

ORIGINAL RESEARCH



MiRNA expression patterns are associated with tumor mutational burden in lung adenocarcinoma

Yufeng Lv^{a*}, Zhong Huang^{a*}, Yan Lin^b, Yuan Fang^a, Zhichao Chen^a, Lili Pan^a, Yanxian Zhang^a, and Zihai Xu^a

^aDepartment of Oncology, Langdong Hospital of Guangxi Medical University,, Nanning, P. R. China; ^bDepartment of Medical Oncology, Affiliated Tumor Hospital of Guangxi Medical University, Nanning, P. R. China

ABSTRACT

Background: Tumor mutational burden (TMB) has emerged as an independent biomarker to predict patient responses to treatment with immune checkpoint inhibitors (ICIs) for lung adenocarcinoma (LUAD). MicroRNAs (miRNAs) have a crucial role in the regulation of anticancer immune responses, but the association of miRNA expression patterns and TMB is not clear in LUAD.

Methods: Differentially expressed miRNAs in samples with high TMB and low TMB samples were screened in the LUAD dataset in The Cancer Genome Atlas. The least absolute shrinkage and selection operator (LASSO) method was applied to develop a miRNA-based signature classifier for predicting TMB levels in the training set. An test set was used to validate this classifier. The correlation between the miRNA-based classifier index and the expression of three immune checkpoints (PD-1, PD-L1, and CTLA-4) were explored. Functional enrichment analysis was carried out of the miRNAs included in the miRNA-based signature classifier.

Results: Twenty-five differentially expressed miRNAs were used to establish a miRNA-based signature classifier for predicting TMB level. The accuracy of the 25-miRNA-based signature classifier was 0.850 in the training set, 0.810 in the test set and 0.840 in the total set. This miRNA-based signature classifier index showed a low correlation with PD-1 and PD-L1, and no correlation with CTLA-4. Enrichment analysis for these 25 miRNA revealed they are involved in many immune-related biological processes and cancer-related pathways.

Conclusion: MiRNA expression patterns are associated with tumor mutational burden and a miRNA-based signature classifier may serve as a biomarker for prediction of TMB levels in LUAD.

ARTICLE HISTORY

Received 10 April 2019
Revised 13 May 2019
Accepted 2 June 2019

KEYWORDS

Tumor mutational burden (TMB); microRNA; non-small cell lung cancer (NSCLC); immunotherapy; lung adenocarcinoma

Introduction

Lung cancer is the most frequently diagnosed cancer and the leading cause of cancer death.¹ Non-small cell lung cancer (NSCLC) represents 85% of all lung cancers, and about 40%–50% of these cancers are lung adenocarcinoma (LUAD).^{2,3} More than 60% of lung cancer patients present with locally advanced or metastatic disease at the time of diagnosis, and surgical resection may not be an option at this stage. Conventional chemotherapy and radiation therapy have been the mainstays of treatment for patients with advanced NSCLC for the past decades.⁴ Within the last decade, targeted therapy based on corresponding driver gene alterations has improved clinical outcomes for certain subsets of NSCLC patients,^{5–10} however, the 5-year survival rate remains less than 20%.^{2,3} In recent years, the introduction of immune checkpoint inhibitors (ICIs) for the treatment of NSCLC has revolutionized treatment algorithms for advanced NSCLC patients and has massively improved survival rates for certain groups of patients.¹¹ The most widely studied checkpoints are the programmed death protein 1 (PD-1)/programmed death receptor ligand 1 (PD-L1) and cytotoxic T-lymphocyte antigen 4 (CTLA-4), and PD-1/PD-L1 inhibitors have been approved by the United States Food and Drug

Administration (FDA) for the treatment of advanced NSCLC.¹² PD-L1 expression can be detected by immunohistochemistry and is used as a biomarker for PD-1/PD-L1 inhibitor therapy to assist in predicting treatment response,¹² however, there is a general need to better identify responders, as only 25–30% of patients under checkpoint treatment show long-term responses and these might not be exclusively identified by PD-L1 expression.^{13–17}

Though it is still debated and randomized trials are needed, tumor mutational burden (TMB) is promising as another effective predictive biomarker for treatment with ICIs and independent to PD-L1 expression.^{18,19} TMB measures the number of somatic mutations within a tumor and was initially assessed by whole exome sequencing (WES),²⁰ and several targeted sequencing panels have recently been developed to efficiently determine TMB,^{21–23} however, the limited amounts of tumor DNA obtained by conventional or fine needle biopsy can make TMB assessment challenging or even impossible due to these large sequencing panels are required larger amounts of tumor DNA.²⁴ High TMB can lead to modifications of the proteins encoded by the mutated genes. The modified proteins may be recognised by the immune system as “non-self” and activate specific

anti-tumor immune responses.^{25,26} The translation of the mutated gene into a modified protein requires post-transcriptional regulation, and microRNAs (miRNAs) are important molecules involved in post-transcriptional regulation.

MiRNAs are small (~21 nucleotides in length) endogenous noncoding RNA molecules, and aberrant expression of miRNAs is often found in cancer.²⁷ As hallmarks of cancer were described,²⁸ miRNAs were included in the regulation of various cancer hallmarks.²⁹ Several studies have suggested miRNAs may be promising outcome predictors for various types of cancers.³⁰ Recently, there has been increasing interest in the role of miRNAs in the regulation of anticancer immune responses.³¹ It is elucidated that miRNAs are involved in mediating and controlling several immune and cancer cell interactions.³² The baseline profile of miRNAs and their dynamic change might be correlative with the efficacy of immunotherapy in advanced NSCLC.³³ Therefore, we hypothesized that different TMB levels may be reflected in the expression pattern of miRNAs, and that miRNA expression patterns could be used as biomarker for predicting TMB levels. To explore our hypothesis, we downloaded the datasets of lung adenocarcinoma (LUAD) from TCGA, including mutation annotation files and miRNA expression profiles, and established a miRNA-based signature classifier for predicting TMB levels.

Materials and methods

Data processing

The mutation annotation files (aligned against the genome of reference GRCh38) for LUAD in TCGA (<https://gdc.cancer.gov/>) were downloaded using the *GDCquery_Maf* function of TCGAAbiolinks package³⁴ in R. The somatic mutation calling workflow used is the MuTect2 pipeline.³⁵ The *read.maf* function was used to read the somatic variants of each sample. Tumor mutational burden was defined as the number of somatic variants per megabase (MB) of genome.¹⁸ A high TMB level was defined as ≥ 10 mutations per MB, and a low TMB level was defined as < 10 mutations per MB.^{18,36} We used 38 MB as the estimate of the exome size.²⁰ The preprocessed LUAD mature miRNA expression profiles in TCGA database, displayed as log₂ converted reads per million (log₂ (RPM + 1)) were downloaded from the UCSC Xena database (<https://xena.ucsc.edu/public>, dataset ID: TCGA.LUAD.sampleMap/miRNA_HiSeq_gene). The miRNA expression profiles contained 495 samples (450 LUAD tissue samples and 45 matched healthy lung tissue samples) based on the IlluminaHiSeq_miRNASeq platform (Illumina Inc., San Diego, CA, USA). A total of 444 samples with both miRNA expression profiles and mutation annotation files were included in this analysis. These samples were randomly assigned to the training set (60%) and test set (40%). The workflow of the present study was shown in Figure 1.

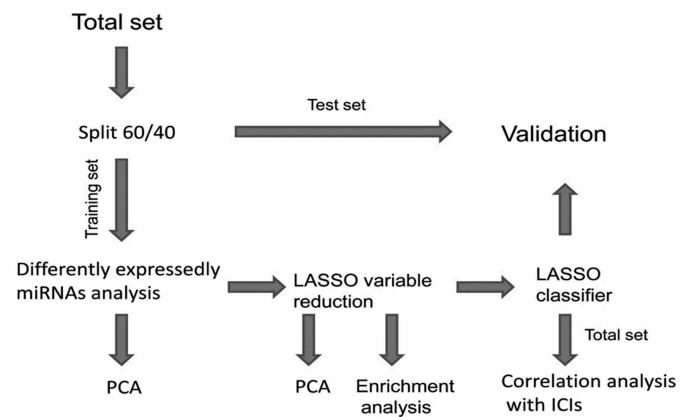


Figure 1. The workflow of the present study.

Screening of differentially expressed miRNAs and bidirectional hierarchical clustering

The miRNAs that were not expressed in $> 10\%$ of the LUAD samples were removed from the training set. The differentially expressed miRNAs between high TMB samples and low TMB samples were analyzed using the “limma” package³⁷ in R. The fold changes (FCs) in the expression of individual miRNAs were calculated, and differentially expressed miRNAs with $|\log_2\text{FC}| > 0.263$ and P (adjusted by false discovery rate) value < 0.01 were considered significant. We performed bidirectional hierarchical clustering³⁸ to these differentially expressed miRNAs based on Euclidean distance and displayed the results as a heat map.

Principal component analysis (PCA) prior to and after least absolute shrinkage and selection operator (LASSO) method feature selection

In the training set, we extracted the expression values of differentially expressed miRNAs for each LUAD sample. The LASSO method with a powerful predictive value and a low correlation between each other to prevent over-fitting³⁹ was applied to select optimal features for the high-dimensional data. The LASSO logistic regression model analysis was performed using the “glmnet” package⁴⁰ in R. The LASSO method was used to select the optimal biomarkers for predicting TMB level. PCA was performed prior to feature selection using the expression profiles of all differentially expressed miRNAs. PCA was subsequently performed using the expression profiles of the optimal differentially expressed miRNAs. Samples were plotted in two-dimensional plots across the first two principal components.

The miRNA-based signature classifier for predicting TMB level

Using the LASSO method, the miRNAs identified with non-zero regression coefficients as optimal miRNAs were used to establish a miRNA-based signature classifier for predicting TMB level. A classifier index for each sample was created using the regression coefficients from the LASSO analysis to

weight the expression value of the selected miRNAs with the following formula:

$$\text{index} = \text{Exp}_{\text{miRNA1}} * \text{Coef1} + \text{Exp}_{\text{miRNA2}} * \text{Coef2} + \text{Exp}_{\text{miRNA3}} * \text{Coef3} + \dots$$

The “Coef” is the regression coefficient of miRNA and is derived from the LASSO Cox regression, and “Exp” indicates the expression values of the miRNAs. The test set was used to validate the robustness and transferability of the classifier. The efficiency of the classifier was assessed by accuracy, sensitivity (Se), specificity (Sp), positive predictive value (PPV), negative predictive value (NPV), and area under the receiver operating characteristic (ROC) curve. These ROC curves were drawn and compared using the “pROC” package⁴¹ in R.

The correlation between the miRNA-based signature classifier index and the expression of three immune checkpoints and functional enrichment analysis

Previous study showed that TMB is independent of PD-L1 expression.⁴² The RNA-sequencing expression profiles (displayed as read counts) were downloaded from TCGA, and the expression values were normalized using VOOM⁴³ function from limma package in R. We explored the correlations between the miRNA-based signature classifier index and the expression of three immune checkpoints (PD-1, PD-L1, and CTLA-4). The starBase⁴⁴ online tool was used to check whether these three immune checkpoints are target genes of any of these miRNAs. In addition, DIANA-mirPath web-server⁴⁵ was used to perform KEGG pathway and gene ontology (GO)⁴⁶ enrichment analysis for these miRNAs. Using the DIANA-mirPath web-server the experimentally validated miRNA-gene interactions derived from TarBase 7.0⁴⁷ were utilized in the present study. KEGG pathways and GO terms with P value < .01 were considered to be significantly enriched.

Statistical analysis

The χ^2 -test was used for categorical data and was applied using IBM SPSS Statistics software version 22.0 (IBM,

Table 1. Summary of patient cohort information.

Characteristic	Training set		Test set		P
	Numbers	%	Numbers	%	
Sex					
Male	125	46.82%	82	46.33%	0.919
Female	142	53.18%	95	53.67%	
Age					
<65 year	122	45.69%	74	41.81%	0.426
≥65 year	136	50.94%	93	52.54%	
Not available	9	3.37%	10	5.65%	
T					
T1-2	236	88.39%	149	84.18%	0.432
T3-4	30	11.24%	26	14.69%	
Tx	1	0.37%	2	1.13%	
N					
N0	177	66.29%	113	63.84%	0.064
N1-3	88	32.96%	57	32.20%	
Nx/Not available	2	0.75%	7	3.95%	
M					
M0	165	61.80%	118	66.67%	0.359
M1	14	5.24%	5	2.82%	
Mx/Not available	88	32.96%	54	30.51%	
Stage					
I-II	211	79.03%	137	77.40%	0.182
III-IV	55	20.60%	36	20.34%	
Not available	1	0.37%	4	2.26%	

Armonk, NY, USA). We analyzed the expression levels of the miRNAs in the high TMB and low TMB group samples using unpaired t-tests provided by limma package. We considered P value <.01 to be statistically significant.

Results

Differentially expressed miRNAs and bidirectional hierarchical clustering

There were no significant differences in routine clinicopathological characteristics between the training and test sets (Table 1). The training set included 66 samples with high TMB level and 201 samples with low TMB level. According to the cutoff criteria ($P < .01$ and $|\log_2\text{FC}| > 0.263$), 49 miRNAs were differentially expressed between the high TMB level and low TMB level samples. These included 44 upregulated miRNAs and 5 downregulated miRNAs in the high TMB level samples. The results of the expression analysis are presented as a heat map (Figure 2),

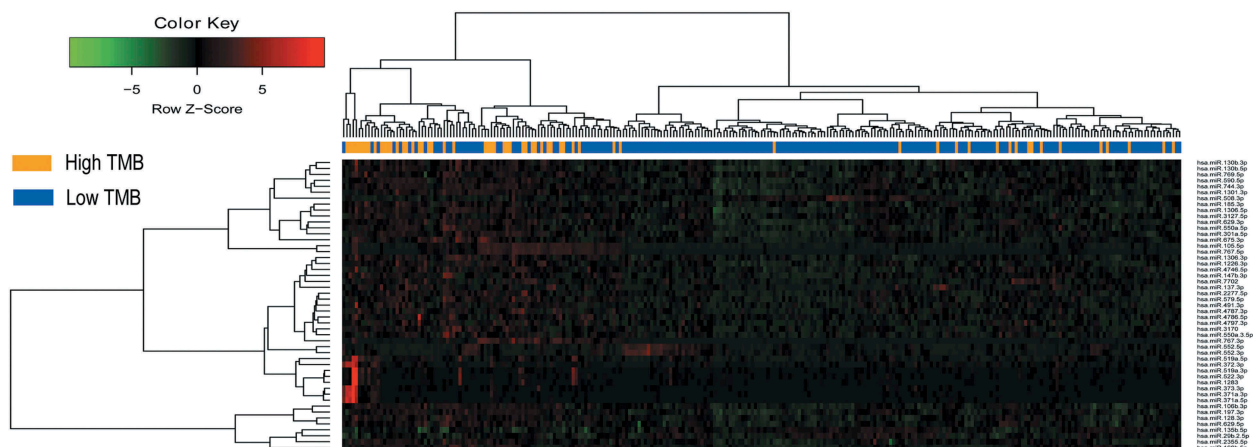


Figure 2. Hierarchical clustering dendrograms of the expression patterns of differentially expressed microRNAs that can basically distinguish between high TMB and low TMB in lung adenocarcinoma.

and the results of hierarchical clustering showed that the expression patterns of these differentially expressed miRNAs could basically distinguish TMB high and TMB low level samples.

PCA and feature selection using LASSO method

For developing a miRNA-based signature classifier for TMB level in LUAD, LASSO logistic regression method was carried out using the expression data of the 49 miRNAs in the training set. We computed the group-wise classifications in 10-fold cross-validations and $\text{type.measure} = \text{"auc"}$ is for two-class logistic regression which gives the AUC curve. Twenty-five miRNAs were identified with non-zero regression coefficients (Figure 3(a)) as optimal features. These 25 miRNAs are miR-767-5p, miR-372-3p, miR-675-3p, miR-508-3p, miR-519a-5p, miR-552-3p, miR-147b-3p, miR-137-3p, miR-7702, miR-2355-5p, miR-106b-3p, miR-371a-3p, miR-29b-2-5p, miR-550a-5p, miR-185-3p, miR-3127-5p, miR-197-3p, miR-769-5p, miR-491-3p, miR-128-3p, miR-1226-3p, miR-2277-5p, miR-4787-3p, miR-550a-3-5p, and miR-4797-3p. Figure 3(b) presents the results of PCA using all 49 differently expressed miRNAs and Figure 3(c) presents the results of PCA using these 25 miRNAs identified by LASSO methods. As demonstrated in Figure 3(c), samples with different TMB level are more easily distinguished using the 25 miRNAs.

The LASSO logistic regression classifier

Using the LASSO method and 10-fold cross-validation, 25 miRNAs were identified with non-zero regression coefficients, and the value of $\text{lambda.min} = 0.008563275$. The miRNA-based classifier index was created as the following formula:

$$\begin{aligned} \text{index} = & \text{miR-767-5p}^*(-0.04755288) + \text{miR-372-3p}^* \\ & (-0.07467899) + \text{miR-675-3p}^*(-0.08205335) + \text{miR-508-} \\ & \text{3p}^*0.20833744 + \text{miR-519a-5p}^*(-0.05549044) + \text{miR-552-} \\ & \text{3p}^*0.49432780 + \text{miR-147b-3p}^*(-0.18427801) + \text{miR-137-3p}^* \\ & (-0.06280235) + \text{miR-7702}^*(-0.21732314) + \text{miR-2355-5p}^* \\ & (-0.29639798) + \text{miR-106b-3p}^*0.06977407 + \text{miR-371a-3p}^* \\ & (-0.21213357) + \text{miR-29b-2-5p}^*0.56165110 + \text{miR-550a-} \\ & \text{5p}^*0.35626234 + \text{miR-185-3p}^*(-0.12431736) + \text{miR-3127-5p}^* \\ & (-0.05386972) + \text{miR-197-3p}^*0.17756353 + \text{miR-769-5p}^* \end{aligned}$$

$$\begin{aligned} & (-0.22362837) + \text{miR-491-3p}^*(-0.57850865) + \text{miR-128-} \\ & \text{3p}^*0.43507960 + \text{miR-1226-3p}^*(-0.21905006) + \text{miR-2277-5p}^* \\ & (-0.37892148) + \text{miR-4787-3p}^*(-0.12085059) + \text{miR-550a-3-5p}^* \\ & (-0.65585909) + \text{miR-4797-3p}^*(-0.45563723). \end{aligned}$$

The accuracy of the 25-miRNA-based classifier was 0.850 in the training set, 0.810 in the test set, and 0.840 in the total set. Based on accuracy, Se, Sp, PPV, NPV and AUC values (Figure 4(a,b)), the sample recognition efficiency of the classifier was high (Table 2). In particular, this classifier has a high specificity. ROC curve analysis revealed that the AUC was 0.895 in the training set and 0.826 in the test set, and their difference was not significant (Delong method⁴⁸ $P = .125$, Figure 4(a)).

Low correlation of the classifier index and three ICIs and enrichment analysis

In the total set, the classifier indexes of all samples were calculated and the correlations of this classifier index with TMB and the expression of three immune checkpoints (PD-1, PD-L1, and CTLA-4) were estimated. Unsurprisingly, as a classifier for predicting TMB, this miRNA-based classifier index showed a high correlation with TMB (Pearson $R = -0.595$, $P = 7.46e-44$, Figure 4(c)), very low correlation with PD-1 (Pearson $R = -0.144$, $P = 2.48e-03$, Figure 4(d)) and PD-L1 (Pearson $R = 0.230$, $P = 1.09e-06$, Figure 4(e)), and no correlation with CTLA-4 (Pearson $R = -0.034$, $P = .473$, Figure 4(f)). Interestingly, according to starBase, only PD-L1 was targeted by 4 miRNAs (miR-372-3p, miR-137-3p, miR-371a-3p, and miR-128-3p), and neither PD-1 nor CTLA4 are target genes of these 25 miRNAs. The results of enrichment analysis for these 25 miRNA shown they involved in many immune-related biological processes (Figure 5(a)) and cancer-related pathways (Figure 5(b)), suggesting that these miRNAs are involved in cancer-related immune processes.

Discussion

TMB has emerged as an independent biomarker to predict patient responses to treatment with ICIs,^{18,19,49,50} and the negative predictive role of low TMB could not be overcome by the

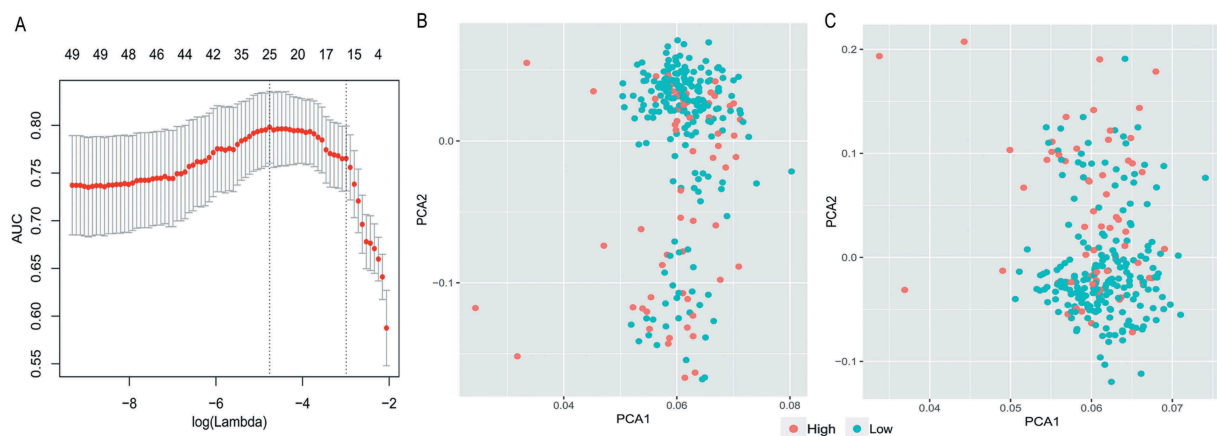


Figure 3. LASSO model and principal component analysis. (a) 10-fold cross-validation for tuning parameter selection in the LASSO model. (b) PCA prior to and (c) after LASSO variable reduction. LASSO, least absolute shrinkage and selection operator; PCA, principal component analysis.

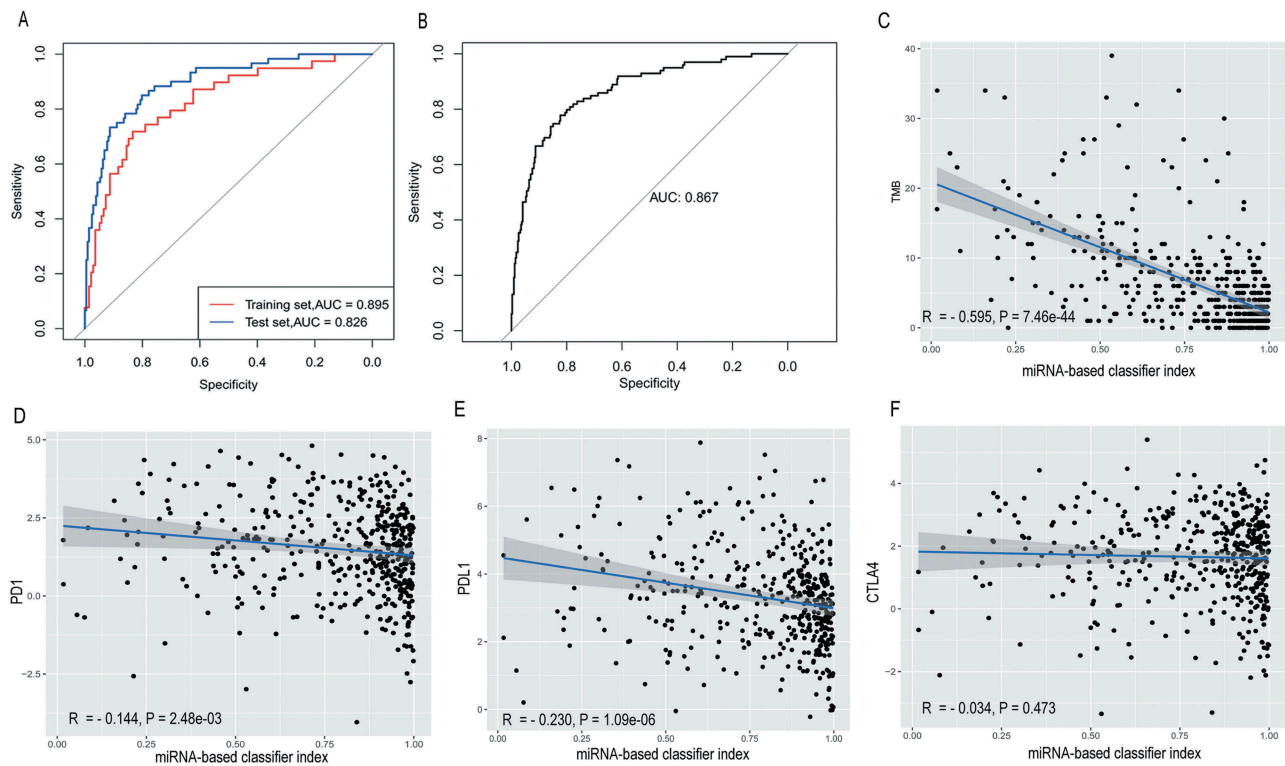


Figure 4. Receiver operating characteristic curves for the 25-miRNA-based signature index and its correlation with TMB, PD-1, PD-L1 and CTLA-4. (a) Receiver operating characteristic analyses in the training and the test set. (b) Receiver operating characteristic analyses in the total set. (c) The 25-miRNA-based signature index is highly correlated with TMB (d) The 25-miRNA-based signature index shows low correlation with PD-1 expression. (e) The 25-miRNA-based signature index shows low correlation with PD-L1 expression. (f) The 25-miRNA-based signature index is not correlated with CTLA4 expression.

Table 2. Performance of 25-miRNA-based classifiers of tumor mutation burden in lung adenocarcinoma.

Cohort	Se	Sp	PPV	NPV	Accuracy	AUC
Training set	0.770	0.960	0.500	0.960	0.850	0.895
Test set	0.670	0.960	0.260	0.960	0.810	0.826
Total set	0.740	0.960	0.400	0.960	0.840	0.867

AUC, area under the receiver operating characteristic curve; NPV, negative predictive value; PPV, positive predictive value; Se, sensitivity; Sp, specificity.

combination immunotherapy.⁵⁰ Several studies have shown to effectively measure TMB from liquid biopsies/blood,⁵¹⁻⁵⁴ this might be an alternative to the biopsies since it is less invasive and easily repeatable. However, TMB assessment by liquid biopsy must face the problem that the circulating DNA deriving from tumor cells is often only a small fraction of the circulating cell free DNA and it needs further investigation

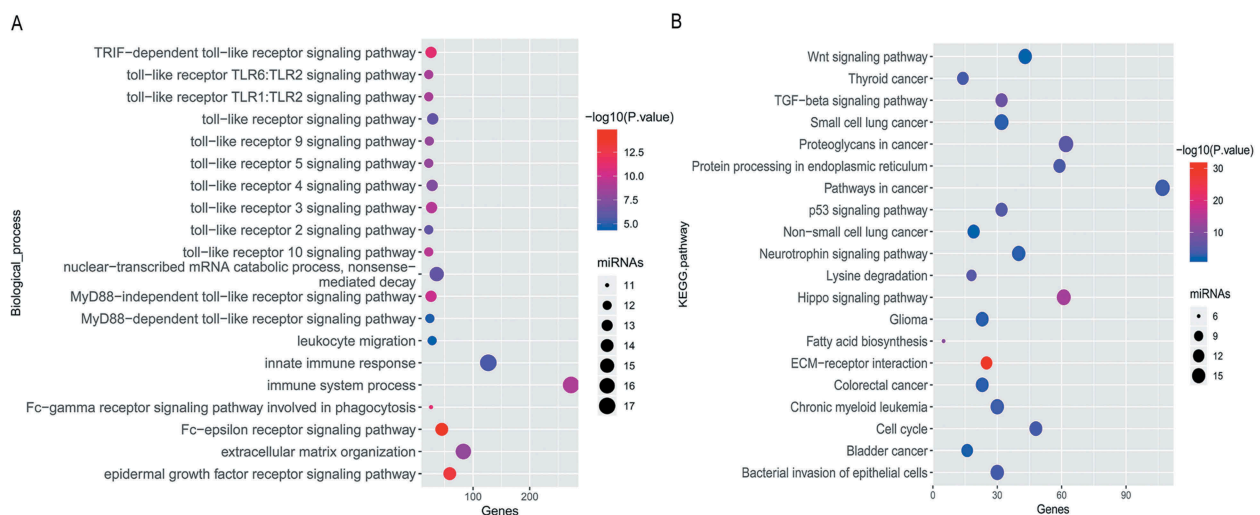


Figure 5. Enrichment analysis of the 25 miRNAs. (a) Significantly enriched immune-related biological process GO terms and (b) Significantly enriched cancer-related KEGG pathways.

and may be clarified within the ongoing clinical trials.⁵⁵ Given the crucial role of miRNAs in tumor-related immune responses,^{33,56} a plasma immune-related miRNA-signature classifier was established that could supplement PD-L1 tumor expression to identify a subgroup of patients with advanced NSCLC.⁷ The association of miRNA expression patterns and TMB was not previously described. In the present study, the differently expressed miRNA between high TMB level and low TMB level samples were identified, and their expression patterns could basically distinguish high TMB level and low TMB level samples. This suggests that changes in genomics may partially lead to changes in transcriptomics. After feature selection, a miRNA-based signature classifier was established in the training set and then validated in an independent test set. The accuracy of the 25-miRNA-based classifier was 0.850 in the training set, 0.810 in the test set and 0.840 in total set. ROC curve analysis revealed that the AUC was 0.895 in the training set and 0.826 in the test set, and this difference was not significant ($P = .125$), indicating that the classifier is robust. In particular, this classifier has strong recognition ability for low TMB, with a high specificity and NPV. However, this classifier has a low PPV, which means that this classifier has a poor recognition ability for high TMB.

For treatment with ICIs, the strongest responses are seen in patients with both high TMB and high PD-L1 expression. It is currently unclear if TMB might be a good addition to existing PD-L1 expression analyses or if TMB level might be able to completely replace PD-L1 testing. It is now widely speculated that high TMB levels lead to an increase in tumor neoantigens may trigger the immune system to attack the tumor.²⁵ The present study suggested that various immune-related miRNAs are differentially expressed between tumors with different TMB levels. The enrichment analysis for the miRNA-signature classifier suggested the 25 miRNAs are involved in immune-related biological processes, such as the “TLR signaling pathway”, leukocyte migration, Fc-gamma receptor signaling pathway involved in phagocytosis. These 25 miRNAs are also involved in cancer-related pathways, including “non-small cell lung cancer”. Although further experimental validation is required, these results suggested that the 25-miRNA-based classifier predicts TMB levels in a biological perspective may be feasible. A study with small sample sizes reported that the expression of plasma exosomal miRNAs might be correlative with the efficacy of immunotherapy in EGFR/ALK wild type advanced NSCLC.³³ In addition, this miRNA-based signature classifier index has a very low correlation with PD-1 and PD-L1, and no correlation with CTLA4.

Although the present study provides a potential surrogate signature for TMB low LUADs, there are some limitations. First, the threshold to distinguish TMB levels may vary for different methods. Second, this classifier has poor recognition ability for high TMB, it may still need to be further confirmed by other methods when a patient is predicted by this classifier to be with high TMB level. Third, further studies are required to validate and even improve the 25-miRNA-based signature classifier in a larger independent cohort of patients. In addition, the molecular mechanisms of these 25 miRNAs in cancer-related immune responses are not yet clear, it is not clear whether these miRNAs are causal or merely markers for response to TMB in LUAD.

Conclusion

In conclusion, LUADs with different TMB levels have different miRNA expression patterns. We established a miRNA-based signature classifier that may be served as biomarker to predict TMB levels in LUAD.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Youth Science Foundation of Guangxi Medical University (Grant number: GXMUYSF 201716), and Guangxi Natural Science Foundation (Grant number: 2018GXNSFBA281091 and 2018GXNSFAA281091).

References

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015;65:87–108. doi:10.3322/caac.21262.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin.* 2017;67:7–30. doi:10.3322/caac.21387.
3. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, et al. 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. doi:10.1097/JTO.0000000000000630.
4. Osmani L, Askin F, Gabrielson E, 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. doi:10.1016/j.semcancer.2017.11.019.
5. Tsao MS, Sakurada A, Cutz J-C, Zhu C-Q, Kamel-Reid S, Squire J, Lorimer I, Zhang T, Liu N, Daneshmand M, et al. Erlotinib in lung cancer - molecular and clinical predictors of outcome. *N Engl J Med.* 2005;353:133–144. doi:10.1056/NEJMoa050736.
6. Mitsudomi T, Yatabe Y. Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer. *Cancer Sci.* 2007;98:1817–1824. doi:10.1111/j.1349-7006.2007.00607.x.
7. Soda M, Schmid KJ, Laouéillé-Duprat S, Pien S, Escobar-Restrepo J-M, Baroux C, Gagliardini V, Page DR, Wolfe KH, Grossniklaus U. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature.* 2007;448:561–566. doi:10.1038/nature05945.
8. Paez JG, Jänne PA, Lee JC, Tracy S, Greulich H, Gabriel S, Herman P, Kaye FJ, Lindeman N, Boggon TJ, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science.* 2004;304:1497–1500. doi:10.1126/science.1099314.
9. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014;511:543–550. doi:10.1038/nature13385.
10. Rehman SSU, Ramalingam SS. Metastatic Lung Cancer: emerging Therapeutic Strategies. *Semin Respir Crit Care Med.* 2016;37:736–749. doi:10.1055/s-0036-1592111.
11. Horn L, Spigel DR, Vokes EE, Holgado E, Ready N, Steins M, Poddubskaya E, Borghaei H, Felip E, Paz-Ares L, et al. Nivolumab versus docetaxel in previously treated patients with advanced non-small-cell lung cancer: two-year outcomes from two randomized, open-label, phase III trials (CheckMate 017 and CheckMate 057). *J Clin Oncol.* 2017;35:3924–3933. doi:10.1200/JCO.2017.74.3062.

12. Kazandjian D, Mikropoulos C, Bancroft EK, Dadaev T, Goh C, Taylor N, Saunders E, Borley N, Keating D, Page EC, et al. FDA approval summary: nivolumab for the treatment of metastatic non-small cell lung cancer with progression on or after platinum-based chemotherapy. *Oncologist*. 2016;21:634–642. doi:10.1634/theoncologist.2015-0507.
13. Olaussen KA, Postel-Vinay S, Li BT, Connell LC, Schindler K, Lacouture ME, Postow MA, Wolchok JD. Predictors of chemotherapy efficacy in non-small-cell lung cancer: a challenging landscape. *Ann Oncol*. 2016;27:2004–2016. doi:10.1093/annonc/mdw321.
14. Ventola CL. Cancer immunotherapy, part 3: challenges and future trends. *P T*. 2017;42:514–521.
15. Sorensen SF, Zhou W, Dolled-Filhart M, Georgsen JB, Wang Z, Emancipator K, Wu D, Busch-Sørensen M, Meldgaard P, Hager H. PD-L1 expression and survival among patients with advanced non-small cell lung cancer treated with chemotherapy. *Transl Oncol*. 2016;9:64–69. doi:10.1016/j.tranon.2016.01.003.
16. Sun JM, Zhou W, Choi Y-L, Choi S-J, Kim SE, Wang Z, Dolled-Filhart M, Emancipator K, Wu D, Weiner R, et al. Prognostic significance of PD-L1 in patients with non-small cell lung cancer: a large cohort study of surgically resected cases. *J Thorac Oncol*. 2016;11:1003–1011. doi:10.1016/j.jtho.2016.04.007.
17. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, Gottfried M, Peled N, Tafreshi A, Cuffe S, et al. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N Engl J Med*. 2016;375:1823–1833. doi:10.1056/NEJMoa1606774.
18. Hellmann MD, Chizema-Kawesha E, Blake A, Chew O, Mwaba J, Zulu G, Poncin M, Rakesh A, Page A-L, Stoitsova S, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N Engl J Med*. 2018;378:2093–2104. doi:10.1056/NEJMoa1801946.
19. Carbone DP, Reck M, Paz-Ares L, Creelan B, Horn L, Steins M, Felip E, van Den Heuvel MM, Ciuleanu T-E, Badin F, et al. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N Engl J Med*. 2017;376:2415–2426. doi:10.1056/NEJMoa1613493.
20. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, Schrock A, Campbell B, Shlien A, Chmielecki J, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med*. 2017;9:34. doi:10.1186/s13073-017-0424-2.
21. Goodman AM, Jovcheva E, Mevellec L, Vialard J, De Lange D, Verhulst T, Paulussen C, Van De Ven K, King P, Freyne E, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther*. 2017;16:2598–2608. doi:10.1158/1535-7163.MCT-17-0386.
22. Garofalo A, Sholl L, Reardon B, Taylor-Weiner A, Amin-Mansour A, Miao D, Liu D, Oliver N, MacConaill L, Ducar M, et al. The impact of tumor profiling approaches and genomic data strategies for cancer precision medicine. *Genome Med*. 2016;8:79. doi:10.1186/s13073-016-0333-9.
23. Campesato LF, Barroso-Sousa R, Jimenez L, Correa BR, Sabbaga J, Hoff PM, Reis LFL, Galante PAF, Camargo AA. Comprehensive cancer-gene panels can be used to estimate mutational load and predict clinical benefit to PD-1 blockade in clinical practice. *Oncotarget*. 2015;6:34221–34227. doi:10.18632/oncotarget.5950.
24. Heeke S, Hofman P. Tumor mutational burden assessment as a predictive biomarker for immunotherapy in lung cancer patients: getting ready for prime-time or not? *Transl Lung Cancer Res*. 2018;7:631–638. doi:10.21037/tlcr.2018.08.04.
25. Schumacher TN, Schreiber RD, Chaudhary N, Page A, Smith D, Thompson J, Vinicius L, Mace R, Migliano AB. Neoantigens in cancer immunotherapy. *Science*. 2015;348:69–74. doi:10.1126/science.aaa4971.
26. Anagnostou V, Smith KN, Forde PM, Niknafs N, Bhattacharya R, White J, Zhang T, Adleff V, Phallen J, Wali N, et al. Evolution of neoantigen landscape during immune checkpoint blockade in non-small cell lung cancer. *Cancer Discov*. 2017;7:264–276. doi:10.1158/2159-8290.CD-16-0828.
27. Calin GA, Croce CM. MicroRNA signatures in human cancers. *Nat Rev Cancer*. 2006;6:857–866. doi:10.1038/nrc1997.
28. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144:646–674. doi:10.1016/j.cell.2011.02.013.
29. Berindan-Neagoe I, Monroig Pdel C, Pasculli B, Calin GA. MicroRNAome genome: a treasure for cancer diagnosis and therapy. *CA Cancer J Clin*. 2014;64:311–336. doi:10.3322/caac.21244.
30. Hayes J, Peruzzi PP, Lawler S. MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol Med*. 2014;20:460–469. doi:10.1016/j.molmed.2014.06.005.
31. Dragomir M, Chen B, Fu X, Calin GA. Key questions about the checkpoint blockade-are microRNAs an answer? *Cancer Biol Med*. 2018;15:103–115. doi:10.20892/j.issn.2095-3941.2018.0006.
32. Mehta A, Baltimore D. MicroRNAs as regulatory elements in immune system logic. *Nat Rev Immunol*. 2016;16:279–294. doi:10.1038/nri.2016.40.
33. Xiao PX, Barrios CH, Kim TM, Cosgriff T, Srimuninnimit V, Pittman K, Sabbatini R, Rha SY, Flaig TW, Page RD, et al. 536PCorrelation of plasma exosomal microRNAs with the efficacy of immunotherapy in EGFR/ALK wild type advanced NSCLC. *Ann Oncol*. 2018;29. doi:10.1093/annonc/mdy425.046.
34. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Carolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res*. 2016;44:e71. doi:10.1093/nar/gkv1507.
35. Cibulskis K, Hu Y-C, Markoulaki S, Welstead GG, Cheng AW, Shivalila CS, Pyntikova T, Dadon DB, Voytas DF, Bogdanove AJ, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31:213–219. doi:10.1038/nbt.2514.
36. Ramalingam SS, Hellmann MD, Awad MM, Borghaei H, Gainor J, Brahmer J, Spigel DR, Reck M, O’Byrne KJ, Paz-Ares L, et al. Abstract CT078: tumor mutational burden (TMB) as a biomarker for clinical benefit from dual immune checkpoint blockade with nivolumab (nivo) + ipilimumab (ipi) in first-line (1L) non-small cell lung cancer (NSCLC): identification of TMB cutoff from CheckMate 568. *Cancer Res*. 2018;78:CT078–CT078.
37. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47. doi:10.1093/nar/gkv007.
38. Szekely GJ, Rizzo ML. Hierarchical clustering via joint between-within distances: extending ward’s minimum variance method. *J Classification*. 2005;22:151–183. doi:10.1007/s00357-005-0012-9.
39. Wu TT, Chen YF, Hastie T, Sobel E, Lange K. Genome-wide association analysis by lasso penalized logistic regression. *Bioinformatics*. 2009;25:714–721. doi:10.1093/bioinformatics/btp041.
40. Friedman JH, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:22. doi:10.18637/jss.v033.i01.
41. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77. doi:10.1186/1471-2105-12-77.
42. Hellmann MD, Callahan MK, Awad MM, Calvo E, Ascierto PA, Atmaca A, Rizvi NA, Hirsch FR, Selvaggi G, Szustakowski JD, et al. Tumor mutational burden and efficacy of nivolumab monotherapy and in combination with ipilimumab in small-cell lung cancer. *Cancer Cell*. 2018;33:853–861 e854. doi:10.1016/j.ccell.2018.04.001.
43. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15:R29. doi:10.1186/gb-2014-15-2-r29.
44. Yang JH, Li J-H, Shao P, Zhou H, Chen Y-Q, Qu L-H. starBase: a database for exploring microRNA-mRNA interaction maps from argonaute CLIP-seq and degradome-seq data. *Nucleic Acids Res*. 2011;39:D202–209. doi:10.1093/nar/gkq1056.

45. Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T, Hatzigeorgiou AG. DIANA-miRPath v3.0: deciphering microRNA function with experimental support. *Nucleic Acids Res.* 2015;43:W460–466. doi:10.1093/nar/gkv403.
46. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet.* 2000;25:25–29. doi:10.1038/75556.
47. Vlachos IS, Paraskevopoulou MD, Karagkouni D, Georgakilas G, Vergoulis T, Kanellos I, Anastasopoulos I-L, Maniou S, Karathanou K, Kalfakakou D, et al. DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA: mRNA interactions. *Nucleic Acids Res.* 2015;43:D153–159. doi:10.1093/nar/gku1215.
48. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44:837–845.
49. Koeppel F, Blanchard S, Jovelet C, Genin B, Marcaillou C, Martin E, Rouleau E, Solary E, Soria J-C, André F, et al. Whole exome sequencing for determination of tumor mutation load in liquid biopsy from advanced cancer patients. *PLoS One.* 2017;12:e0188174. doi:10.1371/journal.pone.0188174.
50. Davis AA, Chae YK, Agte S, Pan A, Simon NI, Taxter TJ, Behdad A, Carneiro BA, Cristofanilli M, Giles FJ. Comparison of tumor mutational burden (TMB) across tumor tissue and circulating tumor DNA (ctDNA). *J Clin Oncol.* 2017;35:e23028–e23028. doi:10.1200/JCO.2017.35.15_suppl.e23028.
51. Gandara DR, Paul SM, Kowanetz M, Schleifman E, Zou W, Li Y, Rittmeyer A, Fehrenbacher L, Otto G, Malboeuf C, et al. Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. *Nat Med.* 2018;24:1441–1448. doi:10.1038/s41591-018-0134-3.
52. Fabrizio D, Lieber D, Malboeuf C, Silterra J, White E, Coyne M, Brennan T, Ma J, Kennedy M, Schleifman E, et al. Abstract 5706: A blood-based next-generation sequencing assay to determine tumor mutational burden (bTMB) is associated with benefit to an anti-PD-L1 inhibitor, atezolizumab. *Cancer Res.* 2018;78:5706. doi:10.1158/1538-7445.am2018-5706.
53. Fenizia F, Pasquale R, Roma C, Bergantino F, Iannaccone A, Normanno N. Measuring tumor mutation burden in non-small cell lung cancer: tissue versus liquid biopsy. *Transl Lung Cancer Res.* 2018;7:668–677. doi:10.21037/tlcr.2018.09.23.
54. Cortez MA, Anfossi S, Ramapriyan R, Menon H, Atalar SC, Aliru M, Welsh J, Calin GA. Role of miRNAs in immune responses and immunotherapy in cancer. *Genes Chromosomes Cancer.* 2019;58:244–253. doi:10.1002/gcc.22725.
55. Boeri M, Milione M, Proto C, Signorelli D, Lo Russo G, Galeone C, Verri C, Mensah M, Centonze G, Martinetti A, et al. Circulating miRNAs and PD-L1 tumor expression are associated with survival in advanced NSCLC patients treated with immunotherapy: a prospective study. *Clin Cancer Res.* 2019;25:2166–2173. doi:10.1158/1078-0432.CCR-18-1981.