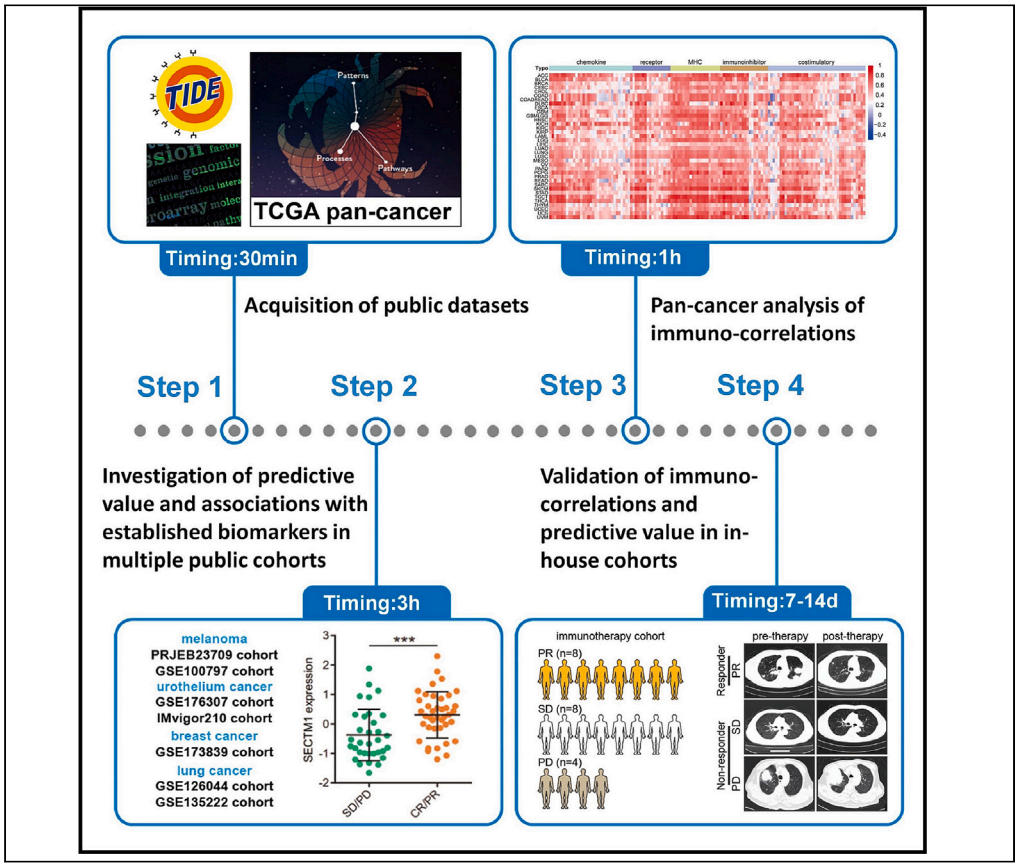
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol to identify novel immunotherapy biomarkers based on transcriptomic data in human cancers



Immune checkpoint inhibitors have transformed the management of advanced cancers, but biomarkers for the prediction of therapeutic responses have not been fully uncovered. Here, we provide a step-by-step approach for the identification of novel biomarkers from public transcriptomic datasets. We comprehensively summarize the available transcriptomic datasets containing immunotherapy information and describe the necessary procedures to evaluate the effectiveness of a novel immunotherapy biomarker, which may accelerate the identification of novel immunotherapy biomarkers.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Jie Mei, Yun Cai, Rui Xu, ..., Wenjun Mao, Junying Xu, Yongmei Yin

maowenjun1@njmu.edu.cn (W.M.)
doctorxjy123@163.com (J.X.)
ymyin@njmu.edu.cn (Y.Y.)

Highlights

Detailed steps to identify novel immunotherapy biomarkers in human cancers

Step-by-step approach to select candidate biomarkers, public and in-house cohorts

Necessary parameters to evaluate the feasibility of immunotherapy biomarkers

Mei et al., STAR Protocols 4, 102258
June 16, 2023
https://doi.org/10.1016/j.xpro.2023.102258

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol to identify novel immunotherapy biomarkers based on transcriptomic data in human cancers

Jie Mei,[1,2,3,9] Yun Cai,[3,9] Rui Xu,[2,9] Yichao Zhu,[4] Xinyuan Zhao,[5] Yan Zhang,[6] Wenjun Mao,[7,*] Junying Xu,[1,10,*] and Yongmei Yin[2,8,11,*]

[1]Department of Oncology, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi 214023, China

[2]Department of Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China

[3]Wuxi Clinical College of Nanjing Medical University, Wuxi 214023, China

[4]Department of Physiology, Nanjing Medical University, Nanjing, Jiangsu 211166, China

[5]Department of Occupational Medicine and Environmental Toxicology, Nantong Key Laboratory of Environmental Toxicology, School of Public Health, Nantong University, Nantong 226019, China

[6]Wuxi Maternal and Child Health Care Hospital, Wuxi Medical Center of Nanjing Medical University, Wuxi 214023, China

[7]Department of Thoracic Surgery, The Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi 214023, China

[8]Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Personalized Cancer Medicine, Nanjing Medical University, Nanjing 211166, China

[9]These authors contributed equally

[10]Technical contact

[11]Lead contact

*Correspondence: maowenjun1@njmu.edu.cn (W.M.), doctorxjy123@163.com (J.X.), ymyin@njmu.edu.cn (Y.Y.)
https://doi.org/10.1016/j.xpro.2023.102258

SUMMARY

Immune checkpoint inhibitors have transformed the management of advanced cancers, but biomarkers for the prediction of therapeutic responses have not been fully uncovered. Here, we provide a step-by-step approach for the identification of novel biomarkers from public transcriptomic datasets. We comprehensively summarize the available transcriptomic datasets containing immunotherapy information and describe the necessary procedures to evaluate the effectiveness of a novel immunotherapy biomarker, which may accelerate the identification of novel immunotherapy biomarkers.

For complete details on the use and execution of this protocol, please refer to Mei et al.[1]

BEFORE YOU BEGIN

Overview

In the past decades, immunotherapy, a revolutionary strategy, has largely transformed the therapeutic situation of human cancers with advanced clinical stages.[2] Although the prognosis of cancer patients with advanced stages has been persistently improved with the application of immunotherapy, not all patients could benefit from the established treatment options.[3,4] It has been well-known that PD-L1 expression is a dominating factor that determines whether a patient responds to anti-PD-1/PD-L1 immunotherapy, but a large group of patients with PD-L1-negative expression could also benefit from immunotherapy.[5,6] Thus, complementary and alternative biomarkers are urgent in clinical practice for the prediction of anti-PD-1/PD-L1 immunotherapeutic responses.

Increasing numbers of scholars are devoted to identifying more effective immunotherapy biomarkers. Tissue biopsy has always been the gold standard for clinical diagnosis and evaluation. With the development of high-throughput technologies, the process of biomarker identification is greatly accelerated. Biomarkers based on genome-wide screening exhibit predominant predictive

values, such as TIDE score[7] and T cell inflamed score,[8] but this is not the mainstream of clinical applications. In addition, the improved detection means of established biomarkers are also widely proposed, such as the deglycosylation detection previously proposed by our research group.[9] However, the complexity of the operations may limit clinical applications. In general, considering the economic burden on patients and the convenience of clinical application, single gene biomarkers have more translational value. Thus, the current protocol aims to provide a standardized procedure that could be used to assess whether a candidate gene could be used as a novel immunotherapy biomarker.

Selection of candidate immunotherapy biomarkers

Timing: 2 h

When users want to follow this process, they need to first select a candidate gene. We offer three screening methods to select a candidate gene.

1. For the identification of pan-cancer biomarkers, we recommend users access the intersection of differentially expressed genes (DEGs) between responders and non-responders in main cancer types suitable for immunotherapy, such as melanoma, non-small cell lung cancer (NSCLC), urothelium cancer, and breast cancer.
2. To identify biomarkers in single cancer type instead of pan-cancer, such as melanoma or NSCLC, accessing the intersection of DEGs between responders and non-responders in different datasets within the same cancer type could narrow the selection range and increase the reliability of candidates.
3. For cancer types without sufficient public datasets including immunotherapy information, the ESTIMATE algorithm is performed initially to assess the relative abundance of tumor-infiltrating immune cells (TIICs).[10] Subsequently, Pearson's coefficient is performed to measure the correlations between transcriptional levels of candidate genes and TIICs. To narrow down the selection range, we set the threshold (for example, Pearson's correlation $\geq$ 0.5 and p-value < 0.05). After the screening, the candidate gene is found to be positively correlated with TIICs. For further validation, Pearson's coefficient is performed to evaluate the correlations between the candidate gene and the relative abundance of TIIC subpopulations assessed by several independent algorithms, such as TIMER,[11] EPIC,[12] MCP-counter,[13] and TISIDB.[14] If the candidate gene is positively correlated with most TIIC subpopulations, the candidate gene will be included in subsequent studies. Totally, evaluating the correlations between candidate genes and TIICs may be helpful to narrow down the selection range, but it is more applicable to a range of candidates, such as the selection of a research object from a gene family.

Note: The criterion for DEGs selection is typically $p < 0.05$, and there is no strict requirement for fold change (FC) value.

Note: We recommend selecting genes expressed in tumor cells rather than immune cells or other cells. Due to the high purity of some tumors, the expression of genes expressed in non-tumor cells may not be evaluated. Thus, it is crucial to analyze expression patterns of candidate genes in various cell subpopulations using single-cell RNA-sequencing data. In addition, the Human Protein Atlas [15] platform provides online visualization of protein-coding genes in tumor tissues, which could also be used as a tool to screen expression patterns of candidate genes. If you want to analyze the single-cell RNA sequencing (scRNA-seq) datasets yourself, the R package Seurat can be used. The R package Seurat is designed for quality control (QC), further analysis, and exploration of scRNA-seq datasets. Notably, the detailed and complete tutorials are available on the website https://satijalab.org/seurat/.

Alternatives: Several online intersection tools could be used for selecting potential biomarkers, such as Venny 2.1.0 (https://bioinfogp.cnb.csic.es/tools/venny/index.html). Other similar tools could also be used as alternatives.

Alternatives: Several online intersection tools could be used for scRNA-seq analysis, such as TISCH (http://tisch.comp-genomics.org/home/).[16]

Selection of transcriptomic datasets

Timing: 2 h

More and more studies have published public transcriptomic datasets containing immunotherapy information. We have collected and summarized datasets with sufficient cases and complete information, most of which could be downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) or the Tumor Immune Dysfunction and Exclusion (TIDE, http://tide.dfci.harvard.edu/) databases. The expression and clinical data of the IMvigor210 cohort could be obtained from the website (http://research-pub.gene.com/IMvigor210CoreBiologies/). In general, the more datasets with expression differences of candidate genes between responders and non-responders, the more stable predictive value of candidate genes.

Note: The full list is as follows: melanoma: PRJEB23709, GSE100797, GSE91061, GSE78220, GSE93157; NSCLC: GSE126044, GSE135222, GSE136961, GSE93157; breast cancer: GSE173839, GSE194040; urothelial cancer: GSE176307, IMvigor210; gastric cancer: PRJEB25780; hepatocellular carcinoma: GSE140901; esophagus cancer: GSE165252.

Preparation of necessary in-house clinical cohorts

Timing: unpredictable

In general, in-house clinical cohorts are needed to validate the predictive performance of candidates. The collection of paraffin-embedded tumor samples is helpful for validation. As far as possible, samples unaffected by unrelated treatments need to be selected for all in-house cohorts, and at least one in-house cohort should contain immunotherapy information, such as Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 and survival data after immunotherapy.

Note: If users decide to collect in-house cohorts containing immunotherapy information, several points may be helpful. Samples should be obtained before immunotherapy, and other treatments should not be received before receiving immunotherapy. Response evaluation using the RECIST 1.1 criterion is usually necessary, and it is also recommended to collect the follow-up information. In addition, more immunotherapy biomarkers could also be collected to evaluate the correlation of candidate biomarkers with these established biomarkers and their predictive values.

⚠ CRITICAL: Although in-house cohorts seem not to be necessary in some published articles, validated results from in-house cohorts could greatly increase the clinical availability of predicted candidates.

Alternatives: Collecting in-house cohorts is challenging, but some commercialized clinical sample libraries could be complementary, such as Outdo BioTech (https://www.superchip.com.cn/biology/tissue.html) and Liaoding BioTech (http://www.shliaoding.com/). However, there is still no commercial cohort containing immunotherapy information. Thus, only the correlations between predicted candidates and tumor immune microenvironment features (such as PD-L1 expression and TIICs levels) could be validated.

Institutional permissions

If in-house cohorts are included, users need to acquire permissions from the Institutional Review Board and acquire signed consent to allow the use of biopsies for research practice in compliance with local regulation.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
CD8 antibody (ready-to-use)	Abcarta	PA067
PD-L1 antibody (ready-to-use)	GeneTech	GT2280
MLH1 antibody (ready-to-use)	GeneTech	GT2304
MSH2 antibody (ready-to-use)	GeneTech	GT2310
MSH6 antibody (ready-to-use)	GeneTech	GT2195
PSM2 antibody (ready-to-use)	GeneTech	GT2149
Biological samples		
lung cancer TMA	Outdo	HLugA060PG02
breast cancer TMA	Outdo	HBreD090PG01
lung cancer samples	Mei et al.[1]	NA
gastric cancer samples	Mei et al.[1]	NA
Deposited data		
TIDE: PRJEB23709 (melanoma)	TIDE	http://tide.dfci.harvard.edu/
TIDE: GSE100797 (melanoma)	TIDE	http://tide.dfci.harvard.edu/
TIDE: GSE91061 (melanoma)	TIDE	http://tide.dfci.harvard.edu/
TIDE: GSE78220 (melanoma)	TIDE	http://tide.dfci.harvard.edu/
GEO: GSE93157 (multiple cancer types)	GEO	https://www.ncbi.nlm.nih.gov/geo/
GEO: GSE126044 (NSCLC)	GEO	https://www.ncbi.nlm.nih.gov/geo/
GEO: GSE135222 (NSCLC)	GEO	https://www.ncbi.nlm.nih.gov/geo/
GEO: GSE136961 (NSCLC)	GEO	https://www.ncbi.nlm.nih.gov/geo/
GEO: GSE173839 (breast cancer)	GEO	https://www.ncbi.nlm.nih.gov/geo/
GEO: GSE194040 (breast cancer)	GEO	https://www.ncbi.nlm.nih.gov/geo/
GEO: GSE176307 (urothelial cancer)	GEO	https://www.ncbi.nlm.nih.gov/geo/
Database: IMvigor210 (urothelial cancer)	IMvigor210	http://research-pub.gene.com/IMvigor210CoreBiologies/
GEO: GSE140901 (hepatocellular carcinoma)	GEO	https://www.ncbi.nlm.nih.gov/geo/
TIDE: PRJEB25780 (gastric cancer)	TIDE	http://tide.dfci.harvard.edu/
GEO: GSE165252 (esophagus cancer)	GEO	https://www.ncbi.nlm.nih.gov/geo/
TCGA: TCGA RNA-seq data	TCGA	https://xenabrowser.net/datapages/
TCGA: TCGA mutation data	TCGA	https://portal.gdc.cancer.gov/
Software and algorithms		
SPSS 26	IBM SPSS	https://www.ibm.com/docs/zh/spss-statistics/
Graphpad Prism 6.0	GraphPad	https://www.graphpad.com/
Sangerbox	Sangerbox	https://vip.sangerbox.com/login.html
R 4.0.2	R project	https://cran.r-project.org
pheatmap 1.0.12	R package	https://mirrors.tuna.tsinghua.edu.cn/CRAN/
maftools 2.6.5	R package	https://mirrors.tuna.tsinghua.edu.cn/CRAN/
RColorBrewer 1.1.2	R package	https://mirrors.tuna.tsinghua.edu.cn/CRAN/
tidyverse 1.3.1	R package	https://mirrors.tuna.tsinghua.edu.cn/CRAN/
limma 3.46.0	R package	http://www.bioconductor.org/
corrplot 0.90	R package	https://mirrors.tuna.tsinghua.edu.cn/CRAN/

Alternatives: Any validated, commercial antibodies are feasible, but ready-to-use antibodies are preferentially recommended.

STEP-BY-STEP METHOD DETAILS

Manual installation of required package components

Timing: 10 min

Install all R packages required for the transcriptomic data analysis.

```
### R packages used in this step

options(stringsAsFactors = F)

install.packages("pheatmap", repos = "https://mirrors.tuna.tsinghua.edu.cn/CRAN/")

install.packages("maftools", repos = "https://mirrors.tuna.tsinghua.edu.cn/CRAN/")

install.packages("RColorBrewer", repos = "https://mirrors.tuna.tsinghua.edu.cn/CRAN/")

install.packages("corrplot ", repos = "https://mirrors.tuna.tsinghua.edu.cn/CRAN/")

install.packages("tidyverse", repos = "https://mirrors.tuna.tsinghua.edu.cn/CRAN/")

if (!require("BiocManager", quietly = TRUE)) install.packages("BiocManager")

BiocManager::install("limma")
```

Acquisition of public datasets

Timing: 30 min

In this step, relative datasets should be downloaded from data storage platforms.

1. Acquisition of public immunotherapy datasets.

   Note: Most datasets comprising RNA-sequencing data from patients receiving immunotherapy could be downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) or the Tumor Immune Dysfunction and Exclusion (TIDE, http://tide.dfci.harvard.edu/) databases. The expression and clinical data of the IMvigor210 cohort could be obtained from the website (http://research-pub.gene.com/IMvigor210CoreBiologies/). The corresponding datasets identifier could be obtained in the key resources table.

2. Acquisition of the TCGA dataset.

   Note: The standardized TCGA pan-cancer dataset (TCGA Pan-Cancer (PANCAN)) could be downloaded from the UCSC (https://xenabrowser.net/) database. The somatic mutation data are obtained from the TCGA (http://cancergenome.nih.gov/) database and then used to calculate the tumor mutation burden (TMB) by R package "maftools".

```
### R packages used in this step

library (maftools)

laml <- read.maf(maf = "TCGA.mutect.maf.gz")

x = tmb(maf = laml)
```

Investigation of predictive value in a discovery cohort

Timing: 1 h

In this step, we first choose one cohort as the discovery cohort containing comprehensive clinical information, such as the PRJEB23709 cohort. The predictive value, and clinical and immune correlation of the candidate gene is defined using the discovery cohort. Here, we show *SECTM1* as an example (Figure 1).

3. Obtain candidate gene expression and evaluate its predictive value and clinical correlation.

   Note: Usually, the discovery cohort should contain as much clinical data as possible. If users aim to select a biomarker in pan-cancer or melanoma, we recommend the PRJEB23709 cohort as the discovery cohort. Concretely speaking, the correlations between candidate biomarkers and immunotherapeutic responses as well as survival time could be assessed, and multiple statistical methods could be used, such as t test, chi-square test, log-rank test, and Cox regression analysis.

   Note: This step is usually done without any programming software, just using widely used softwares such as Office-Excel, GraphPad Prism, and SPSS.

   Note: If the gene name is not stored as the gene symbol, it first needs to be converted to the gene symbol.

   ⚠ CRITICAL: Comparing the predictive value of the candidate biomarker with those of classical markers such as PD-L1 helps to illustrate the importance of the candidate biomarker.

4. Assess the correlations between candidate gene and tumor immune microenvironment features.

```
### R packages used in this step

# Gene expression profile with genes as the row names and samples as the column names

expr = readRDS('expr.rds')

# Tumor immune microenvironment (TME) features, taking chemokine genes for example

features = readRDS('chemokine.rds') # gene vector

# calculation the correlation between candiate gene SECTM1 and chemokine genes

correlation = data.frame(do.call(rbind,lapply(intersect(features,rownames(expr)),
function(x) {

tmp = cor.test(as.numeric(expr['SECTM1',]),as.numeric(expr[x,]),method='pearson')

return(c(x,tmp$estimate,tmp$p.value)) })))

colnames(correlation) = c('feature','correlation','pvalue')

# plot

anno_col = data.frame(SECTM1=as.numeric(expr['SECTM1',]),Sample=colnames(expr)) %>%
arrange(SECTM1) %>% column_to_rownames('Sample')

mat = expr[intersect(features,rownames(expr)),intersect(rownames(anno_col),colnames(expr))]

pheatmap(mat, show_rownames=T, cluster_cols=F, cluster_rows=F, annotation_col=anno_col,
scale='row')
```

   Note: The features of the tumor immune microenvironment include immunomodulators, the activities of the cancer immunity cycle, infiltration levels of TIICs, and the expression of inhibitory immune checkpoints, tumor purity, the detail information could be found in our previous studies.[17,18]

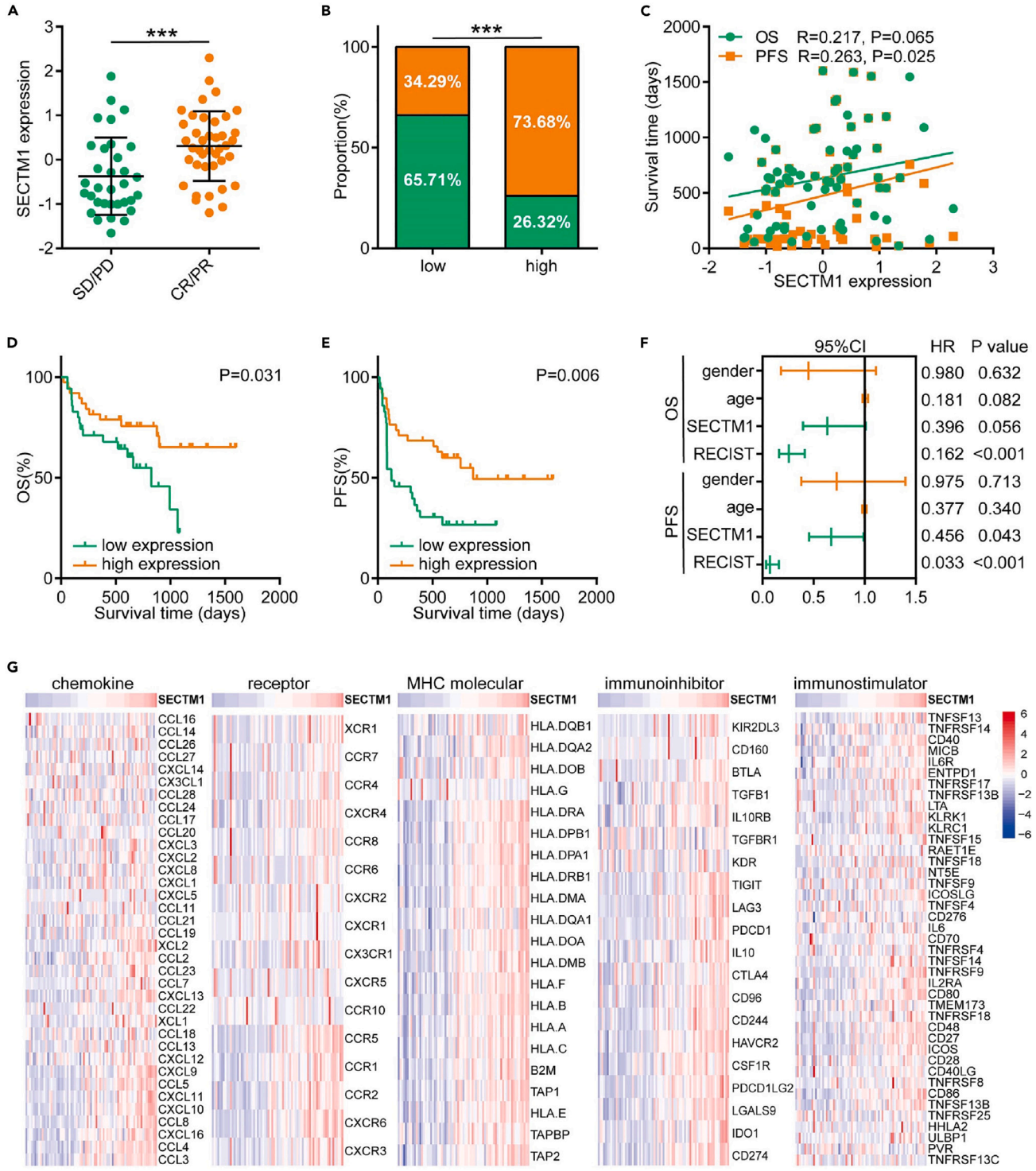


Figure 1. Predictive value and immunological correlations of *SECTM1* in the PRJEB23709 cohort

Reproduced with permission from iScience (Mei et al.[1]).

(A) *SECTM1* expression levels in tumors from patients with different responses. Data presented as mean ± SD. Significance was calculated with Student's t test. ***p < 0.001.

(B) ORR in patients with low and high *SECTM1* expression. Significance was calculated with Pearson's $\chi 2$ test. ***p < 0.001.

(C) Correlations between *SECTM1* expression and OS and PFS time. Significance was calculated with Pearson correlation test.

(D and E) Prognostic values of *SECTM1* in terms of OS and PFS. Median *SECTM1* expression was used as the cut-off value. Significance was calculated with log-rank test.

Figure 1. *Continued*
(F) Cox regression analysis of prognosis-related factors in melanoma patients.
(G) Heatmap showing correlations between *SECTM1* and immunomodulators expression, including chemokines, receptors, MHCs, immunoinhibitors, and immunostimulators. Significance was calculated with Pearson correlation test.

Generalization of results in more public cohorts

Timing: 1 h

In this section, the predictive value of the candidate gene should be explored in as many clinical as possible. In general, the more datasets with expression differences of candidate genes between responders and non-responders, the more stable predictive value of candidate genes. Here, we show SECTM1 as an example, *SECTM1* is dys-regulated in tumors from responders and non-responders in at least six other clinical cohorts (Figure 2).

5. Obtain candidate gene expression and evaluate its predictive values in more cohorts.

To make it easier to load the gene list, the .rds file is used. To generate the .rds file, the function "saveRDS" is performed. The command is as follows: saveRDS(genelist, file = "genelist.rds").

Note: The .rds file is a document format like .txt and .xlsx.

Note: The "genelist" is a variable storing a list of genes (Table 1).

Note: This step is usually done without any programming software, just using widely used softwares such as Office-Excel, GraphPad Prism, and SPSS.

Note: If the gene name is not stored as the gene symbol, it first needs to be converted to the gene symbol.

⚠ CRITICAL: Comparing the predictive value of the candidate biomarker with those of classical biomarkers such as PD-L1 helps to illustrate the importance of the candidate biomarker.

6. Assess the correlations between candidate gene and tumor immune microenvironment features.

```
### R packages used in this step
# path for gene expression profiles
exprPath = c('GSE100797.rds','GSE176307.rds','IMvigor210.rds','GSE173139.rds','GSE126244.rds','GSE135222.rds')
# tumor immune microenvironment (TME) features, taking chemokine genes for example
features = readRDS('chemokine.rds') # gene vector
# calculation the correlation between candiate gene SECTM1 and chemokine genes in each dataset
correlation = data.frame(do.call(cbind,lapply(exprPath,function(path){
expr = readRDS(path)
rs = sapply(intersect(features,rownames(expr)),function(x){
return(cor.test(as.numeric(expr['SECTM1',]),as.numeric(expr[x,]),method='pearson')$estimate) })
return(data.frame(correlation=rs, row.names=intersect(features,rownames(expr)))) })))
colnames(correlation) = gsub('\\..*','',exprPath)
pheatmap(correlation, show_rownames=T, show_colnames=T, cluster_cols=F, cluster_rows=F)
```

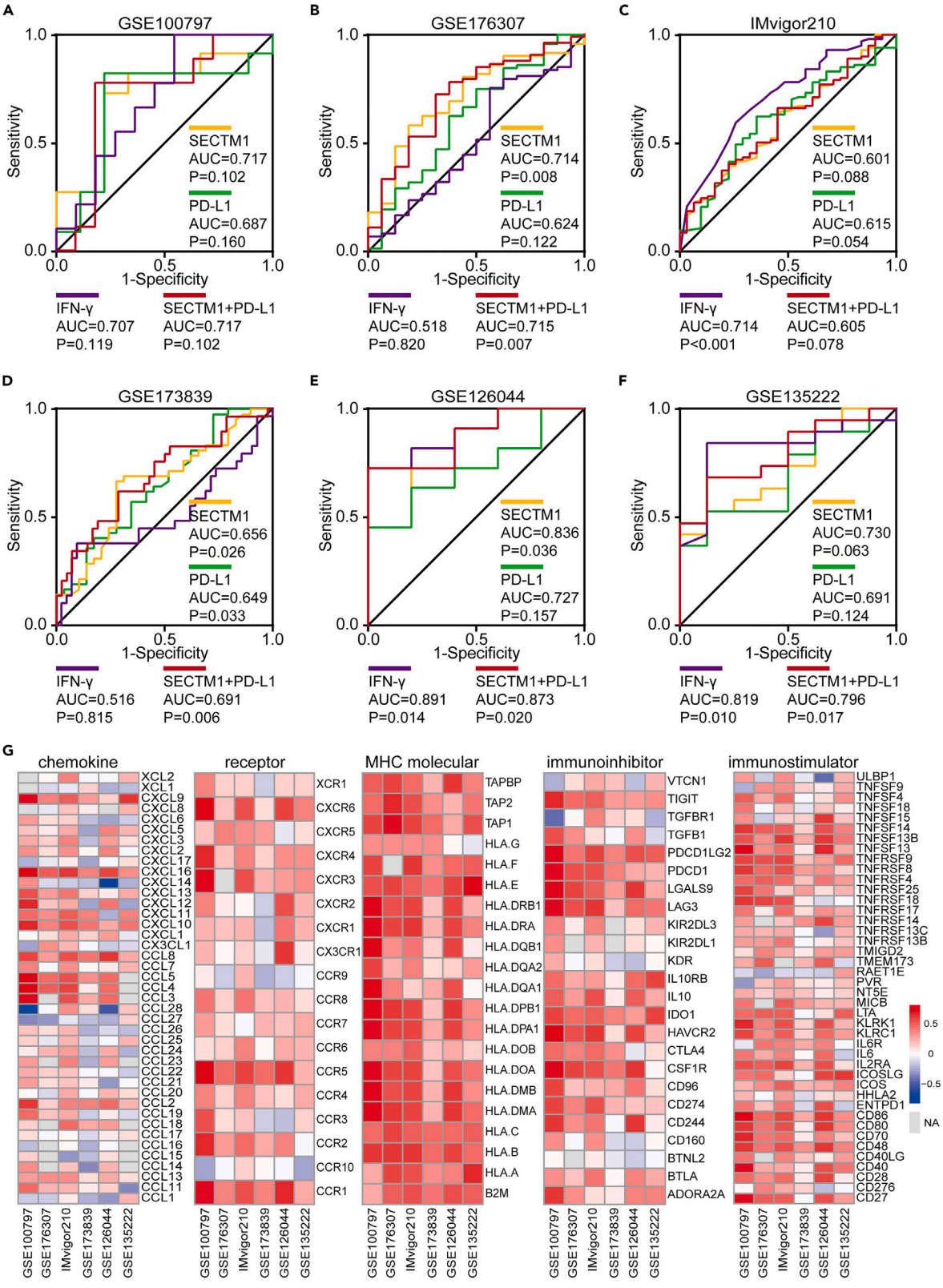

A
GSE100797
SECTM1
AUC=0.717
P=0.102
PD-L1
AUC=0.687
P=0.160
IFN-γ
AUC=0.707
P=0.119
SECTM1+PD-L1
AUC=0.717
P=0.102
B
GSE176307
SECTM1
AUC=0.714
P=0.008
PD-L1
AUC=0.624
P=0.122
IFN-γ
AUC=0.518
P=0.820
SECTM1+PD-L1
AUC=0.715
P=0.007
C
IMvigor210
SECTM1
AUC=0.601
P=0.088
PD-L1
AUC=0.615
P=0.054
IFN-γ
AUC=0.714
P<0.001
SECTM1+PD-L1
AUC=0.605
P=0.078
D
GSE173839
SECTM1
AUC=0.656
P=0.026
PD-L1
AUC=0.649
P=0.033
IFN-γ
AUC=0.516
P=0.815
SECTM1+PD-L1
AUC=0.691
P=0.006
E
GSE126044
SECTM1
AUC=0.836
P=0.036
PD-L1
AUC=0.727
P=0.157
IFN-γ
AUC=0.891
P=0.014
SECTM1+PD-L1
AUC=0.873
P=0.020
F
GSE135222
SECTM1
AUC=0.730
P=0.063
PD-L1
AUC=0.691
P=0.124
IFN-γ
AUC=0.819
P=0.010
SECTM1+PD-L1
AUC=0.796
P=0.017
Sensitivity
1-Specificity
0.0
0.5
1.0
G
chemokine
receptor
MHC molecular
immunoinhibitor
immunostimulator
XCL2
XCL1
CXCL9
CXCL8
CXCL6
CXCL5
CXCL3
CXCL2
CXCL17
CXCL16
CXCL14
CXCL13
CXCL12
CXCL11
CXCL10
CXCL1
CX3CL1
CCL8
CCL7
CCL5
CCL4
CCL3
CCL28
CCL27
CCL26
CCL25
CCL24
CCL23
CCL22
CCL21
CCL20
CCL2
CCL19
CCL18
CCL17
CCL16
CCL15
CCL14
CCL13
CCL11
CCL1
XCR1
CXCR6
CXCR5
CXCR4
CXCR3
CXCR2
CXCR1
CX3CR1
CCR9
CCR8
CCR7
CCR6
CCR5
CCR4
CCR3
CCR2
CCR10
CCR1
TAPBP
TAP2
TAP1
HLA.G
HLA.F
HLA.E
HLA.DRB1
HLA.DRA
HLA.DQB1
HLA.DQA2
HLA.DQA1
HLA.DPB1
HLA.DPA1
HLA.DOB
HLA.DOA
HLA.DMB
HLA.DMA
HLA.C
HLA.B
HLA.A
B2M
VTCN1
TIGIT
TGFBR1
TGFB1
PDCD1LG2
PDCD1
LGALS9
LAG3
KIR2DL3
KIR2DL1
KDR
IL10RB
IL10
IDO1
HAVCR2
CTLA4
CSF1R
CD96
CD274
CD244
CD160
BTNL2
BTLA
ADORA2A
ULBP1
TNFSF9
TNFSF4
TNFSF18
TNFSF15
TNFSF14
TNFSF13B
TNFSF13
TNFRSF9
TNFRSF8
TNFRSF4
TNFRSF25
TNFRSF18
TNFRSF17
TNFRSF14
TNFRSF13C
TNFRSF13B
TMIGD2
TMEM173
RAET1E
PVR
NT5E
MICB
LTA
KLRK1
KLRC1
IL6R
IL6
IL2RA
ICOSLG
ICOS
HHLA2
ENTPD1
CD86
CD80
CD70
CD48
CD40LG
CD40
CD28
CD276
CD27
0.5
0
−0.5
NA
GSE100797
GSE176307
IMvigor210
GSE173839
GSE126044
GSE135222

Figure 2. Predictive value and immunological correlations of *SECTM1* in six cohorts

Reproduced with permission from iScience (Mei et al.[1]).

(A–F) Comparison of predictive values of SECTM1, PD-L1, IFN-γ and the SECTM1/PD-L1 combination for immunotherapy responses in six cohorts. The predictive value of the combination of SECTM1 and PD-L1 was estimated by binary logistic regression using SPSS 26. Receiver-operating characteristic (ROC) analysis was plotted to assess the specificity and sensitivity of the candidate indicator, and the area under the ROC curve (AUC) was generated for diagnostic biomarkers.

(G) Heatmap showing correlations between *SECTM1* and immunomodulators expression, including chemokines, receptors, MHCs, immunoinhibitors, and immunostimulators. Significance was calculated with Pearson correlation test.

Note: The features of the tumor immune microenvironment include immunomodulators, the activities of the cancer immunity cycle, infiltration levels of TIICs, and the expression of inhibitory immune checkpoints, tumor purity, the detail information could be found in our previous studies.[17,18]

Investigation of associations with established immunotherapy biomarkers

Timing: 1 h

In this section, the correlations between candidate gene and established immunotherapy biomarkers should be evaluated using available data in above cohort. Here, we show *SECTM1* as an example (Figure 3).

7. Acquisition of established immunotherapy biomarkers.

Note: Well-established immunotherapy biomarkers include PD-L1 expression, tumor mutation burden (TMB) level, immune infiltration, and microsatellite instability (MSI) status.[19]

8. Evaluate the correlations with established immunotherapy biomarkers.

Note: This step is usually done without any programming software, just using widely used softwares such as Office-Excel, GraphPad Prism, and SPSS.

Pan-cancer analysis of immuno-correlations

Timing: 1 h

Pan-cancer analysis of immuno-correlations of candidate gene is often necessary in this protocol. Since immunotherapy is not strictly tumor specific, pan-cancer analysis can help identify more tumor species for which candidate biomarker are potentially applicable. Here, we show *SECTM1* as an example (Figure 4).

9. Correlations between candidate and tumor immune microenvironment features in pan-cancer.

```
### R packages used in this step

# gene expression profile with genes as the row names and samples as the column names

expr = readRDS('expr_pancancer.rds')

# data holding the tumor type of samples

type = readRDS('type_pancancer.rds')

# tumor immune microenvironment (TME) features

featurePath = c('chemokine.rds','receptor.rds','MHC.rds','immunoinhibitor.rds','immunostimulator.rds')
```

```
features = data.frame(do.call(rbind,lapply(featurePath,function(x){return(data.frame(feature=readRDS(x),type=gsub('\\..*','',x)))})))

# calculation the correlation between candidate gene SECTM1 and chemokine genes in each cancer type

correlation = data.frame(do.call(cbind,lapply(sort(as.character(unique(type$tumor_type))),function(tumor_type){

expr = expr[,intersect(rownames(type)[which(type$tumor_type==tumor_type)],colnames(expr))]

return(data.frame(correlation=sapply(intersect(features$feature,rownames(expr)),function(x){return(cor.test(as.numeric(expr['SECTM1',]),as.numeric(expr[x,]),method='pearson')$estimate) }), row.names=intersect(features$feature,rownames(expr)))) }))

colnames(correlation) = sort(as.character(unique(type$tumor_type)))

anno_col = features %>% column_to_rownames('feature')

pheatmap(correlation, show_rownames=T, show_colnames=T, cluster_cols=F, cluster_rows=F,annotation_col=anno_col)
```

Note: The features of the tumor immune microenvironment include immunomodulators, the activities of the cancer immunity cycle, infiltration levels of TIICs, and the expression of inhibitory immune checkpoints, tumor purity, the detail information could be found in our previous studies.[17,18]

Alternatives: Sangerbox,[20] a user-interactive online tool, could provide R code-free pan-cancer analysis.

10. Correlations between candidate and MSI gene expression.

```
### R packages used in this step

# gene expression profile with genes as the row names and samples as the column names

expr = readRDS('expr.rds')

# MSI genes

msi = readRDS('MSI_gene.rds')

# correlation heatmap

cor = cor(as.matrix(expr[intersect(c('SECTM1',msi),rownames(expr)),]))

p.mat = cor.mtest(as.matrix(mat))

corrplot(cor,col = rev(colorRampPalette(brewer.pal(10,"RdBu"))(20)),p.mat = p.mat, sig.level = 0.05, insig = "blank")
```

Note: If the candidate gene could be used as a pan-cancer biomarker, the association between the candidate gene and MSI status should be validated in gastrointestinal tumors.

Validation of immuno-correlations and predictive value in in-house cohorts

Timing: 1–2 weeks

⚠ CRITICAL: Although the validation of the candidate at protein level is not necessary, it will greatly increase the reliability of the study. There is a certain degree of inconsistency between mRNA and protein.

Table 1. Gene list of chemokines, receptors, MHCs, immunoinhibitors, and immunostimulators

Chemokine	Receptor	MHC molecule	Immunoinhibitor	Immunostimulator
CCL4	XCR1	TAPBP	IDO1	TNFRSF13C
CXCL16	CXCR3	HLA-E	LGALS9	PVR
CCL8	CXCR6	TAP1	PDCD1LG2	ULBP1
CXCL10	CCR2	B2M	CSF1R	HHLA2
CXCL11	CCR1	HLA-C	HAVCR2	TNFRSF25
CCL5	CCR5	HLA-A	CD244	TNFSF13B
CXCL9	CCR10	HLA-B	CD96	CD86
CXCL12	CXCR5	HLA-F	CTLA4	TNFRSF8
CCL13	CX3CR1	HLA-DMB	IL10	CD40LG
CCL18	CXCR1	HLA-DOA	PDCD1	CD28
XCL1	CXCR2	HLA-DQA1	LAG3	ICOS
CCL22	CCR6	HLA-DMA	TIGIT	CD27
CXCL13	CCR8	HLA-DRB1	KDR	CD48
CCL7	CXCR4	HLA-DPA1	TGFBR1	TNFRSF18
CCL23	CCR4	HLA-DPB1	IL10RB	TMEM173
CCL2	CCR7	HLA-DRA	TGFB1	CD80
XCL2	/	HLA-G	BTLA	IL2RA
CCL19	/	HLA-DOB	CD160	TNFRSF9
CCL21	/	HLA-DQA2	KIR2DL3	TNFSF14
CCL11	/	HLA-DQB1	CD274	TNFRSF4
CXCL5	/	TAP2	/	CD70
CXCL1	/	/	/	IL6
CXCL8	/	/	/	CD276
CXCL2	/	/	/	TNFSF4
CXCL3	/	/	/	ICOSLG
CCL20	/	/	/	TNFSF9
CCL17	/	/	/	NT5E
CCL24	/	/	/	TNFSF18
CCL28	/	/	/	RAET1E
CX3CL1	/	/	/	TNFSF15
CXCL14	/	/	/	KLRC1
CCL27	/	/	/	KLRK1
CCL26	/	/	/	LTA
CCL14	/	/	/	TNFRSF13B
CCL16	/	/	/	TNFRSF17
CCL3	/	/	/	ENTPD1
/	/	/	/	IL6R
/	/	/	/	MICB
/	/	/	/	CD40
/	/	/	/	TNFRSF14
/	/	/	/	TNFSF13

In this section, the immuno-correlations and predictive value of candidate gene are validated using in-house cohorts. In general, this section usually consists of two steps to verify the immuno-correlations and predictive value of the candidate gene. As far as possible, samples unaffected by unrelated treatments need to be selected for all in-house cohorts. Here, we show *SECTM1* as an example (Figures 5 and 6).

11. Validate immuno-correlations in in-house cohorts.

Note: The associations between candidate gene and PD-L1 expression as well as $CD8^+$ T cell distribution generally need to be verified. Tumors are discriminated into 3 phenotypes following the spatial distribution of $CD8^+$ T cells, including the inflamed, the excluded, and

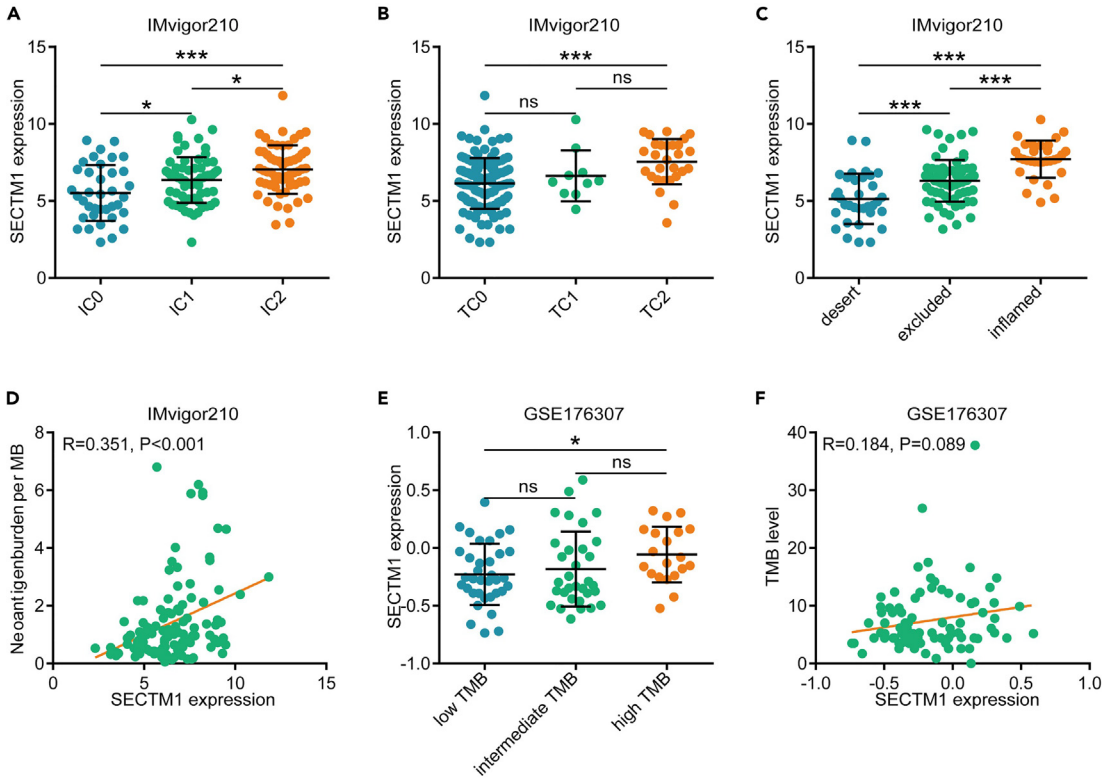


Figure 3. Associations between *SECTM1* expression and established immunotherapy biomarkers

Reproduced with permission from iScience (Mei et al.[1]).

(A) Expression of *SECTM1* in tumors with various PD-L1 IC score. Data are presented as mean ± SD. Significance was calculated with 1-way ANOVA with Tukey's multiple-comparison test. *p < 0.05; ***p < 0.001.

(B) Expression of *SECTM1* in tumors with various PD-L1 TC score. Data are presented as mean ± SD. Significance was calculated with 1-way ANOVA with Tukey's multiple-comparison test. ***p < 0.001.

(C) Expression of *SECTM1* in tumors with various immuno-subtypes. Data are presented as mean ± SD. Significance was calculated with 1-way ANOVA with Tukey's multiple-comparison test. ***p < 0.001.

(D) Correlation between *SECTM1* expression and neoantigen burden. Significance was calculated with Pearson correlation test.

(E) Expression of *SECTM1* in tumors with various TMB levels. Data are presented as mean ± SD. Significance was calculated with 1-way ANOVA with Tukey's multiple-comparison test. *p < 0.05.

(F) Correlation between *SECTM1* expression and TMB levels. Significance was calculated with Pearson correlation test.

the deserted subtypes. The inflamed subtype is considered to be immuno-hot, and both excluded and deserted subtypes are considered to be immuno-cold.[21]

Note: If the candidate gene could be used as a pan-cancer biomarker, the association between the candidate gene and MSI status should be validated in gastrointestinal tumors.

Note: This step is usually done without any programming software, just using widely used softwares such as Office-Excel, GraphPad Prism, and SPSS.

Alternatives: Immunohistochemical (IHC) staining is the most commonly used method,[1] immunofluorescence, flow cytometry, ELISA, Western blotting, etc., can also be used.

⚠ CRITICAL: As flow cytometry, ELISA, Western blotting could not recognize the spatial localization of CD8 expression, there is no further discrimination of immune subtypes. Thus, IHC and immunofluorescence are preferentially recommended.

12. Validate predictive value in in-house cohorts.

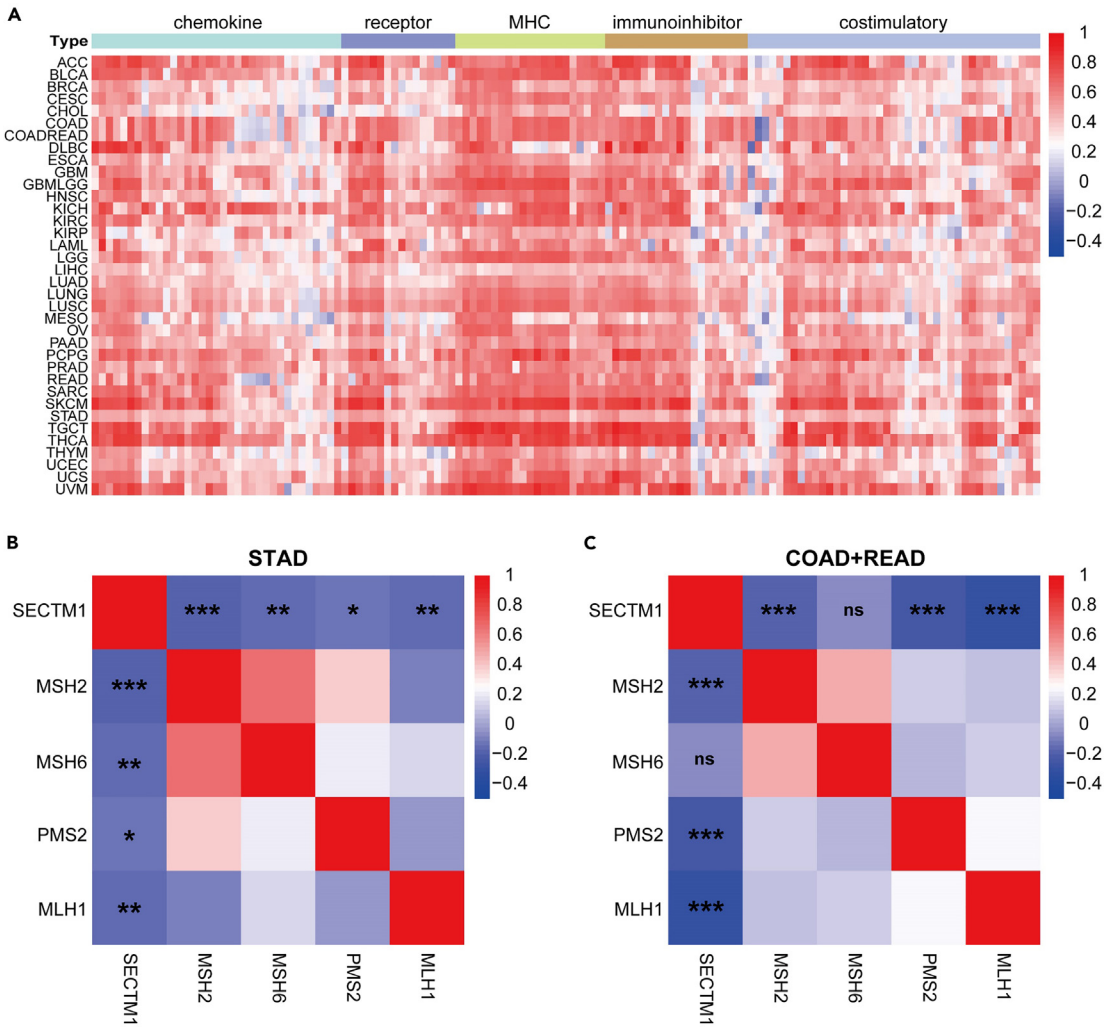


Figure 4. Pan-cancer analysis of immunological correlations of *SECTM1*

Reproduced with permission from iScience (Mei et al.[1]).

(A) Correlations between *SECTM1* and immunomodulators expression in pan-cancer, including chemokine, receptor, MHC, immunoinhibitors, and immunostimulators. Significance was calculated with Pearson correlation test.

(B) *SECTM1* was negatively correlated with DNA mismatch repair genes in gastric cancer. Significance was calculated with Pearson correlation test.*p < 0.05; **p < 0.01; ***p < 0.001.

(C) *SECTM1* was negatively correlated with DNA mismatch repair genes in colorectal cancer. Significance was calculated with Pearson correlation test.***p < 0.001.

Note: This step is usually done without any programming software, just using widely used softwares such as Office-Excel, GraphPad Prism, and SPSS.

Note: If the candidate protein could be secreted by tumor cells, it is recommended to check the level of the candidate protein in serum/plasma.

Alternatives: Immunohistochemical (IHC) staining is the most commonly used method, immunofluorescence, flow cytometry, ELISA, Western blotting, etc., can also be used.

⚠ CRITICAL: Comparing the predictive value of the candidate marker with those of classical markers such as PD-L1 helps to illustrate the importance of the candidate biomarker.

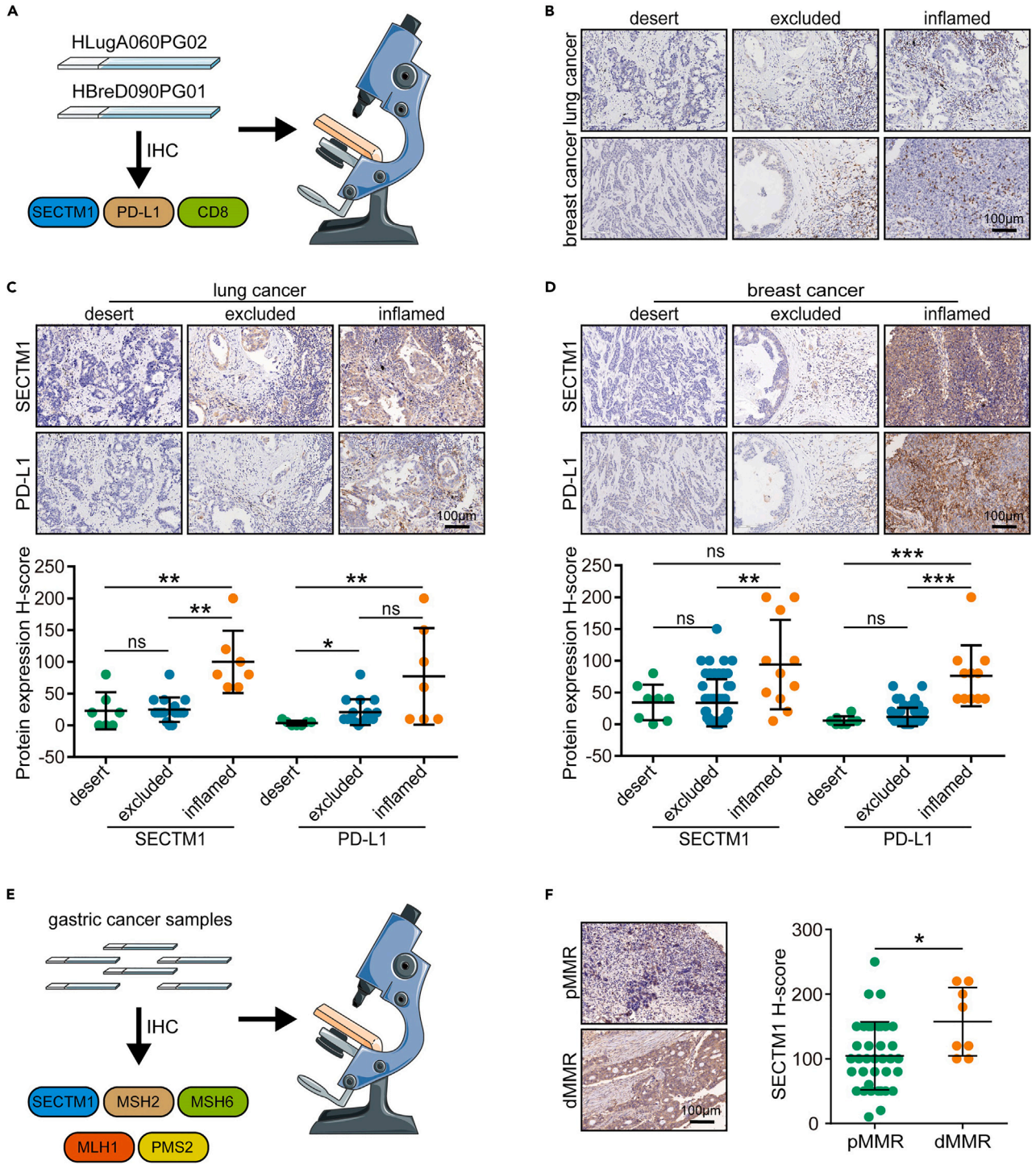


Figure 5. Correlation between SECTM1 expression and immuno-subtypes

Reproduced with permission from iScience (Mei et al.[1]).

(A) Schematic protocol of validation on the TMA cohort.

(B) Representative images revealing the distribution of $CD8^+$ T cells in tumors with different immuno-subtypes. Magnification, 200×.

(C) Representative images revealing SECTM1 and PD-L1 expression in tumors with different immuno-subtypes in lung cancer and semi-quantitative analysis of expression levels of SECTM1 and PD-L1. Magnification, 200×. Data are presented as mean ± SD. Significance was calculated with Kruskal-Wallis test with Dunn's multiple-comparison test. *$p < 0.05$; **$p < 0.01$.

Figure 5. *Continued*

(D) Representative images revealing SECTM1 and PD-L1 expression in tumors with different immuno-subtypes in breast cancer and semi-quantitative analysis of expression levels of SECTM1 and PD-L1. Magnification, 200×. Data are presented as mean ± SD. Significance was calculated with Kruskal-Wallis test with Dunn's multiple-comparison test. **p < 0.01; ***p < 0.001.

(E) Schematic protocol of validation on the recruited gastric cancer cohort.

(F) Representative images revealing SECTM1 expression in tumors with different MMR status in gastric cancer and semi-quantitative analysis of expression levels of SECTM1. Magnification, 200×. Significance was calculated with Mann-Whitney test. *p < 0.05.

EXPECTED OUTCOMES

In the case of *SECTM1*, a valid candidate biomarker should be associated with the most established biomarkers, such as PD-L1, TMB, and MSI. In addition, it is overexpressed in immuno-hot tumors and tumors from patients with well immunotherapeutic responses. Moreover, validated results from in-house cohorts could greatly increase the clinical availability of analyzed results.

QUANTIFICATION AND STATISTICAL ANALYSIS

All data are presented as means ± SDs. The statistical difference of continuous variables between the two groups is evaluated by the Student t test or Mann-Whitney test according to the applicable conditions. The difference between multiple groups is analyzed by one-way ANOVA or Kruskal-Wallis test with multiple comparisons according to the applicable conditions. The chi-square test is used when the categorical variables are assessed. Pearson or Spearman correlation test is used to evaluate the correlation between two variables according to the applicable conditions. The predictive value of the combination of SECTM1 and PD-L1 is estimated by binary logistic regression using SPSS 26. Receiver-operating characteristic (ROC) analysis is plotted to assess the specificity and sensitivity of the candidate indicator, and the area under the ROC curve (AUC) is generated for diagnostic biomarkers. Prognostic values of categorical variables are assessed by log-rank test and Cox regression analysis. For all analyses, p value < 0.05 is deemed to be statistically significant and labeled with *p < 0.05; **p < 0.01; ***p < 0.001.

LIMITATIONS

Admittedly, the current protocol still has some limitations. It is undeniable that immuno-hot tumors are not always related to favorable prognosis and immunotherapeutic response.[22] The protocol may only contribute to the identification of novel immunotherapy biomarkers based on immuno-hot features. This protocol may also be used for the identification of biomarkers for others immunotherapy, but the predictive values should be validated using corresponding clinical cohorts. In addition, validation procedures are more effective when users have a standard clinical immunotherapy trial cohort.

TROUBLESHOOTING

Problem 1

No gene symbol could be identified using the provided R code.

Potential solution

Gene expression profile should be collected into the format of genes as the row names and samples as the column names. If gene name is not stored as gene symbol, it first needs to be converted to gene symbol.

Problem 2

The validated results from the in-house cohorts do not correspond with results from the public cohort.

Potential solution

If you encounter this problem, it cannot be solved. However, we can try to avoid this fatal problem before encountering it. Generally speaking, if a candidate biomarker alone shows immunological correlation and predictive value only in individual cohorts, the biomarker is not worth further

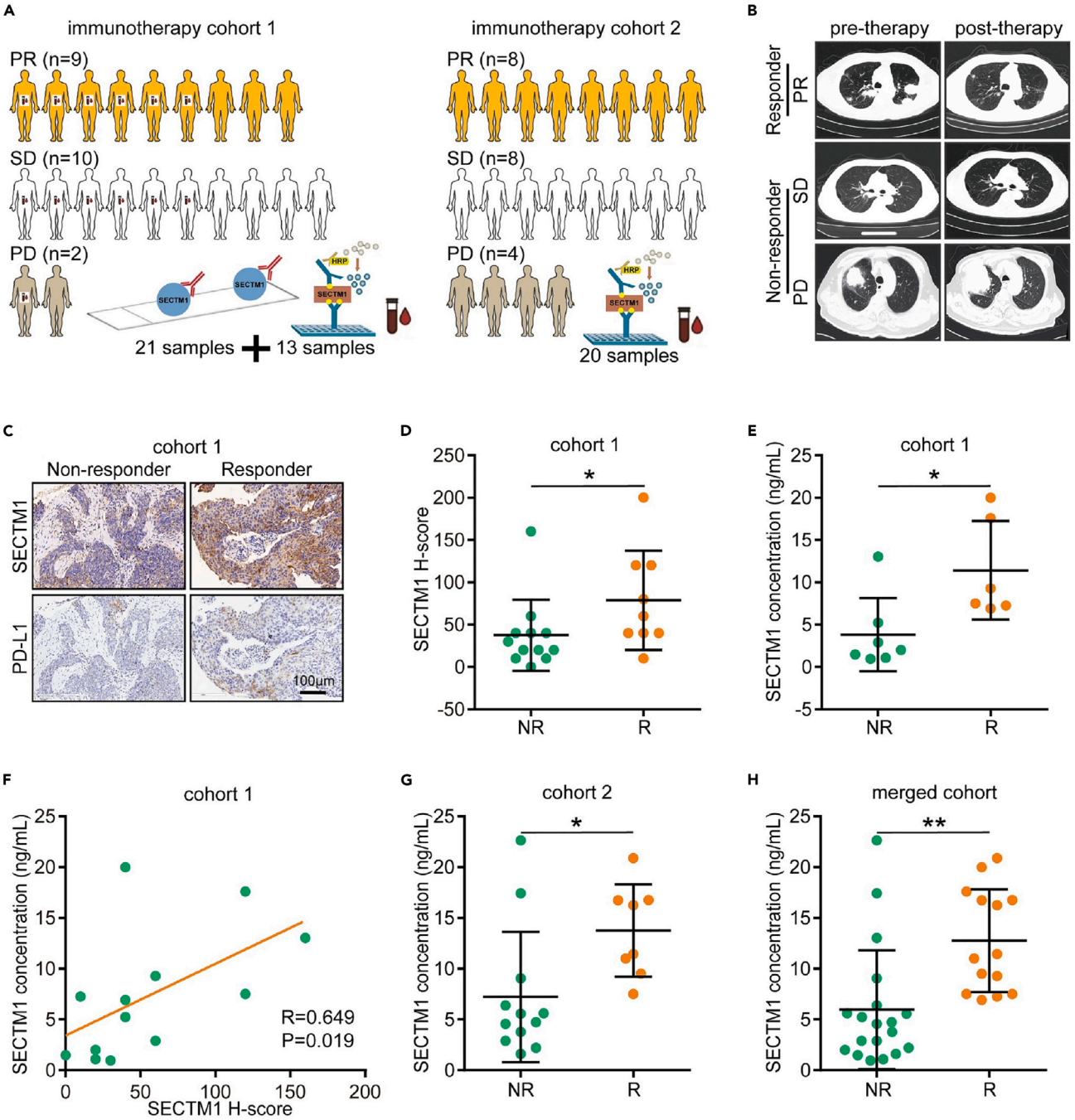


Figure 6. Validation of predictive value of SECTM1 for immunotherapy

Reproduced with permission from iScience (Mei et al.[1]).

(A) Diagram of involved lung cancer cohorts in this research.

(B) Representative CT images showing patients with different therapeutic responses.

(C) Representative images uncovering SECTM1 and PD-L1 expression in tumors from patients with different responses.

(D) Semi-quantitative analysis of expression of SECTM1 in tumors from patients with different responses in cohort 1. Significance was calculated with Mann-Whitney test. *$p < 0.05$.

(E) Circulating SECTM1 levels in patients with different responses in cohort 1. Significance was calculated with Student's t test. *$p < 0.05$.

(F) Correlation between tumor-expressed and circulating SECTM1 in cohort 1. Significance was calculated with Spearman correlation test.

(G and H) Circulating SECTM1 levels in patients with different responses in cohort 2 and merged cohort. Significance was calculated with Student's t test. *$p < 0.05$; **$p < 0.01$.

validation. The more datasets with expression difference of candidate genes between responders and non-responders, the more stable predictive value of candidate genes.

Problem 3

It is difficult to collect any in-house cohort.

Potential solution

Some commercialized clinical sample libraries could be complementary, such as Outdo BioTech (https://www.superchip.com.cn/biology/tissue.html) and Liaoding BioTech (http://www.shliaoding.com/). However, there is still no commercial cohort containing immunotherapy information. In addition, collaboration with other research groups is an encouraging solution.

Problem 4

Problems or errors in IHC assay.

Potential solution

Please follow the manufacturer's instructions, especially the antigen repair reagent corresponding to the antibody, the antibody incubation time and other parameters, which may be a process that requires multiple explorations.

Problem 5

Any other problems or errors are encountered.

Potential solution

Please e-mail us and we are happy to solve any potential problems or errors together.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Yongmei Yin (ymyin@njmu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

These accession numbers for public datasets are listed in the key resources table. The published article includes all datasets/code generated or analyzed during this study. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (No. ZDZX2017ZL-01), the High-level innovation team of Nanjing Medical University (No. JX102GSP201727), the National Natural Science Foundation of China (No. 81972484 & 82173327), Wuxi Science and Technology Bureau Foundation (No. N20202018), the Natural Science Foundation of Jiangsu Province (No. BK20210068), and Beijing Hengji Health Management and Development Foundation (No. HJ-HX-ZLXD-202209-002).

AUTHOR CONTRIBUTIONS

Conceptualization, Y.Y., J.X., W.M.; Methodology, J.M., R.X., Y. Zhu, X.Z., Y. Zhang; Investigation, J.M., Y.C., Y. Zhang; Writing-Original Draft, J.M., Y.C., R.X.; Writing-Review & Editing, Y.Y., J.X., W.M.; Funding Acquisition, Y.Y., J.X., W.M.; Supervision, Y.Y.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Mei, J., Fu, Z., Cai, Y., Song, C., Zhou, J., Zhu, Y., Mao, W., Xu, J., and Yin, Y. (2023). SECTM1 is upregulated in immuno-hot tumors and predicts immunotherapeutic efficacy in multiple cancers. iScience106027. https://doi.org/10.1016/j.isci.2023.106027.

2. Hamilton, P.T., Anholt, B.R., and Nelson, B.H. (2022). Tumour immunotherapy: lessons from predator-prey theory. Nat. Rev. Immunol. *22*, 765–775. https://doi.org/10.1038/s41577-022-00719-y.

3. Cocco, S., Piezzo, M., Calabrese, A., Cianniello, D., Caputo, R., Lauro, V.D., Fusco, G., Gioia, G.D., Licenziato, M., and De Laurentiis, M. (2020). Biomarkers in triple-negative breast cancer: state-of-the-art and future perspectives. Int. J. Mol. Sci. *21*, 4579. https://doi.org/10.3390/ijms21134579.

4. Camidge, D.R., Doebele, R.C., and Kerr, K.M. (2019). Comparing and contrasting predictive biomarkers for immunotherapy and targeted therapy of NSCLC. Nat. Rev. Clin. Oncol. *16*, 341–355. https://doi.org/10.1038/s41571-019-0173-9.

5. Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., et al. (2012). Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. N. Engl. J. Med. *366*, 2443–2454. https://doi.org/10.1056/NEJMoa1200690.

6. Garon, E.B., Rizvi, N.A., Hui, R., Leighl, N., Balmanoukian, A.S., Eder, J.P., Patnaik, A., Aggarwal, C., Gubens, M., Horn, L., et al. (2015). Pembrolizumab for the treatment of non-small-cell lung cancer. N. Engl. J. Med. *372*, 2018–2028. https://doi.org/10.1056/NEJMoa1501824.

7. Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., and Liu, X.S. (2020). Large-scale public data reuse to model immunotherapy response and resistance. Genome Med. *12*, 21. https://doi.org/10.1186/s13073-020-0721-z.

8. Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D.R., Albright, A., Cheng, J.D., Kang, S.P., Shankaran, V., et al. (2017). IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. J. Clin. Invest. *127*, 2930–2940. https://doi.org/10.1172/JCI91190.

9. Mei, J., Xu, J., Yang, X., Gu, D., Zhou, W., Wang, H., and Liu, C. (2021). A comparability study of natural and deglycosylated PD-L1 levels in lung cancer: evidence from immunohistochemical analysis. Mol. Cancer *20*, 11. https://doi.org/10.1186/s12943-020-01304-4.

10. Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat. Commun. *4*, 2612. https://doi.org/10.1038/ncomms3612.

11. Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., Li, B., and Liu, X.S. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res. *48*, W509–W514. https://doi.org/10.1093/nar/gkaa407.

12. Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D.E., and Gfeller, D. (2017). Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. Elife *6*, e26476. https://doi.org/10.7554/eLife.26476.

13. Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H., and de Reyniès, A. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. *17*, 218. https://doi.org/10.1186/s13059-016-1070-5.

14. Ru, B., Wong, C.N., Tong, Y., Zhong, J.Y., Zhong, S.S.W., Wu, W.C., Chu, K.C., Wong, C.Y., Lau, C.Y., Chen, I., et al. (2019). TISIDB: an integrated repository portal for tumor-immune system interactions. Bioinformatics *35*, 4200–4202. https://doi.org/10.1093/bioinformatics/btz210.

15. Colwill, K.; Renewable Protein Binder Working Group, and Gräslund, S. (2011). A roadmap to generate renewable protein binders to the human proteome. Nat. Methods *8*, 551–558. https://doi.org/10.1038/nmeth.1607.

16. Sun, D., Wang, J., Han, Y., Dong, X., Ge, J., Zheng, R., Shi, X., Wang, B., Li, Z., Ren, P., et al. (2021). TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. Nucleic Acids Res. *49*, D1420–D1430. https://doi.org/10.1093/nar/gkaa1020.

17. Mei, J., Cai, Y., Wang, H., Xu, R., Zhou, J., Lu, J., Yang, X., Pan, J., Liu, C., Xu, J., and Zhu, Y. (2023). Formin protein DIAPH1 positively regulates PD-L1 expression and predicts the therapeutic response to anti-PD-1/PD-L1 immunotherapy. Clin. Immunol. *246*, 109204. https://doi.org/10.1016/j.clim.2022.109204.

18. Mei, J., Cai, Y., Xu, R., Yu, X., Han, X., Weng, M., Chen, L., Ma, T., Gao, T., Gao, F., et al. (2022). Angiotensin-converting enzyme 2 identifies immuno-hot tumors suggesting angiotensin-(1-7) as a sensitizer for chemotherapy and immunotherapy in breast cancer. Biol. Proced. Online *24*, 15. https://doi.org/10.1186/s12575-022-00177-9.

19. Pilard, C., Ancion, M., Delvenne, P., Jerusalem, G., Hubert, P., and Herfs, M. (2021). Cancer immunotherapy: it's time to better predict patients' response. Br. J. Cancer *125*, 927–938. https://doi.org/10.1038/s41416-021-01413-x.

20. Shen, W., Song, Z., Zhong, X., Huang, M., Shen, D., Gao, P., Qian, X., Wang, M., He, X., Wang, T., et al. (2022). Sangerbox: a comprehensive, interaction-friendly clinical bioinformatics analysis platform. iMeta *1*, e36. https://doi.org/10.1002/imt2.36.

21. Mao, W., Cai, Y., Chen, D., Jiang, G., Xu, Y., Chen, R., Wang, F., Wang, X., Zheng, M., Zhao, X., and Mei, J. (2022). Statin shapes inflamed tumor microenvironment and enhances immune checkpoint blockade in non-small cell lung cancer. JCI Insight *7*, e161940. https://doi.org/10.1172/jci.insight.161940.

22. Chen, Y.P., Wang, Y.Q., Lv, J.W., Li, Y.Q., Chua, M.L.K., Le, Q.T., Lee, N., Colevas, A.D., Seiwert, T., Hayes, D.N., et al. (2019). Identification and validation of novel microenvironment-based immune molecular subgroups of head and neck squamous cell carcinoma: implications for immunotherapy. Ann. Oncol. *30*, 68–75. https://doi.org/10.1093/annonc/mdy470.