Tang *et al. Genome Medicine* (2025) 17:150
https://doi.org/10.1186/s13073-025-01572-z

Genome Medicine

RESEARCH Open Access

Reference-guided computational framework identifies microenvironment metabolic subtypes and targets using pan-cancer single-cell datasets

Ke Tang[1,2,3†], Ya Han[1,2,3†], Dongqing Sun[1,2,3], Xin Dong[1,2,3], Tong Han[1,2,3], Hailin Wei[1,2,3], Wenwen Shao[1,2,3], Junjie Hu[4], Zhaoyang Liu[1,2,3], Lele Zhang[5], Taiwen Li[6], Peng Zhang[4], Qiu Wu[1,2,3*] and Chenfei Wang[1,2,3,7,8*]

Abstract

Background Metabolic reprogramming is a hallmark of cancer; however, the mechanisms driving metabolic heterogeneity across diverse cell types in the tumor microenvironment remain poorly understood. Most existing methods predict metabolic states at the pathway level but rarely map reaction-level alterations to their upstream regulators, thereby constraining both interpretability and translational relevance.

Methods We developed MetroSCREEN, a reference-guided computational framework that infers reaction-level metabolic flux propensity and nominates upstream regulators from bulk and single-cell transcriptomes. MetroSCREEN uses a fast enrichment-based procedure to quantify reaction-level metabolic activity. To characterize metabolic regulons, it integrates intrinsic gene-regulatory signals with extrinsic cell–cell interaction cues, then applies a robust multi-evidence ranking scheme to combine these information sources, and finally employs a constraint-based causal discovery module to infer regulatory directionality.

Results MetroSCREEN accurately predicts reaction-level metabolic activities and their upstream regulators, as demonstrated using paired transcriptomic–metabolomic datasets from the cancer cell lines. We further validated predicted regulators with in-house single-cell CRISPR screens in PC9 cells targeting metabolic regulators. Applying MetroSCREEN to a pan-cancer single-cell atlas of more than 700,000 fibroblasts and myeloid cells across 36 cancer types, we identified ZNF281 and STAT1 as key regulators of collagen metabolism, which is elevated in extracellular-matrix–associated fibroblasts and macrophages at tumor margins. By contrast, APOE and KLF7 regulate sphingolipid metabolism and antigen presentation in macrophages. Leveraging extensive tumor profiles, MetroSCREEN also delineates metabolic subtypes and regulators associated with patient survival and response to immunotherapy.

†Ke Tang and Ya Han contributed equally to this work.

*Correspondence:
Qiu Wu
qiu_wu@tongji.edu.cn
Chenfei Wang
08chenfeiwang@tongji.edu.cn

Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Conclusions MetroSCREEN is a robust and scalable approach for characterizing metabolic heterogeneity and pinpointing metabolic regulators at single-cell resolution, unveiling novel antitumor targets for future metabolic interventions. The source codes of MetroSCREEN is available at the Github site https://github.com/wanglabtongji/MetroSCREEN.

Keywords Single-cell transcriptomics, Metabolic regulation, Large-scale data integration, Tumor microenvironment, Immunotherapy response

Background

Cell metabolism exhibits remarkable flexibility, allowing it to adapt to the demands of growth and homeostasis [1]. In pathological conditions such as tumors, the environment frequently becomes acidic and nutrient-depleted. These changes significantly affect cellular metabolic pathways and signaling networks, disrupting the expression of metabolic regulators and leading to widespread metabolic reprogramming [2]. A prominent example is the Warburg effect, wherein malignant cells preferentially rely on glycolysis for energy production, generating lactate even in an aerobic environment [3]. This metabolic adaptation satisfies the increased energy requirements essential for the rapid proliferation of malignant cells. Additionally, similar metabolic adaptations can occur in immune and stromal cells within the tumor microenvironment (TME), influencing their maturation, activation, and overall functions [4–6]. Despite the extensive research on malignant cells, investigations into the metabolic reprogramming of other components of the TME and their potential interactions remain scarce.

Metabolic reprogramming is shaped by both intrinsic and extrinsic cellular factors [7–9]. Elucidating the upstream regulatory mechanisms governing this process could uncover novel targets for metabolic therapies. Intrinsic factors encompass transcription regulators (TRs) and signaling transduction pathway components [10, 11]. For example, the TR MYC drives glucose metabolism by upregulating GLUT1 and key glycolytic enzymes such as HK2, PFK-M, and ENO1 [12, 13]. MYC also modulates lipid metabolism by regulating genes involved in fatty acid and cholesterol synthesis, including ATP-citrate lyase, acetyl-CoA carboxylase alpha, and fatty acid synthase [14]. Similarly, intrinsic signaling pathways like PI3K-AKT enhance metabolite uptake and glycolysis through downstream TRs [15]. Extrinsic factors, such as cytokines (e.g., IL-10, TNF-α), further regulate metabolic processes by altering the balance between glycolysis and oxidative phosphorylation (OXPHOS) in macrophages [16–19]. Despite the recognized importance of these factors, systematic elucidation of the regulatory mechanisms and key drivers underlying metabolic reprogramming remains scarce, yet it is crucial for developing potential therapeutic interventions.

Recent technological advancements have significantly advanced our understanding of metabolic heterogeneity. Techniques such as Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry (MALDI-MS) now enable the detection of over 100 metabolites at a cellular resolution of 5 μm, facilitating metabolite imaging and identification at single-cell resolution [20–24]. However, large-scale metabolic profiling at this resolution remains challenging due to low throughput and the time-intensive, technically complex nature of data analysis [25, 26]. Additionally, current methodologies lack the capability to perform simultaneous metabolite measurements and transcriptomic profiling, which is critical for dissecting cell-type-specific metabolic heterogeneity. To address the interconnected nature of multi-omics regulation, two key strategies have emerged for reconstructing metabolic states from single-cell transcriptomic data. The first employs Flux Balance Analysis (FBA), which assumes a quasi-steady metabolic state and identifies reaction flux configurations that optimize biologically relevant objective functions. Representative tools in this category include scFBA [27], scFEA [28], COMPASS [29], and METAFlux [30]. The second strategy leverages statistical methods to infer metabolic states by evaluating gene enrichment within predefined pathways [31–33], exemplified by ssGSEA [31], AUCell [34], and VISION [35]. While existing methods predict metabolic states, they predominantly focus on pathway-level scores and neglect upstream regulatory mechanisms. Furthermore, many approaches suffer from limited scalability, sensitivity, and computational efficiency when applied to single-cell datasets. To overcome these limitations, we developed MetroSCREEN, an efficient computational framework that employs a statistical model to infer reaction-level metabolic states while integrating intrinsic gene regulation and extrinsic cellular interactions to predict metabolic regulators with high accuracy.

To demonstrate MetroSCREEN's scalability on large single-cell datasets, we applied it to a pan-cancer scRNA-seq atlas encompassing 236,510 fibroblasts and 504,608 myeloid cells derived from 735 patient samples spanning 36 cancer types. Through this analysis, we constructed MetroTIME, the first pan-cancer metabolic map which delineates distinct energy utilization patterns and associated biological features of key stromal and immune metabolic subtypes within the TME. MetroTIME reveals robust associations and identified upstream regulators for collagen metabolism and extracellular matrix

(ECM)-remodeling fibroblasts, as well as sphingolipid metabolism and antigen-presenting/phagocytic functions in macrophages. We further investigated the clinical relevance of these subtypes, linking them to patient survival outcomes and immunotherapy response. Our study not only introduces a scalable computational framework for identifying metabolic adaptations and their regulatory drivers but also uncovers actionable targets for future metabolism-targeted anti-cancer therapies.

Methods

Materials

Cancer cell line data

To benchmark MetroSCREEN, we obtained cancer cell-line metabolomics, genome-scale transcriptomic, and drug-sensitivity datasets from the Cancer Cell Line Encyclopedia (CCLE) via the DepMap Public 23Q4 release [36] (https://sites.broadinstitute.org/ccle/). The dataset includes metabolite abundances for 225 metabolites and transcriptomes from 928 cell lines. For quality control, tumor types in which a given metabolite was quantified in fewer than 10 cell lines were excluded. After filtering, 13 tumor types remained. We applied principal component analysis (PCA) for dimensionality reduction and clustered cell lines using the top 10 principal components in Seurat [37] (v4.3.0) with default parameters. Drug-sensitivity data cover 1,500 compounds across 578 cancer cell lines, with sensitivity summarized as the area under the dose–response curve (AUC). Lower AUC values indicate greater sensitivity.

Single-cell perturbation data

To further benchmark the accuracy of the MetroSCREEN in predicting metabolic regulators, we obtained the single-cell perturbation dataset from the SRA under BioProject PRJNA831566 [38] (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA831566). The dataset includes a genome-scale Perturb-seq experiment in K562 cells that used CRISPR interference (CRISPRi) to target 9,426 genes. Raw scRNA-seq reads were processed with Cell Ranger [39] (v1.1.0) for alignment and expression quantification. For perturbation assignment, sgRNAs were mapped to individual cells using SCREE [40] (v1.1.0). An sgRNA was assigned to a cell if it had > 20 reads and represented > 80% of total sgRNA reads in that cell. SCREE produced a three-column table (cell, sgRNA, target gene). The cells without an assigned sgRNA were labeled "blank," and cells with multiple sgRNAs were labeled "multiple." Low-quality cells were removed if they had < 1,000 UMIs, < 200 detected genes, or > 10% mitochondrial gene content.

After preprocessing, scMAGeCK [41] (v1.9.1) was used to quantify transcriptional effects of each perturbation. Using the assigned sgRNA–cell matrix, scMAGeCK modeled expression differences between perturbed and control cells to estimate a regression coefficient (β score) reflecting effect direction and magnitude, and computed p-values for each perturbation–gene association. Target genes with $p < 0.05$ were retained for downstream analyses.

We generated single-cell CRISPR perturbation data in PC9 cells targeting four metabolic regulators using the 10x Genomics platform. The in-house dataset is available from the NIH Gene Expression Omnibus (GEO) under accession GSE236708 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE236708). Data processing followed the same pipeline used for the K562 single-cell perturbation dataset, including read alignment and quantification with Cell Ranger, sgRNA–cell assignment with SCREE, and standard quality-control filtering.

Single-cell transcriptomic data

To apply MetroSCREEN to single-cell transcriptomic data, we obtained cancer-associated scRNA-seq datasets from the Tumor Immune Single-cell Hub (TISCH) [42] (http://tisch.comp-genomics.org/). Variations in throughput, gene coverage, and dataset availability across different scRNA-seq platforms can result in technological noise when integrating datasets. Therefore, our study focuses on datasets produced by 10X Genomics, which offer the largest number of available datasets. Overall, we collected over 4 million cells from 735 patients representing 36 different cancer types. The expression matrices, raw counts or transcripts per million (TPM), were downloaded. To ensure accurate gene symbol mapping across different genome assemblies, we standardized all genes to the GRCh38.p13 assembly.

ICB cohort

To assess the association of metabolic states and regulators with ICB outcomes, we curated melanoma tumor RNA-seq cohorts from patients treated with anti–PD-1 monotherapy or combination anti–PD-1/anti–CTLA-4 therapy [43–47]. RNA-seq data and overall survival (OS) information were obtained from those cohorts. To minimize the impact of bias and randomness, cohorts with fewer than 10 samples were excluded. This criteria resulted in the inclusion of samples from 296 patients; 182 were treated with anti–PD-1 alone, and 114 were treated with both anti–PD-1 and anti-CTLA4. To mitigate the influence of variations among batches, we employed quantile normalization on the gene expression data from all samples in the cohorts, ensuring a standardized distribution [48].

The cancer genome atlas cohort

To assess the association between metabolic states, their regulators, and clinical survival, we obtained tumor

clinical metadata and TPM-normalized expression profiles from The Cancer Genome Atlas (TCGA) data portal (https://www.cancer.gov/tcga). This cohort comprises 9,550 patients spanning 32 cancer types.

In-house single cell perturbation data

Cell culture

HEK293T cells and PC-9 cells were cultured in DMEM medium (Gibco, #11960044) supplemented with 10% FBS (Gibco, #A5669701), 1% penicillin/streptomycin (Gibco, #15070063), and 1% L-Glutamin (Gibco, #25030081). The cells were maintained in a humidified incubator at 37 °C with 5% CO_2.

Generation of Cas9-expressing cell line

PC-9 cells were transduced with lentiviral vector lenti-Cas9-Blast (Addgene, #52962) [49]. The following lentiviral packaging and transduction were used to produce the Cas9-expressing cell line. 48 h post-transduction, the culture medium was changed to one supplemented with 1.4–2.0 μg/ml blasticidin (Gibco, #A1113903) to maintain selection pressure for Cas9. The cells were cultured in the selection medium for approximately 5 days until the selected cells became stably resistant to blasticidin. Subsequently, the cells were subcloned at low seeding density to obtain single-cell-derived monoclonal lines. The expression of Cas9 in the monoclonal lines was verified by Western blotting.

sgRNA cloning and Perturb-seq library preparation

Four to six sgRNA sequences for each target gene, along with four non-targeting sgRNAs, were selected from CRISPick [50, 51] and the A549 demo data released by 10x Genomics Inc (Additional file 1: Table S1). The insert oligos were ordered as separate single-stranded DNA (ssDNA) with overhangs homologous to (Addgene, #122237), which had been digested with BstXI (NEB, #R0113V) and BlpI (Thermo Scientific, #FD0094) [52]. Library vectors were prepared using arrayed cloning and verified by Sanger sequencing of the protospacer regions. The plasmids were combined in equal molar ratios to create a screening library.

Lentiviral packaging and transduction

Lentivirus was produced by co-transfecting HEK293T cells with the plasmid/library and lentiviral packaging plasmids psPAX2 (Addgene, #12260) and pMD2.G (Addgene, #12259), as previously described [53]. Transfection was performed using the X-tremeGENE HP DNA Transfection Reagent (Roche, #6366244001) according to the manufacturer's protocol. Supernatants were collected 48 and 72 h post-transfection.

PC-9 cells stably expressing Cas9 were transduced with the lentiviral library at a multiplicity of infection (MOI) of 0.3–0.5, supplemented with 8 μg/ml of polybrene and blasticidin. 48 h post-transduction, selection was carried out using 0.6 μg/ml of puromycin (Gibco, #A1113803) and 1.4–2.0 μg/ml of blasticidin. The selective medium was renewed every 2–3 days for a total duration of 10 days to ensure efficient editing. Libraries were prepared following the Chromium Single Cell 3' Reagent Kits User Guide (v3.1 Chemistry Dual Index) and the Feature Barcoding Technology for CRISPR Screening User Guide (CG000316), and sequenced on an Illumina NovaSeq 6000.

Multiplex immunohistochemistry

Human tissue specimens were provided by the Shanghai Pulmonary Hospital (for the NSCLC sections) and the West China Hospital of Stomatology at Sichuan University (for the HNSC sections) under the approval of local medical ethics (Additional file 2: Table S2). Tissues were fixed in formalin, embedded in paraffin, sectioned into 4 μm, and positioned onto adhesion microscope slides following routine methods. Automated staining was performed with the Leica BOND-MAX autostainer (Leica, Microsystem). The slices were deparaffinized and pretreated with the Epitope Retrieval Solution 2 (AR9640-CN, LeicaBiosystems) at 100 °C for 20 min. Peroxidase blocking was performed for 10 min using the Bond Polymer (DS9800-CN, Leica) subsequent to rinsing in TBST buffer. Slides were incubated with a primary antibody at room temperature (RT) for 1 h. Primary antibodies included: rabbit CTHRC1 (1:400, ab85739; Abcam), rabbit smooth muscle Myosin heavy chain 11 (MYH11) (1:1000, ab133567; Abcam), mouse PGK1 (1:25, MA5-37712; ThermoFisher), rabbit NDUFS2 (1:200, ab192022; Abcam), rabbit SLPI (1:500, PA5-82990; ThermoFisher), rabbit C1QC (1:100, ab75756; Abcam), mouse PFKP (1:150, MA5-25792; ThermoFisher), rabbit PLD3 (1:100, PA5-104016; ThermoFisher), rabbit CD68 (1:400, BX50031; Biolynx), and rabbit cytokeratin pan (panCK) (1:300, BX50143; Biolynx). All slides were stained with secondary reagents at RT for 10 min and tyramide signal amplification reagents at RT for 10 min. Nuclei were stained with DAPI (Sigma-Aldrich) after all the human antigens had been labeled. Stained slides were scanned at 20× magnification using the Pannoramic MIDI scanner (3DHISTECH Ltd.) and images were analyzed by Halo software (Indica Labs).

MetroSCREEN framework overview

MetroSCREEN is a reference-guided computational framework that infers reaction-level metabolic states and candidate upstream regulators from bulk and single-cell transcriptomic data. MetroSCREEN accepts a single-cell or bulk RNA-seq gene expression matrix as input and produces two output matrices: a MetaModule score

matrix and a MetaRegulon score matrix. A MetaModule is defined as a set of enzyme-coding genes that catalyze a single metabolic reaction. Its score reflects the transcriptional potential of each cell or sample to utilize that reaction. Under non-limiting substrate and cofactor conditions, it can be regarded as an approximate upper limit of the reaction's feasible flux, rather than providing a direct measurement of flux or metabolite abundance. A MetaRegulon is defined by the putative regulons of a MetaModule, which include both cell-intrinsic regulators and extrinsic interactions. The MetaRegulon scores reflect the impact of these regulons on the MetaModule, along with causal directions for each regulon (1 for regulators and 0 for effectors).

The workflow comprises two stages (Additional file 3: Fig. S1). First, MetroSCREEN computes reaction-level MetaModule scores for each cell or sample. Based on the profiles, cells or samples are clustered to define metabolic subtypes. Second, MetaRegulon scoring is performed within each metabolic subtype. For MetaModules upregulated in a given subtype, MetroSCREEN reports MetaRegulon scores and its causal direction for a MetaModule. This prediction proceeds in three steps: (i) inferring the regulons activities based on transcription factor-derived, cellular interaction-derived, and signaling-derived regulons; (ii) screening candidate regulons for each MetaModule using an integration of Pareto-efficiency [54] and Robust Rank Aggregation [55] (RRA); and (iii) inferring candidate direct regulators and downstream effectors with the fast Peter–Clark (PC) causal algorithm [56].

MetaModule scoring and metabolic subtypes identification

In metabolically active cells, expressed metabolic enzymes catalyze reactions that produce numerous metabolites. We sourced 8,033 MetaModules, comprising approximately 3,700 genes and 8,378 metabolites from the Human1 database [57]. These reactions occur in nine distinct compartments: extracellular space, peroxisome, mitochondria, cytosol, lysosome, endoplasmic reticulum, Golgi apparatus, nucleus, and inner mitochondrial membrane.

A metabolic reaction, which involves enzymes converting substrate metabolites into product metabolites, serves as a fundamental unit in metabolic networks. Given the strong correlation between enzyme gene expression levels and associated metabolites, transcriptome data can be used as an alternative method to assess metabolic states [30, 58]. MetroSCREEN infers MetaModule scores for each cell (or sample) using ssGSEA [31] on the gene expression matrix, similar to the approach used in scMetabolism [32]. The difference is that MetroSCREEN calculates the metabolic scores based on individual reactions rather than the entire pathway, enabling finer-resolution identification of specific metabolic subtypes [59, 60]. MetroSCREEN accounts for the gene-protein-reaction (GPR) rules when calculating the MetaModule scores. It first rewrites each reaction's GPR rule into disjunctive normal form (DNF): OR between branches encodes interchangeable isoenzymes/isoforms, whereas AND within a branch denotes the complete set of subunits required to assemble a single enzyme complex. The MetaModule score is calculated as follows:

$$S_{r,b}(i) = ss\mathrm{GSEA}(G_{r,b} \mid x_{i,g}) \tag{1}$$

$$S_{i,r} = \max_{b \in B_r} S_{r,b}(i) \tag{2}$$

$$S = (S_{i,r})_{r \in R,\, i=1,\ldots,N} \in R^{|R| \times N} \tag{3}$$

Here, r denotes a reaction, $r \in R$; $x_{i,g}$ is the expression of gene g in cell or sample i; $G_{r,b}$ is the AND-set of required subunits for a branch b of reaction r; $S_{r,b}$ is the activity score of branch b; and $\mathcal{B}_r$ is the set of OR-branches for reaction r. The reaction-level MetaModule score $S_{i,r}$ is obtained by taking the maximum across branches. Given a gene-by-cell (or sample) expression matrix as input, MetroSCREEN outputs a MetaModule-by-cell (or sample) score matrix S.

This aggregation strategy avoids the dilution that occurs when treating the union of all reaction genes as a single set and, through ssGSEA normalization, reduces bias from heterogeneous set sizes. Using the MetaModule score matrix S, we applied PCA for dimensionality reduction, used the leading principal components to build a k-nearest neighbors (KNN) graph, and ran the Louvain algorithm [61] to identify clusters as metabolic subtypes. We then performed one-vs-rest differential MetaModule testing with the Wilcoxon rank-sum test. Finally, we annotated metabolic subtypes by the most upregulated MetaModules within each cluster, reflecting shared metabolic states.

MetaRegulon scoring

Metabolic reprogramming results from converging signals originating from intrinsic and extrinsic factors. Intrinsic signaling maintains the baseline metabolic subtypes, while extrinsic signals fine-tune the metabolic processes based on the availability of metabolites and the requirements of the cells [9]. MetroSCREEN calculates MetaRegulons score through a combination of intrinsic and extrinsic cellular signals.

Step 1: Inferring the activity of the regulons

Cell-intrinsic Transcription Regulators (TRs) activity MetroSCREEN utilizes public ChIP-seq data with the Regulatory Potential (RP) model to predict the TRs

activity for each query cell or sample, similar to the chromatin landscapes in the "Reference Dataset" of SCRIPro [62]. Subsequently, for each metabolic subtype, MetroSCREEN conducts in silico deletion (ISD) calculations by using the marker genes associated with each metabolic subtype using LISA [63], and then converts the results to negative logarithmic forms used in SCRIPro: $U = -\log(P_{summary})$. Furthermore, MetroSCREEN calculates a z-score relative to the mean for each TR and its corresponding targets within cells.

The chromatin regulatory potential (chrom-RP) of a ChIP-seq sample for a TR at a gene is defined as:

$$R_{i,j} = \sum_{k \in [tss_j - L,\, tss_j + L]} w_k C_{ik} \tag{4}$$

$$w_k = \frac{2}{1+\exp\left(\frac{2\mu|k - tss_j|}{2L}\right)} \tag{5}$$

Here, L is set to 100 kb around the gene transcription start site tss_j; C_{ik} is the ChIP-seq signal of cell or sample i at genomic position k; μ is the weight decay rate; The weight w_k represents the regulatory impact of a locus at position k on the gene j's tss; The ISD method deletes chromatin signals in all 1-kb windows containing at least one peak from a putative regulatory cistrome and recalculates the chromatin regulatory potential (chrom-RP). The RPs from models with and without deletions are compared to produce a ΔRP value for each gene. The one-sided Wilcoxon rank-sum test is then used to rank TRs by comparing the ΔRP of the query gene set to that of the background gene set. The Cauchy combination test is employed to compute a summary *P*-value for each TR.

MetroSCREEN calculates the TR activity as follows:

$$Q_{TR}(i,m) = \frac{x_{i,m} - \overline{x}_m}{sd(x_m)} \tag{6}$$

$$Q_{Target}(i,n) = \frac{1}{a_m} \sum_{n \in T_m} \frac{x_{i,n} - \overline{x}_n}{sd(x_n)} \tag{7}$$

$$Q_{\exp}(i,m) = Q_{TR}(i,m) + Q_{Target}(i,n) \tag{8}$$

$$\underset{\exp}{\hat{Q}}(i,m) = \frac{Q_{\exp}(i,m) - \min\limits_{i'} Q_{\exp}(i',m)}{\max\limits_{i'} Q_{\exp}(i',m) - \min\limits_{i'} Q_{\exp}(i',m)} \tag{9}$$

$$S_{i,m} = U_m \underset{\exp}{\hat{Q}}(i,m) \tag{10}$$

$$S_{TR} = (S_{i,m})_{m \in T,\, i=1,\ldots,N} \in \mathrm{R}^{|T| \times N} \tag{11}$$

Here, $x_{i,m}$ is the expression value of the TR m in cell or sample i ($x_{i,n}$ is the expression value of the TR targets n in cell or sample$i$); a_m represents the number of targets of the TRm; For each ChIP-seq dataset, genes with an RP score > 5 were defined as the targets of the TR. If the number of genes is more than 300, MetroSCREEN considers the top 300 genes in the RP ranking as potential targets. Then $Q_{TR}(i,m)$ and $Q_{Target}(i,n)$ are combined to obtain$Q_{\exp}(i,m)$, which represents the expression level aggregation of the TRs and their targets, which further been min–max normalized to $\underset{\exp}{\hat{Q}}(i,n)$. The TR m activity score $S_{i,m}$ is equal to U_m(the negative logarithmic transformation of the summary P-values of TR m in the metabolic subtype to which the cell or sample i belongs) multiplied by the min–max normalized $\underset{\exp}{\hat{Q}}(i,m)$. This approach ensures that both the expression of the TR and its target genes are included when calculating the final TR activity score. Given a gene-by-cell (or sample) expression matrix as input, MetroSCREEN calculates the TRs-by-cells (or samples) activity matrixS_{TR}.

Cell-extrinsic drivers activity (LRs) MetroSCREEN leverages NicheNet [64] to evaluate the activity of cell-extrinsic drivers by integrating prior knowledge of ligand-receptor interaction networks with gene expression profiles. High ligand activity in a cell indicates a greater expression of the corresponding target genes compared to other cells. NicheNet identifies potentially active ligands based on the expressed receptors in the metabolic subtype and infers ligand activity for each cell using receptor expression and ligand-receptor data. Given a gene-by-cell (or sample) expression matrix as input, MetroSCREEN calculates a ligands-by-cells (or samples) activity matrix S_{LR}. It's important to note that in the CCLE data, the calculated ligand activity signifies the secreted ligands by the cells themselves.

Genome-wide activity (GGs) To systematically capture regulatory influences not explicitly modeled by curated TRs or ligands, MetroSCREEN incorporates a data-driven, genome-wide channel. For intrinsic regulation, this channel accounts for residual effects—such as unannotated TRs/cofactors, chromatin and metabolic feedback, and broader pathway crosstalk—that are not covered by the curated TR set. For extrinsic regulation, it compensates for ligand-mediated signals, particularly autocrine (self-signaling), that are not represented in the curated ligand set. For curated TRs, MetroSCREEN prioritizes candidates under the hypothesis that more highly expressed TRs are more likely to be functionally active [64, 65]. Accordingly, MetroSCREEN ranks genes by their average expression within each metabolic subtype and retains the top 3,000 most highly expressed genes for downstream analyses. Raw Gene expression values are library-size normalized, log1p-transformed, and then

each gene expression was standardized across cells or samples, we defined the standardized gene expression as the genome-wide activity. Given a gene-by-cell (or sample) expression matrix as input, MetroSCREEN calculates a gene-by-cell (or sample) activity matrix S_{GG}.

Step 2: MetaRegulons screening

To screen the regulons for MetaModules from the three aforementioned regulons sources, MetroSCREEN utilizes a two-pronged, multi-objective recommendation scheme that combines Pareto efficiency [54] with RRA [55].

For each MetaModule M and candidate regulon f, MetroSCREEN computes an association score by correlating the MetaModule score S with the regulon activities inferred from three source-S_{TR}, S_{LR} and S_{GG}.

For source $k \in \{TR,LR,GG\}$, the score is calculated as follows:

$$H_{f,M}^{k} = \rho\left(S_{\cdot M}, S_{\cdot f}^{k}\right) \tag{12}$$

Here, ρ is Spearman's rank correlation; Correlations are computed separately within each metabolic subtype (using only cells/samples assigned to that subtype). A larger absolute value of $H_{f,M}^{k}$ indicates a stronger association between the source-specific regulon activity $S_{\cdot f}^{k}$ and MetaModule *M*. After this step, for each MetaModule *M*, we obtain a MetaRegulon score matrix whose rows index candidate regulon *f* (MetaRegulons) and whose columns index the three evidence sources (TR, LR, GG). Each entry $H_{f,M}^{k}$ stores the correlation score defined above. Not all regulons have evidence from all sources—transcription factor-derived regulons and cellular interaction-derived regulons are largely disjoint—but this does not affect downstream analyses: if regulon *f* lacks evidence for source *k*, the corresponding score $H_{f,M}^{k}$ is set to zero.

Pareto optimality

To rank the MetaRegulons by importance, quantified by the $H_{f,M}^{k}$ across the three evidence sources $k \in \{TR,LR,GG\}$, MetroSCREEN adopts the standard dominance relation from multi-objective optimization. Let $h_k(f) = H_{f,M}^{k}$, for two candidates y and z, y Pareto-dominates z iff $h_k(y) \geq h_k(z)$ for all k, and $h_l(y) > h_l(z)$ for at least one $l \in \{TR, LR, GG\}$. The set of candidates that are not dominated by any other candidate forms the first Pareto front F_1. After removing F_1, MetroSCREEN repeatedly extract the non-dominated set from the remaining candidates to obtain F_2, F_3,

We define the Pareto score as followes:

$$S_{Pareto}(f,M) = L(f,M) = \min\{\ell \in \mathbb{N} : f \in F_\ell\} \tag{13}$$

Here, ℓ is the integer label of a front. $L(f,M)$ is the unique index ℓ such that $f \in F_\ell$. Because the fronts are pairwise disjoint and cover the entire candidate set, the set $\{\ell : f \in F_\ell\}$ contains exactly one element; the $\min\{\cdot\}$ operator therefore returns that index. By convention, a smaller $L(f,M)$ indicates a better trade-off across objectives. Candidates with the same L are mutually non-dominating at this step; ties are resolved downstream by RRA.

RRA optimality

Let $r_k(f,M)$ denote the rank of f in source $k \in \{TR, LR, GG\}$ (with rank 1 being best), and let $N_k(M)$ denotes the number of candidates in that source. MetroSCREEN converts raw ranks to fractional ranks $u_k(f,M)\,\epsilon\,(0,1)$:

$$u_k(f,M) = \frac{r_k(f,M)}{N_k(M)} \tag{14}$$

Under the RRA null model, the positions of f across sources are independent and uniformly distributed on (0,1]. If f appears in k sources ($k \leq 3$ here), MetroSCREEN sorts the observed fractions to obtain $u_1 \leq \cdots \leq u_k$. For each position *j*, compute the one-sided probability under the null of random ranks that the $j-\text{th}$ order statistic is at most the observed $u_{(j)}$, and take the minimum over $j = 1 \ldots k$. Then the RRA score is calculated as follows:

$$S_{RRA}(f,M) = \min_{1 \leq j \leq k} Pr\left(U_{(j)} \leq u_{(j)}\right) \tag{15}$$

Smaller values indicate stronger agreement that f ranks near the top in multiple sources.

Combining Pareto and RRA

For each MetaRegulon, MetroSCREEN combines the Pareto front index and RRA to produce the overall priority for (f, M):

$$S_{MR}(f,M) = S_{Pareto}(f,M) \times S_{RRA}(f,M) \tag{16}$$

Smaller values indicate stronger evidence that a regulon (MetaRegulon) f has the potential to regulate the corresponding MetaModule M.

Step 3: Examining the causality between the MetaModule and MetaRegulon

To identify direct regulators and effectors of a MetaModule, MetroSCREEN jointly models the pre-screened MetaRegulons (default: top 50) together with all genes in the MetaModule, and learns a causal graph using the fast PC (Peter–Clark) algorithm [56]. On the MetaRegulon side, MetaRegulons supported solely by

cell-extrinsic drivers represent regulators rather than intracellular effectors. Because their effects propagate through ligand–receptor–signaling cascades, MetroSCREEN treats them as upstream annotations and does not include them as explicit nodes in the receiver-cell causal graph. MetroSCREEN therefore, includes only MetaRegulons with cell-intrinsic source (transcription factor-derived and/or signaling-derived regulons). When both TR-based and gene-derived regulons are available for the same MetaRegulon, MetroSCREEN prioritizes TR activity; otherwise, gene activity is used. On the MetaModule side, MetroSCREEN includes the same set of genes used to compute the MetaModule score. During the PC stage, MetroSCREEN models the MetaModule at gene resolution rather than using a single MetaModule score. Because MetaModule score is computed according to GPR rules, regulation of any constituent gene within a MetaModule can change the module's score. Aggregating into one score conflates heterogeneous gene-level effects and risks information loss. MetroSCREEN therefore use the MetaModule score mainly for upstream screening of MetaRegulons to reduce dimensionality and improve computational efficiency. The resulting node set thus comprises the screened MetaRegulons and the MetaModule genes, from which PC first prunes the skeleton via conditional-independence tests and then orients edges to yield candidate regulators(1) and effectors(0) for MetaModules.

Conditional independence decisions within PC are provided by DCC.gamma [66, 67], enabling efficient detection of nonlinear dependencies. For a MetaRegulon- gene pair (X, Y) and a conditioning set S, MetroSCREEN test:

$$H_0 : X \perp\!\!\!\perp Y \mid S \tag{17}$$

Practically, X and Y are first nonparametrically regressed on S to obtain residuals. A distance-covariance-based statistic T is computed and approximated by a Gamma distribution:

$$T \sim Gamma\,(\alpha, \beta)\,, p = 1 - F_{Gamma(T_{obs};\alpha,\beta)} \tag{18}$$

Here, when the $p > \alpha_{test}$, the X and Y are independent. By using the stable PC procedure, MetroSCREEN removes the undirected edge of the MetaRegulon-gene and learns the graph skeleton.

On the learned skeleton, MetroSCREEN applies the standard PC orientation rules, including collider identification ($X \rightarrow Z \leftarrow Y$) and Meek propagation [56], while enforcing acyclicity. It adopts a relaxed conflict-resolution strategy and produces a directed acyclic graph (DAG). Finally, edges directed from MetaRegulons to MetaModule genes are interpreted as candidate direct regulatory relationships and edges in the opposite direction are treated as downstream effector signals. To ensure efficiency, if a metabolic subtype contains more than 300 cells, MetroSCREEN randomly subsamples 300 cells for analysis. For stability, it repeats the DCC-based tests five times by default.

Optimization of hyperparameters

MetroSCRREN is designed to be hyperparameter-free: thresholds and weights are determined by data-adaptive, rank-based rules, which are scale- and order-invariant. This approach reduces tuning bias and improves reproducibility.

MetaModule scores were computed with ssGSEA using its default normalization, and cells or samples were clustered on the MetaModule scores with hierarchical clustering under Euclidean distance. The number of clusters was chosen per dataset by inspecting silhouette and NMI trends. Marker detection used a one-vs-rest Wilcoxon test with BH-FDR of 0.05.

For regulon activity channels, the TR channel followed standard chrom-RP and ISD settings. We used a ±100 kb window around TSS, 1 kb deletion windows for ISD, and an RP threshold greater than 5. When fewer than 300 targets were obtained, we retained the top 300 genes ranked by RP. Ligand activity via NicheNet used a subtype-specific receptor expression cutoff at the 0.10 quantile and a candidate ligand pool of 50 per subtype. The genome-wide channel ranked genes by average expression within each subtype, retained the top 3,000, applied library-size normalization and log1p transformation, and then z-scored expression within each subtype.

Causal graphs were learned with the PC algorithm using DCC.gamma for conditional-independence tests at α_{test}=5 × 10^{-3} and a maximum conditioning set size of 3. For subtypes with more than 300 cells, we randomly sampled 300 cells and repeated the tests five times, then retained edges with frequency at least 0.6 to ensure stability. These defaults are well supported by prior studies and produced stable results in our analyses. We compared several strategy choices that could influence the results (Additional file 3: Fig. S2a,b).

MetaModule score

MetroSCREEN uses GPR rules by converting each reaction into AND/OR branches. It computes ssGSEA per AND branch (requiring all subunits) and aggregates across OR branches using the maximum. We also evaluated alternatives-aggregating OR branches by mean or sum, and a direct approach that ignores GPR and applies ssGSEA to all reaction genes. The GPR-aware and max-aggregation strategy achieved the best performance across cancer types (Additional file 3: Fig. S2a).

MetaRegulon score

By default, MetroSCREEN reports the top-50 regulators as high-confidence candidates. We compared top-100/150/200 cutoffs and observed slightly higher recall but consistently lower accuracy, precision, and F1-score relative to top-50 (Additional file 3: Fig. S2b). Therefore, we retain top-50 as a balanced and conservative choice.

Performance evaluation in CCLE data

MetaModule score benchmark

For each cell line i, MetroSCREEN computed per-reaction MetaModule scores by enforcing GPR rules in disjunctive normal form (AND for enzyme subunits; OR for isoenzymes). As our MetaModule scores reflect the metabolic flux of a specific reaction, we aggregated MetaModule scores using the production minus consumption propensity. Specifically, for each CCLE metabolite m, we identified the reactions that produce m (denoted $\mathrm{Prod}\,(m)$) and those that consume m (denoted $\mathrm{Cons}\,(m)$), and defined metabolite propensity P_m as follows:

$$P_m\,(i) = \frac{1}{|\mathrm{Prod}(m)|} \sum_{r \in \mathrm{Prod}(m)} A_{r,i} - \frac{1}{|\mathrm{Cons}(m)|} \sum_{r \in \mathrm{Cons}(m)} A_{r,i} \quad (19)$$

Currency metabolites (e.g., ATP/ADP, NAD(H)/NADP(H), CoA, H_2O, H^+) were excluded to reduce degree-driven bias. For AUCell [32], Mean, scFEA [28], and MetaFlux [30], we applied the same metabolite-propensity calculation strategy once reaction-level scores were obtained.

Within each tumor type, we computed the Pearson correlation between measured metabolite abundances across multiple cell lines and the predicted metabolite propensity for that tumor type. This quantifies the concordance between predictions and measurements. It also limits cross-tissue confounding by keeping the context fixed and tests robustness across cell lines.

Importantly, we do not claim that transcriptomes quantitatively predict metabolite concentrations. Metabolite abundance reflects not only enzyme-mediated conversion but also medium composition, uptake/secretion and transport, protein abundance and post-translational/allosteric regulation, compartmentalization, and growth-dilution. MetroSCREEN instead estimates propensities for reaction utilization. Using metabolomics as the benchmark is therefore a stringent test, and results are interpreted as associations rather than absolute quantification. Furthermore, because our ultimate goal is to pinpoint which specific metabolic processes (reactions) are reprogrammed and to identify actionable regulators, reaction-level MetaModule offers higher mechanistic resolution than metabolite readouts. They localize signals to enzyme complexes and isoenzymes under GPR logic, rather than to composite metabolite endpoints influenced by many pathways.

To uncover the metabolic subtypes within the CCLE, we utilized PCA for dimensionality reduction, incorporating the MetaModule scores calculated by MetroSCREEN. Subsequently, we applied the Louvain clustering method using the top 10 principal components to categorize the cell lines into distinct clusters at default resolutions. Further reduction of dimensions and visualization of the clustering outcomes were achieved through Uniform Manifold Approximation and Projection(UMAP).

MetaRegulon score benchmark

For the MetaRegulon benchmark, we defined a correlation-derived ground truth using CCLE metabolite abundance and transcriptome data. For each metabolite, we computed Pearson correlations between its abundance and TR activity (S_{TR}, Eq. 11) and LR activity (S_{LR}) across cell lines, and labeled MetaRegulons with $|r| > 0.3$ and FDR < 0.01 as positives (label as 1) and others as negatives (label as 0). To evaluate MetaRegulons scores, we used a two-step procedure. First, we restricted analysis to MetaModules that produce the metabolite and selected those MetaModules significantly associated with this metabolite. Second, if exactly one MetaModule remained, we evaluated MetroSCREEN ($S_{MR}\,(f, M)$, Eq. 16) (multi-source) and single-source baselines (SCRIPro [62], SCENIC [68] and SparseNCA [69] for TRs ($H^k_{f,M}, k = TR$), NicheNet [64] for LRs($H^k_{f,M}, k = LR$), Eq. 12) against the reference using Area Under the Curve (AUC), precision, recall, F1-score, and accuracy. If multiple MetaModules remained, we aggregated MetaRegulon scores ($S_{MR}\,(f, M)$, Eq. 16) across those MetaModules using RRA to obtain metabolite-level MetaRegulon scores before evaluation. Because SCRIPro, SCENIC, SparceNCA and NicheNet are strong single-source methods in the specific domain, our goal is to demonstrate the advantage of multi-source integration (MetroSCREEN) and to provide a light-coding, comprehensive prioritization for biologists. Note that this benchmark targets associational prediction ($S_{MR}(f, M)$, Eq. 16). MetroSCREEN can further nominate regulator–effector pairs, which cannot be verified in CCLE dataset, and we validated key factors experimentally using public and in-house single-cell perturbation datasets. AUC was computed from the global ranking of all candidate factors, whereas precision, recall, F1, and accuracy were computed on the Top-K subset of factors (K = 50 unless otherwise specified). Based on our hyperparameter tests and wet-lab validations, the Top-50 set provides biologists with more accurate and actionable candidates for downstream experimental verification.

Single-cell transcriptomic data preprocessing

Data processing and Metacell identification

Each dataset were processed using MAESTRO [70] (v1.1.0), which involved quality control, doublets and batch effect removal, and cell clustering. Only cells with a UMI count greater than 1,000 and more than 500 expressed genes were retained for downstream analysis. A total of 4,483,367 cells from 735 patients, encompassing 36 cancer types and covering 103 studies, were retained in the pan-cancer landscape, previously constructed by our previous work TabulaTIME [71]. This dataset includes 236,510 fibroblasts and 504,608 myeloid cells.

For the substantial number of cells collected from diverse sources, TabulaTIME [71] integrated them through a three-step process. First, within each dataset, we standardized the raw counts to TPM to ensure consistent expression level representation across different datasets. Specifically, the expression level $E_{i,j}$ of gene i in cell j was quantified as $log2(\frac{TPM_{i,j}}{10}+1)$. Second, to reduce technical noise and computational demands, single cells were grouped into small clusters (referred to as Metacells) based on their similarity in the KNN graph. This strategy assesses the similarity within each cluster of a specific sample from which the cells originated. The optimal number of cells in each Metacell was determined to be 27–30 based on their performance in terms of gene coverage, variation within the Metacell, and the local inverse Simpson's Index (LISI) score during integration in TabulaTIME. The average log TPM-transformed gene expression of all cells within each Metacell was utilized to represent the Metacell's expression, and the original gene-by-cell expression matrix was converted to the gene by the Metacell expression matrix. Third, canonical correlation analysis (CCA) was employed to reduce data dimensionality and align data batches by capturing the most correlated data features. Subsequently, UMAP was utilized to further reduce dimensionality and visualize the annotation results.

Dimensionality reduction and annotation

Cell type annotation For each lineage, we addressed the challenge of dealing with high-dimensional data by applying PCA to the Metacells. Specifically, we focused on the top 3,000 highly variable genes to capture the most informative features. To ensure consistency with other analyses, we selected the top 30 principal components, aligning with the counts used for KNN and the Louvain algorithm for cluster identification and visualization. To determine the optimal clustering resolution for each cell lineage, we utilized both Silhouette scores [72] and Clustree [73] metrics. This comprehensive approach allowed us to identify clusters that faithfully represent the biological diversity present in the data while avoiding unnecessary fragmentation. With the optimal clustering resolution determined, we proceeded to annotate the clusters based on marker gene expression, providing further insights into their biological identity.

Metabolic subtype annotation To uncover the metabolic diversity of each Metacell within fibroblasts and myeloid lineages, we honed in on the MetaModule scores and leveraged PCA for dimensionality reduction. Mirroring our cell type identification process, we employed the top 30 principal components for KNN analysis and the Louvain algorithm for cluster identification and visualization. In assessing the optimal clustering resolution for each cell lineage, we integrated both Silhouette scores and Clustree metrics. MetroTIME scrutinized the cluster-enriched MetaModules related to metabolism and characterized the metabolic cluster based on the most significantly enriched MetaModules.

Finally, to assess the purity of cell types and metabolic subtypes, we employed ROGUE [74], an entropy-based universal metric. ROGUE quantifies the purity of individual cell populations on a scale of 0 to 1, where a score of 1 represents a completely pure subtype, while 0 indicates a highly heterogeneous population. Our analysis revealed that all annotated cell subtypes and metabolic subtypes with a median ROGUE value greater than 0.9 demonstrate a remarkable level of homogeneity, underscoring the coherence within these defined subgroups.

Association between metabolic subtypes and function

To characterize cellular functions, we gathered gene sets linked to cell function, cancer hallmarks, and KEGG pathways from past studies and the Molecular Signatures Database (MSigDB [75] v6.1.13). Subsequently, we applied the ssGSEA method to calculate the functional signature scores for individual Metacells within the fibroblasts and myeloid lineages (Additional file 4: Table S3). Finally, we evaluated the relationship between function and metabolic subtypes by computing the Pearson correlation between the scores of Metacells enriched in metabolic subtypes and the functional scores.

Trajectory building

To analyze the relationships among cell subtypes within fibroblasts and myeloid lineage, we employed STREAM [76] and Monocle [77] to construct developmental trajectories based on highly variable genes. The root of the fibroblast lineage is defined as qFibro_SAA1 reflecting its quiescent state. Similarly, the root of the macrophage lineage is identified as Mono_FCN1, reflecting the common understanding that monocytes often serve as a primary origin for macrophages. When monocytes are stimulated by signals of inflammation or tissue damage in the blood,

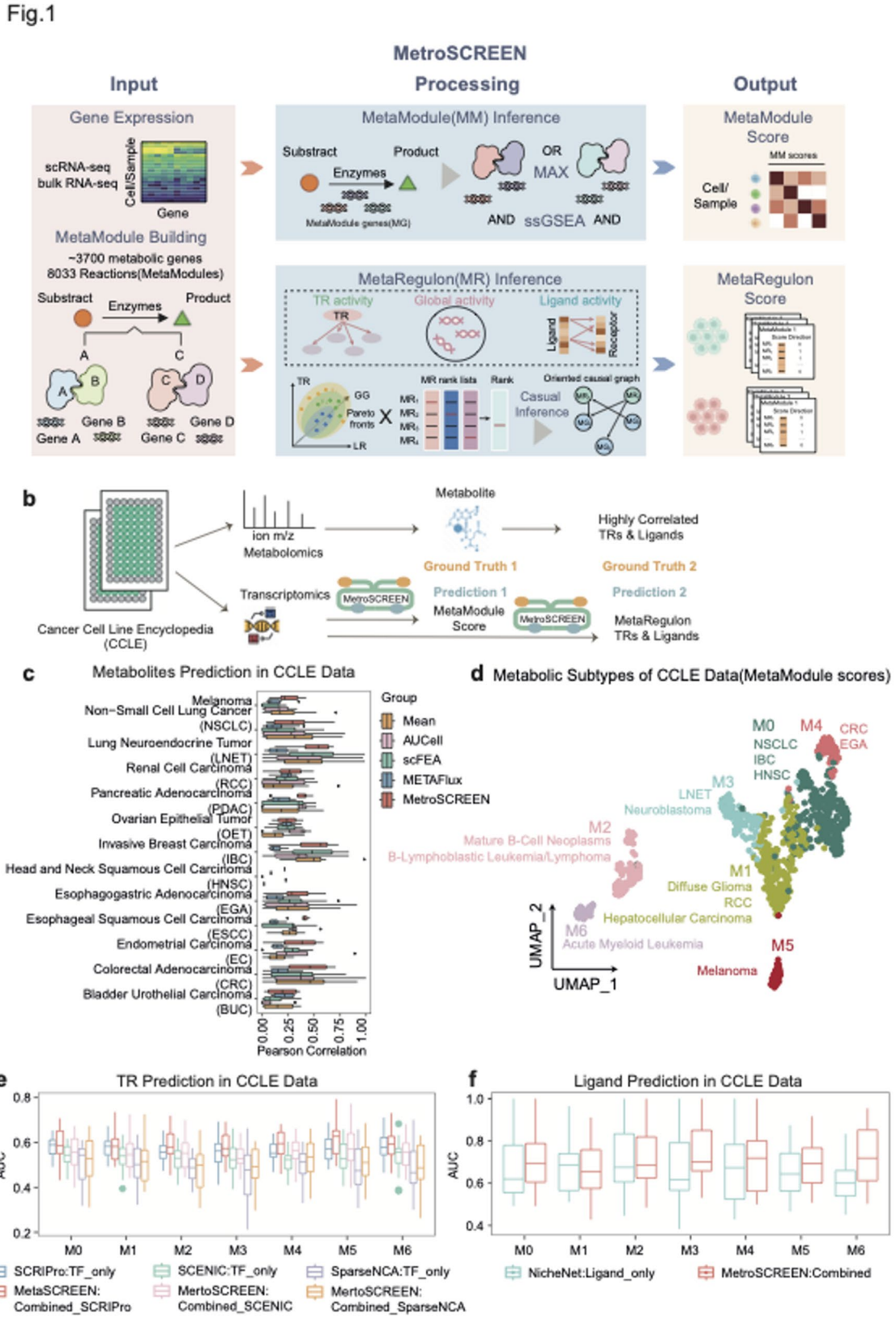

Fig. 1 (See legend on next page.)

(See figure on previous page.)

Fig. 1 MetroSCREEN accurately predicts metabolic states and regulators. **a** Schematic of the MetroSCREEN framework. **b** Schematic of the MetroSCREEN benchmarking strategy with the Cancer Cell Line Encyclopedia (CCLE) dataset. **c** Metabolites prediction in the CCLE dataset. Pearson correlations were computed between predicted metabolite propensities (derived from MetroSCREEN MetaModule scores or reaction-level results from other tools) and measured metabolites abundance across major cancer types. **d** UMAP of MetaModule scores in the CCLE dataset, identifying seven metabolic subtypes within 928 cancer cell lines. **e** Box plots of predictive performance for cell-intrinsic TR predictions in the CCLE dataset. **f** Box plots of predictive performance for cell-extrinsic ligand predictions in the CCLE dataset

they migrate to tissues and differentiate into mature macrophages.

Survival analysis

To analyze the association of MetaMudoles scores and MetaRegulons scores with patient survival, the TCGA cohorts were used. For each cancer type, MetaSCREEN calculated the MetaModules scores for each patient. Univariable Cox proportional hazards regression was used to link MetaModules scores to OS. To better understand the range of variation and the magnitude of the impact of the coefficients, we took the logarithm base 10 of the hazard ratio. Consequently, a value greater than zero indicates an increased mortality risk in the high MetaModules scores group, while a value less than zero indicates a decreased mortality risk.

Kaplan–Meier survival curves were used to directly illustrate the impact of specific MetaModules and their MetaRegulons on OS time. For each MetaModule, the median MetaModule score was utilized to divide the patients into the high MetaModule score group and the low MetaModule score group in each cancer type. Similarly, the median value of gene expression for MetaRegulons was used to categorize samples into high and low MetaRegulon expression groups. All of these analyses were implemented using the R package survival [78] (v3.8–3).

Immunosuppressive signature score

To explore the clinical relevance of MetaModules and MetaRegulons, we analyzed the ICB cohort. Signatures of immune response were collected from the study conducted by Thorsson et al. [79]. First, cancer type-specific effects were regressed out from immune signatures using the R package remef [80] (v1.0.7). Then, independent principal components of these immune signatures were derived to control the confounding effect of cross-correlation between signatures and cancer type-specific effects. The PC1, which gave high weights to immunosuppressive signatures in an unbiased manner, was used to estimate the MetaModules associated with immunosuppressive signatures. MetaModules for each sample were calculated by MetroSCREEN, followed by the calculation of the Pearson correlation between the MetaModules score and PC1.

ICB response

We computed the MetaModules scores for each sample in ICB cohort using MetroSCREEN. Subsequently, we assessed the signature score associated with the response to ICB for each sample. This encompassed scores such as 'Exclusion,' 'MSI.Score,' 'TAM.M2,' 'TIDE,' 'Dysfunction,' 'CAF,' and 'CTL,' which were derived and calculated using the TIDE package [81]. Additionally, we incorporated and computed other gene features that have been reported to predict the response to ICB. To facilitate comparability and stabilize the variance across the scores, we transformed them into logarithmic scale and performed mean subtraction within the same row for each signature. This normalization step centers the data by eliminating the average expression level of each gene, allowing for effective comparison across distinct samples or conditions. Lastly, we calculated the AUC scores for each MetaModule, TIDE signature, and gene expression associated with the ICB response, leveraging the pROC package [82] (v1.18.0). Moreover, we determined the differential scores by subtracting the median score in the non-response group from the median score in the response group.

Results

Overview of the MetroSCREEN framework

To infer reaction-level metabolic states and their associated regulators from bulk and single-cell transcriptomic data, we developed MetroSCREEN, a reference-guided computational framework (Fig. 1a; Additional file 3: Fig. S1, and Methods). MetroSCREEN accepts a single-cell or bulk RNA-seq gene expression matrix as input and produces two output matrices: a MetaModule score matrix and a MetaRegulon score matrix.

A MetaModule is defined as a set of enzyme-coding genes that catalyze a single metabolic reaction. The score is derived from enzyme-coding gene expression via ssGSEA [31] and incorporates gene–protein–reaction (GPR) rules (Methods). Under non-limiting substrate and cofactor conditions, it can be regarded as an approximate upper limit of the reaction's feasible flux. Based on the clustering of MetaModule scores, a metabolic subtype is defined as a cluster of cells or samples that share similar MetaModules, indicating that they exhibit comparable metabolic states (Methods). Reaction-level MetaModule scores provide a higher-resolution delineation of metabolic subtypes.

A MetaRegulon is defined by the putative regulons of a MetaModule. MetroSCREEN first estimates the activities of regulons based on transcription factor-derived, cellular interaction-derived, and signaling-derived source (Methods). To screen the regulons from these heterogeneous sources and prioritize candidates, MetroSCREEN applies Pareto efficiency [54] and Robust Rank Aggregation (RRA) [55], resulting in a robust ranking framework that generates MetaRegulon scores. Next, to determine the causal directions between MetaRegulons and the MetaModule, MetroSCREEN adopts a constraint-based causal discovery strategy using the standard Peter-Clark (PC) rules [56], and output the causal direction (1 as regulators and 0 as effectors) in the MetaRegulon matrix. By predicting a MetaModule's upstream MetaRegulon, MetroSCREEN helps biologists prioritize precise perturbations to modulate metabolic reprogramming and uncover potential therapeutic targets.

MetroSCREEN accurately predicts metabolic states and regulators in CCLE data

To evaluate MetroSCREEN's performance in the inference of MetaModule and MetaRegulon scores, we benchmarked it on the Cancer Cell Line Encyclopedia (CCLE) dataset. This dataset includes metabolomics for 225 metabolites and genome-scale transcriptomics from 928 cell lines (Fig. 1b and Methods).

For the MetaModule scores benchmark, we treated measured metabolite abundance as an external reference (Methods). Compared with the FBA-based method METAFlux [30], the machine-learning method scFEA [28], and statistical baselines (mean expression and AUCell [32]), MetroSCREEN showed superior performance in predicting metabolic abundance in 7 of 13 major cancer types (Fig. 1c). Focusing on specific metabolites, MetroSCREEN also achieved higher correlations than the recent METAFlux (Additional file 3: Fig. S3a). We further assessed whether predicted MetaModule scores capture biologically meaningful structure. Cell lines clustered by MetaModule scores showed greater concordance with tissue or lineage of origin than clusters derived from raw metabolite abundance (NMI = 0.41 versus 0.13), indicating that predicted scores better reflect metabolic heterogeneity across tissues (Fig. 1d and Additional file 3: Fig. S3b-d).

For the MetaRegulon scores benchmark, we defined a correlation-derived ground truth using CCLE metabolite abundance and transcriptome data (Methods). MetroSCREEN is a multi-source integration framework that combines intrinsic gene-regulatory evidence (SCRIPro [62]), extrinsic cell–cell interactions (NicheNet [64]), and genome-wide regulatory cues via a multi-objective optimization procedure. For TR prioritization, in addition to SCRIPro, we evaluated a tree-based gradient-boosting framework (SCENIC [68]) and a network component analysis method (SparseNCA [69]). Substituting different TR-activity evidence sources confirmed that integrated MetroSCREEN consistently outperformed single-source approaches (SCRIPro, SCENIC, or SparseNCA alone) across all seven metabolic subtypes, as assessed by AUC, precision, recall, F1 score, and accuracy (Fig. 1e and Additional file 3: Fig. S3e). Similarly, the multi-source integration in MetroSCREEN outperformed the single-source NicheNet approach (Fig. 1f and Additional file 3: Fig. S3f). These results show that integrating multiple evidence streams yields more accurate MetaRegulon prioritization than relying on any single source.

Validation of predicted glycolysis regulators using single-cell CRISPR screens

To validate the impact of predicted MetaRegulons on their associated MetaModules, we generated in-house single-cell CRISPR knockout data targeting CTNNB1, ATF4, MYC, and JUN, key MetaRegulons for glycolysis, using the NSCLC cell line PC9 (Additional file 1: Table S1 and Methods). Additionally, we analyzed a published genome-wide single-cell CRISPR interference (CRISPRi) dataset from the K562 cell line [38] to explore a broader range of perturbed MetaRegulons (Fig. 2a and Methods). The MetaRegulons for each MetaModule were predicted by MetroSCREEN using PC9 and K562 data in CCLE dataset. For each MetaRegulon perturbation, we calculated the essential score (β-score) for every MetaModule using scMAGeCK [41](Methods). A negative β-score of the MetaModule indicates downregulation of the MetaModule following MetaRegulon perturbation, suggesting that the MetaRegulon positively regulates its targeted MetaModule, and vice versa.

In the genome-wide screen, the top 50 and top 100 MetaRegulons predicted by MetroSCREEN significantly positively regulated the corresponding MetaModules compared to 50 randomly selected MetaRegulons (Fig. 2b and Additional file 3: Fig. S4a,b). Notably, nearly all of the top 50 predicted MetaRegulons were identified as positive regulators of their respective MetaModules (Additional file 3: Fig. S4c). Furthermore, the MetaRegulons identified by MetroSCREEN showed high MetaModule specificity, including well-known regulators such as FOSL1 [83] and HIF1A [84] for glycolysis, and STAT3 [85] for fatty acid oxidation (Fig. 2c). These results demonstrating MetroSCREEN's ability to accurately pinpoint key MetaRegulons for different MetaModules.

To further validate the MetaRegulons for glycolysis, we calculated the β-score of MetaModules following the perturbation of CTNNB1, ATF4, MYC, and JUN. Encouragingly, CTNNB1 and MYC were shown to positively regulate the conversion of pyruvate to L-lactate in glycolysis [86, 87], MYC and ATF4 positively regulate

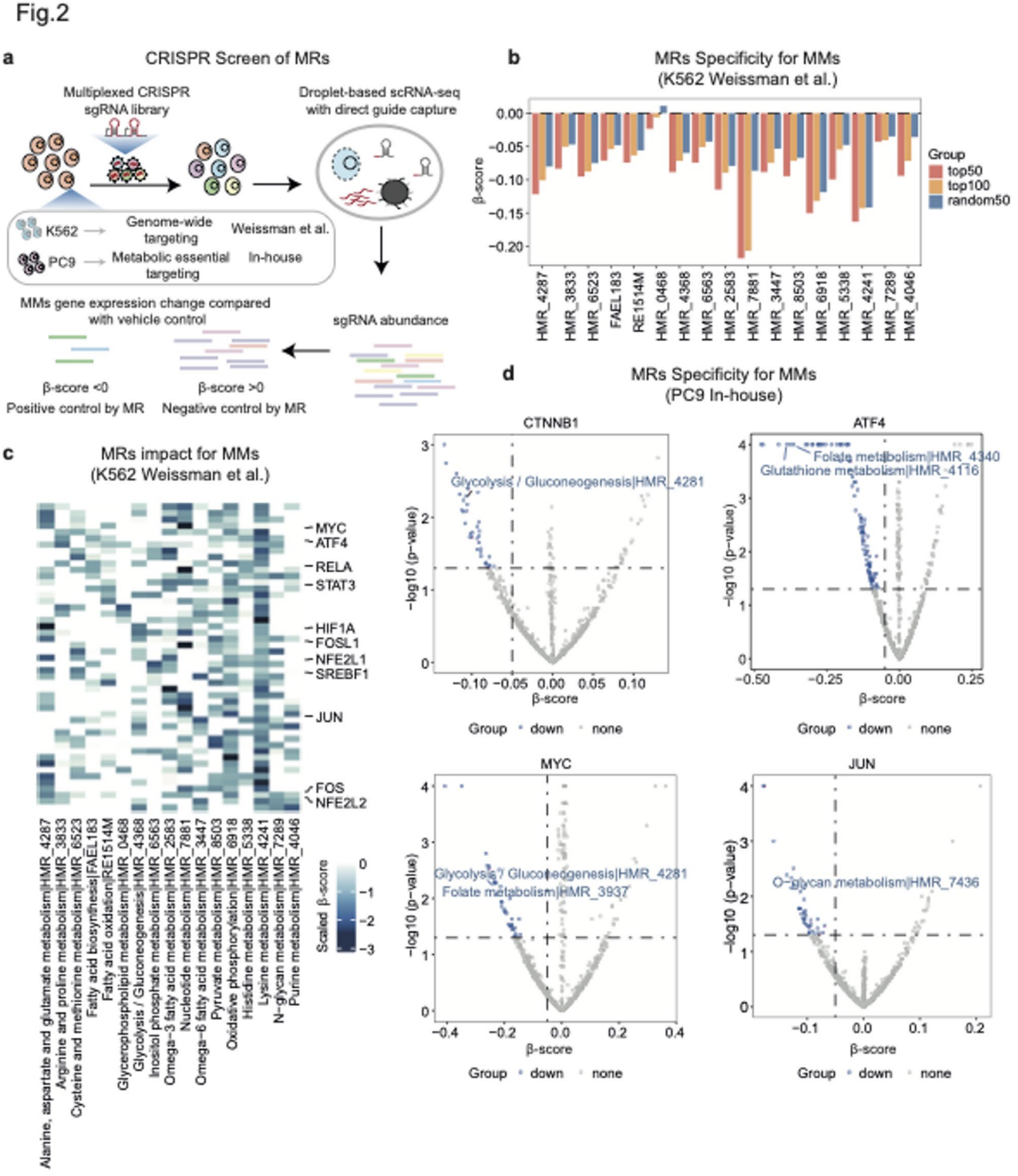

Fig. 2 Performance evaluation of MetaRegulons prediction using single-cell perturbation data. **a** Single-cell CRISPR screening strategy. **b** Bar plot of MetaRegulon specificity for MetaModules in K562. **c** Heatmap of MetaRegulons knockdown effects on MetaModule in K562. **d** Bar plot of MetaRegulon specificity for MetaModules in PC9

folate metabolism [88, 89], ATF4 positively regulates glutathione metabolism, and JUN is involved in O-glycan metabolism (Fig. 2d). These MetaRegulon-targeted MetaModules were top ranked with significant negative β-scores after MetaRegulon perturbation. In summary, these results suggest that MetroSCREEN precisely identifies key MetaRegulons for glycolysis MetaModules.

Pan-cancer analyses of fibroblast and myeloid cell metabolic heterogeneities and regulations

Metabolic reprogramming is a hallmark of cancer; however, most research has primarily focused on the metabolic dysregulation of malignant cells, often neglecting other important components of the TME, such as stromal and immune cells. Since MetroSCREEN can accurately reconstruct the MetaModule scores and their associated MetaRegulons, we applied it to the pan-cancer single-cell atlas, TabulaTIME [71], which we previously developed [42, 71, 90]. This atlas encompasses 4,254,586 cells from 735 patients across 36 cancer types. Our analysis concentrated on two representative TME components: fibroblasts ($n = 236{,}510$) from the stromal lineage and myeloid cells ($n = 504{,}608$) from the immune lineage. We analyzed their corresponding MetaModule dynamics and identified crucial MetaRegulons, leading to the creation of the MetroTIME resource (Metabolism and Regulatory Analyses in the Tumor Immune Microenvironment) (Fig. 3a). This resource was developed through five steps: collection and integration of tumor single-cell profiles, identification of MetaModules and MetaRegulons using MetroSCREEN, pan-cancer analyses of MetaModules, pan-cancer analyses of MetaRegulons, and large-scale cohort exploration for key MetaModules and MetaRegulons (Methods).

Metabolic subtypes of fibroblasts correspond well to cell subtypes and functions

Emerging evidence suggests that fibroblasts are pivotal regulators of tumor metabolism [91, 92]. Metabolic adjustments in fibroblasts are believed to influence adjacent tumor cells they interact with [93, 94]. Therefore, it is crucial to systematically delineate the connections among metabolic heterogeneities and functions of fibroblasts in the TME (Additional file 3: Fig. S5a).

MetroTIME identified seven distinct fibroblast cell subtypes (CS) across various cancer types, classified into four major categories based on their functions: extracellular matrix (ECM)-remodeling fibroblasts (eFibro and myoFibro), immune-regulatory fibroblasts (iFibro), antigen-presenting fibroblasts (apFibro), and quiescent fibroblasts (qFibro) (Fig. 3b and Additional file 3: Fig. S5b-h). Each cell subtype was annotated according to its highly expressed marker genes. Additionally, MetroTIME recognized eight distinct metabolic subtypes (MS) through unsupervised clustering of MetaModule scores (Fig. 3c and Additional file 3: Fig. S6). Among these subtypes, GLY displayed significantly high glycolysis levels, GLYCAN exhibited the highest chondroitin sulfate biosynthesis, PUFA was enriched in arachidonic acid-related metabolism, and LYS showed elevated score in lysine metabolism and inositol phosphate metabolism (Fig. 3d). Notably, the MSs demonstrated a strong correlation with the CSs, indicating that a cell's metabolic score is heavily influenced by its identity (Fig. 3e and Additional file 3: Fig. S7a, ARI = 0.92). Focusing on the major fibroblast categories, eFibro and myoFibro, characterized by high ECM remodeling scores, were significantly enriched in energy metabolism MSs, including OXP, IPM, and GLYCAN (Fig. 3f and Additional file 5: Table S4). In contrast, iFibro and apFibro, which have immune regulatory functions, were predominantly associated with the LYS and PUFA subtypes (Fig. 3f). The iFibro_IL6 was identified as a major source of PUFA subtypes, showing upregulation of TNFα signaling (Additional file 3: Fig. S7b). This finding aligns with previous research demonstrating that polyunsaturated fatty acids, such as prostaglandins, exacerbate inflammation and promote tumorigenesis by increasing the gene expression of TNFα and IL6 [95–98].

To further explore the relationship between cellular functions and MSs, we calculated the Pearson correlation between MetaModules scores and functional signature scores. Consistently, MetaModules enriched in the GLYCAN subtype exhibited a higher correlation with ECM remodeling functions compared to other MetaModules. Additionally, MetaModules enriched in the PUFA and LYS subtypes showed higher correlations with immune-regulatory and antigen-presenting functions, respectively (Additional file 3: Fig. S7c-e). Overall, our results reveal that metabolic subtypes closely align with cell subtypes and functions. This alignment likely arises from similar cell subtypes being governed by comparable intrinsic programs and existing within analogous TME contexts, exposing them to similar extrinsic signals.

Distinct energy usage between ECM-associated fibroblasts

Energy metabolism, the process of generating energy (ATP) from nutrients, is essential for maintaining cellular homeostasis and adapting to varying conditions [48, 99]. Our analyses indicated that eFibro and myoFibro are highly enriched in energy metabolism MSs (Fig. 3f). To explore their differences further, we conducted lineage tracing analysis for all fibroblasts (Methods). Consistent with previous studies, qFibro_SAA1 was located at the root of the trajectory branches, which included three distinct activation paths: the ECM remodeling path (E), the myofibroblast path (M), and the immune-regulatory path (I) (Fig. 4a-b and Additional file 3: Fig. S7f-h). Notably, the E path began with the AA subtype, transitioned

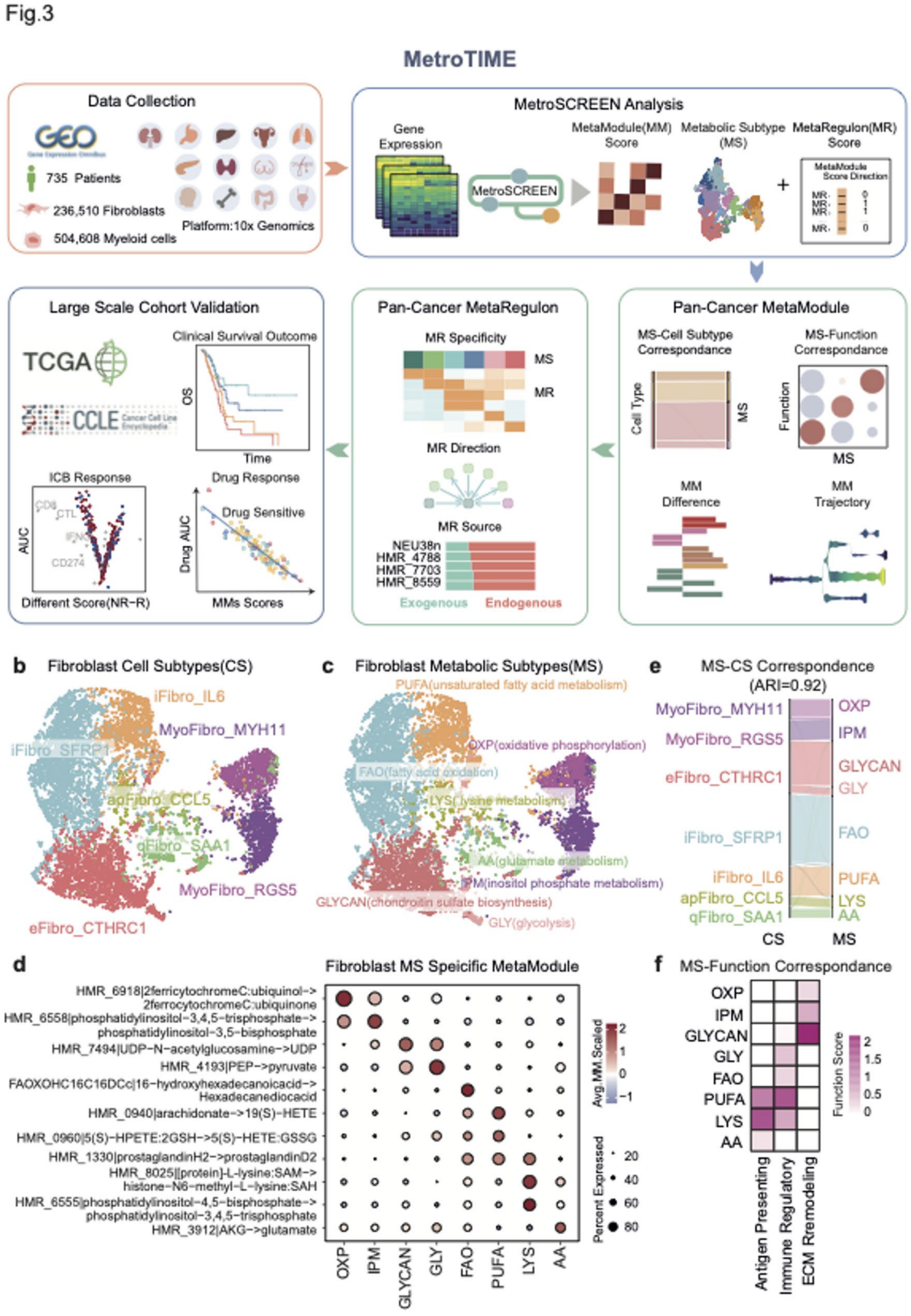

Fig. 3 Linkage of metabolic subtypes, cell phenotypes and cell functions. **a** Schematic illustrating the application of MetroSCREEN to pan-cancer scRNA-seq datasets. **b** UMAP of fibroblast Metacells, colored by cell subtypes. **c** UMAP of fibroblast Metacells, colored by metabolic subtypes. **d** Dot plot showing MetaModules used to annotate fibroblast metabolic subtypes. **e** Alluvial plot showing the correspondence between metabolic subtypes and cell subtypes. **f** Heatmap of enriched functions for each fibroblast metabolic subtype

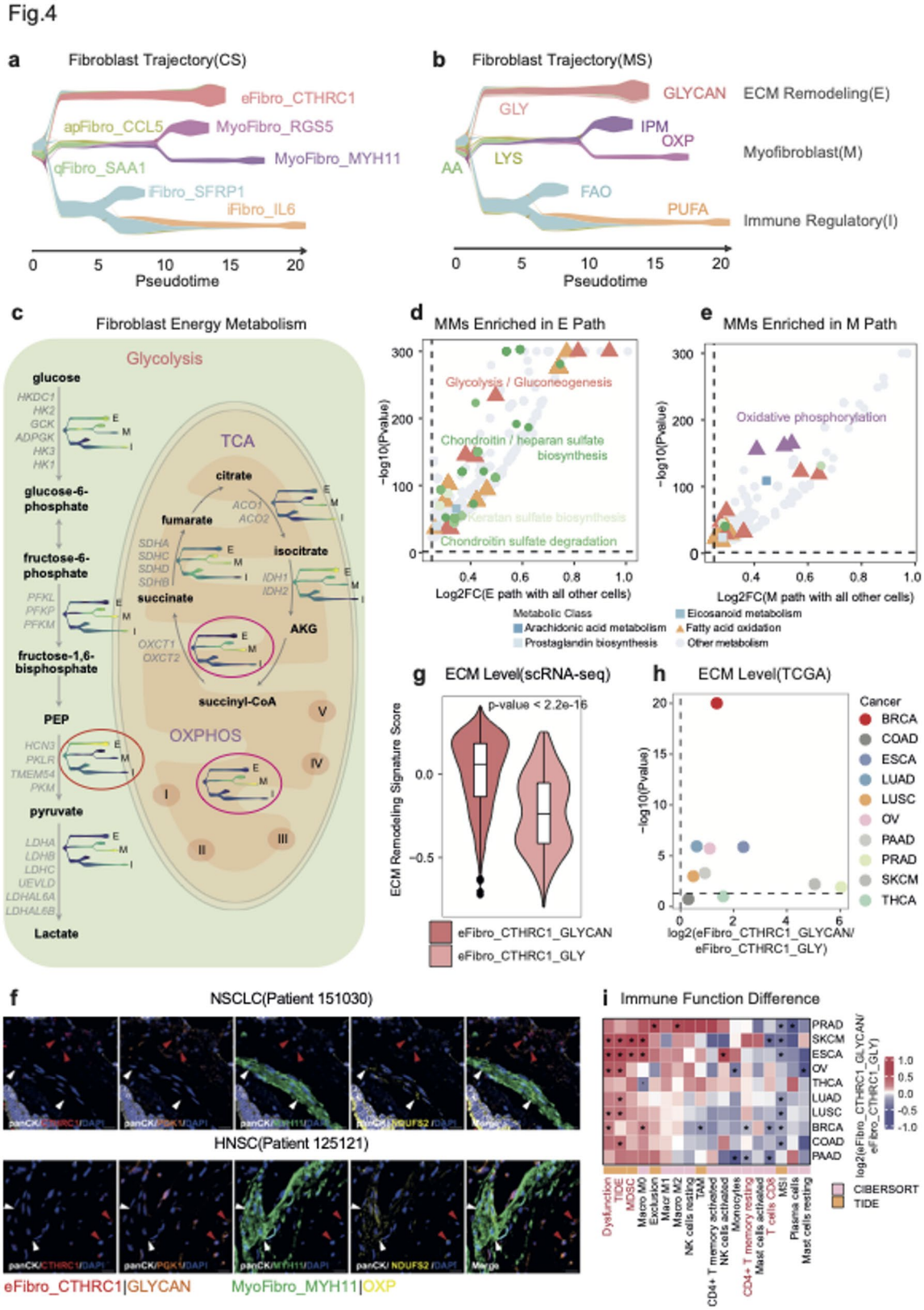

Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Distinct energy usage between ECM-associated fibroblasts. **a** Pseudotime-ordered analysis of fibroblasts Metacells, colored by cell subtypes. **b** Pseudotime-ordered analysis of fibroblasts Metacells, colored by metabolic subtypes. **c** Cartoon summarizing potential energy-supply pathways—glycolysis, TCA cycle, and OXPHOS—for each trajectory branch inferred from MetaModule scores. **d** Dot plot of MetaModules enriched on the E branch. Metabolic classes indicated by markers: energy metabolism (triangle), chondroitin-sulfate metabolism (circle), and unsaturated fatty-acid metabolism (square). **e** Dot plot of MetaModules enriched in the M branch. **f** mIHC staining to validate the energy usage between eFibro and myoFibro. Scale bars, 20 μm. **g** Violin plot of ECM levels within the eFibro_CTHRC1 fibroblast subtype from integrated scRNA-seq data; dark red denotes the ECM-remodeling signature score in the GLYCAN subtype, light red denotes the score in the GLY subtype. **h** Dot plot of the ECM-remodeling signature within the eFibro_CTHRC1 subtype form deconvolved TCGA data. The x-axis shows log2(median ECM score of GLYCAN subtypes – meadian ECM scores of GLY subtype). The y-axis reports the Wilcoxon rank-sum test significance comparing ECM scores between GLYCAN and GLY metabolic subtypes. **i** Heatmap illustrating immune-function differences between GLYCAN and GLY subtypes in TCGA data

to GLY, and culminated in the GLYCAN subtype. The M path progressed through the LYS subtype, reaching the IPM and OXP subtypes, while the I path extended from the FAO to the PUFA subtype. These analyses suggest that fibroblasts in the TME undergo significant energy reprogramming during differentiation.

Interestingly, eFibro and myoFibro diverged into two branches, corresponding to the E and M paths. Mapping detailed MetaModule scores onto the trajectory revealed that the E path exhibited elevated score in MetaModules responsible for converting phosphoenolpyruvate (PEP) to pyruvate-a critical step in converting ADP to ATP and the final rate-limiting stage of glycolysis [100, 101] (Fig. 4c, d). Previous research reported that the ECM relies heavily on glycolysis as a primary source of ATP to support its stiffness [102, 103]. Our findings align with the established metabolic traits of cancer-associated fibroblasts (CAFs), which rely on enhanced glycolysis to drive tumor progression [104]. In contrast, fibroblasts along the M path rely on energy derived from OXPHOS and the tricarboxylic acid (TCA) cycle, which are the main energy sources in most non-malignant cells [105, 106] (Fig. 4c, e). Significantly, mIHC staining for CTHRC1, MYH11, PGK1 (Phosphoglycerate Kinase 1, a key enzyme in glycolysis), and NDUFS2 (NADH Dehydrogenase [Ubiquinone] Iron-Sulfur Protein 2, a component of mitochondrial complex I) in non-small-cell lung cancer (NSCLC) and Head and Neck Squamous Cell Carcinoma (HNSC) patient samples has further confirmed differences in energy usage between eFibro and myoFibro (Fig. 4f; Additional file 3: Fig. S7i). This suggests that distinct fibroblast subtypes, while functionally similar, may generate energy through different metabolic reactions.

Linkage of metabolic subtypes and cellular functions within CTHRC1$^+$ fibroblasts

Unlike other cell subtypes, the eFibro_CTHRC1 subtype exhibited two distinct metabolic subtypes, each associated with different cellular functions (Fig. 3e, f). Compared to the GLY subtype, the GLYCAN subtype showed a significantly higher ECM remodeling signature score (Figs. 3f and 4g). Conversely, the GLY subtype exhibited elevated levels of TNFα signaling via the NF-κB pathway, indicating heightened inflammatory responses (Additional file 3: Fig. S7b). The eFibro_CTHRC1 is known to reside at the tumor's leading edge, forming a physical barrier that may hinder T-cell infiltration into the TME [107]. To further explore the functional differences between these two MSs, we analyzed a large-scale cohort from the Cancer Genome Atlas (TCGA) cohort (Methods). We employed BayesPrism [108] to deconvolute the mixed transcriptomes and isolate the gene expression profile of the eFibro_CTHRC1 subtype from bulk TCGA data. We included ten primary cancer types, as high-quality single-cell RNA-seq datasets from the same tissues were available for reference (Additional file 6: Table S5). To stratify patients into the GLY and GLYCAN subtypes, we calculated the respective MetaModule scores based on the deconvoluted eFibro_CTHRC1 expression profiles. Patients with scores above the median for GLY and GLYCAN were classified as the eFibro_CTHRC1_GLY and eFibro_CTHRC1_GLYCAN groups, respectively. The patients in the eFibro_CTHRC1_GLYCAN group showed higher ECM signature score compared with the patients in the eFibro_CTHRC1_GLY group (Fig. 4h). We then performed TIDE [81] analyses to assess immune function status in different patient groups. Notably, patients in the eFibro_CTHRC1_GLYCAN group demonstrated higher T-cell dysfunction scores, TIDE scores, and MDSC scores (Fig. 4i). This suggests that the elevated GLYCAN subtype score may hinder T-cell functions, contributing to an immunosuppressive TME [81, 109–111]. In contrast, patients in the eFibro_CTHRC1_GLY group exhibited increased infiltration of CD8$^+$ T cells and CD4$^+$ T cells, suggesting a more favorable immune response. These findings highlight the close relationship between metabolic heterogeneity and cellular functions, even within the same cell subtype.

Highly heterogeneous metabolic subtypes of tumor-infiltrated macrophages

Metabolic fluctuations in immune cells are closely linked to their phenotypes and functions [112]. Tumor-associated macrophages (TAMs) and their precursors constitute the largest fraction of the myeloid cells in most human solid tumors. These cells have a significant capacity to uptake intertumoral glucose and play diverse roles

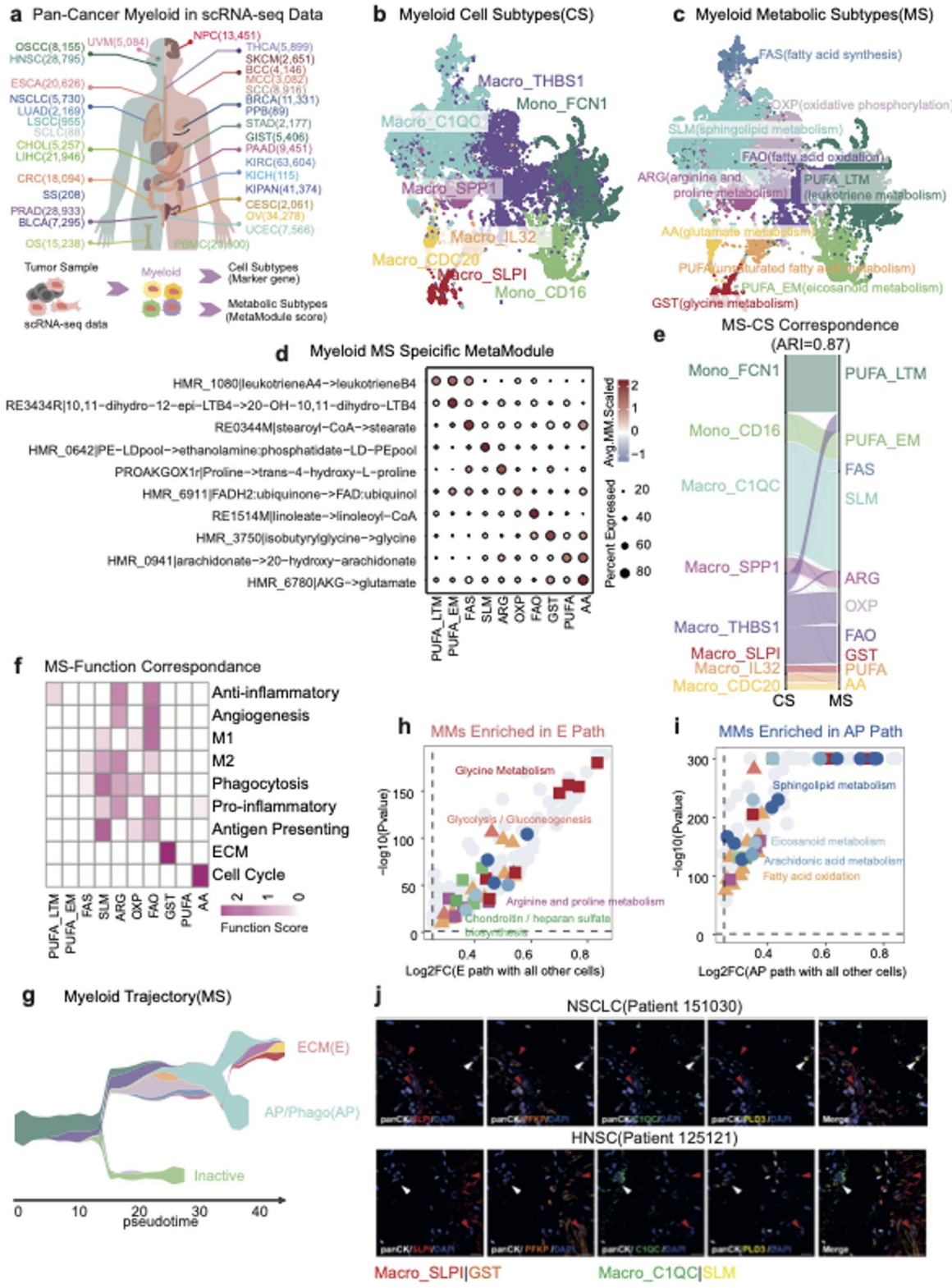

Fig. 5 (See legend on next page.)

(See figure on previous page.)

Fig. 5 Metabolic heterogeneity in myeloid cells. **a** Cancer types included in the pan-cancer myeloid analysis, with schematics of myeloid single-cell subtypes and their metabolic classifications. **b** UMAP of myeloid Metacells, colored by cell subtypes. **c** UMAP of myeloid Metacells, colored by metabolic subtypes. **d** Dot plot of MetaModules used to annotate myeloid metabolic subtypes. **e** Alluvial plot showing the correspondence between metabolic subtypes and cell subtypes. **f** Heatmap of enriched functions for each myeloid metabolic subtype. **g** Pseudotime-ordered analysis of myeloid Metacells, colored by metabolic subtypes. **h** Dot plot of MetaModules enriched in the ECM remodeling branch. Metabolic classes indicated by markers: energy metabolism (triangle), chondroitin sulfate-related metabolism (circle), and unsaturated fatty acid metabolism (square). **i** Dot plot of MetaModules enriched in the antigen-presenting and phagocytosis branch. **j** mIHC staining to validate the metabolism adaptions between E path and AP path. Scale bars, 20 μm

in either promoting or inhibiting tumor progression across different cancer types [113, 114].

We investigated the metabolic subtypes of macrophages and their precursor cells, monocytes, in the TME across 33 cancer types (Fig. 5a and Additional file 5: Table S4). A total of 13 distinct cell subtypes and 10 metabolic subtypes were identified based on marker genes and MetaModule scores (Fig. 5b-c; Additional file 3: Fig. S8 and Additional file 3: Fig. S9). The annotation of metabolic subtypes was validated by MetaModule scores (Fig. 5c, d). For example, the SLM subtype showed increased MetaModule scores in sphingolipid metabolism, ARG exhibited higher MetaModule scores in arginine metabolism, GST showed higher MetaModule scores in glycine metabolism, FAO exhibited higher MetaModule scores in fatty acid metabolism, and PUFA was linked to increased arachidonic acid metabolism. Notably, the PUFA_LTM and PUFA_EM subtypes represent monocyte subsets enriched in polyunsaturated fatty acids, such as leukotriene B4 (LTB4) and eicosanoids which are known to modulate both innate and adaptive immune responses [115–117]. Similar to fibroblasts, a strong correspondence was observed between the MSs and CSs of myeloid cells (Fig. 5e and Additional file 3: Fig. S10a, ARI = 0.87). However, several CSs displayed greater heterogeneity: Macro_C1QC corresponded to both the SLM and FAS subtypes, while Macro_THBS1 primarily aligned with the OXP and FAO subtypes (Fig. 5e). These results highlight the broad spectrum of activation subtypes of TAMs associated with their metabolic profiles and cellular plasticity.

Sphingolipid metabolisms are associated with antigen presentation and phagocytosis in TAMs

To explore the relationship between metabolic subtypes and cellular functions, we calculated the enrichment of macrophage functional signatures across different MSs. The GST subtype demonstrated higher scores for ECM remodeling and epithelial-mesenchymal transition (EMT) hallmarks, indicating its potential pivotal roles in tumor metastasis (Fig. S5f and Additional file 3: Fig. S10b). Notably, while both the OXP and FAO subtypes correspond to the Macro_THBS1 subtype, OXP exhibited an elevated phagocytosis score, whereas FAO showed a higher angiogenesis score (Fig. 5f). This suggests that metabolic subtypes may provide a more accurate definition of cellular functions than cell subtypes. The SLM subtype exhibited higher scores for antigen presentation and phagocytosis, aligning with previous studies that highlight the critical role of sphingolipids in phagocytosis during fungal infections [118–120]. Pearson correlation analysis confirmed that MetaModules enriched in the SLM subtype were more strongly correlated with the antigen presentation score compared to the background (Additional file 3: Fig. S10c-e). This suggests that macrophages with the SLM subtype are crucial in driving anti-tumor activities through antigen presentation and phagocytosis.

We then performed lineage tracing analysis to examine metabolic reprogramming along differentiation trajectories. Starting from monocytes, macrophages primarily follow two distinct paths: the ECM path (E path) and the antigen-presenting/phagocytosis path (AP path) (Fig. 5g and Additional file 3: Fig. S10f-h). Consistent with observations in fibroblasts, MetaModules HMR_4281 and HMR_4358, which contribute to ATP production through glycolysis, correlated with increased ECM remodeling along the trajectory. In contrast, HMR_6916, which facilitates ATP production via OXPHOS, increased alongside elevated antigen presentation and phagocytosis scores (Additional file 3: Fig. S10i). These findings suggest a tight linkage between energy reprogramming from OXPHOS to glycolysis and functional shifts among antigen presentation, phagocytosis, and ECM remodeling in TAMs [121–123]. Furthermore, the E path enriched the glycine metabolism MetaModules, while the AP path enriched the sphingolipid metabolism, aligning with the metabolic subtype functions associated with each path (Fig. 5h-i, $p < 0.001$). Significantly, mIHC staining for SLPI, C1QC, PFKP (Phosphofructokinase, a key enzyme in glycolysis), and PLD3 (Phospholipase D3, a key enzyme involved in sphingolipid metabolism) in NSCLC and HNSC patient samples has further validated differences in energy usage and sphingolipid metabolism between the profibrotic and antigen-presenting/phagocytosis macrophages (Fig. 5j and Additional file 3: Fig. S10j).

MetroSCREEN identifies conserved and metabolic subtype-specific regulators

The pan-cancer MetaModule analyses indicated that metabolic subtypes are closely linked to cellular functions in both fibroblasts and myeloid cells. We then investigated whether there are common upstream regulators for

Fig.6

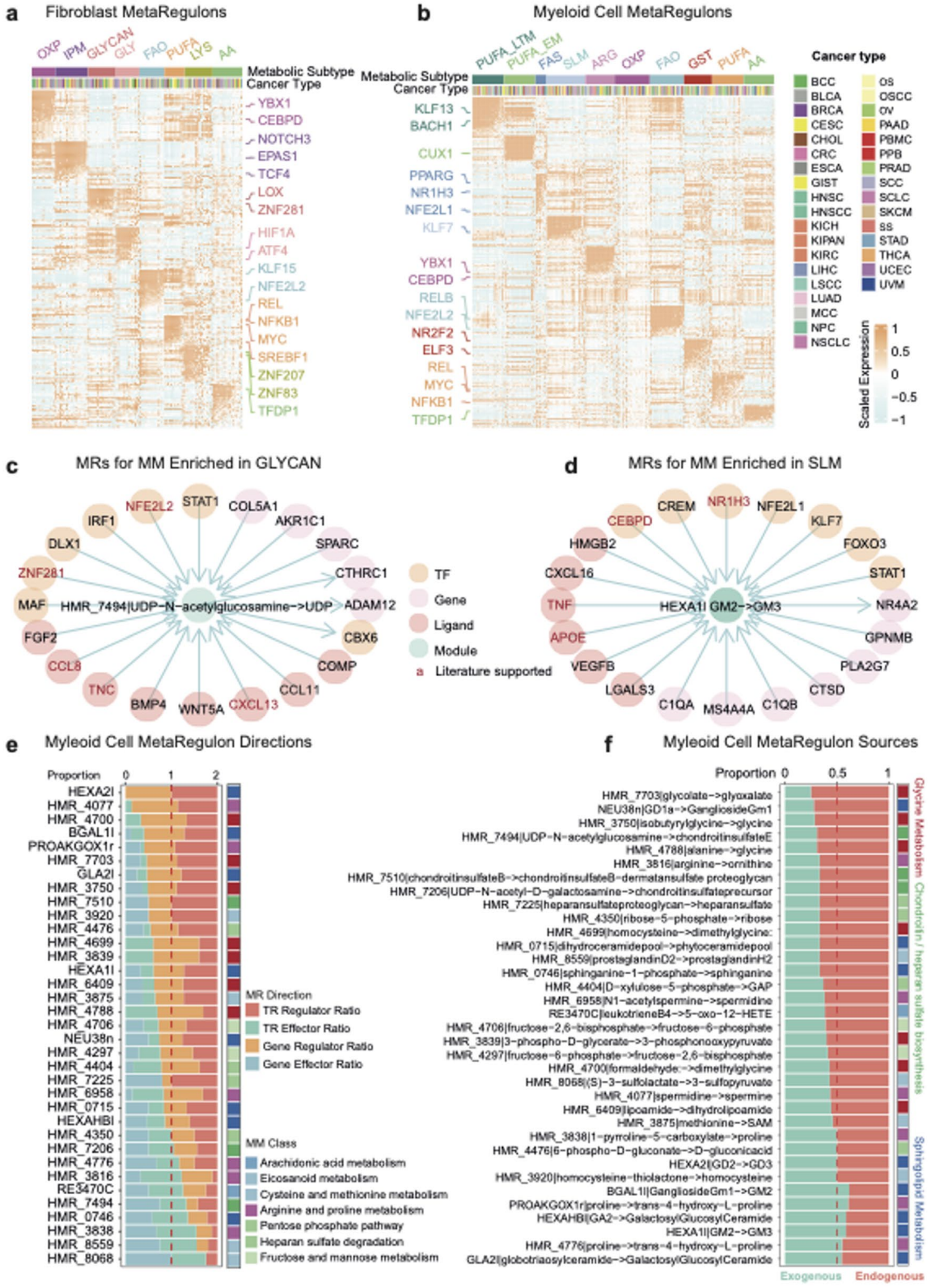

Fig. 6 (See legend on next page.)

(See figure on previous page.)

Fig. 6 MetroSCREEN identifies conserved and subtype-specific metabolic regulators in fibroblasts and myeloid cells. **a** Heatmap showing fibroblast MetaRegulons expression across metabolic subtypes. **b** Heatmap showing myeloid MetaRegulons expression across metabolic subtypes. **c** Causal inference network for HMR_7494 in the GLYCAN fibroblast subtype; literature-supported MetaRegulons are labeled in red. **d** Causal inference network for HEXA1l in the SLM myeloid subtypes; literature-supported MetaRegulons are labeled in red. **e** Proportion plot showing the predicted regulator-to-effector ratio across myeloid metabolic subtypes. **f** Proportion plot showing ratio of the exogenous and endogenous MetaRegulons for the enriched MetaModules in the metabolic subtypes

specific MSs. We applied MetroSCREEN to MetaModules of different MSs in fibroblasts and myeloid cells and identified their corresponding MetaRegulons (Fig. 6a, b). Most inferred MetaRegulons displayed high specificity for their respective MSs across various cancer types and cell lineages, suggesting a relatively conserved metabolic regulation within the TME (Fig. 6a, b). For example, the two FAO subtypes, corresponding to iFibro_SFRP1 and Macro_THBS1, both enriched for NFE2L2, a MetaRegulon known to enhance mitochondrial fatty acid oxidation [124]. Similarly, the MetaRegulons YBX1 and CEBPD were enriched in both OXP subtypes linked to MyoFibro_MYH11 and Macro_THBS1, consistent with their roles in regulating mitochondrial activity [125, 126]. Furthermore, for CSs with heterogenous MSs, MetroSCREEN can identify MS-specific MetaRegulons (Fig. 6a, b). The eFibro_CTHRC1_GLY subtype was regulated by HIF1A and ATF4, two well-known TRs that modulate the expression of metabolic enzymes involved in glycolysis [84, 127]. In contrast, the eFibro_CTHRC1_GLYCAN subtype was predominantly regulated by ZNF281, which is associated with ECM remodeling and fibrosis [128]. Additionally, PPARG, an important regulator of fatty acid synthesis [129], was enriched in the FAS subtype rather than the SLM subtype, even though both metabolic subtypes correspond to Macro_C1QC. These analyses indicate the conserved metabolic subtypes across different cell lineages may be regulated by similar MetaRegulons.

Next, we evaluated MetroSCREEN's capability to construct a metabolic regulation network (MRN) that includes both intrinsic transcription regulators and extrinsic ligands. We focused on building a collagen metabolism-related MRN for the eFibro_CTHRC1_GLYCAN subtype and a sphingolipid metabolism-related MRN for the Macro_C1QC_SLM subtype, considering the enrichment of these metabolisms and their association with the functions of these cell subtypes as identified in our previous findings (Fig. 6c, d). MetroSCREEN identified several MetaRegulons associated with ECM functions in the eFibro_CTHRC1_GLYCAN subtype. For instance, NFE2L2 is known to directly target extracellular matrix gene transcription [130], and silencing of ZNF281 has been shown to abolish the increase in extracellular collagen levels [128]. In addition to MetaRegulons directly involved in ECM remodeling, ligands such as CXCL13 [131], TNC [132], and CCL8 [133] have been reported to influence the expression of collagen-related genes. In the SLM subtype of Macro_C1QC, NR1H3 is recognized for directly regulating sphingolipid synthesis [134], while CEBPD specifically contributes to lipid accumulation [135]. Moreover, ligands such as TNF [136] and APOE [137] have been previously linked to sphingolipid metabolism. In addition to accurately predicting these known potential metabolic regulators, we also identified novel metabolic regulatory factors within each MRN, such as STAT1 in regulating collagen metabolism and KLF7 in sphingolipid metabolism. Taken together, our results suggest that MetroSCREEN can accurately identify both conserved and MS-specific MetaRegulons across different lineages, covering both intrinsic and extrinsic regulatory mechanisms.

MetroSCREEN offers insights into regulator directions and sources

Previous research has primarily relied on correlations to predict regulators mediating metabolite changes [60]. However, correlation does not establish causality between metabolic states and TR activities. MetroSCREEN combines constraint-based algorithms with independence tests to predict causal relationships between MetaRegulons and their associated MetaModules (Methods). Encouragingly, our analyses indicate that a significant proportion of TRs are indeed upstream regulators of MetaModules, with 62.86% in myeloid cells and 71.43% in fibroblasts (Fig. 6e and Additional file 3: Fig. S11a). For MetaRegulons identified through genome-wide correlations, 62.86% of them were still regulators of MetaModules in myeloid cells, while only 13.33% served as regulators in fibroblasts.

We further analyzed the sources of MetaRegulons for each MetaModule (Fig. 6f and Additional file 3: Fig. S11b). MetaModules involved in glycolysis and collagen synthesis are more likely to be regulated by endogenous mechanisms. Conversely, MetaModules associated with sphingolipid metabolism tend to be influenced more by exogenous factors, likely due to their role as primary components of the cell membrane. Overall, the majority of MetaModules were controlled by intrinsic factors, accounting for 82.86% in myeloid cells and 73.33% in fibroblasts. These findings highlight the dynamic regulatory directions and the complex resources of MetaRegulons for each MetaModule.

Fig.7

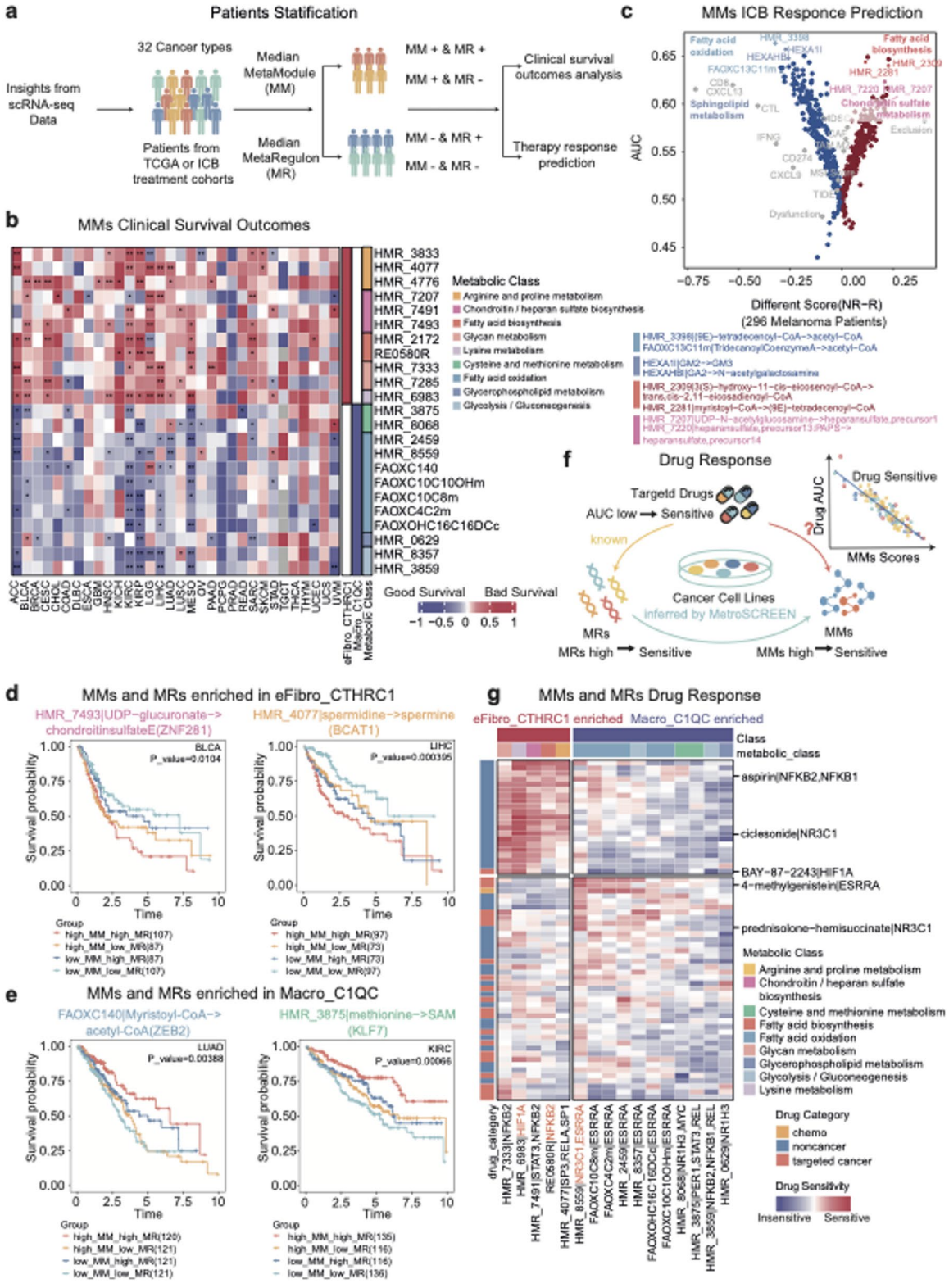

Fig. 7 (See legend on next page.)

(See figure on previous page.)

Fig. 7 MetaModule and MetaRegulon scores are linked to survival and therapy response. **a** Schematic overview of the workflow employed for pan-cancer single-cell analysis of MetaModules and MetaRegulons in clinical data. **b** Survival associations of MetaModules with significant prognostic impact in at least five cancer types. The heatmap highlights MetaModules enriched in the eFibro_CTHRC1_GLYCAN and Macro_C1QC_SLM subtypes. **c** Area under the ROC curve (AUC) for each MetaModule in predicting response to immune checkpoint blockade (ICB) therapy. **d** Kaplan–Meier curves showing overall survival for four stratified groups defined by MetaModule scores and MetaRegulon expression, highlighting eFibro_CTHRC1_GLYCAN–enriched MetaModules associated with poor survival. **e** Kaplan–Meier curves showing overall survival for four stratified groups defined by MetaModule scores and MetaRegulon expression, highlighting Macro_C1QC_SLM–enriched MetaModules associated with good survival. **f** Schematic of drug-sensitivity analysis for MetaModules and MetaRegulons; a negative correlation between drug AUC and MetaModule score indicates greater drug sensitivity. **g** Heatmap of drug-response profiles for MetaModules and MetaRegulons

Metabolic modules are associated with survival and therapy response

We next explored the potential clinical impact of different MetaModules and MetaRegulons identified through pan-cancer single-cell analyses. We stratified 9,550 patients across 32 cancer types from TCGA datasets using the average scores of MetaModules and MetaRegulons (Fig. 7a). Among all MetaModules, collagen metabolism that enriched in eFibro_CTHRC1_GLYCAN, such as chondroitin/heparan sulfate biosynthesis and glycan metabolism, along with arginine and proline metabolism, have been linked to poor survival outcomes in at least five cancer types (Fig. 7b). For example, MetaModule HMR_7493 facilitates the transfer of UDP-glucuronate to chondroitin sulfate E, while HMR_4776 catalyzes the conversion of proline to trans-4-hydroxy-L-proline; both are associated with worse survival outcomes. In contrast, MetaModules enriched in Macro_C1QC_SLM, including those involved in fatty acid oxidation and unsaturated fatty acid pathways such as omega-3 and arachidonic acid metabolism, were linked to better survival outcomes (Fig. 7b). We further investigated whether these MetaModules might regulate immune functions [79]. As expected, MetaModules associated with poor survival showed a positive correlation with immune suppressive scores, while those linked to better survival demonstrated a negative correlation (Additional file 3: Fig. S12a, b).

We also evaluated whether MetaModules could serve as responsive biomarkers to predict immune checkpoint blockade (ICB) therapy responses. We compared the predictive power of MetaModules with previously published ICB response biomarkers across various ICB cohorts, including the melanoma PD1 treatment cohort [48, 81, 138–140] (Methods; Additional file 7: Table S6 and Additional file 8: Table S7). Notably, several MetaModules demonstrated higher predictive power for ICB responses than other biomarkers (Fig. 7c). For instance, fatty acid oxidation and sphingolipid metabolism showed elevated score in responsive patients, while fatty acid synthesis and chondroitin sulfate metabolism were more enriched in non-responsive patients. Interestingly, these metabolic classes are separately enriched in eFibro_CTHRC1_GLYCAN and Macro_C1QC_SLM subtypes (Fig. 7c and Additional file 3: Fig. S12c, d). These results emphasize the critical roles of collagen metabolism and sphingolipid metabolism in tumor progression and immune therapy response. Furthermore, the findings highlight the significance of lipid metabolic reprogramming, especially the dynamic equilibrium between fatty acid oxidation and biosynthesis, in impacting both tumor progression and immune therapy outcomes.

Targeting metabolic modules and regulators for effective cancer treatment

To determine whether the combination of MetaModules and MetaRegulons can enhance predictions of overall survival outcomes, we stratified patients into four groups based on their MetaModule scores and corresponding upstream MetaRegulons (Fig. 7a). Patients with higher scores for both MetaModules and MetaRegulons were associated with the poorest survival, particularly for MetaModules enriched in eFibro_CTHRC1_GLYCAN (Fig. 7d and Additional file 3: Fig. S12e). For instance, the group with elevated scores for ZNF281 and HMR_7493, which involved in collagen metabolism, exhibited the worst survival prognosis. Similarly, the group with elevated scores for BCAT1 and HMR_4077, catalyzing the conversion of spermidine to spermine, also showed a comparable pattern. Conversely, for MetaModules enriched in Macro_C1QC_SLM, such as FAOXC140 and HMR_3875, patients with higher scores for these MetaModules and their regulators ZEB2 and KLF7 had the best survival compared to the other patients (Fig. 7e and Additional file 3: Fig. S12f). These findings underscore the importance of considering both MetaModule scores and their MetaRegulons when investigating clinically significant metabolic pathways.

Next, we explored the potential of targeting MetaModules and MetaRegulons enriched in eFibro_CTHRC1_GLYCAN and Macro_C1QC_SLM within the CCLE dataset. This dataset encompasses approximately 1,500 drug tests across 578 cell lines [141] (Fig. 7f and Additional file 9: Table S8). Drug sensitivity in the CCLE is expressed as AUC values; where lower AUC values indicate increased sensitivity to treatment. Thus, a negative correlation between MetaModules scores and drug AUC values implies that cell lines with higher MetaModule score are more sensitive to the drug. As anticipated, the drug sensitivity of MetaModules mirrored that of their upstream MetaRegulons (Fig. 7g). For example,

the HIF1A-regulated MetaModule HMR_6983, which catalyzes the conversion of L-lysine to procollagen-5-hydroxy-L-lysine and is involved in collagen metabolism, was negatively correlated with BAY-87–2243, a drug known to inhibit HIF1A protein accumulation in cancer [142]. Similarly, NR3C1-regulated MetaModules, such as HMR_8559, which converts prostaglandin D2 to prostaglandin H2 and is involved in fatty acid metabolism, exhibited negative correlations with prednisolone-hemisuccinate targeting NR3C1 [141].

Interestingly, many non-cancer drugs such as aspirin (targeting NFKB2/1) and ciclesonide (targeting NR3C1), show significantly higher inhibiting effects on eFibro_CTHRC1_GLYCAN enriched MetaModules, without affecting the MetaModules in Macro_C1QC_SLM, indicating these non-cancer drugs could be repurposed for cancer therapy [141] (Fig. 7g). Overall, our results suggest that MetaModule and MetaRegulon scores significantly impact patient survival, therapy response, and drug tolerance. Consequently, metabolic-based signatures derived from MetaModule and MetaRegulon deregulation could pave the way for improved personalized therapeutic strategies.

Discussion

Metabolic reprogramming alters cellular functions, influencing tumor initiation, progression, and therapy resistance. While several methods exist to infer metabolic states from transcription data, most rely heavily on constraint-based computational approaches, which are often computationally intensive and time-consuming, making them less suitable for single-cell data. In contrast, the statistical method ssGSEA provides a simple and robust option for quantifying metabolic states at the single-cell level. MetroSCREEN enhances computational resolution by analyzing metabolic states at the reaction level, offering comprehensive insights into single-cell metabolic alterations. When benchmarked against paired transcriptomic and metabolite abundance data from the CCLE, MetroSCREEN outperformed the FBA-based method METAFlux, the machine learning-based method scFEA and statistical baselines (mean expression and AUCell) in predicting metabolite abundance at both the sample and metabolite levels.

MetroSCREEN innovatively identifies comprehensive regulators of metabolic modules, incorporating both cell-intrinsic transcriptional regulation and signal transduction, as well as the extrinsic cellular environment. Despite the limited availability of ground truth datasets, MetroSCREEN demonstrates high consistency with metabolic regulators identified from genome-wide CRISPR interference and knockout screen datasets.

Overall, MetroSCREEN proves to be a versatile and user-friendly R-based package for predicting metabolic scores and identifying metabolic regulators using scRNA-seq or bulk RNA-seq data. However, MetroSCREEN has potential limitations, particularly in determining the direction of reversible metabolic reactions. In our future research, we plan to adapt causal deep neural networks [143–145] to elucidate the directionality of these reactions, thereby enhancing MetroSCREEN's ability to identify causal metabolic regulators.

Our analyses with MetroTIME revealed that distinct cell subtypes exhibit varying metabolic heterogeneities, showing a notable concordance between cellular functions and metabolic subtypes. This suggests that the metabolic subtype of cells is rigorously regulated by both intrinsic and extrinsic factors, maintaining their differentiated subtypes and specific functions. Changes in metabolic pathways may drive or result from a cell's transition between subtypes. Notably, different cell lineages with similar functions, such as ECM-associated eFibro_CTHRC1 and Macro_SLPI, exhibited similar metabolic patterns, with upregulated glycolysis and chondroitin sulfate biosynthesis. This indicates that the connection between metabolic status and cellular function transcends cell identities and could be targeted collectively. Patients stratified by MetaRegulons and MetaModules demonstrated distinct survival outcomes and therapy responses, highlighting new opportunities for personalized treatment strategies.

Conclusions

In this work, we introduce MetroSCREEN, a reference-guided computational framework that infers reaction-level metabolic states and nominates candidate upstream regulators from bulk and single-cell transcriptomic data, capturing both intrinsic gene-regulatory and extrinsic cell–cell interaction drivers of metabolic change. Building on an efficient methodology and a large pan-cancer scRNA-seq compendium, we also developed MetroTIME to delineate metabolic subtypes and their corresponding regulons in fibroblasts and myeloid cells across cancers. Our analyses show that these metabolic subtypes are tightly coupled to cellular functions, and that fine-grained metabolic modules can be leveraged in large cohorts to predict patient survival and therapeutic response. The MetroTIME framework further highlights putative therapeutic targets that could be directly inhibited or combined with future immune-modulating strategies.

All software and resources are available at http://wanglab-compbio.cn/MetroTIME/, where users can search, visualize, and analyze the MetaModules and corresponding MetaRegulons reported here. We anticipate that MetroSCREEN and MetroTIME will accelerate the development of metabolism-guided therapies to improve cancer treatment.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-025-01572-z.

Additional file 1: Table S1. sgRNA design for metabolic regulators

Additional file 2: Table S2. Summary of Patients for mIHC in This Study

Additional file 3: Supplementary figures

Additional file 4: Table S3. Function signatures of fibroblasts and myeloid cells

Additional file 5: Table S4. Metabolic and cell subtypes annotation for Metacell in fibroblasts and myeloid cells

Additional file 6: Table S5. Single-cell dataset information for TCGA deconvolution

Additional file 7: Table S6. Immune Checkpoint BlockadeDatasets

Additional file 8: Table S7. Prediction of ICB response using MetaModule

Additional file 9: Table S8. MetaModule and MetaRegulon Drug response

Acknowledgements
We thank members of the Wang laboratory for insightful comments and suggestions.

Authors' contributions
K.T. and Y.H. are equally contributed to this work. C.W. and Q.W. designed the study. K.T. developed the MetroSCREEN method, processed datasets, and performed computational analyses with input from Y.H., D.S., X.D., J.H., and Z.L. K.T. and Y.H. performed the MetroTIME analyses and constructed the web server. K.T. and W.S. performed the MetroSCREEN benchmark in CCLE data. Q.W. and T.H. generated the single-cell CRISPR screen datasets, K.T. and H.W. analyzed the single-cell CRISPR screen datasets. Q.W. performed the mIHC experiments, L.Z, T.L., and P.Z. facilitated sample preparation. K.T. wrote the manuscript, C.W., Q.W., and Y.H. revised the manuscript. C.W. supervised the whole project. All authors read and approved the final manuscript.

Funding
This work was supported by the National Key R&D Program of China [2022YFA1106000], the National Natural Science Foundation of China [32222026, 32170660, 82521002], the Natural Science Foundation of Shanghai [24ZR1492800], Tongji University Medicine-X Interdisciplinary Research Initiative. AI for Science Program of Shanghai Municipal Education Commission. Tongji University Spark-X Program. Shanghai Pilot Program for Basic Research, the Fundamental Research Funds for the Central Universities [22120240435].

Data availability
Metabolomics, genome-scale transcriptomic, and drug-response datasets were obtained from the Cancer Cell Line Encyclopedia (CCLE) via the DepMap Public 23Q4 release [36] (https://sites.broadinstitute.org/ccle/). The public single-cell perturbation dataset was retrieved from the SRA under BioProject PRJNA831566 [38] (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA831566). Our in-house single-cell perturbation datasets are available from the NIH Gene Expression Omnibus (GEO) under accession GSE236708 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE236708). Cancer-associated scRNA-seq datasets were collected from the Tumor Immune Single-cell Hub [42] (TISCH; http://tisch.comp-genomics.org/). The tumor clinical metadata and TPM-normalized expression profiles were collected from the TCGA data portal (https://www.cancer.gov/tcga). MetroTIME is available for exploration and visualization at http://wanglab-compbio.cn/MetroTIME/ [146] . MetroSCREEN is available at https://github.com/wanglabtongji/MetroSCREEN [147] , and all analysis scripts used in this study are provided at https://github.com/wanglabtongji/MetroTIME [146] .

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Key Laboratory of Spine and Spinal Cord Injury Repair and Regeneration of Ministry of Education, Department of Orthopedics, School of Life Science and Technology, Tongji Hospital, Tongji University, Shanghai 200092, China
[2]Sycamore Research Institute of Life Sciences, Shanghai 201210, China
[3]Frontier Science Center for Stem Cells, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China
[4]Department of Thoracic Surgery, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai 200433, China
[5]Central Laboratory, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai 200433, China
[6]State Key Laboratory of Oral Diseases, National Clinical Research Center for Oral Diseases, Research Unit of Oral Carcinogenesis and Management, Chinese Academy of Medical Sciences, West China Hospital of Stomatology, Sichuan University, Chengdu 610041, China
[7]National Key Laboratory of Autonomous Intelligent Unmanned Systems, Tongji University, Shanghai 201210, China
[8]Frontier Science Center for Intelligent Autonomous Systems, Tongji University, Shanghai 201210, China

Received: 10 May 2025 / Accepted: 3 November 2025
Published online: 19 December 2025

References

1. Faubert B, Solmonson A, DeBerardinis RJ, American Association for the Advancement of Science. Metabolic reprogramming and cancer progression. Science. 2020;368:eaaw5473. https://doi.org/10.1126/science.aaw5473.
2. Lyssiotis CA, Kimmelman AC. Metabolic interactions in the tumor microenvironment. Trends Cell Biol. 2017;27:863–75. https://doi.org/10.1016/j.tcb.2017.06.003.
3. Koppenol WH, Bounds PL, Dang CV. Otto Warburg's contributions to current concepts of cancer metabolism. Nat Rev Cancer. 2011;11:325–37. https://doi.org/10.1038/nrc3038.
4. Ghesquière B, Wong BW, Kuchnio A, Carmeliet P. Metabolism of stromal and immune cells in health and disease. Nature. 2014;511:167–76. https://doi.org/10.1038/nature13312.
5. Cassetta L, Fragkogianni S, Sims AH, Swierczak A, Forrester LM, Zhang H, et al. Human tumor-associated macrophage and monocyte transcriptional landscapes reveal cancer-specific reprogramming, biomarkers, and therapeutic targets. Cancer Cell. 2019;35:588-602.e10. https://doi.org/10.1016/j.ccell.2019.02.009.
6. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015;21:938–45. https://doi.org/10.1038/nm.3909.
7. Heiden MGV, DeBerardinis RJ. Understanding the intersections between metabolism and cancer biology. Cell. 2017;168:657–69. https://doi.org/10.1016/j.cell.2016.12.039.
8. Sun L, Suo C, Li S-T, Zhang H, Gao P. Metabolic reprogramming for cancer cells and their microenvironment: beyond the Warburg effect. Biochimica et Biophysica Acta (BBA). 2018;1870:51–66. https://doi.org/10.1016/j.bbcan.2018.06.005.
9. Dey P, Kimmelman AC, DePinho RA. Metabolic codependencies in the tumor microenvironment. Cancer Discov. 2021;11:1067–81. https://doi.org/10.1158/2159-8290.CD-20-1211.
10. Locasale JW, Cantley LC. Metabolic flux and the regulation of mammalian cell growth. Cell Metab. 2011;14(4):443–51. https://doi.org/10.1016/j.cmet.2011.07.014.

11. Manning BD, Cantley LC. United at last: the tuberous sclerosis complex gene products connect the phosphoinositide 3-kinase/Akt pathway to mammalian target of rapamycin (mTOR) signalling. Biochem Soc Trans. 2003;31:573–8. https://doi.org/10.1042/bst0310573.
12. Zhao X, Petrashen AP, Sanders JA, Peterson AL, Sedivy JM. SLC1A5 glutamine transporter is a target of MYC and mediates reduced mTORC1 signaling and increased fatty acid oxidation in long-lived Myc hypomorphic mice. Aging Cell. 2019;18:e12947. https://doi.org/10.1111/acel.12947.
13. Osthus RC, Shim H, Kim S, Li Q, Reddy R, Mukherjee M, et al. Deregulation of glucose transporter 1 and glycolytic gene expression by c-Myc. J Biol Chem. 2000;275:21797–800. https://doi.org/10.1074/jbc.C000023200.
14. Stine ZE, Walton ZE, Altman BJ, Hsieh AL, Dang CV. MYC, metabolism, and cancer. Cancer Discov. 2015;5:1024–39. https://doi.org/10.1158/2159-8290.CD-15-0507.
15. Hoxhaj G, Manning BD. The PI3K-AKT network at the interface of oncogenic signalling and cancer metabolism. Nat Rev Cancer. 2020;20:74–88. https://doi.org/10.1038/s41568-019-0216-7.
16. Feinstein R, Kanety H, Papa MZ, Lunenfeld B, Karasik A. Tumor necrosis factor-alpha suppresses insulin-induced tyrosine phosphorylation of insulin receptor and its substrates. J Biol Chem. 1993;268:26055–8.
17. Hummel DM, Fetahu IS, Gröschel C, Manhardt T, Kállay E. Role of proinflammatory cytokines on expression of vitamin D metabolism and target genes in colon cancer cells. J Steroid Biochem Mol Biol. 2014;144:91–5. https://doi.org/10.1016/j.jsbmb.2013.09.017.
18. Ip WKE, Hoshi N, Shouval DS, Snapper S, Medzhitov R. Anti-inflammatory effect of IL-10 mediated by metabolic reprogramming of macrophages. Science. 2017;356:513–9. https://doi.org/10.1126/science.aal3535.
19. Komai T, Inoue M, Okamura T, Morita K, Iwasaki Y, Sumitomo S, et al. Transforming growth factor-β and interleukin-10 synergistically regulate humoral immunity via modulating metabolic signals. Front Immunol. 2018;9:1364. https://doi.org/10.3389/fimmu.2018.01364.
20. Rappez L, Stadler M, Triana S, Gathungu RM, Ovchinnikova K, Phapale P, et al. Spacem reveals metabolic states of single cells. Nat Methods. 2021;18:799–805. https://doi.org/10.1038/s41592-021-01198-0.
21. Seydel C. Single-cell metabolomics hits its stride. Nat Methods. 2021;18:1452–6. https://doi.org/10.1038/s41592-021-01333-x.
22. Bien T, Koerfer K, Schwenzfeier J, Dreisewerd K, Soltwisch J. Mass spectrometry imaging to explore molecular heterogeneity in cell culture. Proc Natl Acad Sci U S A. 2022;119:e2114365119. https://doi.org/10.1073/pnas.2114365119.
23. Wang G, Heijs B, Kostidis S, Mahfouz A, Rietjens RGJ, Bijkerk R, et al. Analyzing cell-type-specific dynamics of metabolism in kidney repair. Nat Metab. 2022;4:1109–18. https://doi.org/10.1038/s42255-022-00615-8.
24. Bourceau P, Geier B, Suerdieck V, Bien T, Soltwisch J, Dreisewerd K, et al. Visualization of metabolites and microbes at high spatial resolution using MALDI mass spectrometry imaging and in situ fluorescence labeling. Nat Protoc. 2023;18:3050–79. https://doi.org/10.1038/s41596-023-00864-1.
25. Zamboni N, Saghatelian A, Patti GJ. Defining the metabolome: size, flux, and regulation. Mol Cell. 2015;58:699–706. https://doi.org/10.1016/j.molcel.2015.04.021.
26. X D, Jj A-H, Tj G, M B, Wr H, Dj L. A Checklist for Reproducible Computational Analysis in Clinical Metabolomics Research. Metabolites. 2022;12. https://doi.org/10.3390/metabo12010087.
27. Md F, C D, M V, D M, G M, L A, et al. Single-cell Digital Twins for Cancer Preclinical Investigation. Methods Mol Biol Clifton NJ. 2020;2088. https://doi.org/10.1007/978-1-0716-0159-4_15.
28. Alghamdi N, Chang W, Dang P, Lu X, Wan C, Gampala S, et al. A graph neural network model to estimate cell-wise metabolic flux using single-cell RNA-seq data. Genome Res. 2021;31:1867–84. https://doi.org/10.1101/gr.271205.120.
29. Wagner A, Wang C, Fessler J, DeTomaso D, Avila-Pacheco J, Kaminski J, et al. Metabolic modeling of single Th17 cells reveals regulators of autoimmunity. Cell. 2021;184:4168-4185.e21. https://doi.org/10.1016/j.cell.2021.05.045.
30. Huang Y, Mohanty V, Dede M, Tsai K, Daher M, Li L, et al. Characterizing cancer metabolism from bulk and single-cell RNA-seq data using METAFlux. Nat Commun. 2023;14:4883. https://doi.org/10.1038/s41467-023-40457-w.
31. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. BMC Bioinformatics. 2013;14:7. https://doi.org/10.1186/1471-2105-14-7.
32. Wu Y, Yang S, Ma J, Chen Z, Song G, Rao D, et al. Spatiotemporal immune landscape of colorectal cancer liver metastasis at single-cell level. Cancer Discov. 2022;12:134–53. https://doi.org/10.1158/2159-8290.CD-21-0316.
33. Xiao Z, Dai Z, Locasale JW. Metabolic landscape of the tumor microenvironment at single cell resolution. Nat Commun. 2019;10:3763. https://doi.org/10.1038/s41467-019-11738-0.
34. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. Nat Methods. 2017;14:1083–6. https://doi.org/10.1038/nmeth.4463.
35. DeTomaso D, Jones MG, Subramaniam M, Ashuach T, Ye CJ, Yosef N. Functional interpretation of single cell similarity maps. Nat Commun. 2019;10:4376. https://doi.org/10.1038/s41467-019-12235-0.
36. DepMap, Broad. DepMap 23Q4 Public. Datasets. Figshare+; 2023. https://doi.org/10.25452/figshare.plus.24667905.v2.
37. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, et al. Comprehensive integration of single-cell data. Cell. 2019;177:1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.
38. Replogle JM, Saunders RA, Pogson AN, Hussmann JA, Lenail A, Guna A, et al. Mapping information-rich genotype-phenotype landscapes with genome-scale perturb-seq. Cell. 2022;185:2559-2575.e28. https://doi.org/10.1016/j.cell.2022.05.013.
39. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8:14049. https://doi.org/10.1038/ncomms14049.
40. Wei H, Han T, Li T, Wu Q, Wang C. Scree: a comprehensive pipeline for single-cell multi-modal CRISPR screen data processing and analysis. Brief Bioinform. 2023;24:bbad123. https://doi.org/10.1093/bib/bbad123.
41. Yang L, Zhu Y, Yu H, Cheng X, Chen S, Chu Y, et al. scMAGeCK links genotypes with multiple phenotypes in single-cell CRISPR screens. Genome Biol. 2020;21:19. https://doi.org/10.1186/s13059-020-1928-4.
42. Han Y, Wang Y, Dong X, Sun D, Liu Z, Yue J, et al. TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. Nucleic Acids Res. 2023;51:D1425–31. https://doi.org/10.1093/nar/gkac959.
43. Gide TN, Quek C, Menzies AM, Tasker AT, Shang P, Holst J, et al. Distinct immune cell populations define response to anti-PD-1 monotherapy and anti-PD-1/anti-CTLA-4 combined therapy. Cancer Cell. 2019;35:238-255.e6. https://doi.org/10.1016/j.ccell.2019.01.003.
44. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. Cell. 2016;165:35–44. https://doi.org/10.1016/j.cell.2016.02.065.
45. Liu D, Schilling B, Liu D, Sucker A, Livingstone E, Jerby-Arnon L, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. Nat Med. 2019;25:1916–27. https://doi.org/10.1038/s41591-019-0654-5.
46. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. Cell. 2017;171:934-949.e16. https://doi.org/10.1016/j.cell.2017.09.028.
47. Freeman SS, Sade-Feldman M, Kim J, Stewart C, Gonye ALK, Ravi A, et al. Combined tumor and immune signals from genomes or transcriptomes predict outcomes of checkpoint inhibition in melanoma. Cell Rep Med. 2022;3:100500. https://doi.org/10.1016/j.xcrm.2021.100500.
48. Sahu A, Wang X, Munson P, Klomp JPG, Wang X, Gu SS, et al. Discovery of targets for immune-metabolic antitumor drugs identifies estrogen-related receptor alpha. Cancer Discov. 2023;13:672–701. https://doi.org/10.1158/2159-8290.CD-22-0244.
49. Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J, et al. Pooled CRISPR screening with single-cell transcriptome readout. Nat Methods. 2017;14:297–301. https://doi.org/10.1038/nmeth.4177.
50. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016;34:184–91. https://doi.org/10.1038/nbt.3437.
51. Sanson KR, Hanna RE, Hegde M, Donovan KF, Strand C, Sullender ME, et al. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nat Commun. 2018;9:5416. https://doi.org/10.1038/s41467-018-07901-8.
52. Replogle JM, Norman TM, Xu A, Hussmann JA, Chen J, Cogan JZ, et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. Nat Biotechnol. 2020;38:954–61. https://doi.org/10.1038/s41587-020-0470-y.
53. Xiao T, Li W, Wang X, Xu H, Yang J, Wu Q, et al. Estrogen-regulated feedback loop limits the efficacy of estrogen receptor-targeted breast cancer therapy. Proc Natl Acad Sci U S A. 2018;115:7869–78. https://doi.org/10.1073/pnas.1722617115.

54. Gogleva A, Polychronopoulos D, Pfeifer M, Poroshin V, Ughetto M, Martin MJ, et al. Knowledge graph-based recommendation framework identifies drivers of resistance in EGFR mutant non-small cell lung cancer. Nat Commun. 2022;13:1667. https://doi.org/10.1038/s41467-022-29292-7.
55. Kolde R, Laur S, Adler P, Vilo J. Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics. 2012;28:573–80. https://doi.org/10.1093/bioinformatics/btr709.
56. Glymour C, Zhang K, Spirtes P, Frontiers. Review of causal discovery methods based on graphical models. Front Genet. 2019. https://doi.org/10.3389/fgene.2019.00524.
57. Robinson JL, Kocabaş P, Wang H, Cholley P-E, Cook D, Nilsson A, et al. An atlas of human metabolism. Sci Signal. 2020;13:eaaz1482. https://doi.org/10.1126/scisignal.aaz1482.
58. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, et al. Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. PLoS Comput Biol. 2009;5:e1000489. https://doi.org/10.1371/journal.pcbi.1000489.
59. Hrovatin K, Fischer DS, Theis FJ. Toward modeling metabolic state from single-cell transcriptomics. Mol Metab. 2021;57:101396. https://doi.org/10.1016/j.molmet.2021.101396.
60. Ortmayr K, Dubuis S, Zampieri M, Nature Publishing Group. Metabolic profiling of cancer cells reveals genome-wide crosstalk between transcriptional regulators and metabolism. Nat Commun. 2019;10:1841. https://doi.org/10.1038/s41467-019-09695-9.
61. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019;9:5233. https://doi.org/10.1038/s41598-019-41695-z.
62. Chang Z, Xu Y, Dong X, Gao Y, Wang C. Single-cell and spatial multiomic inference of gene regulatory networks using SCRIPro. Bioinformatics. 2024. https://doi.org/10.1093/bioinformatics/btae466.
63. Qin Q, Fan J, Zheng R, Wan C, Mei S, Wu Q, et al. Lisa: inferring transcriptional regulators through integrative modeling of public chromatin accessibility and ChIP-seq data. Genome Biol. 2020;21:32. https://doi.org/10.1186/s13059-020-1934-6.
64. Browaeys R, Saelens W, Saeys Y. Nichenet: modeling intercellular communication by linking ligands to target genes. Nat Methods. 2020;17:159–62. https://doi.org/10.1038/s41592-019-0667-5.
65. Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. Cell. 2017;171:1611-1624.e24. https://doi.org/10.1016/j.cell.2017.10.044.
66. Le TD, Hoang T, Li J, Liu L, Liu H. A fast PC algorithm for high dimensional causal discovery with multi-core PCs. IEEE ACM Trans Comput Biol Bioinform. 2019;16:1483–95. https://doi.org/10.1109/TCBB.2016.2591526.
67. Wen Y, Huang J, Guo S, Elyahu Y, Monsonego A, Zhang H, et al. Applying causal discovery to single-cell analyses using CausalCell. Elife. 2023;12:e81464. https://doi.org/10.7554/eLife.81464.
68. Bravo González-Blas C, De Winter S, Hulselmans G, Hecker N, Matetovici I, Christiaens V, et al. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. Nat Methods. 2023;20:1355–67. https://doi.org/10.1038/s41592-023-01938-4.
69. Noor A, Ahmad A, Serpedin E. SparseNCA: sparse network component analysis for recovering transcription factor activities with incomplete prior information. IEEE ACM Trans Comput Biol Bioinform. 2018;15:387–95. https://doi.org/10.1109/TCBB.2015.2495224.
70. Wang C, Sun D, Huang X, Wan C, Li Z, Han Y, et al. Integrative analyses of single-cell transcriptome and regulome using MAESTRO. Genome Biol. 2020;21:198. https://doi.org/10.1186/s13059-020-02116-x.
71. Han Y, Zhang L, Sun D, Cao G, Wang Y, Yue J, et al. Spatiotemporal analyses of the pan-cancer single-cell landscape reveal widespread profibrotic ecotypes associated with tumor immunity. Nat Cancer. Nature Publishing Group; 2025;1–19. https://doi.org/10.1038/s43018-025-01039-5.
72. Patterson-Cross RB, Levine AJ, Menon V. Selecting single cell clustering parameter values using subsampling-based robustness metrics. BMC Bioinformatics. 2021;22:39. https://doi.org/10.1186/s12859-021-03957-4.
73. Zappia L, Oshlack A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. Gigascience. 2018;7:giy083. https://doi.org/10.1093/gigascience/giy083.
74. Liu B, Li C, Li Z, Wang D, Ren X, Zhang Z, et al. An entropy-based metric for assessing the purity of single cell populations. Nat Commun. 2020;11:3155. https://doi.org/10.1038/s41467-020-16904-3.
75. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1:417–25. https://doi.org/10.1016/j.cels.2015.12.004.
76. Chen H, Albergante L, Hsu JY, Lareau CA, Lo Bosco G, Guan J, et al. Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. Nat Commun. 2019;10:1903. https://doi.org/10.1038/s41467-019-09670-4.
77. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods. 2017;14:979–82. https://doi.org/10.1038/nmeth.4402.
78. Therneau TM. survival: Survival Analysis. 2024. https://cran.r-project.org/web/packages/survival/index.html.
79. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Yang T-HO, et al. The immune landscape of cancer. Immunity. 2018;48:812-830.e14. https://doi.org/10.1016/j.immuni.2018.03.023.
80. Hohenstein S. hohenstein/remef. 2024. https://github.com/hohenstein/remef.
81. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med. 2018;24:1550–8. https://doi.org/10.1038/s41591-018-0136-1.
82. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77. https://doi.org/10.1186/1471-2105-12-77.
83. Liu G, Wang H, Ran R, Wang Y, Li Y. *FOSL1* transcriptionally regulates *PHLDA2* to promote 5-FU resistance in colon cancer cells. Pathol Res Pract. 2023;246:154496. https://doi.org/10.1016/j.prp.2023.154496.
84. Taylor CT, Scholz CC, Nature Publishing Group. The effect of HIF on metabolism and immunity. Nat Rev Nephrol. 2022;18:573–87. https://doi.org/10.1038/s41581-022-00587-8.
85. Zhang C, Yue C, Herrmann A, Song J, Egelston C, Wang T, et al. STAT3 activation-induced fatty acid oxidation in CD8+ T effector cells is critical for obesity-promoted breast tumor growth. Cell Metab. 2020;31(1):148-161.e5. https://doi.org/10.1016/j.cmet.2019.10.013.
86. Vallée A, Lecarpentier Y, Vallée J-N. The key role of the WNT/β-catenin pathway in metabolic reprogramming in cancers under normoxic conditions. Cancers. 2021;13:5557. https://doi.org/10.3390/cancers13215557.
87. Dong Y, Tu R, Liu H, Qing G, Nature Publishing Group. Regulation of cancer cell metabolism: oncogenic MYC in the driver's seat. Signal Transduct Target Ther. 2020;5:1–11. https://doi.org/10.1038/s41392-020-00235-2.
88. Su A, Ling F, Vaganay C, Sodaro G, Benaksas C, Dal Bello R, et al. The folate cycle enzyme MTHFR is a critical regulator of cell response to MYC-targeting therapies. Cancer Discov. 2020;10:1894–911. https://doi.org/10.1158/2159-8290.CD-19-0970.
89. Celardo I, Lehmann S, Costa AC, Loh SH, Miguel Martins L. dATF4 regulation of mitochondrial folate-mediated one-carbon metabolism is neuroprotective. Cell Death Differ. 2017;24:638–48. https://doi.org/10.1038/cdd.2016.158.
90. Sun D, Wang J, Han Y, Dong X, Ge J, Zheng R, et al. TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. Nucleic Acids Res. 2020;49:D1420–30. https://doi.org/10.1093/nar/gkaa1020.
91. Li Z, Sun C, Qin Z, Ivyspring International Publisher. Metabolic reprogramming of cancer-associated fibroblasts and its effect on cancer cell reprogramming. Theranostics. 2021;11:8322–36. https://doi.org/10.7150/thno.62378.
92. Martínez-Reyes I, Chandel NS. Cancer metabolism: looking forward. Nat Rev Cancer. 2021;21:669–80. https://doi.org/10.1038/s41568-021-00378-6.
93. Sahai E, Astsaturov I, Cukierman E, DeNardo DG, Egeblad M, Evans RM, et al. A framework for advancing our understanding of cancer-associated fibroblasts. Nat Rev Cancer. 2020;20:174–86. https://doi.org/10.1038/s41568-019-0238-1.
94. Santi A, Kugeratski FG, Zanivan S. Cancer associated fibroblasts: the architects of stroma remodeling. Proteomics. 2018;18:e1700167. https://doi.org/10.1002/pmic.201700167.
95. Nie JZ, Wang M-T, Nie D. Regulations of tumor microenvironment by prostaglandins. Cancers. 2023;15:3090. https://doi.org/10.3390/cancers15123090.
96. Paoli A, Cenci L, Pompei P, Sahin N, Bianco A, Neri M, et al. Effects of two months of very low carbohydrate ketogenic diet on body composition, muscle strength, muscle area, and blood parameters in competitive natural body builders. Nutrients. 2021;13:374. https://doi.org/10.3390/nu13020374.
97. Li K, Zhao J, Wang M, Niu L, Wang Y, Li Y, et al. The roles of various prostaglandins in fibrosis: a review. Biomolecules. 2021;11:789. https://doi.org/10.3390/biom11060789.

98. Franks I. Prostaglandin E2 inhibits fibroblast migration in intestinal wound healing. Nat Rev Gastroenterol Hepatol. 2010;7:591–591. https://doi.org/10.1038/nrgastro.2010.164.
99. Zanotelli MR, Zhang J, Reinhart-King CA. Mechanoresponsive metabolism in cancer cell migration and metastasis. Cell Metab. 2021;33:1307–21. https://doi.org/10.1016/j.cmet.2021.04.002.
100. Koendjbiharie JG, van Kranenburg R, Kengen SWM. The PEP-pyruvate-oxaloacetate node: variation at the heart of metabolism. FEMS Microbiol Rev. 2020;45:fuaa061. https://doi.org/10.1093/femsre/fuaa061.
101. Court SJ, Waclaw B, Allen RJ. Lower glycolysis carries a higher flux than any biochemically possible alternative. Nat Commun. 2015;6:8427. https://doi.org/10.1038/ncomms9427.
102. Sullivan WJ, Mullen PJ, Schmid EW, Flores AA, Momcilovic M, Sharpley MS, et al. Extracellular matrix remodeling regulates glucose metabolism through TXNIP destabilization. Cell. 2018;175:117-132.e21. https://doi.org/10.1016/j.cell.2018.08.017.
103. Wieboldt R, Läubli H. Glycosaminoglycans in cancer therapy. Am J Physiol-Cell Physiol. 2022;322:C1187-200. https://doi.org/10.1152/ajpcell.00063.2022.
104. Roy A, Bera S. CAF cellular glycolysis: linking cancer cells with the microenvironment. Tumor Biol. 2016;37:8503–14. https://doi.org/10.1007/s13277-016-5049-3.
105. Gentric G, Mechta-Grigoriou F. Tumor cells and cancer-associated fibroblasts: an updated metabolic perspective. Cancers. 2021;13:399. https://doi.org/10.3390/cancers13030399.
106. El-Botty R, Morriset L, Montaudon E, Tariq Z, Schnitzler A, Bacci M, et al. Oxidative phosphorylation is a metabolic vulnerability of endocrine therapy and palbociclib resistant metastatic breast cancers. Nat Commun. 2023;14:4221. https://doi.org/10.1038/s41467-023-40022-5.
107. Qi J, Sun H, Zhang Y, Wang Z, Xun Z, Li Z, et al. Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer. Nat Commun. 2022;13:1742. https://doi.org/10.1038/s41467-022-29366-6.
108. Chu T, Wang Z, Pe'er D, Danko CG. Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. Nat Cancer. 2022;3:505–17. https://doi.org/10.1038/s43018-022-00356-3.
109. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12:453–7. https://doi.org/10.1038/nmeth.3337.
110. Hu C, Pang B, Lin G, Zhen Y, Yi H, Nature Publishing Group. Energy metabolism manipulates the fate and function of tumour myeloid-derived suppressor cells. Br J Cancer. 2020;122:23–9. https://doi.org/10.1038/s41416-019-0644-x.
111. Wang MM, Coupland SE, Aittokallio T, Figueiredo CR, Nature Publishing Group. Resistance to immune checkpoint therapies by tumour-induced T-cell desertification and exclusion: key mechanisms, prognostication and new therapeutic opportunities. Br J Cancer. 2023;129:1212–24. https://doi.org/10.1038/s41416-023-02361-4.
112. Reina-Campos M, Heeg M, Kennewick K, Mathews IT, Galletti G, Luna V, et al. Metabolic programs of T cell tissue residency empower tumour immunity. Nature. 2023;621:179–87. https://doi.org/10.1038/s41586-023-06483-w.
113. Vitale I, Manic G, Coussens LM, Kroemer G, Galluzzi L. Macrophages and metabolism in the tumor microenvironment. Cell Metab. 2019;30:36–50. https://doi.org/10.1016/j.cmet.2019.06.001.
114. Liu Y, Xu R, Gu H, Zhang E, Qu J, Cao W, et al. Metabolic reprogramming in macrophage responses. Biomark Res. 2021;9:1. https://doi.org/10.1186/s40364-020-00251-y.
115. Le Bel M, Brunet A, Gosselin J. Leukotriene B4, an endogenous stimulator of the innate immune response against pathogens. J Innate Immun. 2014;6:159–68. https://doi.org/10.1159/000353694.
116. Bird L, Nature Publishing Group. Homing in on leukotrienes. Nat Rev Immunol. 2003;3:777–777. https://doi.org/10.1038/nri1218.
117. Wei J, Gronert K. Eicosanoid and specialized proresolving mediator regulation of lymphoid cells. Trends Biochem Sci. 2019;44:214–25. https://doi.org/10.1016/j.tibs.2018.10.007.
118. Bryan AM, Del Poeta M, Luberto C. Sphingolipids as regulators of the phagocytic response to fungal infections. Mediators Inflamm. 2015;2015:640540. https://doi.org/10.1155/2015/640540.
119. Lee M, Lee SY, Bae Y-S, Nature Publishing Group. Functional roles of sphingolipids in immunity and their implication in disease. Exp Mol Med. 2023;55:1110–30. https://doi.org/10.1038/s12276-023-01018-9.
120. Mehendale N, Mallik R, Kamat SS. Mapping sphingolipid metabolism pathways during phagosomal maturation. ACS Chem Biol. 2021;16:2757–65. https://doi.org/10.1021/acschembio.1c00393.
121. Fairley LH, Lai KO, Wong JH, Chong WJ, Vincent AS, D'Agostino G, et al. Mitochondrial control of microglial phagocytosis by the translocator protein and hexokinase 2 in Alzheimer's disease. Proc Natl Acad Sci U S A. 2023;120:e2209177120. https://doi.org/10.1073/pnas.2209177120.
122. Harel M, Ortenberg R, Varanasi SK, Mangalhara KC, Mardamshina M, Markovits E, et al. Proteomics of melanoma response to immunotherapy reveals mitochondrial dependence. Cell. 2019;179:236-250.e18. https://doi.org/10.1016/j.cell.2019.08.012.
123. Tian J, Luo J, Zeng X, Ke C, Wang Y, Liu Z, et al. Targeting oxidative phosphorylation to increase the efficacy of immune-combination therapy in renal cell carcinoma. J Immunother Cancer. 2024;12:e008226. https://doi.org/10.1136/jitc-2023-008226.
124. Pergamon. The emerging role of Nrf2 in mitochondrial function. Free Radic Biol Med. 2015;88:179–88. https://doi.org/10.1016/j.freeradbiomed.2015.04.036.
125. Matsumoto S, Uchiumi T, Tanamachi H, Saito T, Yagi M, Takazaki S, et al. Ribonucleoprotein Y-box-binding protein-1 regulates mitochondrial oxidative phosphorylation (OXPHOS) protein expression after serum stimulation through binding to OXPHOS mRNA. Biochem J. 2012;443:573–84. https://doi.org/10.1042/BJ20111728.
126. Chan T, Chen Y, Tan KT, Wu C, Wu W, Li W, et al. Biological significance of MYC and CEBPD coamplification in urothelial carcinoma: multilayered genomic, transcriptional and posttranscriptional positive feedback loops enhance oncogenic glycolysis. Clin Transl Med. 2021;11:e674. https://doi.org/10.1002/ctm2.674.
127. Liu T, Wen Z, Shao L, Cui Y, Tang X, Miao H, et al. ATF4 knockdown in macrophage impairs glycolysis and mediates immune tolerance by targeting HK2 and HIF-1α ubiquitination in sepsis. Clin Immunol. 2023;254:109698. https://doi.org/10.1016/j.clim.2023.109698.
128. Pierdomenico M, Palone F, Cesi V, Vitali R, Mancuso AB, Cucchiara S, et al. Transcription factor ZNF281: a novel player in intestinal inflammation and fibrosis. Front Immunol. 2018;9:2907. https://doi.org/10.3389/fimmu.2018.02907.
129. Varga T, Czimmerer Z, Nagy L. PPARs are a unique set of fatty acid regulated transcription factors controlling both lipid metabolism and inflammation. Biochim Biophys Acta. 2011;1812:1007–22. https://doi.org/10.1016/j.bbadis.2011.02.014.
130. Elsevier. The Nrf2 transcription factor: a multifaceted regulator of the extracellular matrix. Matrix Biol Plus. 2021;10:100057. https://doi.org/10.1016/j.mbplus.2021.100057.
131. Lisignoli G, Toneguzzi S, Piacentini A, Cristino S, Grassi F, Cavallo C, et al. CXCL12 (SDF-1) and CXCL13 (BCA-1) chemokines significantly induce proliferation and collagen type I expression in osteoblasts from osteoarthritis patients. J Cell Physiol. 2006;206:78–85. https://doi.org/10.1002/jcp.20435.
132. Choi YE, Song MJ, Hara M, Imanaka-Yoshida K, Lee DH, Chung JH, et al. Effects of tenascin C on the integrity of extracellular matrix and skin aging. Int J Mol Sci. 2020;21:8693. https://doi.org/10.3390/ijms21228693.
133. Honda F, Tsuboi H, Ono Y, Abe S, Takahashi H, Ito K, et al. Pathogenic roles and therapeutic potential of the CCL8–CCR8 axis in a murine model of IgG4-related sialadenitis. Arthritis Res Ther. 2021;23:214. https://doi.org/10.1186/s13075-021-02597-6.
134. Waddington KE, Robinson GA, Rubio-Cuesta B, Chrifi-Alaoui E, Andreone S, Poon K-S, et al. LXR directly regulates glycosphingolipid synthesis and affects human CD4+ T cell function. Proc Natl Acad Sci USA. 2021;118:e2017394118. https://doi.org/10.1073/pnas.2017394118.
135. Lai H-Y, Hsu L-W, Tsai H-H, Lo Y-C, Yang S-H, Liu P-Y, et al. CCAAT/enhancer-binding protein delta promotes intracellular lipid accumulation in M1 macrophages of vascular lesions. Cardiovasc Res. 2017;113:1376–88. https://doi.org/10.1093/cvr/cvx134.
136. Maceyka M, Spiegel S, Nature Publishing Group. Sphingolipid metabolites in inflammatory disease. Nature. 2014;510:58–67. https://doi.org/10.1038/nature13475.
137. Uranbileg B, Isago H, Sakai E, Kubota M, Saito Y, Kurano M, et al. Alzheimer's disease manifests abnormal sphingolipid metabolism. Front Aging Neurosci. 2024. https://doi.org/10.3389/fnagi.2024.1368839.
138. Marabelle A, Fakih M, Lopez J, Shah M, Shapira-Frommer R, Nakagawa K, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. Lancet Oncol. 2020;21:1353–65. https://doi.org/10.1016/S1470-2045(20)30445-9.

139. Litchfield K, Reading JL, Puttick C, Thakkar K, Abbosh C, Bentham R, et al. Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. Cell. 2021;184:596-614.e14. https://doi.org/10.1016/j.cell.2021.01.002.
140. Wu C-C, Wang YA, Livingston JA, Zhang J, Futreal PA, Nature Publishing Group. Prediction of biomarkers and therapeutic combinations for anti-PD-1 immunotherapy using the global gene network association. Nat Commun. 2022;13:42. https://doi.org/10.1038/s41467-021-27651-4.
141. Corsello SM, Nagari RT, Spangler RD, Rossen J, Kocak M, Bryan JG, et al. Discovering the anticancer potential of non-oncology drugs by systematic viability profiling. Nat Cancer. 2020;1:235–48. https://doi.org/10.1038/s43018-019-0018-6.
142. Ellinghaus P, Heisler I, Unterschemmann K, Haerter M, Beck H, Greschat S, et al. <scp>BAY</scp>87-2243, a highly potent and selective inhibitor of hypoxia-induced gene activation has antitumor activities by inhibition of mitochondrial complex I. Cancer Med. 2013;2:611–24. https://doi.org/10.1002/cam4.112.
143. Tejada-Lapuerta A, Bertin P, Bauer S, Aliee H, Bengio Y, Theis FJ. Causal machine learning for single-cell genomics. Nat Genet. 2025;57:797–808. https://doi.org/10.1038/s41588-025-02124-2.
144. Li L, Xia R, Chen W, Zhao Q, Tao P, Chen L. Single-cell causal network inferred by cross-mapping entropy. Brief Bioinform. 2023. https://doi.org/10.1093/bib/bbad281.
145. Lopez R, Tagasovska N, Ra S, Cho K, Pritchard JK, Regev A. Learning causal representations of single cells via sparse mechanism shift modeling. arXiv. 2023. https://doi.org/10.48550/arXiv.2211.03553.
146. Tang K, Han Y, Sun D, Dong X, Wu Q, Wang C. Reference-guided computational framework identifies microenvironment metabolic subtypes and targets using pan-cancer single-cell datasets. Github. 2025. https://github.com/wanglabtongji/MetroTIME
147. Tang K, Han Y, Sun D, Dong X, Wu Q, Wang C. Reference-guided computational framework identifies microenvironment metabolic subtypes and targets using pan-cancer single-cell datasets. Github. 2025. https://github.com/wanglabtongji/MetroSCREEN

Publisher's Note

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