In Silico Metabolic Modeling of Single Th17 Cells Reveals Regulators of Autoimmunity

Allon Wagner

Electrical Engineering and Computer Sciences
University of California, Berkeley

Technical Report No. UCB/EECS-2023-23
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2023/EECS-2023-23.html

May 1, 2023

Acknowledgement

I was fortunate to work with wonderful mentors. I thank my thesis advisor, Prof. Nir Yosef, for the wise guidance, support, and incisive criticism where it was due. Before that I was privileged to work with Prof. Eytan Ruppin. They set two very different scientific role models to follow.

I owe immense gratitude to my collaborators and co-authors. The long collaboration with Chao Wang and Vijay K. Kuchroo has been enjoyable as it had been productive. Uri Amit could always be trusted to provide an engaging idea from a clinician’s perspective. An early review paper I wrote with Nir and Aviv Regev has been an extraordinary learning experience.

Finally, I would like to thank my lab mates and the exceptional staff of Berkeley’s EECS department. Go Bears!

In Silico Metabolic Modeling of Single Th17 Cells Reveals Regulators of Autoimmunity

By

Allon Shlomo Wagner

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering and Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Nir Yosef, Chair
Professor Jennifer Listgarten
Professor Ellen Robey

Spring 2021

Abstract

In Silico Metabolic Modeling of Single Th17 Cells Reveals Regulators of Autoimmunity

by

Allon Shlomo Wagner

Doctor of Philosophy in Electrical Engineering and Computer Science

University of California, Berkeley

Professor Nir Yosef, Chair

Metabolism is a major regulator of immune cell function, but it remains difficult to study the metabolic status of individual cells with current technologies. Here, we present Compass, an algorithm to characterize cellular metabolic states based on single-cell RNA sequencing (scRNA-Seq) and flux balance analysis. We applied Compass to associate metabolic states with Th17 functional variability (pathogenic potential) and recovered a metabolic switch between glycolysis and fatty acid oxidation, akin to known Th17/Treg differences, which we validated by metabolic assays. Compass also predicted that Th17 pathogenicity was associated with arginine and downstream polyamine metabolism. Indeed, polyamine-related enzymes expression were enhanced in pathogenic Th17 and suppressed in Treg cells. Chemical and genetic perturbation of polyamine metabolism inhibited Th17 cytokines, promoted Foxp3 expression, and remodeled the transcriptome and epigenome of Th17 cells towards a Treg-like state. *In vivo* perturbations of the polyamine pathway altered the phenotype of encephalitogenic T cells and attenuated tissue inflammation in central nervous system (CNS) autoimmunity.

The introduction highlights the motivation to this study, which stems from the conjunction of two transformative developments of recent years – the emergence of single-cell RNA sequencing technologies and the growing appreciation of cellular metabolism as key player in health and disease. Chapter 1 introduces lays the groundwork by introducing the fields of computational modeling of metabolism, immunometabolism, and single-cell genomics. Chapter 2 then introduces the Compass algorithm that serves as the computational framework to the rest of the study. Chapter 3 turns to T helper 17 (Th17) cells and emphasizes their diverse effector phenotype, which makes them an attractive system to query with computationally-informed metabolic methods. A Compass-based study reveals a parallel diversity of metabolic phenotypes within the Th17 cell type. In the next chapters we demonstrate our computational framework's ability to discover metabolic regulators of Th17 inflammatory potential (*pathogenicity)*. Chapters 4 and 5 complement one another; chapter 4 restricts its analysis to the well-studied central carbon metabolism pathways, whereas chapter 5 presents an unsupervised network-wide analysis that uncovers a novel metabolic regulator in the peripheral polyamine pathway. Chapter 6 follows on this discovery and shows that chemical and genetic

perturbations of the polyamine pathway lead to a sizable shift in both the molecular and effector profiles of Th17 cells. It is suggested that polyamine metabolism might be a novel therapeutic target in autoimmune disorders by showing that *in vivo* inhibition of two different enzymes in the pathway alleviates experimental autoimmune encephalomyelitis (EAE) – a murine model for human multiple sclerosis.

To my parents

Contents

Chapter 3......28

Querying the metabolic states of single T helper 17 (Th17) cells 28

Chapter 4......34

Computationally-informed analysis of central carbon metabolism 34

Chapter 5......45

Computationally-informed analysis of peripheral metabolism......45

Chapter 6......53

The polyamine pathway regulates Th17-induced autoimmunity......53

Chapter 7......61

Conclusions......61

Chapter 8......63

List of Figures

Acknowledgements

I was fortunate to work with wonderful mentors. I thank my thesis advisor, Prof. Nir Yosef, for the wise guidance, support, and incisive criticism where it was due. Before that I was privileged to work with Prof. Eytan Ruppin. They set two very different scientific role models to follow.

I owe immense gratitude to my collaborators and co-authors. The long collaboration with Chao Wang and Vijay K. Kuchroo has been enjoyable as it had been productive – none of the work presented herein would have been possible if it were not for them – and I hope will continue in the years to come. Uri Amit could always be trusted to provide an engaging idea from a clinician's perspective. An early review paper I wrote with Nir and Aviv Regev has been an extraordinary learning experience.

Finally, I would like to thank my lab mates and the exceptional staff of Berkeley's EECS department. Go Bears!

Chapter 1

Introduction

1.1 Setting the table, part 1: metabolism and computation

1.1.1 Metabolism and immunometabolism – an (almost) forgotten field is revived and a new field is born

The early- and mid-twentieth century were a golden era for biochemistry and metabolic research (Chandel, 2014). Seminal discoveries, such as the discovery of the glycolysis pathway, the Krebs cycle, the Warburg effect, the Cori cycle, and the urea cycle were made, and awarded their discoverers with Nobel recognition. However, with the discovery of DNA and growing understanding of and interest in genomics, focus shifted away from metabolism. In addition, many metabolic genes are lethal if deleted in model organisms, and their biallelic deficiency in human is rare and often results in debilitating cognitive and physical impairments (Hamosh et al., 2000; Marsden et al., 2006; Pampols, 2010). These contributed to a common perception of metabolism as a largely homeostatic process that runs parallel to other cellular circuits, more relevant to human disease.

The tides began to shift ca. 2008, first in the discipline of cancer research, where altered energy metabolism gained universal recognition is a hallmark of cancer (Hanahan and Weinberg, 2011) alongside previous hallmarks of the disease (Hanahan and Weinberg, 2000). Other discoveries demonstrated that metabolic dysregulation in cancer was not limited to energy pathways. One key discovery was that mutations in isocitrate dehydrogenase (IDH), common in multiple cancer types, result in neomorphic enzyme activity that converted α-ketoglutarate to the oncometabolite 2-hydroxyglutarate (2HG) (Dang et al., 2010; Ward et al., 2010), which interferes with DNA demethylation and can produce an hypermethylated epigenomic landscape (Figueroa et al., 2010). This early discovery exemplifies the confluence of two major paradigm shifts. First, the realization that altered metabolism is fundamental to tumor survival. Second, the understanding that the metabolic subsystem is not self-contained, and is capable of regulating cell behavior, for example through epigenomic (Mews et al., 2017; Wellen et al., 2009), epitranscriptomic (Cho et al., 2021), or post-transcriptional (Chang et al., 2013) modifications.

Aberrant metabolism thus became universally recognized as a major mediator and therapeutic target in almost any major health concern: cancer, neurodegeneration, cardiovascular disorders, and even normal aging. The stage was set for the emergence of the discipline of *immunometabolism*, the study of metabolism in immune cells. Today, immunometabolism is a copiously active field of research. It has revealed metabolic regulation of inflammation in almost all contexts, including anti-viral immunity, autoimmunity, and anti-tumor response (Buck et al., 2017; Certo et al., 2020; Chapman et al., 2019; Diskin et al., 2021; Elia and Haigis, 2021; Geltink et al., 2018; Ho and Kaech, 2017; Hotamisligil, 2017; Jung et al., 2019; Makowski et al., 2020; O'Neill et al., 2016; Roy et al., 2021; Russell et al., 2019; Varanasi et al., 2020; Voss et al., 2021).

1.1.2 Flux Balance Analysis (FBA) – metabolism begs computation

Cellular metabolism is commonly abstracted as a *network*, with nodes being metabolites and a directed edge *(u,v)* connecting metabolites *u,v* if there exists a reaction with *u* as a substrate and *v* as a product (see below for generalizations of the network abstraction). The metabolic network is highly complex and interconnected, even in unicellular model organisms, and its network theoretical properties have been studied and compared to other natural and man-made networks (Newman, 2018). Standard visual representations of metabolic pathways, as well as common metabolic databases (KEGG, MetaCyc) omit some of the edges for accessibility and simplicity.

Due to the scale and complexity of the metabolic network, a metabolic perturbation can create cascading effects and alter a seemingly distant part of the network or cut across 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.

Chemical stoichiometry describes the conversion of substrates into products (Petrucci et al., 2016), for instance:

$$2 \cdot \mathrm{Mg} + \mathrm{O_2} \longrightarrow 2 \cdot \mathrm{MgO}$$

Here, 2 moles of magnesium (Mg) react with one mole of dioxygen (O_2, with each molecule of dioxygen containing 2 oxygen atoms) to produce 2 moles of magnesium oxide (MgO). The coefficients needed to preserve mass balance in the equation are called the *stoichiometric coefficients*.

In general notation, reactions take the form:

$$a \cdot X + b \cdot Y + \cdots \longrightarrow c \cdot Z + \cdots$$

The mathematical object describing biochemical stoichiometry is the formal representation of the metabolic network. This representation may take many forms depending on the computational goals and contexts.

By neglecting the stoichiometric coefficients and retaining only the ordered relation $\langle\{X, Y\}, \{Z\}\rangle$, we may view the metabolic network as a hypergraph, where nodes are metabolites and directed hyperedges connect substrates to products. The hypergraph is often simplified by transformation into a bipartite (two-mode) graph, where nodes are either reaction or metabolites and directed edges connect substrates to reactions and reactions to products (Jha et al., 2015). Multiple studies have simplified the bipartite graph further to one-mode graphs of reactions or metabolites, where directed edges represent 2-step paths from nodes in the bipartite graph (Palsson, 2015). If this procedure is applied to transform the bipartite graph into a one-mode graph with metabolite nodes, we obtain the standard network representation of metabolism presented above.

A biochemical reaction, or an hyperedge in the metabolic hypergraph, is associated with metadata. The stoichiometric coefficients were already discussed. Under physiological conditions, many of the biochemical reactions are enzymatically catalyzed, and the enzymes or genes coding them are associated with the hyperedge as well. Notably, many reactions can be catalyzed by isozymes (enzymes with different amino-acid sequence capable of catalyzing the same reaction), or require a protein complex, coded by multiple genes, to take place. This naturally gives rise to boolean expressions over gene literal, where *OR* and *AND* relations correspond to genes coding isozymes and parts of protein complexes, respectively. Genetic deletions are then modeled by assigning truth values to gene literals according to their functional status, and discarding reactions whose associated boolean expression is unsatisfied (Heirendt et al., 2019).

Computational treatments of metabolism that take account of the stoichiometric coefficients model *biochemical fluxes*, namely the instantaneous rate by which substrates are converted to products. These methods largely fall into one of two broad categories. One approach is modeling metabolic reaction kinetics by ODEs (Hahl and Kremling, 2016; Palsson, 2011). The major obstacles to this approach are computational tractability, limiting the number of feasibly simulated reactions in the system, and the scarcity of available data concerning pertinent kinetic constants in human physiological conditions. A second approach eliminates kinetic constants by assuming chemical steady-state (Palsson, 2015). Hybrid approaches have been suggested as well (Antoniewicz, 2013; Mahadevan et al., 2002; Martínez et al., 2015; Yugi et al., 2005).

Flux balance analysis (FBA) is a family of algorithms for flux modeling on genome-scale metabolic network. The vector of flux values per reaction (*flux distribution*) is constrained by thermodynamic (e.g., reaction irreversibility) and chemical (e.g., nutrient availability), which limit the space of feasible flux distributions (Orth et al., 2010). FBA then employs constraint-based optimization to solve for points of interest in the high-dimensional flux distributions space (Burgard et al., 2003; Lewis et al., 2010; Segrè et al., 2002; Wagner et al., 2013; Yizhak et al., 2013).

1.2 **Setting the table, part 2: single-cell genomics**

1.2.1 The emergence of genome-scale single-cell assays

Until recently, most genomic profiling studies have analyzed cell populations, although even cells of the same 'type' can exhibit substantial heterogeneity, reflecting finer sub-types, regulated functional variation, or inherent stochasticity (Altschuler and Wu, 2010; Ben-Moshe et al., 2019; Gaublomme et al., 2015; Grun et al., 2015; Halpern et al., 2017; Hildreth et al., 2021; Katzenelenbogen et al., 2020; Kennedy et al., 2020; Kumar et al., 2019; Levine et al., 2021; Shalek et al., 2013, 2014; Zeisel et al., 2015). However, during the 2010's, rapid technological advances have enabled genome-wide profiling of RNA (Hashimshony et al., 2012; Klein et al., 2015; Macosko et al., 2015; Picelli et al., 2013; Ramskold et al., 2012; Tang et al., 2009), DNA (Hou et al., 2012; Leung et al., 2015; Lohr et al., 2014; Navin et al., 2011; Wang et al., 2014; Xu et al., 2012; Zong et al., 2012), protein (Bandura et al., 2009; Bendall et al., 2011; Bodenmiller et al., 2012; Chattopadhyay et al., 2006), epigenetic modifications (Farlik et al., 2015; Guo et al., 2013, 2014; Rotem et al., 2015; Smallwood et al., 2014), chromatin accessibility (Buenrostro et al., 2015; Cusanovich et al., 2015), and other molecular events (Nagano et al., 2013) in single cells. Specialized assays capable of measuring more than one of these data modalities in single cells (Angermueller et al., 2016; Dey et al., 2015; Kearney et al., 2021; Lareau et al., 2021; Macaulay et al., 2015; Mimitou et al., 2019; Stoeckius et al., 2017), measuring single-nucleus RNA (Grindberg et al., 2013; Habib et al., 2016; Lacar et al., 2016; Lake et al., 2016), which is particularly suitable for frozen tissues, have since been introduced. Efforts are also invested towards single-cell assays of RNA in its spatial context. Pioneering technologies conceded the individual cell resolution (Ståhl et al., 2016), or the comprehensiveness of transcriptomic profiling (Chen et al., 2015; Shah et al., 2016). More recent work is now overcoming these barriers with microscopy (Eng et al., 2019) or sequencing systems (Lee et al., 2021; Rodriques et al., 2019; Stickels et al., 2020; Vickovic et al., 2019).

1.2.2 Single-cell RNA sequencing – challenges and opportunities

In this work, we study single-cell RNA-Sequencing (scRNA-Seq), one of the most widely used data modalities in single-cell functional genomics. The scale and precision of scRNA-Seq is continually increasing (Aldridge and Teichmann, 2020; Svensson et al., 2018), reaching millions of cells in massively parallel assays that multiplex samples through antibody (*cell hashing*) (Stoeckius et al., 2018) or lipid tagging (McGinnis et al., 2019). The newfound frontiers allowed addressing profound biological questions that were previously out of reach (Tanay and Regev, 2017) and undertaking the ambitious project to collaboratively create an atlas of all human cell types in health (Regev et al., 2017a; Rozenblatt-Rosen et al., 2021) and disease (Rozenblatt-Rosen et al., 2020). However, single-cell data poses unique challenges owing to its unique characteristics stemming from the underlying biotechnology (Hie et al., 2020; Wagner et al., 2016), and the new questions that it can address need novel computational approaches to be answered (Argelaguet et al., 2021; Lähnemann et al., 2020). Here, we present the challenges and opportunities of single-cell transcriptomics from a computational point of view. The conceptual framework we put forth is largely based on an early and influencial review, in which we anticipated some of the key research direction of the field (Wagner et al., 2016).

1.2.3 The many facets of a cell's identity

Single-cell genomics offers novel data modalities, and requires rethinking and formulating of our conceptual frameworks and mental models of its empirical results. Here, we offer our attempt to address the need to more formally define some concepts required for reasoning about single-cell genomics.

We define a cell's *identity* as the outcome of the instantaneous intersection of all factors that affect it **(Figure 1)**. We refer to the more permanent aspects in a cell's identity as its *type* (e.g., a hepatocyte typically cannot turn into a neuron) and to the more transient elements as its *state*. Cell types are often organized in a hierarchical taxonomy, as types may be further divided into finer subtypes; such taxonomies are often related to a cell fate map, reflecting key steps in differentiation. Cell *states* arise transiently during time-dependent processes, either in a *temporal progression* that is unidirectional (e.g., during differentiation, or following an environmental stimulus) or in a *state vacillation* that is not necessarily unidirectional and in which the cell may return to the origin state. Vacillating processes can be *oscillatory* (e.g., cell-cycle or circadian rhythm) or can transition between states with no predefined order (e.g., due to stochastic, or environmentally controlled, molecular events). These time-dependent processes may occur transiently within a stable cell type (as in a transient environmental response), or may lead to a new, distinct type (as in differentiation). A cell's identity is also affected by its *spatial context*, which includes the cell's absolute *location*, defined as its position in the tissue (e.g., the location of a cell along the dorsal–ventral axis during embryo development determines its exposure to a morphogen gradient), and the cell's *neighborhood*, or the identity of neighboring cells.

The cell's identity is manifested in its molecular contents. Genomic experiments measure these in *molecular profiles*, and computational methods infer information on the cell's identity from the measured molecular profiles (inevitably, the molecular profile also reflects allele-intrinsic and technical variation that must be handled properly by computational methods before any analysis is done). We refer to this as inferring *facets* of the cell's identity, or inferring *factors* that created the cell's identity, to stress that none describes it fully, but each is an important, distinguishable aspect.

By analogy, we relate the biological factors that shape a cell's identity to *basis vectors* that span a space of cell identities. Similarly, the computationally inferred facets of the cell's identity are likened to basis vectors that span a space of cell molecular profiles. In many cases, computational analysis methods find such basis vectors directly (as discussed in main text) and these indeed relate well to biological facets of identity. However, this idealized definition, and the present computational tools, are likely to be insufficient to capture the true nature of this space. In particular, basis vectors in algebra are defined to be independent of each other, but facets of a cell's identity that we would like to distinguish and identify separately—such as its type, location, and state—may be largely dependent on one another. For example, the spatial position of a cell in a solid organ is a fixed element of its identity that is usually distinguished from its 'type' but is nevertheless not independent of cell type. In another example, whereas a cell cycle phase may have invariant characteristics across systems (Kowalczyk et al., 2015; Macosko et al., 2015; Tirosh et al., 2016), the ability of a cell to enter the cell cycle and the duration of the phase can depend on cell type and can influence other temporal processes like differentiation. As the field of single-cell genomics develops, it may be possible to define abstractions, possibly employing data-driven categories rather than ones imposed by prior conceptions, that are mathematically precise and reflect the key biological components.

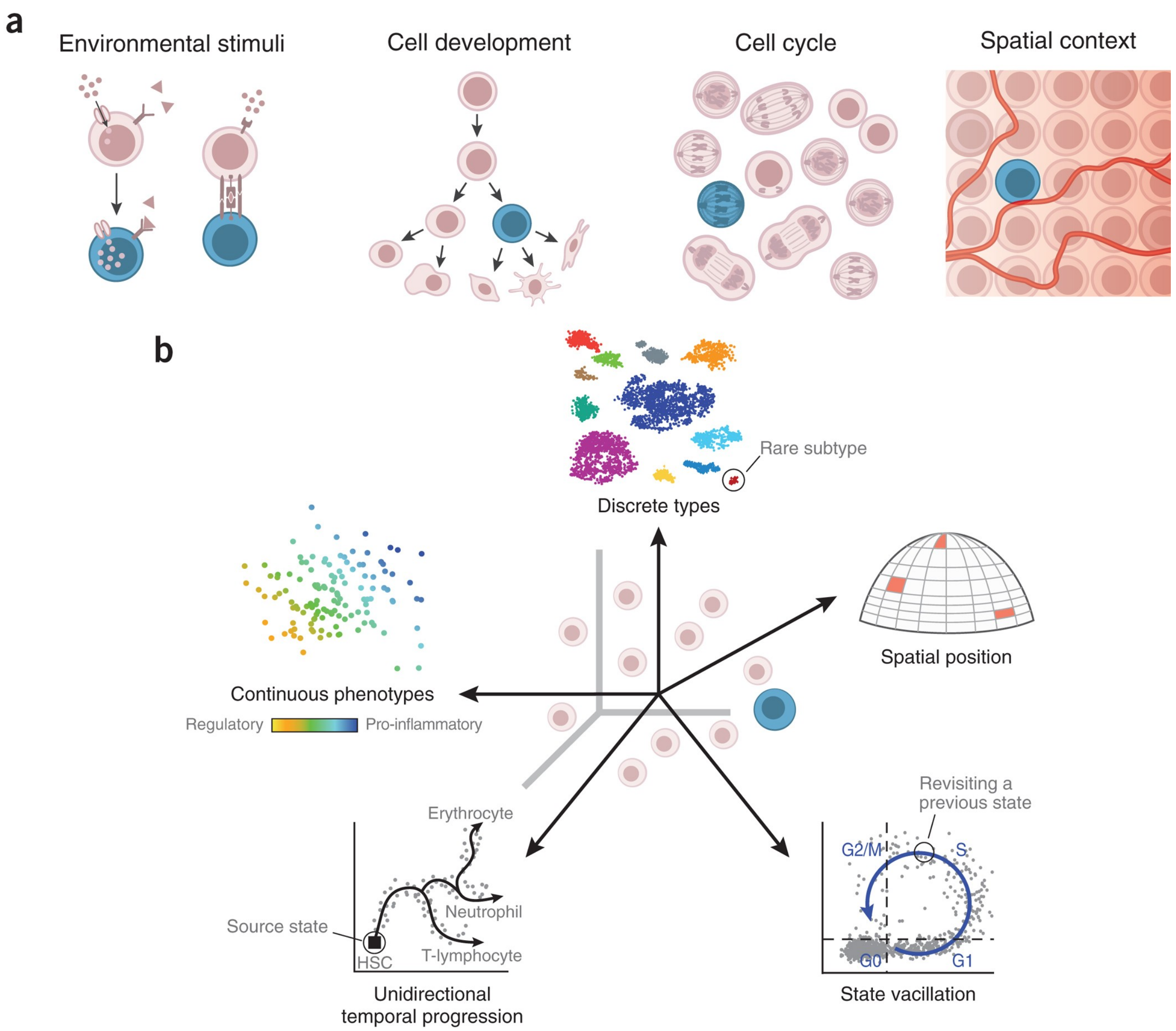


Figure 1. Diverse factors combine to create a cell's unique identity, and computational methods reveal them.

(**A**) A cell participates simultaneously in multiple biological contexts. The illustration depicts a particular cell (blue) as it experiences multiple contexts that shape its identity simultaneously (from left to right): environmental stimuli, such as nutrient availability or the binding of a signaling molecule to a receptor; a specific state on a developmental trajectory; the cell cycle; and a spatial context, which determines its physical environment (e.g., oxygen availability), cellular neighbors, and developmental cues (e.g., morphogen gradients). (**B**) The biological factors affecting the cell combine to create its unique, instantaneous identity, which is captured in the cell's molecular profile. Computational methods dissect the molecular profile and tease apart facets of the cell's identity, which are akin to 'basis vectors' that span a space of possible cellular identities. Key examples include (counterclockwise from top): (1) discrete cell types (e.g., cell populations in the retina (Macosko et al., 2015)); cell type frequency can vary by multiple orders of magnitude from the most abundant to the rarest subtype; (2) continuous phenotypes (e.g., the pro-inflammatory potential of each individual T cell, quantified through a gene expression signature derived from bulk pathogenic T cell profiles (Gaublomme et al., 2015)); (3) unidirectional temporal progression (e.g., normal differentiation, such as hematopoiesis); (4) temporal vacillation between cellular states (e.g., oscillation through cell cycle; (Kowalczyk et al., 2015)); (5) physical location (e.g., a cell's location during embryo development determines its exposure to morphogen gradients. Dividing an organ into discrete spatial bins, combined with independent data on landmark genes, allows inference of spatial bins (highlighted) from which single cells had likely originated (Satija et al., 2015). The scatterplots represent single cells (dots) projected onto two dimensions (e.g., first two principal components or non-linear projections such as t-SNE or UMAP).

1.2.4 Sources of biological and technical variation in single-cell RNA-sequencing

We distinguish three sources of variation in scRNA-seq **(Figure S1, top)**. The first is technical variation, which is due to factors such as differences in cell integrity and lysis, RNA capture and cDNA conversion, and detection (Kim et al., 2015; Kolodziejczyk et al., 2015). The second is allele-intrinsic variation, namely stochastic factors intrinsic to the molecular mechanisms that control gene expression (Raj and van Oudenaarden, 2008; Raj et al., 2006; Stewart-Ornstein et al., 2012). For example, the bursting statistics of transcriptional initiation coupled to variable rates of mRNA degradation can lead to fluctuations in transcript levels over time in one cell, and to differences between otherwise 'identical' cells measured at a single time point. This inherent stochasticity does not correlate between two alleles of the same gene. The third is allele-extrinsic variation, due to factors (Raj and van Oudenaarden, 2008; Swain et al., 2002) extrinsic to the process of transcription, such as the presence of certain regulators or differences in stable chromatin state. These factors contribute to establishing differences between cell types or states, either stably or transiently, but are correlated between two alleles of the same gene (Stewart-Ornstein et al., 2012).

Although most studies aim to understand allele-extrinsic variation and its function, technical and allele-intrinsic variations are major confounders. Some technical variation is common to both scRNA-seq and bulk (population) RNA-Seq, whereas several other factors—including zero inflation due to false negatives, overamplification, and cell doublets—are specific to the technical variation between single-cell profiles. In some cases, the extent of technical variation is affected by biological differences, undermining definitions of quality and limiting our ability to remove technical variation. For example, because smaller cells typically harbor less RNA, they appear to be lower in quality. Similarly, some cell types may be harder to capture or lyse. Finally, some cells are characterized by transcriptional profiles functionally dominated by very high expression of a few transcripts, whereas others have far more complex transcriptomes. Indeed, it was reported that technical quality features were highly correlated with biological cell type (Ilicic et al., 2016).

One has to address the unique analytical and technical challenges of scRNA-Seq before one can fully exploit its wealth of information towards biological discovery. The analytical challenges include (1) designing experiments and performing power analysis (e.g., how many cells do we need to profile for a given task? At what depth?); (2) preprocessing to distinguish biological from technical variation, especially false-negative gene detections (dropouts); (3) inferring the key aspects of a cell atlas, from discrete sub-types to continuous spatiotemporal ordering of cells; and (4) deriving molecular mechanisms from cell-to-cell variation. In each of these areas, we must grapple with common technical challenges, such as noise, sparsity, and false negatives; ever-increasing scale, which defies many traditional implementations of basic tasks in genomics; partial dependencies between the multiple facets of a cell's identity (its type, state, position, etc.), such that variation in one biological dimension may be a confounder for another; and the need for accessible and interpretable visualizations. Novel computational methods are required to overcome these challenges and exploit the biological signals in single-cell data **(Figure S1, bottom)**.

1.2.5 Single-cell transcriptomics allows addressing novel research questions

To understand a cell — the basic unit of life — we must determine the multiple factors that shape its identity. These include its position in a taxonomy of cell types, the progress of multiple time-dependent processes that take place simultaneously, the cell's response to signals from its local environment, and the precise location and neighborhood in which it resides **(Figure 1A)**. The factors that together span the space of possible cell identities can be likened to the basis vectors that span a linear space, yet, unlike basis vectors, they may be intricately dependent on one another.

Large-scale single-cell data allow us to address biological questions that were previously out of reach. First, we can now explore the identity of an individual cell and the factors underlying it through the comprehensive lens of the cell's unique molecular profile **(Figure 1B)**. By decomposing this profile to its separate components, it should be possible to determine, in a data-driven way, the specific physiological and molecular features of each of these factors, without relying on prior definitions, hypotheses, or markers. Second, construction of a comprehensive atlas of all human cell types and sub-types—including their activity states, dynamic transitions, physical locations, and lineage relationships through development—has become a tangible goal (Regev et al., 2017b; Rozenblatt-Rosen et al., 2017, 2021). Even preliminary progress toward such an atlas would help elucidate the organization and function of tissues in health and disease. In addition, single-cell data allow us to study the regulatory circuitry that governs cells at a resolution that had been impossible with data collected from bulk cell populations. Finally, single cells are the basic component of complex tissues. Through deconvolution of a complex sample, such as a tumor biopsy (Katzenelenbogen et al., 2020; Ringel et al., 2020; Tirosh et al., 2016), one may infer its cellular composition and characterize the rare (Grun et al., 2015), functionally important (Choi et al., 2021; Culemann et al., 2019; Mathewson et al., 2021), and unknown (Tasic et al., 2016; Vieira Braga et al., 2019) cell types it contains.

The factors that jointly define a cell's identity may not only be the discrete categories that are often assumed when classifying cells into major types (Zeisel et al., 2015) but may also represent a continuous spectrum (Gaublomme et al., 2015), or a combination of discrete and continuous categories (Antebi et al., 2013; Korem et al., 2015; Patel et al., 2014; Tasic et al., 2016). First, in contexts such as development and physiology, some facets of cellular identity are transient in time and space. Temporal processes may, for example, progress along one or more trajectories (e.g., differentiation, (Marioni and Arendt, 2017)); oscillate continuously between cellular states (e.g., the cell cycle (Kowalczyk et al., 2015; Pollen et al., 2015) or circadian rhythm (Lande-Diner et al., 2015); or be influenced by the physical position and neighborhood of the cell (Adler et al., 2019; Ben-Moshe et al., 2019; Halpern et al., 2017; Kennedy et al., 2020; Nitzan et al., 2019; Satija et al., 2015). Moreover, even within a type, cells may span a continuous range of functional phenotypes (e.g., T cells of a single type, but with a range of inflammatory versus regulatory phenotypes, (Gaublomme et al., 2015; Pompura et al., 2021)). While each such facet of a cell's identity is often considered separately, they are at least partly interdependent. Cataloging sources of biological variation, and understanding how they combine to determine a cell's identity, is an integral task in the compilation of a human cell atlas. This will reopen definitions of a cell's type and will allow a more sophisticated view of a cell, not only as an instance of a predefined type, but rather as the sum of the biological factors that shaped it. We expect that, eventually, the measured genomic profiles of a cell will be used to characterize these factors and to quantify their relative contributions to forming its identity. By analogy, the sources of biological variation that determine a cell's identity are akin to basis vectors that span a linear space—namely, their combinations produce all possible points in the space. However, unlike basis vectors in algebra, they can in fact be dependent, which further complicates their identification and interpretation, and poses a problem of statistical identifiability. For instance, cells in a given position may be more quiescent compared to other positions, which complicates the inference of the biological variation that should be attributed to cell cycle versus spatial position (Buettner et al., 2015).

1.3 Setting the table, part 3: T helper 17 and T regulatory cells – the odd siblings

The immune system has an integral suppressive branch to suppresses its inflammation-inducing branches, which allows resolution of inflammation once a pathogen has been removed limits damage by self-activating lymphocyte and chronic inflammation. CD4+ T regulatory (Treg) cells are thought to be one of the most important parts of the suppressive branch and are required for immune tolerance and homeostasis (Josefowicz et al., 2012; Sakaguchi et al., 2008, 2020). The transcription factor Forkhead box protein P3 (FOXP3) is a master regulator of the Treg gene expression program, although its role in human Treg cells is less understood than in mouse (Vignali et al., 2008). People suffering from genetic mutations that eliminate FOXP3 function develop a severe autoimmune disease called IPEX (immunodysregulation, polyendocrinopathy and enteropathy, X-linked syndrome in early infancy.

In vivo, T regulatory cells can develop in the thymus as a functionally distinct group (tTreg, thymus-derived Treg), or differentiate from naïve T cells in peripheral sites (pTreg, peripheral Treg) (Abbas et al., 2013; Sakaguchi et al., 2020; Shevach and Thornton, 2014). *In vitro*, at least in mouse, Treg cells can be induced from naïve CD4+ T cells (iTreg, induced Treg) by antigen activation in the presence of TGF-β1 (Kretschmer et al., 2005). Perhaps surprisingly, however, *in* vitro antigen activation in the presence of both TGF-β1 and IL-6 induces differentiation into a different cell type – T helper 17 (Th17) that can generate massive tissue inflammation (Korn et al., 2009). Th17 cells are thought to take part in host defense against certain pathogens, but also studied because they are implicated in severe autoimmune disease, including, for example, multiple sclerosis, psoriasis, and lupus.

The shared lineage differentiation factor hints at an intimate connection between the Th17 and Treg types. Indeed, it seems that they are reciprocally regulated, with inhibition of one fate leading to promotion of the other, which gave rise to the paradigm of a Th17/Treg balance (Eisenstein and Williams, 2009; Wu et al., 2018).

More recently, it was discovered by us and others that substantial heterogeneity of effector profiles exists also within the Th17 cell type (Gaublomme et al., 2015; Stockinger and Omenetti, 2017; Wang et al., 2015). Th17 cells. Th17 do not exclusively pro- inflammation, but rather can also protect mucosal tissues by promoting homeostasis, maintaining barrier function, and restraining opportunistic microbiota (Bettelli et al., 2008; Conti et al., 2014; Gaffen et al., 2011; Guglani and Khader, 2010; Korn et al., 2009; Ouyang et al., 2008; Romani, 2011; Yang et al., 2014). The Janus-faced nature of Th17 cells, manifest also in their close relation to the hallmark suppressive Treg type, motivates a single-cell approach to studying them.

1.4 **Putting it all together**

1.4.1 Motivations for the present study

Thanks to technological breakthroughs of the past decade, single-cell transcriptomics is now cost-effective and readily accessible. Immense datasets are already available in the public domain, and ambitious efforts are underway to sequence all human cell types (Regev et al., 2017a). At the same time, extant assays to directly query cellular metabolic states are currently limited in their comprehensiveness, availability, or ability to study individual cells. This, together with the growing interest in cellular metabolism and the benefits afforded by computational analysis in the study of cellular metabolism, motivated us to develop an *in silico* method to predict cellular metabolic profiles based on single-cell transcriptomics.

FBA methods proved particularly useful when contextualized with functional genomic data (Bordbar et al., 2014), usually gene expression of bulk cell populations. The advent of single-cell RNA-Seq (scRNA-Seq) now offers an opportunity to harness FBA to study the metabolic heterogeneity of single cells. Importantly, single-cells form a natural perturbation system, where every cell can be considered a realization of a sampling process from a high-dimensional distribution. The statistical power afforded by the number of observations (cells) in current datasets allows for extensions of FBA that were previously out of reach.

Metabolism is an established regulator of the Th17 and Treg cells, as well their balance with respect to one another (Barbi et al., 2013; Berod et al., 2014; Gerriets et al., 2015; Johnson et al., 2018; Michalek et al., 2011; Pompura et al., 2021; Wang et al., 2015; Watson et al., 2021). We hypothesized that metabolism similarly regulated the effector functions of Th17 subtypes. However, most cellular assays, including metabolic assays, are targeted and difficult to undertake at a single-cell resolution. Furthermore, low cell numbers may prohibit direct metabolic assays, for example, in the study of immune cells present at tissue sites. In contrast, scRNA-Seq is broadly accessible and rapidly collected in concentrated efforts to reach a complete representation of human physiology (Regev et al., 2017a). A computational method is thus required to capitalize on the opportunities afforded by scRNA-Seq for contextualization of metabolic models, while systematically addressing the unique challenges of this data modality (e.g., its sparsity).

1.4.2 Our contribution

Here, we present Compass, an FBA algorithm that uses single-cell transcriptomic profiles to characterize cellular metabolic states at single-cell resolution and with network-wide comprehensiveness. The combination of single-cell granularity and genome-scale (i.e., nontargeted) molecular profiling is presently a singular advantage of scRNA-Seq over other metabolic data modalities, such as metabolomics or proteomics. By leveraging scRNA-Seq, Compass allows an unsupervised, network-wide search for metabolic targets, agnostically of *a priori* pathway boundaries, and including ancillary pathways that are normally less studied, yet important for cellular function (Puleston et al., 2017).

Compass is a step towards two important scientific goals – the integration of massive functional genomic data with molecular-level modeling, and turning human cell atlases from a phenomenological into an actionable resource that can be computationally mined for novel therapeutic targets. FBA algorithms build on manual and semi-automatic curation of metabolic models to make predict the way cells utilize their metabolic networks and configure fluxes through it to meet their function needs. Previous studies used gene expression to inform FBA prediction, but were hampered by the limited number of observations (samples) in a typical dataset. The introduction of single-cell transcriptomics, in which cells are observations in a natural perturbation system (alternatively, realizations of a low-dimensional latent generative process, (Lopez et al., 2018)) allows Compass to utilize FBA without an assumed objective function (Damiani et al., 2019) that might be inappropriate in many cases of mammalian cells (Wagner et al., 2013). Thus, Compass is a step towards bridging massive, atlas-scale transcriptomics, and modeling of molecular processes at a single biochemical level. Moreover, the use of FBA also allows conducting large-scale *in silico* perturbation experiments (Alper et al., 2005; Burgard et al., 2003; Lee et al., 2007; Park et al., 2007; Pharkya et al., 2003, 2004; Ranganathan et al., 2010; Yizhak et al., 2013) to detect metabolic targets that shift a give cell state towards a desired cell state, for example from a disease to a healthy molecular profile (Wagner et al., 2015). Thus, Compass is a step towards an actionable human cell atlas in the realm of cellular metabolism.

We applied Compass to Th17 cells, uncovering substantial immunometabolic diversity associated with their inflammatory effector functions. In addition to differential wiring of central carbon metabolism, we unexpectedly found – based on computational predictions – that the glycolytic reaction catalyzed by phosphoglycerate mutase (PGAM) wan promoting an anti-inflammatory phenotype in Th17 cells. This was a surprising finding since high glycolytic activity is commonly thought to be associated with pro-inflammatory T effector functions (Geltink et al., 2018; MacIver et al., 2013; O'Neill et al., 2016; Pearce et al., 2013; van der Windt and Pearce, 2012). A network-wide Compass analysis identified the polyamine pathway as a critical regulator of Th17 effector function, which we experimentally validated. This demonstrates that Compass is able to uncover novel metabolic targets regulating inflammation in peripheral, understudied metabolic pathways.

Chapter 2

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

The widespread adoption of single-cell transcriptomics (Svensson et al., 2018), combined with its ability to comprehensively profile cellular RNA pools (in contrast to the more targeted nature of extant metabolomic assays) motivated us to develop a computational framework for inference of metabolic profiles based on single-cell RNA-Seq. In this chapter, we offer a detailed presentation of the Compass algorithm, which underlies the rest of this study.

Metabolic dysregulation is now recognized as a key element in many disease conditions that severely affect people worldwide, including cancer, neurodegeneration, cardiovascular disease and even normal aging. We therefore believe that development of computational paradigms for studying metabolism, in our case based on single-cell transcriptomes, is a timely and much needed contribution.

2.1 Definitions

2.1.1 Stoichiometric matrix (S)

A stoichiometric matrix S describes the set of possible metabolic reactions in the system. Its rows correspond to metabolites, columns to reactions, and entries hold the stoichiometric coefficients of the reactions available to the cells. Observe that by transforming the entries into {-1, 0, 1} according to their sign, we obtain the hypergraph representation of the metabolic network.

2.1.2 Metabolic flux (v)

A metabolic flux is the instantaneous rate in which a chemical reaction occurs, and measured in units of mol X gDW^{-1} X hr^{-1}, with gDW denoting gram dry weight. Let x be metabolite concentrations in the system (as a function of time t), and v be the metabolic fluxes (also called *flux distribution*). Then:

$$S \cdot v = \frac{dx}{dt}$$

Compass belongs to the family of static FBA algorithms (Orth et al., 2010), which assumes metabolic steady-state, i.e.

$$S \cdot v = 0.$$

Equivalently, we limit the space of feasible flux distributions to $\ker(S)$.

2.1.3 Genome-scale metabolic model (GSMM)

Compass leverages prior knowledge about the metabolic network of the cells in question in the form of a Genome-Scale Metabolic Model (GSMM), which includes the following components (Heirendt et al., 2019; Monk et al., 2014; O'Brien et al., 2015; Palsson, 2015):

- A stoichiometric matrix *S*
    - Reactions are partitioned into cellular *compartments* representing membrane-enclosed space in which a biochemical reaction may take place, for example the cytosol, golgi, or mitochondria.
    - Metabolites are represented separately for each compartment. For example, cytosolic citrate and mitochondrial citrate correspond to two rows in *S*.
    - Reactions may contain non-zero coefficients for metabolites located in different compartment to represent physiological transportation of metabolites across membranes. These reactions are called *Transport reactions*.
    - The extracellular space is represented as an additional compartment.
    - Reactions may have exactly one non-zero coefficient if the corresponding metabolite is located in the extracellular space. These reactions are called *exchange reactions* and used to import metabolites into or export them from the system. It is standard convention to assign -1 as the non-zero coefficient of exchange reactions.
- Upper and lower bounds on fluxes through the reactions corresponding to columns of *S*.
    - Lower bounds of exchange reactions limit metabolite uptake and used to simulate constraints on nutrient availability. Note these are the lower, and not upper, bounds because the exchange coefficient is -1.
- A set if genes coding enzymes that catalyze reactions in the network
- Gene-to-reaction associations:
    - Every reaction is associated with a boolean expression over gene literals and the $\{\vee, \wedge\}$ operators.
    - Every gene is assigned a truth value based on its presence or absence in the cell's genome. A gene may be absent, for example, in a knockout genotype.
    - The boolean expressions code the dependency of reactions on their catalyzing proteins. Usually, an OR relationship corresponds to isozymes, namely two

enzymes capable of catalyzing the same reaction, and AND relationship corresponds to enzyme complexes.

2.2 **Algorithm overview**

We reasoned that even though the mRNA expression of enzymes is not an accurate proxy for their metabolic activity, a global analysis of the 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 states. This led us to develop the Compass algorithm, which integrates scRNA-Seq profiles with prior knowledge of the metabolic network to infer metabolic states of cells **(Figure 2A)**.

The metabolic network is encoded in a Genome-Scale Metabolic Model (GSMM) that includes reaction stoichiometry, biochemical constraints such as reaction irreversibility, nutrient availability, and gene-enzyme-reaction associations. Here, we use Recon2, which comprises of 7,440 reactions and 2,626 unique metabolites (Thiele et al., 2013). 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. 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 (the rate by which chemical reactions convert substrates to products), through constrained-based optimization (Lewis et al., 2012; O'Brien et al., 2015; Orth et al., 2010; Palsson, 2015). The first step of Compass is agnostic to gene expression and computes, for every metabolic reaction *r*, the maximal flux v_r^{opt} it can carry while imposing only stoichiometry and mass balance constraints. Next, Compass assigns every reaction in every cell a penalty inversely proportional to the mRNA expression associated with the enzyme(s) catalyzing the reaction in that cell. Finally, for every reaction *r* and every cell Compass finds a flux distribution (an assignment of flux values to every reaction in the network) that minimizes the overall penalty incurred, while maintaining a flux of at least $\omega \cdot v_r^{opt}$ (here, $\omega = 0.95$) through *r*. The additive inverse of this penalty term is the reaction score.

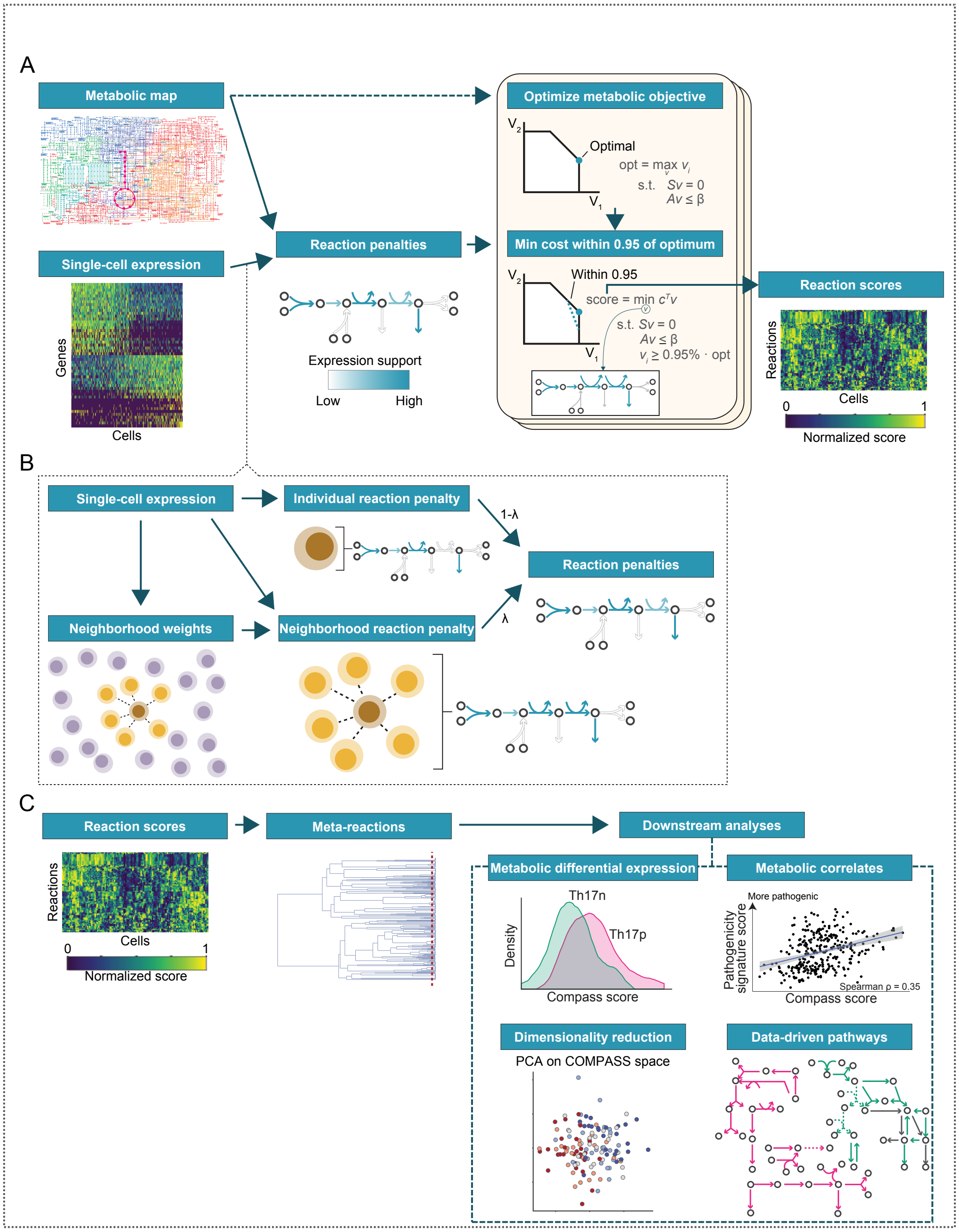

A
Metabolic map
Single-cell expression
Genes
Cells
Reaction penalties
Expression support
Low
High
Optimize metabolic objective
V_2
V_1
Optimal
$\text{opt} = \max_v v_i$
s.t. $Sv = 0$
$Av \le \beta$
Min cost within 0.95 of optimum
Within 0.95
$\text{score} = \min c^T v$
s.t. $Sv = 0$
$Av \le \beta$
$v_i \ge 0.95\% \cdot \text{opt}$
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. Algorithm overview.

(A) Computation of Compass scores matrix. Compass leverages the topology and stoichiometry of the metabolic network to analyze single-cell RNA expression. Briefly, it computes a reaction-penalties matrix, where the penalty of a given reaction is inversely proportional to the expression of 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)** Soft information sharing between a cell and its k-nearest neighbors mitigates technical noise in single-cell data. **(C)** Downstream analysis of the score matrix. Metabolic reactions are hierarchically clustered into meta-reactions; scores are used in differential expression of reactions, detection of reactions correlated with a phenotype of interest, dimensionality reduction, and data-driven network analysis.

The use of genome-scale metabolic networks 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 with metabolic activity and of scRNA-Seq dropouts (Wagner et al., 2016). Compass further mitigates data sparsity effects through information-sharing on a *k*-nearest neighbors graph, similar to other scRNA-Seq algorithms (Baran et al., 2019; van Dijk et al., 2018; Grün, 2019; Haghverdi et al., 2018; Huang et al., 2018; Lun et al., 2016) (**Figure 2B**).

The output of Compass is a quantitative profile for the metabolic state of every cell, which is then subject to downstream analyses **(Figure 2C)**. 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.

2.3 **Notation**

In the following sections we denote:

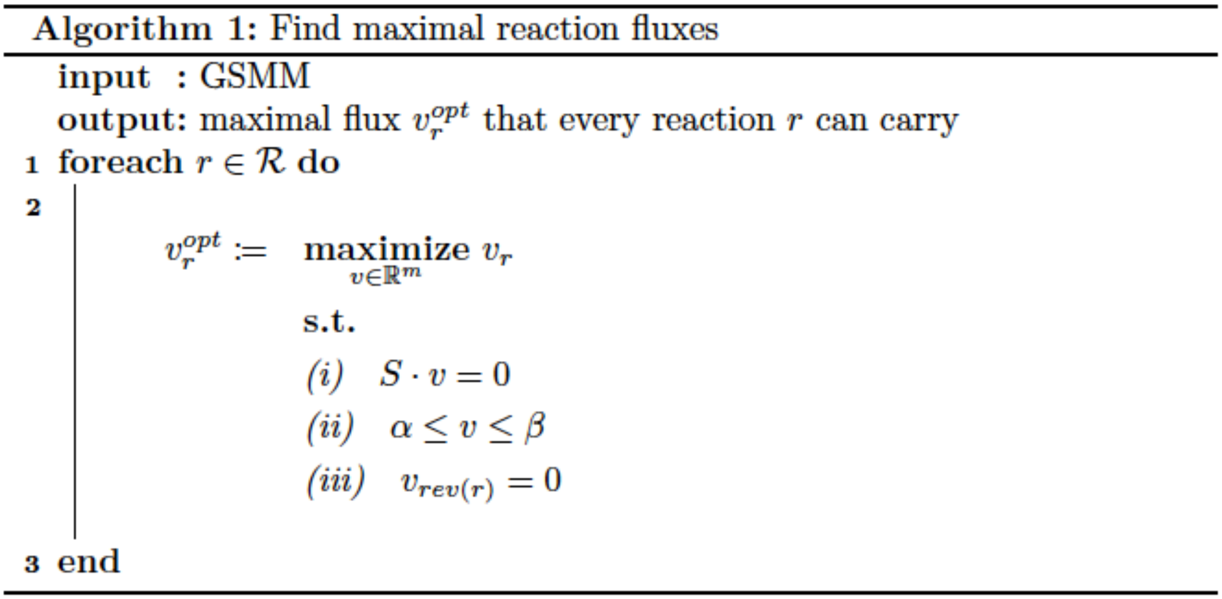
- n: number of cells (or RNA libraries).
- m: number of metabolic reactions in the GSMM.
- C: the set of cells in the data. ($|C| = n$).
- R: the set of metabolic reactions in the GSMM. ($|R| = m$).
- $rev(r)$: the reverse unidirectional reaction of reaction r, which has the same stoichiometry but proceeds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

- *S*: the stoichiometric matrix defined in the GSMM, where rows represent metabolites, columns represent reactions, and entries are stoichiometric 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ we use $f(M)$ to denote (where the intention is obvious from the context) the respective point-wise transformation, namely $f(M) := (f(m_{i,j}))$.

2.4 **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
      $v_r^{opt} := \underset{v \in \mathbb{R}^m}{\text{maximize}}\ v_r$
                    s.t.
                    (i)   $S \cdot v = 0$
                    (ii)  $\alpha \leq v \leq \beta$
                    (iii) $v_{rev(r)} = 0$
  end
```

Constraint *(i)* constrains the system to steady state (Varma and Palsson, 1994). Constraint *(ii)* is interpreted as $\forall i$: $\alpha_i \leq v_i \leq \beta_i$ 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 et al., 2011).

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

2.5 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 (Becker and Palsson, 2008; Shlomi et al., 2008) 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.

2.6 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

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

2.7 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 $\forall i.\ v_i > 0$.

```
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
2.8)
Compute with the procedures described in sections 2.6, 2.7:
    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 $\hat{P}^{(c)} = \left(\hat{P}_{1,c}, \ldots, \hat{P}_{m,c}\right)$
        $y_{r,c} := \underset{v\in\mathbb{R}^m}{\textbf{minimize}}\ \hat{P}^{(c)} \cdot v$
                    s.t.
                    (i)   $S \cdot v = 0$
                    (ii)  $\alpha \leq v \leq \beta$
                    (iii) $v_{rev(r)} = 0$
                    (iv)  $v_r \geq \omega \cdot v_r^{opt}$
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 2.8
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 $C$
```

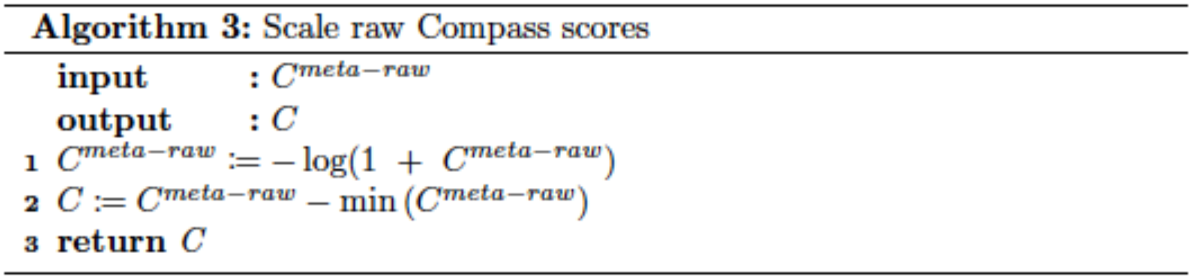
2.8 **Meta-reactions**

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 – ρ, with ρ being the Spearman correlation) by averaging the respective rows. In the present work, we used ρ = 0.98. We denote the row-merged matrix $\{C^{meta-raw}\}_{mXn}$.

2.9 **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 3: Scale raw Compass scores
   input    : $C^{meta-raw}$
   output   : $C$
1  $C^{meta-raw} := -\log(1 + C^{meta-raw})$
2  $C := C^{meta-raw} - \min(C^{meta-raw})$
3  return $C$
```

2.10 **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.

2.11 **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* (DeTomaso et al., 2019) implemented in the `VISION` R package (`https://github.com/YosefLab/VISION`), and meta-cells (Baran et al., 2019) (`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 1.3) non-core reactions (Section 1.11.2), 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.

2.12 **Metabolic model considerations**

2.12.1 Metabolic network

We used the Recon2 GSMM (Thiele et al., 2013) 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.

2.12.2 *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.

2.13 **Gene expression considerations**

2.13.1 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.

2.13.2 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 andignores other transcripts.

2.13.3 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.

2.14 **Code and data availability**

2.14.1 Software

Compass is available at `https://github.com/YosefLab/Compass`.

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 freefor academic use.

2.14.2 Sequencing data

Sequencing data used in this study has been deposited to GEO under accession GSE164999. The data published in (Gaublomme et al., 2015; Wang et al., 2015) and reanalyzed here was deposited as GSE74833.

2.15 **Limitations of the study**

Compass is subject to the limitations of static FBA. The Recon2 metabolic network is incomplete (e.g., lacks annotation of enzyme moonlighting functions), and pertains to a generalized human cell. Consequently, the current version of Compass does not consider the differences between human and mouse metabolism, or tissue-specificity of the metabolic network. The algorithm makes the simplifying assumption of metabolic steady-state, and heuristically aggregates expression of multiple genes that are linked to a reaction. The inference of metabolic programs based on transcriptomes does not consider post-transcriptional and post-translational regulation, which could be particularly important for metabolic adaptations on short timescales. Last, 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. A more accurate representation of the cellular environment should increase the Compass's predictive capabilities.

Chapter 3

Querying the metabolic states of single T helper 17 (Th17) cells

In this chapter, we introduce the T helper 17 (Th17) and the closely related T regulatory (Treg) cells as a model system for heterogenous effector profiles within populations of immune cells. We present an *in vitro* Th17 system that allows controlled inquiry into the molecular underpinnings of their functional heterogeneity, and leverage Compass to conduct a network-wide search for metabolic targets associated with Th17 inflammatory potential (*pathogenicity)*.

3.1 Th17 cells are a model system for effector-metabolic heterogeneity

3.1.1 Th17 manifest heterogeneous effector profiles

We and others have studied transcriptional and functional diversity among T helper 17 cells (Gaublomme et al., 2015) and Foxp3$^+$ regulatory T cells (Tregs) (Miragaia et al., 2019). Th17 and Treg cells share lineage signals, and their balance shapes the outcome of tissue inflammation (Eisenstein and Williams, 2009; Omenetti and Pizarro, 2015). While TGF-β1 alone induces Foxp3$^+$ Tregs *in vitro*, a combination of TGF-β1 and IL-6 induce Th17 cells *in vitro* and *in vivo* (Bettelli et al., 2006; Mangan et al., 2006; McGeachy and Cua, 2008; Veldhoen et al., 2006).

On top of the Th17/Treg balance, there exists functional diversity within the Th17 compartment. Th17 cells may induce disease, but they also protect mucosal tissues by promoting tissue homeostasis, maintaining barrier function, and restraining opportunistic microbiota (Bettelli et al., 2008; Conti et al., 2014; Gaffen et al., 2011; Guglani and Khader, 2010; Korn et al., 2009; Ouyang et al., 2008; Romani, 2011; Yang et al., 2014). Murine Th17 cells induced by TGF-β1+IL-6 *in vitro* produce IL-17 but are incapable of inducing potent tissue inflammation upon adoptive transfer (Jager et al., 2009; McGeachy et al., 2007) without additional stimuli, such as IL-1b and IL-23 (Awasthi et al., 2009; Chung et al., 2009; Cua et al., 2003; Ghoreschi et al., 2010; Lee et al., 2012; McGeachy et al., 2009). Therefore, there appear to be at least two different types of Th17 cells: homeostatic ones that do not promote tissue inflammation, to which we refer as non-pathogenic Th17 cells; and ones that produce IL-17 together with IFNg and GM-CSF and induce tissue inflammation/ autoimmunity, to which we refer as pathogenic Th17 cells (Lee et al., 2014). Distinct types of Th17 cells have also been identified in humans, where cells similar to mouse pathogenic Th17 cells are specific for immune response to *Candida albicans,* while cells similar to the non-pathogenic mouse Th17 are observed with *Staphylococcus aureus* infection (Zielinski et al., 2012). Thus, Treg, non-pathogenic Th17, and pathogenic Th17 cells represent a functional spectrum in tissue homeostasis, infection and tissue inflammation *in vivo* and can be differentiated with different cytokine cocktails *in vitro*.

3.1.2 Motivation for studying Th17 metabolic heterogeneity

Metabolism is an established regulator of the Th17/Treg axis (Barbi et al., 2013). We hypothesized that metabolism similarly regulated the effector functions of Th17 subtypes. However, most cellular assays, including metabolic assays, are targeted and difficult to undertake at a single-cell resolution. Furthermore, low cell numbers may prohibit direct metabolic assays, for example, in the study of immune cells present at tissue sites. In contrast, scRNA-Seq is broadly accessible and rapidly collected in concentrated efforts to reach a complete representation of human physiology (Regev et al., 2017a). A computational method is thus required to capitalize on the opportunities afforded by scRNA-Seq for contextualization of metabolic models, while systematically addressing the unique challenges of this data modality (e.g., its sparsity).

3.2 Th17 functional states are associated with metabolic states

We used Compass to study the metabolism of Th17 cells differentiated *in vitro* into two extreme functional states **(Figure 3A)** – pathogenic (Th17p) and non-pathogenic (Th17n). We analyzed a dataset we generated in a previous study that included 139 Th17p and 151 Th17n cells sorted for IL-17A/GFP+ (Gaublomme et al., 2015; Wang et al., 2015). We computed Compass scores and aggregated reactions that were highly correlated across the entire dataset (Spearman rho ≥ 0.98) into meta-reactions (median of two reactions per meta-reaction; **Figure S2A)** for downstream analysis. We tested the robustness of Compass by running the algorithm on the same gene expression input with added random noise. The deviation of the noised Compass scores from the original output did not exceed the deviation induced by the noise to the gene expression input **(Figures S2B-C)**.

To investigate the main determinants of metabolic heterogeneity between Th17 cells, 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 genes. We performed principal component analysis (PCA) on the meta-reaction matrix, while restricting it to 784 meta-reactions (out of 1,911) that are associated with core metabolism **(Supplementary Methods)**, spanning conserved and well-studied pathways for generation of ATP and synthesis of key biomolecules.

The first two principal components (PCs) were associated with overall metabolic activity and T effector functions **(Figures 3B; S3A-B)**. 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, $p < 4e-10$), as well as a transcriptional signature of late stages of Th17 differentiation over time (Yosef et al., 2013) (**Figure S3C; Supplementary Methods;** Pearson rho = 0.18, $p < 0.003$) **(Supplementary Methods)**. PC2 and PC3 represented a choice between ATP generation through aerobic glycolysis versus fatty acid oxidation, similar to previous observations in comparisons of Th17 to Tregs, or Teff to Tmem (Geltink et al., 2018). Accordingly, these PCs correlated with multiple Th17 pathogenicity markers, as well as a signature of Th17 pathogenicity consisting of cytokines, chemokines and transcription factors (TFs) that are associated with each phenotypic group (Lee et al., 2012) **(Figure S3D,E)**. PC2 and PC3 were also associated with nitrogen metabolism, and were enriched in urea cycle targets whose power to modulate Th17 pathogenicity is demonstrated below.

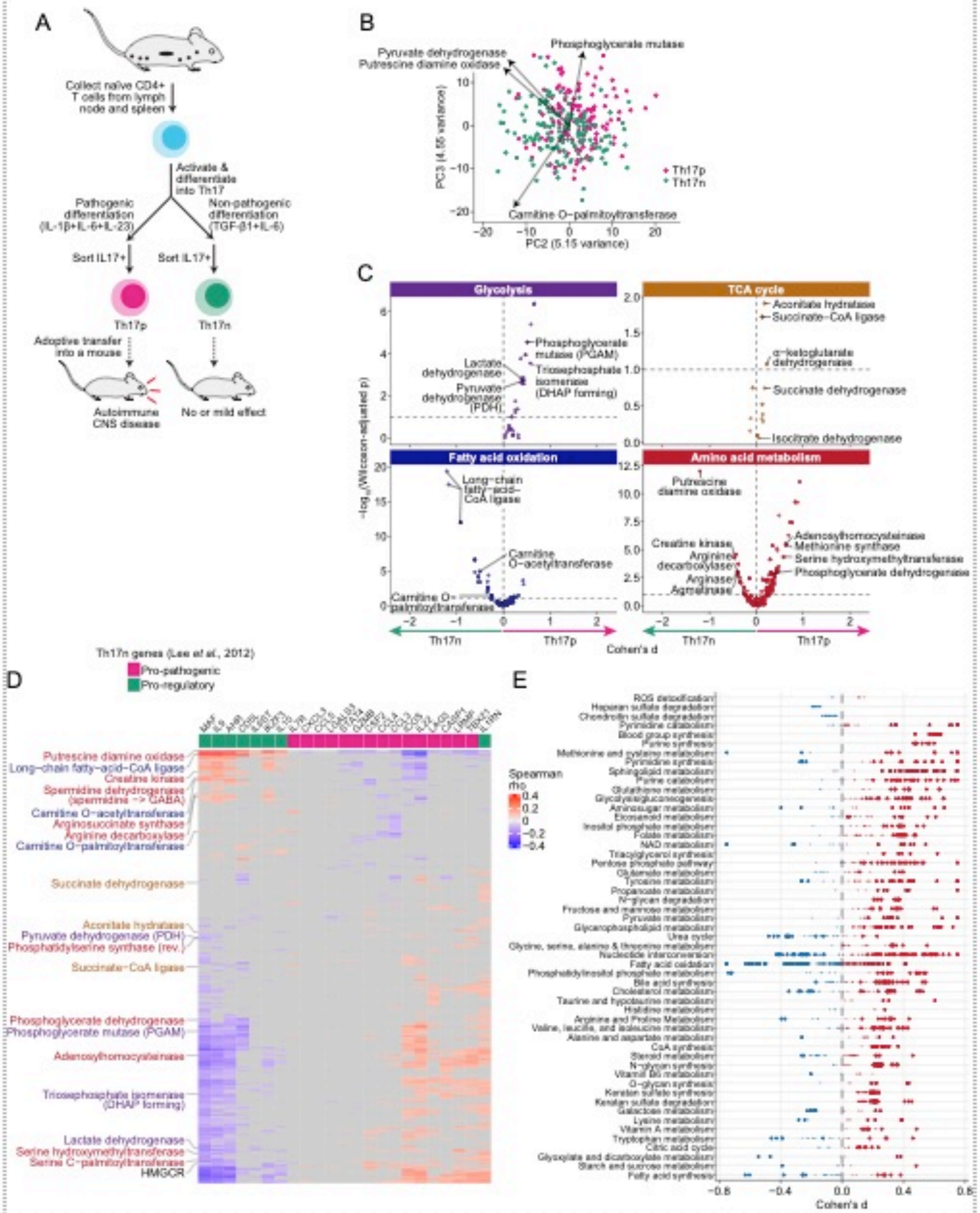

A
Collect naive CD4+ T cells from lymph node and spleen
Activate & differentiate into Th17
Pathogenic differentiation (IL-1β+IL-6+IL-23)
Non-pathogenic differentiation (TGF-β1+IL-6)
Sort IL17+
Sort IL17+
Th17p
Th17n
Adoptive transfer into a mouse
Autoimmune CNS disease
No or mild effect
B
Pyruvate dehydrogenase
Putrescine diamine oxidase
Phosphoglycerate mutase
Carnitine O−palmitoyltransferase
PC3 (4.55 variance)
PC2 (5.15 variance)
Th17p
Th17n
C
Glycolysis
Phosphoglycerate mutase (PGAM)
Lactate dehydrogenase
Triosephosphate isomerase (DHAP forming)
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
Creatine kinase
Arginine decarboxylase
Arginase
Agmatinase
Adenosylhomocysteinase
Methionine synthase
Serine hydroxymethyltransferase
Phosphoglycerate dehydrogenase
−log10(Wilcoxon-adjusted p)
Th17n
Th17p
Cohen's d
D
Th17n genes (Lee et al., 2012)
Pro-pathogenic
Pro-regulatory
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 & 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. Compass-based exploration of metabolic heterogeneity in Th17 cells.

(A) The experimental system. Naive $CD4^+$ T cells differentiated into pathogenic (Th17p) or non-pathogenic (Th17n) IL-17^+ T cells cause severe or mild CNS autoimmunity upon adoptive transfer, respectively. **(B)** PCA of the Compass score matrix with top loadings shown. **(C)** Compass-score differential activity test between Th17p and Th17n cells (**STAR Methods**). **(D)** Spearman correlation of Compass scores with the expression of pro-pathogenic (magenta) or pro-regulatory (green) Th17n genes (Lee et al., 2012). Non-significant correlations (BH-adjusted $p \geq 0.1$) shown in grey. Rows are 489 meta-reactions that belong to core pathways (**STAR Methods**), and significantly correlated or anti-correlated with at least one of the genes. **(E)** Differential activity (as in C) of metabolic reactions. Reactions (dots) are partitioned by Recon2 pathways and colored by the sign of their Cohen d's statistic.

3.3 Compass predicts metabolic regulators of Th17 cell pathogenicity

To detect metabolic targets associated with the pathogenic capacity of individual Th17 cells, we defined pro-pathogenic and pro-regulatory reactions as ones that were significantly differentially active in Th17p or Th17n, respectively (**Figures 3C; S3F;** 1,213 / 6,563 reactions, Benjamini-Hochberg (BH) adjusted Wilcoxon rank sum $p < 0.001$). Many of these reactions were also correlated with the expression of cytokines and TFs relevant for Th17 function **(Figures 3D; S3G;** note that these genes do not code metabolic enzymes and thus were not used by Compass). Notably, many classically defined metabolic pathways included both reactions predicted to be pro-pathogenic and pro-regulatory **(Figure 3E)**, 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 **(Figure S3H)**.

Compass highlighted differences in central carbon metabolism between the Th17p and Th17n states, which mirror those found between Th17 and Treg. The algorithm predicted that glycolytic reactions were generally more active in Th17p than in Th17n **(Figures 3C, 4A)**. This parallels previous results showing that Th17 upregulate glycolysis, and failure to do so promotes a Treg fate (Gerriets et al., 2015; Michalek et al., 2011; Shi et al., 2011). Compass also predicted an increased activity in Th17p through two segments of the TCA cycle, but not the cycle as a whole **(Figures 3C, 4A)**. A similar breakdown of the TCA cycle in relation with pro-inflammatory function has been described in macrophages where M1 polarization divided the TCA cycle at the same two points: isocitrate dehydrogenase (Jha et al., 2015), and succinate dehydrogenase (Mills et al., 2016), which supported macrophage inflammatory functions (Mills and O'Neill, 2014; Shi et al., 2019).

In fatty acid (FA) metabolism, Compass predicted that cytosolic acetyl-CoA carboxylase (ACC1), the committed step towards FA synthesis, was upregulated in Th17p, whereas the first two steps of long-chain FA oxidation (long chain fatty acyl-CoA synthetase and carnitine O-palmitoyltransferase (CPT)) were predicted to be higher in Th17n. These predictions mirror a known metabolic difference between the Th17 and Treg lineages, where Th17 rely on *de novo* FA synthesis (Berod et al., 2014), whereas Tregs scavenge them from their environment and catabolize them and produce ATP through beta-oxidation (Michalek et al., 2011; Pompura et al., 2021). We note, however, that recent evidence suggests that CPT may be upregulated in Treg over Th17, but is not functionally indispensable for Treg to obtain their effector phenotypes (Raud et al., 2018).

Multiple amino-acid metabolism reactions were also differentially active between Th17p and Th17n **(Figure 3C)**. Amino acids are important for Th17 cell differentiation (Sundrud et al., 2009), and Compass adds further granularity to these findings. In particular, it predicted that serine biosynthesis from 3-phosphoglycerate, and 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 Th17 states, **(Figure 3C)**, suggesting that alternative fluxing within this subsystem may be associated with diverging Th17 cell function.

Chapter 4

Computationally-informed analysis of central carbon metabolism

In this chapter, we validate some of the predictions Compass made with respect to differences between the metabolic programs of pathogenic Th17 (Th17p) and non-pathogenic Th17 (Th17n) cells. We show that Th17p preferentially produce ATP by elevating glycolysis and shunting its products into the TCA for oxidative phosphorylation, whereas Th17n cells oxidize fatty acids for that purpose. Next, Compass recovers a specific glycolytic enzyme (PGAM) which is associated with a pro-regulatory Th17 effector profile, contrary to what could be expected according to standard understanding of lymphocyte metabolism. These findings demonstrate that Compass is able to advance understanding even of central and well-studied metabolic pathways.

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

We validated the Compass prediction that Th17p and Th17n differ in their central carbon metabolism (**Figure 4A**). First, we assayed glycolysis and mitochondrial function of Th17 cells (**Figure 4B**). Th17p had higher extracellular acidification rate (ECAR) than Th17n, indicating accumulation of lactate due to aerobic glycolysis. Th17p also generated more ATP in a mitochondria dependent fashion, consistent with the predicted higher entrance of pyruvate into the TCA cycle despite the diversion of some pyruvate towards the lactate fate.

LC/MS metabolomics indicated that glycolytic metabolites were higher in Th17p than in Th17n (**Figure 4C, top**). When further pulsed with fresh media containing glucose for 15 minutes, there was a substantial increase in glycolytic metabolites in Th17p but less so in Th17n. After 3 hours, the level of these metabolites decreased back to steady state (**Figure S4A**). TCA metabolites, apart from succinate, were more abundant at steady-state in Th17p than in Th17n **(Figure 4C, middle)**, consistent with Compass prediction that two parts of the TCA cycle, but not the cycle as a whole were upregulated in Th17p. 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 alternative fates differ between Th17p and Th17n, we performed carbon tracing with 13C-glucose. Consistent with our predictions, Th17p had higher relative abundance of 13C-labeled glycolytic metabolites (**Figure 4D**). Furthermore, Th17p preferentially incorporated glucose-derived carbon into serine (which branches from glycolysis; **Figure S4B**) and its downstream product choline (**Figure 4D**), confirming a Compass prediction (**Figure 3C**). Conversely, Th17p had lower relative abundance of 13C-labeled TCA metabolites (**Figure 4D**), suggesting that the higher level of TCA intermediates observed in Th17p at steady state (**Figure 4C**) might not be supported from glucose, but rather from other sources (Johnson et al., 2018; Pucino et al., 2019). Taken together, our results suggest that Th17p cells have higher overall activity through the TCA cycle at steady-state, but quickly switch to aerobic glycolysis when glucose is readily available in the environment.

We next validated the prediction that Th17n cells prefer beta-oxidation. Metabolomics analysis shows that Th17n were enriched in acyl-carnitine metabolites, particularly short- to medium-length acyl groups **(Figures 4C, S4A)**, indicating active lipid transport through the mitochondrial membrane. When etomoxir was used to block acyl-carnitine transportation, oxygen consumption rate decreased in Th17n but not Th17p (**Figure 4E**). While etomoxir has off-target effects (Divakaruni et al., 2018; Raud et al., 2018), overall our data supports the hypothesis that Th17n cells ultimately divert fatty acid breakdown products into the electron transport chain to generate ATP.

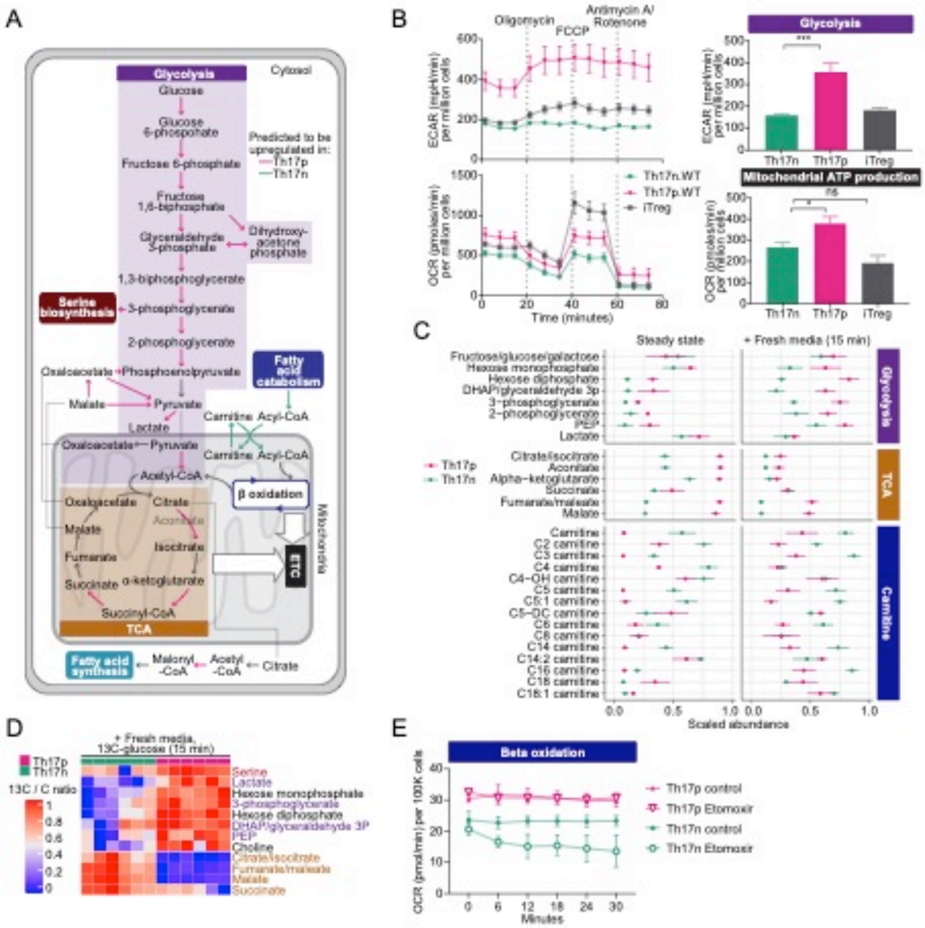
A
Glycolysis
Cytosol
Glucose
Glucose 6-phospohate
Predicted to be upregulated in:
Th17p
Th17n
Fructose 6-phosphate
Fructose 1,6-biphosphate
Glyceraldehyde 3-phosphate
Dihydroxy-acetone phosphate
1,3-biphosphoglycerate
Serine biosynthesis
3-phosphoglycerate
2-phosphoglycerate
Fatty acid catabolism
Oxaloacetate
Phosphoenolpyruvate
Malate
Pyruvate
Carnitine
Acyl-CoA
Lactate
Oxaloacetate
Pyruvate
Carnitine
Acyl-CoA
Acetyl-CoA
β oxidation
Oxaloacetate
Citrate
Aconitate
Malate
Isocitrate
Mitochondria
Fumarate
ETC
Succinate
α-ketoglutarate
Succinyl-CoA
TCA
Fatty acid synthesis
Malonyl-CoA
Acetyl-CoA
Citrate
B
Oligomycin
FCCP
Antimycin A/ Rotenone
ECAR (mpH/min) per million cells
600
400
200
0
OCR (pmoles/min) per million cells
1500
1000
500
0
0
20
40
60
80
Time (minutes)
Th17n.WT
Th17p.WT
iTreg
Glycolysis
***
ECAR (mpH/min) per million cells
500
400
300
200
100
0
Th17n
Th17p
iTreg
Mitochondrial ATP production
ns
*
OCR (pmoles/min) per million cells
500
400
300
200
100
0
Th17n
Th17p
iTreg
C
Steady state
+ Fresh media (15 min)
Fructose/glucose/galactose
Hexose monophosphate
Hexose diphosphate
DHAP/glyceraldehyde 3p
3-phosphoglycerate
2-phosphoglycerate
PEP
Lactate
Glycolysis
Th17p
Th17n
Citrate/isocitrate
Aconitate
Alpha-ketoglutarate
Succinate
Fumarate/maleate
Malate
TCA
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
Carnitine
0.0
0.5
1.0
0.0
0.5
1.0
Scaled abundance
D
+ Fresh media, 13C-glucose (15 min)
Th17p
Th17n
13C / C ratio
1
0.8
0.6
0.4
0.2
0
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
40
30
20
10
0
0
6
12
18
24
30
Minutes
Th17p control
Th17p Etomoxir
Th17n control
Th17n Etomoxir

Figure 4. Differential usage of glycolysis and fatty acid oxidation by pathogenic and non-pathogenic Th17 cells.

(A) Central carbon metabolism overlaid with Compass predictions for differentially active reactions; Th17p vs. Th17n, BH-adjusted Wilcoxon rank sum $p < 0.1$ denoted in non-grey. **(B)** Mitostress test by Seahorse assay of differentiated T cells (68h). **(C-D)** LC/MS metabolomics (n=6 mice; C shows means and SE error bars) of cells harvested at 68h (**C**, left), replated in fresh media with no additives (C, right) or with 13C-tagged glucose for 15 minutes (D). D shows the ratio of 13C-tagged carbon out of the total carbon content associated with the metabolite. **(E)** T cells were measured for their oxygen consumption rate measured in T cells with control or 40uM etomoxir (n=2, each with 6 mouse replicates).

4.2 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 (Gaublomme et al., 2015). 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.

We correlated the Compass scores associated with each reaction with the pathogenicity gene signature scores of the respective cells **(Figures 5A, S5A)**, and found that some glycolytic reactions were predicted to 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, but positively associated with Th17p cells in the inter-population analysis **(Figure 4A)**. This prediction was unexpected because increased glycolysis is generally thought to support pro-inflammatory phenotypes in T cells (Chang et al., 2013; Doedens et al., 2013; Finlay et al., 2012; Gemta et al., 2019; Gerriets et al., 2015; Ho et al., 2015; Kono et al., 2019; Macintyre et al., 2014; Michalek et al., 2011; Peng et al., 2016; Shi et al., 2011).

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 5B**). The inhibitors were shikonin (inhibits PKM2), dehydroepiandrosterone (DHEA, inhibits G6PD), epigallocatechin-3-gallate (EGCG, inhibits PGAM1), and 2,3-dihydroxypropyl-dichloroacetate (DCA, inhibits GK).

We first analyzed the effects of inhibitors on Th17n and Th17p cell differentiation and function using flow cytometry (**Figure 5C**). 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 5C)**. 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 **(Figure S5B).**

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 5D-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 5D**). 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 5E-F)**. A comparison of DHEA to vehicle control identified a large number of effector genes that are modulated. 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). 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 5F)**.

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. 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 5G)**. 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 3C**) and lies directly downstream of 3PG (**Figure 5B**), 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 5H**).

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.

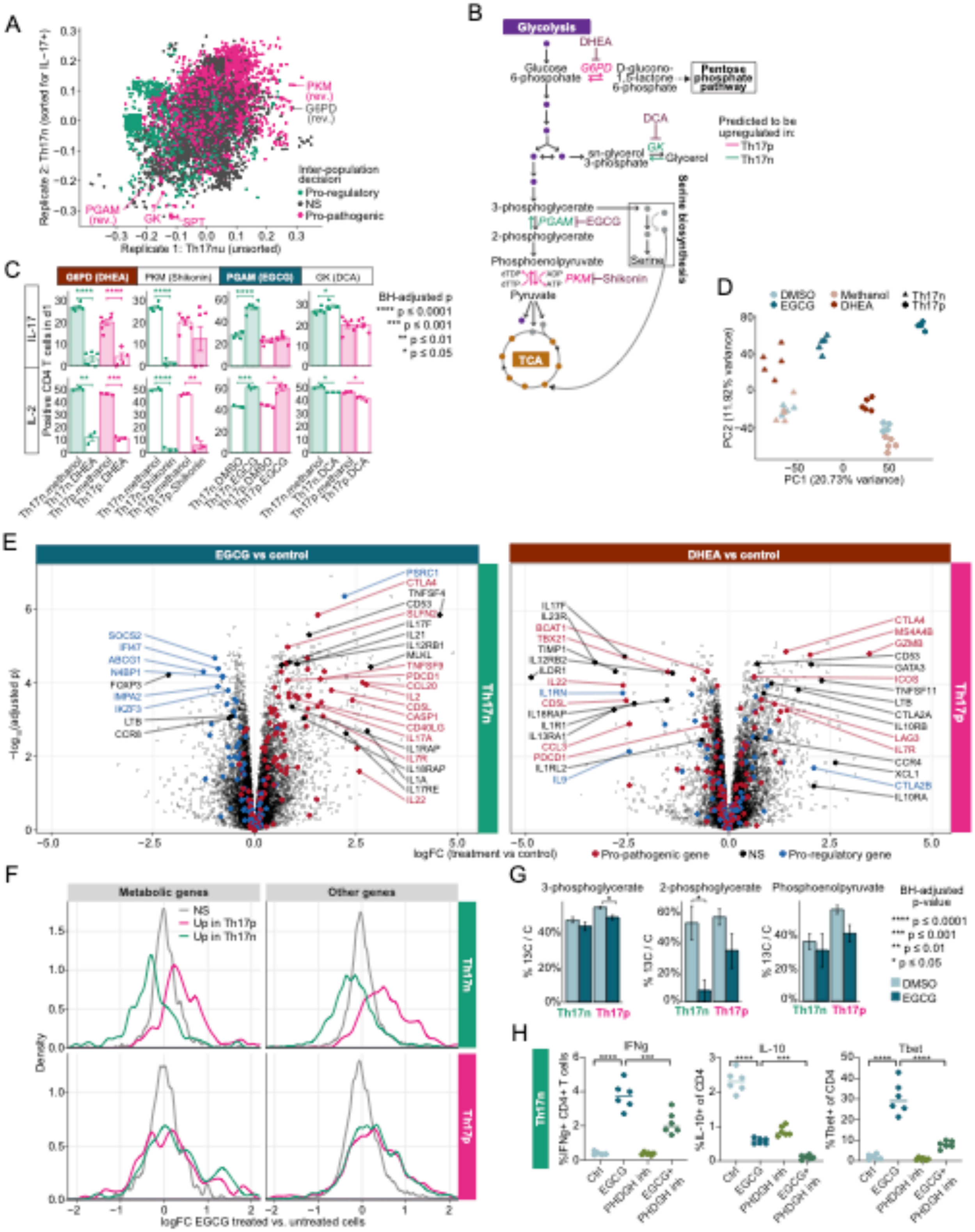

A
Replicate 2: Th17n (sorted for IL−17+)
Replicate 1: Th17nu (unsorted)
PKM (rev.)
G6PD (rev.)
PGAM (rev.)
GK
SPT
Inter-population decision
Pro-regulatory
NS
Pro-pathogenic
B
Glycolysis
DHEA
G6PD
Glucose 6-phosphate
D-glucono-1,5-lactone 6-phosphate
Pentose phosphate pathway
DCA
GK
sn-glycerol 3-phosphate
Glycerol
Predicted to be upregulated in:
Th17p
Th17n
3-phosphoglycerate
PGAM
EGCG
2-phosphoglycerate
Phosphoenolpyruvate
PKM
Shikonin
Pyruvate
Serine
Serine biosynthesis
TCA
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
PC1 (20.73% variance)
PC2 (11.82% variance)
E
EGCG vs control
Th17n
PSRC1
CTLA4
TNFSF4
CD53
SLFN5
IL17F
IL21
IL12RB1
MLKL
TNFSF8
PDCD1
CCL20
IL2
CD5L
CASP1
CD40LG
IL17A
IL1RAP
IL7R
IL18RAP
IL1A
IL17RE
IL22
SOCS2
IFI47
ABCG1
N4BP1
FOXP3
IMPA2
IKZF3
LTB
CCR8
−log10(adjusted p)
DHEA vs control
Th17p
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
logFC (treatment vs control)
Pro-pathogenic gene
NS
Pro-regulatory gene
F
Metabolic genes
Other genes
NS
Up in Th17p
Up in Th17n
Density
Th17n
Th17p
logFC EGCG treated vs. untreated cells
G
3-phosphoglycerate
2-phosphoglycerate
Phosphoenolpyruvate
% 13C / C
Th17n
Th17p
BH-adjusted p-value
**** p ≤ 0.0001
*** p ≤ 0.001
** p ≤ 0.01
* p ≤ 0.05
DMSO
EGCG
H
Th17n
IFNg
%IFNg+ CD4+ T cells
IL-10
%IL-10+ of CD4
Tbet
% Tbet+ of CD4
Ctrl
EGCG
PHDGH inh
EGCG+ PHDGH inh

Figure 5. An unexpected role for PGAM in mediating TGFb-induced Th17 pathogenicity.

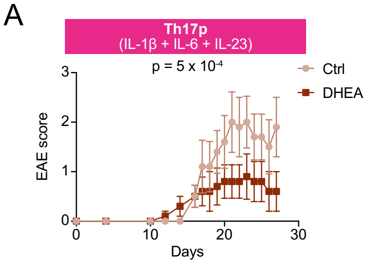
(A) Intra-population analysis in two biological replicates. Th17nu (Th17n unsorted) corresponds to the set of 130 T cells stimulated by the Th17n cytokines but unsorted for IL-17A-GFP+ in (Gaublomme et al., 2015) (the Th17n and Th17nu cell populations, see Figure 3A). Dots are single metabolic reactions, and axes denote Spearman correlation with the pathogenic signature. 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. SPT = serine-pyruvate transaminase (EC 2.6.1.51). Rev = reverse (backwards) direction. (STAR 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 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. **(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 (Supplementary Methods) or to the Th17 pro-inflammatory covariation module defined by (Gaublomme et al., 2015); blue genes are similarly defined). **(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), or the combination. Cells were harvested at 96h for flow cytometry.

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

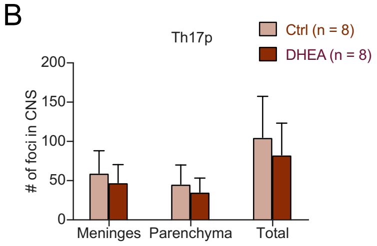
To test the functional relevance of the transcriptome shifts induced by EGCG and DHEA *in vivo*, we used the adoptive T cell transfer system, so that the effect of inhibitors is limited to T cells rather than all cells in the host. We generated Th17n and Th17p cells from naive CD4+ T cells isolated from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

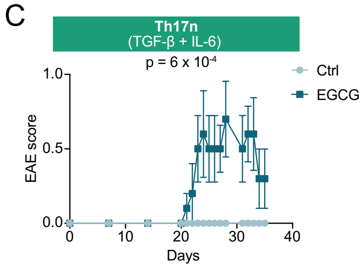
Consistent with Compass prediction, Th17p cells treated with DHEA reduced the severity of disease at peak of EAE in the recipients (**Figure 6A**). By the time the mice were sacrificed, however, the number of lesions in CNS was not significantly different (**Figure 6B**), and surviving mice showed no significant alterations in antigen-specific cytokine secretion in response to MOG, except for increased IL-2 in the DHEA treated group (**Figure S6A**). 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 6C)**. 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 6D**). While there was only a small difference in antigen-specific T cell proliferation (**Figure 6E**), there was a significant increase in secretion of IL-17, IL-17F, IL-22 and IL-6 (**Figure 6F** and **Figure S6B**) 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 (Aggarwal et al., 2003; Awasthi et al., 2009; McGeachy et al., 2009; Zhou et al., 2007). 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 6G**). 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 6H**). 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 6I, 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.



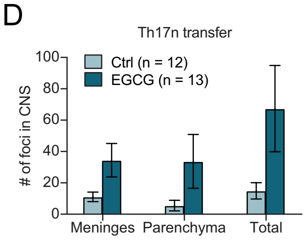
	Incidence	Peak disease	Onset (days)
Methanol	10 (of 10)	2.5+/-1.5	17+/-2
DHEA	7 (of 8)	1.3+/-0.8	17+/-4

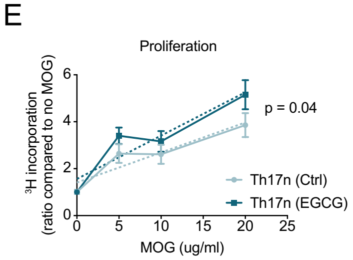


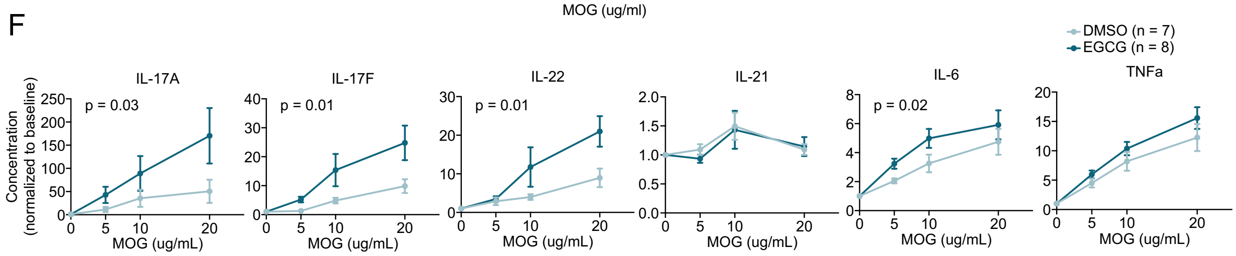


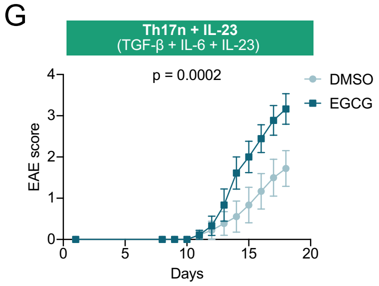
	Incidence	Peak disease	Onset (days)
DMSO	0 (of 12)	0	n/a
EGCG	10 (of 13)	1+/-0.6	25.1+/-4

Fisher's exact p = 1.1 x 10-4









	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	

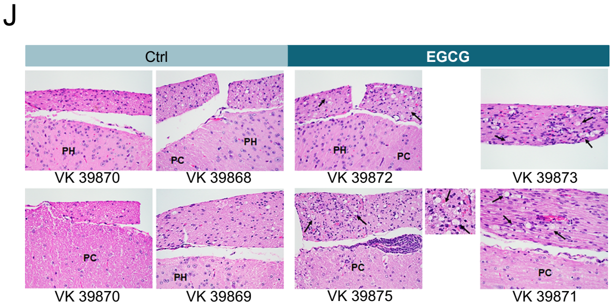


Figure 6. EGCG exacerbates and DHEA ameliorates Th17-induced EAE *in vivo*.

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.

Chapter 5

Computationally-informed analysis of peripheral metabolism

In this chapter, we turn to peripheral (also called ancillary, or secondary) metabolism – a loose term to capture parts of the metabolic network other than the main energy producing pathways, which can be crucial to function of immune cells (Puleston et al., 2017) as well as other organs. Since Compass leverages RNA-Seq inputs, it is able to make network-wide predictions in understudied parts of the metabolic network cut across traditional pathway boundaries. Indeed, this was one of our motivations to develop an *in silico* method that transforms transcriptomes into predicted metabolic profiles in the first place. Here, we show that Compass predicts that arginine metabolism, and particularly its downstream polyamine branch, are associated with Th17 effector function, and validate this prediction through functional assays.

5.1 Identifying the polyamine pathway as a candidate regulator of Th17 function

The polyamine pathway stood out as one of the most significantly associated with differences in Th17 pathogenicity according to Compass predictions (**Figure 7A**). We constructed a data-driven metabolic network anchored around putrescine, the starting metabolite in canonical polyamine synthesis, by including adjacent metabolites whose reactions are predicted to be associated with pathogenicity (**Figure 7B**). While several polyamine-associated genes were differentially expressed between Th17p and Th17n, the network tied the differential polyamine metabolism to differences in upstream and downstream metabolic reactions which could not be captured from differential gene expression directly. Specifically, Compass predicted that Th17n cells are more active in arginine metabolic pathways, upstream of putrescine, and in alternative fates of putrescine (other than conversion to spermidine along the canonical polyamine synthesis pathway).

5.2 Cellular polyamines are suppressed in regulatory T cells and non-pathogenic Th17

We asked whether critical enzymes of the polyamine pathway (**Figure 7C**) were differentially expressed between CD4$^+$ T cell subsets using qPCR. Ornithine decarboxylase 1 (ODC1) and spermidine/spermine N1 acetyltransferase 1 (SAT1) are the rate-limiting enzymes of polyamine biosynthesis and catabolic processes, respectively. SAT1 level was higher in Th17p than in Th17n or Treg, while ODC1 was similarly expressed in Th17n and Th17p, but significantly lower in Treg. Interestingly, argininosuccinate synthetase 1 (ASS1), an enzyme upstream of polyamine biosynthesis was upregulated in Th17n, consistent with Compass-predicted alternative flux in the polyamine neighborhood (**Figure 7D**). Collectively, these data suggest that the polyamine pathway may be associated with functional state in Th17 and other T cell lineages.

We next measured polyamine metabolites using an enzymatic assay and LC/MS metabolomics (**Supplementary Methods**). Compared to Th17p, Treg and Th17n had reduced levels of total polyamines (**Figure 7E**), reflecting reduced import, reduced biosynthesis, or increased export of polyamines in these cells. We then measured metabolites in the polyamine neighborhood (**Figures 7F**, **S7A**). Consistent with Compass's predictions, there was higher creatine content in Th17n than in Th17p. While cellular ornithine (polyamine precursor) was comparable between Th17p and Th17n, Th17p had higher levels of putrescine and acetyl-putrescine (**Figure 7F**). Of note, cellular spermidine and acetyl-spermidine were not different between the conditions, and spermine was not detected (**Figure 7F**). The reduced putrescine and its acetyl form in Th17n are unlikely due to increased export, as we observed very little polyamines in the media of either Th17n or Th17p (**Figure S7A**). These data suggest that polyamines accumulate within Th17p and that the main function of SAT1 in Th17p may be to recycle rather than to export polyamines.

Arginine and citrulline can be used to synthesize the polyamine precursor ornithine. We used carbon- or hydrogen-labeled arginine or citrulline to study polyamine biosynthesis (**Figure S7B, C**). We harvested cells and media for LC/MS at 0, 1, 5 and 24 hours post addition of labeled arginine. While labeled cellular guanidinoacetic acid, a byproduct of arginine conversion into ornithine, was comparable between Th17n and Th17p, over time Th17p cells accumulated more intracellular putrescine, acetyl-putrescine and acetyl-spermidine (**Figure S7B**), consistent with increased polyamine biosynthesis and/or recycling activity in these cells. Conversely, Th17n accumulated more labeled arginine than Th17p. This prompted us to investigate whether Th17n can also better synthesize (as opposed to better uptake) arginine, which would be consistent with higher ASS1 expression (**Figure 7D**) in these cells. We pulsed cells with labeled citrulline, an arginine precursor, and indeed observed higher levels of labeled arginine in Th17n (**Figure S7C**). Collectively, the targeted metabolomics and tracing data suggest that Th17n accumulate arginine, and that Th17p preferentially synthesize or recycle polyamines. We conclude that differences in polyamine biosynthesis are associated with the different Th17 functional states.

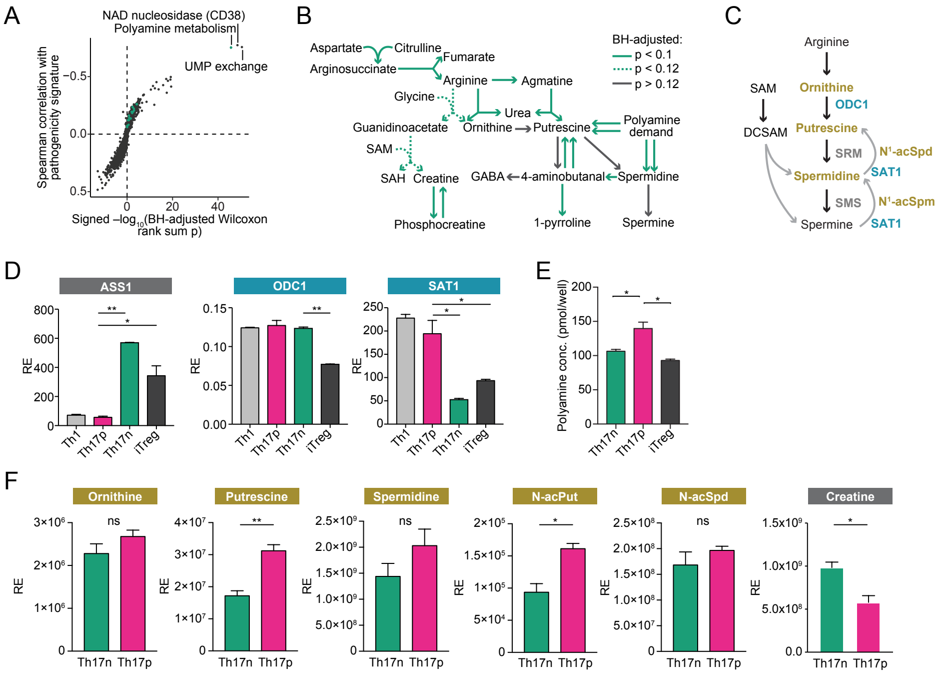

A
NAD nucleosidase (CD38)
Polyamine metabolism
UMP exchange
Spearman correlation with pathogenicity signature
−0.5
0.0
0.5
0
20
40
Signed $-\log_{10}$(BH-adjusted Wilcoxon rank sum p)
B
Aspartate
Citrulline
Fumarate
Arginosuccinate
Arginine
Agmatine
Glycine
Urea
Guanidinoacetate
Ornithine
Putrescine
Polyamine demand
SAM
SAH
Creatine
GABA
4-aminobutanal
Spermidine
Phosphocreatine
1-pyrroline
Spermine
BH-adjusted:
p < 0.1
p < 0.12
p > 0.12
C
Arginine
SAM
Ornithine
ODC1
DCSAM
Putrescine
SRM
N^1-acSpd
Spermidine
SAT1
SMS
N^1-acSpm
Spermine
SAT1
D
ASS1
ODC1
SAT1
RE
800
600
400
200
0
0.15
0.10
0.05
0.00
250
200
150
100
50
0
**
*
Th1
Th17p
Th17n
iTreg
E
Polyamine conc. (pmol/well)
200
150
100
50
0
*
Th17n
Th17p
iTreg
F
Ornithine
Putrescine
Spermidine
N-acPut
N-acSpd
Creatine
RE
3×10^6
2×10^6
1×10^6
4×10^7
3×10^7
2×10^7
1×10^7
2.5×10^9
2.0×10^9
1.5×10^9
1.0×10^9
5.0×10^8
2×10^5
1.5×10^5
1×10^5
5×10^4
2.5×10^8
2.0×10^8
1.5×10^8
1.0×10^8
5.0×10^7
$1.5\text{x}10^9$
$1.0\text{x}10^9$
$5.0\text{x}10^8$
0.0
ns
**
*
Th17n
Th17p

Figure 7. Prediction and metabolic validation of the polyamine pathway as a regulator of Th17 function.

(A) X-axis: Compass-score differential activity test; Th17p vs. Th17n. BH-adjusted Wilcoxon rank sum p signed by the direction of change. Y-axis: Spearman correlation between Compass scores and cell pathogenicity scores **(Supplementary Methods)**. Dots are meta-reactions, green dots are meta-reactions containing at least one reaction that appears in the network of panel B. All reactions comprising the meta-reaction labelled "polyamine metabolism" are denoted in panel B. **(B)** A metabolic network that is preferentially active in Th17n based on Compass results. Green arrows represent reactions predicted to be significantly associated with the Th17n program; double arrows are reactions belonging to the meta-reaction labeled "polyamine metabolism" in panel A. SAM: S-Adenosyl-Methionine; SAH: S-Adenosyl-Homocysteine. GABA: gamma-aminobutyric acid. **(C)** Schematic polyamine pathway based on KEGG. **(D-F)** T cells were differentiated (**Supplementary Methods**; shown is a representative of 2-3 experiments with 3-4 mice each) and harvested at 48h for qPCR (D), and 68h for polyamine enzymatic assay (E) and metabolomics (E-F).

5.3 ODC1 or SAT1 inhibition restricts Th17 function in a putrescine-dependent manner

We studied the effects of polyamine pathway inhibitors on Th17 differentiation *in vitro*. Difluoromethylornithine (DFMO), an irreversible inhibitor of ODC1, suppressed polyamines and inhibited IL-17 expression in both Th17n and Th17p (**Figures 8A-B, S8A**). DFMO also inhibited the expression of other canonical Th17 cytokines such as IL-17A, IL-17F, IL-21 and IL-22, while promoting IL-9 levels in culture supernatant (**Figure 8C**). IL-17 inhibition does not appear to be solely related to IL-2 regulation (Bowlin et al., 1987), as DFMO promoted IL-2 expression in Th17p, but not Th17n (**Figure 8C**). In support of this view, IL-2 neutralization did not rescue the inhibitory effects of DFMO (**Figure S8B-E**). DFMO reduced cell proliferation (data not shown) and this likely contributed to its suppression of Th17 effector functions. However, when considering only cells that had divided exactly once, as means to select cells that are equally potent proliferators, DFMO treatment led to a lower frequency of IL-17$^+$ cells (data not shown). Therefore, DFMO can regulate Th17 also independently of proliferation. The increased secretion of IL-9 (a Th9 cytokine) by DFMO-treated cells also suggests that DFMO does not inhibit Th17 function solely by reducing cell viability.

Turning to Th17 transcription factors, DFMO suppressed Rorgt and Tbet expression in Th17p but not Th17n (**Figure 8D**), suggesting a nuanced effect. Consistently, DFMO decreased pStat3, and not total Stat3 protein levels, only in Th17p (**Figure S8F**). IL-17 inhibition is not due to increased Foxo1 activity, another critical regulator of Th17 function, as DFMO promoted pFoxo1(S256) (**Figure S8F**). Given the reciprocal regulation of Th17 and Treg, and as DMFO also impacted polyamine levels in Tregs (**Figure S8A**), we asked whether DFMO regulates Foxp3 even under Th17 differentiation conditions. DFMO increased the frequency of Foxp3$^+$ cells in Th17n but not Th17p (**Figure 8E**), presumably because TGF-β1 is required for Foxp3 expression.

To determine whether other enzymes of the polyamine pathway also regulate Th17 function, we used inhibitors of spermine synthase (SRM) and spermidine synthase (SMS) (**Figure 8A**). Similar to DFMO, inhibitors of any of the biosynthesis enzymes suppressed IL-17 and promoted IL-9 and Foxp3 expression, the latter in Th17n (**Figure 8F**). Furthermore, using diminazene aceturate to inhibit SAT1, a rate-limiting enzyme of polyamine acetylation and recycling, had similar effects (**Figure 8F**). SAT1 and ODC1 are probably part of a self-regulating feedback loop, as perturbation in one was previously reported to affect the other (Jell et al., 2007; Mounce et al., 2016; Pegg, 2008). Consistent with this finding, inhibition of ODC1 with DFMO suppressed SAT1 expression in Th17 cells (**Figure S8G**).

Finally, we confirmed that the effect of DFMO was through inhibition of ODC1, as addition of putrescine to cells treated with DFMO completely reversed their phenotype (**Figure 8G**). Interestingly, addition of putrescine with SAT1 inhibition partially reversed the upregulation of Foxp3, but not the suppression of IL-17 (**Figure 8H**). This suggests that ODC1 and SAT1 have not only shared but also (despite their proximity in the metabolic network) distinct functions in the regulation of the T effector program.

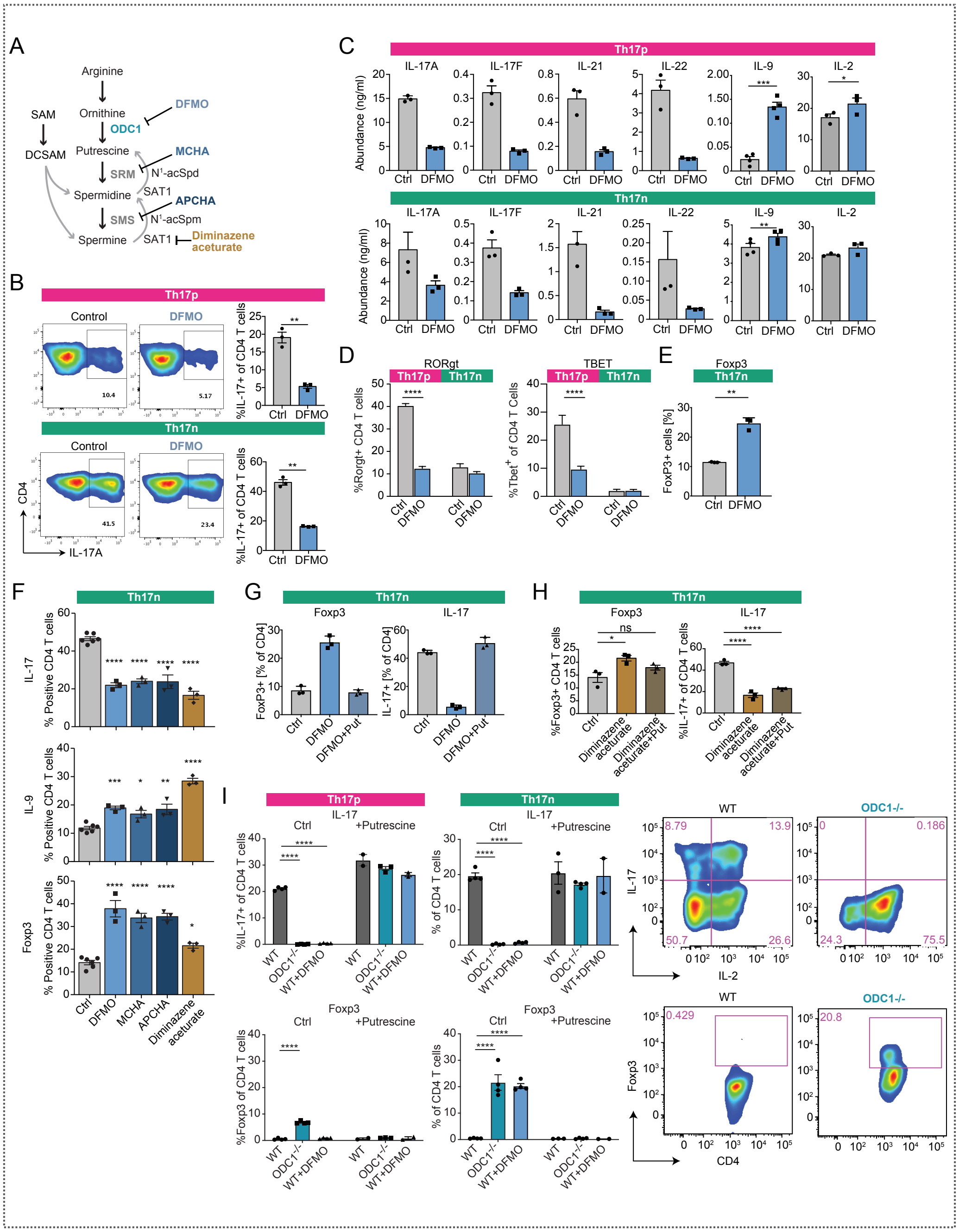

A
Arginine
SAM
Ornithine
DFMO
ODC1
DCSAM
Putrescine
MCHA
SRM
N1-acSpd
SAT1
Spermidine
APCHA
SMS
N1-acSpm
Spermine
SAT1
Diminazene aceturate
B
Th17p
Control
DFMO
10.4
5.17
%IL-17+ of CD4 T cells
Ctrl
DFMO
Th17n
41.5
23.4
CD4
IL-17A
C
Th17p
IL-17A
IL-17F
IL-21
IL-22
IL-9
IL-2
Abundance (ng/ml)
Th17n
D
RORgt
TBET
%Rorgt+ CD4 T cells
%Tbet+ of CD4 T Cells
E
Foxp3
FoxP3+ cells [%]
F
IL-17
IL-9
Foxp3
% Positive CD4 T cells
Ctrl
DFMO
MCHA
APCHA
Diminazene aceturate
G
FoxP3+ [% of CD4]
IL-17+ [% of CD4]
DFMO+Put
H
%Foxp3+ CD4 T cells
%IL-17+ of CD4 T cells
Diminazene aceturate+Put
ns
I
+Putrescine
%IL-17+ of CD4 T cells
% of CD4 T cells
%Foxp3 of CD4 T cells
WT
ODC1-/-
WT+DFMO
WT
ODC1-/-
8.79
13.9
50.7
26.6
0
0.186
24.3
75.5
IL-2
0.429
20.8
CD4

Figure 8. Chemical and genetic perturbations of the polyamine pathway suppress canonical Th17 cytokines.

(A) Polyamine pathway schematic depicting chemical inhibitors. (**B-E**) Flow cytometry of differentiated cells on day 3 (d3); B,D,E intracellular staining; C secreted cytokines by legendplex. DFMO or solvent control (water) was added at d0. (**F**) Inhibition of polyamine pathway targets in *in vitro* differentiated Th17n. (**G-H**) Addition of 2.5mM putrescine has a rescue effect in cultures with inhibited ODC1 (G) or SAT1 (H). (**I**) Flow cytometry of T cells differentiated from WT or ODC1$^{-/-}$ mice under the indicated conditions (n=4 mice). Statistical significance computed by paired t-test or one-way ANOVA, as appropriate for the context, with Bonferroni adjustment for multiple comparisons.

5.4 **ODC1$^{-/-}$ Th17 cells upregulate Foxp3 expression**

To further confirm the effects of polyamine pathway inhibition on T cells, we differentiated Th17 cells from WT and ODC1$^{-/-}$ mice. Similar to DFMO treatment, there was complete inhibition of Th17 canonical cytokines, such as IL-17A, IL-17F and IL-22, but not IFNg, in ODC1$^{-/-}$ Th17 cells (**Figures 8I**, **S8H**). Although ODC1 deficiency did not lead to a decrease in Rorgt expression (data not shown), the loss of Th17 canonical cytokines is consistent with loss of the Th17 program. Furthermore, ODC1$^{-/-}$ Th17n upregulated Foxp3, consistent with promotion of a Treg program (**Figure 8I**). Finally, the observed effects of ODC1$^{-/-}$ were rescued by addition of putrescine (**Figures 8I**, **S8H**).

Chapter 6

The polyamine pathway regulates Th17-induced autoimmunity

In the previous chapter, we established the importance of polyamine metabolism in determining Th17 effector profiles. Here, we support this observation through analysis of *in vitro* molecular profiles and *in vivo* phenotypes of polyamine perturbations. We show that polyamine perturbations substantially shift the transcriptome and chromatin accessibility landscapes of Th17 cells towards a pro-tolerance, Treg-like profile. We find that the transcription factor JMJD3 is required to restrict Th17 differentiation and favor a Treg fate in a polyamine-dependent manner. Deletion of the polyamine synthesizing enzyme ODC1 abrogates the promotion of Th17 differentiation by JMJD3 deficiency. Next, we test two perturbations of the polyamine perturbation *in vivo* – a chemical inhibitor (supplementation of DFMO in drinking water) and a genetic knokcout.

6.1 DFMO restricts Th17-cell transcriptome and epigenome in favor of Treg-like state

To gain mechanistic insight, we profiled bulk RNA from Th17p, Th17n, and iTreg cells treated with DFMO or control (**Supplementary Methods**). DFMO had a profound impact on the transcriptome of all three lineages, driving Th17 cells towards a Treg profile (**Figure 9A**). We next determined the aggregate effect of DFMO on genes that are up-regulated (n=1,284), down-regulated (n=1,255) or comparable (n=8,257) in untreated Th17 vs. Treg (**Figure 9B**). In both Th17p and Th17n, DFMO suppressed the Th17-specific in favor of the Treg-specific transcriptome (**Figure 9C**). Canonical Th17 cell genes such as IL17a, IL17f, and IL23r were suppressed, whereas Treg related genes, such as Foxp3 and CCL3, were upregulated (**Figure 9B**). There was no consistent effect of DFMO on Treg or genes expressed comparably in Th17 and Treg (**Figure 9C**). Thus, the polyamine pathway is important for restricting the Treg-like program in Th17 cells in both Th17p and Th17n states.

DFMO also promoted a subset of genes that are characteristic to Th17n but not to Th17p or iTreg **(Figure S9A)**. Notably, DFMO treatment in both Th17n and Th17p led to elevated levels of the pro-regulatory IL-9 transcript (Lee et al., 2012), consistent with the aforementioned protein analysis **(Figure 8C)** and the initial Compass prediction **(Figure 3D)**. This accords with polyamine metabolism promoting Th17 over Treg fate on the one hand, while being associated with a Th17n fate (marked by high IL-9 expression) over Th17p on the other. Furthermore, DFMO-treated Th17 upregulated the CD5L/AIM transcript, which we showed to be predominantly expressed by Th17n (Wang et al., 2015).

While DFMO profoundly altered the Th17 transcriptome, it did not consistently restrict phosphorylation of key Th17 cell regulators, particularly not in Th17n (**Figure S8F**). We hypothesized that polyamines may impact the epigenome, which was also suggested by the altered expression of many chromatin modifiers in DFMO-treated cells (**Figure S9B**). To test this hypothesis, we measured chromatin accessibility by ATAC-Seq in Th17n and iTreg cells (**Supplementary Methods**). Overall, DFMO treatment resulted in considerable changes in accessible regions in both lineages (**Figure S9C**). We asked whether DFMO preferentially altered accessibility in Th17- and Treg-specific regions by partitioning all accessible peaks into those more accessible (n=10,431), less accessible (n=3,421), or comparably accessible (n=34,591) in untreated Th17n vs. Treg (**Figure 9D**). Consistent with the RNA-Seq data, DFMO restricted accessibility in Th17-specific regions and favored accessibility in Treg-specific regions (**Figure 9D**). Differentially accessible regions were found near loci encoding key effector molecules. For instance, DFMO restricted peaks in the promoter and intergenic regions of IL17a-IL17f that correspond to Rorgt binding sites known to regulate IL-17 expression (**Figure 9E**; based on a ChIP-Seq of Th17 cells by (Xiao et al., 2014)). Thus, DFMO can shape chromatin accessibility in favor of a Treg epigenomic landscape, and this may contribute to the emergence of a Treg-like transcriptional program in DFMO-treated Th17 cells.

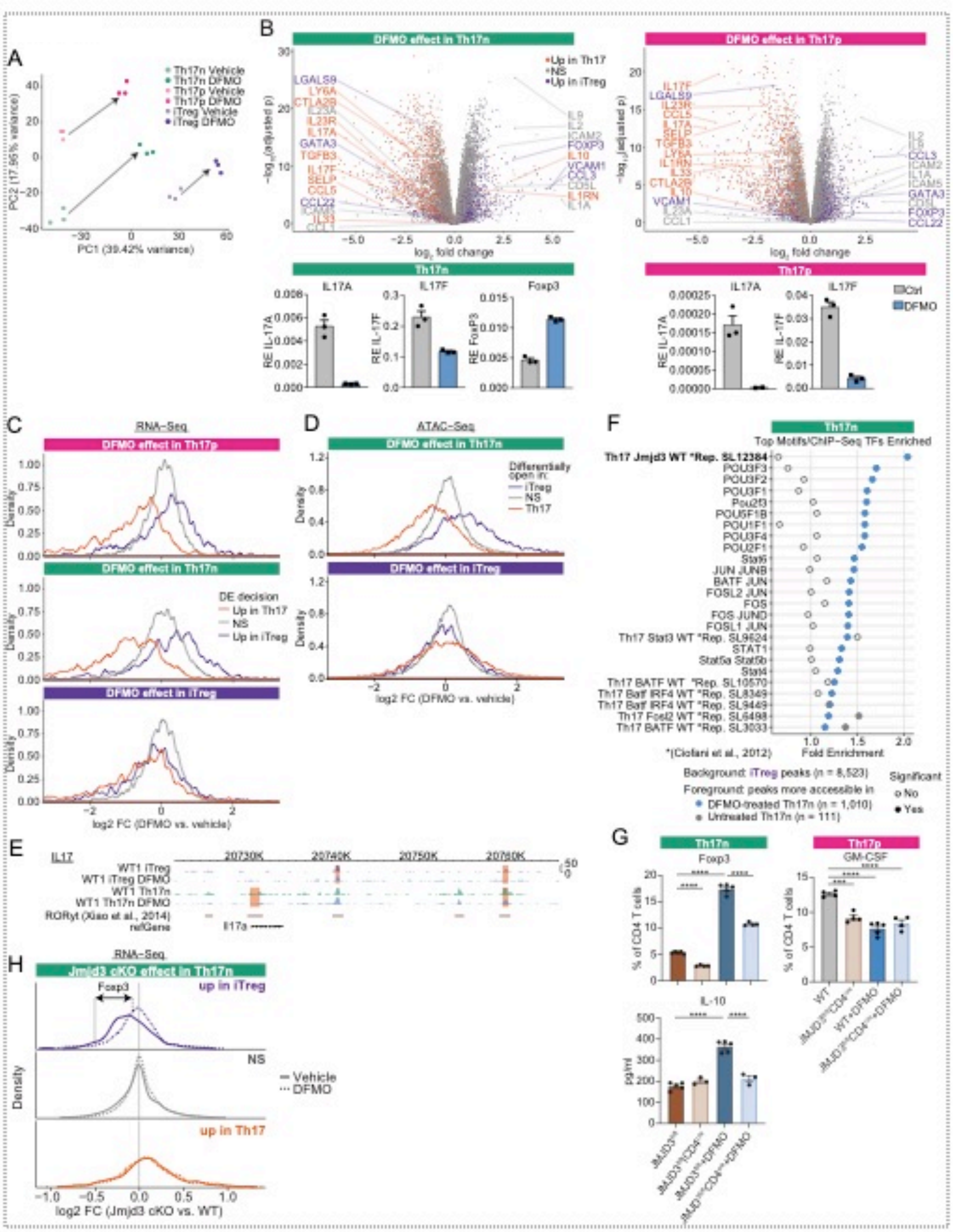

A
Th17n Vehicle
Th17n DFMO
Th17p Vehicle
Th17p DFMO
iTreg Vehicle
iTreg DFMO
PC2 (17.95% variance)
PC1 (39.42% variance)
B
DFMO effect in Th17n
Up in Th17
NS
Up in iTreg
LGALS9
LY6A
CTLA2B
IL23A
IL23R
IL17A
GATA3
TGFB3
IL17F
SELP
CCL5
CCL22
ICAM5
IL33
CCL1
IL9
IL2
ICAM2
FOXP3
IL10
VCAM1
CCL3
CD5L
IL1RN
IL1A
−log10(adjusted p)
log2 fold change
DFMO effect in Th17p
IL17F
LGALS9
IL23R
CCL5
IL17A
SELP
TGFB3
LY6A
IL1RN
IL33
CTLA2B
IL10
VCAM1
IL23A
CCL1
IL2
IL9
CCL3
ICAM2
IL1A
ICAM5
GATA3
CD5L
FOXP3
CCL22
Th17n
IL17A
IL17F
Foxp3
RE IL-17A
RE IL-17F
RE FoxP3
Th17p
Ctrl
DFMO
C
RNA−Seq
DFMO effect in Th17p
DFMO effect in Th17n
DFMO effect in iTreg
DE decision
Up in Th17
NS
Up in iTreg
Density
log2 FC (DFMO vs. vehicle)
D
ATAC−Seq
DFMO effect in Th17n
DFMO effect in iTreg
Differentially open in:
iTreg
NS
Th17
Density
log2 FC (DFMO vs. vehicle)
F
Th17n
Top Motifs/ChIP−Seq TFs Enriched
Th17 Jmjd3 WT *Rep. SL12384
POU3F3
POU3F2
POU3F1
Pou2f3
POU5F1B
POU1F1
POU3F4
POU2F1
Stat6
JUN JUNB
BATF JUN
FOSL2 JUN
FOS
FOS JUND
FOSL1 JUN
Th17 Stat3 WT *Rep. SL9624
STAT1
Stat5a Stat5b
Stat4
Th17 BATF WT *Rep. SL10570
Th17 Batf IRF4 WT *Rep. SL8349
Th17 Batf IRF4 WT *Rep. SL9449
Th17 Fosl2 WT *Rep. SL6498
Th17 BATF WT *Rep. SL3033
*(Ciofani et al., 2012)
Fold Enrichment
Background: iTreg peaks (n = 8,523)
Foreground: peaks more accessible in
DFMO-treated Th17n (n = 1,010)
Untreated Th17n (n = 111)
Significant
No
Yes
E
IL17
20730K
20740K
20750K
20760K
50
0
WT1 iTreg
WT1 iTreg DFMO
WT1 Th17n
WT1 Th17n DFMO
RORγt (Xiao et al., 2014)
refGene
Il17a
G
Th17n
Foxp3
% of CD4 T cells
IL-10
pg/ml
Th17p
GM-CSF
% of CD4 T cells
WT
JMJD3fl/flCD4cre
WT+DFMO
JMJD3fl/flCD4cre+DFMO
JMJD3fl/fl
JMJD3fl/flCD4cre
JMJD3fl/fl+DFMO
JMJD3fl/flCD4cre+DFMO
****
***
H
RNA−Seq
Jmjd3 cKO effect in Th17n
Foxp3
up in iTreg
NS
up in Th17
Vehicle
DFMO
Density
log2 FC (Jmjd3 cKO vs. WT)

Figure 9. DFMO treatment promotes Treg-like transcriptome and epigenome.

T cells were harvested at 68h for live cell sorting and population RNA-Seq (A-C) and ATAC-Seq (D-F). **A**, PCA plot of the RNA-Seq. **B**, Volcano plots (upper) and qPCR validation (lower) of genes affected by DFMO treatment. **C-D,** Density plot showing DFMO effects on transcriptome (C) and chromatin accessibility (D). Genes or loci are divided into those up-regulated or more accessible, respectively, in Th17 cells (orange), Treg (violet) or neither (grey) **(Supplementary Methods)**. **E**, IGV plots of IL17a region; Rorgt binding sites with significantly altered accessibility by DFMO are highlighted; a representative (WT1) of 3 mice is shown. **F**, Rows show hypergeometric enrichment computed for peaks annotated with a genomic feature by Chip-Seq (Ciofani et al., 2012) or DNA binding motifs **(Supplementary Methods)**. Dots are fold enrichment of peaks more accessible in DFMO-treated (blue) or untreated (grey) Th17n against a background of iTreg peaks. **G,** Differentiated cells were rested at 68h and harvested at 120h for analysis of intracellular Foxp3 or GM-CSF, and IL-10 secretion. BH-adjusted $p < 0.0001$ denoted ****; $p < 0.001$ denoted ***. A representative of 2 experiments with 4 mice each. **H,** Cells from **G** were harvested at 68h for RNA-Seq (n=4). Density plots showing JMJD3 deficiency and DFMO treatment effects. Treg and Th17 programs were defined similarly to panel C.

6.2 The chromatin regulator JMJD3 maintains Treg-like state in Th17 cells in a polyamine-dependent manner

To investigate which TFs may be responsible for the suppression of the Th17-specific program and upregulation of the Treg program, we looked for putative binding sites overlapping with regions whose accessibility is modulated by DFMO. When considering genomic regions that are typically more accessible in Tregs compared to Th17 and may be modulated by DFMO (**Figures 9F, S9D**), we found that in Th17n cells, DFMO increased chromatin accessibility near potential binding sites of JMJD3 along with a number of POU-domain containing TFs.

As JMJD3 regulates T cell plasticity (Ciofani et al., 2012; Li et al., 2014; Liu et al., 2015), we tested whether the transcriptional shifts induced by DFMO were altered in JMJD3$^{fl/fl}$CD4^{cre} T cells. JMJD3 deficiency reduced Foxp3 expression and abrogated the upregulation of Foxp3 and IL-10 by DFMO in Th17n (**Figure 9G**). On the other hand, both JMJD3 deficiency and ODC1 inhibition suppressed GM-CSF expression in Th17p (**Figure 9G**). Of note, DFMO did not alter JMJD3 expression in Th17 cells (data not shown). These data suggest that JMJD3 and ODC1 have both distinct and overlapping functions. Consistent with the flow cytometry data, JMJD3 deficiency in Th17n resulted in a global transcriptome-wide shift restricting the Treg and (to a lesser extent) promoting the Th17 program (**Figure 9H**). This agrees with previous findings that JMJD3 ablation *in vivo* promoted intestinal Th17 differentiation (Li et al., 2014). Notably, ODC1 inhibition by DFMO abolished the suppression of the Treg program by JMJD3 deficiency (differences between solid and dotted curves in **Figures 9H, S9E**). However, the milder transcriptomic shift towards the Th17 program in JMJD3$^{fl/fl}$CD4^{cre} mice was not reversed with further ODC1 inhibition by DFMO. In the other direction, JMJD3 had little global impact on the effects of DFMO (**Figure S9F**). We conclude that JMJD3 requires uninhibited ODC1 activity to sustain the Treg program, whereas the promotion of the Th17 program and the suppression of the Treg program by ODC1 activity do not require JMJD3.

6.3 Perturbation of ODC1 and SAT1, key enzymes of the polyamine pathway, alleviates EAE

We investigated the polyamine pathway *in vivo* in the context of EAE, a central nervous system (CNS) autoimmune disease, via two approaches: chemical inhibition of ODC1 and T-cell specific genetic deletion of SAT1 (**Figure 10**). As targeting multiple nodes in the polyamine pathway upregulated Foxp3 during Th17 differentiation *in vitro* (**Figures 8**-**9**), we hypothesized that perturbing rate-limiting enzymes *in vivo* would regulate the induction of EAE.

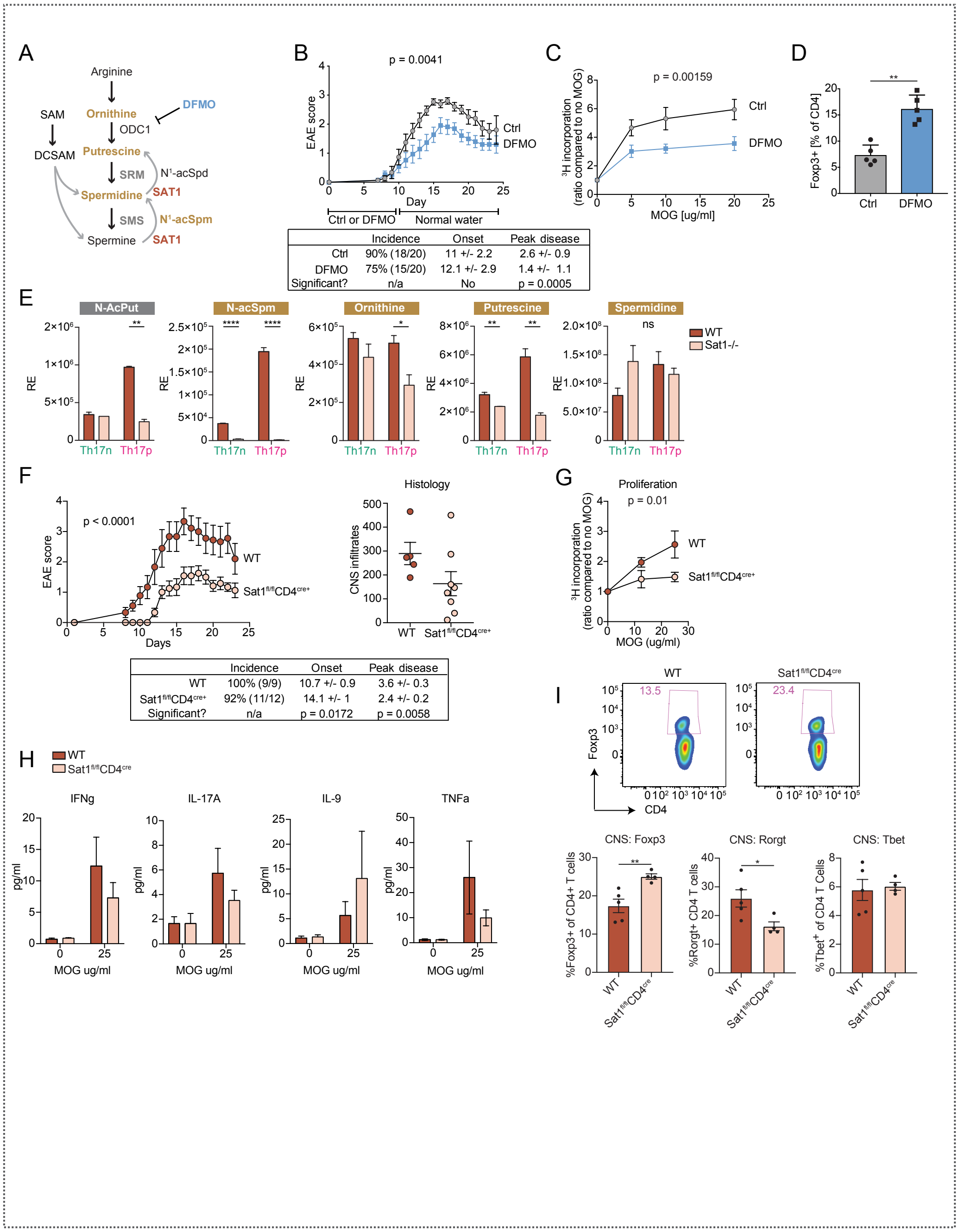
We first tested ODC1 inhibition by adding DFMO to the drinking water of mice immunized with MOG/CFA for EAE induction. DFMO delayed the onset and severity of EAE (**Figure 10B**), and reduced antigen-specific recall responses as measured by T cell proliferation in the draining lymph node (dLN) (**Figure 10C**). Further analysis of T cells isolated from the CNS showed no difference in cytokine production, but increased frequency of Foxp3^{+} CD4^{+} T cells (**Figure 10D** and data not shown). These results agree with polyamine biosynthesis regulating the Th17/Treg balance in favor of Th17, and consequently the induction of autoimmune CNS inflammation.

Since administration of DFMO via drinking water could affect multiple cell types, we also genetically deleted SAT1 in CD4^{+} T cells (SAT1$^{fl/fl}$CD4^{cre}). We confirmed that genetic SAT1 deficiency in Th17 abrogated polyamine acetylation (**Figure 10E**). Notably, loss of SAT1 also reduced (non-acetylated) putrescine levels, supporting a feedback mechanism as discussed above (**Figure S8G**).

We observed delayed onset and severity of EAE in SAT1$^{fl/fl}$CD4^{cre} mice **(Figure 10F**), as well as inhibition of antigen-specific recall responses (**Figure 10G**), similar to the effects observed with DFMO treatment. Although cytokine production was unaltered *ex vivo* by SAT1 deficiency, there was a trend towards decreased antigen-dependent IFNg, IL-17 and TNF production, and increased IL-9 (**Figure 10H, S9A**). Foxp3^{+}CD4^{+} T cells were increased with a concomitant decrease of Rorgt^{+}CD4^{+} T cells isolated from the CNS of SAT1$^{fl/fl}$CD4^{cre} mice (**Figure 10I**). Notably, the frequencies of Foxp3^{+} or Rorgt^{+} cells were not different in dLN (**Figure S10B**), suggesting that the effect of SAT1 on T cells may be amplified in tissue recall responses. In conclusion, we demonstrated using both chemical and genetic perturbations at multiple levels that the polyamine pathway is an important mediator of autoimmune inflammation.

6.4 **Limitations of the study**

Concerning polyamine metabolism, chemical inhibitions of several target enzymes led to similar effects on Th17 cells, but genetic deletions of ODC1 and SAT1 did not produce identical results. While both ODC1 and SAT1 deletion promoted Foxp3 expression, ODC1 but not SAT1 suppressed Th17 cytokines *in vitro* (**Figure 8I** and data not shown). Further studies are necessary to understand the mechanistic roles and reciprocal regulation between the enzymes of the polyamine pathway *in vivo*.



	Incidence	Onset	Peak disease
Ctrl	90% (18/20)	11 +/- 2.2	2.6 +/- 0.9
DFMO	75% (15/20)	12.1 +/- 2.9	1.4 +/- 1.1
Significant?	n/a	No	p = 0.0005

	Incidence	Onset	Peak disease
WT	100% (9/9)	10.7 +/- 0.9	3.6 +/- 0.3
Sat1$^{fl/fl}$CD4^{cre+}	92% (11/12)	14.1 +/- 1	2.4 +/- 0.2
Significant?	n/a	p = 0.0172	p = 0.0058

Figure 10. Targeting ODC1 and SAT1 alleviates experimental autoimmune encephalomyelitis (EAE).

(**A**) Polyamine pathway schematics; (**B-D**) Effects of chemical inhibition of ODC1 by DFMO in drinking water on MOG_{35-55}/CFA induced EAE in wildtype mice. **B**, Clinical score over time. **C**, Antigen-specific proliferation of cells isolated from draining lymph node (dLN) at d23 (**STAR Methods**). **D**, Flow cytometry of T cells isolated from CNS at d15. (**E-I**) Effects of genetic perturbation of SAT1. **E**, Abundance of metabolites in differentiated WT or $SAT1^{-/-}$ T cells. (**F-I**) EAE was induced in WT and $SAT1^{fl/fl}CD4^{cre}$ mice. **F**, Clinical score (left) and histological score (right) showing the number of CNS infiltrates. **G-H**, Antigen-specific response of cells from dLN (d23) is analyzed by proliferation (G) and antigen-specific cytokine secretion by legendplex (H). **I**, Flow cytometry of cells isolated from CNS at d15. Linear regression analysis (b, c, f, g), two-way ANOVA (e) and Student t-test (d, i) were used for statistical analysis.

Chapter 7

Conclusions

We presented Compass, a flux balance analysis (FBA) algorithm to study metabolism in single-cell transcriptomic data and validated a number of its predictions by molecular and functional analyses. These results support the power of transcriptomic-based FBA to make valid predictions in a mammalian system. The network-wide approach enabled Compass to successfully predict metabolic targets in both central and ancillary pathways.

Static FBA assumes that the system under consideration operates in chemical steady state (Varma and Palsson, 1994). Even under this assumption, there remains an infinite number of feasible flux distributions that satisfy the preset biochemical constraints. Therefore, most studies assume that cells aim to optimize some metabolic function, usually biomass production (Damiani et al., 2019). However, while biomass maximization may successfully predict phenotypes in unicellular organisms (Lewis et al., 2010), it is ill-suited for studying mammalian cells (Adler et al., 2019). To overcome this challenge, rather than optimizing a single metabolic objective, 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 coordinate per reaction). A biological signal (e.g., differential reaction activity) can be detected despite the 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.

Compass correctly predicted the role of aerobic glycolysis in the induction of pathogenic Th17 and the role of beta-oxidation in the induction of pro-regulatory Th17, mirroring previous findings in comparisons of Th17 to Treg. 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 (Geltink et al., 2018; MacIver et al., 2013; O’Brien and Finlay, 2019; O’Neill and Pearce, 2016; O’Neill et al., 2016; Pearce and Everts, 2015; Pearce et al., 2013; van der Windt and Pearce, 2012), the role of glycolysis in induction of pro-inflammatory phenotypes may more nuanced (Newton et al., 2016; Van den Bossche et al., 2017).

Compass also predicted that novel targets in peripheral metabolism were associated with Th17 pathogenicity, one of which was the polyamine pathway. We investigated the metabolic circuitry centered around the polyamine pathway, and demonstrated its critical role in promoting pathogenicity and restricting a Treg like program in Th17 cells. Because of reciprocal generation of Th17/Treg cells, the effects observed with the inhibition of polyamine pathway may be unique to diseases where Th17 cells are the effector cells. The significance of the polyamine pathway in autoimmunity contexts is further supported by anecdotal data that polyamine levels are increased in several autoimmune diseases (Hsu et al., 1994; Karouzakis et al., 2012) and it is thought that aberrant polyamine metabolism contributes to autoantigen stabilization (Brooks, 2013). Here, we present a mechanism through which the polyamine pathway can regulate epigenome and thereby Th17/Treg balance and impact development of autoimmunity. We showed that difluoromethylornithine (DFMO), which inhibits a key step of polyamine synthesis, substantially alters the Th17 effector profile. Since DFMO is a well-tolerated, FDA-approved drug (Casero et al., 2018), there might be grounds for drug repurposing of DFMO for use in immune disorders.

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 (Tanay and Regev, 2017). 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.

Chapter 8

Extended methods

8.1 Experimental models

8.1.1 T cell differentiation culture

Naive CD4+CD44-CD62L+CD25-T cells were sorted using BD FACSAria sorter and activated with plate-bound anti-CD3 (1µg/ml) and anti-CD28 antibodies (1µg/ml) in the presence of cytokines at a concentration of 0.5 10^6 cells/ml. For T cell differentiations the following combinations of cytokines were used: pathogenic Th17: 25ng/ml rmIL-6, 20ng/ml rmIL-1b (both Miltenyi Biotec) and 20ng/ml rmIL-23 (RD systems); non-pathogenic Th17: 25ng/ml rmIL-6 and 2ng/ml of rhTGFb1 (Miltenyi Biotec); iTreg: 2ng/ml of rhTGFb1; Th1: 20ng/ml rmIL-12 (RD systems); Th2: 20ng/ml rmIL-4 (Miltenyi Biotec). For differentiation experiments, cells were harvested at 68 hours for RNA analysis or ATAC-Seq and 72-96h for flow cytometry analysis and Seahorse assay.

8.1.2 Mice

C57BL/6 wildtype (WT) mice were obtained from Jackson Laboratory (Bar Harbor, ME). WT 2D2 transgenic mice were bred in house. SAT1flox mice were kindly provided by Dr. Manoocher Soleimani (University of Cincinnati), which we crossed to CD4^{cre} to generate conditional T cell deletion of SAT1. ODC1$^{fl/fl}$CD4^{cre} were gifted by Dr. Erika Pearce (Max Planck Institute). For experiments, mice were matched for sex and age, and most mice were 6–10 weeks old. Littermate WT or Cre^{-} mice were used as controls. All experiments were performed in accordance to the guidelines outlined by the Harvard Medical Area Standing Committee on Animals at the Harvard Medical School or the Brigham and Women's hospital Institutional Animal Care and Use Committees (Boston, MA).

8.1.3 Experimental Autoimmune Encephalomyelitis (EAE)

For adoptive transfer EAE, naive 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.

For active EAE immunization, MOG35-55 peptide was emulsified in complete freund adjuvant (CFA). Equivalent of 40µg MOG peptide was injected per mouse subcutaneously followed by pertussis toxin injection intravenously on day 0 and day 2 of immunization. Mice were treated with 0.5% DFMO in drinking water for 10 days as indicated. DFMO was replenished every third day.

EAE was scored as previously published (Jager et al., 2009).

8.2 Experimental procedures

8.2.1 Flow cytometry

Intracellular cytokine staining was performed after incubation for 4-6h with Cell Stimulation cocktail plus Golgi transport inhibitors (Thermo Fisher Scientific) using the BD Cytofix/Cytoperm buffer set (BD Biosciences) per manufacturer's instructions. Transcription factor staining was performed using the Foxp3/Transcription Factor Staining Buffer Set (eBioscience). Proliferation was assessed by staining with CellTrace Violet (Thermo Fisher Scientific) per manufacturer's instructions. Apoptosis was assessed using Annexin V staining kit (BioLegend). Phosphorylation of proteins to determine cell signaling was performed with BD Phosflow buffer system (BD bioscience) as per manufacturer's instructions.

8.2.2 Seahorse assay

Seahorse assay was performed and seahorse media was prepared following manufacturer instructions (Agilent). Approximately 50,000 T cells were seeded per well in 96 well plate for seahorse assay and readout were normalized against cell count.

8.2.3 Inhibitors and metabolites

Inhibitors are added at the time of differentiation as follows: 100-200µM DFMO, 500µM trans-4-methylcyclohexylamine (MCHA, both Sigma), 500µM N-(3-Amino-propyl)cyclohexylamine (APCHA, Santa Cruz Biotechnology), 50µM diminazene aceturate (Dize, Cayman Chemical) with or without 2.5 mM putrescine (Sigma, P7505) as indicated.

8.2.4 qPCR

RNA was isolated using RNeasy Plus Mini Kit (Qiagen) and reverse transcribed to cDNA with iScript cDNA Synthesis Kit (Bio-Rad). Gene expression was analyzed by quantitative real-time PCR on a ViiA7 System (Thermo Fisher Scientific) using TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific) with the following primer/probe sets:

- *Ass1* (Mm00711256 m1)
- *Odc1* (Mm02019269 g1)
- *Sat1* (Mm00485911 g1)
- *Srm* (Mm00726089 s1)
- *Sms* (Mm00786246 s1)
- *Il-17a* (Mm00439618 m1)
- *Il-17f* (Mm00521423 m1)
- *Foxp3* (Mm00475162 m1),
- *Tead1* (Mm00493507 m1)
- *Taz* (Mm00504978 m1)
- *Actb* (Applied Biosystems)

Expression values were calculated relative to *Actb* detected in the same sample by duplex qPCR.

8.2.5 Polyamine ELISA

Cell pellets of *in vitro* differentiated cells were frozen down and further processed with the Total Polyamine Assay Kit (BioVision Inc.) according to the manufacturer’s instructions.

8.2.6 Legendplex

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).

8.3 **RNA-Seq**

8.3.1 Bulk RNA sequencing

For population (bulk) RNA-seq, *in vitro* differentiated T cells were sorted for live cells and lysed with RLT Plus buffer and RNA was extracted using the RNeasy Plus Mini Kit (Qiagen). Full-length RNA-Seq libraries were prepared as previously described [Singer et al., 2016] and paired-end sequenced (75 bp 2) with a 150 cycle Nextseq 500 high output V2 kit.

8.3.2 Smart-Seq single-cell RNA sequencing

Full experimental details are given in (Gaublomme et al., 2015). Briefly, we sequenced CD4+ naive T cells 48hrs post polarization under the pathogenic (Th17p) or non-pathogenic (Th17n), ultimately retaining after quality tests 151 IL-17A/GFP+ Th17n cells, and 139 IL-17A/GFP+ Th17p cells. Unlike (Gaublomme et al., 2015), in the present study we analyze only IL-17A/GFP+ sorted cells.

8.3.3 Estimation of transcript abundance from RNA libraries

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 [Fletcher et al., 2017]. This computational pipeline is a massively revised and updated version of the one originally used to analyze these libraries (Gaublomme et al., 2015). Batch effects and other nuisance factors were normalized with a model chosen empirically with SCONE (Cole et al., 2017) [Cole et al., 2019]. Bulk RNA-Seq were processed with a modified variant of the same pipeline, or with a Kallisto (Bray et al., 2016) pipeline. Both gave similar results.

8.3.4 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 (Ritchie et al., 2015). For the bulk RNA libraries, differentially expressed genes were called with `limma-trend` or `limma-voom` (Law et al., 2014) depending on the variance of library sizes, as recommended in the `limma` package manual. Differentially expressed genes (DEGs) were decided by cutoffs of BH-adjusted $p \leq 0.05$ fold-change of at least 1.5 in absolute value.

8.3.5 Further bioinformatic analysis of RNA-Seq data

Bulk RNA libraries from DFMO- or vehicle-treated Th17p, Th17n, or Treg were studied with 3 replicates per condition for a total of 18 libraries as shown in Figure 9A. In all subsequent analyses, genes are considered differentially expressed by the cutoffs defined above.

The PCA shown in Figure 9A was computed on the set of 3,414 that were differentially expressed in comparisons of vehicle-treated Th17 (i.e., Th17n or Th17p) vs. with iTreg, or vehicle-treated Th17p vs. Th17n to focus it on the subspace of the transcriptome relevant to Th17 pathogenicity phenotypes.

Genes that are associated with a Th17 or Treg programs (orange and purple, respectively, in Figure 9B-C) were determined by differential expression test between bulk RNA libraries of vehicle-treated Th17 (i.e., Th17n or Th17p) one side and iTreg on the other. Similar results were obtained when the definition of the Th17 program (orange curve) and Treg programs was based on a comparison of only Th17n against iTreg cells.

A similar analysis was performed on the independent bulk-RNA dataset that involved Jmjd3 conditional knockout (Figures 9H, S9E-F). Th17 and Treg programs were defined based on differential expression of the vehicle-treated conditions from the dataset (i.e., not borrowed from the results of comparison within the previous dataset). Note, however, that no Th17p cells were included in this dataset, and therefore the Th17 program (orange curve) in Figures 9H, S9E-F is based on a comparison of Th17n against iTreg cells. Th17 and Treg-associated peaks in ATAC-Seq (Figure 9D) were similarly defined based on the ATAC-Seq dataset.

To further stratify genes into Th17-related programs (Figure S9A; based on the bulk RNA dataset shown in Figure 9A), we applied a heuristic based on four comparisons (all in vehicle-treated cells):

- Th17 (i.e., Th17n or Th17p) vs. iTreg
- Th17p vs. iTreg
- Th17n vs. iTreg
- Th17p vs. Th17n

The comparisons are decided based on differentially expression with the cutoffs defined above. A label is assigned to each gene as follows:

```
1. IF      Th17    vs.   iTreg   ==   iTreg   AND

           Th17p   vs.   iTreg   ==   iTreg   AND
```

```
                    Th17n  vs.  iTreg  ==  iTreg
   THEN RETURN  Treg (purple)
2. ELSE IF          Th17   vs.  iTreg  ≠   iTreg  AND
                    Th17p  vs.  iTreg  ==  Th17p  AND
                    Th17p  vs.  Th17n  ==  Th17p
   THEN RETURN  Th17p (magenta)
3. ELSE IF          Th17   vs.  iTreg  ≠   iTreg  AND
                    Th17n  vs.  iTreg  ==  Th17n  AND
                    Th17p  vs.  Th17n  ==  Th17n
   THEN RETURN  Th17n (green)
4. ELSE IF          Th17   vs.  iTreg  ==  Th17
   THEN RETURN  Th17 (orange)
5. ELSE RETURN  NS (grey)
```

The gene is associated with the Th17n, Th17p, general Th17, or Treg program by the returned label.

8.4 LC/MS metabolomics and carbon tracing

8.4.1 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. At 68hrs, cells were washed and cultured in media supplemented with 8 mM [U-13C]-glucose for 15min or 3hrs. Alternatively, at 68hrs, cells were washed and cultured in media supplemented with arginine (^{13}C6, Sigma, Cat 643440) or aspartic acid (^{13}C4, Sigma, Cat 604852) for 1, 5 and 24 hours.

8.4.2 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:

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

8.5 **Downstream analysis of Compass scores**

8.5.1 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. Core pathways are defined as Recon2 subsystems that have at least 3 core reactions. Metabolic genes are defined as the set of genes annotated in Recon2.

8.5.2 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).

8.5.3 Manual curation of central carbon predictions

We manually curated the significant predictions of the central carbon metabolism pathways discussed in the manuscript (glycolysis, 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 4A). Note that Compass operates independently on the forward and backward directions of every reaction, and that the direction is denoted in the pathway diagrams of this manuscript.

8.6 **Transcriptomic signatures**

8.6.1 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 (Gaublomme et al., 2015; Wang et al., 2015). Briefly, for each cell compute the average z-scored expression (log(1 +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.

8.6.2 A compendium of T cell state signatures

A compendium of T cell state transcriptomic signatures was described in (Gaublomme et al., 2015). 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 (Kanehisa et al., 2017) 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.

8.6.3 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, $p < 3 \cdot 10^{-16}$).

8.6.4 Late-stage Th17 differentiation

We defined a transcriptomic signature for late-stage differentiation of Th17 cells based on microarray data from [Yosef et al., 2013]. 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.

8.7 ATAC-Seq

8.7.1 Library preparation

For population ATAC-Seq, *in vitro* differentiated T-cells were sorted for live cells and stored in Bambanker freezing media (Thermo Fisher Scientific) at -80°C until further processing. Prior to library preparation, cells were thawed at 37°C and washed with PBS. For ATAC-Seq, cell pellets were lysed and tagmented in 1X TD Buffer, 0.2µl TDE1 (Illumina), 0.01% digitonin, and 0.3X PBS in 40µl reaction volume following the protocol described in (Corces et al., 2016). Transposition reactions were incubated at 37for 30 min at 300 rpm. The DNA was purified from the reaction using a MinElute PCR purification kit (QIAGEN). The whole resulting product was then PCR-amplified using indexed primers with NEBNext High-Fidelity 2X PCR Master Mix (NEB). First, we performed 5 cycles of pre-amplification. We sampled 10% of the pre-amplification reaction for SYBR Green quantitative PCR to assess the number of additional cycles needed for final amplification. After purifying the final library with the MinElute PCR purification kit (QIAGEN), the library was quantified for sequencing using qPCR and a Qubit dsDNA HS Assay kit (Invitrogen). Libraries were sequenced on an Illumina NextSeq 550 system with paired-end reads of 37 base pairs in length.

8.7.2 Alignment of ATAC-Seq and peak calling

All ATAC-Seq reads were trimmed using Trimmomatic (Bolger et al., 2014) to remove primer and low-quality bases. Reads smaller than 36bp were dropped. Reads were then passed to FastQC `[http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]` to check the quality of the trimmed reads. The paired-end reads were then aligned to the mm10 reference genome using bowtie2 (Langmead and Salzberg, 2012), allowing maximum insert sizes of 2000 bp, with the `--no-mixed` and `--no-discordant` parameters added. Reads with a mapping quality (MAPQ) below 30 were removed. Duplicates were removed with

PicardTools, and the reads mapping to the blacklist regions and mitochondrial DNA were also removed. Reads mapping to the positive strand were moved +4bp, and reads mapping to the negative strand were moved -5bp following the procedure outlined in (Buenrostro et al., 2013) to account for the binding of the Tn5 transposase.

Peaks were called using macs2 on the aligned fragments (Zhang et al., 2008) with a qvalue cutoff of 0.001 and overlapping peaks among replicates were merged.

8.7.3 Tests of differential accessibility

Differential accessibility was assessed using DESeq2 (Love et al., 2014) with a matrix of peaks (merging all samples) by samples. Similar to common practice in the analysis of differential gene expression, our analysis of differential accessibility was conducted using the number of observed Tn5 cuts (i.e., number of reads).

Peaks that are associated with a Th17 or Treg programs (orange and purple, respectively, in Figure 9D) were determined by differential accessibility test between libraries of vehicle-treated (water) Th17n and Th17p on one side (unpublished dataset) and Treg on the other with BH-adjusted $p \leq 0.05$ and absolute value of log2 fold-change of at least 1.

8.7.4 Reprocessing of published ChIP-Seq data

ChIP-Seq Peaks from (Xiao et al., 2014) were transferred from mm9 to mm10 using the UCSC liftOver tool. ChIP-Seq replicates from (Ciofani et al., 2012) were downloaded and were trimmed using Trimmomatic [Bolger et al., 2014] to remove primer and low-quality bases. Reads were then passed to FastQC `[http://www.bioinformatics.babraham.ac.uk/projects/fastqc/]` to check the quality of the trimmed reads. These single-end reads were then aligned to the mm10 reference genome using bowtie2 (Langmead and Salzberg, 2012) allowing maximum insert sizes of 2000 bp, with the `--no-mixed` and `--no-discordant` parameters added. Reads with a mapping quality (MAPQ) below 30 were removed. Duplicates were removed with PicardTools, and the reads mapping to the blacklist regions and mitochondrial DNA were also removed.

ChIP-Seq peaks were called in each replicate, versus a control sample, using macs2 (Zhang et al., 2008) with a qvalue cutoff of 0.05.

8.7.5 Enrichment of motifs and ChIP-Seq peaks in differentially accessible regions

Peaks were considered differentially accessible if they had a BH-adjusted $p < 0.05$. We calculated fold enrichment of various genomic features in these peaks (described below) versus a background set of peaks. q-values were estimated using q-value `R` package [`http://github.com/jdstorey/qvalue`]

8.7.6 Motifs and annotation tracks

PWM's for motifs were downloaded from the 2018 release of JASPAR (Khan et al., 2018). We used FIMO (Grant et al., 2011) to identify motifs in mm10, and applied the default threshold of 10^{-4}. We also included the following genomic features from the UCSC Genome Browser `[http://genome.ucsc.edu]`: the ORegAnno database (Lesurf et al., 2016), conserved regions annotated by the PHAST package (Siepel et al., 2005), and repeat regions annotated by RepeatMasker `[http://www.repeatmasker.org]`.

8.7.7 GREAT pathways and genes

Loci were associated with genes and pathways using GREAT (McLean et al., 2010), submitted with the rGREAT `R` package `[https://github.com/jokergoo/rGREAT]`. We retrieved pathways found in the MSigDB Immunologic Signatures, MSigDB Pathways, and GO Biological Processes databases.

8.7.8 Statistical Analysis of non-sequencing data

Unless otherwise specified, the tests performed were two-tail Welch t-tests using Prism software. P values were adjusted with the Bonferroni method for multiple comparisons where appropriate. P value less than 0.05 is considered significant ($P < 0.05$ = *; $P < 0.01$ = **; $P < 0.001$ = ***) unless otherwise indicated.

Supplementary Figures

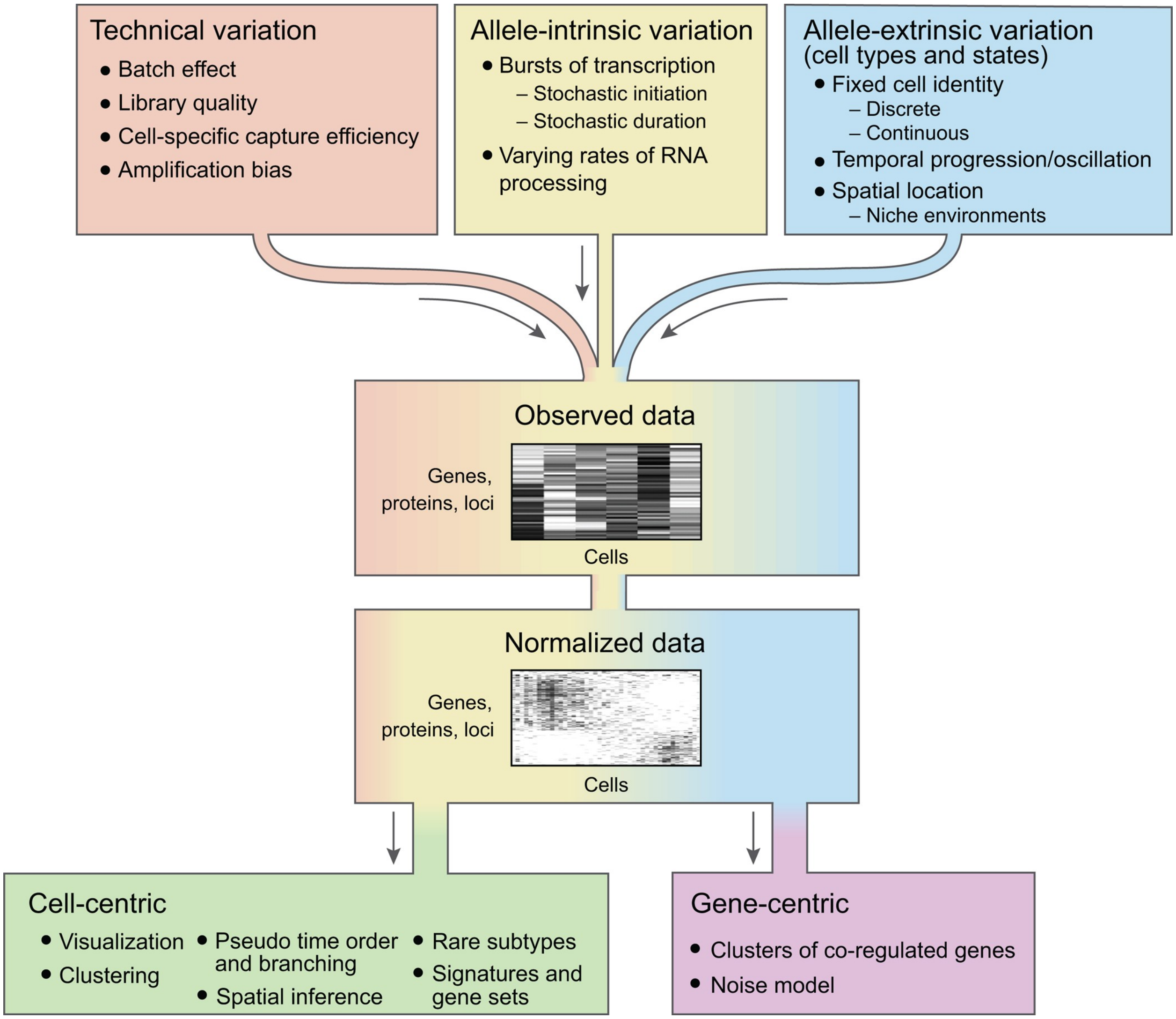
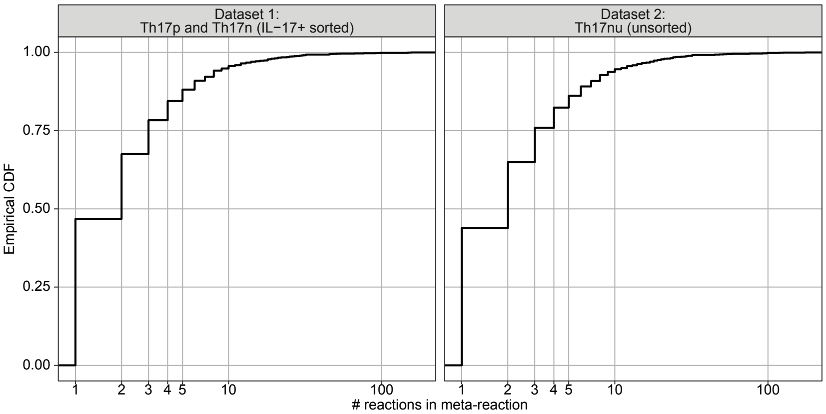


Figure S1. Biological and technical factors combine to determine the measured genomic profiles of single cells; computational methods remove technical effects and tease apart facets of the biological variation, related to Figure 1.

The sources of variation that affect single-cell genomics data are (1) technical factors that reflect variance due to the experimental process (e.g., batch effects); (2) factors that are intrinsic to the process under study (e.g., transcription) and reflect stochastic fluctuations (e.g., transcriptional or translational bursts in mRNA or proteins) that do not correlate between two alleles of the same gene; and (3) factors that are extrinsic to the process under study, reflecting the presence of different cell types and states (e.g., concentrations of key transcription, translation, or metabolic factors). Computational methods are needed to remove the nuisance technical variation (although they typically cannot completely eliminate it) before the biological variation can be confidently explored. Most single-cell studies explore allele-extrinsic factors and can be classified as either cell-centric or gene-centric. Cell-centric analyses aim to catalog the cells into phenotypic groups, whether discrete (e.g., clustering) or continuous (e.g., temporal ordering). Gene-centric analyses aim to understand the dynamics and regulation of the generating mechanisms (e.g., transcriptional circuits).

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

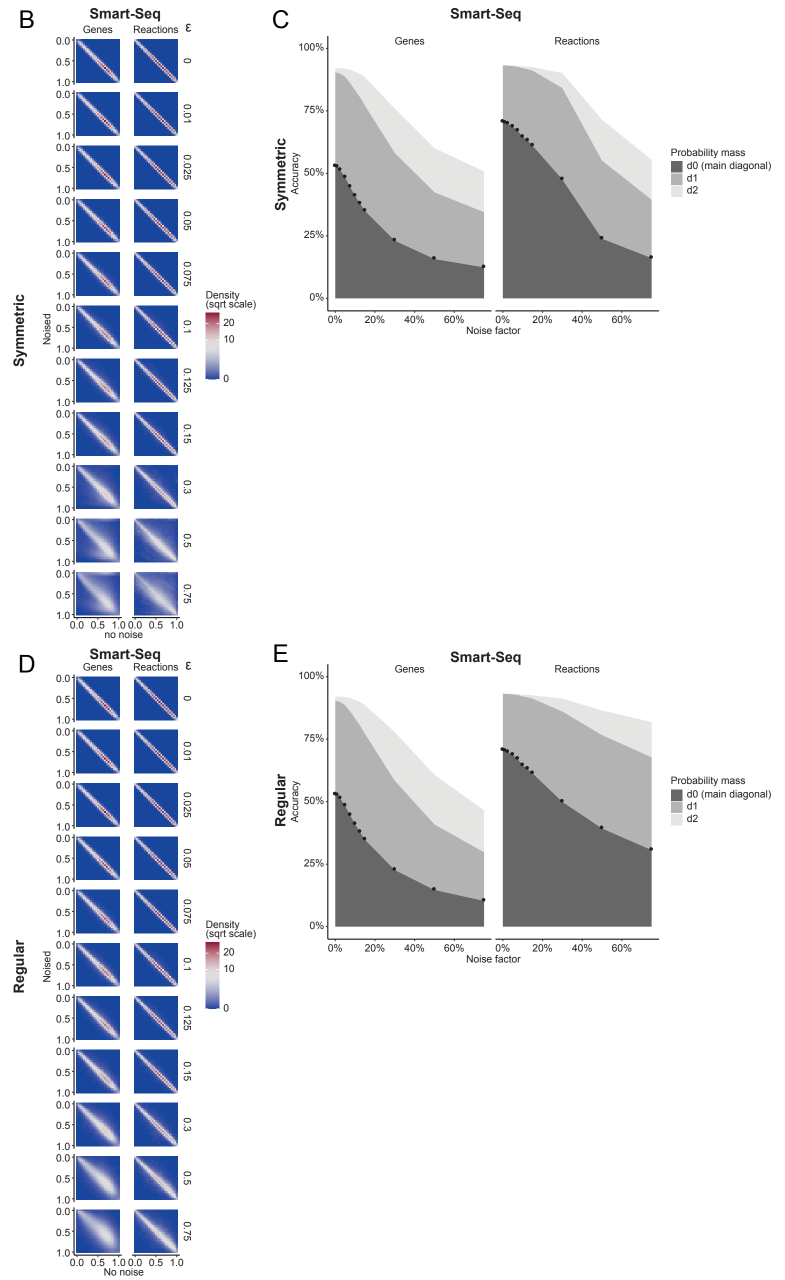
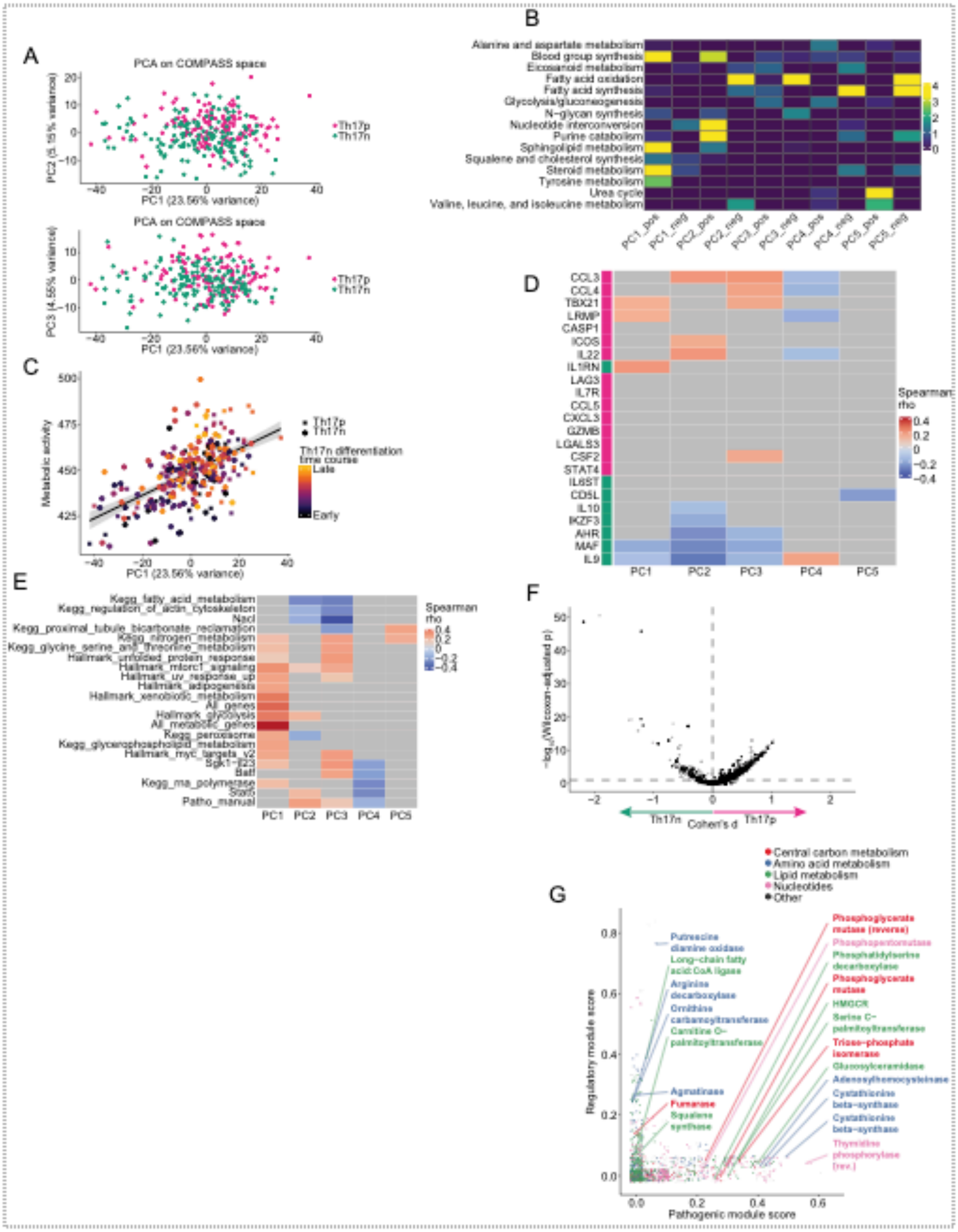

B
Smart-Seq
Genes
Reactions
ε
0
0.01
0.025
0.05
0.075
0.1
0.125
0.15
0.3
0.5
0.75
Symmetric
Noised
no noise
0.0
0.5
1.0
Density
(sqrt scale)
20
10
0
C
Smart-Seq
Genes
Reactions
Symmetric
Accuracy
100%
75%
50%
25%
0%
0%
20%
40%
60%
Noise factor
Probability mass
d0 (main diagonal)
d1
d2
D
Smart-Seq
Genes
Reactions
ε
Regular
Noised
No noise
E
Smart-Seq
Genes
Reactions
Regular
Accuracy
Noise factor
Probability mass
d0 (main diagonal)
d1
d2

Figure S2. Algorithm overview, related to Figure 2.

(A) Cumulative distribution function (CDF) of number of reactions per meta-reaction. (B-C) Random noise ε was added to the input gene expression matrix with two transcription noise models ("symmetric" and "regular") as described in the Supplementary Methods. (B) Left column: expression for every gene was scaled to [0,1] range. Two-dimensional density was computed over the of original and noised expression per gene is shown; right column: similarly, compass scores for every reaction were scaled to [0,1] and 2d density for raw and noised scores is shown. (C) Total probability mass under the 2d density surface was computed for (d0) the main diagonal; (d1) the two diagonals above and below the main diagonal; (d2) similarly, the two diagonals above and below the d1 diagonals.

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
$-\log_{10}$(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.)

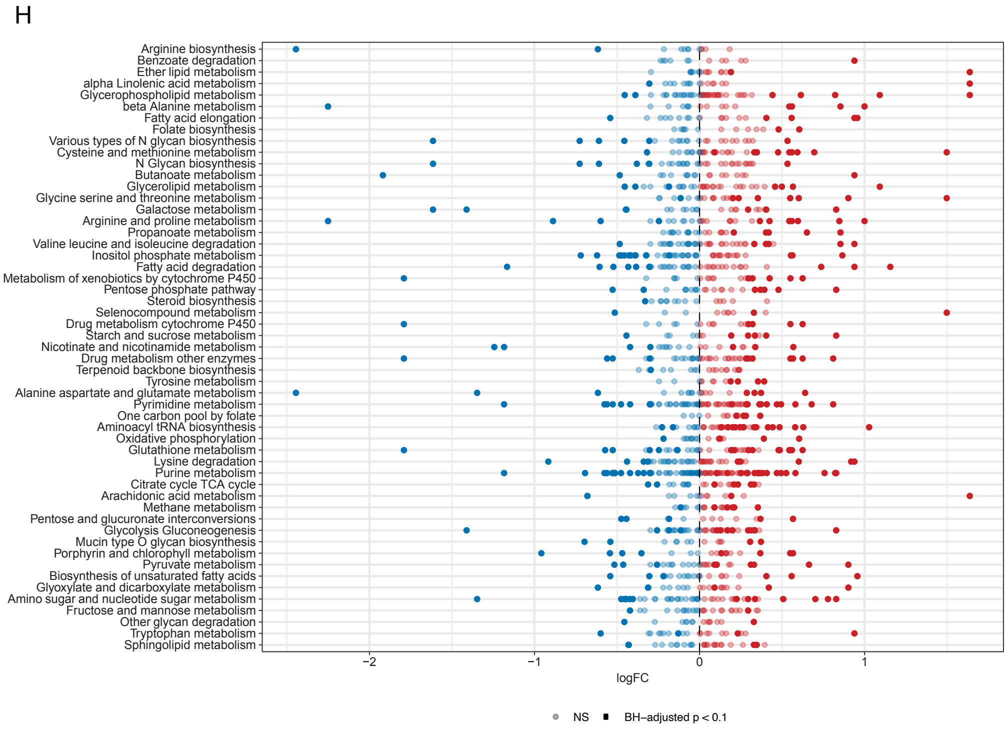

H
Arginine biosynthesis
Benzoate degradation
Ether lipid metabolism
alpha Linolenic acid metabolism
Glycerophospholipid metabolism
beta Alanine metabolism
Fatty acid elongation
Folate biosynthesis
Various types of N glycan biosynthesis
Cysteine and methionine metabolism
N Glycan biosynthesis
Butanoate 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

Figure S3. Compass-based exploration of metabolic heterogeneity within the Th17 compartment, related to Figure 3.

(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 (Subramanian et al., 2005) 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 **(Supplementary 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 **Supplementary Methods** for signature computation. (F) Same analysis as shown in Figure 3c, 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 3D** and **Supplementary Methods**). Similarly, we computed pro-regulatory reaction scores. Only core reactions are shown. (H) Same analysis as shown in Figure 3E, 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).

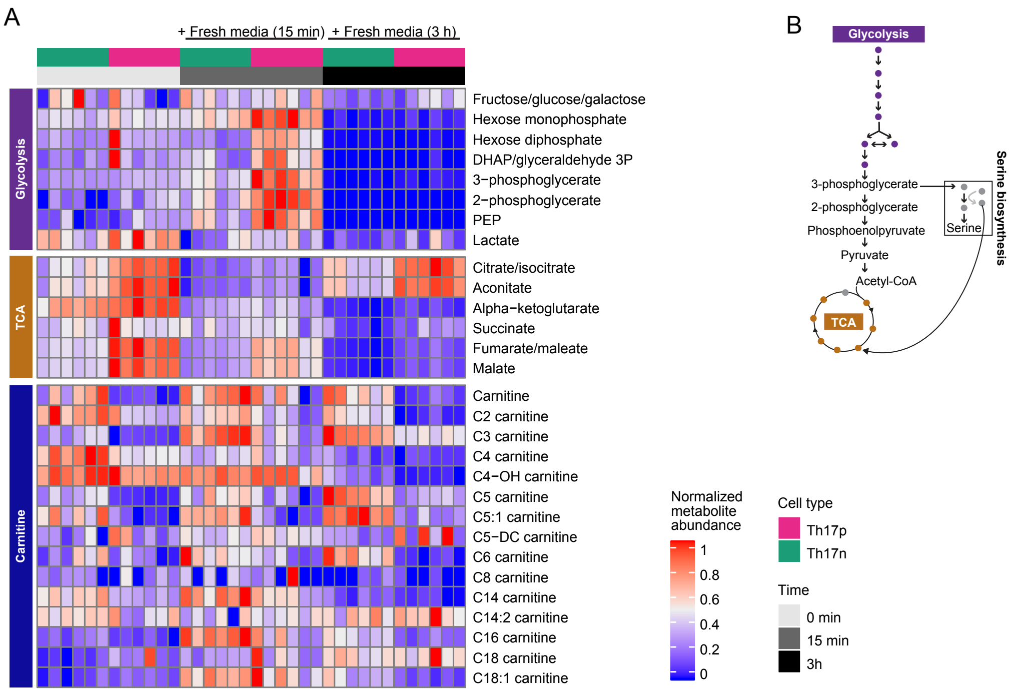

A
+ Fresh media (15 min)
+ Fresh media (3 h)
Glycolysis
TCA
Carnitine
Fructose/glucose/galactose
Hexose monophosphate
Hexose diphosphate
DHAP/glyceraldehyde 3P
3−phosphoglycerate
2−phosphoglycerate
PEP
Lactate
Citrate/isocitrate
Aconitate
Alpha−ketoglutarate
Succinate
Fumarate/maleate
Malate
Carnitine
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
Acetyl-CoA
TCA
Serine
Serine biosynthesis

Figure S4. Differential usage of glycolysis and fatty acid oxidation by pathogenic and non-pathogenic Th17 cells, related to Figure 4.

(A) Parallel of main Figure 4C showing also 3h after fresh media pulse. (B) The glycolysis pathway, as shown in main Figure 5A, highlighting its junction with serine biosynthesis.

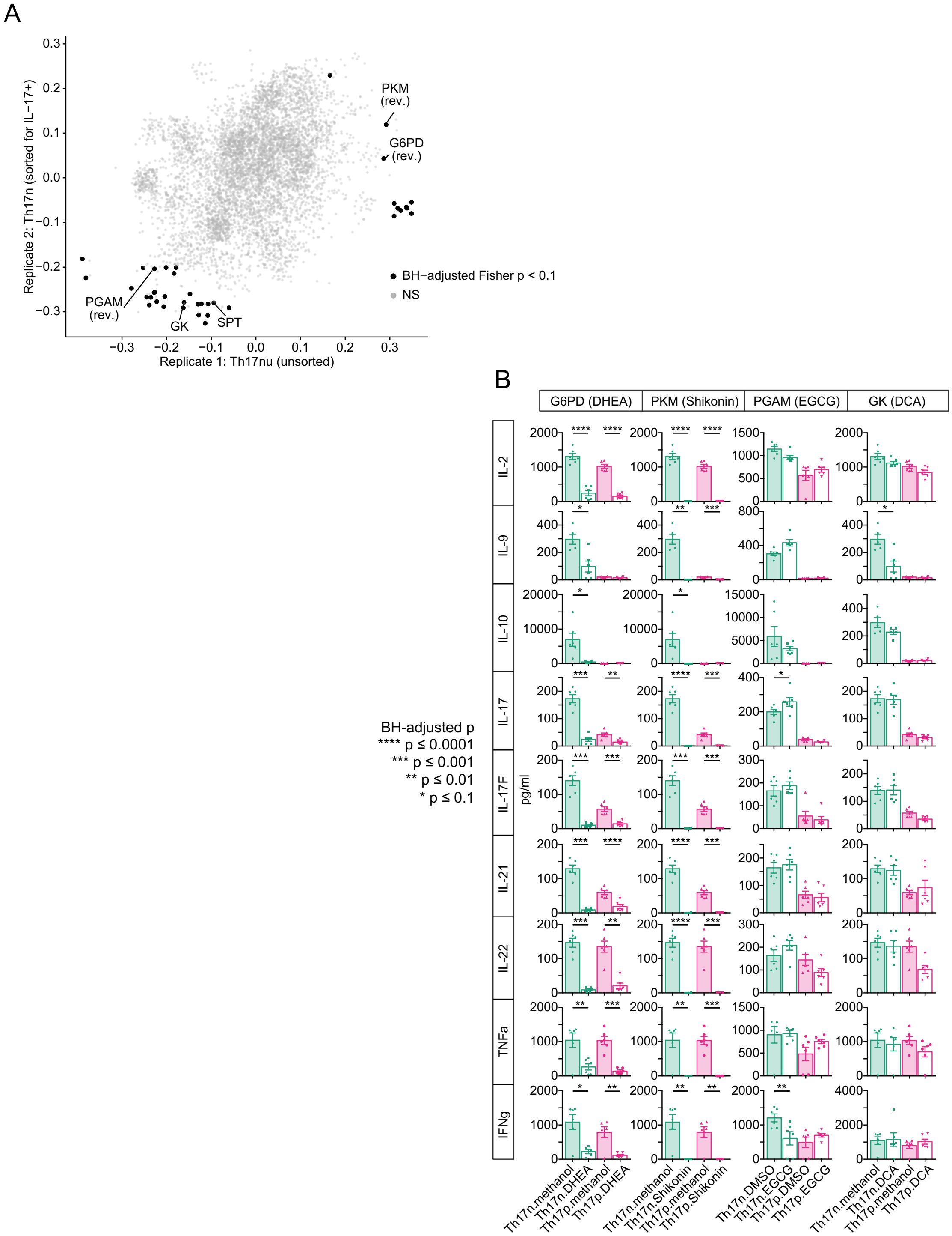

A
Replicate 2: Th17n (sorted for IL−17+)
Replicate 1: Th17nu (unsorted)
PKM (rev.)
G6PD (rev.)
PGAM (rev.)
GK
SPT
BH−adjusted Fisher p < 0.1
NS
B
G6PD (DHEA)
PKM (Shikonin)
PGAM (EGCG)
GK (DCA)
IL-2
IL-9
IL-10
IL-17
IL-17F
IL-21
IL-22
TNFa
IFNg
pg/ml
BH-adjusted p
**** p ≤ 0.0001
*** p ≤ 0.001
** p ≤ 0.01
* p ≤ 0.1
Th17n.methanol
Th17n.DHEA
Th17p.methanol
Th17p.DHEA
Th17n.Shikonin
Th17p.Shikonin
Th17n.DMSO
Th17n.EGCG
Th17p.DMSO
Th17p.EGCG
Th17n.DCA
Th17p.DCA

Figure S5. An unexpected role for PGAM in mediating Th17 pathogenicity, related to Figure 5.

(A) Same data as shown in Figure 5a, highlighting the reactions with significant adjusted Fisher p value in the intra-population analysis **(B)** Supernatant from Th17 cell cultures performed for Figure 5C are harvested for cytokine analysis using Legendplex.

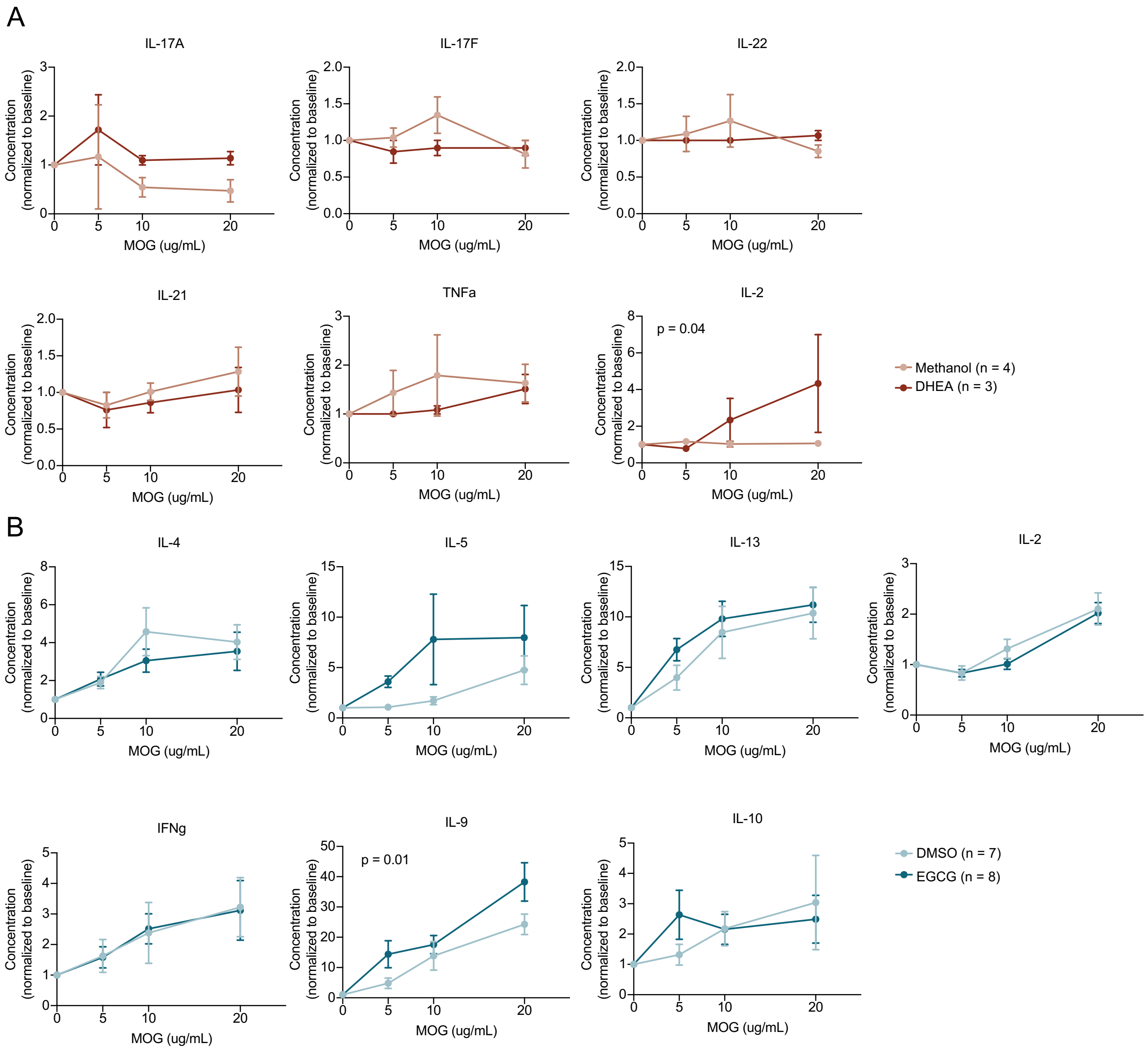

A
IL-17A
IL-17F
IL-22
IL-21
TNFa
IL-2
p = 0.04
Concentration (normalized to baseline)
MOG (ug/mL)
Methanol (n = 4)
DHEA (n = 3)
B
IL-4
IL-5
IL-13
IL-2
IFNg
IL-9
p = 0.01
IL-10
DMSO (n = 7)
EGCG (n = 8)

Figure S6. EGCG exacerbates and DHEA ameliorates Th17-induced EAE *in vivo*, related to Figure 6.

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

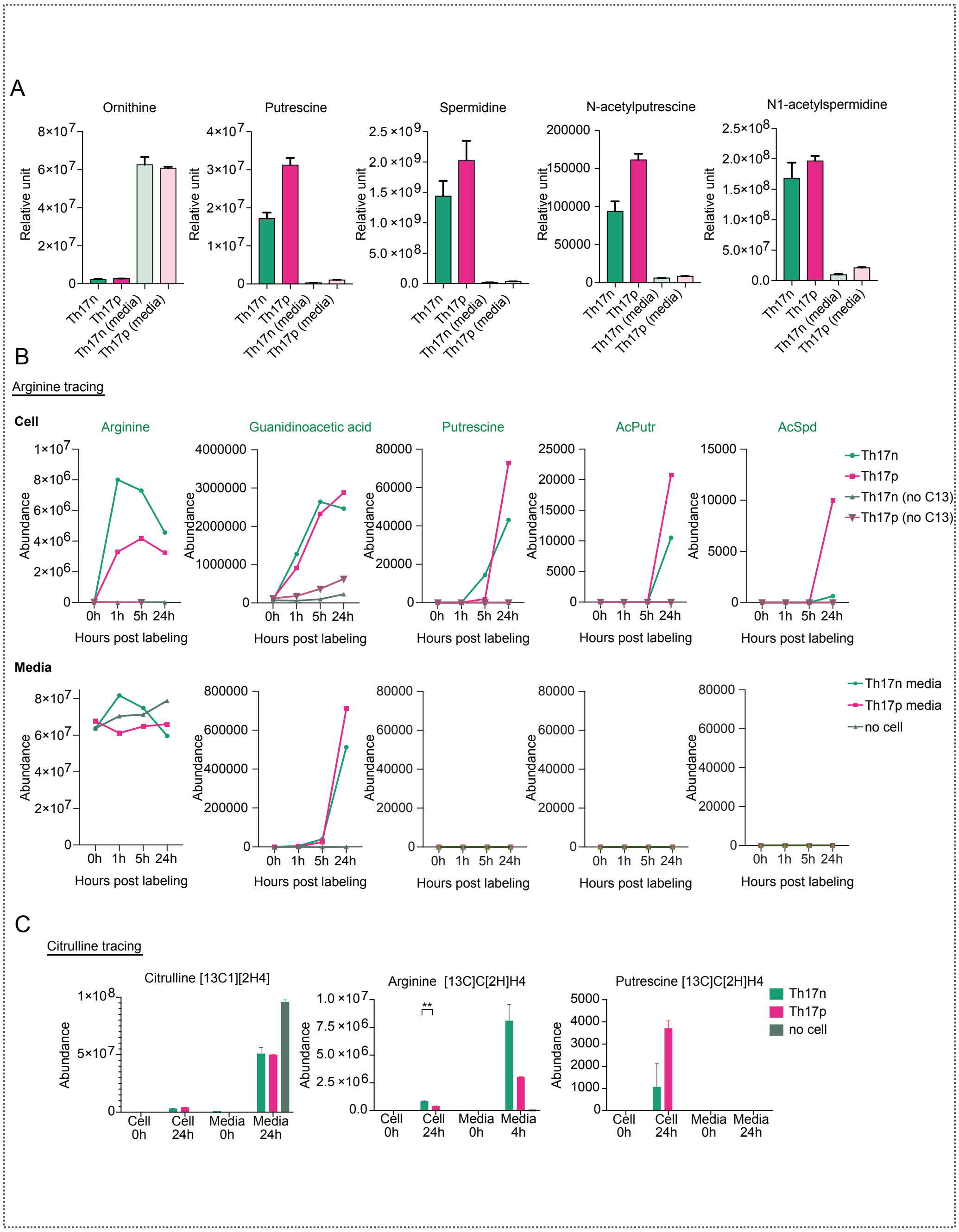
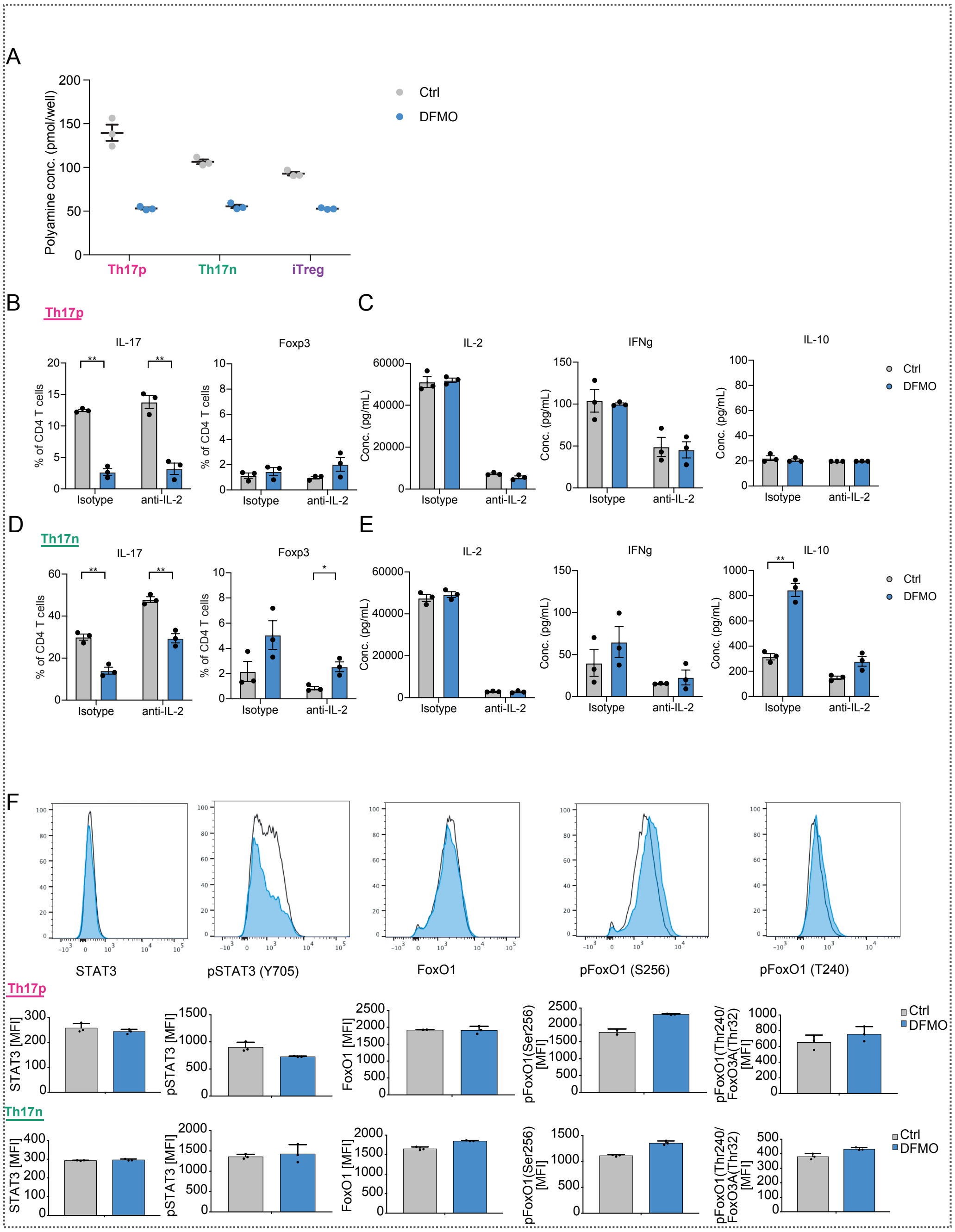

A
Ornithine
Putrescine
Spermidine
N-acetylputrescine
N1-acetylspermidine
Relative unit
Th17n
Th17p
Th17n (media)
Th17p (media)
8×10^7
6×10^7
4×10^7
2×10^7
0
4×10^7
3×10^7
2×10^7
1×10^7
2.5×10^9
2.0×10^9
1.5×10^9
1.0×10^9
5.0×10^8
0.0
200000
150000
100000
50000
2.5×10^8
2.0×10^8
1.5×10^8
1.0×10^8
5.0×10^7
B
Arginine tracing
Cell
Arginine
Guanidinoacetic acid
Putrescine
AcPutr
AcSpd
Abundance
Hours post labeling
0h
1h
5h
24h
1×10^7
8×10^6
6×10^6
4×10^6
2×10^6
4000000
3000000
2000000
1000000
80000
60000
40000
20000
25000
20000
15000
10000
5000
15000
10000
5000
Th17n
Th17p
Th17n (no C13)
Th17p (no C13)
Media
8×10^7
6×10^7
4×10^7
2×10^7
800000
600000
400000
200000
Th17n media
Th17p media
no cell
C
Citrulline tracing
Citrulline [13C1][2H4]
Arginine [13C]C[2H]H4
Putrescine [13C]C[2H]H4
1×10^8
5×10^7
1.0×10^7
7.5×10^6
5.0×10^6
2.5×10^6
5000
4000
3000
2000
1000
**
Cell 0h
Cell 24h
Media 0h
Media 24h
Media 4h
Th17n
Th17p
no cell

Figure S7. Prediction and metabolic validation of the polyamine pathway as a candidate in regulating Th17 cell function, related to Figure 7.

(**A**) Metabolomics analysis of the polyamine pathway as in Figure 2H. Cell lysates as well as media from Th17n and Th17p differentiation cultures are shown. (**B**-**C**) Carbon tracing in the polyamine pathway. Th17n and Th17p cells were differentiated as described (STAR Methods), lifted to rest at 68 hours and pulsed with C13 labeled Arginine (B) or Citrulline (C) followed by LC/MS analysis at time points indicated.


A
Polyamine conc. (pmol/well)
200
150
100
50
0
Th17p
Th17n
iTreg
Ctrl
DFMO
B
Th17p
IL-17
% of CD4 T cells
Foxp3
Isotype
anti-IL-2
**
C
IL-2
IFNg
IL-10
Conc. (pg/mL)
Ctrl
DFMO
D
Th17n
IL-17
Foxp3
*
E
IL-2
IFNg
IL-10
F
STAT3
pSTAT3 (Y705)
FoxO1
pFoxO1 (S256)
pFoxO1 (T240)
Th17p
STAT3 [MFI]
pSTAT3 [MFI]
FoxO1 [MFI]
pFoxO1(Ser256) [MFI]
pFoxO1(Thr240/ FoxO3A(Thr32) [MFI]
Ctrl
DFMO
Th17n

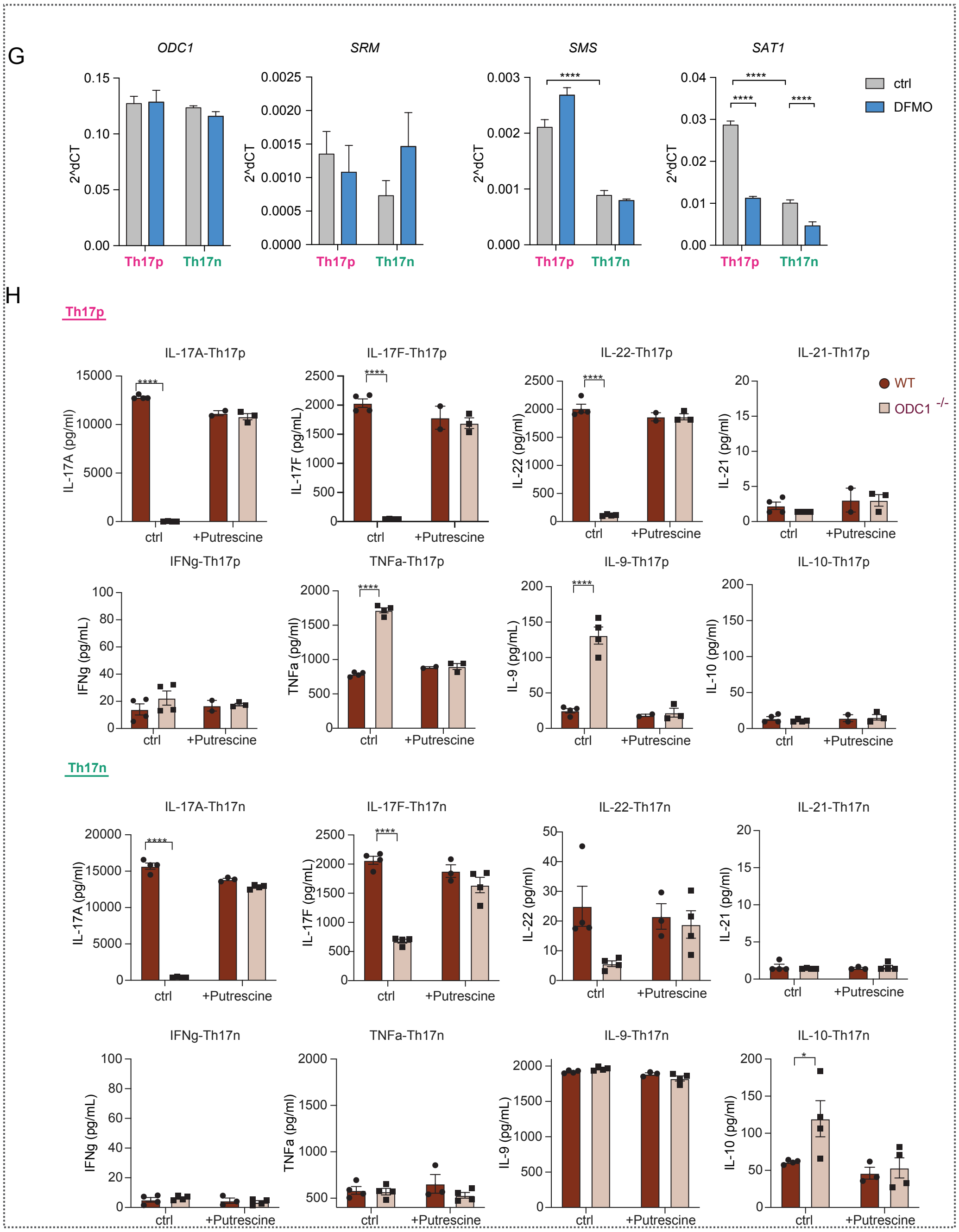

G
ODC1
SRM
SMS
SAT1
2^dCT
0.15
0.10
0.05
0.00
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0.003
0.002
0.001
0.000
0.04
0.03
0.02
0.01
0.00
****
Th17p
Th17n
ctrl
DFMO
H
Th17p
IL-17A-Th17p
IL-17F-Th17p
IL-22-Th17p
IL-21-Th17p
IL-17A (pg/ml)
IL-17F (pg/mL)
IL-22 (pg/ml)
IL-21 (pg/ml)
15000
10000
5000
0
2500
2000
1500
1000
500
20
15
10
5
WT
ODC1 -/-
ctrl
+Putrescine
IFNg-Th17p
TNFa-Th17p
IL-9-Th17p
IL-10-Th17p
IFNg (pg/mL)
TNFa (pg/ml)
IL-9 (pg/mL)
IL-10 (pg/ml)
100
80
60
40
200
150
50
Th17n
IL-17A-Th17n
IL-17F-Th17n
IL-22-Th17n
IL-21-Th17n
20000
30
IFNg-Th17n
TNFa-Th17n
IL-9-Th17n
IL-10-Th17n
*

Figure S8. Chemical and genetic interference with the polyamine pathway suppress canonical Th17 cell cytokines, related to Figure 8.

(**A**) The effect of DFMO on cellular polyamine concentration is measured by an enzymatic assay. Th17p, Th17n and iTregs are differentiated in the presence of DFMO and harvested at 96 hours for analysis. (**B-E**) IL-2 neutralization does not regulate the effect of ODC1 inhibition. 10ug/ml of anti-IL-2 antibody or isotype control were added at the time of Th17n or Th17p cell differentiation with control or DFMO. Cells were analyzed by **B,D,** flow cytometry and **C,E,** supernatant were analyzed for cytokine secretion. Welch t-test significance is denoted. (**F**) Protein and phospho-protein analysis by flow cytometry for Th17n and Th17p cells treated with control of DFMO. (**G**) The effect of DFMO on enzymes in the polyamine pathway as measured by qPCR. Th17p and Th17n cells were differentiated in the presence of control or DFMO and harvested at 48h for RNA extraction and qPCR analysis. (**H**) The effect of genetic perturbation of ODC1 on cytokine production from Th17p (upper panels) and Th17n cells (lower panels). Supernatant from Th17p and Th17n differentiation culture was harvested at 96 hours and analyzed by legendplex for cytokine concentration.

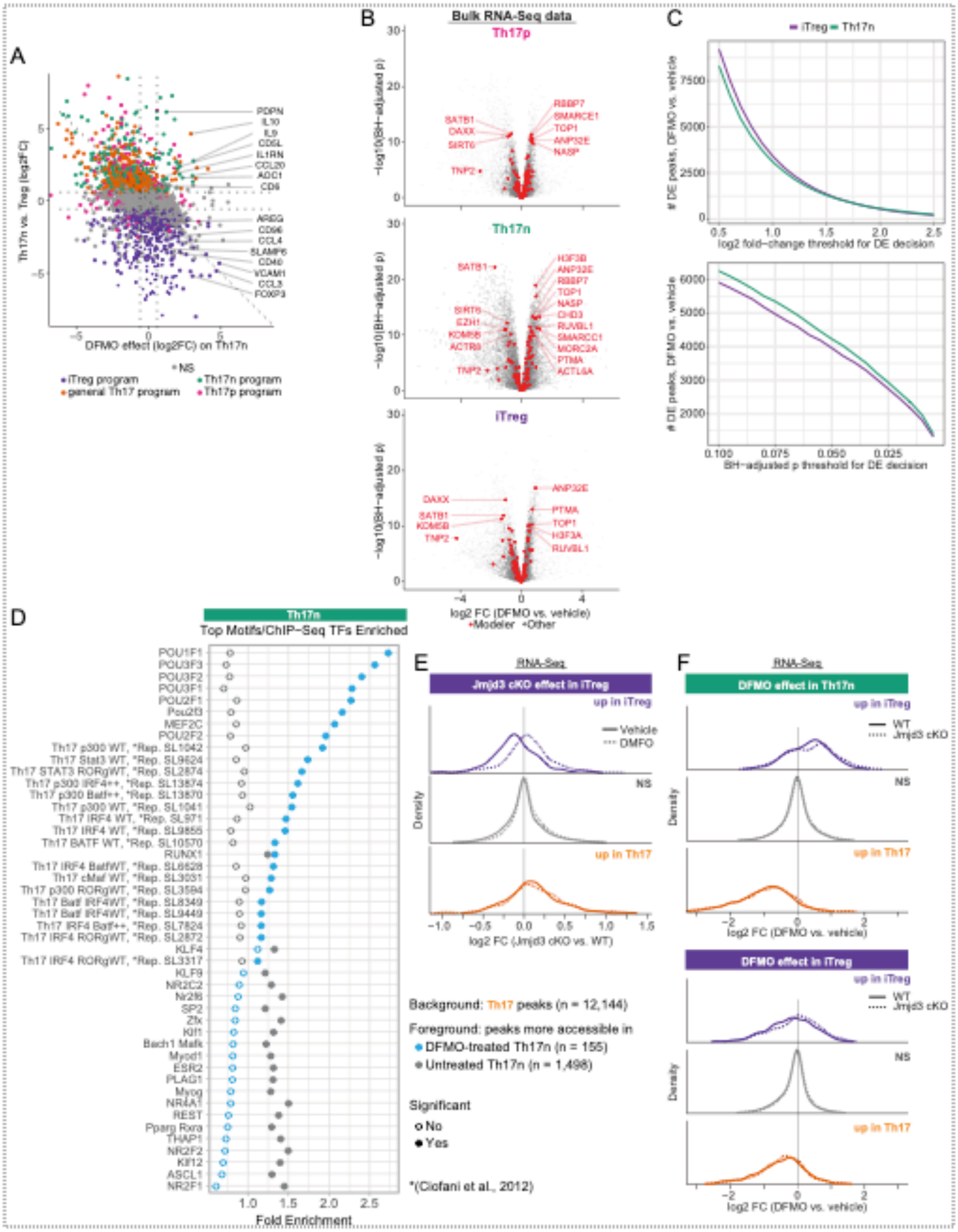

A
PDPN
IL10
IL9
CD5L
IL1RN
CCL20
ADC1
CD6
AREG
CD96
CCL4
SLAMF6
CD40
VCAM1
CCL3
FOXP3
5
0
−5
Th17n vs. Treg (log2FC)
−5
0
5
DFMO effect (log2FC) on Th17n
NS
iTreg program
Th17n program
general Th17 program
Th17p program
B
Bulk RNA-Seq data
Th17p
SATB1
DAXX
SIRT6
TNP2
RBBP7
SMARCE1
TOP1
ANP32E
NASP
Th17n
SATB1
SIRT6
EZH1
KDM5B
ACTR8
TNP2
H3F3B
ANP32E
RBBP7
TOP1
NASP
CHD3
RUVBL1
SMARCC1
MORC2A
PTMA
ACTL6A
iTreg
DAXX
SATB1
KDM5B
TNP2
ANP32E
PTMA
TOP1
H3F3A
RUVBL1
30
20
10
0
−log10(BH−adjusted p)
−4
0
4
log2 FC (DFMO vs. vehicle)
+Modeler +Other
C
iTreg
Th17n
7500
5000
2500
0
DE peaks, DFMO vs. vehicle
0.5
1.0
1.5
2.0
2.5
log2 fold−change threshold for DE decision
6000
5000
4000
3000
2000
DE peaks, DFMO vs. vehicle
0.100
0.075
0.050
0.025
BH−adjusted p threshold for DE decision
D
Th17n
Top Motifs/ChIP−Seq TFs Enriched
POU1F1
POU3F3
POU3F2
POU3F1
POU2F1
Pou2f3
MEF2C
POU2F2
Th17 p300 WT, *Rep. SL1042
Th17 Stat3 WT, *Rep. SL9624
Th17 STAT3 RORgWT, *Rep. SL2874
Th17 p300 IRF4++, *Rep. SL13874
Th17 p300 Batf++, *Rep. SL13870
Th17 p300 WT, *Rep. SL1041
Th17 IRF4 WT, *Rep. SL971
Th17 IRF4 WT, *Rep. SL9855
Th17 BATF WT, *Rep. SL10570
RUNX1
Th17 IRF4 BatfWT, *Rep. SL6628
Th17 cMaf WT, *Rep. SL3031
Th17 p300 RORgWT, *Rep. SL3594
Th17 Batf IRF4WT, *Rep. SL8349
Th17 Batf IRF4WT, *Rep. SL9449
Th17 IRF4 Batf++, *Rep. SL7824
Th17 IRF4 RORgWT, *Rep. SL2872
KLF4
Th17 IRF4 RORgWT, *Rep. SL3317
KLF9
NR2C2
Nr2f6
SP2
Zfx
Klf1
Bach1 Mafk
Myod1
ESR2
PLAG1
Myog
NR4A1
REST
Pparg Rxra
THAP1
NR2F2
Klf12
ASCL1
NR2F1
1.0
1.5
2.0
2.5
Fold Enrichment
E
RNA-Seq
Jmjd3 cKO effect in iTreg
up in iTreg
Vehicle
DMFO
NS
up in Th17
Density
−1.0
−0.5
0.0
0.5
1.0
1.5
log2 FC (Jmjd3 cKO vs. WT)
Background: Th17 peaks (n = 12,144)
Foreground: peaks more accessible in
DFMO-treated Th17n (n = 155)
Untreated Th17n (n = 1,498)
Significant
No
Yes
*(Ciofani et al., 2012)
F
RNA-Seq
DFMO effect in Th17n
up in iTreg
WT
Jmjd3 cKO
NS
up in Th17
Density
−2
0
2
log2 FC (DFMO vs. vehicle)
DFMO effect in iTreg
up in iTreg
WT
Jmjd3 cKO
NS
up in Th17
Density
−2
0
2
log2 FC (DFMO vs. vehicle)

Figure S9. DFMO treatment promotes Treg-like transcriptome and epigenome, related to Figure 9.

(A) Log2 fold changes for a comparison between DFMO-treated and vehicle-treated Th17n cells (x-axis), and Th17n vs. iTreg cells (both vehicle-treated; y-axis) are based on the same data described in manuscript Figure 9A-C. Dots represent genes, which are divided into 5 groups based on differential expression in untreated cells (**Supplementary Methods**). Dashed lines correspond to the logFC threshold used in differential expression calling log2(1.5) (**Supplementary Methods**). Differentially expressed genes by either of the comparisons shown in the axes are opaque and transparent otherwise. **(B)** Volcano plots showing affected chromatin modifiers by DFMO treatment in Th17n, Th17p and iTreg cells. **(C)** Number of differentially expressed (DE) peaks between DFMO and vehicle-treated cells as a function of the significance threshold. Upper panel, log2FC used as threshold; Lower panel, BH-adjusted P used as threshold. **(D)** Similar analysis to Figure 9F, only using Th17 background instead of Treg. I.e., dots represent fold enrichment of peaks more accessible in DFMO-treated (blue) or untreated (grey) Th17n cells against a background of Th17 peaks as described in the manuscript (i.e., this analysis considers only peaks that were differentially more accessible in untreated Th17 compared to untreated iTreg, corresponding to the orange curve in Figure 9D). **(E-F)** Cells were cultured under Th17n or iTreg condition with DFMO or solvent control (water) as in Figure 9 and harvested at 68h for RNA-Seq. Treg and Th17 programs were defined as in **Figure 9H (Supplementary Methods)**. **E,** Histogram showing the effects of JMJD3 conditional knockout (cKO) in control vs DFMO-treated Treg cells. **F,** Histogram showing the effects of DFMO in WT vs JMJD3$^{fl/fl}$CD4^{cre} in Th17n and Treg cells.

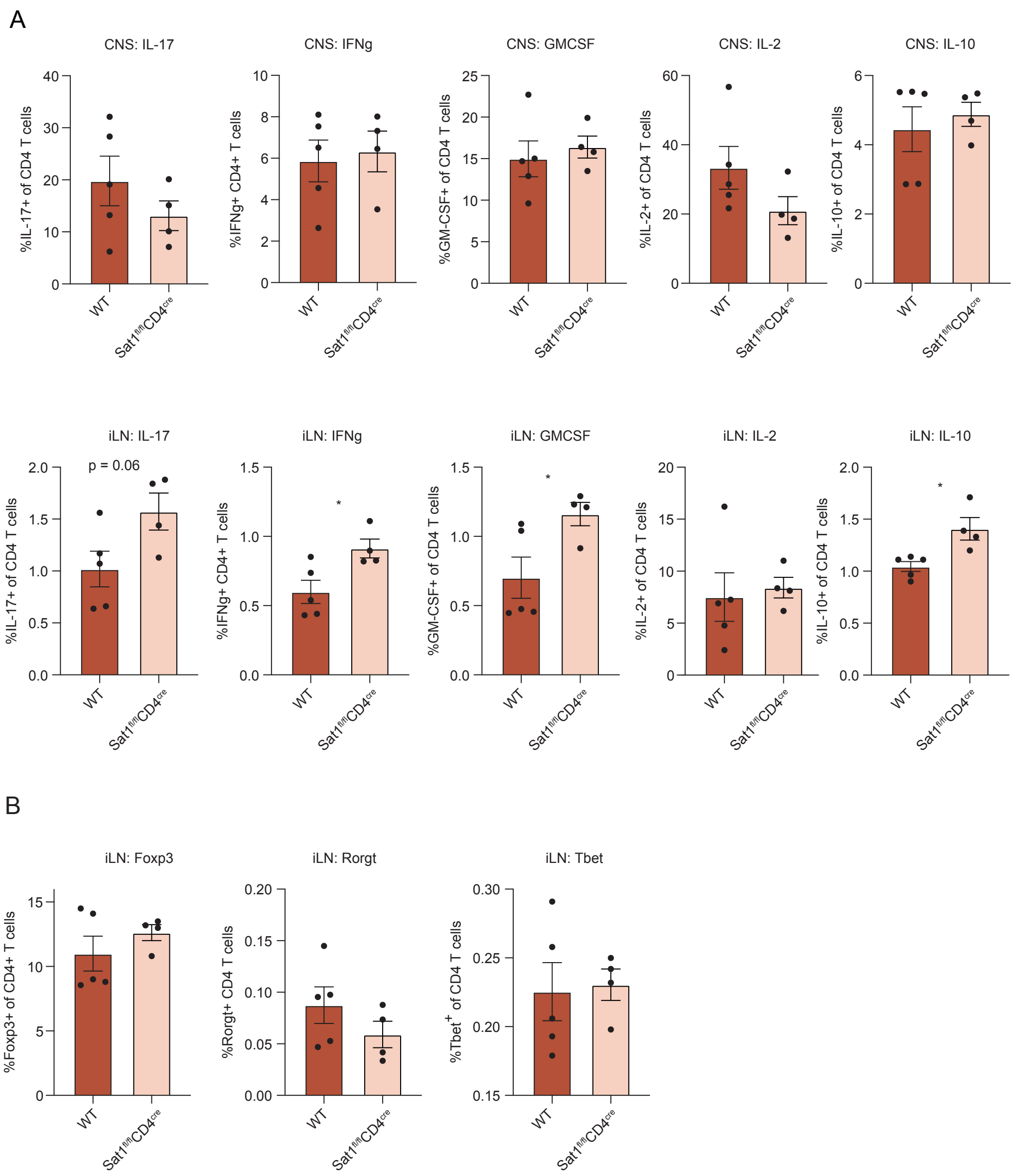

A
CNS: IL-17
%IL-17+ of CD4 T cells
CNS: IFNg
%IFNg+ CD4+ T cells
CNS: GMCSF
%GM-CSF+ of CD4 T cells
CNS: IL-2
%IL-2+ of CD4 T cells
CNS: IL-10
%IL-10+ of CD4 T cells
iLN: IL-17
p = 0.06
%IL-17+ of CD4 T cells
iLN: IFNg
*
%IFNg+ CD4+ T cells
iLN: GMCSF
*
%GM-CSF+ of CD4 T cells
iLN: IL-2
%IL-2+ of CD4 T cells
iLN: IL-10
*
%IL-10+ of CD4 T cells
B
iLN: Foxp3
%Foxp3+ of CD4+ T cells
iLN: Rorgt
%Rorgt+ CD4 T cells
iLN: Tbet
%Tbet+ of CD4 T cells
WT
Sat1fl/flCD4cre

Figure S10. Targeting ODC1 and SAT1 alleviate EAE, related to Figure 10.

Cells were isolated from CNS or inguinal lymph node (iLN) of WT or SAT1$^{fl/fl}$CD4^{cre} mice on day 15 post EAE induction (similar experiments as in Figure 7F). (**A**) Intracellular cytokines were measured by flow cytometry after 4-hour PMA/ionomycine stimulation ex vivo in the presence of brefaldin and monensin. (**B**) Transcription factors were analyzed directly ex vivo by intracellular staining.

References

Abbas, A.K., Benoist, C., Bluestone, J.A., Campbell, D.J., Ghosh, S., Hori, S., Jiang, S., Kuchroo, V.K., Mathis, D., Roncarolo, M.G., et al. (2013). Regulatory T cells: recommendations to simplify the nomenclature. Nat. Immunol. *14*, 307–308.

Adler, M., Korem Kohanim, Y., Tendler, A., Mayo, A., and Alon, U. (2019). Continuum of Gene-Expression Profiles Provides Spatial Division of Labor within a Differentiated Cell Type. Cell Syst *8*, 43-52.e5.

Aggarwal, S., Ghilardi, N., Xie, M.-H., de Sauvage, F.J., and Gurney, A.L. (2003). Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17. J. Biol. Chem. *278*, 1910–1914.

Aldridge, S., and Teichmann, S.A. (2020). Single cell transcriptomics comes of age. Nat. Commun. *11*, 4307.

Alper, H., Jin, Y.-S., Moxley, J.F., and Stephanopoulos, G. (2005). Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli. Metab. Eng. *7*, 155–164.

Altschuler, S.J., and Wu, L.F. (2010). Cellular Heterogeneity: Do Differences Make a Difference? Cell *141*, 559–563.

Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X., Krueger, F., Smallwood, S.A., Ponting, C.P., Voet, T., et al. (2016). Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat. Methods *13*, 229–232.

Antebi, Y.E., Reich-Zeliger, S., Hart, Y., Mayo, A., Eizenberg, I., Rimer, J., Putheti, P., Pe'er, D., and Friedman, N. (2013). Mapping Differentiation under Mixed Culture Conditions Reveals a Tunable Continuum of T Cell Fates. PLoS Biol. *11*, e1001616.

Antoniewicz, M.R. (2013). Dynamic metabolic flux analysis--tools for probing transient states of metabolic networks. Curr. Opin. Biotechnol. *24*, 973–978.

Argelaguet, R., Cuomo, A.S.E., Stegle, O., and Marioni, J.C. (2021). Computational principles and challenges in single-cell data integration. Nat. Biotechnol. 1–14.

Awasthi, A., Riol-Blanco, L., Jäger, A., Korn, T., Pot, C., Galileos, G., Bettelli, E., Kuchroo, V.K., and Oukka, M. (2009). Cutting edge: IL-23 receptor gfp reporter mice reveal distinct populations of IL-17-producing cells. J. Immunol. *182*, 5904–5908.

Bandura, D.R., Baranov, V.I., Ornatsky, O.I., Antonov, A., Kinach, R., Lou, X., Pavlov, S., Vorobiev, S., Dick, J.E., and Tanner, S.D. (2009). Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. Anal. Chem. *81*, 6813–6822.

Baran, Y., Bercovich, A., Sebe-Pedros, A., Lubling, Y., Giladi, A., Chomsky, E., Meir, Z., Hoichman, M., Lifshitz, A., and Tanay, A. (2019). MetaCell: analysis of single-cell RNA-seq data using K-nn graph partitions. Genome Biol. *20*, 206.

Barbi, J., Pardoll, D., and Pan, F. (2013). Metabolic control of the Treg/Th17 axis. Immunol. Rev. *252*, 52–77.

Becker, S.A., and Palsson, B.O. (2008). Context-Specific Metabolic Networks Are Consistent with Experiments. PLoS Comput. Biol. *4*, e1000082.

Bendall, S.C., Simonds, E.F., Qiu, P., Amir, E.-A.D., Krutzik, P.O., Finck, R., Bruggner, R.V., Melamed, R., Trejo, A., Ornatsky, O.I., et al. (2011). Single-Cell Mass Cytometry of Differential Immune and Drug Responses Across a Human Hematopoietic Continuum. Science *332*, 687–696.

Ben-Moshe, S., Shapira, Y., Moor, A.E., Manco, R., Veg, T., Bahar Halpern, K., and Itzkovitz, S. (2019). Spatial sorting enables comprehensive characterization of liver zonation. Nat Metab *1*, 899–911.

Berod, L., Friedrich, C., Nandan, A., Freitag, J., Hagemann, S., Harmrolfs, K., Sandouk, A., Hesse, C., Castro, C.N., Bahre, H., et al. (2014). De novo fatty acid synthesis controls the fate between regulatory T and T helper 17 cells. Nat. Med. *20*, 1327–1333.

Bettelli, E., Carrier, Y., Gao, W., Korn, T., Strom, T.B., Oukka, M., Weiner, H.L., and Kuchroo, V.K. (2006). Reciprocal developmental pathways for the generation of pathogenic effector TH17 and regulatory T cells. Nature *441*, 235–238.

Bettelli, E., Korn, T., Oukka, M., and Kuchroo, V.K. (2008). Induction and effector functions of T(H)17 cells. Nature *453*, 1051–1057.

Bodenmiller, B., Zunder, E.R., Finck, R., Chen, T.J., Savig, E.S., Bruggner, R.V., Simonds, E.F., Bendall, S.C., Sachs, K., Krutzik, P.O., et al. (2012). Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. Nat. Biotechnol. *30*, 858–867.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bordbar, A., Monk, J.M., King, Z.A., and Palsson, B.O. (2014). Constraint-based models predict metabolic and associated cellular functions. Nat. Rev. Genet. *15*, 107–120.

Bowlin, T.L., McKown, B.J., and Sunkara, P.S. (1987). The effect of alpha-difluoromethylornithine, an inhibitor of polyamine biosynthesis, on mitogen-induced interleukin 2 production. Immunopharmacology *13*, 143–147.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

Brooks, W.H. (2013). Increased polyamines alter chromatin and stabilize autoantigens in autoimmune diseases. Front. Immunol. *4*, 91.

Buck, M.D., Sowell, R.T., Kaech, S.M., and Pearce, E.L. (2017). Metabolic Instruction of Immunity. Cell *169*, 570–586.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature *523*, 486–490.

Buettner, F., Natarajan, K.N., Casale, F.P., Proserpio, V., Scialdone, A., Theis, F.J., Teichmann, S.A., Marioni, J.C., and Stegle, O. (2015). Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat. Biotechnol. *33*, 155–160.

Burgard, A.P., Pharkya, P., and Maranas, C.D. (2003). Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. *84*, 647–657.

Casero, R.A., Jr, Murray Stewart, T., and Pegg, A.E. (2018). Polyamine metabolism and cancer: treatments, challenges and opportunities. Nat. Rev. Cancer *18*, 681–695.

Certo, M., Tsai, C.-H., Pucino, V., Ho, P.-C., and Mauro, C. (2020). Lactate modulation of immune responses in inflammatory versus tumour microenvironments. Nat. Rev. Immunol.

Chandel, N. (2014). Navigating Metabolism (New York, NY: Cold Spring Harbor Laboratory Press).

Chang, C.-H., Curtis, J.D., Maggi, L.B., Jr., Faubert, B., Villarino, A.V., O'Sullivan, D., Huang, S.C.-C., van der Windt, G.J.W., Blagih, J., Qiu, J., et al. (2013). Posttranscriptional Control of T Cell Effector Function by Aerobic Glycolysis. Cell *153*, 1239–1251.

Chapman, N.M., Boothby, M.R., and Chi, H. (2019). Metabolic coordination of T cell quiescence and activation. Nat. Rev. Immunol.

Chattopadhyay, P.K., Price, D.A., Harper, T.F., Betts, M.R., Yu, J., Gostick, E., Perfetto, S.P., Goepfert, P., Koup, R.A., De Rosa, S.C., et al. (2006). Quantum dot semiconductor nanocrystals for immunophenotyping by polychromatic flow cytometry. Nat. Med. *12*, 972–977.

Chen, K.H., Boettiger, A.N., Moffitt, J.R., Wang, S., and Zhuang, X. (2015). Spatially resolved, highly multiplexed RNA profiling in single cells. Science *348*.

Cho, S., Lee, G., Pickering, B.F., Jang, C., Park, J.H., He, L., Mathur, L., Kim, S.-S., Jung, S., Tang, H.-W., et al. (2021). mTORC1 promotes cell growth via m6A-dependent mRNA degradation. Mol. Cell.

Choi, J.-E., Sebastian, C., Ferrer, C.M., Lewis, C.A., Sade-Feldman, M., LaSalle, T., Gonye, A., Lopez, B.G.C., Abdelmoula, W.M., Regan, M.S., et al. (2021). A unique subset of glycolytic tumour-propagating cells drives squamous cell carcinoma. Nat Metab *3*, 182–195.

Chung, Y., Chang, S.H., Martinez, G.J., Yang, X.O., Nurieva, R., Kang, H.S., Ma, L., Watowich, S.S., Jetten, A.M., Tian, Q., et al. (2009). Critical regulation of early Th17 cell differentiation by interleukin-1 signaling. Immunity *30*, 576–587.

Ciofani, M., Madar, A., Galan, C., Sellars, M., Mace, K., Pauli, F., Agarwal, A., Huang, W., Parkhurst, C.N., Muratet, M., et al. (2012). A validated regulatory network for Th17 cell specification. Cell *151*, 289–303.

Cole, M.B., Risso, D., Wagner, A., DeTomaso, D., Ngai, J., Purdom, E., Dudoit, S., and Yosef, N. (2017). Performance Assessment and Selection of Normalization Procedures for Single-Cell RNA-Seq.

Conti, H.R., Peterson, A.C., Brane, L., Huppler, A.R., Hernández-Santos, N., Whibley, N., Garg, A.V., Simpson-Abelson, M.R., Gibson, G.A., Mamo, A.J., et al. (2014). Oral-resident natural Th17 cells and γδ T cells control opportunistic Candida albicans infections. J. Exp. Med. *211*, 2075–2084.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat. Genet. *48*, 1193–1203.

Cua, D.J., Sherlock, J., Chen, Y., Murphy, C.A., Joyce, B., Seymour, B., Lucian, L., To, W., Kwan, S., Churakova, T., et al. (2003). Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. Nature *421*, 744–748.

Culemann, S., Grüneboom, A., Nicolás-Ávila, J.Á., Weidner, D., Lämmle, K.F., Rothe, T., Quintana, J.A., Kirchner, P., Krljanac, B., Eberhardt, M., et al. (2019). Locally renewing resident synovial macrophages provide a protective barrier for the joint. Nature *572*, 670–675.

Cusanovich, D.A., Daza, R., Adey, A., Pliner, H.A., Christiansen, L., Gunderson, K.L., Steemers, F.J., Trapnell, C., and Shendure, J. (2015). Multiplex single-cell profiling of chromatin accessibility by combinatorial cellular indexing. Science *348*, 910–914.

Damiani, C., Maspero, D., Di Filippo, M., Colombo, R., Pescini, D., Graudenzi, A., Westerhoff, H.V., Alberghina, L., Vanoni, M., and Mauri, G. (2019). Integration of single-cell RNA-seq data into population models to characterize cancer metabolism. PLoS Comput. Biol. *15*, e1006733.

Dang, L., White, D.W., Gross, S., Bennett, B.D., Bittinger, M.A., Driggers, E.M., Fantin, V.R., Jang, H.G., Jin, S., Keenan, M.C., et al. (2010). Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. Nature *465*, 966.

DeTomaso, D., Jones, M.G., Subramaniam, M., Ashuach, T., Ye, C.J., and Yosef, N. (2019). Functional interpretation of single cell similarity maps. Nat. Commun. *10*, 4376.

Dey, S.S., Kester, L., Spanjaard, B., Bienko, M., and van Oudenaarden, A. (2015). Integrated genome and transcriptome sequencing of the same cell. Nat. Biotechnol. *33*, 285–289.

van Dijk, D., Sharma, R., Nainys, J., Yim, K., Kathail, P., Carr, A.J., Burdziak, C., Moon, K.R., Chaffer, C.L., Pattabiraman, D., et al. (2018). Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. Cell *174*, 716-729.e27.

Diskin, C., Ryan, T.A.J., and O’Neill, L.A.J. (2021). Modification of Proteins by Metabolites in Immunity. Immunity *54*, 19–31.

Divakaruni, A.S., Hsieh, W.Y., Minarrieta, L., Duong, T.N., Kim, K.K.O., Desousa, B.R., Andreyev, A.Y., Bowman, C.E., Caradonna, K., Dranka, B.P., et al. (2018). Etomoxir Inhibits Macrophage Polarization by Disrupting CoA Homeostasis. Cell Metab. *28*, 490-503.e7.

Doedens, A.L., Phan, A.T., Stradner, M.H., Fujimoto, J.K., Nguyen, J.V., Yang, E., Johnson, R.S., and Goldrath, A.W. (2013). Hypoxia-inducible factors enhance the effector responses of CD8(+) T cells to persistent antigen. Nat. Immunol. *14*, 1173–1182.

Eisenstein, E.M., and Williams, C.B. (2009). The T(reg)/Th17 cell balance: a new paradigm for autoimmunity. Pediatr. Res. *65*, 26R-31R.

Elia, I., and Haigis, M.C. (2021). Metabolites and the tumour microenvironment: from cellular mechanisms to systemic metabolism. Nature Metabolism 1–12.

Eng, C.-H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.-C., et al. (2019). Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. Nature *568*, 235–239.

Farlik, M., Sheffield, N.C., Nuzzo, A., Datlinger, P., Schönegger, A., Klughammer, J., and Bock, C. (2015). Single-Cell DNA Methylome Sequencing and Bioinformatic Inference of Epigenomic Cell-State Dynamics. Cell Rep. *10*, 1386–1397.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., et al. (2010). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell *18*, 553–567.

Finlay, D.K., Rosenzweig, E., Sinclair, L.V., Feijoo-Carnero, C., Hukelmann, J.L., Rolf, J., Panteleyev, A.A., Okkenhaug, K., and Cantrell, D.A. (2012). PDK1 regulation of mTOR and hypoxia-inducible factor 1 integrate metabolism and migration of CD8+ T cells. J. Exp. Med. *209*, 2441–2453.

Gaffen, S.L., Hernandez-Santos, N., and Peterson, A.C. (2011). IL-17 signaling in host defense against Candida albicans. Immunol. Res. *50*, 181–187.

Gaublomme, J.T., Yosef, N., Lee, Y., Gertner, R.S., Yang, L.V., Wu, C., Pandolfi, P.P., Mak, T., Satija, R., Shalek, A.K., et al. (2015). Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity. Cell *163*, 1400–1412.

Geltink, R.I.K., Kyle, R.L., and Pearce, E.L. (2018). Unraveling the Complex Interplay Between T Cell Metabolism and Function. Annu. Rev. Immunol. *36*, 461–488.

Gemta, L.F., Siska, P.J., Nelson, M.E., Gao, X., Liu, X., Locasale, J.W., Yagita, H., Slingluff, C.L., Jr, Hoehn, K.L., Rathmell, J.C., et al. (2019). Impaired enolase 1 glycolytic activity restrains effector functions of tumor-infiltrating CD8+ T cells. Sci Immunol *4*.

Gerriets, V.A., Kishton, R.J., Nichols, A.G., Macintyre, A.N., Inoue, M., Ilkayeva, O., Winter, P.S., Liu, X., Priyadharshini, B., Slawinska, M.E., et al. (2015). Metabolic programming and PDHK1 control CD4+ T cell subsets and inflammation. J. Clin. Invest. *125*, 194–207.

Ghoreschi, K., Laurence, A., Yang, X.-P., Tato, C.M., McGeachy, M.J., Konkel, J.E., Ramos, H.L., Wei, L., Davidson, T.S., Bouladoux, N., et al. (2010). Generation of pathogenic T(H)17 cells in the absence of TGF-β signalling. Nature *467*, 967–971.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Grindberg, R.V., Yee-Greenbaum, J.L., McConnell, M.J., Novotny, M., O'Shaughnessy, A.L., Lambert, G.M., Araúzo-Bravo, M.J., Lee, J., Fishman, M., Robbins, G.E., et al. (2013). RNA-sequencing from single nuclei. Proc. Natl. Acad. Sci. U. S. A. *110*, 19802–19807.

Grün, D. (2019). Revealing dynamics of gene expression variability in cell state space. Nat. Methods.

Grun, D., Lyubimova, A., Kester, L., Wiebrands, K., Basak, O., Sasaki, N., Clevers, H., and van Oudenaarden, A. (2015). Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature *525*, 251–255.

Guglani, L., and Khader, S.A. (2010). Th17 cytokines in mucosal immunity and inflammation. Curr. Opin. HIV AIDS *5*, 120–127.

Guo, H., Zhu, P., Wu, X., Li, X., Wen, L., and Tang, F. (2013). Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. Genome Res. *23*, 2126–2135.

Guo, H., Zhu, P., Yan, L., Li, R., Hu, B., Lian, Y., Yan, J., Ren, X., Lin, S., Li, J., et al. (2014). The DNA methylation landscape of human early embryos. Nature *511*, 606–610.

Habib, N., Li, Y., Heidenreich, M., Swiech, L., Avraham-Davidi, I., Trombetta, J.J., Hession, C., Zhang, F., and Regev, A. (2016). Div-Seq: Single-nucleus RNA-Seq reveals dynamics of rare adult newborn neurons. Science *353*, 925 LP – 928.

Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. Nat. Biotechnol. *36*, 421–427.

Hahl, S.K., and Kremling, A. (2016). A comparison of deterministic and stochastic modeling approaches for biochemical reaction systems: On fixed points, means, and modes. Front. Genet. *7*, 157.

Halpern, K.B., Shenhav, R., Matcovitch-Natan, O., Tóth, B., Lemze, D., Golan, M., Massasa, E.E., Baydatch, S., Landen, S., Moor, A.E., et al. (2017). Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature *542*, 352–356.

Hamosh, A., Scott, A.F., Amberger, J., Valle, D., and McKusick, V.A. (2000). Online Mendelian Inheritance in Man (OMIM). Hum. Mutat. *15*, 57–61.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. Cell *144*, 646–674.

Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. Cell Rep. *2*, 666–673.

Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S.N., Richelle, A., Heinken, A., Haraldsdóttir, H.S., Wachowiak, J., Keating, S.M., Vlasov, V., et al. (2019). Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. Nat. Protoc. *14*, 639–702.

Hie, B., Peters, J., Nyquist, S.K., Shalek, A.K., Berger, B., and Bryson, B.D. (2020). Computational Methods for Single-Cell RNA Sequencing. Annu. Rev. Biomed. Data Sci.

Hildreth, A.D., Ma, F., Wong, Y.Y., Sun, R., Pellegrini, M., and O'Sullivan, T.E. (2021). Single-cell sequencing of human white adipose tissue identifies new cell states in health and obesity. Nat. Immunol.

Ho, P.-C., and Kaech, S.M. (2017). Reenergizing T cell anti-tumor immunity by harnessing immunometabolic checkpoints and machineries. Curr. Opin. Immunol. *46*, 38–44.

Ho, P.-C., Bihuniak, J.D., Macintyre, A.N., Staron, M., Liu, X., Amezquita, R., Tsui, Y.-C., Cui, G., Micevic, G., Perales, J.C., et al. (2015). Phosphoenolpyruvate Is a Metabolic Checkpoint of Anti-tumor T Cell Responses. Cell *162*, 1217–1228.

Hotamisligil, G.S. (2017). Foundations of Immunometabolism and Implications for Metabolic Health and Disease. Immunity *47*, 406–420.

Hou, Y., Song, L., Zhu, P., Zhang, B., Tao, Y., Xu, X., Li, F., Wu, K., Liang, J., Shao, D., et al. (2012). Single-Cell Exome Sequencing and Monoclonal Evolution of a JAK2-Negative Myeloproliferative Neoplasm. Cell *148*, 873–885.

Hsu, H.C., Seibold, J.R., and Thomas, T.J. (1994). Regulation of ornithine decarboxylase in the kidney of autoimmune mice with the lpr gene. Autoimmunity *19*, 253–264.

Huang, M., Wang, J., Torre, E., Dueck, H., Shaffer, S., Bonasio, R., Murray, J.I., Raj, A., Li, M., and Zhang, N.R. (2018). SAVER: gene expression recovery for single-cell RNA sequencing. Nat. Methods *15*, 539–542.

Ilicic, T., Kim, J.K., Kolodziejczyk, A.A., Bagger, F.O., McCarthy, D.J., Marioni, J.C., and Teichmann, S.A. (2016). Classification of low quality cells from single-cell RNA-seq data. Genome Biol. *17*, 1–15.

Jager, A., Dardalhon, V., Sobel, R.A., Bettelli, E., and Kuchroo, V.K. (2009). Th1, Th17, and Th9 effector cells induce experimental autoimmune encephalomyelitis with different pathological phenotypes. J. Immunol. *183*, 7169–7177.

Jell, J., Merali, S., Hensen, M.L., Mazurchuk, R., Spernyak, J.A., Diegelman, P., Kisiel, N.D., Barrero, C., Deeb, K.K., Alhonen, L., et al. (2007). Genetically altered expression of spermidine/spermine N1-acetyltransferase affects fat metabolism in mice via acetyl-CoA. J. Biol. Chem. *282*, 8404–8413.

Jha, A.K., Huang, S.C.-C., Sergushichev, A., Lampropoulou, V., Ivanova, Y., Loginicheva, E., Chmielewski, K., Stewart, K.M., Ashall, J., Everts, B., et al. (2015). Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization. Immunity *42*, 419–430.

Johnson, M.O., Wolf, M.M., Madden, M.Z., Andrejeva, G., Sugiura, A., Contreras, D.C., Maseda, D., Liberti, M.V., Paz, K., Kishton, R.J., et al. (2018). Distinct Regulation of Th17 and Th1 Cell Differentiation by Glutaminase-Dependent Metabolism. Cell *175*, 1780-1795.e19.

Josefowicz, S.Z., Lu, L.-F., and Rudensky, A.Y. (2012). Regulatory T cells: mechanisms of differentiation and function. Annu. Rev. Immunol. *30*, 531–564.

Jung, J., Zeng, H., and Horng, T. (2019). Metabolism as a guiding force for immunity. Nat. Cell Biol. *21*, 85–93.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. *45*, D353–D361.

Karouzakis, E., Gay, R.E., Gay, S., and Neidhart, M. (2012). Increased recycling of polyamines is associated with global DNA hypomethylation in rheumatoid arthritis synovial fibroblasts. Arthritis Rheum. *64*, 1809–1817.

Katzenelenbogen, Y., Sheban, F., Yalin, A., Yofe, I., Svetlichnyy, D., Jaitin, D.A., Bornstein, C., Moshe, A., Keren-Shaul, H., Cohen, M., et al. (2020). Coupled scRNA-Seq and Intracellular Protein Activity Reveal an Immunosuppressive Role of TREM2 in Cancer. Cell.

Kearney, C.J., Vervoort, S.J., Ramsbottom, K.M., Todorovski, I., Lelliott, E.J., Zethoven, M., Pijpers, L., Martin, B.P., Semple, T., Martelotto, L., et al. (2021). SUGAR-seq enables simultaneous detection of glycans, epitopes, and the transcriptome in single cells. Sci Adv *7*.

Kennedy, D.E., Okoreeh, M.K., Maienschein-Cline, M., Ai, J., Veselits, M., McLean, K.C., Dhungana, Y., Wang, H., Peng, J., Chi, H., et al. (2020). Novel specialized cell state and spatial compartments within the germinal center. Nat. Immunol. *21*, 660–670.

Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R., Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., et al. (2018). JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. *46*, D1284.

Kim, J.K., Kolodziejczyk, A.A., Illicic, T., Teichmann, S.A., and Marioni, J.C. (2015). Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. Nat. Commun. *6*.

Klein, A.M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D.A., and Kirschner, M.W. (2015). Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. Cell *161*, 1187–1201.

Kolodziejczyk, A.A., Kim, J.K., Svensson, V., Marioni, J.C., and Teichmann, S.A. (2015). The technology and biology of single-cell RNA sequencing. Mol. Cell *58*.

Kono, M., Maeda, K., Stocton-Gavanescu, I., Pan, W., Umeda, M., Katsuyama, E., Burbano, C., Orite, S.Y.K., Vukelic, M., Tsokos, M.G., et al. (2019). Pyruvate kinase M2 is requisite for Th1 and Th17 differentiation. JCI Insight *4*.

Korem, Y., Szekely, P., Hart, Y., Sheftel, H., Hausser, J., Mayo, A., Rothenberg, M.E., Kalisky, T., and Alon, U. (2015). Geometry of the Gene Expression Space of Individual Cells. PLoS Comput. Biol. *11*, e1004224.

Korn, T., Bettelli, E., Oukka, M., and Kuchroo, V.K. (2009). IL-17 and Th17 Cells. Annu. Rev. Immunol. *27*, 485–517.

Kowalczyk, M.S., Tirosh, I., Heckl, D., Nageswara Rao, T., Dixit, A., Haas, B.J., Schneider, R., Wagers, A.J., Ebert, B.L., and Regev, A. (2015). Single cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res.

Kretschmer, K., Apostolou, I., Hawiger, D., Khazaie, K., Nussenzweig, M.C., and von Boehmer, H. (2005). Inducing and expanding regulatory T cell populations by foreign antigen. Nat. Immunol. *6*, 1219–1227.

Kumar, S., Sharife, H., Kreisel, T., Mogilevsky, M., Bar-Lev, L., Grunewald, M., Aizenshtein, E., Karni, R., Paldor, I., Shlomi, T., et al. (2019). Intra-Tumoral Metabolic Zonation and Resultant Phenotypic Diversification Are Dictated by Blood Vessel Proximity. Cell Metab. *30*, 201-211.e6.

Lacar, B., Linker, S.B., Jaeger, B.N., Krishnaswami, S.R., Barron, J.J., Kelder, M.J.E., Parylak, S.L., Paquola, A.C.M., Venepally, P., Novotny, M., et al. (2016). Nuclear RNA-seq of single neurons reveals molecular signatures of activation. Nat. Commun. *7*, 11022.

Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D.J., Hicks, S.C., Robinson, M.D., Vallejos, C.A., Campbell, K.R., Beerenwinkel, N., Mahfouz, A., et al. (2020). Eleven grand challenges in single-cell data science. Genome Biol. *21*, 31.

Lake, B.B., Ai, R., Kaeser, G.E., Salathia, N.S., Yung, Y.C., Liu, R., Wildberg, A., Gao, D., Fung, H.-L., Chen, S., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science *352*, 1586 LP – 1590.

Lande-Diner, L., Stewart-Ornstein, J., Weitza, C.J., and Lahavb, G. (2015). Single cell analysis of circadian dynamics in tissue explants. Mol. Biol. Cell.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lareau, C.A., Ludwig, L.S., Muus, C., Gohil, S.H., Zhao, T., Chiang, Z., Pelka, K., Verboon, J.M., Luo, W., Christian, E., et al. (2021). Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. Nat. Biotechnol. *39*, 451–461.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, 1–17.

Lee, K.H., Park, J.H., Kim, T.Y., Kim, H.U., and Lee, S.Y. (2007). Systems metabolic engineering of Escherichia coli for L-threonine production. Mol. Syst. Biol. *3*.

Lee, Y., Awasthi, A., Yosef, N., Quintana, F.J., Xiao, S., Peters, A., Wu, C., Kleinewietfeld, M., Kunder, S., Hafler, D.A., et al. (2012). Induction and molecular signature of pathogenic TH17 cells. Nat. Immunol. *13*, 991–999.

Lee, Y., Collins, M., and Kuchroo, V.K. (2014). Unexpected targets and triggers of autoimmunity. J. Clin. Immunol. *34 Suppl 1*, S56-60.

Lee, Y., Bogdanoff, D., Wang, Y., Hartoularos, G.C., Woo, J.M., Mowery, C.T., Nisonoff, H.M., Lee, D.S., Sun, Y., Lee, J., et al. (2021). XYZeq: Spatially resolved single-cell RNA sequencing reveals expression heterogeneity in the tumor microenvironment. Sci Adv *7*.

Lesurf, R., Cotto, K.C., Wang, G., Griffith, M., Kasaian, K., Jones, S.J.M., Montgomery, S.B., Griffith, O.L., and Open Regulatory Annotation Consortium (2016). ORegAnno 3.0: a community-driven resource for curated regulatory annotation. Nucleic Acids Res. *44*, D126-32.

Leung, M., Wang, Y., Waters, J., and Navin, N. (2015). SNES: single nucleus exome sequencing. Genome Biol. *16*, 55.

Levine, L.S., Hiam-Galvez, K.J., Marquez, D.M., Tenvooren, I., Madden, M.Z., Contreras, D.C., Dahunsi, D.O., Irish, J.M., Oluwole, O.O., Rathmell, J.C., et al. (2021). Single-cell analysis by mass cytometry reveals metabolic states of early-activated CD8+ T cells during the primary immune response. Immunity.

Lewis, N.E., Hixson, K.K., Conrad, T.M., Lerman, J.A., Charusanti, P., Polpitiya, A.D., Adkins, J.N., Schramm, G., Purvine, S.O., Lopez-Ferrer, D., et al. (2010). Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. Mol. Syst. Biol. *6*.

Lewis, N.E., Nagarajan, H., and Palsson, B.O. (2012). Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. Nat. Rev. Microbiol. *10*, 291–305.

Li, Q., Zou, J., Wang, M., Ding, X., Chepelev, I., Zhou, X., Zhao, W., Wei, G., Cui, J., Zhao, K., et al. (2014). Critical role of histone demethylase Jmjd3 in the regulation of CD4+ T-cell differentiation. Nat. Commun. *5*, 5780.

Liu, Z., Cao, W., Xu, L., Chen, X., Zhan, Y., Yang, Q., Liu, S., Chen, P., Jiang, Y., Sun, X., et al. (2015). The histone H3 lysine-27 demethylase Jmjd3 plays a critical role in specific regulation of Th17 cell differentiation. J. Mol. Cell Biol. *7*, 505–516.

Lohr, J.G., Adalsteinsson, V.A., Cibulskis, K., Choudhury, A.D., Rosenberg, M., Cruz-Gordillo, P., Francis, J.M., Zhang, C.-Z., Shalek, A.K., Satija, R., et al. (2014). Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat. Biotechnol. *32*, 479–484.

Lopez, R., Regier, J., Cole, M.B., Jordan, M.I., and Yosef, N. (2018). Deep generative modeling for single-cell transcriptomics. Nat. Methods *15*, 1053–1058.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lun, A.T.L., Bach, K., and Marioni, J.C. (2016). Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. Genome Biol. *17*, 1–14.

Macaulay, I.C., Haerty, W., Kumar, P., Li, Y.I., Hu, T.X., Teng, M.J., Goolam, M., Saurat, N., Coupland, P., Shirley, L.M., et al. (2015). G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. Nat. Methods *12*, 519–522.

Macintyre, A.N., Gerriets, V.A., Nichols, A.G., Michalek, R.D., Rudolph, M.C., Deoliveira, D., Anderson, S.M., Abel, E.D., Chen, B.J., Hale, L.P., et al. (2014). The glucose transporter Glut1 is selectively essential for CD4 T cell activation and effector function. Cell Metab. *20*, 61–72.

MacIver, N.J., Michalek, R.D., and Rathmell, J.C. (2013). Metabolic Regulation of T Lymphocytes. Annu. Rev. Immunol. *31*, 259–283.

Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202–1214.

Mahadevan, R., Edwards, J.S., and Doyle, F.J., III (2002). Dynamic flux balance analysis of diauxic growth in Escherichia coli. Biophys. J. *83*, 1331–1340.

Makowski, L., Chaib, M., and Rathmell, J.C. (2020). Immunometabolism: From basic mechanisms to translation. Immunol. Rev. *295*, 5–14.

Mangan, P.R., Harrington, L.E., O'Quinn, D.B., Helms, W.S., Bullard, D.C., Elson, C.O., Hatton, R.D., Wahl, S.M., Schoeb, T.R., and Weaver, C.T. (2006). Transforming growth factor-beta induces development of the T(H)17 lineage. Nature *441*, 231–234.

Marioni, J.C., and Arendt, D. (2017). How single-cell genomics is changing evolutionary and developmental biology. Annu. Rev. Cell Dev. Biol. *33*, 537–553.

Marsden, D., Larson, C., and Levy, H.L. (2006). Newborn screening for metabolic disorders. J. Pediatr. *148*, 577–584.

Martínez, V.S., Buchsteiner, M., Gray, P., Nielsen, L.K., and Quek, L.-E. (2015). Dynamic metabolic flux analysis using B-splines to study the effects of temperature shift on CHO cell metabolism. Metab. Eng. Commun. *2*, 46–57.

Mathewson, N.D., Ashenberg, O., Tirosh, I., Gritsch, S., Perez, E.M., Marx, S., Jerby-Arnon, L., Chanoch-Myers, R., Hara, T., Richman, A.R., et al. (2021). Inhibitory CD161 receptor identified in glioma-infiltrating T cells by single-cell analysis. Cell *184*, 1281-1298.e26.

McGeachy, M.J., and Cua, D.J. (2008). Th17 cell differentiation: the long and winding road. Immunity *28*, 445–453.

McGeachy, M.J., Bak-Jensen, K.S., Chen, Y., Tato, C.M., Blumenschein, W., McClanahan, T., and Cua, D.J. (2007). TGF-beta and IL-6 drive the production of IL-17 and IL-10 by T cells and restrain T(H)-17 cell-mediated pathology. Nat. Immunol. *8*, 1390–1397.

McGeachy, M.J., Chen, Y., Tato, C.M., Laurence, A., Joyce-Shaikh, B., Blumenschein, W.M., McClanahan, T.K., O'Shea, J.J., and Cua, D.J. (2009). The interleukin 23 receptor is essential for the terminal differentiation of interleukin 17-producing effector T helper cells in vivo. Nat. Immunol. *10*, 314–324.

McGinnis, C.S., Patterson, D.M., Winkler, J., Conrad, D.N., Hein, M.Y., Srivastava, V., Hu, J.L., Murrow, L.M., Weissman, J.S., Werb, Z., et al. (2019). MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. Nat. Methods *16*, 619–626.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Mews, P., Donahue, G., Drake, A.M., Luczak, V., Abel, T., and Berger, S.L. (2017). Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. Nature *546*, 381–386.

Michalek, R.D., Gerriets, V.A., Jacobs, S.R., Macintyre, A.N., MacIver, N.J., Mason, E.F., Sullivan, S.A., Nichols, A.G., and Rathmell, J.C. (2011). Cutting edge: distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4+ T cell subsets. The Journal of Immunology *186*, 3299–3303.

Mills, E., and O'Neill, L.A.J. (2014). Succinate: a metabolic signal in inflammation. Trends Cell Biol. *24*, 313–320.

Mills, E.L., Kelly, B., Logan, A., Costa, A.S.H., Varma, M., Bryant, C.E., Tourlomousis, P., Däbritz, J.H.M., Gottlieb, E., Latorre, I., et al. (2016). Succinate Dehydrogenase Supports Metabolic Repurposing of Mitochondria to Drive Inflammatory Macrophages. Cell *167*, 457-470.e13.

Mimitou, E.P., Cheng, A., Montalbano, A., Hao, S., Stoeckius, M., Legut, M., Roush, T., Herrera, A., Papalexi, E., Ouyang, Z., et al. (2019). Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. Nat. Methods *16*, 409–412.

Miragaia, R.J., Gomes, T., Chomka, A., Jardine, L., Riedel, A., Hegazy, A.N., Whibley, N., Tucci, A., Chen, X., Lindeman, I., et al. (2019). Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. Immunity *50*, 493-504.e7.

Monk, J., Nogales, J., and Palsson, B.O. (2014). Optimizing genome-scale network reconstructions. Nat. Biotechnol. *32*, 447–452.

Mounce, B.C., Poirier, E.Z., Passoni, G., Simon-Loriere, E., Cesaro, T., Prot, M., Stapleford, K.A., Moratorio, G., Sakuntabhai, A., Levraud, J.P., et al. (2016). Interferon-Induced Spermidine-Spermine Acetyltransferase and Polyamine Depletion Restrict Zika and Chikungunya Viruses. Cell Host Microbe *20*, 167–177.

Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature *502*, 59–64.

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., et al. (2011). Tumour evolution inferred by single-cell sequencing. Nature *472*, 90–94.

Newman, M. (2018). Biological networks (Oxford University Press).

Newton, R., Priyadharshini, B., and Turka, L.A. (2016). Immunometabolism of regulatory T cells. Nat. Immunol. *17*, 618–625.

Nitzan, M., Karaiskos, N., Friedman, N., and Rajewsky, N. (2019). Gene expression cartography. Nature *576*, 132–137.

O’Brien, K.L., and Finlay, D.K. (2019). Immunometabolism and natural killer cell responses. Nat. Rev. Immunol. *19*, 282–290.

O’Brien, E.J., Monk, J.M., and Palsson, B.O. (2015). Using Genome-scale Models to Predict Biological Capabilities. Cell *161*, 971–987.

Omenetti, S., and Pizarro, T.T. (2015). The Treg/Th17 Axis: A Dynamic Balance Regulated by the Gut Microbiome. Front. Immunol. *6*, 639.

O'Neill, L.A.J., and Pearce, E.J. (2016). Immunometabolism governs dendritic cell and macrophage function. J. Exp. Med. *213*, 15–23.

O'Neill, L.A.J., Kishton, R.J., and Rathmell, J. (2016). A guide to immunometabolism for immunologists. Nat. Rev. Immunol. *16*, 553–565.

Orth, J.D., Thiele, I., and Palsson, B.O. (2010). What is flux balance analysis? Nat. Biotechnol. *28*, 245–248.

Ouyang, W., Kolls, J.K., and Zheng, Y. (2008). The biological functions of T helper 17 cell effector cytokines in inflammation. Immunity *28*, 454–467.

Palsson, B.O. (2011). Systems biology: Simulation of dynamic network states (Cambridge, England: Cambridge University Press).

Palsson, B.Ø. (2015). Systems Biology: Constraint-based Reconstruction and Analysis (Cambridge, England: Cambridge University Press).

Pampols, T. (2010). Inherited metabolic rare disease. Adv. Exp. Med. Biol. *686*, 397–431.

Park, J.H., Lee, K.H., Kim, T.Y., and Lee, S.Y. (2007). Metabolic engineering of Escherichia coli for the production of l-valine based on transcriptome analysis and in silico gene knockout simulation. Proceedings of the National Academy of Sciences *104*, 7797–7802.

Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401.

Pearce, E.J., and Everts, B. (2015). Dendritic cell metabolism. Nat. Rev. Immunol. *15*, 18–29.

Pearce, E.L., Poffenberger, M.C., Chang, C.-H., and Jones, R.G. (2013). Fueling Immunity: Insights into Metabolism and Lymphocyte Function. Science *342*.

Pegg, A.E. (2008). Spermidine/spermine-N(1)-acetyltransferase: a key metabolic regulator. American Journal of Physiology-Endocrinology and Metabolism *294*, E995-1010.

Peng, M., Yin, N., Chhangawala, S., Xu, K., Leslie, C.S., and Li, M.O. (2016). Aerobic glycolysis promotes T helper 1 cell differentiation through an epigenetic mechanism. Science *354*, 481–484.

Petrucci, R.H., Herring, F.G., Madura, J.D., and Bissonnette, C. (2016). General chemistry (Pearson).

Pharkya, P., Burgard, A.P., and Maranas, C.D. (2003). Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. Biotechnol. Bioeng. *84*, 887–899.

Pharkya, P., Burgard, A.P., and Maranas, C.D. (2004). OptStrain: a computational framework for redesign of microbial production systems. Genome Research *14*, 2367–2376.

Picelli, S., Bjorklund, A.K., Faridani, O.R., Sagasser, S., Winberg, G., and Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods *10*, 1096–1098.

Pollen, A.A., Nowakowski, T.J., Chen, J., Retallack, H., Sandoval-Espinosa, C., Nicholas, C.R., Shuga, J., Liu, S.J., Oldham, M.C., Diaz, A., et al. (2015). Molecular Identity of Human Outer Radial Glia during Cortical Development. Cell *163*, 55–67.

Pompura, S.L., Wagner, A., Kitz, A., LaPerche, J., Yosef, N., Dominguez-Villar, M., and Hafler, D.A. (2021). Oleic acid restores suppressive defects in tissue-resident FOXP3 Tregs from patients with multiple sclerosis. J. Clin. Invest. *131*, e138519.

Pucino, V., Certo, M., Bulusu, V., Cucchi, D., Goldmann, K., Pontarini, E., Haas, R., Smith, J., Headland, S.E., Blighe, K., et al. (2019). Lactate Buildup at the Site of Chronic Inflammation Promotes Disease by Inducing CD4+ T Cell Metabolic Rewiring. Cell Metab. *30*, 1055-1074.e8.

Puleston, D.J., Villa, M., and Pearce, E.L. (2017). Ancillary Activity: Beyond Core Metabolism in Immune Cells. Cell Metab. *26*, 131–141.

Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. Cell *135*, 216–226.

Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA Synthesis in Mammalian Cells. PLoS Biol. *4*, e309.

Ramskold, D., Luo, S., Wang, Y.-C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtukova, I., Loring, J.F., Laurent, L.C., et al. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat. Biotechnol. *30*, 777–782.

Ranganathan, S., Suthers, P.F., and Maranas, C.D. (2010). OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. PLoS Comput. Biol. *6*, e1000744.

Raud, B., Roy, D.G., Divakaruni, A.S., Tarasenko, T.N., Franke, R., Ma, E.H., Samborska, B., Hsieh, W.Y., Wong, A.H., Stüve, P., et al. (2018). Etomoxir Actions on Regulatory and Memory T Cells Are Independent of Cpt1a-Mediated Fatty Acid Oxidation. Cell Metab. *28*, 504-515.e7.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017a). Science forum: the human cell atlas. Elife *6*, e27041.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017b). The Human Cell Atlas. Elife *6*.

Ringel, A.E., Drijvers, J.M., Baker, G.J., Catozzi, A., García-Cañaveras, J.C., Gassaway, B.M., Miller, B.C., Juneja, V.R., Nguyen, T.H., Joshi, S., et al. (2020). Obesity Shapes Metabolism in the Tumor Microenvironment to Suppress Anti-Tumor Immunity. Cell *183*, 1848-1866.e26.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47–e47.

Rodriques, S.G., Stickels, R.R., Goeva, A., Martin, C.A., Murray, E., Vanderburg, C.R., Welch, J., Chen, L.M., Chen, F., and Macosko, E.Z. (2019). Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. Science *363*, 1463–1467.

Romani, L. (2011). Immunity to fungal infections. Nat. Rev. Immunol. *11*, 275–288.

Rotem, A., Ram, O., Shoresh, N., Sperling, R.A., Goren, A., Weitz, D.A., and Bernstein, B.E. (2015). Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat. Biotechnol. *33*, 1165–1172.

Roy, D.G., Kaymak, I., Williams, K.S., Ma, E.H., and Jones, R.G. (2021). Immunometabolism in the Tumor Microenvironment. Annu. Rev. Cancer Biol. *5*, 137–159.

Rozenblatt-Rosen, O., Stubbington, M.J.T., Regev, A., and Teichmann, S.A. (2017). The Human Cell Atlas: from vision to reality. Nature *550*, 451–453.

Rozenblatt-Rosen, O., Regev, A., Oberdoerffer, P., Nawy, T., Hupalowska, A., Rood, J.E., Ashenberg, O., Cerami, E., Coffey, R.J., Demir, E., et al. (2020). The Human Tumor Atlas

Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. Cell *181*, 236–249.

Rozenblatt-Rosen, O., Shin, J.W., Rood, J.E., Hupalowska, A., Ardlie, K., Clatworthy, M., Carninci, P., Enard, W., Greenleaf, W., Heyn, H., et al. (2021). Building a high-quality Human Cell Atlas. Nat. Biotechnol.

Russell, D.G., Huang, L., and VanderVen, B.C. (2019). Immunometabolism at the interface between macrophages and pathogens. Nat. Rev. Immunol.

Sakaguchi, S., Yamaguchi, T., Nomura, T., and Ono, M. (2008). Regulatory T cells and immune tolerance. Cell *133*, 775–787.

Sakaguchi, S., Mikami, N., Wing, J.B., Tanaka, A., Ichiyama, K., and Ohkura, N. (2020). Regulatory T cells and human disease. Annu. Rev. Immunol. *38*, 541–566.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat. Biotechnol. *33*, 495–502.

Schellenberger, J., Lewis, N.E., and Palsson, B.Ø. (2011). Elimination of Thermodynamically Infeasible Loops in Steady-State Metabolic Models. Biophys. J. *100*, 544–553.

Segrè, D., Vitkup, D., and Church, G.M. (2002). Analysis of optimality in natural and perturbed metabolic networks. Proceedings of the National Academy of Sciences *99*, 15112–15117.

Shah, S., Lubeck, E., Zhou, W., and Cai, L. (2016). In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. Neuron *92*, 342–357.

Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., et al. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature *498*, 236–240.

Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaublomme, J.T., Yosef, N., et al. (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature *510*, 363–369.

Shevach, E.M., and Thornton, A.M. (2014). tTregs, pTregs, and iTregs: similarities and differences. Immunol. Rev. *259*, 88–102.

Shi, L., Jiang, Q., Bushkin, Y., Subbian, S., and Tyagi, S. (2019). Biphasic Dynamics of Macrophage Immunometabolism during Mycobacterium tuberculosis Infection. MBio *10*.

Shi, L.Z., Wang, R., Huang, G., Vogel, P., Neale, G., Green, D.R., and Chi, H. (2011). HIF1alpha-dependent glycolytic pathway orchestrates a metabolic checkpoint for the differentiation of TH17 and Treg cells. J. Exp. Med. *208*, 1367–1376.

Shlomi, T., Cabili, M.N., Herrgard, M.J., Palsson, B.O., and Ruppin, E. (2008). Network-based prediction of human tissue-specific metabolism. Nat. Biotechnol. *26*, 1003–1010.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. *15*, 1034–1050.

Smallwood, S.A., Lee, H.J., Angermueller, C., Krueger, F., Saadeh, H., Peat, J., Andrews, S.R., Stegle, O., Reik, W., and Kelsey, G. (2014). Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat. Methods *11*, 817–820.

Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science *353*, 78–82.

Stewart-Ornstein, J., Weissman, J.S., and El-Samad, H. (2012). Cellular Noise Regulons Underlie Fluctuations in Saccharomyces cerevisiae. Mol. Cell *45*, 483–493.

Stickels, R.R., Murray, E., Kumar, P., Li, J., Marshall, J.L., Di Bella, D.J., Arlotta, P., Macosko, E.Z., and Chen, F. (2020). Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-seqV2. Nat. Biotechnol.

Stockinger, B., and Omenetti, S. (2017). The dichotomous nature of T helper 17 cells. Nat. Rev. Immunol. *17*, 535–544.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods *14*, 865–868.

Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck, W.M., 3rd, Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol. *19*, 224.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. *102*, 15545–15550.

Sundrud, M.S., Koralov, S.B., Feuerer, M., Calado, D.P., Kozhaya, A.E., Rhule-Smith, A., Lefebvre, R.E., Unutmaz, D., Mazitschek, R., Waldner, H., et al. (2009). Halofuginone Inhibits T(H)17 Cell Differentiation by Activating the Amino Acid Starvation Response. Science *324*, 1334–1338.

Svensson, V., Vento-Tormo, R., and Teichmann, S.A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. Nat. Protoc. *13*, 599–604.

Swain, P.S., Elowitz, M.B., and Siggia, E.D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. Proceedings of the National Academy of Sciences *99*, 12795–12800.

Tanay, A., and Regev, A. (2017). Scaling single-cell genomics from phenomenology to mechanism. Nature *541*, 331–338.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B.B., Siddiqui, A., et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods *6*, 377–382.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., et al. (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci. *19*, 335–346.

Thiele, I., Swainston, N., Fleming, R.M.T., Hoppe, A., Sahoo, S., Aurich, M.K., Haraldsdottir, H., Mo, M.L., Rolfsson, O., Stobbe, M.D., et al. (2013). A community-driven global reconstruction of human metabolism. Nat. Biotechnol. *31*, 419–425.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

Van den Bossche, J., O’Neill, L.A., and Menon, D. (2017). Macrophage Immunometabolism: Where Are We (Going)? Trends Immunol. *38*, 395–406.

Varanasi, S.K., Kumar, S.V., and Rouse, B.T. (2020). Determinants of Tissue-Specific Metabolic Adaptation of T Cells. Cell Metab. *32*, 908–919.

Varma, A., and Palsson, B.O. (1994). Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use. Nature Biotechnology *12*, 994–998.

Veldhoen, M., Hocking, R.J., Atkins, C.J., Locksley, R.M., and Stockinger, B. (2006). TGFbeta in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. Immunity *24*, 179–189.

Vickovic, S., Eraslan, G., Salmén, F., Klughammer, J., Stenbeck, L., Schapiro, D., Äijö, T., Bonneau, R., Bergenstråhle, L., Navarro, J.F., et al. (2019). High-definition spatial transcriptomics for in situ tissue profiling. Nat. Methods *16*, 987–990.

Vieira Braga, F.A., Kar, G., Berg, M., Carpaij, O.A., Polanski, K., Simon, L.M., Brouwer, S., Gomes, T., Hesse, L., Jiang, J., et al. (2019). A cellular census of human lungs identifies novel cell states in health and in asthma. Nat. Med. *25*, 1153–1163.

Vignali, D.A.A., Collison, L.W., and Workman, C.J. (2008). How regulatory T cells work. Nat. Rev. Immunol. *8*, 523–532.

Voss, K., Hong, H.S., Bader, J.E., Sugiura, A., Lyssiotis, C.A., and Rathmell, J.C. (2021). A guide to interrogating immunometabolism. Nat. Rev. Immunol.

Wagner, A., Zarecki, R., Reshef, L., Gochev, C., Sorek, R., Gophna, U., and Ruppin, E. (2013). Computational evaluation of cellular metabolic costs successfully predicts genes whose expression is deleterious. Proceedings of the National Academy of Sciences *110*, 19166–19171.

Wagner, A., Cohen, N., Kelder, T., Amit, U., Liebman, E., Steinberg, D.M., Radonjic, M., and Ruppin, E. (2015). Drugs that reverse disease transcriptomic signatures are more effective in a mouse model of dyslipidemia. Mol. Syst. Biol. *11*, 791.

Wagner, A., Regev, A., and Yosef, N. (2016). Revealing the vectors of cellular identity with single-cell genomics. Nat. Biotechnol. *34*, 1145–1160.

Wang, C., Yosef, N., Gaublomme, J., Wu, C., Lee, Y., Clish, C.B., Kaminski, J., Xiao, S., Horste, G.M.Z., Pawlak, M., et al. (2015). CD5L/AIM Regulates Lipid Biosynthesis and Restrains Th17 Cell Pathogenicity. Cell *163*, 1413–1427.

Wang, Y., Waters, J., Leung, M.L., Unruh, A., Roh, W., Shi, X., Chen, K., Scheet, P., Vattathil, S., Liang, H., et al. (2014). Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature *512*, 155–160.

Ward, P.S., Patel, J., Wise, D.R., Abdel-Wahab, O., Bennett, B.D., Coller, H.A., Cross, J.R., Fantin, V.R., Hedvat, C.V., Perl, A.E., et al. (2010). The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. Cancer Cell *17*, 225–234.

Watson, M.J., Vignali, P.D.A., Mullett, S.J., Overacre-Delgoffe, A.E., Peralta, R.M., Grebinoski, S., Menk, A.V., Rittenhouse, N.L., DePeaux, K., Whetstone, R.D., et al. (2021). Metabolic support of tumour-infiltrating regulatory T cells by lactic acid. Nature.

Wellen, K.E., Hatzivassiliou, G., Sachdeva, U.M., Bui, T.V., Cross, J.R., and Thompson, C.B. (2009). ATP-citrate lyase links cellular metabolism to histone acetylation. Science *324*, 1076–1080.

van der Windt, G.J.W., and Pearce, E.L. (2012). Metabolic switching and fuel choice during T-cell differentiation and memory development. Immunol. Rev. *249*, 27–42.

Wu, C., Chen, Z., Xiao, S., Thalhamer, T., Madi, A., Han, T., and Kuchroo, V. (2018). SGK1 Governs the Reciprocal Development of Th17 and Regulatory T Cells. Cell Rep. *22*, 653–665.

Xiao, S., Yosef, N., Yang, J., Wang, Y., Zhou, L., Zhu, C., Wu, C., Baloglu, E., Schmidt, D., Ramesh, R., et al. (2014). Small-molecule RORgammat antagonists inhibit T helper 17 cell transcriptional network by divergent mechanisms. Immunity *40*, 477–489.

Xu, X., Hou, Y., Yin, X., Bao, L., Tang, A., Song, L., Li, F., Tsang, S., Wu, K., Wu, H., et al. (2012). Single-Cell Exome Sequencing Reveals Single-Nucleotide Mutation Characteristics of a Kidney Tumor. Cell *148*, 886–895.

Yang, Y., Torchinsky, M.B., Gobert, M., Xiong, H., Xu, M., Linehan, J.L., Alonzo, F., Ng, C., Chen, A., Lin, X., et al. (2014). Focused specificity of intestinal TH17 cells towards commensal bacterial antigens. Nature *510*, 152–156.

Yizhak, K., Gabay, O., Cohen, H., and Ruppin, E. (2013). Model-based identification of drug targets that revert disrupted metabolism and its application to ageing. Nat. Commun. *4*, 2632.

Yosef, N., Shalek, A.K., Gaublomme, J.T., Jin, H., Lee, Y., Awasthi, A., Wu, C., Karwacz, K., Xiao, S., Jorgolli, M., et al. (2013). Dynamic regulatory network controlling TH17 cell differentiation. Nature *496*, 461–468.

Yugi, K., Nakayama, Y., Kinoshita, A., and Tomita, M. (2005). Hybrid dynamic/static method for large-scale simulation of metabolism. Theor. Biol. Med. Model. *2*, 42.

Zeisel, A., Manchado, A.B.M., Codeluppi, S., Lönnerberg, P., La Manno, G., Juréus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., et al. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhou, L., Ivanov, I.I., Spolski, R., Min, R., Shenderov, K., Egawa, T., Levy, D.E., Leonard, W.J., and Littman, D.R. (2007). IL-6 programs T(H)-17 cell differentiation by promoting sequential engagement of the IL-21 and IL-23 pathways. Nat. Immunol. *8*, 967–974.

Zielinski, C.E., Mele, F., Aschenbrenner, D., Jarrossay, D., Ronchi, F., Gattorno, M., Monticelli, S., Lanzavecchia, A., and Sallusto, F. (2012). Pathogen-induced human TH17 cells produce IFN-gamma or IL-10 and are regulated by IL-1beta. Nature *484*, 514–518.

Zong, C., Lu, S., Chapman, A.R., and Xie, X.S. (2012). Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. Science *338*, 1622–1626.