Autoantibody Signatures Discovered by HuProt Protein Microarray to Enhance the Diagnosis of Lung Cancer

Yulin Wang
Zhengzhou University
Jiaqi Li
Zhengzhou University
Xue Zhang
Zhengzhou University
Man Liu
Zhengzhou University
Longtao Ji
Zhengzhou University
Ting Yang
Zhengzhou University
Kaijuan Wang
Zhengzhou University
Chunhua Song
Zhengzhou University
Peng Wang
Zhengzhou University
Hua Ye
Zhengzhou University
Jiangxiang Shi
Zhengzhou University
Liping Dai (lpdai@zzu.edu.cn)
Zhengzhou University https://orcid.org/0000-0001-9384-949X

Research Article

Keywords: lung cancer, diagnostic model, autoantibodies, HuProt protein microarray

Posted Date: December 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1104087/v1

Abstract

Background: This study aims to comprehensively discover novel autoantibodies (TAAbs) against tumor-associated antigens (TAAs) and establish diagnostic models for assisting in the diagnosis of lung cancer (LC) and discrimination of pulmonary nodules (PN).

Methods: HuProt human microarray was used to discover the candidate TAAs and Enzyme-linked immunosorbent assay (ELISA) was performed to detect the level of TAAbs in 634 participants of two independent validation cohorts. Logistic regression analysis was used to construct models. Receiver operating characteristic curve (ROC) analysis was utilized to assess the diagnostic value of models.

Results: Eleven TAAs were discovered by means of protein microarray and data analysis. The level of ten TAAbs (anti-SARS, anti-ZPR1, anti-FAM131A, anti-GGA3, anti-PRKCZ, anti-HDAC1, anti-GOLPH3, anti-NSG1, anti-CD84 and anti-EEA1) was higher in LC patients than that in NC of validation cohort 1 (***P*<0.05**). The model 1 comprising 4 TAAbs (anti-ZPR1, anti-PRKCZ, anti-NSG1 and anti-CD84) and CEA reached an AUC of 0.813 (***95%CI*: 0.762-0.864**) for diagnosing LC from normal individuals. 5 of 10 TAAbs (anti-SARS, anti-GOLPH3, anti-NSG1, anti-CD84 and anti-EEA1) existed a significant difference between malignant pulmonary nodules (MPN) and benign pulmonary nodules (BPN) patients in validation cohort 2 (***P*<0.05**). Model 2 consisting of anti-EEA1, traditional biomarkers (CEA, CYFRA211 and CA125) and 3 CT characteristics (vascular notch sign, lobulation sign, mediastinal lymph node enlargement) could distinguish MPN from BPN patients with an AUC of 0.845 (sensitivity: 58.3%, specificity: 96.6%).

Conclusions: High-throughput protein microarray is an efficient approach to discovering novel TAAbs which could increase the accuracy of lung cancer diagnosis in the clinic.

1. Introduction

More than 2.2 million new lung cancer (LC) cases and 1.8 million deaths were estimated to occur in 2021, which accounts for 12.4% of total new cancer cases and 21.6% of total new death cases, respectively(1). The 5-year survival rate for metastatic disease is 6% while it could be up to 57% for localized LC(2). Low-dose computed tomography (LDCT), as an effective LC screening approach, could significantly reduce the mortality of LC compared to X-ray examination (3, 4). Owing to heavy economic burden and high false positive ratio, its application as a routine diagnostic method for LC was not unrealistic (5, 6). Given the mental and physical pressure on patients and their family, numerous studies are mainly committed to look for a kind of biomarker with lower cost and excellent diagnostic utility which could discriminate LC patients at an earlier stage.

Tumor-associated antigens (TAAs) refer to the aberrantly expressed or mutated proteins which stimulate humoral immune response and the corresponding antibody generated by immune system was known as autoantibodies against TAAs (TAAbs). Because of its stability and specificity in serum, TAAbs emerged as promising biomarkers was extensively investigated.

To date, a great deal of technologies was commonly applied in the identification of novel TAAs, such as serological analysis of recombination cDNA expression libraries (SEREX), serological proteome analysis (SERPA) and protein microarray. Although SEREX has made huge effects on the TAA identification, its application was limited. Since only highly abundant TAAs with linear epitopes can be identified, other epitopes might be omitted. SERPA, combined 2-dimensional gel electrophoresis (2D-GE) with mass spectrometry (MS), has been used to explore TAAs with post-translational modifications and overexpression in the long term. Since 2D-GE was utilized, SERPA was restricted by 2D-GE technology, which can be challenging to obtain reproducible 2D gels (7, 8). Protein microarray, focused on correlations among proteins, peptides on a large scale, accelerated dramatically high-throughput identification of TAAs for the development of tumor biomarkers (7, 9). This approach was widely applied to the identification of autoantibody signatures in a number of cancers (10–14). Our previous study illustrated that a panel of seven TAAbs discovered via focused protein array encoded by 154 cancer driver genes manifested prominent diagnostic capacity with the area under the receiver operating characteristic curve (AUC) of 0.897, the sensitivity of 94.4% and the specificity of 84.9% for the detection of LC (15).

The HuProt human microarray (CDI Laboratories) was originally developed by the Zhu laboratory at Johns Hopkins University and contained 16,368 unique full-length human proteins, representing 12,586 protein-coding genes(16). Pan and his colleagues employed HuProt v3.0 array consisted of 20,240 human full-length proteins to discover expressed differently biomarkers between early LC patients and normal control (NC) (14). In our research, HuProt v3.1 protein microarray based on 21,216 human proteome proteins was applied to screen TAAs with sera from LC patients and normal individuals. After verification of TAAbs by ELISA in different sets of samples, the diagnostic value in LC of TAAbs were evaluated, followed by combining traditional biomarkers and CT parameters to construct models for distinguishing malignant pulmonary nodules (MPN) and benign pulmonary nodules (BPN).

2. Materials And Methods

2.1 Study population and serum collection

This study consisted of 654 serum samples were derived from the Biological Specimen Bank of Henan Key Laboratory of Tumor Epidemiology. The characteristics of all participants were described in **Table 1**. Three independent cohorts (**discovery cohort**, **validation cohort 1 and 2**) were used in this study. In **discovery cohort**, 10 LC serum and 10 NC serum matched by age and gender were detected through protein microarray. Serum samples from 212 LC patients and 212 matched NC, 105 MPN patients and 105 BPN patients were involved in **validation cohort 1** and **2**, respectively.

All LC patients and 105 BPN patients were recruited from November 2016 to November 2019 at the time of initial diagnosis with histopathology, while all patients had not received treatment with any chemotherapy or radiation therapy. Normal individuals were recruited from the physical examination population from May 2019 to June 2019. Five milliliter peripheral blood sample was drawn and separated by centrifugation at 3000 rpm for 5 min and then stored at -80 ℃ for further experiments.

CEA, CYFRA211 and CA125 concentration in serum was detected by electro-chemiluminescence immunoassay (Roche, USA). The 12 characteristic data of CT (number, diameter, edge, spiculation, vascular notch sign, lobulation, spines, pleural indentation, mediastinal lymph node enlargement, emphysema and calcification) were judged by two professional radiologists. The study protocol was approved by the Ethics Committee at Zhengzhou University and all the participants have signed informed consent.

2.2 Huprot protein microarray and gene enrichment analysis

HuProt^{TM} v3.1 protein microarray was purchased from BCBIO technology (Guangzhou, China). Protein microarray was used to screen the candidate autoantibody for the diagnosis of LC. Initially, protein microarray was blocked by 3% BSA at room temperature for 1 h before incubation. Subsequently, the blocked microarray was incubated with 1:200 dilution of serum sample as primary antibody at 4℃ overnight. After washing with PBST, the microarray was incubated with 1:1000 dilution of secondary antibody at room temperature for 1h in the dark. After washing with PBST and ddH_2O, the microarray was dried and then scanned with LuxScan 10K-A (CapitalBio).

The medians of foreground (**F532 Median**) and background (**B532 Median**) intensity of each protein at 532 nm were observed by scanning instrument. The ratio of F532 Median to B532 Median was F/B defined as signal-to-noise ratio (SNR) for the normalization of microarray. The normalization among microarrays was conducted by z-score. The positive ratio of each TAA in LC or NC refers to the ratio of the number whose SNR are higher than 4 (cutoff) to the total in LC or NC group. The analysis method for screening candidate protein was as follow: (a) Mann-Whitney U test, single-tail, $P<0.05$. (b) Fold change (FC: the ratio of median of LC to NC) >1.2. (c) The positive ratio of LC $\geq$50% while that of NC $\leq$10 %.

The Gene Ontology (GO) term and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, based on Metascape software (https://metascape.org/), were applied for the exploration of pathways in which candidate genes of microarray significantly enriched. Adjusted $P < 0.01$ was considered as significantly enriched.

2.3 Indirect Enzyme-Linked Immunosorbent Assay (ELISA)

All commercial purified proteins were purchased from CLOUD-CLONE (Wuhan, China) and CUSBIO technology (Wuhan, China). Indirect ELISA was used to detect the level of autoantibodies against TAAs. The 11 purified recombinant proteins were diluted in coating buffer at optimal concentrations (SARS, ZPR1, FAM131A, GGA3, PRKCZ, HDAC1, GOLPH3, NSG1, DAB1, CD84, EEA1 at final concentrations of 0.125, 0.25, 0.125, 0.25, 0.25, 0.25, 0.25, 0.25, 0.125, 0.25, and 0.25 μg/ml). The detailed experiment was described in the previous study (17-20). All participants' serum samples were randomly dispensed on the plates. Blank control was set on each plate in order to ensure the assay quality.

2.4 Statistical analysis

GenePix Pro 6.0 was applied to obtain the original data from Huprot protein microarray. All data from ELISA was analyzed or visualized by SPSS 25.0, and GraphPad Prism 8.0 software. Non-parametric test was used to compare the levels of TAAbs between LC patients and NC, MPN and BPN patients. The χ^2 test was performed to compare the frequency differences in characteristics of each cohort. Logistic regression analysis was used in the construction of models. ROC analysis was performed to evaluate the capability of each autoantibody and models in the diagnosis of LC and discrimination of PN. The OD value of each autoantibody that shows the highest sensitivity with more than 90% specificity was defined as the cutoff. The cutoff of model was the predicted probability with maximum Youden's Index (YI). The two-tailed P value < 0.05 was statistically significant.

3. Results

3.1 Study design

The design of the present study was illustrated in **Fig.1**. In the discovery phase, Huprot protein microarray was tested in 10 LC and 10 NC samples and 11 TAAbs were identified as selected TAAbs according to the selected criteria. ELISA was then applied to examine the level of selected TAAbs in samples from 212 LC and 212 NC in validation cohort 1 of validation phase, where 10 TAAbs were validated and a diagnostic model with 4 TAAbs and CEA was established. In the following validation cohort 2, 5 out of 10 TAAbs were further validated in 105 MPN and 105 BPN patients, and a model with TAAb and clinical characteristics was constructed for distinguishing malignant and benign PNs.

3.2 Candidate TAAs based on Huprot protein microarray

182 candidate TAAs were screened via protein microarray. The GO and KEGG analysis of 182 geneswas showed in **Figure S1.** We found thatthese genes were closely associated with terms in biological pathways through GO analysis, such as immune system process, biological adhesion, positive regulation of biological process and so on(**Figure S1a**). Besides, these genes revealed enrichment in KEGG analysis mainly relative to natural killer cell mediated cytotoxicity, glycosphingolipid biosynthesis, measles, cell adhesion molecules and so forth (**Figure S1b**).

Eleven TAAbs used for ELISA were selected based on the follow criterion. (a) FC>2, (b) The positive ratio of LC ≥50% while that of NC ≤10 %, (c) The positive ratio of LC - positive ratio of NC ≥60%. The detailed information of 11 TAAs (SARS, ZPR1, FAM131A, GGA3, PRKCZ, HDAC1, GOLPH3, NSG1, DAB1, CD84, EEA1) with **FC>2**, ***P*<0.05** and **positive ratio LC-NC≥60%** were showed in **Table S1. Figure S2** exhibited the specific SNR of each TAAbs in 20 individuals.

3.3 Autoantibodies against TAAs in validation cohort 1

As shown in **Fig.2**, the level of 10 TAAbs (anti-SARS, anti-ZPR1, anti-FAM131A, anti-GGA3, anti-PRKCZ, anti-HDAC1, anti-GOLPH3, anti-NSG1, anti-CD84 and anti-EEA1) showed a significant difference between LC patients and NC (***P*<0.05**). **Fig.3a-3j** displayed the ROC analysis of 10 significant TAAbs in validation cohort 1. Anti-CD84 possessed the highest diagnostic ability with the AUC of 0.693 (0.643-0.744) while anti-

HDAC1 yielded the lowest diagnostic ability with the AUC of 0.579 (0.524-0.634) for LC. The range of sensitivity and specificity were 8.96%-27.4% and 90.1%-99.5%, respectively (**Table 3**). The level of each TAAbs in different group was showed in **Table S2**. There were no significant difference among the expression of each TAAbs in LC patients with different characteristics (***P*<0.05**).

3.4 Establishment and evaluation of the diagnostic model 1

In validation cohort 1, 320 samples (198 NC samples and 122 LC samples) with the result of CEA were selected for the construction of diagnostic model. The AUC of CEA was 0.654 (*95%CI*: 0.587-0.721). Based on these samples, 4 TAAbs panel (anti-ZPR1, anti-PRKCZ, anti-NSG1 and anti-CD84) built by forward logistic regression analysis could distinguish LC patients from NC with an AUC of 0.747 (*95%CI*: 0.691-0.802) (**Fig.3k**). When combined 4 TAAbs with CEA, the AUC of the diagnostic model 1 was up to 0.813 (*95%CI*: 0.762-0.864) with the sensitivity and specificity of 68.9% and 83.8% (**Fig.3k**, **Table 4**). The predicted possibility for diagnosis as LC was PRE (LC, 4TAAbs+CEA)=1/(1+EXP(-(-10.599×anti-ZPR1+4.033×anti-PRKCZ+4.700×anti-NSG1+7.531×anti-CD84+0.276×CEA-3.028))).

For the purpose of evaluating the diagnostic capability of model 1 in LC patients with different clinical character, LC patients were divided into different subgroups according to the clinical characteristics of tumor stage, pathological type, lymph node metastasis (LM) and distant metastasis (DM). **Fig.4** and **Table 4** demonstrated the diagnostic performance of the model in each subgroup. The model could distinguish LC patients from NC in each subgroup. No significant difference was discovered for the diagnostic capability of the model 1 among NSCLC, SCLC, DM+ and DM- patients (***P*>0.05**, **Fig.4a**). To ensure the same specificity in all subgroups (**cutoff: 0.392**), the sensitivity of each subgroup varied from 47.8% to 73.6%, and the accuracy of each subgroup was from 78.1% to 82.6% (**Table 4**).

The diagnostic model owned a higher diagnostic ability in LC patients at advanced stage (AUC: 0.840, 95%*CI*: 0.788-0.892, **Fig 4d**) and with positive lymph node metastasis (AUC: 0.827, 95%*CI*: 0.769-0.885, **Fig.4h**) than LC patients at early stage (AUC 0.695, 95%*CI*: 0.570-0.820, **Fig.4c**) and with negative lymph node metastasis (AUC: 0.770, 95%*CI*: 0.677-0.863, **Fig.4g**). The model 1 showed no significant difference in diagnosing NSCLC patients (AUC: 0.818, 95%*CI*: 0.761-0.878, **Fig.4e**) compared with SCLC patients (AUC: 0.816, 95%*CI*: 0.718-0.915, **Fig.4f**), in DM+ patients (AUC: 0.808, 95%*CI*: 0.740-0.872, **Fig.4j**) compared with DM- patients (AUC: 0.816, 95%*CI*: 0.742-0.890, **Fig.4i**) from NC.

3.5 Autoantibodies against TAAs differentially expressed in MPN and BPN patients

Table 2 presented 12 nodular characteristic of CT in validation cohort 2. All of the CT characteristics were distributed differently between MPN and BPN (P<0.05) except for diameter, edge and emphysema.

In order to understand the performance of TAAbs in MPN and BPN, we selected 7 TAAbs (anti-SARS, anti-FAM131A, anti-PRKCZ, anti-GOLPH3, anti-NSG1, anti-CD84 and anti-EEA1) according to the criterion as ***P*<0.05** and **AUC>0.6 in the validation cohort 1** to be tested by using ELISA in 105 BPN patients and 105 MPN patients of validation cohort 2. The level of 5 TAAbs (anti-SARS, anti-GOLPH3, anti-NSG1, anti-CD84 and anti-EEA1) was significantly higher in MPN patients than that in BPN patients (***P*<0.05**, **Fig.5a**). The

AUC (95%CI) of 5 TAAbs was from 0.580 (95%*CI*: 0.503-0.657) to 0.630 (95%*CI*: 0.555-0.705) and the sensitivity was from 12.4% to 21.9% with a specificity over 90% (**Fig.5b-5f**, **Table 3**).

3.6 Establishment and evaluation of the model in discriminating MPN from BPN

In **validation cohort 2**, 5 significant TAAbs (anti-SARS, anti-GOLPH3, anti-NSG1, anti-CD84, anti-EEA1), 3 traditional biomarkers (CEA, CYFRA211 and CA125) and 9 nodular characteristics of CT (number, cavity, spicule sign, vascular notch sign, lobulation sign, spines, pleural indentation, mediastinal lymph node enlargement, calcification) were applied to establish model used for distinguishing MPN from BPN patients through logistic regression analysis. One hundred eighteen samples (60 MPN samples and 58 BPN samples) with the result of traditional biomarkers and CT were selected for the further research. **Fig.5j** presented that a model (**model 2**) with anti-EEA1, traditional biomarkers (CEA, CYFRA211 and CA125) and 3 nodular characteristics of CT (vascular notch sign, lobulation sign and mediastinal lymph node enlargement) could improve the discrimination capability compared to any single diagnostic method. The model 2 owned the AUC of 0.845 with the sensitivity and specificity of 58.3% and 96.6% for the discrimination of PN (**Fig.5j**, **Table 5**). The predicted possibility for discrimination as MPN was PRE (MPN, TAAbs+3CT+3biomarkers) = 1/(1+EXP(-(+4.497×anti-EEA1+1.264×vascular notch sign +0.921×lobulation sign +0.999×mediastinal lymph node enlargement +0.167×CEA+0.002×CA125+0.203×CYFRA21-1-3.860))).

60 MPN patients were layered by the clinical characteristics of tumor stage, nodular diameter, LM and DM. The discriminating performance of the model 2 for MPN patients with different characteristics was described in **Fig.6** and **Table 5**. The model could discriminate MPN patients in each subgroup from BPN patients (**Fig 6a**). Furthermore, the AUC of model 2 in all subgroups ranged from 0.683 to 0.983 (**Fig.6b-6j, Table5**). The sensitivity of each subgroup was from 26.9% to 94.4% with the same specificity of 96.6% (**Table 5**). To ensure same specificity in all subgroups (**cutoff: 0.392**), the sensitivity varied from 47.8% to 73.6%. Moreover, the accuracy of each subgroup was from 78.1% to 82.6% (**Table 5**).

The model 2 exhibited a higher distinguishing ability in MPN patients at advanced stage (AUC: 0.942, 95%*CI*: 0.893-0.990), LM+ patients (AUC: 0.950, 95%*CI*: 0.910-0.999) and DM+ patients (AUC: 0.983, 95%CI: 0.952-1) than MPN patients at early stage (AUC: 0.713, 95% CI: 0.592-0.834), LM- MPN patients (AUC: 0.683, 95%*CI*: 0.545-0.820) and DM- MPN patients (AUC: 0.731, 95%*CI*: 0.615-0.848) (**Fig.6c-6h, Table 5**). The AUC of patients with more than or equal to 3cm in diameter (AUC: 0.938, 95%*CI*: 0.862-1) was higher than patients with less than 3 cm in diameter (AUC: 0.808, 95%*CI*: 0.722-0.894) (**Fig.6i-6j**, **Table 5**).

4. Discussions

Recently, an increasing number of microarray technology was widely used in a broad range of applications, including biomarker discovery, profiling of immune responses, identification of enzyme substrates, and quantifying protein-small molecule, protein-protein and protein-DNA/RNA interactions(16). A great deal of studies on the basis of protein microarray has confirmed TAAbs in serum that can screen high-risk patients from normal individuals (14, 15). In our research, 11 candidate autoantibodies to TAAs

(SARS, ZPR1, FAM131A, GGA3, PRKCZ, HDAC1, GOLPH3, NSG1, CD84, DAB1 and EEA1) were screened out for the diagnosis of LC by Huprot protein microarray in 20 serum samples. Subsequently, 10 TAAbs (anti-DAB1 was excluded) validated by ELISA were considered as promising biomarkers for the detection of LC and these TAAbs could diagnose LC from healthy individuals with a sensitivity range from 8.96 % to 27.4 % with specificity over 90%. In addition, 5 of 10 TAAbs were proved to distinguish MPN from BPN patients. These results proved that protein microarray is an efficient way to screen novel TAAbs for the detection of LC and MPN patients.

Seryl-tRNA synthetases (SARS) are ubiquitously expressed enzymes that play an essential role in the maintenance of the genetic code by coupling proteinogenic amino acids to their cognate tRNA(s)(21). ZPR1 is a zinc finger protein that binds to the cytoplasmic tyrosine kinase domain of the epidermal growth factor receptor (EGFR) (22) . Furthermore, it has essential roles in embryonic development and cell apoptosis. ZPR1 can translocate to the nucleus during the S phase to regulate the cell cycle progression (23, 24) and cell proliferation (25-27). ZPR1 promotes breast cancer cells invasion and migration via ERK/GSK3β/snail signaling (27). PRKCZ is a member of the PKC family that serves important roles in cell growth, metabolism and other associated signal transduction pathways (28). Positive PRKCZ expression in ADC was associated with cell invasion, lymph node metastasis, and the expression of MMP-2 and MMP-9 (29). NSG1, also designated D4S234E or NEEP21, is an endosomal protein expressed in neuronal cells under normal conditions and a direct transcriptional target gene of the tumor suppressor p53 (30, 31).

CD84, also named SLAM5, is the member of the signaling lymphocyte activation molecule (SLAM) family which contains homophilic and heterophilic receptors that modulate both adaptive and innate immune responses (32, 33). Treatment of human T cells with CD84-Ig enhances TCR-induced IFN-γ secretion presumably through homophilic engagement of cell surface CD84 (34). Early Endosome Antigen 1 (EEA1) is a protein responsible for vesicle budding, transporting, tethering, and docking events in early endosomes, and has also been demonstrated to be involved in PIK3 pathway (35, 36). HDAC1is tightly controlled by a balance between the opposing activities of histone acetyltransferases and histone deacetylases(37). Overexpression of HDAC was found in several cancers (37-39)and thus influences their development and treatment. HDAC1 knockdown inhibits invasion and induces apoptosis in non-small cell lung cancer cells (40).

Among the above TAAbs, only anti-NSG1 was proved to be highly expressed in LC (14) and esophageal squamous cell carcinoma (ESCC) patients (41) while the others have not reported in the serological marker of LC. It exhibited an AUC of 0.784 with sensitivity of 30.0% and 95.5% in the diagnosis of ESCC (41). The sensitivity and specificity of anti-NSG1 was 27.6% and 91.5% in discovery stage and 29.5% and 92.3% in validation stage with the same AUC of 0.740 in diagnosing LC patients. In our study, it performed an AUC of 0.639 with the sensitivity of 23.6% and the specificity of 91.5% in the detection of LC. Ten TAAbs (anti-SARS, anti-ZPR1, anti-FAM131A, anti-GGA3, anti-PRKCZ, anti-HDAC1, anti-GOLPH3, anti-NSG1, anti-CD84 and anti-EEA1) could be considered as promising biomarkers for the diagnosis of LC with the sensitivity of 8.96%-27.4% and the specificity of 90.1%-94.3% .

Numerous studies have proved that combination of markers from different levels could improve significantly the accuracy of diagnostic test (4, 18, 19). Logistic regression is a common method used for the construction of diagnostic model and the combination of multiple level markers in various cancers (10, 17, 18). In our previous study, we employed logistic regression to combine CEA and anti-EGFR for gaining a higher diagnostic capability (17). The AUC of CEA and anti-EGFR is 0.681 and 0.703, respectively, while the combination of CEA and anti-EGFR reached an AUC of 0.784(17). Similarly, logistic regression was utilized to improve the diagnostic capability of 4 TAAb panel (anti-ZPR1, anti-PRKCZ, anti-NSG1 and anti-CD84) with AUC of 0.747 in validation cohort 1 in our study. When combined with CEA, the AUC of combination is up to 0.813 (95%CI: 0.762-0.864). Moreover, the model with four TAAbs and CEA could distinguish LC patients at early stage from NC with the AUC of 0.695.

CT, an efficient radiological technology, was proved to reduce lung cancer-related mortality effectively(3). Increasing models based on CT indicators and serum biomarkers with excellent diagnostic performance were used for the differentiation of PN (4, 42, 43). Hence, we established a model with an AUC of 0.845 (95%CI: 0.775-0.914) comprised of anti-EEA1, CEA, CYFRA211, CA125, vascular notch sign, lobulation sign and mediastinal lymph node enlargement to differentiate between MPN and BPN patients. It not only showed a high AUC in early LC detection, but also exhibited a better performance for advanced and LM MPN patients. The excellent diagnostic performance of the model in different subgroup indicated its practicability in the clinical application.

Several improvements have been made in the current study. Firstly, the protein microarray used for discovery cohort is one of the world's largest collections of full-length human proteins, covering 81% of the human proteome. Subsequently, in contrast to Jiang's (15) and Pan's research (14), we not only tested the level of candidate TAAbs in NC and LC patients, but also validated their expression in MPN and BPN patients of validation phase. Last but not least, we combined TAAbs with traditional biomarkers and CT indicators in order to establish economical models applied in the diagnosis of LC and the discrimination of PN for alleviating economic burden of patients. These models manifest prospective diagnostic capability in LC and MPN patients with different clinical features.

The limitation of our study is that these models are not validated in a larger sample size cohort and failed to validate their stability on another independent groups. It is also regretful that we fail to evaluate our model's differential ability in comparison with other models published previously by other researchers.

To sum up, we applied Huprot protein microarray to discovery promising TAAbs for the detection of LC and the discrimination of PN. The established models which own excellent ability in screening LC and MPN patients are expected to perform in clinical applications.

Declarations

Ethics approval and Consent to participate and publication

The study protocol was approved by the Ethics Committee at Zhengzhou University. Written informed consent to participate in this study and for publication was provided by all the participants of the study.

Availability of data and materials

On a reasonable request, the data supporting study's findings are available from the corresponding author.

Conflicts of Interest

The authors have declared that there is no competing interests

Funding

This work was supported by the Leading Talents of Science and Technology Innovation in Henan Province (Grant Number 20420051008), the National Natural Science Foundation of China (Grant Number 8167291), the Major Project of Science and Technology in Henan Province (Grant Number 16110311400), the Key Project of Discipline Construction of Zhengzhou University (Grant Number XKZDQY202009), and the Project of Basic Research Fund of Henan Institute of Medical and Pharmacological Sciences (Grant Number 2020BP0202).

Author's contributions

LD conceived and designed the study. YW conducted most experiments, analyzed data and wrote the manuscript. JL, ML, TY and LJ performed parts of the experiments. KW, CS and PW revised the manuscript, HY and JS provided the assistance for statistical analysis. All the authors listed have approved the manuscript that is enclosed.

Code availability

Not applicable.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, 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):209–49.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70(1):7–30.
3. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395–409.
4. Lin Y, Leng Q, Jiang Z, Guarnera MA, Zhou Y, Chen X, et al. A classifier integrating plasma biomarkers and radiological characteristics for distinguishing malignant from benign pulmonary nodules. Int J Cancer. 2017;141(6):1240–8.

5. Dai L, Li J, Tsay JJ, Yie TA, Munger JS, Pass H, et al. Identification of autoantibodies to ECH1 and HNRNPA2B1 as potential biomarkers in the early detection of lung cancer. Oncoimmunology. 2017;6(5):e1310359.
6. Berrington de Gonzalez A, Mahesh M, Kim KP, Bhargavan M, Lewis R, Mettler F, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. Arch Intern Med. 2009;169(22):2071–7.
7. Tan HT, Low J, Lim SG, Chung MC. Serum autoantibodies as biomarkers for early cancer detection. FEBS J. 2009;276(23):6880–904.
8. Solassol J, Maudelonde T, Mange A, Pujol JL. Clinical relevance of autoantibody detection in lung cancer. J Thorac Oncol. 2011;6(5):955–62.
9. Kijanka G, Murphy D. Protein arrays as tools for serum autoantibody marker discovery in cancer. J Proteomics. 2009;72(6):936–44.
10. Sun G, Ye H, Wang X, Cheng L, Ren P, Shi J, et al. Identification of novel autoantibodies based on the protein chip encoded by cancer-driving genes in detection of esophageal squamous cell carcinoma. Oncoimmunology. 2020;9(1):1814515.
11. Yang Q, Qin J, Sun G, Qiu C, Jiang D, Ye H, et al. Discovery and Validation of Serum Autoantibodies Against Tumor-Associated Antigens as Biomarkers in Gastric Adenocarcinoma Based on the Focused Protein Arrays. Clin Transl Gastroenterol. 2020;12(1):e00284.
12. Qiu C, Wang B, Wang P, Wang X, Ma Y, Dai L, et al. Identification of novel autoantibody signatures and evaluation of a panel of autoantibodies in breast cancer. Cancer Sci. 2021.
13. Cui C, Duan Y, Qiu C, Wang P, Sun G, Ye H, et al. Identification of Novel Autoantibodies Based on the Human Proteomic Chips and Evaluation of Their Performance in the Detection of Gastric Cancer. Front Oncol. 2021;11:637871.
14. Pan J, Song G, Chen D, Li Y, Liu S, Hu S, et al. Identification of Serological Biomarkers for Early Diagnosis of Lung Cancer Using a Protein Array-Based Approach. Mol Cell Proteomics. 2017;16(12):2069–78.
15. Jiang D, Zhang X, Liu M, Wang Y, Wang T, Pei L, et al. Discovering Panel of Autoantibodies for Early Detection of Lung Cancer Based on Focused Protein Array. Front Immunol. 2021;12:658922.
16. Duarte JG, Blackburn JM. Advances in the development of human protein microarrays. Expert Rev Proteomics. 2017;14(7):627–41.
17. Wang Y, Liu F, OuYang S, Liu M, Zhang X, Wang P, et al. Humoral immune response to epidermal growth factor receptor in lung cancer. Immunol Res. 2021;69(1):71–80.
18. Wang T, Liu H, Pei L, Wang K, Song C, Wang P, et al. Screening of tumor-associated antigens based on Oncomine database and evaluation of diagnostic value of autoantibodies in lung cancer. Clin Immunol. 2020;210:108262.
19. Pei L, Liu H, Ouyang S, Zhao C, Liu M, Wang T, et al. Discovering novel lung cancer associated antigens and the utilization of their autoantibodies in detection of lung cancer. Immunobiology. 2020;225(2):151891.

20. Jiang D, Wang Y, Liu M, Si Q, Wang T, Pei L, et al. A panel of autoantibodies against tumor-associated antigens in the early immunodiagnosis of lung cancer. Immunobiology. 2020;225(1):151848.
21. Ibba M, Soll D. Aminoacyl-tRNA synthesis. Annu Rev Biochem. 2000;69:617–50.
22. Galcheva-Gargova Z, Konstantinov KN, Wu IH, Klier FG, Barrett T, Davis RJ. Binding of zinc finger protein ZPR1 to the epidermal growth factor receptor. Science. 1996;272(5269):1797–802.
23. Dyczkowski J, Vingron M. Comparative analysis of cell cycle regulated genes in eukaryotes. Genome Inform. 2005;16(1):125–31.
24. Gangwani L. Deficiency of the zinc finger protein ZPR1 causes defects in transcription and cell cycle progression. J Biol Chem. 2006;281(52):40330–40.
25. Galcheva-Gargova Z, Gangwani L, Konstantinov KN, Mikrut M, Theroux SJ, Enoch T, et al. The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus of proliferating cells. Mol Biol Cell. 1998;9(10):2963–71.
26. Gangwani L, Mikrut M, Galcheva-Gargova Z, Davis RJ. Interaction of ZPR1 with translation elongation factor-1alpha in proliferating cells. J Cell Biol. 1998;143(6):1471–84.
27. Gangwani L, Flavell RA, Davis RJ. ZPR1 is essential for survival and is required for localization of the survival motor neurons (SMN) protein to Cajal bodies. Mol Cell Biol. 2005;25(7):2744–56.
28. Wu J, Zhang B, Wu M, Li H, Niu R, Ying G, et al. Screening of a PKC zeta-specific kinase inhibitor PKCzI257.3 which inhibits EGF-induced breast cancer cell chemotaxis. Invest New Drugs. 2010;28(3):268–75.
29. Cai X, Zhu H, Li Y. PKCzeta, MMP2 and MMP9 expression in lung adenocarcinoma and association with a metastatic phenotype. Mol Med Rep. 2017;16(6):8301–6.
30. Ohnishi S, Futamura M, Kamino H, Nakamura Y, Kitamura N, Miyamoto Y, et al. Identification of NEEP21, encoding neuron-enriched endosomal protein of 21 kDa, as a transcriptional target of tumor suppressor p53. Int J Oncol. 2010;37(5):1133–41.
31. Kudoh T, Kimura J, Lu ZG, Miki Y, Yoshida K. D4S234E, a novel p53-responsive gene, induces apoptosis in response to DNA damage. Exp Cell Res. 2010;316(17):2849–58.
32. Veillette A. Immune regulation by SLAM family receptors and SAP-related adaptors. Nat Rev Immunol. 2006;6(1):56–66.
33. Yan Q, Malashkevich VN, Fedorov A, Fedorov E, Cao E, Lary JW, et al. Structure of CD84 provides insight into SLAM family function. Proc Natl Acad Sci U S A. 2007;104(25):10583–8.
34. Martin M, Romero X, de la Fuente MA, Tovar V, Zapater N, Esplugues E, et al. CD84 functions as a homophilic adhesion molecule and enhances IFN-gamma secretion: adhesion is mediated by Ig-like domain 1. J Immunol. 2001;167(7):3668–76.
35. Christoforidis S, McBride HM, Burgoyne RD, Zerial M. The Rab5 effector EEA1 is a core component of endosome docking. Nature. 1999;397(6720):621–5.
36. Simonsen A, Lippé R, Christoforidis S, Gaullier JM, Brech A, Callaghan J, et al. EEA1 links PI(3)K function to Rab5 regulation of endosome fusion. Nature. 1998;394(6692):494–8.

37. Li Y, Seto E. HDACs and HDAC Inhibitors in Cancer Development and Therapy. Cold Spring Harb Perspect Med. 2016;6(10).
38. Minamiya Y, Ono T, Saito H, Takahashi N, Ito M, Mitsui M, et al. Expression of histone deacetylase 1 correlates with a poor prognosis in patients with adenocarcinoma of the lung. Lung Cancer. 2011;74(2):300–4.
39. Burdelski C, Ruge OM, Melling N, Koop C, Simon R, Steurer S, et al. HDAC1 overexpression independently predicts biochemical recurrence and is associated with rapid tumor cell proliferation and genomic instability in prostate cancer. Exp Mol Pathol. 2015;98(3):419–26.
40. Zhang L, Bu L, Hu J, Xu Z, Ruan L, Fang Y, et al. HDAC1 knockdown inhibits invasion and induces apoptosis in non-small cell lung cancer cells. Biol Chem. 2018;399(6):603–10.
41. Pan J, Zheng QZ, Li Y, Yu LL, Wu QW, Zheng JY, et al. Discovery and Validation of a Serologic Autoantibody Panel for Early Diagnosis of Esophageal Squamous Cell Carcinoma. Cancer Epidemiol Biomarkers Prev. 2019;28(9):1454–60.
42. Swensen SJ, Silverstein MD, Ilstrup DM, Schleck CD, Edell ES. The probability of malignancy in solitary pulmonary nodules. Application to small radiologically indeterminate nodules. Arch Intern Med. 1997;157(8):849–55.
43. Li Y, Chen KZ, Wang J. Development and validation of a clinical prediction model to estimate the probability of malignancy in solitary pulmonary nodules in Chinese people. Clin Lung Cancer. 2011;12(5):313–9.

Tables

Table 1.Characteristics of participants in discovery cohort,validation cohort 1 and validation cohort 2.

	Discovery cohort		Validation cohort 1		Validation cohort 2	
	LC	NC	LC	NC	MPN	BPN
N	10	10	212	212	105	105
Age（year						
Mean ± SD	63±12	57±10	59±12	56±12	55±9	57±9
Range	43-82	39-70	26-85	28-89	26-72	31-81
Gender						
Male (%)	6(60.0)	6(60.0)	110(51.9)	110(51.9)	62(59.0)	62(59.0)
Female (%)	4(40.0)	4(40.0)	102(48.1)	102(48.1)	43(41.0)	43(41.0)
Smoking						
Yes (%)			96(45.3)		32(30.5)	30(28.6)
No (%)			104(49.1)		65(61.9)	66(62.8)
Unknown (%)			12(5.6)		8(7.6)	9(8.6)
Drinking						
Yes (%)			52(24.5)		17(16.2)	20(19.0)
No (%)			148(69.8)		80(76.2)	76(72.4)
Unknown (%)			12(5.7)		8(7.6)	9(8.6)
Clinical stage						
I (%)	3(30.0)		36(17.0)		30(28.6)	
II (%)	1(10.0)		15(7.1)		14(13.3)	
III (%)	3(30.0)		69(32.5)		15(14.3)	
IV (%)	3(30.0)		79(37.3)		34(32.4)	
Unknown (%)			13(6.1)		12(11.4)	
Histologic type						
NSCLC						
ADC (%)	6(60.0)		71(33.5)		69(65.7)	
SCC (%)	4(40.0)		95(44.8)		14(13.3)	
LCLC (%)			2(0.9)		1(1.0)	
SCLC			34(16.0)		6(5.7)	
Unknown (%)			10(4.7)		15(14.3)	
Lymph node Metastasis						
Yes (%)			131(61.7)		49(46.7)	
No (%)			68(32.2)		44(41.9)	
Unknown (%)			13(6.1)		12(11.4)	
Distant metastasis						
Yes (%)			80(37.7)		34(32.4)	
No (%)			119(56.2)		59(56.2)	
Unknown (%)			13(6.1)		12(11.4)	
CEA						
0−5 ng/mL(%)			80(37.7)	187(88.2)	56(53.3)	62(59.0)
>5 ng/mL (%)			42(19.8)	11(5.2)	23(21.9)	4(3.8)
Unknown (%)			90(42.5)	14(6.6)	26(24.8)	39(37.2)
CA125						
0-35 U/ml (%)					44(41.9)	52(49.5)
>35 U/ml (%)					27(25.7)	11(10.5)
Unknown (%)					34(32.4)	42(40.0)
CYFRA211						
0-3.3 ng/mL (%)					46(43.8)	54(51.4)

>3.3 ng/mL (%)	30(28.6)	7(6.7)
Unknown (%)	29(27.6)	44(41.9)

NC: normal control, LC: lung cancer, SD: standard deviation, ADC: adenocarcinoma, SCC: squamous cell carcinoma, NSCLC: non-small cell lung cancer, SCLC: small cell lung cancer, MPN: malignant pulmonary nodule, BPN: benign pulmonary nodule

Table 2. Nodular characteristic of CT in validation cohort 2.

Nodular Characteristic of CT	MPN (%) (n=105)	BPN (%) (n=105)	P
Number			**<0.001***
1	87(82.9)	60(57.1)	
>1	18(17.1)	45(42.9)	
Diameter（cm）			
Mean±SD	74(70.5)	83(79.0)	0.153
Range	31(29.5)	22(21.0)	
Edge			0.072
Yes (%)	26(24.8)	38(36.2)	
No (%)	79(75.2)	67(63.8)	
Empty			**0.023***
Yes	8(7.6)	19(18.1)	
No	97(92.4)	86(81.9)	
Spicule sign			**0.004***
Yes	30(28.6)	13(12.4)	
No	75(71.4)	92(87.6)	
Vascular notch sign			**<0.001***
Yes	80(76.2)	48(45.7)	
No	25(23.8)	57(54.3)	
Lobulation sign			**<0.001***
Yes	55(52.4)	13(12.4)	
No	50(47.6)	92(87.6)	
Spines			**0.009***
Yes (%)	18(17.6)	6(5.7)	
No (%)	87(82.9)	99(94.3)	
Pleural indentation			**0.004***
Yes	35(33.3)	17(16.2)	
No	70(66.7)	88(83.8)	
Mediastinal lymph node enlargement			**0.003***
Yes	36(34.3)	17(16.2)	
No	69(65.7)	88(83.8)	
Emphysema			0.408
Yes	21(20.0)	26(24.8)	
No	84(80.0)	79(75.2)	
Calcification			**<0.001***
Yes	3(2.9)	25(23.8)	
No	102(97.1)	80(76.2)	

LDCT: low-dose computed tomography, BPN: benign pulmonary nodule, MPN: malignant pulmonary nodule, *: there are significant difference between BPN and MPN patients.

Table 3. Diagnostic value of 11 TAAbs in validation cohort 1 and validation cohort 2.

TAAbs	Validation cohort 1					Validation cohort 2				
	Se(%)	Sp(%)	P	Accuracy (%)	Cutoff	Se(%)	Sp(%)	P	Accuracy (%)	Cutoff
SARS	21.7	91.5	<0.001	56.6	0.363	12.4	94.3	0.021	53.3	0.539
ZPR1	12.7	94.3	0.001	53.5	0.429	-	-	-	-	-
FAM131A	19.8	92.9	<0.001	56.4	0.287	-	-	0.071	-	-
GGA3	8.96	99.5	0.001	54.2	0.491	-	-	-	-	-
PRKCZ	22.2	92.5	<0.001	57.3	0.356	-	-	0.199	-	-
HDAC1	25.0	90.1	0.005	57.5	0.334	-	-	-	-	-
GOLPH3	25.5	90.1	<0.001	57.8	0.258	19.0	90.5	0.007	54.8	0.417
NSG1	23.6	91.5	<0.001	57.5	0.316	21.0	90.5	0.02	55.7	0.345
DAB1	-	-	0.487	-	-	-	-	-	-	-
CD84	27.4	91.0	<0.001	59.2	0.402	15.2	90.5	0.045	52.8	0.502
EEA1	23.1	90.6	<0.001	56.8	0.381	21.9	90.5	<0.001	56.2	0.392

Se: sensitivity, Sp: specificity, AUC: area under curve, 95%CI: 95% confidence interval.

Table 4. Diagnostic value of model 1 in LC patients with different clinical stages.

Group	n	Se (%)	Sp(%)	+LR	-LR	PPV (%)	NPV (%)	YI	Accuracy (%)	Cutoff
All LC vs. NC	122	68.9	83.8	4.25	0.37	72.4	81.4	0.53	78.1	0.392
NSCLC vs. NC	94	70.2	83.8	4.33	0.36	67.3	85.6	0.54	79.5	0.392
SCLC vs. NC	21	71.4	83.8	4.41	0.34	31.9	96.5	0.55	82.6	0.392
Early LC vs. NC	23	47.8	83.8	2.95	0.62	25.6	93.3	0.32	80.1	0.392
Advanced LC vs. NC	91	73.6	83.8	4.54	0.32	67.7	87.4	0.57	80.6	0.392
LM(-) LC vs. NC	33	60.6	83.8	3.74	0.47	38.5	92.7	0.44	80.5	0.392
LM(+) LC vs. NC	81	71.6	83.8	4.42	0.34	64.4	87.8	0.55	80.3	0.392
DM(-) LC vs. NC	63	66.7	83.8	4.12	0.40	56.8	88.8	0.51	79.7	0.392
DM(+) LC vs. NC	51	70.6	83.8	4.36	0.35	52.9	91.7	0.54	81.1	0.392

Se: sensitivity, Sp: specificity, AUC: area under curve, 95%CI: 95% confidence interval, +LR: positive likelihood ratio, -LR: negative likelihood ratio, PPV: positive predictive value, NPV: negative predictive value, YI: Youden's index, LM(-): negative lymph node metastasis, LM(+): positive lymph node metastasis, DM(+): positive distant metastasis, DM(-): negative distant metastasis, NC: normal control, LC: lung cancer, NSCLC: non-small cell lung cancer, SCLC: small cell lung cancer

Table 5. Diagnostic value of model 2 in MPN subgroups and BPN.

Group	n	Se(%)	Sp(%)	+LR	-LR	PPV(%)	NPV(%)	YI	Accuracy(%)	Cutoff
All MPN vs.BPN	60	58.3	96.6	17.1	0.43	94.6	69.1	0.55	77.1	0.676
Early MPN vs.BPN	26	26.9	96.6	7.9	0.76	77.8	74.7	0.24	75.0	0.676
Advanced MPN vs.BPN	31	80.6	96.6	23.7	0.20	92.6	90.3	0.77	91.0	0.676
LM(-) MPN vs.BPN	20	20.0	96.6	5.8	0.83	66.7	77.8	0.17	77.9	0.676
LM(+) MPN vs.BPN	24	83.3	96.6	24.5	0.17	90.9	93.3	0.80	92.7	0.676
DM(-) MPN vs.BPN	27	29.6	96.6	8.7	0.73	80.0	74.7	0.26	75.2	0.676
DM(+) MPN vs.BPN	18	94.4	96.6	27.7	0.06	89.5	98.2	0.91	96.1	0.676
≥3cm MPN vs.BPN	43	48.8	96.6	14.40	0.53	90.9	70.9	0.45	75.2	0.676
<3cm MPN vs.BPN	17	82.4	96.6	24.20	0.18	87.5	94.9	0.79	93.3	0.676

Se: sensitivity, Sp: specificity, AUC: area under curve, 95%CI: 95% confidence interval, +LR: positive likelihood ratio, -LR: negative likelihood ratio, PPV: positive predictive value, NPV: negative predictive value, YI: Youden's index, LM(-): negative lymph node metastasis, LM(+): positive lymph node metastasis, DM(+): positive distant metastasis, DM(-): negative distant metastasis, MPN: malignant pulmonary nodule, BPN: benign pulmonary nodule.

Figures

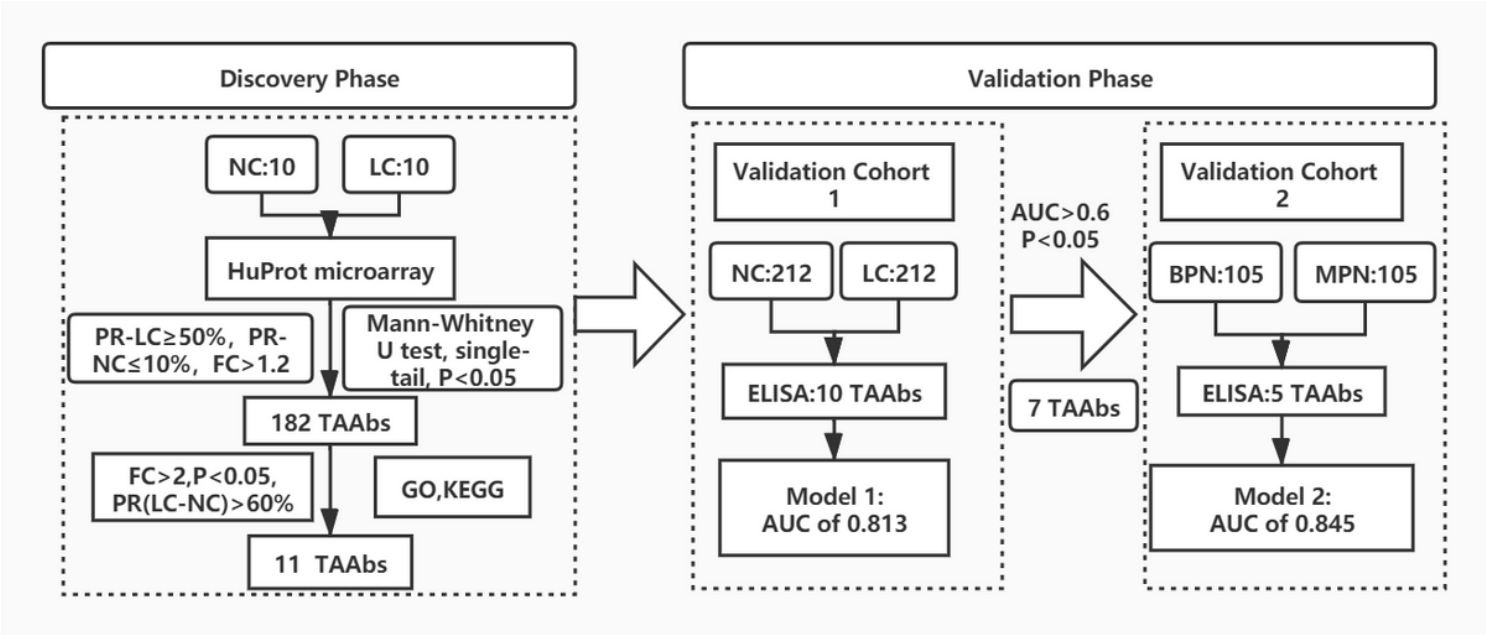

Figure 1

The flow diagram of this study. NC: normal control, LC: lung cancer, FC: fold change, PR: positive ratio, BPN: benign pulmonary nodule, MPN: malignant pulmonary nodule, Model 1: anti-ZPR1, anti-PRKCZ, anti-NSG1, anti-CD84 and CEA, Model 2: anti-EEA1, CEA, CYFRA211, CA125, vascular notch sign, lobulation sign and mediastinal lymph node enlargement

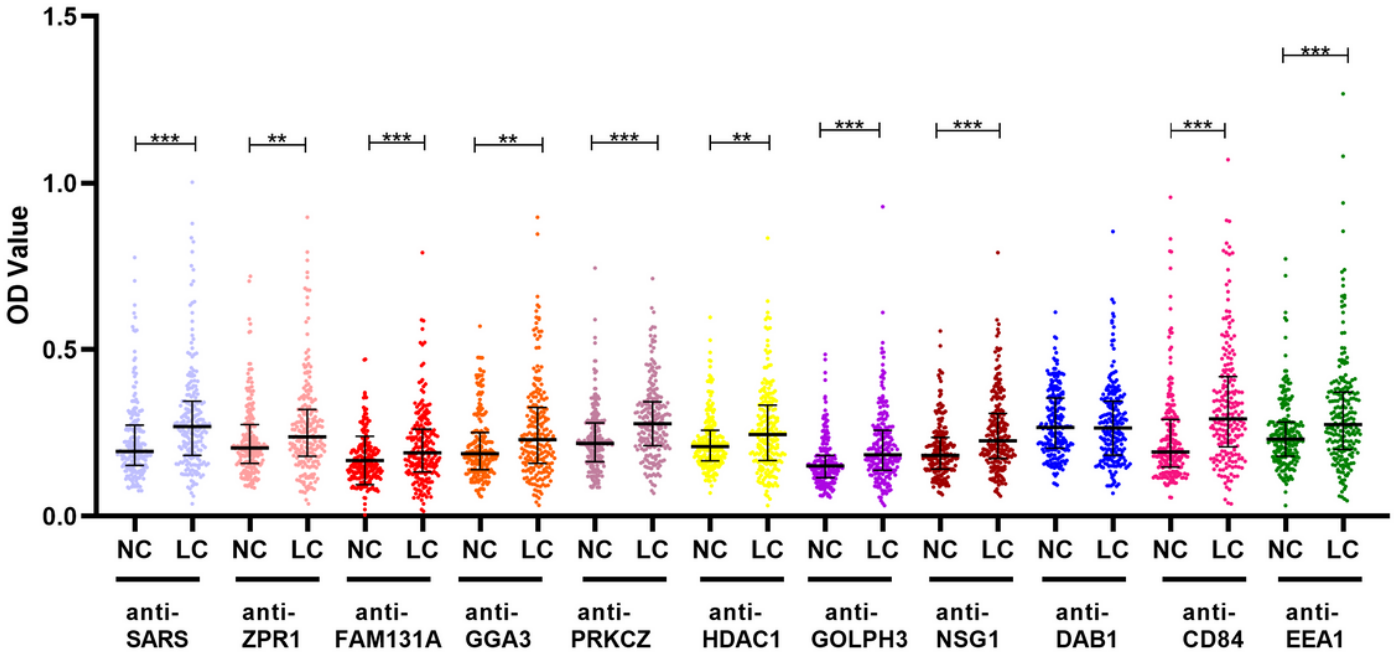

Figure 2

The scatter plot of 11 selected TAAbs in LC patients and NC. NC: normal control, LC: lung cancer, *: P<0.05, **: P<0.01, ***: P<0.001, Lines represented median and quartile range.

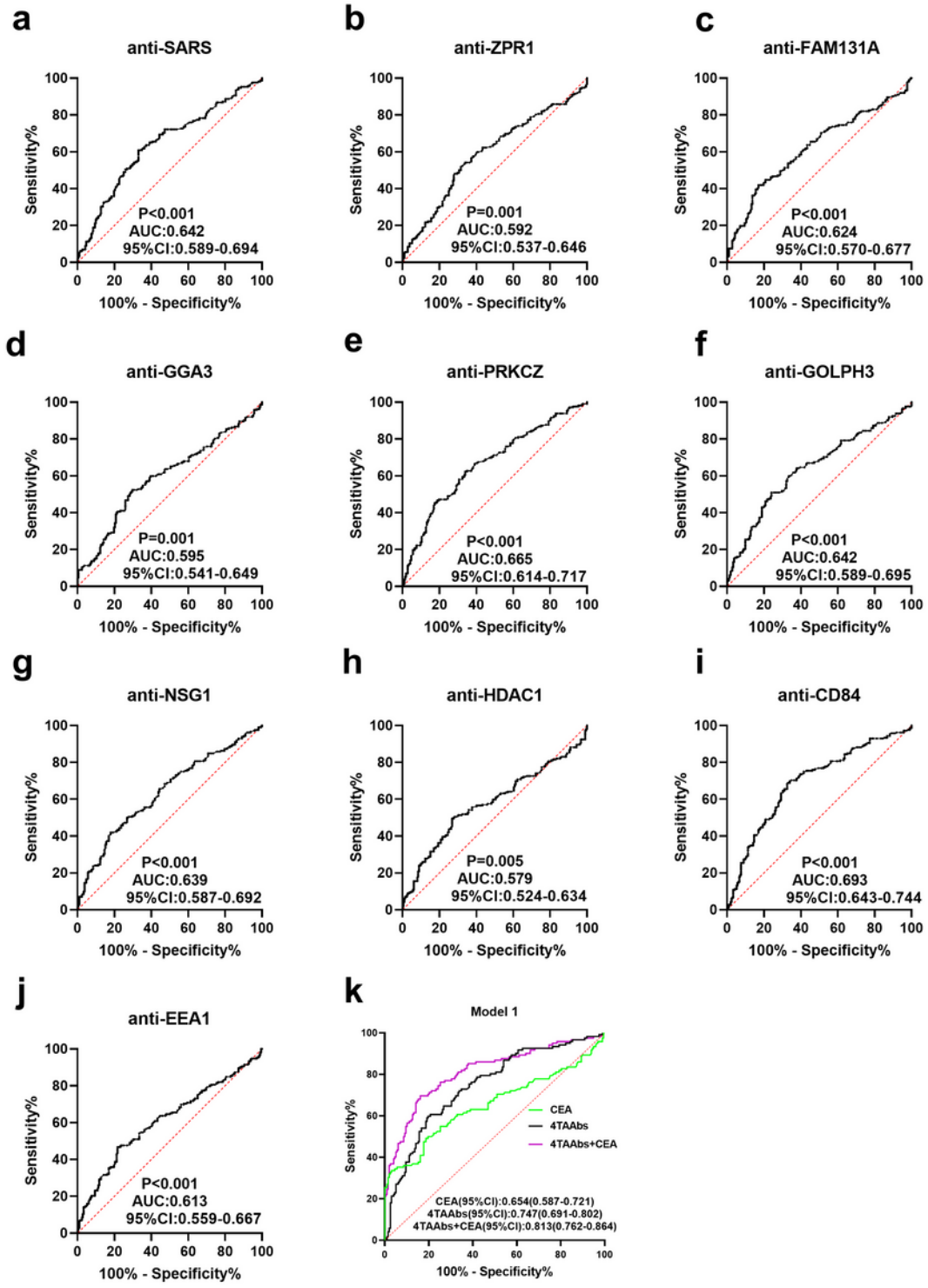

Figure 3

Diagnostic value of 10 significant TAAbs and several combination in LC patients and NC. a-j: ROC of 10 significant TAAbs for discriminating LC patients from NC. k: ROC comparison of CEA, 4TAAbs and model 1 in LC detection. LC: lung cancer, NC: normal control, 4TAAbs: anti-ZPR1, anti-PRKCZ, anti-NSG1 and anti-CD84

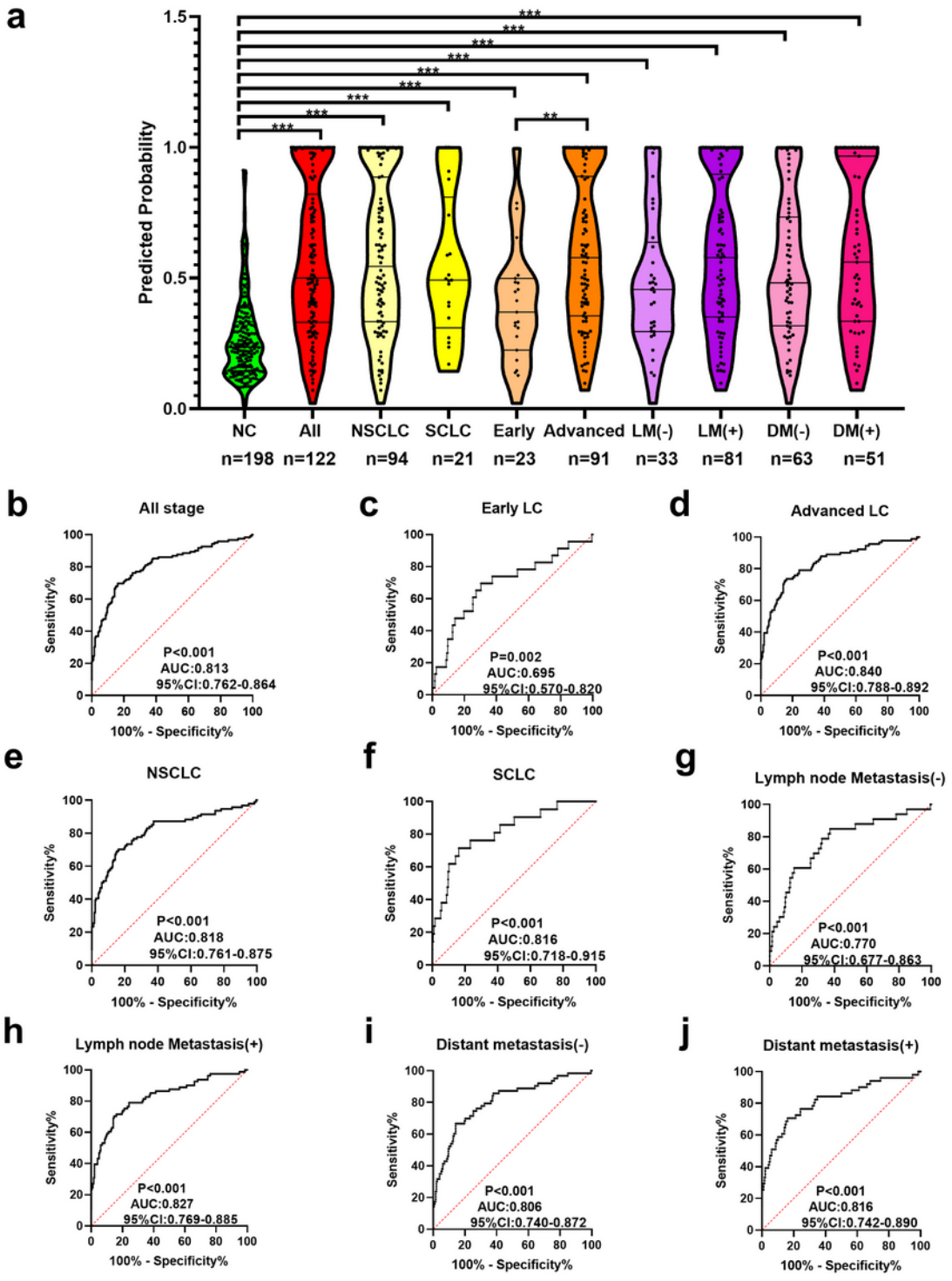

Figure 4

Diagnostic value of model 1 in LC patients with different clinical stages. a.violin plot and scatter plot of predicted probability of model in LC patients with different clinical character and NC. b-j. ROC of model for distinguishing LC patients with different clinical character from NC. n:number, NC: normal control, NSCLC: non-small cell lung cancer, SCLC: small cell lung cancer, LM: lymph node metastasis, DM: distant

metastasis, ALL: all LC, Early: early LC, Advanced: advanced LC.*: P<0.05, **: P<0.01, ***: P<0.001, Lines represented median and quartile range

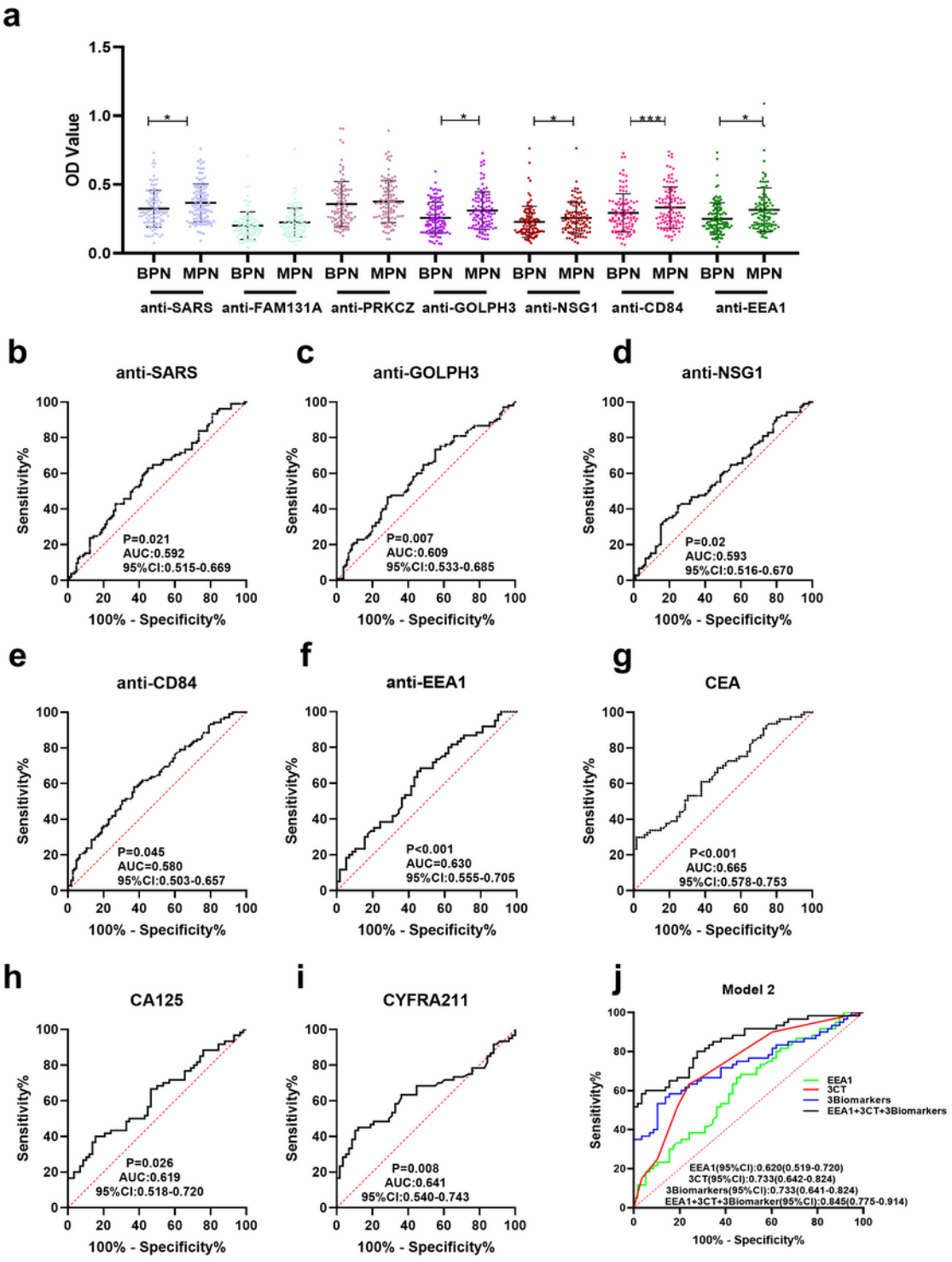

Figure 5

Diagnostic performance of 7 TAAbs and 3 traditional biomarkers in BPN and MPN patients. a. scatter plot of OD value of 8 TAAbs in MPN patients and BPN patients. b-i. ROC of 5 significant TAAbs and 3 biomarkers for distinguishing MPN patients from BPN patients. j. ROC comparison of anti-EEA1,

3biomarkers, 3CT and model in MPN patients detection. BPN: benign pulmonary nodule, MPN: malignant pulmonary nodule, 3biomarker: CEA, CYFRA211 and CA125, 3CT: vascular notch sign, lobulation sign and mediastinal lymph node enlargement, *: P<0.05, **: P<0.01, ***: P<0.001, Lines represented median and quartile range

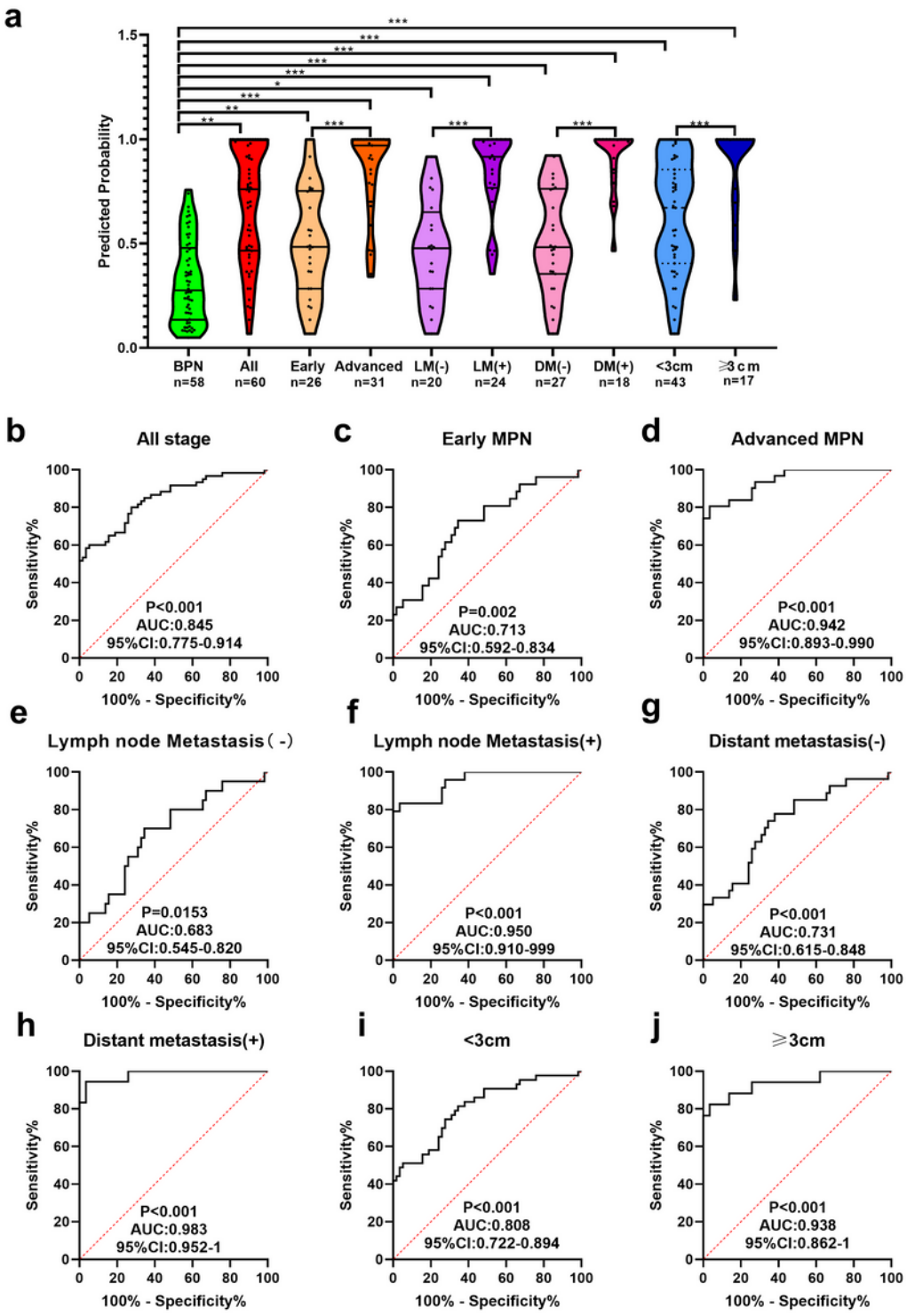

Figure 6

Diagnostic performance of model 2 in validation cohort 2. a: violin plot and scatter plot of predicted probability of model in MPN patients with different clinical character and BPN patients. b-j: ROC of model for distinguishing MPN patients with different clinical character from BPN. n: number, ALL: all LC, Early: early LC, Advanced: advanced LC, LM(-): Lymph node metastasis negative, LM(+): Lymph node metastasis positive, DM(-): distant metastasis negative, DM(+): distant metastasis positive, <3cm: diameter<3cm, ≥3cm: diameter≥3cm, *: P<0.05, **: P<0.01, ***: P<0.001, Lines represented median and quartile range.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementary.doc