Multiomic phenotyping of the airways in health and disease at single-cell resolution to discover molecular mechanisms of asthma

Wojciech Lasoń

Wolfson College
University of Oxford

Thesis submitted for the degree of
Doctor of Philosophy

Michaelmas 2024

Dedication

Dla kochających Rodziców i Babci,
w podziękowaniu za nieprzerwaną wiarę we mnie

ABSTRACT

Asthma is a heterogenous disease of the airways showing phenotypic diversity across patients, requiring varied approaches to treatment. However, the underlying pathophysiological processes which drive the differences have not been comprehensively studied.

To profile the varied molecular pathways involved in its pathogenesis, I present a collection of single cell and spatial transcriptomic datasets, together spanning over 400,000 cells captured from upper and lower airway tissue sites from 62 individuals (asthma patients and healthy controls) across multiple molecular modalities (transcriptome, immune repertoire, protein expression). In addition to profiling the cell surface proteome, I developed a technique to interrogate antigen binding in allergic asthma with house dust-mite sensitisation. This rich omics data resource is accompanied by detailed clinical phenotyping, thereby providing the largest single-cell atlas of asthma and its key phenotypes.

Analysis of this resource reveals that in patients with allergic sensitisation, contrary to past concepts of allergic asthma pathogenesis, IgE-producing plasma cells were not found in the lung. However, IgE-binding cells were detected (e.g. basophils and mast cells), suggesting that IgE may be entering from the periphery. Notably, the allergens thought to cause IgE secretion were found to bind to immune cells, including antigen-presenting cells, suggesting the IgE secretion may be localised to lymphoid tissues. Profiling the differences between asthmatic and healthy individuals in the spatial context reveals probable mast cell - plasma cell and $CD4^+$ - B cell niches, which may contribute to disease mechanism.

Differential abundance and gene expression analyses between the whole asthma cohort and healthy individuals showed features predominantly related to type 2 eosinophilic inflammation. Those changes correlated with clinical variables describing lung function, such as FEV_1 and FVC. In contrast, comparisons between the asthma phenotypes allowed identification of features associated with neutrophilic inflammation. These features included IL-1 signalling, antibacterial defence, and inflammasome activation.

My findings underscore critical importance of careful phenotyping to uncover underlying molecular changes in asthma, as well as providing a rich resource to study it further.

STATEMENT OF ORIGINALITY

I hereby declare that this Thesis is my own work and that the experiments described in the following pages were performed by myself, unless otherwise stated. The Thesis has not been submitted for any other degree or professional qualification.

Wojciech Lasoń

January 2025

ACKNOWLEDGEMENTS

In these final moments of writing and revisions of what constitutes a summary of my last four years at Oxford, I am feeling a sense of nervousness, excitement, but also accomplishment. I am full of gratitude for my time here and the people I have met, who made this time very special.

First and foremost, I want to thank my supervisors, Timothy Hinks, Calliope Dendrou, and Julian Knight, for embarking with me on this academic journey, providing support, engaging in discussions, and showing me what it means to be a scientist. Special thanks are owed to Callope Dendrou for helping me at some of the most challenging times of my DPhil studies, providing guidance and reassurance that everything will work out, and for the many evenings spent discussing my data.

I met some of the brightest minds studying here. Thank you to Devika Agarwal, the most amazing bioinformatician, and the kindest heart, for enduring my mood swings and the cynicism, and for showing me how to do amazing science with code.

St Peter's College has become a home away from home. I want to thank Ashita Alag, a colleague turned a dear friend, for teaching me how to be just, brave, and kind. Thank you for keeping my chakras aligned when times were rough. Thanks are also due to Tim Mawson for supporting us both at St Peter's College, and to the wider College staff.

"Thank you to the Wolfson College Boat Club rowing friends – James Neenan, Adam Ferris, Jorn Reniers, Kit Boyett, Tom Monahan, and Sam Wiese – for welcoming me into the club with open hearts and giving me the chance to become an athlete, something I had never experienced before in my life

Thank you to Dalia Gala for her nonsense-free approach to academia, for the mentorship, and for the giggles. You are dearly missed at Oxford.

I want to thank my lifelong friend, Asia Nowaczyk, for always believing in me, and for sticking around for what has been ten years of a beautiful friendship. Here is to another 100 years.

I would not be here were it not for my first chemistry teacher, Krystyna Harańczyk, who encouraged a little boy in a small town in the south of Poland to dream big, and generously gifted many, many hours of her time to show me the beautiful world of atoms and molecules.

My special thanks are owed to William Bartlett, the smartest costume designer there is, for offering me kindness, teaching me self-love, and for the amazing memories we have made together over the last three years. You are special.

Finally, I want to thank my amazing family – my mum Katarzyna Lasoń, my dad Marcin Lasoń, my grandma Lucyna Paluch, and my godmother Anna Bieda – for offering me the unconditional love and support, which enabled me to thrive. Thank you for the books, chemistry sets, microscopes, and for indulging my love for science. I am beyond grateful for everything you have done for me, and I will never be able to repay you.

IMPACT OF COVID-19 ON MY WORK

The experimental projects described in Chapters 3 to 5 depended on the collection of lung samples obtained during research bronchoscopies, which are invasive clinical procedures. The COVID-19 pandemic significantly impacted this work, as due to the SARS-CoV2 being a respiratory virus, and the social distancing regulations, it became significantly more difficult, and at times impossible, to perform bronchoscopies. Additionally, global supply chain disruptions delayed reagent shipments, with some single cell reagents being completely unavailable. These factors collectively delayed the sample collection for my main project by approximately 24 months (planned sample collection completion: end of 2022; actual completion: end of 2024), significantly shortening the planned time available for analysis from 2 years to approximately 1 year. In addition, unavailability of samples meant that I could not run the required optimisation experiments (sample processing and staining optimisation, scATACseq and scRNAseq protocol optimisation, development of QC and library preparation methods on a liquid handling robot). Due to the delays in sample collection, the data presented in Chapter 5 represents a subset of the planned dataset, which was originally intended to include 60 patients, but was limited to the samples available as of January 2024, so that sufficient time remained for analysis.

During the pandemic, I contributed to many other projects taking place in the group – I was involved in optimisation, sample extraction, library preparation, and sequencing of blood samples for a GWAS project in asthma, as well as optimisation, extraction, library preparation, and sequencing of RNA samples obtained from the nasal transcriptome of patients with SARS-CoV-2 infection. In addition, I worked on the Office of National Statistics' COVID-19 Infection Survey project and RECOVERY clinical trial taking place at the John Radcliffe Hospital. The Infection Survey generated COVID-19 infection statistics reported to the Government's Scientific Advisory Group for Emergencies. I was responsible for handling blood and serum samples and operation of the automated liquid handlers. My involvement led to being listed as a co-author in two publications:

Convalescent plasma in patients admitted to hospital with COVID-19 (RECOVERY):
a randomised controlled, open-label, platform trial
RECOVERY Collaborative Group, *Lancet*, 2021

Casirivimab and imdevimab in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial
RECOVERY Collaborative Group, *Lancet*, 2022

PUBLICATIONS

During my time at Oxford, I was involved in many interdisciplinary collaborations, and contributed to the following publications:

Panpipes: a pipeline for multiomic single-cell and spatial transcriptomic data analysis
Fabiola Curion, Charlotte Rich-Griffin, Devika Agarwal, Sarah Ouologuem, Kevin Rue-Albrecht, Lilly May, Giulia Garcia, Lukas Heumos, Tom Thomas, Wojciech Lason, David Sims, Fabian Theis, Calliope Dendrou
Genome Biology, 2024

Contributed to software development and testing.

MAIT cells protect against sterile lung injury
Xiawei Zhang, Shuailin Li, Wojciech Lason, Maria Greco, Paul Klenerman, Timothy SC Hinks
Cell Reports, 2025

Performed and helped with the design of single-cell experiments.

THESIS CONTRIBUTIONS

I am grateful to all my collaborators, without whom the work presented in this Thesis would not have been possible:

Contribution	Contributors
Patient recruitment for the clinical studies, sample acquisition, provision of clinical metadata, and definition of patient phenotype groups	Timothy Hinks Maisha Jabeen James Melhorn Anastasia Fries Gabriel Laovie Simon Couillard
Assistance with sample processing for single cell assays on the day of the bronchoscopy	Maisha Jabeen Bianka Brassenyi
Cell sorting service	Helen Ferry Paul Sopp Joanna Hester
Assistance with sequencing library preparation, quality control, and automation of the sequencing process using the liquid handling robot	Maria Greco Bianka Brassenyi MRC WIMM Advanced Single Cell Omics Facility
Tissue sectioning and placement on Xenium slides and H&E staining of reference slides	Esther Bridges Jeanette Norman
Sequencing service	Novogene UK Ltd

LIST OF ABBREVIATIONS

7-AAD	7-Aminoactinomycin D
APC	Antigen Presenting Cell
ASM	Airway Smooth Muscle
BAL	Bronchioalveolar Lavage
CD	Cluster of Differentiation[number]
cDC	Conventional Dendritic Cell
COPD	Chronic Obstructive Pulmonary Disease
DNA	Deoxyribonucleic Acid
EMT	Epithelial-Mesenchymal Transition
FeNO	Fraction of Exhaled Nitric Oxide
FEV1	Forced Expiratory Volume in 1 second
FVC	Forced Vital Capacity
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
GWAS	Genome-Wide Association Study
H&E	Haematoxylin and Eosin
HDM	House Dust Mite
HLA	Human Leukocyte Antigen
HLCA	Human Lung Cell Atlas
HVG	Highly Variable Gene
ICS	Inhaled Corticosteroid
IFN	Interferon
Ig	Immunoglobulin

IL	Interleukin
LISI	Local Inverse Simpson's Index
mRNA	Messenger RNA
NF	Nuclear Factor
PNEC	Pulmonary Neuroendocrine Cell
pDC	Plasmocytoid Dendritic Cell
PC	Principal Component
PCA	Principal Component Analysis
RBC	Red Blood Cell
RNA	Ribonucleic Acid
RPMI	Roswell Park Memorial Institute (cell culture medium)
SD	Standard Deviation
SMG	Sub-Mucosal Gland
TGF	Tumour Growth Factor
Th	T-Helper Cell
TLR	Toll-Like Receptor
TNF	Tumour Necrosis Factor
UMAP	Uniform Manifold Approximation and Projection for Dimension Reduction
UMI	Unique Molecular Identifier

LIST OF FIGURES

Figure 1-4, Figure 1-5, Figure 3-1, Figure 3-2, Figure 4-1, Figure 5-1, and Figure 5-3 created with BioRender.com.
Figure 4-5 generated with 10x Genomics Xenium Explorer 3.2.0

Gene names in Figures refer to RNA transcripts unless otherwise stated.

LIST OF TABLES

LIST OF CONTENTS

1 INTRODUCTION

1.1 INTRODUCTION TO ASTHMA

Asthma is a heterogenous inflammatory disease of the airways which affects an estimated 262 million people worldwide. The UK is amongst the countries with the highest prevalence of asthma and allergy in the world, with approximately 4.3 million (6.5% of the adult population) affected [1,2]. Its hallmark features are chronic airway inflammation, which manifests itself clinically through cough, wheeze, shortness of breath, and feeling of tightness in the chest. These result from the overproduction and thickening of mucus, airway wall remodelling and airway narrowing at the microscopic scale [3,4].

In a proportion of patients, asthma symptoms can be controlled by a combination of an inhaled corticosteroid (which suppresses many aspects of the inflammation) and a short- or long-acting β_2-adrenergic agonist (which relaxes the smooth muscle, thus opening the airway) [3]. However, approximately 10% of adults and 2.5% of children with asthma present with severe disease, with a reduced quality of life and an increased risk of exacerbations, as well as complications, including the need for hospitalisation and the risk of death [5]. A proportion of these hospital admissions are caused by viral infection of the respiratory tract, most commonly with rhinovirus [3].

This population of severe asthmatics, which either requires repeated use of corticosteroids or, as we are discovering using fractional exhaled nitric oxide (FeNO) measurements, does not respond to these drugs [6], would likely benefit from targeted immunologic therapy specific to their disease biology. This is exemplified by the success of monoclonal antibodies against specific components of type 2 airways inflammation [7]. Whilst the first licenced biologic, omalizumab, targeted IgE, much greater efficacy and broader application have come with biologics targeting interleukin (IL)-5 and IL-5RA in eosinophilic asthma: mepolizumab, reslizumab and benralizumab [7]. More recent biologics target the IL-4R-alpha (dupilumab) and alarmins, particularly thymic stromal lymphopoietin (TSLP; tezepelumab). However, these biologics are costly, require lifelong therapy, and are not effective for all phenotypes, so novel approaches to targeted treatment for asthma therefore remain an urgent research priority.

In addition to patient suffering, on the population scale, asthma healthcare incurs a large direct cost, with estimates of annual cost of £1.1 billion in the UK and $18 billion in the USA [2,3], in addition to secondary economic costs from lost productivity.

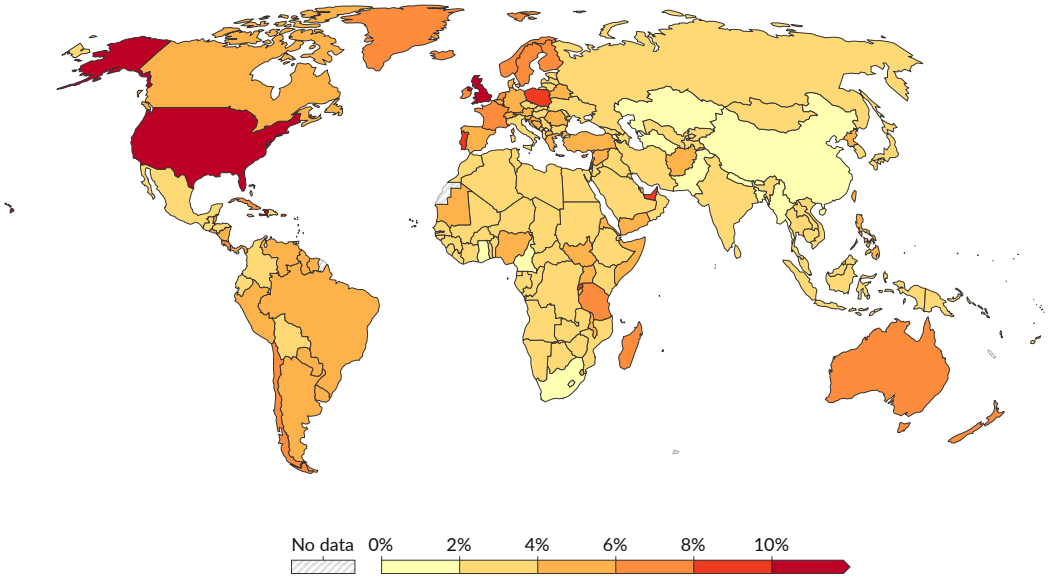

Figure 1-1 Global prevalence of asthma in 2021
The share of the country's population with asthma. Disease prevalence is age-standardized to account for changes in the age structure of a population over time and between countries.
Data source: IHME, Global Burden of Disease (2024) with major processing by Our World in Data, reproduced under Creative Commons license CC BY

1.2 EFFECTS OF GENOTYPE AND ENVIRONMENT ON PATHOGENESIS OF ASTHMA

Asthma incidence has been steadily growing over the past 100 years [8,9], with the number of individuals affected continuing to grow as the global population increases. Effective targeted therapies are relatively costly and of scarce availability [7], leading to high mortality in countries with low sociodemographic index [10].

Asthma is thought to be triggered by a variety of environmental factors acting on the epithelium, such as allergens, air pollutants (particulate matter and gaseous pollutants e.g. ozone, nitrogen dioxide, and sulphur dioxide, and reactive oxygen species) [11,12], as well as microbes and viruses [4]. The air pollutants in particular increase in prevalence with urbanization [12], however, animal allergens (e.g.

mouse) are found in inner-city homes at log-fold higher concentrations than in suburban areas (100-1000x)- levels known to elicit symptoms in animal facility workers [13,14].

In addition to the rising level of pollution and socioeconomic inequality, climate change has also been shown to exacerbate asthma [15]. Indeed, it appears that extreme temperatures can trigger new asthma cases: based on the data from the last 30 years, it is estimated that with each increase of maximum air temperature variability by 1°C the global risk of asthma rises by 5.0% [16]. The maximum variability takes into account both extreme cold and heat, and the increase in asthma incidence was predicted to be particularly localised to regions of high altitude.

1.3 HETEROGENEITY OF ASTHMA

Asthma is now thought to be a collection of common symptoms resulting from varied molecular mechanisms leading to airway inflammation, rather than a uniform disease unit with well-defined pathophysiology. It has been suggested that 'asthma' as a disease is an outdated term and more detailed definitions supported by clinically measurable variables and biologically sound mechanisms should be encouraged instead [17]. Observable asthma characteristics form a phenotype, defined as "observable properties produced by the interactions of the genotype and the environment", whereas the endotype is a mechanistically uniform phenotype where a specific biological pathway driving the disease has been identified at

cellular and molecular level [18,19]. Several asthma endotypes have been proposed, including classification by driving cell type [20] and cytokine [21], however no broad consensus exists and the definitions are evolving with our understanding of the disease. In clinical practice, phenotypes based on observable traits such as triggers of inflammation (allergens, viruses, chronic conditions), abundance of cell types, and markers of inflammation such as exhaled nitric oxide, are used to inform diagnosis and treatment [6,18,19]. These tend to be dualistic discriminators (e.g. allergic versus non-allergic, eosinophilic versus non-eosinophilic, adult-onset versus childhood-onset).

1.3.1 GENETIC INFLUENCES IN ASTHMA

It is agreed that asthma likely develops as a result of a combination of genetic predisposition [22–24] and exposure to environmental factors, such as the previously mentioned pollutants and allergens [4,12]. However, relative contribution of each of those factors remains poorly understood and is an active research area. Estimates of heritability in asthma range from 35% and 95% [25], with most twin studies placing estimates at 70-85% [26–28][26–28].

The first genome-wide association study (GWAS) of asthma identified the *ORMDL3* candidate gene in allergic eosinophilic asthma [25,29], a member of a gene cluster encoding transmembrane proteins located in the endoplasmic reticulum. Although the exact pathophysiological role of ORMDL3 in asthma remains disputed, recent hypotheses propose that ORMDL3 acts to upregulate ceramide levels which

exacerbate inflammatory responses, and lead to mucus overproduction and airway hyperresponsiveness [30]. Since then, many other genes at the same GWAS locus have also been implicated, most prominently *GSDMB*, encoding gasdermin B, which upregulates mediators of airway remodelling *in vitro* (*TGFB1*, *ALOX5*, *MMP9*) and induces airway hyperresponsiveness [31]. Due to the effect of linkage disequilibrium on the SNPs in the *ORMDL3* locus and co-regulation of the expression of many of the genes in the locus, it is not possible to delineate between the effects of these genes in GWAS studies, which are commonly referred to as the 17q21 asthma locus, and indeed effects on more than one gene may drive the association signal at this locus [32]. For example, interestingly, the *CDHR3* gene also located at the same locus (17q12–21) was reported to be the elusive receptor for rhinovirus (RV)-C, a virus thought to cause asthma exacerbations [33]. Since the first GWAS study, more than 3,000 SNPs have been implicated in asthma (between 3,291 [34] and 3,691 [22]) across 194 further GWAS studies, identifying 179 independent asthma-associated loci [35], with a subset of key loci summarised in Table 1-1.

Functional validation of GWAS findings and the identification of causal genes and signalling pathways remain significant challenges. Among the loci that have been biochemically characterized, several are implicated in key inflammatory pathways, including IL-4, IL-13, and IL-33, which are associated with inflammation; antigen presentation, mediated by HLA genes; cell activation, through CD28; and innate immune responses to the lung microbiota via TLR1 and TLR6 [35], as seen in in Table

1-1. Transcription factors such as GATA3 and STAT5, which regulate Th2/ILC2 cell differentiation and broader transcriptional shifts, have also been highlighted [35,36].

Additionally, pathways related to airway structure and remodelling have emerged as critical, including mechanisms of epithelial cell homeostasis mediated by MUC5AC, which drives mucus production and goblet cell metaplasia [37]. Enzymes released by mast cells, such as tryptases, chymases, and carboxypeptidase A3, further contribute to airway remodelling [38]. Collectively, the characterised genes provide evidence of a central role for epithelial-immune interactions as key drivers of asthma pathophysiology.

Notably, the largest GWASes of asthma to date have predominantly relied on a broad asthma diagnosis [39,40]. Smaller studies have sought to address disease heterogeneity by refining inclusion criteria, such as by comparing adult-onset to childhood-onset asthma [41] or using the number of asthma exacerbations as a proxy for disease severity and limiting the population to childhood-onset cases [42]. Whilst these studies have offered some supporting evidence for genetic associations more relevant to specific asthma subtypes, they have generally failed to identify many novel, phenotype-specific genetic associations [35]. It is therefore proposed that future studies should incorporate asthma subtypes defined through clinical phenotyping and/or biomarker measurements. This approach may facilitate the identification of new genetic drivers specific to asthma subgroups, enabling a deeper understanding of the disease and development of personalized treatments [35,40].

Table 1-1 GWAS regions with 10 or more asthma-associated loci

Chromosomal region	Associations asthma* (# studies)	Associations allergy (# studies)	Genes
6p21.3	28 (10)	6 (5)	**HLA-DQB1**, **HLA-DRB1**, **HLA-DRB6**, **HLA-DQA1**, **HLA-DPA1**, **HLA-B**, **HLA-C**, MICA, MICB, COL11A1, TCP11, SCUBE3, **HLA-DOB**, HCP5, MCCD1
10p14	15 (7)	3 (2)	**GATA3**, CELF2, SFTA1P, LOC101928272, RP11
2q12.1	13 (7)	2 (2)	**IL1R1**, **IL1RL1**, IL1RL2, IL18R1, IL18RAP, MIR4772, SLC9A2, SLC9A4
5q22.1	12 (7)	3 (2)	CAMK4, WDR36, SLC25A46, TMEM232, **TSLP**
5q31.1	11 (9)	1 (1)	C5ORF56, SLC22A5, IRF1, KIF3A, **IL4**, CCNI2, **IL13**, RAD50, SEPT8
9p24.1	11 (8)	3 (2)	RANBP6, **IL33**, KIAA2026, MIR4665, TPD52L3, GLDC, UHRF2
17q12–21	11 (7)	1 (1)	**ORMDL3**, ZPBP2, ERBB2, MED1, CSF3, ERBB2, GRB7, GSDMA, **GSDMB**, IKZF3, LRRC3C, MED24, MIEN1, MIR4728, MIR6884, PGAP3, PNMT, PSMD3, SNORD124, STARD3, TCAP
11q13.5	10 (7)	3 (3)	LRRC32, C11ORF30, EMSY, THAP12, WNT11, PRKRIR
15q22.33	10 (6)	1 (1)	SMAD3, SMAD6, AAGAB

Chromosomal locations of 28 loci associated with asthma or allergy in at least 5 GWASes of asthma or at least two GWASes of an allergic disease. Regions with 10 or more associations are listed. Genes listed were reported in each study for the association, with genes known to be implicated in immune processes in asthma in bold.

*Number includes multiple independent associations at the same locus.

1.4 CLINICAL PHENOTYPES OF ASTHMA

1.4.1 ALLERGIC AND NON-ALLERGIC ASTHMA (EARLY AND LATE ONSET ASTHMA)

Early attempts to define clinical asthma phenotypes in the 1940s led to a broad classification of allergic ('extrinsic') and non-allergic ('intrinsic') asthma [43]. Allergic asthma could be characterised by early age of onset and sensitization to an identifiable allergen, as defined by production of IgE specific to that allergen (e.g. dust, pollen) and positive skin prick test, and possibly co-incidence of other autoimmune diseases, such as eczema or allergic rhinitis. Conversely, non-allergic asthma was thought to develop later in life (past 40 years of age), often coincided with aspirin-exacerbated respiratory disease, but was mechanistically poorly understood [18,43].

1.4.2 TYPE 2-HIGH AND -LOW ASTHMA (EOSINOPHILIC AND NEUTROPHILIC ASTHMA)

In the 1990s, we began to understand that in many asthmatics, chronic airway inflammation is driven by cytokines IL-4, IL-5, and IL-13 [4,44,45]. These are commonly referred to as "type 2" cytokines, as they are mostly produced by Th2 cells and ILC2s[4,45].

Type 2 cytokines promote hallmark features of allergic asthma such as eosinophilia, mucus hypersecretion, and IgE production, and this "type 2-high" asthma

phenotype can be targeted with inhaled corticosteroids [4,21]. High sputum eosinophil count is a common diagnostic trait of “type 2-high” asthma [46]. However, only about 50-60% of asthmatics are diagnosed with the “type 2-high” steroid-responsive disease [21], whereas the remaining patients suffer from the “type 2-low” asthma is more associated with increased neutrophil counts, smoking, obesity, and environmental factors changing the inflammatory landscape of the body and influencing cytokine signalling in immune cells [21,47,48]. Due to these changes in inflammation signalling and emergence of neutrophils as key drivers of inflammation, the “type 2-low” patients are not responsive to steroid treatment [4,21]. It now emerges that the resistance to treatment is likely mediated by, amongst other mechanisms, gastrointestinal hormones leptin and adiponectin [47,48]; given this with the administration of novel anti-obesity treatments may offer these patients a new therapeutic avenue [49].

1.4.3 FENO SUPPRESSORS AND NON-SUPPRESSORS

FeNO, typically reported in ppb (parts per billion), can quickly and non-invasively be measured in the clinic [50]. It is a marker of type 2 inflammation and its levels are elevated in the air exhaled by “type 2-high" asthmatics, with its levels also associated with increased likelihood of exacerbations [6]. FeNO measurement over time whilst administering steroid treatment (“FeNO suppression test”) can be employed to distinguish between patients who are non-adherent to treatment and those who do not respond to it. FeNO measurements decrease over time in non-adherent patients as they use a microchip-enabled inhaler and the steroid

treatment reduces inflammation in the airway [51]; these patients are referred to as "FeNO suppressors". Conversely, "FeNO non-suppressors" are patients with severe asthma who are resistant to steroid treatment and may benefit from biological therapies [50,51]. The mechanisms of steroid resistance remain poorly understood and are of therapeutic interest.

1.4.4 OTHER CLINICAL ASTHMA PHENOTYPES

Beyond the phenotype definitions adopted in this thesis, it should be pointed out for completeness that other classifications of asthma have also been proposed, which are either a subset or fall outside of our defined phenotypes. These are briefly summarised below.

1.4.4.1 EXERCISE-INDUCED ASTHMA

Exercise-induced asthma is a poorly characterised phenotype defined as bronchoconstriction experienced in response to sustained exercise, particularly in dry and cold conditions [18], and as a result of chlorine exposure in swimmers [52]. There is evidence which supports "type 2-high" eosinophilic origin [53] as well as "type 2-low" inflammation driven by neutrophils and macrophages [54]. A few mechanisms are proposed including: lowering of airway temperature as a result of increased ventilation, with bronchial blood vessels expanding and causing bronchial wall oedema and subsequently bronchoconstriction as the airway returns to normal temperature; increased ventilation causing dehydration of the

airways and subsequent release of inflammatory mediators; and stress of increased ventilation causing injury to the airway epithelium leading to airway hyperresponsiveness [55].

1.4.4.2 OBESITY-RELATED ASTHMA

Until recently, the causal link between asthma and obesity could not be established, with experts debating whether obesity is a comorbidity or part of pathological mechanism of asthma [18,19]. The disease exhibits sex differences, with higher prevalence in females (in adults) [56]. The disease is non-eosinophilic, and as previously discussed, signalling seems to be driven by neutrophils [47], but the inflammation landscape is distinct, with Th17 and ILC3 cells being implicated [57], which is supported by the immunomodulatory effect on ILCs and the CD4+ T cell compartment by adipose tissue [58].

1.5 KEY CLINICAL VARIABLES IMPORTANT IN ASTHMA

As outlined in section 1.3, inclusion of clinical phenotypes and relevant biomarkers in new studies of asthma seems necessary to discover key drivers of each subtype, as asthma is a heterogenous disease. Experimental variables available for measurement in asthma clinical studies are summarised in Szefler *et al.* [59], with key variables measured in the clinic for the studies I carried out listed below.

1.5.1 TOTAL IGE AND ALLERGEN-SPECIFIC IGE

Both total IgE levels and the presence of allergen-specific IgE antibodies in serum are important biomarkers for characterizing the patient's asthma phenotype, however, the reported correlations of treatment with either of these variables are varied [60,61].

An increase in total serum IgE has been associated with asthma [62]. However, considerable overlap in IgE levels between atopic and non-atopic populations limits its utility in identifying atopy [59]. In addition, whilst total IgE is useful in identifying patients who could benefit from anti-IgE therapy, it cannot predict the degree of response to treatment [61].

In contrast, an increase in IgE specific to aeroallergens is more diagnostically relevant and defines a patient as having allergic asthma phenotype [18,19], with levels of specific IgE in serum inversely correlating with lung function parameters [63]. Screening for sensitization to common aeroallergens aids in diagnosing atopy and improves disease management by allowing the patient to avoid allergen exposure [59].

1.5.2 EOSINOPHIL COUNT

The association between increased blood eosinophil counts and asthma was observed over a 100 years ago, shortly after this cell type was first described [64]. While not all patients with asthma exhibit blood eosinophilia [21], those who do

tend to experience greater airway remodelling and more frequent exacerbations [65]. In contrast, patients without eosinophilia are more likely to present with increased airway obstruction [65]. In adults, blood eosinophil counts range from 0.015 to 0.65 × 10^9/L [66], with 0.30 × 10^9/L as the cut-off recommended by NICE to start biological treatment of eosinophilic asthma [67]

Sputum eosinophil count can also be used to distinguish between patients with eosinophilic and non-eosinophilic asthma, with greater relevance to airway inflammation, but also increased variability due to the multiple experimental protocols in use [59].

1.5.3 EXHALED FRACTIONAL NITRIC OXIDE (FENO)

FeNO is a reproducible and easily measurable biomarker that serves as a reliable predictor of response to steroid treatment [6] and is increased in patients with type 2 inflammation [50,61]. However, its levels can be influenced by several confounding factors, including demographics, smoking, atopy, and diet [61]. According to the American Thoracic Society guidelines, FeNO measurements above 50 ppb indicate eosinophilic inflammation, while values below 25 ppb suggest that eosinophilic inflammation and responsiveness to corticosteroids are less likely [68].

1.5.4 PULMONARY FUNCTION MEASUREMENTS

Lung function is readily assessed in clinical settings using spirometry, which measures several parameters based on the volume of air exhaled during a forced respiratory manoeuvre. These include FEV_1 (forced expiratory volume in 1 second), FVC (forced vital capacity), and their ratio (FEV_1/FVC), as summarized in Figure 1-2. A reduced FEV_1/FVC ratio indicates airflow obstruction. In contrast, a reduced FVC with a preserved FEV_1/FVC ratio suggests a reduced capacity for lung expansion, as seen in diseases such as idiopathic pulmonary fibrosis [35,69].

These lung function parameters have been reported to have significant associations with asthma, and for example, a putative missense *GATA5* variant is associated with lower FEV_1 and increased risk of asthma [22].

1.6 APPROACHES TO TREATMENT OF ASTHMA

“Type2-high" inflammation in asthma can be treated by corticosteroids in FeNO suppressors, resulting in a decrease in FeNO (regulated by IL-13 signalling) when inhaled corticosteroids are prescribed, and a decrease in blood eosinophil counts (regulated by IL5- signalling) with the use of oral corticosteroids [7]. The FeNO non-suppressor group of patients does not respond to steroid treatment, with FeNO and eosinophil count biomarkers remaining high despite high doses of oral or inhaled corticosteroids [50,70]. Despite that, systemic (oral) corticosteroids remain the first line of treatment and are commonly used for the management of severe asthma, as

either short-term courses or long-term daily regimens [7]. Regardless of the heterogeneity of responses, these patients remain included in the treatment guidelines, owing partly to high accessibility, low cost, and familiarity [71]. However, even at low doses, they are associated with acute and chronic adverse effects, such as pneumonia and increased susceptibility to infection [71]. For example, while asthma is not associated with an increased severity of COVID-19 [72], the use of oral glucocorticoids for uncontrolled severe asthma has been linked to increased COVID-19–related mortality [73]. Beyond steroid treatment, a number of biologic treatments are now available to patients with severe asthma and high levels of eosinophilic inflammation, which target well characterised cytokines and signalling molecules in asthma, predominantly type 2 cytokines and alarmins. Approved biological treatments for asthma are summarised in Table 1-2.

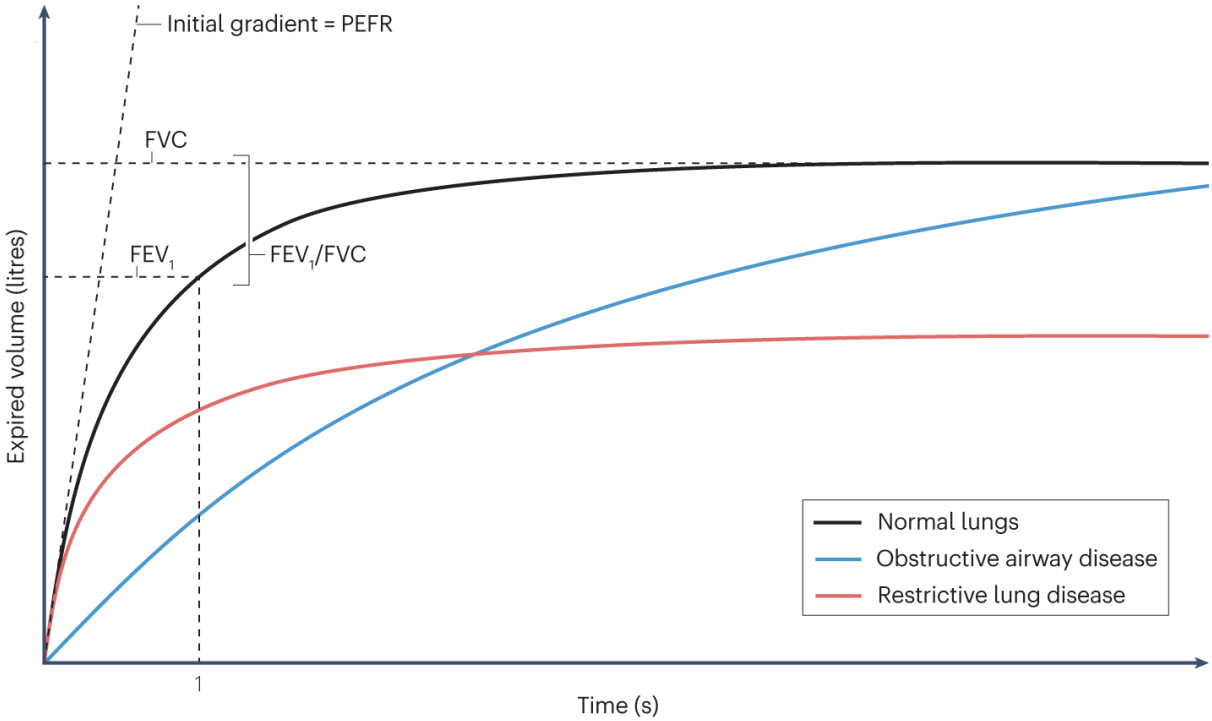

Figure 1-2 Forced expired volumes and lung disease

Spirometry is a clinical test which measures the expired volume of air over time. Forced expiratory volumes – FEV1, the volume of air expelled from the lungs in one second; and forced vital capacity, FVC, is the total volume of air expelled from the lungs; and their ratio (FEV1/FVC) – decrease in patients with obstructive airway disease (blue) or restrictive lung disease (red) compared to values obtained in healthy individuals (black).

PEFR = peak expiratory flow rate (PEFR)

Table 1-2 Biological treatments in asthma

Monoclonal antibody	Antibody target and mechanism of action	Asthma phenotype indication	Efficacy	Safety concerns
Benralizumab	interleukin-5Rα; antibody binds to interleukin-5Rα on eosinophils and basophils, depleting them through antibody-dependent, cell-mediated cytotoxicity	Severe eosinophilic asthma	Reduced exacerbations, reduced symptoms, small or moderate effect on FEV1; decrease or withdrawal of steroids if blood eosinophils >150/µl; improved quality of life	Helminthic infections, hypersensitivity reactions, abrupt discontinuation of steroids
Dupilumab	interleukin-4Rα; antibody binds to interleukin-4Rα, inhibiting interleukin-4 and interleukin-13 signalling in immune cells (B and T cells, eosinophils, epithelial cells, and airway smooth-muscle cells	Severe eosinophilic asthma (FDA), severe type-2 high asthma (EMA), steroid-dependent asthma	Reduced exacerbations, reduced symptoms, improved lung function; decrease or withdrawal of steroid treatment, irrespective of blood eosinophil count at baseline; improved quality of life	Helminthic infections, hypersensitivity reactions, abrupt discontinuation of steroids, hypereosinophilic conditions, conjunctivitis
Mepolizumab	interleukin-5; antibody binds to circulating interleukin-5	Severe eosinophilic asthma	Reduced exacerbations, reduced symptoms, small or moderate effect on FEV1; reduction or withdrawal of OGs if blood eosinophils >150/µl; improved quality of life	Helminthic infections, hypersensitivity reactions, abrupt discontinuation of steroids, herpes zoster infections (rare)

Omalizumab	IgE; antibody binds to Fc part of free IgE, inhibiting binding of IgE to FcεRI on mast cells and basophils and FcεRII on dendritic cells and eosinophils	Severe allergic asthma	Reduced exacerbations, reduced symptoms, small effect on FEV1; improved quality of life	Serum sickness, hypereosinophilic conditions, abrupt discontinuation of steroids; black-box warning for anaphylaxis (occurring in ±0.2% of patients)
Reslizumab	interleukin-5; antibody binds to circulating interleukin-5	Severe eosinophilic asthma	Reduced exacerbations, reduced symptoms, small or moderate effect on FEV1; improved quality of life	Helminthic infections, abrupt discontinuation of steroids; black-box warning for anaphylaxis (occurring in ±0.3% of patients)
Tezepelumab	thymic stromal lymphopoietin; antibody binds to TSLP, inhibiting signalling in a wide array of immune and stromal cells (dendritic cells, T and B cells, NK cells, regulatory T cells, ILC2s, eosinophils, basophils, mast cells, monocytes, macrophages, airway smooth muscle cells, and fibroblasts)	Severe asthma	Reduced exacerbations, reduced symptoms, improved lung function; improved quality of life	Pharyngitis, arthralgia, back pain
Reproduced with permission from Brusselle & Koppelman (2022). Copyright Massachusetts Medical Society.				

1.7 CELL TYPES IN THE RESPIRATORY SYSTEM

The respiratory system's primary role is maintaining physiological homeostasis by facilitating the exchange of gases necessary for aerobic metabolism. It supplies oxygen to the bloodstream while eliminating carbon dioxide, a by-product of cellular respiration, through its interaction with the cardiovascular system. Anatomically, the respiratory system is divided into the upper and lower respiratory tracts. The upper respiratory tract includes the nose, nasal cavity, oral cavity, pharynx, and larynx. The lower respiratory tract comprises the trachea and lungs, which are further subdivided into terminal and respiratory airways, as well as the alveoli. The alveoli serve as the primary site of gas exchange, facilitating the diffusion of oxygen and carbon dioxide across the alveolar-capillary membrane [74]. The structure of the lower airway is summarised in Figure 1-3A.

The respiratory tract is lined with epithelial cells, supported by stromal cells, and interspersed with immune cells whose type and distribution vary across anatomical regions and tissue layers. Notable differences exist between the upper and lower airways, as well as between the epithelial and stromal compartments. The dynamic interplay among epithelial, stromal, and immune cells plays a key role in regulating airway function and immune responses. Due to the number of cell types involved, I will characterise the immune and non-immune (stromal and epithelial) compartments separately.

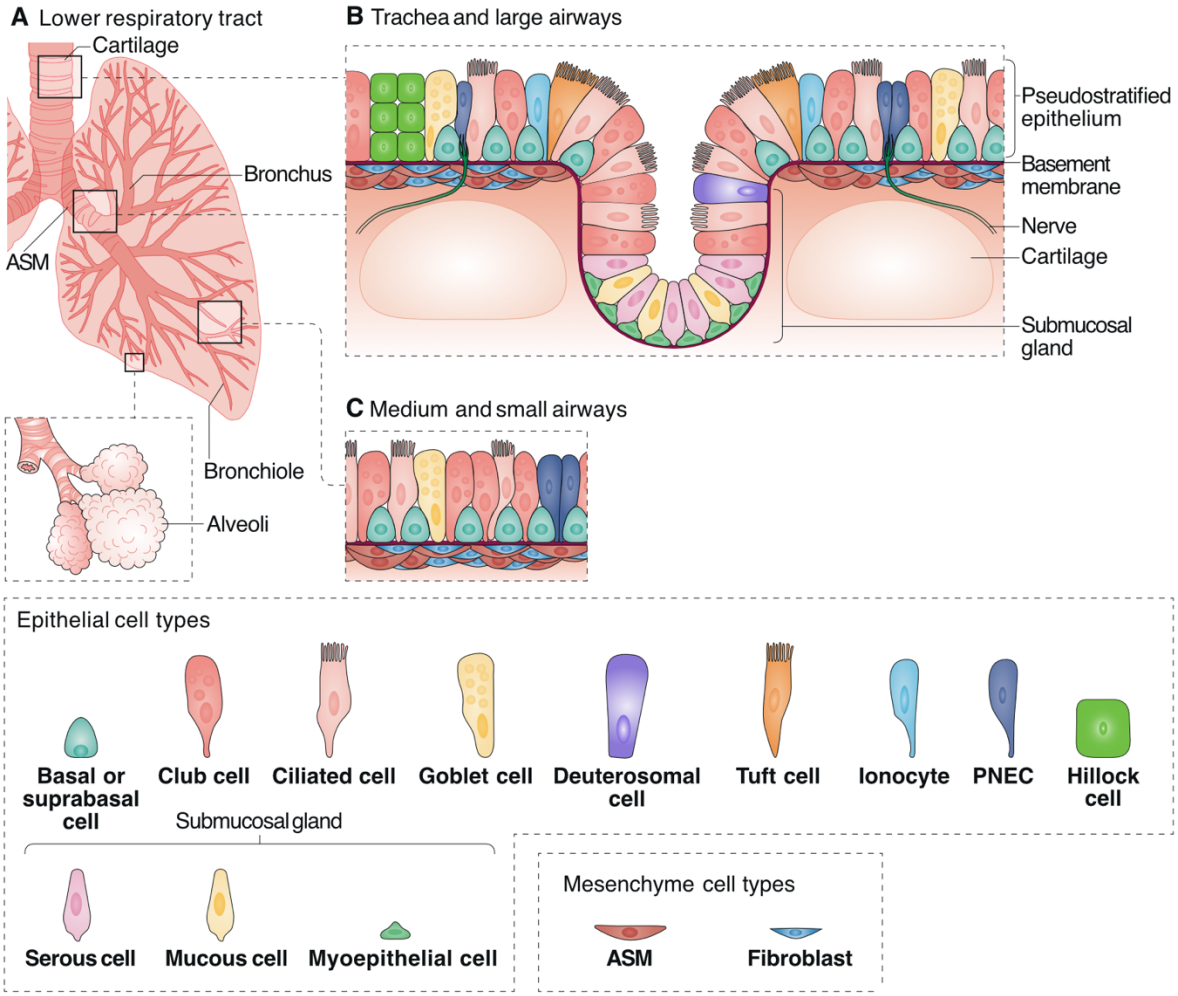

Figure 1-3 Anatomy of the airways and known cell types in large and small airways

A The lower respiratory tract consists of the *respiratory tree*, which comprises large and small airways that branch and terminate in alveoli. The airway tubes are surrounded by mesenchymal tissue, containing airway smooth muscle supported by fibroblasts.

B-C The respiratory epithelium comprises diverse epithelial cell types with specialized functions. Goblet cells secrete mucous to trap inhaled particulates, while ciliated cells expel it along with trapped particles. Deuterosomal cells generate ciliary apparatus and serve as progenitors for ciliated cells. Club cells secrete protective and immunomodulatory factors and aid in detoxification, whereas basal and suprabasal cells act as progenitors for the airway epithelium. Pulmonary neuroendocrine cells regulate smooth muscle tone and immune responses. Rare cell types include tuft cells, involved in allergen-driven type-2 immune responses, and ionocytes, responsible for ion transport via CFTR channels. Stromal cells – airway smooth muscle and fibroblasts – control epithelial turnover and airway diameter. The trachea and large airways feature cartilaginous rings and submucosal glands, which contain mucous, serous, and myoepithelial cells, contributing to mucus production alongside goblet cells.

ASM = airway smooth muscle, PNEC = pulmonary neuroendocrine cell

1.7.1 EPITHELIAL COMPARTMENT

The trachea and proximal airways (Figure 1-3A) are lined with pseudostratified epithelium—a single layer of epithelial cells in which all cells contact the basal lamina, but their nuclei are positioned at different levels, creating the appearance of multiple layers. This epithelium comprises diverse lineages, classically defined as ciliated cells, secretory cells, goblet cells, and basal cells, with new, rare cell types recently discovered [75,76].

The anatomical origin of cell populations within the respiratory tract is now recognized as a key determinant of their transcriptional signatures. Single-cell atlases have revealed distinct region-specific populations of suprabasal, secretory, and ciliated cells between the upper and lower airways [77,78].

1.7.1.1 MAJOR EPITHELIAL CELL TYPES

As a critical component of the airway, the respiratory epithelium serves as a barrier that protects the host from environmental insults. It works in concert with the immune system to respond to microbes, toxins, pollutants, and allergens that breach the mucociliary barrier, playing a central role in both innate and adaptive immune defence [79]. Understanding how these environmental stimuli drive chronic lung diseases, such as asthma and COPD, is crucial to developing strategies to address these conditions.

Ciliated cells line the airway wall and have cilia on their apical surface that move inhaled particulates and mucus out of the airways, clearing them towards the upper airway for removal [74]. In addition to the airway cells, two mucus-producing cell types are present in the airway – goblet and club cells. Goblet cells secrete mucins and other glycoproteins that contribute to the viscosity of mucus, which plays a key role in innate immunity against inhaled particulates and lung-invading microbes [80]. The most prominent mucins are MUC5AC and MUC5B, which also serve as markers for these cells (Figure 1-4). Club cells secrete protease inhibitors and immunomodulatory substances such as SCGB1A1 (a club cell marker protein which suppresses inflammation by inhibiting phospholipase A2 and IFNγ signalling), and likely detoxify and eliminate exogeneous substances [80,81] (Figure 1-4). Recent studies have shown that club cells can detect apoptotic cells caused by allergens, pathogens, or pollutants through the AXL receptor and clear them via RAC1-dependent phagocytosis [79,82].

Basal cells act as a reservoir of quiescent multipotent stem cells and proliferating progenitors, replenishing both secretory and ciliated cells during homeostasis and following injury (Figure 1-4). Their abundant cytoskeletal and adhesive proteins anchor the epithelium to the matrix and protect the underlying stroma from the external environment [83].

The trachea and airways are lined with mesenchymal tissue, including cartilage surrounding the trachea, smooth muscle, and fibroblasts, which act as support and regulate airway diameter [74].

The alveoli, found in distal airways, differ structurally from the more proximal airways (Figure 1-3). The alveolar epithelium consists of two major cell types: alveolar type 1 (AT1) and alveolar type 2 (AT2) epithelial cells, also referred to as pneumocytes. AT1 cells are flattened, squamous cells that cover more than 95% of the lung's gas exchange surface. These cells form an intimate association with the underlying endothelial capillary plexus, creating the thin, gas-permeable interface essential for efficient diffusion. In contrast, AT2 cells are cuboidal and secrete pulmonary surfactant, which reduces surface tension and prevents alveolar collapse during respiration, and they are the stem cells of the alveolar epithelium [74,84]. These cells are not present in the typical bronchoscopy samples [85] and instead require material from deceased individuals or a surgery in order to be detected (Figure 1-5).

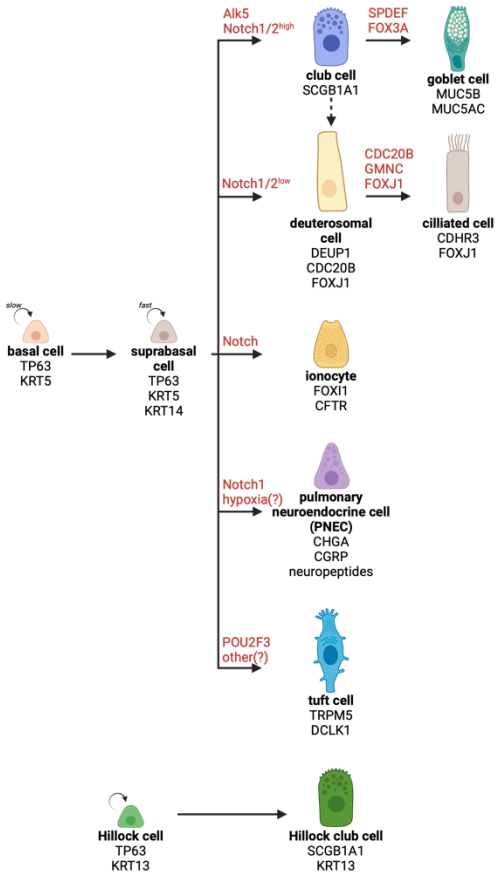

Figure 1-4 Airway epithelium differentiation pathways

Basal cells are the stem cells of the airways, and upon activation, increase the proliferation rate and are referred to as supra/parabasal. Notch and other signalling pathways determine further fate. Secretory cells (club and goblet) differentiate as a result of Notch2 signalling, whilst ciliated cells form through an intermediate, deuterosomal cell/deuterocyte, activating formation of centrioles. Rare cell types, such as ionocytes, PNECs, and tuft cells, also originate from basal cells, but the pathways are less well described. Recently characterised Hillocks form a separate population of stem cell reservoir within squamous epithelium.

Based on Davis & Wypych (2021), Ruysseveldt *et al*. (2021), and Montoro *et al* (2018); adapted and updated.

1.7.1.2 RARE AND NOVEL EPITHELIAL CELL TYPES

The traditional understanding of airway cellular composition has been significantly enhanced by single-cell omics profiling techniques, which have led to the identification of several novel cell types, such as ionocytes [86], deuterocytes [87], and hillock cells [88], as well as a population of sensory cells (PNECs), as seen in Figure 1-3 and Figure 1-4. It is hypothesised that there is also additional yet undiscovered heterogeneity within the epithelial compartment caused by the environmental factors, with unique cell states emerging due to smoking or exposure to allergens or pollutants [79].

In contrast to all other epithelial cells, PNECs are innervated and produce neurotransmitters, including amines and peptides. This enables them to function as sensory cells, signalling to the nervous system through secretory granules containing neuropeptides and neurotransmitters, and also influencing the immune response [89].

1.7.2 IMMUNE COMPARTMENT

The respiratory system is continuously exposed to environmental factors, airborne pathogens, and immunostimulatory molecules released by microbes [90,91]. It employs innate and adaptive immune mechanisms to fight these.

Most primitive defence mechanisms of the lung include the presence of antibacterial peptides (e.g. defensins), mannose binding protein, lysozyme, lactoperoxidase, and airway

surfactant [92]. These innate mechanisms are supplemented by immunoglobulins secreted by B and plasma cells – the airway mucous contains approximately equal concentrations of IgA and IgG in addition to smaller amounts of IgM [93]. In pathogenic responses such as asthma, IgE is also present [94].

Cellular mechanisms of innate immunity include the presence of phagocytic cells (alveolar macrophages, neutrophils, eosinophils) and natural killer cells, which recognize and remove bacteria and virus-infected host cells through PAMP and DAMP receptors [92]. Aberrant function of these cell types can contribute to the development of asthma [4].

Macrophages, interspersed through the respiratory system, act as immune sentinels of the airways, with the ability to uptake and present antigens and mediate Th1 immune response [95]. They can be divided into two distinct populations based on anatomical location – alveolar and interstitial macrophages. In animal models, depletion of alveolar macrophages led to increased inflammatory response to house dust mite [96] and raise in IgE levels [97]. In comparison, interstitial macrophages have regulatory phenotype, are characterised by secretion of IL-10 [91,98], and associated with regulating Th17-mediated neutrophilic inflammation in asthma [99].

Dendritic cells, the other antigen presenting cell type along macrophages, bridge innate and adaptive immunity and enable the involvement of B and T cells in immune response. The dendritic cells can be broadly divided into plasmocytoid (pDC) and conventional (cDC) dendritic cells, with cDCs further divided into DC1 and DC2 populations. The main function

of cDC1s is antigen cross-presentation, whilst cDC2s respond to various danger signals, such as ranging from nucleotides and polysaccharides, through engagement of Toll-like receptors found on their surface [100]. Conversely, pDCs' main role is in antiviral immunity, mediated through their ability to produce type I interferon [101].

Adaptive immune responses in the lung lead to formation of inducible bronchus-associated lymphoid tissue (iBALT), which consists of T- and B-cell aggregates coinciding with antigen-presenting cells and supported by populations of stromal cells [102,103]. In mice models of allergic inflammation, iBALT was demonstrated to host pathogenic memory Th2 cells [104].

1.7.3 GRANULOCYTES IN SINGLE-CELL GENOMICS

To date, high-quality single-cell transcriptomes of granulocytes have predominantly been obtained from non-solid tissues, the blood and bone marrow [105,106]. In humans, substantial numbers of human eosinophils have only been successfully sequenced from the blood [107], with microwell-based methods (such as BD Rhapsody) showing better performance in general compared to microfluidics-based systems (e.g. 10X Genomics Chromium). In other tissues, negligible eosinophil count figures have been reported in a small number of studies – 13 eosinophils out of 71,752 total cells in foetal lung [108] (0.02%) and 321 eosinophils out of 103,228 total cells in foetal bone marrow [109] (0.31%).

The key factors that likely underlie the challenge of capturing granulocytes from fresh tissues include the shearing forces generated during tissue dissociation and within the

microfluidics-based single-cell platforms that can compromise granulocyte membrane integrity, and the fact that these cells are prone to degranulation in response to activating/stress signals [110].

1.8 SINGLE-CELL SEQUENCING AND THE RESPIRATORY SYSTEM

1.8.1 SINGLE-CELL GENOMICS

The development of single-cell genomics and the rapid decrease of the sequencing costs enabled simultaneous profiling of thousands to millions of cells in any given sample at the same time. While early studies focused on the readily accessible blood tissue, more recent research has shifted attention to primary tissues, with an aim to increase significance to disease pathology. For instance, the foetal lung cell atlas traced the differentiation trajectories of epithelial cell types and identified previously uncharacterized cell states, including developmental-specific secretory progenitors [108].

In addition to individual efforts, ambitious international, multi-centre consortia such as Human Cell Atlas [111] have aimed to profile every tissue in the body at single-cell resolution, in health and disease. Single-cell profiling of the lung brought many exciting discoveries, such as the identification of pulmonary ionocytes [86], presence of gland-associated

immune niches [112], and alveolar epithelial niches [113] important in diseases such as COPD and fibrosis.

Data from individual studies is often harmonised and integrated, which creates large tissue atlases, such as the Human Lung Cell Atlas, consisting of an impressive 2.4 million cells across the airway [78]. In addition to sample collection, the technological advances necessitated new computational methods able to efficiently deal with large datasets whilst answering biologically relevant questions, leading to over 1,400 tools to analyse scRNAseq data [114]. Moreover, these tools need to be used in a reproducible way and at a large scale, leading to the development of single-cell analysis pipelines, such as *panpipes* [115] and nf-core *scrnaseq* [116], with the goal of streamlining the standard preprocessing tasks.

Whilst early iterations of single-cell sequencing were limited to capturing the polyadenylated mRNA transcripts, subsequent advances enabled the capture of more than one molecular modality (layer of information) from a single cell. These modalities include antibodies binding to cell surface proteins, akin to flow cytometry, as well as chromatin accessibility and the immune (V(D)J) repertoire (T cell and B cell receptor sequences), and spatial location in a sample of the tissue [117].

1.8.2 DIFFERENCES IN CELL TYPES DETECTED BASED ON SAMPLING TECHNIQUE AND ANATOMICAL LOCATION

Sampling of the airway poses unique challenges due to the paucity of the material that can be collected from living individuals [79]. The anatomical location (proximal versus distal airway) and the sampling technique, influence the type and number of cells collected, as summarised in Figure 1-5.

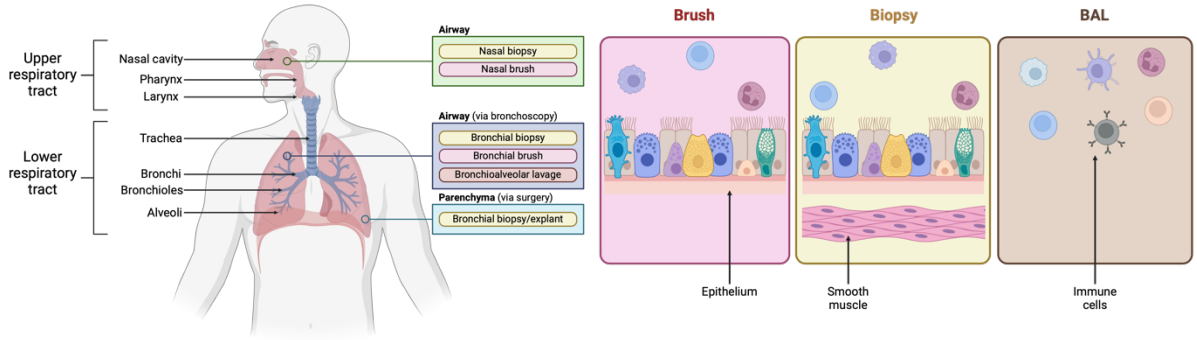

Figure 1-5 Anatomical locations and sampling strategies in upper and lower airway
The three main sample types obtained during a bronchoscopy are bronchoalveolar lavage (BAL), brushings, and biopsies. Sampling of the alveolar region requires parenchymal lung tissue obtained through surgical biopsy or explants. The choice of sampling technique significantly influences the cell types collected: lavage yields mostly immune cells, brushings collect both epithelial and immune cells, while biopsies capture epithelial and immune cells as well as stromal cells.

Adapted from Hewitt and Lloyd (2021). Copyright Springer Nature, reproduced with permission (license 5945441012604).

Typically, samples from healthy and asthmatic patients are collected via fibreoptic bronchoscopy during routine procedures [77,118] or in the context of a clinical trial and, in our experience, yield no more than a few tens of thousands of cells. This contrasts sharply with samples from deceased lungs, surgical biopsies, or explanted tissue, which often provide access to several million cells. This sparsity of research material prevented detailed immunophenotyping in the past, limiting the available techniques to microscopy, flow cytometry, and cytospins. Samples of the upper airway can also be obtained through a

straightforward and non-invasive nasal brushing [77,79]. However, whilst there are structural and functional similarities, the biology of the upper airway may differ significantly from that of the lower airway, presenting a critical challenge when investigating airway diseases. Addressing this gap in research is essential to determine the extent to which upper airway samples can serve as reliable proxies for studying lower airway diseases, such as asthma and COPD. Finally, samples with defined tissue structure, such as biopsies, require enzymatic and/or mechanical dissociation to create a single-cell suspension required for transcriptome profiling, which was shown to introduce batch effects by upregulating *JUN*/*FOS* and *HSP* gene signatures [119,120].

1.9 THESIS OBJECTIVES

Asthma is a heterogenous disease with differences in molecular mechanisms between the phenotypes, sparsity of information about granulocytes in single-cell datasets, and unknown tissue heterogeneity both between the tissue sites (upper versus lower airway) and within the tissues (cell types, cell locations). To address these challenges:

- In chapter 3, I focus on a defined asthma phenotype, severe allergic eosinophilic asthma with house dust mite (HDM)-sensitisation, which is the most common asthma phenotype. I characterise the differences in cellular composition of common sample types available to study the respiratory system using single-cell techniques, I employ a novel technique to track the immune cells binding the allergen (HDM), and profile HDM and IgE binding to immune cells. I also characterise the differences in abundance and gene expression across the cell types compared to healthy individuals
- In chapter 4, I use state-of-the-art spatial transcriptomics techniques to investigate cell co-localisation and existence of tissue niches in patients with severe allergic eosinophilic asthma with HDM-sensitization compared to healthy controls, and expand to include gene expression differences in epithelial cell types compared to immune cell types characterised in chapter 3

- In chapter 5, I include a further three asthma phenotypes to build a comprehensive atlas of asthma at single-cell resolution, and profile key differences between severe asthma and health, as well as between the asthma phenotypes

2 MATERIALS AND METHODS

2.1 PATIENT RECRUITMENT

All patient samples were collected as part of the Oxford Airways Study (ethical approval REC number 18/SC/0361). Bronchoscopies were performed at John Radcliffe Hospital and Churchill Hospital in Oxford, UK, by an authorised clinician, using standardised protocol in accordance with American Thoracic Society recommendations. Samples collected included nasal brushings, endobronchial brushings, biopsies, and bronchoalveolar lavage (BAL).

2.2 SINGLE CELL GENOMICS

All tissue processing was carried out in class II biosafety cabinet previously treated with UV light for 20min, with equipment and surfaces wiped with RNaseZap (Thermo, cat. AM9780).

2.2.1 TISSUE DISSOCIATION AND STAINING WITH ANTIBODIES

Endobronchial biopsies: Biopsies (8-10 samples, approximately 2 mm in size) were collected into 4 mL of colourless RPMI (Gibco cat. 11835030). Collagenase dispersal was performed using 1 mg/ml type 1 collagenase from *Clostridium histolyticum* (Sigma cat. C0130) in RPMI for 1 hr at 37°C with magnetic stirring, samples were be depleted of RBCs using 30 s incubation in sterile nuclease-free water, dispersed using 23G needle and 70 µm mesh filter and stained. The staining panel consisted of a mixture of flow cytometry antibodies, which were used for sorting and purity assessment, and TotalSeq™-C antibodies. CD45 was used to balance the immune and non-immune compartments, such that an approximately equal number of CD45+ and CD45- cells was collected. Hashing antibody (different for every patient-tissue combination) was used to demarcate sample ID and tissue of origin (nasal brush, BAL, or endobronchial biopsy). The detailed panel designs are given in Table 2-1 and Table 2-2.

Nasal brushings: Nasal brushings were collected by rotating cervical cytology brushes in the nasal cavity 5 times and inserted into 4 mL of colourless RPMI. Cells were dissociated from the collection brushes by vigorous aspiration with 1000 µl pipette and rinsed three times with 1 mL of RPMI. The cells were passed through 23G needle and 70 µm mesh filter to create single cell suspension, treated with 5 µl Human TruStain FcX™ for 10 min on ice (to prevent non-specific binding) and stained. The detailed panel designs are given in Table 2-1and Table 2-2.

Bronchioalveolar lavage (BAL): BAL samples were collected using 50 mL of PBS (Sigma cat. D8537), passed through 23G needle and 70 µm mesh filter to create single cell suspension, depleted of RBCred blood cells (RBCs) using 30 s incubation in sterile nuclease-free water, treated with 5 µl Human TruStain FcX™ for 10 min on ice to prevent non-specific binding, and stained. The detailed panel design is indicated in Table 2-2.

One of the aims of the study was to profile differences between the biology of upper (nasal brush) and lower (BAL, endobronchial biopsy) airway in healthy and asthmatic patients. For the experiments described in Chapter 3, cells hashed cells obtained from BAL, endobronchial biopsy, and nasal brush samples from a single patient were pooled together and run on a single 10X Genomics Chromium chip. For the experiments described in Chapter 5, cells obtained from endobronchial biopsy and nasal brush samples from a single patient were pooled together and run on a single 10X Genomics Chromium chip. This multiplexing strategy was used to accommodate the need to process samples fresh and the

variable number of patients undergoing bronchoscopies on different occasions (ranging from 1 to 4 patients). In ideal circumstances, samples from different patients would be combined instead to take advantage of genetic demultiplexing algorithms and thus potentially reduce the number of cells of ambiguous tissue origin.

2.2.2 ANTIBODY PANELS

Antibody panels used for staining are given in Table 2-1 (Chapter 3) and Table 2-2 (Chapter 5), with antibodies common between the two panels listed in bold. The differences in panels are due to different timelines and commercial founders of each study. The CD45+ sorting strategy used in Chapter 3, which enriched for immune cells, was employed to increase the statistical power of the small study, as immune cells are rare compared to epithelial cells. The staining results are presented in Figure 2-1.

Table 2-1 Antibody panel used in the study of asthma phenotypes (nasal bush, biopsy)

Live cells were sorted in approximately equal proportions of CD45+ and CD45- populations.

Flow cytometry	**TotalSeq™-C**
CD161 BV786	CD45
CD25 BV711	**HLA-DR**
CD19 BV650	**CD3**
CD45 BV605	**CD16**
CD4 BV570	**CD19**
MAIT tetramer BV421	**CD14**
CD16 PerCP	**CD4**
CD3 AF488	**CD8**
CRTH2 PE-Cy7	**CD193**
CD14 PE-TxR	**TCR Va7.2**
CD8 APC-Cy7	CD161
CD127 APC-R700	**CD38**
CD193 APC	CD31
7-AAD (viability dye)	**TCR Va24-Ja18**
	CD138
	Mouse IgG1, κ (isotype control)
	Mouse IgG2a, κ (isotype control)
	Hashing antibody

Table 2-2 Antibody panel used in the study of HDM-sensitized allergic eosinophilic asthma
Live CD45+ cells were sorted.

Flow cytometry	**TotalSeq™-C**
CD45 APC-Fire450	**HLA-DR**
7-AAD (viability dye)	**CD3**
	CD16
	CD19
	CD14
	CD4
	CD8
	CD193
	TCR Va7.2
	CD38
	CD138
	Polyhistidine tag (“His-tag”)
	CD117
	TCR Va24-Ja18
	IgE
	DerP1-oligo conjugate
	DerP2-oligo conjugate
	Mouse IgG1, κ (isotype control)
	Mouse IgG2a, κ (isotype control)
	Hashing antibody

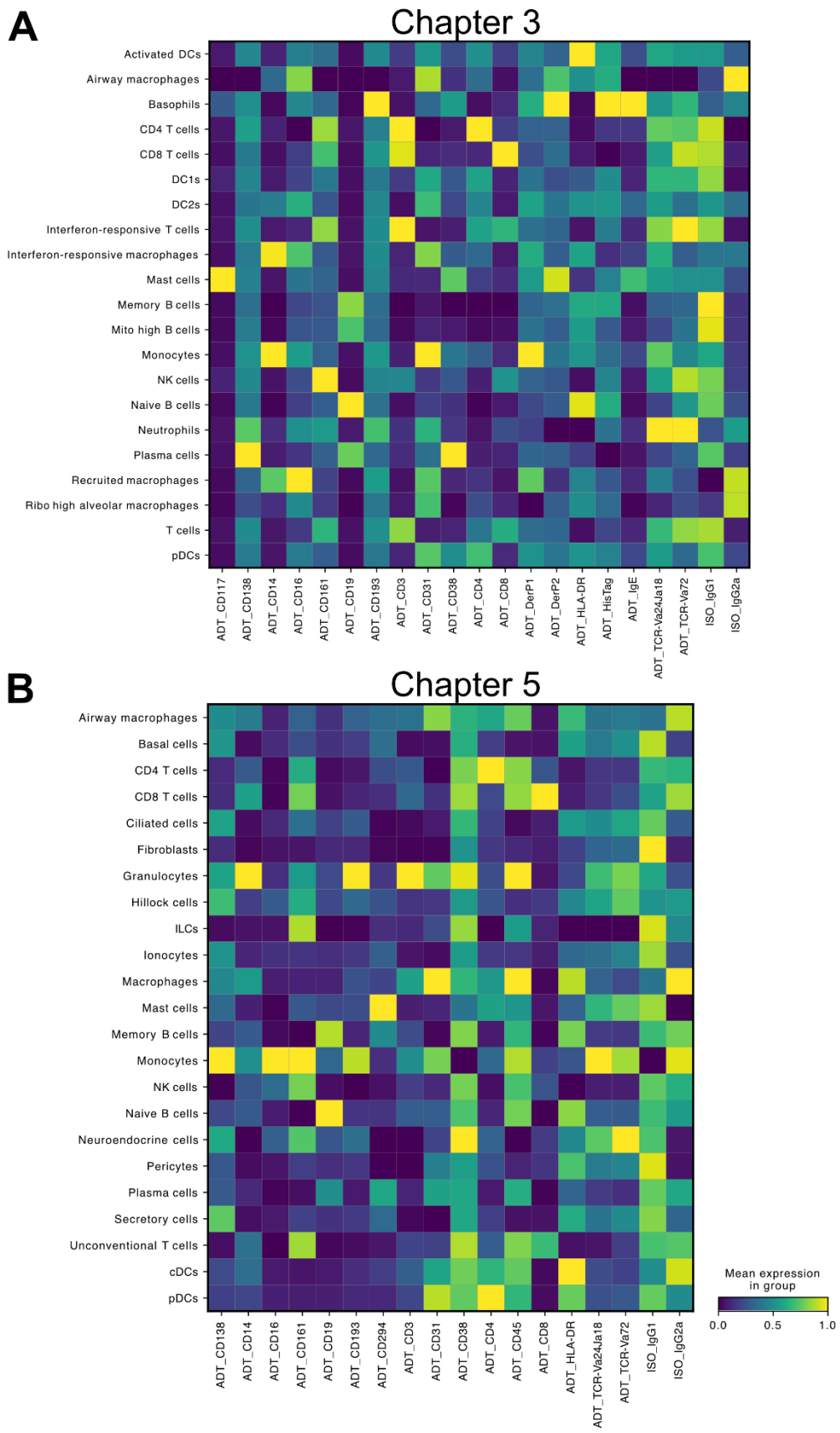

Figure 2-1 Overview of the dsb-normalised protein signal in Chapters 3 and 5

The dsb-normalised antibody staining signal for antibodies used (ADT) and antibody isotype controls (ISO) for single cell studies described in Chapter 3 (A) and Chapter 5(B).

2.2.3 OLIGO-CONJUGATED ALLERGENS

Recombinant House Dust Mite DerP1 (Cambridge Bioscience cat. 230-00023-50) and DerP2 (Cambridge Bioscience cat. 230-00024-50) proteins were concentrated using buffer exchange spin columns (abcam cat. ab102778) and conjugated to 10X Genomics-compatible DNA oligonucleotides using a commercial kit (5' Feature Barcode Antibody Conjugation Kit, abcam cat. ab270703). These conjugates were added to the cell suspension and incubated for 10 min on ice.

2.2.4 LIBRARY PREPARATION AND SEQUENCING

Sequencing libraries were generated using Chromium Next GEM Single Cell 5' Reagent kit v2 (Dual Index) following manufacturer's instructions (protocol CG000330). Library generation was automated using Biomek FXP Laboratory Automation Workstation (Beckman Coulter) at MRC WIMM Advanced Single Cell Omics Facility (WASCOF, University of Oxford). The automation methods were written and tested by me and Dr Maria Greco. Library quality and concentration was assessed using a Bioanalyzer (Agilent) and Qubit 2.0 Fluorometer (Thermo Fisher Scientific), respectively. Sequencing was carried out by a commercial supplier, Novogene UK Limited. Libraries were sequenced on NovaSeq6000 and X Plus instruments (Illumina Inc) to a requested depth of 400 M read pairs for transcriptome (RNA) libraries, 160 M read pairs for feature barcoding (ADT), and 80 M read pairs for immune repertoire (TCR and BCR) libraries.

2.2.5 DATA ANALYSIS

R 4.3.2 and Python 3.10 programming languages were used for the analysis.

Scanpy's anndata [121] and *muon*'s mudata [122] frameworks were used for data manipulation and plotting in Python. The analyses were performed using *scanpy* (single cell) [121] and *squidpy* [123] and supplemented with task-specific tools, as indicated below.

For packages written in R, *SingleCellExperiment* [124] was used as data container. *zellkonverter* was used to convert between anndata/mudata and SingleCellExperiment objects [125].

2.2.5.1 PREPROCESSING

CellRanger analysis pipeline (cellranger-multi 7.1.0) was used to carry out unique molecular identifier (UMI)-based read collapsing, read alignment and generate raw count matrices. Reads from gene expression libraries were aligned to GRCh38-2020-A reference, and immune repertoire to GRCh38-alts-ensembl-7.0.0, both provided by 10X Genomics. Custom references based on barcode sequences provided by BioLegend and abcam were created for TotalSeq™-C and custom conjugate libraries.

2.2.5.2 CELL HASH DEMULTIPLEXING

cellhashR package (v1.0.3) [126] was used to assign the cells to tissues of origin using consensus mode with agreement rate of at least 50% (chapter 3) and 100% (chapter 5) – i.e., at least half or all the algorithms which were able to be called on a given sample had to be in agreement for the result to be deemed unambiguous. The algorithms used were: "htodemux", "multiseq", "dropletutils", "demuxem", "demuxmix", "bff_raw", and "bff_cluster".

2.2.5.3 QUALITY CONTROL, FILTERING, AND DIMENSIONALITY REDUCTION

panpipes (v.1.1.0) pipelines [115] was used to carry out QC steps, filtering, dimensionality reduction, and clustering. Cells were filtered based on minimum (to exclude debris) and maximum (to exclude multiplets, predominantly doublets) UMI counts, percentage of mitochondrial counts (to exclude dead and dying cells), doublet scores (to exclude homotypic doublets), cellhashR doublet calls (to exclude heterotypic doublets). Highly variable genes (HVGs) were identified using Seurat algorithm (version 3), principal components (PCs) were calculated for the dataset and dimensionality was reduced using UMAP algorithm. UMAP embedding was corrected for batch effect using an array of methods (harmony, scvi, bbknn, scanorama) and assessed visually and through calculating LISI scores. Leiden clusters were calculated at resolutions 0.1 to 2.0, cell number permitting. Annotation was performed on broad cell type groups (“compartments”): T and NK cells, B

and plasma cells, myeloid cells, non-immune cells. These compartments had HVGs, PCs, and batch-corrected UMAP embeddings and clusters recalculated before proceeding to annotation based on expert input, a set of canonical markers, and Human Lung Cell Atlas [78] (core dataset) reference annotations mapped using *scvi-tools* [127].

I was an active contributor to the panpipes package during its development and the set of pipelines are now published as a peer-reviewed article in Genome Biology.

2.2.5.4 PSEUDOBULKING

Early approaches to differential gene expression analysis treated each single cell as an independent observation, inflating the number of statistically significant differentially expressed genes [128]. To circumnavigate this problem, pseudobulk RNA samples were created by summing the gene counts in individual cell types over patients using *decoupleR* [129]. Immune receptors whose expression is highly variable within a given cell type (each B/T cell clone expresses a unique V(D)J combination) were excluded from the pseudobulk aggregates using the regular expression recognising any gene name that starts with the following: "IGHV", "IGLV", "IGKV", "AC233755", "TRAJ", "TRAV", "TRDV", "TRBJ", "TRBV", "TRGV", "TRGJ", "TRG-".

2.2.5.5 CELL ABUNDANCE CALCULATION

Cell type abundance in a given patient was calculated as the ratio of number of cells with a given annotation to all the cells with the same annotation at annotation level 0 (compartment annotation).

2.2.5.6 PRINCIPAL COMPONENT ASSOCIATION ANALYSIS

Association of abundance and gene expression with experimental variables was performed with *decoupleR* [129]. Correlation scores were calculated between pseudobulk gene expression values/abundance values and clinical variables, and ANOVA was used to test for association. ANOVA p values were corrected using Benjamini-Hochberg method.

2.2.5.7 DIFFERENTIAL GENE EXPRESSION ANALYSIS

Cluster gene markers for annotation purposes were calculated using Wilcoxon test. Differential gene expression was analysed using *MAST* [130], providing confounding variables as covariates to the model and excluding genes with zero cellular detection rate. 0.25 logFC and p_{adj} <0.05 were used as thresholds to determine significant changes in expression.

2.2.5.8 DIFFERENTIAL ABUNDANCE ANALYSIS

Differential abundance analysis was performed using *miloR* [131]. Nearest -neighbours graph was calculated using PCs given by *panpipes*, with parameters optimised to reach neighbourhood size approximately equal 4-5 times the number of samples in the dataset or 100-200, whichever was larger.

2.2.5.9 GENE SET ENRICHMENT ANALYSIS

ClusterProfiler [132] was used to test for gene set enrichment analysis against MSigDB "h.all.v2024.1.Hs.symbols" reference. Genes were ordered according to the score calculated as $-\log(p_{adj}) \times \text{sign}(\log FC)$ values from the MAST algorithm. GSEA was performed using the compareCluster function. A small pseudocount (1E-300) was added to every p-value to avoid errors resulting from taking a logarithm of 0.

2.2.5.10 CELL-CELL INTERACTION ANALYSIS

LIANA+ [133] was used for cell-cell interaction analysis, with all *MAST* genes with $p_{adj} < 0.05$ taken as potential ligands for differential interaction analysis.

2.2.5.11 CLONALITY ANALYSIS

B and T cell clones were called by cellranger-multi as described in section 2.2.5.1. Clonality analysis of T cells was carried out at annotation level 2 using *immunarch* [134] with clone sizes set as 1, 2-10, >11. Gene expression data was used to establish cell type annotations. Percentage of clonal repertoire was calculated as percentage of a given clone size in relation to all the clones called using VDJ data. The clone counts in two groups (healthy controls, asthmatic patients) were compared using Wilcoxon rank sum test with p values adjusted for multiple comparisons.

2.3 IN SITU SPATIAL TRANSCRIPTOMICS

2.3.1 TISSUE PROCESSING

Biopsies (1-2 samples, approximately 2 mm in size) for spatial transcriptomics and imagining experiments were collected into 4 mL of phosphate-buffered formalin solution (Sigma cat. HT5011). The biopsies were paraffin embedded using a robot at the clinical facility (OCHRe, Oxford Centre for Histopathology Research) producing FFPE (formalin-fixed, paraffin embedded) blocks. Some archival biopsies stored as part of OAS were included in the study.

Biopsies were sectioned to according to protocol CG000578 by 10X Genomics. Prior to starting, all work surfaces were treated with RNaseZap. Briefly, FFPE blocks were kept at 4℃ at all times, sectioned using a microtome to 5 µm slices, floated in RNA-free water bath at 42℃ and placed on the Xenium slide. Slides were then dried at 42℃ for 3h in histological oven and placed in a desiccator.

On the day of the experiment, the slides were processed according to protocol CG000760 by 10X Genomics. Briefly, a cocktail of priming DNA oligos was added onto the tissue and allowed to hybridise to the RNA in the tissue, the RNA template was then released from the oligo and enzymatically digested using RNase, and the proprietary “polishing” step was performed to prepare the tissue for probe binding. The RNA-detecting probes were then added to detect gene expression, ligated to form a circular DNA template, and amplified

through rolling circle amplification. Finally, cell membranes were stained with a proprietary fluorophore cocktail, the autofluorescence was quenched using chemical treatment, the nuclei were visualised with DAPI stain, and the tissue was analysed using Xenium Analyzer instrument.

The slide was retrieved, and H&E staining was performed according to protocol CG000613 by 10X Genomics. Briefly, the autofluorescence quencher was removed using 10 mM solution of sodium hydrosulphite, stained with haematoxylin solution and bluing solution. The slide was washed three times in purified water between the steps. The slide was then dehydrated in 70% and 95% ethanol solutions, stained with eosin, and the tissue was dehydrated by submersion in a series of alcohol solutions (two times 95%, two times 100%). The alcohol was replaced with xylene by submersion in two xylene jars and mounted with a cover slip.

The slide was visualised using ZEISS Axioscan 7 microscope at the Oxford-ZEISS Centre of Excellence in Biomedical Imaging at 20x magnification.

2.3.2 DATA ANALYSIS

2.3.2.1 PREPROCESSING

Tissue regions were selected according to patient diagnosis (healthy, asthma, allergy) based on the slide map created during the sectioning stage and exported by the Xenium Analyzer at the data acquisition stage. The cell segmentation was performed by the Xenium Analyzer based on 5 µm expansion of the DAPI nuclear stain, boundary stain using a mixture of ATP1A1, CD45, and E-cadherin staining, and interior 18S RNA stain.

2.3.2.2 QUALITY CONTROL, FILTERING, AND DIMENSIONALITY REDUCTION

Gene count matrix and spatial information was extracted using SpatialData i/o functions [135]. The resulting anndata objects were fed to *panpipes* and processed in a similar manner to single cell genomics data. Cells were filtered based on minimum (to exclude debris) and maximum (to exclude aggregates) cell and nucleus area, number of negative control probes detected (which was set to 0), and nucleus count (1). The minimum and maximum number of transcripts filter consistent with the gene panel size was also used. Highly variable genes (HVGs) were identified using Seurat algorithm (version 3), principal components (PCs) were calculated for the dataset and dimensionality was reduced using UMAP algorithm. UMAP embedding was corrected for batch effect using an array of methods (harmony, scvi, bbknn, scanorama) and assessed visually and through calculating LISI scores. Annotation was

performed based on expert input and Human Lung Cell Atlas [78] (core dataset) reference annotations mapped using *scvi-tools*.

2.3.2.3 PSEUDOTIME ANALYSIS

Pseudotime analysis was performed using *Palantir* [136] on a subset of cells of interest, with preprocessing (HVG identification, PC calculation, dimensionality reduction) performed in *scanpy* to mirror the processing of the whole dataset conducted in *panpipes*.

2.3.2.4 DIFFERENTIAL GENE EXPRESSION ANALYSIS

Differential gene expression was performed using *MAST*, similarly to single cell expression data, correcting for cellular detection rate but not including any additional covariates.

2.3.2.5 SPATIAL NICHE ANALYSIS

Squidpy was used to perform neighbourhood analysis [121][123]. Neighbourhood graph of cells was calculated using Delaunay triangulation and plotted as an average z-score across 1000 permutations. Hierarchical clustering was used to group the scores and niches were annotated manually based on known biological function of individual cell types.

3 CHARACTERISATION OF IMMUNE CELLS ACROSS TISSUES IN ALLERGIC ASTHMA AND HEALTH

3.1 INTRODUCTION

Allergens of animal (e.g. house dust mite) and plant (grass and birch pollen) are a major cause of asthma worldwide. They contain proteins with enzymatic activity, which are able to cleave intercellular junctions, leading to the loss of cell-cell contacts and consequently to altered immune responses [137] (Figure 3-1). Nonetheless, enzymatic activity of environmental proteins has been demonstrated to be insufficient to cause IgE-driven allergic reactions, and other compounds derived from allergenic sources co-delivered with the allergens were shown to provide an adjuvant activity needed to develop allergic sensitisation [138].

In allergic asthma, allergen activity provides the type 2 inflammatory signal, through the release of alarmins IL-25, IL-33 and TSLP at mucosal surfaces [139,140] (Figure 3-1). This signal, in the right genetic background, primes and amplifies airway inflammation and respiratory disease [140,141]. This is supported by the molecular function of the proteins contained in dust mites. HDM allergens are termed using the "*DerP"* prefix with a number, with approximately 40 distinct proteins identified [142]. Typically, allergic sensitisation starts with *DerP 1*, *DerP 2*, and *DerP 23* proteins, which are also the most prevalent allergens, with approximately 50-70% of patients producing IgE against them. Later in life, *DerP*

4/5/7/21 follows, with 15-30% sensitization rates, and lastly, *DerP* 11/14/15/16/18 reactivity emerges around 5 years of age, with prevalence 10%. Presence of IgE antibodies to *DerP* 1 or *DerP* 23 at 5 years of age or less is a predictor of childhood onset asthma [143]. *DerP* 1, which is a cysteine protease, induces epithelial cytokine production and facilitates transepithelial delivery of allergens by disrupting tight junctions [142], whereas *DerP* 2, promotes airway inflammation via TLR4 engagement, resulting in epithelial production of IL-1α, which then triggers granulocyte GM-CSF and IL-33 release in an autocrine fashion. This results in the recruitment and activation of antigen-presenting dendritic cells as well as innate lymphoid cells, forming a link between the innate and adaptive immunity in HDM allergy [144], thereby driving a strong type 2 immune response (Figure 3-1).

HDM extracts prepared from either whole mites, mite bodies, or their faeces, have been found to be contaminated with immunomodulators such as LPS, β-glucans, and chitin, as well as microbial and fungal immunostimulatory substances, such as PAMP receptor activating bacterial DNA and endotoxins [138,145]. As an alternative approach, large amounts of recombinant *DerP* proteins can be expressed in *E. coli*, which can be purified to a high standard using standard molecular biology approaches [146].

Upon antigen presentation and Th2 TCR receptor engagement, the Th2 cells release type 2 cytokines (IL-4, IL-5, and IL-13), which act as a signal for the maturation of B cells, promoting development into plasma cells and release of IgE antibodies (Figure 3-1) [4] IgE production mainly occurs in the lymphoid organs [4], however, it has been suggested that it can also

occur locally in the mucosa of asthmatic patients [147]. IgE has many targets in the inflamed airway and its role is not completely understood. For instance, it can bind to the high-affinity FcεRI and the low-affinity CD23 (FcεRII) receptors, which are differentially expressed across many cell types in the airway [148,149]. FcεRI is expressed on multiple myeloid-lineage cell types (basophils, mast cells, eosinophils, and DCs), as well as epithelial-lineage cells (smooth muscle, endothelium, and epithelium), and is thought to be preliminarily involved in signalling [150]. Conversely, CD23 is mostly expressed on B and plasma cells and thought to be involved in controlling the amounts of soluble IgE [149]. In allergic response, IgE is only produced intermittently by a pool of memory IgG B cells. These cells rapidly switch isotype and expand into short-lived IgE+ plasmablasts, which are sustained by type 2 cytokines such as IL-5 and IL-13 [151].

Allergen-binding B cells are important drivers of IgE-mediated inflammation, activating mast cell degranulation and propagating inflammation in the airway [4,140] (Figure 3-1). Recent development in single-cell genomics saw the invention of techniques to detect antigen binding, with LIBRAseq (community developed, Ivelin Georgiev lab) and BEAMseq (10X Genomics) being the most notable examples. These techniques allow mapping of protein binding cells by tagging the protein of interest with a DNA barcode with appropriate amplification sequences to enable construction of sequencing library, with counts being captured as surface protein (LIBRAseq) or separate (BEAMseq) library [152]. LIBRAseq has been demonstrated to robustly identify B cells recognising SARS-CoV-2 antigen [152].

Nonetheless, the utility of this technique to map protein binding to other cell types, through Fc receptor or non-specific interaction, was not explored.

Studies have been carried out to profile the effect of aeroallergen in allergic asthma at single-cell level [153]. However, these have been limited to small patient numbers and they lack the cross-tissue and cross-cell compartment context (upper versus lower airway, epithelium versus stroma), as well as the inclusion of modalities other than RNA (allergen binding, T and B cell clonality). I aimed to design an experiment to address these problems and expand the understanding of the role of immune cells in allergic asthma.

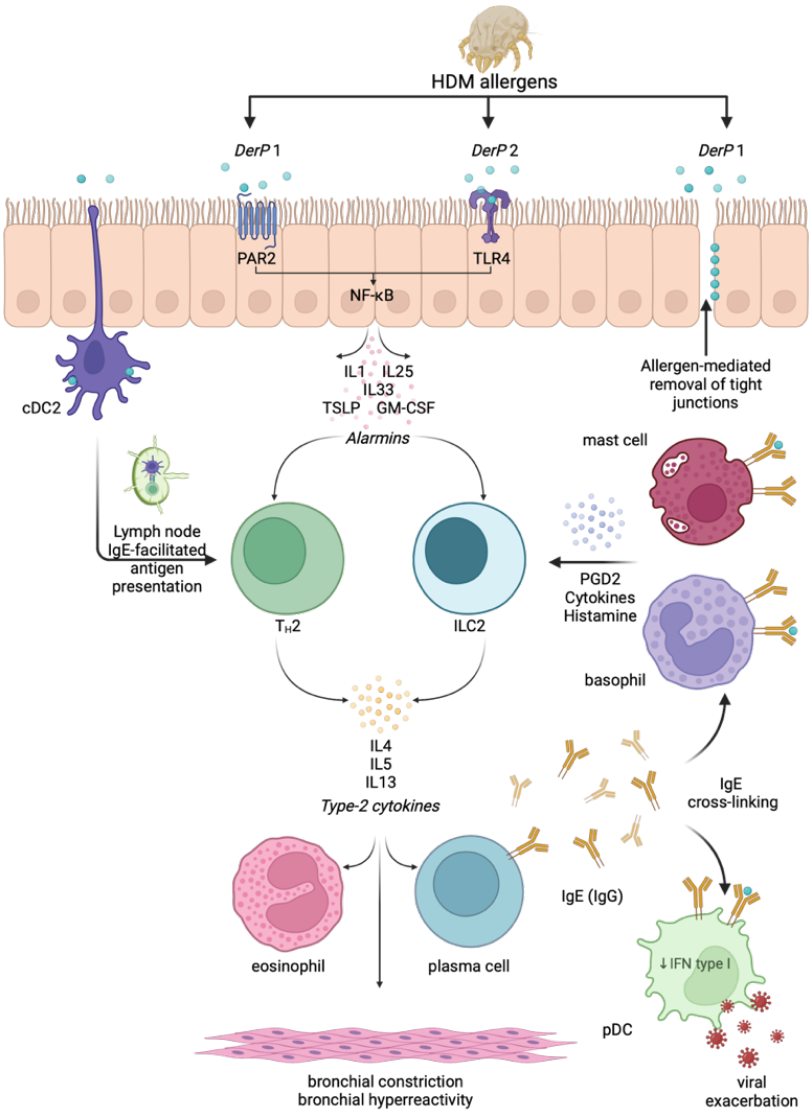

Figure 3-1 Immune cells in HDM-sensitized allergic asthma

DerP 1 is a proteinase capable of cleaving tight junctions and binds to PAMP receptors, whilst DerP2 is a TLR4 agonist. Both allergens trigger the release of alarmins, which in turn trigger type-2 immune response, driven by Th2 and ILC2 cells, and facilitated by antigen-presenting cDC2s. Type 2 cytokines promote IgE secretion, bronchial constriction, and eosinophil activation. IgE cross-linking on mast cells and basophils leads to release of pro-inflammatory cytokines and decrease in anti-viral response.

Adapted from Hammad & Lambrecht (2021) and Hendriks (2020) and updated.

3.2 HYPOTHESIS

In allergic asthma, upon inhalation of an aeroallergen, various immune cells bind the allergen and elicit an immune response against it. One of the responses is production of IgE, which is also bound by immune cells in the innate and adaptive compartments. As a result of binding the allergen, IgE, or cytokine signalling, or a combination of those factors, relative cell proportions and/or transcription profiles change, leading to the development of disease. The immune response differs across the tissues (BAL, biopsy, and nasal brush).

3.3 AIMS OF THIS CHAPTER

- To determine if DNA-conjugated proteins can be used to identify aeroallergen (HDM)-binding immune cells in the airway
- To examine the immune cell types and states present in different sampling locations in the airway (BAL, nasal brush, endobronchial biopsy) in allergic asthma patients and controls
- To identify IgE-secreting and IgE-binding cells in the airway
- To investigate the changes in cell type abundance in asthma and health and across tissue sites
- To delineate the differences in gene expression between cell types and between healthy and asthmatic patients at the single-cell level
- To quantify the clonal expansion of B cells or T cells specific to the aeroallergen

3.4 RESULTS

3.4.1 COHORT CHARACTERISTICS

I designed an observational study to characterize the cellular composition and phenotype of airway cells to obtain a deeper understanding and insight into allergic inflammation, both on its own and in the context of eosinophilic asthma, compared to healthy controls. The study includes three phenotypes: (1) patients with allergic asthma who are sensitized to HDM protein (as characterised by high HDM-specific IgE titres), (2) one patient with HDM sensitization who did not develop asthma, and (3) healthy controls without HDM and few other allergic sensitisations, as summarised in Table 3-1. Patients of phenotype (3), allergic asthmatics, were older and had higher BMI compared to healthy controls – this is on one hand due to the difficulty in recruiting healthy age-matched volunteers willing to undergo a bronchoscopy, and due to obesity being one of the features of asthma [57]. The original study design consisted of phenotypes (1) and (3) only, however, a patient initially diagnosed as having asthma was later found to be misdiagnosed. More patients with phenotype (2) are included in spatial transcriptomics experiments in Chapter 4, hence I decided to retain this sample in the analysis.

BAL, endobronchial biopsies, and nasal brushes were collected from patients during fibreoptic bronchoscopy by a trained clinician. The samples were processed to create single-cell suspensions, as outlined in the methods, stained with an antibody panel, and

processed using 10X Genomics 5prime kit. Four modalities were collected for each cell – RNA expression, surface protein expression, TCR sequence, BCR sequence (Figure 3-2A).

The protocol I developed allows the generation of high-quality single-cell data from samples with very low cell count (typically below 100,000 cells) and viability (typically ~50% viable cells, as quantified by 7-AAD staining). I opted for processing of fresh tissue on the day of collection, as my pilot experiments with frozen samples yielded small recovery rates and poor cell survival, accepting a degree of batch effect may be introduced to the data and may have to be accounted for at the stage of computational data analysis.

In order to reliably detect cells which have bound the aeroallergen patients are sensitised to, I adapted the existing LIBRAseq[152] protocol and used direct conjugation of *DerP* 1 and *DerP*2 proteins expressed in *E. coli* to TotalSeq-compatible DNA barcodes and removed the biotin-streptavidin sorting stain present in LIBRAseq (Figure 3-2B). This allowed me to visualise all the cells binding the antigen, including specialised cells such as B cells (through the B cell receptor), as well as other cells, such as antigen presenting cells (dendritic cells, macrophages) and those with a TLR receptor.

Table 3-1 Cohort characteristics for allergic eosinophilic asthma projects

	Healthy control (n=4)	HDM sensitization (n=1)	Allergic asthma + HDM sensitization (n=3)
Age, years (mean ± SD)	25.3 ± 1.5	29	43.7± 18.2
Male/female	3/1	1/0	2/1
BMI, kg/m2 (mean ± SD)	24.03 ± 3.44	29	31.10 ± 6.01
Eosinophil count [on the day], counts x 109/ml (mean ± SD)	0.15 ± 0.17	0.45	0.29 ± 0.00
Eosinophil count [highest ever], counts x 109/ml (mean ± SD)	0.18 ± 0.22	0.50	0.59 ± 0.41
HDM at screening, kUA/l (mean ± SD)* Detection limits: 0.01-0.35	0.04 ± 0.05	33.7	22.51 ± 27.34
Number of Sensitizations [6 tested: HDM, Grass + Trees, Cat, Dog, Moulds, Aspergillus] (mean ± SD)	0.50 ± 1.00	4	4.33 ± 1.15
Total IgE at screening IU/mL **	134.77 ± 223.67	1655	499.67 ± 274.45
Smoker (%)	50%	0%	33%
Current smoker (%)	50%	0%	33%
Smoking status unknown (%)	50%	0%	0%
* - if the result was below detection limit, it was recorded as 0.01 ** - if the result was below detection limit, it was recorded as 2			

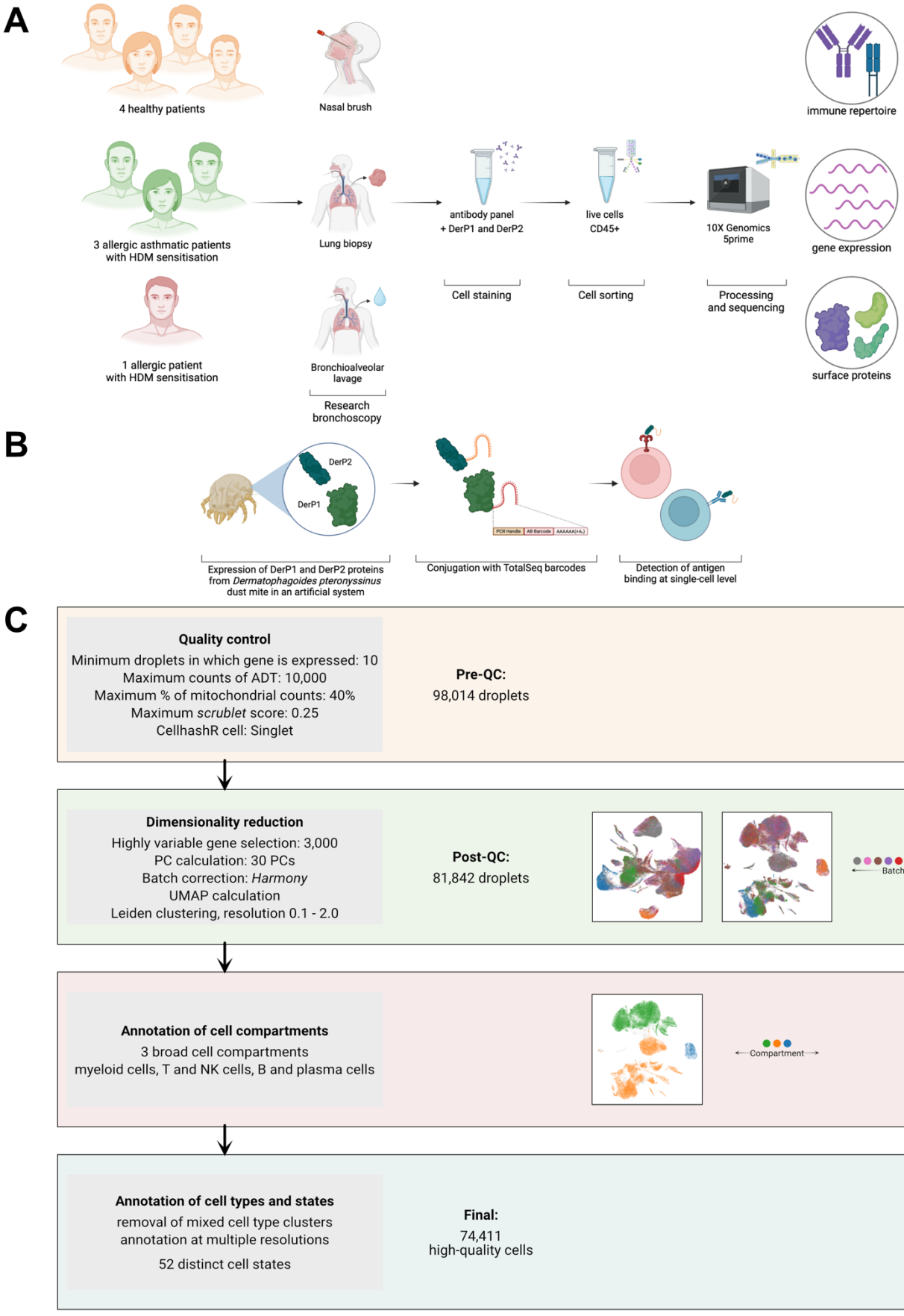

Figure 3-2 Allergic asthma study design and data processing overview

A Overview of the patient phenotypes (asthma, allergic, healthy) and the experimental processing, including sample collection, antibody staining, and single cell genomics library preparation. **B** Synthesis of custom allergen-DNA barcode conjugates using *DerP* 1 and *DerP* 2 proteins expressed in *E. coli*. **C** Summary of key preprocessing steps with associated cell and droplet counts.

3.4.2 DATA PREPROCESSING AND METHOD DEVELOPMENT

Sequencing yielded 98,014 cells (Figure 3-2C). The QC entailed exclusion of cells with mitochondrial content higher than 40%, which could be indicative of physically damaged or dying cells, whose cytoplasmic mRNA leaked through a broken cell membrane, leaving behind only mRNA located in the mitochondria [154]. The mitochondrial content threshold of 40% is more permissive than the typically recommended values (5-10%) and was chosen to avoid removal of potentially biologically relevant cells. Of note, in benchmark studies, it has been shown that such low threshold fails to preserve healthy cells in 13 of 44 human tissues, with metabolically active healthy cells, such as cardiomyocytes, exhibiting up to 30% mitochondrial reads [155,156] and up to 60% in some tissues such as the human gut [157]. In addition, droplet-based sequencing methods lead to generation of doublets and multiplets, i.e. droplets containing two or more cells. These violate the assumption that each barcode represents the transcriptome of a single cell and need to be removed, although new emerging methods are now available to make use of this data in interesting ways, such as to interrogate immune synapses and cell-to-cell interactions [158]. Cell hashing, a method where single cells are labelled with surface antibodies to universally expressed cell-membrane proteins (such as CD45 or components of the MHC), enables not only cost reduction, by running multiple patient samples as part of a single reaction, but also detection of heterotypic doublets through modelling the detection rates of each bound antibody hashtag [159]. To this end, I used cell hashing to encode the tissue of origin in each sample (nasal brush, biopsy, or BAL) and used *cellhashR* [126] to remove heterotypic

doublets, i.e. those droplets where more than one hashing barcode was detected (Figure 3-2C). I identified 13,443 doublets which were excluded from further analysis, 996 negative cells which did not have detectable levels of hashtags, 5,220 discordant cells where it was not possible to assign the tissue with high confidence, and 78,355 singlets, of which 43,739 cells from the BAL, 25,143 from the biopsies, and 9,473 from nasal brushes (Figure 3-2C and Figure 3-3A). As this approach would not identify homotypic doublets, i.e. those where both cells in the droplet share the barcode, I additionally used *scrublet* [160] to model a distribution of artificial doublets, and excluded cells with the score of 0.25 or above – this excluded 2,331 further cells (Figure 3-2C).

Of the 98,014 cells, 81,842 passed QC and were used for annotation, where further cells were excluded either due to forming donor-specific clusters or lacking distinct gene markers to produce 74,411 annotated as high-quality cells (Figure 3-2C).

Genes were filtered to remove those with a low detection rate, by removing genes which were expressed in less than 10 cells in the dataset. That resulted in a dataset of 26,483 expressed genes. To best capture sources of variation in the dataset, one or more dimensionality reduction techniques are typically applied to the single cell data [154]. I calculated 50 principal components on the dataset and identified the first 30 to be explaining a significant portion of the variance in the dataset using a common ‘elbow plot’ approach [154]. I then used these to calculate a final embedding using Uniform Manifold

Approximation and Projection for Dimension Reduction (UMAP) [161] and 4 different data integration algorithms (Figure 3-2C and Figure 3-3B,C).

Batch correction was then applied to the data to account for various technical effects, such as the day of the processing, the operator, the sequencing batch and instrument. As it was not possible to predict and collect the data on all the possible confounding variables, I used the sample ID (OAS patient ID) to model these random effects. As part of *panpipes*, I tested four data integration approaches, namely: harmony, bbknn, and scanorama (Figure 3-3B). Visual assessment of integration as well as Local Inverse Simpson's Index (LISI) score was used to choose the optimal integration method to use on the data (Figure 3-3C), with higher density of LISI scores indicative of better integration [162]. Harmony was chosen as the most optimal method of integration, also reflected in its highest LISI scores.

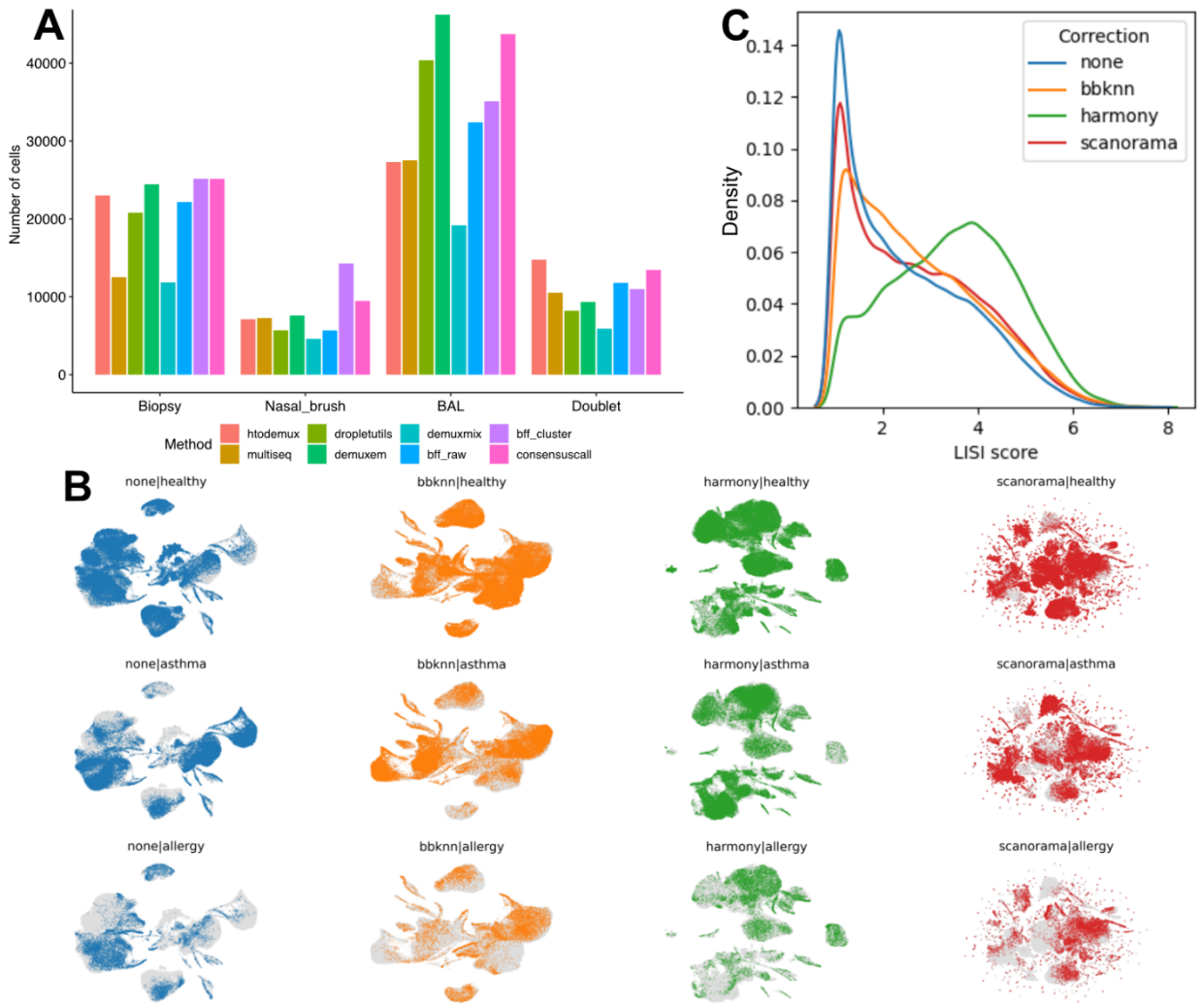

Figure 3-3 De-hashing algorithm and integration method selection

A Summary of the cell calls of 7 different de-hashing algorithms and 50% consensus majority vote method ('consensuscall'). **B** Comparison of three data integration methods on the UMAP embedding coloured by patient phenotype **C** Local Inverse Simpson's Index (LISI) scores calculated for each integration method.

3.4.3 CELL ANNOTATION

To distinguish between the cell types, cells were split into broad cell types based on their differentiation origin [163] (B and plasma cells, NK and T cells, myeloid cells, see Figure 3-4 A) using a combination of canonical cell type markers (such as *CD19*, *CD3*, *HLA-DR*) and markers calculated for each cluster using Student's t test with overestimated variance from the *scanpy* package [121]. I supplemented this with the mapping of BCR and TCR sequences (Figure 3-4 B) which were integrated with the gene expression data. Next, I re-calculated principal components, re-integrated data using Harmony, and annotated the cell states using canonical markers sourced from literature, as well as markers previously identified for the lung in single cell and spatial transcriptomics [112] (Figure 3-4 C-E). The information from the RNA portion of the assay was supplemented with a targeted panel of antibodies designed to identify the immune cells previously described in the lung, including rare populations, such as MAIT cells. Our annotation identified 52 unique cell states (Figure 3-4 F) which I compared to Human Lung Cell Atlas (HLCA) [78] by reference mapping using *scvi-tools*, an approach using transfer learning [164] (Figure 3-4 G).

While there was a broader consensus, some cell states were found to be incorrectly annotated – for instance, neutrophils expressing *FCGR3B*, *CXCL8*, *PLAUR*, *IFITM2* were mapped as classical monocytes, mast cells expressing *CPA3* and *TPSAB1* as 'None', and cells not expected in CD45+ sorted data, such as contaminating club cells mapped as monocytes from HLCA references. I found the semi-supervised approach of clustering and cell marker annotation to be a better representation of the underlying biology. In total, for the 52 identified unique cell states I used the knowledge of lung immunology to define three annotation resolutions differing in the level of cell state granularity, closely mirroring the approach taken by the HLCA.

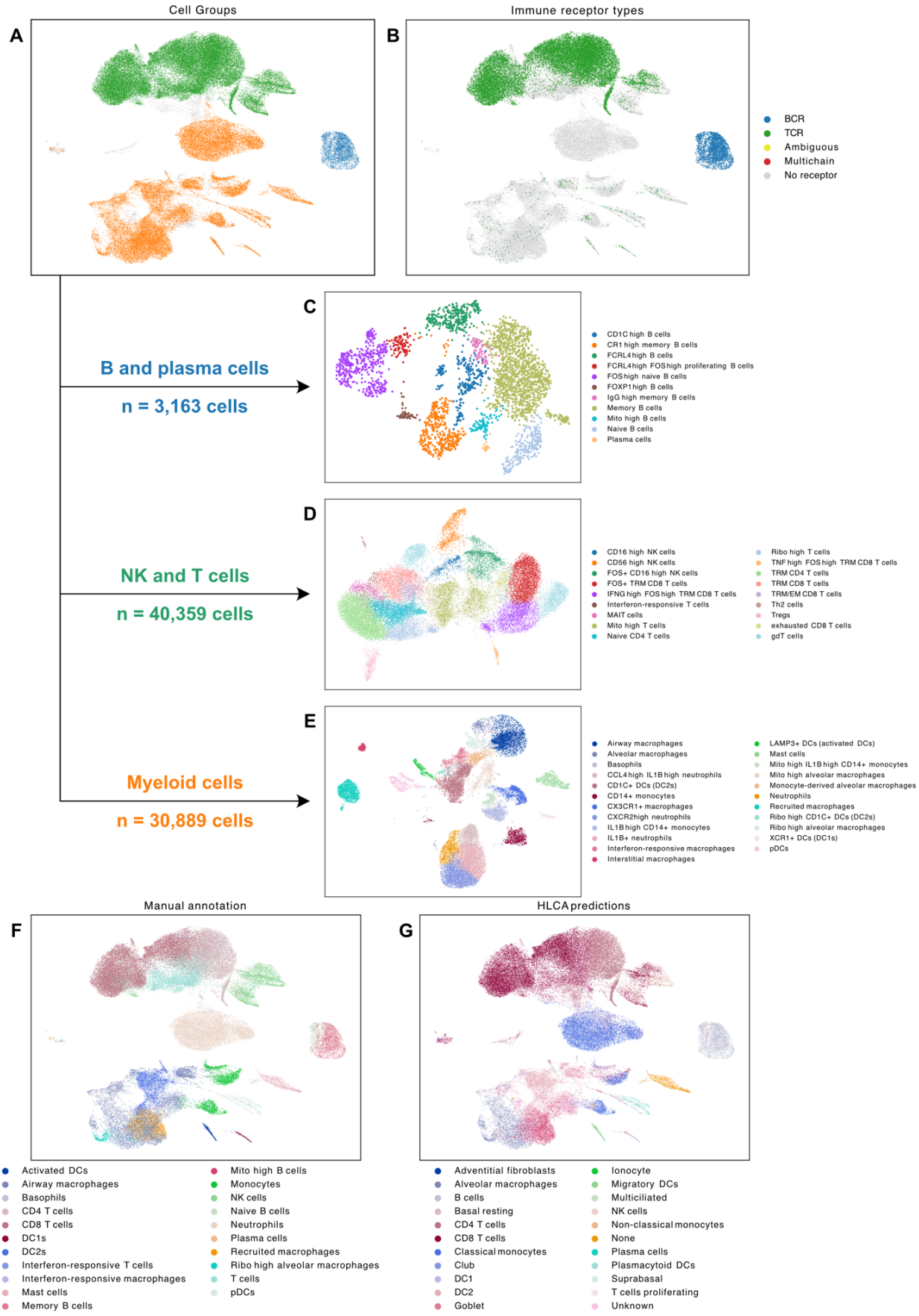

Figure 3-4 Annotation of cell states across compartments and comparison with Human Lung Cell Atlas

A Location of the cell compartments (B and plasma, NK and T, myeloid) cells annotated according to their developmental origin. **B** Mapping of the immune cell receptors to confirm the location of B, T, and plasma cells on the embedding. **C-E** Embeddings of individual cell compartments and corresponding cell state annotations. **F** Mapping of all the annotations to the overall embedding. **G** Mapping of the overall embedding to the Human Lung Cell Atlas reference.

3.4.4 CELL STATES IN THE HUMORAL IMMUNITY COMPARTMENT

Through the subclustering of B and plasma cells, I was able to identify cells across the B-to-plasma developmental progression [165] (Figure 3-5 A). In particular, I identified naïve B cells (*TCL1A*, p_{adj}=1.56E-66, expressing IgM and IgD), memory B cells (*CD27*, p_{adj}=8.26E-178 and *TNFRSF13B*, p_{adj}=1.46E-185), as well as plasma cells (*MZB1*, p_{adj}=1.94E-9) (Figure 3-5 A-C). I did not identify plasmablasts, which have previously been observed in circulation of patients with allergic asthma and allergic sensitisation [166]. In addition, I observed that memory, activated (*JUN*/*FOS*-high) plasma cells, but not naïve B cells, expressed high levels of anti-IgE antibody binding, even though the *IGHE* transcript expression was low across the dataset (Figure 3-5 B-C). At the level of BCR sequences, there was evidence of class switching in the dataset, with the disappearance of IgM clones and decrease of IgM/IgD, shifting to IgA, secretion of which is characteristic of mucosal surfaces in the lung [167].

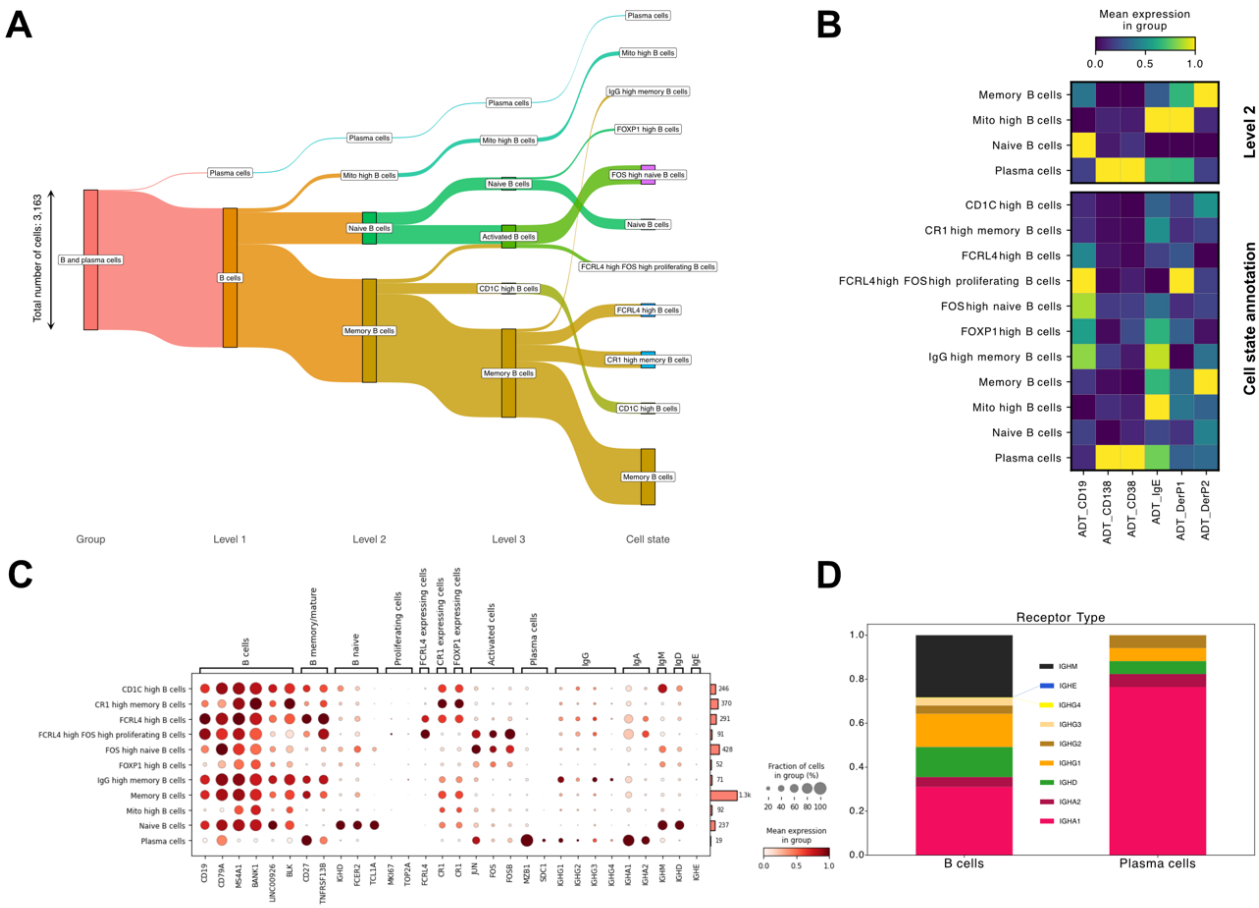

Figure 3-5 Characterisation of B and plasma cell compartment

A Sankey diagram of B and plasma cell populations and states across increasingly detailed annotation levels. **B** Normalised surface protein staining signal across annotation level 2 and cell state annotation. **C** Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars. **D** Normalised abundance of B cell receptor types in B and plasma cells.

3.4.5 CELL STATES IN THE MYELOID COMPARTMENT

In the myeloid compartment, I was able to identify the granulocytes, professional antigen presenting cells (dendritic cells and macrophages), as well as the mast cells, key drivers of inflammation in allergic asthma (Figure 3-6 A). In the granulocyte compartment, I saw neutrophils (*FCGR3B*, p_{adj}=2.99E-12; *CXCL8*, p_{adj}=1.27E-246; *PLAUR*, p_{adj}=4.11E-279) and basophils (*HDC*, p_{adj}=2.65E-17; *IL3RA*, p_{adj}=1.28E-10; *CLC*, p_{adj}=3.47E-05; *CPA3*, p_{adj}=1.67E-06), but not eosinophils (Figure 3-6 B-C). Of those cells, basophils bound IgE (likely through FcεRI [168]) and *DerP*2, but not *DerP*1 protein (Figure 3-6 B). I identified a population of monocytes (CD14+; *S100A8*, *S100A9*, *LYZ*, p_{adj}<1E-300 for all) which bound *DerP*1, but not *DerP*2 (Figure 3-6 B-C). Finally, I identified another population with FcεRI receptor, mast cells, which are thought to be the key drivers of inflammation in allergic asthma (*CPA3* and *TPSAB1*, p_{adj}<1E-300 for both), which bound the *DerP*2 protein. Of the professional antigen presenting cells, I was able to identify type 1 (*XCR1*, p_{adj}=2.84E-101; *CLEC9A* p_{adj}=9.12E-143) and type 2 (*CD1C* p_{adj}=3.50E-3; *CLEC10A* p_{adj}=9.12E-143) conventional dendritic cells, as well as an activated subset of DC1s expressing *LAMP3* (*LAMP3*, p_{adj}=1.05E-232), none of which showed high levels of allergen binding.

I also saw tissue-resident (*MARCO*, *MCEMP1*, *INHBA*, *FABP4*, all p_{adj}<1E-300 for all) and monocyte-derived alveolar macrophages (*MARCO*, p_{adj}<1E-300; *MAFB*, p_{adj}=4.81E-97), interstitial macrophages (CD14+, *F13A1* p_{adj}=4.25E-83; *PLA2G7* p_{adj}=8.80E-55; *MMP9* p_{adj}=5.31E-09), and a transition-like subset of alveolar macrophages which I termed 'airway macrophages', as well as interferon-responsive subset of those cells (expressing interferon-response genes such as *GBP1*, *GBP4*, *GBP5*, *CXCL9*, *CXCL10*, *CXCL11*). Of those, interstitial macrophages seem to weakly bind DerP1, whereas alveolar/airway macrophages - *DerP*2.

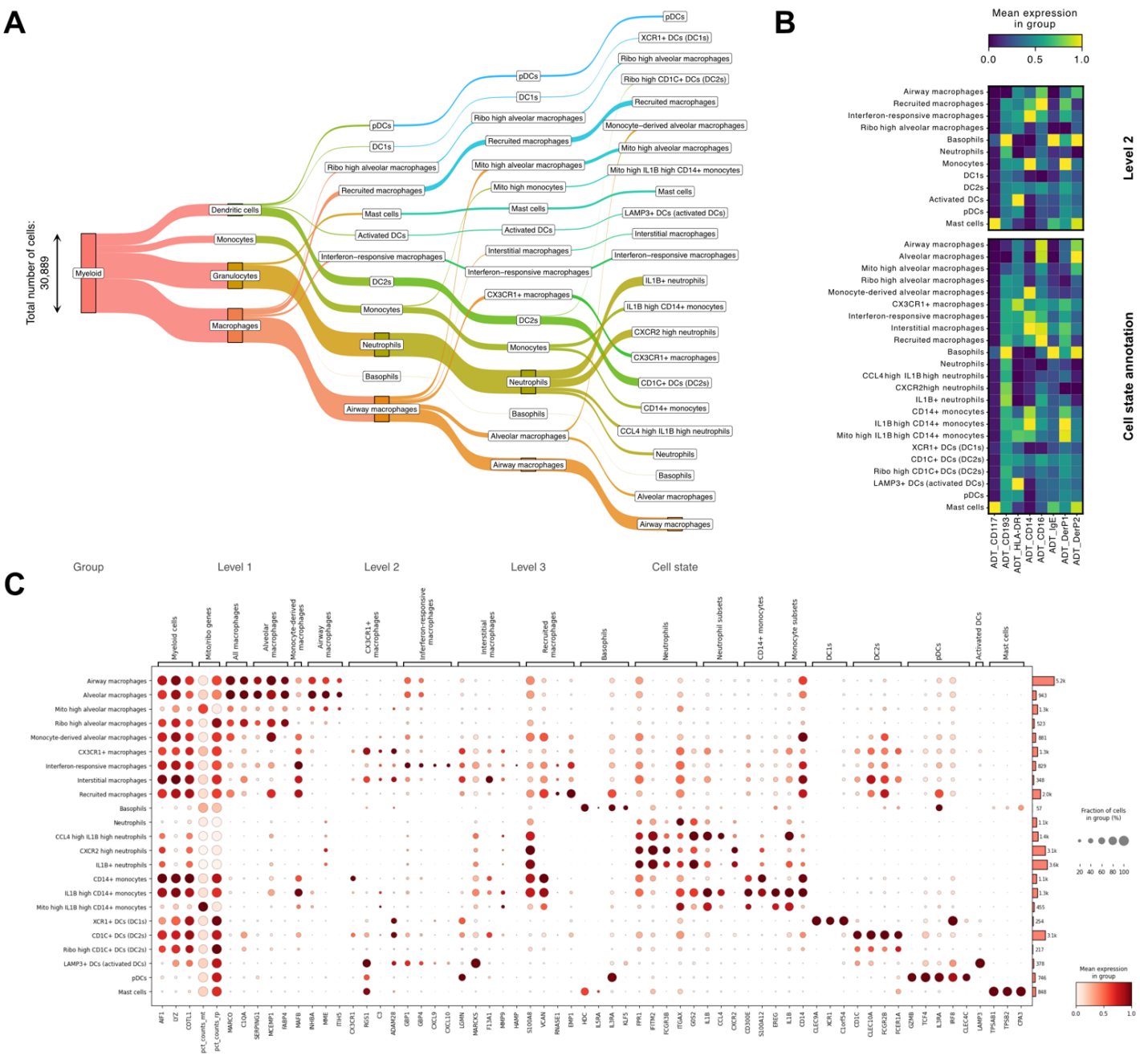

Figure 3-6 Characterisation of myeloid cell compartment

A Sankey diagram of myeloid cell populations and states across increasingly detailed annotation levels. **B** Normalised surface protein staining signal across annotation level 2 and cell state annotation. **C** Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars.

3.4.6 CELL STATES IN NK AND T CELL COMPARTMENT

In the NK and T cell compartment, I was able to identify naïve and tissue-resident cells, as well as effector memory T cells, in both CD4+ and CD8+ subsets (Figure 3-7 A). I identified CD4+ T_H cells (naïve: *LEF1*, p_{adj}<1E-300; *SELL,* p_{adj}<2.51E-75, tissue-resident memory: *CXCR6* and *GZMA* p_{adj}<1E-300; CCR6 p_{adj}=3.09E-67) and CD8+ cytotoxic T cells (tissue-resident memory: *GZMB*, p_{adj}<1E-300; *CXCR6* p_{adj}=1.16E-283; *CCL4*, p_{adj}=2.77E-139, effector memory: *GZMK*, p_{adj}<1E-300; *EOMES*, p_{adj}=5.02E-115; *DTHD1* p_{adj}=2.80E-58). I also identified regulatory T cells expressing lineage-defining transcription factor and a gene encoding immune checkpoint (*FOXP3*, p_{adj}=9.33E-299; CTLA4, p_{adj}=5.91E-250). In addition, I identified rare non-conventional subsets, two with particular importance in asthma: MAIT cells (expressing TCR-Vα7.2 protein; *KLRB1*, p_{adj}=1.98E-277; *SLC4A10*, p_{adj}=1.09E-121; *TRAV1-2* p_{adj}=1.57E-110) and T_H2 cells (*GATA3* p_{adj}=3.52E-33; *IL5*, p_{adj}=1.71 E-4, *IL13*, p_{adj}=6.46E-6), as well as gamma-delta T cells (*TRDV1*, p_{adj}<1E-300). In the NK compartment, I identified CD56-high (otherwise known as “CD56 bright”) NK cells (*NCAM1*, p_{adj}=5.04E-90; *XCL1*, p_{adj}<1E-300) and CD16-high NK cells (*NKG7*, p_{adj}<1E-300; *GNLY*, p_{adj}=3.10E-293). Out of all the cell types identified, IgE was bound at high annotation resolution by highly activated TNF-high FOS-high pro-inflammatory CD8+ T cells, likely an artefact of overall activation, as well as CD4+ T-helper cells, as described before in mice, whose CD4+ T cells express FcεR1 [148] (Figure 3-7 B). *DerP* 2 protein was bound primarily by CD8+ cytotoxic T cells, whereas *DerP* 1 – by interferon-responsive T cells (Figure 3-7 B), which upregulate expression of interferon associated genes, such as the OAS1/2/3 gene cluster, *MX1*, and *IFI6* (Figure 3-7 C).

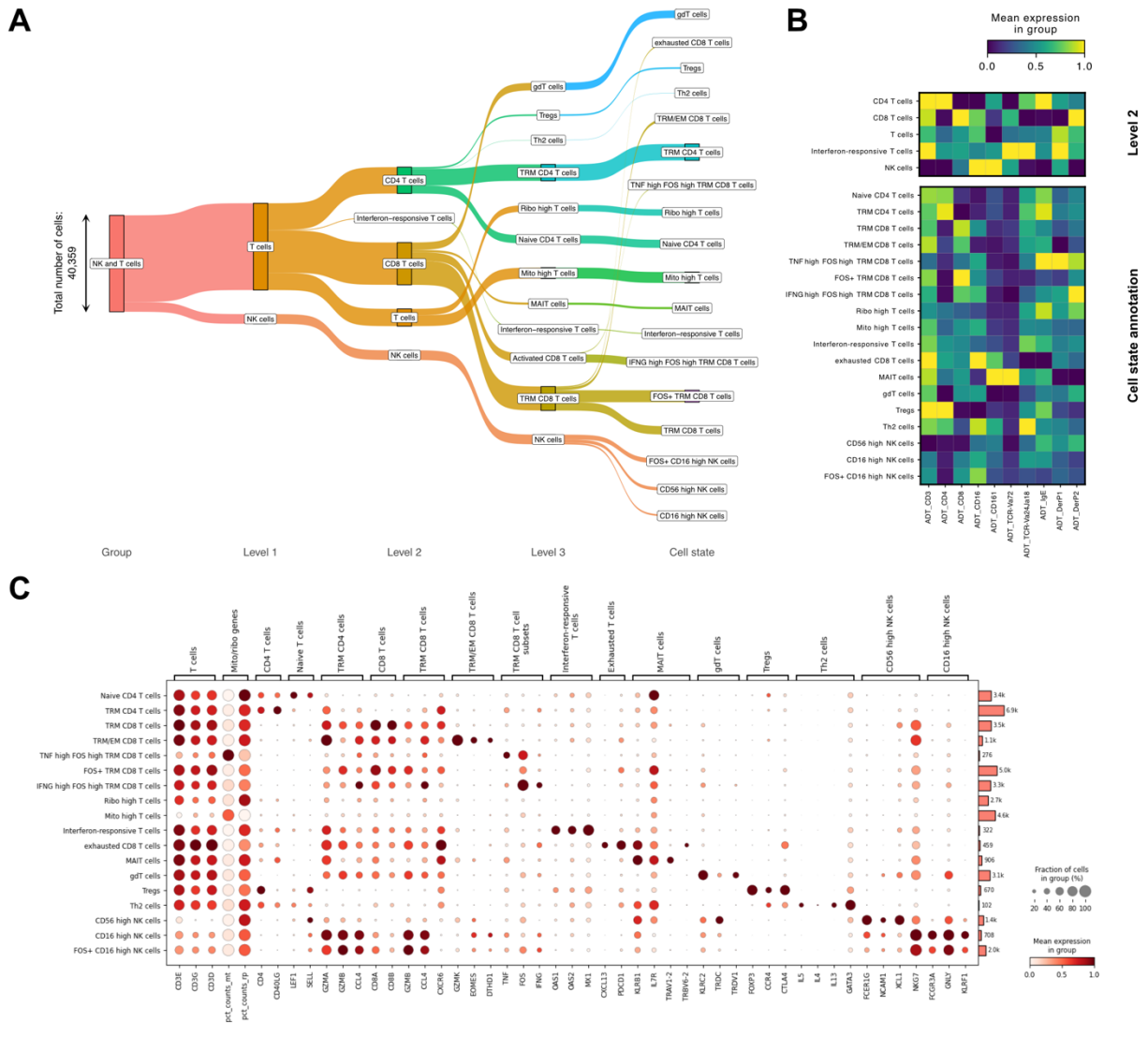

Figure 3-7 Characterisation of NK and T cell compartment

A Sankey diagram of myeloid cell populations and states across increasingly detailed annotation levels. **B** Normalised surface protein staining signal across annotation level 2 and cell state annotation. **C** Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars.

3.4.7 ASSOCIATION OF GENE EXPRESSION WITH CLINICAL VARIABLES

To identify the sources of variation in the dataset, as well as potential confounding variables that would require correcting for, I calculated 50 principal components for each cell type-patient pseudobulk aggregate and performed ANOVA test for correlation of the clinical variables with the PCs (Figure 3-8 A), using the first 40 PCs, which collectively explain >85% of the dataset variance (Figure 3-8 B). Across the components, I observed a gradient of PC scores, rather than a sharp bimodal distinction (Figure 3-8 C). Over multiple components (PCs 13, 17, 18, 20-22, 26, 28, 32-33, 37, 40) I observed association with covariates known to be involved in asthma pathogenesis, such as related to type 2 inflammation (IgE titre, eosinophil count, allergen sensitization, FeNO, eczema), lung function (FEV, FVC), as well as confounding variables (sex, age, BMI, ethnicity). In the differential gene expression analyses presented below, non-pathology related variables (age, sex, BMI, ethnicity) were accounted for in the models. As an example, the principal component with highest correlation values (PC21, Figure 3-8 D) had extreme distances for basophils, DC2s, recruited macrophages, as well as mito-high B cells, (which I demonstrated before (Figure 3-5 and Figure 3-6)), basophils and B cells binding anti-IgE antibody, and macrophages binding the allergens (Figure 3-6). Upon testing for differentially expressed genes at annotation level 2, between asthma and healthy phenotype, I observed numerous differentially expressed genes (Figure 3-8 E), of which the highest number of significantly up and downregulated genes was present in airway macrophages, followed by monocytes, T cells, NK cells, and DC2s (Figure 3-8 E).

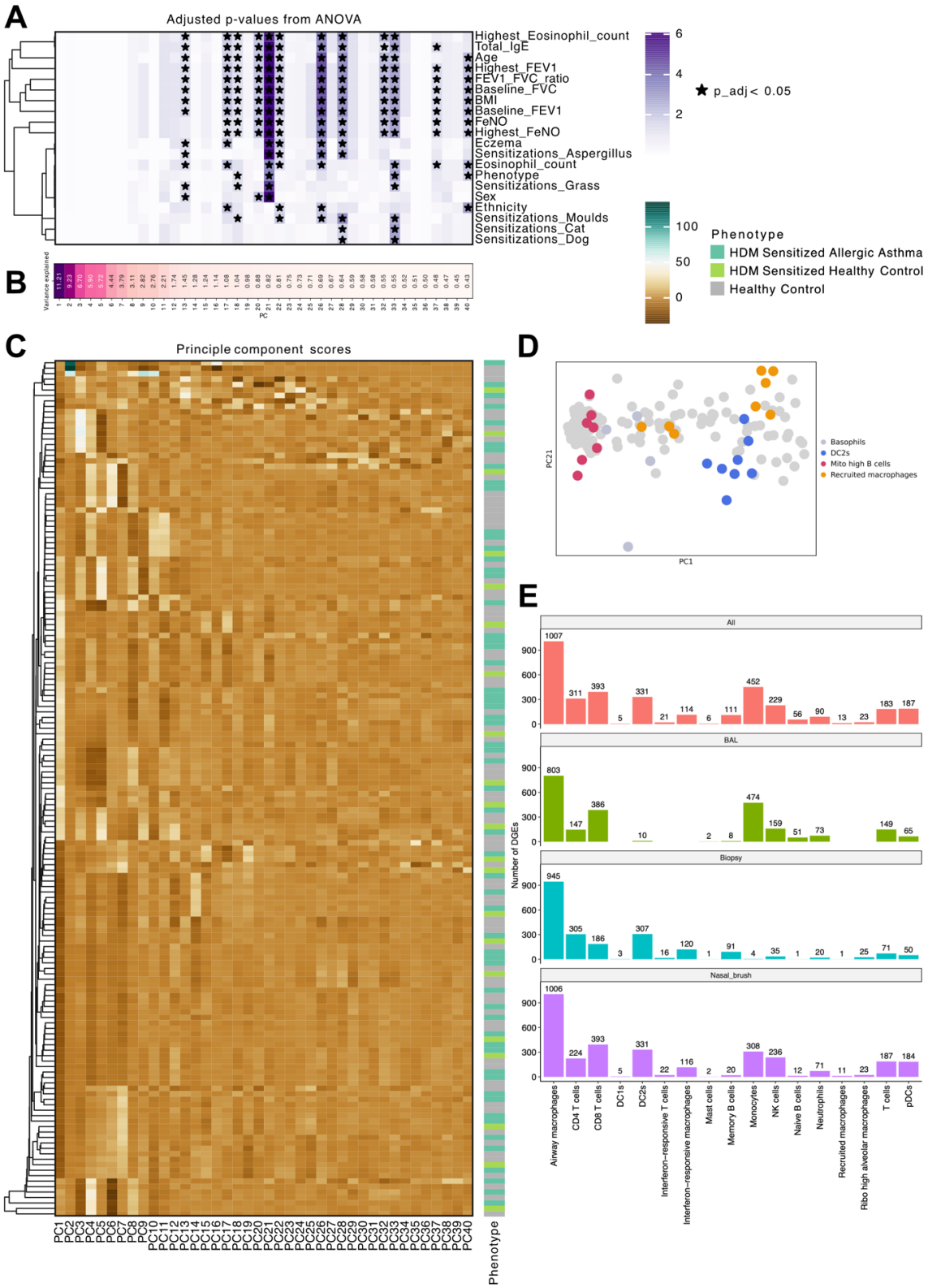

Figure 3-8 Association of principal components calculated from pseduobulk gene expression values with clinical variables

A Correlation score and ANOVA test result for clinical variables collected for the cohort. **B** Percentage of variance explained for each principal component. **C** Principal component scores calculated for each cell type-patient combination. **D** Cell types with the extreme (positive and negative) PC21 values (cell types being sources of greatest variance of PC21). **E** Number of statistically significant genes for all the cells tested and in tissue-specific comparisons when adjusting for correlated clinical variables not related to disease (sex, age, BMI, ethnicity).

3.4.8 ASSOCIATION OF CELL ABUNDANCE WITH CLINICAL VARIABLES

As with testing for the associations with gene expression profiles, I tested for associations of clinical variables with relative cell numbers ('cell abundances'), normalised as a proportion of each differentiation compartment (B and plasma cells, myeloid cells, NK and T cells), to account for the fact that the samples were sorted using flow cytometry and hence the numbers may be biased. Only one component had statistically significant association, PC2 which was correlated with FEV_1 at the time of sampling (baseline FEV_1) and highest recorded FEV_1 (Figure 3-9 A). PC2 explained approximately 19% of dataset variance (Figure 3-9 B). I also observed more profound differences in the principal component scores with PC2 having negative scores in healthy individuals and positive scores in those with asthma (Figure 3-9 C). The PC loading revealed that this positive association in patients with asthma is driven by the abundance scores of airway macrophages and naïve B cells, whereas negative scores are driven by neutrophils and memory B cells (Figure 3-9 D). Comparing mean relative abundances of cell types did not reveal any obvious differences, however it confirmed some clues identified through the PC loadings, namely a modest increase in memory B cell abundance in healthy patients across tissues and neutrophils in nasal brushes compared to allergy and asthma, as well as increase in naïve B cell abundance in biopsies in asthma compared to health (Figure 3-9 E). In order to confirm those observations, whilst accounting for tissue effect, and make robust conclusions informed by statistical significance, we employed differential abundance analysis on KNN graph, as outlined in section 3.4.9.

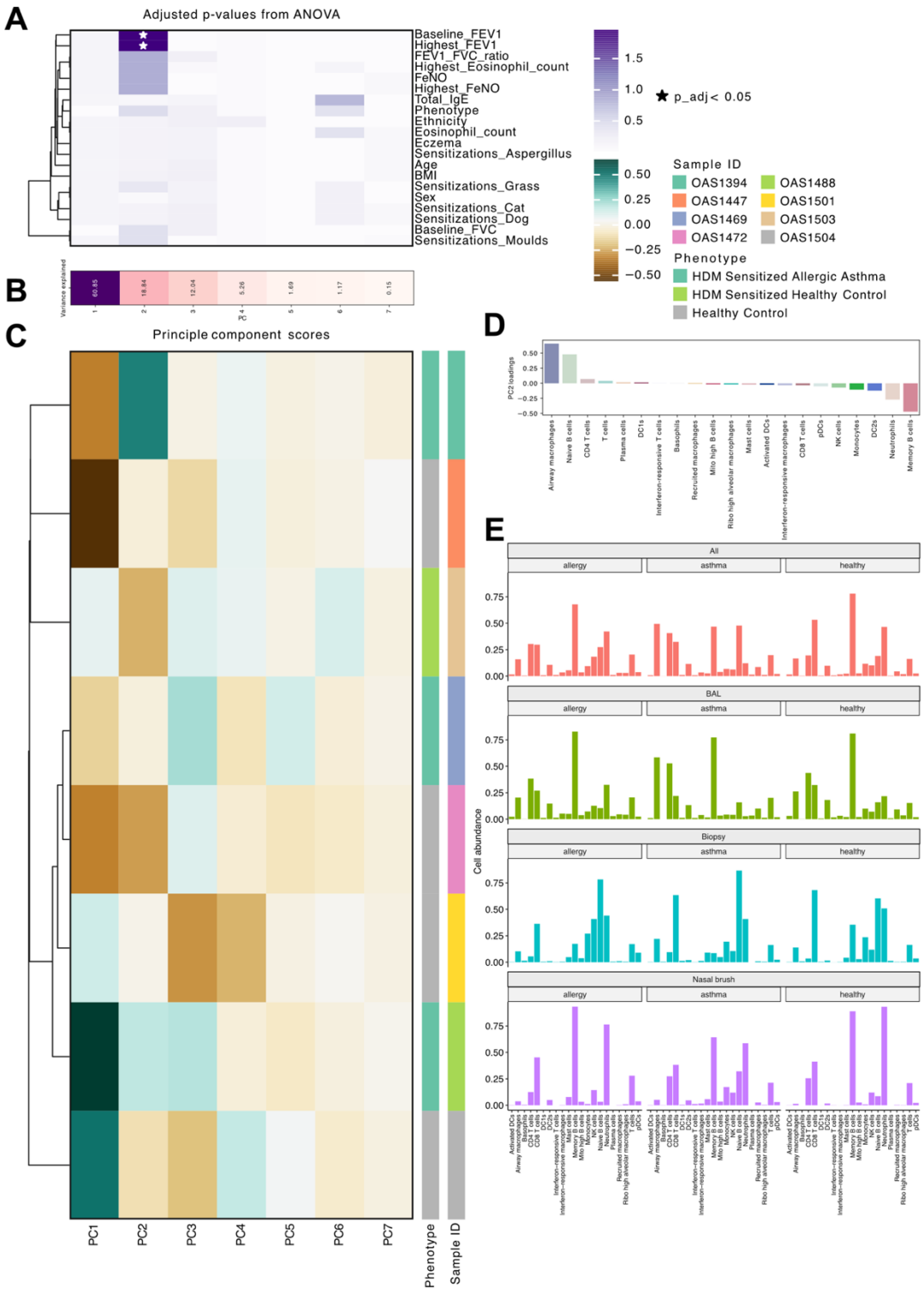

Figure 3-9 Association of principal components calculated from cell abundance values with clinical variables and cell abundance analysis

A Correlation score and ANOVA test result for clinical variables collected for the cohort. **B** Percentage of variance explained for each principal component. **C** Principal component scores calculated for each cell type-patient combination. **D** Principal component loadings for each cell type – positive loadings associate with healthy patients and negative loadings with asthmatic patients. **E** Mean abundance of each cell type (annotation level 2) normalised by compartment across tissues.

3.4.9 CELL ABUNDANCE CHANGES ACROSS TISSUES IN HEALTH AND DISEASE

To investigate the changes in cellular composition across the tissues in health and disease in more detail, I employed *miloR*, a software package which creates cellular neighbourhoods using KNN graph clustering, and assigns those neighbourhoods to cell types using majority vote algorithm [131]. The neighbourhoods identified in the data are shown in Figure 3-10 A, and, on average, consist of samples of approximately 50-100 cells (mean: 79.4, SD: 31.8). I removed the single allergy patient from the data and opted to test for changes in cell abundance across the three sampling sites (tissues – BAL, biopsy, nasal brush) and accounting for the disease status, as well as to test for changes between asthma and healthy accounting for the tissue origin of the sample, and that decision was made for two reasons: i) due to small sample size of data, the inclusion of tissue and phenotype in a single model increases the power to find meaningful and robust changes in the tissue composition and avoids the need of subsetting the data to a single tissue location ii) the phenotype (healthy or asthmatic) is a surrogate for other clinical variables, such as FEV, which were significantly correlated with PCs in the association analysis. Plotting the log fold change of cell abundance in those neighbourhoods per tissue (Figure 3-10 B) revealed that whilst the abundance of some cell types remains consistent across the tissues (pDCs, basophils, monocytes, mast cells), others followed distinct patterns. Of note, the BAL was identified as highly abundant in antigen presenting cells present on the surface of the epithelium, such as recruited and airway macrophages (as previously described [79,169]),

and dendritic cells (DC2s). The biopsies were enriched in NK and T cells (CD8+ as well as CD4+), whereas the nasal brushes showed significant enrichment of abundance of memory B cells, consistent with being the first site of response to inhaled allergens. When contrasting asthma with healthy patients (Figure 3-10 C), I discovered an increase in abundance of airway and recruited macrophages, as well as DC1s, and a decrease of CD8+ T cell abundance, in patients with asthma, across the airway.

As can be seen collectively from Figure 3-10 B and C, whilst some cell types follow a distinct pattern, others form subpopulations within a given cell type, in which one subset increases in abundance whilst another decreases. Taken together, this shows that there are changes in both cellular composition across the sampling sites and the disease status, but also that some cells may form subpopulations based on the local microenvironment, something which could be further investigated using spatial data or trajectory analysis.

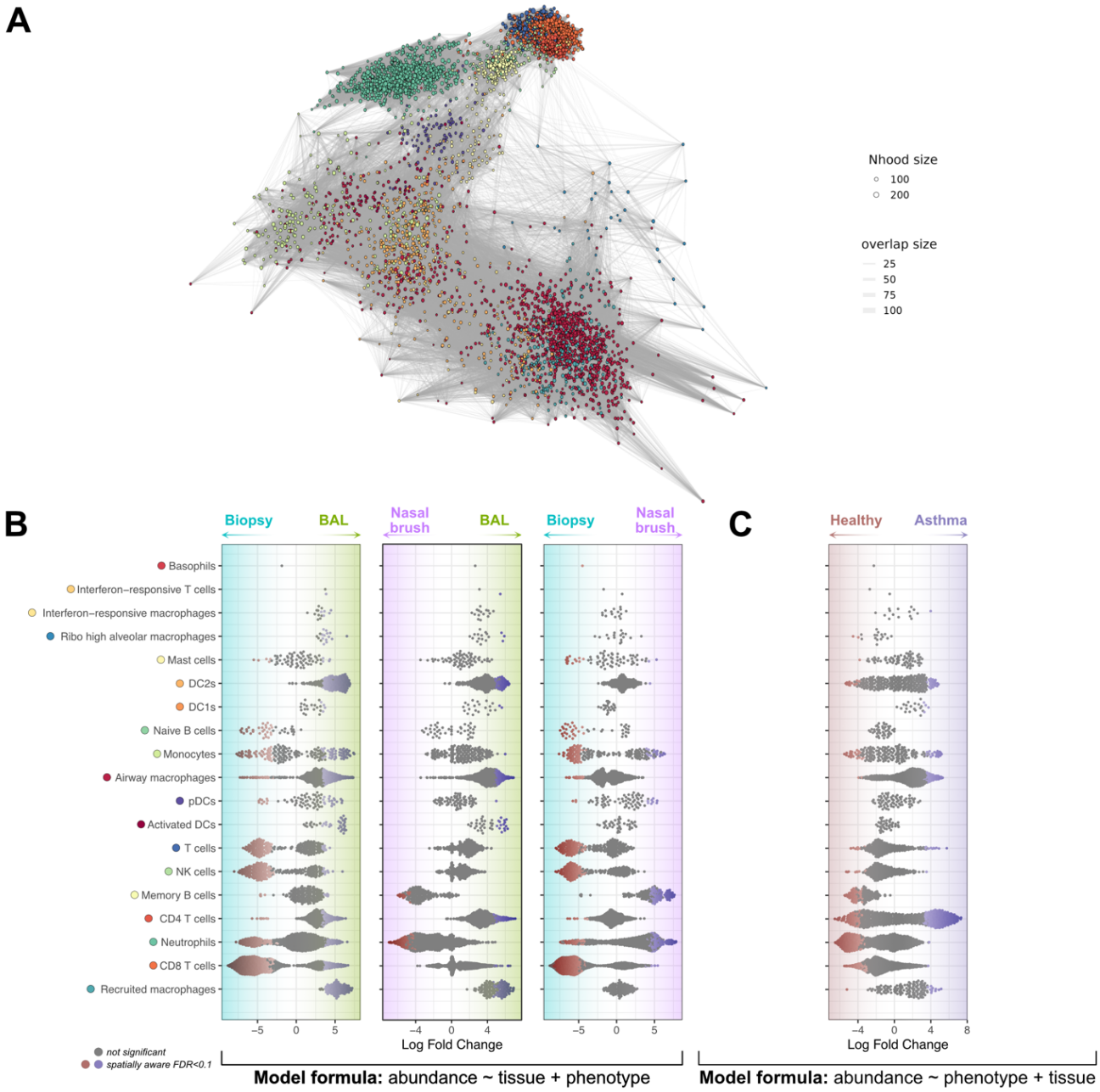

Figure 3-10 Differential abundance of immune cells across respiratory tissues and in health and disease
A Neighbourhood graph showing cell neighbourhoods and relative neighbourhood distances at annotation level 2. **B** Differential abundance as log fold change in cell proportions across the three tissues (biopsy, BAL, nasal brush). **C** Differential abundance as log fold change in cell proportions across the three tissues between health and disease.

3.4.10 DIFFERENTIAL GENE EXPRESSION BETWEEN HEALTH AND DISEASE

To understand the changes in gene expression profiles between health and disease, I contrasted gene expression using a two-part, generalized linear model (GLM) designed for zero-inflated single cell gene expression data, MAST [130]. The GLM formula allowed to correct for covariates identified in the correlation analysis carried out in section 3.4.7, which were age, sex, BMI, and ethnicity. Whilst sex is associated with prevalence of asthma [170] and thus may be confounded with gene expression profiles, it was not the focus of this study. In addition, MAST also corrects for cellular detection rate, defined as the fraction of genes expressed in a given cell. The results are summarised in Figure 3-11 for cell types with highest number of DGEs identified, as seen in section 3.4.7 and Figure 3-8.

In the airway macrophages, genes related to MHC class II antigen presentation, such as *HLA-DRB5*, *HLA-DQB1*, and MHC class I expression (*HLA-C*) are upregulated in patients with allergic asthma compared to healthy controls (Figure 3-11 A), which may suggest upregulation of the antigen presentation machinery in response to the aeroallergen uptake. The MHC class II upregulation was also observed on other immune cells, such as cDC2s (known to be involved in antigen presentation in asthma) and memory B cells (hypothesised to be present in low number in the airway and secreting IgE only in response to allergen challenge), as seen in Figure 3-11 D and F.

Monocytes upregulated CD52 (Figure 3-11 B). CD52 has been trialled as a target for immunotherapy in HDM-sensitised asthma in mouse models of asthma [171], with the mechanism of action thought to relate to the targeting of Th2 cells and ILC2s. However, the observed upregulation in monocytes may suggest that efficacy of this targeting in reducing airway inflammation may be due to an effect on multiple cell types, as CD52 is broadly expressed on T and B cells, and at a lower level on monocytes, mast cells, cDCs, neutrophils, and eosinophils [171].

Neutrophils, a cell type important in some phenotypes of asthma, have not previously been identified in single cell lung datasets [78]. They exhibited a general activation phenotype, with expression of *S100A8, S100A9, S100A11*; these genes encode the S100 proteins that have previously been demonstrated to promote neutrophil motility and adhesion [172] (Figure 3-11 E).

CRIP1, a transcription factor, was identified as a differentially expressed gene in CD4+ T cells, cDC2s, and memory B cells, and has been demonstrated to regulate cytokine-mediated immune response. For instance, it upregulates IL-6 and IL-10 and a downregulates in IFN-γ and IL-2 upon in LPS-stimulated mice [173]. What is more interesting, the *CRIP1* promoter has consensus sequences for glucocorticoid response elements and imparts responsiveness to glucocorticoids, such as dexamethasone, *in vitro* and in rats [174]. It could thus be hypothesised that there is an immunological memory of steroid use in the immune cells of asthma patients, even if they are not on steroid therapy. This would

further be supported to the epigenetic changes (hypermethylation of CpG islands) in the *CRIP1* region discovered in the epithelium of patients with severe asthma [175].

To better understand which cellular processes correspond to the gene signatures observed, I performed gene-set enrichment analysis on a per cell-type level, sorting genes by log(padj) × sign(logFC) metric, with results shown in Figure 3-11 G. Most strikingly, multiple cell types, such as DC1s, monocytes, naïve B cells, and pDCs, showed upregulation of genes involved in NFκB-mediated TNF signalling (Figure 3-11 G). TNF signalling has been demonstrated to have pleiotropic effects in the airway biology and in severe asthma [176], such as mediating airway hyperresponsiveness and neutrophilic inflammation [177,178]. Biologic therapies targeting TNF have been trialled in asthma, however, perhaps due to the broad asthma phenotyping of 'severe, uncontrolled asthma' and not specifically targeting allergic asthma, no benefit in was observed in the patients included in the trial [179].

In addition, signalling by different types of IFNs was upregulated, namely, IFNα in dendritic cells (DC2s and pDCs) and IFNɣ in a small population of interferon-responsive macrophages. IFNα signalling in pDCs is likely due to rhinovirus infection, a common co-morbidity in patients with severe asthma [180,181]. However, previous studies have suggested that this signalling pathway is impaired in allergic asthma, but my findings do not support this. One explanation is that the patients in this study had controlled asthma and were not challenged with HDM allergen before sample collection.

Airway remodelling in asthmatics is thought to be promoted by the type 2 epithelial-mesenchymal transition (EMT), a process wherein epithelial cells are phenotypically transformed into mesenchymal cells, and which results in airway fibrosis [182]. Whilst functional validation would be required, gene set enrichment analysis suggests CD4+ T cells may be involved in signalling pathways mediating EMT, promoting transcriptional changes well characterised fibroblasts and myofibroblasts [183]. A study of EMT in HDM-sensitised asthma would further support this notion, where presence of both HDM as well as an external TGFβ source was shown to be necessary to induce EMT in cell culture [184].

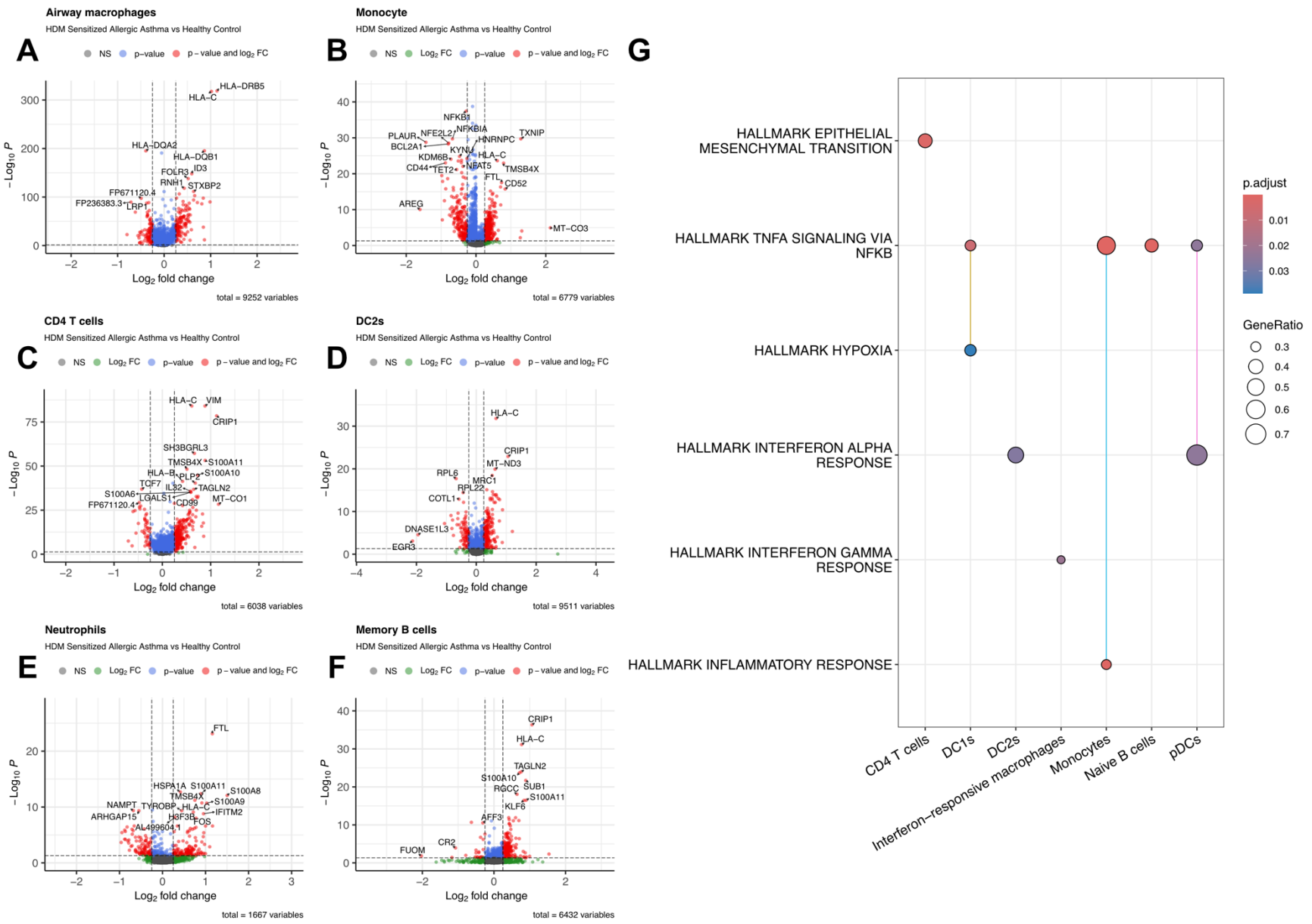

Figure 3-11 Differential gene expression at cell-type level between asthma and health

A-F Volcano plots depicting results of differential gene expression analysis in key cell types at annotation level 2. Log2 fold change is shown on x-axis and negative log10(p_{adj}) on the y-axis.

B Results of MSigDB Hallmark dataset gene-set enrichment analysis at per-cell level at annotation level 2. GeneRatio is the ratio of genes enriched to the overall Hallmark gene set size.

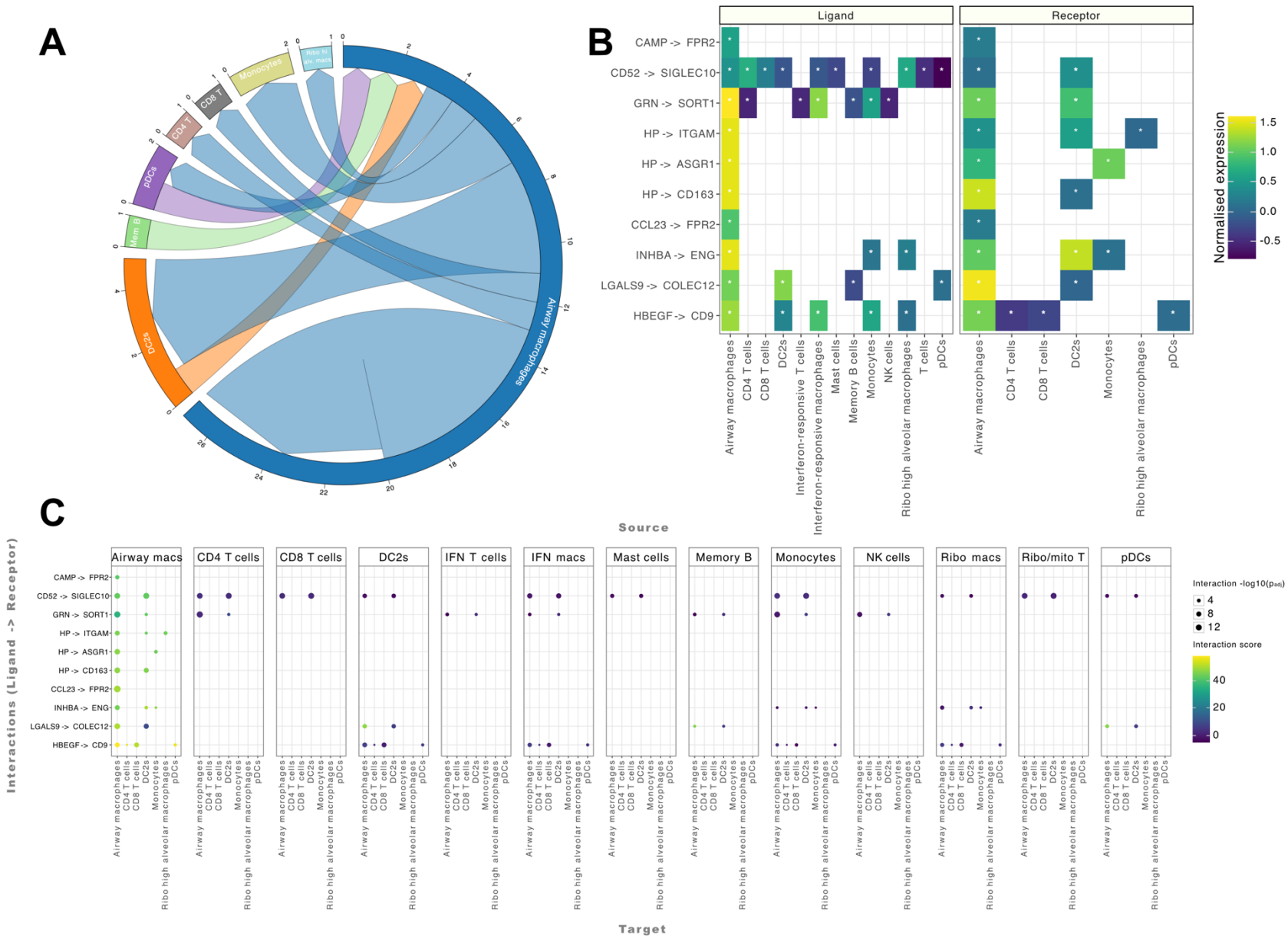

Figure 3-12 Cell-cell differential interaction analysis between asthma and health

A Circos plot summarising all the statistically significant (p_{adj}<0.05) interactions between the cell types. Numbers around the circumference of the plot indicate the interaction count. **B** Normalised expression levels of genes involved in top 10 ligand-interaction pairs, sorted by LIANA+ interaction score. **C** Interaction scores of top 10 ligand-interaction pairs per cell type.

3.4.11 IMMUNE CELL INTERACTIONS IN ALLERGIC ASTHMA

As multiple cells can be involved in immunological processes in asthma, I opted to analyse the cell-cell interactions in the immune compartment to understand the cellular communication patterns. I used the differentially expressed genes in each cell type and the *LIANA+* framework [133] to map the differential interactions between asthma and health, with results presented in Figure 3-12. Overall, consistent with the highest number of differentially expressed genes, the airway macrophages showed the greatest absolute number of cell-cell statistically significant interactions, 27, with both autocrine and paracrine signalling in approximately equal proportions (Figure 3-12 A). The interactions between dendritic cell subsets (DC2s and pDCs) and airway macrophages were bidirectional, with each cell type acting as both a sender and a receiver for distinct ligand-receptor pairs. Memory B cells also interacted with airway macrophages. Overall, the cell-cell interaction analysis highlights a potentially novel role for airway macrophages in airway inflammation associated with allergic asthma. The validity of this analysis is supported by the identification of known interaction patterns, such as the well-documented involvement of cDC2s and B cells in the pathogenesis of airway inflammation [4].

To further understand the nature of individual interactions, I analysed the gene expression patterns of the ligand-target pairs and used the interaction score to prioritise them (Figure 3-12 B and C). *CD52* was upregulated on 10 cell types at annotation level 2, including CD4+, CD8+ and mito- and ribo-high T cells ('T cells'), with its receptor *SIGLEC10* being differentially expressed on the surface of airway macrophages and DC2s (Figure 3-12 B).

Human and mouse antigen-activated T cells with high expression of the lymphocyte surface marker CD52 were found to suppress other T cells [171]. As noted in the cited study, SIGLEC10, which interacts with CD52, is expressed on various immune cells involved in both innate and adaptive immunity [185], thus its effects may extend beyond T cells to influence other immune cell types.

FPR2 was found to interact with *CCL23* and *CAMP*. FPR2 is a versatile receptor which can trigger both pro-inflammatory and anti-inflammatory responses, and is involved in sensing microbial peptides, as well as endogenous lipids, peptides, and proteins [186]. One of its known roles is promoting the migration of dendritic cells into the bronchioalveolar area during allergic airway inflammation in mice [187]. In general, FPR2 agonists have recently attracted interest as targets for novel medications in asthma [188].

In addition to the *FPR2* and *CD52* interactions, a GRN-sortilin signalling pathway has been identified with the *SORT1* gene being upregulated in cDC2s and airway macrophages. The membrane glycoprotein encoded by the *SORT1* gene regulates lipid metabolism and is involved in pathways promoting inflammation and vascular calcification [189,190]. In cardiac tissue, sortilin promotes fibrosis and calcification by shuttling alkaline phosphatase into vesicles secreted by smooth muscle cells [189,190]. Beyond orchestrating inflammation and fibrosis in cardiac disease, *SORT1* may also have a role in the development of fibrosis in the airway.

3.4.12 TYPE 2 INFLAMMATION AND IGE SIGNAL PROFILING IN ALLERGIC ASTHMA

Type 2 cytokines are the key mediators of inflammation in allergic, and more broadly in eosinophilic asthma. To locate the cellular source of type 2 cytokines in allergy, allergic asthma, and health, I mapped the expression of genes encoding type 2 cytokines in a subset of cells hypothesised to be involved in the inflammation, as seen in Figure 3-13 A. I observed that *IL4* is produced by basophils in homeostasis. In contrast, in allergic asthma, but not allergy alone, *IL4* expression appears to have two cellular sources – not only the basophils, but also the Th2 cells. A different expression pattern was observed for the cytokines *IL5* and *IL13*. These were expressed by Th2 and mast cells at steady state but were upregulated in Th2 cells in allergy and asthma (Figure 3-13 A).

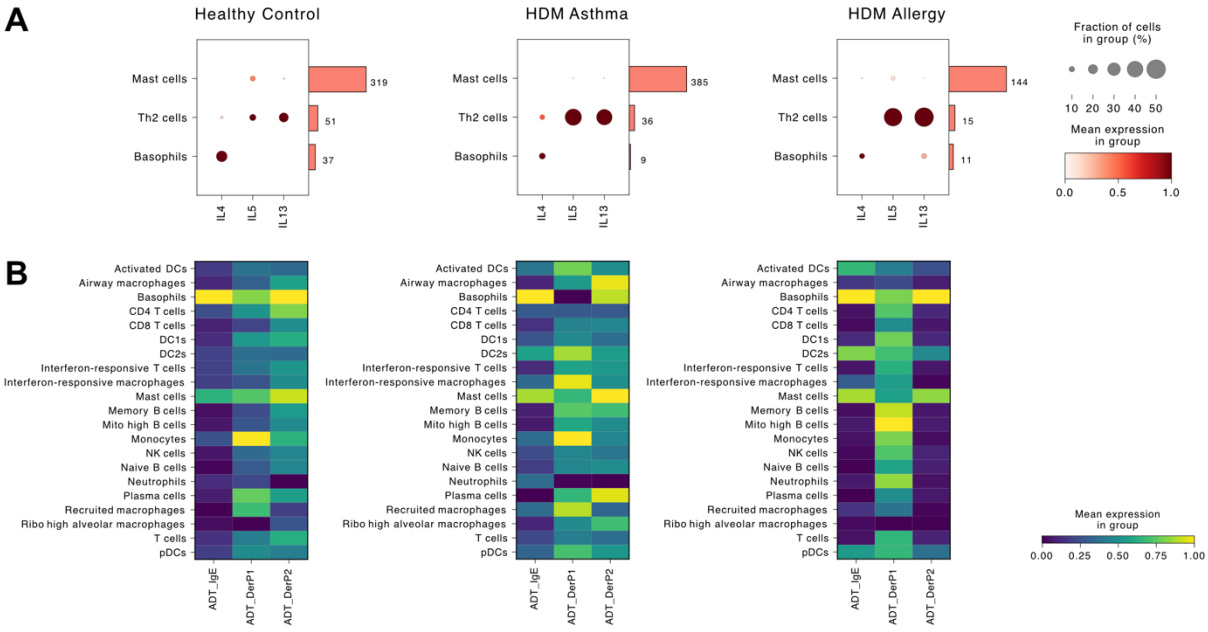

Figure 3-13 Type-2 cytokine, IgE, and allergen profiling in healthy, allergic, and allergic asthmatic patients

A Dotplots showing the magnitude of expression (dot colour), the percentage of cells expressing (dot size) and the total number of cells (bar height) of three hallmark cytokines of type-2 inflammation (IL4, IL5, IL13). **B** Heatmaps illustrating the normalised antibody staining levels of three antibodies important in severe allergic asthma with HDM sensitization (IgE, *DerP*1, *DerP*2) at annotation level 2.

Upon binding of the aeroallergen in the airway, and its presentation by Th2 cells to B cells, the B cells become activated, and as short-lived plasmablasts secrete IgE. To map which cells bind the aeroallergen, as well as the IgE secreted as the result of antigen-induced inflammation, I examined the binding patterns of the allergen and IgE molecules in allergy, allergic asthma, and health, as seen in Figure 3-13 B. FcεRI receptor, which binds IgE, is constitutively expressed on the surface of mast cells and basophils, but in humans can also be found on dendritic cells, both conventional and plasmacytoid, and monocytes [191]. As expected, across the phenotypes, I observed that basophils and mast cells displayed the highest levels of binding of IgE, however in asthma and allergy, I also saw evidence of IgE binding by DC2s, activated DCs, and to a lesser extent pDCs. Binding of IgE to mast cells and basophils, but also activated dendritic cells and cDC2s would be consistent with its role in driving inflammation in allergy and asthma. *DerP* 1 was bound strongly by monocytes and macrophages, with other cells also exhibiting antigen binding to varying degrees in other phenotypes. It has been shown that during inflammation, monocytes become activated and acquire some DC-like features such as expression of CD11C and MHCII [192], which could enable binding and presentation of the aeroallergen. In contrast, *DerP 2* was bound by a comparatively smaller subset of cells, mast cells and basophils, possibly through IgE accumulated on their surfaces (Figure 3-13 B). Of interest, the binding patterns were drastically different in individuals with allergic asthma, with evidence of binding by other cell types, such as plasma cells and airway macrophages (Figure 3-13 B).

It must be acknowledged that both IgE and allergen signal showed a high level of background, despite the use of isotype controls and background-aware normalisation method (dsb). Thus neither non-specific binding of IgE to Fc receptors on various cell types, nor allergen adhering to cell surface, cannot be completely excluded, even taking into account the fact that the experimental protocol incorporates the blocking of Fc receptor before staining with antibodies. An improved experimental design would include negative controls for allergen binding, such as influenza or SARS-CoV-2 antigen, which should be mutually exclusive on immune cells. Samples from patients who have undergone allergen challenge before bronchoscopy could be a positive control. Further, antibodies against other immunoglobulin classes (IgM, IgG, IgA, IgD) could be added to the staining panel as negative controls to demonstrate specificity of IgE binding to basophils and mast cells.

3.4.13 CLONAL EXPANSION OF T CELLS IN ALLERGIC ASTHMA

As T cells are key regulators of immune response in asthma, and display broad changes in gene expression, I decided to further investigate and verify if the T cells clonally expand upon antigen engagement (Figure 3-14). The number of T cells captured did not enable me to call clonotypes in individual tissues, however, I was able to conduct a preliminary analysis of clonal expansion of cells using *immunarch* suite of tools [134].

Investigating individual patients, I found that 2 out of 4 healthy controls (50%) did not show any signs of clonal expansion. The remaining two did, at a similar proportion (Figure 3-14 A). Conversely, in HDM-sensitised allergic asthmatics, each patient showed 40-50% of cells with at least 2 clones, and one patient had a small proportion of cells with 11 or more clones (Figure 3-14 A).

Splitting these T cells according to annotation level 2 and contrasting asthma with healthy controls revealed no statistically significant clonal expansion in any given subset (Figure 3-14 B). I hypothesize that the reason for this is two-fold: the number of T cells with captured TCR sequences in the dataset is small (16,229 cells) compared to the number of transcriptomes (36,229 cells), decreasing the power of statistical analysis, further compounded by the biological signal as only a subset of T cells would likely be expanded: it has been previously reported that there is CD4+ T cell expansion in asthma (consistent with antigen presentation and B cell activation mechanism, as well as type 2 cytokine production) but not CD8+ T cell expansion [193]. Consistent with these observations, the

CD4+ compartment exhibited some of the lowest p values in asthma-health comparisons (even though none reached statistical significance). Due to the low number of BCR sequences captured (2,794 BCRs out of 3,144 B cells in the dataset), it was not possible to profile the B cell clonality. Mouse studies of HDM sensitisation suggest that the likely number B and plasma cells needed to detect differences is on the order of tens to hundreds of thousands [194]. Moreover, and ideal scenario would compare lung tissue with the blood to establish where the IgE+ cells home to.

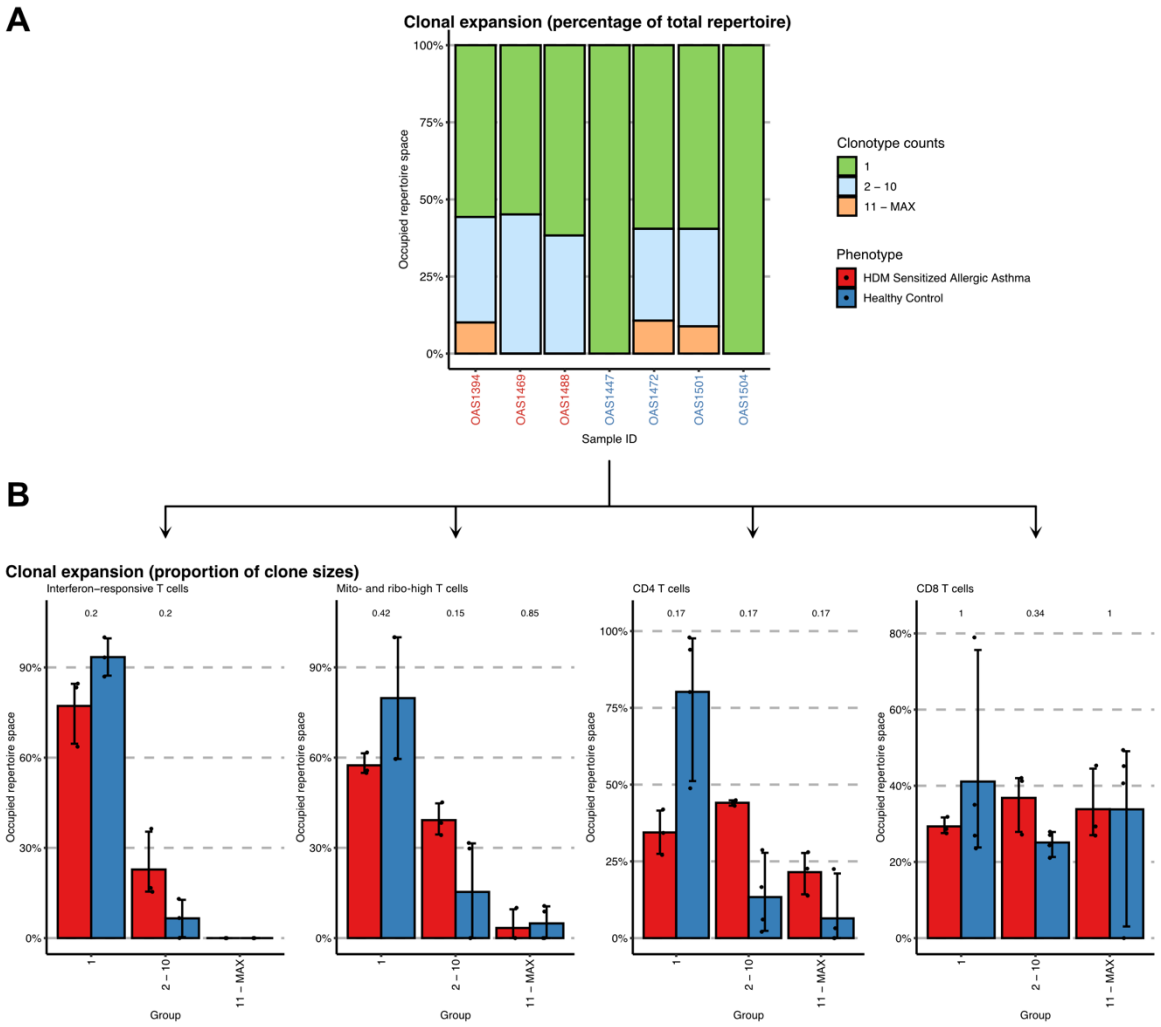

Figure 3-14 Clonotype analysis of T cell compartment in health and disease

A. Per-patient clonotype analysis across all T cell subtypes with expansion brackets 1, 2-10, 11-more cells with identical TCR. The colour of patient ID corresponds to phenotype (healthy, asthmatic). **B.** Per cell-type clonotype analysis with the same expansion buckets. The number above the bar pairs indicates Wilcoxon signed-rank test p value, the error bars represent 95% confidence intervals.

3.5 DISCUSSION

In this Chapter, I present a tri-modal dataset of patients with severe allergic eosinophilic asthma with defined allergen sensitization and healthy controls, together with detailed clinical metadata and cell-type annotations. The dataset presented is a comprehensive single-cell resource of severe allergic asthma, which surpasses the only existing single-cell dataset of allergic asthma with HDM sensitisation [153] in terms of size (with ~40% more cells in this study), annotation detail, and number of modalities (RNA versus RNA/VDJ/protein) collected. The dataset is of high quality, with minimal ambiguous cell clusters excluded from analysis and distinct gene expression profiles across cell types. The dataset has some limitations – there may be technical artefacts due to enzymatic dissociation and the VDJ clonotypes identified are low in number.

I profiled the immune cell types in the lung, defining a total of 52 unique cell states, and capturing a greater heterogeneity of immune cell states compared to larger lung atlasing initiatives given the immune cell enrichment performed prior to single-cell sequencing. I also demonstrated that reference mapping for cell type annotation cannot always be relied on and can mislabel key cell types, such as neutrophils, highlighting the importance of expert curation of initial labels. It is worth noting that the annotation approach I have undertaken also has its limitations, as labelling clusters generated by an unsupervised algorithm (Leiden algorithm) can sometimes lead to overclustering and produce artificial populations of cells with no biological relevance, such that downstream validation of cell

types and states through spatial methods and/or protein validation would likely be beneficial.

I profiled the changes of relative abundance of cell types across sample types commonly obtained from research bronchoscopies, including biopsy, nasal brushing, and BAL. There are profound differences between the upper and lower airway, with increased numbers of memory B cells and neutrophils in the nasal samples and mast cells and DC2s in the lung samples. Nasal samples and biopsies contain more diverse cell types compared to BAL, which predominantly comprises airway and recruited macrophages. Interestingly, BAL also had an increased number of CD4+ T cells.

I used an antibody specific to IgE, in addition to gene expression data, to profile not only the cells producing IgE, but also the effector cells binding it. Cells which bind IgE are of clinical relevance, as IgE+ plasma cells/plasmablasts are understood to release IgE in response to allergen stimulation in allergic asthma [4]. The findings reveal that there is low level expression of *IGHE* in B and plasma cells without antigen challenge and demonstrate binding of IgE on the surface of mast cells and basophils, likely through the high-affinity IgE receptor FcεRI [149,195]. These findings are consistent with recent reports indicating the presence of only a transient population of IgE-secreting cells in the airway following stimulation, along with the emergence of IgG+ memory B cells during steady-state conditions [151].

This chapter also introduces a novel technique to map allergen binding to specific cell types, which differs from commercially available solutions in that it offers direct conjugation of the DNA barcode to the protein of interest. Most strikingly, I demonstrated that the binding profiles of allergen proteins not only differ between asthma and health, but that they also depend on the protein in question, with *DerP* 1 binding to many more cell types than *DerP* 2. This can be explained by the biology of these proteins, with *DerP* 1 being a proteinase capable of cleaving tight junctions, whereas *DerP* 2 likely binds to the surface of the immune cells as a TLR agonist, although some unspecific binding of *DerP* 1 cannot be completely ruled out and requires further confirmation. A good approach to exclude non-specific binding would be to use a negative control antigen, such as that of influenza or another aeroallergen patients were not sensitised to, in the antibody panel. The allergen staining data in immune cells supports the notion of allergic inflammation being a key driver of inflammation in allergic asthma, with antigen-presenting cells, including dendritic cells and monocytes/macrophages, binding the proteins responsible for inflammation.

Due to poor definition and resulting large heterogeneity of asthma, multiple studies report individual clinical parameters having a stronger predictive value than the diagnosis itself [62,22]. Multiple clinical variables, such as FEV_1, FeNO, IgE measurements, as well as history of atopy, associate with cell type abundance, gene expression profiles, or both, underlining the importance of defining novel and consistent numerical criteria to classify asthma and its phenotypes.

Consistent with previous studies, I observed an increase in abundance of pathogenic airway macrophages in asthma, which also display a unique transcriptional signature. I characterised transcriptional changes likely resulting from ICS treatment, such as the upregulation of *CRIP1*, and those that may be indicative of novel biological pathways, such as MHC class II antigen presentation machinery and regulatory molecule CD52. I further profiled the changes by mapping out cell-cell interactions and show that the airway macrophages increase both autocrine and paracrine signalling, with some key ligands associated with asthma, such as CD52 and SIGLEC10 being involved, providing new potential targets for experimental validation.

Probing the relative abundances of immune cell types in asthma and health reveals modest changes in populations of most immune cell types, but a pronounced increase in the abundance of airway and recruited macrophages. In further chapters I demonstrated that this apparent lack of changes in abundance may be attributed to small sample size and lack of non-immune compartment.

Finally, I profiled the VDJ repertoire of T cells and show interesting trends in the clonality of CD4+ T cells, with some evidence of clonal expansion of those cells in asthma, presumably due to continued and repeated allergen exposure. Unfortunately, limited number of captured VDJ sequences and overall cell number preclude me from more in-depth analysis of T cell repertoire and make the analysis of B cell repertoire impossible, highlighting the need of large number of cells for those types of studies, often only obtainable from the blood.

Bronchoscopy samples of individuals with asthma are extremely rare and difficult to obtain, in particular from patients with severe, active disease of the respiratory system. There are inherent risks of invasive research procedures like bronchoscopy, as well as the need to manage the disease and avoid exacerbations. The samples are small in size, with less than 100,000 cells normally collected from the biopsies, with further cell loss of 50-70% during handling and sorting. Immune cells, including rare cell types such as Th2 cells and ILCs, form the minority of cells in the airway, but drive inflammation in asthma. To enable appropriately powered statistical analysis, it was therefore decided to limit the study to the immune cells, removing the non-immune cell types from the dataset. The non-immune cells, the stroma and the epithelium, remain interesting populations to characterise, something I attempt in Chapter 5.

Given the challenges of obtaining airway samples, patients were not BMI, sex or age matched. However, my specific conclusions are unlikely to be influenced by age and sex. Moreover, association of patient characteristics with the gene expression and cell type abundance was tested, and a correction was made to account for differences where those associations were found to be statistically significant (testing for differential gene expression). Due to the same challenges, the patients were not subject to allergen challenge before collecting the sample, as is common in mouse models of asthma [196] – instead, bronchoscopy was performed on patients who had well-controlled severe asthma. Thus, a state of chronic low-level inflammation in those patients is described, with allergen stimulation *ex vivo* shortly before transcriptome profiling.

Eosinophils and their IL5 signalling are an important part of the understood pathophysiology of asthma, as described in Chapter 1, yet they are not present in the dataset I generated, leading to limitations in drawing conclusions about relative importance of cell types driving inflammation in asthma through mechanisms such as differences in abundance, gene expression, and intercellular interactions. This cell type has not successfully been captured using any commercial fluidics-based platform, capturing either RNA or DNA modality (RNAseq, ATACseq). In humans, eosinophils only seem to be identifiable with gentle gravity-based cell capture approaches, as opposed to microfluidics methods, and only in the blood [107,197].

In summary, I provide a rich resource for studying severe allergic eosinophilic asthma at single-cell level, as well as understanding the mechanisms of allergen action in disease. I show changes at gene expression, abundance, cell-cell interaction, and clonal expansion levels, generating new hypotheses which will hopefully start new avenues of research, and highlighting important limitations of single cell studies in tissue.

4 EXPLORATORY ANALYSIS OF LUNG BIOPSIES IN ALLERGIC ASTHMA AND HEALTH WITH SPATIAL TRANSCRIPTOMICS APPROACHES

4.1 INTRODUCTION

Whilst the cell types in the are being characterised at increasing resolutions, with multiple new cell types and inflammatory mechanisms reported in the recent years, there is still limited understanding of the three-dimensional organisation in the airway, and how the architecture enables the respiratory system to perform its role, mostly stemming from conventional techniques such as microscopy, which offer limited analyte plexity and resolution.

In this Chapter, I complement the characterisation of immune cell types in allergic asthma and healthy individuals by a description of changes occurring in the epithelium and the underlying stromal cells. The use of both *in situ* and sequencing-based techniques to profile lung samples has been recently reviewed [198]. Here, I test the feasibility of obtaining high-resolution spatial transcriptomics data from FFPE biopsies commonly collected in the clinic, using a large probe panel targetting 5,000 genes. I describe the cellular organisation of the airway and define distinct structures, such as submucosal glands (SMGs), invaginations consisting of mucous and serous cells, which contribute to airway homeostasis and antimicrobial protection through secretion of the mucus [74].

4.2 HYPHOTESIS

Lung samples can be studied at single cell level in the spatial complex using *in situ* spatial transcriptomics, which captures the structural and compositional diversity of the airway. Through high gene-plexity, spatial transcriptomics enables categorisation of multiple rare and common cell types in the airway. Spatial differences exist in the organisation of lung tissue between asthmatic and healthy individuals. In addition to structural changes, gene expression changes also can be captured between healthy controls and asthma patients.

4.3 AIMS OF THIS CHAPTER

- To characterise the cell types and anatomical structures captured with *in situ* spatial transcriptomics
- To verify the presence of immune cells involved in type-2 inflammation (mast cells, plasma cells, eosinophils) in the FFPE-preserved lung biopsies
- To interrogate how T cells and B cells communicate spatially in the airway mucosa
- To identify immune and epithelial niches present in the lung in healthy and diseased individuals

4.4 RESULTS

4.4.1 COHORT CHARACTERISTICS

To further characterise the differences between severe allergic eosinophilic asthma and healthy individuals, and expand the comparisons to address differences in non-immune cells, with the help of the clinical team we recruited further patients and accessed archival tissue and conducted *in situ* spatial transcriptomic profiling of a panel of 5,001 genes on an expanded cohort including patients from Chapter 3, as well as new patients with matching phenotypes, as summarised in Figure 4-1 A. The patients were well-matched with regards to biological sex, smoking status, with healthy individuals younger and with lower BMI, as seen in Table 4-1. The reason for lower age and BMI of healthy controls is discussed in Chapter 3. The mean values of parameters relating to allergic inflammation (total IgE, eosinophil counts) and atopy (number of sensitizations) were higher in patients with allergy or asthma compared to healthy controls.

Table 4-1 Cohort characteristics for spatial transcriptomics project

	Healthy control (n=7)	HDM sensitization (n=8)	Allergic asthma + HDM sensitization (n=8)
Age, years (mean ± SD)	26.3 ± 5.0	38.6 ± 14.6	45.9 ± 13.2
Male/female	2/5	3/5	4/4
BMI, kg/m^2 (mean ± SD)	23.29 ± 2.75	24.86 ± 2.85	30.54 ± 4.97
Eosinophil count [on the day], counts x 10^9/ml (mean ± SD)	1.10 ± 0.12	1.65 ± 0.15	2.02 ± 0.13
Eosinophil count [highest ever], counts x 10^9/ml (mean ± SD)	1.43 ± 0.17	2.37 ± 0.15	5.00 ± 0.48
HDM at screening, kUA/l (mean ± SD)* Detection limits: 0.01-0.35	0.05 ± 0.09	14.25 ± 13.64	39.84 ± 31.06
Number of Sensitizations [6 tested: HDM, Grass + Trees, Cat, Dog, Moulds, Aspergillus] (mean ± SD)	0.29 ± 0.49	2.13 ± 0.99	4.13 ± 1.36
Total IgE at screening IU/mL **	87.33 ± 187.69	378.21 ± 549.32	913.23 ± 834.42
Smoker (%)	14%	13%	13%
Current smoker (%)	14%	0%	13%
Ex-smoker (%)	0%	13%	0%
Smoking status unknown (%)	29%	0%	0%
* - if the result was below detection limit, it was recorded as 0.01 ** - if the result was below detection limit, it was recorded as 2			

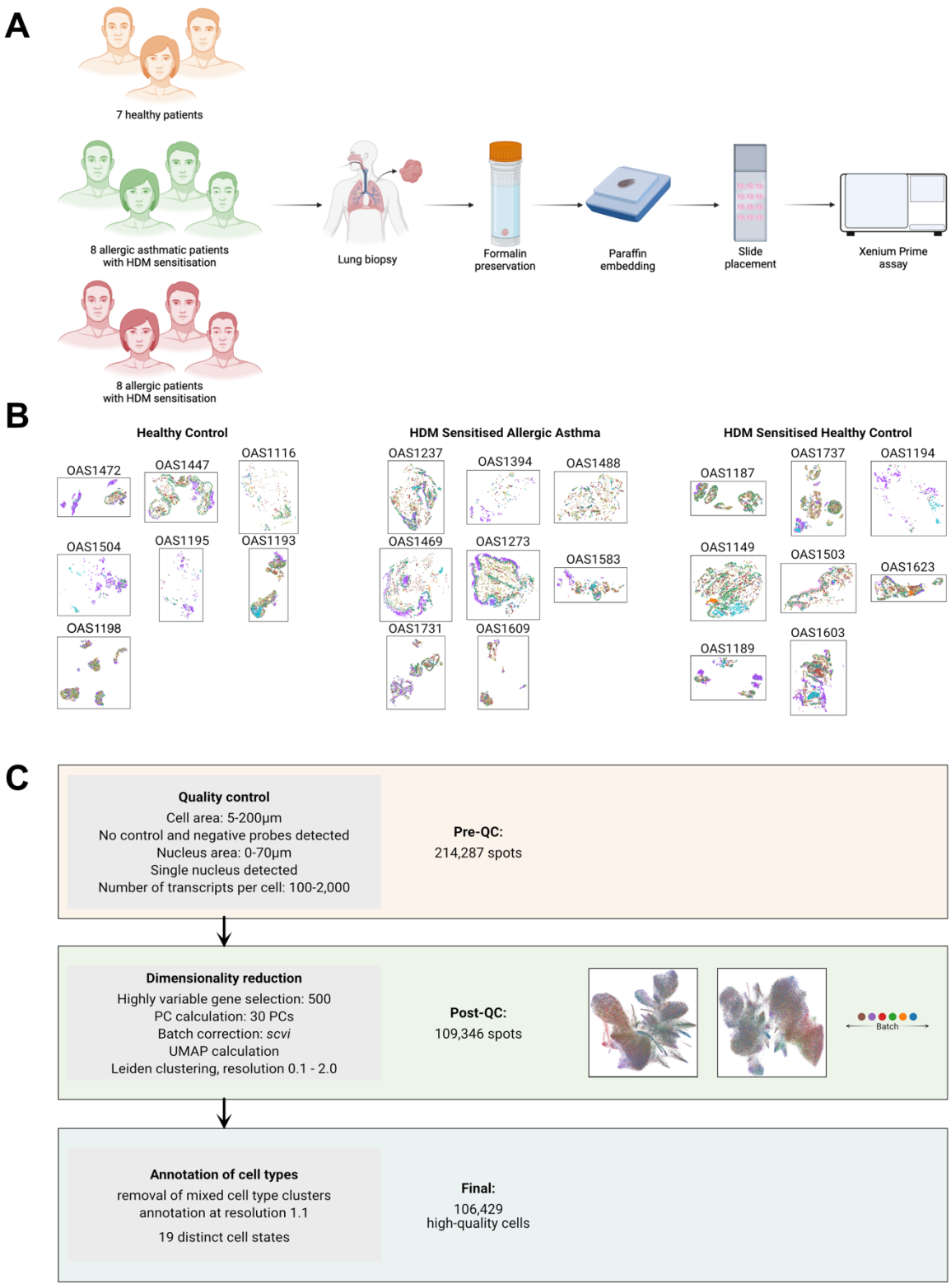

Figure 4-1 Spatial transcriptomics study design and data processing overview

A Overview of the patient phenotypes (allergic with HDM sensitization, healthy with HDM sensitization, healthy control) and the experimental processing, consisting of biopsy collection as part of a research bronchoscopy, FFPE preservation and embedding, placement on the Xenium slide, and processing in Xenium instrument. **B** Overview of tissue images acquired per patient (coloured by cell type). **C** Summary of key preprocessing steps with associated cell and spot counts. The key steps are QC filtering, dimensionality reduction with batch correction (batch = slide ID + condition) and annotation of cell types.

4.4.2 DATA PREPROCESSING

A total of 23 sections were captured across two Xenium slides, with patient phenotypes randomly distributed between the slides. The samples were of variable quality, with some displaying excellent cell capture rates matching the expected tissue architecture (OAS1198, OAS1273), whilst others only showed a proportion of high-quality cells (OAS1194), as seen in Figure 4-1 B. 214,287 spots were collected in total (Figure 4-1 C). Almost all the spots in the dataset were segmented using either an interior 18S stain, or the boundary stain (67.1% and 31.5%, respectively, see Figure 4-2 A), with less than 2% segmented using the more conventional nuclear expansion method, underlining that both technical and computational advances were needed to solve the segmentation problem, which posed a significant obstacle to generating high-quality results in the past [199].

The count matrices were extracted with *squidpy* and processed with *panpipes* to ensure consistency with Chapter 3.

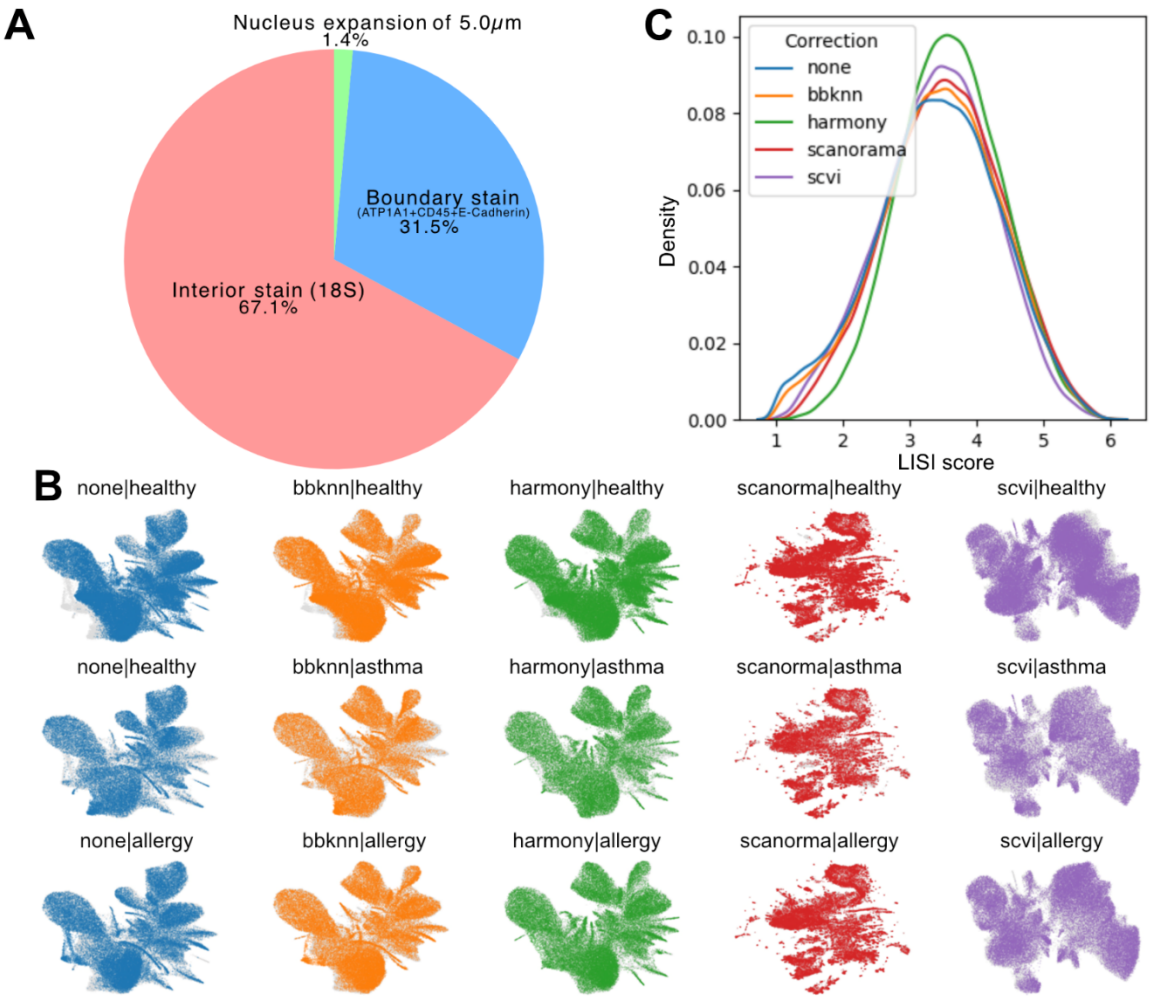

Figure 4-2 Summary of cell boundary detection results and optimisation of batch correction algorithms

A Summary of the methods of cell boundary detection used by Xenium algorithm on the dataset.

B Comparison of four data integration methods on the UMAP embedding coloured by patient diagnosis (healthy control, allergy, and asthma). **C** Local Inverse Simpson's Index (LISI)

Cells unusually small or large were excluded. Epithelial cell diameter ranges from 200 μm^3 for goblet to 300 μm^3 for ciliated cells [75], implying an area in the range of 40 - 55 μm^2 (assuming a perfect sphere), with immune cells displaying a greater variability in size, with large plasma cells having an area of approximately 300 μm^2 [200]. Thus, generous filtering threshold of cell area ranging from 5 to 200 μm^2 and nuclear area smaller than 70 μm^2 was applied, with nucleus size filter applied in addition to cell size filter to exclude doublets. The total number of transcripts detected ranged from 2 to 6059, with a median value of 122 and a standard deviation of 236, and a generous threshold of 100-2,000 transcripts per spot applied to maximise the number of cells in the analysis. These filtering steps, summarised in Figure 4-1 C, yielded 109,346 spots. The following processing steps were consistent with Chapter 3, with the number of HVGs decreased to 500 (Figure 4-1 C). I made the decision to use 500 HVGs (~10% of the panel size) to more closely mirror analyses performed in the literature to-date [201] with smaller panel sizes on one hand, and to use a similar proportion of overall dataset to that taken in single cell genomics (typically between 2,000 and 3,000 out of approximately 20,000 protein-coding genes expressed in human transcriptome [202]. In addition, the commercially available gene panel is designed for use with many tissue types, so it can reasonably be assumed that many genes will not be expressed in the lung, hence it would not be justified to use all genes for dimensionality reduction.

Out of the integration methods tested (bbknn, harmony, scanorama, scvi), visually the best result was produced by scvi, with dataset splitting distinctly into two lobes – immune (left) and non-immune (right), and good overall mixing of batches without donor-specific orphan clusters (Figure 4-2 B). Interestingly, the LISI score was comparable for every method, with peak score in the range of 3-4 (Figure 4-2 C), possibly pointing to its limited utility in assessing the integration in spatial datasets.

The cells were annotated with aid of HLCA mapping, with further low-quality cell excluded, producing a final dataset consisting of 106,429 high-quality cells (Figure 4-2 C), about 50% of the original dataset size, underlying the importance of manual data curation in spatial transcriptomics studies.

4.4.3 CELL TYPE ANNOTATION

Due to the smaller gene space captured in *in situ* methods compared to single-cell and sequencing-based methods, only limited cell type heterogeneity can be captured. Nonetheless, cell types identified recapitulated broad cell types annotated in Chapter 3 at annotation level 1, with some rare cell types present in the epithelium also identified due to their unique markers that did not overlap with other cell types. The annotation strategy along with the key markers is presented in Figure 4-3 A.

In the immune compartment, I identified the major lymphocytes – B and plasma cells responsible for IgE production in type-2 inflammation (B cells: *MS4A1*, p_{adj}<1E-300; *CD19*, p_{adj}=2.87E-163; *CD79B*, p_{adj}=1.64E-33 and plasma cells: *XBP1*, p_{adj}<1E300, *TNFRSF17*, p_{adj}=2.36E-186) and tissue-resident memory follicular helper CD4+ T cells (*CD4*, p_{adj}=5.60E-79; *CXCR4*, p_{adj}=2.27E-75) and cytotoxic CD8+ T cells (*CD8A*, p_{adj}<1E-300; *CXCR6*, p_{adj}=6.89E-32). In the myeloid cell compartment, I observed mast cells, effectors in allergic response (*MS4A2*, p_{adj}=1.30E-226; *SLC18A2*, p_{adj}<1E-300; *KIT*, p_{adj}<1E-300) and antigen-presenting cells – macrophages (*CYP27A1*, p_{adj}=4.34E-56; *MARCO*, p_{adj}=9.45E-70) and dendritic cells (*XCR1*, p_{adj}=0.03; *CD1C*, p_{adj}=6.51E-20; *IRF8*, p_{adj}=2.48E-24).

I was able to identify cells present in the circulatory system, namely arterial (*PODXL*) and venous (*CCL14*) endothelium (both p_{adj}<1E-300), as well as the neighbouring pericytes (*PDGFRB*, p_{adj}=4.52E-278, *CCL19*, p_{adj}=4.90E-173).

In epithelial compartment, I identified the stem cells, basal cells (*BCAM* and *DST*, both p_{adj}<1E-300; *CCN1*, p_{adj}=6.46E-188), along with closely related, suprabasal cells (*TP63* p_{adj}=1.66E-107, *NOTCH3* p_{adj}=1.86E-96). I identified secretory goblet cells (*MUC5AC*, *MUC5B*, *TSPAN8*, all p_{adj}<1E-300), deuterosomal cells (CDC20, p_{adj}=4.47E-52; *CCNO* p_{adj}=9.79E-92), and the ciliated cells of the epithelium (*FOXJ1* and *DNAH9*, both p_{adj}<1E-300). Interestingly, I observed submucosal glands in the biopsies, with the cells located in them displaying a unique transcriptional profile (*XBP1*, *FCGBP*, *MUC5B*, all p_{adj}<1E-300; *S100A1,* p_{adj}=1.17E-241).

I identified stromal cell types – fibroblasts (*PI16* p_{adj}=9.86E-271; *PDGFRA*, p_{adj}<1E-300; *LGR5*, p_{adj}=1.14E-43) and airway smooth muscle expressing myosin and actin encoding genes (*MYLK*, p_{adj}<1E-300; *ACTN1*, p_{adj}=1.69E-234) – as well as rare PNECs, which expressed neuropeptides such as gastrin-releasing peptide (*GRP*, p_{adj}=8.23E-202) and neuropeptide precursor secretogranin 2 (*SCG2*, p_{adj}=7.00E-115).

Overall, the populations of basal cells and ciliated cells were approximately equal in size (19,669 and 16,717, respectively, see Figure 4-3 A, right margin), consistent with the structure of airway epithelium described in chapter 1. Large numbers of stromal cells were captured, with 11,537 fibroblasts and 3,215 smooth muscle cells captured in total (Figure 4-3 A). This permitted detailed analysis of the non-immune populations in severe allergic eosinophilic asthma and expand the observations made in chapter 3.

Comparison of the manual annotations with automated mapping is particularly interesting, as the gene pool analysed is much smaller than that examined in single cell sequencing studies. HLCA annotation is shown in Figure 4-3 A with corresponding manual annotations given in Figure 4-3 C. Using automated approach, the plasma cell cluster is missed, potentially hindering the analysis of humoral response mechanisms in allergy and allergic asthma. Similarly, suprabasal cells and goblet cells are jointly annotated as 'goblet cells', removing the heterogeneity of these two distinct cell types, one being a stem cell and the other secreting mucus [88]. Similarly to suprabasal cells, the glandular cells were also missed from the annotation and instead annotated as 'goblet cells', which is correct to an extent, as submucosal glands have a role in secreting mucus, but further reduces dataset heterogeneity.

In summary, the HLCA-guided manual annotation seems to be optimal for annotation of spatial dataset, as the automated methods can miss key cell types, likely due to limited number of genes available, lack of cell types in the reference, or the necessity to use non-canonical markers of cell types, as *in situ* panels cannot contain genes with very high expression levels due to the risk of optical crowding [203] .

The cellular composition of the biopsies was similar across the three phenotypes studied – patients with HDM-sensitized allergic asthma, HDM allergy, and healthy controls – was similar, as depicted in Figure 4-4. Compared to results obtained from single-cell studies described in Chapter 5, the cell-type proportions remain consistent phenotype-to-

phenotype and sample-to-sample, suggesting more robust results can be obtained with *in situ* technologies, even from archival tissue of sub-optimal quality.

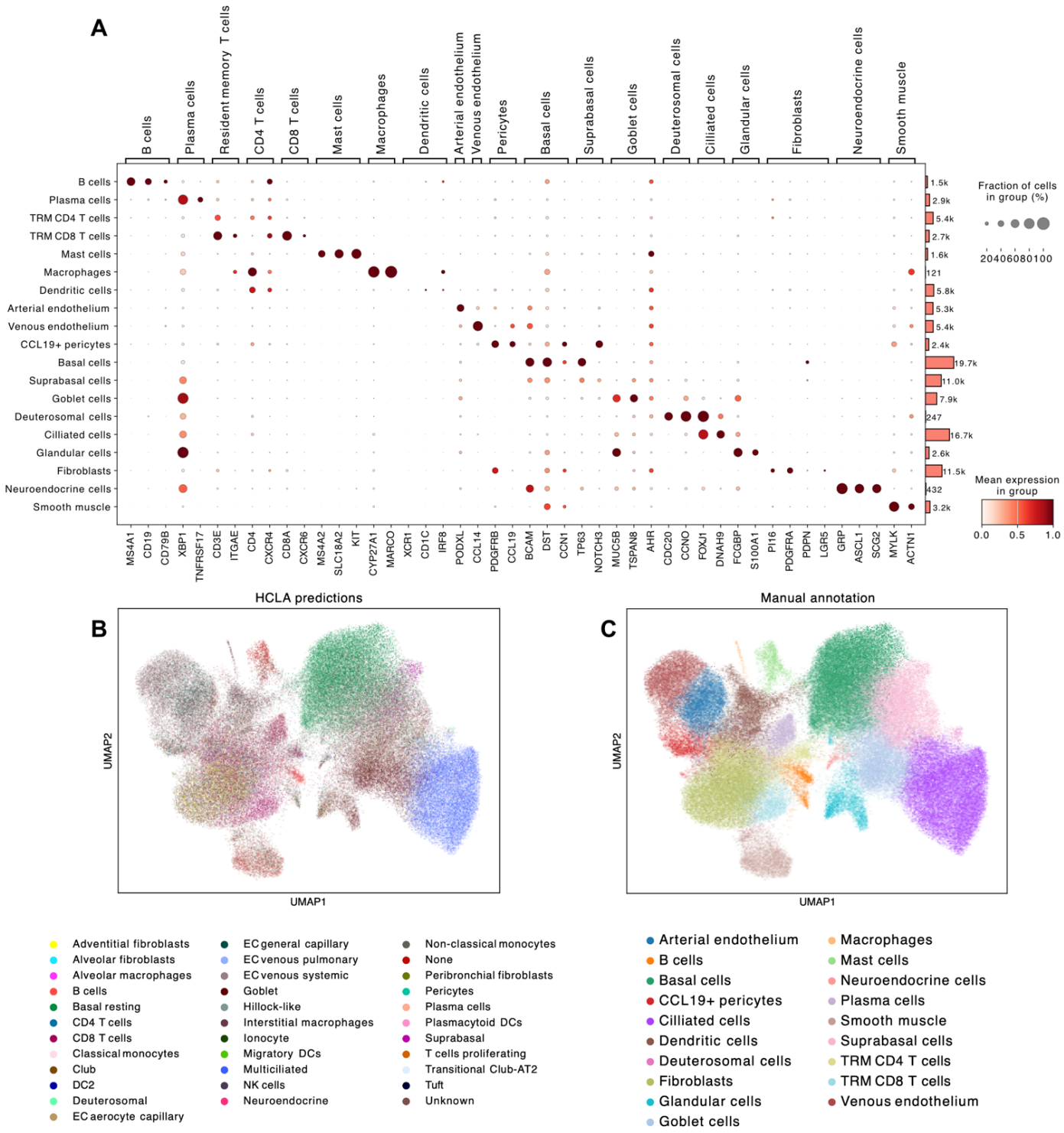

Figure 4-3 Annotation of cell types using manual and automated methods

A Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars. **B** UMAP embedding of spatial dataset coloured by the cell labels mapped from the HLCA at annotation level 4. **C** Refined manual annotation based on cell type markers and datasets in Chapters 3 and 5.

4.4.4 MAPPING OUT SPATIAL STRUCTURES IN LUNG BIOPSIES

To gain a better understanding of how the cell types are distributed in space in the lung of healthy and asthmatic patients, cell-type annotations were mapped onto images obtained during spatial transcriptomics experiment and the H&E staining generated post-run. As seen in Figure 4-5 A, the biopsies retained epithelial structure, with epithelial compartment lined with basal, suprabasal and ciliated cells attached to the basement membrane, and stromal compartment infiltrated by the innate and adaptive immune cells (B and T lymphocytes, macrophages).

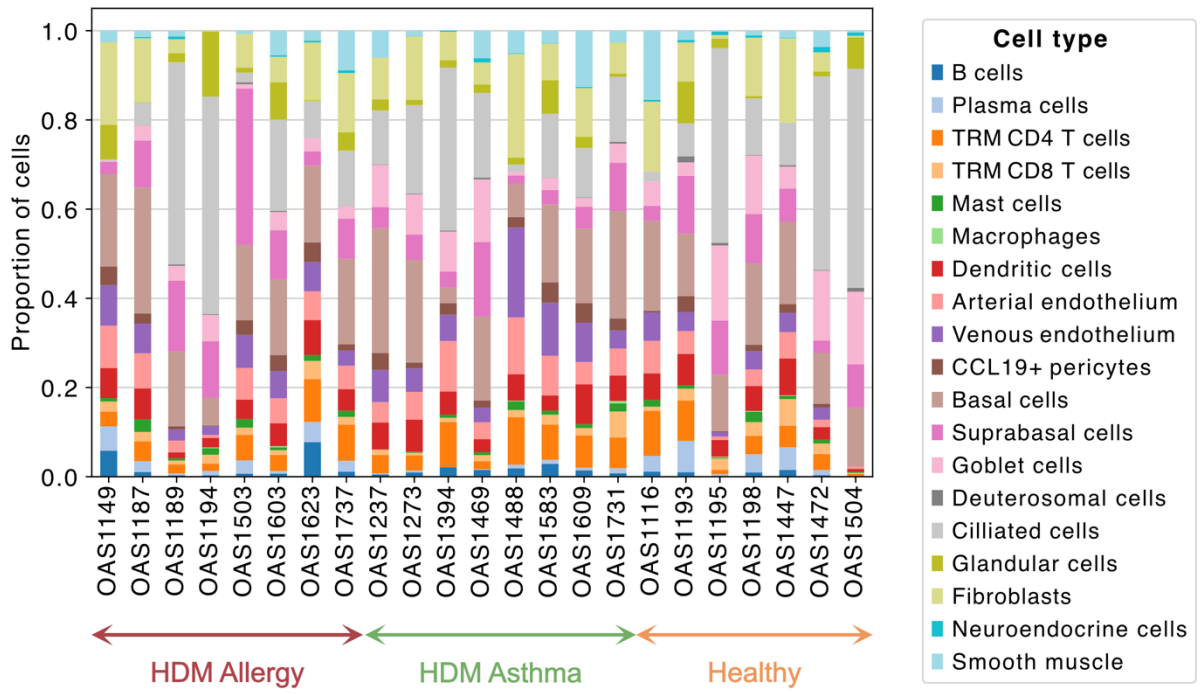

Figure 4-4 Cell type proportions in biopsy samples in patients with allergic asthma, allergy, and healthy controls

Cell type proportions normalised to 1 in sample obtained from each patient, sorted by phenotype (HDM-sensitized allergic asthma, HDM allergy, and healthy control). Cells are coloured according to shared developmental origin.

Of note, in some biopsies, glandular structures were identified, shown in Figure 4-5 B. Hinks group [204] and others [205,206] reported increased size and number of submucosal glands in patients with asthma. Increase in size of these mucous-secreting glands, taken together with the previously reported increase in infiltration by mast cells, neutrophils, and eosinophils[204,205], suggests that they may play an important role in asthma pathogenesis and airway remodelling, which should be further explored.

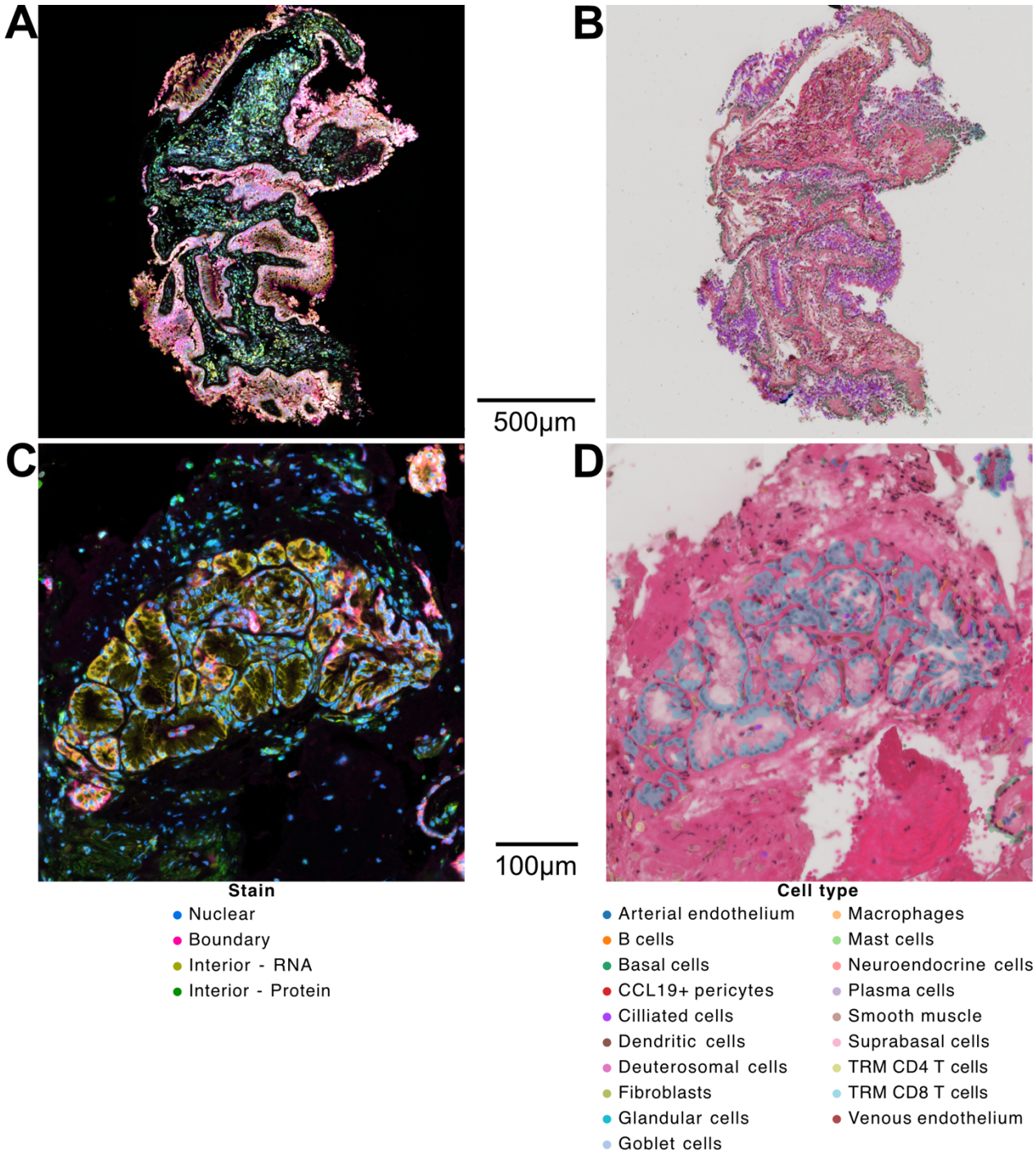

Figure 4-5 Anatomical structures present in lung biopsies and their cellular makeup

A Representative lung biopsy from the healthy individual coloured by the staining signal. **B** Corresponding H&E stain of the tissue with overlaid cell type annotation, showing epithelial and immune regions. Scale bar for **A** and **B** represents 500µm. **C** Submucosal gland present in a biopsy of allergic asthmatic patients with **D** cell type annotation overlaid. Scale bar represents 100µm.

4.4.5 EPITHELIAL DIFFERENTIATION PATHWAYS IN LUNG BIOPSIES

To confirm the cell type annotations as well as to further investigate the epithelial tissue structure identified in section 4.4.4, I conducted pseudotime analysis of epithelial cell types using *Palantir* [136], a tool treating differentiation as a probabilistic process, with a defined cell type of origin Mapping out spatial structures in lung biopsies– basal cells.

I subsetted the dataset to the cells of interest and recalculated PCs, HVGs, and the UMAP embedding, as seen in Figure 4-6 A. The cells included were basal and suprabasal cells, which form the stem cell reservoir of the airway, as well as differentiated states – ciliated, glandular, deuterosomal, glandular, and goblet cells, as well as PNECs, which all stem from the same progenitor, as described in chapter 1. The Palantir entropy, which is a measure of the cell plasticity [136], was observed as high in basal, suprabasal, and goblet cells, and decreased in the differentiated states – glandular, a subset of suprabasal, ciliated, and neuroendocrine cells. Identification of two termini of suprabasal cell differentiation may suggest heterogeneity in this cell subset, which could be caused by existence of self-renewing and differentiated suprabasal state within the airway [88]. Interestingly, entropy was still relatively high in goblet cells in the terminal state of differentiation, suggesting their plasticity, which perhaps may be due to goblet cell metaplasia observed in the lungs of asthmatic patients [37].

Examining the pseudotime trajectories (Figure 4-6 C) and comparing them with the known marker genes for epithelial cell types (Figure 4-6 D), confirmed the accuracy of my results,

with distinct differentiation branches identified. Neuroendocrine cell differentiation pathway also suggests that it is a separate entity branching off early along the differentiation route, possibly owing to distinct transcription regulators needed for development of this unique immunomodulatory and chemosensory cell type [89].

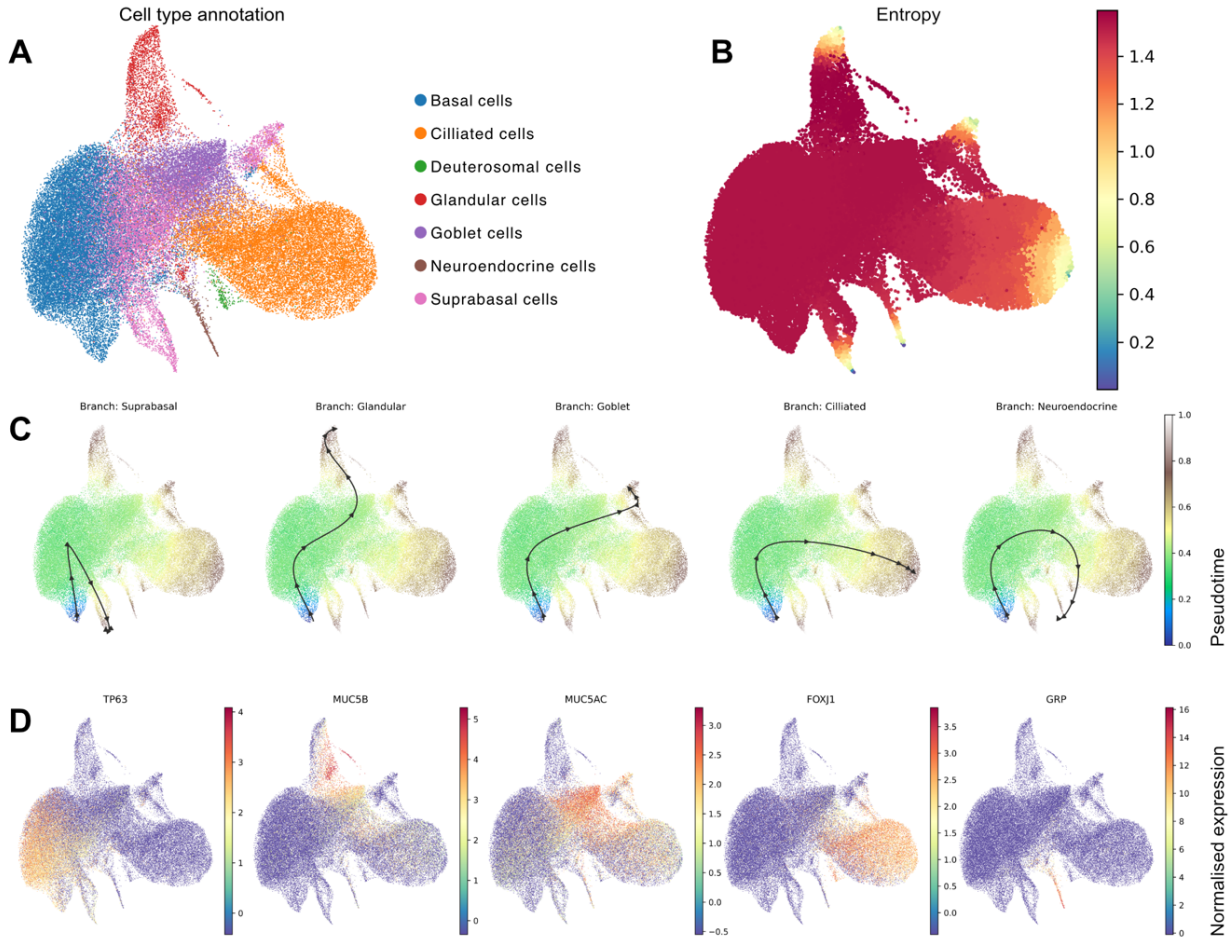

Figure 4-6 Pseudotime analysis of epithelial cell differentiation pathways
A UMAP embedding of the subset of spatial data consisting of epithelial cells. **B** Palantir state entropy based on the multiscale diffusion map. **C** Palantir pseudotime metric for each cell type (branch), for a differentiation pathway starting with basal cells. **D** Key marker genes for each cell type given in C.

4.4.6 DIFFERENTIAL GENE EXPRESSION IN EPITHELIUM IN ASTHMA VERSUS HEALTH

To build upon the analysis conducted in Chapter 3, where non-immune cells were excluded through CD45+ sorting to enhance statistical power given the small patient cohort, I compare the gene expression patterns in epithelial cells, as well as stromal and vascular cells, between allergic asthmatic patients and healthy controls. The results of this analysis are presented in Figure 4-7 and whilst providing insight, should be interpreted with caution due to small number of genes with non-zero expression levels tested, ranging from 385 in smooth muscle to 1336 in deuterosomal cells.

Different gene expression patterns are observed in closely related basal and suprabasal cells (Figure 4-7 A-B). In basal cells, a known marker of type-2 inflammation in asthma involved in allergen-induced eosinophil recruitment, periostin (encoded by the gene *POSTN*) [207]. CCN1 and CCN2, genes encoding cysteine-rich proteins which in mouse models have been shown to be upregulated in lungs of mice infected with non-typable Hemophilus influenza a common pathogen in asthma [208,209]. Whilst basal cells upregulated periostic, suprabasal cells showed increased expression of *NOS2*, a nitric oxide synthase responsible for synthesis of type-2 asthma biomarker, nitric oxide [54].Gene encoding 15 lipoxygenase 1 (15LO1, *ALOX15*) induced by type-2 cytokines was upregulated in suprabasal cells, but not basal cells, and was also found in epithelial samples described in chapter 5. Interestingly, despite the large number of transcripts detected in deuterocytes (Figure 4-7 C), no differences except for a single gene were observed at the level of statistical significance.

Conversely, ciliated and goblet cells (Figure 4-7 D-E) differentially expressed *SOX2-OT*, a transcriptional regulator reported to be involved in airway inflammation through in *in vitro* experiments [210,211].

In the smooth muscle cells, two genes upregulated in chronic intermittent hypoxia, *HIF1* and *THBS1*, were upregulated in patients with asthma (Figure 4-7 F). These genes form a HIF-1α/THBS1/BECN1 signalling pathway recently demonstrated to contribute to pathogenic EMT resulting in airway remodelling in asthma [212]. Fibroblasts (Figure 4-7 G) also upregulated *THBS1*, whilst downregulating periostin.

In the vascular compartment, endothelium (Figure 4-7 H-I) showed upregulated nitric oxide synthase 3, *NOS3*, a different source of nitric oxide also associated with FeNO readouts in asthmatic patients [213]. Interestingly, pericytes (Figure 4-7 G) showed the same *HIF1, THBS1* pattern of expression indicative of EMT events as the smooth muscle cells.

Finally, multiple cell types, such as basal, suprabasal, and goblet cells, fibroblasts, as well as pericytes and endothelium, upregulated *NR4A1*, a gene which does not have obvious links to airway inflammation but has been show to regulates TGFβ signalling and be involved in inflammatory processes and fibrosis [214]. Therefore, it could potentially be an attractive new biomarker in asthma.

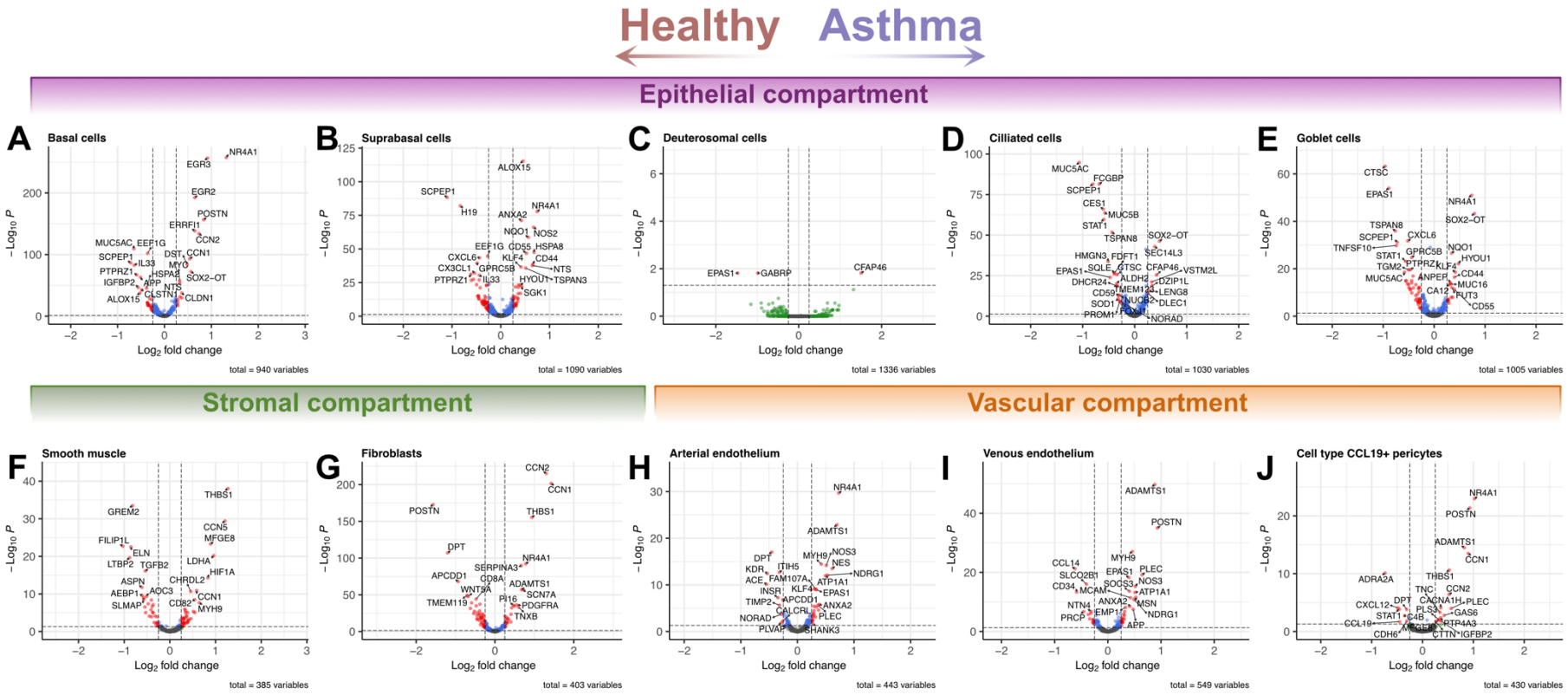

Figure 4-7 Differential gene expression in epithelial, stromal, and vascular compartments in allergic asthmatic patients compared to healthy controls
Volcano plots showing gene expression changes in **A-E** epithelial compartment, **F**-G stromal compartment, and **H**-J vascular compartment in spatial transcriptomics biopsy samples in patients with HDM-sensitized allergic asthma compared to healthy controls.

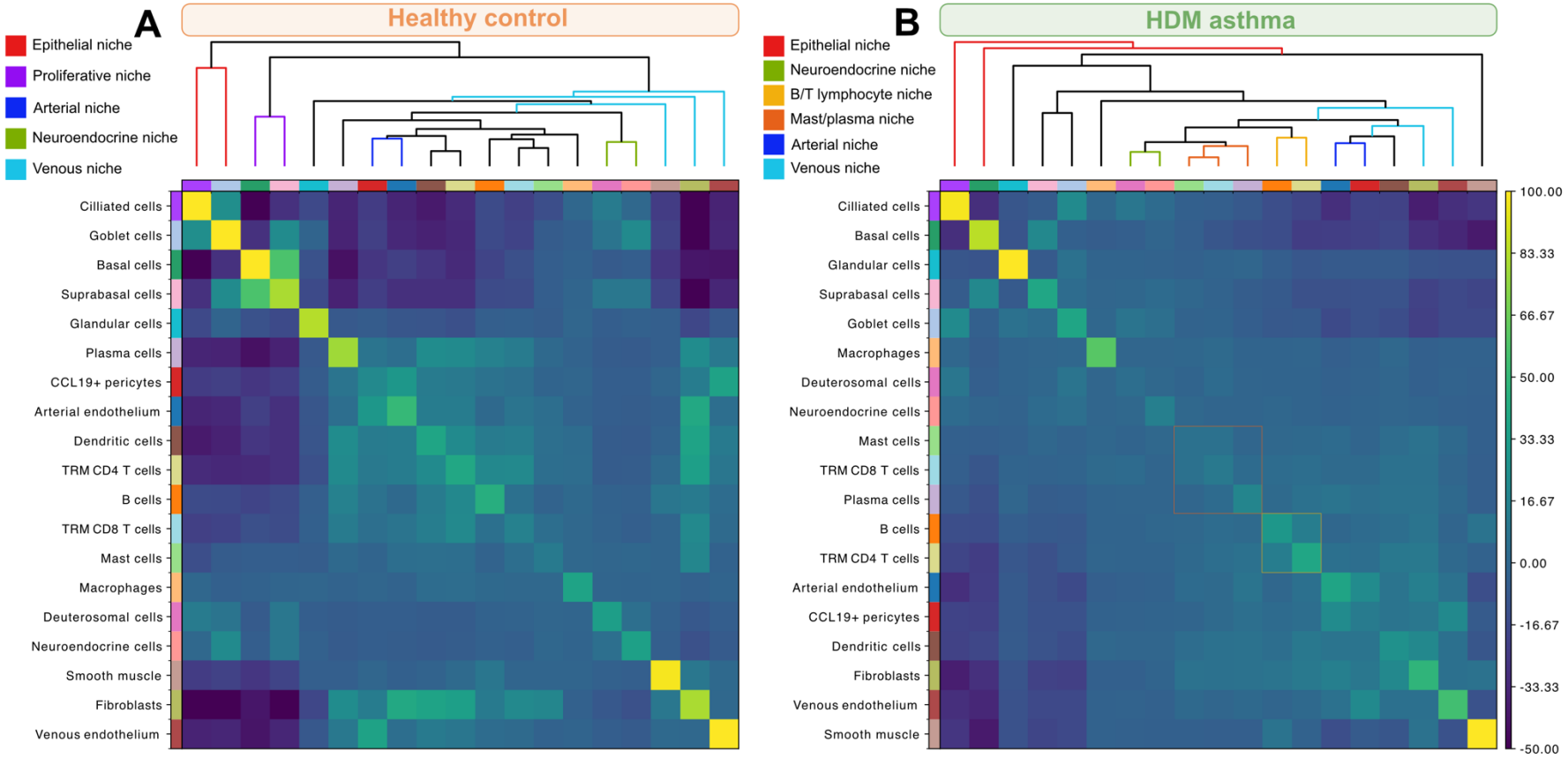

Figure 4-8 Hierarchical clustering of between cell type neighbourhood enrichment scores in biopsies of healthy individuals and allergic asthma patients
Heatmap of neighbourhood score based on neighbourhood graph calculated using Delaunay triangulation and given as an average z-score across 1000 permutations in **A** healthy controls and **B** asthmatic patients. Select niches are annotated based on the cell types. B/T lymphocyte niche and mast/plasma cell niche have been highlighted in asthmatic patients.

4.4.7 CHANGES OF IMMUNE MICROENVIRONMENTS IN ALLERGIC ASTHMA IN SPATIAL CONTEXT

Asthma is driven by distinct interactions between the epithelial barrier and the immune system. To understand how immune cells, profiled in detail in the previous chapter (Chapter 3), and the epithelial cells interact, I conducted neighbourhood enrichment analysis on the spatial data, with resulting interaction matrices shown in Figure 4-8.

In healthy individuals, I could identify an epithelial niche, consisting of co-occurring goblet and ciliated cells, stem-cell populated proliferative niche consisting of basal and suprabasal cells, venous and arterial niches characterised by the presence of stromal cells and vasculature, as well as a neuroendocrine niche, where neuroendocrine cells were interestingly co-localised with deuterosomal cells. These niches are seen in Figure 4-8 A.

In patients with asthma, broadly similar niches could be observed (Figure 4-8 B), except for proliferative niche. Instead, basal cells were present in the neighbourhood of ciliated cells. In addition, two other immune niches could be seen in allergic asthma. One niche was observed to consist of plasma cells, CD8+ T cells, and mast cells, which could be speculated to be in proximity due to IgE secretion and binding to mast cells, and another, consisting of CD4+ T cells and B cells, which could be interacting as a result of antigen presentation by CD4+ T cells in the context of allergic inflammation.

In summary, disease-specific immune niches seem to be present in allergic asthma, consistent with known mechanisms of inflammation. It should be added that care must

be taken when interpreting this data, as the conclusions may be subjective and not statistically robust. New computational methods are needed to assess statistical significance of neighbourhood enrichment findings.

4.5 DISCUSSION

In summary, I am describing the first of its kind dataset to profile allergic asthma and allergic inflammation in lung tissue at spatial resolution. I was able to profile over 100,000 cells across 23 patients. Detailed annotation shows that whilst limited cell type distinction is possible (19 cell types), there cell type bias observed in single cell studies seems to be less potent in *in situ* spatial transcriptomic studies.

My initial analysis shows that the lung tissue sections captured recapitulate known anatomy, represented by separate epithelial and stromal-immune layers, as well as biology, represented by recreation of established differentiation pathways of epithelial cells, starting from basal cells and diverging into multiple cell types. What is more, I identify submucosal glands, both through the anatomy and unsupervised clustering, in the patients with allergic asthma.

The differential gene expression analysis reveals new and known patterns of inflammation in allergic asthma, with upregulation of nitric oxide synthase 2 and 3 in the airway samples, signs of tissue remodelling, and *NR4A1* being a potential novel biomarker candidate in asthma.

Finally, through neighbourhood analysis, I recapitulate immune niches consistent with known pathophysiology of allergic asthma, demonstrating CD4+ T-B lymphocyte interaction and a niche enriched in plasma and mast cells.

The study has its limitations. It was not possible to locate club cells in the dataset, suggesting they may be present in another cell cluster, most likely goblet cells. This can likely be mitigated by using the subclustering approach I demonstrate to be effective in cell type annotation in other tissues. I could also not locate eosinophils, which are expected to be infiltrating the lung, with spatial transcriptomics offering an attractive avenue to study them, as they are known to be lost in microfluidics-based single cell approaches. Given that they can be identified using microscopy approaches, I hypothesise the harsh conditions of FFPE preservation may lead to the degradation of genetic material. Likely due to similar reasons, the other granulocyte subset (neutrophils) was not identified in the data. In addition, no submucosal glands could be located in samples from healthy individuals, limiting the possibility of comparisons between glandular cells in asthma and health.

Beyond caveats in cell type detection, the detection of cellular neighbours can be affected by technical artifacts. As physical distance is being analysed, placement of the tissues on the slide, cutting angle chosen during sectioning, as well as tissue folding during sectioning or dethatching during imaging can affect the metrics. New analysis tools are being developed to circumvent these problems [215].

In summary, I provide a resource to study allergic asthma at single cell resolution and demonstrate its utility by recapitulating known biological processes in the lung using a collection of tools. This is the first study of its type, applying large gene panel consisting of 5,001 targets to over 20 lung biopsies of carefully phenotyped asthma patients, allergic, and healthy controls. With it, I set out new hypotheses about type-2 inflammation and fibrosis which I hope can be tested in more depth in future projects.

5 CHARACTERISATION OF IMMUNE CELLS AND EPITHELIAL CELLS ACROSS ASTHMA PHENOTYPES

5.1 INTRODUCTION

Asthma is a heterogenous disease of the airway, with varied molecular mechanisms and cell types involved, unified by its defining clinical features – bronchoconstriction, airway hyperresponsiveness, shortness of breath – which stem from persistent inflammation [1,11]. In Chapters 3 and 4, I extensively profiled severe allergic eosinophilic asthma with HDM sensitisation, a single, well-defined phenotype. In this Chapter, I detail the single-cell profiling and comparison of 4 major asthma phenotypes across all cell compartments present in the airway –immune, epithelial and stromal.

The main phenotypes, which are approximately equal in prevalence, are type 2-high (eosinophilic) and type 2-low (non-eosinophilic), and are driven by different cell types. These show varied treatment response, with most treatments targeting severe eosinophilic asthma [4,7]. In type 2-high asthma, characterised by eosinophilia in peripheral blood and airways, as well as elevated levels of type 2 cytokines (IL-4, IL-5, and IL-13), at least two distinct mechanisms exist which drive the inflammation [3,19]. Major known asthma phenotypes, along with their subdivisions, are outlined in Figure 5-1.

In allergic asthma, aeroallergens (often with enzymatic activity) such as pet dander, spores, pollen, or HDM proteins, present in the airway stimulate the release of alarmins (IL-1, IL-25, IL-33, TSLP, and GM-CSF) from the airway epithelium [216]. These alarmins activate ILC2s and Th2 lymphocytes, which subsequently secrete type 2 cytokines [216,217]. By contrast, in the much less common, but more severe phenotype of adult-onset eosinophilic asthma, the primary pathology is eosinophilic airways inflammation, driven by IL-5, though the source of this IL-5 and the aetiopathogenic mechanism for its overproduction remain unknown [218]. Regardless of the cause of the eosinophilia, type

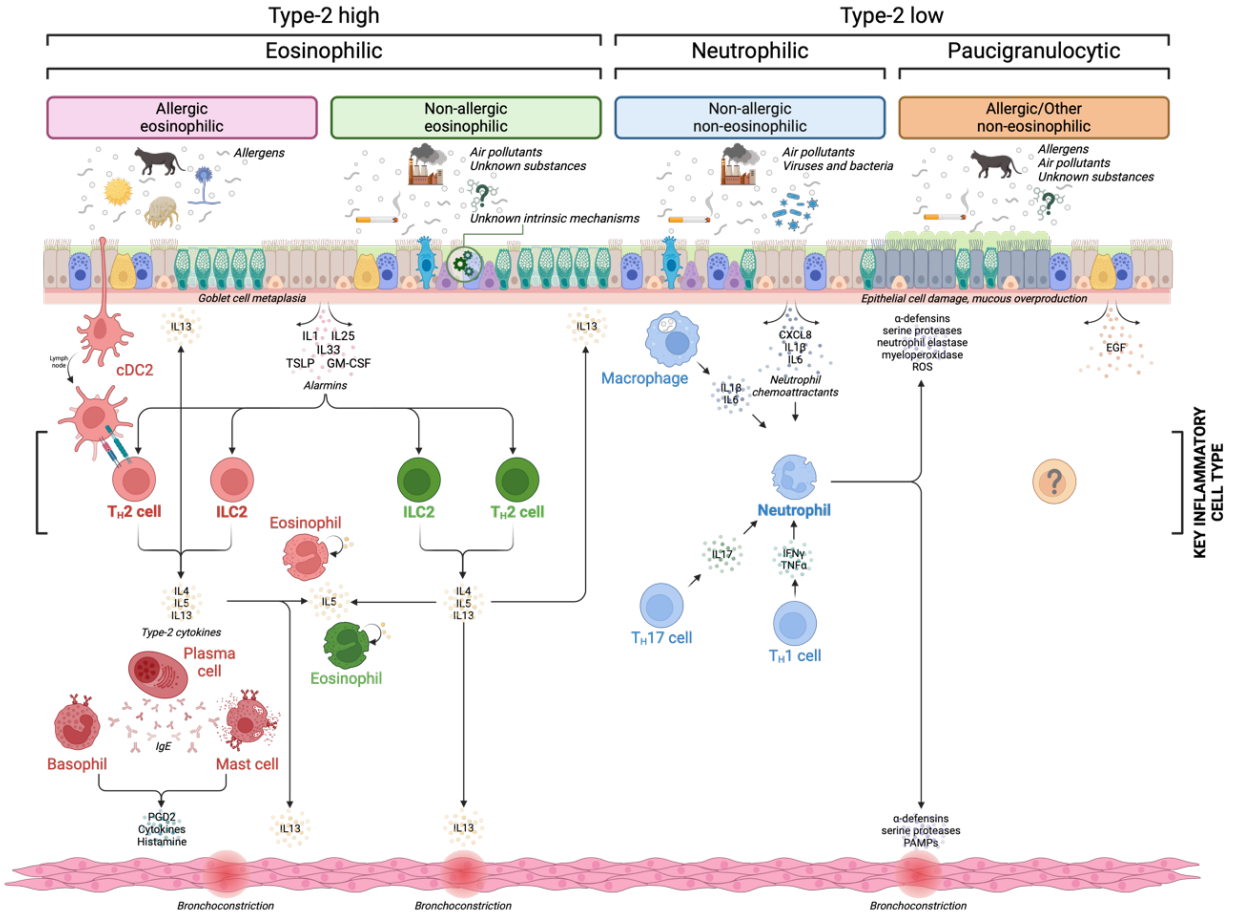

Figure 5-1 Summary of the main inflammatory cells and pathways across known asthma phenotypes
Asthma phenotypes are heterogeneous and can be broadly divided into type-2 high and type-2 low asthma. In type-2 high (eosinophilic) asthma, airway inflammation can result from exposure to allergens or other unknown environmental factors, with mechanisms of inflammation converging at the Th2 lymphocytes and ILC2s. Type-2 low asthma includes neutrophilic asthma driven by neutrophilic inflammation, as well as other less well characterised subtypes, such as obesity-related and allergic non-eosinophilic asthma.

Based on figures by Oppenheimer *et al* (2022) and Godar *et al* (2017), adapted and updated.

2-high airway inflammation leads to a cascade of pathological changes, including bronchoconstriction, airway mucus hypersecretion, increased mucus viscosity, mucus plugging, airway inflammation with oedema, and potential airway remodelling [219]. Airway remodelling is exacerbated by eosinophil-derived cytokines, such as TGFβ and leukotriene C4 [220].

Whilst Th2 lymphocytes are mainly involved in the development of eosinophilic asthma, other Th cells, such as Th17 cells characterised by expression IL-17 family cytokines (e.g. IL-17A and IL-17F), can induce airway neutrophilic inflammation, often present in most severe cases of asthma phenotypes [221]. Levels of IL-17A and IL-17F expression correlate with asthma severity in patients with neutrophilic disease resistant to steroid treatment [222].

Environmental factors such as cigarette smoke, pollutant particles, and allergens can trigger Th17-mediated airway inflammation in patients with asthma, with cigarette smoking correlating with bronchial neutrophilia, asthma severity, and corticosteroid insensitivity [223,224]. In addition to inhaled particles, microbial colonisation by fungi, bacteria, and viruses have been implicated in pathogenesis of asthma [225]. Microbiota can trigger NLRP3 inflammasome assembly in the epithelial and immune cells and subsequent IL-1β release, which promotes Th17 differentiation [224].

Other non-eosinophilic asthma subtypes likely exist (as summarised in Chapter 1), and display mild-to-moderate phenotype, such as obesity-related asthma thought to be mediated by IL-6 and TNF [18]

Previous efforts to characterise airway immunology at the single-cell level, such as the Human Lung Cell Atlas [78], which consists of 584,444 cells from 107 patients in the core atlas and 2.4 million cells from 486 patients in the extended version, did not include patients with asthma. Similarly, the “lung census” project [118] analysed biopsies from a limited cohort of six patients with broadly defined mild-to-moderate, childhood-onset asthma, generating data from only 15,033 cells. In contrast, my study represents the largest asthmatic lung single-cell dataset to date, consisting of 219,222 high-quality cells across 39 patients with severe asthma, with rigorous phenotypic characterisation to enable precise insights into the airway immunology of each asthma phenotype.

5.2 HYPOTHESIS

Distinct asthma phenotypes exhibit unique molecular mechanisms of inflammation, which vary based on the anatomical region of the airway (upper or lower). Inflammation in type 2-high asthma is driven by alarmins, whereas type 2-low asthma has a different molecular mechanism. Inflammation induces alterations in both immune cell populations within the respiratory system and the airway epithelium. These alterations manifest as changes in gene expression profiles and/or relative cell type proportions. While some changes are shared across multiple cell types, others are specific to individual cell types. Application of single-cell sequencing to airway tissues in well characterised individuals with specific phenotypes will provide novel mechanistic insight into the mechanisms of eosinophilic asthma and of steroid-resistant airways inflammation.

5.3 AIMS OF THIS CHAPTER

- To examine the epithelial and immune cell types present in the two sampling locations (nasal brush, endobronchial biopsy) in patients with asthma and healthy controls and compare the results with published datasets.
- To investigate the changes in cell type abundance in asthmatic patients, healthy controls, and across asthma phenotypes
- To profile the differences in gene expression in asthmatic patients, healthy controls, and across asthma phenotypes
- To identify candidate signalling pathways and gene markers specific to asthma phenotypes and link molecular mechanisms to phenotypic outcomes
- To examine the role of epithelial alarmins across asthma phenotypes
- To examine the role of immune type 2 cytokines across asthma phenotypes

5.4 RESULTS

5.4.1 COHORT CHARACTERISTICS

To compare the varied molecular and clinical manifestations of asthma, I undertook a clinical study to profile, compare, and contrast patients with severe asthma versus healthy controls, as well as patients with different types of asthma. With the help of Hinks group clinical team, we defined 4 broad study categories (allergic eosinophilic asthma, allergic non-eosinophilic asthma, non-allergic eosinophilic asthma, non-allergic non-

eosinophilic asthma), which were further divided or grouped together (type 2-high (eosinophilic) and type 2-low (non-eosinophilic) for more in-depth comparisons. Figure 5-3 details the clinical phenotypes defined and the corresponding immune classifications derived from the characterisation of the sputum samples.

Patient characteristics are summarised in Table 5-1. In total, 39 patient samples were collected, with non-allergic and non-eosinophilic patients slightly older than the control group, however most of the study was well matched. Caucasian individuals formed the majority of the patient groups, and a small proportion had a history of tobacco smoking, however, none were current smokers. The FeNO measurements were consistent with the correlation of this biomarker with type 2 (eosinophilic) inflammation, with mean FeNO 3-5-fold higher in allergic eosinophilic and non-allergic eosinophilic groups compared to non-eosinophilic groups and healthy patients.

Similarly to study described in chapter 3, patients underwent fibreoptic bronchoscopy and up to 10 biopsies were collected along with brushings of the a brushing of their nasal epithelium (Figure 5-3 A). The biopsy was enzymatically dissociated, and both the biopsy and the nasal brush were mechanically dissociated using a syringe and needle. A mix of flow cytometry and sequencing antibodies allowed me to sort live cells and correct for the low abundance of immune cell types, by sorting the samples in such way that the immune and non-immune cells (CD45+ and CD45-, respectively) were approximately equal in cell numbers (Figure 5-3 A). Sequencing libraries were constructed for the transcriptome, BCR and TCR sequences, and the panel of antibodies (Figure 5-3 A).

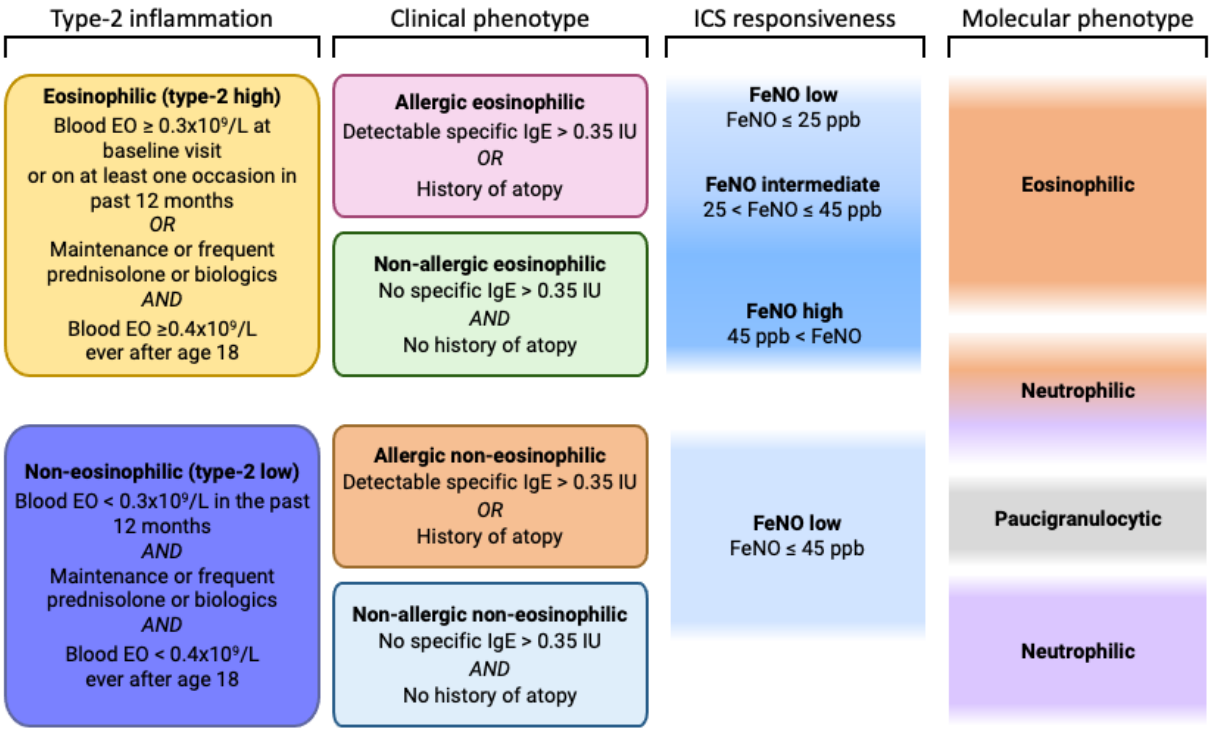

Figure 5-2 Clinical phenotypes in the MORSE study of severe asthma

Patients were stratified by the clinical team into four phenotypes— allergic eosinophilic asthma, allergic non-eosinophilic asthma, non-allergic eosinophilic asthma, non-allergic non-eosinophilic asthma —based on the biochemical markers and clinical history detailed in the figure. These clinical phenotypes correspond to the molecular phenotypes depicted on the right. ICS = inhaled corticosteroids

Original figure created by Timothy Hinks, Maisha Jabeen, and Anastasia Fries; modified for

Table 5-1 Cohort characteristics for MORSE project

	Healthy control (n=10)	Allergic eosinophilic (n=13)	Allergic non-eosinophilic (n=6)	Non-allergic eosinophilic (n=7)	Non-allergic non-eosinophilic (n=3)
Age, years (mean ± SD)	36.44 ± 12.49	44.08 ± 16.40	52.83 ± 18.95	63.14 ± 17.22	73.00 ± 3.46
Male/female	3/7	7/6	3/3	5/2	2/1
Ethnicity (Caucasian, Asian)	9, 1	12, 1	6, 0	7, 0	3, 0
BMI, kg/m^2 (mean ± SD)	25.90 ± 6.33	28.08 ± 6.34	28.75 ± 9.00	24.86 ± 2.61	31.00 ± 7.94
Eosinophil count, counts x 10^9/ml (mean ± SD)	1.05 ± 0.53	5.28 ± 0.27	1 ± 0.08	2.96 ± 0.30	0.35 ± 0.11
Fractional exhaled nitric oxide [FeNO] (ppm)	18.86 ± 6.67	74.70 ± 54.50	21.00 ± 11.39	96.14 ± 39.49	20.33 ± 9.50
Smoker (%)	10%	23%	0%	14%	33%
Current smoker (%)	0%	0%	0%	0%	0%
Ex-smoker (%)	30%	0%	17%	14%	33%
Smoking status unknown (%)	0%	8%	17%	14%	33%

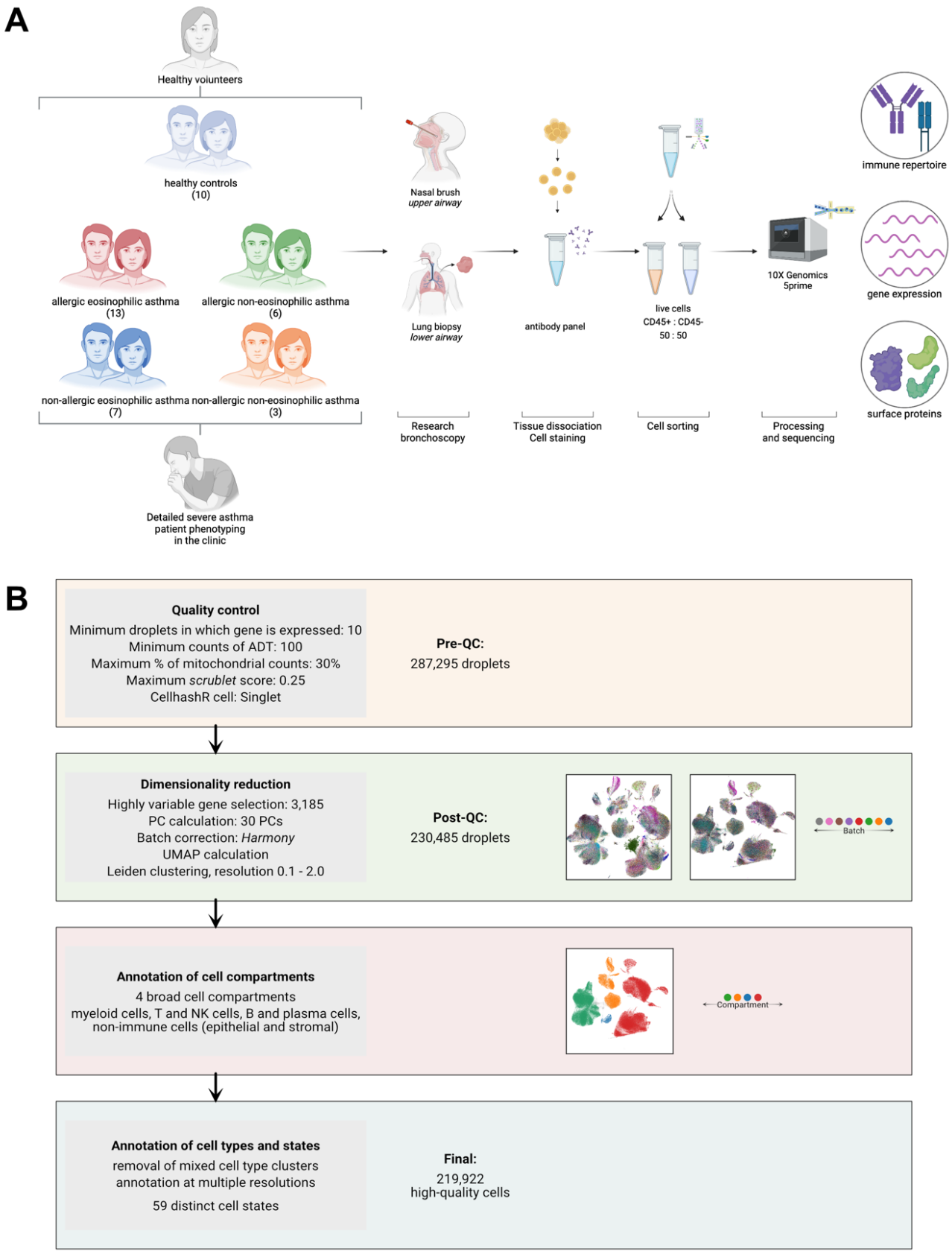

Figure 5-3 MORSE study design and data processing overview

A Overview of the patient phenotypes (allergic eosinophilic, non-allergic eosinophilic, allergic non-eosinophilic, non-allergic non-eosinophilic, healthy control) and the experimental processing, consisting of sample collection from upper (nasal brush) and lower (biopsy) airway, physical (nasal brush) and enzymatic (biopsy) dissociation, antibody staining and flow cytometry sort to equalise immune (CD45+) and non-immune (CD45-) populations, and single cell genomics library preparation collecting four modalities (gene expression, B and T cell receptor sequence, antibody staining). **B** Summary of key preprocessing steps with associated cell and droplet counts. The key steps are QC filtering, dimensionality reduction with batch correction (batch = patient ID), annotation of broad compartments, and detailed cell type annotation.

5.4.2 DATA PREPROCESSING AND METHOD DEVELOPMENT

Data was processed with *panpipes* to increase reproducibility and streamline the preprocessing workflow. 287,295 raw droplets were obtained (Figure 5-3B). Only genes expressed in more than 10 droplets were considered, leaving 31,913 genes. Droplets with less than 100 protein counts, indicating poor antibody staining, were excluded from the analysis in the protein modality, but the RNA modality was retained (Figure 5-3B). Droplets with more than 30% mitochondrial gene counts, indicative of dead or dying cells as described in chapter 3, were excluded (Figure 5-3B). Similar doublet exclusion approach was taken as described in the chapter 3, with heterotypic doublets identified using hashing oligos, and homotypic doublets identified using *scrublet*, an algorithm comparing droplets to simulated doublets. A consensus of multiple algorithms was used to call the tissues (biopsy, nasal brush) versus the doublets (Figure 5-4 A). Taken together, these approaches removed further 56,810 poor-quality droplets, leaving 230,485 cells in the dataset (Figure 5-3B). I selected highly variable genes using *scanpy* implementation of *Seurat* highly variable gene calling function, which automatically selects the optimal number of HVGs to capture the dataset variance, determined to be 3,185 genes. These genes were fed to the PC algorithm and 30 PCs were used to calculate the UMAP embedding and batch correction (Figure 5-3 B and Figure 5-4 B-C). Patient ID was used as the batch to aggregate technical variables, which can be approximated to the random effects introduced by the clinician performing the bronchoscopy, the operator processing the tissue, the day of the experiment, the sequencing batch and instrument, amongst others. By assessing the UMAP embeddings visually and comparing the density

of LISI scores (Figure 5-4 B-C), I decided *Harmony* performed best at the batch integration task.

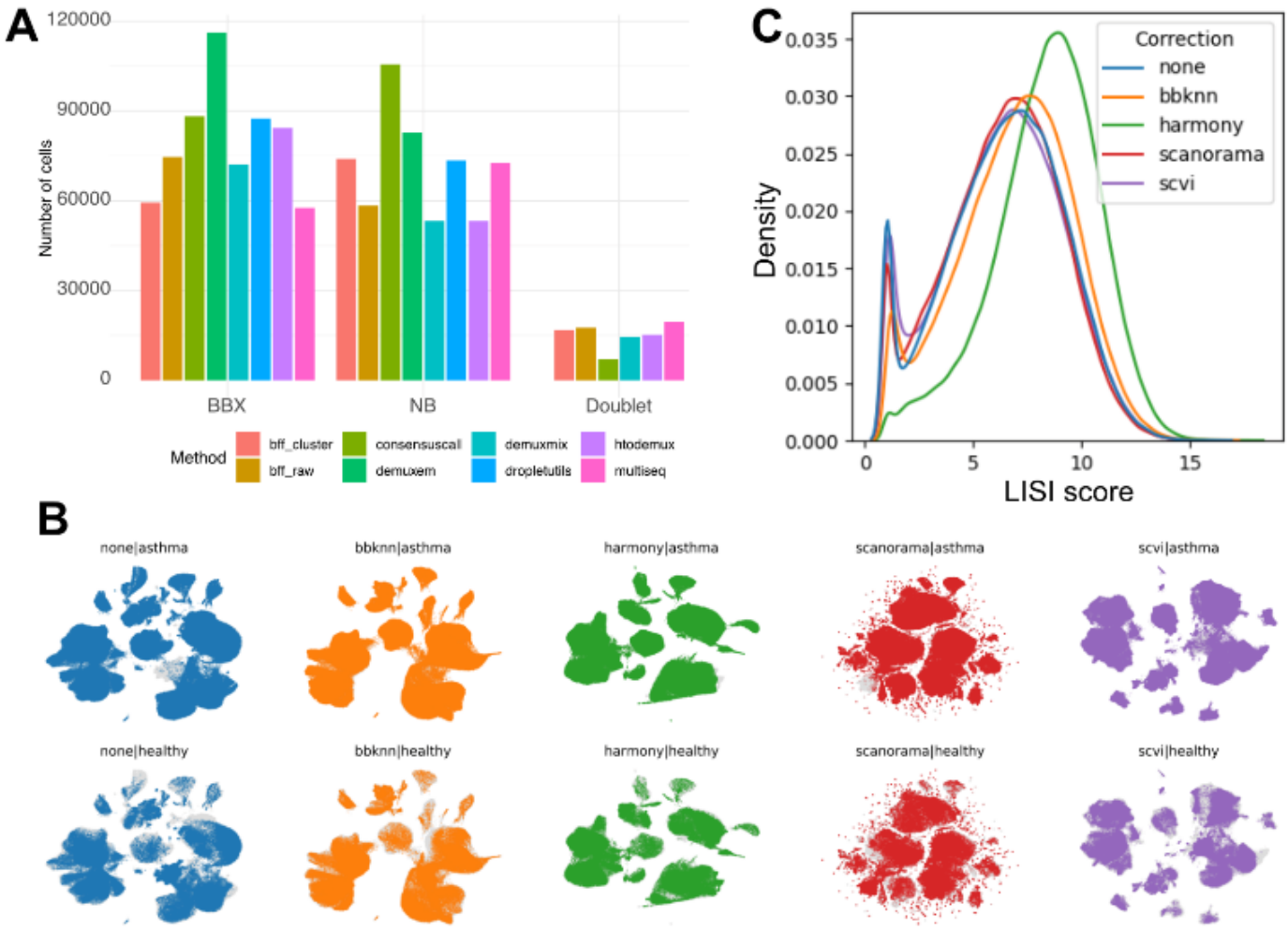

Figure 5-4 Summary of optimisation of de-hashing and batch correction algorithms
A Summary of the cell calls of 7 different de-hashing algorithms and 100% consensus majority vote method ('consensuscall'). BBX=biopsy, NB=nasal brush **B** Comparison of four data integration methods on the UMAP embedding coloured by patient diagnosis (healthy control or any of the asthma phenotypes). **C** Local Inverse Simpson's Index (LISI) scores calculated for each integration method.

5.4.3 CELL ANNOTATION

The *Harmony* clusters were separated into immune (protein CD45+) and non-immune (protein CD45-, *EPCAM*+) cell subsets, with the immune subsets further divided according to haematopoietic origin to myeloid, NK and T, and B and plasma cell compartments, using cluster markers, BCR/TCR sequences, and antibody staining for cell compartment assignment (Figure 5-3 B and Figure 5-5 A-B). The HVGs and embedding was re-calculated for each subset and Leiden clusters were calculated at multiple resolutions (0.1-2.0, cell numbers permitting). I annotated the cells by amalgamating HLCA annotations (Figure 5-6), literature reported cell type markers, and expert input, which resulted in a total of 59 distinct cell states. Clusters with mixed cell types, representing further doublets or ambient RNA contamination, were removed, as is the best practice [154], resulting in the final number of 219,922 high-quality cells (Figure 5-3 B and Figure 5-5). Of those, 2,531 were in B and plasma compartment (*CD79B*, *CD19*, Figure 5-5 C), 79,029 were in the NK and T cell compartment (*CD3E*, *CD8A*, *CD4* Figure 5-5 D), 29,940 in myeloid compartment (*S100A8*, *S100A9*, *FCER1G*, *CLEC7A*, *CP3A*, Figure 5-5 E), 108,422 in the epithelial compartment (*EPCAM*, Figure 5-5 F). In summary, 111,500 immune and 108,422 non-immune cells were captured, demonstrating the efficacy of the sorting strategy (50.70% vs 49.3% ratio). On average, approximately 5,640 cells were captured per patient, compared to 30,000 – 50,000 loaded cells (11 – 19% efficacy).

Comparing the final annotations with those generated by reference mapping to the HLCA (Figure 5-6), I observed broad agreement, however, I was able to capture substantial

granularity in the T cell compartment (MAIT cells, gamma-delta T cells, and multiple CD8+ T cell states), B cell compartment (naïve, activated, and memory B cells and plasma cells), and myeloid compartment (multiple monocyte and airway macrophage populations, pDCs, cDC2s and LAMP3+ DCs). Notably, in my dataset I was able to identify a rare LAMP3+ dendritic cell population, but not a DC1 population. In addition, I did not observe capillary or lymphatic epithelial cells, submucosal glands, or tuft cells, which were present in the HLCA mapping. In summary, whilst the HLCA was a useful tool to guide cell type annotation, further manual curation was necessary to remove noise and enhance granularity of the annotations, likely due to the absence of asthma-specific data in the HLCA.

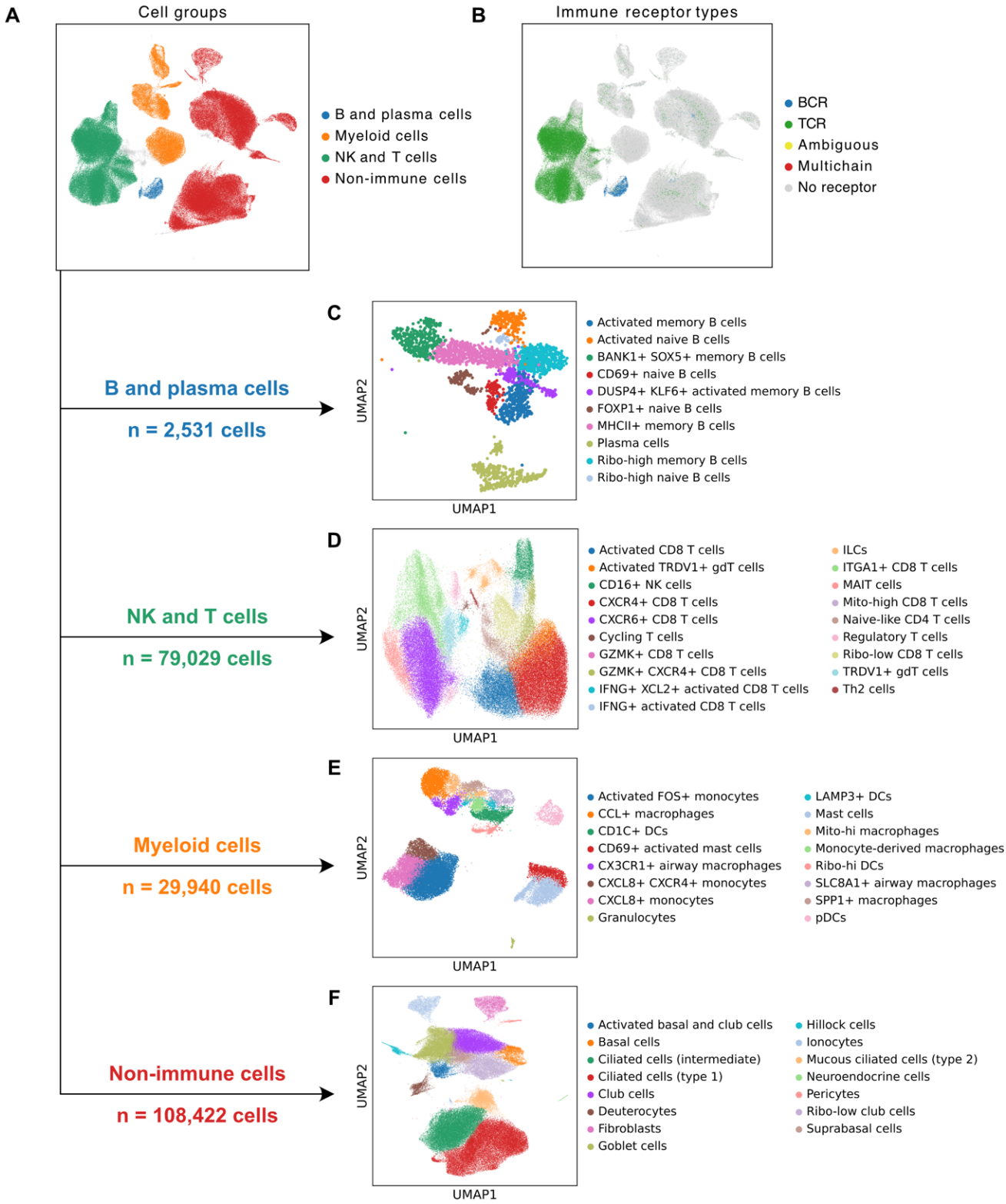

Figure 5-5 Details of annotation of cell states across compartments

A Location of the cell compartments (B and plasma, NK and T, myeloid, non-immune) cells annotated according to their developmental origin. **B** Mapping of the immune cell receptors to confirm the location of B, T, and plasma cells on the UMAP embedding. **C-F** Embeddings of individual cell compartments and corresponding cell state annotations.

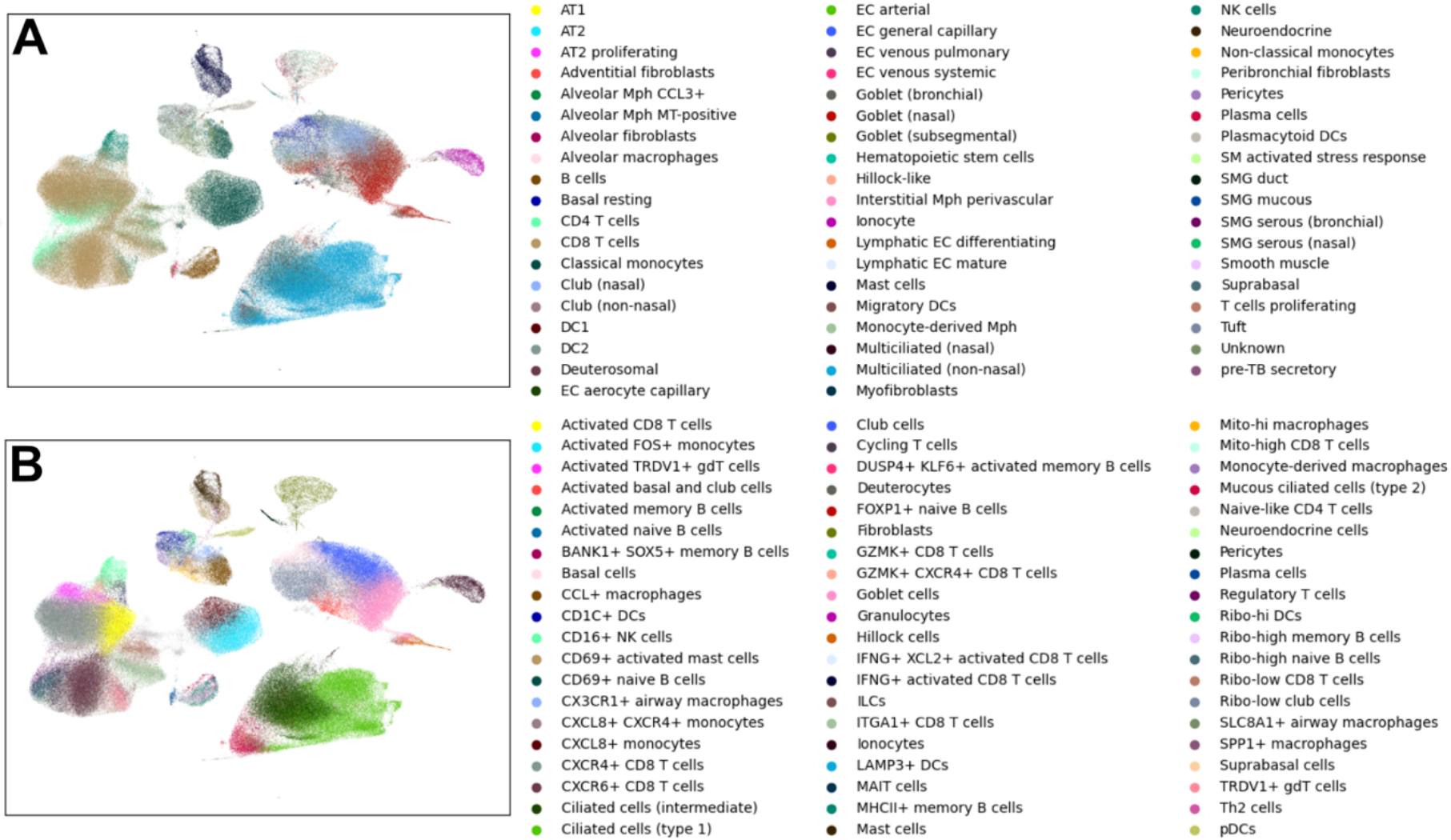

Figure 5-6 Comparison of HLCA and manual cell type annotations

UMAP embedding coloured by the finest level of annotation available from **A** HLCA mapping performed using *scvi* at resolution 'annotation_finest_level' and **B** manual annotation using sub-grouping approach at cell-state (highest) resolution.

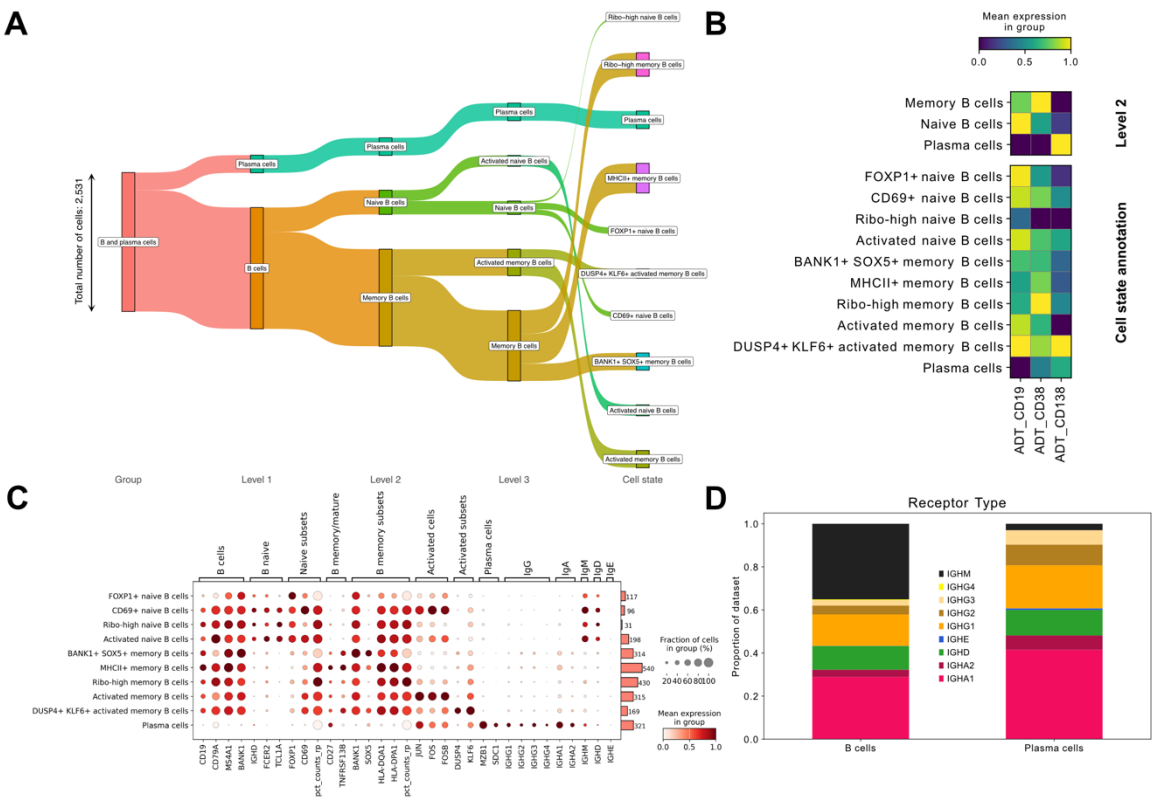

Figure 5-7 Characterisation of B and plasma cell compartment

A Sankey diagram of B and plasma cell populations and states across increasingly detailed annotation levels. **B** Normalised surface protein staining signal across annotation level 2 and cell state annotation. **C** Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars. **D** Normalised abundance of B cell receptor types in B and plasma cells.

5.4.4 CELL STATES IN THE HUMORAL IMMUNITY COMPARTMENT

In the B and plasma compartment, I was able to identify all states in the developmental pathway (Figure 5-7 A): naïve B cells (*IGHD*, p_{adj}=2.72E-9 / 7.72E-4; *FCER2*, p_{adj}=1.15E-13 /2.45E-4; *TCL1A*, p_{adj}=9.91E-14/n.s. in the *CD69*+ / *FOXP1*+ subset, respectively), memory B cells (*CD27*, p_{adj}=1.04E-7/8.41E-43; *TNFRSF13B*, p_{adj}=4.96E-43 / 5.92E-102 in the *BANK1*+ *SOX1*+ / MHCII+ subset, respectively), and plasma cells (*MZB1*, p_{adj}= 1.00E-123; *SDC1*, p_{adj}=1.75E-8). Whilst naïve and memory cells expressed CD19 and CD38 surface proteins, the plasma cells exclusively bound anti-CD138 antibody (Figure 5-7 B). Of note, both memory B cells and plasma cells expressed high levels of IgA (*IGHA1* and *IGHA2*, p_{adj}<0.05 for activated memory, *DUSP4*+ *KLF6*+ activated memory, *BANK1*+ *SOX5*+ memory, MHCII+ memory B cells, and plasma cells, see Figure 5-7 C-D), with previous reports of IgA+ memory B cells having increased abundance in asthma and diseases resulting in small airway dysfunction [226]. Overall, the clonotype switching typical of mucosal surfaces was seen, with an increase in IgA+ and decrease in IgD/IgM between B and plasma cells (Figure 5-7 D). Only a very small proportion of B and plasma cells were the IgE-expressing cells, with only a single cell each called as *IGHE*+ in the immune repertoire library (Figure 5-7 D). Activated B cells in both naïve and memory compartment had upregulated *JUN* and *FOS* transcription factors, likely resulting from a combination of biological activation [227] and cellular stress due to handling and dissociation, particularly enzymatic dissociation in physiological (37 °C) temperatures [119]. Compared to chapter 3, the B and plasma cell numbers captured were similar, even though the number of patients was increased 5-fold (39 versus 8), showing the benefit of sorting to increase the abundance of immune cell populations in small studies.

5.4.5 CELL STATES IN THE MYELOID COMPARTMENT

In the myeloid compartment, I observed monocytes, macrophages, dendritic cells, as well as mast cells and granulocytes (Figure 5-8 A). I confidently identified conventional type 2 dendritic cells involved in epithelial antigen presentation (cDC2s or CD1C+ DCs, *CD1C*, p_{adj}<1E-300; CLEC10A, p_{adj}=3.31E-241) as well as plasmacytoid dendritic cells responsive to viral infections (*TCF4*; *IRF7*; *GZMB*, p_{adj}<1E-300 for all) (Figure 5-8 B-C). I also identified a rare population of LAMP3+ DCs (also known as mregDCs [228], *LAMP3*, p_{adj}=1.06E-72; *CCR7*, p_{adj}=1.11E-54) (Figure 5-8 B-C). The macrophages had dual origin – either airway resident which included four subsets: CCL+ (*CD68* and *IL1B*, p_{adj}<1E-300 for both), SPP1+ (*SPP1*, p_{adj}=1.22E-41; *MARCO* p_{adj}=8.19E-56), CX3CR1+ airway (*CX3CR1*, p_{adj}=3.59E-5; IL1B, p_{adj}=7.75E-269), SLC8A1+ airway macrophages (*SLC8A1*, p_{adj}<1E-300; *MEF2C*, p_{adj}=1.19E-161) or recruited from the blood – monocyte derived macrophages (*C1QA*, p_{adj}=1.44E-179; *CLEC10A*, p_{adj}=1.62E-120) (Figure 5-8 B-C). Both dendritic cells and macrophages expressed HLA proteins at high levels on their surface (Figure 5-8 B), as would be expected of professional antigen presenting cells. In the monocyte group, I identified a CXCL8+ subset (*CXCL8*, p_{adj}<1E-300) and a CXCR4+ subset (*CXCR4*, 8.00E-129), as well as an activated subset expressing *JUN* and *FOS* (p_{adj}<1E-300 for both). The monocytes were CD14+ and/or CD16+, suggesting the presence of classical and non-classical cells (Figure 5-8 B). The granulocytes expressed *IL3RA* (p_{adj}=1.12E-53, basophils) and *CR1* (p_{adj}=1.08E-23) and are likely a mixed cluster with basophils being the majority of cells (as neutrophils are poorly captured by 10X fluidics), see Figure 5-8 C. Due to larger mast cell numbers compared to the dataset in Chapter 3 (4,486 versus 848), I was able to identify both the mast cell cluster (*TPSAB1*, *TPSB2*, *CPA3*, p_{adj}<1E-300

for all) and an activated mast cell population, which, in addition to these markers, expressed CD69 and other activation markers (*CD69*, p_{adj}=1.26E-204; *LIF*, p_{adj}=8.75E-65; *CD83*, p_{adj}=3.26E-108; Figure 5-8 C), previously described to be upregulated in mast cells binding IgE or having FcεRI receptor cross-linked in a different way [23].

Consistent with previous lung atlasing efforts [78,112,229], I was not able to identify eosinophils in the granulocyte subset, despite their known involvement in the pathophysiology of asthma. Moreover, in contrast to Chapter 3, I could not identify a separate neutrophil cluster, which is likely due to a combination of technical and methodological factors – poor capture rate in the 10X Chromium chip and low-count cell filtering at the QC stage.

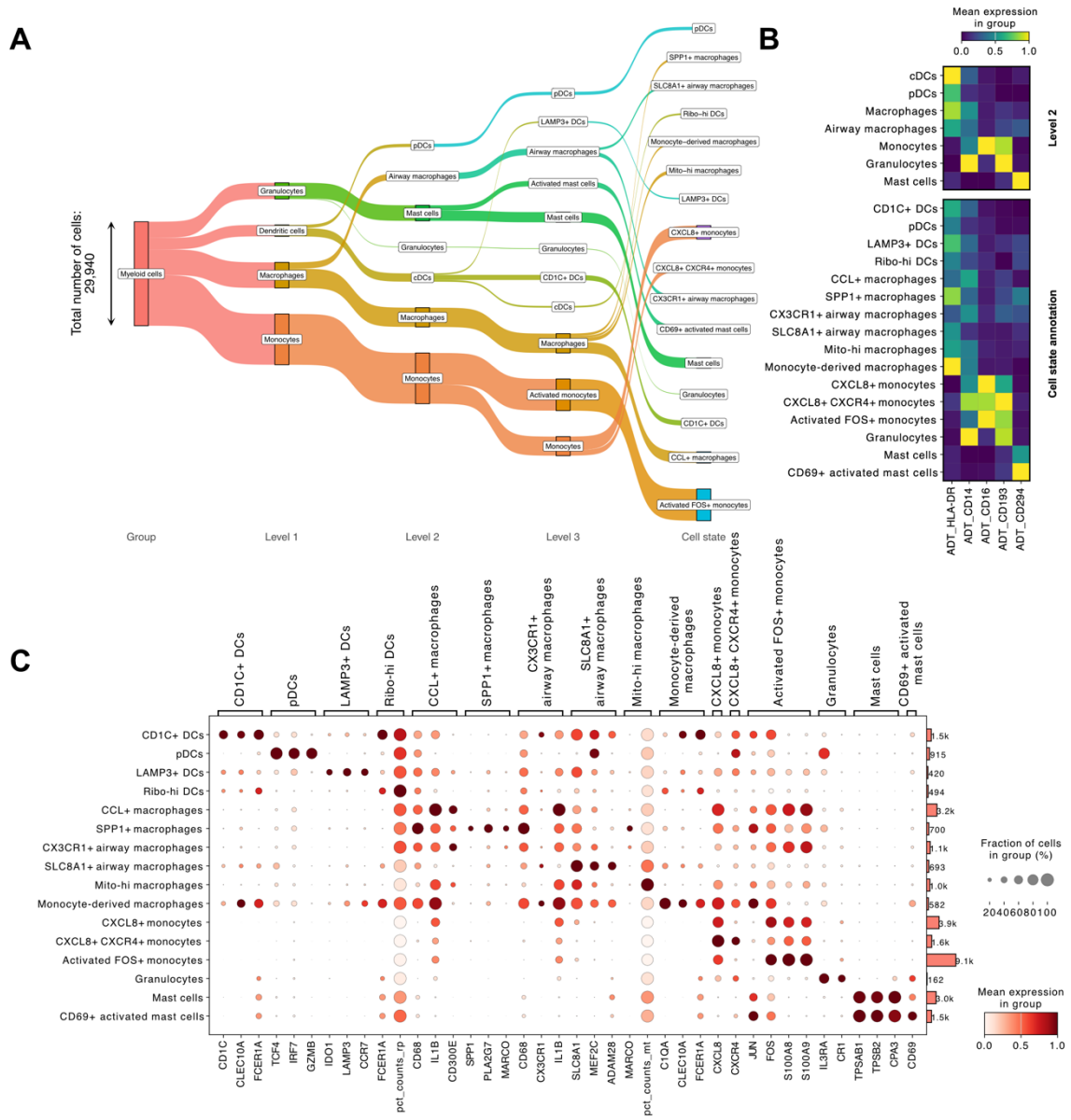

Figure 5-8 Characterisation of the myeloid compartment

A Sankey diagram of myeloid cell populations and states across increasingly detailed annotation levels. **B** Normalised surface protein staining signal across annotation level 2 and cell state annotation. **C** Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars.

5.4.6 CELL STATES IN THE NK AND T CELL COMPARTMENT

In the T and NK cell compartment, I identified CD4+ and CD8+ T cells, as well as NK cells, and unconventional subsets – gamma-delta T cells, and MAIT cells, as seen in Figure 5-9 A. CD4+ T cells constituted a cluster with naïve-like phenotype (*CD4*; p_{adj}=8.66E-21, *CCR7*; p_{adj}<1E-300, *LEF1*; p_{adj}=8.51E-196), which also expressed high levels of ribosomal proteins, regulatory T cell cluster (*FOXP3*, p_{adj}=3.96E-65; *CCR4*, p_{adj}=1.34E-25;*CTLA4*, p_{adj}=5.63E-91), and Th2 cluster (*IL5*, p_{adj}=4.04E-8; *IL13*, p_{adj}=3.23E-36; *GATA3* p_{adj}=6.79E-64) as seen in Figure 5-9 B-C.

In the CD8+ T cell compartment, I observed *JUN* and *FOS* expressing activated T cells, with distinct subsets characterised by the expression of interferon (*IFNG*, p_{adj}=1.44E-88), *XCL2* (p_{adj}=1.33E-150), *CXCR6* (p_{adj}<1E-300), *ITGA1* (p_{adj}<1E-300), *GZMK* and *CXCR4* (p_{adj}<1E-300 for both) subsets (Figure 5-9 B-C), with granzyme K (*GZMK*) expressing cells typically being considered effector memory T cells [112,230]. I also identified ribosomal-high and mitochondrial-high clusters of CD8+ T cells. In the unconventional T cell compartment, I located MAIT cells (*KLRB1*, p_{adj}<1E-300; *IL7R* p_{adj}=3.35E-241), which expressed TCR- Vα7.2 protein (Figure 5-9 B). The gamma-delta T cells expressed *KLRC2* (p_{adj}<1E-300) and *TRDC* (p_{adj}=2.94E-46), and could be divided into two clusters, one of which showing activation signature. In the innate compartment, I characterised ILCs (*SOX4*, p_{adj}=2.96E-18; *KIT*, p_{adj}=1.68E-12), which also expressed *GATA3* (p_{adj}=1.75E-20) and are thus likely ILC2s, as well as NK cells (*FCERG1*, *GLNY*, p_{adj}<1E-300 for both), which also expressed CD16 protein (Figure 5-9 B-C). I did not identify iNKT cells, and the anti-Vα24 -Jα18 antibody only showed binding to T cell subsets.

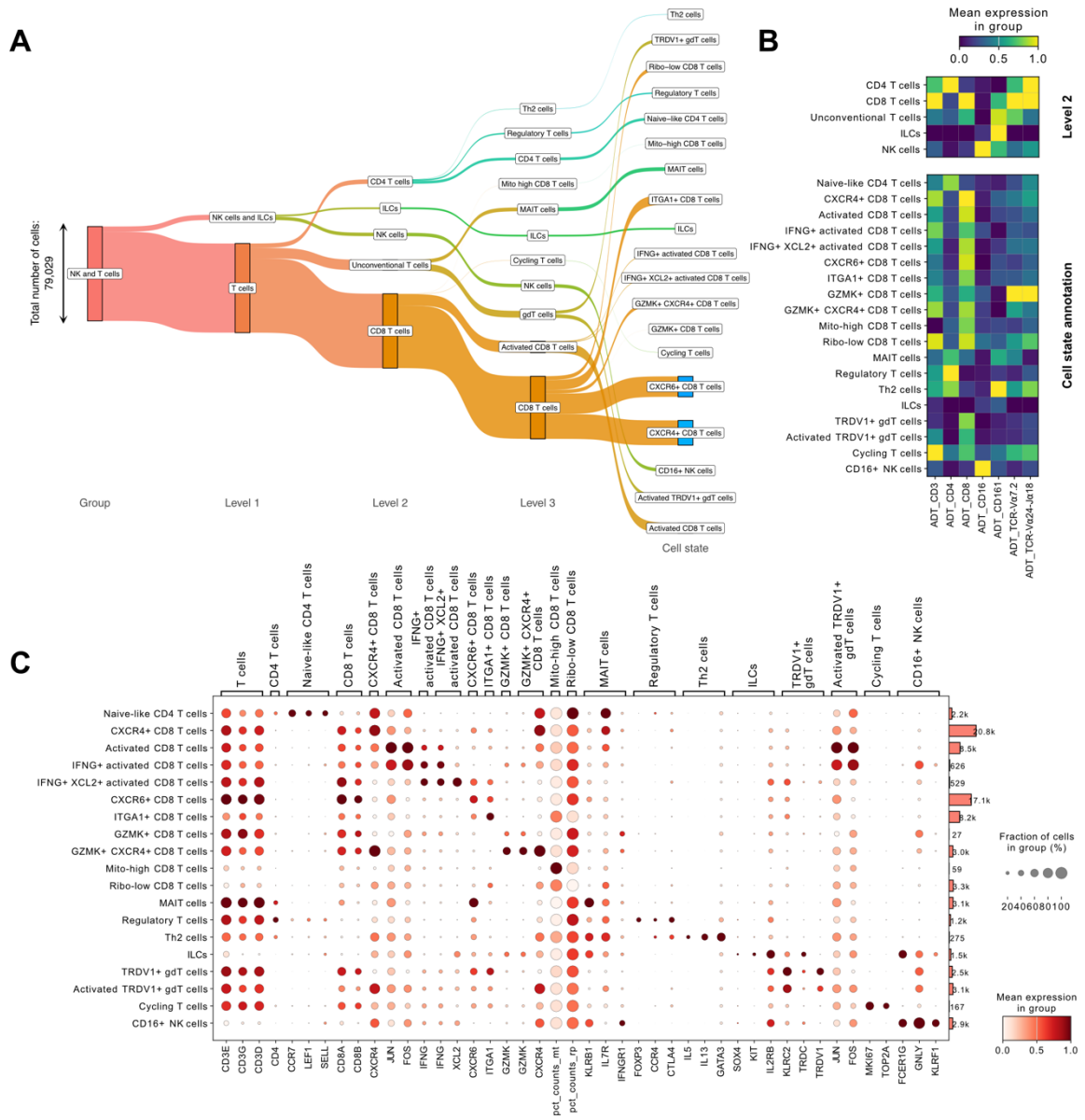

Figure 5-9 Characterisation of NK and T cell compartment

A Sankey diagram of NK and T cell populations and states across increasingly detailed annotation levels. **B** Normalised surface protein staining signal across annotation level 2 and cell state annotation. **C** Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars.

5.4.7 CELL STATES IN THE EPITHELIAL AND STROMAL COMPARTMENTS

Within the epithelial and stromal compartment, I observed a range of stromal and epithelial cell subtypes (Figure 5-10 A). The epithelium consisted of progenitor basal (*KRT5*, *TP63*, p_{adj}<1E-300 for both) and suprabasal cells (*KRT19*, p_{adj}<1E-300), together with their activated state, characterised by expression of *JUN* and *FOS* transcription factors (Figure 5-10 A-B). I also identified cells which differentiate from these progenitors, including club cells (*SCGB3A1*, p_{adj}=6.49E-123; *TSPAN8*, p_{adj}<1E-300), deuterosomal cells (*CDC20B*, p_{adj}<1E-300; *FOXN4*, p_{adj}=1.27E-97), and mucous-secreting goblet cells (*MUC5AC*, *CEACAM5*, p_{adj}<1E-300 for both), see Figure 5-10 A-B. I identified three subgroups of the ciliated cells, which shared expression of *FOXJ1* and *PIFO* (both p_{adj}<1E-190 in every subset), but formed distinct clusters on the UMAP, with type 2 ('mucous') cells expressing high levels of *MUC5AC* (p_{adj}=5.16E-13) (Figure 5-10 B and Figure 5-5 F). I also discovered rare cell types, such as ionocytes (*FOXI1*, *CFTR*, p_{adj}<1E-300 for both) and Hillock cells (sometimes referred to as 'Hillock-like'), which expressed *KRT4* (p_{adj}=1.94E-130) and *KRT13* (p_{adj}=4.09E-165), thought to be injury-resistant reservoirs of stem cells in the airway [231] (Figure 5-10 B). I also identified neuroendocrine cells, secreting neuropeptides such as *CHGA* (p_{adj}=3.71E-19) and *CALCA* (p_{adj}=1.25E-11) (Figure 5-10 B). Finally, I identified -stromal cells – fibroblasts (*PI16*, *CXCL14*, p_{adj}<1E-300 for both) and pericytes (*ACTA2*, p_{adj}=1.55E-146; *TAGLN*, p_{adj}=2.28E-148) (Figure 5-10 B). All the anticipated cell types have been captured in the annotation, except the tuft cells, which I was not able to reliably distinguish. Tuft cells have been previously described by some (e.g. HLCA [78]), but not other single cell studies [112].

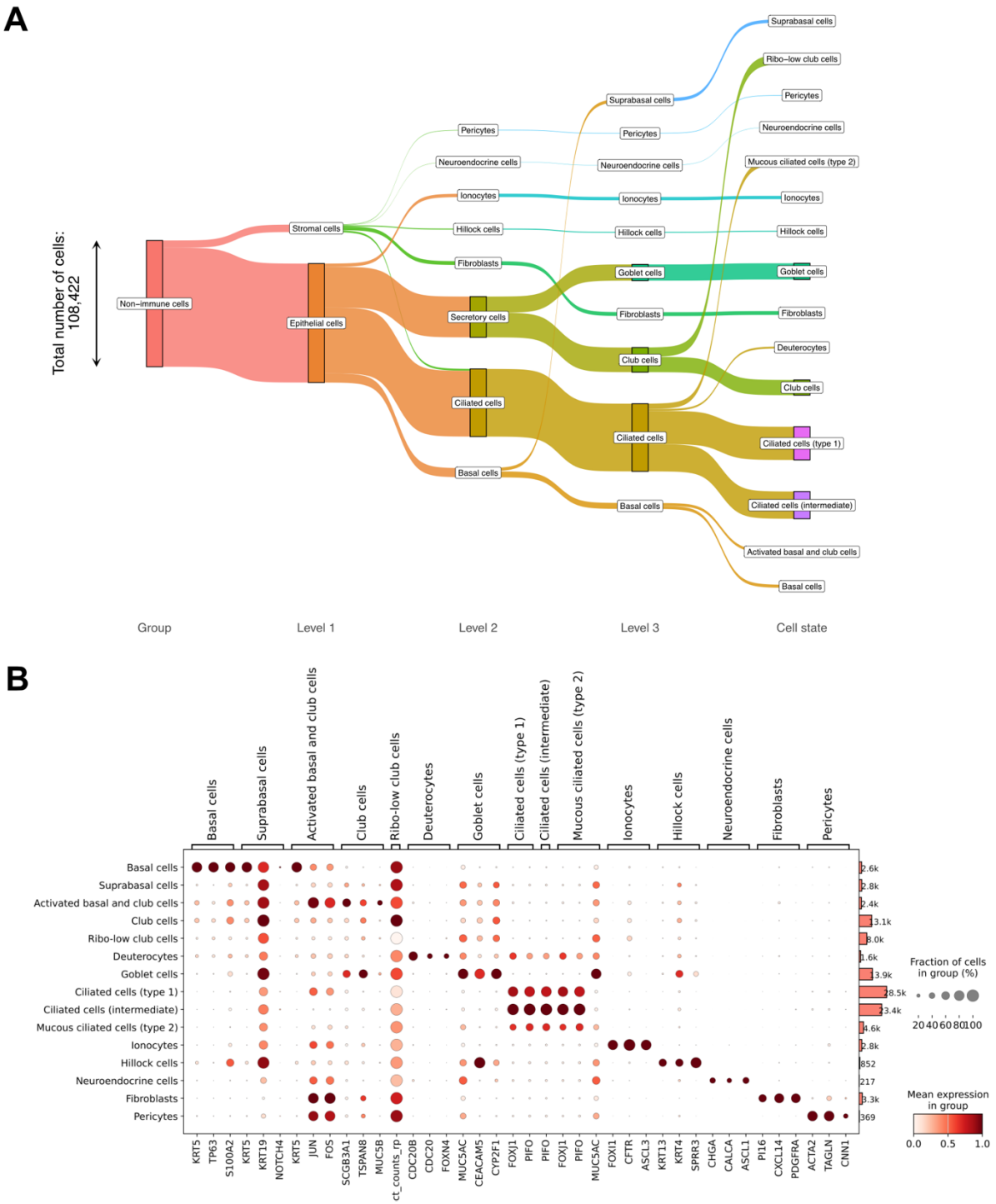

Figure 5-10 Characterisation of non-immune cell compartment

A Sankey diagram of NK and T cell populations and states across increasingly detailed annotation levels. **B** Dotplot demonstrating normalised expression levels of key markers used in cell state annotations, with the per cell-type cell count indicated as the bars.

5.4.8 ASSOCIATION OF GENE EXPRESSION WITH CLINICAL VARIABLES

To understand how the clinical data correlates with the gene expression, as well as to identify any spurious correlations to correct for at the differential gene expression testing stage, the association of pseudobulk-level principal component scores with clinical variables was calculated, as seen in Figure 5-11. The first few components, which explain the most variance, do not associate with clinical variables (Figure 5-11 A-B), but do separate patients into homogenous clusters (especially PCs1-5, Figure 5-11 C). Starting with principal component 9, multiple PCs correlated with clinical variables relevant in asthma, such as eosinophil count, FeNO measurement, presence of specific IgE, or clinical phenotype (Figure 5-11 A). In addition, some PCs, such as 17, 18, 28, associated with other confounding variables, such as sex, BMI, or smoking status (Figure 5-11 A). PC17 was driven by expression of molecules such as *GNAO* and *NRXN1*, both previously associated with neurodegenerative diseases [232,233], thus it can be hypothesised that neuroendocrine cells are involved, as well as IL-6, known to be secreted by dendritic cells and macrophages in severe asthma [234] (Figure 5-11 D). PC18 changes were driven by *FCER1A*, which encodes high affinity IgE receptor, FcεRI, present on mast cells and basophils [4], and *TNFSF8*, encoding CD30, a TNF receptor (Figure 5-11 E). *FCER1A* polymorphisms were demonstrated to influence the IgE levels in allergic asthma [195], whilst CD30 expression was linked to increased inflammation, serum IgE levels, and type-2 cytokines in mouse models [235]and humans [236].

Whilst the variables such as age, sex, BMI, and smoking status are interesting on their own, and have all been linked to asthma [3], they may obscure the picture in the gene expression testing. Together with ethnicity, they have been included as confounding variables in the differential gene expression analysis. The DGE analysis was conducted for the cells in the upper airway (nasal brush), the lower airway (biopsy), and disregarding the tissue of origin (all), to increase the statistical power and capture general trends, and the results are summarised in Figure 5-11 F. Overall, the patterns were broadly consistent across the upper and lower airway, with monocytes, airway macrophages, and cDCs showing the greatest amount of significant DGEs, consistent with the disease mechanism. In the lower airway, mast cells exhibited a greater number of DGEs than in the upper airway (351 vs 143). Conversely, fibroblasts showed more of DGEs in the upper airway compared to lower airway (362 vs 29). Some cell types (T cells, basal cells, macrophages, plasma cells) showed few or no differentially expressed genes, despite their known role in asthma pathogenesis.

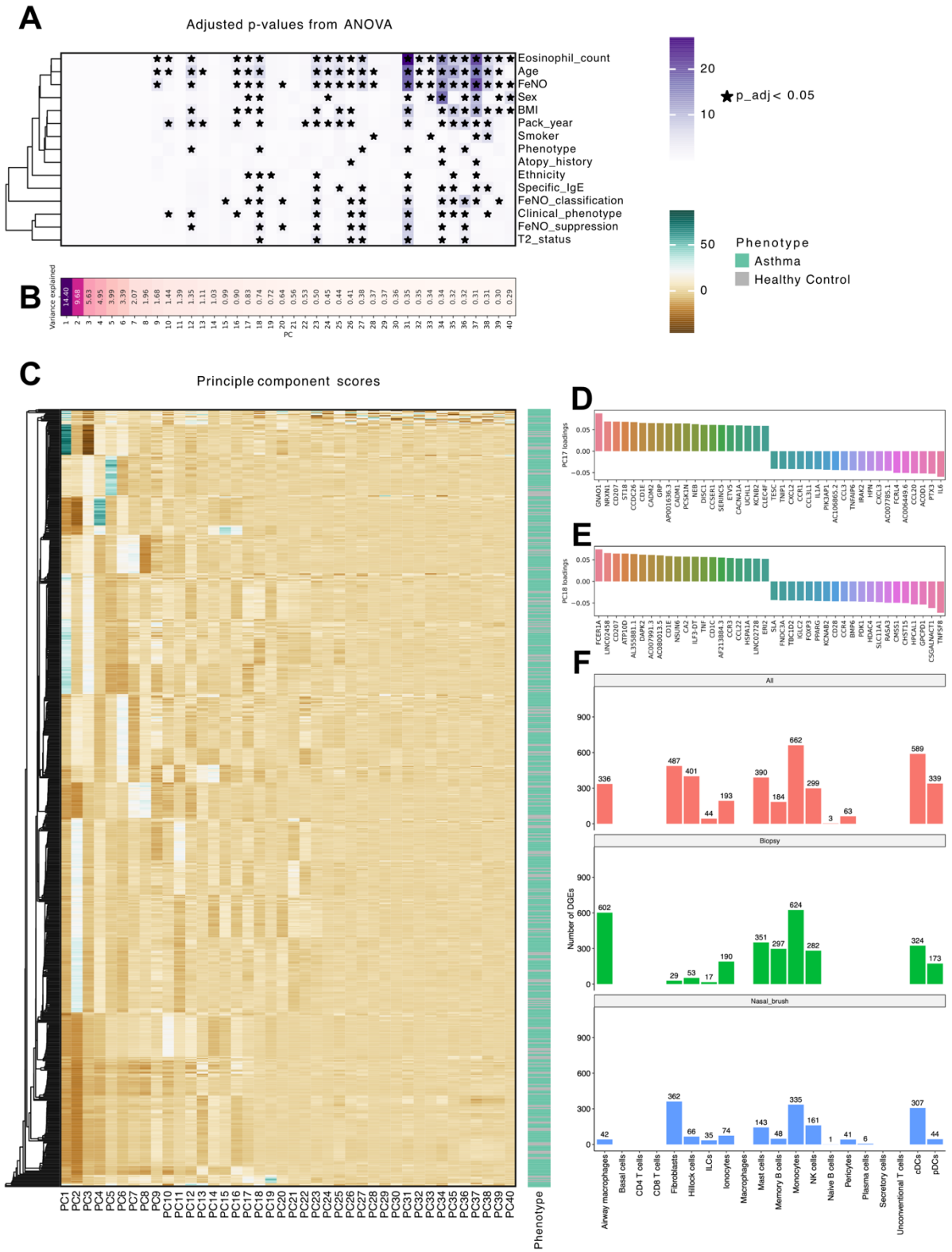

Figure 5-11 Association of principal components calculated from pseduobulk gene expression values with clinical variables

A Correlation score and ANOVA test result for clinical variables collected for the cohort. **B** Percentage of variance explained for each principal component. **C** Principal component scores calculated for each gene – patient combination at annotation level 2 with patient diagnosis included as 'Phenotype'. **D-E** Top and bottom 20 loadings of PC17 and PC18, example PCs with high clinical variable association. **F** Number of statistically significant genes for all the cells tested and in tissue-specific comparisons when adjusting for correlated clinical variables not related to disease (sex, age, BMI, ethnicity, smoker status) at annotation level 2.

5.4.9 ASSOCIATION OF CELL ABUNDANCE WITH CLINICAL VARIABLES

Similarly to gene expression, I investigated correlations of clinical variables with the cell abundance calculated per compartment (annotation level 0), as seen in Figure 5-12. Only two principal components – PC2 and PC3, explaining 21% and 15% of dataset variance, respectively – were associated with the clinical variables measured (Figure 5-12 A-B). There was no obvious clustering between healthy and asthmatic patients (Figure 5-12 C). PC2 associated with presence of IgE specific to an aeroallergen (Figure 5-12 A), and was most influenced by monocytes and epithelial (secretory and ciliated) cells (Figure 5-12 D). The contribution of epithelial and mucous-secreting cells would be consistent with damage to the epithelial barrier, discussed in Chapter 1, whereas contribution of monocytes supports high number of DGEs expressed between healthy and asthmatic patients. Conversely, PC3 values were most influenced by memory B cells, plasma cells, and mast cells numbers, and it associated with a measure of airway type-2 inflammation (FeNO value and FeNO classification – 'low', 'intermediate', and 'high'), as seen in Figure 5-12 E. Overall comparison of cell type abundances across the upper and lower airway again reveals similar patterns, with only a few cells changing the proportion – for example, mast cells and ciliated cells increase in abundance in asthma, whereas memory B cells decrease in proportion (Figure 5-12 F). In summary, no confounding factors appear to be associated with the cell abundances, beyond clinically relevant variables. To draw robust conclusions about relative cell proportions, statistical modelling of cell type proportions is introduced in section 5.4.10.

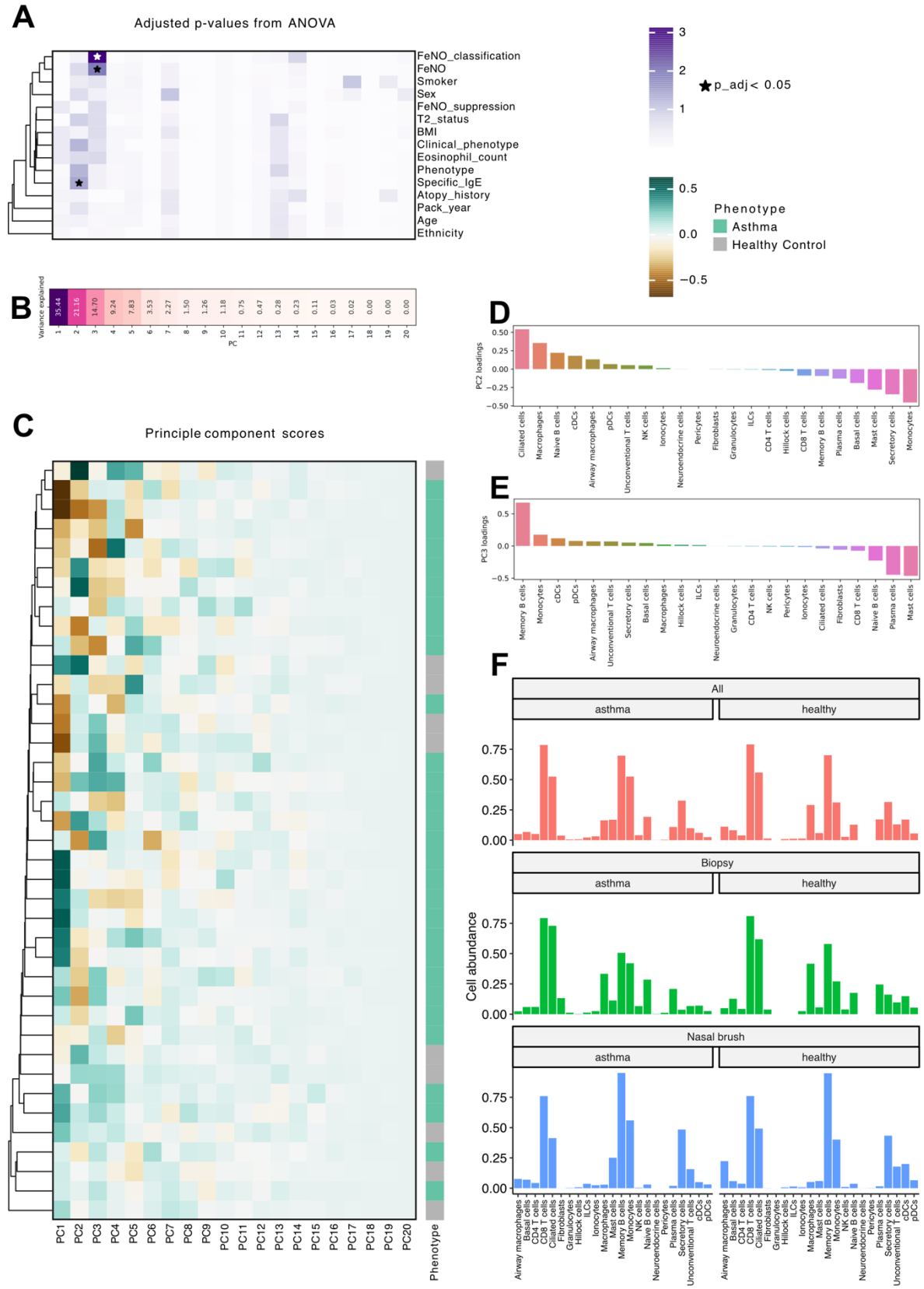

Figure 5-12 Association of principal components calculated from cell abundance values with clinical variables

A Correlation score and ANOVA test result for clinical variables collected for the cohort. **B** Percentage of variance explained for each principal component. **C** Principal component scores calculated for each cell type - patient combination at annotation level 2 with patient diagnosis indicated as 'Phenotype'. **D-E** Loadings of PC2 and PC3, PCs with significant clinical variable association. **F** Cell type abundance calculated per compartment in asthma and health across all the cells, in the cells from nasal brushes, and in the cells from biopsies.

5.4.10 CELL ABUNDANCE CHANGES ACROSS TISSUES IN ASTHMA AND HEALTH

To understand broad patterns in cellular composition of samples collected from healthy and asthmatic individuals, I calculated the proportion of cells in each compartment (annotation level 0: NK and T cells, myeloid cells, B and plasma cells, non-immune cells) and clustered them using Bray-Curtis dissimilarity index, taking into consideration the sample origin (upper or lower airway) and overall (Figure 5-13 Cellular composition of

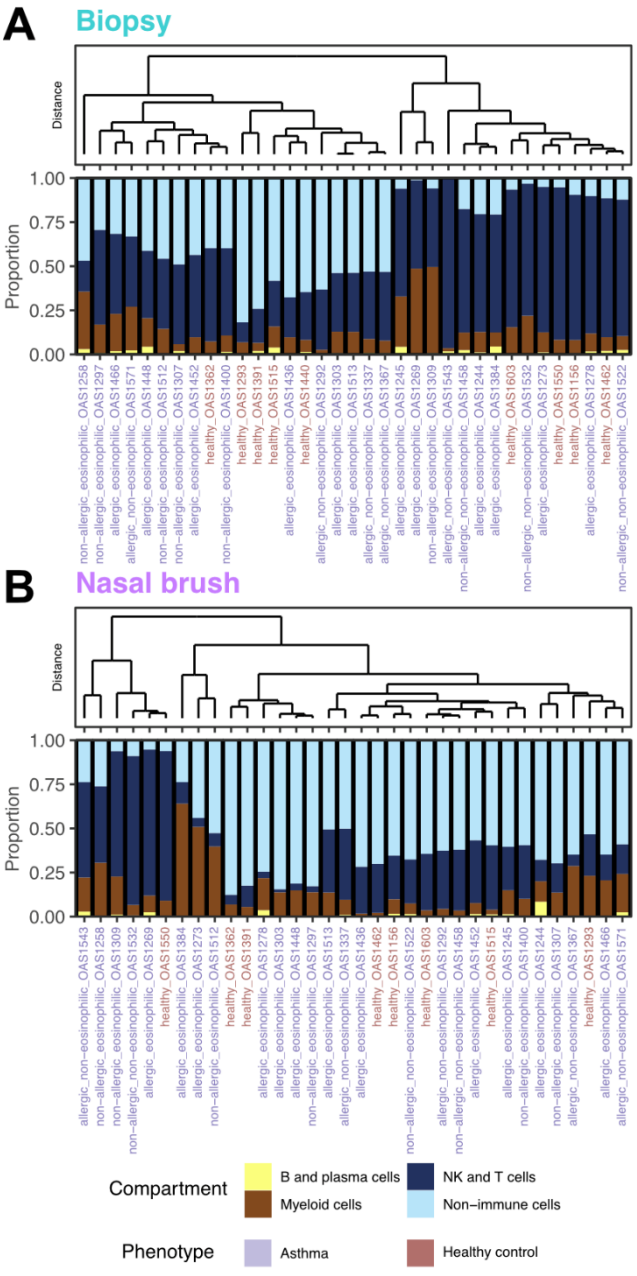

Figure 5-13 Cellular composition of samples in upper and lower airway in asthmatic and healthy patients
Stacked barplots showing proportion of cells in each compartment (NK and T cells, myeloid cells, B and plasma cells, non-immune cells) within individual patient samples in the **A** biopsies and **B** nasal brushes. Samples are ordered according the UPGMA (unweighted pair group method with arithmetic mean) clustering on Bray-Curtis dissimilarity index and coloured by the diagnosis (asthma, healthy).

samples in upper and lower airway in asthmatic and healthFigure 5-13). Overall, the biopsies contained more immune cells compared to the nasal brushes, where the epithelial cells were more represented (Figure 5-13 Cellular composition of samples in upper and lower airway in asthmatic and healthFigure 5-13). In the biopsies, the samples divide into two subgroups, one with balanced epithelial/immune compartment, and another with characterised by high abundance of NK and T cells – as the samples in either group do not map to any distinct patient diagnosis (asthmatic or healthy) or phenotype, this could be explained by the sampling bias and the effect of sampling location and the operator on cellular composition of the biopsy. Nonetheless, a smaller cluster predominantly composed of myeloid cells can be identified in the data (patients OAS1245, 1269, 1309, see Figure 5-13 A), which consists of eosinophilic asthmatic patients. Similar patterns could be observed in the nasal brushes, with a broad divide into two subsets – predominantly NK and T infiltrated, and a balanced cluster with larger presence of epithelial cells (Figure 5-13 B). Again, a small cluster of eosinophilic asthmatic patients with large myeloid infiltrates could be identified (OAS1273, 1384, 1512, see Figure 5-13 B). Whilst these broad patterns help with understanding of overall cellular composition of airway samples and may point to a sampling bias, a more robust statistical framework is needed to draw conclusions about the individual cell types.

To further characterise the changes in abundance of each cell type in immune and epithelial tissue across asthma phenotypes, I again employed *miloR* to characterise cellular neighbourhoods, as shown in Figure 5-14. The KNN graph, with the average neighbourhood size in the range of 100-200 (mean: 186.9, SD: 64.7), showed separation of cell compartments – B/plasma, myeloid, NK/T, and non-immune subset, with non-

immune subset further splitting into epithelial and stromal compartments, as can be seen in Figure 5-14 A. Comparing the asthmatic patients with healthy controls revealed that the relative proportions of most cells remained stable (Figure 5-14 B). A low number of Th2 cells identified (274) unfortunately precluded any neighbourhoods consisting of Th2 cells from being identified. The other key driver of eosinophilic inflammation, ILCs, showed an increase in abundance in the upper, but not the lower airway. ILC2s have been reported to be involved in the pathogenesis of eosinophilic chronic rhinosinusitis with nasal polyps (CRSwNPs), another disease driven by type-2 cytokines [237]. The abundance of mast cells was consistently higher in patients with asthma across the upper and lower airway. Interestingly, some rare cell populations, such as gamma-delta T cells, Hillock cells, were decreased across the samples, with other cell types, such as memory B cells, airway macrophages, and pDCs being present in lesser abundance in biopsies only. Goblet cell metaplasia is a widely reported feature of asthma [11], however my data suggests that there is a general decrease in cell types (goblet, ciliated, club) in asthma compared to health, which may result from the breakdown of the epithelial barrier.

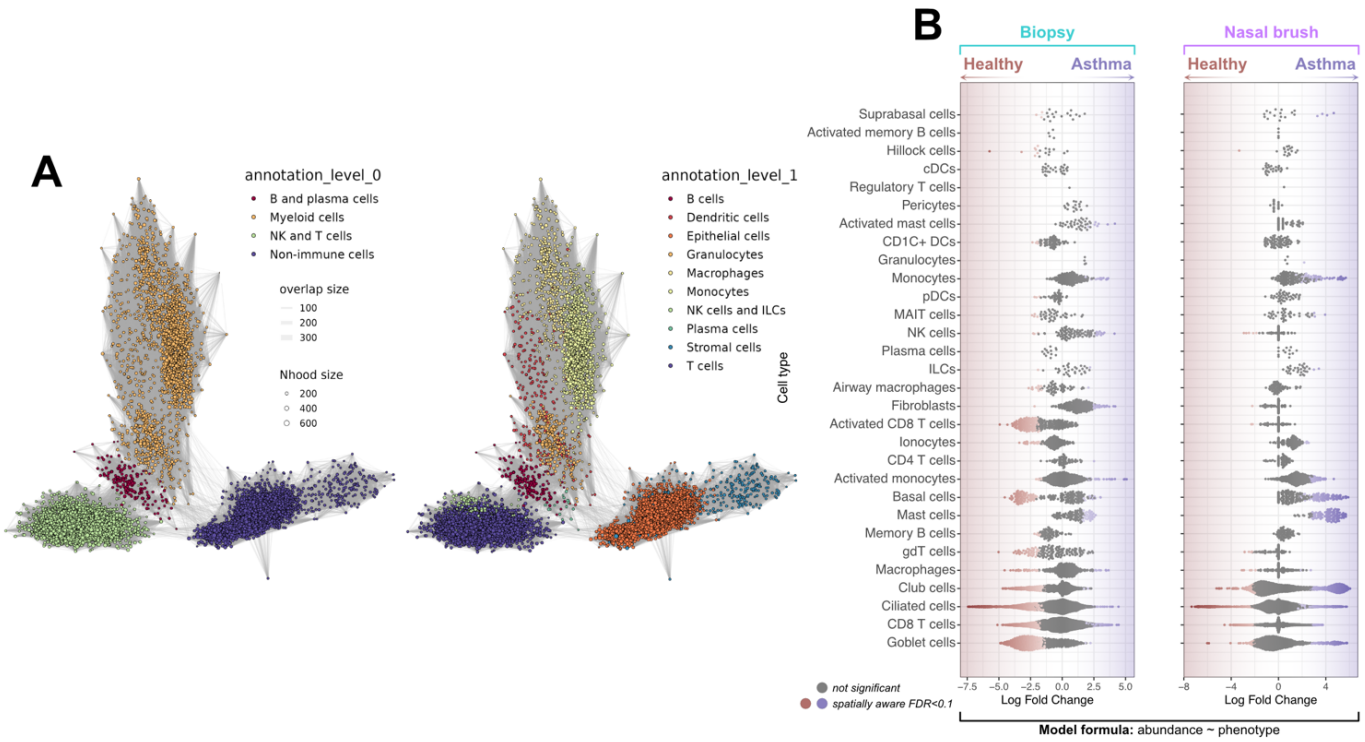

Figure 5-14 Differential abundance of epithelial, stromal, and immune cells in severe asthma and health
A Neighbourhood graphs showing cell neighbourhoods and relative neighbourhood distances, coloured by annotation levels 0 (compartment) and 1 (broad cell type annotation).**B** Differential abundance as log fold change in cell proportions between asthma and health in lower (biopsy) and upper (nasal brush) airway at annotation level 3.

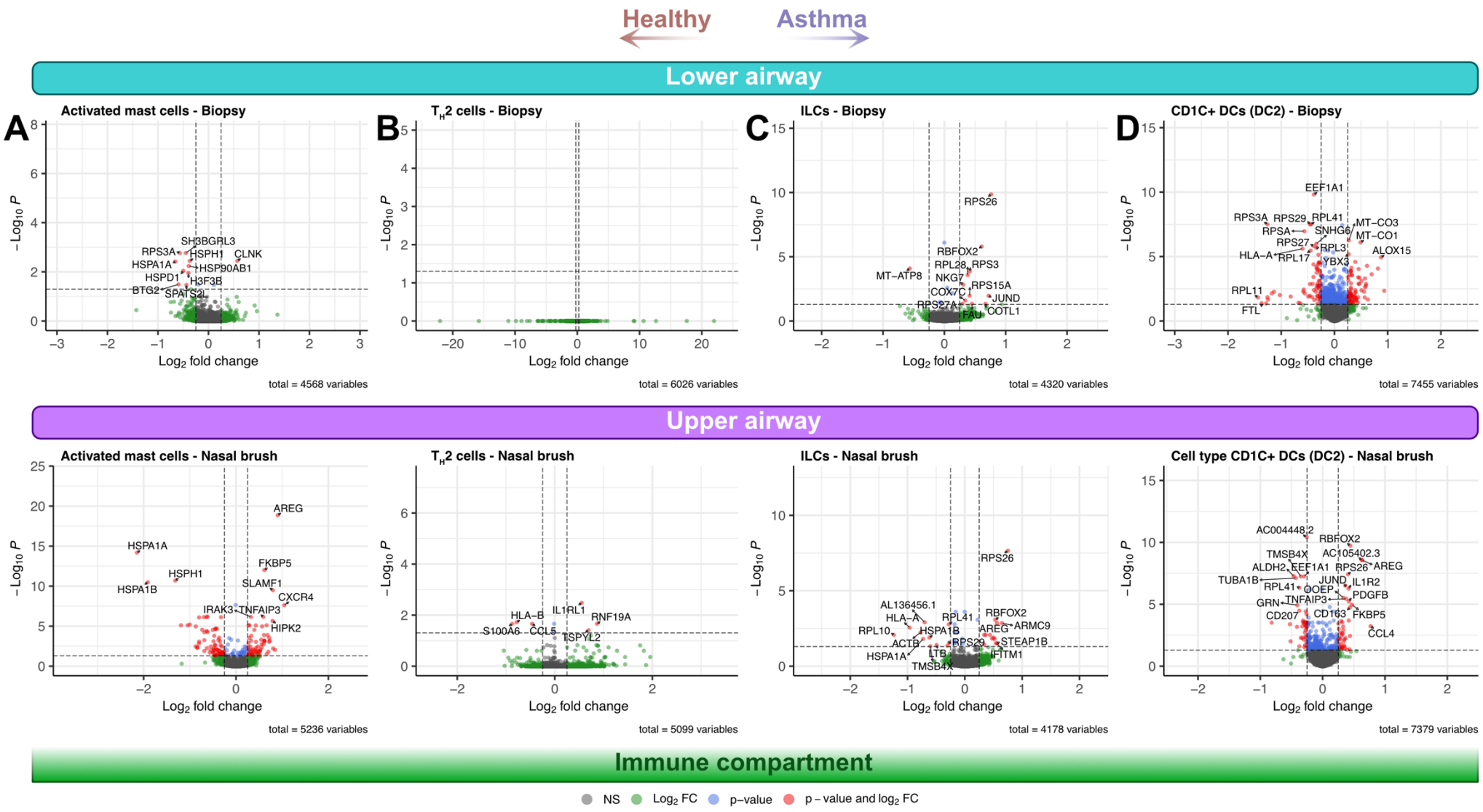

Figure 5-15 Gene expression changes across the tissues in upper and lower airway in the immune compartment in asthma versus health

A-D Volcano plots showing gene expression changes in select immune cell types are key effector cells in type-2 inflammation at annotation level 3 in samples from the upper (nasal brush) and lower (biopsy) airway.

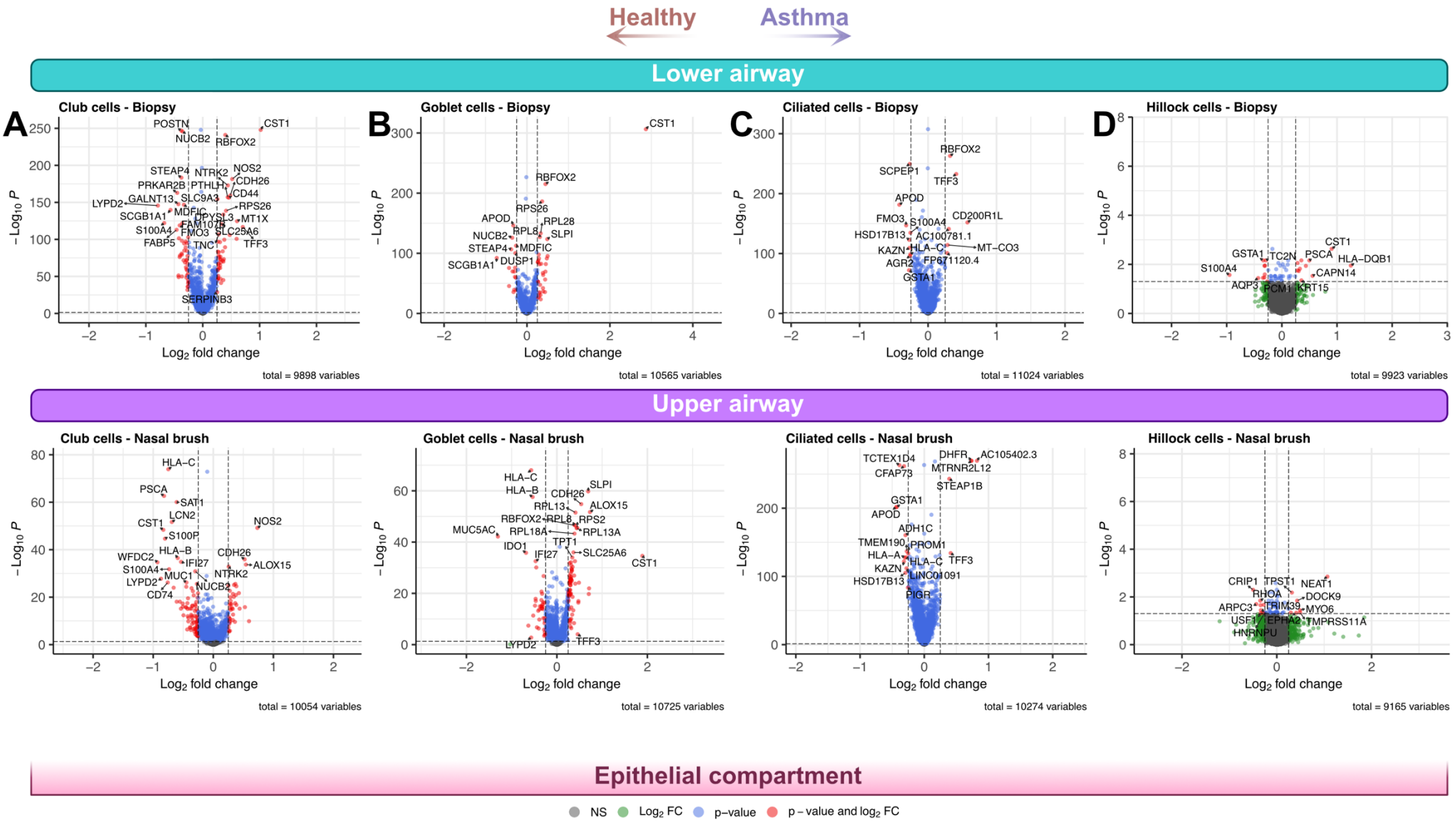

Figure 5-16 Gene expression changes across the tissues in upper and lower airway in the epithelial compartment in asthma versus health

A-D Volcano plots showing gene expression changes in major epithelial cell types (club, goblet, and ciliated) cells and recently discovered Hillock cells in samples from the upper (nasal brush) and lower (biopsy) airway.

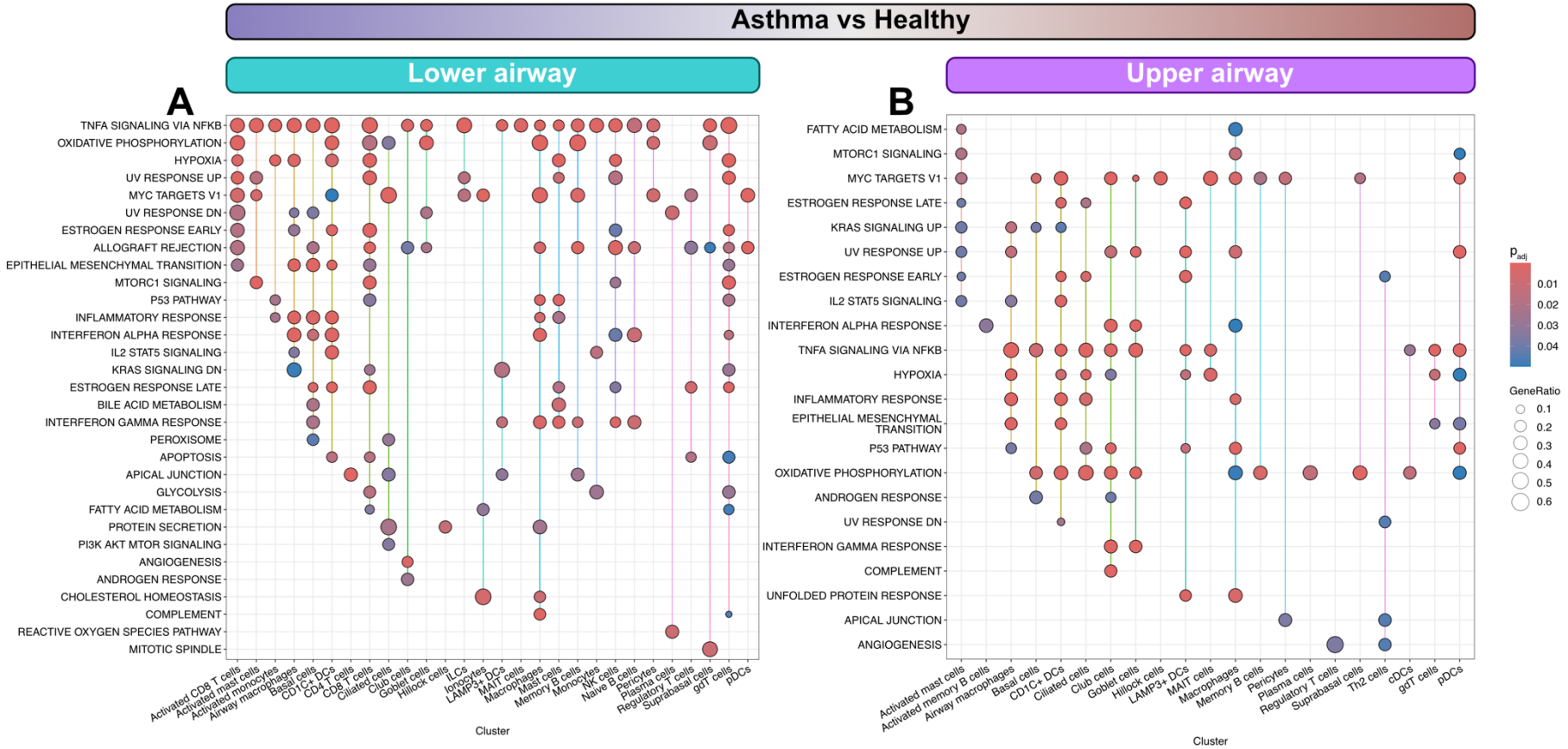

Figure 5-17 MSigDB Hallmark dataset gene-set enrichment analysis in upper and lower airway in asthma versus health
Statistically significant pathways are shown per cell type at annotation level 3 in biopsy (**A**) and nasal brush (**B**) samples.

5.4.11 GENE EXPRESSION CHANGES IN ASTHMA AND HEALTH

5.4.11.1 IMMUNE COMPARTMENT

To understand the transcriptional changes resulting from asthma, I investigated differential gene expression patterns across the upper (nasal brush) and lower (biopsy) airway samples and identified notable differences between these anatomical regions, and between asthma and health. After carrying out a differential gene expression analysis on per-cell type calculated pseudobulk profiles using MAST [130], accounting for confounding variables, including age, sex, ethnicity, BMI, and smoking status, several cell types associated with type-2 inflammation revealed intriguing insights (Figure 5-15). In activated mast cells, which differed in gene expression profile from the mast cells, *SLAMF1* and *AREG* were upregulated in cells obtained from the nasal brushes (Figure 5-15 A).

SLAM receptors are transmembrane proteins expressed on many immune cell types, such as monocytes, macrophages, NK cells, CD8+ T and B lymphocytes, and dendritic cells [238,239]. SLAMF1 plays a role in CD8+ T cell activation and interferon production [239] and incudes production of type-2 cytokines, such as IL4 [240]. In human macrophages, SLAMF1 has been demonstrated to regulate TLR4-mediated response to bacteria by forming a protein complex which promotes IFNβ secretion [241]. TLR4 is a receptor for HDM proteins and potentially other allergens [242], which I speculate may highlight the influence of allergens on other cell types beyond epithelium in type-2 high asthma. Upregulation of *SLAMF1* in the upper airway would be consistent with it being the first site of exposure to inhaled allergens.

Amphiregulin, encoded by *AREG*, is found at increased concentration in the blood of patients with asthma, and has been proposed as a biomarker. Whilst its role in airway inflammation is debated, with some proposing its involvement in ILC2-mediated resolution of inflammation [243], recent research demonstrates that AREG mediates the TGFβ-induced epithelial-mesenchymal transition in airway epithelium via a mechanism involving EGFR/JNK/AP-1 activation [244].

Alarmins play a central role in across phenotypes of eosinophilic asthma, with IL33 being one of the key alarmins (Figure 5-1). IL33 is a member of the IL1 family, and together with its cognate receptor, interleukin-1 receptor like-1 (IL-1RL1 or ST2) , both are susceptibility genes for childhood-onset asthma identified in many GWASes, as summarised in chapter 1, and a regulator of IL33 concentration in the lung parenchyma [245]. Its upregulation in Th2 cells in the upper airway may suggest that IL33 is also driving inflammation in the upper airway (Figure 5-15 B). Of note, no genes were found to be differentially expressed in Th2 cells in the lower airway when broad asthma-health comparison was used.

Among ILCs (Figure 5-15 C), the patterns of differential gene expression were in broad agreement between the upper and lower airway. Amphiregulin, *AREG*, was again upregulated, potentially highlighting multiple sources of this molecule acting on the airway epithelium. Other DEGs mostly included ribosomal genes, indicative of either general activation pattern, or an artifact of cell handling and dissociation [120].

In DC2s (Figure 5-15 D), important for antigen presentation in asthma, *ALOX15*, a gene encoding 15 lipoxygenase 1 (15LO1) protein and induced by type-2 cytokines, was upregulated in the lower airway, with evidence of 15LO1 metabolites secreted by immune cells leading to airway epithelial injury in animal models of asthma [246].

Overall, the immune cell types expressed genes with previous links to asthma through either genome-wide association or mechanistic studies, with some potential biomarkers expressed across multiple cell types, such as amphiregulin, identified as potential biomarkers of asthma in the clinic.

5.4.11.2 EPITHELIAL COMPARTMENT

Next, to investigate the gene expression profiles in asthma in the epithelial compartment, I focused on the predominant airway epithelial cell types – club, goblet, and ciliated cells – as well as rare cell types, such as Hillock cells, where notable changes were observed, as seen in Figure 5-16.

In club cells (Figure 5-16 A), I observed the upregulation of nitric oxide synthase 2 (*NOS2*), both in the upper and lower airway, with nitric oxide being used as a biomarker of type-2 inflammation in asthma in the clinic [6,68], but with limited understanding of its cellular source in the epithelium. In addition, similarly to *ALOX15* upregulation in the immune compartment, it was also upregulated in club cells in the upper airway. *CST1*, one of the top DGEs both in terms of log-fold change and p value, was reported to promote airway

eosinophilic inflammation in asthma via the *AKT* signalling pathway [247] and inhibit HDM protease activity in allergic asthma [248].

Similar to club cells, other secretory cells, goblet cells (Figure 5-16 B), displayed upregulation of *CST1* and *ALOX15* genes. In addition, *SLPI* gene was upregulated in both upper and lower airway. It encodes secretory leukocyte protease inhibitor, which is capable of broad spectrum inhibition of mast cell and leukocyte serine proteases [249]. In animal models of asthma, SLPI inhibits allergen-induced inflammation, and may be part of the physiological regulatory response of the airway in allergic asthma [250,251].

In ciliated cells (Figure 5-16 C), *CD200R1* was upregulated in the biopsy samples. This anti-inflammatory receptor was previously thought to be mainly expressed on non-epithelial cell types [252], with reports of dysregulation in asthma, and is thought to contribute to the resolution of inflammation in the lung [253]. One of the top differentially expressed genes in the nasal brushes and biopsies was *TFF3*, previously reported to be present at elevated levels in the sputum of patients with COPD and asthma. TFF3 was shown to have an immune-resolving effect via downregulation of TNFα, TLR4, and NF-κB [254], as well as wound healing throughout mucosal tissues [255], including in the airway [256].

In recently characterised Hillock cells (Figure 5-16 D), *CST1* was again observed to be upregulated, in the biopsies. Also in the lower airway, *CAPN14*, a gene in the locus epigenetically regulated by type-2 cytokines and associated with eosinophilic inflammation, was identified [257].

In summary, I identified secretory cells (goblet and club cells) as the likely cellular sources of nitric oxide in asthma, and characterised broad gene changes related to the resolution of inflammation in the airway epithelium, pointing to the pro-inflammatory role of immune cells and regulatory role of the epithelium in asthma. Broadly, the changes observed in contrasts of asthma patients without phenotyping and healthy controls skew towards type-2 inflammation, with few genes being identified relating to type-2 low inflammatory response.

5.4.11.3 PATHWAY ANALYSIS

To understand the overall signalling landscape across the cell types in asthma when compared to health, I performed GSEA using MSigDB database, similarly to Chapter 3 (Figure 5-17). Overall, enrichment of more pathways was observed in the lower airway compared to upper airway. TNFα signalling was upregulated across both tissues and in most cell types. While anti-TNFα therapies were trialled for asthma in the late 2000s [179], the treatment had to be discontinued early due to an unfavourable risk-benefit profile. The inefficacy of treatment may stem from the widespread upregulation of TNFα signalling across multiple cell types, making cell-type precise targeting challenging when designing drug candidates. Nevertheless, TNFα remains a key pro-inflammatory cytokine in asthma pathogenesis.

Epithelial-mesenchymal transition (EMT) term was significantly enriched across both tissue types in immune (CD8+ T cells, macrophages) and epithelial (ciliated, basal cells) compartments, and was also evident at DGE-level, consistent with findings reported in

chapter Chapter 3. In a healthy airway, EMT contributes to tissue healing by generating mesenchymal cells necessary for regeneration. However, the context of chronic inflammation, excessive EMT can lead to tissue fibrosis and lung damage, a process where fibroblasts accumulate and secrete collagen fibres [182].

Finally, interferon α and ɣ signalling terms were both upregulated in patients with asthma, which perhaps can be tied to type-2 low responses, characterised by upregulation of inflammasomes, as well as type I, II and III interferons, and IL17 signalling [258], or a broader feature of chronic inflammation present in the airway.

In summary, the enrichment analysis shows various sources of inflammation across the immune cell types in asthma, along with evidence of airway fibrosis.

5.4.12 CELL ABUNDANCE CHANGES IN EOSINOPHILIC VERSUS NON-EOSINOPHILIC ASTHMA

Broad patterns in cell abundance changes can be observed by comparing healthy individuals to those with asthma. To better understand the heterogeneity within asthma phenotypes and examine the changes associated with type 2 immune response, I contrasted eosinophilic asthma with non-eosinophilic asthma in the upper and lower

airway. The results of the differential abundance analysis across the tissues are presented in Figure 5-18.

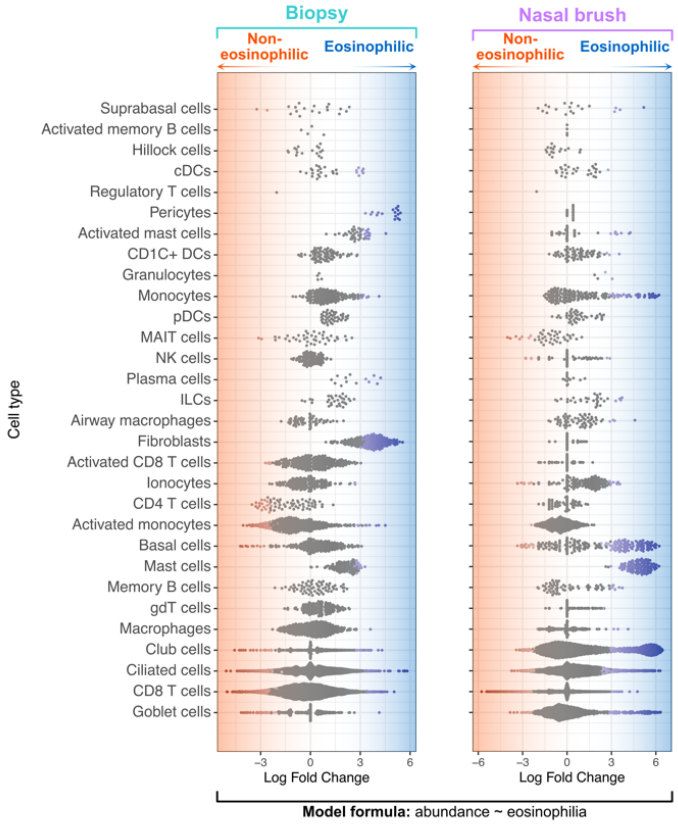

Figure 5-18 Differential abundance of epithelial, stromal, and immune cells in eosinophilic and non-eosinophilic asthma

Differential abundance as log fold change in cell proportions between eosinophilic and non-eosinophilic asthma in lower (biopsy) and upper (nasal brush) airway at annotation level 3.

In the biopsies, there was an increased abundance of fibroblasts, pericytes, mast cells and activated mast cells, cDCs, and plasma cells. The increased presence of mast cells, plasma cells, and dendritic cells is consistent with the type-2 inflammation. The increased abundance of fibroblasts and pericytes may be attributed to the EMT or fibrosis, further suggesting it may play a role in eosinophilic asthma. Some of those shifts in airway cell populations are mirrored in the upper airway – mast cells and activated mast, as well as cDCs cells increase in abundance – and there is a marked increase in the monocytes and granulocytes. Changes in epithelial composition do not seem to be present in the upper airway except for an increase in the abundance of goblet cells,

consistent with goblet cell metaplasia, when comparing eosinophilic and non-eosinophilic asthma (Figure 5-18).

Conversely, in non-eosinophilic asthma, in the biopsies there is an increase in the presence of CD4+ T and MAIT cells, as well as basal and suprabasal cells. MAIT cells have been described, by Hinks group [259] and others [260], to be involved in the pathogenesis of neutrophilic asthma and correlate with the number of exacerbations in severe asthma. The increase in abundance of MAIT cells in non-eosinophilic asthma is recapitulated in the upper airway (Figure 5-18). In addition, a subset of CD8+ T, club, and ciliated cells also seem to increase in abundance in non-eosinophilic asthma, however, further investigations would be needed to establish how it differs from the subset increased in eosinophilic phenotype.

Overall, distinct patterns of cell abundance can be observed when contrasting eosinophilic and non-eosinophilic asthma, with the involvement of mast cells, DCs, and plasma cells in eosinophilic type, and CD4+ T and MAIT cells in non-eosinophilic asthma.

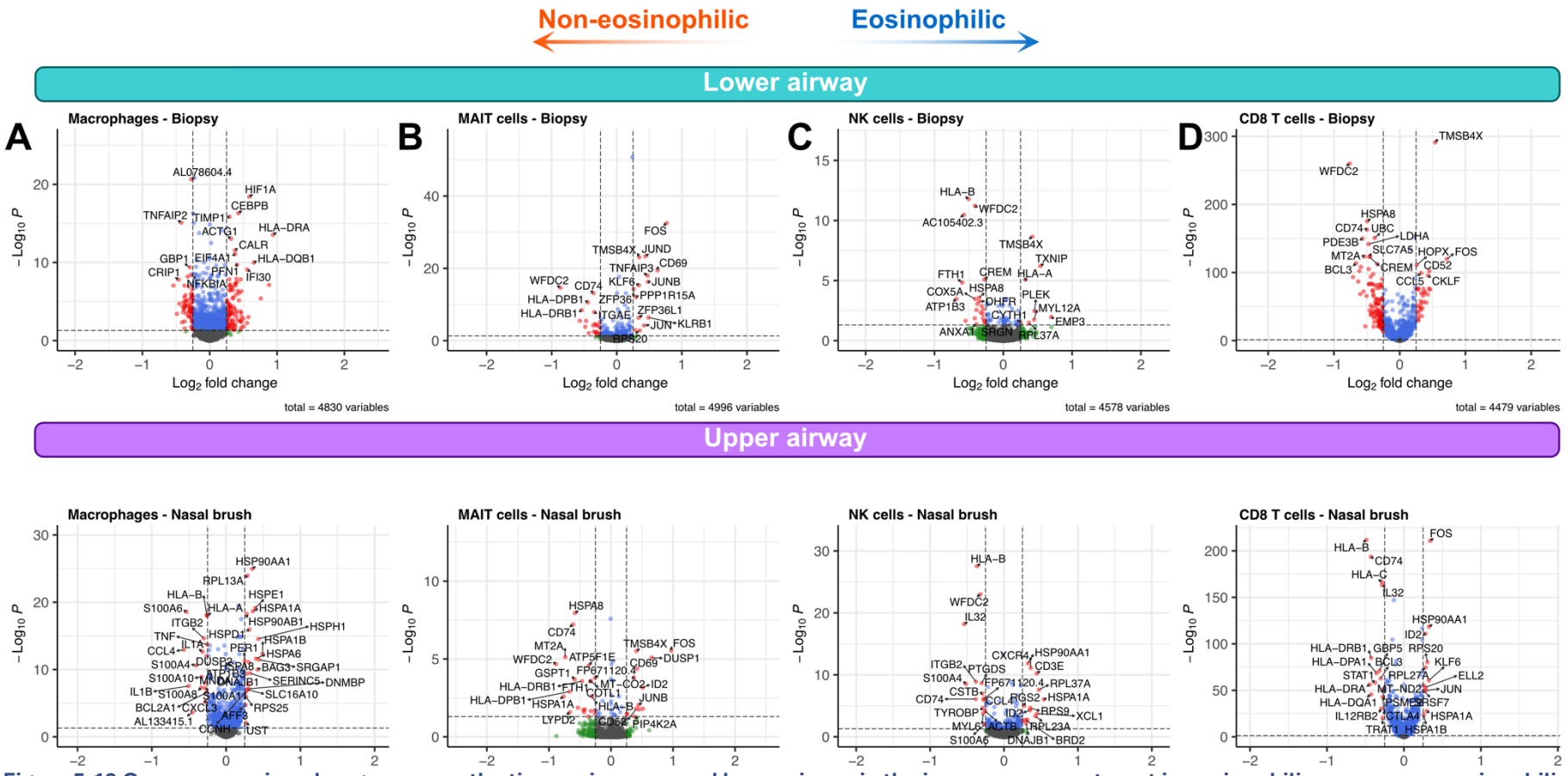

Figure 5-19 Gene expression changes across the tissues in upper and lower airway in the immune compartment in eosinophilic versus non-eosinophilic asthma

A-D Volcano plots showing gene expression changes in select immune cell relevant in neutrophilic inflammation at annotation level 3 in samples from the upper (nasal brush) and lower (biopsy) airway.

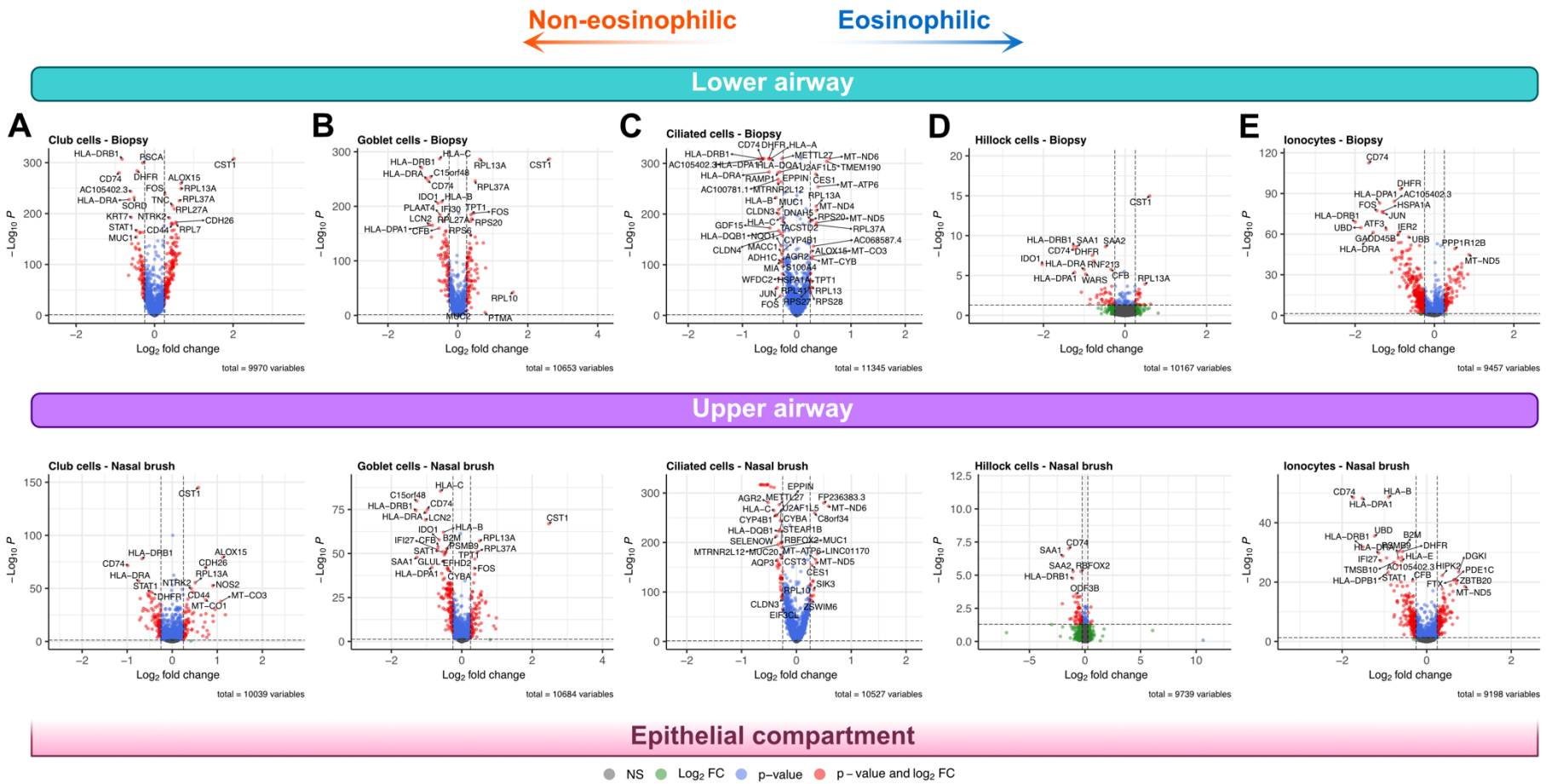

Figure 5-20 Gene expression changes across the tissues in upper and lower airway in the epithelial compartment in eosinophilic versus non-eosinophilic asthma

A-D Volcano plots showing gene expression changes in major epithelial cell types (club, goblet, and ciliated) cells and recently discovered Hillock cells and ionocytes in samples from the upper (nasal brush) and lower (biopsy) airway.

5.4.13 GENE EXPRESSION CHANGES IN EOSINOPHILIC AND NON-EOSINOPHILIC ASTHMA

As demonstrated in section 5.4.11, the overall differential gene expression profile comparing asthmatic patients to healthy controls, disregarding of phenotype, is reminiscent of eosinophilic asthma and type-2 inflammation. To better understand the differences in type-2 high and type-2 low inflammation, I contrasted gene expression in eosinophilic versus non-eosinophilic asthma.

5.4.13.1 IMMUNE COMPARTMENT

The results of differential gene expression analysis in key cell types relevant in non-eosinophilic asthma, as summarised in Figure 5-1 in the introduction to this chapter, are presented in Figure 5-19.

As seen in Figure 5-19 A, macrophages upregulated class II MHC genes *HLA-DRA* and *HLA-DQB1*. MHC II genes have been identified in GWASes with polymorphisms linked to IgE responses to common allergens [261]. In the upper airway, genes encoding heat-shock proteins are upregulated, likely related to general stress response and/or activation due to dissociation [120]. Conversely, in non-eosinophilic asthma, MHC class I gene *HLA-B* is differentially expressed in nasal samples. These differences in associations of asthma phenotypes with HLA genotypes were previously reported [262], however, they have never been linked to a cell type or molecular mechanism. In addition, in non-eosinophilic asthma, a steroid-response gene *CRIP1* was identified as a DGE in the biopsies [174], as described in chapter 3. In the upper airway, IL1-encoding genes,

IL1A and *IL1B*, were differentially expressed. IL-1 locus was previously reported to be differentially methylated in asthma [263] and to correlate with neutrophilia and inflammasome activation in asthma and is associated with disease exacerbations [264].

MAIT cells also upregulated MHC class II genes, both in biopsy and nasal brush samples in non-eosinophilic asthma (Figure 5-19 B). *WFDC2* and *CD74* were two other genes upregulated across both tissues. WFDC2 is a peptide previously has been reported to have antibacterial activity against bacteria such as *Staphylococcus aureus*, *Salmonella enterica* and *Pseudomonas aeruginosa* [265], with the later pathogen present in some patients with asthma and associated with increased number of exacerbations and poor long-term prognosis [266]. Therefore, MAIT cells may be involved in antimicrobial protection in asthma. CD74, which is also differentially expressed by other cell types such as NK cells and CD8+ T cells (Figure 5-19), serves multiple roles. It acts as a chaperone in the assembly of MHC class II molecules and functions as a receptor for the cytokine MIF. MIF has been shown to induce pro-inflammatory mediators, including TNF, while also modulating the expression of TLR4 [267], both important mechanisms in asthma.

NK cells also upregulated the previously mentioned genes (Figure 5-19 C), including MHC class I (*HLA-B*) across the sites, and *WFDC2* and *CD74* in the nasal samples. In addition, *IL32* was also upregulated in nasal brushes. IL-32 exhibits anti-inflammatory properties in animal models of asthma but its role remains disputed in humans, with reports of IL-32 inhibiting airway remodelling, but also increasing the likelihood of exacerbations and increased sera levels of pro-inflammatory, with effects hypothesised

to be dependent on asthma phenotype [268]. Thus, IL32 would be an interesting candidate for further explorations of a potential phenotype-specific therapeutic target.

CD8 cells also exhibited similar patterns (Figure 5-19 D), with upregulation of *WFDC2* and *CD74* in the lower airway and class I ands II MHC in the upper airway in non-eosinophilic asthma. Moreover, in the lower airway *PDE3B* was upregulated in biopsies, which was demonstrated contribute to airway inflammation in HDM-sensitized mouse models of asthma [269] and to regulate NLRP3 inflammasome and upregulate TNF and IL1β in adipose tissue [270]. In summary, those genes may be hypothesised to drive obesity-associated non-eosinophilic asthma subtype.

In summary, genes differentially expressed in non-eosinophilic asthma are involved in antigen processing, bacterial response, inflammasome activation and to a by extension, neutrophilic inflammation.

5.4.14 EPITHELIAL COMPARTMENT

I contrasted gene expression in common and rare cell types, similarly to section 5.4.11, but focusing on eosinophilic versus non-eosinophilic asthma, with changes summarised in Figure 5-20.

In non-eosinophilic asthma, club cells (Figure 5-20 A), in addition to upregulating previously described MHC class II genes and *CD74* in nasal brushes as well as the biopsies, also showed differential expression of *STAT1* in both of those tissues. STAT1 is a key signalling molecule involved in Th1 lymphocyte differentiation and IFNγ production

[271]. STAT1 phosphorylation induced by IFN-γ remained enhanced in the presence of corticosteroid treatment, thus mediating resistance to ICS treatment in asthma [272,273]. STAT1 signalling also regulates nitric oxide synthase *NOS2* [274], which is a marker of steroid-resistant "FeNO non-suppressor" phenotype. Given these observations, epithelium may be another important tissue mediating steroid resistance, in addition to Th cells.

Goblet cells (Figure 5-20 B) again displayed upregulation of MHC class I and II genes and *CD74*. Interestingly, across upper and lower airway, they also upregulated *LCN2* and *SAA1*, genes previously demonstrated to be blood biomarkers of Th17-mediated steroid-resistant neutrophilic asthma [275]. The expression of lipocalin-2, encoded by *LCN2*, is regulated by IL-17 [276] and dexamethasone (in other tissues) [277]. By sequestering iron, it has been implicated in resolving bacterial infections [278] and promoting lung inflammation by inducing a state of oxidative stress in macrophages in mice [279]. Serum amyloid A (SAA) are proteins encoded by a family of genes: *SAA1*, *SAA2*, *SAA3*, and *SAA4*. SAA1 and SAA2, acute-phase response proteins, promote pathogenic Th17 differentiation and IL-17A production in CD4+ T cells [280], and induce neutrophilic inflammation in the lung [281].

Ciliated cells upregulated claudins – *CLDN3* and *CLDN4* in the lower airway, and *CLDN3* in the upper airway – in non-eosinophilic asthma (Figure 5-20 C). Claudins are a large family of structural proteins that form tight junctions within the epithelial barrier across various tissues, including the lung [282,283]. Their diversity allows for the regulation of selective transport and paracellular permeability in epithelial layers [282]. Claudin 4 one

of the major claudins expressed in the lung, with increased blood plasma concentrations observed in patients experiencing asthma exacerbations and in mouse models of asthma [284]. Inhibition of claudin-4 has been shown to decrease transepithelial electrical resistance in both rats and humans [285]. Collectively, these findings suggest that claudins may play a role in barrier dysfunction in asthma.

Interestingly, Hillock cells (Figure 5-20 D) found in nasal brushes upregulated *SAA1* and *SAA2*, similarly to previously described goblet cells. This was in addition to the set of genes common across the epithelium, including *CD74* and MHC II genes. Whilst mechanistic studies would be needed to establish the relevance of this finding, it suggest SAA proteins may be used as biomarkers of non-eosinophilic asthma in the upper airway.

Ionocytes (Figure 5-20 E) were found to express previously described *STAT1*, along with MHC class II genes, *CD74*, and markers of activation (*JUN* and *FOS*). In the lower airway, ubiquitin-encoding genes *UBB* and *UBD* were upregulated. Notably, *UBD* was previously reported to be upregulated in biopsies from patients with neutrophilic asthma, along with increased IL17A signalling, and confirmed in a mouse model [286]. However, the original study did not identify the cell type of origin, whereas my analysis suggests ionocytes as a potential source of this signal.

In summary, I identify gene expression changes common across many epithelial cell types (*STAT1*, *CD74*, *HLA*) and specific to a given cell type or subset (*SAA*, *UBD*). I identify club cells as the likely source of exhaled nitric oxide in non-eosinophilic asthma and hypothesise about novel mechanisms in Th1/Th17-mediated neutrophilic asthma.

5.4.15 TYPE 2 CYTOKINE AND ALARMIN PROFILING IN EOSINOPHILIC AND NON--EOSINOPHILIC ASTHMA

Alarmins and type-2 cytokines play a critical role in the pathogenesis of eosinophilic asthma [4], as illustrated in Figure 5-1. To determine whether these molecules act through changes in cell numbers or gene expression, I profiled the expression levels and quantified cell numbers in the epithelium (for alarmins) and mast cells, Th2 cells, and granulocytes (for type-2 cytokines). The results of this analysis are presented in Figure 5-21.

All alarmins were identified except IL25, for which no transcripts were detected (Figure 5-21 A). Compared to healthy individuals and patients with non-eosinophilic asthma, a shift in the expression pattern of *IL33* was observed, with a larger proportion of basal cells expressing it in eosinophilic asthma patients. In addition, basal cells also seem to upregulate *TSLP*. In previous sections, I demonstrated number of basal cells was significantly increased in some tissues when compared to non-eosinophilic asthma and healthy controls (Figure 5-14 and Figure 5-18). These findings suggest that in the epithelium, basal cells are significant contributors to the secretion of alarmins in type-2 high asthma.

No differences in type-2 cytokine secretion were observed across the three phenotypes profiled. However, the number of mast cells increased nearly 13-fold compared to healthy controls, as seen in Figure 5-21. In rare cell types, the expression levels of *IL5* and *IL13* increased in Th2 lymphocytes, and expression of *IL4* decreased in granulocytes, but again, their numbers increased approximately 12 times compared to healthy

individuals. Thus, it seems like whilst mast cells and granulocytes increase in numbers, the Th2 cells upregulate expression of type-2 cytokines in eosinophilic asthma. These changes agree with the patterns observed in allergic eosinophilic asthma in chapter 3.

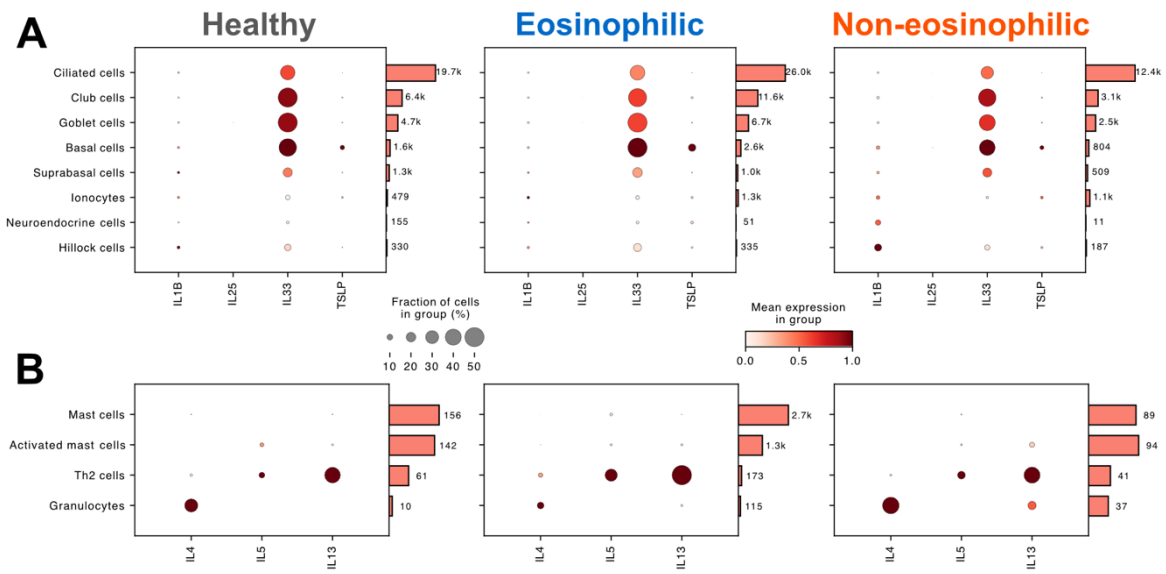

Figure 5-21 Alarmin and type-2 cytokine profiling in healthy, eosinophilic, and non-eosinophilic patients

A Dotplots showing the magnitude of expression (dot colour), the percentage of cells expressing (dot size) and the total number of cells (bar height) of **A** canonical epithelial alarmins in asthma and **B** three hallmark cytokines of type-2 inflammation (IL4, IL5, IL13) at annotation level 3.

5.5 DISCUSSION

In this chapter, I present a first of its kind study of four distinct phenotypes of severe asthma at the single-cell level, providing a dataset consisting of 220,000 cells annotated at high granularity, with 59 distinct cell states present. Clinical phenotyping, along with a full panel of blood and respiratory tests, supplement the data and provide a comprehensive resource for assessing associations between cellular and molecular signatures and clinical variables. Defined phenotyping criteria ensure clinically homogenous groups of patients within each of the four severe asthma phenotypes profiled.

Association testing demonstrated links between cell type abundance and gene expression profiles with clinical metadata, but not the asthma diagnosis or phenotype *per se*, further highlighting the likely value of utilising unbiased cellular/molecular markers to refine patient classification. This is consistent with patterns identified in phenome-wide associatcion studies, with linkage found to traits such as wheezing, the FEV_1/FVC ratio, and diagnosis of atopy (allergic rhinitis or eczema), and eosinophil counts [287].

When contrasting patients with a diagnosis of asthma to healthy individuals, both differential abundance analysis and differential gene expression profiling demonstrated changes associated with type 2-high disease, such as the upregulation of alarmins and an increase in mast cell abundance. Some of these changes were restricted to a single tissue site (upper or lower airway). The signalling landscape consisted of pro-

inflammatory mediators such as TGFβ, with some suggestion that epithelial-mesenchymal transition, with signalling contributions from both immune and non-immune cells, may mediate fibrotic processes in asthma.

In comparison, when contrasting type 2-high with type 2-low asthma (eosinophilic versus non-eosinophilic asthma), changes characteristic of Th2-mediated and Th17-mediated disease begin to emerge, with shifts in MAIT and CD4+ T cells and some epithelial cell subsets in the non-eosinophilic compartment. This demonstrates that detailed phenotyping in necessary to understand the varied molecular mechanisms in asthma, and that the changes in type 2-high inflammation are likely most pronounced and can thus obscure other differences when considering mixed cohorts of asthma patients. As the sample collection for this dataset is ongoing, even more detailed comparisons will be possible in the future, such as neutrophilic versus other non-eosinophilic asthma types. Indeed, as patients have the FeNO measured as part of the clinical panel, comparison between steroid treatment responders ('FeNO suppressors') and non-responders ('FeNO non-suppressors') would be of particular interest. In the future, it would be exciting to better understand the mechanisms of action of biologic treatments on the airway at single-cell resolution, in addition to steroid response.

Some phenotypes of non-eosinophilic asthma are defined by neutrophilic inflammation. Neutrophils were not identified in this dataset, which is a limitation. As the non-immune cells were profiled along the immune cells in this study, neutrophil numbers in the sample were likely too low for detection (further compounded by the poor neutrophil capture rate by microfluidic technologies). Whilst the neutrophils were not identified,

features characteristic of neutrophilic inflammation, such as upregulation of genes involved in antimicrobial response and NLRP3 inflammasome assembly, provide indirect clues that neutrophils are involved in inflammation in non-eosinophilic asthma.

In addition to mapping out potentially novel mechanisms of inflammation in asthma, my analysis highlights the likely importance of rare epithelial cell types, such as the recently identified Hillock cells, in the pathogenesis of disease, as is the case in non-eosinophilic asthma, where those cells express genes encoding acute-phase proteins SAA1 and SAA2. As the existence of those cells was not appreciated until very recently, being able to capture them in this study opens new avenues of research to understand their involvement in disease.

In summary, this dataset, unique in terms of the patient phenotypes, the size, and the number of modalities collected, will enable detailed profiling of molecular mechanisms of asthma. Its on-going expansion by a further 20 patients, resulting in a full dataset of 60 individuals, 12 of each phenotype, along with 12 healthy controls, will enable robust comparisons across phenotypes. The existing annotation is a precious resource, as the samples are collected from multiple tissues of patients with varied asthma phenotypes and are not limited to healthy individuals or to a single lung tissue site, which hopefully would lead to improved reference mapping results, and remove the need for time-intensive manual annotation.

6 DISCUSSION

In this thesis, I present three large atlasing resources to study asthma at single-cell resolution. Together, they encompass over 400,000 cells, with detailed cell –state-level annotation, and rich clinical metadata. I demonstrate that new insights about molecular mechanisms of inflammation can be obtained by comparing and contrasting distinct asthma phenotypes - describing rare epithelial and inflammatory cell types, as well as epithelial structures, and expanding the profiling of asthma beyond the changes in gene expression by adding new modalities, antibody binding profiles, and immune receptor repertoires.

In Chapter 3, I introduce a novel method to profile aeroallergens in asthma by implementing direct conjugation with DNA oligonucleotide barcodes. With this tool, I demonstrate that the binding profiles of two major allergens described in HDM sensitisation, *DerP* 1 and *DerP*2 , show differential binding profiles, consistent with their different molecular functions as a protease and a TLR agonist, respectively. I demonstrate that whilst IgE-producing plasmablasts are not present in the airway in the state of controlled disease, various cell types binding IgE, likely through FcεRI surface receptor, including mast cells and basophils, remain present in the airway. I also more broadly profile the three common sampling sites in human airways – BAL, nasal brushings, and biopsies – and show that there are profound differences in their cellular composition, an observation important when designing bronchoscopic studies of disease. I contrast gene expression profiles of immune cells, revealing features of type 2 inflammation, TGFβ signalling, and enrichment of pathways relating to EMT, pointing to

possible sources of fibrosis. Finally, I demonstrate using ligand-receptor interaction analysis that airway macrophages may play an important role in inflammation as they signal to multiple other immune cell types.

In Chapter 4 I expand the immune cell profiling carried out in Chapter 3 by profiling biopsies of the same or phenotype-matched patients using *in situ* spatial transcriptomics. I present a first-of-a-kind dataset capturing epithelial, immune, stromal, and vascular cells. Using pseudotime analysis I recapitulate differentiation trajectories of epithelial cell types, profiling potential goblet cell plasticity, as evidenced by high pseudotime entropy. I describe immune niches which seem to only be present in patients with allergic asthma and are consistent with the mechanisms of inflammation – a niche of B and CD4+ T lymphocytes which may arise from the antigen presentation or co-stimulation of B cells in the lung, and a mast cell-plasma niche, also enriched in CD8+ T cells, which could suggest low-level expression of IgE and its binding to the surface of mast cells, or that IgE signalling is absent in steady state but close co-localisation enables rapid response upon allergen stimulation.

Two major phenotypes of asthma are driven by neutrophilic and eosinophilic inflammation, yet the datasets I generated are sparse in those important granulocytes. My datasets are not an isolated case, capture of granulocytes has proven to pose difficulties in single-cell research, likely due to their fragile nature. In general, gravity-based microwell technologies have shown better rates of in capturing granulocytes compared to microfluidics-based platforms [197].

Neutrophils are commonly excluded at quality-control stages due to low cell counts, or get lost during sample handling, as activated neutrophils attach to laboratory ware and wells of microfluidic chips. My data suggest that neutrophils can be captured at greater rates when immune cells are sorted from the samples, which could simply be due to the sort increasing their relative concentration in the input material.

Whilst eosinophils can be captured from the blood with relative ease, they have not yet been described in meaningful numbers (beyond a few isolated cells) in solid tissues, however, they can readily be captured in the lungs of patients with asthma using traditional techniques, such as immunohistochemistry [288]. The gene panel used in my study had limited power to discriminate eosinophils, however, I did not observe any *IL5*-expressing cells in my samples. It could be that those fragile cells get destroyed during FFPE preservation, thus, it can be hypothesised that those cells can perhaps be identified using spatially-resolved profiling techniques utilising tissue preserved by snap-freezing.

Having extensively profiled allergic asthma across omics platforms and modalities, in Chapter 5 I expand my analysis to three other asthma phenotypes. When comparing asthmatic individuals to healthy controls, I observed an increase in the abundance of mast cells and upregulation of genes associated with type 2 inflammation, including upregulation of nitric oxide synthase, an enzyme responsible for production of NO detected in breath of asthmatic patients [50]. The pathway analysis, similarly to Chapter 3, revealed a landscape of inflammation primarily mediated by TGFβ, with evidence of EMT signalling. Although only approximately half of the asthmatic patients in this comparison

had a diagnosis of eosinophilic asthma, most of the changes identified could be associated with eosinophilic inflammation, and resembled those identified in Chapter 3. This could be because patients with type 2 inflammation show the most extreme patterns in gene expression and abundance changes, obscuring the smaller changes present in individuals with non-eosinophilic inflammation.

Indeed, contrasting eosinophilic patients with non-eosinophilic individuals revealed patterns of neutrophilic inflammation, including IL-1 and IL-32 signalling, inflammasome assembly and upregulation of genes involved in response to bacterial pathogens. When the full dataset is collected for the broader study described in Chapter 5, more detailed comparisons, including between patients responsive and resistant to steroid treatment, will become possible, highlighting the value of detailed phenotyping in the single-cell atlas of asthma that I have created.

Interestingly, rare cell types in the epithelium seemed to be involved in inflammation and airway remodelling in asthma, such as Hillock cells, which express serum amyloid A proteins, which promote pathogenic Th17 differentiation. The advent of single-cell technologies enabled us to first identify and now profile those previously unknown cell types. Another rare cell type, the ionocyte, was implicated in the pathogenesis of cystic fibrosis [86,88], suggesting it may play an important role in disease pathogenesis, even if the ionocyte relative abundance in the epithelium is low.

The studies described in this thesis have limitations. Technical caveats include low number of cells and immune repertoire sequences acquired in Chapter 3 and thus limited discovery power of the study, possible activation signatures due to enzymatic dissociation detected in cells annotated in Chapter 3 and 5, and technical artefacts introduced by tissue sectioning in Chapter 4. Chapter 3 contains an experimental group consisting of a single patient, future studies could be improved by increasing the number of patients and better matching them to healthy controls in terms of age, sex, and body weight. The number of immune receptor sequences captured could be improved by including paired blood samples in the study design, as the number of required immune receptor sequences is likely on the order of tens of thousands. Importantly, the granulocytes, neutrophils and eosinophils, which are key in driving neutrophilic and eosinophilic/type-2 high asthma, respectively, are only captured in low numbers (neutrophils) or not at all (eosinophils) by current single cell and spatial omics techniques. As they can be visualised using conventional microscopy approaches, adding staining specific for those cell types in Chapter 4 could help us understand why thir transcriptomes are not being captured.

Despite the caveats, together with other datasets I am working to profile asthma at molecular level, including single-cell ATAC data and whole genome profiling, the large omics datasets described in this thesis have the potential to transform our understanding of severe asthma phenotypes, and improve the treatment.

BIBLIOGRAPHY

[1] Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 2020;396:1204–22. https://doi.org/10.1016/S0140-6736(20)30925-9.

[2] Mukherjee M, Stoddart A, Gupta RP, Nwaru BI, Farr A, Heaven M, et al. The epidemiology, healthcare and societal burden and costs of asthma in the UK and its member nations: analyses of standalone and linked national databases. BMC Med 2016;14:113. https://doi.org/10.1186/s12916-016-0657-8.

[3] Lambrecht BN, Hammad H. The immunology of asthma. Nat Immunol 2015;16:45–56. https://doi.org/10.1038/ni.3049.

[4] Hammad H, Lambrecht BN. The basic immunology of asthma. Cell 2021;184:1469–85. https://doi.org/10.1016/j.cell.2021.02.016.

[5] Settipane RA, Kreindler JL, Chung Y, Tkacz J. Evaluating direct costs and productivity losses of patients with asthma receiving GINA 4/5 therapy in the United States. Annals of Allergy, Asthma & Immunology 2019;123:564-572.e3. https://doi.org/10.1016/j.anai.2019.08.462.

[6] Smith AD, Cowan JO, Brassett KP, Filsell S, McLachlan C, Monti-Sheehan G, et al. Exhaled Nitric Oxide: A Predictor of Steroid Response. Am J Respir Crit Care Med 2005;172:453–9. https://doi.org/10.1164/rccm.200411-1498OC.

[7] Brusselle GG, Koppelman GH. Biologic Therapies for Severe Asthma. N Engl J Med 2022;386:157–71. https://doi.org/10.1056/NEJMra2032506.

[8] Holgate ST, Wenzel S, Postma DS, Weiss ST, Renz H, Sly PD. Asthma. Nat Rev Dis Primers 2015;1:15025. https://doi.org/10.1038/nrdp.2015.25.

[9] Eder W, Ege MJ, Von Mutius E. The Asthma Epidemic. N Engl J Med 2006;355:2226–35. https://doi.org/10.1056/NEJMra054308.

[10] Cao Y, Chen S, Chen X, Zou W, Liu Z, Wu Y, et al. Global trends in the incidence and mortality of asthma from 1990 to 2019: An age-period-cohort analysis using the global burden of disease study 2019. Front Public Health 2022;10:1036674. https://doi.org/10.3389/fpubh.2022.1036674.

[11] Lambrecht BN, Hammad H. The airway epithelium in asthma. Nat Med 2012;18:684–92. https://doi.org/10.1038/nm.2737.

[12] Guarnieri M, Balmes JR. Outdoor air pollution and asthma. The Lancet 2014;383:1581–92. https://doi.org/10.1016/S0140-6736(14)60617-6.

[13] Matsui EC, Wood RA, Rand C, Kanchanaraksa S, Swartz L, Eggleston PA. Mouse allergen exposure and mouse skin test sensitivity in suburban, middle-class children with asthma☆. Journal of Allergy and Clinical Immunology 2004;113:910–5. https://doi.org/10.1016/j.jaci.2004.02.034.

[14] Byrd RS, Joad JP. Urban asthma. Current Opinion in Pulmonary Medicine 2006;12:68–74. https://doi.org/10.1097/01.mcp.0000199001.68908.45.

[15] Han A, Deng S, Yu J, Zhang Y, Jalaludin B, Huang C. Asthma triggered by extreme temperatures: From epidemiological evidence to biological plausibility. Environmental Research 2023;216:114489. https://doi.org/10.1016/j.envres.2022.114489.

[16] Xu Q, Zhou Q, Chen J, Li T, Ma J, Du R, et al. The incidence of asthma attributable to temperature variability: An ecological study based on 1990–2019 GBD data. Science of The Total Environment 2023;904:166726. https://doi.org/10.1016/j.scitotenv.2023.166726.

[17] The Lancet. A plea to abandon asthma as a disease concept. The Lancet 2006;368:705. https://doi.org/10.1016/S0140-6736(06)69257-X.

[18] Wenzel SE. Asthma phenotypes: the evolution from clinical to molecular approaches. Nat Med 2012;18:716–25. https://doi.org/10.1038/nm.2678.

[19] Kuruvilla ME, Lee FE-H, Lee GB. Understanding Asthma Phenotypes, Endotypes, and Mechanisms of Disease. Clinic Rev Allerg Immunol 2019;56:219–33. https://doi.org/10.1007/s12016-018-8712-1.

[20] Papi A, Brightling C, Pedersen SE, Reddel HK. Asthma. The Lancet 2018;391:783–800. https://doi.org/10.1016/S0140-6736(17)33311-1.

[21] Wenzel SE, Schwartz LB, Langmack EL, Halliday JL, Trudeau JB, Gibbs RL, et al. Evidence That Severe Asthma Can Be Divided Pathologically into Two Inflammatory Subtypes with Distinct Physiologic and Clinical Characteristics. Am J Respir Crit Care Med 1999;160:1001–8. https://doi.org/10.1164/ajrccm.160.3.9812110.

[22] Shrine N, Izquierdo AG, Chen J, Packer R, Hall RJ, Guyatt AL, et al. Multi-ancestry genome-wide association analyses improve resolution of genes and pathways influencing lung function and chronic obstructive pulmonary disease risk. Nat Genet 2023;55:410–22. https://doi.org/10.1038/s41588-023-01314-0.

[23] Jayapal M, Tay HK, Reghunathan R, Zhi L, Chow KK, Rauff M, et al. Genome-wide gene expression profiling of human mast cells stimulated by IgE or FcεRI-aggregation reveals a complex network of genes involved in inflammatory responses. BMC Genomics 2006;7:210. https://doi.org/10.1186/1471-2164-7-210.

[24] Han Y, Jia Q, Jahani PS, Hurrell BP, Pan C, Huang P, et al. Genome-wide analysis highlights contribution of immune system pathways to the genetic architecture of

asthma. Nat Commun 2020;11:1776. https://doi.org/10.1038/s41467-020-15649-3.

[25] Ober C, Yao T. The genetics of asthma and allergic disease: a 21st century perspective. Immunological Reviews 2011;242:10–30. https://doi.org/10.1111/j.1600-065X.2011.01029.x.

[26] Skadhauge L, Christensen K, Kyvik K, Sigsgaard T. Genetic and environmental influence on asthma: a population-based study of 11,688 Danish twin pairs. Eur Respir J 1999;13:8–14. https://doi.org/10.1183/09031936.99.13100899.

[27] Duffy DL, Martin NG, Battistutta D, Hopper JL, Mathews JD. Genetics of Asthma and Hay Fever in Australian Twins. Am Rev Respir Dis 1990;142:1351–8. https://doi.org/10.1164/ajrccm/142.6_Pt_1.1351.

[28] Laitinen T, Räsänen M, Kaprio J, Koskenvuo M, Laitinen LA. Importance of Genetic Factors in Adolescent Asthma: A Population-based Twin-family Study. Am J Respir Crit Care Med 1998;157:1073–8. https://doi.org/10.1164/ajrccm.157.4.9704041.

[29] Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 2007;448:470–3. https://doi.org/10.1038/nature06014.

[30] James B, Milstien S, Spiegel S. ORMDL3 and allergic asthma: From physiology to pathology. Journal of Allergy and Clinical Immunology 2019;144:634–40. https://doi.org/10.1016/j.jaci.2019.07.023.

[31] Das S, Miller M, Beppu AK, Mueller J, McGeough MD, Vuong C, et al. GSDMB induces an asthma phenotype characterized by increased airway responsiveness and remodeling without lung inflammation. Proc Natl Acad Sci USA 2016;113:13132–7. https://doi.org/10.1073/pnas.1610433113.

[32] Stein MM, Thompson EE, Schoettler N, Helling BA, Magnaye KM, Stanhope C, et al. A decade of research on the 17q12-21 asthma locus: Piecing together the puzzle. J Allergy Clin Immunol 2018;142:749-764.e3. https://doi.org/10.1016/j.jaci.2017.12.974.

[33] Bochkov YA, Watters K, Ashraf S, Griggs TF, Devries MK, Jackson DJ, et al. Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication. Proc Natl Acad Sci USA 2015;112:5485–90. https://doi.org/10.1073/pnas.1421178112.

[34] Tsuo K, Zhou W, Wang Y, Kanai M, Namba S, Gupta R, et al. Multi-ancestry meta-analysis of asthma identifies novel associations and highlights the value of increased power and diversity. Cell Genomics 2022;2:100212. https://doi.org/10.1016/j.xgen.2022.100212.

[35] Sayers I, John C, Chen J, Hall IP. Genetics of chronic respiratory disease. Nat Rev Genet 2024;25:534–47. https://doi.org/10.1038/s41576-024-00695-0.

[36] Fu Y, Wang J, Panangipalli G, Ulrich BJ, Koh B, Xu C, et al. STAT5 promotes accessibility and is required for BATF-mediated plasticity at the Il9 locus. Nat Commun 2020;11:4882. https://doi.org/10.1038/s41467-020-18648-6.

[37] Young HWJ, Williams OW, Chandra D, Bellinghausen LK, Pérez G, Suárez A, et al. Central Role of Muc5ac Expression in Mucous Metaplasia and Its Regulation by Conserved 5' Elements. Am J Respir Cell Mol Biol 2007;37:273–90. https://doi.org/10.1165/rcmb.2005-0460OC.

[38] Pejler G. The emerging role of mast cell proteases in asthma. Eur Respir J 2019;54:1900685. https://doi.org/10.1183/13993003.00685-2019.

[39] Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A Large-Scale, Consortium-Based Genomewide Association Study of Asthma. N Engl J Med 2010;363:1211–21. https://doi.org/10.1056/NEJMoa0906312.

[40] Schoettler N, Rodríguez E, Weidinger S, Ober C. Advances in asthma and allergic disease genetics: Is bigger always better? Journal of Allergy and Clinical Immunology 2019;144:1495–506. https://doi.org/10.1016/j.jaci.2019.10.023.

[41] Ferreira MAR, Mathur R, Vonk JM, Szwajda A, Brumpton B, Granell R, et al. Genetic Architectures of Childhood- and Adult-Onset Asthma Are Partly Distinct. The American Journal of Human Genetics 2019;104:665–84. https://doi.org/10.1016/j.ajhg.2019.02.022.

[42] Bønnelykke K, Sleiman P, Nielsen K, Kreiner-Møller E, Mercader JM, Belgrave D, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. Nat Genet 2014;46:51–5. https://doi.org/10.1038/ng.2830.

[43] Rackemann FM. A working classification of asthma. The American Journal of Medicine 1947;3:601–6. https://doi.org/10.1016/0002-9343(47)90204-0.

[44] Humbert M, Durham SR, Ying S, Kimmitt P, Barkans J, Assoufi B, et al. IL-4 and IL-5 mRNA and protein in bronchial biopsies from patients with atopic and nonatopic asthma: evidence against “intrinsic” asthma being a distinct immunopathologic entity. Am J Respir Crit Care Med 1996;154:1497–504. https://doi.org/10.1164/ajrccm.154.5.8912771.

[45] Wills-Karp M, Luyimbazi J, Xu X, Schofield B, Neben TY, Karp CL, et al. Interleukin-13: Central Mediator of Allergic Asthma. Science 1998;282:2258–61. https://doi.org/10.1126/science.282.5397.2258.

[46] Green RH, Brightling CE, McKenna S, Hargadon B, Parker D, Bradding P, et al. Asthma exacerbations and sputum eosinophil counts: a randomised controlled trial. The Lancet 2002;360:1715–21. https://doi.org/10.1016/S0140-6736(02)11679-5.

[47] Telenga ED, Tideman SW, Kerstjens HAM, Hacken NHTT, Timens W, Postma DS, et al. Obesity in asthma: more neutrophilic inflammation as a possible explanation for a reduced treatment response. Allergy 2012;67:1060–8. https://doi.org/10.1111/j.1398-9995.2012.02855.x.

[48] Rastogi D, Fraser S, Oh J, Huber AM, Schulman Y, Bhagtani RH, et al. Inflammation, Metabolic Dysregulation, and Pulmonary Function among Obese Urban Adolescents with Asthma. Am J Respir Crit Care Med 2015;191:149–60. https://doi.org/10.1164/rccm.201409-1587OC.

[49] Wu AY, Peebles RS. The GLP-1 receptor in airway inflammation in asthma: a promising novel target? Expert Review of Clinical Immunology 2021;17:1053–7. https://doi.org/10.1080/1744666X.2021.1971973.

[50] Rupani H, Kent BD. Using Fractional Exhaled Nitric Oxide Measurement in Clinical Asthma Management. Chest 2022;161:906–17. https://doi.org/10.1016/j.chest.2021.10.015.

[51] Butler CA, Heaney LG. Fractional exhaled nitric oxide and asthma treatment adherence. Current Opinion in Allergy & Clinical Immunology 2021;21:59–64. https://doi.org/10.1097/ACI.0000000000000704.

[52] Couto M, Kurowski M, Moreira A, Bullens DMA, Carlsen K -H., Delgado L, et al. Mechanisms of exercise-induced bronchoconstriction in athletes: Current perspectives and future challenges. Allergy 2018;73:8–16. https://doi.org/10.1111/all.13224.

[53] Karjalainen E-M, Laitinen A, Sue-Chu M, Altraja A, Bjermer L, Laitinen LA. Evidence of Airway Inflammation and Remodeling in Ski Athletes with and without Bronchial Hyperresponsiveness to Methacholine. Am J Respir Crit Care Med 2000;161:2086–91. https://doi.org/10.1164/ajrccm.161.6.9907025.

[54] Porsbjerg C, Lund TK, Pedersen L, Backer V. Inflammatory Subtypes in Asthma are Related to Airway Hyperresponsiveness to Mannitol and Exhaled NO. Journal of Asthma 2009;46:606–12. https://doi.org/10.1080/02770900903015654.

[55] Rasmussen SM, Hansen ESH, Backer V. Asthma in elite athletes – do they have Type 2 or non-Type 2 disease? A new insight on the endotypes among elite athletes. Front Allergy 2022;3:973004. https://doi.org/10.3389/falgy.2022.973004.

[56] The ENFUMOSA Study Group. The ENFUMOSA cross-sectional European multicentre study of the clinical phenotype of chronic severe asthma. Eur Respir J 2003;22:470–7. https://doi.org/10.1183/09031936.03.00261903.

[57] Peters U, Dixon AE, Forno E. Obesity and asthma. Journal of Allergy and Clinical Immunology 2018;141:1169–79. https://doi.org/10.1016/j.jaci.2018.02.004.

[58] Dixon AE, Pratley RE, Forgione PM, Kaminsky DA, Whittaker-Leclair LA, Griffes LA, et al. Effects of obesity and bariatric surgery on airway hyperresponsiveness,

asthma control, and inflammation. Journal of Allergy and Clinical Immunology 2011;128:508-515.e2. https://doi.org/10.1016/j.jaci.2011.06.009.

[59] Szefler SJ, Wenzel S, Brown R, Erzurum SC, Fahy JV, Hamilton RG, et al. Asthma outcomes: Biomarkers. Journal of Allergy and Clinical Immunology 2012;129:S9–23. https://doi.org/10.1016/j.jaci.2011.12.979.

[60] Wahn U, Martin C, Freeman P, Blogg M, Jimenez P. Relationship between pretreatment specific IgE and the response to omalizumab therapy. Allergy 2009;64:1780–7. https://doi.org/10.1111/j.1398-9995.2009.02119.x.

[61] Diamant Z, Vijverberg S, Alving K, Bakirtas A, Bjermer L, Custovic A, et al. Toward clinically applicable biomarkers for asthma: An EAACI position paper. Allergy 2019;74:1835–51. https://doi.org/10.1111/all.13806.

[62] Burrows B, Martinez FD, Halonen M, Barbee RA, Cline MG. Association of Asthma with Serum IgE Levels and Skin-Test Reactivity to Allergens. N Engl J Med 1989;320:271–7. https://doi.org/10.1056/NEJM198902023200502.

[63] Simpson A, Soderstrom L, Ahlstedt S, Murray CS, Woodcock A, Custovic A. IgE antibody quantification and the probability of wheeze in preschool children. Journal of Allergy and Clinical Immunology 2005;116:744–9. https://doi.org/10.1016/j.jaci.2005.06.032.

[64] Ellis AG. The pathological anatomy of bronchial asthma. The American Journal of the Medical Sciences 1908;136:407–28. https://doi.org/10.1097/00000441-190809000-00009.

[65] Wenzel SE. Eosinophils in Asthma — Closing the Loop or Opening the Door? N Engl J Med 2009;360:1026–8. https://doi.org/10.1056/NEJMe0900334.

[66] Krause JR, Boggs DR. Search for eosinopenia in hospitalized patients with normal blood leukocyte concentration. American J Hematol 1987;24:55–63. https://doi.org/10.1002/ajh.2830240108.

[67] Mepolizumab for treating severe eosinophilic asthma. National Institute for Health and Care Excellence; 2021.

[68] Dweik RA, Boggs PB, Erzurum SC, Irvin CG, Leigh MW, Lundberg JO, et al. An Official ATS Clinical Practice Guideline: Interpretation of Exhaled Nitric Oxide Levels (F_{ENO}) for Clinical Applications. Am J Respir Crit Care Med 2011;184:602–15. https://doi.org/10.1164/rccm.9120-11ST.

[69] Knox-Brown B, Mulhern O, Feary J, Amaral AFS. Spirometry parameters used to define small airways obstruction in population-based studies: systematic review. Respir Res 2022;23:67. https://doi.org/10.1186/s12931-022-01990-2.

[70] Dunican EM, Fahy JV. Asthma and corticosteroids: time for a more precise approach to treatment. Eur Respir J 2017;49:1701167. https://doi.org/10.1183/13993003.01167-2017.

[71] Bleecker ER, Menzies-Gow AN, Price DB, Bourdin A, Sweet S, Martin AL, et al. Systematic Literature Review of Systemic Corticosteroid Use for Asthma Management. Am J Respir Crit Care Med 2020;201:276–93. https://doi.org/10.1164/rccm.201904-0903SO.

[72] Liu S, Cao Y, Du T, Zhi Y. Prevalence of Comorbid Asthma and Related Outcomes in COVID-19: A Systematic Review and Meta-Analysis. The Journal of Allergy and Clinical Immunology: In Practice 2021;9:693–701. https://doi.org/10.1016/j.jaip.2020.11.054.

[73] Williamson EJ, Walker AJ, Bhaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature 2020;584:430–6. https://doi.org/10.1038/s41586-020-2521-4.

[74] Zepp JA, Morrisey EE. Cellular crosstalk in the development and regeneration of the respiratory system. Nat Rev Mol Cell Biol 2019;20:551–66. https://doi.org/10.1038/s41580-019-0141-3.

[75] Mercer RR, Russell ML, Roggli VL, Crapo JD. Cell number and distribution in human and rat airways. Am J Respir Cell Mol Biol 1994;10:613–24. https://doi.org/10.1165/ajrcmb.10.6.8003339.

[76] Montoro DT, Haber AL, Rood JE, Regev A, Rajagopal J. A Synthesis Concerning Conservation and Divergence of Cell Types across Epithelia. Cold Spring Harb Perspect Biol 2020;12:a035733. https://doi.org/10.1101/cshperspect.a035733.

[77] Deprez M, Zaragosi L-E, Truchi M, Becavin C, Ruiz García S, Arguel M-J, et al. A Single-Cell Atlas of the Human Healthy Airways. Am J Respir Crit Care Med 2020;202:1636–45. https://doi.org/10.1164/rccm.201911-2199OC.

[78] Sikkema L, Ramírez-Suástegui C, Strobl DC, Gillett TE, Zappia L, Madissoon E, et al. An integrated cell atlas of the lung in health and disease. Nat Med 2023;29:1563–77. https://doi.org/10.1038/s41591-023-02327-2.

[79] Hewitt RJ, Lloyd CM. Regulation of immune responses by the airway epithelial cell landscape. Nat Rev Immunol 2021;21:347–62. https://doi.org/10.1038/s41577-020-00477-9.

[80] Russell RJ, Boulet L-P, Brightling CE, Pavord ID, Porsbjerg C, Dorscheid D, et al. The airway epithelium: an orchestrator of inflammation, a key structural barrier and a therapeutic target in severe asthma. Eur Respir J 2024;63:2301397. https://doi.org/10.1183/13993003.01397-2023.

[81] Blackburn JB, Li NF, Bartlett NW, Richmond BW. An update in club cell biology and its potential relevance to chronic obstructive pulmonary disease. American Journal

of Physiology-Lung Cellular and Molecular Physiology 2023;324:L652–65. https://doi.org/10.1152/ajplung.00192.2022.

[82] Fujino N, Brand OJ, Morgan DJ, Fujimori T, Grabiec AM, Jagger CP, et al. Sensing of apoptotic cells through Axl causes lung basal cell proliferation in inflammatory diseases. Journal of Experimental Medicine 2019;216:2184–201. https://doi.org/10.1084/jem.20171978.

[83] Rock JR, Randell SH, Hogan BLM. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. Disease Models & Mechanisms 2010;3:545–56. https://doi.org/10.1242/dmm.006031.

[84] Zacharias WJ, Frank DB, Zepp JA, Morley MP, Alkhaleel FA, Kong J, et al. Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. Nature 2018;555:251–5. https://doi.org/10.1038/nature25786.

[85] Varma R, Soleas JP, Waddell TK, Karoubi G, McGuigan AP. Current strategies and opportunities to manufacture cells for modeling human lungs. Advanced Drug Delivery Reviews 2020;161–162:90–109. https://doi.org/10.1016/j.addr.2020.08.005.

[86] Plasschaert LW, Žilionis R, Choo-Wing R, Savova V, Knehr J, Roma G, et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. Nature 2018;560:377–81. https://doi.org/10.1038/s41586-018-0394-6.

[87] Ruiz García S, Deprez M, Lebrigand K, Cavard A, Paquet A, Arguel M-J, et al. Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. Development 2019;146:dev177428. https://doi.org/10.1242/dev.177428.

[88] Montoro DT, Haber AL, Biton M, Vinarsky V, Lin B, Birket SE, et al. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature 2018;560:319–24. https://doi.org/10.1038/s41586-018-0393-7.

[89] Branchfield K, Nantie L, Verheyden JM, Sui P, Wienhold MD, Sun X. Pulmonary neuroendocrine cells function as airway sensors to control lung immune response. Science 2016;351:707–10. https://doi.org/10.1126/science.aad7969.

[90] Curtis JL. Cell-mediated adaptive immune defense of the lungs. Proc Am Thorac Soc 2005;2:412–6. https://doi.org/10.1513/pats.200507-070JS.

[91] Sabatel C, Bureau F. The innate immune brakes of the lung. Front Immunol 2023;14:1111298. https://doi.org/10.3389/fimmu.2023.1111298.

[92] Medzhitov R, Janeway CA. Innate Immunity: The Virtues of a Nonclonal System of Recognition. Cell 1997;91:295–8. https://doi.org/10.1016/s0092-8674(00)80412-2.

[93] Brandtzaeg P, Farstad IN, Johansen FE, Morton HC, Norderhaug IN, Yamanaka T. The B-cell system of human mucosae and exocrine glands. Immunol Rev 1999;171:45–87. https://doi.org/10.1111/j.1600-065x.1999.tb01342.x.

[94] Eguiluz-Gracia I, Layhadi JA, Rondon C, Shamji MH. Mucosal IgE immune responses in respiratory diseases. Curr Opin Pharmacol 2019;46:100–7. https://doi.org/10.1016/j.coph.2019.05.009.

[95] Desmedt M, Rottiers P, Dooms H, Fiers W, Grooten J. Macrophages Induce Cellular Immunity by Activating Th1 Cell Responses and Suppressing Th2 Cell Responses. The Journal of Immunology 1998;160:5300–8. https://doi.org/10.4049/jimmunol.160.11.5300.

[96] Mathie SA, Dixon KL, Walker SA, Tyrrell V, Mondhe M, O'Donnell VB, et al. Alveolar macrophages are sentinels of murine pulmonary homeostasis following inhaled antigen challenge. Allergy 2015;70:80–9. https://doi.org/10.1111/all.12536.

[97] Thepen T, McMENAMIN C, Girn B, Kraal G, Holt PG. Regulation of IgE production in pre-sensitized animals: *in vivo* elimination of alveolar macrophages preferentially increases IgE responses to inhaled allergen*. Clin Experimental Allergy 1992;22:1107–14. https://doi.org/10.1111/j.1365-2222.1992.tb00137.x.

[98] Sabatel C, Radermecker C, Fievez L, Paulissen G, Chakarov S, Fernandes C, et al. Exposure to Bacterial CpG DNA Protects from Airway Allergic Inflammation by Expanding Regulatory Lung Interstitial Macrophages. Immunity 2017;46:457–73. https://doi.org/10.1016/j.immuni.2017.02.016.

[99] Kawano H, Kayama H, Nakama T, Hashimoto T, Umemoto E, Takeda K. IL-10-producing lung interstitial macrophages prevent neutrophilic asthma. INTIMM 2016;28:489–501. https://doi.org/10.1093/intimm/dxw012.

[100] Dendritic cell subsets and locations. International Review of Cell and Molecular Biology, Elsevier; 2019, p. 1–68. https://doi.org/10.1016/bs.ircmb.2019.07.004.

[101] Reizis B. Plasmacytoid Dendritic Cells: Development, Regulation, and Function. Immunity 2019;50:37–50. https://doi.org/10.1016/j.immuni.2018.12.027.

[102] Silva-Sanchez A, Randall TD. Role of iBALT in Respiratory Immunity. Curr Top Microbiol Immunol 2020;426:21–43. https://doi.org/10.1007/82_2019_191.

[103] Naderi W, Schreiner D, King CG. T-cell–B-cell collaboration in the lung. Current Opinion in Immunology 2023;81:102284. https://doi.org/10.1016/j.coi.2023.102284.

[104] Hwang JY, Silva-Sanchez A, Carragher DM, Garcia-Hernandez MDLL, Rangel–Moreno J, Randall TD. Inducible Bronchus–Associated Lymphoid Tissue (iBALT) Attenuates Pulmonary Pathology in a Mouse Model of Allergic Airway Disease. Front Immunol 2020;11. https://doi.org/10.3389/fimmu.2020.570661.

[105] Miyake K, Ito J, Nakabayashi J, Shichino S, Ishiwata K, Karasuyama H. Single cell transcriptomics clarifies the basophil differentiation trajectory and identifies pre-basophils upstream of mature basophils. Nat Commun 2023;14:2694. https://doi.org/10.1038/s41467-023-38356-1.

[106] Ballesteros I, Rubio-Ponce A, Genua M, Lusito E, Kwok I, Fernández-Calvo G, et al. Co-option of Neutrophil Fates by Tissue Environments. Cell 2020;183:1282-1297.e18. https://doi.org/10.1016/j.cell.2020.10.003.

[107] Rodrigo-Muñoz JM, Naharro-González S, Callejas S, Relaño-Ruperez C, Torroja C, Benguria A, et al. Single-cell RNA sequencing of human blood eosinophils reveals plasticity and absence of canonical cell subsets. Allergy 2024:all.16213. https://doi.org/10.1111/all.16213.

[108] He P, Lim K, Sun D, Pett JP, Jeng Q, Polanski K, et al. A human fetal lung cell atlas uncovers proximal-distal gradients of differentiation and key regulators of epithelial fates. Cell 2022;185:4841-4860.e25. https://doi.org/10.1016/j.cell.2022.11.005.

[109] Jardine L, Webb S, Goh I, Quiroga Londoño M, Reynolds G, Mather M, et al. Blood and immune development in human fetal bone marrow and Down syndrome. Nature 2021;598:327–31. https://doi.org/10.1038/s41586-021-03929-x.

[110] Borrelli C, Gurtner A, Arnold IC, Moor AE. Stress-free single-cell transcriptomic profiling and functional genomics of murine eosinophils. Nat Protoc 2024;19:1679–709. https://doi.org/10.1038/s41596-024-00967-3.

[111] Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The Human Cell Atlas. eLife 2017;6:e27041. https://doi.org/10.7554/eLife.27041.

[112] Madissoon E, Oliver AJ, Kleshchevnikov V, Wilbrey-Clark A, Polanski K, Richoz N, et al. A spatially resolved atlas of the human lung characterizes a gland-associated immune niche. Nat Genet 2023;55:66–77. https://doi.org/10.1038/s41588-022-01243-4.

[113] Sauler M, McDonough JE, Adams TS, Kothapalli N, Barnthaler T, Werder RB, et al. Characterization of the COPD alveolar niche using single-cell RNA sequencing. Nat Commun 2022;13:494. https://doi.org/10.1038/s41467-022-28062-9.

[114] Zappia L, Theis FJ. Over 1000 tools reveal trends in the single-cell RNA-seq analysis landscape. Genome Biol 2021;22:301. https://doi.org/10.1186/s13059-021-02519-4.

[115] Curion F, Rich-Griffin C, Agarwal D, Ouologuem S, Rue-Albrecht K, May L, et al. Panpipes: a pipeline for multiomic single-cell and spatial transcriptomic data analysis. Genome Biol 2024;25:181. https://doi.org/10.1186/s13059-024-03322-7.

[116] Alexander Peltzer, Felipe Marques de Almeida, Gregor Sturm, heylf, Olga Botvinnik, nf-core bot, et al. nf-core/scrnaseq: 3.0.0 2024. https://doi.org/10.5281/ZENODO.14360028.

[117] Heumos L, Schaar AC, Lance C, Litinetskaya A, Drost F, Zappia L, et al. Best practices for single-cell analysis across modalities. Nat Rev Genet 2023;24:550–72. https://doi.org/10.1038/s41576-023-00586-w.

[118] Vieira Braga FA, Kar G, Berg M, Carpaij OA, Polanski K, Simon LM, et al. A cellular census of human lungs identifies novel cell states in health and in asthma. Nat Med 2019;25:1153–63. https://doi.org/10.1038/s41591-019-0468-5.

[119] Denisenko E, Guo BB, Jones M, Hou R, De Kock L, Lassmann T, et al. Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. Genome Biol 2020;21:130. https://doi.org/10.1186/s13059-020-02048-6.

[120] Van Den Brink SC, Sage F, Vértesy Á, Spanjaard B, Peterson-Maduro J, Baron CS, et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nat Methods 2017;14:935–6. https://doi.org/10.1038/nmeth.4437.

[121] Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol 2018;19:15. https://doi.org/10.1186/s13059-017-1382-0.

[122] Bredikhin D, Kats I, Stegle O. MUON: multimodal omics analysis framework. Genome Biol 2022;23:42. https://doi.org/10.1186/s13059-021-02577-8.

[123] Palla G, Spitzer H, Klein M, Fischer D, Schaar AC, Kuemmerle LB, et al. Squidpy: a scalable framework for spatial omics analysis. Nat Methods 2022;19:171–8. https://doi.org/10.1038/s41592-021-01358-2.

[124] Amezquita RA, Lun ATL, Becht E, Carey VJ, Carpp LN, Geistlinger L, et al. Orchestrating single-cell analysis with Bioconductor. Nat Methods 2020;17:137–45. https://doi.org/10.1038/s41592-019-0654-x.

[125] Luke Zappia, Aaron Lun. zellkonverter n.d. https://doi.org/10.18129/B9.BIOC.ZELLKONVERTER.

[126] Boggy GJ, McElfresh GW, Mahyari E, Ventura AB, Hansen SG, Picker LJ, et al. BFF and cellhashR: analysis tools for accurate demultiplexing of cell hashing data. Bioinformatics 2022;38:2791–801. https://doi.org/10.1093/bioinformatics/btac213.

[127] Lotfollahi M, Naghipourfar M, Luecken MD, Khajavi M, Büttner M, Wagenstetter M, et al. Mapping single-cell data to reference atlases by transfer learning. Nat Biotechnol 2022;40:121–30. https://doi.org/10.1038/s41587-021-01001-7.

[128] Squair JW, Gautier M, Kathe C, Anderson MA, James ND, Hutson TH, et al. Confronting false discoveries in single-cell differential expression. Nat Commun 2021;12:5692. https://doi.org/10.1038/s41467-021-25960-2.

[129] Badia-i-Mompel P, Vélez Santiago J, Braunger J, Geiss C, Dimitrov D, Müller-Dott S, et al. decoupleR: ensemble of computational methods to infer biological activities from omics data. Bioinformatics Advances 2022;2:vbac016. https://doi.org/10.1093/bioadv/vbac016.

[130] Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol 2015;16:278. https://doi.org/10.1186/s13059-015-0844-5.

[131] Dann E, Henderson NC, Teichmann SA, Morgan MD, Marioni JC. Differential abundance testing on single-cell data using k-nearest neighbor graphs. Nat Biotechnol 2022;40:245–53. https://doi.org/10.1038/s41587-021-01033-z.

[132] Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology 2012;16:284–7. https://doi.org/10.1089/omi.2011.0118.

[133] Dimitrov D, Schäfer PSL, Farr E, Rodriguez-Mier P, Lobentanzer S, Badia-i-Mompel P, et al. LIANA+ provides an all-in-one framework for cell–cell communication inference. Nat Cell Biol 2024;26:1613–22. https://doi.org/10.1038/s41556-024-01469-w.

[134] Nazarov VI, Tsvetkov VO, Fiadziushchanka S, Rumynskiy E, Popov AA, Balashov I, et al. immunarch: Bioinformatics Analysis of T-Cell and B-Cell Immune Repertoires n.d.

[135] Marconato L, Palla G, Yamauchi KA, Virshup I, Heidari E, Treis T, et al. SpatialData: an open and universal data framework for spatial omics. Nat Methods 2024. https://doi.org/10.1038/s41592-024-02212-x.

[136] Setty M, Kiseliovas V, Levine J, Gayoso A, Mazutis L, Pe'er D. Characterization of cell fate probabilities in single-cell data with Palantir. Nat Biotechnol 2019;37:451–60. https://doi.org/10.1038/s41587-019-0068-4.

[137] Grosse-Kathoefer S, Aglas L, Ferreira F, Pointner L. What inhalant allergens can do and not do?—The cooperation of allergens and their source in Th2 polarization and allergic sensitization. Allergo J Int 2023;32:258–68. https://doi.org/10.1007/s40629-023-00262-9.

[138] Miller JD. The Role of Dust Mites in Allergy. Clinic Rev Allerg Immunol 2019;57:312–29. https://doi.org/10.1007/s12016-018-8693-0.

[139] Corrigan CJ, Wang W, Meng Q, Fang C, Eid G, Caballero MR, et al. Allergen-induced expression of IL-25 and IL-25 receptor in atopic asthmatic airways and late-phase

cutaneous responses. Journal of Allergy and Clinical Immunology 2011;128:116–24. https://doi.org/10.1016/j.jaci.2011.03.043.

[140] Smole U, Gour N, Phelan J, Hofer G, Köhler C, Kratzer B, et al. Serum amyloid A is a soluble pattern recognition receptor that drives type 2 immunity. Nat Immunol 2020;21:756–65. https://doi.org/10.1038/s41590-020-0698-1.

[141] Schuijs MJ, Willart MA, Vergote K, Gras D, Deswarte K, Ege MJ, et al. Farm dust and endotoxin protect against allergy through A20 induction in lung epithelial cells. Science 2015;349:1106–10. https://doi.org/10.1126/science.aac6623.

[142] Huang H-J, Sarzsinszky E, Vrtala S. House dust mite allergy: The importance of house dust mite allergens for diagnosis and immunotherapy. Molecular Immunology 2023;158:54–67. https://doi.org/10.1016/j.molimm.2023.04.008.

[143] Posa D, Perna S, Resch Y, Lupinek C, Panetta V, Hofmaier S, et al. Evolution and predictive value of IgE responses toward a comprehensive panel of house dust mite allergens during the first 2 decades of life. Journal of Allergy and Clinical Immunology 2017;139:541-549.e8. https://doi.org/10.1016/j.jaci.2016.08.014.

[144] Willart MAM, Deswarte K, Pouliot P, Braun H, Beyaert R, Lambrecht BN, et al. Interleukin-1α controls allergic sensitization to inhaled house dust mite via the epithelial release of GM-CSF and IL-33. Journal of Experimental Medicine 2012;209:1505–17. https://doi.org/10.1084/jem.20112691.

[145] Klimov P, Molva V, Nesvorna M, Pekar S, Shcherbachenko E, Erban T, et al. Dynamics of the microbial community during growth of the house dust mite Dermatophagoides farinae in culture. FEMS Microbiology Ecology 2019;95:fiz153. https://doi.org/10.1093/femsec/fiz153.

[146] Cao T, Zhang Z, Liu Z-G, Dou X, Zhang J, Zhang W, et al. High-level expression and purification of the major house dust mite allergen Der p 2 in Escherichia coli. Protein Expression and Purification 2016;121:97–102. https://doi.org/10.1016/j.pep.2016.01.012.

[147] Manise M, Holtappels G, Van Crombruggen K, Schleich F, Bachert C, Louis R. Sputum IgE and Cytokines in Asthma: Relationship with Sputum Cellular Profile. PLoS ONE 2013;8:e58388. https://doi.org/10.1371/journal.pone.0058388.

[148] Wang J, Lindholt JS, Sukhova GK, Shi MA, Xia M, Chen H, et al. IgE actions on CD4+ T cells, mast cells, and macrophages participate in the pathogenesis of experimental abdominal aortic aneurysms. EMBO Mol Med 2014;6:952–69. https://doi.org/10.15252/emmm.201303811.

[149] McDonnell JM, Dhaliwal B, Sutton BJ, Gould HJ. IgE, IgE Receptors and Anti-IgE Biologics: Protein Structures and Mechanisms of Action. Annu Rev Immunol 2023;41:255–75. https://doi.org/10.1146/annurev-immunol-061020-053712.

[150] Redhu NS, Gounni AS. The high affinity IgE receptor (FcεRI) expression and function in airway smooth muscle. Pulmonary Pharmacology & Therapeutics 2013;26:86–94. https://doi.org/10.1016/j.pupt.2012.04.004.

[151] Hoof I, Schulten V, Layhadi JA, Stranzl T, Christensen LH, Herrera De La Mata S, et al. Allergen-specific IgG+ memory B cells are temporally linked to IgE memory responses. Journal of Allergy and Clinical Immunology 2020;146:180–91. https://doi.org/10.1016/j.jaci.2019.11.046.

[152] Setliff I, Shiakolas AR, Pilewski KA, Murji AA, Mapengo RE, Janowska K, et al. High-Throughput Mapping of B Cell Receptor Sequences to Antigen Specificity. Cell 2019;179:1636-1646.e15. https://doi.org/10.1016/j.cell.2019.11.003.

[153] Alladina J, Smith NP, Kooistra T, Slowikowski K, Kernin IJ, Deguine J, et al. A human model of asthma exacerbation reveals transcriptional programs and cell circuits specific to allergic asthma. Sci Immunol 2023;8:eabq6352. https://doi.org/10.1126/sciimmunol.abq6352.

[154] Luecken MD, Theis FJ. Current best practices in single-cell RNA-seq analysis: a tutorial. Molecular Systems Biology 2019;15:e8746. https://doi.org/10.15252/msb.20188746.

[155] Osorio D, Cai JJ. Systematic determination of the mitochondrial proportion in human and mice tissues for single-cell RNA-sequencing data quality control. Bioinformatics 2021;37:963–7. https://doi.org/10.1093/bioinformatics/btaa751.

[156] Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood A-MJ, et al. The Human Mitochondrial Transcriptome. Cell 2011;146:645–58. https://doi.org/10.1016/j.cell.2011.06.051.

[157] Thomas T, Friedrich M, Rich-Griffin C, Pohin M, Agarwal D, Pakpoor J, et al. A longitudinal single-cell atlas of anti-tumour necrosis factor treatment in inflammatory bowel disease. Nat Immunol 2024;25:2152–65. https://doi.org/10.1038/s41590-024-01994-8.

[158] Giladi A, Cohen M, Medaglia C, Baran Y, Li B, Zada M, et al. Dissecting cellular crosstalk by sequencing physically interacting cells. Nat Biotechnol 2020;38:629–37. https://doi.org/10.1038/s41587-020-0442-2.

[159] Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM, et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. Genome Biol 2018;19:224. https://doi.org/10.1186/s13059-018-1603-1.

[160] Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Systems 2019;8:281-291.e9. https://doi.org/10.1016/j.cels.2018.11.005.

[161] McInnes L, Healy J, Saul N, Großberger L. UMAP: Uniform Manifold Approximation and Projection. JOSS 2018;3:861. https://doi.org/10.21105/joss.00861.

[162] Luecken MD, Büttner M, Chaichoompu K, Danese A, Interlandi M, Mueller MF, et al. Benchmarking atlas-level data integration in single-cell genomics. Nat Methods 2022;19:41–50. https://doi.org/10.1038/s41592-021-01336-8.

[163] Wang LD, Wagers AJ. Dynamic niches in the origination and differentiation of haematopoietic stem cells. Nat Rev Mol Cell Biol 2011;12:643–55. https://doi.org/10.1038/nrm3184.

[164] Gayoso A, Lopez R, Xing G, Boyeau P, Valiollah Pour Amiri V, Hong J, et al. A Python library for probabilistic analysis of single-cell omics data. Nat Biotechnol 2022;40:163–6. https://doi.org/10.1038/s41587-021-01206-w.

[165] Shapiro-Shelef M, Calame K. Regulation of plasma-cell development. Nat Rev Immunol 2005;5:230–42. https://doi.org/10.1038/nri1572.

[166] Heeringa JJ, Rijvers L, Arends NJ, Driessen GJ, Pasmans SG, Van Dongen JJM, et al. IgE-expressing memory B cells and plasmablasts are increased in blood of children with asthma, food allergy, and atopic dermatitis. Allergy 2018;73:1331–6. https://doi.org/10.1111/all.13421.

[167] Bertrand Y, Sánchez-Montalvo A, Hox V, Froidure A, Pilette C. IgA-producing B cells in lung homeostasis and disease. Front Immunol 2023;14:1117749. https://doi.org/10.3389/fimmu.2023.1117749.

[168] Sokol CL, Medzhitov R. Emerging functions of basophils in protective and allergic immune responses. Mucosal Immunology 2010;3:129–37. https://doi.org/10.1038/mi.2009.137.

[169] Yu Y-RA, Hotten DF, Malakhau Y, Volker E, Ghio AJ, Noble PW, et al. Flow Cytometric Analysis of Myeloid Cells in Human Blood, Bronchoalveolar Lavage, and Lung Tissues. Am J Respir Cell Mol Biol 2016;54:13–24. https://doi.org/10.1165/rcmb.2015-0146OC.

[170] Chowdhury NU, Guntur VP, Newcomb DC, Wechsler ME. Sex and gender in asthma. Eur Respir Rev 2021;30:210067. https://doi.org/10.1183/16000617.0067-2021.

[171] Shafiei-Jahani P, Helou DG, Hurrell BP, Galle-Treger L, Howard E, Quach C, et al. CD52-targeted depletion by Alemtuzumab ameliorates allergic airway hyperreactivity and lung inflammation. Mucosal Immunology 2021;14:899–911. https://doi.org/10.1038/s41385-021-00388-5.

[172] Ryckman C, Vandal K, Rouleau P, Talbot M, Tessier PA. Proinflammatory Activities of S100: Proteins S100A8, S100A9, and S100A8/A9 Induce Neutrophil Chemotaxis and Adhesion. The Journal of Immunology 2003;170:3233–42. https://doi.org/10.4049/jimmunol.170.6.3233.

[173] Lanningham-Foster L, Green CL, Langkamp-Henken B, Davis BA, Nguyen KT, Bender BS, et al. Overexpression of CRIP in transgenic mice alters cytokine patterns and the immune response. American Journal of Physiology-Endocrinology and Metabolism 2002;282:E1197–203. https://doi.org/10.1152/ajpendo.00508.2001.

[174] Levenson CW, Shay NF, Lee-Ambrose LM, Cousins RJ. Regulation of cysteine-rich intestinal protein by dexamethasone in the neonatal rat. Proc Natl Acad Sci USA 1993;90:712–5. https://doi.org/10.1073/pnas.90.2.712.

[175] Stefanowicz D, Hackett T-L, Garmaroudi FS, Günther OP, Neumann S, Sutanto EN, et al. DNA Methylation Profiles of Airway Epithelial Cells and PBMCs from Healthy, Atopic and Asthmatic Children. PLoS ONE 2012;7:e44213. https://doi.org/10.1371/journal.pone.0044213.

[176] Berry M, Brightling C, Pavord I, Wardlaw A. TNF-α in asthma. Current Opinion in Pharmacology 2007;7:279–82. https://doi.org/10.1016/j.coph.2007.03.001.

[177] Thomas PS, Yates DH, Barnes PJ. Tumor necrosis factor-alpha increases airway responsiveness and sputum neutrophilia in normal human subjects. Am J Respir Crit Care Med 1995;152:76–80. https://doi.org/10.1164/ajrccm.152.1.7599866.

[178] Thomas PS. Effects of inhaled tumour necrosis factor alpha in subjects with mild asthma. Thorax 2002;57:774–8. https://doi.org/10.1136/thorax.57.9.774.

[179] Wenzel SE, Barnes PJ, Bleecker ER, Bousquet J, Busse W, Dahlén S-E, et al. A Randomized, Double-blind, Placebo-controlled Study of Tumor Necrosis Factor-α Blockade in Severe Persistent Asthma. Am J Respir Crit Care Med 2009;179:549–58. https://doi.org/10.1164/rccm.200809-1512OC.

[180] Lin T, Lo C, Tsao K, Chang P, Kuo CS, Lo Y, et al. Impaired interferon-α expression in plasmacytoid dendritic cells in asthma. Immunity Inflam & Disease 2021;9:183–95. https://doi.org/10.1002/iid3.376.

[181] Gill MA. The role of dendritic cells in asthma. Journal of Allergy and Clinical Immunology 2012;129:889–901. https://doi.org/10.1016/j.jaci.2012.02.028.

[182] Bartis D, Mise N, Mahida RY, Eickelberg O, Thickett DR. Epithelial–mesenchymal transition in lung development and disease: does it exist and is it important? Thorax 2014;69:760–5. https://doi.org/10.1136/thoraxjnl-2013-204608.

[183] Nihlberg K, Andersson-Sjoland A, Tufvesson E, Erjefalt JS, Bjermer L, Westergren-Thorsson G. Altered matrix production in the distal airways of individuals with asthma. Thorax 2010;65:670–6. https://doi.org/10.1136/thx.2009.129320.

[184] Heijink IH, Postma DS, Noordhoek JA, Broekema M, Kapus A. House Dust Mite–Promoted Epithelial-to-Mesenchymal Transition in Human Bronchial Epithelium. Am J Respir Cell Mol Biol 2010;42:69–79. https://doi.org/10.1165/rcmb.2008-0449OC.

[185] Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. Nat Rev Immunol 2007;7:255–66. https://doi.org/10.1038/nri2056.

[186] Lee C, Han J, Jung Y. Formyl peptide receptor 2 is an emerging modulator of inflammation in the liver. Exp Mol Med 2023;55:325–32. https://doi.org/10.1038/s12276-023-00941-1.

[187] Chen K, Liu M, Liu Y, Wang C, Yoshimura T, Gong W, et al. Signal relay by CC chemokine receptor 2 (CCR2) and formylpeptide receptor 2 (Fpr2) in the recruitment of monocyte-derived dendritic cells in allergic airway inflammation. J Biol Chem 2013;288:16262–73. https://doi.org/10.1074/jbc.M113.450635.

[188] Maciuszek M, Cacace A, Brennan E, Godson C, Chapman TM. Recent advances in the design and development of formyl peptide receptor 2 (FPR2/ALX) agonists as pro-resolving agents with diverse therapeutic potential. European Journal of Medicinal Chemistry 2021;213:113167. https://doi.org/10.1016/j.ejmech.2021.113167.

[189] Goettsch C, Hutcheson JD, Aikawa M, Iwata H, Pham T, Nykjaer A, et al. Sortilin mediates vascular calcification via its recruitment into extracellular vesicles. Journal of Clinical Investigation 2016;126:1323–36. https://doi.org/10.1172/JCI80851.

[190] Iqbal F, Schlotter F, Becker-Greene D, Lupieri A, Goettsch C, Hutcheson JD, et al. Sortilin enhances fibrosis and calcification in aortic valve disease by inducing interstitial cell heterogeneity. European Heart Journal 2023;44:885–98. https://doi.org/10.1093/eurheartj/ehac818.

[191] Shin J-S, Greer AM. The role of FcεRI expressed in dendritic cells and monocytes. Cell Mol Life Sci 2015;72:2349–60. https://doi.org/10.1007/s00018-015-1870-x.

[192] Kamphorst AO, Guermonprez P, Dudziak D, Nussenzweig MC. Route of Antigen Uptake Differentially Impacts Presentation by Dendritic Cells and Activated Monocytes. The Journal of Immunology 2010;185:3426–35. https://doi.org/10.4049/jimmunol.1001205.

[193] Choi G, Kim B-S, Park Y-J, Shim I, Chung Y. Clonal Expansion of Allergen-specific $CD4^+$ T Cell in the Lung in the Absence of Lymph Nodes. Immune Netw 2017;17:163. https://doi.org/10.4110/in.2017.17.3.163.

[194] Vecchione A, Devlin JC, Tasker C, Ramnarayan VR, Haase P, Conde E, et al. IgE plasma cells are transcriptionally and functionally distinct from other isotypes. Sci Immunol 2024;9:eadm8964. https://doi.org/10.1126/sciimmunol.adm8964.

[195] Potaczek DP, Michel S, Sharma V, Zeilinger S, Vogelberg C, Von Berg A, et al. Different *FCER 1 A* polymorphisms influence I g E levels in asthmatics and non-asthmatics. Pediatric Allergy Immunology 2013;24:441–9. https://doi.org/10.1111/pai.12083.

[196] Nials AT, Uddin S. Mouse models of allergic asthma: acute and chronic allergen challenge. Dis Model Mech 2008;1:213–20. https://doi.org/10.1242/dmm.000323.

[197] Ben-Baruch Morgenstern N, Rochman M, Kotliar M, Dunn JLM, Mack L, Besse J, et al. Single-cell RNA-sequencing of human eosinophils in allergic inflammation in the esophagus. Journal of Allergy and Clinical Immunology 2024;154:974–87. https://doi.org/10.1016/j.jaci.2024.05.029.

[198] Megas S, Wilbrey-Clark A, Maartens A, Teichmann SA, Meyer KB. Spatial Transcriptomics of the Respiratory System. Annual Review of Physiology 2024. https://doi.org/10.1146/annurev-physiol-022724-105144.

[199] Atta L, Fan J. Computational challenges and opportunities in spatially resolved transcriptomic data analysis. Nat Commun 2021;12:5283. https://doi.org/10.1038/s41467-021-25557-9.

[200] Wickramasinghe S, Porwit A, Erber W. Normal bone marrow cells. Blood and Bone Marrow Pathology, Elsevier; 2011, p. 19–44. https://doi.org/10.1016/B978-0-7020-3147-2.00002-X.

[201] Vannan A, Lyu R, Williams AL, Negretti NM, Mee ED, Hirsh J, et al. Spatial transcriptomics identifies molecular niche dysregulation associated with distal lung remodeling in pulmonary fibrosis. Nat Genet 2025;57:647–58. https://doi.org/10.1038/s41588-025-02080-x.

[202] Breschi A, Gingeras TR, Guigó R. Comparative transcriptomics in human and mouse. Nat Rev Genet 2017;18:425–40. https://doi.org/10.1038/nrg.2017.19.

[203] Eng C-HL, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, et al. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature 2019;568:235–9. https://doi.org/10.1038/s41586-019-1049-y.

[204] Khalfaoui L, Symon FA, Couillard S, Hargadon B, Chaudhuri R, Bicknell S, et al. Airway remodelling rather than cellular infiltration characterizes both type2 cytokine biomarker-high and -low severe asthma. Allergy 2022;77:2974–86. https://doi.org/10.1111/all.15376.

[205] Green FHY, Williams DJ, James A, McPhee LJ, Mitchell I, Mauad T. Increased myoepithelial cells of bronchial submucosal glands in fatal asthma. Thorax 2010;65:32–8. https://doi.org/10.1136/thx.2008.111435.

[206] Carroll NG, Mutavdzic S, James AL. Increased mast cells and neutrophils in submucosal mucous glands and mucus plugging in patients with asthma. Thorax 2002;57:677–82. https://doi.org/10.1136/thorax.57.8.677.

[207] Li W, Gao P, Zhi Y, Xu W, Wu Y, Yin J, et al. Periostin: its role in asthma and its potential as a diagnostic or therapeutic target. Respir Res 2015;16:57. https://doi.org/10.1186/s12931-015-0218-2.

[208] Ackland J, Barber C, Heinson A, Azim A, Cleary DW, Christodoulides M, et al. Nontypeable *Haemophilus influenzae* infection of pulmonary macrophages drives neutrophilic inflammation in severe asthma. Allergy 2022;77:2961–73. https://doi.org/10.1111/all.15375.

[209] Gueugnon F, Thibault VC, Kearley J, Petit-Courty A, Vallet A, Guillon A, et al. Altered expression of the CCN genes in the lungs of mice in response to cigarette smoke exposure and viral and bacterial infections. Gene 2016;586:176–83. https://doi.org/10.1016/j.gene.2016.04.022.

[210] Herrera-Solorio AM, Peralta-Arrieta I, Armas López L, Hernández-Cigala N, Mendoza Milla C, Ortiz Quintero B, et al. LncRNA SOX2-OT regulates AKT/ERK and SOX2/GLI-1 expression, hinders therapy, and worsens clinical prognosis in malignant lung diseases. Molecular Oncology 2021;15:1110–29. https://doi.org/10.1002/1878-0261.12875.

[211] Yi C, Gu T, Li Y, Zhang Q. Depression of long non-coding RNA SOX2 overlapping transcript attenuates lipopolysaccharide-induced injury in bronchial epithelial cells via miR-455-3p/phosphatase and tensin homolog axis and phosphatidylinositol 3-kinase/protein kinase B pathway. Bioengineered 2022;13:13643–53. https://doi.org/10.1080/21655979.2022.2083820.

[212] Zhou JP, Wang Y, Li SQ, Zhang JQ, Lin YN, Sun XW, et al. Exogenous Ang-(1-7) inhibits autophagy via HIF-1α/THBS1/BECN1 axis to alleviate chronic intermittent hypoxia-enhanced airway remodelling of asthma. Cell Death Discov 2023;9:366. https://doi.org/10.1038/s41420-023-01662-0.

[213] Van'S Gravesande KS, Wechsler ME, Grasemann H, Silverman ES, Le L, Palmer LJ, et al. Association of a Missense Mutation in the *NOS3* Gene with Exhaled Nitric Oxide Levels. Am J Respir Crit Care Med 2003;168:228–31. https://doi.org/10.1164/rccm.200212-1491BC.

[214] Palumbo-Zerr K, Zerr P, Distler A, Fliehr J, Mancuso R, Huang J, et al. Orphan nuclear receptor NR4A1 regulates transforming growth factor-β signaling and fibrosis. Nat Med 2015;21:150–8. https://doi.org/10.1038/nm.3777.

[215] Fu YC, Das A, Wang D, Braun R, Yi R. scHolography: a computational method for single-cell spatial neighborhood reconstruction and analysis. Genome Biol 2024;25:164. https://doi.org/10.1186/s13059-024-03299-3.

[216] Lambrecht BN, Hammad H. Allergens and the airway epithelium response: Gateway to allergic sensitization. Journal of Allergy and Clinical Immunology 2014;134:499–507. https://doi.org/10.1016/j.jaci.2014.06.036.

[217] Kubo M. Innate and adaptive type 2 immunity in lung allergic inflammation. Immunol Rev 2017;278:162–72. https://doi.org/10.1111/imr.12557.

[218] Amelink M, De Groot JC, De Nijs SB, Lutter R, Zwinderman AH, Sterk PJ, et al. Severe adult-onset asthma: A distinct phenotype. Journal of Allergy and Clinical Immunology 2013;132:336–41. https://doi.org/10.1016/j.jaci.2013.04.052.

[219] Oppenheimer J, Hoyte FCL, Phipatanakul W, Silver J, Howarth P, Lugogo NL. Allergic and eosinophilic asthma in the era of biomarkers and biologics: similarities, differences and misconceptions. Annals of Allergy, Asthma & Immunology 2022;129:169–80. https://doi.org/10.1016/j.anai.2022.02.021.

[220] Kariyawasam HH, Robinson DS. The role of eosinophils in airway tissue remodelling in asthma. Current Opinion in Immunology 2007;19:681–6. https://doi.org/10.1016/j.coi.2007.07.021.

[221] Newcomb DC, Peebles RS. Th17-mediated inflammation in asthma. Current Opinion in Immunology 2013;25:755–60. https://doi.org/10.1016/j.coi.2013.08.002.

[222] Al-Ramli W, Préfontaine D, Chouiali F, Martin JG, Olivenstein R, Lemière C, et al. TH17-associated cytokines (IL-17A and IL-17F) in severe asthma. Journal of Allergy and Clinical Immunology 2009;123:1185–7. https://doi.org/10.1016/j.jaci.2009.02.024.

[223] Polosa R, Thomson NC. Smoking and asthma: dangerous liaisons. Eur Respir J 2013;41:716–26. https://doi.org/10.1183/09031936.00073312.

[224] Pelaia G, Vatrella A, Busceti MT, Gallelli L, Calabrese C, Terracciano R, et al. Cellular Mechanisms Underlying Eosinophilic and Neutrophilic Airway Inflammation in Asthma. Mediators of Inflammation 2015;2015:879783. https://doi.org/10.1155/2015/879783.

[225] Edwards MR, Bartlett NW, Hussell T, Openshaw P, Johnston SL. The microbiology of asthma. Nat Rev Microbiol 2012;10:459–71. https://doi.org/10.1038/nrmicro2801.

[226] Habener A, Grychtol R, Gaedcke S, DeLuca D, Dittrich A-M, Happle C, et al. IgA^+ memory B-cells are significantly increased in patients with asthma and small airway dysfunction. Eur Respir J 2022;60:2102130. https://doi.org/10.1183/13993003.02130-2021.

[227] Zenz R, Eferl R, Scheinecker C, Redlich K, Smolen J, Schonthaler HB, et al. Activator protein 1 (Fos/Jun) functions in inflammatory bone and skin disease. Arthritis Res Ther 2007;10:201. https://doi.org/10.1186/ar2338.

[228] Li J, Zhou J, Huang H, Jiang J, Zhang T, Ni C. Mature dendritic cells enriched in immunoregulatory molecules (mregDCs): A novel population in the tumour microenvironment and immunotherapy target. Clinical & Translational Med 2023;13:e1199. https://doi.org/10.1002/ctm2.1199.

[229] Travaglini KJ, Nabhan AN, Penland L, Sinha R, Gillich A, Sit RV, et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature 2020;587:619–25. https://doi.org/10.1038/s41586-020-2922-4.

[230] Tiberti S, Catozzi C, Croci O, Ballerini M, Cagnina D, Soriani C, et al. GZMKhigh CD8+ T effector memory cells are associated with CD15high neutrophil abundance in non-metastatic colorectal tumors and predict poor clinical outcome. Nat Commun 2022;13:6752. https://doi.org/10.1038/s41467-022-34467-3.

[231] Lin B, Shah VS, Chernoff C, Sun J, Shipkovenska GG, Vinarsky V, et al. Airway hillocks are injury-resistant reservoirs of unique plastic stem cells. Nature 2024;629:869–77. https://doi.org/10.1038/s41586-024-07377-1.

[232] Feng H, Khalil S, Neubig RR, Sidiropoulos C. A mechanistic review on GNAO1-associated movement disorder. Neurobiology of Disease 2018;116:131–41. https://doi.org/10.1016/j.nbd.2018.05.005.

[233] Sebastian R, Jin K, Pavon N, Bansal R, Potter A, Song Y, et al. Schizophrenia-associated NRXN1 deletions induce developmental-timing- and cell-type-specific vulnerabilities in human brain organoids. Nat Commun 2023;14:3770. https://doi.org/10.1038/s41467-023-39420-6.

[234] Gubernatorova EO, Namakanova OA, Gorshkova EkaterinaA, Medvedovskaya AD, Nedospasov SA, Drutskaya MS. Novel Anti-Cytokine Strategies for Prevention and Treatment of Respiratory Allergic Diseases. Front Immunol 2021;12:601842. https://doi.org/10.3389/fimmu.2021.601842.

[235] Polte T, Behrendt A, Hansen G. Direct evidence for a critical role of CD30 in the development of allergic asthma. Journal of Allergy and Clinical Immunology 2006;118:942–8. https://doi.org/10.1016/j.jaci.2006.07.014.

[236] Del Prete G, De Carli M, D'Elios MM, Daniel KC, Almerigogna F, Alderson M, et al. CD30-mediated signaling promotes the development of human T helper type 2-like T cells. The Journal of Experimental Medicine 1995;182:1655–61. https://doi.org/10.1084/jem.182.6.1655.

[237] Kim DH, Lim JY, Jang JY, Gwak J, Joo HA, Ryu S, et al. Distinct subsets of innate lymphoid cells in nasal polyp. Allergology International 2023;72:151–60. https://doi.org/10.1016/j.alit.2022.06.007.

[238] Farhangnia P, Ghomi SM, Mollazadehghomi S, Nickho H, Akbarpour M, Delbandi A-A. SLAM-family receptors come of age as a potential molecular target in cancer immunotherapy. Front Immunol 2023;14:1174138. https://doi.org/10.3389/fimmu.2023.1174138.

[239] Howie D, Okamoto S, Rietdijk S, Clarke K, Wang N, Gullo C, et al. The role of SAP in murine CD150 (SLAM)-mediated T-cell proliferation and interferon γ production. Blood 2002;100:2899–907. https://doi.org/10.1182/blood-2002-02-0445.

[240] Yusuf I, Kageyama R, Monticelli L, Johnston RJ, DiToro D, Hansen K, et al. Germinal Center T Follicular Helper Cell IL-4 Production Is Dependent on Signaling Lymphocytic Activation Molecule Receptor (CD150). The Journal of Immunology 2010;185:190–202. https://doi.org/10.4049/jimmunol.0903505.

[241] Yurchenko M, Skjesol A, Ryan L, Richard GM, Kandasamy RK, Wang N, et al. SLAMF1 is required for TLR4-mediated TRAM-TRIF–dependent signaling in human macrophages. Journal of Cell Biology 2018;217:1411–29. https://doi.org/10.1083/jcb.201707027.

[242] Hammad H, Chieppa M, Perros F, Willart MA, Germain RN, Lambrecht BN. House dust mite allergen induces asthma via Toll-like receptor 4 triggering of airway structural cells. Nat Med 2009;15:410–6. https://doi.org/10.1038/nm.1946.

[243] Kim HY, Umetsu DT, Dekruyff RH. Innate lymphoid cells in asthma: Will they take your breath away? Eur J Immunol 2016;46:795–806. https://doi.org/10.1002/eji.201444557.

[244] Cheng W-H, Kao S-Y, Chen C-L, Yuliani FS, Lin L-Y, Lin C-H, et al. Amphiregulin induces CCN2 and fibronectin expression by TGF-β through EGFR-dependent pathway in lung epithelial cells. Respir Res 2022;23:381. https://doi.org/10.1186/s12931-022-02285-2.

[245] Gordon ED, Palandra J, Wesolowska-Andersen A, Ringel L, Rios CL, Lachowicz-Scroggins ME, et al. IL1RL1 asthma risk variants regulate airway type 2 inflammation. JCI Insight 2016;1. https://doi.org/10.1172/jci.insight.87871.

[246] Mabalirajan U, Rehman R, Ahmad T, Kumar S, Leishangthem GD, Singh S, et al. 12/15-lipoxygenase expressed in non-epithelial cells causes airway epithelial injury in asthma. Sci Rep 2013;3:1540. https://doi.org/10.1038/srep01540.

[247] Du L, Xu C, Tang K, Shi J, Tang L, Lisha X, et al. Epithelial CST1 Promotes Airway Eosinophilic Inflammation in Asthma via the AKT Signaling Pathway. Allergy Asthma Immunol Res 2023;15:374. https://doi.org/10.4168/aair.2023.15.3.374.

[248] Yao L, Yuan X, Fu H, Guo Q, Wu Y, Xuan S, et al. Epithelium-derived cystatin SN inhibits house dust mite protease activity in allergic asthma. Allergy 2023;78:1507–23. https://doi.org/10.1111/all.15739.

[249] Thompson RC, Ohlsson K. Isolation, properties, and complete amino acid sequence of human secretory leukocyte protease inhibitor, a potent inhibitor of leukocyte elastase. Proc Natl Acad Sci USA 1986;83:6692–6. https://doi.org/10.1073/pnas.83.18.6692.

[250] Wright CD, Havill AM, Middleton SC, Kashem MA, Lee PA, Dripps DJ, et al. Secretory Leukocyte Protease Inhibitor Prevents Allergen-Induced Pulmonary Responses in Animal Models of Asthma. The Journal of Pharmacology and Experimental Therapeutics 1999;289:1007–14. https://doi.org/10.1016/S0022-3565(24)38229-1.

[251] Marino R, Thuraisingam T, Camateros P, Kanagaratham C, Xu YZ, Henri J, et al. Secretory leukocyte protease inhibitor plays an important role in the regulation of allergic asthma in mice. J Immunol 2011;186:4433–42. https://doi.org/10.4049/jimmunol.1001539.

[252] Snelgrove RJ, Goulding J, Didierlaurent AM, Lyonga D, Vekaria S, Edwards L, et al. A critical function for CD200 in lung immune homeostasis and the severity of influenza infection. Nat Immunol 2008;9:1074–83. https://doi.org/10.1038/ni.1637.

[253] Lauzon-Joset J-F, Marsolais D, Tardif-Pellerin É, Patoine D, Bissonnette EY. CD200 in asthma. The International Journal of Biochemistry & Cell Biology 2019;112:141–4. https://doi.org/10.1016/j.biocel.2019.05.003.

[254] Teng X, Xu L-F, Zhou P, Sun H-W, Sun M. Effects of Trefoil Peptide 3 on Expression of TNF-α, TLR4, and NF-κB in Trinitrobenzene Sulphonic Acid Induced Colitis Mice. Inflammation 2009;32:120–9. https://doi.org/10.1007/s10753-009-9110-x.

[255] Belle NM, Ji Y, Herbine K, Wei Y, Park J, Zullo K, et al. TFF3 interacts with LINGO2 to regulate EGFR activation for protection against colitis and gastrointestinal helminths. Nat Commun 2019;10:4408. https://doi.org/10.1038/s41467-019-12315-1.

[256] Greeley MA, Van Winkle LS, Edwards PC, Plopper CG. Airway Trefoil Factor Expression during Naphthalene Injury and Repair. Toxicological Sciences 2010;113:453–67. https://doi.org/10.1093/toxsci/kfp268.

[257] Litosh VA, Rochman M, Rymer JK, Porollo A, Kottyan LC, Rothenberg ME. Calpain-14 and its association with eosinophilic esophagitis. J Allergy Clin Immunol 2017;139:1762-1771.e7. https://doi.org/10.1016/j.jaci.2016.09.027.

[258] Liu T, Woodruff PG, Zhou X. Advances in non-type 2 severe asthma: from molecular insights to novel treatment strategies. Eur Respir J 2024;64:2300826. https://doi.org/10.1183/13993003.00826-2023.

[259] Hinks TSC, Zhou X, Staples KJ, Dimitrov BD, Manta A, Petrossian T, et al. Innate and adaptive T cells in asthmatic patients: Relationship to severity and disease mechanisms. Journal of Allergy and Clinical Immunology 2015;136:323–33. https://doi.org/10.1016/j.jaci.2015.01.014.

[260] Lezmi G, Abou Taam R, Dietrich C, Chatenoud L, De Blic J, Leite-de-Moraes M. Circulating IL-17-producing mucosal-associated invariant T cells (MAIT) are associated with symptoms in children with asthma. Clinical Immunology 2018;188:7–11. https://doi.org/10.1016/j.clim.2017.11.009.

[261] Young RP, Dekker JW, Wordsworth BP, Schou C, Pile KD, Matthiesen F, et al. HLA-DR and HLA-DP genotypes and immunoglobulin E responses to common major allergens. Clin Experimental Allergy 1994;24:431–9. https://doi.org/10.1111/j.1365-2222.1994.tb00931.x.

[262] Takejima P, Agondi RC, Rodrigues H, Aun MV, Kalil J, Giavina-Bianchi P. Allergic and Nonallergic Asthma Have Distinct Phenotypic and Genotypic Features. Int Arch Allergy Immunol 2017;172:150–60. https://doi.org/10.1159/000458151.

[263] Gagné-Ouellet V, Guay S-P, Boucher-Lafleur A-M, Bouchard L, Laprise C. DNA methylation signature of interleukin 1 receptor type II in asthma. Clin Epigenet 2015;7:80. https://doi.org/10.1186/s13148-015-0114-0.

[264] Baines K, Fu J, McDonald V, Gibson P. Airway gene expression of IL-1 pathway mediators predicts exacerbation risk in obstructive airway disease. COPD 2017;Volume 12:541–50. https://doi.org/10.2147/COPD.S119443.

[265] Watt AP, Sharp JA, Lefevre C, Nicholas KR. WFDC2 is differentially expressed in the mammary gland of the tammar wallaby and provides immune protection to the mammary gland and the developing pouch young. Developmental & Comparative Immunology 2012;36:584–90. https://doi.org/10.1016/j.dci.2011.10.001.

[266] Eklöf J, Alispahic IA, Sivapalan P, Wilcke T, Seersholm N, Armbruster K, et al. Targeted AntiBiotics for Chronic pulmonary diseases (TARGET ABC): can targeted antibiotic therapy improve the prognosis of Pseudomonas aeruginosa-infected patients with chronic pulmonary obstructive disease, non-cystic fibrosis bronchiectasis, and asthma? A multicenter, randomized, controlled, open-label trial. Trials 2022;23:817. https://doi.org/10.1186/s13063-022-06720-z.

[267] Schröder B. The multifaceted roles of the invariant chain CD74 — More than just a chaperone. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 2016;1863:1269–81. https://doi.org/10.1016/j.bbamcr.2016.03.026.

[268] Xin T, Chen M, Duan L, Xu Y, Gao P. Interleukin-32: its role in asthma and potential as a therapeutic agent. Respir Res 2018;19:124. https://doi.org/10.1186/s12931-018-0832-x.

[269] Beute J, Lukkes M, Koekoek EP, Nastiti H, Ganesh K, de Bruijn MJ, et al. A pathophysiological role of PDE3 in allergic airway inflammation. JCI Insight 2018;3:e94888, 94888. https://doi.org/10.1172/jci.insight.94888.

[270] Ahmad F, Chung YW, Tang Y, Hockman SC, Liu S, Khan Y, et al. Phosphodiesterase 3B (PDE3B) regulates NLRP3 inflammasome in adipose tissue. Sci Rep 2016;6:28056. https://doi.org/10.1038/srep28056.

[271] Annunziato F, Cosmi L, Liotta F, Maggi E, Romagnani S. Human T helper type 1 dichotomy: origin, phenotype and biological activities. Immunology 2015;144:343–51. https://doi.org/10.1111/imm.12399.

[272] Gauthier M, Chakraborty K, Oriss TB, Raundhal M, Das S, Chen J, et al. Severe asthma in humans and mouse model suggests a CXCL10 signature underlies corticosteroid-resistant Th1 bias. JCI Insight 2017;2:e94580. https://doi.org/10.1172/jci.insight.94580.

[273] Britt RD, Thompson MA, Sasse S, Pabelick CM, Gerber AN, Prakash YS. Th1 cytokines TNF-α and IFN-γ promote corticosteroid resistance in developing human airway smooth muscle. American Journal of Physiology-Lung Cellular and Molecular Physiology 2019;316:L71–81. https://doi.org/10.1152/ajplung.00547.2017.

[274] Guo FH, Comhair SAA, Zheng S, Dweik RA, Eissa NT, Thomassen MJ, et al. Molecular Mechanisms of Increased Nitric Oxide (NO) in Asthma: Evidence for Transcriptional and Post-Translational Regulation of NO Synthesis. The Journal of Immunology 2000;164:5970–80. https://doi.org/10.4049/jimmunol.164.11.5970.

[275] Hong L, Herjan T, Bulek K, Xiao J, Comhair SAA, Erzurum SC, et al. Mechanisms of Corticosteroid Resistance in Type 17 Asthma. The Journal of Immunology 2022;209:1860–9. https://doi.org/10.4049/jimmunol.2200288.

[276] Ferreira MC, Whibley N, Mamo AJ, Siebenlist U, Chan YR, Gaffen SL. Interleukin-17-Induced Protein Lipocalin 2 Is Dispensable for Immunity to Oral Candidiasis. Infect Immun 2014;82:1030–5. https://doi.org/10.1128/IAI.01389-13.

[277] Owen HC, Roberts SJ, Ahmed SF, Farquharson C. Dexamethasone-induced expression of the glucocorticoid response gene lipocalin 2 in chondrocytes. American Journal of Physiology-Endocrinology and Metabolism 2008;294:E1023–34. https://doi.org/10.1152/ajpendo.00586.2007.

[278] Flo TH, Smith KD, Sato S, Rodriguez DJ, Holmes MA, Strong RK, et al. Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. Nature 2004;432:917–21. https://doi.org/10.1038/nature03104.

[279] An HS, Yoo J-W, Jeong JH, Heo M, Hwang SH, Jang HM, et al. Lipocalin-2 promotes acute lung inflammation and oxidative stress by enhancing macrophage iron accumulation. Int J Biol Sci 2023;19:1163–77. https://doi.org/10.7150/ijbs.79915.

[280] Lee J-Y, Hall JA, Kroehling L, Wu L, Najar T, Nguyen HH, et al. Serum Amyloid A Proteins Induce Pathogenic Th17 Cells and Promote Inflammatory Disease. Cell 2020;180:79-91.e16. https://doi.org/10.1016/j.cell.2019.11.026.

[281] Anthony D, Seow HJ, Uddin M, Thompson M, Dousha L, Vlahos R, et al. Serum Amyloid A Promotes Lung Neutrophilia by Increasing IL-17A Levels in the Mucosa and γδ T Cells. Am J Respir Crit Care Med 2013;188:179–86. https://doi.org/10.1164/rccm.201211-2139OC.

[282] Angelow S, Ahlstrom R, Yu ASL. Biology of claudins. Am J Physiol Renal Physiol 2008;295:F867-876. https://doi.org/10.1152/ajprenal.90264.2008.

[283] Soini Y. Claudins in lung diseases. Respir Res 2011;12:70. https://doi.org/10.1186/1465-9921-12-70.

[284] Lee P-H, Kim B-G, Lee S-H, Lee J-H, Park S-W, Kim D-J, et al. Alteration in Claudin-4 Contributes to Airway Inflammation and Responsiveness in Asthma. Allergy Asthma Immunol Res 2018;10:25. https://doi.org/10.4168/aair.2018.10.1.25.

[285] Wray C, Mao Y, Pan J, Chandrasena A, Piasta F, Frank JA. Claudin-4 augments alveolar epithelial barrier function and is induced in acute lung injury. American Journal of Physiology-Lung Cellular and Molecular Physiology 2009;297:L219–27. https://doi.org/10.1152/ajplung.00043.2009.

[286] Liu Y, Cheng K, Sun M, Ding C, Li T, Jia Y, et al. UBD participates in neutrophilic asthma by promoting the activation of IL-17 signaling. International Journal of Biological Macromolecules 2024;264:130581. https://doi.org/10.1016/j.ijbiomac.2024.130581.

[287] Dapas M, Lee YL, Wentworth-Sheilds W, Im HK, Ober C, Schoettler N. Revealing polygenic pleiotropy using genetic risk scores for asthma. Human Genetics and Genomics Advances 2023;4:100233. https://doi.org/10.1016/j.xhgg.2023.100233.

[288] Al-Shaikhly T, Murphy RC, Parker A, Lai Y, Altman MC, Larmore M, et al. Location of eosinophils in the airway wall is critical for specific features of airway hyperresponsiveness and T2 inflammation in asthma. Eur Respir J 2022;60:2101865. https://doi.org/10.1183/13993003.01865-2021.