Artificial Intelligence Defines Protein-Based Classification of Thyroid Nodules

Dr. Yaoting Sun
孙耀庭

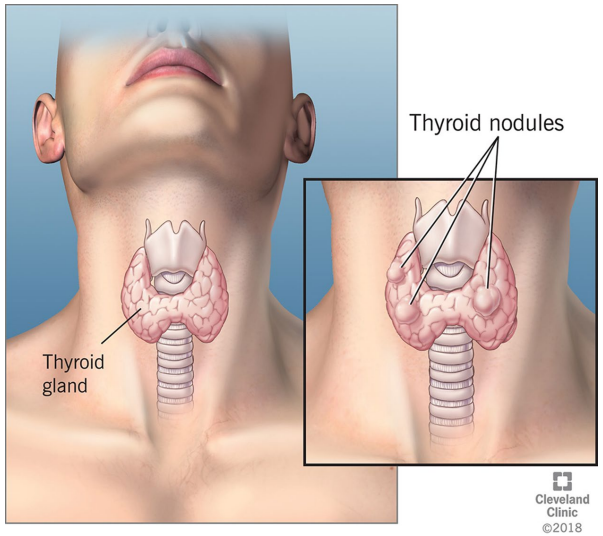
Laboratory of Big Proteomic Data
www.guomics.com
2022.08.04

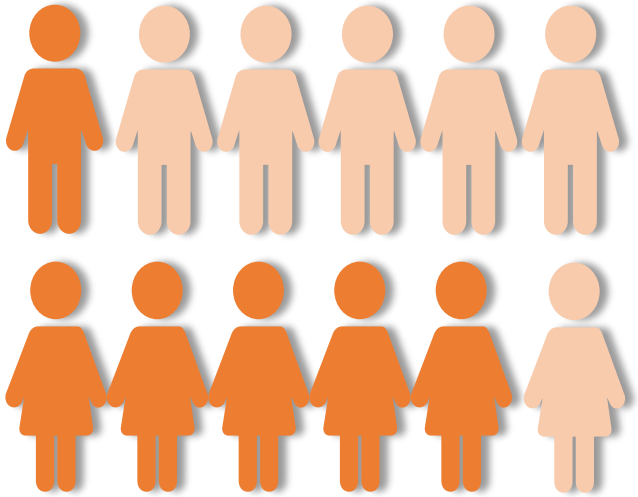
西湖大學
WESTLAKE UNIVERSITY

Background of thyroid nodules

In epidemiological surveys, about 4%-7% of the population have palpable thyroid nodules, compared with 17%-46% on ultrasound, and nearly 60% of the population have thyroid nodules reported in autopsy findings.

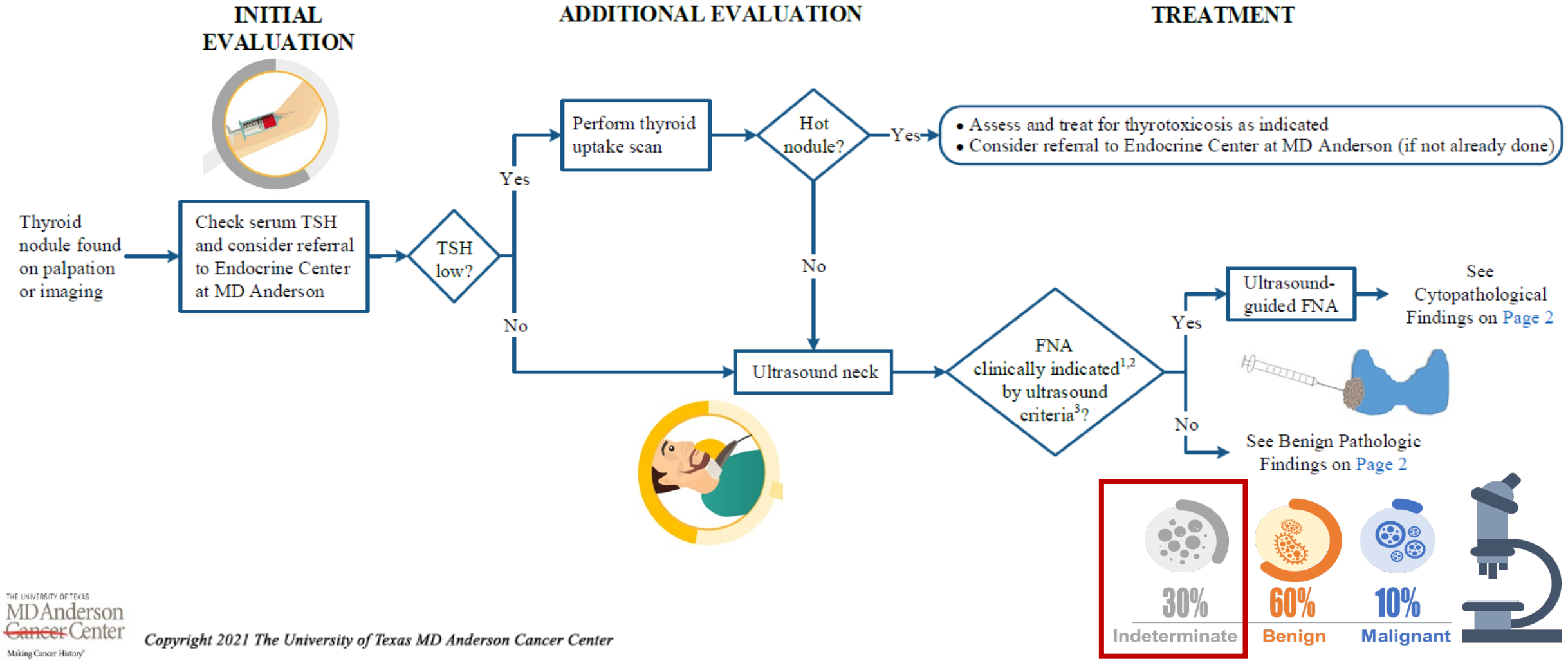
Despite of high incidence of thyroid nodules, only 7-15% of them are malignant.





[1] Fagin JA & Wells SA, Jr. (2016). N Engl J Med 375, 1054-1067
[2] Li Y, et al. (2021). Front Endocrinol (Lausanne) 12, 676144
[3] Zhou J, et al.(2020) Endocrine 70, 256-279
[4] K. D. Burman, L. Wartofsky. (2015). N Engl J Med 373, 2347-2356
[5] C. Durante et al. (2018). JAMA 319, 914-924

Clinical evaluation of patients with thyroid nodules



THE UNIVERSITY OF TEXAS MD Anderson Cancer Center
Making Cancer History®

Nucleic acid-based molecular testing for thyroid nodules

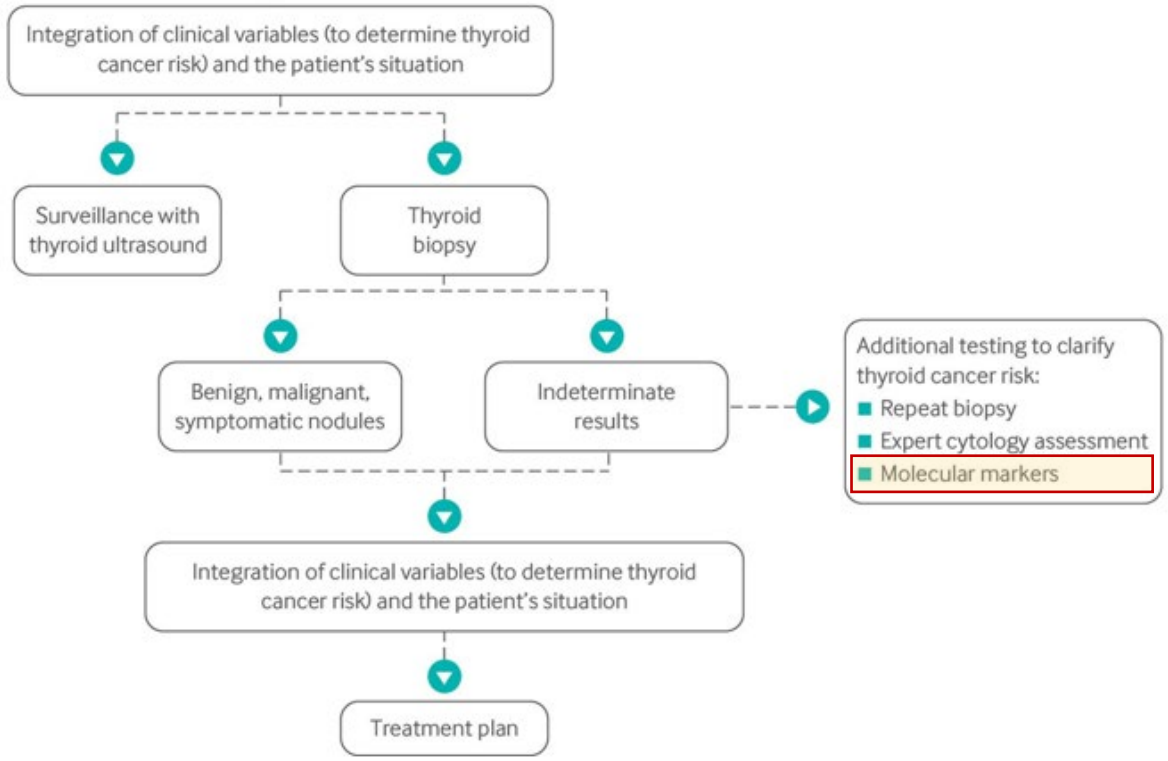


Table 1 | **Commercially available molecular tests for cytologically indeterminate thyroid nodules**

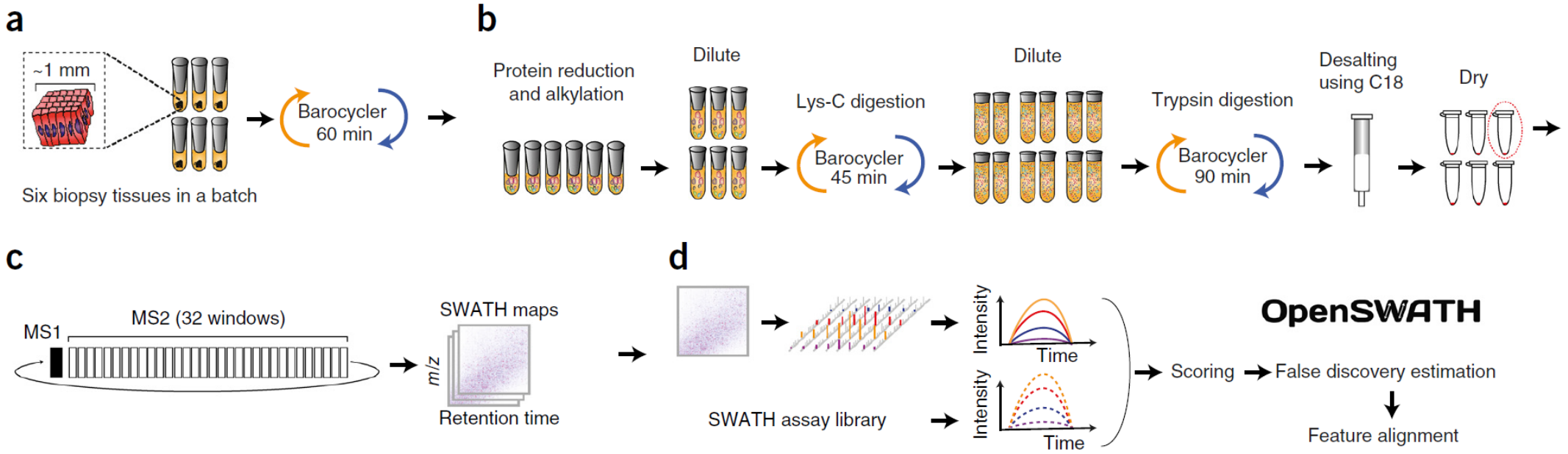
Test	Cytology (Bethesda)	Number of cases	SN	SP	PPV	NPV	Test result	Clinical utility	Refs
GEC	III	129	90	53	38	95	Benign	Active surveillance	19
							Suspicious	Surgery[a]	19
	IV	81	90	49	37	94			
ThyroSeq v2	III	96	91	92	77	97			
	IV	143	90	93	83	96			
ThyroSeq v3	III	154	91	85	64	97			
	IV	93	97	75	68	98	Negative	Active surveillance	25
							Positive	Surgery[a]	25
RosettaGX Reveal	III and IV	150	74	74	43	92	Benign	Active surveillance	113
							Suspicious	Surgery[a]	113
ThyGenX or ThyraMIR	III	58	94	80	68	97	Negative	Surgery or active surveillance	114
							Positive	Surgery[a]	114
	IV	51	82	91	82	91	Negative	Surgery or active surveillance	114
							Positive	Surgery[a]	114

GEC, gene expression classifier testing; NPV, negative predictive value; PPV, positive predictive value; SN, sensitivity; SP, specificity.
[a]Extent of surgery (thyroid lobectomy or total thyroidectomy should be determined by clinical variables, ultrasonographic features and patient preference).

indeterminate thyroid nodules[19]. Subsequent studies have demonstrated significant variation with regard to positive predictive values reported in surgical histopathology in GEC-suspicious nodules (16–57%); the sensitivity (83–100%) and specificity (10–52%) of GEC testing vary across studies, but many reports lacked long-term follow-up of GEC-benign nodules and/or histopathological correlation of GEC findings[17,19,22,23]. Overall, the low specificity excludes its use as a rule-in test, and thyroid lobectomy, at a minimum, should be performed for diagnostic purposes for GEC-suspicious thyroid nodules (TABLE 1).

[1] Singh Ospina N, *et al*. (2020). BMJ 368, l6670
[2] Wang TS & Sosa JA (2018). Nat Rev Endocrinol 14, 670-683

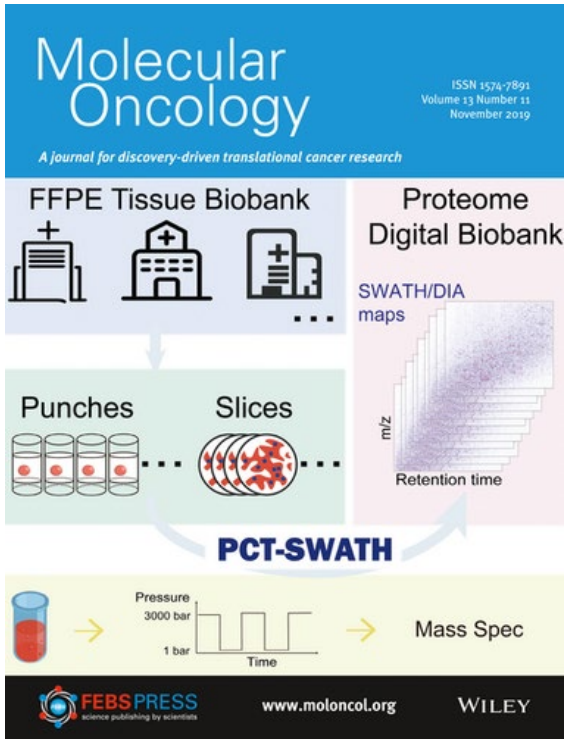
Pressure cycling technology assisted sample preparation

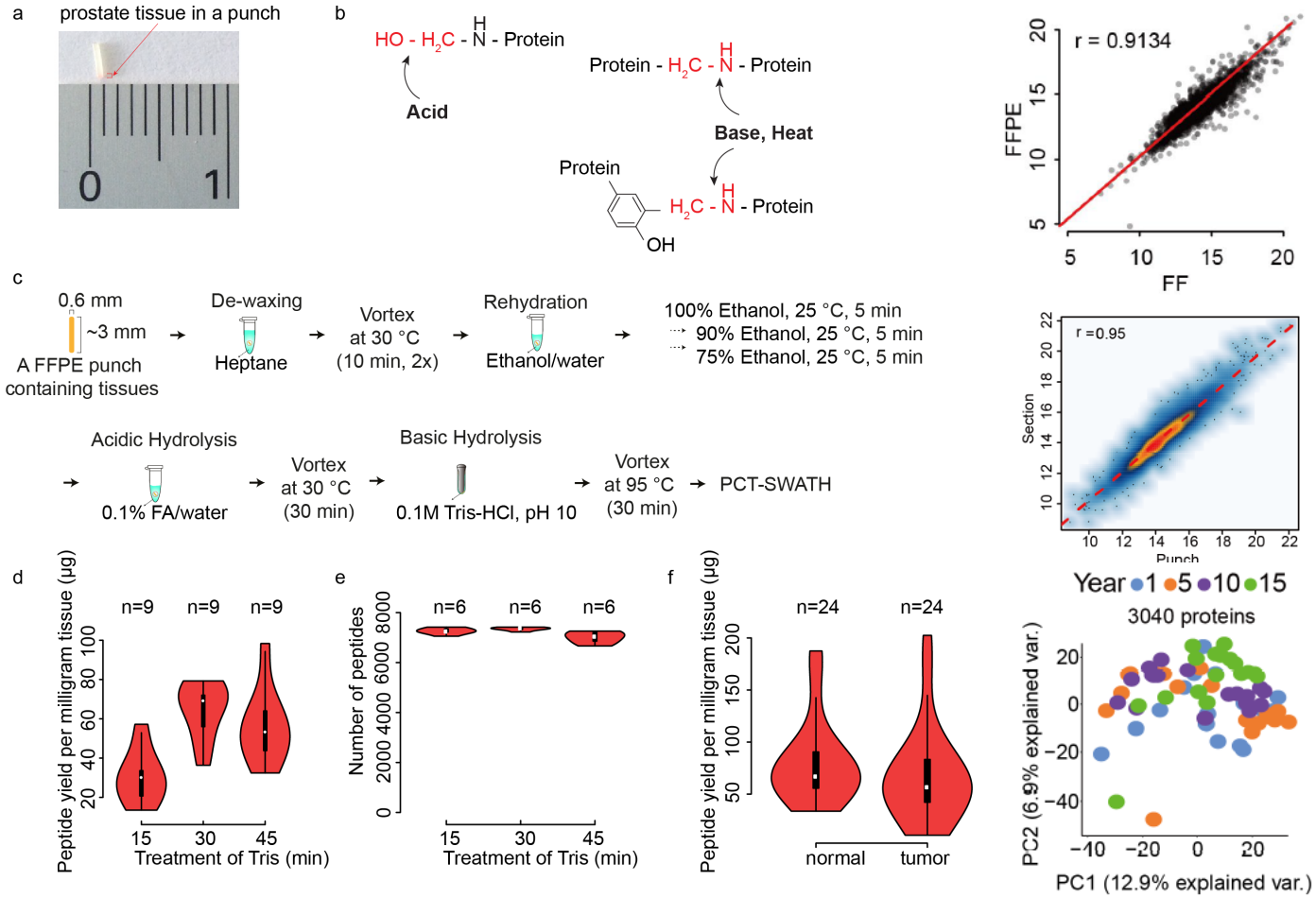


1. High-throughput
2. Small amount of samples(<1mg tissue weight)
3. Fast （100s-1000s）
4. Robustness
5. Potential for quantification of entire proteome

[1] Gillet LC, ..., Bonner R & Aebersold R (2012). Mol Cell Proteomics 11, O111.016717
[2] Guo T, ..., Jochum W & Aebersold R (2015). Nat Med 21, 407-413
[3] Powell K (2018) Technology to watch in 2018. Nature 553, 531-534

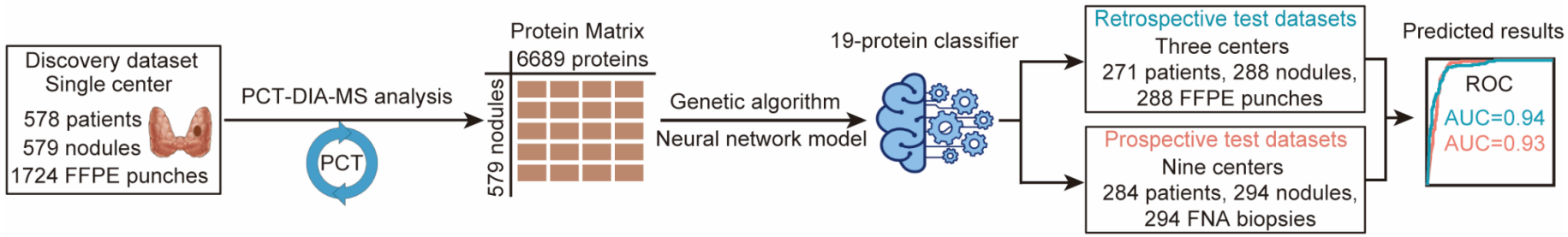
PCT assisted sample preparation from FFPE tissues





Zhu Y, ..., Aebersold R & Guo T (2019). Mol Oncol 13, 2305-2328

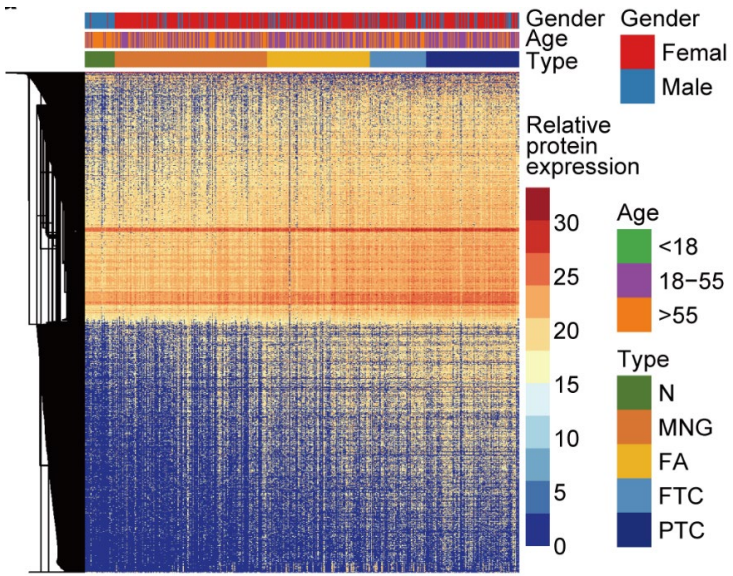
Study design and patient characteristics

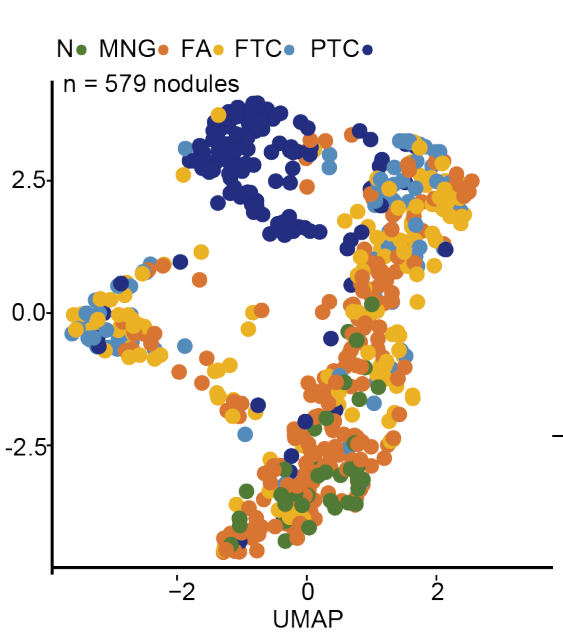


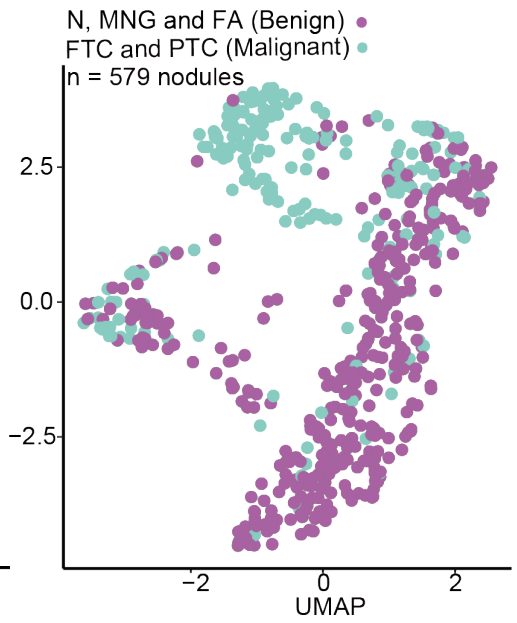
Clinico-pathologic characteristics of the study cohorts

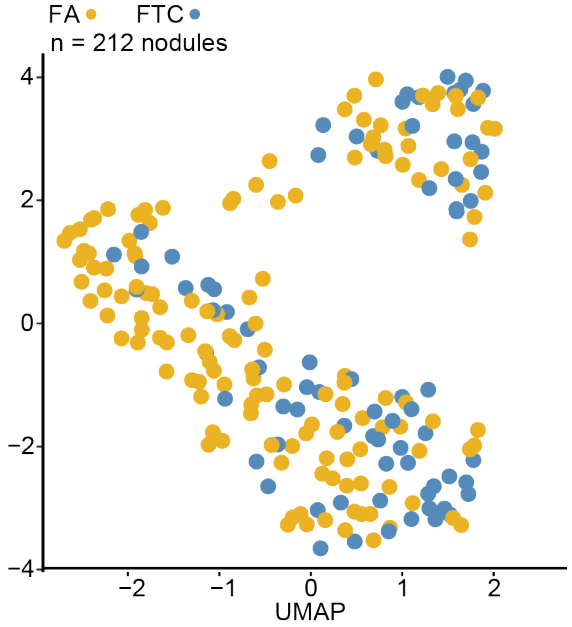
		Discovery dataset (FFPE)	Independent test datasets		All
			Retrospective test datasets (FFPE)	Prospective test datasets (FNA)	
Total no.					
	Clinical centers	1	3	9	12[c]
	Patients	578	271	284	1133
	Nodules	579	288	294	1161
	FFPE cores	1724	288	0	2012
	Fine needle asperation biopsies	0	0	294	294
	DIA files	1780	576	294	2650
Histopathology diagnosis					
	Normal tissues / lymphocytic thyroiditis	40 (6.9%)	16 (5.6%)	6 (2.7%)	64 (5.5%)
	Moltinodular goiter	203 (35.1%)	44 (15.3%)	62 (21.1%)	309 (26.6%)
	Follicular adenoma[a]	137 (23.7%)	84 (29.2%)	23 (7.8%)	244 (21.0%)
	Follicular thyroid carcinoma[a]	75 (13.0%)	52 (18.1%)	4 (1.4%)	131 (11.3%)
	Papillary thyroid carcinoma	124 (21.4%)	92 (31.9%)	197 (67.0%)	413 (35.6%)
Bethesda classification					
	I	-	-	32 (10.9%)	-
	II	-	-	41 (13.9%)	-
	III	-	-	52 (17.7%)	-
	IV	-	-	22 (7.5%)	-
	V	-	-	20 (6.8%)	-
	VI	-	-	127 (43.2%)	-

Global thyroid proteome profile





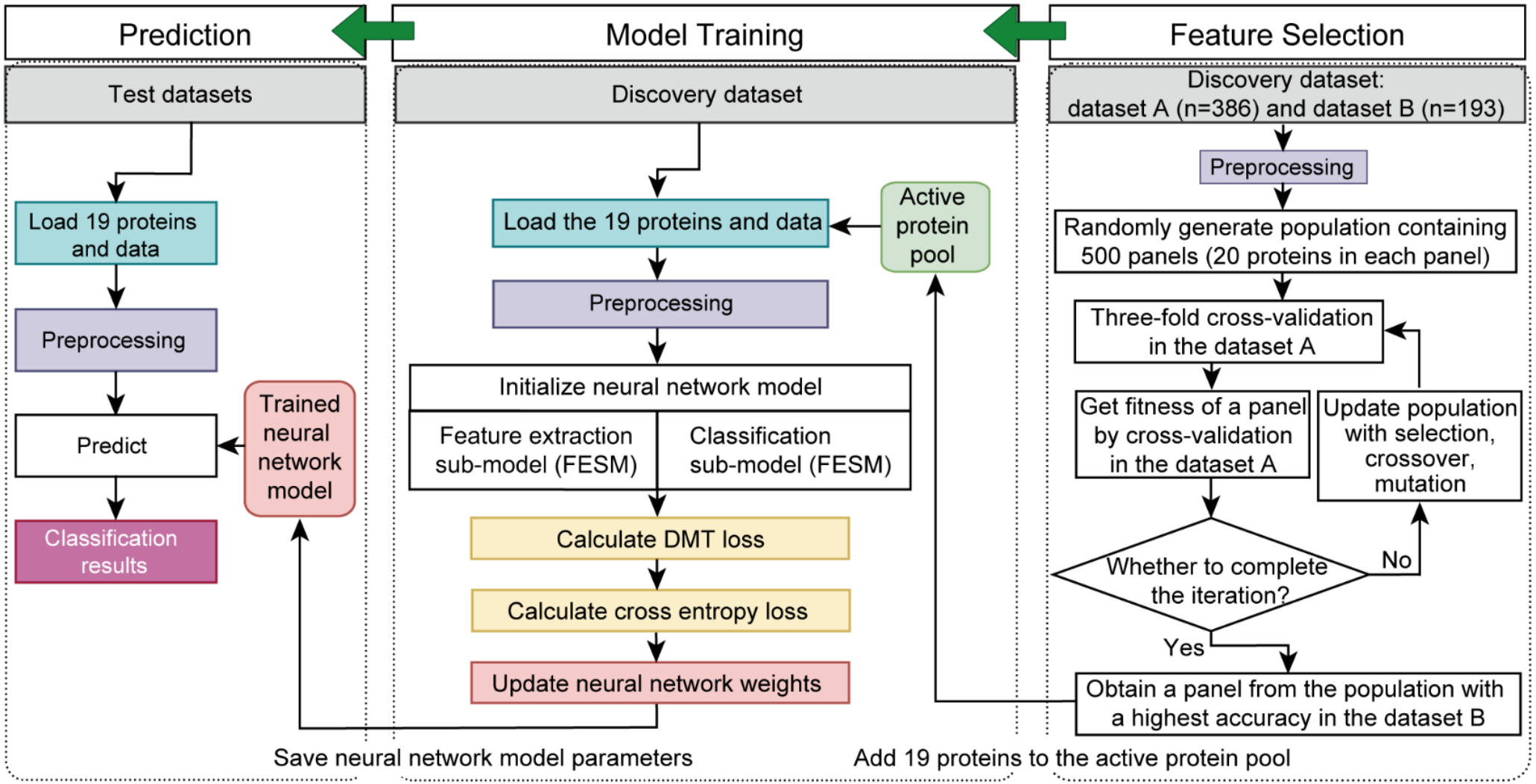




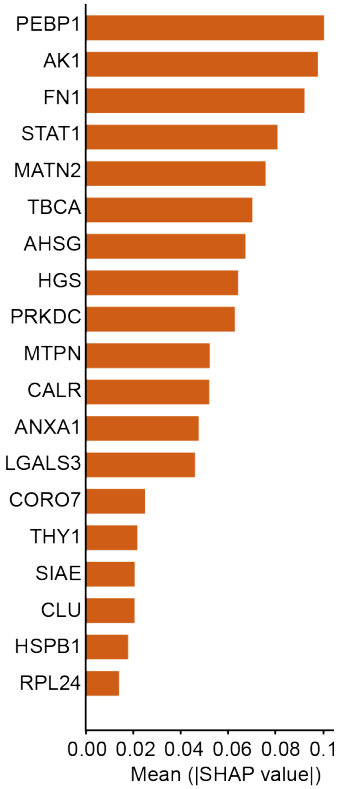
579 thyroid tissue specimens
5312 proteins (rows) are clustered without supervision
Samples (columns) are ordered based on the tissue types

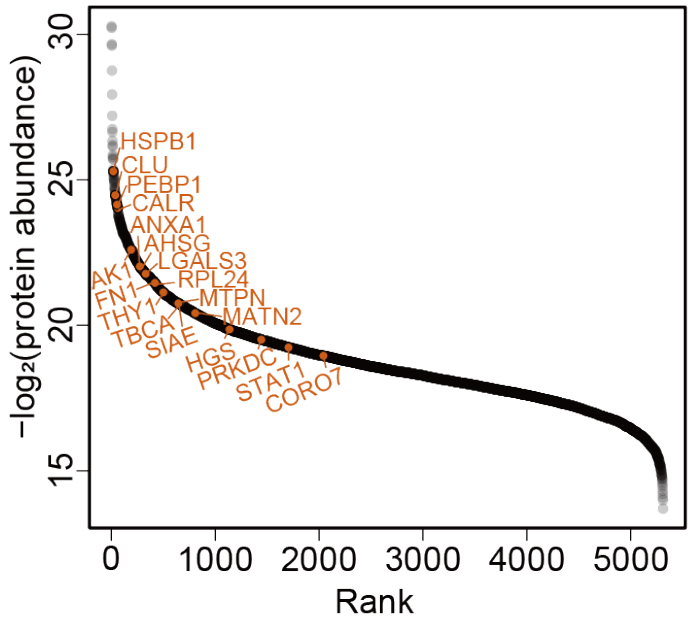
UMAP plots showing global snapshots comparing the indicated types of thyroid tissues using 5312 proteins for all subtypes; malignant versus benign nodules;FA versus FTC.

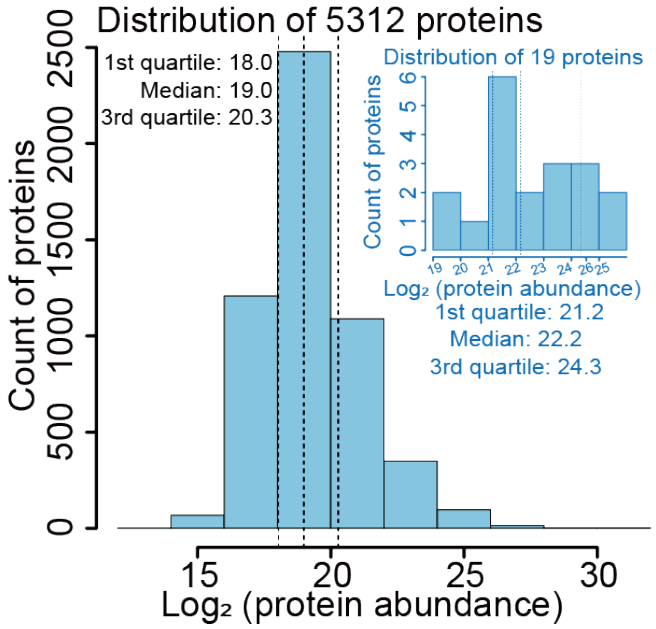
Schematic of principal classifier model



19 selected protein features



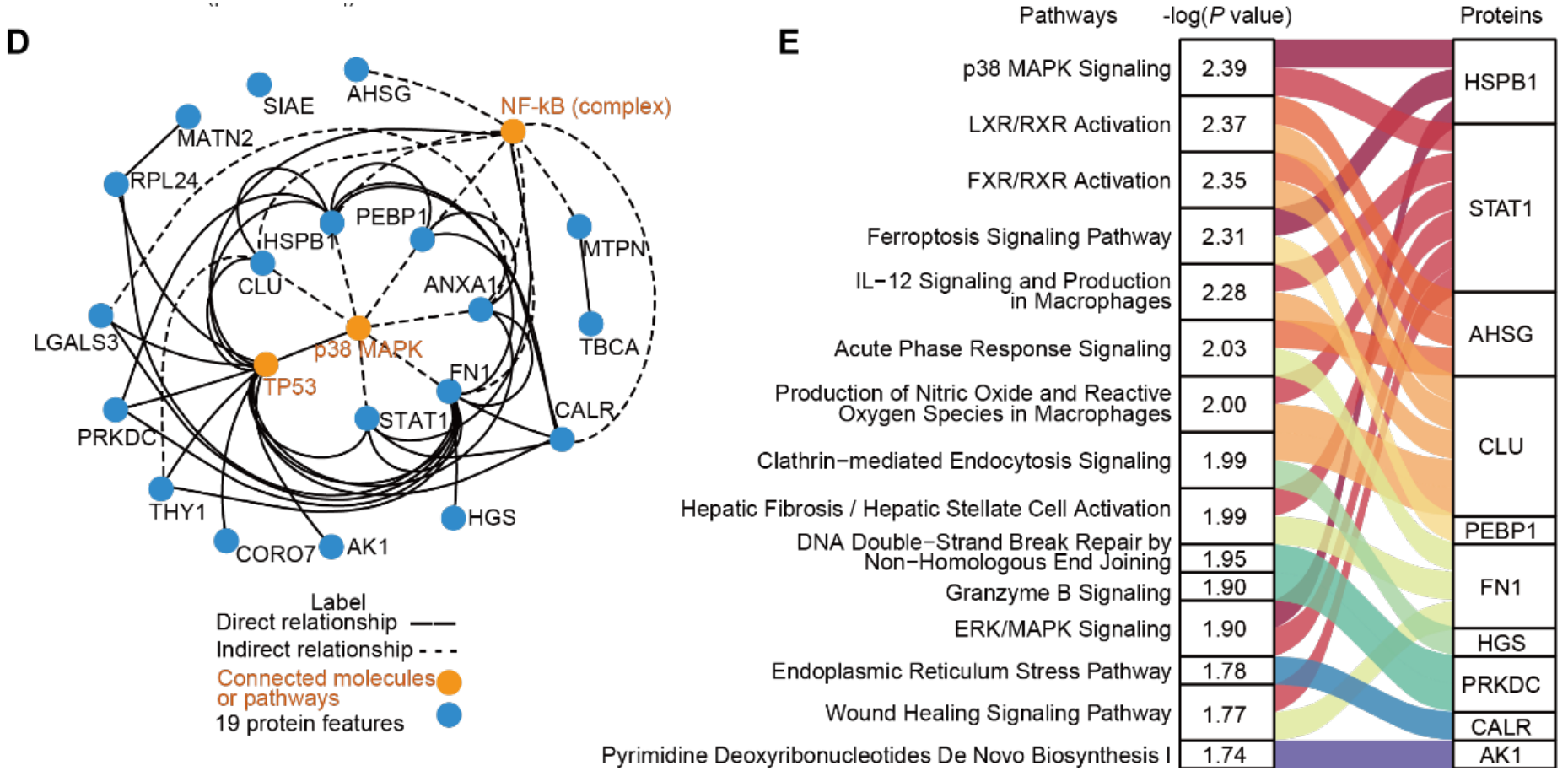




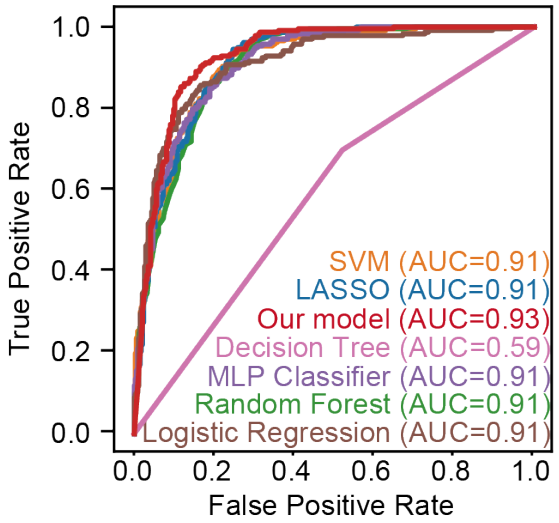
The importance rank of the selected 19 protein features was interpreted by SHapley Additive exPlanations (SHAP) algorithm.

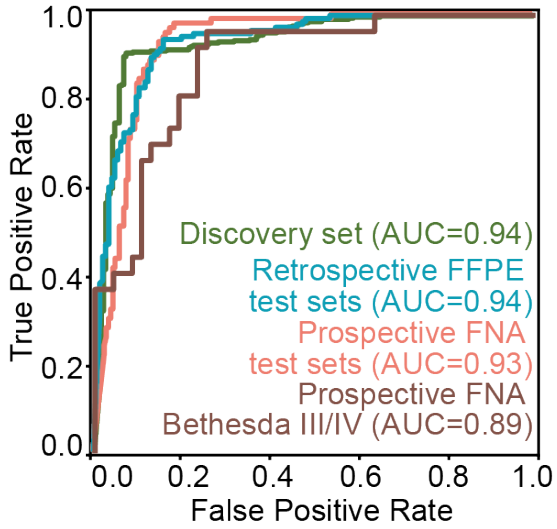
Protein abundance distribution of the 19 features

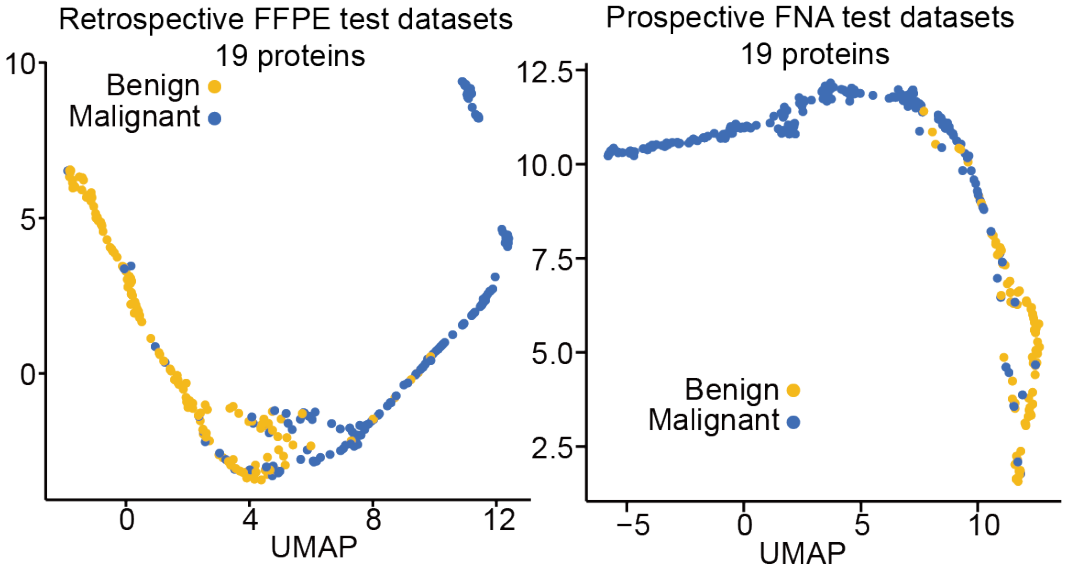
Biological insights on 19 selected features



Multicenter Clinical Evaluation of patients with thyroid nodules







Compared with other six different machine models of 19 selected features

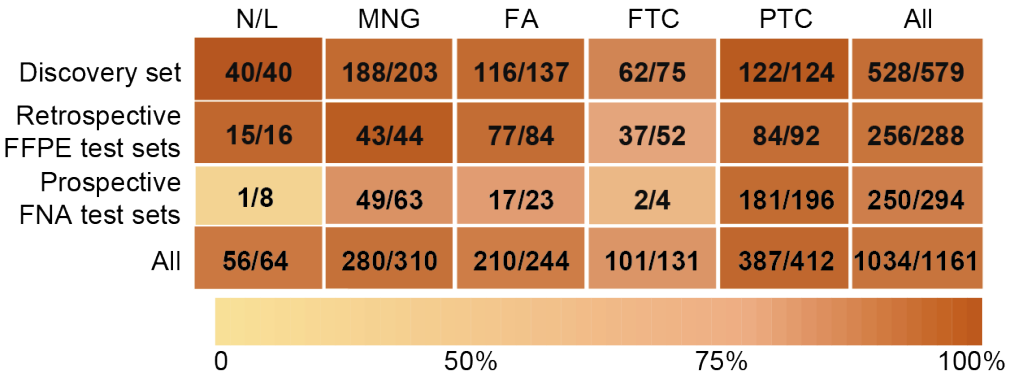
Retrospective sets
FFPE, 3 hospitals, 288 nodules

Prospective sets
FNA, 9 hospitals, 294 nodules

Multicenter Clinical Evaluation of patients with thyroid nodules

Performance of ThyroProt classifier

Method	Prevalence	Sensitivity	Specificity	NPV	PPV	Accuracy
ThyroProt_FFPE	50.00%	84.03%	93.75%	93.08%	85.44%	88.89%
ThyroProt_FNA	68.03%	91.50%	71.28%	87.14%	79.76%	85.03%
ThyroSeq V3[1]	29.57%	93.42%	81.22%	96.71%	67.62%	84.82%
GSC[2]	23.68%	91.11%	68.28%	96.12%	47.13%	73.68%
GEC[3]	32.08%	91.76%	51.67%	93.00%	47.27%	64.53%



[1] Steward, David L., et al. ***JAMA oncology,*** 2019
[2] Kepal N. Patel., et al. **JAMA Surgery,** 2018
[3] Alexander, Erik K., et al. ***New England Journal of Medicine,*** 2012

Acknowledgements

西湖大學 WESTLAKE UNIVERSITY

西湖欧米 WESTLAKE OMICS

He Wang
Weigang Ge
Guan Ruan
Qiushi Zhang

UOMICS -since 2017

Prof. Stan Z. Li
Zelin Zang

All labmates in Laboratory of Big Proteomic Data

ThermoFisher SCIENTIFIC

PBI Pressure BioSciences Inc.

Hospitals and clinical centers

National Cancer Centre Singapore (Prof. Oi Lian Kon; Prof. Gopalakrishna Iyer);
the First Affiliated Hospital, Zhejiang University (Prof. Yijun Wu, Xiaodong Teng, Weibin Wang);
the Second Affiliated Hospital, Zhejiang University (Dr. Shu Zheng, Dr. Lirong Chen);
the Second Hospital of Dalian Medical University (Prof. Yongfu Zhao);
the First Hospital of China Medical University (Prof. Hao Zhang, Zhihong Wang, Wei Sun);
Affiliated Hangzhou First People's Hospital (Prof. Dingcun Luo);
Sir Run Run Shaw Hospital, Zhejiang University (Prof. Jiang Zhu, Jianbiao Wang);
Zhejiang Provincial People's Hospital (Prof. Minghua Ge,Xiaohong Wu, Wanyuan Chen);
Guangdong Provincial People's Hospital (Prof. Haixia Guan);
PLA General Hospital (Prof. Wen Tian);
Fudan University Shanghai Cancer Center (Prof. Yu Wang, Wenjun Wei)
ngsu Provincial Hospital,the First Affiliated Hospital with Nanjing Medical University (Dr.Yan Si)
Cancer Hospital of The University of Chinese Academy of Science (Prof. Dong Xu)
Yantai Yuhuangding Hospital (Dr. Chi Ma);
Renmin Hospital of Wuhan University, Hubei General Hospital (Prof. Chuang Chen)
Zhongnan Hospital of Wuhan University (Dr. Lewei Zheng);
the First Hospital of Shanxi Medical University (Dr. Shujing Li);
Zhongshan Hospital of Xiamen University (Dr. Ende Lin);
Jiangsu Province Traditional and Western Hospital (Dr. Jianhua Wang);
The First Hospital of Jilin University (Dr. Qiyu Lu)
General Hospital of Tianjin Medical University (Dr. Ke Zhao)

THANK YOU

西湖大學
WESTLAKE UNIVERSITY