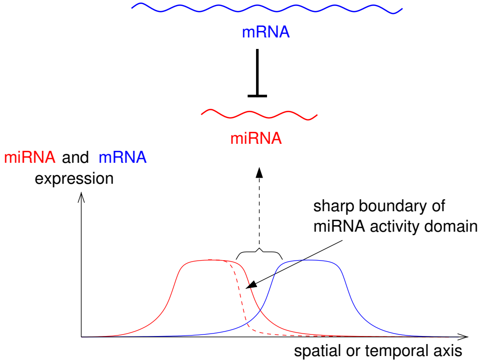
# Published evidence

Expression domains for miRNAs and their predicted targets overlap at their boundaries.



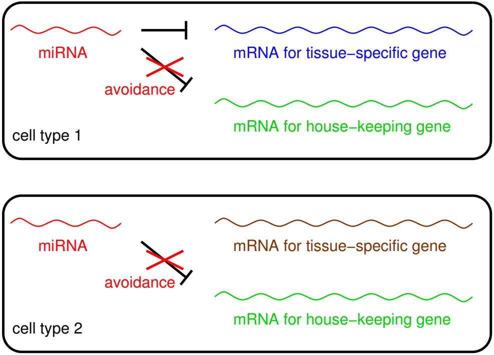
# Published evidence

Expression domains for miRNAs and their predicted targets overlap at their boundaries.



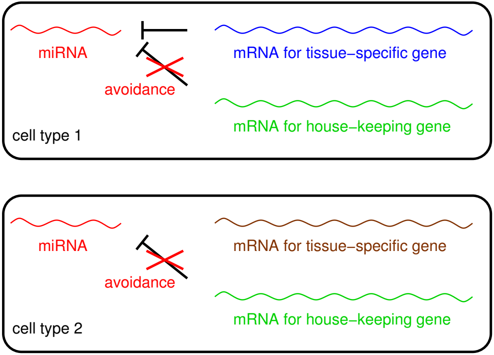
# Published evidence

House-keeping genes are rarely predicted to be targeted.



# Published evidence

House-keeping genes are rarely predicted to be targeted.



# Reported examples of pseudo-targets

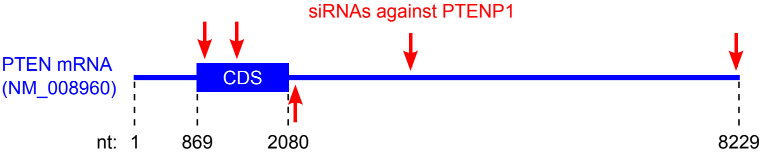
A proposed pseudo-target: PTENP1, that de-repressed PTEN (Poliseno *et al.*, 2010).

# Reported examples of pseudo-targets

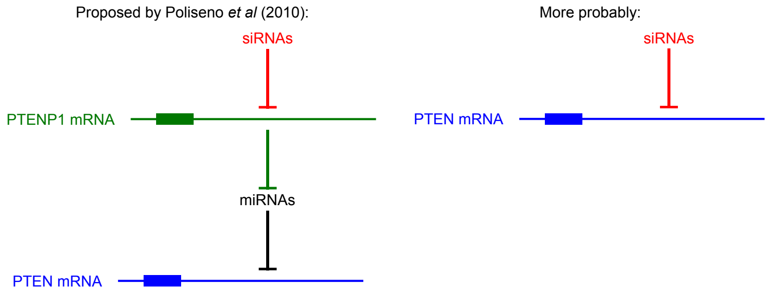
A proposed pseudo-target: PTENP1, that de-repressed PTEN (Poliseno *et al.*, 2010).

But PTENP1 mRNA is $\approx$ 100 times less abundant than the PTEN mRNA (Ebert and Sharp, 2010).

# Reported examples of pseudo-targets



# Reported examples of pseudo-targets



Proposed by Poliseno *et al* (2010):

siRNAs

PTENP1 mRNA

miRNAs

PTEN mRNA

More probably:

siRNAs

PTEN mRNA

# Reported examples of pseudo-targets

```
siRNA #2 against PTENP1:      5′ UAAUAAUCAUCAUUCUGGC 3′
                                  ||||||    ||| |||
PTEN mRNA (nt 948-961):       3′ GUUAUUA----UAA-ACCU 5′
```

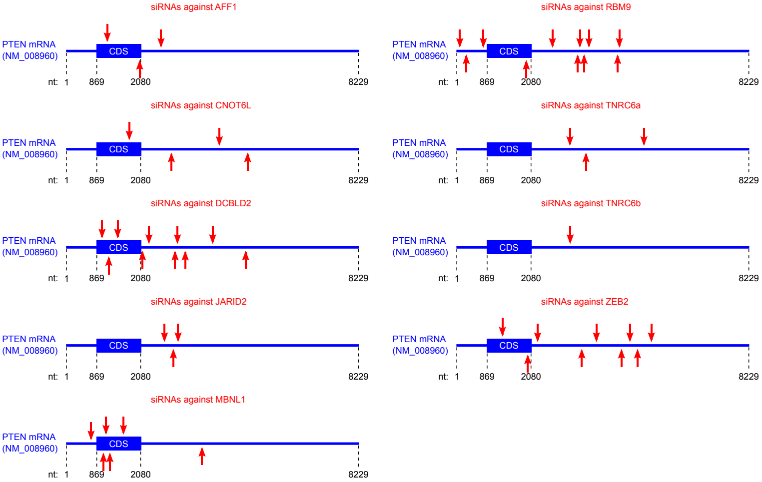
```
siRNA #2 against PTENP1:      5′ UAAUAAUCAUCAUUCUGGC 3′
                                  ||||||   |||  :| |
PTEN mRNA (nt 1383-1401):     3′ AUUAUUAUAUGUAUCGCGG 5′
```

```
siRNA #2 against PTENP1:      5′ UAAUAAUCAUCAUUCUGGC 3′
                                  ||||||       ||||
PTEN mRNA (nt 3769-3782):     3′ AUUAUUACC-----GACCU 5′
```

```
siRNA #2 against PTENP1:      5′ UAAUAAUC-AUCAUUCUGGC 3′
                                  ||||||  |:|    |::
PTEN mRNA (nt 8189-8208):     3′ UUUAUUACUUGGAAAAAUUA 5′
```

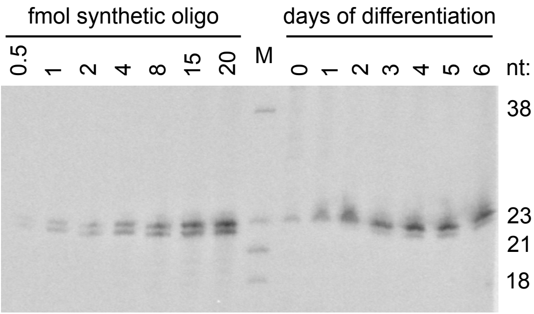
```
siRNA #3 against PTENP1:      5′ UCCUAUA--UGAUCUCUGAUG 3′
                                  ||||||  ||   |||||::
PTEN mRNA (nt 2192-2212):     3′ AGGAUAUUGACGUUAGACUGU 5′
```

# Reported examples of pseudo-targets



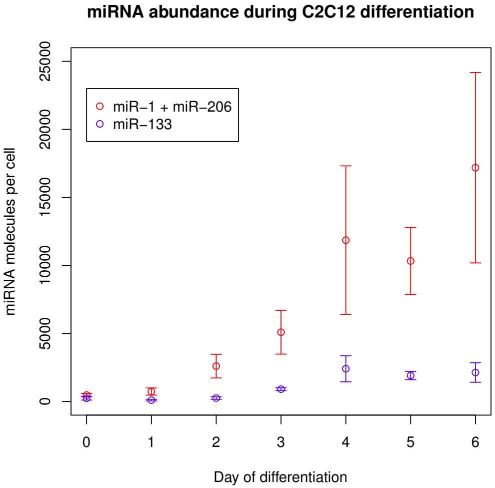
Return

# miRNA quantification



(quantification of miR-1 and miR-206)

# miRNA quantification



# mRNA quantification

# mRNA quantification

Very deep sequencing: three time points (day 0, day 3, day 6) in triplicate; each replicate: between 267 and 333 million transcriptome-matching reads.

▸ Statistics

华大基因
BGI

# mRNA quantification

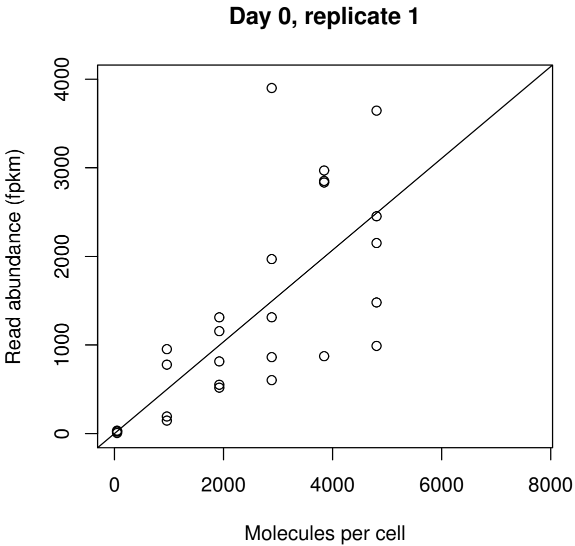
Very deep sequencing: three time points (day 0, day 3, day 6) in triplicate; each replicate: between 267 and 333 million transcriptome-matching reads.

▸ Statistics

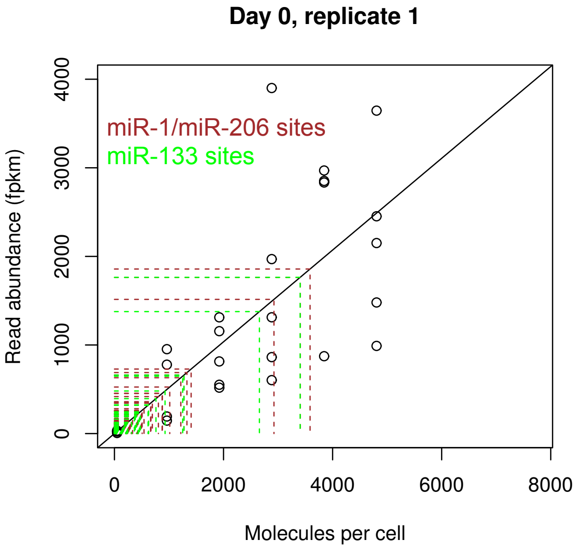
华大基因
BGI

27 synthetic spike-ins, for calibration.

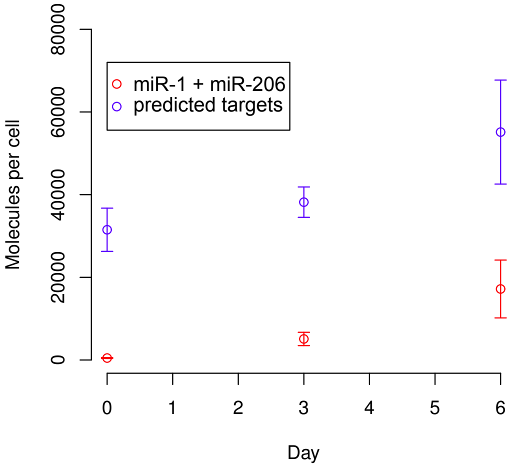
# mRNA quantification



# mRNA quantification

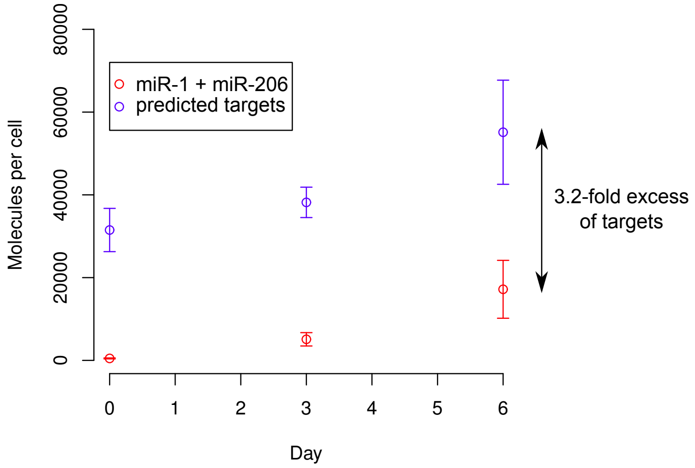


# mRNA quantification



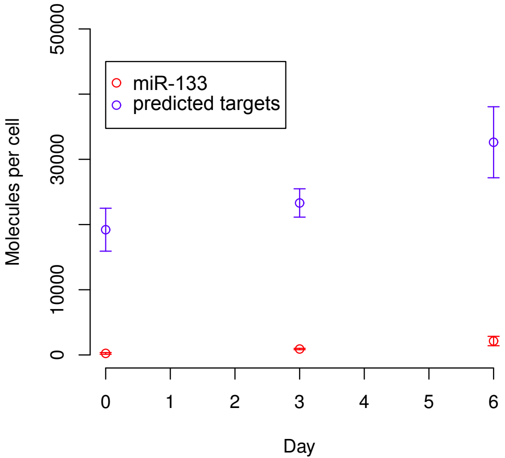
◂ Return

# mRNA quantification



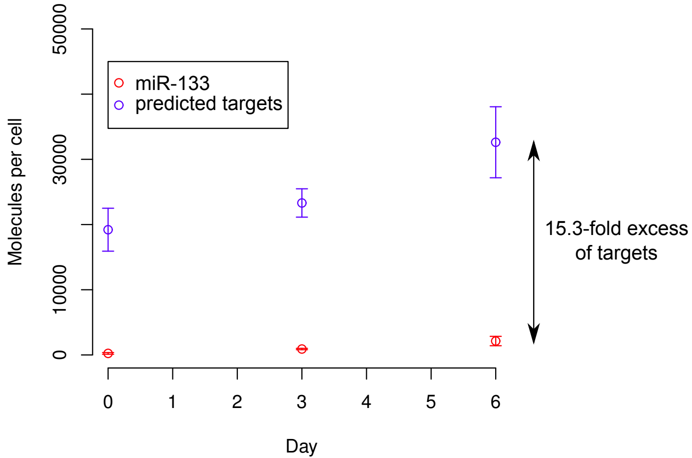
◂ Return

# mRNA quantification



◂ Return

# mRNA quantification



◂ Return

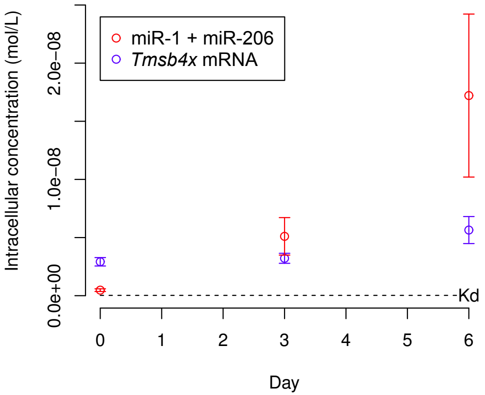
# Deep-sequencing statistics

Number of reads per kb (median transcript, excluding spike-ins):

| | | |
|---|---|---|
| Day 0 | Replicate 1 | 159.26 |
| | Replicate 2 | 140.53 |
| | Replicate 3 | 150.03 |
| Day 3 | Replicate 1 | 213.09 |
| | Replicate 2 | 167.77 |
| | Replicate 3 | 198.48 |
| Day 6 | Replicate 1 | 223.57 |
| | Replicate 2 | 243.86 |
| | Replicate 3 | 220.23 |

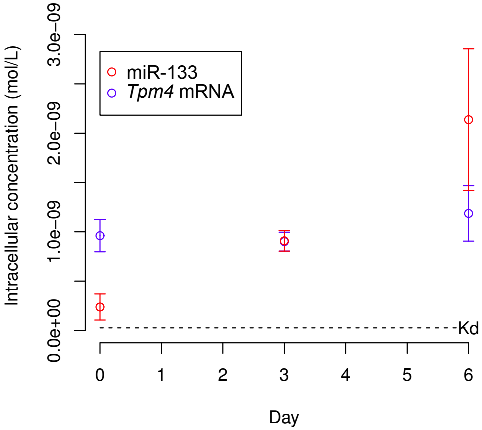
◂ Return

# Concentration effects



◂ Return

# Concentration effects



Return

# Pseudo-targets for other regulators?

- Transcription factors: experimentally identified binding sites are poorly conserved among vertebrates (Odom *et al.*, 2007 and Schmidt *et al.*, 2010).

# Pseudo-targets for other regulators?

- Transcription factors: experimentally identified binding sites are poorly conserved among vertebrates (Odom *et al.*, 2007 and Schmidt *et al.*, 2010).
- RNA-binding proteins are poorly specific (thousands of experimentally validated targets for each analyzed protein: Hafner *et al.*, 2010; Lebedeva *et al.*, 2011 and Hafner *et al.*, 2013).

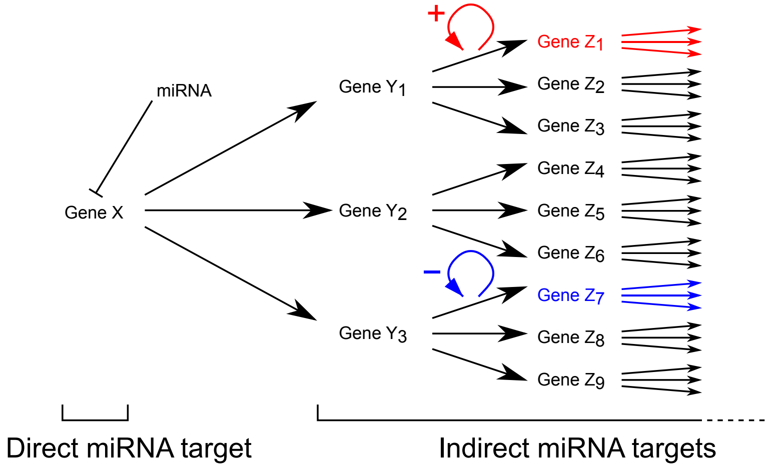
# Pseudo-targets for other regulators?

- Transcription factors: experimentally identified binding sites are poorly conserved among vertebrates (Odom *et al.*, 2007 and Schmidt *et al.*, 2010).
- RNA-binding proteins are poorly specific (thousands of experimentally validated targets for each analyzed protein: Hafner *et al.*, 2010; Lebedeva *et al.*, 2011 and Hafner *et al.*, 2013).

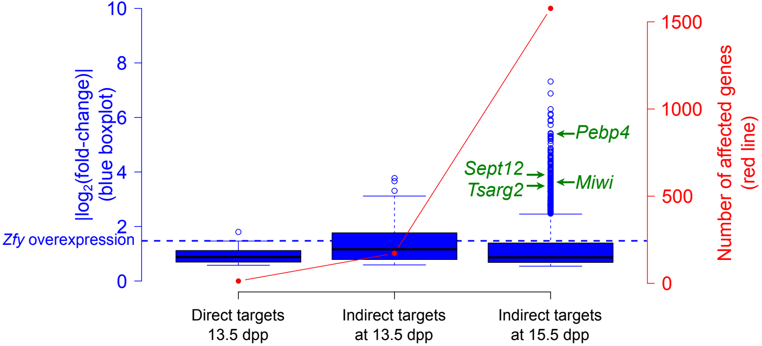
Real molecular events, which are neutral in evolutionary terms?

◂ Return

# Propagation of gene expression perturbation



# Propagation of gene expression perturbation



(in collaboration with H. Royo and J. Turner, MRC, London)