Partir à la chasse de nouveaux marqueurs, cibles et médicaments contre le cancer avec des armes de petits ARN

25 janvier 2019

UE #2 - Cancer et génomique : Big data et modèles prédictifs

Major Lab
Département d'informatique et de recherche opérationnelle
Institute de recherche en immunologie et en cancérologie (IRIC)
Université de Montréal

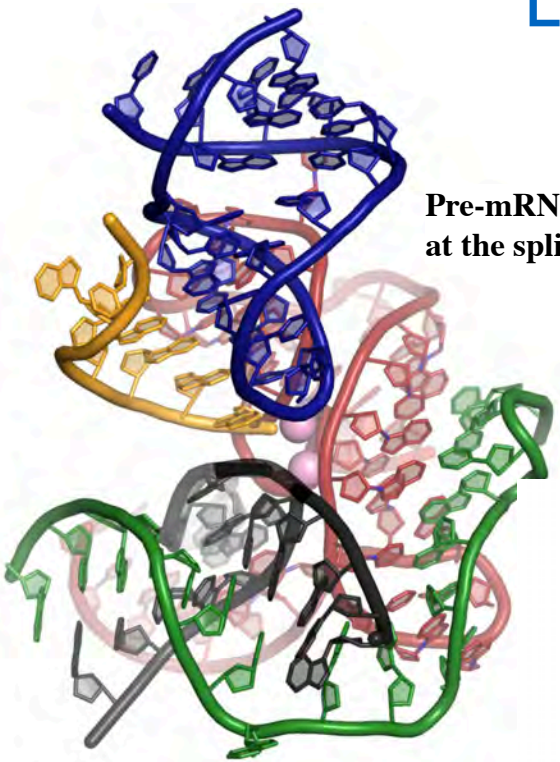
Chaire d'Alembert
U. Paris-sud

Pourquoi analyser les données NGS
(next-generation sequencing)

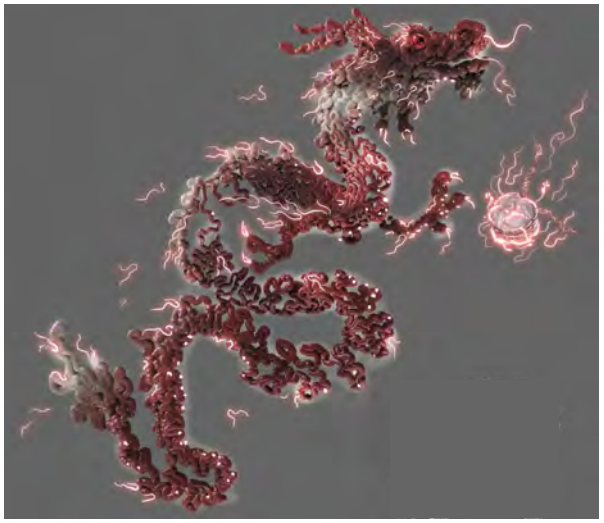
- Suivre l'expression des gènes
- Détecter des variants somatiques, e.g. mutations, transcrits alternatifs
- Apprendre des signatures de variants cliniques

➡ **Étudier la régulation de l'expression des gènes**

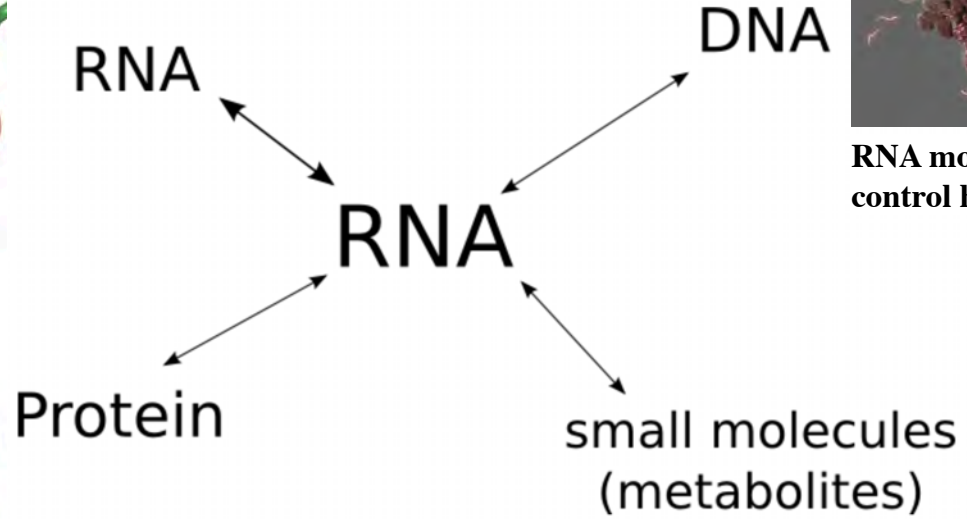
Les ARN exprimés interagissent

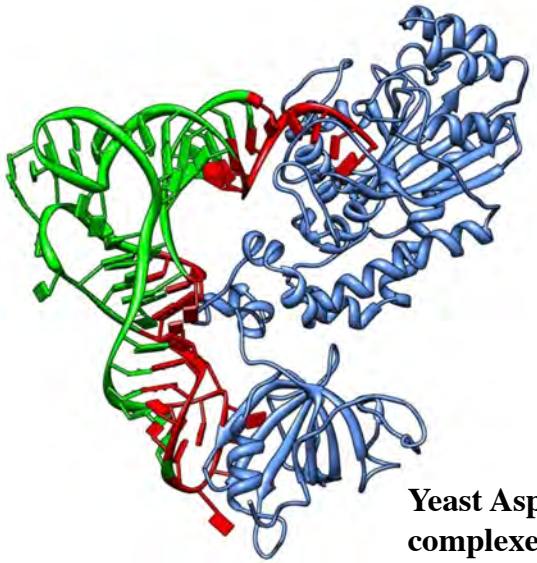


Pre-mRNA substrate and snRNAs at the spliceosomal active site



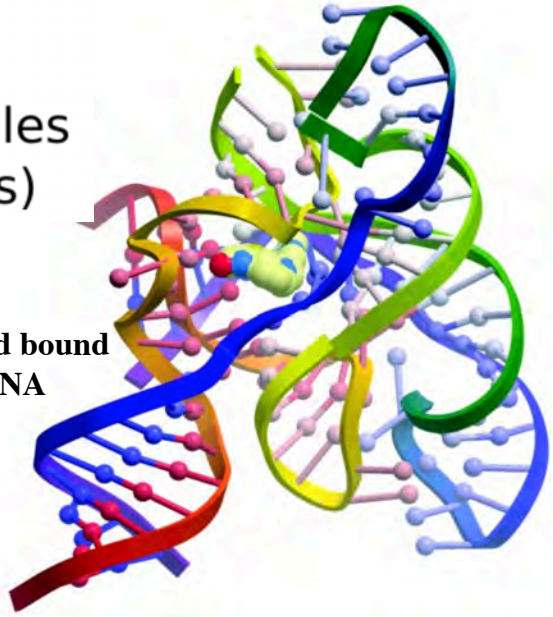
RNA molecules attach to particular DNA to control how much protein these genes produce



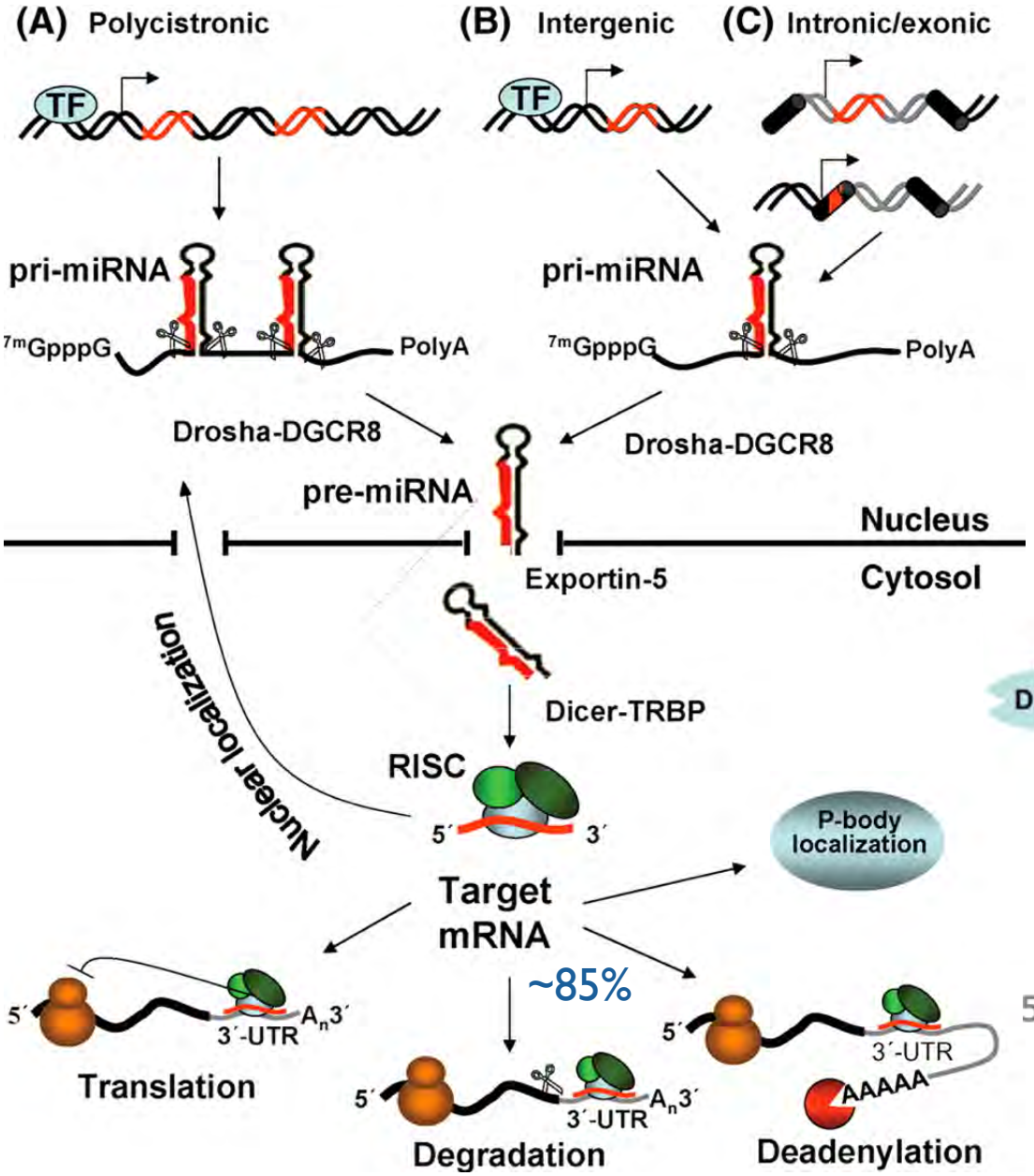


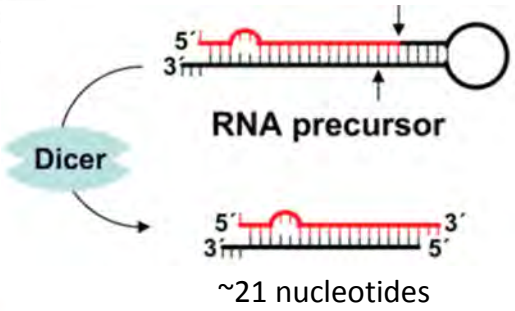
Yeast Aspartyl-tRNA synthetase complexed with $tRNA^{Asp}$

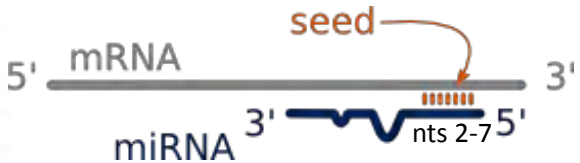
A model of Merck's ribocil compound bound to a bacterial riboflavin riboswitch RNA



MicroRNAs interactions act to decrease translational efficiency and mRNA levels







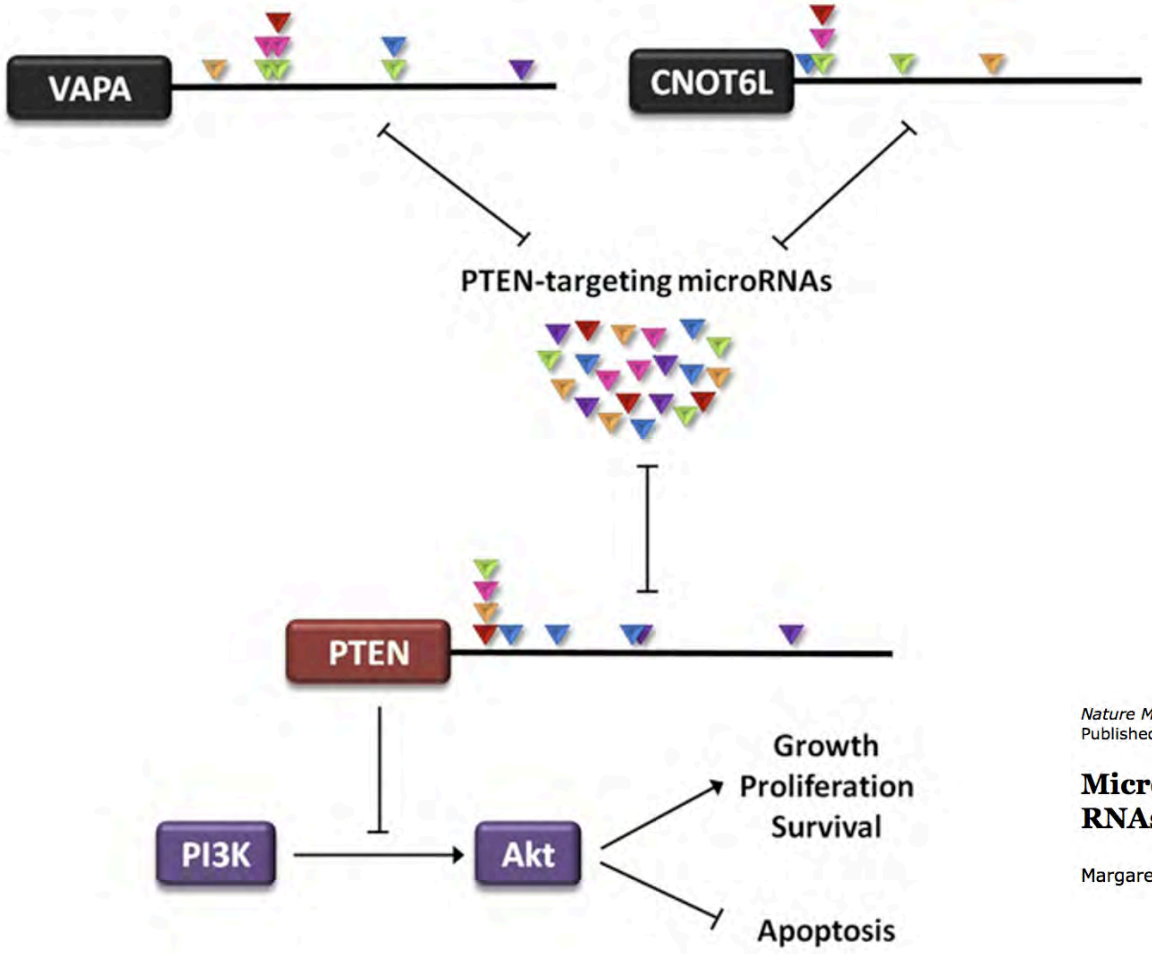
nature
International journal of science

Mammalian microRNAs predominantly act to decrease target mRNA levels

Huili Guo, Nicholas T. Ingolia, Jonathan S. Weissman & David P. Bartel

Figure from Fazi & Nervi (2008) *Cardiovasc Res*

MicroRNAs crosstalks regulate (up or down) competitive mRNA levels



nature methods
Techniques for life scientists and chemists

Nature Methods - **4**, 721 - 726 (2007)
Published online: 12 August 2007; | doi:10.1038/nmeth1079

MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells

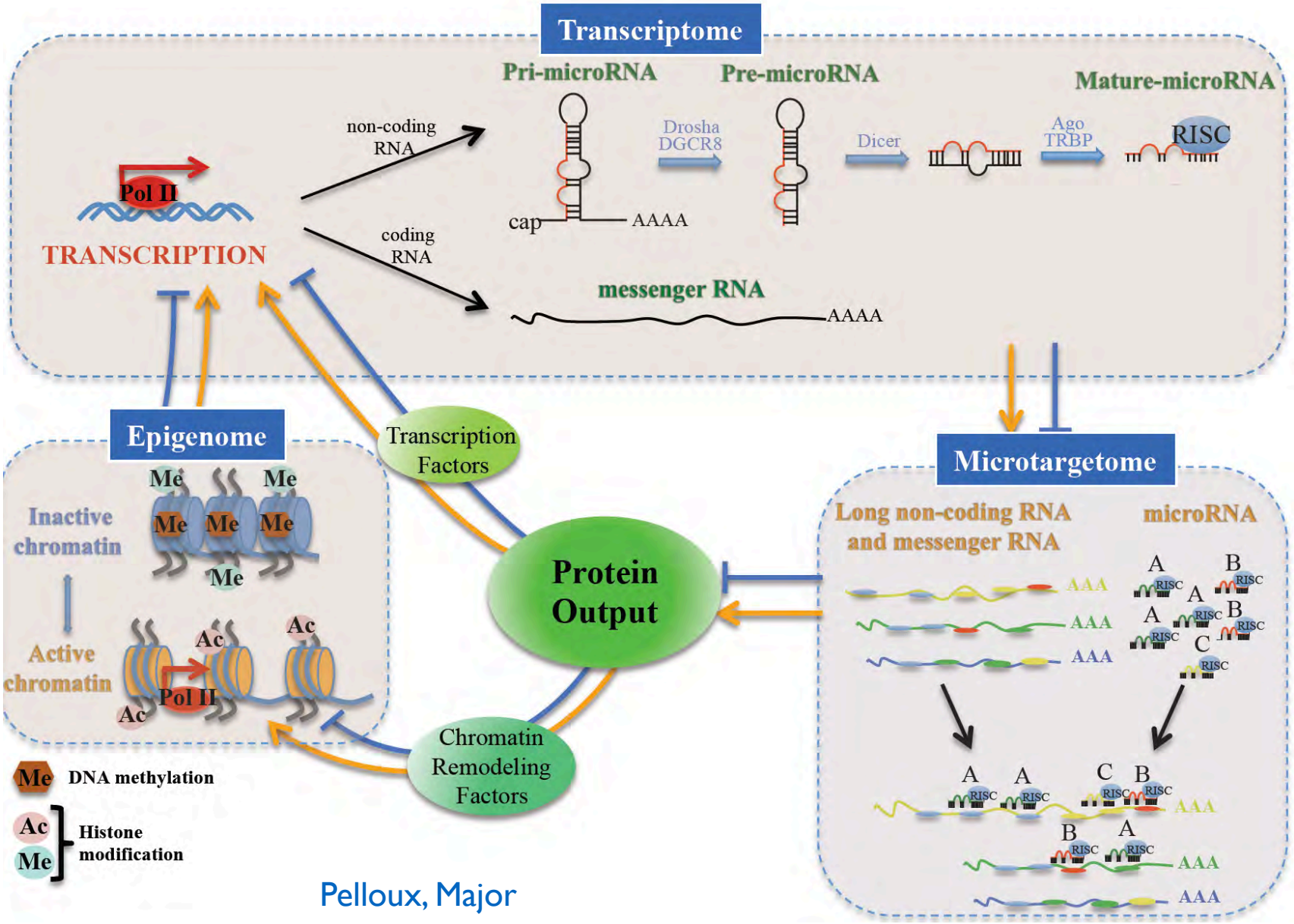
Margaret S Ebert, Joel R Neilson & Phillip A Sharp

Cell

Coding-Independent Regulation of the Tumor Suppressor PTEN by Competing Endogenous mRNAs p344

Yvonne Tay, Lev Kats, Leonardo Salmena, Dror Weiss, Shen Mynn Tan, Ugo Ala, Florian Karreth, Laura Poliseno, Paolo Provero, Ferdinando Di Cunto, Judy Lieberman, Isidore Rigoutsos, Pier Paolo Pandolfi

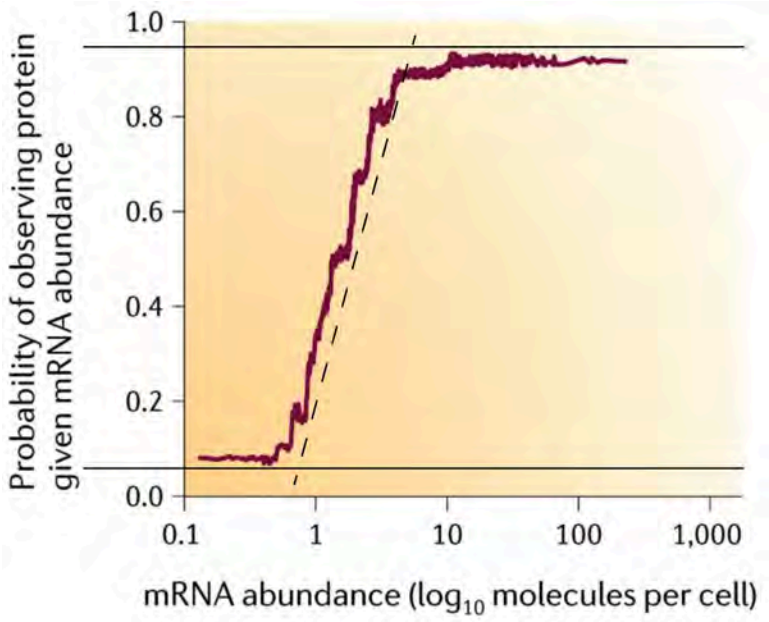
Gene expression systems are interrelated



Pelloux, Major

Transcription defines the activation and the order of magnitude of protein expression

- Cells produce the proteome from DNA, involving transcription and translation, post-transcriptional and translational regulatory processes, as well as related to protein degradation.
- Although mRNA abundance broadly explains protein levels, these factors imperfectly correlate.
- RNA expression in eucaryotes may act as an activation device of protein expression: the detection of proteins by mass spectrometry is almost null when mRNA levels are low, and goes up sharply starting at a certain threshold.



mRNA-protein abundance relationship in yeast
The slope (dashed line) and horizontal asymptotes (plain lines) that define protein abundance as a sigmoid function: $f(x) = e^x/(e^x+1)$.

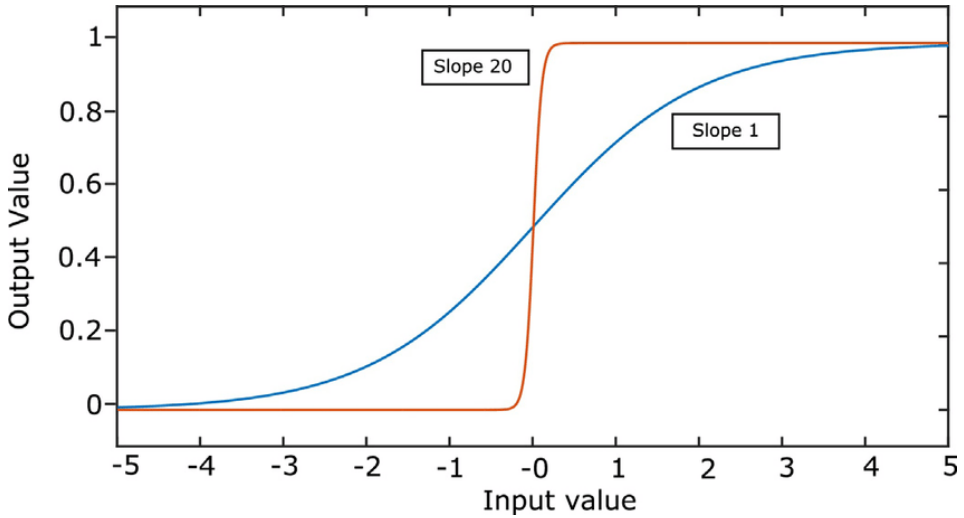
Modified from Vogel & Marcotte 2010 *Nat. Rev. Genet.*

The downstream processes fine-tune protein levels

(adjust the slope of the sigmoid function)

MiRNAs have been shown to act as a rheostat of protein abundances.

Selbach et al. 2008 *Nature*
Baek et al. 2008 *Nature*



By fine-tuning protein expression levels, miRNA-induced silencing (miS) is a key player in numerous cellular functions, development, and diseases.

Ambros 2001 *Cell*
Ambros 2004 *Nature*

Numerous experimental techniques provide partial microtargetome data

Luciferase reporter assays

Grimson et al. 2007 Mol Cell (Bartel)
Wu et al. 2010 Oncogene (He)
Zhou et al. 2013 PLoS One (Jiang)

Microarrays

DeRisi et al. 1997 Science (Brown)
Lim et al 2005 Nature (Johnson)

RNA-seq

Wang et al. 2009 Nat Rev Genet (Snyder)

Variants of SILAC

Baek et al. 2008 Nature (Bartel)
Selbach et al. 2008 Nature (Rajewsky)

Variants of CLIP

Chi et al. 2009 Nature (Darnell, HITS-CLIP)
Hafner et al. 2010 Cell (Tuschl, PAR-CLIP)
Helwak et al. 2013 Cell (Tollervey, CLASH)
Broughton et al. 2016 Mol Cell (Pasquinelli, iCLIP)

Results from these studies showed that:

(i) Each miRNA can target several mRNAs;

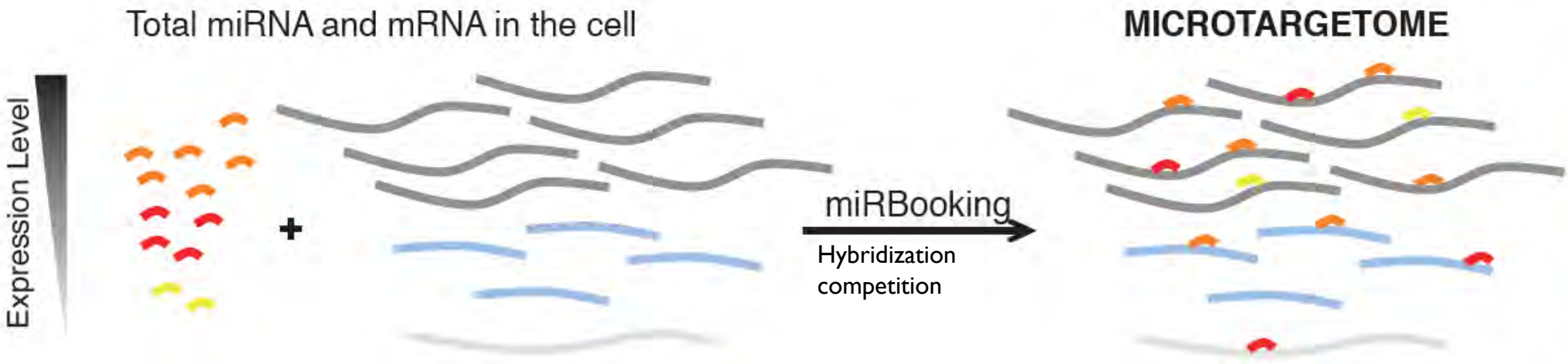
(ii) Each mRNA can be the target of several miRNAs;

(iii) Most mRNAs are subject to miRNA-induced silencing; and,

(iv) MiRNA-induced silencing is subject to cellular conditions.

Mukherji et al. 2011 Nat Genet (Sharp & Oudenaarden)
Arvey et al. 2010 Mol Syst Biol (Marks)
Salmena et al. 2011 Cell (Pandolfi)
Sood et al. 2006 PNAS (Rajewsky)

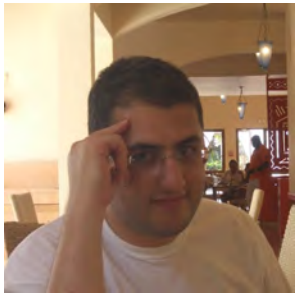
Given the RNA content of a cell, miRBooking predicts its *microtargetome*: the miRNA-mRNA interaction network



Nucleic Acids Research 2015

MiRBooking simulates the stoichiometric mode of action of microRNAs

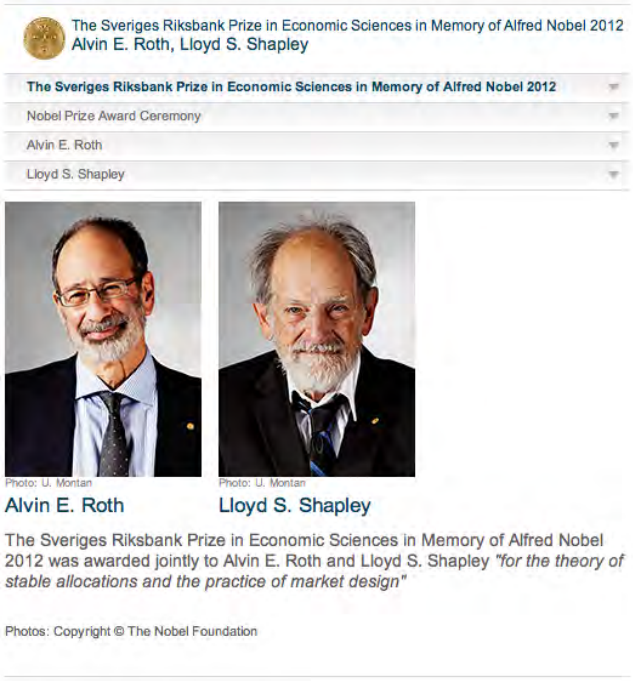
Nathanaël Weill, Véronique Lisi, Nicolas Scott, Paul Dallaire, Julie Pelloux and François Major*

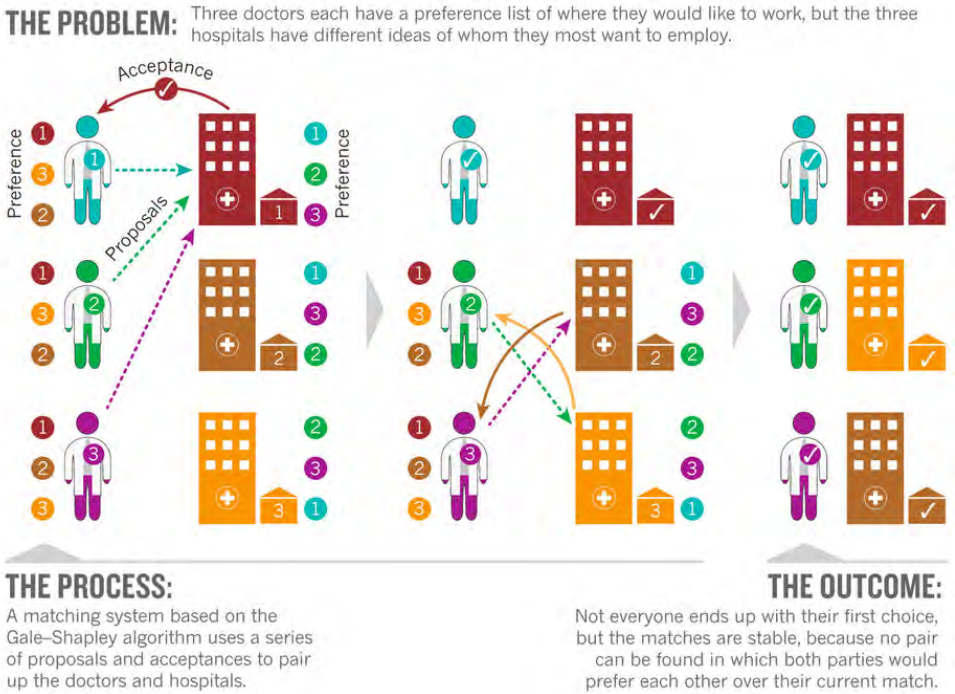


First miRNA target prediction program that:

- solves the entire miRNA-mRNA network
- simulates the stoichiometric mode of action of miRNAs
- allows us to simulate experiments with abundance changes
- designs and predicts the effects of artificial sequences

We defined the hybridization competition as a Gale-Shapley stable marriage problem





Chen & Goeree (2012) *Nature*

```
function stableMatching {
  Initialize all m ∈ M and w ∈ W to free
  while ∃ free man m who still has a woman w to propose to {
     w = m's highest ranked such woman who he has not proposed to yet
     if w is free
       (m, w) become engaged
     else some pair (m', w) already exists
       if w prefers m to m'
         (m, w) become engaged
         m' becomes free
       else
         (m', w) remain engaged
  } // while
} // function
```

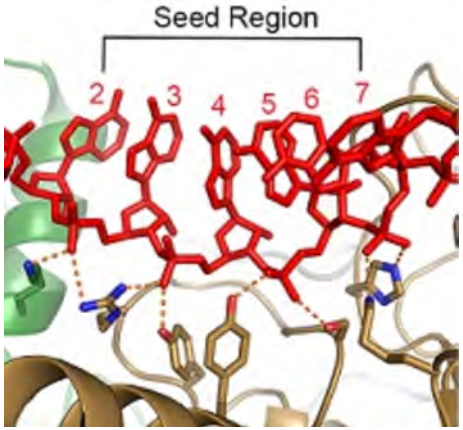
The stable marriage algorithm:

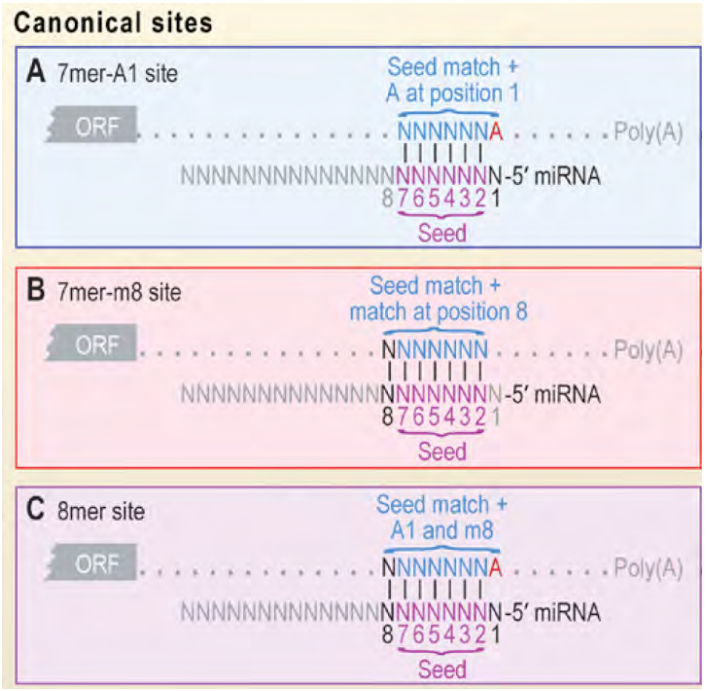
- **Solves the problem**
- **Is simple**
- **Runs in $O(n^2)$**
- **Favors the proposers**
- **Guarantees that no pair can be broken so to favor both partners**

Weill et al. 2015 *NAR*

MiRBooking uses simple preferences determined by RNA abundances and seed matching

- The proposers are the seeds; the acceptors the MREs.
- The preferences are defined by MRE and seed abundances, and HP.
- *Q* mariages (bindings) between seeds and MREs are pronounced at each iteration.
- α is a dilution factor to prevent that all seeds of a species bind the same mRNA.





Bartel (2009) Cell

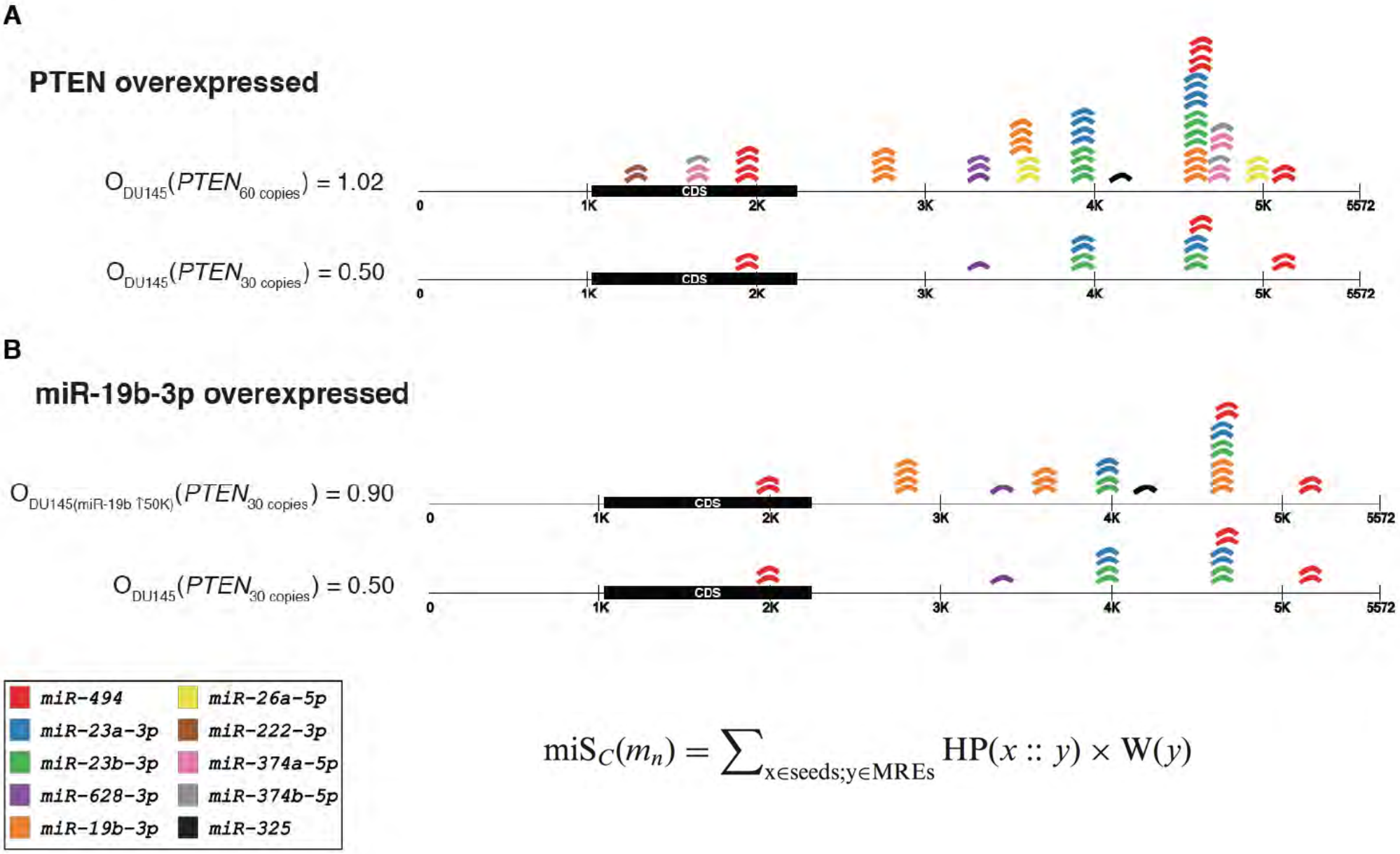
$$\mathrm{HP}(\mathrm{seed}::\mathrm{MRE}) = \frac{e^{\frac{-\Delta G(\mathrm{seed}::\mathrm{MRE})}{kT}}}{\sum_{h \in heptamers} e^{\frac{-\Delta G(seed::h)}{kT}}}$$

$$Q(\mathrm{seed},\mathrm{MRE}) = q(\mathrm{MRE}) \times log_{\alpha}\ q(\mathrm{seed}) \times HP(\mathrm{seed}::\mathrm{MRE})$$

- The miRNA::mRNA competition results in networks that depend on cellular contexts
- The silencing contribution of a given miRNA on:
  - a highly abundant mRNA is infinitesimal (because of collective targeting)
  - a lowly abundant mRNA can be significant

Weill et al. 2015 *NAR*

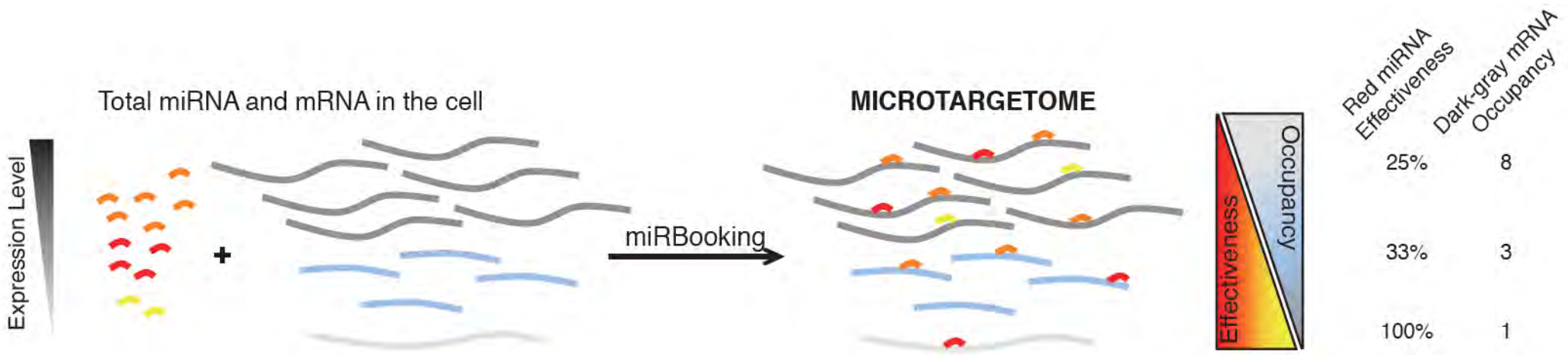
PTEN occupancies change in various DU145 conditions



$$\mathrm{miS}_C(m_n) = \sum_{x \in \mathrm{seeds}; y \in \mathrm{MREs}} \mathrm{HP}(x :: y) \times \mathrm{W}(y)$$

Weill et al. 2015 *NAR*

MiRNA efficiencies to silence a given gene decreases as its expression increases

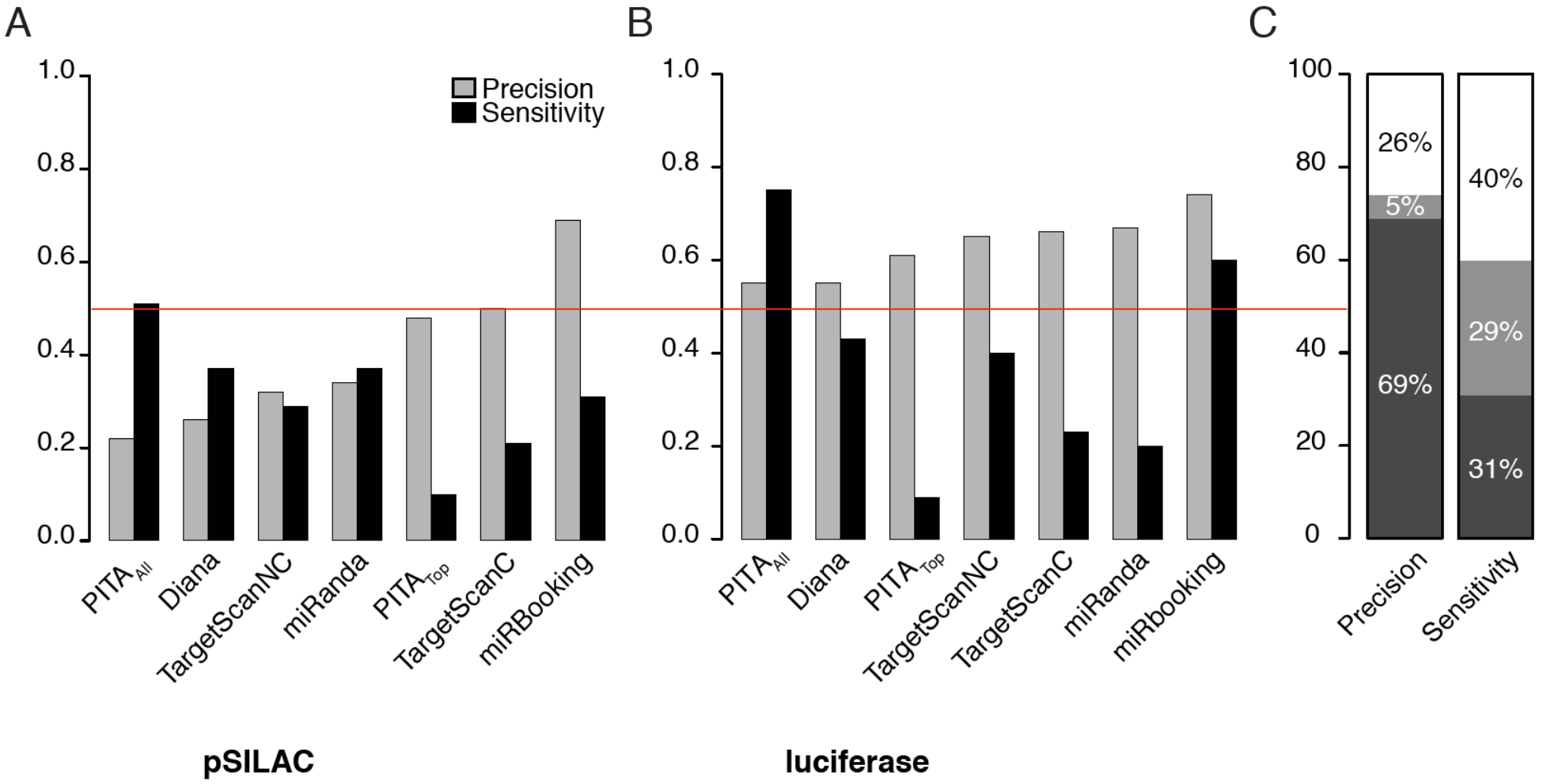


$$\mathrm{miS}_C(m_n) = \sum_{\mathrm{x}\in\mathrm{seeds};\mathrm{y}\in\mathrm{MREs}} \mathrm{HP}(x :: y) \times \mathrm{W}(y)$$

$$\mathrm{efficiency}(x, m_n) = \Sigma_{x;y\in\,\mathrm{MREs}}\, \mathrm{HP}(x::y)\ \mathrm{x}\ \mathrm{W}(y) \,/\, \mathrm{miS}_C(m_n)$$

Weill et al. 2015 *NAR*

Solving the hybridization competition using stable matching based on seed matching and abundances improves miRNA target prediction accuracy



Weill et al. NAR 2015

MiRBookiong2: we characterized the microtargetome by a set of enzymatic reactions, where the miRNA and mRNA concentrations are estimated by RNAseq data

$$[E_m] + [S_{t,p}] \underset{k_r}{\overset{k_f}{\rightleftharpoons}} [E_m S_{t,p}] \xrightarrow{k_{cat} + k_{other}} [E_m] + [P_t]$$

Each reaction is represented by the Michaelis-Menten equation
The energy is similar to miRBooking1

$$\frac{\partial [E_m S_{t,p}]}{\partial t} = k_f [E_m][S_{t,p}] - k_r [E_m S_{t,p}] - k_{cat} [E_m S_{t,p}] - \underbrace{\left(\sum_{m',p'} \frac{k_{cat} [E_{m'} S_{t,p'}]}{[S_t]} \right)}_{k_{other}} [E_m S_{t,p}]$$

Each reaction has an associated differential equation that characterizes its kinetic

$$k_f [E][S] = k_r [ES] + k_{cat} [ES] + k_{other} [ES]$$

The steady state is an equilibrium point between formation and dissociation where forward and reverse states are equal:

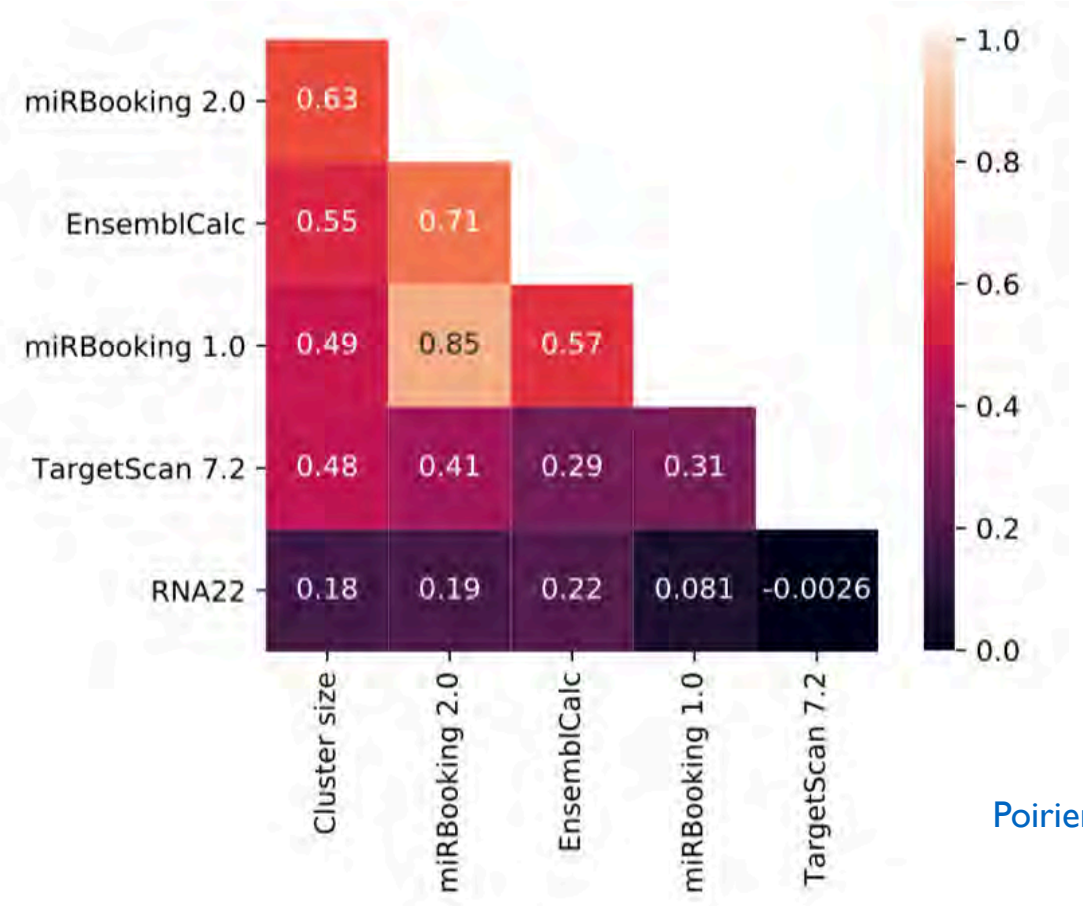
$$\frac{\partial [ES]}{\partial t} = 0$$

Poirier-Morency

We benchmarked miRBooking2 using HITS-CLIP data

Dataset of ~60000 points:

- Noisy data (r ~ 0.4 across technical replicates)
- RNA abundances extrapolated from ENCODE (HeLa S3)
- Molar concentration are grossly estimated.



Poirier-Morency

MiRBooking summary

- Simulating miRNA::mRNA hybridization using a stable matching algorithm using preferences defined by seed matching and abundances increases miRNA target predictions
- It produces miRNA::mRNA networks that depend on cellular contexts
- The efficiency of a given miRNA to silence a highly abundant mRNA is low because abundant mRNAs are heavily targeted by many miRNAs
- A low abundant mRNA, however, can be efficiently targeted by a single miRNA
- A consequence of the crosstalk mechanism is that miRNAs participate in the synchronization of the expression of groups and subgroups of mRNAs (not discussed in this presentation)

MiRBooking allows us to design small artificial (smart) RNAs

- Smart RNAs target multiple genes efficiently with minimum effects on off-targets

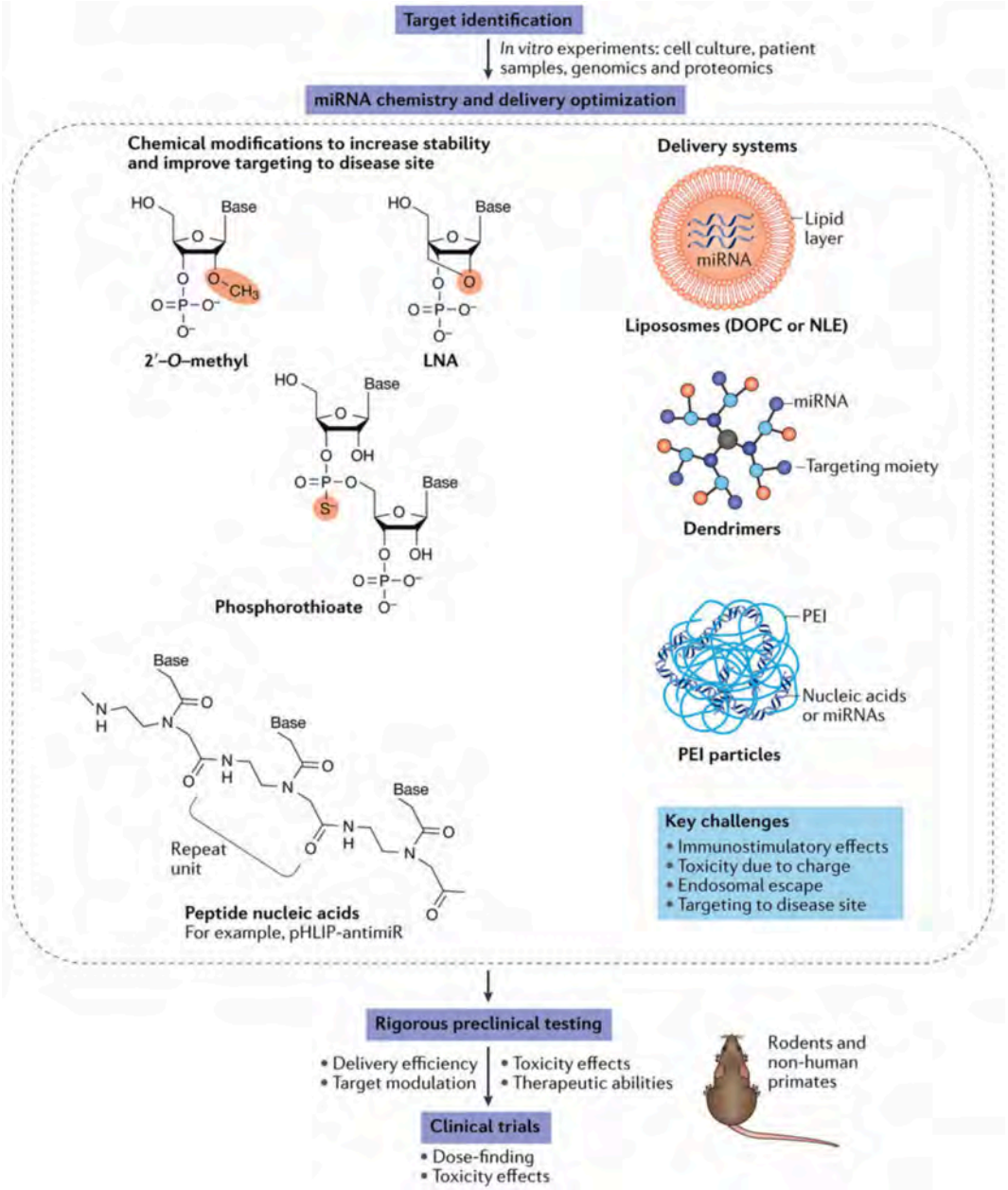
- Smart RNAs can be used in RNAi-based therapeutics of complex diseases, requiring the control of the expression of multiple genes

First-Ever FDA Approval of an RNAi Therapeutic, ONPATTRO™ (patisiran) for the Treatment of the Polyneuropathy of Hereditary Transthyretin-Mediated Amyloidosis in Adults

10 August 2018

Small interfering RNA (siRNA), the molecules that mediate RNAi and comprise Alnylam's RNAi therapeutic platform, **function upstream** of today's medicines by potently silencing messenger RNA (mRNA)

Small RNAs can be delivered safely and efficiently



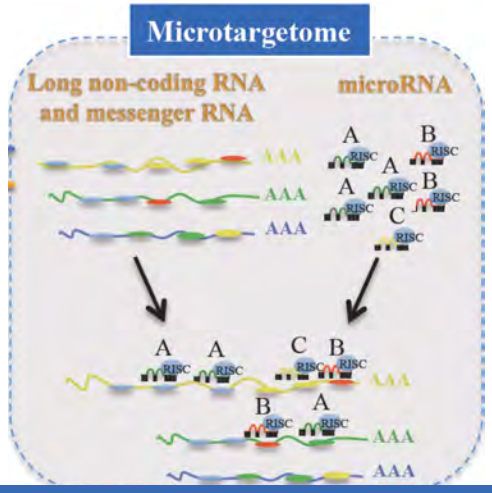
Rupaimoole & Slack (2017) *Nat Reviews*

What to consider for a rational design of smart RNAs

MiRNAs and mRNAs form couples based on a wysiNwyg principle (seed and abundance), as the time they stay married (translational repression efficiency) depends on their complete complementarity, i.e. beyond what they present at each others at first sight (beyond the seed)

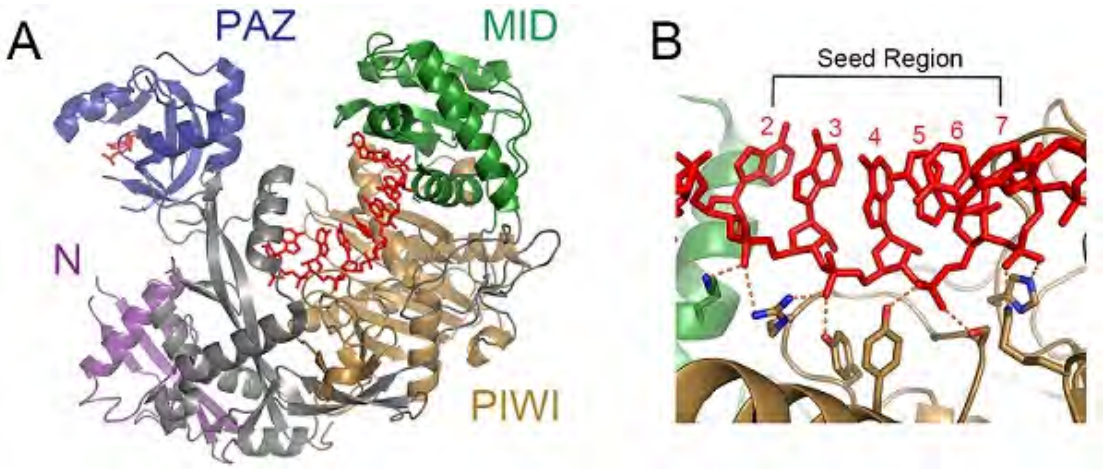
> *miRBooking*: finds the stable marriages in a transcriptome (microtargetome) from expression data and seed complementarity

> *miScore*: estimates the duration of marriages from complete sequence complementarity



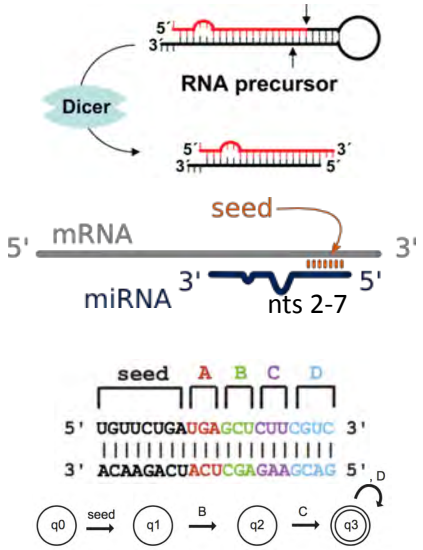
The miRNA::mRNA interaction network

Weill et al. (2015) *Nucl Acids Res*



Only the seed is exposed and determines miRNA::mRNA interactions

Ameres et al. (2007) *Cell*



Yan et al. (2018) *Nucl Acids Res*

We investigated the pairing interactions between a miRNA and target to corroborate a hierarchical recognition model where the seed binds first and then nucleotides in the 3' region are able to bind the target

Yifei Yan

Mismatches near the centre of the guide-target duplex are tolerated

Doench et al. 2003 Genes & Dev (Sharp)

Mismatches at the 3'-end of the duplex were found to enhance miRISC-mediated gene silencing

De et al. 2013 Mol Cell (MacRae)

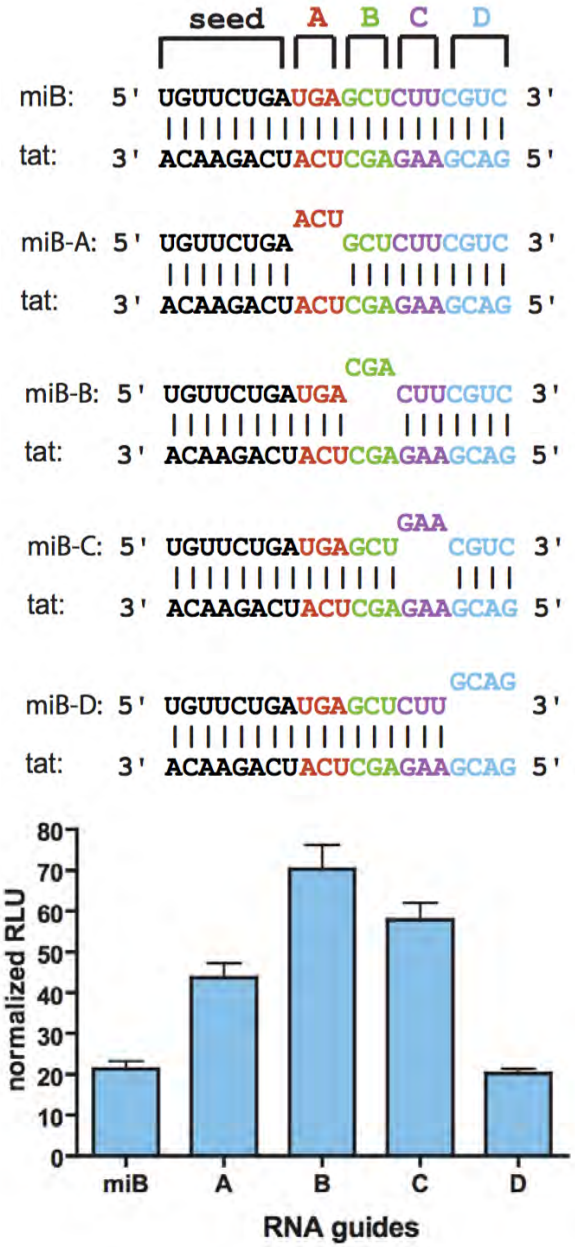
Sequence and miRNA::mRNA structure influence efficiency (but provide little improvement)

Kiriakidou et al. 2004 Genes & Dev
Ye et al. 2008 PloS One

Hypothesis: Understanding and integrating beyond-the-seed effects may improve miS estimations...

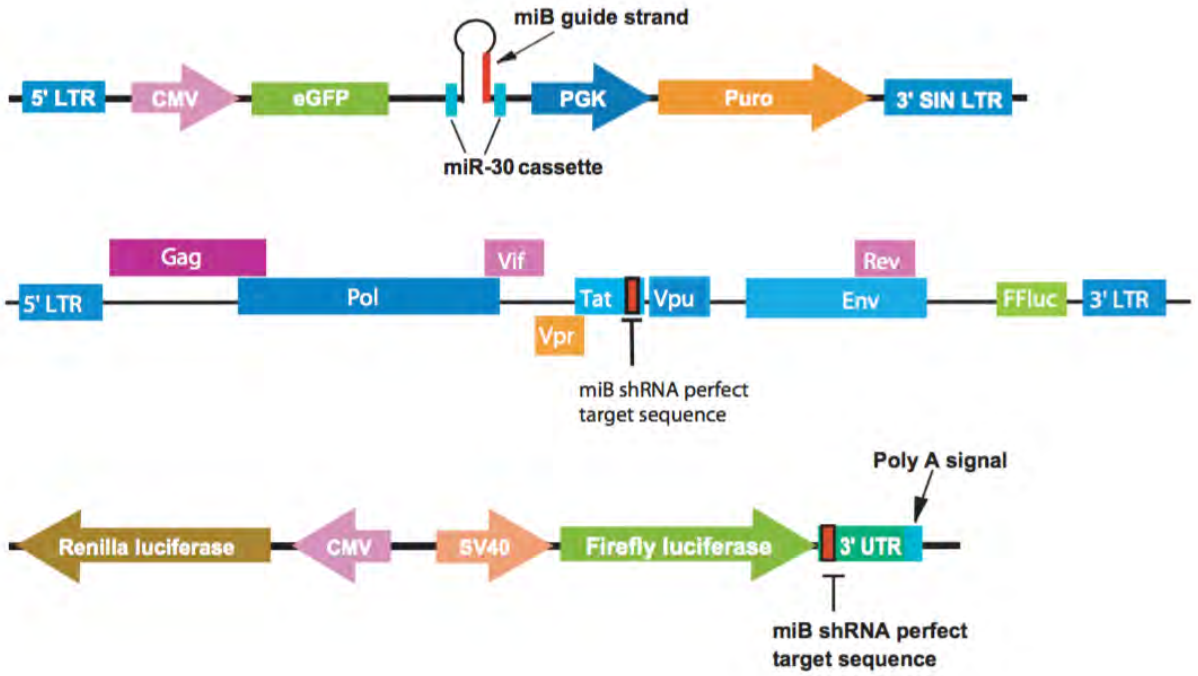
Luciferase reporters were used to tile the region beyond-the-seed

(Boxes: seed, A, B, C, D)



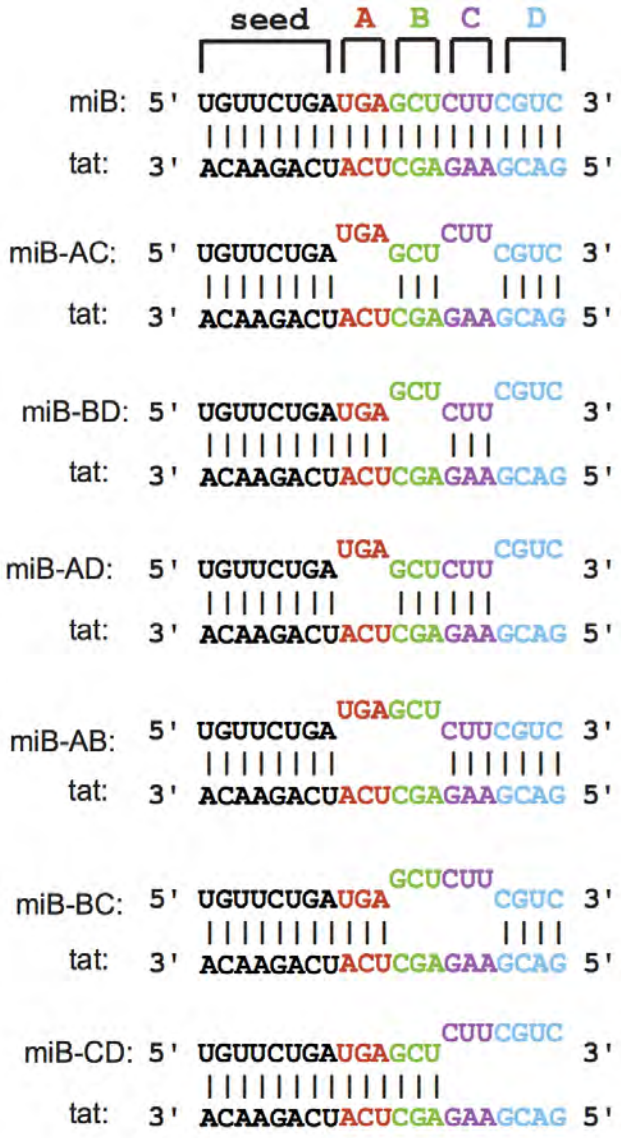
miB is a perfectly complementary shRNA against the *tat* gene of HIV

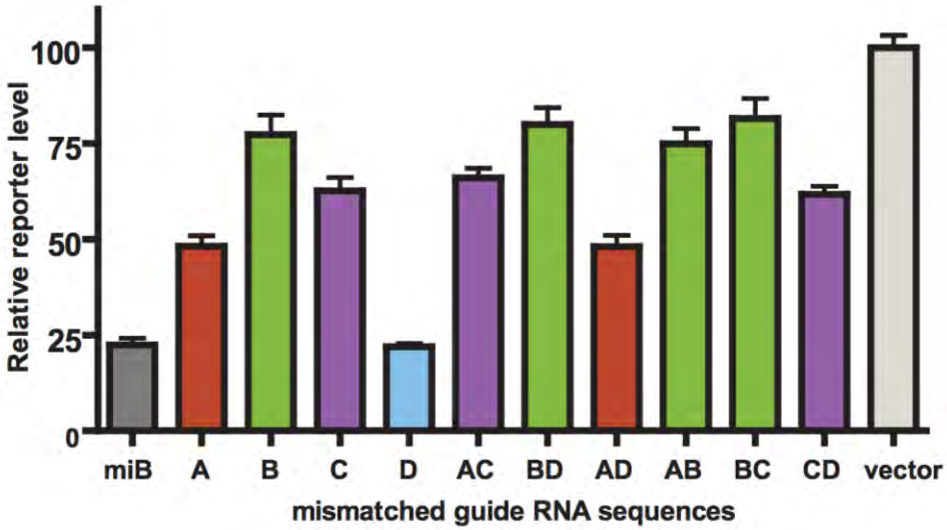
Boden et al. 2003 J Virol
Boden 2004 NAR

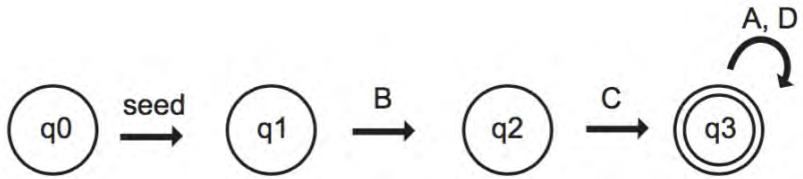


Yan et al. 2018 *NAR*

Dual-box reporter assays reveal a hierarchical pattern
(finite-state machine)

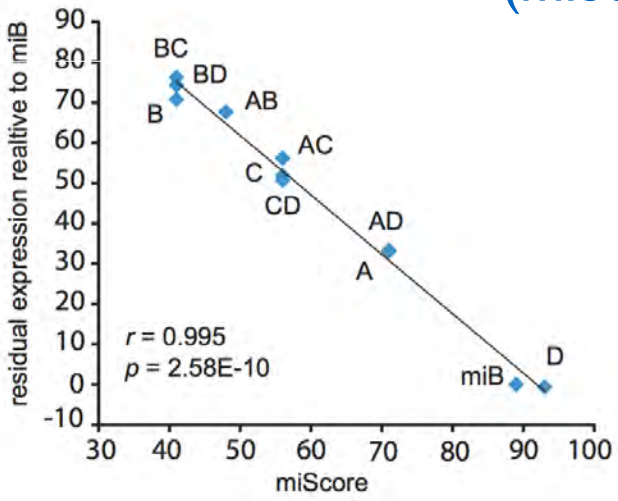


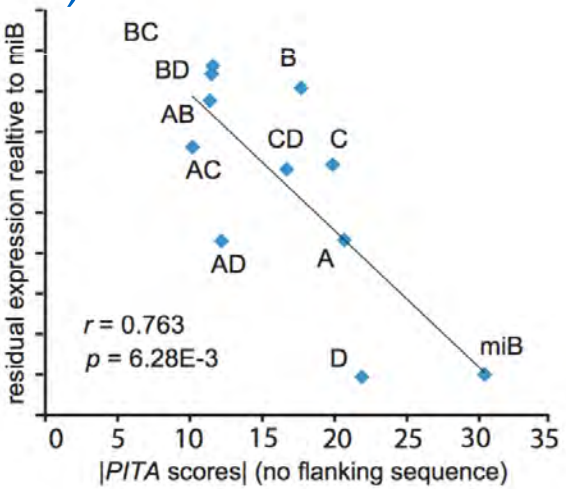




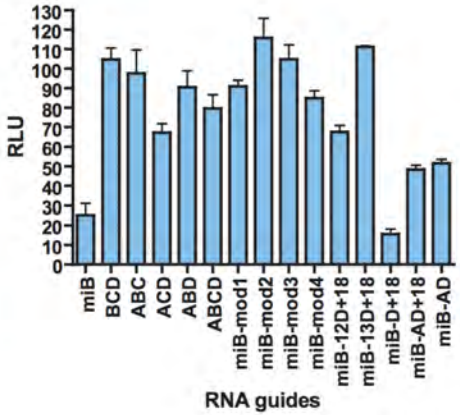
Yan et al. 2018 *NAR*

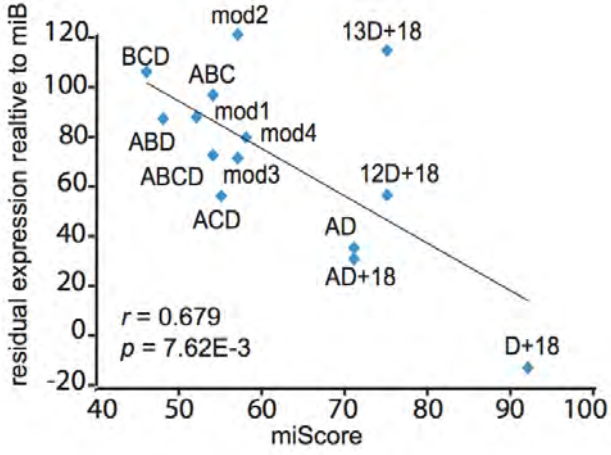
Beyond-the-seed silencing efficiency scores correlate with experimental data (miScores)

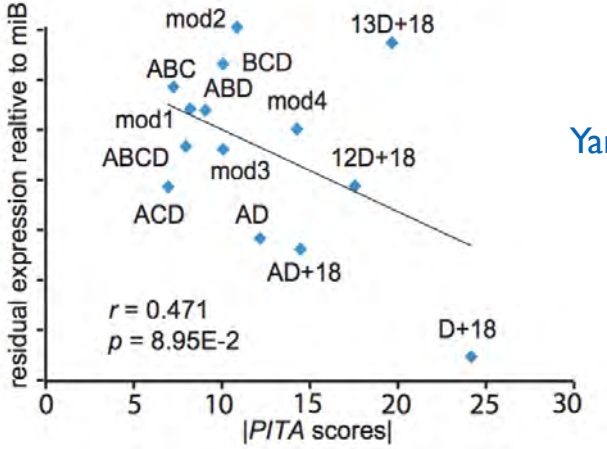




Oligo Name	5'->3' guide sequences
miB	UGUUCUGAUGAGCUCUUCGUC
miB-BCD	UGUUCUGAUGACGAGAAGCAG
miB-ACD	UGUUCUGAACUGCUGAAGCAG
miB-ABD	UGUUCUGAACUCGACUUGCAC
miB-ABC	UGUUCUGAACUCGAGAACGUC
miB-ABCD	UGUUCUGAACUCGAGAAGCAG
miB-mod1	UGUUCUGAUAAACGCUGAGUC
miB-mod2	UGUUCUGAUGGAGUCUUAGAG
miB-mod3	UGUUCUGAAGACCUAUCCGUC
miB-mod4	UGUUCUGAUGGCCUCCCAGUC
miB-12D+18	UGUUCUGAUGACCUCUUCCAG
miB-13D+18	UGUUCUGAUGAGGUCUUCCAG
miB-D+18	UGUUCUGAUGAGCUCUUCCAG
miB-AD+18	UGUUCUGAACUGCUCUUCCAG







Yan et al. 2018 *NAR*

MiScores predict public data and efficient smartRNAs against E2F1-3 to reduce PC3 cell proliferation

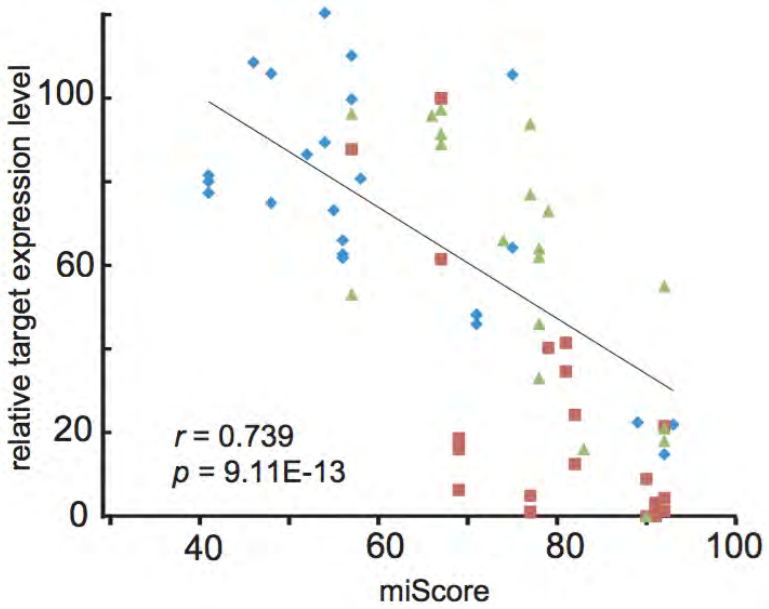
Catalytic efficiency (K_{cat}/K_m) measured for AGO2 *in vitro*, where mismatches were systematically generated in the guide RNA

Wee et al., 2012 Cell (Zamore)

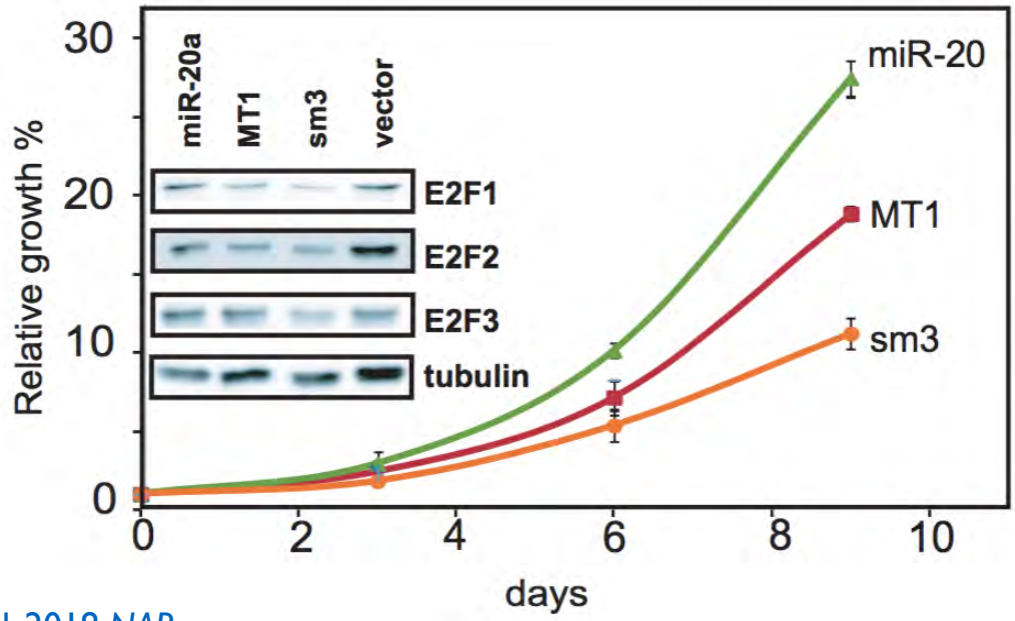
MiRNA sponges engineered with dinucleotide mismatches tiling the entire non-seed region

Robertson et al., 2010 Silence (Dharmacon)

Against public data

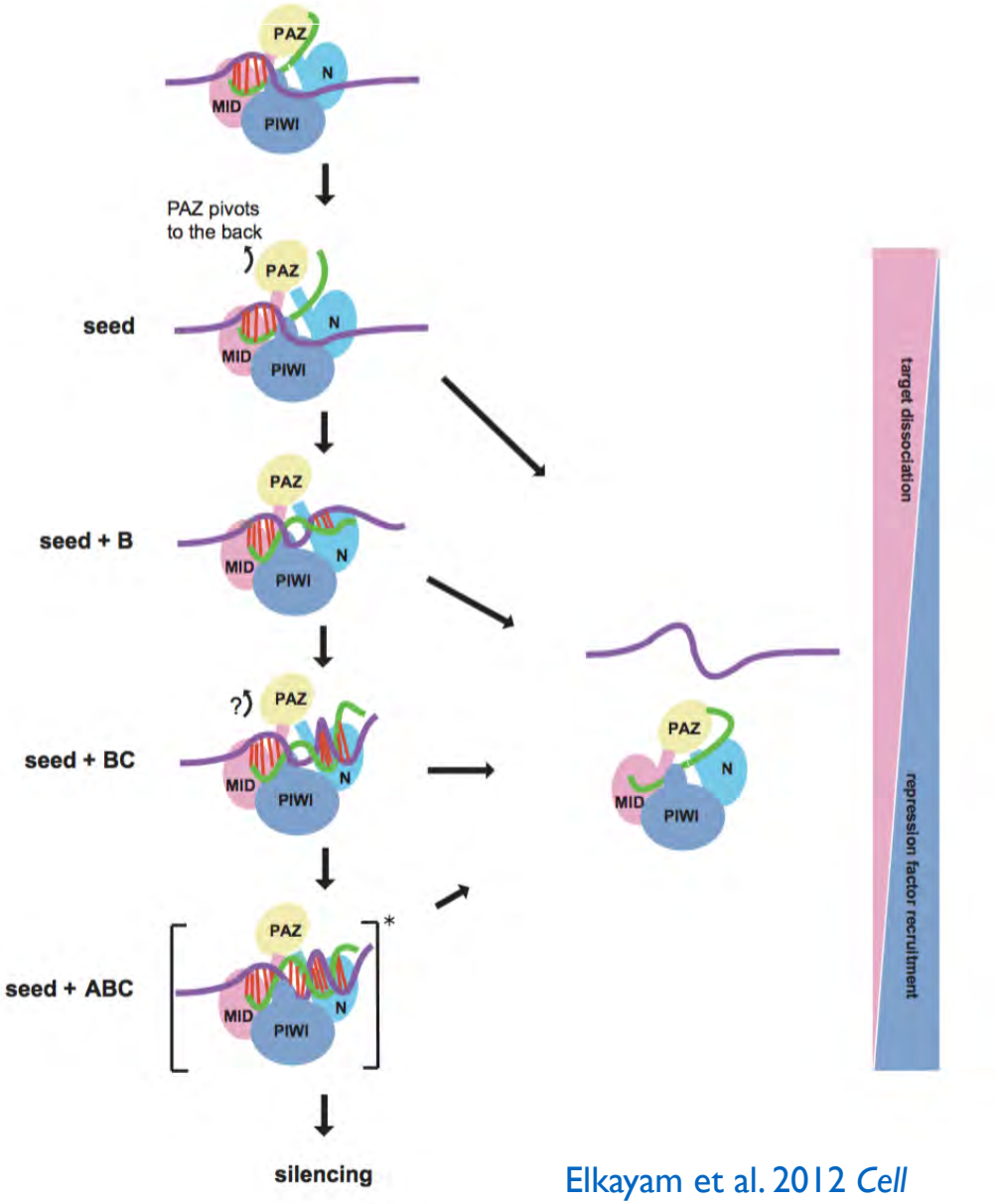


@work targeting E2F1-3 in PC3 cells



Yan et al. 2018 *NAR*
De Guire et al. 2010 *NAR*

Sequential base pairing on both flanks of the scissile nucleotide leads to efficient silencing



Yan et al. 2018 *NAR*

Elkayam et al. 2012 *Cell*
Schirle & MacRae 2012 *Science*
Schirle et al. 2014 *Science*

The silencing complex sequentially allows for base pairing on both sides of the scissile nucleotide leading to efficient silencing.

"The agents of natural genome editing"
Witzany 2011 *J Mol Cell Biol*

Similar step-wise binding pathways in:

- The *par* RNAI and RNAII in *Enterococcus faecalis* Greenfield 2001 Mol Microbiol
- Clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR associated proteins (Cas) system in Bacteria and Archaea Semenova et al. 2011 PNAS
- RNA-editing enzymes in the Euglenozoa phylum Kiethega et al. 2013 RNA Biol
  Stuart et al. 1997 Microbiol Mol Biol Rev
- Hammerhead ribozymes in all domains of life
  Perreault et al. 2011 PLoS Comput Biol (Breaker)

Application to study the Epithelial-Mesenchymal Transition (EMT)

- Breast cancer is the second most important cancer in the world and the most frequent among women
- It is the second leading cause of death for women (in Canada) after lung cancer
- The mortality related to the disease is mainly caused by metastases
- One of the steps in the formation of metastases is the epithelial-mesenchymal transition (EMT)
- The EMT is a dynamic and reversible process that refers to the passage of epithelial cells to a mesenchymal phenotype leading to the invasion of cancer cells in the body through the bloodstream, and to the development of cancer in other organs
- This transition results in a phenotypic difference caused by the loss of intercellular junctions, a depolarization of epithelial cells, and the acquisition of cell motility capacities
- Epithelial cells are characterized by a narrow mobility caused by the presence of E-cadherin, a molecule that maintains cell junctions
- Mesenchymal cells are not polarized and differ from epithelial cells by the loss of adhesion and the absence of apical and lateral membrane, characterized by invasive properties of tumor cells to surrounding tissues while expressing mesenchymal markers such as vimentin
- The EMT is mainly triggered by stress conditions and signals from the microenvironment, such as hypoxia or inflammation
- New signaling pathways are then activated including the Wnt, Hedgehoh, and Notch, as well as pathways taken by cytokines and growth factors, such as the TGF-β/BMP, PDGF, and HGF. In turn, these pathways activate the transcription factors Twist, Snail, and Zeb, which induce the EMT.

Rational design of smartRNA-based therapeutics

① **Establish condition specific microtargetome (e.g. hypoxia-induced melanoma; EMT-induced metastases) from transcriptome and mirnome data using miRBooking ($$$)**

② **Identify/Select target genes (e.g. hypoxia resistance and EMT involved genes) using expression data, literature, pathway analysis, and so on (🕑🕑🕑)**

③ **Design smartRNAs to target the selected genes (💻)**

④ **Measure the effects of designed smartRNAs in primary tumors (👷)**

① Establish condition specific microtargetome (e.g. hypoxia-induced melanoma; EMT-induced metastases) from transcriptome and mirnome data using miRBooking ($$$)

- Using miRBooking, we built the microtargetome of the A549 cells
- The A549 cell line is widely used to study the EMT
- A549 cells are derived from pulmonary adenocarcinoma.

score = k_{cat}/K_m (enzymatic efficiency)

A549-control-mirbooking-with-enzymatic-efficiency

	target_accession	target_name	target_quantity	position	mirna_accession	mirna_name	mirna_quantity	score	quantity
0	ENST00000387347.2	MT-RNR2	70,792.28	7	MIMAT0000080	hsa-miR-24-3p	345.51	7.270418364107398e-09	0.08
1	ENST00000387347.2	MT-RNR2	70,792.28	25	MIMAT0000072	hsa-miR-18a-5p	474.25	3.129019241512698e-07	33.63
2	ENST00000387347.2	MT-RNR2	70,792.28	33	MIMAT0000732	hsa-miR-378a-3p	84.72	6.598778071270652e-09	0.11
3	ENST00000387347.2	MT-RNR2	70,792.28	41	MIMAT0000072	hsa-miR-18a-5p	474.25	1.2379320963296862e-08	1.33
4	ENST00000387347.2	MT-RNR2	70,792.28	45	MIMAT0000080	hsa-miR-24-3p	345.51	2.3523161983433816e-08	0.24
5	ENST00000387347.2	MT-RNR2	70,792.28	99	MIMAT0000429	hsa-miR-137-3p	232.36	2.1370242115938908e-08	43.50
6	ENST00000387347.2	MT-RNR2	70,792.28	118	MIMAT0019208	hsa-miR-3074-5p	172.71	3.084876124871158e-09	0.05
7	ENST00000387347.2	MT-RNR2	70,792.28	119	MIMAT0019208	hsa-miR-3074-5p	172.71	2.5479348828528548e-08	0.45
8	ENST00000387347.2	MT-RNR2	70,792.28	172	MIMAT0000072	hsa-miR-18a-5p	474.25	1.2580225144095996e-08	1.36
9	ENST00000387347.2	MT-RNR2	70,792.28	176	MIMAT0000101	hsa-miR-103a-3p	330.80	3.3450195307327847e-09	0.03
10	ENST00000387347.2	MT-RNR2	70,792.28	213	MIMAT0000080	hsa-miR-24-3p	345.51	1.6939343035812595e-09	0.02
11	ENST00000387347.2	MT-RNR2	70,792.28	214	MIMAT0000080	hsa-miR-24-3p	345.51	1.1907998801261455e-06	12.20
12	ENST00000387347.2	MT-RNR2	70,792.28	220	MIMAT0000076	hsa-miR-21-5p	9,131.92	8.010076676458986e-09	102.29
13	ENST00000387347.2	MT-RNR2	70,792.28	237	MIMAT0000732	hsa-miR-378a-3p	84.72	6.493005815298334e-09	0.11
14	ENST00000387347.2	MT-RNR2	70,792.28	243	MIMAT0000076	hsa-miR-21-5p	9,131.92	5.6180827356317535e-11	0.72
15	ENST00000387347.2	MT-RNR2	70,792.28	244	MIMAT0000076	hsa-miR-21-5p	9,131.92	1.0708668444008166e-08	136.67
16	ENST00000387347.2	MT-RNR2	70,792.28	257	MIMAT0000430	hsa-miR-138-5p	636.80	6.9265028779619465e-09	0.12
17	ENST00000387347.2	MT-RNR2	70,792.28	258	MIMAT0000077	hsa-miR-22-3p	47.34	1.1941819455612259e-07	1.16
18	ENST00000387347.2	MT-RNR2	70,792.28	258	MIMAT0000076	hsa-miR-21-5p	9,131.92	9.920983500164317e-10	12.66
19	ENST00000387347.2	MT-RNR2	70,792.28	277	MIMAT0000757	hsa-miR-151a-3p	165.22	6.185731219153498e-09	1.79

② Identify/Select target genes (e.g. hypoxia resistance and EMT involved genes) using expression data, literature, pathway analysis, and so on (🕒🕒🕒)

Based on the microtargetome of the A549 cells, your **assignment** will consist in identifying:

- ***OncomiRs***, i.e. miRNA that target tumor-suppressor protein genes
- ***Tumor-suppressor-miRs***, i.e. miRNA that target protein oncogenes

1.You will first find a list of cancer genes (oncogenes and tumor suppressors)
2.You will also learn about EMT genes (specific to the case of pulmonary cancer)
3.You will then identify in the microtargetome which miRNAs target these cancer genes

Therapeutically:

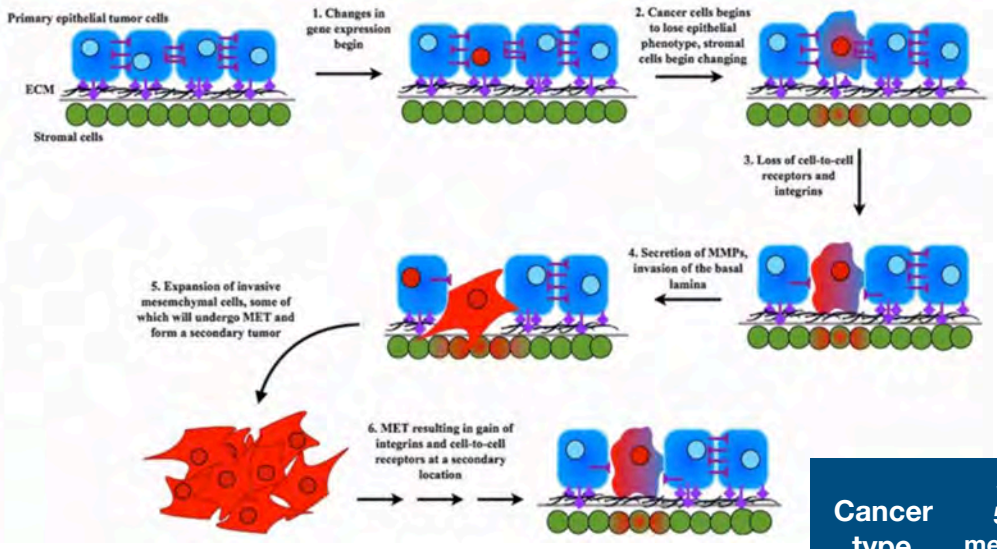
Oncomirs can be eliminated using antimiRs (RNA antisense)
Tumor-suppressor-miRs are direct therapeutics

Both antimiRs and tumor-suppressor-miRs can be chemically optimized and stabilized, and delivered (see slide 19)

An alternative is to design smart RNAs against a selection of targets

Let's take a look at some EMT markers and players...

The EMT triggers tumor metastasis



At least two studies suggest that targeting the EMT may inhibit both cancer metastasis and chemoresistance

Fischer et al. (2015) *Nature*
Zheng et al. (2015) *Nature*

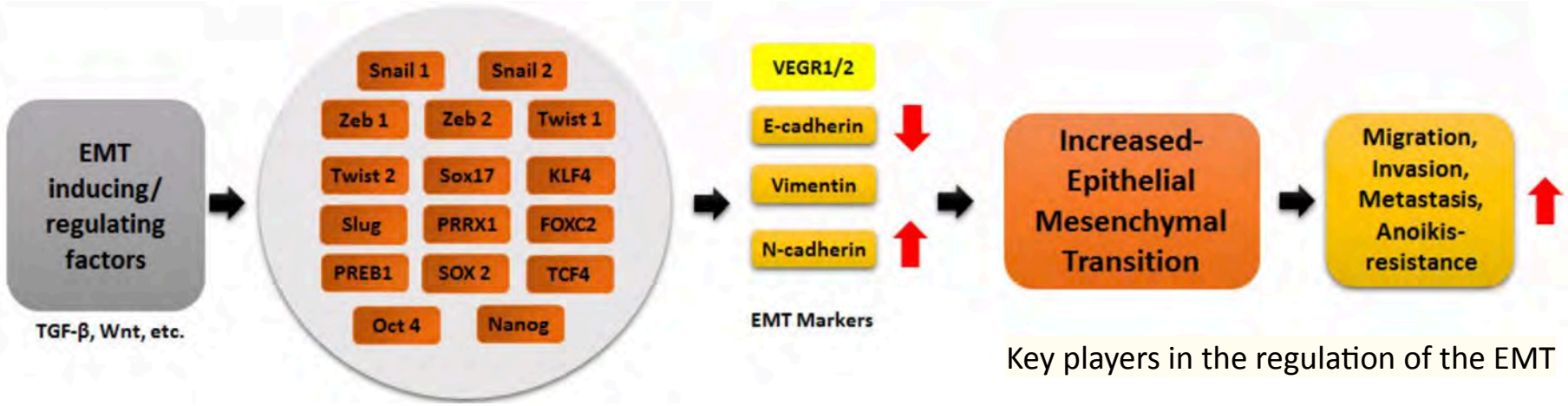
miRNA	Effect on EMT	Target
miR-9	Promote	CDH1
miR-15b	Suppress	BMI1
miR-27	Promote	APC
miR-29a	Promote	TTP
miR-30a	Suppress	Snail
miR-103/107	Promote	DICER1
miR-155	Promote	RHOA
miR-194	Suppress	BMI1
miR-200	Suppress	ZEB1/2, Sec23a
miR-205	Suppress	ZEB1/2
miR-204	Suppress	TGF-BR2, SNAIL2
miR-221/222	Promote	TRPS1, ESR1, DICER1
miR-661	Promote	StarD10, Nectin-1

Cancer type	Survival 5-y after metastasized (%)	Survival 5-y after diagnosis (%)	EMT Markers
Pancreas	2.3	6.7	Snail, Twist, Zeb1, Zeb2, E-cadherin, β-catenin Brachyury, HDAC1,2,3, **miR-34**, **miR-200**
Liver	2.8	16.6	Snail, Twist, Zeb1, Zeb2, TGF-β, EZH2, HDAC1,2,3, **miR-101**, STAT3, SUZ12
Lung	4.0	16.8	Snail, Zeb1, Zeb2, E-cadherin, vimentin, α-catenin, EZH2, BMI1, Brachyury, Claudin-1, Cytokeratins, G9a, HDAC1,2,3, LSD1, **miR-34**, **miR-101**, **miR-205**, Periostin, Slug, SUZ12, TTF-1, versican, N-cadherin
Bladder	5.5	77.4	Twist, Zeb1, Zeb2, N-cadherin, EZH2, Fibronectin, LSD1, **miRs-1/133a/218**, **miR-19a**, **miRs-30a-3p/133a/199a**, **miR-34**, **miR-99a/100**, **miR-101**, **miR-125b**, **miR-129**, **miR-145/133a**, **miR-200**, **miR-205**, **miR-221**
Colorectal	12.9	64.7	Snail, Twist, vimentin, Zeb1, Zeb2, β-catenin, Brachyury, CD44, E-cadherin, EZH2, FGFR4, Fibronectin, HDAC1,2,3, LSD1, **miR-34**, p16INK4a, SIRT1, Slug, SUZ12, SUV39H1
Skin melanoma	16.1	91.3	TGF-β, MITF, N-cadherin, **miR-205**
Breast	25.0	89.2	Snail, Zeb1, Zeb2, vimentin, β-catenin, E-cadherin, BMI1, Brachyury, Claudin, EZH2, HDAC1,2,3, Klf8, LSD1, miR-9 (2); **miR-10b**, **miR-34**, Slug, SUZ12, Twist, versican
Prostate	28.0	98.9	Twist, Zeb1, N-cadherin, APC, Cyclin D2, collagen, decorin, E47, E-cadherin, ER, EZH2, Fibronectin, GSTP1, HDAC1,2,3, **Let-7a**, LSD1, **miR-1**, **miR-7**, **miR-15a-16** cluster, **miR-20a**, **miR-21**, **miR-24**, **miR-32**, **miR-34a**, **miR-34c**, **miR-101**, **miR-106b**, **miR-107**, **miR125b**, **miR-143**, **miR-145**, **miR-146a**, **miR-148a**, **miR-205**, **miR-221**, **miR-222**, **miR-331-3P**, **miR-449a**, **miR-521**, **miR-1296**, Notch-1, RAR-β2, RASSF1A, versican

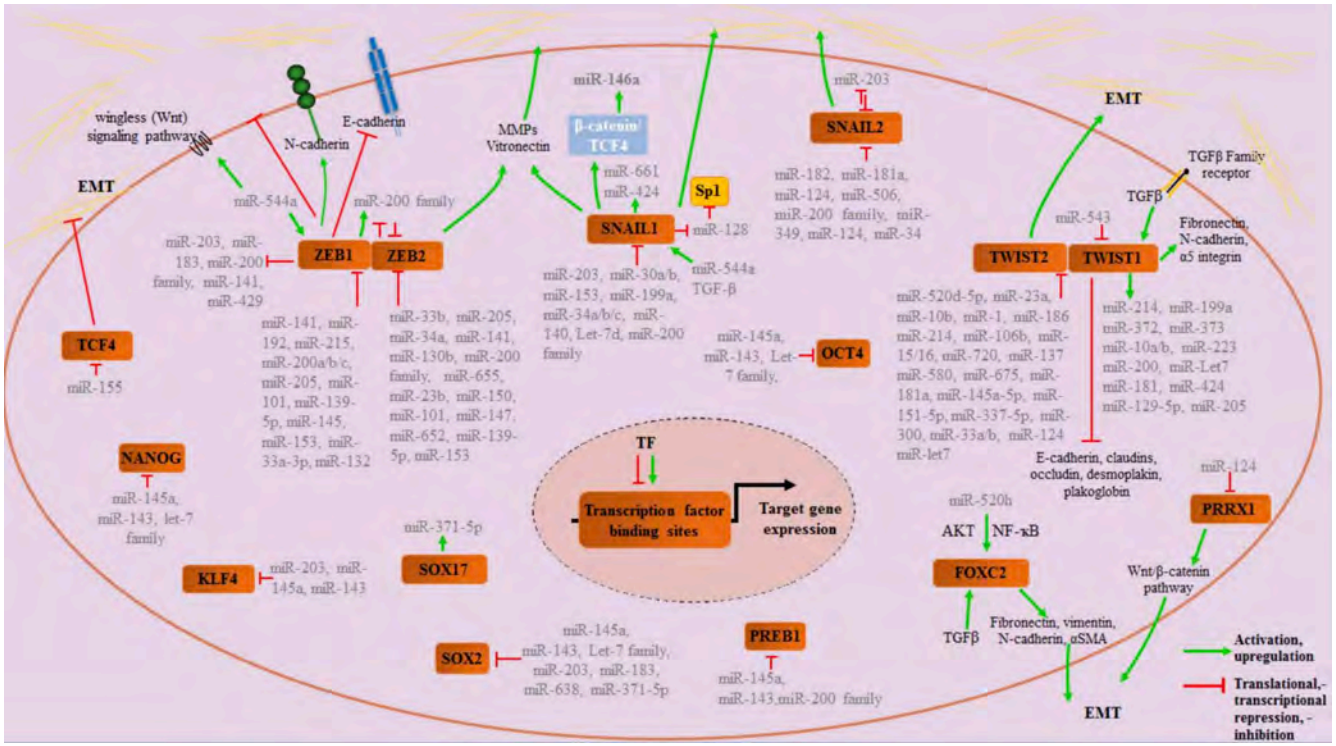
Zhang & Ma (2012) *Cancer Metastasis Rev*

Heerboth et al. (2015) *Clin Transl Med*

The fore(wo)men of the EMT are miRNAs?



Key players in the regulation of the EMT



A group of developmental transcription factors form the backbone of the EMT cascade and a large body of evidence shows that microRNAs are heavily involved in the successful coordination of mesenchymal transformation and *vice versa*, either by suppressing the expression of different groups of transcription factors, or otherwise acting as their functional mediators in orchestrating EMT.

Abba et al. (2016) *J Clin Med*

Significant miRNAs in the regulation of the EMT players

③ Design smartRNAs against hypoxia resistance

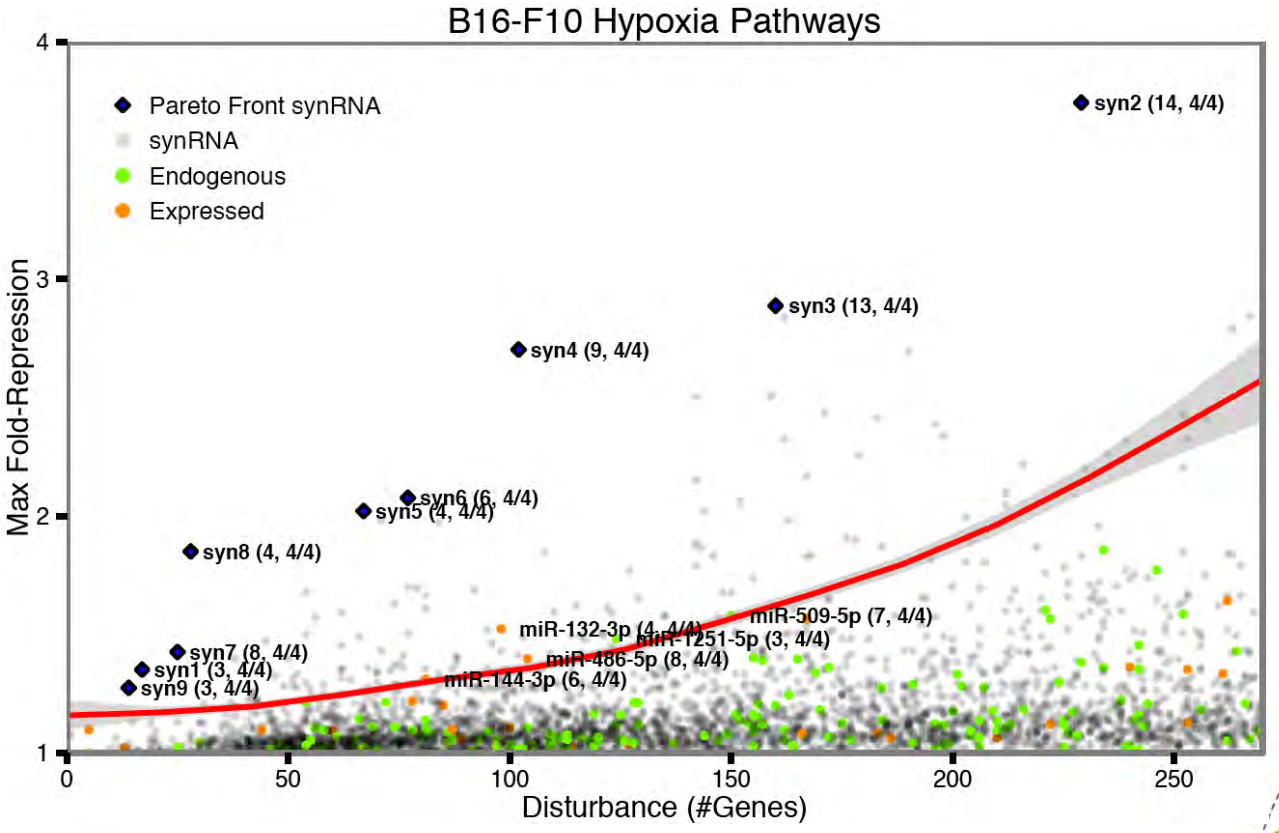
Run 2^7=16,384 miRBooking simulations
(all possible seed expression in melanoma tumor cells)

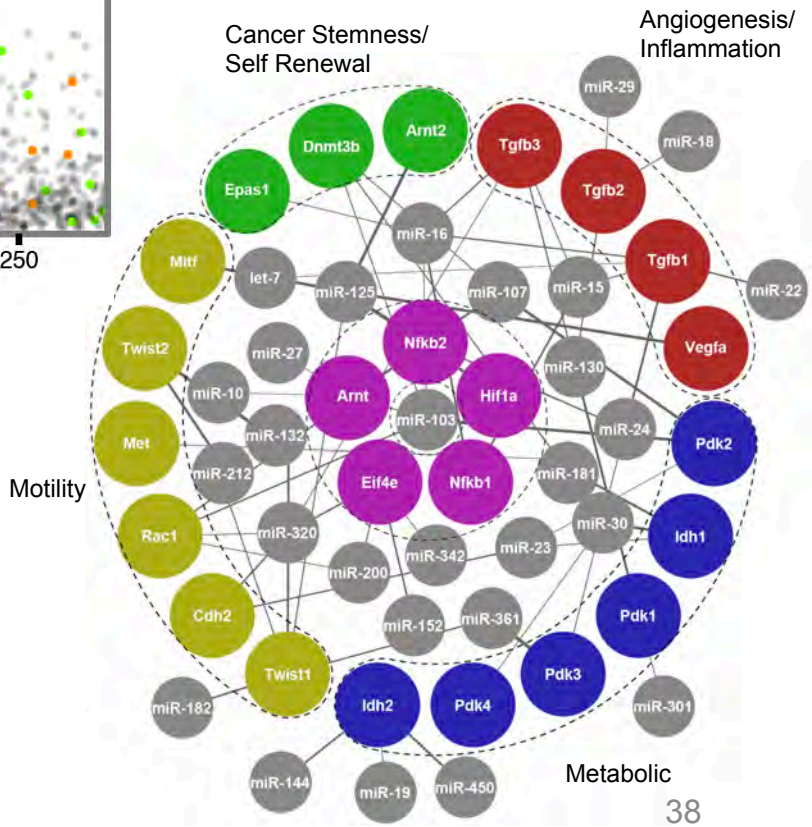
Determine the best seeds to TARGET hypoxia resistance pathways

Filter the best seeds to avoid NON TARGET effects (Disturbance)

Compute 3' sequence to optimize silencing efficiency
a. siRNA on HIF-1α and miRNA on all other targets
b. MiScores for efficiency

- Less than 3,000 endogenous mature miRNAs have been reported in humans so far
- MiRDesign screens 16,384 seeds (giving rise to 1.7 x 10^{13} possible mature miRNA sequences)
- Preliminary data indicate that efficient smartRNAs with relatively small disturbance can be found, regardless of the context or pathways/genes targeted
- MiRNAs are natural compounds and it is expected that the payers and public in general should see such therapeutics positively


B16-F10 Hypoxia Pathways
Pareto Front synRNA
synRNA
Endogenous
Expressed
Max Fold-Repression
Disturbance (#Genes)
syn2 (14, 4/4)
syn3 (13, 4/4)
syn4 (9, 4/4)
syn6 (6, 4/4)
syn5 (4, 4/4)
syn8 (4, 4/4)
syn7 (8, 4/4)
syn1 (3, 4/4)
syn9 (3, 4/4)
miR-132-3p (4, 4/4)
miR-509-5p (7, 4/4)
miR-1251-5p (3, 4/4)
miR-486-5p (8, 4/4)
miR-144-3p (6, 4/4)

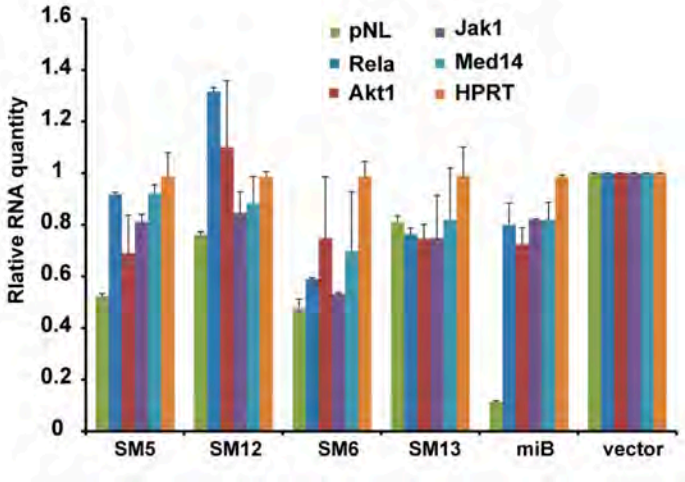

Cancer Stemness/
Self Renewal
Angiogenesis/
Inflammation
Motility
Metabolic
miR-29
Dnmt3b
Arnt2
Tgfb3
miR-18
Epas1
Tgfb2
miR-16
Mitf
let-7
miR-125
miR-107
miR-15
Tgfb1
miR-22
Twist2
miR-27
Nfkb2
Vegfa
miR-130
miR-10
Arnt
Hif1a
miR-103
miR-24
Met
miR-132
Pdk2
miR-212
Eif4e
Nfkb1
miR-181
Rac1
miR-320
miR-30
Idh1
miR-342
miR-23
miR-200
Cdh2
Pdk1
miR-152
miR-361
Twist1
Pdk3
miR-182
Idh2
Pdk4
miR-301
miR-144
miR-19
miR-450

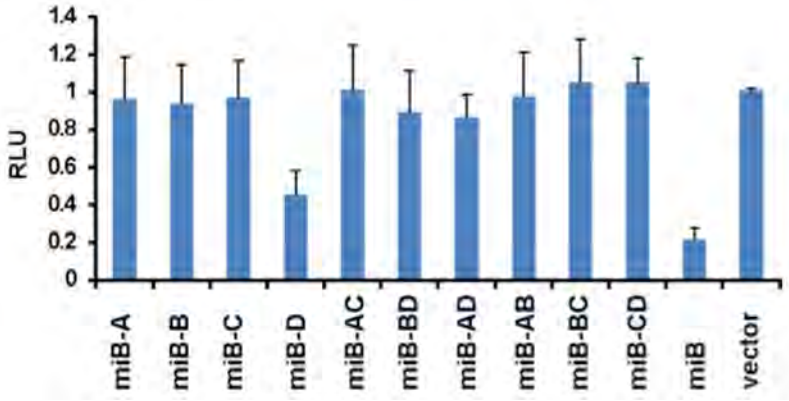
Rational design vs existing shRNA libraries

- There are 4 nucleotides (if we do not consider the modified ones)
- The number of off-targets predicted to by affected by more than two-fold by one miRNA (disturbance) varies from 0 to 400 (almost uniform distribution between 50 and 250)

- shRNA libraries were designed to target efficiently one mRNA by using complete and perfect complementarity
- For the disturbance, siRNAs (and shRNAs) can function as miRNAs Doench et al. (2003) *Genes Dev*

- For smart RNAs, we use perfect seed (nts 2-8) complementarity (4^7 = 16,384 sequences) on multiple predetermined targets
- It is difficult to find perfect seed matching on more than three predetermined targets
- We optimize the complementarity of modules ABC (4^9 = 262,144 sequences)
- It is yet to be seen how many genes can be targeted efficiently by a single smart RNA
- Breaking the complementarity in any of modules ABC is deadly





Yan et al. (in prep.)

MiRDesign summary

- We developed a strategy to design cell-specific smartRNA-based therapeutics
- It uses miRBooking to predict the microtargetome of the specific cells
- It uses miScores to compute the smartRNA sequence beyond-the-seed

Toward the promise of microRNAs
Enhancing reproducibility and rigor in microRNA research

- **Areas of microRNA confusion and pitfalls to avoid:**
  - microRNAs are expressed by cells, not tissues
  - microRNA expression levels will vary in tissue when the cellular composition of the tissue has changed in disease or malignancy
  - For the measurement of microRNAs, methodological details matter
  - microRNA-seq normalization
  - microRNA-seq alignment
  - microRNA-seq library preparation
  - Extremely large microRNA fold changes can suggest poor normalization, not interesting biology
  - microRNAs must be present to work: The importance of minimal thresholds of expression
  - Too much of a good thing: Supraphysiologic microRNA overexpression can cloud interpretation
  - The tyranny of numbers: Factors beyond microRNA concentration also affect gene regulation
  - Target levels
  - Interactions
  - Subcellular location
  - Active vs inactive microRNA complexes
  - Public datasets and methods – Use! (but with caution)

Witwer & Halushka (2016) *RNA Biology*

Acknowledgments

Hydro Québec

No carbon dioxide was emitted during the calculations

Major Lab
Mathieu Dupont (CS)
Roqaya Imane (MB)
Olivier Mailhot (BINF)
Guillaume Poirier-Morency (CS)
Jordan Quenneville (MB)

Past members
Véronique Lisi (MB)
Julie Pelloux (MB)
Julie Robitaille (BCH)
Zohra Saci (BINF)
Nicolas Scott (CS)
Nathanaël Weill (BINF)
Yifei Yan (BCH)

Collaborators
Gerardo Ferbeyre, BCH
Etienne Gagnon, IRIC
Sylvie Mader, IRIC

CIHR IRSC Canadian Institutes of Health Research

NSERC CRSNG

National Institutes of Health

Canada Foundation for Innovation
Fondation canadienne pour l'innovation

The Cancer Research Society

GenomeCanada