

Breast Cancer

The most global cancer incidence in women

Rank	Cancer	New cases diagnosed in 2018	% of all cancers (excl. non-melanoma skin cancer)
1	Breast	2,088,849	25.4
2	Colorectal	794,958	9.7
3	Lung	725,352	8.8
4	Cervix uteri	569,847	6.9



- Breast cancer causes **the greatest number of cancer-related deaths among women**. In 2018, it is estimated that 627,000 women died from breast cancer – that is approximately 15% of all cancer deaths among women.
- While breast cancer rates are higher among women in more developed regions, **rates are increasing in nearly every region globally**.



Diversity of Breast Cancer: Subtypes

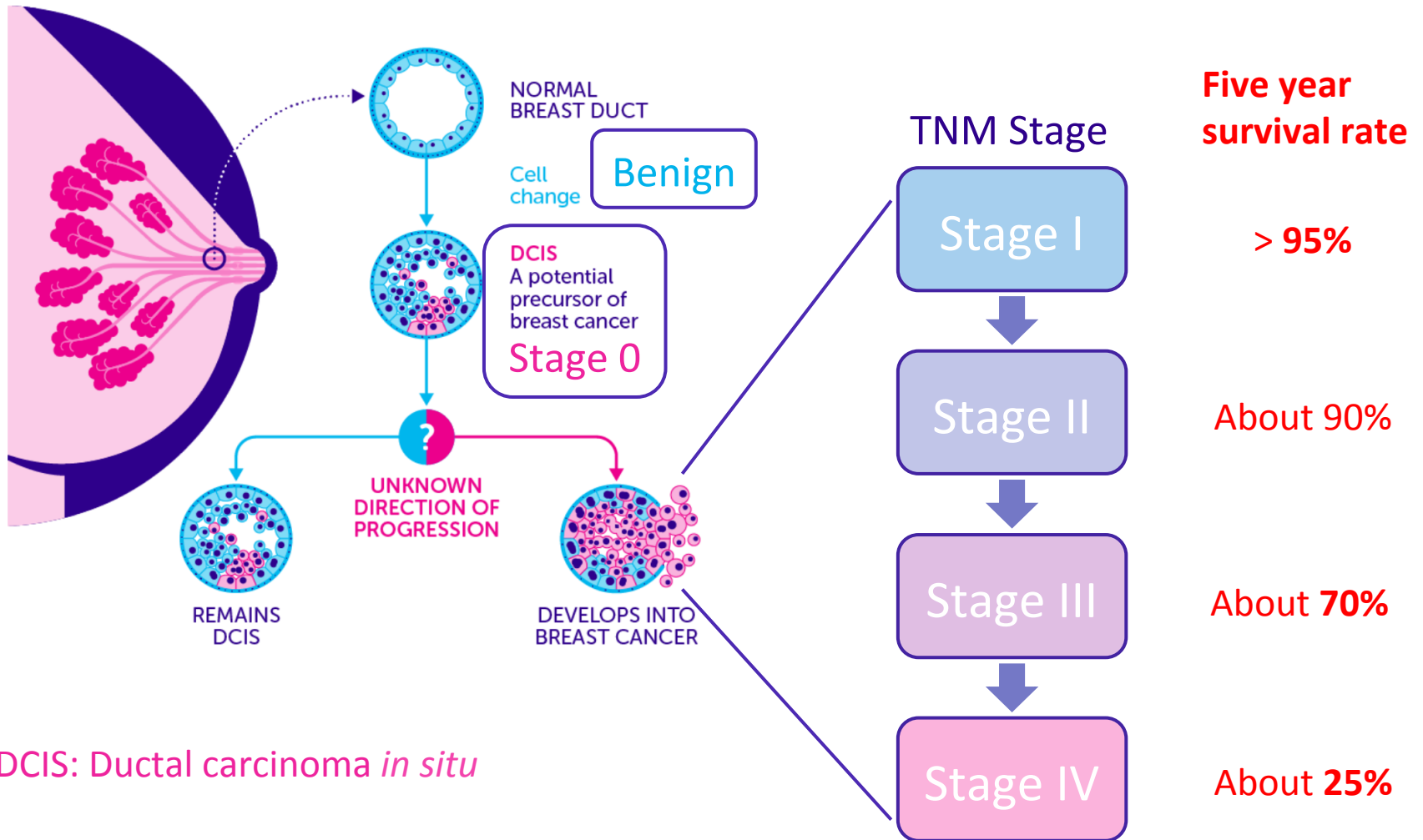
	Luminal A	Luminal B	HER2 Positive	Triple Negative
Percentage at diagnosis	40%	20%	10-15%	15-20%
Receptor expression	Estrogens and progesterone	Human epidermal growth factor 2		
Treatment strategies	Chemotherapy			
		Anti-HER2 therapies		
	Hormonal therapies			
	Novel target therapies			

Catalanotti, V., Bertaglia, V., Tariq, N., Califano, R. (2014). Treatment of Advanced Breast Cancer (ABC): The Expanding Landscape of Targeted Therapies. *J Cancer Biol Res*, 2(1), 1036.

Clinical unmet needs in breast cancer

- 1. Early detection**
- 2. New progression mechanism for targeted therapies (e.g. TNBC)**
- 3. Heterogeneity**
- 4. Drug Resistance**
- 5. Companion diagnosis**
- 6. Racial disparity**

Progression & Diversity of Breast Cancer



DCIS: Ductal carcinoma *in situ*

TNM Stage

T = size of primary tumor

N = the extent of spread to nearby lymph nodes

M = presence or absence of distant metastases

Diversity of Breast Cancer: Racial Disparity

Figure 6a. Trends in Female Breast Cancer Incidence Rates by Race/Ethnicity, 1975-2014, US

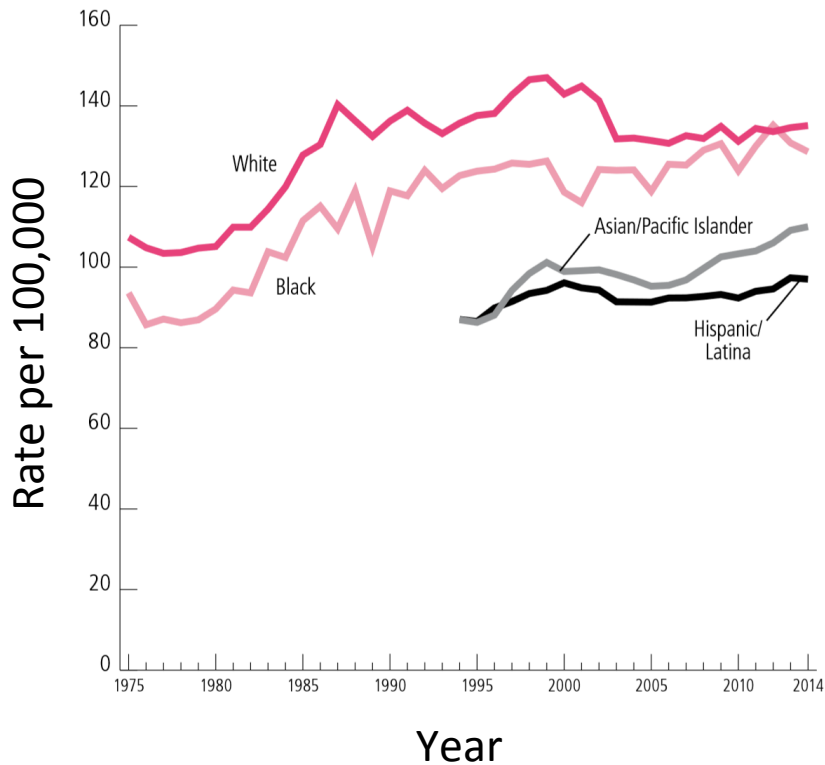
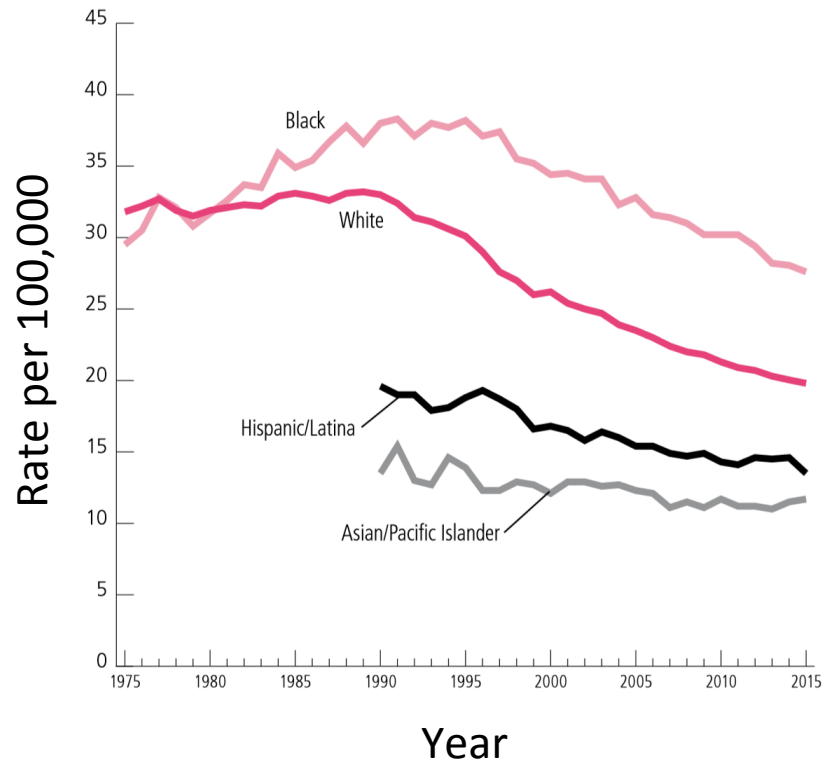
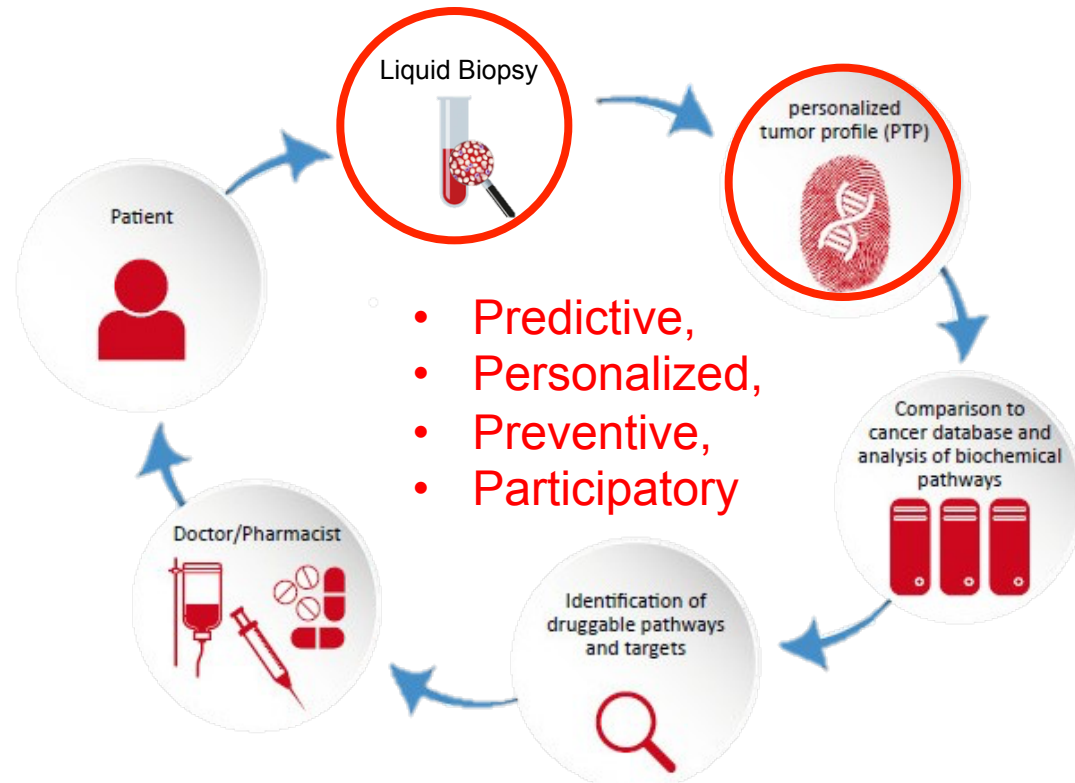


Figure 6b. Trends in Female Breast Cancer Death Rates by Race/Ethnicity, 1975-2015, US



Disease Biomarkers

From Liquid Biopsy to Precision Medicine



“Doctors have always recognized that every patient is unique, and doctors have always tried to tailor their treatments as best they can to individuals. You can match a blood transfusion to a blood type — that was an important discovery. What if matching a cancer cure to our genetic code was just as easy, just as standard? What if figuring out the right dose of medicine was as simple as taking our temperature?”

- President Obama, January 30, 2015

OPINION

nature
REVIEWS **CLINICAL ONCOLOGY**

Predictive, personalized, preventive, participatory (P4) cancer medicine

Leroy Hood and Stephen H. Friend

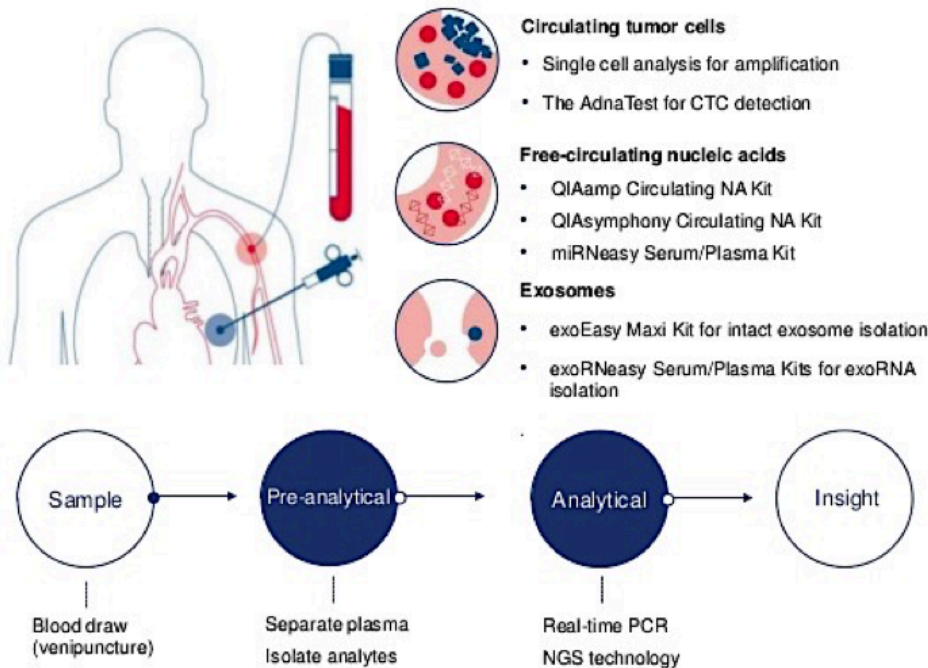
Published: 02 March 2011

Liquid Biopsy:

- Less invasive,
- less costly,
- less risky,
- contain more dynamic information than conventional tissue biopsies.

Trend in Cancer Precision Medicine: Liquid Biopsy: CTC, ctDNA, Exosome

For non-invasive biomarker discovery: Sample to Insight



Liquid Biopsy:

Variety of biomarkers existed in bodily fluids such as blood, saliva, urine, ascites, etc

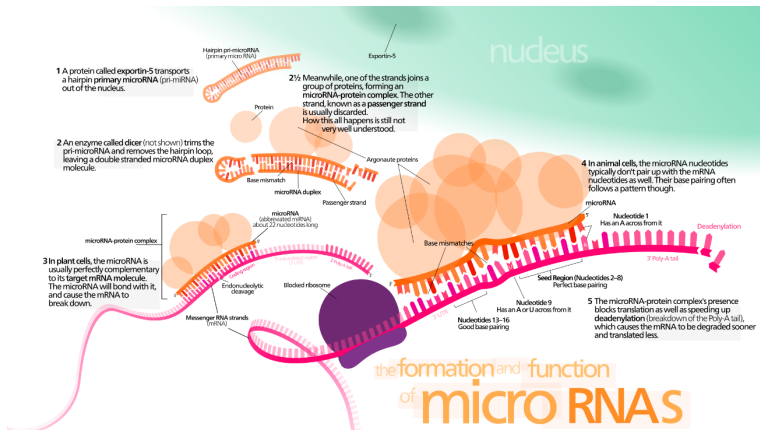
Main areas of Liquid Biopsy:

- Circulating tumor cells (CTC)
- Circulating tumor DNA (ctDNA)
- Exosomes
 - mRNA, Protein, miRNA, lncRNA

miRNA Biomarkers in Liquid Biopsy

MicroRNA(miRNA)

1. Non-coding RNA
2. 17-25 nucleotides
3. Post-transcriptional regulation: base-pair with mRNAs and silence those mRNAs
4. Appear to target about 60% of the genes of human



Evaluation of serum microRNA biomarkers for gastric cancer based on blood and tissue pools profiling: the importance of *miR-21* and *miR-331*

British Journal of Cancer (2017) 117, 266-273. doi: 10.1038/bjc.2017.190

Marek Sierzega^{*1}, Marcin Kaczor², Piotr Kolodziejczyk¹, Jan Kulig¹, Marek Sanak² and Piotr Richter¹

miR-331 and miR-21 => gastric cancer

Serum MicroRNA profile in patients with colon adenomas or cancer

BMC Medical Genomics (2017) 10:23
DOI 10.1186/s12920-017-0260-7

Yajie Zhang^{1,2}, Min Li⁴, Yijiang Ding³, Zhimin Fan³, Jinchun Zhang⁵, Hongying Zhang⁶, Bin Jiang^{3*} and Yong Zhu^{3*}

8 miRNAs => colon cancer

Circulating microRNAs in breast cancer: novel diagnostic and prognostic biomarkers

Cell Death & Disease **8**, e3045 (2017)

Rimi Hamam, Dana Hamam, Khalid A Alsaleh, Moustapha Kassem, Waleed Zaher, Musaad Alfayez, Abdullah Aldahmash & Nehad M Alajez ✉

Problems to be solved

1. Identify possible biomarkers for early detection of breast cancer from liquid biopsy (e.g. miRNAs).
2. Generate profiling on the various stages of breast cancer. Identify more insightful information regarding the critical factors that progress cancer to the next stage.

Comprehensive data (containing all variation)
+ High-throughput techniques (high sensitivity)
+ Proper data analysis (low sample number issue)
+ Artificial intelligence for optimization (predictable)

Current biomarkers for breast cancer

Breast Cancer: 3 tumor markers

- cancer antigen 15-3 (CA 15-3),
- cancer antigen 27.29 (CA 27.29), and
- carcinoembryonic antigen (CEA)

Tumor Marker	N (Total)	Sensitivity
CA 15-3	35 (145)	24.1%
CA 27.29	37 (145)	25.5%
CEA	27 (145)	18.6 %

Hou, MF, et al. (1999). Evaluation of serum CA27.29, CA15-3 and CEA in patients with breast cancer. *Kaohsiung J Med Sci.*, 23(1), pp.88-93.

Comprehensive data



Noncancer	Benign	10	DCIS	20
-----------	--------	----	------	----

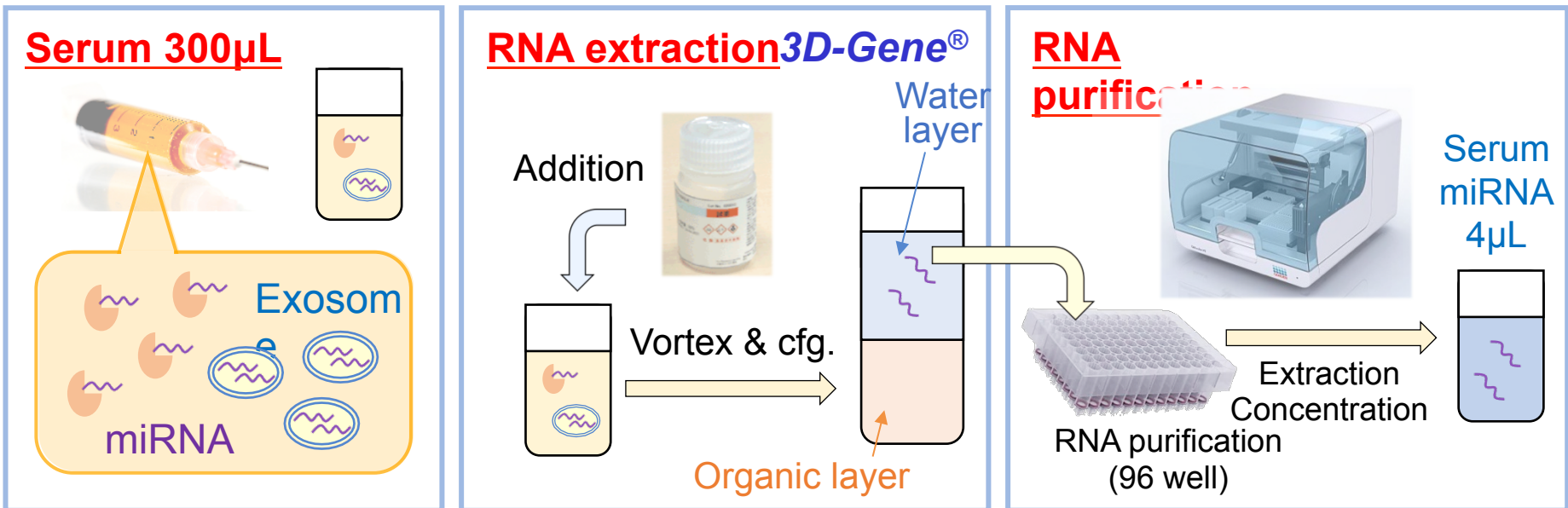
Cancer		Stage				Total
		I	II	III	IV	
Subtypes	Luminal A	8	8	12	5	33
	Luminal B	8	8	12	5	33
	HER2 ⁺	8	8	12	5	33
	Triple Negative	8	8	12	4	32
Total		32	32	48	19	131

High-throughput techniques

1. miRNA extraction from serum
2. miRNA analysis with microarray



Takahiro Ochiya, PhD

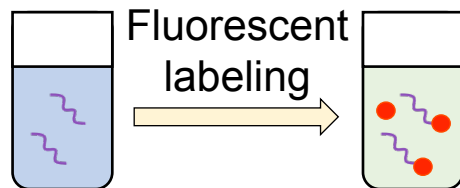


High-throughput techniques

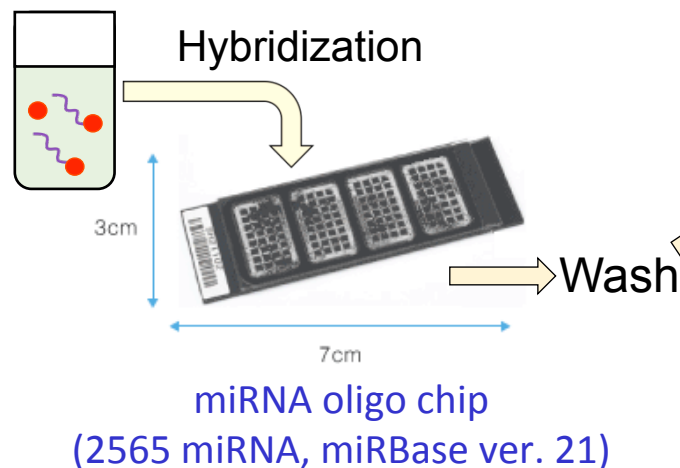
1. miRNA extraction from serum
2. miRNA analysis with microarray

RNA labeling

RNA 2 μ L

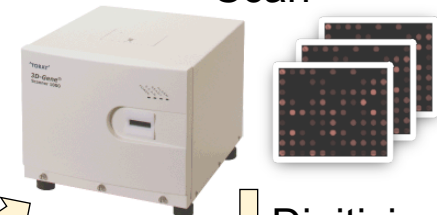


DNA chip analysis 3D-Gene[®]



Analysis

Scan

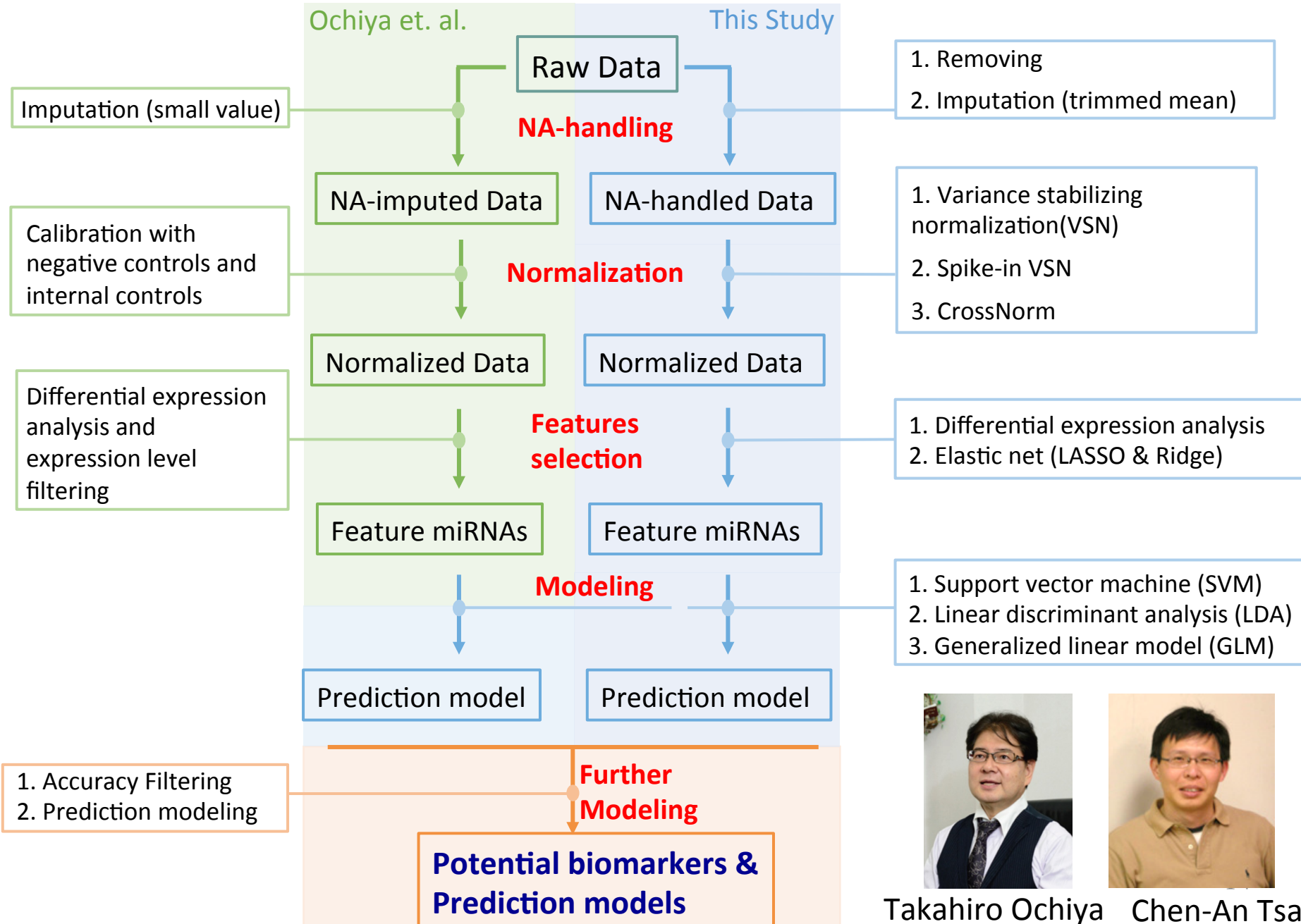


Digitizing



miRNA expression data
(Raw data)

Data Analysis Pipeline

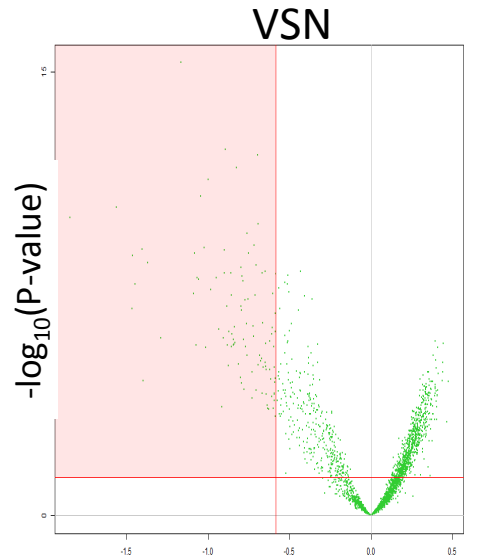


Takahiro Ochiya

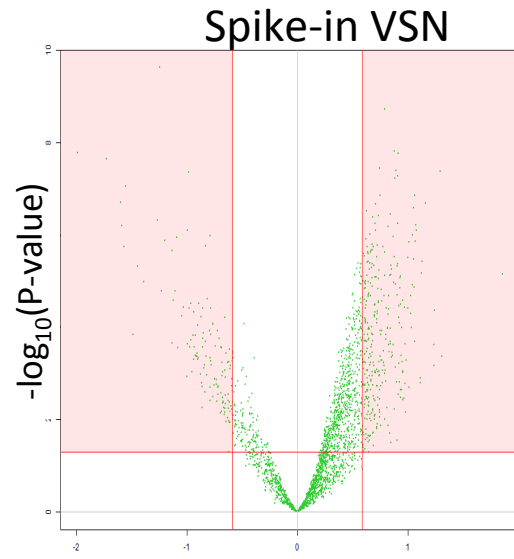


Chen-An Tsai

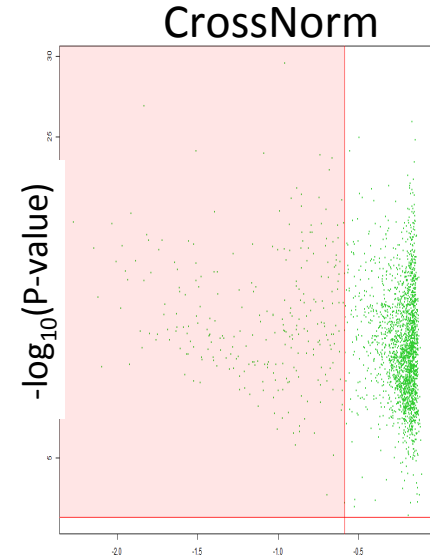
Differential Expression Analyses



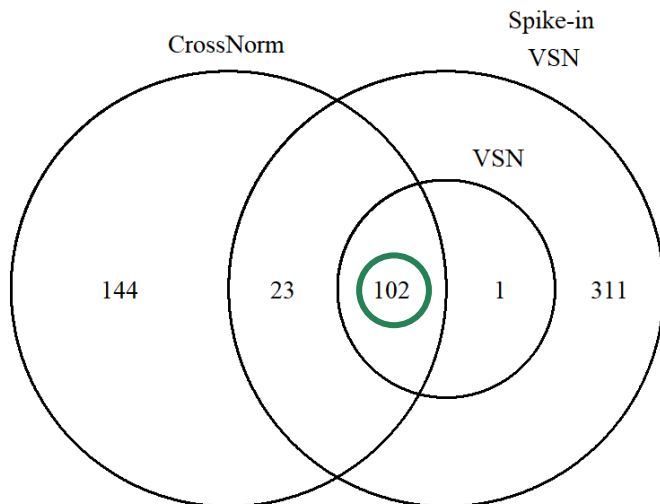
$\log_2(\text{fold change})$
103 probes are selected



$\log_2(\text{fold change})$
437 probes are selected

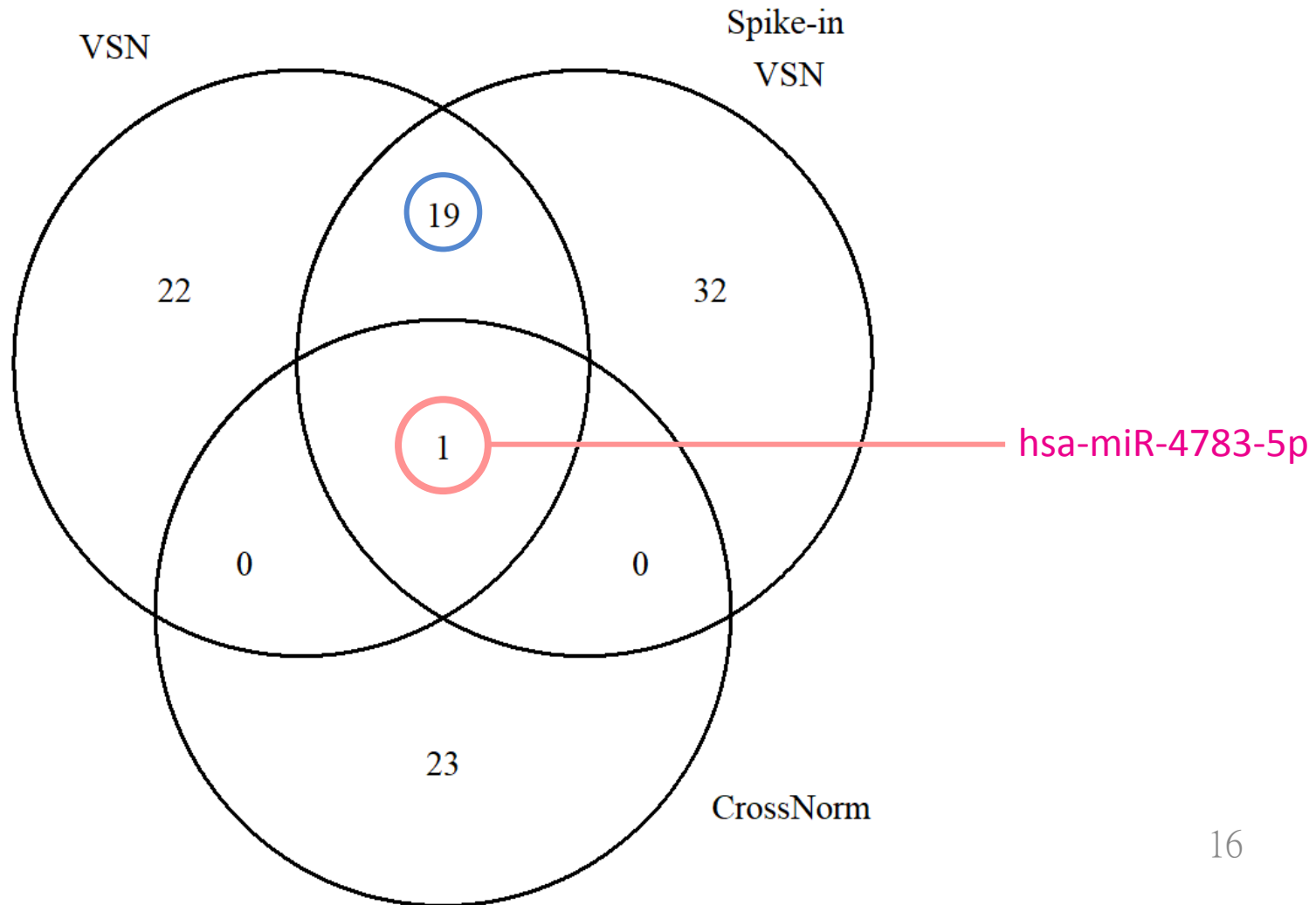


$\log_2(\text{fold change})$
269 probes are selected

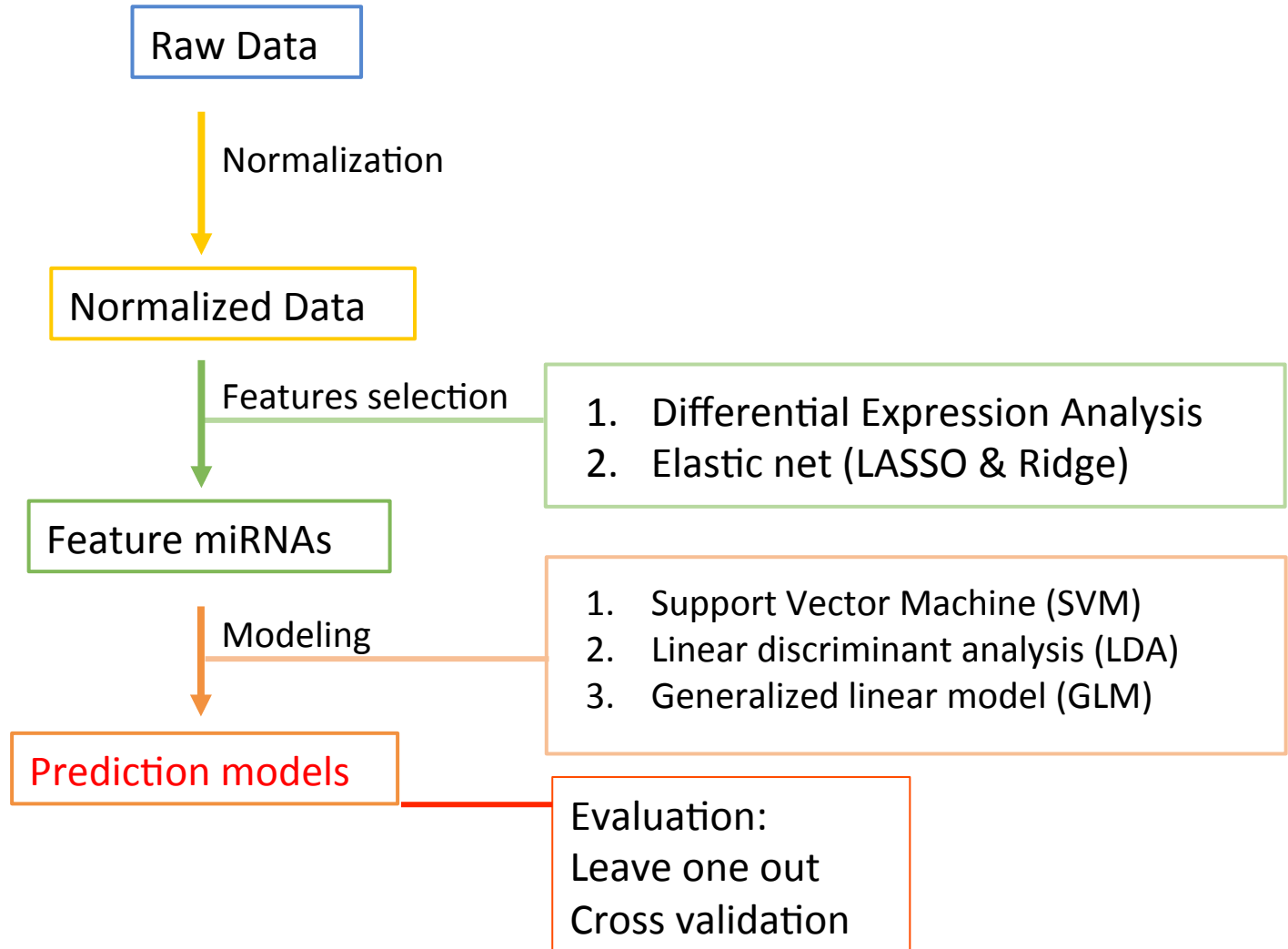


Adjusted P-values < 0.05
(Benjamini–Hochberg procedure)
Significant \log_2 fold change = 1.5

miRNAs Selected by Elastic Net Regression



miRNA Analysis Flow Chart



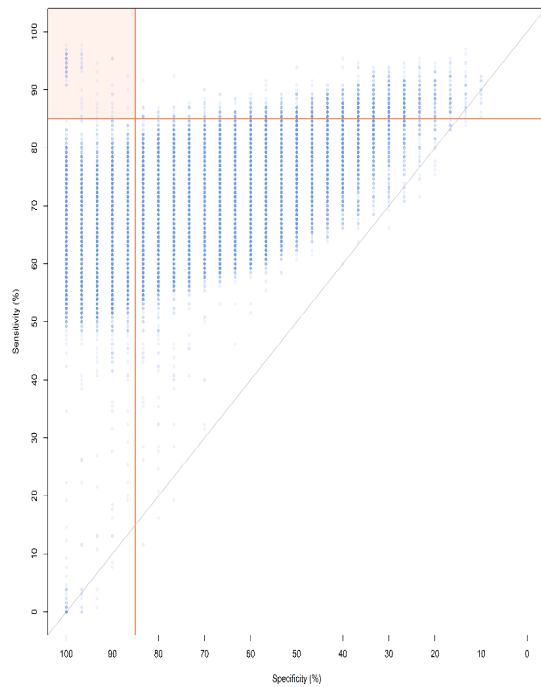
Results of miRNA through Analysis Pipeline

Whole miRNA		2565		
After removing NA		2462		
Normalization		VSN	Spike-in VSN	CrossNorm
Differential Expression		0	307	0
		103	130	269
Elastic Net		32	50	0
		31	14	22
Modeling with at most 3 selected miRNA or each single miRNAs -> Construct 17,423,153 models -> Evaluate with 10-fold cross validation				

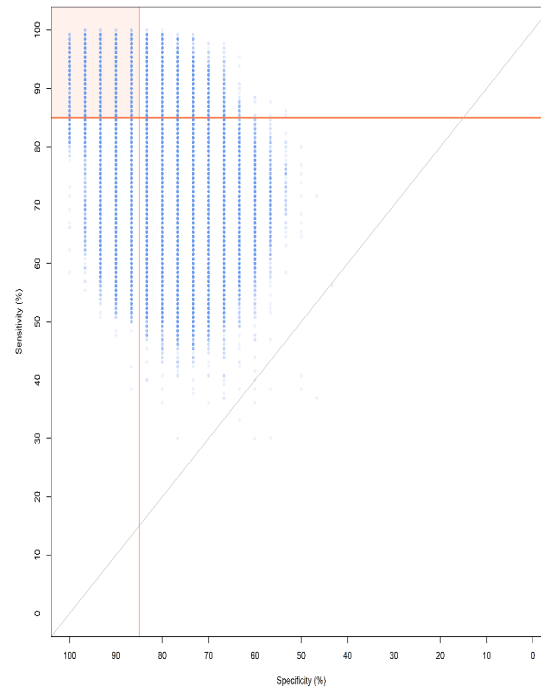
Modeling with Selected miRNAs

Selecting the miRNAs with prediction ability

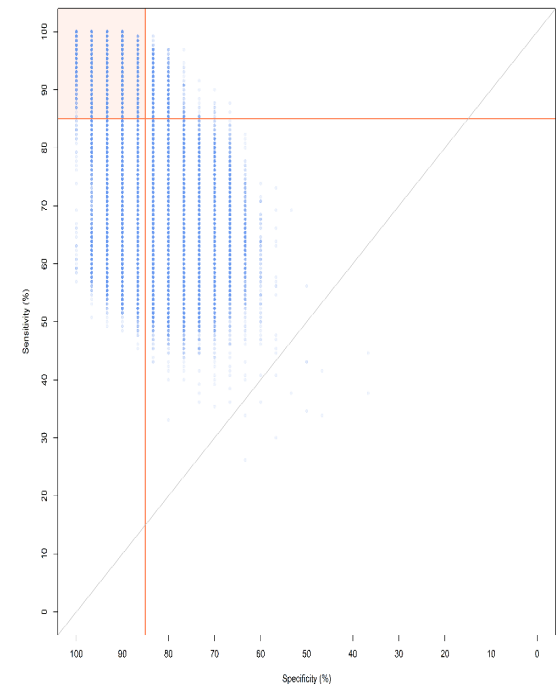
sensitivity & specificity > 85%



SVM

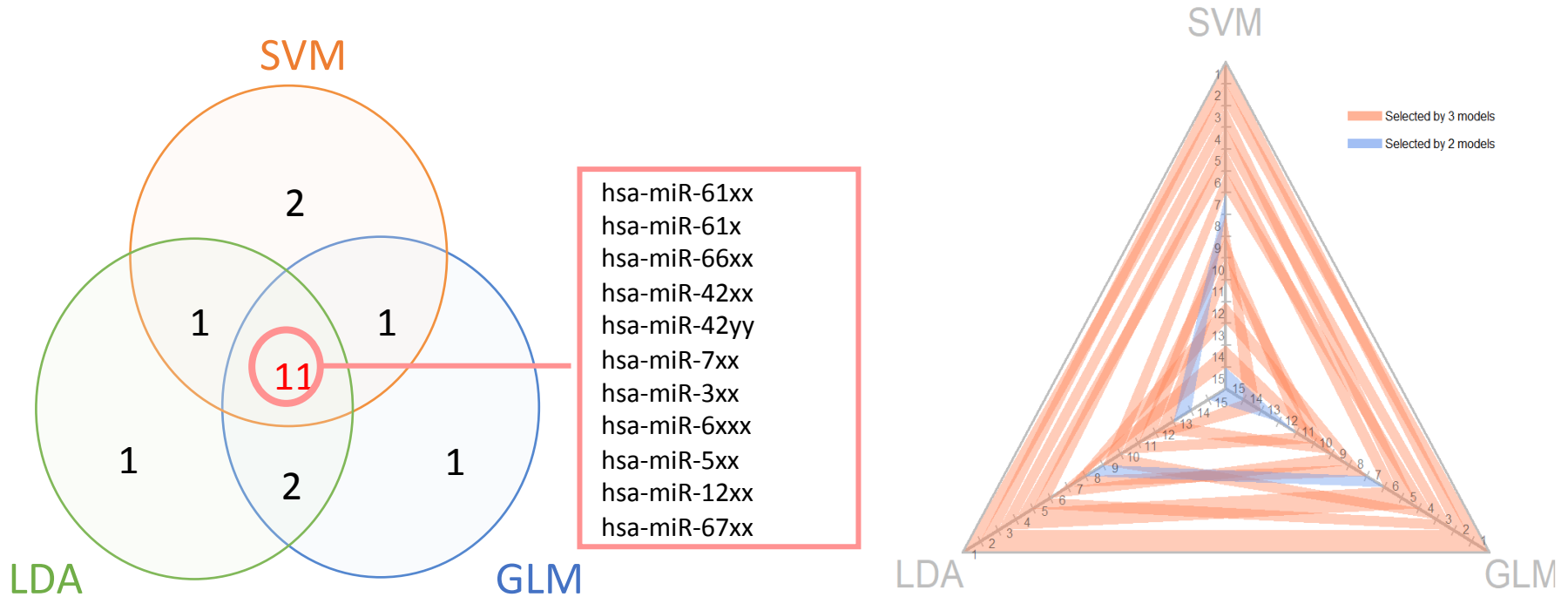


LDA



GLM

Overlap and Consistency of Each Modeling Method

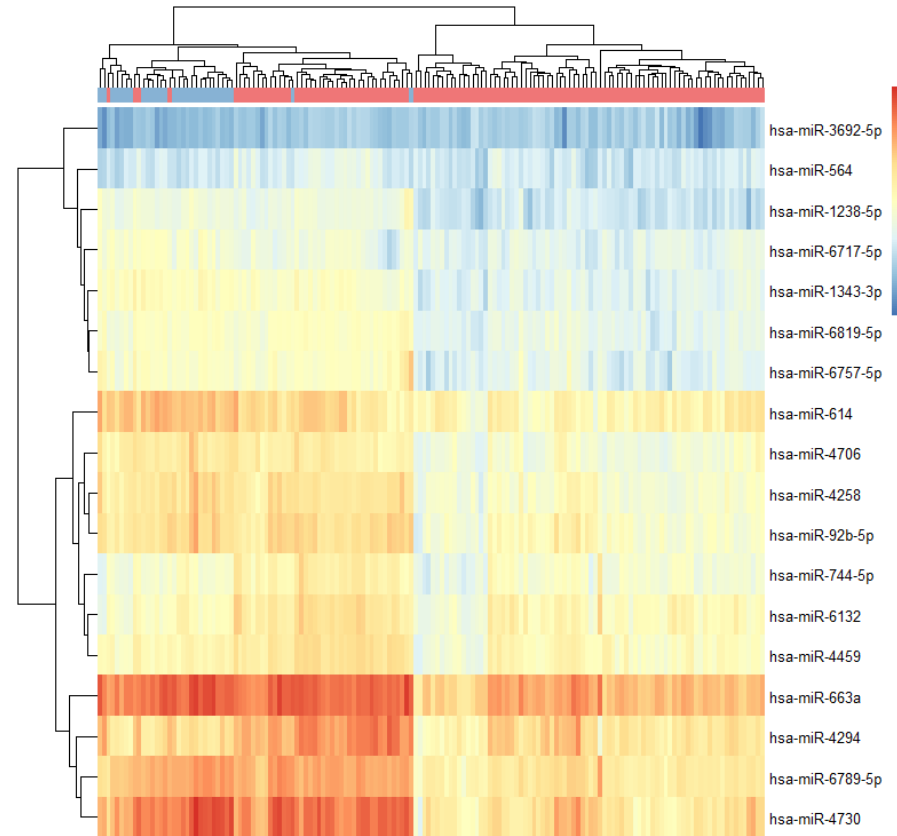
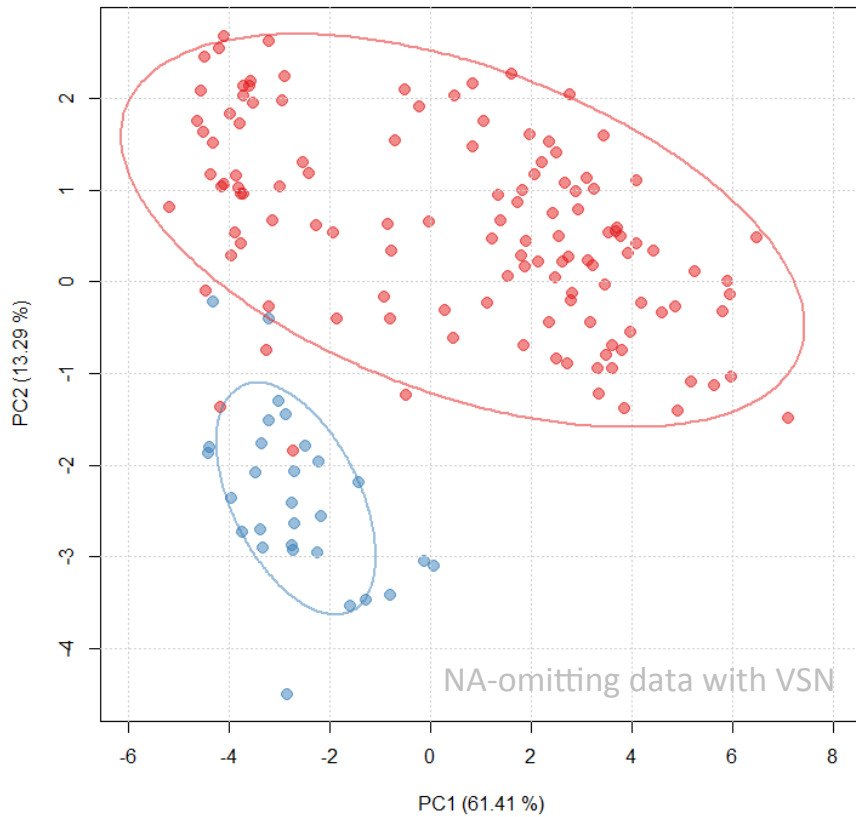


Selected miRNAs with Prediction Ability

⇒ 18 miRNAs (overlap from 4 methods)

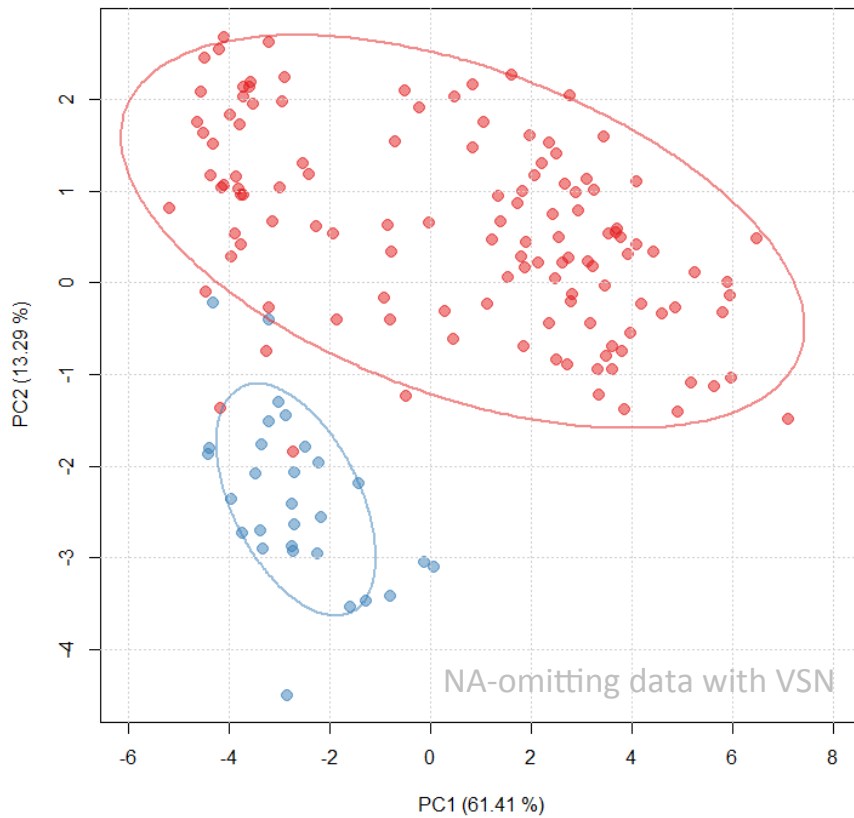
PCA and heatmap with Selected 18 miRNAs

Principal Component Analysis (Cancer v.s. Noncancer)

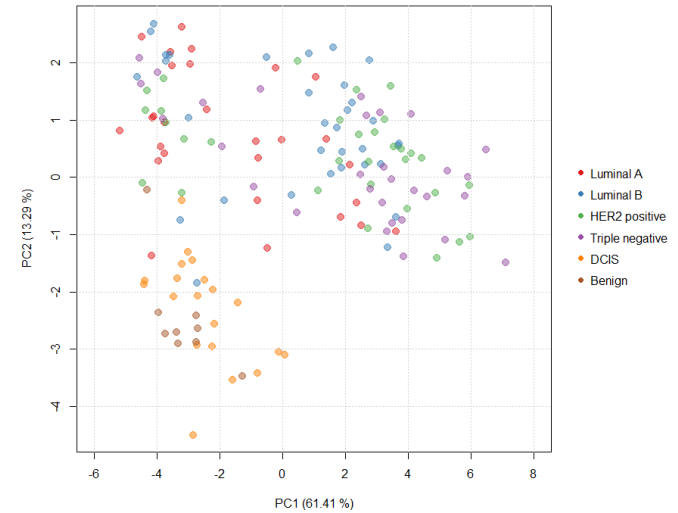


PCA plots show selected 18 miRNAs fit-in early detection

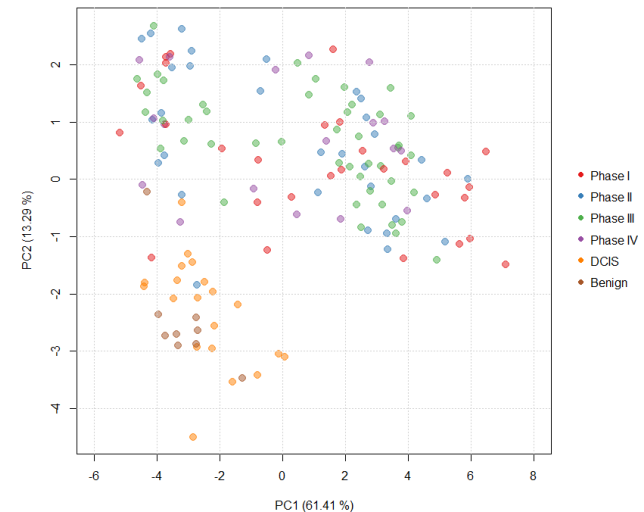
Principal Component Analysis (Cancer v.s. Noncancer)



Principal Component Analysis (Subtype)



Principal Component Analysis (Phase)



Further modeling

After previous prediction modeling, we used the union of selection from each pipeline to build more prediction models. Fisher's linear discriminant analysis (LDA) was performed with each of these miRNA marker or a combination of at most six miRNA markers.

To evaluation the prediction performance, 10-fold cross validation were applied to each model. We separated the data into 10 groups, built the model with 9 groups and used the residual group as testing cohort to calculate the prediction accuracy, sensitivity and specificity. After repeating the estimation process with different testing group 10 times, the average values of each test result were calculated for model evaluation.

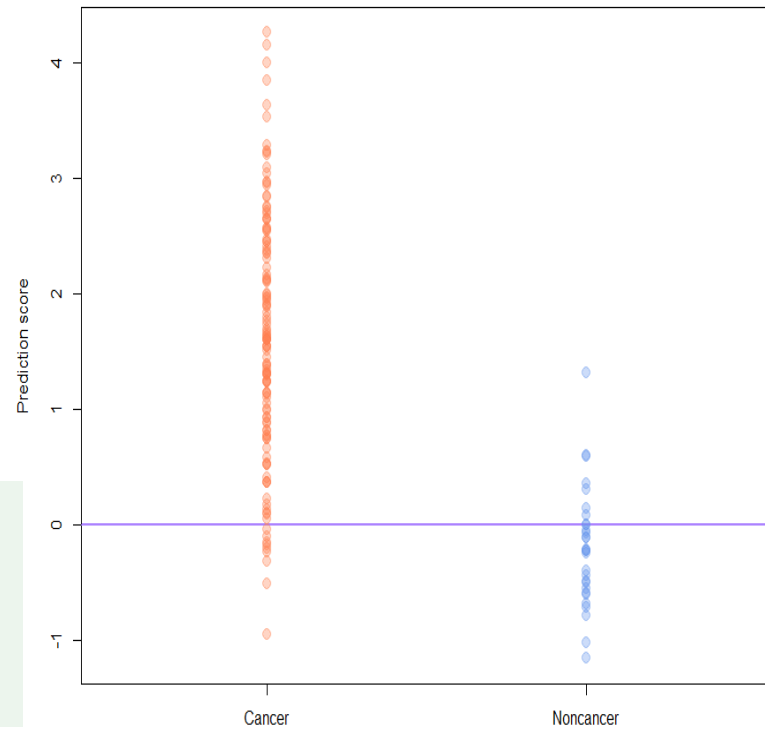
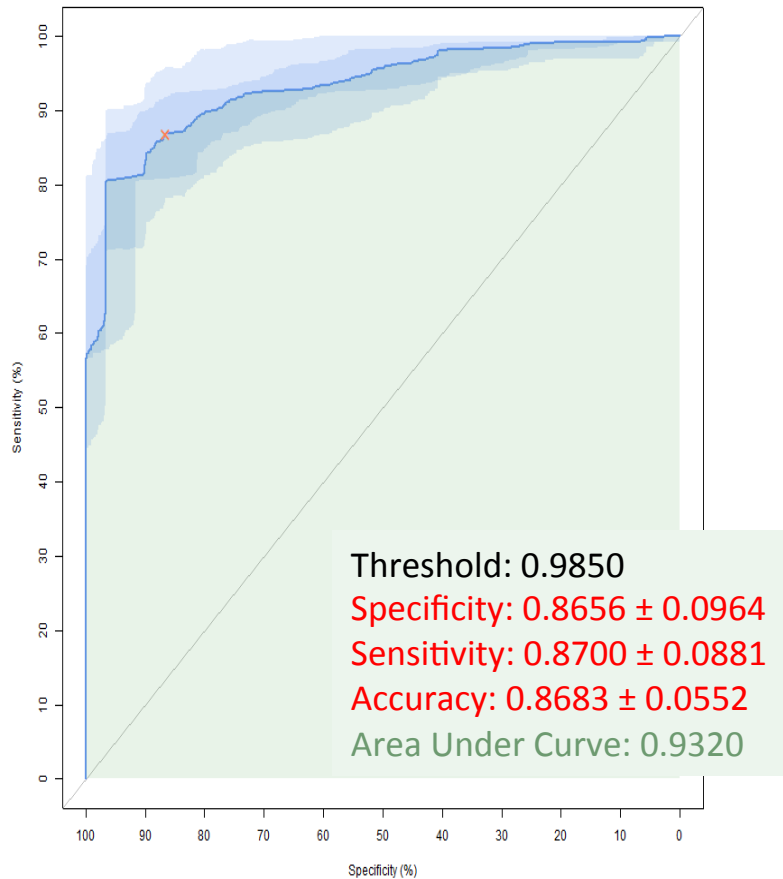
The resulting values of the discriminant functions were used to prepare the diagnostic index.

Index score ≥ 0 : breast cancer

Index score < 0 : non-breast cancer or other clinical conditions

Further Modeling (with 1 miRNA)

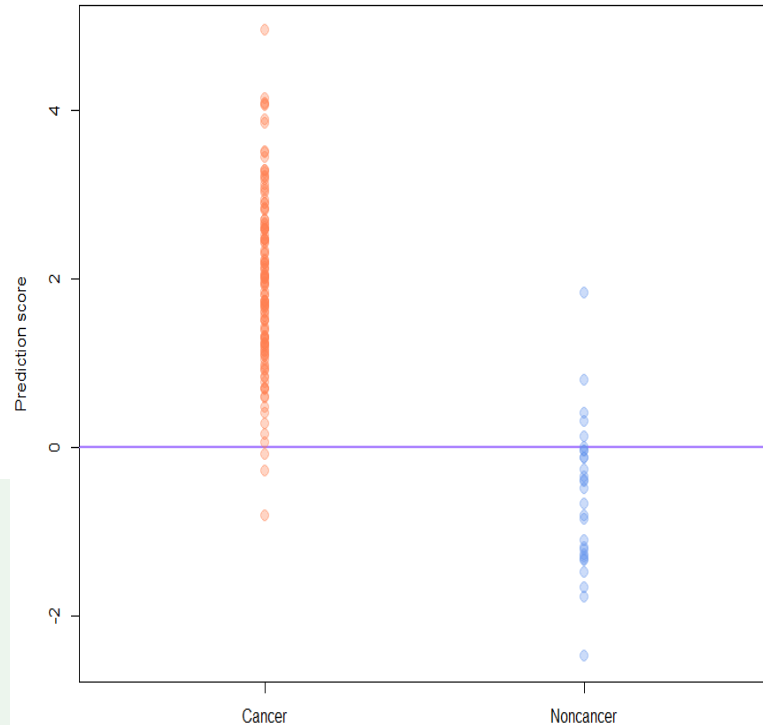
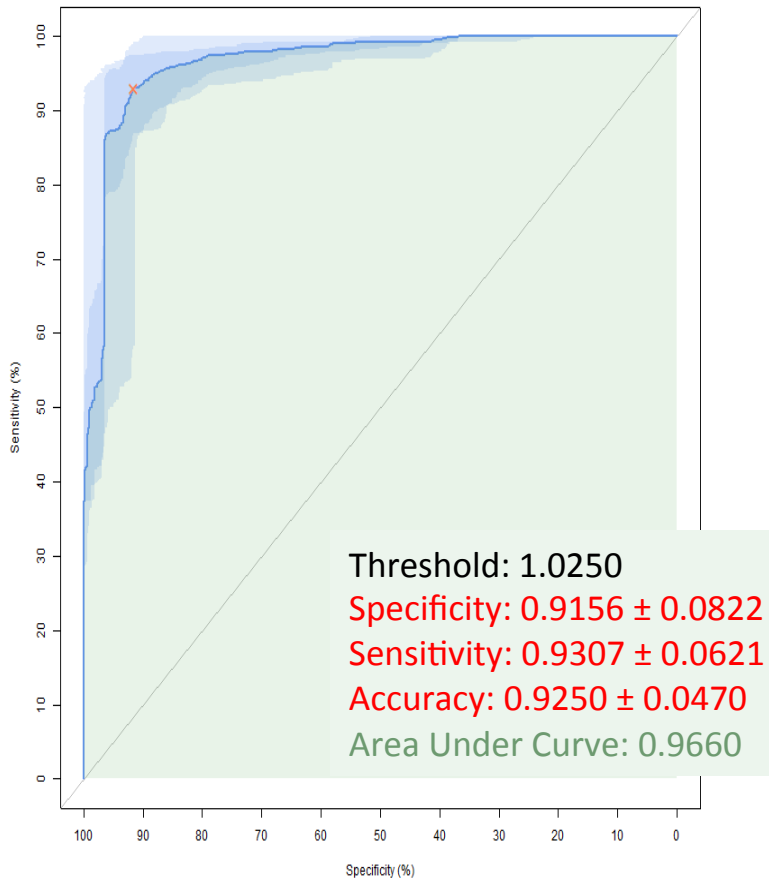
Modeling with 1 miRNA (hsa-miR-614)



$$\text{Prediction score} = 11.3262 - 1.5682 \times \text{hsa-miR-614}$$

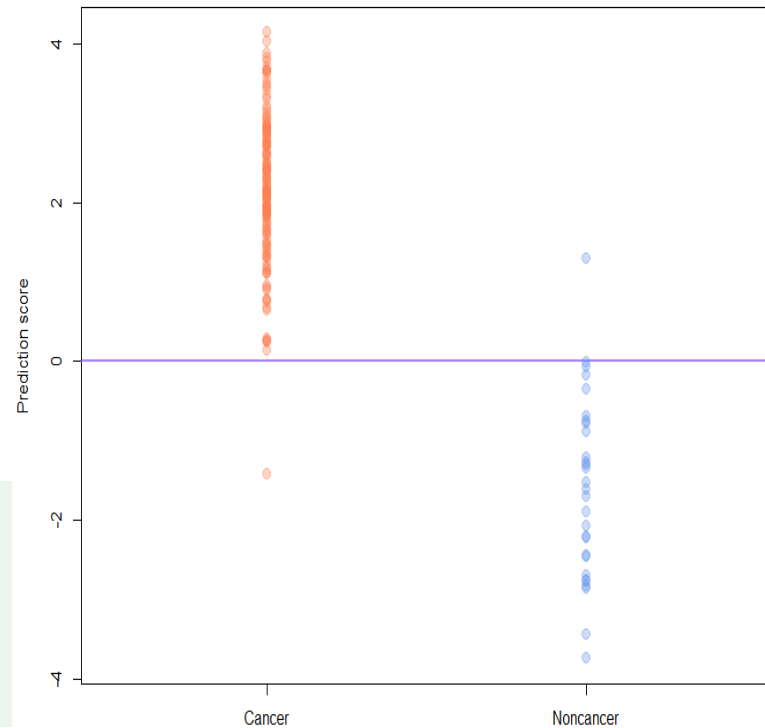
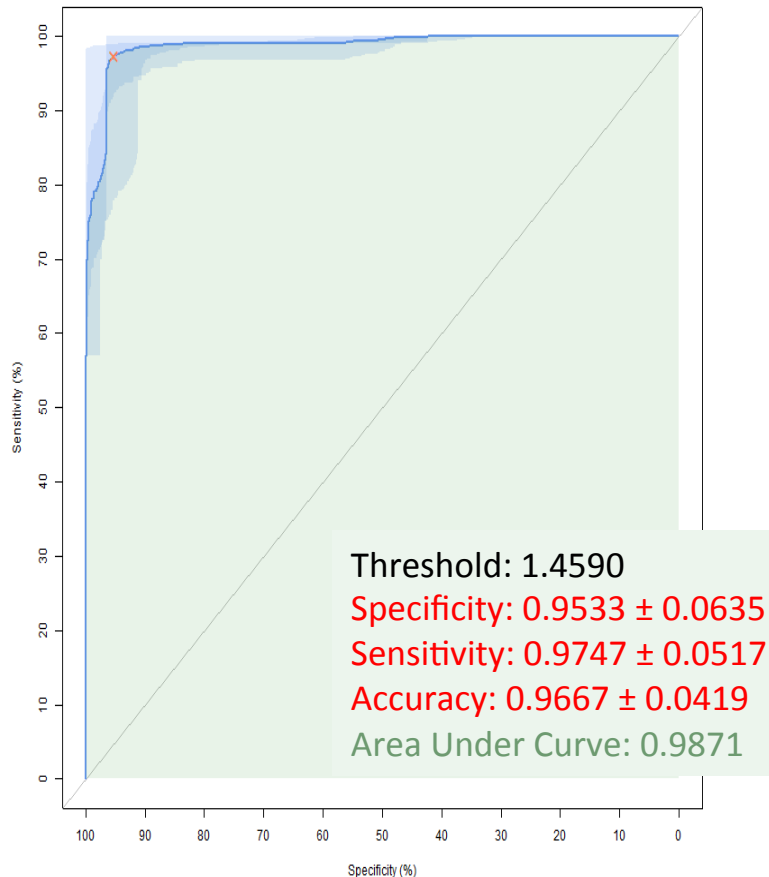
Further Modeling (with 3 miRNAs)

Prediction score = $5.4049 - 1.7271 \times \text{hsa-miR-614} + 0.0937 \times \text{hsa-miR-42XX} + 1.1171 \times \text{hsa-miR-61XX}$



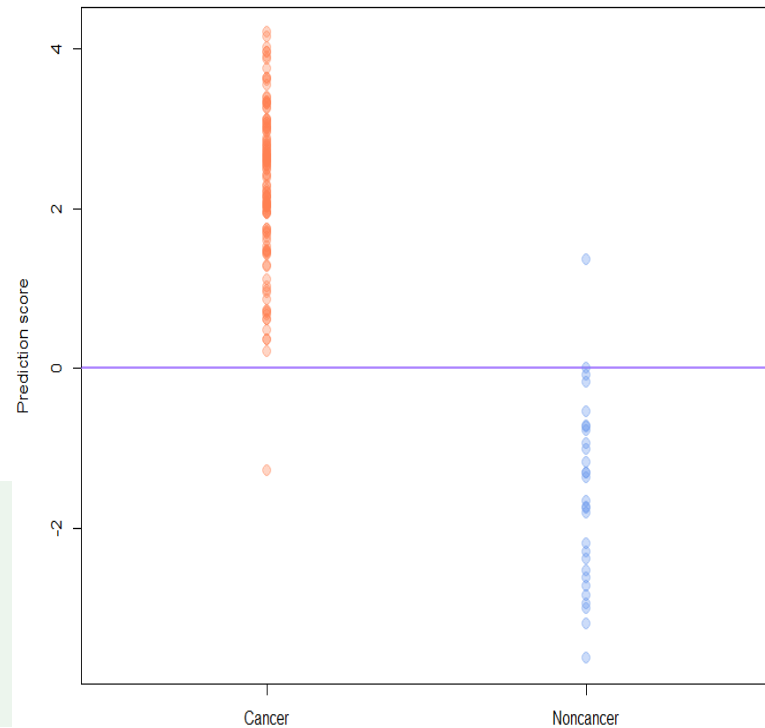
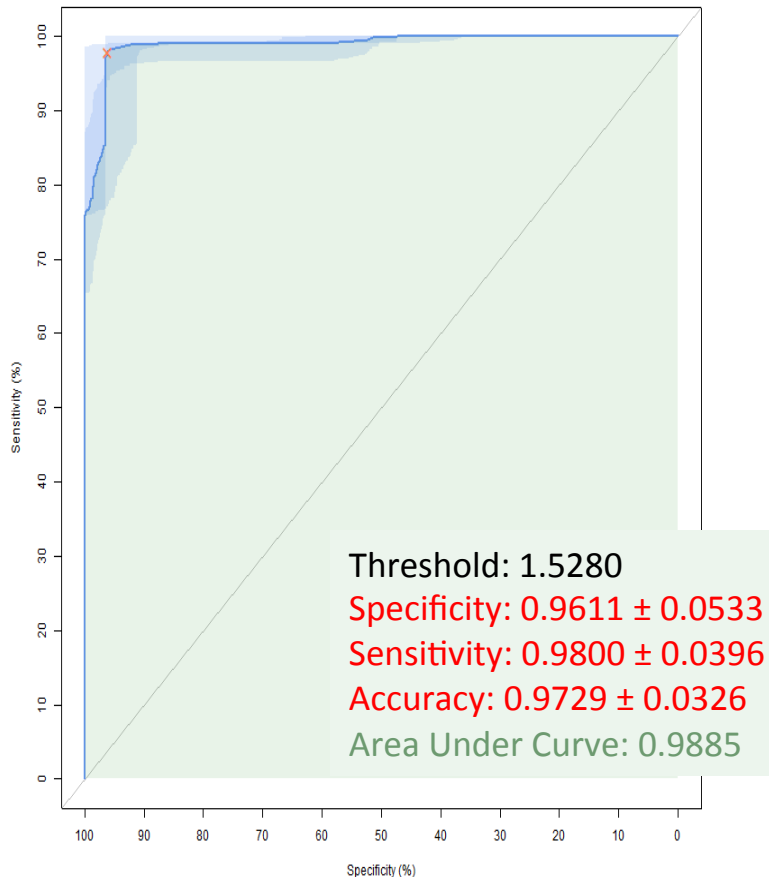
Further Modeling (with 4 miRNAs)

Prediction score = $9.225 - 0.9554 \times \text{hsa-miR-614} + 0.8076 \times \text{hsa-miR-42xx} + 1.4167 \times \text{hsa-miR-61xx} - 1.9153 \times \text{hsa-miR-66XX}$



Further Modeling (with 5 miRNAs)

Prediction score = $8.763 - 0.665 \times \text{miR-614} + 0.865 \times \text{miR-42xx} + 1.413 \times \text{miR-61xx} - 1.697 \times \text{miR-66xx} - 0.716 \times \text{miR-1xxx}$



Validation

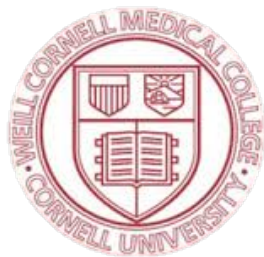
- Patient Serum Collection: Healthy, Benign, pre-Cancer & Cancer
- Workflow:
 1. Isolate RNA
 2. Optimize primer
 3. Reverse transcription
 4. Amplify cDNA
 5. Run qPCR / nanostring/(liquid) chip
 6. Analyze

Conclusions

1. While applying different analysis pipeline would get quiet different outcomes, there are some overlaps, which show the consistency of these methods.
2. Several serum miRNAs are enough to identify the group (cancer or noncancer) of a patient at a high accuracy level. Thus, these selected miRNAs could be viewed as potential biomarkers for implementing early detection of breast cancer.

Perspectives

1. The models seem to be precise enough to fit early detection, more validations are still required to establish robust criteria.
2. The established useful analysis pipeline enables applying for other different expression data derived from other diseases.
3. The mechanisms of these selected miRNAs related are unknown. It is much more meaningful and critical for the understanding of these identified biomarkers. By comprehending the molecular mechanisms underlying these biomarkers, the developing effective treatments and translational research would be promoted.



Acknowledgements

Lab members

Bao-Hong Lee, PhD

Yu-Ling Tai, PhD

Yumi Yasua, PhD

Emily Lo, PhD

Shion-Shan Shen, PhD

Shang-Jer Kuo (PhD student)

Chia-Yu Yang (Master)

Ru-Ying Fang (Master)

Yi-Ting Chen (Master)

Dan-Jung Chuang (Master)

Shi-Chen Wang (Master)

Mitch Lin (Master)

Angela Wang (RA)

Jeff Ku (RA)

NTU

King-Jen Chang, MD, PhD

Wen-Hong Kuo, MD, PhD

Ming-Yang Wang, MD, PhD

Chen-Chih Hsu, PhD

Chen-An Tsai, PhD

Japan

Takahiro Ochiya, PhD (NCC)

Hidetoshi Tahara, PhD (U Hiroshima)

Koji Ueda, PhD (Cancer Found.)

Cornell

David C. Lyden, MD, PhD

Ayuko Hoshino, PhD

Héctor Peinado, PhD (CNIO, Spain)

Others

Arthur Lander, MD, PhD (UC Irvine)

Kuo-Kan Liang (AS)

Ali Mortazabi, PhD (UC Irvine)

Zhongmin Zhao, PhD (UT Health)

