Prioritization of prostate cancer to immune checkpoint therapy by ranking tumors along IFN-γ axis and identification of immune resistance mechanisms

Boris Reva[1*], Tatiana Omelchenko[3], Anna Calinawan[1], Sujit Nair[2], Eric Schadt[1,4], AshTewari[2]

[1]Department of Genetics and Genomic Sciences at Icahn School of Medicine at Mount Sinai, New York, 10029, USA
[2]Department of Urology at Icahn School of Medicine at Mount Sinai, New York, 10029, USA
[3]Cell Biology Program, Sloan Kettering Institute at Memorial Sloan Kettering Cancer Center, New York, 10065, USA
[4]Sema4, a Mount Sinai venture, New York, 10029, USA

*To whom correspondence should be addressed: Tel: +1-212-824-9663; Fax: +1-212-824-9699; Email: boris.reva@mssm.edu

Abstract

We propose priotirization of prostate cancer patients to PD-L1 checkpoint therapy by assessing activity of IFN-γ/PD-L1 signaling in tumor from transcriptional profile. To this end, we introduced a new approach for inferring pathway activity and suppression (IPAS) by assessing significance of positioning pathway's genes expression levels at top (activation) or bottom (suppression) in gene expression profile of a given tumor. By ordering tumors along IFN-γ/PD-L1 axis, we determined distinct "IFN-γ-depleted" and "IFN-γ-enriched" immune subtypes, genes involved in immune evasion and potential targets for combination therapy. Using IPAS scoring method, we proposed biomarker panels for accurate ranking tumors along IFN-γ/PD-L1 axis.

1. Introduction

The diversity of cancer driving genomic alterations and related tumor microenvironment is the biggest obstacle in development and identifying efficient cancer therapy[1]. To treat cancer, it is necessary to have many highly specific drugs and developed accurate diagnostics for targeted assignment of particular drugs or, more likely, combinations of drugs. Molecular profiling of tumors (TCGA [2]) can provide the most complete data for construction of clinically relevant diagnostics, however, the molecular biomarkers for diagnostics can hardly be learned directly from even large clinical trials, because of the low success rate of the application of the specific drugs to tumors driven by diverse molecular alterations. This is especially true for emerging targeted immune therapies, which show fantastic results for a small percentage of patients and fails for the majority of cases [3,4]. Despite multiple testing efforts and despite the fact that about 8% of primary tumors and 32% of metastatic castration-resistant tumors express PD-L1 [5], prostate cancer remains resistant to immunotherapy with PD-1/PD-L1 or CTLA-4 blockade (immune checkpoint blockade) [6-10]. It has been proposed that combination therapy, including targeting potential compensatory mechanisms in immunotherapy-resistant prostate tumors, can potentially assist in the reactivation of anti-tumor immune response [11,12] . Thus, it becomes clear that a success of treatment depends greatly on developed diagnostics for accurate prediction of a sensitivity of a given tumor to a particular therapy. However, it still remains unclear how to identify tumors, which will likely respond to checkpoint therapy, how to identify tumors, which will respond to available combination therapies, and what new combination therapies have to be developed to treat the remaining tumors?

To fill this clinical gap, we propose development of a hypothesis-driven diagnostic framework, which combines biological knowledge, analysis of large sets of cancer molecular profiles and prediction algorithms that can be applied to molecular data of a particular tumor. In this approach, molecular data of tumors linked to the results of clinical trials can be used to test hypotheses, and facilitating the identification of tumors sensitive to a particular targeted therapy. Currently, there is no accurate analytic protocol for determining which tumors will respond to checkpoint inhibition, and also for determining what combination therapy should be used to sensitize a tumor to treatment.

In this study, we relied on the assumption that the sensitivity of prostate cancer tumors to immune therapy can be deduced from expression levels of genes of the relevant molecular pathways. We assumed that computational analysis of large RNAseq datasets of primary prostate cancers of TCGA [2] and recently published metastatic set of MSKCC [13] can help narrow down a list of potential biomarkers that can be tested for prioritization of tumors for immune therapy. Specifically, in the development of the computational protocol for constructing molecular diagnostics for checkpoint immune therapy, we focused on the following questions: - How should tumors be ranked and optimally classified into potentially “sensitive” and “non-sensitive” groups? What are the relevant genes for assessing tumor sensitivity to a particular therapy? Should one use only a target gene or an entire molecular pathway? If a molecular pathway is chosen, what is the optimal scoring function for assessment of an activity of this pathway? And, finally, how can tumor classification be performed in practical setting, i.e. what are biomarkers and classification signatures that can be tested in practice? Answering these questions resulted in development the computational protocol for ranking tumors by sensitivity to checkpoint immune therapy; application of this protocol to molecular profiles of prostate cancer is the focus of this work.

2. Results

IFN-γ/PD-L1 biological axis. Multiple studies showed that the immune responsiveness of tumor cells to immune checkpoint blockade is driven by interferon-gamma (IFN-γ) signaling [14,15]. Interferon-gamma (IFN-γ) is a cytokine released by activated T cells into the tumor microenvironment (TME) in response to neoantigens presented on tumor cells [11,16]. IFN-γ has both immunostimulatory and immunosuppressive functions that modulate the TME [17]. Disruption of the balance between stimulatory and suppressive mechanisms of IFN-γ signaling can result in tumor immune evasion [18]. Loss-of-function and genomic mutations in IFN-γ signaling genes were found associated with immune checkpoint blockade in clinical studies [19-21] and in genetic screens [22-24]. Chronic IFN-γ signaling in tumor cells increased resistance to immune checkpoint blockade by multiple inhibitory pathways [25].

The IFN-γ signaling is driven by binding of IFN-γ from the microenvironment to tumor IFN-γ receptor subunits IFNGR1 and IFNGR2, inducing expressions of MHC-I [26] and PD-L1 (CD274) [27] (Fig. 1). Binding of IFN-γ to its receptor results in recruitment and activation of the Janus kinases, JAK1 and JAK2, and subsequent phosphorylation, dimerization, and activation of a transcription factor known as signal transducer and activator of transcription, STAT1. STAT1 homodimers then translocate to the nucleus, where they bind to specific promoter elements and modulate transcription of IFN-γ-regulated genes [15,16,28].

Insert Fig.1 here

Thus, based on the key role of IFN-γ in modulation of immune and inflammatory responses and tumor immunosurveillance, we chose “IFN-γ/PD-L1 axis” as the primary pathway controlling tumor sensitivity to PD-L1/2 therapy - (Fig.1) We settled with this signaling pathway because it has all necessary components – a cytokine mediator, receptors, signal transduction, transcription factors and final effectors (PD-L1/2, MHC-I), and all those genes are expressed in cancer cells. We hypothesized that non-responder tumors are either tumors with downregulation of IFN-γ axis genes, which make them invisible to immune cells, or tumors with upregulated IFN-γ axis genes and another specifically activated mechanism of immune evasion preventing cancer cell death. Then, ranking tumors by activation of IFN-γ axis enabled us to classify tumors by sensitivity to checkpoint inhibition.

IFN-γ axes and methods for inferring pathway activity. Comparison of axes of a single CD274, positive regulators of IFN-γ/PD-L1 pathway and combination of positive and negative regulators of IFN- γ/PD-L1 pathway are presented in Table 1.

Insert Table 1 here

We used three methods for inferring pathway activity across tumor samples: averaged expression of pathways genes and two enrichment scores ssGSEA and IPAS (Supplement 1). A combination activity of a pathway with positive and negative components was assessed as a difference between independently assessed activities of positive and negative components. For better comparisons, we put together correlation distributions computed for both primary and metastatic tumors. As a reference axis, we took the ordering of tumors by expression level of CD274 genes. There is a drastic difference between primary and metastatic tumors revealed by a number of genes correlated with expression level of CD274. With an increase of correlation threshold, only a few highly correlated genes remain in metastatic cancers as compared to

hundreds in primary tumors. For positive regulators, the enrichment scores have significantly more correlations as compared to CD274 axis and average expression axis for both primary and metastatic tumors. Surprisingly, CD274 axis produced more correlated genes as compared to the axis formed by averaging expression levels. The IPAS axis produced more highly correlated genes as compared to ssGSEA axis for both primary and metastatic tumors. The similar trend is seen in combination of scoring functions computed for positive and negative components of IFN-γ/PD-L1 pathway. However, the decrease of highly correlated genes is stronger for ssGSEA score as compared to the averaging score.

Thus, ranking tumors along IFN-γ/PD-L1 signaling axis using the pathway enrichment scores revealed more correlated genes and, in that way, more biological differences between tumors as compares to ranking tumors along the axis of CD274 expression level or the axis formed by averaging expression levels. Thus, the pathway scoring functions can be more informative for diagnostics as compared to the use of single biomarkers. The IPAS score revealed more high correlated genes as compared to ssGSEA score. Therefore, in further analysis – revealing potential immune subtypes, characteristic genes and pathways (Fig.2) we used IPAS scores for ranking tumors along IFN-γ/PD-L1 signaling axis.

Determining immune subtypes. We proposed to determine disease subtypes by determining a point of the maximal difference between two tumor sets aligned along IFN-γ/PD-L1 signaling axis (Methods, Fig.6, Supplement 2). We assumed here that the true biologically different disease subtypes have to have the maximal difference between gene expression distributions. However, position of such a point of the maximal difference may depend on a number of genes used in computation of the combined score (Supplement 2). Therefore, following the protocol of Fig.6, we examined different gene sets filtered by the absolute value of correlation with the IFN-γ/PD-L1 axis score (Fig.2). For primary prostate tumors, when all genes were taken into account (correlation threshold 0); we found a distinct minimum, which separated the most significantly IFN-γ-depleted tumors from the rest (Fig.2A). With increase of the correlation threshold, the combined scoring function became rather shallow, while preserving position of the minimum (Fig.2A). Thus, for primary tumors, we considered two separation into IFN-γ-depleted ("cold") and IFN-γ enriched ("hot") subtypes. The first separation defines "truly" IFN-γ-depleted tumors; the second separation defines IFN-γ-enriched tumors. For metastatic prostate tumors, the scoring functions remained shallow and pointed to the same position of the minimum for all correlation thresholds up to 0.5 (Fig.2B). At correlation threshold of 0.6, position of the global minimum had shifted to right along IFN-γ axis and a new local minimum emerged in IFN-γ-enriched region. Because a number of correlated genes at threshold of 0.6 was small, we decided to choose for current analysis the consistent position of the minima obtained for correlation threshold of 0.0-0.5 as in Fig.2B.

Insert Fig.2 here

Identification of key genes, pathways and cell type signatures for primary prostate tumors. Prostate cancer is considered poorly immunogenic and characterized by low tumor infiltration with cytotoxic T-cells [29]. To gain insight into the mechanisms of low immunogenicity and immune evasion in prostate cancer, we analyzed genes and pathways in primary tumors with the highest increase of inferred activation of the IFN-γ/PD-L1 axis. We compared two groups ranked high or low along the IFN-γ/PD-L1 axis, which we labeled "IFN-γ-depleted" or "IFN-γ-enriched" tumor subtypes, respectively. From this analysis, we observed that IFN-γ-

enriched primary tumors are inflamed and enriched with genes and pathway involved in multiple immune evasion mechanisms.

In the IFN-γ-enriched subtype we obtained 426 genes and 714 pathways including prostate specific pathways: prostate cancer pathways (p = 9.1E-12), miRNA regulation of prostate cancer signaling pathways (p = 5.6E-06) (cutoff, 2-fold change (FC) with p < 0.00001). Among the top 200 genes the highest increase was identified in genes involved in signal transduction (30 genes), immune system (28 genes) and in inflammatory response (24 genes). In particular, we found genes mediating G-protein coupled receptor, cytokine- and chemokine-mediated signaling pathways (27 genes) and neutrophil degranulation (16 genes). Among the top 50 pathways the highest increase was detected in the immune system pathway, neutrophil degranulation, cytokine and GPCR signaling. The identified pathways were consistent with the functions of the top 200 upregulated genes demonstrating the inflammatory environment of the IFN-γ-enriched subtype.

We also identified upregulated immune checkpoints in IFN-γ-enriched tumors, known to reduce inflammation under normal physiological conditions and that are hijacked during tumor immune evasion [30]. We found increase in canonical immune checkpoints: CTLA4 inhibitory signaling pathway (p= 2.3E-10), CTLA4 gene (2.4FC, p = 0); PD-L1 expression and PD-1 checkpoint pathway in cancer (p = 0), genes CD274 (PD-L1) (2FC, p = 0) and PDCD1LG2 (PD-L2) (2FC, p = 0). Among alternative immune checkpoints, the most upregulated were TIGIT and BTLA (both 3FC, p = 0) and VISTA (VSIR) gene (2FC, p = 0); LAG3 was 2FC, p = 4.4E-13.

Insert Figure 3 here

Resistance to the immune checkpoint blockade is associated with the tumor microenvironment enriched in immune suppressive cells and cytokines [31]. We identified gene signatures of multiple cell types in the IFN-γ-enriched subtype with the highest upregulation in gene signature of fibroblasts (p = 1.84E-11, Fibroblasts Tirosh xcell complete signatures) and macrophages (p = 0, Macrophages Tirosh xcell complete signatures) with both populations of M1 (p = 0) and M2 (p = 7.19E-13). We found upregulation of immunosuppressive cytokines that recruit regulatory T cells (Tregs) and myeloid-derived suppressor cells (MDSC) (CXCR2 2.4FC, p = 8.88E-16; CXCR4, 2FC, p = 0; CCL22, 2.4FC, p = 0, CCL5, 2.5FC, p = 0) that was consistent with increase in regulatory T cell population (Tregs Charoentong xcell complete signatures, p = 0). We observed the prevalence of immunosuppressive cells over cytotoxic lymphocytes (CD8+ Tem BLUEPRINT 3 xcell, p = 0; CD8+ Tem Charoentong xcell complete signatures, p = 0; CD4+ Tem NOVERSHTERN 1 xcell, p = 0; CD4+ Tcm Charoentong xcell complete signatures, p = 0; NK cells FANTOM 1 xcell, p = 0; NK cells Rooney xcell complete signatures, p = 8.4E-12; Natural killer cell mediated cytotoxicity pathway, p =0).

Among known immunosuppressive pathways associated with resistance to immunotherapy through recruitment of immature dendritic cells, MDSC and Treg cells, we detected upregulation of VEGF signaling pathway [30] VEGFA-VEGFR2 Signaling Pathway, p = 2.57E-13, PRKCB (2FC, p = 3.3E-16), CCL2 (2.4FC, p = 0) and CXCL17 (4.8FC, p = 0), TGF-beta signaling pathway (p = 9.9E-09) and IL-10 anti-inflammatory signaling pathway (p = 0), IL10 (2FC, p = 4.4E-16). Significant upregulation of CXCL17 positioned this gene among the top three upregulated genes along the IFN-γ/PD-L1 axis. CXCL17 regulates T cell homeostasis and it is expressed under inflammatory conditions functioning as a chemoattractant for macrophages, dendritic cells and is capable of attenuating inflammation through the recruitment of myeloid derived suppressor cells [32,33].

Tumor cell-intrinsic signaling pathways also shape the tumor immune landscape [15]. We identified upregulation of IFN-γ responsive genes IDO1 (3.3FC, p = 0) and TDO2 (2.8FC, p = 3.1E-8), responsible for activation of MDSCs and suppression of T cells. As tumor inflammation causes hypoxia we detected significant upregulation in hypoxia (Hypoxia, p = 0) positioned among the top 50 activated pathways. One of the consequences of hypoxia is suppression of T cell function by upredulation of adenosine-mediated inhibitory pathway of immune evasion and consistent with this hypothesis, we identified upregulation of genes NT5E, ADORA2A.

Interestingly, upregulation of genes involved in development of the prostate gland and genitalia TP63, HOXD13, FGFR2 and Developmental Biology pathway in the IFN-γ-enriched subtype indicates on the involvement of tumor dedifferentiation and possible stemness mechanisms of immune evasion [15].

Alteration in iron availability and iron metabolism can be immunosuppressive leading to tumor immune evasion [34]. Among the top 15 genes of IFN-γ/PD-L1 axis, four of them were genes regulating iron transport: LTF, 7FC, p = 0; CP, 4.2FC, p = 0; LCN2, 3.8FC, p = 1.1E-16; ABP1 (AOC1), 3.7FC, p = 3.5E-11. Consistent with activation of the iron homeostasis pathways (Mtb iron assimilation by chelation (p = 0); iron uptake and transport (p = 0), LTF danger signal response pathway (p = 0), observed upregulation in the IFN-γ-enriched subtype indicates the iron depletion in the tumor microenvironment. Iron deficiency impairs immune response [35] including anti-tumor immunity [36]. Similar to the oxygen depletion (hypoxia), deficiency in iron activates HIF1-alpha and its responsive genes required for cell survival under stress [37]. As iron is needed for IFN-γ-induced tumor ferroptosis [38] it was proposed to target this pathway in combination with immune checkpoint blockade treatment.

These data confirm presence of established immune suppressive cell-mediated mechanisms of resistance to immunotherapy in the IFN-γ-enriched subtype and reveal genes with anti-inflammatory, immunoregulative and protective activities such as CXCL17 [33], LTF [39] and CP [36,37] as promising anti-tumor immune therapy targets in this subtype.

On the opposite side of the IFN-γ/PD-L1 axis, in the IFN-γ-depleted subtype we obtained 200 genes and 446 pathway (cutoff, 1.3-fold change (FC) with p < 0.00001). Among 200 genes, the highest increase (2.2 - 1.3FC) was in genes involved in translation (13 genes), rRNA processing (14 genes), translation initiation (13 genes), transcription (13 genes), oxidation-reduction process (8 genes) and signal transduction (8 genes). Pathway with the highest increase (18K - 6FC) were translation, RNA metabolism and cell cycle. We found a number of pathways driving translation with the initiation step in translation at the top of the list (Cap-dependent translation initiation pathway, p = 2.7E-06). The critical role of transcription initiation in prostate tumorigenesis is recognized and currently under investigation [40-44]. We identified multiple oncogenic factors that were translationally regulated in the IFN-γ-depleted subtype including cell cycle control and anti-apoptotic molecules (MYC targets V1 pathway (p = 7.4E-06), Cell Cycle pathway (p = 9.6E-07), genes MYC, NKX3.1, SOX9, TBX3, CDK19, PAK4, HPN).

Increased tumor cell oxidative metabolism is associated with repressed T cell function during T cell tumor infiltration and poor survival after anti-PD-1 immunotherapy in melanoma [45]. Consistent with the hypothesis that oxidative axis represents an important barrier to effective immunotherapy we found significant upregulation of oxidative phosphorylation in the IFN-γ-depleted subtype: oxidative phosphorylation pathway (p = 8.8E-04), mitochondrial biogenesis (p = 1.5E-11), respiratory electron transport (p = 2.6E-05).

Due to the narrow fold change range (2.2-1.5FC), arrangement by the highest average expression revealed the TARP gene (1.5FC, p = 9.0E-09). As a highly expressed protein in primary and metastatic prostate tumors, TARP was used as a tumor antigen to develop a successful vaccine that increased patient survival [46]. The other top list gene was MCCC2 (1.5FC, 5.0E-08), methylcrotonoyl-CoA carboxylase 2, which is involved in tumor formation [47,48].

We found a number of uncharacterized proteins with prognostic and therapeutic potential (genes: LOC100130872, C2orf72, C19orf48). In particular, C19orf48 encodes a multi-drug resistance protein overexpressed in androgen receptor - dependent manner in prostate cancer [49]. C19orf48, a minor histocompatibility antigen recognized by cytotoxic T cells was demonstrated to have an antitumor effects in kidney cancer patients [50].

We identified significant enrichment in the gene signature for pericytes (p = 8.1E-05). Due to expression of PD-L1/2 and immunosuppressive properties (NO, PGE2, and TGF-β), pericytes drive immune evasion in gliomas [51]. The angiogenic role of pericytes in prostate tumor microenvironment [52,53] in the IFN-γ-depleted subtype deserves further investigation.

Among top 200 genes with highest fold change we also identified upregulated genes involved in negative regulation of immune response (C1QTNF3), in activation-induced cell death of T cells (RPS6), negative regulation of macrophage chemotaxis (MIF), VEGF and angiogenesis (MAPK12, ADM2), that contribute to the phenotype of the IFN-γ-depleted tumors.

Characteristic genes and pathways obtained for metastatic tumors. Tumor-induced systemic inflammation has been associated with metastasis in multiple cancers including prostate cancer [54]. To understand how the status of IFN-γ affects immune evasion mechanisms of metastatic prostate cancer, we compared the transcriptomes of the two groups – metastatic tumors of the IFN-γ-enriched subtype with metastatic tumors of the IFN-γ-depleted subtype.

In the IFN-γ-enriched subtype in metastatic tumors we obtained 203 genes and 776 pathway (cutoff, 2-FC with p < 0.00001). The transcriptome of the top 200 genes of the IFN-γ-enriched subtype was dominated by genes regulating neutrophil degranulation (37 genes), which is consistent with an increased pool of neutrophils due to the systemic inflammation [55] and with the promoting role of neutrophils in metastatic spread [56]. Among the top 200 genes were also genes promoting tumor cell invasion and metastasis formation including cell adhesion (16 genes), proteolysis (15 genes), extracellular matrix organization and degradation (12 genes), leukocyte migration (11 genes), negative regulation of apoptosis (11 genes), cell-cell signaling (10 genes), cell migration (10 genes) and angiogenesis (10 genes). Among the top upregulated 50 pathways, we found significant increase in the fold change of the neutrophil degranulation (p = 0), cytokine signaling (p = 0), epithelial-to-mesenchymal transition pathway (p = 6.3E-11), antigen processing-cross presentation (p = 0), TCR signaling (p = 0), extracellular matrix organization (p = 7.3E-15).

Among established immune evasion mechanisms contributing to metastatic disease, we identified significant increase in selected immune checkpoints including PD-1 (PDCD1, 3.6FC, p= 1.9E-15), PD-L1 expression and PD-1 checkpoint pathway in cancer (p = 0), VISTA (VSIR) gene (3.2FC, p = 0) and in CD274 (PD-L1) (2.3FC, p = 2E-13). We observed upregulation of epithelial-to-mesenchymal transition genes TAGLN (2.1FC, p = 5.5E-8) and AXL (2.3FC, p =0) involved in resistance to the immune checkpoint blockade[57].

The features of TME of the metastatic IFN-γ-enriched subtype included enrichment in fibroblast gene signature (Fibroblasts Tirosh xcell complete signatures, p = 2.6E-11) and increase in macrophages (Macrophages Tirosh xcell complete signatures, p = 0) and monocytes (Monocytes FANTOM 2 xcell, p = 0) and their chemotactic genes CCL8 (3FC, p = 3.8E-13) and inhibitory cytokines (CXCR2, 3.3FC, p = 2.0E-08), CCL5 (5FC, p = 0). We also identified signatures of endothelial cells (Endothelial cells Tirosh xcell complete signatures, p = 1.6E-14), Th17 cells (Th17 cell differentiation pathway, p = 0), Tγδ cells (Tgd cells HPCA 3 xcell, p = 2.5E-11) and pericytes (p = 5.9E-06).

Upregulation of signaling by VEGF (p = 3.9E-13), VEGFA-VEGFR2 pathway (p = 7.9E-13), VEGFA gene (1.7FC, p =3.3E-06) and angiogenesis genes (TGFBI, CCL2, MMP14, MMP2, ANPEP, ECSCR, APLNR, HMOX1 and TNFAIP2, all 2-3FC with p < E-07) indicate immunosuppression mechanism by VEGF in addition to promotion of tumor angiogenesis [30].

In summary, these data reveal the active role of the immunosuppressive tumor microenvironment with the major role of the immunosuppressive myeloid cells in modulation of immune response in metastatic IFN-γ-enriched subtype of prostate tumors.

IFN-γ-depleted subtype in metastatic tumors we obtained 215 genes and 303 pathway (cutoff, 1.5-FC with p < 0.001). The top 200 upregulated genes were involved in transcription (39 genes), cell proliferation (27 genes), signal transduction (16 genes), oxidation-reduction process (12 genes), cell adhesion (12 genes) and hypoxia (8 genes). Among the top 50 pathways we identified metabolism of RNA (p = 3.8E-06), translation (p = 3.3E-05), transcription (p = 3.4E-11), cell cycle (p = 1.1E-03), mitochondrial biogenesis (p = 1.8E-08), DNA repair (p = 1.8E-08), ubiquitination (p = 6.7E-10) and cilia assembly (p = 2.6E-14).

We identified upregulation of the Wnt signaling pathway (WNT mediated activation of DVL pathway p = 1.8E-07), genes DVL2 (1.3FC, p = 1.2E-05), AXIN1 (1.3FC, p = 8.5E-05)), and its positive regulator SOX4 (1.7FC, p = 2.9E-08) and its mediator PAK4 (1.7FC, p = 7E-08) confirming the role of increased Wnt signaling in developing resistance to immune checkpoint blockade [15].

Upregulation of SOX9 gene (1.7FC, p = 6.1E-06) in the metastatic IFN-γ-depleted subtype is consistent with its role in promoting metastasis in prostate cancer [58]. SOX9 is an important regulator of prostate gland development and its overexpression in advanced lung adenocarcinoma metastasis confers resistance to natural killer-mediated cell killing [59] pointing toward SOX9-mediated evasion of immune response in the metastatic IFN-γ-depleted subtype. These data demonstrate the presence of immune evasion mechanisms in the metastatic IFN-γ-depleted subtype that are mediated by upregulation of Wnt signaling and stem-like program including SOX9.

We identified pericytes cell type signature (Pericytes ENCODE 3 xcell p = 0.5E-04)and mesenchymal stem cell signature (MSC HPCA 3 xcell, p = 2.4E-04). Identified neuronal signature (Neurons ENCODE 2 xcell, p = 8.0E-07) in the metastatic IFN-γ-depleted subtype indicates presence of perineural invasion[60], an indicator of aggressive prostate cancer[61,62].

Classification of prostate tumors into IFN-γ-enriched and IFN-γ-depleted subtypes. To assign immune therapy correctly and, to assign, when necessary, additional specific drugs to neutralize immune evasion, one needs accurate diagnostics. The results presented in previous sections showed significant difference between IFN-γ-enriched and IFN-γ-depleted tumors both in primary and metastatic tumors. Therefore, it is not difficult to propose a set of biomarker

genes for accurate classification of tumors by immune subtype or along IFN-γ/PD-L1 axis. However, the sequencing platforms and computational pipeline used to assess gene expression levels are subject to batch effect, while correction of batch effect needs large data sets that is problematic in practice. Therefore, typically, one cannot used directly expression levels of biomarker genes derived in one study to classification of transcriptional profiles produced in another study. We propose a solution to this problem by transition from profiles of gene expression levels to profiles of pathway activity scores. The multiple classification tests run with MSM [63] algorithm showed that ssGSEA and IPAS scoring methods, which assess collective up/down-regulation of pathway genes based on ranking genes in each of tumors, can be used in classification individual tumors given whole-genome transcriptional profiles (Supplement 3). However, getting whole-genome transcriptional profile for each of tumors, is not the best practical solution. - Can accurate diagnostics be developed with a smaller number of genes and with no complications caused by batch effect? – To this end, we propose biomarker panels composed of two gene sets, which are inversely regulated in IFN-γ-enriched and in IFN-γ-depleted subtypes, i.e. one "positive" gene set is upregulated in IFN-γ enriched and downregulated in IFN-γ-depleted subtype, another "negative" set is up-regulated in IFN-γ-depleted and down regulated in IFN-γ-enriched subtype. Thus, genes of such a panel can be treated as a pathway composed of positive and negative branches. Using the IPAS scoring method, we computed two scores, an activation score for a "positive set", and a suppression score for a negative set, and used these scores to feed in our classification algorithm [63]. Assessing classification accuracies, we simulated a practical situation, when based on a scoring value, one needs to classify a particular sample. In particular, we assessed accumulation of misclassification error independently for IFN-γ-depleted and IFN-γ enriched subtypes as functions of the classification score. Thus, for primary prostate tumors ~50% of tumors were classified correctly in IFN-γ-enriched and IFN-γ-depleted subtypes with error less than 5%; ~70% of tumors were classified correctly with error less than 10%;for metastatic tumors, ~70% and 80% of tumors were classified correctly with errors less than 5% and 10%, respectively. Taking into account that no special optimization was done in selection of biomarkers, we believe that the proposed approach has a significant potential for development accurate diagnostic panels with a small number of biomarker genes.

3. Methods

The schema of the computational protocol used in this work is presented in Fig.6. We proposed to order tumors along a "biological axis", which can be presented by an expression level of a target gene (in our case, CD274) or by inferred activity of a relevant molecular pathway (a pathway score) or by experimentally derived sensitivities to drug treatment (IC50, EC50). Then, we (i) rank different biological axes by a number of correlated genes; (ii) select a subset of correlated genes for (iii) determining two most distinct sets of tumors (potential disease subtypes) separated by a position on a biological axis. Two-class separation of tumors make possible to determine specific genes and pathways – potential biomarkers - to better characterization biological differences and determine biomarkers for classification of tumors along the axis (Supplement 2).

Insert Fig.6 here

In this approach, we assume that a "true" biological axis has to have the highest number of genes correlated with this axis; on the contrary, a biologically irrelevant axis correlates with any of the genes only by random. Therefore, for each of axes, we computed distributions of Pearson

correlations between the axis and gene expression levels. In computational analysis we used gene expression levels of primary tumors of TCGA [2] and metastatic tumors of MSKCC [13]. We tested three genes sets as clinically relevant axes and considered three types of pathway scoring. For nomination prostate tumors for PD-L1 therapy, we took the expression level of CD274 gene and inferred activities of the "IFN-γ/PD-L1 axis" pathway (Fig.1). This pathway is comprised of 29 genes, which can be classified as "positive" and "negative" regulators of CD274 expression. Assessing activity of the pathway, we used separately a subset of 17 positive regulators and combination of both positive and negative regulator genes (Fig.1).

The simplest way of scoring pathway activities is to compute a normalized sum of expression levels of all pathway genes; in more sophisticated approach – single sample gene set enrichment analysis (ssGSEA [64]), one computes an enrichment score by assessing a shift of the actual distribution of pathway genes from the uniform distribution towards top (or the bottom) in the ordered gene list levels [64]. In both of the approaches it is assumed that gene's role in a pathway is entirely depends on its expression level; no specific context of expression levels of other genes or activities of molecular pathways is taken into account. One of the problems in such assessments is a potential presence of two brunches of genes in a pathway – positive and negative regulators. In our current tests of the pathway scores, we simply took the difference between scores of two brunches. To infer pathway activation more accurately, we introduced an entropy based assessment of deviation from uniform distribution our new entropy based score – IPAS (Inferred Pathway Activity score, Supplement 1). The main idea of the approach is to assess deviation from uniform distribution by computing two probabilities: one probability assesses enrichment of pathway genes at the top of ordered list of gene expression levels, the other one – from the bottom. Both probabilities can be significant, in case that a pathway has both significantly upregulated and significantly downregulated genes. The other essential difference is that a pathway can be assessed with a high scoring value even when only a few of pathway genes are significantly above the median of gene distribution, while the majority of pathways genes can be below the median. Thus, our scoring method is genuinely biased towards those pathways which have truly top (or bottom) regulated genes. In this study, we tested all above pathway scoring methods to choose the most accurate method for ranking prostate cancer tumors by inferred sensitivity to PD-L1 therapy.

The important part of the computation protocol is selection of characteristic genes and pathways and derivation of classification signatures. To this end, we used our previously developed Molecular Signature Method [63] and software.

4. Discussion and Conclusion

In this work, we hypothesized that response to PD-L1 checkpoint therapy depends primarily on activation of IFN-γ/PD-L1 signaling pathway (IFNG axis). This hypothesis means that non-responder tumors are either tumors with downregulated IFNG axis genes, which became invisible to immune cells, or tumors with upregulated IFNG axis genes and specifically activated mechanisms of immune evasion that prevent cancer cell death.

To test this hypothesis, we ranked tumors by computationally assessed activity of IFNG axis and assessed the biological relevance of the axis computing correlations between gene expression levels and the IFNG axis score. The more genes correlate with the axis, the deeper biological separation between tumors aligned along the axis. We also used the IFNG axis scores to nominate potential disease subtypes by determining a point of the maximal difference between two tumor sets with lower and higher levels of the IFNG axis score. Thus, we classified

primary and metastatic prostate tumors into “IFN-γ-depleted” and “IFN-γ-enriched” subtypes along inferred activity of IFN-γ/PD-L1 signaling axis and determined genes and pathways differentially expressed and activated between “IFN-γ depleted” and “IFN-γ enriched” tumors.

We examined four scoring of IFNG axis, including axis of PD-L1 (CD274) expression level alone and a new scoring function, IPAS, introduced for assessment of activation of molecular pathways. Based on correlation assessments, we concluded that taken alone an expression level of CD274 is not likely be a correct marker for immune therapy, and this is especially true for metastatic tumors. We settled on combination of IPAS scores computed for positive and negative regulators of IFN-γ signaling as the scoring axis, which produced the deepest separation of tumors into biologically different disease subtypes.

We obtained a number of positive controls in proposing IFNG signaling axis as clinically relevant. In particular, we found significant correlation between IFNG axis and immune system related pathways of all major pathway databases used [65-69], we also found significant correlation between the axis score and well known immune evasion genes (IDO1, ENTPD1). The genomic signatures showed significant prevalence of immunosuppressive cells over cytotoxic lymphocytes in “IFN-γ-enriched” tumors.

The consistency of results obtained for primary set and independently for metastatic set of prostate tumors suggest that overall, the revealed characteristic genes and pathways represent true biological differences between “IFN-γ-depleted” and “IFN-γ-enriched” subtypes of prostate cancer tumors.

The further analysis suggests optimization of weights of positive and negative regulators in IFNG axis scoring function, adding mutation burden as a co-factor in ranking tumors, identifying potentially missed regulator genes to the genes of IFNG axis – all these are new questions set by this study.

The obtained results set a problem of practical classification of prostate cancer tumors into “IFN-γ-enriched” and immune “IFN-γ-depleted” subtypes. The fact that big cohorts of prostate cancer tumors can be separated into different biological subtypes does not mean that such separation can be easily made for any new tumor expression profile. We proposed a solution to this problem, a biomarker panel composed of two gene sets, which are inversely regulated in IFN-γ-enriched and in IFN-γ-depleted subtypes. The IPAS scoring function can naturally combined two scores and produce an accurate ranking of tumors.

Thus, this study provides a basis for a rational design of novel quantitative platform for classification of immune phenotypes of prostate cancer tumors and development of combination therapy strategies for patients. It worth noting also that proposed approach can be applied not only to classification of immune phenotypes, but also to classification of tumors with known or hypothetic molecular mechanism of drug action, e.g. classification metabolic phenotypes along inferred activity of a drug specific metabolic pathway and also along experimentally derived drug sensitivity axis.

References

1. Hegde PS, Chen DS. Top 10 Challenges in Cancer Immunotherapy. *Immunity.* 2020;52(1):17-35.
2. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell.* 2015;163(4):1011-1025.
3. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell.* 2015;27(4):450-461.
4. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell.* 2017;168(4):707-723.
5. Haffner MC, Guner G, Taheri D, et al. Comprehensive Evaluation of Programmed Death-Ligand 1 Expression in Primary and Metastatic Prostate Cancer. *Am J Pathol.* 2018;188(6):1478-1485.
6. Brahmer JR, Drake CG, Wollner I, et al. Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates. *J Clin Oncol.* 2010;28(19):3167-3175.
7. Kwon ED, Drake CG, Scher HI, et al. Ipilimumab versus placebo after radiotherapy in patients with metastatic castration-resistant prostate cancer that had progressed after docetaxel chemotherapy (CA184-043): a multicentre, randomised, double-blind, phase 3 trial. *Lancet Oncol.* 2014;15(7):700-712.
8. Beer TM, Kwon ED, Drake CG, et al. Randomized, Double-Blind, Phase III Trial of Ipilimumab Versus Placebo in Asymptomatic or Minimally Symptomatic Patients With Metastatic Chemotherapy-Naive Castration-Resistant Prostate Cancer. *J Clin Oncol.* 2017;35(1):40-47.
9. Hansen AR, Massard C, Ott PA, et al. Pembrolizumab for advanced prostate adenocarcinoma: findings of the KEYNOTE-028 study. *Ann Oncol.* 2018;29(8):1807-1813.
10. Boudadi K, Suzman DL, Anagnostou V, et al. Ipilimumab plus nivolumab and DNA-repair defects in AR-V7-expressing metastatic prostate cancer. *Oncotarget.* 2018;9(47):28561-28571.
11. Venturini NJ, Drake CG. Immunotherapy for Prostate Cancer. *Cold Spring Harb Perspect Med.* 2019;9(5).
12. Zappasodi R, Merghoub T, Wolchok JD. Emerging Concepts for Immune Checkpoint Blockade-Based Combination Therapies. *Cancer Cell.* 2018;34(4):690.
13. Abida W, Cyrta J, Heller G, et al. Genomic correlates of clinical outcome in advanced prostate cancer. *Proc Natl Acad Sci U S A.* 2019;116(23):11428-11436.
14. Abril-Rodriguez G, Ribas A. SnapShot: Immune Checkpoint Inhibitors. *Cancer Cell.* 2017;31(6):848-848 e841.
15. Kalbasi A, Ribas A. Tumour-intrinsic resistance to immune checkpoint blockade. *Nat Rev Immunol.* 2020;20(1):25-39.
16. Ikeda H, Old LJ, Schreiber RD. The roles of IFN gamma in protection against tumor development and cancer immunoediting. *Cytokine Growth Factor Rev.* 2002;13(2):95-109.
17. Ivashkiv LB. IFNgamma: signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy. *Nat Rev Immunol.* 2018;18(9):545-558.
18. Bai X, Fisher DE, Flaherty KT. Cell-state dynamics and therapeutic resistance in melanoma from the perspective of MITF and IFNgamma pathways. *Nat Rev Clin Oncol.* 2019;16(9):549-562.
19. Gao J, Shi LZ, Zhao H, et al. Loss of IFN-gamma Pathway Genes in Tumor Cells as a Mechanism of Resistance to Anti-CTLA-4 Therapy. *Cell.* 2016;167(2):397-404 e399.
20. Shin DS, Zaretsky JM, Escuin-Ordinas H, et al. Primary Resistance to PD-1 Blockade Mediated by JAK1/2 Mutations. *Cancer Discov.* 2017;7(2):188-201.
21. Zaretsky JM, Garcia-Diaz A, Shin DS, et al. Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N Engl J Med.* 2016;375(9):819-829.
22. Manguso RT, Pope HW, Zimmer MD, et al. In vivo CRISPR screening identifies Ptpn2 as a cancer immunotherapy target. *Nature.* 2017;547(7664):413-418.
23. Pan D, Kobayashi A, Jiang P, et al. A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science.* 2018;359(6377):770-775.

24. Patel SJ, Sanjana NE, Kishton RJ, et al. Identification of essential genes for cancer immunotherapy. *Nature.* 2017;548(7669):537-542.
25. Benci JL, Xu B, Qiu Y, et al. Tumor Interferon Signaling Regulates a Multigenic Resistance Program to Immune Checkpoint Blockade. *Cell.* 2016;167(6):1540-1554 e1512.
26. Dighe AS, Richards E, Old LJ, Schreiber RD. Enhanced in vivo growth and resistance to rejection of tumor cells expressing dominant negative IFN gamma receptors. *Immunity.* 1994;1(6):447-456.
27. Garcia-Diaz A, Shin DS, Moreno BH, et al. Interferon Receptor Signaling Pathways Regulating PD-L1 and PD-L2 Expression. *Cell Rep.* 2017;19(6):1189-1201.
28. Boehm U, Klamp T, Groot M, Howard JC. Cellular responses to interferon-gamma. *Annu Rev Immunol.* 1997;15:749-795.
29. Sfanos KS, Yegnasubramanian S, Nelson WG, De Marzo AM. The inflammatory microenvironment and microbiome in prostate cancer development. *Nat Rev Urol.* 2018;15(1):11-24.
30. Fares CM, Van Allen EM, Drake CG, Allison JP, Hu-Lieskovan S. Mechanisms of Resistance to Immune Checkpoint Blockade: Why Does Checkpoint Inhibitor Immunotherapy Not Work for All Patients? *Am Soc Clin Oncol Educ Book.* 2019;39:147-164.
31. Joyce JA, Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. *Science.* 2015;348(6230):74-80.
32. Oka T, Sugaya M, Takahashi N, et al. CXCL17 Attenuates Imiquimod-Induced Psoriasis-like Skin Inflammation by Recruiting Myeloid-Derived Suppressor Cells and Regulatory T Cells. *J Immunol.* 2017;198(10):3897-3908.
33. Hernandez-Ruiz M, Othy S, Herrera C, et al. Cxcl17(-/-) mice develop exacerbated disease in a T cell-dependent autoimmune model. *J Leukoc Biol.* 2019;105(5):1027-1039.
34. Brown RAM, Richardson KL, Kabir TD, Trinder D, Ganss R, Leedman PJ. Altered Iron Metabolism and Impact in Cancer Biology, Metastasis, and Immunology. *Front Oncol.* 2020;10:476.
35. Nairz M, Haschka D, Demetz E, Weiss G. Iron at the interface of immunity and infection. *Front Pharmacol.* 2014;5:152.
36. Pfeifhofer-Obermair C, Tymoszuk P, Petzer V, Weiss G, Nairz M. Iron in the Tumor Microenvironment-Connecting the Dots. *Front Oncol.* 2018;8:549.
37. Linder MC. Ceruloplasmin and other copper binding components of blood plasma and their functions: an update. *Metallomics.* 2016;8(9):887-905.
38. Wang W, Green M, Choi JE, et al. CD8(+) T cells regulate tumour ferroptosis during cancer immunotherapy. *Nature.* 2019;569(7755):270-274.
39. Kruzel ML, Zimecki M, Actor JK. Lactoferrin in a Context of Inflammation-Induced Pathology. *Front Immunol.* 2017;8:1438.
40. Graff JR, Konicek BW, Lynch RL, et al. eIF4E activation is commonly elevated in advanced human prostate cancers and significantly related to reduced patient survival. *Cancer Res.* 2009;69(9):3866-3873.
41. Hsieh AC, Nguyen HG, Wen L, et al. Cell type-specific abundance of 4EBP1 primes prostate cancer sensitivity or resistance to PI3K pathway inhibitors. *Sci Signal.* 2015;8(403):ra116.
42. Furic L, Rong L, Larsson O, et al. eIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression. *Proc Natl Acad Sci U S A.* 2010;107(32):14134-14139.
43. Liu Y, Horn JL, Banda K, et al. The androgen receptor regulates a druggable translational regulon in advanced prostate cancer. *Sci Transl Med.* 2019;11(503).
44. Robichaud N, Sonenberg N, Ruggero D, Schneider RJ. Translational Control in Cancer. *Cold Spring Harb Perspect Biol.* 2019;11(7).
45. Najjar YG, Menk AV, Sander C, et al. Tumor cell oxidative metabolism as a barrier to PD-1 blockade immunotherapy in melanoma. *JCI Insight.* 2019;4(5).

46. Wood LV, Fojo A, Roberson BD, et al. TARP vaccination is associated with slowing in PSA velocity and decreasing tumor growth rates in patients with Stage D0 prostate cancer. *Oncoimmunology.* 2016;5(8):e1197459.
47. Pang J, Liu WP, Liu XP, et al. Profiling protein markers associated with lymph node metastasis in prostate cancer by DIGE-based proteomics analysis. *J Proteome Res.* 2010;9(1):216-226.
48. Marques RB, Dits NF, Erkens-Schulze S, van Ijcken WF, van Weerden WM, Jenster G. Modulation of androgen receptor signaling in hormonal therapy-resistant prostate cancer cell lines. *PLoS One.* 2011;6(8):e23144.
49. He J, Lan X, Duan HL, Luo H, Zhou XD. CA916798 affects growth and metastasis of androgen-dependent prostate cancer cells. *Eur Rev Med Pharmacol Sci.* 2018;22(14):4477-4487.
50. Tykodi SS, Fujii N, Vigneron N, et al. C19orf48 encodes a minor histocompatibility antigen recognized by CD8+ cytotoxic T cells from renal cell carcinoma patients. *Clin Cancer Res.* 2008;14(16):5260-5269.
51. Navarro R, Compte M, Alvarez-Vallina L, Sanz L. Immune Regulation by Pericytes: Modulating Innate and Adaptive Immunity. *Front Immunol.* 2016;7:480.
52. Killingsworth MC, Wu X. Vascular pericyte density and angiogenesis associated with adenocarcinoma of the prostate. *Pathobiology.* 2011;78(1):24-34.
53. Welen K, Jennbacken K, Tesan T, Damber JE. Pericyte coverage decreases invasion of tumour cells into blood vessels in prostate cancer xenografts. *Prostate Cancer Prostatic Dis.* 2009;12(1):41-46.
54. Garner H, de Visser KE. Immune crosstalk in cancer progression and metastatic spread: a complex conversation. *Nat Rev Immunol.* 2020.
55. Coffelt SB, Wellenstein MD, de Visser KE. Neutrophils in cancer: neutral no more. *Nat Rev Cancer.* 2016;16(7):431-446.
56. Wculek SK, Malanchi I. Neutrophils support lung colonization of metastasis-initiating breast cancer cells. *Nature.* 2015;528(7582):413-417.
57. Hugo W, Zaretsky JM, Sun L, et al. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell.* 2016;165(1):35-44.
58. Francis JC, Capper A, Ning J, Knight E, de Bono J, Swain A. SOX9 is a driver of aggressive prostate cancer by promoting invasion, cell fate and cytoskeleton alterations and epithelial to mesenchymal transition. *Oncotarget.* 2018;9(7):7604-7615.
59. Laughney AM, Hu J, Campbell NR, et al. Regenerative lineages and immune-mediated pruning in lung cancer metastasis. *Nat Med.* 2020;26(2):259-269.
60. Zahalka AH, Frenette PS. Nerves in cancer. *Nat Rev Cancer.* 2020;20(3):143-157.
61. Kraus RD, Barsky A, Ji L, et al. The Perineural Invasion Paradox: Is Perineural Invasion an Independent Prognostic Indicator of Biochemical Recurrence Risk in Patients With pT2N0R0 Prostate Cancer? A Multi-Institutional Study. *Adv Radiat Oncol.* 2019;4(1):96-102.
62. Barsky AR, Kraus RD, Carmona R, et al. Investigating association of perineural invasion on prostate biopsy with Gleason score upgrading at prostatectomy: A multi-institutional analysis. *Cancer Med.* 2020;9(10):3383-3389.
63. Rykunov D, Beckmann ND, Li H, Uzilov A, Schadt EE, Reva B. A new molecular signature method for prediction of driver cancer pathways from transcriptional data. *Nucleic Acids Res.* 2016;44(11):e110.
64. Barbie DA, Tamayo P, Boehm JS, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature.* 2009;462(7269):108-112.
65. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol.* 2017;18(1):220.
66. Jassal B, Matthews L, Viteri G, et al. The reactome pathway knowledgebase. *Nucleic Acids Res.* 2020;48(D1):D498-D503.

67. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27-30.
68. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 2015;1(6):417-425.
69. Slenter DN, Kutmon M, Hanspers K, et al. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.* 2018;46(D1):D661-D667.
70. Endo TA, Masuhara M, Yokouchi M, et al. A new protein containing an SH2 domain that inhibits JAK kinases. *Nature.* 1997;387(6636):921-924.
71. Hodges C, Kirkland JG, Crabtree GR. The Many Roles of BAF (mSWI/SNF) and PBAF Complexes in Cancer. *Cold Spring Harb Perspect Med.* 2016;6(8).
72. Rabellino A, Andreani C, Scaglioni PP. The Role of PIAS SUMO E3-Ligases in Cancer. *Cancer Res.* 2017;77(7):1542-1547.
73. Huang L, Malu S, McKenzie JA, et al. The RNA-binding Protein MEX3B Mediates Resistance to Cancer Immunotherapy by Downregulating HLA-A Expression. *Clin Cancer Res.* 2018;24(14):3366-3376.

ACKNOWLEDGMENTS:

Author contributions: BR, ES and AT designed the project; BR and ES developed the IPAS algorithm; BR and AC developed the software and performed data analysis; TO, BR, SN and AT performed biological and clinical interpretations of results; all wrote the paper.

Competing interests: The authors report no competing interests.

Data and materials availability:
Supplementary:
Supplement 1. IPAS - a new method for comparison of tumors by Inferred Pathway Activation and Suppression Score.
Supplement 2. Algorithm for separation of tumors into 2 distinct classes along given scoring axis.
Supplement 3. Classification of different subtypes of primary and metastatic prostate tumors made with MSM algorithm.
Supplement 4. Statistical characteristics of genes and pathways differentially expressed between IFN-γ enriched and IFN-γ depleted subtypes (excel table).

Computational research reported in this paper was supported by the Office of Research Infrastructure of the National Institutes of Health under award numbers S10OD018522 and S10OD026880. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

TABLE AND FIGURES LEGENDS

Table 1. Statistics of Pearson correlation coefficients computed between gene expression levels of prostate cancer tumors and scoring functions representing activation of IFN-γ/PD-L1 axis[$].

[$]Numbers of genes with the absolute value of Pearson correlation coefficient computed between gene's expression levels of TCGA and MSKSS sets of prostate tumors and values of the pathway scoring functions above or equal to given correlation thresholds (0.1, 0.2,...0.8); [$$]Scoring functions used in computing correlations: CD274 stands for expression levels of the gene CD274; IPAS($IFNG_P$) and ssGSEA($IFNG_P$) stand, respectively, for IPAS and ssGSEA scoring functions computed for genes, which positively regulate expression of CD274 in IFN-γ pathway of Fig.1; <$IFNG_P$> represent average expression levels of positive regulators of IFN-γ pathway of Fig.1; the combined activity of IFN-γ pathway, which include both positive and negative regulators, was represented by the differences of IPAS and ssGSEA scoring functions computed, respectively, for sets of positive ($IFNG_P$) and negative ($IFNG_N$) regulator genes and also by the difference of average gene expression levels of positive, <$IFNG_P$> , and negative, <$IFNG_N$>, regulators; ; the maximum numbers of correlated genes obtained for each of thresholds are shown in bold.

Figure 1. Schematic of IFN-γ/PD-L1 signaling axis. A reactive T cell recognizes a cancer cell antigens via MHC class molecules and releases IFN-γ (IFNG) that binds by the IFN-γ receptors (IFNGR1, INFGR2) on the cancer cell resulting in activation of the JAK-STAT pathway (JAK1, JAK2, STAT1, STAT3). This activates the transcription factor IRF1, which then activates the transcription of CD274 (PD-L1 is a product of CD274). Expression of PD-L1/2 on the surface of the cancer cell negatively regulates the anti-cancer T cell response [14,15]. IFN-γ signaling is positively regulated by the above genes and by receptor apelin (APLNR) and negatively regulated by non-receptor tyrosine-protein phosphatases PTPN2, PTPN6, PTPN11 [22-24] and by suppressor of cytokine signaling (SOCS) proteins [70]. Transcription of IFN-γ target genes is negatively regulated by the chromatin remodeling complex PBAF (BRD7, ARID2, PBRM1) [71] and SUMO, PIAS1 [72]. IFN-γ dependent transcriptional regulation of MHC class I (HLA-A) is inhibited by MEX3B [73].

Figure 2. Determining position of the maximal difference between IFN-γ-depleted and IFN-γ-enriched subtypes for primary and metastati**c** tumor sets. We identify two IFN-γ-depleted subtypes of 82 and 189 tumors of total 497 primary tumors **(A)** and a subtype of 174 tumors of 270 of total metastatic tumors **(B)**. The position of the maximal difference between tumors ranked by increase of inferred activity of IFN-γ/PD-L1 signaling is determined by the minimum of the combined scoring function (Supplement 2). We compare two axes: the positive IPAS ($IFNG_P$) and combined IPAS ($IFNG_P$) – IPAS ($IFNG_N$). The combined axis makes deeper separation of tumors (lower value of the combined scoring function). The genes of IFN-γ/PD-L1 pathway are omitted in computations to minimize the bias from the axis genes. The right vertical axis corresponds to green plots when all genes are taken into account (correlation threshold equal to zero); the depth of the minimum increases for sets of higher correlated genes **(A, B)**. The vertical blue lines show positions of the determined minima **(A, B)**.

Figure 3. (A) Heatmap of major genes with known or proposed role in immune evasion determined as differentially expressed between IFN-γ-enriched and IFN-γ-depleted subtypes of primary prostate cancers. (B) Heatmap of major pathways determined as differentially activated between IFN-γ-enriched and IFN-γ-depleted subtypes of primary prostate cancers.

Figure 4. (A) Heatmap of major genes with known or proposed role in immune evasion determined as differentially expressed between IFN-γ-enriched and IFN-γ-depleted subtypes of

metastatic prostate cancers. (B) Heatmap of major pathways determined as differentially activated between IFN-γ-enriched and IFN-γ-depleted subtypes of metastatic prostate cancers.

Figure 5. Biomarker genes proposed for classification of primary and metastatic prostate tumors. Heatmaps of expression levels of proposed biomarker genes for classification of primary and metastatic prostate tumors **(A, C)** and related classification summary plots **(B, D)**. The classification scores were computed using IPAS method (Supplement 2) applied to "positive" and "negative" biomarker genes. "Positive" and "negative" biomarkers were defined based on, respectfully, significant upregulation or downregulation of expression levels in IFN-γ enriched immune subtype as compared to "IFN-γ-depleted" immune subtype. Thus, given only relative positions of "positive" and "negative" genes in a set of 14 (primary) or 11 (metastatic) biomarker genes, one can classify tumors based on a value of the IPAS score. Panels **(C)** and **(D)** present percentages of tumors which were classified within a given misclassification threshold; in particular, ~50% of primary tumors and ~70% of metastatic tumors were classified in immune IFN-γ-enriched or IFN-γ-depleted subtypes with percentage of misclassifications less than 5%. As accuracy of classification depends on a value the classification score, percentages of misclassifications were computed separately for each of the immune subtypes as function of the classification score. Selecting genes for the panel, we used the following criteria: genes upregulated in IFN-γ-enriched or IFN-γ-depleted subtypes were sorted by P values (<0.0001) and then by fold changes (>2), and, finally, we excluded genes with very low or very high average expression level.

Figure 6. Computational protocol for classification of tumors into clinically distinct subtypes along a given biological axis. The protocol combines three algorithms (red diamonds): "Two class separation" (Supplement 1); the pathway scoring method "Pathway Assessor" (ssGSEA [64] and IPAS, Supplement 2); and the classification algorithm (Molecular Signature Method, MSM [63]). Given as input gene sets (blue rectangles) from "Pathway Databases", "Biological Axis", "Expression Table" (ET), the "Pathway Assessor" computes scores (grey rectangles) "Pathway scores" and "Biological Axis scores" used for ranking tumors along this axis. "Biological Axis scores" can be input also as independent variable. The ET genes are been filtered by a given correlation (both positive and negative) with the BA scores, and, after this, given filtered the ET and the BA scores used by the "Two class separation" algorithm that finds the most distinct classes. Finally, the MSM [63] takes as input either the ET genes or Pathway scores and determines characteristic genes and pathways and their classification signatures. The obtained characteristic genes and pathways are been matched with genes available from genomic analysis of clinical data. In application to immune therapy, we use known genes and pathways involved in immune evasion.

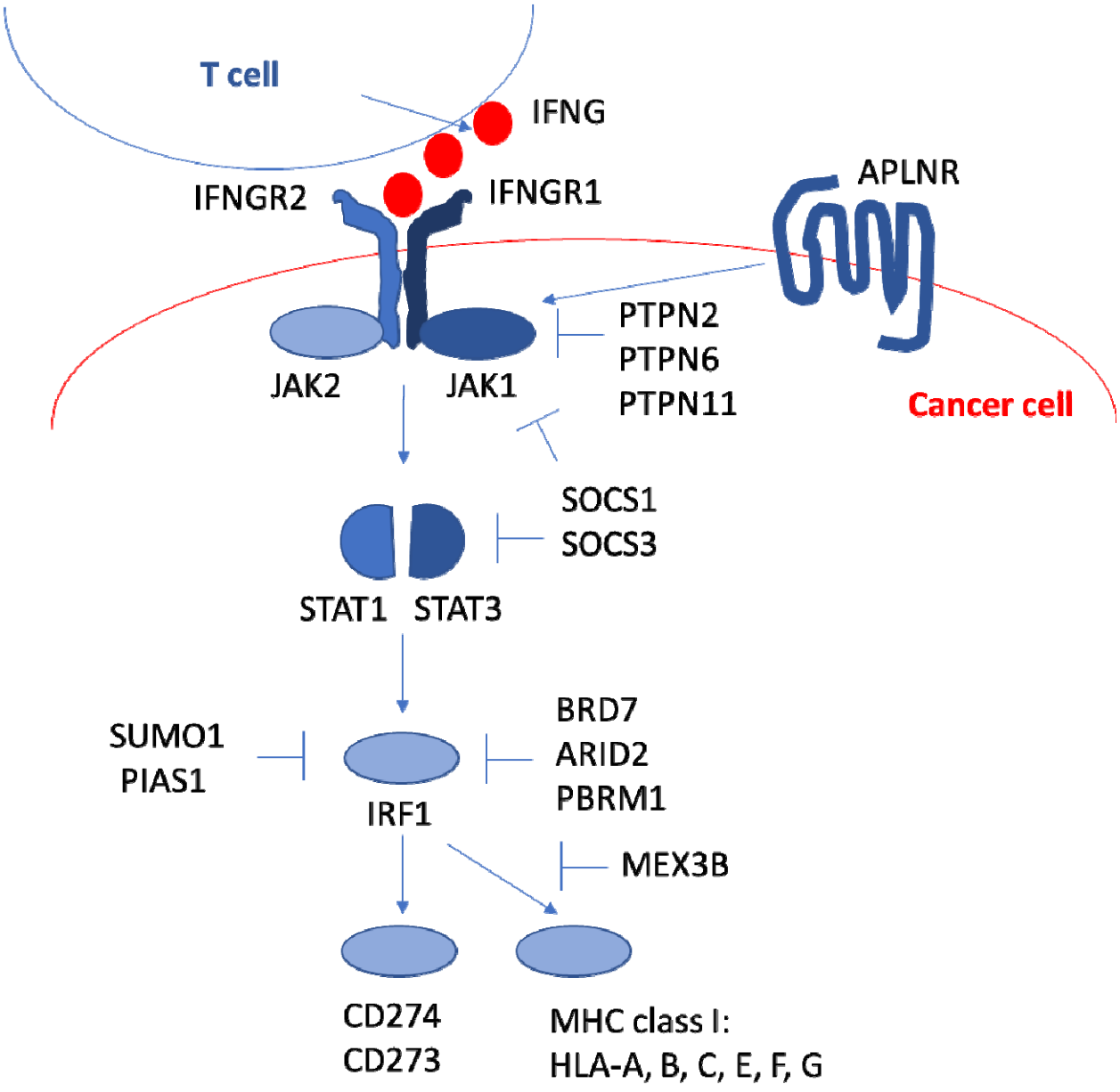

Figure 1. IFN-γ/PD-L1 signaling axis.

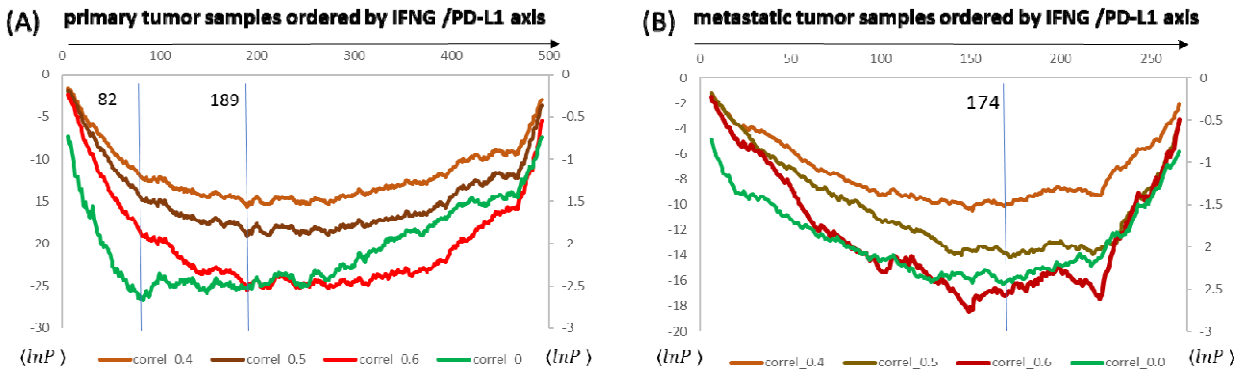

Figure 2. Determining position of the maximal difference between IFN-γ-depleted and IFN-γ-enriched subtypes for primary **(A)** and metastatic **(B)** tumors.

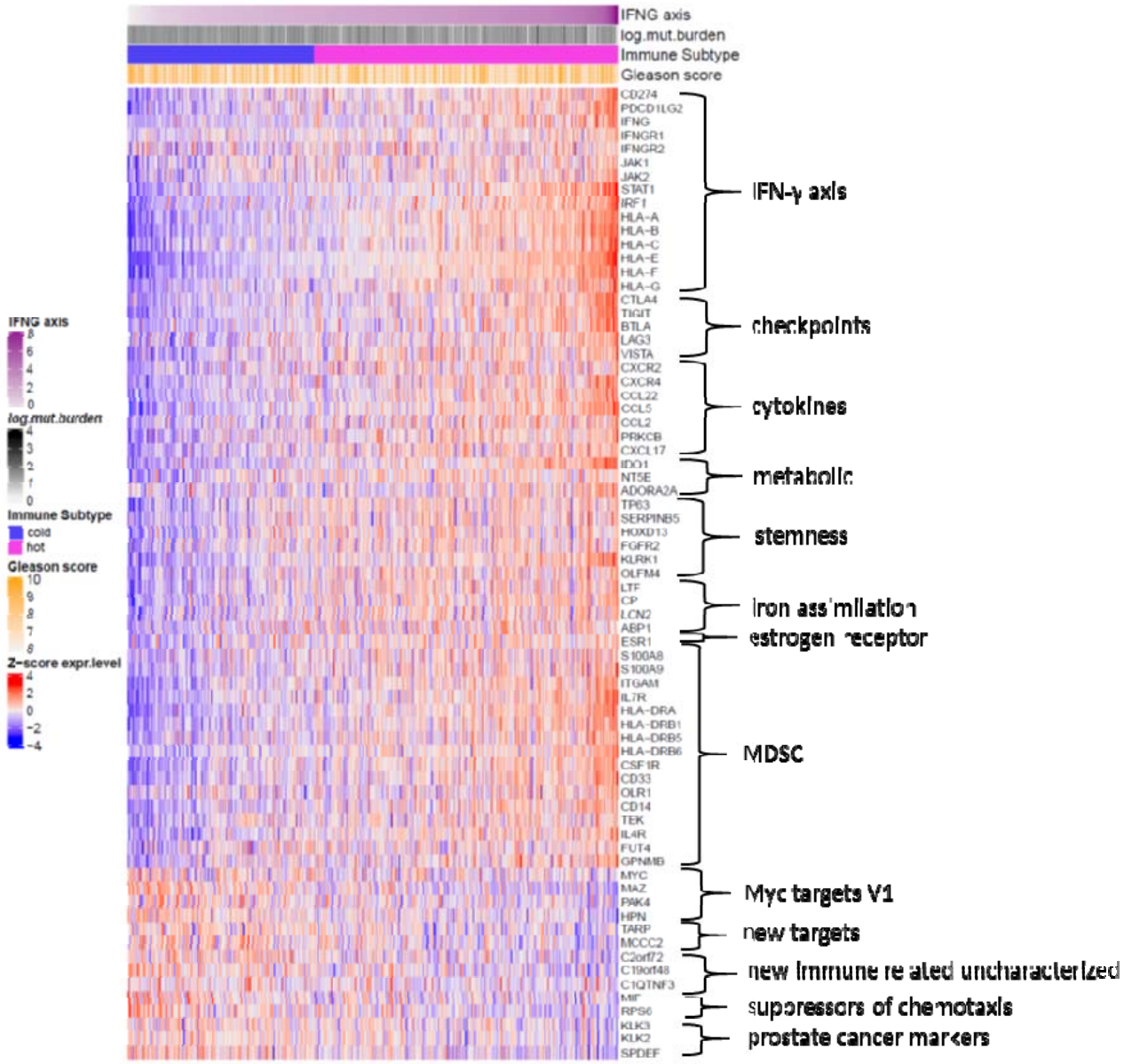

Figure 3A. Heatmap of major genes with known or proposed role in immune evasion determined as differentially expressed between IFN-γ-enriched and IFN-γ-depleted subtypes of primary prostate cancers.

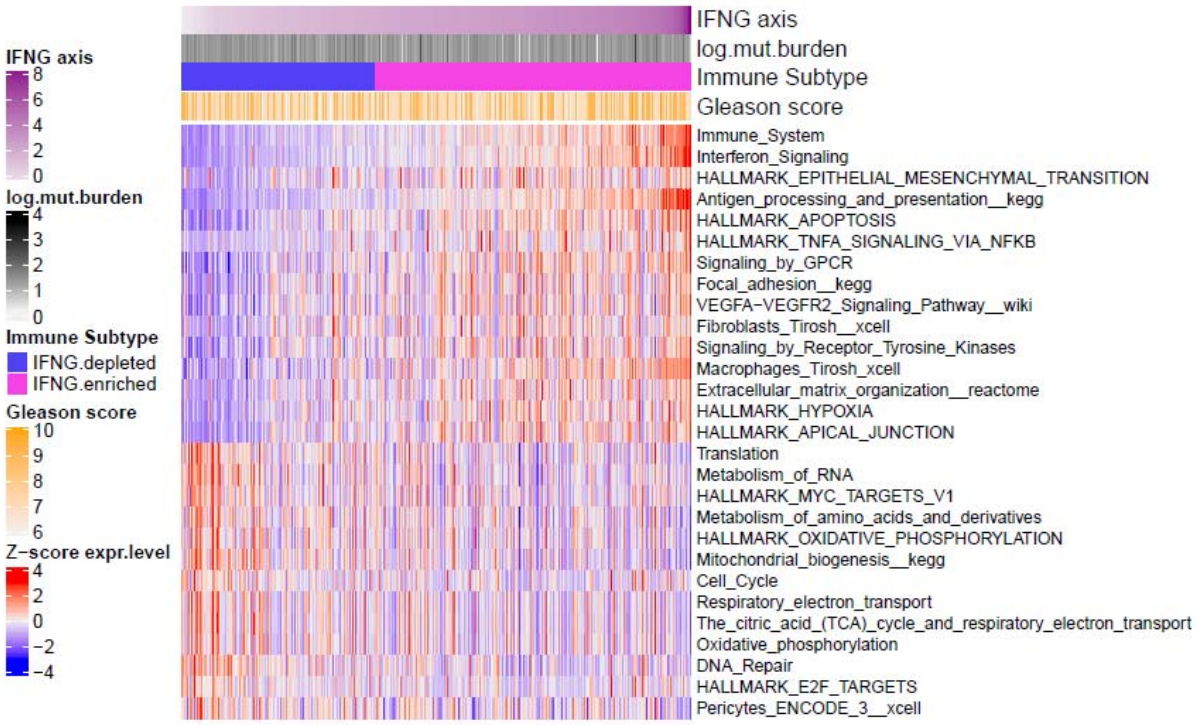

Fig.3B. Heatmap of major pathways determined as differentially activated between IFN-γ-enriched and IFN-γ-depleted subtypes of primary prostate cancers.

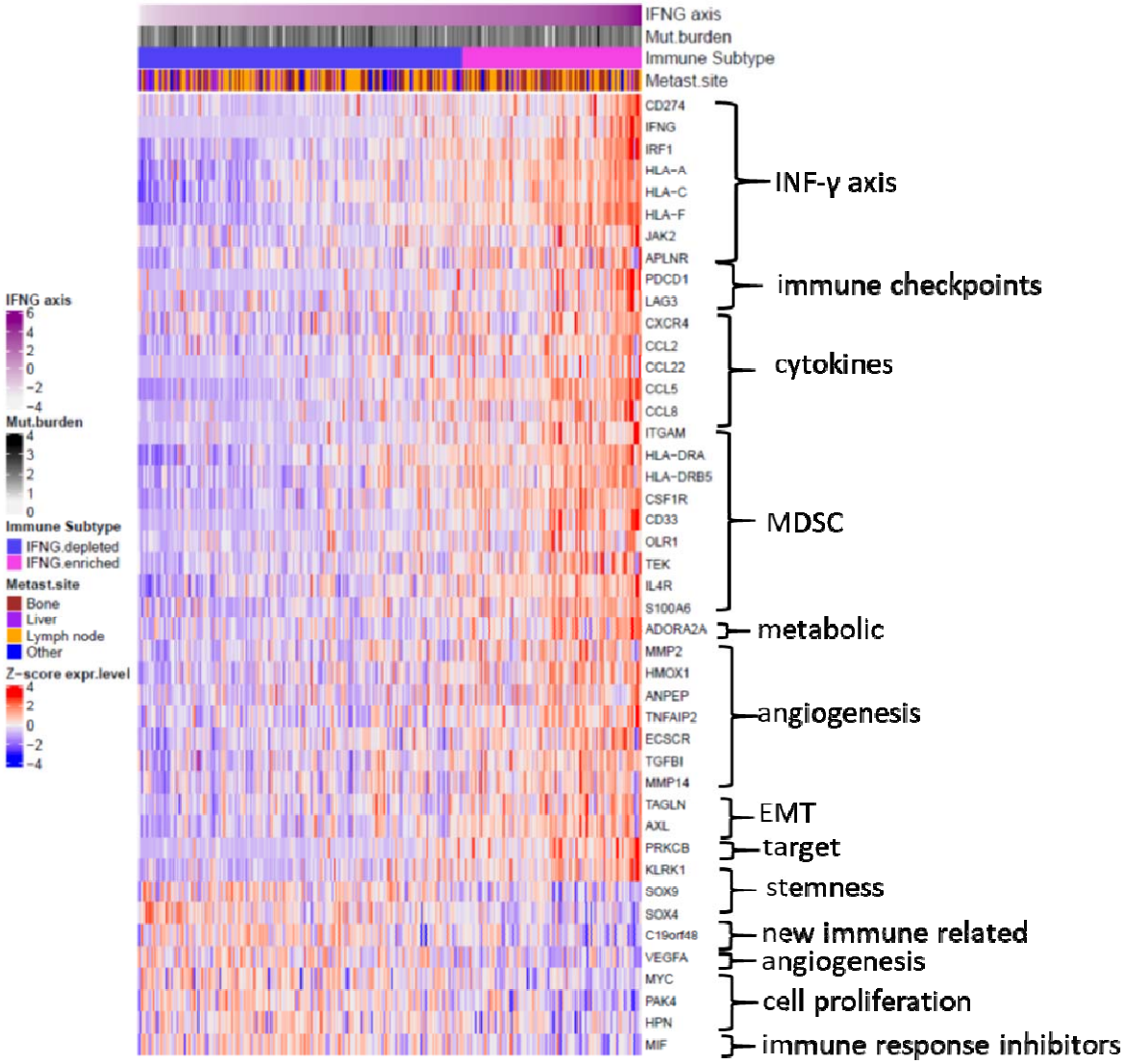

Figure 4A. Heatmap of major genes with known or proposed role in immune evasion determined as differentially expressed between IFN-γ-enriched and IFN-γ-depleted subtypes of metastatic prostate cancers.

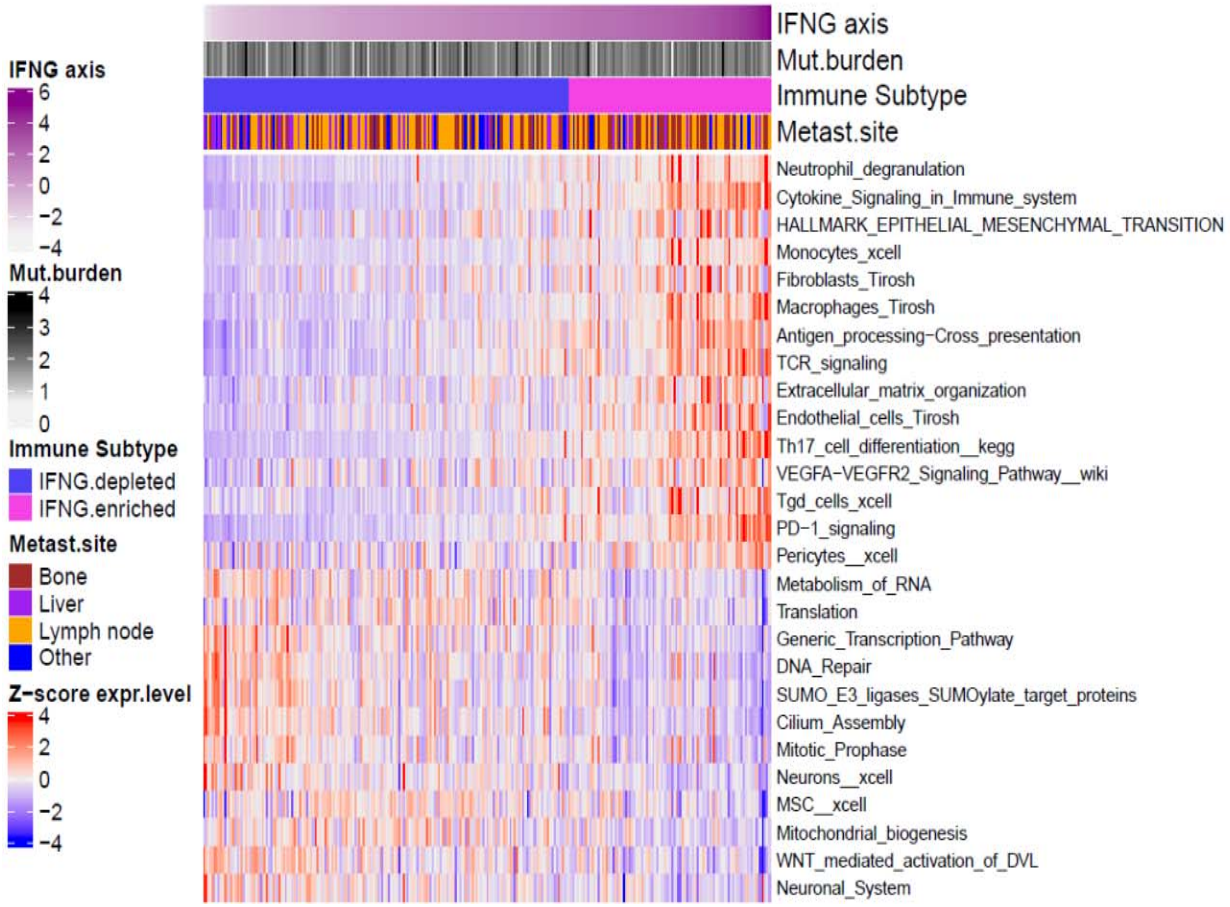

Figure 4B. Heatmap of major pathways determined as differentially activated between IFN-γ-enriched and IFN-γ-depleted subtypes of metastatic prostate cancers.

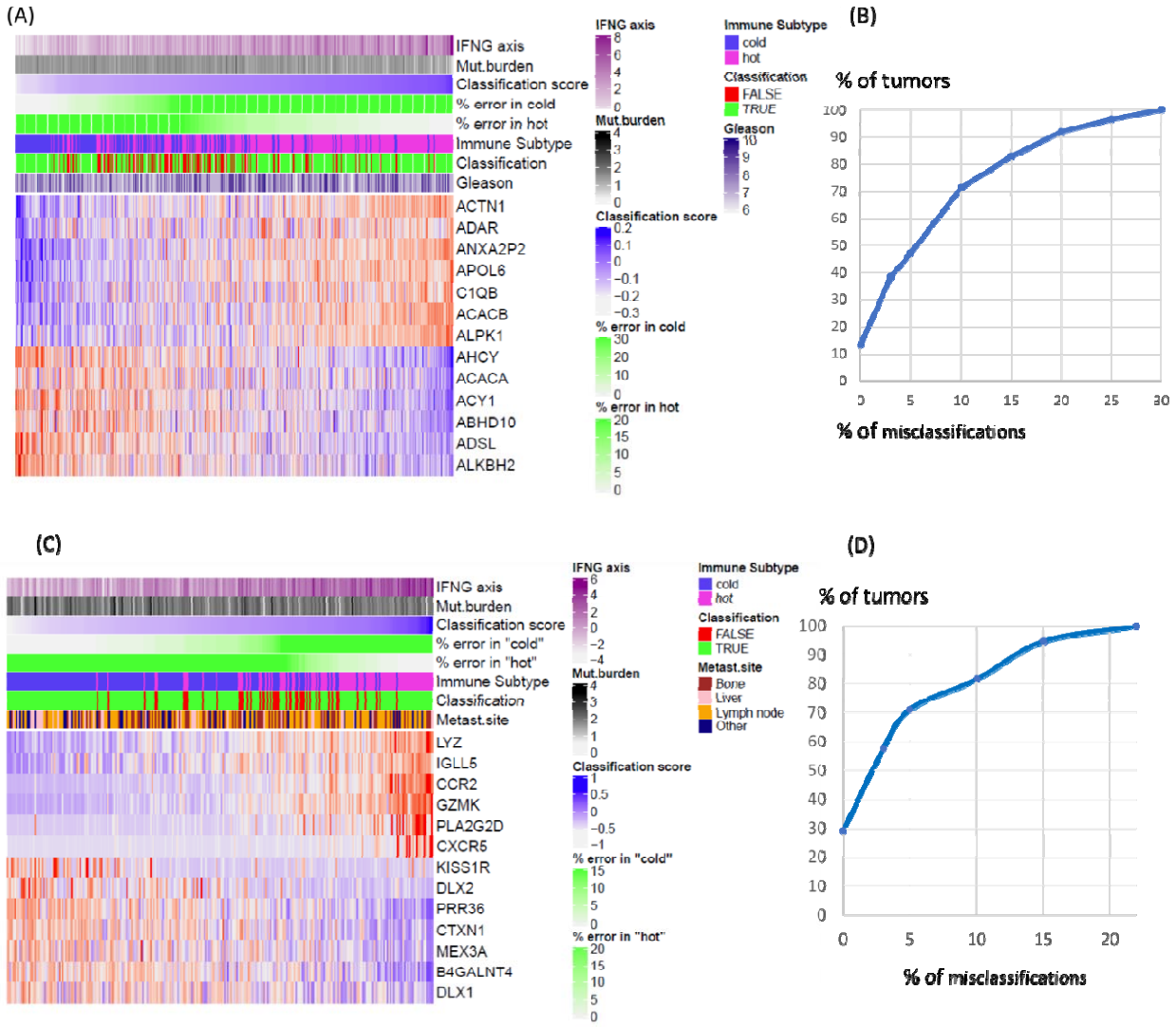

Fig.5. Heatmaps of expression levels of biomarker genes proposed for classification of primary and metastatic prostate tumors (panels A and C) and related classification summary plots (panels C and D).

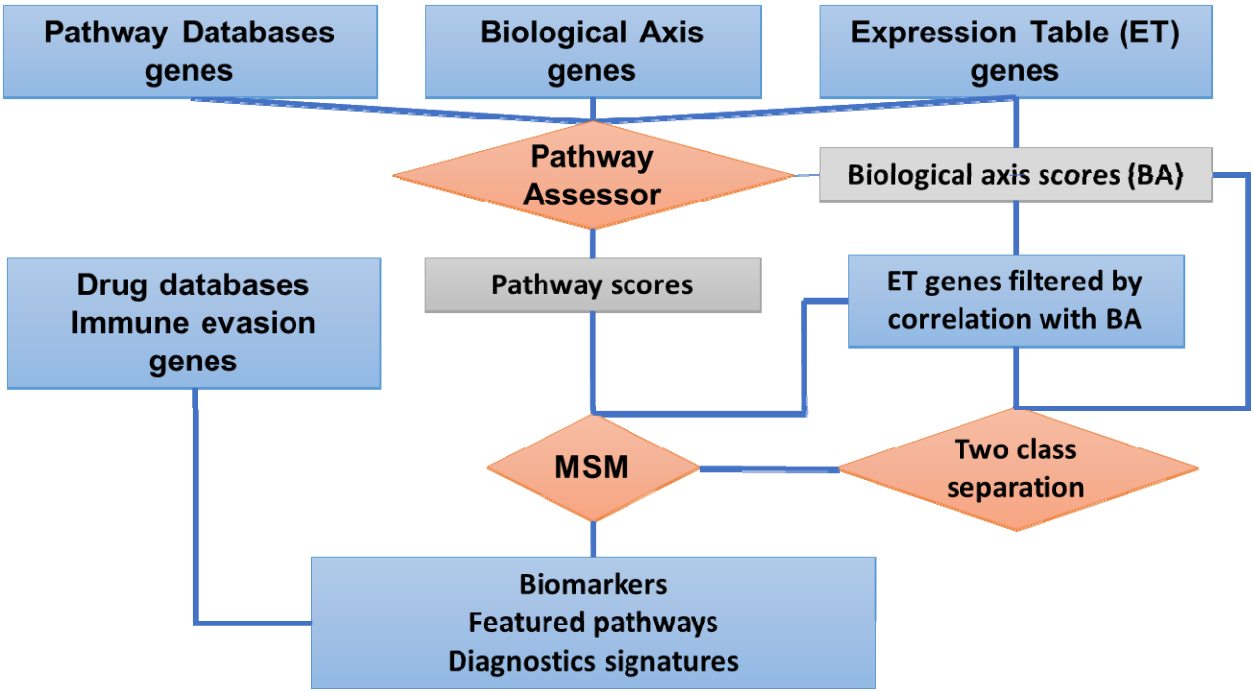

Fig.6. Computational protocol for classification of tumors into clinically distinct subtypes along a given biological axis.

Tables

Table 1. Statistics of Pearson correlation coefficients computed between gene expression levels of prostate cancer tumors and scoring functions representing activation of IFN-γ/PD-L1 axis[$].

	primary tumors (TCGA)								metastatic tumors (MSKCC)							
Correlation threshold / Scoring functions[$$]	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.1	0.2	0.3	0	0.5	0.6	0.7	0.8
CD274	10337	5357	1819	732	401	170	**36**	0	475	80	18	9	5	4	2	0
IPAS($IFNG_P$)	11461	**6409**	**2912**	1042	**511**	**189**	17	0	8243	4748	2156	731	**292**	**107**	**15**	0
ssGSEA($IFNG_P$)	**11543**	6406	2876	**1070**	466	119	8	0	**8446**	**4822**	**2256**	**748**	226	61	5	0
<$IFNG_P$>	8228	2579	867	398	115	24	11	**6**	7116	2962	938	330	116	20	3	0
IPAS($IFNG_P$) - IPAS($IFNG_N$)	9620	4017	1502	**751**	**383**	**82**	**14**	**1**	**9116**	**4383**	**1755**	**520**	**158**	**24**	0	0
ssGSEA($IFNG_P$) - ssGSEA($IFNG_N$)	**10735**	**5302**	**1614**	193	19	2	0	0	8221	3798	1497	363	45	0	0	0
<$IFNG_P$> - <$IFNG_N$>	8516	2722	818	345	92	21	10	6	7059	2895	870	299	102	17	3	0

[$]Numbers of genes with the absolute value of Pearson correlation coefficient computed between gene's expression levels of TCGA and MSKSS sets of prostate tumors and values of the pathway scoring functions above or equal to given correlation thresholds (0.1, 0.2,...0.8); [$$]Scoring functions used in computing correlations: CD274 stands for expression levels of the gene CD274; IPAS($IFNG_P$) and ssGSEA($IFNG_P$) stand, respectively, for IPAS and ssGSEA scoring functions computed for genes, which positively regulate expression of CD274 in IFN-γ pathway of Fig.1; <$IFNG_P$> represent average expression levels of positive regulators of IFN-γ pathway of Fig.1; the combined activity of IFN-γ pathway, which include both positive and negative regulators, was represented by the differences of IPAS and ssGSEA scoring functions computed, respectively, for sets of positive ($IFNG_P$) and negative ($IFNG_N$) regulator genes and also by the difference of average gene expression levels of positive, <$IFNG_P$> , and negative, <$IFNG_N$>, regulators; ; the maximum numbers of correlated genes obtained for each of thresholds are shown in bold.

Supplement 1. IPAS - a new method for comparison of tumors by Inferred Pathway Activation / Suppression Score.

Understanding the activation status of molecular pathways in tumors can help nominate disease driver pathways, identify targets for therapeutic intervention, and determine clinically relevant subtypes and diagnostic biomarkers. We introduce a new method to infer pathway activation and suppression by examining under- and over-representation of pathway genes in tumor genes ranked by gene expression levels.

Like in related ssGSEA method [1], we assume that biological activity of a pathway is regulated by collective changes of expression levels of majority of genes involved in this pathway; then a difference in a pathway activity between tumors can be assessed by a difference in positioning of expression levels of genes involved in this pathway in ranked list of expression levels of all genes in each of tumors (Fig.1).

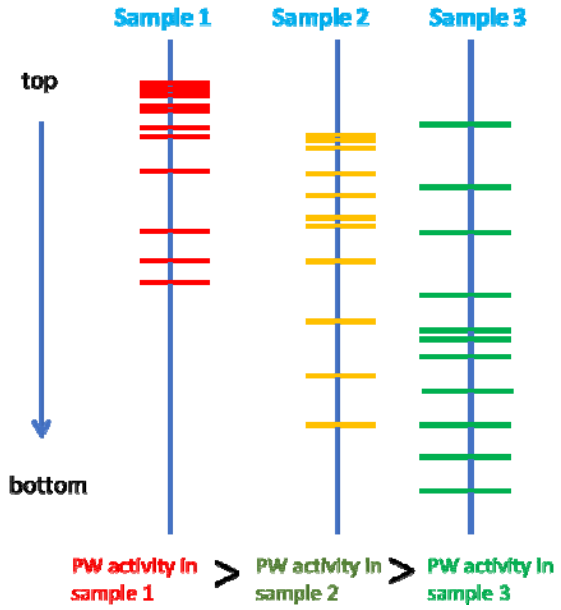

Fig.1 Tumors differ by distributions of gene expression levels. Expression levels of an illustrative pathway in different samples are shown by red, yellow and green bars. We hypothesize that activity of a given molecular PW in different tumors can be compared by levels of overrepresentation of pathways genes among top upregulated genes in a tumor. The roles and expression levels of individual genes are not taken into account in this approach.

The novelty of our approach rests in defining the pathway scoring function as an entropy difference associated with deviation of the gene expression levels from equilibrium or random distribution because of either activation or suppression. Following this idea, the entropy difference or scoring function can be presented as a log of ratio of two probabilities: a probability to occupy by random the observed positions in a list of tumor genes ranked by expression levels from the top to the bottom and, similarly, a probability to occupy by random the observed positions in a list of expression levels ranked from the bottom to the top. Thus, for a pathway of a single gene, its relative activity across tumors is defined as a negative log of ratio of two numbers: a number of genes with expression level bigger than an expression level of given gene, and a number of genes with expression levels less than an expression level of given gene.

For pathways of multiple genes, the “top” and “bottom” probabilities can be computed accurately by using statistical mechanics of one-dimension systems [2]. The complexity of this algorithms is assessed as ~$M*N^2$, where M is a number of genes in a pathway and N is a number of genes in expression table. However, direct accurate computations of partition function become time consuming and impractical for M~200, N ~20K and ~5K pathways of major pathway databases.

Therefore, we assessed probabilities as geometrically averaged P values computed for each of genes using Fisher’s exact test [3], given gene’s ranks in a list of pathway genes and in a list of ranked genes of a tumor, a number of genes in a pathway, and the total number of genes with the assessed expression level in a given tumor (Fig.2).

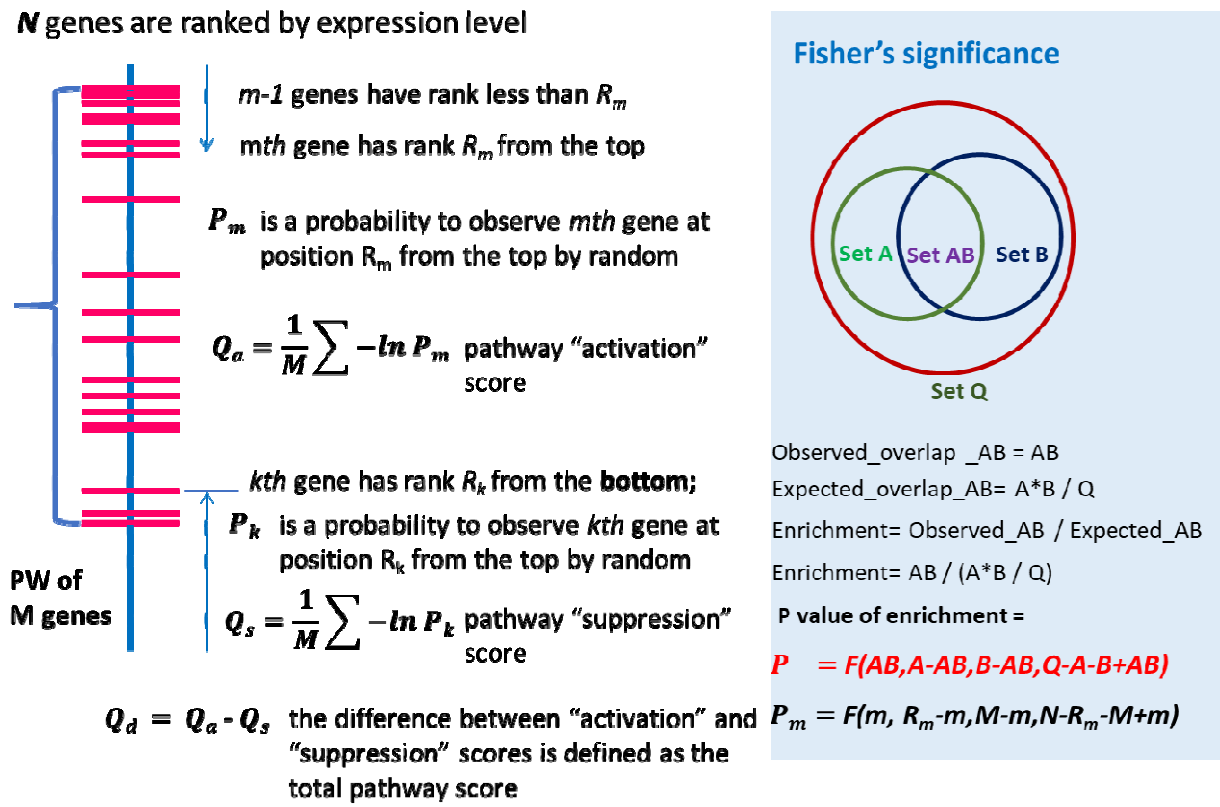

Fig.2. Algorithm for computing IPAS scoring function.

The scoring function Q_d is positive, when expression levels of pathway's genes are overrepresented among top expressed genes of a tumor, and it is negative, when pathway's genes are at the bottom of expressed genes of a tumor; the scoring function is close to zero, when expression levels are distributed by random or approximately equally split between genes at the top and genes at the bottom. Thus, considering independently activation, Q_a, and suppression, Q_s, scores, this approach makes possible to detect dysregulated pathways, which have gene expression levels overrepresented both at the top and at the bottom of a ranked gene list. Given a pathway with both positively and negatively regulated branches (e.g. cancer driver module which becomes active because both of upregulated expression of an oncogene and downregulation expression of tumor suppressor) one can assess its overall activity by combining inferred activities (or suppressions) of positive and negative pathway branches.

Software and data access

To facilitate IPAS analysis, we have developed a web application accessible at: http://pathwayassessor.org. This tool allows researchers to perform IPAS analysis on a given expression table, without the need to install applications or write code. The user uploads an expression table, where row names are gene symbols and column names are sample identifiers. The user then selects several parameters, including: Pathway Database, Mode, and Direction. The final output is a table, or collection of tables, where the row names are pathway names, column names are sample identifiers, and each record is the final IPAS score.

The Pathway Database parameter determines the collection of pathways or gene sets for which to calculate IPAS scores. Pathway databases included are KEGG (Kyoto Encyclopedia of Genes and Genomes), Reactome, Hallmark gene sets from the Broad Institute, and HMDB/SMPDB (Human Metabolome Database and Small Molecule Pathway Database).

The Mode parameter determines the method of aggregating p-value scores for gene rankings to produce the final score. A user selects harmonic averaging, geometric averaging, or selection of the minimum p-value.

Direction options include “ascending”, “descending”, and “difference”. The options for “ascending” and “descending” indicate the sorting direction of a sample’s expression profile for determining gene rankings, for assessing the underactivity and overactivity respectively of genes in a pathway. The “difference” option produces a third table of scores with the difference between the overactivity and underactivity of genes in a given pathway.

When submitting the expression table and selected parameters, the user also enters their email address to receive analysis results. The analysis is then queued for execution on a remote server, which emails the results to the user as TSV format file attachments.

The source code for this web application is completely open-source and can be accessed on Github at: https://github.com/RevaLab/PathwayAssessor. For researchers who would like to include this analysis as part of a computational pipeline, the code is also available as a Python package, available through the Python Package Index (pypi) at: https://pypi.org/project/pathway-assessor/.

References

1. Barbie, D.A. *et al.* Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* **462**, 108-112 (2009)
2. Finkelstein and Reva. Search for the stable state of a short chain in a molecular field. Protein Engineering vol.5 no 7 pp.617-624, 1992.
3. McDonald, J.H. 2014. Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, Baltimore, Maryland. pp 77-85.

Supplement 2. Algorithm for separation of tumors into two distinct classes along given scoring axis.

Tumors ranked by a given scoring function. For a set of N tumors, there are N-1 possible separations into two classes (Fig.1). Which of those tumor classes are the most distinct?

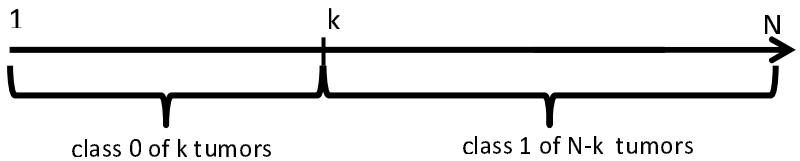

Fig.1. Separation of tumors into 2 classes along a given scoring function; there are N-1 separations into 2 classes.

Let *M* be a number of genes in a dataset; for a separation k, a difference between distribution of gene expression values in classes 0 and 1 can be assessed by a P value, P_{km}. Then the statistical difference between gene expression distributions in two given classes of k and N-k tumors is assessed by a harmonically averaged P-value of individual genes:

It is assumed that the minimal value of determines the most distinct separation of tumors along a given axis. For control, we also performed random selection of tumors in classes 0 and 1, computed, $\{P_{mrk}\}$ and assess a random separation of tumors by harmonically averaged P-values $\{P_{rk}\}$,where r =1, 2,… 10^4 .
For computing P_{mrk} , we implemented and tested several methods. The most robust method, which never gives P=0 for practical gene expression distribution is describe below in Fig.2.

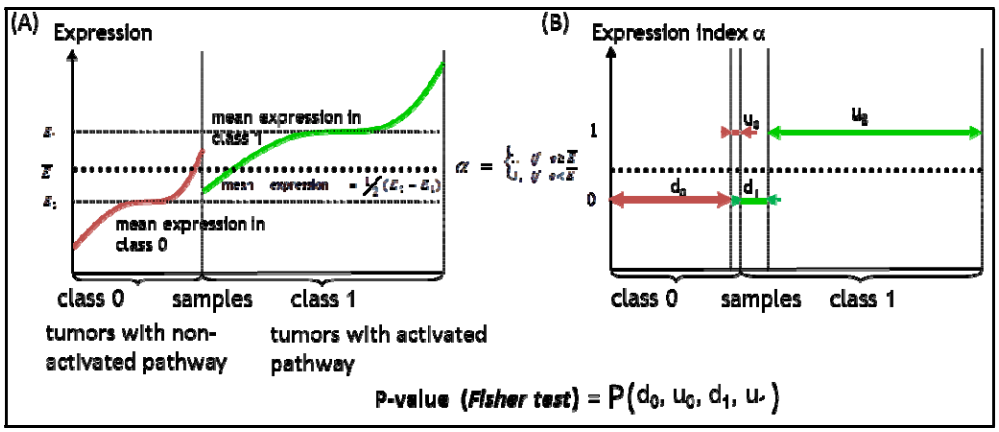

Fig.2. Computing a P-value between two distributions of gene expressions using the discrete approximation. (A) For each gene, the mean expression values, E_0 and E_1, are computed for sample classes 0 and 1, respectively. The discrete approximation threshold, $\overline{E}$, is defined as the average value of E_0 and E_1. (B) A discrete index "u" or "1" is assigned to an expression value e , if $e \geq \overline{E}$; an index "d" or "0" is assigned to e , if $e < \overline{E}$. Thus, depending on a gene's expression values, samples of class 0 are separated into two groups, d_0 and u_0, and samples of class 1 are separated into groups d_1 and u_1, respectively. For clarity, expression values are sorted within each of the sample classes. The statistical significance of association between expression values and tumor classes is assessed by Fisher test [1].

1. McDonald, J.H. 2014. Handbook of Biological Statistics (3rd ed.). Sparky House Publishing, Baltimore, Maryland. pp 77-85.

Supplement 3. Classification of different subtypes of primary and metastatic prostate tumors made with MSM algorithm.

Given a table with gene expression levels and two tumor classes, the algorithm [1] uses halve of tumor samples as a training set to determine (i) biomarker genes, (ii) a weighted combination of biomarker expression levels (classification score) and (iii) optimal threshold for separation tumors of training set into given classes. Then, the determined biomarker genes, their weights and threshold were used in classification of remaining halve of tumor samples. After 1024 rounds of training/test classifications, for each of samples its probability of been correctly classified in training and test set was determined as well as averaged value of classification function. The classification accuracies were assessed by area under receiver-operating curve (AUC). Two AUC values reported in figures for both training and test classifications [1] were derived from, respectively, probabilities of each of samples to be classified in correct class; and from the average classification scores.

The figures below show classification accuracies of different disease subtypes computed for gene expression values of individual genes and for pathways scores; probabilities and averaged values of scoring functions are shown only for classification of test sets; red and blue circles represent tumors of different classes.

The presented classification plots show that revealed immune subtypes (Fig.S3.1, FigS3.3) are distinct as much as genomic subtypes in primary tumors (Fig.S3.4) or clinical subtypes in metastatic tumors (Figs.S3.5-8). They also show that the immune subtypes (Fig.S3.1, FigS3.3) can be classified with using relative expression values of a small number of biomarker genes (Fig.S3.2 and Fig.S3.4) practically as accurate as with using gene expression values or pathways scores assessed from ~20K expression values for each of the samples. Finally, we demonstrated the competiveness of classification based on pathway scoring as compared to individual gene expression levels and competiveness of a new IPAS scoring method as compared to ssGSEA. The pathway scoring is based on relative position of genes expression levels and therefore can be potentially platform independent.

1. Rykunov D, Beckmann ND, Li H, Uzilov A, Schadt EE, Reva B. A new molecular signature method for prediction of driver cancer pathways from transcriptional data. *Nucleic Acids Res.* 2016;44(11):e110.

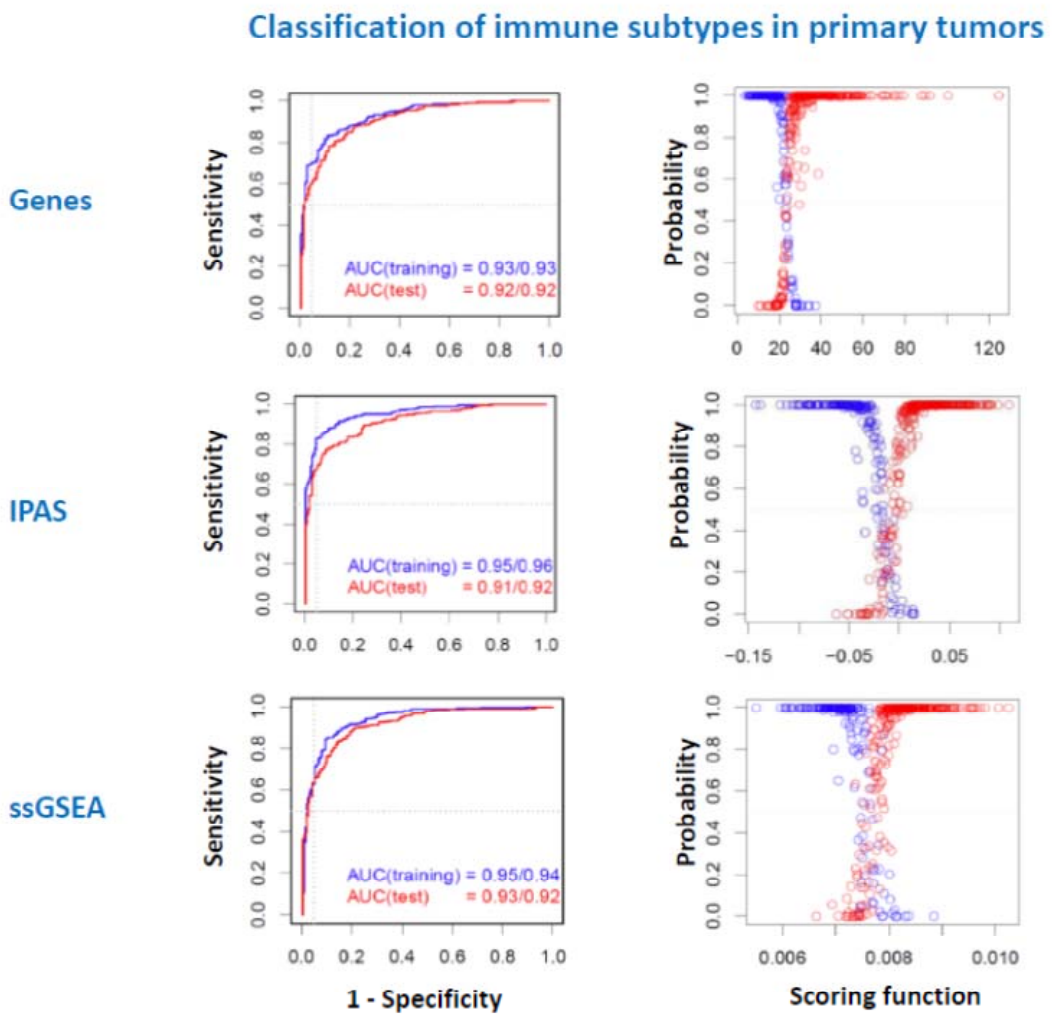
Classification of immune subtypes in primary tumors
Genes
IPAS
ssGSEA
Sensitivity
Probability
AUC(training) = 0.93/0.93
AUC(test) = 0.92/0.92
AUC(training) = 0.95/0.96
AUC(test) = 0.91/0.92
AUC(training) = 0.95/0.94
AUC(test) = 0.93/0.92
1 - Specificity
Scoring function

Figure S3.1

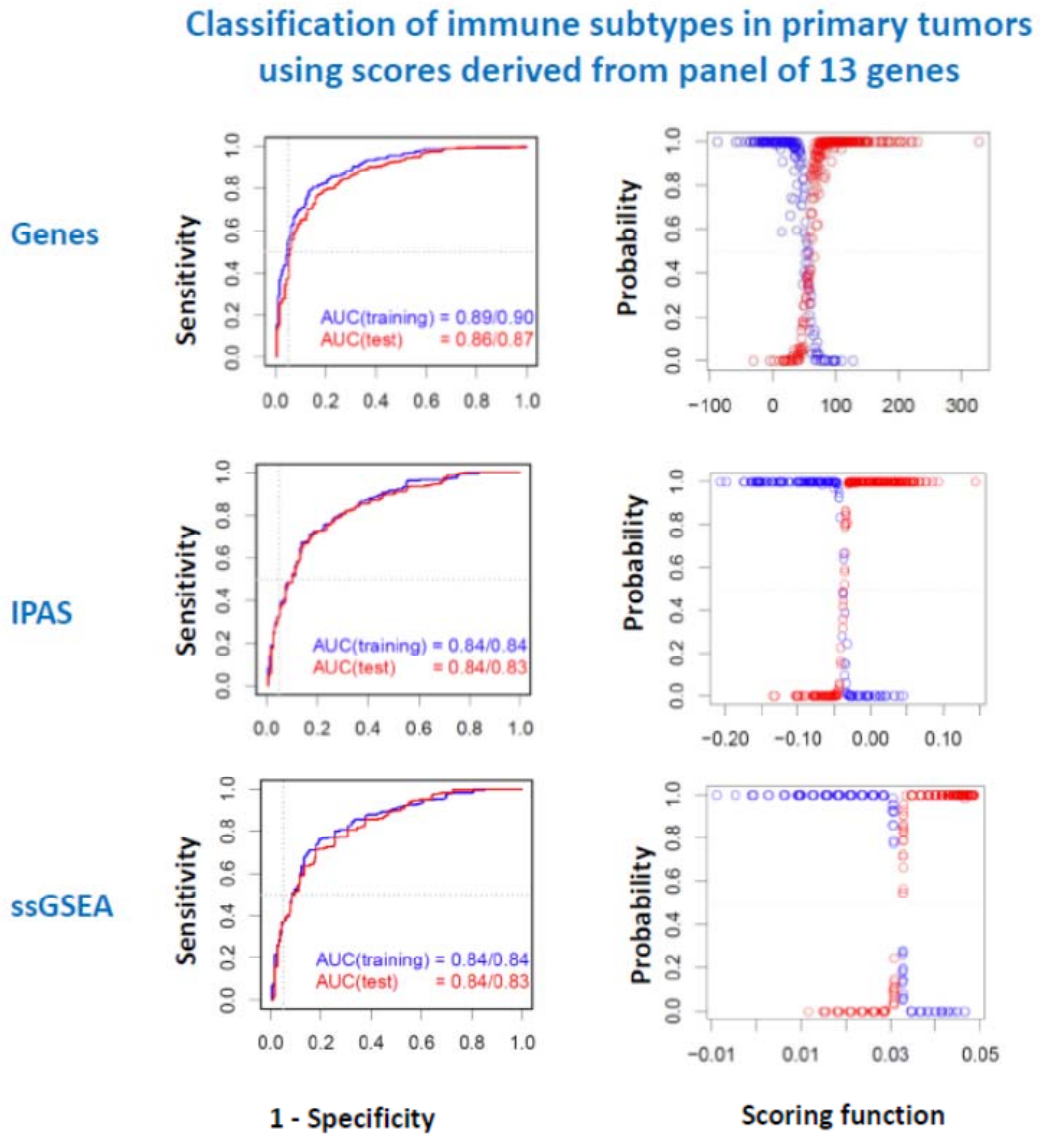
Classification of immune subtypes in primary tumors
using scores derived from panel of 13 genes
Genes
IPAS
ssGSEA
Sensitivity
Probability
AUC(training) = 0.89/0.90
AUC(test) = 0.86/0.87
AUC(training) = 0.84/0.84
AUC(test) = 0.84/0.83
AUC(training) = 0.84/0.84
AUC(test) = 0.84/0.83
1 - Specificity
Scoring function

Fig.S3.2

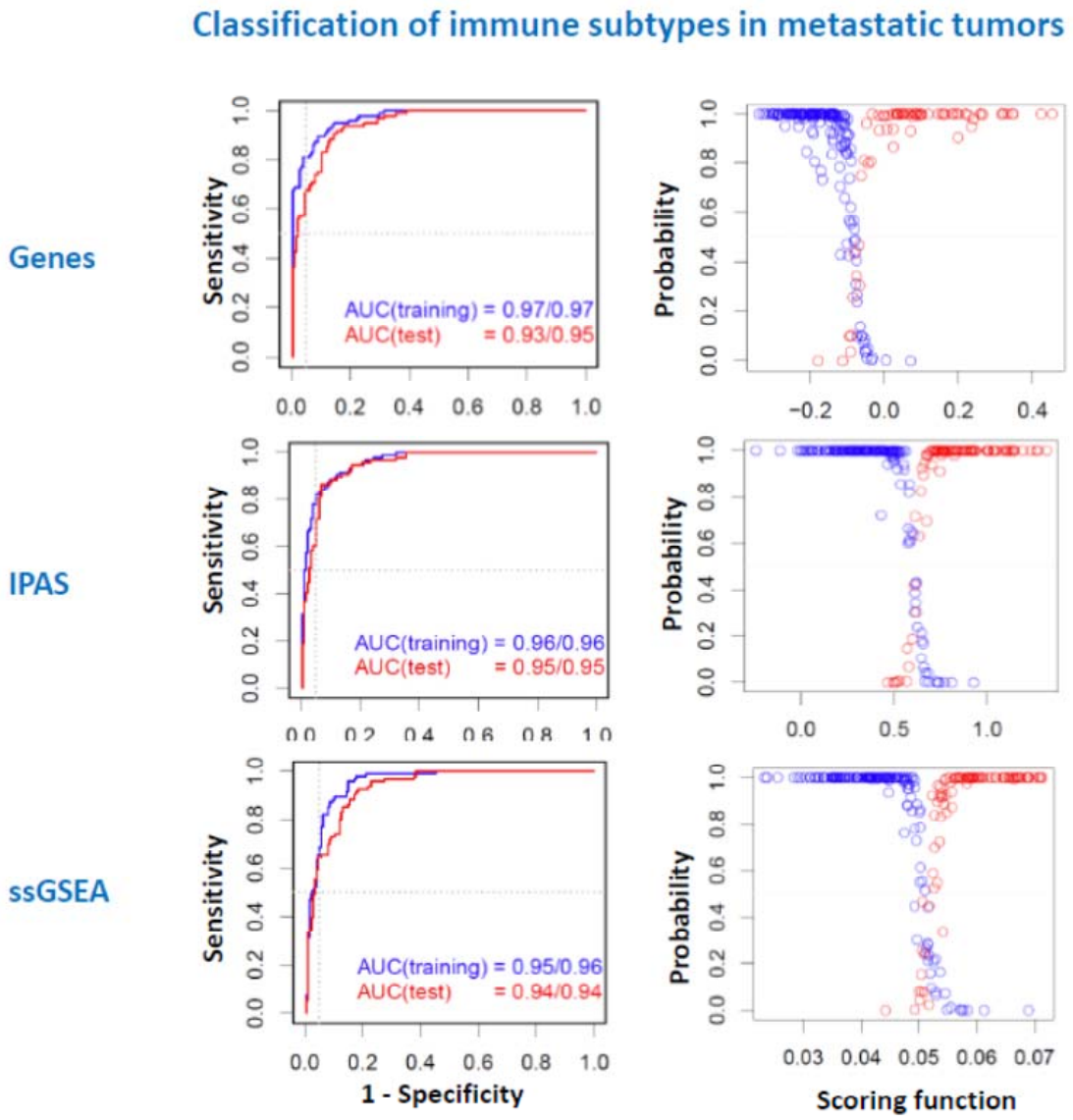
Classification of immune subtypes in metastatic tumors
Genes
IPAS
ssGSEA
Sensitivity
Probability
AUC(training) = 0.97/0.97
AUC(test) = 0.93/0.95
AUC(training) = 0.96/0.96
AUC(test) = 0.95/0.95
AUC(training) = 0.95/0.96
AUC(test) = 0.94/0.94
1 - Specificity
Scoring function

Figure S3.3.

Classification of immune subtypes in metastatic tumors using scores derived from panel of 13 genes

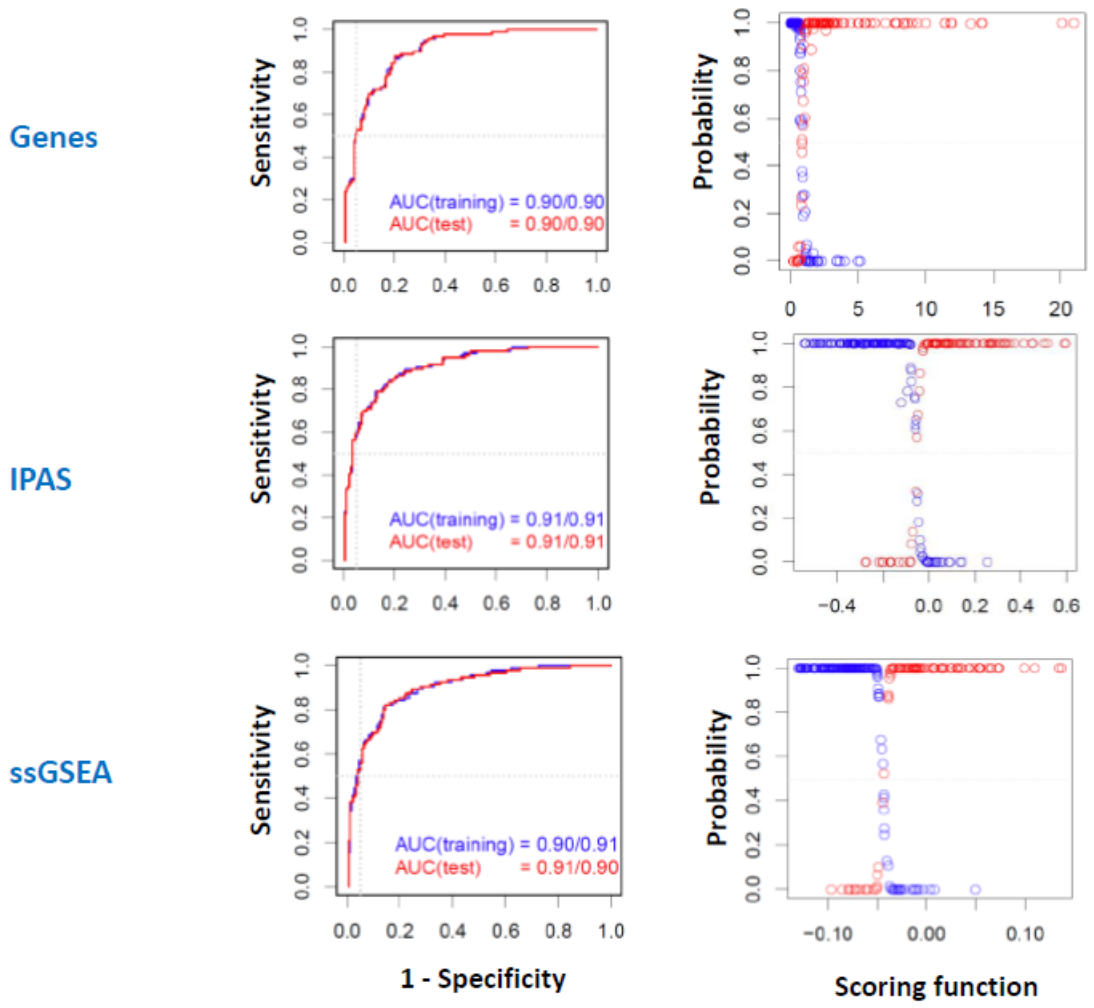

Fig.S3.4

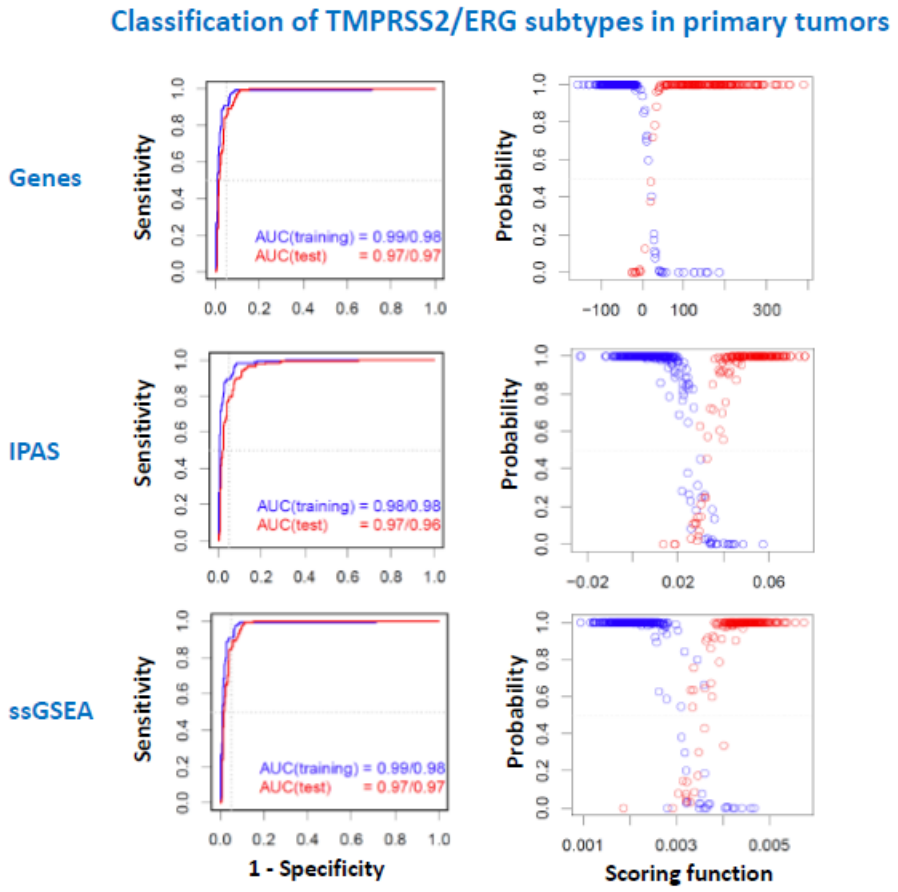
Classification of TMPRSS2/ERG subtypes in primary tumors
Genes
IPAS
ssGSEA
Sensitivity
Probability
AUC(training) = 0.99/0.98
AUC(test) = 0.97/0.97
AUC(training) = 0.98/0.98
AUC(test) = 0.97/0.96
AUC(training) = 0.99/0.98
AUC(test) = 0.97/0.97
1 - Specificity
Scoring function

Fig.S3.5

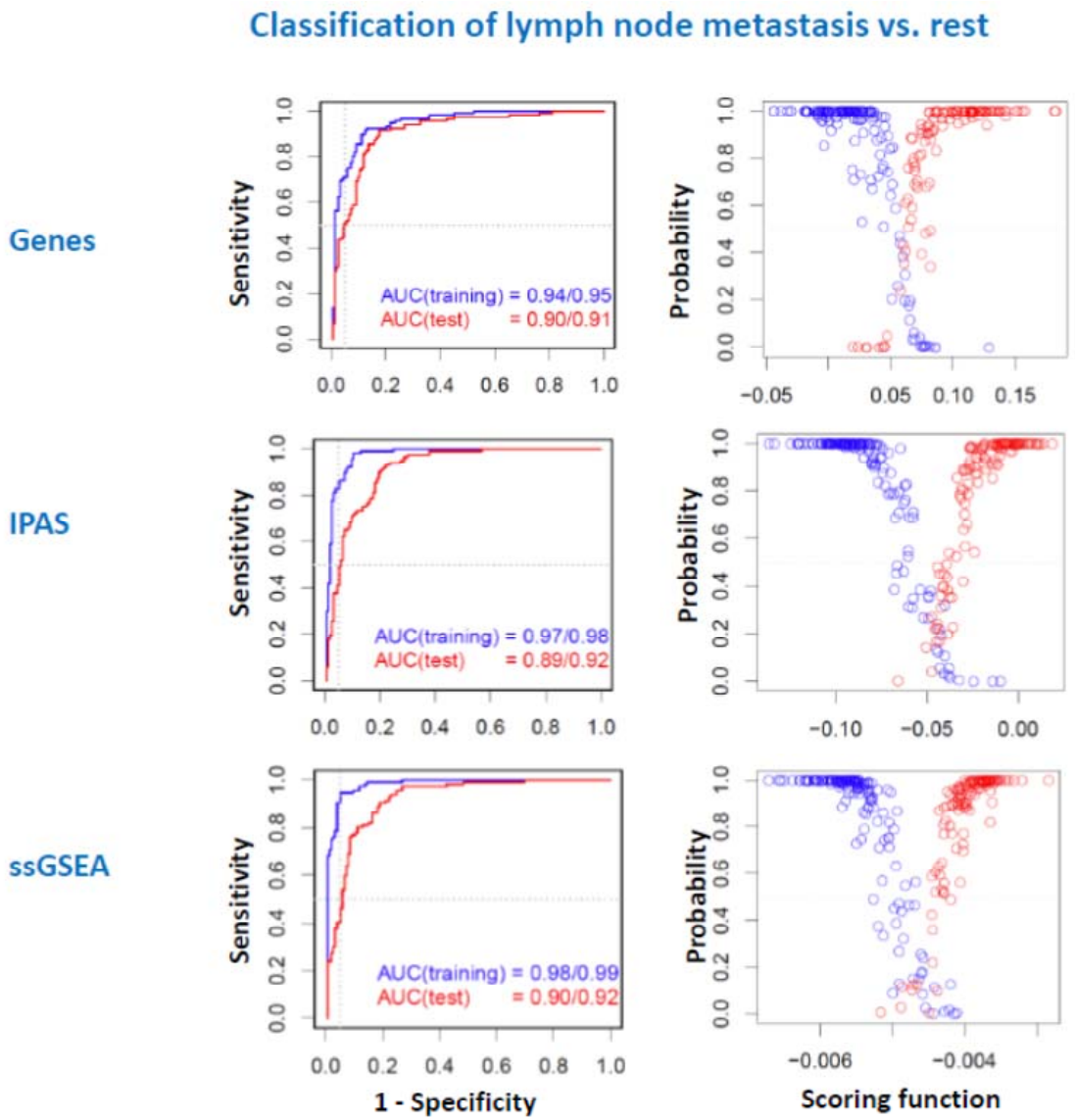
Classification of lymph node metastasis vs. rest
Genes
IPAS
ssGSEA
Sensitivity
Probability
AUC(training) = 0.94/0.95
AUC(test) = 0.90/0.91
AUC(training) = 0.97/0.98
AUC(test) = 0.89/0.92
AUC(training) = 0.98/0.99
AUC(test) = 0.90/0.92
1 - Specificity
Scoring function

Fig.S3.6

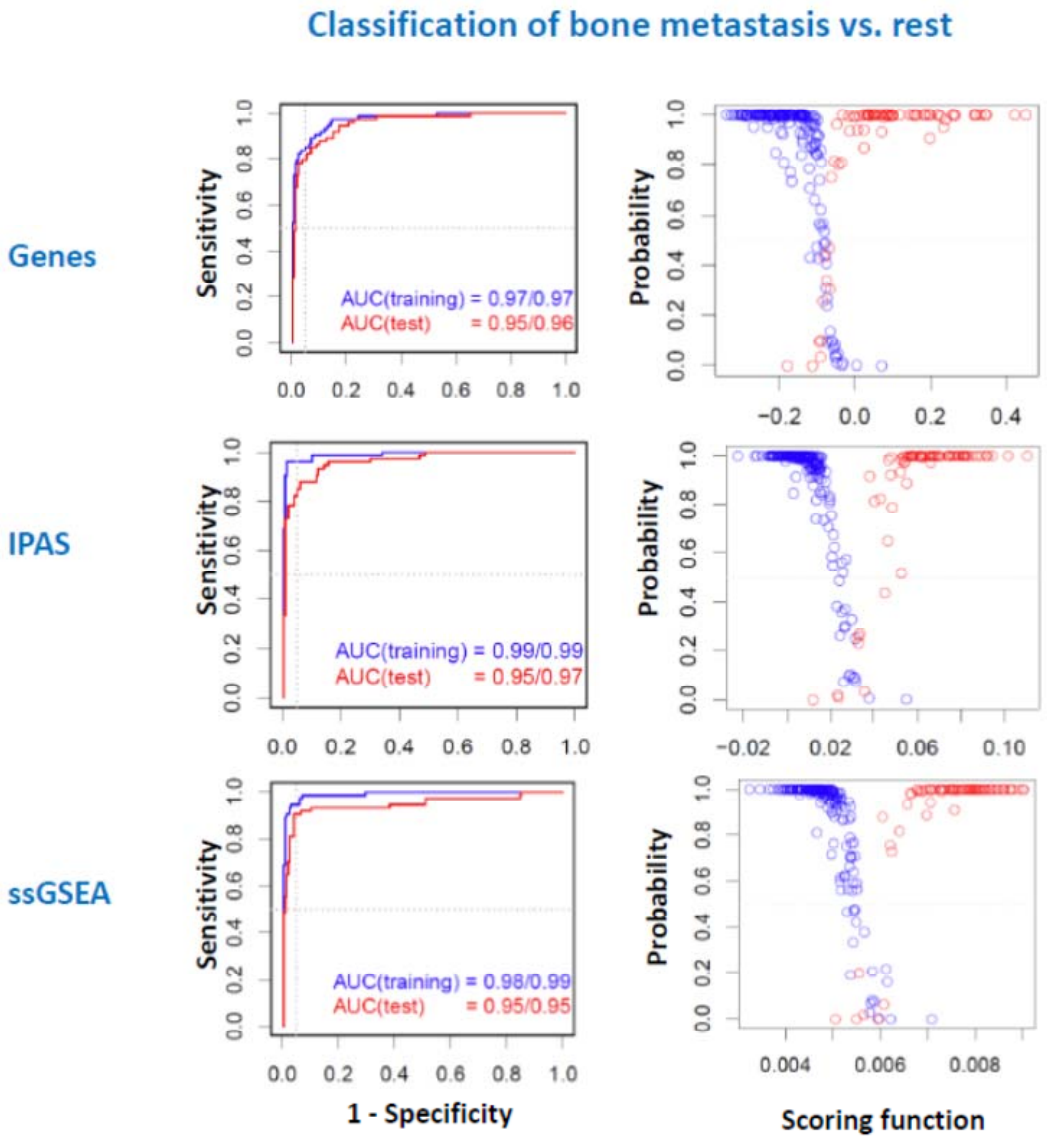
Classification of bone metastasis vs. rest
Genes
IPAS
ssGSEA
Sensitivity
Probability
AUC(training) = 0.97/0.97
AUC(test) = 0.95/0.96
AUC(training) = 0.99/0.99
AUC(test) = 0.95/0.97
AUC(training) = 0.98/0.99
AUC(test) = 0.95/0.95
1 - Specificity
Scoring function

Fig.S3.7

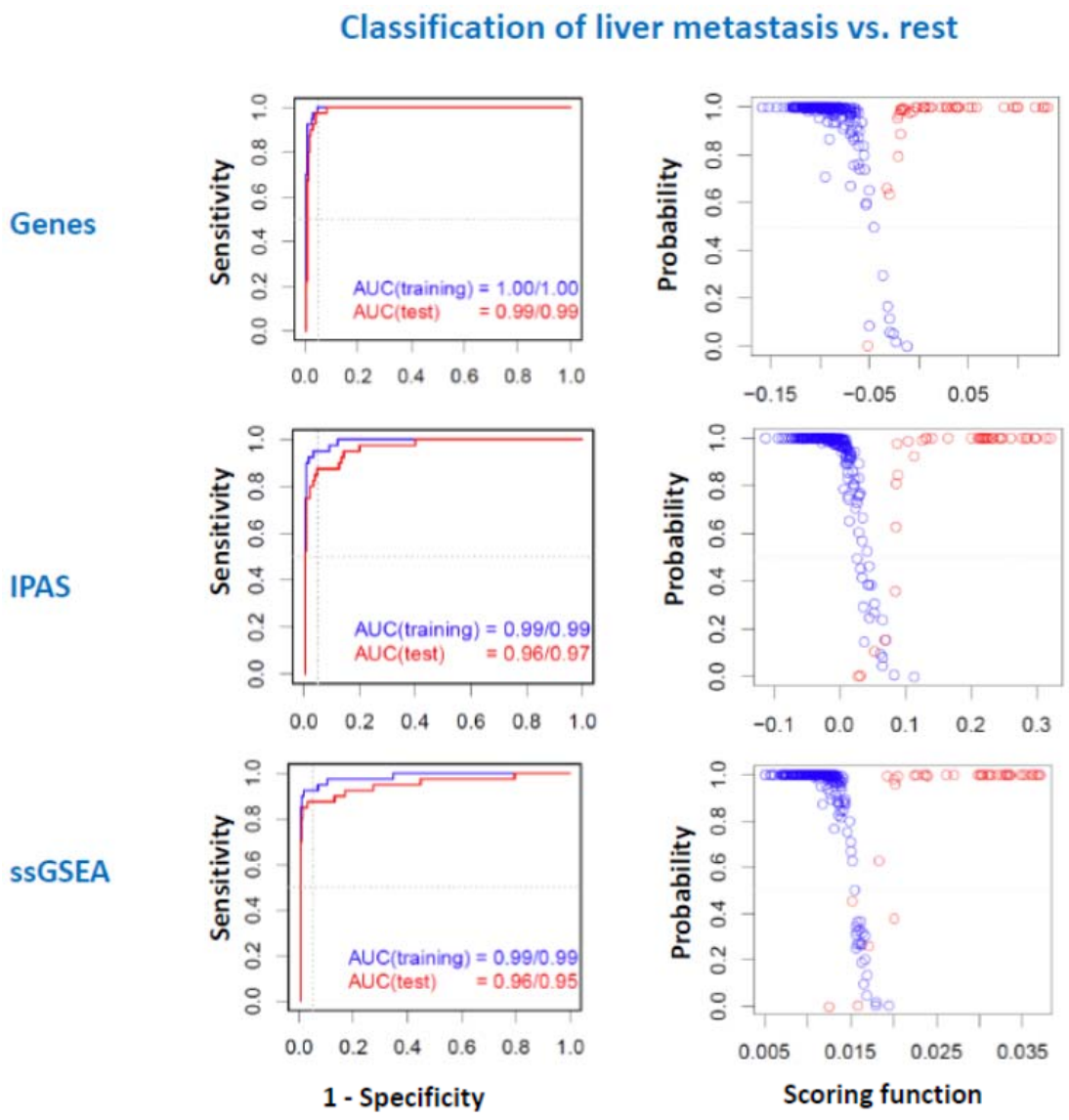
Classification of liver metastasis vs. rest
Genes
IPAS
ssGSEA
Sensitivity
Probability
AUC(training) = 1.00/1.00
AUC(test) = 0.99/0.99
AUC(training) = 0.99/0.99
AUC(test) = 0.96/0.97
AUC(training) = 0.99/0.99
AUC(test) = 0.96/0.95
1 - Specificity
Scoring function

Fig.S3.8