



**Universiteit
Leiden**
The Netherlands

A balance score between immune stimulatory and suppressive microenvironments identifies mediators of tumour immunity and predicts pan-cancer survival

Turan, T.; Kongpachith, S.; Halliwill, K.; Roelands, J.; Hendrickx, W.; Marincola, F.M.; ... ; Ceccarelli, M.

Citation

Turan, T., Kongpachith, S., Halliwill, K., Roelands, J., Hendrickx, W., Marincola, F. M., ... Ceccarelli, M. (2022). A balance score between immune stimulatory and suppressive microenvironments identifies mediators of tumour immunity and predicts pan-cancer survival. *British Journal Of Cancer*, (124), 760-769. doi:10.1038/s41416-020-01145-4

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3627398>

Note: To cite this publication please use the final published version (if applicable).



ARTICLE

Translational Therapeutics

A balance score between immune stimulatory and suppressive microenvironments identifies mediators of tumour immunity and predicts pan-cancer survival

Tolga Turan¹, Sarah Kongpachith¹, Kyle Halliwill¹, Jessica Roelands^{1,2,3}, Wouter Hendrickx², Francesco M. Marincola⁴, Thomas J. Hudson¹, Howard J. Jacob⁵, Davide Bedognetti^{2,6}, Josue Samayoa¹ and Michele Ceccarelli^{7,8}

BACKGROUND: The balance between immune-stimulatory and immune-suppressive mechanisms in the tumour microenvironment is associated with tumour rejection and can predict the efficacy of immune checkpoint-inhibition therapies. **METHODS:** We consider the observed differences between the transcriptional programmes associated with cancer types where the levels of immune infiltration predict a favourable prognosis versus those in which the immune infiltration predicts an unfavourable prognosis and defined a score named **Mediators of Immune Response Against Cancer in soLid microEnvironments (MIRACLE)**. MIRACLE deconvolves T cell infiltration, from inhibitory mechanisms, such as TGFβ, EMT and PI3Kγ signatures. **RESULTS:** Our score outperforms current state-of-the-art immune signatures as a predictive marker of survival in TCGA ($n = 9305$, HR: 0.043, p value: 6.7×10^{-36}). In a validation cohort ($n = 7623$), MIRACLE predicts better survival compared to other immune metrics (HR: 0.1985, p value: 2.73×10^{-38}). MIRACLE also predicts response to checkpoint-inhibitor therapies ($n = 333$). The tumour-intrinsic factors inversely associated with the reported score such as EGFR, PRKAR1A and MAP3K1 are frequently associated with immune-suppressive phenotypes. **CONCLUSIONS:** The association of cancer outcome with the level of infiltrating immune cells is mediated by the balance of activatory and suppressive factors. MIRACLE accounts for this balance and predicts favourable cancer outcomes.

British Journal of Cancer (2021) 124:760–769; <https://doi.org/10.1038/s41416-020-01145-4>

BACKGROUND

The presence of an active immune microenvironment with a high density of activated T cells associates with favourable prognosis and responsiveness to immunotherapy.¹ Genetic factors preventing the development of a favourable immune milieu and/or responsiveness to immunotherapy include limited expression of neoantigens,^{2,3} which is partially a function of mutational load, and high tumour aneuploidy. However, in the “immune-active” tumours, upregulation of pro-inflammatory signals is accompanied by the expression of interferon-gamma-inducible immune regulatory molecules like programmed death ligand-1 (PD-L1) and indoleamine 2,3-dioxygenase (IDO), whose levels correlate with response to anti-PD-1 and anti-cytotoxic T-lymphocyte-associated protein 4 (anti-CTLA4) treatment. The expression of such molecules as well as of other immune-regulatory markers (e.g., FOXP3, CTLA4 and PD-1) characterise a *compensatory immune resistance*, reflecting the presence of counter-regulatory mechanisms that follow, rather than precede,^{4,5} the recognition of tumour antigens by T cells and the subsequent amplification of the inflammatory response.^{6,7} Correlative studies in humans and experimental models suggest that

checkpoint inhibitors are less effective in tumours characterised by a primary immune suppression (also called as “primary immune ignorance”), including the ones with low mutational load,^{1,8} or dominated by the genomic dysregulations of oncogenic pathways leading to T cell exclusion such as WNT/beta-catenin,^{9–11} mitogen-activated protein kinase^{6,7,12} and transforming growth factor-β (TGFβ) pathways^{13–15}. The dichotomy between “immune active” (associated with the displaying of compensatory immune resistance) and “immune silent” (typified by the presence of a primary immune suppression) might be useful to explain a general phenomenon but do not reflect the high level of inter-patient heterogeneity and do not take into account the contribution of antagonist signals involved in primary immune suppression. In fact, one of the major limitations of the transcriptomic studies performed so far is the use of gene signatures or modules that capture only a dominant process or a group of processes tightly interconnected or correlated among each other.

We speculate that the concurrent measurement of stimulatory and suppressive immunologic signals might better define the true immune state of the tumour–immune microenvironment and thus develop the **Mediators of Immune Response Against Cancer** in

¹Computational Immunology and Oncology (CIAO), AbbVie, Redwood City, CA, USA; ²Cancer Research Department, Sidra Medicine, Doha, Qatar; ³Department of Surgery, Leiden University Medical Center (LUMC), Leiden, The Netherlands; ⁴Refuge Biotechnologies, Menlo Park, CA, USA; ⁵Genomics Research Center (GRC), AbbVie, Lake County, IL, USA; ⁶Dipartimento di Medicina Interna e Specialità Mediche, Università degli Studi di Genova, Genova, Italy; ⁷Department of Electrical Engineering and Information Technology, University of Naples “Federico II”, Naples, Italy and ⁸Istituto di Ricerche Genetiche “G. Salvatore”, Biogem s.c.ar.l. 83031 Ariano Irpino, Italy
Correspondence: Davide Bedognetti (dbedognetti@sidra.org) or Josue Samayoa (josue.samayoa@abbvie.com) or Michele Ceccarelli (michele.ceccarelli@unina.it)

Received: 18 March 2020 Revised: 26 September 2020 Accepted: 15 October 2020
Published online: 5 November 2020

soLid microEnvironments (MIRACLE), a score that captures the balance between opposite immune-related signals.

We show that MIRACLE is a better prognostic factor than the current immune-associated signatures. We also evaluate MIRACLE in the context of response to immune checkpoint inhibitor (ICI) therapy. Further, in contrast to other measures of the tumour-immune microenvironment,^{16,17} MIRACLE is an absolute measure that can be applied to a single sample and does not need any data set-specific normalisation.

METHODS

Data collection

All The Cancer Genome Atlas (TCGA) data (31 solid tumour cohorts) were downloaded using the TCGAbiolinks package and from Genomic Data Commons.¹⁸ RNAseq expression data were downloaded in the form of raw counts and survival attributes were downloaded in the form of Clinical Data Tables. All cohorts except TCGA-SKCM were exclusively primary tumours. Cancer types, Pheochromocytoma and Paraganglioma (PCPG), Thymoma (THYM) and Testicular Germ cell tumours (TGCT), were excluded from the analysis as the number of deaths in the comparison groups was too small for survival estimation.¹⁹

Besides TCGA, publicly available data sets were downloaded from NCBI Gene Expression Omnibus (GEO), European Nucleotide Archive or Short Read Archive. The clinical phenotypes were harmonised between raw or processed data and the associated publications. GEO data sets (or data sets from other Public Repositories) were treated as separate studies and were not merged, even when they are annotated as the same Solid Tumour Cohort.

For microarray data sets, processed data with acceptable normalisation (quantile, RMA or equivalent) were downloaded as series matrix files from GEO, when available. If processed data were not available, raw data were downloaded. For RNAseq data sets, raw gene counts were downloaded when available, otherwise raw.fastq files were downloaded and processed (Table S1 lists the public data sets used in this study.).

Data processing

All data processing was done in R version 3.5.2 except when indicated.

RNA sequencing. Raw fastq files were adapter- and quality-trimmed using Trim_Galore/CutAdapt version 3.5.2. The trimmed reads were aligned to hg19 using STAR RNAseq aligner (STAR_2.5.3a). Raw gene counts were generated using summarizeOverlaps functions of the GenomicAlignments Bioconductor package. Raw counts were then normalised (within lane and between lane) using EDASeq Bioconductor package. All heatmaps and geneset enrichment analyses were performed on the EDASeq normalised matrices.

Microarray. Affymetrix.CEL files were downloaded from GEO using "GEOquery" Bioconductor package and normalised using justRMA function of GEOquery.

Geneset enrichment. For single sample gene set enrichment, we used the yaGST package.²⁰

Precalculated TCGA "Immune Landscape of Cancer" scores were downloaded from the respective publication.¹³ The gene signatures used to calculate remaining enrichment scores can be found in Supplementary Table S2.⁵ ESTIMATE-immune and -stromal signatures were downloaded from the respective publication.²¹ The signatures from these sources (83 signatures/predictors) are further filtered due to missing data or outlier behaviour. A total of 76 signatures were used to benchmark against MIRACLE within 28 solid TCGA cohorts in survival and correlation analyses.

Stratification of cohorts based on immunologic constant of rejection (ICR)–survival association

Normalised enrichment scores (NESs) for ICR gene signature were used as a continuous metric to identify survival association in 28 solid TCGA cohorts individually. The cohorts with positive survival association are designated as ICR-Enabled (IE; hazard ratio (HR) < 1 with a *p* value < 0.1) and the cohorts with significant negative survival association are designated as ICR-Disabled (ID) cohorts (HR > 1 with a *p* value < 0.1)

Differential expression

To identify the genomic features correlated with ICR in IE and ID cohorts, we performed supervised differential expression analysis between samples with high enrichment of ICR genes versus samples with low enrichment of ICR, pooling together all the samples in each cohort (IE and ID). High and low enrichment levels were defined according to the first and third tertile of the NES. We used the non-parametric Wilcoxon test (R function wilcox.test) and filtered the genes with log₂ Fold Change > 1, *p* value < 0.05 and adjusted *p* value < 0.1; these genes constitute the IE and ID genesets (Supplementary Table S3).

Computation of the MIRACLE score

MIRACLE is the ratio between the single-sample enrichment scores of two gene sets, the top genes of ΔR_{IE} and ΔR_{ID} in a given sample. We used the NES of the Mann–Whitney–Wilcoxon Gene Set test (mww-GST) available in the yaGST package.²⁰ Briefly, NES is an estimate of the probability that the expression of a gene in the geneset is greater than the expression of a gene outside this set:

$$NES = 1 - \frac{U}{mn},$$

where *m* is the number of genes in a gene set, *n* is the number of those outside the gene set, $U = mn + m(m + 1) - T$ and *T* is the sum of the ranks of the genes in the geneset.

Gene ontology (GO) and network analysis

GO analyses were performed using the BiNGO application in the Cytoscape Suite version 3.7.1. GO enrichments are calculated starting from top IE genes (*n* = 605) and ID genes (*n* = 377) using the hypergeometric test. To visualise the enriched GO terms in a network graph, the Cytoscape EnrichmentMap application was used. To exclude very general or too granular GO terms, BiNGO output was filtered removing GO terms corresponding to gene sets with >600 features or <10 features. In the network images (EnrichmentMap), the size of the nodes is proportional to the geneset size while edge thickness is proportional to the relatedness of the GO terms (Jaccard similarity coefficient).

Survival analysis

To calculate survival statistics for continuous or categorical variables, coxph function from the survival R package was used. When testing the predictive value of a signature in a single cohort, univariate analysis is used. In pan-cancer analysis of multiple cohorts, "Cohort" variable is added to the model (Survival ~ Predictor + Cohort). In multivariate analysis, variables "Age", "Gender" and "Stage" are added to the model. To compare different models and additive value of each variable, analysis of variance method was used.

Model 1: Survival ~ Predictor + Cohort

Model 2: Survival ~ Predictor + Cohort + Age + Gender + Stage

Model 3: Survival ~ Cohort + Age + Gender + Stage

When testing large number (*n* > 10) of survival models (i.e. multiple cohorts for a single predictor or multiple predictors for single cohort), adjusted *p* values are used (Benjamin–Hochberg).

ICI response definitions and data sets

Any patient classified as “Progressive Disease” based on RECIST criteria was considered as a “Non-Responder”. Any patient classified as “Complete Response”, “Partial Response” or “Stable Disease” was considered as a “Responder”. Receiver operating characteristic (ROC) curves and area under the curve (AUC) values are calculated using “pROC” R package based on above response definitions and associated MIRACLE scores.

RESULTS

A score balancing immune-stimulatory and immune-suppressive microenvironments outperforms state-of-the-art immune signatures in predicting pan-cancer survival

Here we characterise samples with a highly active immune phenotype using the ICR.^{6,22–24} The ICR gene signature includes genes involved in T helper type 1 (Th1) and interferon signalling (IFNG, TBX21, CD8A/B, IL12B, STAT1 and IRF1), Th1 chemoattraction (such as the CXCR3 and CCR5 ligands, respectively, CXCL9 and CXCL10, and CCL5) and cytotoxic functions (GNLY, PRF, GZMA, GZMB, and GZMH). ICR signature also includes markers of immuno-suppressive mechanisms such as CD274/PD-L1, PDCD1, CTLA4, FOXP3 and IDO1. We developed a balance score using the approach in Fig. 1a based on ICR. We first identify cancer types from TCGA database in which a highly active immune phenotype is associated with favourable survival (HR < 1 with a p value < 0.1; Fig. S1) and cancer types in which this phenotype is associated with decreased survival (HR > 1 with a p value < 0.1; Fig. S1). We categorised these two groups of cancer types in ICR-enabled (BRCA, SKCM, UCEC, SARC, LIHC, HNSC, OV; abbreviated as “IE”) and ICR-disabled (UVM, LGG, PAAD, KIRP; abbreviated as “ID”) groups, respectively as previously described.²⁴ All other cancer types in which ICR did not show an association or trend were categorised as ICR-neutral. Supervised analysis between samples expressing high levels of the ICR signature versus samples expressing low levels of the ICR signature in enabled tumour types resulted in 1493 differentially expressed genes (DEGs) (\log_2 fold change ≥ 1 , p value ≤ 0.05 and false discovery rate < 10%). The same supervised analysis was performed in disabled tumour types obtaining 1265 DEGs associated with high levels of the ICR signature that promote worse prognosis. As expected, the two lists were highly overlapping with 888 common genes, whereas we had 605 genes positively correlated with ICR signature exclusively in the IE cohorts and 377 genes positively correlated with ICR exclusively in ID cohorts (Fig. 1a). We focussed on these last two lists and ranked genes according to their differential expression in the two supervised analyses and scored each gene according to its differential rank in the two lists (see “Methods”). In particular, for each gene of the IE-specific list we compute $\Delta R_{IE} = R_{IE} - R_{ID}$ where R_{IE} and R_{ID} are the ranks of the gene in enabled and disabled differential expression lists, respectively. The top genes re-ranked by ΔR_{IE} were retained for further analysis. In the same way, the genes in the ID-specific list were scored according to the opposite $\Delta R_{ID} = R_{ID} - R_{IE}$. Given a transcriptomics profile, the MIRACLE score is the ratio between the NESs of the whole-gene expression profiles with top genes ordered by ΔR_{IE} and ΔR_{ID} in IE- and ID-specific gene sets, respectively (Fig. 1b).

MIRACLE score was highly prognostic (HR: 0.043, confidence interval: [0.027, 0.071], p value: 6.7×10^{-36}) in the TCGA cohort (9305 cases, 28 cohorts) (Fig. 1c). The association with survival remained highly significant even if we excluded the 3127 samples used for the identification of the MIRACLE gene signatures (Table S4). When compared with 76 different immune scores recently reported in pan-cancer studies,^{5,13} MIRACLE resulted in the most significant predictor of survival across all the TCGA. We also evaluated MIRACLE at the cohort level, and it was significantly and positively associated with survival in 15 cohorts out of 28 (Fig. S2). Being the largest data set in the ID cohorts, LGG

contributes to ID-specific geneset the most and has the highest median value of MIRACLE compared to other cohorts (Fig. S3)

In addition to MIRACLE, 25 other immune-related metrics were significantly found to be associated with better survival (Fig. 1c, Left). However, the significance level of MIRACLE is several orders of magnitude higher than these predictors. On the other hand, 30 other metrics such as “Proliferation”, “Wound Healing”, “Th2 Cells”, “TGFB.Resp” and “Epithelial-to-Mesenchymal Transition” and others were found to be significantly associated with worse survival (Fig. 1c, right).

In order to evaluate whether MIRACLE is an independent survival predictor, we computed likelihood ratio test between different multivariate models including clinical attributes, such as Age, Gender and Stage (see “Methods”). When the survival model, including MIRACLE and cohort as variable (Model 1), was compared to the model including Age, Gender, Stage (Model 2), the clinical attributes did not increase the prognostic value of the model (p value: 0.422). On the other hand, MIRACLE significantly improved the significance when we compared Model 2 with the model containing clinical variables only (Model 3) with a p value of 1.2×10^{-35} .

Using PanCancer subtypes, we observed that MIRACLE and survival association is improved in TCGA-BRCA and TCGA-LUAD data sets when the subtypes are included in the survival models. PAM50 classification for Breast Cancer²⁵ improved MIRACLE-survival association p value from 5.15×10^{-3} to 2.04×10^{-4} and iCluster classifications for Lung Cancer²⁶ improved the survival prediction p value from 9.37×10^{-3} to 8.54×10^{-4} . We also observed the effect of PAM50 classification on MIRACLE-survival association in Breast Cancer in METABRIC²⁷ and GEO data sets,²⁸ in addition to TCGA (data not shown).

We correlated the MIRACLE scores for TCGA cases profiled with both RNAseq and Microarray platforms and found that MIRACLE RNAseq values correlated highly with MIRACLE Microarray values (Affy U133A Platform) (TCGA-LUSC: $n = 132$ patients, $r = 0.76$, p value = 1.7×10^{-26} ; TCGA-OV: $n = 367$, $r = 0.61$, p value = 2.2×10^{-39} ; TCGA-GBM: $n = 157$, $r = 0.7$, p value = 3.2×10^{-26}), showing independence of MIRACLE score from the gene expression platforms used.

In order to validate MIRACLE’s prognostic value in independent data sets, we collected 55 studies with a total of 7623 patients from GEO. As study selection criteria, we included the microarray platforms with MIRACLE compatible feature IDs (i.e. Affymetrix, Illumina, Entrez IDs or Hugo Gene Symbols as explained in the “Methods” section) but excluded haematological malignancies and cohorts identified as outliers in TCGA analysis (Germ Cell Tumours, Thymomas). The majority (44/55) of the studies used here are also present in the PRECOG database.¹⁴ In a Cox proportional hazards survival analysis, MIRACLE was strongly associated with better clinical outcome compared to other predictors such as ICR, Estimate²¹ and TIS,²⁹ with an HR of 0.1985 and p value of 2.73×10^{-38} (Fig. S4). Overall, the above results validate and show that MIRACLE outperforms state-of-the-art immune signatures in predicting pan-cancer survival.

MIRACLE score decouples inflammation from stromal score/TGF β signalling pathways and other immuno-suppressive signatures. Since MIRACLE was derived from the DEGs induced by the ICR signature, we analysed the difference between these two scores by evaluating how they correlate with a set of immune-related indices (Fig. 2).^{5,13} As expected, ICR (Fig. 2a) highly correlates with other signatures of T cell infiltration such as Tcell,³⁰ TIS,²⁹ NFkB³¹ and Checkpoint,^{32,33} as well as ESTIMATE²¹ and Lymph.Inflit.Sig.Sc,^{5,13} across all solid tumours (with a mean correlation coefficient of 0.85 across all mentioned signatures in 28 cohorts). In addition, ICR moderately but consistently correlates with immuno-suppressive signatures such as TGF β Response,¹⁵ Cancer-Associated Fibroblasts (CAF),^{34,35} ESTIMATE-Stroma,²¹ Epithelial-Mesenchymal Transition (EMT),^{36,37} PI3KY^{38–40} and LAP-Helios (with a mean correlation

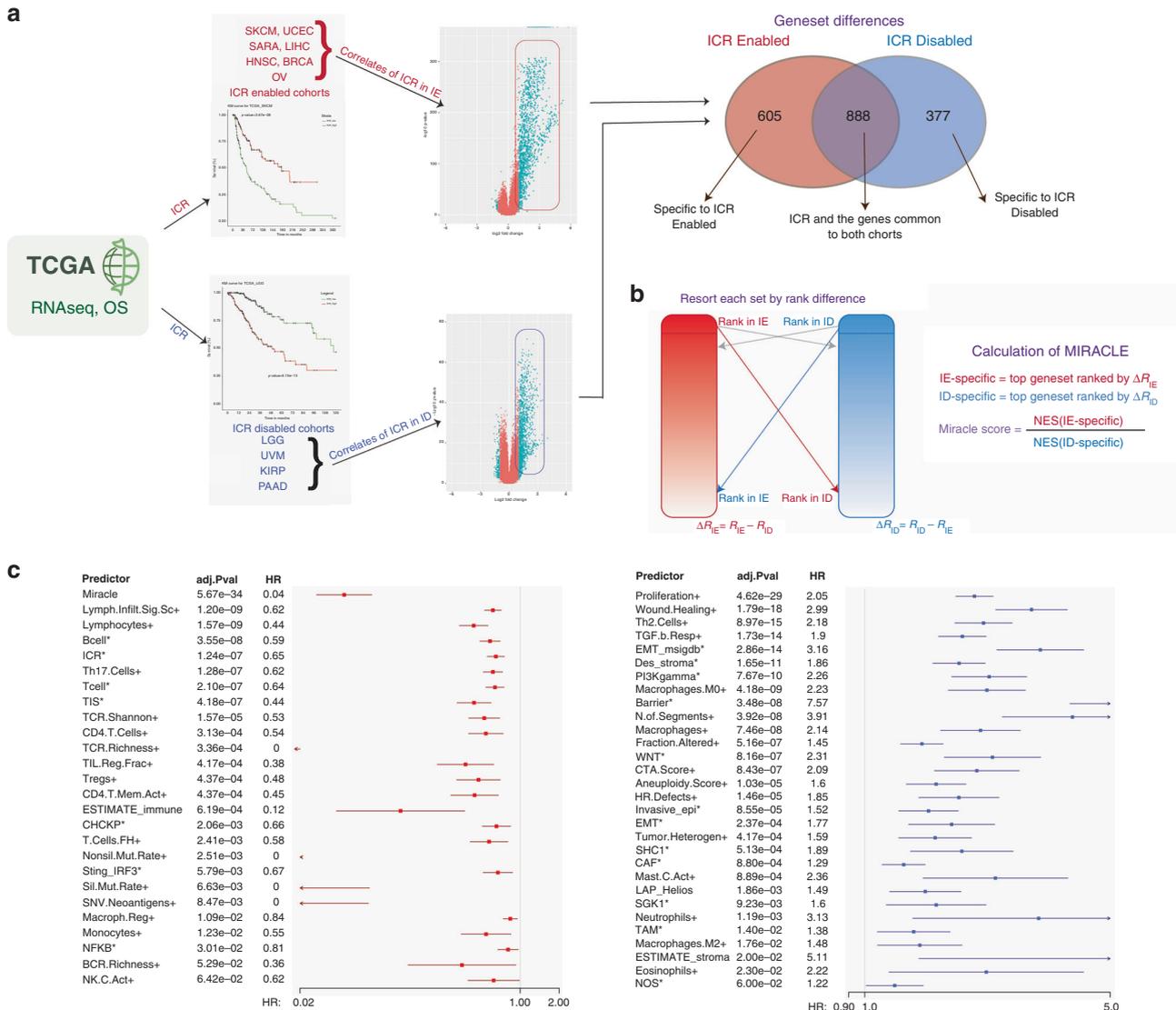


Fig. 1 Development of balance score between immune stimulatory and suppressive microenvironments. **a** The rationale and step-by-step workflow of generation of MIRACLE genesets. Venn diagram comparisons show the significant sets in ICR enabled (IE) cohorts (the cohorts where ICR is significantly associated with good prognosis) and ICR disabled (ID) cohorts (the cohorts where ICR is significantly associated with worse prognosis). The genesets are determined by the differential expression between high and low ICR tertiles (log FoldChange >1, p value < 0.05 and FDR < 10%). **b** Differentially expressed gene matrices (initially sorted by p value) are represented by red and blue rectangles. The genes are then re-ordered using the rank difference in the two lists. The features missed due to hard cut-offs (grey arrows) become lower in rank and are filtered out. IE-specific (red arrows) and ID-specific (blue arrows) genesets are the top genes after ΔR reordering. MIRACLE score is the ratio between enrichment scores of the two lists (MIRACLE is short for Mediators of Immune Response Against Cancer in soLid microEnvironments). **c** MIRACLE is compared to a variety of immune-related signatures to predict prognosis in TCGA ($n = 76$ signatures, only significant associations are shown). Hazard ratio for death forest plots represent top performing positive (left) and negative predictors (right) of survival.

coefficient of 0.31).⁴¹ Intriguingly, MIRACLE (Fig. 2b) score is either not correlated or anti-correlated with these inhibitory signatures (mean correlation coefficient: -0.1 , Table S5), while it is highly correlated with all the hallmarks of Th1-polarised T cell infiltration (Fig. 2b, Table S5). MIRACLE score correlates with T cell infiltration signatures in all cancer types except for Glioma (LGG and GBM) and Kidney renal papillary cell carcinoma (KIRP). The inverse correlation with Glioma depends on the fact that the immuno-suppressive mechanisms driven by the high presence of myeloid cells,^{42,43} especially in the IDH wild-type tumours having a worse survival,⁴⁴ tend to enrich the denominator of the MIRACLE. Similarly, a large proportion (about 42%) of KIRP tumours is characterised by the presence of high immune infiltration concomitant with a suppressive microenvironment.⁴⁵

To better clarify the functional role of the signatures that define the MIRACLE score, we performed enrichment analysis of GO categories of the list of genes specific of ICR-Enabled tumour types (605 genes) and the list specific of ICR-Disabled list (377 genes) (Fig. 3). We found that the ICR-Enabled list was associated with "T cell activation", "Regulation of Lymphocyte Mediated Immunity" and "Adaptive Immune Response", as well as "Positive Regulation of Cell Killing", among others. On the other hand, the ICR-Disabled list was associated with GO terms including "Negative Regulation of Cell Death", "TGF β Signalling Pathway" and "Epithelial-to-Mesenchymal Transition" (EMT), which are involved in tumour progression with metastatic expansion, and the generation of tumour cells with stem cell properties that play a major role in resistance to cancer treatment.^{46,47} In the self-organising network

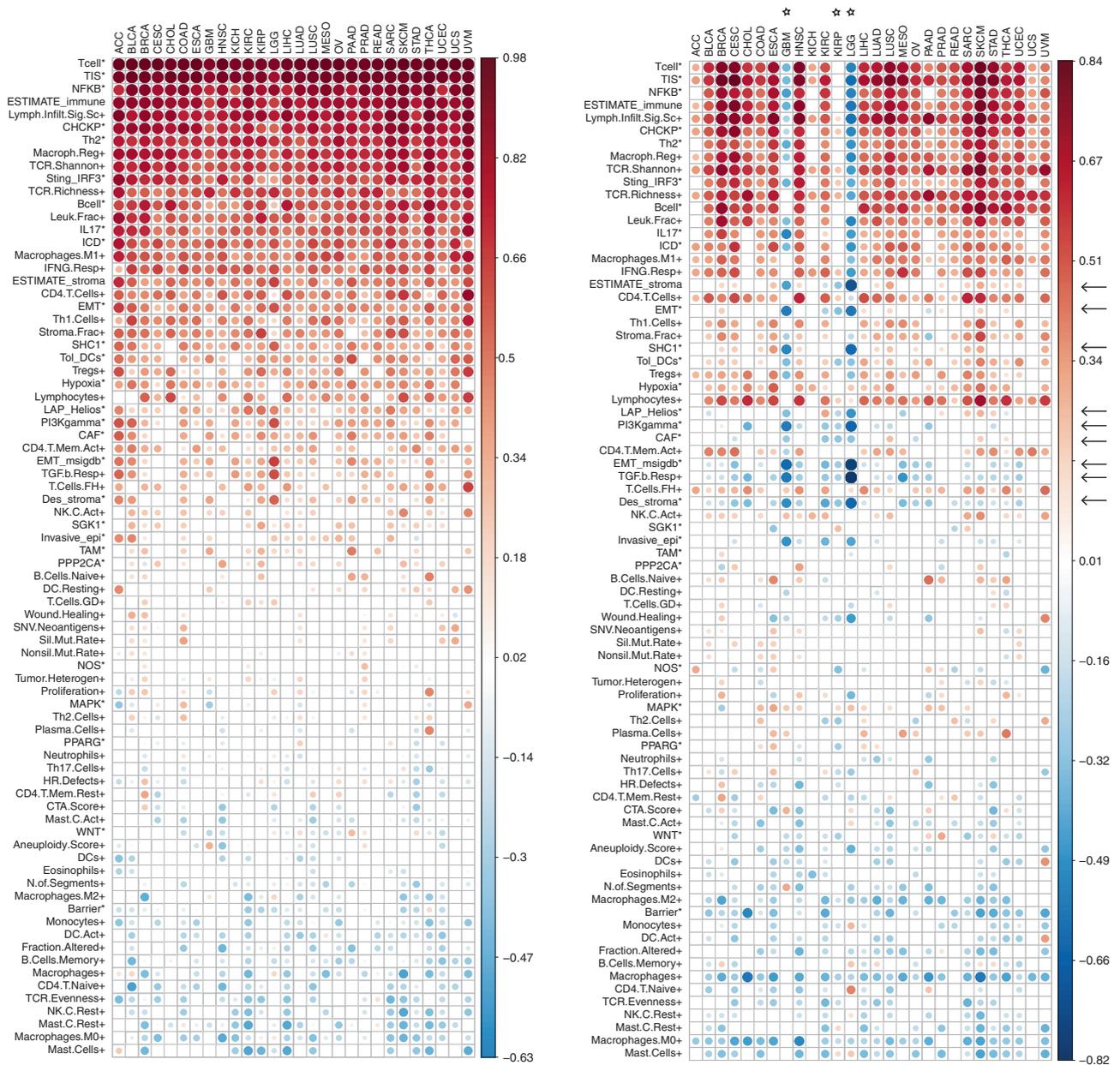


Fig. 2 Immune pathways correlate to MIRACLE score. Correlation of ICR enrichment score (left) and MIRACLE score (right) with a set ($n = 76$) of benchmark signatures and metrics (rows) within 28 TCGA cohorts (columns). The colours reflect Pearson correlation coefficients (empty when correlation p value is not significant). The rows of both panels are sorted by average correlation coefficients across 28 cohorts of the left (ICR) heatmap. The arrows indicate signatures for which correlation with MIRACLE score deviates from correlation with ICR. Similarly, stars indicate the cohorts where results between MIRACLE and ICR are most divergent.

plot of these GO terms, IE and ID signatures reflect two opposing networks, clustering separately (Fig. 3). Interestingly the GO terms enriched in both IE and ID lists include “Chemotaxis” and “Taxis”, which are functions that result in the presence of immune infiltrate (hence high ICR in both IE and ID cohorts) but do not reflect activatory or suppressive mediators of immune response against cancer. These results suggest that, within a given sample, MIRACLE score measures the existing balance between two opposing immune-related phenotypes.

Genomics events associated with MIRACLE score

We then explored the association of specific oncogenic mutations with MIRACLE score. We first selected a set of 470 frequently mutated genes in cancer,⁴⁸ then trained an elastic net model to

predict our score as a function of mutations in each sample and using the cohort as covariate.⁴⁹ The positive coefficients of the trained model were used to identify genes whose mutations are associated with an increase of the MIRACLE score and negative coefficients identify the genes whose mutations are associated with a decrease of the MIRACLE score (Fig. 4). We evaluated the accuracy of the model in a tenfold cross validation computing the correlation between the model prediction and the true MIRACLE scores obtaining a Spearman correlation of 0.78 ± 0.017 (p value ~ 0). Genes associated with a decrease of MIRACLE score include: EGFR, PRKAR1A, ZNF814, PTEN, MAP3K1, CHEK2, FRG1, KRAS, MAP4K1, NF1, PIK3CA, and others. Interestingly, several of these genes, such as EGFR,⁵⁰ have previously been associated with tumour-intrinsic factors modulating cancer immunity activating

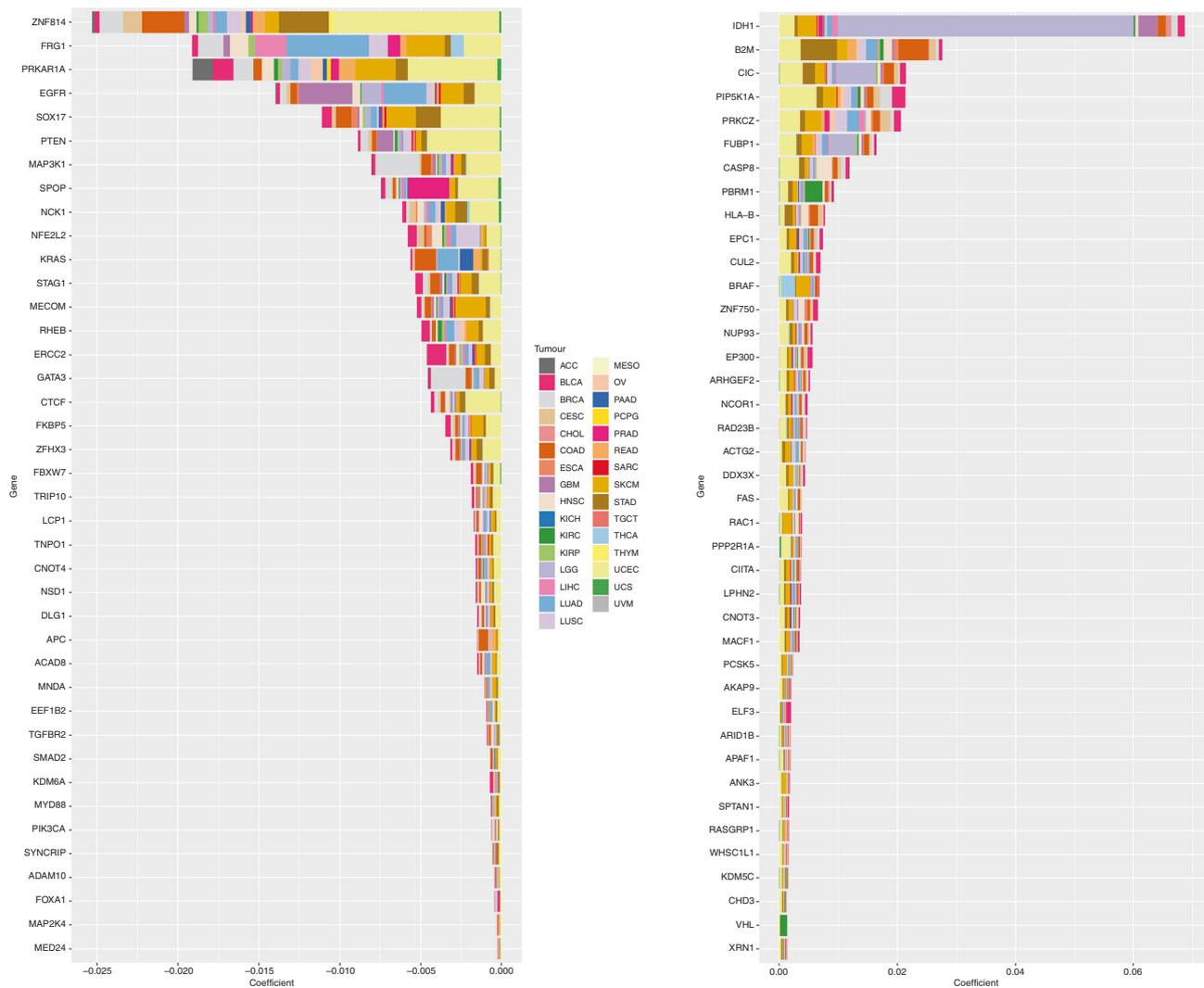


Fig. 4 Genomic events positively and negatively associated with MIRACLE. Top 40 of the mutated genes with negative non-zero coefficients of a trained elastic net model. One hundred and eighty-six identified genes whose mutation is associated with a decrease of the MIRACLE (left panel). Top 40 mutated genes with a positive association with MIRACLE score in pan-cancer trained model (right panel). Contributions of each individual cancer type to the coefficient in trained elastic net model are proportionally indicated by the size of the bars.

from anti-PD10-treated patients, where MIRACLE was associated with response (p value: 0.042, ref. 58). Overall, considering the various stratifications of the publicly available response data sets, MIRACLE, ICR and TIS were predictive with the most number of data sets when compared to other immune-related signatures (Fig. S5).

DISCUSSION

Here we have shown that a balance score encapsulating both immune-active and immune-repressive cancer phenotype defines a continuum of clinically and biologically relevant classes of tumours. We derived our score by evaluating the differences between the transcriptional programmes associated with the presence of a pro-survival Th1-polarised immune infiltration and the transcriptional programmes activated in tumour types where the immune microenvironment has a negative effect on prognosis. Our score displayed the highest positive association with survival when compared with several immune phenotyping scores reported recently in large scale studies.¹³ and was significantly associated with better survival in 15 out 28 TCGA

indications; its prognostic features were confirmed in 55 data sets spanning 11 tumour types and 7623 cases. It decouples the typical immune-regulatory mechanisms that often are simultaneously active in T cell-inflamed microenvironments. This explains why it always results in a positive HR. Indeed, immune homeostasis is a subtle balance between the immune defence against foreign pathogens and suppression of the immune system to maintain self-tolerance and prevent autoimmune disease.⁵⁹

We interrogated the association between the tumour composition and the increase or decrease of our score. Interestingly, by using an elastic net regression approach we found known and novel alterations consistently associated with a desert immune phenotype. The top genes whose mutations are associated with a decrease of MIRACLE were EGFR, PRKAR1A, ZNF814, PTEN, TRIP10, MECOM, MAP3K1, MYD88, CNOT4 and NCK1. The association of EGFR with an immune-desert tumour phenotype with poor response to checkpoint inhibitor therapies has been observed in several clinical trials.^{60,61} Mutations of PRKAR1A have been recently demonstrated to drive a drastic remodelling of the tumour microenvironment. Loss of PTEN expression has been shown to prime the microenvironment for tumour development

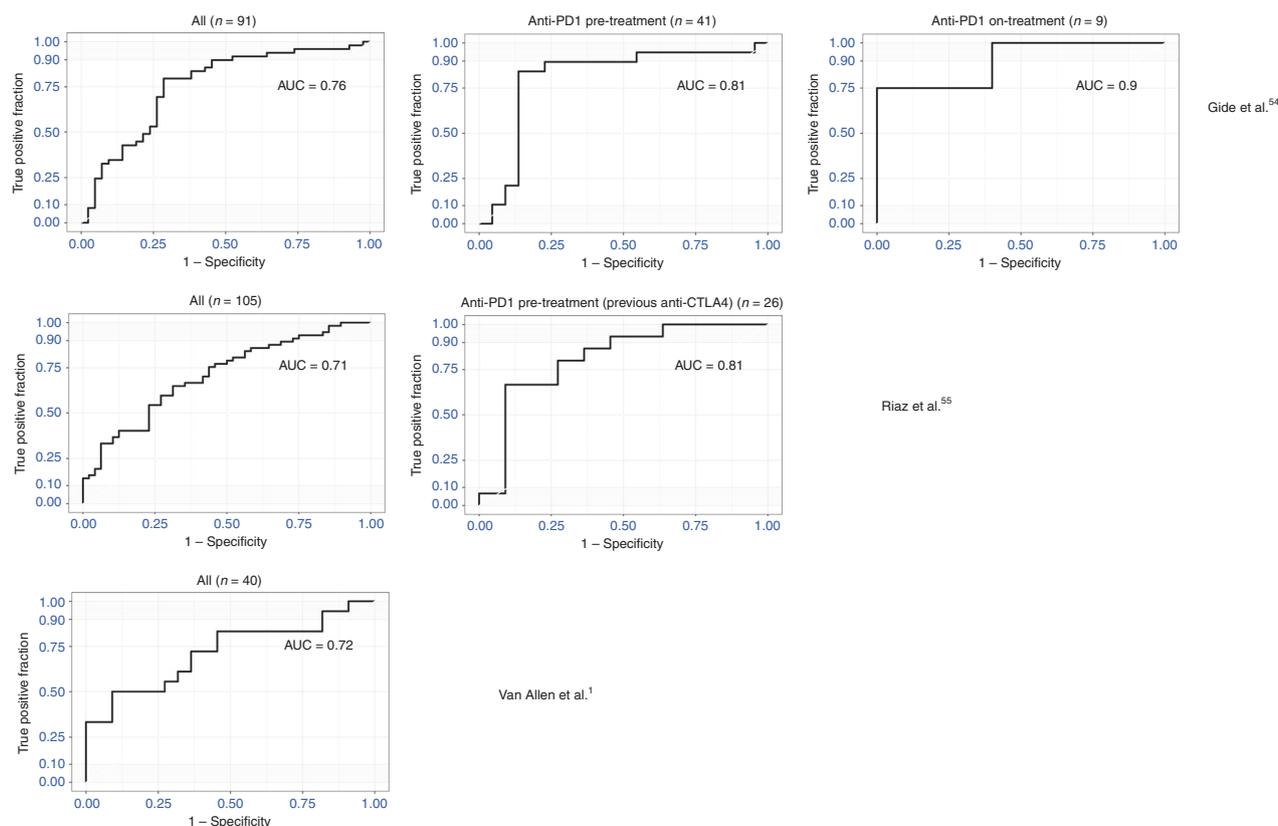


Fig. 5 MIRACLE predicts ICI response. ROC curves showing predictive value of MIRACLE in four melanoma immunotherapy response data sets. Responders are defined as complete response, partial response, or stable disease, and non-responders are defined as patients with progressive disease. Top: Gide et al., all samples (left), anti-PD1 pre-treatment (middle) and anti-PD1 on-treatment samples (right).⁵³ Middle: Riaz et al. all samples (left) and samples from patients previously treated with anti-CTLA4.⁵⁴ Bottom: Van Allen et al., data set. The patient counts are indicated for each cohort in parentheses (*n*).

through several mechanisms, such as modification of the pattern of cytokine secretion creating an immune-suppressive microenvironment with increase of immune cell populations that can promote tumour progression.⁶² Enrichment of PTEN mutations driving an immuno-suppressive microenvironment with higher expression of M2 macrophages have been observed in non-responder to checkpoint inhibitor therapies on glioma.⁶³ Mutations of MAP3K1 are associated with immune-desert phenotype in breast.⁶ The reported association could represent novel tumour-intrinsic mechanisms deserving further experimental studies.

We have also reported that MIRACLE score is a predictive biomarker of response to ICIs in public data sets. Several transcriptional-based signatures have been recently proposed and validated in clinical trials⁶⁴ or retrospectively.^{16,17} One of the main differences between the proposed score and other similar approaches^{16,17} is that MIRACLE score does not need any data set-dependent normalisation and can be independently and incrementally applied patient by patient as in a typical clinical trial scenario to stratify patients who will be more likely to benefit from PD-1 or CTLA-4 checkpoint immunotherapies. Several limitations and possible extensions of the analysis reported in this paper should also be mentioned. First, our study was primarily focussed on gene expression biomarkers. Other data types such as protein levels and somatic mutations could eventually improve the ability of our method to predict ICI response, and therefore one of the possible extensions of the method to multi-omics variables will be considered in the follow-up. Another limitation that could eventually lead to further improvements is that the signatures that define the MIRACLE score were selected just considering the

effect of immune system polarisation on survival. As more and more response data sets are being made available from clinical trials, these data could be retrospectively used for refining the immune signatures and designing better response predictors.

To enable testing of MIRACLE by clinicians and the public, we created a web application for response prediction using transcriptome profiles at https://miracle.shinyapps.io/miracle_shinyapp/. MIRACLE can potentially help physicians to predict survival and ICI response for defining patient selection strategies.

In conclusion, MIRACLE features have been derived in a genome-wide and pan-cancer manner and not been limited to any prior knowledge of gene features or cancer types other than measurable immune phenotypes and patient outcomes. Thus MIRACLE can be applied to any bulk RNA measurement to stratify patients to predict outcomes and checkpoint inhibitor therapy responses.

ACKNOWLEDGEMENTS

We are grateful to Zoltan Dezso, PhD for his input on Transcription Factor Networks regulating MIRACLE-high and MIRACLE-low cancers as well as to William Randall Henner, MSc for his comments on statistical comparison of survival models.

AUTHOR CONTRIBUTIONS

Conceived and supervised the study: T.T., J.S., M.C. Developed the method: T.T., M.C. Analysed and discussed the data: S.K., K.H., F.M.M., T.J.H., H.J.J. Contributed with statistics and code: J.R., W.H., D.B. Wrote the manuscript: T.T., J.S., D.B., M.C.

ADDITIONAL INFORMATION

Ethics approval and consent to participate All data were from other publications. The ethical approvals obtained for each data set are listed in the original publications.

Consent for publication Not applicable.

Data availability Raw and source data availability: all raw data used in the study (weblinks, references, cohort sizes and other meta-data attributes) are listed in Table S1.

Code availability The source code and installation instructions for MIRACLE R package can be found publicly at GitHub (<https://github.com/tolgaturan-github/Miracle>). MIRACLE score can be calculated for a normalised RNAseq or Microarray gene expression matrix allowing the use of various platform identifiers including Ensembl Gene Ids (preferred), U133p2 and U133A from Affymetrix, IlluminaHTv3 and IlluminaHTv4 from Illumina and other expression data sets with features mapped to Entrez IDs or Hugo Gene Symbols. A shinyapp is also available at the following address https://miracle.shinyapps.io/miracle_shinyapp/, to enable the use of MIRACLE by non-bioinformaticians.

Competing interests T.T., S.K., K.H., T.J.H., H.J.J. and J.S. are all employees of AbbVie. F.M.M. was an employee of AbbVie at the time of the study. M.C. was an employee of AbbVie at the time of the study and at present continues to have consulting relation with AbbVie. J.R., W.H. and D.B. have no current or prior affiliation with AbbVie beyond the current study. The design, study conduct and financial support for this research were provided by AbbVie. AbbVie participated in the interpretation of data, review and approval of the publication.

Funding information The research was funded by AbbVie, Inc. The research leading to these results has also received funding from AIRC under IG 2018 - ID. 21846 project —P.I. Ceccarelli Michele and from Italian Ministry of Research Grant PRIN 2017XJ38A4_004.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41416-020-01145-4>.

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

REFERENCES

1. Van Allen, E. M., Miao, D., Schilling, B., Shukla, S. A., Blank, C., Zimmer, L. et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* **350**, 207–211 (2015).
2. Snyder, A., Makarov, V., Merghoub, T., Yuan, J., Zaretsky, J. M., Desrichard, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med.* **371**, 2189–2199 (2014).
3. Le, D. T., Uram, J. N., Wang, H., Bartlett, B. R., Kemberling, H., Eyring, A. D. et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
4. Spranger, S., Spaepen, R. M., Zha, Y., Williams, J., Meng, Y., Ha, T. T. et al. Upregulation of PD-L1, IDO, and T(regs) in the melanoma tumor microenvironment is driven by CD8⁺ T cells. *Sci. Transl. Med.* **5**, 200ra116 (2013).
5. Turan, T., Kannan, D., Patel, M., Matthew Barnes, J., Tanlimco, S. G., Lu, R. et al. Immune oncology, immune responsiveness and the theory of everything. *J. Immunother. Cancer* **6**, 50 (2018).
6. Hendrickx, W., Simeone, I., Anjum, S., Mokraby, Y., Bertucci, F., Finetti, P. et al. Identification of genetic determinants of breast cancer immune phenotypes by integrative genome-scale analysis. *Oncimmunology* **6**, e1253654 (2017).
7. Bedognetti, D., Hendrickx, W., Ceccarelli, M., Miller, L. D. & Seliger, B. Disentangling the relationship between tumor genetic programs and immune responsiveness. *Curr. Opin. Immunol.* **39**, 150–158 (2016).
8. Snyder, A., Wolchok, J. D. & Chan, T. A. Genetic basis for clinical response to CTLA-4 blockade. *N. Engl. J. Med.* **372**, 783 (2015).
9. Ganesh, S., Shui, X., Craig, K. P., Park, J., Wang, W., Brown, B. D. et al. RNAi-mediated β -catenin inhibition promotes T cell infiltration and antitumor activity in combination with immune checkpoint blockade. *Mol. Ther.* **26**, 2567–2579 (2018).
10. Peng, W., Chen, J. Q., Liu, C., Malu, S., Creasy, C., Tetzlaff, M. T. et al. Loss of PTEN promotes resistance to T cell-mediated immunotherapy. *Cancer Discov.* **6**, 202–216 (2016).

11. Spranger, S., Bao, R. & Gajewski, T. F. Melanoma-intrinsic β -catenin signalling prevents anti-tumour immunity. *Nature* **523**, 231–235 (2015).
12. Dushyanthen, S., Teo, Z. L., Caramia, F., Savas, P., Mintoff, C. P., Virasamy, B. et al. Agonist immunotherapy restores T cell function following MEK inhibition improving efficacy in breast cancer. *Nat. Commun.* **8**, 606 (2017).
13. Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T.-H. et al. The immune landscape of cancer. *Immunity* **48**, 812.e14–830.e14 (2018).
14. Gentles, A. J., Newman, A. M., Liu, C. L., Bratman, S. V., Feng, W., Kim, D. et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.* **21**, 938–945 (2015).
15. Mariathasan, S., Turley, S. J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y. et al. TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* **554**, 544–548 (2018).
16. Auslander, N., Zhang, G., Lee, J. S., Frederick, D. T., Miao, B., Moll, T. et al. Robust prediction of response to immune checkpoint blockade therapy in metastatic melanoma. *Nat. Med.* **24**, 1545–1549 (2018).
17. Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X. et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **24**, 1550–1558 (2018).
18. Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D. et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71 (2016).
19. Liu, J., Lichtenberg, T., Hoadley, K. A., Poisson, L. M., Lazar, A. J., Cherniack, A. D. et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* **173**, 400.e11–416.e11 (2018).
20. Frattini, V., Pagnotta, S. M., Tala, F., Fan, J. J., Russo, M. V., Lee, S. B. et al. A metabolic function of FGFR3-TACC3 gene fusions in cancer. *Nature* **553**, 222–227 (2018).
21. Yoshihara, K., Shahmoradgol, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W. et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013).
22. Galon, J., Angell, H. K., Bedognetti, D. & Marincola, F. M. The continuum of cancer immunosurveillance: prognostic, predictive, and mechanistic signatures. *Immunity* **39**, 11–26 (2013).
23. Bertucci, F., Finetti, P., Simeone, I., Hendrickx, W., Wang, E., Marincola, F. M. et al. The immunologic constant of rejection classification refines the prognostic value of conventional prognostic signatures in breast cancer. *Br. J. Cancer* **119**, 1383–1391 (2018).
24. Roelands, J., Hendrickx, W., Zoppoli, G., Mall, R., Saad, M., Halliwill, K. et al. Oncogenic states dictate the prognostic and predictive connotations of intratumoral immune response. *J. Immunother. Cancer* **8**, e000617 (2020).
25. Parker, J. S., Mullins, M., Cheang, M. C. U., Leung, S., Voduc, D., Vickery, T. et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
26. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **511**, 543–550 (2014).
27. Curtis, C., Shah, S. P., Chin, S.-F., Turashvili, G., Rueda, O. M., Dunning, M. J. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
28. Nagalla, S., Chou, J. W., Willingham, M. C., Ruiz, J., Vaughn, J. P., Dubey, P. et al. Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. *Genome Biol.* **14**, R34 (2013).
29. Danaher, P., Warren, S., Lu, R., Samayoa, J., Sullivan, A., Pekker, I. et al. Pan-cancer adaptive immune resistance as defined by the tumor inflammation signature (TIS): results from The Cancer Genome Atlas (TCGA). *J. Immunother. Cancer* **6**, 63 (2018).
30. Saleh, S. M. I., Bertos, N., Gruosso, T., Gigoux, M., Souleimanova, M., Zhao, H. et al. Identification of interacting stromal axes in triple-negative breast cancer. *Cancer Res.* **77**, 4673–4683 (2017).
31. Hopewell, E. L., Zhao, W., Fulp, W. J., Bronk, C. C., Lopez, A. S., Massengill, M. et al. Lung tumor NF- κ B signaling promotes T cell-mediated immune surveillance. *J. Clin. Invest.* **123**, 2509–2522 (2013).
32. Benci, J. L., Xu, B., Qiu, Y., Wu, T. J., Dada, H., Twyman-Saint Victor, C. et al. Tumor interferon signaling regulates a multigenic resistance program to immune checkpoint blockade. *Cell* **167**, 1540.e12–1554.e12 (2016).
33. Koyama, S., Akbay, E. A., Li, Y. Y., Herter-Sprie, G. S., Buczkowski, K. A., Richards, W. G. et al. Adaptive resistance to therapeutic PD-1 blockade is associated with upregulation of alternative immune checkpoints. *Nat. Commun.* **7**, 10501 (2016).
34. Kraman, M., Bambrough, P. J., Arnold, J. N., Roberts, E. W., Magiera, L., Jones, J. O. et al. Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein- α . *Science* **330**, 827–830 (2010).
35. Öhlund, D., Handly-Santana, A., Biffi, G., Elyada, E., Almeida, A. S., Ponz-Sarvisse, M. et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **214**, 579–596 (2017).

36. Herbst, R. S., Soria, J.-C., Kowanetz, M., Fine, G. D., Hamid, O., Gordon, M. S. et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. *Nature* **515**, 563–567 (2014).
37. Shields, B. D., Mahmoud, F., Taylor, E. M., Byrum, S. D., Sengupta, D., Koss, B. et al. Indicators of responsiveness to immune checkpoint inhibitors. *Sci. Rep.* **7**, 807 (2017).
38. Karlsson, E., Veenstra, C., Emin, S., Dutta, C., Pérez-Tenorio, G., Nordenskjöld, B. et al. Loss of protein tyrosine phosphatase, non-receptor type 2 is associated with activation of AKT and tamoxifen resistance in breast cancer. *Breast Cancer Res. Treat.* **153**, 31–40 (2015).
39. Daragmeh, J., Barriah, W., Saad, B. & Zaid, H. Analysis of PI3K pathway components in human cancers. *Oncol. Lett.* **11**, 2913–2918 (2016).
40. De Henau, O., Rausch, M., Winkler, D., Campesato, L. F., Liu, C., Cymerman, D. H. et al. Overcoming resistance to checkpoint blockade therapy by targeting PI3K γ in myeloid cells. *Nature* **539**, 443–447 (2016).
41. Abd Al Samid, M., Chaudhary, B., Khaled, Y. S., Ammori, B. J. & Elkord, E. Combining FoxP3 and Helios with GARP/LAP markers can identify expanded Treg subsets in cancer patients. *Oncotarget* **7**, 14083–14094 (2016).
42. Klemm, F., Maas, R. R., Bowman, R. L., Kornete, M., Soukup, K., Nassiri, S. et al. Interrogation of the microenvironmental landscape in brain tumors reveals disease-specific alterations of immune cells. *Cell* **181**, 1643.e17–1660.e17 (2020).
43. Caruso, F., Garofano, L., D'Angelo, F., Yu, K., Tang, F., Yuan, J. et al. A MAP of tumor-host interactions in glioma at single cell resolution. *Gigascience* **9**, 1–14 (2020).
44. Ceccarelli, M., Barthel, F. P., Malta, T. M., Sabedot, T. S., Salama, S. R., Murray, B. A. et al. Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma. *Cell* **164**, 550–563 (2016).
45. Zhang, J., Caruso, F. P., Sa, J. K., Justesen, S., Nam, D.-H., Sims, P. et al. The combination of neoantigen quality and T lymphocyte infiltrates identifies glioblastomas with the longest survival. *Commun. Biol.* **2**, 135 (2019).
46. Lambert, A. W., Pattabiraman, D. R. & Weinberg, R. A. Emerging biological principles of metastasis. *Cell* **168**, 670–691 (2017).
47. Moustakas, A. & de Herreros, A. G. Epithelial-mesenchymal transition in cancer. *Mol. Oncol.* **11**, 715–717 (2017).
48. Iorio, F., Knijnenburg, T. A., Vis, D. J., Bignell, G. R., Menden, M. P., Schubert, M. et al. A landscape of pharmacogenomic interactions in cancer. *Cell* **166**, 740–754 (2016).
49. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. B* **67**, 301–320 (2005).
50. Hastings, K., Yu, H. A., Wei, W., Sanchez-Vega, F., DeVeaux, M., Choi, J. et al. EGFR mutation subtypes and response to immune checkpoint blockade treatment in non-small-cell lung cancer. *Ann. Oncol.* **30**, 1311–1320 (2019).
51. Codina, A., Renauer, P. A., Wang, G., Chow, R. D., Park, J. J., Ye, H. et al. Convergent identification and interrogation of tumor-intrinsic factors that modulate cancer immunity in vivo. *Cell Syst.* **8**, 136.e7–151.e7 (2019).
52. 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).
53. Gide, T. N., Quek, C., Menzies, A. M., Tasker, A. T., Shang, P., Holst, J. 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.e6 (2019).
54. Riaz, N., Havel, J. J., Makarov, V., Desrichard, A., Urba, W. J., Sims, J. S. et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* **171**, 934–949.e16 (2017).
55. Hugo, W., Zaretsky, J. M., Sun, L., Song, C., Moreno, B. H., Hu-Lieskovan, S. et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* **168**, 542 (2017).
56. Fridman, W. H., Pagès, F., Sautès-Fridman, C. & Galon, J. The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer* **12**, 298–306 (2012).
57. Ulloa-Montoya, F., Louahed, J., Dizier, B., Gruselle, O., Spiessens, B., Lehmann, F. F. et al. Predictive gene signature in MAGE-A3 antigen-specific cancer immunotherapy. *J. Clin. Oncol.* **31**, 2388–2395 (2013).
58. Kim, K., Park, S., Park, S. Y., Kim, G., Park, S. M., Cho, J.-W. et al. Single-cell transcriptome analysis reveals TOX as a promoting factor for T cell exhaustion and a predictor for anti-PD-1 responses in human cancer. *Genome Med.* **12**, 22 (2020).
59. Wrzesinski, S. H., Wan, Y. Y. & Flavell, R. A. Transforming growth factor-beta and the immune response: implications for anticancer therapy. *Clin. Cancer Res.* **13**, 5262–5270 (2007).
60. Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., Ready, N. E. et al. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* **373**, 1627–1639 (2015).
61. Herbst, R. S., Baas, P., Kim, D.-W., Felip, E., Pérez-Gracia, J. L., Han, J.-Y. et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet* **387**, 1540–1550 (2016).
62. Cretella, D., Digiacomo, G., Giovannetti, E. & Cavazzoni, A. PTEN alterations as a potential mechanism for tumor cell escape from PD-1/PD-L1 inhibition. *Cancers* **11**, 1318 (2019).
63. Zhao, J., Chen, A. X., Gartrell, R. D., Silverman, A. M., Aparicio, L., Chu, T. et al. Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma. *Nat. Med.* **25**, 462–469 (2019).
64. Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D. R. et al. IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Investig.* **127**, 2930–2940 (2017).



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/>.