

# Implications of RNA Quality on miRNA Microarray Data

# *Overview of the Microarray Study*

Determine the effects of RNA degradation on the data generated by :

Genisphere HSR labeling

Affymetrix miRNA Galaxy array

- At what point does good data become bad?
- How should we handle degraded RNA samples?
- How does normalization affect this type of data?
- How do 9 miRNA's selected for validation compare between real time qPCR and Microarray?

# *The miRNA Genechip*

**The array contains 46,228 probes comprising 7,815 probe sets**

- Coverage of 71 organisms
- 11um features
- 4 Probes/miRNA (not 25bp)
- up to 11 probes for Sno/ScaRNA



**Content is derived from the**

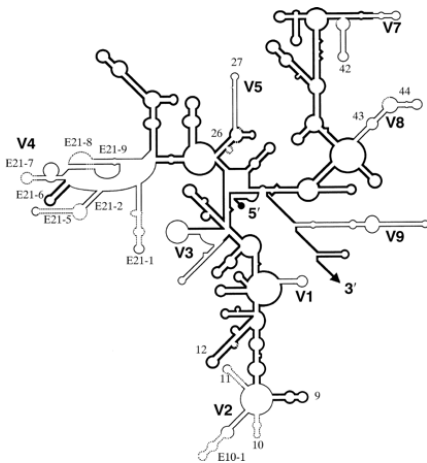
- Sanger miRBase miRNA database v11
- snoRNABase
- Ensembl

**Includes several types of ncRNAs**

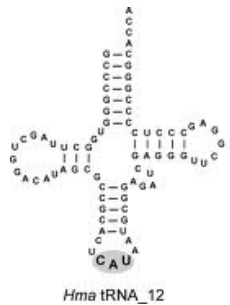
- miRNA
- short rRNA
- SnoRNA
  - C/D box
  - H/ACA Box
  - ScaRNA

# Review of RNA

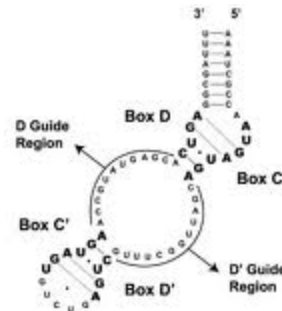
➤ <b>rRNA (LSU, SSU, 5.8s, 5S)</b>	<b>80–85 %</b>	<b>~1.5-5 kb</b>	
➤ <b>mRNAs</b>		<b>1–5 %</b>	<b>~0.5-9 kb</b>
➤ <b>tRNAs</b>		<b>15–20 %</b>	<b>~73-95 b</b>
➤ <b>ncRNAs</b>		<b>0-2 %</b>	
• miRNA, siRNA, piRNA		~19-25 bp	Regulation
• H/ACA box snoRNA		~120-200 (140)	2'-OH ribose Methylation of pre-rRNA
• C/D box snoRNAs	~70-150 (75)		Pseudouridylation of pre-rRNA
• ScaRNA snoRNA		~100-500 (207)	Cajal localized SnoRNA (H/ACA +/- C/D)
• SnoRNA (ENSG)		~60-225 (125)	Ensemble predicted



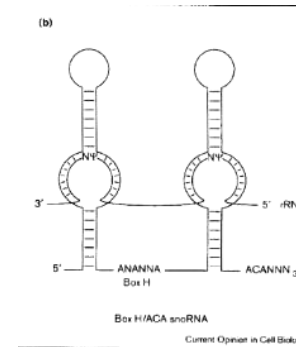
**18s rRNA**



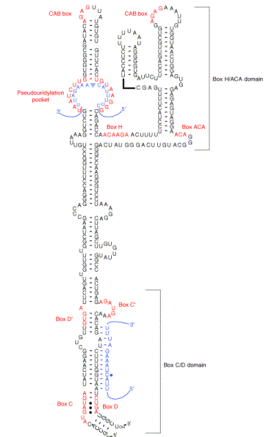
**tRNA**



**c/d Box Sno**



**H/ACA Box Sno**



**ScaRNA**



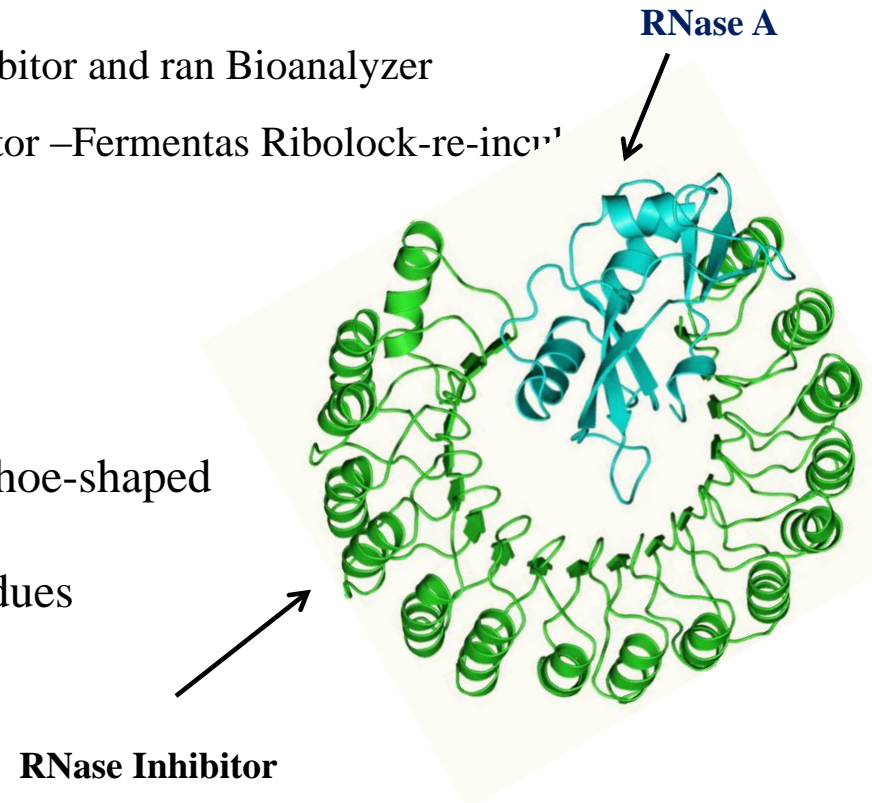
## RNase A

**Endonuclease**-hydrolyzes RNA at C and U residues within the RNA molecule -ssRNA

- Used at 0.03U/ul
- Added a total of 0.3U/50ug of HBR RNA
- Incubated at 37C for 1 hr to 6 days
- Withdrew aliquot added RNase Inhibitor and ran Bioanalyzer
- Stopped reaction with RNase inhibitor –Fermentas Ribolock-re-incu'
- Took many trial runs!!!

## RNase Inhibitor

- Ribonuclease inhibitor is horseshoe-shaped
- RNase Trapped by arginine residues

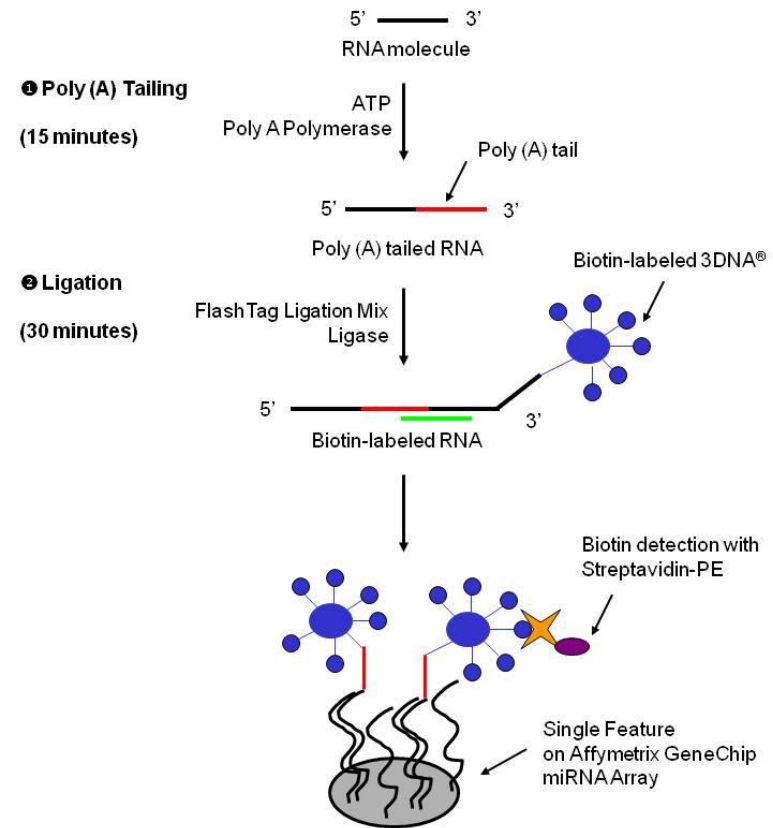


# Target Preparation

## Microarray

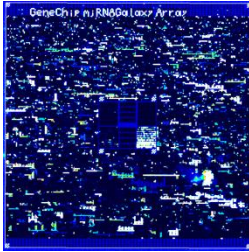
- Four labs, 3 different RINS each (N=4)
- All labs used Standardized Input Volume eqv to 500ng- Ave of five Nanodrop readings and Qubit
- Standard operating procedures with the Genisphere FlashTag HSR kit with 500 ng input
  - Poly A polymerase method
  - Labels all degraded “smaller” fragments
- Scanned with 2x Gain

### FlashTagHSR: Procedure Overview

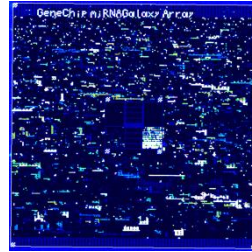


# Raw Data : Review of Genechip Image

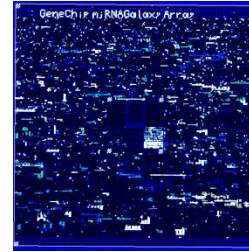
RIN 8



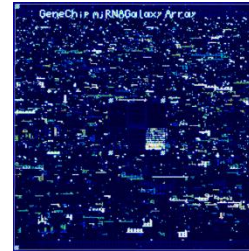
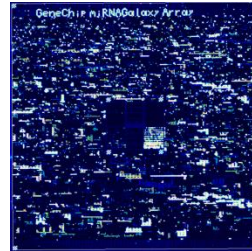
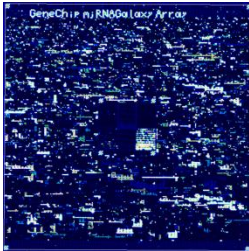
RIN 4



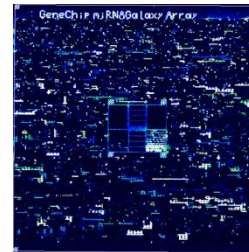
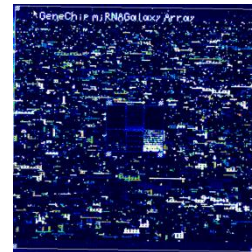
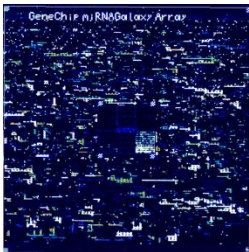
RIN 2



500 ng

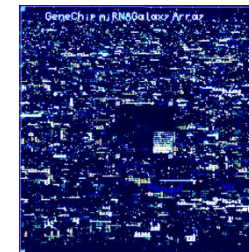
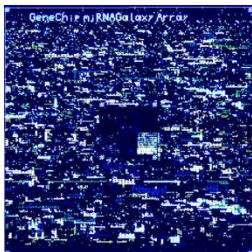
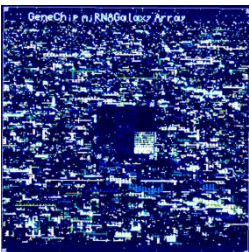


500 ng



500 ng

The loss of signal is immediately noticeable in the chip image

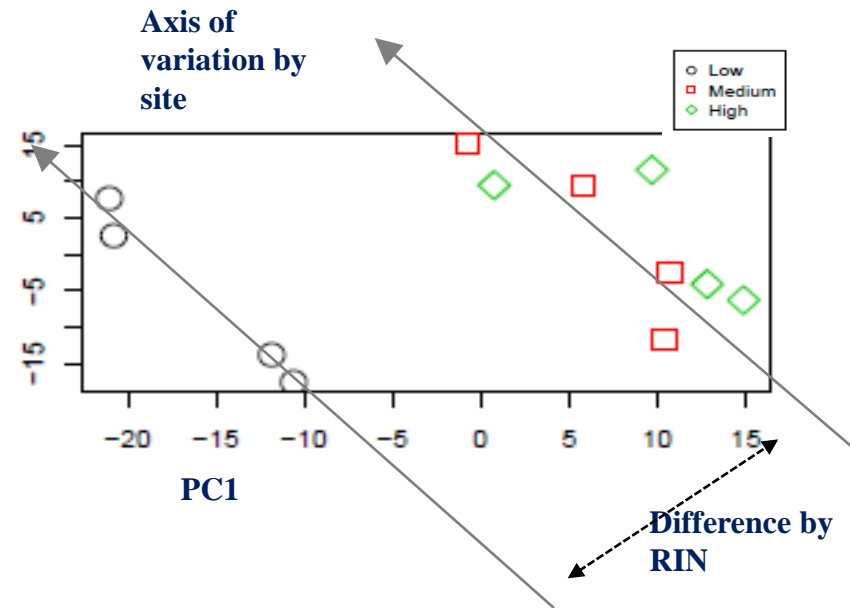


1000 ng (as a reference)

# QC: PC Plots and Pearson's Correlation

- Both plots suggest large differences between RIN 2 and 8 and much less with RIN 4 vs 8
  - Variation by site
  - Difference by RIN is significant for RIN 2 vs 8
  - Clear differences are observed for both site and RIN value

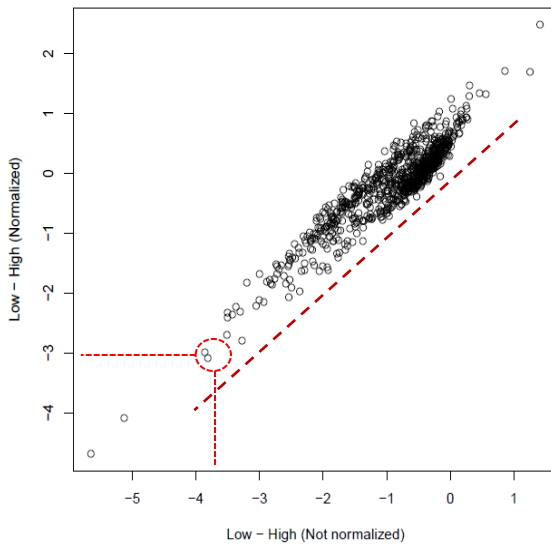
	N1-RIN2	N1-RIN4	N1-RIN8	A1-RIN2	A1-RIN4	A1-RIN8	P1-RIN2	P1-RIN4	P1-RIN8	V1-RIN2	V1-RIN4	V1-RIN8
N1-RIN2	1.00	0.95	0.91	0.92	0.88	0.85	0.97	0.86	0.89	0.91	0.88	0.87
N1-RIN4	0.95	1.00	0.98	0.86	0.95	0.95	0.92	0.96	0.97	0.86	0.89	0.94
N1-RIN8	0.91			A1-RIN2	A1-RIN4	A1-RIN8	0.96	0.81	0.84	0.91		
A1-RIN2	0.92	A1-RIN2	1	0.87	0.81	0.82	0.99	0.92	0.87			
A1-RIN4	0.88	A1-RIN4	0.87	1	0.99	0.98	0.82	0.89	0.96			
A1-RIN7	0.85	A1-RIN8	0.81	0.99	1	0.90	0.99	1.00	0.84	0.90	0.96	
P1-RIN2	0.97									0.97	0.92	0.91
P1-RIN4	0.86									0.82	0.88	0.96
P1-RIN8	0.89	0.97	0.96	0.82	0.97	0.98	0.90	0.99	1.00	0.84	0.90	0.96
V1-RIN2	0.91	0.86	0.81	0.99	0.88	0.82	0.97	0.82	0.84	1.00	0.95	0.89
V1-RIN4	0.88	0.89	0.84	0.92	0.92	0.89	0.92	0.88	0.90	0.95	1.00	0.95
V1-RIN7	0.87	0.94	0.91	0.87	0.98	0.96	0.91	0.96	0.96	0.89	0.95	1.00



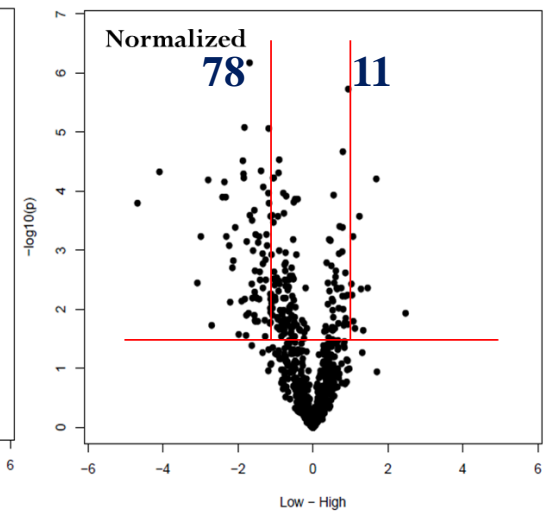
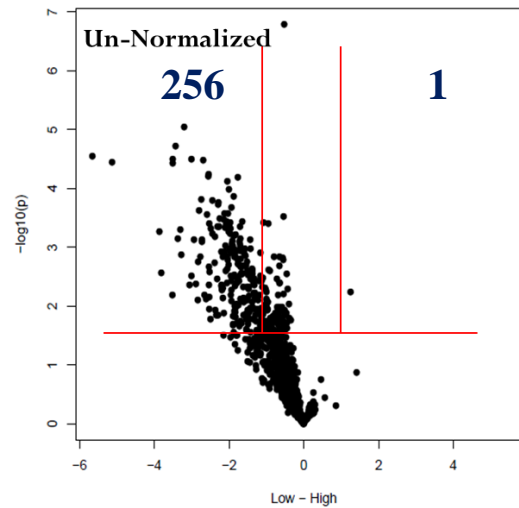
# Data Normalization

## What should you normalize to in a degradation experiment?

- Standard normalization leads to possible false increases of signal for large  $\Delta$  RINs
- Normalizing to the Spike-in oligos may be more appropriate for large  $\Delta$  RINs
- A scaling would result in a vertical shift as shown to the left.
- If there is differential detection, then normalization must effect LFC



**DOD Plot**



**Volcano Plots**

## *Fold Change and Detection Data*

- RIN 2 RNA has noticeable loss of signal for ncRNA as compared to the RIN 8
- False positives detection was negligible, unless normalization was applied

Comparison	Fold Change	Un-Normalized	Normalized
RIN 2 vs 8	2 Up	1	11
	2 Down	256	78
RIN 4 vs 8	2 Up	0	0
	2 Down	4	0

FDR < 0.1

- Degradation had a greater effect on miRNA vs SnoRNA
  - Is differential detection length or secondary structure related?

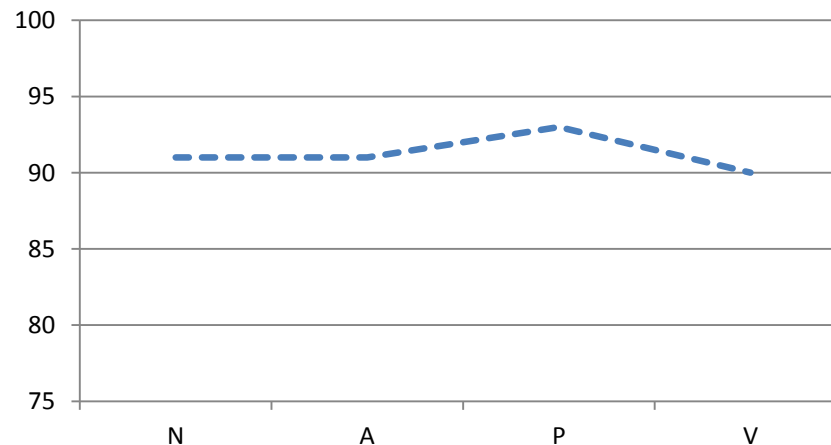
Type	Percentage differentially quantified
All	70
Human miRNA	80
Human SnoRNA	35

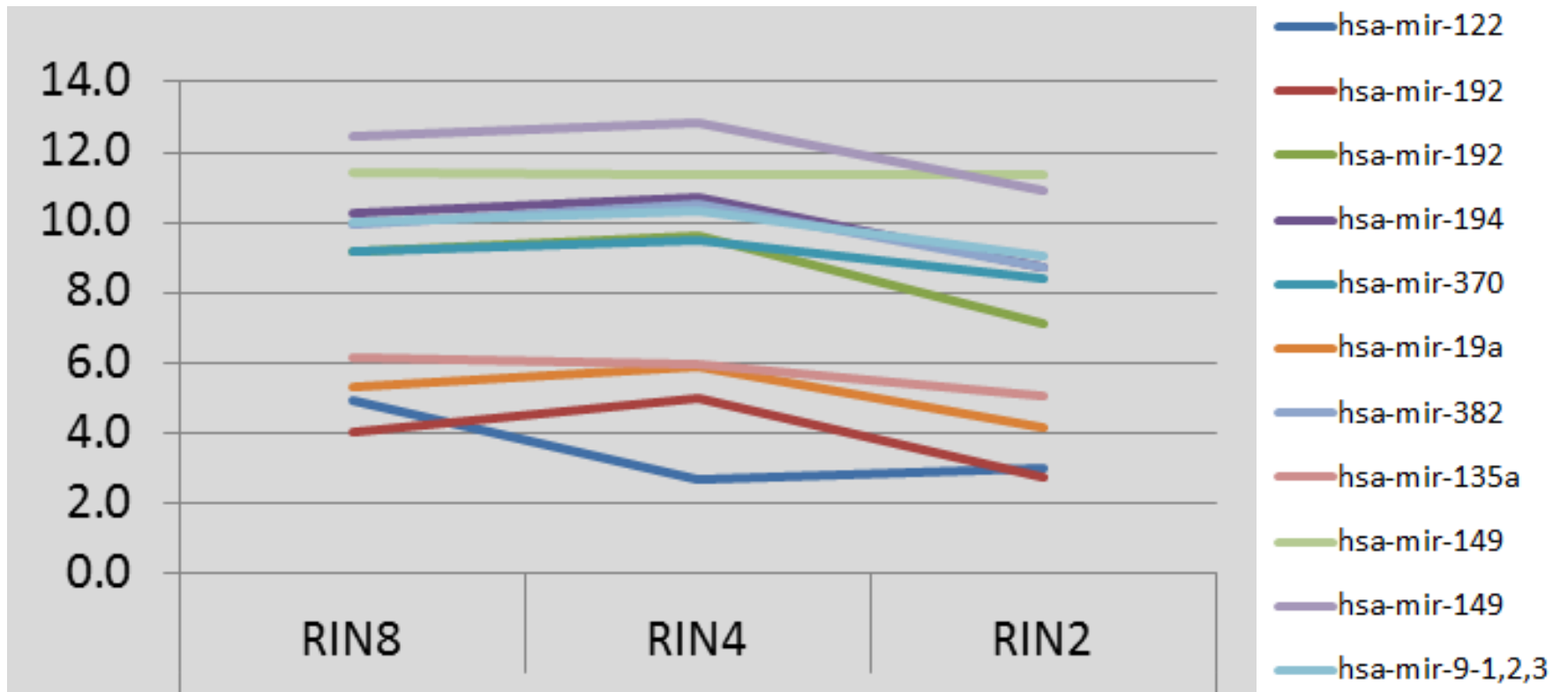
## *Fold Change and Detection Data*

- Percent Agreement for RIN8 RNA across sites (Phi coefficients  $\geq 0.80$ )

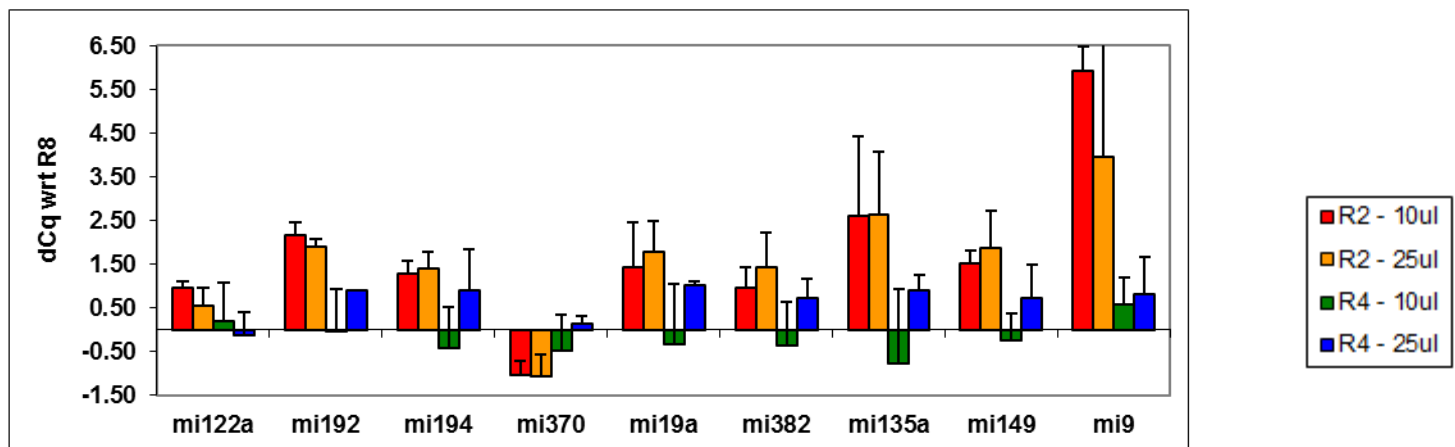
	N	A	P	V
N	100			
A	92	100		
P	93	92	100	
V	90	94	90	100

- Agreement on differential ncRNA detection between RIN8 and RIN2 across sites



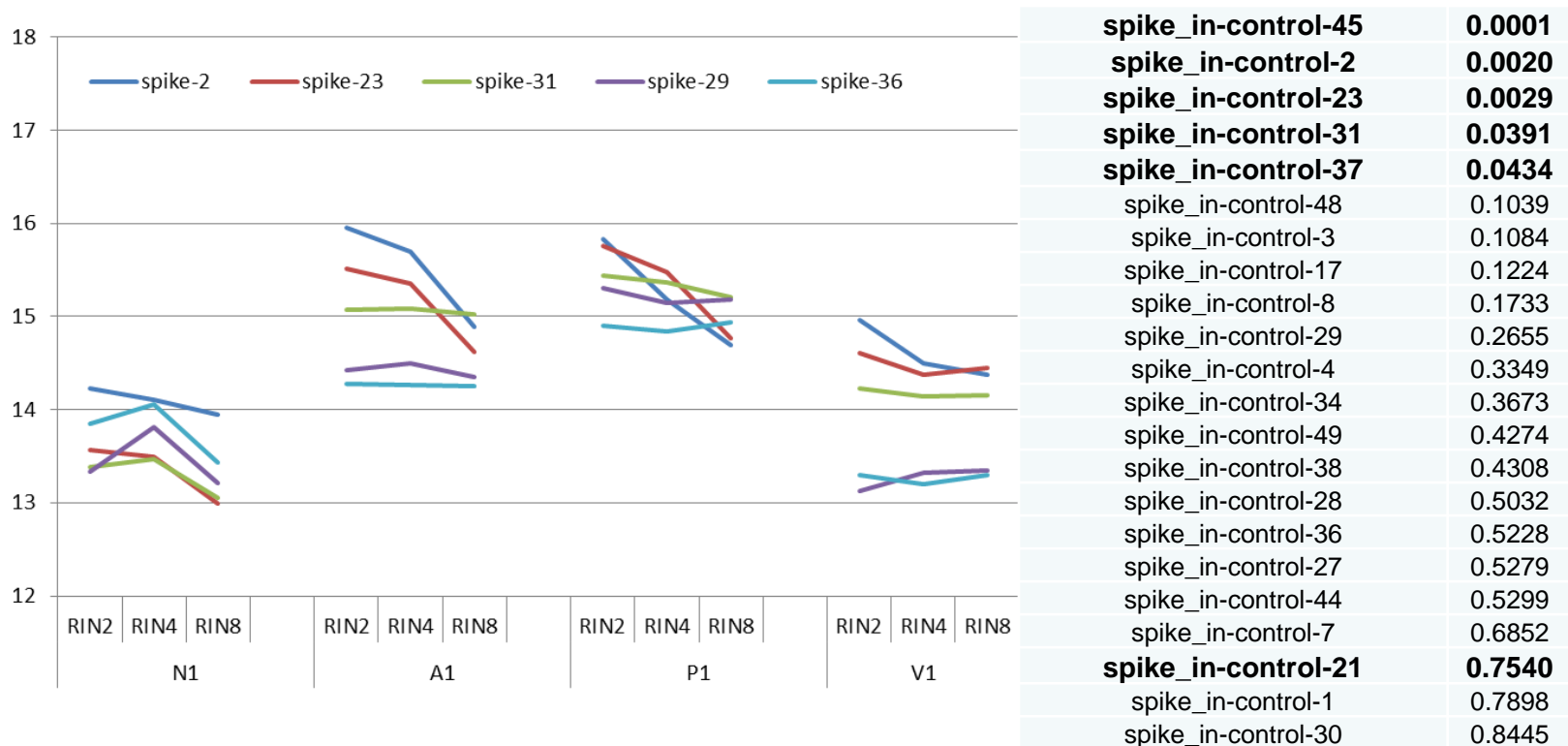


### Rt-qPCR Data



# Spike in Oligos Controls

- Maybe a good alternative as a normalization method
- Not enough data to make a call



Thank what might be happening

**miRNA are degraded by RNase A**

**Interference by high concentrations of degraded fragments**

**Digested RNA fragments are excessive and maybe starving the reagents**

# Summary and Conclusions

- **Detection of ncRNA using the Genisphere HSR labeling and Affymetrix miRNA chip was robust with respect to variation in RNA integrity, however RIN2 RNA demonstrated an expected signal loss.**
- **RIN 8 and 4 have very similar miRNA detection profiles and may be used little concern. RIN matching is always optimal**
- **Normalization methods need to be considered for mixed RIN comparisons situations like FFPE**
  - Quantile**
  - None**
  - Spike-in's**
- **Microarray data mimics RT-qPCR validation data**

# Acknowledgements

## ➤ **ABRF Nucleic Acids Research Group**

Russ Carmical	UTMB
Sridar Chittur	Ualbany (Co-chair)
Jennifer Holbrok	Nemours
Vijay Nadella	Ohio Univ
Scott Tighe	UVM (Co-chair)



## ➤ **Labs that did the microarrays (Donated GeneChips, Reagents, and Time)**

Herbert Auer	IRB Barcelona
Don Baldwin	UPenn
Sridar Chittur	UAlbany
Nalini Raghavachari	NIH
Scott Tighe	UVM

## ➤ **Bioinformatics**

Jeff Bond	UVM
Julie Dragon	UVM

## ➤ **Vendor support for qPCR studies (reagents)**

Applied Biosystems (Lifetech), SABiosciences (Qiagen)

## ➤ **Vermont Genetics Network-** Partial funding for RNA