

## Tumor microenvironment remodeling plus immunotherapy could be used in mesenchymal-like tumor with high tumor residual and drug resistant rate

Shuai Shen<sup>1,4</sup>, Xing Liu<sup>1,4</sup>, Qing Guo<sup>1,4</sup>, Qingyu Liang<sup>2</sup>, Jianqi Wu<sup>1</sup>, Gefei Guan<sup>2</sup>, Cunyi Zou<sup>2</sup>, Chen Zhu<sup>2</sup>, Zihao Yan<sup>2</sup>, Tianqi Liu<sup>1</sup>, Ling Chen<sup>3</sup>, Peng Cheng<sup>2</sup>, Wen Cheng<sup>1</sup> & Anhua Wu<sup>1</sup>

Epithelial-mesenchymal transition (EMT) is a common process during tumor progression and is always related to residual tumor, drug resistance and immune suppression. However, considering the heterogeneity in EMT process, there is still a need to establish robust EMT classification system with reasonable molecular, biological and clinical implications to investigate whether these unfavorable survival factors are common or unique in different individuals. In our work, we classify tumors with four EMT status, that is, EMT<sup>low</sup>, EMT<sup>mid</sup>, EMT<sup>high</sup>-NOS (Not Otherwise Specified), and EMT<sup>high</sup>-AKT (AKT pathway overactivation) subtypes. We find that EMT<sup>high</sup>-NOS subtype is driven by intrinsic somatic alterations. While, EMT<sup>high</sup>-AKT subtype is maintained by extrinsic cellular interplay between tumor cells and macrophages in an AKT-dependent manner. EMT<sup>high</sup>-AKT subtype is both unresectable and drug resistant while EMT<sup>high</sup>-NOS subtype can be treated with cell cycle related drugs. Importantly, AKT activation in EMT<sup>high</sup>-AKT not only enhances EMT process, but also contributes to the immunosuppressive microenvironment. By remodeling tumor immune-microenvironment by AKT inhibition, EMT<sup>high</sup>-AKT can be treated by immune checkpoint blockade therapies. Meanwhile, we develop TumorMT website (<http://tumormt.neuroscience.org.cn/>) to apply this EMT classification and provide reasonable therapeutic guidance.

<sup>1</sup>Department of Neurosurgery, Shengji Hospital of China Medical University, Shenyang, Liaoning, China. <sup>2</sup>Department of Neurosurgery, The First Hospital of China Medical University, Shenyang, Liaoning, China. <sup>3</sup>Department of Neurosurgery, Chinese People's Liberation Army of China (PLA) General Hospital, Medical School of Chinese PLA, Institute of Neurosurgery of Chinese PLA, Beijing, China. <sup>4</sup>These authors contributed equally: Shuai Shen, Xing Liu, Qing Guo. ✉email: [chengpengcmu@cmu.edu.cn](mailto:chengpengcmu@cmu.edu.cn); [cmu071207@163.com](mailto:cmu071207@163.com); [ahwu@cmu.edu.cn](mailto:ahwu@cmu.edu.cn)

Drug resistance and tumor residue continue to be the principal limiting factors for patient's survival<sup>1,2</sup>. Clinically, tumor residue and treatment resistance are always related to epithelial-mesenchymal transition (EMT)<sup>3–5</sup>. Unfortunately, most tumors undergo EMT during tumor development<sup>3–5</sup>. Following EMT, tumors exhibit high invasive ability<sup>3</sup> and unlimited proliferation<sup>6</sup> which is responsible for treatment resistance and high residual rate. For now, somatic alteration has been regarded as primary EMT driving mechanism and gives tumor cells aggressive phenotype<sup>3,5,7,8</sup>. For example, E-cadherin mutation could directly trigger EMT process<sup>3,5</sup>. Besides, copy number variation (CNV) loss like *BRCA1* could downregulate the expression of E-cadherin<sup>7</sup>. Also, over-activated oncogenic pathway like *FOXD1* could promote EMT in glioma cells<sup>8</sup>. These somatic alterations are called intrinsic driving force of EMT. Recently, the tumor microenvironment (TME) factor has been recognized as an alternative EMT driving force. The interplay between tumor cells and non-tumor cells facilitates the EMT process of tumor cells and EMT-ed tumor cells shape TME into immunosuppressive subsequently<sup>9,10</sup>. Considering the heterogeneity of the driving force (genomic alteration or TME factor) in triggering EMT process, high EMT subtypes need to be established to investigate whether EMT-ed tumors process homogeneous or heterogeneous biological characteristics like overall survival time.

In this study, the heterogeneity of EMT status was firstly classified at the pan-cancer level through machine learning-based method. Four EMT subtypes (EMT<sup>low</sup>, EMT<sup>mid</sup>, EMT<sup>high</sup>-NOS, and EMT<sup>high</sup>-AKT) were recognized with distinct molecular, biological, and clinical features. We noticed that although with similar EMT enhancement, two mesenchymal like subtypes (EMT<sup>high</sup>-NOS and EMT<sup>high</sup>-AKT) were driven by intrinsic genomic alteration or extrinsic microenvironment interaction, respectively. EMT<sup>high</sup>-NOS was characterized with extensive genetic alteration and could be treated by cell cycle targeting drugs. As for EMT<sup>high</sup>-AKT subtype, EMT was driven by the interplay between tumor cell and macrophages through AKT pathway which is drug resistant and unresectable. Importantly, this interplay not only enhance EMT status, but also lead to the dysfunction of anti-tumor T cells which could be reverted by AKT inhibition. By remodeling tumor immune-microenvironment by AKT pathway inhibition, EMT<sup>high</sup>-AKT could be treated by immune checkpoint blockade (ICB) therapies. Moreover, we introduced TumorMT (<http://tumormt.neuroscience.org.cn>) to conduct tumor EMT classification and provide reasonable therapeutic options for patients. Taken together, we established a novel EMT classification system, which would help us clarify the EMT driving mechanisms and provide valuable treatment guidance.

## Results

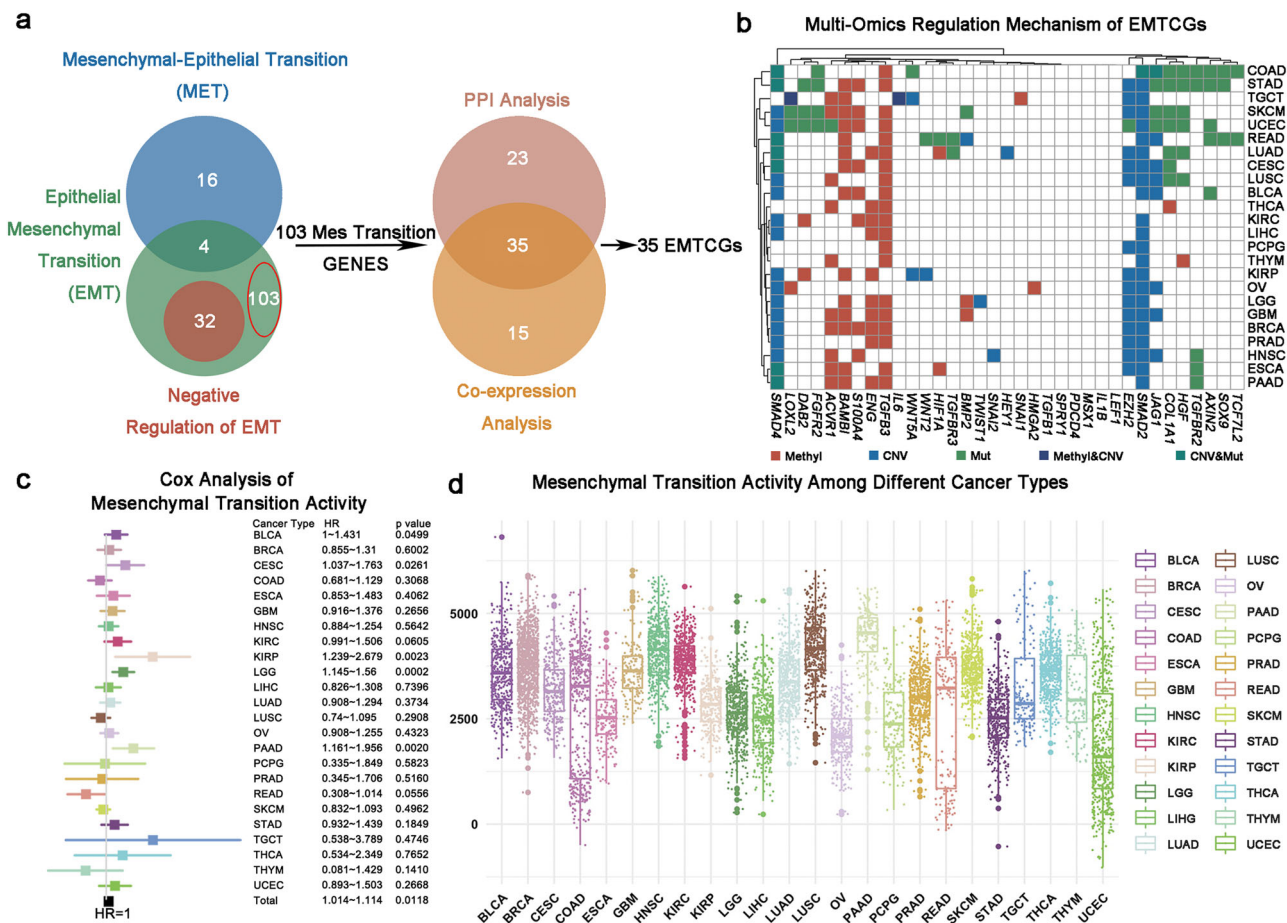
**Epithelial Mesenchymal Transition Core Gene (EMTCG) signature construction.** To summarize the core genes associated with EMT process, we collected three well-established EMT related Gene Ontology (GO) gene sets from GSEA website (<http://www.gseamsigdb.org/gsea/index.jsp>). Among these, one gene set (Go\_Epithelial\_to\_Mesenchymal\_Transition GO: 0001837) was annotated in positive correlation with EMT process, whereas the other two sets (Go\_Mesenchymal\_to\_Epithelial\_Transition GO: 0060231 and Go\_Negative\_Regulation\_of\_Epithelial\_to\_Mesenchymal\_Transition GO: 0010719) represented the negative direction (Supplementary Data 1). By excluding genes participating in GO: 0060231 and GO: 0010719 from GO: 0001837, a total of 103 genes were filtered out as being positively correlated with EMT. Then, we conducted protein-protein interaction analysis using STRING website (<https://www.string-db.org>) and identified 58 genes exhibiting strong protein

interactions (Supplementary Data 2). Meanwhile, co-expression networks of the 103 genes were also depicted using COEXPEDIA website (<https://www.coexpedia.org>), eliciting 50 hub genes with substantial expression associations (Supplementary Data 2). Genes with protein and expression relationships were intersected to obtain a list of 35 genes, namely the EMT core genes (EMTCGs, Fig. 1a and Supplementary Data 2).

To test our signature's accuracy, we downloaded other 76 EMT related signatures from EMTome<sup>11</sup> (<http://www.emtome.org>). We calculated the correlation between our signature and these 76 signatures based on ssGSEA score. We found out that our signature was strongly positively correlated with these 76 signatures (mean correlation = 0.71, Std = 0.14, Supplementary Fig. 1a). Besides, we also calculated the correlation among each two signatures of 76 signatures and we found that the correlation (mean correlation = 0.71, Std = 0.14) between our signature and other 76 signatures is better than the correlation (mean correlation = 0.63, Std = 0.20) among each two signatures of 76 signatures (Supplementary Fig. 1b, c). Next, we compared the ssGSEA score of our EMTCG signature between mesenchymal and non-mesenchymal subtype tumors in four cancer types<sup>12–15</sup>. Firstly, we found out that our signature was significantly upregulated in mesenchymal glioma in four cohorts (Supplementary Fig. 1d). Secondly, we found out that our signature was significantly upregulated in mesenchymal colorectal cancer (CRC) in 10 cohorts (Supplementary Fig. 1e). Thirdly, we found out that our signature was significantly upregulated in mesenchymal breast cancer in two cohorts (Supplementary Fig. 1f). Finally, we found out that our signature was significantly upregulated in mesenchymal gastric cancer in six cohorts (Supplementary Fig. 1g). These results demonstrated the validity of our signature.

**EMT status could be disturbed due to multi-omics mechanisms.** Next, we tried to investigate the underlying molecular mechanisms that could affect EMT status. Firstly, genomic mutation frequency of the EMTCGs was compared to the baseline mutation frequency in each cancer type. In colon adenocarcinoma (COAD), rectum adenocarcinoma (READ), skin cutaneous melanoma (SKCM), and ovarian serous cystadenocarcinoma (OV), more EMTCGs were frequently mutant compared to other cancer types (frequency >5% and fold change >1.5) (Supplementary Fig. 2a and Supplementary Data 3). Secondly, exploration of the effect of CNVs on the abnormal expression of EMTCGs indicated that the upregulation of *EZH2*, *JAG1*, *SMAD2*, and *SMAD4* might be due to the amplified copy number of these genes ( $R > 0.3$ , frequency >30%, false discovery rate (FDR) < 0.0001) (Supplementary Fig. 2b and Supplementary Data 4). Finally, we found that reduced DNA methylation levels likely contributed to the upregulation of *ACVRI*, *BAMBI*, *ENG*, *S100A4*, and *TGFB3* ( $R < -0.3$ , FDR < 0.0001) (Supplementary Fig. 2c and Supplementary Data 5). Overall, 23.3% (196/840) cases were affected by genomic mechanisms, in which 5.6% (11/196) were mediated by multiple mechanisms (Fig. 1b). Taken together, these results offered a quantitative view of how different multi-omics mechanisms might disturb mesenchymal transition in tumor samples.

**EMT represents robust clinically relevant patterns and serves as an aggressive factor.** Subsequently we assessed prognostic implications of EMT based on EMTCG score (Supplementary Data 6). By univariate COX model, we found a significant negative correlation (hazard ratio (HR): 1.014–1.114,  $p = 0.0118$ ) between the mesenchymal transition and OS times when 9415 samples were considered as a whole. Next, we conducted the



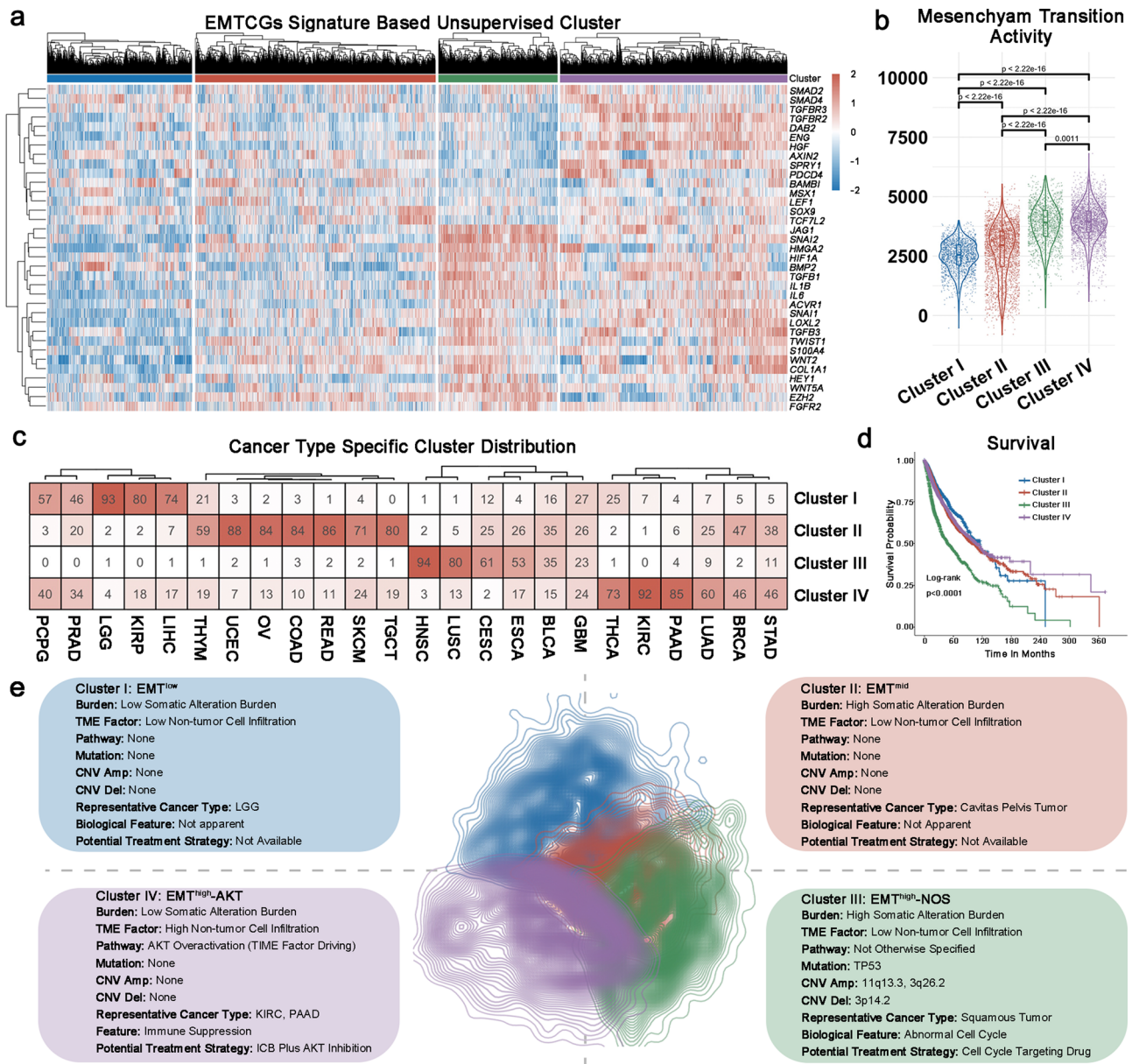
**Fig. 1** The landscape of EMTCG signature. **a** Flowchart of construction of EMTCG signature. Left panel: Venn diagram shown that 103 genes involving only in the positive direction of mesenchymal transition were summarized. Right panel: Identification of 35 EMTCGs by protein-protein interaction network analysis and co-expression analysis. **b** Multi-omics regulation mechanism including somatic mutation, copy number variation (CNV) or DNA methylation of EMTCGs. Significant correlation ( $p < 0.0001$ ) was represented on the heat-map. **c** Univariate Cox regression analysis access Hazard Ratio and prognostic significance of mesenchymal transition activity based on ssGSEA across different cancer types. **d** Boxplots of mesenchymal transition activity across 24 cancer types, as inferred using ssGSEA based on EMTCGs signature.

Cox regression analysis within each cancer type. Notably, in bladder urothelial carcinoma (BLCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), kidney renal papillary cell carcinoma (KIRP), brain lower grade glioma (LGG), and pancreatic adenocarcinoma (PAAD), a significant negative correlation remained between mesenchymal transition and OS time (Fig. 1c and Supplementary Data 7). Furthermore, we also grouped patients based on the median EMTCG score. Higher mesenchymal transition status conferred poor prognosis at a pan-cancer scale ( $p < 0.0001$ ), as did in BLCA, LGG, liver hepatocellular carcinoma (LIHC), and uterine corpus endometrial carcinoma (UCEC) (Supplementary Fig. 3). Among all the cancer types surveyed, the EMTCG score showed the most significant survival value in LGG (HR = 1.787,  $p = 0.00035$ ).

As high EMT status could reduce patients' survival, we next investigated whether mesenchymal transition was consistently enhanced in tumors (Supplementary Fig. 4). First, the EMTCG score was compared between tumor samples and their paired normal samples. Although the EMT activity was enhanced in most tumor compartments, the mesenchymal transition activity was inhibited compared to paired normal samples in some tumor compartments. For example, in head and neck squamous cell carcinoma (HNSC), the mesenchymal transition activity levels of 81.4% (35/43) of the tumor samples were significantly enhanced,

whereas the levels were inhibited compared to normal tissues in 18.6% (8/43) of the samples. Next, we calculated the coefficient of variance (CV) of EMT activity within each cancer type to explore the fluctuation status. UCEC, READ, and COAD were found to exhibit high mesenchymal transition activity fluctuation ( $CV > 50\%$ ), whereas HNSC, lung squamous cell carcinoma (LUSC), and PAAD showed low mesenchymal transition activity fluctuation ( $CV < 20\%$ ). Besides, we also found a strong negative correlation ( $R = -0.682$ ,  $p = 0.0002$ ) between the mean activity and the CV, which indicated that in the case of general low mesenchymal transition tumor types, individuals with high mesenchymal transition activity could also be observed (Fig. 1d and Supplementary Data 8). Together, these findings indicated that the malignant mesenchymal transition was heterogeneous between tumor samples, highlighting the need for further studies to define different mesenchymal-like groups.

**EMT clustering system based on EMTCG.** To discriminate tumor samples according to mesenchymal transition status, unsupervised consensus cluster was used to classify 9415 samples across 24 cancer types based on transcriptomic expression of 35 genes consisting EMTCG. Following evaluation of cumulative distribution function, Delta area, and consensus cluster data, four



**Fig. 2 Epithelial origin tumors exhibit four EMT subtypes.** **a** Unsupervised consensus cluster of 9415 samples across 24 cancer-types based on 35 EMTCGs classified tumors into four subtypes. **b** Comparison of mesenchymal transition activity among four clusters. The mesenchymal transition activity was significantly up-regulated in Cluster III and IV; the Student's t test was used to analyze statistical significance; Cluster I,  $n = 1864$ ; Cluster II,  $n = 3094$ ; Cluster III,  $n = 1532$ ; Cluster IV,  $n = 2925$ . **c** Distribution of individual cancer samples among four different EMT clusters. **d** Kaplan-Meier curve of overall survival of patients with different EMT subtypes for TCGA datasets; Cluster I,  $n = 1864$ ; Cluster II,  $n = 3094$ ; Cluster III,  $n = 1532$ ; Cluster IV,  $n = 2925$ ; Cluster III was statistically associated with unfavorable survival outcome. **e** Summarization of biological features among four EMT clusters.

EMT clusters were established (Fig. 2a and Supplementary Fig. 5a–d). PCA analysis revealed that before clustering the distribution of 24 cancer types were in chaos (Supplementary Fig. 5e) while consensus cluster could clearly classify tumor samples into four cluster (Supplementary Fig. 5f). Comparison of the EMTCG score among the four clusters indicated that Cluster I was characterized by an attenuated mesenchymal transition tendency (EMT<sup>low</sup>), Cluster II by moderate mesenchymal transition activation (EMT<sup>mid</sup>), Cluster III and IV exhibited enhanced mesenchymal transition profiles (EMT<sup>high</sup>) (Fig. 2b). Since the sample number in each cancer type varied dramatically (for example, whereas 1102 cases of breast invasive carcinoma (BRCA) were available, only 120 cases were available for thymoma (THYM)), we sampled 10 patients of each cancer in each cluster and merged them together to eliminate the bias due to

cancer type specificity for 100 times<sup>16</sup>. For 100% (100/100) iterations, Cluster III and IV exhibited enhanced mesenchymal transition profiles (Supplementary Fig. 6a).

The clinical relevance of this mesenchymal transition classification was also investigated. Specifically, squamous carcinoma including HNSC, LUSC, CESC, and esophageal carcinoma (ESCA) tended to be enriched in Cluster III, whereas adenocarcinomas including thyroid carcinoma (THCA), PAAD, lung adenocarcinoma (LUAD), BRCA, and stomach adenocarcinoma (STAD) tended to be enriched in Cluster IV (Fig. 2c). Besides, the survival time was also compared, showing that Cluster III had the shortest survival (median survival = 48.6 months), whereas Cluster I, Cluster II, and Cluster IV exhibited relatively long survival (median survival for Cluster I = 115.7 months, median survival for Cluster II = 101.4 months, median survival for Cluster

IV = 108.6 months) (Fig. 2d). To eliminate the bias due to cancer type specificity, we also introduced sampling strategy mentioned above. After sampling analysis, in 96% (96/100) iterations, Cluster III still had the shortest survival (Supplementary Fig. 6b). These analysis indicated two EMT<sup>high</sup> subtype might exist which urged us to investigate the heterogeneity between different EMT state tumors.

**Intrinsic and extrinsic features in EMT<sup>low</sup> and EMT<sup>mid</sup> tumors.** To explore potential mechanisms driving mesenchymal transition process, both intrinsic somatic alteration and extrinsic TME factor were explored for this purpose. The somatic mutation data (299 genes, which were considered as tumor driver mutations), somatic CNV data (84 loci, which were considered as tumor driver CNVs)<sup>17,18</sup>, and tumor purity were analyzed among different EMT subtypes. In this part, all results were acquired with sampling analysis for 100 iterations.

In EMT<sup>low</sup> subtype (Fig. 2e), no characteristic mutation (Supplementary Data 9) or CNVs (Supplementary Data 10) were detected. From the extrinsic viewpoint (Supplementary Fig. 7a–c), these tumors had the lowest level of non-tumor cell infiltration for 100% (100/100) iterations.

In EMT<sup>mid</sup> subtype (Fig. 2e), no characteristic mutation (Supplementary Data 9) or CNVs (Supplementary Data 10) were detected. From the extrinsic perspective (Supplementary Fig. 7a–c), the EMT<sup>mid</sup> subtype exhibited a low fraction of non-tumor cells infiltration in 100% (100/100) iterations.

**Two EMT<sup>high</sup> subtypes are driven by distinct mechanisms.** Here, we established two mesenchymal like (EMT<sup>high</sup>) subtypes with similar high EMT status. It should be important to explore the clinical, molecular and biological features between these two mesenchymal like subtypes. The clinical features between these two EMT<sup>high</sup> subtype was firstly investigated. Since the sample number in each clinical aspect varied drastically (for example, whereas 1102 cases of BRCA were available, only 120 cases were available for THYM), we made normalization of sample number to make sure our analysis does not be skewed by sample sizes as shown in Supplementary Data 11<sup>16</sup>. For example, in histological type part, the sample number of each cancer type was adjusted into 100 to calculate the sample number of every cancer type in each cluster. Firstly, we compared the difference of tumor composition fraction (Fig. 3a–c and Supplementary Fig. 8a, b). Cluster III was mainly composed of squamous carcinoma, including HNSC (24.3%), LUSC (20.8%), CESC (15.8%) and ESCA (13.7%). In total, these four cancer types composed 74.7% of Cluster III samples. As for Cluster IV type, top composition cancer types are all adenocarcinoma, including kidney renal clear cell carcinoma (KIRC, 13.3%), PAAD (12.4%) and THCA (10.55%). Secondly, we found that Cluster III was mainly composed of recurrent or primary samples while the portion distribution was relatively even in Cluster IV. Thirdly, we discovered that with histological grade progression, high EMT tumor samples tended to enrich in Cluster IV type (from Grade I to Grade IV: 15.4%, 18.5%, 17.7%, and 48.4%) instead of Cluster III type (from Grade I to Grade IV: 33.6%, 36.0%, 21.3%, and 9.1%). Lastly, age and gender distribution patterns were relatively similar between these two mesenchymal like subtypes. These results demonstrated there are substantial histopathological differences between Cluster III and IV with similar enhanced mesenchymal transition status.

Next, we tried to identify over-activated pathway in two EMT<sup>high</sup> subtypes. To do this, we first mapped the 35 genes in EMTCG to highlight their enrichment pathways using KEGG analysis. A total of 9 pathways were acquired (Supplementary Data 12), including well established pathway facilitating EMT

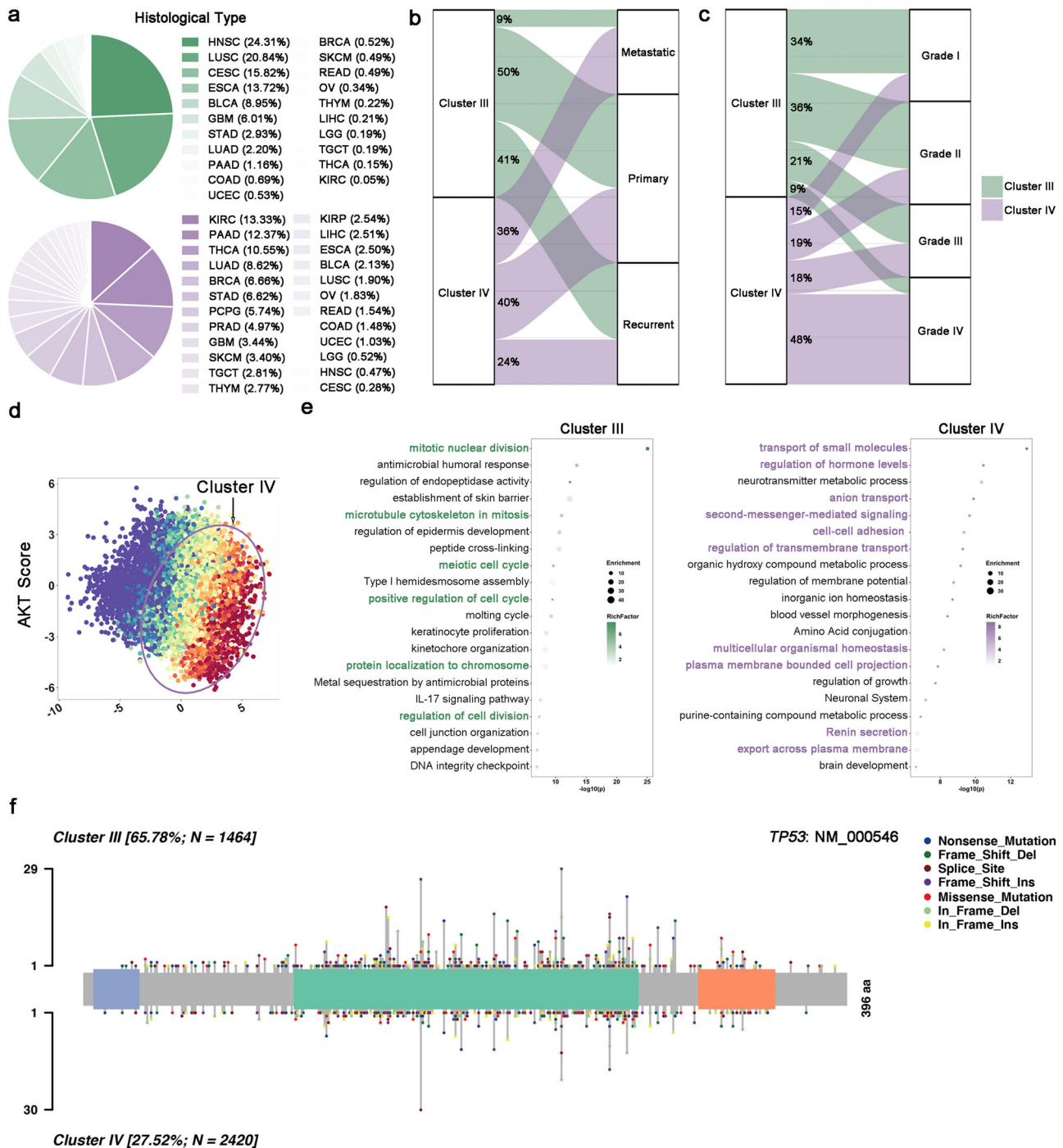
process, including TGF- $\beta$  (norm  $p = 1.5 \times 10^{-8}$ ), YAP (norm  $p = 4.645 \times 10^{-12}$ ), and AKT (norm  $p = 0.032$ )<sup>19,20</sup>. Compared to Cluster I or Cluster II, all 9 pathways were over-activated in two EMT<sup>high</sup> subtypes. When pathway activity was compared between two mesenchymal like subtypes, we found that AKT pathway was over-activated in Cluster IV in 100% (100/100) iterations (Fig. 3d and Supplementary Fig. 9i). However, we didn't find specific over-activated pathway in Cluster III (Supplementary Fig. 9). Besides, we also classified cell lines ( $n = 629$ ) from CCLE dataset into four EMT clusters (Supplementary Fig. 10a and Supplementary Data 13). Consistently, the AKT pathway were significantly enriched in Cluster IV (Supplementary Fig. 10b). To further demonstrate Cluster IV is specified with AKT pathway over-activation, a cohort of glioma samples were used for IHC analysis. The result indicated the phosphorylation level of AKT pathway was significantly upregulated in Cluster IV (Supplementary Fig. 11). Therefore, Cluster III was thus termed as the EMT<sup>high</sup>-NOS (Not Otherwise Specified) subtype, whereas Cluster IV was termed as the EMT<sup>high</sup>-AKT subtype.

Next, we attempted to further distinguish these two mesenchymal like subtypes to investigate the difference and similarity between them. To do this, we firstly performed different expression gene analysis between two EMT<sup>high</sup> subtypes using Limma R package and found out 548 genes were significantly upregulated in EMT<sup>high</sup>-NOS subtype while 558 genes were significantly upregulated in EMT<sup>high</sup>-AKT subtype (fold change  $>2$ , adjusted  $p < 0.05$ , Supplementary Data 14). Then, we performed GO analysis using genes upregulated in each mesenchymal like subtype separately. Genes in EMT<sup>high</sup>-NOS subtype were mainly enriched in GO terms related to cell cycle (Fig. 3e and Supplementary Data 15). Genes in EMT<sup>high</sup>-AKT subtype were mainly enriched in GO terms related to multi-cellular crosstalk (Fig. 3e and Supplementary Data 15).

Finally, we tried to explore the differential mesenchymal transition driving mechanisms between EMT<sup>high</sup>-NOS and -AKT subtype. In EMT<sup>high</sup>-NOS subtype, *TP53* was a recurrent mutation event in 86% (86/100) iterations (median OR = 2.20, 1.50–3.26). Compare to 27.52% mutation rate in EMT<sup>high</sup>-AKT subtype, *TP53* mutation rate is as high as 65.78% in EMT<sup>high</sup>-NOS subtype (Fig. 3f). Besides, recurrent deletion of 3p14.2 (in 86% (86/100) iterations, median OR = 1.65), and amplification of 3q26.2 (in 98% (98/100) iterations, median OR = 1.73) and amplification of 11q13.3 (in 97% (97/100) iterations, median OR = 1.85) are representative CNVs in EMT<sup>high</sup>-NOS subtype. *TP53* mutation<sup>21</sup> and 11q13.3 amplification<sup>22</sup> explain why EMT<sup>high</sup>-NOS subtype is featured with abnormal cell cycle. In EMT<sup>high</sup>-AKT subtype, no characteristic mutation or CNV was identified. From extrinsic view, EMT<sup>high</sup>-AKT subtype had the highest non-tumor cell infiltration in 100% (100/100) iterations. Besides, we found out that tumor purity negatively correlated with AKT pathway activation in EMT<sup>high</sup>-AKT subtype, indicating the role of non-tumor cells in promoting EMT of this subtype (Supplementary Fig. 7a–d).

Taken together, we demonstrated that two mesenchymal like subtypes might be viewed as distinct EMT<sup>high</sup> subtypes in future clinical practice.

**Only EMT<sup>high</sup>-AKT subtype but not EMT<sup>high</sup>-NOS was resistant to targeting drugs and shown high residual tumor rate.** EMT was always related with drug resistance<sup>20</sup>. Thus, we tried to investigate drug responsiveness between four EMT subtypes using GDSC database<sup>23</sup>. After comparing the IC<sub>50</sub> of different drugs among these clusters, we discovered that EMT<sup>high</sup>-NOS was actually sensitive to targeting drugs (Supplementary Fig. 12a). Among 445 drugs compared, 129 drugs (29%) were sensitive,



**Fig. 3 Multi-scale comparison between EMT<sup>high</sup>-NOS subtype and EMT<sup>high</sup>-AKT subtype.** **a** Histological type composition of EMT<sup>high</sup>-NOS subtype and EMT<sup>high</sup>-AKT subtype. EMT<sup>high</sup>-NOS subtype was mainly composed of squamous carcinoma while EMT<sup>high</sup>-AKT subtype was mainly composed of adenocarcinoma. **b** Sample portion composition of EMT<sup>high</sup>-NOS subtype and EMT<sup>high</sup>-AKT subtype. EMT<sup>high</sup>-NOS subtype was mainly composed of recurrent or primary samples while the portion distribution was relatively even in EMT<sup>high</sup>-AKT subtype. **c** WHO grade composition of EMT<sup>high</sup>-NOS subtype and EMT<sup>high</sup>-AKT subtype. With histological grade progression, mesenchymal transition tumor samples tended to enrich in EMT<sup>high</sup>-AKT subtype. **d** AKT pathway activity was significantly upregulated in EMT<sup>high</sup>-AKT subtype. **e** Gene ontology analysis of EMT<sup>high</sup>-NOS subtype and EMT<sup>high</sup>-AKT subtype. **f** Compared to EMT<sup>high</sup>-AKT subtype, the mutation rate of TP53 is significantly higher in EMT<sup>high</sup>-NOS subtype.

while only 6 (1%) were resistant in this subtype. Among 129 EMT<sup>high</sup>-NOS sensitive drugs, 50 (39%) drugs were related to cell proliferation, consistent with the abnormal cell cycle characteristic of EMT<sup>high</sup>-NOS subtype (Supplementary Fig. 12b). Therefore, although EMT<sup>high</sup>-NOS subtype had the worst survival status (Fig. 2d), it might be treated by targeting drugs in the future.

Next, although EMT<sup>high</sup>-AKT subtype had similar survival status to EMT<sup>low</sup> and EMT<sup>mid</sup> subtype, we compared the residual tumor rate to see whether EMT<sup>high</sup>-AKT subtype needed more non-surgical strategies compare to these low mesenchymal transition subtypes. The EMT<sup>high</sup>-AKT subtype had the highest residual tumor rate (19%, Supplementary Fig. 12c). Since EMT<sup>high</sup>-AKT subtype was characterized as AKT over-activation, we firstly

compared overall survival in this subtype between low AKT activation and high AKT activation samples. The survival status didn't alter between high AKT group and low AKT group in EMT<sup>high</sup>-AKT subtype, indicating only inhibition of AKT pathway couldn't prolong survival in EMT<sup>high</sup>-AKT subtype (Supplementary Fig. 12d). Besides, we also found that the IC50 of relative resistant drugs in EMT<sup>high</sup>-AKT subtype were not correlated with AKT pathway activity, indicating targeting drugs couldn't be sensitized by AKT inhibition (Supplementary Fig. 12e, f). Taken together, these results indicating that EMT<sup>high</sup>-AKT subtype was resistant to targeting drugs and other strategies should be developed.

**Only EMT<sup>high</sup>-AKT subtypes is associated with immunosuppressive microenvironment.** EMT is consistently associated with impaired anti-tumor immunity<sup>24</sup>. We next assessed the immune reaction among the four subtypes. The *IFNG* signature was adopted as adaptive anti-tumor immunity marker<sup>25</sup>, whereas the *TNF* signature as innate pro-inflammatory indicator<sup>26</sup>. Compared to EMT<sup>low</sup> subtype, both *IFNG* activity and *TNF* activity were upregulated in the two EMT<sup>high</sup> subtypes. However, when compared to EMT<sup>mid</sup> subtype, the *IFNG* activity is similar between EMT<sup>high</sup>-AKT subtype and EMT<sup>mid</sup> subtype in 54% (54/100) iterations, whereas the *IFNG* activity is higher in EMT<sup>high</sup>-NOS subtype than in EMT<sup>mid</sup> subtype in 100% (100/100) iterations (Supplementary Fig. 13a). As for pro-inflammatory status, EMT<sup>high</sup>-AKT subtype exhibits the highest *TNF* activity in 100% (100/100) iterations (Supplementary Fig. 13b). The unparallel innate and adoptive anti-tumor immunity in EMT<sup>high</sup>-AKT subtype suggested that adaptive anti-tumor immunity might be inhibited in EMT<sup>high</sup>-AKT subtype. To verify whether anti-tumor immunity was suppressed in the EMT<sup>high</sup>-AKT subtype, we used gene set enrichment analysis (GSEA) to profile four immunosuppressive gene signatures. We found that these signatures were all enriched in the EMT<sup>high</sup>-AKT subtype compared to non EMT<sup>high</sup>-AKT subtype patients (Supplementary Fig. 13c).

Next, we further explored the immunosuppressive features in EMT<sup>high</sup>-AKT patients. A lymphocyte signature<sup>27</sup> was applied to profile the CD8 + T cell fraction in the TME. Compared to other subtypes, the EMT<sup>high</sup>-AKT subtype exhibited similar CD8 + T cell infiltration level in 97% (97/100) iterations (Supplementary Fig. 13d). The T cell dysfunction signature<sup>25</sup> was then adapted to evaluate the T cell function, which revealed that the EMT<sup>high</sup>-AKT subtype presented the highest dysfunction score level in 100% (100/100) iterations (Supplementary Fig. 13e). As the major immune regulatory cells contributing to T cell dysfunction, the proportions of regulatory T cells (Tregs) and tumor associated macrophages (TAMs) were profiled among EMT subtypes. We found that TAM abundance and immunosuppressive M2-TAMs were significantly enriched in the EMT<sup>high</sup>-AKT subtype compared to other EMT subtypes in 100% (100/100) iterations (Supplementary Fig. 13f), suggesting the important role of TAMs in facilitating immunosuppression in EMT<sup>high</sup>-AKT subtype. As for Tregs, the abundance in EMT<sup>high</sup>-AKT type was not significantly enhanced compared to that in EMT<sup>mid</sup> or EMT<sup>high</sup>-NOS subtype in 98% (98/100) iterations, suggesting Tregs might not participate in shaping immunosuppression in EMT<sup>high</sup>-AKT subtype (Supplementary Fig. 13g).

**AKT activation dependent tumor-TAM feedback is important for shaping immunosuppression only in EMT<sup>high</sup>-AKT subtype.** EMT related immunosuppression is always related with TAMs<sup>9,28</sup>. We next tried to investigate the role of AKT pathway in regulating the crosstalk between mesenchymal transition and TAM manipulation in different EMT subtypes. Breast cancer cell

lines and melanoma cell lines were used for this purpose, since breast cancer had the largest sample number ( $n = 1102$ ) in TCGA cohort while melanoma is the first cancer type approved by FDA to use immune checkpoint blockade (ICB) therapy<sup>29</sup>. Four breast cancer cell lines (T-47D: EMT<sup>low</sup> subtype, Luminal A; MDA-MB-468: EMT<sup>mid</sup> subtype, basal like; HCC38: EMT<sup>high</sup>-NOS subtype, claudin-low; MDA-MB-231: EMT<sup>high</sup>-AKT subtype, claudin-low) and four melanoma cancer cell lines (MeWo: EMT<sup>low</sup> subtype; A-375: EMT<sup>mid</sup> subtype; SK-MEL-3: EMT<sup>high</sup>-NOS subtype; WM-115 EMT<sup>high</sup>-AKT subtype) were used for this purpose.

We firstly tested the mesenchymal transition status of tumor cells after being co-cultured with macrophages. Two mesenchymal markers, N-cadherin (N-cad) and Vimentin (VIM) were tested. N-cad and VIM were consistently upregulated in MDA-MB-231 cell after macrophage co-culturing. To test the role of the AKT pathway in this phenomenon, tumor cells were pre-treated with an AKT inhibitor (MK2206). MK2206 could only inhibited the facilitating mesenchymal transition promoting effects of TAMs on MDA-MB-231 cell line (Fig. 4a-d). These results indicated a specific mesenchymal transition facilitating role of TAMs on EMT<sup>high</sup>-AKT subtype mediated by AKT pathway.

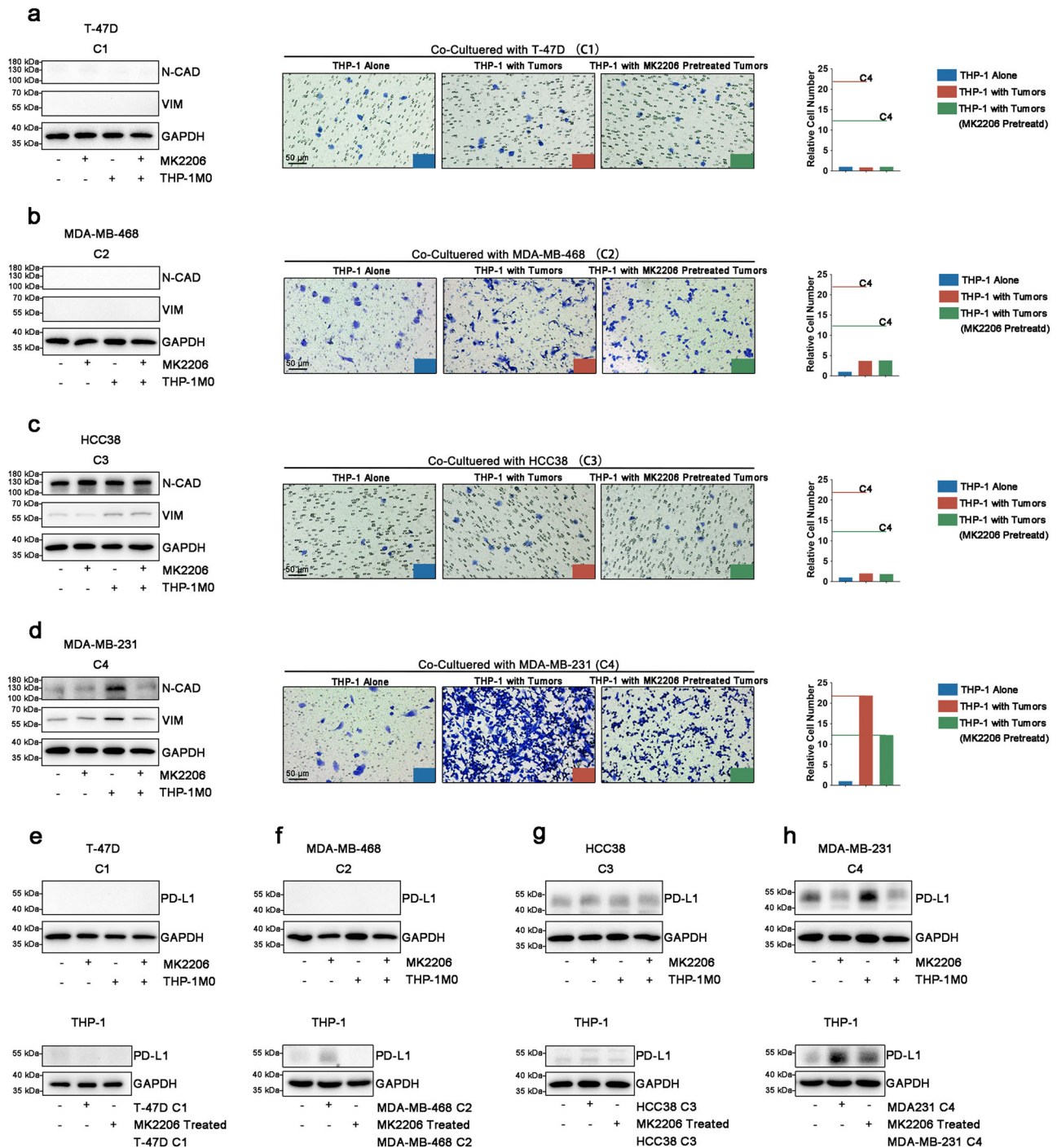
Next, we detected the role of different EMT subtypes on TAM infiltration alteration. Following co-culture with different subtypes of breast cancer cell lines, we detected the infiltration ability of macrophages. Although MDA-MB-468 (fold change 3.7) and HCC38 (fold change 2) also enhanced the infiltration of macrophages, MDA-MB-231 dramatically increased the infiltration (fold change 21.9). Besides, MK2206 could only abolish TAM promoting effect in the MDA-MB-231 cell line (Fig. 4a-d).

Furthermore, PD-L1, as major effector leading to the dysfunction of T cell, was analyzed on both tumor cells and macrophages. Using breast cancer cell line, we found THP-1 could only up-regulate expression of PD-L1 in MDA-MB-231 but not in other EMT subtype breast cancer cell through AKT pathway (Fig. 4e-h, upper panel). Additionally, MDA-MB-231 cells could up-regulate the PD-L1 expression of THP-1 more intensively than other EMT subtype breast cancer cell through AKT pathway (Fig. 4e-h, lower panel).

By using melanoma cell lines, similar tumor-TAM feedback loop was also observed in WM-115 cell line (EMT<sup>high</sup>-AKT subtype, Supplementary Fig. 14). These results suggested that the AKT pathway was the specific hub mediating immunosuppression in EMT<sup>high</sup>-AKT subtype by regulating the crosstalk between mesenchymal transition and TAM manipulation.

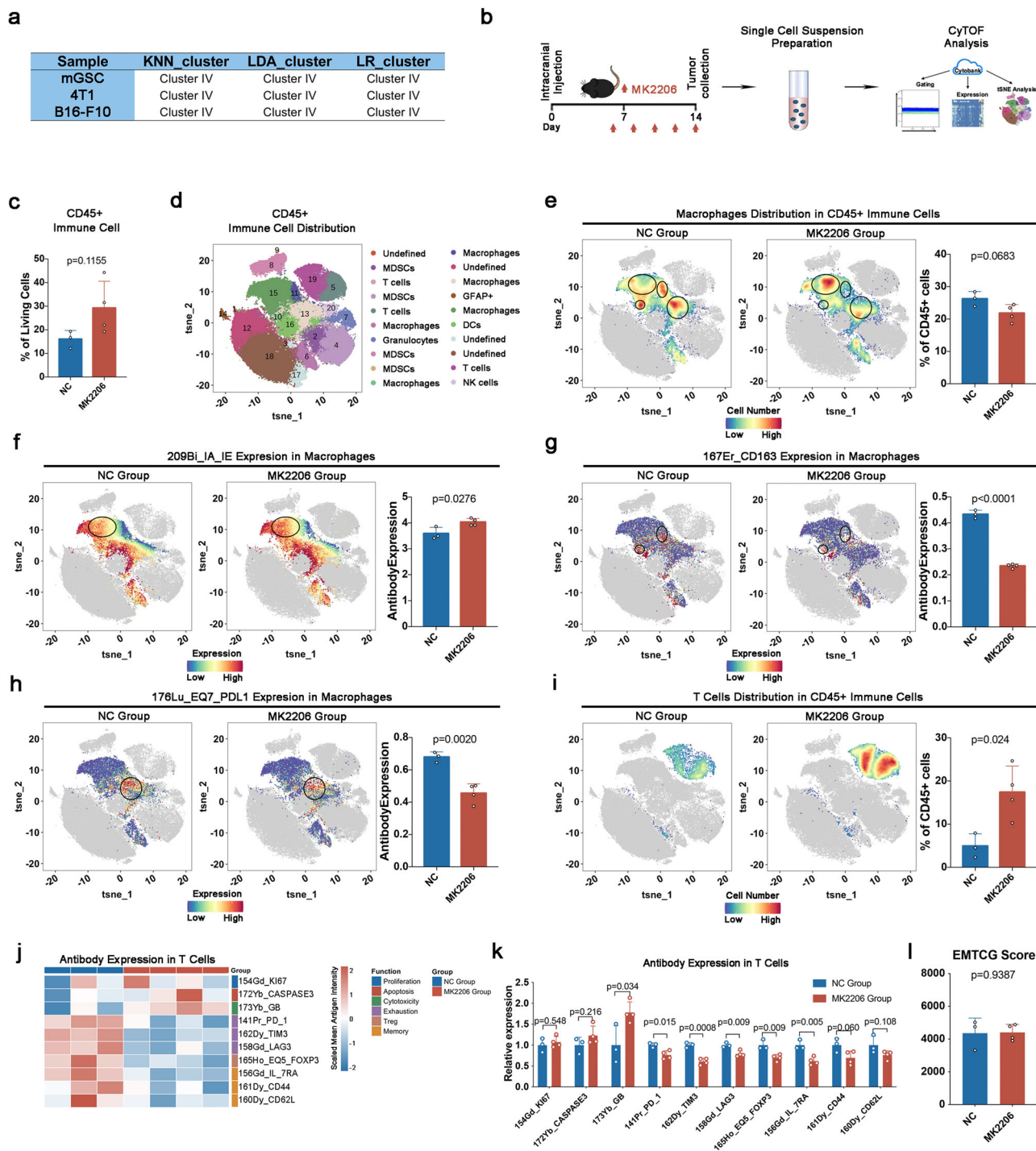
**AKT inhibition recover the immunosuppressive microenvironment of EMT<sup>high</sup>-AKT tumor in vivo.** Next, we tried to test whether AKT signaling was responsible for immunosuppression of EMT<sup>high</sup>-AKT subtype in vivo. To do this, we first conducted RNA-seq in three mouse cancer cell lines including mGSC (glioma), 4T1 (breast cancer) and B16-F10 (melanoma) to identify mouse EMT<sup>high</sup>-AKT subtype tumor. Using three machine learning methods, all these three cell lines were classified into EMT<sup>high</sup>-AKT subtype (Fig. 5a). Since mGSC was a spontaneous primary cancer cell line, which intimates more to primary human cancer than established cancer cell lines, mGSC was used for CyTOF analysis.

As indicated in Fig. 5b, an intracranial xenograft glioma model was established using mGSC cell line to harvest single-cell suspension for the CyTOF analysis with 37 antibodies (Supplementary Data 16). Firstly, we found that by inhibiting AKT signaling pathway, the immune cell infiltration tends to increase (Fig. 5c). Next, we gated into CD45+ immune cells and acquired 20 cell populations for further analysis (Fig. 5d). We observed a tendency of few macrophage infiltration in MK2206 group (Fig. 5e).



**Fig. 4 AKT pathway is important for mediating mesenchymal transition and immunosuppression only in EMT<sup>high</sup>-AKT subtype.** **a** Left panel: Immunoblotting analysis of N-Cadherin and Vimentin indicated that macrophage could not promote mesenchymal transition of T-47D. Right panel: Transwell analysis indicated T-47D could not enhance infiltration ability of macrophages. **b** Left panel: Immunoblotting analysis of N-Cadherin and Vimentin indicated that macrophage could not promote mesenchymal transition of MDA-MB-468. Right panel: Transwell analysis indicated MDA-MB-468 enhanced infiltration ability of macrophages mildly and could not be inhibited by AKT inhibition. **c** Left panel: Immunoblotting analysis of N-Cadherin and Vimentin indicated that macrophage could partially enhance mesenchymal transition of HCC38 (Vimentin) and could not be inhibited by AKT inhibition. Right panel: Transwell analysis indicated HCC38 enhanced infiltration ability of macrophages mildly and could not be inhibited by AKT inhibition. **d** Left panel: Immunoblotting analysis of N-Cadherin and Vimentin indicated that macrophage could enhance mesenchymal transition of MDA-MB-231 which could be inhibited by AKT inhibition. Right panel: Transwell analysis indicated MDA-MB-231 enhanced infiltration ability of macrophages dramatically and could be inhibited by AKT inhibition. **e** PD-L1 expression did not up-regulate in co-cultured T-47D and macrophages. **f** PD-L1 expression did not up-regulate in co-cultured MDA-MB-468 and macrophages. **g** PD-L1 expression did not up-regulate in co-cultured HCC38 and macrophages. **h** PD-L1 expression up-regulated in co-cultured MDA-MB-231 and macrophages which could be inhibited by AKT inhibition. For all transwell images in this figure, scale bar = 50  $\mu$ m.





**Fig. 5** AKT inhibition could reverse the immune-suppressive phenotype of the EMT<sup>high</sup>-AKT subtype in vivo. **a** EMT classification of mGSC, 4T1 and B16-F10 cell line using machine learning method. **b** Flowchart of CyTOF Analysis. **c** Comparison of CD45<sup>+</sup> immune cell's percentage between NC group and MK2206 group. **d** 20 clusters were acquired in CD45<sup>+</sup> immune cells. **e** TAM's infiltration has a tendency to decrease in MK2206 group compared to NC group. **f** IA<sub>IE</sub> expression is upregulated in TAM's of MK2206 group. **g** CD163 expression is downregulated in TAM's of MK2206 group. **h** PDL1 expression is downregulated in TAM's of MK2206 group. **i** T cell's infiltration increased in MK2206 group compared to NC group. **j** Expression heatmap of T cell's function markers. **k** T cell's dysfunction status was reversed by MK2206 without increase of memory T cells. **l** Mesenchymal transition status did not alter after MK2206 inhibition. Error bars indicate SD; the Student's *t* test was used to analyze statistical significance between 2 groups; In NC group, *n* = 3 biological replicates; In MK2206 group, *n* = 4 biological replicates.

Importantly, macrophage distribution pattern (illustrated by circles in Fig. 5e) was different between NC and MK2206 groups, indicating macrophage function might differ between these two groups. IA<sub>IE</sub> (M1), CD163 (M2), and PDL1 were used to evaluate macrophage functional phenotype. The expression of IA<sub>IE</sub> was

upregulated in the MK2206 group, while CD163 and PDL1 were downregulated in MK2206 group (Fig. 5f-h).

To further explore the functional status of TAMs, mRNA expression level of TAM's proinflammatory (M1) or anti-inflammatory (M2) markers were also detected using RT-qPCR

test. As shown in Supplementary Fig. 15, M1 markers including Ifng (Supplementary Fig. 15a), Tnfa (Supplementary Fig. 15b) and Cd86 (Supplementary Fig. 15c) didn't alter while M2 markers including Cd206 (Supplementary Fig. 15d), Arg1 (Supplementary Fig. 15e) and Trem2 (Supplementary Fig. 15f) downregulated in mGSC xenograft model after MK2206 administration. These results indicated that AKT pathway inhibition could block M2 transformation of macrophages.

Next, we investigated whether inhibiting AKT pathway was helpful for antitumor effects. Compared to the NC group, more T cells infiltrated into tumor tissue in MK2206 group (Fig. 5i). The cytotoxic ability (GB) of T cell was increased in the MK2206 group (Fig. 5j, k). T cell exhaustion markers (PD\_1, TIM3, and LAG3) were downregulated in the MK2206 group (Fig. 5j, k). These results indicated that T cell dysfunction was reversed by AKT inhibition. Meanwhile, the proportion of peripheral macrophages, microglia, dendritic cells, NK cells, and MDSCs were also investigated. More dendritic cells were infiltrated into the MK2206 group (Supplementary Fig. 16a–d).

To rule out the possibility that AKT inhibition induced cancer cell differentiating into a more epithelial state to recover immunosuppressive microenvironment, mGSC xenograft tissues were sequenced. The result shown MK2206 didn't alter the mesenchymal status of cancer cell (Fig. 5l). Taken together, these findings demonstrated that AKT signaling inhibition could reverse the immunosuppressive phenotype of the EMT<sup>high</sup>-AKT subtype *in vivo*.

### Immune checkpoint blockade plus TME remodeling by AKT inhibition could prolong survival in EMT<sup>high</sup>-AKT subtype.

Considering the immunosuppression in EMT<sup>high</sup>-AKT subtype, we tried to test whether ICB therapy would work for these patients. Firstly, we investigated whether single ICB would prolong survival in this subtype. Four ICB clinical trial cohorts were used to do this. In Imvigor210 dataset<sup>30</sup>, one-way ANOVA analysis revealed EMT<sup>high</sup>-AKT subtype exhibiting the lowest response fraction (17%,  $p = 0.0053$ , Fig. 6a). Besides, t test revealed that, ICB response rate in EMT<sup>high</sup>-AKT subtype was dependent on AKT pathway activity (Fig. 6b). In GSE78220<sup>31</sup>, although patients were not classified into all 4 EMT subtypes due to limited sample numbers, we found that response rate was also dependent on AKT pathway activity (Fig. 6c). Next, two anti-CTLA4 trials (phs000452<sup>32</sup> and SRP067586<sup>33</sup>) were combined together to test the response rate of anti-CTLA4 therapy. Consistently, EMT<sup>high</sup>-AKT subtype exhibited the lowest response fraction (26%,  $p = 0.0056$ , Fig. 6d). Besides these trials, we applied the TIDE algorithm<sup>25</sup> to evaluate the immunotherapeutic response and compared the response rate between the four EMT subtypes in TCGA dataset. Similarly, the EMT<sup>high</sup>-AKT subtype exhibited the lowest fraction of response (22%) dependent on AKT pathway activity (Fig. 6e, f). Then, we applied 4T1, B16-F10 and mGSC xenograft model to test the ICB therapy response. Single ICB administration could not reduce tumor burden in 4T1, B16-F10 and mGSC cell lines and could not prolong survival in B16-F10 and mGSC cell lines (Fig. 6g–o). For 4T1, no death due to tumor was recorded under the observation period. These findings suggested that EMT<sup>high</sup>-AKT subtype was resistance to single administration of immunotherapy.

Since the immunosuppressive microenvironment of EMT<sup>high</sup>-AKT subtype was shape by AKT pathway, we next investigated whether AKT inhibition could enhance the treatment value of ICB therapy in EMT<sup>high</sup>-AKT subtype cell lines. The drug administration strategy was illustrated in Supplementary Fig. 17. Combination of ICBs (including anti-PD-L1, anti-PD-1 and anti-CTLA-4) and MK2206 significantly inhibited tumor growth

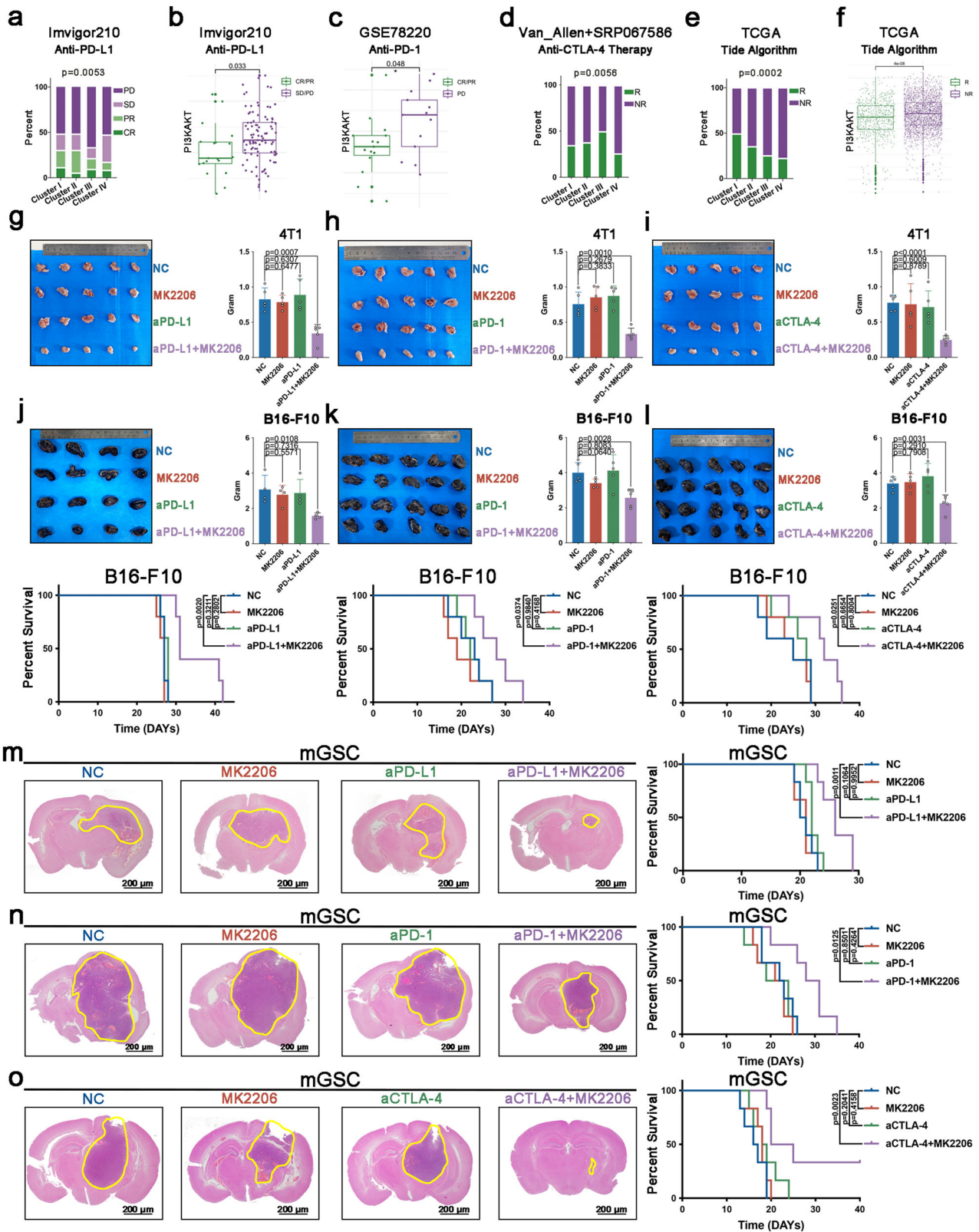
in 4T1, B16-F10 and mGSC cell lines and prolonged survival in B16-F10 and mGSC cell lines (Fig. 6g–o). For 4T1, no death due to tumor was recorded under the observation period. Besides, MK2206 did not cause apparent structural alteration or toxicity in liver or kidney (Supplementary Fig. 18). Therefore, AKT inhibition relieved the immunosuppression and therapeutic resistance to ICB treatment in EMT<sup>high</sup>-AKT subtype.

**TumorMT: A comprehensive resource for exploring tumor mesenchymal transition status.** Our results highlighted the necessity for precisely identifying mesenchymal transition status among patient samples (Supplementary Fig. 19). Therefore, we developed the TumorMT website (<http://tumormt.neuroscience.org.cn>) (Supplementary Fig. 20a–e). The in-house glioma dataset, Chinese Glioma Genome Atlas (CGGA), was used as a paradigm and validation of the EMT subtype assignment accuracy and practicability of the website. The website classified the 388 Glioblastoma Multiforme (GBM) samples of CGGA into four EMT clusters (Supplementary Fig. 21a–c). Consistent with the TCGA cohort, EMT<sup>high</sup>-AKT subtype GBM samples in the CGGA cohort exhibited the lowest tumor purity level and the highest level of non-tumor compartment infiltration (Supplementary Fig. 21d–f). Moreover, anti-tumor immunity was impaired in this subtype owing to high M2-TAM infiltration, which led to T cell dysfunction (Supplementary Fig. 21g, h). These results demonstrated the accuracy of our findings and the practicability of the TumorMT website.

### Discussion

For now, several studies have established reasonable clustering system for pan-cancer patients in the field of epithelial mesenchymal transition<sup>34,35</sup>. However, how to discern EMT driving force from intrinsic and extrinsic perspectives remains challenging and raised questions regarding the true role of EMT in tumor biology<sup>36</sup>. In this study, we generated an EMTTCG signature that robustly classified tumors into four EMT subtypes. Especially, two EMT<sup>high</sup> subtypes were found to be driven by either genomic alteration or TAM crosstalk which spurred us to explore the heterogeneity in high EMT status separately. The unresectable, drug resistant EMT<sup>high</sup>-AKT subtype was immunosuppressive and could be treated by combination of blocking ATK pathway and immune checkpoint blockade.

Routinely, two approaches have been applied to evaluate tumor EMT status<sup>5,6,8,10,34,35</sup>. The first approach incorporates experimental methods to detect the expression of individual mesenchymal markers<sup>8,10</sup>. The second approach utilizes score-based bioinformatic approaches to detect the overall enrichment of EMT genes<sup>34,35</sup>. However, these methods have intrinsic limitations. On one hand, as EMT is a multi-mechanism hybrid process, using a single experimental marker might introduce selection bias into the evaluation. On the other hand, score-based bioinformatic approaches evaluate EMT activity by comparing the score within the same cohort and might lose efficacy when applied toward independent clinical patient. Our work solved these problems to some degree. Firstly, our EMTTCG signature was constructed by using co-expression analysis and protein-protein-interaction analysis to include all key genes in EMT process and thus avoid marker selection bias. Secondly, by using unsupervised consensus cluster, the characteristic cluster center of the four EMT subtypes could be acquired. By calculating the similarity distances to the cluster centers of the four EMT subtypes, individual clinical patient could also get categorical information regarding to our classification. Moreover, we developed TumorMT website, which may provide precise classification and treatment information based on our findings.



Although a lot of tumors have been featured by enhanced mesenchymal phenotypes, we still have few knowledge about the heterogeneity among these patients<sup>36</sup>. Especially, although both genetic alteration and TME could promote mesenchymal transition, existing methods are not be able to identify dominant driving mechanism in individual mesenchymal-like samples which hinders in choosing the most proper EMT targeting

treatment strategy<sup>36</sup>. Our work established two mesenchymal like subtypes (EMT<sup>high</sup>-NOS and EMT<sup>high</sup>-AKT), which were found to have distinct biological and clinical features. Compared to EMT<sup>high</sup>-NOS subtype, EMT<sup>high</sup>-AKT is drug resistant, and unresectable. Importantly, our findings suggested that the EMT<sup>high</sup>-NOS subtype was driven by intrinsic somatic alterations which could be treated by cell cycle targeting strategies, whereas

**Fig. 6 ICB plus AKT inhibition could prolong survival in EMT<sup>high</sup>-AKT subtype.** **a** Comparison of anti-PD-L1 therapy response rate among the four EMT clusters in IMvigor210 cohort; Complete response (CR), Partial response (PR), Stable disease (SD), Progressive disease (PD). **b** Comparison of AKT pathway activity in EMT<sup>high</sup>-AKT subtype patients with different immunotherapy responses in IMvigor210 cohort; CR/PR group,  $n = 19$ ; SD/PD group,  $n = 95$ . **c** Comparison of AKT pathway activity with different immunotherapy responses using anti-PD-1 in GSE78220 cohort; CR/PR group,  $n = 15$ ; PD group,  $n = 13$ . **d** Comparison of anti-CTLA-4 therapy response rate among the four EMT clusters in Van\_allen+SRP067586 cohort; Response (R), Nonresponse (NR). **e** Comparison of ICB therapy response rate among the four EMT clusters in TCGA cohort. Response (R) and non-response (NR). **f** Comparison of AKT pathway activity in EMT<sup>high</sup>-AKT subtype patients with different immunotherapy responses in TCGA cohort; R group,  $n = 646$ ; NR group,  $n = 2278$ . **g** MK2206 could overcome anti-PD-L1 resistance in 4T1 xenograft tumors;  $n = 5$  biological replicates. **h** MK2206 could overcome anti-PD-1 resistance in 4T1 xenograft tumors;  $n = 5$  biological replicates. **i** MK2206 could overcome anti-CTLA-4 resistance in 4T1 xenograft tumors;  $n = 5$  biological replicates. **j** MK2206 could overcome anti-PD-L1 resistance in B16-F10 xenograft tumors;  $n = 4$  biological replicates for tumor burden assay;  $n = 5$  biological replicates for survival assay. **k** MK2206 could overcome anti-PD-1 resistance in B16-F10 xenograft tumors;  $n = 5$  biological replicates. **l** MK2206 could overcome anti-CTLA-4 resistance in B16-F10 xenograft tumors;  $n = 5$  biological replicates. **m** MK2206 could overcome anti-PD-L1 resistance in mGSC intracranial xenograft tumors;  $n = 6$  biological replicates. **n** MK2206 could overcome anti-PD-1 resistance in mGSC intracranial xenograft tumors;  $n = 6$  biological replicates. **o** MK2206 could overcome anti-CTLA-4 resistance in mGSC intracranial xenograft tumors;  $n = 6$  biological replicates. Error bars indicate SD; the Student's *t* test or Kaplan-Meier analysis was used to analyze statistical significance between 2 groups; For **(m-o)**, scale bar = 200  $\mu\text{m}$ .

the EMT<sup>high</sup>-AKT subtype was driven by the crosstalk between tumor and TME which could be treated by ICB combined with TME remodeling. In the EMT<sup>high</sup>-NOS subtype, somatic alterations occurred in cell cycle-related factors such as *TP53* and *Cyclin D1*. Their abnormality might cause tumor cells to undergo devastating cellular mitosis. In the EMT<sup>high</sup>-AKT subtype, M2-TAMs comprised the representative compartments in TME and promoted mesenchymal transition. In EMT<sup>high</sup>-AKT subtype, microenvironment factors, especially TAMs, weighs more in shaping mesenchymal transition than genetic variation factors. Therefore, intrinsic and extrinsic mesenchymal transition enhancing mechanisms might lead to distinct biological and clinical outcome, which need to be managed differently.

Increasing evidences have suggested that there are complex interactions between tumor mesenchymal transition and immunosuppressive phenotype. Here, we observed a specific AKT pathway depending tumor cell and TAM interaction facilitating both mesenchymal transition and immunosuppression in EMT<sup>high</sup>-AKT subtype tumors. Marks et al.<sup>37</sup> also demonstrated AKT inhibition was related with favorable immune profile, which was consistent with our CyTOF detection results. However, as was in Fig. 6g–o, AKT inhibition could not prolong overall survival of tumor patients. Our CyTOF analysis answered why AKT inhibition could not prolong overall survival to some extent. Although AKT inhibition reverted immunosuppressive status of TAMs and increased cytotoxic ability of T cells, AKT inhibition also tended to decrease the ration of memory T cells. On the other hand, ICBs have been demonstrated to increase the number ration of memory T cells<sup>38</sup>. However, ICBs didn't work in EMT<sup>high</sup>-AKT subtype, since the substantial tumor-TAM interaction significantly upregulated the expression of immune checkpoints in EMT<sup>high</sup>-AKT subtype (Fig. 4e–h). Previous studies have demonstrated that ICBs only work to the status that overexpression of immune checkpoints were due to T cell cytotoxicity<sup>39,40</sup>. Blocking non-CD8 + T cell origin-induced PD-L1 expression could help to enhance response rate of ICB therapy<sup>41</sup>. As shown in Fig. 4e–h, AKT inhibition could decrease non-CD8 + T cell origin-induced PD-L1 expression which indicated to increase ICB response rates. Here, by using three cell models and three ICB regimens, we demonstrated the universe feasibility of ICB treatment plus AKT inhibition. In this strategy, AKT inhibition was used to block PD-L1 overexpression and to revert immune suppressive microenvironment while ICBs were used to increase the ration of memory T cells.

Our findings highlighted the distinct molecular, biological and clinical features among patients with enhanced mesenchymal transition, indicating that different treatment strategies should be applied.

## Methods

**Data acquisition.** TCGA multi-omics data (clinical information '2018-09-13 version', CNV profile '2016-08-16 version', gene mutation profile '2016-12-29 version', DNA methylation profile 'Methylation450K 2016-12-29 version', transcriptomic expression profile '2016-12-29 version, RNA-seq', and reverse-phase protein array profile '2016-08-16 version') were downloaded from the Xena Website (<https://xenabrowser.net/>). The cancer cell line transcriptomic expression profile '20180929 version, RNA seq' was downloaded from Cancer Cell Line Encyclopedia (CCLE) Website (<https://portals.broadinstitute.org/ccle/>). Colorectal cancer datasets (KFSYSCC cohort, FRENCH cohort, GSE2109 cohort, GSE37892 cohort, GSE35896 cohort, GSE23878 cohort, GSE20916 cohort, GSE17536 cohort, and GSE13067 cohort; Transcriptome) were downloaded as the authors indicated<sup>13</sup>. Breast cancer dataset (FUSCCTNBC cohort; Transcriptome) was downloaded as the authors indicated<sup>14</sup>. Gastric cancer datasets (ACRG cohort, KUCM cohort, KUGH cohort, MDACC cohort and SMC cohort; Transcriptome) were downloaded as the author indicated<sup>15</sup>. IMvigor210CoreBiologies data (RNA seq) were downloaded using the R package (version 1.0.0) provided by the following website (<http://research-pub.gene.com/IMvigor210CoreBiologies/packageVersions/>). GSE78220 (RNA-seq) data was downloaded from GEO website; Anti-CTLA4 clinical trial data (RNA seq) was acquired from dbGaP: phs000452 ([https://github.com/vanallenlab/VanAllen\\_CTLA4\\_Science\\_RNASeq\\_TPM/](https://github.com/vanallenlab/VanAllen_CTLA4_Science_RNASeq_TPM/) commit/3d1793629716cc1fd8e7334ea3bf593a20e6fe07)<sup>32</sup> and SRA: SRP067586 cohort<sup>33</sup>. We used pre-processed data provided by these authors. The details of the sample information included in this work are summarized in Supplementary Data 17.

**CGGA patient samples.** A total of 388 GBM samples were included in our study<sup>42,43</sup>. These samples were acquired by CGGA. Each sample was diagnosed and independently confirmed by two neuropathologists based on the 2007 WHO classification guidelines<sup>44</sup>. Only samples consisting of > 80% tumor cells were selected for transcriptomic profiling (RNA-seq). The details of the sample information included in this work are summarized in Supplementary Data 17. The CGGA cohort data could be obtained from the CGGA database (<http://www.cgga.org.cn>).

**Wu lab glioma samples.** Glioma tissue samples from patients were acquired from our institution with the assistance of neuronavigation between October 10, 2020 and August 18, 2022. Samples were sent for RNA-seq and classified into four EMT subtypes according to transcriptomic profile. 19 samples with sufficient tissue block were sent for immunohistochemistry

staining. The sample information of Wu lab could be found in Supplementary Data 17.

**Antibodies and reagents.** Detailed information on the antibodies and reagents used in this study is listed in Supplementary Data 18.

**PPI analysis.** PPI was performed using String Website (<https://string-db.org/>). Gene ID was used as input data. The interaction network using combined score was then imported into Cytoscape 'version 3.6.1' (<https://cytoscape.org>) for network analysis. The genes that had a degree (K) within the top two-thirds of all genes were considered node genes.

**Co-expression analysis.** Co-expression analysis was performed using the Coexpedia Website (<http://www.coexpedia.org/>). Gene ID was used as input data. The genes with a score within the top two-thirds of all genes were considered co-expression genes.

**Mutation analysis.** The baseline mutation frequency in each cancer type was calculated first (Baseline mutation frequency is defined as the mutation rate of the whole genome in a certain cancer type by calculating the ratio between positive mutation event number and total observed event number). We then calculated the ratio between mutation frequency of EMTCGs with the baseline mutation frequency as fold change in each cancer type. A mutation frequency >3% and fold change >1.5 was considered significant.

**DNA methylation analysis.** R software 'version 3.5.1' was used to analyze the Pearson correlation between the expression of EMTCGs and DNA methylation level ( $\beta$ -value). Three criteria were used to identify the abnormal expression of EMTCGs affected by DNA methylation for each cancer type: [1] the probe with the largest average  $\beta$  value was taken into account; [2] Pearson correlation between expression and DNA methylation probe < -0.3; [3] FDR < 0.05 was considered statistically significant. FDR adjust was done using Benjamin and Hochberg's method<sup>45</sup>.

**CNV analysis.** CNV values were acquired by the Xena Website using TCGA FIREHOSE pipeline based on GISTIC2. R software 'version 3.5.1' was used to analyze the Pearson correlation between the expression of EMTCGs and CNV. Three criteria were used to identify the abnormal expression of EMTCGs affected by CNV for each cancer type: [1] More than 30% of the tumor samples had the amplification or deletion of loci; [2] Pearson correlation between expression and CNV > 0.3; [3] FDR < 0.05 was considered statistically significant. FDR adjust was done using Benjamin and Hochberg's method<sup>45</sup>.

**Mesenchymal transition activity evaluation.** Mesenchymal transition activity of individual samples was evaluated using the ssGSEA method supported by GSVA R package 'version 1.28.0'<sup>46</sup>. The accuracy of ssGSEA to predict pathway activity has already been validated by other work<sup>47</sup>. Transcriptomic expression profile was used as input "expr" for the analysis and EMTCG signature was used as input "gset.idx.list" for the analysis. Following parameters are used: method=ssgsea, kcdf=Gaussian, abs.ranking=FALSE.

**Survival analysis.** Univariate Cox regression analysis was chosen to assess the hazard ratio of mesenchymal transition. 'coxph' function in 'survival' R package 'version 3.1-11' was used to do

univariate cox regression analysis. Survival time is used as time value, survival status is used as status and ssGSEA score of EMTCG signature is used as covariate<sup>47</sup>. HR > 1 was considered a harm factor, while HR < 1 was considered a protective factor.  $p < 0.05$  was considered statistically significant.

**Machine learning based cluster analysis.** The ConsensusClusterPlus R package 'version 1.46.0' was used for K-Means cluster analysis of the TCGA cohort<sup>48</sup>. The expression profile of 35 EMTCGs was first normalized using the scale function of R software and then used as the input for cluster analysis. Machine learning parameters were as follows: Permutations = 100, Distance = Euclidean, InnerLinkage = Average, FinalLinkage = Average, and ClusterAlg = km.

**Sample number normalization between different clusters.** The number of samples for each categorical variable (histological type, sample portion, WHO grade, age and gender) were normalized into 100<sup>16</sup>. We then grouped samples into each EMT subtype.

**Sampling analysis.** To make sure our analysis didn't skew by sample number variation, we sampled 10 patients of each cancer in each cluster and merged them together to eliminate the bias due to cancer type specificity. Sampling would be repeated for 100 times to remove false positive results.

**KEGG analysis.** KEGG analysis was performed using the ClueGo plug-in unit 'version 2.5.3' provided by Cytoscape. Medium was chosen as Network Specificity. Two-sided hypergeometric test was chosen as statistical method. Bonferroni step down was chosen for p value correction and an FDR < 0.05 was considered statistically significant. FDR was done by Cytoscape automatically.

**Removal of batch effect.** The ComBat function of the SVA R package 'version 3.30.1' was used to remove the batch effect between different data cohorts<sup>49</sup>. Cohort variable was considered as batch variable. Covariable (mod parameter) was not considered in this analysis. Par.prior parameter was set as true. PCA analysis was used to test the result of removal of the batch effect.

**Identification of EMT Subtypes of Non-TCGA Samples.** Identification of non-TCGA samples was done with following steps. For human samples: (1) The batch effect between external cohorts and TCGA cohort was removed; (2) Each individual sample in external cohort was scaled with TCGA cohort separately; (3) Euclidian distance between individual samples and four TCGA EMT cluster centers was calculated using the expression of 35 EMTCG genes. Euclidian distance =  $\sqrt{\varepsilon(\text{nonTCGA}_i - \text{TCGA}_i)^2}$ ,  $i = 35$ ,  $\varepsilon$  means summation; (4) Individual samples were classified into the corresponding subtype with the shortest distance.

For mouse samples: (1) The batch effect between external cohorts and TCGA cohort was removed; (2) three machine learning method, K-NearestNeighbor (KNN), Linear Discriminant Analysis (LDA) and Logistic Regression (LR) were used to identify samples into relative clusters. The sequencing data of mGSC (mouse spontaneous glioma cell) and 4T1 were acquired using xenograft tissue from our group and the sequencing data of B16-F10 was acquired from GSE174724.

**Identification of driver mutations.** The driver mutation list was downloaded from Bailey et al.<sup>18</sup>. The Maftools R package 'version 1.8.10' was used to determine driver mutations in each EMT subtype. Mutations with adjPval < 0.05 and fold change > 1.5 were

considered driver mutations. adjPval was done by Maftools R package automatically.

**Identification of driver copy number variations.** The driver CNV list was downloaded from Hoadley et al.<sup>17</sup>. The chi-square test was used to determine driver mutations in each EMT subtype. CNV with FDR < 0.05 and fold change > 1.5 were considered as driver CNV. FDR adjust was done using Benjamin and Hochberg's method<sup>45</sup>.

**Tumor purity analysis.** Stromal and immune scores were calculated using the ESTIMATE R package 'version 1.0.13' (<https://sourceforge.net/projects/estimateproject/>), and tumor purity was estimated according to this formula:  $\text{purity} = \text{COS} (0.6049872018 + 0.0001467884 * \text{estimatescore})$ <sup>28</sup>.

**IFNG activity and TNFA activity calculation.** IFNG activity was acquired by calculating the geometric mean of marker genes defined by Jiang et al.'s work<sup>25</sup>. TNFA activity was acquired using the ssGSEA method using marker genes defined by Wei et al.'s work<sup>26</sup>.

**GSEA analysis.** Gene Set Enrichment Analysis (GSEA, <http://www.broadinstitute.org/gsea/index.jsp>, version 4.3.2) was performed to explore whether the identified sets of genes showed statistical differences between the two groups using default parameters<sup>50</sup>. Expression profile was used as expression dataset set. EMT<sup>high</sup>-AKT group or non-EMT<sup>high</sup>-AKT group were used as phenotype label. Gene sets were acquired from previous published work<sup>51</sup>. Number of permutations = 100. Collapse dataset to gene symbols = FALSE. Permutation Type = phenotype. NES and FDR were used to determine statistical significance.

**Immune cell infiltration fraction analysis.** Immune cell infiltration fraction analysis was performed using the ssGSEA method based on the LM-22 immune cell marker signature<sup>27</sup>. Transcriptomic expression profiles were used as the input for analysis.

**TIP analysis.** Tracking Tumor Immunophenotype (TIP) analysis was performed using the website tool (<http://biocc.hrbmu.edu.cn/TIP/index.jsp>)<sup>52</sup>. Transcriptomic expression profiles were used as the input for analysis.

**Immunohistochemistry.** For immunohistochemistry, tissue sections were fixed in paraformaldehyde (4%), embedded in paraffin, and cut into 4 μm sections. Streptavidin-biotin immunostaining method was employed<sup>28</sup>.

**Construction of mouse spontaneous glioma cell using sleeping beauty transposon.** Construction of spontaneous glioma cells (mGSC) was done using sleeping beauty transposon<sup>53</sup>. In brief, PEI/DNA complexes were prepared according to nucleic acid delivery protocol: vivo-jetPEI = 0.08 μl, total of DNA = 0.5 μg (PT2/C-Luc//PGK-SB13 = 0.1 μg, pT/Nestin-SV40-LgT = 0.2 μg, pT/Nestin-NRASV12 = 0.2 μg). Two microliters of PEI/DNA complexes (0.5 μg/μl) were administered to neonatal mouse brain. One month later, mGSC was cultured in DMEM/F12 (Gibco), supplemented with 2% B27 supplement (Gibco), 20 ng/mL EGF (Peprotech, Rocky Hill, NJ, USA), 20 ng/mL basic-FGF (Peprotech), and 1% penicillin/streptomycin (Gibco) at 37 °C with 5% CO<sub>2</sub>.

**Cell lines and cell culture.** The human breast cancer cell lines MDA-MB-231 and MDA-MB-468 were purchased from the

Chinese Academy of Sciences cell bank (Shanghai, China). The human breast cancer cell lines T-47D, HCC38, human melanoma cell lines A-375 and WM-115 were purchased from iCell (Shanghai, China). Human melanoma cell line MeWo was purchased from FENGHUIHENGWU (Hunan, China). Human melanoma cell line SK-MEL-3 was purchased from COBIER (Jiangsu, China). The human monocyte cell line THP-1 was purchased from the Chinese Academy of Sciences cell bank (Shanghai, China). The mouse breast cancer cell line 4T1 and mouse melanoma cell line B16-F10 were purchased from Procell Life Science&Technology Co.,Ltd (Wuhan, China). The human breast cancer cell lines T-47D, MDA-MB-468 and human melanoma cell line A-375 were maintained in Dulbecco's Modified Eagle's Medium (DMEM, Gibco, USA) containing 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin at 37 °C with 5% CO<sub>2</sub>. The human breast cancer cell line MDA-MB-231 was maintained in Leibovitz's L-15 medium containing 10% FBS and 1% penicillin/streptomycin at 37 °C without CO<sub>2</sub>. The human breast cancer cell line HCC38 was cultured in RPMI-1640 medium (Gibco, USA) containing 10% FBS and 1% penicillin/streptomycin (Gibco) at 37 °C with 5% CO<sub>2</sub>. The human monocyte cell line THP-1 was cultured in RPMI-1640 medium (Gibco, USA) containing 10% FBS and 1% penicillin/streptomycin (Gibco) at 37 °C with 5% CO<sub>2</sub>. Human melanoma cell line MeWo and WM-115 were cultured in MEM medium (Gibco, USA) containing 10% FBS and 1% penicillin/streptomycin (Gibco) at 37 °C with 5% CO<sub>2</sub>. Human melanoma cell line SK-MEL-3 was cultured in McCoy's 5a medium (Gibco, USA) containing 15% FBS and 1% penicillin/streptomycin (Gibco) at 37 °C with 5% CO<sub>2</sub>. The mouse breast cancer cell line 4T1 was cultured in RPMI-1640 medium (Gibco, USA) containing 10% FBS and 1% penicillin/streptomycin (Gibco) at 37 °C with 5% CO<sub>2</sub>. Mouse melanoma cell line B16-F10 was cultured in RPMI-1640 medium (Gibco, USA) containing 10% FBS and 1% penicillin/streptomycin (Gibco) at 37 °C with 5% CO<sub>2</sub>.

**Cell treatment.** For AKT inhibition, the cells were pre-incubated with 10 μM MK2206 (MCE, USA) for 48 h before cell co-culture were performed.

**Cell migration assay.** 8 μm transwell chamber was used for cell migration in vitro. After co-culture with tumor cells, THP-1 cells were re-suspended as single cells in RPMI-1640 medium containing 0.2% FBS and seeded into the upper chambers at a density of  $4 \times 10^4$  cells/200 μL. 800 μL RPMI-1640 medium containing 20% FBS was added into the lower chamber. After 24 h of incubation, cells on the upper side of the membrane were removed using cotton swab. Cells that migrated to the lower side of the membrane were stained with 1% crystal violet solution after methanol fixation.

**Protein extraction and western blotting.** Total proteins were extracted using whole cell lysis buffer (Beyotime Biotechnology, Beijing, China) and quantified using bicinchoninic acid (BCA) method<sup>54</sup>. 20 μg protein was loaded, electrophoresed using 10% SDS-PAGE and transferred to a polyvinylidene difluoride (PVDF) membrane (0.45 μm; Millipore, Burlington, MA, USA). After blocking with skimmed milk (5%), the PVDF membranes were incubated with the primary antibodies overnight at 4 °C. Then, the PVDF membranes were incubated with the secondary antibodies at 25 °C for 1 h. Chemiluminescence ECL reagent (Tanon, Woburn, MA, USA) was used for protein visualization.

**Intracranial xenograft transplantation.** Male C57BL/6 N mice (6–8 weeks of age) were purchased from Beijing Vital River

Laboratory Animal Technology. Mice were raised in laminar flow cabinets under specific pathogen-free (SPF) conditions. For orthotopic transplantations, 3  $\mu$ L glioma cell suspension (mGSC sphere:  $1 \times 10^4$ ) was injected into the mouse brains at a depth of 3.0 mm using stereotactic devices.

**Subcutaneous xenograft transplantation.** Male C57BL/6 N mice (6–8 weeks of age) were purchased from Beijing Vital River Laboratory Animal Technology. Approximately  $1 \times 10^6$  B16-F10 cells were transplanted into the right flank of the animals in inoculation volumes of 200  $\mu$ L sterile media. Female BALB/C mice (6–8 weeks of age) were purchased from Beijing Vital River Laboratory Animal Technology. Approximately  $2 \times 10^6$  4T1 cells were transplanted into the right flank of the animals in inoculation volumes of 200  $\mu$ L sterile media. The tumor mass was weighted after tumor harvest.

**In vivo inhibitor experiments.** Three days after tumor xenograft transplantation, the mice were randomly divided into control, anti-ICBs, MK2206, and anti-ICBs+MK2206 groups. MK2206 was orally administered on days 6, 8, 10, 12, 14, 16, 18, and 20 at a dose of 150  $\mu$ g/g body weight. DMSO was orally administered as a control for MK2206. Anti-ICBs were intraperitoneally injected on days 7, 10, and 13 at a dose of 10  $\mu$ g/g body weight. ICBs including anti-PD-L1, anti-PD-1 or anti-CTLA-4 which were bought from BioXcell, USA.

**RNA isolation and reverse-transcription quantitative PCR (RT-qPCR).** Trizol (Invitrogen, USA) was used to extract total RNA according to manufacture's protocol. Total RNA was reversely transcribed into cDNA using Prime-Script RT Master Mix Kit (TaKaRa, Japan). RT-qPCR was performed with TB Green Premix Ex Taq II (TaKaRa, Japan). The PCR primer sequences were as follows:

Ifng Forward Primer: CAGCAACAGCAAGGCGAAAAAGG  
 Ifng Reverse Primer: TTTCCGCTTCCTGAGGCTGGAT  
 Tnfa Forward Primer: GGTGCCTATGTCTCAGCCTCTT  
 Tnfa Reverse Primer: GCCATGAAGTGTGATGAGAGGGAG  
 Cd86 Forward Primer: ACGTATTGGAAGGAGATTACAGCT  
 Cd86 Reverse Primer: TCTGTACGCGTTACTATCCCGC  
 Cd206 Forward Primer: GTTCACCTGGAGTGATGGTTCTC  
 Cd206 Reverse Primer: AGGACATGCCAGGGTACCTTT  
 Arg1 Forward Primer: CATTGGCTTGGAGACGTAGAC  
 Arg1 Reverse Primer: GCTGAAGGTCTCTTCCATCAC  
 Trem2 Forward Primer: CTACAGTGTGAGTCTCCGA  
 Trem2 Reverse Primer: CCTCGAAACTCGATGACTCCTC  
 Gapdh Forward Primer: CATCACTGCCACCCAGAAGACTG  
 Gapdh Reverse Primer: ATGCCAGTGAGCTTCCCGTTCAG

**CyTOF analysis of immune cells.** CyTOF analysis was performed by Novogene Co., Ltd. (Beijing, China). In brief, tumor tissue was digested by DNAase and Tyrisin to make a single-cell suspension. Cells were enriched using Percoll density gradient media, and red blood cells were removed by Red Blood Cell Lysis Buffer. The types of immune cells were identified by tSNE, followed by KNN clustering. After preliminary analysis, one sample in NC group was obviously distinct from others and was removed in further analysis. For the panels in Fig. 4, the t-SNE coordinates were determined before the sample in NC group was removed. For graphing, one sample in MK2206 group was randomly removed to keep balance.

**Construction of the tumor MT resource.** The Tumor MT website was developed in JSP using a Laravel framework and was deployed on a Nginx Web server that ran under Ubuntu system.

All data in Tumor MT was stored and managed using MySQL. jQuery was used to manage the result views. Tumor MT was thoroughly tested using Google Chrome.

**Statistics and reproducibility.** Statistical analyses were conducted using Prism 7 and R 3.5.1 unless otherwise stated. Definite statistical methods of various statistical tests are described and referenced in their respective sections. Otherwise, a two-tailed *t* test or one-way ANOVA was used, and *p* < 0.05 was considered statistically significant. FDR adjust was done using Benjamin and Hochberg's method<sup>45</sup>.

**Reporting summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

### Data availability

RNA-seq data of mice glioma support the findings in this manuscript are available in BioProject database at <https://ngdc.cncb.ac.cn/bioproject/>, with BIGD ID: PRJCA021222. RNA-seq data of human glioma support the findings in this manuscript are available in BioProject database at <https://ngdc.cncb.ac.cn/bioproject/>, with BIGD ID: PRJCA021271. Numerical source data for graphs and charts in this manuscript could be found in Supplementary Data 19. Uncropped and unedited blot images could be found in Supplementary Fig. 22. All other data are available from the corresponding author upon reasonable request.

Received: 19 November 2022; Accepted: 4 December 2023;

Published online: 18 December 2023

### References

- Vasan, N., Baselga, J. & Hyman, D. M. A view on drug resistance in cancer. *Nature* **575**, 299–309 (2019).
- Almenawer, S. A. et al. Biopsy versus partial versus gross total resection in older patients with high-grade glioma: a systematic review and meta-analysis. *Neuro Oncol.* **17**, 868–881 (2015).
- Thiery, J. Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* **2**, 442–454 (2002).
- Thiery, J., Acloque, H., Huang, R. & Nieto, M. Epithelial-mesenchymal transitions in development and disease. *Cell* **139**, 871–890 (2009).
- Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).
- Mani, S. et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).
- Wang, H. et al. BRCA1/FANCD2/BRG1-Driven DNA repair stabilizes the differentiation state of human mammary epithelial cells. *Mol. Cell* **63**, 277–292 (2016).
- Cheng, P. et al. FOXD1-ALDH1A3 signaling is a determinant for the self-renewal and tumorigenicity of mesenchymal glioma stem cells. *Cancer Res.* **76**, 7219–7230 (2016).
- Su, S. et al. A positive feedback loop between mesenchymal-like cancer cells and macrophages is essential to breast cancer metastasis. *Cancer Cell* **25**, 605–620 (2014).
- Wu, J. et al. Chemerin enhances mesenchymal features of glioblastoma by establishing autocrine and paracrine networks in a CMKLR1-dependent manner. *Oncogene* **41**, 3024–3036 (2022).
- Vasaikar, S. V. et al. EMTope: a resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. *Br. J. Cancer* **124**, 259–269 (2021).
- Wang, Q. et al. Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. *Cancer Cell* **32**, 42–56.e46 (2017).
- Guinney, J. et al. The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015).
- Jiang, Y. Z. et al. Genomic and transcriptomic landscape of triple-negative breast cancers: subtypes and treatment strategies. *Cancer Cell* **35**, 428–440.e425 (2019).
- Oh, S. C. et al. Clinical and genomic landscape of gastric cancer with a mesenchymal phenotype. *Nat. Commun.* **9**, 1777 (2018).
- Berger, A. C. et al. A comprehensive pan-cancer molecular study of gynecologic and breast cancers. *Cancer Cell* **33**, 690–705.e699 (2018).

17. Hoadley, K. A. et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* **158**, 929–944 (2014).
18. Bailey, M. H. et al. Comprehensive characterization of cancer driver genes and mutations. *Cell* **173**, 371–385.e318 (2018).
19. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).
20. Dongre, A. & Weinberg, R. A. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* **20**, 69–84 (2019).
21. Chen, X. et al. Molecular profiling identifies distinct subtypes across TP53 mutant tumors. *JCI Insight* **7**, <https://doi.org/10.1172/jci.insight.156485> (2022).
22. Shi, Z. Z. et al. Consistent and differential genetic aberrations between esophageal dysplasia and squamous cell carcinoma detected by array comparative genomic hybridization. *Clin. Cancer Res.* **19**, 5867–5878 (2013).
23. Yang, W. et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **41**, D955–961 (2013).
24. Jiang, Y. & Zhan, H. Communication between EMT and PD-L1 signaling: new insights into tumor immune evasion. *Cancer Lett.* **468**, 72–81 (2020).
25. Jiang, P. et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **24**, 1550–1558 (2018).
26. Wei, Y. et al. The local immune landscape determines tumor PD-L1 heterogeneity and sensitivity to therapy. *J. Clin. Investig.* **129**, 3347–3360 (2019).
27. Newman, A. M. et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* **12**, 453–457 (2015).
28. Zhang, C. et al. Tumor purity as an underlying key factor in glioma. *Clin. Cancer Res.* **23**, 6279–6291 (2017).
29. Hodi, F. S. et al. TMB and inflammatory gene expression associated with clinical outcomes following immunotherapy in advanced melanoma. *Cancer Immunol. Res.* **9**, 1202–1213 (2021).
30. Mariathasan, S. et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).
31. Hugo, W. et al. Genomic and transcriptomic features of response to Anti-PD-1 therapy in metastatic melanoma. *Cell* **165**, 35–44 (2016).
32. Van Allen, E. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
33. Nathanson, T. et al. Somatic mutations and neopeptide homology in melanomas treated with CTLA-4 blockade. *Cancer Immunol. Res.* **5**, 84–91 (2017).
34. Mak, M. et al. A patient-derived, pan-cancer EMT signature identifies global molecular alterations and immune target enrichment following epithelial-to-mesenchymal transition. *Clin. Cancer Res.* **22**, 609–620 (2016).
35. Gibbons, D. & Creighton, C. Pan-cancer survey of epithelial-mesenchymal transition markers across the cancer genome Atlas. *Dev. Dyn.* **247**, 555–564 (2018).
36. Ye, X. & Weinberg, R. Epithelial-mesenchymal plasticity: a central regulator of cancer progression. *Trends Cell Biol.* **25**, 675–686 (2015).
37. Marks, D. K. et al. Akt inhibition is associated with favorable immune profile changes within the tumor microenvironment of hormone receptor positive, HER2 negative breast cancer. *Front. Oncol.* **10**, 968 (2020).
38. Kadiyala, P. et al. Inhibition of 2-Hydroxyglutarate Elicits Metabolic-reprogramming and Mutant IDH1 Glioma Immunity in Mice. *J. Clin. Investig.* <https://doi.org/10.1172/jci.139542> (2020).
39. Zhang, J. et al. A novel model of controlling PD-L1 expression in ALK anaplastic large cell lymphoma revealed by CRISPR screening. *Blood* **134**, 171–185 (2019).
40. Cerezo, M. et al. Translational control of tumor immune escape via the eIF4F-STAT1-PD-L1 axis in melanoma. *Nat. Med.* **24**, 1877–1886 (2018).
41. Zhu, Y. et al. Disruption of tumour-associated macrophage trafficking by the osteopontin-induced colony-stimulating factor-1 signalling sensitises hepatocellular carcinoma to anti-PD-L1 blockade. *Gut* **68**, 1653–1666 (2019).
42. Jiang, T. et al. CGCG clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett.* **375**, 263–273 (2016).
43. Jiang, T. et al. Clinical practice guidelines for the management of adult diffuse gliomas. *Cancer Lett.* **499**, 60–72 (2021).
44. Louis, D. et al. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* **114**, 97–109 (2007).
45. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B: Methodol.* **57**, 289–300 (1995).
46. Hanzelmann, S., Castelo, R. & Guinney, J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* **14**, 7 (2013).
47. Bagaev, A. et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* **39**, 845–865.e847 (2021).
48. Wilkerson, M. & Hayes, D. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* **26**, 1572–1573 (2010).
49. Leek, J., Johnson, W., Parker, H., Jaffe, A. & Storey, J. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **28**, 882–883 (2012).
50. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
51. Doucette, T. et al. Immune heterogeneity of glioblastoma subtypes: extrapolation from the cancer genome atlas. *Cancer Immunol. Res.* **1**, 112–122 (2013).
52. Xu, L. et al. TIP: a web server for resolving tumor immunophenotype profiling. *Cancer Res.* **78**, 6575–6580 (2018).
53. Wiesner, S. et al. De novo induction of genetically engineered brain tumors in mice using plasmid DNA. *Cancer Res.* **69**, 431–439 (2009).
54. Guo, Q. et al. Overexpression of oncostatin M receptor regulates local immune response in glioblastoma. *J. Cell. Physiol.* <https://doi.org/10.1002/jcp.28197> (2019).

## Acknowledgements

The authors would like to acknowledge all the members in Dr. Wu AH's laboratory for help with this study. This work was supported by the National Natural Science Foundation of China (nos. U20A20380, 81172409, 81472360, and 81872054 to A.W., no. 81872057 to P.C., no. 81902546 to W.C., no. 82103464 to Q.L. and no. 81672824 to L.C.), Liaoning Science and Technology Plan Projects (no. 2011225034 to A.W.), the National Science Foundation of Liaoning Province (no. 20180550063 to P.C.), the National Postdoctoral Program for Innovative Talents (no. BX20180384 to W.C.), the China Postdoctoral Science Foundation (no. 2022M713484 to S.S. and no. 2019M651169 to W.C.), and Liao Ning Revitalization Talents Program (no. XLYC1807255 to W.C.), the Liaoning Province Science and Technology Plan Project (no. 2020-BS-117 to Q.G.)

## Author contributions

S.S., X.L., Q.G., P.C., W.C. and A.W. contributed to the study conceptualization. Q.L., X.L., Z.Y. and Q.G. contributed to data curation. A.W., W.C., P.C. and L.C. contributed to funding acquisition. J.W., C.Zou., Q.G. and C.Zhu. took charge of the experiments. G.G., L.C. and P.C. took charge of the methodology. T.L., C.Zou. and Q.G. contributed to project administration. W.C. and A.W. supervised the study and took charge of the review and editing. S.S. took charge of validation. S.S. and T.L. contributed to writing the original draft. All authors contributed to the article and approved the submitted version.

## Competing interests

The authors declare no competing interests.

## Ethical statement

The experimental protocol was approved by the ethics committee of The First Hospital of China Medical University. Collection of tumor tissue and clinicopathologic information was obtained with informed consent. All ethical regulations relevant to human research participants were followed. Animal experiments were conducted in accordance with the China Medical University Animal Care and Use Committee guidelines and approved by the Institutional Review Board of the First Hospital of China Medical University. We have complied with all relevant ethical regulations for animal use.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05667-4>.

**Correspondence** and requests for materials should be addressed to Peng Cheng, Wen Cheng or Anhua Wu.

**Peer review information** *Communications Biology* thanks Muh-Hwa Yang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Ruby Huang, Zhijuan Qiu and Johannes Stortz. A peer review file is available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023