Am J Transl Res 2021;13(5):4464-4476
www.ajtr.org /ISSN:1943-8141/AJTR0123472

Original Article

The value of AGR2 and KRT5 as an immunomarker combination in distinguishing lung squamous cell carcinoma from adenocarcinoma

Bo Pan[1*], Zi-Xin Wei[1*], Ju-Xuan Zhang[2*], Xin Li[2], Qing-Wei Meng[1], Ying-Yue Cao[1], Li-Shuang Qi[2], Yan Yu[1]

[1]*Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China;* [2]*College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China.* [*]*Co-first authors.*

Received September 30, 2020; Accepted March 2, 2021; Epub May 15, 2021; Published May 30, 2021

Abstract: With the advancement of tumor subtype-specific treatments, precise histopathologic distinction between adenocarcinoma (ADC) and squamous cell carcinoma (SCC) is of significant clinical importance. Nevertheless, the current markers are insufficiently precise in poorly differentiated tissue. This study aimed to establish a histology-specific immunomarker combination to subclassify non-small cell lung cancer (NSCLC) specimens. Based on previous work, we assessed the differential expression of anterior gradient 2 (AGR2) and keratin 5 (KRT5) in ADC and SCC by analyzing public datasets and postoperative specimens. Subsequently, we established a train set (n = 188) and a validation set (n = 42) comprised of NSCLC surgical specimens for training and verifying the subtype-identification capabilities of the two biomarkers separately and in combination, and contrasted the diagnostic utility of AGR2-KRT5 with that of the classic immunomarker combination, TTF1-P40. Differential expression of the two genes was statistically significant in ADC and SCC samples, both at the mRNA and protein levels. The specificity and sensitivity of AGR2 to detect ADC in the training set were 97.0% and 94.4%, while the sensitivity and specificity of KRT5 to determine SCC were 93.9% and 98.9%, respectively. The accuracies of AGR2-KRT5 in ADC, SCC, and across all samples were 93.3%, 92.0% and 92.6% respectively. In the validation cohort, the predictive accuracy of AGR2-KRT5 was up to 100% for ADC and 86.7% for SCC. Compared with TTF1-P40 in ADC samples, AGR2-KRT5 had 8.4% higher accuracy. In summary, the AGR2-KRT5 immunomarker combination reliably distinguished SCC from ADC, and was more accurate than TTF1-P40 in ADC.

Keywords: Anterior gradient 2, keratin 5, histological subtype, immunohistochemistry, non-small cell lung cancer

Introduction

Lung cancer is a leading cause of global cancer-related mortality, and is responsible for one-quarter of all cancer deaths [1]. Non-small cell lung cancer (NSCLC) constitutes more than 85% of lung cancer diagnoses. Adenocarcinoma (ADC) and squamous cell carcinoma (SCC) are the most prevalent NSCLC histological phenotypes [2-4]. Despite having several similar biological features, ADC and SCC are distinct diseases that develop through unique molecular mechanisms [5]. They differ in several attributes, such as cell of origin, primary sites, and tumor progression [6, 7], suggesting that certain targeted therapies for NSCLC depend on histology. Due to the emergence of precision medicine and therapies with differential therapeutic efficacies and toxicities in specific histological tumor subtypes [8-12], accurate subtyping of NSCLC is essential to guide treatment decisions.

Because the resection rate for lung cancer is approximately 30% [3], pathological diagnosis is often made by evaluating small biopsy samples. Although the majority of NSCLCs can be subtyped through the examination of hematoxylin and eosin (H&E)-stained slides, morphological findings alone may be insufficient to refine the diagnosis of poorly differentiated and heterogenous tumors [5], especially if specimen quantities are limited. Consequently, the 2015 World Health Organization (WHO) Lung Cancer Classification Guidelines [2] recommend the use of immunohistochemistry (IHC) in sub-classifying NSCLCs. The combination of multiple ADC and SCC immunomarkers may optimize the accurate and precise diagnosis of NSCLC histological subtypes, and the National Comprehensive Cancer Network (NCCN) guidelines also recommend a panel of thyroid transcription factor-1 (TTF1) and P40 to facilitate precision diagnosis [13-15]. Nevertheless, their performance on poorly differentiated tissue remains far from ideal. For example, the diagnostic accuracy of TTF1 to detect ADC in specimens lacking specific morphological features may be as low as 60% [16]. Subtypes cannot be determined accurately with the current IHC marker combinations in ambiguous cases, necessitating second pathological assessments and causing treatment delays. Additionally, the WHO guidelines suggest that tissue samples for IHC staining should be used sparingly to ensure sufficient samples for further molecular testing [3]. For instance, the selection of patient candidates for anaplastic lymphoma-kinase and programmed death ligand 1 inhibitor treatments requires IHC analyses [17, 18]. Consequently, discovering new markers with enhanced diagnostic precision in discerning NSCLC pathological subtypes has tremendous relevance.

We previously developed a qualitative transcriptional signature to identify histologic classifications of ADC and SCC. The signature comprised two genes, anterior gradient 2 (*AGR2*) and keratin 5 (*KRT5*) [19], both of which have been used previously as histology-specific biomarkers for discriminating between ADC and SCC at the protein level [20, 21]. In this study, we trained and validated the ability of AGR2 and KRT5 for determining ADC and SCC, and contrasted the diagnostic accuracy of AGR2-KRT5 with that of the classic immunomarker combination, TTF1-P40, whose use in histologic subclassification of NSCLC has been proposed in the NCCN guidelines [14]. The aim of this study was to assess the value of AGR2-KRT5 in distinguishing SCC from ADC.

Materials and methods

Study design

The study design is illustrated in **Figure 1**. First, the differential mRNA expression of AGR2 and KRT5 along with TTF1/P40 and other related biomarkers used for the differential diagnosis of ADC and SCC, were determined by analyzing NSCLC tissue results in six independent public gene expression datasets. Then, we applied IHC to evaluate the differences in protein expression levels of the two genes in a set of resection specimens obtained from NSCLC patients in our hospital (n = 188). Subsequently, we applied clinical data as a training set to "train" the two biomarkers (separately and in combination) to discriminate among NSCLC histological subtypes. We validated the ability of AGR2 and KRT5 to distinguish ADC and SCC in a set of 42 resected poorly differentiated NSCLC tissues (the validation set), and compared it with that of the commonly utilized IHC markers, TTF1 and P40 (alone and in combination, by retrieving the IHC staining results from the postoperative pathological reports of 42 patients). Lastly, the association of AGR2/KRT5 expression and clinicopathological parameters including patient survival was studied in the training set and public databases.

Gene signature for histological classification of NSCLC

Details of signature construction have been described previously [19]. Simply put, the genes that were differentially expressed in both ADC and SCC were extracted, and displayed opposite dysregulated directions after drawing comparisons with healthy tissue (termed subtype-opposite genes). From the subtype-opposite genes, we identified gene pairs whose relative expression ordering patterns (E*a* > E*b*) occurred significantly more frequently in ADC than in SCC samples. E*a* and E*b* represent the mRNA expression of gene *a* and gene *b,* respectively. The gene pair of *AGR2* and *KRT5* exhibited the highest accuracy (98.43%) in a stepwise forward selection procedure, and was chosen as the signature for distinguishing SCC from ADC. As a rule, AGR2 mRNA expression higher than that of KRT5 was pathognomic of ADC; specimens without this finding were designated as SCC.

mRNA expression analyses of the two signature genes and other diagnostic markers

Six public gene expression datasets of NSCLC tissues (Supplementary Table 1), which were not the training sets for the signature, were downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/). We evaluated whether the mRNA of *AGR2* and *KRT5* was differentially expressed between

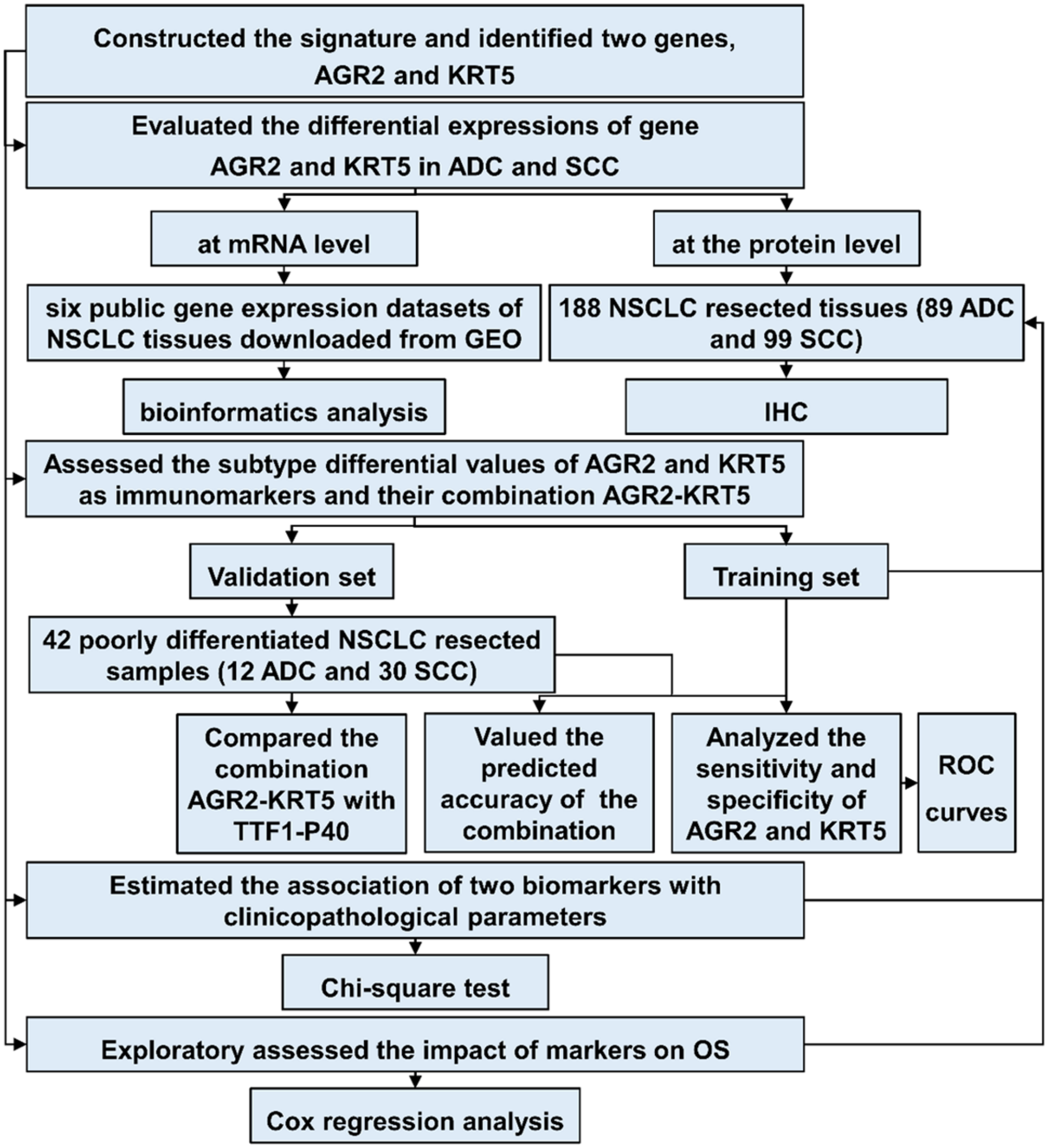

Constructed the signature and identified two genes, AGR2 and KRT5
Evaluated the differential expressions of gene AGR2 and KRT5 in ADC and SCC
at mRNA level
at the protein level
six public gene expression datasets of NSCLC tissues downloaded from GEO
188 NSCLC resected tissues (89 ADC and 99 SCC)
bioinformatics analysis
IHC
Assessed the subtype differential values of AGR2 and KRT5 as immunomarkers and their combination AGR2-KRT5
Validation set
Training set
42 poorly differentiated NSCLC resected samples (12 ADC and 30 SCC)
Compared the combination AGR2-KRT5 with TTF1-P40
Valued the predicted accuracy of the combination
Analyzed the sensitivity and specificity of AGR2 and KRT5
ROC curves
Estimated the association of two biomarkers with clinicopathological parameters
Chi-square test
Exploratory assessed the impact of markers on OS
Cox regression analysis

Figure 1. Study design. Differences in mRNA expression of the biomarker genes were analyzed using the Student's t-test. The subtype identification abilities of the two markers were evaluated by IHC staining in both the training and validation sets. AGR2, anterior gradient 2; KRT5, keratin 5; ADC, adenocarcinoma; SCC, squamous cell carcinoma; NSCLC, non-small cell lung cancer; GEO: Gene Expression Omnibus; IHC, immunohistochemistry; ROC, receiver operator characteristic; OS, overall survival.

ADC and SCC samples. We also assessed the mRNA expression of TTF1/P40 and other subtype-specific biomarkers in the six datasets. It is worth mentioning that P40 is an isoform of the gene *P63* transcript-ΔNp63 that has a variant structure of the N-terminal domain [22]. Therefore, mRNA expression levels of P40 and P63 were identical; consequently, their results were combined for display.

Patients and specimens

A total of 230 resected, formalin-fixed, paraffin-embedded NSCLC tissue specimens from patients who had undergone surgical treatment at our hospital between January 2006 and December 2014 (n = 188; the training set) or between October 2015 and March 2019 (n = 42; the validation set) were obtained. The training set included specimens from patients with postoperative diagnoses of ADC (n = 89) or SCC (n = 99). The validation set comprised specimens from patients with postoperative diagnoses of poorly differentiated NSCLC by H&E staining. Initial diagnoses were subsequently refined as ADC (n = 12) or SCC (n = 30) by IHC staining, utilizing TTF1, P40, Napsin A, P63 and other classic immunomarkers. Clinicopathological and prognostic data of the patients in the training set were retrieved from the medical electronic database. The study was approved by the Ethics Committee of Harbin Medical University Cancer Hospital (KY2017-12). All patients provided written informed consent for the use of their tissue for research purposes. Tumors were staged and classified on the basis of the 8th Edition of Lung Cancer Stage Classification [23].

Immunohistochemistry

Formalin-fixed, paraffin-embedded specimens were cut into 4 μm thick sections. The tissue slices were heated at 70°C overnight, then deparaffinized in xylene and rehydrated in an ethanol gradient. To block endogenous peroxidase activity, the slides were treated with 3% hydrogen peroxide and pressure-cooked in antigen retrieval solution (0.01 mmol/L citrate buffer, pH 6.0) for 5 mins. The specimens were then incubated with rabbit monoclonal anti-AGR2 (ab76473, Abcam, Cambridge, UK; dilution 1:400) and rabbit monoclonal anti-KRT5 (ab52635, Abcam; dilution 1:400) antibodies overnight at 4°C. Subsequently, the sections were washed and probed with a secondary antibody (Zhongshan Golden Bridge Biotechnology, Beijing, China) for 20 mins at 25°C. Specimens were then incubated for two mins after being treated with 3,3'-diaminobenzidine (Zhongshan Golden Bridge Biotechnology, Beijing, China), and counterstained with hematoxylin solution for 30 secs. Human colon and squamous cell lung carcinoma tissues were used as the positive controls for AGR2 and KRT5 IHC staining, and tissues without primary antibody treatment were utilized as negative controls.

Immunohistochemistry evaluation

Immunostaining was blindly evaluated by two pathologists utilized a semi-quantitative scoring system. Discrepant results were reevaluated by the two pathologists. The consensus among the two pathologists scoring AGR2 and KRT5 expression with this semi-quantitative system was 96%. Five high-magnification visual fields were chosen randomly in each slice under an optical microscope (400×; Leica, Wetzlar, Germany) to observe positive staining results of resected NSCLC tissue. The intensity of the staining was graded thus: 0, negative; 1, mild positive; 2, moderate positive; or 3, strong positive, while the proportion of staining-positive cells was categorized as follows: 0, none; 1, 1-25%; 2, 26-50%; 3, 51-75%, or 4, 76-100% cells stained. The combination of the intensity and percentage scores were obtained [24, 25], and the mean of the five chosen fields represented the final scores. The integer portion of the averages was taken for convenience. Total scores <3 were considered negative (-) while scores ≥3, were deemed positive (+). Only the cytoplasmic staining was assessed; membranous staining was not considered positive. Detailed information on the expression of AGR2 and KRT5 along with the final IHC scores for each sample in training and validation sets are displayed in the supplementary material (Supplementary Table 2).

Assessment criteria for subtype identification using the AGR2-KRT5 biomarker panel

The utility of AGR2-KRT5 for discriminating ADC from SCC was evaluated using two methods: comparing the total IHC score, and assessing the positive/negative expression of each biomarker. Using the former method, a sample was assigned as ADC if the IHC score for AGR2 was higher than that for KRT5; otherwise, it was categorized as SCC. With the latter technique, only samples with AGR2+ and KRT5- expressions were identified as ADC; conversely, SCC was diagnosed when expression patterns were AGR2+ and KRT5-.

Statistical analyses

Statistical analyses were conducted in SPSS 24.0 (IBM SPSS, Armonk, NY, USA). Student's t-test was utilized in analyzing the mRNA expression of the subtype-specific markers in public gene expression datasets, while Pearson's Chi-square test was applied for assessing the relationship between marker expression and clinicopathological features. Receiver operator characteristic (ROC) curves were constructed and the area under the curve (AUC) for each

ROC curve was computed in evaluating the diagnostic value of markers. Survival curves were plotted using univariate Cox regression analysis. The overall survival (OS) time of the patients in the training set was defined commencing from the surgical operation to moment of death or, if the patient was still alive, the last follow-up (the final follow up date was on 20 November 2018). Differences with $P<0.05$ were considered statistically significant.

Results

mRNA expression of markers in ADC and SCC

In the GSE50081 dataset, which included 127 ADC and 43 SCC samples, AGR2 mRNA expression was considerably higher in ADC than in SCC (Student's t-test, P = 3.46E-05, **Figure 2A**). In contrast, patients with SCC had a significantly higher mRNA expression of KRT5 than those with ADC (Student's t-test, P = 6.14E-17, **Figure 2A**). Similar results were recorded in the remaining five gene expression datasets (**Figure 2B-F**). These results revealed that the differential mRNA expression patterns of the two genes between ADC and SCC samples were extremely reproducible across various datasets, indicating the robustness of AGR2 and KRT5 for differentiating SCC from ADC. Furthermore, the mRNA expressions of TTF1 and NAPSA exhibited trends similar to AGR2. The expressions of P40/P63 tended to be higher in SCC than in ADC, similar to KRT5.

Protein expression of AGR2 and KRT5 in ADC and SCC

The expression of AGR2 was abundant in ADC samples, in which it was localized predominately in the cytoplasm, and was negligible or weak in SCC samples. In contrast, KRT5 was more highly expressed in SCC than in ADC tissues, and was localized in the endochylema. Representative IHC staining of AGR2 and KRT5 with variable expression levels are showed in **Figure 3**. Resected specimens from four patients were selected as representative samples. The training set images were from patients 1 and 2, while the validation set were selected from the remaining two patients. Patient 1 and 3 were diagnosed with ADC by pathology, with AGR2+ expression as well as negative expression for KRT5 (AGR2+/KRT5-). Patient 2 and 4 were diagnosed with SCC, and exhibited the reversed expression pattern (AGR2-/KRT5+).

Subtype discriminatory abilities of the biomarkers and their combination

In the training set, 94.4% and 1% of ADC and SCC samples were AGR2+, while 2% and 93% of ADC and SCC samples were KRT5+, respectively. The ROC curves were analyzed for evaluating both the sensitivity and specificity of AGR2 for ADC and KRT5 for SCC (**Figure 4**). Both markers had acceptable discriminatory capacity, with AUC ROC of 0.975 and 0.986 for AGR2 and KRT5, respectively. Notwithstanding, the optimal cutoff value for the IHC staining score determined by ROC curve analysis was 2.5, which could not be applied in clinical procedure. Therefore, we selected the nearest integers above and below the cutoff value, and computed the associated sensitivity, specificity, and Youden index (data not shown). After comparison, 3 was selected as the optimal cutoff value for dividing the protein expression of our two markers, with 94.4% sensitivity and 97.0% specificity for AGR2, and 93.9% sensitivity and 98.9% specificity for KRT5. The results of this analysis are provided in Supplementary Table 3.

Subsequently, we assessed the ability of the AGR2-KRT5 combination to discern the two NSCLC subtypes. Specimens were initially diagnosed as either ADC or SCC through comparison of the IHC staining scores of the two markers. AGR2-KRT5 had high accuracy in ADC (97.8%), SCC (96.0%), as well as in the entire sample set (96.8%). Despite its accuracy, comparison of IHC staining scores is demanding, time consuming, and exhibits sub-standard reproducibility, thus impeding its clinical implementation. Therefore, we reclassified the tissues based on positive or negative expression of AGR2 and KRT5. This method was equally effective, with accuracies of up to 93.3%, 92.0%, and 92.6% in ADC, SCC, and in all samples, respectively. Based on its higher clinical feasibility, we applied biomarker expression to examine all subsequent outcomes. The abilities of the two methods to distinguish histological subtypes of lung carcinoma are compared in **Table 1**.

Validation of AGR2-KRT5 marker combination for the diagnosis of ADC and SCC

Table 2 presents the sensitivity and specificity of AGR2, KRT5, TTF1, and P40 for the diagnosis of ADC and SCC in the validation set. AGR2 was

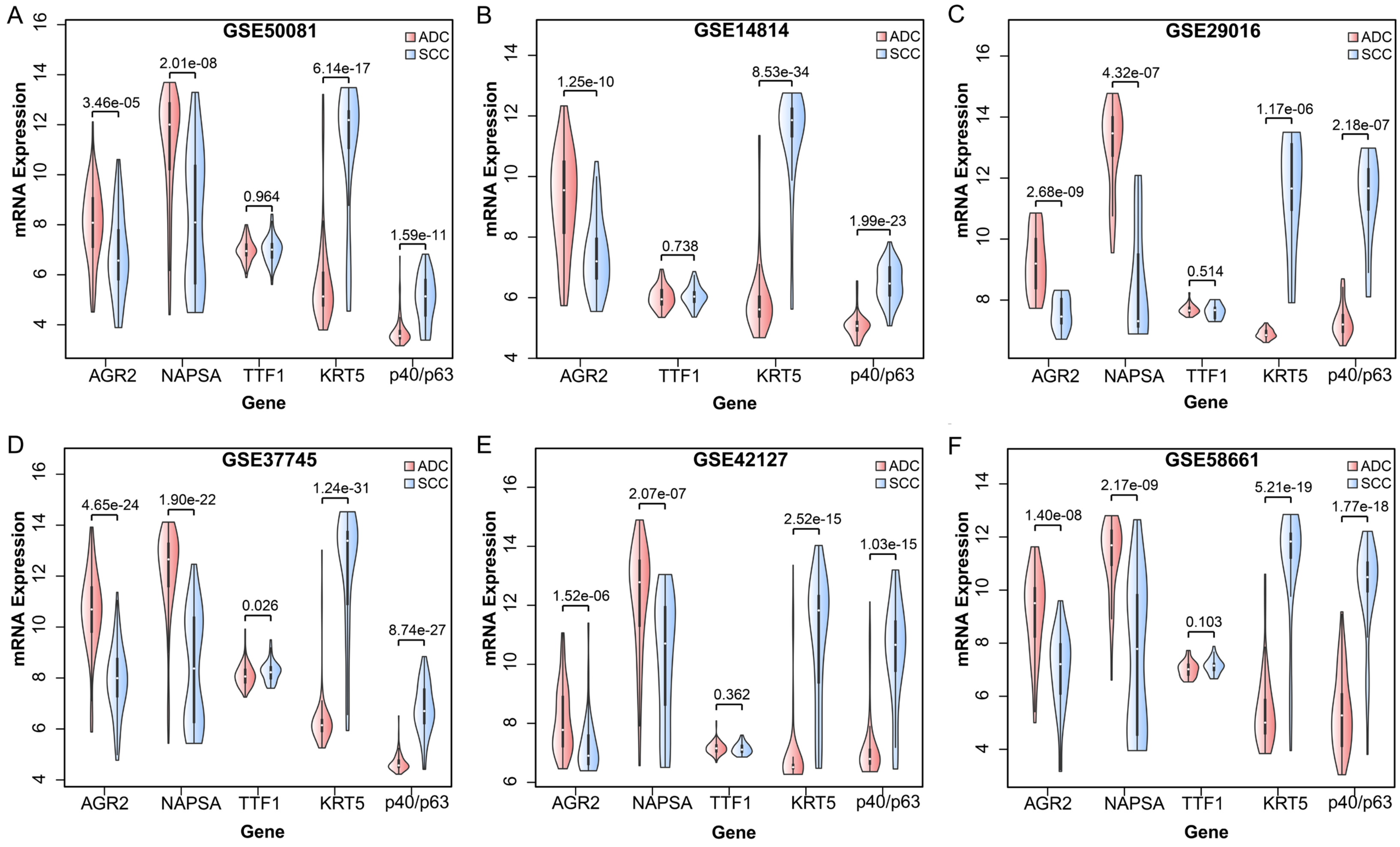


Figure 2. Differential mRNA expression of subtype-specific marker genes in public gene expression datasets. A. GSE50081 dataset. B. GSE14814 dataset. C. GSE29016 dataset. D. GSE37745 dataset. E. GSE42127 dataset. F. GSE58661 dataset.

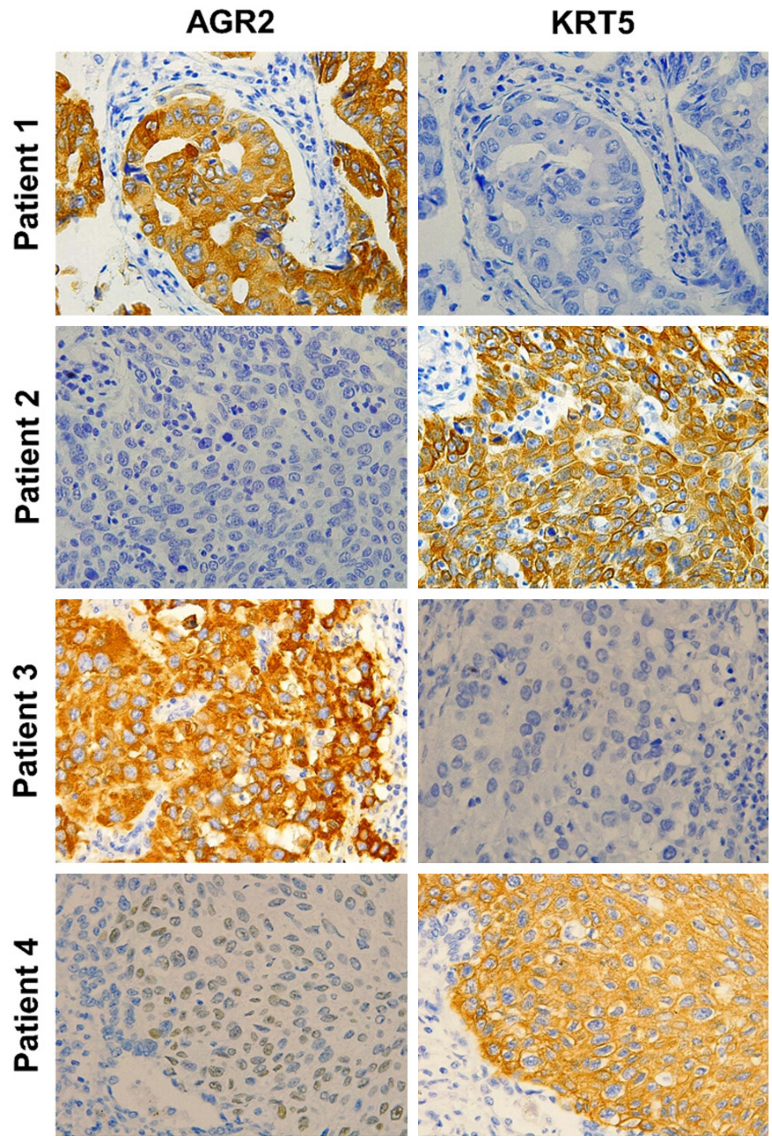


Figure 3. IHC staining for AGR2 and KRT5 (×400). The training set comprised patients 1 and 2, and the validation set included patients 3 and 4. The histopathologic diagnosis for patients 1 and 3 was ADC. The expression pattern of the two IHC markers was AGR2+/KRT5-. Patient 2 and 4 were diagnosed with SCC, with the expression pattern of AGR2-/KRT5+.

superior to TTF1, however, KRT5 sensitivity was marginally worse than P40. Subsequently, we re-identified the histopathologic diagnoses of the validation set tissues using AGR2-KRT5. Among the 42 cases observed, 12 were ADC, 26 were SCC, and the remaining 4 cases could not be categorized. The accuracy of our marker combination was 100% in ADC, 86.7% in SCC, and 90.5% across all samples. The accuracy of TTF1-P40 combination was 91.6% in ADC, 90% in SCC, and 90.5% in all specimens. AGR2-KRT5 was 8.4% more accurate in ADC, and 3.3% less accurate in SCC specimens than TTF1-P40. The discriminatory abilities of AGR2-KRT5 and TTF1-P40 are contrasted in Supplementary Table 4.

Association between biomarker expression and clinicopathological features

The correlation between AGR2/KRT5 protein expression levels and the clinicopathological characteristics of the training set are presented in **Table 3**. The protein expressions of both markers were associated with gender, smoking history, tumor differentiation, histologic type, lymph node metastasis status, and pathological stage (*P*<0.05). We chose GSE50081 from GEO as the representative dataset and analyzed mRNA expressions of AGR2 and KRT5 and their associations with clinicopathological features. These results are presented in Supplementary Table 5.

Association between biomarker expression and patient survival

We initially investigated the prognostic value of AGR2 and KRT5 among the patients in the training set, for which follow-up data of 181 patients were available. Patients with negative AGR2 expression had poorer prognoses compared to those with positive AGR2 expression (Supplementary Figure 1A). Furthermore, KRT5 protein expression and OS were inversely related (Supplementary Figure 1B). Nevertheless, patient survival was unrelated to AGR2 and KRT5 expression. Moreover, the relationship between AGR2/KRT5 mRNA expression and patient prognosis in the GEO database was estimated. Five datasets that record patient survival information were selected for analyzing the relationship between expression

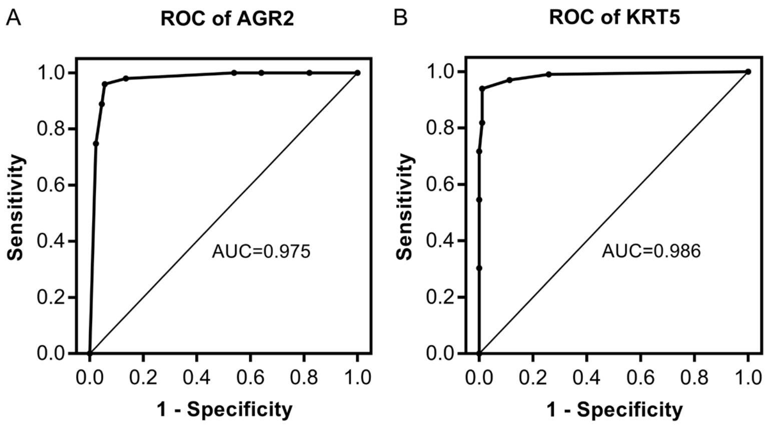


Figure 4. ROC curve assessment of the value of AGR2 and KRT5 in distinguishing ADC and SCC (2.5 taken as the cutoff value). A. AGR2 (sensitivity 94.4%, specificity 96.0%, AUC 0.975). B. KRT5 (sensitivity 93.9%, specificity 98.9%, AUC 0.986).

patterns and OS. Only KRT5 was associated with prognosis in GSE14814 and GSE42127. Unfortunately, these two trends were contradictory. Supplementary Figure 2 illustrates these results.

Discussion

We applied the *AGR2-KRT5* gene pair as a transcriptional signature to distinguish SCC from ADC, and analyzed its predictive performance at the protein level. To our knowledge, this is the first study examining the role of AGR2-KRT5 alone as the IHC marker combination for differentiating ADC from SCC.

AGR2 is an adenocarcinoma antigen that is overexpressed extensively in multiple human cancers, such as breast [26], lung [27, 28], ovarian [29], and nasopharyngeal carcinoma [30] and prostate cancer [31]. AGR2 overexpression is associated with poor differentiation, deep invasion, and lymph node metastasis in various cancers [32, 33], which corroborates our findings. Moreover, AGR2 has been proposed as a potential IHC marker for ADC [21, 27]. The other gene in our transcriptional signature, KRT5, is primarily expressed in epidermal basal keratinocytes. Its overexpression is a distinctive feature of SCC [34, 35]. KRT5, also known as cytokeratin 5 (CK5), was used either as a solitary marker or in combination with CK6 (CK5/6) as part of an SCC-specific IHC diagnostic panel for evaluating NSCLC [36, 37]. In addition, CK5 and CK5/6 demonstrated varying sensitivities and specificities in breast cancer tissue [38]; further studies are required to determine whether CK5 (KRT5) is superior to CK5/6 in differentiating SCC from ADC. Congruent with the results of previous work, we discovered that AGR2 is preferentially expressed in ADC, while KRT5 has a higher tendency to be expressed in SCC, providing biological evidence for the utility of the AGR2-KRT5 panel in distinguishing NSCLC subtypes. Although several previous studies have confirmed that AGR2 and KRT5 are sensitive and specific for ADC and SCC detection respectively, and that either could be the components of NSCLC subclassification immunomarker panels, to the best of our knowledge the IHC marker combination comprised of only AGR2 and KRT5 for determining NSCLC histological subtypes is reported here for the first time.

In clinical practice, the combination of TTF1, Napsin A, CK5/6 and P40 (p63) comprises the most common panel used to diagnose ADC and SCC. Because most NSCLC patients present with advanced disease, a small biopsy specimen represents the sole sample available for subtype classification. To guarantee adequate samples for further molecular assessment, minimal tissue aliquots should be used for IHC staining, which implies the importance of reducing the number of immunomarkers. For small and limited specimens, the NCCN guidelines state that measuring TTF1 and P40 levels is adequate to refine the diagnosis of NSCLC as either ADC or SCC [14]. Nevertheless, these markers have certain limitations and shortcomings. Previous reports have suggested that a portion of SCCs (3-21%) are TTF1-positive [39], which could have been reflected in our TTF1 expression results. Moreover, small biopsies may bias results toward ADCs by raising the prevalence of TTF1 expression [16]. P40 positivity has been reported in ADC and other tumor types [40]. Consistent with our results, previous studies have shown that the sensitivity of P40 is less than its specificity [41, 42]. This has led to the combination of TTF1 and P40 with other IHC markers in clinical practice. Herein,

Table 1. Distinguishing SCC from ADC using various methods

Histological type	No. of cases	By scores[a]				By expression[b]			
		A[c]>K[d]	A<K	A = K	Rate[e]	A+/K-	A-/K+	Others	Rate
ADC	89	87	0	2	97.8%	83	0	6	93.3%
SCC	99	1	95	3	96.0%	1	91	7	92.0%

Notes: [a]distinguished by comparing IHC scores, [b]distinguished by comparing biomarker expression, [c]AGR2, [d]KRT5, [e]Accuracy after classification.

Table 2. Sensitivity, specificity, PPV, and NPV of IHC markers in the validation set

Markers	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
AGR2	100% (12/12)	93.4% (28/30)	85.7% (12/14)	100% (28/28)
TTF1	91.7% (11/12)	93.4% (28/30)	84.6% (11/13)	96.6% (28/29)
KRT5	93.4% (28/30)	100% (12/12)	100% (28/28)	85.7% (12/14)
P40	96.7% (29/30)	100% (12/12)	100% (29/29)	92.3% (12/13)

Notes: Sensitivity = TP/TP+FN; Specificity = TN/TN+FP; Positive predictive value (PPV) = TP/ TP+FP; Negative predictive value (NPV) = TN/TN+FN. FN indicates false negatives; FP, false positives; TN, true negatives; TP, true positives.

Table 3. Associations between marker protein expression and clinicopathological parameters

Variable	No. of cases	AGR2		*P*	KRT5		*P*
		No. of pos. (%)	No. of neg. (%)		No. of pos. (%)	No. of neg. (%)	
Patient age							
<60 years	116	50 (43.1%)	66 (56.9%)	0.196	60 (51.7%)	56 (48.3%)	0.548
≥60 years	72	38 (52.8%)	34 (47.2%)		34 (47.2%)	38 (57.8%)	
Gender							
Male	136	46 (33.8%)	90 (66.2%)	<0.001	84 (61.8%)	52 (38.2%)	<0.001
Female	52	42 (80.8%)	10 (19.2%)		10 (19.2%)	42 (80.8%)	
Smoking history							
Yes	125	40 (32.0%)	85 (68.0%)	<0.001	78 (62.4%)	47 (38.6%)	<0.001
No	63	48 (76.2%)	15 (23.8%)		16 (25.4%)	47 (74.6%)	
Family history							
Yes	38	16 (42.1%)	22 (57.9%)	0.515	20 (52.6%)	18 (47.4%)	0.716
No	150	72 (48.0%)	78 (52.0%)		74 (49.3%)	76 (50.7%)	
Differentiation							
Well-differentiated	47	40 (85.1%)	7 (14.9%)	<0.001	6 (12.8%)	41 (87.2%)	<0.001
Moderately or poorly differentiated	141	48 (34.0%)	93 (66.0%)		88 (62.4%)	53 (37.6%)	
Histology							
ADC	89	84 (94.4%)	5 (5.6%)	<0.001	1 (1.1%)	88 (98.9%)	<0.001
SCC	99	4 (4.0%)	95 (96.0%)		93 (93.9%)	6 (6.1%)	
pT-status							
pT1-2	165	78 (50.3%)	77 (49.7%)	0.036	73 (47.1%)	82 (52.9%)	0.084
pT3-4	33	10 (30.3%)	23 (69.7%)		21 (63.6%)	12 (36.4%)	
pN-status							
pN0-1	76	70 (92.1%)	6 (7.9%)	<0.001	4 (5.3%)	72 (94.7%)	<0.001
pN2-3	112	18 (16.1%)	94 (83.9%)		90 (80.4%)	22 (19.6%)	
TNM stage							
I-II	71	67 (94.4%)	4 (5.6%)	<0.001	1 (1.4%)	70 (98.6%)	<0.001
III-IV	117	21 (17.9%)	96 (82.1%)		93 (79.5%)	24 (20.5%)	

Notes: Differences with *P*<0.05 were considered statistically significant.

AGR2-KRT5 and TTF1-P40 demonstrated comparable abilities to identify ADC and SCC in poorly differentiated tissues, with equivalent precision across the validation set samples. Furthermore, in the ADC samples, AGR2-KRT5 had a higher accuracy (up to 100%) than TTF1-P40.

However, AGR2-KRT5 performance in the validation set was less than that in the training set, and identified only 86.7% of the SCC tissues correctly. This may be due to several reasons. Firstly, the validation set specimens were initially classified as poorly differentiated NSCLC by H&E staining, and this histological feature confounds diagnosis. Secondly, the number of samples in the validation set was relatively small, which may have biased our statistical analyses. Besides the aforementioned reasons, this inconsistency may also be related to the fact that diagnoses in the validation set were confirmed by using IHC staining in which essential components were TTF1 and P40. Therefore, the final diagnosis would depend to a large extent on TTF1 and P40 expression. We evaluated the diagnostic efficacy of TTF1 and P40 through a direct retrieval of IHC staining results from postoperative pathological reports and contrasted it with that of AGR2 and KRT5. Through this method, the diagnostic accuracies of TTF1 and P40 were significantly improved. Notwithstanding, AGR2-KRT5 was still 8.4% more accurate in ADC samples than TTF1-P40. This further highlights the advantages of utilizing AGR2-KRT5 to distinguish pathological subtypes of poorly differentiated NSCLC, particularly ADCs.

AGR2 and KRT5 protein expression was not significantly associated with OS in patients with NSCLC. Nevertheless, AGR2 negativity was associated with unfavorable prognosis, while the inverse trend was observed for KRT5. These findings suggest that SCC patients have shorter OS than those with ADC, which has been previously demonstrated by us and other authors [19, 43], and indirectly prove the diagnostic accuracy of our markers.

As a result of technical constraints, the small biopsy tissues were unavailable for this study, thereby diminishing the investigation quality. To address this constraint, poorly differentiated NSCLC tissues were chosen as the validation set components to verify the discriminatory capabilities of AGR2-KRT5. We endeavor to conduct a follow-up study by establishing a biopsy specimen repository to demonstrate the differential diagnostic utility of AGR2-KRT5, as well as to clarify its clinical value. Large-scale prospective clinical trials are also warranted to further validate the use of this marker combination.

In conclusion, the AGR2-KRT5 immunomarker combination has the capacity to distinguish ADC and SCC. In poorly differentiated tumor specimens, the diagnostic accuracy of this combination compared well with that of the conventional TTF1-P40 signature that is widely recognized as accurate. Furthermore, AGR2-KRT5 showed better diagnostic performance in ADC compared to TTF1-P40. Nevertheless, an inevitable limitation of IHC is that it is impossible to standardized immunostaining result interpretation. Subsequent studies are required to develop more accurate methodologies for NSCLC subtype classification. Nonetheless, our results provided new directions to reduce the number of IHC markers, and exhibited higher accuracy than TTF1-P40 at equivalent marker densities in ADC. With subsequent prospective research, AGR2-KRT5 has the potential of becoming a new clinical option for distinguishing lung ADC and SCC.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (815-72824, 81672931 and 81872396), CSCO-Merck Serono Cancer Research Fund (Y-MT-2014-015), Haiyan Foundation of Harbin Medical University Cancer Hospital (JJZD2016-01), National Cancer Center Climbing Fund (NCC201808B017), China Postdoctoral Science Foundation (2016M591552) and Hei Long Jiang Postdoctoral Foundation (LBH-Z18196).

Disclosure of conflict of interest

None.

Address correspondence to: Yan Yu, Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China. Tel: +86-451-86298727; Fax: +86-451-86298727; E-mail: yuyan@hrbmu.edu.cn; Li-Shuang Qi, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China. Tel: +86-451-86615933; Fax: +86-

451-86669617; E-mail: qilishuang7@ems.hrbmu.edu.cn

References

[1] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020; 70: 7-30.

[2] Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, Geisinger K, Hirsch FR, Ishikawa Y, Kerr KM, Noguchi M, Pelosi G, Powell CA, Tsao MS and Wistuba I. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. J Thorac Oncol 2015; 10: 1243-1260.

[3] Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, Ishikawa Y, Wistuba I, Flieder DB, Franklin W, Gazdar A, Hasleton PS, Henderson DW, Kerr KM, Petersen I, Roggli V, Thunnissen E and Tsao M. Diagnosis of lung cancer in small biopsies and cytology: implications of the 2011 international association for the study of lung cancer/American thoracic society/European respiratory rociety classification. Arch Pathol Lab Med 2013; 137: 668-684.

[4] Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, Beer DG, Powell CA, Riely GJ, Van Schil PE, Garg K, Austin JH, Asamura H, Rusch VW, Hirsch FR, Scagliotti G, Mitsudomi T, Huber RM, Ishikawa Y, Jett J, Sanchez-Cespedes M, Sculier JP, Takahashi T, Tsuboi M, Vansteenkiste J, Wistuba I, Yang PC, Aberle D, Brambilla C, Flieder D, Franklin W, Gazdar A, Gould M, Hasleton P, Henderson D, Johnson B, Johnson D, Kerr K, Kuriyama K, Lee JS, Miller VA, Petersen I, Roggli V, Rosell R, Saijo N, Thunnissen E, Tsao M and Yankelewitz D. International association for the study of lung cancer/American thoracic society/European respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol 2011; 6: 244-285.

[5] Osmani L, Askin F, Gabrielson E and Li QK. Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): moving from targeted therapy to immunotherapy. Semin Cancer Biol 2018; 52: 103-109.

[6] Chen Z, Fillmore CM, Hammerman PS, Kim CF and Wong KK. Non-small-cell lung cancers: a heterogeneous set of diseases. Nat Rev Cancer 2014; 14: 535-546.

[7] Tian S. Classification and survival prediction for early-stage lung adenocarcinoma and squamous cell carcinoma patients. Oncol Lett 2017; 14: 5464-5470.

[8] Chen X, Yang S and Ma S. Drug induced hepatotoxicity in targeted therapy for lung cancer. Zhongguo Fei Ai Za Zhi 2014; 17: 685-688.

[9] Tan DSW, Yom SS, Tsao MS, Pass HI, Kelly K, Peled N, Yung RC, Wistuba II, Yatabe Y, Unger M, Mack PC, Wynes MW, Mitsudomi T, Weder W, Yankelevitz D, Herbst RS, Gandara DR, Carbone DP, Bunn PA, Mok TS and Hirsch FR. The International Association for the study of lung cancer consensus statement on ptimizing management of EGFR mutation-positive non-small cell lung cancer: status in 2016. J Thorac Oncol 2016; 11: 946-963.

[10] Pao W and Girard N. New driver mutations in non-small-cell lung cancer. Lancet Oncol 2011; 12: 175-180.

[11] Jänne PA, Shaw AT, Camidge DR, Giaccone G, Shreeve SM, Tang Y, Goldberg Z, Martini JF, Xu H, James LP and Solomon BJ. Combined Pan-HER and ALK/ROS1/MET inhibition with dacomitinib and crizotinib in advanced non-small cell lung cancer: results of a phase I study. J Thorac Oncol 2016; 11: 737-747.

[12] Subbiah V, Berry J, Roxas M, Guha-Thakurta N, Subbiah IM, Ali SM, McMahon C, Miller V, Cascone T, Pai S, Tang Z and Heymach JV. Systemic and CNS activity of the RET inhibitor vandetanib combined with the mTOR inhibitor everolimus in KIF5B-RET re-arranged non-small cell lung cancer with brain metastases. Lung Cancer 2015; 89: 76-79.

[13] Zhan C, Yan L, Wang L, Sun Y, Wang X, Lin Z, Zhang Y, Shi Y, Jiang W and Wang Q. Identification of immunohistochemical markers for distinguishing lung adenocarcinoma from squamous cell carcinoma. J Thorac Dis 2015; 7: 1398-1405.

[14] National Comprehensive Cancer Network: NCCN Clinical Practice Guidelines in Oncology: Non-small cell lung cancer, version 3. 2020 https://www.nccn.org/professionals/physician_gls/default.aspx#nscl/; 2020. Accessed 06 May 2020.

[15] Argon A, Nart D and Veral A. The value of cytokeratin 5/6, p63 and thyroid transcription factor-1 in adenocarcinoma, squamous cell carcinoma and non-small-cell lung cancer of the lung. Turk Patoloji Derg 2015; 31: 81-88.

[16] Loo PS, Thomas SC, Nicolson MC, Fyfe MN and Kerr KM. Subtyping of undifferentiated non-small cell carcinomas in bronchial biopsy specimens. J Thorac Oncol 2010; 5: 442-447.

[17] von Laffert M, Warth A, Penzel R, Schirmacher P, Kerr KM, Elmberger G, Schildhaus HU, Buttner R, Lopez-Rios F, Reu S, Kirchner T, Pauwels P, Specht K, Drecoll E, Hofler H, Aust D, Baretton G, Bubendorf L, Stallmann S, Fisseler-Eckhoff A, Soltermann A, Tischler V, Moch H, Penault-Llorca F, Hager H, Schaper F,

Lenze D, Hummel M and Dietel M. Multicenter immunohistochemical ALK-testing of non-small-cell lung cancer shows high concordance after harmonization of techniques and interpretation criteria. J Thorac Oncol 2014; 9: 1685-1692.

[18] Ilie M, Hofman V, Dietel M, Soria JC and Hofman P. Assessment of the PD-L1 status by immunohistochemistry: challenges and perspectives for therapeutic strategies in lung cancer patients. Virchows Arch 2016; 468: 511-525.

[19] Li X, Shi G, Chu Q, Jiang W, Liu Y, Zhang S, Zhang Z, Wei Z, He F, Guo Z and Qi L. A qualitative transcriptional signature for the histological reclassification of lung squamous cell carcinomas and adenocarcinomas. BMC Genomics 2019; 20: 881.

[20] Xiao J, Lu X, Chen X, Zou Y, Liu A, Li W, He B, He S and Chen Q. Eight potential biomarkers for distinguishing between lung adenocarcinoma and squamous cell carcinoma. Oncotarget 2017; 8: 71759-71771.

[21] Fritzsche FR, Dahl E, Dankof A, Burkhardt M, Pahl S, Petersen I, Dietel M and Kristiansen G. Expression of AGR2 in non-small cell lung cancer. Histol Histopathol 2007; 22: 703-708.

[22] Bishop JA, Teruya-Feldstein J, Westra WH, Pelosi G, Travis WD and Rekhtman N. p40 (ΔNp63) is superior to p63 for the diagnosis of pulmonary squamous cell carcinoma. Mod Pathol 2012; 25: 405-415.

[23] Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WEE, Nicholson AG, Groome P, Mitchell A and Bolejack V. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (Eighth) edition of the TNM classification for lung cancer. J Thorac Oncol 2016; 11: 39-51.

[24] Allred DC, Harvey JM, Berardo M and Clark GM. Prognostic and predictive factors in breast cancer by immunohistochemical analysis. Mod Pathol 1998; 11: 155-168.

[25] Meyerholz DK and Beck AP. Principles and approaches for reproducible scoring of tissue stains in research. Lab Invest 2018; 98: 844-855.

[26] Salmans ML, Zhao F and Andersen B. The estrogen-regulated anterior gradient 2 (AGR2) protein in breast cancer: a potential drug target and biomarker. Breast Cancer Res 2013; 15: 204.

[27] Pizzi M, Fassan M, Balistreri M, Galligioni A, Rea F and Rugge M. Anterior gradient 2 overexpression in lung adenocarcinoma. Appl Immunohistochem Mol Morphol 2012; 20: 31-36.

[28] Chung K, Nishiyama N, Wanibuchi H, Yamano S, Hanada S, Wei M, Suehiro S and Kakehashi A. AGR2 as a potential biomarker of human lung adenocarcinoma. Osaka City Med J 2012; 58: 13-24.

[29] Park K, Chung YJ, So H, Kim K, Park J, Oh M, Jo M, Choi K, Lee EJ, Choi YL, Song SY, Bae DS, Kim BG and Lee JH. AGR2, a mucinous ovarian cancer marker, promotes cell proliferation and migration. Exp Mol Med 2011; 43: 91-100.

[30] Li Y, Lu J, Peng Z, Tan G, Liu N, Huang D, Zhang Z, Duan C, Tang X and Tang F. N,N'-dinitrosopiperazine-mediated AGR2 is involved in metastasis of nasopharyngeal carcinoma. PLoS One 2014; 9: e92081.

[31] Bu H, Bormann S, Schafer G, Horninger W, Massoner P, Neeb A, Lakshmanan VK, Maddalo D, Nestl A, Sultmann H, Cato AC and Klocker H. The anterior gradient 2 (AGR2) gene is overexpressed in prostate cancer and may be useful as a urine sediment marker for prostate cancer detection. Prostate 2011; 71: 575-587.

[32] Vanderlaag KE, Hudak S, Bald L, Fayadat-Dilman L, Sathe M, Grein J and Janatpour MJ. Anterior gradient-2 plays a critical role in breast cancer cell growth and survival by modulating cyclin D1, estrogen receptor-alpha and survivin. Breast Cancer Res 2010; 12: R32.

[33] Wang Z, Hao Y and Lowe AW. The adenocarcinoma-associated antigen, AGR2, promotes tumor growth, cell migration, and cellular transformation. Cancer Res 2008; 68: 492-497.

[34] Chu PG and Weiss LM. Expression of cytokeratin 5/6 in epithelial neoplasms: an immunohistochemical study of 509 cases. Mod Pathol 2002; 15: 6-10.

[35] Chen Y, Cui T, Yang L, Mireskandari M, Knoesel T, Zhang Q, Pacyna-Gengelbach M and Petersen I. The diagnostic value of cytokeratin 5/6, 14, 17, and 18 expression in human non-small cell lung cancer. Oncology 2011; 80: 333-340.

[36] Bernardi FDC, Bernardi MDC, Takagaki T, Siqueira SAC and Dolhnikoff M. Lung cancer biopsy: can diagnosis be changed after immunohistochemistry when the H&E-Based morphology corresponds to a specific tumor subtype? Clinics (Sao Paulo, Brazil) 2018; 73: e361.

[37] Micke P, Mattsson JS, Djureinovic D, Nodin B, Jirström K, Tran L, Jönsson P, Planck M, Botling J and Brunnström H. The impact of the fourth edition of the WHO classification of lung tumours on histological classification of resected pulmonary NSCCs. J Thorac Oncol 2016; 11: 862-872.

[38] Bhargava R, Beriwal S, McManus K and Dabbs DJ. CK5 is more sensitive than CK5/6 in identifying the "basal-like" phenotype of breast carcinoma. Am J Clin Pathol 2008; 130: 724-730.

[39] Chang YL, Lee YC, Liao WY and Wu CT. The utility and limitation of thyroid transcription factor-1 protein in primary and metastatic pulmonary neoplasms. Lung Cancer 2004; 44: 149-157.

[40] Yatabe Y, Dacic S, Borczuk AC, Warth A, Russell PA, Lantuejoul S, Beasley MB, Thunnissen E, Pelosi G, Rekhtman N, Bubendorf L, Mino-Kenudson M, Yoshida A, Geisinger KR, Noguchi M, Chirieac LR, Bolting J, Chung JH, Chou TY, Chen G, Poleri C, Lopez-Rios F, Papotti M, Sholl LM, Roden AC, Travis WD, Hirsch FR, Kerr KM, Tsao MS, Nicholson AG, Wistuba I and Moreira AL. Best practices recommendations for diagnostic immunohistochemistry in lung cancer. J Thorac Oncol 2019; 14: 377-407.

[41] Lilo MT, Allison D, Wang Y, Ao M, Gabrielson E, Geddes S, Zhang H, Askin F and Li QK. Expression of P40 and P63 in lung cancers using fine needle aspiration cases. Understanding clinical pitfalls and limitations. J Am Soc Cytopathol 2016; 5: 123-132.

[42] Sharma R, Wang Y, Chen L, Gurda GT, Geddes S, Gabrielson E, Askin F and Li QK. Utility of a novel triple marker (combination of thyroid transcription factor 1, Napsin A, and P40) in the subclassification of non-small cell lung carcinomas using fine-needle aspiration cases. Hum Pathol 2016; 54: 8-16.

[43] Cooke DT, Nguyen DV, Yang Y, Chen SL, Yu C and Calhoun RF. Survival comparison of adenosquamous, squamous cell, and adenocarcinoma of the lung after lobectomy. Ann Thorac Surg 2010; 90: 943-948.

Supplementary Table 1. Information about datasets applied in this study

Data Source	Type	Platform	pSCC	pADC
GSE50081	frozen	Affy. Plus 2.0	43	127
GSE37745	frozen	Affy. Plus 2.0	24	40
GSE14814	frozen	Affy. U133A	26	32
GSE29016	frozen	Illu. HT	13	37
GSE42127	frozen	Illu. WG	33	94
Total	frozen		139	330
GSE58661	biopsy	Merck RSTA	36	42

Notes: frozen, frozen tissues; biopsy, small biopsy specimens; pADC, pathologically-determined ADC; pSCC, pathologically-determined SCC; Affy. Plus 2.0, Affymetrix Plus 2.0; Affy. U133A, Affymetrix U133A; Illu. WG, Illumina Human WG; Illu. HT, Illumina Human HT.

Supplementary Table 2. The protein expression of AGR2 and KRT5 evaluated by IHC for samples in training and validation sets

No.	Sample set	Pathological lable	AGR2		KRT5	
			expression results	IHC scores	expression results	IHC scores
1	Training set	ADC	+	4	-	1
2	Training set	ADC	+	7	-	0
3	Training set	ADC	+	4	-	0
4	Training set	ADC	+	4	-	0
5	Training set	ADC	+	7	-	2
6	Training set	ADC	+	4	-	0
7	Training set	ADC	-	1	-	0
8	Training set	ADC	+	3	-	1
9	Training set	ADC	+	3	-	0
10	Training set	ADC	+	4	-	0
11	Training set	ADC	+	3	-	0
12	Training set	ADC	+	4	-	0
13	Training set	ADC	+	4	-	1
14	Training set	ADC	+	5	-	0
15	Training set	ADC	+	3	-	0
16	Training set	ADC	+	6	-	2
17	Training set	ADC	+	4	-	0
18	Training set	ADC	+	4	-	0
19	Training set	ADC	+	5	-	0
20	Training set	ADC	+	4	-	1
21	Training set	ADC	+	4	-	0
22	Training set	ADC	+	4	-	0
23	Training set	ADC	-	0	-	0
24	Training set	ADC	+	5	-	2
25	Training set	ADC	+	4	-	0
26	Training set	ADC	+	6	-	0
27	Training set	ADC	+	4	-	0
28	Training set	ADC	+	4	-	0
29	Training set	ADC	+	6	-	0
30	Training set	ADC	+	5	-	0

31	Training set	ADC	+	6	-	0
32	Training set	ADC	+	6	-	0
33	Training set	ADC	+	4	-	0
34	Training set	ADC	+	7	-	0
35	Training set	ADC	+	7	-	2
36	Training set	ADC	+	7	-	2
37	Training set	ADC	+	7	-	0
38	Training set	ADC	-	1	-	0
39	Training set	ADC	+	6	-	0
40	Training set	ADC	-	0	-	0
41	Training set	ADC	+	7	-	0
42	Training set	ADC	+	7	-	0
43	Training set	ADC	+	6	-	0
44	Training set	ADC	+	4	-	0
45	Training set	ADC	+	6	-	2
46	Training set	ADC	+	4	-	0
47	Training set	ADC	+	5	-	0
48	Training set	ADC	+	7	-	1
49	Training set	ADC	+	6	-	1
50	Training set	ADC	+	7	-	0
51	Training set	ADC	+	4	-	0
52	Training set	ADC	+	4	-	0
53	Training set	ADC	+	4	-	1
54	Training set	ADC	+	5	-	0
55	Training set	ADC	+	6	-	0
56	Training set	ADC	+	4	-	0
57	Training set	ADC	+	6	-	1
58	Training set	ADC	+	4	-	0
59	Training set	ADC	+	4	-	0
60	Training set	ADC	+	7	+	4
61	Training set	ADC	+	7	-	0
62	Training set	ADC	+	7	-	0
63	Training set	ADC	+	6	-	2
64	Training set	ADC	+	7	-	1
65	Training set	ADC	+	4	-	0
66	Training set	ADC	+	4	-	2
67	Training set	ADC	+	4	-	2
68	Training set	ADC	+	4	-	0
69	Training set	ADC	+	6	-	0
70	Training set	ADC	+	4	-	1
71	Training set	ADC	+	4	-	0
72	Training set	ADC	+	3	-	0
73	Training set	ADC	+	4	-	0
74	Training set	ADC	+	6	-	0
75	Training set	ADC	+	5	-	0
76	Training set	ADC	+	4	-	0
77	Training set	ADC	-	2	-	0
78	Training set	ADC	+	4	-	0
79	Training set	ADC	+	5	-	0
80	Training set	ADC	+	4	-	0

81	Training set	ADC	+	3	-	0
82	Training set	ADC	+	6	-	1
83	Training set	ADC	+	7	-	0
84	Training set	ADC	+	6	-	0
85	Training set	ADC	+	3	-	1
86	Training set	ADC	+	7	-	0
87	Training set	ADC	+	4	-	0
88	Training set	ADC	+	5	-	1
89	Training set	ADC	+	4	-	0
90	Training set	SCC	-	0	+	7
91	Training set	SCC	-	0	+	5
92	Training set	SCC	-	1	+	5
93	Training set	SCC	-	0	+	7
94	Training set	SCC	-	0	+	7
95	Training set	SCC	-	0	+	6
96	Training set	SCC	-	1	+	4
97	Training set	SCC	-	0	+	7
98	Training set	SCC	-	1	-	1
99	Training set	SCC	-	0	-	1
100	Training set	SCC	-	0	-	0
101	Training set	SCC	-	0	+	7
102	Training set	SCC	-	0	+	7
103	Training set	SCC	-	1	+	7
104	Training set	SCC	-	0	+	6
105	Training set	SCC	-	0	+	7
106	Training set	SCC	-	0	+	7
107	Training set	SCC	-	0	+	3
108	Training set	SCC	-	0	+	5
109	Training set	SCC	-	0	+	6
110	Training set	SCC	+	3	-	2
111	Training set	SCC	-	0	+	6
112	Training set	SCC	-	0	+	6
113	Training set	SCC	-	0	+	6
114	Training set	SCC	-	0	+	7
115	Training set	SCC	-	1	+	7
116	Training set	SCC	-	0	+	3
117	Training set	SCC	-	1	+	6
118	Training set	SCC	-	0	+	7
119	Training set	SCC	-	0	+	7
120	Training set	SCC	-	1	-	2
121	Training set	SCC	-	0	+	6
122	Training set	SCC	-	0	+	5
123	Training set	SCC	-	0	+	5
124	Training set	SCC	-	1	+	5
125	Training set	SCC	-	0	+	6
126	Training set	SCC	-	2	+	5
127	Training set	SCC	-	0	+	6
128	Training set	SCC	-	0	+	5
129	Training set	SCC	-	0	+	3
130	Training set	SCC	-	0	+	3

131	Training set	SCC	-	2	+	7
132	Training set	SCC	-	2	+	5
133	Training set	SCC	-	0	+	6
134	Training set	SCC	-	0	+	5
135	Training set	SCC	-	0	+	7
136	Training set	SCC	-	1	+	7
137	Training set	SCC	-	0	+	6
138	Training set	SCC	-	0	+	7
139	Training set	SCC	-	1	+	7
140	Training set	SCC	-	0	+	6
141	Training set	SCC	-	2	+	5
142	Training set	SCC	-	0	+	6
143	Training set	SCC	-	0	+	3
144	Training set	SCC	-	0	+	6
145	Training set	SCC	-	0	+	6
146	Training set	SCC	-	0	+	3
147	Training set	SCC	-	0	+	4
148	Training set	SCC	-	2	+	7
149	Training set	SCC	-	0	+	7
150	Training set	SCC	-	2	+	6
151	Training set	SCC	-	0	+	6
152	Training set	SCC	-	0	+	3
153	Training set	SCC	-	0	-	2
154	Training set	SCC	-	0	+	7
155	Training set	SCC	-	0	+	6
156	Training set	SCC	-	0	+	7
157	Training set	SCC	-	1	+	7
158	Training set	SCC	-	0	+	5
159	Training set	SCC	-	1	+	5
160	Training set	SCC	-	0	+	4
161	Training set	SCC	-	0	+	3
162	Training set	SCC	-	0	+	3
163	Training set	SCC	-	0	+	7
164	Training set	SCC	-	0	+	6
165	Training set	SCC	-	0	+	6
166	Training set	SCC	-	0	+	7
167	Training set	SCC	+	3	+	7
168	Training set	SCC	-	0	+	3
169	Training set	SCC	-	0	+	4
170	Training set	SCC	-	0	+	4
171	Training set	SCC	-	0	+	6
172	Training set	SCC	-	0	+	7
173	Training set	SCC	-	0	+	5
174	Training set	SCC	-	0	+	5
175	Training set	SCC	-	0	+	3
176	Training set	SCC	-	2	+	5
177	Training set	SCC	-	0	+	4
178	Training set	SCC	-	0	+	7
179	Training set	SCC	-	0	+	4
180	Training set	SCC	+	4	+	4

181	Training set	SCC	-	0	+	3
182	Training set	SCC	-	0	+	5
183	Training set	SCC	-	0	+	6
184	Training set	SCC	-	2	+	7
185	Training set	SCC	-	0	+	6
186	Training set	SCC	-	0	+	4
187	Training set	SCC	-	1	+	4
188	Training set	SCC	-	1	+	7
189	Validation set	ADC	+	4	-	0
190	Validation set	SCC	-	0	+	4
191	Validation set	SCC	-	0	+	6
192	Validation set	ADC	+	3	-	0
193	Validation set	ADC	+	5	-	0
194	Validation set	SCC	-	1	+	7
195	Validation set	ADC	+	3	-	0
196	Validation set	SCC	-	0	+	4
197	Validation set	SCC	-	0	+	3
198	Validation set	SCC	-	0	-	1
199	Validation set	ADC	+	5	-	0
200	Validation set	SCC	-	1	+	4
201	Validation set	SCC	+	3	+	4
202	Validation set	ADC	+	6	-	0
203	Validation set	SCC	-	0	+	4
204	Validation set	ADC	+	4	-	0
205	Validation set	SCC	-	2	+	7
206	Validation set	SCC	-	2	+	6
207	Validation set	SCC	-	2	+	4
208	Validation set	SCC	-	0	+	3
209	Validation set	ADC	+	5	-	0
210	Validation set	SCC	-	0	+	5
211	Validation set	ADC	+	7	-	0
212	Validation set	SCC	-	0	+	5
213	Validation set	SCC	-	2	+	3
214	Validation set	ADC	+	6	-	0
215	Validation set	SCC	-	0	+	3
216	Validation set	SCC	-	0	+	4
217	Validation set	SCC	-	1	+	6
218	Validation set	SCC	-	2	+	4
219	Validation set	SCC	-	0	-	2
220	Validation set	SCC	-	1	+	4
221	Validation set	SCC	-	0	+	4
222	Validation set	ADC	+	5	-	0
223	Validation set	SCC	-	2	+	4
224	Validation set	SCC	-	2	+	3
225	Validation set	SCC	-	1	+	3
226	Validation set	ADC	+	5	-	0
227	Validation set	SCC	-	0	+	3
228	Validation set	SCC	-	2	+	6
229	Validation set	SCC	+	3	+	4
230	Validation set	SCC	-	0	+	4

Supplementary Table 3. Sensitivity, specificity, PPV and NPV of IHC markers in the training set (%)

Markers	Sensitivity	Specificity	PPV	NPV
AGR2	94.4% (84/89)	97.0% (96/99)	97.7% (84/87)	95.1% (96/101)
KRT5	93.9% (93/99)	98.9% (88/89)	98.9% (93/94)	93.6% (88/94)

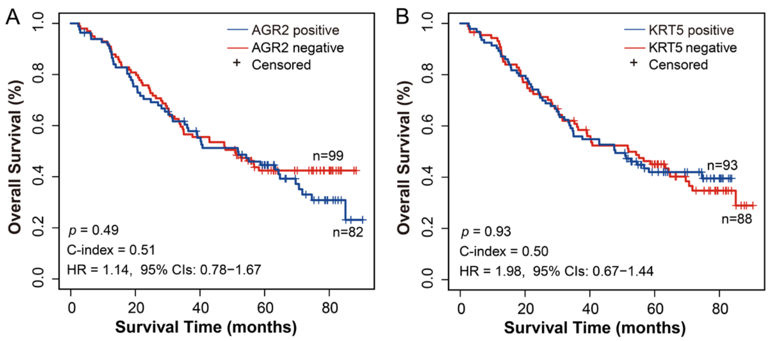
Notes: In this table, we took IHC staining score 3 as the cutoff value. Sensitivity = TP/TP+FN; Specificity = TN/TN+FP; Positive predictive value (PPV) = TP/TP+FP; Negative predictive value (NPV) = TN/TN+FN. FN indicates false negatives; FP, false positives; TN, true negatives; TP, true positives.

Supplementary Table 4. The comparison of two IHC marker combinations in the validation set

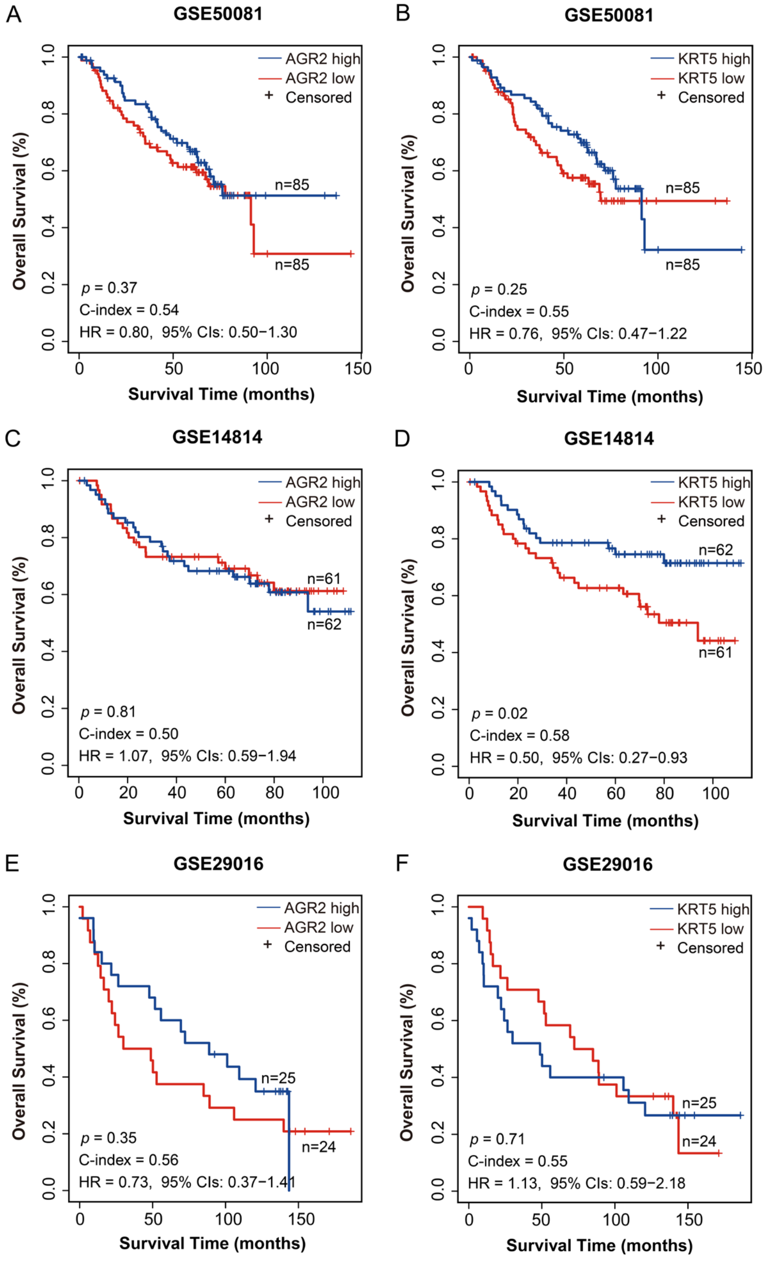
AGR2-KRT5	TTF1-P40		Total
	no. of correct diagnosis	no. of false diagnosis	
no. of correct diagnosis	37	1	38
no. of false diagnosis	1	3	4
Total	38	4	42

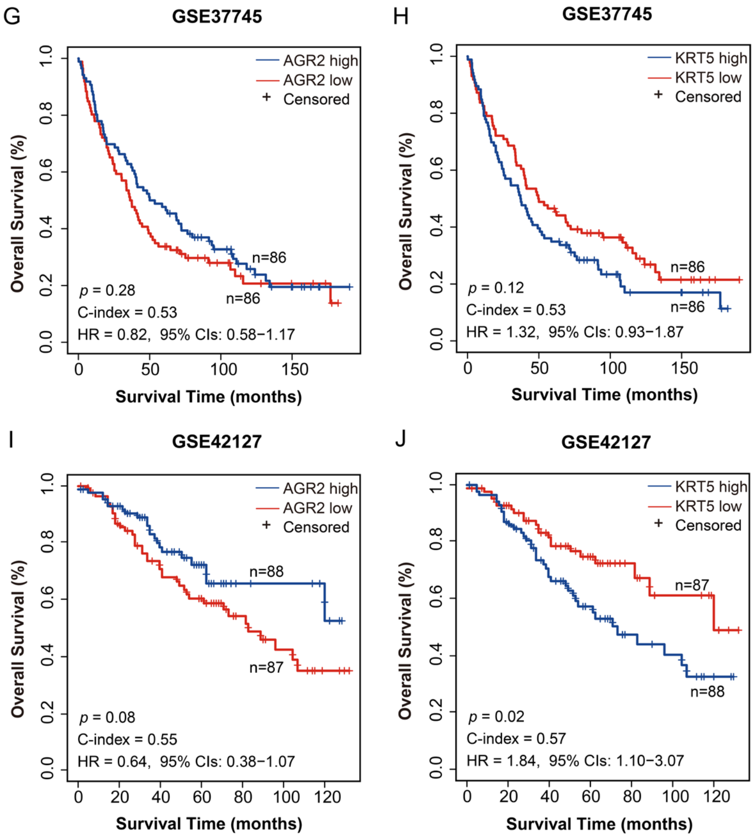
Supplementary Table 5. Associations between marker mRNA expression and clinicopathological parameters in GSE50081

Variable	No. of cases	AGR2		*P*	KRT5		*P*
		No. of pos. (%)	No. of neg. (%)		No. of pos. (%)	No. of neg. (%)	
Patient age							
<60 years	25	8 (32.00%)	17 (68.00%)	0.082	10 (40.00%)	15 (60.00%)	0.387
≥60 years	145	77 (53.10%)	68 (46.90%)		75 (51.72%)	70 (48.28%)	
Gender							
Male	90	51 (56.67%)	39 (43.33%)	0.091	44 (48.89%)	46 (51.11%)	0.878
Female	80	34 (42.50%)	46 (57.50%)		41 (51.25%)	39 (48.75%)	
Smoking history							
Yes	126	64 (50.79%)	62 (49.21%)	0.967	65 (51.59%)	61 (48.41%)	0.791
No	24	11 (45.83%)	13 (54.17%)		11 (45.83%)	13 (54.17%)	
Unable to determine	20	10 (50.00%)	10 (50.00%)		9 (45.00%)	11 (55.00%)	
Histology							
ADC	127	74 (58.27%)	53 (41.73%)	<0.001	46 (36.22%)	81 (63.78%)	<0.001
SCC	43	11 (25.58%)	32 (74.42%)		39 (90.70%)	9.30%	
pT-status							
pT1-2	168	85 (50.60%)	83 (49.40%)	0.497	85 (50.60%)	83 (49.40%)	0.497
pT3-4	2	0	2 (100%)		0	2 (100%)	
pN-status							
pN0-1	170	85 (50.00%)	85 (50.00%)	1	85 (50.00%)	85 (50.00%)	1
pN2-3	0	0	0		0	0	
TNM stage							
I-II	170	85 (50.00%)	85 (50.00%)	1	85 (50.00%)	85 (50.00%)	1
III-IV	0	0	0		0	0	



Supplementary Figure 1. Survival curves of training set based on AGR2 and KRT5 protein expression. Survival curves of positive and negative expressions for AGR2 (A) and KRT5 (B).





Supplementary Figure 2. Survival curves of patients with expression of AGR2 and KRT5 in GEO datasets. Survival curves of overall survival (OS) accordingly for the high and low AGR2 expression groups in GSE50081 (A), GSE14814 (C), GSE29016 (E), GSE37745 (G) and GSE42127 (I). Survival curves of OS respectively for the diverse expression groups of KRT5 in GSE50081 (B), GSE14814 (D), GSE29016 (F), GSE37745 (H) and GSE42127 (J). The high and low expression groups of marker genes were categorized according to the median of the gene expression.