In Silico Modeling of Metabolic State in Single Th17 Cells Reveals Novel Regulators of Inflammation and Autoimmunity

Allon Wagner[1,2,13,*], Chao Wang[3,4,*], David DeTomaso[2], Julian Avila-Pacheco[3], Sarah Zaghouani[4], Johannes Fessler[4], Elliot Akama-Garren[4], Kerry Pierce[3], Noga Ron-Harel[5], Vivian Paraskevi Douglas[6], Marcia Haigis[5,7], Raymond A. Sobel[8], Clary Clish[3], Aviv Regev[3,9,10], Vijay K. Kuchroo[3,4,13,#], & Nir Yosef[1,2, 11,12,13,#]

[1] Department of electrical engineering and computer science, University of California, Berkeley, CA 94720, USA
[2] Center for computational biology, University of California, Berkeley, CA 94720, USA
[3] Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[4] Evergrande center for immunologic diseases, Harvard Medical School and Brigham and Women's Hospital, Boston, MA 02115, USA
[5] Department of cell biology, Harvard Medical School, Boston, MA 02115, USA
[6] Massachusetts Eye and Ear, Department of Ophthalmology, Harvard Medical School, Boston, MA 02114, USA
[7] Blavatnik Institute, Medical School, Boston, MA 02115, USA
[8] Palo Alto Veteran's Administration Health Care System and Department of Pathology, Stanford University School of Medicine, CA 94305, USA
[9] Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[10] Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02140, USA
[11] Chan-Zuckerberg Biohub, San Francisco, CA 94158, USA
[12] Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139, USA

* These authors contributed equally
# These authors contributed equally.
[13] Correspondence to:
allonwag@berkeley.edu, vkuchroo@evergrande.hms.harvard.edu, niryosef@berkeley.edu

ABSTRACT

Cellular metabolism, a key regulator of immune responses, is difficult to study with current technologies in individual cells Here, we present Compass, an algorithm to characterize the metabolic state of cells based on single-cell RNA-Seq and flux balance analysis. We applied Compass to associate metabolic states with functional variability (pathogenic potential) amongst Th17 cells and recovered a metabolic switch between glycolysis and fatty acid oxidation, akin to known differences between Th17 and Treg cells, as well as novel targets in amino-acid pathways, which we tested through targeted metabolic assays. Compass further predicted a particular glycolytic reaction (phosphoglycerate mutase — PGAM) that promotes an anti-inflammatory Th17 phenotype, contrary to the common understanding of glycolysis as pro-inflammatory. We demonstrate that PGAM inhibition leads non-pathogenic Th17 cells to adopt a pro-inflammatory transcriptome and induce autoimmunity in vivo. Compass is broadly applicable for characterizing metabolic states of cells and relating metabolic heterogeneity to other cellular phenotypes.

INTRODUCTION

Cellular metabolism is both a mediator and a regulator of cellular functions. Metabolic activities are key in normal cellular processes such as activation, expansion and differentiation, but also play an important role in the pathogenesis of multiple disease conditions including autoimmunity, cancer, cardiovascular disease, neurodegeneration, and aging. Recently, the study of metabolism in immune cells (immunometabolism) has gained particular attention as a major regulator of almost all aspects of immune responses including anti-viral immunity, autoimmunity, and cancer [1]–[9].

Due to the scale and complexity of the metabolic network, a metabolic perturbation may create cascading effects and eventually alter a seemingly distant part of the network, while cross-cutting traditional pathway definitions. Therefore, computational tools are needed to contextualize observations on specific reactions or enzymes into a systems-level understanding of metabolism and its dysregulation in disease. One successful framework has been Flux Balance Analysis (FBA), which translates curated knowledge on the network's topology and stoichiometry into mathematical objects and uses them to make *in silico* predictions on metabolic fluxes [10]–[13]. FBA methods have proven particularly useful when contextualized with functional genomics data, including gene expression [14].

While such metabolic models aim to represent the behavior of individual cells, their contextualization has generally relied on information collected from bulk population data. However, the advent of single-cell RNA-Seq (scRNA-Seq) has highlighted the substantial extent of cell-to-cell diversity that is often missed by bulk profiles [15], [16], and can be especially

prominent in immune cells and associated with their functional diversity [8], [17]–[28]. One of the earliest examples has been the diversity among T helper 17 (Th17) cells [29]. On the one hand, IL-17 producing Th17 cells can be potent inducers of tissue inflammation in autoimmune disorders [30], [31] but on the other hand, these cells are critical in host defense against pathogens [32], [33] and can promote mucosal homeostasis and barrier functions [34]–[36]. Th17 cells with distinct effector functions can be found in patients and animal models and can also be generated in vitro with different combinations of differentiation cytokines, as we have previously demonstrated [34]. We have previously shown that such functional diversity can be captured by studying transcriptional diversity at the single cell level with scRNA-Seq, and enabled the discovery of novel regulators that are otherwise difficult to detect in bulk RNA-Seq analysis [29], [37].

We hypothesized that a similar spectrum of diversity may exist at the immunometabolic level and relate to cell function. However, most cellular assays, including metabolic assays, are normally done in a targeted manner and difficult to undertake at single-cell level. Furthermore, low cell numbers frequently prohibit direct metabolic assays, for example, in the study of immune cells that are present at tissue sites. In contrast, scRNA-Seq, is broadly accessible and rapidly collected across the human body [38], and should allow, in principle, to contextualize metabolic models to the single cell level. A computational method is thus required to systematically address the unique challenges of scRNA-Seq, such as data sparsity, and to capitalize on its opportunities, for example by treating cell populations as natural perturbation systems with a rapidly increasing scale [39].

Here, we present Compass, an FBA algorithm to characterize and interpret the metabolic heterogeneity among cells, which uses available knowledge of the metabolic network in conjunction with RNA expression of metabolic enzymes. Compass uses single cell transcriptomic profiles to characterize cellular metabolic states at single-cell resolution and with network-wide comprehensiveness. It allows detection of metabolic targets across the entire metabolic network, agnostically of pre-defined metabolic pathway boundaries, and including ancillary pathways that are normally less studied, yet could play an important role in the determining cell function [40]. We applied Compass to Th17 cells, uncovering substantial immunometabolic diversity associated with their inflammatory effector functions. In addition to the expected glycolytic shift, we found diversity in amino acid metabolism, and highlighted a unique and surprising role for the glycolytic reaction catalyzed by phosphoglycerate mutase (PGAM) in promoting an anti-inflammatory phenotype in Th17 cells. Compass is a broadly applicable tool for studying metabolic diversity at the single cell level, and its relationship to the functional diversity between cells.

RESULTS

Compass — an algorithm for comprehensive characterization of single-cell metabolism

We reasoned that even though the mRNA expression of individual enzymes does not necessarily provide an accurate proxy for their metabolic activity, a global analysis the entire metabolic network (as enabled by RNA-Seq) in the context of a large sample set (as offered by single cell genomics) coupled with strict criteria for hypotheses testing, would provide an effective framework for predicting cellular metabolic status of the cell. This led us to develop the Compass algorithm, which integrates scRNA-Seq profiles with prior knowledge of the metabolic network to infer a metabolic state of the cell **(Figure 1A)**.

The metabolic network is encoded in a Genome-Scale Metabolic Model (GSMM) that includes reaction stoichiometry, biochemical constraints such as reaction irreversibility and nutrient availability, and gene-enzyme-reaction associations. Here, we use Recon2, which comprises of 7,440 reactions and 2,626 unique metabolites [41]. To explore the metabolic capabilities of each cell, Compass solves a series of constraint-based optimization problems (formalized as linear programs) that produce a set of numeric scores, one per reaction **(Methods)**. Intuitively, the score of each reaction in each cell reflects how well adjusted is the cell's overall transcriptome to maintaining high flux through that reaction. Henceforth, we refer to the scores as quantifying the "potential activity" of a metabolic reaction (or "activity" in short when it is clear from the context that Compass predictions are discussed).

Compass belongs to the family of Flux Balance Analysis (FBA) algorithms that model metabolic fluxes, namely the rate by which chemical reactions convert substrates to products and apply constrained optimization methods to find flux distributions that satisfy desired properties (a flux distribution is an assignment of flux value to every reaction in the network) [10]–[13]. In the first step, Compass is agnostic to any measurement of gene expression levels and computes, for every metabolic reaction *r*, the maximal flux v_r^{opt} it can carry without imposing any constraints on top of those imposed by stoichiometry and mass balance. Next, Compass relies on the assumption that mRNA expression of an enzyme coding gene should preferably correlate with the flux through the metabolic reaction(s) it catalyzes. It thus assigns every reaction in every cell a penalty inversely proportional to the mRNA expression associated with its enzyme(s) in that cell. Compass then finds a flux distribution which minimizes the overall penalty incurred in any given cell *i* (summing over all reactions), while maintaining a flux of at least $\omega \cdot v_r^{opt}$ (here $\omega = 0.95$) in *r*. The Compass score of reaction *r* in cell *i* is the negative of that minimal penalty (so that lower scores correspond to lower potential metabolic activity).

Using genome-scale metabolic network allows the entire metabolic transcriptome to impact the computed score for any particular reaction, rather than just the mRNA coding for the enzymes that catalyze it. We reasoned that this helps reduce the effect of instances where mRNA expression does not correlate well with metabolic activity, for example due to post-transcriptional or post-translational modifications. This also mitigates the effects of data sparsity, which is characteristic of scRNA-Seq data. The low transcript signal in scRNA-Seq, which results in the extreme case in false-negative gene detections, magnifies the repercussions of sampling bias and transcription stochasticity, and leads to an overestimation of the variance of lowly expressed genes, which in turn leads in turn to false-positive calling of differentially expressed genes [15]. Compass further mitigates data sparsity effects with an information-sharing approach, similar to other scRNA-Seq algorithms [42]–[50]. Instead of treating each cell in isolation, the score vector for each cell is determined by a combined objective that balances the effects in the cell in question with those in its *k*-nearest neighbors (based on similarity of their RNA profiles; here, using $k = 10$; **Figure 1B; Methods**).

The output of Compass is a quantitative profile for the metabolic state of every cell, which is then subject to downstream analyses **(Figure 1C)**. These include finding metabolic reactions that are differentially active between cell types or that correlate with continuous properties of cell state (e.g., expression of a certain group of cytokines). It can also provide unsupervised insights into cellular diversity by projecting cells into a low-dimensional space of metabolic activity (e.g., for visual exploration). The statistical power afforded by the large number of individual cells in a typical scRNA-Seq study adds robustness and allows these downstream analyses to gain biological insight despite the high dimension of the metabolic space in which Compass embeds cells. Finally, because Compass does not rely on a predetermined set of metabolic pathways (or gene sets) such as Reactome [51] or KEGG [52], it allows unsupervised derivation of cell-specific metabolic pathways in a data driven way.

Th17 cell metabolic diversity reflects a balance between glycolysis and fatty acid oxidation, which is associated with pathogenicity

To demonstrate Compass, we applied it to scRNA-Seq data from Th17 cells, differentiated *in vitro* from naïve CD4+ T into two extreme functional states [34], [53] **(Figure 2A)**. Differentiation with IL-1β+IL-6+IL-23 creates Th17 cells that upon adoptive transfer into recipient mice induce severe neuroinflammation in the form of experimental autoimmune encephalomyelitis (EAE). We refer to these cells as pathogenic Th17 (Th17p). However, when naïve CD4+ T cells are cultured with TGF-β1+IL-6, the resulting Th17 induce only mild-to-none EAE when adoptively transferred to recipient mice. We refer to these cells as non-pathogenic Th17 (Th17n). We performed a Compass analysis of a dataset we generated in a previous study that included 139 Th17p and 151 Th17n cells sorted for IL-17A/GFP+ and profiled using Fluidigm C1 and SMART-Seq2 [29], [37].

We first computed a Compass score for each metabolic reaction in each cell **(Methods).** We then aggregated reactions that were highly correlated across the entire dataset (Spearman rho ≥ 0.98) into meta-reactions (median of two reactions per meta-reaction; **Supplementary Figure 1**) for downstream analysis.

To investigate the main determinants of Th17 cell-to-cell metabolic heterogeneity, we first analyzed the Compass output as a high dimensional representation of the cells which parallels the one produced by scRNA-Seq, but with features corresponding to metabolic meta-reaction rather than transcripts. We performed principal component analysis (PCA) on the meta-reaction matrix, while restricting it to 784 meta-reactions (out of 1,911) associated with core metabolism (**Methods**) that span conserved and well-studied pathways for generation of ATP and synthesis of key biomolecules.

The first two principal components (PCs) of the core metabolism subspace were associated both with overall metabolic activity and T effector functions (**Figure 2B, Supplementary Figure 2A,B**, **Supplementary Table 1**). PC1 correlated with the cell's total metabolic activity, defined as the expression ratio of genes coding metabolic enzymes out of the total protein coding genes (Pearson rho = 0.36, $p < 4*10^{-10}$), as well as a transcriptional signature of late stages of Th17 differentiation over time [54] (**Supplementary Figure 2C,** Pearson rho = 0.18, $p < 0.003$) (**Methods**). PC2 and PC3 represented a choice between ATP generation through aerobic glycolysis versus fatty acid oxidation, which is a prominent finding in immunometabolism while comparing activated Th17 to Tregs, or Teff vs. Tmem cells [3]. Accordingly, they were correlated with multiple Th17 pathogenicity markers, as well as a signature of Th17 pathogenicity consisting of cytokines, chemokines and transcription factors that are associated with each phenotypic group [29], [34] (**Supplementary Figure 2D,E**). PC2 and PC3 were also noticeably associated with nitrogen metabolism, and were enriched in urea cycle targets whose power to modulate Th17 pathogenicity is demonstrated below and in an accompanying manuscript (Wang et al., biorxiv preprint).

<u>Compass predicts metabolic regulators of Th17 cell pathogenicity</u>

To directly search for metabolic targets that are associated with the pathogenic capacity of individual Th17 cells, we searched for biochemical reactions with differential predicted activity between the Th17p and Th17n conditions according to Wilcoxon's rank sum p value and Cohen's d effect size statistics) and defined pro-pathogenic and pro-regulatory reactions as ones that were significantly different in the Th17p or Th17n direction, respectively (**Figure 2C; Supplementary Figure 2F; Methods; Supplementary Table 2**). We next discuss several key predictions, which we follow up on in the rest of the manuscript as well as in the accompanying manuscript (Wang et al., biorxiv preprint).

Metabolic reactions in both primary and ancillary pathways were associated with Th17 cell pathogenicity (1,213 or 3,362 reactions out of 6,563 reactions, Benjamini-Hochberg (BH) adjusted Wilcoxon $p < 0.001$ or 0.1, respectively). Many of these reactions are also significantly correlated with the expression of signature genes for Th17 functional activity, which code cytokines and transcription factors [34] (**Figure 2D; Supplementary Figure 2G; Supplementary Table 3**), but not metabolic enzymes. Notably, many classically defined metabolic pathways partially overlapped both with reactions predicted to be pro-pathogenic and reactions predicted to be pro-regulatory (**Figure 2E**), highlighting the value in examining single reactions within a global network rather than conducting a pathway-level analysis. A similar result is obtained at the gene expression level — many metabolic pathways included both genes that were upregulated and genes that were downregulated in Th17p compared to Th17n (**Supplementary Figure 2H, Supplementary Table 4**).

Compass highlighted distinctions in central carbon and fatty acid metabolism between the Th17p and Th17n states, which mirror those found between Th17 and Foxp3$^+$ T regulatory (Treg) cells. In central carbon metabolism, Compass predicted that glycolytic reactions, ending with the conversion of pyruvate to lactate are generally more active in the pro-inflammatory Th17p than in the Th17n state **(Figure 2C and 3A)**. This parallels previous results showing that Th17 cells upregulate glycolysis even in the presence of oxygen (hence “aerobic glycolysis”), and that interference with this process promotes a Treg fate [55]–[57]. Compass also predicted an increased activity in Th17p through two segments of the TCA cycle, but not the cycle as a whole (**Figures 2C** and **3A**). A similar breakdown of the TCA cycle in relation with pro-inflammatory function has been shown in macrophages where M1 polarization divided the TCA cycle at the same two points: at isocitrate dehydrogenase (IDH) [58], and at succinate dehydrogenase (SDH) [59], which supported their inflammatory functions [60], [61].

In fatty acid metabolism, Compass predicted that cytosolic acetyl-CoA carboxylase (ACC1), the committed step towards fatty acid synthesis, is upregulated in Th17p, whereas the first two steps of long-chain fatty acid oxidation (long chain fatty acyl-CoA synthetase and carnitine O-palmitoyltransferase (CPT)) were predicted to be significantly higher in Th17n. These predictions mirror a known metabolic difference between the Th17 and Treg lineages, where Th17 cells rely more on *de novo* fatty acid synthesis [62], whereas Tregs scavenge them from their environment and catabolize them and produce ATP through beta-oxidation [56]. We note, however, that recent evidence suggests that CPT may be upregulated in Treg over Th17, but is not functionally indispensable for Treg cells to obtain their effector phenotypes [63].

Among ancillary metabolic pathways, Compass highlighted multiple reactions of amino-acid metabolism that are differentially active between Th17p and Th17n cells **(Figure 2C, Supplementary Table 2)**. It was previously shown that amino acids are important for Th17 cell

differentiation [64], and Compass adds further granularity to these findings. In particular, it predicted that serine biosynthesis from 3-phosphoglycerate, as well as three downstream serine fates — sphingosines, choline, and S-adenosyl-methionine (SAM) — were higher in Th17p. On the other hand, parts of urea cycle and arginine metabolism are significantly associated with both pro-Th17p and pro-Th17n states, (**Figure 2C**), suggesting that alternative fluxing within this sub-system may be associated with diverging Th17 cell function. We pursue these predictions and study this subsystem in detail in a companion manuscript (Wang et al., biorxiv preprint). In the following sections we validate the other predictions discussed thus far and build on them to find novel metabolic regulators of Th17 functional states.

Pathogenic Th17 cells maintain higher aerobic glycolysis and TCA activity, whereas Non-pathogenic Th17 cells oxidize fatty acids to produce ATP

We validated the Compass prediction that pathogenic and non-pathogenic Th17 functional states differ in their central carbon metabolism (**Figure 3a**), using Seahorse assays and liquid-chromatography mass spectrometry (LC/MS) based metabolomics.

First, we compared glycolysis and mitochondrial function of Th17p and Th17n cells. A Seahorse assay (which involves culturing cells with glucose-rich media) confirmed that Th17p cells caused significantly higher extracellular acidification (ECAR) than Th17n, indicating accumulation of lactic acid due to aerobic glycolysis **(Figure 3B, top**). Th17p cells also generated significantly more ATP in a mitochondria dependent fashion (**Figure 3B, bottom**), consistent with the predicted higher entrance of pyruvate into the TCA cycle despite the diversion of some pyruvate towards the lactate fate.

Next, we directly measured metabolites within the glycolysis pathway and TCA cycle using LC/MS based metabolomics. When pulsed with fresh media containing glucose (and rested for 15 minutes), there is a substantial increase in glycolytic metabolites in Th17p but less so in Th17n cells (**Figure 3C, top**). Conversely, steady state (pre-pulsing) Th17p and Th17n cells show no apparent difference in glycolytic metabolites, likely due to alternative nutrients or shunting of the glycolytic metabolites into alternative fates. Indeed, after 3 hours with glucose pulsing, the increased level of such metabolites in Th17p return to steady state (**Supplementary Figure 3**).

Interestingly, Compass predicted that two parts of the TCA cycle, but not the cycle as a whole, were upregulated in Th17p: the conversion of citrate to isocitrate and of alpha-ketoglutarate to succinate (mirroring previous findings in macrophages, see above and [58], [59]). LC/MS Metabolomics analysis of cells at steady state revealed that TCA metabolites were generally more abundant in Th17p than in Th17n, apart from succinate **(Figure 3C, middle)**. Therefore, both Compass and the metabolomics data point to succinate as a potential metabolic control point.

To test whether not only absolute metabolite levels, but also the relative allocation of carbon into its possible fates differ between Th17p and Th17n cells, we performed a carbon tracing assay with 13C-glucose. We augmented fresh media with 13C-labeled glucose and computed the ratio of the 13C isotope out of the total carbon for each metabolite. Consistent with our predictions, Th17p had significantly higher relative abundance of 13C-labeled glycolytic metabolites than Th17n (**Figure 3D**). Furthermore, Th17p preferentially incorporated glucose-derived carbon into serine (which branches from glycolysis; **Supplementary Figure 3B**) and its downstream product choline (**Figure 3D**), consistent with shunting of glycolytic metabolites into alternative fates by Th17p cells. This also conforms to the Compass prediction of elevated serine synthesis in Th17p (**Figure 2C**). Th17p cells also had significantly lower relative abundance of 13C-labeled TCA metabolites (**Figure 3D**), suggesting that the higher level of TCA intermediates observed in Th17p at steady state (**Figure 3C**) might not be supported from glucose, but rather from other sources, such as catabolism of amino acids. Taken together, our results suggest that Th17p cells have a higher overall activity through the TCA cycle at steady-state, but quickly switch to aerobic glycolysis when glucose is readily forged from the environment, as observed in the Seahorse, the fresh-media pulsing metabolome assay, and the carbon tracing assay.

We next validated that Th17n cells prefer beta oxidation as predicted by Compass. Metabolomics analysis shows that Th17n cells were enriched in acyl-carnitine metabolites **(Figure 3C, bottom)**, indicative of active lipid transport through the mitochondrial membrane. This could be a result of either increased lipid biosynthesis or increased catabolism (via beta-oxidation), since acyl-carnitines are intermediates of both processes. However, acyl-carnitines are noticeably more abundant in Th17n, and short- to medium-length acyl groups are particularly more abundant in the steady state and three hours post glucose pulsing **(Supplementary Figure 3A)**. This supports the hypothesis that under glucose-poor conditions, Th17n cells, more than Th17p, break fatty acids to produce energy (a process which involves the progressive degradation of long-chain fatty acids into shorter acyl-CoA chains, two carbon atoms at a time). Indeed, when etomoxir was used to block acyl-carnitine transportation across mitochondrial membranes, oxygen consumption rate decreased in Th17n but not Th17p cells, as measured by Seahorse assay (**Figure 3E**). Although etomoxir has off-target effects [63], [65], overall our data supports the hypothesis that Th17n cells ultimately divert fatty acid breakdown products into the electron transport chain to generate ATP, which utilizes oxygen as an electron acceptor.

PDK4-deficient Th17p cells adopt a non-pathogenic-like central carbon program, but retain a pathogenic-like amino acid phenotype

Pyruvate dehydrogenase (PDH) is a critical metabolic juncture through which glycolysis-derived pyruvate enters the TCA cycle (**Supplementary Figure 3B**). Previous studies have shown that the PDH inhibiting kinase 1 (PDK1) is expressed at higher levels in Th17 cells compared to Th1 or

FoxP3^{+} Tregs. Consistently, inhibition of PDK1 suppressed Th17 cells but increased the abundance of Tregs [57], whereas PDH activation by PDH phosphatase catalytic subunit 2 (PDP2) had the opposite effects [66]. Compass's prediction (**Figure 2D**) of increased glycolytic activity, along with these previous studies, prompted us to ask whether PDH has a parallel role in regulating Th17p vs. Th17n states mirroring the reports for Th17 vs. Treg cells [57]. Among PDH inhibitors, PDK4 is of particular interest in immunometabolism, because it plays a role in the cellular starvation response [67], [68].

To determine whether increased glycolysis, regulated by PDH enzymes, in Th17p cells is important for their global metabolic phenotype, we used PDK4$^{-/-}$ mice for perturbation. Despite the low expression of PDK4 mRNA in Th17 cells **(Supplementary Figure 3C)**, Th17 cells differentiated from naive T cells from PDK4$^{-/-}$ mice had reduced conversion of pyruvate to lactate as measured by ECAR in Th17p but not Th17n conditions **(Figure 3F)**. This suggests that PDK4-deficiency increases the alternative pyruvate fate, namely entrance into the TCA only in Th17p cells. It also suggests that Th17n cells control pyruvate entrance to the TCA cycle by other means than PDK4.

To further determine the global transcriptional and metabolic changes induced by PDK4 perturbation, we profiled 146 WT and 132 PDK4$^{-/-}$ Th17p cells and 236 WT and 307 PDK4$^{-/-}$ Th17n cells by scRNA-seq using SMART-Seq2. Consistent with the size of the effects on lactate secretion (**Figure 3F**), we observed a considerably larger effect of PDK4-deficiency on the transcriptome of Th17p cells **(Figure 3G, Supplementary Table 5)**. The genes differentially expressed in Th17p cells were primarily enriched in central-carbon metabolism (**Supplementary Figure 3D** and **Supplementary Table 6**). LC/MS metabolomics showed that PDK4-deficient Th17p cells had notably higher levels of acyl-carnitine, indicating elevated fatty acid transport across mitochondrial membranes (**Supplementary Figure 3E**), similar to WT Th17n cells. Nevertheless, the metabolic phenotype of PDK4-deficient Th17p did not fully shift towards that of Th17n cells. PDK4-deficient Th17p cells retained the WT Th17p metabolome in pathways other than central carbon metabolism, for instance in amino-acid pathways (**Supplementary Figure 3F**). Furthermore, we did not observe significant differences in the expression of key cytokines or transcription factors that are associated with the effector function of these cells.

Seeing that PDK4-deficiency had partially shifted Th17p central carbon metabolism towards the Th17n state *in vitro*, we next tested the pertaining effects *in vivo*. To this end, we studied the impact of PDK4- deficiency on the development of EAE, an autoimmune disease induced by pathogenic Th17 cells. Consistent with previous studies that glycolysis promotes inflammation [57], [69]–[71], mice with global knockout of PDK4 developed less severe disease as determined by the clinical disease scores **(Figure 3H)** with decrease in Th17 cells and increase in the infiltration of Foxp3^{+} Tregs in the CNS of the mice undergoing EAE (**Figure 3I**). We therefore

conclude that PDK4-deficient Th17p cells resemble Th17n in their central carbon metabolic state, but not in other metabolic pathways. These results prompted us to interrogate metabolic differences outside of central carbon pathways, which are presented in the companion manuscript (Wang et al., biorxiv preprint).

The glycolytic enzyme phosphoglycerate mutase (PGAM) suppresses Th17 cell pathogenicity

Thus far, our analysis relied on an *inter-population* comparison between the extreme states of Th17n and Th17p cells. However, we have previously shown that there is also considerable continuous variation in the transcriptomes of Th17n cells, which spans into pathogenic-like states [29]. To explore the relationship between metabolic heterogeneity and pathogenic potential within the Th17n subset, we next performed an intra-population analysis of Th17n cells. This also demonstrates that Compass can be applied to scRNA-Seq data in cases where the states of interest (e.g., Th17n vs. Th17p) are either unknown or cannot be experimentally partitioned into discrete types. To perform an intra-population Compass analysis of single Th17n cells, we correlated the Compass scores associated with each reaction with the pathogenicity gene signature scores of the respective cells **(Methods)**.

While the resulting correlations of individual reactions with the pathogenicity score were largely consistent with the results of the inter-population analysis (Th17p vs. Th17n) **(Figure 4A, Supplementary Tables 7-8**), the intra-population analysis predicted that some glycolytic reactions may be negatively, rather than positively (as in the inter-population analysis), associated with Th17 pathogenicity. The most notable of these reactions was the one catalyzed by the enzyme phosphoglycerate mutase (PGAM), which was negatively associated with pathogenicity in the intra-population analysis of Th17n cells **(Figure 4A)**, but positively associated with Th17p cells in the inter-population analysis **(Figure 2B,C)**. This prediction was unexpected because increased glycolysis is generally known to support pro-inflammatory phenotypes in T cells (see also **Discussion**) [55]–[57], [69], [72]–[78].

To functionally validate the glycolytic targets associated with Th17 cell pathogenicity by the intra-population analysis, we used chemical inhibitors against enzymes driving the top two glycolytic reactions that were most positively correlated (regulated by pyruvate kinase muscle isozyme [PKM], and glucose-6-phosphate dehydrogenase [G6PD]) and top two that were most negatively correlated (phosphoglycerate mutase [PGAM], and glucokinase [GK]) with the pathogenicity score (**Figure 4B**). The inhibitors were shikonin (inhibits PKM2), dehydroepiandrosterone (DHEA, inhibits G6PD), epigallocatechin-3-gallate (EGCG, inhibits PGAM1), and 2,3-dihydroxypropyl-dichloroacetate (DCA, inhibits GK) (**Methods**).

We first analyzed the effects of inhibitors on Th17n and Th17p cell differentiation and function using flow cytometry (**Figure 4C**). Due to the possibly deleterious effects of blocking these central

reactions on cell viability, we used the highest dose of each inhibitor that did not affect cell viability (compared to solvent alone). We further used flow cytometry to restrict the analysis to cells that had undergone one division (d1) so as to exclude arrested cells or cells that have been blocked from activation and expansion. In addition, since two different solvents (DMSO and methanol) were needed for different inhibitors, every treatment group was matched with an appropriate vehicle control. We found that IL-17 expression conformed to the prediction made by Compass. It was significantly upregulated by chemical inhibition of the two enzymes (PGAM or GK) predicted to suppress pathogenicity, and downregulated by chemical inhibition of the two enzymes (G6PD or PKM) predicted to promote pathogenicity **(Figure 4C)**. This was further confirmed when profiling a larger set of cytokines secreted by Th17 cells: inhibition of PKM or G6PD curtailed all cytokine production suggesting that these enzymes are important for overall T effector functions. In contrast, cells with PGAM or GK inhibition, at the optimal concentration, mostly retained their cytokine profile with a few exceptions **(Supplementary Figure 4).**

To analyze the impact of perturbing glycolytic enzymes on the transcriptome, we used bulk RNA-Seq to profile Th17n and Th17p cells grown in the presence of either the predicted pro-regulatory inhibitor DHEA (inhibiting G6PD) or the predicted pro-inflammatory inhibitor EGCG (inhibiting PGAM) (**Figure 4D-F**). The first principal component (PC1), which is the main axis of variation in the data, represented as expected, the pathogenicity phenotype. In both Th17n and Th17p cells, EGCG shifted cell profiles towards a more pathogenic state on PC1, whereas DHEA shifted them to a less pathogenic state (**Figure 4D**). The difference between the two vehicle controls was inconsequential compared to cell type and interventions.

To better interpret the drug-induced transcriptional changes, we examined individual genes whose expression is associated with either Th17n or Th17p effector function as wells as global transcriptomic shifts **(Figure 4E,F** and **Supplementary Table 9)**. A comparison of DHEA to vehicle control identified a large number of effector genes that are modulated **(Figure 4E, right)**. These include a significant decrease in *IL23R* and *TBX21* transcripts in both Th17p and Th17n, two genes critical for Th17 cell pathogenicity, and in IL9 and IL1RN, two genes highly expressed in non-pathogenic Th17 cells (Lee et al. 2012). Conversely, EGCG clearly strengthened the pathogenic transcriptional program in Th17n, globally upregulating pro-inflammatory genes (e.g., IL22, IL7R, and CASP1) and (to a more limited extent) downregulating pro-regulatory ones (e.g., IKZF3) **(Figure 4E, left and 4F)**. The global shift towards the pro-inflammatory Th17 program was observed both in metabolic and non-metabolic transcripts, supporting the hypothesis that PGAM inhibition by EGCG effected a network-wide metabolic shift that mediated emergence of a pro-inflammatory Th17 program **(Figure 4F)**.

To verify that the effect of EGCG was mediated by a specific inhibition of PGAM (rather than an off-target effect) we conducted a carbon tracing assay in which the cell's medium was

supplemented with 13C-glucose **(Methods)**. PGAM inhibition with EGCG led to a sharp decrease (from 51% 13C ratio to 7% in Th17n and from 55% to 33% in Th17p) in 13C contents of 2PG (PGAM's product) but not 3PG (PGAM's substrate) or any other glycolytic metabolite that we were able to measure **(Figure 4G)**. Interestingly, 13C ratio of PEP (one step downstream of 2PG) was not changed as well. This suggests that the effect of the inhibitor is restricted (at least within glycolysis) to the PGAM reaction that lies directly downstream of 3PG.

As the serine biosynthesis pathway is more active in Th17p than in Th17n (**Figure 2C**) and lies directly downstream of 3PG (**Figure 4B**), we asked whether inhibiting serine biosynthesis can rescue the effect of PGAM inhibition. To this end, we treated Th17n cells with inhibitors to PGAM and PHGDH (phosphoglycerate dehydrogenase), alone or in combination. We found that further inhibiting PHDGH rescued the upregulation of Tbet and IFNg induced by EGCG but not its impact on IL-10 suppression (**Figure 4H**).

Taken together, an intra-population Compass analysis predicted that within the Th17n compartment, the glycolytic PGAM reaction inhibits, rather than promotes, pathogenicity. This prediction relied on heterogeneity within the Th17n population, yielding results that are contrary to those from inter-population comparisons of Th17 to Treg or Th17p to Th17n. EGCG specifically inhibited this reaction, and promoted a transcriptional state indicative of a more pro-inflammatory potential, as evidenced by a global shift in the transcriptome toward a Th17p-like profile. RNA-Seq further supported the hypothesis that EGCG mediates its effects by altering the cellular metabolic profile.

PGAM inhibition exacerbates, whereas G6PD inhibition ameliorates, Th17-mediated neuroinflammation *in vivo*

To test the functional relevance of the transcriptome shifts induced by EGCG and DHEA *in vivo*, we used the adoptive T cell transfer system, so that the effect of inhibitors is limited to T cells rather than all cells in the host. We generated Th17n and Th17p cells from naive CD4+ T cells isolated from 2D2 TCR-transgenic mice, with specificity for MOG 35-55, and transferred them into wildtype mice to induce EAE.

Consistent with Compass prediction, Th17p cells treated with DHEA reduced the severity of disease at peak of EAE in the recipients (**Figure 5A**). By the time the mice were sacrificed, however, the number of lesions in CNS was not significantly different (**Figure 5B**), and surviving mice showed no significant alterations in antigen-specific cytokine secretion in response to MOG, except for increased IL-2 in the DHEA treated group (**Supplementary Figure 5A**). More interestingly, and in agreement with Compass predictions, EGCG-treated Th17n cells induced EAE, albeit in mild form, whereas solvent treated cells failed to produce any consequential neuroinflammation **(Figure 5C)**. Recipients of EGCG-treated Th17n cells had a significantly higher

EAE incidence rate (10/12) compared with the control group (0/12, Fisher's exact $p = 1.1*10^{-4}$). Consistent with the clinical disease, histological analyses revealed an increased number of CNS lesions in in both the meninges and the parenchyma of mice that were injected with EGCG-treated Th17 cells (**Figure 5D**). While there was only a small difference in antigen-specific T cell proliferation (**Figure 5E**), there was a significant increase in secretion of IL-17, IL-17F, IL-22 and IL-6 (**Figure 5F** and **Supplementary Figure 5B**) in response to antigen in cells isolated from draining lymph node of mice transferred with EGCG-treated Th17n cells. As EGCG treated non-pathogenic Th17n cells induced only mild EAE, we asked whether EGCG will further enhance encephalitogenicity of Th17 cells if IL-23 is included in the differentiation cultures, which stabilizes the Th17 phenotype [79]–[82]. IL-23- treatment indeed enhanced EAE disease severity, but still Th17n cells treated with IL-23+EGCG induced significantly more severe EAE than their IL-23+solvent-treated counterparts (**Figure 5G**). Histopathology across all experiments revealed that EGCG treatment of Th17n cells promoted, whereas DHEA treatment Th17p cells restricted, optic neuritis/perineuritis in host mice (**Figure 5H**). Interestingly, mice transferred with EGCG-treated Th17 cells (Th17n or Th17n with IL-23) were the only experimental group to produce Wallerian degeneration in proximal spinal nerve roots (**Figure 5I, J)**.

In conclusion, Compass correctly predicted metabolic targets including glycolytic pathways whose deletion affected Th17 function. Importantly, it was able to pinpoint a glycolytic reaction that suppresses Th17 pathogenicity, which runs contrary to the current understanding that aerobic glycolysis as a whole is associated with a pro-inflammatory phenotype in Th17 cells.

Discussion

We presented Compass — a flux balance analysis (FBA) algorithm for the study of metabolic heterogeneity among cells based on single-cell transcriptome profiles and validated a number of predictions by metabolome and functional analyses. Compass successfully predicted metabolic targets in both central and ancillary pathways based on its network approach. These results support the power of transcriptomic-based FBA to make valid predictions in a mammalian system.

Glycolysis is a central regulator of T cell function. Compass predicted an association between aerobic glycolysis and Th17 pathogenicity, which accords with multiple previous results tying elevated glycolysis with T cell inflammatory functions. However, a Compass-based data-driven analysis based on scRNAseq unexpectedly revealed that not all glycolytic reactions promote the pro-inflammatory phenotype in Th17 cells. This result was obtained via an intra-population analysis of individual cells. It serves as a further example to the power of studying single-cell

heterogeneity within seemingly homogenous populations (here, Th17n), which allowed us to identify a novel regulator that would have otherwise been missed at a population level (here, a comparison of Th17p and Th17n). The computational prediction and the data corroborating it also demonstrate that despite the common assumption that glycolysis promotes pro-inflammatory functions in Th17 cells and other immune compartments [2], [3], [83]–[87], the role of glycolysis in induction of pro-inflammatory phenotypes may more nuanced [88], [89].

Static FBA algorithms assume that the system under consideration operates in chemical steady state [90]. Even under this assumption, there remains an infinite number of feasible flux distributions that satisfy the preset biochemical constraints. Therefore, most studies assume that the system (here, a cell) aims to optimize some metabolic function, usually production of biomass or ATP [91]. However, whereas such objectives may successfully predict phenotypes of a unicellular organism [92], they are ill-suited for studying mammalian cells [93]. To overcome this challenge, rather than optimizing a single metabolic objective function, Compass optimizes a set of objective functions, each estimating the degree to which a cell's transcriptome supports carrying the maximal theoretical flux through a given reaction. The result is a high dimensional representation of the cell's metabolic potential (one number per reaction). A biological signal (e.g., differences in reaction potential) can be detected in this high-dimension owing to the statistical power afforded by the large number of cells in a typical scRNA-Seq dataset. Nonetheless, there is no inherent limitation preventing one from applying Compass to study bulk (i.e., non-single-cell) transcriptomic data.

The database of metabolic reactions we used pertains to human cells, and as such our study does not address differences between human and mouse metabolism. In addition, the database provides a global view of the metabolic capabilities of a human cell, accrued from various sources and in diverse cell types. Not all reactions may be functional in a studied cell type, or under particular physiological conditions. This concern can be addressed to some extent by procedures for deriving organ-specific metabolic models [94]. Moreover, the metabolic state of a cell depends on the nutrients available in its environment, which are often poorly characterized. Here, our computations assume an environment rich with nutrients, which accords with the studied *in vitro* growth media. Modifying this to better represent physiological conditions should increase the algorithm's predictive capabilities, especially for cells derived *in vivo*, where nutrient scarcity may be a limiting factor, and nutrient availability may vary between tissues.

One of the outstanding challenges in the field of single cell genomics is translating the vast data sets presented in cell atlases into an actionable knowledge resource, i.e. using observed cell states to deduce molecular mechanisms and targets [16]. Compass was designed with this challenge in mind, and addresses it in the metabolic cellular subsystem, which can be tractably modeled *in silico*. In light of the wide appreciation of cellular metabolism as a critical regulator of

physiological processes in health and disease, we expect Compass to be useful in predicting cell metabolic states, as well as actionable metabolic targets, in diverse physiological and pathologic contexts.

Code and data availability

We are working on making an open-source and free for academic use implementation available on Github, and will update this preprint when it is available. In the meantime, code is available upon request. A GEO deposition of RNA-Seq data is underway as well.

Author Contributions

AW, CW, AR, VKK, and NY conceived and designed the study. AW, DD, and NY designed the algorithm and performed computational analyses. CW designed, performed, and analyzed experiments, with significant contributions by JAP, KP, and CC in LC/MS metabolomics, NRH and MH in Seahorse assays, and further assistance by SZ, JF, EAG and VPD. VKK and NY supervised the study. AW, CW, VKK, and NY wrote the manuscript with contributions from all authors. All authors read and approved the final manuscript

Acknowledgments

We thank Eytan Ruppin and Michael B. Cole for fruitful discussions and Christina Usher for artwork. NY and AW were supported by the Chan Zuckerberg Biohub and by a National Institute of Mental Health (NIMH) grant NIH5U19MH114821. This work was supported by grants from the National Institutes of Health (RO1NS30843, R01AI144166, R01NS045937, P01AI073748, P01AI039671, P01AI056299) awarded to VKK, and by a Career Transitional Fellowship from the National Multiple Sclerosis Society awarded to CW. JF was supported by a Max Kade fellowship awarded by the Austrian Acadamy of Science (ÖAW).

Declaration of Interests

VKK has an ownership interest and is a member of the SAB for Tizona Therapeutics, and is a co-founder of and has an ownership interest in Celsius Therapeutics. VKK is an inventor on patents related to Th17 cells and immunometabolism. VKK's interests were reviewed and managed by the Brigham and Women's Hospital and Partners Healthcare in accordance with their conflict of

interest policies. AR is a SAB member of ThermoFisher Scientific, Neogene Therapeutics, Asimov and Syros Pharmaceuticals. AR is a cofounder of and equity holder in Celsius Therapeutics and an equity holder in Immunitas. AW, CW, JF, AR, VKK, and NY are co-inventors on a provisional patent application directed to inventions relating to methods for modulating metabolic regulators of T cell pathogenicity as described in this manuscript, filed by The Broad Institute, Brigham and Women's Hospital, and the Regents of California. All other authors declare no competing interests.

Figure Legends

Figure 1 - Algorithm overview

(A) Computation of Compass scores matrix. Compass leverages prior knowledge on metabolic topology and stoichiometry (encoded in a GSMM, see main text) to analyze single-cell RNA-Seq expression. Briefly, it computes a reaction-penalties matrix, where the penalty of a given reaction is inversely proportional to the expression its respective enzyme-coding genes. The reaction-penalties matrix is the input to a set of flux-balance linear programs that produce a score for every reaction in every cell, namely the Compass score matrix. (B) To compute the reaction penalties matrix, Compass allows soft information sharing between a cell and its k-nearest neighbors to mitigate technical noise in single-cell library preparation. (C) Downstream analysis of the score matrix. Rows are hierarchically clustered into meta-reactions (agnostically of canonical pathway definitions). The scores are then amenable to common genomics procedures including differential expression of meta-reactions, detecting meta-reactions correlating with a phenotype of interest, dimensionality reduction, and data-driven network analysis (the latter pursued in the accompanying manuscript (Wang et al., biorxiv preprint)).

Figure 2 - Compass-based exploration of metabolic heterogeneity within the Th17 compartment

(A) The experimental system. Naive CD4+ T cells are collected and differentiated into Th17p or Th17n cells, which are IL-17+ T cells that cause severe or mild-to-none CNS autoimmunity upon adoptive transfer. Th17nu cells are Th17n cells which were not sorted by IL-17 and exhibit higher variability [29]. (B) PCA of the Compass scores matrix (restricted to core metabolism, see main text), with select top loadings shown. (C) Dots represent a single biochemical reactions, Cohen's d and Wilcoxon rank sum p values computed as described in the main text for a comparison of Th17p vs. Th17n. This computation is done over meta-reactions, and every meta-reaction is expanded into its constituent single reactions (**Methods**), each shown as a separate dot. Only

core reactions (**Methods**) are shown. Reactions are partitioned by Recon2 pathways; bottom-right panel groups together all Recon2 subsystems associated with amino-acid metabolism. (D) Spearman correlation of Compass scores of single reactions with the expression of pro-pathogenic (magenta) or pro-regulatory Th17n (green) genes [29], none of which is metabolic and therefore none of them directly serves as a Compass input. Only significant correlations (BH-adjusted $p < 0.1$) are shown in color and non-significant correlation coefficients are greyed out. The rows represent 489 meta-reactions that belong to core pathways (defined as Recon2 subsystems that have at least 3 core reactions), and significantly correlated (or anti-correlated) with at least one of the genes. Key reactions (rows) in pathways discussed in the manuscript are highlighted according to the meta-reaction to which they belong. (E) Dots represent single biochemical reactions, Cohen's d and Wilcoxon rank sum p values computed as described in panel C. Only core reactions are shown. Reactions are partitioned by Recon2 pathways. Reactions are colored by the sign of their Cohen d's statistic, and are opaque or transparent according to statistical significance.

Figure 3 - Differential usage of glycolysis and fatty acid oxidation by pathogenic an non-pathogenic Th17 cells

(A) A diagram of central carbon metabolism, overlaid with Compass prediction for differential potential activity between Th17p and Th17n. Differentially active reactions (BH-adjusted Wilcoxon $p < 0.1$) are colored in magenta (pro-Th17p) or green (pro-Th17n), non-significantly different reactions are colored in grey. (B) Th17n, Th17p and Treg cells were differentiated as described (**Methods**) and replated with Seahorse media at 68h for Seahorse assay. Extracellular acidification rate (ECAR) and oxygen consumption rate (OCR) are reported in response to mitostress test. (C) Th17p and Th17n cells were differentiated and harvested at 68h (left columns) or replated in fresh media with no TCR stimulation or cytokine for 15 minutes (right columns) and subject to LC/MS based metabolomics. (D) Cells were harvested as in C and pulsed with 13C-tagged glucose for 15 minutes. Shown is the ratio of 13C-tagged carbon out of the total carbon content associated with the metabolite (**Methods**). (E) Th17n and Th17p cells were measured for their oxygen consumption rate in the presence of control or 40uM etomoxir. (F) Th17n and Th17p cells from either WT or PDK4-deficient mice were differentiated as described (**Methods**) and replated with Seahorse media at 68h for Seahorse assay. Extracellular acidification rate (ECAR) is reported in response to mitostress test. (G) Number of differentially expressed (DE) genes between PDK4-deficient and WT cells as a function of the significance threshold. (H-I) WT and PDK4-/- mice were immunized with MOG35-55 to induce EAE. (H) EAE clinical score was followed for 21 days. (I) Cells were harvested from CNS at day 15 post immunization for intracellular cytokine or transcription factor analysis.

Figure 4 - An unexpected role for PGAM in mediating TGFb-induced Th17 pathogenicity

(A) Intra-population analysis in two biological replicates (the Th17n and Th17nu cell populations, see **Figure 2a**). Dots are single metabolic reactions, and axes denote their correlation with the pathogenic signature in the Th17nu and Th17n groups. Colors denote whether the reaction was decided as pro-inflammatory, pro-regulatory, or non-significantly (NS) associated with either state by the inter-population analysis. PGAM, GK, PKM, and G6PD are reactions discussed in the manuscript (see Figure 4b). SPT = serine-pyruvate transaminase (EC 2.6.1.51). Rev = reverse (backwards) direction. (**Methods**). (B) Schematics of central carbon metabolism, the highlighted magenta and green reactions are the two predicted to be most correlated and anti-correlated with the computational pathogenicity score within the Th17n compartment, respectively. Reported inhibitors of these reactions are denoted. (C) Effects of inhibiting candidate genes on Th17 cytokines as measured by flow cytometry are shown. Naïve T cells were differentiated under pathogenic (Th17p) and non-pathogenic (Th17n) Th17 cell conditions (**Methods**) in the presence of control solvent or inhibitors. Cells were pre-labeled with division dye and protein expression is reported for cells that have gone through one division (d1) to exclude arrested cells. (D) PCA of bulk RNA-Seq of d1 Th17 cells as in B. (E) Differential gene expression due to EGCG and DHEA treatment. Red and blue dots represent genes associated with the pro-pathogenic and pro-regulatory Th17 transcriptional programs, respectively (red genes are ones belonging either to the list of pro-pathogenic Th17 markers (Figure 2d and **Methods**) or to the Th17 pro-inflammatory covariation module defined by [29]; blue genes are similarly defined). Highly differential genes associated with surface receptors, cytokine activity, or that are otherwise of interest are labelled by name. (F) Histograms of the logFC per gene in differential expression of EGCG- vs. DMSO-treated cells. A separate histogram is shown for Th17p-associated (magenta), Th17n-associated (green), and non-significantly associated (grey) genes. Genes were partitioned into these three groups by differential expression in bulk RNA-Seq (same libraries as shown in panel D) between DMSO-treated Th17p and Th17n cells with significance threshold of BH-adjusted $p \leq 0.05$ and log2 fold-change ≥ 1.5 in absolute value. (G) ratio of 13C-tagged carbon to total carbon in Th17 cells cultured for 15 minutes in the presence of 13C-glucose. Three metabolites are shown: PGAM's substrate (3-phosphoglycerate), product (2-phosphoglycerate), and the next downstream metabolite along the glycolytic pathway (phosphoenolpyruvate). (H) Th17n cells were differentiated in the presence of solvent alone, EGCG, PHDGH inhibitor (PKUMDL-WQ-2101, **Methods**), or the combination. Cells were harvested at 96h for flow cytometric analysis.

Figure 5 - EGCG exacerbates and DHEA ameliorates Th17-induced EAE *in vivo*

2D2 TCR–transgenic Th17 cells were adoptively transferred after differentiation in vitro in the presence of an inhibitor or vehicle as indicated. (A, C, G) Clinical outcome of EAE; (B, D) Histological score based on cell infiltrates in meninges and parenchyma of CNS; (E, F) Draining lymph node (cervical) from respective mice were isolated and pulsed with increasing dose of MOG_{35-55} peptide for 3 days and (E) subjected to thymidine incorporation assay; or (F) measurement of cytokine secretion by Legendplex and flow cytometry. Concentrations were normalized through division by the respective response to no antigen control (H-I) Independent pathological report of CNS isolated from mice with EAE at end point (d35 for EGCG experiments; d28 for DHEA experiment); Optic nerves were not found in the histologic section from one animal in the EGCG+IL-23 group. (J) Representative histology of spinal cord and spinal nerve roots. There is greater meningeal inflammation and Wallerian degeneration (digestion chambers, arrows) in posterior spinal nerve roots in EGCG vs. Control mice. PC, posterior column; PH, posterior horn. Individual mouse numbers are indicated. The smaller panel shows VK 39875 mouse section at higher magnification. All are H. & E., 40X objective. Three similar experiments were performed.

Supplementary Figure Legends

Supplementary Figure 1

Cumulative distribution function (CDF) of number of reactions per meta-reaction.

Supplementary Figure 2

(A-E) PCA of Compass space restricted to core meta-reactions, see main text. (A) PC1 scores plotted against PC2 and PC3 scores. (B) Enrichment of metabolic pathways in the positive or negative directions of top principal components. Enrichment is computed with GSEA [95] over single reactions (rather than genes, as in the common applications). Colors are -log10(BH-adjusted p), truncated at 4, with p being the GSEA p value. Pathways correspond to Recon2 subsystems. (C) PC1 scores plotted against computational signatures of cellular metabolic activity and Th17 differentiation time course (**Methods**). (D) Spearman correlation of top PCs with known pro-pathogenic (magenta) and pro-regulatory (green) marker genes, none of which is metabolic.

Only significant correlations (BH-adjusted $p < 0.1$) are shown in color. (E) Spearman correlation of computational transcriptome signatures with the top principal components. Only significant correlations (BH-adjusted $p < 0.1$) are shown in color and non-significant correlation coefficients are greyed out. See **Methods** for signature computation. (F) Same analysis as shown in Figure 2c, but showing all reactions (and not just ones belonging to certain pathways, as in the main figure). (G) We computed a pro-pathogenic score for each reaction by taking the ratio of pro-pathogenic and pro-regulatory markers with which it correlates and anti-correlates, respectively (BH-adjusted $p < 0.1$ for a Spearman correlation) out of the 23 marker genes (listed in **Figure 2d** and **Methods**). Similarly, we computed pro-regulatory reaction scores. Only core reactions are shown. (H) Same analysis as shown in Figure 2E, only at the gene expression level (and not reaction level based on Compass scores). Genes are grouped by KEGG pathways (and may be annotated as belonging to more than one pathway).

Supplementary Figure 3

(A) Parallel of main figures 3c showing also 3h after fresh media pulse. (B) The glycolysis pathway, as shown in main figure 4a, highlighting PDH and associated reactions. (C) PDK4 transcript expression in the experiment described in main Figure 4c. (D) Dots are transcriptomic computational signatures (**Methods**), axes correspond to the fold-change in the signature's value in comparisons of PDK4-deficient cells vs. WT cells in Th17p (x-axis) and Th17n (y-axis). (E) Th17 cells from PDK4-/- and WT mice were subject to LC/MS metabolomics as in main Figure 3c, having been replated for 15 minutes. (F) metabolites associated with amino-acid metabolic pathways in the assay described in main Figure 3c.

Supplementary Figure 4

(A) Same data as shown in Figure 4a, highlighting the reactions with significant adjusted Fisher p value in the intra-population analysis; every reaction is assigned a combined Fisher p-value of the two p-values measuring the significance of the correlation with the two axes (**Methods**). Search space was limited to core reactions. (B-C) Hypergeometric enrichment of the targets identified by the inter-population analysis (reactions with differential potential activity between Th17p and Th17n, decided by a BH-adjusted p cutoff) in targets identified by the intra-population analysis (reactions identified by a BH-adjusted Fisher p cutoff) while varying the cutoffs. (D) Supernatant from Th17 cell cultures performed for main Figure 4c are harvested for cytokine analysis using Legendplex.

Supplementary Figure 5

Cytokine secretion after three days of culture with increasing dose of MOG_{35-55} peptide from cells isolated from draining lymph node (cervical) of mice transferred with (A) methanol or DHEA treated Th17p cells as in Figure 5A or (B) DMSO or EGCG as in Figure 5C. Concentrations were normalized through division by the respective response to no antigen control.

References

[1] G. S. Hotamisligil, “Foundations of Immunometabolism and Implications for Metabolic Health and Disease,” *Immunity*, vol. 47, no. 3, pp. 406–420, Sep. 2017.

[2] L. A. J. O’Neill, R. J. Kishton, and J. Rathmell, “A guide to immunometabolism for immunologists,” *Nat. Rev. Immunol.*, vol. 16, no. 9, pp. 553–565, Sep. 2016.

[3] R. I. K. Geltink, R. L. Kyle, and E. L. Pearce, “Unraveling the Complex Interplay Between T Cell Metabolism and Function,” *Annu. Rev. Immunol.*, vol. 36, pp. 461–488, Apr. 2018.

[4] D. G. Russell, L. Huang, and B. C. VanderVen, “Immunometabolism at the interface between macrophages and pathogens,” *Nat. Rev. Immunol.*, Jan. 2019.

[5] M. D. Buck, R. T. Sowell, S. M. Kaech, and E. L. Pearce, “Metabolic Instruction of Immunity,” *Cell*, vol. 169, no. 4, pp. 570–586, May 2017.

[6] P.-C. Ho and S. M. Kaech, “Reenergizing T cell anti-tumor immunity by harnessing immunometabolic checkpoints and machineries,” *Curr. Opin. Immunol.*, vol. 46, pp. 38–44, Jun. 2017.

[7] N. M. Chapman, M. R. Boothby, and H. Chi, “Metabolic coordination of T cell quiescence and activation,” *Nat. Rev. Immunol.*, Aug. 2019.

[8] P. W. F. Karmaus *et al.*, “Metabolic heterogeneity underlies reciprocal fates of TH17 cell stemness and plasticity,” *Nature*, vol. 565, no. 7737, pp. 101–105, Jan. 2019.

[9] D. R. Green, L. Galluzzi, and G. Kroemer, “Metabolic control of cell death,” *Science*, vol. 345, no. 6203, Sep. 2014.

[10] J. D. Orth, I. Thiele, and B. O. Palsson, “What is flux balance analysis?,” *Nat. Biotechnol.*, vol. 28, no. 3, pp. 245–248, Mar. 2010.

[11] E. J. O’Brien, J. M. Monk, and B. O. Palsson, “Using Genome-scale Models to Predict Biological Capabilities,” *Cell*, vol. 161, no. 5, pp. 971–987, May 2015.

[12] N. E. Lewis, H. Nagarajan, and B. O. Palsson, “Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods,” *Nat. Rev. Microbiol.*, vol. 10, no. 4, pp. 291–305, Apr. 2012.

[13] B. Ø. Palsson, *Systems Biology: Constraint-based Reconstruction and Analysis*, 2nd ed. edition. Cambridge University Press, 2015.

[14] A. Bordbar, J. M. Monk, Z. A. King, and B. O. Palsson, “Constraint-based models predict metabolic and associated cellular functions,” *Nat. Rev. Genet.*, vol. 15, no. 2, pp. 107–120, Feb. 2014.

[15] A. Wagner, A. Regev, and N. Yosef, “Revealing the vectors of cellular identity with single-cell genomics,” *Nat. Biotechnol.*, vol. 34, no. 11, pp. 1145–1160, Nov. 2016.

[16] A. Tanay and A. Regev, “Scaling single-cell genomics from phenomenology to mechanism,” *Nature*, vol. 541, no. 7637, pp. 331–338, Jan. 2017.

[17] N. B. Ben-Moshe *et al.*, “Predicting bacterial infection outcomes using single cell RNA-sequencing analysis of human immune cells,” *Nat. Commun.*, vol. 10, no. 1, p. 3266, 2019.

[18] F. A. Vieira Braga *et al.*, “A cellular census of human lungs identifies novel cell states in health and in asthma,” *Nat. Med.*, vol. 25, no. 7, pp. 1153–1163, Jul. 2019.

[19] E. Azizi *et al.*, “Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment,” *Cell*, vol. 174, no. 5, pp. 1293-1308.e36, Aug. 2018.

[20] M. Sade-Feldman *et al.*, "Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma," *Cell*, vol. 175, no. 4, pp. 998-1013.e20, Nov. 2018.
[21] H. Keren-Shaul *et al.*, "A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease," *Cell*, vol. 169, no. 7, pp. 1276-1290.e17, Jun. 2017.
[22] R. Vento-Tormo *et al.*, "Single-cell reconstruction of the early maternal-fetal interface in humans," *Nature*, vol. 563, no. 7731, pp. 347–353, Nov. 2018.
[23] F. Paul *et al.*, "Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors," *Cell*, vol. 163, no. 7, pp. 1663–1677, Dec. 2015.
[24] R. Soldatov *et al.*, "Spatiotemporal structure of cell fate decisions in murine neural crest," *Science*, vol. 364, no. 6444, Jun. 2019.
[25] R. J. Miragaia *et al.*, "Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation," *Immunity*, vol. 50, no. 2, pp. 493-504.e7, Feb. 2019.
[26] F. Zanini *et al.*, "Virus-inclusive single-cell RNA sequencing reveals the molecular signature of progression to severe dengue," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 52, pp. E12363–E12369, Dec. 2018.
[27] C. C. Brown *et al.*, "Transcriptional Basis of Mouse and Human Dendritic Cell Heterogeneity," *Cell*, vol. 179, no. 4, pp. 846-863.e24, Oct. 2019.
[28] H. Van Hove *et al.*, "A single-cell atlas of mouse brain macrophages reveals unique transcriptional identities shaped by ontogeny and tissue environment," *Nat. Neurosci.*, vol. 22, no. 6, pp. 1021–1035, Jun. 2019.
[29] J. T. Gaublomme *et al.*, "Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity," *Cell*, vol. 163, no. 6, pp. 1400–1412, Dec. 2015.
[30] T. Korn, E. Bettelli, M. Oukka, and V. K. Kuchroo, "IL-17 and Th17 Cells," *Annu. Rev. Immunol.*, vol. 27, no. 1, pp. 485–517, Mar. 2009.
[31] L. A. Tesmer, S. K. Lundy, S. Sarkar, and D. A. Fox, "Th17 cells in human disease," *Immunol. Rev.*, vol. 223, pp. 87–113, Jun. 2008.
[32] S. L. Gaffen, N. Hernández-Santos, and A. C. Peterson, "IL-17 signaling in host defense against Candida albicans," *Immunol. Res.*, vol. 50, no. 2–3, pp. 181–187, Aug. 2011.
[33] L. Romani, "Immunity to fungal infections," *Nat. Rev. Immunol.*, vol. 11, no. 4, pp. 275–288, Apr. 2011.
[34] Y. Lee *et al.*, "Induction and molecular signature of pathogenic TH17 cells," *Nat. Immunol.*, vol. 13, no. 10, pp. 991–999, Oct. 2012.
[35] B. Stockinger and S. Omenetti, "The dichotomous nature of T helper 17 cells," *Nat. Rev. Immunol.*, vol. 17, no. 9, pp. 535–544, Sep. 2017.
[36] X. Wu, J. Tian, and S. Wang, "Insight Into Non-Pathogenic Th17 Cells in Autoimmune Diseases," *Front. Immunol.*, vol. 9, p. 1112, May 2018.
[37] C. Wang *et al.*, "CD5L/AIM Regulates Lipid Biosynthesis and Restrains Th17 Cell Pathogenicity," *Cell*, vol. 163, no. 6, pp. 1413–1427, Dec. 2015.
[38] A. Regev *et al.*, "Science forum: the human cell atlas," *Elife*, vol. 6, p. e27041, 2017.
[39] V. Svensson, R. Vento-Tormo, and S. A. Teichmann, "Exponential scaling of single-cell RNA-seq in the past decade," *Nat. Protoc.*, vol. 13, no. 4, pp. 599–604, Apr. 2018.
[40] D. J. Puleston, M. Villa, and E. L. Pearce, "Ancillary Activity: Beyond Core Metabolism in Immune Cells," *Cell Metab.*, vol. 26, no. 1, pp. 131–141, Jul. 2017.

[41] I. Thiele *et al.*, “A community-driven global reconstruction of human metabolism,” *Nat. Biotechnol.*, vol. 31, no. 5, pp. 419–425, May 2013.
[42] C. A. Vallejos, J. C. Marioni, and S. Richardson, “BASiCS: Bayesian Analysis of Single-Cell Sequencing Data,” *PLoS Comput. Biol.*, vol. 11, no. 6, p. e1004333, Jun. 2015.
[43] R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, and A. Regev, “Spatial reconstruction of single-cell gene expression data,” *Nat. Biotechnol.*, vol. 33, no. 5, pp. 495–502, May 2015.
[44] A. T. L. Lun, K. Bach, and J. C. Marioni, “Pooling across cells to normalize single-cell RNA sequencing data with many zero counts,” *Genome Biol.*, vol. 17, no. 1, pp. 1–14, 2016.
[45] L. Haghverdi, A. T. L. Lun, M. D. Morgan, and J. C. Marioni, “Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors,” *Nat. Biotechnol.*, vol. 36, no. 5, pp. 421–427, Jun. 2018.
[46] F. Wagner, Y. Yan, and I. Yanai, “K-nearest neighbor smoothing for high-throughput single-cell RNA-Seq data,” *bioRxiv*, p. 217737, 09-Apr-2018.
[47] M. Huang *et al.*, “SAVER: gene expression recovery for single-cell RNA sequencing,” *Nat. Methods*, vol. 15, no. 7, pp. 539–542, Jul. 2018.
[48] D. van Dijk *et al.*, “Recovering Gene Interactions from Single-Cell Data Using Data Diffusion,” *Cell*, vol. 174, no. 3, pp. 716-729.e27, Jul. 2018.
[49] Y. Baran *et al.*, “MetaCell: analysis of single-cell RNA-seq data using K-nn graph partitions,” *Genome Biol.*, vol. 20, no. 1, p. 206, Oct. 2019.
[50] D. Grün, “Revealing dynamics of gene expression variability in cell state space,” *Nat. Methods*, Nov. 2019.
[51] A. Fabregat *et al.*, “The Reactome Pathway Knowledgebase,” *Nucleic Acids Res.*, vol. 46, no. D1, pp. D649–D655, Jan. 2018.
[52] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, “KEGG: new perspectives on genomes, pathways, diseases and drugs,” *Nucleic Acids Res.*, vol. 45, no. D1, pp. D353–D361, Jan. 2017.
[53] K. Ghoreschi *et al.*, “Generation of pathogenic T(H)17 cells in the absence of TGF-β signalling,” *Nature*, vol. 467, no. 7318, pp. 967–971, Oct. 2010.
[54] N. Yosef *et al.*, “Dynamic regulatory network controlling TH17 cell differentiation,” *Nature*, vol. 496, no. 7446, pp. 461–468, Apr. 2013.
[55] L. Z. Shi *et al.*, “HIF1alpha-dependent glycolytic pathway orchestrates a metabolic checkpoint for the differentiation of TH17 and Treg cells,” *J. Exp. Med.*, vol. 208, no. 7, pp. 1367–1376, Jul. 2011.
[56] R. D. Michalek *et al.*, “Cutting edge: distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4+ T cell subsets,” *J. Immunol.*, vol. 186, no. 6, pp. 3299–3303, Mar. 2011.
[57] V. A. Gerriets *et al.*, “Metabolic programming and PDHK1 control CD4+ T cell subsets and inflammation,” *J. Clin. Invest.*, vol. 125, no. 1, pp. 194–207, Jan. 2015.
[58] A. K. Jha *et al.*, “Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization,” *Immunity*, vol. 42, no. 3, pp. 419–430, Mar. 2015.
[59] E. L. Mills *et al.*, “Succinate Dehydrogenase Supports Metabolic Repurposing of Mitochondria to Drive Inflammatory Macrophages,” *Cell*, vol. 167, no. 2, pp. 457-470.e13, Oct. 2016.

[60] L. Shi, Q. Jiang, Y. Bushkin, S. Subbian, and S. Tyagi, “Biphasic Dynamics of Macrophage Immunometabolism during Mycobacterium tuberculosis Infection,” *MBio*, vol. 10, no. 2, Mar. 2019.
[61] E. Mills and L. A. J. O’Neill, “Succinate: a metabolic signal in inflammation,” *Trends Cell Biol.*, vol. 24, no. 5, pp. 313–320, May 2014.
[62] L. Berod *et al.*, “De novo fatty acid synthesis controls the fate between regulatory T and T helper 17 cells,” *Nat. Med.*, vol. 20, no. 11, pp. 1327–1333, Nov. 2014.
[63] B. Raud *et al.*, “Etomoxir Actions on Regulatory and Memory T Cells Are Independent of Cpt1a-Mediated Fatty Acid Oxidation,” *Cell Metab.*, vol. 28, no. 3, pp. 504-515.e7, Sep. 2018.
[64] M. S. Sundrud *et al.*, “Halofuginone Inhibits T(H)17 Cell Differentiation by Activating the Amino Acid Starvation Response,” *Science*, vol. 324, no. 5932, pp. 1334–1338, Jun. 2009.
[65] A. S. Divakaruni *et al.*, “Etomoxir Inhibits Macrophage Polarization by Disrupting CoA Homeostasis,” *Cell Metab.*, vol. 28, no. 3, pp. 490-503.e7, Sep. 2018.
[66] M. Kono *et al.*, “Pyruvate dehydrogenase phosphatase catalytic subunit 2 limits Th17 differentiation,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 37, pp. 9288–9293, Sep. 2018.
[67] P. Wu, P. V. Blair, J. Sato, J. Jaskiewicz, K. M. Popov, and R. A. Harris, “Starvation increases the amount of pyruvate dehydrogenase kinase in several mammalian tissues,” *Arch. Biochem. Biophys.*, vol. 381, no. 1, pp. 1–7, Sep. 2000.
[68] P. Wu, J. M. Peters, and R. A. Harris, “Adaptive increase in pyruvate dehydrogenase kinase 4 during starvation is mediated by peroxisome proliferator-activated receptor alpha,” *Biochem. Biophys. Res. Commun.*, vol. 287, no. 2, pp. 391–396, Sep. 2001.
[69] L. F. Gemta *et al.*, “Impaired enolase 1 glycolytic activity restrains effector functions of tumor-infiltrating CD8+ T cells,” *Sci Immunol*, vol. 4, no. 31, Jan. 2019.
[70] K. E. Beckermann, S. O. Dudzinski, and J. C. Rathmell, “Dysfunctional T cell metabolism in the tumor microenvironment,” *Cytokine Growth Factor Rev.*, vol. 35, pp. 7–14, Jun. 2017.
[71] J. P. Rhoads, A. S. Major, and J. C. Rathmell, “Fine tuning of immunometabolism for the treatment of rheumatic diseases,” *Nat. Rev. Rheumatol.*, vol. 13, no. 5, pp. 313–320, May 2017.
[72] C.-H. Chang *et al.*, “Posttranscriptional Control of T Cell Effector Function by Aerobic Glycolysis,” *Cell*, vol. 153, no. 6, pp. 1239–1251, Jun. 2013.
[73] P.-C. Ho *et al.*, “Phosphoenolpyruvate Is a Metabolic Checkpoint of Anti-tumor T Cell Responses,” *Cell*, vol. 162, no. 6, pp. 1217–1228, Sep. 2015.
[74] A. L. Doedens *et al.*, “Hypoxia-inducible factors enhance the effector responses of CD8(+) T cells to persistent antigen,” *Nat. Immunol.*, vol. 14, no. 11, pp. 1173–1182, Nov. 2013.
[75] D. K. Finlay *et al.*, “PDK1 regulation of mTOR and hypoxia-inducible factor 1 integrate metabolism and migration of CD8+ T cells,” *J. Exp. Med.*, vol. 209, no. 13, pp. 2441–2453, Dec. 2012.
[76] A. N. Macintyre *et al.*, “The glucose transporter Glut1 is selectively essential for CD4 T cell activation and effector function,” *Cell Metab.*, vol. 20, no. 1, pp. 61–72, Jul. 2014.
[77] M. Kono *et al.*, “Pyruvate kinase M2 is requisite for Th1 and Th17 differentiation,” *JCI Insight*, vol. 4, no. 12, Jun. 2019.
[78] M. Peng, N. Yin, S. Chhangawala, K. Xu, C. S. Leslie, and M. O. Li, “Aerobic glycolysis promotes T helper 1 cell differentiation through an epigenetic mechanism,” *Science*, vol. 354, no. 6311, pp. 481–484, Oct. 2016.

[79] A. Awasthi *et al.*, “Cutting edge: IL-23 receptor gfp reporter mice reveal distinct populations of IL-17-producing cells,” *J. Immunol.*, vol. 182, no. 10, pp. 5904–5908, May 2009.

[80] M. J. McGeachy *et al.*, “The interleukin 23 receptor is essential for the terminal differentiation of interleukin 17-producing effector T helper cells in vivo,” *Nat. Immunol.*, vol. 10, no. 3, pp. 314–324, Mar. 2009.

[81] L. Zhou *et al.*, “IL-6 programs T(H)-17 cell differentiation by promoting sequential engagement of the IL-21 and IL-23 pathways,” *Nat. Immunol.*, vol. 8, no. 9, pp. 967–974, Sep. 2007.

[82] S. Aggarwal, N. Ghilardi, M.-H. Xie, F. J. de Sauvage, and A. L. Gurney, “Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17,” *J. Biol. Chem.*, vol. 278, no. 3, pp. 1910–1914, Jan. 2003.

[83] E. L. Pearce, M. C. Poffenberger, C.-H. Chang, and R. G. Jones, “Fueling Immunity: Insights into Metabolism and Lymphocyte Function,” *Science*, vol. 342, no. 6155, Oct. 2013.

[84] E. J. Pearce and B. Everts, “Dendritic cell metabolism,” *Nat. Rev. Immunol.*, vol. 15, no. 1, pp. 18–29, Jan. 2015.

[85] N. J. MacIver, R. D. Michalek, and J. C. Rathmell, “Metabolic Regulation of T Lymphocytes,” *Annu. Rev. Immunol.*, vol. 31, no. 1, pp. 259–283, Mar. 2013.

[86] L. A. J. O’Neill and E. J. Pearce, “Immunometabolism governs dendritic cell and macrophage function,” *J. Exp. Med.*, vol. 213, no. 1, pp. 15–23, Jan. 2016.

[87] K. L. O’Brien and D. K. Finlay, “Immunometabolism and natural killer cell responses,” *Nat. Rev. Immunol.*, vol. 19, no. 5, pp. 282–290, May 2019.

[88] J. Van den Bossche, L. A. O’Neill, and D. Menon, “Macrophage Immunometabolism: Where Are We (Going)?,” *Trends Immunol.*, vol. 38, no. 6, pp. 395–406, Jun. 2017.

[89] R. Newton, B. Priyadharshini, and L. A. Turka, “Immunometabolism of regulatory T cells,” *Nat. Immunol.*, vol. 17, no. 6, pp. 618–625, Jun. 2016.

[90] A. Varma and B. O. Palsson, “Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use,” *Nature Biotechnology*, vol. 12, no. 10, pp. 994–998, Oct. 1994.

[91] C. Damiani *et al.*, “Integration of single-cell RNA-seq data into population models to characterize cancer metabolism,” *PLoS Comput. Biol.*, vol. 15, no. 2, p. e1006733, Feb. 2019.

[92] N. E. Lewis *et al.*, “Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models,” *Mol. Syst. Biol.*, vol. 6, Jul. 2010.

[93] M. Adler, Y. Korem Kohanim, A. Tendler, A. Mayo, and U. Alon, “Continuum of Gene-Expression Profiles Provides Spatial Division of Labor within a Differentiated Cell Type,” *Cell Syst*, vol. 8, no. 1, pp. 43-52.e5, Jan. 2019.

[94] S. Opdam, A. Richelle, B. Kellman, S. Li, D. C. Zielinski, and N. E. Lewis, “A Systematic Evaluation of Methods for Tailoring Genome-Scale Metabolic Models,” *Cell Systems*, vol. 4, no. 3, pp. 318-329.e6, Mar. 2017.

[95] A. Subramanian *et al.*, “Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 43, pp. 15545–15550, Oct. 2005.

1 Experimental procedures

1.1 T cell differentiation culture

Naïve CD4+CD44-CD62L+CD25- T cells were sorted using BD FACSAria sorter and activated with plate-bound anti-CD3 and anti-CD28 antibodies (both at 1μg/ml) in the presence of cytokines at a concentration of 0.5×10^6 cells/ml. For Th17 differentiation: 2ng/ml of rhTGFb1, 25ng/ml rmIL-6, 20ng/ml rmIL-1b (all from Miltenyi Biotec) and 20ng/ml rmIL-23 (R & D systems) were used at various combinations as specified in figures. For differentiation experiments, cells were harvested at 68 hours for RNA analysis and 72-96h for flow cytometry analysis and Seahorse assay.

Seahorse assay was performed and seahorse media was prepared following manufacturer instructions (Agilent). Cells were re-stimulated with PMA/ionomycin for four hours in the presence of brefaldin and monensin before analysis for cytokines by intracellular cytokine staining. Cytokine concentrations in supernatants of *in vitro* cultures were analyzed by the LegendPlex Mouse Th Cytokine Panel (13-plex) (BioLegend) according to the manufacturer's instructions and analyzed on a FACS LSR II (BD Biosciences).

1.2 Smart-Seq Single-cell RNA sequencing

Full experimental details are given in [1]. Briefly, we sequenced CD4+ naive T cells 48hrs post polarization under one of these conditions, ultimately retaining after quality tests 130 Th17n cells unsorted for IL-17 (denoted Th17nu in the present manuscript), 151 IL-17A/GFP+ Th17n cells, and 139 IL-17A/GFP+ Th17p cells. Unlike [1], in the present study we analyzed the unsorted and sorted cells independently from one another. The sorted cells (Th17n and Th17p) were used for the inter-population analysis, and the unsorted cells (Th17nu) were used for the intra-population analysis.

1.3 Estimation of transcript abundance from RNA sequencing

We aligned single-cell SMART-Seq libraries with Bowtie2, quantified TPM gene expression with RSEM, and performed QC as we described in detail in a previous publication [2]. This computational pipeline is a massively revised and updated version of the one originally used to analyze these libraries [1]. Batch effects and other nuisance factors were normalized with a model chosen empirically with SCONE [3]. Bulk RNA-Seq were processed with a modified variant of the same pipeline.

1.4 Differential gene expression

For the Smart-Seq libraries, due the absence of UMIs in the dataset, differentially expressed genes were called through a linear model fitted to TPM values

with the `limma` R package and with a mean-variance trend added to the empirical Bayes prior [4]. For the bulk RNA libraries, differentially expressed genes were called with `limma-trend` or `limma-voom` [5] depending on the variance of library sizes, as recommended in the `limma` package manual [6].

1.5 Inhibitors

All chemical inhibitors were purchased from Sigma with the exception of EGCG (Selleck Chemicals) and tested in a wide range of dose (20nM-200uM) on Th17 cells. The lowest dose that resulted in minimal impact on cell viability is used for functional evaluation: EGCG, the inhibitor for PGAM1 [7], was used at 20-50uM; DHEA, inhibitor for G6PD [8], was used at 50uM; DCA, inhibitor for GK [9], [10], was used at 40uM; Shikonin, inhibitor for PKM2 [11], [12], was used at 10uM; and PKUMDL-WQ-2101, inhibitor for PHGDH [13], was used at 12.5uM.

1.6 Mice

C57BL/6 wildtype (WT) and PDK4$^{-/-}$ mice were obtained from Jackson Laboratory (Bar Harbor, ME). WT 2D2 transgenic mice were bred in house. All experiments were performed in accordance to the guidelines outlined by the Harvard Medical Area Standing Committee on Animals at the Harvard Medical School (Boston, MA).

1.7 Experimental Autoimmune Encephalomyelitis (EAE)

For adoptive transfer EAE, naïve T cells (CD4+CD44-CD62L+CD25-) were isolated from 2D2 TCR-transgenic mice and activated with anti-CD3 (1μg /ml) and anti-CD28 (1μg /ml) in the presence of differentiation cytokines for 68h. Cells were rested for 2 days and restimulated with plate-bound anti-CD3 (0.5μg /ml for pathogenic condition; 1μg /ml for non-pathogenic condition) and anti-CD28 (1μg /ml) for 2 days prior to transfer. Equal number (2 to 8 million) cells were transferred per mouse intravenously. EAE is scored as previously published [14].

1.8 LC/MS metabolomics and carbon tracing

1.8.1 Assays

For untargeted metabolomics, Th17 cells were differentiated as described. Culture media were snap frozen. Cells were harvested at 96h. 1×10^6 cells per sample were snap frozen and extracted in either 80% methanol (for fatty acids and oxylipids) or isopropanol (for polar and nonpolar lipids). Two liquid chromatography tandem mass spectrometry (LC-MS) methods were used to measure fatty acids and lipids in cell extracts.

For carbon tracing experiments Th17 cells were differentiated as described. Thereafter, cells were washed and cultured in media supplemented with 8 mM [U-13C]-glucose for 15min or 3hrs.

1.8.2 Statistical analysis

Differentially abundant metabolites were found with Student's t-test and a significance threshold of BH-adjusted $p < 0.1$.

To find metabolites with differential 13C relative abundance, we computed the ratio $y_{i,j}$ of 13C out of the total carbon contents for each metabolite i in sample j. Let $|C_i|$ be the number of carbon atoms in metabolite i, and let $x_{c,i,j}$ be the measured signal of metabolite i in sample j (subsequent to all normalization and QC procedures) in which there are exactly c 13C atoms. We define the 13C/C ratio:

$$y_{i,j} = \frac{\sum_{t=0}^{|C_i|} t \cdot x_{t,i,j}}{|C_i| \cdot \sum_{t=0}^{|C_i|} x_{t,i,j}}$$

2 Downstream analysis of Compass scores

2.1 Core metabolic reactions and meta-reactions

In this work, we define we core metabolism based on reaction metadata included in the Recon2 database. Recon2 assigns a confidence score to each reaction based on the level of evidence supporting it between 1 (no evidence) and 4 (biochemical evidence), with 0 denoting reactions whose confidence was not evaluated. Since pathways generally considered part of primary metabolism are also the best studied ones, we define a reaction as belonging to core metabolism if (a) its Recon2 confidence is either 0 or 4; *and* (b) it is annotated with an EC (Enzyme Commission) number. We chose to label reactions with unevaluated confidence (*i.e.*, Recon2 confidence score of 0) as part of core metabolism because some of them were found to be key reactions in primary metabolic pathways based on manual correction. Our definition of core metabolism is equivalent to taking the set of all metabolic reactions in Recon2, but excluding reactions that either don't have an annotated EC number or for which the Recon2 curators explicitly specified they do not have direct biochemical support.

We define a meta-reaction as belonging to core metabolism if it contains at least one core reaction. Metabolic genes are defined as the set of genes annotated in Recond2 (Section 4.7)

2.2 Inter-population analysis: finding reactions with differential potential activity

To test for differential potential-activity of reactions based on Compass predictions, we computed for each meta-reaction M the Wilcoxon's rank sum between the Compass scores of M in the two populations of interest (here, Th17p and

Th17n). Effect size were further assessed with Cohen's d statistic, defined as the difference between the sample means over the pooled sample standard deviation. Let n_1, x_1, s_1 be the number of observations in population 1, and the sample mean and standard deviation of their scores in a given meta-reaction, respectively (with a similar notation for population 2). Then

$$d = \frac{(\bar{x_1} - \bar{x_2})}{s}$$

with

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}).$$

The resulting p values are adjusted with the Benjamini-Hochberg (BH) method. Note that so far, the computation was done for meta-reactions. We assigned all reactions $r \in M$ the Cohen's d and Wilcoxon's p value that were computed for M. We call a reaction differentially active if its adjusted p is smaller than 0.1. The computation was done on all reactions in the network (namely, both core and non-core reactions).

2.3 Transcriptomic signatures

2.3.1 Th17 Pathogenicity and other T cell state signatures

We used a transcriptomic signature that we have previously shown to capture a Th17 cell's pathogenic capacity [1], [15]. Briefly, for each cell compute the average z-scored expression ($\log(1 + \text{TPM})$) of pro-pathogenic markers (CASP1, CCL3, CCL4, CCL5, CSF2, CXCL3, GZMB, ICOS, IL22, IL7R, LAG3, LGALS3, LRMP, STAT4, TBX21) and pro-regulatory markers (AHR, IKZF3, IL10, IL1RN, IL6ST, IL9, MAF), with the latter group multiplied by -1.

2.3.2 A compendium of T cell state signatures

A compendium of T cell state transcriptomic signatures was described in [1]. Every signature consists of two gene subsets: a set of positively associated genes and an optionally empty set of negatively associated genes. A scalar signature value is computed for every cell based on its transcriptome profile as described above for pathogenicity. Signatures that are based on KEGG [16] pathways or similar resources are constructed by defining the set of positively-associated genes as the ones belonging to the pathway and defining the set of negatively-associated genes as an empty set.

2.3.3 Total metabolic activity of a cell

We defined the total metabolic activity of a cell as the sum expression of metabolic enzyme coding genes over the sum expression of all protein coding genes in log-scale TPM (transcripts per million) units. We computed the partial

correlation between this quantity and cell PC1 coordinates, while controlling for the sum expression of all protein coding genes in the cells (the aforementioned divisor) to verify the correlation does not arise from the ratio of protein-coding to non-protein coding RNA in the RNA libraries. The correlation was more significant when not controlling for the covariate (Pearson rho $= 0.56$, $p < 3{\cdot}10^{-16}$).

2.3.4 Late-stage Th17 differentiation

We defined a transcriptomic signature for late-stage differentiation of Th17 cells based on microarray data from [17]. We assigned microarrays into three differentiation stages as described in that paper into early (up to 4h), intermediate (6-16h) and late (20-72h) and fitted with the `limma` R package a linear model for the discrete 3-level stage covariate. We called differentially expressed genes (BH-adjusted $p < 0.05$ and log2 fold-change ≥ 3) and used them to define a transcriptomic signature as described above.

computed genes differentially expressed between the late

2.4 Intra-population analysis: correlation with a quantitative cellular trait

Here, the population was Th17n cells, in one of two biological replicates, and the quantitative trait was a transcriptomic pathogenicity signature. For every meta-reaction M, we computed Spearman correlation of its Compass scores and the pathogenicity scores across all cells. We assigned all reactions $r \in M$ the Spearman correlation and its statistical significance that were computed for M. Note that the division of reactions to meta-reactions is dataset-specific and therefore a reaction can belong to different meta-reactions in each of the replicates. So far, the computation was done independently for the two biological replicates. We then computed for each reaction r the Fisher combined p value of the two p values corresponding to the statistical significance of its Spearman correlation with the pathogenicity scores in the two replicates. The combined Fisher p values were adjusted with the Benjamini-Hochberg (BH) method. We call a reaction significantly correlated (or anti-correlated) with the pathogenicity score if its adjusted Fisher combined p is smaller than 0.1. The search space was limited to core reactions.

2.5 Manual curation of central carbon predictions

We manually curated the significant predictions of the central carbon metabolism pathways discussed in the manuscript (glycolyss, TCA cycle, and fatty acid synthesis/oxidation). Recon2 takes account of metabolite localization, and reactions may be functional in more than one cellular compartment. For every reaction, we picked the prediction corresponding to the pertinent cellular compartment (here, cytosol or mitochondria, as shown in **Figure 3a**). Note that Compass operates independently on the forward and backward directions of ev-

ery reaction, and that the direction is denoted in the pathway diagrams of this manuscript.

3 Running Compass

3.1 Software

We are working on making an open-source and free for academic use implementation available on Github, and will update this preprint when it is available. In the meantime, code is available upon request. A GEO deposition of RNA-Seq data is underway as well.

The algorithm is highly parallelizable. It currently supports execution on multiple threads in a single machine, submission to a Torque queue, and execution on a single machine on Amazon Web Services (AWS). The current implementation relies on the IBM ILOG CPLEX Optimization Studio, which is free for academic use.

3.2 Gene expression input

The main input is gene expression matrix G in which rows correspond to genes and columns to RNA libraries. We assume that G is (i) already normalized to remove batch and other nuisance effects; (ii) scaled to CPMs or TPMs. In the present work we used TPMs; *and* (iii) in linear (*i.e.*, not log) scale..

3.3 Running Compass on bulk (i.e., non-single-cell) inputs

The current manuscript presents the algorithm in the context of single cells, where Compass leverages the statistical power afforded by the large number of observations (cells). Nevertheless, there is no inherent limitation preventing one from applying Compass to study bulk (i.e., non-single-cell) transcriptomic data. In this case, we recommend disabling the information-sharing feature by setting $lambda = 0$ in Algorithm 2. There is also no limitation preventing one from applying Compass to non-RNA-Seq transcriptomic data, such as microarrays.

3.4 Scalability

For prohibitively large datasets, the number of cells (observations) can be reduced by partitioning the cells into small clusters and treating the average of each cluster as an observation in downstream analysis. Two implementations of this approach are *micropools* [18], implemented in the `VISION` R package (`https://github.com/YosefLab/VISION`), and *metacells* [19] (`https://tanaylab.github.io/metacell`). No pooling was necessary for the analysis presented in this manuscript (*i.e.*, the results are on a single cell level). If cell clusters are large enough, one may choose to skip the information-sharing procedure, which is equivalent to setting the parameter $\lambda = 0$ in Algorithm 2.

In addition, the number of reactions in the GSMM can be reduced as well by not executing Algorithm 2 on blocked reactions (Section 4.5), non-core reactions (Section 2.1), or reactions outside a predetermined set of metabolic pathways that are of interest. We note that we do not suggest excluding non-blocked reactions from the network altogether (which would result in neglecting their effects on reactions that are of interest), but rather only excluding them from the $R(G)$ matrices in Algorithm 2.

4 The Compass algorithm in detail

4.1 Metabolic network

We used the Recon2 GSMM [20], which we transformed into a unidirectional network by replacing bidirectional reactions with the respective pair of unidirectional reactions. Consequently, flux values are always non-negative.

4.2 Metabolic genes

Throughout this work, metabolic genes are defined as the set of genes annotated in Recon2. Note that Compass uses only the expression of metabolic genes and ignores other transcripts.

4.3 *In silico* growth medium

The results of flux balance analysis significantly depend on the nutrients made available to the GSMM, referred to as the *in silico* growth medium. Since exact medium composition is mostly unknown even for common *in vitro* protocols and *in vivo* models, we chose a rich *in silico* medium where all nutrients for which a transporter exits are made available in an unlimiting quantity.

4.4 Notation

In the following sections we denote:

- n: number of cells (or RNA libraries).
- m: number of metabolic reactions in the GSMM.
- $\mathcal{C}$: the set of cells in the data. ($|\mathcal{C}|= n$).
- $\mathcal{R}$: the set of metabolic reactions in the GSMM. ($|\mathcal{R}|= m$).
- $rev(r)$: the reverse unidirectional reaction of reaction r, which has the same stoichiometry but proceeds in the opposite direction. g: number of genes in a given transcriptome dataset.

- S: he stoichiometric matrix defined in the GSMM, where rows represent metabolites, columns represent reactions, and entries are stoichiometrical coefficients for the reactions comprising the metabolic network. Reactions for uptake and secretion of a metabolite are encoded as having only a coefficient of 1 and -1 in the metabolite's row entry, respectively, and 0 otherwise.

For a matrix $M = (m_{i,j})$ and a function $f\colon \mathbb{R} \to \mathbb{R}$ we use $f(M)$ to denote (where the intention is obvious from the context) the respective point-wise transformation, namely $f(M) \coloneqq (f(m_{i,j}))$.

4.5 Transcriptome-agnostic preparatory step

For a given GSMM (here, Recon2), we run once a preparatory step that does not depend on transcriptome data and cache the results **(Algorithm 1)**.

Algorithm 1: Find maximal reaction fluxes

input : GSMM
output: maximal flux v_r^{opt} that every reaction r can carry

foreach $r \in \mathcal{R}$ **do**

$$
\begin{aligned}
v_r^{opt} \coloneqq \quad & \underset{v \in \mathbb{R}^m}{\textbf{maximize}}\ v_r \\
& \textbf{s.t.} \\
& \textit{(i)} \quad S \cdot v = 0 \\
& \textit{(ii)} \quad \alpha \leq v \leq \beta \\
& \textit{(iii)} \quad v_{rev(r)} = 0
\end{aligned}
$$

end

Constraint *(i)* constrains the system to steady state (Varma and Palsson 1994). Constraint *(ii)* is interpreted as $\forall i\colon \alpha_i \leq v_i \leq \beta_i$ and encodes directionality and capacity limits for reactions, including uptake and secretion limits. Constraint *(iii)* ensures that when evaluating the maximum flux for each reaction, its reverse reaction carries no flux to avoid the creation of a futile cycle. This does not prevent futile cycles longer than 2 edges, which can be avoided only by more time-consuming computations (Schellenberger, Lewis, and Palsson 2011).

Note that the GSMM may contain blocked reactions ($v_r^{opt} = 0$) that can be excluded from the next steps to speed the computation.

4.6 From gene expression to reaction expression

By reaction expression, we denote a matrix $\{R(G)\}_{m \times n}$ that is conceptually similar to the gene expression matrix $\{G\}_{g \times n}$. The columns are the same RNA

libraries (*e.g.*, cells) as in G, but rows represent single metabolic reactions rather than transcripts. An entry $R_{r,j}$ in the matrix $R(G)$ is a quantitative proxy for the activity of reaction r in cell j. We omit the dependence on gene expression matrix and denote simply R when G is obvious from the context.

The reaction expression matrix is created by using the boolean gene-to-reaction mapping included in the GSMM, similar to the approach taken by [21], [22]. Let $G = \{x_{i,j}\}$ and consider a particular reaction r in a particular cell j. If a single gene with linear-scale expression x is associated with r, then the reaction's expression will be $R_{r,j} = \log_2(x+1)$. If no genes are associated with r then $R_{r,j} = 0$.

If the reaction is associated with more than one gene, then this association is expressed as a boolean relationship. For example, two genes which encode different subunits of a reaction's enzyme are associated using an AND relationship as both are required to be expressed for the reaction to be catalyzed. Alternately, if multiple enzymes can catalyze a reaction, the genes involved in each will be associated via an OR relationship. For reactions associated with multiple genes in this manner, the boolean expression is evaluated by taking the sum or the mean of linear-scale expression values x when genes are associated via an OR or AND relationship, respectively. This way, the full gene(s)-to-reaction associations is evaluated to arrive at a single summary expression value for each reaction in the GSMM.

4.7 Information sharing between single cells (smoothing)

To mitigate the sparseness and stochasticity of single-cell measurements, Compass allows for a degree of information-sharing between cells with similar transcriptional profiles. Given a gene expression G, we compute k-nearest neighbors (kNN) graph based Euclidean distances in reduced dimension, obtained by taking the top 20 principal components of G. The PCA is computed over all the genes in G, not only metabolic ones.

Let $R(G) = \{r_{i,j}\}$ and

$$w_{i,j} = \begin{cases} \frac{1}{k}, & \text{if cell } j \text{ is in the } k\text{-nearest-neighborhood of cell } i, \\ 0, & \text{otherwise} \end{cases}$$

then $R^N(G) = \{r^N_{i,j}\}$ where

$$r^N_{i,j} = \sum_{c \in \mathcal{C}} w_{j,c} r_{i,c}$$

4.8 Main algorithm

Compass transforms a gene expression matrix $\{G\}_{g \times n}$, where rows represent genes and columns represent RNA libraries (usually, single cells, although bulk RNA can also be used as discussed below) into a matrix $\{C\}_{m \times n}$ of scores where rows represent metabolic reactions, columns are the same RNA libraries

as in the gene expression, and an entry quantifies a proxy for potential reaction activity. More precisely, the entry quantifies the propensity of the cell to use that reaction.

The algorithm is summarized in **(Algorithm 2)**. First, we convert the gene expression matrix $G_{g\times n}$ into a *reaction expression* matrix $R_{m\times n}$ which is parallel to the gene expression matrix, but with rows representing single metabolic reactions rather than transcripts. We convert R into a *penalty matrix* $P_{m\times n}$ by point-wise inversion. Whereas R represents gene expression support that a reaction is functional in the cell, P represents the lack thereof (which will be used in a linear program below). The computation of R and P occurs also for the neighborhood of each cell for to smooth results and mitigate single-cell technical noise. Then, we solve a linear program for every reaction r in every cell i to find the minimal resistance of cell i to carry maximal flux through r. Last, we scale the scores, which also entails negating them such that that larger scores will represent larger potential activities (instead of larger penalties, hence smaller potential activity). The final scores indicative of a cell's propensity to use a certain reaction. We interpret it as a proxy for the potential activity of the reaction in that cell.

In step 10 of Algorithm 2, a high penalty $y_{r,c}$ indicates that cell c is unlikely, judged by transcriptomic evidence, to use reaction r. Cells whose transcriptome are overall more aligned with an ability to carry flux through a reaction will be assigned a lower penalty $y_{r,c}$. With regards to the correctness of the step, recall that that the GSMM is unidirectional and therefore $\forall i.\ v_i > 0$.

4.9 Meta-reactions

Rows in the C^{raw} matrix that correspond to reactions that are topologically close in the metabolic network can be highly correlated. We therefore hierarchically cluster C^{raw} rows by Spearman distance. We call the resulting clusters meta-reactions and each represents a set of closely correlated metabolic reactions. Note that the division into meta-reactions is data-driven and does not rely on canonical metabolic pathway definitions. Therefore, the division is dataset-dependent - for example, two reactions might be closely correlated and clustered in the same meta-reaction in one cell type, but not in another.

After computing the hierarchical clusters over rows of C^{raw}, we merged leaves in which Spearman similarity (namely $1-\rho$, with ρ being the Spearman correlation) by averaging the respective rows. In the present work, we used $\rho = 0.98$. We denote the row-merged matrix $\{C^{meta-raw}\}_{m'\times n}$

4.10 Scaling raw compass scores

By definition, all entries in $C^{meta-raw}$ are non-negative. We scale it in Algorithm 3 (the min in the second step denotes matrix-wide minimal entry)

Algorithm 2: Compass

input : Gene expression matrix $\{G\}_{g\times n}$
GSMM
Pre-computed maximal fluxes $\{v_r^{opt} : r \in \mathcal{R}\}$

output : Compass scores matrix $\{C\}_{m'\times n}$ $(m' \leq m)$

parameters: Smoothing parameter $\lambda \in [0,1]$ (here, $\lambda = 0.25$)
Nearest neighbor parameter k (here, $k = 10$)
Optimality slack parameter $\omega \in [0,1]$ (here, $\omega = 0.95$)
Penalty function $p(x)$ (here, $p(x) = 1/(1+x)$)
Meta-reaction merging threshold $\rho \in [0,1]$ (see section 4.9)

Compute with the procedures described in sections 4.6, 4.7:
  a reaction expression matrix $\{R(G)\}_{m\times n}$
  a neighborhood reaction expression matrix $\{R^N(G)\}_{m\times n}$
Transform reaction expressions to penalties:

$$P := p(R(G))$$
$$P^N := p(R^N(G))$$

$\hat{P} := (1-\lambda)P + \lambda P^N$
foreach $r \in \mathcal{R}, c \in \mathcal{C}$ **do**
  Let $\hat{P}^{(c)} = \left(\hat{P}_{1,c}, \ldots, \hat{P}_{m,c}\right)$

$$\begin{aligned} y_{r,c} := \quad & \underset{v\in\mathbb{R}^m}{\textbf{minimize}}\ \hat{P}^{(c)} \cdot v \\ & \textbf{s.t.} \\ & \textit{(i)} \quad S \cdot v = 0 \\ & \textit{(ii)} \quad \alpha \leq v \leq \beta \\ & \textit{(iii)} \quad v_{rev(r)} = 0 \\ & \textit{(iv)} \quad v_r \geq \omega \cdot v_r^{opt} \end{aligned}$$

end
Let $C^{raw} = \{y_{r,c}\}_{m\times n}$
Compute meta-reaction scores $C^{meta-raw} = \{y'_{r,c}\}_{m''\times n}$ $(m'' \leq m)$ as described in section 4.9
Use Algorithm 3 to scale $C^{meta-raw}$ and obtain C
Remove constant rows from C, defined as rows in which the difference between largest and smallest score is less than $\varepsilon = 10^{-3}$
return C

Algorithm 3: Scale raw Compass scores

input : $C^{meta-raw}$
output : C

$C^{meta-raw} := -\log(1 + C^{meta-raw})$
$C := C^{meta-raw} - \min(C^{meta-raw})$
return C

5 Algorithm generalization

One of the intuitions behind Compass is that the statistical power afforded by the number of observations (cells) in single-cell RNA-Seq allows *increasing* dimensionality by computing a new feature set based on the gene expression data and the GSMM. Here, we used an intuitive set of objective functions — for each reaction in the network, we defined one objective function which is to maximize the flux it carries (recall that the network is unidirectional and therefore all reactions carry non-negative fluxes). This allows intuitive interpretation of the Compass scores as quantitative proxies to reaction activities. However, the algorithm can be generalized by using an arbitrary set of linear objective functions that pertain to cellular metabolism.

References

[1] J. T. Gaublomme, N. Yosef, Y. Lee, R. S. Gertner, L. V. Yang, C. Wu, P. P. Pandolfi, T. Mak, R. Satija, A. K. Shalek, V. K. Kuchroo, H. Park, A. Regev, H. Park, C. Wu, P. P. Pandolfi, T. Mak, R. Satija, A. K. Shalek, V. K. Kuchroo, H. Park, A. Regev, and H. Park, "Single-Cell genomics unveils critical regulators of th17 cell pathogenicity," *Cell*, vol. 163, no. 6, pp. 1400–1412, Dec. 2015.

[2] R. B. Fletcher, D. Das, L. Gadye, K. N. Street, A. Baudhuin, A. Wagner, M. B. Cole, Q. Flores, Y. G. Choi, N. Yosef, E. Purdom, S. Dudoit, D. Risso, and J. Ngai, "Deconstructing olfactory stem cell trajectories at Single-Cell resolution," *Cell Stem Cell*, vol. 20, no. 6, 817–830.e8, Jun. 2017.

[3] M. B. Cole, D. Risso, A. Wagner, D. DeTomaso, J. Ngai, E. Purdom, S. Dudoit, and N. Yosef, "Performance assessment and selection of normalization procedures for Single-Cell RNA-Seq," en, *Cell Syst*, vol. 8, no. 4, 315–328.e8, Apr. 2019.

[4] M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth, "Limma powers differential expression analyses for RNA-sequencing and microarray studies," en, *Nucleic Acids Res.*, vol. 43, no. 7, e47, Apr. 2015.

[5] C. W. Law, Y. Chen, W. Shi, and G. K. Smyth, "Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts," en, *Genome Biol.*, vol. 15, no. 2, R29, Feb. 2014.

[6] G. Smyth, *Limma Bioconductor Package*, `https://bioconductor.org/packages/release/bioc/html/limma.html`, [Bioconductor release 3.10; Last accessed Dec. 2019], 2019.

[7] X. Li, S. Tang, Q.-Q. Wang, E. L.-H. Leung, H. Jin, Y. Huang, J. Liu, M. Geng, M. Huang, S. Yuan, X.-J. Yao, and J. Ding, "Identification of epigallocatechin-3- gallate as an inhibitor of phosphoglycerate mutase 1," en, *Front. Pharmacol.*, vol. 8, p. 325, May 2017.

[8] A. G. Schwartz and L. L. Pashko, "Dehydroepiandrosterone, glucose-6-phosphate dehydrogenase, and longevity," en, *Ageing Res. Rev.*, vol. 3, no. 2, pp. 171–187, Apr. 2004.

[9] N. Westergaard, P. Madsen, and K. Lundgren, "Characterization of glycerol uptake and glycerol kinase activity in rat hepatocytes cultured under different hormonal conditions," en, *Biochim. Biophys. Acta*, vol. 1402, no. 3, pp. 261–268, Apr. 1998.

[10] M. J. Tisdale and M. D. Threadgill, "(+/-)2,3-dihydroxypropyl dichloroacetate, an inhibitor of glycerol kinase," en, *Cancer Biochem. Biophys.*, vol. 7, no. 3, pp. 253–259, Sep. 1984.

[11] X. Zhao, Y. Zhu, J. Hu, L. Jiang, L. Li, S. Jia, and K. Zen, "Shikonin inhibits tumor growth in mice by suppressing pyruvate kinase m2-mediated aerobic glycolysis," en, *Sci. Rep.*, vol. 8, no. 1, p. 14 517, Sep. 2018.

[12] J. Chen, J. Xie, Z. Jiang, B. Wang, Y. Wang, and X. Hu, "Shikonin and its analogs inhibit cancer cell glycolysis by targeting tumor pyruvate kinase-m2," en, *Oncogene*, vol. 30, no. 42, pp. 4297–4306, Oct. 2011.

[13] Q. Wang, M. V. Liberti, P. Liu, X. Deng, Y. Liu, J. W. Locasale, and L. Lai, "Rational design of selective allosteric inhibitors of PHGDH and serine synthesis with anti-tumor activity," en, *Cell Chem Biol*, vol. 24, no. 1, pp. 55–65, Jan. 2017.

[14] A. Jäger, V. Dardalhon, R. A. Sobel, E. Bettelli, and V. K. Kuchroo, "Th1, th17, and th9 effector cells induce experimental autoimmune encephalomyelitis with different pathological phenotypes," en, *J. Immunol.*, vol. 183, no. 11, pp. 7169–7177, Dec. 2009.

[15] C. Wang, N. Yosef, J. Gaublomme, C. Wu, Y. Lee, C. B. Clish, J. Kaminski, S. Xiao, G. M. Z. Horste, M. Pawlak, Y. Kishi, N. Joller, K. Karwacz, C. Zhu, M. Ordovas-Montanes, A. Madi, I. Wortman, T. Miyazaki, R. A. Sobel, H. Park, A. Regev, V. K. Kuchroo, G. M. Zu Horste, M. Pawlak, Y. Kishi, N. Joller, K. Karwacz, C. Zhu, M. Ordovas-Montanes, A. Madi, I. Wortman, T. Miyazaki, R. A. Sobel, H. Park, A. Regev, and V. K. Kuchroo, "CD5L/AIM regulates lipid biosynthesis and restrains th17 cell pathogenicity," *Cell*, vol. 163, no. 6, pp. 1413–1427, Dec. 2015.

[16] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: New perspectives on genomes, pathways, diseases and drugs," en, *Nucleic Acids Res.*, vol. 45, no. D1, pp. D353–D361, Jan. 2017.

[17] N. Yosef, A. K. Shalek, J. T. Gaublomme, H. Jin, Y. Lee, A. Awasthi, C. Wu, K. Karwacz, S. Xiao, M. Jorgolli, D. Gennert, R. Satija, A. Shakya, D. Y. Lu, J. J. Trombetta, M. R. Pillai, P. J. Ratcliffe, M. L. Coleman, M. Bix, D. Tantin, H. Park, V. K. Kuchroo, and A. Regev, "Dynamic regulatory network controlling TH17 cell differentiation," en, *Nature*, vol. 496, no. 7446, pp. 461–468, Apr. 2013.

[18] D. DeTomaso, M. G. Jones, M. Subramaniam, T. Ashuach, C. J. Ye, and N. Yosef, "Functional interpretation of single cell similarity maps," en, *Nat. Commun.*, vol. 10, no. 1, p. 4376, Sep. 2019.

[19] Y. Baran, A. Bercovich, A. Sebe-Pedros, Y. Lubling, A. Giladi, E. Chomsky, Z. Meir, M. Hoichman, A. Lifshitz, and A. Tanay, "MetaCell: Analysis of single-cell RNA-seq data using k-nn graph partitions," en, *Genome Biol.*, vol. 20, no. 1, p. 206, Oct. 2019.

[20] I. Thiele, N. Swainston, R. M. T. Fleming, A. Hoppe, S. Sahoo, M. K. Aurich, H. Haraldsdottir, M. L. Mo, O. Rolfsson, M. D. Stobbe, S. G. Thorleifsson, R. Agren, C. Bölling, S. Bordel, A. K. Chavali, P. Dobson, W. B. Dunn, L. Endler, D. Hala, M. Hucka, D. Hull, D. Jameson, N. Jamshidi, J. J. Jonsson, N. Juty, S. Keating, I. Nookaew, N. Le Novère, N. Malys, A. Mazein, J. A. Papin, N. D. Price, E. Selkov Sr, M. I. Sigurdsson, E. Simeonidis, N. Sonnenschein, K. Smallbone, A. Sorokin, J. H. G. M. van Beek, D. Weichart, I. Goryanin, J. Nielsen, H. V. Westerhoff, D. B. Kell, P. Mendes, and B. Ø. Palsson, "A community-driven global reconstruction of human metabolism," en, *Nat. Biotechnol.*, vol. 31, no. 5, pp. 419–425, May 2013.

[21] S. A. Becker and B. O. Palsson, "Context-Specific metabolic networks are consistent with experiments," *PLoS Comput. Biol.*, vol. 4, no. 5, e1000082, May 2008.

[22] T. Shlomi, M. N. Cabili, M. J. Herrgard, B. O. Palsson, and E. Ruppin, "Network-based prediction of human tissue-specific metabolism," *Nat. Biotechnol.*, vol. 26, no. 9, pp. 1003–1010, Sep. 2008.

Figure 1

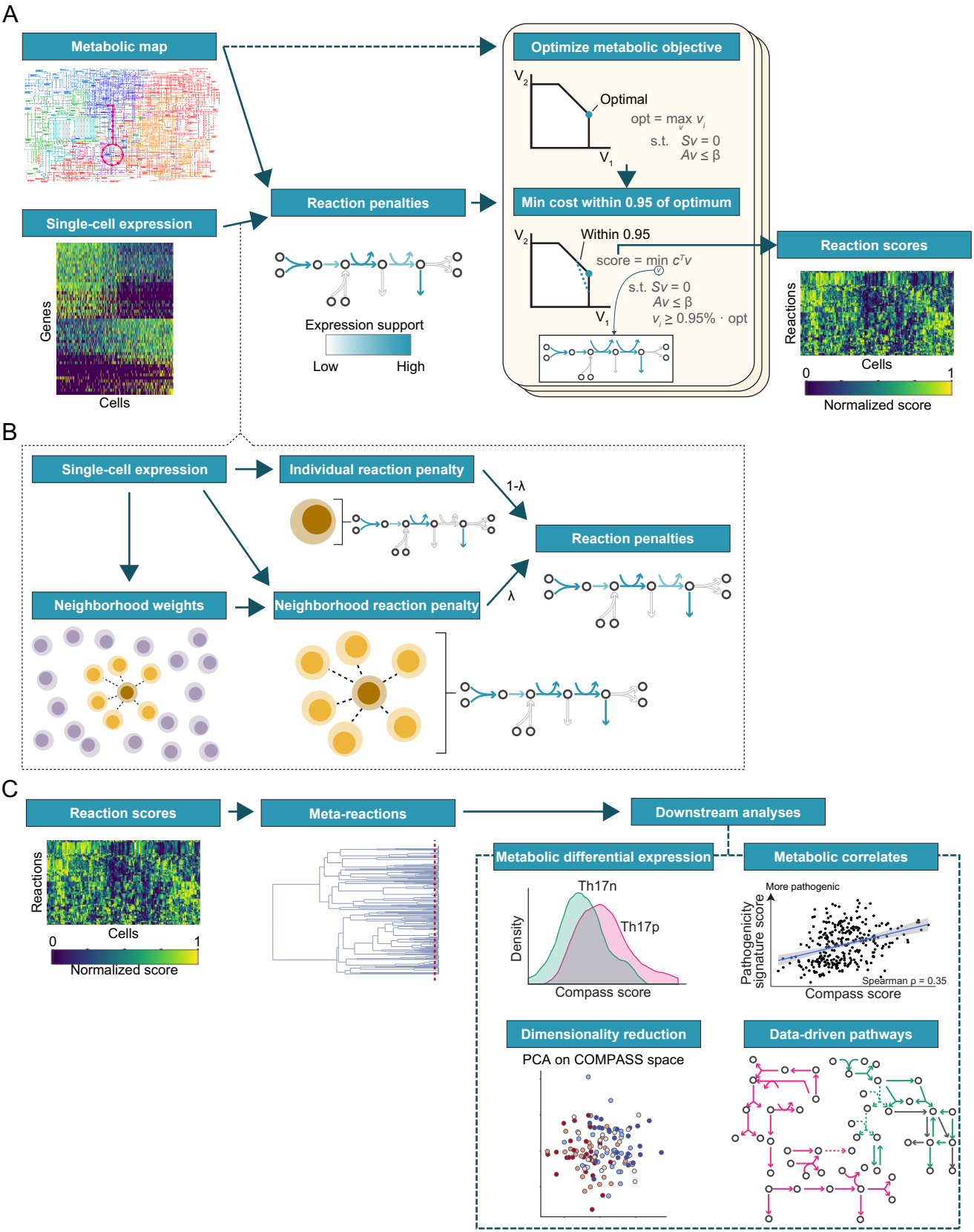

A
Metabolic map
Single-cell expression
Genes
Cells
Reaction penalties
Expression support
Low
High
Optimize metabolic objective
V2
V1
Optimal
opt = max vi
s.t. Sv = 0
Av ≤ β
Min cost within 0.95 of optimum
Within 0.95
score = min cTv
s.t. Sv = 0
Av ≤ β
vi ≥ 0.95% · opt
Reaction scores
Reactions
Cells
0
1
Normalized score
B
Single-cell expression
Individual reaction penalty
1-λ
Reaction penalties
Neighborhood weights
Neighborhood reaction penalty
λ
C
Reaction scores
Reactions
Cells
0
1
Normalized score
Meta-reactions
Downstream analyses
Metabolic differential expression
Th17n
Th17p
Density
Compass score
Metabolic correlates
More pathogenic
Pathogenicity signature score
Spearman ρ = 0.35
Compass score
Dimensionality reduction
PCA on COMPASS space
Data-driven pathways

Figure 2

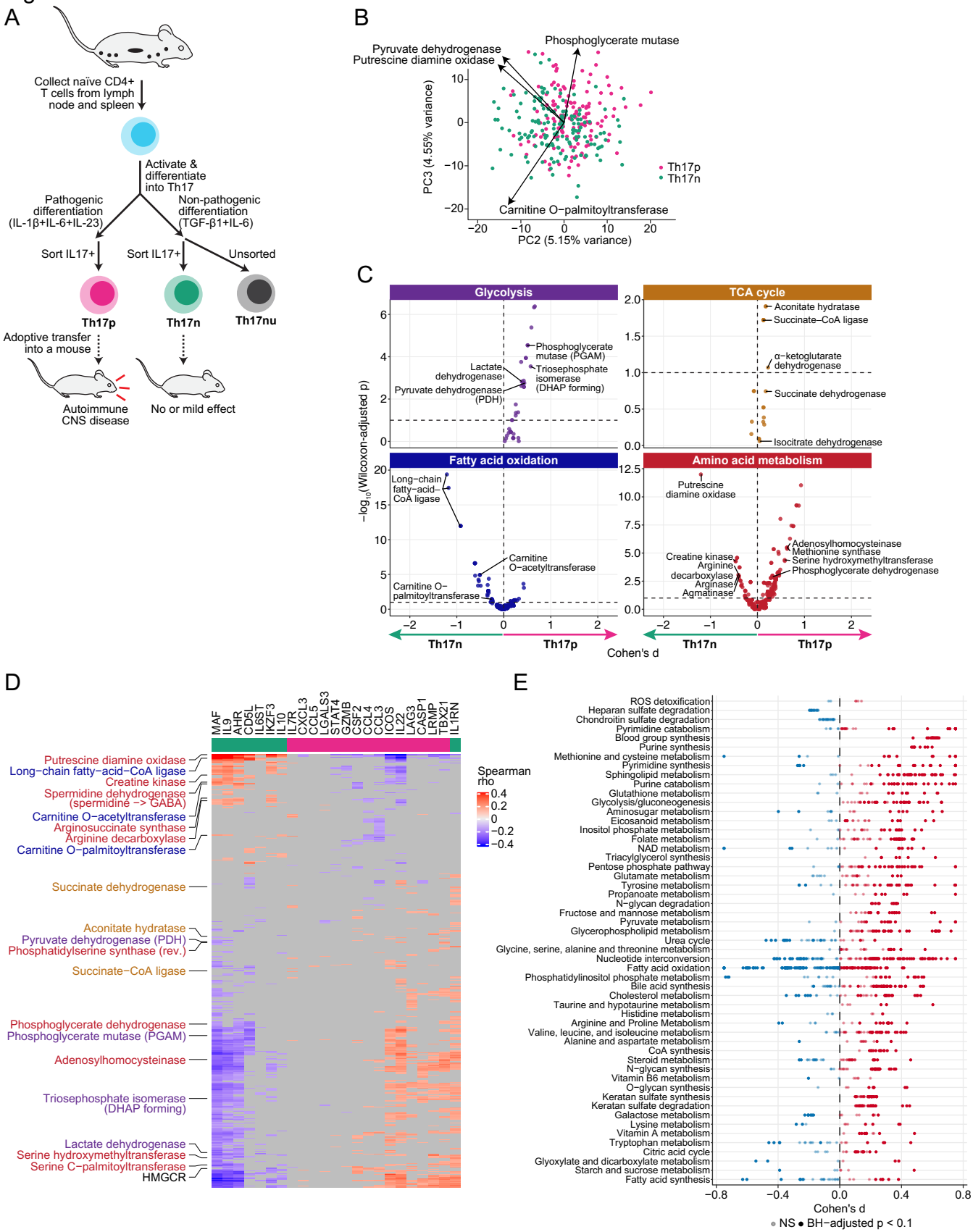

A
Collect naïve CD4+ T cells from lymph node and spleen
Activate & differentiate into Th17
Pathogenic differentiation (IL-1β+IL-6+IL-23)
Non-pathogenic differentiation (TGF-β1+IL-6)
Sort IL17+
Sort IL17+
Unsorted
Th17p
Th17n
Th17nu
Adoptive transfer into a mouse
Autoimmune CNS disease
No or mild effect
B
Pyruvate dehydrogenase
Putrescine diamine oxidase
Phosphoglycerate mutase
Carnitine O−palmitoyltransferase
PC3 (4.55% variance)
PC2 (5.15% variance)
Th17p
Th17n
C
Glycolysis
Phosphoglycerate mutase (PGAM)
Triosephosphate isomerase (DHAP forming)
Lactate dehydrogenase
Pyruvate dehydrogenase (PDH)
TCA cycle
Aconitate hydratase
Succinate−CoA ligase
α−ketoglutarate dehydrogenase
Succinate dehydrogenase
Isocitrate dehydrogenase
Fatty acid oxidation
Long−chain fatty−acid−CoA ligase
Carnitine O−acetyltransferase
Carnitine O−palmitoyltransferase
Amino acid metabolism
Putrescine diamine oxidase
Adenosylhomocysteinase
Methionine synthase
Serine hydroxymethyltransferase
Phosphoglycerate dehydrogenase
Creatine kinase
Arginine decarboxylase
Arginase
Agmatinase
−log10(Wilcoxon-adjusted p)
Th17n
Th17p
Cohen's d
D
MAF
IL9
AHR
CD5L
IL6ST
IKZF3
IL10
IL7R
CXCL3
CCL5
LGALS3
STAT4
GZMB
CSF2
CCL4
CCL3
ICOS
IL22
LAG3
CASP1
LRMP
TBX21
IL1RN
Putrescine diamine oxidase
Long−chain fatty−acid−CoA ligase
Creatine kinase
Spermidine dehydrogenase (spermidine −> GABA)
Carnitine O−acetyltransferase
Arginosuccinate synthase
Arginine decarboxylase
Carnitine O−palmitoyltransferase
Succinate dehydrogenase
Aconitate hydratase
Pyruvate dehydrogenase (PDH)
Phosphatidylserine synthase (rev.)
Succinate−CoA ligase
Phosphoglycerate dehydrogenase
Phosphoglycerate mutase (PGAM)
Adenosylhomocysteinase
Triosephosphate isomerase (DHAP forming)
Lactate dehydrogenase
Serine hydroxymethyltransferase
Serine C−palmitoyltransferase
HMGCR
Spearman rho
E
ROS detoxification
Heparan sulfate degradation
Chondroitin sulfate degradation
Pyrimidine catabolism
Blood group synthesis
Purine synthesis
Methionine and cysteine metabolism
Pyrimidine synthesis
Sphingolipid metabolism
Purine catabolism
Glutathione metabolism
Glycolysis/gluconeogenesis
Aminosugar metabolism
Eicosanoid metabolism
Inositol phosphate metabolism
Folate metabolism
NAD metabolism
Triacylglycerol synthesis
Pentose phosphate pathway
Glutamate metabolism
Tyrosine metabolism
Propanoate metabolism
N−glycan degradation
Fructose and mannose metabolism
Pyruvate metabolism
Glycerophospholipid metabolism
Urea cycle
Glycine, serine, alanine and threonine metabolism
Nucleotide interconversion
Fatty acid oxidation
Phosphatidylinositol phosphate metabolism
Bile acid synthesis
Cholesterol metabolism
Taurine and hypotaurine metabolism
Histidine metabolism
Arginine and Proline Metabolism
Valine, leucine, and isoleucine metabolism
Alanine and aspartate metabolism
CoA synthesis
Steroid metabolism
N−glycan synthesis
Vitamin B6 metabolism
O−glycan synthesis
Keratan sulfate synthesis
Keratan sulfate degradation
Galactose metabolism
Lysine metabolism
Vitamin A metabolism
Tryptophan metabolism
Citric acid cycle
Glyoxylate and dicarboxylate metabolism
Starch and sucrose metabolism
Fatty acid synthesis
Cohen's d
NS
BH−adjusted p < 0.1

Figure 3

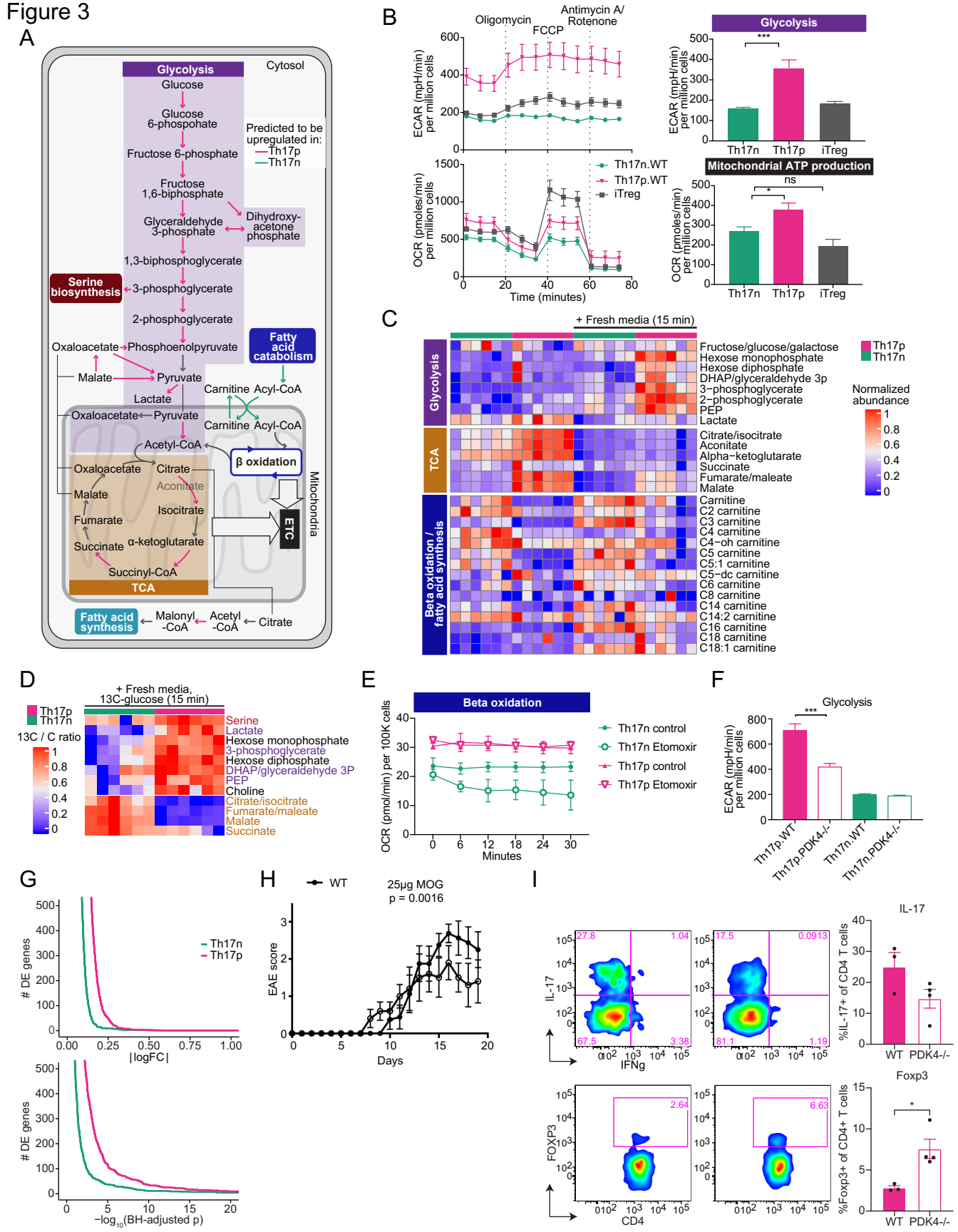

A
Glycolysis
Cytosol
Glucose
Glucose 6-phospohate
Fructose 6-phosphate
Fructose 1,6-biphosphate
Glyceraldehyde 3-phosphate
Dihydroxy-acetone phosphate
1,3-biphosphoglycerate
3-phosphoglycerate
2-phosphoglycerate
Phosphoenolpyruvate
Pyruvate
Lactate
Predicted to be upregulated in:
Th17p
Th17n
Serine biosynthesis
Fatty acid catabolism
Oxaloacetate
Malate
Carnitine
Acyl-CoA
Acetyl-CoA
β oxidation
Citrate
Aconitate
Isocitrate
α-ketoglutarate
Succinyl-CoA
Succinate
Fumarate
TCA
ETC
Mitochondria
Fatty acid synthesis
Malonyl -CoA
Acetyl -CoA
B
Oligomycin
FCCP
Antimycin A/ Rotenone
ECAR (mpH/min) per million cells
OCR (pmoles/min) per million cells
Time (minutes)
Th17n.WT
Th17p.WT
iTreg
Glycolysis
Mitochondrial ATP production
Th17n
Th17p
iTreg
ns
C
+ Fresh media (15 min)
Glycolysis
TCA
Beta oxidation / fatty acid synthesis
Fructose/glucose/galactose
Hexose monophosphate
Hexose diphosphate
DHAP/glyceraldehyde 3p
3−phosphoglycerate
2−phosphoglycerate
PEP
Lactate
Citrate/isocitrate
Aconitate
Alpha−ketoglutarate
Succinate
Fumarate/maleate
Malate
Carnitine
C2 carnitine
C3 carnitine
C4 carnitine
C4−oh carnitine
C5 carnitine
C5:1 carnitine
C5−dc carnitine
C6 carnitine
C8 carnitine
C14 carnitine
C14:2 carnitine
C16 carnitine
C18 carnitine
C18:1 carnitine
Th17p
Th17n
Normalized abundance
D
+ Fresh media, 13C-glucose (15 min)
Th17p
Th17n
13C / C ratio
Serine
Lactate
Hexose monophosphate
3-phosphoglycerate
Hexose diphosphate
DHAP/glyceraldehyde 3P
PEP
Choline
Citrate/isocitrate
Fumarate/maleate
Malate
Succinate
E
Beta oxidation
OCR (pmol/min) per 100K cells
Minutes
Th17n control
Th17n Etomoxir
Th17p control
Th17p Etomoxir
F
Glycolysis
ECAR (mpH/min) per million cells
Th17p.WT
Th17p.PDK4-/-
Th17n.WT
Th17n.PDK4-/-
G
DE genes
|logFC|
Th17n
Th17p
−log10(BH-adjusted p)
H
WT
25μg MOG
p = 0.0016
EAE score
Days
I
IL-17
IFNg
FOXP3
CD4
%IL-17+ of CD4 T cells
%Foxp3+ of CD4+ T cells
WT
PDK4-/-

Figure 4

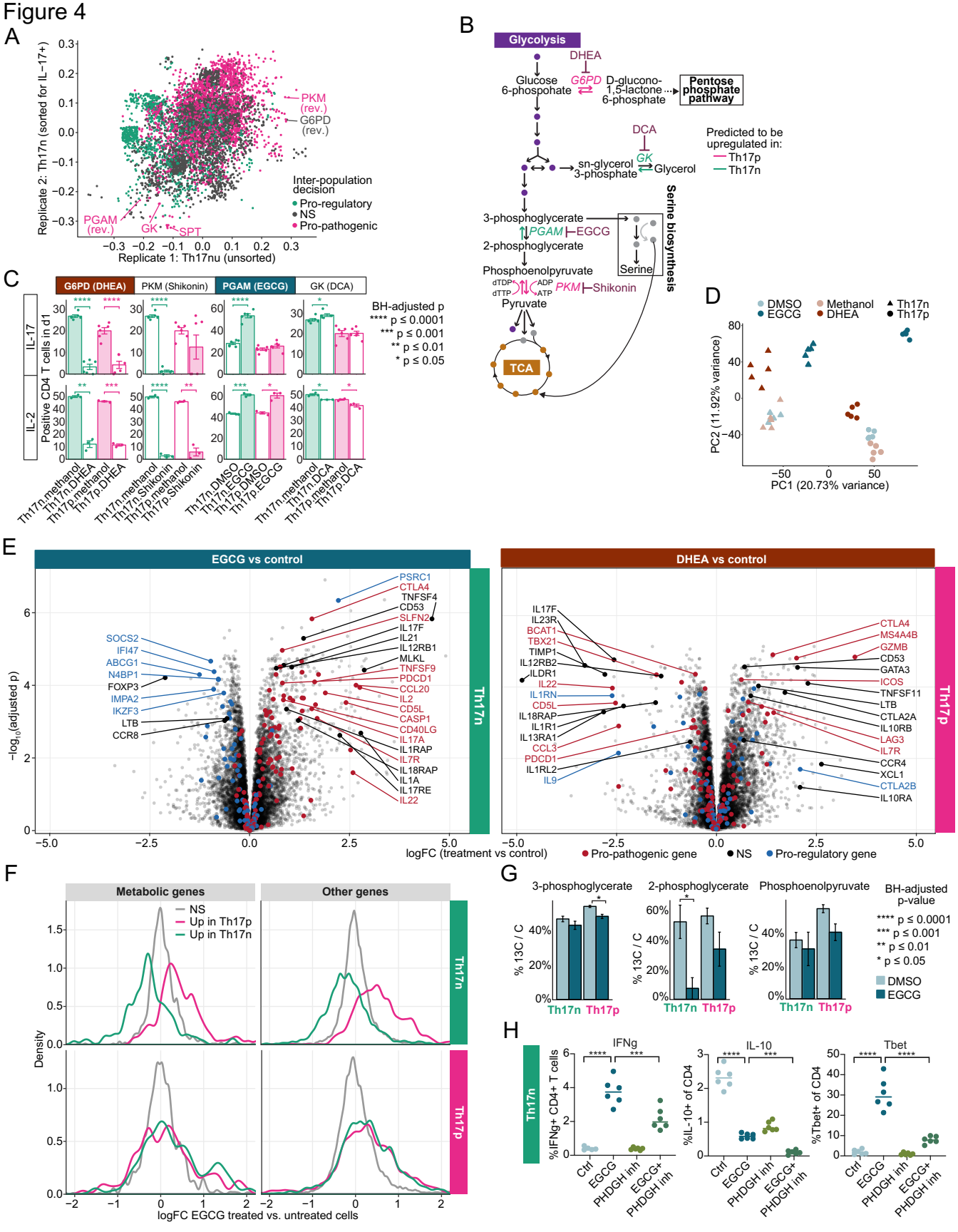
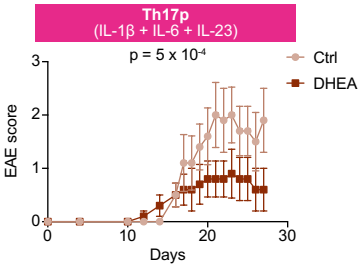

A
Replicate 2: Th17n (sorted for IL−17+)
Replicate 1: Th17nu (unsorted)
PKM (rev.)
G6PD (rev.)
PGAM (rev.)
GK
SPT
Inter-population decision
Pro-regulatory
NS
Pro-pathogenic
B
Glycolysis
DHEA
Glucose 6-phospohate
G6PD
D-glucono-1,5-lactone 6-phosphate
Pentose phosphate pathway
DCA
GK
sn-glycerol 3-phosphate
Glycerol
Predicted to be upregulated in:
Th17p
Th17n
3-phosphoglycerate
PGAM
EGCG
2-phosphoglycerate
Phosphoenolpyruvate
PKM
Shikonin
Pyruvate
TCA
Serine
Serine biosynthesis
C
G6PD (DHEA)
PKM (Shikonin)
PGAM (EGCG)
GK (DCA)
IL-17
IL-2
Positive CD4 T cells in d1
BH-adjusted p
**** p ≤ 0.0001
*** p ≤ 0.001
** p ≤ 0.01
* p ≤ 0.05
Th17n.methanol
Th17n.DHEA
Th17p.methanol
Th17p.DHEA
Th17n.methanol
Th17n.Shikonin
Th17p.methanol
Th17p.Shikonin
Th17n.DMSO
Th17n.EGCG
Th17p.DMSO
Th17p.EGCG
Th17n.methanol
Th17n.DCA
Th17p.methanol
Th17p.DCA
D
DMSO
EGCG
Methanol
DHEA
Th17n
Th17p
PC2 (11.92% variance)
PC1 (20.73% variance)
E
EGCG vs control
DHEA vs control
−log10(adjusted p)
logFC (treatment vs control)
Th17n
Th17p
PSRC1
CTLA4
TNFSF4
CD53
SLFN2
IL17F
IL21
IL12RB1
MLKL
TNFSF9
PDCD1
CCL20
IL2
CD5L
CASP1
CD40LG
IL17A
IL1RAP
IL7R
IL18RAP
IL1A
IL17RE
IL22
SOCS2
IFI47
ABCG1
N4BP1
FOXP3
IMPA2
IKZF3
LTB
CCR8
IL17F
IL23R
BCAT1
TBX21
TIMP1
IL12RB2
ILDR1
IL22
IL1RN
CD5L
IL18RAP
IL1R1
IL13RA1
CCL3
PDCD1
IL1RL2
IL9
CTLA4
MS4A4B
GZMB
CD53
GATA3
ICOS
TNFSF11
LTB
CTLA2A
IL10RB
LAG3
IL7R
CCR4
XCL1
CTLA2B
IL10RA
Pro-pathogenic gene
NS
Pro-regulatory gene
F
Metabolic genes
Other genes
NS
Up in Th17p
Up in Th17n
Density
logFC EGCG treated vs. untreated cells
G
3-phosphoglycerate
2-phosphoglycerate
Phosphoenolpyruvate
% 13C / C
BH-adjusted p-value
**** p ≤ 0.0001
*** p ≤ 0.001
** p ≤ 0.01
* p ≤ 0.05
DMSO
EGCG
H
IFNg
IL-10
Tbet
%IFNg+ CD4+ T cells
%IL-10+ of CD4
%Tbet+ of CD4
Th17n
Ctrl
EGCG
PHDGH inh
EGCG+ PHDGH inh

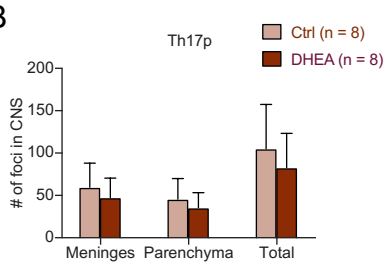
Figure 5

A

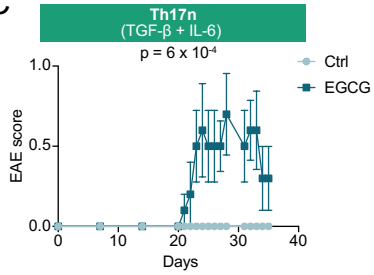


	Incidence	Peak disease	Onset (days)
Methanol	10 (of 10)	2.5+/-1.5	17+/-2
DHEA	7 (of 8)	1.3+/-0.8	17+/-4

B



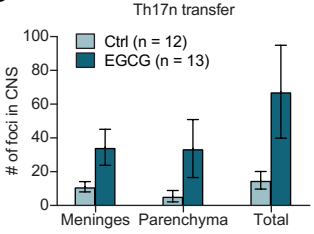
C



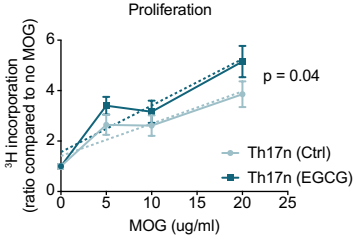
	Incidence	Peak disease	Onset (days)
DMSO	0 (of 12)	0	n/a
EGCG	10 (of 13)	1+/-0.6	25.1+/-4

Fisher's exact p = 1.1 x 10-4

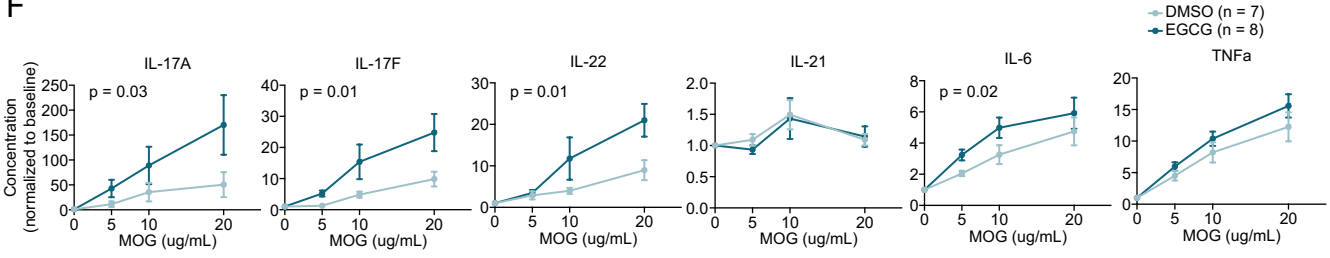
D



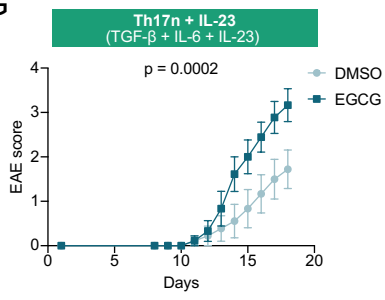
E



F



G



	Incidence	Peak disease	Onset (days)
DMSO	8 (of 9)	1.7+/-1.3	15+/-2
EGCG	9 (of 9)	3.3+/-1.1	13.4+/-1

H

Optic neuritis / perineuritis

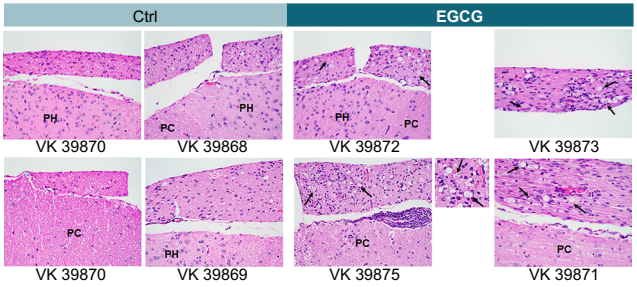
	Treatment	Incidence	Fisher's exact p
Th17n	DMSO	3/12	0.23
	EGCG	7/13	
Th17n + IL-23	DMSO	9/9	1
	EGCG	8/8	
Th17p	Methanol	9/9	0.03
	DHEA	4/8	

I

Wallerian degeneration

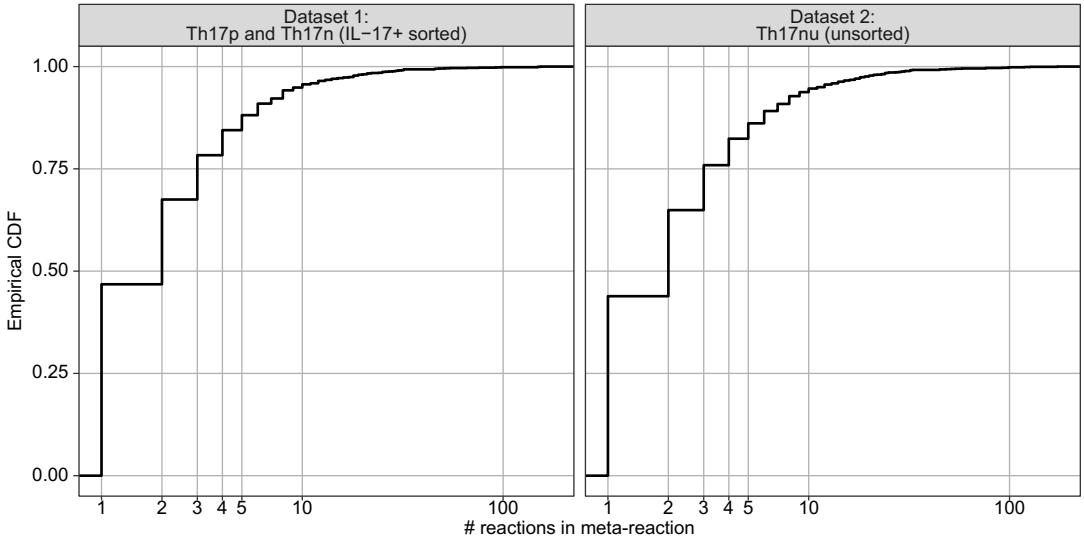
	Treatment	Incidence	Fisher's exact p
Th17n	DMSO	0/12	0.005
	EGCG	7/13	
Th17n + IL-23	DMSO	0/9	0.08
	EGCG	4/9	
Th17p	Methanol	0/9	0.47
	DHEA	1/8	

J

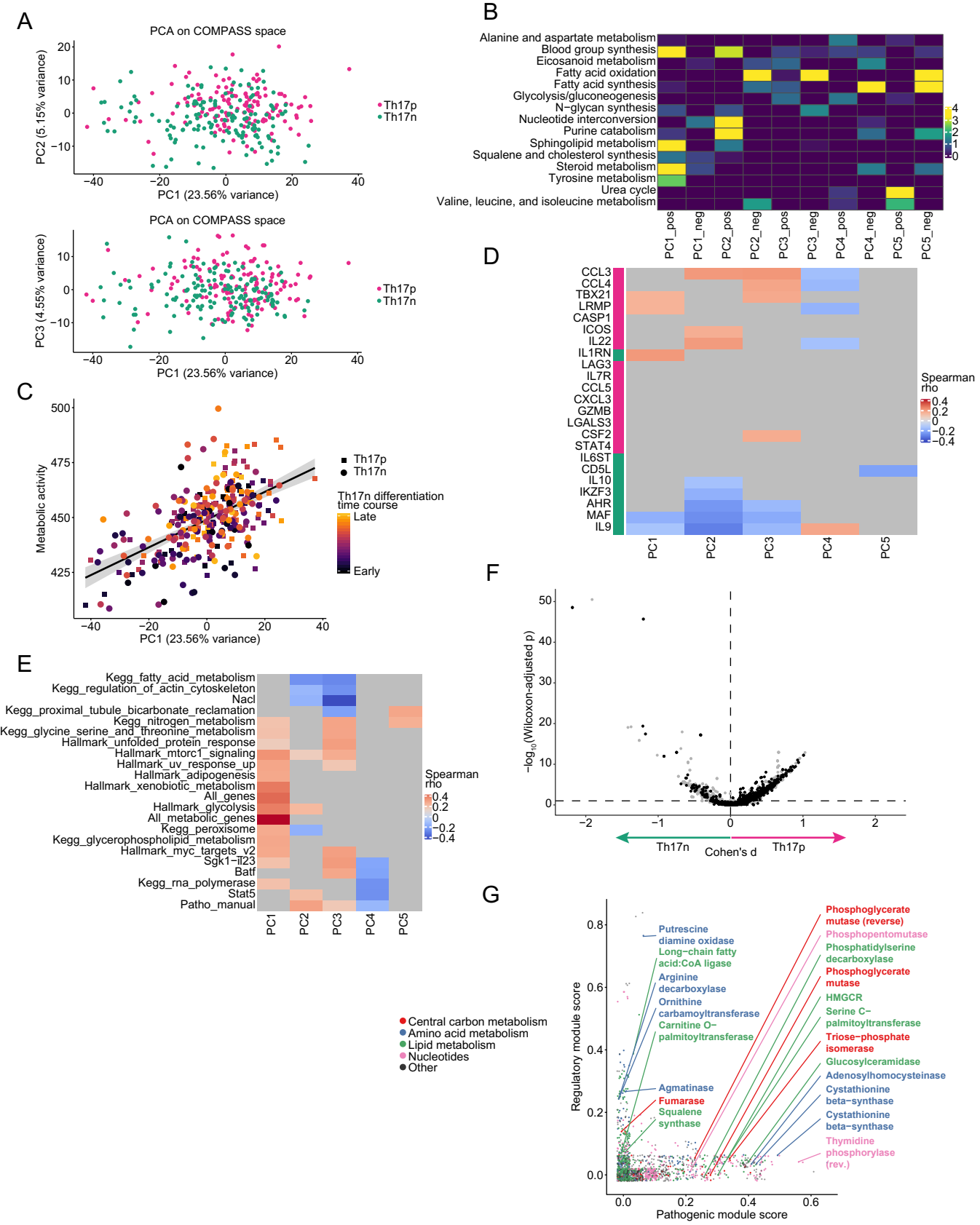


S1

A

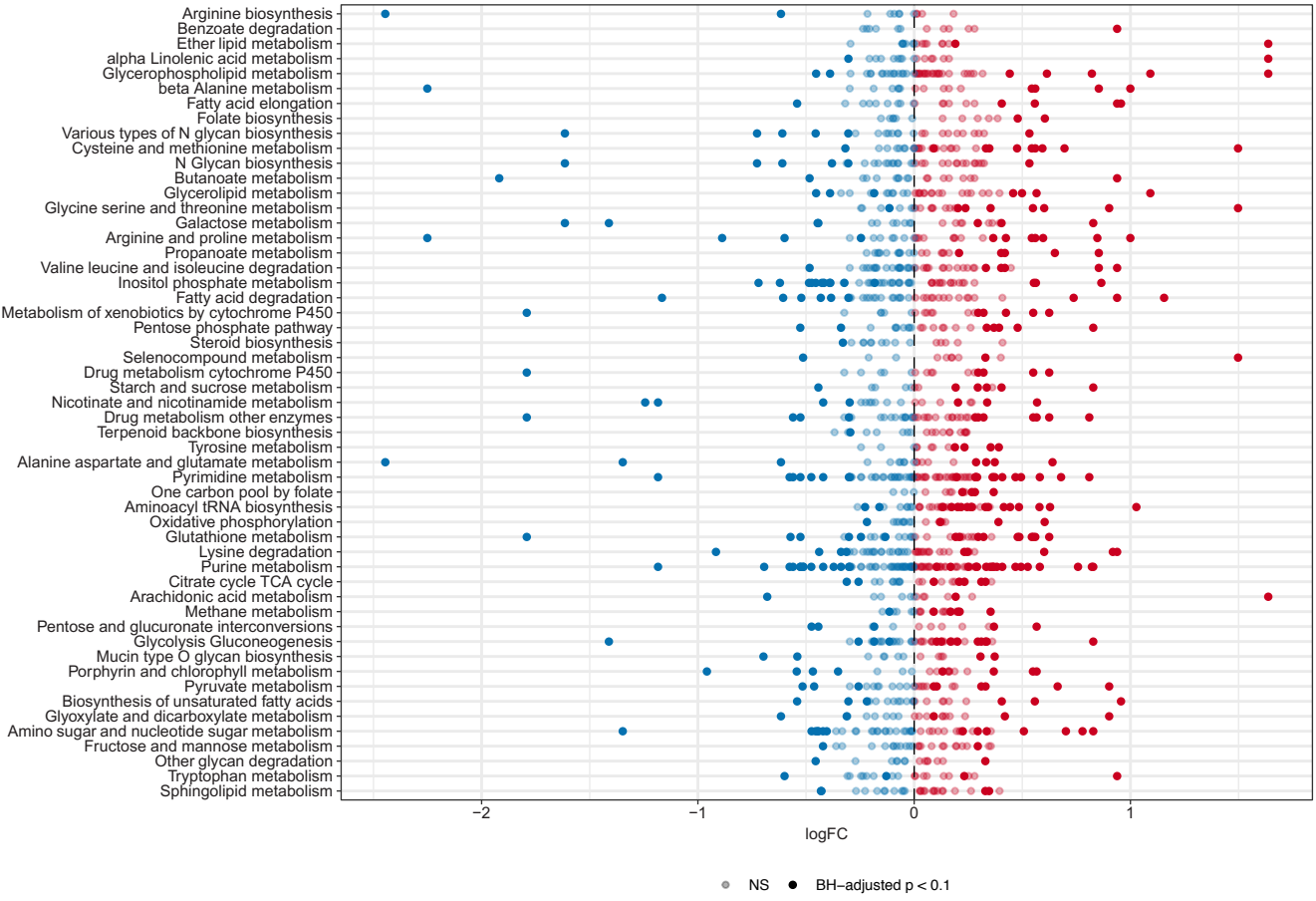

Dataset 1:
Th17p and Th17n (IL−17+ sorted)
Dataset 2:
Th17nu (unsorted)
Empirical CDF
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
10
100
reactions in meta-reaction

S2 Part 1


A
PCA on COMPASS space
PC2 (5.15% variance)
PC1 (23.56% variance)
Th17p
Th17n
PCA on COMPASS space
PC3 (4.55% variance)
PC1 (23.56% variance)
Th17p
Th17n
B
Alanine and aspartate metabolism
Blood group synthesis
Eicosanoid metabolism
Fatty acid oxidation
Fatty acid synthesis
Glycolysis/gluconeogenesis
N−glycan synthesis
Nucleotide interconversion
Purine catabolism
Sphingolipid metabolism
Squalene and cholesterol synthesis
Steroid metabolism
Tyrosine metabolism
Urea cycle
Valine, leucine, and isoleucine metabolism
PC1_pos
PC1_neg
PC2_pos
PC2_neg
PC3_pos
PC3_neg
PC4_pos
PC4_neg
PC5_pos
PC5_neg
C
Metabolic activity
PC1 (23.56% variance)
Th17p
Th17n
Th17n differentiation time course
Late
Early
D
CCL3
CCL4
TBX21
LRMP
CASP1
ICOS
IL22
IL1RN
LAG3
IL7R
CCL5
CXCL3
GZMB
LGALS3
CSF2
STAT4
IL6ST
CD5L
IL10
IKZF3
AHR
MAF
IL9
PC1
PC2
PC3
PC4
PC5
Spearman rho
E
Kegg_fatty_acid_metabolism
Kegg_regulation_of_actin_cytoskeleton
Nacl
Kegg_proximal_tubule_bicarbonate_reclamation
Kegg_nitrogen_metabolism
Kegg_glycine_serine_and_threonine_metabolism
Hallmark_unfolded_protein_response
Hallmark_mtorc1_signaling
Hallmark_uv_response_up
Hallmark_adipogenesis
Hallmark_xenobiotic_metabolism
All_genes
Hallmark_glycolysis
All_metabolic_genes
Kegg_peroxisome
Kegg_glycerophospholipid_metabolism
Hallmark_myc_targets_v2
Sgk1−il23
Batf
Kegg_rna_polymerase
Stat5
Patho_manual
PC1
PC2
PC3
PC4
PC5
Spearman rho
F
−log10(Wilcoxon-adjusted p)
Th17n
Cohen's d
Th17p
G
Central carbon metabolism
Amino acid metabolism
Lipid metabolism
Nucleotides
Other
Regulatory module score
Pathogenic module score
Putrescine diamine oxidase
Long−chain fatty acid:CoA ligase
Arginine decarboxylase
Ornithine carbamoyltransferase
Carnitine O−palmitoyltransferase
Agmatinase
Fumarase
Squalene synthase
Phosphoglycerate mutase (reverse)
Phosphopentomutase
Phosphatidylserine decarboxylase
Phosphoglycerate mutase
HMGCR
Serine C−palmitoyltransferase
Triose−phosphate isomerase
Glucosylceramidase
Adenosylhomocysteinase
Cystathionine beta−synthase
Cystathionine beta−synthase
Thymidine phosphorylase (rev.)

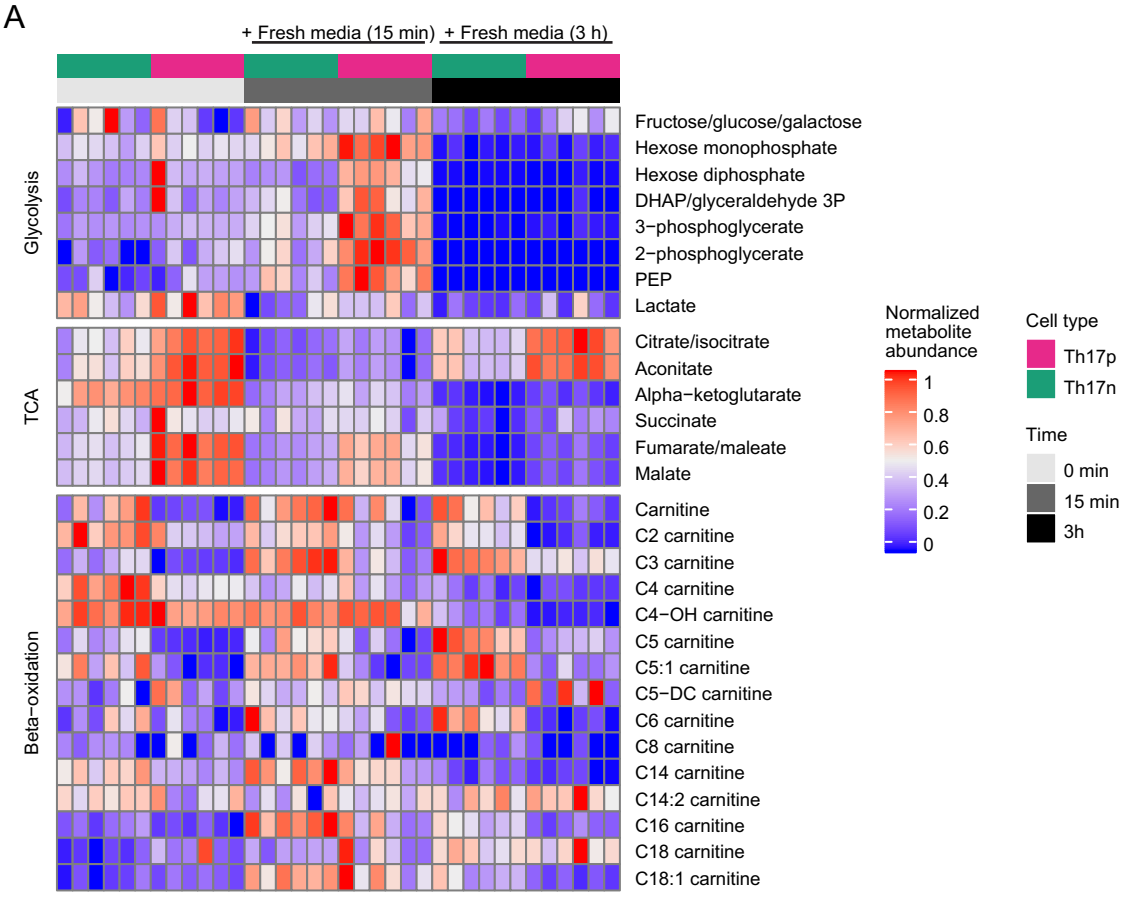
S2 Part 2

H

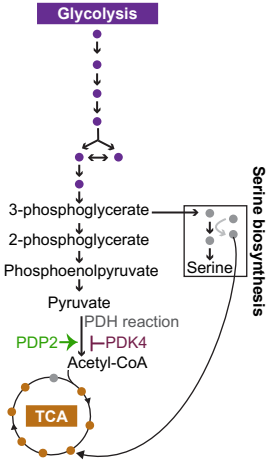

Arginine biosynthesis
Benzoate degradation
Ether lipid metabolism
alpha Linolenic acid metabolism
Glycerophospholipid metabolism
beta Alanine metabolism
Fatty acid elongation
Folate biosynthesis
Various types of N glycan biosynthesis
Cysteine and methionine metabolism
N Glycan biosynthesis
Butanoate metabolism
Glycerolipid metabolism
Glycine serine and threonine metabolism
Galactose metabolism
Arginine and proline metabolism
Propanoate metabolism
Valine leucine and isoleucine degradation
Inositol phosphate metabolism
Fatty acid degradation
Metabolism of xenobiotics by cytochrome P450
Pentose phosphate pathway
Steroid biosynthesis
Selenocompound metabolism
Drug metabolism cytochrome P450
Starch and sucrose metabolism
Nicotinate and nicotinamide metabolism
Drug metabolism other enzymes
Terpenoid backbone biosynthesis
Tyrosine metabolism
Alanine aspartate and glutamate metabolism
Pyrimidine metabolism
One carbon pool by folate
Aminoacyl tRNA biosynthesis
Oxidative phosphorylation
Glutathione metabolism
Lysine degradation
Purine metabolism
Citrate cycle TCA cycle
Arachidonic acid metabolism
Methane metabolism
Pentose and glucuronate interconversions
Glycolysis Gluconeogenesis
Mucin type O glycan biosynthesis
Porphyrin and chlorophyll metabolism
Pyruvate metabolism
Biosynthesis of unsaturated fatty acids
Glyoxylate and dicarboxylate metabolism
Amino sugar and nucleotide sugar metabolism
Fructose and mannose metabolism
Other glycan degradation
Tryptophan metabolism
Sphingolipid metabolism
−2
−1
0
1
logFC
NS
BH−adjusted p < 0.1

S3 Part 1

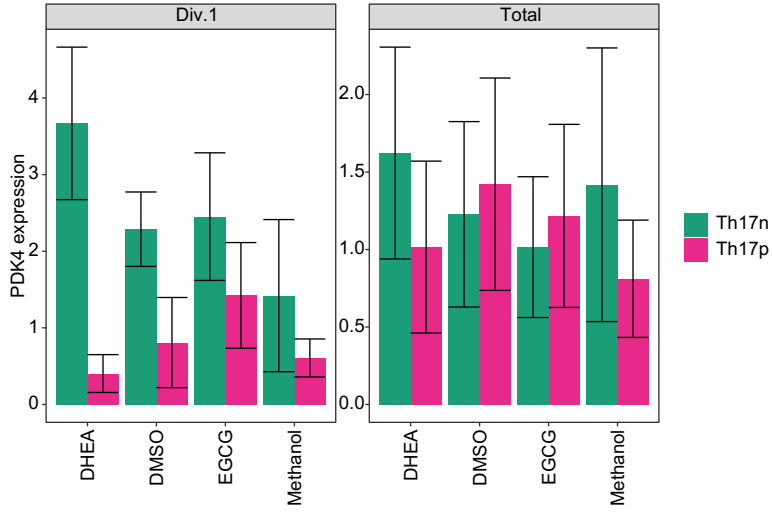
A


+ Fresh media (15 min)
+ Fresh media (3 h)
Glycolysis
TCA
Beta−oxidation
Fructose/glucose/galactose
Hexose monophosphate
Hexose diphosphate
DHAP/glyceraldehyde 3P
3−phosphoglycerate
2−phosphoglycerate
PEP
Lactate
Citrate/isocitrate
Aconitate
Alpha−ketoglutarate
Succinate
Fumarate/maleate
Malate
Carnitine
C2 carnitine
C3 carnitine
C4 carnitine
C4−OH carnitine
C5 carnitine
C5:1 carnitine
C5−DC carnitine
C6 carnitine
C8 carnitine
C14 carnitine
C14:2 carnitine
C16 carnitine
C18 carnitine
C18:1 carnitine
Normalized metabolite abundance
1
0.8
0.6
0.4
0.2
0
Cell type
Th17p
Th17n
Time
0 min
15 min
3h

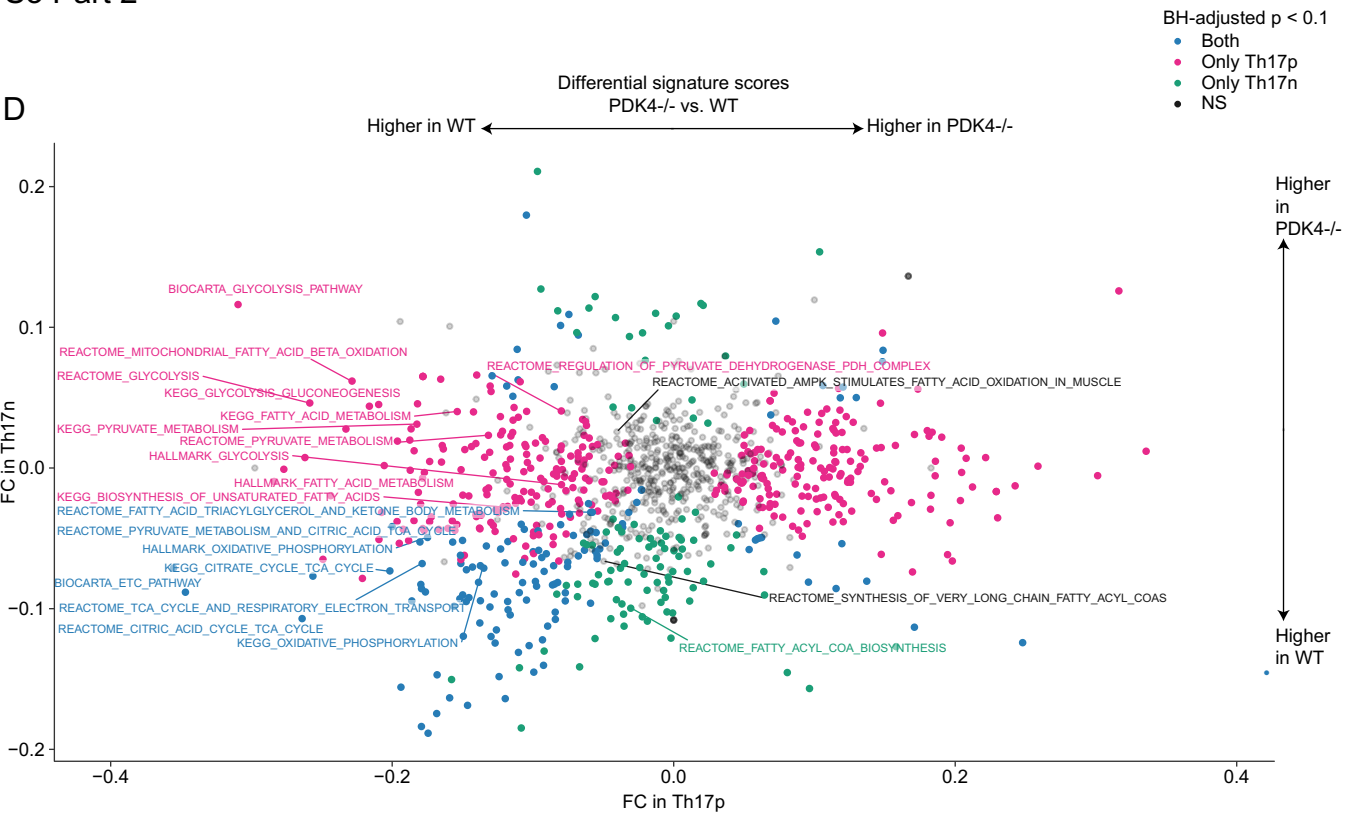
B

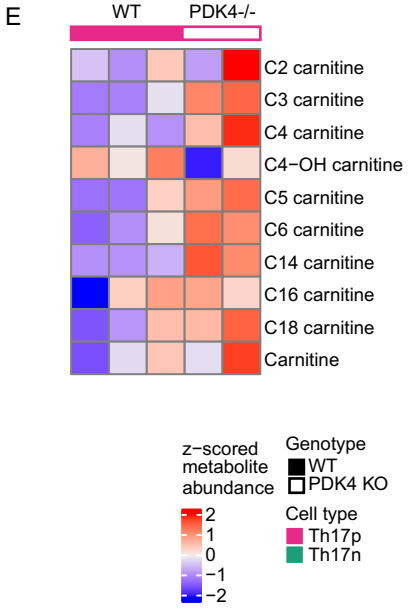

Glycolysis
3-phosphoglycerate
2-phosphoglycerate
Phosphoenolpyruvate
Pyruvate
PDH reaction
PDP2
PDK4
Acetyl-CoA
TCA
Serine
Serine biosynthesis

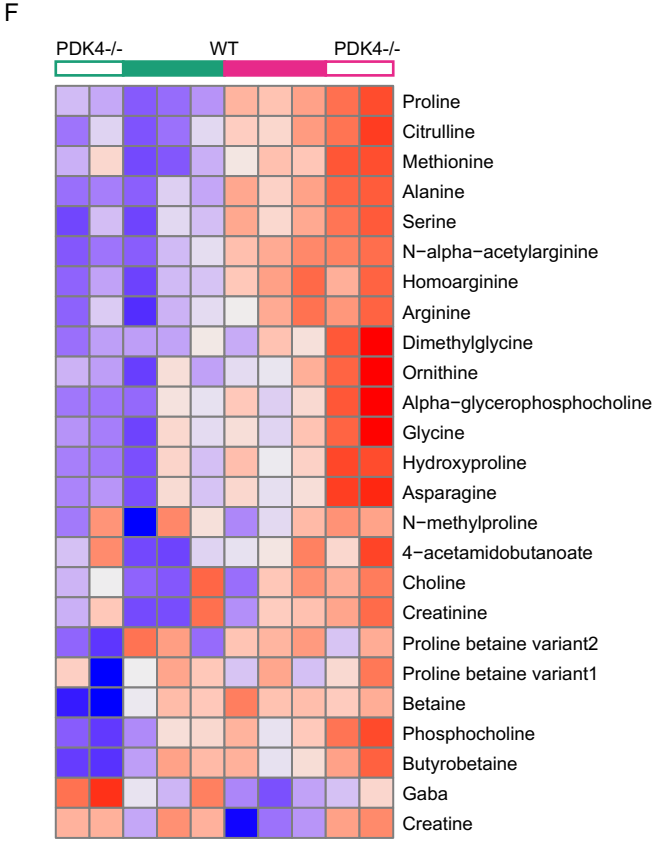
C


Div.1
Total
PDK4 expression
0
1
2
3
4
0.0
0.5
1.0
1.5
2.0
DHEA
DMSO
EGCG
Methanol
Th17n
Th17p

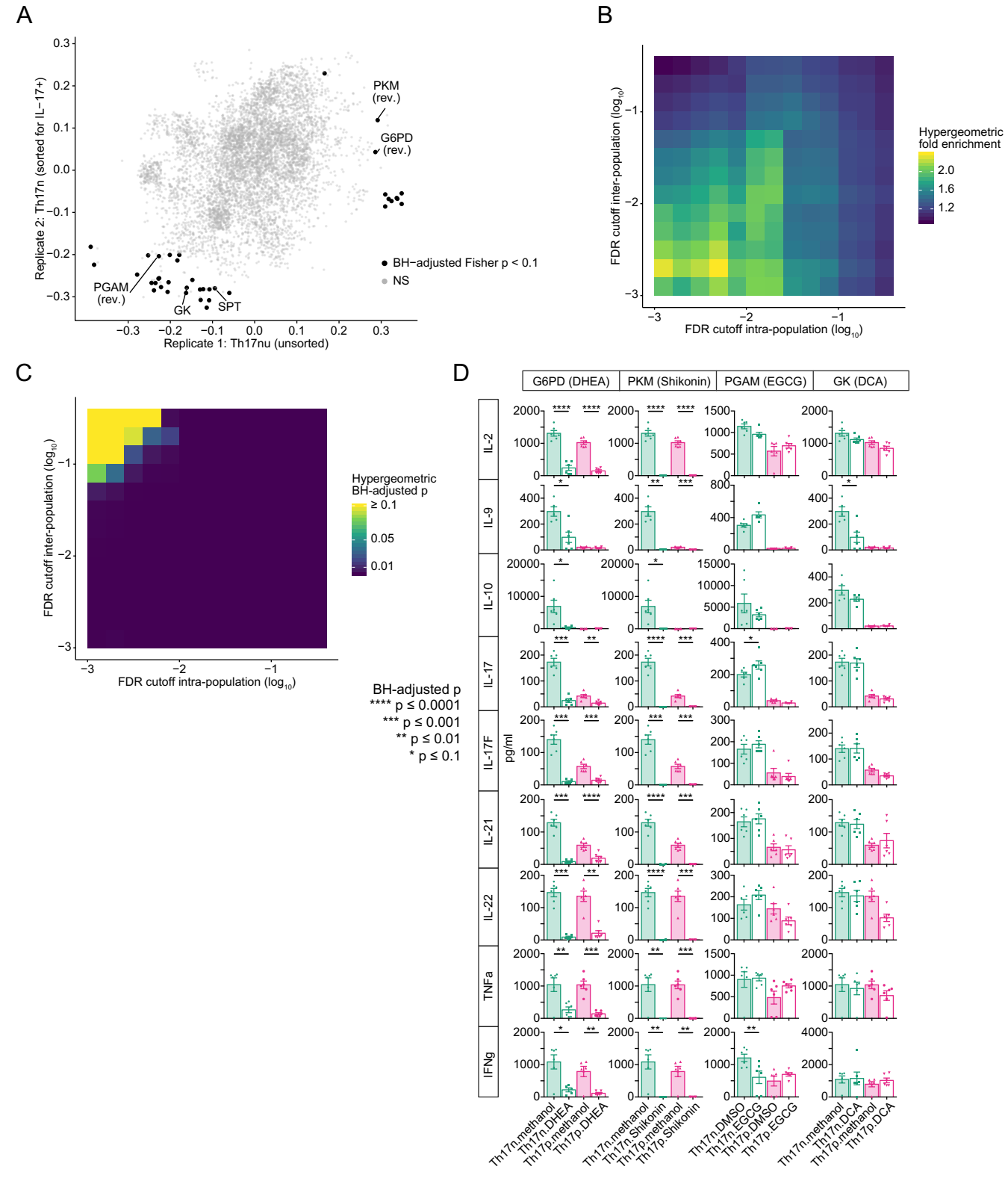
S3 Part 2


D
BH-adjusted p < 0.1
Both
Only Th17p
Only Th17n
NS
Differential signature scores
PDK4-/- vs. WT
Higher in WT
Higher in PDK4-/-
Higher in PDK4-/-
Higher in WT
BIOCARTA_GLYCOLYSIS_PATHWAY
REACTOME_MITOCHONDRIAL_FATTY_ACID_BETA_OXIDATION
REACTOME_GLYCOLYSIS
KEGG_GLYCOLYSIS_GLUCONEOGENESIS
KEGG_FATTY_ACID_METABOLISM
KEGG_PYRUVATE_METABOLISM
REACTOME_PYRUVATE_METABOLISM
HALLMARK_GLYCOLYSIS
HALLMARK_FATTY_ACID_METABOLISM
KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS
REACTOME_FATTY_ACID_TRIACYLGLYCEROL_AND_KETONE_BODY_METABOLISM
REACTOME_PYRUVATE_METABOLISM_AND_CITRIC_ACID_TCA_CYCLE
HALLMARK_OXIDATIVE_PHOSPHORYLATION
KEGG_CITRATE_CYCLE_TCA_CYCLE
BIOCARTA_ETC_PATHWAY
REACTOME_TCA_CYCLE_AND_RESPIRATORY_ELECTRON_TRANSPORT
REACTOME_CITRIC_ACID_CYCLE_TCA_CYCLE
KEGG_OXIDATIVE_PHOSPHORYLATION
REACTOME_REGULATION_OF_PYRUVATE_DEHYDROGENASE_PDH_COMPLEX
REACTOME_ACTIVATED_AMPK_STIMULATES_FATTY_ACID_OXIDATION_IN_MUSCLE
REACTOME_SYNTHESIS_OF_VERY_LONG_CHAIN_FATTY_ACYL_COAS
REACTOME_FATTY_ACYL_COA_BIOSYNTHESIS
FC in Th17n
FC in Th17p
0.2
0.1
0.0
−0.1
−0.2
−0.4
−0.2
0.0
0.2
0.4


E
WT
PDK4-/-
C2 carnitine
C3 carnitine
C4 carnitine
C4−OH carnitine
C5 carnitine
C6 carnitine
C14 carnitine
C16 carnitine
C18 carnitine
Carnitine
z−scored metabolite abundance
2
1
0
−1
−2
Genotype
WT
PDK4 KO
Cell type
Th17p
Th17n

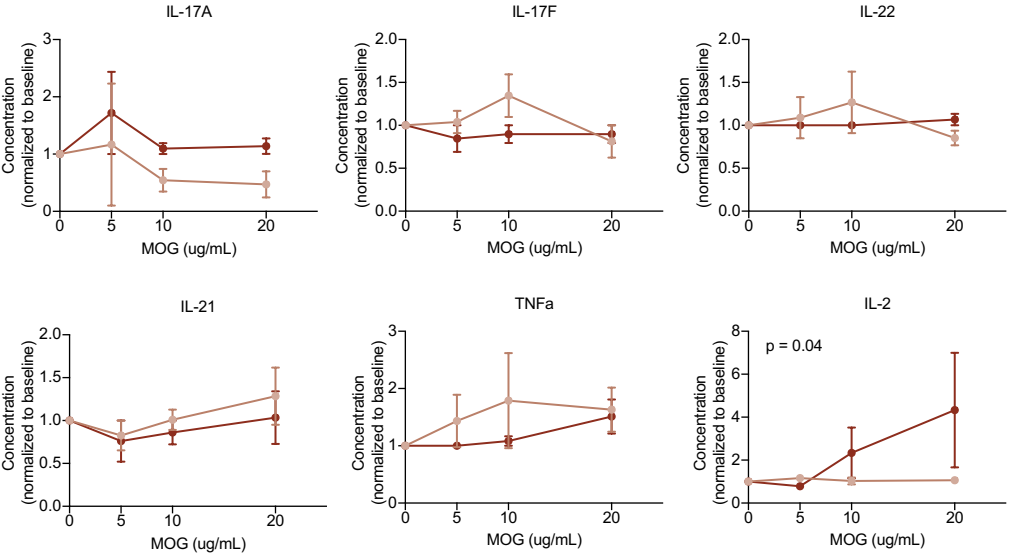

F
PDK4-/-
WT
PDK4-/-
Proline
Citrulline
Methionine
Alanine
Serine
N−alpha−acetylarginine
Homoarginine
Arginine
Dimethylglycine
Ornithine
Alpha−glycerophosphocholine
Glycine
Hydroxyproline
Asparagine
N−methylproline
4−acetamidobutanoate
Choline
Creatinine
Proline betaine variant2
Proline betaine variant1
Betaine
Phosphocholine
Butyrobetaine
Gaba
Creatine

S4

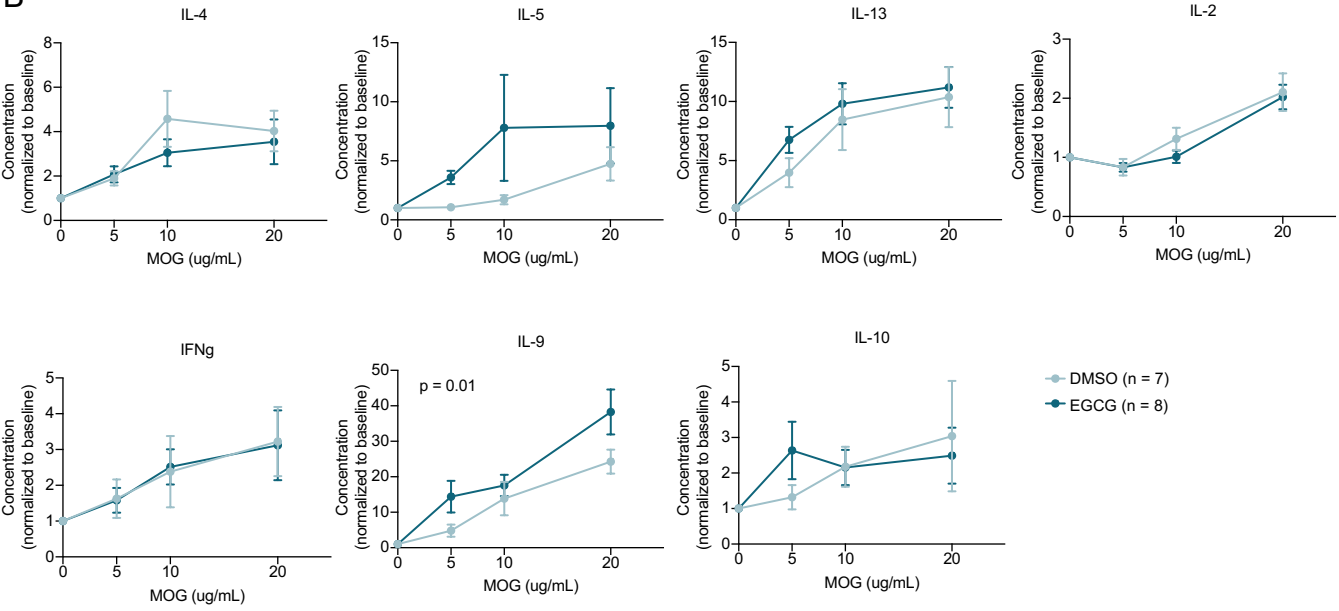

A
Replicate 2: Th17n (sorted for IL−17+)
Replicate 1: Th17nu (unsorted)
PKM (rev.)
G6PD (rev.)
PGAM (rev.)
GK
SPT
BH−adjusted Fisher p < 0.1
NS
B
FDR cutoff inter-population (log10)
FDR cutoff intra-population (log10)
Hypergeometric fold enrichment
2.0
1.6
1.2
C
FDR cutoff inter-population (log10)
FDR cutoff intra-population (log10)
Hypergeometric BH-adjusted p
≥ 0.1
0.05
0.01
BH-adjusted p
**** p ≤ 0.0001
*** p ≤ 0.001
** p ≤ 0.01
* p ≤ 0.1
D
G6PD (DHEA)
PKM (Shikonin)
PGAM (EGCG)
GK (DCA)
IL-2
IL-9
IL-10
IL-17
IL-17F
IL-21
IL-22
TNFa
IFNg
pg/ml
Th17n.methanol
Th17n.DHEA
Th17p.methanol
Th17p.DHEA
Th17n.methanol
Th17n.Shikonin
Th17p.methanol
Th17p.Shikonin
Th17n.DMSO
Th17n.EGCG
Th17p.DMSO
Th17p.EGCG
Th17n.methanol
Th17n.DCA
Th17p.methanol
Th17p.DCA

S5

A


IL-17A
IL-17F
IL-22
IL-21
TNFa
IL-2
p = 0.04
Concentration (normalized to baseline)
MOG (ug/mL)
Methanol (n = 4)
DHEA (n = 3)

B


IL-4
IL-5
IL-13
IL-2
IFNg
IL-9
p = 0.01
IL-10
Concentration (normalized to baseline)
MOG (ug/mL)
DMSO (n = 7)
EGCG (n = 8)