Identification of metastasis driver mutations by integrating genomics and transcriptomics data

Xing-Ming Zhao

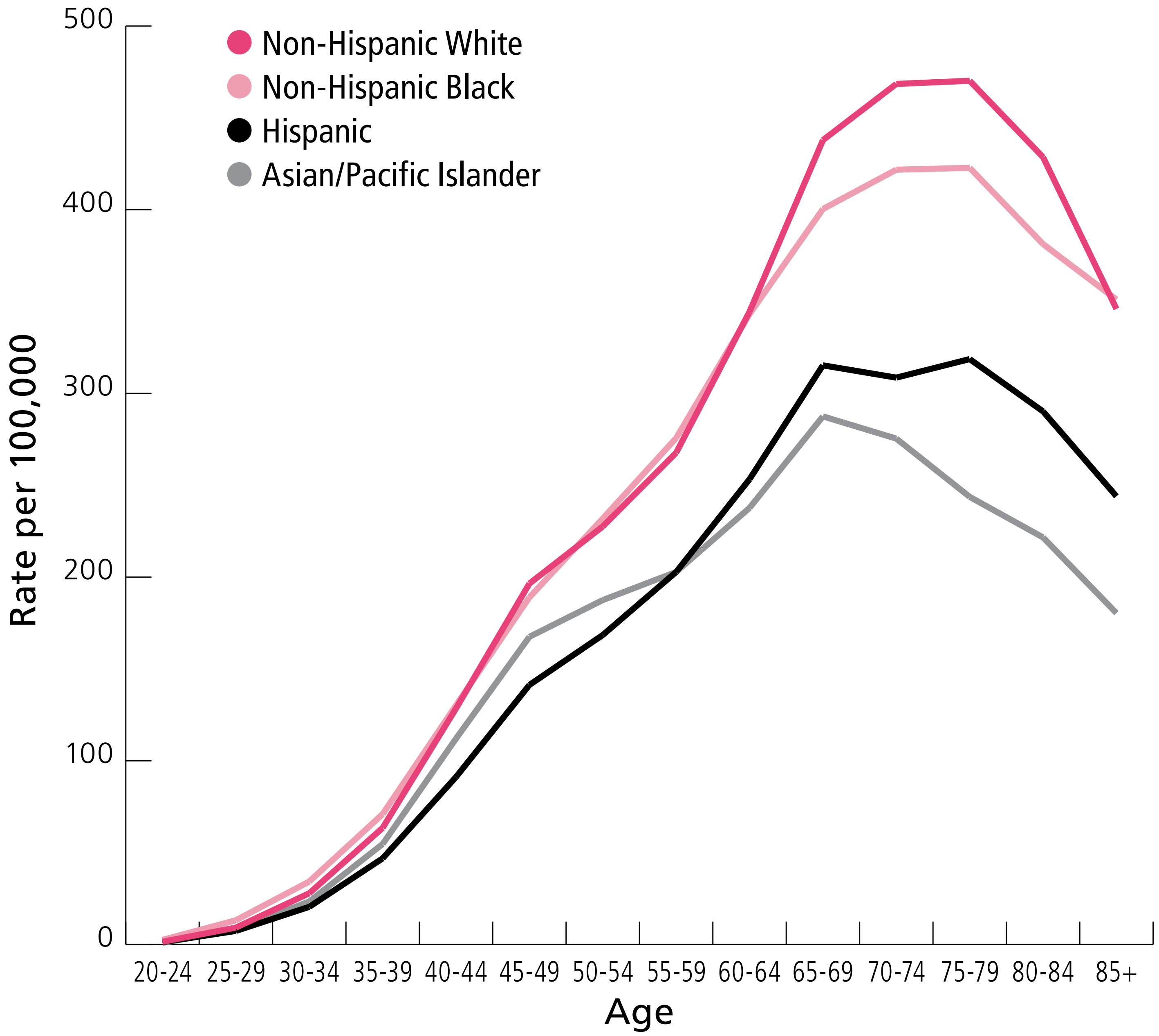
Institute of Science and Technology for Brain-Inspired Intelligence
Fudan University

http://www.comp-sysbio.org/

ISTBI 复旦大学类脑智能科学与技术研究院
Institute of Science and Technology for Brain-Inspired Intelligence

Breast cancer is the most common cancer in women

Figure 1. Age-specific Female Breast Cancer Incidence Rates by Race/Ethnicity, 2010-2014, US



About 1 in 8 U.S. women (about 12.4%) will develop invasive breast cancer over the course of her lifetime.

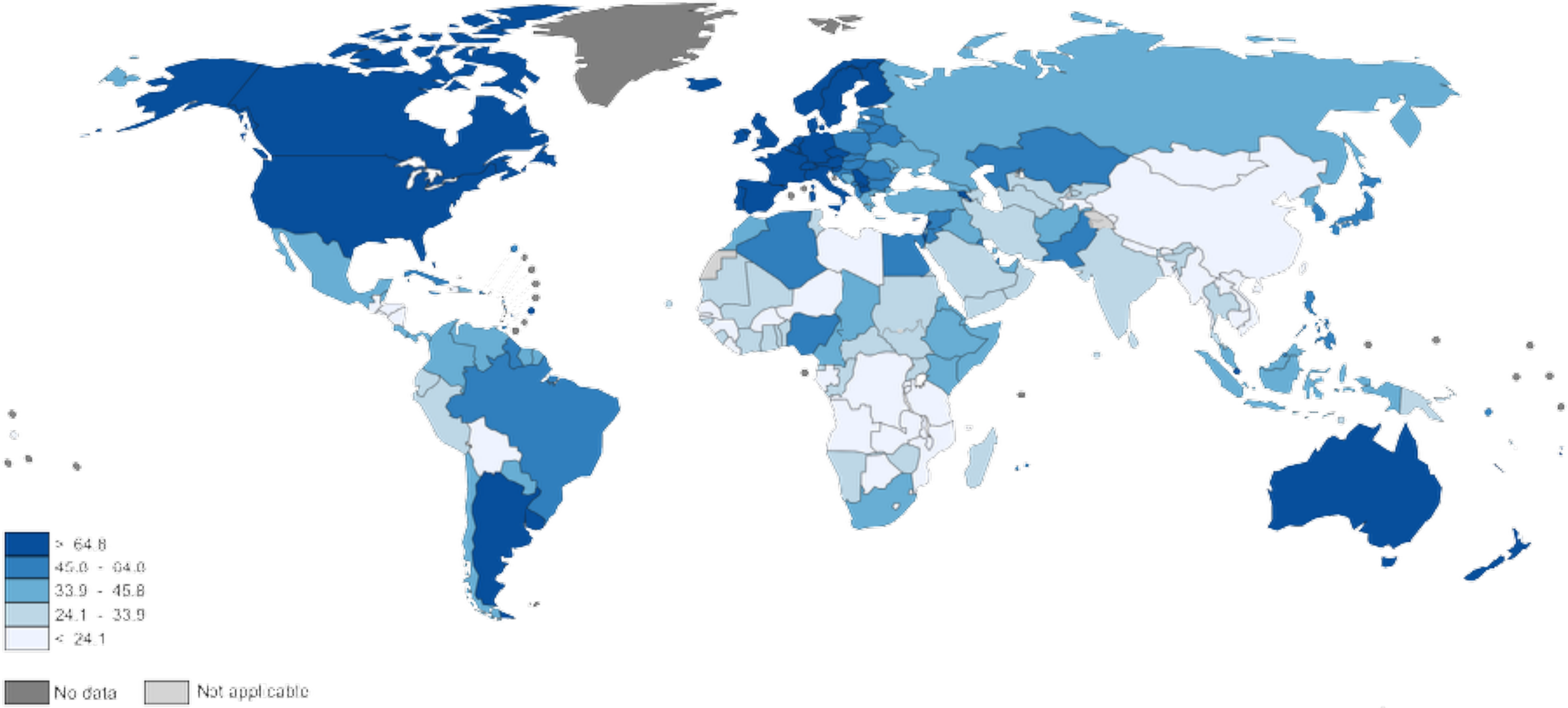
BREASTCANCER.ORG

FUDAN UNIVERSITY 1905

Breast cancer is the most common cancer in women

- Leading cancer in women (23% of all new cases)
- Leading cause of cancer death in women (14% of all cases)
- 1.2 mi. new cases and 500,000 death every year with an increasing trend.

Breast cancer is the most common cancer in women

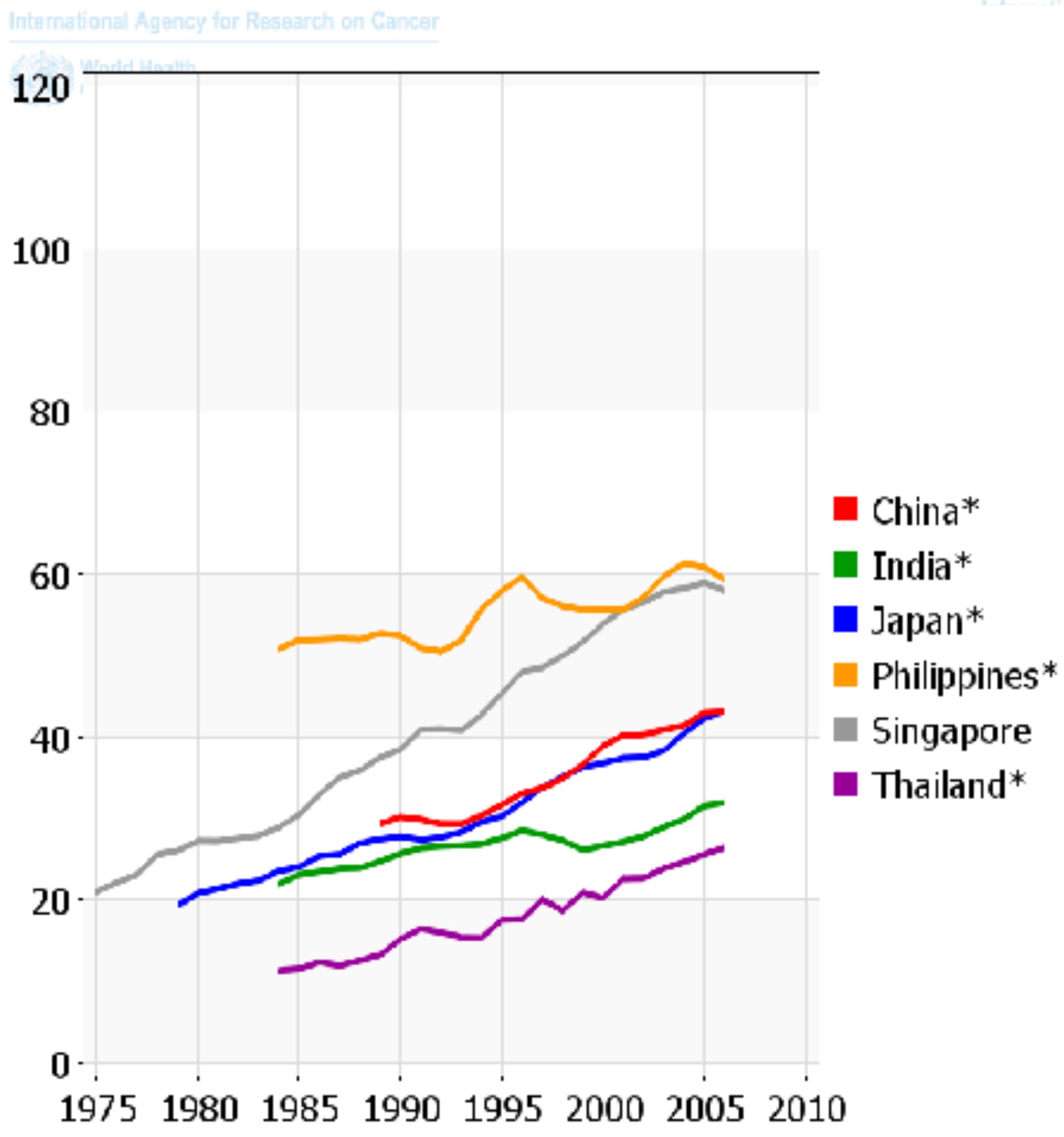


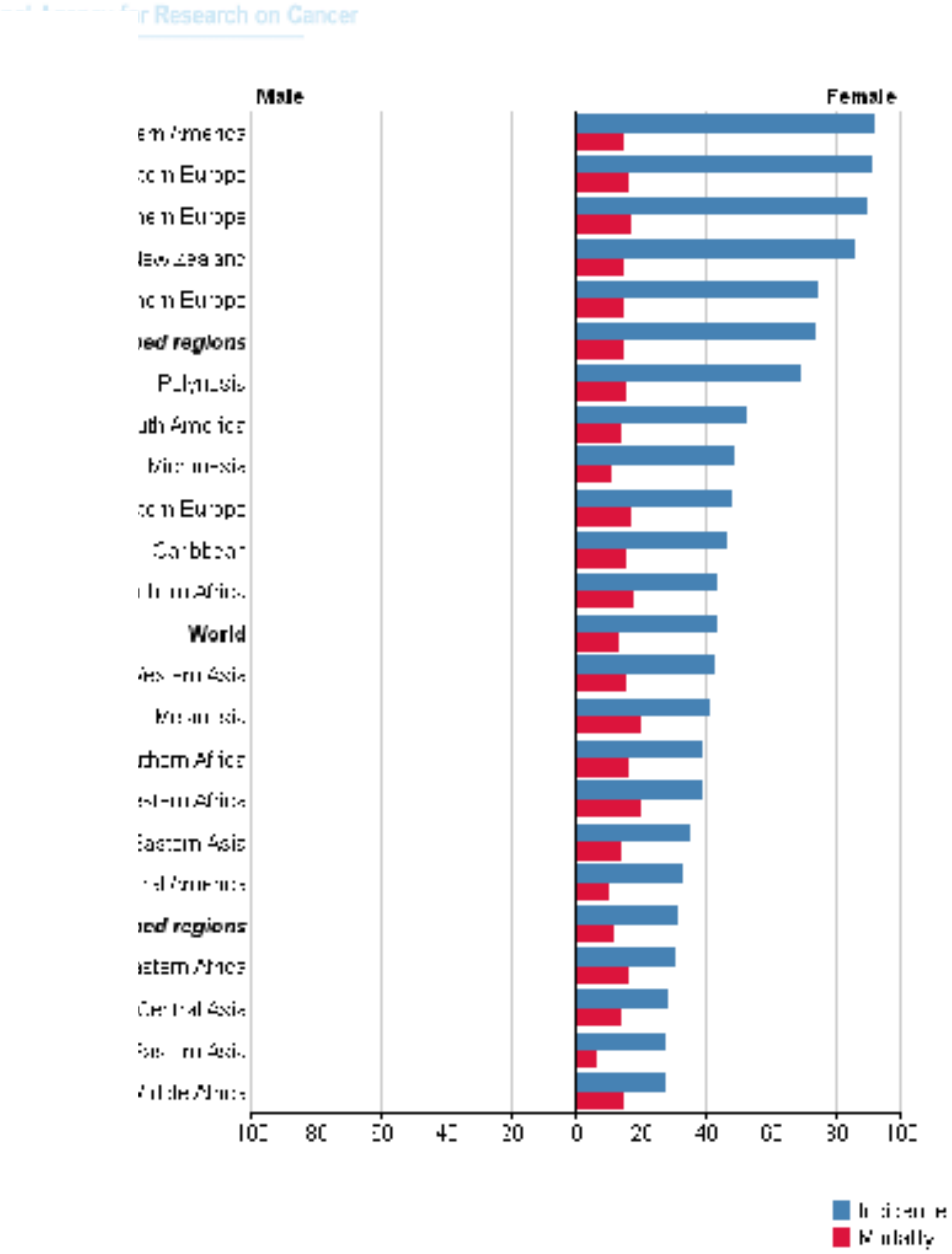
The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

Data source: GLOBOCAN 2012
Map production: IARC
World Health Organization

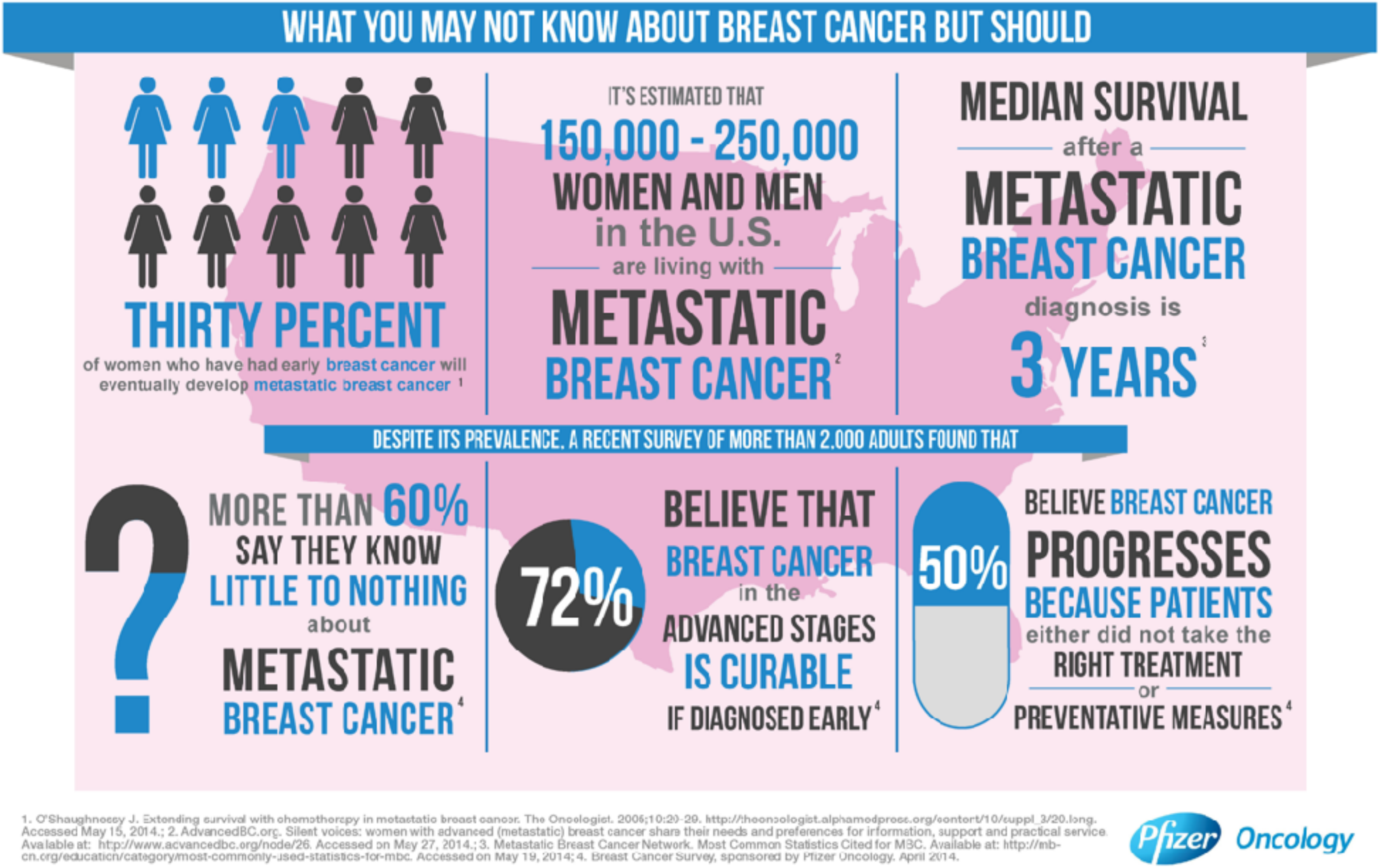
World Health Organization

Breast cancer is the most common cancer in women





Metastatic breast cancer



WHAT YOU MAY NOT KNOW ABOUT BREAST CANCER BUT SHOULD

THIRTY PERCENT
of women who have had early breast cancer will eventually develop metastatic breast cancer[1]

IT'S ESTIMATED THAT
150,000 - 250,000
WOMEN AND MEN
in the U.S.
are living with
METASTATIC BREAST CANCER[2]

MEDIAN SURVIVAL
after a
METASTATIC BREAST CANCER
diagnosis is
3 YEARS[3]

DESPITE ITS PREVALENCE, A RECENT SURVEY OF MORE THAN 2,000 ADULTS FOUND THAT

MORE THAN **60%**
SAY THEY KNOW
LITTLE TO NOTHING
about
METASTATIC BREAST CANCER[4]

72%
BELIEVE THAT
BREAST CANCER
in the
ADVANCED STAGES
IS CURABLE
IF DIAGNOSED EARLY[4]

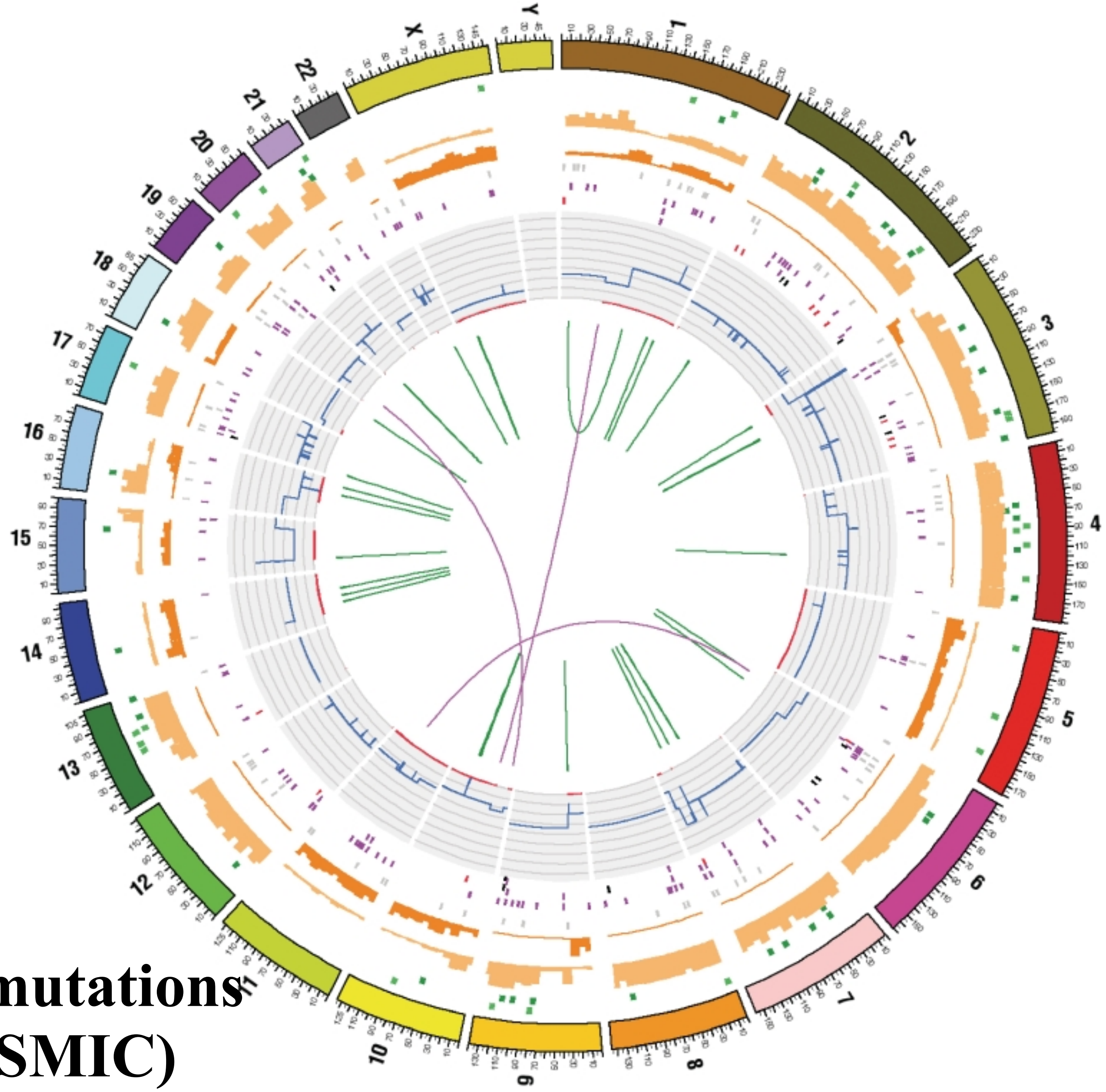
50%
BELIEVE BREAST CANCER
PROGRESSES
BECAUSE PATIENTS
either did not take the
RIGHT TREATMENT
or
PREVENTATIVE MEASURES[4]

1. O'Shaughnessy J. Extending survival with chemotherapy in metastatic breast cancer. The Oncologist. 2005;10:20-29. http://theoncologist.alphamedpress.org/content/10/suppl_3/20.long. Accessed May 15, 2014.; 2. AdvancedBC.org. Silent voices: women with advanced (metastatic) breast cancer share their needs and preferences for information, support and practical service. Available at: http://www.advancedbc.org/node/26. Accessed on May 27, 2014.; 3. Metastatic Breast Cancer Network. Most Common Statistics Cited for MBC. Available at: http://mbcn.org/education/category/most-commonly-used-statistics-for-mbc. Accessed on May 19, 2014; 4. Breast Cancer Survey, sponsored by Pfizer Oncology. April 2014.

Pfizer Oncology

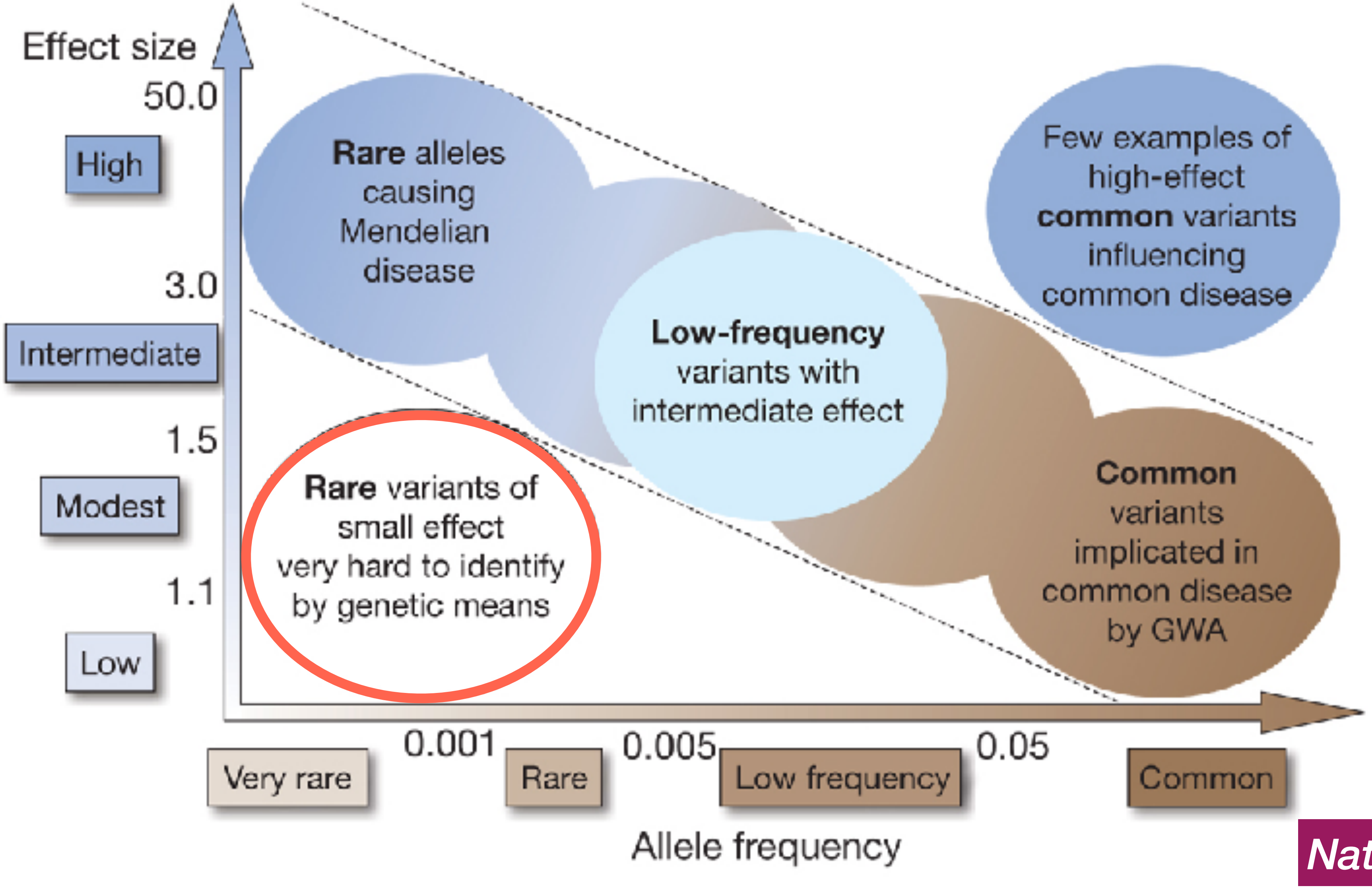
Metastatic breast cancer

The missing heritability of breast cancer



5731490 coding mutations in cancer (COSMIC)

The missing heritability of breast cancer



Nature. 2009. 461(7265): 747–753.

Breast cancer metastasis markers

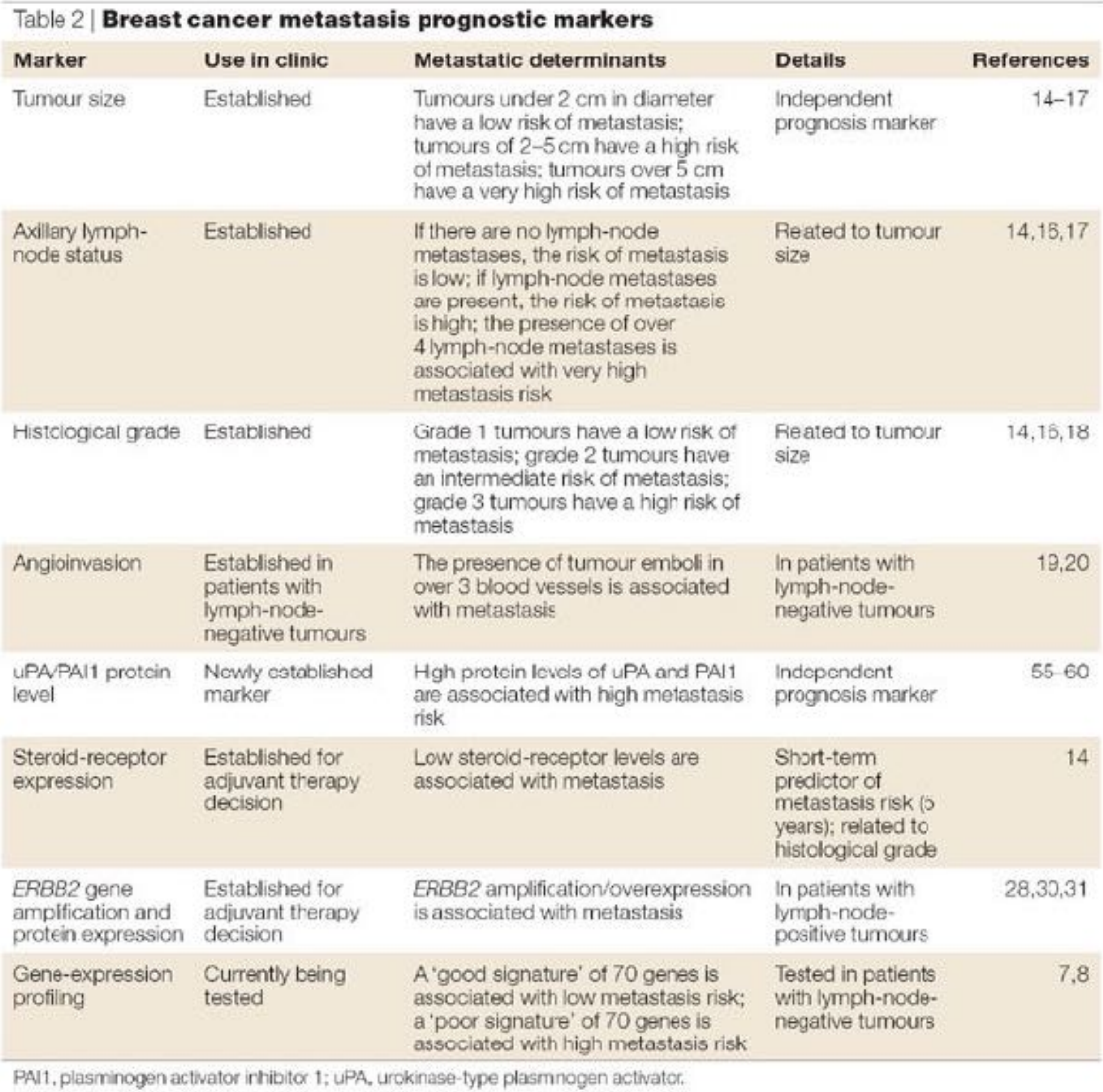


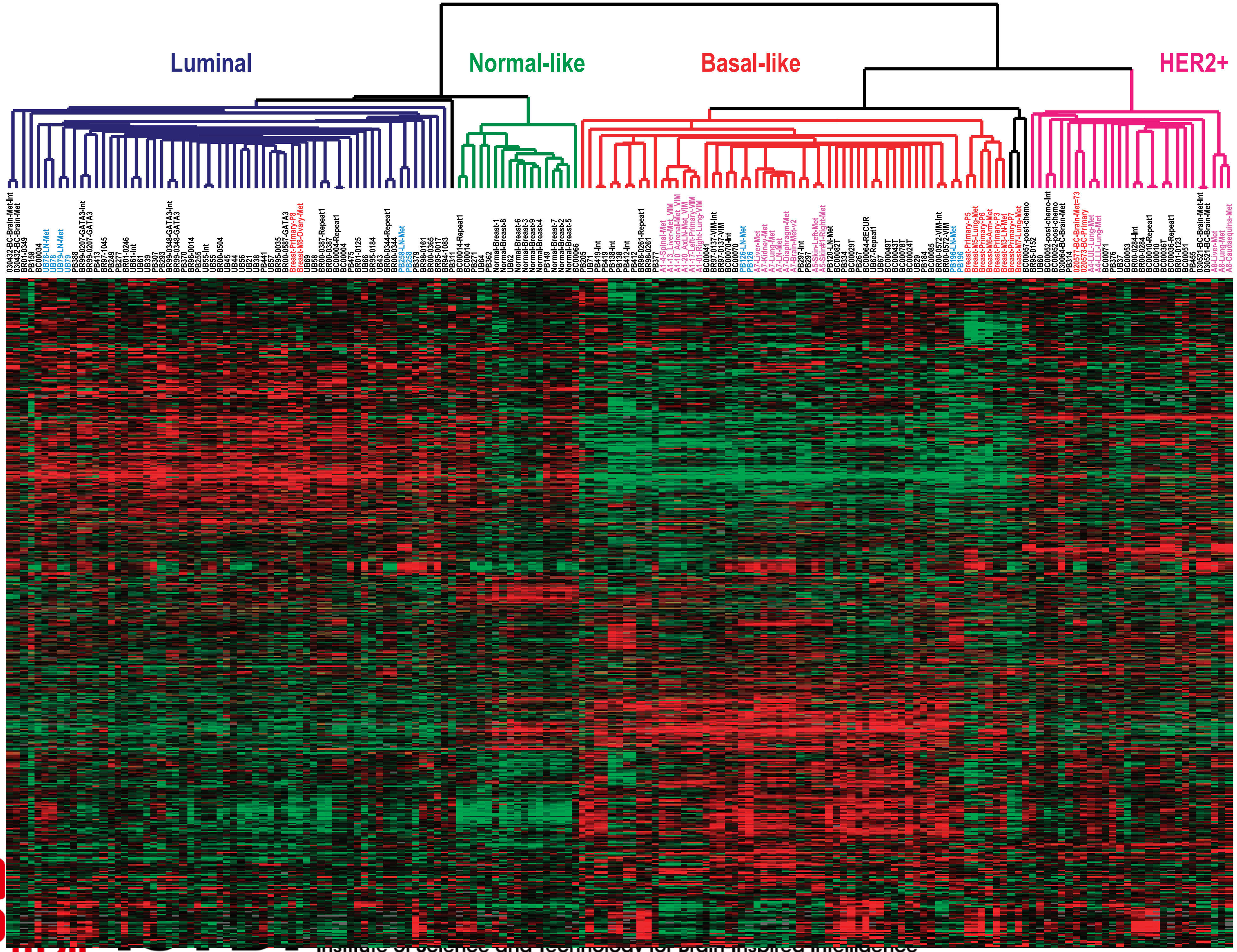
Table 2 | **Breast cancer metastasis prognostic markers**

Marker	Use in clinic	Metastatic determinants	Details	References
Tumour size	Established	Tumours under 2 cm in diameter have a low risk of metastasis; tumours of 2–5 cm have a high risk of metastasis; tumours over 5 cm have a very high risk of metastasis	Independent prognosis marker	14–17
Axillary lymph-node status	Established	If there are no lymph-node metastases, the risk of metastasis is low; if lymph-node metastases are present, the risk of metastasis is high; the presence of over 4 lymph-node metastases is associated with very high metastasis risk	Related to tumour size	14,16,17
Histological grade	Established	Grade 1 tumours have a low risk of metastasis; grade 2 tumours have an intermediate risk of metastasis; grade 3 tumours have a high risk of metastasis	Related to tumour size	14,16,18
Angioinvasion	Established in patients with lymph-node-negative tumours	The presence of tumour emboli in over 3 blood vessels is associated with metastasis	In patients with lymph-node-negative tumours	19,20
uPA/PAI1 protein level	Newly established marker	High protein levels of uPA and PAI1 are associated with high metastasis risk	Independent prognosis marker	55–60
Steroid-receptor expression	Established for adjuvant therapy decision	Low steroid-receptor levels are associated with metastasis	Short-term predictor of metastasis risk (5 years); related to histological grade	14
ERBB2 gene amplification and protein expression	Established for adjuvant therapy decision	*ERBB2* amplification/overexpression is associated with metastasis	In patients with lymph-node-positive tumours	28,30,31
Gene-expression profiling	Currently being tested	A 'good signature' of 70 genes is associated with low metastasis risk; a 'poor signature' of 70 genes is associated with high metastasis risk	Tested in patients with lymph-node-negative tumours	7,8

PAI1, plasminogen activator inhibitor 1; uPA, urokinase-type plasminogen activator.

ISTB. Institute of Science and Technology for Brain-Inspired Intelligence

Breast cancer metastasis markers



Cancer Res.
2005;65(20):9155-8

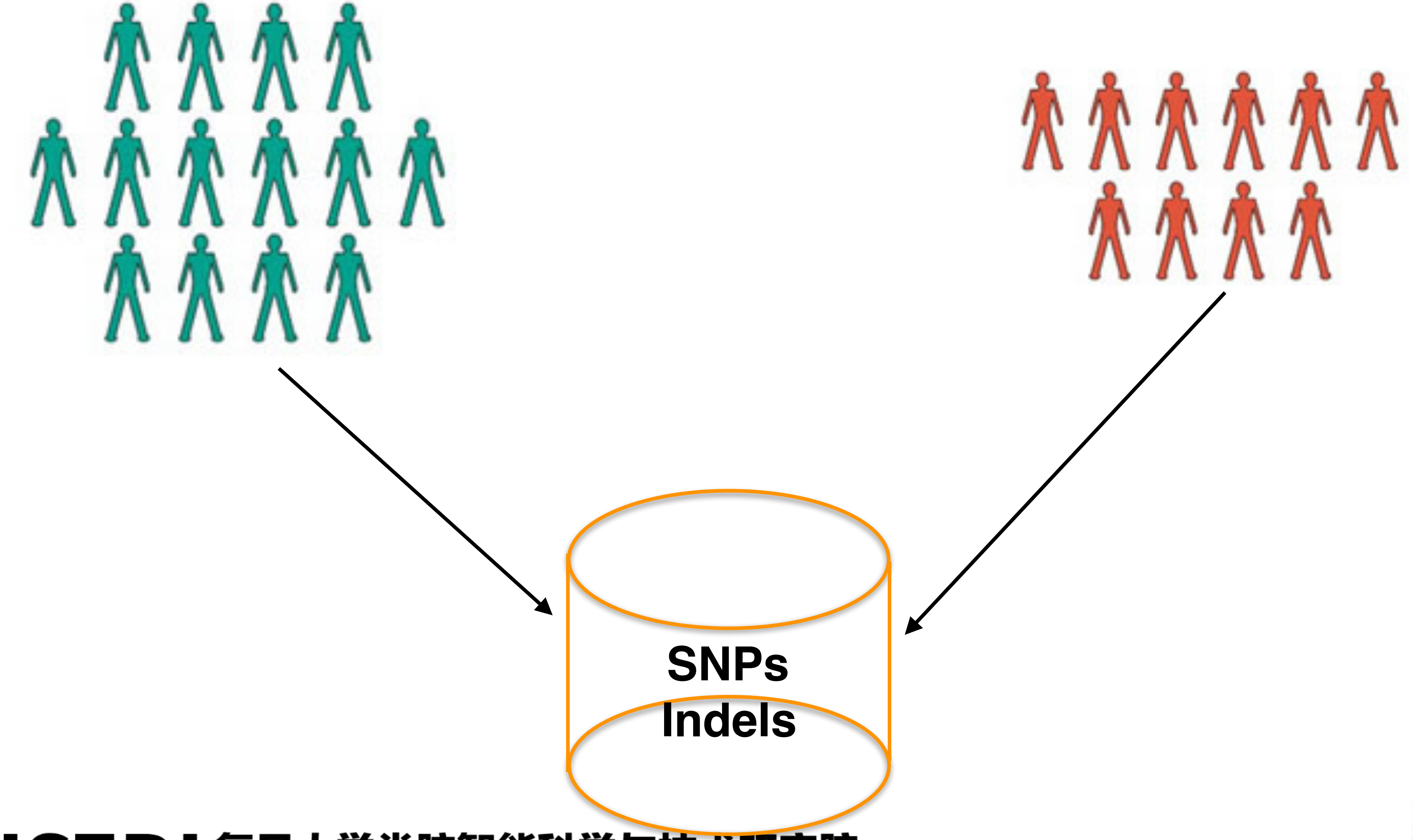
Exome sequencing of breast cancers

56 LRM

(low-risk for distant metastasis)

22 HRM

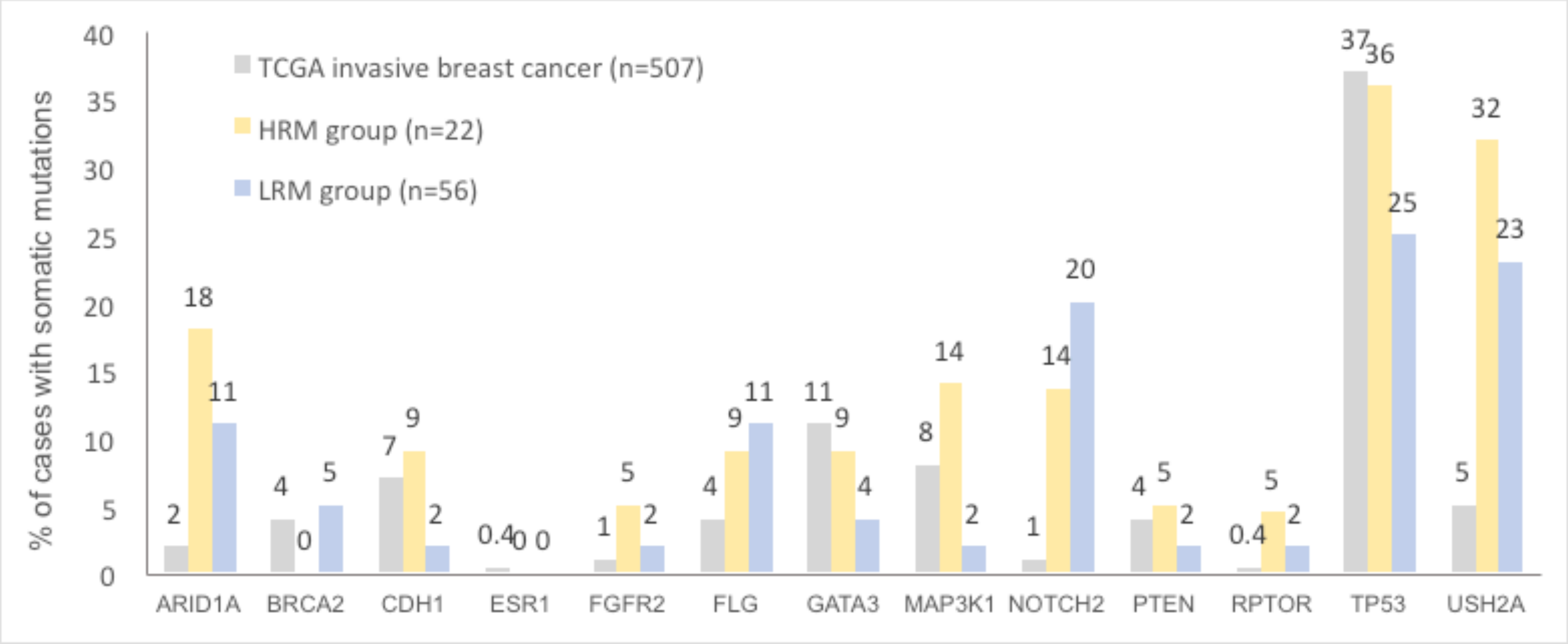
(high-risk for distant metastasis)



SNPs

Indels

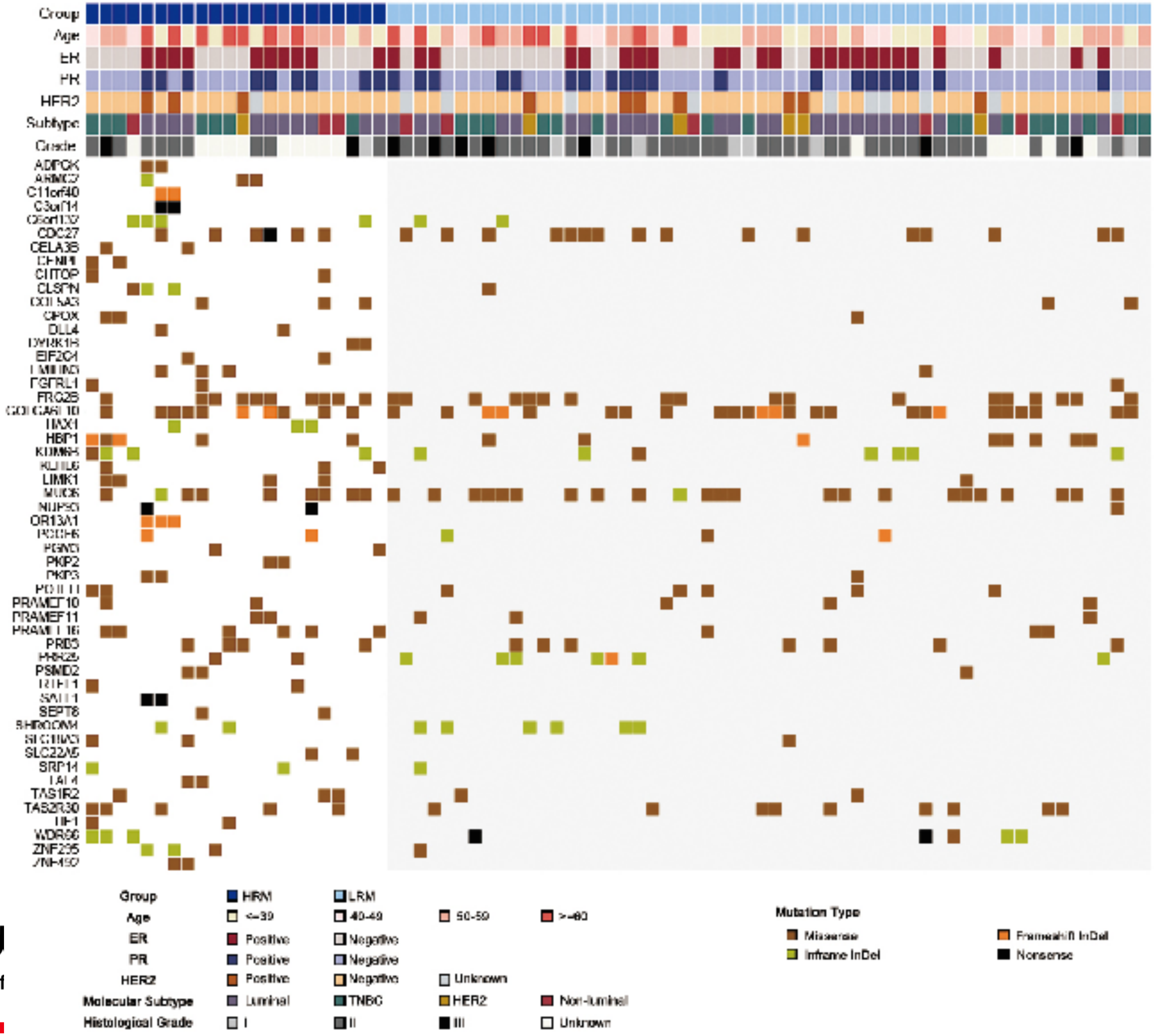
de novo mutations



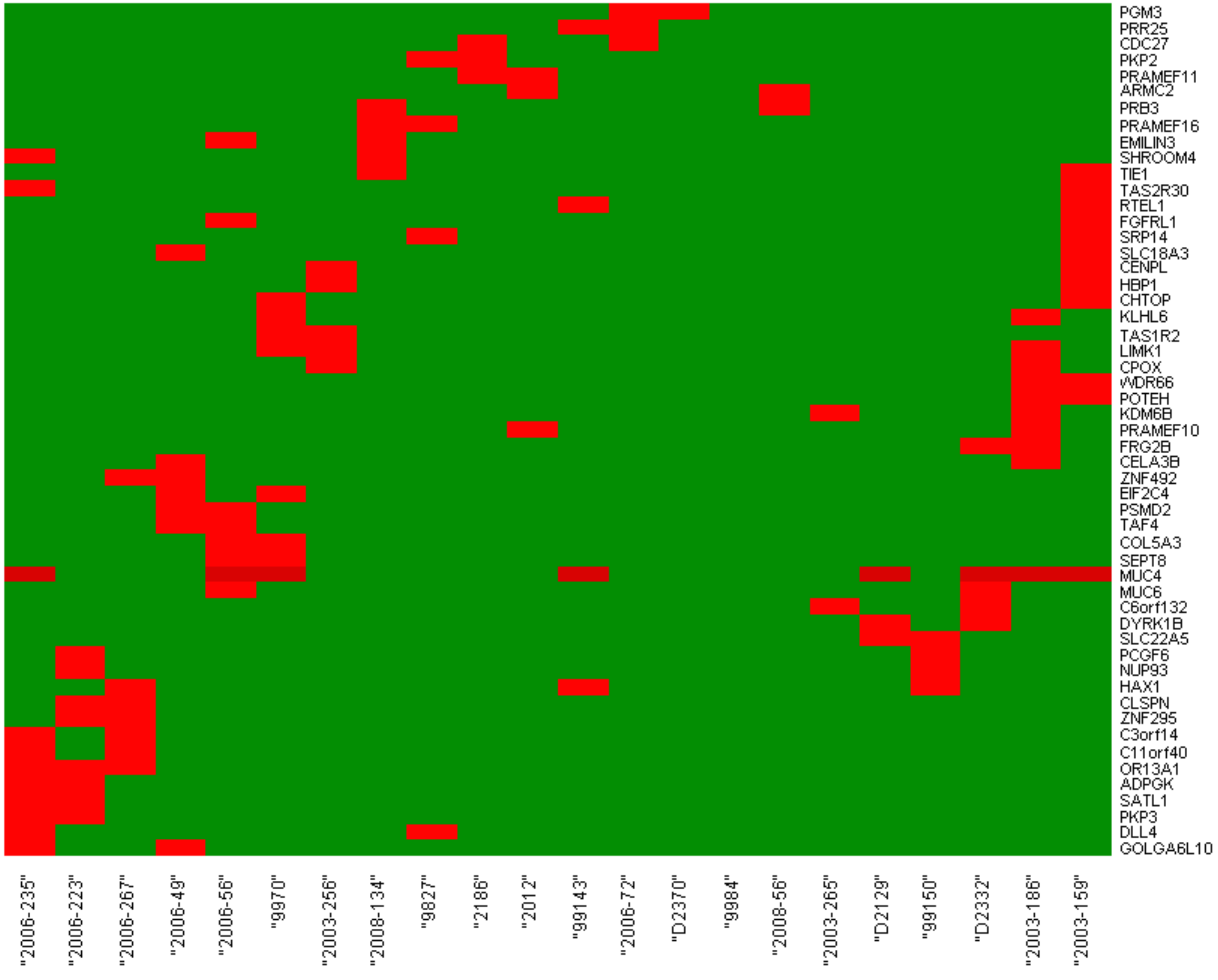
ISTBI 复旦大学类脑智能科学与技术研究院
Institute of Science and Technology for Brain-Inspired Intelligence

Distribution of mutations

53 HRM-specific mutations detected



Distribution of HRM mutations



Candidate driver mutations

$$\max \sum_{j=1}^{n} \frac{c_j x_j}{M} - \frac{\lambda}{K} \cdot \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{c_{ij} x_j x_i}{c_i \cdot K}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} x_j = K, x_j \in \{0,1\}$$

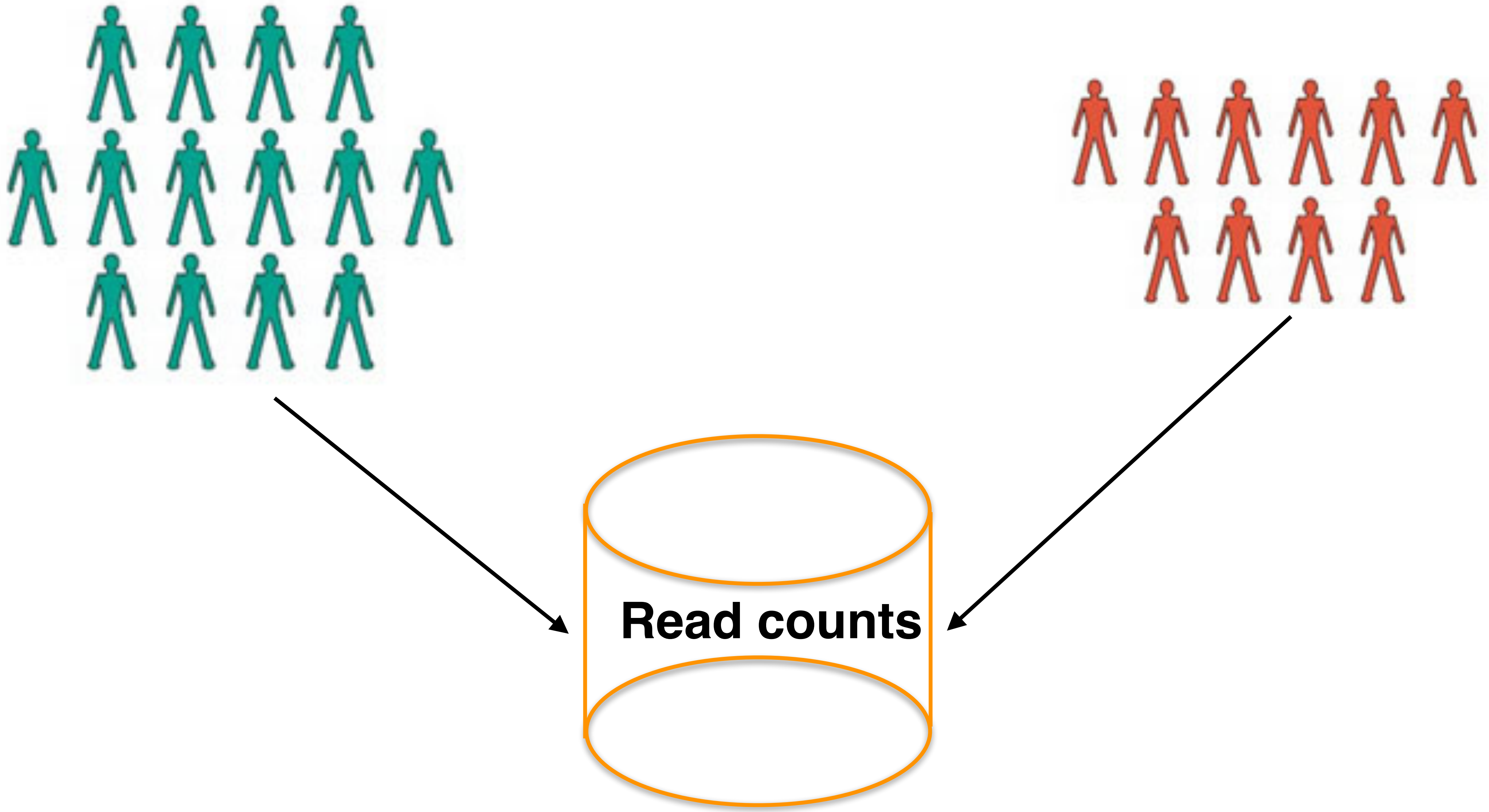
RNA sequencing breast cancers

56 LRM

(low-risk for distant metastasis)

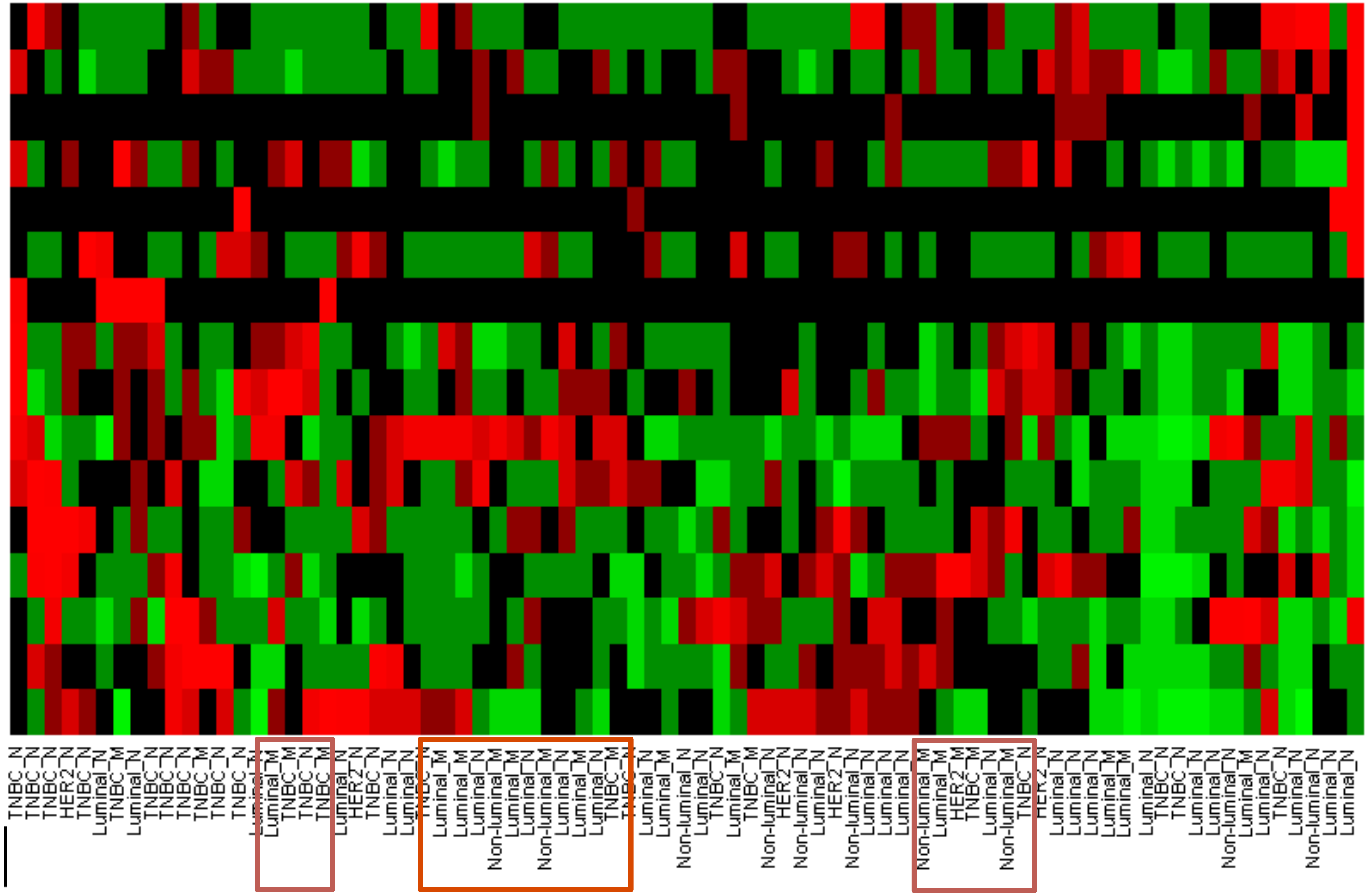
22 HRM

(high-risk for distant metastasis)

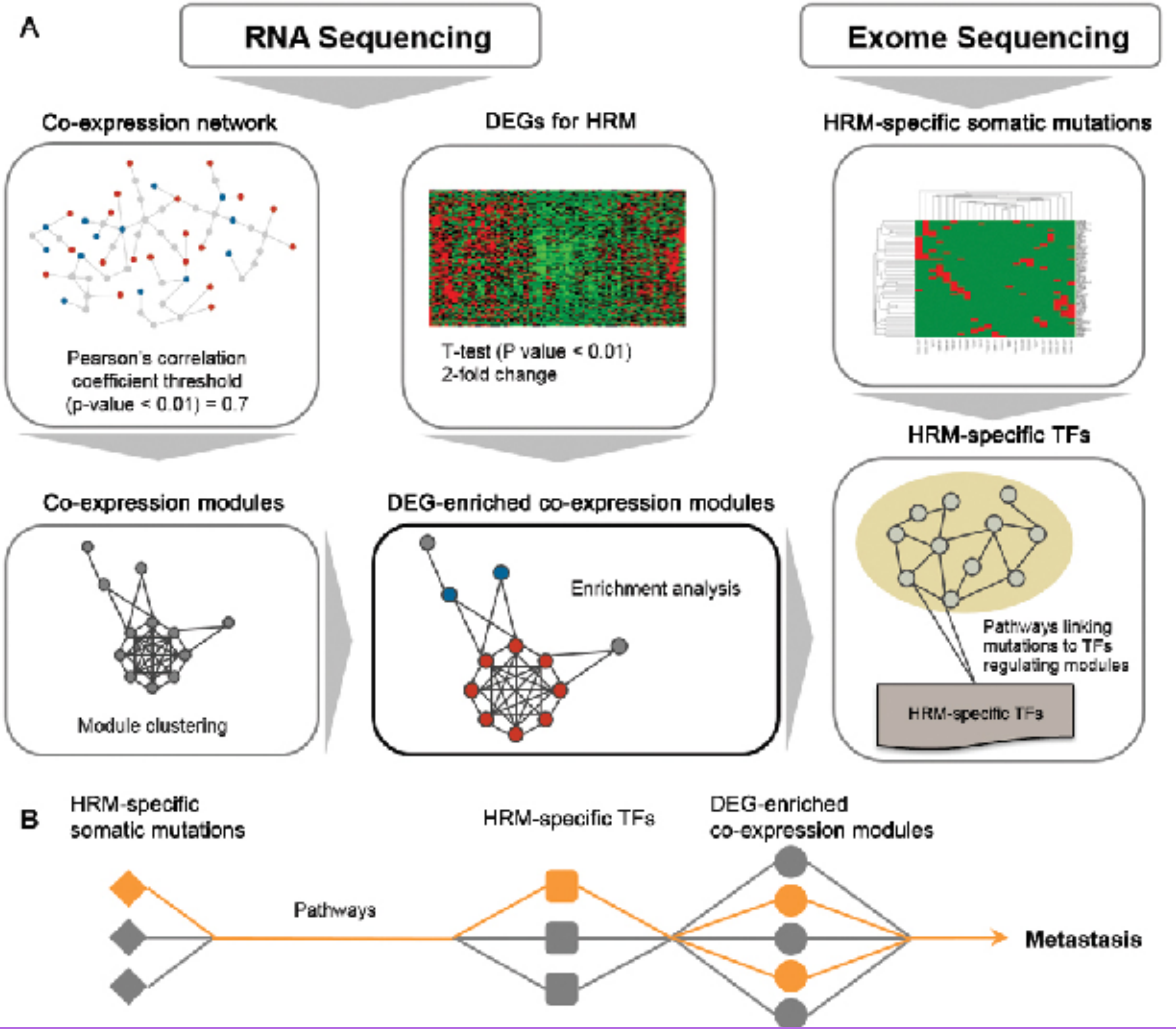


Read counts

Inconsistency between mutations and gene expression



Identification of driver mutations of breast cancer metastasis



Integrative analysis of mutational and transcriptional profiles reveals driver mutations of metastatic breast cancers. ***Cell Discovery***, 2:16025, 2016.

Identification of driver mutations of breast cancer metastasis

$$\underset{\{x_i, y_{ij}, z_{ij}\}}{Max} \sum_{i\in V\cup\{s,t\}} \sum_{j\in V\cup\{s,t\}} w_{ij} y_{ij} - \lambda \sum_{i\in V} \sum_{j\in V} y_{ij} \tag{1}$$

$$s.t. \quad y_{ij} \le x_i, \tag{2}$$

$$y_{ij} \le x_j, \tag{3}$$

$$\sum_{j\in V\cup\{s,t\}} y_{ij} \ge 1, \quad \text{if } i \text{ is } s \text{ or } t, \tag{4}$$

$$\sum_{j\in V\cup\{s,t\}} y_{ij} \ge 2x_i, \quad \text{if } i \text{ is not } s \text{ or } t, \tag{5}$$

$$\sum_{j\in V\cup\{t\}} z_{ij} = R+1, \tag{6}$$

$$\sum_{i\in V\cup\{s,t\}} z_{ij} - \sum_{k\in V\cup\{t\}} z_{jk} = x_j, \quad \text{for } j \in V \cup \{t\}, \tag{7}$$

$$\sum_{i\in V\cup\{s,t\}} z_{ij} \le (R+1) x_j, \quad \text{for } j \in V \cup \{t\}, \tag{8}$$

$$x_i = 1, \quad \text{if } i \text{ is } S \text{ or } TF, \tag{9}$$

$$x_i \in \{0,1\}, \quad i \in V \cup \{s,t\}, \tag{10}$$

$$y_{ij} \in \{0,1\}, \quad i, j \in V \cup \{s,t\}, \tag{11}$$

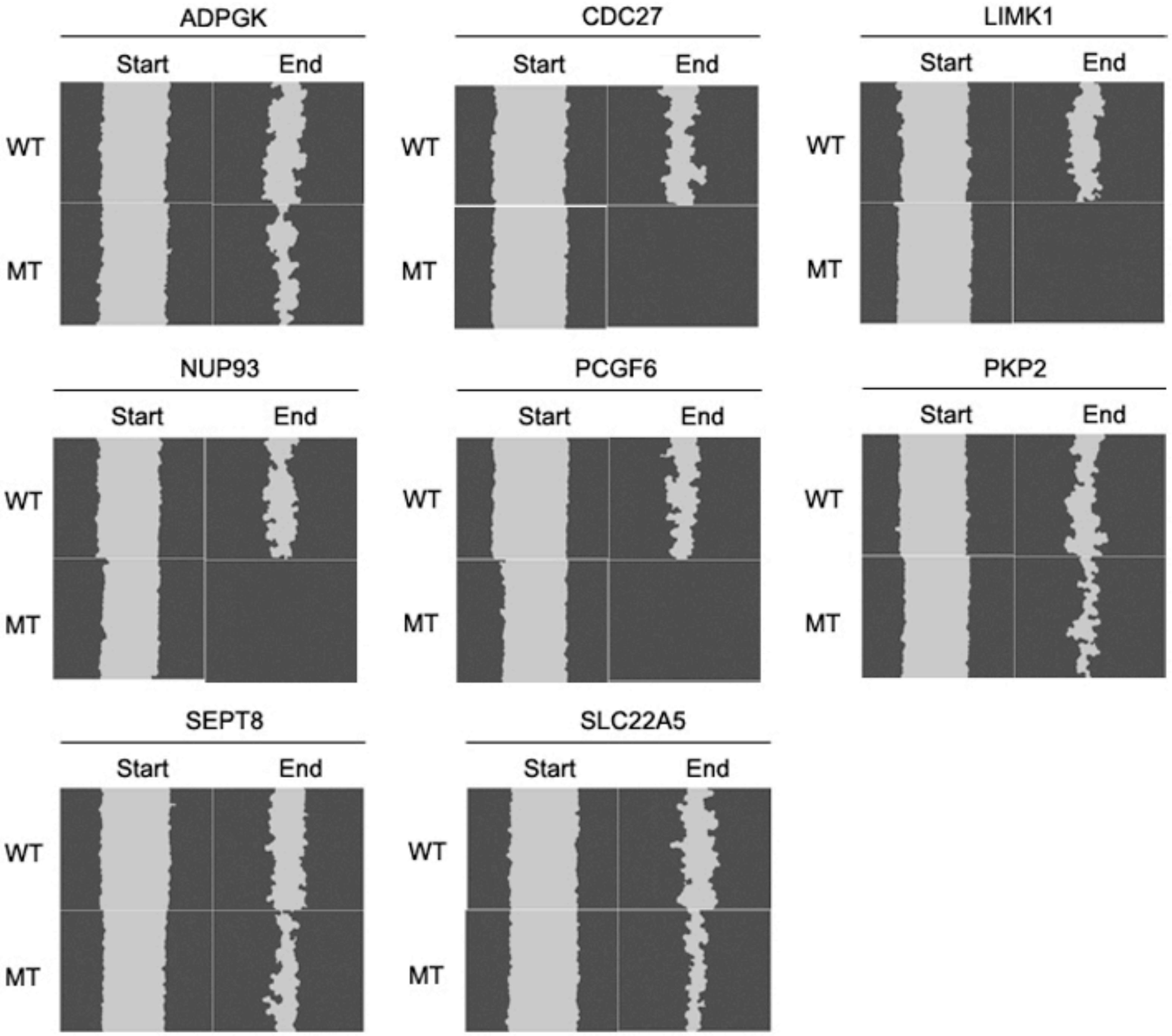
$$z_{ij} \in \{0,1\}, \quad i \in V \cup \{s,t\}, j \in V \cup \{t\} \tag{12}$$

Identification of driver mutations of breast cancer metastasis

Rank	Gene	Target TF	In the list of mathematic model?
1	ADPGK	PAX5	
2	SLC22A5	PAX5	
3	PCGF6	PPARG	Yes
4	DYRK1B	SP1	
5	NUP93	E2F6	Yes
6	PKP2	USF2	Yes
7	LIMK1	PPARG	Yes
8	TIE1	PPARG	Yes
9	CDC27	CEBPA	Yes
10	DLL4	ELK1	Yes
11	HAX1	CEBPA	
12	SEPT8	IRF8	
13	KLHL6	E2F6	

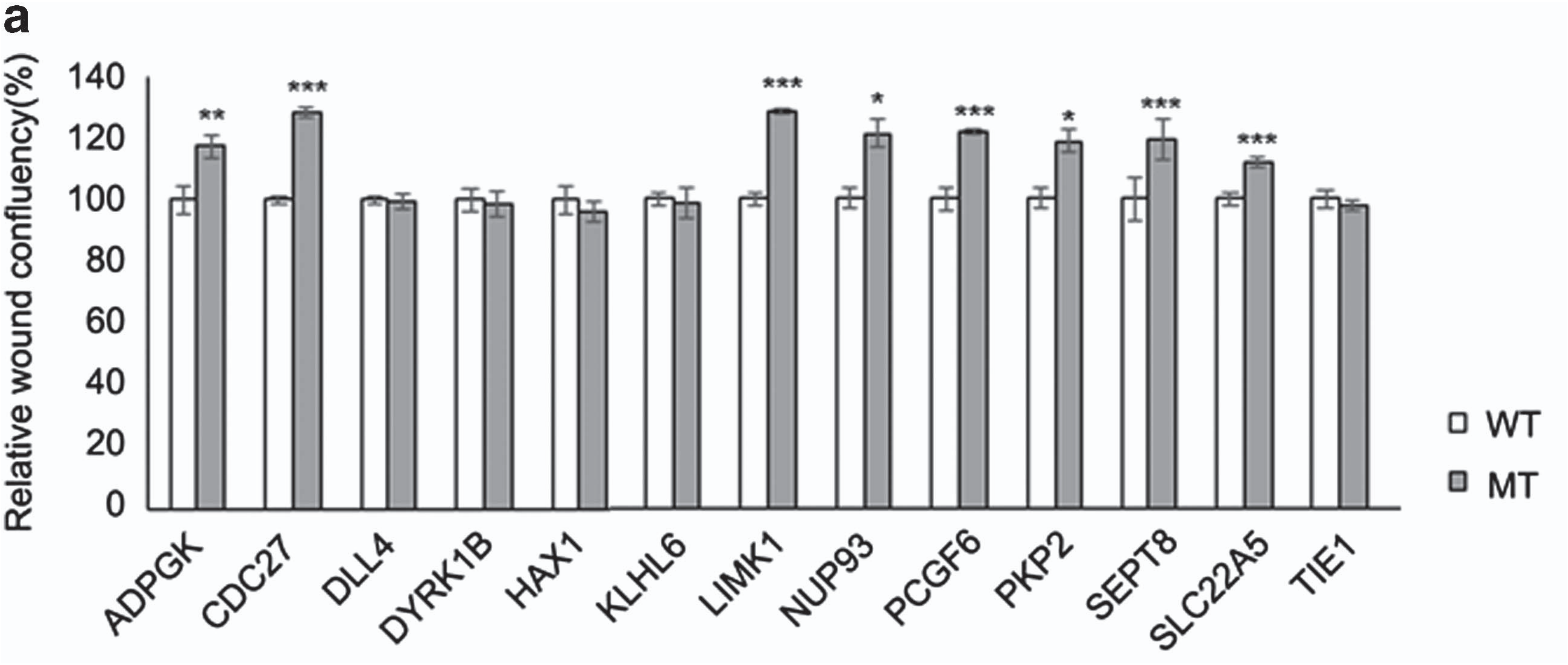
Experimental validation of driver mutations

wound healing assay



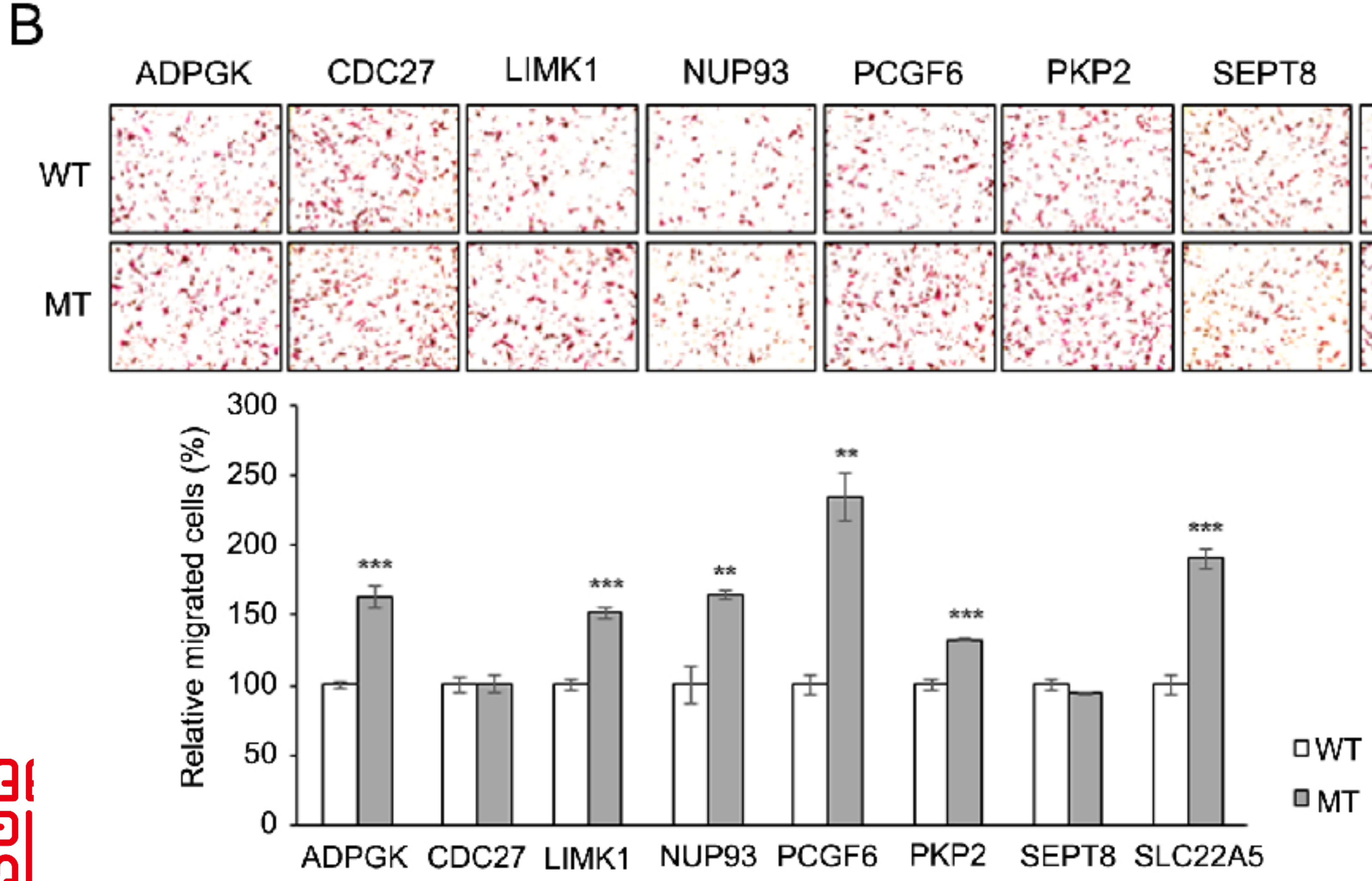
Experimental validation of driver mutations

wound healing assay



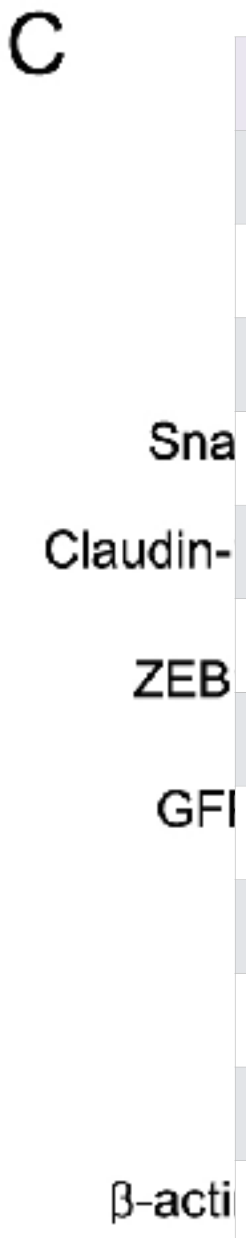
Experimental validation of driver mutations

transwell migration assay



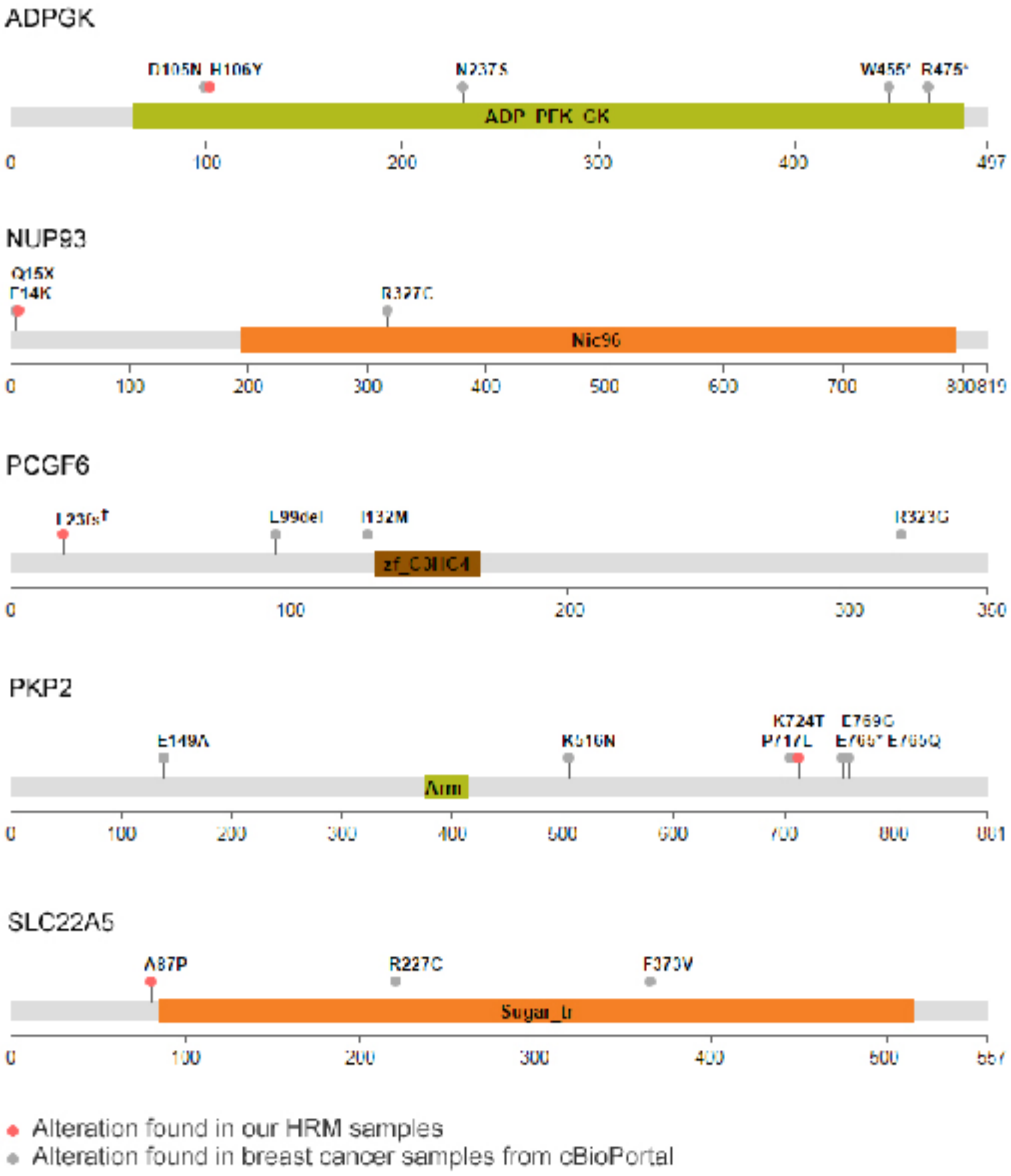
Experimental validation of driver mutations

markers for metastasis

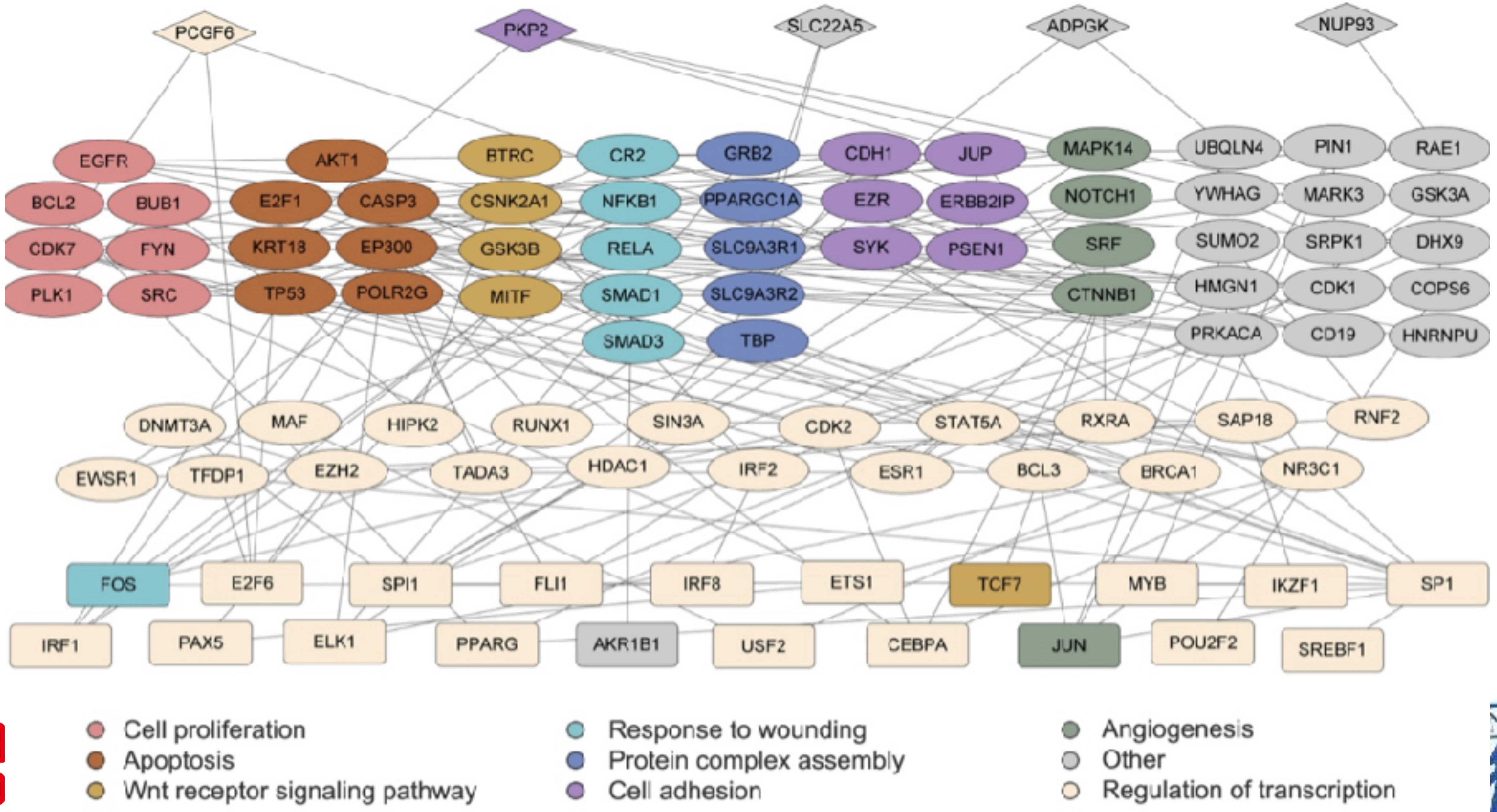


Rank	Gene	Target TF	In the list of mathematic model?
1	ADPGK	PAX5	
2	SLC22A5	PAX5	
3	PCGF6	PPARG	Yes
4	DYRK1B	SP1	
5	NUP93	E2F6	Yes
6	PKP2	USF2	Yes
7	LIMK1	PPARG	Yes
8	TIE1	PPARG	Yes
9	CDC27	CEBPA	Yes
10	DLL4	ELK1	Yes
11	HAX1	CEBPA	
12	SEPT8	IRF8	
13	KLHL6	E2F6	

Validation in other datasets



Potential metastasis driver pathways



Acknowledgements

ZhaoGroup
Yin-Ying Wang

Seoul National University
Sunghoon Kim
Jihyun Lee

Fundings
NSFC
Shanghai Pujiang Talents
Innovation Program of SHEC

ISTBI 复旦大学类脑智能科学与技术研究院
Institute of Science and Technology for Brain-Inspired Intelligence