Received: 7 July 2021 | Revised: 12 November 2021 | Accepted: 21 November 2021

DOI: 10.1111/cas.15227

Cancer Science WILEY

ORIGINAL ARTICLE

High-throughput and multi-phases identification of autoantibodies in diagnosing early-stage breast cancer and subtypes

Rongrong Luo[1] | Cuiling Zheng[1] | Wenya Song[2] | Qiaoyun Tan[2] | Yuankai Shi[2] | Xiaohong Han[3]

[1]Department of Clinical Laboratory, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[2]Department of Medical Oncology, Beijing Key Laboratory of Clinical Study on Anticancer Molecular Targeted Drugs, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[3]Clinical Pharmacology Research Center, Peking Union Medical College Hospital, State Key Laboratory of Complex Severe and Rare Diseases, NMPA Key Laboratory for Clinical Research and Evaluation of Drug, Beijing Key Laboratory of Clinical PK & PD Investigation for Innovative Drugs, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

Correspondence
Xiaohong Han, Clinical Pharmacology Research Center, Peking Union Medical College Hospital, State Key Laboratory of Complex Severe and Rare Diseases, NMPA Key Laboratory for Clinical Research and Evaluation of Drug, Beijing Key Laboratory of Clinical PK & PD Investigation for Innovative Drugs, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 41 Damucang Hutong, Xicheng District, Beijing 100032, China.
Email: hanxiaohong@pumch.cn

Yuankai Shi, Department of Medical Oncology, Beijing Key Laboratory of Clinical Study on Anticancer Molecular Targeted Drugs, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 17 Panjiayuan Nanli, Chaoyang District, Beijing 100021, China.
Email: syuankai@cicams.ac.cn

Funding information
China National Major Project for New Drug Innovation, Grant/Award Number: 2017ZX09304015, 2019ZX09201-002

Abstract

Autoantibodies (AAbs) targeted tumor-associated antigens (TAAs) have the potential for early detection of breast cancer. Here, 574 early-stage breast cancer (ES-BC) patients containing 4 subtypes (Luminal A, Luminal B, HER2+, TN), 126 benign breast disease (BBD) patients, and 199 normal healthy controls (NHC) were separated into three-phases to discover, verify, and validate AAbs. In discovery phase using high-throughput protein microarray, 37 AAbs with sensitivity of 31.25%-86.25% and specificity over 73% in ES-BC, and 40 AAbs with different positive rates between subtypes were identified as candidates. In verification phase, 18 AAbs were significantly increased compared with the Control (BBD and NHC) in focused array. Ten out of 18 AAbs exhibited a significant difference between subtypes ($P < .05$). In ELISA validation phase, 5 novel AAbs (anti-KJ901215, -FAM49B, -HYI, -GARS, -CRLF3) exhibited significantly higher levels in ES-BC compared with BBD/NHC ($P < .05$). The sensitivities of individual AAb and a 5-AAbs panel were 20.41%-28.57% and 38.78%, whereas the specificities were over 90% and 85.94%. Simultaneously, 4 AAbs except anti-GARS differed significantly between TN and non-TN subtype ($P < .05$). We constructed 3 random forest classifier models based on AAbs to discriminant ES-BC from Control or BBD, and to discern TN subtype, which yielded an area under the curve of 0.870, 0.860, and 0.875, respectively. Biological interaction analysis revealed 4 TAAs, except for KJ901215, that were associated with well known proteins of BC.

Abbreviations: AAbs, autoantibodies; AUC, area under the curve; BBD, benign breast disease; BC, breast cancer; BI-RADS, breast imaging reporting and data system; Control, benign breast disease and normal healthy controls; ELISA, enzyme-linked immunosorbent assay; ES-BC, early-stage breast cancer; GO, gene ontology; NHC, normal healthy controls; non-TN, non-triple negative; OD, optical density; PPI, protein-to-protein interaction; RF, random forest; ROC, receiver operation curve; TAAs, tumor-associated antigens; TN, triple negative.

This study discovered and stepwise validated 5 novel AAbs with the potential to diagnose ES-BC and discern TN subtype, indicating easy-to-detect and minimally invasive diagnostic value of serum AAbs ahead of biopsy for future application.

KEYWORDS
autoantibodies, diagnosis, early-stage breast cancer, protein microarrays, subtypes

1 | INTRODUCTION

According to GLOBOCAN 2020, BC has risen to be the most prevalent malignant tumor and remains the first leading cause of cancer-related death in women worldwide.[1] China accounts for 18.4% of the global incidence and 13.2% of the global death. Early diagnosis and treatment are critical to improving BC survival, as the 5-y relative survival rate for 44% of patients with Stage I BC approaches 100%.[2] Mammography is the contemporary detective modality but cannot detect carcinoma grown within normal breast architecture and encounters low specificity in woman with dense breasts.[3,4] Therefore, rapid and cost-effective blood-based biomarkers that could be detected early are fostering the potential as an investigation supplementary to mammography.

Autoantibodies (AAbs) against TAAs have become one of the emerging filed in cancer diagnostic biomarkers for the following advantages: (a) AAbs reflect the information on both cancer cells and immune status; (b) AAbs with higher concentration compared with TAAs are easier to detect for the magnifying generation from cloning B cells and longer half-life period; (c) AAbs produced early during the tumorigenesis are present in serum several months, even years, before clinical diagnosis.[5] Currently, biomarkers reported often in BC are p53, MUC1, and HER2/Neu AAbs. Anti-p53 is the mostly studied with ubiquitous presence in various cancers.[6] Anti-MUC1 is a classic AAb that frequently appears in BC and other cancers and that was reported to show no significant difference between BC patients and controls.[7] Anti-HER2 was found to be increased in BC before and at the time point of diagnosis in a rigorously designated study but with a relatively small cohort.[8] The positive rate of the 3 and other AAbs ranged from 10% to 20%, with a specificity ~90%. It is acknowledged that individual AAbs encounter low sensitivity problems, so that a combinatorial AAbs panel approach is likely to be a better approach for early detection of BC. However, AAbs panels performed variably from different studies even with similar AAbs. AAbs against 7 TAAs (p53, *c*-MYC, HER2, NY-ESO-1, BRCA1, BRCA2, and MUC1) in a European cohort reached a sensitivity of 60% or 45% to distinguish primary BC or ductal carcinoma in situ from controls at 85% specificity.[9] More recently, 6 AAbs in combination (p53, Cyclin B1, p16, p62, 14-3-3ξ, survivin) used to construct discriminant models showed a high sensitivity of 69.5%-78.2% at 64.8%-89.0% specificity in a Chinese cohort.[10] Nevertheless, multiple AAbs studies in BC have focused on a few to a dozen commonly reported markers in various cancers, however few studies have investigated de novo identification of new candidates. Anderson et al[11] used the NAPPA protein array comprising 4988 antigens to discover and validate a 28-AAbs panel with a sensitivity of 80.8% and specificity of 61.6% (AUC = 0.756), but in a small validation cohort.

Additionally, BC is highly heterogeneous for the expression difference of estrogen receptor (ER), progesterone receptor (PR), and epidermal growth factor receptor ERBB2/HER2, which are now divided into 4 accepted subtypes: Luminal A, Luminal B, HER2-enriched, and basal-like (most are triple negative).[12,13] There are still no validated serum biomarkers to characterize different subtypes and only rare studies have investigated AAbs in some specific subtype.[14]

In this study, we utilized a comprehensively high-throughput protein microarray containing ~20 K proteins from the human proteome to survey novel AAbs to diagnose ES-BC and characterize subtypes. After screening and verification in 329 ES-BC samples and 197 control samples, including BBD and NHC, we eventually focused and validated a 5-AAbs panel containing KJ901215, FAM49B, HYI, GARS, and CRLF3 in an independent cohort of 373 samples comprising 245 ES-BC patients and 128 controls in an easy-to-detect ELISA. Furthermore, we established differentiating classifier models based on the 5AAbs for diagnosis of ES-BC and subtyping.

2 | MATERIALS AND METHODS

2.1 | Participants and samples

Patients diagnosed as BC with biopsy confirmation at the Cancer Hospital Chinese Academy of Medical Sciences in Beijing from May 2016 to July 2020, were sequentially included in this study. In total, 574 BC patients were diagnosed as early-stage including Stages 0, IA, IB, IIA, IIB according to the TNM staging system of the *American Joint Committee on Cancer*, 8th edition. Only invasive carcinoma (Stages IA, IB, IIA, and IIB) were included in subtype analysis. All of the ES-BC patients were classified into 4 subtypes based on the expression status of ER, PR, HER2, Ki-67 in immunohistochemical (IHC) staining. To interpret the 4 indices we referred to the *National Comprehensive Cancer Network (NCCN 2020) Clinical Practice Guidelines in Oncology* and *ASCO/CAP guidelines*.[15,16] Ambiguous IHC results for HER2 "2+" were designated to take the *HER2* gene amplification result tested using FISH as the final conclusion. Here, ER+ and/or PR+, HER2−, Ki-67 < 14% were considered as Luminal A, while ER+ and/or PR+, HER2−, Ki−67 > 14%, and ER+ and/or PR+, HER2+ ("3+" in IHC or the *HER2* gene amplified in FISH) were classified as Luminal B. Because the determination

threshold of Ki-67 labeling index varied in different pathological experimental centers, we set the threshold at 14% based on the situation in Department of Pathology at our center and referring to previous studies.[17,18] The HER2+ subtype was ER−, PR−, HER2+, and the triple-negative subtype was ER−, PR−, HER2−. In total, the 126 BBD patients included 118 patients with fibroadenoma and adenosis of the breast and 8 patients with hyperplasia of the mammary glands. All the patients with ES-BC or BBD were treated as naïve, and serum samples were obtained at the time of diagnosis. The NHC classification was given to individuals who had had regular physical examinations with no abnormal laboratory and imaging results. All serum samples were stored at −80℃ and were used with the approval of the Ethics Committee of the National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College (Permission No. 19-019/1804). Waivers of informed consent were requested as the serum samples used in this study were left-over from routine clinical tests.

2.2 | Construction of high-density microarrays and serum profiling assays

High-density microarrays, HuProt™ version 3.0, were provided by CDI Laboratories, Inc. HuProt™ library clones from public opening reading frames (ORFs) or independently synthesized were expressed in proteins with the GST-His6 tag through a yeast expression system.[19] HuProt™ v.3.0 contained 21 888 proteins covering >81% of canonically expressed proteins defined by the *Human Protein Atlas* and 21 888 proteins plus 2304 controls were printed as 24 blocks onto glass slides.

The experiment procedures for AAbs profiling have been described in previous studies.[20] Briefly, microarrays were blocked with 5% BSA diluted in PBS at room temperature for 1.5 h. After discarding the BSA, microarrays were incubated with serum samples diluted with 5% BSA at a 1:1000-fold ratio, for 1 h. After washing, Alexa fluor 647 goat anti-human IgG (Jackson) diluted in 5% BSA at 1:1000-fold ratio was added to microarrays with 0.1% PBS, and incubated at room temperature in darkness for 1 h. After thorough washing with PBST, microarrays were dried naturally and scanned using a GenePix 4000B microarray scanner (Grace Bio-Labs) with a 635 nm excitation laser. GenePix Pro v.6.0 software (Molecular Devices) was used to obtain signal intensities of the foreground signal divided by the background signal (F/B). Positive hits were defined as average signal intensities above the cut-off, set as the mean + 6SD of all the signal points per chip after block correction and *Z*-score normalization.[21]

2.3 | Construction of focused arrays and serum profiling assays

Candidate proteins from HuProt™ selection and the literature were picked to fabricate focused arrays designated as 2 × 7 subarrays by 14-chamber rubber gasket separation. The hybridization process, scanning, and data acquisition were similar to that of the high-density microarray except that the blocking and dilution buffer was changing to 3% BSA.

2.4 | ELISA assay

Recombinant proteins (CDI) with the GST tag were used to detect serum AAbs according to a protocol described in previous studies.[22] Briefly, 50 ng recombinant proteins were coated onto each well of 96-well plates (Corning) at 4℃ overnight. After blocking with 5% skimmed milk for 2 h and washing with 0.2% PBST, 50 μL of each serum sample diluted in 1:100-fold were added and incubated at 37℃ for 1 h. Next, 50 μL 1:20 000-fold diluted peroxidase goat anti-human IgG antibody (Jackson) was added at 37℃ for 1 h, then the chromogenic reaction was conducted at room temperature for 15 min and then the reaction was stopped. Plates were scanned on a Multiskan GO automatic microplate reader (Thermo), OD value of the blank control was subtracted from the OD of each well.

FAM49B, HYI, CRLF3, and GARS proteins were detected using ELISA kits (INS) following the manufacturer's instructions. mRNA data for the 4 proteins were analyzed on the GEPIA website (http://gepia.cancer-pku.cn/).

2.5 | Statistical analysis

Data processing and analysis were carried out in R v.4.0.2 (www.r-project.org). Chi-square (χ^2) test and Fisher exact test were used to compare the positive rates between groups. Wilcoxon test was used to compare levels of AAbs in different groups. Correlations were evaluated using Pearson analysis and Spearman analysis. A *P*-value < .05 was considered statistically significant and could be adjusted using Bonferroni correction when needed. RF modeling was conducted using *caret* package v.6.0-86 (https://github.com/topepo/caret/) after oversampling with the "SMOTE" function in the *DMwR* v.0.4.1 package (http://www.dcc.fc.up.pt/~ltorgo/DataMiningWithR).

3 | RESULTS

3.1 | Study design and objects

In total, 899 sera from 574 ES-BC and 126 BBD patients, and 199 NHS participants were collected to conduct the high-density HuProt™ array, low-density focused array and ELISA detection for novel AAbs discovery, verification, and validation, respectively (Figure 1).

Most patients with ES-BC were over 50 y old, with Stage I and Stage IIA, invasive histological type, negative lymph nodes metastasis, and Luminal A and Luminal B subtypes (Table 1).

3.2 | Discovery of candidate AAbs in high-density protein microarray

In the discovery cohort consisting of 80 ES-BC patients, 20 BBD patients, and 19 NHC participants, the high-density microarray with a linear correlation of 0.93 between parallel duplicates ensured the reproducibility of the detection of serum IgG AAbs (Figure S1A). Using a stringent cut-off Z-score ≥ 6 to determine positive AAbs, 37 differential IgG AAbs candidates were identified for comparisons of ES-BC vs BBD/NHC according to the filtration criteria: Fisher exact test $P < .05$ for ES-BC vs BBD/NHC, positive rate over 30% in ES-BC and specificity ranked at the top. The positive rate of 37 AAbs in patients with ES-BC ranged from 31.25% to 86.25%, while the positive rate of the majority of AAbs was no more than 15% in BBD/NHC (Table S1). In total, 40 AAbs were selected by considering the top-ranking positive rate ratio and $P < .05$ in pairwise comparisons between the ES-BC subgroups (Table S2). Profiles of the differential AAbs showed higher levels in ES-BC compared with NHC, and a relatively higher level in the HER2+ subtype, but the difference was less obvious between subtypes and warranted later validation (Figure 2). Finally, the aforementioned 76 AAbs from inter-group and inner-group comparisons in ES-BC, and other 24 available AAbs formed a 100 AAbs panel for focused array fabrication (Table S3).

3.3 | Verification of AAbs in focused protein microarray

Signals of negative control between ES-BC and Control (BBD + NHC) exhibited a similar distribution ~1.0, indicating neglectable non-specific binding (Figure S1B,C). We expanded samples to contain 249 ES-BC patients (46 Luminal A, 109 Luminal B,

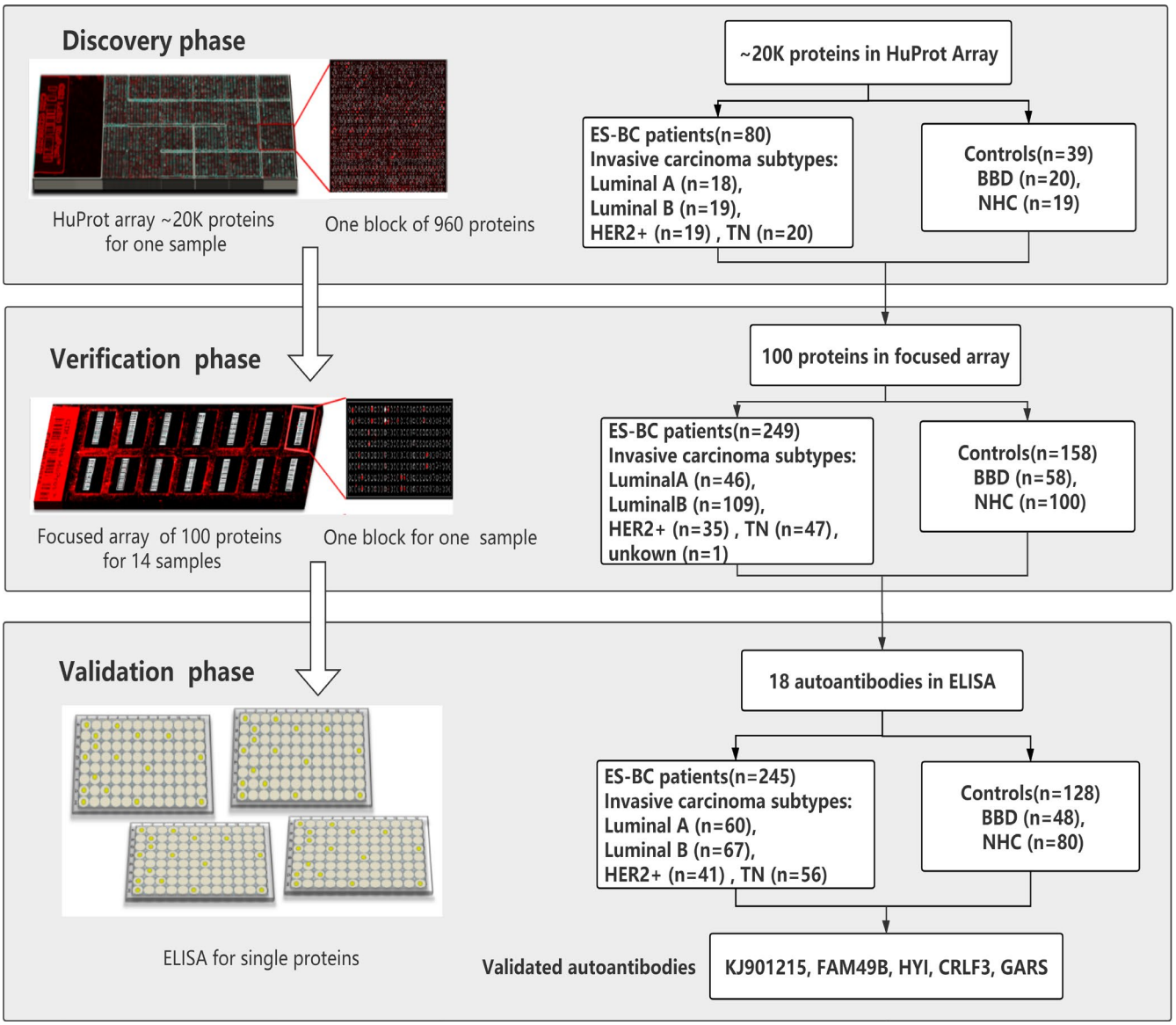


FIGURE 1 Study design. BBD, benign breast disease; ES-BC, early-stage breast cancer; NHC, normal healthy controls

TABLE 1 Clinic characteristics of study subjects

	Discovery phase			Verification phase			Validation phase		
Variable	**ES-BC**	**BBD**	**NHC**	**ES-BC**	**BBD**	**NHC**	**ES-BC**	**BBD**	**NHC**
Number	80	20	19	249	58	100	245	48	80
Female, n (%)	20 (100)	20 (100)	19 (100)	249 (100)	58 (100)	100 (100)	245 (100)	48 (100)	80 (100)
Age, y									
Mean ± SD	51.8 ± 9.0	43.1 ± 13.0	43.0 ± 11.4	51.1 ± 10.1	42.5 ± 11.0	50.3 ± 9.9	53.2 ± 11.4	51.0 ± 7.4	53.2 ± 7.3
Range	35-72	19-71	21-69	24-77	17-75	24-75	28-89	36-77	35-71
Subtypes of invasive cancer									
Luminal A, n (%)	18 (23.7)			46 (19.3)			60 (26.8)		
Luminal B, n (%)	19 (25)			109 (45.8)			67 (29.9)		
HER2+, n (%)	19 (25)			35 (14.7)			41 (18.3)		
TN, n (%)	20 (26.3)			47 (19.8)			56 (25.0)		
Unknown				1 (0.4)					
TNM stage									
0	4 (5)			11 (4.4)			21 (8.6)		
IA	40 (50)			114 (45.8)			125 (51.0)		
IB				1 (0.4)					
IIA	31 (38.8)			88 (35.3)			72 (29.4)		
IIB	5 (6.2)			35 (14.1)			27 (11.0)		
Lymph nodes metastasis, n (%)									
Positive	21 (26.2)			73 (29.3)			40 (16.3)		
Negative	59 (73.8)			176 (70.7)			205 (83.7)		
Histological type, n (%)									
Invasive	76 (95)			238 (95.6)			224 (91.4)		
Noninvasive	4 (5)			11 (4.4)			21 (8.6)		
BI-RADS									
3	1 (1.2)	3 (15)		3 (1.2)	8 (13.8)		5 (2.0)	9 (18.8)	
4a	4 (5.0)	9 (45)		11 (4.5)	12 (20.7)		19 (7.8)	18 (37.5)	
4b	12 (15.0)	1 (5)		26 (10.4)	4 (6.9)		37 (15.1)	16 (33.3)	
4c	12 (15.0)			42 (16.9)	3 (5.2)		73 (29.8)	3 (6.3)	
4	3 (3.8)	7 (35)		15 (6.0)	8 (13.8)				
5	38 (47.5)			121 (48.6)	1 (1.7)		88 (35.9)	2 (4.1)	
6	6 (7.5)			15 (6.0)			23 (9.4)		
NA	4 (5)			16 (6.4)	22 (37.9)				

Abbreviations: BBD, benign breast disease; BI-RADS, breast imaging reporting and data system; ES-BC, early-stage breast cancer; NHC, normal healthy controls.

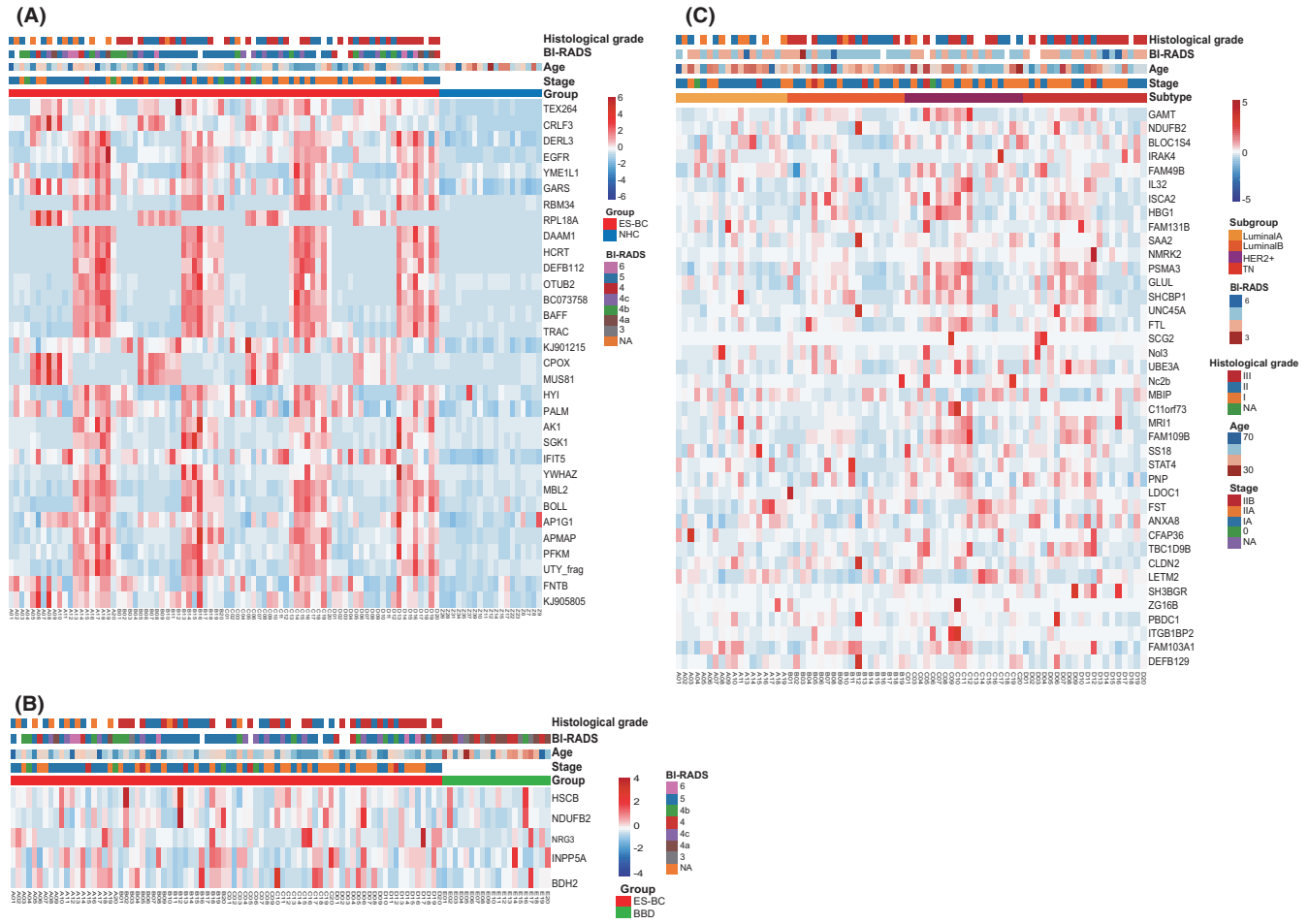


FIGURE 2 Heatmap of differential AAbs in discovery cohort. A-C, Heatmaps of the signal intensities of 32 AAbs in ES-BC vs NHC, 5 AAbs in ES-BC vs BBD, 40 AAbs between subgroups. AAbs, autoantibodies; BBD, benign breast disease; ES-BC, early-stage breast cancer

35 HER2+, 47 TN and 1unknown subtypes), 58 BBD patients, and 100 NHC participants to hybridize with 100 proteins in a focused array. Examples of focused arrays images for ES-BC, BBD, and NHC showed some distinction (Figure 3A). We narrowed the 100 AAbs down to 18 that met the criteria: fold-change (FC), ie, signal intensities of ES-BC divided by that of the Control, over 1.2; stringent cut-off of $P < .0005$ adjusted using the Bonferroni method because of 100 times of Wilcoxon test (Table S4). In total, 18 AAbs were taken as preliminarily validated biomarkers highlighted in a volcano plot (Figure 3B). Levels of the 18 AAbs were significantly higher than those in the Control (Figure 3C). Generally, the majority of the 18 AAbs exhibited a gradually descending trend from ES-BC to BBD to NHC and a significant difference was observed in ES-BC vs NHC groups. Compared with BBD, anti-GARS remained significantly higher in patients with ES-BC (Figure S1D). In comparisons between subtypes and Control shown in Figure 3D, HER2+ vs Control contained the greatest number of differential AAbs (n = 18, left panel bar plot), and 5 AAbs presented simultaneously in 3 comparisons (upper panel bar plot and down panel dot-line intersection). In comparisons between subtypes, Luminal A vs HER2+ contained the greatest number of differential AAbs (n = 10; Figure 3D,E).

3.4 | Validation of AAbs in ELISA

We further conducted validation in ELISA in another independent cohort of 245 ES-BC patients (60 Luminal A, 67 Luminal B, 41 HER2+, and 56 TN), 48 BBD patients, and 80 NHC. We eventually identified 5 novel differential AAbs of anti-KJ901215, anti-FAM49B, anti-HYI, anti-GARS, and anti-CRLF3. Whether it was compared with BBD or NHC or Control separately, the levels of 5 AAbs both significantly increased in ES-BC ($P < .05$; Figure 4A). Using the mean + 2SD of OD in NHC group as cut-off, the OD value of 5 AAbs displayed different distributions above the cut-off value in ES-BC/BBD/NHC and the 5 AAbs hit obviously higher positivity in ES-BC compared with BBD/NHC (Figure 4B,C). The sensitivity of individual AAbs ranged from 20.41% to 28.57% with specificity above 93% to discriminate ES-BC from BBD or NHC (Figure 4D and Table S5). Anti-KJ901215 and anti-CRLF3 had the highest sensitivity of 28.57%, anti-FAM49B and anti-HYI had the highest specificity of 96.09%. Furthermore, we combined 5 AAbs as a panel and defined positivity of panel as at least 1 of the 5 AAbs with greater intensity than the corresponding cut-off values. For the 5-AAbs panel, the sensitivity promoted to 38.78% while the specificity slightly decreased to 87.50% or 85.00% to distinguish

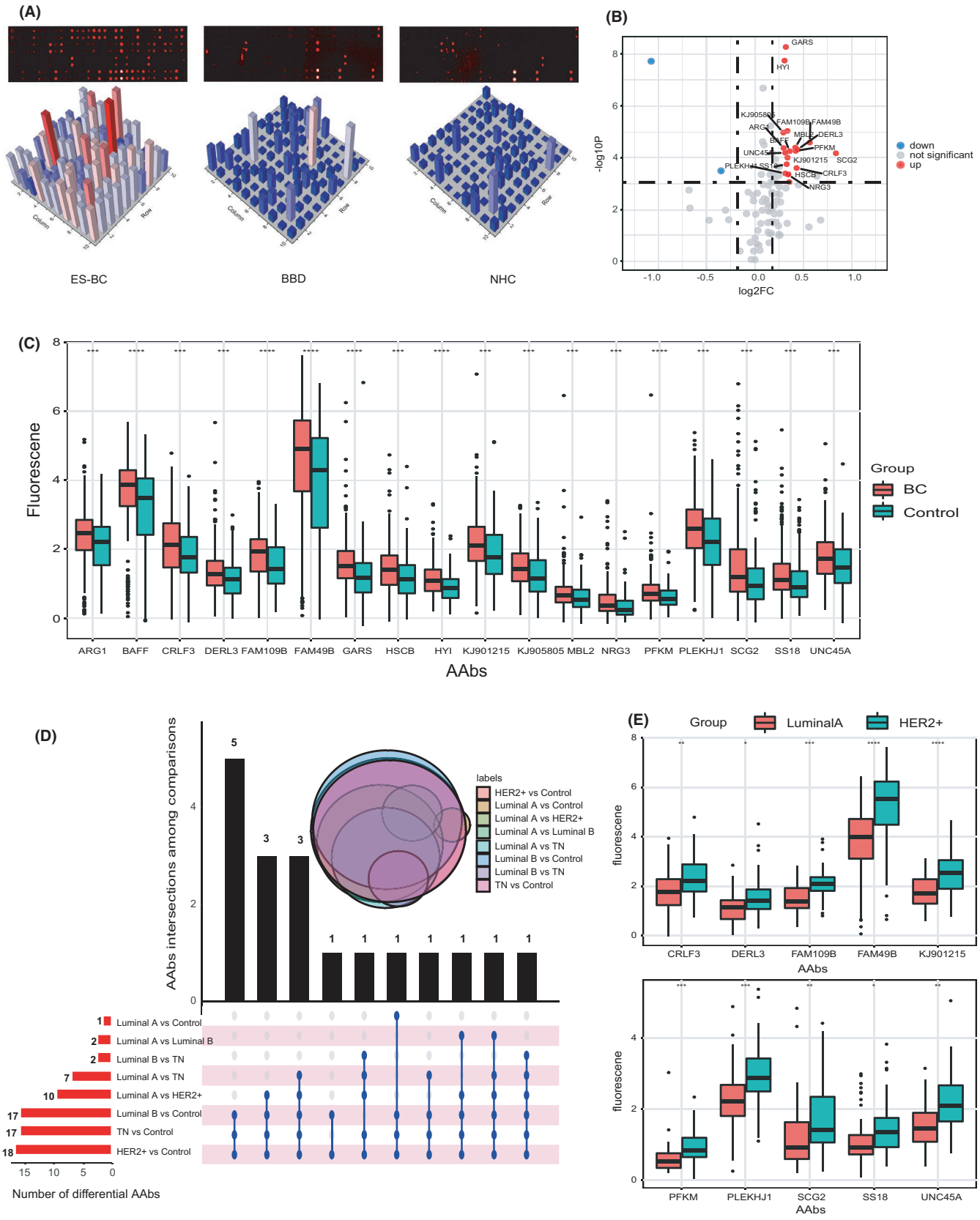


FIGURE 3 AAbs verified using focused array. A, Signal intensities from 1 ES-BC, 1 BBD, and 1 NHC serum incubated with focused array. B, 18 AAbs with $P < .0005$ and FC > 1.2 in ES-BC vs Control (BBD + NHC). C, 18 AAbs exhibited significant higher levels in ES-BC compared with the Control. D, Intersections of significantly differential AAbs in different comparisons. E, 10 AAbs had significant higher levels in HER2+ than Luminal A. *$P < .05$, **$P < .01$, ***$P < .001$, ****$P < .001$. AAbs, autoantibodies; BBD, benign breast disease; ES-BC, early-stage breast cancer; NHC, normal healthy controls

ES-BC from BBD or NHC, respectively (Figure 4E and Table S5). This is a newly found AAbs panel for diagnosing ES-BC apart from reported combinational AAbs in BC. In addition, the BI-RADS category of breast imaging of ES-BC tended to have higher grades (4, 5, and 6) than that in BBD (3 and 4) as expected (Figure 5A). We also observed that the levels of 5AAbs generally increased with the increasing category of BI-RADS, indicating positive relevance between them and underlying the mutually complementary diagnostic value (Figure 5 A). Moreover, the 5 AAbs were not associated with clinical characteristics including age, stage, lymph node invasion, vascular invasion, nerve invasion, and EGFR in IHC, even though ES-BC patients at Stage IIB tended to have higher AAbs levels (Figure 5B).

We then investigated the 5-AAbs profiles in subtypes and BBD/NHC. First, compared with BBD/NHC, Luminal A, Luminal B and HER2+ subtypes had significantly increased levels of the 5 AAbs (P <.05). The 5 AAbs in the TN subtype also tended to increase but without significance (Figure 5C and Table S6). Second, in comparisons between subtypes, Luminal B and HER2+ displayed higher levels, while TN had the lowest levels. Significant differences were observed in anti-KJ901215 for Luminal A vs TN, and in anti-KJ901215, anti-FAM49B, anti-HYI, and anti-CRLF3 for Luminal B vs TN, and in anti-KJ901215, and anti-CRLF3 for HER2+ vs TN ($P < .05$; Figure 5C and Table S6). Therefore, we compared non-TN subtype (Luminal A plus Luminal B plus HER2+) with TN subtype. In total, 4 AAbs showed significantly higher levels ($P < .05$) in non-TN than those in the TN subtype and anti-GARS displayed marginal significance ($P = .050$) between non-TN and TN subtypes (Figure 6A). The positivity of the 5 AAbs in subtypes varied from 14.29% in TN to 36.59% in HER2+, the positive rate of the anti-KJ901215, anti-FAM49B, and anti-CRLF3 subtype was >30% in 1 or more non-TN subtypes (Figure 6B and Table S7).

3.5 | Classifier models and interaction network analysis

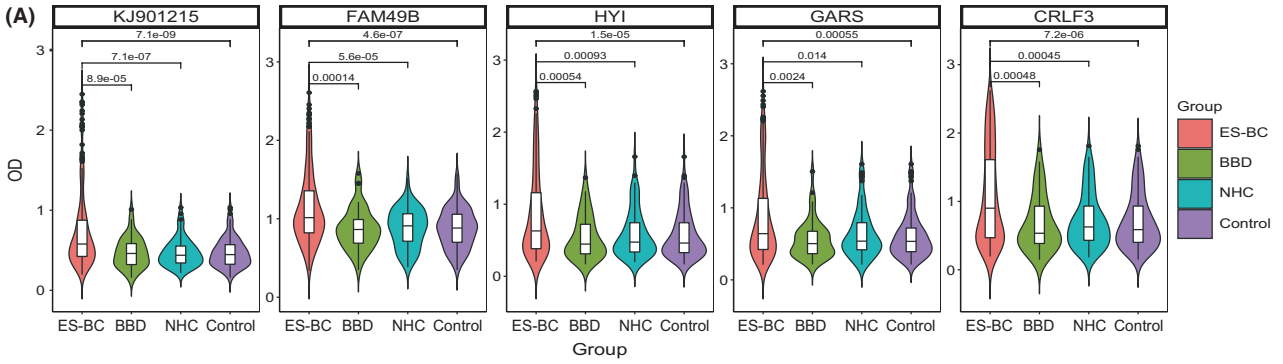
We further established classifier models for diagnosis of ES-BC using RF based on the ELISA cohort. The cohort was randomly partitioned into training set and testing set after oversampling at the ratio of 70%:30%. Ten-fold cross-validation was proceeded in training set, followed by validation in the testing set. First, we built an RF model based on the 5 AAbs, the classifier model displayed a sensitivity of 76.3%, specificity of 86.8% and AUC of 0.870 to discriminant the ES-BC group from the Control (Figure 6C(1)). We then built a classifier model combining BI-RADS and the 5 AAbs in ES-BC and BBD patients. Compared with BI-RADS alone, the model combining the 5 AAbs and BI-RADS yielded an improved AUC from 0.860 to 0.970 to distinguish ES-BC from BBD (Figure 6C(2,3)). Serum AAbs could promote the sensitivity from 74.7% to 85.1% and specificity from 89.6% to 95.8% when the BI-RADS category was combined with the 5-AAbs panel. Furthermore, we constructed an RF model based on the 5 AAbs in the subtype and the model yielded a moderate ability of AUC = 0.875, sensitivity = 82.0%, and specificity = 84.0% for differentiating non-TN and TN patients (Figure 6C(4)).

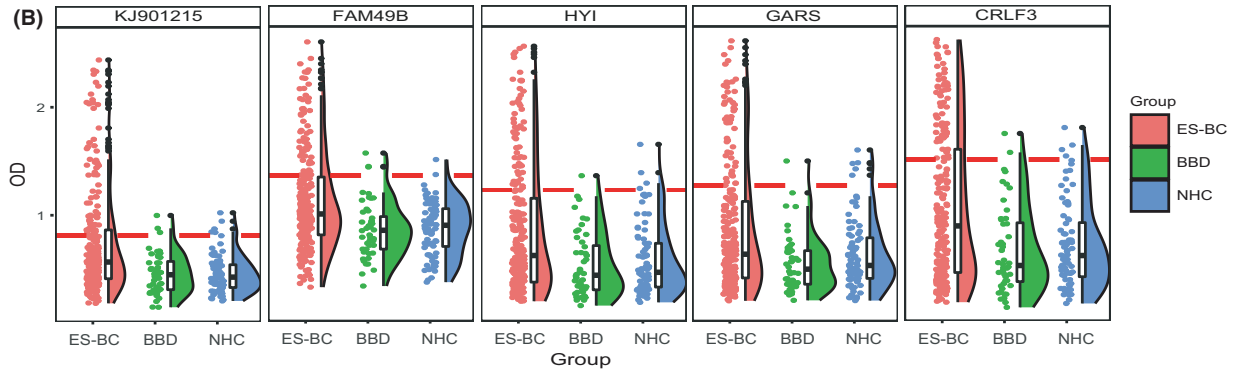
Intriguingly, the 5 AAbs correlated with each other, indicating underlying inner biological relevance (Figure S2A), which inspired the investigation involving an interaction network analysis. First, in total, 215 interactors of 4 corresponding proteins except for KJ901215 without the UniProt ID were explored in the Bio-Grid database and were subsequently used to conduct GO enrichment, KEGG pathways enrichment, and PPI analysis. Aminoacyl-tRNA ligase activity, protein processing in endoplasmic reticulum, and Epstein-Barr virus infection were enriched in the GO biological process (BP) of GARS and its interactors, and central carbon metabolism in cancer and biosynthesis of amino acids were enriched in the GO BP of FAM49B and its interactors (Figure 6D). Top KEGG pathways such as metabolic process, signaling, and cellular component organization or biogenesis were enriched for FAM49B, GARS and CRLF3 (Figure 6E). Finally, the 3 proteins FAM49B, HYI and CRLF3 formed the center, circled by their interactors to give a three-circles PPI network intertwined with each other (Figure 6F). Notably, we uncovered that 5 known proteins in BC, such as BRACA1, CDK2, EGFR, GATD3A, and GSK3B, indirectly interacted with FAM49B, HYI and CRLF3 through corresponding interactors (green nodes) in the 3 circles. A GARS PPI network containing incoming associations (orange edges) and outgoing associations (purple edges) with GARS (GARS1) and different colors for the edges represented different subsets with similar biological functions categorized using the MCODE algorithm. Importantly, 6 proteins, such as BRACA1, CDK2, GATD3A, GSK3B, CTNNA1, and MYC, engaged in BP in BC directly connected to GARS, implying potential relevance of GARS to BC (Figure 6G).

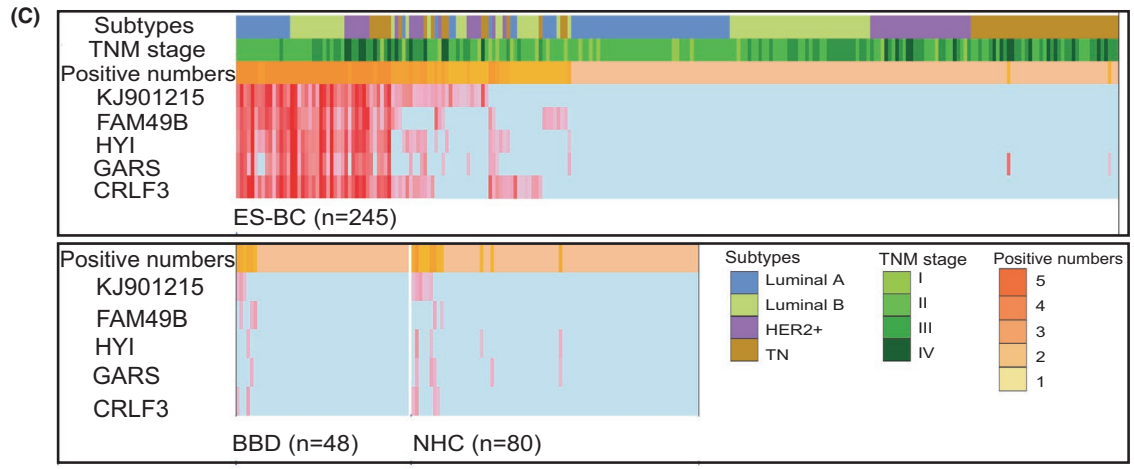
Finally, we found that the distribution of 4 proteins (FAM49B, HYI, CRLF3, and GARS) were consistent with the distribution of AAbs in ES-BC and BBD/NHC/Control, as well as in the non-TN and TN subtypes (Figure S2B,C). A weak, but significant, positive correlation (R:0.12-0.19, $P < .05$) was observed between levels of AAbs and proteins (Figure S2D). Similarly, except for HYI in opposite trend but without significance, the mRNA data for the 3 proteins from tissues in TCGA and the GTEx databases were higher in BC compared with those in the normal controls (Figure S2E).

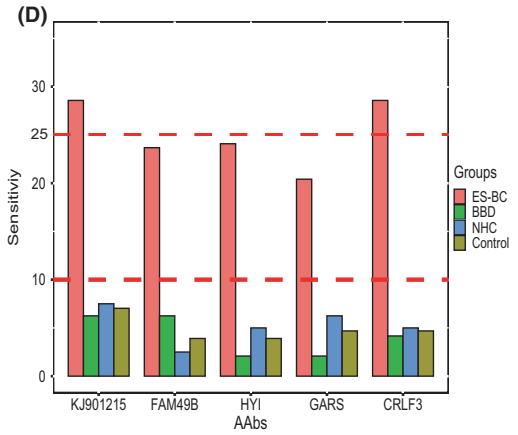
4 | DISCUSSION

The lack of validated serum molecular biomarkers to detect ES-BC more accurately is the utmost challenge to prolong patient survival, therefore it is imperative to discover novel biomarkers.[23] Easily detectable AAbs have been recognized as potential serum biomarkers for the early diagnosis of cancers and high-throughput protein microarrays facilitate novel profiling of AAbs.[20] Taking the small sample size in the discovery phase into consideration, this study utilized another two-phase stepwise verification strategy in an independent cohort comprised of more subjects to profile AAbs in ES-BC. Here, 5 AAbs, namely anti-KJ901215, anti-FAM49B, anti-HYI, anti-GARS,


(A)
KJ901215
FAM49B
HYI
GARS
CRLF3
7.1e-09
7.1e-07
8.9e-05
4.6e-07
5.6e-05
0.00014
1.5e-05
0.00093
0.00054
0.00055
0.014
0.0024
7.2e-06
0.00045
0.00048
OD
ES-BC BBD NHC Control
Group
Group
ES-BC
BBD
NHC
Control


(B)
KJ901215
FAM49B
HYI
GARS
CRLF3
OD
ES-BC BBD NHC
Group
Group
ES-BC
BBD
NHC


(C)
Subtypes
TNM stage
Positive numbers
KJ901215
FAM49B
HYI
GARS
CRLF3
ES-BC (n=245)
BBD (n=48)
NHC (n=80)
Subtypes
Luminal A
Luminal B
HER2+
TN
TNM stage
I
II
III
IV
Positive numbers
5
4
3
2
1


(D)
Sensitivity
KJ901215 FAM49B HYI GARS CRLF3
AAbs
Groups
ES-BC
BBD
NHC
Control

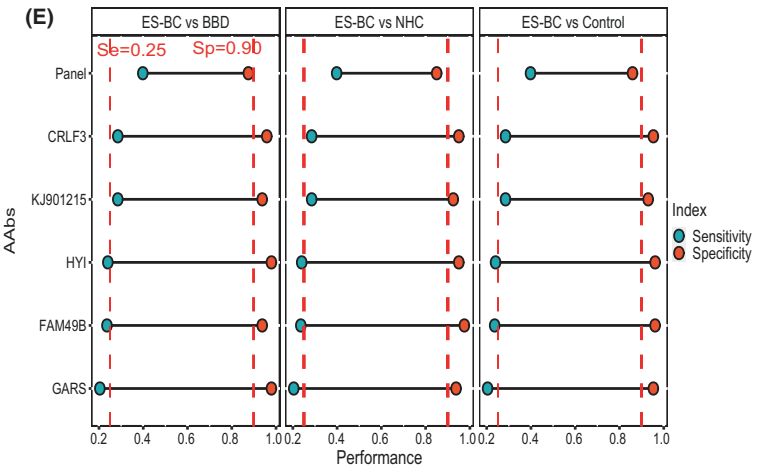

(E)
ES-BC vs BBD
ES-BC vs NHC
ES-BC vs Control
Se=0.25
Sp=0.90
Panel
CRLF3
KJ901215
HYI
FAM49B
GARS
AAbs
Performance
Index
Sensitivity
Specificity

FIGURE 4 Five AAbs validated by ELISA. A, 5 AAbs showed significant difference between ES-BC and BBD/NHC. B, Distribution of 5 AAbs. Red dashed line represented cutoffs set as mean + 2SD of OD value in NHC. C, Heatmaps of positivity of 5 AAbs in ES-BC, BBD, and NHC. D, Positive rate of 5 AAbs in ES-BC, BBD, and NHC. E, Comparisons of the sensitivity and specificity of individual AAb and 5 AAbs panel. AAbs, autoantibodies; BBD, benign breast disease; ES-BC, early-stage breast cancer; NHC, normal healthy controls

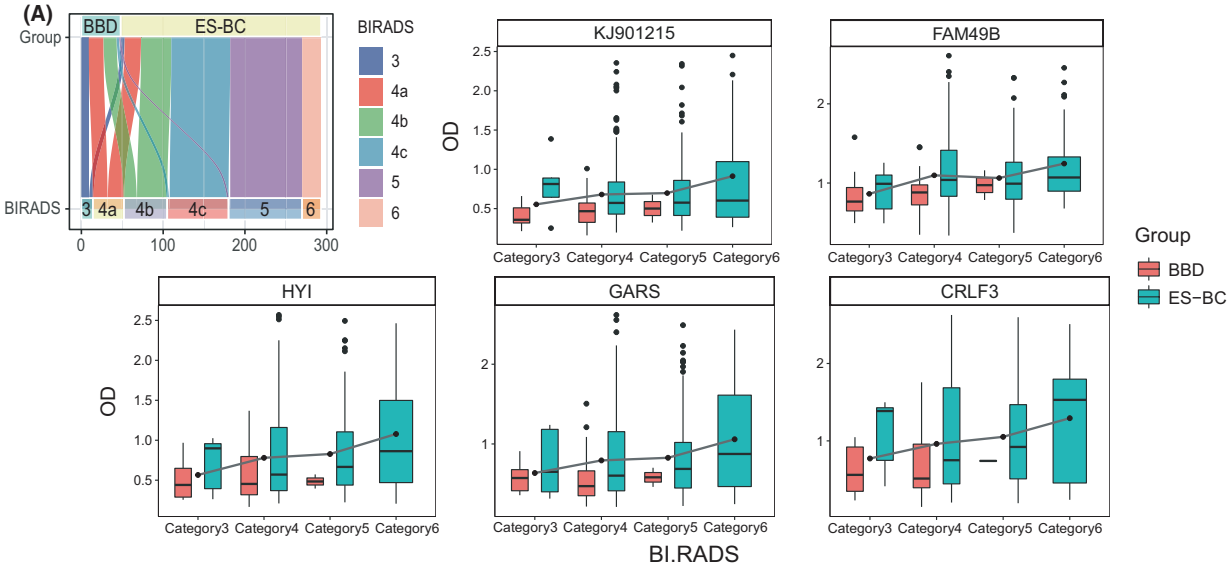
and anti-CRLF3, were found for the first time specifically in ES-BC and were associated with different subtypes. Classifier modeling also unraveled the potential of these 5 AAbs for diagnosing ES-BC and discerning the TN subtype. GO and pathways analyses illustrated their biological relevance to BC. To the best of our knowledge, this is the first study to report these 5 AAbs in ES-BC.

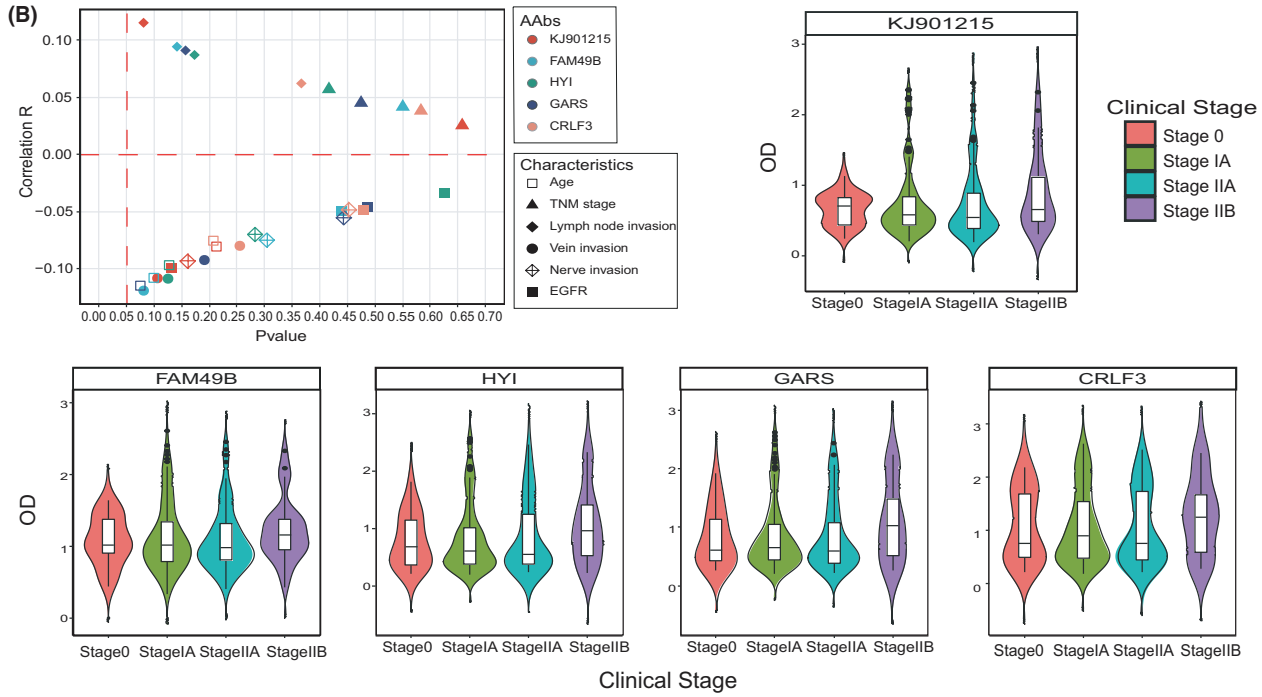
Studies have demonstrated either single or combined AAbs in BC with various results. Previously known onco-proteins such as p53, p16, *c*-myc etc have been mainly investigated using ELISA, and are present universally in many cancers.[24-26] The latest study using ELISA to detect 5 AAbs to common TAAs (p53, RalA, p90, NYESO-1, and HSP70) in a Japanese cohort illustrated a sensitivity of 5%-10% at over 90% specificity for individual AAbs, and 5 AAbs in combination improved the sensitivity to 38%.[24] Few studies have investigated unknown AAbs identified using serologic proteome analysis (SERPA) or serological identification of antigens by recombinant expression cloning (SEREX) or microchip. Such studies were performed without a BBD control, some reported individual AAbs with a sensitivity of ~80% and specificity of 56%-60%,[27] and some demonstrated multiple AAbs with a sensitivity of 35%-66% and specificity of 74%-95% in a small sample size.[28-30] This study used a high-density chip that covered 81% of the ORF of the human genome and gave the opportunity to explore novel AAbs not reported in other cancers[31] Here, 5 validated AAbs showed sensitivities ranging from 20.41%-28.57% and specificities of over 93%, which was not inferior to similar studies.[11,24] Consistent with previous studies, the positive rate of the AAbs panel could be improved to 38.78%. Following the high-throughput discovery phase, this study took advantage of 2 stepwise phase confirmation of the 5 AAbs using tailored protein arrays and ELISA in a larger cohort that focused only on the early stages of BC with the addition of complete controls including both NHC participants and BBD. Additionally, we obtained a moderate value for AUC of 0.870 when discriminating ES-BC from the Control through an RF model using the 5 AAbs. Previous studies have demonstrated a combinational proteomic biomarker assay, "Videssa Breast," that included 15 known AAbs and reported that serum protein biomarkers were detected to establish a model with an AUC of 0.89.[32] Compared with "Videssa Breast" that tested 15 biomarkers simultaneously, our model illustrated the potential of only 5 AAbs for possible future clinical translation. Consistent with a later study using "Videssa Breast" that combined BI-RADS and serum biomarkers, we also obtained an improved AUC of 0.970 compared with BI-RADS (AUC = 0.860),[29] that facilitated better differentiation of ES-BC because of the high specificity of the AAbs that were complementary to mammary imaging.

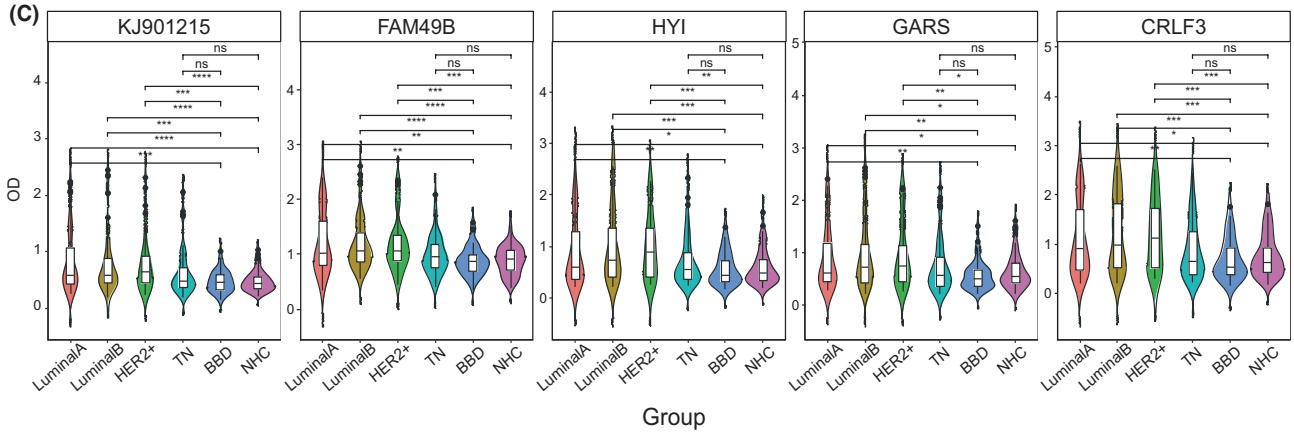
In addition, the 5 AAbs maintained increased levels in subtypes compared with BBD/NHC, implying a positive relationship between the high levels of these AAbs and carcinogenesis. Nevertheless, no individual AAb could differentiate a certain subtype from the others simultaneously because of the high heterogeneity of BC. This was in accordance with other studies on the proteomics of BC.[12,33] Notably, the lower levels of all 5 AAbs, especially anti-KJ901215 and anti-CRLF3, in TN compared with others was probably a result of the "cold" tumor immune milieu in TN.[34] This was significant for its ability to discriminate the TN subtype due to its association with a poor prognosis. Previous studies using the NAPPA microchip with 10 000 antigens identified a 13-AAbs panel to differentiate basal-like BC (most were TN) controls with 33% sensitivity and 98% specificity, but the inestimable effect of therapy on AAbs was an added caveat because only 52% of basal-like BC were pre-treated.[14] Rare studies that profile the AAbs in baseline serum samples in all subtypes at the same time make this study a possible supplement to current understanding. Furthermore, the model distinguished the TN from the non-TN subtype with an AUC of 0.875 expanding and deepening the relevance of the 5 AAbs in future clinical translation.

In light of this new identification of the 5 AAbs in ES-BC, the BP of their corresponding proteins also intrigued us. There has been little evidence on the KJ901215 protein because it has not yet been fully investigated and annotated in public databases. GARS1, a cytosolic enzyme secreted by macrophages, also known as SMAD1, is involved in MEK/ERK/SMAD1 cascade activation, sustaining proliferation in BC cells.[35,36] Higher GARS levels enhanced its chaperone role in neddylation conjugation to the NEDD8 protein, and therefore promoted cellular proliferation via degradation of the cell-cycle inhibitor p27.[37] Increasing levels of anti-GARS were perhaps reactively accompanied with over-expressed GARS

FIGURE 5 Relevance of 5 AAbs to BI-RADS and clinical characteristics, and performance of 5 AAbs in subtypes. A, Distribution of BI-RADS category in ES-BC and BBD in the left-top "Sanky plot," and general ascending trend of 5 AAbs levels along with increasing BI-RADS in boxplots. B, No significant association of AAbs with clinical characteristics. C, Higher levels of the 5 AAbs in subtypes than those in BBD and NHC. AAbs, autoantibodies; BBD, benign breast disease; BC, breast cancer; BI-RADS, breast imaging reporting and data system; NHC, normal healthy controls


(A)
Group
BBD
ES-BC
BIRADS
3
4a
4b
4c
5
6
0
100
200
300
KJ901215
FAM49B
HYI
GARS
CRLF3
OD
Category3
Category4
Category5
Category6
Group
BBD
ES-BC
BI.RADS


(B)
Correlation R
Pvalue
AAbs
KJ901215
FAM49B
HYI
GARS
CRLF3
Characteristics
Age
TNM stage
Lymph node invasion
Vein invasion
Nerve invasion
EGFR
KJ901215
OD
Stage0
StageIA
StageIIA
StageIIB
Clinical Stage
Stage 0
Stage IA
Stage IIA
Stage IIB
FAM49B
HYI
GARS
CRLF3
Clinical Stage


(C)
KJ901215
FAM49B
HYI
GARS
CRLF3
OD
ns
LuminalA
LuminalB
HER2+
TN
BBD
NHC
Group

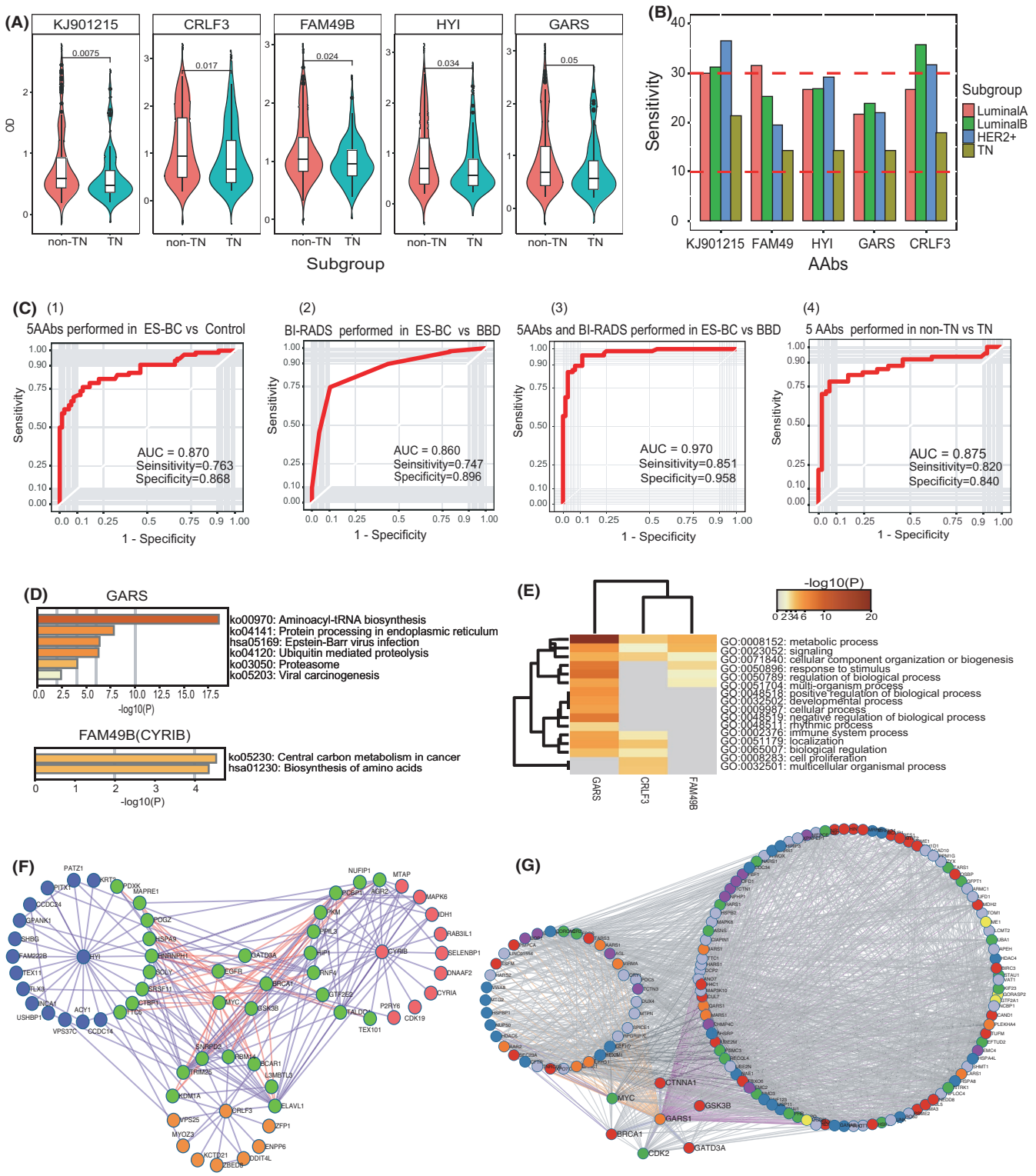


FIGURE 6 Performance of 5 AAbs in subtypes and RF modelling evaluation, and biological interpretation of corresponding TAAs. A, Distribution of 5 AAbs in non-TN subtype (Luminal A, Luminal B and HER2+) and TN subtype. B, Positive rate of 5 AAbs in subtypes. C(1) (2) (3), 5 AAbs or/and BI-RADS had AUCs of 0.870, 0.860, and 0.970 to discriminate ES-BC from Controls, to distinguish ES-BC from BBD patients, using RF modelling and/or ROC analysis. C(4), 5 AAbs with AUC of 0.875 in distinguishing non-TN patients from TN patients using RF modelling. D, GO biological process enrichment of GARS and FAM49B (CYRIB) proteins, and their interactors. E, KEGG pathways enrichment of GARS, CRLF3, and FAM49B, and their interactors. F, PPI networks of HYI, CRLF3, FAM49B. HYI, CRLF3, FAM49B protein encircled by corresponding interactors. Green nodes interacted with the well known proteins of BC. G, PPI network of GARS. Different colors of nodes represented different modules using the MOCODE algorithm. The left and right circle in orange and purple edges represented incoming and outgoing association to GARS. GARS directly linked to the well known proteins of BC in the middle circle. AAbs, autoantibodies; AUC, area under the curve; BBD, benign breast disease; BC, breast cancer; BI-RADS, breast imaging reporting and data system; ES-BC, early-stage breast cancer; NHC, normal healthy controls; RF, random forest; TAAs, tumor-associated antigens; TN, triple negative

in BC cells. A one-sided perception of FAM49B in the past did not include its function, until recent studies uncovered its contradictory function in cancer. FAM49B suppressed cancer cell proliferation via the regulation of mitochondrial fission and actin nucleation dynamics in pancreatic ductal adenocarcinoma or promoted cancer cell proliferation and metastasis by upregulation through the PI3K/AKT pathway in gallbladder cancer, or by inactivating T cells by repressing Rac activity and modulating cytoskeleton reorganization.[38-41] Here, our results of elevated levels of reactive production of anti-FAM49B tended to support its cancer-promoting role in BC. HYI, a putative hydroxy-pyruvate isomerase may be involved in carbohydrate transport and metabolism. Little information is known other than on the non-coding mutations of HYI transcription in melanoma and BC cell lines.[42] CRLF3 as a class I helical cytokine receptor binds erythropoietin-like cytokines to react to injury and physiological challenges through pleiotropic cellular reactions.[43] Increased expression of CRLF3 has only been reported in cutaneous squamous cell carcinoma.[44] AAbs are probably produced as a result of altered or disturbed expression of HYI or CRLF3 caused by unknown mechanisms. Here, metabolic processes, signaling and cellular component organization or biogenesis pathways were enriched by 3 TAAs (FAM49B, GARS, CRLF3) and their interactors were closely related to comprehensive BPs of carcinogenesis. Interestingly, among the listed BP GARS, accumulated evidence has demonstrated that Epstein-Barr virus infection promoted malignant transformation and increased the risk of BC.[45] Finally, interaction networks offered an intuitive view of the intercommunicating links, and revealed the important relationship between 4 target TAAs and well known oncoproteins including BRACA1, CDK2, CTNNA1, EGFR, GATD3A, and GSK3B, implying their close relevance to the development of BC.

Minimally invasive serum AAbs characterization prior to invasive biopsy is meaningful for early diagnosis as it is both more cost effective and a more acceptable method of detection for patients. In addition, use of a comprehensive proteome-wide microchip facilitated the de novo identification of novel AAbs, and stepwise validation ensured the reliability of the AAbs. More importantly, the convenient and rapid, easy and widely accessible ELISA detection suggested potential future clinical translation.

However, serum samples obtained from single center in this retrospective case-control study were subjected to selection bias, small sample size, and lack of in-depth molecular function experiments. Additional efforts are warranted especially in prospective multicenter studies with larger sample sizes to extrapolate the 5 AAbs to the general population and further investigate their biological function.

Altogether, to the best of our knowledge, our discovery and stepwise validation of 5 AAbs is reported here for the first time in ES-BC, and we characterized different subtypes of ES-BC with AAbs. The findings here suggest the potential of the 5 AAbs to distinguish BC at early stages and for the TN subtype, further explanation of their roles in diagnosis is needed in the future.

ACKNOWLEDGMENTS

We thank all the participants in this study and thank Dr. Heng Zhu's laboratory at Johns Hopkins University for providing microarrays. This work was supported by the China National Major Project for New Drug Innovation (2017ZX09304015, 2019ZX09201-002).

DISCLOSURE

The authors have no conflict of interest.

ORCID

Rongrong Luo https://orcid.org/0000-0003-0702-8471
Yuankai Shi https://orcid.org/0000-0002-8685-6744
Xiaohong Han https://orcid.org/0000-0001-9190-0167

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249.
2. Miller KD, Nogueira L, Mariotto AB, et al. Cancer treatment and survivorship statistics, 2019. *CA Cancer J Clin*. 2019;69(5):363-385.
3. Oeffinger KC, Fontham ET, Etzioni R, et al. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *JAMA*. 2015;314(15):1599-1614.
4. Melnikow J, Fenton JJ, Whitlock EP, et al. Supplemental screening for breast cancer in women with dense breasts: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2016;164(4):268-278.
5. Rauf F, Anderson KS, LaBaer J. Autoantibodies in early detection of breast cancer. *Cancer Epidemiol Biomarkers Prev*. 2020;29(12):2475-2485.
6. Sobhani N, D'Angelo A, Wang X, et al. Mutant p53 as an antigen in cancer immunotherapy. *Int J Mol Sci*. 2020;21(11):4087.
7. Burford B, Gentry-Maharaj A, Graham R, et al. Autoantibodies to MUC1 glycopeptides cannot be used as a screening assay for early detection of breast, ovarian, lung or pancreatic cancer. *Br J Cancer*. 2013;108(10):2045-2055.
8. Lu H, Ladd J, Feng Z, et al. Evaluation of known oncoantibodies, HER2, p53, and cyclin B1, in prediagnostic breast cancer sera. *Cancer Prev Res (Phila)*. 2012;5(8):1036-1043.
9. Chapman C, Murray A, Chakrabarti J, et al. Autoantibodies in breast cancer: their use as an aid to early diagnosis. *Ann Oncol*. 2007;18(5):868-873.
10. Qiu C, Wang P, Wang B, et al. Establishment and validation of an immunodiagnostic model for prediction of breast cancer. *Oncoimmunology*. 2020;9(1):1682382.
11. Anderson KS, Sibani S, Wallstrom G, et al. Protein microarray signature of autoantibody biomarkers for the early detection of breast cancer. *J Proteome Res*. 2011;10(1):85-96.
12. Tyanova S, Albrechtsen R, Kronqvist P, et al. Proteomic maps of breast cancer subtypes. *Nat Commun*. 2016;7:10259.
13. Prat A, Perou CM. Deconstructing the molecular portraits of breast cancer. *Mol Oncol*. 2011;5(1):5-23.
14. Wang J, Figueroa JD, Wallstrom G, et al. Plasma autoantibodies associated with basal-like breast cancers. *Cancer Epidemiol Biomarkers Prev*. 2015;24(9):1332-1340.
15. Allison KH, Hammond MEH, Dowsett M, et al. Estrogen and progesterone receptor testing in breast cancer: ASCO/CAP guideline update. *J Clin Oncol*. 2020;38(12):1346-1366.
16. Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists

Clinical Practice Guideline Focused Update. *Arch Pathol Lab Med*. 2018;142(11):1364-1382.
17. Goddard ET, Bassale S, Schedin T, et al. Association between postpartum breast cancer diagnosis and metastasis and the clinical features underlying risk. *JAMA Netw Open*. 2019;2(1):e186997.
18. Valkonen M, Isola J, Ylinen O, et al. Cytokeratin-supervised deep learning for automatic recognition of epithelial cells in breast cancers stained for ER, PR, and Ki-67. *IEEE Trans Med Imaging*. 2020;39(2):534-542.
19. Hu S, Xie Z, Onishi A, et al. Profiling the human protein-DNA interactome reveals ERK2 as a transcriptional repressor of interferon signaling. *Cell*. 2009;139(3):610-622.
20. Li K, Mo W, Wu L, et al. Novel autoantibodies identified in ACPA-negative rheumatoid arthritis. *Ann Rheum Dis*. 2021;80(6):739-747.
21. Pan J, Song G, Chen D, et al. Identification of serological biomarkers for early diagnosis of lung cancer using a protein array-based approach. *Mol Cell Proteomics*. 2017;16(12):2069-2078.
22. Tan Q, Wang D, Yang J, et al. Autoantibody profiling identifies predictive biomarkers of response to anti-PD1 therapy in cancer patients. *Theranostics*. 2020;10(14):6399-6410.
23. Li J, Guan X, Fan Z, et al. Non-invasive biomarkers for early detection of breast cancer. *Cancers (Basel)*. 2020;12(10):2767.
24. Sumazaki M, Ogata H, Nabeya Y, et al. Multipanel assay of 17 tumor-associated antibodies for serological detection of stage 0/I breast cancer. *Cancer Sci*. 2021;112(5):1955-1962.
25. Sullivan FM, Mair FS, Anderson W, et al. Earlier diagnosis of lung cancer in a randomised trial of an autoantibody blood test followed by imaging. *Eur Respir J*. 2021;57(1):2000670.
26. Zhang S, Liu Y, Chen J, et al. Autoantibody signature in hepatocellular carcinoma using seromics. *J Hematol Oncol*. 2020;13(1):85.
27. Dong X, Yang M, Sun H, et al. Combined measurement of CA 15-3 with novel autoantibodies improves diagnostic accuracy for breast cancer. *Onco Targets Ther*. 2013;6:273-279.
28. Lacombe J, Mange A, Bougnoux AC, et al. A multiparametric serum marker panel as a complementary test to mammography for the diagnosis of node-negative early-stage breast cancer and DCIS in young women. *Cancer Epidemiol Biomarkers Prev*. 2014;23(9):1834-1842.
29. Lourenco AP, Benson KL, Henderson MC, et al. A noninvasive blood-based combinatorial proteomic biomarker assay to detect breast cancer in women under the age of 50 years. *Clin Breast Cancer*. 2017;17(7):516-525.e6.
30. Ladd JJ, Chao T, Johnson MM, et al. Autoantibody signatures involving glycolysis and splicesome proteins precede a diagnosis of breast cancer among postmenopausal women. *Cancer Res*. 2013;73(5):1502-1513.
31. Li Y, Li CQ, Guo SJ, et al. Longitudinal serum autoantibody repertoire profiling identifies surgery-associated biomarkers in lung adenocarcinoma. *EBioMedicine*. 2020;53:102674.
32. Henderson MC, Hollingsworth AB, Gordon K, et al. Integration of serum protein biomarker and tumor associated autoantibody expression data increases the ability of a blood-based proteomic assay to identify breast cancer. *PLoS One*. 2016;11(8):e0157692.
33. Mueller C, Haymond A, Davis JB, et al. Protein biomarkers for subtyping breast cancer and implications for future research. *Expert Rev Proteomics*. 2018;15(2):131-152.
34. Chen C, Li A, Sun P, et al. Efficiently restoring the tumoricidal immunity against resistant malignancies via an immune nanomodulator. *J Control Release*. 2020;324:574-585.
35. Ren J, Wang Y, Ware T, et al. Reactivation of BMP signaling by suboptimal concentrations of MEK inhibitor and FK506 reduces organ-specific breast cancer metastasis. *Cancer Lett*. 2020;493:41-54.
36. Nokin MJ, Bellier J, Durieux F, et al. Methylglyoxal, a glycolysis metabolite, triggers metastasis through MEK/ERK/SMAD1 pathway activation in breast cancer. *Breast Cancer Res*. 2019;21(1):11.
37. Rahane CS, Kutzner A, Heese K. Establishing a human adrenocortical carcinoma (ACC)-specific gene mutation signature. *Cancer Genet*. 2019;230:1-12.
38. Zhang Y, Du P, Li Y, et al. TASP1 promotes gallbladder cancer cell proliferation and metastasis by up-regulating FAM49B via PI3K/AKT pathway. *Int J Biol Sci*. 2020;16(5):739-751.
39. Chattaragada MS, Riganti C, Sassoe M, et al. FAM49B, a novel regulator of mitochondrial function and integrity that suppresses tumor metastasis. *Oncogene*. 2018;37(6):697-709.
40. Kaplan E, Stone R, Hume PJ, et al. Structure of CYRI-B (FAM49B), a key regulator of cellular actin assembly. *Acta Crystallogr D Struct Biol*. 2020;76(Pt 10):1015-1024.
41. Shang W, Jiang Y, Boettcher M, et al. Genome-wide CRISPR screen identifies FAM49B as a key regulator of actin dynamics and T cell activation. *Proc Natl Acad Sci USA*. 2018;115(17):E4051-E4060.
42. Zhang W, Bojorquez-Gomez A, Velez DO, et al. A global transcriptional network connecting noncoding mutations to changes in tumor gene expression. *Nat Genet*. 2018;50(4):613-620.
43. Hahn N, Buschgens L, Schwedhelm-Domeyer N, et al. The orphan cytokine receptor CRLF3 emerged with the origin of the nervous system and is a neuroprotective erythropoietin receptor in locusts. *Front Mol Neurosci*. 2019;12:251.
44. Dang C, Gottschling M, Manning K, et al. Identification of dysregulated genes in cutaneous squamous cell carcinoma. *Oncol Rep*. 2006;16(3):513-519.
45. Hu H, Luo ML, Desmedt C, et al. Epstein-Barr virus infection of mammary epithelial cells promotes malignant transformation. *EBioMedicine*. 2016;9:148-160.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Luo R, Zheng C, Song W, Tan Q, Shi Y, Han X. High-throughput and multi-phases identification of autoantibodies in diagnosing early-stage breast cancer and subtypes. *Cancer Sci*. 2022;113:770–783. doi:10.1111/cas.15227