Research paper

The prognostic landscape of interactive biological processes presents treatment responses in cancer

Bin He a,1, Rui Gao a,b,c,1, Dekang Lv b, Yalu Wen d, Luyao Song b, Xi Wang a, Suxia Lin a, Qitao Huang a, Ziqian Deng b, Zifeng Wang a, Min Yan a, Feimeng Zheng e, Eric W.-F. Lam f, Keith W. Kelley g, Zhiguang Li b,*, Quentin Liu a,b,⁎⁎

a State Key Laboratory of Oncology in South China, Collaborative Innovation Center of Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou 510060, PR China
b Institute of Cancer Stem Cell, Dalian Medical University, Dalian 116044, PR China
c Department of Medical Oncology, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen 510275, PR China
d Department of Statistics, University of Auckland, New Zealand
e Department of Medical Oncology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510060, PR China
f Department of Surgery and Cancer, Imperial College London, Hammersmith Hospital Campus, London, W12 ONN, UK
g Laboratory of Immunophysiology, Department of Animal Sciences, College of ACES and Department of Pathology, College of Medicine, University of Illinois at Urbana-Champaign, Urbana, USA

A R T I C L E  I N F O

Article history:
Received 1 October 2018
Received in revised form 15 January 2019
Accepted 31 January 2019
Available online xxx

Keywords:
Prognosis
Biological processes
Treatment response
Cell-proliferation
Immune processes

A B S Т R А C Т

Background: Differential gene expression patterns are commonly used as biomarkers to predict treatment responses among heterogeneous tumors. However, the link between response biomarkers and treatment-targeting biological processes remain poorly understood. Here, we develop a prognosis-guided approach to establish the determinants of treatment response.

Methods: The prognoses of biological processes were evaluated by integrating the transcriptomes and clinical outcomes of ~26,000 cases across 39 malignancies. Gene-prognosis scores of 39 malignancies (GEO datasets) were used for examining the prognoses, and TCGA datasets were selected for validation. The Oncomine and GEO datasets were used to establish and validate transcriptional signatures for treatment responses.

Findings: The prognostic landscape of biological processes was established across 39 malignancies. Notably, the prognostic patterns distinguished response to treatment targeting specific biological processes. Applying this metric, we found that low tumor proliferation rates predicted favorable prognosis, whereas elevated cellular stress response signatures signified resistance to anti-proliferation treatment. Moreover, while high immune activities were associated with favorable prognosis, enhanced lipid metabolism signatures distinguished immunotherapy resistant patients.

Interpretation: These findings between prognosis and treatment response provide further insights into patient stratification for precision treatments, providing opportunities for further experimental and clinical validations.

Fund: National Natural Science Foundation, Innovative Research Team in University of Ministry of Education of China, National Key Research and Development Program, National Science Foundation of Guangdong, Science and Technology Planning Project of Guangzhou, MRC, CRUK, Breast Cancer Now, Imperial ECMC, NIHR Imperial BRC and NIH.

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Heterogeneity among tumors, which leads to differential treatment responses, remains the leading challenge for effective treatment [1]. Extensive efforts have been devoted to distinguish treatment responders from non-responders. Individual genes or gene signatures are commonly used as the determinants of treatment response [2]. Yet, each gene may be involved in many distinct biological processes, rendering its biological function highly context-dependent [3,4]. Gene signatures...
that contain a group of genes are more informative since they represent biological functions in defined contexts. Indeed, as genes in a signature are not functionally equal, the biological processes underlying these signatures make a major contribution to determining treatment responses [2]. Therefore, establishing gene signatures and the underlying biological processes is key to determining treatment response. However, how these response biomarkers are linked to treatment targeting biological processes remains poorly understood, which impedes the process in identifying determinants of effective treatment.

To link response biomarkers and treatment targeting biological processes, gene set-based approaches can be used [3]. Gene sets associated with cellular response to chemical/genetic perturbations have previously been developed in the Molecular Signatures Database (MSigDB) [5]. As treatments act by targeting specific biological processes, biological process gene sets (BPGSs) associated with particular clinical outcomes may reveal molecular features of treatment response. The prognoses of BPGSs among distinct cancer cohorts provide an opportunity to identify the underlying biological processes for treatment response. Prognostic evaluation in cancer datasets has hitherto been achieved by single-gene or signature-based approaches [2]. Recently, the PREdiction of Clinical Outcomes from Genomic (PRECOG) [6] and the Human Pathology Atlas [7] have provided a systematic approach to define the prognosis of individual genes, yet it is difficult to mine out the biological insights. Gene set analysis of gene-prognosis profiles could be an effective way to systematically evaluate the prognoses of BPGSs. The prognostic patterns of BPGSs among heterogeneous cancer types thus provide an approach to learn the determinant of prognosis as well as treatment response.

Here we report a gene-prognosis ranked Gene Set Enrichment Analysis (GSEA) [3] approach to evaluate the prognostic values of BPGSs in a variety of cancers. Based on the gene-prognosis information established from the tumor transcriptomes and clinical outcomes of ~26,000 patients, we systematically evaluated the prognostic values of BPGSs across 39 distinct cancer types. We found the prognoses of BPGSs vary among cancer types, and that the transcriptional contexts are responsible for the intertumor heterogeneity caused prognostic variations. Notably, we demonstrate that the transcriptional signatures and underlying biological processes associated with prognoses of BPGSs can distinguish treatment response for both anti-proliferation treatments and immunotherapy. The prognostic landscape of BPGSs reveals interactive biological processes for treatment responses, providing a novel approach for developing precision treatment strategies through available cancer datasets.

2. Materials and methods

2.1. TCGA, PRECOG z-scores, pathology atlas and gene expression data

TCGA datasets for 36 malignancies were downloaded from the BROAD GDAC Firehouse (https://gdac.broadinstitute.org/). Genome-wide gene-prognosis data were organized according to the meta-z score method [6]. Specifically, for each dataset, RNA-seq and clinical data were downloaded and matched. The association of each gene with survival outcomes was assessed via Cox proportional hazards regression using the ‘coxph’ function of the R ‘survival’ package. P values, z-score cox coefficients and hazard ratios with 95% confidence intervals were obtained for each gene. P values for each gene were transformed into meta-z scores. Weighted meta-z-scores were collapsed into a global meta-z-score for each gene as described (Table S1). In this manner, each gene is assigned a specific meta-z-score for each type of cancer, as well as a global meta-z-score across all cancer types. The z-score represents the capability of a gene to differentiate two subsets of patients with distinct survival outcomes. The ‘meta-z-scores’ for prognostic outcome of each gene in 39 malignancies were downloaded from the PRECOG project (http://pre cog.stanford.edu) [6]. Survival analysis results for all protein-coding genes in 17 TCGA major cancer types were downloaded from the pathology atlas database (http://www.proteinatlas.org/pathology) [7]. Gene expression values for patients with distinct sensitivity to therapies were downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo). Gene expression levels (mRNA Expression z-Scores (RNA Seq V2 RSEM)) in the TCGA datasets were downloaded from the cbioPortal database (http://www.cbioportal.org/) [8,9].

2.2. Gene set enrichment analysis and single sample GSEA

Genes were ordered in a decreasing fashion according to pan-cancer (unweighted) or individual cancer prognostic z-scores. Pre-defined gene sets (eight major collections, H and C1-C7) were obtained from the Molecular Signatures Database (MSigDB) [5]. Chemical and genetic perturbations (CGP) gene sets containing “target” or “response” as a
key word were selected to generate a "gmt" gene set file for analysis. A customized gene set was defined as a set of genes in specific gene signature. Ranked lists were submitted to GSEA using the 'PreRanked' tool of GSEA software [3]. Interpretation of enrichment score (ES) to prognostic value is similar to the gene expression based GSEA. For a randomly distributed gene set, ES will be relatively small, but if it is concentrated at the top (adverse prognosis) or bottom (favorable prognosis) of the list, or otherwise non-randomly distributed, then the ES will be correspondingly high. For GSEA on CCLE, cell lines were grouped as sensitive or resistant according to their sensitivity to cell-proliferation targeting compounds. Enrichment of gene sets in both groups was determined. For GSEA of GEO datasets, patients were grouped as sensitive or resistant according to the authors' instructions, and then analyzed with candidate gene sets. Significantly enriched gene sets were defined using a False Discovery Rate (FDR) q-value <0.25 and a nominal P value <0.05. All analyses were performed using GSEA v2.2.1 software with the pre-ranked list and 1000 data permutations. Leading edge genes were defined by GSEA as genes in the gene set that appear in the ranked list at, or before the point where the running sum reaches its maximum deviation from zero, interpreted as the core of a gene set that accounts for the enrichment signal.

To perform single-sample gene set enrichment (ssGSEA), normalized gene expression data (downloaded from the CCLE portal) were submitted to the GenePattern platform. The ssGSEA Projection program was used to calculate separate enrichment scores for each pairing of a submitted gene set expression, normalized by rank, and the weighting exponent was set as 0.75. All analyses were performed using GSEA v2.2.1 software with the pre-ranked list and 1000 data permutations. Leading edge genes were defined by GSEA as genes in the gene set that appear in the ranked list at, or before the point where the running sum reaches its maximum deviation from zero, interpreted as the core of a gene set that accounts for the enrichment signal.

2.4. Hierarchical clustering

Normalized enrichment scores (NES) of each hallmark gene set for individual cancers (Table S3) were subjected to Cluster 3.0 software and both gene set and cancer type were clustered by average linkage. A clustered heat map was analyzed and visualized by TreeView.

2.5. Validation in the TCPA dataset

Markers for enriched processes, such as cell proliferation and the immune system, were examined by the webpage survival analysis of The Cancer Proteome Atlas project [12] (TCPA, http://app1.bioinformatics.mudanderson.org/tcpa/_design/basic/index.html). Datasets of different cancer types were examined as follows: bladder urothelial carcinoma (TCA), breast invasive carcinoma (TCA), colon adenocarcinoma (TCA), glioblastoma multiforme (TCA), head and neck squamous cell carcinoma (TCA), kidney renal clear cell carcinoma (TCA), brain lower grade glioma (TCA), lung adenocarcinoma (TCA), lung squamous cell carcinoma (TCA), endometrial carcinoma (MDACC), ovarian carcinoma (Japan), stomach adenocarcinoma (TCA), uterine corpus endometrioid carcinoma (TCA) and TCGA women’s cancer (combined breast, uterine and ovarian cancer). Survival plots and log-rank P-values were downloaded for further analysis.

2.6. Oncomine based signature analysis

Differential expression analyses (Cancer vs Cancer Analysis) were performed in the Oncomine database [13]. Datasets in this study included: Astrocytoma (ASTR: Bredel Brain 2, Cancer Res, 2005; Freij Brain, Cancer Res, 2004; Shai Brain, Oncogene, 2003; Sun Brain, Cancer Cell, 2006; vandenBoom Brain, Am J Pathol, 2003; Yamanaka Brain, Oncogene, 2006), Glioblastoma multiforme (GBM: Bredel Brain 2, Cancer Res, 2005; Freij Brain, Cancer Res, 2004; Liang Brain, Proc Natl Acad Sci, 2005; Nutt Brain, Cancer Res, 2003; Ramaswamy Multicancer, Proc Natl Acad Sci, 2001; Shai Brain, Oncogene, 2003; Sun Brain, Cancer Cell, 2006; vandenBoom Brain, Am J Pathol, 2003; Yamanaka Brain, Oncogene, 2006).

Please cite this article as: B. He, R. Gao, D. Lv, et al., The prognostic landscape of interactive biological processes presents treatment responses in cancer, EBioMedicine, https://doi.org/10.1016/j.ebiom.2019.01.064
(all known human gene symbols) and n is the number of genes in the query set. FDR q-value is the false discovery rate analog of hypergeometric P-value after correction for multiple hypothesis testing according to Benjamini and Hochberg, and those genes sets with FDR q < .05 were selected.

2.8. Prognostic analysis of gene signatures

To assess the prognostic values of signatures, the normalized counts were scaled to a to z-score for each signature-gene, using the following function: \( z = \frac{(\text{normalized counts} - \text{mean counts across all samples})}{(\text{standard deviation of counts across all samples})} \). Subsequently, we used the average z-scores of all signature-genes to define a signature score for each single sample. Finally, Kaplan-Meier analysis was performed in all cases with stratification of risk sub-groups based on the median value of signature scores.

2.9. ImmuneNet analysis

For gene set analysis, the top 200 favorable prognostic gene signatures were analyzed using ImmuneNet [14] gene set analysis (http://immunet.princeton.edu/geneset/). The immune networks of ImmuneNet were used as features in a SVM classifier to generate gene predictions for the genes that were investigated. The resulting gene predictions and their probabilities are provided as Table S5. The proportion of genes in the submitted gene set (GS = 1) those not in the set (GS = 0) and their probabilities are provided as Table S5. The proportion of genes in the set with the CCLE were downloaded from the cBioPortal [8,9] database. Unsuspected consensus clustering was performed with nonnegative matrix factorization [16] (NMF, v 0.20.6, default Brunet algorithm, 50 and 100 iterations for the rank survey and clustering runs, respectively). A preferred cluster result was selected by considering proportion of genes in the top 200 favorable prognostic genes were scaled to a to z-score for clustering solutions between 2 and 5 clusters. Profiles of genes or samples were generated by reordering with silhouette widths in the consensus clusters.

2.10. Expression profile analysis in CCLE

Candidate genes were analyzed using the GENE-E tool of the CCLE database [15] (http://www.broadinstitute.org/ccle/data/browsee-analyses). To identify lymphocyte specific genes, expression data in the CCLE were downloaded from the cbioPortal [8,9] database. Unsupervised consensus clustering was performed with nonnegative matrix factorization [16] (NMF, v 0.20.6, default Brunet algorithm, 50 and 100 iterations for the rank survey and clustering runs, respectively). A preferred cluster result was selected by considering profiles of the cophenetic score for clustering solutions between 2 and 5 clusters. Profiles of genes or samples were generated by reordering with silhouette widths in the consensus clusters.

2.11. Functional immune genes

Genes in functional immune pathways were identified using the Gene Expression Commons database (https://gexc.stanford.edu). Gene sets of the mouse hematopoiesis model were selected by functional immune key words (e.g., dendritic, T cell, NK cell, etc.) inquiries.

2.12. Vector constructions, lentivirus production and transfection

To knockdown PLAT, PLAU and SERPINE1 in human cancer cell lines, we constructed shRNA or scramble (shGFP) in pLKO-Tet-On vector (addgene #21915, RRID: Addgene_21915). Briefly, vector was digested by AgeI and EcoRI (Thermo Scientific) for 15 min at 37 °C and recovered from agarose gel by DNA recovery kit (Tiangen). Annealed oligos were ligated with digested vector with T4 DNA ligase (Thermo Scientific) for 2 h at room temperature. Competent SbiB E. coli cells were transformed with the constructs, selected and identified by Sanger sequencing. The lentiviral vector (10 μg) and the two packaging viral vectors, pMD.2G (3.3 μg) and pspPax2 (6.7 μg), were co-transfected into 293 T (RRID: CVCL_0045) cells using Lipofectamine 2000 (ThermoFisher, #11668019, 1 μg/μL). The virus supernatant was collected at 48 h. MDA-MB-231 (bought from ATCC, RRID: CVCL_0062) cells were transfected with virus and selected with puromycin (Sigma-Aldrich, #P8833, 2 μg/ml) for six days.

2.13. Reverse transcription PCR (RT-PCR)

Cells were treated with doxycycline (Sigma-Aldrich, #D9891, 200 ng/ml) for 72 h. Total RNA was isolated using SuperPrep Cell Lysis kit (TOYOBO, #SCQ–101) as described by manufacturer and reverse transcribed with SuperPrep RT kit for qPCR (TOYOBO, #SCQ–101) for cDNA synthesis. Target genes were detected by real time PCR in CFX96 Touch™ according to manufacturer’s protocol. Transcripts were quantified relative to the housekeeping gene, GAPDH. The probes used for this study were listed in Table S7.

2.14. Cell viability assay

Cells were maintained with doxycycline (200 ng/ml) for 72 h before vehicle (DMSO), MLN8237 (Selleckchem, #S1133, 200 nM) or Palbociclib (Sigma-Aldrich, #PZ0199-PD0332991, 500 nM) treatment. Cells were seeded in 96-well plates at a density of 3 × 10^3 cells/well, allowed to adhere overnight, and treated with vehicle (DMSO) or drug for 72 h, after which they were processed for cell viability using Cell Counting Kit-8 reagent (Dojindo Molecular Technologies, #CK04). Each condition was performed in replicates of eight wells.

2.15. Statistical analysis

Statistical analyses were performed using SPSS version 13.0 (SPSS Inc.). Kaplan–Meier statistic and log-rank tests were performed to estimate the relevance of candidate markers in overall and disease-free survival of patients. Multivariate analysis was performed using the Cox proportional hazards model. Other P values were determined using the two-tailed Student’s t-test. Significant results were defined as P < .05.

3. Results

3.1. A prognosis-ranked GSEA for evaluating biological processes

To identify the interactive pattern of treatment targeting biological processes, we used cancer datasets to develop a stepwise bioinformatic approach (Fig. 1a). In this model, we proposed that response to a treatment in heterogeneous patients depends on two factors: 1) inactivation of the target and 2) prognostic contribution of the targeting biological process. An alternative treatment that successfully inactivate the target rescues type II, but not type I resistance. The mechanisms for type I resistance is key to stratifying patients and effective treatment. To this end, the prognostic values of BPGs were used to define patient cohorts where BPGs indicate favorable (BPG-Fav) or adverse prognosis (BPG-Adv). A comparative analysis between BPG-Fav and BPG-Adv patients was followed to establish gene signatures that distinguish BPG-Fav from BPG-Adv patients. Gene signatures and interacting biological processes were tested in treatment response datasets to determine their capacity to distinguish responding patients (Fig. 1b). The prognoses for individual genes (gene-prognosis scores) were examined by integrating gene expression patterns and clinical outcomes of cancer patients [6]. The 39 cancer types whose gene-prognosis scores generated by PRECOG from microarray gene expression data (Fig. 1b, Methods) were used for examining the prognoses of gene sets. In addition, 19 TCGA datasets (PRECOG matched cancer types, in-house generated gene-prognosis scores) were selected for validation (Table S1). To evaluate the prognostic values of gene sets, we performed an enrichment analysis of genes sets in the gene list decreasingly ranked by gene-prognosis scores, termed prognosis-ranked GSEA here. This prognosis-ranked GSEA defined two distinct prognostic classes indicating favorable and adverse prognosis, respectively (Fig. 1b) [3,4]. Pan-cancer gene-prognosis scores showed a high concordance between the TCGA and PRECOG datasets (Figs. S1a and S1b). Consistently, normalized enrichment score (NES) of the gene sets (hallmark in MSigDB)}
3.2. Prognoses of biological processes vary across cancer types

To assess the prognoses of key biological processes in cancer, we performed pan-cancer prognosis-ranked GSEA in 50 hallmark gene sets (MSigDB). Consistent with previous reports [6,7], cell-proliferation programs represented by cell-cycle related gene sets were found to be associated with adverse prognoses and enriched in AdvP, as 6 of the top 10 AdvP-enriched gene sets were cell-proliferation programs (gene sets with \( P < 0.05 \), FDR \( q < 0.25 \), Figs. 2a, left, 2a). In contrast, immune processes represented by immune-related gene sets were associated with favorable prognoses and enriched in FavP, as 6 of the top 10 FavP-enriched gene sets were immune processes (gene sets with \( P < 0.05 \), FDR \( q < 0.25 \), Figs. 2a, right, 2a). Next, the prognostic association of these processes was validated by a gene signature-based survival analysis in the PRC0geneV2 cancer datasets. Consistent with prognosis-ranked GSEA, leading edge genes of cell-proliferation programs (49 genes combining leading edge genes of E2F_TARGETS and MYC_TARGETS.V1 gene sets) were associated with adverse prognoses in 43 of 47 datasets (log-rank \( P < 0.05 \), Figs. 2b and 2d). In contrast, leading edge genes of immune processes (46 leading edge genes from INTERFERON_GAMMA_RESPONSE and INTERFERON_ALPHA_RESPONSE gene sets) forecasted favorable prognoses in 13 of 15 datasets (log-rank \( P < 0.05 \), Figs. 2c and 2d). In addition, “E2F targets” associated gene signature was coupled with adverse prognoses (Fig. 2b), while “IFNG response” and “IFNA response” gene signatures defined favorable prognoses (Fig. 2c). Collectively, these data demonstrate that cell-proliferation programs and immune processes are core processes of prognostic significance in cancer.

3.3. Cell-proliferation programs point to treatment response

The cancer-type dependent prognostic variations provide an opportunity to examine whether transcriptional context of BPGS prognosis could distinguish treatment response. To determine the representative cancer types, we defined malignancies that displayed the consistent variations in the cell-proliferation program, which is a principal treatment target in a variety of cancers. Using data from the Human Pathology Atlas [7], we found that cancer-type dependent prognostic variations in the cell-proliferation program were consistent with prognostic NES (Fig. 2e). Specifically, the prognoses of individual cell-proliferation leading edge genes were adverse in \( 11/17 \) cancer types, while favorable prognoses were observed in malignancies such as glioma and colon cancer (Fig. 3a). These prognostic patterns were also validated in TCPA datasets [12], where cell-cycle markers (cyclin B1/E) showed adverse prognoses in several cancers, but favorable in colon cancer (Figs. S3a and S3b). Similarly, cell-proliferation leading edge genes displayed elevated expression in “high risk” in breast cancer (Cell-cycle Con cancer, Fig. S3c), but displayed lower expression in “high...
**Working model:** Treatment efficacy = Inactivation\textsuperscript{Target} × Prognosis\textsuperscript{Target}

**Analysis:** Prognosis\textsuperscript{Target} = favorable

**Prognostic landscape:**

- 

**Cohort 1**

**PrognosisTarget = favorable**

**Gene signatures**

- **TCGA RNA-seq datasets**
  - Patients survival data
  - KM survival analysis (R 'survfit')
  - Survival associations: Metagene signatures GO sets, MSigDB, v5.1

**Drug response data**

- **TCGA RNA-seq datasets**
  - Patients survival data
  - Cox regression (R 'coxph')
  - Gene-prognosis data (Prognostic gene sets)
  - Signature validations

**Cell proliferation/immune markers**

- **TCGA RNA-seq datasets**
  - Patients survival data
  - Signature enrichments: cell line responses

**Cancer datasets:**

- **Gene signatures**
  - **Prognosis** Favorable
  - **Prognosis** Adverse

Please cite this article as: B. He, R. Gao, D. Lv, et al., The prognostic landscape of interactive biological processes presents treatment responses in cancer, EBioMedicine, https://doi.org/10.1016/j.ebiom.2019.01.064
To determine whether PVS distinguish responders from non-responders in patients, we inspected the transcriptional profiles of pretreated patients in eight anti-proliferation treatment-response cancer datasets (GSE106977, GSE19697, GSE41998, GSE50948, GSE58280, GSE66667, GSE66999, GSE76360) [20] [21]. GSEA in the transcriptional profiles confirmed that gene sets including CycleR, EMT, hypoxia and angiogenesis were enriched in the non-responders (Fig. 3e). Indeed, CycleR was upregulated in non-responders (NES < 0, Fig. 3f), and was significantly enriched in non-responders in three of them (GSEA nominal P < .05 in GSE58280, GSE76360 and GSE106977, Fig. 3f). In addition, stress response CycleR genes distinguished resistant patients from sensitive ones in the treatment-response dataset [22] (GSE5820 and GSE32962, Fig. S3i). Therefore, transcriptional signatures that were defined by prognostic variation could distinguish response to anti-proliferation treatments.

Co-enriched transcriptional signatures of tumors that distinguish tumor subgroups have been proposed to facilitate patient stratification [23]. To determine whether the cell-proliferation PVS is co-enriched in tumors, we explored the gene expression patterns in other cancer datasets [8,9] and found that PVS were co-enriched in tumor subgroups in multiple malignancies (Figs. 3g and 3h). Consistent with the prognostic NES patterns, the majority of glioblastoma (Cell-cycle Rev cancer) patients expressed low levels of CycleC genes, and a larger proportion of pancreatic cancer (Cell-cycle Con cancer) patients were found to be in the CycleC subgroup (Fig. 3g). Subsets of CycleR co-enriched tumors were observed in glioblastoma, pancreatic cancer and additional tumor cohorts (Figs. 3h and S3j). To determine whether CycleC and CycleR are associated with established molecular subtyping, we examined the correlation of CycleC and CycleR with molecular subtypes and IDH1 mutation in the TCGA glioblastoma dataset [24]. Interestingly, CycleC and CycleR showed a reverse pattern of correlation with glioblastoma subtypes. CycleC was expressed at higher levels in neural and proneural and lower in mesenchymal subtype, whereas CycleR was enriched in mesenchymal and lower in neural and proneural subtypes (Figs. 3i and S3k). These results are consistent with previous findings that mesenchymal subtype is associated with treatment resistance and relapse in glioblastoma [25]. In addition, CycleC was enriched IDH1 mutant subset, while CycleR was enriched in IDH1 wild-type subtype (Figs. 3j and S3l), consistent with previous report that IDH1 mutant glioblastoma patients were more sensitive to radiochemotherapy [26]. Thus, transcriptional signatures that were associated with both cell-proliferation prognosis and treatment response stratified patient subsets in multiple tumors.

3.4. Immune prognoses exhibit intrinsic intratumor variation

We next examined this prognosis-guided strategy in immunotherapy, which is at present hampered by their low response rates and a
Hallmark Sets enriched in AdvP

FAT_XX, ACYCLIC METABOLISM
ESTROGEN_RESPONSE_LATE
PKC_ATK/MITOR, SIGNALING
SOMATIC MUTATION
REACTIVE_OXIDANT SPECIES PATHWAY
ANOEAFIRNEN
CHOLESTEROL_HOMEOSTASIS
LR_RESPONSE_UP
NOREFATIN
OXIDATIVE PHOSPHORYLATION
EPITHELIAL_MESenchymal_TRANSITION
MITOTIC_SPINDLE
DNA_REPAIR
UNFOLDED_PROTEIN_RESPONSE
MTORC1_SIGNALING
MIC_TARGETS_1
COM_CHEKPOINT
MIC_TARGETS_2
ESF_TARGETS

Hallmark Sets enriched in FavP

HYPOGENESIS
COAGULATION
XENOBIOTIC, METABOLISM
ANGIOGENESIS
MTORC1_SIGNALING
APOPTOSIS
BLE_ACE, METABOLISM
COMPLEMENT
BLE_ACE, METABOLISM
TXN, SIGNALING, VIA_VN, FIB
ESTROGEN, RESPONSE, EARLY
IL2, SIGNALING, ON
SLJ, SIGNALING, ON
KRAST, SIGNALING, UP
INTERFERON_ALPHA, RESPONSE
INFLAMMATORY, RESPONSE
INHIBITION, CHAPERONE, RESPONSE
ALLOGRAFT, REJECTION

CycleLead

Cancer datasets: adverse prognosis
Cancer datasets: favorable prognosis

GO Cell-cycle programs

MITOSIS
MITOTIC_CELL_CYCLE
CELL_CYCLE_PHASE
CELL_CYCLE_GO_0007049
CHROMOSOMESEGREGATION
REGULATION_OF_MITOsis
SISTER_CHROMATID_SEGREGATION
DNADEPENDENT DNA REPLICATION
CELL_CYCLE_CHECKPOINT
DNA_METABOLIC_PROCESS

GO Immune processes

IMMUNE_SYSTEM_PROCESS
IMMUNE_RESPONSE
T_CELL_ACTIVATION
LEUKOCYTE_ACTIVATION
POS_IMMUNE_SYSTEM_PROCESS
DEFENSE_RESPONSE
CELLULAR_DEFENSE_RESPONSE
ADAPTIVE_IMMUNE_RESPONSE
REG_IMMUNE_SYSTEM_PROCESS
POS T_CELL_ACTIVATION

Please cite this article as: B. He, R. Gao, D. Lv, et al., The prognostic landscape of interactive biological processes presents treatment responses in cancer, EBioMedicine, https://doi.org/10.1016/j.ebiom.2019.01.064
lack of treatment biomarkers. We found that immune processes were the highest ranked biological processes associated with favorable prognoses (higher gene expression in patients with better survival outcomes). Core immune gene sets, such as antigen presentation, lymphocyte development and activation, T cell response and natural killer (NK) cell-mediated cytotoxicity were enriched in the FavP ($P < 0.05$ in GSEA, Fig. S4a). In addition, Gene Ontology (GO) annotation of the top 200 favorable genes (PRECOG pan-cancer) demonstrated that immune processes were highly enriched (Fig. S4b). These top favorable genes were found to be involved in well-defined immune networks (determined by ImmuNet [14]), such as those involved in antigen presentation, lymphocyte activation and NK cell mediated killing pathways (Fig. S4c and Table S5). Pan-cancer prognosis-ranked GSEA for gene sets representing 28 immune cell subpopulations and 10 immune cell processes indicated that the prognostic NES between TCGA and PRECOG were positively correlated (Pearson $r = 0.7295$, $P < 0.001$, Figs. S4d and S4e). Specifically, dendritic cells, CD8 T cells, B cells and NK cells were associated with favorable prognoses, whereas macrophage and regulatory T cells indicated adverse prognoses (Fig. S4d).

We then assessed the cancer type specific patterns of these immune processes by prognosis-ranked GSEA in cancer type specific gene-prognosis scores. Hierarchical clustering of prognostic NES for individual cancer types confirmed that dendritic cells, CD8 T cells, B cells and NK cells were associated with favorable prognoses in the majority of the 39 malignancies (Fig. S4f). However, adverse prognoses were observed in a subset of cancers, including pancreatic cancer, glioma and glioblastoma (red labels, Fig. S4f). TCGA datasets confirmed the favorable prognosis of immune processes in the majority of cancers, as well as the variations across different types of cancers (Fig. S4g). These results reveal the overall favorable prognoses of core immune processes and cancer-type dependent variations.

As inflammatory signaling in cancer cells promotes cancer progression [27], inflammatory genes from immune gene sets that are expressed in cancer cells might confer an adverse immune prognosis. To inspect this possibility, we defined lymphocyte specific genes by expression-based non-negative matrix factorization (NMF) clustering of immune leading edge genes in the CCLE dataset (Cluster 2, left panel, Fig. S5a). A notable portion of lymphocyte-specific leading edge genes exhibited favorable prognoses in malignancies with adverse immune prognostic NES, including breast cancer and lung adenocarcinoma, but not in cancers such as glioma (right panel, Fig. S5a). We examined the Gene EnrichmentProfiler database, performed NMF clustering and collected a subset of 32 Lymphocyte Specific Functional genes (LSF genes, Fig. S5b). Similar to the prognoses of immune cell gene sets (Figs. S4f and S4g), the prognoses of favorable LSF gene sets again varied among cancer types (Fig. 4a). The prognostic patterns of immune process were substantiated in representative cancer datasets from the SurvExpress [11], with immune leading edge genes indicating "low risk" in lung squamous cell carcinoma patients (ImmuneCon cancer, Fig. S5c), but "high risk" in glioblastoma (ImmuneRev cancer, Fig. S5d). Moreover, survival outcomes of individual immune leading edge genes were favorable in 13/17 cancer types whereas adverse prognoses were observed in malignancies such as glioma and pancreatic cancer according to datasets from the Human Pathology Atlas [7] (Fig. S5e). Therefore, non-hematopoietic specific gene sets partially contribute to adverse immune prognoses in certain situations, but not in malignancies such as brain cancers.

3.5. Immune interactive processes for treatment response

To uncover key factors that were responsible for immune prognostic variations, we established the immune PVS by combining transcriptional signatures of tumors from representative cancer types (MEDU, NEUB, LUSC and LUAD) in the Oncomine database [13]. These two pairs of immune representative cancer types (MEDU-NEUB and LUSC-LUAD) exhibited similar prognostic patterns for the 50 hallmark gene sets, with the exception of immune processes (Fig. 2e). The top 300 up-regulated and down-regulated genes of representative cancer types were combined to generate the gene signatures associated with prognoses (Fig. S5f and Table S6). Genes up-regulated in malignancies with adverse immune prognoses were defined as ImmuneRev signature (ImmuneR), and genes up-regulated in malignancies with favorable immune prognoses as ImmuneCon signature (ImmuneC). GO annotation of these signatures indicated a significant enrichment of lipid metabolism, regulatory T cell (Treg) and macrophage processes in the ImmuneR, whereas T-cell activation process was in the ImmuneC (Fig. S5g). To test whether patients stratified by PVS displayed distinct prognoses of immune processes, we performed survival analysis of immune genes (GO) in patient subsets classified by the ImmuneC and ImmuneR signatures, respectively. Consistently, the immune processes were associated with adverse prognoses in the ImmuneRhigh patients (Fig. 4b), whereas the immune processes were correlated with favorable prognoses in the ImmuneChigh patients (Fig. 4c).

Encouraged by results from the cell-proliferation program, we speculated that immune PVS may inform treatment response in immunotherapies. To test this possibility, we performed ssGSEA on transcriptional profiles of individual patients treated with anti-PD-1 immunotherapy (GSE78220 [23], GSE67501 [28], GSE81061 and A95951 [29]). In agreement with our hypothesis, enrichment scores for ImmuneC/ImmuneR (combined ssGSEA score for ImmuneC and ImmuneR) were significantly higher in responders than non-responders (Student’s t-test, $P = 0.097, 0.040, 0.008$ and $0.0201$, respectively, Fig. 4d). We carried out GSEA of ImmuneC, ImmuneR and hallmark gene sets in grouped patients (responders vs non-responders) in each dataset. Consistent with the report that EMT is associated with immunotherapy resistance, EMT was found to be enriched in the immunotherapy non-responder phenotype (GSE78220, $P = 0.004$ in GSEA, Fig. S5h) [23]. The ImmuneC leading edge genes, especially MGLL that is involved in membrane lipid metabolism processes, were up-regulated in immunotherapy-resistant patients (Fig. 4e). Notably, ImmuneR was enriched in the immunotherapy non-responder phenotype ($P = 0.040, 0.046$ and 0.081 in GSEA, upper panel, Fig. 4f), whereas ImmuneC was enriched in the responder population ($P = 0.108, 0.053$ and 0.155 in GSEA, lower panel, Fig. 4f). Indeed, among hallmark gene sets, ImmuneC, ImmuneR and lipid metabolic...
Glioblastoma TCGA N=166

Pancreatic cancer TCGA N=179

CycleR genes

CycleC genes

Normalized gene expression -1 1

Enrichment Score -10 10
GO_REG_NF_ΚB_IMPORT_INTO_NUCLEU
GO_STAT_CASCADE
GO_NEG_DEFENSE_RESPONSE
GO_ADAPTIVE_IMMUNE_RESPONSE
GO_RESPONSE_TO_GRAVITY
GO_ANGIOGENESIS
GO_RESPONSE_TO_CORTICOSTERONE
GO_RESPONSE_TO_ALCOHOL
GO_RESPONSE_TO_DEXAMETHASONE
GO_RESPONSE_TO_WOUNDING
GO_WOUND_HEALING
GO_RESPONSE_TO_ESTROGEN
GO_RESPONSE_TO_NUTRIENT
GO_REGENERATION
GO_RESPONSE_TO_MONOAMINE
GO_EPITHELIAL_TO_MESENCHYMAL_TRANSITION
GO_RESPONSE_TO_PLATELET_DERIVED_GROWTH_FACTOR
GO_RESPONSE_TO_ACID_CHEMICAL
GO_RESPONSE_TO_ORGANIC_CYCLIC_COMPOUND
GO_RESPONSE_TO_KETONE
GO_RESPONSE_TO_TTGF_BETA
GO_REGULARIZATION

Response No_Response Response No_Response Response No_Response Response No_Response

GSE106977 GSE19697 GSE41998 GSE50948 GSE5820 GSE6667 GSE66999 GSE76360

GSE106977: BRCA to taxanes
NES = -1.558
P = 0.014
FDR q = 0.269

GSE76360: BRCA to trastuzumab
NES = -1.361
P = 0.002
FDR q = 1.0

Please cite this article as: B. He, R. Gao, D. Lv, et al., The prognostic landscape of interactive biological processes presents treatment responses in cancer, EBioMedicine, https://doi.org/10.1016/j.ebiom.2019.01.064
signatures were consistently enriched in these datasets [Fig. 4g]. These results led us to propose that transcriptional signatures and the underlying biological processes that define immune prognostic variations are associated with clinical response to immunotherapies.

To test whether these immune PVS distinguish patient subsets in a general fashion, we explored expression of Immuc and Immur signatures in a panel of TCGA datasets. Indeed, Immuc genes were co-expressed and classified into distinct transcriptional subsets in multiple types of cancer (Figs. 4h and S5i). The Immur signature, which was expressed at low levels in melanoma patients (Fig. S5j), defined an Immur high (high expression of Immur signature) subset in lung cancers [Fig. 4i]. Notably, the proportions of Immuc high, Immur low, Immur high, and Immur low patient subsets were consistent with the prognoses of immune processes in those cancer types. For instance, the Immune Rev cancers (e.g., glioblastoma) consisted of a small Immuc low subset and a large Immur high subset, whereas the majority of Immune Con cancers (e.g., lung squamous cell carcinoma) were composed of the Immur low subset [Figs. 4h and 4i]. These findings established that the prognoses of immune processes are dependent on specific transcriptional contexts of tumors that are applicable to distinguish immunotherapy response. In summary, these conglomerate analyses reinforce the concept that prognosis of treatment targeting biological process is an indicator for treatment response. By integrating transcriptomic profiles and clinical parameters, this study presents an approach to establish transcriptional signatures and underlying biological processes as determinants that distinguish treatment responders across multiple cancers [Fig. 4j].

4. Discussion

Distinguishing treatment responders from non-responders is the leading challenge to effective treatment of cancer patients. Intensive experimental and data-mining efforts has been devoted to establish differentially expressed gene patterns to distinguish responsive patients. However, how this response biomarkers are linked to treatment targeting biological processes remains poorly understood. Inspired by the fact that clinical outcomes are the primary readout for the treatment evaluation efficacy, we have developed a prognosis-based strategy to define treatment biomarkers. Current approaches for prognoses of gene sets can be affected by the absolute level of gene expression, leading to biased results that are dominated by genes that are abundantly expressed in the gene set [4]. Here we linked gene-prognosis profiles to biological processes through prognosis-ranked GSEA. The prognosis-ranked GSEA approach is validated by the finding that biologically paired gene sets are consistently enriched, both in pan-cancer and cancer specific gene-prognosis profiles. This systematic approach establishes a prognostic landscape of biological processes that uncovers both general and cancer type specific prognostic patterns across 39 distinct malignancies. Based on the prognostic patterns, a comparative analysis of patient cohorts with distinct prognostic patterns establishes transcriptional signatures that distinguish patients cohorts. Notably, transcriptional signatures that define intertumor prognostic variations of treatment targeting biological processes distinguish treatment response, a concept that is proved in both cell-proliferation and immune processes.

The propositions of biological processes showed both general and cancer-type specific patterns that are consistent with the oncogenic or tumor suppressive roles of these processes. Biological processes, such as cell-proliferation programs, mTORC1 signaling, the unfolded protein response and glycolysis, are frequently associated with adverse prognoses. By contrast, immune processes, TP53 pathway and apoptosis process are linked to favorable prognoses. Notably, the prognoses of all gene sets are context dependent among different cancer types. Cell-proliferation programs, which are the primary targets for current anti-cancer therapeutics, generally confer poor outcomes but signify favorable prognoses in glioblastoma and colon cancer. Similarly, immune processes indicate favorable outcomes in cancers, such as melanoma and lung cancer, yet the same immune processes suggest poor prognoses in pancreatic and a number of brain cancers. These prognostic variations may result from differences in cell-of-origin, oncogenic drivers and microenvironment among different cancer types [1]. For instance, enriched cancer stem-like cells may contribute to favorable prognoses of cell-proliferation markers in brain cancer types, as stem-like cells are slow-cycling (low cell-proliferating activity) and responsible for treatment resistance and relapse [30]. Both tumor intrinsic and microenvironmental heterogeneity may contribute to the prognostic variations of immune processes. Pancreatic cancer and glioblastoma are characterized by infiltrating immune suppressive cells, which may contribute to adverse prognoses of immune processes in cancer types [31–33]. The prognostic variations not only strengthen the importance of our cancer type-specific prognostic resource, but also provide an opportunity to establish the determinants for prognosis and treatment response.

Tumor heterogeneity presents daunting challenges to the development of biomarkers and treatments for cancer [1]. To examine the potential of our approach in identifying treatment biomarkers, we focused on the cell-proliferation program, the primary target of cancer treatment [34]. Enrichment of stress response processes in CycleR is in accordance with the concept that cell-cycle activation in the context of stress impairs cell survival [35,36]. The enrichment of CycleR and stress response processes in the anti-proliferation non-responders phenotype in multiple datasets demonstrates the consistency of our approach. A consistent correlation with molecular subtypes of glioblastoma further confirms the clinical significance of these signatures [24]. In addition, co-enrichment of PVS in patients indicates that transcriptional programs distinguishing prognostic patterns and treatment response exist across multiple cancers. As anti-proliferative treatments are the primary anti-cancer strategies [37–39], these transcriptional signatures...
The prognostic landscape of interactive biological processes presents treatment responses in cancer.
and underlying biological processes have significant implications for patient stratification and effective treatment. The concept that PVS distinguishes treatment response has now been proven, and this is not limited to cell-proliferation program. This approach may also be applicable to treatments targeting other biological processes, especially treatments that are newly approved or under preclinical development, thereby providing a strategy for patient stratification based on existing cancer datasets.

Immunotherapy is now being evaluated in a wide range of cancers [38,39], yet the predictors and mechanisms of response remains poorly understood. For example, mutational loads have been demonstrated to benefit anti-CTLA4 responses in melanoma [40] and anti-PD1 response in lung cancer [41,42], while another study reports that mutational loads benefit survival, but not anti-PD-1 response [23]. The limited patient numbers may contribute to the substantial inconsistencies between these studies [43–52]. In contrast, our approach establishes transcriptional signatures and the underlying biological processes based on the molecular profiles of ~26,000 cases, and are applicable in multiple independent datasets. This high levels of consistency indicates a general mechanism of immunotherapy responses. Specifically, the T cell activation processes that are critical for successful immunotherapy are enriched in ImmuC. Similarly, enrichment of Treg and lipid metabolism processes in ImmuR is consistent with the immunosuppressive function of Treg [54] and requirement for lipid uptake and metabolism for survival of tissue-resident memory T cells [55]. Indeed, lipid metabolism controls the functional switch between immune protection and disease state in T cells [56]. Interestingly, cancer types with favorable immune prognoses, such as melanoma and lung cancer, are responsive to immunotherapies [38,39,57], whereas cancer types with adverse immune prognoses, such as pancreatic and several brain cancers, show a poor response to immunotherapy [58]. These findings further confirm a link between prognosis of treatment targeting process and treatment response.

The findings that prognoses are linked to treatment responses for specific biological processes may be useful for repurposing currently available targets and treatments based on the prognoses of related biological processes. The prognosis-repurposed vulnerabilities provide a resource of both therapeutic targets (e.g., MYC, EZH2, CDK4) and interventions (e.g., Rapamycin, Aplidin, Darapladib) that will facilitate development of novel therapeutic approaches across 39 human malignancies. This cancer type-dependent property of treatment response establishes resources that repurpose established treatments in specific cancer types. Identifying therapeutic vulnerabilities for cancers, especially those malignancies for which there are currently no effective treatments, is urgently needed. Therefore, our analysis also has the important advantage of presenting valuable resources indicative of potential treatment approaches for distinct malignancies. Although we have tried our best to optimize these analyses by keeping the versions of data consistent among different online tools and databases, there’re still several limitations in this study. Firstly, the prognostic analyses here are based on overall survival because the lack of recurrence-free survival information, which is more relevant to treatment response in cancer datasets. Secondly, the prognostic landscape built by GSEA is more exploratory and less accurate as predictive biomarkers here, and more efforts are needed before their clinical application [4]. Thirdly, although several findings from our analyses are consistent with previous experimental publications, we provided only cell-based experimental validations of CycloR signature. Since both tumor-specific and immune cell-specific factors should be considered in responses to immunotherapy [43–52], immune signatures are not experimentally validated because responses to immunotherapy are not well recapitulated in cell based experiments in vitro. Future in vivo studies will be considered for cell-proliferation, immune processes and other biological processes of potential interest.

Collectively, by integrating transcriptomic profiles, clinical outcomes and biological processes, we have now characterized the prognostic landscape of biological processes for a variety of cancers. These prognoses show intertumor variation patterns, which is a valuable property needed to define representative cancer types and transcriptional determinants of prognostic variation. This link between prognosis and treatment response establishes a strategy to stratify sensitive and resistant patients using transcriptional signatures that define prognostic variations. The concept that prognosis of target distinguishes treatment response may be applicable to a variety of treatments, providing potential biomarkers for further experimental and clinical validations. These findings highlight the utility of bioinformatic resources to develop biomarkers and treatment strategies that will allow more precisely targeted and immunotherapies for cancer patients.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2019.01.064.

Acknowledgements

We thank all members of Liu laboratory and Li laboratory for their critical comments and technical support.

Funding sources

This work was supported by the National Natural Science Foundation of China (81630005, 81130040, and 81573025 to Q.L., 81773166 to Z.W.), the Innovative Research Team in University of Ministry of Education of China (IRT13049 to Q.L.), the National Key Research and Development Program of China (2017YFA0505600-04 to Q.L.), the National Science Foundation of Guangdong (2016A030311038 and 2017A030313608 to Q.L., 2017A020215098 to Z.W.), the Science and Technology Planning Project of Guangzhou (201604020163 to Q.L.), and MRC (MR/N012097/1), CRUK (A12011), Breast Cancer Now (2012MayPR070; 2012NovPhD016), the Cancer Research UK Imperial Centre, Imperial ECMC and NHIR Imperial BRC to E. W.-F. Lam and the NIH grant (RO1 SUB UT00000712 to K.W. Kelley).

Fig. 4. Immune interactive processes indicate immunotherapy response (a) The prognostic NES of favorable LSF gene set (32 genes that are specifically expressed in lymphocytes that exhibited favorable prognostic z-scores in > 60% (24/39) PRECOG cancer datasets) in 19 matched TCGA (blue bars) and PRECOG (black bars) cancer types. (b) and (c) Kaplan Meier survival curves of patients with high (H) and low (L) gene expression (low, L; high, H) were defined according to the median value of gene set expression of the cohort (n = 118/117/117/117 for four groups in SKCM, and n = 125/125/125/125 for four groups in LUSC). P values were determined by log-rank test. (d) The combined ssGSEA enrichment scores of ImmuC and ImmuR gene sets (score ImmuC-score ImmuR) in individual patients from four independent datasets (GSE87501, GSE78220, GSE91061 and A5951). Patients were classified as responders (Res, black dots) and non-responders (Non, red dots) to immunotherapy according to their clinical outcomes in each dataset. P values were calculated using a two-tailed Student’s t-test. (e) Expression of MCL in individual patients from four independent datasets (GSE87501, GSE78220, GSE91061 and A5951). Patients were classified as responders (Res, labeled in black) or non-responders (Non, labeled in red) to immunotherapy according to their clinical outcomes. P values were calculated using a two-tailed Student’s t-test. (f) The NES of ImmuR (upper panel) and ImmuC (lower panel) in three independent GEO datasets. GSEA were performed in transcription of patients classified as responders and non-responders according to their response to immunotherapy. (g) The NES of ImmuC, ImmuR and Hallmark gene sets in three GEO datasets (black, GSE67501; blue, GSE91061; red, GSE78220; grey, A5951). GSEA were performed in transcription of patients classified as responders and non-responders according to their response to treatment. (h) and (i) Hierarchical clustering of ImmuR (h) and ImmuC (i) genes in three TCGA datasets, including glioblastoma (upper panel), lung adenocarcinoma (middle panel) and lung squamous cell carcinoma (lower panel). Yellow boxes indicate distinct clustered subsets with high gene expression. Sidebars, gene expression index. (j) A summary of the strategy to identify interactive biological processes for treatment responses and to stratify treatment responders from non-responders. The prognoses of BPGSs were systematically evaluated among cancer datasets. Differentially expressed genes between BPGS-favorable and BPGS-adverse patients were defined to identify the underlying biological processes and to distinguish responder patients. The concept has now been proven in treating treatments targeting both tumor (anti-proliferation treatment) and non-tumor cells (immunotherapy).
Declarations of interests

The authors declare that no conflict of interest exists.

Author contributions

Bin He and Rui Gao participated in design, acquisition, analysis and interpretation of data and manuscript preparation for the entire project. Dekang Lv, Yalu Wen, Luyao Song and Ziqian Deng performed the assembly of TCGA datasets, NMF analysis and statistical analysis of validation datasets. Xi Wang, Suxia Lin, Qiaoao Huang, Min Yan, Zifeng Wang, Feiming Zheng, Eric W.-F. Lam, Keith W. Kelley and Zhiguang Li participated in manuscript editing and making critical revisions. Quentin Liu participated in conception, design, analysis and interpretation of data in the entire project and made critical revisions in the manuscript.

References

[1] Bedard PL, Hansen AR, Ratain MJ, et al. Tumour heterogeneity in the clinic. Nature 2013;501(7467):355–64.
[2] Nevins JR, Potti A. Mining gene expression profiles: expression signatures as cancer phenotypes. Nat Rev Genet 2007;8(8):561–9.
[3] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102(43):15545–50.
[4] de Loouw CA, Neale BM, Heskes T, et al. The statistical properties of gene-set analysis. Nat Rev Genet 2016;17(7):353–64.
[5] Liberzon A, Birger C, Thorvaldsdottir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1(6):417–25.
[6] Gentles AJ, Newman AM, Liu CL, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med 2015;21(8):938–45.
[7] Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome. Nature 2015;526(7577):343–7.
[8] Lim M, Xia Y, Bettegowda C, et al. Current state of immunotherapy for glioblastoma. Cancer Cell 2010;17(1):98–105.
[9] Saygin C, Materi D, Majeti R, et al. Targeting cancer stemness in the clinic: from hype to hope. Cell Stem Cell 2019;24(1):25–40.
[10] Rees MG, Cheah JH, Seashore-Ludlow B, et al. Harnessing connectivity in a large-scale small-molecule sensitivity dataset. Cancer Discov 2015;5(11):1091–1108.
[11] Saygin C, Matei D, Majeti R, et al. Cell cycle activation linked to neuronal migration in the cerebral cortex. J Neurosci 2003;23(6):2131–40.
[12] Shukla SA, Bachireddy P, Schilling B, et al. Cancer-germline antigen expression dissects T cell immune repertoire across human cancers. Nat Rev Immunol 2010;10(7):490–500.
[13] Hamid O, Robert C, Daud A, et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. N Engl J Med 2013;369(2):134–45.
[14] Royal RE, Levy C, Turner K, et al. Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. J Immunother 2010;33(8):826–33.

Please cite this article as: B. He, R. Gao, D. Lv, et al., The prognostic landscape of interactive biological processes presents treatment responses in cancer, EBioMedicine, https://doi.org/10.1016/j.ebiom.2019.01.064