Supplemental information

Mesothelioma survival prediction based

on a six-gene transcriptomic signature

Kiarash Behrouzfar, Steve E. Mutsaers, Wee Loong Chin, Kimberley Patrick, Isaac
Trinstern Ng, Fiona J. Pixley, Grant Morahan, Richard A. Lake, and Scott A. Fisher
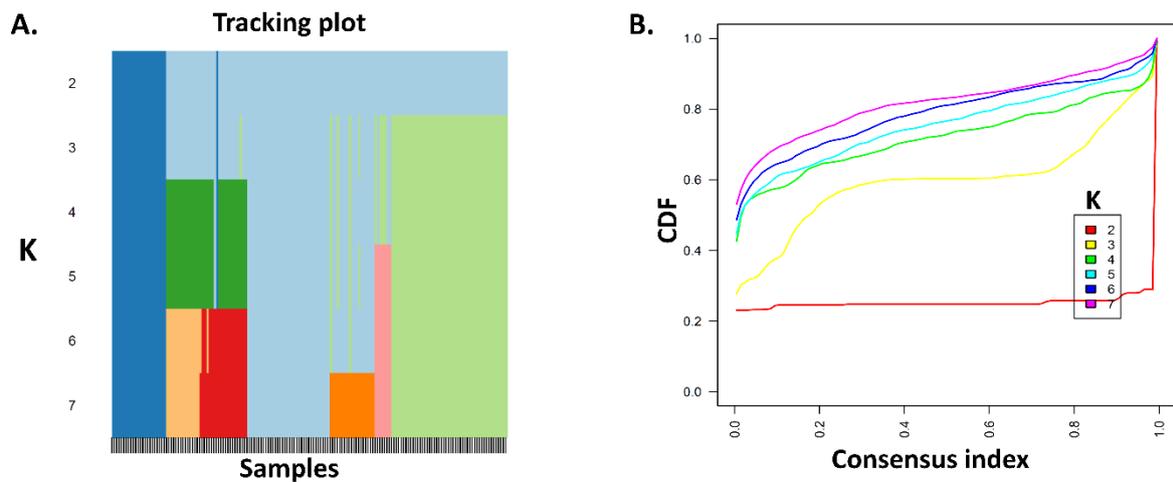
Fig. S1. Identification of the optimal cluster number using the Consensus clustering methodology
(A) Tracking bar graph shows the distribution of samples (columns) across different cluster number or K value (rows). Less changing colours within a row (K) indicates the stability of that cluster number (K) compared to other cluster number. (**B**) Cumulative distribution function (CDF) curves are shown in different colours for each cluster number assignment. A cluster number curve with the lowest alterations is considered the most stable cluster number.
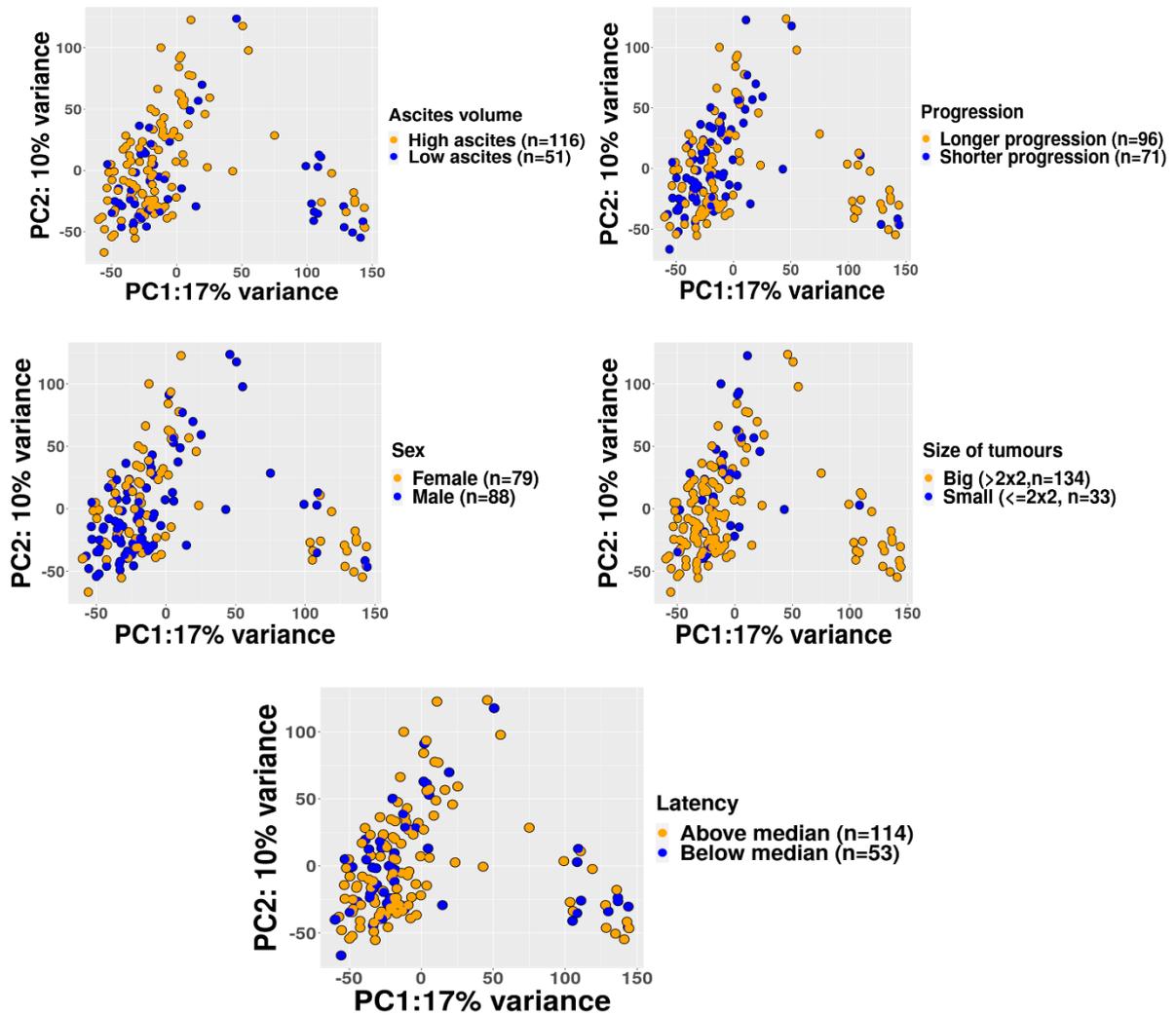
Fig. S2. PCA plots of RNA sequencing data. CC-MexTAg phenotypic traits did not demonstrate any association with tumour gene expression profiles. Each dot represents an individual tumour and colour coded based on the respective trait for each PCA plot.
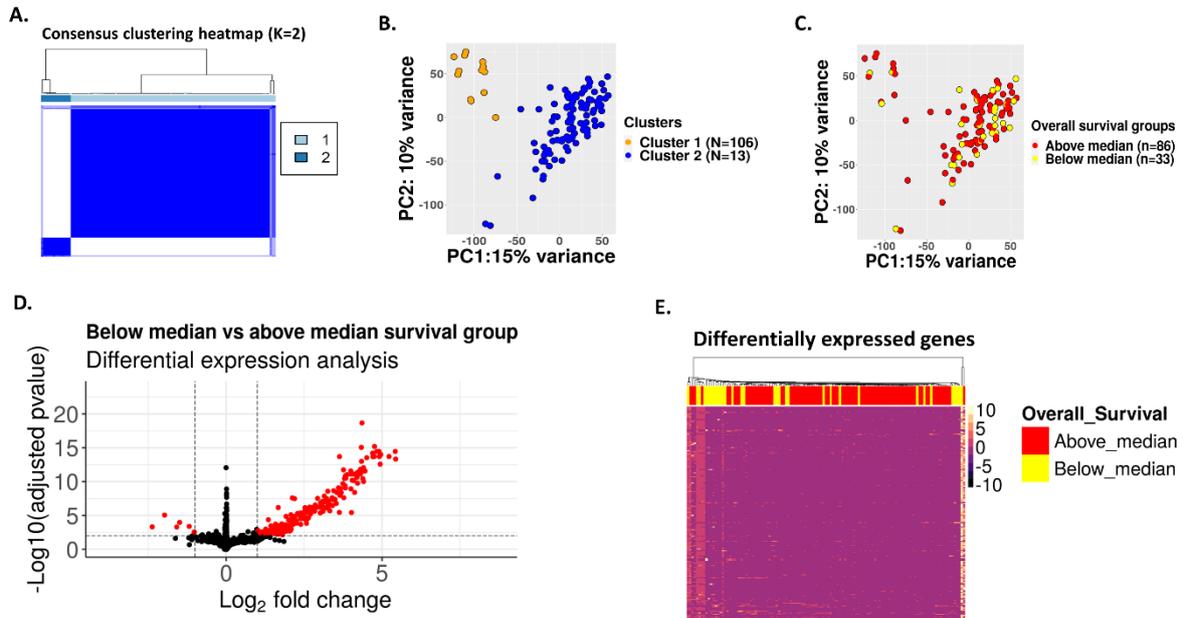
Fig. S3. Gene expression analysis using a single tumour from each mouse demonstrated two distinct tumour clusters that were not associated with overall survival of CCMT mice. (A) Hierarchical Consensus clustering heatmap generated from ConsensusClusterPlus package for K=2 selected based on the minimum overlap between two clusters using the 5000 most highly variable genes from RNAseq data. (**B**) Principal component analysis (PCA) plot of RNAseq data coloured based on consensus clustering. (**C**) PCA plot of single mouse tumours coloured in red and yellow corresponding to above median and below median survival groups, respectively. (**D and E**) Volcano plot and heatmap of differentially expressed genes in tumours from below median survival group versus above median survival group.
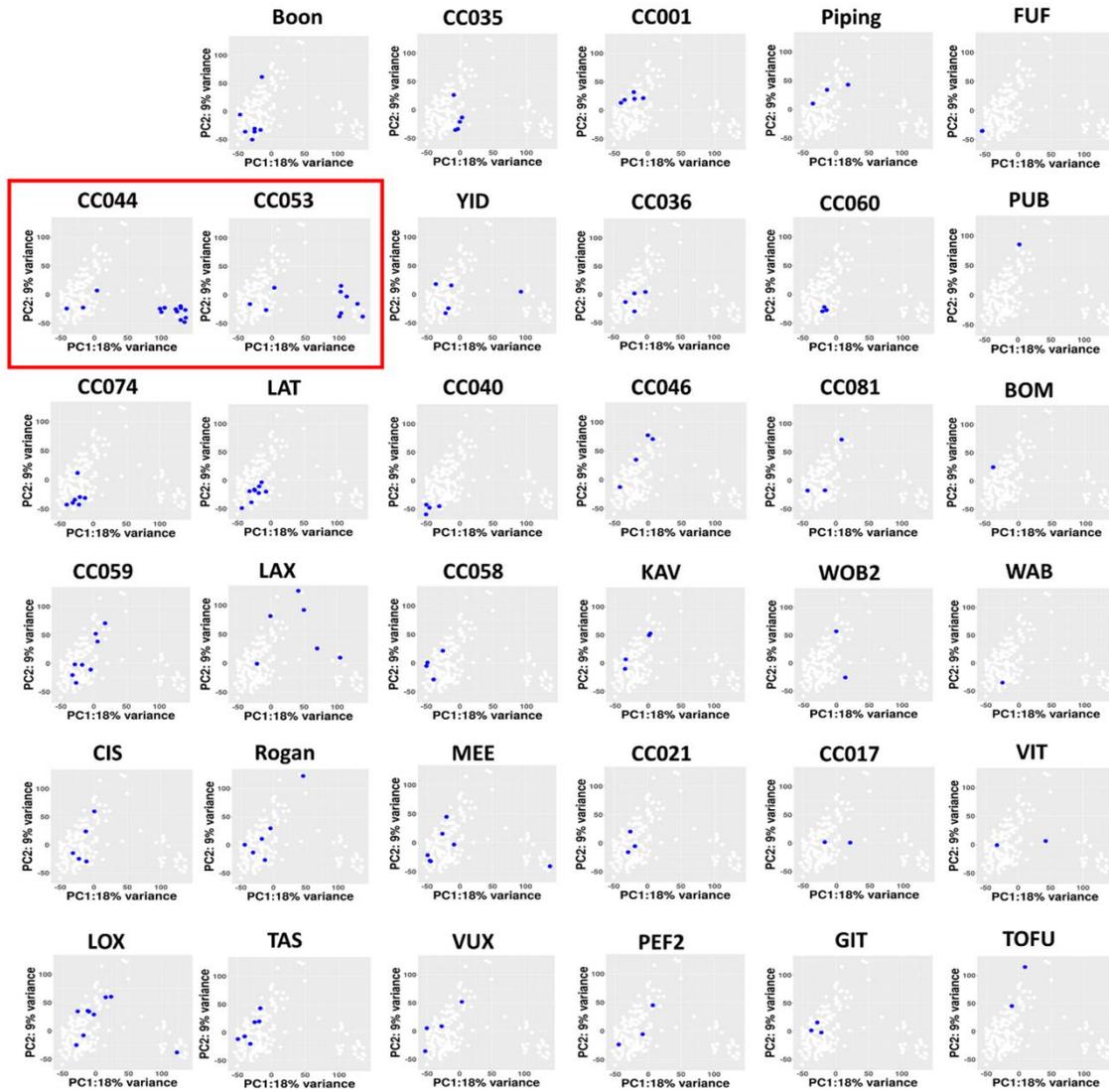
Fig. S4. Cluster 2 (immune) tumours are predominantly from CC044 and CC053 CCMT strains. CCMT stain specific PCA plots are shown. All dots from the same strain are coloured in blue. PCA plots highlighted by red outline indicate mouse strains where their tumours predominately constitute the distinct cluster (Cluster 2).
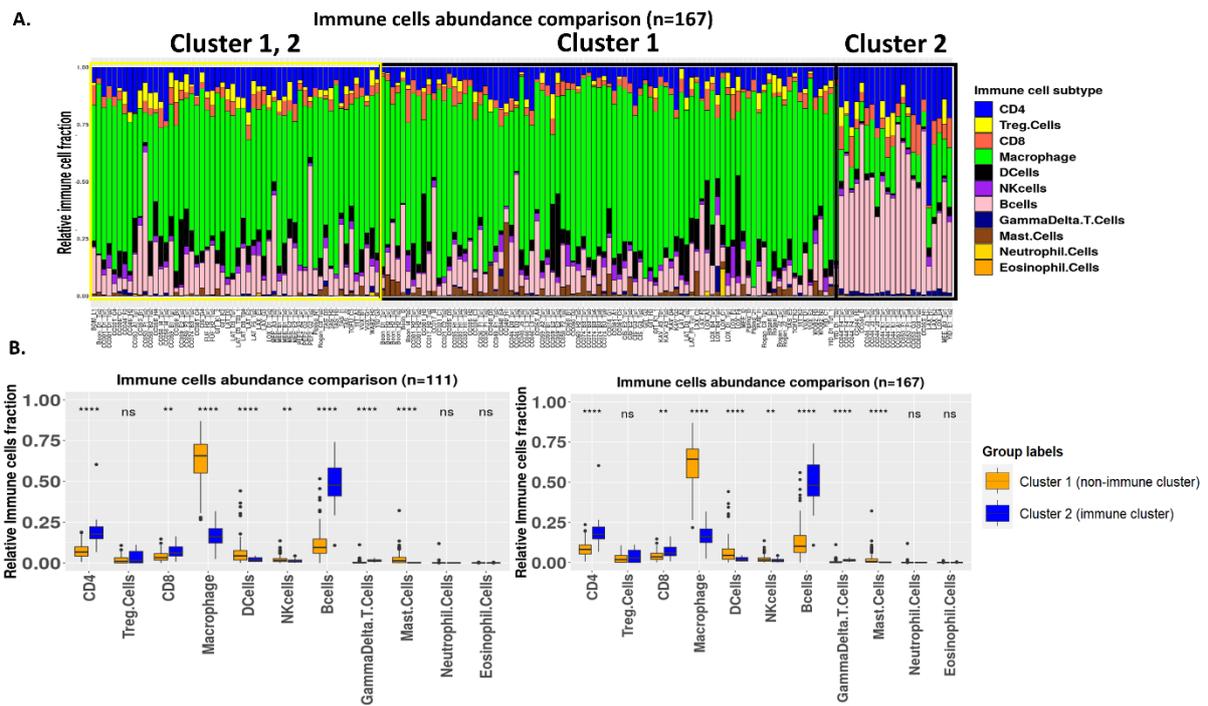
Fig. S5. Comparison of CCMT tumour cluster immune cell abundance. (A) CIBERSORT results of all 167 samples. All samples selected by Yellow rectangle correspond to samples with low deconvolution confidence of CIBERSORT algorithm, p ≥0.05. All immune cell types are shown in different colours. (**B**) Comparison between the number of immune cells in Cluster 1 and Cluster 2 tumours. Box plots on the left represent immune cells fractions using samples with significant deconvolution confidence (n=111), while Box plots on the right represent immune cells fraction using all samples (n=167). All immune cells except Treg, neutrophils and eosinophiles shown by 'ns' are significantly different between Cluster 1 and 2 using student T test. P value less than 0.05 was considered significant. *P* <0.05 *, *P* <0.01 **, *P* <0.001 ***, *P* <0.0001 ****. ns = not significant. Box and whiskers plots show median and interquartile range.
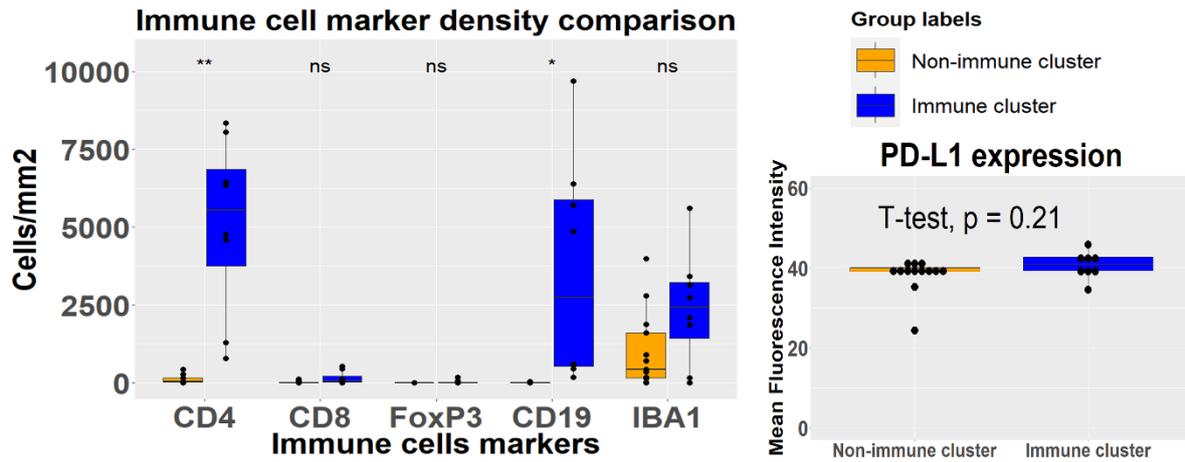
Fig. S6. Tumours from Cluster 2 (immune cluster) had significantly higher infiltration of CD4+ T and CD19+ B cells compared to Cluster 1 (non-immune cluster). Quantification of immune cell density and PD-L1 expression in whole tumour sections from immune and non-immune clusters. Density of cells expressing CD4, CD8 and FoxP3 as T cells markers, CD19 as B cell marker and IBA1 as pan macrophage marker, expressed as cells/mm^2. Expression of PD-L1 was measured in Mean Fluorescence Intensity (MFI). P <0.05 was considered significant. Box and whiskers plots show median and interquartile range.

A.

MSigDB (cluster 2)

IL2 PATHWAY
CSK PATHWAY
IL2RB PATHWAY
TCR PATHWAY
GH PATHWAY

NES

No. of significant genes
10
50
100
300

q score
0.25
0.10
0.00

B.

Anti-tumour immune response cytokines

Normalized gene expression

**** **** * *** ns ns **** **** **** ns

Tnf Ifng Tgfb1 Il1b Ifna1 Il15 Il18 Il21 Il10 Csf2
Cytokine genes

Group labels

Cluster 1 (non-immune cluster)
Cluster 2 (Immune cluster)
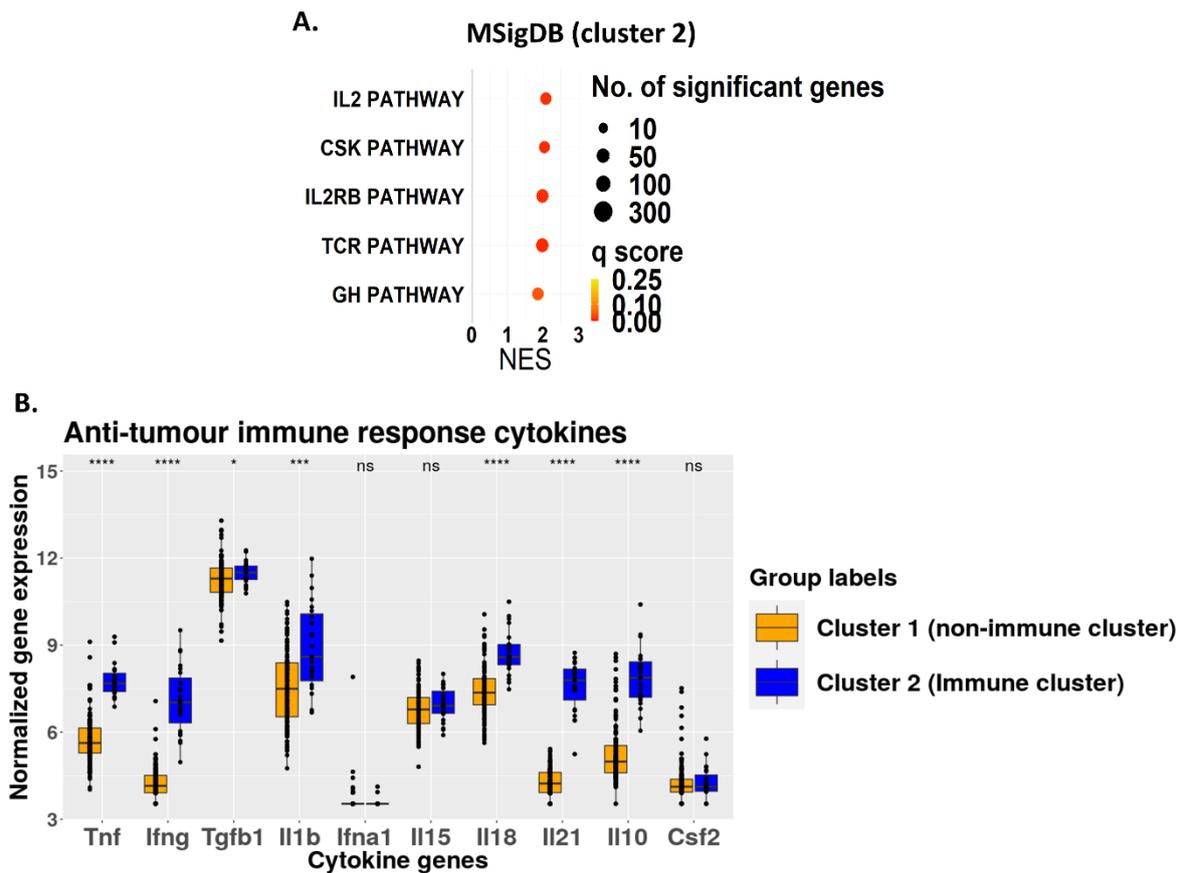
Fig. S7. Cytokines associated with anti-tumour immune responses were highly expressed in the gene expression profile of immune cluster tumours. (A) Gene set enrichment analyses (GSEA) depicting the top 5 significant gene sets enriched in tumours from immune cluster (Cluster 2) using BioCarta database. The number of genes per gene set is represented by the size of circles and the colour of the circle corresponds to the q score (false discovery rate) value for each gene set signature. **(B)** Box plots of normalised gene expression to compare the expression of 10 cytokine genes associated with anti-tumour immune response between immune and non-immune cluster tumours. *P* <0.05 was considered significant. *P* <0.05 *, *P* <0.01 **, *P*<0.001 ***, *P* <0.0001 ****. Box and whiskers plots show median and interquartile range.

A.

Non-immune cluster(n=13)

Vimentin PDPN COL1A1 E-Cadherin

62.0X

Immune cluster (n=8)

Vimentin PDPN COL1A1 E-Cadherin

66.0X

57.0X

B.

ECM/EMT expression comparison

Group labels
Cluster 1 (non-immune cluster)
Cluster 2 (immune cluster)
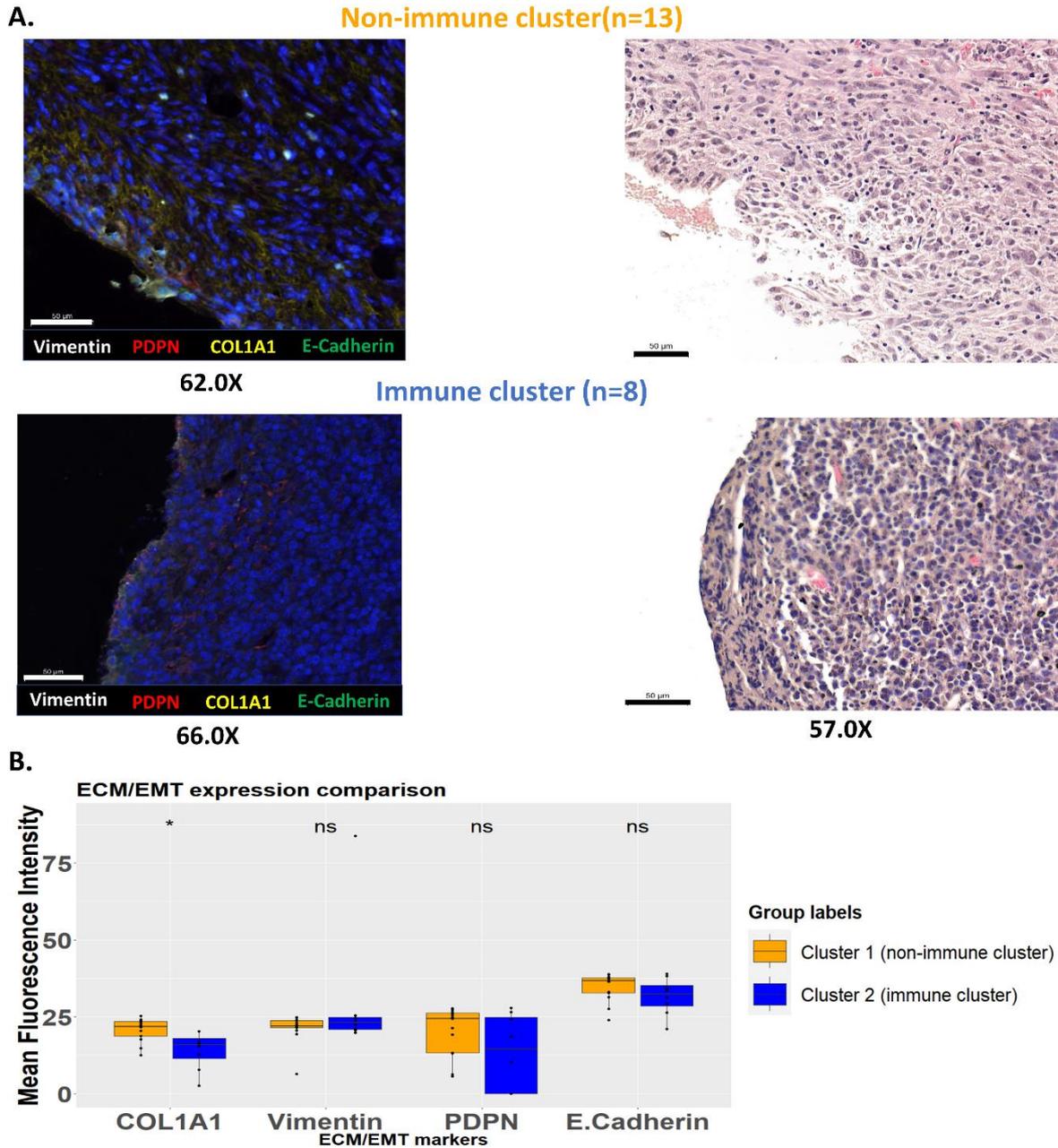
Fig. S8. Cluster 1 (non-immune) tumours expressed significantly higher levels of COL1A1 compared to Cluster 2 (immune) tumours. (A) Representative immunofluorescence and Haemotoxylin and Eosin (H&E) images of Panel 3 markers (Table S1) comparing the expression of epithelial mesenchymal transition (EMT) and extracellular matrix (ECM) markers of tumours from immune and non-immune (EMT/ECM) clusters identified in RNAseq data of CCMT mice tumours. **(B)** Immunofluorescence analysis of EMT and ECM markers between immune (Blue) and non-immune (orange) tumours. Mean Fluorescence Intensity (MFI) was used to compare the expression of each marker. Student T-test was calculated and $P < 0.05$ was considered significant. Scale bars equal 50 μm. Box and whiskers plots show median and interquartile range.

A.

EPITHELIAL MESENCHYMAL TRANSITION

TGF BETA SIGNALING

ANGIOGENESIS

UV RESPONSE DN

No. of significant genes
● 50
● 75
● 100
● 150

q score
0.4
0.3
0.2
0.1
0.0

0 1 2
NES

B.

ALLOGRAFT REJECTION

INTERFERON GAMMA RESPONSE

E2F TARGETS

SPERMATOGENESIS

No. of significant genes
● 50
● 75
● 100
● 150

q score
0.4
0.3
0.2
0.1
0.0

0 1 2
NES

C.

Module trait relationship

	Overall survival	Ascites volume	Progression	Sex	Size of tumour	Immune cluster	Non-immune cluster
MEdarkturquoise	0.032 (0.7)	-0.34 (2e-04)	0.087 (0.3)	0.035 (0.7)	0.13 (0.2)	0.52 (2e-09)	-0.52 (2e-09)
MEgrey60	-0.063 (0.5)	-0.3 (0.001)	-0.081 (0.4)	0.11 (0.2)	0.14 (0.1)	0.32 (5e-04)	-0.32 (5e-04)
MEsaddlebrown	-0.12 (0.2)	-0.16 (0.08)	-0.063 (0.5)	0.13 (0.2)	0.26 (0.004)	0.042 (0.7)	-0.042 (0.7)
MEblue	0.086 (0.4)	-0.17 (0.06)	0.11 (0.2)	-0.2 (0.03)	-0.06 (0.5)	0.8 (2e-27)	-0.8 (2e-27)
MEdarkred	0.023 (0.8)	0.27 (0.004)	-0.02 (0.8)	-0.064 (0.5)	-0.17 (0.07)	-0.17 (0.06)	0.17 (0.06)
MEorange	0.013 (0.9)	-0.0055 (1)	-0.2 (0.03)	-0.044 (0.6)	-0.24 (0.008)	-0.21 (0.02)	0.21 (0.02)
MEdarkorange	0.042 (0.6)	0.25 (0.007)	-0.016 (0.9)	-0.17 (0.07)	-0.21 (0.02)	-0.29 (0.002)	0.29 (0.002)
MEcyan	-0.062 (0.5)	0.011 (0.9)	-0.22 (0.02)	0.038 (0.7)	-0.16 (0.02)	-0.21 (0.02)	0.21 (0.02)
MEdarkgreen	0.011 (0.9)	0.031 (0.7)	-0.011 (0.9)	0.017 (0.8)	-0.23 (0.01)	0.014 (0.9)	-0.014 (0.9)
MEdarkgrey	0.011 (0.9)	0.095 (0.3)	0.16 (0.09)	0.036 (0.7)	0.12 (0.2)	-0.14 (0.1)	0.14 (0.1)
MEyellow	-0.016 (0.9)	-0.26 (0.005)	0.2 (0.04)	0.029 (0.8)	0.28 (0.002)	0.46 (2e-07)	-0.46 (2e-07)
MEred	0.13 (0.2)	-0.066 (0.5)	0.095 (0.04)	-0.1 (0.3)	-0.047 (0.6)	0.45 (3e-07)	-0.45 (3e-07)
MElightcyan	0.072 (0.4)	0.27 (0.003)	-0.018 (0.8)	0.031 (0.7)	-0.16 (0.002)	-0.57 (9e-12)	0.57 (9e-12)
MEtan	0.1 (0.3)	0.092 (0.3)	0.018 (0.8)	-0.021 (0.8)	-0.075 (0.4)	-0.0049 (1)	0.0049 (1)
MEturquoise	-0.069 (0.5)	0.28 (0.002)	-0.13 (0.1)	0.13 (0.2)	0.012 (0.9)	-0.86 (9e-37)	0.86 (9e-37)
MEviolet	-0.1 (0.3)	0.11 (0.2)	0.087 (0.3)	-0.045 (0.6)	0.2 (0.03)	-0.13 (0.2)	0.13 (0.2)

D.

Blue module
(Immune related module)

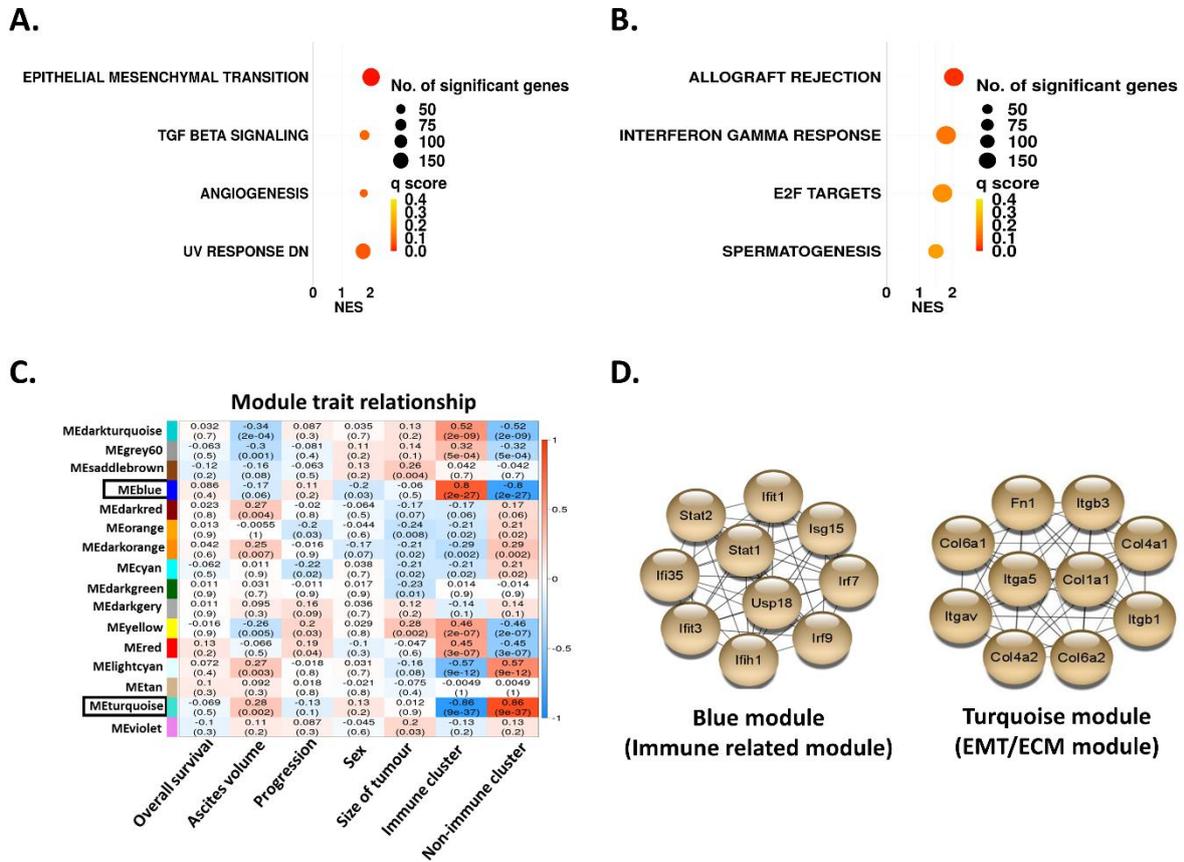**Turquoise module**
(EMT/ECM module)

Fig. S9. Gene set enrichment analysis (GSEA) and weighted correlation network analysis (WGCNA) performed on single tumours from each mouse indicate similar genetic profile and hub genes for all tumours. (A and **B)** GSEA on a single tumour per mouse using MSigDB database. Number of genes per gene set is shown by size of circles and colours of circle correspond to q score (FDR) value per each gene set signature. (**C)** Correlation matrix of modules constructed by WGCNA and correlated with mice phenotypic traits and immune and non-immune cluster. **(D)** Hub genes identified from significant WGCNA modules, Turquoise and Blue modules.
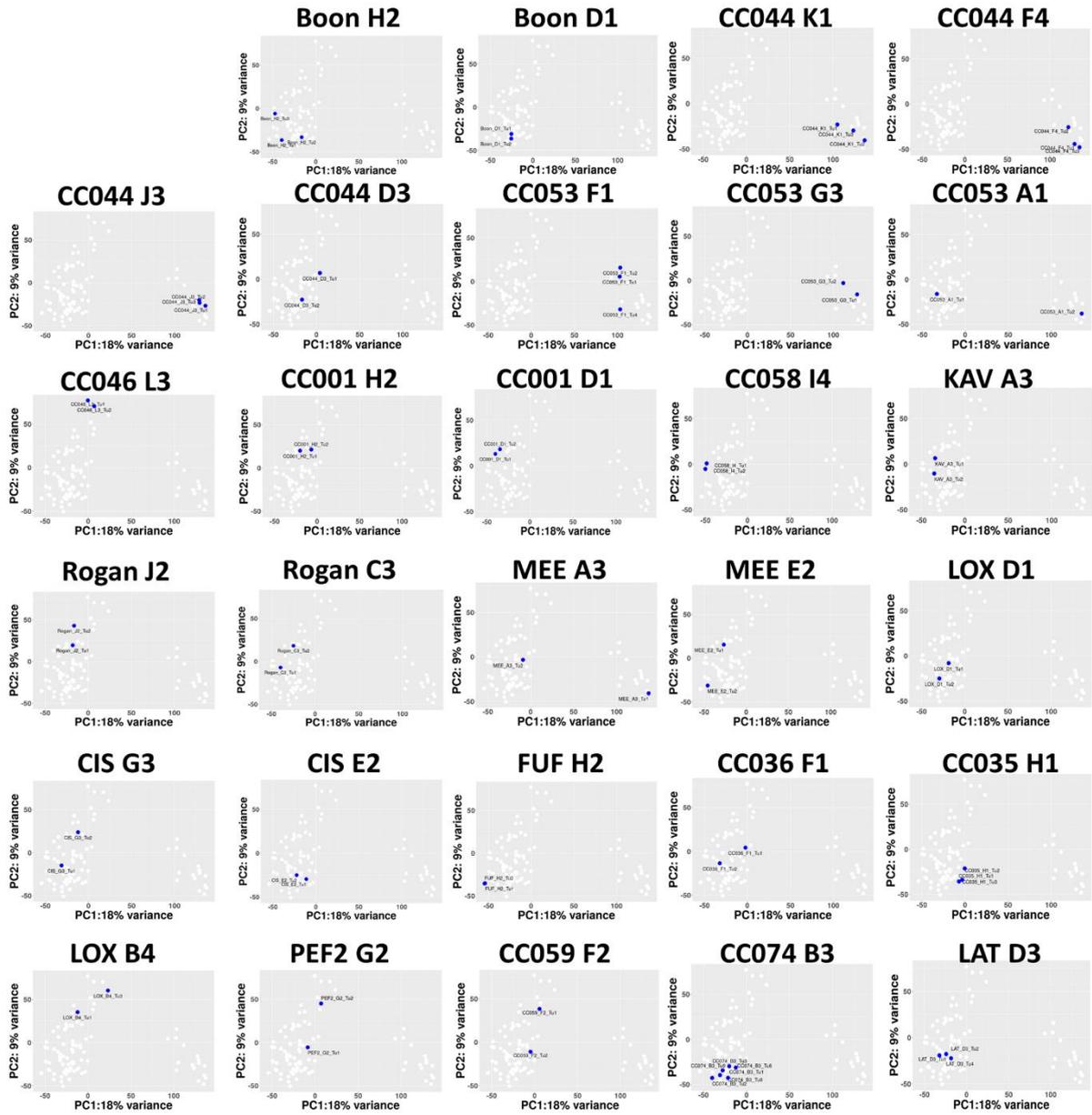
Fig. S10. RNA sequencing data from tumours harvested from mice with multiple tumours show similar gene expression profiles within each mouse. CCMT mouse-specific PCA plots demonstrating tumours harvested from the same mouse (blue dots).
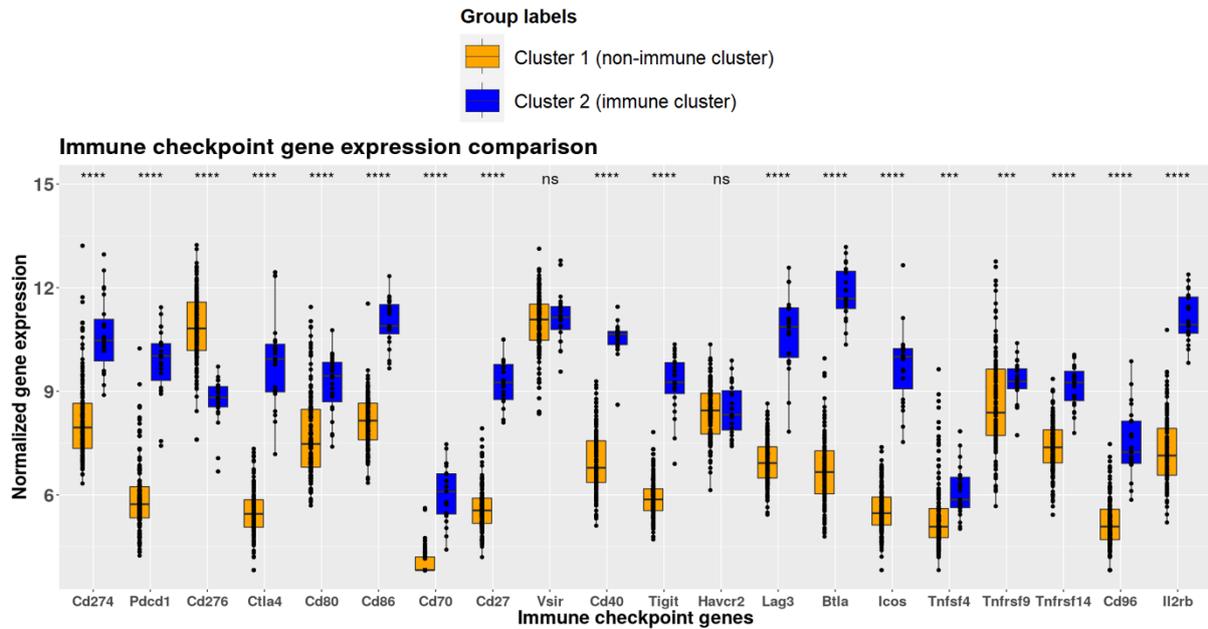
Fig. S11. Differential expression of immune checkpoint genes between CCMT tumour clusters. Normalised gene count was used to compare ICB related gene expression profiles between tumours from Cluster 1 (orange) with Cluster 2 (blue). P <0.05 was considered significant. *P* <0.05 *, *P* <0.01 **, *P*<0.001 ***, *P* <0.0001 ****. n.s. not significant. Box and whiskers plots show median and interquartile range.
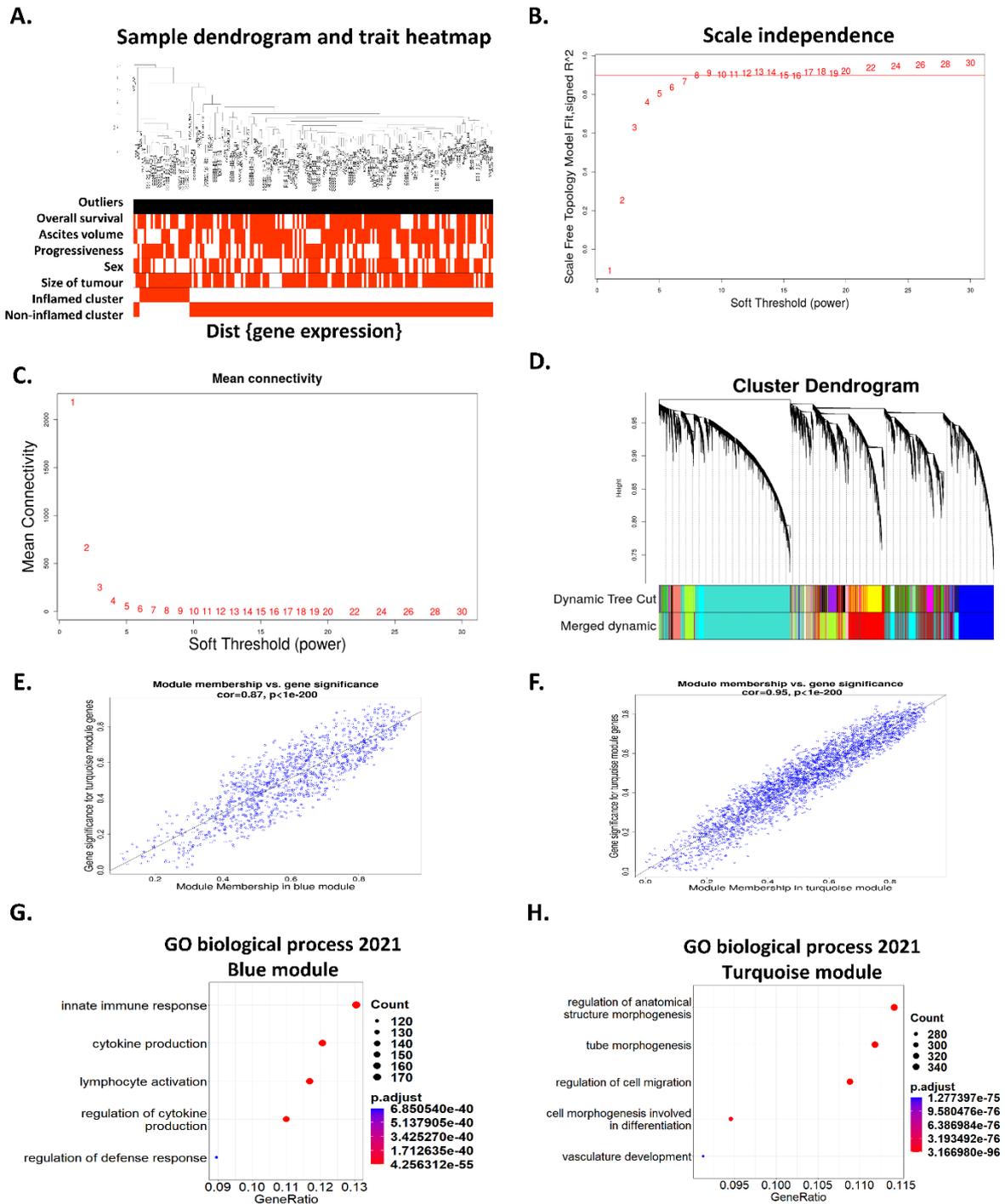
Fig. S12. Weighted correlation network analysis (WGCNA) demonstrated two significant modules associated with CCMT immune tumour cluster. (A) Samples dendrogram. Red and white bars correspond to variations in phenotypic traits in immune and non-immune tumour clusters. **(B)** Selection of "soft threshold power" based on scale-free topology analysis results to ensure scale-free network for subsequent analysis. **(C)** Mean connectivity for various soft-thresholding powers. **(D)** WGCNA cluster dendrogram demonstrates 10 colour coded modules identified by merging highly correlated modules that were primarily determined by the Dynamic Tree Cut algorithm. Grey colour corresponds to genes were not assigned to a module**. (E and F)** Module membership versus gene significance plots of Blue and Turquoise modules. **(G and H)** Functional annotation of the top 10 most significant terms identified in the Blue and Turquoise modules respectively.
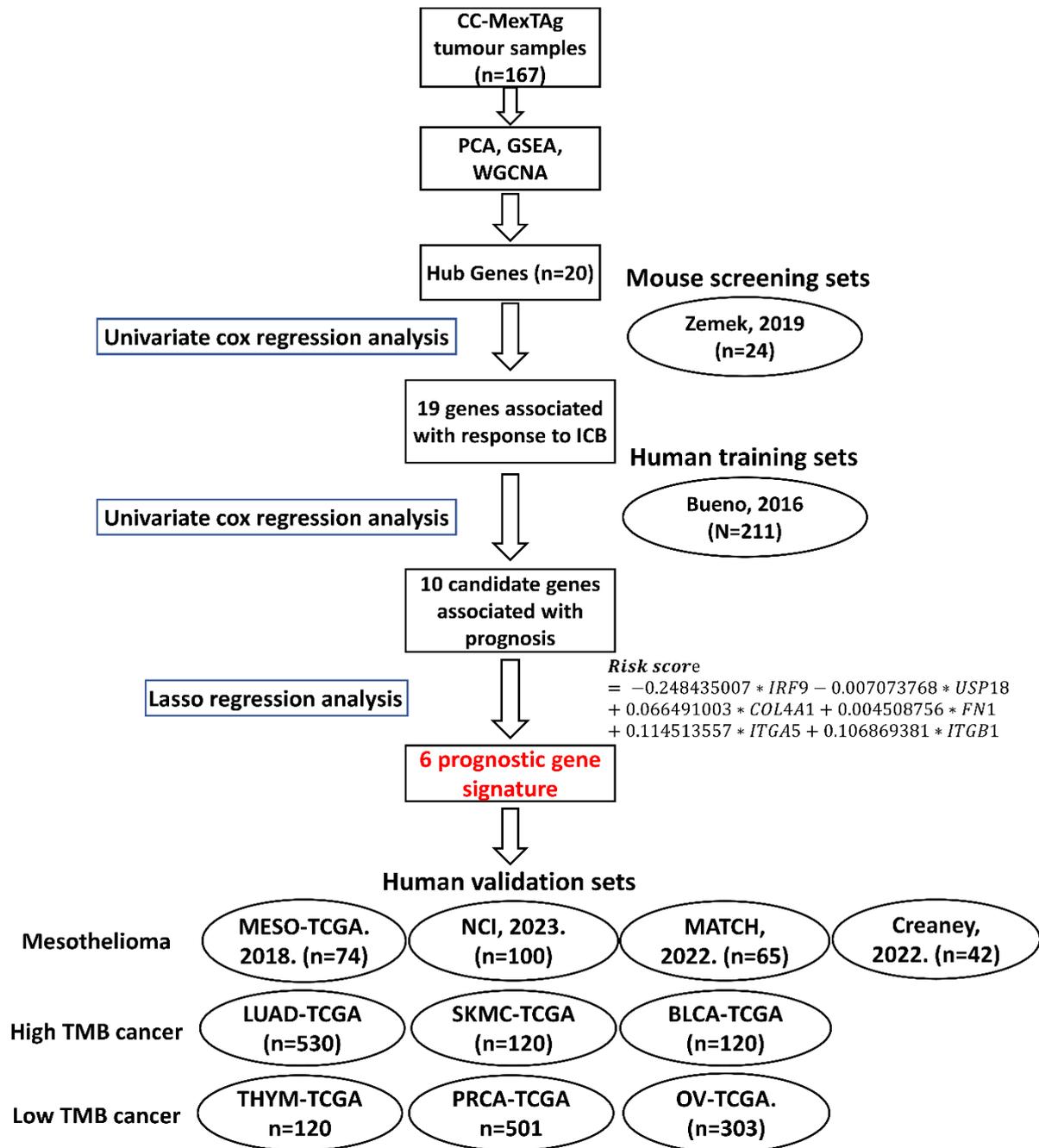
CC-MexTAg
tumour samples
(n=167)

↓

PCA, GSEA,
WGCNA

↓

Hub Genes (n=20)

Mouse screening sets

Univariate cox regression analysis

Zemek, 2019
(n=24)

↓

**19 genes associated
with response to ICB**

Human training sets

Univariate cox regression analysis

Bueno, 2016
(N=211)

↓

**10 candidate genes
associated with
prognosis**

$$Risk\ score = -0.248435007*IRF9 - 0.007073768*USP18 + 0.066491003*COL4A1 + 0.004508756*FN1 + 0.114513557*ITGA5 + 0.106869381*ITGB1$$

Lasso regression analysis

↓

**6 prognostic gene
signature**

↓

Human validation sets

Mesothelioma

MESO-TCGA.
2018. (n=74)

NCI, 2023.
(n=100)

MATCH,
2022. (n=65)

Creaney,
2022. (n=42)

High TMB cancer

LUAD-TCGA
(n=530)

SKMC-TCGA
(n=120)

BLCA-TCGA
(n=120)

Low TMB cancer

THYM-TCGA
n=120

PRCA-TCGA
n=501

OV-TCGA.
(n=303)

Fig. S13. Six-gene mesothelioma specific gene signature study workflow. Hub genes identified in CCMT tumours were screened against immune checkpoint blockade (ICB) treated mouse mesothelioma tumours. The Bueno human mesothelioma dataset was used as a training set to develop a 6-gene signature to predict mesothelioma prognosis. Four independent mesothelioma cohorts were used as validation sets to evaluate predictive performance of the gene signature. This model was also evaluated with other high and low TMB cancers.

A. MATCH (n=65)

Strata — Risk_group=High_risk — Risk_group=Low_risk

p = 0.068

Survival probability / Time (years)

Number at risk

33	4	2	1
32	13	6	0

ROC plot

True positive / False positive

AUC of 1-year survival 0.61
AUC of 2-year survival 0.69
AUC of 3-year survival 0.69

B. Creaney (n=42)

Strata — Risk_group=High_risk — Risk_group=Low_risk

p = 0.018

Survival probability / Time (years)

Number at risk

21	9	3	2	0	0
21	15	7	3	2	2

ROC plot

True positive / False positive

AUC of 1-year survival 0.71
AUC of 2-year survival 0.64
AUC of 3-year survival 0.58

C.

Variable		N	Hazard ratio		p
6-Gene signature	High_risk	33	■	Reference	
	Low_risk	32	⊢■⊣	0.85 (0.49, 1.48)	0.57
Sex	F	16	■	Reference	
	M	49	⊢■⊣	1.02 (0.57, 1.82)	0.95
Histological subtypes	eMPM	43	■	Reference	
	non-eMPM	22	⊢■⊣	2.16 (1.16, 4.00)	0.01
age		65	■	1.05 (1.01, 1.08)	0.01

0.5 1 1.5 2 2.5 3.5

D.

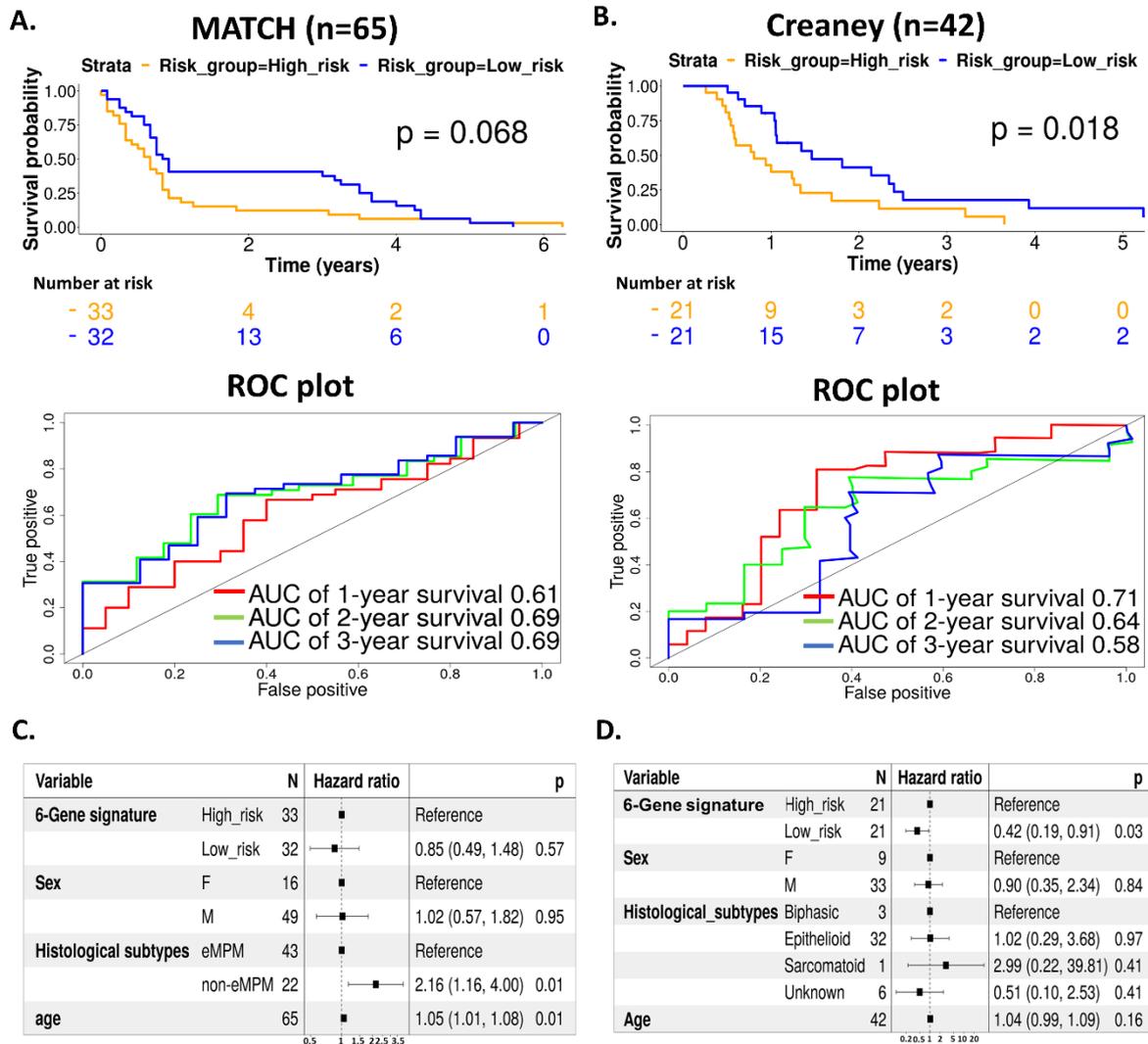| Variable		N	Hazard ratio		p
6-Gene signature	High_risk	21	■	Reference	
	Low_risk	21	⊢■⊣	0.42 (0.19, 0.91)	0.03
Sex	F	9	■	Reference	
	M	33	⊢■⊣	0.90 (0.35, 2.34)	0.84
Histological_subtypes	Biphasic	3	■	Reference	
	Epithelioid	32	⊢■⊣	1.02 (0.29, 3.68)	0.97
	Sarcomatoid	1	⊢■———⊣	2.99 (0.22, 39.81)	0.41
	Unknown	6	⊢■⊣	0.51 (0.10, 2.53)	0.41
Age		42	■	1.04 (0.99, 1.09)	0.16

0.2 0.5 1 2 5 10 20

Fig. S14. Kaplan Meier and ROC plots of MATCH and Creaney cohorts demonstrated differences in capacity of the 6-gene signature to predict outcome. (A and B) Kaplan Meier curve and ROC plots of MATCH (n=65) and Creaney cohorts (n=42) comparing high (orange) versus low (blue) risk groups of mesothelioma patients. ROC plot of MATCH (left) and Creaney (right) are shown for 1, 2, 3-year survival. (**C and D**) Forrest plot of multivariate Cox regression analysis of the gene signature in MATCH (left) and Creaney (right) datasets. Multivariate analyses were adjusted for sex, histological subtypes and age of patients. For MATCH datasets, histological subtypes were only defined as epithelioid (eMPM) and non-epithelioid (non-eMPM) subtypes. Hazard ratio, (confidence interval) and P values are shown on the right end column of the table.
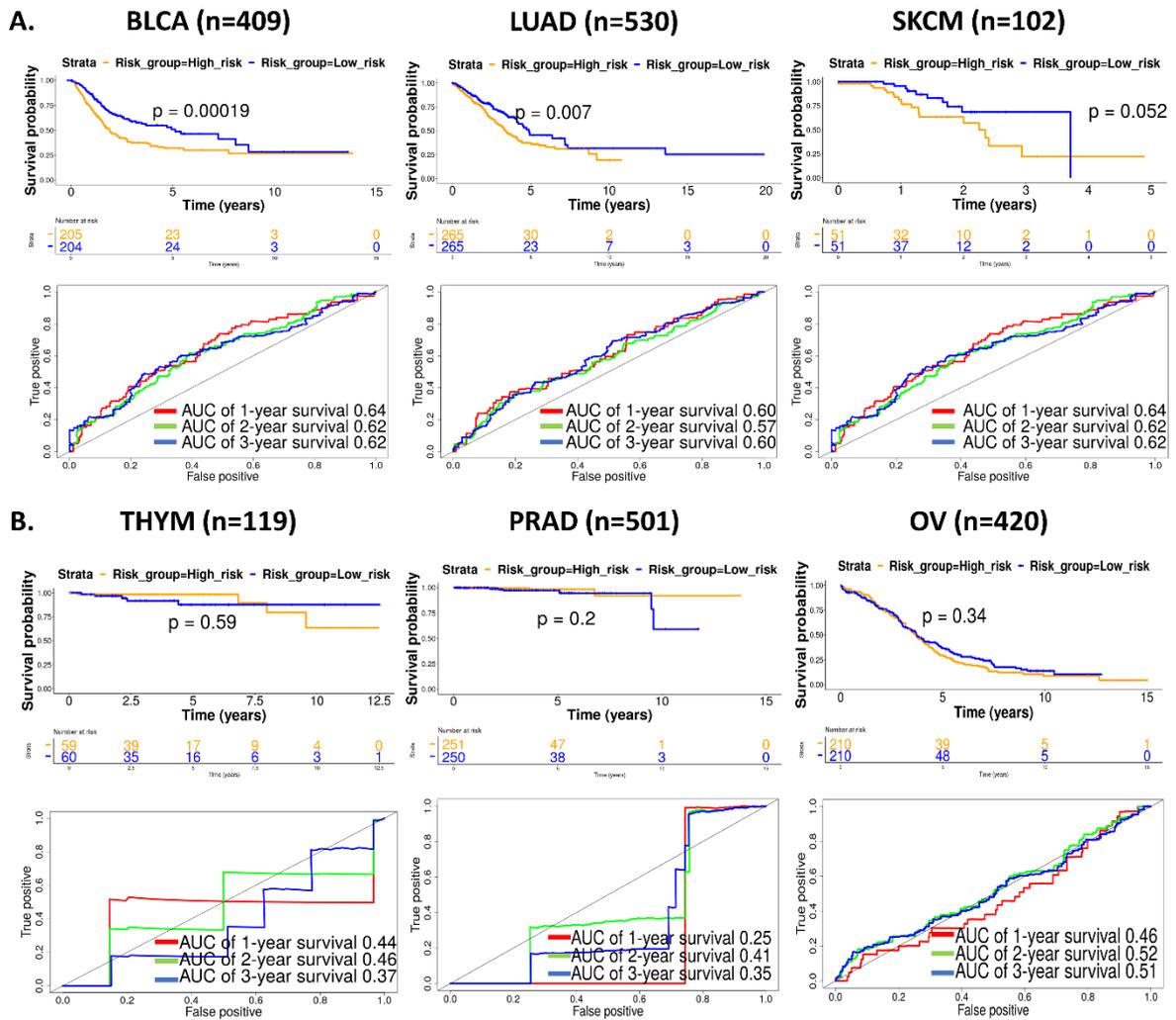
Fig. S15. Comparison of the prognostic efficacy of our 6-gene mesothelioma signature in other non-mesothelioma cancers. (A and B) Kaplan Meier and ROC curves of survival prediction by gene signature model using TCGA datasets of high TMB cancers, Bladder cancer (BLCA), Lung adenocarcinoma (LUAD) and Skin cutaneous melanoma (SKCM). (**C and D)** Kaplan Meier and ROC curves of survival prediction by gene signature model using TCGA datasets of low TMB cancers, Thymoma (THYM), Prostate adenocarcinoma (PRAD) and Ovarian cancer (OV).
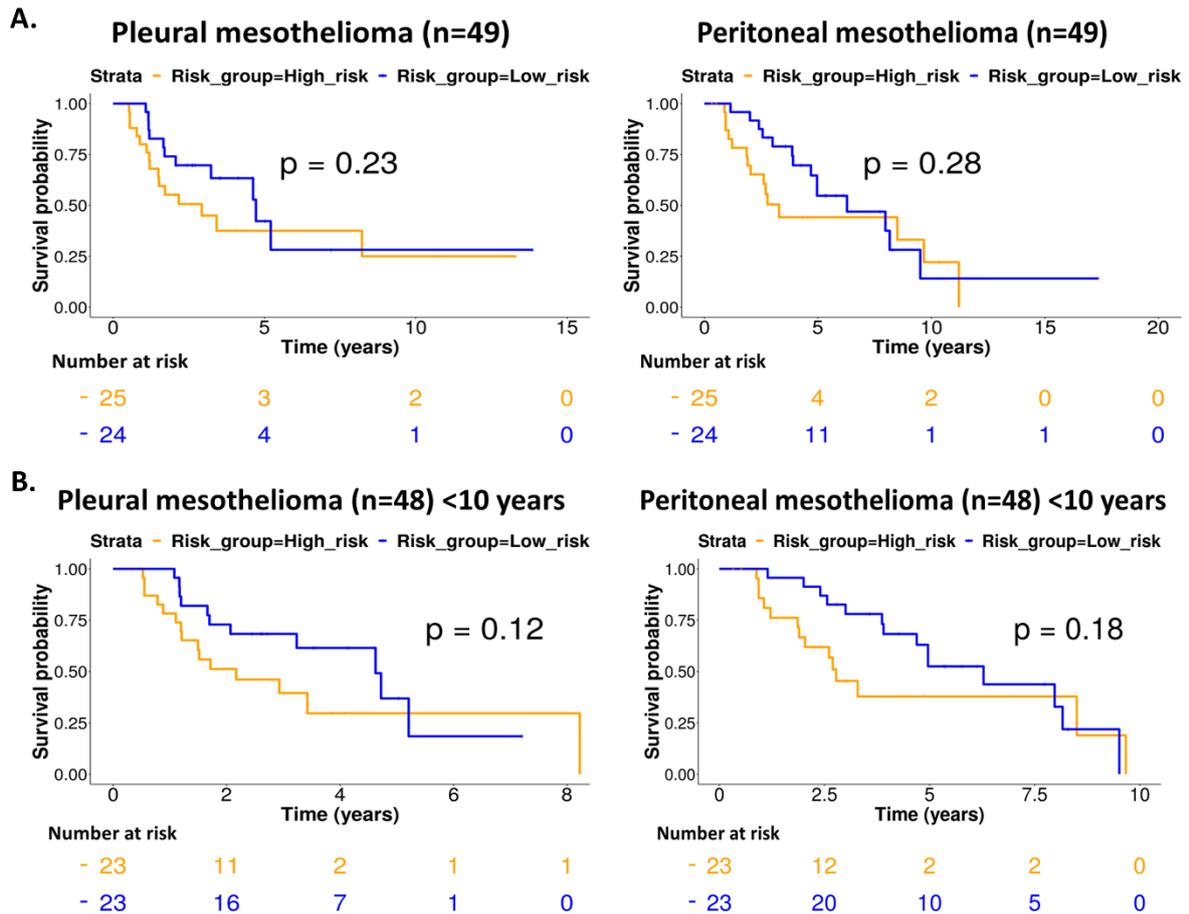
Fig. S16. Mesothelioma-specific 6-gene signature could not predict survival stratified on anatomical location. (A) Kaplan Meier curves of survival prediction by the 6-gene signature in pleural and peritoneal groups of all patients and (**B)** patients with less than 10 years survival in NCI mesothelioma cohort.

Table S1. Antibody specification, protein target and staining parameters for each multiplex immunofluorescence (mIF) panel. Goat anti-Rabbit IgG (H+L) was used as a secondary antibody.

Antibody target	Catalog number, Manufacturer	Clone number
CD4	25229S, Cell signaling	D7D2Z
CD8a	98941, Cell signaling	D4W2Z
FoxP3	12653, Cell signaling	D6O8R
PD-L1	ab213480, Abcam	EPR20529
CD19	90176, Cell signaling	D4V4B
IBA1	019-19741, FUJIFILM Wako Pure Chemical Corporation	Polyclonal
COL1A1	72026T, Cell signaling	E8F4L
Vimentin	5741T, Cell signaling	D21H3
PDPN	ab109059, Abcam	Polyclonal
E-cadherin	3195S, Cell signaling	24E10

Panel 1 (T cell and checkpoint)

Function	Antigen	Step	Dilution (working conc ug/ml)	TSA-Fluorophores reagent (dilution), incubation time
T cell markers	CD4	1	1/800 (0.19)	Cy5 (1/200), 5 min
	CD8a	2	1/800 (0.125)	AF594 (1/200), 5 min
	FoxP3	3	1/400 (1.25)	Cy3 (1/100), 10 min
Immune checkpoint	PD-L1	4	1/400 (1.27)	FITC (1/50), 10 min

Panel 2 (immune)

Function	Antigen	Step	Dilution (working conc ug/ml)	TSA-Fluorophores reagent (dilution), incubation time
T cell markers	CD4	1	1/800 (0.19)	Cy5 (1/200), 5 min
	CD8a	2	1/800 (0.125)	AF594 (1/200), 5 min
B cell marker	CD19	3	1/400 (0.1)	FITC (1/50), 10 min
Pan macrophage marker	IBA1	4	1/800 (1.27)	Cy3 (1/200), 5 min

Panel 3 (EMT/ECM/Mesothelioma)

Function	Antigen	Step	Dilution (working conc ug/ml)	TSA-Fluorophores reagent (dilution), incubation time
ECM marker	COL1A1	1	1/400 (0.03)	Cy5 (1/100), 10 min
EMT marker	Vimentin	2	1/400 (0.11)	Cy3 (1/50), 10 min
Mesothelioma marker	PDPN	3	1/400 (2.5)	AF594 (1/50), 10 min
EMT marker	E-cadherin	4	1/800 (0.06)	FITC (1/50), 10 min

Table S2. List of immune checkpoint genes for gene expression analysis.

RNAseq Gene symbols	Aliases	Function of encoded protein in cancer
Cd274	PD-L1, B7-H1	Transmembrane protein expressed on tumour and APC suppressing T cell mediated immune response
Pdcd1	PD-1, Pdc1	Immune inhibitory receptor expressed on T and B cells suppressing T cell mediated immune response
Cd276	B7-H3, B7RP-2	Transmembrane protein expressed on Tumour cells. Contributes to T cell mediated immune response
Ctla4	Cd152	Protein receptor mostly upregulated on activated T cells. Contributes to inhibition of T cell mediated immune response
Cd80	B7-1	Membrane protein expressed on T cell, B cells and APC and is a ligand for Ctla4 and CD28 protein
Cd86	B7-2	Membrane protein expressed on B cells, macrophages, dendritic cells and is a ligand for Ctla4 and CD28 protein
Cd70	Cd27l, Tnfsf7	Cytokine protein ligand binds to CD27 protein
Cd27	Tnfrsf7, S152, Tp55	TNF-receptor protein acts as a co-stimulatory immune checkpoint protein expressed on T , NK and B cells
Vsir	Vista, B7-H5	Immunoregulatory receptor protein inhibits T cell response and cytokine production expressed on T and B cells
Cd40	Bp50, Tnfrsf5	TNF-receptor protein mediating immune and inflammatory response
Tigit	Vsig9, Vstm3	Transmembrane receptor, acts as a co-inhibitory immune checkpoint expressed on T and NK cells
Havcr2	Tim3, Cd366	Transmembrane protein involved in CD8 T cell exhaustion and expressed on T cells and Myeloid cells
Lag3	Cd223	Transmembrane protein involved in CD8 T cell exhaustion and Treg activation expressed on T, B, NK and Dendritic cells
Btla	Cd272	Transmembrane glycoprotein contributed to immune suppression and expressed on T, B and dendritic cells
Icos	Ailim, Cd278	Cell surface receptor for Ctla4 and Cd28, expressed on activated T cells
Tnfsf4	Cd252, Ox40l	Cytokine binds to TNFRSF4, functions as Immune checkpoint co-stimulator mediating T cells binding to endothelial cells
Tnfrsf9	Cd137, 4-1BB	TNF family receptor protein function as immune checkpoint co-stimulator expressed on activated T cells
Tnfrsf14	Hvem, Tr2, Cd270	Cell surface receptor protein binds to TNF receptor associated factor (TRAF) activating T-cell immune response
Cd96	Tactile	Cell surface receptor protein expressed on T and NK cells and may play role on activation of immune cells adhesion
Il2rb	IL15RB, Cd122	Cell surface receptor protein that is involved in T cell immune response

Table S3. Tables of Area Under the ROC Curve (AUC) of gene signatures for survival years (year 1, 2, 3) and p values of the comparison between AUCs of the 6-gene signature against the other gene signatures. Each table refers to AUCs and p values of gene signatures calculated within each mesothelioma datasets (Bueno, TCGA, NCI, MATCH and Creaney). AUC values of every survival year are compared within each datasets using the iid-representation of the AUC estimator from TimeROC package. P value <0.05 *, Pvalue <0.01 **.

Bueno

Gene signature	AUC	Year	P value
6-gene signature	0.68	Year 1	NA
6-gene signature	0.71	Year 2	NA
6-gene signature	0.75	Year 3	NA
44-gene signature	0.69	Year 1	0.99
44-gene signature	0.76	Year 2	0.56
44-gene signature	0.78	Year 3	0.95
Bai et.al signature	0.71	Year 1	0.79
Bai et.al signature	0.75	Year 2	0.64
Bai et.al signature	0.79	Year 3	0.90
Zhou et.al signature	0.62	Year 1	0.43
Zhou et.al signature	0.65	Year 2	0.64
Zhou et.al signature	0.62	Year 3	0.19
Zhang et.al signature	0.63	Year 1	0.24
Zhang et.al signature	0.72	Year 2	0.98
Zhang et.al signature	0.75	Year 3	0.99
Shi et.al signature	0.61	Year 1	0.24
Shi et.al signature	0.66	Year 2	0.64
Shi et.al signature	0.66	Year 3	0.38
Xiao et.al signature	0.68	Year 1	0.99
Xiao et.al signature	0.68	Year 2	0.86
Xiao et.al signature	0.68	Year 3	0.43
CCNB1 expression	0.65	Year 1	0.84
CCNB1 expression	0.73	Year 2	0.97
CCNB1 expression	0.77	Year 3	0.99

TCGA

Gene signature	AUC	Year	P value
6-gene signature	0.68	Year 1	NA
6-gene signature	0.82	Year 2	NA
6-gene signature	0.87	Year 3	NA
44-gene signature	0.74	Year 1	0.777
44-gene signature	0.83	Year 2	0.994
44-gene signature	0.90	Year 3	0.955
Bai et.al signature	0.80	Year 1	0.123
Bai et.al signature	0.84	Year 2	0.939
Bai et.al signature	0.90	Year 3	0.926
Zhou et.al signature	0.66	Year 1	0.986
Zhou et.al signature	0.63	Year 2	0.056
Zhou et.al signature	0.67	Year 3	**0.008****
Zhang et.al signature	0.80	Year 1	0.059
Zhang et.al signature	0.84	Year 2	0.982
Zhang et.al signature	0.85	Year 3	0.952
Shi et.al signature	0.78	Year 1	0.422
Shi et.al signature	0.77	Year 2	0.764
Shi et.al signature	0.84	Year 3	0.949
Xiao et.al signature	0.79	Year 1	0.183
Xiao et.al signature	0.87	Year 2	0.660
Xiao et.al signature	0.89	Year 3	0.970
CCNB1 expression	0.75	Year 1	0.697
CCNB1 expression	0.79	Year 2	0.967
CCNB1 expression	0.86	Year 3	0.994

NCI

Gene signature	AUC	Year	P value
6-gene signature	0.84	Year 1	NA
6-gene signature	0.71	Year 2	NA
6-gene signature	0.72	Year 3	NA
44-gene signature	0.89	Year 1	0.693
44-gene signature	0.77	Year 2	0.680
44-gene signature	0.82	Year 3	0.180
Bai et.al signature	0.89	Year 1	0.666
Bai et.al signature	0.73	Year 2	0.983
Bai et.al signature	0.78	Year 3	0.603
Zhou et.al signature	0.75	Year 1	0.739
Zhou et.al signature	0.67	Year 2	0.934
Zhou et.al signature	0.67	Year 3	0.893
Zhang et.al signature	0.83	Year 1	0.987
Zhang et.al signature	0.68	Year 2	0.860
Zhang et.al signature	0.75	Year 3	0.888
Shi et.al signature	0.83	Year 1	0.995
Shi et.al signature	0.70	Year 2	0.995
Shi et.al signature	0.75	Year 3	0.956
Xiao et.al signature	0.82	Year 1	0.970
Xiao et.al signature	0.72	Year 2	0.992
Xiao et.al signature	0.76	Year 3	0.918
CCNB1 expression	0.83	Year 1	0.994
CCNB1 expression	0.69	Year 2	0.992
CCNB1 expression	0.75	Year 3	0.936

MATCH

Gene signature	AUC	Year	P value
6-gene signature	0.61	Year 1	NA
6-gene signature	0.69	Year 2	NA
6-gene signature	0.69	Year 3	NA
44-gene signature	0.70	Year 1	0.379
44-gene signature	0.74	Year 2	0.739
44-gene signature	0.74	Year 3	0.701
Bai et.al signature	0.73	Year 1	0.119
Bai et.al signature	0.78	Year 2	0.246
Bai et.al signature	0.79	Year 3	0.253
Zhou et.al signature	0.62	Year 1	0.993
Zhou et.al signature	0.62	Year 2	0.612
Zhou et.al signature	0.62	Year 3	0.596
Zhang et.al signature	0.61	Year 1	0.999
Zhang et.al signature	0.68	Year 2	0.984
Zhang et.al signature	0.69	Year 3	1.000
Shi et.al signature	0.64	Year 1	0.932
Shi et.al signature	0.60	Year 2	0.441
Shi et.al signature	0.58	Year 3	0.300
Xiao et.al signature	0.64	Year 1	0.864
Xiao et.al signature	0.69	Year 2	1.000
Xiao et.al signature	0.69	Year 3	0.995
CCNB1 expression	0.62	Year 1	0.990
CCNB1 expression	0.65	Year 2	0.741
CCNB1 expression	0.65	Year 3	0.755

Creaney

Gene signature	AUC	Year	P value
6-gene signature	0.71	Year 1	NA
6-gene signature	0.65	Year 2	NA
6-gene signature	0.57	Year 3	NA
44-gene signature	0.60	Year 1	0.498
44-gene signature	0.56	Year 2	0.561
44-gene signature	0.47	Year 3	0.453
Bai et.al signature	0.62	Year 1	0.643
Bai et.al signature	0.58	Year 2	0.835
Bai et.al signature	0.44	Year 3	0.676
Zhou et.al signature	0.64	Year 1	0.880
Zhou et.al signature	0.53	Year 2	0.732
Zhou et.al signature	0.35	Year 3	0.376
Zhang et.al signature	0.66	Year 1	0.715
Zhang et.al signature	0.67	Year 2	0.976
Zhang et.al signature	0.57	Year 3	1.000
Shi et.al signature	0.55	Year 1	0.361
Shi et.al signature	0.66	Year 2	0.997
Shi et.al signature	0.58	Year 3	0.999
Xiao et.al signature	0.51	Year 1	0.182
Xiao et.al signature	0.50	Year 2	0.415
Xiao et.al signature	0.32	Year 3	0.193
CCNB1 expression	0.50	Year 1	0.080
CCNB1 expression	0.55	Year 2	0.538
CCNB1 expression	0.40	Year 3	0.212

Table S4. Tables of Area Under the ROC Curve (AUC) of gene signatures, p values and confidence intervals (95%) of comparison between the performance of 6-gene signature against the other gene signatures for prediction of response to immune checkpoint blockade (ICB) therapy and chemotherapy (CTX) in NCI mesothelioma datasets. AUCs of the performance of gene signatures to predict response to **(A)** ICB therapy and **(B)** Chemotherapy are shown for univariate prediction and multivariate prediction that was adjusted for the effect of age, sex and site of disease (pleural or peritoneal). P values and confidence intervals of the comparison between 6-gene signature against other gene signatures were calculated using roc.test function and 'delong' method from the pROC package. Pvalue <0.05*, pvalue<0.01 **.

A.

Univariate

Gene signature	AUC	P value	Confidence interval
6-gene signature	0.846154	NA	NA
44-gene signature	0.846154	1.000	(-0.15, 0.15)
SELECT prediction	0.910256	0.749	(-0.45, 0.32)
Bai et.al signature	0.794872	0.513	(-0.10, 0.20)
Zhou et.al signature	0.589744	**0.021***	(0.03, 0.47)
Zhang et.al signature	0.769231	0.201	(-0.04, 0.19)
Shi et.al signature	0.794872	0.455	(-0.08, 0.18)
Xiao et.al signature	0.794872	0.810	(-0.36, 0.46)
CCNB1 expression	0.641026	**0.019***	(0.03, 0.37)

Multivariate

Gene signature	AUC	P value	Confidence interval
6-gene signature	0.48	NA	NA
44-gene signature	0.51	0.75	(-0.18, 0.13)
SELECT prediction	0.64	0.29	(-0.44, 0.13)
Bai et.al signature	0.66	0.09	(-0.38, 0.03)
Zhou et.al signature	0.43	0.62	(-0.15, 0.25)
Zhang et.al signature	0.51	0.92	(-0.58, 0.53)
Shi et.al signature	0.51	0.92	(-0.58, 0.53)
Xiao et.al signature	0.51	0.82	(-0.25, 0.20)
CCNB1 expression	0.66	0.09	(-0.38, 0.03)

B.

Univariate

Gene signature	AUC	P value	Confidence interval
6-gene signature	0.57	NA	NA
44-gene signature	0.56	0.917	(-0.24, 0.27)
SELECT prediction	0.67	0.556	(-0.43, 0.23)
Bai et.al signature	0.57	1.000	(-0.28,0.28)
Zhou et.al signature	0.46	0.622	(-0.33, 0.55)
Zhang et.al signature	0.56	0.961	(-0.53,0.56)
Shi et.al signature	0.49	0.650	(-0.25, 0.40)
Xiao et.al signature	0.63	0.790	(-0.52, 0.39)
CCNB1 expression	0.54	0.855	(-0.27, 0.32)

Multivariate

Gene signature	AUC	P value	Confidence interval
6-gene signature	0.58	NA	NA
44-gene signature	0.49	0.754	(-0.51, 0.70)
SELECT prediction	0.60	0.865	(-0.17,0.14)
Bai et.al signature	0.56	0.870	(-0.22,0.26)
Zhou et.al signature	0.50	0.788	(-0.52,0.69)
Zhang et.al signature	0.49	0.759	(-0.52, 0.71)
Shi et.al signature	0.62	0.748	(-0.24, 0.17)
Xiao et.al signature	0.52	0.848	(-0.57, 0.70)
CCNB1 expression	0.44	0.609	(-0.41, 0.70)