

ARTICLE OPEN



Pan-cancer characterization of lncRNA modifiers of immune microenvironment reveals clinically distinct de novo tumor subtypes

Zicheng Zhang^{1,3}, Congcong Yan^{1,3}, Ke Li¹, Siqi Bao¹, Lei Li², Lu Chen¹, Jingting Zhao¹, Jie Sun¹ and Meng Zhou¹

The emerging field of long noncoding RNA (lncRNA)-immunity has provided a new perspective on cancer immunity and immunotherapies. The lncRNA modifiers of infiltrating immune cells in the tumor immune microenvironment (TIME) and their impact on tumor behavior and disease prognosis remain largely uncharacterized. In the present study, a systems immunology framework integrating the noncoding transcriptome and immunogenomics profiles of 9549 tumor samples across 30 solid cancer types was used, and 36 lncRNAs were identified as modifier candidates underlying immune cell infiltration in the TIME at the pan-cancer level. These TIME lncRNA modifiers (TIL-lncRNAs) were able to subclassify various tumors into three de novo pan-cancer subtypes characterized by distinct immunological features, biological behaviors, and disease prognoses. Finally, a TIL-lncRNA-derived immune state index (TISI) that was reflective of immunological and oncogenic states but also predictive of patients' prognosis was proposed. Furthermore, the TISI provided additional prognostic value for existing tumor immunological and molecular subtypes. By applying the TISI to tumors from different clinical immunotherapy cohorts, the TISI was found to be significantly negatively correlated with immune-checkpoint genes and to have the ability to predict the effectiveness of immunotherapy. In conclusion, the present study provided comprehensive resources and insights for future functional and mechanistic studies on lncRNA-mediated cancer immunity and highlighted the potential of the clinical application of lncRNA-based immunotherapeutic strategies in precision immunotherapy.

npj Genomic Medicine (2021)6:52; <https://doi.org/10.1038/s41525-021-00215-7>

INTRODUCTION

Cancer is a highly complicated and delicate disease influenced by not only genetic/epigenetic changes in the tumor cells but also the surrounding complex and dynamic milieu known as the tumor microenvironment (TME)^{1,2}. The mutual and dynamic crosstalk among cellular and molecular components of the TME serves profound roles in tumor initiation, progression, and metastasis^{3,4}. The immunological components of the TME have been recognized as essential hallmark features of the TME through the formation of a vital specialized microenvironment known as the tumor immune microenvironment (TIME)⁵. Recent advances in the precise dissection of the TIME have demonstrated that infiltrating immune cells exert multiple functions in the complex ecosystem of the TME, and their complexity and diversity within the TME can exert both pro- and antitumorigenic effects, as well as affect a variety of clinical outcomes and therapeutic responses, particularly the response to immunotherapy^{6,7}.

Considerable evidence has suggested that the TIME is highly dynamic and plastic during tumor progression and therapeutic interventions, and is determined and remodeled by genetic alterations of oncogenic signaling, genetic and epigenetic regulators, and cellular metabolism^{8–10}. Over the last few years, long noncoding RNAs (lncRNAs) have emerged as critical players in gene regulatory networks, affecting diverse biological and physiological processes^{11,12}. Recent progress in functional studies has also highlighted the crucial roles of lncRNAs in the development and functions of the immune system, and their

potential to regulate all aspects of immunity^{13–15}. lncRNAs are expressed in various immune cell types, preferentially in a lineage-specific manner, and contribute to immune cell development, differentiation, and activation^{16–18}. Increasing evidence also shows that lncRNAs can function as communicators and mediators, being directly and/or indirectly involved in the crosstalk between tumor cells and infiltrating immune cells within the TIME to participate in cancer onset and progression^{19–24}. Furthermore, certain lncRNAs have been shown to influence and regulate the migration and infiltration of immune cells within the TIME, and are associated with tumor immune evasion and prognosis^{24,25}. For example, Yang et al.²⁶ identified lncRNA *EPIC1* as a regulator of tumor immune evasion and response by suppressing tumor cell antigen presentation. lncRNA *TCL6* was found to be correlated with immune cell infiltration and poor survival in breast cancer²⁷.

Despite the evidence that suggests the involvement of lncRNAs in remodeling the TIME, the known TIL-lncRNAs of infiltrating immune cells are limited and need to be characterized further. In the present study, a systems immunology framework was used to identify potential lncRNA modifier candidates of infiltrating immune cells within the TIME through an integrative analysis of the noncoding transcriptome and immunogenomics profile of 9549 tumor samples across 30 solid cancer types. The specific impact of TIL-lncRNAs on tumor behavior and disease prognosis was also investigated herein.

¹School of Biomedical Engineering, School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou, China. ²Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen, China. ³These authors contributed equally: Zicheng Zhang, Congcong Yan. ✉email: suncarajie@wmu.edu.cn; zhoumeng@wmu.edu.cn

RESULTS

Derivation of de novo pan-cancer subtypes associated with distinct immunological features from the perspective of immune cell infiltration-lncRNA crosstalk

In the present study, a systems immunology framework was proposed to identify potential lncRNA modifier candidates of infiltrating immune cells within the TIME by considering the correlation between lncRNA expression and immune-related molecular and cellular components of the TIME at the pan-cancer level (Fig. 1a). A total of 36 lncRNAs were identified as modifier candidates underlying immune cell infiltrating in the TIME (referred to as TII-lncRNAs; Supplementary Table 1). To explore whether molecular subgroups could be discovered by these TII-lncRNAs, as opposed to pan-cancer subtypes previously identified using other transcriptomic features, consensus clustering was conducted for all TCGA patients based on the expression pattern of 36 TII-lncRNAs, which uncovered three major pan-cancer subtypes (referred to here as TII-C1 to -C3; Fig. 1b). The patient distribution of each cancer type across the three subtypes was examined, and the ES was calculated by comparing the number of patients with a given cancer type in each subtype with those of any cancer types in this subtype, and by performing the hypergeometric test to assess statistical significance, as reported in Xie's study²⁸. As shown in Fig. 1c, 17/30 cancer types were over-represented in the TII-C1 subtype and 12 cancer type were over-represented in the TII-C2 subtype. Brain tumors (GBM and brain lower grade glioma) were over-represented in the TII-C3 subtype. Consequently, the subsequent experiments focused on the TII-C1 and TII-C2 subtypes, where cancer types were relatively uniformly distributed.

Next, the characteristics of the TIME were examined by comparing infiltrating stromal and immune cell populations, immune effector activity, immune-mediated tissue-specific destruction, expression of immune-checkpoint genes (ICGs), and immune pathway activity between the TII-C1 and TII-C2 subtypes. As shown in Fig. 1d, the TII-C1 subtype showed significantly higher immune and lower stromal cell scores compared with the TII-C2 subtype, as determined by Wilcoxon rank-sum test ($P < 0.001$; Fig. 1d). The ICR and CYT scores were also significantly higher in the TII-C1 subtype compared with those in the TII-C2 subtype, as determined by the Wilcoxon rank-sum test ($P < 0.001$; Fig. 1d). Further detailed analysis of the relative abundance of 22 immune cell populations using the deconvolution method revealed that the TII-C1 subtype had significantly higher levels of adaptive and activated immune cell infiltration, whereas the TII-C2 subtype demonstrated a low infiltration of adaptive and activated immune cells and high infiltration of inactivated immune cells. Furthermore, ICGs tended to be significantly upregulated in the TII-C1 subtype compared to those in the TII-C2 subtype, as determined by Wilcoxon rank-sum test ($P < 0.001$; Fig. 1d). The pathway enrichment analysis by ssGSEA showed that the TII-C1 subtype could be characterized by upregulated immune pathways, as compared with the TII-C2 subtype (Fig. 1d). Overall, these results demonstrated that the TII-C1 subtype may be associated with an immune-active microenvironment phenotype and the TII-C2 subtype with an immune-silent microenvironment phenotype.

To confirm the stability and reliability of TII-lncRNA-derived subtype classification, randomized analysis was performed by randomly removing 40 or 60% of samples and conducting the same consensus clustering for the remaining 60 or 40% of samples based on the expression of 36 TII-lncRNAs (Fig. S1). The results of the randomized analysis were the same as those of the analysis of the entire TCGA pan-cancer samples, indicating the classification effectiveness and robustness with 36 TII-lncRNAs.

Biological and clinical differences across TII-lncRNA-derived pan-cancer subtypes

Next, the clinical relevance of different TII-lncRNA-derived subtypes was explored. Pan-cancer survival analysis revealed that the TII-lncRNA-derived subtypes were significantly associated with patient survival (HR = 1.872, 95% CI = 1.709–2.05, $P < 0.001$; Fig. 2a), and that the TII-C2 subtype had a significantly superior overall survival than the TII-C1-subtype (log-rank test $P < 0.001$; Fig. 2a). Cancer-specific survival analysis also revealed an association between the TII-lncRNA-derived subtypes and overall survival in multiple cancer types, with a similar prognostic trend as that observed in the pan-cancer analysis (Fig. 2b).

It was then further examined whether tumors in different TII-lncRNA-derived subtypes had different biological features, including EMT process, DNA damage response (DDR), and hallmark gene expression. By examining the EMT status of the tumors, it was found that EMT signature scores of tumors were significantly lower (and more epithelial) in the TII-C1 subtype than in the TII-C2 subtype. A total of 81/94 epithelial markers were upregulated in tumors from the TII-C1 subtype, whereas 74/111 mesenchymal markers were upregulated in the TII-C2 subtype. A total of 169/309 DDR genes were found to be upregulated in the TII-C1 subtype compared with the TII-C2 subtype (Fig. 2c). Furthermore, a significant difference in the overall number of DDR gene mutations was observed between the TII-C1 and TII-C2 subtypes, as determined by the Wilcoxon rank-sum test ($P < 0.001$; Fig. 2d). Clear links between genomic alteration features and the TIME have been reported. Therefore, genomic alteration features were further investigated in the two TII-lncRNA-derived pan-cancer subtypes. Overall, the TII-C1 subtype was characterized by a high level of tumor mutational burden, microsatellite instability, and copy number variation (Fig. 2d). Geneset enrichment analysis by ssGSEA showed that the TII-C1 subtype was associated with high activation of nearly all cancer hallmark pathways (Fig. 2e). In combination, these findings suggested that the biological and clinical behaviors of tumors in different TII-lncRNA-derived subtypes were heterogeneous.

Development of a TISI

Considering the association of TII-lncRNA-derived subtypes with the TIME and biological and clinical behaviors, a TISI was constructed by calculating the mean expression levels of 36 TII-lncRNAs to reflect the affinity immune phenotype of a tumor. As shown in Fig. 3a, cancer types that were over-represented in the TII-C1-subtype exhibited a lower TISI, whereas those over-represented in the TII-C3-subtype exhibited a higher TISI. Furthermore, the TISI exhibited a significant negative correlation with the immune ($R = -0.12$, $P < 0.001$), CYT ($R = -0.19$, $P < 0.001$), and ICR ($R = -0.26$, $P < 0.001$) scores, and a significant positive correlation with the stromal ($R = 0.058$, $P < 0.001$) and EMT ($R = 0.43$, $P < 0.001$) scores (Fig. 3b).

Next, the association between the infiltration status of different immune cell populations and the TISI was examined, and the TISI was found to be negatively correlated with the abundance of adaptive immune cell populations, such as activated CD4 T cells, $\gamma\delta$ T cells, central memory CD8 T cells, and central memory CD4 T cells, and positively correlated with the abundance of innate immune cell populations, such as natural killer cells, monocytes, and mast cells (Fig. 3c). This observation suggested that the TISI has great potential in capturing the antitumor immunity of tumors. Further hallmark pathway enrichment analysis also revealed the enrichment of pathways involved in immunosuppression, such as glycolysis, late estrogen response, and mTOR complex 1 signaling, which were found to be significantly positively correlated with the TISI, whereas the enrichment of most cancer-related pathways was found to be negatively correlated with the TISI. These results demonstrated that the TISI

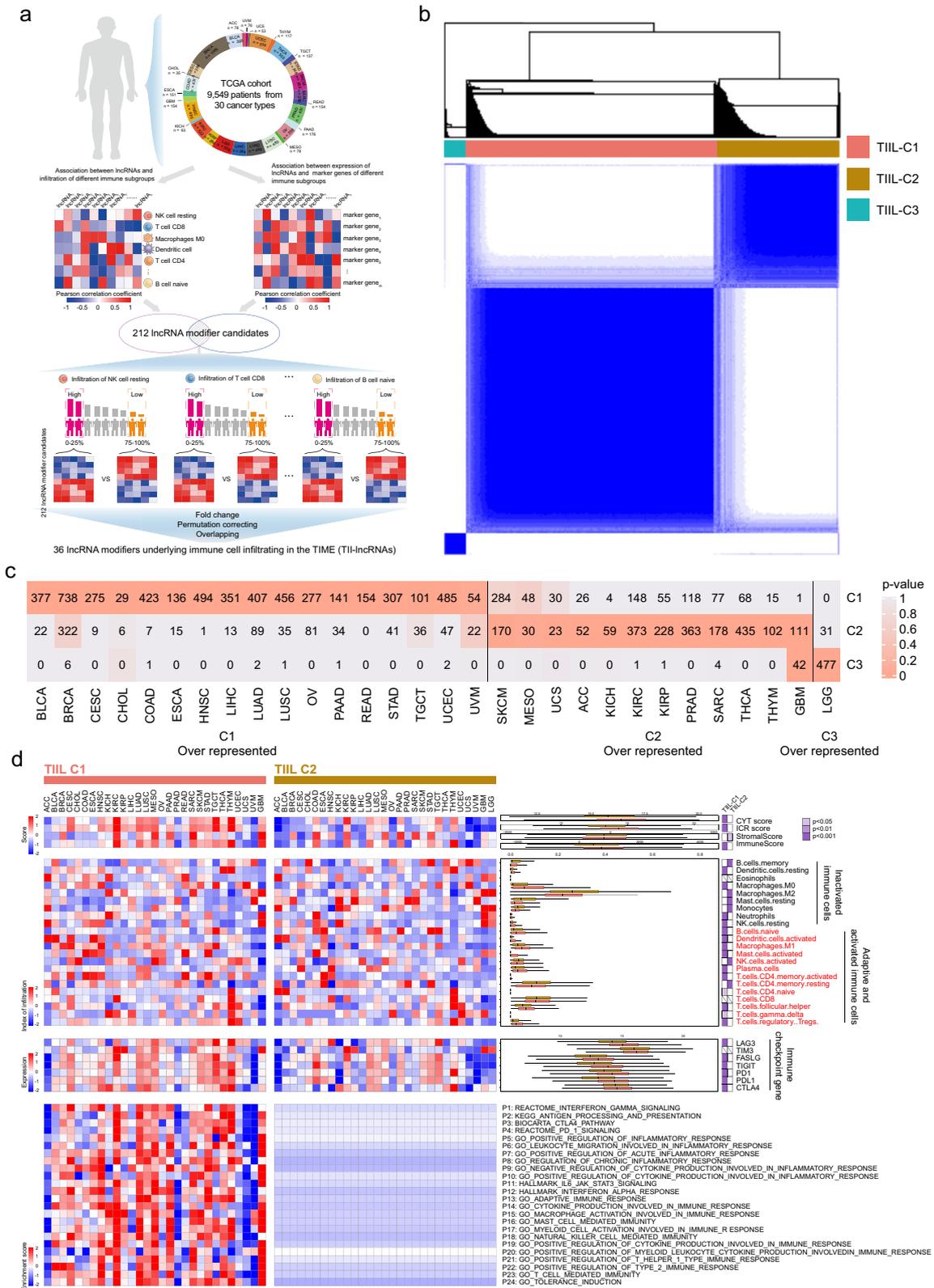


Fig. 1 Derivation of de novo pan-cancer subtypes based on TII-lncRNAs. **a** Schematic illustration of the systems immunology framework for identifying potential lncRNA modifier candidates of infiltrating immune cells within the TIME. **b** Consensus clustering matrix of 9549 TCGA pan-cancer samples based on RNA-seq expression values of 36 TII-lncRNAs. **c** The number of cases in each cluster across tumor types. Colors in the boxes represent the P -value calculated from a hypergeometric test that compared the proportion of samples with a known cancer type in a cluster to the overall proportion of samples in that cluster. Red represents the enrichment level. **d** Comparison of immunological features within the TIME between de novo pan-cancer subtypes defined by TII-lncRNAs. lncRNA long noncoding RNAs, TII-lncRNAs TIME lncRNA modifiers, TCGA The Cancer Genome Atlas.

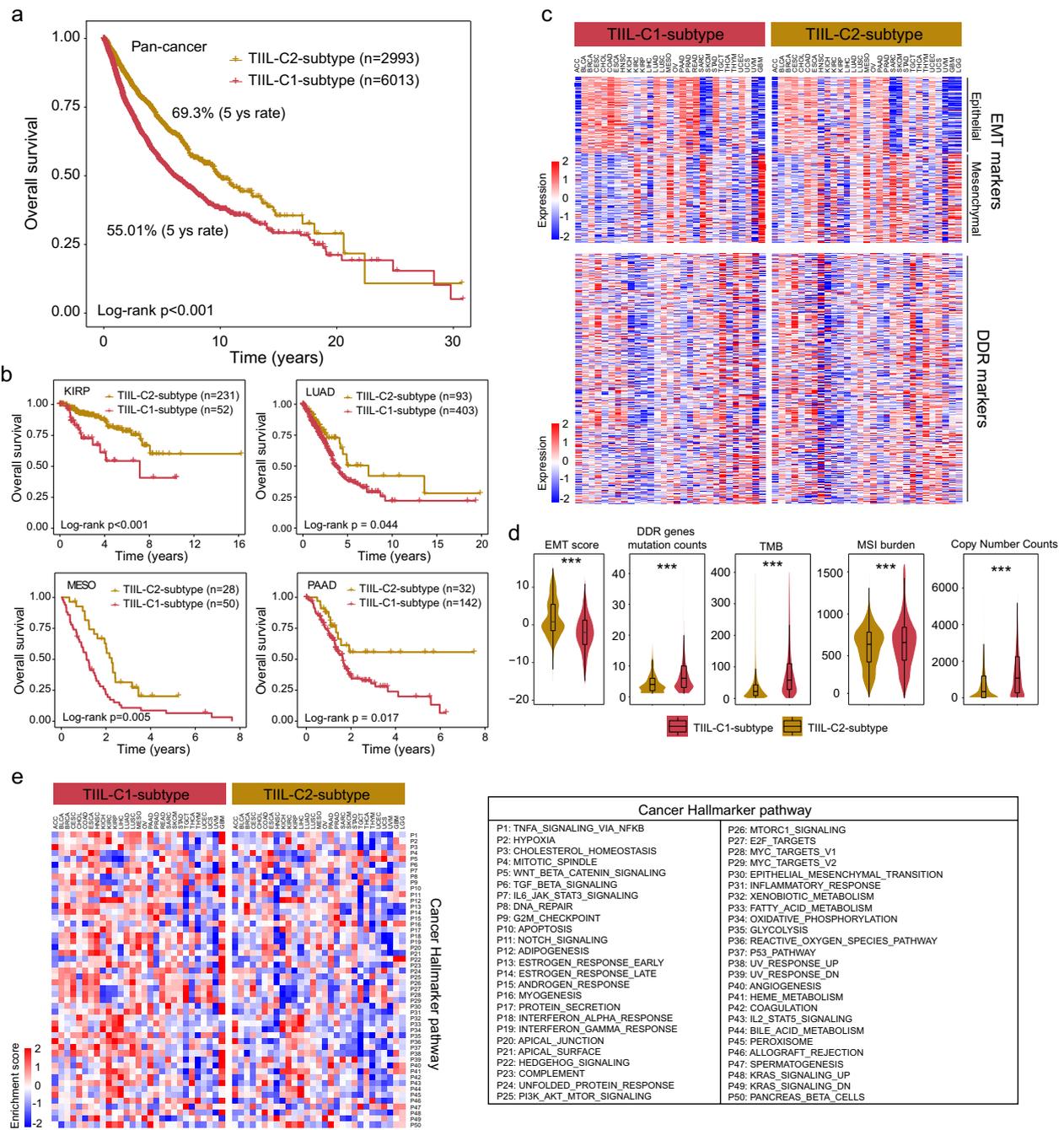


Fig. 2 Biological and clinical heterogeneity across TII-lncRNA-derived pan-cancer subtypes. **a** Pan-cancer Kaplan–Meier survival curves of tumors in patients with the TII-C2 and TII-C1 subtypes. The P -value was calculated by the log-rank test. **b** Kaplan–Meier survival curves of tumor patients with KIRP, LUAD, MESO, and PAAD. The P -value was calculated by the log-rank test. **c** Heatmap of the expression of EMT and DDR markers between the TII-C1 and TII-C2 subtypes. The P -value was calculated by two-sided Wilcoxon rank-sum tests ($***P < 0.001$). **d** Boxplot for genomic alteration features between the TII-C2 and TII-C1 subtypes. The P -value was calculated by two-sided Wilcoxon rank-sum tests ($***P < 0.001$). **e** Heatmap of ssGSEA enrichment score of cancer hallmark pathways between the TII-C1- and TII-C2-subtype. IncRNA long noncoding RNAs, TII-lncRNAs TIME lncRNA modifiers, KIRP kidney renal papillary cell carcinoma, LUAD lung adenocarcinoma, MESO mesothelioma, PAAD pancreatic adenocarcinoma, EMT epithelial-mesenchymal transition, DDR DNA Damage Response, ssGSEA single-sample Gene Set Enrichment Analysis.

was not only associated with intertumoral immune states but also reflected oncogenic states, which were predictive of patient outcomes.

Therefore, the prognostic significance of the TISI in predicting pan-cancer survival was examined. Pan-cancer survival analysis revealed that the TISI was significantly associated with patient survival (HR = 1.627, 95% CI = 1.495–1.770, $P < 0.001$; Fig. 3d). As shown in Fig. 3d, patients with a high TISI had a significantly

improved survival compared with those with a low TISI (log-rank $P < 0.001$). Cancer-specific survival analyses also revealed a significant association between the TISI and survival in multiple cancer types, with a similar prognostic trend observed in the pan-cancer analysis (Fig. 3d). Furthermore, the TISI maintained a significant association with survival after adjusting other standard clinical features in the pan-cancer analysis (Fig. 3e).

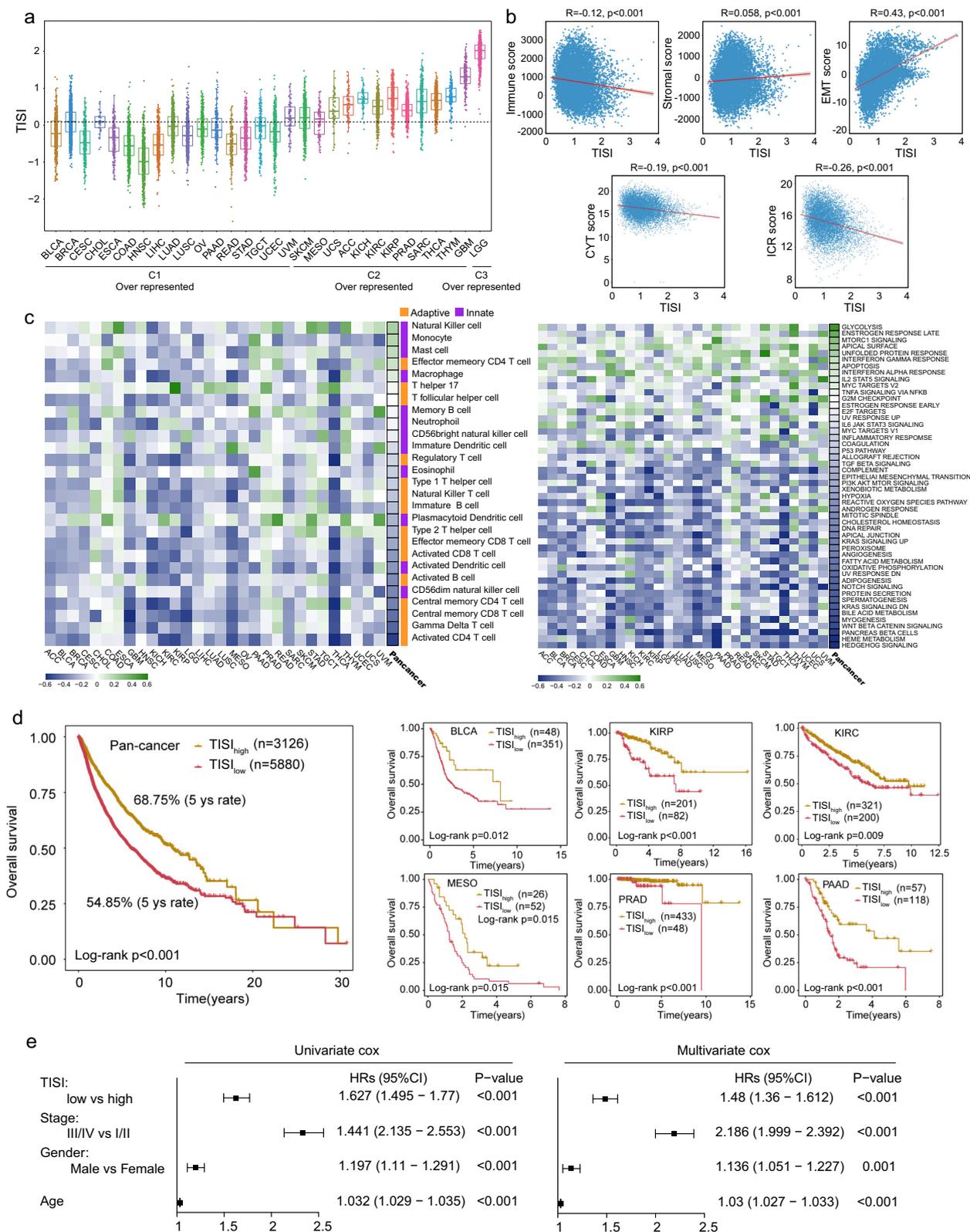


Fig. 3 Development and validation of the TISI in pan-cancer. **a** Distribution of TISI in different cancer types. The dotted line represents the median TISI in all samples. **b** Correlation between cancer-related events and the TISI. **c** Association of the infiltration status of different immune cell populations and cancer hallmark pathways with the TISI. **d** Pan-cancer and cancer-specific Kaplan–Meier survival curves of tumor patients between TISI_{high} and TISI_{low} groups. The *P*-value was calculated by the log-rank test. **e** Univariate and multivariate analysis of the TISI with other standard clinical features in the pan-cancer. *P*-value was calculated by the Cox proportional hazard model. IncRNA long noncoding RNAs, TISI TIL-IncRNA-derived immune state index.

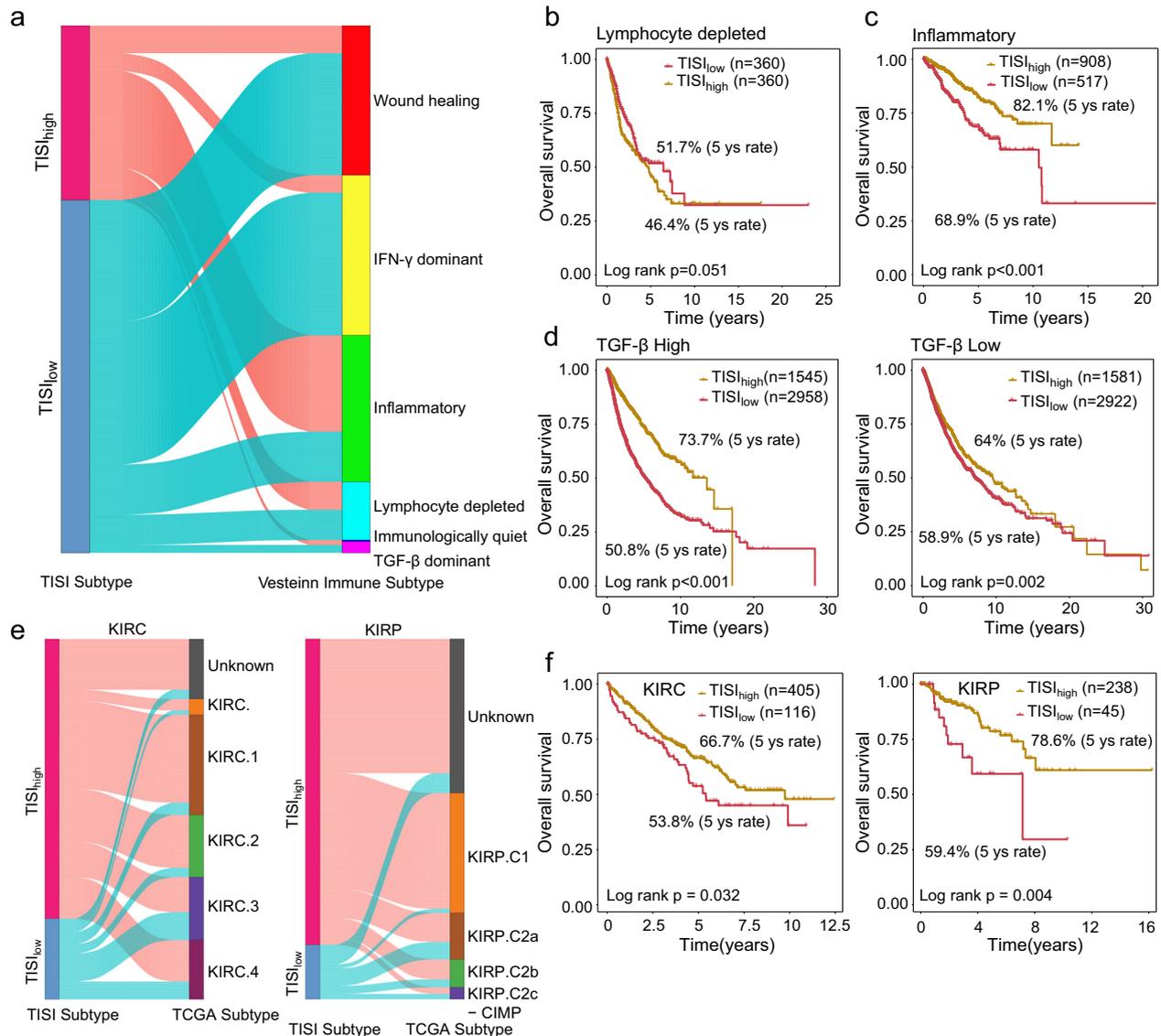


Fig. 4 Concordance between TISi-defined risk groups and immunological and molecular subtypes. **a** Alluvial diagram showing the association between TISi-defined risk groups and immunological subtypes. **b–d** Kaplan–Meier survival curves of tumor patients between the TISi_{high} and TISi_{low} groups in several immunological subtypes. The *P*-value was calculated by the log-rank test. **e** Alluvial diagram showing the association between TISi-defined risk groups and molecular subtypes of KIRC and KIRP. **f** Kaplan–Meier survival curves of tumor patients in the TISi_{high} and TISi_{low} groups in KIRC and KIRP. The *P*-value was calculated by the log-rank test. TISi TIL-lncRNA-derived immune state index, KIRC kidney renal clear cell carcinoma, KIRP kidney renal papillary cell carcinoma.

The TISi provides additional prognostic value for tumor immunological and molecular subtypes

The correlation between TISi-defined risk groups and immune subtypes (ISs) defined by immune signature sets was further investigated²⁹. As shown in Fig. 4a, TISi-defined risk groups were present in all ISs, but their relative distribution differed among ISs. The C1-wound healing, C2-interferon (IFN)- γ dominant and C6-transforming growth factor- β (TGF- β) dominant ISs were particularly dominant in the low TISi group, whereas the high TISi group was enriched in the C3-inflammatory and C5-immunologically quiet ISs. A relatively equal distribution for TISi-defined risk groups was observed in C4-lymphocyte depleted ISs, which was consistent with the fact that C4-lymphocyte depleted ISs include subtypes with mixed signatures, whose prognosis is dependent on immune contexts (Fig. 4a). Further observation indicated that some of the low TISi tumors also occurred in the C3-inflammatory ISs, with some high TISi tumors occurring in the C6-TGF- β

dominant ISs. Therefore, stratification analysis of the TISi for C3-inflammatory, C4-lymphocyte depleted, and C6-TGF- β dominant ISs was conducted. Although the C3-inflammatory IS has been reported to have the best prognosis, the TISi was still able to effectively identify a subgroup of patients at high risk for poor prognosis from the C3-inflammatory IS (log-rank $P < 0.001$; Fig. 4c). Although tumors in the C4-lymphocyte depleted and C6-TGF- β dominant ISs are known to have the least favorable outcome, the TISi revealed a good prognostic prediction ability and could effectively distinguish high- and low-risk patients for the same C4-lymphocyte depleted (log-rank $P = 0.051$) and C6-TGF- β dominant ISs (Fig. 4b and d). For patients with high and low TGF- β signaling, a high TISi was significantly associated with improved survival compared with low TISi (log-rank $P < 0.001$ for high TGF- β signaling, and $P = 0.002$ for low TGF- β signaling; Fig. 4d). The opposite trend was observed in the C4-lymphocyte depleted IS, in

which high TISI was marginally significantly associated with poor survival than low TISI (Fig. 4b).

The interplay of TISI-defined risk groups with the previously defined mRNA-based consensus molecular subtypes (CMSs) was next assessed, and kidney renal clear cell carcinoma (KIRC) and kidney renal papillary cell carcinoma (KIRP) were selected as case studies. It was found that TISI-defined risk groups spanned across CMSs and had no substantial heterogeneity in the distribution of CMS subgroups (Fig. 4e). However, the TISI could effectively stratify patients in the different CMS subgroups of KIRC and KIRP into different risk groups with significantly different overall survival rates (Fig. 4f). As shown in Fig. 4f, patients with a high TISI exhibited a significantly improved survival compared with those with a low TISI (log-rank $P = 0.032$ for KIRC and log-rank $P = 0.004$ for KIRP).

Potential of TISI as a predictor of immunotherapy response

By examining the association between the TISI and the expression of ICGs, the TISI was found to be significantly negatively correlated with the expression of 6 ICGs [$R = -0.25$ and $P < 0.001$ for lymphocyte-activation gene 3, $R = -0.15$ and $P < 0.001$ for CTLA-4, $R = -0.15$ and $P < 0.001$ for Fas ligand, $R = -0.25$ and $P < 0.001$ for T cell immunoreceptor with Ig and ITIM domains, $R = -0.14$ and $P < 0.001$ for PD-1, and $R = -0.29$ and $P < 0.001$ for programmed death-ligand 1 (PD-L1); Fig. 5a]. Therefore, it was next investigated whether the TISI had the potential as a genomic tool to predict treatment response to immune-checkpoint inhibitors (ICIs). The prognostic value of the TISI was evaluated by univariate Cox regression analysis within each of the four patient cohorts receiving anti-PD-1 and anti-CTLA-4 treatment, and was integrated using meta-analysis to estimate an overall prognostic effect. The increased TISI was significantly associated with a favorable prognosis following ICI treatment (HR = 0.46, 95% CI = 0.26–0.81, $P = 0.0074$; Fig. 5b). The Zhao and Miao cohorts, which contained detailed clinical information, were further selected as case studies for the validation of the predictive value of the TISI in ICI response. Analysis of the expression pattern of 36 TIL-IncRNAs in these two patient cohorts revealed two patient clusters with different immunotherapy responses, which was consistent with the findings of TISI classification (Fig. 5c). Furthermore, the TISI was not only significantly negatively correlated with the immune score of patients ($R = -0.48$ and $P = 0.061$ for the Miao cohort, and $R = -0.54$ and $P < 0.001$ for the Zhao cohort; Fig. 5d) but also stratified patients into TISI^{high} and TISI^{low} groups with an apparently different survival (log-rank $P = 0.038$ for the Miao cohort and log-rank $P = 0.084$ for the Zhao cohort; Fig. 5e). The tendency regarding the association between TISI and ICI response indicated that a low TISI revealed significant enrichment in patients that were generally more responsive to ICIs, and a high TISI revealed a predominant enrichment trend in patients that might be more resistant to ICIs (Fig. 5f). Moreover, ROC analysis suggested that the TISI exhibited a predictive superiority or comparable performance (AUC, 0.699 and 0.625) in predicting the response to ICI therapy, as compared with traditional immune biomarkers PD-1, PD-L1, and CTLA-4 (Fig. 5g). These results supported that the TISI is a potential predictive biomarker for the effectiveness of ICI therapy.

DISCUSSION

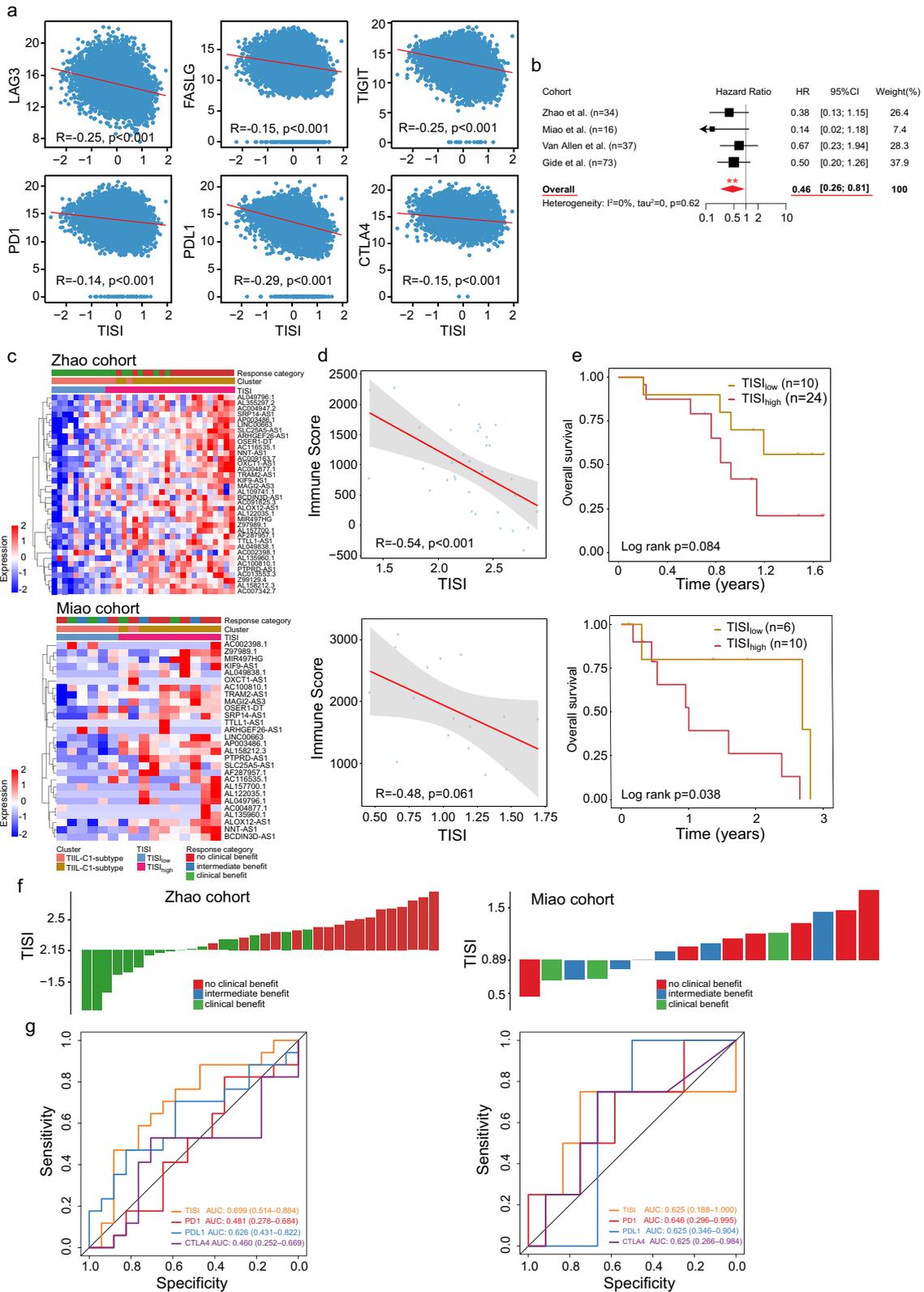
The immune system in the TME has been characterized extensively during the past few years, and the understanding of the complexity and dynamics of immunological compositions within the TIME has shed new light on the mechanisms of immune evasion of cancer cells, as well as the discovery of new immunotherapeutic strategies and biomarkers for clinical benefit^{30–32}. The establishment and maintenance of specific

immunological compositions in the TIME could be affected and regulated by genetic makeup, as well as transcriptional and epigenetic regulators^{7,33,34}. Recently, the emerging field of lncRNA-immunity has provided a new perspective on cancer immunity and immunotherapies. Although lncRNAs have been shown to be crucially involved in regulating immune cell functions and diversity within the TIME, lncRNA modifiers of infiltrating immune cells and their contribution to the immunological and clinical phenotype of patients with cancer remain largely uncharacterized.

Given the compelling body of knowledge regarding lncRNAs and tumor-infiltrating immune cells, a systems immunology framework was developed in the present study to identify potential lncRNA modifiers of infiltrating immune cells within the TIME. These modifiers were found to have been strongly correlated with the expression levels of marker genes and infiltrating levels of at least one immune cell type; a significantly differential expression pattern of these modifiers was also observed between tumors with a high and low infiltration for at least one immune cell type. Through integrative analysis of the noncoding transcriptome and immunogenomics profile of 9549 tumor samples across 30 solid cancer types, 36 lncRNAs were identified as modifier candidates underlying immune cell infiltration in the TIME at the pan-cancer level. The results presented herein provided a framework and comprehensive catalog for elucidating the emerging functional roles of lncRNAs in immune modulation.

Previous studies have demonstrated that lncRNA-based molecular subtyping can provide valuable insights into the molecular landscape of cancer, separately from previous protein-coding gene-centered views^{35–38}. By focusing on the expression pattern of 36 TIL-IncRNAs, various tumors were subclassified into three pan-cancer subtypes with distinct immunological features, biological behaviors, and disease prognoses. This subtyping scheme expanded the previously existing molecular subtype classification system, and may have clinical implications for TME lncRNA modifiers in terms of prognosis stratification and therapy response prediction. Thus, a TISI index that can predict intratumoral immune and oncogenic states across different cancer types was developed. In addition, the TISI index also had a prognostic value at the pan-cancer level and was able to distinguish between patients with improved and poor survival outcomes. A total of 6 ISs spanning multiple tumor types have previously been reported. Herein, the intersection between the TISI subgroup and ISs was examined, and it was found that TISI_{low} patients tended to be preferentially distributed in the wound healing and IFN- γ dominant subtypes, whereas TISI_{high} patients were preferentially distributed in the inflammatory and immunologically quiet subtypes. For the other two immune types, the TISI was able to stratify patients from the same immune subtype into two risk groups with differential survival time. For transcriptomic molecular subtypes, it was not only shown that the TISI-defined risk groups spanned across CMSs and had no substantial heterogeneity in CMS distribution, but also allowed for a more precise categorization of patients within the same molecular subtype. These findings demonstrated that the TISI stratification scheme is not only primarily driven by the immunological or transcriptomic features alone but also revealed that there is a crosstalk between these features in the TIME.

Immunotherapy has shown great potential as an innovative form of cancer treatment³⁹. However, immunotherapeutic responsiveness varies across different cancer types, and even across different patients with the same type of cancer. Emerging evidence has indicated the potential of lncRNAs to predict and guide immunotherapeutic responsiveness^{40–42}. Given the association of the TISI with immune and oncogenic states, the clinical relevance of the TISI in cancer immunotherapy was further determined. A sufficient correlation was observed between the



TISi and the outcome of and response to cancer immunotherapy in several types of cancer. Furthermore, the TISi revealed a higher predictive performance compared to previous biomarkers (such as PD-1/PD-L1 and CTLA-4) in certain cases, likely due to the additional information on the crosstalk between transcriptomic

and immunogenomic features provided by the TISi. These findings further supported the potential of lncRNAs as predictive biomarkers or therapeutic targets in cancer immunotherapy.

In the present study, the impact of lncRNA expression on immune-related molecular and cellular components in the TIME at

Fig. 5 Association between the TISI and the clinical response to immunotherapy. **a** Association between the TISI and expression of immune-checkpoint genes. **b** Forest plot visualizing the HRs of univariate survival analysis of the TISI in four immunotherapy cohorts. Significance was determined using the Cox proportional hazard model. The red diamond shows the random-effects meta-analysis summary of HRs over four immunotherapy cohorts (HR = 0.46; 95% CI: 0.26–0.81, $P = 0.0074$). **c** Unsupervised clustering of tumors receiving ICI therapy based on the expression pattern of 36 TII-lncRNAs. **d** Correlation between the TISI with the immune score in the Zhao and Miao cohorts. **e** Kaplan–Meier survival curves of tumor patients between TISI_{high} and TISI_{low} groups in the Zhao and Miao cohorts. **f** Association between TISI and ICI responses. Tumors were sorted according to their TISI. **g** ROC curves for ICB response for the TISI, PD-1, PD-L1, and CTLA-4 in two independent immunotherapy cohorts. TISI TIL-lncRNA-derived immune state index, HRs hazard ratios, CI confidence interval, lncRNAs long noncoding RNAs, ICB immune-checkpoint blockade, PD-1 programmed cell death protein 1, PD-L1 programmed death-ligand 1, CTLA-4 cytotoxic T-lymphocyte-associated protein 4, ICIs immune-checkpoint inhibitors, ROC receiver operating characteristic.

the pan-cancer level was systematically investigated, and lncRNA modifier candidates of infiltrating immune cells within the TIME were identified, providing a comprehensive resource and view for future functional and mechanistic investigations of lncRNA-mediated cancer immunity. In addition, the clinical relevance of these TIL-lncRNAs with survival and response to cancer immunotherapy presented herein highlighted the future potential of the clinical application of lncRNA-based immunotherapeutic strategies in precision immunotherapy.

METHODS

Pan-cancer patient and immunotherapy data

The Cancer Genome Atlas (TCGA) multi-omics data were obtained from UCSC Xena (<https://xenabrowser.net/datapages/>)⁴³, including RNA-seq with HiSeq Illumina platform transformed by $\log(x + 1)$, copy number variation with the Illumina platform estimated using the GISTIC2 method, and somatic mutation with the Illumina platform. A total of 9549 tumor samples across 30 solid cancer types were analyzed in this study.

Transcriptomic data (HiSeq Illumina platform) and clinical information from patients with tumors treated with programmed cell death protein 1 (PD-1)/cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) blockade therapy were obtained from previously published prospective clinical trials, including from 16 patients with clear metastatic cell renal cell carcinoma (ccRCC) from Miao's study (referred to as the Miao cohort)⁴⁴, 34 patients with glioblastoma multiforme (GBM) from Zhao's study (referred to as the Zhao cohort)⁴⁵, 37 patients with metastatic melanoma from VanAllen's study (referred to as the VanAllen cohort)⁴⁶, and 76 patients with melanoma from Gide's study (referred to as the Gide cohort)⁴⁷. Ethical approval was not needed because these datasets are allowed to be publicly available.

Transcriptome deconvolution of the TIME

The abundance of infiltrating immune cell populations in the TIME was estimated by deconvolution methods using the CIBERSORT with LM22 signature matrix⁴⁸. The fraction of stromal and immune cells in the tumor samples was calculated by the single-sample geneset enrichment analysis (ssGSEA) using ESTIMATE⁴⁹.

Identification of lncRNA modifiers of tumor-infiltrating immune cells

lncRNA expression profiles of TCGA pan-cancer cases were obtained based on the lncRNA annotation of the GENCODE project⁵⁰. Following the removal of the lncRNAs with no expression in >20% of total patients⁵¹, a total of 4755 lncRNAs were kept for further analysis. lncRNA modifiers of tumor-infiltrating immune cells (TII-lncRNAs) were identified using a systems immunology framework as follows (Fig. 1a): (i) All lncRNAs were ranked based on their co-expression relationship with immune marker genes, and lncRNAs with a significantly higher Pearson correlation coefficient (PCC) of 0.3 and $P < 0.05$ were considered as immune gene-related lncRNAs; (ii) these immune gene-related lncRNAs were ranked based on the correlation between their expression and the abundance of a given infiltrating immune cell population, as calculated by PCC, and those with a significantly higher PCC were considered as candidate TII-lncRNAs; (iii) samples were classified into high and low immune infiltration groups using the top and bottom quartiles for a given immune cell population. Candidate TII-lncRNAs that were significantly differentially expressed

between tumors with a high and low immune infiltration were defined as TII-lncRNAs.

Computational index of cancer-related events and immunomodulation

To quantify the role and dynamics of epithelial-mesenchymal transition (EMT) in each patient, 94 mesenchymal and 111 epithelial cell marker genes were obtained from a previous study⁵², and the EMT score was calculated using the Student's *t*-test score between epithelial cell marker genes and the expression of mesenchymal cell marker genes⁵³. To characterize the occurrence of immune-mediated tissue-specific destruction, the Immunologic Constant of Rejection (ICR) score was calculated using the mean of the normalized log2 transformed expression values of 20 ICR signature genes from Roeland's study⁵⁴. To quantify the immune effector activity in solid tumors, the cytolytic activity (CYT) score was calculated using the geometric mean of two key cytolytic effectors, granzyme A, and perforin, in each patient⁵⁵.

Development of TII-lncRNA-derived immune state index (TISI)

The TISI was calculated using the mean expression levels of 36 TII-lncRNAs. Samples with a higher TISI exhibited a low-affinity immune phenotype, whereas a lower TISI reflected a high-affinity immune phenotype.

Enrichment analysis of functional genesets

Single-sample Gene Set Enrichment Analysis (ssGSEA) was performed to calculate the enrichment score (ES) of each patient using R package 'GSVA'⁵⁶ and identify up- or downregulated interested genesets or pathways in different subtypes within each tumor type. The immunologic signature and hallmark genesets were obtained from the Molecular Signatures Database (MSigDB, V7.2)⁵⁷.

Statistical analysis

Consensus clustering analysis was performed on the R package 'ConsensusClusterPlus' using the K-means method and Euclidean distances to identify the optimum number of clusters in pan-cancer based on the expression pattern of TII-lncRNAs. Univariate and multivariate Cox proportional hazards regression models were used to assess the association between the TISI and overall survival with/without clinical variables. The hazard ratio (HR) and 95% confidence interval (CI) were calculated. Two-sided Wilcoxon rank-sum tests were used to compare two groups. The Kaplan–Meier method and log-rank test were conducted to compare survival differences between two tumor groups. Receiver operating characteristic (ROC) curves were used to evaluate the predictive performance for the response to immunotherapy, and the area under the curve (AUC) was calculated. All statistical analysis was performed using R/Bioconductor (version 3.6.1).

Reporting summary

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

DATA AVAILABILITY

The Cancer Genome Atlas (TCGA) pan-cancer data were obtained from UCSC Xena (https://gdc-hub.s3.us-east-1.amazonaws.com/download/GDC-PANCAN.htseq_fpkm-ucsv.gz), including adrenocortical carcinoma (ACC, $n = 78$), bladder urothelial carcinoma (BLCA, $n = 399$), breast invasive carcinoma (BRCA, $n = 1066$), cervical

squamous cell carcinoma and endocervical adenocarcinoma (CESC, $n=284$), cholangiocarcinoma (CHOL, $n=35$), colon adenocarcinoma (COAD, $n=431$), esophageal carcinoma (ESCA, $n=151$), glioblastoma multiforme (GBM, $n=154$), head and neck squamous cell carcinoma (HNSC, $n=495$), kidney chromophobe (KICH, $n=63$), kidney renal clear cell carcinoma (KIRC, $n=522$), kidney renal papillary cell carcinoma (KIRP, $n=284$), brain lower grade glioma (LGG, $n=508$), liver hepatocellular carcinoma (LIHC, $n=364$), lung adenocarcinoma (LUAD, $n=498$), lung squamous cell carcinoma (LUSC, $n=489$), mesothelioma (MESO, $n=78$), ovarian serous cystadenocarcinoma (OV, $n=358$), pancreatic adenocarcinoma (PAAD, $n=176$), prostate adenocarcinoma (PRAD, $n=481$), rectum adenocarcinoma (READ, $n=154$), sarcoma (SARC, $n=259$), skin cutaneous melanoma (SKCM, $n=454$), stomach adenocarcinoma (STAD, $n=348$), testicular germ cell tumors (TGCT, $n=137$), thyroid carcinoma (THCA, $n=503$), thymoma (THYM, $n=117$), uterine corpus endometrial carcinoma (UCEC, $n=534$), uterine carcinosarcoma (UCS, $n=53$), uveal melanoma (UVM, $n=76$). Transcriptomic data (HiSeq Illumina platform) and clinical information of four immunotherapy datasets were obtained from previously published prospective clinical trials, including Miao's study⁴⁴, Zhao's study⁴⁵, VanAllen's study⁴⁶, and Gide's study⁴⁷.

Received: 26 March 2021; Accepted: 25 May 2021;

Published online: 17 June 2021

REFERENCES

1. Baghban, R. et al. Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Commun. Signal* **18**, 59 (2020).
2. Hamidi, H. & Ivaska, J. Every step of the way: integrins in cancer progression and metastasis. *Nat. Rev. Cancer* **18**, 533–548 (2018).
3. Anderson, N. M. & Simon, M. C. The tumor microenvironment. *Curr. Biol.* **30**, R921–R925 (2020).
4. Chen, F. et al. New horizons in tumor microenvironment biology: challenges and opportunities. *BMC Med.* **13**, 45 (2015).
5. Jin, M. Z. & Jin, W. L. The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct. Target. Ther.* **5**, 166 (2020).
6. Zhang, Z. et al. Computational principles and practice for decoding immune contexture in the tumor microenvironment. *Brief Bioinformatics* <https://doi.org/10.1093/bib/bbaa075> (2020).
7. Binnewies, M. et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–550 (2018).
8. Bolouri, H. Network dynamics in the tumor microenvironment. *Semin. Cancer Biol.* **30**, 52–59 (2015).
9. Duan, Q., Zhang, H., Zheng, J. & Zhang, L. Turning cold into hot: firing up the tumor microenvironment. *Trends Cancer* **6**, 605–618 (2020).
10. Flister, M. J. & Bergom, C. Genetic modifiers of the breast tumor microenvironment. *Trends Cancer* **4**, 429–444 (2018).
11. Yao, R. W., Wang, Y. & Chen, L. L. Cellular functions of long noncoding RNAs. *Nat. Cell Biol.* **21**, 542–551 (2019).
12. Statello, L., Guo, C. J., Chen, L. L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).
13. Atianand, M. K., Caffrey, D. R. & Fitzgerald, K. A. Immunobiology of long non-coding RNAs. *Annu. Rev. Immunol.* **35**, 177–198 (2017).
14. Chen, Y. G., Satpathy, A. T. & Chang, H. Y. Gene regulation in the immune system by long noncoding RNAs. *Nat. Immunol.* **18**, 962–972 (2017).
15. Geng, H. & Tan, X. D. Functional diversity of long non-coding RNAs in immune regulation. *Genes Dis.* **3**, 72–81 (2016).
16. Elling, R., Chan, J. & Fitzgerald, K. A. Emerging role of long noncoding RNAs as regulators of innate immune cell development and inflammatory gene expression. *Eur. J. Immunol.* **46**, 504–512 (2016).
17. Ranzani, V. et al. The long intergenic noncoding RNA landscape of human lymphocytes highlights the regulation of T cell differentiation by linc-MAF-4. *Nat. Immunol.* **16**, 318–325 (2015).
18. Wang, L. et al. Integrative genome-wide analysis of long noncoding RNAs in diverse immune cell types of melanoma patients. *Cancer Res.* **78**, 4411–4423 (2018).
19. Agirre, X. et al. Long non-coding RNAs discriminate the stages and gene regulatory states of human humoral immune response. *Nat. Commun.* **10**, 821 (2019).
20. Chen, D. et al. Long non-coding RNAs as communicators and mediators between the tumor microenvironment and cancer cells. *Front. Oncol.* **9**, 739 (2019).
21. de Lima, D. S. et al. Long noncoding RNAs are involved in multiple immunological pathways in response to vaccination. *Proc. Natl Acad. Sci. USA* **116**, 17121–17126 (2019).
22. Del Vecchio, F. et al. Long non-coding RNAs within the tumour microenvironment and their role in tumour-stroma cross-talk. *Cancer Lett.* **421**, 94–102 (2018).
23. Lin, Y. H., Wu, M. H., Yeh, C. T. & Lin, K. H. Long non-coding RNAs as mediators of tumor microenvironment and liver cancer cell communication. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms19123742> (2018).
24. Zhang, Y., Liu, Q. & Liao, Q. Long noncoding RNA: a dazzling dancer in tumor immune microenvironment. *J. Exp. Clin. Cancer Res.* **39**, 231 (2020).
25. Yu, W. D., Wang, H., He, Q. F., Xu, Y. & Wang, X. C. Long noncoding RNAs in cancer-immunity cycle. *J. Cell Physiol.* **233**, 6518–6523 (2018).
26. Guo, W. et al. LincRNA-immunity landscape analysis identifies EPIC1 as a regulator of tumor immune evasion and immunotherapy resistance. *Sci. Adv.* <https://doi.org/10.1126/sciadv.abb3555> (2021).
27. Zhang, Y. et al. lncRNA TCL6 correlates with immune cell infiltration and indicates worse survival in breast cancer. *Breast Cancer* **27**, 573–585 (2020).
28. Xie, F. et al. Multifactorial deep learning reveals pan-cancer genomic tumor clusters with distinct immunogenomic landscape and response to immunotherapy. *Clin. Cancer Res.* **26**, 2908–2920 (2020).
29. Thorsson, V. et al. The immune landscape of cancer. *Immunity* **51**, 411–412 (2019).
30. Gajewski, T. F., Schreiber, H. & Fu, Y. X. Innate and adaptive immune cells in the tumor microenvironment. *Nat. Immunol.* **14**, 1014–1022 (2013).
31. Zhang, Y. & Zhang, Z. The history and advances in cancer immunotherapy: understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell Mol. Immunol.* **17**, 807–821 (2020).
32. Bedognetti, D. et al. Toward a comprehensive view of cancer immune responsiveness: a synopsis from the SITC workshop. *J. Immunother. Cancer* **7**, 131 (2019).
33. Wellenstein, M. D. & de Visser, K. E. Cancer-cell-intrinsic mechanisms shaping the tumor immune landscape. *Immunity* **48**, 399–416 (2018).
34. Chen, Y. P. et al. Single-cell transcriptomics reveals regulators underlying immune cell diversity and immune subtypes associated with prognosis in nasopharyngeal carcinoma. *Cell Res.* **30**, 1024–1042 (2020).
35. Niknafs, Y. S. et al. The lncRNA landscape of breast cancer reveals a role for DSCAM-AS1 in breast cancer progression. *Nat. Commun.* **7**, 12791 (2016).
36. Chen, H. et al. Long noncoding RNA profiles identify five distinct molecular subtypes of colorectal cancer with clinical relevance. *Mol. Oncol.* **8**, 1393–1403 (2014).
37. de Jong, J. J. et al. Long non-coding RNAs identify a subset of luminal muscle-invasive bladder cancer patients with favorable prognosis. *Genome Med.* **11**, 60 (2019).
38. Zhou, M. et al. Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol. Cancer* **16**, 16 (2017).
39. Kruger, S. et al. Advances in cancer immunotherapy 2019—latest trends. *J. Exp. Clin. Cancer Res.* **38**, 268 (2019).
40. Li, Y. et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. *Nat. Commun.* **11**, 1000 (2020).
41. Sun, J. et al. Identification of tumor immune infiltration-associated lncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. *J. Immunother. Cancer* <https://doi.org/10.1136/jitc-2019-000110> (2020).
42. Yu, Y. et al. Association of long noncoding RNA biomarkers with clinical immune subtype and prediction of immunotherapy response in patients with cancer. *JAMA Netw. Open* **3**, e202149 (2020).
43. Goldman, M. J. et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678 (2020).
44. Miao, D. et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science* **359**, 801–806 (2018).
45. Zhao, J. et al. Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma. *Nat. Med.* **25**, 462–469 (2019).
46. Van Allen, E. M. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
47. Gide, T. N. et al. Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. *Cancer Cell* **35**, 238–255 (2019).
48. Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M. & Alizadeh, A. A. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* **1711**, 243–259 (2018).
49. Yoshihara, K. et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013).
50. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* **47**, D766–D773 (2019).
51. Biswas, D. et al. A clonal expression biomarker associates with lung cancer mortality. *Nat. Med.* **25**, 1540–1548 (2019).

52. Tan, T. Z. et al. Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol. Med.* **6**, 1279–1293 (2014).
53. Zhao, Z. et al. CtBP promotes metastasis of breast cancer through repressing cholesterol and activating TGF- β signaling. *Oncogene* **38**, 2076–2091 (2019).
54. Roelands, J. et al. Oncogenic states dictate the prognostic and predictive connotations of intratumoral immune response. *J. Immunother. Cancer* <https://doi.org/10.1136/jitc-2020-000617> (2020).
55. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, 48–61 (2015).
56. Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
57. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).

ACKNOWLEDGEMENTS

This study was supported by the National Natural Science Foundation of China (Grant No. 61973240 and 62072341) and the Fundamental Research Funds for the Provincial Universities of Zhejiang (Grant No. KYW202001). The funders had no roles in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We also thank Xena project.

AUTHOR CONTRIBUTIONS

M.Z. and J.S. conceived and designed the study. Z.C.Z., C.C.Y., K.L., S.Q.B., L.L., L.C. and J.T.Z. prepared and carried out all analyses, including the development of their statistical framework, and interpreting the data. Z.C.Z. and C.C.Y. contributed equally to this work. M.Z. and J.S. drafted the manuscript. All authors participated in the interpretation and discussion of the results and in the version of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41525-021-00215-7>.

Correspondence and requests for materials should be addressed to J.S. or M.Z.

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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021