Metabolic landscape of the tumor microenvironment

Zhengtao Xiao, Ziwei Dai*, Jason W. Locasale*

Department of Pharmacology and Cancer Biology, Duke University School of Medicine, Durham, NC 27710, USA

***Corresponding author:** Jason W. Locasale. Email: jason.locasale@duke.edu

Ziwei Dai. Email: ziwei.dai@duke.edu

Abstract

The tumor milieu consists of numerous cell types with each cell existing in a different nutritional environment. However, a characterization of intratumoral metabolic heterogeneity at the single-cell level in human cancer is not established. Here, we develop a computational pipeline to analyze metabolic gene expression programs of single cells within human tumors. In two representative cancer types, melanoma and head and neck, we apply this algorithm to define the single-cell metabolic landscape of human tumors. We find that malignant cells in general have higher metabolic activity and higher metabolic variation than previously observed from studies of bulk tumor comparisons. Indeed, most of the observed metabolic variation of single tumor and normal cells were found to be inconsistent with comparisons with bulk tumor samples. Variation in the expression of mitochondrial programs is the major contributor to intratumoral metabolic heterogeneity. Surprisingly, the expression of both glycolytic and mitochondrial programs strongly correlates with hypoxia in almost all cell types. Immune and stromal cells could also be distinguished by their metabolic features. Taken together this analysis establishes a computational framework for characterizing metabolism using single cell expression data and defines principles of intratumoral metabolic heterogeneity.

Introduction

Metabolic reprogramming in cancer cells supports specific demands for energy, biomass, redox maintenance, and cellular communication[1]. Cellular metabolism is shaped by both genetic and environmental factors including the somatic driver mutations selected during tumor evolution, the tissue of origin, and the local nutrient environment[1-5]. Given the importance of environmental factors, tumor metabolism exhibits grossly different properties in laboratory cell culture settings than *in vivo*. Also, differences in the metabolism of humans and model organisms are also substantial. Nevertheless, most conclusions surrounding tumor metabolism have been obtained using cell culture models and model organisms and the number of direct observations of cellular metabolism are few. Indeed, nearly all observations of human tumor metabolism in vivo have been conducted using measurements obtained from bulk tumors. These findings have advanced our understanding of tumor metabolism tremendously. Nevertheless, they do result from a population average over the genetic and environmental variables of each cell.

There are numerous sources of intratumoral heterogeneity. Tumor cells exist within a microenvironment consisting of stromal cells such as cancer-associated fibroblasts (CAFs), immune cells, endothelial cells and many others. Each of these cell types takes an active role in tumor cell proliferation. For example, CAFs may release cytokines and growth factors that are received by and function to signal in cancer cells[6]. The immune compartment of an established malignant tumor is collectively immunosuppressive[7]. Endothelial cells provide vasculature to provide nutritional support in challenging environments[8]. Each of the cells has unique metabolic demands that enable specific function. In addition to the unique metabolic demands of each cell type, each cancer experiences a distinct nutrient environment, distinct engagement of extracellular signals, and may derive from a different cell of origin thus

possibly having distinct mutational patterns[9]. Therefore at the cellular level, each cell within the tumor is likely to have a different metabolic status[10-12]. Nevertheless, direct observations of cellular metabolism in vivo at the level of single cells is difficult. Most conclusions about the nature of the tumor microenvironment have relied on in vitro models such as co-culture systems[13-15] or measurements of single variables such as immunohistochemical staining for the expression of a metabolic enzyme[16].

Metabolism and its associated phenotypic biology are governed by the concentrations of metabolites and the rates or fluxes by which one metabolite is converted to another. A comprehensive understanding of metabolism requires knowledge of both concentrations and fluxes. These measurements are difficult to obtain in humans and haves so far been exclusively conducted in bulk tumors. Global gene expression however is readily measurable and the advent of single cell sequencing technologies enables expression profiling of individual cells within entire tissues or tumors[17-22]. It is also an indirect means of assessing metabolism. Nevertheless, gene expression has provided many insights into metabolic pathway activity and in many documented instances the gene expression is predictive of metabolic flux[23,24]. Thus single cell sequencing could provide some insight into metabolism at the single cell level in human tumors.

In this study, we analyze metabolic gene expression profiles of more than 9,000 single cells from two representative human tumor types including melanoma[19] and squamous cell carcinoma of the head and neck (HNSCC)[21]. We find that activities of metabolic pathways in malignant cells are in general more active and plastic than those in non-malignant cells, and the metabolic features of single cancer cells are poorly captured by measurements done with bulk tumors. Variation in mitochondrial activity is the major contributor to the metabolic heterogeneities among both malignant cells and non-malignant cells, and, strikingly, the

activities of glycolysis and oxidative phosphorylation both correlate with hypoxia at the single-cell level. We also identify metabolic features of different immune and stromal cell subtypes and find patterns distinct from behaviors of these cells in ex vivo culture conditions. These findings begin to unravel principles of how malignant transformation affects the metabolic phenotypes of tumor and non-tumor cells within the tumor microenvironment.

Results

Landscape of metabolic gene expression at the single-cell level

We first developed a computational pipeline for analyzing metabolic gene expression profiles at the single-cell level (Fig. 1a, Methods). Briefly, we applied missing data imputation and data normalization to gene expression profiles to account for the influence of technical noise. We then characterized the global structure of single cell metabolic programs using clustering analysis, identified cell type-specific metabolic programs using quantitative metrics we developed, and designed algorithms for quantitation of metabolic heterogeneity of malignant and non-malignant cells. We applied this pipeline to two scRNA-seq datasets for human melanoma[19] and squamous cell carcinoma of the head and neck (HNSCC)[21], which include an expansive set of gene expression of 4,054 cells and 5,502 cells respectively (Methods). These datasets were selected because they covered the largest numbers of cells and included detailed annotation of the cell types, while many of the currently available scRNA-seq datasets for human cancer are limited by cell number and sequencing depth. Each dataset covers both malignant and non-malignant cells isolated from patient-derived human tumors with different genotypic and phenotypic backgrounds (Fig. 1b, 1c), thus enabling an in-depth investigation of the expression of metabolic genes and pathways in each cell type.

We first analyzed the global structure of metabolic gene expression profiles in these two datasets using t-distributed stochastic neighbor embedding (t-SNE)[25] based on expression levels of 1566 metabolic genes (Methods). Imputation of zero values was performed before applying t-SNE to correct for the influence of the high frequency of dropout events (i.e. failure in detecting expressed genes due to low sequencing depth) in certain cell types (Methods). The imputation effectively reduced the dropout rates in all cell types (Supplementary Fig. 1a-d) without changing the pattern of the metabolic gene expression (Supplementary Fig. 1e, f). Clustering analysis after dimensionality reduction with t-SNE showed that metabolic gene expression profiles of the malignant cells formed distinct clusters that corresponded to their tumors of origin (i.e. from which tumor the cell was derived) for both melanoma (Fig. 1d) and HNSCC (Fig. 1e), suggesting that metabolic gene expression in malignant cells is largely set by patient-specific factors. This was further corroborated by the higher correlation coefficients of metabolic gene expression profiles between malignant cells from tumors within the same patient (average Spearman's correlation = 0.91) than those between cells from tumors from different patients (average Spearman's correlation = 0.79, Fig. 1f, g). Cells from patients of the same genotypic background also show higher similarity than those from different genotypic backgrounds (average Spearman's correlation = 0.87 compared to 0.79 for melanoma, 0.87 compared to 0.77 for HNSCC, Fig. 1f, g). In contrast, metabolic gene expression profiles of non-malignant cells showed no distinguishable differences between patients (average difference between intratumoral and intertumoral Spearman's correlation = 0.01 for non-malignant cells in melanoma and 0.007 for those in HNSCC, Fig. 1h, i), indicating that metabolism of these normal cells in the tumor microenvironment exhibits no observable interpatient heterogeneity. Similar trends were also observed for the distributions of correlation coefficients of metabolic gene expression between cells from the same patient or different patients, in which only malignant cells

showed significantly stronger intratumoral correlation than intertumoral correlation (Supplementary Fig. 2). Taken together, these results suggest that malignant cells in general exhibit higher metabolic plasticity which likely leads to patient-specific metabolic reprogramming of cancer cells but not the supporting cells in the tumor microenvironment.

Cell type-specific metabolic reprogramming in the tumor microenvironment

We next sought to identify the overall features of metabolic pathway variation among the different cell types, especially between malignant and non-malignant cells. To quantify the activity of a metabolic pathway, we developed a pathway activity score that we defined as the relative gene expression value averaged over all genes in this pathway and all cells of this type (Fig. 2a, b, Methods). To enable comparison between different cell types, we normalized the gene expression values using a deconvolution method[26] which resulted in the highest similarity of normalized expression value distributions between cell types among four commonly used data normalization methods (Supplementary Fig. S3, Methods). Among the 80 metabolic pathways with at least 5 genes included, over 70 were more highly expressed (pathway activity score >1 and permutation test p-value <0.01) in at least one cell type compared to other cell types, indicating that the activities of most metabolic pathways are determined by cell type. To further characterize the cell type specific metabolic features, we grouped the metabolic pathways into 11 categories based on KEGG classifications (Supplementary Fig. 4a, b), and assessed enrichment of each category in pathways up-regulated in each cell type using a one-sided Fisher's exact test. These categories reflected different aspects of cellular metabolism such as carbohydrate metabolism, amino acid metabolism and nucleotide metabolism. To our surprise, most categories (10 out of 11 for melanoma and 8 out of 11 for HNSCC) were not enriched in pathways up-regulated in a specific cell type (*i.e.* one-sided Fisher's exact p-value > 0.05 for all cell types,

Supplementary Fig. 4c, d), indicating that each cell type undergoes global up- or down-regulation of metabolic pathways in all categories. Malignant cells had the largest number of metabolic pathways significantly up-regulated in both melanoma (37 pathways up-regulated in malignant cells compared to 23 in cancer-associated fibroblasts, 16 in macrophages and less than 10 in all other cell types, Fig. 2a) and HNSCC (51 pathways up-regulated in malignant cells compared to 27 in fibroblasts and less than 10 in all other cell types, Fig. 2b), and the up-regulated pathways included many different parts of cellular metabolism such as central carbon metabolism, one carbon metabolism, methionine metabolism, steroid biosynthesis and beta-alanine metabolism. Thus, compared to non-malignant cells, malignant cells undergo a global up-regulation of metabolic activity. Interestingly, comparison of pathway activities between melanoma and HNSCC in the cell types shared by the two datasets (Supplementary Fig. 4e) revealed high concordance between the two tumor types in the metabolic features of endothelial cells (Pearson's R=0.49, p-value=1.9e-5), macrophages (Pearson's R=0.75, p-value=9.2e-14) and fibroblasts (Pearson's R=0.45, p-value=9.3e-5), while metabolic pathway activities of B cells, T cells and malignant cells correlated poorly between the two tumor types (Pearson's R<0.1, p-value>0.05 for all three cell types), suggesting that metabolism of these cell types is more sensitive to environmental factors and the effector status of these immune cells.

To evaluate whether patterns of pathway activities in single malignant cells were consistent with the more commonly used transcriptomic profiling of bulk tissue samples, we computed pathway activity scores based on RNA-seq data for bulk tumor samples and matched normal tissue samples from The Cancer Genome Atlas (TCGA)[27] and compared the results to our analysis in single malignant cells. Since there are no matched normal samples for melanoma in TCGA, here we only considered the HNSCC dataset which contains 43 paired tumor and normal samples[28]. We found 7 pathways up-regulated and 34 pathways

down-regulated in tumor samples compared to normal samples (Permutation test p-value<0.01, Fig 2c). Among these pathways, only 13 showed a consistent pattern of activity in single malignant cells and bulk tumors (Fig. 2c). Notably, 23 out of the 51 pathways up-regulated in the single malignant cells were identified as down-regulated in tumors based on the bulk RNA-seq, and the pathway activities correlated poorly (Pearson's R=-0.2, p-value=0.04) between bulk tumors and single malignant cells (Fig. 2d). The discrepancy between single-cell and bulk RNA-seq in identifying tumor-associated metabolic pathway activities is likely due to the intrinsic heterogeneity in cellular composition of the tumors (Fig. 2e). Bulk RNA-seq measures the average expression levels over a mixture of different cell types thus masking the difference between cell types in the same sample. This was further supported by an analysis of the distributions of pathway activities in single cells and bulk samples showing higher variation of pathway activities between different types of single cells than between bulk tumors and normal tissues (average standard deviation of pathway activities = 0.39 for single cells compared to 0.14 for bulk samples, Fig. 2f). Taken together, these results reveal a global enhancement of metabolic activity in single malignant cells which can only be detected with gene expression profiling at the single cell level.

Intratumoral metabolic heterogeneity of malignant cells

In addition to the common routes of metabolic reprogramming in malignant cells, metabolism is also affected by a location-specific, fluctuating nutrient supply and interactions with other neighboring cells in space. It is thus intriguing to investigate what parts of cellular metabolism are impacted by these environmental factors. To identify major contributors to intratumoral metabolic heterogeneity of malignant cells (i.e. variation of metabolism among malignant cells from the same tumor), we performed principal component analysis (PCA) and gene set enrichment analysis (GSEA)[29] to identify metabolic pathways enriched in genes

explaining most of the variance among malignant cells in each tumor (Fig. 3a, Supplementary Fig. 5). We found that for both melanoma and HNSCC, oxidative phosphorylation (OXPHOS) was the top-scoring pathway in most tumors (Fig. 3b, c). Similarly, tricarboxylic acid cycle (TCA cycle) gene expression also showed a substantial contribution to metabolic heterogeneity in several tumors, indicating that variation in mitochondrial activity is the major contributor to intratumoral metabolic heterogeneity of malignant cells.

We next sought to investigate the coupling between mitochondrial activity and environmental factors such as oxygen and other nutritional supplies in single malignant cells. Since direct measurements of nutritional status of the individual cells are not available, we used the average expression level of a set of genes known to respond to hypoxia as a metric of oxygen supply which is an environmental factor known to have great impact on cellular metabolism (Methods). We used the average expression level of genes in OXPHOS as a surrogate of mitochondrial activity. We also considered glycolysis since it is another important pathway in supplying energy and material for cell survival and proliferation, and its relationship with OXPHOS in cancer metabolism is still a matter of interest. We found that activity of glycolysis and the hypoxia signature were highly correlated in both melanoma and HNSCC (Pearson's R=0.72 for melanoma and 0.52 for HNSCC, Fig. 3d, e), which agrees with previous studies showing that hypoxia increases glycolytic activity[30,31]. To our surprise, we also found that OXPHOS significantly correlated with glycolysis (Pearson's R=0.67 for melanoma and 0.53 for HNSCC, Fig. 3d, e) and the response to hypoxia (Pearson's R=0.49 for melanoma and 0.27 for HNSCC, Fig. 3d, e). To exclude the possibility that the positive correlations between pathways were driven by a few correlated genes with high variation in their expression levels while most other genes were not correlated, we confirmed that most pairwise correlation coefficients of expression levels between individual genes from the pathways were also positive (Supplementary Fig. S6). These results together

indicate that mitochondrial OXPHOS expression is the most important contributor to intratumoral metabolic heterogeneity, and OXPHOS is not mutually exclusive with glycolysis as routes for energy production in adapting to the tumor microenvironment with varying oxygen supply.

Metabolic heterogeneity of non-malignant cells

We next explored the heterogeneity of metabolic pathway activities in non-malignant cells. Since metabolic gene expression profiles of non-malignant cells were clustered according to cell type, and there were no significant differences between different tumors for the same cell type (Fig. 1h, j and Supplementary Fig. 2a, b). Thus we focused on identifying the major contributors to metabolic heterogeneity in each non-malignant cell type. For each non-malignant cell type, we repeated the PCA and GSEA analyses as we did for malignant cells and found that variations in OXPHOS also dominated the metabolic heterogeneity in all non-malignant cell types (Fig. 4a, b). In addition to OXPHOS, TCA cycle expression also substantially contributed to the metabolic heterogeneity of all non-malignant cell types. These results demonstrate that similar to the case of malignant cells, mitochondrial activity is also the major contributor to metabolic heterogeneity in non-malignant cells.

To evaluate how mitochondrial activity in non-malignant cells relates to oxygen availability, we correlated activities of OXPHOS and glycolysis with the hypoxia-associated feature in each cell type (Fig 4c, d). Similar to the case of malignant cells, glycolysis, OXPHOS and the hypoxia feature were also significantly correlated in almost all cell types except for macrophages in the HNSCC dataset, in which OXPHOS and hypoxia showed no significant correlation (Pearson's R= -0.02, p-value= 0.87). These findings, together with the results for malignant cells, challenge the long-standing concept that metabolic reprogramming of central carbon metabolism frequently takes the form of a switch between

glycolysis and mitochondrial respiration. At the single-cell level, it is plausible that cells facing more severe oxygen deprivation tend to up-regulate both glycolysis and mitochondrial OXPHOS, which may help cells to more effectively compete with other cells for limited resources. Such a positive correlation between the hypoxia signature and energy-producing pathways is conserved in almost all cell types included in the single-cell RNA-seq datasets.

Metabolic features of non-malignant cell subtypes

Non-malignant cells such as immune and stromal cells are important constituents of the tumor microenvironment. These cells are known to differentiate into subtypes with distinct roles, and this process involves metabolic reprogramming to satisfy their cell autonomous metabolic demands and enable interactions with other cell types[10,11,32,33]. Next, we use scRNA-seq to characterize the metabolic features of T cell and fibroblast subpopulations which together constitute the largest non-malignant cell populations in the melanoma and HNSCC datasets. T cells were first separated into $CD8^+$ and $CD4^+$ subtypes based on the expression of CD4 and CD8A[34] (Fig. 5a, Methods). The $CD4^+$ T cells were further classified into regulatory T cells (Tregs) and T helper cells (Ths) based on expression levels of FOXP3 and CD25 which are known to be specifically expressed in these specific cell types[35] (Fig. 5a, Methods). We then performed GSEA analysis to identify metabolic pathways enriched in each subtype. We found that OXPHOS was the most important metabolic pathway distinguishing T cell subtypes: $CD4^+$ T cells exhibited significantly higher levels of OXPHOS compared to $CD8^+$ T cells in both melanoma (GSEA p-value=0.002) and HNSCC (GSEA p-value=0, Fig. 5b, c). Interestingly, compared to Ths, Tregs exhibited up-regulation of glycolysis (GSEA p-value=0 for both tumors) in addition to OXPHOS (GSEA p-value=0 for both tumors, Fig. 5d, e). This appears to contradict with previous studies showing that among immune cells derived from healthy mice not bearing tumors, Ths tend to be more

glycolytic compared to Tregs[36]. On the other hand, the OXPHOS preference of $CD4^+$ T cells and Tregs is consistent with previous studies[37,38], highlighting enhanced mitochondrial oxidative metabolism as a universal metabolic feature of these T cell subtypes in different contexts. These results suggest that subpopulations of immune cells in the tumor microenvironment have metabolic features that differ from their behaviors in normal tissues.

We next characterized the metabolic features of subpopulations of fibroblasts which serve as a major component of tumor stroma and have diverse roles in both normal functions such as wound healing and tumor-promoting functions such as remodeling the extracellular matrix and interacting with tumor cells to support their growth[39]. According to a previous study[21], the 1,422 fibroblasts in the HNSCC dataset formed two major subpopulations exhibiting gene expression corresponding to cancer-associated fibroblasts (CAFs) or myofibroblasts, respectively (Fig. 5f, Methods). We thus performed GSEA analysis to compare metabolic gene expression between the two fibroblast subtypes (Fig. 5g, Supplementary Table 1). We found significant up-regulation of 15 metabolic pathways (GSEA p-value<0.05) that distinguished CAFs from myofibroblasts. On the other hand, inositol phosphate metabolism was the only metabolic pathway up-regulated in myofibroblasts (p-value=0), indicating that CAFs are more metabolically active compared to myofibroblasts. Glycolysis was significantly up-regulated in CAFs (GSEA p-value=0.048), which is in line with a hypothesized metabolic feature of CAFs, in which CAFs exhibit enhanced glycolysis which produces excess lactate that can be utilized by adjacent tumor cells to support growth[40,41]. Notably, we also found several other groups of metabolic pathways up-regulated in CAFs. These pathways included arachidonic acid metabolism (GSEA p-value=0) and linoleic acid metabolism (GSEA p-value=0.002), which are known to produce inflammatory mediators[42], and a group of pathways related to glycan biosynthesis and degradation. Up-regulation of these pathways in CAFs may support the function of CAFs

in secreting small molecule compounds and proteins to remodel the tumor microenvironment. Taken together, these results identify a metabolic phenotype of CAFs that potentially helps establish their roles in interacting with other cell types and modulating the tumor microenvironment.

Discussion

We have characterized the rewiring of metabolic pathways in single malignant cells compared to their normal partners in the same tumor microenvironment. We found that, compared to non-malignant cells, malignant cells not only exhibit high metabolic plasticity that allows them to adapt their metabolism to different genotypic and environmental contexts, but also follow a common pattern of global up-regulation of activities of metabolic pathways in almost all functional categories. These results point to the principle that metabolism of cancer cells is in general more flexible and active than that of non-malignant cells. Notably, most of the metabolic changes detected in single malignant cells compared to single non-malignant cells were not captured by comparing expression levels of metabolic genes between bulk tumor and normal samples, implying that comparison of metabolic configurations between tumors and normal tissues based on bulk measurements tends to underestimate the differences between malignant and non-malignant cells due to the highly complicated cellular composition of the bulk samples. Consistently, our results appear to be different from previous studies comparing metabolic network expression in tumor and normal tissues[43,44].

There are several interesting findings around the role of mitochondrial activity in shaping the metabolic heterogeneity of tumors. First, variation in activity of OXPHOS genes is the most important contributor to the metabolic heterogeneity among malignant cells from the same tumor and that among non-malignant cells of the same type. The high variation in

OXPHOS activity suggests that this pathway might be responsible for adapting to alterations in the environment. It is thus interesting to investigate how such variability in OXPHOS activity contributes to tumor progression.

The role of mitochondria (OXPHOS and TCA cycle) in cancer is still a matter of debate. In addition to the well-known Warburg effect[45], several studies comparing metabolic gene expression between bulk tumors and normal tissues have also identified suppression of OXPHOS as a recurrent metabolic phenotype in tumors[43,46-48]. However, there are also numerous studies showing that active OXPHOS activity is in fact required for cancer progression. Mitochondrial inhibitors such as metformin are known to suppress cancer cell growth[49-51]. In this study, we found that OXPHOS gene expression levels were in general higher in single malignant cells (Fig 2a, b), which appears to contradict observations based on bulk gene expression levels[43,46-48]. Further work is needed to resolve the discrepancy between single-cell and bulk RNA-seq in evaluating the role of OXPHOS in tumors, but it is likely due to the complexity of cellular composition of tumors that is almost impossible to be dissected by bulk measurements.

Another interesting finding about OXPHOS activity in single cells is that it is correlated with both glycolysis and response to hypoxia in almost all cell types. This at the first glance is counterintuitive because hypoxia activates signal transduction pathways that induce glycolysis and suppress OXPHOS and other mitochondria-associated pathways[31]. Nevertheless, OXPHOS also has established role in mediating the response to hypoxia by serving as a sensor of oxygen availability through stabilization of hypoxia-induced factors (HIF)[52,53]. Therefore, the interplay between glycolysis, OXPHOS, and hypoxia is highly dynamic in living cells, and the quantitative relationship between them is at least partly determined by the interaction between the inhibitory effects of the HIF signaling pathway and

the positive feedback from OXPHOS activity to HIF signaling in response to oxygen availability. Our analysis of single cell transcriptomic profiles clarifies that activities of these pathways tend to be positively correlated in single cells from the tumor microenvironment with scarce and fluctuating oxygen supply. Whether such coupling at single cell level exists in other types of tumor and whether it benefits the cancer cells needs further investigation.

With the gene expression profiles of single cells from tumors, we were able to identify metabolic features that distinguish subpopulations of immune and stromal cells. This approach has the advantage of providing a direct snapshot of the metabolic landscape of tumors and their microenvironment consisting of numerous known and unknown types of cells whose metabolism is greatly influenced by the interactions between them and the shortage of nutrients in the tumor microenvironment[11,54,55]. We found that some non-malignant cell subpopulations, Ths and Tregs for instance, adopt metabolic phenotypes distinct from what they show in *ex vivo* culture conditions[36]. Metabolic reprogramming of CAFs compared to myofibroblasts was also shown to involve more pathways than what is currently known. These results highlight the great impact of the tumor microenvironment on cellular metabolism. It is worth noting that currently the ability to characterize metabolic phenotypes of cell subpopulations is still limited by the number of single cells that can be profiled at the same time due to the diversity of cell types and noisy gene expression in single cells. Improvement in single cell omics techniques will help address this issue and provide higher resolution in identifying cell subpopulations with different metabolic phenotypes.

To summarize, this study offers a global picture of metabolic gene expression in single tumor and non-tumor cells from the highly complex tumor microenvironment. These cells display metabolic activities distinct from the average pattern of their metabolism at the bulk level. Although this study only focused on two tumor types consisting of the highest

quality data at this time that allowed for the current scope of analysis, the principles about metabolic landscape of single cells in the tumors – the metabolic plasticity and activity of malignant cells, the dominant role of mitochondrial programs in shaping metabolic heterogeneity of malignant and non-malignant cells, and the metabolic features of immune cell subtypes - were applicable to both tumor types, and the data analysis pipeline that we developed here can easily be extended to datasets of other tumor types. With the rapid development of novel single-cell omics techniques and accumulation of data in more tumor types and patients, we are optimistic that a comprehensive portrait of metabolic features of every unique tumor cell will emerge in the near future.

Methods

Data acquisition and processing

Processed gene expression profiles for melanoma[19] and HNSCC[21] were retrieved from Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/gds) under accession numbers GSE72056 and GSE103322. Gene expression levels were quantified using the metric log2 (TPM +1). Tumors and non-malignant cell types containing less than 50 cells were excluded from the downstream analysis. Missing gene expression values were imputed using the scImpute algorithm[56] with default parameters and TPM values and gene lengths (for a gene associated with multiple transcripts, the length of the longest transcript was used) as the input. Imputation was only applied to genes with dropout rates (*i.e.* the fraction of cells in which the corresponding gene has zero expression value) larger than 50% to avoid over-imputation[56]. Lists of metabolic genes and pathways were obtained from the KEGG database (http://www.kegg.jp). The imputed expression values were then used in clustering analysis using the tSNE method[25] implemented in the Rtsne package[57] with default parameters. Bulk

RNA-seq data for matched HNSCC tumor and normal samples were downloaded from TCGA database (https://portal.gdc.cancer.gov/).

Evaluation of normalization methods

The relative log expression (RLE) method[58] was implemented using the 'estimateSizeFactorsForMatrix' function in DESeq2[59]. The trimmed mean of M-values (TMM)[60] and upper quartile[61] methods were performed using the 'calcNormFactors' function in the edgeR package[60]. For deconvolution normalization[26] for scRNA data with annotated cell type information, the 'computeSumFactors' function in the scran package was used to compute cell type specific size factors[62]. TPM values were transformed to read counts by multiplying TPM values and gene lengths (for genes with multiple transcripts, the length of the longest transcript was used). Normalized gene counts were computed by dividing read counts by the size factor corresponding to the cell and then transforming back to TPM by dividing the gene lengths. To avoid noise caused by low expressed and undetected genes, only genes with dropout rate < 0.75 (*i.e.* genes with non-zero expression levels in at least 25% of the cells) were used as the reference genes to do normalization. The distributions of relative gene expression values (defined in 'Calculation of pathway activity') in different cell types were used to evaluate the performances of these methods. The method minimizing differences in distributions of relative gene expression levels between cell types was selected for the following analysis.

Calculation of pathway activity

For the *i*-th metabolic gene, we first calculated its mean expression level across cells of the *j*-th cell type:

$$E_{i,j} = \frac{\sum_{k=1}^{n_j} g_{i,k}}{n_j}, i \in 1 \dots M, j \in 1 \dots N \quad (1)$$

In which n_j is the number of cells in the *j*-th cell type, $g_{i,k}$ is the expression level of the *i*-th gene in the *k*-th cell in this cell type, *M* is the number of metabolic genes, and *N* is the number of cell types. The relative expression level of the *i*-th gene in the *j*-th cell type was then defined as the ratio of $E_{i,j}$ to its average over all cell types:

$$r_{i,j} = \frac{E_{i,j}}{\frac{1}{N}\sum_{j}^{N} E_{i,j}} \quad (2)$$

Here $r_{i,j}$ quantifies the relative expression level of gene *i* in cell type *j* comparing to the average expression level of this gene in all cell types. A $r_{i,j}$ value larger than 1 means that expression level of gene *i* is higher in cell type *j* compared to its average expression level over all cell types. The pathway activity score for the *t*-th pathway and the *j*-th cell type was then defined as the weighted average of $r_{i,j}$ over all genes included in this pathway:

$$p_{t,j} = \frac{\sum_{i=1}^{m_t} w_i \times r_{i,j}}{\sum_{i=1}^{m_t} w_i} \quad (3)$$

Where $p_{t,j}$ represents the activity of the *t*-th pathway in the *j*-th cell type, m_t is the number of genes in the pathway *t*, w_i is the weighting factor equal to the reciprocal of number of pathways that include the *i*-th gene. Statistical significance of higher or lower pathway activity in a specific cell type was evaluated by a random permutation test, in which the cell type labels were randomly shuffled for 5,000 (for the scRNA datasets) or 1,000 times (for the TCGA data) to simulate a null distribution of the pathway activity scores and compare to the pathway activity scores in the original, non-shuffled dataset. For the pathway activity score $p_{t,j}$, we then calculated a p-value defined as the fraction of random pathway activity scores

larger than $p_{t,j}$ (if $p_{t,j}$ is larger than 1) or smaller than $p_{t,j}$ (if $p_{t,j}$ is smaller than 1) to assess if activity of this pathway is significantly higher or lower in this cell type than average.

Analyzing heterogeneity of metabolic pathways

The PCA analysis was applied on the raw gene expression data without imputation of missing values. The function 'prcomp' in R was used to perform the PCA analysis. For each metabolic gene, we computed its PCA score defined as the sum of absolute values of the loadings of this gene in the top PCs that in total account for at least 80% of the variance to measure variability of gene expression across cells. We then sorted the PCA scores of the genes in descending order and applied GSEA analysis to the ranked list of genes to identify metabolic pathways enriched in genes with highest variability. GSEA analysis was done by the software javaGSEA available at http://software.broadinstitute.org/gsea/downloads.jsp with the option "Pre-ranked" and default parameters. The hypoxia signature genes were retrieved from the gene set 'HALLMARK_HYPOXIA' in the molecular signature database (MSigDB) available at http://software.broadinstitute.org/gsea/msigdb/index.jsp.

Analysis of non-malignant cell subtypes

T cells were classified as $CD4^+$ or $CD8^+$ based on expression levels of CD4 and CD8A. T cells with CD4 expression level higher than 1 and CD8A expression level lower than 1 were classified as $CD4^+$ T cells, while those with CD4 expression level lower than 1 and CD8A expression level higher than 1 were classified as $CD8^+$ T cells. Cells with CD4 and CD8A expression levels both higher than 1 were excluded from the following analysis. $CD4^+$ T cells with the total expression level of FOX3P and CD25 higher than 2 were further defined as Tregs, while $CD4^+$ T cells without these two genes expressed (i.e. both genes have zero expression values in these cells) were defined as Ths. For fibroblast cells, k-means clustering

analysis was performed on the expression levels of a set of gene markers (Fig. 5f) to classify them into CAFs and myofibroblasts. The metabolic gene expression profiles were then compared between different cell subtypes using GSEA with the following parameters: nperm=1000, metric=Diff_of_Classes, permute=gene_set, set_max=500, set_min=5. The metabolic pathways with GSEA nominal p-value<0.05 were considered as significant.

Data availability

Raw and processed gene expression values and annotation of cell and tumor types used in this study are available at: https://doi.org/10.6084/m9.figshare.7174922.

Code availability

Computer codes used in this study are available at the Github page of the Locasale Lab: https://github.com/LocasaleLab/Single-Cell-Metabolic-Landscape.

Authors' contributions

Z.X, Z.D and J.W.L designed the study and wrote the manuscript. Z.X performed the data analysis with help from Z.D. All authors have read and approved the final manuscript.

Acknowledgments

We thank members of the Locasale laboratory for helpful discussions. Z.D. thanks Annamarie Allen for help with the text. Support from National Institutes of Health (R01CA193256, R00CA168997), the American Cancer Society (RSG-16-214-01-TBE), and support from the Duke Compute Cluster and Data Commons Storage for computational resources are gratefully acknowledged.

Competing interests

The authors declare no competing interests at this time.

Figures

Figure 1 | Landscape of metabolic gene expression at the single-cell level.

(a) Schematic representation of the scRNA-seq data analysis pipeline.

(b) Numbers of each type of cells in the melanoma dataset.

(c) Numbers of each type of cells in the head and neck squamous cell carcinoma (HNSCC) dataset.

(d) t-SNE plot of metabolic gene expression profiles of malignant cells from the melanoma dataset. The color of each dot indicates the tumor which the cell comes from.

(e) Same as in (d) but for the HNSCC dataset.

(f) Clustered correlation matrix showing Spearman's rank correlation coefficients of metabolic gene expression profiles between malignant cells in the melanoma dataset.

(g) Same as in (f) but for the HNSCC dataset.

(h) t-SNE plot of metabolic gene expression profiles of non-malignant cells from the melanoma dataset. The color of each dot indicates the tumor which the cell comes from.

(i) Same as in (h) but for the HNSCC dataset.

Figure 2 | Cell type-specific metabolic reprogramming in the tumor microenvironment.

(a) Metabolic pathway activities in cell types in the melanoma dataset.

(b) Metabolic pathway activities in cell types in the HNSCC dataset.

(c) Metabolic pathway activities in HNSCC tumor samples and matched adjacent normal samples from TCGA computed based on bulk RNA-seq data. The color bar on the top marks the pathways with similar activity changes in single malignant cells compared to single non-malignant cells and bulk tumors compared to normal tissue samples.

(d) Scatter plot comparing pathway activities between bulk HNSCC tumors in TCGA and single malignant cells in the HNSCC scRNA-seq dataset.

(e) Difference between bulk and single cell RNA-seq in characterizing gene expression profiles in tumors.

(f) Distributions of pathway activities in different cell types from the HNSCC scRNA-seq dataset (left) and in bulk tumors and normal samples from TCGA (right).

Figure 3 | Intratumoral metabolic heterogeneity of malignant cells.

(a) Workflow for quantitating metabolic heterogeneity of malignant cells.

(b) Metabolic pathways enriched in genes with highest contribution to the metabolic heterogeneities among malignant cells from different tumors in the melanoma dataset.

(c) Same as in (b) but for the HNSCC dataset.

(d) Scatter plots comparing activities of glycolysis, OXPHOS and response to hypoxia in single malignant cells from the melanoma dataset.

(e) Same as in (d) but for the HNSCC dataset.

Figure 4 | Metabolic heterogeneity of non-malignant cells.

(a) Metabolic pathways enriched in genes with highest contribution to the metabolic heterogeneities among different types of non-malignant cells from the melanoma dataset.

(b) Pearson’s correlation coefficients between the activities of glycolysis, OXPHOS and response to hypoxia in non-malignant cells from the melanoma dataset.

(c) Same as in (a) but for the HNSCC dataset.

(d) Same as in (b) but for the HNSCC dataset.

Figure 5 | Metabolic features of non-malignant cell subtypes.

(a) Left panel: classification of T cells into $CD4^+$, $CD8^+$, regulatory T cells (Tregs) and T helper cells (Ths). Middle and right: expression levels of the gene markers used for separating T cell subtypes in melanoma (middle) and HNSCC (right) datasets.

(b) Top 10 metabolic pathways enriched in $CD4^+$ or $CD8^+$ T cells in the melanoma dataset. Significantly enriched pathways with GSEA p-value < 0.05 are highlighted in red (higher in $CD8^+$) or blue (higher in $CD4^+$).

(c) Same as in (b) but for the HNSCC dataset.

(d) Top 10 metabolic pathways enriched in Tregs or Ths in the melanoma dataset. Significantly enriched pathways with GSEA p-value < 0.05 are highlighted in red (higher in Th) or blue (higher in Treg).

(e) Same as in (d) but for the HNSCC dataset.

(f) Gene markers and their expression levels used for classifying fibroblast cells in the HNSCC dataset into CAFs and myofibroblasts.

(g) Top 10 metabolic pathways enriched in CAFs or myofibroblasts in the HNSCC dataset. Significantly enriched pathways with GSEA p-value < 0.05 are highlighted in red (higher in myofibroblasts) or blue (higher in CAFs).

References

1 DeBerardinis, R. J. & Chandel, N. S. Fundamentals of cancer metabolism. *Sci Adv* **2**, e1600200, doi:10.1126/sciadv.1600200 (2016).

2 Boroughs, L. K. & DeBerardinis, R. J. Metabolic pathways promoting cancer cell survival and growth. *Nat Cell Biol* **17**, 351-359, doi:10.1038/ncb3124 (2015).

3 Pavlova, N. N. & Thompson, C. B. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab* **23**, 27-47, doi:10.1016/j.cmet.2015.12.006 (2016).

4 Palm, W. & Thompson, C. B. Nutrient acquisition strategies of mammalian cells. *Nature* **546**, 234-242, doi:10.1038/nature22379 (2017).

5 Vander Heiden, M. G. & DeBerardinis, R. J. Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **168**, 657-669, doi:10.1016/j.cell.2016.12.039 (2017).

6 Curtis, M. *et al.* Fibroblasts Mobilize Tumor Cell Glycogen to Promote Proliferation and Metastasis. *Cell Metab*, doi:10.1016/j.cmet.2018.08.007 (2018).

7 Joyce, J. A. & Fearon, D. T. T cell exclusion, immune privilege, and the tumor microenvironment. *Science* **348**, 74-80, doi:10.1126/science.aaa6204 (2015).

8 Rohlenova, K., Veys, K., Miranda-Santos, I., De Bock, K. & Carmeliet, P. Endothelial Cell Metabolism in Health and Disease. *Trends Cell Biol* **28**, 224-236, doi:10.1016/j.tcb.2017.10.010 (2018).

9 Sottoriva, A. *et al.* A Big Bang model of human colorectal tumor growth. *Nat Genet* **47**, 209-216, doi:10.1038/ng.3214 (2015).

10 Ghesquiere, B., Wong, B. W., Kuchnio, A. & Carmeliet, P. Metabolism of stromal and immune cells in health and disease. *Nature* **511**, 167-176, doi:10.1038/nature13312 (2014).

11 Lyssiotis, C. A. & Kimmelman, A. C. Metabolic Interactions in the Tumor Microenvironment. *Trends Cell Biol* **27**, 863-875, doi:10.1016/j.tcb.2017.06.003 (2017).

12 Reina-Campos, M., Moscat, J. & Diaz-Meco, M. Metabolism shapes the tumor microenvironment. *Curr Opin Cell Biol* **48**, 47-53, doi:10.1016/j.ceb.2017.05.006 (2017).

13 Carmona-Fontaine, C. *et al.* Metabolic origins of spatial organization in the tumor microenvironment. *Proc Natl Acad Sci U S A* **114**, 2934-2939, doi:10.1073/pnas.1700600114 (2017).

14 Park, S. *et al.* ERRalpha-Regulated Lactate Metabolism Contributes to Resistance to Targeted Therapies in Breast Cancer. *Cell Rep* **15**, 323-335, doi:10.1016/j.celrep.2016.03.026 (2016).

15 Liu, X. *et al.* Acetate Production from Glucose and Coupling to Mitochondrial Metabolism in Mammals. *Cell* **175**, 502-513 e513, doi:10.1016/j.cell.2018.08.040 (2018).

16 Miller, A. *et al.* Exploring Metabolic Configurations of Single Cells within Complex Tissue Microenvironments. *Cell Metab* **26**, 788-800 e786, doi:10.1016/j.cmet.2017.08.014 (2017).

17 Levitin, H. M., Yuan, J. & Sims, P. A. Single-Cell Transcriptomic Analysis of Tumor Heterogeneity. *Trends Cancer* **4**, 264-268, doi:10.1016/j.trecan.2018.02.003 (2018).

18 Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396-1401, doi:10.1126/science.1254257 (2014).

19 Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189-196, doi:10.1126/science.aad0501 (2016).

20 Chung, W. *et al.* Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun* **8**, 15081, doi:10.1038/ncomms15081 (2017).

21 Puram, S. V. *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* **171**, 1611-1624 e1624, doi:10.1016/j.cell.2017.10.044 (2017).

22 Venteicher, A. S. *et al.* Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* **355**, doi:10.1126/science.aai8478 (2017).

23 Mehrmohamadi, M., Liu, X., Shestov, A. A. & Locasale, J. W. Characterization of the usage of the serine metabolic network in human cancer. *Cell Rep* **9**, 1507-1519, doi:10.1016/j.celrep.2014.10.026 (2014).

24 Lee, D. *et al.* Improving metabolic flux predictions using absolute gene expression data. *BMC Syst Biol* **6**, 73, doi:10.1186/1752-0509-6-73 (2012).

25 Maaten., L. v. d. & Hinton., G. Visualizing Data using t-SNE. *Journal of Machine Learning Research* **9**, 26 (2008).

26 Lun, A. T., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol* **17**, 75, doi:10.1186/s13059-016-0947-7 (2016).

27 Cancer Genome Atlas Research, N. *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45**, 1113-1120, doi:10.1038/ng.2764 (2013).

28 Cancer Genome Atlas, N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576-582, doi:10.1038/nature14129 (2015).

29 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

30 Gatenby, R. A. & Gillies, R. J. Why do cancers have high aerobic glycolysis? *Nature Reviews Cancer* **4**, 891-899, doi:10.1038/nrc1478 (2004).

31 Nakazawa, M. S., Keith, B. & Simon, M. C. Oxygen availability and metabolic adaptations. *Nat Rev Cancer* **16**, 663-673, doi:10.1038/nrc.2016.84 (2016).

32 Gerriets, V. A. & Rathmell, J. C. Metabolic pathways in T cell fate and function. *Trends Immunol* **33**, 168-173, doi:10.1016/j.it.2012.01.010 (2012).

33 Andrejeva, G. & Rathmell, J. C. Similarities and Distinctions of Cancer and Immune Metabolism in Inflammation and Tumors. *Cell Metab* **26**, 49-70, doi:10.1016/j.cmet.2017.06.004 (2017).

34 O'Shea, J. J. & Paul, W. E. Mechanisms underlying lineage commitment and plasticity of helper CD4+ T cells. *Science* **327**, 1098-1102, doi:10.1126/science.1178334 (2010).

35 Sakaguchi, S., Yamaguchi, T., Nomura, T. & Ono, M. Regulatory T cells and immune tolerance. *Cell* **133**, 775-787, doi:10.1016/j.cell.2008.05.009 (2008).

36 Michalek, R. D. *et al.* Cutting edge: distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4+ T cell subsets. *J Immunol* **186**, 3299-3303, doi:10.4049/jimmunol.1003613 (2011).

37 Angelin, A. *et al.* Foxp3 Reprograms T Cell Metabolism to Function in Low-Glucose, High-Lactate Environments. *Cell Metab* **25**, 1282-1293 e1287, doi:10.1016/j.cmet.2016.12.018 (2017).

38 Buck, M. D., Sowell, R. T., Kaech, S. M. & Pearce, E. L. Metabolic Instruction of Immunity. *Cell* **169**, 570-586, doi:10.1016/j.cell.2017.04.004 (2017).

39 Kalluri, R. The biology and function of fibroblasts in cancer. *Nat Rev Cancer* **16**, 582-598, doi:10.1038/nrc.2016.73 (2016).

40 Guido, C. *et al.* Metabolic reprogramming of cancer-associated fibroblasts by TGF-beta drives tumor growth: connecting TGF-beta signaling with "Warburg-like" cancer metabolism and L-lactate production. *Cell Cycle* **11**, 3019-3035, doi:10.4161/cc.21384 (2012).

41 Zhang, D. *et al.* Metabolic reprogramming of cancer-associated fibroblasts by IDH3alpha downregulation. *Cell Rep* **10**, 1335-1348, doi:10.1016/j.celrep.2015.02.006 (2015).

42 Funk, C. D. Prostaglandins and leukotrienes: advances in eicosanoid biology. *Science* **294**, 1871-1875, doi:10.1126/science.294.5548.1871 (2001).

43 Gaude, E. & Frezza, C. Tissue-specific and convergent metabolic transformation of cancer correlates with metastatic potential and patient survival. *Nat Commun* **7**, 13041, doi:10.1038/ncomms13041 (2016).

44 Hu, J. *et al.* Heterogeneity of tumor-induced gene expression changes in the human metabolic network. *Nat Biotechnol* **31**, 522-529, doi:10.1038/nbt.2530 (2013).

45 Liberti, M. V. & Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci* **41**, 211-218, doi:10.1016/j.tibs.2015.12.001 (2016).

46 Reznik, E. *et al.* Mitochondrial DNA copy number variation across human cancers. *Elife* **5**, doi:10.7554/eLife.10769 (2016).

47 Reznik, E., Wang, Q., La, K., Schultz, N. & Sander, C. Mitochondrial respiratory gene expression is suppressed in many cancers. *Elife* **6**, doi:10.7554/eLife.21592 (2017).

48 Hakimi, A. A. *et al.* An Integrated Metabolic Atlas of Clear Cell Renal Cell Carcinoma. *Cancer Cell* **29**, 104-116, doi:10.1016/j.ccell.2015.12.004 (2016).

49 Wheaton, W. W. *et al.* Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *Elife* **3**, e02242, doi:10.7554/eLife.02242 (2014).

50 Liu, X., Romero, I. L., Litchfield, L. M., Lengyel, E. & Locasale, J. W. Metformin Targets Central Carbon Metabolism and Reveals Mitochondrial Requirements in Human Cancers. *Cell Metab* **24**, 728-739, doi:10.1016/j.cmet.2016.09.005 (2016).

51 Lord, S. R. *et al.* Integrated Pharmacodynamic Analysis Identifies Two Metabolic Adaption Pathways to Metformin in Breast Cancer. *Cell Metab* **28**, 679-688 e674, doi:10.1016/j.cmet.2018.08.021 (2018).

52 Guzy, R. D. *et al.* Mitochondrial complex III is required for hypoxia-induced ROS production and cellular oxygen sensing. *Cell Metab* **1**, 401-408, doi:10.1016/j.cmet.2005.05.001 (2005).

53 Vyas, S., Zaganjor, E. & Haigis, M. C. Mitochondria and Cancer. *Cell* **166**, 555-566, doi:10.1016/j.cell.2016.07.002 (2016).

54 Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674, doi:10.1016/j.cell.2011.02.013 (2011).

55 Ecker, C. & Riley, J. L. Translating In Vitro T Cell Metabolic Findings to In Vivo Tumor Models of Nutrient Competition. *Cell Metab* **28**, 190-195, doi:10.1016/j.cmet.2018.07.009 (2018).

56 Li, W. V. & Li, J. J. An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nat Commun* **9**, 997, doi:10.1038/s41467-018-03405-7 (2018).

57 Rtsne: T-Distributed Stochastic Neighbor Embedding using Barnes-Hut Implementation (2015).

58 Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11**, R106, doi:10.1186/gb-2010-11-10-r106 (2010).

59 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

60 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

61 Bullard, J. H., Purdom, E., Hansen, K. D. & Dudoit, S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94, doi:10.1186/1471-2105-11-94 (2010).

62 Lun, A. T., McCarthy, D. J. & Marioni, J. C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res* **5**, 2122, doi:10.12688/f1000research.9501.2 (2016).

Figure 1

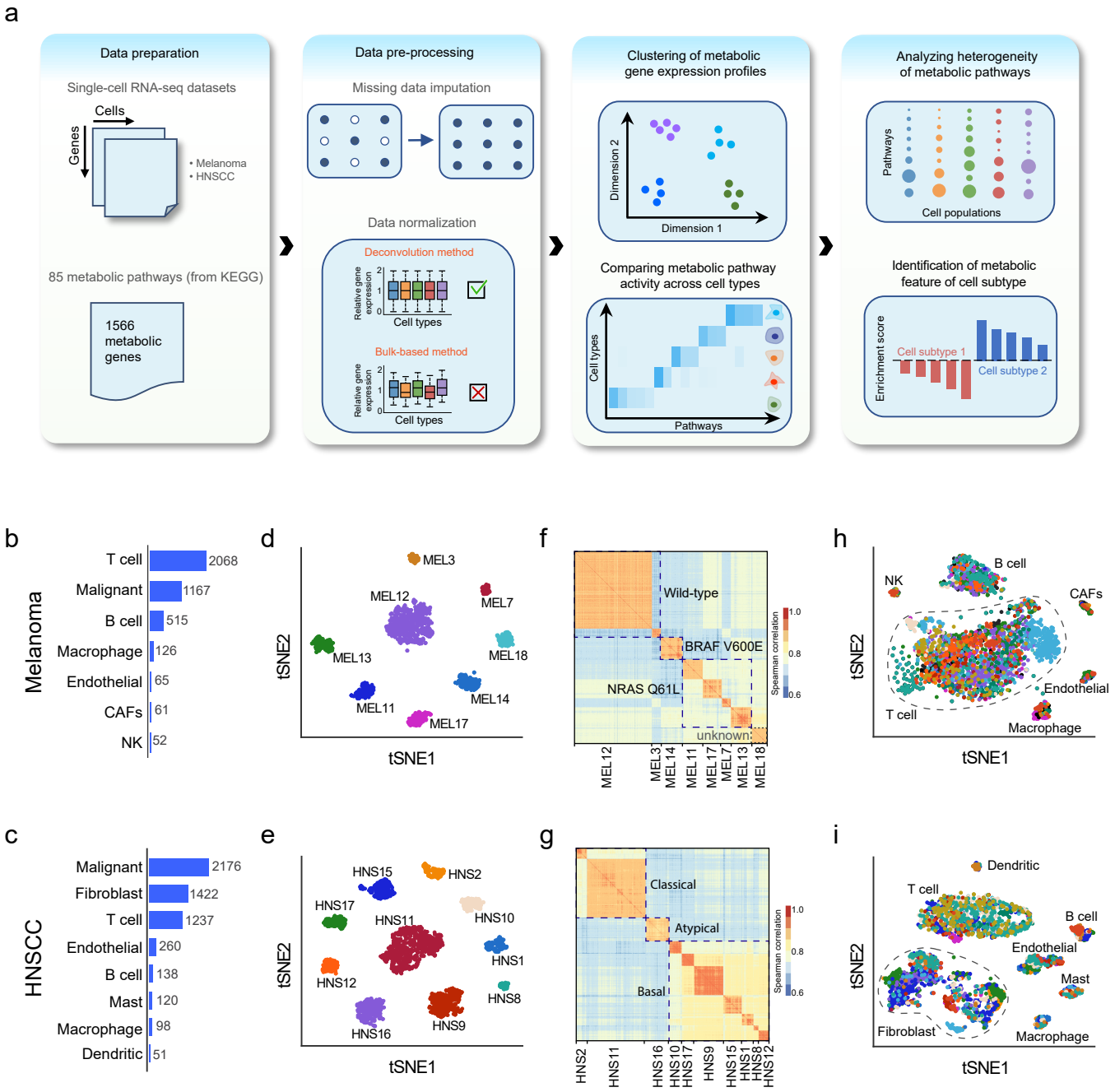

a
Data preparation
Single-cell RNA-seq datasets
Cells
Genes
• Melanoma
• HNSCC
85 metabolic pathways (from KEGG)
1566 metabolic genes
Data pre-processing
Missing data imputation
Data normalization
Deconvolution method
Relative gene expression
Cell types
Bulk-based method
Clustering of metabolic gene expression profiles
Dimension 2
Dimension 1
Comparing metabolic pathway activity across cell types
Cell types
Pathways
Analyzing heterogeneity of metabolic pathways
Pathways
Cell populations
Identification of metabolic feature of cell subtype
Enrichment score
Cell subtype 1
Cell subtype 2
b
Melanoma
T cell 2068
Malignant 1167
B cell 515
Macrophage 126
Endothelial 65
CAFs 61
NK 52
c
HNSCC
Malignant 2176
Fibroblast 1422
T cell 1237
Endothelial 260
B cell 138
Mast 120
Macrophage 98
Dendritic 51
d
MEL3
MEL12
MEL7
MEL13
MEL18
MEL11
MEL14
MEL17
tSNE2
tSNE1
e
HNS15
HNS2
HNS17
HNS11
HNS10
HNS1
HNS12
HNS8
HNS16
HNS9
tSNE2
tSNE1
f
Wild-type
BRAF V600E
NRAS Q61L
unknown
Spearman correlation
1.0
0.8
0.6
MEL12
MEL3
MEL14
MEL11
MEL17
MEL7
MEL13
MEL18
g
Classical
Atypical
Basal
Spearman correlation
1.0
0.8
0.6
HNS2
HNS11
HNS16
HNS10
HNS17
HNS9
HNS15
HNS1
HNS8
HNS12
h
NK
B cell
CAFs
Endothelial
T cell
Macrophage
tSNE2
tSNE1
i
Dendritic
T cell
B cell
Endothelial
Mast
Fibroblast
Macrophage
tSNE2
tSNE1

Figure 2

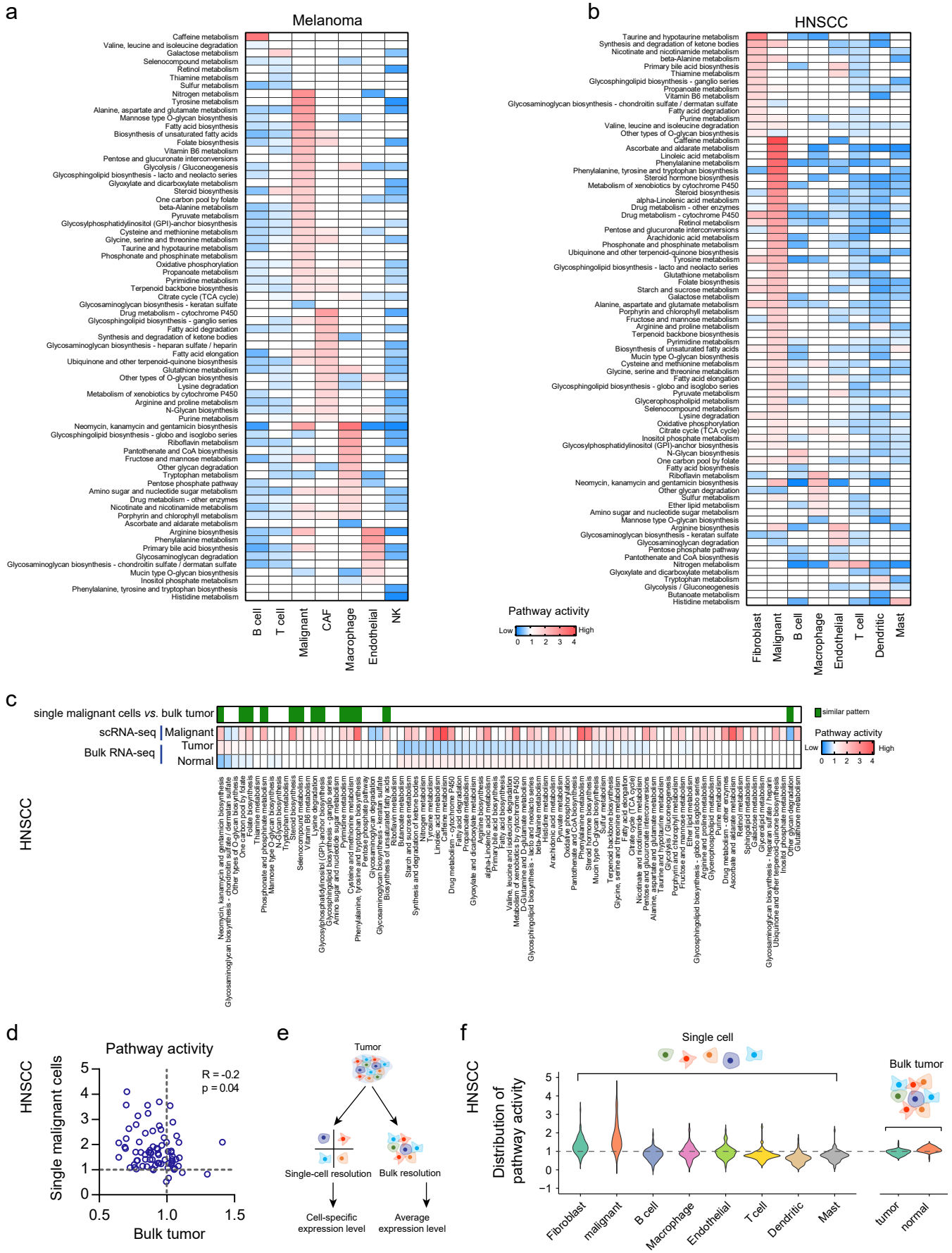

a
Melanoma
B cell
T cell
Malignant
CAF
Macrophage
Endothelial
NK
Pathway activity
Low
High
b
HNSCC
Fibroblast
Malignant
B cell
Macrophage
Endothelial
T cell
Dendritic
Mast
c
single malignant cells vs. bulk tumor
similar pattern
scRNA-seq
Malignant
Bulk RNA-seq
Tumor
Normal
Pathway activity
Low
High
HNSCC
d
Pathway activity
HNSCC
Single malignant cells
R = -0.2
p = 0.04
Bulk tumor
e
Tumor
Single-cell resolution
Bulk resolution
Cell-specific expression level
Average expression level
f
Single cell
HNSCC
Distribution of pathway activity
Fibroblast
malignant
B cell
Macrophage
Endothelial
T cell
Dendritic
Mast
Bulk tumor
tumor
normal

Figure 3

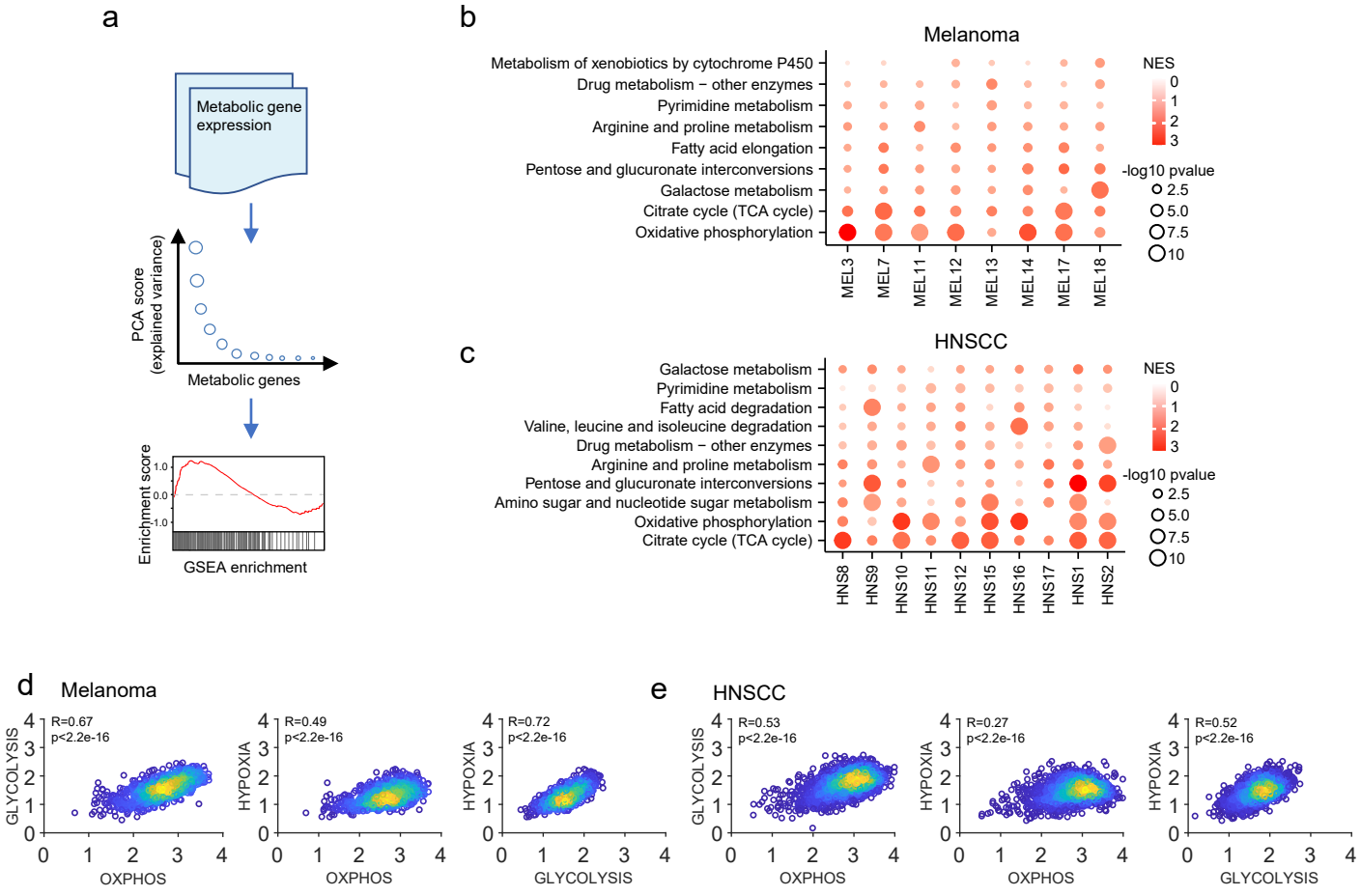

a
Metabolic gene expression
PCA score (explained variance)
Metabolic genes
Enrichment score
GSEA enrichment
b
Melanoma
Metabolism of xenobiotics by cytochrome P450
Drug metabolism − other enzymes
Pyrimidine metabolism
Arginine and proline metabolism
Fatty acid elongation
Pentose and glucuronate interconversions
Galactose metabolism
Citrate cycle (TCA cycle)
Oxidative phosphorylation
MEL3
MEL7
MEL11
MEL12
MEL13
MEL14
MEL17
MEL18
NES
-log10 pvalue
c
HNSCC
Galactose metabolism
Pyrimidine metabolism
Fatty acid degradation
Valine, leucine and isoleucine degradation
Drug metabolism − other enzymes
Arginine and proline metabolism
Pentose and glucuronate interconversions
Amino sugar and nucleotide sugar metabolism
Oxidative phosphorylation
Citrate cycle (TCA cycle)
HNS8
HNS9
HNS10
HNS11
HNS12
HNS15
HNS16
HNS17
HNS1
HNS2
NES
-log10 pvalue
d
Melanoma
R=0.67
p<2.2e-16
GLYCOLYSIS
OXPHOS
R=0.49
p<2.2e-16
HYPOXIA
OXPHOS
R=0.72
p<2.2e-16
HYPOXIA
GLYCOLYSIS
e
HNSCC
R=0.53
p<2.2e-16
GLYCOLYSIS
OXPHOS
R=0.27
p<2.2e-16
HYPOXIA
OXPHOS
R=0.52
p<2.2e-16
HYPOXIA
GLYCOLYSIS

Figure 4

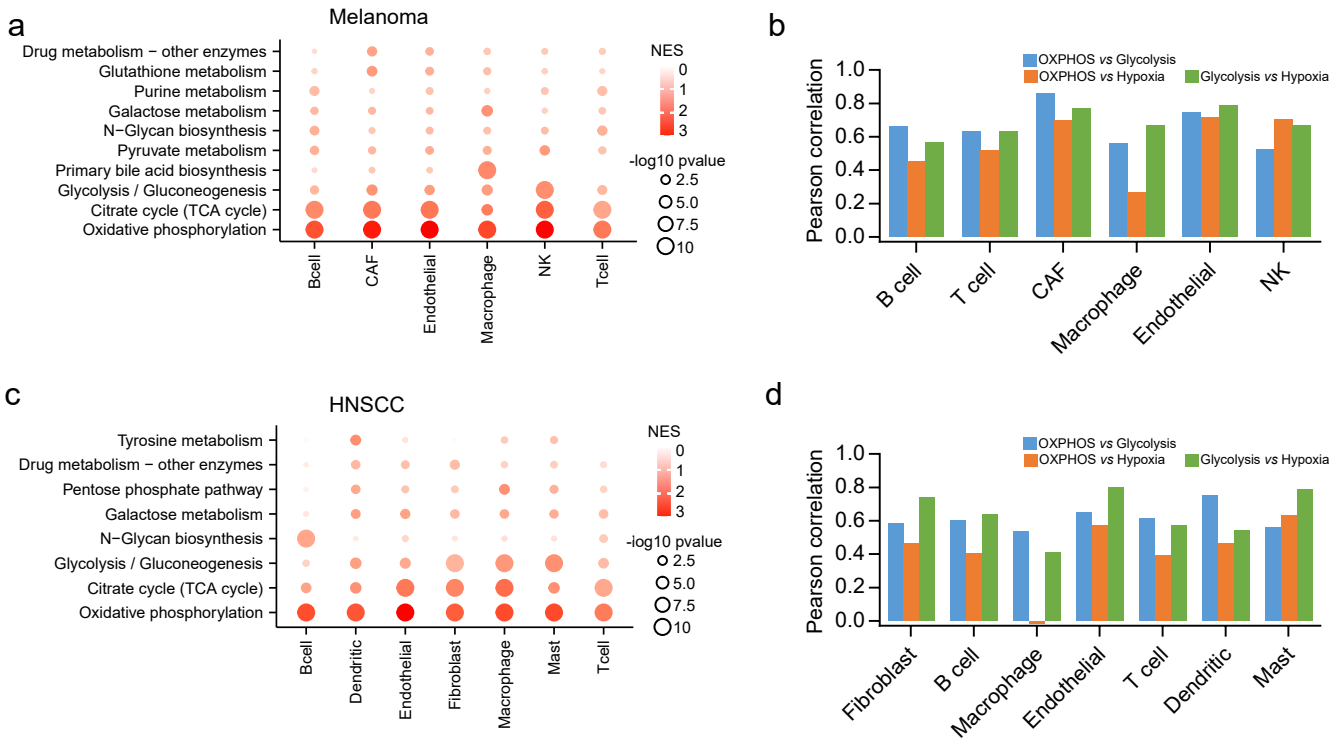
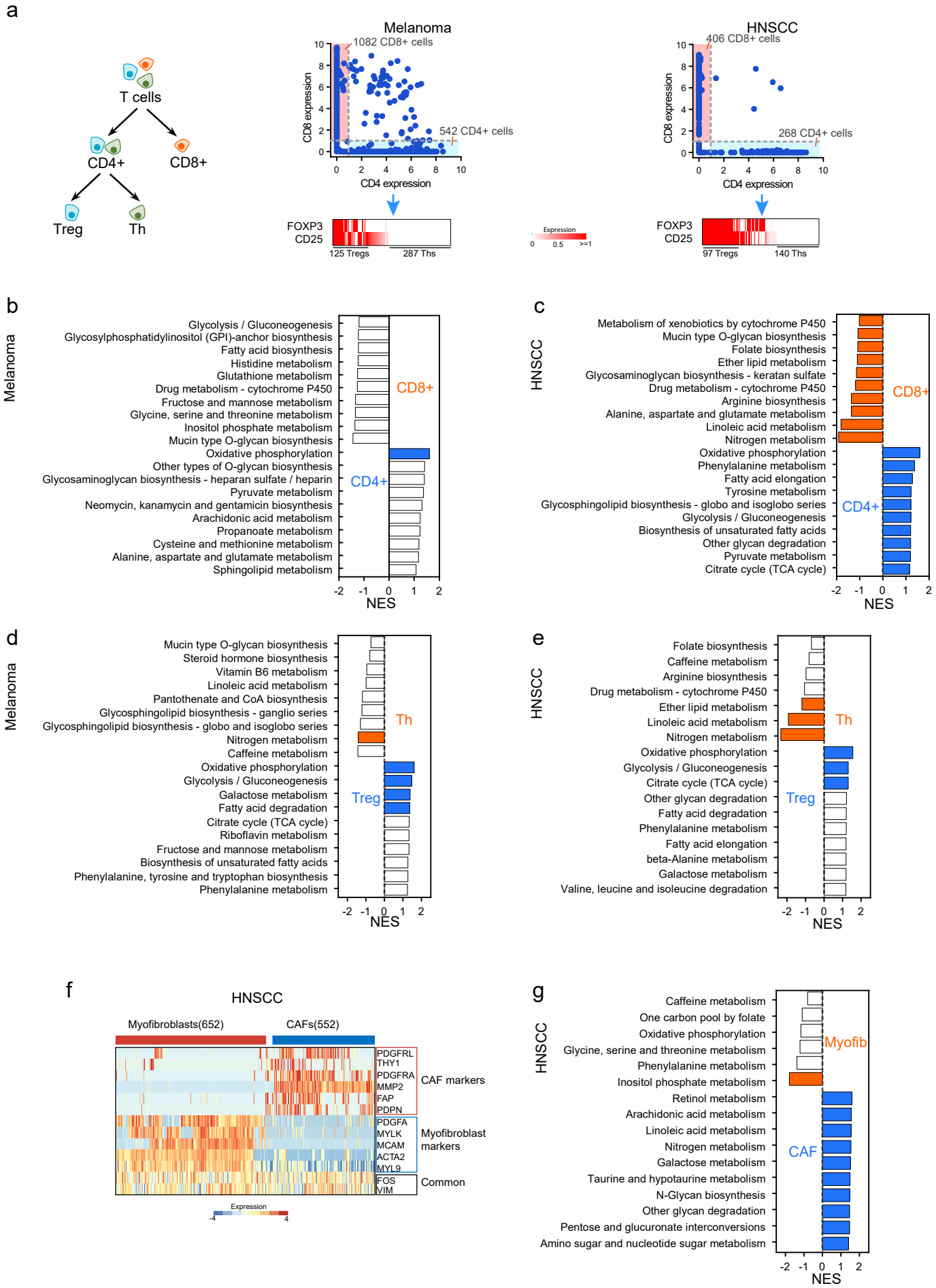

a
Melanoma
Drug metabolism − other enzymes
Glutathione metabolism
Purine metabolism
Galactose metabolism
N−Glycan biosynthesis
Pyruvate metabolism
Primary bile acid biosynthesis
Glycolysis / Gluconeogenesis
Citrate cycle (TCA cycle)
Oxidative phosphorylation
Bcell
CAF
Endothelial
Macrophage
NK
Tcell
NES
0
1
2
3
-log10 pvalue
2.5
5.0
7.5
10
b
Pearson correlation
1.0
0.8
0.6
0.4
0.2
0.0
OXPHOS vs Glycolysis
OXPHOS vs Hypoxia
Glycolysis vs Hypoxia
B cell
T cell
CAF
Macrophage
Endothelial
NK
c
HNSCC
Tyrosine metabolism
Drug metabolism − other enzymes
Pentose phosphate pathway
Galactose metabolism
N−Glycan biosynthesis
Glycolysis / Gluconeogenesis
Citrate cycle (TCA cycle)
Oxidative phosphorylation
Bcell
Dendritic
Endothelial
Fibroblast
Macrophage
Mast
Tcell
NES
0
1
2
3
-log10 pvalue
2.5
5.0
7.5
10
d
Pearson correlation
1.0
0.8
0.6
0.4
0.2
0.0
OXPHOS vs Glycolysis
OXPHOS vs Hypoxia
Glycolysis vs Hypoxia
Fibroblast
B cell
Macrophage
Endothelial
T cell
Dendritic
Mast

Figure 5


a
T cells
CD4+
CD8+
Treg
Th
Melanoma
1082 CD8+ cells
542 CD4+ cells
CD8 expression
CD4 expression
FOXP3
CD25
125 Tregs
287 Ths
HNSCC
406 CD8+ cells
268 CD4+ cells
97 Tregs
140 Ths
Expression
b
Melanoma
Glycolysis / Gluconeogenesis
Glycosylphosphatidylinositol (GPI)-anchor biosynthesis
Fatty acid biosynthesis
Histidine metabolism
Glutathione metabolism
Drug metabolism - cytochrome P450
Fructose and mannose metabolism
Glycine, serine and threonine metabolism
Inositol phosphate metabolism
Mucin type O-glycan biosynthesis
Oxidative phosphorylation
Other types of O-glycan biosynthesis
Glycosaminoglycan biosynthesis - heparan sulfate / heparin
Pyruvate metabolism
Neomycin, kanamycin and gentamicin biosynthesis
Arachidonic acid metabolism
Propanoate metabolism
Cysteine and methionine metabolism
Alanine, aspartate and glutamate metabolism
Sphingolipid metabolism
CD8+
CD4+
NES
c
HNSCC
Metabolism of xenobiotics by cytochrome P450
Mucin type O-glycan biosynthesis
Folate biosynthesis
Ether lipid metabolism
Glycosaminoglycan biosynthesis - keratan sulfate
Drug metabolism - cytochrome P450
Arginine biosynthesis
Alanine, aspartate and glutamate metabolism
Linoleic acid metabolism
Nitrogen metabolism
Oxidative phosphorylation
Phenylalanine metabolism
Fatty acid elongation
Tyrosine metabolism
Glycosphingolipid biosynthesis - globo and isoglobo series
Glycolysis / Gluconeogenesis
Biosynthesis of unsaturated fatty acids
Other glycan degradation
Pyruvate metabolism
Citrate cycle (TCA cycle)
CD8+
CD4+
NES
d
Melanoma
Mucin type O-glycan biosynthesis
Steroid hormone biosynthesis
Vitamin B6 metabolism
Linoleic acid metabolism
Pantothenate and CoA biosynthesis
Glycosphingolipid biosynthesis - ganglio series
Glycosphingolipid biosynthesis - globo and isoglobo series
Nitrogen metabolism
Caffeine metabolism
Oxidative phosphorylation
Glycolysis / Gluconeogenesis
Galactose metabolism
Fatty acid degradation
Citrate cycle (TCA cycle)
Riboflavin metabolism
Fructose and mannose metabolism
Biosynthesis of unsaturated fatty acids
Phenylalanine, tyrosine and tryptophan biosynthesis
Phenylalanine metabolism
Th
Treg
NES
e
HNSCC
Folate biosynthesis
Caffeine metabolism
Arginine biosynthesis
Drug metabolism - cytochrome P450
Ether lipid metabolism
Linoleic acid metabolism
Nitrogen metabolism
Oxidative phosphorylation
Glycolysis / Gluconeogenesis
Citrate cycle (TCA cycle)
Other glycan degradation
Fatty acid degradation
Phenylalanine metabolism
Fatty acid elongation
beta-Alanine metabolism
Galactose metabolism
Valine, leucine and isoleucine degradation
Th
Treg
NES
f
HNSCC
Myofibroblasts(652)
CAFs(552)
PDGFRL
THY1
PDGFRA
MMP2
FAP
PDPN
CAF markers
PDGFA
MYLK
MCAM
ACTA2
MYL9
Myofibroblast markers
FOS
VIM
Common
Expression
g
HNSCC
Caffeine metabolism
One carbon pool by folate
Oxidative phosphorylation
Glycine, serine and threonine metabolism
Phenylalanine metabolism
Inositol phosphate metabolism
Retinol metabolism
Arachidonic acid metabolism
Linoleic acid metabolism
Nitrogen metabolism
Galactose metabolism
Taurine and hypotaurine metabolism
N-Glycan biosynthesis
Other glycan degradation
Pentose and glucuronate interconversions
Amino sugar and nucleotide sugar metabolism
Myofib
CAF
NES