PhD-FSTM-2023-014
The Faculty of Science, Technology and Medicine

DISSERTATION

Defence held on 09/02/2023 in Esch-sur-Alzette
to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG EN BIOLOGIE

by

Kartikeya SINGH
Born on 29th September 1993 in city of New Delhi, India

DECIPHERING THE ROLE OF CELL-CELL COMMUNICATION IN HEALTH AND DISEASE - USING SYSTEMS BIOLOGY BASED COMPUTATIONAL MODELLING

Dissertation defence committee

Dr. Antonio del Sol, Dissertation supervisor
Professor, Université du Luxembourg

Dr. Thomas Sauter, Chairman
Professor, Université du Luxembourg

Dr. Feng He
Luxembourg Institute of Health

Dr. Juan Anguita
Professor, CIC bioGUNE - Centro de Investigación Cooperativa en Biociencias

Dr. Catharina Schütz
Professor, Universitätsklinikum Carl Gustav Carus Dresden

Affidavit

I hereby confirm that the Ph.D. thesis entitled “DECIPHERING THE ROLE OF CELL-CELL COMMUNICATION IN HEALTH AND DISEASE - USING SYSTEMS BIOLOGY BASED COMPUTATIONAL MODELLING” has been written independently and without any other sources than cited.

Luxembourg, 09/02/2023
Kartikeya Singh

Acknowledgments

First, I would like to thank my PhD supervisor Prof. Antonio del Sol for the opportunity he gave me when accepting me as a PhD candidate in his research group. I would like to thank him for his help and the guidance he provided me during my PhD studies. I have learned to think critically, to become more independent in my research and to improve my skills and knowledge in the field thanks to the training I had during these years.

Then, I would like to thank my CET members, Prof. Thomas Sauter and Dr. Feng He for their yearly follow-up, the scientific discussions and the valuable advice they gave me to improve my research projects. I would also like to thank the Fond National de la Recherche Luxembourg for funding my research and all the colleagues of my DTU.

I would also like to thank my collaborators at the Luxembourg Institute of Health (LIH), Dr. Guillem Montamat, Dr. Cathy Leonard and Prof. Markus Ollert at the Department of Infection and Immunity for their help and support in my projects.

Finally, I would like to thank Dr. Srikanth Ravichandran and Dr. Sascha Jung for their co-supervision of my PhD projects. Thank you so much for all the time you took to guide me, provide me with valuable feedback and suggestions, for all the stimulating scientific discussions and, most of all for your advice to help me grow as a researcher. I would also like to thank all my CBG colleagues for all the help and support they provided me during these years.

At last, I must thank all the people I lost or found during these years of PhD and especially those who remained a constant.

"*When I first read a biology textbook, it was like reading a thriller.*

Every page brought a new shock.

As a physicist, I was used to studying matter that obeys precise mathematical laws.

But cells are matter that dances.

Structures spontaneously assemble, perform elaborate biochemical functions,

***and vanish effortlessly when their work is done*."**

– Uri Alon, Chapter 1- An Introduction to Systems Biology: Design Principles of Biological Circuits

Table of contents

List of figures

Introduction

Manuscript 1

Manuscript 2

Manuscript 3

List of Abbreviations

ATAC-seq: Assay for transposase-accessible chromatin sequencing
BP: Biological process
DNA: Deoxyribonucleic acid
DEG: Differentially expressed gene
GRN: Gene regulatory network
RNA: Ribonucleic acid
mRNA: Messenger ribonucleic acid
scRNA-seq: single-cell assay for ribonucleic acid sequencing
TF: Transcription factor
UMAP: Uniform manifold approximation and projection

Summary

Cell-cell communication plays a significant role in shaping the functionality of the cells. Communication between the cells is also responsible for maintaining the physiological state of the cells and the tissue. Therefore, it is important to study the different ways by which cell-cell communication impacts the functional state of cells. Alterations in cell-cell communication can contribute to the development of disease conditions. In this thesis, we present two computational tools and a study to explore the different aspects of cell-cell communication. In the first manuscript, FunRes was developed to leverage the cell-cell communication model to investigate functional heterogeneity in cell types and characterize cell states based on the integration of inter- and intra-cellular signalling. This tool utilizes a combination of receptors and transcription factors (TFs) based on the reconstructed cell-cell communication network to split the cell types into functional states. The tool was applied to the *TabulaMurisSenis* atlas to discover functional cell states in both young and old mouse datasets. In addition, we compared our tool with state-of-the-art tools and validated the cell states using available markers from the literature. Secondly, we studied the evolution of gene expression in developing astrocytes under normal and inflammatory conditions. We characterized these cells using both transcriptional and chromatin accessibility data which were integrated to reconstruct the gene regulatory networks (GRNs) specific to the condition and timepoints. The GRNs were then topologically analyzed to identify key regulators of the developmental process under both normal and inflammatory conditions. In the final manuscript, we developed a computational tool that identified regulators of allergy and tolerance in a mouse model. The tool works by first reconstructing the cell-cell communication network and then analyzing the network for feedback loops. These feedback loops are important as they contribute to the sustenance of the tissue's state. Identification of the feedback loops allows for the discovery of important molecules by comparative analysis of these feedback loops between various conditions.

In summary, this thesis encompasses various ways of cellular regulation using cell-cell communication in a tissue. These studies contribute to a better understanding of the role cell-cell communication plays in health and disease along with the identification of therapeutic targets to design novel strategies against diseases.

1. Introduction

1.1. Systems biology: a perspective

Systems biology at its core is meant as an inverse to the concept of reductionism that took a foothold in the scientific progress of the 19th and 20th centuries. Reductionism is the idea that a complex problem can be broken down into its constituent parts and can then be studied individually to understand the whole problem (Brigandt 2022). This found its use in the breaking down of scientific theories and concepts into their constituents to understand them. Substantial examples of those can be found in the theories of atoms being split into individual particles to conceptualize their individual roles in an atom. The systems biology approach gained traction in the early part of the 21st century which states that a system is more than the sum of its parts and breaking down into constituent parts leads to the loss of information that was composed in the original system (Wanjek 2011). The systems approach lends itself well to the discipline of biology, as biology is composed of a multitude of complex individual parts that can interact with each other in a myriad number of ways.

Studies focused on the reductionist approach have helped understand the role of different organs and cell types independently. At the molecular level, individual genes have been studied for their roles in physiology and the correlation of their expression with diseases (Pinero, Ramirez-Anguita et al. 2020). The cells exert their roles through multiple levels of organization including the genome, transcriptome, epigenome, proteome, and metabolome levels (TA 2002). Each with its responsibility in the working of biological systems. The systems approach intends to use the knowledge generated from these individual levels in an integrated manner to explore the information that cannot be garnered by studying them separately. This integrated approach has been found useful in the development of therapies against diseases (Mulder, Hamidi et al. 2018, Xu, Ware et al. 2019). For example, the integration of transcriptomics, which signifies the number of transcripts along with DNA accessibility data from epigenetics could aid to unravel key regulatory mechanisms that govern the expression of genes in the cells and identify key regulators. Therefore, to gain a holistic understanding of biology it is important to utilize the systems approach of integration of these various levels. Biological systems are inherently robust to perturbations but under disease conditions perturbations could occur at multiple levels in the cells which can push the system into a dysfunctional state (del Sol, Balling et al. 2010). As the disruption occurs at

various levels it is important to study these various levels and the relationships between them. This holistic view of biology can then be utilized to develop approaches aimed at tackling various diseases.

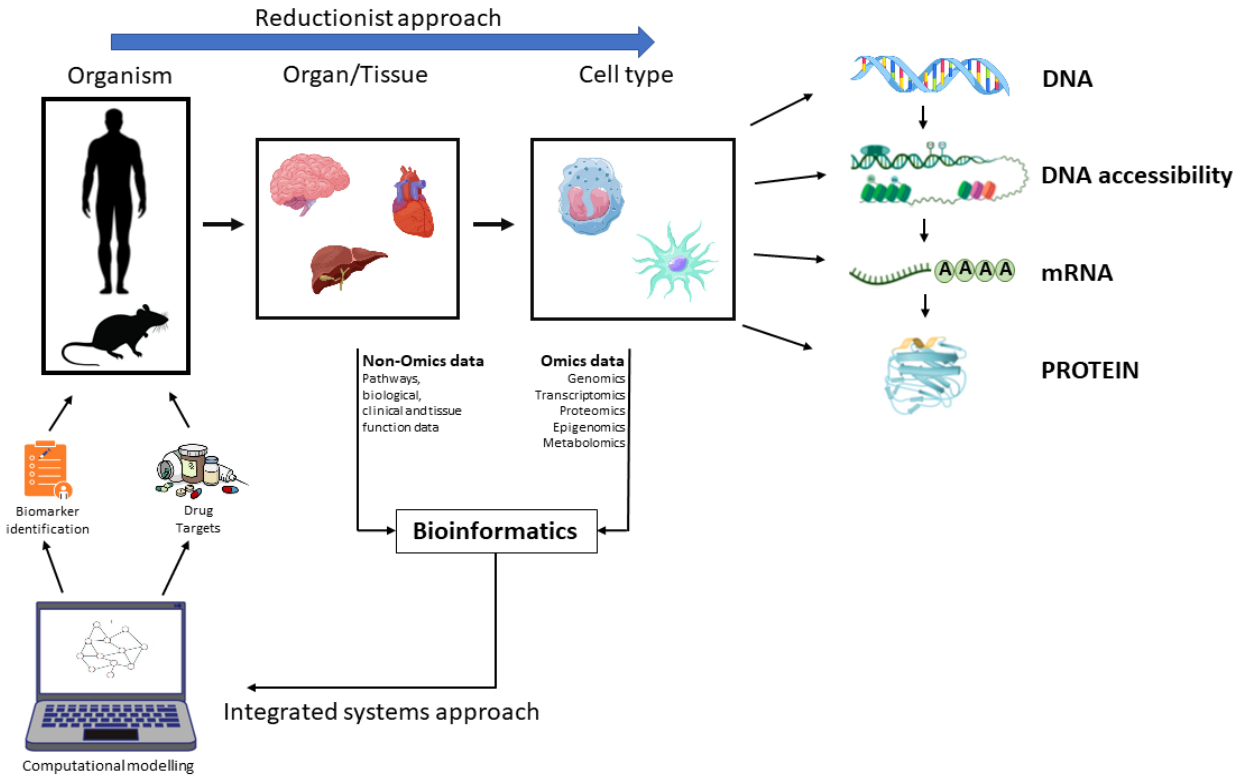


Figure 1. Diagrammatic representation of Systems Biology and reductionist approaches in biology. The figure summarises the difference in approach between the reductionist and systems biology approaches. Under the reductionist approach, the system is broken down into smaller levels and constituents and studied individually. In the systems biology approach, we integrate data from various OMICS technologies with biological knowledge to model biological processes. These models help in the prediction of disease biomarkers and drug targets. (Figure inspired by Schaub et al. 2020)

In cellular biology, this interaction between multiple levels of networks of signalling pathways, gene regulatory networks and cell-cell interaction networks provide a set of highly connected and interdependent networks that share information and studying them in the disconnected form leads to loss of information. Communication between cells is organized by a combination of cell-cell interactions, intracellular signalling and gene regulatory networks (Hunter 2000). Thus, accurate comprehension of cell-cell communication requires a systems approach that relies on the interconnected roles of these different levels of information in cells. A better

understanding of cell-cell communication could aid in the development of strategies for diseases involving multiple cell types by finding the cells that wield a dominant role in the progression and establishment of these diseases (Armingol, Officer et al. 2021).

1.2. Cell-cell communication and its role in cellular function

Cells interact with their environment to receive signals, process them and then provide appropriate responses thus leading to the dissemination of biological information between cells. These interactions are important for their survival, proliferation and regulation of their functions as required. Cell-cell communication provides a collective behaviour to the cells of the tissue and coordinates the development and confers environmental adaptation to cells (Li and Elowitz 2019). The contribution of these communications has been studied under physiological conditions (Ruch 2002, Park, Musson et al. 2017, Valls and Esposito 2022). Ligand-Receptor (L-R) based communication plays a major role in the communication that happens between cells, soluble molecules known as ligands are secreted by the cells and interact by binding with proteins on the surface of cells based on their conjugate structures also known as receptors. This initiates signal transduction downstream via signalling molecules ending in transcription factors (Heldin, Lu et al. 2016). These signals coming from the outside environment to the regulatory network of the cell shape the gene expression of transcription factors and their target genes which provide the functionality to the cells (Wray, Hahn et al. 2003). This regulation of genes also leads to the production of ligands downstream which relay the information back to other cells. Thus, this communication between cells allows sharing of information about the environment.

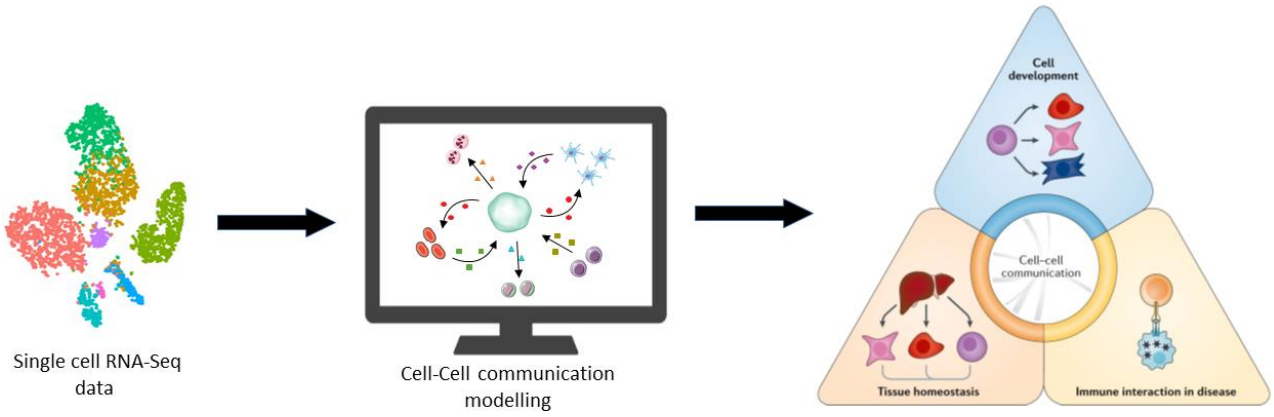


Figure 2. Applications of cell-cell communication modelling using scRNA-Seq data.

The figure showcases the various avenues of research for modelling approaches based on cell-cell communication networks. (Figure modified from Armingol et al. 2020)

The critical balance of these interactions in regulating the specific function of cells and coordinating responses in the case of disease has been widely studied (Huang, Xu et al. 2022, Overton and Mastracci 2022). The balance of these communications is what maintains the physiological state of the cells in the tissue and modelling of cell-cell communication can be crucial in comprehending the role of different cells in various processes (Figure 2). Disturbance of this essential communication between cells can lead to a breakdown in the homeostatic state of the cells by dysregulation of the cellular functional state contributing to disease induction. Thus, understanding which signals between the cells regulates specific changes in the functional states would contribute to specifying the roles of various cell types and cell states in sustaining homeostasis and the development of disease.

Modelling of cell-cell communication allows the unravelling of the role of cell types along with ligands, receptors and TFs that constitute the interactions of these cell types with other cells. Under the condition of disease parts of these communications get altered or new interactions can be established that contribute to the dysfunction of these cells.

1.3. Computational modelling of cell-cell communication

Experimentally cell-cell communication can be studied on a few cell types and a handful of genes due to technical and monetary constraints. Computational modelling aids in the identification of genes playing a crucial role in cell-cell communication. In recent years, the advent of single-cell sequencing technologies has led to rapid growth in the availability of more and more data about the basic processes in biology. This has also led to research in utilising these new and high-resolution data to identify and understand biological function which has ushered a boom in the number of computational tools (D'Argenio 2018). These new computational tools have led to rapid enhancement in our understanding of biological processes (Hill and Gerner 2021). Single-cell RNA sequencing (scRNA-seq) platforms have supported high throughput access to the transcriptome of cells. This has brought an era of development of computational tools based on high throughput sequencing. Transcription is a part of the “central dogma of replication, transcription and translation” and is responsible for the transcription of genes to mRNA which is then translated into proteins that mediate cellular functions.

These tools further utilise the ligand-receptor information to establish the communication between cell types. Since then, several new tools have been published that utilize different types of algorithms to decipher cell-cell communication. Most recent of these tools are CellCall (Zhang, Liu et al. 2021), NicheNet (Browaeys, Saelens et al. 2020), CellChat (Huang, Xu et al. 2022), CellphoneDB (Vento-Tormo, Efremova et al. 2018, Efremova, Vento-Tormo et al. 2020), SpaOTsc (Cang and Nie 2020) and NCEM (Fischer, Schaar et al. 2022) which also represent the state-of-the-art modelling in cell-cell communication. All these tools utilize publicly available databases of Ligand-receptor and intracellular signalling with some also integrating gene regulatory data to further substantiate their results. Of these tools CellphoneDB and CellChat use protein subunit information, taking into consideration multi-subunit receptor complexes while most other tools rely on single ligand-single receptor interactions. While all other tools utilize only transcriptomics data, SpaOTsc and NCEM can integrate spatial information by using spatial transcriptomics to model cell-cell communication.

CellPhoneDB was one of the first published tools to model cell-cell communication. These tools utilize protein complex information to model receptor complex formation. Then based on gene expression of ligands and receptors to calculate communication scores between various cell types. The study was able to recover important ligands-receptors based communications that play a significant role in the development of maternal-fetal communication in early pregnancy (Vento-Tormo, Efremova et al. 2018). The study was further able to shed light on the roles of different cell types in the placenta and decipher the difference in cell communication channels between different subtypes of cells like natural killer cells. Further, the tool provides fundamental information regarding communication that is responsible for immunomodulation in the decidual layers of the placenta. Recently the tool saw an upgrade (Efremova, Vento-Tormo et al. 2020), which enables the reduction of time and memory utilization by the tool and further allows the addition of new interacting molecules to the analysis.

NicheNet, build up on the simple L-R based communication network by integrating signalling and gene regulatory networks. This method provides a model that establishes a relationship of ligands that regulate the expression of a target gene in the GRN. For this purpose, NicheNet uses a prior model of ligand-target along with gene expression data to establish these relations. The tool provides both ligand and target gene activity scores as a way of prioritizing

genes involved in the communication and regulation of cells. The study utilized the head and neck squamous cell carcinoma (HNSCC) tumour dataset to predict the genes involved in the partial epithelial-to-mesenchymal transition (p-EMT) program in malignant cells (Browaeys, Saelens et al. 2020). Recently, NicheNet got an update referred to as Differential NicheNet (Guilliams, Bonnardel et al. 2022) that was used to compare cell-cell interactions across different niches and prioritise niche specific ligands and receptors.

CellChat (Huang, Xu et al. 2022) builds upon previous work by using multimeric complexes as in CellPhoneDB. Then combining it with network analysis and pattern recognition to identify important cell types that provide signals and the coordination between different sender cells. Following this, they can be put together in groups based on the communication computed by CellChat. The tool was applied to embryonic development and wounded skin datasets and the inferred networks were then confirmed using in-situ assays.

CellCall integrates ligand-receptor and TF activity to infer both the inter- and intracellular signalling in the cells (Zhang, Liu et al. 2021). It further embeds the signalling information of the cells by using pathway activity based on the Jaccard similarity coefficient index, to identify crucial pathways involved in the communication. The Jaccard index statistic is used to measure the similarity and diversity of sets. The inference of the cell communication is based on ligand-receptor-TF by combining activity scores of TFs with the L2norm of the L-R interaction. The TF activity score is based on the enrichment of expression of the TF regulons. The tool was implemented on human testicular cells dataset identifying important pathways involved in communication during spermatogenesis, which was further confirmed using immunostaining. The tool was further utilized to identify pathways and TFs, on tumour immune microenvironment (TIME) immune datasets involved in monocyte/macrophage development, activation, polarization and recruitment and The Cancer Genome Atlas (TCGA) pancancer data those associated with patient survival.

Spatially optimal transporting of the single cells (SpaOTsc) tool was designed to utilize gene expression from spatial imaging datasets (Cang and Nie 2020). SpaOTsc provides multiple capabilities – mapping between scRNA-seq and spatial data, inferring spatial distances between single cells, quantitatively comparing spatial gene expression patterns, reconstructing spatial cell-cell communications and identifying gene pairs that potentially regulate each other. The integration

of scRNA-seq and imaging datasets is performed by using the optimal transport algorithm where the transport cost is based on the dissimilarity in the expression of shared genes between the two datasets. The spatial metric for the scRNA-seq data is then calculated by using the cell-cell distance from the spatial data. This metric is used to calculate optimal transport between the distribution of sender and receiver cells. The tool mimics the physical process of ligand release and consumption by senders and receivers respectively to characterize the communication. In this method, the spatial range of signalling is also estimated, by using an ensemble of trees based machine learning model on the initial cell-cell communication network. The tool also infers the direct and indirect influence among genes across different cells by using partial information decomposition. This tool was implemented on datasets from zebrafish embryo, *Drosophila* embryo and mouse visual cortex. These datasets were integrated with spatial counterparts to create cell-cell communication networks. The study utilized data generated from Slide-Seq (Rodriques, Stickels et al. 2019) and RNAseqFISH+ (Eng, Lawson et al. 2019) technologies along with scRNA-seq dataset of mouse olfactory bulb to compare inference of cell-cell communication with and without integration of scRNA-seq data with spatial transcriptomics.

Node-centric expression models (NCEM) was recently proposed to infer cell communication (Fischer, Schaar et al. 2022). The algorithm improves upon ligand-receptor based cell communication inference by using a spatial graph of cells. NCEM is defined as a graph neural network that predicts the gene expression vector of a cell based on its celltype, the niche and its spatial constraints. The study utilizes datasets from multiple spatial assay technologies including MERFISH (Xia, Fan et al. 2019) and 10x Visium from multiple tissues. The tool was able to identify spatial dependencies between cells as well as predict the length scale of interactions based on spatial information.

While these tools have been developed to pursue various avenues of research in cell-cell communication utilising a range of algorithms to define its significance in health and disease. There still is a need for improvement in the level of biological insights that can be provided by these tools.

1.4. Characterizing Functional heterogeneity and niche induced cell states

Multicellular organisms are composed of cells of various types which vary in shape, size, and gene expression all specialized for their function (Prasad and Alizadeh 2019). This specialization in cells is contributed by the cell's lineage and environment. Cell lineage broadly defines the functional role of these cells during development (Hartenstein, Omoto et al. 2021) in addition to this, the environment of the cell shapes the function of the cells. The effect of the environment is conducted in multiple ways on the cells which include vesicles, ligands, physical contact, etc. based on interactions with other cells in the environment. This external communication of the cell leads to changes in gene expression of the cells which further impacts the functionality of the cell (Figure 3).

Cells are typically defined based on their well-defined functions and are known as cell types which may appear similar but still contain an inherent heterogeneity in the gene expression of individual cells (Altschuler and Wu 2010). With the boom in the generation of high throughput single-cell data, this heterogeneity of cells has become much more accessible to study and thus research in the role of functional states of cell types has gained more interest. In biological systems, heterogeneity contributes to the robustness and adaptive response (Kucharavy, Rubinstein et al. 2018). Many cell types show plasticity in their function which corresponds to the different roles the cells must play under various conditions. This heterogeneity of cells has been known to play a crucial role in the proper function of cells which for meaningful biological states can be defined by multiple dimensions like cell cycle stages (McDavid, Finak et al. 2016), gene expression modules, etc.

Tissue resident specific differences in the expression of cell types across tissues further consolidate the idea of niche based functional specialization of the cells. Cells residing in different tissues are exposed to varying microenvironments known as the niche of the cell. This niche is composed of other cells of the tissue and cells interact with each other thus leading to a different combination of interactions for every single cell. These cells further differ in the expression of genes that compose the signalling of the cells and contribute to different transcription factor profiles across cells of the same cell type. These differences may arise from spatial or temporal differences as well as conditions of the tissue. For example, immune cells like macrophages and dendritic cells can be found across various tissues of the cells and based on the conditions of niche can acquire cell states of M1/M2 for macrophages (Zhang, Yang et al. 2021) and DCs could be in

activated/quiescent states or other known subsets (Boltjes and van Wijk 2014). While immune cells as they travel can acquire peripheral and tissue resident specific characteristics also. Moreover, the developmental process in cells also contributes to heterogeneity, as development is a continuous process a celltype can contain cells at different stages of development such as mature and immature cells. Thus, cell-cell communication plays an essential role in shaping the functional states of cell types based on their niche.

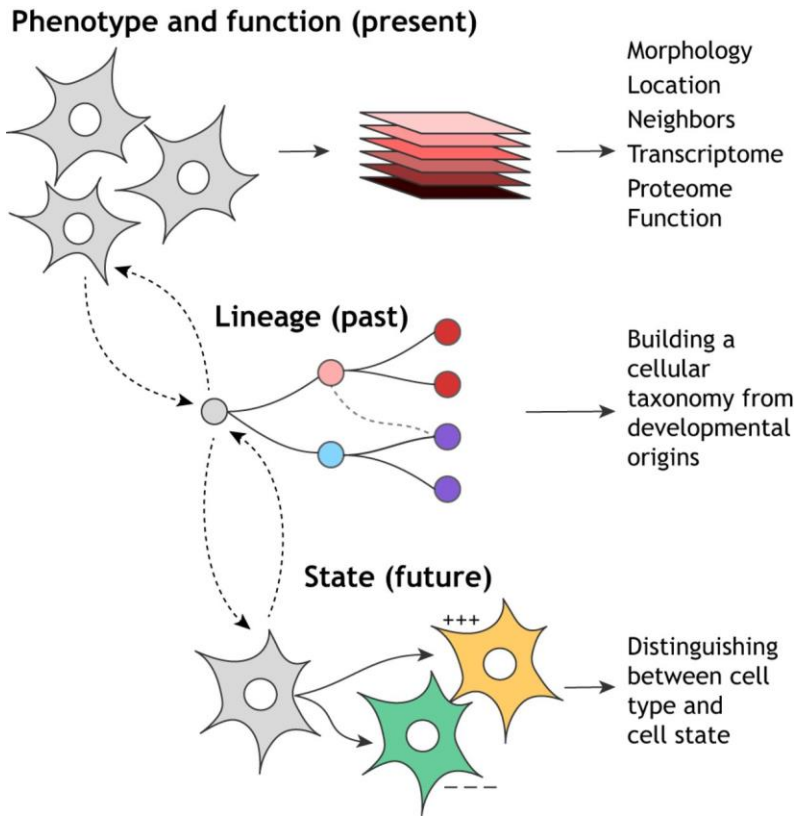


Figure 3. Pillars of identity This figure demonstrates the role of lineage and environment on the phenotype and function of cells. (Figure from Morris 2019)

The role of this cellular heterogeneity also becomes quite apparent under conditions of disease and ageing. During ageing, induction of senescence (Mylonas and O'Loghlen 2022) or quiescence (Tumpel and Rudolph 2019) plays a detrimental effect as it reduces the number of stem cells available to replenish the pool of functionally active cells. In diseases, like infections, the immune cells shift towards inflammatory cell states which are required to clear the infections, while at the same time under cancer the tolerogenic cell states are relevant as they allow cancerous cells to escape the immune system by suppressing its response via anti-inflammatory molecules.

This heterogeneity of cells and cell states. These examples underline the impact of cell-cell communication in both healthy and disease conditions thus making it important to study their existence and roles.

1.5. Identifying impacts of perturbation in cell-cell communication

Development is a highly synchronized process that requires precise combination and cooperation between signals from the environment and the intrinsic gene expression (Li and Elowitz 2019). During the process of differentiation, cells change their morphology and cellular functions, based on the changes in gene expression that are required as part of the developmental process. These genes consist of key TFs and their regulated targets constituting the gene regulatory networks that govern the cell specification. Perturbations in the environment could lead to dysregulation in the developmental process due to changes in the GRN and have been extensively studied as they provide valuable information about the role of specific environmental signals and their effects on genes involved in development (Levine and Davidson 2005, Weidemuller, Kholmatov et al. 2021). During development, the regulatory changes in the GRN are not limited to the level of gene expression but also at the epigenomics level as explained below.

Regulation of gene expression happens at two levels, first, the expression of genes is strictly controlled by the binding of transcription factors in the gene's promoter region, and the combination of various factors binding in the promoter controls the level of a gene expressed in a particular cell. Secondly, the expression is controlled at the epigenetic level by chromatin accessibility, methylation and acylation patterns which dictate the possibility of TFs binding to the promoter region by controlling the steric hindrance in the binding regions. Epigenetic patterns are important throughout the development of cells as they are specific to certain lineages supporting the idea that certain sets of genes provide lineage specific roles to cells (Kluger, Tuck et al. 2004, Rommelfanger and MacLean 2021). Also, as cells move from pluripotent to multipotent and differentiated cells the epigenetic patterns for other lineages become inaccessible for TF binding (Cedar and Bergman 2008, Boland, Nazor et al. 2014). These factors shed light on the need to study both gene expression and epigenetics simultaneously in an integrated way. Moreover, this method provides insights into the changes that occur during development or cases of perturbations to these systems leading to changes in the functional state of cells. While changes in gene expression lead to rapid changes in the cells as a response to environmental perturbations the effect

on the epigenetic markers could lead to more long term ramifications for the cells during their development.

Communication between cells is a well-regulated process and small changes in the concentration of ligand molecules could lead to dramatic changes in the cellular functional states that could contribute to loss of homeostasis at the tissue level. These changes occur as different molecules convey different information regarding the environment to the cells (Turner, Nedjai et al. 2014, Liu, Chu et al. 2021). These signals are produced by different cells in the environment as part of the response to changes like injury or disease and are meant to prepare the cells for an appropriate response to them. This communication between cells during injury/disease may lead to inflammation. Inflammation is an intrinsic mechanism that plays a significant role in the regulation of development, homeostasis and immune response (Bennett, Reeves et al. 2018, Chen, Deng et al. 2018, Dostert, Grusdat et al. 2019). Inflammation is mediated via a specific set of cytokines that bind to their cognate receptors and activate a different set of downstream signalling that cascade into changes in the expression of the genes that are at the core of gene regulatory networks of the cells. These changes in GRN are meant to activate anti-inflammatory responses for the cells as well as produce signals downstream that further communicate the information of environmental inflammation for other cells in the environment. As these cytokines linked with inflammation can create functional changes in cells by perturbing the gene expression of key genes thus making it essential to study their role in biology.

The role of cytokines is vital during development, in the orchestration of specific changes in gene expression as the cell develops and changes its functional nature of the cell (Robb 2007, Metcalf 2008). For example, during the development of immune cells, specific signals are required at certain time points during the development of the cells which then leads to the proper development of functional cells (Banyer, Hamilton et al. 2000, Arango Duque and Descoteaux 2014, Wu, Xie et al. 2014). Perturbation in these communication signals during any stage of development can alter the intrinsic gene expression of the genes that are required for the cell's specific functions (Jiang, Zhang et al. 2021).

Previous studies have inquired into the role of inflammatory cytokines in disease (Kany, Vollrath et al. 2019). These studies have showcased that inflammatory molecules if encountered by the cells when not essential for physiological response or development can lead to dysfunction

in the developing cells thus also affecting other cells in the environment that rely on the communication from these cells for maintaining their functional states. This dysfunction during development has also been involved in various disease conditions (Okin and Medzhitov 2012) and research in cytokine perturbation during cellular development can provide new avenues to study disease induction and thus help in the discovery of therapeutical targets for diseases.

1.6. Feedback loops based regulation in cell-cell communication

Connecting output signals back to inputs leads to the creation of feedback loops and has been recognized as an essential concept of biology playing roles in homeostasis, pattern formation, transcriptional regulation and recently as part of intercellular signalling as they account for the ability of system to continuously regulate itself (Brandman and Meyer 2008). Intercellular communication creates a complex network between cells and the presence of multiple feedback loops in these networks points to their role in communication.

Feedback loops have been identified in various tissues across cell types and have been studied for the precise role they play in maintaining specific functions in the tissue. In the heart, interactions between cardiomyocytes, fibroblasts and endothelial cells using multiple growth factors like Vascular endothelial growth factor (Vegf) and Neuregulin 1 (Nrg1), shape coronary vessel growth and maintenance (Tirziu, Giordano et al. 2010). Similarly, stromal cells which include fibroblasts, myofibroblasts, bone marrow stromal cells, etc interact with immune cells across tissues via various cytokines including Interleukin 6 (IL-6) playing role in both homeostasis and coordination of immune response (West 2019). In the liver, hepatocytes form feedback loops with liver sinusoidal endothelial cells and hepatic stellate cells using growth factors regulating the cellular phenotypes in homeostatic conditions (Iwakiri, Shah et al. 2014, Weiskirchen and Tacke 2014). In the brain, interactions between neurons and microglia indicate their role in the regulation/maintenance of microglial activation and synaptic development of neurons via feedback (Eyo and Wu 2013), microglia also interact with astrocytes for the regulation of immune functions in the brain (Matejuk and Ransohoff 2020). In the immune system both the innate and adaptive response is dependent upon interactions between various immune cells such as macrophage, dendritic cell, natural killer cell, T cell and regulatory T cell, etc to regulate the response to injury/disease (Malyshev, Manukhina et al. 2014, Rahman, Tiwari et al. 2018).

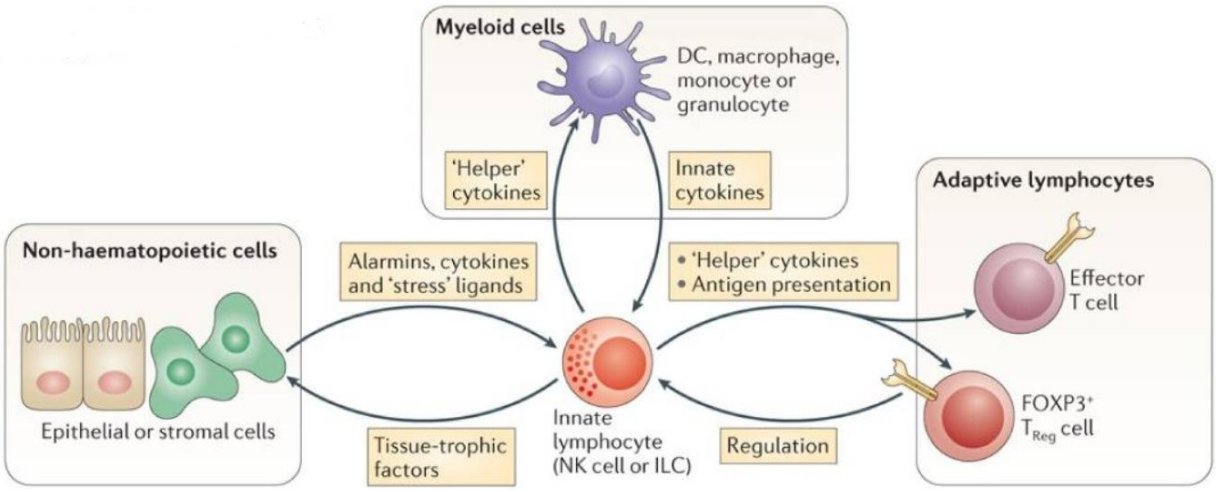


Figure 4. Interaction between cells from the innate and adaptive immune system with other cell types. (Figure from Gasteiger and Rudensky 2014)

The immune system is meant to provide a robust response in case of infection or injury and plays an essential role in the communication between cells to regulate appropriate responses during any such event (Figure 4). This response is tightly regulated as the immune system consists of cells that play roles in inducing inflammation during the early phase of the response and other cells that can induce tolerance in the system once the issue with the injury/infection is resolved by the removal of inflammatory molecules or producing more tolerogenic molecules. Thus, the balance between these two types of immune cells and the level of expression of pro-/anti-inflammatory molecules creates feedback of regulation between cells. This regulation is both spatial and temporal as specific inflammatory signals are required to start the response which is then taken over by signals that are meant to maintain the state of inflammation till the resolution of the injury/infection. This removes the signals in the environment that initiated the inflammatory immune response and inflammation is steadily removed from the system by the degradation of pro-inflammatory molecules over time. Similarly, to build tolerance there needs to be a communication of tolerogenic signals that could create feedback loops between cell types and thus create a state of tolerance.

In healthy individuals or animals, these feedback loops consist of ligands and receptors involved in the basic function of these cells like growth and development. While in the case of disease, these signals from the tissue resident cells and immune cells are involved in feedback loops that support the dysfunctional cell states of the cells. As in the case of allergic diseases, the

systemic response from the immune system after encountering specific allergens. This hypersensitive response to allergens is accompanied by inflammation which is shaped by the ligands and receptors creating feedback loops between different immune cells that sustains this state of inflammation in the tissues (Galli, Tsai et al. 2008). To counter the inflammatory effects of allergic diseases various therapies have been developed that attempt to cure the disease by developing a tolerance state in the tissues (Wisniewski, Agrawal et al. 2013). To accomplish this, it is important to understand the role of different regulatory cells and the feedback created between them which could then sustain this state of tolerance.

Modelling of feedback loops in cell-cell communication is an important and new avenue of research into the role of different cell types in both health and disease. As these feedbacks are responsible for regulating the overall state of the tissues their better understanding could aid in deciphering the molecules that are involved. This allows the development of better therapeutic targets for diseases by combining the various levels of complexities in the biological system as well as considering that it is important to study a composite picture of both health and disease as it involves the coordination and cooperation of many cell types and hundreds of molecules that work together to make the tissue and regulate its functions.

2 Scope and aims of the thesis

2.1 Scope

This Ph.D. thesis is focused on using systems biology approaches to decipher the role of cell-cell communication in health and disease. Systems biology is an emerging discipline that helps better our understanding of complex biological processes by integrating the various networks and organizational levels in biology. This thesis includes the development of computational methods and analysis of newly produced datasets to discover and understand the role of cell-cell communication in the regulation of both functional cell types and tissue homeostatic states. These tools allow us to better understand the underlying role of different molecules involved in this communication. Using these tools we study the changes that functional cell states, the effect of signal perturbations and the role of feedback loops in maintaining the state of a tissue. The development of these novel computational tools using single-cell technologies along with the integration of various networks in biology allow for a more holistic approach to modelling of health and disease.

2.2 Aims

These Ph.D. projects are aimed at developing computational tools and studying newly produced datasets to model health and disease to improve our understanding of the role of cell-cell communication in regulating cell and tissue states with the intent of discovering therapeutic strategies against diseases.

Aim 1: **Development of a computational method to resolve functional cell states using cell-cell communication model.** With this study, we focused on developing a novel tool of cell-cell communication that resolves functional cell states based on signals from the tissue niche. Several studies have showcased that cells in a celltype show functional heterogeneity which gets shaped by the signals coming to these cells from their niche. Therefore, we aimed at developing a cell-cell communication model by integration of signalling, gene regulatory and ligand-receptor networks to characterize these functional states. In addition, we aimed at studying cell types and states across young and old tissue atlases to study the differences and similarities between them.

Aim 2: **Study the development of astrocytes under inflammatory by Integrating transcriptional and chromatin accessibility.** In this study, we first characterized the developing astrocytes using gene expression and chromatin accessibility in both normal and inflammatory conditions at global and individual genes levels. We explored the changes that occur during development and the impact of inflammatory conditions during development. Our main focus remained on the integration of transcriptomic and chromatin accessibility data to generate the gene regulatory networks of astrocytes developing from neural stem cells under normal and inflammatory conditions to discover the roles of various genes during the developmental process. The study provides a preliminary step in the development of therapeutical targets for neurological diseases with an inflammatory component.

Aim 3: **Development of a model of feedback loops in cell-cell communication network to discover key molecules in allergy and tolerance.** With this study we aimed at developing a novel computational tool that first builds a cell-cell communication model network in different tissues and then discovers feedback loops between the cells involved in communication. Several studies have shown the role of these feedback loops in regulating the state of the tissue thus we applied the tool to a novel single cell dataset of immune cells during allergy and tolerance conditions of the tissue generated under the project. We leveraged these feedback loops to discover novel therapeutic targets for allergic inflammation and tolerance induction.

2.3 Originality

The studies presented in this thesis were designed to decipher the roles cell-cell communication plays in shaping cell and tissue states. In the first study, we developed a novel tool to identify cell sates using reconstructed cell-cell communication networks. In the second study, we generated transcriptomic and epigenomic data for developing astrocytes under normal and inflammatory conditions. Our analysis provided valuable insight into the impact of inflammation on development. In the last study, we generated single cell resolution gene expression data from immune cells to study allergy and tolerance in a mouse model. Under this project, we also created a new computational tool that identifies key regulators of allergic inflammation and tolerance induction using feedback loops in cell-cell communication networks. The focus of these studies was to improve our understanding of biological processes and aid in development of new therapeutic approaches against diseases.

3 Materials and Methods

Materials and methods details are presented in the results section of this thesis for the three manuscripts (sections 4.1 to 4.3). A summary is described below for each of them –

In "FunRes: resolving tissue-specific functional cell states based on a cell–cell communication network model" (section 4.1) we developed a cell-cell communication network based model to resolve functional cell states in cell types. The cell-cell communication network is reconstructed in a bottom-up approach for individual celltypes, at first identifying the conserved set of transcription factors. The signalling network upstream of these transcription factors is then modelled based on a Markov chain to identify sustained signals in the celltype. Next, the transcription factors via these signalling networks are then connected upstream to the set of conserved receptors expressed by the celltype. The cell-cell communication network is then established using a scaffold of ligand-receptor interactions collected from the literature. The communication network is pruned, and only high confidence interactions are retained as the final output by comparison with the background of all interactions. The final output is then used to establish combinations of receptor and transcription factors that are conserved together in a significant fraction of the celltype population and are involved in the communication network via the receptor providing sustained signals to the transcription factor. The cell states are resolved by creating an incidence matrix of the cells in columns and receptor-TF combinations as rows. Hierarchical clustering is then performed on the incidence matrix to resolve states represented by a split in columns and rows which represent different states and processes based on their similar profiles. Functional enrichment (Yu, Wang et al. 2012) is then performed on the row clusters in the incidence matrix to identify the biological processes that differentiate the cell states. Markers were then identified in these states using Seurat's *FindMarkers* function.

We implemented the tool on the *TabulaMurisSenis* atlas to identify these cell states across tissues from young and old mice. Literature evidence was then used to establish support for the cell states identified using the tool. Comparative analysis was performed with state-of-the-art tools by identification of cell states to demonstrate the advantage of using cell-cell communication in establishing and understanding the cell states induced by the niche.

In “Transcriptional and chromatin accessibility profiling of neural stem cells differentiating into astrocytes reveal dynamic signatures affected under inflammatory conditions” we studied the transcriptional and chromatin accessibility changes during the development of astrocytes under normal and inflammatory conditions. The Neural stem cells (NSCs) differentiating into astrocytes were first characterized using qPCR to check for known astrocytic markers under both conditions. Bulk RNA-Seq and ATAC-Seq data were generated at 24h and 72h timepoints for both conditions along with control (0h). The global gene expression and chromatin accessibility were then studied to identify changes during normal development and under inflammation. The changing transcriptional and accessibility profiles were also explored using functional enrichment algorithms. Further, the gene expression changes of genes involved in astrocytic functional genes were also examined. We performed transcription factor binding motif analysis on the chromatin accessibility data using the HOMER (Hypergeometric Optimization of Motif EnRichment) tool. Through this analysis, we identified Nfe2l2 (aka Nrf2) as an important TF involved in response to inflammation. Finally, we reconstructed the GRNs for developing astrocytes under both conditions for both 24h and 72h timepoints by performing contextualisation (GRNOptR) over the networks created using differentially expressed genes. Topological analysis was then performed on the GRNs to identify key regulators of the development of astrocytes under normal and inflammatory conditions.

In “Feedback loops in cell-cell communication identify key regulators of allergy and tolerance”, we utilized an AIT (Allergen Immunotherapy) model of the mouse to identify key regulators of allergy and tolerance by reconstructing the cell-cell communication network between a subset of immune cells involved in the allergy and tolerance state of the tissue. We generated single cell RNA (scRNA-Seq) sequencing data for specific immune cells (cDCs, pDCs, B cells, NK cells, Regulatory T cells and effector T cells) that have been characterized before as playing significant roles in the development and resolution of allergy as well as the induction of tolerance in a tissue. The cells were extracted from the Spleen and Peritoneal cavity of the mouse at different stages of the experiment.

Downstream analysis of pre-processing and quality control was performed using the Seurat package in R, which was further utilized to find cell clusters in the data. The cell clusters were then characterized using known markers. After annotation, the cell-cell communication network

was reconstructed for individual tissues and timepoints. For the reconstruction of the cell-cell communication networks, we utilized our recently published tool FunRes, which reconstructs the network in a bottom-up approach of identifying the conserved transcription factors and then linking them upstream by using a Markov chain model of the signalling network. Finally, the cell-cell communication network is established by using conserved receptors and cognate ligand-receptor information. We then applied our new computational tool C3Loop for feedback analysis on the cell-cell communication networks. The tool first identifies the feedback loops in the network and then performs a sensitivity analysis on loops by varying various parameters in the tool. The sensitivity analysis allows capturing molecules that are consistent even while the variation of parameters changes the feedback loops thus providing high confidence regulator genes. The tool then performs a comparative analysis between conditions to discover molecular targets against allergic inflammation and tolerance induction. We used literature based evidence to verify the effectiveness of our tool by accessing the molecules selected and the role they could play in allergy and tolerance.

4 Results

4.1 Characterization of functional heterogeneity and niche induced cell states

4.1.1 Preface

This manuscript entitled “FunRes: resolving tissue-specific functional cell states based on a cell–cell communication network model” was published in Briefings in Bioinformatics in July 2021 accessible with DOI: 10.1093/bib/bbaa283. The pre-print version has been presented in this thesis and the supplementary data is accessible on the publisher’s website.

In recent years studies have shed light on the role of tissue microenvironments in the functionality of cells. The tissue microenvironment or niche is made up of all the cells constituting the tissue. These cells communicate with each other using molecules that bind to cognate receptors on other cells creating networks of communication between different cell types. Cell-cell communication plays a vital role in shaping the function of cells and the differences in the interactions between the cells contribute to the heterogeneity of the cells. These differences in interactions could also lead to changes in the function of cells that could be important to study and understand their role across cell types.

In recent years, many computational tools have been developed to discover novel cell subtypes but they have not taken into consideration the role of ell-cell communication. Understanding the role of cell-cell communication in the maintenance of cell subtypes could be important in discerning the role of specific functional cell subtypes found in diseases.

This study is the first proposed tool to resolve functional cell states using an underlying cell-cell communication network. The method uses a bottom-up approach to building cell-cell communication using conserved expression of transcription factors. The tool utilizes a Markov chain based model of signaling networks to identify sustained signals in the cells. These conserved transcription factors are then connected upstream to the conserved set of receptors via the sustained signals identified. Using a scaffold on ligand-receptor interactions which were manually curated using literature evidence the cell-cell communication network is reconstructed in the tissue. The tool was applied to 177 celltypes of 10 tissues from the Tabula Muris Senis atlas and disease datasets. This dataset contains cells from young and old mice helping in the identification of functional cell types. The application of the tool on the dataset allowed the characterization of emerging and vanishing cell states in aging and disease. The cell states were then evaluated, and

their possible functional roles were accessed based on the molecules involved in the cell-cell communication and known celltype markers. Further, the tool was compared with the state-of-the-art clustering tools that are meant to identify cell clusters in data and showcased the tool identified the functional cell states more accurately, whereas the other tools were not satisfactory. Finally, the tool and its code were made publicly available to the community in a user-friendly way.

Contribution: I implemented the computational algorithm, collected the data, processed the data, performed the analysis, and contributed to the manuscript and its figures.

4.1.2 Published Manuscript

FunRes: resolving tissue-specific functional cell states based on a cell–cell communication network model

Sascha Jung[1], Kartikeya Singh[2] and Antonio del Sol[1,2,3]

[1]Computational Biology Group, CIC bioGUNE-BRTA (Basque Research and Technology Alliance), Derio, Spain

[2]Computational Biology Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

[3]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Abstract

The functional specialization of cell types arises during development and is shaped by cell–cell communication networks determining a distribution of functional cell states that are collectively important for tissue functioning. However, the identification of these tissue-specific functional cell states remains challenging. Although a plethora of computational approaches have been successful in detecting cell types and subtypes, they fail in resolving tissue-specific functional cell states. To address this issue, we present FunRes, a computational method designed for the identification of functional cell states. FunRes relies on scRNA-seq data of a tissue to initially reconstruct the functional cell–cell communication network, which is leveraged for partitioning each cell type into functional cell states. We applied FunRes to 177 cell types in 10 different tissues and demonstrated that the detected states correspond to known functional cell states of various cell

types, which cannot be recapitulated by existing computational tools. Finally, we characterize emerging and vanishing functional cell states in aging and disease and demonstrate their involvement in key tissue functions. Thus, we believe that FunRes will be of great utility in the characterization of the functional landscape of cell types and the identification of dysfunctional cell states in aging and disease.

Introduction

In multicellular organisms, the functional specification of cell types arises during development and is further shaped by signals from other cells. In particular, the exchange of these signals through receptor–ligand-mediated cell–cell communication networks determines a distribution of different functional cell states that are collectively relevant for tissue functioning [1]. Thus, cell types are composed of a variety of functional cell states that are shaped by their tissue environment. More specifically, in response to different stimuli, cells of the same cell type can exhibit different phenotypes defined by physical, molecular and functional characteristics, called functional cell states. Great efforts have been devoted to characterize tissues pecific functional differences imparted by cell–cell interactions, which has led to the identification of functional cell states in various cell types. For instance, comparison of tissue-resident macrophages in different organs revealed functionally relevant differences in their gene expression programs [2]. While ileal and colonic macrophages showed higher expression of CD74 compared to other tissue-resident macrophages, TGFB2 is exclusively expressed by peritoneal macrophages [2]. Moreover, transplantation of these macrophages to other organs reprograms their expression profile towards tissue-resident cells, which demonstrates that the observed differences are largely imparted by their environment [2]. In addition, previous studies investigated the effect of the environment on the functional specification of stem cells with respect to their activity and the tissue compartment in which they reside. For example, neural stem cells in the subgranular and subventricular zones show subtle phenotypic differences driven by the niche. While subgranular zone stem cells express the transcription factor (TF) HES5, which is induced by Notch signaling, subventricular zone stem cells express Id proteins, which are induced by BMP signaling, to maintain their function [3, 4]. Although these studies have enabled the characterization of certain functional cell states determined by cell–cell interactions in specific cell types, the development of computational

methods would greatly aid the systematic identification and characterization of such states. The identification of cell types and subtypes has been traditionally performed by molecular biology approaches based on the shape and size of cells, while, more recently, these populations have been characterized by their expression of cellular membrane proteins. However, due to the limited number of surface proteins, it is likely that important functional differences between cell states cannot be described by combinations of these proteins alone. Advances in single-cell RNA sequencing (scRNA-seq) technologies have provided an unprecedented view on the cellular heterogeneity within individual cell types. However, the exploitation of these datasets requires computational tools for processing and clustering the data. A plethora of computational tools have been developed in recent years for the unsupervised clustering of scRNA-seq profiles [5–8]. Although these tools have enabled the identification of several novel cell subtypes, they do not specifically consider the effect of cell– cell interactions on tissue-specific functional processes, and therefore fail in resolving the heterogeneity in functional states of different cell types. In order to address this challenge, we present FunRes, a computational method for the identification of tissue-specific functional cell states that are induced by receptor–ligand-mediated cell–cell interactions. FunRes employs single-cell RNA-seq data to reconstruct the cell–cell communication network among cell types. In particular, FunRes exclusively detects interactions that are functionally relevant to each cell type by warranting the compatibility of the signal with the intracellular signaling and transcriptional network. Next, FunRes leverages this network to identify functional cell states by partitioning cell types based on cell–cell interactions and the functional annotation of downstream target TFs. We applied FunRes to 177 cell types in 10 different tissues and demonstrate that the detected states correspond to known functional cell states of various immune and non-immune cell types, including macrophages, NK and endothelial cells, which cannot be recapitulated by existing computational tools. Moreover, comparison of the identified functional cell states of the same cell types in different tissues showed conserved states carrying out essential cell-type functions as well as the existence of states with unique functionalities. Finally, we assessed the effect of aging and disease on the composition of functional cell states. Although the overall number of functional cell states does not significantly change, certain cell types, such as enterocytes of the large intestine, hepatic sinusoidal epithelial cells and pancreatic alpha cells, display vast differences in their composition of functional states. Moreover, we validate the detected functional cell states in aged and pathologic tissues by providing evidence

for their involvement in pathological tissue functions. In summary, FunRes constitutes the first computational method specifically designed for identifying functional cell states induced by receptor-ligand mediated cell–cell interactions and complements current clustering methodologies that successfully detect cell types and subtypes. We demonstrate that FunRes is applicable to a wide range of tissue conditions and can accurately detect functional cell states. Thus, we believe that FunRes will be of great utility in the characterization of the functional landscape of cell types and the identification of dysfunctional cell states in aging and disease.

Material and Methods

Assembly of cell–cell communication scaffolds, intracellular signaling and TF-gene regulatory interactions

A cell–cell communication scaffold was generated for human and mouse on the basis of a previously published dataset including manually curated, validated and predicted intercellular interactions (Figure 1A) [9]. These interactions have been collected from different databases, i.e. DLRP [10], HPMR [11], IUPHAR [12], HPRD [13] and String DB [14] or manually validated in previous studies. The interactions were further filtered based on UniProt [15] annotations to include only ligands annotated to be 'Secreted'. Since the original dataset only consists of human data, these interactions were mapped to mouse orthologs using BioMart from Ensembl [16]. The intracellular signaling network is composed of pathway interactions included in Omnipath [17], Reactome [18] and MetaCore from Thomson Reuters. In particular, all pathways from MetaCore were obtained including all signal transduction interactions while discarding transcriptional gene regulatory interactions. MetaCore objects were mapped to gene symbols using the provided mapping table. In case MetaCore objects of the regulator or regulated gene mapped to multiple gene symbols, all possible interactions between the associated gene symbols were generated using a mapping table provided by MetaCore. Gene regulatory interactions were obtained from MetaCore from Thomson Reuters, a manually curated resource of gene– gene interactions, on 01 April 2019 for human and mouse genes. Only transcriptional regulatory interactions with known effects, i.e. activation or inhibition, were selected by filtering for 'direct interactions' with reported effects 'activation' or 'inhibition'. MetaCore objects were again mapped to gene symbols using the provided mapping table. As in case of the intracellular signaling network, if MetaCore objects

of the regulator or regulated gene mapped to multiple gene symbols, all possible interactions between the associated gene symbols were generated using an in-house script.

Selection of preserved TFs

To select preserved TFs, FunRes identifies transcription factors that are expressed in at least a user-defined fraction of cells. For the analysis presented in this manuscript, we selected a permissive cutoff of 10% for all samples. For that, FunRes transforms the expression data into a binary format in which TFs with at least one count become '1' while not-expressed TFs become '0' and aggregates the binary data by cell type. Finally, FunRes selects in each subpopulation the TFs that are expressed in the top five-percentile of cells.

Detecting receptors inducing preserved TFs

To detect receptors inducing the expression of the selected TFs, FunRes employs a Markov Chain model of intracellular signaling, called SigHotSpotter, to identify high-probability intermediate molecules (Figure 1A) [19]. In brief, SigHotSpotter uses singlecell RNA-seq data of a subpopulation and the assembled intracellular signaling network to create a state transition matrix representing the traversal of a signal through the network. The state transition matrix represents a finite discrete Markov Chain and is subsequently evolved to create the stationary distribution. The stationary distribution displays the intermediate molecules exhibiting the highest steady state probabilities. Afterwards, SigHotSpotter connects the intermediate molecules to interface TFs, i.e. the first transcription factors in the signal transduction chain, and characterizes the compatibility with their downstream targets. Calculating all shortest paths from highprobability intermediate molecules to downstream preserved TFs defines a compatibility score that classifies each molecule as being active or inactive. Here, a high-probability gene is deemed compatible with its downstream target if the expression of the gene and target TF agrees with the sign of the interaction path, i.e. an even number of inhibitions is an activating path while all other paths are inhibiting. In case of an activation, the intermediate and target genes have to be expressed whereas, in case of an inhibition, the target gene must not be expressed. Next, a gene is compatible, if a significant number of its targets is compatible. Significance is assessed using a hypergeometric test with P-value cutoff 0.05. Following the same rationale, receptors are identified that target the compatible high-probability intermediates.

Inference of the cell–cell communication network

The main algorithm consists of four steps. First, preserved TFs and the receptors regulating them are selected in each cell population, as described before. Second, ligands expressed in a userdefined fraction of cells are selected in each cell population. For this study, a fraction of 10% was selected. Third, ligand–receptor interactions are established between two cell populations if (i) the receptor was selected in the first step for the first population, (ii) the ligand was selected in the second step for the second population and (iii) the receptor-ligand interaction is contained in cell–cell communication scaffold. Every interaction between receptor r in subpopulation p1 and ligand l in subpopulation p2 is augmented with an interaction strength $s_{r,l,p1,p2}$ defined by:

$$s_{r,l,p1,p2} = \left(\frac{1}{|\{x_{r,p1} | x_{r,p1} > 0\}|} \sum_{\{x_{r,p1} | x_{r,p1} > 0\}} x_{r,p1} \right) \cdot \left(\frac{1}{|\{x_{l,p2} | x_{l,p2} > 0\}|} \sum_{\{x_{l,p2} | x_{l,p2} > 0\}} x_{l,p2} \right) \quad (1)$$

Informally, the score is the product of average non-zero receptor expression values and average non-zero ligand expression values with respect to the receptor r and its expressing population p1 as well as to the ligand l and its expressing population p2. Significance of each interaction is determined by comparing the score of an interaction between two cell types against a background distribution of scores between these cell types based on all interactions in the scaffold. Interactions in the 90th percentile that have at least one significant downstream TF target are retained in the final cell–cell communication network.

Identification of functional cell states

After inferring the cell–cell communication network, FunRes creates an incidence matrix of receptor/downstream TF expression for each cell type (Figure 1B). In particular, each cell in the cell type under study is represented by a binary vector in which each entry corresponds to a receptor/downstream TF pair. If both the receptor and downstream TF are expressed in a cell, the corresponding entry in the vector will be ‘1’ and ‘0’ otherwise. Hierarchical clustering is performed on the resulting incidence matrix using the ‘hclust’ R function with Euclidean distance and complete linkage. The optimal number of clusters is computed based on the Dunn index, an internal cluster evaluation metric comparing within-cluster with the between-cluster distances.

Finally, Gene Ontology [20, 21] enrichment of biological processes is performed using the DOSE R package (Figure 1B) [22]. Terms with false discovery rate lower than 0.01 are considered significant. Clusters having identical functional annotations are subsequently merged. The resulting clusters constitute the identified functional cell states. Importantly, FunRes does not require the proteins within the enriched GO terms to be identical, but only the terms themselves.

Comparison with state-of-the-art clustering methods

We selected Seurat [5, 6], SC3 [7] and SINCERA [8] for assessing their ability to detect functional cell states. All methods were employed in a standard workflow with default parameters in R. In particular, for Seurat, the 2000 most variable genes were detected using the 'FindVariableFeatures' method with selection method 'vst'. Subsequently, the data were scaled ('ScaleData' function) and principal component analysis (PCA) ('RunPCA' function) was performed. Finally, the data were clustered by employing the 'FindNeighbors' and 'FindClusters' functions on the first 10 principal components. SC3 was invoked using the 'sc3' function and the maximum number of clusters was set to 10. SINCERA was invoked without specific pre-processing steps. All analyses were carried out in R v3.6.1.

Results

Identification of functional cell states based on cell–cell communication networks

For the purpose of identifying niche-induced functional cell states, FunRes initially reconstructs the cell–cell communication networks among all cell types in a dataset (Figure 1A). In particular, it integrates transcriptional and signaling networks with extracellular ligand–receptor interactions following a bottom-up approach by first selecting transcription factors (TFs) whose expression is preserved across cells. Next, FunRes identifies receptors that regulate these preserved TFs by employing a previously introduced Markov Chain model that assesses signal transduction probabilities from receptors to their co-expressed TFs [19]. Finally, FunRes finds cognate ligands for these receptors from a previously curated set of ligand–receptor interactions [9]. Based on a previous study about the proportions of functional cell states in heterogeneous T cell populations, we required throughout this study both the ligands and receptors to be expressed in more than 10% of cells of the secreting and receiving cell population, respectively [23] After reconstructing the functional cell–cell communication network, FunRes partitions each cell type into different

functional cell states based on the identified cell–cell interactions and their downstream TF targets (Figure 1B). More specifically, hierarchical clustering is performed to group cells by their function based on the downstream effect of cell–cell interactions they are participating in. Finally, the optimal number of functional states in each cell type is determined by evaluating cluster consistency. Notably, the optimal number of states could be equal to one, which results in no partitioning of the cell type under consideration, suggesting no niche-induced functional heterogeneity.

Functional evaluation of niche-induced cell states

We employed FunRes to identify niche-induced functional cell states of 177 cell types in 10 mouse tissues from Tabula Muris Senis [24]. On average, each cell type was partitioned into 2.49 functional states (median: 2), with a maximum of nine identified states in enterocytes of the intestinal epithelium and pancreatic alpha cells (Figure 2A). Nevertheless, the vast majority of cell types is partitioned into at most three functional cell states with 16% of cell types remaining unpartitioned. To assess the performance of our method, we evaluated the function of the identified cell states with known cell state markers compiled from literature and compared the performance of FunRes against current clustering methods. As a result, we were able to collect evidence for multiple functional cell states in the heart, kidney and liver. Firstly, we identified functional cell states of cardiac coronary vascular endothelial cells. As a result, FunRes detected two clusters differentiated by Bmpr2 signaling. While the Bmpr2-positive state corresponds to quiescent coronary vascular endothelial cells, the Bmpr2-negative state consists of activated cells undergoing endothelial mesenchymal transition, which is required for cardiac functioning [25]. In this regard, FosB is activated by Bmpr2 according to FunRes and is downregulated in active vascular endothelial cells [26]. Thus, Bmpr2 and FosB serve as markers for quiescent and active coronary vascular endothelial cells. In order to assess the ability of current clustering methods in identifying functional cell states, we compared the results from FunRes against Seurat [5, 6], SC3 [7] and SINCERA [8]. As a result, SINCERA and SC3 detected 19 and 9 clusters, respectively, none of which could be distinguished using the marker genes of active and quiescent endothelial cells. Similar to FunRes, Seurat detected three clusters. However, no significant difference in the expression of these cell state markers could be detected. Consequently, none of the clusters corresponds to the functional cell states identified by FunRes (Supplementary Figure S1).

Secondly, we assessed the functional cell states of kidney macrophages and identified two states that are differentiated based on Nrp1 signaling. In particular, the first cluster expresses Nrp1 and Il10, all of which are known markers of M2 macrophages, an immunosuppressive cell state necessary for wound healing and cessation of the inflammatory response to pathogens (Figure 2B) [27, 28]. Previous reports showed that activation of Nrp1, which is involved in the cell–cell interactions differentiating the two clusters, induces this functional state, which underscores the accuracy of FunRes in detecting functionally relevant cell–cell interactions. In addition, the other functional cell state is characterized by active Tlr2 and Tlr4 signaling, which results in the polarization into M1 macrophages [29]. Indeed, cells belonging to this cluster express the M1 markers Nfatc1, Tnfsf9 and Il16 [30–32]. Thus, the identified cell states correspond to M1 and M2 polarization of macrophages, respectively. In contrast, current clustering methodologies cannot resemble this finding. Seurat was unable to cluster the data due to the failure of PCA with default parameters as the number of kidney macrophage cells is below the number of principal components to be computed. Sincera partitioned the cells into 35 different clusters having on average 1.09 cells whereas SC3 was the only tool providing a reasonable number of clusters. However, these populations do not correspond to M1 and M2 macrophages as known marker genes are expressed across clusters (Supplementary Figure S2). Thirdly, we investigated the functional states of liver B cells and identified two states that are distinguished by IL4R signaling [33]. Previous studies elucidated the role of IL4R signaling and found that it induces autophagy, a key mechanism for the maintenance of B cell memory against pathogens and synthesis of antigens, such as IgE [34]. Moreover, B cells with active IL4R signaling express the conventional B cell genes Tnfrsf13b and Tnfrsf13c (Figure 2C) [35, 36]. In contrast, the other identified cell state is deficient of IL4R signaling and express the plasma cell markers Sec61a1 (Figure 2C) [37]. Comparison of the cell states identified by FunRes to current clustering methodologies reveals that all of the other methods fail to detect the underlying functional heterogeneity. While Seurat was unable to perform clustering due to the failure of PCA with default parameters as the number of liver B cells is below the number of principal components to be computed, Sincera partitioned the data such that each cell belongs to a different cluster. In addition, SC3 did not detect any functional heterogeneity in liver B cells and leaves the data unpartitioned (Supplementary Figure S3). Finally, we applied our method to liver NK cells and detected two functional NK cell states defined by differential downstream signaling targets. In particular, cells of the second cluster specifically

activate Ikzf3 (Figure 2D), whereas other downstream signaling targets, such as Stat3, Stat1 and Ets1, are commonly activated in both clusters. Indeed, the second cluster is characterized by Ikzf2 and Ikzf3 expression, two markers of peripheral NK cells (Figure 2E) [38]. In contrast, the first cluster shows high expression of the adaptive NK cell markers Klrc1 and Klrc2 (Figure 2F) [39, 40]. Similar to the case of kidney macrophages, Sincera produced an excess of clusters with, on average, only 1.38 cells whereas Seurat produced no clusters due to the failure of PCA with default parameters as the number of liver NK cells is below the number of principal components to be computed. Although SC3 detected 2 clusters like FunRes, they could not be attributed to different functional NK cell states (Supplementary Figure S4). In summary, FunRes was able to dissect the functional heterogeneity of cell types in various tissues and accurately identified known functional states, while current clustering methods failed to provide satisfactory results. Moreover, comparison of the clusterings identified by each method shows only a moderate agreement between the functional cell states detected by FunRes and the results of Seurat, SC3 and SINCERA (Supplementary Figure S5).

Comparison of functional cell states in different tissues

Next, we set out to characterize the niche-induced functional states of macrophages and B cells in different tissues. As expected, we detected conserved functional states of kidney and limb muscle macrophages with activated Tlr2, Tlr4 and Tnfrsf1a signaling pathways, which constitute the key signaling cascades for the recognition of antigens and the activation of macrophages [41, 42]. However, while kidney macrophages show an Nrp1-positive subpopulation corresponding to M2 polarization, Nrp1 signaling is inactive in all limb muscle macrophages. However, a subset of cells in the limb muscle display activity of the Il1 decoy receptor Il1r2, which is unique to M2 polarized macrophages. Moreover, these cells display an activated Il6 signaling pathway, which has been previously shown to be an alternative mechanism for M2 macrophage polarization [43]. This demonstrates that M2 macrophages reside in both tissues but are alternatively activated. In contrast to macrophages, B cells were detected in the spleen, liver, limb muscle and lung having vastly different functional states that greatly differ between the spleen and liver/limb muscle/lung tissues (Figure 3A). In particular, the identified functional B cell states in the liver and limb muscle are commonly characterized by Il2 receptor signaling activity, which is involved in B cell proliferation and plasma cell differentiation [44, 45]. However, unique functional cell states are acquired

through Il4ra and Cxcr4 signaling activity in these tissues. As previously described, Il4ra promotes B cell autophagy required for memory maintenance and antigen presentation [33]. In contrast, Cxcr4 signaling is uniquely activated in the functional state of limb muscle B cells demonstrating the differentiation of these cells into plasma cells [46]. While B cells lose their responsiveness to Cxcl12, the cognate ligand of Cxcr4, during development, they regain their sensitivity upon differentiation into mature B cells [46]. Due to the involvement of Il4ra and Cxcr4 in plasma cell differentiation, we compared the downstream targets of these signaling cascades identified by FunRes and found Foxo1 to be the only common TF. Previous studies already revealed the key role of Foxo1 in plasma cell function and differentiation as an activator of Prdm1, a plasma cell master regulator [47]. Therefore, the identified functional B cell states unique to liver and limb muscle tissue, respectively, induce plasma cell differentiation through distinct signaling pathways. Opposed to liver and limb muscle tissues, the functional B cell states in the lung are characterized by Il2 and interferon gamma signaling, which is the only commonality with B cells from the spleen. Whereas a B cell state expressing Ifngr1 was detected in both spleen and lung tissues, their functions vary due to the presence of co-stimulatory signals. As expected, the identified functional B cell states in the spleen are characterized by active Cd40 and interferon gamma signaling corresponding to a proliferative state [48]. In contrast, no costimulatory signals were detected in pulmonary B cell states, which suggests an inhibition of B cell activity [48]. This finding is consistent with the second functional state identified in pulmonary B cells, which shows Il2 receptor signaling activity as well as the absence of interferon gamma signaling, similar to the functional states of limb muscle and liver B cells [44, 45].

Functional heterogeneity underlying aging and disease identifies emerging cell states

Finally, we investigated the influence of different tissue conditions, such as aging and disease, on the functional states of cell types. In order to interrogate how aging affects the functional cell states of macrophages, we compared the identified states in kidney and limb muscle tissue in young and old mice. Based on our analysis, we observed a significant decline in macrophage polarization due to the inactivation of Nrp1 and Il6r signaling pathways. In particular, while a functional state of kidney macrophages exists in old mice that is characterized by the ability to respond to antigens through Tlr2 signaling, the capacity of limb muscle macrophages to antigen response is significantly reduced. Moreover, the heterogeneous activation pattern of macrophages in young

tissues has generally declined, which results in more homogeneous, quiescent states. This observation is consistent with previous studies reporting a significant decline of macrophage activation during aging, which is mediated by reduced Tlr signal transduction capacity due to dysfunctional MyD88 [49]. In order to determine whether the loss of niche-induced functional states is a general characteristic of aging tissues, we compared the number of identified functional states for each cell type (Figure 3B). However, no statistically significant difference could be detected (two-sided Wilcoxon-signed-rank test, P-value: 0.64). While in more than 85% of cell types, the number of states differed by at most two, three cell types differed in more than six niche-induced functional states. For example, the increase of detected functional states in hepatic sinusoidal endothelial cells is largely driven by the specific activation of the transcription factors Mef2c, Nr3c1 and Jun through distinct signaling cascades of the same receptors. Indeed, these TFs carry out protective functions in hepatic endothelial cells by stimulating cellular survival, modulating the immune response and preventing the breakdown of the extracellular matrix [50–52]. Similarly, an increase in nicheinduced functional states can be observed in pancreatic alpha cells due to an elevated selectivity in the activation of Bmpr1a, Itgav and Ddr1. These receptors regulate key functions of alpha cells, such as glucagon secretion [53], even though their combinatorial effect remains elusive. Furthermore, we observe a few of cellular processes unique to young and old tissues, respectively (Figure 3C and D). In contrast to hepatic sinusoidal endothelial and pancreatic alpha cells, enterocytes of the large intestinal epithelium show a marked reduction of heterogeneity with age, which is associated to the impairment of key signaling pathways, including protective signals, such as Lpa and Somatostatin, as well as fat uptake through the sortilin receptor. Likewise, we investigated the effects of pathologic tissue conditions on the functional states of cell types exemplified in a dataset of human liver cirrhosis [54]. Hepatic cirrhosis is a chronic disease characterized by the progressive formation of permanent scar tissue through fibrosis. Hepatocytes are the main parenchymal cells of the liver, making up 50–60% of the tissue and are involved in key functions, such as lipid synthesis and detoxification. However, their dysregulation in the context of liver fibrosis and the functional states they attain remains elusive. Therefore, we applied FunRes, to healthy and cirrhotic human liver samples for identifying the functional states of hepatocytes. As a result, our method detected two functional states in healthy liver tissue characterized by Sdc1 and Cd74 signaling. Indeed, Sdc1 is necessary for the uptake and degradation of triglyceride-rich lipoproteins in the liver and is therefore involved

in a key function of the liver [55]. In addition, Cd74 is a key component of the response to acute liver injury by antigen processing and host defense [56]. In contrast, a great diversity of functional states can be observed in cirrhotic liver tissue. In total, FunRes identified six functional states that are characterized by interferon gamma, leptin receptor, Cxcr4, Tnf, Cd40 and Caveolin-1 signaling (Figure 3E). While two functional states show a highly pro-inflammatory phenotype mediated by interferon gamma, Tnfrsf1a and Cxcr4, three states are characterized by Lepr, Cd40 and Cav1 signaling. The function of leptin receptor in the context of cirrhosis has been subject to previous studies that demonstrated the production of Tnf and Il1 due to leptin signaling activation [57]. Moreover, in a rat model of liver cirrhosis, no significant changes in the expression of Lepr have been found, which is consistent with our data. Nevertheless, FunRes only detects significant signaling activity in pathologic livers highlighting the significance of this method. On the contrary, previous studies suggested that the functional hepatocyte state characterized by active Cd40 signaling likely amplifies Fas-dependent apoptosis, thus contributing to the progressive formation of scar tissue in response to injury [58]. Finally, functional states with Caveolin-1 signaling activity regulate endocytosis, energy metabolism and fatty acid uptake as well as inhibit NOS3, which promotes liver fibrosis [59].

Discussion and Conclusion

In this study, we presented FunRes, a computational method for resolving tissue-specific functional cell states that are determined by receptor–ligand-mediated cell–cell interactions. This method accounts for the effect of cell–cell interactions on specific cellular processes, which have been shown to play a fundamental role in the specification of functional cell states [60–62]. Namely, FunRes identifies functional cell states based on the downstream target genes of cell–cell interactions and the cellular processes they participate in. Hence, our approach is conceptually different from current clustering methods solely relying on differences in the expression of the most variable genes regardless of their cellular function. As a result, unlike other methods, FunRes was able to resolve functional states of various cell types. Importantly, the cell–cell interactions underlying functional cell states were largely supported by previous studies, indicating that FunRes can inform about the cues inducing these states. The validation of our method was further corroborated by the comparison of functional cell states in different tissues and tissue conditions. In particular, FunRes identified previously reported differences in the functional states of

macrophages and B cells, including the appearance of specific functional cell states, such as wound-healing macrophages in the limb muscle. Furthermore, we applied FunRes to several examples of aging and disease to disentangle the composition of functional cell states in these conditions. For instance, in the case of liver cirrhosis, results show the emergence of novel functional hepatocyte states sustained by Caveolin-1 and Leptin signaling that are characterized by dysregulated metabolic processes contributing to pathological tissue functioning. A limitation of FunRes is that it solely considers receptor– ligand-mediated cell–cell interactions whereas it ignores other ways of cellular communication, such as via exosomes. In addition, this method does not incorporate the effect of environmental factors, such as metabolites, on the functional cell states [63]. In this regard, the method could be extended to overcome some of these limitations. In summary, FunRes is the first computational method that systematically resolves tissue-specific functional cell states, thereby providing a functional characterization of the identified states. Thus, we believe FunRes will be of great utility in the characterization of the tissue-specific functional landscape and in the identification of dysfunctional cell states underlying aging and disease.

Author Contributions

S.J. implemented the software, conducted the analysis and wrote the manuscript. K.S. implemented the software and conducted the analysis. A.d.S. conceived the idea, supervised the work and wrote the manuscript.

Data Availability

The data underlying this article are available in figshare. The weblinks and identifiers of the individual datasets can be found in Supplementary Table S1. The source code of FunRes is deposited in a GitHub repository (https://git-r3lab.uni.lu/kartikeya.singh/funres).

Supplementary Data

Supplementary data are available online at Briefings in Bioinformatics. Acknowledgement We thank Ilya Potapov for fruitful discussions about the cell– cell communication network inference. Funding K.S. is supported by the Luxembourg National Research Fund (FNR) under the PRIDE program (Project code: 11012546) within the NextImmune DTU. Conflict of interest The authors declare no conflict of interest.

References

1. Gartner ZJ, Prescher JA, Lavis LD. Unraveling cell-to-cell signaling networks with chemical biology. Nat Chem Biol 2017;13:564–8.

2. Lavin Y, Winter D, Blecher-Gonen R, et al. Tissue-resident macrophage enhancer landscapes are shaped by the local microenvironment. Cell 2014;159:1312–26.

3. Lugert S, Basak O, Knuckles P, et al. Quiescent and active hippocampal neural stem cells with distinct morphologies respond selectively to physiological and pathological stimuli and aging. Cell Stem Cell 2010;6:445–56.

4. Hsieh J. Orchestrating transcriptional control of adult neurogenesis. Genes Dev 2012;26:1010–21.

5. Stuart T, Butler A, Hoffman P, et al. Comprehensive integration of single-cell data. Cell 2019;177:1888–1902.e21. 6. Butler A, Hoffman P, Smibert P, et al. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol 2018;36:411–20.

7. Kiselev VY, Kirschner K, Schaub MT, et al. SC3: consensus clustering of single-cell RNA-seq data. Nat Methods 2017;14:483–6.

8. Guo M, Wang H, Potter SS, et al. SINCERA: a pipeline for single-cell RNA-Seq profiling analysis. PLoS Comput Biol 2015;11:e1004575.

9. Ramilowski JA, Goldberg T, Harshbarger J, et al. A draft network of ligand–receptor-mediated multicellular signalling in human. Nat Commun 2015;6:7866.

10. Salwinski L. The database of interacting proteins: 2004 update. Nucleic Acids Res 2004;32:449D–51.

11. Ben-Shlomo I, Yu Hsu S, Rauch R, et al. Signaling receptome: a genomic and evolutionary perspective of plasma membrane receptors involved in signal transduction. Sci Signal 2003;2003:re9–9. 12. Armstrong JF, Faccenda E, Harding SD, et al. The IUPHAR/BPS guide to Pharmacology in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV guide to MALARIA PHARMACOLOGY. Nucleic Acids Res 2020;48: D1006–21.

13. Keshava Prasad TS, Goel R, Kandasamy K, et al. Human protein reference database–2009 update. Nucleic Acids Res 2009;37:D767–72.

14. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: proteinprotein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019;47:D607–13.

15. Consortium TU. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 2019;47:D506–15.

16. Cunningham F, Achuthan P, Akanni W, et al. Ensembl 2019. Nucleic Acids Res 2019;47:D745–51. 17. Türei D, Korcsmáros T, Saez-Rodriguez J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. Nat Methods 2016;13:966–7.

18. Jassal B, Matthews L, Viteri G, et al. The reactome pathway knowledgebase. Nucleic Acids Res 2020;48:D498–503.

19. Ravichandran S, Hartmann A, del Sol A. SigHotSpotter: scRNA-seq-based computational tool to control cell subpopulation phenotypes for cellular rejuvenation strategies. Bioinformatics 2019;36:1963–5.

20. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium Nat Genet 2000;25:25–9.

21. The Gene Ontology Consortium. The gene ontology resource: 20 years and still GOing strong. Nucleic Acids Res 2019;47:D330–8.

22. Yu G, Wang L-G, Yan G-R, et al. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. Bioinformatics 2015;31:608–9.

23. Zemmour D, Zilionis R, Kiner E, et al. Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. Nat Immunol 2018;19:291–301.

24. Pisco AO,McGeever A, Schaum N, et al. A single cell transcriptomic atlas characterizes aging tissues in the mouse. bioRxiv 2020;661728.

25. Kovacic JC, Dimmeler S, Harvey RP, et al. Endothelial to mesenchymal transition in cardiovascular disease. J Am Coll Cardiol 2019;73:190–209.

26. Franscini N, Bachli EB, Blau N, et al. Gene expression profiling of inflamed human endothelial cells and influence of activated protein C. Circulation 2004;110:2903–9.

27. Dai X, Okon I, Liu Z, et al. A novel role for myeloid cell-specific neuropilin 1 in mitigating sepsis. FASEB J 2017;31:2881–92.

28. Roszer T. Understanding the mysterious M2 macrophage ″ through activation markers and effector mechanisms. Mediators Inflamm 2015;2015:1–16.

29. Schlaepfer E, Rochat M-A, Duo L, et al. Triggering TLR2, -3, -4, -5, and -8 reinforces the restrictive nature of M1- and M2-polarized macrophages to HIV. J Virol 2014;88:9769–81.

30. Zhao Z, Hou X, Yin X, et al. TNF induction of NF-κB RelB enhances RANKL-induced Osteoclastogenesis by promoting inflammatory macrophage differentiation but also limits it through suppression of NFATc1 expression. PLoS One 2015;10:e0135728.

31. Wu J, Wang Y, Jiang Z. TNFSF9 is a prognostic biomarker and correlated with immune infiltrates in pancreatic cancer. J Gastrointest Cancer 2020. 32. Huang Y, Du KL, Guo PY, et al. IL-16 regulates macrophage polarization as a target gene of mir-145-3p. Mol Immunol 2019;107:1–9.

33. Xia F, Deng C, Jiang Y, et al. IL4 (interleukin 4) induces autophagy in B cells leading to exacerbated asthma. Autophagy 2018;14:450–64.

34. Chen M, Hong MJ, Sun H, et al. Essential role for autophagy in the maintenance of immunological memory against influenza infection. Nat Med 2014;20:503–10.

35. Yan M, Brady JR, Chan B, et al. Identification of a novel receptor for B lymphocyte stimulator that is mutated in a mouse strain with severe B cell deficiency. Curr Biol 2001;11:1547–52.

36. Schneider P, MacKay F, Steiner V, et al. BAFF, a novel ligand of the tumor necrosis factor family, stimulates B cell growth. J Exp Med 1999;189:1747–56.

37. Schubert D, Klein M-C, Hassdenteufel S, et al. Plasma cell deficiency in human subjects with heterozygous mutations in Sec61 translocon alpha 1 subunit (SEC61A1). J Allergy Clin Immunol 2018;141:1427–38.

38. Holmes ML, Huntington ND, Thong RP, et al. Peripheral natural killer cell maturation depends on the transcription factor Aiolos. EMBO J 2014;33:2721–34.

39. Lunemann S, Langeneckert AE, Martrus G, et al. Human liver-derived CXCR6 + NK cells are predominantly educated through NKG2A and show reduced cytokine production. J Leukoc Biol 2019;105:1331–40.

40. Abel AM, Yang C, Thakar MS, et al. Natural killer cells: development, maturation, and clinical utilization. Front Immunology 2018;9:1869.

41. Janssens S, Beyaert R. Role of toll-like receptors in pathogen recognition. Clin Microbiol Rev 2003;16:637–46.

42. Parameswaran N, Patial S. Tumor necrosis factor-α Signaling in macrophages. Crit Rev Eukaryot Gene Expr 2010;20: 87–103.

43. Mauer J, Chaurasia B, Goldau J, et al. Signaling by IL-6 promotes alternative activation of macrophages to limit endotoxemia and obesity-associated resistance to insulin. Nat Immunol 2014;15:423–30.

44. Mingari MC, Gerosa F, Carra G, et al. Human interleukin- 2 promotes proliferation of activated B cells via surface receptors similar to those of activated T cells. Nature 1984;312:641–3.

45. Le Gallou S, Caron G, Delaloy C, et al. IL-2 requirement for human plasma cell generation: coupling differentiation and proliferation by enhancing MAPK–ERK Signaling. J Immunol 2012;189:161–73.

46. Mcheik S, Van Eeckhout N, De Poorter C, et al. Coexpression of CCR7 and CXCR4 during B cell development controls CXCR4 responsiveness and bone marrow homing. Front Immunol 2019;10:2970.

47. Vogel MJ, Xie L, Guan H, et al. FOXO1 repression contributes to block of plasma cell differentiation in classical Hodgkin lymphoma. Blood 2014;124:3118–29.

48. Vazquez MI, Catalan-Dibene J, Zlotnik A. B cells responses and cytokine production are regulated by their immune microenvironment. Cytokine 2015;74:318–26.

49. Dunston CR, Griffiths HR. The effect of ageing on macrophage toll-like receptor-mediated responses in the fight against pathogens. Clin Exp Immunol 2010;161:407–16.

50. Salameh A, Galvagni F, Anselmi F, et al. Growth factor stimulation induces cell survival by c-Jun·ATF2-dependent activation of Bcl-X L. J Biol Chem 2010;285:23096–104.

51. Sturtzel C, Testori J, Schweighofer B, et al. The transcription factor MEF2C negatively controls Angiogenic sprouting of endothelial cells depending on oxygen. PLoS One 2014;9:e101521.

52. Goodwin JE, Feng Y, Velazquez H, et al. Endothelial glucocorticoid receptor is required for protection against sepsis. Proc Natl Acad Sci 2013;110:306–11.

53. Nielsen SS, Christensen GL, Holst JJ, et al. Regulation of pancreatic α-cell function and proliferation by bone morphogenetic protein 4 (BMP4) in vitro. Endocrinology 2016;157:3809–20.

54. Ramachandran P, Dobie R, Wilson-Kanamori JR, et al. Resolving the fibrotic niche of human liver cirrhosis at single-cell level. Nature 2019;575:512–8.

55. Stanford KI, Bishop JR, Foley EM, et al. Syndecan-1 is the primary heparan sulfate proteoglycan mediating hepatic clearance of triglyceride-rich lipoproteins in mice. J Clin Invest 2009;119:3236–45.

56. Koch KS, Leffert HL. Ectopic expression of CD74 in Ikkβ-deleted mouse hepatocytes. Acta Histochem 2011;113: 428–35.

57. Otte C, Otte J-M, Strodthoff D, et al. Expression of leptin and leptin receptor during the development of liver fibrosis and cirrhosis. Exp Clin Endocrinol Diabetes 2004;112: 10–7.

58. Afford SC, Randhawa S, Eliopoulos AG, et al. CD40 activation induces apoptosis in cultured human hepatocytes via induction of cell surface Fas ligand expression and amplifies Fasmediated hepatocyte death during allograft rejection. J Exp Med 1999;189:441–6.

59. Fernandez-Rojo MA, Ramm GA. Caveolin-1 function in liver physiology and disease. Trends Mol Med 2016;22: 889–904.

60. Schrier SB, Hill AS, Plana D, et al. Synergistic communication between CD4+ T cells and monocytes impacts the cytokine environment. Sci Rep 2016;6:34942.

61. Shalek AK, Satija R, Shuga J, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 2014;510:363–9.

62. Xue Q, Lu Y, Eisele MR, et al. Analysis of single-cell cytokine secretion reveals a role for paracrine signaling in coordinating macrophage responses to TLR4 stimulation. Sci Signal 2015;8:ra59–9.

63. Vishwakarma A, Rouwkema J, Jones PA, et al. The need to study, mimic, and target stem cell niches. Biol Eng Stem Cell Niches 2017;3–13.

Figures

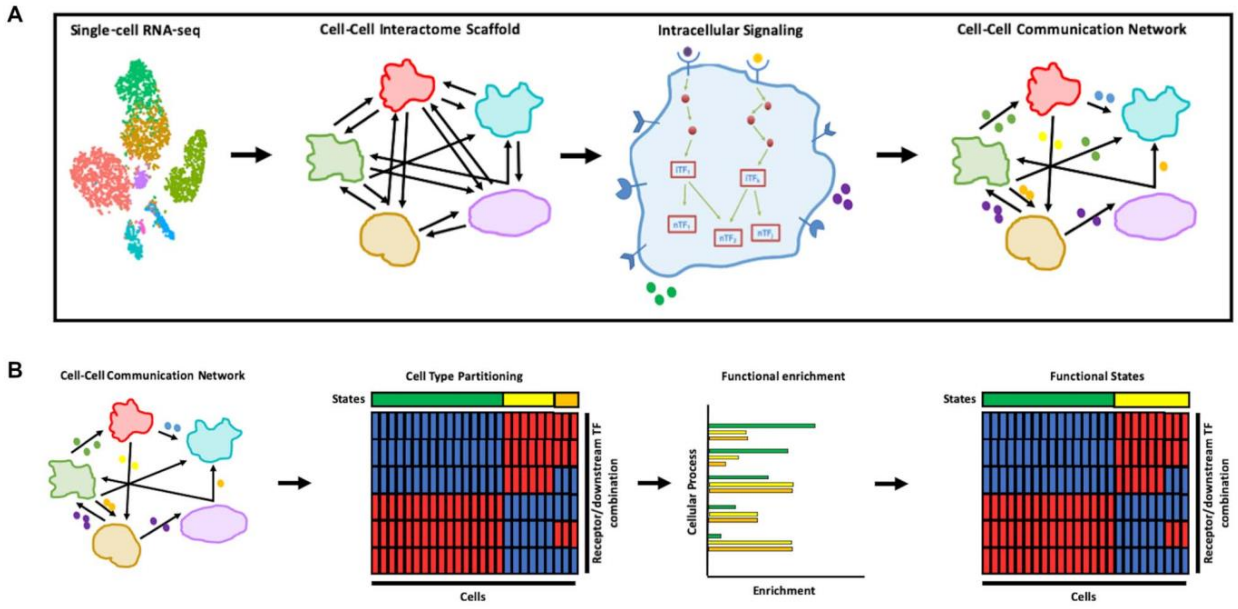


Figure 1. Method overview. (A) Based on single-cell RNA-seq data of a tissue, FunRes infers active receptors in each cell population. Based on these receptors, FunRes reconstructs a cell–cell communication network by identifying interactions having a significantly higher score compared to other interactions in the interaction scaffold. (B) Given a reconstructed cell–cell communication network of a tissue, FunRes identifies functional cell states for each cell type individually. First, cells are represented as a binary vector of receptor/downstream TF expression. If a receptor and its corresponding downstream TF is expressed in a cell, it will be represented as '1' (red) and '0' (blue) otherwise. An optimal number of clusters is determined, which will be subjected to functional enrichment analysis. If two clusters are enriched in the same processes, they will be merged. The merged clusters define the functional cell states.\

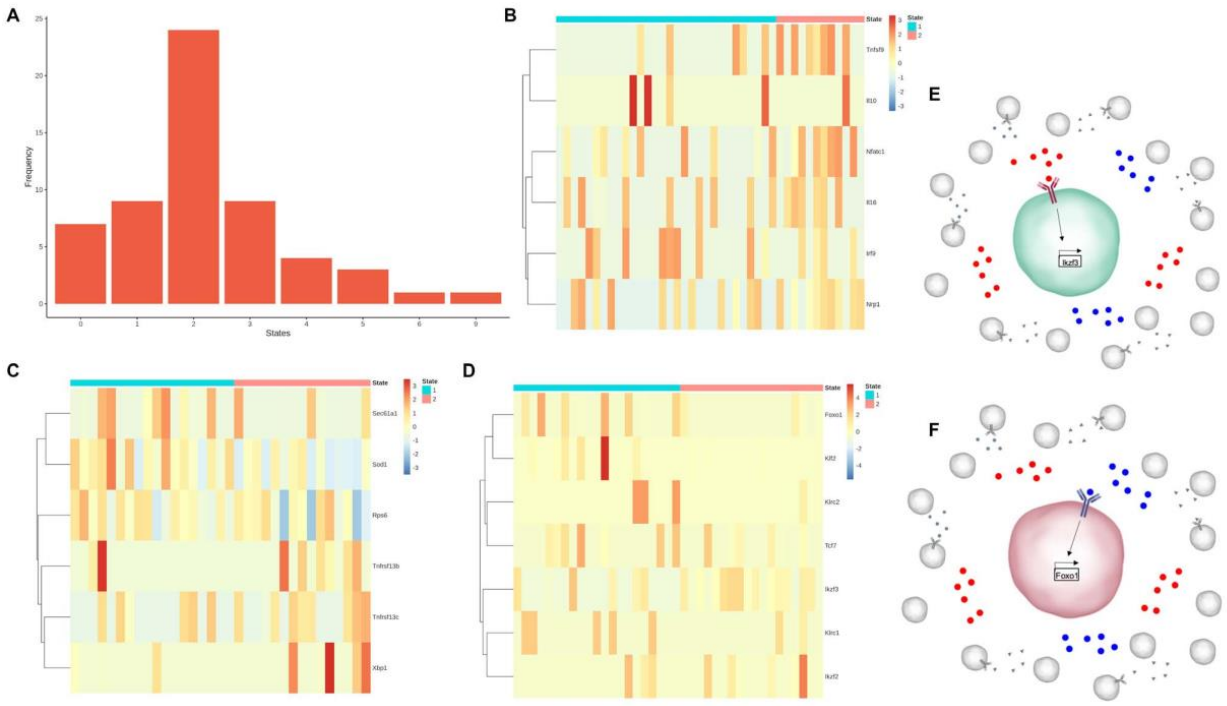


Figure 2. Validation of FunRes and comparison to the state-of-the-art methods. (**A**) Histogram of the number of identified functional cell states in all cell types of all tissues. (**B**–**D**) Heatmaps of genes important for the function of the identified functional cell states in kidney macrophages (B), liver B cells (C) and liver NK cells (D). The expression values are scaled for each gene individually. (**E** and **F**) Visualization of how different cues from the environment induce downstream marker genes of peripheral (E) and adaptive (F) NK cell states.

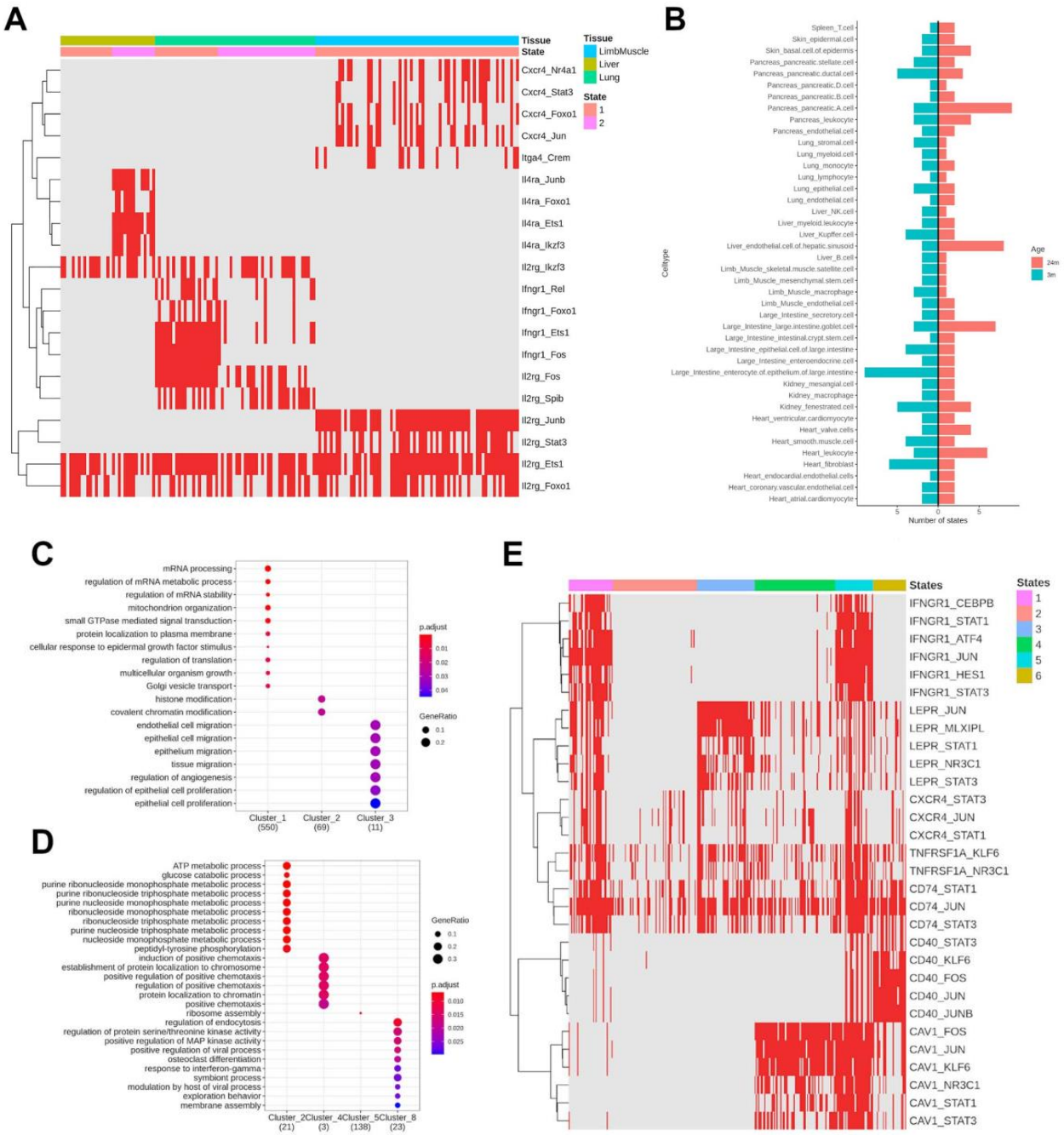


Figure 3. Comparison of functional states across tissues and tissue conditions. (**A**) Heatmap of active signaling cascades induced by cell–cell interactions in B cells of three different tissues. Each row corresponds to a pair of receptor and downstream TF target, whereas columns correspond to cells. Expression of both the receptor and downstream TF target in a cell is depicted in red. Failure to express any of the two genes is depicted in grey. B cells from all tissues share active Il2 signaling cascades targeting common (Ets1, Foxo1) and unique downstream TFs (Stat3, Junb, Spib, Fos). (**B**) Bar chart of identified functional cell states per cell type in young and old tissues. The number of functional states does not significantly change in aging, except for certain cell types, such as pancreatic alpha cells, endothelial cells of the hepatic

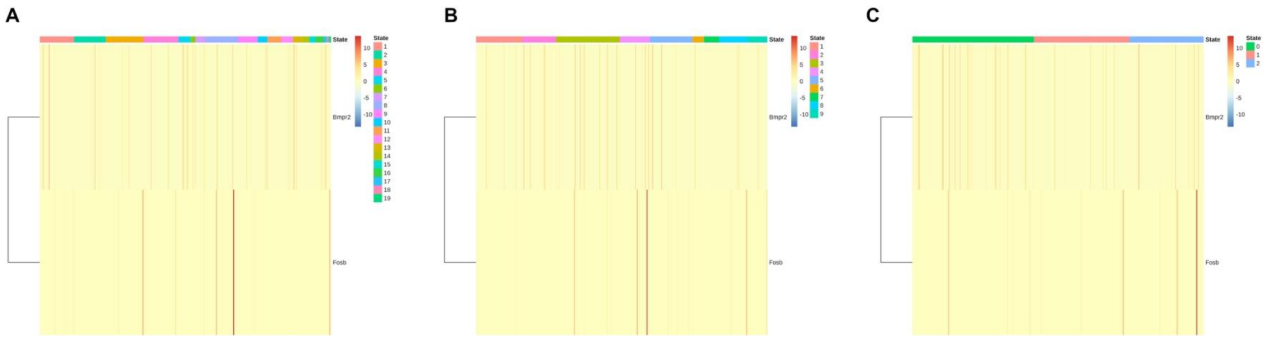
sinusoid and enterocytes of the large intestine epithelium. (**C** and **D**) Cellular processes unique to pancreatic alpha cells in young and old tissues. The size of the dots represents the fraction of genes in the category that are expressed in the corresponding state, whereas the color indicates significance from blue (less significant) to red (most significant). (**E**) Heatmap of active signaling cascades induced by cell–cell interactions in hepatocytes of cirrhotic liver tissue. Each row corresponds to a pair of receptor and downstream TF target, whereas columns correspond to cells. Expression of both the receptor and downstream TF target in a cell is depicted in red. Six functional cell states were identified that are characterized by Interferon-gamma, CD40, Cav1 and Leptr signaling, respectively.

4.1.3 Supplementary Information

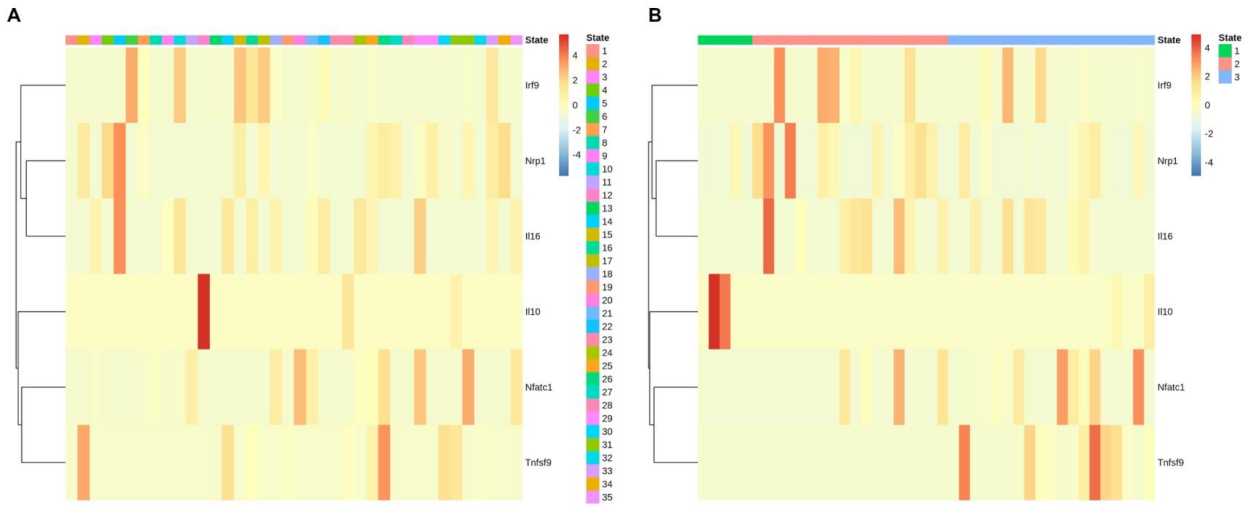
Supplementary Table

Tissue	Age	Sequencing type	Source	Identifier
Heart	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Heart_facs.h5ad
Heart	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Heart_facs.h5ad
Kidney	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Kidney_facs.h5ad
Kidney	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Kidney_facs.h5ad
Large Intestine	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Large_Intestine_facs.h5ad
Large Intestine	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Large_Intestine_facs.h5ad
Limb Muscle	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Limb_Muscle_facs.h5ad
Limb Muscle	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Limb_Muscle_facs.h5ad
Liver	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Liver_facs.h5ad
Liver	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Liver_facs.h5ad
Lung	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Lung_facs.h5ad
Lung	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Lung_facs.h5ad
Pancreas	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Pancreas_facs.h5ad
Pancreas	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Pancreas_facs.h5ad
Skin	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Skin_facs.h5ad
Skin	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Skin_facs.h5ad
Spleen	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Spleen_facs.h5ad
Spleen	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Spleen_facs.h5ad
Thymus	3m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Thymus_facs.h5ad
Thymus	24m	Smart-seq	https://doi.org/10.6084/m9.figshare.8273102.v2	Thymus_facs.h5ad

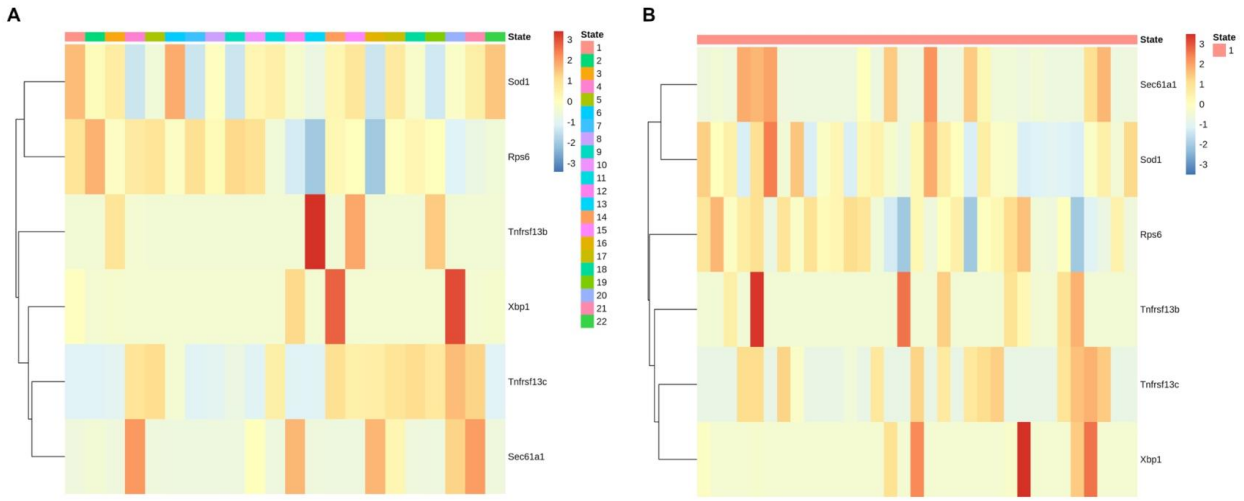
Supplementary Figures



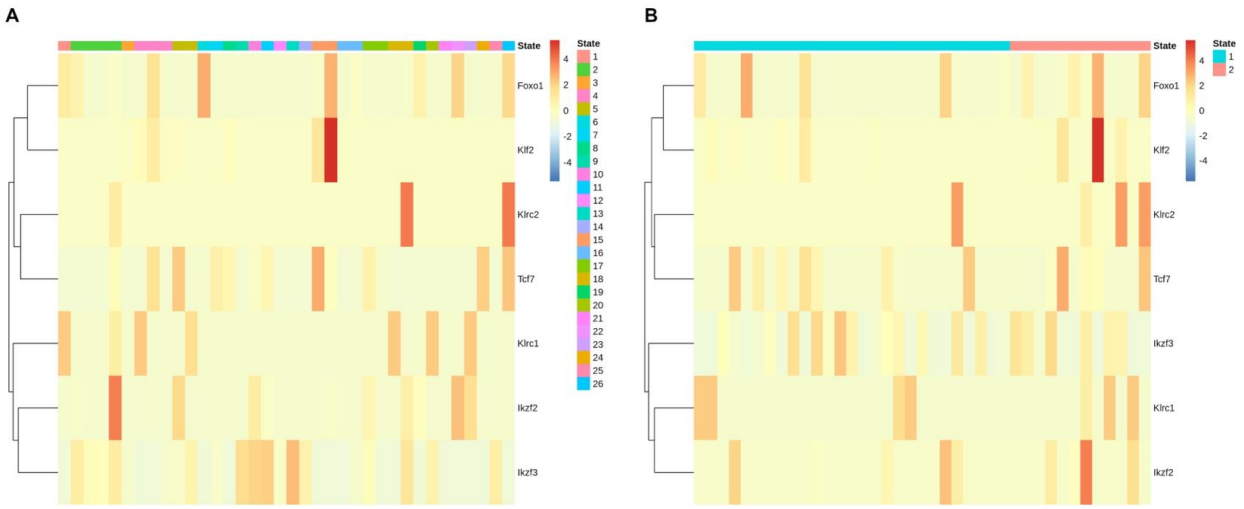
Supplementary Figure S1. A-C Heatmaps of genes important for the function of coronary vascular endothelial cells grouped by the clusters identified by (A) SINCERA, (B) SC3 and (C) Seurat. Expression values are scaled per gene. The identified clusters do not group highly expressed genes together.



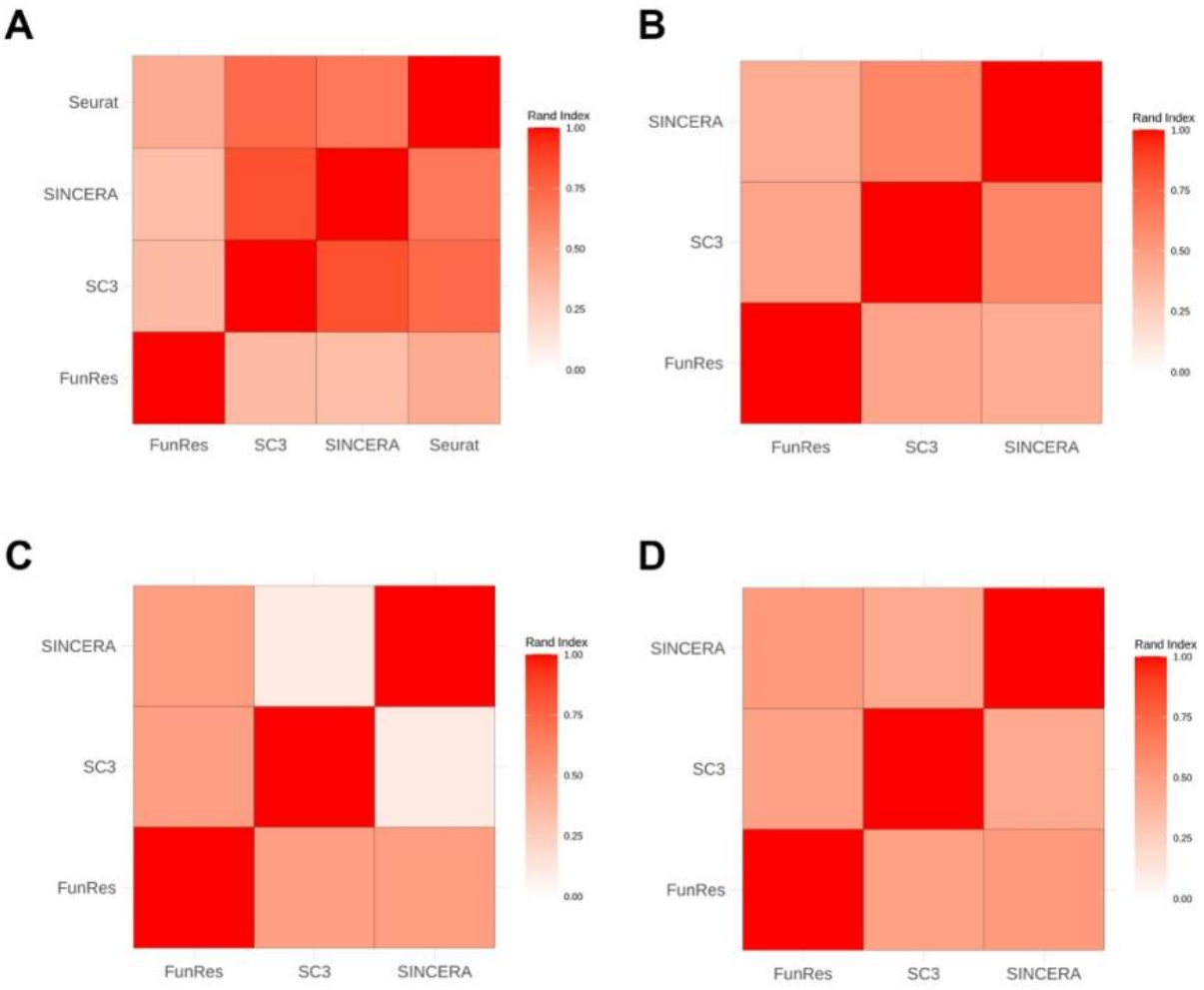
Supplementary Figure S2. A-B Heatmaps of genes important for the function of kidney macrophages grouped by the clusters identified by (A) SINCERA and (B) SC3. Expression values are scaled per gene. The identified clusters by SINCERA do not group highly expressed genes together. The clusters identified by SC3 are similar to the ones identified by FunRes. Seurat was unable to cluster the data due to a failure in principal component analysis.



Supplementary Figure S3. A-B Heatmaps of genes important for the function of liver B cells grouped by the clusters identified by (A) SINCERA and (B) SC3. Expression values are scaled per gene. The identified clusters by SINCERA do not group highly expressed genes together while SC3 did not detect any functional heterogeneity. Seurat was unable to cluster the data due to a failure in principal component analysis.



Supplementary Figure S4. A-B Heatmaps of genes important for the function of liver NK cells grouped by the clusters identified by (A) SINCERA and (B) SC3. Expression values are scaled pergene. The identified clusters by SINCERA and SC3 do not group highly expressed genes together. Seurat was unable to cluster the data due to a failure in principal component analysis.



Supplementary Figure S5. Similarity of clusterings produced by FunRes, SC3, SINCERA and Seurat for (A) coronary vascular endothelial cells, (B) kidney macrophages, (C) liver B cells and (D) liver NK cells. Similarity was measured by the Rand index of the pairwise clusterings.

4.2 Identifying impacts of perturbation in cell-cell communication

4.2.1 Preface

In the previous study, we demonstrated the role of cell-cell communication and tissue niche in the function of celltypes. These interactions are an important factor in the maintenance of tissue physiology and alteration in the cell-cell communication in the tissue could lead to loss/alteration of function in various cell types. In this study "Transcriptional and chromatin accessibility profiling of neural stem cells differentiating into astrocytes reveal dynamic signatures affected under inflammatory conditions" we explored the effect of inflammatory conditions on astrocytes during their development. This study would help us understand the impact of acute inflammation on a tightly regulated process such as the development of cells from pluripotent to differentiated cells.

Inflammation is a regulatory mechanism employed by biological systems during various important processes including development and immune responses. While the role of inflammation has been well studied as part of various disorders including neurological disorders like Alzheimer's disease and epilepsy. Astrocytes are specialized cells of the central nervous system that are involved in major functions including neurogenesis and synaptogenesis. Loss of function in astrocytes has been implicated in neurodegenerative disorders. While the effect of inflammatory molecules on astrocytes has been studied but these studies did not explore the effect of inflammation during development. As astrocytic progenitors move postnatally to the brain and then develop into mature astrocytes the conditions of the area during that duration would shape the astrocytes.

This study had two novelties, firstly we characterized the development of astrocytes from neural stem cells at various timepoints under normal and inflammatory conditions. We employed TNF a major inflammatory cytokine for inflammatory conditions. As the role of astrocytes has been characterized in various diseases owing to their role in the regulation of a healthy state in the CNS. Also, inflammation has been identified as a constituent of many neurological diseases thus providing importance to the study from both perspectives of studying the development of astrocytes under both normal and inflammatory conditions. Secondly, we explored the major regulatory genes involved in the developmental process aiding in the identification of possible therapeutical targets for various neurological disorders. This was performed by the integration of

RNA-Seq and ATAC-Seq data generated under this study. The RNA-Seq data provides information of the transcriptomic changes that occur under the development of cells under inflammation compared to normal conditions. Further, the change in the epigenetic landscape of the cells was inferred using the ATAC-Seq data generated.

The bulk RNA-Seq and ATAC-Seq data generated during this study have been made available to the community by submitting them to open-source database. These datasets will support the community's efforts to disseminate the role of inflammation in various disorders.

Contribution: I pre-processed the RNA-Seq and ATAC-Seq data, analysed the data, and inferred the results. Further, I wrote the manuscript and created the figures for the manuscript.

4.2.2 Manuscript

Transcriptional and chromatin accessibility profiling of neural stem cells differentiating into astrocytes reveal dynamic signatures affected under inflammatory conditions

Maria Angeliki S. Pavlou[1,2*], Kartikeya Singh[3*], Srikanth Ravichandran[3], Rashi Halder[4], Nathalie Nicot[5,6], Cindy Birck[1], Luc Grandbarbe[1], Antonio del Sol[3,7,8#], Alessandro Michelucci[2,4#]

[1]Department of Life Sciences and Medicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

[2]Neuro-Immunology Group, Department of Cancer Research, Luxembourg Institute of Health, Luxembourg, Luxembourg

[3]Computational Biology Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

[4]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

[5]Translational Medicine Operations Hub, Luxembourg Institute of Health, Dudelange, Luxembourg

[6]LuxGen Genome Center, Luxembourg Institute of Health & Laboratoire national de santé, Dudelange, Luxembourg

[7]Computational Biology Group, CIC bioGUNE-BRTA (Basque Research and Technology Alliance), Derio, Spain

[8]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

* Contributed equally

# Contributed equally

ǂ Corresponding author

Abstract

Astrocytes arise from multipotent neural stem cells (NSCs) and represent the most abundant cell type of the central nervous system (CNS) playing key roles in the developing and adult brain. Since the differentiation of NSCs towards a gliogenic fate is a precisely timed and regulated process, its perturbation gives rise to dysfunctional astrocytic phenotypes. Inflammation, which often underlies neurological disorders, including neurodevelopmental disorders and brain tumors, disrupts the accurate developmental process of NSCs. However, the specific consequences of an inflammatory environment on the epigenetic and transcriptional programs underlying NSCs differentiation into astrocytes is unexplored yet. Here, we address this gap by profiling glial precursors from neural tissue derived from early embryonic stages along their astrocytic differentiation trajectory in the presence or absence of tumor necrosis factor-alpha (TNF), a master pro-inflammatory cytokine. By using a combination of RNA- and ATAC-sequencing approaches, together with footprint and integrated gene regulatory network analyses, we here identify key differences during the differentiation of NSCs into astrocytes under physiological and inflammatory settings. Notably, in agreement with its role to turn cells resistant to inflammatory challenges, we detect *Nrf2* as a master transcription factor supporting the astrocytic differentiation under TNF exposure. Further, under these conditions, we unravel additional transcriptional regulatory hubs, including *Stat3*, *Smad3*, *Cebpb* and *Nfkb2*, highlighting the interplay between

signaling pathways underlying physiological astrocytic developmental processes and those involved in inflammatory responses.

Overall, our study reports key transcriptional and epigenetic changes leading to the identification of molecular regulators of the astrocytic differentiation. Furthermore, our analyses provide a valuable resource for understanding inflammation-induced astrocytic phenotypes that might contribute to the development and progression of CNS disorders with an inflammatory component.

Introduction

In the developing brain, multipotent neural stem cells (NSCs) can give rise to neurons, oligodendrocytes and astrocytes. The process of differentiation towards a neurogenic or gliogenic fate is finely timed and regulated through mammalian brain development. During the expansion phase, NSCs self-replicate via symmetric cell division to expand their own pool. Later on, during mid-gestation they undergo asymmetric cell division and receive cues to differentiate into neurons. Lastly, from late-gestation to perinatal periods they enter the gliogenic phase and differentiate into astrocytes and oligodendrocytes (Hirabayashi and Gotoh 2005, Takouda, Katada et al. 2017). Importantly, undifferentiated NSCs still reside in two niches of the adult brain, namely the subventricular zone (SVZ) of the lateral ventricles and the subgranular zone (SGZ) of the dentate gyrus (Gotz, Nakafuku et al. 2016).

Astrocytes are active supporters of the neuronal network providing neurons with nutrients and metabolites, regulating neuronal synaptic transmission, participating in the formation and functioning of the synapses as well as directly communicating with neurons (Pavlou, Grandbarbe et al. 2019). Further, during development and homeostasis, astrocytes support the immune system and contribute to the formation of the blood brain barrier. Also, they contribute to oligodendrocyte differentiation and maturation as well as support myelination (Domingues, Portugal et al. 2016). Therefore, aberrant astrocytic phenotypes can give rise to neurological diseases, including

neurodevelopmental disorders (Pekny, Pekna et al. 2016). For example, dominant gain-of-function mutations in the gene coding for glial fibrillary acidic protein (GFAP), the hallmark intermediate filament in astrocytes, elicit the accumulation of cytoplasmic GFAP aggregates resulting in macrocephaly, white matter degeneration and developmental delay in Alexander disease (Quinlan, Brenner et al. 2007, Messing, Brenner et al. 2012). During inflammatory events in the CNS, several immune mediators, including cytokines and free radicals, affect intrinsic NSC properties and neurogenesis (Arvidsson, Covacu et al. 2015). Injuries of the CNS largely elicit proliferation and migration of NSCs toward the injury site, suggesting the activation of a putative regenerative response. However, detrimental cues encountered within the inflammatory niche are responsible, at least partly, of an incomplete regeneration from NSCs (Arvidsson, Covacu et al. 2015). Thus, in traumatic brain injuries, for example, attempts promoting a regenerative-favoring environment using endogenous or transplanted NSCs for reparative purposes represent critical challenges (Hess and Borlongan 2008, Lozano, Gonzales-Portillo et al. 2015). In a different context, perturbations of the fine-tuned process of NSCs differentiation under inflammatory conditions can also give rise to brain tumors. Indeed, inflammation is linked to gliomagenesis (Galvao and Zong 2013) and evidence indicates that NSCs can give rise to gliomas (Sanai, Alvarez-Buylla et al. 2005, Alcantara Llaguno and Parada 2016). Hence, defective differentiation abilities of NSCs and aberrant neoplastic astrocytic phenotypes in gliomas are attributed, at least partly, to the inflammatory tumor microenvironment.

Tumor necrosis factor-alpha (TNF) is a crucial inflammatory cytokine mainly produced by microglia during neuroinflammatory processes. TNF induces various signaling responses within the cells that eventually result in apoptosis or necrosis (Idriss and Naismith 2000), cell survival, differentiation, and inflammation in cells expressing the corresponding receptors (Sun and Fink 2007). Notably, the crosstalk between interferons and TNF affects the reprogramming of the macrophage epigenome and their inflammatory responses (Park, Kang et al. 2017). In this context, the role of epigenetic processes in the establishment, maintenance and differentiation of NSCs is emerging (Adefuin, Kimura et al. 2014). For example, deficiency of the chromatin remodeler gene *Cdh5* in NSCs leads to their improper activation altering cell fate decisions and favoring an astroglial over a neuronal identity (Hwang, Jaganathan et al. 2018).

Here, we sought to investigate the dynamics of the transcriptional and chromatin accessibility changes underlying NSCs differentiation into astrocytes under normal and inflammatory

conditions. For this, we adopted an integrated genome-wide approach combining RNA-sequencing (RNA-seq) and assay for transposase-accessible chromatin followed by deep sequencing (ATAC-seq) (Buenrostro, Giresi et al. 2013) to investigate the astrocytic differentiation process from murine neurospheres. We uncovered that normal differentiated astrocytes display mature cell-specific characteristics, whereas differentiating cells exposed to TNF do not reach the same level of maturity. In line with these observations and in agreement with a more restricted lineage-specific transcriptional program, we identified an overall reduction of chromatin accessibility in differentiated astrocytes compared to multipotent undifferentiated NSCs. Notably, this restriction was less pronounced under inflammatory conditions. Lastly, we reconstructed gene regulatory networks to identify transcription factors involved in the regulation of differentiation under normal and inflammatory settings.

Overall, our findings shed light on the molecular mechanisms underlying NSC differentiation into astrocytes and the effect of an inflammatory environment during this developmental process, which might pave the way to novel therapeutic targets restoring appropriate astrocytic phenotypes in CNS disorders with an inflammatory component, including neurodevelopmental diseases and brain tumors.

Materials and Methods

Animals and ethics

C57BL/6J mice, both wild type and transgenic mice expressing green fluorescent protein (GFP) under the control of glial fibrillary acidic protein (GFAP) promoter, were housed in individually ventilated cages (IVC) in a conventional animal facility at the University of Luxembourg in accordance with the EU Directive 2010/63/EU and Commission recommendation 2007/526/EC. We kept mice in groups under a dark-light cycle with *ad libitum* access to water and food. All the animal work of the present study has been conducted and reported in accordance with the ARRIVE (Animal Research: Reporting of In Vivo Experiments) guidelines to improve the design, analysis and reporting of research using animals, maximizing information published and minimizing unnecessary studies. We conducted all animal procedures in accordance with the local Committee for Care and Use of Laboratory Animals.

Cell culture

We isolated primary NSCs from the ventricular zone of single wild type mouse brains (C57BL/6J,

Harlan, The Netherlands) at embryonic day 14 (E14) as described before (Grandbarbe, Bouissac et al. 2003, Birck, Ginolhac et al. 2021). Primary cultures of neurospheres (NSPs) were kept under proliferating conditions in neurobasal medium (DMEM F12; Lonza) supplemented with 1% B27 without vitamin A (Life Technologies), penicillin (100 U/ml; Lonza), streptomycin (100 g/ml; Lonza), and 20 ng/ml EGF (Epidermal Growth Factor; Life Technologies). We differentiated NSPs into astrocytes on 6-well or 12-well plates coated with poly-L-ornithine and by exchanging the proliferation medium with DMEM containing 10% FBS (Fetal Bovine Serum; Gibco), penicillin (100 U/ml; Lonza), and streptomycin (100 g/ml; Lonza). We differentiated cells into astrocytes at different time points (24, 48, 72 hours and 1 week) at 37°C in 5% CO_2 and we kept them under normal conditions or treated with TNF (50 ng/ml; R&D Systems).

Selection of astrocytes based on antibody-coated beads

We used NSPs and differentiating NSPs at different time points (24, 48 and 72 hours) to isolate astrocytes by magnetic separation. We enriched astrocytic populations by magnetic separation using ACSA1+ (GLAST) beads (MACS Miltenyi Biotec) according to the manufacturer's recommendations.

Real-Time qPCR (RT-qPCR)

We kept murine NSPs under proliferating conditions or we differentiated them into astrocytes for 24, 48 and 72 hours or 1 week with or without TNF treatment. We extracted total RNA by the innuPREP RNA Mini Kit (Westburg) according to manufacturer's recommendations. Complementary DNA (cDNA) was synthetized by using the ImProm-II Reverse Transcription System (Promega). We conducted cDNA synthesis for 5 min at 25^{o}C, followed by 1 hour at 42^{o}C and then 15 minutes at 70^{o}C. We measured gene expression levels using SYBR Green Supermix (Promega) following the manufacturer's recommendations and by using CFX Connect Real-Time PCR Detection System (Bio-Rad). The expression levels of mRNA were expressed as $2^{-\Delta Ct}$ and normalized to beta actin levels. Primer sequences were as follows: *Mki67* forward primer: 5'-TTCCTTCAGCAAGCCTGAG-3'; *Mki67* reverse primer: 5'-GTATTAGGAGGCAAGTT-3; *Ccnb1* forward primer: 5'-AGAGGTGGAACTTGCTGAGCCT-3'; *Ccnb1* reverse primer: 5'-GCACATCCAGATGTTTCCATCGG-3'; *Egr1* forward primer: 5' AGCCGAGCGAACAACCCTAT-3'; *Egr1* reverse primer: 5'-TGTCAGAAAAGGACTCTGTGGTCA-3'; *Gfap* forward primer: 5'-GGTTGAATCGCTGGAGGAG-3'; *Gfap* reverse primer: 5'-CTGTGAGGTCTGGCTTGG-3';

Gpld1 forward primer: 5'-ACCCTAACCCAAGTCCTGCT-3'; *Gpld1* reverse primer: 5'-CAGGTCAGTCAGGTGCAGAA-3'; *Nfkbia* forward primer: 5'-GCCAGTGTAGCAGTCTTGAC-3'; *Nfkbia* reverse primer: 5'-GCCAGGTAGCCGTGAGTAG-3'; *Actb* forward primer: 5'-AGGGAAATCGTGCGTGACATCAAAGAG-3'; *Actb* reverse primer: 5'-GGAGGAAGAGGATGCGGCAGTGG-3'.

Immunocytochemistry

NSPs and differentiating cells were cultivated on poly-L-ornithine coated coverslips and fixed with 4% paraformaldehyde in PBS, followed by a permeabilization with 0.05% Triton X-100 in PBS. We performed a blocking step in PBS containing 3% BSA at room temperature for 1 hour. We conducted immunostainings by incubating cells with primary antibodies diluted in blocking solution overnight at 4°C, followed by incubation with the corresponding secondary antibodies diluted in the blocking solution for 1 hour at room temperature (1:1000; Jackson ImmunoResearch). We used the following antibodies: rabbit anti-CD44 IgG (1:100; Abcam), rabbit anti-Kir4.1 IgG (1:100; Alomone Labs), rabbit anti-GLT1 (SLC1A2) IgG (1:100; Abcam), rabbit anti-GLAST (SLC1A3) IgG (1:100; Abcam), mouse anti-MAP2 IgG (1:200; Chemicon) and mouse anti-O4 IgM (1:100; Bio-Techne). For GFAP staining, we used a cyanine 3-conjugated mouse anti-GFAP IgG antibody (1:400; Sigma). Cells were then washed and mounted with DAPI-Fluoromount G (SouthernBiotech, USA). Images were collected by confocal microscopy using a LSM 510 confocal microscope (Carl Zeiss Micro Imaging) and analyzed on Adobe Photoshop (San Jose, CA, USA) software.

Preparation of mouse forebrain cell suspensions from GFAP-GFP transgenic mice and flow cytometry analyses

GFAP-GFP transgenic mice (Gfap-eGFP mice (FVB/N- Tg(GFAPGFP)14Mes/J) were purchased from The Jackson Laboratory (Bar Harbor, USA). The genetic identification of transgenic mice was determined by the analysis of DNA extracted from tails of 3-week old mice. DNA was extracted from tail samples (<5 mm) using "DirectPCR Lysis Reagent" (Viagen Biotech, USA) containing freshly prepared 20 mg/ml of proteinase K (Invitrogen, Belgium). After a lysis incubation step at 55°C for 5 hours under gentle shaking (550 rpm), proteinase K was inactivated by incubating lysates at 85°C for 45 minutes. DNA analysis was performed by Polymerase Chain

Reaction (PCR). The PCR reaction mixture, with a total volume of 25 ml, contained 1 ml of DNA lysate, 12.5 ml of a PCR master mix (50 U/ml Taq DNA polymerase, 400 μM of each dNTP, 3 mM MgCl2; Promega; The Netherlands), and 11.5 ml of primers mix (Eurogentec, Belgium). Primer sequences were as follows: GFP forward primer: 5'-AAGTTCATCTGCACCACCG-3'; GFP reverse primer: 5'-TCCTTGAAGAAGATGGTGCG-3'; internal positive control forward primer: 5'-CTAGGCCACAGAATTGAAAGATCT-3'; internal positive control reverse primer: 5'-GTAGGTGGAAATTCTAGCATCATCC-3'. Initial denaturation at 94°C for 1 minute was followed by 35 cycles of denaturation at 94°C for 30 seconds, annealing at 60°C for 1 minute and extension at 72°C for 1 minute, with final extension at 72°C for 2 minutes. PCR samples were analyzed on a 2% agarose gel (ThermoFisher Scientific, Belgium), revealed on a ChemiDoc XRS+ System (Bio-Rad, Belgium) and visualized using ImageLab software (Bio-Rad).

We collected forebrains from GFAP-GFP transgenic mice in calcium/magnesium-free HBSS at different postnatal developmental days (P4, P10 or P21). Tissue was diced and papain digested at 33°C for 90 minutes (20 U/ml, Sigma) in dissociation buffer [EBSS (Sigma), D(+)-glucose 22.5 mM, NaHCO3 26 mM and DNaseI 125 U/ml with EDTA 0.5 mM and L-cysteine–HCl 1 mM (Sigma)] and washed 3 times in dissociation buffer with BSA (1 mg/ml, Sigma) and trypsin inhibitor (1 mg/ml, Sigma) before mechanical dissociation through 5 ml and fire-polished Pasteur pipettes to a single cell suspension. Cells were pelleted, re-suspended in cold phosphate-buffered saline (PBS) with DNaseI at 1×10^6 cell/ml, passed through a 70-μm filter and 7-aminoactinomycin D (7-AAD, Sigma) added. FITC-positive/PE-Cy5-negative cells were taken into account for quantification. We performed flow cytometry analyses using a FACS Aria I SORP running FACS Diva6.3 software (BD Biosciences).

RNA-sequencing analyses

For RNA-sequencing analyses, murine NSPs were kept under proliferation conditions or were differentiated into astrocytes for 24 and 72 hours without or with TNF treatment and for 1 week under normal conditions. We extracted total RNA with TRIzol (Invitrogen) following the manufacturer's recommendations. We added a DNase treatment to the RNA extraction using DNA-free kit DNase Treatment and Removal Reagents (Ambion-Thermofisher cat #AM1906) following manufacturer's instructions. We checked the RNA quality using a Fragment Analyzer Systems (Agilent Technologies), RQN were between 3.4 and 10. We converted total RNA into Stranded Total RNA Library for RNA-sequencing by using the TruSeq Stranded Total RNA

Library Prep workflow with Ribo-Zero Gold kit according to Illumina's instructions. Briefly, ribosomal RNA was removed from 300 ng of total RNA using biotinylated, target-specific oligos combined with Ribo-Zero Gold beads that deplete cytoplasmic and mitochondrial rRNA. Then, RNA was fragmented using divalent cations under elevated temperature. Cleaved RNA fragments are double stranded cDNA converted keeping strand-specificity. A single index adapter was ligated to each sample and DNA fragments were enriched using a polymerase. All libraries have been quantified using a Qubit HS dsDNA kit (Thermofisher) and the library quality check has been performed using a High Sensitivity NGS Fragment Analysis Kit on a Fragment Analyzer System (Agilent Technologies). We normalized indexed DNA libraries and pooled them at 10 nM. Libraries were paired-end sequenced at LuxGen Genome Center with 75bp length on a NextSeq Illumina sequencer.

Fastq files contain raw sequenced data and FastQC tool was used for quality check. The fastq data was trimmed to remove bad quality reads using Trim Galore and trimmed reads were then aligned to the transcriptome using STAR aligner. STAR produces unsorted BAM files as output, which were then sorted using Samtools. We used the sorted files as input for Feature Counts to get gene expression counts. We conducted differential gene expression analyses using DESeq2 package in R.

ATAC-sequencing analyses

Cells grown as either free floating spheres (NSPs) or as astrocytes (for 24h and 72h non-treated or treated with TNF and for 1 week without TNF treatment) in poly-L-ornithine coated plates were washed once with ice-cold PBS and were spun at 500 g for 5min in a pre-chilled (4°C) fixed-angle centrifuge. We then re-suspended the cells in 200 µl of lysis buffer (5 mM PIPES pH 8, 85 mM KCl, 0.5% NP-40) and incubated them for 20 min on ice. Next, 800 µl of resuspension buffer (10 mM Tris pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Tween-20) were added to the cells. Following cell counting, we centrifuged 50,000 cells at 500 g for 10 min in a pre-chilled (4°C) fixed-angle centrifuge. The cell pellet was re-suspended in 50 µl of transposition mix (25 µl 2× TD buffer, 2.5 µl transposase, 17 µl PBS, 0.5 µl 10% Tween-20 and 5 µl water). Transposition reactions were incubated at 37°C for 40 min in a thermomixer with shaking at 1'000 rpm. Reactions were cleaned up with MinElute Raction purification kit (Qiagen) following manufacturer's recommendations. Following purification, library fragments were amplified using 1× NEBnext PCR master mix and 0.6 µM of custom PCR primers 1 and 2, using the following PCR conditions: 72°C for 5 min; 98°C

for 30 s; and thermocycling at 98°C for 10 s, 63°C for 30 s and 72°C for 1 min. Full libraries were amplified for five cycles, after which an aliquot of the PCR reaction was used together with 9 µl of the PCR cocktail with Sybr Green at a final concentration of 0.05×. We ran this reaction for 30 cycles to determine the additional number of cycles needed for the remaining reaction. Libraries were purified using AMPure XP magnetic beads (Beckman Coulter), quantified using qubit fluorimeter and diluted to make 4nM final concentration. We sequenced pooled libraries on NextSeq500 75bp paired end reads. We used the FastQC tool to perform quality check on fastq files containing raw sequenced reads and bad quality reads were removed using Trim Galore. The trimmed fastq reads were aligned to the genome using Bowtie2. We used Samtools to convert the SAM output files from Bowtie2 into sorted BAM files and to remove reads aligned to mitochondrial chromosomes. We used MACS2 to call chromatin accessibility peaks and the peaks that were found in the blacklisted regions of the genome (ENCODE) were removed before further analyses. Peaks were annotated using annotate peaks tool from HOMER and differential accessibility analysis was performed using DiffBind package in R.

Transcription factor footprint analysis

We used the Regulatory Genomic Toolbox to study the changes in the transcription factor binding motifs in open chromatin regions of the genome. Using the HINT tool (Li, Schulz et al. 2019), we identified transcription factor motifs in the epigenomic landscape. We conducted differential motif analysis using the rgt-motif analysis tool to identify transcription factor binding motifs differentially present across two conditions, thus providing information of underlying differences in gene regulation.

Reconstruction of gene regulatory networks

We re-constructed gene regulatory networks and identified hubs based on topology analyses. First, we inferred gene regulatory networks (GRNs) at each time point by using GRN inference tool GENIE3 (Huynh-Thu, Irrthum et al. 2010), which utilises Random forest-based ensemble methods to identify regulatory links between genes. We performed bootstrapping on these interactions and only selected interactions with significant value of more than zero in more than 50 percent of runs. To obtain high confidence networks, we further overlaid the re-constructed networks from GENIE3 with manually curated interactions from Metacore database. We further pruned the GRN by contextualising it using a tool developed in our group (https://gitlab.lcsb.uni.lu/andras.hartmann/GRNOptR). Briefly, the tool utilizes a parent network

and differential gene expression data with respect to the preceding time-point to create contextualised gene expression networks. Contextualization removes nodes whose expression cannot be justified by the nodes regulating it as input. We then identified the regulatory hubs by performing topological analyses using the igraph package in R. We further used the clusterprofiler package in R to perform Gene Ontology pathways enrichment of transcription factors included in the re-constructed GRNs. We visualized the resultant networks using Cytoscape.

Statistical analyses

Data were analyzed using GraphPad Prism 8 software (GraphPad software, La Jolla, CA, USA) and R environment (R Core Team, Vienna, Austria). Unless otherwise indicated, all data are presented as mean ± standard deviation (SD) or standard error of the mean (SEM) of at least three independent biological experiments. We performed statistical analyses using unpaired t test or two-way ANOVA. All differences were considered significantly different at p value <0.05 and were annotated as follows: *<0.05, **<0.01.

Results

Characterization of an *in vitro* model of astrocytic differentiation from murine neurospheres

To investigate the molecular mechanisms underlying NSCs differentiation into astrocytes under inflammatory conditions, we first characterized our *in vitro* cellular model of primary neurospheres (NSPs) derived from NSCs isolated from the ventricular zone at embryonic day-14 (E14) of the mouse embryos. Briefly, we differentiated NSPs into astrocytes in the presence of 10% fetal calf serum (FCS) for one or two weeks and analyzed them at specific intermediate stages, including 24, 48 and 72 hours (**Fig. 1A**). Under these conditions, differentiating NSPs showed progressive increased expression levels of the pan-astrocytic immature marker GFAP and a concomitant decreased expression of the stem cell marker CD44 (**Fig. 1B**). After 2 weeks of differentiation, mature astrocytic markers, such as Kir4.1 (*Kcnj10*), GLT1 (*Slc1a2*) and GLAST (*Slc1a3*) were highly expressed (**Fig. S1A**). Notably, we did not detect neurons and oligodendrocytes in these culture conditions (**Fig. S1B-C**), while multipotentiality of the NSCs was confirmed by the appearance of these cells upon addition of 1% B27 and 2 ng/ml EGF to 0,5% FCS medium (**Fig. S1D-E**). At the transcriptional level, differentiating astrocytes at all stages showed significant decreased expression levels of the proliferating (*Mki67* and *Ccnb1*) and stem cell (*Egr1*) markers compared to NSPs (**Fig. 1C**). On the other hand, the expression levels of *Gfap* and the astrocyte-enriched gene *Gpld1* were mainly up-regulated in differentiating cells at the different time points. We detected high *Gfap* expression levels at 48 and 72 hours, which dropped at 1 week, while *Gpld1* expression was not detectable in NSPs and showed progressively increased levels up to 1 week of differentiation (**Fig. 1D**). In order to corroborate these results *in vivo*, we investigated the decrease of *Gfap* expression along the astrocytic maturation by isolating GFP+ cells from GFAP-GFP transgenic mice. For this, we separated GFP+ cells by FACS at postnatal day 4 (P4), P10 and P21, with P4 representing an immature phase, P10 signifying an intermediate stage, while P21 corresponding to a mature astrocytic stage. In line with the *in vitro* model, the amount of GFAP-GFP+ cells gradually decreased along these time points (P4: 39.2 ± 3.6%; P10: 21.0 ± 0.9%; P21: 3.5 ± 1.3%) (**Fig. 1E**). Notably, we confirmed the upregulation of *Gpld1* as a marker of astrocytic maturation *in vivo* from our previous transcriptomics analyses of GFP+ cells sorted from ALDH1L1-GFP transgenic mice when comparing cells isolated at P4, P10 and P21 (**Fig. S2A**) (Michelucci, Bithell et al. 2016). We similarly verified their progressive exit from the cell cycle by detecting the decrease of *Ccnb1* expression levels (**Fig. S2B**). Lastly, to verify the purity of our

in vitro cultures, we sorted enriched astrocytic populations by MACS using an antibody against the pan-astrocytic marker GLAST (*Slc1a3*). GLAST+ cells exhibited similar expression levels of *Mki67, Ccnb1, Gfap* and *Gpld1* to unsorted cells (**Fig. S2C-D**). Hence, this common gene expression pattern observed in GLAST+ and GLAST- cells suggests that our culture conditions give rise to pure astrocytic populations expressing GLAST levels consistent with their maturation state.

Taken together, these results demonstrate that our *in vitro* model of NSPs differentiation into astrocytes represent a reproducible method enabling enrichment of astrocytic populations expressing markers of *ex vivo*-isolated astrocytes.

Exposure of NSPs to TNF under differentiating conditions modulates the expression of specific NSC and astrocytic markers

Next, we took advantage of our model to investigate the effect of inflammation during the differentiation of NSPs into astrocytes. For this, we cultivated the NSPs in differentiation medium with or without TNF (50 ng/ml), as in our previous study [21]. As expected, TNF treatment induced NF-κB activation in differentiating NSPs as shown by enhanced *Nfkbia* mRNA levels compared to untreated cells. This was particularly evident after 24, 48 and 72h, while *Nfkbia* mRNA levels were less enhanced at 1 week (**Fig. 2A**). While the decrease of the stem cell marker *Egr1* was less significant in the presence of TNF when compared to untreated samples (**Fig. 2B**), TNF treatment did not cause significant changes in the expression levels of *Ccnb1* (**Fig. 2C**) and *Mki67* (**Fig. 2D**). Further, *Gfap* and *Gpld1* were down-regulated in TNF-treated cells at 48 and 72 hours of differentiation for the former, and at 72 hours and 1 week for the latter (**Fig. 2E-F**).

These targeted gene expression studies, together with our recent analyses , show that exposure of differentiating NSPs to TNF induces the activation of the NF-κB pathway without affecting the exit of the differentiating astrocytes from the cell cycle. On the other hand, TNF treatment causes transcriptional changes in the expression levels of specific stem cell and astrocytic markers, thus affecting the physiologic astrocytic differentiation process.

Transcriptomics analyses under normal and inflammatory conditions reveal discrete populations of differentiating astrocytes

As TNF affected the expression of specific NSC and astrocytic genes during differentiation of NSPs into astrocytes, we sought to expand our analyses by conducting a genome-wide RNA-sequencing analysis of the astrocytic transcriptome at different time points, both under normal and inflammatory conditions. Principal component analysis (PCA) using the top 1'000 variable genes showed distinct clustering corresponding to the different conditions (**Fig. 3A**). We identified differentially expressed genes ($p < 0.05$) between different conditions and compared them at 24 and 72h under normal or inflammatory conditions. Specifically, we detected 4'187 overexpressed genes and 4'544 down-regulated genes at 24h compared to undifferentiated NSPs, both with or without TNF. On the other hand, we detected 1'005 and 756 genes, respectively exclusively up- or down-regulated at 24h under control conditions, while 1'227 and 922 genes were correspondingly enhanced or decreased only under TNF treatments. We obtained similar results at 72h (**Fig. 3B**). Hierarchical gene clustering of the differentially expressed genes during normal differentiation (up- and down-regulated genes, adjusted p value < 0.05) revealed five major clusters differentially represented across the specific time points (**Fig. 3C**). Corresponding Gene Ontology (GO) enrichment analyses indicated that proliferating NSPs were mainly associated with cluster 1, largely characterized by RNA processing. Clusters 2 and 4, largely represented in NSPs and differentiating astrocytes at 24h, were characterized by processes associated with regulation of cell morphogenesis, motility and migration. Clusters 3 and 5, mostly enriched in differentiating astrocytes at 72h and 1 week, were characterized by terms linked to cell projection assembly and cilium organization (**Fig. 3D**).

We conducted similar analyses to investigate the effect of TNF treatment at 24 and 72h. Hierarchical gene clustering of the differentially expressed genes during differentiation under normal and inflammatory conditions (up- and down-regulated genes, adjusted p value < 0.05) identified four major clusters differentially represented across the specific settings (**Fig. 3E**). Clusters 1 and 4, characterized by terms associated with response to molecule of bacterial origin and regulation of inflammatory response, were enriched under TNF treatment, at both 24 and 72h. On the contrary, cluster 2 described by terms linked to regulation of cell motility and migration, was under-represented under inflammatory conditions. Notably, we observed that the presence of TNF, especially when comparing conditions at 24h, decreased cluster 3, related to gliogenesis, glial cell differentiation and development (**Fig. 3F**).

Next, we used heat maps to visualize gene expression levels across different time points between normal and inflammatory conditions of specific genes associated with different astrocytic categories. Specifically, we analyzed markers linked to "developing and differentiated astrocytes", "reactive astrocytes", "A1 - neurotoxic - astrocytes" and "A2 - neuroprotective - astrocytes" taking advantage of different available datasets (Lovatt, Sonnewald et al. 2007, Cahoy, Emery et al. 2008, Liddelow, Guttenplan et al. 2017, Escartin, Galea et al. 2021) (**Fig. S3A-D**). Notably, astrocytic gene markers, including *Gfap*, *Slc14a1* and *Gja1* were up-regulated at 72h in the absence of TNF (**Fig. S3A**), while reactive astrocyte marker genes, including *Timp1, Lcn2, Icam1* and *Ptx3* were enhanced under inflammatory conditions, mainly at 24h or *Sbno2, Cxcl10* and *Cd109* (Zamanian, Xu et al. 2012, Grill, Syme et al. 2015, Rothhammer and Quintana 2015) at 72h (**Fig. S3B**). Further, we showed that the obtained astrocytic populations express a mixture of A1 and A2 markers, both under normal and inflammatory conditions **(Fig. S3C-D)**. Lastly, due to the importance of glycogen and glucose metabolism in astrocytes, we also looked at the expression levels of genes associated with "glycogen activity" and "glucose metabolism" (**Fig. S3E-F**). Along the differentiation process, we detected increased expression levels of genes related to glycogen activity (*Ugp2, Pygb, Ppp1r3c*) and glucose metabolism (*Lmbrd1, Pid1, Esr1, Upk3b, Lep),* which were decreased under TNF exposure at 72h (**Fig. S3E-F**).

Taken together, these results show that genes linked to gliogenesis and glial cell development are down-regulated in the presence of TNF, suggesting that an inflammatory environment alters the physiological astrocytic differentiation, thus affecting critical associated functional activities, including glycogen and glucose metabolic pathways.

Chromatin accessibility profiling detects extensive chromatin remodeling along the astrocytic differentiation affected under TNF exposure

To understand further the mechanisms underlying changes in gene expression under normal and inflammatory conditions, we sought to examine the corresponding chromatin accessibility profiles, which affect, at least partly, gene expression depending on their more open or closed states at specific loci. To do this, we performed ATAC-sequencing analyses under the same conditions as the transcriptome studies. A first PCA analysis using differential peaks discriminated the various time points and highlighted differences between treated and untreated samples. Of note, similarly to the PCA analyses of the RNA-seq samples, PC1 seemed to separate samples according to the

developmental stage, while PC2 divided them according to stage and treatment (**Fig. 4A**). Next, we conducted our analyses based on differentially accessible peaks in the promoter region nearest to the transcription start site (TSS). We identified differentially accessible peaks ($p < 0.05$) up- and down-regulated between different conditions and compared them at 24 and 72h under normal or inflammatory conditions. Specifically, we detected 545 up-regulated and 9'081 down-regulated accessible peaks at 24h compared to undifferentiated NSPs, both with or without TNF. On the other hand, we detected 102 and 960 accessible peaks, respectively exclusively up- or down-regulated at 24h under control conditions, while 297 and 385 accessible peaks were correspondingly enhanced or decreased only under TNF treatments. We found similar proportions at 72h (**Fig. 4B**). Hierarchical gene clustering of the differentially accessible peaks across the analyzed normal differentiation time points (up- and down-accessible peaks, adjusted p value < 0.05) revealed five major clusters differentially represented across the specific time points (**Fig. 4C**). Clusters 1, 3 and 5, mainly associated to proliferating NSPs, were the highest represented in the heatmap, thus indicating that multipotent NSPs have higher chromatin accessibility when compared to differentiated astrocytes (**Fig. 4C**). NSPs were characterized by GO terms related to RNA processing, histone and covalent chromatin modifications (clusters 1 and 3), synapse organization and axonogenesis (clusters 3 and 5) as well as lipid and monovalent inorganic cation transport (cluster 5). On the other hand, cluster 2, mainly enriched in NSPs and 1 week-differentiated astrocytes, was represented by GO terms related to response to chemical stress and external stimulus. Lastly, cluster 4, mostly enhanced in late-stage astrocytes, showed terms linked to lipid localization and transport (**Fig. 4D**).

Under inflammatory conditions, we detected increased chromatin accessibility in TNF treated cells when compared to untreated cells. More precisely, following a TNF treatment at both 24 and 72h, prevalent clusters 1,2 and 3 were accessible (**Fig. 4E**). Enrichment analysis of these clusters specifically showed involvement of the NF-κB signaling pathway (cluster 2), results that come to agreement with the RNA-seq data. Cluster 4, which was under-represented under inflammatory conditions, was associated with GO terms related to cell junction assembly, while cluster 5, mainly enriched in 24h-differentiated astrocytes, was linked to regulation of developmental growth (**Fig. 4F**). Lastly, we used volcano plots to analyze differentially accessible peaks across time points between normal and inflammatory conditions. Here, a comparison between TNF treated and

untreated samples showed a skewed distribution towards peaks that were more open under inflammatory conditions, both at 24 and 72h (**Fig. S3G**).

Overall, we observed significant changes in chromatin accessibility upon astrocytic differentiation compared to undifferentiated cells, with genes associated, for example, with lipid localization and transport displaying increased accessible peaks in their promoter region. Upon TNF treatment, we detected a strong shift towards enhanced chromatin accessibility compared to normal conditions, with several gene clusters related to inflammation.

Transcription factor binding analyses uncover dynamic footprints associated with astrocyte specification and maturation under normal and inflammatory conditions

Next, we investigated changes in the transcription factor binding motifs across accessible chromatin regions by footprint analyses using HINT-ATAC (Li, Schulz et al. 2019) to detect transcription factors (TFs) associated with astrocyte specification and maturation with or without TNF treatment.

We first analyzed differences in the TF binding motifs between NSPs and differentiating astrocytes at 24 and 72h under normal conditions. At 24h, we detected an enriched activity score for *Atoh1*, *Dlx1*, *Rarg* and *Arid3a* with a concomitant decreased score for *Dmbx1* and *Tcf21* TFs when compared to NSPs (**Fig. 5A**). Similar analyses conducted at 72h showed increased activity score for *Hoxd8*, *Lhx3*, *Rarb*, *Rarg* and *Dlx1*, together with a decreased score for *Hes1*, *Tcf21* and *Tcfl5* TFs compared to NSPs (**Fig. 5B**). Hence, based on these analyses, ATOH1 and ARID3A represent TFs associated with astrocytic specification (24h), HOXD8, LHX3 and RARB represent TFs linked to astrocyte maturation (72h), and DLX1 and RARG are TFs linked to both processes (**Fig. 5A-B**).

By conducting similar analyses under inflammatory conditions, we detected at 24h increased activity score for *Nfe2l2 (Nrf2)* and a decreased score for *Dmbx1*, *Arid3a*, *Arid3b* and *Atoh1* TFs when comparing to normal conditions (**Fig. 5C**). Lastly, at 72h *Atoh1* and *Rarg* TFs showed an enhanced score by comparing TNF-treated versus untreated cells, while the score for *Tcf21* and *Dmbx1* was decreased (**Fig. 5D**). Due to the importance of *Nfe2l2 (Nrf2)* in regulating inflammatory responses, mainly by inducing the expression of antioxidants and cytoprotective genes, we analyzed its binding motifs across accessible chromatin regions at the genome-wide level. We detected enriched footprints across the whole genome in TNF-treated cells at 24h (**Fig.**

5E), thus suggesting that NRF2 represents a master TF driving the astrocytic differentiation in the presence of TNF.

Transcriptional and chromatin accessibility states positively correlate at specific NSC and astrocytic gene loci

Next, we sought to study the correlation between transcriptional and chromatin accessibility states at the genome-wide level to explore at which extent the chromatin states are linked to the transcriptional activities at specific gene loci. For this, we took advantage of our RNA-seq and ATAC-seq data to conduct correlation analyses between gene expression levels and their chromatin accessibility states at 24h. We used scatterplots to show the number of genes differentially expressed and exhibiting differentially accessible peaks at their promoter regions under normal (**Fig. 6A**) and inflammatory (**Fig. 6B**) conditions. By using Pearson's correlation analysis for log fold change in gene expression and accessibility, we identified similar correlation coefficients (normal: R=0.38; inflammation: R=0.51) to those detected in analogous studies comparing RNA-seq and ATAC-seq datasets (Ackermann et al. 2016, Starks et al., 2019). In a second step, we conducted a targeted correlation analysis at specific NSC, resting and reactive astrocytic gene loci (Lovatt, Sonnewald et al. 2007, Cahoy, Emery et al. 2008, Liddelow, Guttenplan et al. 2017, Escartin, Galea et al. 2021). Here, we calculated at 24h a higher correlation score (R=0.64) compared to genome-wide levels. Under normal conditions, we showed that specific NSC (e.g. *Ccnb1*, *Mki67*, *Egr1*) or astrocytic (e.g. *Gfap, Gpld1*, *Clcf1*) markers exhibited respectively decreased or increased gene expression levels and accessibility (**Fig. 6C**). Under TNF exposure, we detected enhanced gene expression levels and accessibility for reactive astrocyte markers (e.g. *Lcn2*, *Cxcl10*, *Osmr*, *Sbno2*) (**Fig. 6D**). Interestingly, *Nfkbia*, a pan-inflammatory marker, exhibited both enhanced gene expression levels and accessibility states (**Fig. 6D**).

Taken together, these results indicate that at specific NSC and astrocytic gene loci, gene expression levels are associated with chromatin accessibility states, thus indicating a close link between chromatin remodeling and gene expression at lineage-specific gene loci.

Inference of gene regulatory networks enables the identification of key transcription factors associated with astrocytic differentiation under physiological and inflammatory conditions

To systematically identifying regulators of the astrocytic differentiation under physiological and inflammatory conditions, we reconstructed the corresponding gene regulatory network (GRN) for each analyzed time-point. The GRN after 24h of physiological differentiation was represented by 262 interactions between 96 TFs, which, based on the topological analysis, include *Esr1, Nr3c1, Foxa1, Gata2, Stat5a, Fosl2, Stat1, Nfia* and *Arid5b* (**Fig. 7A**). Corresponding GO analyses highlighted terms associated to developmental processes, including "regulation of developmental process", "tissue development" and "regulation of cell differentiation" (**Fig. S4A**). At 72h, when compared to 24h, the GRN consisted of 91 interactions between 46 TFs, thus reflecting a lineage-committed network made up of less than half TFs when compared to earlier stages (**Fig. 7B**). The top regulatory hubs include *Rxra, Rarg, Cebpb, Myc, Smad3, Rbpj* and *Esr1*. Subsequent GO analyses showed enriched terms related to cell responses and communication, such as "cellular response to stimulus", "signaling" and "signal transduction" (**Fig. S4B**), thus indicating that at earlier stages differentiated cells went through the developmental process (astrocyte specification) and later acquired their functional astrocytic properties (astrocyte maturation). Interestingly, 10 TFs, including *Esr1, Foxp1* and *Nfe2l2*, were shared between the two GRNs (**Fig. S4C**). Out of these, we identified *Esr1* and *Foxp1* as regulatory hubs at both time points.

Similarly, we reconstructed GRNs under inflammatory conditions. Following a 24h differentiation, the GRN consists of 107 TFs and 393 interactions (**Fig. 8A**). We identified *Stat3, Smad3, Nfkb1, Runx1, Ar, Cebpb, Relb, Nfkb2* and *Atf3* as regulatory hubs. Out of these TFs, 75 of them, such as *Esr1, Nr3c1, Gata2, Stat5a, Fosl2, Sat1, Junb, Mef2c, Arid5b, Mef2a* and *Foxo3*, were shared with the corresponding GRN under normal conditions (**Fig. S5A**), suggesting that the developmental process and the inflammatory response converge during the acquisition of the astrocytic phenotype. GO analyses of the TFs underlying the GRN at 24h under inflammatory conditions strengthen this observation, with terms associated to both developmental processes, such as "regulation of developmental process", and inflammatory responses, including "immune system process", be highly represented (**Fig. S5B**). Lastly, following a 72h differentiation under TNF treatment, the GRN is composed of 18 interactions between 15 TFs, including *Rxra, Myc, Stat5b* and *Rorc* (**Fig. 8B**). Corresponding GO analyses identified "response to chemical", "response to hormone" and "response to lipid" as main enriched terms (**Fig. S5C**). Among the resultant TFs, 12 of them, including *Rxra, Rarg, Stat5b* and *Myc* were in common with the

corresponding network under physiological conditions, while 3 TFs (*Thrb, Zfp217, Zbtb37*) were unique to the inflammatory condition (**Fig. S5D**).

Taken together, reconstructed GRNs enabled the identification of regulatory hub TFs playing critical roles along the differentiation of astrocytes in the presence or absence of TNF. These analyses allowed establishing the effect of an inflammatory environment along the astrocytic development, indicating that it affects the underlying transcriptional signatures skewing astrocytes towards a discrete astrocytic phenotype when compared to physiological settings.

Discussion

Astrogliogenesis depends on both extracellular cues and intrinsic dynamic changes of the epigenome. Perturbations during astrocytic development may result in neurological diseases, including neurodevelopmental disorders, such as Rett syndrome and fragile X mental retardation (Molofsky, Krencik et al. 2012) or Down syndrome and autism spectrum disorders (Sloan and Barres 2014). In recent years, computational approaches to study astrocytes and their development using next-generation sequencing technologies have enabled a better understanding of their phenotypic and functional role in the CNS (Magistri, Khoury et al. 2016, Goodnight, Kremsky et al. 2019). Further, the generation of novel *in vitro* models, such as astrocytes derived from induced pluripotent stem cells, has opened up a new area for studying neurological diseases *in vitro*. However, as the brain hosts a range of astrocyte populations, it is critical to develop standardized protocols for the *in vitro* generation of astrocytes with defined maturity states and phenotypic properties(Chandrasekaran, Avci et al. 2016).

In vitro model characterization

In this context, our *in vitro* NSP differentiation model enabled us to conduct kinetic analyses along the process of astrocytic differentiation. In this model, under specific culture conditions NSCs retain their multipotential properties, thus being able to give rise to neurons, astrocytes and oligodendrocytes. For our experiments, we differentiated NSPs into astrocytic populations by cultivating them in the presence of serum, as previously described(Birck, Ginolhac et al. 2021). In these conditions, we detected the highest *Gfap* expression levels at 48 and 72 hours, whereas its expression decreased after 1 week. *Gfap* expression pattern is supported by previous *in vivo* studies, documenting that although GFAP is a reliable marker of reactive astrocytes, it is mainly

not detectable in mature astrocytes of the healthy central nervous system (Sofroniew and Vinters 2010).

Targeted gene expression and genome-wide transcriptional analyses

We have previously shown that TNF is a key regulator of astrocytic reprogramming into neural progenitor-like cells (Gabel, Koncina et al. 2016). In addition, TNF treatment of *in vitro* differentiated murine astrocytes resulted in modifications of their histone chromatin profile at the promoters of genes related to cell cycle, stemness or neuronal fate, thus indicating that inflammatory events intervene in such regulatory pathways . In the present study, RNA-seq analyses showed major gene expression changes along the physiological process of the astrocytic differentiation, which are modulated by TNF, suggesting the establishment of discrete astrocytic populations between normal and inflammatory conditions. More specifically, TNF-treated NSC-derived astrocytes exhibited significant increased expression levels of classical reactive astrocytic marker genes, including *Timp1, Lcn2, Icam1* and *Ptx3*. In line with these observations, reactive astrocytes in inflammatory environments display increased expression levels of various proteins that are absent or weakly expressed in resting states. Among those proteins, LCN2 and TIMP1 are highly up-regulated in reactive astrocytes, with LCN2 considered as a potential marker of astrogliosis . ICAM-1 is aberrantly expressed by astrocytes in CNS pathologies, such as multiple sclerosis, experimental allergic encephalomyelitis and Alzheimer's disease, suggesting a possible role for ICAM-1 in these disorders (Lee, Drabik et al. 2000). Lastly, *Ptx3* is considered as a marker of A2 reactive astrocytes (Liddelow, Guttenplan et al. 2017, Neal, Luo et al. 2018). Hence, TNF-treated NSC-derived astrocytes express markers of both A1 and A2 reactive types, thus representing a distinct type of astrocytes that do not exclusively relate to the described dichotomous distribution.

Notably, the loss of normal astrocytic functions and not only reactivity contributes to the early pathologic events of neurological diseases, as described for example in multiple sclerosis (Parratt and Prineas 2010, Brosnan and Raine 2013). In this context, we sought to examine genes associated with important astrocytic functions, including glucose import from blood, the primary energy source for generating ATP in the brain, and the metabolism of glycogen, which represents an important energy reserve synthesized from glucose in astrocytes(Belanger, Allaman et al. 2011). Indeed, astrocytes represent an exclusive reserve of glycogen, which is transported to axons for

ensuring neuronal functions (Brown and Ransom 2007, Brunet, Allaman et al. 2010, Iglesias, Morales et al. 2017, Bak, Walls et al. 2018). Thus, the detection of decreased expression levels of genes associated with glucose and glycogen metabolism under TNF exposure suggests an impairment of key functional activities when NSCs differentiate into astrocytes under inflammatory conditions.

Genome-wide chromatin accessibility profiling

The ATAC-seq method enabled us to analyze the accessible chromatin landscape of differentiating astrocytes and the extent of its modifications upon inflammation, while providing insights into gene transcription complexity. We observed a general decrease of chromatin accessibility along NSC differentiation into astrocytes, thus suggesting a more lineage-restricted profile. However, upon TNF treatment, chromatin accessibility at the promoter region of specific NSC, astrocytic and inflammatory markers was up-regulated, indicating that inflammation-induced astrocytes retain some stemness vestige. The next step would be to identify which epigenetic mechanisms, such as DNA methylation or histone modifications, may be linked to the transcriptional and chromatin accessibility alterations that we identified either in the physiologic differentiating astrocytes or under inflammatory conditions. Intriguingly, a recent study showed that aberrant transcription in gliomas does not result from DNA methylation mechanisms, but rather from DNA methylation-independent modifications (Court, Le Boiteux et al. 2019).

Footprint analyses

Footprint analyses identified, among several modulated TF binding motifs across accessible chromatin regions, decreased levels of the downstream effector of the Notch pathway, HES1, when comparing differentiated astrocytes at 72h with NSCs. Since HES1 is involved in the repression of pro-neural gene expression and maintenance of NSCs (Ahmed, Gan et al. 2009, Imayoshi, Sakamoto et al. 2010, Yun, Byun et al. 2010), the preservation of its binding motifs under inflammatory conditions is in line with a less lineage-restricted fate of inflammation-induced astrocytes when compared to the corresponding naïve cells. Further, we detected increased levels of NRF2 binding motifs across accessible chromatin regions at 24h under inflammatory conditions. NRF2 is a master regulator of stress induced antioxidant response inducing the expression of antioxidants and cytoprotective genes(Habas, Hahn et al. 2013). In the brain, NRF2

is involved in astrocytic neuroprotective pathways via their interactions with neurons (Habas, Hahn et al. 2013, Liddell 2017). Interestingly, in our dataset, the NRF2 target gene *Cebpb* (Hayes and Dinkova-Kostova 2014) was both upregulated and accessible in the presence of TNF. Taken together, our results point towards NRF2 as a key TF orchestrating immune responses in differentiating astrocytes exposed to TNF.

RNA-seq & ATAC-seq correlation analyses

Our correlation analyses are consistent with previous correlation studies addressing gene expression and chromatin accessibility, thus providing further support to our data . Notably, correlation levels were enhanced when looking at NSC, astrocytic and inflammatory markers. Further, our chromatin accessibility and gene expression analyses resulted in similar biological processes and pathways, thus indicating concordant modules of regulation at these molecular levels.

Reconstructed gene regulatory networks

Gene regulatory networks (GRNs) inferred from our transcriptional analyses enabled the identification of TFs underlying the astrocytic differentiation under physiological and inflammatory conditions. Different pathways are known to be sequentially involved in the astrocytic differentiation process, with the Notch, STAT3 and BMP signaling pathways being the most important ones implicated in this process (Wen, Li et al. 2009). At 24h of differentiation under normal conditions, several TFs related to these pathways were highly represented in the GRN. These results validate once again the relevance of our *in vitro* model, which is often described in the literature as very different from the differentiation processes obtained *in vivo* (Lattke, Goldstone et al. 2021). Indeed, under our established differentiation conditions in the absence of growth factors and mitogen cocktails found in other studies (Tiwari, Pataskar et al. 2018, Lattke, Goldstone et al. 2021), we obtained an efficient and reproducible process of astrocytic differentiation. Among the key detected regulators, *Esr1*, *Nr3c1*, *Foxp1*, *Runx2* and *Nfia* appears particularly interesting as these TFs have been ascribed to this process. For example, *Esr1*-dependent signaling in astrocytes is involved in the regulation of several important pathways, such as PI3K-AKT, ERK and JAK-STAT, and plays a crucial role on their metabolic control of

providing nutrients to the neurons (Spence, Hamby et al. 2011, Fuente-Martin, Garcia-Caceres et al. 2013) . *Nr3c1*, also known as glucocorticoid receptor, is an important component of the metabolic functions of astrocytes and is required for memory formation(Tertil, Skupio et al. 2018) . NSCs express *Foxp1* promoting their maintenance and renewal (Pearson, Moore et al. 2020), and is required for differentiation towards astrocytes (Braccioli, Vervoort et al. 2017) . *Runx2* promotes maturation-associated transcriptional changes in astrocytic cell culture models(Tiwari, Pataskar et al. 2018). Lastly, *Nfia* contributes to the gliogenic switch enabling rapid derivation of functional human astrocytes from pluripotent stem cells (Tchieu, Calder et al. 2019). Its expression is required for the maintenance of the astrocytic morphological complexity and its deficiency compromises the astrocyte-neuron communication (Zhang, Xu et al. 2021). Notably, most of the detected TFs show reduced chromatin accessibility after 24h of differentiation, thus suggesting that they are implicated in the first step of differentiation, but not in the astrocytic maturation. The TFs identified at 72h relate to the maturation and development of astrocytic functions. Compared to earlier time points, the network is less complex, characterized by the reduction of the number of TFs composing the GRN with a robust representation of TFs related to retinoic acid (RA). In astrocytes, GFAP expression induced by RA is mediated by the activation of RA receptor alpha (Rara), which downstream activates the PI3K pathway and binds to the STAT3-p300/CBP-SMAD complex (Herrera, Chen et al. 2010). The PI3K-AKT pathway plays also crucial roles in glutamate transport (Wu, Kihara et al. 2010) and synthesis of neuroprotective RANTES in astrocytes (Lin, Sun et al. 2011). The presence of *Sox2* in the network reflects a certain degree of immaturity still at 72h of differentiation. In fact, SOX2 acts upstream the Notch pathway to maintain cell proliferation potential and its expression in the brain is restricted to neural stem and progenitor cells, glial precursors and proliferating astrocytes (Bani-Yaghoub, Tremblay et al. 2006). This observation is corroborated by the presence in the GRN of the effector of the Notch pathway, *Rbpj*, whose reduced accessibility suggests that the Notch pathway has lost part of its activity. Interestingly, reduction of Notch signaling in astrocytes promotes neurogenesis and absence of the Notch coactivator RBPJ activates neurogenesis in astrocytes (Santopolo, Magnusson et al. 2020). Lastly, the TGFB1-SMAD signaling, with *Smad3* being among the TFs whose accessibility is increased, is involved in the fate commitment process of astrocytes (Stipursky and Gomes 2007).

Under inflammation, the TFs that are involved in the GRN at 24h represent both astrocytic development and inflammatory responses. This is reflected by the overlap of regulatory hubs from

the normal condition and TFs (e.g. *Stat3, Nfkb1, Nfkb2, Runx1, Atf3, Relb*) representing the inflammatory part of the process. STAT3 and NFkB regulate the function and development of cells under inflammation (Herrmann, Imura et al. 2008, Lattke, Reichel et al. 2017, Platanitis and Decker 2018). The roles of NF-kB and STAT3 in colon, gastric, and liver cancers have been extensively investigated and the activation and interaction between STAT3 and NF-κB plays a fundamental role in the control of the communication between cancer cells and inflammatory cells (Grivennikov and Karin 2010, Fan, Mao et al. 2013). Specifically, STAT3 plays an important role in the development of astrocytes, such as it binds to the promoter region of GFAP, and its absence results in defects during astrogliogenesis (Hong and Song 2014). The decrease of chromatin accessibility at the promoter region of *Stat3* may reflect the partial inhibition of the astrocytic differentiation induced by NFkB activation. NFKB1 and NFKB2 are critical effectors of the canonical NFkB pathway regulating expression of genes involved in inflammatory and cell-survival responses (Lawrence 2009). Further, *Runx1* codes for an important TF involved in the astrocytic development and its overexpression is linked to differentiation along astrocytic and neuronal lineages (Logan, Rusnak et al. 2015). Lastly, *Cebpb* is involved in the regulation of multiple pathways associated with inflammatory responses in astrocytes (Fields and Ghorpade 2012). *Atf3* expression in astrocytes is induced by NRF2, one of the major regulators of inflammatory responses (Hunt, Raivich et al. 2012). It is interesting to note that several TFs, such as *Esr1,* are present in the GRN both under normal and inflammatory conditions at 24h of differentiation, thus suggesting that although inflammation affects the astrocytic differentiation, patterns of the normal process still drive the process of differentiation.

At 72h under inflammation, the GRN is composed by TFs involved in multiple metabolic functions. For example, *Rarg, Rorc* and *Rxra* are involved in the RA pathway, while *Thrb* is the thyroid hormone receptor. Astrocytes are the mediators of thyroid metabolism, which is linked to the proliferation and differentiation of astrocytes and also promotes neuronal development via astrocytes (Trentin 2006). The corresponding GO terms support these processes related to metabolism (e.g. “response to steroid hormone”, “response to corticosteroid” and “response to glucocorticoid”). Overall, the identified TFs support the tight interplay between development and inflammation influencing astrocytic functional activities.

Conflict of interest

The authors declare no competing interests.

Author contributions

LG, AdS and AM designed project and obtained financial support; MASP, RH, NN, CB and AM performed experiments; MASP, KS, SR, LG and AM analyzed experiments; KS and SR conducted bioinformatics analyses; LG, AdS and AM supervised research; MASP, KS and AM wrote the manuscript. All the authors edited and approved the manuscript.

Funding

MASP, LG, AdS and AM were supported by the University of Luxembourg (IRP ASTROSYS). AM was supported by the Luxembourg Institute of Health and the Luxembourg Centre for Systems Biomedicine (MIGLISYS). KS was supported by the Luxembourg National Research Fund (FNR) through the FNR-PRIDE program NEXTIMMUNE for doctoral education (PRIDE/11012546/NEXTIMMUNE).

Acknowledgments

We thank Prof Noel J. Buckley for the critical revision of the manuscript.

References

1. Adefuin, A. M., A. Kimura, H. Noguchi, K. Nakashima and M. Namihira (2014). "Epigenetic mechanisms regulating differentiation of neural stem/precursor cells." Epigenomics 6(6): 637-649.

2. Ahmed, S., H. T. Gan, C. S. Lam, A. Poonepalli, S. Ramasamy, Y. Tay, M. Tham and Y. H. Yu (2009). "Transcription factors and neural stem cell self-renewal, growth and differentiation." Cell Adh Migr 3(4): 412-424.

3. Alcantara Llaguno, S. R. and L. F. Parada (2016). "Cell of origin of glioma: biological and clinical implications." Br J Cancer 115(12): 1445-1450.

4. Arvidsson, L., R. Covacu, C. P. Estrada, S. R. Sankavaram, M. Svensson and L. Brundin (2015). "Long-distance effects of inflammation on differentiation of adult spinal cord neural stem/progenitor cells." J Neuroimmunol 288: 47-55.

5. Bak, L. K., A. B. Walls, A. Schousboe and H. S. Waagepetersen (2018). "Astrocytic glycogen metabolism in the healthy and diseased brain." J Biol Chem 293(19): 7108-7116.

6. Bani-Yaghoub, M., R. G. Tremblay, J. X. Lei, D. Zhang, B. Zurakowski, J. K. Sandhu, B. Smith, M. Ribecco-Lutkiewicz, J. Kennedy, P. R. Walker and M. Sikorska (2006). "Role of Sox2 in the development of the mouse neocortex." Dev Biol 295(1): 52-66.

7. Belanger, M., I. Allaman and P. J. Magistretti (2011). "Brain energy metabolism: focus on astrocyte-neuron metabolic cooperation." Cell Metab 14(6): 724-738.

8. Birck, C., A. Ginolhac, M. A. S. Pavlou, A. Michelucci, P. Heuschling and L. Grandbarbe (2021). "NF-kappaB and TNF Affect the Astrocytic Differentiation from Neural Stem Cells." Cells 10(4).

9. Braccioli, L., S. J. Vervoort, Y. Adolfs, C. J. Heijnen, O. Basak, R. J. Pasterkamp, C. H. Nijboer and P. J. Coffer (2017). "FOXP1 Promotes Embryonic Neural Stem Cell Differentiation by Repressing Jagged1 Expression." Stem Cell Reports 9(5): 1530-1545.

10. Brosnan, C. F. and C. S. Raine (2013). "The astrocyte in multiple sclerosis revisited." Glia 61(4): 453-465.

11. Brown, A. M. and B. R. Ransom (2007). "Astrocyte glycogen and brain energy metabolism." Glia 55(12): 1263-1271.

12. Brunet, J. F., I. Allaman, P. J. Magistretti and L. Pellerin (2010). "Glycogen metabolism as a marker of astrocyte differentiation." J Cereb Blood Flow Metab 30(1): 51-55.

13. Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf (2013). "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position." Nat Methods 10(12): 1213-1218.

14. Cahoy, J. D., B. Emery, A. Kaushal, L. C. Foo, J. L. Zamanian, K. S. Christopherson, Y. Xing, J. L. Lubischer, P. A. Krieg, S. A. Krupenko, W. J. Thompson and B. A. Barres (2008). "A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function." J Neurosci 28(1): 264-278.

15. Chandrasekaran, A., H. X. Avci, M. Leist, J. Kobolak and A. Dinnyes (2016). "Astrocyte Differentiation of Human Pluripotent Stem Cells: New Tools for Neurological Disorder Research." Front Cell Neurosci 10: 215.

16. Court, F., E. Le Boiteux, A. Fogli, M. Muller-Barthelemy, C. Vaurs-Barriere, E. Chautard, B. Pereira, J. Biau, J. L. Kemeny, T. Khalil, L. Karayan-Tapon, P. Verrelle and P. Arnaud (2019). "Transcriptional alterations in glioma result primarily from DNA methylation-independent mechanisms." Genome Res 29(10): 1605-1621.

17. Domingues, H. S., C. C. Portugal, R. Socodato and J. B. Relvas (2016). "Corrigendum: Oligodendrocyte, Astrocyte and Microglia Crosstalk in Myelin Development, Damage, and Repair." Front Cell Dev Biol 4: 79.

18. Escartin, C., E. Galea, A. Lakatos, J. P. O'Callaghan, G. C. Petzold, A. Serrano-Pozo, C. Steinhauser, A. Volterra, G. Carmignoto, A. Agarwal, N. J. Allen, A. Araque, L. Barbeito, A. Barzilai, D. E. Bergles, G. Bonvento, A. M. Butt, W. T. Chen, M. Cohen-Salmon, C. Cunningham, B. Deneen, B. De Strooper, B. Diaz-Castro, C. Farina, M. Freeman, V. Gallo, J. E. Goldman, S. A. Goldman, M. Gotz, A. Gutierrez, P. G. Haydon, D. H. Heiland, E. M. Hol, M. G. Holt, M. Iino, K. V. Kastanenka, H. Kettenmann, B. S. Khakh, S. Koizumi, C. J. Lee, S. A. Liddelow, B. A. MacVicar, P. Magistretti, A. Messing, A. Mishra, A. V. Molofsky, K. K. Murai, C. M. Norris, S. Okada, S. H. R. Oliet, J. F. Oliveira, A. Panatier, V. Parpura, M. Pekna, M. Pekny, L. Pellerin, G. Perea, B. G. Perez-Nievas, F. W. Pfrieger, K. E. Poskanzer, F. J. Quintana, R. M. Ransohoff, M.

Riquelme-Perez, S. Robel, C. R. Rose, J. D. Rothstein, N. Rouach, D. H. Rowitch, A. Semyanov, S. Sirko, H. Sontheimer, R. A. Swanson, J. Vitorica, I. B. Wanner, L. B. Wood, J. Wu, B. Zheng, E. R. Zimmer, R. Zorec, M. V. Sofroniew and A. Verkhratsky (2021). "Reactive astrocyte nomenclature, definitions, and future directions." Nat Neurosci 24(3): 312-325.

19. Fan, Y., R. Mao and J. Yang (2013). "NF-kappaB and STAT3 signaling pathways collaboratively link inflammation to cancer." Protein Cell 4(3): 176-185.

20. Fields, J. and A. Ghorpade (2012). "C/EBPbeta regulates multiple IL-1beta-induced human astrocyte inflammatory genes." J Neuroinflammation 9: 177.

21. Fuente-Martin, E., C. Garcia-Caceres, E. Morselli, D. J. Clegg, J. A. Chowen, B. Finan, R. D. Brinton and M. H. Tschop (2013). "Estrogen, astrocytes and the neuroendocrine control of metabolism." Rev Endocr Metab Disord 14(4): 331-338.

22. Gabel, S., E. Koncina, G. Dorban, T. Heurtaux, C. Birck, E. Glaab, A. Michelucci, P. Heuschling and L. Grandbarbe (2016). "Inflammation Promotes a Conversion of Astrocytes into Neural Progenitor Cells via NF-kappaB Activation." Mol Neurobiol 53(8): 5041-5055.

23. Galvao, R. P. and H. Zong (2013). "Inflammation and Gliomagenesis: Bi-Directional Communication at Early and Late Stages of Tumor Progression." Curr Pathobiol Rep 1(1): 19-28.

24. Goodnight, A. V., I. Kremsky, S. Khampang, Y. H. Jung, J. M. Billingsley, S. E. Bosinger, V. G. Corces and A. W. S. Chan (2019). "Chromatin accessibility and transcription dynamics during in vitro astrocyte differentiation of Huntington's Disease Monkey pluripotent stem cells." Epigenetics Chromatin 12(1): 67.

25. Gotz, M., M. Nakafuku and D. Petrik (2016). "Neurogenesis in the Developing and Adult Brain-Similarities and Key Differences." Cold Spring Harb Perspect Biol 8(7).

26. Grandbarbe, L., J. Bouissac, M. Rand, M. Hrabe de Angelis, S. Artavanis-Tsakonas and E. Mohier (2003). "Delta-Notch signaling controls the generation of neurons/glia from neural stem cells in a stepwise process." Development 130(7): 1391-1402.

27. Grill, M., T. E. Syme, A. L. Nocon, A. Z. Lu, D. Hancock, S. Rose-John and I. L. Campbell (2015). "Strawberry notch homolog 2 is a novel inflammatory response factor predominantly but not exclusively expressed by astrocytes in the central nervous system." Glia 63(10): 1738-1752.

28. Grivennikov, S. I. and M. Karin (2010). "Dangerous liaisons: STAT3 and NF-kappaB collaboration and crosstalk in cancer." Cytokine Growth Factor Rev 21(1): 11-19.

29. Habas, A., J. Hahn, X. Wang and M. Margeta (2013). "Neuronal activity regulates astrocytic Nrf2 signaling." Proc Natl Acad Sci U S A 110(45): 18291-18296.

30. Hayes, J. D. and A. T. Dinkova-Kostova (2014). "The Nrf2 regulatory network provides an interface between redox and intermediary metabolism." Trends Biochem Sci 39(4): 199-218.

31. Herrera, F., Q. Chen and D. Schubert (2010). "Synergistic effect of retinoic acid and cytokines on the regulation of glial fibrillary acidic protein expression." J Biol Chem 285(50): 38915-38922.

32. Herrmann, J. E., T. Imura, B. Song, J. Qi, Y. Ao, T. K. Nguyen, R. A. Korsak, K. Takeda, S. Akira and M. V. Sofroniew (2008). "STAT3 is a critical regulator of astrogliosis and scar formation after spinal cord injury." J Neurosci 28(28): 7231-7243.

33. Hess, D. C. and C. V. Borlongan (2008). "Stem cells and neurological diseases." Cell Prolif 41 Suppl 1: 94-114.

34. Hirabayashi, Y. and Y. Gotoh (2005). "Stage-dependent fate determination of neural precursor cells in mouse forebrain." Neurosci Res 51(4): 331-336.

35. Hong, S. and M. R. Song (2014). "STAT3 but not STAT1 is required for astrocyte differentiation." PLoS One 9(1): e86851.

36. Hunt, D., G. Raivich and P. N. Anderson (2012). "Activating transcription factor 3 and the nervous system." Front Mol Neurosci 5: 7.

37. Huynh-Thu, V. A., A. Irrthum, L. Wehenkel and P. Geurts (2010). "Inferring regulatory networks from expression data using tree-based methods." PLoS One 5(9).

38. Hwang, D. W., A. Jaganathan, P. Shrestha, Y. Jin, N. El-Amine, S. H. Wang, M. Hammell and A. A. Mills (2018). "Chromatin-mediated translational control is essential for neural cell fate specification." Life Sci Alliance 1(4): e201700016.

39. Idriss, H. T. and J. H. Naismith (2000). "TNF alpha and the TNF receptor superfamily: structure-function relationship(s)." Microsc Res Tech 50(3): 184-195.

40. Iglesias, J., L. Morales and G. E. Barreto (2017). "Metabolic and Inflammatory Adaptation of Reactive Astrocytes: Role of PPARs." Mol Neurobiol 54(4): 2518-2538.

41. Imayoshi, I., M. Sakamoto, M. Yamaguchi, K. Mori and R. Kageyama (2010). "Essential roles of Notch signaling in maintenance of neural stem cells in developing and adult brains." J Neurosci 30(9): 3489-3498.

42. Lattke, M., R. Goldstone, J. K. Ellis, S. Boeing, J. Jurado-Arjona, N. Marichal, J. I. MacRae, B. Berninger and F. Guillemot (2021). "Extensive transcriptional and chromatin changes underlie astrocyte maturation in vivo and in culture." Nat Commun 12(1): 4335.

43. Lattke, M., S. N. Reichel and B. Baumann (2017). "NF-kappaB-mediated astrocyte dysfunction initiates neurodegeneration." Oncotarget 8(31): 50329-50330.

44. Lawrence, T. (2009). "The nuclear factor NF-kappaB pathway in inflammation." Cold Spring Harb Perspect Biol 1(6): a001651.

45. Lee, S. J., K. Drabik, N. J. Van Wagoner, S. Lee, C. Choi, Y. Dong and E. N. Benveniste (2000). "ICAM-1-induced expression of proinflammatory cytokines in astrocytes: involvement of extracellular signal-regulated kinase and p38 mitogen-activated protein kinase pathways." J Immunol 165(8): 4658-4666.

46. Li, Z., M. H. Schulz, T. Look, M. Begemann, M. Zenke and I. G. Costa (2019). "Identification of transcription factor binding sites using ATAC-seq." Genome Biol 20(1): 45.

47. Liddell, J. R. (2017). "Are Astrocytes the Predominant Cell Type for Activation of Nrf2 in Aging and Neurodegeneration?" Antioxidants (Basel) 6(3).

48. Liddelow, S. A., K. A. Guttenplan, L. E. Clarke, F. C. Bennett, C. J. Bohlen, L. Schirmer, M. L. Bennett, A. E. Munch, W. S. Chung, T. C. Peterson, D. K. Wilton, A. Frouin, B. A. Napier, N. Panicker, M. Kumar, M. S. Buckwalter, D. H. Rowitch, V. L. Dawson, T. M. Dawson, B. Stevens and B. A. Barres (2017). "Neurotoxic reactive astrocytes are induced by activated microglia." Nature 541(7638): 481-487.

49. Lin, M. S., Y. Y. Sun, W. T. Chiu, C. C. Hung, C. Y. Chang, F. S. Shie, S. H. Tsai, J. W. Lin, K. S. Hung and Y. H. Lee (2011). "Curcumin attenuates the expression and secretion of RANTES after spinal cord injury in vivo and lipopolysaccharide-induced astrocyte reactivation in vitro." J Neurotrauma 28(7): 1259-1269.

50. Logan, T. T., M. Rusnak and A. J. Symes (2015). "Runx1 promotes proliferation and neuronal differentiation in adult mouse neurosphere cultures." Stem Cell Res 15(3): 554-564.

51. Lovatt, D., U. Sonnewald, H. S. Waagepetersen, A. Schousboe, W. He, J. H. Lin, X. Han, T. Takano, S. Wang, F. J. Sim, S. A. Goldman and M. Nedergaard (2007). "The transcriptome and metabolic gene signature of protoplasmic astrocytes in the adult murine cortex." J Neurosci 27(45): 12255-12266.

52. Lozano, D., G. S. Gonzales-Portillo, S. Acosta, I. de la Pena, N. Tajiri, Y. Kaneko and C. V. Borlongan (2015). "Neuroinflammatory responses to traumatic brain injury: etiology, clinical consequences, and therapeutic opportunities." Neuropsychiatr Dis Treat 11: 97-106.

53. Magistri, M., N. Khoury, E. M. Mazza, D. Velmeshev, J. K. Lee, S. Bicciato, P. Tsoulfas and M. A. Faghihi (2016). "A comparative transcriptomic analysis of astrocytes differentiation from human neural progenitor cells." Eur J Neurosci 44(10): 2858-2870.

54. Messing, A., M. Brenner, M. B. Feany, M. Nedergaard and J. E. Goldman (2012). "Alexander disease." J Neurosci 32(15): 5017-5023.

55. Michelucci, A., A. Bithell, M. J. Burney, C. E. Johnston, K. Y. Wong, S. W. Teng, J. Desai, N. Gumbleton, G. Anderson, L. W. Stanton, B. P. Williams and N. J. Buckley (2016). "The Neurogenic Potential of Astrocytes Is Regulated by Inflammatory Signals." Mol Neurobiol 53(6): 3724-3739.

56. Molofsky, A. V., R. Krencik, E. M. Ullian, H. H. Tsai, B. Deneen, W. D. Richardson, B. A. Barres and D. H. Rowitch (2012). "Astrocytes and disease: a neurodevelopmental perspective." Genes Dev 26(9): 891-907.

57. Neal, M., J. Luo, D. S. Harischandra, R. Gordon, S. Sarkar, H. Jin, V. Anantharam, L. Desaubry, A. Kanthasamy and A. Kanthasamy (2018). "Prokineticin-2 promotes chemotaxis and alternative A2 reactivity of astrocytes." Glia 66(10): 2137-2157.

58. Park, S. H., K. Kang, E. Giannopoulou, Y. Qiao, K. Kang, G. Kim, K. H. Park-Min and L. B. Ivashkiv (2017). "Type I interferons and the cytokine TNF cooperatively reprogram the macrophage epigenome to promote inflammatory activation." Nat Immunol 18(10): 1104-1116.

59. Parratt, J. D. and J. W. Prineas (2010). "Neuromyelitis optica: a demyelinating disease characterized by acute destruction and regeneration of perivascular astrocytes." Mult Scler 16(10): 1156-1172.

60. Pavlou, M. A. S., L. Grandbarbe, N. J. Buckley, S. P. Niclou and A. Michelucci (2019). "Transcriptional and epigenetic mechanisms underlying astrocyte identity." Prog Neurobiol 174: 36-52.

61. Pearson, C. A., D. M. Moore, H. O. Tucker, J. D. Dekker, H. Hu, A. Miquelajauregui and B. G. Novitch (2020). "Foxp1 Regulates Neural Stem Cell Self-Renewal and Bias Toward Deep Layer Cortical Fates." Cell Rep 30(6): 1964-1981 e1963.

62. Pekny, M., M. Pekna, A. Messing, C. Steinhauser, J. M. Lee, V. Parpura, E. M. Hol, M. V. Sofroniew and A. Verkhratsky (2016). "Astrocytes: a central element in neurological diseases." Acta Neuropathol 131(3): 323-345.

63. Platanitis, E. and T. Decker (2018). "Regulatory Networks Involving STATs, IRFs, and NFkappaB in Inflammation." Front Immunol 9: 2542.

64. Quinlan, R. A., M. Brenner, J. E. Goldman and A. Messing (2007). "GFAP and its role in Alexander disease." Exp Cell Res 313(10): 2077-2087.

65. Rothhammer, V. and F. J. Quintana (2015). "Control of autoimmune CNS inflammation by astrocytes." Semin Immunopathol 37(6): 625-638.

66. Sanai, N., A. Alvarez-Buylla and M. S. Berger (2005). "Neural stem cells and the origin of gliomas." N Engl J Med 353(8): 811-822.

67. Santopolo, G., J. P. Magnusson, O. Lindvall, Z. Kokaia and J. Frisen (2020). "Blocking Notch-Signaling Increases Neurogenesis in the Striatum after Stroke." Cells 9(7).

68. Sloan, S. A. and B. A. Barres (2014). "Mechanisms of astrocyte development and their contributions to neurodevelopmental disorders." Curr Opin Neurobiol 27: 75-81.

69. Sofroniew, M. V. and H. V. Vinters (2010). "Astrocytes: biology and pathology." Acta Neuropathol 119(1): 7-35.

70. Spence, R. D., M. E. Hamby, E. Umeda, N. Itoh, S. Du, A. J. Wisdom, Y. Cao, G. Bondar, J. Lam, Y. Ao, F. Sandoval, S. Suriany, M. V. Sofroniew and R. R. Voskuhl (2011). "Neuroprotection mediated through estrogen receptor-alpha in astrocytes." Proc Natl Acad Sci U S A 108(21): 8867-8872.

71. Stipursky, J. and F. C. Gomes (2007). "TGF-beta1/SMAD signaling induces astrocyte fate commitment in vitro: implications for radial glia development." Glia 55(10): 1023-1033.

72. Sun, M. and P. J. Fink (2007). "A new class of reverse signaling costimulators belongs to the TNF family." J Immunol 179(7): 4307-4312.

73. Takouda, J., S. Katada and K. Nakashima (2017). "Emerging mechanisms underlying astrogenesis in the developing mammalian brain." Proc Jpn Acad Ser B Phys Biol Sci 93(6): 386-398.

74. Tchieu, J., E. L. Calder, S. R. Guttikonda, E. M. Gutzwiller, K. A. Aromolaran, J. A. Steinbeck, P. A. Goldstein and L. Studer (2019). "NFIA is a gliogenic switch enabling rapid derivation of functional human astrocytes from pluripotent stem cells." Nat Biotechnol 37(3): 267-275.

75. Tertil, M., U. Skupio, J. Barut, V. Dubovyk, A. Wawrzczak-Bargiela, Z. Soltys, S. Golda, L. Kudla, L. Wiktorowska, K. Szklarczyk, M. Korostynski, R. Przewlocki and M. Slezak (2018). "Glucocorticoid receptor signaling in astrocytes is required for aversive memory formation." Transl Psychiatry 8(1): 255.

76. Tiwari, N., A. Pataskar, S. Peron, S. Thakurela, S. K. Sahu, M. Figueres-Onate, N. Marichal, L. Lopez-Mascaraque, V. K. Tiwari and B. Berninger (2018). "Stage-Specific Transcription Factors Drive Astrogliogenesis by Remodeling Gene Regulatory Landscapes." Cell Stem Cell 23(4): 557-571 e558.

77. Trentin, A. G. (2006). "Thyroid hormone and astrocyte morphogenesis." J Endocrinol 189(2): 189-197.

78. Wen, S., H. Li and J. Liu (2009). "Dynamic signaling for neural stem cell fate determination." Cell Adh Migr 3(1): 107-117.

79. Wu, X., T. Kihara, A. Akaike, T. Niidome and H. Sugimoto (2010). "PI3K/Akt/mTOR signaling regulates glutamate transporter 1 in astrocytes." Biochem Biophys Res Commun 393(3): 514-518.

80. Yun, S. J., K. Byun, J. Bhin, J. H. Oh, T. H. Nhung le, D. Hwang and B. Lee (2010). "Transcriptional regulatory networks associated with self-renewal and differentiation of neural stem cells." J Cell Physiol 225(2): 337-347.

81. Zamanian, J. L., L. Xu, L. C. Foo, N. Nouri, L. Zhou, R. G. Giffard and B. A. Barres (2012). "Genomic analysis of reactive astrogliosis." J Neurosci 32(18): 6391-6410.

82. Zhang, J., X. Xu, H. Liu, L. Jin, X. Shen, C. Xie, W. Xiang, D. Yang, W. Feng, J. Wang, M. Wang, T. Dong, H. Qiu, L. Wu, Y. Wang, X. Zhang and Z. Huang (2021). "Astrocytic YAP prevents the demyelination through promoting expression of cholesterol synthesis genes in experimental autoimmune encephalomyelitis." Cell Death Dis 12(10): 907.

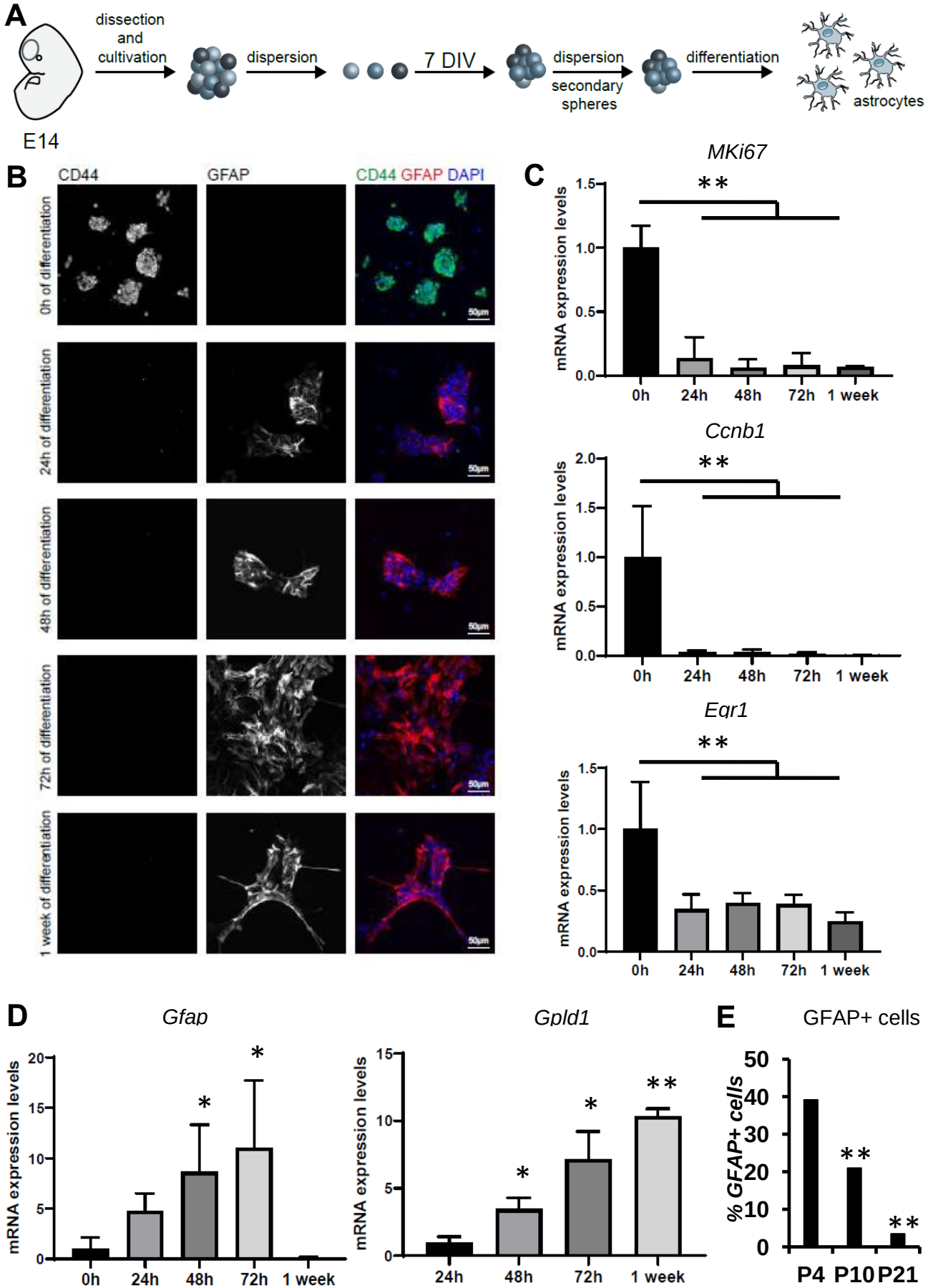

A
dissection and cultivation
dispersion
7 DIV
dispersion
secondary spheres
differentiation
astrocytes
E14
B
CD44
GFAP
CD44 GFAP DAPI
0h of differentiation
24h of differentiation
48h of differentiation
72h of differentiation
1 week of differentiation
50µm
C
MKi67
Ccnb1
Egr1
mRNA expression levels
0h
24h
48h
72h
1 week
D
Gfap
Gpld1
E
GFAP+ cells
% GFAP+ cells
P4
P10
P21

Figure 1. *In vitro* differentiation of neurospheres (NSPs) into astrocytes. (A) Schematic showing the *in vitro* cellular model of primary neurospheres (NSPs) derived from NSCs isolated from the ventricular zone at embryonic day-14 (E14) of the mouse embryos and the subsequent passages to differentiate them into astrocytes. **(B)** Immunostaining analyses of CD44 (green) and GFAP (red) of NSPs and differentiating astrocytes obtained from primary murine cultures of NSPs cultivated under proliferation conditions (DMEM F12, 1% B27 (- vitamin A), antibiotics and 20 ng/ml EGF) or differentiated into astrocytes (DMEM, 10% FBS and antibiotics) for 24h, 48h, 72h and 1 week. Nuclei were counterstained with DAPI (in blue). Scale bar: 50 μm. **(C-D)** RT-qPCR analyses showing mRNA expression levels of **(C)** the cell cycle genes *Mki67* and *Ccnb1*, the NSC marker *Egr1*, **(D)** the astrocytic marker *Gfap* and the astrocyte-enriched gene *Gpld1*. Results are expressed as mean ± standard deviation. *$p < 0.05$; **$p < 0.01$, $n \geq 4$. **(E)** Quantification of GFAP+ cells from mouse forebrain cell suspensions of GFAP-GFP transgenic mice by FACS at postnatal day 4 (P4), P10 and P21. Bars represent the mean ± SEM. **$p < 0.01$, $n = 3$.

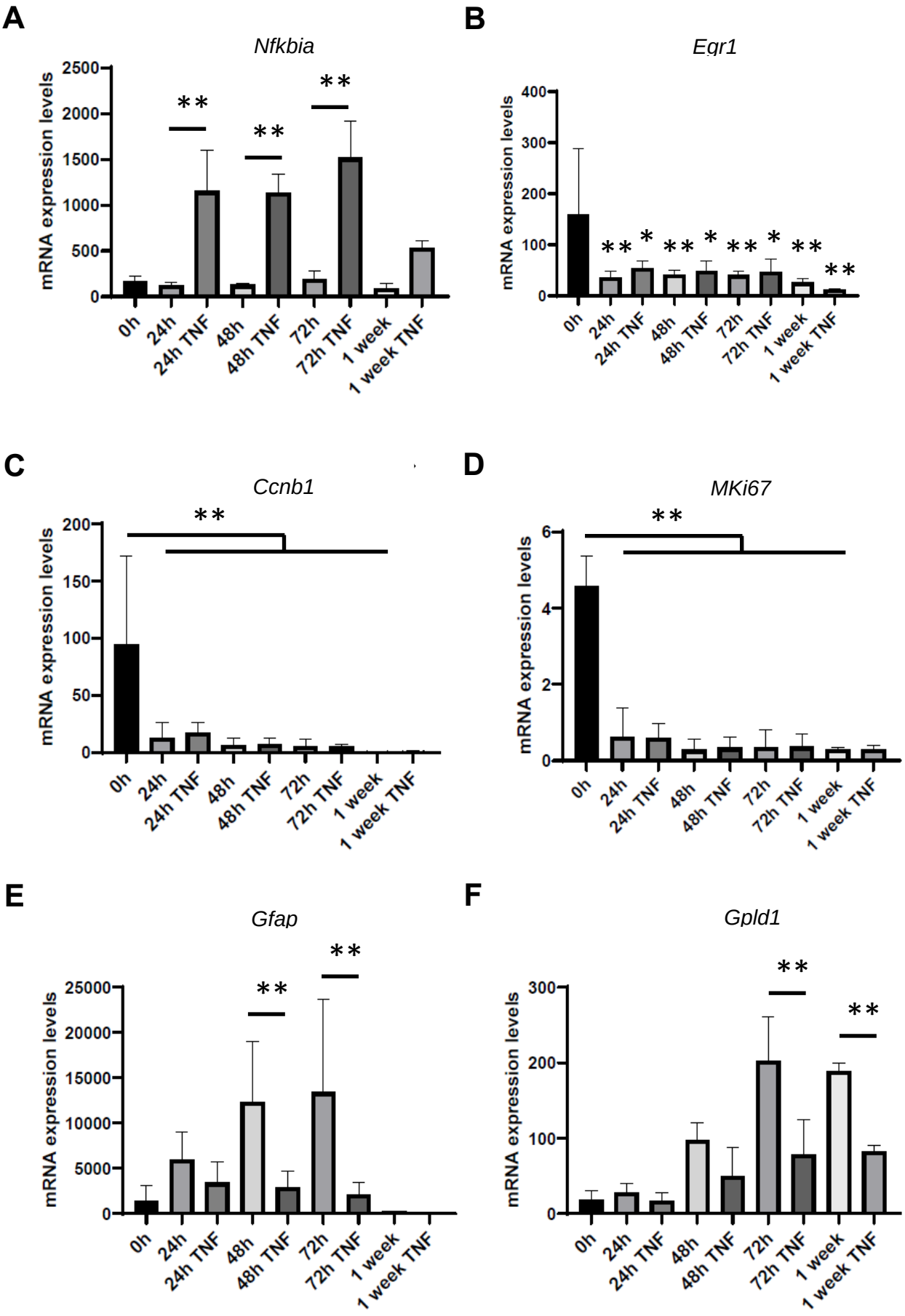

A
Nfkbia
mRNA expression levels
2500
2000
1500
1000
500
0
**
**
**
0h
24h
24h TNF
48h
48h TNF
72h
72h TNF
1 week
1 week TNF
B
Egr1
mRNA expression levels
400
300
200
100
0
**
*
**
*
**
*
**
**
C
Ccnb1
mRNA expression levels
200
150
100
50
0
**
D
MKi67
mRNA expression levels
6
4
2
0
**
E
Gfap
mRNA expression levels
25000
20000
15000
10000
5000
0
**
**
F
Gpld1
mRNA expression levels
300
200
100
0
**
**

Figure 2. Gene expression levels of NSC and astrocyte markers following treatment of NSCs with TNF upon astrocytic differentiation. NSPs were differentiated into astrocytes with or without exposure to TNF (50 ng/ml) for 24 h, 48 h, 72 h and 1 week. RT-qPCR analyses show the mRNA expression levels of **(A)** *Nfkbia*, **(B)** *Egr1*, **(C)** *Ccnb1*, **(D)** *Mki67,* **(E)** *Gfap* and **(F)** *Gpld1*. Results are expressed as mean ± standard deviation. *$p < 0.05$; **$p < 0.01$, $n \geq 4$.

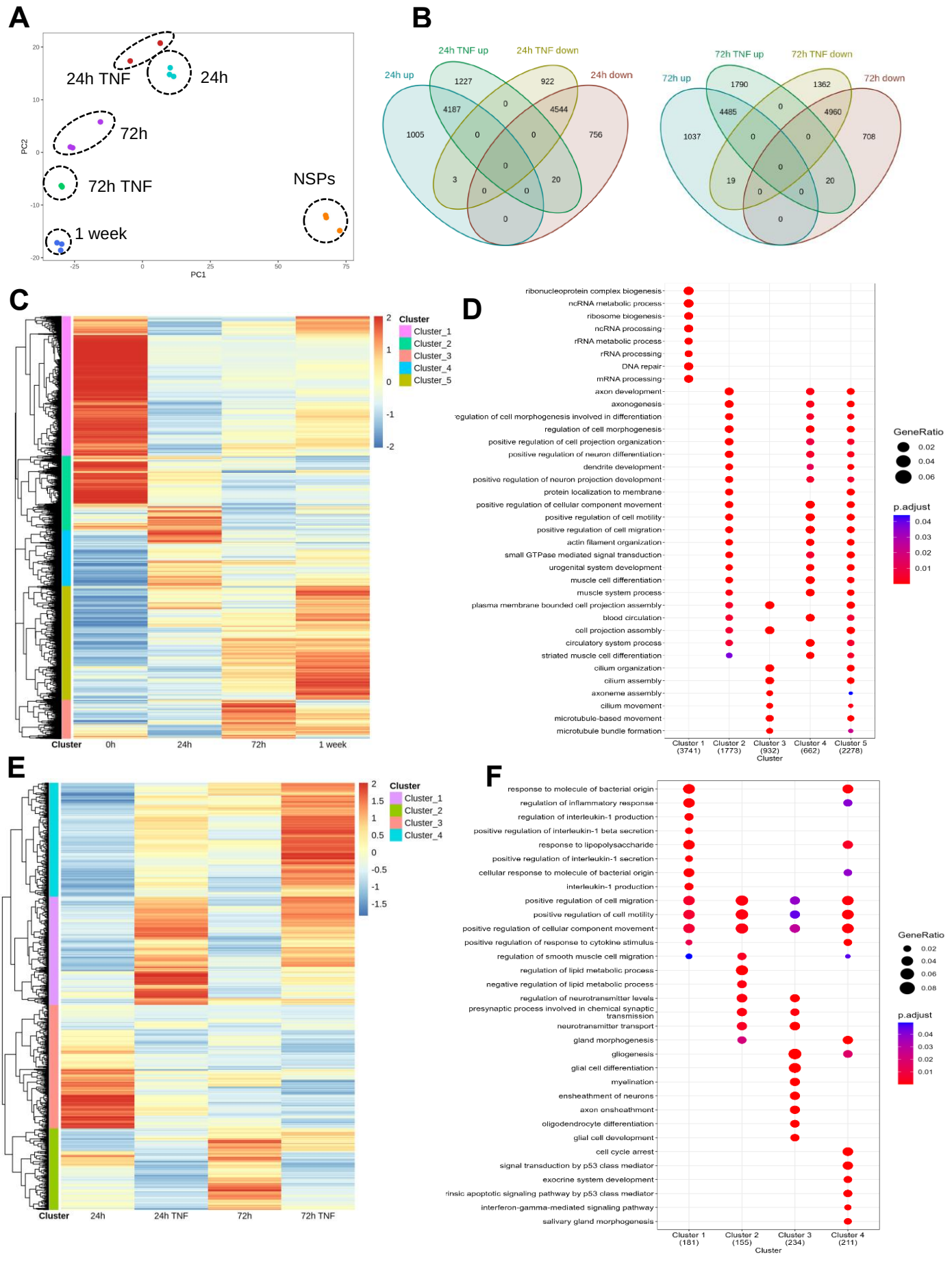

A
24h TNF
24h
72h
72h TNF
NSPs
1 week
PC1
PC2
B
24h up
24h TNF up
24h TNF down
24h down
1005
1227
922
756
4187
4544
3
20
72h up
72h TNF up
72h TNF down
72h down
1037
1790
1362
708
4485
4960
19
20
C
Cluster
Cluster_1
Cluster_2
Cluster_3
Cluster_4
Cluster_5
0h
24h
72h
1 week
D
ribonucleoprotein complex biogenesis
ncRNA metabolic process
ribosome biogenesis
ncRNA processing
rRNA metabolic process
rRNA processing
DNA repair
mRNA processing
axon development
axonogenesis
regulation of cell morphogenesis involved in differentiation
regulation of cell morphogenesis
positive regulation of cell projection organization
positive regulation of neuron differentiation
dendrite development
positive regulation of neuron projection development
protein localization to membrane
positive regulation of cellular component movement
positive regulation of cell motility
positive regulation of cell migration
actin filament organization
small GTPase mediated signal transduction
urogenital system development
muscle cell differentiation
muscle system process
plasma membrane bounded cell projection assembly
blood circulation
cell projection assembly
circulatory system process
striated muscle cell differentiation
cilium organization
cilium assembly
axoneme assembly
cilium movement
microtubule-based movement
microtubule bundle formation
Cluster 1 (3741)
Cluster 2 (1773)
Cluster 3 (932)
Cluster 4 (662)
Cluster 5 (2278)
Cluster
GeneRatio
p.adjust
E
Cluster
Cluster_1
Cluster_2
Cluster_3
Cluster_4
24h
24h TNF
72h
72h TNF
F
response to molecule of bacterial origin
regulation of inflammatory response
regulation of interleukin-1 production
positive regulation of interleukin-1 beta secretion
response to lipopolysaccharide
positive regulation of interleukin-1 secretion
cellular response to molecule of bacterial origin
interleukin-1 production
positive regulation of cell migration
positive regulation of cell motility
positive regulation of cellular component movement
positive regulation of response to cytokine stimulus
regulation of smooth muscle cell migration
regulation of lipid metabolic process
negative regulation of lipid metabolic process
regulation of neurotransmitter levels
presynaptic process involved in chemical synaptic transmission
neurotransmitter transport
gland morphogenesis
gliogenesis
glial cell differentiation
myelination
ensheathment of neurons
axon ensheathment
oligodendrocyte differentiation
glial cell development
cell cycle arrest
signal transduction by p53 class mediator
exocrine system development
rinsic apoptotic signaling pathway by p53 class mediator
interferon-gamma-mediated signaling pathway
salivary gland morphogenesis
Cluster 1 (181)
Cluster 2 (155)
Cluster 3 (234)
Cluster 4 (211)
Cluster
GeneRatio
p.adjust

Figure 3. Transcriptomic signature of differentiating astrocytes under nomal and inflammatory conditions. (A) Representation of first two principal components (PC) analysis of the top 1'000 most variable genes across all the conditions analyzed by RNA-seq. **(B)** Venn diagrams showing the number of significantly differentially expressed genes ($p < 0.05$) and their overlap across comparisons 24h TNF vs 24 h (left) and 72h TNF vs 72h (right). **(C)** Heatmap showing normalized (z scores) raw gene expression counts of differentially expressed genes under normal conditions based on absolute log2 fold change from comparisons between the different time points (1 week vs 72h, 72h vs 24h and 24h vs 0h). **(D)** Dot plot showing gene set enrichment analysis of genes in clusters identified in the heatmap (**Fig. 3C**). **(E)** Heatmap showing normalized (z scores) raw gene expression counts under inflammatory conditions (24h TNF vs 24h and 72h TNF vs 72h) based on absolute log2 fold change. **(F)** Dot plot showing gene set enrichment analysis of genes in clusters identified in the heatmap (**Fig. 3E**).

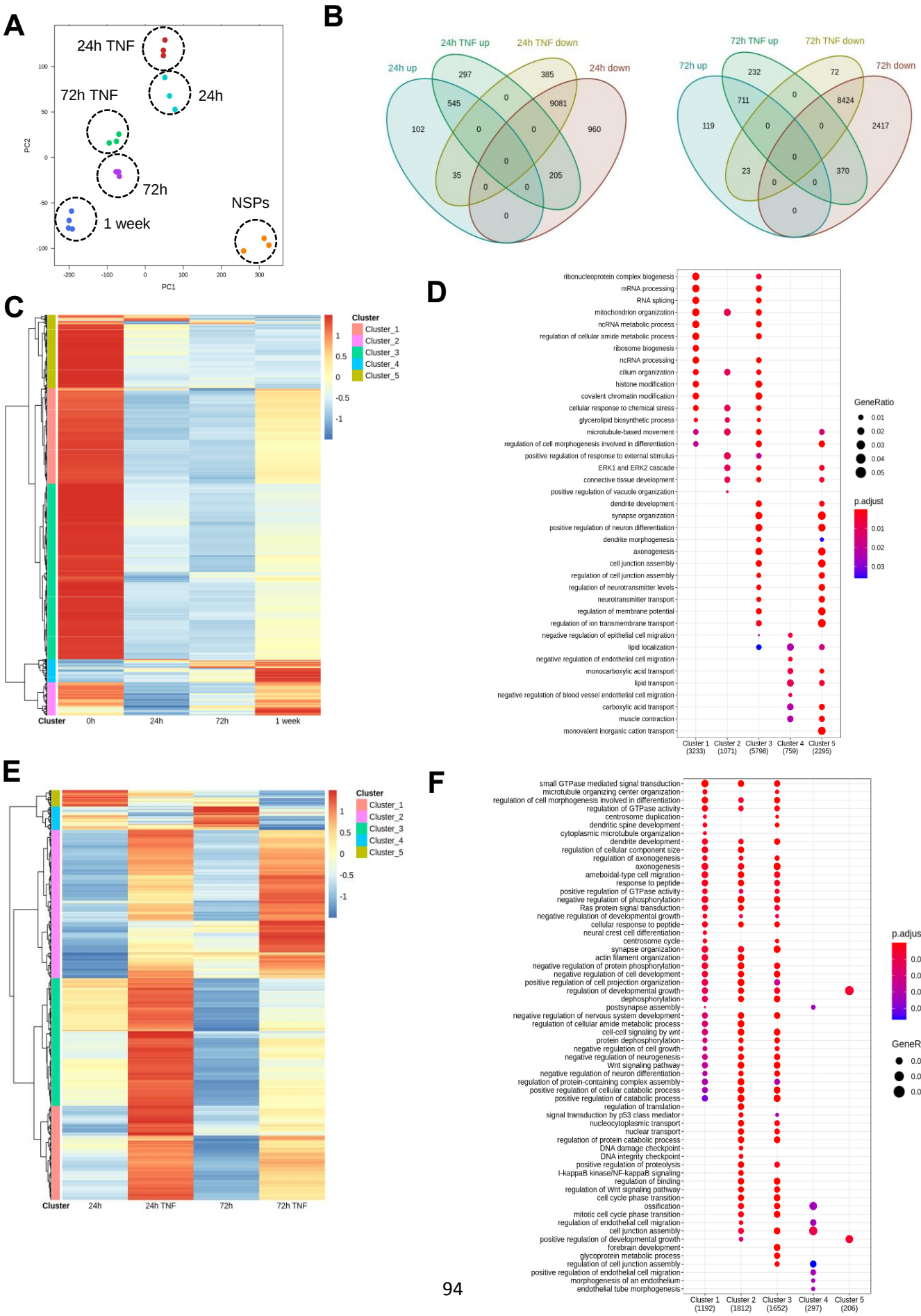

A
24h TNF
72h TNF
24h
72h
1 week
NSPs
PC1
PC2
B
24h up
24h TNF up
24h TNF down
24h down
72h up
72h TNF up
72h TNF down
72h down
C
Cluster
Cluster_1
Cluster_2
Cluster_3
Cluster_4
Cluster_5
Cluster 0h 24h 72h 1 week
D
ribonucleoprotein complex biogenesis
mRNA processing
RNA splicing
mitochondrion organization
ncRNA metabolic process
regulation of cellular amide metabolic process
ribosome biogenesis
ncRNA processing
cilium organization
histone modification
covalent chromatin modification
cellular response to chemical stress
glycerolipid biosynthetic process
microtubule-based movement
regulation of cell morphogenesis involved in differentiation
positive regulation of response to external stimulus
ERK1 and ERK2 cascade
connective tissue development
positive regulation of vacuole organization
dendrite development
synapse organization
positive regulation of neuron differentiation
dendrite morphogenesis
axonogenesis
cell junction assembly
regulation of cell junction assembly
regulation of neurotransmitter levels
neurotransmitter transport
regulation of membrane potential
regulation of ion transmembrane transport
negative regulation of epithelial cell migration
lipid localization
negative regulation of endothelial cell migration
monocarboxylic acid transport
lipid transport
negative regulation of blood vessel endothelial cell migration
carboxylic acid transport
muscle contraction
monovalent inorganic cation transport
Cluster 1 (3233)
Cluster 2 (1071)
Cluster 3 (5798)
Cluster 4 (759)
Cluster 5 (2295)
GeneRatio
p.adjust
E
Cluster
Cluster_1
Cluster_2
Cluster_3
Cluster_4
Cluster_5
Cluster 24h 24h TNF 72h 72h TNF
F
small GTPase mediated signal transduction
microtubule organizing center organization
regulation of cell morphogenesis involved in differentiation
regulation of GTPase activity
centrosome duplication
dendritic spine development
cytoplasmic microtubule organization
dendrite development
regulation of cellular component size
regulation of axonogenesis
axonogenesis
ameboidal-type cell migration
response to peptide
positive regulation of GTPase activity
negative regulation of phosphorylation
Ras protein signal transduction
negative regulation of developmental growth
cellular response to peptide
neural crest cell differentiation
centrosome cycle
synapse organization
actin filament organization
negative regulation of protein phosphorylation
negative regulation of cell development
positive regulation of cell projection organization
regulation of developmental growth
dephosphorylation
postsynapse assembly
negative regulation of nervous system development
regulation of cellular amide metabolic process
cell-cell signaling by wnt
protein dephosphorylation
negative regulation of cell growth
negative regulation of neurogenesis
Wnt signaling pathway
negative regulation of neuron differentiation
regulation of protein-containing complex assembly
positive regulation of cellular catabolic process
positive regulation of catabolic process
regulation of translation
signal transduction by p53 class mediator
nucleocytoplasmic transport
nuclear transport
regulation of protein catabolic process
DNA damage checkpoint
DNA integrity checkpoint
positive regulation of proteolysis
I-kappaB kinase/NF-kappaB signaling
regulation of binding
regulation of Wnt signaling pathway
cell cycle phase transition
ossification
mitotic cell cycle phase transition
regulation of endothelial cell migration
cell junction assembly
positive regulation of developmental growth
forebrain development
glycoprotein metabolic process
regulation of cell junction assembly
positive regulation of endothelial cell migration
morphogenesis of an endothelium
endothelial tube morphogenesis
Cluster 1 (1192)
Cluster 2 (1812)
Cluster 3 (1652)
Cluster 4 (297)
Cluster 5 (206)
p.adjust
GeneRatio

Figure 4. Chromatin accessibility profile of differentiating astrocytes under normal and inflammatory conditions. (A) Representation of first two principal components (PC) analysis of the top 1'000 most variable genes across all the conditions analyzed by ATAC-seq. **(B)** Venn diagrams showing the number of significantly differentially accessible peaks ($p < 0.05$) and their overlap across comparisons 24h TNF vs 24h (left) and 72h TNF vs 72h (right). **(C)** Heatmap showing normalized (z scores) raw expression of differential peaks under normal conditions based on absolute log2 fold change from comparisons between the different time points (1 week vs 72h, 72h vs 24h and 24h vs 0h). **(D)** Dot plot showing gene set enrichment analysis of genes in clusters identified in the heatmap (**Fig. 4C**). **(E)** Heatmap showing normalized (z scores) raw expression peak counts under inflammatory conditions (2h TNF vs 24h and 72h TNF vs 72h) based on absolute log2 fold change. **(F)** Dot plot showing gene set enrichment analysis of genes in clusters identified in the heatmap **(Fig. 4E)**.

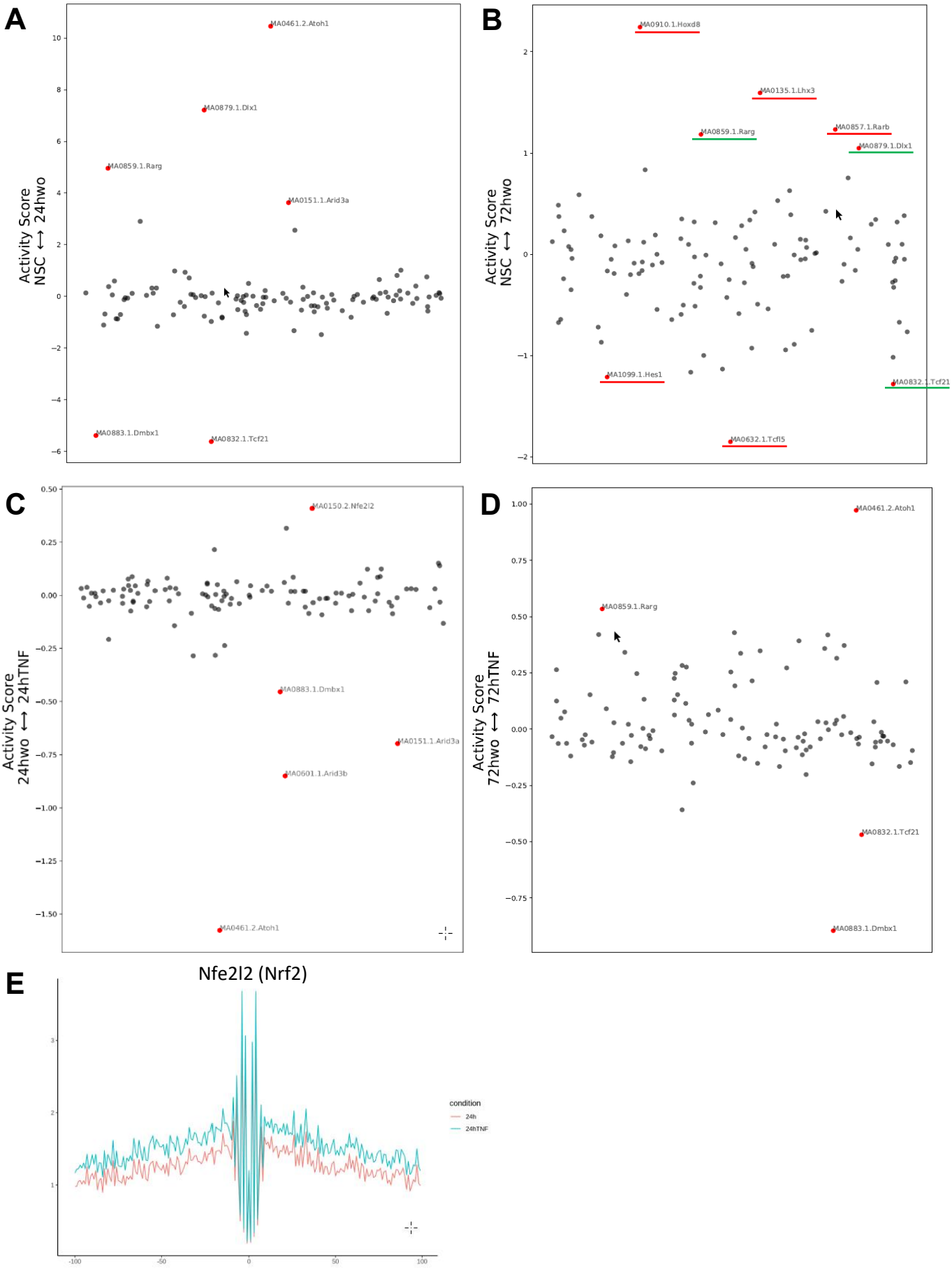

A
MA0461.2.Atoh1
MA0879.1.Dlx1
MA0859.1.Rarg
MA0151.1.Arid3a
MA0883.1.Dmbx1
MA0832.1.Tcf21
Activity Score
NSC ↔ 24hwo
B
MA0910.1.Hoxd8
MA0135.1.Lhx3
MA0859.1.Rarg
MA0857.1.Rarb
MA0879.1.Dlx1
MA1099.1.Hes1
MA0832.1.Tcf21
MA0632.1.Tcfl5
Activity Score
NSC ↔ 72hwo
C
MA0150.2.Nfe2l2
MA0883.1.Dmbx1
MA0151.1.Arid3a
MA0601.1.Arid3b
MA0461.2.Atoh1
Activity Score
24hwo ↔ 24hTNF
D
MA0461.2.Atoh1
MA0859.1.Rarg
MA0832.1.Tcf21
MA0883.1.Dmbx1
Activity Score
72hwo ↔ 72hTNF
E
Nfe2l2 (Nrf2)
condition
24h
24hTNF

Figure 5. Chromatin footprinting activity under normal and inflammatory conditions. (A-D) Activity score plots for different time point under normal and inflammatory conditions. **(E)** Comparison of the activity of *Nfe2l2 (Nrf2)* gene based on motif enrichment analysis at 24h between normal (red line) and inflammatory (blue line) conditions.

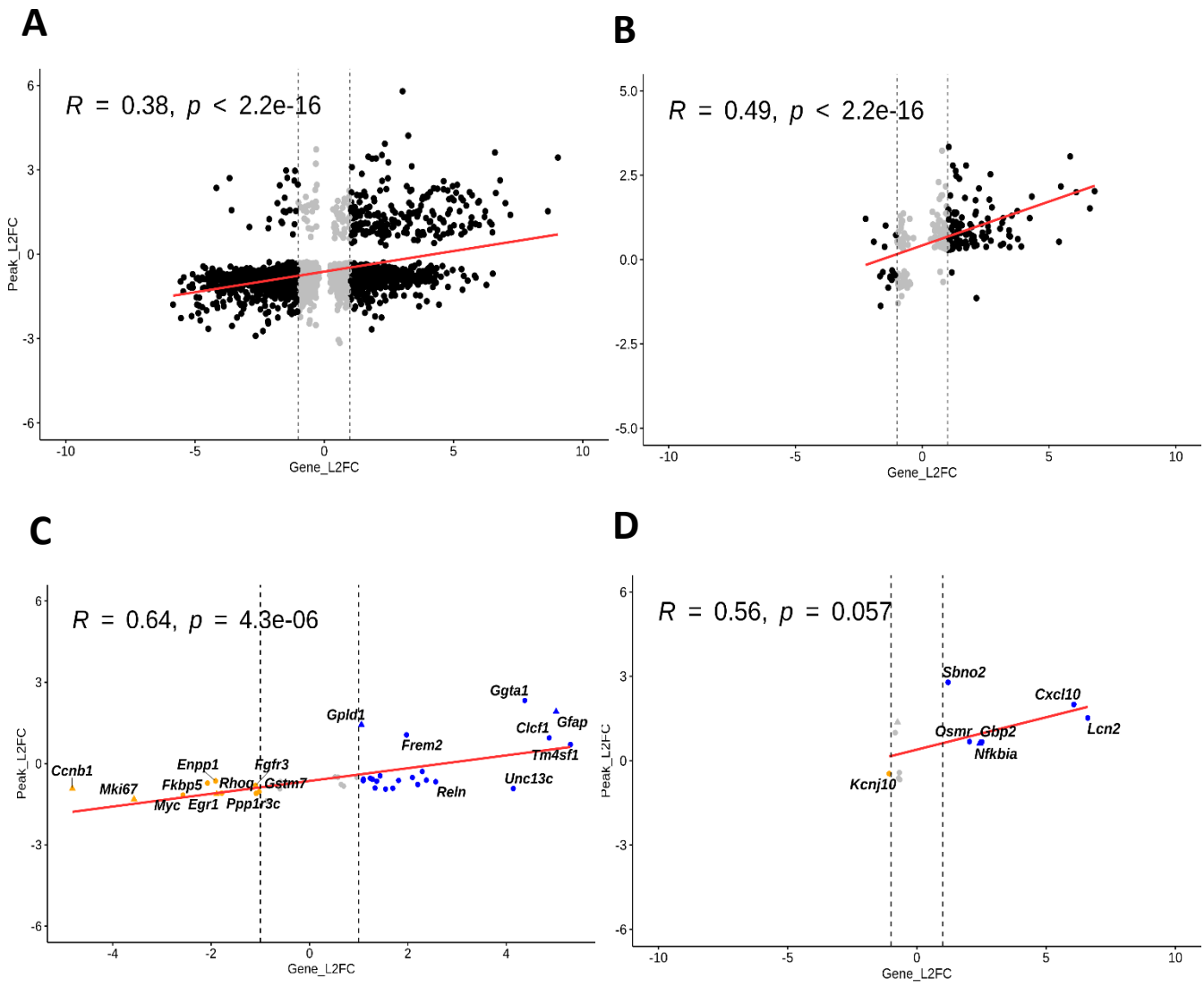


Figure 6. Correlation analyses of transcriptional and chromatin accessibility changes upon the differentiation process under normal and inflammatory conditions. Scatter plots for combined RNA-seq and ATAC-seq data **(A, B)** at the genome-wide level and **(C, D)** at specific NSC, resting and reactive astrocytic gene loci at 24h under **(A, C)** normal and **(B, D)** inflammatory conditions. Plots display Pearson's correlation coefficient indicating correlation levels between RNA-seq and ATAC-seq data, along with corresponding p-values (considering differentially

expressed genes with adjusted p value < 0.05 and differentially accessible peaks with false discovery rate (FDR) < 0.05). Blue and orange dots represent marker genes, which are significantly differentially expressed (adjusted p value < 0.05), have more than 2 fold positive change (upregulation) or more than 2 fold negative change (downregulation) respectively in gene expression and have a significantly differentially accessible peak (FDR < 0.05) in their promoter region. Red lines represent the regression line of correlation between fold change in gene expression and accessibility.

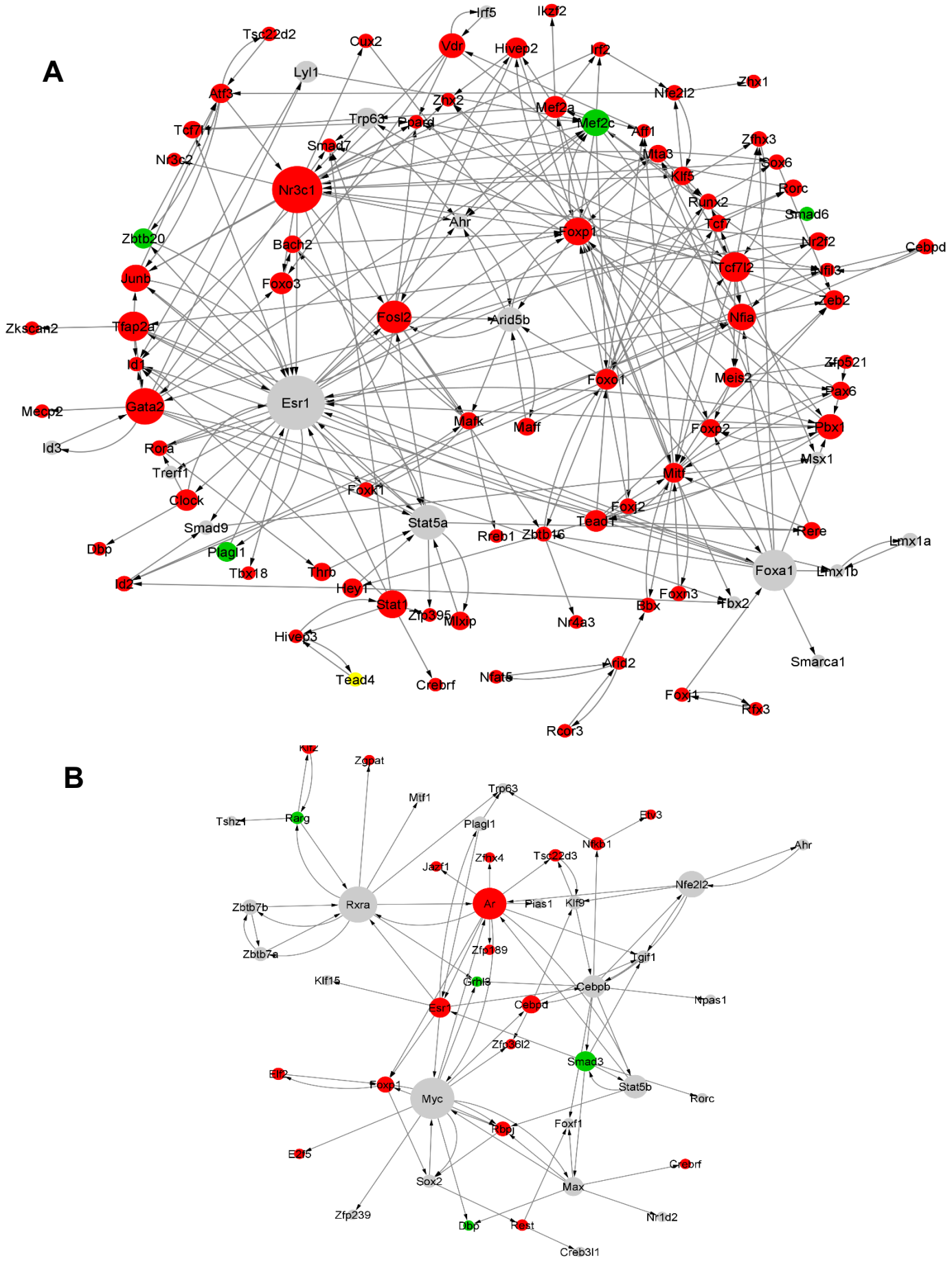

A
Irf5
Ikzf2
Tsc22d2
Cux2
Vdr
Hivep2
Irf2
Lyl1
Atf3
Nfe2l2
Zhx1
Zhx2
Mef2a
Trp63
Pparg
Mef2c
Tcf7l1
Aff1
Smad7
Zfhx3
Nr3c2
Mta3
Sox6
Nr3c1
Klf5
Rorc
Runx2
Ahr
Smad6
Tcf7
Zbtb20
Foxp1
Nr2f2
Bach2
Cebpd
Tcf7l2
Junb
Foxo3
Ifil3
Zeb2
Fosl2
Arid5b
Nfia
Zkscan2
Tfap2a
Id1
Zfp521
Foxo1
Meis2
Esr1
Pax6
Mecp2
Gata2
Mafk
Maff
Foxp2
Pbx1
Id3
Rora
Msx1
Trerf1
Mitf
Foxk1
Clock
Foxi2
Smad9
Stat5a
Tead1
Rere
Rreb1
Zbtb16
Dbp
Lmx1a
Plagl1
Foxa1
Tbx18
Thrb
Lmx1b
Id2
Hey1
Foxn3
Stat1
Bbx
Tbx2
Zfp395
Mlxip
Nr4a3
Hivep1
Smarca1
Arid2
Tead4
Crebrf
Nfat5
Foxj1
Rfx3
Rcor3
B
Klf2
Zgpat
Mtf1
Trp63
Tshz1
Rarg
Plagl1
Etv3
Nfkb1
Ahr
Jazf1
Zfp704
Tsc22d3
Nfe2l2
Rxra
Ar
Pias1
Klf9
Zbtb7b
Zbtb7a
Zfp189
Tgif1
Klf15
Grhl3
Cebpb
Npas1
Esr1
Cebpd
Zfp36l2
Smad3
Elf2
Foxp1
Myc
Stat5b
Rorc
Foxf1
Rbpj
E2f5
Crebrf
Sox2
Max
Zfp239
Nr1d2
Dbp
Rest
Creb3l1

Figure 7. Reconstructed gene regulatory networks under normal conditions. TF-TF regulatory network at **(A)** 24h and **(B)** 72h time points. Color legend represents differential accessibility of the TFs (green: upregulated; red: downregulated; grey: not differentially accessible).

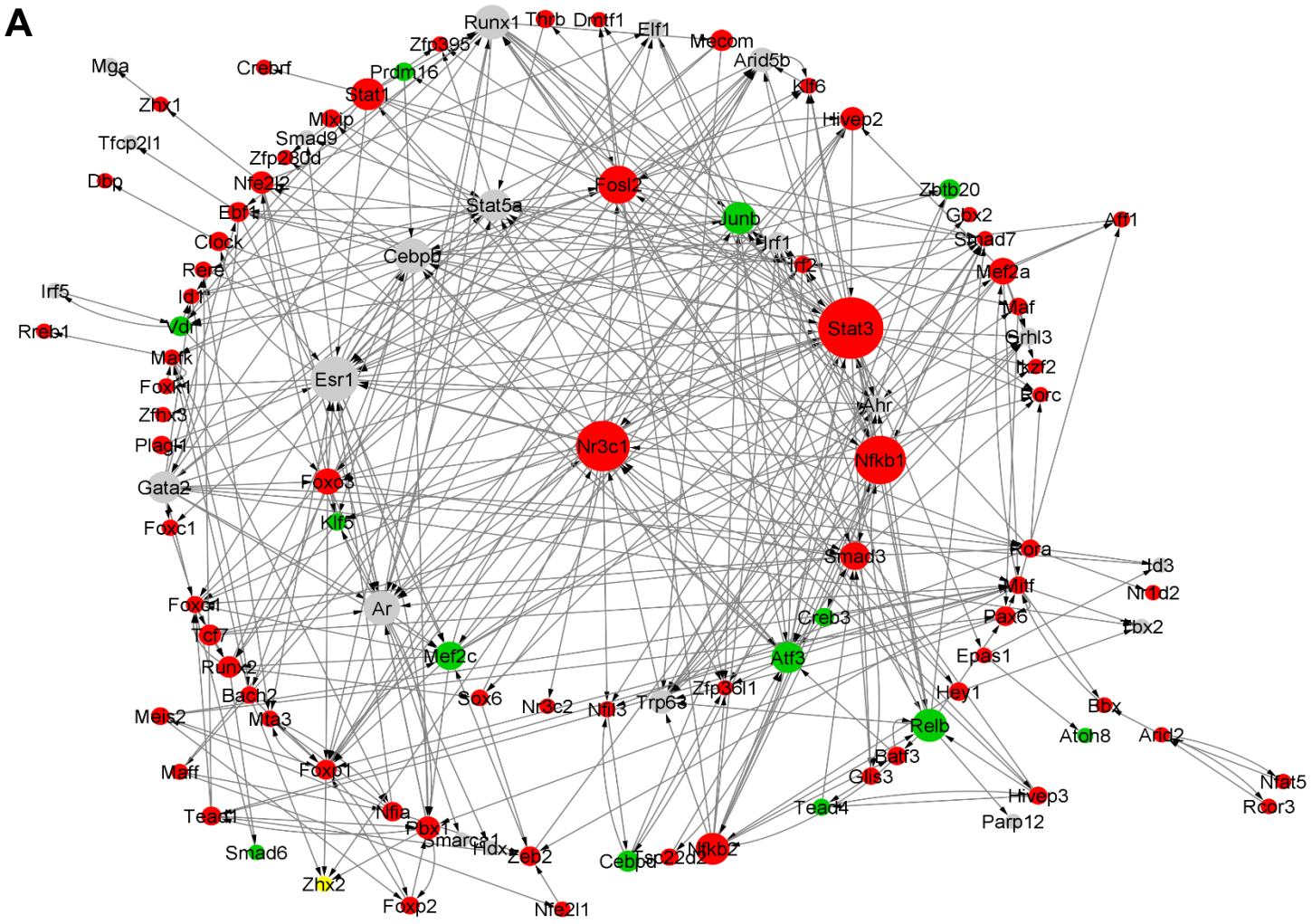
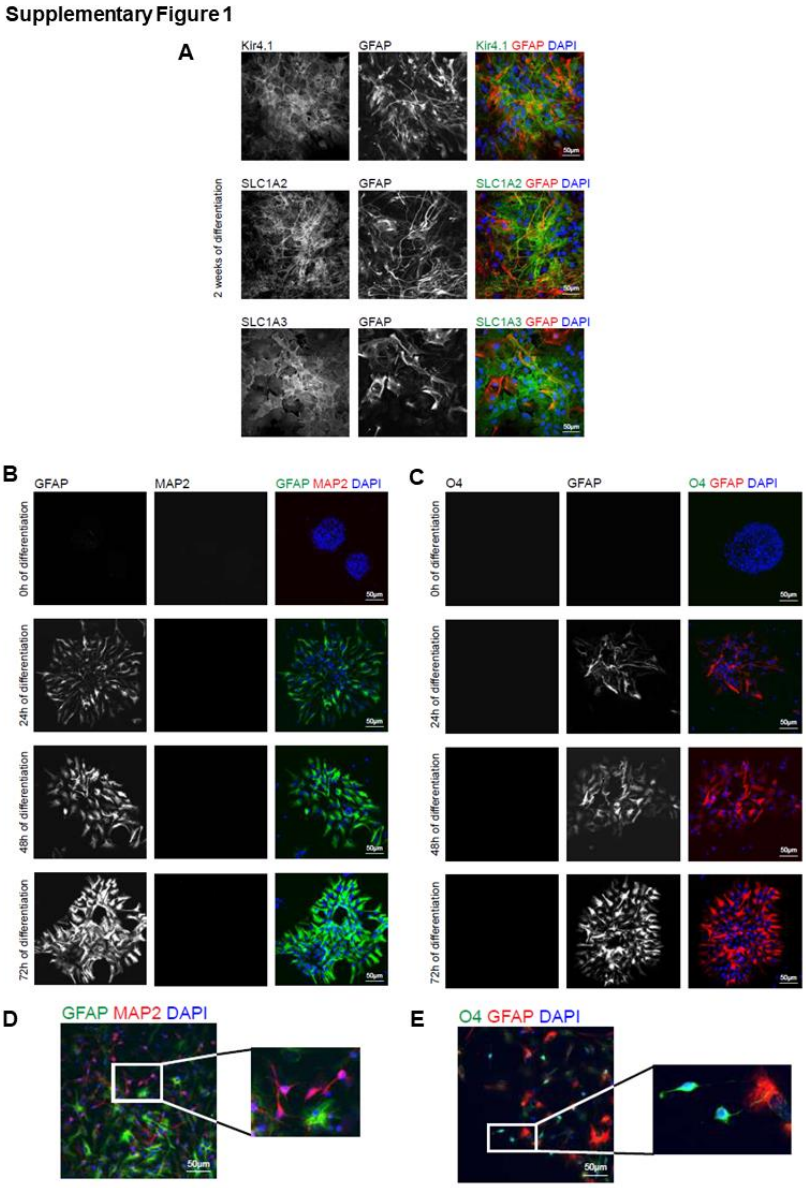

A
Runx1
Tfrb
Dnttf1
Elf1
Mecom
Arid5b
Klf6
Hivep2
Zfp395
Prdm16
Stat1
Creb3rf
Mga
Zhx1
Mlxip
Smad9
Tfcp2l1
Zfp238d
Dbp
Nfe2l2
Ebf1
Clock
Rere
Irf5
Rreb1
Vdr
Mafk
Foxk1
Zfhx3
Plagl1
Gata2
Foxc1
Foxo3
Klf5
Esr1
Ar
Foxa1
Tcf7
Runx2
Bach2
Meis2
Mta3
Maff
Foxp1
Tead1
Smad6
Znx2
Foxp2
Nfia
Pbx1
Smarca1
Hdx
Zeb2
Nfe2l1
Cebpb
Tsc22d2
Nfkb2
Fosl2
Stat5a
Junb
Irf1
Irf2
Cebpd
Stat3
Nr3c1
Nfkb1
Ahr
Smad3
Creb3
Atf3
Mef2c
Sox6
Nr3c2
Nfil3
Trp63
Zfp36l1
Zbtb20
Gbx2
Smad7
Mef2a
Maf
Srhl3
Rxra
Rorc
Id3
Nr1d2
Mitf
Pax6
Tbx2
Epas1
Hey1
Relb
Batf3
Glis3
Tead4
Hivep3
Parp12
Bbx
Atoh8
Arid2
Ntat5
Rcor3
Atf1
B
Zfp217
Myc
Dbp
Zfp36l2
Tsc22d3
Ar
Rorc
Stat5b
Rxra
Thrb
Klf9
Rarg
Mtf1
Zbtb7a
Zbtb37

Figure 8. Reconstructed gene regulatory networks under inflammatory conditions. TF-TF regulatory network at **(A)** 24h and **(B)** 72h time points. Color legend represents differential accessibility of the TFs (green: upregulated; red: downregulated; grey: not differentially accessible).

.

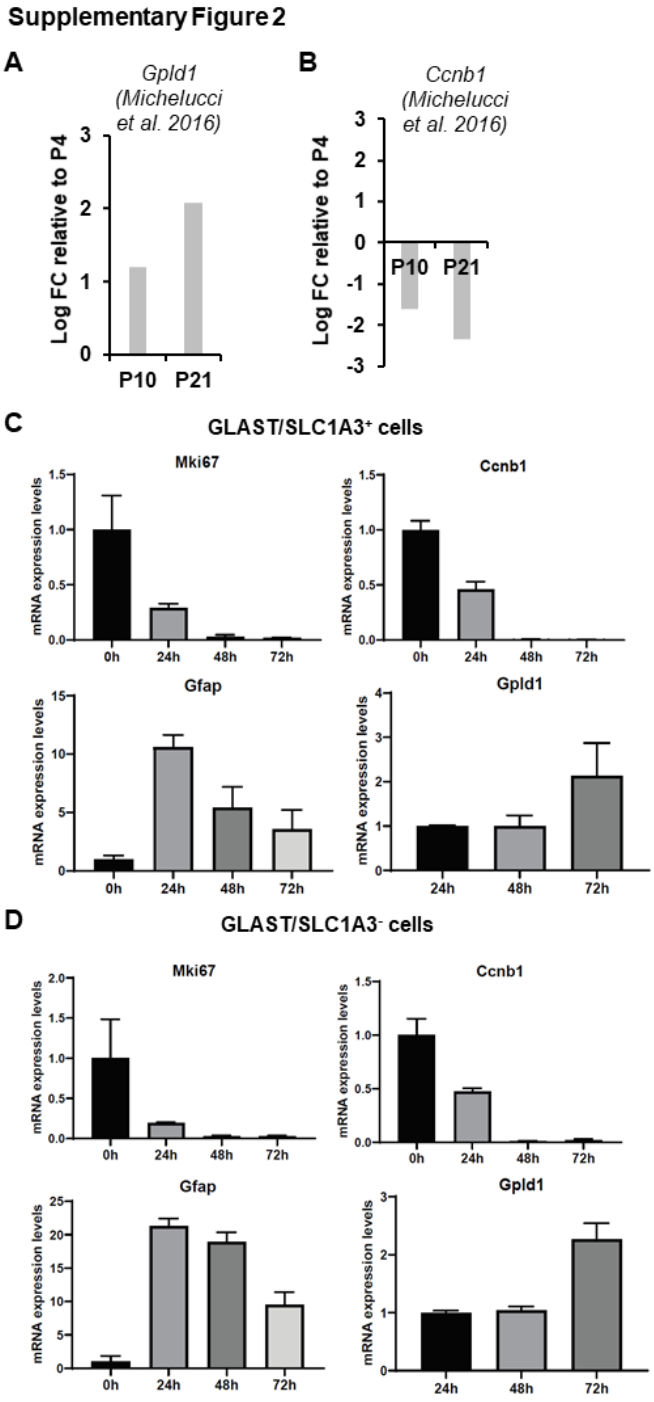
4.2.3 Supplementary Information

Supplementary figures

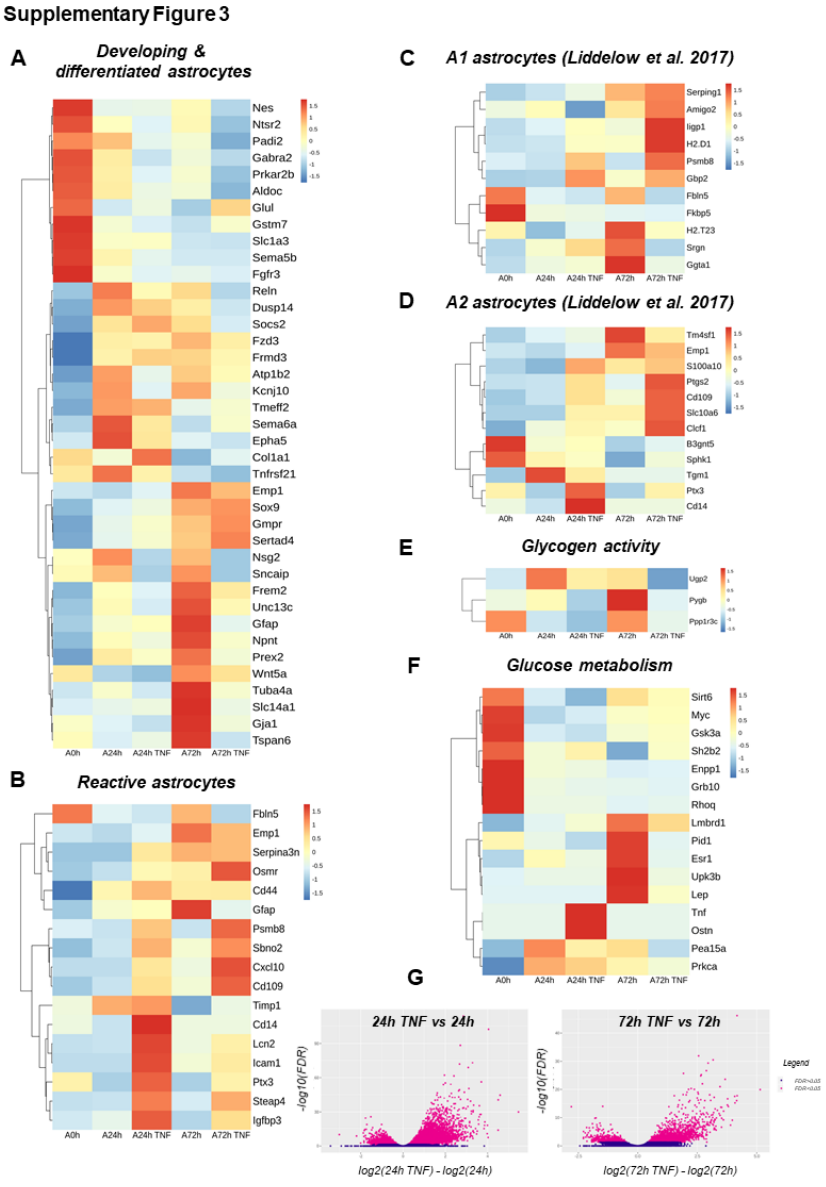


Supplementary Figure 1. Generation of pure and mature astrocytic populations. (A) Primary murine neurospheres (NSPs) were differentiated into astrocytes (DMEM, 10% FBS) for 2 weeks. Pictures show immunostaining of differentiated astrocytes labeled with Kir4.1, SLC1A2 or SLC1A3 (in green) together with GFAP (in red). Nuclei were counterstained with DAPI (in blue). Scale bar: 50 µm. **(B-C)** Primary murine neurospheres (NSPs) were cultivated for 72h in an altered differentiation medium (DMEM, 0.5% FBS, 1% B27 (- vitamin A), 2 ng/ml EGF). Pictures show

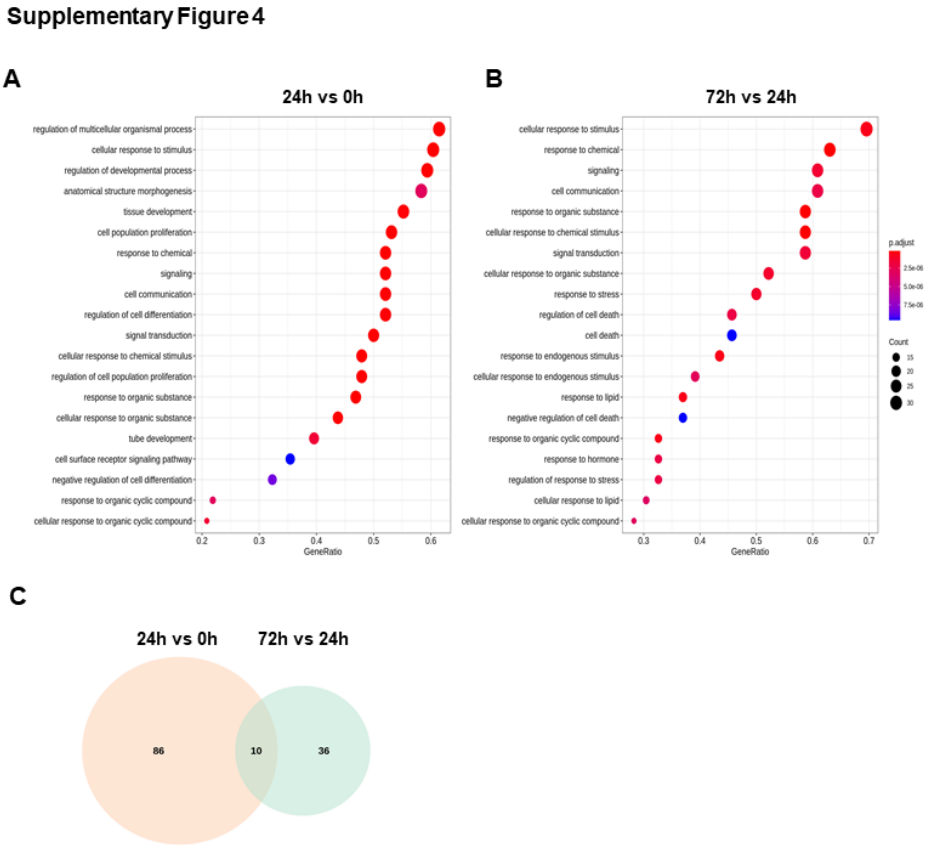
immunostaining of undifferentiated and differentiated NSPs at different time points (24h, 48h and 72h) labeled with antibodies against **(B)** GFAP (green) and MAP2 (red) or **(C)** O4 (green) and GFAP (red). Nuclei were counterstained with DAPI (in blue). Scale bar: 50 µm. **(D-E)** Magnification of **(D)** MAP2+ neurons (red) and **(E)** O4+ oligodendrocytes (green). Scale bar: 50 µm.



Supplementary Figure 2. Assessment of the validity of the *in vitro* model. (A-B) Gene expression levels of **(A)** *Gpld1* and **(B)** *Ccnb1* obtained from the transcriptomics analyses of GFP+ cells sorted from ALDH1L1-GFP transgenic mice comparing cells isolated at P4, P10 and P21 [25]. Results are shown as log FC relative to P4. **(C-D)** Primary murine neurospheres (NSPs) were cultivated under proliferation conditions or differentiated into astrocytes for 24h, 48h and 72h. We used GLAST magnetic beads to separate GLAST/SLC1A3+ and GLAST/SLC1A3- cells by MACS. Graphs show RT-qPCR results of *Mki67*, *Ccnb1*, *Gfap* and *Gpld1* in **(C)** GLAST/SLC1A3+ and **(D)** GLAST/SLC1A3- cells. Bars represent the mean ± SD of 2 independent experiments.

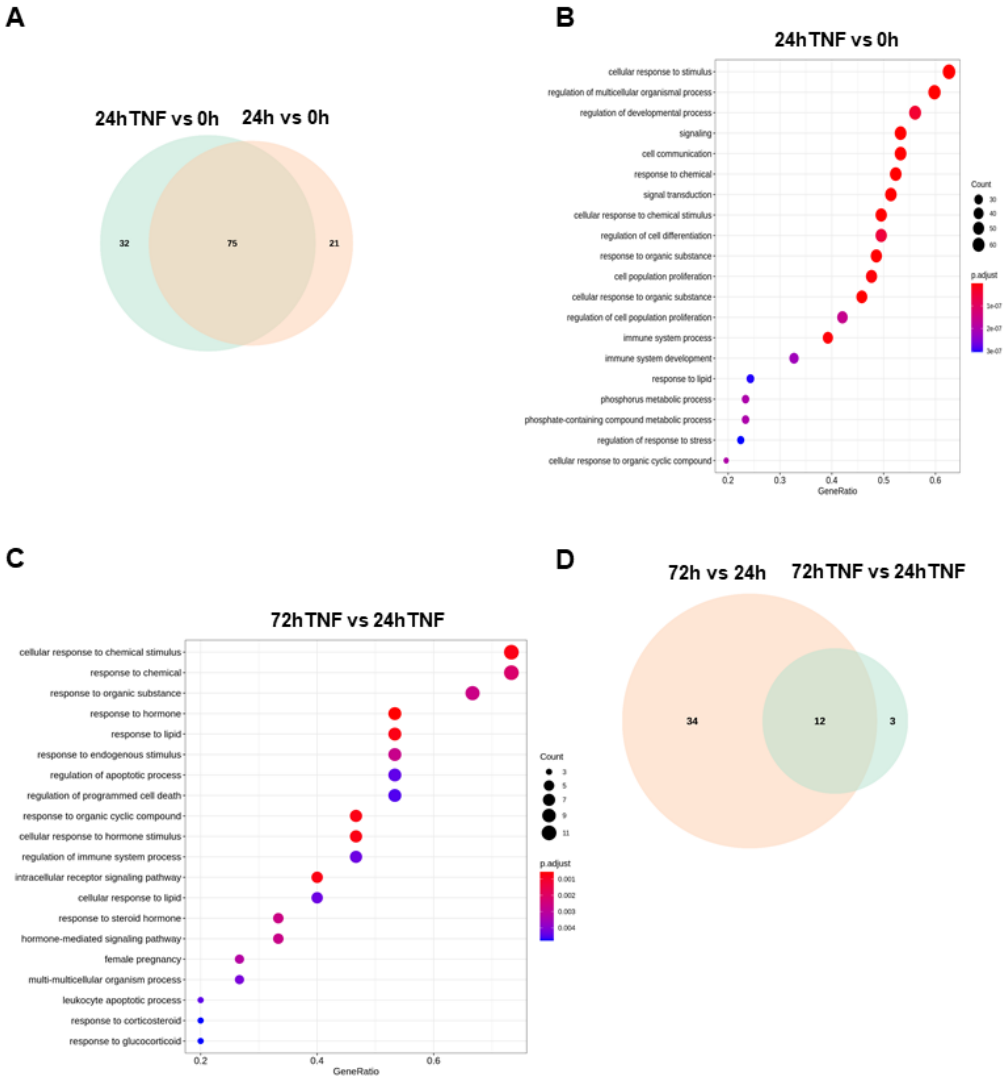


Supplementary Figure 3. Gene expression analyses for specific astrocytic markers and chromatin peak distribution comparing normal and inflammatory conditions along NSP differentiation. (A-F) Heat maps showing gene expression of specific astrocytic markers classified in various categories: **(A)** "developing and differentiated astrocytes", **(B)** "reactive astrocytes", **(C)** "A1 astrocytes", **(D)** "A2 astrocytes", **(E)** "glycogen activity" and **(F)** "glucose metabolism". Heat maps represent normalized (z scores) raw gene expression counts of differentially expressed genes (adjusted p value < 0.05) in normal (24h vs 0h, 72h vs 24h) and inflammatory (24h TNF vs 24h and 72h TNF vs 72h) conditions. **(G)** Volcano plots showing the distribution of open and closed differentially accessible peaks at different time points upon astrocytic development between treated and untreated NSPs (24h TNF vs 24h and 72h TNF vs 72h).



Supplementary Figure 4. Enrichment analysis of transcription factors in gene regulatory networks under normal conditions. (A-B) Gene Ontology terms for **(A)** 24h and **(B)** 72h time points. **(C)** Venn diagram showing the comparison between 24h and 72h.

Supplementary Figure 5



Supplementary Figure 5. Enrichment analysis of transcription factors in gene regulatory networks under inflammatory conditions. (A) Venn diagram showing the comparison of astrocytic differentiation between physiological and inflammatory conditions at 24h. **(B-C)** Gene Ontology terms for **(B)** 24h and **(C)** 72h time points. **(D)** Venn diagram showing the comparison of astrocytic differentiation between physiological and inflammatory conditions at 72h.

4.3 Feedback loops based regulation in cell-cell communication

4.3.1 Preface

In the study "Feedback loops in cell-cell communication identify key regulators of allergy and tolerance" we aimed to utilize cell-cell communication based network to identify regulators of allergy and tolerance using feedback loops analysis. Allergic diseases have seen a dramatic rise in their prevalence over the past several decades and these diseases put a lot of pressure on the medical system. Allergies are caused by persistent or repetitive exposure to environmental entities known as allergens resulting in the development of allergic inflammation. Further, tolerance can be induced to counter this inflammation and has been used to develop therapies for allergic diseases.

In this study we present a new computational tool C3Loop (Cell-Cell Communication based feedback Loop), to find key molecular regulators by identifying feedback loops in cell-cell communication networks. We made two novel contributions to this study, firstly this was the first time immune cells have been characterized from different tissues during various timepoints including the development of allergy, its resolution and induction of tolerance from this mouse model of allergy and AIT. These cells are responsible for the hyper-inflammatory response in the tissue upon encountering an allergen. Secondly, a novel computational tool is developed to identify feedback loops in the cell-cell communication network in the tissues. These feedback loops sustain the state of the cells in a tissue. These feedback loops can be both pro- and anti-inflammatory based on the type of molecules involved in the loops. We used this information about the feedback loops to identify and characterize them across tissues and timepoints. Sensitivity analysis is then performed using variation in computational parameters to the tool to remove bias due to ad hoc parameters. This sensitivity analysis also allows for curation by the selection of targets that are consistent across parameters. Comparisons are then performed between conditions to determine the regulators involved in feedback loops. These regulatory targets identified were then curated based on their possible role in allergy and tolerance. We propose a set of novel genes expressed by specific immune cell types that can be utilized to develop therapies and explore the mechanisms.

Contribution: I developed and implemented the algorithm. Planned the scRNA-Seq experiments, pre-processed the scRNA-Seq data, and analysed the data. I further wrote the manuscript and created the figures for the manuscript.

4.3.2 Manuscript

Feedback regulation in cell-cell communication identifies key regulators of allergy and tolerance

Kartikeya Singh[1]*, Guillem Montamat[2]*, Cathy Leonard[2], Feng He[2], Markus Ollert[2†], Antonio del Sol[1,3,4†]

[1]Computational Biology Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

[2]Department of Infection and Immunity, Luxembourg Institute of Health

[3]Computational Biology Group, CIC bioGUNE-BRTA (Basque Research and Technology Alliance), Derio, Spain

[4]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

* Co-authors

† Co-corresponding authors

Abstract

Over the last decades the incidence of allergic disease has risen proving to exert considerable social and economic burdens on society. Allergen sensitization involve different cell to cell communication mechanisms and inflammatory process. Allergen specific immunotherapies (AIT) stands as the only curative treatment for allergic diseases that is able to reinstate immune homeostasis. As allergen sensitization, AIT is known to involve immune cell crosstalk leading to the induction of tolerance. Cell-cell communication plays a vital role in immune responses by shaping the functionality of the cells and the coordination between them. Feedback loops between the cells of a tissue are responsible for the maintenance the tight inflammatory/tolerance equilibrium. Here we developed C3Loop (cell-cell communication based feedback loop), a novel computational tool for the identification of feedback loops in the cell-cell communication between cells in allergic reaction and in a highly successful model of AIT. Using these feedback loops, key regulators of the communication networks that play vital role in maintenance of allergen inflammation and tolerance in tissues were identified.

Introduction

The prevalence of allergic diseases is rising globally. Industrialized countries show a higher prevalence of allergy than non-developed countries, yet low- and medium-income nations are experiencing a sharp rise as they become more industrialized over time (Pawankar 2014). In a healthy state, the type 2 T helper response and allergen-specific immunoglobulin (Ig) E (IgE) production are necessary and beneficial to defend the organism against extracellular parasites, such as helminths. However, when the Th2 response is dysregulated or misdirected, allergic diseases appear. Allergy occurs when harmless antigens from foreign sources such as food, pets, pollen and others are not tolerated by the immune system of the host (Bluestone 2011, Giardino, Gallo et al. 2016). Allergic diseases develop in two clear and distinctive steps. In the first phase, known as the sensitization phase, the immune system meets the allergen for the first time and remains poised to react to it in future encounters. During this phase, allergen specific IgE is produced by plasma cells and accumulated on the surface of allergy effector cells such as eosinophils and mast cells (Lambrecht and Hammad 2014). During the second phase, or effector phase, allergen re-exposure

allows allergen binding to the IgE pre-loaded on the surface of allergy effector cells causing them to release soluble mediators that will cause the allergic reaction and leading toa variety of symptoms depending on the organ or tissue where this occurs (Yao, Chen et al. 2021). Due to the immunological nature of allergy, it is crucial that therapies for allergic diseases not only treat symptoms but also tackle the underlying stratum of the disease, the Th2 cells. Indeed, classically, treatments of allergic diseases have been focused on reducing allergic symptoms by blocking the activity of the main allergic mediators. However, blocking allergy mediators is often not enough to fully treat the symptoms. Indeed, the most prominent downside of symptomatic drugs is that upon treatment discontinuation, symptoms commonly relapse (Chung, La Grenade et al. 2019). Therefore, curative therapies such as allergen-specific immunotherapy (AIT) often represent the best treatment option in the context of allergic diseases. AIT has long-term effects in providing sustained relief of symptoms and reducing the need for symptomatic medication (Akdis and Akdis 2015). AIT stablishes immune tolerance by modifying the T and B cell compartment which are vital for implementing response towards the causative allergen (Marogna, Spadolini et al. 2010). One of the main and most described mechanism of AIT is the induction of T regulatory cells (Tregs) which has been described in both mice and humans (Palomares, Martin-Fontecha et al. 2014). For this study, we used an adjuvated-AIT to treat experimental allergic asthma to cat. This model was published as a highly successful method to treat allergy in mice and mimicked many of the human features of disease and treatment (Leonard, Montamat et al. 2021). AIT effectively changes the immune system response to allergens to which coordination via cell-cell communication between the different specialized cell types is crucial.

The role of cell-cell communication is intrinsic to the immune response, activation and resilience of the immune system depend on the appropriate regulation of this communication between immune cells and resident tissues. Under various studies the role of cell-cell communication has been studied and established across cell types and tissues (Ruch 2002, Huang, Xu et al. 2022, Overton and Mastracci 2022, Valls and Esposito 2022). Ligands and receptors form a vital fraction of this communication. Ligands produced by cells form the signal which interacts with surface bound molecules called receptors on the cells (Heldin, Lu et al. 2016). Interactions between cognate ligands and receptors lead to the initiation of signal transduction via signalling networks in the cells that conclude into gene regulatory networks downstream. These signalling

changes end in transcription factors that lead to regulation of gene expression and production of ligands as response to the incoming signals (Weidemuller, Kholmatov et al. 2021).

Immune cells communicate with each other to understand and regulate their environment. The immune system is composed of various cells with their specialized functions in the relay of this communication. This division of labor has been well studied under various physiological and disease conditions (Swanson, Pelanda et al. 2013). For example, macrophages play the role of sentinels as part of the innate immune system keeping vigil in the body to find any points of dysfunction or damage that could be harmful. They communicate that information via cells like dendritic cells to cells of the adaptive immune system like B and T cells (Garcia 2019). Cells interact with each other mainly via cytokines and physical contacts to mount a response and shape the function of immune cells. This network of interactions contains loops between different cell types which provide regulatory control to the immune response. These regulatory mechanisms are required as to control the proliferation and function of the immune cells such that the cells do not damage other cells and tissues in the vicinity. Feedback loops are constructed when outputs of a system are connected to its inputs. In communication these loops are formed by connecting ligands secreted to their cognate receptors. Feedback loops form the basis of regulation between immune cells in various tissues as they play a vital role of coordination during immune responses (Rahman, Tiwari et al. 2018, Zhou, Fu et al. 2022). Both allergic inflammation and tolerance in the tissues require these feedback loops to maintain the state of the cells and the tissues.

In this study we employ a novel tool C3Loop, that identifies feedback loops in cell-cell communication between the cell types of a tissue. Prior studies have employed computational frameworks for investigating feedback loops (Jung, Potapov et al. 2021, Xin, Lyu et al. 2022), however, these tools are designed to work on small networks. As the number of interactions increase in a cell-cell communication, the possible feedback loops would increase exponentially thus making the computation of such large and complex systems very time expensive. In our tool we break down this problem by deconstructing the cell-cell communication network into strongly connected component (SCC) clusters and then using the bridges between the clusters to find feedback loops. The tool also takes into consideration both pro- and anti- inflammatory loops which have not previously been explored, hence taking a more holistic approach to disease modelling. Finally, the tool creates a ranked list of potential gene candidate targets of inflammation

or tolerance by comparing feedback loops between conditions. These targets are based on the number of both pro- and anti-inflammatory feedback loops that a certain gene is involved in. Based on which feedback loops, pro- or anti-inflammatory the genes are more involved in, we identify 2 types of targets - allergy or tolerance respectively. The future aim of follow-up studies will be to interfere with those gene targets to prove the accuracy of our method and to identify them as potential candidates for new immunomodulatory therapies.

In this study, we generated a single cell RNA sequencing data set from specific immune cells from the Spleen and the peritoneal cavity (PC) of a highly successful CpG-AIT mouse model. These cells were extracted from 4 different timepoints at different stages during the AIT model to understand which feedback loops where important in each key phase of allergy modulation. The feedback loops-based analysis was performed by reconstructing the cell-cell communication network between the cells of a tissue. Key regulatory molecules and cell types were then identified using these feedback loops by performing topological analysis.

Materials & Methods

Animal welfare

All experimental animal procedures and experiments were reviewed and approved internally by the animal welfare structure (AWS) of the Luxembourg Institute of Health (LIH) and ultimately authorized by the competent authority of the Luxembourgish Ministry of Agriculture under the European Directive 2010/63/EU. The protocols approved and used was **DII-2019-05**.

High-dose CpG-ODN allergen-specific immunotherapy (hCpG-AIT) mouse model

The hCpG-AIT model consisted of three main stages: allergen sensitization, treatment (hCpG-AIT) and challenge. During the sensitization phase, the mice received three i.p. injections of 200 µl every two weeks (days 0, 14 and 28) containing Fel d 1 (10 µg) and aluminum hydroxide (500 µg), both diluted in PBS (Figure 1). Control mice received only aluminum hydroxide diluted in PBS. Two weeks after the last sensitization injection, the treated group of mice received three i.p. injections of 200 µl on days 42, 56 and 70 consisting of hCpG-AIT treatment containing CpG-

ODN (21 µg) and Fel d 1 (10 µg) diluted in PBS. Non-treated mice received a vehicle containing only PBS (Fig. 1). Two weeks after the last treatment course, the mice underwent a daily allergen challenge in the form of a 50 µl nasal instillation (NI) of Fel d 1 (5µg) diluted in PBS on days 83, 84 and 85 (Fig.1). All reagents and buffers used for the preparation of these solutions were free of endotoxin to prevent co-activation of any other too-like receptor than TLR9 during treatment and resemble the Good Manufacturing Practice (GMP) conditions necessary for human pharmaceutical products. Female BALB/cJRj mice were purchased from Janvier labs France (Strasbourg).

Sorting and single-cell sequencing of key immune cells

Mice were subjected to the protocol described above but sacrificed at key time points; day 43 (one day after first AIT injection, t1), day 82 (before nasal instillation challenge, t2) and day 86 (after full protocol completion, t3). Key immune cells previously identified in this model by mass cytometry (Leonard, Montamat et al. 2021) were sorted by fluorescence-activated cell sorting. B cells ($CD19^{+}$), NK cells ($CD49b^{+}NKp46^{+}$), macrophages ($CD11b^{+}F4/80^{+}CD11c^{-}$), classical dendritic cells (cDCs, $CD11c^{+}CD317^{-}CD11b^{-}$) and plasmacytoid DC (pDCs, $CD11c^{+}CD317^{+}CD11b^{-}$) were sorted from the peritoneal cavity (PC). The same immune cell types were sorted from the Spleen with the addition of T regulatory cells ($CD3^{+}CD4^{+}CD25^{hi}CD127^{lo}$) and T effector/memory cells ($CD3^{+}CD4^{+}CD44^{+}$). To note that pDCs were only sorted from the PC from the hCpG-AIT condition since too few pDC are found in a steady state PC. Cells from two mice were pooled for each time point (t1, t2 and t3) and group of mice (control, allergic and hCpG-AIT mice) (Figure 1). Cells were stained extracellularly and sorted in a BD FACSAria™ III Cell Sorter individually in a cooled 384 well plate containing mineral oil provided by Single Cell Discoveries (Netherlands). Plates were immediately frozen at -80^{o}C and shipped in dry ice to Single Cell Discoveries (Netherlands) for processing.

Single cell RNA Sequencing

Single-cell mRNA sequencing was performed according to an adapted version of the SORT-seq protocol (Muraro, Dharmadhikari et al. 2016). In brief, single cells were FACS-sorted into 384-well plates containing 384 primers and mineral oil (Sigma). After sorting, plates were snap-frozen on dry ice and stored at −80 °C. For amplification, cells were heat-lysed at 65 °C followed by cDNA synthesis using the CEL-Seq2 protocol and robotic liquid-handling platforms

(Hashimshony, Senderovich et al. 2016). After second-strand cDNA synthesis, the barcoded material was pooled into libraries of 384 cells and amplified using in vitro transcription. Following amplification, the rest of the CEL-seq2 protocol was followed for preparation of the amplified cDNA library using TruSeq small RNA primers (Illumina). The DNA library was paired-end sequenced on an Illumina Nextseq 500. Reads were mapped to the reference transcriptome of mm10 with Burrows–Wheeler Aligner and counts were generated.

Data cleaning and clustering

Raw counts files were combined and integrated based on the condition of the mice from which the cells were extracted and sequenced. Integration is performed based on workflow recommended for Seurat integration (Satija, Farrell et al. 2015, Butler, Hoffman et al. 2018, Stuart, Butler et al. 2019, Hao, Hao et al. 2021). Anchors were first identified in the data which are then used to integrate all the conditions to remove any bias in the data based on sequencing other technical/biological variations such as batch effects that could occur during sequencing. Clustering is then performed on the integrated data using Seurat and clusters are identified based on expression of known markers.

Cell-cell communication network reconstruction

Reconstruction of cell-cell communication was performed using the FunRes tool (Jung, Singh et al. 2021). FunRes first identifies conserved set of transcription factors and receptors in each cellular population and then the sustained intracellular signalling are recognised using a Markov chain model (Ravichandran, Hartmann et al. 2019). The active signalling genes are identified using this approach are then connected to the conserved receptors upstream and conserved transcription factors downstream of these signalling genes. The intercellular communication is then ascertained by linking the conserved set of receptors and ligands produced by the cellular populations. Bootstrapping is performed on the ligand-receptor network to get high confidence ligand-receptor interactions between cellular populations.

Feedback regulation in cell-cell communication network

To create feedback loops in the network the ligands are associated with the known set of transcription factors (TFs) identified in the previous steps. The association is done by calculating the shortest paths from the TFs to the expressed ligands in each cell type thus completing the

intracellular signalling starting from receptors and ending in ligands. These ligands are then connected to complete the communication network. This provides the basis of network for calculating feedback loops. Feedback loops are described as such when a path has the same starting and end points. For further details of the feedback loops algorithm refer to Supplementary data.

Sensitivity analysis

Computational tools have an inherent bias because of the choice of variables for the different parameters in the tool. We implemented a sensitivity analysis by performing the aforementioned feedback loop analysis using a range of parameters for variables that govern the generation of feedback loops in the communication networks. These variables include the path probability from signalling gene to conserved TFs, bootstrapping significance variable and feedback generation based on TF to ligand paths. This leads to production of communication networks that are distinct as variation in variables lead to loss of important nodes in the network.

Target identification

Immune feedback loops involve both inflammatory and anti-inflammatory interactions between cells. Similarly, the overall feedback in the tissues is responsible for either transformation of the tissue towards an inflammatory/disease or tolerance functional state for the tissue. We annotate the feedback loops by annotating individual genes as “pro-inflammatory”, “anti-inflammatory” or both. Feedback loops must consist of at least 2 of these interactions and the loops are annotated as pro- or ant-inflammatory based on the overall sum of pro-inflammatory and anti-inflammatory interactions in the loop. Individual interactions between the genes are annotated by using an OR logic between the genes. To ascertain targets for allergic and tolerance state of the tissue we apply a comparative analysis between feedback loops. Under the comparative analysis the targets are identified based on their involvement in feedback loops based on the type of target we wish to get. We identify the targets for allergic inflammation by comparing feedback loops in allergy vs control condition by calculating genes in more pro-inflammatory loops in allergy condition and less in the control condition. Similarly, tolerance targets are identified by ranking based on the number of anti-inflammatory feedback loops the gene is involved in. Further curation of targets was performed by identifying the cell type and level of gene expression across various conditions for the genes.

Results

Identification of cell types and gene expression profiles

Selected immune cells were extracted from the Spleen and peritoneal cavity (PC) from the three groups of mice- control, allergic and AIT- at various timepoints during the different stages of the mouse AIT (Fig. 1). Extracted cells were then sorted at a single-cell level using known surface markers by FACS and sequenced at Single Cell Discoveries (Netherlands). The SORT-Seq platform, an automated version of the CEL-Seq2 protocol (Hashimshony, Senderovich et al. 2016, Muraro, Dharmadhikari et al. 2016) was used. Quality measurement and pre-processing were performed using the Seurat package in R. The cell and transcript numbers per sample plate, distribution of detected genes were also analysed. After pre-processing, 9947 cells remained which were then clustered downstream using the recommended workflow. A mean of 4450 transcripts were expressed by the cells, which represent 1497 genes on average (SupplFig. 1A). The gene transcripts count, and genes expressed showed a correlation of 0.93 (SupplFig. 1B). Clustering was then performed by using anchor integration method in the Seurat package of data from the different samples. The number of clusters was then optimised using the silhouette score analysis of the clusters. Using the *FindAllMarkers* function all the top expressed genes for each population were investigated and dotplot was created from Seurat package in R to determine the expression of various markers in the identification of the cell clusters.

We identified and annotated 19 clusters using known literature genes based on CD markers. One of the clusters (Cluster 12) was removed from downstream analysis due to lack of known immune phenotypical markers in the population. The following cell populations were identified in our dataset: Natural killer cells (NK cell, $Cd3^{-}$ $Nkp46^{+}$), $Cxcr7^{+}$ Macrophage ($Cd11b^{hi}$ $F480^{hi}$ $Cxcr7^{+}$ $Ccr2^{-}$ $MHCII^{+}$), $Ccr2^{+}$ Macrophage ($Cd11b^{mid}$ $F480^{mid}$ $Cxcr7^{-}$ $Ccr2^{+}$ $MHC2^{+}$), $Cxcr2^{+}$ Macrophage ($CD11b^{+}$ $F480^{-}$ $Cd64^{+}$ $Cxcr2^{+}$ $MHCII^{-}$), pDC ($Cd11c_{+}$ $Cd11b_{-}$ $Cd317_{+}$ $MHCII_{+}$), $Cd4_{+}$ $Ccr7_{+}$ DC ($Cd11c_{+}$ $Cd11b_{mid}$ $Cd4_{+}$ $Ccr7_{+}$ $MHC2_{hi}$), cDC1 ($Cd11c^{+}$ $Cd11b^{-}$ $F480^{mid}$ $MHC2^{+}$), cDC2 ($Cd11c^{+}$ $Cd11b^{+}$ $MCHII^{lo}$), DC ($Cd11c^{+}$ $Cd11b^{mid}$ $MHCII^{hi}$), Non-follicular B cell ($Cd19^{+}$ $Cxcr5^{-}$ $MHCII^{+}$), Follicular B cell ($Cd19^{+}$ $Cxcr5^{+}$ $MHCII^{+}$), Central memory $Cd4^{+}$ T cell ($Cd3^{+}$ $Cd4^{+}$ $Foxp3^{-}$ $Gata3^{-}$ $Cd44^{mid}$ $Cd62l^{mid}$), Effector memory $Cd4^{+}$ T cell ($Cd3^{+}$ $Cd4^{+}$ $Foxp3^{-}$ $Gata3^{-}$ $Cd44^{+}$ $Cd62l^{lo}$), Effector memory $Cd4^{+}$ Th2 cell ($Cd3^{+}$ $Cd4^{+}$ $Foxp3^{-}$ $Gata3^{+}$ $Cd44^{+}$ Cd62llo), Naive Treg ($Cd3^{+}$ $Cd4^{+}$ $Foxp3^{+}$ $Cd25^{hi}$ $Gata3^{-}$ $Ccr7^{mid}$), Active Treg ($Cd3^{+}$ $Cd4^{+}$ $Foxp3^{+}$ $Cd25^{hi}$

Gata3^{-} Ccr7^{-}), NKT cell (Cd3^{+} Ncr1^{+} Cd4^{+} Foxp3^{+} Gata3^{+}) and Mast cell (Gata2^{+}, Gp49a^{+}) (Fig. 2A,B).

Effect of CpG-AIT on key cell types

Some of the identified populations were found specifically enriched/depleted under certain conditions/timepoints. In peritoneal cavity (SupplFig. 2A), Non-follicular B cell, Cxcr7^{+} Macrophage and NK cell clusters were reduced at the CpG-AIT t1 timepoint. On the contrary, the NKT cell, pDC, Cxcr2^{+} Macrophage and Cd4^{+} Ccr7^{+} DC clusters were enriched in treated mice at t1 when compared to both non-treated groups. Indeed, a clear shift in the macrophage phenotype can be observed in the site of injection 24h after the treatment. We observed a switch from lower number of Cxcr7-expressing macrophages towards an enrichment of Cxcr2 macrophages. The Cxcr7^{+} Macrophage cluster was also depleted at Allergic t3, in favor of Ccr2^{+} Macrophages. Regarding population changes in the Spleen (SupplFig. 2B), we found depletion of the Cd4^{+} Ccr7^{+} DC, Activated cDC2 and cDC2 clusters, as well as a slight decrease in pDC and NK cell numbers for CpG-AIT t1 timepoint. Simultaneously, we observed that Central memory Cd4^{+} T cells, Effector memory Cd4^{+} Th2 cells, Mast cells, NKT cells and Cxcr2^{+} Macrophages were enriched, the latter having the same tendency as in the peritoneal cavity. The lower pDC number observed in t1 is still found at t3 in CpG-AIT vs Allergic. Similarly, CpG-AIT induces a persistent reduction of Central memory Cd4^{+} T cell in the CpG-AIT for t2 and t3 when compared to non-treated conditions.

Reconstructed cell-cell communication networks

Cell-cell communication networks were reconstructed for each time point, condition and tissue from which cells were extracted and sequenced. To compare the communication networks, a reconstruction was performed using default parameters of FunRes (Jung, Singh et al. 2021). The average number of populations in the communication network was 5 and 12 for peritoneal cavity and Spleen respectively. The communication networks consisted of 215 and 651 interactions on average respectively for PC and Spleen.

We compared the Ligand-Receptor (L-R) interactions for various networks (Control vs Allergic and Allergic vs CpG-AIT) to identify the combinations that are unique and shared

between the conditions. The comparison aids in identification of differential L-R interactions between the conditions of concern at various timepoints. We also inquired the role of individual ligands and receptors in the networks by analysing the fraction of interactions they were involved in. As expected, the top ligands and receptors perform basic physiological functions of regulating immune cells, cellular adhesion, and migration of cells. This obvious but important finding validates our tool since indeed these are genes that compose the fundamental core of cell-cell interaction between immune cells.

For all the timepoints (t1, t2 and t3) we performed two comparisons – Allergic vs Control and CpG-AIT vs Allergic, for both PC and Spleen. For t1, in the comparison Allergic vs Control, the cell-cell communication network for PC has 5 and 4 unique interactions in Allergic and Control condition respectively. We observed that Spp1 ligand interacts with Cd44 in Allergic whereas it interacts with Itga4 and S1pr1 in Control. For the comparison CpG-AIT vs Allergic, in the PC several receptors were unique to both conditions. Plaur is unique to Allergic while Il1rap, Il1r2, Tnfrsf1a and Ifngr2 were some of the unique receptors for CpG-AIT network (Fig 3A). In Spleen, for the Allergic condition comparison, the cell-cell communication network had unique interactions with Ccr1, Lifr and Nrp1 receptors compared to Control. For CpG-AIT condition the unique interaction included Il1rn-Il1r2, Ifng-Ifngr2 and multiple chemokines (Cxcl9, Cxcl10, Cxcl11) with receptor Cxcr3. While Ebi3-IL27ra and Tnf-Traf2 interactions were present in Allergic conditions (Fig 3B).

For t2 timepoint in PC, several interactions were found for the Allergic vs Control comparison with the receptors Ccr1, Ccr2 and Ccr5 being unique for Allergic condition. While for the CpG-AIT vs Allergic comparison Ifng interaction with both Ifngr1 and Ifngr2 were unique to CpG-AIT, interaction of Ccl7-Ccr2 and Ccl12-Ccr2 were only found in the Allergic conditions (SupplFig 3A). For the Spleen, in the Allergic vs Control comparison Cxcl10-Ccr3 is unique to Allergic conditions. In the CpG-AIT comparison, Tnfrsf1b interactions with Lta and Tnf are unique to CpG-AIT while multiple interactions with receptors Ccr2, Ccr3, Ifngr1, Ifngr2, Il2rg and Cxcr3 were unique to Allergic conditions (SupplFig. 3B).

At timepoint t3 in PC multiple cytokines interact with Ccr1, Ccr2 and Ccr5 receptors in Allergic conditions while the Cxcl13-Cxcr5 pair was found unique to Control. For the CpG-AIT vs Allergic comparison, Ccl7-Ccr2, Ccl2-Ccr1 and Ccl2-Ccr5 are unique to the Allergic conditions

while Spp1-S1pr1, Il1b-Il1rap, Cxcl13-Cxcr5 and Ifng-Ifngr2 are unique to CpG-AIT conditions (SupplFig 3C). In Spleen, the interactions Tnf-Tnfrs1a, Ifng-Ifngr1, Ifng-Ifngr2 along with others were unique to the Allergic conditions when compared to the communication network in Control. Lta-Tnfrsf1b, Ccl5-Sdc4 and Lpl-Lrp1 are unique to Allergic conditions compared to CpG-AIT while Ebi3-Il27ra, Ccl5-Cxcr3, Il6-Il6ra are unique to CpG-AIT, along with multiple cytokines interacting with Ccr1 were unique to CpG-AIT in contrast to Allergic conditions (SupplFig 3D).

We observed many interactions shared across the conditions at different timepoints and tissues in the re-constructed networks. These common interactions between the conditions demonstrates the common biological processes between the conditions which would represent basic functions required in the tissue. Similarly, we prove that a very well-studied immune disease such as allergy, and a highly efficient successful immunomodulatory treatment have distinct cell-cell communication dynamics showcased by the distinctive feedback loops of each condition.

Regulators of feedback loops

To identify the key unique central regulators of allergy and tolerance we investigated the feedback loops in our reconstructed cell-cell communication networks using C3Loop. For this purpose, the intracellular signalling was extended from receptor to downstream ligands via the conserved set of receptors in each population. After the intracellular signalling was completed, these ligands were connected to their cognate receptors in all the cell populations in the network to form loops. The molecular regulators were then identified by detecting the feedback loops in the networks by using our algorithm. As the study is concerned with the pro- and anti- inflammatory processes, the feedback loops are also annotated as such by using literature based evidence of the inflammatory nature of each molecule involved. We next compared these feedback loops across conditions for a given timepoint to identify the regulators of allergic inflammation and tolerance. For this purpose, the Allergic vs Control and CpG-AIT vs Allergic comparisons pairs were performed respectively. The final set of targets were then ranked based on their involvement in pro- (Allergy targets) or anti- (Tolerance targets) inflammatory feedback loops across timepoints. The top ranked genes were considered as candidates and then curated based on biological knowledge of their role in allergy/tolerance. These candidates were considered across timepoints and tissues as these provide for higher confidence in the predictions. In addition, novelty, and feasibility to be translated to a

future treatment were also factored in the final selection of target genes for experimental validation.

The top candidates include various cytokines/chemokines along with their receptors, integrins and various phenotypic markers vital for cellular function of cells. We identified a final list of candidate targets based on their function, expression profile (Fig. 4 for Spleen and SupplFig. 4 for PC) and novelty being predicted to be tested in subsequent validation experiments. The targets for allergic inflammation were predicted using the comparison Allergic vs Control conditions across tissues and timepoints and were then curated based on their role in allergy. For allergy the final predicted targets include Tnfrsf1a (TNFR1), Klrd1 (CD94), Fn1, Lpl/Lrp1, Adam17 and Mmp9.

We also identified the populations that expressed the key regulators of feedback loops in order to discover their possible roles during allergy and tolerance. Under allergic conditions in Spleen, Tnfrsf1a was relatively higher expressed in Effector memory $Cd4^+$ Th2 cells in allergic t2 and $Cd4^+$ $Ccr7^+$ DCs at Allergic t3 timepoint. Tnfrsf1a is also expressed in NKT cell, activated cDC2 and macrophage populations across various timepoints in both PC and Spleen. Klrd1 was expressed by pDC in Spleen as well as in NK cell in both PC and Spleen with higher expression in different Allergic and CpG-AIT timepoints. Fn1 was expressed across multiple cell types in both PC and Spleen with higher expression for t1 and t2 Allergic timepoints. Both Lpl and Lrp1 were found to be expressed in many cell types mainly in macrophage and dendritic populations with significantly higher expression at Allergic timepoints in both PC and Spleen, Adam17 was expressed in most populations in both PC and Spleen with variable expressions across timepoints. Mmp9 was expressed in Effector memory $Cd4^+$ Th2 population at Allergic t2 timepoint in Spleen while having relatively higher expression in NK, NKT and Follicular B cells in CpG-AIT timepoints.

Next, using the comparison CpG-AIT vs Allergic conditions we were able to identify regulators of tolerance in the cell-cell communication networks using our feedback loops algorithm. These genes can be targeted to inhibit the tolerance state induced by the cell-cell communication network and hence prove their role in the success of the therapy. Indeed, predicting and proving that these genes are important for reinstating tolerance in the context of allergy would make them ideal targets to be enhanced in AIT and therefore improve the effectiveness of the

therapy. The most interesting of these targets are as follows – S1pr1, Il1r2/ Il1rn, Lifr/Il6st, Il18r1. These targets were selected for their role in tolerance, expression across condition/timepoints (Fig. 5 for Spleen and SupplFig. 5 for PC) and their novelty.

S1pr1 was highly expressed across B cell populations and macrophage in PC in CpG-AIT t1 and t2 timepoints as well as in Non-follicular B cell and Central memory $Cd4^+$ T cell in Spleen in CpG-AIT t2 timepoint. Both Il1r2 and Il1rn were expressed mainly in macrophage populations across PC and Spleen with higher expression at CpG-AIT timepoints. Il6st and Lifr were expressed in dendritic populations in Spleen with higher expression in CpG-AIT while being expressed across multiple cell populations in PC. The expression of Il18r1 was higher at CpG-AIT t1/t2 timepoints in the B and T cell populations of PC. In Spleen, Il18r1 was expressed at relatively higher expression at CpG-AIT t1 timepoint in Effector memory $Cd4^+$ Th2 cells.

Apart from these top interesting candidates there was a plethora of attractive candidates which are involved in many biological processes. These genes perform basic functions for different cellular populations and targeting these genes could result in loss of functionality of the cells and may lead to disruption in immune responses. These genes can be classified in 3 categories, functional cell surface markers (B2m, C3, Cd19, Cd14, Cd3g, Cd247, etc), chemokines ligand/receptors (Ccl3, Ccl5, Ccr5, etc) and integrins (Itgam, Itgal, Itgav, Itgb7, etc). The functional marker genes are vital for different cellular populations and suppression of any of these genes would result is loss of functionality by their respective cell populations. These chemokines and their respective receptors have previously been studied for both pro- and anti-inflammatory roles. As chemokines play role in both migration and function of immune cells, inhibition of their genes can damage fundamental processes of the immune system. Lastly, the aforementioned integrins are responsible for migration and residence of cells in tissues and loss of these integrin proteins in tissues can dismantle the tissue structures as well. Therefore, these genes were not considered further as targets as their role is vital to immune system as well as to basic cellular functions.

Discussion

The development of allergy occurs when an abnormal immune response is generated against an harmless environmental substance. The allergic response is accompanied by robust inflammation which may have a debilitating impact over an individual's health. With the rising burden of allergies across the globe, it has become pertinent to understand the mechanisms involved in allergic inflammation to develop therapeutic strategies to counteract the inflammatory response either by inhibiting cellular or molecular entities that are at its origin. Allergic inflammation has been linked to the responses mediated by Th2 cells which secrete pro-inflammatory molecules that lead to hyperresponsiveness against the allergen. Th2 cells, while vital for understanding allergic immune responses, are part of a compendium of immune cells that are involved in these responses including dendritic cells, macrophages, regulatory T cells, other T cell subsets, etc. Under healthy conditions, these cell populations communicate with each other to regulate their response and maintain their functionality. Feedback loops between these cells are essential for maintenance of these functional states and the genes involved in these feedback loops constitute the key regulators of inflammation/tolerance between the cells. These regulators can then be purposed to develop therapeutic interventions against allergic inflammation and/or for tolerance induction. In this study, we generated gene expression data at single cell resolution using cells extracted at various timepoints of the sensitization, tolerance induction and challenge stages of a highly successful adjuvanted-AIT mouse model (Leonard, Montamat et al. 2021). Indeed, including the immunostimulatory CpG adjuvant in AIT strikingly reduces the burden of allergy, hence, it is the ideal model in which study how cell-cell communication feedback loops transition from an inflammatory and dysregulated state back to homeostasis and tolerance.

We used FunRes (Jung, Singh et al. 2021) to model cell-cell communication networks and used our new tool C3Loop to study feedback loops. This tool represents the next step in understanding not just the role that cell-cell communication plays in regulation of cellular functions but also the role of feedback in maintenance of cellular and tissues states. In order to identify these feedback loops, our tool starts first with completing the intracellular signalling between surface receptors and secreted ligands. This leads to completion of the cell-cell communication networks which can be used afterwards for feedback loop identification. Next, we use a deconstructing framework to make the process of this identification tractable both in time

and computationally. Identification of feedback loops allows for their annotation based on their role in inflammation which can be used to understand the role of different genes and cellular populations. These pro- and anti- inflammatory feedback loops play antagonistic roles and the balance between the two dictates the progress or remission of disease. Therefore, this tool can be utilized to identify target genes whose inhibition could disrupt different processes involved in the induction, progression or maintenance of diseases.

Identification of immune cell types

We identified various immune populations in the cells extracted across different stages of the experiment. Unexpectedly, although mast cells and NKT cells, specifically $CD4^+$ NKT cells, were not included in the original sorting strategy (material and methods) we detected them in the downstream analysis of the scRNA-seq data in both the Spleen and PC. NKT cells could have been included among the sorted cell when sorting T cell and NK cells since they share many markers. In respect to mast cells, they could have been wrongly sorted when sorting myeloid cells such as macrophages, pDCs, DCs due to the complexity of this cell linage.

We have found that macrophages change their phenotype upon treatment. Indeed, we found an enrichment of $Cxcr2^+$ macrophages in detriment of $Cxcr7^+$ macrophages after only 24h CpG-AIT at t1. Interestingly, the expression of Cxcr7 on macrophages have been associated with M1 (pro-inflammatory) phenotype (Zhang, Zhang et al. 2020) and inflammation induction in several disease contexts such as atherosclerosis (Ma, Liu et al. 2013, Cai, Chen et al. 2020). Conversely, Cxcr2-expressing macrophages (M2, anti-inflammatory) have been associated with a tolerogenic and regenerative role (Di Mitri, Mirenda et al. 2019). Indeed, these two distinct phenotypes of macrophages fit the overall interpretation of the data as $Cxcr2^+$ anti-inflammatory and regenerative macrophages are induced by CpG-AIT while $Cxcr7^+$ inflammatory macrophages are dampened. Similar to macrophages, we have also observed crucial modulation of the DC compartment. Plasmacytoid DCs, crucial for tolerance induction under CpG-modulation conditions (Volpi, Fallarino et al. 2013), were shown to be increased in peritoneal cavity but decreased in the Spleen at t1. These phenomena could indicate a local recruitment to the site of injection but systemic migration of pDC to secondary immune organs, such as lymph nodes, to induce allergen tolerance. Further in timeline of the CpG-AIT model, pDCs are increased in both peritoneal cavity and Spleen at t3, solely in the CpG-AIT conditions. The interpretation of this result becomes complex since

no injection is performed in the peritoneal cavity at that point, but a nasal allergen challenge is applied. However, this generalized induction of pDC numbers could be due a specific tolerance reaction occurring only in previously treated mice due to the imprinting of innate immune cells trained immunity as described in other settings (Netea, Dominguez-Andres et al. 2020). Other DCs subtypes were also modulated by CpG-AIT. One of the most notable changes in DCs was the decrease of both cDC2 and activated cDC2 cells at t1 after CpG-AIT but not in the non-treated conditions. This result fits our previous observation (Leonard, Montamat et al. 2021) and brings another layer of evidence by which CpG-AIT can reduce allergy burden since cDC2 have been implicated in antigen presentation for the generation of Th2 cells (Kumar, Jeong et al. 2019). The $Cd4^+$ $Ccr7^+$ DC cluster was increased in PC while dramatically reduced in Spleen at t1 after treatment. This cluster of DCs follows a similar pattern as pDCs, which could indicate an increased migratory state. Indeed, $Ccr7^+$ expression in DCs has been linked with migration (Jang, Sougawa et al. 2006) and favours the meeting of DCs with naïve T cells for antigen presentation (Brandum, Jorgensen et al. 2021).

In terms of adaptive immunity, we observed decreased non-follicular B cells in the peritoneal cavity indicating a possible migration of this cell type. However, the biggest effect in the immune adaptive immune system was found in $CD4^+$ T cells. We observed higher numbers of central memory and effector memory Th2 $CD4^+$ T cells at t1 only in the CpG-AIT conditions. This phenomenon is most likely derived from the fact that the CpG-AIT includes the causative allergen, hence antigen-specific $CD4^+$ T cells, specially Th2 cells, would react and expand when presented with the allergen. Interestingly, the increase in central memory $CD4^+$ T cells at t1 was followed by a persistent decrease of the same cells in the treated mice at t2 and t3, suggesting that CpG-AIT could be depleting part of the central memory and long living CD4 T cells pool responsible for perpetuating the allergic reaction. Surprisingly, we did not observe any changes in the number of both naïve and memory CD4 Treg cell subtypes.

Finally, and unexpectedly, we found that although NKT cells are not know to express TLR9 they are modulated by CpG-AIT, probably due to the indirect effect of TLR9-activated DCs and pDCs (Paget, Mallevaey et al. 2007). Contrary to the literature that attributes NKT a pro-allergic role (Iwamura and Nakayama 2018), we found NKT cells induced by the CpG-adjuvanted AIT 24h after the first injection of the treatment in both the peritoneal cavity and Spleen,

suggesting that this cell type could be contributing to the success of the therapy. Indeed, these NKT cells induced by CpG-AIT are very different from the NKT cell types associated with allergy since they express FoxP3 and Gata3, transcription factors usually attributed to Treg and Th2 cells respectively (Fig. 2C) (Geginat, Paroni et al. 2014). FoxP3$^-$expressing NKT have been described when induced by TGF-β (Monteiro, Almeida et al. 2010). These cells have been implicated in tolerance situations such as cancer (Li, Peng et al. 2015), but the literature on this cell type is scarce. Interestingly, double positive Gata3$^+$ FoxP3$^+$ biTregs resembling the NKT cells we have found have been described to specifically supress Th2 cells and therefore could be crucial for the success of AIT (Sjaastad, Owen et al. 2021). Indeed, we have previously shown that such biTreg cell type is induced by CpG-AIT at t3 (Leonard, Montamat et al. 2021). To our knowledge, this is the first time that a FoxP3$^+$ NKT cell have been described in the context of re-introduction of tolerance through AIT. In addition, to our best of knowledge this is the first time that double positive FoxP3$^+$ Gata3$^+$ NKT have been described in the mouse system and their possible contribution to an allergen tolerization process.

Reconstructed cell-cell communication networks and ligand-receptor pairs

Re-constructing the cell-cell communication networks allowed us to identify the ligand-receptor interactions between the different immune cell populations across timepoints and tissues in our experiments. We compared ligand-receptor interactions across conditions for all the timepoints and tissues. Using this comparison, we were able to identify most important ligand-receptor combinations.

Spp1 (Osteopontin) has been recognised as a selective marker for a macrophage subpopulation implicated for its role in infiltration and M2 polarization. It has also been associated with adverse outcome of various cancers (Dong, Wu et al. 2021). Cd44 is a transmembrane glycoprotein which is expressed by some leukocytes and has been studied for its involvement in both health and disease. Among other function, Cd44 plays important in leukocyte recruitment and may have both protective or disease promoting roles in asthma (Rothenberg 2003, Katoh 2021). Nrp1 is a transmembrane protein playing important roles in axon guidance, angiogenesis, and cell migration. Expression of Nrp1 on T cell and dendritic populations provides capacity to suppress anti-tumor responses. On Treg cells Nrp1 plays suppressive role in incidence of inflammatory diseases (Gaddis, Padgett et al. 2019). Plaur (Urokinase plasminogen activator receptor), which is

the cognate receptor for Plau, is a serine protease receptor and is involved in airway remodelling a clinical feature of asthma (Ierodiakonou, Portelli et al. 2016). Plaur also plays role in cancer-associated processes such as metastasis, proliferation, cell differentiation and angiogenesis via cell-cell and cell-extracellular matrix interactions (Liu, Chen et al. 2021). Tnf is a pleiotropic cytokine, and its functions are mediated via two receptors Tnfr1 and Tnfr2 which act via NF-kB pathways. Interestingly for the reinstalment of tolerance, Tnf-Tnfr2 interactions plays role in activation, expansion, function, and stability of Treg cells. Tnfr2 expressing Treg cells have also been implicated in suppressive role in cancer immune invasion (Yang, Islam et al. 2021). Traf2 mediates crosstalk between Tnfr1 and Tnfr2 receptors and thus dictates the outcomes of TNF stimulation. Traf2 is involved in negative regulation of apoptosis and necroptosis via Tnfr1 signalling and is also necessary to keep the non-canonical NF-kB pathway inactive (Borghi, Verstrepen et al. 2016). Cxcl9, Cxcl10 and Cxcl11 are chemokines that act via Cxcr3 receptor expressed by a multitude of immune cells. Transduction of signals through Cxcr3 in the cells leads to activation of two major transcription factors Stat1 and NF-kB via PI3K/MAPK pathways. These ligands have been implicated in migration, metastasis, and tumour metabolism. Cxcl10/Cxcr3 interactions particularly drives Th1 polarization while Cxcl11/Cxcr3 induces immunotolerance state in T cells (Li, Han et al. 2022). Further, higher affinity of Cxcl11 to bind with Cxc3 compared to Cxcl10 also suggests potential to restrain inflammation (Zohar, Wildbaum et al. 2014, Li, Han et al. 2022). Ccr1 is a chemokine receptor expressed by many immune cells which leads to activation and accumulation of cells in allergic airway inflammation. Deficiency of Ccr1 in mouse models of Asthma has indicated its role in regulating the levels of Th2 chemokines (Feng, Ju et al. 2020). Ccr2 is involved in promotion of Th1 immune responses and is implicated in various inflammatory diseases. It is also an important receptor for mediation of monocyte recruitment and is required for pathogenesis of T cell- and macrophage- mediated inflammatory diseases (Izikson, Klein et al. 2000). Ccr3 plays critical role in allergic inflammation via the PI3K/AKT pathway and loss of Ccr3 gene has been implicated in reduction of eosinophilic inflammation and Th2 immune responses (Yuan, Liu et al. 2022). Ccr5 is a marker of Th1 cells, and its abrogation can alter Th2 induced inflammatory response with studies demonstrating its role in development of OVA-induced airway hyperresponsiveness (AHR) in a mouse model (Fuchimoto, Kanehiro et al. 2011). Il1r2 is associated with many aspects of pro-inflammatory Il-1 signalling, it competes with other Il1 receptors (Il1r1, Il1r3) for interaction with its ligands. Additionally, Il1r2 can sequester Il-1

ligands by its soluble form created by its enzymatic cleavage or alternative splicing (Supino, Minute et al. 2022). Ifng is an important regulator of both innate and adaptive immunity and is vital for regulation of inflammation. While Ifng is known as a major pro- inflammatory cytokine produced by a range of immune cells it has also been studied for its induction of tolerogenic characteristics in dendritic cells. In Tregs, Ifng plays intricate part in its regulatory functions (Rozman and Svajger 2018). Ifng signalling via Ifngr2 also causes switching from Th2 to Th1 phenotype as Ifng has growth inhibitory and proapoptotic effects on Th2 cells (Schroder, Hertzog et al. 2004). IL-27 is a cytokine which influences both innate and adaptive immune responses and has both pro- and anti-inflammatory roles. IL-27 is also crucial in maintaining the balance of Th1 and Th2 immunity. The role of IL-27 is dependent on the cells expressing and the specific disease environments (Povroznik and Robinson 2020). Ebi3 is a subunit of Il27 which has potential of inducing proliferative response, Th1 polarization and Ifng production in naïve $Cd4^+$ T cells (Pflanz, Timans et al. 2002).

Regulators of feedback loops – allergy targets

The key regulators and targets for allergy were identified on the basis of the number of pro-inflammatory feedback loops the genes were involved in. We also took into consideration their involvement in anti-inflammatory feedback loops. Allergy targets were selected as such due to their predominant involvement in pro-inflammatory loops but not anti-inflammatory loops. TNFR1 is a cognate receptor for TNF-alpha, a cytokine produced by a range of cell types and associated with most inflammatory diseases. Anti-TNF therapies have been successful in diseases like rheumatoid arthritis. Indeed, genetic inhibition of TNFR1 leads to blockade of TNF-a induced inflammation which reduces allergen-induced response (Sousa Garcia, Chen et al. 2017). Further, therapy based on local inhibition of TNFR1 has been established against allergic diseases by restricting allergen or antigen presentation (European patent – EP2746396A1). Klrd1(CD94) is a surface marker expressed as heterodimer CD94 with different isoforms of NKG2. CD94/NKG2 regulate the effector functions and cell survival of NK and $CD8^+$ T cells. CD94/NKG2 interacts with its ligand HLA-E in humans and its homolog Qa-1 in mouse (Gunturi, Berg et al. 2004). CD94 has been studied in tumor resistance (Kamiya, Seow et al. 2019) and NKG2A immune checkpoint inhibition has been used in development of cancer immunotherapy (van Hall, Andre et al. 2019). Interestingly for controlling inflammatory responses, CD94 has been shown key in

supressing CD4 effector T cells via CD8 T regulatory cells (Kim and Cantor 2011). Fn1 (Fibronectin) is an extracellular matrix protein which plays key role in communication between intra and extra cellular environment and thus regulates cell behaviour. It has also been involved in systemic immune responses and has been studied for its role in eradication of pathogens (Sandig, McDonald et al. 2009). Moreover, it plays a critical role in development of fibrosis in heart, kidneys etc including during the remodelling of airways in allergen induced rhinitis and thus has been recognised as an interesting target for fibrosis. (Hirshoren, Kohan et al. 2013). Lpl (Lipoprotein lipase) is involved in the removal of lipoproteins and triglycerides from circulation and has been implicated in immunomodulatory role in allergic and skin diseases (Trakaki and Marsche 2020). Further, LRP1 (Low-density lipoprotein receptor–related protein 1), a cognate receptor of Lpl, plays vital role as a negative regulator for DC-mediated adaptive immune response in HDM induced airway inflammation (Mishra, Yao et al. 2018). ADAM17 also known as TACE cleaves cell surface proteins like cytokines, cytokine receptors and adhesion proteins. Adam17 governs various pro- and anti-inflammatory responses required for appropriate response to damage and stress. The enzymatic function of Adam17 has been studied as a potential target for cancer immunotherapy (Wu, Mishra et al. 2019) and Th2 mediated diseases like allergies (Chen, Cheng et al. 2021). Matrix metalloproteinases (MMPs) are a family of enzymes involved in many physiological and pathological processes. MMP9 is involved in cell trafficking and inflammation by interacting with ECM proteins. Mmp9 protein levels activity has shown significant association with severity and inflammation in allergic asthma patients (Belleguic, Corbel et al. 2002). Inhibition of MMP9 has been demonstrated to decrease lymphatic inflammation and infiltration in allergen induced Asthma (Cataldo, Tournoy et al. 2002).

Regulators of feedback loops – allergen immunotherapy targets

Similarly, immunotherapy targets were selected based on their involvement in anti-inflammatory loops. These anti-inflammatory loops are responsible for countering inflammation in tissues. The anti-inflammatory loops are induced and sustained as part of inflammatory signals to the cells. Thus, in order to make sure that these targets are able to perturb tolerance we also take into consideration the pro-inflammatory feedback loops these genes are involved.

Sphingosine-1-phosphate (S1p) together with its receptor (S1pr1) are involved in lymphocyte trafficking, hence able to modulate immune responses (Syed, Raue et al. 2019). The

S1PR1-KLF2 axis also regulates migration of T cell out of the lymph nodes (Tindemans, van Schoonhoven et al. 2020). Further, in models of colitis S1PR1 signalling has been studied to govern the number and function of tumour associated T cell and Tregs (Priceman, Shen et al. 2014). IL-1 pathway is involved in the generation of inflammation, angiogenesis, and haematopoiesis. Inflammation mediated through IL-1 contributes to many diseases like rheumatoid arthritis, type 2 diabetes, systolic heart failure and pustular psoriasis and inhibition of IL-1 signalling is considered a major target for development of therapies (Peters, Joesting et al. 2013). IL-1RA is the main cognate receptor for IL-1 ligand. Il1r2 functions as a receptor antagonist to IL-1RA acting as a decoy receptor for IL-1 signal transduction. Similarly, Il1rn modulates the activity of IL-1 thus acting as a potent anti-inflammatory molecule and the equilibrium between IL-1 and Ilrn levels in local tissues influences the localized inflammatory effect (Godwin, Reeder et al. 2019, Tahtinen, Tong et al. 2022). Sortilin (Sort1) acts as receptor, co-receptor and trafficking regulator mediating immune responses via IL-6 signalling. Sort1 regulates secretion of cytokines involved in cytotoxicity and inflammation thus effecting immune functions across various cells. Sortilin expression has been implicated in Th1 response, antigen clearance and graft tolerance. Further, the absence of Sortilin leads to defect in induction of macrophages and T cells leading to aberrant immune response against infections (Talbot, Saada et al. 2018). Leukemia inhibitory factor receptor (LIFR) is part of IL-6 receptor family which is an heterodimer formed of LIFRβ and gp130 (Il6st) as its subunits. Studies have shown that expression of this receptor is upregulated in circulating immune cells of untreated MS patients (Janssens, Van den Haute et al. 2015). Through its interaction with Lif, this receptor is involved in proliferation of T cells and induction of tolerance by Tregs (Metcalfe 2011). Il18r1 is a subunit of the Il18r receptor which is a member of the Il1 receptor family. Its cognate ligand Il18 along with Il18r have been implicated in several inflammatory diseases like Asthma by regulating both innate and adaptive immunity (Zhu, Whyte et al. 2008) plays important functions in the pathogenesis of asthma. The loss of Il18r1 from Treg cells leads to failure in controlling T cell mediated colitis in animal models (Harrison, Srinivasan et al. 2015). Further, reports have noted that IL-18 itself did not induce the differentiation of the Il18r1+ Tregs suggesting a prior polarization of the cells (Alvarez, Al-Aubodah et al. 2020).

Our study predicts key regulators of allergic inflammation and tolerance induction/maintenance. The cell-cell communication networks of two states, immune dysregulation and immune tolerance, were re-constructed utilizing our previously published tool.

The tool is then extended to complete the intracellular signalling and reconstruct the complete cell-cell communication network. This completed cell-cell communication network can then be used to identify sustained feedback loops in each of the analysed conditions. While feedback loops have been studied for small networks with few nodes, larger networks pose an exponentially larger problem. Direct methods of feedback loop identification in these larger networks are both computationally expensive and have high time complexity as the number of possible feedback loops explodes exponentially based on connectivity of the network exponentially as the number of nodes and edges expand. We utilized a novel algorithm for detection of feedback loops in the communication networks between the cells. This algorithm is able to reduce both the computational time and memory usage by breaking down the larger network into smaller parts. The algorithm finds bridges between the clusters and therefore decreases the interaction and space that needs to be explored to identify all possible feedback loops in the networks. Topological analysis of the feedback loops allows the prediction of its regulators. Integrating biological information with the key regulators in these feedback loops, we predicted genes which can be inhibited to suppress the allergic inflammation or establishment of tolerance through AIT.

We predicted targets for inhibition of both allergic inflammation and tolerance by performing comparisons of Allergic vs Control and CpG-AIT vs Allergic respectively. The top targets predicted for allergy are TNFR1, Klrd1, Fn1, Lpl/Lrp1, Adam17 and Mmp9. All these genes have been previously identified to play crucial roles in inflammation. TNFR1 has been established as a therapeutic target for allergy, while the other predicted genes form a novel set of targets that could be explored for treatment of allergic diseases. Similarly, for inhibiting tolerance we predicted – S1pr1, Il1r2/ Il1rn, Lifr/Il6st, Il18r1. These genes are involved in induction and maintenance of tolerance by either generation of subsequent anti-inflammatory molecules or quenching pro-inflammatory molecules. While these genes have been studied for their roles in inflammation and for their expression on various cell types, they remain unexplored for tolerance induction and maintenance in allergy. Therefore, they have the potential be developed for novel therapeutic approaches. In addition, our results support the validity of our algorithm in identifying known and predicting new key nodes of cell-cell interaction in feedback loops in immune dysregulated environments as well as intolerance-inducing therapies. We expect this and similar tools to be used in future studies to screen available and newly produced single cell datasets to

better understand molecular disease pathophysiology and identify the key pathways to be modulated with a translational aim.

Figure Legends

Figure 1. Schematic diagram of mouse model and single-cell RNA sequencing experiment. BALB/c mice were subjected the sensitization on days 0, 14, and 28 followed by CpG-AIT on days 42, 56 and 60, finally allergen challenge was performed on days 83, 84 and 85. The indicated cell tyoes sorted from the Peritoneal cavity and/or the Spleen of the three groups mice.

Figure 2. (A) UMAP plot with annotated cell clusters (B) Dotplot of marker genes expression across all cell populations (C) Dotplot for expression of genes Foxp3, Ncr1 and Gata3 in cell populations.

Figure 3. Setmaps for all unique L-R interactions across the conditions – Control, Allergic and CpG-AIT for (A) t1 PC (B) t1 Spleen.

Figure 4. Scaled gene expression of allergy targets expressed in Spleen (A) Tnfrsf1a (B) Klrd1 (C) Fn1 (D) Lpl (E) Adam17 (F) Lrp1 (G) Mmp9

Figure 5. Scaled gene expression of immunotherapy targets expressed in Spleen (A) S1pr1 (B) Il1r2 (C) Il1rn (D) Il6st (E) Lifr (F) Il18r1.

Supplementary Figure Legends

Supplementary Figure 1. (A) Distribution of number of genes and transcripts expressed by the cells. (B) Correlation plot between number of genes (Y-axis) and transcripts expressed (X-axis) by all cells in the data. (C) Distribution of number of genes (Left) and transcripts expressed (Right) for all the 384-well sequenced plates (X-axis).

Supplementary Figure 2. Number of cells for all the cell populations across all timepoints and conditions in (A) Peritoneal cavity (PC) (B) Spleen.

Supplementary Figure 3. Setmaps for all unique L-R interactions across the conditions – Control, Allergic and CpG-AIT for (A) t2 PC (B) t2 Spleen (C) t3 PC (D) t3 Spleen.

Supplementary Figure 4. Scaled gene expression of allergy targets expressed in PC (A) Tnfrsf1a (B) Fn1 (C) Lpl (D) Klrd1 (E) Adam17

Supplementary Figure 5. Scaled gene expression of immunotherapy targets expressed in PC (A) Il1r2 (B) Il1rn (C) S1pr1.

References

Akdis, C. A. and M. Akdis (2015). "Advances in allergen immunotherapy: aiming for complete tolerance to allergens." Sci Transl Med 7(280): 280ps286.

Alvarez, F., T. A. Al-Aubodah, Y. H. Yang and C. A. Piccirillo (2020). "Mechanisms of TREG cell adaptation to inflammation." J Leukoc Biol 108(2): 559-571.

Belleguic, C., M. Corbel, N. Germain, H. Lena, E. Boichot, P. H. Delaval and V. Lagente (2002). "Increased release of matrix metalloproteinase-9 in the plasma of acute severe asthmatic patients." Clin Exp Allergy 32(2): 217-223.

Bluestone, J. A. (2011). "Mechanisms of tolerance." Immunol Rev 241(1): 5-19.

Borghi, A., L. Verstrepen and R. Beyaert (2016). "TRAF2 multitasking in TNF receptor-induced signaling to NF-kappaB, MAP kinases and cell death." Biochem Pharmacol 116: 1-10.

Brandum, E. P., A. S. Jorgensen, M. M. Rosenkilde and G. M. Hjorto (2021). "Dendritic Cells and CCR7 Expression: An Important Factor for Autoimmune Diseases, Chronic Inflammation, and Cancer." Int J Mol Sci 22(15).

Butler, A., P. Hoffman, P. Smibert, E. Papalexi and R. Satija (2018). "Integrating single-cell transcriptomic data across different conditions, technologies, and species." Nat Biotechnol 36(5): 411-420.

Cai, X., R. Chen, K. Ma, F. Wang, Y. Zhou, Y. Wang and T. Jiang (2020). "Identification of the CXCL12–CXCR4/CXCR7 axis as a potential therapeutic target for immunomodulating macrophage polarization and foreign body response to implanted biomaterials." Applied Materials Today 18.

Cataldo, D. D., K. G. Tournoy, K. Vermaelen, C. Munaut, J. M. Foidart, R. Louis, A. Noel and R. A. Pauwels (2002). "Matrix metalloproteinase-9 deficiency impairs cellular infiltration and bronchial hyperresponsiveness during allergen-induced airway inflammation." Am J Pathol 161(2): 491-498.

Chen, J. Y., W. H. Cheng, K. Y. Lee, H. P. Kuo, K. F. Chung, C. L. Chen, B. C. Chen and C. H. Lin (2021). "Abnormal ADAM17 expression causes airway fibrosis in chronic obstructive asthma." Biomed Pharmacother 140: 111701.

Chung, A. H., L. La Grenade and L. M. Harinstein (2019). "Pruritus after discontinuation of cetirizine." Ther Adv Drug Saf 10: 2042098619859996.

Di Mitri, D., M. Mirenda, J. Vasilevska, A. Calcinotto, N. Delaleu, A. Revandkar, V. Gil, G. Boysen, M. Losa, S. Mosole, E. Pasquini, R. D'Antuono, M. Masetti, E. Zagato, G. Chiorino, P. Ostano, A. Rinaldi, L. Gnetti, M. Graupera, A. R. Martins Figueiredo Fonseca, R. Pereira Mestre, D. Waugh, S. Barry, J. De Bono and A. Alimonti (2019). "Re-education of Tumor-Associated Macrophages by CXCR2 Blockade Drives Senescence and Tumor Inhibition in Advanced Prostate Cancer." Cell Rep 28(8): 2156-2168 e2155.

Dong, B., C. Wu, L. Huang and Y. Qi (2021). "Macrophage-Related SPP1 as a Potential Biomarker for Early Lymph Node Metastasis in Lung Adenocarcinoma." Front Cell Dev Biol 9: 739358.

Feng, S., L. Ju, Z. Shao, M. Grzanna, L. Jia and M. Liu (2020). "Therapeutic Effect of C-C Chemokine Receptor Type 1 (CCR1) Antagonist BX471 on Allergic Rhinitis." J Inflamm Res 13: 343-356.

Fuchimoto, Y., A. Kanehiro, N. Miyahara, H. Koga, G. Ikeda, K. Waseda, Y. Tanimoto, S. Ueha, M. Kataoka, E. W. Gelfand and M. Tanimoto (2011). "Requirement for chemokine receptor 5 in the development of allergen-induced airway hyperresponsiveness and inflammation." Am J Respir Cell Mol Biol 45(6): 1248-1255.

Gaddis, D. E., L. E. Padgett, R. Wu and C. C. Hedrick (2019). "Neuropilin-1 Expression on CD4 T Cells Is Atherogenic and Facilitates T Cell Migration to the Aorta in Atherosclerosis." J Immunol 203(12): 3237-3246.

Garcia, K. C. (2019). "Dual Arms of Adaptive Immunity: Division of Labor and Collaboration between B and T Cells." Cell 179(1): 3-7.

Geginat, J., M. Paroni, S. Maglie, J. S. Alfen, I. Kastirr, P. Gruarin, M. De Simone, M. Pagani and S. Abrignani (2014). "Plasticity of human CD4 T cell subsets." Front Immunol 5: 630.

Giardino, G., V. Gallo, R. Prencipe, G. Gaudino, R. Romano, M. De Cataldis, P. Lorello, L. Palamaro, C. Di Giacomo, D. Capalbo, E. Cirillo, R. D'Assante and C. Pignata (2016). "Unbalanced Immune System: Immunodeficiencies and Autoimmunity." Front Pediatr 4: 107.

Godwin, M. S., K. M. Reeder, J. M. Garth, J. P. Blackburn, M. Jones, Z. Yu, S. Matalon, A. T. Hastie, D. A. Meyers and C. Steele (2019). "IL-1RA regulates immunopathogenesis during fungal-associated allergic airway inflammation." JCI Insight 4(21).

Gunturi, A., R. E. Berg and J. Forman (2004). "The role of CD94/NKG2 in innate and adaptive immunity." Immunol Res 30(1): 29-34.

Hao, Y., S. Hao, E. Andersen-Nissen, W. M. Mauck, 3rd, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert and R. Satija (2021). "Integrated analysis of multimodal single-cell data." Cell 184(13): 3573-3587 e3529.

Harrison, O. J., N. Srinivasan, J. Pott, C. Schiering, T. Krausgruber, N. E. Ilott and K. J. Maloy (2015). "Epithelial-derived IL-18 regulates Th17 cell differentiation and Foxp3(+) Treg cell function in the intestine." Mucosal Immunol 8(6): 1226-1236.

Hashimshony, T., N. Senderovich, G. Avital, A. Klochendler, Y. de Leeuw, L. Anavy, D. Gennert, S. Li, K. J. Livak, O. Rozenblatt-Rosen, Y. Dor, A. Regev and I. Yanai (2016). "CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq." Genome Biol 17: 77.

Heldin, C. H., B. Lu, R. Evans and J. S. Gutkind (2016). "Signals and Receptors." Cold Spring Harb Perspect Biol 8(4): a005900.

Hirshoren, N., M. Kohan, M. Assayag, T. Neuman, F. Vernea, A. Muro, R. Eliashar and N. Berkman (2013). "Extra domain-A fibronectin is necessary for the development of nasal remodeling in chronic allergen-induced rhinitis." Ann Allergy Asthma Immunol 110(5): 322-327.

Huang, M., L. Xu, J. Liu, P. Huang, Y. Tan and S. Chen (2022). "Cell-Cell Communication Alterations via Intercellular Signaling Pathways in Substantia Nigra of Parkinson's Disease." Front Aging Neurosci 14: 828457.

Ierodiakonou, D., M. A. Portelli, D. S. Postma, G. H. Koppelman, J. Gerritsen, N. H. Ten Hacken, W. Timens, H. M. Boezen, J. M. Vonk and I. Sayers (2016). "Urokinase plasminogen activator receptor polymorphisms and airway remodelling in asthma." Eur Respir J 47(5): 1568-1571.

Iwamura, C. and T. Nakayama (2018). "Role of CD1d- and MR1-Restricted T Cells in Asthma." Front Immunol 9: 1942.

Izikson, L., R. S. Klein, I. F. Charo, H. L. Weiner and A. D. Luster (2000). "Resistance to experimental autoimmune encephalomyelitis in mice lacking the CC chemokine receptor (CCR)2." J Exp Med 192(7): 1075-1080.

Jang, M. H., N. Sougawa, T. Tanaka, T. Hirata, T. Hiroi, K. Tohya, Z. Guo, E. Umemoto, Y. Ebisuno, B. G. Yang, J. Y. Seoh, M. Lipp, H. Kiyono and M. Miyasaka (2006). "CCR7 is critically important for migration of dendritic cells in intestinal lamina propria to mesenteric lymph nodes." J Immunol 176(2): 803-810.

Janssens, K., C. Van den Haute, V. Baekelandt, S. Lucas, J. van Horssen, V. Somers, B. Van Wijmeersch, P. Stinissen, J. J. Hendriks, H. Slaets and N. Hellings (2015). "Leukemia inhibitory factor tips the immune balance towards regulatory T cells in multiple sclerosis." Brain Behav Immun 45: 180-188.

Jung, S., I. Potapov, S. Chillara and A. Del Sol (2021). "Leveraging systems biology for predicting modulators of inflammation in patients with COVID-19." Sci Adv 7(6).

Jung, S., K. Singh and A. Del Sol (2021). "FunRes: resolving tissue-specific functional cell states based on a cell-cell communication network model." Brief Bioinform 22(4).

Kamiya, T., S. V. Seow, D. Wong, M. Robinson and D. Campana (2019). "Blocking expression of inhibitory receptor NKG2A overcomes tumor resistance to NK cells." J Clin Invest 129(5): 2094-2106.

Katoh, S. (2021). "Critical Involvement of CD44 in T Helper Type 2 Cell-Mediated Eosinophilic Airway Inflammation in a Mouse Model of Acute Asthma." Front Immunol 12: 811600.

Kim, H. J. and H. Cantor (2011). "Regulation of self-tolerance by Qa-1-restricted CD8(+) regulatory T cells." Semin Immunol 23(6): 446-452.

Kumar, S., Y. Jeong, M. U. Ashraf and Y. S. Bae (2019). "Dendritic Cell-Mediated Th2 Immunity and Immune Disorders." Int J Mol Sci 20(9).

Lambrecht, B. N. and H. Hammad (2014). "Allergens and the airway epithelium response: gateway to allergic sensitization." J Allergy Clin Immunol 134(3): 499-507.

Leonard, C., G. Montamat, C. Davril, O. Domingues, O. Hunewald, D. Revets, C. Guerin, S. Blank, J. Heckendorn, G. Jardon, F. Hentges and M. Ollert (2021). "Comprehensive mapping of immune tolerance yields a regulatory TNF receptor 2 signature in a murine model of successful Fel d 1-specific immunotherapy using high-dose CpG adjuvant." Allergy 76(7): 2153-2165.

Li, X., J. Peng, Y. Pang, S. Yu, X. Yu, P. Chen, W. Wang, W. Han, J. Zhang, Y. Yin and Y. Zhang (2015). "Identification of a FOXP3(+)CD3(+)CD56(+) population with immunosuppressive function in cancer tissues of human hepatocellular carcinoma." Sci Rep 5: 14757.

Li, Y., S. Han, B. Wu, C. Zhong, Y. Shi, C. Lv, L. Fu, Y. Zhang, Q. Lang, Z. Liang, Y. Yu and Y. Tian (2022). "CXCL11 Correlates with Immune Infiltration and Impacts Patient Immunotherapy Efficacy: A Pan-Cancer Analysis." Front Immunol 13: 951247.

Liu, M., S. Chen, A. Zhang, Q. Zheng and J. Fu (2021). "PLAUR as a Potential Biomarker Associated with Immune Infiltration in Bladder Urothelial Carcinoma." J Inflamm Res 14: 4629-4641.

Ma, W., Y. Liu, N. Ellison and J. Shen (2013). "Induction of C-X-C chemokine receptor type 7 (CXCR7) switches stromal cell-derived factor-1 (SDF-1) signaling and phagocytic activity in macrophages linked to atherosclerosis." J Biol Chem 288(22): 15481-15494.

Marogna, M., I. Spadolini, A. Massolo, G. W. Canonica and G. Passalacqua (2010). "Long-lasting effects of sublingual immunotherapy according to its duration: a 15-year prospective study." J Allergy Clin Immunol 126(5): 969-975.

Metcalfe, S. M. (2011). "LIF in the regulation of T-cell fate and as a potential therapeutic." Genes Immun 12(3): 157-168.

Mishra, A., X. Yao, A. Saxena, E. M. Gordon, M. Kaler, R. A. Cuento, A. V. Barochia, P. K. Dagur, J. P. McCoy, K. J. Keeran, K. R. Jeffries, X. Qu, Z. X. Yu and S. J. Levine (2018). "Low-density lipoprotein receptor-related protein 1 attenuates house dust mite-induced eosinophilic airway inflammation by suppressing dendritic cell-mediated adaptive immune responses." J Allergy Clin Immunol 142(4): 1066-1079 e1066.

Monteiro, M., C. F. Almeida, M. Caridade, J. C. Ribot, J. Duarte, A. Agua-Doce, I. Wollenberg, B. Silva-Santos and L. Graca (2010). "Identification of regulatory Foxp3+ invariant NKT cells induced by TGF-beta." J Immunol 185(4): 2157-2163.

Muraro, M. J., G. Dharmadhikari, D. Grun, N. Groen, T. Dielen, E. Jansen, L. van Gurp, M. A. Engelse, F. Carlotti, E. J. de Koning and A. van Oudenaarden (2016). "A Single-Cell Transcriptome Atlas of the Human Pancreas." Cell Syst 3(4): 385-394 e383.

Netea, M. G., J. Dominguez-Andres, L. B. Barreiro, T. Chavakis, M. Divangahi, E. Fuchs, L. A. B. Joosten, J. W. M. van der Meer, M. M. Mhlanga, W. J. M. Mulder, N. P. Riksen, A. Schlitzer, J. L. Schultze, C. Stabell Benn, J. C. Sun, R. J. Xavier and E. Latz (2020). "Defining trained immunity and its role in health and disease." Nat Rev Immunol 20(6): 375-388.

Overton, D. L. and T. L. Mastracci (2022). "Exocrine-Endocrine Crosstalk: The Influence of Pancreatic Cellular Communications on Organ Growth, Function and Disease." Front Endocrinol (Lausanne) 13: 904004.

Paget, C., T. Mallevaey, A. O. Speak, D. Torres, J. Fontaine, K. C. Sheehan, M. Capron, B. Ryffel, C. Faveeuw, M. Leite de Moraes, F. Platt and F. Trottein (2007). "Activation of invariant NKT cells by toll-like receptor 9-stimulated dendritic cells requires type I interferon and charged glycosphingolipids." Immunity 27(4): 597-609.

Palomares, O., M. Martin-Fontecha, R. Lauener, C. Traidl-Hoffmann, O. Cavkaytar, M. Akdis and C. A. Akdis (2014). "Regulatory T cells and immune regulation of allergic diseases: roles of IL-10 and TGF-beta." Genes Immun 15(8): 511-520.

Pawankar, R. (2014). "Allergic diseases and asthma: a global public health concern and a call to action." World Allergy Organ J 7(1): 12.

Peters, V. A., J. J. Joesting and G. G. Freund (2013). "IL-1 receptor 2 (IL-1R2) and its role in immune regulation." Brain Behav Immun 32: 1-8.

Pflanz, S., J. C. Timans, J. Cheung, R. Rosales, H. Kanzler, J. Gilbert, L. Hibbert, T. Churakova, M. Travis, E. Vaisberg, W. M. Blumenschein, J. D. Mattson, J. L. Wagner, W. To, S. Zurawski, T. K. McClanahan, D. M. Gorman, J. F. Bazan, R. de Waal Malefyt, D. Rennick and R. A. Kastelein (2002). "IL-27, a heterodimeric cytokine composed of EBI3 and p28 protein, induces proliferation of naive CD4+ T cells." Immunity 16(6): 779-790.

Povroznik, J. M. and C. M. Robinson (2020). "IL-27 regulation of innate immunity and control of microbial growth." Future Sci OA 6(7): FSO588.

Priceman, S. J., S. Shen, L. Wang, J. Deng, C. Yue, M. Kujawski and H. Yu (2014). "S1PR1 is crucial for accumulation of regulatory T cells in tumors via STAT3." Cell Rep 6(6): 992-999.

Rahman, A., A. Tiwari, J. Narula and T. Hickling (2018). "Importance of Feedback and Feedforward Loops to Adaptive Immune Response Modeling." CPT Pharmacometrics Syst Pharmacol 7(10): 621-628.

Ravichandran, S., A. Hartmann and A. Del Sol (2019). "SigHotSpotter: scRNA-seq-based computational tool to control cell subpopulation phenotypes for cellular rejuvenation strategies." Bioinformatics.

Rothenberg, M. E. (2003). "CD44--a sticky target for asthma." J Clin Invest 111(10): 1460-1462.

Rozman, P. and U. Svajger (2018). "The tolerogenic role of IFN-gamma." Cytokine Growth Factor Rev 41: 40-53.

Ruch, R. J. (2002). "Intercellular communication, homeostasis, and toxicology." Toxicol Sci 68(2): 265-266.

Sandig, H., J. McDonald, J. Gilmour, M. Arno, T. H. Lee and D. J. Cousins (2009). "Fibronectin is a TH1-specific molecule in human subjects." J Allergy Clin Immunol 124(3): 528-535, 535 e521-525.

Satija, R., J. A. Farrell, D. Gennert, A. F. Schier and A. Regev (2015). "Spatial reconstruction of single-cell gene expression data." Nat Biotechnol 33(5): 495-502.

Schroder, K., P. J. Hertzog, T. Ravasi and D. A. Hume (2004). "Interferon-gamma: an overview of signals, mechanisms and functions." J Leukoc Biol 75(2): 163-189.

Sjaastad, L. E., D. L. Owen, S. I. Tracy and M. A. Farrar (2021). "Phenotypic and Functional Diversity in Regulatory T Cells." Front Cell Dev Biol 9: 715901.

Sousa Garcia, D., M. Chen, A. K. Smith, P. R. Lazarini and A. P. Lane (2017). "Role of the type I tumor necrosis factor receptor in inflammation-associated olfactory dysfunction." Int Forum Allergy Rhinol 7(2): 160-168.

Stuart, T., A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, 3rd, Y. Hao, M. Stoeckius, P. Smibert and R. Satija (2019). "Comprehensive Integration of Single-Cell Data." Cell 177(7): 1888-1902 e1821.

Supino, D., L. Minute, A. Mariancini, F. Riva, E. Magrini and C. Garlanda (2022). "Negative Regulation of the IL-1 System by IL-1R2 and IL-1R8: Relevance in Pathophysiology and Disease." Front Immunol 13: 804641.

Swanson, C. L., R. Pelanda and R. M. Torres (2013). "Division of labor during primary humoral immunity." Immunol Res 55(1-3): 277-286.

Syed, S. N., R. Raue, A. Weigert, A. von Knethen and B. Brune (2019). "Macrophage S1PR1 Signaling Alters Angiogenesis and Lymphangiogenesis During Skin Inflammation." Cells 8(8).

Tahtinen, S., A. J. Tong, P. Himmels, J. Oh, A. Paler-Martinez, L. Kim, S. Wichner, Y. Oei, M. J. McCarron, E. C. Freund, Z. A. Amir, C. C. de la Cruz, B. Haley, C. Blanchette, J. M. Schartner, W. Ye, M. Yadav, U. Sahin, L. Delamarre and I. Mellman (2022). "IL-1 and IL-1ra are key regulators of the inflammatory response to RNA vaccines." Nat Immunol 23(4): 532-542.

Talbot, H., S. Saada, T. Naves, P. F. Gallet, A. L. Fauchais and M. O. Jauberteau (2018). "Regulatory Roles of Sortilin and SorLA in Immune-Related Processes." Front Pharmacol 9: 1507.

Tindemans, I., A. van Schoonhoven, A. KleinJan, M. J. de Bruijn, M. Lukkes, M. van Nimwegen, A. van den Branden, I. M. Bergen, O. B. Corneth, I. W. F. van, R. Stadhouders and R. W. Hendriks (2020). "Notch signaling licenses allergic airway inflammation by promoting Th2 cell lymph node egress." J Clin Invest 130(7): 3576-3591.

Trakaki, A. and G. Marsche (2020). "High-Density Lipoprotein (HDL) in Allergy and Skin Diseases: Focus on Immunomodulating Functions." Biomedicines 8(12).

Valls, P. O. and A. Esposito (2022). "Signalling dynamics, cell decisions, and homeostatic control in health and disease." Curr Opin Cell Biol 75: 102066.

van Hall, T., P. Andre, A. Horowitz, D. F. Ruan, L. Borst, R. Zerbib, E. Narni-Mancinelli, S. H. van der Burg and E. Vivier (2019). "Monalizumab: inhibiting the novel immune checkpoint NKG2A." J Immunother Cancer 7(1): 263.

Volpi, C., F. Fallarino, M. T. Pallotta, R. Bianchi, C. Vacca, M. L. Belladonna, C. Orabona, A. De Luca, L. Boon, L. Romani, U. Grohmann and P. Puccetti (2013). "High doses of CpG oligodeoxynucleotides stimulate a tolerogenic TLR9-TRIF pathway." Nat Commun 4: 1852.

Weidemuller, P., M. Kholmatov, E. Petsalaki and J. B. Zaugg (2021). "Transcription factors: Bridge between cell signaling and gene regulation." Proteomics 21(23-24): e2000034.

Wu, J., H. K. Mishra and B. Walcheck (2019). "Role of ADAM17 as a regulatory checkpoint of CD16A in NK cells and as a potential target for cancer immunotherapy." J Leukoc Biol 105(6): 1297-1303.

Xin, Y., P. Lyu, J. Jiang, F. Zhou, J. Wang, S. Blackshaw and J. Qian (2022). "LRLoop: a method to predict feedback loops in cell-cell communication." Bioinformatics 38(17): 4117-4126.

Yang, Y., M. S. Islam, Y. Hu and X. Chen (2021). "TNFR2: Role in Cancer Immunology and Immunotherapy." Immunotargets Ther 10: 103-122.

Yao, Y., C. L. Chen, D. Yu and Z. Liu (2021). "Roles of follicular helper and regulatory T cells in allergic diseases and allergen immunotherapy." Allergy 76(2): 456-470.

Yuan, J., Y. Liu, J. Yu, M. Dai, Y. Zhu, Y. Bao, H. Peng, K. Liu and X. Zhu (2022). "Gene knockdown of CCR3 reduces eosinophilic inflammation and the Th2 immune response by inhibiting the PI3K/AKT pathway in allergic rhinitis mice." Sci Rep 12(1): 5411.

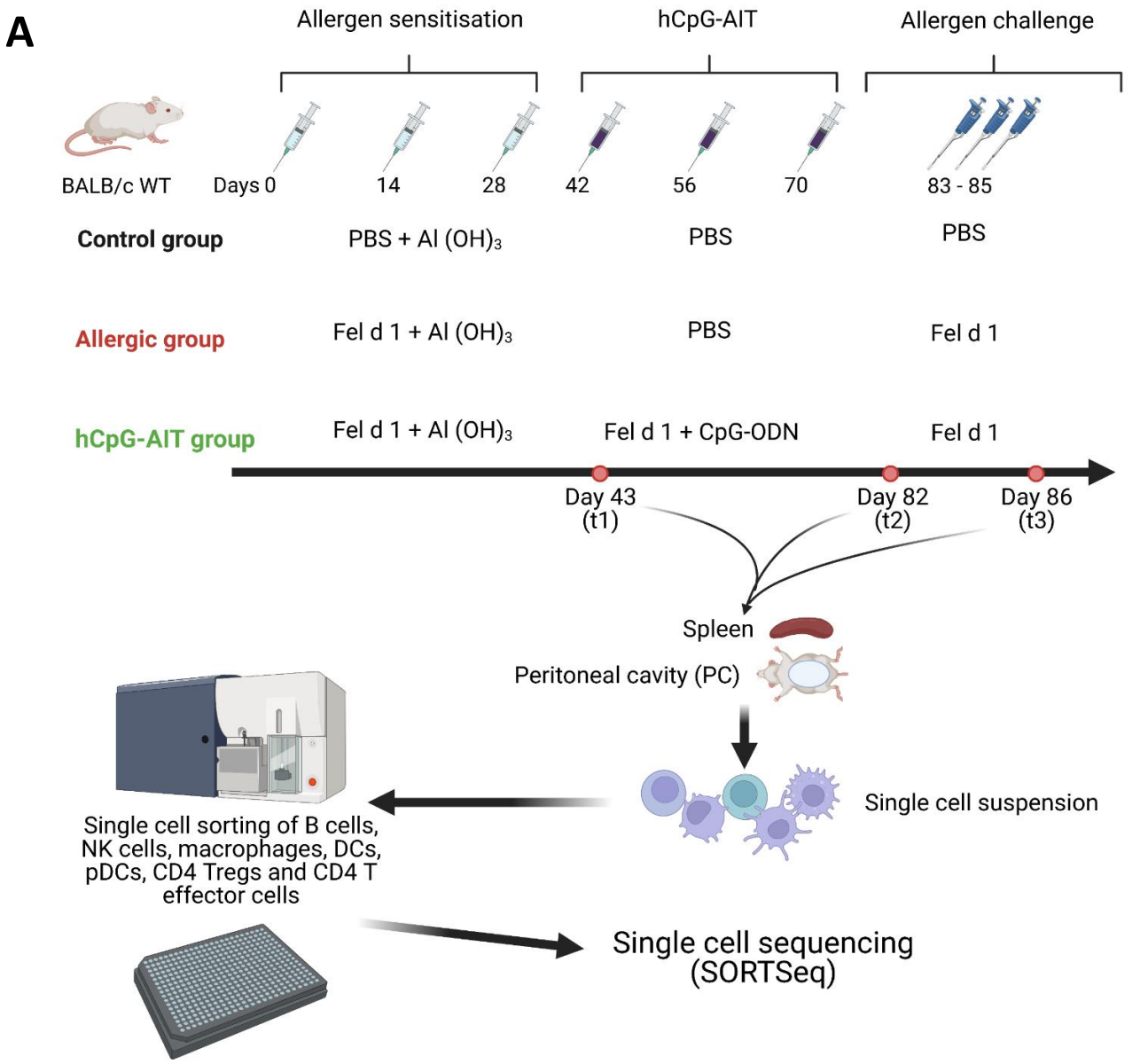
Zhang, J., Y. Zhang, S. Xin, M. Wu, Y. Zhang and L. Sun (2020). "CXCR7 suppression modulates macrophage phenotype and function to ameliorate post-myocardial infarction injury." Inflamm Res 69(5): 523-532.

Zhou, L., F. Fu, Y. Wang and L. Yang (2022). "Interlocked feedback loops balance the adaptive immune response." Math Biosci Eng 19(4): 4084-4100.

Zhu, G., M. K. Whyte, J. Vestbo, K. Carlsen, K. H. Carlsen, W. Lenney, M. Silverman, P. Helms and S. G. Pillai (2008). "Interleukin 18 receptor 1 gene polymorphisms are associated with asthma." Eur J Hum Genet 16(9): 1083-1090.

Zohar, Y., G. Wildbaum, R. Novak, A. L. Salzman, M. Thelen, R. Alon, Y. Barsheshet, C. L. Karp and N. Karin (2014). "CXCL11-dependent induction of FOXP3-negative regulatory T cells suppresses autoimmune encephalomyelitis." J Clin Invest 124(5): 2009-2022.

Figure 1


A
Allergen sensitisation
hCpG-AIT
Allergen challenge
BALB/c WT
Days 0
14
28
42
56
70
83 - 85
Control group
PBS + Al (OH)3
PBS
PBS
Allergic group
Fel d 1 + Al (OH)3
PBS
Fel d 1
hCpG-AIT group
Fel d 1 + Al (OH)3
Fel d 1 + CpG-ODN
Fel d 1
Day 43 (t1)
Day 82 (t2)
Day 86 (t3)
Spleen
Peritoneal cavity (PC)
Single cell suspension
Single cell sorting of B cells, NK cells, macrophages, DCs, pDCs, CD4 Tregs and CD4 T effector cells
Single cell sequencing (SORTSeq)

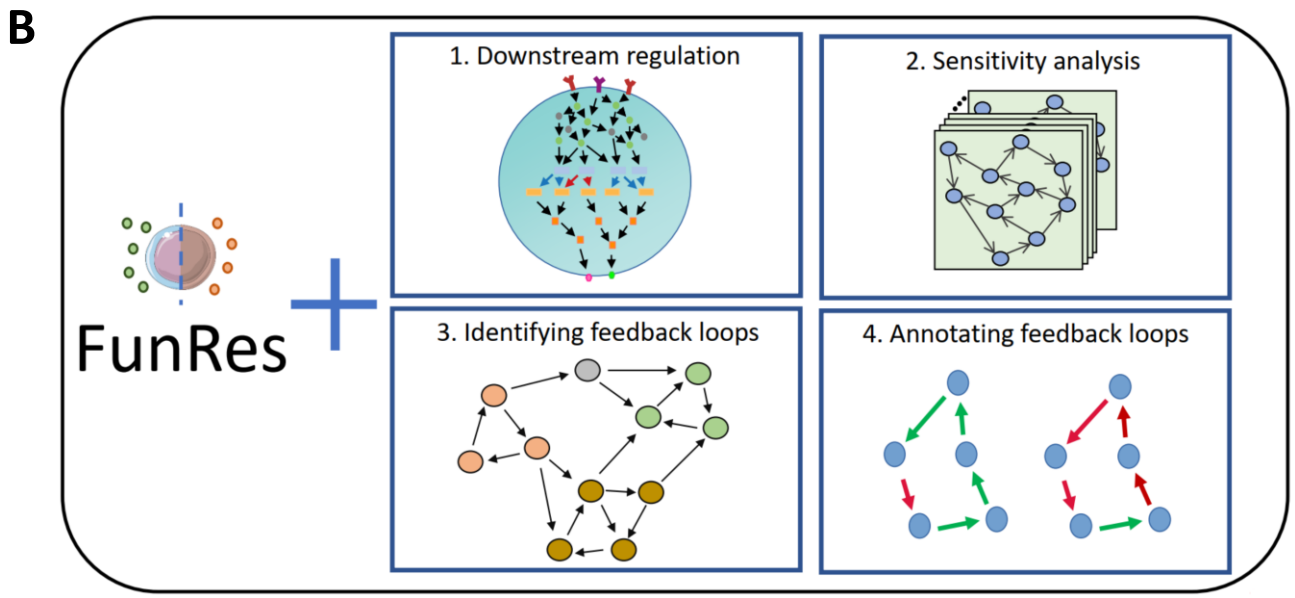
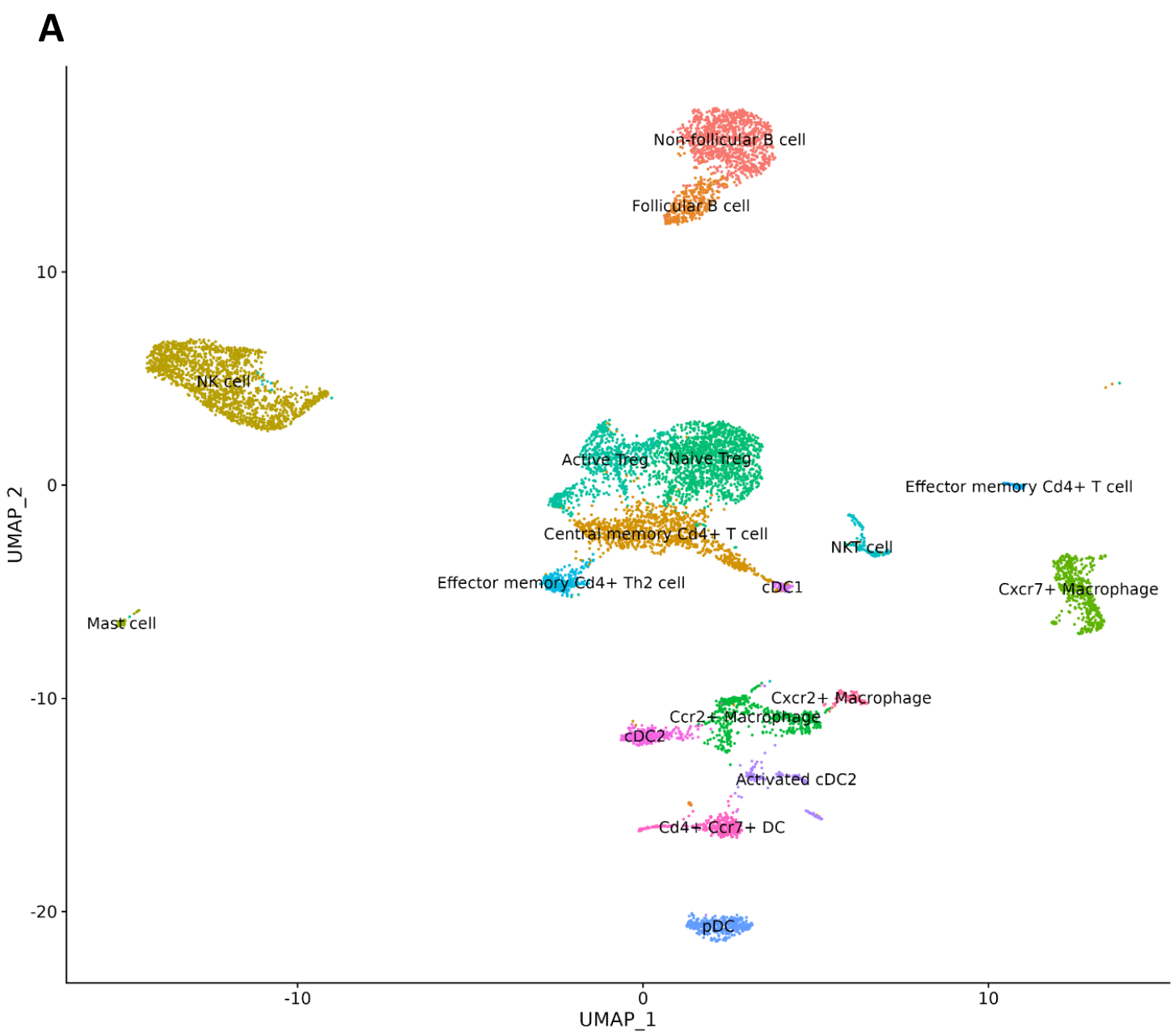
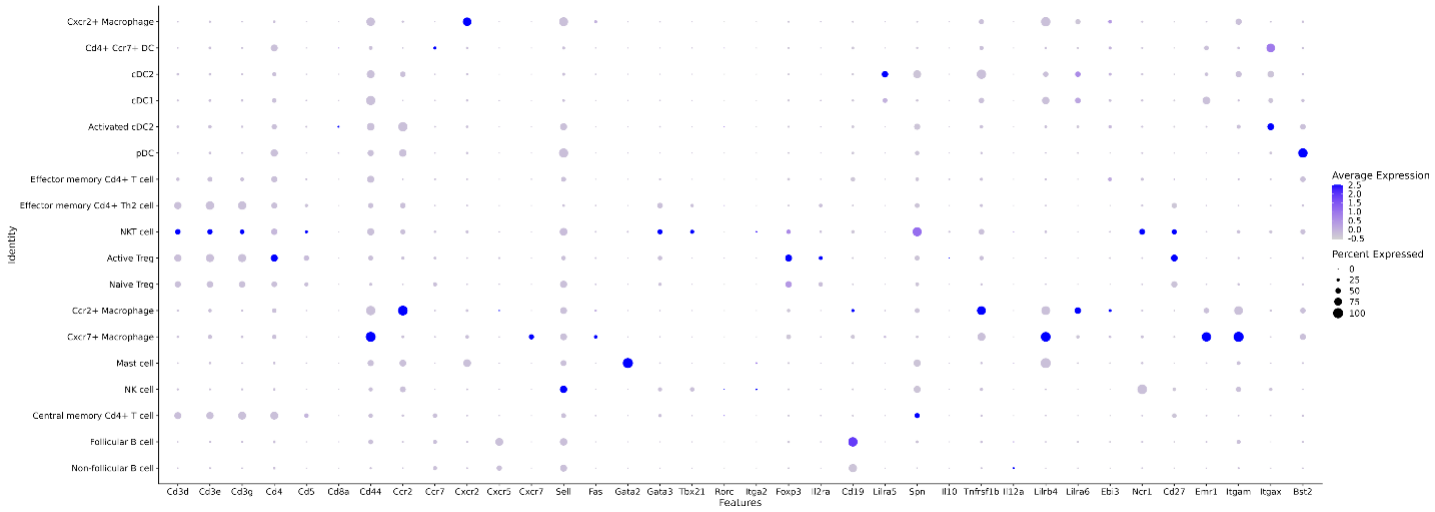

B
FunRes
1. Downstream regulation
2. Sensitivity analysis
3. Identifying feedback loops
4. Annotating feedback loops

Figure 2


A
Non-follicular B cell
Follicular B cell
NK cell
Active Treg
Naive Treg
Effector memory Cd4+ T cell
Central memory Cd4+ T cell
NKT cell
Effector memory Cd4+ Th2 cell
cDC1
Cxcr7+ Macrophage
Mast cell
Cxcr2+ Macrophage
Ccr2+ Macrophage
cDC2
Activated cDC2
Cd4+ Ccr7+ DC
pDC
10
0
-10
-20
UMAP_2
-10
0
10
UMAP_1

B


Cxcr2+ Macrophage
Cd4+ Ccr7+ DC
cDC2
cDC1
Activated cDC2
pDC
Effector memory Cd4+ T cell
Effector memory Cd4+ Th2 cell
NKT cell
Active Treg
Naive Treg
Ccr2+ Macrophage
Cxcr7+ Macrophage
Mast cell
NK cell
Central memory Cd4+ T cell
Follicular B cell
Non-follicular B cell
Identity
Cd3d Cd3e Cd3g Cd4 Cd5 Cd8a Cd44 Ccr2 Ccr7 Cxcr2 Cxcr5 Cxcr7 Sell Fas Gata2 Gata3 Tbx21 Rorc Itga2 Foxp3 Il2ra Cd19 Lilra5 Spn Il10 Tnfrsf1b Il12a Lilrb4 Lilra6 Ebi3 Ncr1 Cd27 Emr1 Itgam Itgax Bst2
Features
Average Expression
2.5
2.0
1.5
1.0
0.5
0.0
Percent Expressed
0
25
50
75
100

C

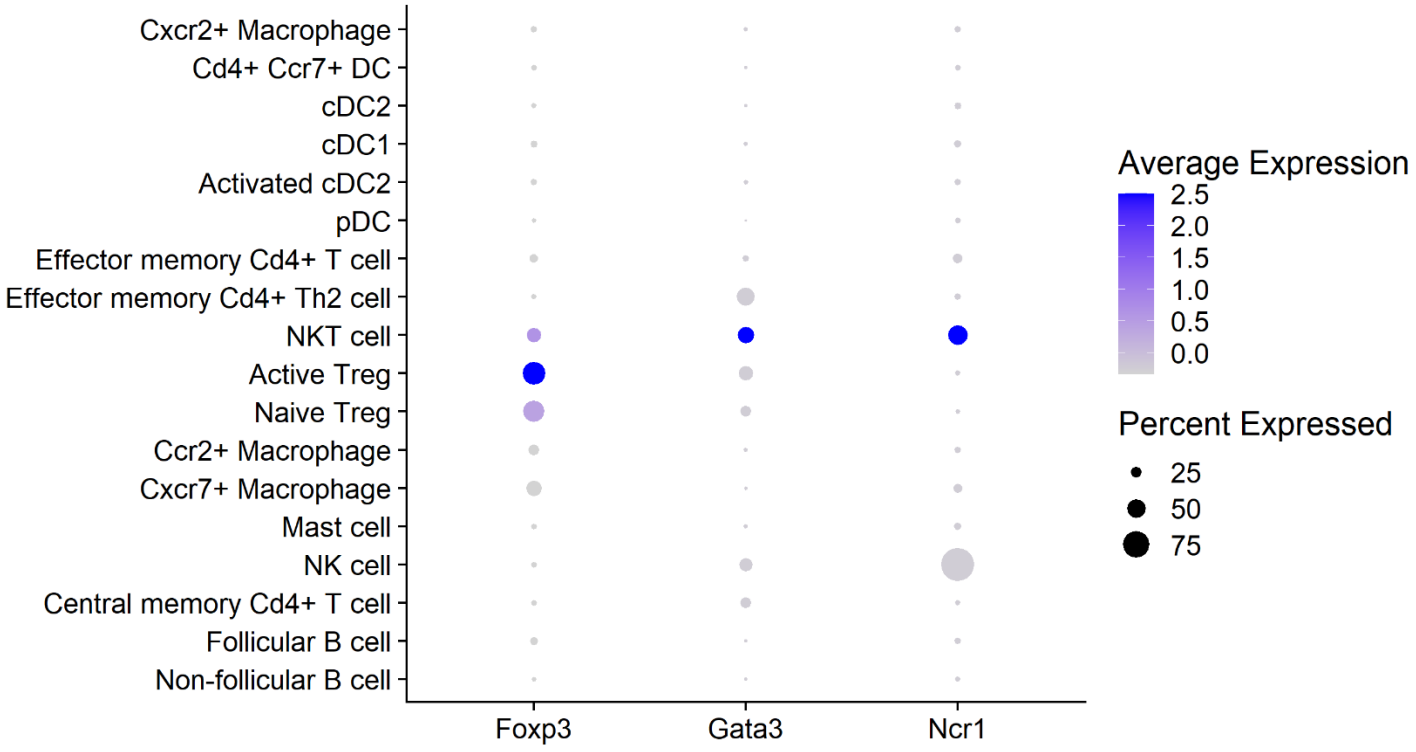
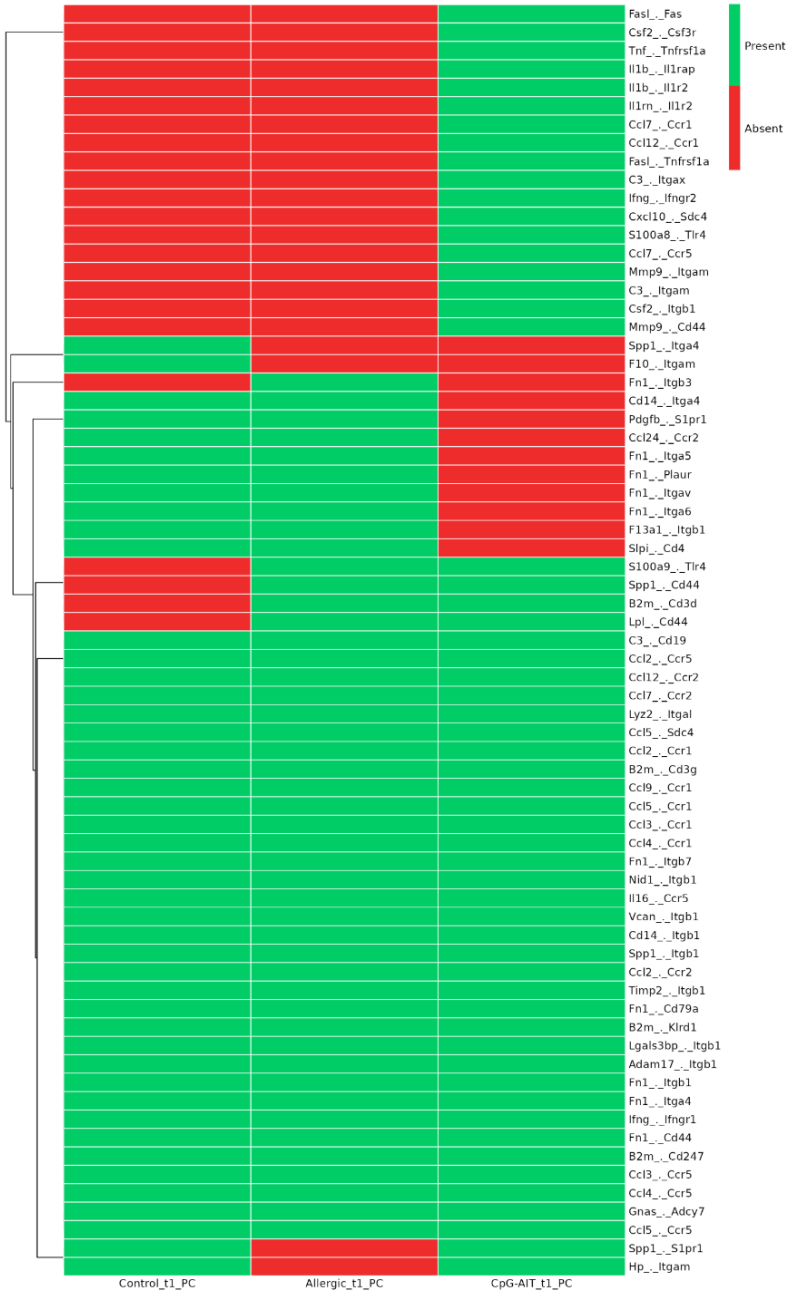

Cxcr2+ Macrophage
Cd4+ Ccr7+ DC
cDC2
cDC1
Activated cDC2
pDC
Effector memory Cd4+ T cell
Effector memory Cd4+ Th2 cell
NKT cell
Active Treg
Naive Treg
Ccr2+ Macrophage
Cxcr7+ Macrophage
Mast cell
NK cell
Central memory Cd4+ T cell
Follicular B cell
Non-follicular B cell
Foxp3
Gata3
Ncr1
Average Expression
2.5
2.0
1.5
1.0
0.5
0.0
Percent Expressed
25
50
75

Figure 3

A


Present
Absent
Fasl_._Fas
Csf2_._Csf3r
Tnf_._Tnfrsf1a
Il1b_._Il1rap
Il1b_._Il1r2
Il1rn_._Il1r2
Ccl7_._Ccr1
Ccl12_._Ccr1
Fasl_._Tnfrsf1a
C3_._Itgax
Ifng_._Ifngr2
Cxcl10_._Sdc4
S100a8_._Tlr4
Ccl7_._Ccr5
Mmp9_._Itgam
C3_._Itgam
Csf2_._Itgb1
Mmp9_._Cd44
Spp1_._Itga4
F10_._Itgam
Fn1_._Itgb3
Cd14_._Itga4
Pdgfb_._S1pr1
Ccl24_._Ccr2
Fn1_._Itga5
Fn1_._Plaur
Fn1_._Itgav
Fn1_._Itga6
F13a1_._Itgb1
Slpi_._Cd4
S100a9_._Tlr4
Spp1_._Cd44
B2m_._Cd3d
Lpl_._Cd44
C3_._Cd19
Ccl2_._Ccr5
Ccl12_._Ccr2
Ccl7_._Ccr2
Lyz2_._Itgal
Ccl5_._Sdc4
Ccl2_._Ccr1
B2m_._Cd3g
Ccl9_._Ccr1
Ccl5_._Ccr1
Ccl3_._Ccr1
Ccl4_._Ccr1
Fn1_._Itgb7
Nid1_._Itgb1
Il16_._Ccr5
Vcan_._Itgb1
Cd14_._Itgb1
Spp1_._Itgb1
Ccl2_._Ccr2
Timp2_._Itgb1
Fn1_._Cd79a
B2m_._Klrd1
Lgals3bp_._Itgb1
Adam17_._Itgb1
Fn1_._Itgb1
Fn1_._Itga4
Ifng_._Ifngr1
Fn1_._Cd44
B2m_._Cd247
Ccl3_._Ccr5
Ccl4_._Ccr5
Gnas_._Adcy7
Ccl5_._Ccr5
Spp1_._S1pr1
Hp_._Itgam
Control_t1_PC
Allergic_t1_PC
CpG-AIT_t1_PC

B


Present
Absent
Lama5_._Sdc1
Lpl_._Sdc1
Fn1_._Cd79a
Ccl5_._Sdc1
Vcan_._Itga4
Il1rn_._Il1r2
Saa3_._Fpr1
Adam17_._Notch1
Anxa1_._Fpr1
Gpi1_._Amfr
Cxcl9_._Cxcr3
Cxcl11_._Cxcr3
Cxcl10_._Cxcr3
Reln_._Lrp8
Cxcl10_._Sdc4
Apoe_._Lrp8
Reln_._Itgb1
Mmp9_._Lrp1
S100a8_._Tlr4
S100a9_._Tlr4
Fasl_._Tnfrsf1a
Camp_._Igf1r
Mmp9_._Cd44
Il15_._Il2rg
Cd14_._Itgb1
C3_._Itgam
Mmp9_._Itgam
Hp_._Itgam
Ccl24_._Ccr3
Gas6_._Axl
Mdk_._Lrp1
Pros1_._Axl
Lpl_._Lrp1
Ccl9_._Ccr3
Pdgfb_._Lrp1
Ccl4_._Ccr3
Ccl3_._Ccr3
Ccl5_._Ccr3
Ccl24_._Ccr2
Cd14a1_._Cd44
Il12a_._Cd28
Apoe_._Lrp5
Adam23_._Itgb3
Fn1_._Itgb3
Vegfb_._Nrp1
Lama5_._Itgb1
Adam23_._Itga5
Ccl9_._Ccr1
Ccl4_._Ccr1
Ccl3_._Ccr1
Ccl5_._Ccr1
Il4_._Il2rg
Osm_._Lifr
Fn1_._Itgav
Ebi3_._Il27ra
Tnf_._Traf2
Timp2_._Itgb1
Pdgfb_._S1pr1
Adam17_._Itgb1
Mdk_._Itga4
C1qb_._Lrp1
Apoe_._Lrp1
Psap_._Lrp1
Igfbp4_._Lrp6
Psap_._Sort1
C3_._Itgax
Ccl5_._Sdc4
Ccl3_._Ccr5
Il16_._Ccr5
Ccl4_._Ccr5
Ccl5_._Ccr5
Csf1_._Csf1r
Tnf_._Tnfrsf1a
Tnf_._Tnfrsf1b
Ccl5_._Cxcr3
Lpl_._Cd44
Fn1_._Cd44
Il16_._Cd4
Slpi_._Cd4
C3_._Cd19
Fn1_._Itgb7
B2m_._Cd3g
B2m_._Cd3d
F13a1_._Itgb1
F13a1_._Itga4
Lyz2_._Itgal
Fn1_._Itgb1
Fn1_._Itga4
B2m_._Klrd1
Lgals3bp_._Itgb1
Gnas_._Adcy7
B2m_._Cd247
C3_._Lrp1
Ifng_._Ifngr1
Il1b_._Il1r2
Cd14_._Itga4
Ifng_._Ifngr2
Control_t1_Spleen
Allergic_t1_Spleen
CpG-AIT_t1_Spleen

Figure 4

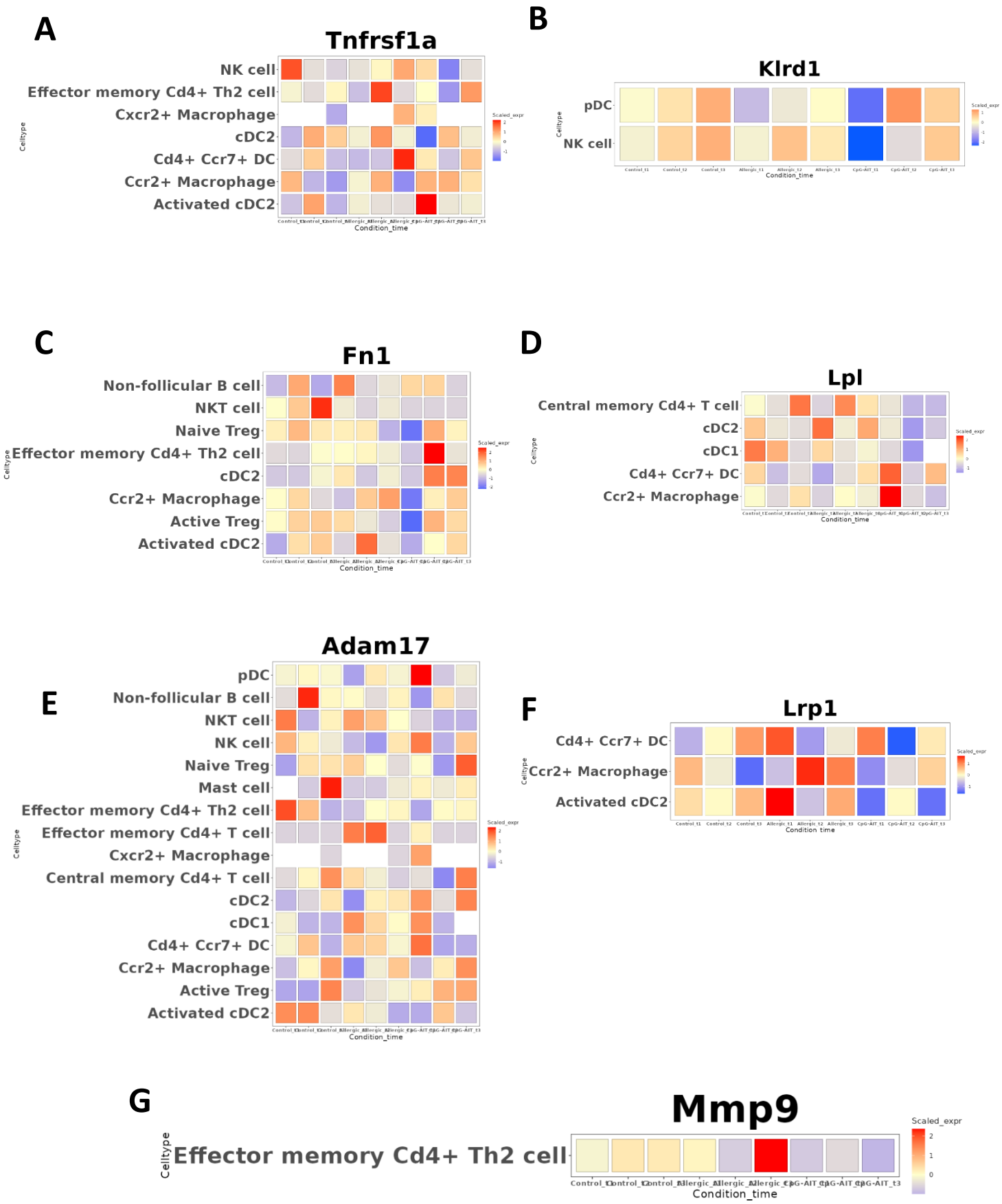
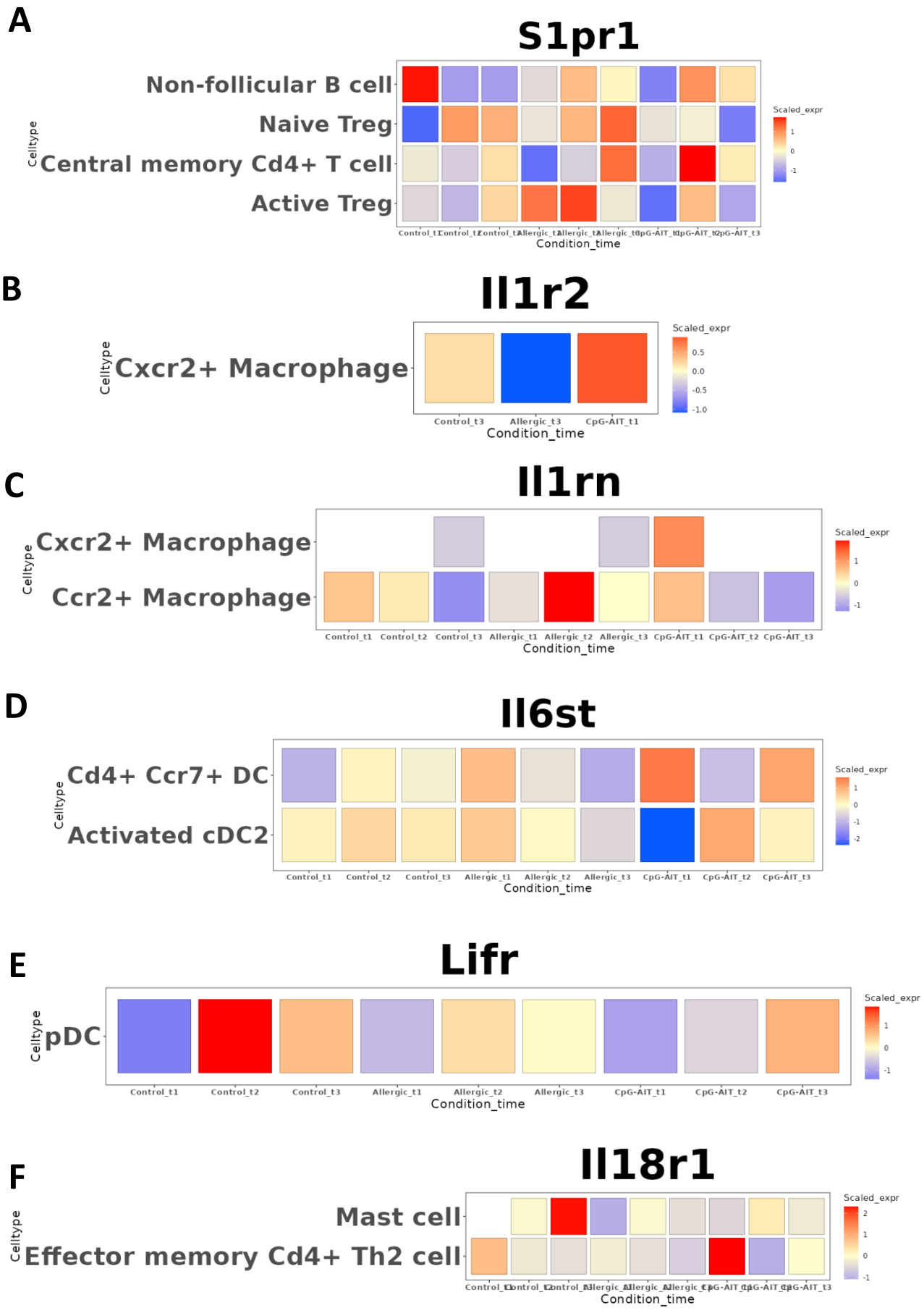
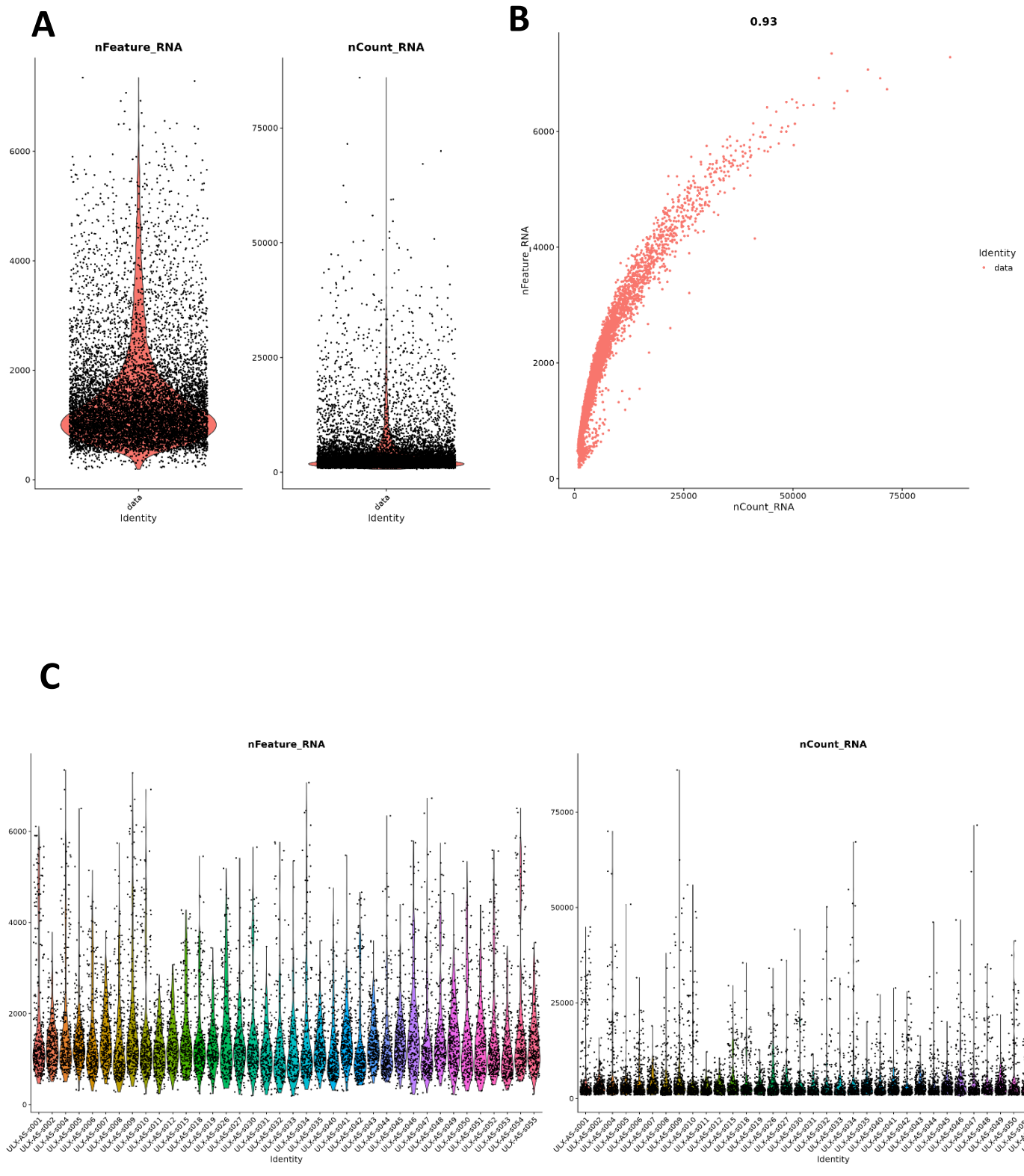

A
Tnfrsf1a
NK cell
Effector memory Cd4+ Th2 cell
Cxcr2+ Macrophage
cDC2
Cd4+ Ccr7+ DC
Ccr2+ Macrophage
Activated cDC2
Celltype
Condition_time
Scaled_expr
B
Klrd1
pDC
NK cell
Celltype
Condition_time
Scaled_expr
C
Fn1
Non-follicular B cell
NKT cell
Naive Treg
Effector memory Cd4+ Th2 cell
cDC2
Ccr2+ Macrophage
Active Treg
Activated cDC2
Celltype
Condition_time
Scaled_expr
D
Lpl
Central memory Cd4+ T cell
cDC2
cDC1
Cd4+ Ccr7+ DC
Ccr2+ Macrophage
Celltype
Condition_time
Scaled_expr
E
Adam17
pDC
Non-follicular B cell
NKT cell
NK cell
Naive Treg
Mast cell
Effector memory Cd4+ Th2 cell
Effector memory Cd4+ T cell
Cxcr2+ Macrophage
Central memory Cd4+ T cell
cDC2
cDC1
Cd4+ Ccr7+ DC
Ccr2+ Macrophage
Active Treg
Activated cDC2
Celltype
Condition_time
Scaled_expr
F
Lrp1
Cd4+ Ccr7+ DC
Ccr2+ Macrophage
Activated cDC2
Celltype
Condition time
Scaled_expr
G
Mmp9
Effector memory Cd4+ Th2 cell
Celltype
Condition_time
Scaled_expr

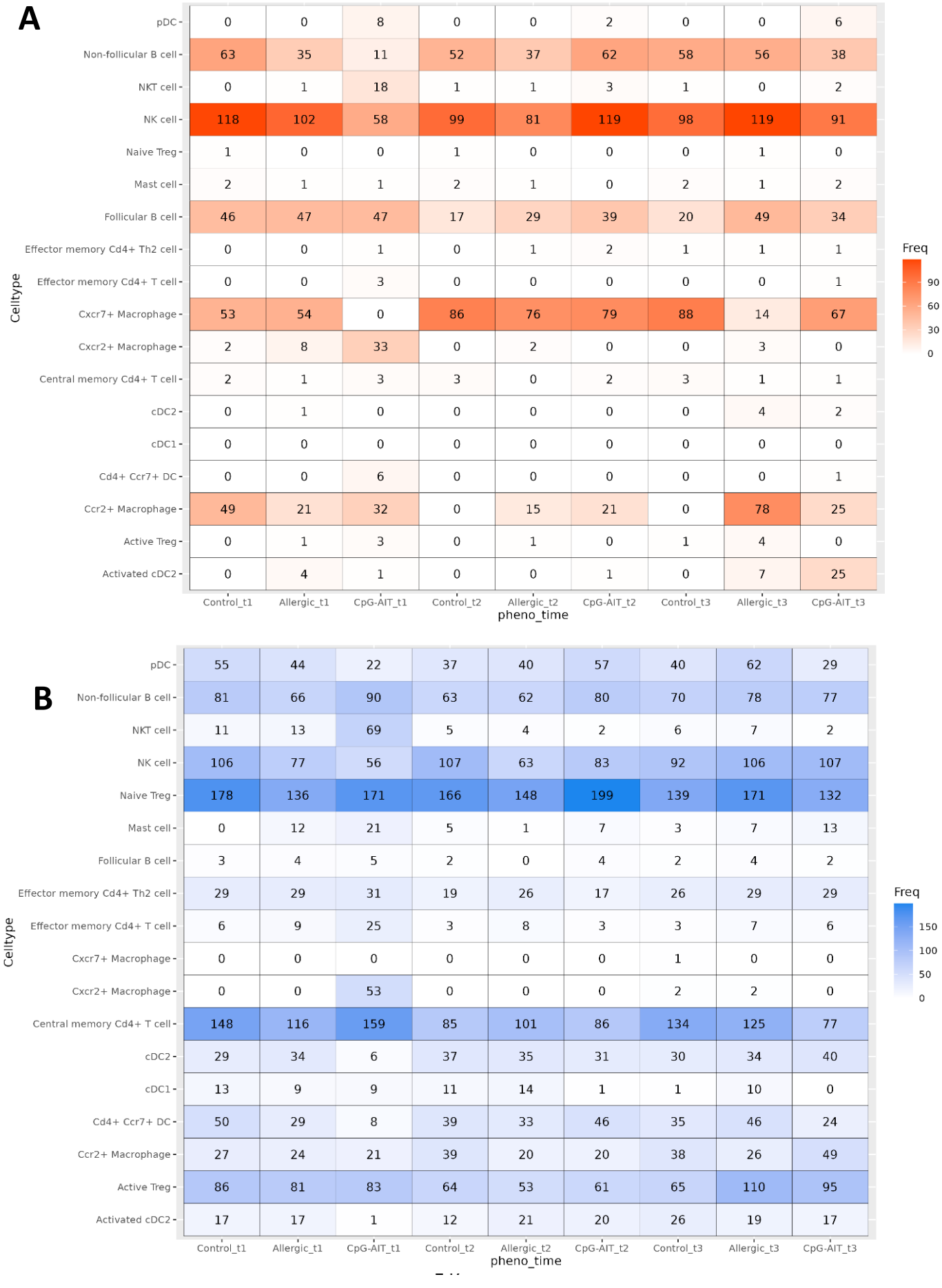
Figure 5


A
S1pr1
Non-follicular B cell
Naive Treg
Central memory Cd4+ T cell
Active Treg
Celltype
Scaled_expr
Control_t1 Control_t2 Control_t3 Allergic_t1 Allergic_t2 Allergic_t3 CpG-AIT_t1 CpG-AIT_t2 CpG-AIT_t3
Condition_time
B
Il1r2
Cxcr2+ Macrophage
Celltype
Scaled_expr
Control_t3 Allergic_t3 CpG-AIT_t1
Condition_time
C
Il1rn
Cxcr2+ Macrophage
Ccr2+ Macrophage
Celltype
Scaled_expr
Control_t1 Control_t2 Control_t3 Allergic_t1 Allergic_t2 Allergic_t3 CpG-AIT_t1 CpG-AIT_t2 CpG-AIT_t3
Condition_time
D
Il6st
Cd4+ Ccr7+ DC
Activated cDC2
Celltype
Scaled_expr
Control_t1 Control_t2 Control_t3 Allergic_t1 Allergic_t2 Allergic_t3 CpG-AIT_t1 CpG-AIT_t2 CpG-AIT_t3
Condition_time
E
Lifr
pDC
Celltype
Scaled_expr
Control_t1 Control_t2 Control_t3 Allergic_t1 Allergic_t2 Allergic_t3 CpG-AIT_t1 CpG-AIT_t2 CpG-AIT_t3
Condition_time
F
Il18r1
Mast cell
Effector memory Cd4+ Th2 cell
Celltype
Scaled_expr
Condition_time

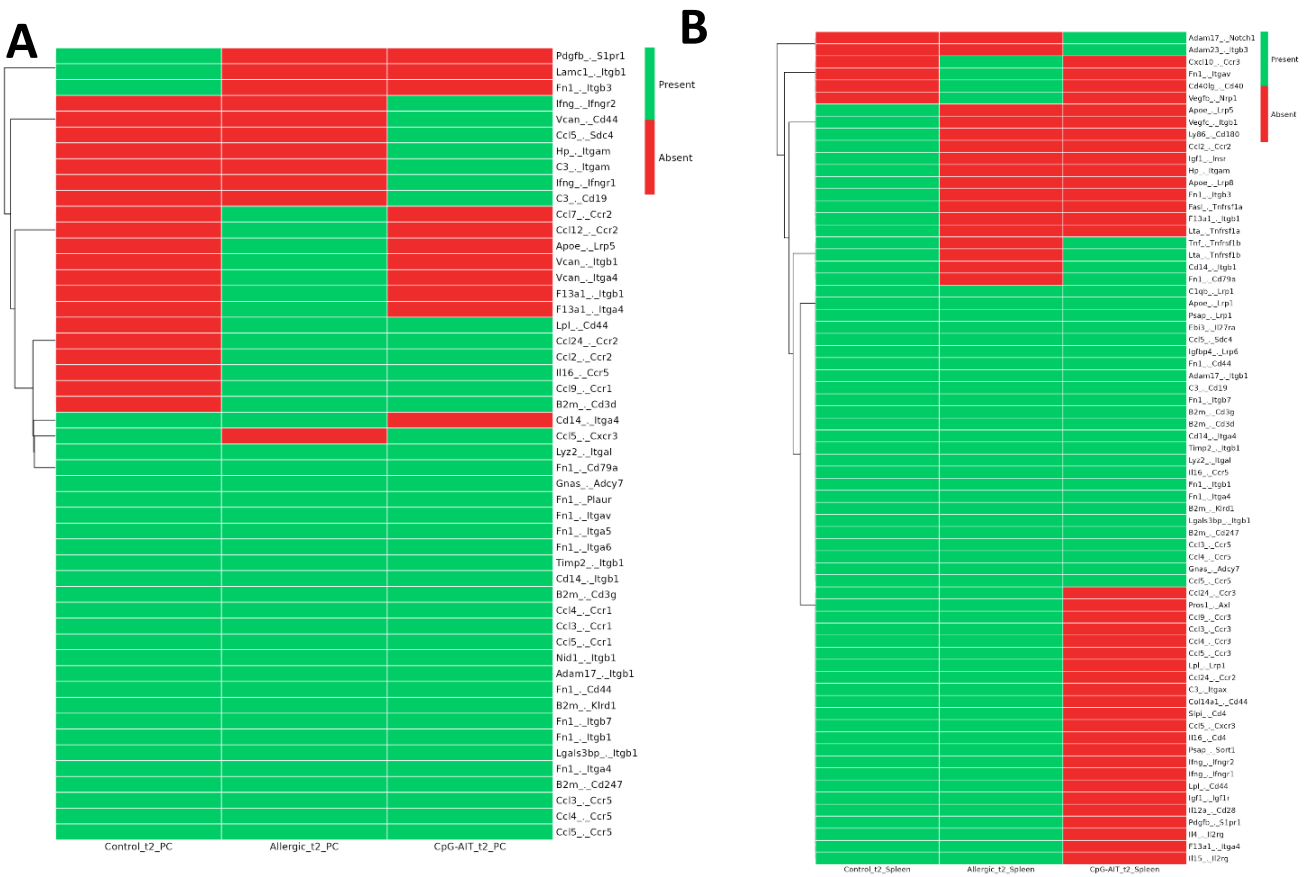
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



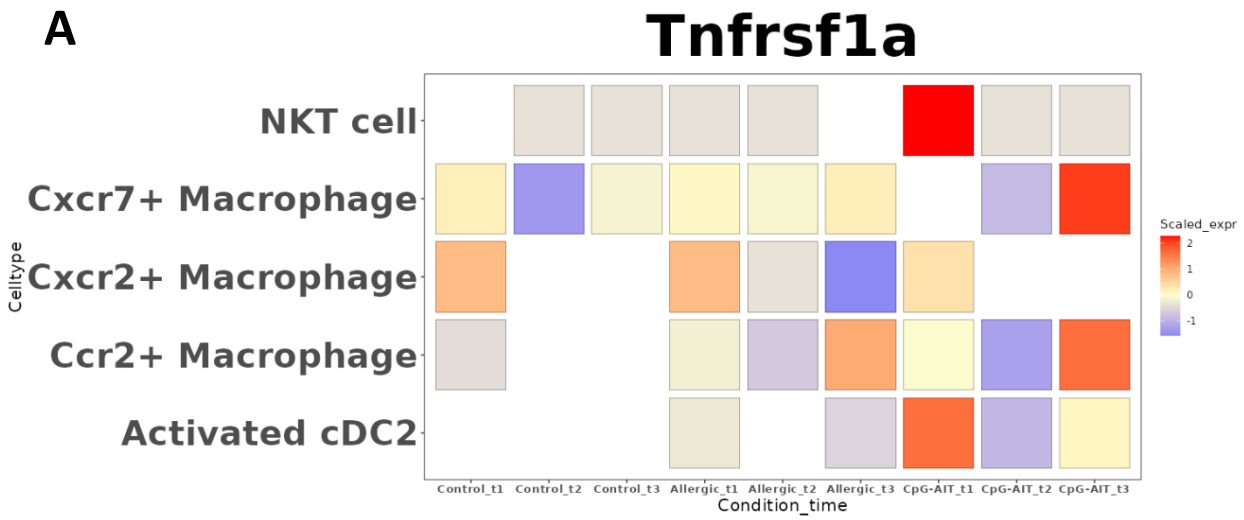
C

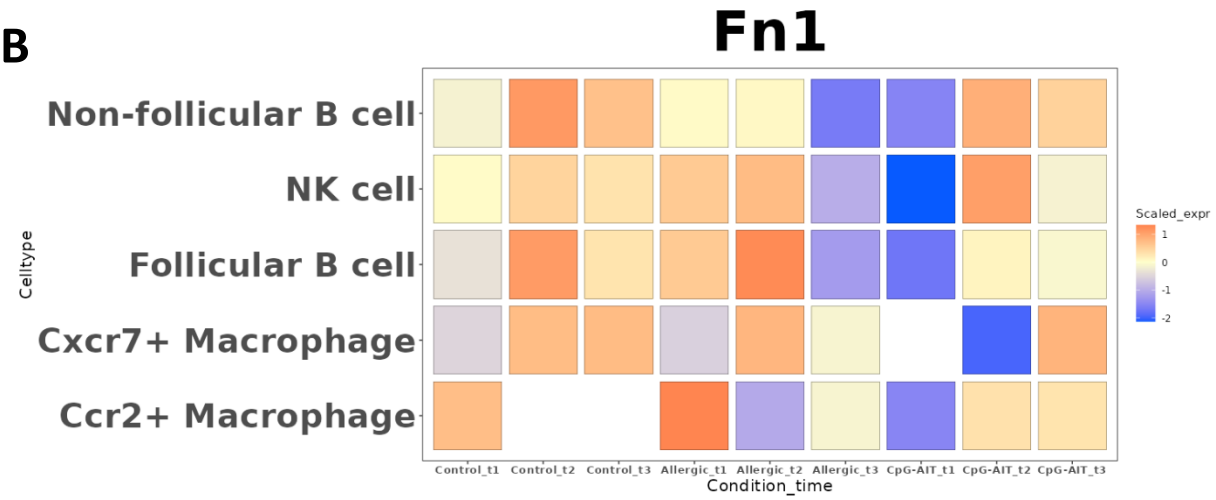

Present
Absent
Cxcl13_:_Cxcr5
Il1b_:_Il1rap
Lpl_:_Lrp1
Psap_:_Lrp1
C1qb_:_Lrp1
Apoe_:_Lrp1
Ccl12_:_Ccr2
Ly86_:_Cd180
Ccl5_:_Sdc4
Psap_:_Sort1
Plau_:_Itgam
C3_:_Itgam
B2m_:_Cd3d
Spp1_:_S1pr1
Ifng_:_Ifngr2
Cd14_:_Itga4
Spp1_:_Itga4
Vcan_:_Itga4
F10_:_Itgam
Vegfc_:_Itgb1
Angpt1_:_Itgb1
Fn1_:_Cd79a
Fn1_:_Itgav
Fn1_:_Itga6
Ccl5_:_Ccr1
Fn1_:_Itgb7
Fn1_:_Itga4
Nid1_:_Itgb1
Timp2_:_Itgb1
Lgals3bp_:_Itgb1
Fn1_:_Itgb1
Adam17_:_Itgb1
B2m_:_Cd247
Ccl4_:_Ccr5
Ccl5_:_Ccr5
C3_:_Cd19
Hp_:_Itgam
Spp1_:_Cd44
Vcan_:_Cd44
Fn1_:_Plaur
Lpl_:_Cd44
Ccl24_:_Ccr2
Vcan_:_Itgb1
Spp1_:_Itgb1
Cd14_:_Itgb1
Fn1_:_Itga5
Ccl9_:_Ccr1
Ccl4_:_Ccr1
Ccl3_:_Ccr1
Pdgfb_:_S1pr1
Lamc1_:_Itgb1
F13a1_:_Itgb1
F13a1_:_Itga4
Ccl2_:_Ccr2
Fn1_:_Cd44
Lyz2_:_Itgal
B2m_:_Klrd1
Ifng_:_Ifngr1
Ccl3_:_Ccr5
Gnas_:_Adcy7
Il16_:_Ccr5
B2m_:_Cd3g
Ccl7_:_Ccr2
Col1a1_:_Cd44
Ccl2_:_Ccr1
Fn1_:_Itgb3
Col1a1_:_Itgb1
Ccl2_:_Ccr5
Control_t3_PC
Allergic_t3_PC
CpG AIT_t3_PC

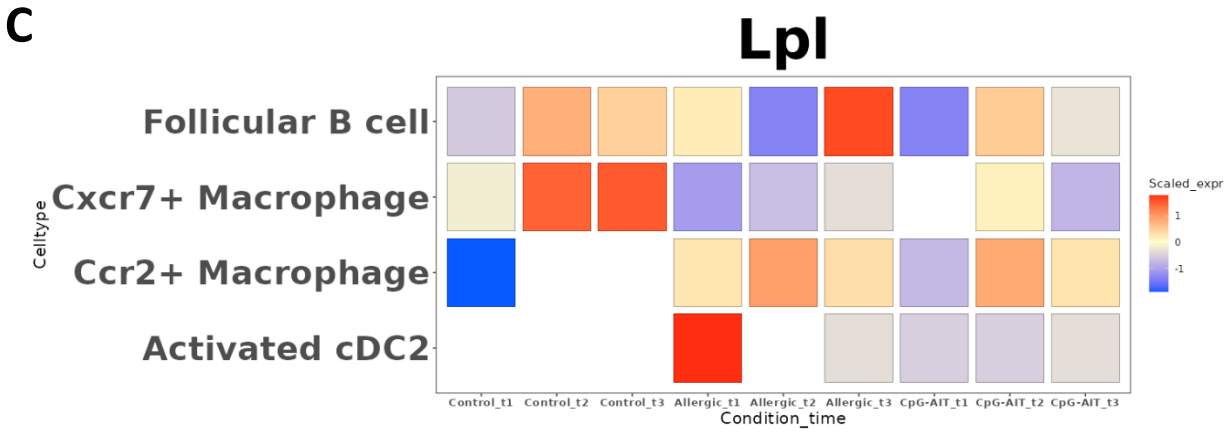
D


Present
Absent
Hp_:_Itgam
Vegfc_:_Itgb1
Ebi3_:_Il27ra
Ccl5_:_Cxcr3
Osm_:_Il6st
Il6_:_Il6st
Cd14_:_Itgb1
Fasl_:_Tnfrsf1a
Tnf_:_Traf2
F13a1_:_Itgb1
Osm_:_Lifr
Apoe_:_Lrp8
Anxa1_:_Egfr
Il6_:_Il6ra
Ccl9_:_Ccr1
Ccl4_:_Ccr1
Ccl3_:_Ccr1
Ccl5_:_Ccr1
Csf1_:_Csf1r
Il12a_:_Cd28
Lama5_:_Itgb1
Adam23_:_Itgb3
Apoe_:_Ldlr
Mmp9_:_Lrp1
Mmp9_:_Itgam
Lpl_:_Lrp1
Igfbp4_:_Lrp6
Ccl5_:_Sdc4
Mmp9_:_Cd44
Lta_:_Tnfrsf1b
Fn1_:_Itgb7
F13a1_:_Itga4
Lpl_:_Cd44
Psap_:_Sort1
Lyz2_:_Itgal
Tnf_:_Tnfrsf1b
Fn1_:_Cd44
Gnas_:_Adcy7
Adam17_:_Itgb1
Slpi_:_Cd4
Il16_:_Cd4
C3_:_Cd19
B2m_:_Cd3g
B2m_:_Cd3d
Lgals3bp_:_Itgb1
Cd14_:_Itga4
Fn1_:_Itgb1
Fn1_:_Itga4
B2m_:_Klrd1
B2m_:_Cd247
Ccl3_:_Ccr5
Ccl5_:_Ccr5
C3_:_Itgax
Tnf_:_Tnfrsf1a
C1qb_:_Lrp1
Apoe_:_Lrp1
Psap_:_Lrp1
Pdgfb_:_Lrp1
Ifng_:_Ifngr1
Ifng_:_Ifngr2
Pdgfb_:_S1pr1
Il4_:_Il2rg
Il16_:_Ccr5
Ccl4_:_Ccr5
Timp2_:_Itgb1
Control_t3_Spleen
Allergic_t3_Spleen
CpG-AIT_t3_Spleen

Supplementary Figure 4

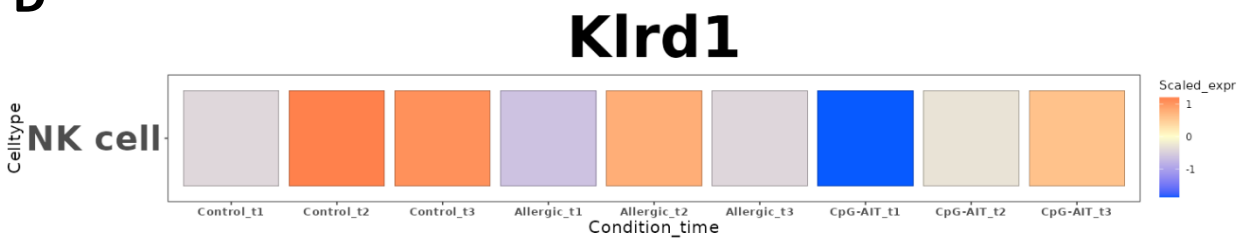




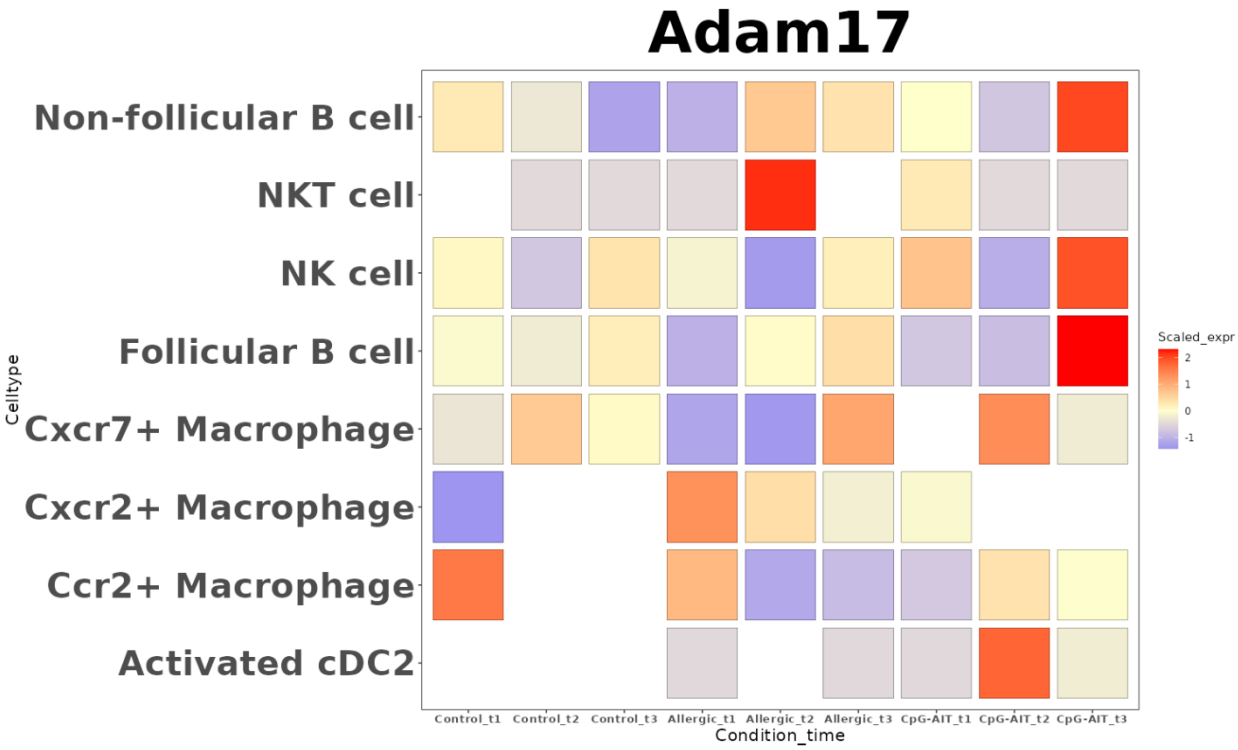


Supplementary Figure 4

D

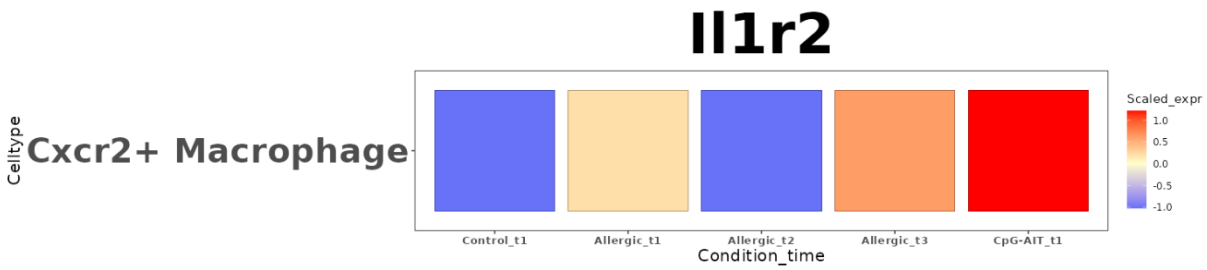


E

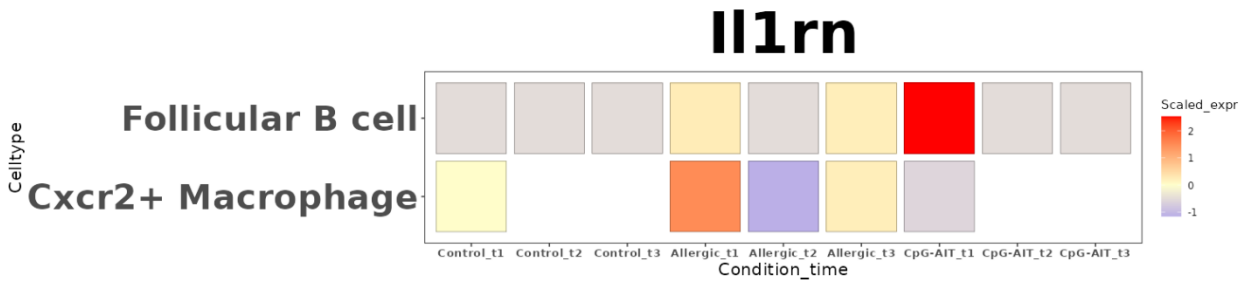


Supplementary Figure 5

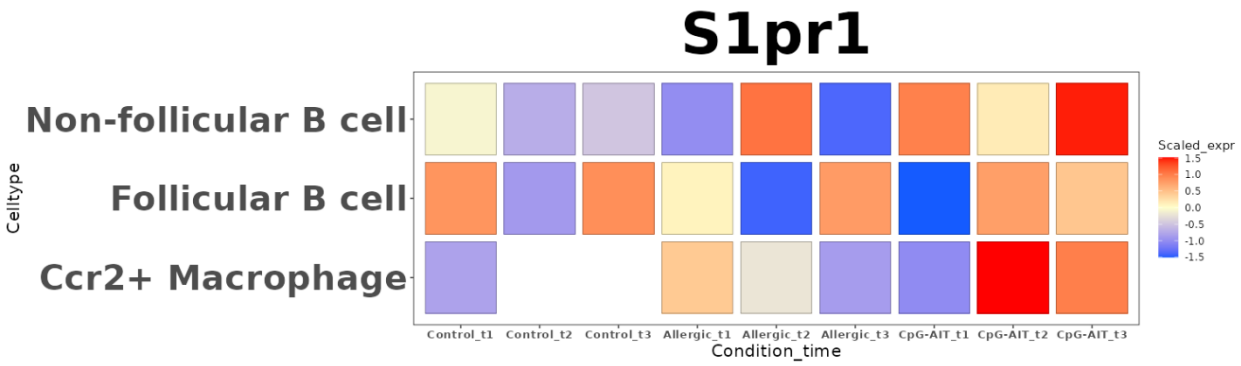
A



B



C



5 Discussion

Multicellular organisms are made up of cells each with its specialized function. This specialization is a combination of a cell lineage and its environment (Morris 2019). The cell receives signals from its environment which regulate the expression of genes that impart the cell its functionality (Gosselin, Link et al. 2014, Chacon-Martinez, Koester et al. 2018). These environments are made up of the neighbouring cells that via multiple ways can influence the cell including soluble molecules, physical contact, etc (Alberts B 2002). Cells secrete soluble molecules known as ligands which can bind to structures called receptors on the surface of other cells it is called a Ligand-Receptor (L-R) interaction and is an important part of cell-cell communication. Early research in the field required overexpression/inhibition experiments of ligands/receptors to study their role in biology (Singh, Phillips et al. 2000, Schneider, Willen et al. 2014). As ligand-receptor interactions are promiscuous which means multiple ligands can bind to multiple receptors with receptors also having subunits makes their study experimentally problematic. Recent advances in high-throughput data generation have provided an important opportunity to create computational models of cell-cell communication. These computational models have been helpful not only in studying these interactions under physiological conditions but also have been utilized to understand the role various genes play in the dysfunction of cells due to disease. While this growth in data has fuelled advancement in our understanding of the role of cell-cell communication in health and disease conditions. There remain many limitations in the field that still need to be addressed. This thesis presents three studies that study cell-cell communication and its influence on cells including the role of signals from the niche in creating functional cell states within celltypes, the effect of ligand perturbation on the transcriptome and epigenome of the cells and the role of feedback loops in maintaining the state of tissues. These studies focused on both aspects of filling the gap in our basic understanding of the role of cells in heath as well as the identification of therapeutic targets in case of disease.

5.1 Characterization of functional heterogeneity and niche induced cell states

Functional heterogeneity in cells serves a vital purpose in the maintenance of homeostasis as these cells can be used to play various roles as required under conditions of stress. This functional heterogeneity is shaped and maintained by the signals coming from a cell's niche. The niche in a tissue is composed of all the other cells in the tissue providing sustained signals and shaping the

functional state of the cells by contributing to the signalling and transcription factor profile of these cells. Cell-cell communication provides a new and promising approach to understanding the role of tissue in regulating these functional states. This approach also helps gain an understanding of the role other cell types play in maintaining the cell states thus opening new avenues in developing strategies to revolve the detrimental effects of aging and disease on various cell types.

We designed a new pipeline and produced a computational tool to identify these functional cell states using the cell-cell communication network of the tissue. The tool relies on celltype annotated single cell gene expression data to re-construct these cell-cell communication networks. We identified the states across celltypes and tissues from various datasets to identify their functional role in the tissues.

5.1.1 Strengths

Leveraging the advantages provided by single cell resolution of emerging techniques the tool ventures into the applications of cell-cell communication. FunRes uses transcription factor (TF) conservation as the starting point of its pipeline as the transcription factors play a dominant role in molding the function of the cells. The pipeline also uses a previously established model of signaling to re-enforce the cell-cell communication model. The final part of reconstructing the cell-cell communication model uses a scoring method comparing individual interactions with the background of all the interactions. These steps used in the reconstruction of cell-cell communication add to the confidence of the tool's predictions. The high confidence level of these predictions is also supported by the manually curated protein-protein interaction and ligand-receptor scaffold used in the pipeline.

We next demonstrated the effectiveness of the tool by applying it to multiple datasets of aging and disease. The datasets in total consisted of 177 cell types which provide a broad range of the tool's application. The cell states were then validated using literature evidence providing support to the tool's results. The tool provides for an effective tool for the reconstruction of cell-cell communication network which can also be useful in developing our understanding of the changes that occur during aging or disease. The variation in cell states can help devise strategies to counter the adverse effects of aging or disease in cells. The validation also accompanies a comparison of cell states for cell types across tissues. This cross-tissue analysis provides insights into the states shared and specialized to tissues, which have been previously studied as well.

Finally, we compared FunRes with state-of-the-art clustering methods (SC3, Seurat and SINCERA) to establish the effectiveness of the tool. We observed higher accuracy compared to all the tools in identifying functional cell states by FunRes. Demonstrating the usefulness of using a tool based on cell-cell communication in the identification of functional cell states.

5.1.2 Limitations

As the tool is based on single cell RNA sequencing (scRNA-Seq) data it is also accompanied by limitations of these technologies. scRNA-Seq technologies provide a wide range of depth in their sequencing which is a measure of the number of transcripts required by the sequencer to produce counts which leads to dropouts in the data. Further, the interactions data that we use in the reconstruction of the cell-cell communication network are not complete and biased as certain molecules have been much more studied in comparison to others. As part of the characterization of the cell states identified we utilize known markers. But these markers were not designed for cell states and have been designated as cell type markers under various contexts. Thus, the idea of using them for characterizing cell states is far from perfect but these are the best available literature evidence to support the study.

5.2 Identifying impacts of perturbation in cell-cell communication

Cells reside in different microenvironments which can regulate their function and changes in the environment can lead to shifts in the function of the cells. These kinds of changes are even more prominent when they occur during highly controlled processes like development which require a specific set of signals to be received by the cells from their microenvironment. These signals are then responsible for changes in the transcriptional and epigenetic landscape of the cells which can lead to dysregulation of cellular function as they modify the developmental path taken by these cells.

Here we studied the changes in the transcriptional and epigenetic profiles of astrocytes developing in normal and inflammatory conditions. As inflammation is a component in many neurological disorders and astrocytes are a major component of the central nervous system the study uses omics technologies to study the effects of inflammation on these cells.

5.2.1 Strengths

The study intends to fill the gaps in our current knowledge regarding the development of astrocytes under inflammation. For this purpose, we used a previously established model of Astrocyte development and studied the role of TNF- a master inflammatory cytokine known to be involved in various inflammatory processes both in healthy and disease conditions. This study involved the usage of a multi-omics method to study the gene expression and chromatin accessibility of the astrocytes during the developmental process. We employed the bulk RNA-sequencing and ATAC-sequencing techniques to explore the changes in astrocytes at a genome wide level. Using the generated data, we re-constructed the gene regulatory networks (GRNs) for various timepoints. The GRNs are re-constructed using a manually curated TF-TF interaction scaffold and then further contextualised using the gene expression data to create high confidence regulatory networks. The topological analysis of these networks assists in the identification of regulatory genes involved in astrocyte development. Owing to the role astrocytes play in various neurological disorders with an inflammatory component these genes can be used in advancing therapeutic targets for these disorders.

5.2.2 Limitations

While the study uses advanced omics technology to study the development of astrocytes, these techniques have their limitations. First, the study utilised an in-vitro model of the development of astrocytes, while these models provide an important method to examine the role of inflammation in their development the model cannot be translated directly into therapies without establishing them in in-vivo models. The RNA-Seq provides a snapshot of the transcriptome thus not favorable to studying the dynamics of development further the transcriptome expression does not correlate with the protein expression of the genes adding to the limitations of the technologies involved. ATAC-Seq technology intends to capture the accessibility of transcription factor binding to the promoter regions in the genome. While this technique does offer information about the epigenetic landscape of the cells, it cannot be used to study other epigenetic markers like histone modifications. The GRN re-construction is also constricted by the fact that the TF-TF interactions database used in the study while proving high confidence interactions remains incomplete. Further,

it would be important to test our predicted targets using knockdown experiments to establish their functional role in astrocytic development.

5.3 Feedback loops based regulation in cell-cell communication

Allergic diseases pose a growing challenge in health across the world and are caused when the immune system responds aberrantly to certain environmental factors called allergens. The immune system is composed of various cell types and the coordination between them is important to provide appropriate immune responses. Feedback is an important aspect of the maintenance of cellular states by a constant exchange of information in form of signals between the cells.

In this study, we explored the role of feedback in cell-cell communication by implementing a computational tool that identifies feedback loops after re-constructing the cell-cell communication network in a tissue. We utilized this tool on a model of allergy and tolerance to identify the molecular and cellular regulators of allergy and tolerance. This study generated scRNA-Seq data from immune cells which play specific roles in the immune system in the context of allergic disease and its therapy. We propose novel targets that can be utilized moving forward to understand the mechanisms of allergic diseases and develop therapies.

5.3.1 Strengths

In the biological context, the study utilizes a novel model of allergy induction and its resolution by installing a tolerant state in the tissue. We studied a subset of immune cells for this purpose by generating single cell resolution gene expression data from the cells across two tissues from multiple timepoints across the experiment. On the computational front, the study introduces a novel computational tool with identifies feedback loops in re-constructed cell-cell communication networks. The cell-cell communication network is reconstructed using an upgraded model of a previously published cell-cell communication tool. The strengths of the computational tool also include the use of manually curated interactions to model the TF-TF interactions and cell-cell interactions. Further, the tool resolves the issue of the time taken in calculating the feedback for larger networks by using an algorithm that can break the problem down into smaller pieces and solve them. These feedback loops aid in the identification of molecular and cellular targets of allergy and tolerance. The targets were then validated using literature to create a list of possible

therapeutic targets. This study has novelties in both computational and biological sections along with strengthening the conceptual role of cell-cell communication and feedback regulation in the maintenance of allergy and tolerance states in tissues. Also, the code for the feedback tool and data generated by this study will be made available using open-source databases for the research community.

5.3.2 Limitations

This study intends to push forward our understanding of cell-cell communication and the role feedback play in the maintenance of tissue states. But still there remain certain limitations to this study. First, the in-vivo study of the mouse model is intended to develop strategies for human allergies while mice provide a good model there still are known differences between mouse models and human physiology. The study also is constrained by other technical limitations of the technology used to generate the gene expression data along with changes that occur during the processing of the cells. The study in its current form is also limited in its applications because while our molecular targets have been validated using literature evidence it is necessary to provide experimental evidence. We are currently in process of resolving this limitation by performing perturbation experiments for the therapeutic targets identified by the tool.

5.4 Outlook

This thesis through its research projects explores important aspects of cell-cell communication. But as previously mentioned this research work accompanies certain limitations. Here we discuss the future avenues of further research from different perspectives – technical, biological and conceptual. Implementation of such new frameworks should allow for better optimization in modelling processes involving communication between cells. This would allow a better and more accurate understanding of biology and could help develop therapeutic strategies that remain unexplored.

Overcoming technical limitations

The computational tools developed in this thesis and the wider research community depend upon the evolution of high-throughput technologies. The data produced by these technologies serve as

the backbone of the computational tools being developed (Berger, Peng et al. 2013). In the projects "*Characterization of functional heterogeneity and niche induced cell states*" (sections 4.1) and "*Feedback loops based regulation in cell-cell communication*" (sections 4.3) we have utilized single cell RNA-sequencing datasets which rely on the technologies developed recently that count the number of transcripts in individual cells (Jovic, Liang et al. 2022). These technologies are limited by the technical resolution in the depth of capturing the number of genes and minimum transcripts that can be processed using these technologies. Further, the data generated in the study "*Identifying impacts of perturbation in cell-cell communication*" (section 4.2) is bulk RNA-seq and ATAC-seq data to study developing astrocytes, usage of single cell resolution data in similar time series based study could help understand the developmental process of important cells like astrocytes in CNS, as well as changes in their functional cell states under inflammatory conditions thus providing a better understanding of core developmental/differentiation processes and the impact of inflammation has on these processes. This study also relies on the integration of transcriptomic and epigenetic information of cells, but the data does not belong to the same cells. Only recently, technologies have been developed to study both transcriptomics and epigenomics in the same cell at single cell level (Reyes, Billman et al. 2019). Usage of these new technologies can greatly exemplify the knowledge gained during such experiments providing a more systematic understanding of the developmental process in cells that perform key functions like astrocytes in CNS. Also, ATAC-seq data can only retrieve knowledge of chromatin accessibility in cells but is unable to take into consideration other epigenetic markers of DNA acylation and methylation that control the expression of genes (Jaenisch and Bird 2003). Therefore, to capture information of the regulatory mechanisms in cells it is essential to develop techniques to capture multi-OMICS data from the same cell.

Incompleteness of biological knowledge in literature

The studies in this thesis utilize knowledge of biological interactions from the literature to model ligand-receptor interactions as well as signalling and gene regulatory networks. This knowledge has been acquired over decades of experimental research that has yielded high confidence interaction knowledge of components of all the networks that in combination make cell-cell communication feasible. Expansion in the knowledge of these interactions could help identify previously unknown channels of communication between cells and the possible role they play in

the functional regulation of cells. In the study "*Characterization of functional heterogeneity and niche induced cell states*" (sections 4.1) we utilize known markers to identify functional cell states. In absence of relevant cell state/subtype markers for the celltypes in our datasets, we had to use celltype markers to identify these cell states (Trapnell 2015). Aggregation of cell state markers could be highly useful to characterize the exact cell states/subtypes like those discovered using our FunRes tool. Investigation using better cell state markers could further facilitate understanding the states that appear or vanish during aging/disease conditions (Ponting 2019). The tool also provides the receptor-TF combinations that regulate these cell states thus facilitating gene targets for developing therapeutic approaches by modulating cell states. Also, in the study "*Feedback loops based regulation in cell-cell communication*" (section 4.3) we use known markers to annotate cell clusters identified in the project. The study is the first sequencing experiment for cells extracted from mouse treated by hCpG-AIT. The cell clusters identified in this study could potentially help in better annotation of cells in further experiments using hCpG-AIT treatment against allergic inflammation. In the same study the annotation of feedback loops is done as pro-/anti- or both based on literature evidence. Many of the molecules were annotated to be involved in both processes, while this may be fundamental to biological processes a better determination of involvement of the genes in either pro- or anti- inflammatory processes could be advantageous in assessing possible targets of allergic inflammation and tolerance induction thus advancing work on identification of therapeutic targets against allergic diseases as well as other disease involving inflammation/tolerance.

New conceptual frameworks for biological processes

The conceptual frameworks implemented in various computational tools are an essential part of understanding biological processes. In the study "*Characterization of functional heterogeneity and niche induced cell states*" (sections 4.1) we used the knowledge that cells are shaped by the signals they receive from their niche. The identity of these cells belonging to specific celltypes are based on certain markers but in a different niche, the same cell types acquire different functional states. This framework was able to infer cell states in celltypes across tissues, with shared and distinct receptor-TF combinations demonstrating the advantage of using our new framework. For the study "*Feedback loops based regulation in cell-cell communication*" (section 4.3), we developed a novel framework involving feedback loops in cell-cell communication. The role of feedback has been

already established in the maintenance of cell and tissue states (Brandman and Meyer 2008, O'Brien 2022). We implemented this new framework to identify gene targets for allergy and tolerance. These studies demonstrate the importance of developing these new conceptual frameworks to understand the underlying biology. While these studies have explored some aspects of cell-cell communication, its role in many fundamental biological processes such as homeostasis (Valls and Esposito 2022) and self-organization (Hiraiwa 2020, Japon, Jimenez-Morales et al. 2022) remains unexplored. The development of computational tools that could model these processes would be beneficial in understanding the role of different cellular signals and responses in biological systems.

6 Conclusion

Development of therapeutic targets against diseases requires an accurate characterization of the role of different celltypes and genes involved in the dysregulation of cells. Cell-cell communication provides a new avenue of research to determine the role of different cellular populations and the genes that regulate their functionality. This thesis aimed to perform studies and develop computational tools that improve our understanding of the role cell-cell communication plays in cellular function. These projects focus on the limitations of existing tools and fill the gap in current knowledge in the literature.

In summary, this thesis provides the following contributions:

- **Development of a new method to resolve cell states:** FunRes is a novel computational tool designed to utilize a reconstructed cell-cell communication network to resolve functional cell states. It was applied to ageing and disease datasets to determine appearing/vanishing functional cell states. The tool has been developed in R and its code for FunRes is publicly available at: https://git-r3lab.uni.lu/CBG/FunRes.
- **Studying the development of astrocytes under normal and inflammatory conditions:** The project was designed to study the development of astrocytes under normal and inflammatory conditions. Bulk transcriptomics and epigenomics data were generated during the project. Using an integrated approach GRNs were reconstructed and key genes

were identified. The code used to analyse the data and generate manuscript figures is in the private repository: https://git-r3lab.uni.lu/CBG/Astrosys. The bulk RNA-seq and ATAC-seq data along with the manuscript code would be made public after publication.

- **Implemented a new method to identify key regulators of allergy and tolerance using feedback loops:** C3Loop is a new computational tool that reconstructs cell-cell communication network and identifies feedback loops. The tool was utilized to find key regulators of allergic inflammation and tolerance induction in a mouse model. The code for C3Loop has been submitted to a private repository: https://git-r3lab.uni.lu/CBG/C3Loop. The single cell RNA-seq data generated during the project along with the tool's code would be made publicly available upon publication.

Through the projects in this thesis, we explored the role of cell-cell communication through the development of an analytical study and two novel computational tools. These projects were envisioned to explore various aspects of cell-cell communication. Determining the part various cell types and states play in health and disease could provide a potential first step toward the development of new therapeutic approaches. These projects provide insights into biological processes by application of systems biology based computational modelling.

7 References

1. Alberts B, J. A., Lewis J, et al. (2002). General Principles of Cell Communication. Molecular Biology of the Cell. New York, Garland Science.
2. Altschuler, S. J. and L. F. Wu (2010). "Cellular heterogeneity: do differences make a difference?" Cell **141**(4): 559-563.
3. Arango Duque, G. and A. Descoteaux (2014). "Macrophage cytokines: involvement in immunity and infectious diseases." Front Immunol **5**: 491.
4. Armingol, E., A. Officer, O. Harismendy and N. E. Lewis (2021). "Deciphering cell-cell interactions and communication from gene expression." Nat Rev Genet **22**(2): 71-88.
5. Banyer, J. L., N. H. Hamilton, I. A. Ramshaw and A. J. Ramsay (2000). "Cytokines in innate and adaptive immunity." Rev Immunogenet **2**(3): 359-373.
6. Bennett, J. M., G. Reeves, G. E. Billman and J. P. Sturmberg (2018). "Inflammation-Nature's Way to Efficiently Respond to All Types of Challenges: Implications for Understanding and Managing "the Epidemic" of Chronic Diseases." Front Med (Lausanne) **5**: 316.
7. Berger, B., J. Peng and M. Singh (2013). "Computational solutions for omics data." Nat Rev Genet **14**(5): 333-346.
8. Boland, M. J., K. L. Nazor and J. F. Loring (2014). "Epigenetic regulation of pluripotency and differentiation." Circ Res **115**(2): 311-324.
9. Boltjes, A. and F. van Wijk (2014). "Human dendritic cell functional specialization in steady-state and inflammation." Front Immunol **5**: 131.
10. Brandman, O. and T. Meyer (2008). "Feedback loops shape cellular signals in space and time." Science **322**(5900): 390-395.
11. Brigandt, I. a. L., Alan. (2022). "Reductionism in Biology." from https://plato.stanford.edu/archives/sum2022/entries/reduction-biology/.
12. Browaeys, R., W. Saelens and Y. Saeys (2020). "NicheNet: modeling intercellular communication by linking ligands to target genes." Nat Methods **17**(2): 159-162.
13. Cang, Z. and Q. Nie (2020). "Inferring spatial and signaling relationships between cells from single cell transcriptomic data." Nat Commun **11**(1): 2084.
14. Cedar, H. and Y. Bergman (2008). Epigenetic silencing during early lineage commitment. StemBook. Cambridge (MA).
15. Chacon-Martinez, C. A., J. Koester and S. A. Wickstrom (2018). "Signaling in the stem cell niche: regulating cell fate, function and plasticity." Development **145**(15).
16. Chen, L., H. Deng, H. Cui, J. Fang, Z. Zuo, J. Deng, Y. Li, X. Wang and L. Zhao (2018). "Inflammatory responses and inflammation-associated diseases in organs." Oncotarget **9**(6): 7204-7218.
17. D'Argenio, V. (2018). "The High-Throughput Analyses Era: Are We Ready for the Data Struggle?" High Throughput **7**(1).
18. del Sol, A., R. Balling, L. Hood and D. Galas (2010). "Diseases as network perturbations." Curr Opin Biotechnol **21**(4): 566-571.
19. Dostert, C., M. Grusdat, E. Letellier and D. Brenner (2019). "The TNF Family of Ligands and Receptors: Communication Modules in the Immune System and Beyond." Physiol Rev **99**(1): 115-160.
20. Efremova, M., M. Vento-Tormo, S. A. Teichmann and R. Vento-Tormo (2020). "CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes." Nat Protoc **15**(4): 1484-1506.

21. Eng, C. L., M. Lawson, Q. Zhu, R. Dries, N. Koulena, Y. Takei, J. Yun, C. Cronin, C. Karp, G. C. Yuan and L. Cai (2019). "Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH." Nature **568**(7751): 235-239.
22. Eyo, U. B. and L. J. Wu (2013). "Bidirectional microglia-neuron communication in the healthy brain." Neural Plast **2013**: 456857.
23. Fischer, D. S., A. C. Schaar and F. J. Theis (2022). "Modeling intercellular communication in tissues using spatial graphs of cells." Nat Biotechnol.
24. Gosselin, D., V. M. Link, C. E. Romanoski, G. J. Fonseca, D. Z. Eichenfield, N. J. Spann, J. D. Stender, H. B. Chun, H. Garner, F. Geissmann and C. K. Glass (2014). "Environment drives selection and function of enhancers controlling tissue-specific macrophage identities." Cell **159**(6): 1327-1340.
25. Guilliams, M., J. Bonnardel, B. Haest, B. Vanderborght, C. Wagner, A. Remmerie, A. Bujko, L. Martens, T. Thone, R. Browaeys, F. F. De Ponti, B. Vanneste, C. Zwicker, F. R. Svedberg, T. Vanhalewyn, A. Goncalves, S. Lippens, B. Devriendt, E. Cox, G. Ferrero, V. Wittamer, A. Willaert, S. J. F. Kaptein, J. Neyts, K. Dallmeier, P. Geldhof, S. Casaert, B. Deplancke, P. Ten Dijke, A. Hoorens, A. Vanlander, F. Berrevoet, Y. Van Nieuwenhove, Y. Saeys, W. Saelens, H. Van Vlierberghe, L. Devisscher and C. L. Scott (2022). "Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches." Cell **185**(2): 379-396 e338.
26. Heldin, C. H., B. Lu, R. Evans and J. S. Gutkind (2016). "Signals and Receptors." Cold Spring Harb Perspect Biol **8**(4): a005900.
27. Hill, M. M. and C. Gerner (2021). "Integrative Multi-Omics in Biomedical Research." Biomolecules **11**(10).
28. Hiraiwa, T. (2020). "Dynamic Self-Organization of Idealized Migrating Cells by Contact Communication." Phys Rev Lett **125**(26): 268104.
29. Huang, M., L. Xu, J. Liu, P. Huang, Y. Tan and S. Chen (2022). "Cell-Cell Communication Alterations via Intercellular Signaling Pathways in Substantia Nigra of Parkinson's Disease." Front Aging Neurosci **14**: 828457.
30. Hunter, T. (2000). "Signaling--2000 and beyond." Cell **100**(1): 113-127.
31. Iwakiri, Y., V. Shah and D. C. Rockey (2014). "Vascular pathobiology in chronic liver disease and cirrhosis - current status and future directions." J Hepatol **61**(4): 912-924.
32. Jaenisch, R. and A. Bird (2003). "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals." Nat Genet **33 Suppl**: 245-254.
33. Japon, P., F. Jimenez-Morales and F. Casares (2022). "Intercellular communication and the organization of simple multicellular animals." Cells Dev **169**: 203726.
34. Jiang, P., Y. Zhang, B. Ru, Y. Yang, T. Vu, R. Paul, A. Mirza, G. Altan-Bonnet, L. Liu, E. Ruppin, L. Wakefield and K. W. Wucherpfennig (2021). "Systematic investigation of cytokine signaling activity at the tissue and single-cell levels." Nat Methods **18**(10): 1181-1191.
35. Jovic, D., X. Liang, H. Zeng, L. Lin, F. Xu and Y. Luo (2022). "Single-cell RNA sequencing technologies and applications: A brief overview." Clin Transl Med **12**(3): e694.
36. Kany, S., J. T. Vollrath and B. Relja (2019). "Cytokines in Inflammatory Disease." Int J Mol Sci **20**(23).
37. Kluger, Y., D. P. Tuck, J. T. Chang, Y. Nakayama, R. Poddar, N. Kohya, Z. Lian, A. Ben Nasr, H. R. Halaban, D. S. Krause, X. Zhang, P. E. Newburger and S. M. Weissman (2004). "Lineage specificity of gene expression patterns." Proc Natl Acad Sci U S A **101**(17): 6508-6513.
38. Kucharavy, A., B. Rubinstein, J. Zhu and R. Li (2018). "Robustness and evolvability of heterogeneous cell populations." Mol Biol Cell **29**(11): 1400-1409.
39. Levine, M. and E. H. Davidson (2005). "Gene regulatory networks for development." Proc Natl Acad Sci U S A **102**(14): 4936-4942.

40. Li, P. and M. B. Elowitz (2019). "Communication codes in developmental signaling pathways." Development **146**(12).
41. Liu, C., D. Chu, K. Kalantar-Zadeh, J. George, H. A. Young and G. Liu (2021). "Cytokines: From Clinical Significance to Quantification." Adv Sci (Weinh) **8**(15): e2004433.
42. Malyshev, I. Y., E. B. Manukhina and Y. I. Malyshev (2014). "Physiological organization of immune response based on the homeostatic mechanism of matrix reprogramming: implication in tumor and biotechnology." Med Hypotheses **82**(6): 754-765.
43. Matejuk, A. and R. M. Ransohoff (2020). "Crosstalk Between Astrocytes and Microglia: An Overview." Front Immunol **11**: 1416.
44. McDavid, A., G. Finak and R. Gottardo (2016). "The contribution of cell cycle to heterogeneity in single-cell RNA-seq data." Nat Biotechnol **34**(6): 591-593.
45. Morris, S. A. (2019). "The evolving concept of cell identity in the single cell era." Development **146**(12).
46. Mulder, S., H. Hamidi, M. Kretzler and W. Ju (2018). "An integrative systems biology approach for precision medicine in diabetic kidney disease." Diabetes Obes Metab **20 Suppl 3**(Suppl 3): 6-13.
47. Mylonas, A. and A. O'Loghlen (2022). "Cellular Senescence and Ageing: Mechanisms and Interventions." Front Aging **3**: 866718.
48. O'Brien, L. E. (2022). "Tissue Homeostasis and Non-Homeostasis: From Cell Life Cycles to Organ States." Annu Rev Cell Dev Biol **38**: 395-418.
49. Okin, D. and R. Medzhitov (2012). "Evolution of inflammatory diseases." Curr Biol **22**(17): R733-740.
50. Overton, D. L. and T. L. Mastracci (2022). "Exocrine-Endocrine Crosstalk: The Influence of Pancreatic Cellular Communications on Organ Growth, Function and Disease." Front Endocrinol (Lausanne) **13**: 904004.
51. Park, Y. E., D. S. Musson, D. Naot and J. Cornish (2017). "Cell-cell communication in bone development and whole-body homeostasis and pharmacological avenues for bone disorders." Curr Opin Pharmacol **34**: 21-35.
52. Pinero, J., J. M. Ramirez-Anguita, J. Sauch-Pitarch, F. Ronzano, E. Centeno, F. Sanz and L. I. Furlong (2020). "The DisGeNET knowledge platform for disease genomics: 2019 update." Nucleic Acids Res **48**(D1): D845-D855.
53. Ponting, C. P. (2019). "The Human Cell Atlas: making 'cell space' for disease." Dis Model Mech **12**(2).
54. Rahman, A., A. Tiwari, J. Narula and T. Hickling (2018). "Importance of Feedback and Feedforward Loops to Adaptive Immune Response Modeling." CPT Pharmacometrics Syst Pharmacol **7**(10): 621-628.
55. Reyes, M., K. Billman, N. Hacohen and P. C. Blainey (2019). "Simultaneous profiling of gene expression and chromatin accessibility in single cells." Adv Biosyst **3**(11).
56. Rodriques, S. G., R. R. Stickels, A. Goeva, C. A. Martin, E. Murray, C. R. Vanderburg, J. Welch, L. M. Chen, F. Chen and E. Z. Macosko (2019). "Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution." Science **363**(6434): 1463-1467.
57. Rommelfanger, M. K. and A. L. MacLean (2021). "A single-cell resolved cell-cell communication model explains lineage commitment in hematopoiesis." Development **148**(24).
58. Ruch, R. J. (2002). "Intercellular communication, homeostasis, and toxicology." Toxicol Sci **68**(2): 265-266.
59. Schneider, P., L. Willen and C. R. Smulski (2014). "Tools and techniques to study ligand-receptor interactions and receptor activation by TNF superfamily members." Methods Enzymol **545**: 103-125.

60. Singh, N., R. A. Phillips, N. N. Iscove and S. E. Egan (2000). "Expression of notch receptors, notch ligands, and fringe genes in hematopoiesis." Exp Hematol **28**(5): 527-534.
61. TA, B. (2002). Transcriptomes and Proteomes. Genomes, Oxford: Wiley-Liss. **2nd edition**.
62. Tirziu, D., F. J. Giordano and M. Simons (2010). "Cell communications in the heart." Circulation **122**(9): 928-937.
63. Trapnell, C. (2015). "Defining cell types and states with single-cell genomics." Genome Res **25**(10): 1491-1498.
64. Tumpel, S. and K. L. Rudolph (2019). "Quiescence: Good and Bad of Stem Cell Aging." Trends Cell Biol **29**(8): 672-685.
65. Turner, M. D., B. Nedjai, T. Hurst and D. J. Pennington (2014). "Cytokines and chemokines: At the crossroads of cell signalling and inflammatory disease." Biochim Biophys Acta **1843**(11): 2563-2582.
66. Valls, P. O. and A. Esposito (2022). "Signalling dynamics, cell decisions, and homeostatic control in health and disease." Curr Opin Cell Biol **75**: 102066.
67. Vento-Tormo, R., M. Efremova, R. A. Botting, M. Y. Turco, M. Vento-Tormo, K. B. Meyer, J. E. Park, E. Stephenson, K. Polanski, A. Goncalves, L. Gardner, S. Holmqvist, J. Henriksson, A. Zou, A. M. Sharkey, B. Millar, B. Innes, L. Wood, A. Wilbrey-Clark, R. P. Payne, M. A. Ivarsson, S. Lisgo, A. Filby, D. H. Rowitch, J. N. Bulmer, G. J. Wright, M. J. T. Stubbington, M. Haniffa, A. Moffett and S. A. Teichmann (2018). "Single-cell reconstruction of the early maternal-fetal interface in humans." Nature **563**(7731): 347-353.
68. Wanjek, C. (2011). "Systems Biology as Defined by NIH." The NIH Catalyst Volume 19 Issue 6. from https://irp.nih.gov/catalyst/19/6/systems-biology-as-defined-by-nih.
69. Weidemuller, P., M. Kholmatov, E. Petsalaki and J. B. Zaugg (2021). "Transcription factors: Bridge between cell signaling and gene regulation." Proteomics **21**(23-24): e2000034.
70. Weiskirchen, R. and F. Tacke (2014). "Cellular and molecular functions of hepatic stellate cells in inflammatory responses and liver immunology." Hepatobiliary Surg Nutr **3**(6): 344-363.
71. West, N. R. (2019). "Coordination of Immune-Stroma Crosstalk by IL-6 Family Cytokines." Front Immunol **10**: 1093.
72. Wisniewski, J., R. Agrawal and J. A. Woodfolk (2013). "Mechanisms of tolerance induction in allergic disease: integrating current and emerging concepts." Clin Exp Allergy **43**(2): 164-176.
73. Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman and L. A. Romano (2003). "The evolution of transcriptional regulation in eukaryotes." Mol Biol Evol **20**(9): 1377-1419.
74. Wu, J., A. Xie and W. Chen (2014). "Cytokine regulation of immune tolerance." Burns Trauma **2**(1): 11-17.
75. Xia, C., J. Fan, G. Emanuel, J. Hao and X. Zhuang (2019). "Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression." Proc Natl Acad Sci U S A **116**(39): 19490-19499.
76. Xu, S., K. E. Ware, Y. Ding, S. Y. Kim, M. U. Sheth, S. Rao, W. Chan, A. J. Armstrong, W. C. Eward, M. K. Jolly and J. A. Somarelli (2019). "An Integrative Systems Biology and Experimental Approach Identifies Convergence of Epithelial Plasticity, Metabolism, and Autophagy to Promote Chemoresistance." J Clin Med **8**(2).
77. Yu, G., L. G. Wang, Y. Han and Q. Y. He (2012). "clusterProfiler: an R package for comparing biological themes among gene clusters." OMICS **16**(5): 284-287.
78. Zhang, C., M. Yang and A. C. Ericsson (2021). "Function of Macrophages in Disease: Current Understanding on Molecular Mechanisms." Front Immunol **12**: 620510.

79. Zhang, Y., T. Liu, X. Hu, M. Wang, J. Wang, B. Zou, P. Tan, T. Cui, Y. Dou, L. Ning, Y. Huang, S. Rao, D. Wang and X. Zhao (2021). "CellCall: integrating paired ligand-receptor and transcription factor activities for cell-cell communication." Nucleic Acids Res **49**(15): 8520-8534.