Conditional activation of immune-related signatures and prognostic significance: a pan-cancer analysis

Jessica Roelands[1,2*], Wouter Hendrickx[1,3*#], Raghvendra Mall[4], Mohamad Saad[4], Kyle Halliwill[5], Gabriele Zoppoli[6], Giuseppe Curigliano[7,8], Darawan Rinchai[1], Julie Decock[3,9], Lucia G Delogu[10], Lotfi Chouchane[11], Ena Wang[12], Peter Kuppen[2], Pascal Finetti[13], Francois Bertucci[14], Lance D Miller[15], Jerome Galon[16], Francesco M Marincola[17], Michele Ceccarelli[5#], Davide Bedognetti[1,3#]

1. Department of Immunology, Inflammation and Metabolism, Division of Translational Medicine, Research Branch, Sidra Medicine, Doha, Qatar
2. Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands
3. College of Health and Life Sciences (CHLS), Hamad Bin Khalifa University (HBKU), Doha, Qatar.
4. Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar
5. Computational Biology, Computational Oncology and Immunology (CIAO), AbbVie Biotherapeutics, Inc, Redwood City, California, United States of America
6. Department of Internal Medicine, University of Genoa and Ospedale Policlinico San Martino, Genoa, Italy
7. University of Milano, Department of Oncology and Hemato-Oncology
8. European Institute of Oncology, IRCCS, Milano, Italy
9. Cancer research center, Qatar Biomedical Research Institute (QBRI), Qatar Foundation, Doha, Qatar.
10. Istituto di Ricerca Pediatrica, Fondazione Città della Speranza, Padua, Italy
11. Laboratory of Genetic Medicine and Immunology, Weill Cornell Medicine-Qatar, Qatar
12. Allogene Therapeutics, South San Francisco, California
13. Equipe Oncologie Prédictive, Centre de Recherche en Cancérologie de Marseille (CRCM), Institut Paoli-Calmettes, INSERM UMR1068, CNRS UMR725, Marseille, France.
14. Predictive Oncology laboratory, Marseille Cancer Research Center (CRCM), U1068 INSERM, U7258 CNRS, Institut Paoli-Calmettes, Marseille, France.
15. Wake Forest Comprehensive Cancer Center, Winston-Salem, NC, USA.
16. INSERM, Laboratory of Integrative Cancer Immunology, Sorbonne Université, Sorbonne Paris Cité, Université Paris Descartes, Université Paris Diderot; Centre de Recherche des Cordeliers, 75006 Paris, France.

17. Refuge Biotechnologies, Menlo Park, California, 94025

[#] Corresponding authors

Davide Bedognetti: dbdognetti@sidra.org

Wouter Hendrickx: whendrickx@sidra.org

Michele Ceccarelli: mceccarelli@abbvie.com

* Equal contribution

Abstract

Background

It is becoming clear that tumor immune T cell infiltration and its functional orientation have substantial effect on cancer progression, influencing both response to therapy and prognosis. In this pan-cancer study, the previously described Immunologic Constant of Rejection (ICR) signature is used to define opposing immune phenotypes (i.e., immune-active and immune-silent) across 31 different histologies. We systematically analyze the interconnections between the genetic programming of neoplasms and their immune orientation across different histologies, and the prognostic impact of such interplay. Moreover, we investigated the predictive value of ICR classification across various public datasets of immune checkpoint inhibition therapy.

Methods

RNA-seq data of samples from a total of 9,282 patient tumor samples representing 31 cancer types were obtained from The Cancer Genome Atlas (TCGA). We classified each cancer type based on the expression of the ICR gene signature. Oncogenic pathway gene set enrichment and mutational status were analyzed in relation to ICR phenotypes. To explore whether tumor-intrinsic attributes associate with the prognostic value of ICR across cancers, we compared mutational load, oncogenic alterations and expression of oncogenic pathways between cancer types using an integrative bioinformatic pipeline.

Results

Our analyses identified a distinct prognostic connotation of ICR depending on cancer histology. We identified several oncogenic pathways whose enrichment inversely correlated with ICR in multiple tumor types. We found several cancer specific pathways that were differentially enriched between tumors in which ICR had a prognostic impact versus the ones in which ICR did not bear any prognostic connotation such as proliferation and TGF-beta signaling. Importantly, this conditional impact of ICR was also validated in the context of immune checkpoint inhibition treatment.

Conclusions

We identified tumor-intrinsic attributes that correlate with immune phenotypes and potentially influence their development. In addition, a relationship was observed between the enrichment of oncogenic pathways and the prognostic significance of the ICR and its predictive value for patients treated with anti-CTLA4 immune checkpoint inhibition. Such information can be used to prioritize potential candidates for therapies aimed at converting immune-silent into immune-active tumors and to refine stratification algorithms.

Keywords

- Pan-cancer
- Immunologic Constant of Rejection
- Prognosis
- Mutational load
- Neoantigen load
- Aneuploidy
- Tumor intrinsic pathway enrichment
- Cancer Immunotherapy

Background

Evidence of the effects of anti-tumoral immunity on cancer progression has accumulated over the last decades. The identification of tumor immune escape mechanisms, most importantly the upregulation of immune checkpoints, has led to major advances in immunotherapy. Immune checkpoint inhibitors have dramatically improved clinical outcome for a subset of patients across multiple cancer types. Despite this progress, the majority of patients (60-80%) still fail to respond. In addition, for some cancer types, immunotherapy remains ineffective. A pre-existing anti-tumor immune orientation has been associated with improved survival, both in treated and un-treated settings (Galon et al., 2013). However, in some cancer types, or specific subsets of cancer types, this association seems to be less clear, or even absent. For this reason, it is crucial to identify determinants of cancer immune orientation with its prognostic and predictive value.

Multiple studies have reported differences in the association between measures of intratumoral immune activity and overall survival across different cancer types (Charoentong et al., 2017; Danaher et al., 2018; Tamborero et al., 2018; Thorsson et al.; Varn et al., 2017). Cancer type-specific differences in tumor mutational load are only partially responsible for this observation. It has been shown that cancer types with a high mutational burden have an increased responsiveness to anti-PD-1 or anti-PD-L1 therapy (Yarchoan et al., 2017). The biologic rationale behind this observation is that a high number of mutations increases the likelihood of generating neoantigens recognized as non-self by the adaptive immune system. Indeed, increased mutational load has been associated with increased number of predicted neoantigens and cytolytic activity across various cancer types (Rooney et al., 2015). In breast cancer a positive association between survival and density of tumor infiltrating lymphocytes, as estimated by transcriptomic data, is restricted to tumors displaying a high mutational load or an aggressive/high proliferative phenotype (Miller et al., 2016; Nagalla et al., 2013; Thomas et al., 2018).

Over the years, multiple tumor immune classification systems have been introduced that aim to reflect antitumor immune activity based on intratumoral gene expression. Proposed classifications range from a measure of cytolytic activity by mean expression of *GZMA* and *PRF1* genes (Rooney et al., 2015), to reflections of immune cell infiltration by cell-specific transcriptional profiles (Bindea et al., 2013; Nagalla et al., 2013), or gene signatures reflecting molecular components of an antitumor immune response, including Major Histocompatibility Complex (MHC), co-stimulatory or immunomodulatory molecules (Ayers et al., 2017; Charoentong et al., 2017; Wang et al., 2008). Reported signatures typically show overlapping genes or genes involved in identical immune processes. In this study, we characterize expression of the Immunologic Constant of Rejection (ICR) signature to define immune

phenotypes, which incorporates IFN-stimulated genes driven by transcription factors *IRF1* and *STAT1*, *CCR5* and *CXCR3* ligands, immune effector molecules, and counter-activated immune regulatory genes (Hendrickx et al., 2017; Turan et al., 2018; Wang et al., 2008, 2013). Previously, we observed significantly prolonged survival of patients with a high expression of ICR genes in breast cancer. Moreover, we identified genetic determinants of ICR immune phenotypes, in particular, we noticed transcriptional deregulation of the MAPK pathway in tumors with low ICR gene expression (Hendrickx et al., 2017). We also observed that the ICR signature refines and improves the prognostic value of conventional prognostic signatures adopted in breast cancer (Bertucci et al., 2018).

In this study, we have extended this analysis to encompass cancers from 31 different cancer types from the TCGA database. Using a pan-cancer approach, we identified relationships between tumor genetic programs and immune orientation. In addition, we confirmed different associations between ICR classification and overall survival across cancer types. Thereafter, we systemically analyzed in which (molecular) contexts ICR has prognostic value and in which ones it does not. Combination of immune orientation with tumor intrinsic attributes that interact with the prognostic significance of this immune orientation could refine tumor-immune classifications.

Results

Prognostic impact of ICR classification is different between cancer types

RNA-seq data of samples from a total of 9,282 patients across 31 distinct solid cancer types were obtained from TCGA. To classify cancer samples based on their immune orientation, we performed unsupervised consensus clustering for each cancer type separately based on the expression of the ICR immune gene signature. This signature consists of 20 genes that reflect activation of Th1-signaling (*IFNG*, *TXB21*, *CD8B*, *CD8A*, *IL12B*, *STAT1,* and *IRF1*), CXCR3/CCR5 chemokine ligands (*CXCL9*, *CXCL10,* and *CCL5*), cytotoxic effector molecules (*GNLY*, *PRF1*, *GZMA*, *GZMB,* and *GZMH*) and compensatory immune regulators (*CD274/PD-L1, PDCD1, CTLA4, FOXP,3* and *IDO1*) (**Figure 1A**). Expression of these genes showed a positive correlation with each other across all cancer types (**Supplementary Figure 1**). The ICR signature highly correlates with other immune signatures that aim to reflect a highly active immune tumor microenvironment, including the Tumor Inflammation Signatures (TIS) (r = 0.97) (Danaher et al., 2018) (**Supplementary Figure 2**). As a representative example, consensus clustering and cluster assignment of skin cutaneous melanoma (SKCM) is shown in **Figure 1A**. Analogous figures for each of the 31 cancer types are available as cancer datasheets at figshare.com.

As shown in **Figure 1B**, the mean expression of ICR genes, or ICR score, varies between cancer types, reflecting general differences in tumor immunogenicity between cancers. While brain tumors (brain lower grade glioma's (LGG) and glioblastoma multiforme (GBM)) typically display low immune activity (McGranahan et al., 2017), skin cutaneous melanoma (SKCM) and head and neck squamous cell carcinoma (HNSC) display high levels of immune activation (Economopoulou et al., 2016; Passarelli et al., 2017). In addition, the distribution of ICR scores among patients and the difference between the highest and lowest ICR scores varies between cancers. Accordingly, the proportions of patients assigned to specific ICR clusters are dependent on the cancer type. Even more clinically relevant, the relation of the different immune phenotypes to survival is dissimilar among cancer types (**Figure 1C-D**). While the ICR High phenotype shows a significant survival benefit compared with the ICR Low phenotype for various cancer types (BRCA, SKCM, UCEC, SARC), the ICR High cluster is associated with significantly reduced overall survival in other cancer types (UVM, LGG, PAAD, KIRC) (**Figure 1C**). Similar results were obtained when Cox regression analysis was performed on ICR score as a continuous variable (**Supplementary Table 1**). To explore biological differences between cancer types in which a highly active immune phenotype is mostly associated with favorable survival and cancer types in which this phenotype is mostly associated with decreased survival, we categorized cancer types in ICR-*enabled* (BRCA, SKCM, UCEC, SARC, LIHC, HNSC, STAD, BLCA) and ICR-*disabled* (UVM,

LGG, PAAD, KIRC) groups, respectively (**Figure 1C**). All other cancer types in which ICR did not show an association or trend were categorized as ICR-*neutral*. Of important note, this classification was used for explorative purposes, a role of the immune mediated tumor rejection cannot be precluded in ICR-neutral cancer types.

First, we explored whether the ICR scores and their distributions were different among these defined groups of cancer types. Mean ICR score is low for most ICR-disabled (ranging from 3.97 to 8.34) compared to ICR-enabled cancer types (ranging from 7.26 to 8.36) (**Supplementary Figure 3A**). This observation is most noticeable for ICR-disabled cancer types LGG and UVM. Moreover, the difference (delta) between ICR scores in ICR High compared to ICR Low groups is higher in ICR-enabled cancer types (range: 2.98-4.97) compared to ICR-neutral (range: 1.48-4.49) and ICR-disabled cancer types (range: 2.29-3.35) (**Supplementary Figure 3B**). These factors could underlie, at least partially, the observed divergent associations with survival. However, it is likely that additional factors influence the prognostic significance of the ICR, as ICR parameters *per se* cannot perfectly segregate ICR-enabled from ICR-disabled cancer types.

To define whether tumor pathologic stage might interact with the association between ICR- and overall survival (OS), we fitted a Cox proportional hazards model for each group of ICR-enabled, ICR-neutral and ICR-disabled cancer types (**Table 1**). Overall, including ICR High and ICR Low samples from all cancer types, ICR has significant prognostic value independent of AJCC pathologic stage. For ICR-enabled cancer types, the ICR High group also remains significantly associated with improved survival after adjusting for tumor pathologic stage. For ICR-disabled cancer types, ICR High was associated with worse survival in univariate analysis (HR <1). However, in multivariate models this negative prognostic value of ICR was lost (HR=1.054; 95% CI= 0.7702- 1.443). Kaplan-Meier plots stratified by pathologic stage showed that within individual pathologic stages, ICR was not associated with OS for ICR-disabled cancers (**Supplementary Figure 4**). For ICR-neutral cancer types, while ICR was not associated with survival in univariate analysis, multivariate analysis indeed identified a positive prognostic value of the ICR classification, though less robust than observed for ICR-enabled cancer types.

ICR reflects anti-tumor immune activity and is inversely correlated with tumor-related pathways associated with immune escape

To further explore differences between cancer types, we aimed to compare immune cell infiltration between ICR High and Low samples across cancers. Gene expression signatures specific to 24 cell types (Bindea et al., 2013) were used to deconvolute the abundance of immune cells in tumor samples by performing single sample gene set enrichment analysis (ssGSEA) (Barbie et al., 2009). Cell-specific enrichment scores (ES) for

each patient demonstrated a clear enrichment of transcripts specific to T- and B cells in ICR High patients (**Figure 2A**). More specifically, ICR High samples showed increased expression of transcripts associated with cytotoxic T cells, T-regulatory (T-reg) cells, Th1 cells, NK CD56dim cells, activated dendritic cells (aDC) and macrophages, compared with ICR Medium and ICR Low samples. This observation is consistent across cancer types, in both ICR-enabled and ICR-disabled cancers. So, in addition to the immune functional molecular orientation, the ICR gene signature is a good reflection of anti-tumor immune cell infiltration (Lu et al., 2017). To quantitatively compare immune cell enrichment between individual cancer types, the mean ES was calculated for each cancer type (**Supplementary Figure 5**). Overall, no single consistent difference in terms of immune cell enrichment can be observed that can discriminate ICR-enabled from ICR-disabled cancer types. LGG and UVM show an overall low immune infiltration, consistent with our reported low ICR scores.

We then proceeded to examine which tumor intrinsic attributes correlate with immune phenotype as reflected by ICR gene expression. We performed ssGSEA to identify enrichment of transcripts of common tumor-related pathways (Hendrickx et al., 2017; Lu et al., 2017; Salerno et al., 2016). Not surprisingly, immune-related pathways including TNFR1 Signaling and immunogenic cell death showed a strong positive correlation with expression of ICR genes (**Figure 2B**). This implies that our immune signature captures the anti-tumoral immunological processes well across a wide range of cancer types. Interestingly, various pathways were identified that inversely correlated with ICR gene expression, potentially representing mechanisms by which immune silent tumors develop their phenotype. Of these, WNT-β catenin (Corrales et al., 2017; Spranger et al., 2015), barrier genes (Salerno et al., 2016), AMPK signaling (Dandapani and Hardie, 2013), and dysregulated KRAS signaling (Bedognetti et al., 2017) have previously been reported in the context of cancer immune resistance (Lu et al., 2017). Of special note, genes that we previously found to be upregulated in MAP3K1/MAP2K4-mutated vs wild-type (wt) breast cancer which perfectly segregated ICR High versus Low samples in the BRCA TCGA cohort (Hendrickx et al., 2017), were also inversely correlated with ICR in ACC, THYM, GBM, LGG and TGCT.

Characterization of tumor mutational load and aneuploidy in relation to ICR immune phenotypes

Next, we aimed to identify genomic attributes related to the ICR immune phenotypes. We used the non-silent mutation rate and predicted neoantigen data from the recent TCGA study by Thorsson *et al* (Thorsson et al., 2018) to explore their association with cancer immune phenotypes. As anticipated, mean neoantigen count of each cancer type strongly correlated with mean mutation rate (**Supplementary Figure 6A-B**). While mean non-silent mutation rate is significantly higher in ICR High tumors for some cancer types, no clear association was

observed in most cancer types. Results for predicted neoantigen load are similar (**Figure 3A-B** and **Supplementary Figure 6C-D**). Overall, mean non-silent mutation rate and mean neoantigen load are higher in ICR-enabled cancers compared with ICR-disabled cancers. However, these differences cannot fully explain the divergent association of ICR with survival, as values for ICR-enabled cancers SARC and BRCA are in the same range as ICR-disabled cancers LGG, PAAD and KIRC.

Similarly, we studied the association between genomic instabilities, or aneuploidy, and ICR. Specifically, we compared the individual tumor aneuploidy scores and the ICR score across cohorts. Aneuploidy score was calculated as in Taylor *et al* (Taylor et al., 2018). As has been reported previously, we found a broad negative association between aneuploidy and raw or tumor purity adjusted ICR score (Davoli et al., 2017) (**Figure 3C**). Interestingly, this negative association was most strongly supported in ICR-enabled cancers, with 6 cancers out of 8 showing a significant negative association between aneuploidy score and purity adjusted ICR ($P < 0.01$). In ICR-neutral cancers, a small fraction of cancer types showed a negative association (4 of 18, with an additional 4 showing a non-significant but suggestive negative association). Three cohorts (GBM, KICH and PRAD) showed a suggestive positive association. Similarly, in the ICR-disabled cohorts only KIRC showed a significant negative association, while LGG showed a strongly significant positive association (p-value $< 10^{-8}$). These results suggest that the previously reported relationship between aneuploidy and tumor-immune context is in fact tissue and cancer type dependent, with specific cancers showing a robust association, other cancers showing no relationship, and a minority of cancers (LGG in particular) showing a positive relationship.

Specific mutations associate with ICR immune phenotypes

To define the association of specific oncogenic mutations with ICR immune phenotypes, we first selected a set of 470 frequently mutated genes in cancer (Iorio et al., 2016), then trained an elastic net (Zou and Hastie, 2005) model to predict the ICR score as function of mutations in each sample and using the tumor-type as covariate. The positive non-zero coefficients of the trained model were used to identify genes whose mutation are associated with an increase of the ICR and negative non-zero coefficients identify the genes whose mutations are associated to a decrease of the ICR score (**Figure 4A**). The use of tumor-type as covariate tends to limit the effect of the enrichment of mutations in specific cancer-types and their correlation with ICR score. The coefficients of the tumor-type were all different from zero, with the exception of BLCA, BRCA, CHOL, COAD, READ and SARC and retained in the final model. We evaluated the accuracy of the model in a ten-fold cross-validation computing the correlation between the model prediction and the true ICR scores and obtaining a Spearman correlation of 0.669 ± 0.012 (p-value $< 10^{-400}$). Genes associated

with a decrease of ICR score include, as expected, *IDH1/IDH2* which are typical of cold tumors such as LGG and GBM. Interestingly MAP3K1 mutations, whose effect on ICR Low has been described in breast cancer (Hendrickx et al., 2017), are associated to ICR Low tumors pan-cancer. The figure also shows that, with the exception of CNS tumors, there are mutations which predispose the tumor to be ICR High (associated to positive coefficient) and mutations which predispose the tumor to be ICR low in the same cancer types.

To better compare the association between specific mutations and ICR groups within individual cancer types, we calculated, for each of the identified genes, the mean ICR score in the mutated group divided by the mean ICR score in the wild type (WT) within each individual cancer type. For most cancer types, the genes with a positive coefficient consistently showed a higher ICR score in mutated samples, supporting their association with an ICR High phenotype (**Figure 4B**). On the other hand, genes with a negative coefficient (genes associated with an ICR Low phenotype) as identified at the pan-cancer level, do show some clear deviations between cancer types. While for most cancer types, ICR score is indeed lower in the mutated group, results for cancer types COAD, UCEC and STAD show the reverse (**Figure 4B**). Interestingly, a common characteristic of these three cancer types is frequent hypermutation as a consequence of microsatellite instability (MSI) (Cortes-Ciriano et al., 2017). This hypermutator phenotype could be responsible for the observed increased ICR score in the mutated group, as the genes with negative coefficient could be mutated in the context of hypermutation. We indeed observed an increased ICR score in the MSI-high group compared to MSI-low and microsatellite stable (MSS) groups in COAD and STAD datasets for which sufficient data on MSI status were available (Cortes-Ciriano et al., 2017) (**Supplementary Figure 7A-B**).

We did not observe an enrichment of mutations in genes of pathways that inversely correlate with ICR (**Figure 2B**) suggesting that more complex biological processes are involved in upregulation of these tumor intrinsic pathways. Additionally, mutated genes were frequently part of multiple pathways, suggesting impact on various tumor biological systems (**Supplementary Figure 8**).

Prognostic impact of ICR classification is dependent on the expression of cancer-related pathways

Although we have observed interesting differences between ICR High and ICR Low immune phenotypes across different cancer types, these do not explain the divergent association between immune phenotype and ICR as we observed in ICR-enabled versus ICR-disabled cancer types (**Figure 1C-D**). As previously stated, an active immune phenotype has different impacts on survival depending on molecular subtype (for e.g. breast cancer (Miller et al., 2016)). To examine tumor intrinsic differences between ICR-enabled and ICR-disabled

cancer types, we compared the enrichment of tumor intrinsic pathways between these two groups. Differentially enriched pathways (t-test; FDR <0.05; **Supplementary Table 2**) between ICR-enabled and disabled cancer types were selected and used for pan-cancer hierarchical clustering. Interestingly, a wide variety of pathways were differentially enriched between both groups. Whereas enrichment for pathways involved in proliferation were mostly upregulated in ICR-enabled cancer types (proliferation metagene (Miller et al., 2016), E2F targets, G2M checkpoints), a large number of tumor intrinsic pathways (n=43) were enriched in ICR-disabled cancer types. Visualization of ES for these pathways across different cancer types in a heatmap confirms these findings. Hierarchical clustering based on ES of tumor intrinsic pathways differentially dysregulated by ICR-enabled and ICR-disabled cancer types segregates specimens into two main clusters (**Figure 5A**). As anticipated, pan-cancer survival analysis of all samples that formed a cluster along with samples of the ICR-disabled cancer types, named the *ICR non-beneficial cluster*, revealed no survival benefit of a high ICR expression. On the other hand, survival analysis of all samples in the other cluster, named the *ICR beneficial cluster*, showed a clear survival benefit for ICR High samples (**Figure 5B**). Of note, the prognostic significance of ICR was higher in this ICR beneficial cluster (HR = 1.82; p-value = 4.13^{-9}; 95% CI = 1.49-2.23) compared to the prognostic significance of all samples of ICR-enabled cancer types combined (HR = 1.63, p = 2.26^{-8}; 95% CI = 0.88-1.14), suggesting that tumor intrinsic attributes beyond the tumor site of origin are important to determine the relevance of cancer immune phenotypes. Interestingly, samples from ICR-neutral cancers, in which no clear trend was observed between ICR and survival (**Figure 1C**), and which were not used in calculation of differentially enriched pathways, were divided across the *ICR beneficial* and *ICR non-beneficial clusters*. To evaluate whether the prognostic impact of the ICR was relevant to a subset of samples from ICR-neutral cancer types, subgroup analysis was performed for samples of ICR-neutral cancer types. Indeed, for all samples from ICR-neutral cancer types that clustered to the *ICR non-beneficial cluster*, ICR was not associated with survival. On the other hand, for samples of ICR-neutral cancer types which clustered to the *ICR beneficial cluster*, ICR showed a significant positive association with survival (**Figure 5C**), indicating that the ICR has prognostic relevance in this subgroup of cancer patients as well.

To better clarify this concept, we selected two of the differentially expressed pathways that were of special interest. Firstly, the “Proliferation” signature was used to classify all samples independent of tumor origin in “Proliferation High” and “Proliferation Low” categories, defined as an ES value >median or <median of all samples, respectively. This 52-gene cluster described by Nagalla *et al* (Nagalla et al., 2013) has previously been associated with the prognostic value of immune gene signatures in breast cancer (Miller et al., 2016). As represented by a histogram, the proportion of samples with high proliferation signature

enrichment was larger in ICR-enabled cancer types compared with ICR-disabled cancers (**Figure 6A**). This very basic binary classification was already capable of segregating samples in a group in which ICR has a positive prognostic value from a group in which ICR is not associated with survival (**Figure 6B**). As a second illustration, “TGF beta signaling” was used to classify samples based on this pathway using the same approach. For this oncogenic pathway, ICR-enabled cancer types typically had a lower enrichment of this pathway compared to ICR-disabled cancer types (**Figure 6C**). This classification could also divide samples in a group in which ICR has a positive association with survival and a group in which this association is absent (**Figure 6D**).

As proliferation positively correlates with tumor mutational load (Pearson’s correlation coefficient = 0.49) (**Supplementary Figure 9**), we investigated whether tumor proliferation independently contributes to the prognostic value of ICR. Therefore, we segregated pan-cancer samples in four categories based on both mutation rate and proliferation (**Supplementary Figure 10**). Interestingly, in the proliferation high group, ICR High was associated with significantly improved survival independent of mutation rate. A similar observation is made for the mutation rate high group, ICR High is associated with better survival independent of proliferation. These finding suggest that mutation rate and enrichment of proliferation-related transcripts provide additive information to define the prognostic value of ICR. Furthermore, in a multivariate Cox proportional hazards model including ICR classification, proliferation enrichment, TGF beta signaling enrichment, tumor mutation rate, all parameters remain significant (**Supplementary Figure 11**). This implies that ICR, proliferation rate, TGF beta signaling and tumor mutation rate all have independent prognostic value.

We then continued by verifying whether these tumor intrinsic attributes that interact with the prognostic impact of ICR when evaluated pan-cancer, could also translate to individual cancer types. For each individual cancer type, samples were divided by median ES for each of the selected pathways. ICR HRs (ICR Low vs. ICR High) were compared between each pathway-High and pathway-Low group for each cancer type (**Supplementary Figure S12A-B**). Overall, we indeed observed an increased HR for samples with a high enrichment of ICR enabling pathways for most cancer types. For samples with a high enrichment of ICR disabling pathways, the HR was indeed lower (**Supplementary Figure S12C**).

These data confirm an association between the prognostic impact of ICR classification and enrichment of oncogenic pathways in individual cancer types as well as pan-cancer. Of note, these interactions between the prognostic significance of ICR and tumor intrinsic pathways were mostly present in enabled and neutral cancer types. Within disabled cancer types, with the exception of KIRC, similar associations were not found.

Predictive value of ICR score in immune checkpoint therapy is dependent on proliferation and TGF-beta signaling

To define the clinical relevance of classification of ICR immune phenotypes, in the setting of immune checkpoint treatment, we first evaluated the predictive value of ICR score across multiple public datasets of anti-CTLA4 and anti-PD1 treatment. A significantly increased expression of ICR in responders compared to non-responders was observed across most of the datasets (**Figure 7A**) (Chen et al., 2016; Hugo et al., 2016; Prat et al., 2017; Riaz et al., 2017; Van Allen et al., 2015). The conditional activation of the prognostic impact of the ICR was tested in the Van Allen dataset, which was the only one for which survival information were available. Strikingly, in the proliferation high subgroup, ICR score was significantly higher in pre-treatment samples of patients with long-survival or response (p=0.021), whereas this difference was not significant in proliferation low samples (**Figure 7B**). Cohort dichotomization based on TGF-beta signaling, again demonstrated the reverse trend: a significant difference in ICR score was only observed in the TGF-beta signaling low group (p=0.0044), not in the TGF-beta high group. Stratified survival analysis in these categories confirmed that the prognostic impact of ICR depends on proliferation and TGF-beta signaling (**Figure 7C**). These findings confirm a conditional prognostic and predictive impact of ICR based immune infiltration estimates in the setting of immune checkpoint treatment and demonstrate that these findings might have important clinical implications.

Methods

Data acquisition and normalization

RNA-seq data from The Cancer Genome Atlas (TCGA) were downloaded and processed using TCGA Assembler (v2.0.3). Gene symbols were converted to official HGNC gene symbols and genes without symbol or gene information were excluded. RNA-seq data from as wide as possible sample set of the total of 33 available cancer types of tissue types Primary Solid Tumor (TP), Recurrent Solid Tumor (TR), Additional-New Primary (TAP), Metastatic (TM), Additional Metastatic (TAM) and Solid Tissue Normal (NT) were used to generate a *pan-cancer* normalized dataset. Normalization was performed within lanes, to correct for gene-specific effects (including GC-content and gene length) and between lanes, to correct for sample-related differences (including sequencing depth) using R package EDASeq (v2.12.0) and quantile normalized using preprocessCore (v1.36.0). After normalization, samples were extracted to obtain a single primary tumor tissue (TP) sample per patient. For SKCM patients without available TP sample, a metastatic sample (TM) was included. Finally, the pan-cancer normalized dataset was filtered to remove duplicate patients and samples that did not pass assay-specific QCs (Thorsson) data was log2 transformed. Clinical data were sourced from the TCGA Pan-Cancer Clinical Data Resource (Liu et al., 2018). Mutation rate and predicted neoantigen load were obtained from the recent immunogenomic analysis by Thorsson *et al* (Thorsson et al.). The dataset published by Ellrott *et al* was used for mutation data analysis (Ellrott et al., 2018). Hematological cancer types LAML and DLBC were excluded from analysis.

Raw fastq files of datasets GSE78220 (Hugo et al., 2016) and GSE78220 (Riaz et al., 2017) were downloaded from NCBI SRA servers, quality control and adapter trimming was performed using Trim_Galore (https://github.com/FelixKrueger/TrimGalore). Reads were aligned to hg19 using STAR (Dobin et al., 2013). GenomicFeatures and GenomicAlignments Bioconductor packages were used to generate row counts. The raw counts were normalized with EDASeq (Risso et al., 2011) and log2 transformed. The dataset phs000452.v2.p1 (Van Allen et al., 2015) was downloaded, already normalized, from http://tide.dfci.harvard.edu/.

ICR classification

Consensus clustering based on the 20 ICR genes (**Figure 1A**) was performed for each cancer type separately using the ConsensusClusterPlus (v1.42.0) R package with the following parameters: 5,000 repeats, a maximum of six clusters, and agglomerative hierarchical clustering with ward criterion (Ward.D2) inner and complete outer linkage. The optimal number of clusters (≥ 3) for best segregation of samples based on the ICR signature was determined heuristically using the Calinski-Harabasz criterion (Caliński and Harabasz, 1974) (source

function available on GitHub repository, see cancer datasheets for plots with local maximum). As we were interested to compare cancer samples with a highly active immune phenotype with those that have not, the cluster with the highest expression of ICR genes was designated as “ICR High”, while the cluster with the lowest ICR gene expression was designated “ICR Low”. All samples in intermediate cluster(s) were defined as “ICR Medium”. Samples were annotated with an ICR score, defined as the mean of the normalized, log2 transformed expression values of the ICR genes. For generation of the ICR Heatmaps (**Figure 1C** and the cancer datasheets), a modified version of heatmap.3 function was used (source function).

Survival analysis

Overall survival (OS) from the TCGA Pan-Cancer Clinical Data Resource (Liu et al., 2018) was used to generate Kaplan-Meier curves using a modified version of the ggkm function (Abhijit). Patients with less than one day of follow-up were excluded and survival data were censored after a follow-up duration of 10 years. Hazard ratios (HR) between ICR Low and ICR High groups, including corresponding p-values based on chi-squared test, and confidence interval were calculated using the R package survival (v2.41-3). The forest plot (**Figure 1C**) was generated using the R package forestplot (v1.7.2). Cancer types PCPG, THYM, TGCT, and KICH were excluded before generation of the plot, as the number of deaths in the comparison groups was too small for calculations. Cancer types with a HR > 1 with a p-value < 0.1 were termed ICR-enabled and cancer types with a HR < 1 with a p-value < 0.1 were termed ICR-disabled. The forest plot was annotated manually with indicators for ICR-enabled and ICR-disabled cancer types. Cox proportional hazards regression analysis was performed with the R package survival with the AJCC pathologic tumor stage as described in the TCGA Pan-Cancer Clinical Data Resource (Liu et al., 2018). For simplification, stage categories were reduced to “Stage I”, “Stage II”, “Stage III” and “Stage IV” for subcategories (e.g. Stage IIA, Stage IIB, Stage IIC etc). In multivariate analysis, pathologic stage was entered as a semi-continuous (ordinal) variable.

Gene Set Enrichment Analysis

To define the enrichment of specific gene sets, either reflecting immune cell types (**Figure 2A**) or specific oncogenic pathways (**Figure 2B**), single sample GSEA (Barbie et al., 2009) was performed on the log2 transformed, normalized expression data. Immune cell-specific signatures as described in Bindea *et al* (Bindea et al., 2013) were used as gene sets using this method to deconvolute immune cell abundance. Gene sets to define enrichment of specific tumor-related pathways were obtained from the multiple sources. We started with a selection of 24 Hallmark pathways (Liberzon et al., 2015) which are regularly expressed in cancer. Subsequently, we added 21 non-redundant Ingenuity Pathway Analysis (IPA)

pathways (http://www.ingenuity.com, Ingenuity System Inc., Redwood City, CA, USA). Finally, several pathways were added that have previously been hypothesized to associate with cancer immune phenotypes, including Hypoxia/Adenosine Immune Cell Suppression, Immunogenic Cell Death (ICD), NOS1 Signature, PI3Kgamma signature, and SHC1/pSTAT3 signatures as described by Lu *et al* (Lu et al., 2017), barrier genes as described by Salerno *et al* (Salerno et al., 2016), the proliferation metagene as described by Miller *et al* (Miller et al., 2016) and genes upregulated in MAPK mutated breast cancer (Hendrickx et al., 2017).

Correlation matrix

The correlation matrices of ICR genes (**Supplementary Figure** 1) and correlation between ICR score and ES of selected pathways (**Figure 2B**) were calculated using Pearson test and plotted using "corrplot" version 0.84.

Mutational Analysis

Mutation rate and predicted neoantigen count data (Thorsson et al.) were log10-transformed and distribution across ICR clusters was plotted using R package "ggplot2". Differences between ICR High, Medium and Low clusters were calculated through t-test, using a cut-off p-value of < 0.05. For specific mutation analysis, a set of 470 frequently mutated genes in cancer (Iorio et al., 2016), was selected. An elastic net regularized (Zou and Hastie, 2005) model was built to predict the ICR score as function of mutations in each sample and using the tumor-type as a covariate. The accuracy of the model was evaluated in a ten-fold cross-validation setting computing the correlation between the model prediction and the true ICR scores, finally obtaining a Spearman correlation of 0.669 ± 0.012 (p-value $< 10^{-400}$).

The R package "ComplexHeatmap" was used to plot ICR score ratios between mutated versus wild-type groups. For cancer type/ gene combinations with a number of samples of <3 in the mutated group, ratios were not calculated (NA; grey color in plot). A ratio >1 implies that the ICR score is higher in the mutated group compared with WT, while a ratio <1 implies that the ICR score is higher in subset of tumors without mutation.

Aneuploidy

Aneuploidy scores for each individual cancer were taken from Taylor *et al* (Taylor et al., 2018). Briefly, each tumor was scored for the presence of aneuploid chromosome arms after accounting for tumor ploidy. Tumor aneuploidy scores for each cohort were then compared to ICR scores via linear model with and without purity adjustment. Purity adjustment entailed correlating ICR score and tumor purity (as estimated via ABSOLUTE) and using the residuals to evaluate the post-adjustment relationship between ICR score and tumor aneuploidy. In particular we made use of the precomputed aneuploidy scores and ABSOLUTE tumor purity

values. Raw ICR and aneuploidy score associations were evaluated by linear model in R via the *lm()* function for each cohort independently. Adjusted ICR and aneuploidy score associations were evaluated by first modeling ICR score by tumor purity, then taking the ICR score residuals and assessing the association with aneuploidy score via linear model. Cohorts with model p-values below 0.01 for adjusted or unadjusted ICR score and aneuploidy, regardless of the directionality of the association, were included in **Figure 3C**.

Differential GSEA and stratified survival analysis

Differential ES analysis between samples of ICR-enabled and those of ICR-disabled cancer types was performed using t-tests, with a cut-off of FDR-adjusted p-value (i.e., q-value) < 0.05 (**Supplementary Table 2**). Tumor intrinsic pathways that were differentially enriched between ICR-enabled and disabled cancer types were selected. The heatmap used for visualization of these differences was generated using the adapted heatmap.3 function (source function). For each of these selected pathways, samples were categorized pan-cancer as pathway-High (ES > median ES) or pathway-Low (ES < median ES). Associations between ICR and survival were defined for each pathway "High" and pathway "Low" group separately using the survival analysis methodology as described above. Pathways for which a significant association between ICR and survival was present in one group, but not in the other one, were selected (**Supplementary Table 3**). Similarly, these pathways were used to categorize samples per individual cancer type in pathway-High (ES > cancer specific median ES) and pathway-Low (ES < cancer specific median ES). Differences between HRs of groups in individual cancer types were calculated and plotted using "ComplexHeatmap" (v1.17.1).

Predictive value ICR score in immune checkpoint datasets

ICR scores, or the mean expression of ICR genes, were compared between responders and non-responders to immune checkpoint therapy. For the Chen et al dataset, performed on Nanostring platform, scores were calculated using the 17 ICR genes available in the nanostring panel. Difference in mean ICR score between groups was tested using two-side t-test (cutoff <0.95) (**Fig 7A**). For datasets, GSE78220 (Riaz et al., 2017), GSE78220 (Hugo et al., 2016) and Prat *et al* (Prat et al., 2017), the response category includes both partial and complete clinical responders according to respective publications. For Chen et al, clinical responders also included stable disease, as described by Chen *et al* (Chen et al., 2016). Dataset van Allen *et al*, response was defined as patients with clinical response or long-term survival after treatment (Van Allen et al., 2015). Samples of van Allen dataset were dichotomized based on median ssGSEA of 1) genes of the proliferation metagene and 2) TGF beta signaling signature. Stratified analysis was performed in each of the categories. ICR High, Medium and Low groups were defined according to ICR score tertiles, to obtain groups

of sufficient size. Stratified survival analysis was performed using the same approach as applied to the TCGA data.

Discussion

The immunologic landscape of human cancer can broadly be divided into immune-silent and immune-active phenotypes. Transcriptional signatures used to define the continuum of cancer immune surveillance within the tumor microenvironment typically reflect identical immune processes and include largely overlapping genes (Ayers et al., 2017; Hendrickx et al., 2017; Wang et al., 2008). In this study, we defined cancer immune phenotypes based on the expression of genes related with a Th-1 anti-tumor immune response, called ICR genes, across 31 different cancer types in a total of 9282 patients. We confirmed that high expression of these transcripts, referred to as an ICR High phenotype, was associated with favorable prognosis in multiple cancer types including SKCM, BRCA, UCEC, LIHC, SARC, HNSC, STAD, and BLCA. The expression of ICR genes had positive prognostic value in more cancer types compared with leukocyte infiltration, suggesting that it is crucial to consider the molecular orientation of the immune infiltrate to determine the impact on patient outcome. Similar to previous observations (Charoentong et al., 2017; Danaher et al., 2018; Tamborero et al., 2018; Thorsson et al.; Varn et al., 2017), the prognostic value of an ICR High immune phenotype is restricted to specific cancer types. In fact, the ICR High phenotype was associated with unfavorable prognosis for UVM, LGG, PAAD, and KIRC (**Figure 1C**). These observations prompted us to systematically investigate the association between immune phenotypes and tumor intrinsic attributes and the effect of this interplay on the prognostic value of the ICR immune phenotypes.

To compare cancer types based on the prognostic value of ICR, we categorized cancer types into two groups: one for which ICR High was associated with increased OS and one for which ICR was associated with worse OS. For the first group, Cox proportional hazards regression analysis confirmed a positive prognostic value of ICR independent of pathologic tumor stage. SKCM, BRCA, UCEC, LIHC, SARC, HNSC, STAD, and BLCA are consequently referred to as **ICR-enabled** cancer types. For the second group, including UVM, LGG, PAAD, and KIRC, survival analysis adjusted for tumor stage showed that within individual tumor stages, ICR was, in fact, *not* associated with survival. Therefore, we refer to these as **ICR-disabled** cancer types. A correlation between an active cancer immune phenotype and tumor size has also previously been reported in gliomas (Tamborero et al., 2018), explaining the initially observed negative association between ICR and OS.

The first notable difference we observed between ICR-enabled and -disabled cancer types is the overall lower ICR in disabled cancer types. In particular for UVM and LGG, this low ICR could be a partial explanation for the lack of positive prognostic value of the ICR. On the other hand, mean ICR score of PAAD and KIRC was not different compared with the other cancer types. Therefore, other factors must have an effect on the prognostic value of the ICR.

Estimation of immune cell subsets across cancer immune phenotypes (e.g., ICR Low, - Medium, High) by ssGSEA using immune cell signatures (Bindea et al., 2013) confirmed that ICR High tumors were enriched in transcripts related to immune cells that are typically associated with a Th-1 immune response (CD8 T cells, DCs, T-regs, macrophages, NK and NKT cells). No deviations were observed that could explain the different associations between ICR and survival.

We then proceeded to analyse global genomic aberrations, including overall mutational load, predicted neoantigen load, and genomic instability in relation to ICR immune phenotypes and their prognostic value. Mutational load was significantly higher in the ICR High tumors compared with ICR Low for a subset of cancer types. These included cancer types COAD, STAD and UCEC for which hypermutation as a consequence of microsatellite instability and associated high immunogenicity is present in a subset of the samples. However, for most cancer types, no significant difference in mutational load was found between ICR High and ICR Low. The results confirm previous findings of Danaher *et al*, that demonstrated that mutational load only correlated minimally with expression of the Tumor Inflammation Signature (TIS) for a small number of cancer types (Danaher et al., 2018). As neoantigen load might have a stronger association with tumor immunogenicity, we additionally compared predicted neoantigen load between ICR High and ICR Low. We found a significant increase in ICR High compared to ICR Low tumors in 12 and 14 cancer types for mutational load and neoantigen load, respectively. However, for both mutational load and neoantigen load, relative differences between ICR High and ICR Low groups are not large and range widely within ICR subgroups. These findings support the fact that additional factors beyond mutational load and neoantigen load determine immune phenotypes. Although overall mutational load and neoantigen load were higher in ICR-enabled cancer types, these attributes alone cannot clearly discriminate ICR-enabled from -disabled cancer types.

As for genomic instability, tumors with high aneuploidy are associated with decreased ICR score in a major subset of cancer types. This finding confirms previous observations in which aneuploidy was inversely correlated with cytotoxic immune infiltrates in a pan-cancer analysis (Davoli et al., 2017). This observation is also in agreement with negative association of a chromosome-instable type with an immune signature that predicts response to immunotherapy with MAGE-A3 antigen as well as response to anti-CTLA-4 treatment in melanoma (Ock et al., 2017). The only exceptions we found were brain tumors LGG and GBM in which a positive association between aneuploidy and ICR score was found. Although LGG and GBM are both ICR-disabled cancer types, this correlation is not consistently observed in ICR-disabled cancer types. In fact, for KIRC aneuploidy inversely correlates with ICR.

We continued our investigation to mutations of specific genes. A recent pan-cancer study by Thorsson *et al* investigated mutations associated with changes in leukocyte fraction

(Thorsson et al.). Here, we confirm that *IDH1*, *NSD1* and *GATA3* mutations were associated with an ICR Low immune phenotype pan-cancer. Using our approach, we found additional mutations associated with low ICR, including *IDH2*, *FOXA2*, *WNT5A*, *KRAS*, *NRAS* and *EGFR*. *WNT5A* mutations might impact intrinsic WNT-ß catenin signaling, for which a role in tumor immune escape has been defined (Spranger and Gajewski, 2015). Interestingly, *MAP3K1* mutations, which we previously found to be associated with low ICR in breast cancer (Hendrickx et al., 2017), was found in the top 10 mutated genes associated with a low ICR pan-cancer. Per cancer analysis showed that *MAP3K1* mutation is also associated with low ICR in BLCA, ESCA, GBM, KIRC and OV (**Figure 4B**). The role of *MAP3K1* mutation and MAPK signaling on cancer immune phenotypes warrants further investigation in these cancer types. Similar to previous observations, we found *HLA-A*, *HLA-B*, *B2M*, *CASP-8* and *FAS* to be associated with an ICR High immune phenotype (Ock et al., 2017; Rooney et al., 2015; Shukla et al., 2015; Siemers et al., 2017; Thorsson et al., 2018). Specific mutations associated with ICR High and ICR Low were quite similar in ICR-enabled versus ICR-disabled cancer types. Similarly, we found that genomic alterations of oncogenic pathways (Sanchez-Vega et al., 2018) did not show clear differences between ICR-enabled and - disabled cancer types (data not shown).

Finally, we investigated tumor transcriptomics in relation to ICR and its prognostic value. We found that enrichment of transcripts of several pathways was associated with a low ICR score, including WNT-ß catenin, barrier genes, AMPK signaling, and dysregulated KRAS signaling. For several of the identified pathways negatively correlating with ICR, a role in tumor immune escape has been described previously (Bedognetti et al., 2016; Corrales et al., 2017; Dandapani and Hardie, 2013; Salerno et al., 2016; Spranger et al., 2015). Identification of specific mutations or regulation of tumor intrinsic pathways associated with cancer immune phenotypes is vital to be able to design therapeutic strategies to overcome tumor immune resistance. For example, inhibition of the TGF beta signaling pathway has shown to be able to revert cancer immune phenotype to an immune active one, and could therefore increase efficacy of cancer immunotherapy (Ganesh and Massagué, 2018; Mariathasan et al., 2018; Tauriello and Batlle, 2016). In a similar way, immunomodulatory effects of inhibition of MAPK signaling pathways could be effectively used in combination with immunotherapy (Bedognetti et al., 2017).

Next, we compared enrichment of oncogenic pathways in ICR-enabled and -disabled cancer types. Strikingly, we identified major differences in enrichment of transcripts of tumor intrinsic pathways between these groups of cancer types. While ICR-enabled cancer types are typically more enriched in proliferation-related signatures, ICR-disabled cancer types have high enrichment of pathways generally attributed to tumor signaling. Hierarchical clustering based on enrichment of transcripts of these differentially enriched pathways segregated most

samples of ICR-enabled cancer types from samples of ICR-disabled cancer types. Interestingly, this clustering was even relevant to samples of ICR neutral cancer types. Pan-cancer survival analysis of samples of ICR neutral cancer types showed that for samples that co-clustered with samples of ICR-enabled cancer types (the ICR beneficial cluster), ICR High was associated with significant prolonged survival. Moreover, for samples of ICR neutral cancer types clustered to the ICR non-beneficial cluster, ICR lost its prognostic value.

The clinical relevance of the observed conditional impact of ICR was confirmed in the setting of anti-CTLA4 treatment. The predictive value of ICR was demonstrated to be dependent on tumor intrinsic pathways, such as TGF-beta and proliferation.

To the best of our knowledge, we are the first to report an interaction between tumor intrinsic pathways and the prognostic value of immune phenotypes in a pan-cancer analysis. Most importantly, we found that the impact of tumor intrinsic pathways on the prognostic significance of the ICR is independent of tumor mutation rate and neoantigen load. Analysis within individual cancer types confirmed an association between the prognostic significance of the ICR and expression of tumor intrinsic pathways.

An association between proliferation and the prognostic value of immune phenotypes has previously been identified in breast cancer (Miller et al., 2016). Additionally, in non-small cell lung cancer (NSCLC), cell proliferation was shown to improve prediction of immune checkpoint inhibitors response in PD-L1 positive samples as recently presented at SITC annual meeting 2018 (SITC annual meeting 2018, abstract, Jason Zhu *et al*). Our study clearly demonstrates that such interactions between tumor intrinsic attributes and prognostic and potentially predictive value of immune phenotypes are also relevant in a pan-cancer context. Moreover, we defined additional tumor intrinsic attributes beyond tumor proliferation to correlate with the prognostic significance of the ICR. These findings suggest that prognostication algorithms should be refined by inclusion of tumor intrinsic attributes in order to define the prognostic impact of the ICR.

Limitations of this study include the fact that the directionality of tumor intrinsic pathways cannot be directly derived from enrichment of pathways transcripts. However, we must note that the majority of genes included in our signatures are positive mediators of the pathways. In addition, though our analyzed signatures are typical cancer-related pathways, the origin of transcripts could not be determined in this bulk-sample analysis. Furthermore, infiltration of stromal cells to the tumor site affects the relative expression of tumor derived transcripts. Single cell gene expression profiling approaches should be applied to future cohorts in order to specifically dissect origin of transcripts. Despite these limitations, we observed a strong correlation between expression of the proliferation genes and expression of *MKI67* or Ki-67, an established marker of tumor proliferation, suggesting that our signature enrichment indeed reflects tumor cell-specific attributes.

To summarize, we have explored whether tumor intrinsic attributes, including tumor genomics and transcriptomics, have an impact on the prognostic value of ICR classification. We systematically compared these tumor intrinsic attributes between ICR-enabled and -disabled cancer types. Although overall mutational load and neoantigen load were higher is ICR-enabled cancer types, these attributes did not clearly discriminate ICR-enabled from -disabled cancer types. Similarly, mutations of specific genes or oncogenic alteration of pathways could not distinguish these different groups of cancer types. Remarkably, enrichment tumor intrinsic pathways did show clear differences between ICR-enabled and -disabled cancer types. As much as 43 of 54 analyzed pathways showed differential enrichment between ICR-enabled and ICR-disabled cancer types. We confirmed an impact of enrichment of these pathways on the prognostic value of the ICR in ICR neutral cancer types, as well as within individual cancer types. Overall, our findings indicate that for prognostic purposes, one should not only take the ICR immune phenotype into account, but rather use this in combination with the expression of tumor intrinsic pathways. The concept of the conditional activation of the prognostic role of the ICR was validated in the context of checkpoint inhibition.

In conclusion, we observed a clear relation between enrichment of tumor intrinsic pathways and the prognostic and predictive significance of the ICR gene signatures. This information can be used to prioritize candidates for immunogenic conversion and to refine stratification algorithms.

Declarations

Ethics approval and consent to participate

Not applicable.

Availability of data and material

Scripts for analysis are shared on GitHub. Data generated or analyzed during this study are included in this article.

Competing interests

The authors do not have any competing interests to disclose.

Funding

Support for analysis and manuscript writing was provided by Qatar Foundation, Qatar National Research Fund (grant number: JSREP07-010-3-005).

References

Abhijit An enhanced Kaplan-Meier plot.

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

Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108–112.

Bedognetti, D., Hendrickx, W., Ceccarelli, M., Miller, L.D., and Seliger, B. (2016). Disentangling the relationship between tumor genetic programs and immune responsiveness. Curr. Opin. Immunol. *39*, 150–158.

Bedognetti, D., Roelands, J., Decock, J., Wang, E., and Hendrickx, W. (2017). The MAPK hypothesis: immune-regulatory effects of MAPK-pathway genetic dysregulations and implications for breast cancer immunotherapy. Emerg. Top. Life Sci. *1*, 429–445.

Bertucci, F., Finetti, P., Simeone, I., Hendrickx, W., Wang, E., Marincola, F.M., Viens, P., Mamessier, E., Ceccarelli, M., Birnbaum, D., et al. (2018). The immunologic constant of rejection classification refines the prognostic value of conventional prognostic signatures in breast cancer. Br. J. Cancer.

Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenauf, A.C., Angell, H., Fredriksen, T., Lafontaine, L., Berger, A., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity *39*, 782–795.

Buccitelli, C., Salgueiro, L., Rowald, K., Sotillo, R., Mardin, B.R., and Korbel, J.O. (2017). Pan-cancer analysis distinguishes transcriptional changes of aneuploidy from proliferation. Genome Res. *27*, 501–511.

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl, H., and Trajanoski, Z. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. *18*, 248–262.

Corrales, L., Matson, V., Flood, B., Spranger, S., and Gajewski, T.F. (2017). Innate immune signaling and regulation in cancer immunotherapy. Cell Res. *27*, 96–108.

Cortes-Ciriano, I., Lee, S., Park, W.-Y., Kim, T.-M., and Park, P.J. (2017). A molecular portrait of microsatellite instability across multiple cancers. Nat. Commun. *8*.

Danaher, P., Warren, S., Lu, R., Samayoa, J., Sullivan, A., Pekker, I., Wallden, B., Marincola, F.M., and Cesano, A. (2018). 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.

Dandapani, M., and Hardie, D.G. (2013). AMPK: opposing the metabolic changes in both tumour cells and inflammatory cells? Biochem. Soc. Trans. *41*, 687–693.

Davoli, T., Uno, H., Wooten, E.C., and Elledge, S.J. (2017). Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. Science *355*, eaaf8399.

Economopoulou, P., Perisanidis, C., Giotakis, E.I., and Psyrri, A. (2016). The emerging role of immunotherapy in head and neck squamous cell carcinoma (HNSCC): anti-tumor immunity and clinical applications. Ann. Transl. Med. *4*.

Ellrott, K., Bailey, M.H., Saksena, G., Covington, K.R., Kandoth, C., Stewart, C., Hess, J., Ma, S., Chiotti, K.E., McLellan, M., et al. (2018). Scalable Open Science Approach for

Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. Cell Syst. *6*, 271-281.e7.

Fridman, W.-H., Dieu-Nosjean, M.-C., Pagès, F., Cremer, I., Damotte, D., Sautès-Fridman, C., and Galon, J. (2012). The Immune Microenvironment of Human Tumors: General Significance and Clinical Impact. Cancer Microenviron. *6*, 117–122.

Galon, J., Angell, H.K., Bedognetti, D., and Marincola, F.M. (2013). The Continuum of Cancer Immunosurveillance: Prognostic, Predictive, and Mechanistic Signatures. Immunity *39*, 11–26.

Ganesh, K., and Massagué, J. (2018). TGF-β Inhibition and Immunotherapy: Checkmate. Immunity *48*, 626–628.

Hendrickx, W., Simeone, I., Anjum, S., Mokrab, Y., Bertucci, F., Finetti, P., Curigliano, G., Seliger, B., Cerulo, L., Tomei, S., et al. (2017). Identification of genetic determinants of breast cancer immune phenotypes by integrative genome-scale analysis. OncoImmunology *0*, 00–00.

Iorio, F., Knijnenburg, T.A., Vis, D.J., Bignell, G.R., Menden, M.P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., et al. (2016). A Landscape of Pharmacogenomic Interactions in Cancer. Cell *166*, 740–754.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. *1*, 417–425.

Liu, J., Lichtenberg, T., Hoadley, K.A., Poisson, L.M., Lazar, A.J., Cherniack, A.D., Kovatich, A.J., Benz, C.C., Levine, D.A., Lee, A.V., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell *173*, 400-416.e11.

Lu, R., Turan, T., Samayoa, J., and Marincola, F.M. (2017). Cancer immune resistance: can theories converge? Emerg. Top. Life Sci. *1*, 411–419.

Mandruzzato, S., Brasseur, F., Andry, G., Boon, T., and van der Bruggen, P. (1997). A CASP-8 Mutation Recognized by Cytolytic T Lymphocytes on a Human Head and Neck Carcinoma. J. Exp. Med. *186*, 785–793.

Mariathasan, S., Turley, S.J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., Kadel, E.E., Koeppen, H., Astarita, J.L., Cubas, R., et al. (2018). TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature *554*, 544–548.

McGranahan, T., Li, G., and Nagpal, S. (2017). History and current state of immunotherapy in glioma and brain metastasis. Ther. Adv. Med. Oncol. *9*, 347–368.

Miller, L.D., Chou, J.A., Black, M.A., Print, C., Chifman, J., Alistar, A., Putti, T., Zhou, X., Bedognetti, D., Hendrickx, W., et al. (2016). Immunogenic Subtypes of Breast Cancer Delineated by Gene Classifiers of Immune Responsiveness. Cancer Immunol. Res.

Miller, L.D., Routh, E.D., Chou, J.W., and Thomas, A. (2017). Abstract 4715: Protective anti-tumor immune responses in breast cancer depend on tumor mutation rate. Cancer Res. *77*, 4715–4715.

Nagalla, S., Chou, J.W., Willingham, M.C., Ruiz, J., Vaughn, J.P., Dubey, P., Lash, T.L., Hamilton-Dutoit, S.J., Bergh, J., Sotiriou, C., et al. (2013). Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. Genome Biol. *14*, R34.

Ock, C.-Y., Hwang, J.-E., Keam, B., Kim, S.-B., Shim, J.-J., Jang, H.-J., Park, S., Sohn, B.H., Cha, M., Ajani, J.A., et al. (2017). Genomic landscape associated with potential response to anti-CTLA-4 treatment in cancers. Nat. Commun. *8*, 1050.

Passarelli, A., Mannavola, F., Stucci, L.S., Tucci, M., and Silvestris, F. (2017). Immune

system and melanoma biology: a balance between immunosurveillance and immune escape. Oncotarget *8*, 106132–106142.

Provenzale, J.M., Wang, G.R., Brenner, T., Petrella, J.R., and Sorensen, A.G. (2002). Comparison of permeability in high-grade and low-grade brain tumors using dynamic susceptibility contrast MR imaging. AJR Am. J. Roentgenol. *178*, 711–716.

Rooney, M.S., Shukla, S.A., Wu, C.J., Getz, G., and Hacohen, N. (2015). Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. Cell *160*, 48–61.

Salerno, E.P., Bedognetti, D., Mauldin, I.S., Deacon, D.H., Shea, S.M., Pinczewski, J., Obeid, J.M., Coukos, G., Wang, E., Gajewski, T.F., et al. (2016). Human melanomas and ovarian cancers overexpressing mechanical barrier molecule genes lack immune signatures and have increased patient mortality risk. OncoImmunology *5*, e1240857.

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W.K., Luna, A., La, K.C., Dimitriadoy, S., Liu, D.L., Kantheti, H.S., Saghafinia, S., et al. (2018). Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell *173*, 321-337.e10.

Shukla, S.A., Rooney, M.S., Rajasagi, M., Tiao, G., Dixon, P.M., Lawrence, M.S., Stevens, J., Lane, W.J., Dellagatta, J.L., Steelman, S., et al. (2015). Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. Nat. Biotechnol. *33*, 1152–1158.

Siemers, N.O., Holloway, J.L., Chang, H., Chasalow, S.D., Ross-MacDonald, P.B., Voliva, C.F., and Szustakowski, J.D. (2017). Genome-wide association analysis identifies genetic correlates of immune infiltrates in solid tumors. PLoS ONE *12*.

Spranger, S., Bao, R., and Gajewski, T.F. (2015). Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. Nature *523*, 231–235.

Tamborero, D., Rubio-Perez, C., Muiños, F., Sabarinathan, R., Piulats, J.M., Muntasell, A., Dienstmann, R., Lopez-Bigas, N., and Gonzalez-Perez, A. (2018). A pan-cancer landscape of interactions between solid tumors and infiltrating immune cell populations. Clin. Cancer Res. clincanres.3509.2017.

Tauriello, D.V.F., and Batlle, E. (2016). Targeting the Microenvironment in Advanced Colorectal Cancer. Trends Cancer *2*, 495–504.

Taylor, A.M., Shih, J., Ha, G., Gao, G.F., Zhang, X., Berger, A.C., Schumacher, S.E., Wang, C., Hu, H., Liu, J., et al. (2018). Genomic and Functional Approaches to Understanding Cancer Aneuploidy. Cancer Cell *33*, 676-689.e3.

Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.-H., Porta-Pardo, E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. The Immune Landscape of Cancer. Immunity.

Varn, F.S., Wang, Y., Mullins, D.W., Fiering, S., and Cheng, C. (2017). Systematic pan-cancer analysis reveals immune cell interactions in the tumor microenvironment. Cancer Res. *77*, 1271–1282.

Wang, E., Worschech, A., and Marincola, F.M. (2008). The immunologic constant of rejection. Trends Immunol. *29*, 256–262.

Wang, E., Bedognetti, D., and Marincola, F.M. (2013). Prediction of Response to Anticancer Immunotherapy Using Gene Signatures. J. Clin. Oncol. *31*, 2369–2371.

Yarchoan, M., Hopkins, A., and Jaffee, E.M. (2017). Tumor Mutational Burden and Response Rate to PD-1 Inhibition. N. Engl. J. Med. *377*, 2500–2501.

Yatim, N., Jusforgues-Saklani, H., Orozco, S., Schulz, O., Silva, R.B. da, Sousa, C.R. e, Green, D.R., Oberst, A., and Albert, M.L. (2015). RIPK1 and NF-κB signaling in dying cells determines cross-priming of CD8+ T cells. Science *350*, 328–334.

Zou, H., and Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net. J. R. Stat. Soc. Ser. B Stat. Methodol. *67*, 301–320.

SITC 2018.

Figure Legends

Figure 1: Immunologic classification of 31 cancer types based on expression of ICR gene signature A. Consensus cluster matrix of SKCM samples based on RNA-seq expression values of ICR genes (left panel). RNA-seq expression heatmap of ICR genes annotated with ICR consensus clusters (n = 469). Clusters with intermediate ICR gene expression levels (ICR Medium1 and ICR Medium2) were combined to obtain ICR High, Medium and Low groups (HML classification). ICR genes reflect 4 components of immune mediated tissue rejection: Th1 signaling, CXCR3/CCR5 chemokines, immune effectors and immune regulatory functions (right panel). **B.** Boxplot of ICR scores across ICR clusters in 31 cancer types. Cancer types are ordered by mean ICR score per cancer. **C.** Forest plot showing HRs (overall survival) of ICR Low versus High, p-value and number of patients (N) for each of the cancer types. ICR-enabled cancer types (HR > 1; $p < 0.1$) are indicated with orange asterisks and ICR-disabled cancer types (HR < 1; $p > 0.1$) are indicated with purple asterisks. Cancer types PCPG, THYM and TGCT are excluded from the plot, because confidence intervals ranged from 0 to infinite due to low number of deaths in these cancer types. **D.** Kaplan Meier curves showing OS across two three different ICR groups in ICR-enabled and ICR-disabled cancer types.
(Figures of panel A and Kaplan Meier curves for each individual cancer type are available in the cancer datasheets).

Figure 2: Deconvolution of immune cell populations and enrichment of oncogenic pathways through single sample GSEA. **A.** Heatmap of enrichment values for cell-specific immune-signatures as described by Bindea *et al*. Samples are ordered by ICR cluster and ordered by cancer type within ICR clusters. **B.** Pearson coefficient of correlation between ICR score and enrichment scores of oncogenic pathways per cancer. Pathways that have a positive correlation with ICR are green and those with an inverse correlation are blue.

Figure 3: Association of ICR with nonsilent mutation rate, predicted neoantigen load, and tumor aneuploidy. A. Scatter plot of log transformed nonsilent mutation count per ICR cluster for each cancer type. **B.** Log transformed predicated neoantigen load per ICR cluster for each cancer type. **A.B.** Red crossbar represents the mean value per ICR cluster. Cancer types are ordered by mean nonsilent mutation count per cancer. Nonsilent mutation rate and predicted neoantigen load were obtained from Thorsson et al (Thorsson et al., 2018). **C.** Correlation between aneuploidy score and raw/purity adjusted ICR score for all cohorts with significant relationships between ICR and aneuploidy.

Figure 4: Relationship between ICR score and mutations in individual genes. A. Top 35 of mutated genes with negative non-zero coefficients of a trained elastic net model identified genes whose mutation is associated with a decrease of the ICR (left panel). Top 35 mutated genes with a positive association with ICR 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 size of the bars. **B.** Ratio of mean ICR score in mutated samples and ICR score in WT samples. Cancer types are ordered manually based on patterns of calculated ratios.

Figure 5: Pan-cancer clustering based on oncogenic pathway enrichment segregates ICR-enabled and ICR-disabled cancer types. A. Heatmap of enrichment scores of selected oncogenic pathways, samples are hierarchically clustered in two main clusters: one cluster consists mostly of ICR-enabled cancer types (ICR beneficial cluster), while the second cluster contains all samples from ICR-disabled cancer types (ICR non-beneficial cluster). **B.** Kaplan-Meier OS curves for ICR High, Medium, and Low clusters for samples in the ICR beneficial and ICR non-beneficial cluster separately. **C.** Subgroup survival analysis of all samples of ICR-neutral cancer types clustered in the ICR beneficial cluster and ICR non-beneficial cluster.

Figure 6: Examples of pan-cancer binary classifications based on enrichment of individual tumor intrinsic gene signatures and corresponding stratified pan-cancer survival analysis. A. Histogram showing pan-cancer classification based on median pan-cancer enrichment value of the proliferation signature as described by Miller *et al* (Miller et al., 2016) (Proliferation Low: ES is lower than median ES observed pan-cancer; Proliferation High: ES is higher or equal to median ES observed pan-cancer). **B.** Pan-cancer Kaplan Meier curves of ICR groups stratified by Proliferation High (left panel) and Proliferation Low (right panel) groups corresponding to classification as shown in panel A. **C.** Histogram showing pan-cancer classification based on pan-cancer enrichment values of the Hallmark pathway TGF beta signaling. **D.** Pan-cancer Kaplan Meier curves stratified by TGF beta signaling Low (left panel) and TGF beta signaling High (right panel) groups corresponding to classification as shown in panel C.

Figure 7: Conditional predictive value of ICR for response to immune checkpoint treatment. A. Predictive value of ICR across public datasets with response to immune checkpoint treatment indicated by p-value of two-sided t-test comparing ICR score in samples

of responding versus non-responding patients. ICR score was highest in response group for all significant comparisons. Response was defined as long-survival or response in the van Allen dataset, stable disease, partial response (PR) and complete response (CR) in the Chen dataset, and as PRCR in Riaz, Hugo and Prat datasets. **B.** Boxplot of ICR score in "nonresponse" compared with "long-survival or response" to anti-CTLA4 treatment in van Allen dataset (left). Boxplots of subgroup analysis of proliferation groups (middle) and TGF-beta signaling groups (right). P-value of t-test comparing means are indicated in the plot. **C.** Kaplan Meier curves showing OS across ICR tertiles in all samples (left), across proliferation (middle), and TGF-beta signaling subgroups (left).

Supplementary Figures

Supplementary Figure 1: Pearson correlation between RNA-seq expression values of ICR genes for each of the 31 cancer types.

Supplementary Figure 2: Scatterplot showing correlation between ICR score and TIS score (Danaher et al., 2018) (**A**), ICR score and leukocyte fraction (**B**), and ICR score and TIL percentage (**C**). Leukocyte fraction and TIL percentage values were obtained from Thorsson *et al* (Thorsson et al., 2018). Each dot represents a single sample.

Supplementary Figure 3: A. Boxplot showing mean ICR score for each cancer type per group of cancer types: ICR-enabled, ICR-neutral and ICR-disabled. A single dot represents a single cancer type. **B.** Boxplot showing delta between mean ICR score in ICR High cluster compared with mean ICR score in ICR Low cluster. A single dot represents a single cancer type.

Supplementary Figure 4: Pan-cancer Kaplan-Meier curves in ICR-disabled (top left panel) and ICR-enabled (top right panel) groups and stratified analysis by AJCC pathologic stage I & II (middle panels) and stage III & IV (bottom panels).

Supplementary Figure 5: Dotted heatmap showing mean ES for each immune cell population per cancer type, mean ES scores were z-scored per row.

Supplementary Figure 6: A. Scatterplot of mean mutation rate versus mean neoantigen load per cancer type. **B.** Ratio of nonsilent mutation rate between ICR High and ICR Low groups versus the ratio of predicted neoantigen load between in ICR High compared to ICR Low groups. **C.** Ratio of nonsilent mutation rate between ICR High and ICR Low groups

versus mean nonsilent mutation rate. **D.** Ratio of predicted neoantigen load between ICR High compared to ICR Low groups versus mean predicated neoantigen load.

Supplementary Figure 7: A. Boxplot of ICR score by MSI status in COAD (left panel) and STAD (right panel). P-values of t-test to compare mean ICR score per MSI group are indicated in the plot. **B.** Boxplot of number of mutated genes with a negative coefficient in ICR trained elastic net model by MSI status in COAD (left panel) and STAD (right panel). P-values of t-test to compare mean number of mutations per MSI group are indicated in the plot.

Supplementary Figure 8: Table to check overlap between tumor intrinsic pathways genes and frequently mutated genes. When a gene (columns) is part of a gene signature (rows), this is indicated by “YES”, if not, it is indicated by “NO”. Genes that have a negative coefficient in trained model are shown in blue, pathways that inversely correlate with ICR (**Figure 2B**) are indicated in blue. Genes that have a positive coefficient in trained model are shown in red, pathways that positively correlate with ICR (**Figure 2B**) are indicated in red.

Supplementary Figure 9: Scatterplots of each of the combinations of: 1) ICR scores, 2) proliferation ES, 3) TGF beta signaling ES, and 4) mutation rate (n = 4452). Pearson’s correlation coefficient and regression line (red) are indicated in the plots.

Supplementary Figure 10: Pan-cancer Kaplan Meier curves of ICR groups stratified by both Proliferation High (left panels) and Proliferation Low (right panels) groups (corresponding to classification of shown in **Figure 6A**) and by Mutation rate High (top panels) and Mutation rate Low (bottom panels) based on pan-cancer median mutation rate.

Supplementary Figure 11: Multivariate Cox proportional hazards model including ICR classification, proliferation enrichment, TGF beta signaling enrichment, and tumor mutation rate.

Supplementary Figure 12: Survival analysis of ICR Low versus ICR High in pathway enrichment categories across 40 pathway signatures (rows) for each cancer type (columns). HRs (hazard ratios) for death in high enrichment categories (**A**) are compared with HRs in low enrichment categories (**B**). **C.** Differences in prognostic impact of ICR classification between pathway signature enrichment categories for each cancer type. HR of ICR Low vs. ICR High was calculated per category from binary classification of enrichment of oncogenic pathway signatures (rows) within individual cancer types (columns). The delta between HR

in the highly enriched group and the HR in the group with low enrichment was calculated for each signature/cancer type combination.

Supplementary Tables

Supplementary Table 1: Association of ICR with OS across 31 cancer types with ICR as a categorical variable ICR Low versus ICR High (first and second column; yellow), and ICR as continuous variable (third and fourth column; blue). HR, hazard ratio for death.

Supplementary Table 2: Comparison of mean ES of samples from ICR-disabled cancer types with mean ES of samples from ICR-enabled cancer types for 54 oncogenic pathway gene signatures.

Supplementary Table 3: Pan-cancer survival analysis stratified by binary classification based on enrichment of selected oncogenic pathway signatures. HR, hazard ratio for death.

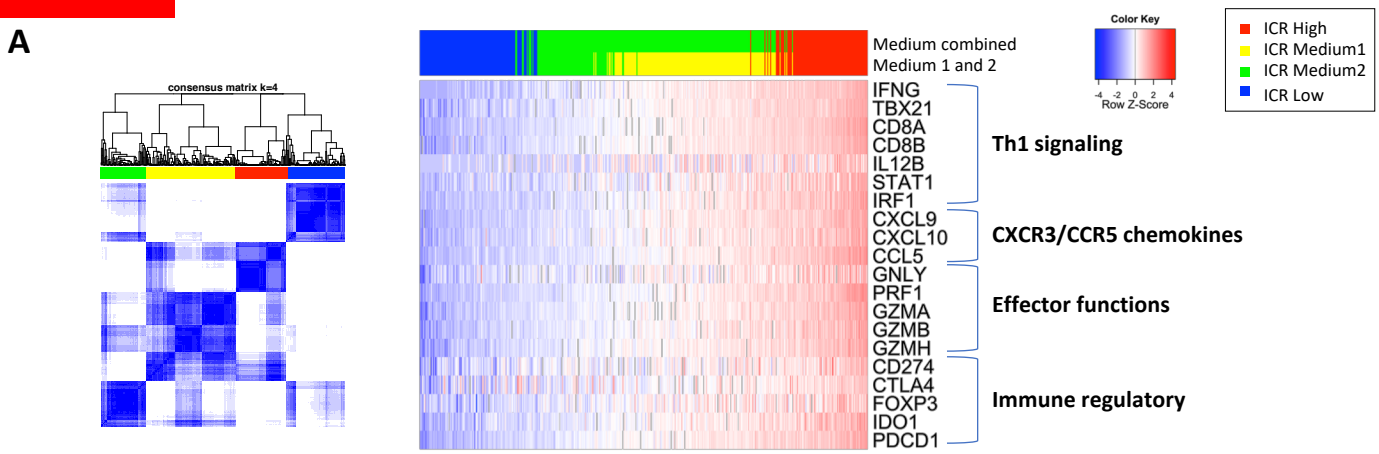
Table 1. Cox proportional hazards regression for association with overall survival in ICR-enabled and ICR-disabled tumors: ICR High and ICR Low samples included (ICR Medium samples excluded)

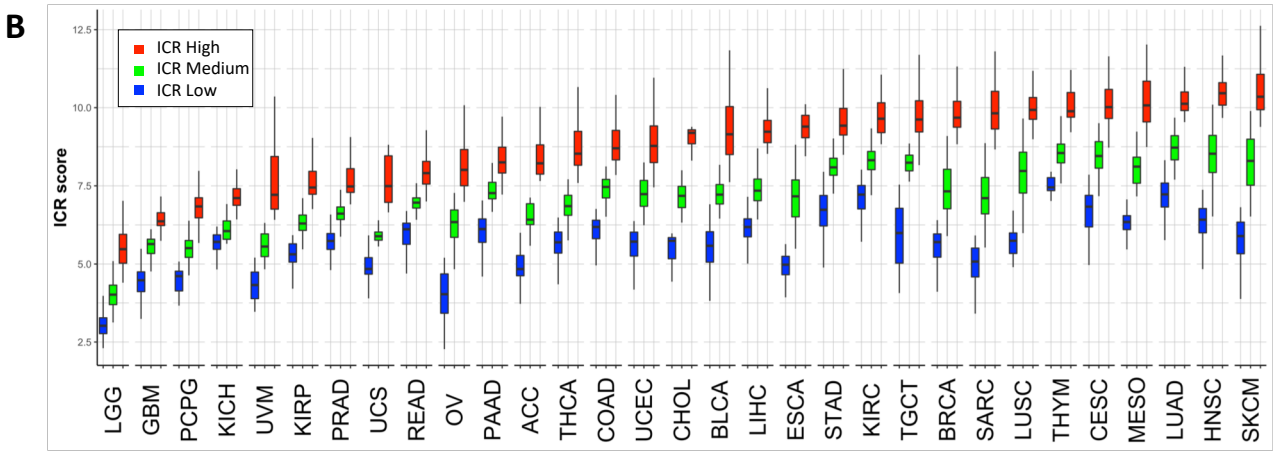
	Univariable		**Multivariable**	
Variables	**HR (95% CI)**	***p***	**HR (95% CI)**	***p***
ICR overall (n = 4735)				
~ ICR cluster (ICR Low vs. High)	1.203 (1.081-1.339)	0.00073 ***	1.343 (1.180- 1.528)	7.85e-06 ***
~ Pathologic stage (Stage I, II, III, IV)	1.72 (1.615-1.832)	<2e-16 ***	1.716 (1.611- 1.827)	<2e-16 ***
Samples from ICR-enabled cancer types (n = 1742)				
~ ICR cluster (ICR Low vs. High)	1.631(1.374-1.937)	2.26e-8 ***	1.488 (1.233- 1.795)	3.35e-05 ***
~ Pathologic stage (Stage I, II, III, IV)	1.817 (1.644-2.008)	<2e-16 ***	1.798 (1.628- 1.987)	<2e-16 ***
Samples from ICR-disabled cancer types (n = 721)				
~ ICR cluster (ICR Low vs. High)	0.6194 (0.4801-0.7992)	0.000229 ***	1.054 (0.7702- 1.443)	0.742
~ Pathologic stage (Stage I, II, III, IV)	1.55 (1.351-1.778)	4.22e-10 ***	1.560 (1.3520- 1.801)	1.19e-9 ***
Samples from ICR neutral cancer types (n = 2272)				
~ ICR cluster (ICR Low vs. High)	1.160 (0.983-1.369)	0.0789	1.336 (1.065- 1.676)	0.0122 *
~ Pathologic stage (Stage I, II, III, IV)	1.665 (1.5-1.848)	<2e-16 ***	1.640 (1.477- 1.821)	<2e-16 ***

Signif. codes: *** <0.001; ** <0.01; * <0.05

ICR cluster entered as categorical (factor) variable (factor levels: “ICR High”, “ICR Low”)
Pathologic stage as semi-continuous variable (Stage I = 1; Stage II = 2; Stage III = 3; Stage IV = 4)

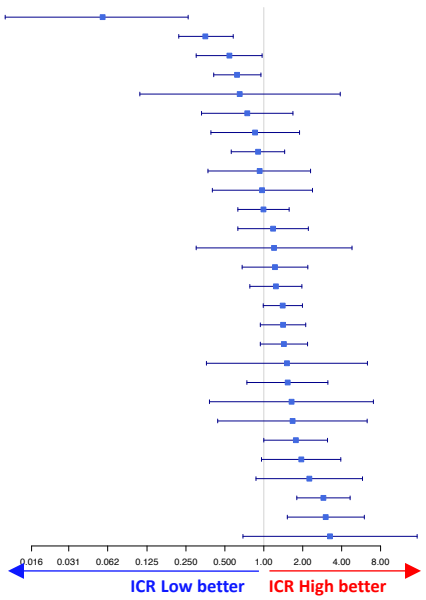
FIG1





C

	Cancer	N	N ICR High	N ICR Low	p–value	HR
ICR disabled	* UVM	80	28	30	0.000221	0.0567
	* LGG	513	101	189	3.29e–05	0.353
	* PAAD	156	23	82	0.0379	0.542
	* KIRC	515	134	139	0.0273	0.621
	PRAD	493	94	218	0.638	0.651
	KIRP	285	87	60	0.477	0.745
	MESO	87	11	25	0.7	0.856
	GBM	153	34	62	0.675	0.903
	ESCA	182	28	22	0.873	0.929
	UCS	57	13	27	0.946	0.969
	LUAD	510	70	234	0.978	0.994
	COAD	276	80	81	0.603	1.18
	READ	89	46	20	0.799	1.2
	LUSC	486	76	47	0.503	1.22
	OV	302	125	43	0.356	1.24
ICR enabled	* BLCA	404	148	146	0.0573	1.4
	* STAD	401	112	181	0.0994	1.41
	* HNSC	514	95	118	0.0949	1.43
	KICH	65	38	15	0.575	1.51
	CESC	301	99	68	0.249	1.53
	ACC	78	7	46	0.504	1.64
	CHOL	36	7	12	0.451	1.67
	* LIHC	368	59	92	0.0483	1.77
	* SARC	254	48	49	0.0643	1.95
	* UCEC	173	59	45	0.0946	2.25
	* SKCM	469	96	114	1.2e–05	2.89
	* BRCA	1082	152	266	0.00163	3.01
	THCA	500	128	159	0.136	3.25



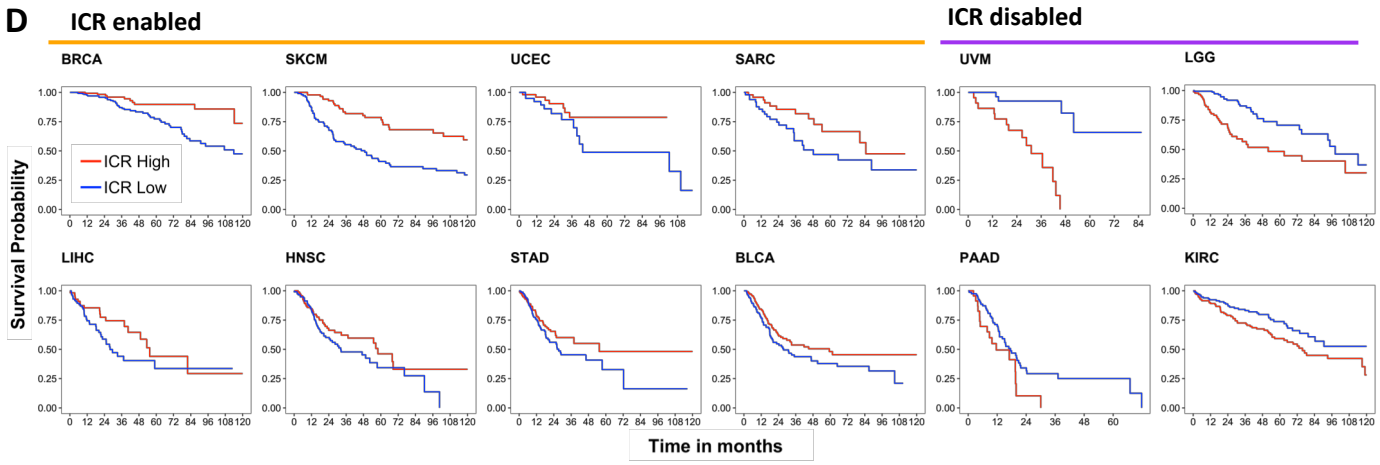
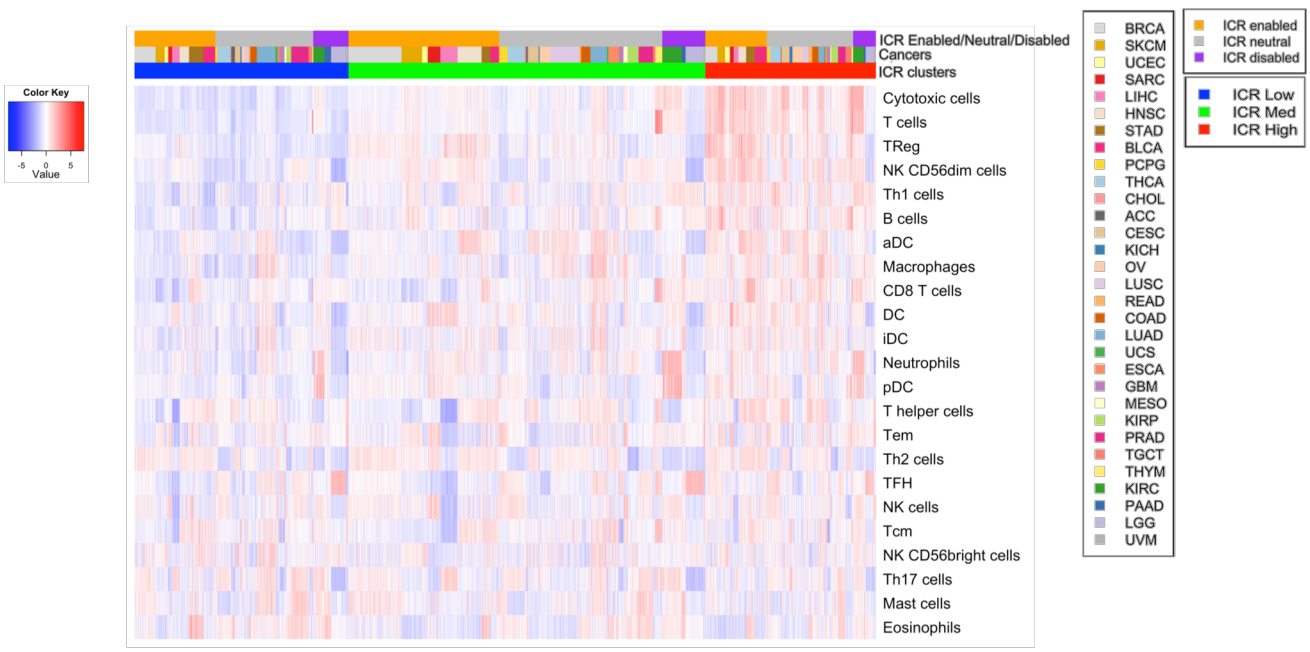


FIG2

A


Color Key
-5 0 5
Value
ICR Enabled/Neutral/Disabled
Cancers
ICR clusters
Cytotoxic cells
T cells
TReg
NK CD56dim cells
Th1 cells
B cells
aDC
Macrophages
CD8 T cells
DC
iDC
Neutrophils
pDC
T helper cells
Tem
Th2 cells
TFH
NK cells
Tcm
NK CD56bright cells
Th17 cells
Mast cells
Eosinophils
BRCA
SKCM
UCEC
SARC
LIHC
HNSC
STAD
BLCA
PCPG
THCA
CHOL
ACC
CESC
KICH
OV
LUSC
READ
COAD
LUAD
UCS
ESCA
GBM
MESO
KIRP
PRAD
TGCT
THYM
KIRC
PAAD
LGG
UVM
ICR enabled
ICR neutral
ICR disabled
ICR Low
ICR Med
ICR High

B

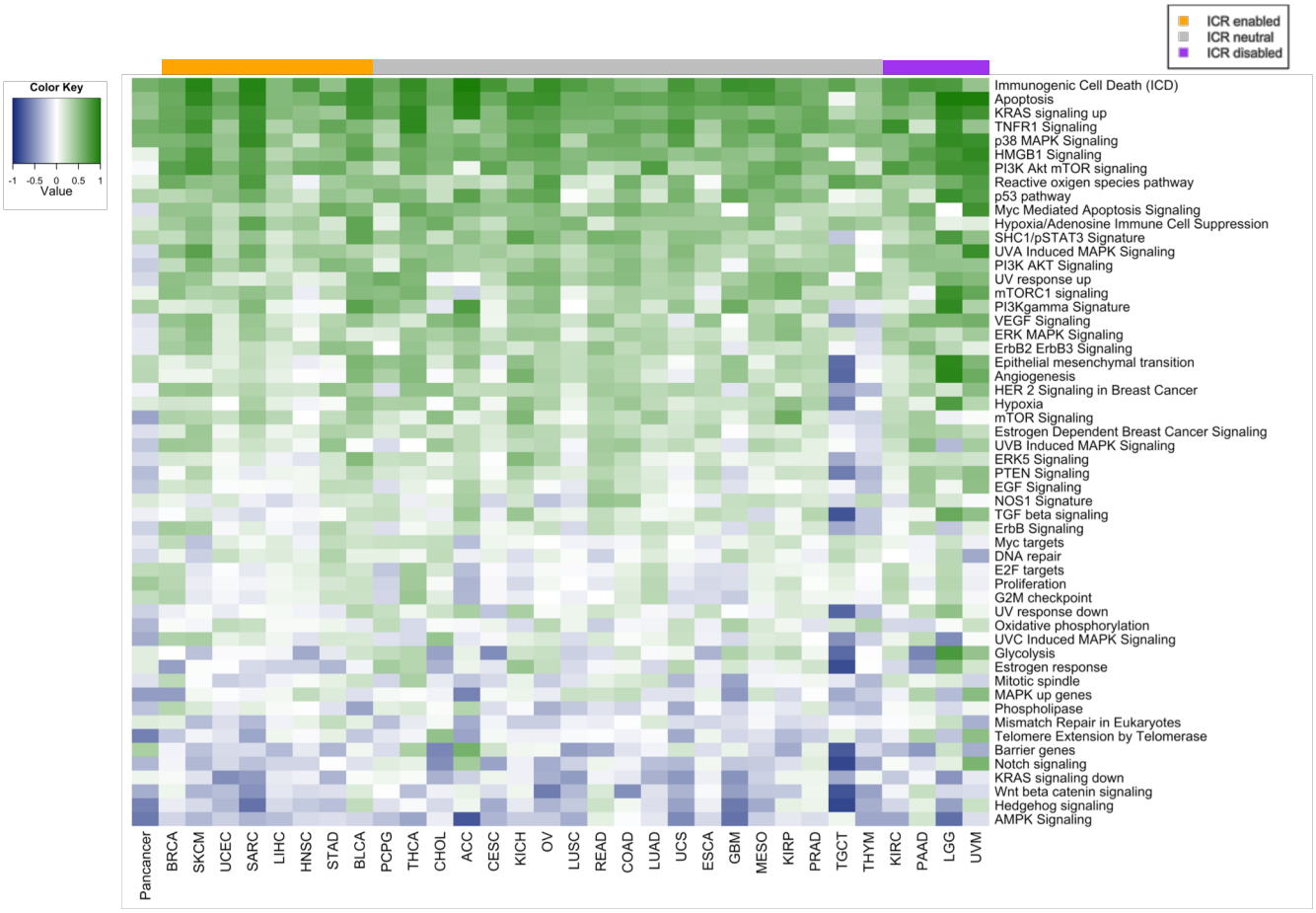
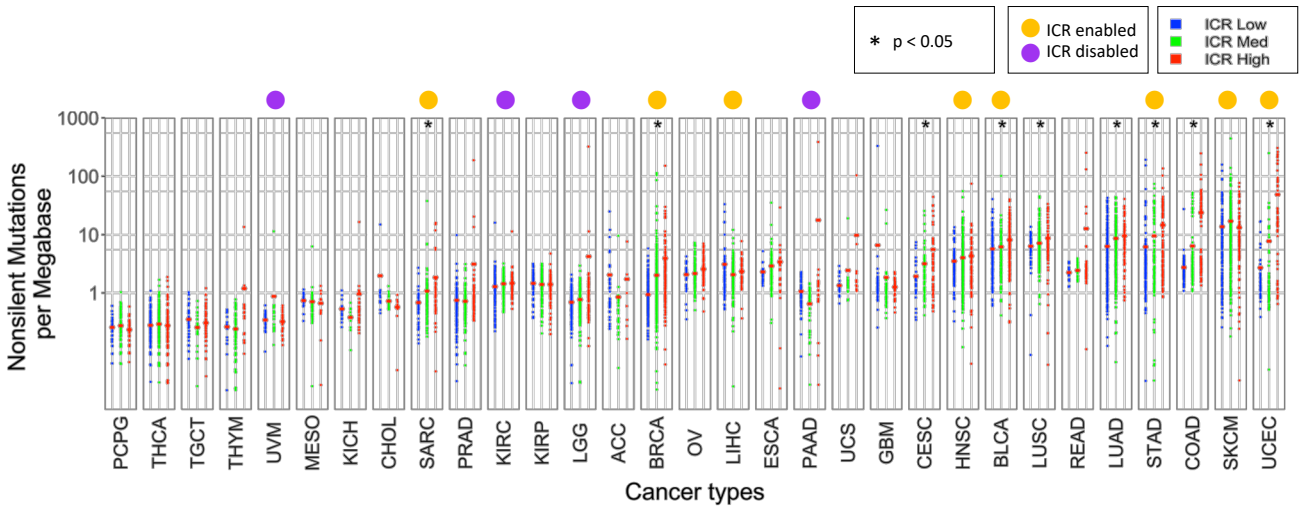
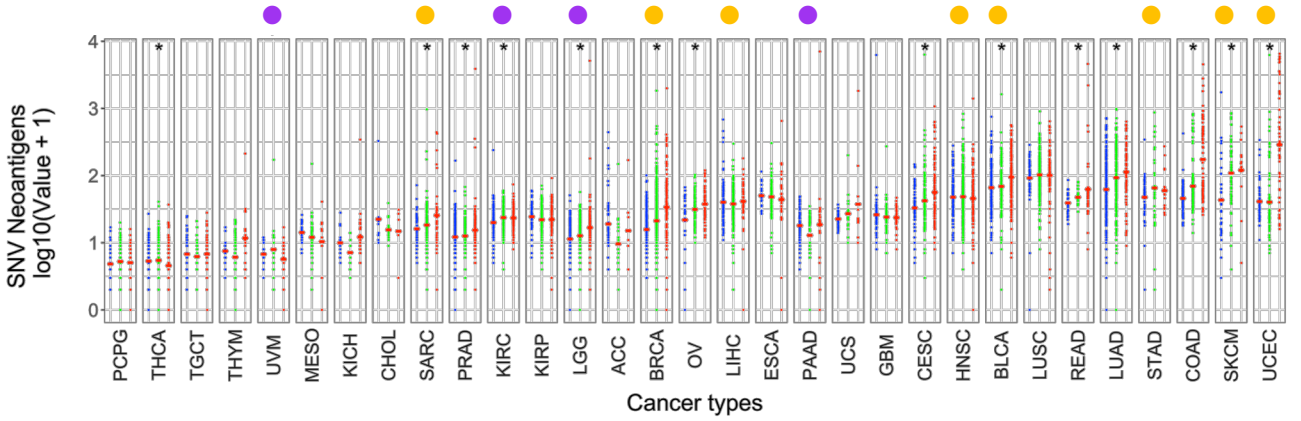

ICR enabled
ICR neutral
ICR disabled
Color Key
-1 -0.5 0 0.5 1
Value
Immunogenic Cell Death (ICD)
Apoptosis
KRAS signaling up
TNFR1 Signaling
p38 MAPK Signaling
HMGB1 Signaling
PI3K Akt mTOR signaling
Reactive oxigen species pathway
p53 pathway
Myc Mediated Apoptosis Signaling
Hypoxia/Adenosine Immune Cell Suppression
SHC1/pSTAT3 Signature
UVA Induced MAPK Signaling
PI3K AKT Signaling
UV response up
mTORC1 signaling
PI3Kgamma Signature
VEGF Signaling
ERK MAPK Signaling
ErbB2 ErbB3 Signaling
Epithelial mesenchymal transition
Angiogenesis
HER 2 Signaling in Breast Cancer
Hypoxia
mTOR Signaling
Estrogen Dependent Breast Cancer Signaling
UVB Induced MAPK Signaling
ERK5 Signaling
PTEN Signaling
EGF Signaling
NOS1 Signature
TGF beta signaling
ErbB Signaling
Myc targets
DNA repair
E2F targets
Proliferation
G2M checkpoint
UV response down
Oxidative phosphorylation
UVC Induced MAPK Signaling
Glycolysis
Estrogen response
Mitotic spindle
MAPK up genes
Phospholipase
Mismatch Repair in Eukaryotes
Telomere Extension by Telomerase
Barrier genes
Notch signaling
KRAS signaling down
Wnt beta catenin signaling
Hedgehog signaling
AMPK Signaling
Pancancer
BRCA
SKCM
UCEC
SARC
LIHC
HNSC
STAD
BLCA
PCPG
THCA
CHOL
ACC
CESC
KICH
OV
LUSC
READ
COAD
LUAD
UCS
ESCA
GBM
MESO
KIRP
PRAD
TGCT
THYM
KIRC
PAAD
LGG
UVM

FIG3

A


* p < 0.05
ICR enabled
ICR disabled
ICR Low
ICR Med
ICR High
Nonsilent Mutations per Megabase
1000
100
10
1
PCPG THCA TGCT THYM UVM MESO KICH CHOL SARC PRAD KIRC KIRP LGG ACC BRCA OV LIHC ESCA PAAD UCS GBM CESC HNSC BLCA LUSC READ LUAD STAD COAD SKCM UCEC
Cancer types

B


SNV Neoantigens log10(Value + 1)
4
3
2
1
0
PCPG THCA TGCT THYM UVM MESO KICH CHOL SARC PRAD KIRC KIRP LGG ACC BRCA OV LIHC ESCA PAAD UCS GBM CESC HNSC BLCA LUSC READ LUAD STAD COAD SKCM UCEC
Cancer types

C

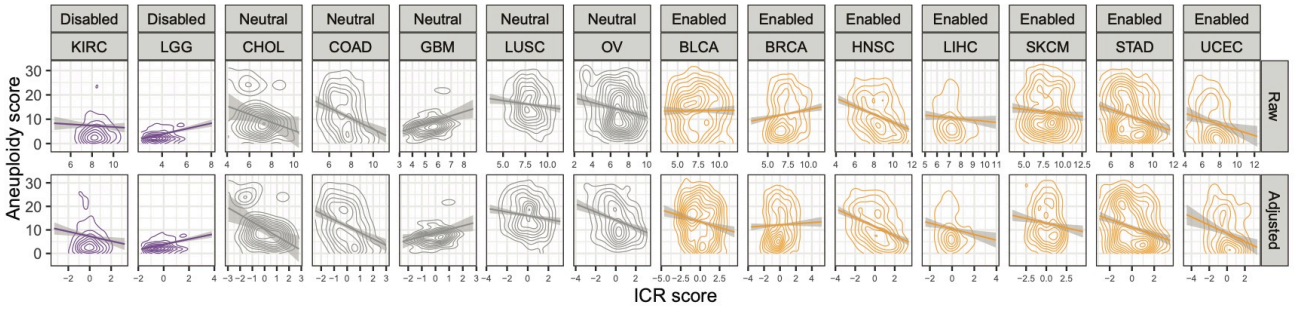
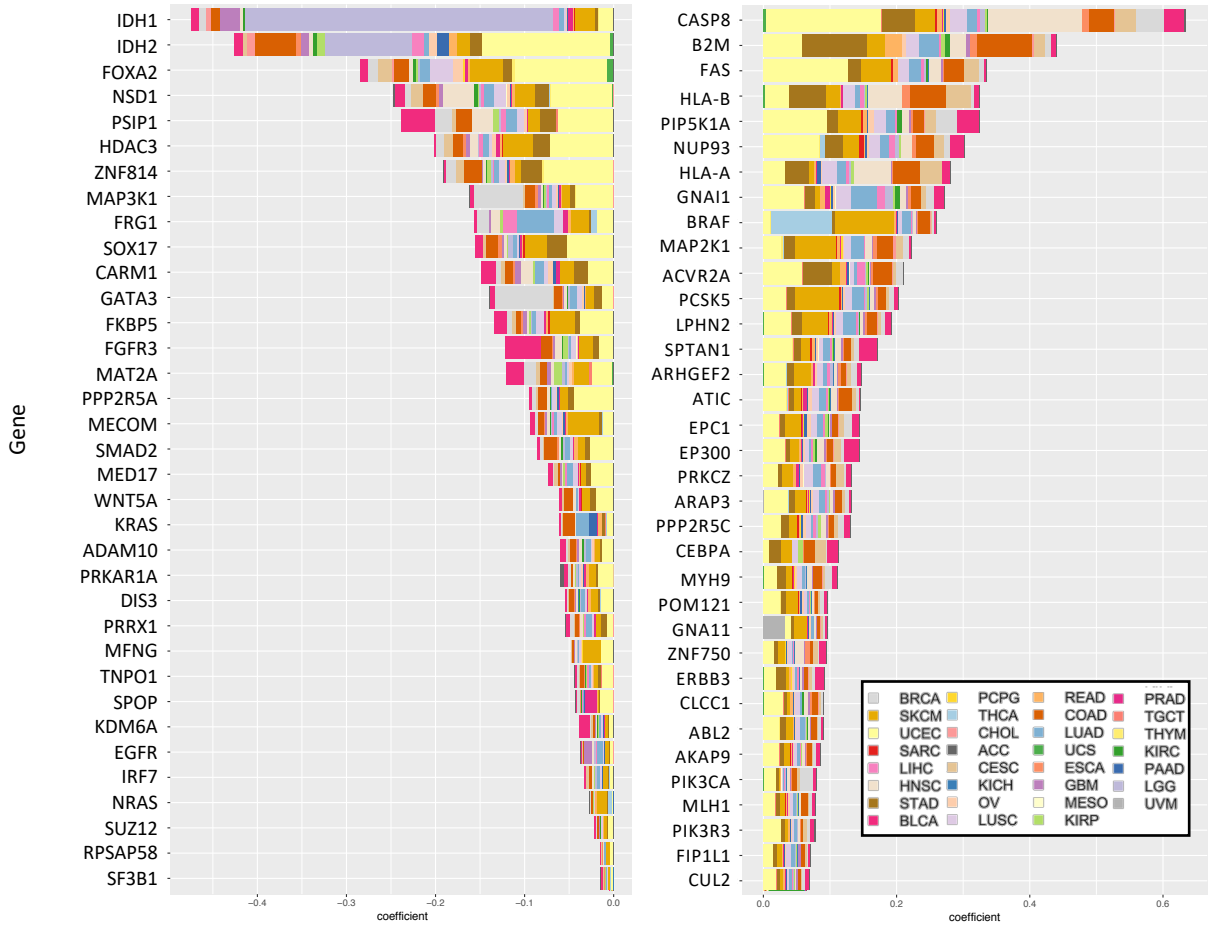

Disabled KIRC
Disabled LGG
Neutral CHOL
Neutral COAD
Neutral GBM
Neutral LUSC
Neutral OV
Enabled BLCA
Enabled BRCA
Enabled HNSC
Enabled LIHC
Enabled SKCM
Enabled STAD
Enabled UCEC
Raw
Adjusted
Aneuploidy score
ICR score

FIG4

A


Gene
IDH1
IDH2
FOXA2
NSD1
PSIP1
HDAC3
ZNF814
MAP3K1
FRG1
SOX17
CARM1
GATA3
FKBP5
FGFR3
MAT2A
PPP2R5A
MECOM
SMAD2
MED17
WNT5A
KRAS
ADAM10
PRKAR1A
DIS3
PRRX1
MFNG
TNPO1
SPOP
KDM6A
EGFR
IRF7
NRAS
SUZ12
RPSAP58
SF3B1
−0.4
−0.3
−0.2
−0.1
0.0
coefficient
CASP8
B2M
FAS
HLA-B
PIP5K1A
NUP93
HLA-A
GNAI1
BRAF
MAP2K1
ACVR2A
PCSK5
LPHN2
SPTAN1
ARHGEF2
ATIC
EPC1
EP300
PRKCZ
ARAP3
PPP2R5C
CEBPA
MYH9
POM121
GNA11
ZNF750
ERBB3
CLCC1
ABL2
AKAP9
PIK3CA
MLH1
PIK3R3
FIP1L1
CUL2
0.0
0.2
0.4
0.6
coefficient
BRCA
SKCM
UCEC
SARC
LIHC
HNSC
STAD
BLCA
PCPG
THCA
CHOL
ACC
CESC
KICH
OV
LUSC
READ
COAD
LUAD
UCS
ESCA
GBM
MESO
KIRP
PRAD
TGCT
THYM
KIRC
PAAD
LGG
UVM

B

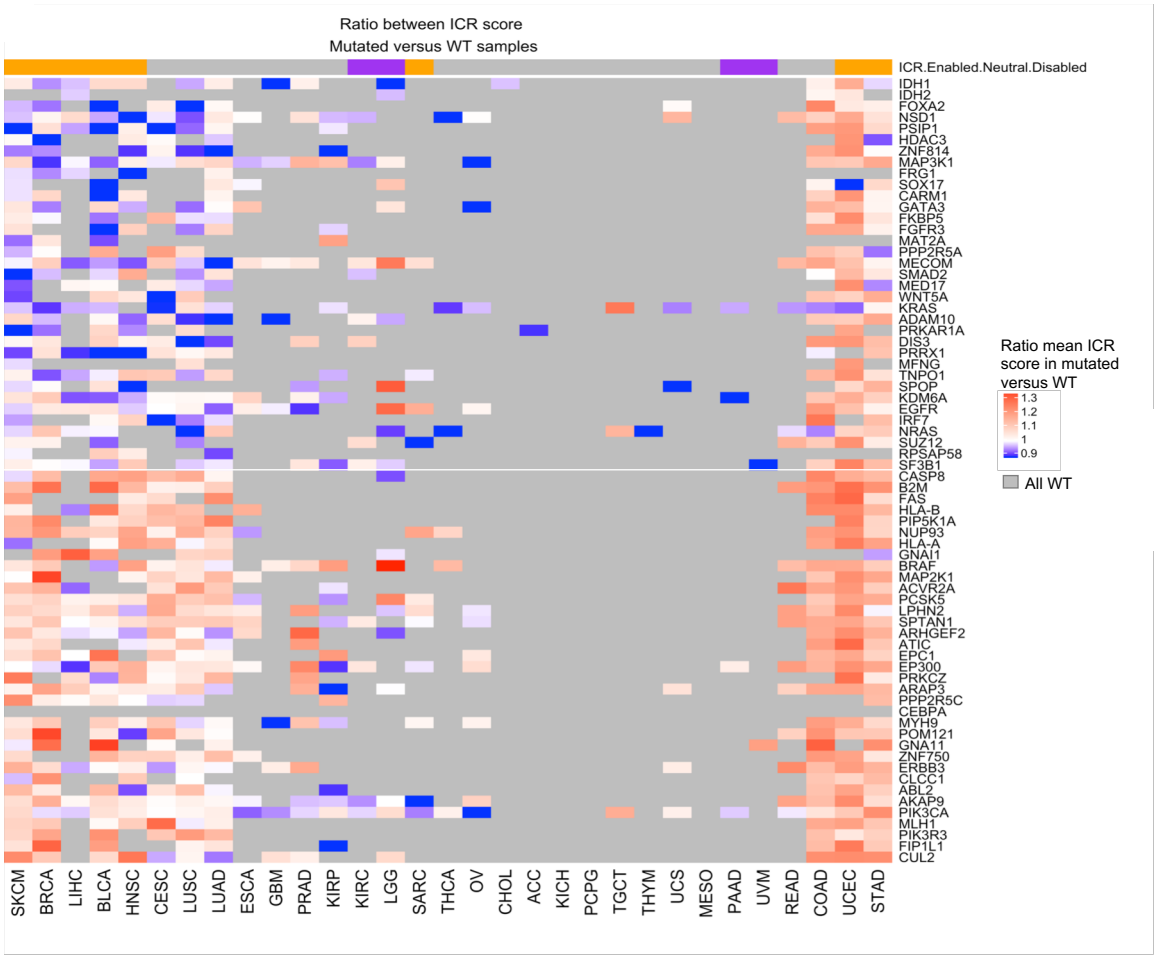
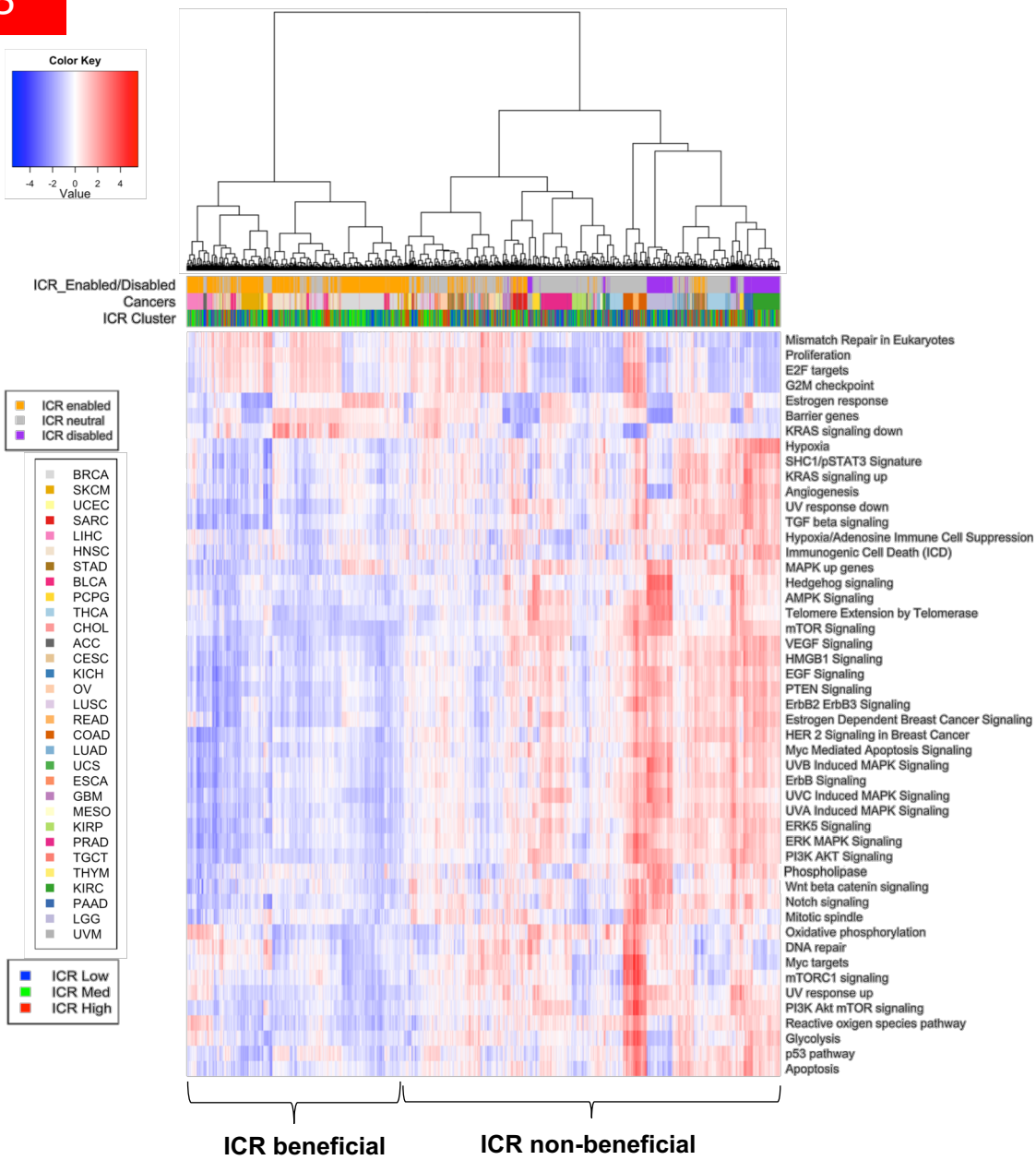
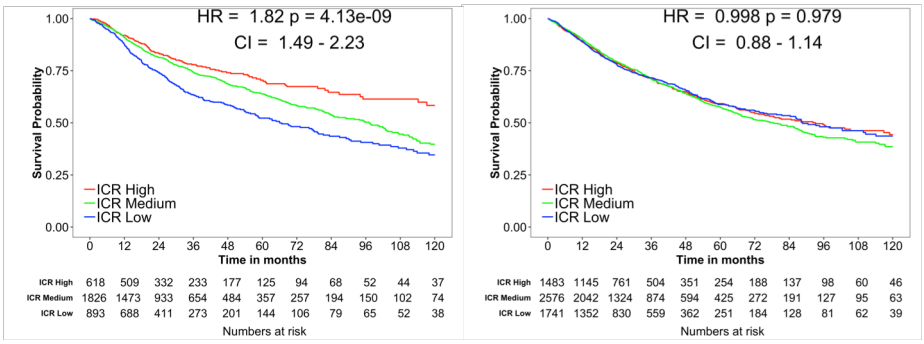

Ratio between ICR score
Mutated versus WT samples
ICR.Enabled.Neutral.Disabled
IDH1
IDH2
FOXA2
NSD1
PSIP1
HDAC3
ZNF814
MAP3K1
FRG1
SOX17
CARM1
GATA3
FKBP5
FGFR3
MAT2A
PPP2R5A
MECOM
SMAD2
MED17
WNT5A
KRAS
ADAM10
PRKAR1A
DIS3
PRRX1
MFNG
TNPO1
SPOP
KDM6A
EGFR
IRF7
NRAS
SUZ12
RPSAP58
SF3B1
CASP8
B2M
FAS
HLA-B
PIP5K1A
NUP93
HLA-A
GNAI1
BRAF
MAP2K1
ACVR2A
PCSK5
LPHN2
SPTAN1
ARHGEF2
ATIC
EPC1
EP300
PRKCZ
ARAP3
PPP2R5C
CEBPA
MYH9
POM121
GNA11
ZNF750
ERBB3
CLCC1
ABL2
AKAP9
PIK3CA
MLH1
PIK3R3
FIP1L1
CUL2
SKCM
BRCA
LIHC
BLCA
HNSC
CESC
LUSC
LUAD
ESCA
GBM
PRAD
KIRP
KIRC
LGG
SARC
THCA
OV
CHOL
ACC
KICH
PCPG
TGCT
THYM
UCS
MESO
PAAD
UVM
READ
COAD
UCEC
STAD
Ratio mean ICR score in mutated versus WT
1.3
1.2
1.1
1
0.9
All WT

FIG5

A

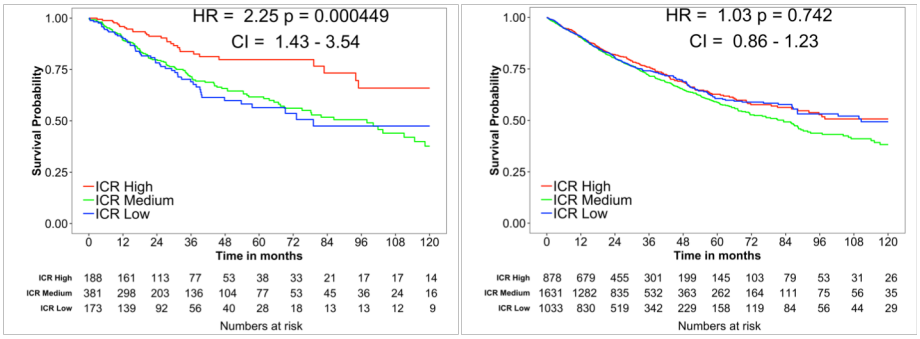


B



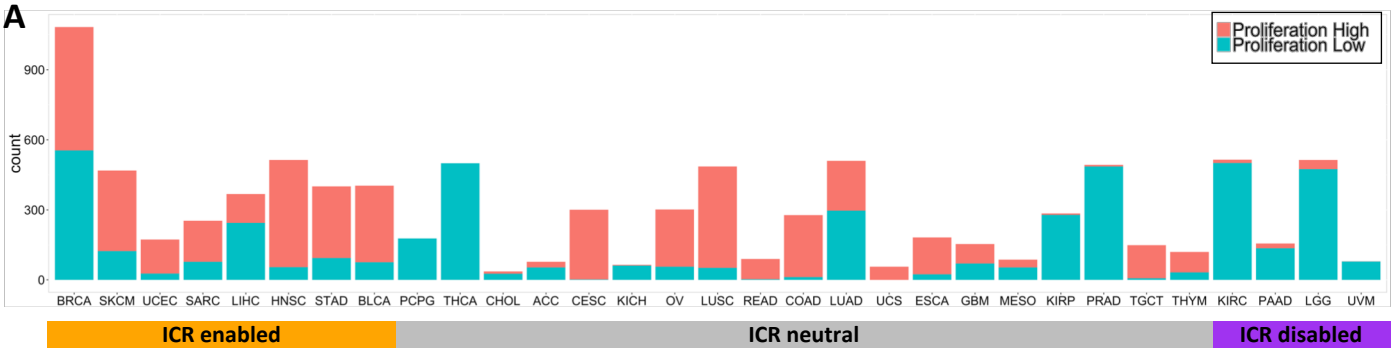
All cancer types

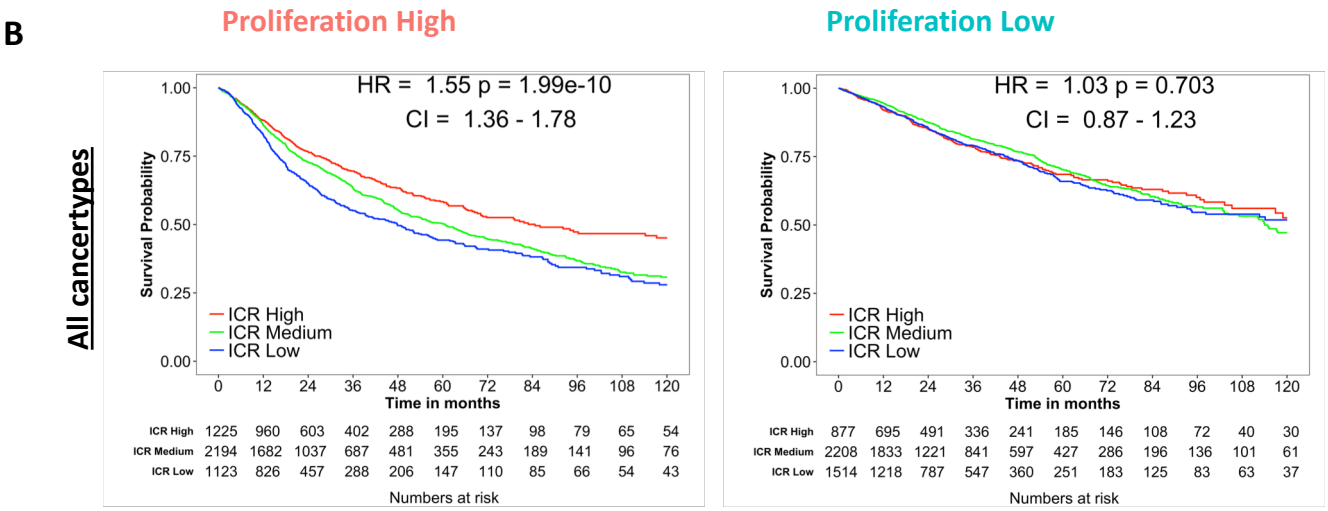
C

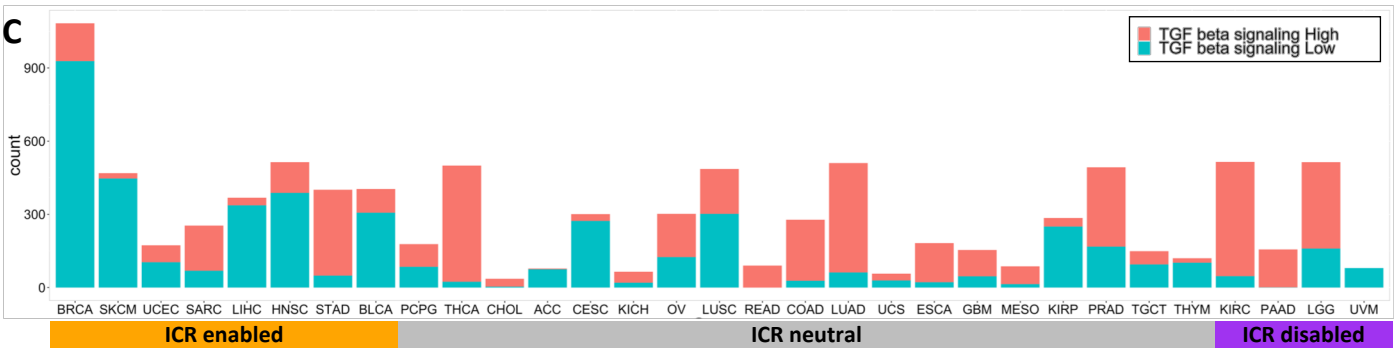


ICR neutral cancer types

FIG6


A
Proliferation High
Proliferation Low
count
900
600
300
0
BRCA SKCM UCEC SARC LIHC HNSC STAD BLCA PCPG THCA CHOL ACC CESC KICH OV LUSC READ COAD LUAD UCS ESCA GBM MESO KIRP PRAD TGCT THYM KIRC PAAD LGG UVM
ICR enabled
ICR neutral
ICR disabled


B
Proliferation High
Proliferation Low
All cancertypes
HR = 1.55 p = 1.99e-10
CI = 1.36 - 1.78
HR = 1.03 p = 0.703
CI = 0.87 - 1.23
Survival Probability
1.00
0.75
0.50
0.25
0.00
ICR High
ICR Medium
ICR Low
0 12 24 36 48 60 72 84 96 108 120
Time in months
ICR High 1225 960 603 402 288 195 137 98 79 65 54
ICR Medium 2194 1682 1037 687 481 355 243 189 141 96 76
ICR Low 1123 826 457 288 206 147 110 85 66 54 43
Numbers at risk
ICR High 877 695 491 336 241 185 146 108 72 40 30
ICR Medium 2208 1833 1221 841 597 427 286 196 136 101 61
ICR Low 1514 1218 787 547 360 251 183 125 83 63 37
Numbers at risk


C
TGF beta signaling High
TGF beta signaling Low
count
900
600
300
0
BRCA SKCM UCEC SARC LIHC HNSC STAD BLCA PCPG THCA CHOL ACC CESC KICH OV LUSC READ COAD LUAD UCS ESCA GBM MESO KIRP PRAD TGCT THYM KIRC PAAD LGG UVM
ICR enabled
ICR neutral
ICR disabled

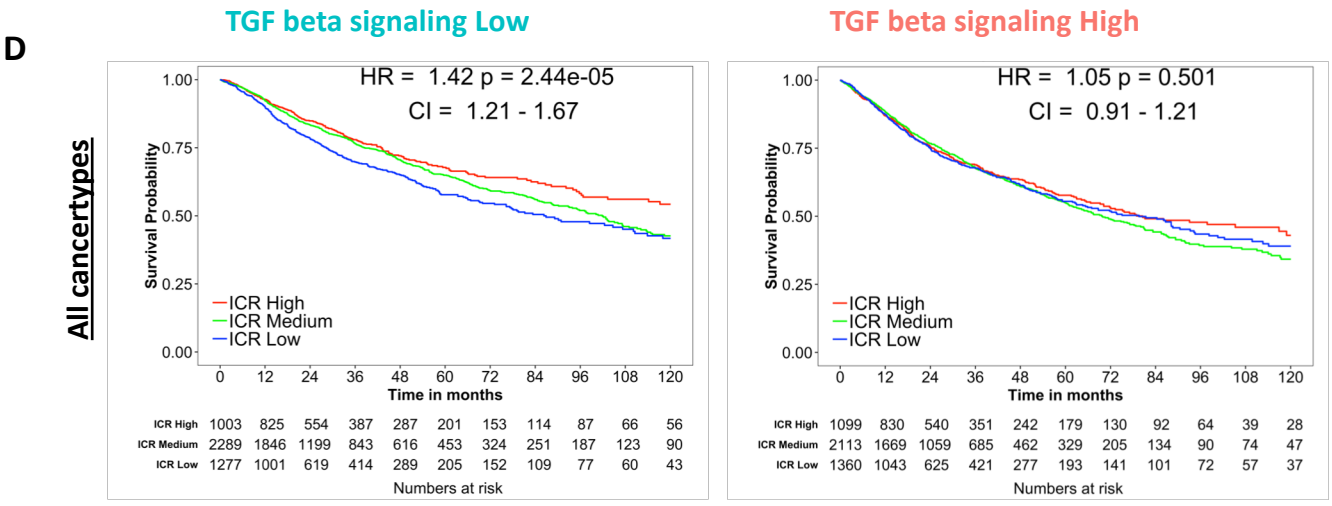
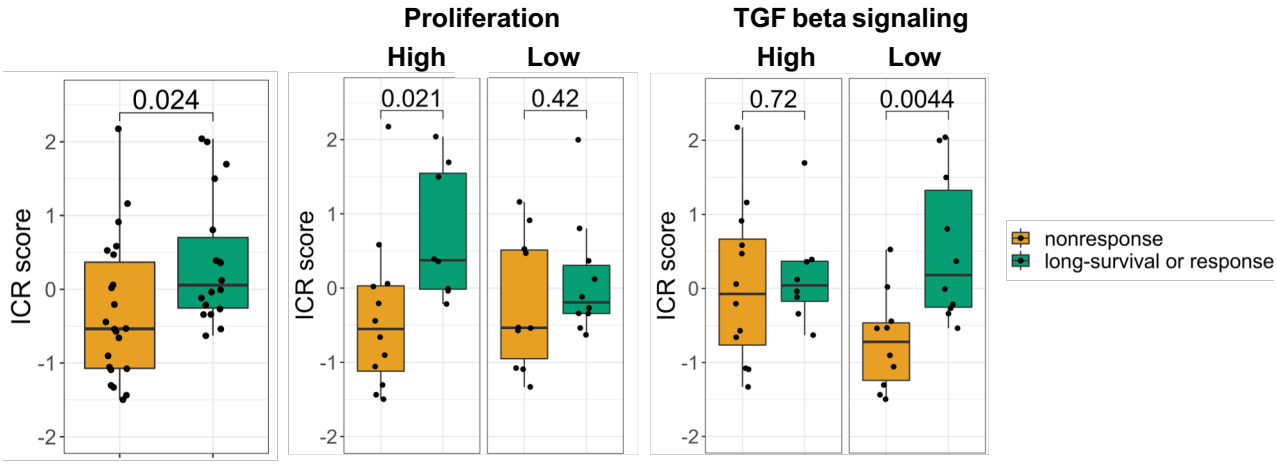
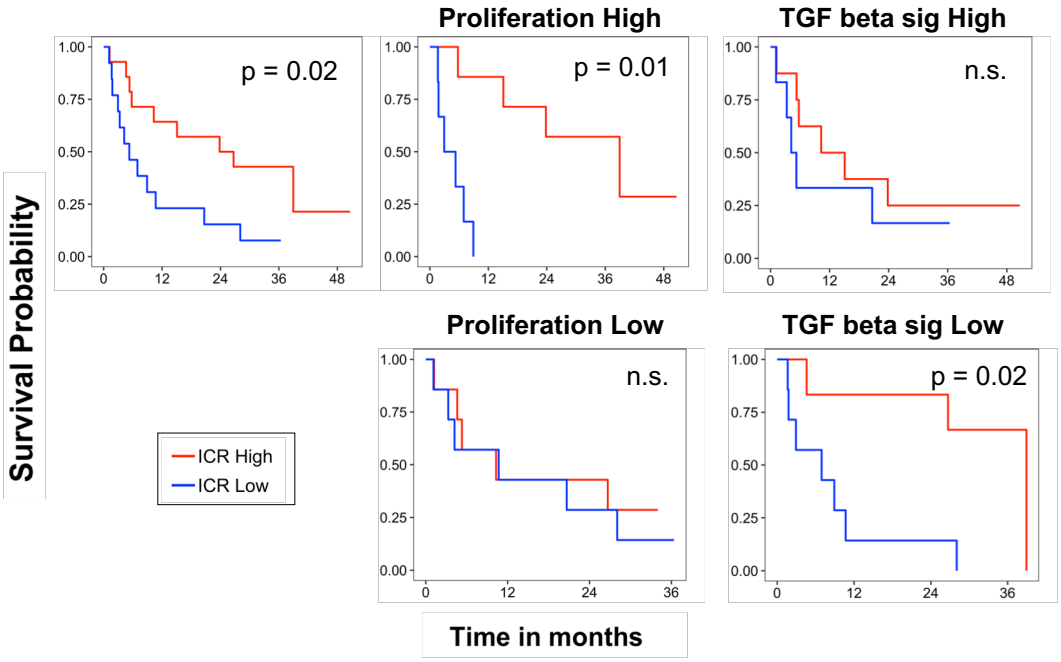

D
TGF beta signaling Low
TGF beta signaling High
All cancertypes
HR = 1.42 p = 2.44e-05
CI = 1.21 - 1.67
HR = 1.05 p = 0.501
CI = 0.91 - 1.21
Survival Probability
1.00
0.75
0.50
0.25
0.00
ICR High
ICR Medium
ICR Low
0 12 24 36 48 60 72 84 96 108 120
Time in months
ICR High 1003 825 554 387 287 201 153 114 87 66 56
ICR Medium 2289 1846 1199 843 616 453 324 251 187 123 90
ICR Low 1277 1001 619 414 289 205 152 109 77 60 43
Numbers at risk
ICR High 1099 830 540 351 242 179 130 92 64 39 28
ICR Medium 2113 1669 1059 685 462 329 205 134 90 74 47
ICR Low 1360 1043 625 421 277 193 141 101 72 57 37
Numbers at risk

FIG7

A

Study	Platform	Treatment type	Time of biopsy	ICR score in Responders > non-responders p value
van Allen	RNASeq	anti-CTLA4	pre-treatment	0.024 (*)
Chen et al	Nanostring	anti-CTLA4	pre-treatment	0.95
		anti-CTLA4	on-treatment	0.028 (*)
		anti-PD1	pre-treatment	0.69
		anti-PD1	on-treatment	0.00054 (***)
Riaz et al (GSE91061)	RNASeq	anti-PD1	pre-treatment	0.14
		anti-PD1	on-treatment	0.015 (*)
Hugo et al (GSE78220)	RNASeq	anti-PD1	pre-treatment	0.92
Prat et al	Nanostring	anti-PD1	pre-treatment	0.05 (*)

B



C